

**Structural characterization of the minimal human  
RISC-loading complex**

**Dissertation**

for the award of the degree

“Doctor rerum naturalium”

of the Georg-August-Universität Göttingen

submitted by

**Stephanie Schell**

born in Nordhausen

Göttingen 2013

**Members of the Thesis Committee:**

**Prof. Dr. Ralf Ficner** (Reviewer)

Department of Molecular and Structural Biology  
Institute for Microbiology and Genetics  
Georg-August-Universität Göttingen

**Prof. Dr. Holger Stark** (Reviewer)

Department of Electron Microscopy  
Max Planck Institute for Biophysical Chemistry  
Göttingen

**Prof. Dr. Kai Tittmann**

Department of Bioanalytics  
Albrecht von Haller Institute  
Georg-August-Universität Göttingen

**Members of the Extended Examination Board:**

**Prof. Dr. Marina Rodnina**

Department of Physical Biochemistry  
Max Planck Institute for Biophysical Chemistry  
Göttingen

**Prof. Dr. Heinz Neumann**

Department of Applied Synthetic Biology  
Institute for Microbiology and Genetics  
Georg-August-Universität Göttingen

**Dr. Jochen Hub**

Computational Molecular Biophysics Group  
Institute for Microbiology and Genetics  
Georg-August-Universität Göttingen

**Day of the oral examination: 11.03.2013**

## **Affidavit**

Herewith I declare that I prepared this thesis “Structural characterization of the minimal human RISC-loading complex” independently and with no other sources and aids than quoted.

Göttingen, 02.02.2013

Stephanie Schell



# **1 Table of contents**

<b>1</b>	<b>TABLE OF CONTENTS</b>	<b>1</b>
<b>2</b>	<b>ABBREVIATIONS</b>	<b>7</b>
<b>3</b>	<b>SUMMARY</b>	<b>11</b>
<b>4</b>	<b>INTRODUCTION</b>	<b>13</b>
<b>4.1</b>	<b>RNA INTERFERENCE</b>	<b>13</b>
4.1.1	SMALL INTERFERING RNAs (siRNAs)	14
4.1.1.1	Source of siRNA Precursors	15
4.1.1.2	SiRNA biogenesis	15
4.1.1.3	Posttranscriptional silencing by siRNAs	17
4.1.1.4	Transcriptional gene silencing by siRNAs	17
4.1.2	MICRO RNAs (miRNAs)	18
4.1.2.1	Canonical miRNA biogenesis	19
4.1.2.2	Alternative miRNA biogenesis pathways	20
4.1.2.3	RNA target recognition and miRNA function	21
<b>4.2</b>	<b>PROTEINS INVOLVED IN RNAI</b>	<b>23</b>
4.2.1	DICER PROTEINS	23
4.2.1.1	Human Dicer	23
4.2.1.2	Generation of precise siRNA and miRNA products by Dicer	25
4.2.1.3	Involvement of additional Dicer domains in RNA processing	26
4.2.1.4	Interaction of human Dicer with other RNAi proteins	27
4.2.2	ARGONAUTE PROTEINS	28
4.2.2.1	The Argonaute protein family	28
4.2.2.2	Human Argonaute proteins	28
4.2.2.3	Structure of Argonaute proteins	29
4.2.3	TRBP	32
4.2.3.1	Structure of TRBP dsRBDs	32
4.2.3.2	TRBP interaction with other proteins	33
<b>4.3</b>	<b>RISC ASSEMBLY</b>	<b>34</b>
4.3.1	RISC-LOADING COMPLEXES	34
4.3.2	EM STRUCTURE OF THE MINIMAL HUMAN RISC-LOADING COMPLEX	36
<b>4.4</b>	<b>GOAL OF THE THESIS</b>	<b>38</b>

<b>5 MATERIAL AND METHODS</b>	<b>39</b>
<b>5.1 MATERIALS</b>	<b>39</b>
5.1.1 CHEMICALS AND REAGENTS	39
5.1.2 EQUIPMENT AND INSTRUMENTATION	39
5.1.3 CHROMATOGRAPHIC COLUMNS	40
5.1.4 KIT SYSTEMS	41
5.1.5 MARKER	41
5.1.6 ENZYMES AND INHIBITORS	41
5.1.7 ANTIBIOTICS WITH WORKING CONCENTRATIONS	41
5.1.8 ANTIBODIES	42
5.1.9 CROSS-LINKER	42
5.1.10 BACTERIAL STRAINS AND CELL LINES	42
5.1.11 DNA-OLIGONUCLEOTIDES	42
5.1.12 PLASMIDS	43
5.1.13 RNAs	44
5.1.14 SOLUTIONS FOR CRYSTALLIZATION SCREENS AND FLUORESCENCE-BASED THERMAL SHIFT ASSAY	44
5.1.15 COMPUTER PROGRAMS	45
5.1.16 CELL CULTURE MEDIUM AND TRANSFECTION REAGENTS	46
<b>5.2 METHODS</b>	<b>46</b>
5.2.1 MOLECULAR BIOLOGICAL METHODS	46
5.2.1.1 General cloning strategies	46
5.2.1.2 Polymerase chain reaction	47
5.2.1.3 DNA isolation	47
5.2.1.4 DNA restriction digestion	48
5.2.1.5 DNA ligation	48
5.2.1.6 DNA Sequencing	48
5.2.1.7 Concentration determination of DNA and RNA	48
5.2.1.8 Agarose gel electrophoresis and visualization of DNA and RNA	49
5.2.1.9 Electrophoretic mobility shift assay (EMSA)	49
5.2.1.10 Urea polyacrylamide gel electrophoresis (urea-PAGE) and visualization of RNA	49
5.2.1.11 In vitro Dicer assay	50
5.2.1.12 In vitro RISC cleavage assay	51
5.2.2 CELL BIOLOGICAL METHODS	52
5.2.2.1 Preparation of medium and agar plates for cultivation of E. coli	52
5.2.2.2 Preparation of chemical competent E. coli cells	53
5.2.2.3 Plasmid Transformation	54

5.2.2.4	Plasmid Isolation	54
5.2.2.5	Heterologous protein production in E. coli	55
5.2.2.6	Protein expression using insect cells	56
5.2.2.6.1	Generation of recombinant bacmids for transfection	56
5.2.2.6.2	Culturing of insect cells	56
5.2.2.6.3	Thawing of insect cells	56
5.2.2.6.4	Freezing of insect cells	57
5.2.2.6.5	Transfection of Sf9 and Sf 21 cells in a 6-well format	57
5.2.2.6.6	Virus production using Sf9 cells	58
5.2.2.6.7	Protein production using Sf9 and High Five cells	58
5.2.2.7	Preparation of cell extracts for protein purification	59
5.2.3	PROTEIN CHEMISTRY	59
5.2.3.1	Concentration determination of proteins	59
5.2.3.2	Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE)	60
5.2.3.3	Concentration of protein solutions	60
5.2.3.4	Purification of proteins and protein complexes from E. coli cells	60
5.2.3.5	Purification of proteins and protein complexes from insect cells	62
5.2.3.6	Co-purification of full length human Dicer-TRBP2 complex	63
5.2.3.7	Analytical size exclusion chromatography	64
5.2.3.8	Western blot	64
5.2.3.9	Limited proteolysis	64
5.2.3.10	Pull down assay	65
5.2.4	SPECIAL METHODS/ BIOPHYSICAL METHODS	65
5.2.4.1	Multi-angle light scattering combined with size exclusion chromatography	65
5.2.4.2	Fluorescence-based thermal stability assay	65
5.2.4.3	Isothermal titration calorimetry	66
5.2.4.4	Characterization of the ATPase activity	66
5.2.4.5	GraFix preparation of macromolecular complexes	66
5.2.4.6	Electron microscopy	67
5.2.4.6.1	Preparation and imaging of negative-stained particles	67
5.2.4.6.2	Preparation and imaging of cryo samples	67
5.2.4.7	Small angle X-ray scattering (SAXS)	68
5.2.4.8	Crystallization	69
5.2.4.8.1	General crystallization setup	69
5.2.4.8.2	In situ proteolysis	69
5.2.4.8.3	Crystallization and structure determination of the second dsRNA-binding domain of hTRBP2	69

<b>6 RESULTS</b>	<b>71</b>
<b>6.1 RISC-LOADING COMPLEX (RLC)</b>	<b>71</b>
6.1.1 PREPARATION OF THE HUMAN RISC-LOADING COMPLEX	71
6.1.1.1 Preparation of hAgo2	72
6.1.1.2 Preparation of hTRBP2	73
6.1.1.3 Preparation of hDicer	74
6.1.1.4 In vitro reconstitution of a minimal human RLC	76
6.1.2 FUNCTIONAL ANALYSIS OF THE RLC PROTEINS hDICER, hAGO2 AND hTRBP2	79
6.1.2.1 Endonuclease activities of human Dicer and Argonaute2	79
6.1.2.2 SiRNA-binding of human TRBP2	80
6.1.2.3 RISC-loading ability of the in vitro reconstituted RLC	83
6.1.3 MALS AND THERMAL SHIFT ASSAYS OF THE MINIMAL HUMAN RLC	85
6.1.3.1 MALS analysis of the RLC	85
6.1.3.2 Fluorescence-based thermal shift assays of the RLC	86
6.1.4 CRYSTALLIZATION TRIALS OF THE RLC	88
6.1.5 EM-ANALYSIS OF THE RLC	91
6.1.5.1 Optimization of EM image quality	92
6.1.5.2 Preliminary EM structure of the RLC	96
6.1.5.3 EM analysis of RLC subcomplexes and Dicer	98
<b>6.2 DICER-TRBP2 COMPLEXES</b>	<b>101</b>
6.2.1 PURIFICATION OF FULL LENGTH HUMAN DICER-TRBP2 COMPLEX	101
6.2.2 PREPARATION OF MINIMAL HUMAN DICER-TRBP2 COMPLEXES	104
6.2.3 SiRNA-BINDING PROPERTIES OF DTB1-T1 COMPLEX COMPARED TO hTRBP2 AND hTRBP2 FRAGMENTS	106
6.2.4 CRYSTALLIZATION TRIALS OF MINIMAL HUMAN DICER-TRBP2 COMPLEXES	108
<b>6.3 PREPARATION AND CRYSTALLIZATION OF hTRBP2-DOMAINS</b>	<b>112</b>
6.3.1 DETERMINATION OF THE CRYSTAL STRUCTURE OF THE SECOND DSRBD OF hTRBP2	114
6.3.2 OVERALL STRUCTURE OF THE SECOND DSRBD OF HUMAN TRBP2	115
6.3.3 CRYSTALLIZATION ARTIFACT OR DIMERIZATION INTERFACE?	117
<b>7 DISCUSSION</b>	<b>119</b>
<b>7.1 THE HUMAN RISC-LOADING COMPLEX</b>	<b>119</b>
7.1.1 PURIFICATION AND FUNCTIONALITY OF hAGO2	119
7.1.2 PURIFICATION AND ENDONUCLEASE ACTIVITY OF THE hDICER	120
7.1.3 PREPARATION AND RNA-BINDING ABILITIES OF THE DSRBP hTRBP2	121
7.1.4 PURIFICATION AND HOMOGENEITY ANALYSIS OF RLC	122
7.1.5 SiRNA-LOADING ABILITY OF THE RLC	124

7.1.6	IN VITRO RECONSTITUTION OF MINIMAL hDICER-hTRBP2 COMPLEXES	126
7.1.7	EM STRUCTURE OF HUMAN RLC	127
<b>7.2</b>	<b>CRYSTAL STRUCTURE OF DSRBD2 OF HUMAN TRBP2</b>	<b>133</b>
7.2.1	DSRBD2 OF hTRBP2 HAS A CANONICAL DSRBD FOLD THAT HAS HIGH FLEXIBILITY IN THE RNA-BINDING REGIONS	133
7.2.2	THE CRYSTAL STRUCTURE OF DSRBD2 OF hTRBP2 REVEALS TWO POSSIBLE DIMERIZATION INTERFACES	138
<b>8</b>	<b>REFERENCES</b>	<b>141</b>
<b>9</b>	<b>ACKNOWLEDGEMENTS</b>	<b>163</b>
<b>10</b>	<b>APPENDIX</b>	<b>165</b>
<b>11</b>	<b>CURRICULUM VITAE</b>	<b>179</b>

Table of contents

## 2 Abbreviations

<b>3D</b>	three-dimensional
<b>3'-UTR</b>	3'-untranslated region
<b>5'-UTR</b>	5'-untranslated region
<b>°C</b>	degree centigrade
<b>α</b>	alpha or anti
<b>ADP</b>	Adenosinediphosphate
<b>Ago</b>	Argonaute
<b>APS</b>	Ammonium peroxydisulfate
<b><i>A. thaliana</i></b>	<i>Arabidopsis thaliana</i>
<b>ATP</b>	Adenosine triphosphate
<b>bp</b>	base pair
<b>C</b>	Cytosine, carbon or carboxy
<b>ca.</b>	circa
<b>cDNA</b>	complementary DNA
<b><i>C. elegance (C.e.)</i></b>	<i>Caenorhabditis elegans</i>
<b><i>C. thermophilum (C.t.)</i></b>	<i>Chaetomium thermophilum</i>
<b>CStrep</b>	carboxy-terminal Strep-tag
<b>Da</b>	Dalton
<b>dd</b>	double distilled
<b>ddH<sub>2</sub>O</b>	double distilled water
<b>DExD/H, DEAD</b>	Consensus sequences of helicases
<b>DFDNB</b>	1,5-Difluoro-2,4-dinitrobenzene
<b>DGCR8</b>	DiGeorge syndrome critical region gene 8
<b>DMSO</b>	Dimethylsulfoxide
<b>DNA</b>	deoxyribonucleic acid
<b>dNTP</b>	deoxynucleotide-5'-triphosphate
<b><i>D. melanogaster (D.m.)</i></b>	<i>Drosophila melanogaster</i>
<b>ds</b>	double stranded
<b>dsRNA</b>	double stranded RNA
<b>DTT</b>	dithiothreitol
<b>DUF</b>	domain of unknown function
<b><i>E. coli</i></b>	<i>Escherichia coli</i>
<b>EDC</b>	1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide
<b>EDTA</b>	ethylenediaminetetraacetate
<b>EM</b>	electron microscopy
<b>Exp-5</b>	Exportin-5
<b>GA</b>	Glutaraldehyde
<b>GDP</b>	guanosine diphosphate
<b>GW182</b>	glycine-tryptophan repeat-containing protein of 182 kDa
<b>GraFix</b>	Gradient Fixation
<b>GST</b>	Glutathione-S-transferase
<b>GTP</b>	guanosine triphosphate
<b>HEPES</b>	2-[4-(2-Hydroxyethyl)-1piperazino-1]-ethane sulfonic acid

## Abbreviations

<b>H3K9</b>	Histone 3 Lysine 9
<b>hAgo2</b>	<i>Homo sapiens</i> Argonaute2
<b>hDicer</b>	<i>Homo sapiens</i> Dicer
<b><i>H. sapiens (H.s.)</i></b>	<i>Homo sapiens</i>
<b>HSP</b>	heat shock protein
<b>hTRBP</b>	<i>Homo sapiens</i> TAR RNA-binding protein
<b>IPTG</b>	Isopropyl $\beta$ -D-1-thiogalactopyranosid
<b>ITC</b>	isothermal titration calorimetry
<b>l</b>	Liter
<b>LB</b>	lysogeny broth
<b>M</b>	Molar (mol/l)
<b>MALS</b>	Multi-angle light scattering
<b>mg</b>	Milligram
<b>ml</b>	Milliliter
<b>mM</b>	milli molar
<b><i>M. musculus (M.m.)</i></b>	<i>Mus musculus</i>
<b>miRNA</b>	micro-RNA
<b>mRNA</b>	messenger RNA
<b>miRNP</b>	micro-ribonucleoprotein
<b>miPDC</b>	miRNA precursor deposit complex
<b>MWCO</b>	Molecular weight cut off
<b>N</b>	Amino
<b>NEB</b>	New England Biolabs
<b><i>N. crassa (N.c.)</i></b>	<i>Neurospora crassa</i>
<b>ng</b>	nano grams
<b>N-hDicer</b>	NHis <sub>6</sub> -CStrep-humanDicer
<b>NHis<sub>6</sub> /N-h</b>	amino-terminal hexahistidine tag
<b>nm</b>	nano meter
<b>NMR</b>	nuclear magnetic resonance
<b>nt</b>	Nucleotide
<b>OD</b>	optical density
<b>OB</b>	Oligonucleotide binding
<b>OH</b>	Hydroxyl
<b>ORF</b>	open reading frame
<b>PABP</b>	poly(A)-binding protein
<b>PACT</b>	protein activator of PKR
<b>PAGE</b>	polyacrylamide gel electrophoresis
<b>PAZ</b>	Piwi-argonaute-Zwille
<b>PBS</b>	phosphate buffered saline
<b>p bodies</b>	processing bodies
<b>PCR</b>	polymerase chain reaction
<b>PEG</b>	Polyethylene glycol
<b>piRNA</b>	piwi-interacting RNA
<b>PIWI</b>	P-element induced wimpy testis
<b>PMPI</b>	N-(p-Maleimidophenyl) isocyanate
<b>PMSF</b>	phenylmethylsulfonylfluoride
<b>pre-miRNA</b>	precursor micro RNA
<b>pre-mRNA</b>	precursor messenger RNA

<b>PTGS</b>	Posttranscriptional gene silencing
<b>RBD</b>	RNA-binding domain
<b>RBP</b>	RNA-binding protein
<b>RIG-I</b>	retinoic acid-induced gene-I
<b>RISC</b>	RNA induced silencing complex
<b>RITS</b>	RNA-induced initiator of transcriptional gene silencing
<b>RLC</b>	RISC-loading complex
<b>RMSD</b>	root-mean-square deviation
<b>RNA</b>	Ribonucleic acid
<b>RdRP</b>	RNA-directed RNA polymerase
<b>Rg</b>	radius of gyration
<b>RNAi</b>	RNA interference
<b>RNase</b>	ribonuclease
<b>RNP</b>	ribonucleoprotein
<b>rRNA</b>	ribosomal RNA
<b>SAXS</b>	Small angle x-ray scattering
<b><i>S. cerevisiae</i> (S.c.)</b>	<i>Saccharomyces cerevisiae</i>
<b>SDS</b>	Sodium dodecyl sulfate
<b>SDS-PAGE</b>	Sodium dodecyl sulfate – polyacrylamide gel electrophoresis
<b>SEC</b>	size exclusion chromatography
<b>siRNA</b>	small interfering RNA
<b>SMCC</b>	Succinimidyl 4-(N maleimidomethyl) cyclohexane-1-carboxylate
<b>snoRNA</b>	small nucleolar RNA
<b>ss</b>	single-stranded
<b>TBS</b>	Tris buffered saline
<b>TEMED</b>	N,N,N',N'-tetramethylethylenediamine
<b>TGS</b>	Transcriptional gene silencing
<b>TRBP</b>	TAR RNA-binding protein
<b>Tris</b>	Tris-(hydroxymethyl)amino methane
<b><i>T. thermophilus</i></b>	<i>Thermus thermophilus</i>
<b>urea-PAGE</b>	Urea polyacrylamide gel electrophoresis
<b>UTP</b>	uridine triphosphate
<b>USA</b>	United States of America
<b>v/v</b>	volume per volume
<b>w/v</b>	weight per volume
<b>WAGO</b>	worm-specific argonaute
<b>µg</b>	micro gram
<b>µl</b>	micro liter

## Abbreviations

### 3 Summary

Gene expression is the prerequisite for protein biosynthesis in all cells. Rapid and fine-tuned control of gene expression in response to environmental changes is of great importance. A major gene regulation mechanism in eukaryotes is RNA interference (RNAi). RNAi is initiated by the generation of small 21-23 nucleotide-containing double stranded RNAs (dsRNAs) within RNA induced silencing complexes (RISCs). Originally small RNAs are double stranded, but the two strands have to be separated in order to base pair with their target mRNA leading to degradation or translational repression. The human RISC-loading complex (RLC) composed of Dicer, Ago2 and TRBP2 couples the processing of precursor RNA substrate by Dicer to the loading and duplex unwinding of the small RNA product onto Ago2. Within the RLC Dicer and Ago2 act as endoribonucleases, whereas the dsRNA-binding protein TRBP2 recruits the Dicer complex to Ago2 and is important for loading of the appropriate small RNA duplex into the complex and onto Ago2. Structural information of the RLC, subcomplexes as well as the single RLC-proteins is rare and the RNA-transfer mechanism within the RLC is poorly understood.

In order to tackle this problem an expression system for the individual recombinant proteins was established. Human Dicer and Ago2 proteins were expressed as N-terminal hexahistidine-tagged proteins in Sf9 and High Five insect cells, respectively, and human TRBP2 was expressed as N-terminal GST-tagged protein in *E.coli* BL21 (DE3) Star cells. In a subsequent step RLC assembly was established and the RLC and its components were structurally analyzed by means of macromolecular X-ray crystallography and single-particle electron microscopy. As crystallization of such a big and highly flexible complex is challenging, the main structural work on the RLC was carried out by single-particle electron microscopy studies. A three dimensional structure of the human RLC could be reconstructed at 22.8 Å resolution. The RLC has a C-shaped form, with highly flexible regions. The comparison with an earlier reconstruction of the human RLC showed that the RNA-bound complex analyzed in this study adopts a far more open conformation. To map the RLC components the aim was to calculate reference density maps of the subcomplexes and Dicer alone. Due to the high heterogeneity of the subcomplexes and Dicer, reliable reconstructions were not possible to date. However, more structural details can be seen in the RLC reconstruction of this work. Based on those as well as previous results from other groups a model for the RNA-transfer mechanism is discussed.

The large human Dicer protein contains an N-terminal DExD/H-helicase domain. The function of this domain is still obscure. Interestingly, this domain interacts with TRBP2 and the endonuclease activity of Dicer is activated upon their interaction. In order to understand this interaction in detail, a minimal Dicer-TRBP2-binding domain was defined. This fragment of Dicer reduced binding of hTRBP to single stranded RNA, suggesting a role in TRBPs substrate

## Summary

selection. An atomic model of the interaction surface using various complexes could not be obtained.

Human TRBP2 consist of three dsRBDs, which are important for RNA-binding and protein interaction as well as homodimerization. The crystal structure of the second dsRNA-binding domain (dsRBD2) of human TRBP2 could be solved and refined at 2.28 Å resolution showing a common  $\alpha$ - $\beta$ - $\beta$ - $\beta$ - $\alpha$  dsRBD fold. The asymmetric unit contains four molecules of the dsRBD2 which form two different dimerization interfaces. However, SAXS and MALS measurements revealed that dsRBD2 exists as a monomer in solution. Thus, these results show that dsRBD2 seems not to be involved in dimerization and suggests that the loop regions between the dsRBDs of human TRBP2 are important for its dimerization. Comparison of the structure with the previously solved structure of the dsRBD2 in complex with a short CG-duplex reveals, that the domain may undergo structural rearrangements upon RNA-binding. Additionally, the SAXS structure of the dsRBD2 uncover that the regions important for RNA-binding are highly flexible, hence allowing the binding of various RNA substrates.

## 4 Introduction

Gene expression leading to protein biosynthesis is a central process in all cells. Hereby the genetic information of the DNA is transcribed into RNA and RNA is subsequently translated into protein. In contrast to prokaryotes, eukaryotic cells are subdivided into different, membrane-enclosed compartments. The central compartment, the nucleus, contains the genetic information encoded in DNA. It is the location where the first step of the protein biosynthesis, the transcription of DNA into messenger RNA precursors (pre-mRNAs), occurs. pre-mRNA contains coding (exons) and non-coding regions (introns). Within the nucleus the pre-mRNA is post-transcriptionally modified and a large ribonucleoprotein machinery, called the spliceosome, removes the introns. After additional 5'-and 3'-end processing, the generated mature mRNA is exported to the cytoplasm, where the encoded protein sequence in the mRNA is translated by the ribosome into a poly-peptide. Hence, in eukaryotes transcription and translation are spatially and temporarily separated allowing for various possible regulatory mechanisms during every step of protein production. The transcribed DNA can be modified allowing for post-transcriptional regulation and transcription itself is regulated by various transcription factors. The processing of the pre-mRNA into mRNA and the mRNA lifetime (mRNA stability) represent other levels of regulation of gene expression. Finally, whether or not a mRNA is translated, is regulated by the translation initiation machinery and the mRNA degradation complexes. This various and complex gene regulation is necessary, when thinking about the complexity of multicellular organisms. One process that is able to regulate eukaryotic gene expression at the transcriptional, post-transcriptional and/or translational level is RNA interference (RNAi).

### 4.1 RNA interference

In the early 1990s, a mechanism called gene (or RNA) silencing was described for plants, whereby a synthetic gene incorporated into the genome inhibits the expression of the homologous sequence transcriptionally (transcriptional gene silencing, TGS) or post transcriptionally (posttranscriptional gene silencing, PTGS) (Matzke *et al.*, 1989; Wassenegger *et al.*, 1994; Park *et al.*, 1996; Napoli *et al.*, 1990; van der Krol *et al.*, 1990; Smithe *et al.*, 1990; de Carvalho *et al.*, 1992; van Blokland *et al.*, 1994). It became clear that RNA plays a key role in gene silencing and that the introduction of single stranded RNA into cells can be used to interfere with the function of an endogenous gene (Izant & Harold, 1984; Guo & Kemphues, 1995). However, the mechanism leading to this gene silencing effect was not understood until Fire *et al.* 1998 explained the phenomenon of RNAi on a molecular level. Surprisingly, they found that the injection of a long dsRNA into *Caenorhabditis elegans* led to a 10-fold higher efficient knock down of the complementary gene than single stranded sense or antisense RNA alone (Fire *et al.*,

## Introduction

1998). During the following years RNAi was described in fruit fly, zebra fish and mammals demonstrating that RNAi and related gene silencing pathways are widespread mechanisms to regulate gene expression in eukaryotes (Kennerdell & Carthew, 1998; Wargelius *et al.*, 1999; Lie *et al.*, 2000; Svoboda *et al.*, 2000; Wianny & Zernicka-Goetz, 2000). RNAi is initiated by the generation of small 21-23 nucleotide-containing dsRNAs and leads to repression of translation and transcription (Zamore *et al.*, 2000; Bernstein *et al.*, 2001; Saxena *et al.*, 2003; Zeng *et al.*, 2003; Doench & Sharp, 2004; Mette *et al.*, 2000; Sijen *et al.*, 2001; Volpe *et al.*, 2002). Additionally, it was early identified that cells use it to combat viral infections, ensure genome stability by keeping mobile elements like transposons silent and to regulate development (Covey *et al.*, 1997; Ratcliff *et al.*, 1997; Ketting *et al.*, 1999; Ketting *et al.*, 2001; Tabara *et al.*, 1999; Bernstein *et al.*, 2003). RNAi has become a powerful tool to study gene function, because it has several advantages to classical gene knockdown. Gene silencing by RNA only requires the injection of the appropriate RNA and might allow the knockdown of genes that cannot be knocked out by other general methods or that are lethal with classical methods. Moreover, it can be used to check gene function *in vivo* and to knockdown a gene during different developmental phases (Lawrence & Pikaard, 2003; Miki *et al.*, 2005; Kurreck, 2009).

RNAi is mediated by small RNAs. However small RNAs cannot catalyze any reactions by themselves, they rather become part of ribo-nucleoprotein (RNP) complexes, which are called RNA-induced silencing complexes (RISCs). The core of a RISC is composed of a small RNA (miRNA or siRNA) and a member of the Argonaute (Ago) protein family, which is the main effector protein of the complex providing a unique platform for target recognition and silencing. Additional proteins are known to associate, extend or modify the function of the RISC. Small non-coding RNAs that play a role in RNAi are classified based on their structure, biogenesis and function. The three biggest classes build the small interfering RNAs (siRNAs), the micro RNAs (miRNAs) and the PIWI protein interacting RNAs (piRNAs). RNAs of all classes have a length of 19-31 nucleotides (nts) and associate with an Ago protein (Rana, 2007; Hutvagner & Simard, 2008; Carthew & Sontheimer, 2009; Ghildiyal & Zamore, 2009; Kim *et al.*, 2009). Within the subsequent paragraphs, the characteristics and biogenesis of the major small RNAs in eukaryotes, the siRNAs and the miRNAs, are outlined.

### **4.1.1 Small interfering RNAs (siRNAs)**

Small interfering RNAs (siRNAs) are 20- to 25-nt long RNAs, that are the specific determinants for endonucleolytic cleavage of RNA targets by a RISC (Zamore *et al.*, 2000; Hammond *et al.*, 2000). By mediating transcriptional and posttranscriptional silencing these small RNAs have diverse biological roles in eukaryotes, including transposon and transgene silencing as well as antiviral and genome defense (Lippman & Martienssen, 2004; Carthew & Sontheimer, 2009; Ghildiyal & Zamore, 2009).

#### 4.1.1.1 Source of siRNA Precursors

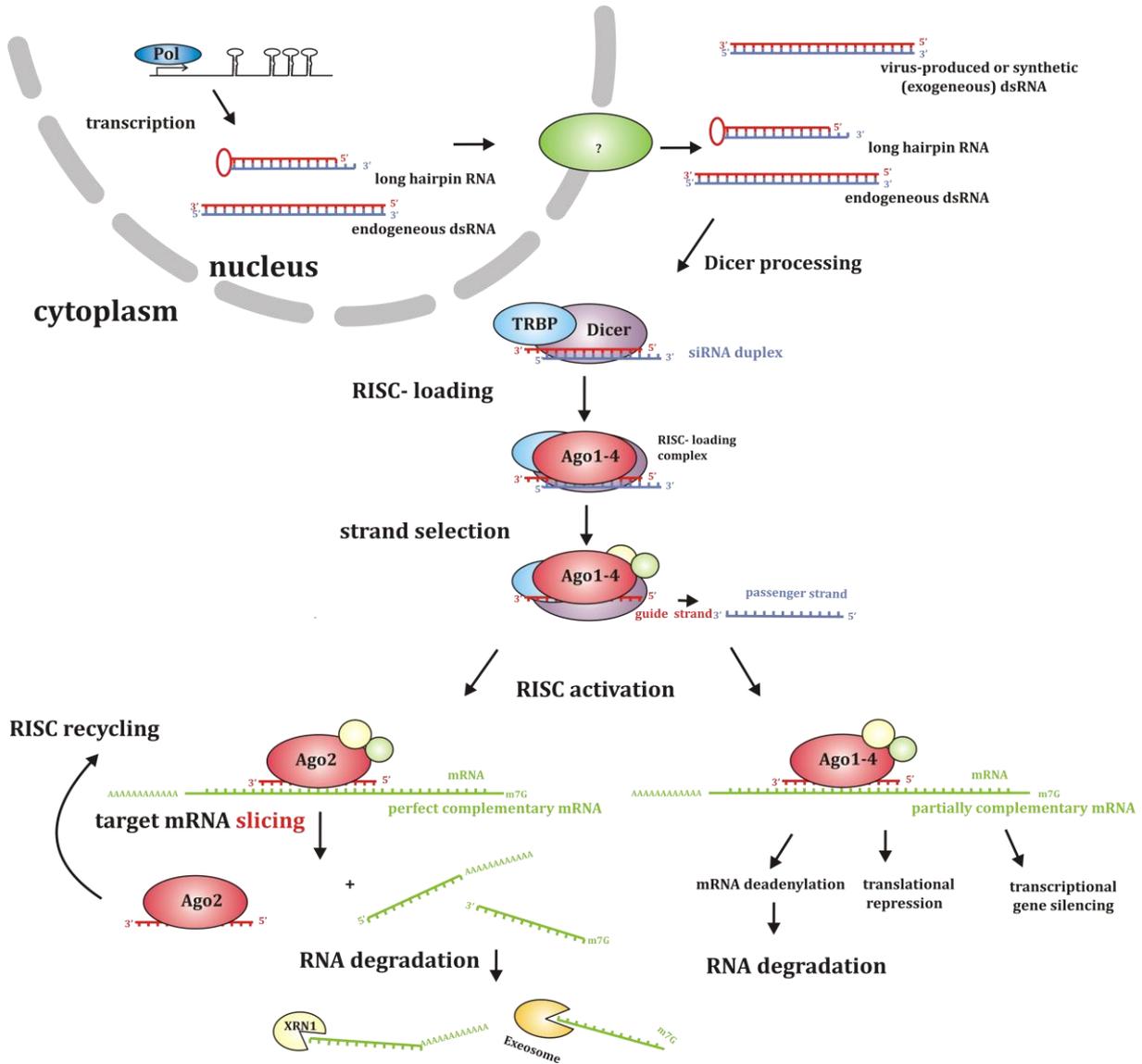
siRNAs were first discovered during transgene- and virus-induced silencing in plants and were later also detected in animals (Waterhouse *et al.*, 1998; Hamilton and Baulcombe, 1999; Zamore *et al.*, 2000; Hammond *et al.*, 2000; Elbashir *et al.*, 2001b; Elbashir *et al.*, 2001c). Originally, siRNAs have been considered extra genomic in origin and these siRNAs are therefore called exogenous siRNAs (exo-siRNAs). Additionally, endogenous siRNAs (endo-siRNAs) were identified in yeast, plants and *C. elegans*. The biogenesis of these siRNAs depend on the activity of RNA-dependent RNA polymerase (RdRp), which catalyze the replication of RNA from a RNA template (Cogoni & Macino, 2000; Ruby *et al.*, 2006; Pak & Fire, 2007; Vaucheret, 200). In flies or mice several endo-siRNA are generated from naturally occurring dsRNAs in an RdRP-independent manner. These dsRNA precursors include hairpin-dsRNAs, transposons, *trans*-natural antisense transcript RNAs (*trans*-nat-dsRNAs) and *cis*-natural antisense transcript RNAs (*cis*-nat-dsRNAs) (Ghildiyal *et al.*, 2008; Czech *et al.*, 2008; Chung *et al.*, 2008; Okamura *et al.*, 2008b, Nilsen, 2008; Golden *et al.*, 2008; Okamura & Lai, 2008).

#### 4.1.1.2 siRNA biogenesis

Although single stranded siRNAs can be directly loaded onto Ago proteins *in vitro*, general siRNA biogenesis is initiated by cleavage of an exogenous or endogenous, linear or hairpin structured long dsRNAs into double stranded siRNAs by the RNase III endonuclease Dicer in the cytoplasm (see Figure 1; Rivas *et al.*, 2005, Zamore *et al.*, 2000; Bernstein *et al.*, 2001; Billy *et al.*, 2001). The produced siRNA duplex is around 21-23 nts long, perfectly complementary, carries characteristic 5'-phosphate and 3'-hydroxyl groups and has 2-nt overhangs at the 3'-ends (Zamore *et al.*, 2001; Elbashir *et al.*, 2001a; Elbashir *et al.*, 2001b). In *H. sapiens*, a protein named Dicer1 catalyzes the cleavage of long dsRNAs into small siRNA duplexes, while its efficiency can be influenced by double stranded RNA-binding proteins like TRBP (TAR RNA-binding protein) and PACT (protein activator of PKR) (Chendrimada *et al.*, 2005; Haase *et al.*, 2005; Lee *et al.*, 2006; for more detail see later section 4.2.1.4,4.2.3.2 and Figure 1). Assisted by TRBP the siRNA product is loaded onto an Ago protein within the RISC-loading complex (RLC). The relative thermodynamic stabilities of the 5'-ends of the two siRNA strand in the produced duplex determines the identity of the guide and passenger strand, but the mechanisms of strand selection are mostly unclear in humans (Khvorova *et al.*, 2003; Schwarz *et al.*, 2003), for further detail see section 4.3.1). General unwinding of the siRNA duplex depends on endonucleolytic cleavage of the passenger strand by human Ago2, which is also called slicer due to its endonucleolytic activity. During this process, Ago2 nicks the siRNA duplex by cleaving the passenger strand phosphodiester bond facing the bond between nucleotides 10 and 11 of the guide strand. As a consequence, the thermodynamic stability of the duplex is reduced and it can be unwound efficiently. This slicer-dependent unwinding is typical to all siRNA duplexes

## Introduction

(Matranga *et al.*, 2005; Rand *et al.*, 2005; Miyoshi *et al.*, 2005; Leuschner *et al.*, 2006). Within a RISC the passenger strand is released and the guide strand directs the RISC to its perfectly complementary target mRNA for the initiation of the actual silencing effect (see Figure 1 and following sections 4.1.1.3, 4.1.1.4).



**Figure 1: Origin and maturation of exo- and endo-siRNAs.**

Detailed description of the depicted processes can be found in section 4.1.1. The figure was prepared according to Jinek & Doudna, 2009 and Röther & Meister, 2011.

#### **4.1.1.3 Posttranscriptional silencing by siRNAs**

During the canonical RNAi pathway, the siRNA guide strand directs the RISC to its perfectly complementary target mRNA, which in turn is then silenced by cleavage and degraded by further recruitment of the RNA degradation machinery (see Figure 1). The 5'-end of the guide strand sets the ruler for target-RNA cleavage, because cleavage occurs between nucleotide 10 and 11 upstream of the 5'-end and products with 5'-monophosphate and 3'-hydroxyl termini are produced (Tomari & Zamore, 2005). The products are attacked by cellular exonucleases to complete the degradation (Orban & Izaurralde, 2005). Additionally, the products with 3'-hydroxyl termini can be substrates for oligouridylation, which can also promote their exonucleolytic degradation (Shen & Goodman, 2004). After cleavage the target dissociates from the siRNA, freeing the RISC to cleave additional targets (see RISC recycling in Figure 1). If endonucleolytic cleavage of the RNA target cannot occur, like in cases where the cleavage is suppressed by mismatches near the center of the siRNA/target duplex or when the siRNA is loaded into a Ago protein that lack endonucleolytic cleavage activity, these RISCs can still silence at the posttranscriptional level. In such cases, translational repression or exonucleolytic degradation similar to miRNA silencing occurs (for more detail see later section 4.1.2.3 and 4.2.2.2).

In general, the described effector steps primarily occur in the cytoplasm, but siRNA-binding can also induce the localization of Ago proteins/effector RISCs to subcellular loci called P-bodies (processing-bodies) in mammals, which were identified to be involved in storage and degradation of translationally repressed mRNAs (Liu *et al.*, 2005; Eulalio *et al.*, 2007). Although P-body localization is not required for effective RNAi silencing, this compartmentalization may increase silencing efficiency, as the P-bodies are components enriched in mRNA degradation factors (Chu & Rana, 2006). Additionally, siRNA-binding to RISCs can induce their nuclear translocation (Ohrt *et al.*, 2008; Guang *et al.*, 2008). Finally, the recently discovered endo-siRNAs can also repress the expression of targeted genes at posttranscriptional levels by inducing mRNA degradation and/or translational suppression as described above (Vaucheret, 2006; Martienssen *et al.*, 2005; Yigit *et al.*, 2006).

#### **4.1.1.4 Transcriptional gene silencing by siRNAs**

Although RNAi is mainly a posttranscriptional mechanism, components of the RNAi machinery can also be involved in nuclear processes, leading to heterochromatin formation and thus direct TGS. These processes are known as RNA-induced initiation of transcriptional gene silencing (RITS) or nuclear RNAi. The mechanism of RNA-induced heterochromatin formation is a natural epigenetic gene regulation mechanism that plays a role in defense against foreign DNA like retro elements or transposons (Lippman & Martienssen, 2004; Verdell *et al.*, 2009; Lejeune *et al.*, 2011; Fagegaltier *et al.*, 2009). Additionally, genes involved in cell development or

## Introduction

chromosome segregation during cell division seem to be regulated by nuclear RNAi. The RNA-mediated heterochromatin formation is best understood in *Schizosaccharomyces pombe* (see reviews by: Wasseregger, 2005; Creamer & Partridge, 2011; Goto & Nakayama, 2011; Reyes-Turcu & Grewal, 2012). This phenomena of has also been discovered in human cells where Ago1 and sequence-specific siRNAs were demonstrated to be essential components of a RITS during RNA polymerase II (Pol-II) dependent siRNA (Kim *et al.*, 2006; Morris *et al.*, 2004). More evidence for nuclear RNAi is given by the fact, that human Ago1 and Ago2 proteins can be found in the nucleus (Robb *et al.*, 2005). Additionally, continuous suppression of simian immunodeficiency virus (SIV) and human immunodeficiency virus type-1 (HIV-1) infections could be shown by siRNA-mediated TGS (Lim *et al.*, 2008; Suzuki *et al.*, 2008). These processes are associated with histone and chromatin modifications, however, the molecular mechanisms underlying the processes of TGS in mammals are still being defined.

### **4.1.2 Micro RNAs (miRNAs)**

Besides, endo-siRNAs and exo-siRNAs, an additional class of small RNAs involved in RNAi, the miRNAs, were discovered. miRNAs are short 20- to 25-nt containing endogenous single stranded RNAs, which are generated from double stranded precursors of eukaryotic nuclear DNA (for more detail see section 4.1.2.1). They interact with members of the Ago protein family to form the so-called miRNA-induced silencing complex (miRISC), usually resulting in gene silencing via translational repression or mRNA degradation (Bartel, 2009). The first described miRNA, lin-4, had been discovered in *C. elegans* as endogenous regulator of genes that control larval developmental timing (Lee *et al.*, 1993). About 7 years later the second miRNA, let-7 was discovered. This miRNA also regulates developmental timing in *C. elegans* (Reinhart *et al.*, 2000). Within the following years thousands of miRNAs have been discovered in diverse organisms like plants and mammals as well as viruses. Now it is known that miRNAs are ubiquitously found, regulating elements within the mammalian genome. More than 1000 human miRNAs have been identified so far, whereby each miRNA can regulate several hundred mRNA targets. The fact that more than 50% of the human transcriptome are subject to miRNA regulation underlines the fundamental role of small RNAs in human gene regulation mechanisms and as core regulators in development and cellular functions (Bentwich *et al.*, 2005; Lewis *et al.*, 2005; Friedman *et al.*, 2009; Kloosterman & Plasterk, 2006; Bartel, 2009; Carthew & Sontheimer, 2009). Furthermore miRNAs have been found to be involved in a wide spectrum of human diseases including cancer as well as cardiovascular and autoimmune inflammatory conditions (Erson & Petty, 2009; Melo *et al.*, 2009; Kloosterman & Plasterk, 2006, Kwak *et al.*, 2010).

#### 4.1.2.1 Canonical miRNA biogenesis

The canonical and best-studied pathway for miRNA biogenesis is conserved among vertebrates and invertebrates and is illustrated in Figure 2 (review: Kim *et al.*, 2009; Winter *et al.*, 2009). miRNAs generated by the canonical biogenesis pathway are directly transcribed in the nucleus from intergenic, intronic or polycistronic genomic loci by RNA Polymerase-II into long, often polycistronic, pri-miRNA (primary miRNA) transcripts (Lee *et al.*, 2002; Lee *et al.*, 2004). Additionally, some miRNAs can be transcribed by RNA Polymerase III (Borchert *et al.*, 2006). pri-miRNAs can form a highly complex structure containing several stem loops in a row. Furthermore, these long transcripts (up to several kilobases) largely resemble protein-coding genes being posttranscriptionally capped and polyadenylated. Interestingly, pri-miRNAs often contain introns (Cai *et al.*, 2004; Rodriguez *et al.*, 2004).

The first step of miRNA maturation is an endonucleolytic cleavage at the stem of the hairpin structure of the pri-miRNA by the RNase III type endoribonuclease Drosha (Lee *et al.*, 2003). It was shown that in *Homo sapiens* and *Drosophila melanogaster* Drosha requires the double stranded RNA-binding domain (dsRBD)-containing cofactor DGCR8 (DiGeorge syndrome critical region gene 8), for efficient and precise processing (Denli *et al.*, 2004; Gregory *et al.*, 2004; Han *et al.*, 2004; Yeom *et al.*, 2006). The complex of Drosha and DGCR8 is also described as microprocessor complex. DGCR8 directly interacts with the pri-miRNA and functions as a molecular ruler to determine the precise cleavage site for Drosha that cleaves the 5'- and 3'-arms 11 base pairs away from the single stranded RNA/double stranded RNA junction at the basis of the hairpin stem. By this cleavage, the microprocessor complex generates a long hairpin-structured precursor miRNA (pre-miRNA) containing 60-100 nucleotides with a 2-nt 3'-overhang (Han *et al.*, 2006; Han *et al.*, 2004).

All dsRNAs with a stem loop longer than 14 bp and 3'-overhangs or blunt ends are recognized by the nuclear export factor Exportin 5 (Exp5) for Ran-GTP dependent transfer to the cytoplasm (Yi *et al.*, 2003; Bohnsack *et al.*, 2004, Lund *et al.*, 2004, Zeng *et al.*, 2004). In the cytoplasm the pre-miRNA is released upon GTP hydrolysis and further processed by Dicer. Human Dicer cleaves pre-miRNAs near the terminal loop, releasing an about 22-nt miRNA duplex composed of two strands which are not necessarily perfectly complimentary to each other, but rather include unpaired regions called bulges (Bernstein *et al.*, 2001, Hutvagner 2001, Ketting *et al.*, 2001; Grishok *et al.*, 2001; Knight & Bass, 2001; Zhang *et al.*, 2004). Like Drosha, Dicer associates with dsRNA-binding proteins (dsRBPs). In humans, Dicers interaction partners are the two closely related proteins, TRBP and PACT. These dsRBPs are not required for correct processing activity, but were shown to have processing activation (TRBP) and repression functions (PACT) (Haase *et al.*, 2005; Chakravarthy *et al.*, 2010; Melo *et al.*, 2009; Lee *et al.*, 2006). After cleavage by Dicer the short miRNA duplex, with a 2-nt overhang on both 3'-ends, is loaded onto an Ago protein to generate the effector complex (see sections 4.1.2.3 and

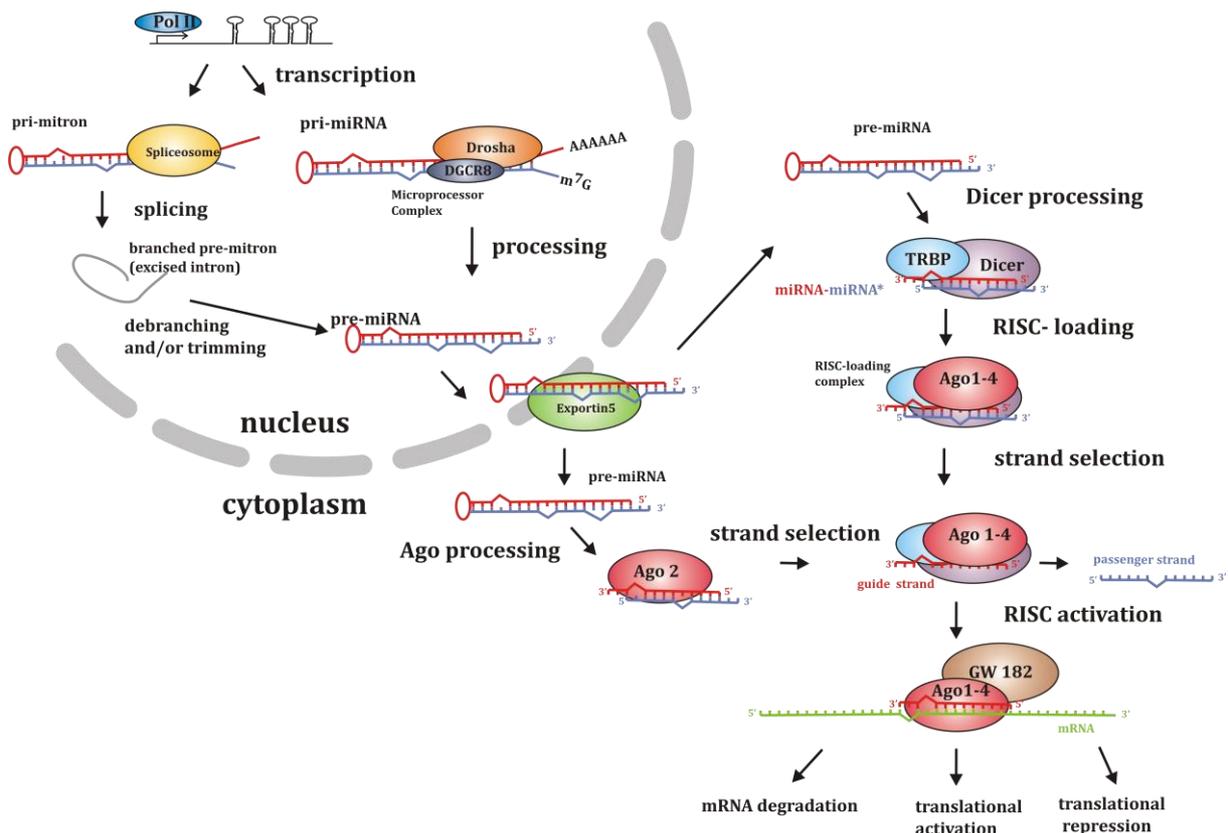
## Introduction

4.3.1). It was shown that hTRBP2 recruits the Dicer complex to an Ago protein for the formation of a RISC in humans. Thus, hTRBP2 seems to be important for dicer processing and Ago2 loading in humans (Chendrimada *et al.*, 2005; Liu *et al.*, 2007). Furthermore loading of miRNA duplexes into Ago proteins is assisted by the Hsc70-Hsp90 chaperone machinery (Iwasaki *et al.*, 2010; Maniataki & Mourelatos, 2005; Yoda *et al.*, 2010). After association with Ago, one strand of the miRNA duplex (the guide strand or miRNA) remains bound to the protein, while the other strand (passenger strand or miRNA\*) is released or degraded. A miRNA is asymmetric and the strand with a thermodynamically less stable 5'-end becomes the miRNA guide, while the passenger strand is released or degraded (Khvorova *et al.*, 2003; Schwarz *et al.*, 2003). More recent findings identified mismatches present in the seed (guide position 2-8) or 3'-mid (guide positions 12-16) of the miRNA duplexes as factors that facilitate unwinding (Kawamata *et al.*, 2009; Matranga *et al.*, 2005; Yoda *et al.*, 2010). Furthermore, the N-terminal domain of Ago proteins itself was shown to drive duplex unwinding (Kwak & Tomari, 2012; for more detail see section 4.2.2.1). Subsequently, the guide strand of the miRNA guides the Argonaute/RISC (miRISC) to its complementary mRNA target for translational repression or degradation of the targeted mRNA (Djuranovic *et al.*, 2011; Hutvagner *et al.*, 2001; Martinez *et al.*, 2002; Mourelatos *et al.*, 2002; Siomi & Siomi, 2009). The miRISC is also called major RISC, when containing the single stranded miRNA guide. The core of this miRISC contains the single stranded miRNA guide, an Ago protein and a glycine-tryptophan repeat-containing protein of 182 kDa (GW182) (Lian *et al.*, 2009; Eulalio *et al.*, 2009).

### **4.1.2.2 Alternative miRNA biogenesis pathways**

Apart from the canonical miRNA biogenesis pathway (described above) several alternative pathways have been discovered, where one or the other processing step can be circumvented. Some miRNAs are produced, independent of the Drosha dependent processing pathway, directly from pre-miRNA-resembling introns, which have been termed mitrons. Within this pathway pre-miRNAs are directly generated by the spliceosome from very short hairpin structured introns of mRNA coding genes. The splicing products are nonlinear intermediates that must be debranched by the lariat debranching enzymes and/or trimmed by exonucleases to generate a proper pre-miRNA (see Figure 2; Okamura *et al.*, 2007; Ruby *et al.*, 2007; Berezikov *et al.*, 2007; Flynt *et al.*, 2010). These pre-miRNAs can then enter the canonical biogenesis pathway as Exp-5 cargo and subsequent Dicer substrates (see Figure 2). In addition to mitrons, a few transfer RNA (tRNA) and small nucleolar RNA (snoRNA) derived fragments with miRNA-like functions have been shown to be loaded onto the RISC independently of the microprocessor complex (Ender *et al.*, 2008; Brameier *et al.*, 2011; Saraiya *et al.*, 2008; Bogerd *et al.*, 2010; Haussecker *et al.*, 2010). Generally, the derived precursors enter the miRNA biogenesis pathways as Dicer substrates (Yang & Lai, 2010; Yang & Lai, 2011). Recently, a Dicer-independent non-canonical pathway for

the generation of miRNAs was uncovered. In 2007, Diederichs and Haber found that a subset of pre-miRNAs can be cleaved by Ago2 (see Figure 2). *Mir-451* homologs from human, mouse and zebrafish are matured in a Drosha-DGCR8 dependent, but Dicer-independent manner. After cleavage of the pri-mir-451 by the Drosha-DGCR8 complex and export to the cytoplasm the pre-mir-451 was directly loaded onto Ago proteins (Diederichs and Haber, 2007; Tan *et al.*, 2009; Cheloufi *et al.*, 2010; Cifuentes *et al.*, 2010; Yang & Lai 2010; Yang & Lai, 2011).



**Figure 2: Canonical and alternative miRNA pathways.**

Detailed descriptions of the depicted processes can be found in section 4.1.2. The figure was prepared according to Jinek & Doudna, 2009 and Röther & Meister, 2011.

#### 4.1.2.3 RNA target recognition and miRNA function

miRNAs within a miRISC recognize their targets by Watson-Crick base pairing. Like siRNAs, miRNAs can induce mRNA target endonucleolytic cleavage when the miRNA sequence is fully complementary to an mRNA target, but miRNAs rarely do so (see section 4.1.1.3; Hutvagner & Zamore 2002; Yekta *et al.*, 2004; Zeng *et al.*, 2003). The vast majority of animal miRNAs recognize partially complementary binding sites, which are generally located within the 3'-UTR of the mRNA targets. Perfect complementary base pairing between nts 2-7 counting from the 5'-end of the miRNA (the 'seed' region) and the target site seem to be necessary and sufficient to trigger silencing (Bartel, 2009; Rigoutsos, 2009; Doench & Sharp, 2004). However, there are many exceptions where effective silencing could be shown without perfect seed base pairing and

## Introduction

instead perfect pairing at the 3'-end, the middle region, or different regions of the miRNA with the target RNA (Rigoutsos, 2009; Reinhart *et al.*, 2000; Tay *et al.*, 2008). Additionally, one 3'-UTR often contains multiple miRNA recognition sites and when situated in an optimal distance to each other increase silencing efficiency (Grimson *et al.*, 2007; Doench & Sharp, 2004). Moreover, a single miRNA can silence hundreds of targets, spreading silencing, however its regulation is not understood. It seems that other factors than base pairing like protein-protein interactions, also contribute to functional target interaction.

After target recognition the mechanisms of how miRNAs silence their mRNA targets are still under debate, but various experimental models suggest multiple mechanisms including target mRNA degradation, translational repression and activation of gene expression (Fabian *et al.*, 2012; Fabian *et al.*, 2010; Chekulaeva & Filipowicz, 2009; Petersen *et al.*, 2006; Filipowicz *et al.*, 2008; Pillai *et al.*, 2007). Recently, mass spectrometry and transcriptome analysis revealed that mammalian miRNAs predominantly regulate gene expression by inducing degradation of the target mRNAs (Selbach *et al.*, 2008; Baek *et al.*, 2008; Hendrickson *et al.*, 2009; Guo *et al.*, 2010). Ago proteins were shown to interact with the integral P-body protein GW182 to localize to P-bodies and this interaction was shown to be crucial for Ago proteins functioning in translational repression and degradation (Lian *et al.*, 2009; Eulalio *et al.*, 2009).

Degradation of the mRNA is induced by a miRNA causes recruitment of the CCR4- NOT or the PAN2-PAN3 deadenylation complexes by the miRISC components (Chen *et al.*, 2009; Behm-Ansmant *et al.*, 2006; Fabian *et al.*, 2009; Piao *et al.*, 2010). In this process the GW182 protein within the miRISC serve as a docking platform for both deadenylase complexes and these interactions drive GW182-assisted deadenylation (Fabian *et al.*, 2009; Chekulaeva *et al.*, 2011; Fabian *et al.*, 2011; Braun *et al.*, 2011). The subsequent removal of the poly (A) tail makes the mRNA accessible for cellular exonucleases that degrade the mRNA (Behm-Ansmant *et al.*, 2006; Braun *et al.*, 2011; Chekulaeva *et al.*, 2011; Fabian *et al.*, 2011; Giraldez *et al.*, 2006; Wu & Belasco, 2008). Some miRNAs only induce a reduction of protein levels instead of mRNA levels, and therefore lead to translational repression. The actual mechanism underlying this phenomenon is not known, but the protein production was found to be repressed by inhibition of translational initiation and elongation as well as direct proteolysis of the peptide synthesized from the target mRNA or CCR4-NOT dependent deadenylation (Carthew & Sontheimer, 2009; Wu & Belasco, 2008; Eulalio *et al.*, 2008; Filipowicz *et al.*, 2008). Finally, some miRNAs have been found to stimulate the translation of the mRNA target (Orom *et al.*, 2008; Vasudevan *et al.*, 2007; Mortensen *et al.*, 2011).

## 4.2 Proteins involved in RNAi

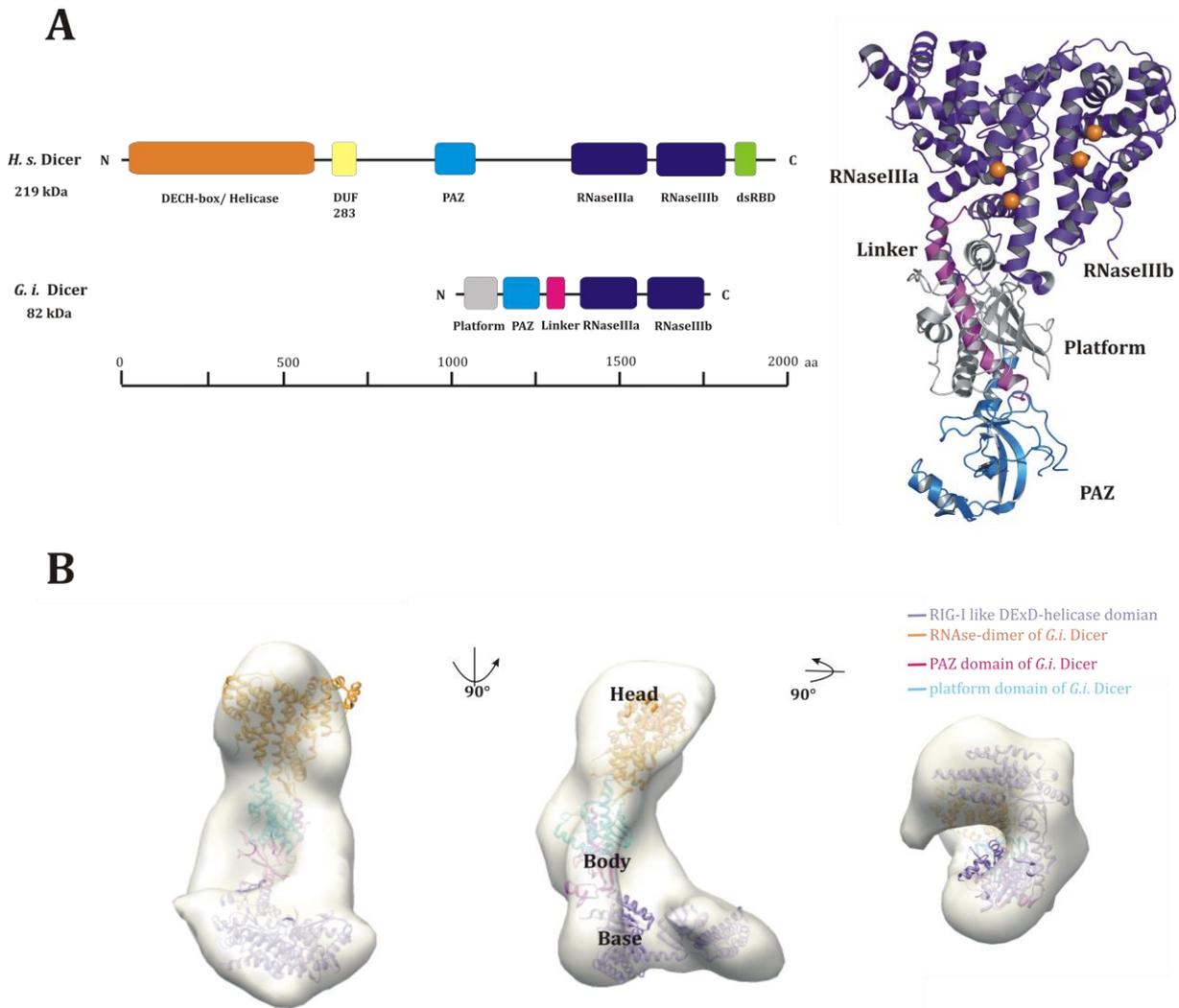
In the sections above, the origin, maturation and effects of siRNAs and miRNAs were described. As indicated there, several proteins and complexes mediate small RNA function. In the following sections proteins involved in the RISC-loading processes are described.

### 4.2.1 Dicer proteins

Dicer proteins are specialized RNase III enzymes that cleave long dsRNA and pre-miRNA into 21- to 27-nt long RNAs (Bernstein *et al.*, 2001; see also Figure 1 and Figure 2). Additionally, Dicer helps to load these RNA products into RISCs (Maniataki *et al.*, 2005; Gregory 2005; Pham 2004; Tomari *et al.*, 2004). Dicer proteins are evolutionary highly conserved and occur in nearly all eukaryotes with the remarkable exception of the baker's yeast *Saccharomyces cerevisiae*. Mammalian Dicer functions can be linked to a wide range of developmental processes including early development centromeric silencing, in embryonic stem cells, oocyte maturation, stem cell proliferation and differentiation of various tissues (Bernstein *et al.*, 2003; Kanellopoulou *et al.*, 2005; Murchison *et al.*, 2007; Tang *et al.*, 2007; Murchison *et al.*, 2005; Cuellar *et al.*, 2008; Korolov *et al.*, 2008; Zehir *et al.*, 2010). Initially RNAi was discovered as a cytoplasmic process and therefore the idea that Dicer is present solely in the cytoplasm has prevailed (Billy *et al.*, 2001; Provost *et al.*, 2002; Lee *et al.*, 2006; Daniels *et al.*, 2009). More recent findings link Dicer to heterochromatin formation and transcriptional regulation of an intergenic region and the human Dicer association with ribosomal DNA chromatin on the mitotic chromosomes made clear that Dicer can also localize and function in the nucleus (Fukagawa *et al.*, 2004; Haussecker *et al.*, 2005; Giles *et al.*, 2009; Sinkkonen *et al.*, 2010). Apart from its role in RNAi pathways, Dicer was also shown to be involved in DNA-damage response processes (Francia *et al.*, 2012; Tang *et al.*, 2012).

#### 4.2.1.1 Human Dicer

Human Dicer is a 219 kDa multidomain protein. Its catalytic core is composed of a RNaseIIIa and RNaseIIIb heterodimer. On top to these RNase III domains, human Dicer possesses a C-terminal dsRNA-binding domain (dsRBD) and a large N-terminal extension including a DExH-box helicase/ATPase domain, a domain of unknown function (DUF238) and a PIWI-Argonaute-Zwille (PAZ) domain (Figure 3). Recently, a three-dimensional EM reconstruction of human Dicer was published, revealing an L-shaped structure of the protein (see Figure 3; Lau *et al.*, 2009; Wang *et al.*, 2009b). The main investigations how Dicer may achieve dsRNA recognition, dsRNA cleavage and product repositioning are described within the following sections.



**Figure 3: Domain organization of Dicer proteins, crystal structure of *G. intestinales* Dicer and EM-structure of *H. sapiens* Dicer-TRBP2 complex.**

**A:** Domain organization of *H. sapiens* (*H.s.*) Dicer and *G. intestinales* (*G.i.*) Dicer and crystal structure of *G. intestinales* Dicer. *H. sapiens* Dicer (1922 aa, 219 kDa) consists of a RIG-I like DExD/H-box helicase domain, the DUF 283 domain, a PAZ domain, the RNaseIIIa and RNaseIIIb domains and a C-terminal dsRBD. In comparison, the *G. intestinales* Dicer (754 aa, 82 kDa) solely consists of a platform domain, a PAZ domain, a long linker helix and the RNaseIIIa and RNaseIIIb domains. **B:** A cartoon representation of the crystal structure of *G. intestinales* Dicer is shown. The color code is as in the domain organization (modified from MacRae *et al.*, 2006). **C:** Three-dimensional reconstruction of human Dicer-TRBP2 complex is shown in different orientations. The atomic models of human RIG-I like helicase domain (purple, PDB-ID: 4AP2) and *G. intestinales* Dicer (domains colored as indicated in the figure, PDB-ID: 2QVW) are docked into the structure in an orientation that the RNase III heterodimer is located in the head, the PAZ domain in the platform and the helicase domain in the base (modified from Lau *et al.*, 2009).

#### 4.2.1.2 Generation of precise siRNA and miRNA products by Dicer

The RNaseIIIa and RNaseIIIb domains of Dicer were shown to be the catalytic core for siRNA and miRNA duplex generation. For miRNA (and siRNA) biogenesis the precise selection of the cleavage site by Dicer is critical, because alterations in the cleavage site can change the abundance and/or targeting specificity of the miRNA (or siRNA). In 2006, the crystal structure of a minimal Dicer protein from the unicellular eukaryote *Giardia intestinalis* elucidated the putative structural requirements for the generation of a small RNA with a defined length (MacRae *et al.*, 2006). While *G. intestinalis* Dicer lacks the DEXD/H box-helicase domain, the DUF and the dsRBD, the PAZ domain that is connected with the tandem RNase III domains via a connector-platform domain could be crystalized (Figure 3). The PAZ domain adopts an OB-fold and specifically recognizes the 2-nt overhangs of the RNA substrate, whereas the intramolecular RNase III dimer cleaves the RNA in an Mg<sup>2+</sup>-dependent manner, generating a specific 5'-phosphate and 2-nt 3'-overhang. The unique spatial arrangement of the PAZ domain relative to the tandem RNase domains mediated by a long alpha-helical linker (connector helix) defines the specific product size. The surface of the platform creates a large positively charged region guiding the negative charged dsRNA substrate. The distance between the 3'-overhang binding pocket within the PAZ domain and the active center within the RNase III heterodimer is 65 Å, which correspond to a 25-nt dsRNA length. *G. intestinalis* Dicer measures a fixed distance from the 3'-end of the terminus, a procedure known as the 3'-counting model generating a dsRNA of 24- to 26-nts in length. It was thought that Dicer proteins from higher eukaryotes use a similar mechanism to generate RNA products with a defined length. However, the connector helix, which acts as a molecular ruler, is not conserved in these species (MacRae *et al.*, 2006; MacRae *et al.*, 2007). In 2011, Park *et al.* showed that the PAZ domain of human Dicer anchors the 3'-and the 5'-termini of the substrate RNA to determine the product size of cleavage (Park *et al.*, 2011). They identified a new basic motif within the PAZ domain critical for specific recognition of the 5'-phosphorylated end of a dsRNA substrate and could show that human Dicer determines the cleavage site mainly by the distance from the 5'-end. This procedure is known as 5'-counting rule and seems to be conserved for *D. melanogaster* Dicer1 (Park *et al.*, 2011), which in complex with the co-factor Loqs-PB processes pre-miRNAs (Jiang *et al.*, 2005; Saito *et al.*, 2005; Forstemann *et al.*, 2005). Interestingly, in *D. melanogaster* Dicer2 that acts in complex with R2D2 to generate siRNAs, the residues specific for 5'-binding are missing. Thus, the 5'-binding motif is only conserved in Dicers with pre-miRNA processing activity, demonstrating that the 5'-counting mechanism seems to be very important for efficient and accurate miRNA maturation (Park *et al.*, 2011).

The two RNase domains (RNaseIIIa and RNaseIIIb) form the catalytically active center of Dicer for small RNA production by heterodimerization. The formed heterodimer exhibits a structural similarity compared to the homodimer formed by bacterial RNaseIII domains (Zhang

## Introduction

*et al.*, 2004). Within this active center, two Mg<sup>2+</sup>-ions are coordinated by four conserved residues of each domain. Therefore it is thought that Dicer cleaves dsRNAs in a metal-dependent manner (Takeshita *et al.*, 2007). To date, no crystal structure of a Dicer-RNA complex confirming this hypothesis could be solved, but mutation of the conserved Mg<sup>2+</sup>-ion binding residues, abolished RNA-substrate cleavage (MacRae *et al.*, 2007). An additional argument for the Mg<sup>2+</sup>-dependent cleavage reaction is the fact, that the distance between the bound Mg<sup>2+</sup>-ions at the inner surface of the catalytic cleft averages 17.5 Å, fitting the width of the major groove of an RNA double helix (Review: Jinek *et al.*, 2009).

### **4.2.1.3 Involvement of additional Dicer domains in RNA processing**

Although the minimal *G. intestinalis* Dicer is capable of dsRNA processing (dicing) and can compensate the lack of functional Dicer in *S. pombe* (MacRae *et al.*, 2006), most eukaryotes Dicer proteins contain additional domains including an N-terminal DExD/D-box helicase domain, a DUF238 and a C-terminal dsRBD. These domains are important for recruiting specific RNA substrates and dsRBPs as shown in the following.

The large N-terminal helicase domain of Dicer belongs to the retinoic acid-induced gene-I (RIG-I-like) family within superfamily 2 (SF2) of RNA helicases (Fairman-Williams *et al.*, 2010). This domain is composed of three predicted domains: HEL1, HEL2i and HEL2 (Zou *et al.*, 2009). The structures of human, mouse and duck RIG-I with and without bound dsRNA could be solved recently giving new insight into the mechanism of RIG-I activation (Review: Kolakofsky *et al.*, 2012). An actual helicase activity of Dicer could not be shown to date, but several intriguing properties of this domain have been discovered while their roles in small RNA processing seem to differ from organism to organism (Lee *et al.*, 2004; Welker *et al.*, 2010; Welker *et al.*, 2011; Cenik *et al.*, 2011; Tsutsumi *et al.*, 2011). The helicase domain of human Dicer acts as an autoinhibitory module. Kinetic analysis revealed that wild type human Dicer processes hairpin RNAs substrates (pre-miRNAs) faster than long dsRNAs substrates (pre-siRNAs) *in vitro* (Ma *et al.*, 2008; Chakravarthy *et al.*, 2010). Deletion or mutations of the helicase domain enhances the processing activity of Dicer for pre-siRNA substrates and does not affect pre-miRNA substrate cleavage *in vitro* (Ma *et al.*, 2008; Chakravarthy *et al.*, 2010). It was also shown that the helicase domain inhibits catalysis and does not affect RNA-binding affinity (Ma *et al.*, 2008). Mutations in the helicase domain also lead to changes in dsRNA-processing *in vivo* (Soifer *et al.*, 2008). Additionally, the presence of the DExD/H box helicase domain correlates with a requirement for ATP by invertebrate Dicers, however mammalian Dicer processes ATP-independent (Provost *et al.*, 2002; Zhang *et al.*, 2002).

Little is known about the dsRBD of human Dicer. Deletion of the C-terminal dsRBD leads to a 1.9- to 4-fold lower activity compared to wild-type enzyme (Zhang, *et al.*, 2004). Additional dissection of Dicer confirmed the role of the dsRBD for RNA-binding and identified long dsRNAs

substrates as binding partners, but hairpin pre-miRNAs as well as RNA products cannot be bound by this domain alone (Ma *et al.*, 2012). The C-terminal portion of Dicer1 from *S. pombe* includes a dsRBD and a short motif called C33 had been subject of NMR analysis. The structure reveals a typical dsRBD fold with a novel zinc-binding motif, which is formed by residues of the dsRBD and the C33 region. A proper zinc coordination of this zinc-binding domain is required for nuclear localization of *S. pombe* Dicer1 and RNAi-mediated heterochromatin assembly. Although this domain has a high dsRNA affinity, the RNA-binding is dispensable for nucleocytoplasmic trafficking. It is thought that this new class of dsRBDs functions in nucleocytoplasmic trafficking rather than substrate binding. Interestingly, this novel zinc-binding motif is conserved in pathogenic yeast Dicers and could be a new antifungal target (Barraud *et al.*, 2011).

Initially the domain of unknown function 238 (DUF238) following the N-terminal ATPase/helicase domain of human Dicer1 was predicted to adopt a dsRBD-fold. Biochemical data suggested that this domain might be critical for RNA processing, because mutants lacking this domain lost pre-miRNA processing activity for *H. sapiens* or *D. melanogaster* Dicer1 (Lee *et al.*, 2006; Ye *et al.*, 2007). Controversially, a similar DUF238 domain deletion construct of human Dicer shows little impact on pre-siRNA or pre-miRNA cleavage activity (Ma *et al.*, 2008). In agreement with that, the structure of the DUF238 of *A. thaliana* Dicer4 uncovered a novel RNA-binding domain, which is required for protein-protein interactions rather than RNA-binding properties. At least the *A. thaliana* DUF283 domain seems to play a significant role for protein partner selection during RNA processing (Qin *et al.*, 2010), while the role of other DUF238 has been uncovered.

#### **4.2.1.4 Interaction of human Dicer with other RNAi proteins**

Dicer proteins are not only RNA-processing enzymes generating miRNAs and siRNAs, but they also transport dsRNAs and load their dsRNA-products into RISCs and thereby interacting with different proteins (see 4.3). Recently, it could be demonstrated that Ago2, PACT and TRBP2 are required of efficient Dicer functioning in human cells (Koscianska *et al.*, 2011). During RNA-processing, human Dicer was found to interact with human TRBP and PACT, which both influence its processing activity (Lee *et al.*, 2006). The TRBP2 binding site in Dicer is located between the ATPase and helicase domains and comprises amino acids 267-431 (Daniels *et al.*, 2009). The Medipal-domain of TRBP2 (aa 228-366) was identified to be responsible for Merlin, Dicer and PACT binding (see also following section 4.2.3.2; Laraki *et al.*, 2007). Residues 287-366 of TRBP2 were mapped to be sufficient for interaction with PACT and amino acids 298-366 of TRBP2 are needed for Dicer binding (Laraki *et al.*, 2007; Daniels *et al.*, 2009). Dicer is able to form a heterodimer with TRBP or PACT or a heterotrimeric complex with both proteins. An initial complex formation/interaction between Dicer and Ago (eIF2C) could be shown using

immunoprecipitations in mammals (Doi *et al.*, 2003). For some time it was believed that these PAZ domain containing proteins interact through these domains, till Tahbaz *et al.* could demonstrate that a fragment of Dicer solely containing the RNaseIIIa, RNaseIIIb heterodimer and the dsRBD is capable to bind to the PIWI lobe of Ago. The Ago-binding site is conserved within vertebrate Dicers and their interaction seems to facilitate the RNA-transfer from Dicer to Ago (Sasaki & Shimizu, 2007; Tahbaz *et al.*, 2004).

### **4.2.2 Argonaute proteins**

The Argonaute (Ago) proteins have been named after an AGO-knockout in *A. thaliana*, the leave morphology of which is reminiscent of the pelagic octopus *Argonauta argo* (Bohmert *et al.*, 1998). Ago proteins are the direct binding partners of small regulatory RNAs and have been implicated as the catalytic core of the RNAi effector complex. It has been shown that a minimal human RISC comprises an Ago protein with a small RNA (Rivas *et al.*, 2005). In the RISCs, the function of the small RNA is the recruitment of the RISC to the target mRNA and the Ago proteins recruit additional protein factors to determine the fate of the target mRNA (Pillai *et al.*, 2004; Chekulaeva *et al.*, 2009). Within the following paragraphs the structure and function of Ago proteins is outlined.

#### **4.2.2.1 The Argonaute protein family**

The Ago protein family is evolutionary highly conserved (Carmell *et al.*, 2002) and Ago proteins have been identified in all kingdoms of life, ranging from archaea (*Archeoglobus fulgidus*, *Pyrococcus furiosus*), bacteria (*Aquifex aeolicus*) to eukaryotes (from *A. thaliana* to *H. sapiens*) (Carmell, *et al.*, 2002; Hutvagner, 2008). However, the number of encoded Ago genes differs between species ranging from one in *S. pombe*, two in *N. crassa*, five in *D. melanogaster*, eight in *H. sapiens* (Sasaki *et al.*, 2003) and *M. musculus*, ten in *A. thaliana* to twenty-seven in *C. elegans* (Tolia & Joshua-Tor, 2007). The Ago proteins are classified into three analogous groups: Argonaute-like proteins, PIWI (P-element induced wimpy testis)-like proteins and the *C. elegans*-specific group 3 Argonautes, which are also called WAGOs (worm-specific Argonautes) (Yigit *et al.*, 2006). The classification is based on sequence homologies to Ago1 from *A. thaliana* and PIWI from *D. melanogaster* (Carmell *et al.*, 2002). The representatives of the WAGO clade contain no catalytically amino acid triade, which is responsible for the endonuclease activity. This consequent functional changes of the protein, led to a new family subclass (Yigit *et al.*, 2006).

#### **4.2.2.2 Human Argonaute proteins**

The eight Ago proteins encoded in the human genome can be divided into four Ago-like proteins named Ago1, Ago2, Ago3 and Ago4 and four PIWI-like proteins named HIWI, HILI,

HIWI3 and HIWI2 (Sasaki *et al.*, 2003). PIWI protein expression is mainly restricted to male germ cells and is important for germline stability and repression of mobile DNA elements (called transposons). In contrast, the four human Ago proteins are ubiquitously expressed and each Ago protein binds to different types of RNA with varying affinities resulting in the formation of various distinct RISCs. All four human Ago proteins typically associate with miRNAs mediating translational repression of target mRNAs in a similar manner without endonucleolytic cleavage (Pillai *et al.*, 2004; Wu *et al.*, 2008). All four human Ago proteins bind to a similar repertoire of miRNAs, which are enriched in 5'-uridines and have similar structural preferences like mismatches at guide positions 9-11 (Hu *et al.*, 2009; Seitz *et al.*, 2011; Yoda *et al.*, 2010; Frank *et al.*, 2010). Additionally, Ago2 binds siRNA duplexes and is able to cleave the passenger strand and assembles the guide strand into RISCs (Leuschner *et al.*, 2006; Matranga *et al.*, 2005; Miyoshi *et al.*, 2005; Rand *et al.*, 2005). siRNA directed Ago2 proteins also cleave the phosphodiester bond of a perfect complementary target RNA. The endonucleolytic cleavage by Ago2 is also called slicing and therefore Ago2 is also called slicer (Höck & Meister, 2008).

Human Agos are highly basic proteins having sizes of around 100 kDa and consist of four conserved domains: an N-terminal domain, a PIWI-Argonaute Zwiille (PAZ) domain for 3'-RNA-binding, a MID domain for 5'-RNA-binding and an RNase H like PIWI domain harboring the catalytic core that is responsible for cleavage of target RNA or the passenger strand during slicer dependent unwinding (see Figure 4 C, Song *et al.*, 2004; Ma *et al.*, 2005; Wang *et al.*, 2008; Frank *et al.*, 2010; Frank *et al.*, 2011; Frank *et al.*, 2012; Boland *et al.*, 2010; Parker *et al.*, 2005; Ma *et al.*, 2004). The PAZ domain has an OB-fold that contains a hydrophobic cavity and is capable of anchoring the 3'-end terminal two nucleotides of the guide strand of a small RNA (Ma *et al.*, 2004; Wang *et al.*, 2008). The MID domain resembles a Rossmann-fold and interacts with the 5'-end of the associated guide RNA (Ma *et al.*, 2005; Parker *et al.*, 2005). The crystal structure of the MID domain of human Ago2 in complex with different nucleotide mimics revealed a nucleotide preference (U>A>>G = C) at the 5'-end of a guide RNA (Frank *et al.*, 2010). The N-terminus of Ago proteins drives duplex unwinding during RISC assembly (Kwak, *et al.*, 2012).

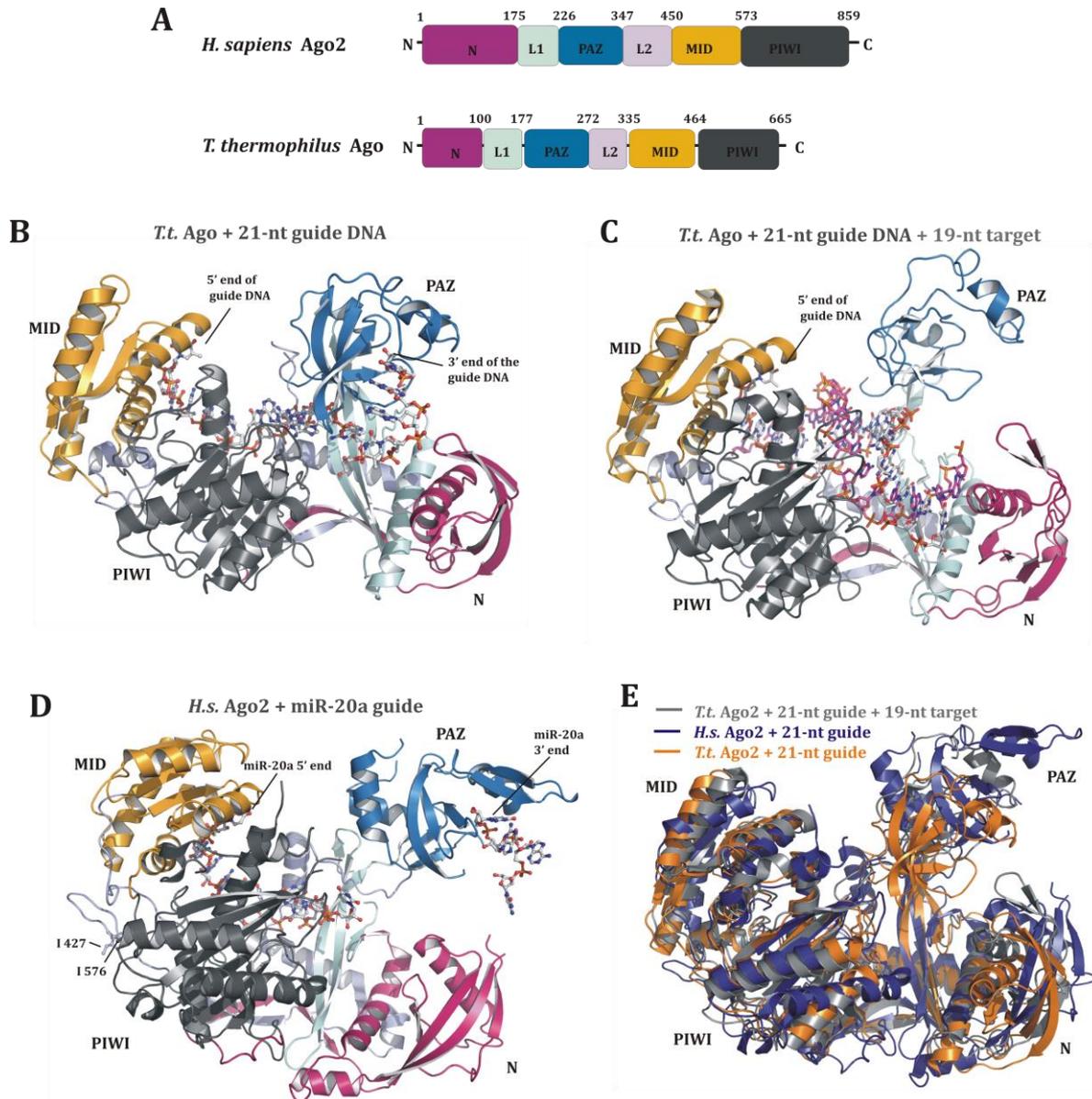
#### **4.2.2.3 Structure of Argonaute proteins**

Several crystal structures of isolated domains of eukaryotic Argonautes and structures of full length archaeal and bacterial Argonautes formed the basis of structural understanding of the Ago family until recently the crystal structure of full length human Ago2 could be solved. In contrast to eukaryotic Ago proteins, prokaryotic Agos use DNA as guides to silence genes (Ma *et al.*, 2005; Song *et al.*, 2004; Parker *et al.*, 2005; Yuan *et al.*, 2005; Wang *et al.*, 2008; Wang *et al.*, 2008; Wang *et al.*, 2009; Frank *et al.*, 2010; Boland *et al.*, 2011; Schirle and MacRae, 2012; Elkayam *et al.*, 2012). When comparing the domain organization of prokaryotic and eukaryotic Ago proteins, the isolated domains are highly similar, while eukaryotic Agos have elongated loop

## Introduction

regions (Figure 4 A). Based on the structures of full length Ago proteins, Agos have a bi-lobed structure with the N-terminal and the PAZ domain located in one lobe and the MID and the PIWI domain located in the other lobe (see Figure 4). The PAZ domain contains a specific binding pocket that anchors the 2-nt 3'-overhangs of a siRNA or a miRNA (Song *et al.*, 2003; Yan *et al.*, 2003; Ma *et al.*, 2004). The phosphorylated 5'-end of the siRNA or the miRNA is buried in a deep pocket within the MID domain (Frank *et al.*, 2010). This binding also explains the unimportance of that nucleotide for seed formation with the target mRNA. Additionally, a possible phosphorylation of Tyr529 within the 5'-binding pocket of the MID domain of human Ago2 seems to regulate guide binding (Rüdel *et al.*, 2010). The nucleotides 2-7 counting from the 5'-end of the RNA guide are of major importance for target recognition by perfect complementary base pairing (Bartel, 2009; Doench & Sharp, 2004). Nucleotides 2-6 counting from the 5'-end of the DNA guide strand (in *T.t.* Ago) or the RNA guide strand (in *H.s.* Ago2) actually contact the Ago proteins via their sugar-phosphate backbone. This interaction takes place in a positively charged tunnel that is formed by the two lobes of the Ago protein across the MID/PIWI/L2/L1 interface (Figure 4). As the backbone of the guide strand is bound by the Ago protein, the bases are free for hydrogen bonding with the complementary target mRNA. The comparison of the crystal structures of *T.t.* Ago shown in Figure 4 A and B clearly shows, that target mRNA-binding to the Ago-guide strand complex induces a movement of the lobe containing the N-terminal and the PAZ domain away from the MID and PIWI domains. Thus the Ago protein reaches a more open conformation allowing the mRNA target to bind (Wang *et al.*, 2008a; Wang *et al.*, 2008b; Wang *et al.*, 2009; Figure 4 B and C). In contrast, the human Ago2 when bound to a RNA guide already has a preformed open conformation probably allowing a mRNA target to bind (see comparison of Ago structures in Figure 4 E; Elkayam *et al.*, 2012).

The actual catalytic center of an Ago protein is located within the PIWI domain, which adopts an RNase H like fold (Parker *et al.*, 2005; Song *et al.*, 2004; Wang *et al.*, 2008). Endonucleolytic active Ago proteins cleave a target mRNA between nucleotides 10 and 11 counting from the 5'-end of the RNA guide and thereby produce an RNA fragment with a 3'-hydroxyl group and a fragment with a 5'-phosphate (Elbashir *et al.*, 2001b; Martinez & Tuschl, 2004; Schwarz *et al.*, 2004). The residues Asp579, Asp669 and His807 located within the PIWI domain were identified as the catalytic triade within human Ago2 being responsible for the endonucleolytic cleavage in an Mg<sup>2+</sup>- and ATP-dependent manner (Song *et al.*, 2004; Rivas *et al.*, 2005).



**Figure 4: Domain organizations and structures of *T. thermophilus* and *H. sapiens* Ago.**

**A:** Schematic view of the domain organization of *T. thermophilus* Ago and *H. sapiens* Ago2. **B:** Crystal structure of *T. thermophilus* Ago in complex with DNA guide strand (PDB-ID: 3DLH) **C:** Crystal structure of *T. thermophilus* Ago N478 mutant in complex with DNA guide strand and 19-nt RNA target strand (PDB-ID: 3HK2). **D:** Crystal structure of *H. sapiens* Ago2 in complex with miR-20a guide strand (PDB-ID: 4F3T). In figure B, C and D the N-terminus is colored in pink, the L1 linker region in pale-cyan, the PAZ-domain in sky-blue, the L2 linker region in light blue, the MID domain in orange, the PIWI domain is shown in grey and the nucleotides of the guide strand (colored in grey) and the target RNA strand (colored in pink) are shown in ball and stick representation. The 5'-end of the guide strand binds to the MID domain, the following nucleotides extend across the MID/PIWI/L2/L1 interface and the 3'-end nucleotides bind to the PAZ domain. **E:** Superposition of the crystal structures from C, D and E reveal that human Ago2 is the most open one with the PAZ domain moved away from the other domains.

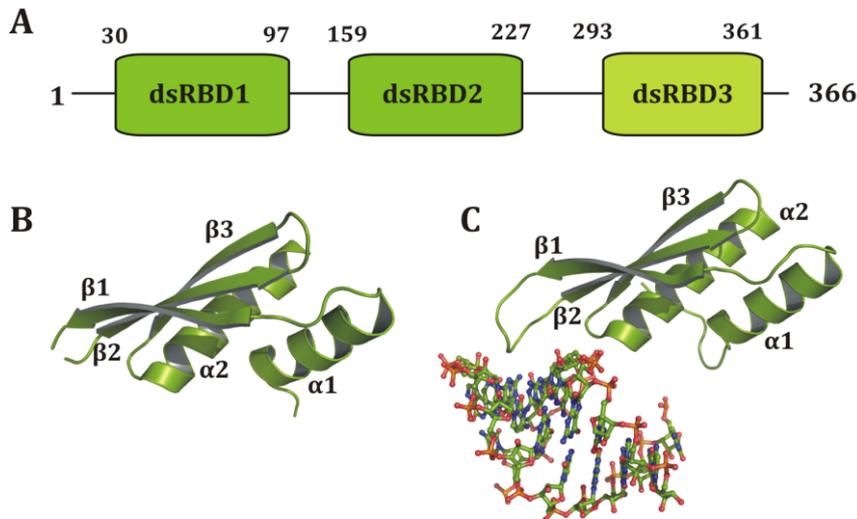
### **4.2.3 TRBP**

The transactivation response (TAR) RNA-binding protein (TRBP) has been first identified because of its ability to bind TAR RNA structures present in human immunodeficiency virus (HIV-1) transcripts and later on found to participate in RNAi (Gatignol *et al.*, 1991; Gatignol *et al.*, 2005). Within human cells two isoforms of TRBP exist, TRBP1 (345 aa) and TRBP2 (366 aa). Both TRBPs are transcribed from the *tarbp2* gene and due to alternative first exon usage differ in 21 additional amino acids within the N-terminus of TRBP2 (Bannwarth *et al.*, 2001). So far no difference in function has been reported between the two proteins (Duarte *et al.*, 2000). TRBPs are located within the cytoplasm as well as in the nucleus and the perinuclear space fulfilling several functions (Lee *et al.*, 2006; Kok *et al.*, 2007; Laraki *et al.*, 2008). Linked to its RNA and protein binding abilities and involvement in different pathways TRBP has functional activities during development, in spermatogenesis, cell growth, oncogenesis and viral replication (Lee *et al.*, 2004; Christensen *et al.*, 2007; Ong *et al.*, 2005; Zhong *et al.*, 1999; Garre *et al.*, 2010; Melo *et al.*, 2009).

#### **4.2.3.1 Structure of TRBP dsRBDs**

TRBPs belong to the family of dsRBPs (St Johnston *et al.*, 1992; Barber, 2009). It has two dsRBDs, which are important for dsRNA-binding and a half dsRBD or dsRBD type B domain, which is also called Medipal (Merlin, Dicer, protein kinase R (PKR) activator (PACT) liaison) and is mainly responsible for protein-protein interactions (Daviet *et al.*, 2000; Erard *et al.*, 1998; Gatignol *et al.*, 1991; Laraki *et al.*, 2008; see also Figure 5). TRBP was found to preferably bind to ds GC-rich sequences (Gatignol *et al.*, 1991; Lee *et al.*, 1996) and it binds siRNA duplexes independent of sequence length with high affinities (Gredell *et al.*, 2010; Kini & Walton, 2009; Parker *et al.*, 2008).

The first two consecutive dsRBDs have similar amino acid sequences and can bind dsRNA. The second dsRBD (dsRBD2) contains a characteristic Lysine-Arginine region, a so called KR-helix motif, for high affinity RNA-binding (Daviet *et al.*, 2000). The crystal structures of the first dsRBD (dsRBD1) of TRBP2, the NMR structure of dsRBD2 and the crystal structure of dsRBD2 in complex with a 10-nt long dsRNA have been solved recently (Yang *et al.*, 2010; Yamashita *et al.*, 2011). The crystal structures of dsRBD1 and dsRBD2 in complex with dsRNA are depicted in Figure 5. Each dsRBD has a common  $\alpha$ - $\beta$ - $\beta$ - $\beta$ - $\alpha$  fold with the two  $\alpha$ -helices packing against one face of the three-stranded antiparallel  $\beta$  sheet (Figure 5; Ryter & Schultz 1998; Yang *et al.*, 2010; Yamashita *et al.*, 2011).



**Figure 5: Domain organization and crystal structures of human TRBP2s dsRBD1 and dsRBD2.**

**A:** Domain organization of *H. sapiens* TRBP2. Human TRBP2 is a 366 amino acid protein containing three dsRBDs. **B and C:** The crystal structures of dsRBD1 of human TRBP2 (B, PDB-ID: 3LLH) and the crystal structure of dsRBD2 of human TRBP2 with bound dsRNA (C, PDB-ID: 3ADL) are depicted in cartoon representation and the RNA is shown in ball and stick mode.

#### 4.2.3.2 TRBP interaction with other proteins

Additionally to their ability to bind dsRNAs, the first and the second dsRBD of TRBP2 bind to PKR (protein kinase R) and PACT. The third dsRBD /the Medipal domain binds Merlin, Dicer and PACT (Lee *et al.*, 2004; Lee *et al.*, 2006; Chendrimada *et al.*, 2005; Daniels *et al.*, 2009; Haase *et al.*, 2005; Laraki *et al.*, 2008). TRBP2 inactivates PKR, which is an important contributor to the innate response to viral infections, by preventing its autophosphorylation (Benkirane *et al.*, 1997).

Within the RNAi pathway TRBP interacts with Dicer as well as with substrates and products of Dicer. Human Dicer can cleave substrate RNAs also in the absence of TRBP. However, the interaction with TRBP2 stabilizes Dicer and enhances its processivity for pre-miRNAs as well as pre-siRNAs 4- to 5-fold under multiple turnover conditions (Kok *et al.*, 2007; Ma *et al.*, 2008; Melo *et al.*, 2009; Chakravarthy *et al.*, 2010). In 2010, it was shown that the third dsRBD of TRBP2 is sufficient for Dicer interaction, but the stimulatory effect of TRBP2 on Dicer processing requires in addition the two N-terminal dsRBDs of TRBP2 (Chakravarthy *et al.*, 2010). The same authors could show that the substrate-dependent differences in dicing activity are mediated by Dicer itself as well as the structure of the substrate RNAs (Chakravarthy *et al.*, 2010). In addition to its stimulatory function, TRBP2 helps to load Dicer products into Ago containing RISCs (Castanotto *et al.*, 2007; Parker *et al.*, 2008). TRBP2 is required for RNAi function mediated by both siRNAs and miRNAs (Chendrimada *et al.*, 2005; Haase *et al.*, 2005; Rossi *et al.*, 2005; Daniels *et al.*, 2009). It acts as one of the sensors for selection and incorporation of the antisense (guide) siRNA (Castanotto *et al.*, 2007; Parker *et al.*, 2008). Additionally, TRBP2 build RNA-

independent heterodimers with PACT, which can be localized within the perinuclear space of human cells (Laraki *et al.*, 2008). Human TRBP2 and PACT directly interact with each other and associate with Dicer to stimulate its production of small interfering RNAs (Kok *et al.*, 2007).

### **4.3 RISC assembly**

Small RNAs are double stranded in origin, but as they have to base pair with their mRNA targets the two strands have to be separated. In this step one strand (the passenger strand) is discarded and the other one (the guide strand) is used to form a functional RISC. Therefore RISC assembly was defined as a process comprising RISC-loading and duplex unwinding (Review: Kawamata *et al.*, 2010). More recently, RISC assembly was redefined as a tripartite process composed of RISC-loading, wedging and unwinding (Kwak *et al.*, 2012). During RISC-loading, the Ago protein undergoes conformational opening and loads a small RNA duplex (Matranga *et al.*, 2005; Rand *et al.*, 2005; Miyoshi *et al.*, 2005; Leuschner *et al.*, 2006; Tomari *et al.*, 2007; Kawamata *et al.*, 2009; Yoda *et al.*, 2010). It is assumed that the Hsc70-Hsp90 chaperone machinery, during ATP hydrolysis, drives the conformational opening of an Ago protein to load a small RNA duplex (Kawamata *et al.*, 2010; Iki *et al.*, 2010; Iwasaki *et al.*, 2010; Miyoshi *et al.*, 2010; Johnston *et al.*, 2010). A duplex siRNA or miRNA associated with an Ago protein is called precursor RISC (pre-RISC). After loading, one end of the duplex is opened and contorted through active wedging by the N-terminal domain of the Ago protein. In case of a miRNA, the duplex is partially unwound and in case of a perfectly complementary siRNA duplex the duplex is properly positioned for passenger strand cleavage. Final unwinding, in which the passenger strand is removed slicer-independent (miRNAs) or slicer-dependent (siRNAs), forms the mature RISC containing the guide strand of the RNA duplex (Kwak *et al.*, 2012). As the RISC-loading process is of special interest within this thesis, approaches concerning the preparation of RISC-loading complexes (RLCs) made so far will be outlined in the following section in detail.

#### **4.3.1 RISC-loading complexes**

Although, Ago proteins generally need the RLCs to incorporate small RNA duplexes, they can also be loaded with single stranded small RNAs via a bypass loading (Martinez *et al.*, 2002; Rivas *et al.*, 2005; Miyoshi *et al.*, 2005; Liu *et al.*, 2004; Yoda *et al.*, 2010). The processes of RISC-loading are best understood in *D. melanogaster*, which encodes two Dicer proteins, Dicer1 that is specialized for miRNA production and Dicer2 that is selective for siRNA biogenesis. Additionally, in *D. melanogaster* miRNAs are selectively sorted into Ago1 and siRNAs into Ago2 proteins (Förstemann *et al.*, 2007; Okamura *et al.*, 2004; Tomari *et al.*, 2007). For siRNA-loading an RLC composed of Dicer2, R2D2 and Ago2 is formed (Liu *et al.*, 2003; Tomari *et al.*, 2004; Pham *et al.*, 2004). After substrate cleavage by Dicer2, the Dicer2-R2D2 heterodimer initiates the formation of an RLC and loads the siRNA duplex onto Ago2. Additionally, the Dicer2-R2D2 heterodimer

senses the thermodynamic asymmetry of the siRNA duplex. R2D2 orients a siRNA duplex by binding to the more stable end, which positions Dicer2 at the opposite, less stable end, before handing it over to Ago2 in this prearranged orientation (Tomari *et al.*, 2004). Thus Dicer2 has a dual role: substrate cleavage and Ago2 loading. Additionally, it was shown that fly Ago2 requires the Dicer2-R2D2 heterodimer for siRNA duplex initiated target cleavage, so it is thought that this heterodimer resembles the minimal fly Ago2-RLC (Kawamata *et al.*, 2010).

However, the RISC-loading process seems to be different for other Ago proteins than the fly Ago2. For example, fly Dicer1-Loqs-PB heterodimer and human Dicer-TRBP heterodimer were also assumed to be responsible for Ago loading, but neither of the Dicer proteins is actually required for loading. Additionally, it has been shown that the Dicer1-Loqs-PB complex releases miRNA duplexes and the Dicer2-R2D2 complex sorts them into Ago (Kawamata *et al.*, 2010). In contrast to *D. melanogaster* only one Dicer protein exists in human, which produces both miRNA and siRNA duplexes. The minimal human RLC is referred to be composed of Dicer, TRBP2 and Ago2 and assembles even in the absence of a dsRNA trigger (Gregory *et al.*, 2005; Maniataki & Mourelatos, 2005). Remarkably, within the complexes the protein composition differs. Maniataki and Mourelatos as well as Gregory *et al.* showed that the miRLC in mammals is composed of Dicer, miRNA-free Ago, Hsp90 and TRBP2, MacRae *et al.* described that a minimal RLC is composed of a ternary complex containing Dicer, Ago2 (RNA bound-form, which RNA was not specified by the authors) and TRBP2 (Maniataki & Mourelatos, 2005; Gregory *et al.*, 2005; MacRae *et al.*, 2008). This RLC is able to recognize a dsRNA substrate, dice it into a siRNA, load the siRNA into Ago2 and discard the passenger strand to generate a functional/ mature RISC (MacRae *et al.*, 2008). Recently, another group assembled this mammalian RISC with a miRNA duplex using cell lysate *in vitro* (Yoda *et al.*, 2010). It was shown that hTRBP2 recruits the Dicer complex to an Ago protein for the formation of a RISC in humans (Chendrimada *et al.*, 2005; Liu *et al.*, 2007). Additionally, human TRBP2 seems to be important for Dicer processing and stability (Chendrimada *et al.*, 2005; Liu *et al.*, 2007; Wang *et al.*, 2009b). The mechanisms for strand selection are mostly unclear in mammals. Within the last years the Dicer interaction partner TRBP was suggested to sense siRNA asymmetry, but these results are in contradiction to later studies where Dicer seems to have a role in strand selection (Gredell *et al.*, 2010; Noland *et al.*, 2011; Betancur & Tomari, 2012). Moreover, it was found that a subset of pre-miRNAs can directly be cleaved by Ago2 in an Dicer-independent manner (Diederichs & Haber, 2007; Tan *et al.*, 2009; Cheloufi *et al.*, 2010; Cifuentes *et al.*, 2010; Yang *et al.*, 2010). The existence of those differently reconstituted miRLCs raises the question, which of the identified complexes resembles the main or a bypass mechanism of RISC assembly (+/- pre-miRNA processing). In order to define the molecular steps of miRNA assembly of both Dicer dependent and Dicer-independent miRNAs, Liu *et al.*, assembled mammalian RISCs with a pre-miRNAs using cell lysates *in vitro* (Liu *et al.*, 2012). They identified a complex composed of human Dicer, Ago and

## Introduction

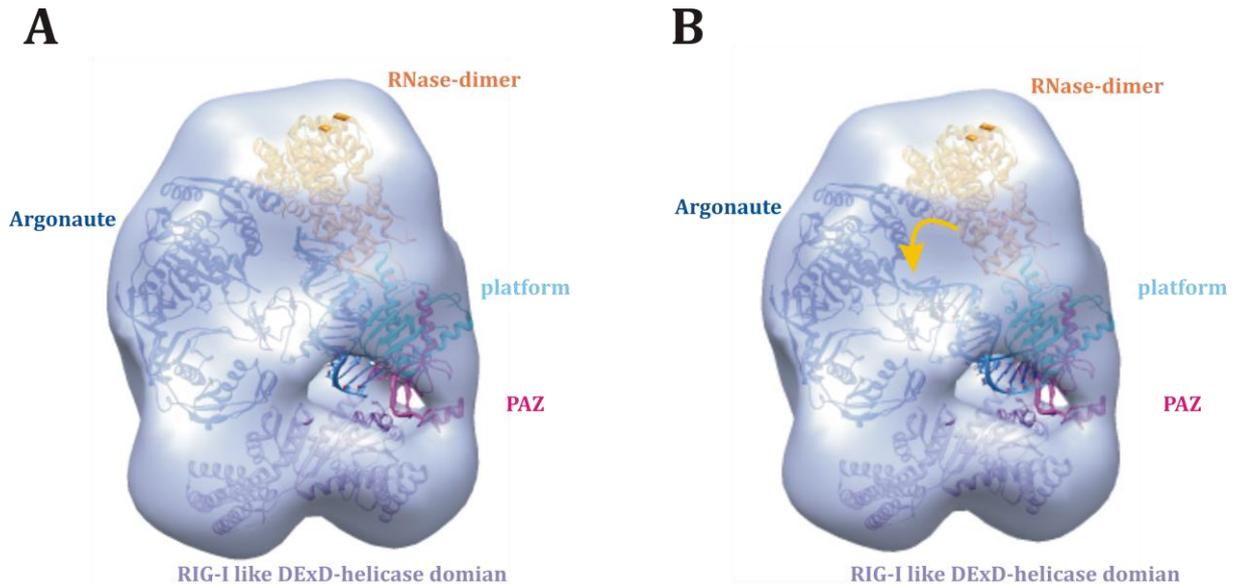
TRBP2 as miRLC (Maniataki & Mourelatos, 2005; Gregory *et al.*, 2005; MacRae *et al.*, 2008; Liu *et al.*, 2012). Liu *et al.* showed that the miRLC contains pre-miRNA, miRNA duplexes and also single stranded miRNAs; demonstrating that pre-miRNA processing and miRNA loading occur within the miRLC. In contrast, pre-miRNAs that are processed independent of Dicer by Ago2 (like pre-miR-451) are bound in a complex named miRNA precursor deposit complex (miPDC) prior to Ago2 endonucleolytic cleavage (Liu *et al.*, 2012). Pre-miR-451, like miRNA duplex, loading into Ago2 is mediated by the Hsp70 (Hsp90) machinery (Iki *et al.*, 2010; Iwasaki *et al.*, 2010; Miyoshi *et al.*, 2010; Liu *et al.*, 2012). Liu *et al.* investigated that some pre-miRNAs (like pre-miR-16 and pre-miR-21) assemble both miRLC and miPDC, whereas others (e.g., pre-miR-28 and pre-miR-27a) assemble only miRLC *in vivo*. They claimed that the miPDC serves (at least for some pre-miRNAs) as a precursor complex to miRLC. However, they confirmed that the miRLC is the primary machinery of pre-miRNA processing *in vitro* and *in vivo* (Liu *et al.*, 2012).

### **4.3.2 EM structure of the minimal human RISC-loading complex**

In 2008, MacRae *et al.* were able to stably reconstitute a minimal human RLC comprising Dicer, Ago2 and TRBP2 *in vitro*. This trimeric complex was assembled from individually expressed and purified recombinant proteins. The stoichiometry of that Dicer-Ago2-TRBP2 complex was suspected to be 1:1:1 (MacRae *et al.*, 2008). In 2009, this recombinant RLC was subjected to EM analysis and a three-dimensional reconstruction with a resolution of about 34 Å could be obtained using a GraFix prepared RLC (Wang *et al.*, 2009b; Figure 6). When compared with the reconstructions of Dicer alone, the reconstruction of the RLC contains an agglomeration of density in the middle of the molecule that connects the top of the platform with the tip of the base branch (Figure 6 A and B). Ago2 can easily be docked into this density and seems to connect the top of Dicer platform and the tip of Dicer base branch, giving the RLC a triangular architecture (Wang *et al.*, 2009b). A clear density peak that corresponds to TRBP2 could not be calculated, but it is known that TRBP2 interacts with the DExD-box helicase domain of Dicer and their data also indicated that TRBP is flexibly bound to the Dicer DExH/D box domain within the RLC (Haase *et al.*, 2005; Kok *et al.*, 2007; Daniels *et al.*, 2009; Wang *et al.*, 2009b). Remarkably, the assigned density of the DExH/D domain region of Dicer was substantially different within the RLC than in Dicer alone suggesting a structural reengagement within the DExH/D helicase domain of Dicer upon binding of TRBP2 and Ago2 (see Figure 6 A and B).

Using the EM reconstruction of the minimal RLC and docking of the atomic models of a human RIG-I-domain (purple, PDB-ID: 4T2P), the *G. intestinalis* Dicer (PDB-ID: 2QVW) and the *T. thermophilus* Ago (PDB-ID: 3HK2) as well as an siRNA duplex in the density map, the authors proposed a working model for the human RLC (Figure 6). In Figure 6 A, a proposed state of the RLC during dicing, the location of the siRNA duplex between Dicers PAZ and RNase domains, is illustrated according to the authors. In this model the siRNA is located vertically between Dicer's

RNase and Ago's PAZ domain so that the newly diced end of the siRNA (the generated 3'-2-nt overhang) can be directly engaged by the PAZ domain of Ago2. Moreover, the distance between Ago's and Dicer's PAZ domains allows a perfect accommodation of a 22-nt siRNA between them. Thus, the authors suggest a hypothetical state after transfer of the newly diced siRNAs onto the PAZ domain of Ago2 while the other end remains bound to Dicer's PAZ domain (see Figure 6 B). The flexible TRBP2 could be important to enhance transfer efficiency and/or for correct strand selection, but cannot be mapped in the EM reconstruction (Wang *et al.*, 2009b).



**Figure 6: Three-dimensional EM reconstruction and proposed working model of the minimal human RLC**

The 3D density map of the minimal human RLC is shown as a purple semitransparent isosurface. The atomic models of a human RIG-I domain (PDB-ID: 4A2P) *G. intestinalis* Dicer (PDB-ID: 2QVW) and the *T. thermophilus* Ago (PDB ID: 3HK2) are docked in the density map. Additionally, an atomic model of a siRNA duplex is docked into the reconstruction. In figure A the siRNA is located vertically between Dicer's RNase and Ago's PAZ domain so that the newly diced end of the siRNA (generated 3'-2-nt overhang) can be directly engaged by the PAZ domain of Ago2, which is shown in figure B and this movement marked with a yellow arrow (modified from Wang *et al.*, 2009b).

#### **4.4 Goal of the thesis**

RNAi is a gene silencing mechanism that can regulate gene expression at the transcriptional, posttranscriptional and translational level, thus influencing gene expression in literally every cellular process. Apart from low-resolution EM structures of the human Dicer-TRBP2 complex and the human minimal RLC other high resolution structural information of eukaryotic RISC proteins and the RLC complex itself are rare. The architecture of eukaryotic RISC proteins and complexes as well as the RNA-recognition, binding and transfer mechanisms are poorly understood. The determination of high-resolution structures of eukaryotic RNAi related proteins and complexes represent the next major steps for the understanding of RNAi on a structural level.

The first major goal of this thesis was the establishment of an expression system for the individual recombinant proteins and the preparation of the minimal human RISC-loading complex (RLC) comprising the proteins Dicer1, Argonaute2 and TRBP2. The purified 350 kDa complex RLC, should be analyzed structurally by means of macromolecular X-ray crystallography, single-particle electron microscopy (EM) as well as small-angle X-ray scattering (SAXS). The crystallization of such a big and highly flexible complex is challenging for a number of reasons. Thus, the main structural work on the RLC and its subcomplexes would preferably be carried out by single-particle electron microscopy studies. One of the main questions that is still poorly understood, is the transfer of an Dicer product RNA onto an Ago protein, thus the actual RISC-loading. A second goal within this thesis was the preparation of a siRNA duplex containing RLC and its structural analysis.

Although, the interaction sites between human Dicer and TRBP2 are known, the actual interaction surface is not known. Furthermore the actual function of human Dicers DExD/H-helicase domain, its interaction and activation with and by hTRBP2 is not understood. Therefore one aim was to prepare and structurally analyze hDicer-hTRBP2 complexes (full length and minimal complexes). The human TRBP2 is a dsRBP that contains a type B dsRBD (also called Medipal-domain) for protein-protein interaction. Structural information of several dsRBDs and type B dsRBDs are rare and therefore full length human TRBP2 as well as fragments in the presence or absence of RNA should be crystallized and the structures solved by X-ray methods.

## 5 *Material and Methods*

### 5.1 *Materials*

#### 5.1.1 *Chemicals and reagents*

All chemicals and organic substances have been purchased from the companies AppliChem (Darmstadt), Fluka (Buchs), BioRad (Munich), Merck Millipore (Darmstadt), Sigma-Aldrich (Steinheim), MWG Biotech (Munich), Oxoid (Basingstoke, England), IBA (Göttingen) and Roth (Karlsruhe) and exhibit the purity “pro analysis”. In most cases the favorable beneficial provider was chosen.

#### 5.1.2 *Equipment and instrumentation*

24-well crystallization plates	Hampton Research, Aliso Viejo (USA)
AbiPrism 3100 DNA sequencer	Applied Biosystems, Darmstadt
Agarose gel electrophoresis chamber	BioRad, Munich
Äkta Prime/Prime plus	GE Healthcare, Munich
Äkta purifier	GE Healthcare, Munich
Alchemist HT-robot	Rigaku/MSK, Japan
Amicon Ultra concentrators	Merck-Millipore, Darmstadt
Autoclave HTS 4-5-8	Zirbus, Bad Grund
Binocular	Carl Zeiss, Jena
Prettl TelStar Type Bio II A	TelStar, Frankfurt
Bradfordreagenz (5x)	BioRad, Munich
Centrifuge Allegra 21R	Beckman Coulter, Krefeld
Centrifuge Avanti J-20 XPIJA-20	Beckman Coulter, Krefeld
Centrifuge Avanti J-30 I	Beckman Coulter, Krefeld
Centrifuge Avanti JA-20	Beckman Coulter, Krefeld
Centrifugation tubes	Beckman Coulter, Krefeld
C1000 Thermal Cycler with CFX96 Real time System	BioRad, Munich
Crystal Clear Tape	Henkel, Aachen
GelDoc gel documentation device	BioRad, Munich
Gel shaker Promax 1020	Heidolph, Schwabach
Glass ware	Merck, Darmstadt
Gradient Master	BioComp Instruments, Inc., Fredericton, NB, Canada
Fraction Collector Frac-900	GE Healthcare, Munich
Heating bath	IKA, Staufen
Heating block Dri-Block CB-2A	Techne, Minneapolis, USA
Innova 4230 incubation shaker	New Brunswick Scientific, Nürtingen
Infors Unitron incubation shaker	Infors, Einsbach
Infors Multitron incubation shaker	Infors, Einsbach

## Material and Methods

Incubator Mytron	Schütt, Göttingen
JA30-centrifugation tubes	Beckman Coulter, Krefeld
Magnetic stirrer IKAMAG REO	IKA, Staufen
mar345dtb detector system	mar Research, Norderstedt
Micro fluidizer 110S	Microfluidics, Newton (USA)
mini DAWN™ TREOS multi-angle light scattering detector	WYATT Technology Europe
Minisart Syringe filter (5 µm)	Millipore, Sartorius
Mytron	Mytron, Heilbad Heiligenstadt
NanoDrop 2000	Thermo Scientific, UK
Nylon loops for protein crystals	Hampton Research, USA
Nylon loops for protein crystals	Molecular Dimensions, England
PCR-Thermocycler	Biometra, Göttingen
pH-electrode Beckman	Beckman Coulter, Krefeld
Phoenix/RE	Art Robbins, USA
Photometer	Biometra, Göttingen
Pipette tips	Sarstedt, Nümbrecht
Pipettes (adjustable)	Eppendorf, Hamburg
Pipetting helper Accu-Jet	Brand, Wertheim
Reaction vessel (0.5-2 ml)	Sarstedt, Nümbrecht
Reaction vessel (15-50 ml)	Sarstedt, Nümbrecht
Roller mixer RM5	Schütt, Göttingen
Rotation shaker	Karl-Hecht, Staufen
Rotor JA-20/JA-30.50 Ti	Beckmann-Coulter, Krefeld
Rotor JLA 8.1000	Beckmann-Coulter, Krefeld
Rotor S4180	Beckmann-Coulter, Krefeld
SDS-PAGE-System Hoefer miniVE	Amersham Pharmacia Biotech, Freiburg
Special accuracy weighing machine	Sartorius, Göttingen
Sorvall Centrifuge Evolution RC	Thermo Electron, Langenselbold, Germany
Sorvall Th 660 rotor	Thermo Electron, Langenselbold, Germany
Thermomixer comfort	Eppendorf, Hamburg
Table centrifuge 5417 R	Eppendorf, Hamburg
Table centrifuge <i>Micro centrifuge II</i>	Sylvania, Ohio, USA
Vivaspin concentrator 5301	Vivascience, Göttingen
Vortex	Schütt, Göttingen
X-ray diffractometer RU-H3R	Rigaku/Japan
X-ray diffractometer 007	Rigaku/Japan

### 5.1.3 Chromatographic columns

G200 gel filtration columns	GE Healthcare, Munich
GSH-Sepharose, 5 ml	GE Healthcare, Munich
Glutathione Sepharose 4B column material	GE Healthcare, Munich
His Trap 5ml Ni-NTA-Sepharose	GE Healthcare, Munich
HiPrep Desalting, 5ml	GE Healthcare, Munich
HiPrep Heparin, 5 ml	GE Healthcare, Munich
StrepTactin HP Sepharose, 5 ml	GE Healthcare, Munich
Superdex 75 (16/60 and 26/60)	Amersham Pharmacia Biotech, Freiburg

Superdex 200 (16/60 and 26/60)	Amersham Pharmacia Biotech, Freiburg
Superdex 200 (10/300) GL	GE Healthcare, Munich
Superose 6 (10/300) GL	GE Healthcare, Munich
Superose 12 (10/300) GL	GE Healthcare, Munich
Zeba Spin desalting columns (0.5 ml)	Thermo Scientific, Rockford (USA)

#### **5.1.4 Kit systems**

Big Dye Terminator v 1.1 Mix	Applied Biosystems, Darmstadt
NucleoSpin plasmid (Miniprep)	Macherey-Nagel, Düren
NucleoSpin Extract II (PCR purification)	Macherey-Nagel, Düren
PeqGOLD Cycle-Pure Kit	Peqlab, Erlangen
PeqGOLD Gel Extraction Kit	Peqlab, Erlangen
PeqGOLD Plasmid Miniprep Kit	Peqlab, Erlangen
QIAquick gelextraction kit	Qiagen, Hilden
QIAquick PCR purification kit	Qiagen, Hilden
QIAquick plasmid Mini, Midi, Maxi kit	Qiagen, Hilden

#### **5.1.5 Marker**

Unstained Protein Molecular Weight Marker	Fermentas, St.-Leon-Rot
PageRuler™ Prestained Protein Ladder	Fermentas, St.-Leon-Rot
Gene Ruler™ 1kb DNA ladder"	Fermentas, St.-Leon-Rot

#### **5.1.6 Enzymes and inhibitors**

ALP-protease inhibitor mix	MSB, University Göttingen
DNase	AppliChem, Darmstadt
KOD Hot start DNA polymerase	Merck Millipore (Novagene), Darmstadt
Phusion-DNA polymerase	Finnzymes, Finland
Pfu-DNA polymerase	Fermentas, St.-Leon-Rot
PreScission-protease	GE Healthcare, Munich
Restriction enzymes	Fermentas, St.-Leon-Rot
Restriction enzymes	NEB, Frankfurt
Complete protease inhibitor tablets	Roche, Kulmbach
T4-DNA ligase	Fermentas, St.-Leon-Rot
Taq-DNA polymerase	University Göttingen
TEV protease	University Göttingen

#### **5.1.7 Antibiotics with working concentrations**

Ampicillin (100µg/ml)	AppliChem, Darmstadt ; Roth, Karlsruhe
Chloramphenicol (30µg/ml)	Roth, Karlsruhe
Gentamycin (10µg/ml)	Roth, Karlsruhe
Kanamycin (50µg/ml)	Roth, Karlsruhe
Tetracycline (10µg/ml)	Roth, Karlsruhe

## Material and Methods

### 5.1.8 Antibodies

Mouse $\alpha$ -penta-His antibody	Qiagen, Hilden
Horse-radish peroxidase conjugated goat $\alpha$ -mouse	Dianova, Hamburg
Alkaline phosphatase conjugated goat $\alpha$ -mouse	Merck Millipore, Darmstadt

### 5.1.9 Cross-linker

DFDNB (1,5-Difluoro-2,4-dinitrobenzene)	Pierce (Thermo Scientific, Bonn)
EDC (1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide)	Pierce (Thermo Scientific, Bonn)
Glutaraldehyde	Pierce (Thermo Scientific, Bonn)
PMPI (N-(p-Maleimidophenyl) isocyanate )	Pierce (Thermo Scientific, Bonn)
SMCC (Succinimidyl 4-(N maleimidomethyl) cyclohexane-1-carboxylate)	Pierce (Thermo Scientific, Bonn)

### 5.1.10 Bacterial strains and cell lines

<i>E. coli</i> Arctic Express	MSB, University Göttingen
<i>E. coli</i> BL21 (DE3)	MSB, University Göttingen
<i>E. coli</i> BL21 (DE3) Star	MSB, University Göttingen
<i>E. coli</i> BL21 (DE3) RIL Codon plus	MSB, University Göttingen
<i>E. coli</i> DH10Bac	MSB, University Göttingen
<i>E. coli</i> DH5 $\alpha$	MSB, University Göttingen
<i>E. coli</i> Rosetta 2 (DE3)	MSB, University Göttingen
<i>E. coli</i> SoluBL21 (DE3)	MSB, University Göttingen
<i>E. coli</i> Top10	Invitrogen
<i>E. coli</i> Top 10 (for StarGate cloning)	IBA, Göttingen
<i>E. coli</i> XL1-Blue	MSB, University Göttingen
Sf21 insect cells	Invitrogen
Sf9 insect cells	Invitrogen
High Five insect cells	Invitrogen

### 5.1.11 DNA-oligonucleotides

DNA-oligonucleotides were ordered from MWG (Munich) and Sigma-Aldrich (Steinheim) and are depicted within the appendix (Table 3).

**5.1.12 Plasmids**

<b>Plasmid</b>	<b>Insert</b>	<b>Cloning-sites</b>	<b>Affinity Sequence</b>	<b>Expression</b>
<b>hAgo2 (NP_036286.2, NM_012154.3)</b>				
pGex6P1-Ago2	<i>H.s.</i> Ago2 gene sequence for aa 1-859	<i>EcoRI/NotI</i>	N-GST	<i>E. coli</i> BL21 (DE3) RIL <sup>+</sup>
pFastBacHTb-Ago2	<i>H.s.</i> Ago2 gene sequence for aa 1-859	<i>BamHI/HindI</i> II	N-His	Transfer vector for expression in Sf9 or Hi5 cells
<b>hTRBP2 (NP_599150.1, NM_134323.1)</b>				
pGex6P1-TRBP2 2	<i>H.s.</i> TRBP2 gene sequence for aa 1-366	<i>BamHI/NotI</i>	N-GST	<i>E. coli</i> BL21(DE3) Star
pGex6P3-T1	<i>H.s.</i> TRBP2 gene sequence for aa 157-366	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3) Star
pGex6P3-T2	<i>H.s.</i> TRBP2 gene sequence for aa 226-366	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3) Star
pGex6P3-T3	<i>H.s.</i> TRBP2 gene sequence for aa 291-366	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3) Star
pGex6P3-T4	<i>H.s.</i> TRBP2 gene sequence for aa 22-99	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3) Star
pGex6P3-T5	<i>H.s.</i> TRBP2 gene sequence for aa 22-227	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3) Star
pGex6P3-T7	<i>H.s.</i> TRBP2 gene sequence for aa 157-227	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3) Star
<b>hDicer (NP_803187.1, NM_177438.2)</b>				
pMACSiBac1-Dicer	<i>H.s.</i> Dicer gene sequence for aa 1-1922	<i>FseI/AcsI</i>	none	Only used as template DNA
pFastBacHTb-hDicer	<i>H.s.</i> Dicer gene sequence for aa 1-1922	<i>Sall/NotI</i>	N-His	Transfer vector for expression in Sf9 insect cells
<b>hDicer (optimized synthetic gene (see appendix) for NP_803187.1)</b>				
0900381_Dicer1_pMA	<i>H.s.</i> Dicer synthetic gene for aa1-1922 additional C-terminal Strep-tag	<i>BamHI/XbaI</i>	C-Strep	Only used as template DNA
pFBDM-hDicer	<i>H.s.</i> Dicer synthetic gene for aa1-1922	<i>BamHI/XbaI</i>	C- Strep	Transfer vector for protein expression in Sf9 insect cells
pFastBacHTb-hDicer	<i>H.s.</i> Dicer synthetic gene for aa1-1922	<i>BamHI/XbaI</i>	N-His and C- Strep	Transfer vector for protein expression in Sf9 insect cells
pGex6P3-DTB1	<i>H.s.</i> Dicer synthetic gene for aa 261-432	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3)Star
pGex6P3-DTB2	<i>H.s.</i> Dicer synthetic gene for aa 261-439	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3)Star
pGex6P3-DTB3	<i>H.s.</i> Dicer synthetic gene for aa 261-414	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3)Star
pGex6P3-DTB4	<i>H.s.</i> Dicer synthetic gene for aa 261-415	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3)Star
pGex6P3-DTB5	<i>H.s.</i> Dicer synthetic gene for aa 264-414	<i>BamHI/XhoI</i>	N-GST	<i>E. coli</i> BL21(DE3)Star

**5.1.13 RNAs**

All antisense siRNAs and siRNA duplexes used in this study were supplied by Roche Kulmbach and are listed in the following table. Additionally *in vitro* transcribed RNAs for activity assays (5.2.1.11; 5.2.1.12) are listed:

<b>Name</b>	<b>Sequence</b>	<b>Origin</b>
siRNA duplex1	GCCCGACAACCACUACCUGAG GACGGGCUGUUGGUGAUGGAC	Roche, Kulmbach
siRNA duplex2	ACAUGAAGCAGCAGCAGCUACU GGUGUACUUCGUCGUGCUGAU	Roche, Kulmbach
Antisense siRNA1	GACGGGCUGUUGGUGAUGGAC	Roche, Kulmbach
Antisense siRNA2 (let7A)	pUGAGGUAGUAGGUUGUAUAGUU	Roche, Kulmbach
Antisense siRNA3 (let7A)	pUUGUCUCUGGUCCUACUU	Roche, Kulmbach
Antisense siRNA4 (complementary to lucII, 21.27)	UCGAAGUAUUCGCGUACGUdT	G. Meister, Munich
40-nt RNA	CCUACGGAGCUGGUGGCGUAGGUUUUUUU UUUUUUUUUUUUUU	Andreas Schmitt, MSB
Pre-miR27a	TTAATACGACTCACTATAGCTGAGGAG CAGGGCTTAGCTGCTTGTGAGCAGGGTCCA CACCAAGTAGTGTT <b>CACAGTGGCTAAGTT</b> <b>CCGCCCCCAGC</b>	G. Meister, Munich
lucI II ( or 86.1)	(Martinez <i>et al.</i> , 2002; Meister <i>et al.</i> , 2004)	G. Meister, Munich

**5.1.14 Solutions for crystallization screens and fluorescence-based thermal shift assay**

All solutions for the crystallization screens in 24-well format were prepared manually. The crystallization solutions in 96-well format except from the MIDAS screen were prepared by the Alchemist HT-robot. The MIDAS screen and the screens for fluorescence-based thermal shift assay were bought ready to use from the manufacturer.

**Crystallization screens in 24-well format**

Clear Strategy Screen	Molecular Dimensions, UK
Crystal Screen 1&2	Hampton Research, USA
Crystal Screen lite 1&2	Hampton Research, USA
Crystal Screen PEG/ION	Hampton Research, USA
Footprint Screen 1-3	Stura <i>et al.</i> , 1992
JB Screens 1-10	Jena Bioscience, Jena
Magic Screen 1-3	Biogenova, Canada
Structure Screen	Molecular Dimensions, UK

**Crystallization screens in 96-well format**

Ammoniumsulfate Suite	Qiagen, Hilden
JB NucPro	Jena Bioscience, Jena
JB Screens (1,2,4,5)	Jena Bioscience, Jena
JB Screens (6,7,8,10)	Jena Bioscience, Jena
JCSG	Newman <i>et al.</i> , 2005
Morpheus	Molecular Dimensions, UK
MIDAS	Molecular Dimensions, UK
NatrixHT	Hampton Research, USA
PGA	Molecular Dimensions, UK
ProPlex	Radaev <i>et al.</i> , 2006

**Screens for fluorescence-based thermal shift assay (in 96-well format)**

Hampton Additive Screen	Hampton Research, USA
pHAT Screen	Emerald Biostructures, USA

**5.1.15 Computer programs**

Astra 5/Astra 6	WYATT Technology Europe
CCP4 suite	Winn <i>et al.</i> , 2011
CCD	Mooij <i>et al.</i> , 2009
CDD	Marchler-Bauer <i>et al.</i> , 2011
Chimera	<a href="http://www.cgl.ucsf.edu/chimera/">http://www.cgl.ucsf.edu/chimera/</a>
ClustalW	Larkin <i>et al.</i> , 2007
Coot	Emsley <i>et al.</i> , 2010
ESPrpt	Gouet <i>et al.</i> , 1999; Gouet <i>et al.</i> , 2003
Expasy, ProtParam	<a href="http://web.expasy.org/protparam">http://web.expasy.org/protparam</a>
MaxStaf software	Developed within the group of Prof. Holger Stark (MPI, Göttingen) <a href="http://molprobity.biochem.duke.edu/">http://molprobity.biochem.duke.edu/</a> Davis <i>et al.</i> , 2004; Davis <i>et al.</i> , 2007
MolProbity	McCoy <i>et al.</i> , 2007
Phaser	Adams <i>et al.</i> , 2010
Phoenix suite	Bryson <i>et al.</i> , 2005
PSIPRED	Kabsch, 2010
XDS	

### **5.1.16 Cell culture medium and transfection reagents**

The serum free medium Sf-900 III SFM for Sf9 and Sf21 cell lines and Express Five ® SFM for the High Five™ cell line were purchased from Invitrogen/Life Technologies (Darmstadt). The transfection reagent Fugene 6 was obtained from Roche (Kulmbach).

## **5.2 Methods**

### **5.2.1 Molecular biological methods**

#### **5.2.1.1 General cloning strategies**

The plasmid pMacSiBac containing the gene DICER1 coding for human Dicer (NP\_085124.2) as well as the pGex6P3 vectors containing the gene sequence encoding full length human Ago2 (NP\_036286.2) or full length human TRBP2 (NP\_599150.1) were kindly provided by G. Meister (5.1.12). Coding sequences for hDicer or hAgo2 were amplified by PCR and cloned into pFastBacHtb transfer vectors via *Sall/NotI* or *BamHI/HindIII* restriction sites, respectively (5.2.1.2; 5.2.1.4; 5.2.1.5). Successful cloning was verified by DNA sequencing (5.2.1.6) and the coding sequences transferred into the bacmid DNA for expression using the baculovirus expression system (5.2.2.6).

Since the expression level of hDicer in insect cells using the original gene sequence was low, a sequence optimized for expression in Sf9 cells was ordered from Genearth. This sequence was free of any commercial restriction sites, contains a C-terminal Strep-tag and had a N-terminal *BamHI* and a C-terminal *XbaI* restriction site (for sequence see 10). The hDicer coding sequence was transferred into the pFBDM, the pFastBac and the pFastBacHtb vector using restriction digestion (5.2.1.4). Successful cloning was verified by DNA sequencing (5.2.1.6) and the coding sequences transferred into the bacmid DNA for expression using the baculovirus expression system (5.2.2.6). Other constructs encoding portions of hDicer were amplified from the Genearth plasmid using PCR and cloned into the pFastBacHtb or pGex6P3-vector via suitable restriction sites (5.1.11; 5.2.1.2; 5.2.1.4; 5.2.1.5; 5.2.1.6).

All constructs encoding portions of hTRBP2 were amplified from the pGex6P3-hTRBP2 plasmid and cloned into pGex6P3 vectors via *BamHI/XhoI* restriction sites, allowing the production of N-terminal GST-tagged fusion proteins (5.1.12; 5.2.1.2; 5.2.1.4; 5.2.1.6).

For co-expression of hDicer-DEXD box domain fragment and hTRBP2 or C-terminal hTRBP2-dsRDB3 fragment appropriate gene constructs were cloned into the pETDuet-1 vector. To equip the hDicer-DEXD box domain fragments with a N-terminal Strep-tag, hDicer constructs were first cloned into a modified pET-22b vector via *BamHI* and *XhoI* restriction sites. The hTRBP2 or C-terminal hTRBP2 fragments were amplified by PCR from the pGex6P3-hTRBP2 plasmid and

cloned into the pETDuet-1 vector via *Sall/HindIII* restriction sites (5.2.1.2; 5.2.1.4). The hDicer-DExD box domain fragments with N-terminal Strep-tag coding region were finally cloned via *NdeI* and *XhoI* restriction sites into the pETDuet-1 containing hTRBP2 or hTRBP2 C-terminal constructs (5.1.11; 5.1.12; 5.2.1.2; 5.2.1.4; 5.2.1.5; 5.2.1.6).

### 5.2.1.2 Polymerase chain reaction

The polymerase chain reaction (PCR) was used to amplify a gene of interest from cDNA, from another plasmid and to verify the successful insertion of a gene of interest into a plasmid. Within the following paragraph DNA polymerases, a standard PCR reaction and a typical PCR protocol used in this work are listed.

#### Thermostable DNA polymerases

Polymerase	Organism	Time and temperature optima		
		Denaturation	Annealing	Elongation
KOD Hot Start	<i>Thermococcus kodakaraensis</i>	95 °C; 15 sec	T <sub>m</sub> ; 15 sec	72°C; 10 sec /1 kb
Phusion	<i>Pyrococcus furiosus</i>	98 °C; 20 sec	T <sub>m</sub> ; 20 sec	72°C; 30 sec /1 kb
<i>Pfu</i>	<i>Pyrococcus furiosus</i>	98 °C ; 30 sec	T <sub>m</sub> -5 °C; 30 sec	72°C;120 sec/1 kb
<i>Taq</i>	<i>Thermus aquaticus</i>	98 °C ; 30 sec	T <sub>m</sub> -5 °C; 30 sec	72°C; 60 sec /1 kb

#### PCR reaction

1-150 ng	template-DNA
1x	PCR-Puffer
10 mM	of dATP, dGTP, dCTP, dTTP
0-10 % (v/v)	DMSO
0-2mM	MgCl <sub>2</sub>
10 pmol	forward primer
10 pmol	reverse primer
1-2 U	DNA-polymerase
ad 50 µl	ddH <sub>2</sub> O

#### PCR-program

Step	Temperature	Time	
Initial denaturation	95- 98 °C	5 minutes	
Denaturation	95- 98 °C	15-30 seconds	20-45 cycles
<i>primer</i> -hybridization	x °C	15-30 seconds	
Elongation	y °C	z seconds	
Final elongation	y °C	10 minutes	

For x a temperature 3 to 5 °C below the melting temperature (T<sub>m</sub>) or the T<sub>m</sub> of both primer according to the table of DNA polymerases depicted above was used. The temperature (y) of the elongation as well as the duration of elongation (z) depended on the used DNA polymerase (see table of DNA polymerases above).

### 5.2.1.3 DNA isolation

For the purification of PCR-products the kit systems QIAquick PCR Purification Kit (Qiagen, Hilden), the NucleoSpin Extract (Macherey-Nagel, Düren) or the peq Gold Cycle Pure Kit (Peqlab,

## Material and Methods

Erlangen) were used according to the manufactures instructions. Extraction of DNA out of agarose gels was performed according to the manufactures instructions of the used kit systems QIAquick Gel Extraction Kit (Qiagen, Hilden), the NucleoSpin Extract (Macherey-Nagel, Düren) or the peq Gold Gel Extraction Kit (Peqlab, Erlangen).

### **5.2.1.4 DNA restriction digestion**

The restriction digestions was carried out according to the manufactures instructions for the respective restriction enzymes (Fermentas/Thermo Scientific or NEB). Generally the incubation time was 1 hour and the reaction was stopped by heat shock at 95 °C for 10-20 minutes or through separation of the reaction products using agarose gel electrophoresis (5.2.1.8).

### **5.2.1.5 DNA ligation**

Restriction digested DNA of a gene of interest was ligated into a linearized restriction digested plasmid (5.2.1.4). For ligation, the T4-DNA ligase (5.1.6) was used according to manufacturers instructions. In general a 3-5 times molar excess of insert to plasmid was used and the reaction was carried out at 4 °C over night.

### **5.2.1.6 DNA Sequencing**

All DNA constructs used and prepared during this work were analyzed for the correct sequence of the respective inserts by DNA sequencing. A standard DNA sequencing reaction contained 200-400 ng plasmid, 10 pmol sequencing primer, 1.5 µl Big Dye Terminator v 1.1 mix, 1.5 µl sequencing buffer adjusted with double distilled water to a final volume of 10 µl. The temperature protocol included an initial denaturation step at 96 °C for 2 minutes, another denaturation step at 95 °C for 5 sec, a primer annealing step at 55 °C for 15 seconds and an elongation step at 60 °C for 4 minutes. The denaturation, annealing and elongation steps were repeated 25 times. Afterwards the DNA was precipitated by addition of 2 µl 3 M NaOAc, 2 µl 125 mM EDTA and 50 µl ethanol (98% v/v). The precipitated DNA was pelletized by centrifugation (20000x g, 20 min), washed with 70 µl ethanol (70% v/v) and dried at room temperature. The DNA pellet was resuspended in 100 % (v/v) formamide and stored at -20 °C until sequencing was performed by Andreas Nolte (Department of Developmental Biochemistry, University of Göttingen).

### **5.2.1.7 Concentration determination of DNA and RNA**

Nucleic acids have absorption maxima at 260 nm and therefore DNA or RNA concentration of a solution was determined at 260 nm. Absorption of 1 at 260 nm corresponds to 50 µg/ml double stranded DNA (= 0.15 mM), 33 µg/ml single stranded DNA (=0.1 mM) and 33 µg/ml single stranded RNA (=0.11 mM). The purity of the nucleic acid solution was determined through the

$A_{260}/A_{280}$  ratio. Pure DNA or RNA preparations have ratios of 1.8 and 2.2 respectively. Solutions containing protein or phenol impurities have a lower value.

#### **5.2.1.8 Agarose gel electrophoresis and visualization of DNA and RNA**

DNA- and RNA-fragments were analyzed and separated by size using agarose gel electrophoresis. 1% (w/v) agarose was boiled in 1× TAE and transferred to a gel chamber with a comb. After the gel was polymerized, it was transferred to an electrophoretic chamber filled with 1× TAE and the DNA/RNA samples supplemented with loading dye are loaded onto the gel. A voltage of 80 V was applied to the gel for 45 to 90 minutes. After separation of the DNA or RNA fragments, the gel was transferred into a solution containing 1 µg/ml ethidium bromide or GelRed and incubated for 5-10 minutes. The visualization of bands under UV light was carried out using the gel documentation system GelDoc EQ system from BioRad.

<b>50 × TAE</b>		<b>6 × loading dye</b>	
2M	Tris acetate, pH 8.0	0.25%	Bromphenolblue
50 mM	EDTA	0.25%	Xylene cyanol FF
		30% (v/v)	Glycerol

#### **5.2.1.9 Electrophoretic mobility shift assay (EMSA)**

The electrophoretic mobility shift assay (EMSA) was used to study antisense siRNAs and double stranded siRNAs binding to RNA-binding proteins. The reactions were incubated for 30 minutes in binding buffer (generally the gel filtration buffer of each protein supplemented with 2 mM MgCl<sub>2</sub>) on ice. The reaction was mixed with 6x loading dye and subjected to agarose gel electrophoresis and visualized as described above (5.2.1.8).

#### **5.2.1.10 Urea polyacrylamide gel electrophoresis (urea-PAGE) and visualization of RNA**

RNA samples were analyzed by urea polyacrylamide gel electrophoresis (urea-PAGE). A typical gel was composed of 8% polyacrylamide and 8 M Urea in 1× TBE and the gel was polymerized by addition of 1% (v/v) APS and 1% (v/v) TEMED. The samples were dissolved in 2× RNA loading buffer (90% (v/v) formamide, 0.025% (w/v) xylenecyanol, 0.025% (w/v) bromphenolblue) and heated to 95 °C for 5 minutes. Each gel was pre-run 10-15 minutes at 35 mA in 1× TBE and the pockets rinsed thoroughly with running buffer. After loading, the gel was run in 1× TBE at 35 mA for 1-2 hours. For visualization of the RNA, the gel was transferred into a 1× TBE solution containing 1 µg/ml ethidium bromide or GelRed and incubated for 1 minute. The visualization of bands under UV light was carried out in a gel documentation system GelDoc EQ system of BioRad (BioRad, Munich).

## Material and Methods

All Urea gels for hDicer and RISC assays within the laboratory of Prof. Dr. G. Meister (MPI, Munich) were prepared using the SequaGel Sequencing System Kit (National diagnostics). For the RISC assay (Argonaute assay) 8% acrylamide gels and for the hDicer assay 10 and 15% acrylamide gels were prepared. Each gel was pre-run 10-15 minutes at 300 V in 1× TBE and the pockets rinsed thoroughly with the running buffer. All samples were dissolved in 2× RNA loading buffer (90% (v/v) formamide, 0.025% (w/v) xylene cyanol, 0.025% (w/v) bromphenolblue) and heated to 95 °C for 5 minutes prior to loading. For the RISC assay the gel was run at 65 W and for the Dicer assay the gels were run at 300 V.

<b>10× TBE</b>		<b>2× RNA loading buffer</b>	
2 M	Tris acetate, pH 8.0	0.025%	Bromphenolblue
50 mM	EDTA	0.025%	Xylene-cyanol FF
		90% (v/v)	Formamide

### **5.2.1.11 *In vitro* Dicer assay**

To determine human Dicer activity an *in vitro* transcribed and radioactively labeled pri-miR-27a substrate was used (Meister *et al.*, 2005; Landthaler *et al.*, 2004). The assay was performed together with Anne Frohn in the laboratory of Prof. Dr. G. Meister in Munich. For *in vitro* transcription of the  $\alpha$ [<sup>32</sup>P] UTP labeled pre-miR-27a a 40  $\mu$ l reaction consisting of 3  $\mu$ l PCR template (Landthaler *et al.*, 2004), 8  $\mu$ l 5x NTP Mix (A/C/G/U = 5/5/8/0.1 mM), 8  $\mu$ l 5x T7 RNA polymerase buffer (40 mM Tris/HCl pH 7.9; 6 mM MgCl<sub>2</sub>, 2 mM spermidine, 10 mM NaCl), 0.2  $\mu$ l DTT (1M), 1  $\mu$ l T7 RNA polymerase, 5  $\mu$ l  $\alpha$ [<sup>32</sup>P] UTP (10  $\mu$ Ci/ $\mu$ l) and 14.8  $\mu$ l autoclaved dd H<sub>2</sub>O was incubated at 37°C for 2 hours. 40  $\mu$ l of RNA loading buffer (5.2.1.10) was added, the solution incubated at 95°C for 5 minutes and separated by 10% urea-PAGE (5.2.1.10). The RNA was detected by autoradiography, cutted out and eluted from the gel by addition of 400  $\mu$ l RNA elution buffer (300 mM NaCl, 2 mM EDTA) shaking at 1400 rpm and 4°C overnight. The RNA was precipitated by addition of 0.1 volume of NaOAc (pH 5.2) and 2.5 volumes of 100% ethanol at -20°C over night. The RNA was pelleted by centrifugation at 13000 rpm for 1 hour at 4°C. The RNA pellet was washed with 70% ethanol and dissolved in 20  $\mu$ l autoclaved ddH<sub>2</sub>O. The concentration of the RNA was measured using a NanoDrop and the RNA was stored at -20 °C before usage (5.2.1.7).

For the cleavage reaction 0.25  $\mu$ g to 5  $\mu$ g recombinant human N-hDicer was incubated with 2 Bq/cm<sup>2</sup> *in vitro* transcribed pri-mi-27a RNA substrate in 20  $\mu$ l PBS supplemented with 5 mM ATP, 7.5 mM MgCl<sub>2</sub> and 10 U/ml RiboLock (Fermentas) at 37 °C for 1 hour. The reaction was stopped by addition of 40  $\mu$ g proteinase K in 200  $\mu$ l proteinase K buffer (300 mM NaCl, 200 mM Tris/HCl pH 7.5, 25 mM EDTA, 2% SDS). The protein digestion reaction was incubated at 65 °C for 15 minutes. The RNA was extracted using 200  $\mu$ l Phenol/Chlorophorm/Isoamylalcohol (25:24:1) solution. For subsequent RNA precipitation, the aqueous phase was mixed with 0.1

volume of NaOAc (pH 5.2) and 2.5 volumes of 100% ethanol and incubated at -20°C over night. The RNA was pelleted by centrifugation at 17000× g for 1 hour at 4°C, the pellet air dried at room temperature, dissolved in autoclaved ddH<sub>2</sub>O and separated by 15% urea-PAGE (5.2.1.10) and signals were detected by autoradiography.

#### **5.2.1.12 *In vitro* RISC cleavage assay**

To assay for the endonucleolytic cleavage of a target RNA (slicer activity) of recombinant human N-His<sub>6</sub>-Ago2 and hAgo2, an *in vitro* RISC cleavage assay was performed together with Anne Frohn in the laboratory of Prof. Dr. G. Meister in Munich. As target RNA a  $\alpha$ [<sup>32</sup>P] GTP-cap labeled 170 nucleotide long RNA corresponding to the firefly luciferase sequence called lucII (also named 86.1) was used (5.1.13; Martinez *et al.*, 2002; Meister *et al.*, 2004). This RNA contains a sequence that is perfectly complementary to a specific single stranded small RNA that was used for recombinant Argonaute loading (antisense siRNA4 in 5.1.13, Meister *et al.*, 2004)

The target mRNA was *in vitro* transcribed by incubating 5  $\mu$ l PCR template (Martinez *et al.*, 2002; Meister *et al.*, 2004) with 20  $\mu$ l 5× NTP Mix (A/C/G/U = 5/5/8/2 mM), 20  $\mu$ l 5× T7 RNA polymerase buffer (40 mM Tris/HCl pH 7.9, 6 mM MgCl<sub>2</sub>, 2 mM spermidine, 10 mM NaCl), 0.5  $\mu$ l DTT (1M), 1  $\mu$ l T7 RNA polymerase and 53.5  $\mu$ l autoclaved dd H<sub>2</sub>O at 37 °C for 2 hours. 100  $\mu$ l of RNA loading buffer (90% (v/v) formamide, 0.025% (w/v) xylene-cyanol, 0.025% (w/v) bromphenolblue) was added, the solution incubated at 95 °C for 5 minutes and separated by 8 % urea-PAGE (5.2.1.10). The RNA was detected by UV shadowing, cut out from the gel and eluted from the gel by addition of 400  $\mu$ l RNA elution buffer (300 mM NaCl, 2 mM EDTA) shaking at 1400 rpm and 4 °C overnight. The RNA was precipitated by addition of 0.1 volume of NaOAc (pH 5.2) and 2.5 volumes of 100% ethanol at -20 °C over night. The RNA was pelleted by centrifugation at 13000 rpm for 1 hour at 4 °C. The RNA pellet was washed with 70% ethanol and dissolved in 20  $\mu$ l autoclaved ddH<sub>2</sub>O. The concentration of the RNA was measured using a NanoDrop by detection of 260 nm absorption and the RNA stored at -20 °C before further usage (5.2.1.7).

For  $\alpha$ [<sup>32</sup>P] GTP-cap labeling 20 pmol lucII-RNA (9  $\mu$ l) were incubated with 2  $\mu$ l  $\alpha$ [<sup>32</sup>P] GTP (3000 Ci/mmol), 2  $\mu$ l Guanylyltransferase, 1  $\mu$ l S-Adenosylmethionin (500 mM), 0.25  $\mu$ l RiboLock (RNase inhibitor, 20 U/ $\mu$ l) in 20  $\mu$ l of Guanylyltransferase reaction buffer (40 mM Tris/HCl pH 8.0, 6 mM MgCl<sub>2</sub>, 10 mM DTT and 2 mM spermidine) for 2 hours at 37 °C. One volume of RNA loading buffer (5.2.1.10) was added, the solution incubated at 95 °C for 5 minutes and separated by 8% urea-PAGE (5.2.1.10). The RNA was detected by autoradiography, cut out from the gel and eluted from the gel using ethanol precipitation as described above and diluted to a concentration of 1.0-1.2 Bq/cm<sup>2</sup> per  $\mu$ l.

## Material and Methods

For the cleavage reaction 0.5 µg to 10 µg recombinant human N-His<sub>6</sub>-Ago2 and hAgo2 was incubated with 1.5 µl of the α[<sup>32</sup>P] GTP luci II target RNA substrate and 2.5 µl of complementary single stranded siRNA in 25 µl buffer containing 40 mM KCl, 5 mM DTT, 1 mM ATP, 0.2 mM GTP, 10 mM MgCl<sub>2</sub> and 1.2 U RiboLock (Fermentas) at 37 °C for 1 hour. The reaction was stopped by addition of 40 µg proteinase K in 200 µl proteinase K buffer (300 mM NaCl, 200 mM Tris/HCl pH 7.5, 25 mM EDTA, 2% SDS) and incubation at 65 °C for 15 minutes. The RNA was extracted using 200 µl Phenol/Chlorophorm/Isoamylalcohol (25:24:1). For subsequent RNA precipitation, the aqueous phase was mixed with 0.1 volume of NaOAc (pH 5.2) and 2.5 volumes of 100% ethanol and incubated at -20 °C over night. The RNA was pelleted by centrifugation at 17000× g for 1 hour at 4 °C, the pellet air dried at room temperature, dissolved in autoclaved ddH<sub>2</sub>O and separated by 8% urea-PAGE (5.2.1.10) and signals were detected by autoradiography.

### 5.2.2 Cell biological methods

#### 5.2.2.1 Preparation of medium and agar plates for cultivation of *E. coli*

LB-, 2YT- and TB-media were used for the cultivation of the *E. coli* cells, for the amplification of plasmids or for protein expression and SOC-medium was used for plasmid transformations (5.2.2.2; 5.2.2.3; 5.2.2.4; 5.2.2.5; 5.2.2.6). Prior to usage all media and agar plates were sterilized by autoclaving and according to the cell type and the containing plasmid/plasmids/bacmids appropriate antibiotics were added (5.1.7). Agar plates were used to grow single bacterial colonies and were prepared under sterile conditions.

<b>LB-medium</b>		<b>LB-agar plates</b>	
10 g/l	Tryptone	500 ml	LB medium
5 g/l	Yeast extract	1.5 % (w/v)	Agar-agar
10 g/l	NaCl		
<b>2YT-medium</b>		<b>2YT-agar plates</b>	
16 g/l	Tryptone	500 ml	2YT medium
10 g/l	Yeast extract	1.5 % (w/v)	Agar-agar
5 g/l	NaCl		
<b>TB-medium</b>		<b>SOC-medium</b>	
12 g/l	Tryptone	20 g/l	Tryptone
24 g/l	Yeast extract	5 g/l	Yeast extract
4 ml	Glycerol	10 mM	NaCl
2.13 g/l	KH <sub>2</sub> PO <sub>4</sub>	2.5 mM	KCl
12.54 g/l	K <sub>2</sub> HPO <sub>4</sub>	10 mM	MgCl <sub>2</sub>
		10 mM	MgSO <sub>4</sub>
		20 mM	Glucose

The following media were prepared and after autoclaving used for expression of proteins by autoinduction (5.2.2.5).

<b>Zy-medium</b>		<b>5052-medium (50×)</b>		<b>M-medium (50×)</b>	
1% (w/v)	Tryptone	25% (v/v)	Glycerol	1.25 M	K <sub>2</sub> HPO <sub>4</sub>
0.5% (w/v)	Yeast extract	2.5% (w/v)	Glucose	1.25 M	NaH <sub>2</sub> PO <sub>4</sub> × H <sub>2</sub> O
		10 % (w/v)	D-Lactose	2.5 M	NH <sub>4</sub> Cl
				0.25 M	Na <sub>2</sub> SO <sub>4</sub> ×10 H <sub>2</sub> O

Furthermore a 1000× Mg<sup>2+</sup>-solution as well as the following 5000× trace-elements solution was prepared and sterile filtered for usage as component of the autoinduction media (Studier *et al.*, 2005; 5.2.2.5).

#### **Trace-elements solution (5000×)**

50 mM	FeCl <sub>3</sub> × 6 H <sub>2</sub> O
20 mM	CaCl <sub>2</sub> × 2 H <sub>2</sub> O
10 mM	MnCl <sub>2</sub> × 4 H <sub>2</sub> O
10 mM	ZnSO <sub>4</sub> × 7 H <sub>2</sub> O
2 mM	CoCl <sub>2</sub> × 6 H <sub>2</sub> O
2 mM	CuCl <sub>2</sub> × 2 H <sub>2</sub> O
2 mM	NiSO <sub>4</sub> × 6 H <sub>2</sub> O
2 mM	Na <sub>2</sub> MoO <sub>4</sub> × 2 H <sub>2</sub> O
2 mM	Na <sub>2</sub> SeO <sub>3</sub>
2 mM	H <sub>3</sub> BO <sub>3</sub>

#### **5.2.2.2 Preparation of chemical competent *E. coli* cells**

For the generation of chemically competent *E. coli* cells, 500 ml LB-medium containing the appropriate antibiotics were inoculated with 500 µl of an overnight culture (5.1.7; 5.2.2.1). Bacteria were grown to the mid-logarithmic phase of OD<sub>600</sub> = 0.5 at 37 °C and subsequently pelleted by centrifugation at 3000× g, 4 °C for 10 minutes. The cells were resuspended in 150 ml ice cold TFB1 buffer and after additional 5 minutes on ice the bacteria were centrifuged at with 3500× g, at 4°C for 10 minutes. Finally, the cell pellet was carefully resuspended in 5 ml TFB2 buffer, aliquoted, shock frozen in liquid nitrogen and stored at -80 °C. Alternatively another buffer system was used, which had the advantage that the buffer could be autoclaved prior to usage. Therefore, after harvesting the cells were resuspended thoroughly in 125 ml ice cold 100 mM MgCl<sub>2</sub> and centrifuged for 10 minutes at 3000 rpm and 4°C. The cell pellet was resuspended in 250 ml 100 mM CaCl<sub>2</sub> and centrifuged for 10 minutes at 3000 rpm and 4 °C. The cell pellet was resolved in 5 ml buffer containing 100 mM CaCl<sub>2</sub> and 15% (v/v) glycerol, aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

## Material and Methods

TFB1 buffer		TFB2 buffer	
30 mM	KOAc, pH 7.0	10 mM	NaMOPS pH 7.2
50 mM	MnCl <sub>2</sub>	75 mM	CaCl <sub>2</sub>
10 mM	CaCl <sub>2</sub>	10 mM	RbCl <sub>2</sub>
100 mM	RbCl <sub>2</sub>	15% (v/v)	Glycerol
15% (v/v)	Glycerol		
pH 5.8 (acetic acid)		pH 6.5 (NaOH)	

### 5.2.2.3 Plasmid Transformation

To transfer plasmid DNA into different chemically competent *E. coli* strains 10-100 ng plasmid-DNA or a ligation reaction was added to 50 µl of chemically competent *E. coli* cells (5.1.10; 5.2.2.2) and incubated on ice for 20 min. For transformation the chemically competent bacteria were heat-shocked for 45 s at 42 °C and subsequently cooled on ice for 2 min. Afterwards 450 µl LB-medium (5.2.2.1) were added and the bacteria were incubated on a shaker for further 45 min at 37 °C. Finally, the cells were plated on LB-agar plates supplemented with appropriate antibiotics and incubated upside down at 37 °C over night.

To transfer a pFastBac donor plasmid into chemically competent *E. coli* DH10Bac cells, 1 µg of plasmid DNA was added to 100 µl of cells on ice. After incubation for 20 minutes on ice the heat shock was carried out for 45 s at 42 °C, the mixture was chilled on ice for 2 min, 900 µl SOC medium (5.2.2.1) was added and the bacteria were incubated shaking for further 5 to 6 hours at 37 °C. The cells were plated on LB-agar plates containing Gentamycin, Kanamycin, Tetracycline, IPTG as well as X-Gal for blue-white screening and incubated upside down for 48 hours at 37°C.

### 5.2.2.4 Plasmid Isolation

Plasmid DNA was isolated from *E. coli* cultures by alkaline lysis using the QIAprep Miniprep kit from Qiagen, the NucleoSpin Extract II Kit from Macherey-Nagel or the PeqGOLD Plasmid Miniprep Kit from Peqlab according to the manufacturers instructions. The yielded plasmid DNA was used for subcloning, sequencing or transformations. In contrast, bacmid-DNA was isolated from *E. coli* cells using the Midi Prep kit from Invitrogen according to the manufacturers instructions. Alternatively, the bacmid DNA was isolated via direct precipitation. Therefore, a 5 ml *E. coli* culture was pelleted by centrifugation at 4000 rpm for 10 minutes at 4 °C and the supernatant was discarded. The pellet was resuspended in 250 µl of P1 buffer containing 50 mM Tris/HCl pH 8.0, 10 mM EDTA, 100 µg/ml RNase A and the solution transferred to a 1.5 ml reaction tube. Then 250 µl of P2 buffer (200 mM NaOH, 1 % SDS) was added and the mixture homogenized by gentle inverting. Afterwards 350 µl of P3 buffer (3M KOAc, pH 5.5) was added, the mixture homogenized by inverting and centrifuged for 10 minutes at 13.000 rpm and 4 °C. After centrifugation the supernatant was collected, transferred to a new 1.5 ml reaction tube and

again centrifuged for 20 minutes at 13.000 rpm and 4 °C. The supernatant, which should be free of any particles or impurities, was collected and transferred to a new 2 ml reaction tube. For DNA precipitation 800 µl of 100% (v/v) isopropanol was added, the mixture homogenized by inverting and centrifuged for 20 minutes at 13.000 rpm and 4 °C. Afterwards the isopropanol was removed, the DNA pellet washed with 500 µl 70% (v/v) ethanol and again centrifuged for 10 minutes at 13.000 rpm at 4 °C. Finally, the ethanol was removed and either the DNA resolved in 20 µl ddH<sub>2</sub>O for usage as template in a control PCR or 50 µl ethanol were added and the tube transferred to the sterile bench of the insect cell culture for transfections (5.2.1.2; 5.2.2.6.5). For storing the bacmid, after cell lysis the supernatant was mixed with isopropanol as described above and the mixture stored at -20 °C.

### **5.2.2.5 Heterologous protein production in *E. coli***

Prior to protein expression *E. coli* BL21 (DE3) Star cells were transformed with the appropriate plasmid DNA carrying the protein coding open reading frame (ORF) (5.1.12; 5.2.2.3). All *H. sapiens* Dicer and TRBP2 fragments showed best expression rates in *E. coli* BL21 (DE3) Star cells and were therefore expressed in those cells within this study. For large-scale protein expression 1 Liter 2YT- or TB-medium containing the appropriate antibiotics was inoculated with 1 ml of an overnight culture. Bacteria were grown to OD<sub>600</sub> = 0.4 at 37 °C and then the temperature was lowered to 16 °C. Cells were further grown to an OD<sub>600</sub> = 0.8 at 16 °C and protein expression induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After 16 hours, the cells were harvested by centrifugation at 4 °C with 4800 × g for 15 minutes. The cell pellets were washed in 1× PBS (137 mM NaCl, 10 mM K-Phosphate, 2.7 mM KCl, pH 7.4) or the appropriate lysis buffer (5.2.3.4) and again centrifuged at 4 °C with 4800 × g for 15 minutes. The cell pellets were either directly used for cell lysis (5.2.2.7) or flash-frozen in liquid nitrogen and stored at -20 °C. For SDS-PAGE analysis, cell samples corresponding to 500 µl of a culture with OD<sub>600</sub> = 0.8 were taken prior to and after induction. Samples were centrifuged and cell pellets resuspended in 80 µl of protein loading dye and 5 to 8 µl applied to SDS-PAGE analysis (5.2.3.2).

Human TRBP2 was expressed using autoinduction as N-terminal GST-fusion proteins in *E. coli* BL21 (DE3) Star cells. A main culture containing 250 ml ZY-medium, 5 ml M-medium (50x), 250 µl Mg<sup>2+</sup> (1000x), 50 µl micronutrients (5000x), 5 ml 5052-medium (50x) and 100 µg/ml Ampicillin was inoculated with 250 µl of an over night culture (5.2.2.1). Bacteria were grown to an OD<sub>600</sub> = 0.4 at 37 °C and then the temperature was reduced to 16 °C for 48 hours. After incubation the cultures reached an OD<sub>600</sub> between 10 and 20 and were harvested by centrifugation at 4 °C and 4800x g for 15 minute. The cells were washed in 1× PBS (137 mM NaCl, 10 mM K-Phosphate, 2.7 mM KCl, pH 7.4) or the appropriate lysis buffer (5.2.3.4) and again

centrifuged at 4 °C and 4800× g for 15 minutes. The cell pellets were either directly used for cell lysis (5.2.2.7) or flash-frozen in liquid nitrogen and stored at -20 °C. Samples for SDS-PAGE were taken prior to and after induction (5.2.3.2).

### **5.2.2.6 Protein expression using insect cells**

#### **5.2.2.6.1 Generation of recombinant bacmids for transfection**

First the gene of interest was cloned into a pFastBacHtb donor plasmid (see 5.1.12). After verification of a correct insertion by DNA sequencing (5.2.1.6) the recombinant plasmid was transformed into *E. coli* DH10Bac competent cells (5.1.10). Colonies containing recombinant bacmids were identified by disruption of the lacZa gene appearing white on blue-white screening agar plates (5.2.2.1). The DNA of the selected *E. coli* clones containing the recombinant bacmid was either isolated by the MIDI prep Kit from Invitrogen according to manufacturers instructions or by direct precipitation (5.2.1.3). Before the DNA was used for the transfection of insect cells a control-PCR was carried out (5.2.1.2).

#### **5.2.2.6.2 Culturing of insect cells**

Cell culture work with Sf9, Sf21 and High Five (H5) insect cells was performed under sterile conditions using the laminar flow hood Prettl TelStar Typ Bio II A (TelStar, Frankfurt). Cells were counted, cell viability checked and virus infection controlled using the Axiovert 25 inverse microscope (Carl Zeiss, Jena). Sf9 and Sf21 cells were used for transfections, virus generation and recombinant protein expression. High Five cells were used for recombinant protein expression. Within this work insect cells were grown in monolayer and suspension cultures using serum-free medium at 27°C in the dark (5.1.16). Monolayer cultures were cultivated in 25 cm<sup>2</sup>, 75 cm<sup>2</sup> or 175 cm<sup>2</sup> flasks in a Mytron incubator and passaged every 2 to 3 days. Suspension cultures were counted every day and kept at 0.5 × 10<sup>6</sup> cells/ml in appropriate flask shaking at 100 rpm in a Multitron incubation shaker.

#### **5.2.2.6.3 Thawing of insect cells**

A vial of frozen cells was placed into a 37 °C water bath, thawed rapidly with gentle agitation until the cells were almost thawed and quickly removed from the water bath. The outside of the cell vial was decontaminated by treatment with 70% (v/v) ethanol and dried. The thawed 1 ml cell suspension was transferred to a 25 cm<sup>2</sup> flask, which was filled with 4 ml complete Sf-900 III SFM or Express Five SFM medium (5.1.16). To allow the cells to attach, the flask was transferred to a 27 °C incubator for 30-45 minutes. After the cells were attached, the medium was gently removed and the cells were fed with 5 ml of fresh medium for 24 hours and afterwards the medium was exchanged. When the cells had a viability of greater than 70% and had formed a confluent monolayer the cells were sub-cultured.

#### 5.2.2.6.4 Freezing of insect cells

Cells were frozen from suspension or adherent culture, whereby cells of a low passage number and a high viability were used. Therefore, sterile cryo vials were set up on ice and labeled. Cells were counted and centrifuged at 500× g for 10 minutes at room temperature. The supernatant was removed. The pelleted cells were resuspended to the given density in the freezing medium listed in the following table and aliquoted in 1 ml sterile cryo vials. The vials were placed at -20°C for 1 hour, then transferred to -80 °C for 24-48 hours and stored in liquid nitrogen.

Cell Line	Freezing Medium	Density (cells/ml)
Sf9	60% Sf900-II Insect Medium, 30% FBS, 10% DMSO	$1 \times 10^7$
Sf21	60% Sf900-II Insect Medium, 30% FBS, 10% DMSO	$1 \times 10^7$
High Five™	42.5% conditioned Express Five ® SFM, 42.5% fresh Express Five ® SFM, 10% DMSO, 5% FBS	$3 \times 10^6$

#### 5.2.2.6.5 Transfection of Sf9 and Sf 21 cells in a 6-well format

The bacmid DNA prepared from a 5 ml pre-culture (5.2.2.4) is used for the transfection of 2 wells of a 6-well culture plate (Thermo Scientific or Nunc). The prepared bacmid DNA was taken to the sterile hood in 50 µl 70% ethanol. After disinfecting the tube outside, the tube was opened inside the hood to allow the ethanol to evaporate completely. Once the bacmid was free of ethanol, 20 µl of autoclaved ddH<sub>2</sub>O was added to dissolve the DNA. In the mean time, a mix containing the transfection reagent was prepared. Therefore, 100 µl of SF900III medium and 10 µl of Fugene 6 were mixed. The resolved bacmid was diluted with 200 µl of SF900 III medium to a final volume of 220 µl. Afterwards 110 µl of the transfection reagent master mix was added to each tube containing 220 µl of diluted bacmid. The mixture was incubated for 30 to 60 minutes at room temperature. While the transfection reagent was incubating with the bacmid, 3 ml of Sf9 or Sf21 cells at  $0.3 \times 10^6$  cells/ml were seeded in 3 wells of the 6 well plate. The cells used for transfections had a viability of more than 95% and one of the 3 wells was used as a cell control. 165 µl of bacmid and transfection reagent mix was added drop-wise to one well of a 6 well plate containing  $0.3 \times 10^6$  Sf9 or Sf21 cells. The plates were incubated at 27 °C protected from light for 60 hours. Then the supernatant that contains the V0 virus was collected and kept at 4 °C or used directly to produce V1 virus. After the V0 virus was collected, 3 ml more medium were added to the cells within the 6 well plate and incubate another 2 days to test for protein expression using SDS-PAGE (5.2.3.2).

### **5.2.2.6.6 Virus production using Sf9 cells**

One day before infection with the V0 virus, 250 ml shaker flasks containing 25 ml of SF9 cells at  $0.5 \times 10^6$  cells/ml were prepared. At the day of the infection with V0, the cells in each flask were split to  $0.5 \times 10^6$  cells per ml, leaving 25 ml cells in the flask. The cells used for virus production were >95% viable. 3 to 6 ml of the V0 virus was added to the cells. After 24 hours the cells were counted to check if they are still proliferating. If the cells were proliferated, they were diluted to  $0.5 \times 10^6$  cells /ml with medium. This procedure was repeated until the cells stop proliferating. Sometimes it was necessary to change the 250 ml flask to a larger one. 60 hours after the cells stopped differentiating, the V1 was collected by centrifuging at 2000 rpm for 10 minutes. The V1 virus (supernatant after centrifugation) was stored at 4 °C in the dark or directly used for protein expressions. The pellet was analyzed for protein expression using SDS-PAGE (5.2.3.2).

### **5.2.2.6.7 Protein production using Sf9 and High Five cells**

Sf9 cells were used for expression of full length N-hDicer for 36 to 40 hours and full length N-hAgo2 was expressed in High five cells for 48 hours. For test protein expression of full length N-hAgo2, 300 ml shaker flasks with 50 ml of High Five at  $0.5 \times 10^6$  cells/ml were prepared and infected with different amounts of fresh V1 virus (0.5 ml, 1 ml, 1.5 ml and 2 ml). The best amount of virus added is the one where the cells stop proliferating immediately and start expressing the protein. 24 hours after the virus was added, the cells were counted every 4 hours and samples were taken for SDS-PAGE analysis (5.2.3.2). For large-scale expression of full length N-hAgo2 2 L shaker flasks with 500 ml of High Five at  $0.5 \times 10^6$  cells/ml were prepared and infected with the amount of V1 virus, which gave the best result during test expression and incubated for 48 hours. After incubation, the cultures were counted, samples for SDS-PAGE analysis were taken and the cells were harvested by centrifugation at 4 °C with  $4800 \times g$  for 15 minutes. The cell pellet was washed in 1× PBS (137 mM NaCl, 10 mM K-Phosphate, 2.7 mM KCl, pH 7.4) or the appropriate lysis buffer (5.2.3.5) and again centrifuged at 4 °C with  $4800 \times g$  for 30 minutes. The cell pellets were either directly used for cell lysis (5.2.2.7) or frozen in liquid nitrogen and stored at -20 °C.

For test-expression of full length N-hDicer 300 ml shaker flasks with 50 ml of Sf9 cells at  $1 \times 10^6$  cells/ml were prepared and infected with different amounts of fresh V1 virus (0.5 ml, 1 ml, 1.5 ml and 2 ml). 24 hours after infection, the cells were counted every 4 hours and samples were taken for SDS-PAGE analysis (5.2.3.2). For large-scale expression of full length N-hDicer 2 L shaker flasks with 500 ml of Sf9 cells at  $1 \times 10^6$  cells/ml were prepared and infected the amount of V1 virus, which gave the best result during test-expression and incubated for 36 to 40 hours. After incubation, the cultures were counted, samples for SDS-PAGE analysis (5.2.3.2) were taken and the cells were harvested by centrifugation at 4 °C and  $4800 \times g$  for 15 minutes. The cell

pellet was washed in 1× PBS (137 mM NaCl, 10 mM K-Phosphate, 2.7 mM KCl, pH 7.4) or the appropriate lysis buffer (5.2.3.5) and again centrifuged at 4 °C with 4800× g for 30 minutes. The cell pellets were either directly used for cell lysis (5.2.2.7) or frozen in liquid nitrogen and stored at -20 °C.

#### **5.2.2.7 Preparation of cell extracts for protein purification**

Prior to cell lysis *E. coli* cells were resuspended in the appropriate lysis buffer that was supplemented with protease inhibitor (5.1.6; 5.2.3.4; 5.2.3.5; 5.2.3.6). The cells were lysed using an equilibrated and ice chilled micro fluidizer for 2-6 times with a pressure of 0.55 MPa. Clarification of the cell lysate was done by centrifugation at 30000× g and 4 °C for 30 minutes. Additionally a sample was taken for SDS-PAGE analysis (5.2.3.2). The soluble protein solution within the supernatant was used for protein purification and SDS-PAGE analysis (5.2.3.2).

Insect cell pellets were grinded in a mortar chilled with liquid nitrogen. After the cells were pestled to powder, this powder was resuspended in lysis buffer supplemented with complete protease inhibitor tablets (Roche) and a sample was taken for SDS-PAGE analysis (5.2.3.2). The thawed suspension was cleared by centrifugation at 100000 × g and 4 °C for 1 hour. Samples of the supernatant and the pellet were taken for SDS-PAGE and the supernatant used for protein purification (5.2.3.2; 5.2.3.4; 5.2.3.5; 5.2.3.6).

### **5.2.3 Protein chemistry**

#### **5.2.3.1 Concentration determination of proteins**

The concentration of proteins and RNA-protein complexes was determined in a colorimetric assay relying on the method of Bradford (Bradford, 1976). For Bradford measurements 20 µl protein or protein-RNA solution was diluted in 980 µl Bradford reagent (BioRad, Munich), incubated for 3 minutes at room temperature and measured at 595 nm in cuvettes.

In some cases the concentration of proteins was also determined spectroscopically. For that purpose the absorbance at 280 nm was verified using the NanoDrop (Thermo scientific). Molar extinction coefficients for proteins under study were calculated using ProtParam (5.1.15). Additionally, the coefficient  $A_{280}/A_{260}$  was calculated to check each protein preparation for contaminations with nucleic acids. For exact protein quantifications and for MALS calculations the native extinction coefficient of that protein was determined according to Gill and Hippel (Gill and Hippel, 1989).

### **5.2.3.2 Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE)**

Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) under denaturing conditions was used to analyze protein solutions. SDS-PAGE was carried out according to standard procedures using the Hoefer miniVE SDS-PAGE-System. SDS gels contained a stacking part, cast of 4% acrylamide, 125 mM Tris (pH 6.8) and 0.1% (w/v) SDS polymerized with 0.05% (w/v) APS and 0.1% (v/v) TEMED. Proteins were resolved on the lower gel, containing 6-15% acrylamide, 375 mM Tris (pH 8.8) and 0.1% (w/v) SDS polymerized with 0.05% APS and 0.1% TEMED. Protein samples were mixed with 2× SDS loading buffer (50% (v/v) glycerol, 70 mM SDS, 62.5 mM Tris, pH 6.8, 0.1% (w/v) bromphenolblue, 5% (v/v) 2-mercaptoethanol) and loaded on the stacking part of the gel. Gels were run at 1 mA/cm<sup>2</sup> in SDS running buffer (25 mM Tris-HCl, 200 mM glycine und 0.1% (w/v) SDS) until the front of the loading dye reached the bottom of the gel. After protein separation the gels were stained with 0,002% (w/v) Coomassie Brilliant Blue in 10% (v/v) ethanol and 5% acetic acid for 30 min and destained first in 10% (v/v) ethanol and 5% (v/v) acetic acid and finally in water.

### **5.2.3.3 Concentration of protein solutions**

Generally, the purified proteins were concentrated to 1-20 mg/ml using Sartorius or Millipore concentrators, which contain an ultrafiltration membrane with a specific molecular weight cut off (MWCO) (5.1.2). A MWCO 30% smaller than the protein size was used. The samples were passed through the membrane by centrifugal force and resuspended from time to time. The concentrated proteins were transferred to a fresh tube, centrifuged for 15 minutes at 4 °C with 17000 rpm using a table centrifuge and the supernatant was carefully pipetted into a new tube. The protein concentration of this solution was determined (5.2.3.1) and the protein solution directly used for further experiments or frozen in liquid nitrogen and stored at -80 °C.

### **5.2.3.4 Purification of proteins and protein complexes from *E. coli* cells**

All chromatographic purification steps were performed on Äkta prime, purifier or explorer systems at 4 °C unless stated otherwise.

Human TRBP2, TRBP2-fragments and Dicer fragments were expressed as GST-fusion proteins (5.1.12; 5.2.1.1; 5.2.2.5). To purify minimal hDicer-hTRBP2 complexes the pellets of the separate expressed proteins were mixed prior to cell lysis, centrifugation and purification (5.2.2.7; 5.2.3.6). Within the following a general purification protocol of a GST-fusion protein or GST-fusion protein complex is described and the buffers used for purification of different proteins and complexes are depicted in tables below. After cell lysis and centrifugation the supernatant was loaded onto 2 × 5 mL GSTrap columns (GE Healthcare), which were equilibrated with lysis buffer (see table below). After loading the sample, the columns were washed with wash buffer to remove LiCl and the GST-fusion protein subsequently eluted with

wash buffer containing 30 mM reduced glutathione. Fractions containing the fusion protein were pooled and incubated with PreScission protease in a 100:1 ratio of protein to protease overnight at 4°C. In case of full length hTRBP2 purification a Heparin Sepharose purification step was included in order to remove cleaved GST as well as contaminating nucleic acid. After protease cleavage or Heparin Sepharose purification step the protein solution was loaded onto a Superdex S75 (16/60) or (26/60) column equilibrated with gel filtration buffer. Finally, to remove GST and PreScission protease a second step of glutathione affinity chromatography using gel filtration buffer was carried out. The purification of GST-fusion proteins for pull down assays was carried out according to the purification of the proteins without in PreScission cleavage step and the final GSH-Sepharose step.

#### The buffers used for the purification of GST-hTRBP2 and hTRBP2

<b>Lysis buffer</b>		<b>Wash buffer</b>		<b>Elution buffer</b>	
300 mM	NaCl	300 mM	NaCl	300 mM	NaCl
50 mM	HEPES, pH 8.0	50 mM	HEPES, pH 8.0	50 mM	HEPES, pH 8.0
5 % (v/v)	Glycerol	5 % (v/v)	Glycerol	5 % (v/v)	Glycerol
2 mM	EDTA	2 mM	EDTA	2 mM	EDTA
1 M	LiCl	2 mM	DTT	2 mM	DTT
2 mM	DTT			30 mM	Red. Glutathione

<b>Heparin buffer A</b>		<b>Heparin buffer B</b>		<b>Gel filtration buffer</b>	
50 mM	NaCl	1 M	NaCl	150 mM	NaCl
50 mM	HEPES, pH 8.0	50 mM	HEPES, pH 8.0	20 mM	HEPES, pH 7.5
5 % (v/v)	Glycerol	5 % (v/v)	Glycerol	2 % (v/v)	Glycerol
2 mM	DTT	2 mM	DTT	1 mM	DTT

#### Buffer for the purification of hTRBP2 fragments

<b>Lysis buffer</b>		<b>Wash buffer</b>		<b>Elution buffer</b>	
400 mM	NaCl	400 mM	NaCl	400 mM	NaCl
50 mM	HEPES, pH 8.0	50 mM	HEPES, pH 8.0	50 mM	HEPES, pH 8.0
5 % (v/v)	Glycerol	5 % (v/v)	Glycerol	5 % (v/v)	Glycerol
2 mM	EDTA	2 mM	EDTA	2 mM	EDTA
1 M	LiCl	2 mM	DTT	2 mM	DTT
2 mM	DTT			30 mM	reduced Glutathione

<b>Gel filtration buffer</b>	
150 mM	NaCl
20 mM	HEPES, pH 7.5
2 mM	DTT

**Buffer for the co purification of minimal hDicer-hTRBP2 complexes**

<b>Lysis buffer</b>		<b>Wash buffer</b>		<b>Chaperone wash buffer</b>	
400 mM	NaCl	400 mM	NaCl	400 mM	NaCl
50 mM	HEPES, pH 8.0	50 mM	HEPES, pH 8.0	50 mM	HEPES, pH 8.0
10% (v/v)	Glycerol	10%(v/v)	Glycerol	10% (v/v)	Glycerol
2 mM	EDTA	2 mM	EDTA	2 mM	DTT
1 M	LiCl	2 mM	DTT	2 mM	ATP
2 mM	DTT			10 mM	MgCl <sub>2</sub>
				200 mM	KCl
<b>Elution buffer</b>		<b>Gel filtration buffer</b>			
400 mM	NaCl	150 mM	NaCl		
50 mM	HEPES, pH 8.0	20 mM	HEPES, pH 7.5		
10% (v/v)	Glycerol	3 % (v/v)	Glycerol		
2 mM	EDTA	2 mM	DTT		
2 mM	DTT				
30 mM	reduced Glutathione				

**5.2.3.5 Purification of proteins and protein complexes from insect cells**

Human Ago2 and Dicer were expressed as hexahistidine-fusion proteins (N-hAgo2 and N-hDicer) in High Five and Sf9 cells respectively. The cells were harvested and lysed as described in 5.2.2.6.7 and 5.2.2.7. For co-purification of N-hAgo2 and N-hDicer the insect cell pellets of the separately expressed proteins were pestled together in a mortar using liquid nitrogen and the powder was dissolved in lysis buffer (see tables below). The fusion proteins unspecifically bind RNAs from insect cells (own observations; Elkayam *et al.*, 2012; Schirle *et al.*, 2012) and therefore the first established purification strategy implied RNase A addition during cell lysis, but did not involve a specific step for later removal of RNase A. The RNase A treatment prohibits protein-RNA interactions and therefore an alternative purification strategy was established. Two alternative strategies to remove unspecifically bound RNA were shown to be the addition of 1-2 M LiCl during cell lysis or a 1-2 M LiCl-washing step during the first affinity chromatography column. As both procedures gave the same result only one purification strategy (with 1 M LiCl within the lysisbuffer) is described.

After cell lysis and centrifugation the clarified supernatant was loaded onto in lysis buffer equilibrated 2 × 5 ml HisTrap-columns. The columns were washed with wash buffer and an optional chaperone wash buffer and the fusion protein/proteins was/were eluted using a step gradient to 300 mM Imidazole (see elution buffer tables below). Optionally the His<sub>6</sub>-tag was removed after this step by incubation with TEV-protease in a 100:1 molar ration of protein to protease at 4 °C over night. Finally, the protein or protein complex was polished using a Superdex 200 (16/60) or (26/60) gel filtration column equilibrated with gel filtration buffer (see tables below).

**Buffer for the purification of hAgo2, hDicer or the hAgo2-hDicer complex**

<b>Lysis buffer</b>		<b>Wash buffer</b>		<b>Chaperone wash buffer (optional)</b>	
300 mM	NaCl or KCl	300 mM	NaCl or KCl	300 mM	NaCl or KCl
50 mM	HEPES, pH 7.0-8.0	50 mM	HEPES, pH 7.0-8.0	50 mM	HEPES, pH 7.0-8.0
5% (v/v)	Glycerol	5% (v/v)	Glycerol	5% (v/v)	Glycerol
1 M	LiCl	35 mM	Imidazole	35 mM	Imidazole
35 mM	Imidazole	2 mM	2-mercapto-ethanol	2 mM	2-mercapto-ethanol
2 mM	2-mercapto-ethanol			2 mM	ATP
				10 mM	MgCl <sub>2</sub>
				200 mM	KCl
<b>Elution buffer</b>		<b>Gel filtration buffer</b>			
300 mM	NaCl/KCl	150 mM	NaCl/KCl		
50 mM	HEPES, pH 7.0-8.0	20 mM	HEPES, pH 7.5 or Imidazole 6.8 - 7.0		
5% (v/v)	Glycerol	5% (v/v)	Glycerol		
300 mM	Imidazole	2 mM	DTT		
2 mM	2-mercapto-ethanol				

**5.2.3.6 Co-purification of full length human Dicer-TRBP2 complex**

A cell pellet of NGST-hTRBP2 was lysed as described in 5.2.2.7 in buffer containing 300 mM NaCl, 50 mM HEPES-NaOH pH 8.0, 10% glycerol, 0-2 mM MgCl<sub>2</sub>, 0-2 mM EDTA, 5 mM DTT and protease inhibitors. Insect cell pellets of N-hDicer were pestled in mortar using liquid nitrogen and the powder dissolved in NGST-TRBP2 lysate. The lysate containing NGST-hTRBP2 and N-hDicer was centrifuged at 100000 × g for 1 hour at 4 °C and loaded on an equilibrated GSH-Sepharose column. After loading a 1M LiCl wash step was included with a buffer containing 1 M LiCl, 300 mM NaCl, 50 mM HEPES/NaOH pH 8.0, 10 % glycerol, 0-2 mM MgCl<sub>2</sub>, 0-2 mM EDTA, 5 mM DTT. The GST-fusion protein as well as bound proteins were eluted with a buffer containing 300 mM NaCl, 50 mM HEPES/NaOH pH 8.0, 10 % glycerol, 0-2 mM MgCl<sub>2</sub>, 0-2 mM EDTA and 5 mM DTT and 30 mM reduced glutathione. Fractions containing the fusion protein and bound proteins (N-hDicer) were pooled and incubated with PreScission protease overnight at 4 °C. To remove excess of hTRBP2 and GST from the N-hDicer-hTRBP2 complex a Superdex 200 (16/60) or (26/60) gel filtration column equilibrated in a buffer containing 150 mM NaCl or KCl, 20 mM HEPES pH 7.5, 5% (v/v) glycerol, 0-2 mM MgCl<sub>2</sub> and 2 mM DTT was used (5.1.3).

### **5.2.3.7 Analytical size exclusion chromatography**

Analytical size exclusion chromatography (analytical SEC) was performed to analyze the protein size and to verify possible protein-protein or protein-RNA complex formations. The chromatography steps were performed on purifier or explorer systems at 4 °C or 20 °C according to manufactures instructions (5.1.2; 5.1.3). Columns were pre-equilibrated with 1.5 to 3 column volumes of running buffer, 400 µl sample were injected, the sample run with a flow rate of 0.3 to 0.5 ml/min, fractions were taken and analyzed on SDS-PAGE, urea-PAGE and agarose gels (5.2.3.2; 5.2.1.10; 5.2.1.8).

### **5.2.3.8 Western blot**

For western blot analysis, proteins were separated by 6-15% SDS-PAGE (5.2.3.2) and transferred to a nitrocellulose membrane by wet blot in transfer buffer (25 mM Tris base, 150 mM glycine, 20% (v/v) methanol) at 90 V and 4°C for 4 hours using a BioRad blotting chamber (BioRad, Munich). For the detection of His-tagged proteins, the membrane was blocked with 1% casein in TBST buffer (20 mM Tris/HCl pH 8.0, 150 mM NaCl, 0.05% (v/v) Tween) for 1 hour shaking. After three ten-minute washes with TBST, the first antibody (*mouse α-His* from Qiagen, 5.1.8) was diluted 1:1000 in TBST and incubated with the membrane shaking at 4 °C over night. The excess of primary antibody was washed away by three ten-minute washes with TBST. The membrane was incubated with the peroxidase conjugated secondary antibody (*goat α-mouse* PO-conjugated, Biomol) or the alkaline phosphatase (AP) conjugated secondary antibody (*goat α-mouse* AP-conjugated, Biomol) for 1 hour. When incubated with peroxidase conjugated secondary antibody, the membrane was subjected to three five-minute washes in TBST buffer, then incubated 2 minutes in chemiluminescence detection reagent ECL (Pierce, Rockford) and analyzed using a digital imaging system (BioRad, Munich). Membranes incubated with AP conjugated secondary antibody were applied to three five-minute washes in buffer A (100 mM Tris pH 9.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>) and the color reaction was started by addition of 33 µl BCIP (5% (w/v) 5-bromo-4-chloro-3-indolylphosphat in 70% dimethylformamide) and 66 µl NTB (7.5% nitroblue-tetrazoliumchlorid in 70 % dimethylformamide).

Strep-tagged fusion proteins were detected with a Strep-Tactin alkaline phosphatase (AP) conjugate from IBA (Göttingen) according to manufacturer instructions.

### **5.2.3.9 Limited proteolysis**

Limited proteolysis of proteins was used to find stable protein fragments or complexes suitable for crystallization. Several proteases were employed and an experiment was carried out with a fixed protease concentration at a series of time points at 4 °C in gel filtration buffer (5.2.3.4; 5.2.3.5; 5.2.3.6). The reactions were stopped by addition of SDS loading dye and the samples were analyzed by SDS-PAGE (5.2.3.2).

#### **5.2.3.10 Pull down assay**

To assay direct binding of proteins to each other pull down assays were performed. GST-fusion proteins were bound to equilibrated Glutathione Sepharose (GE Healthcare) for 1 h at 4 °C on a rotating wheel. Samples were washed three times with gel filtration buffer (5.2.3.4; 5.2.3.5; 5.2.3.6) to remove unbound protein or protein excess. A 2- to 3-molar excess of protein binding partner/partners were added to 20 µl of protein-coated beads and the samples were incubated at 4 °C on a rotating wheel. After 2 hours, the samples were washed three times with binding buffer and proteins eluted with gel filtration buffer (5.2.3.4; 5.2.3.5; 5.2.3.6) supplemented with 30 mM reduced glutathione. Samples from each step were analyzed by SDS-PAGE (5.2.3.2).

### **5.2.4 Special Methods/ Biophysical methods**

#### **5.2.4.1 Multi-angle light scattering combined with size exclusion chromatography**

Multi-angle light scattering (MALS) coupled to size exclusion chromatography (SEC) was used to measure the absolute molar mass and size of proteins or complexes. Usually, a protein sample (0.5-1 mg/ml, 0.4 ml) was injected onto an equilibrated Superdex 200 (10/300) size exclusion column attached to an Äkta purifier (GE Healthcare) at a flow rate of 0.5 ml/min in gel filtration buffer. The elution was recorded by a Wyatt miniDAWN TREOS multi-angle light scattering detector. The data were evaluated in ASTRA5 or ASTRA56 software using the Zimm plot model.

#### **5.2.4.2 Fluorescence-based thermal stability assay**

The influence of different buffer conditions and ions on the stability of proteins, protein-protein and protein-RNA complexes was analyzed by the fluorescence-based thermal shift assay. This method uses the fluorescence dye Sypro-orange that undergoes a signal change while the protein, protein-protein or protein-RNA complex undergoes thermal unfolding. Sypro-orange in a folded protein solution is exposed to an aqueous environment and its fluorescence signal quenched, whereby it exhibits a strong fluorescence emission upon intercalation into the hydrophobic core of an unfolding or unfolded protein (Ericsson *et al.*, 2006). Fluorescence-based thermal stability measurements were recorded from 20 to 95 °C in 1 °C steps using the C1000 Thermal Cycler with CFX96 Real time System (BioRad, Munich) with an excitation of 490 nm and an emission of 575 nm in a 96 well plate. The total sample volume in each well of 20 µl contained 16 µl of protein or protein complex stock solution (0.1-10 µM in gel filtration buffer), 2 µl Sypro orange (100 x) and 2 µl of screening condition (5.1.14). Samples were tested with this assay in doublets. The unfolding curves were normalized and the apparent unfolding temperature was determined with the MaxStaf software (5.1.15).

#### **5.2.4.3 Isothermal titration calorimetry**

Isothermal titration calorimetry (ITC) allows the direct and quantitative characterization of the thermodynamics of binding events. From one optimal ITC experiment the association constant, the binding enthalpy, the binding entropy and the stoichiometry of a binding event can be calculated. For our studies we utilized the VP-ITC system (MicroCal Inc.) using the power compensation method. ITC was used to investigate the interaction between the RNA-binding protein hTRBP2 with siRNA duplex1 (5.1.13). The purified hTRBP2 (5-20 mg/ml) (5.2.2.5) and the substrate siRNA duplex1 (20 mg/ml) were dialyzed for 16 h at 4 °C against 150 mM NaCl and 10 mM HEPES/NaOH pH 7.5. The hTRBP2 or siRNA duplex1 concentration was determined based on absorbance at 280 nm or 260 nm, respectively, using a NanoDrop. A typical experiment involved 20  $\mu$ M siRNA duplex1 in a 300  $\mu$ l injection syringe and 1300  $\mu$ l of TRBP2 at 1  $\mu$ M in the sample cell, whereby the ITC analysis was carried out at 20 °C. Plain buffer was injected into the hTRBP2 solution as a control. Prior to data evaluation the heat releases from the control experiment was subtracted from the experimental data. Data analysis was carried out with Microcal Origin 7.0 to determine the enthalpy of binding ( $\Delta H$ ), the entropy of binding ( $\Delta S$ ), the equilibrium dissociation constant ( $K_d$ ) and the interaction stoichiometry ( $n$ ).

#### **5.2.4.4 Characterization of the ATPase activity**

The ATP hydrolysis rate was determined using the EnzCheck®Phosphate Assay Kit (Invitrogen, USA). The purine nucleoside phosphorylase (PNP) catalyzes the reaction of inorganic phosphate with 2-amino-6-mercapto-7-methylpurine riboside (MESG), forming a methyl purine with an absorption maximum at 360 nm. The increase of absorption at 360 nm is thus directly proportional to the hydrolysis of ATP to ADP and inorganic phosphate by a protein with ATPase activity (Webb, 1992). The 100  $\mu$ l reaction mixture contained 10 mM Tris/HCl pH 7.5, 100 mM NaCl, 2 mM  $MgCl_2$  and 1  $\mu$ M protein as well as 1  $\mu$ l PNP, 20  $\mu$ l MESG and 5  $\mu$ M RNA. The mixture was incubated at 20 °C for 10 min prior to the experiment. The measurement was started upon addition of ATP in concentrations between 0 and 500  $\mu$ M, the absorbance at 360 nm was measured every two seconds during ten minutes at 20 °C using a Ultrospec 2100pro Photometer (GE Healthcare, Munich).

#### **5.2.4.5 GraFix preparation of macromolecular complexes**

The GraFix (gradient fixation) protocol was applied to *H. sapiens* Dicer-Ago2-TRBP2-complex (RLC), Dicer-Ago2 complex, Dicer-TRBP2 complex and Dicer supplemented with siRNA duplex1 or without (5.1.13; Kastner *et al.*, 2009). The samples were run on a GraFix gradient of glycerol or sucrose ranging from 10-30% glycerol or sucrose (top to bottom) and a fixation reagent gradient ranging from 0-0.02% (top to bottom). The cross-linker was added to GraFix gradient solutions immediately before preparation of the gradients. A 4 ml gradient was

prepared by layering 2 ml of the 30% glycerol or sucrose buffer under the 10% buffer using a syringe and a spinocan needle. The gradients were subsequently formed by a gradient mixer (Gradient Master, BioComp Instruments, Inc., Fredericton, NB, Canada). Prior to centrifugation the gradients were pre chilled at 4 °C for 1 hour and 200 µl of the protein or complex solution was loaded on top of the gradient. The prepared gradients were centrifuged in a Sorvall Centrifuge Evolution RC (Thermo Electron, Langenselbold, Germany) using a Sorvall Th 660 rotor (Thermo Electron, Langenselbold, Germany) for generally 18 hours with 38000 rpm at 4 °C. After centrifugation, the gradients were fractionated from the bottom and used for SDS-PAGE analysis and electron microscopy analysis (5.2.3.2; 5.2.4.6).

#### **5.2.4.6 Electron microscopy**

The structures of the *H. sapiens* RLC, Dicer-Ago2 complex, Dicer-TRBP2 complex and Dicer supplemented with siRNA duplex1 (5.1.13) or without were analyzed on a low resolution level by negative-stain electron microscopy. The RLC and RLC-siRNA duplex1 complexes were in addition subjected to cryo electron microscopy. David Haselbach, who is a member of the 3D Electron Cryo-Microscopy Group of Prof. Dr. Holger Stark at the Max-Planck-Institute for Biophysical Chemistry in Göttingen, carried out all such experiments. Prior to preparation and imaging, the samples were applied to the GraFix protocol (5.2.4.5). All image processing steps were carried out by David Haselbach and are not described in detail here.

##### **5.2.4.6.1 Preparation and imaging of negative-stained particles**

Negative-staining was carried out by the double carbon film method with uranyl formate. Therefore the GraFix prepared samples (5.2.4.5) were allowed to absorb to a thin carbon film in solution. The immobilizes complexes (or proteins) were then transferred to a copper grid covered with a perforated carbon support film and stained for 1 to 2 minutes in a 2% (w/v) uranyl formate solution (Jensen *et al.*, 2010). The stain solution was then removed by blotting from the side with filter paper (Whatman) and grids were quick-dried under a light bulb. Negative-stain images were recorded two times binned with a Philips CM 200 FEG electron microscope (Philips, Eindhoven, Netherlands) equipped with 4k x 4k CCD camera (TVIPS, Munich) at an acceleration voltage of 160 kV and a magnification of 88k corresponding to a pixel size of 2.45 Å/pixel.

##### **5.2.4.6.2 Preparation and imaging of cryo samples**

The appropriate GraFix fractions containing the fixated RLC or RLC-siRNA duplex1 complex were submitted to a buffer exchange procedure via a 0.5 ml Zeba Spin desalting column (Thermo Scientific, Rockford, USA) or GE Healthcare G200 gel filtration column (GE Healthcare)

## Material and Methods

according to manufactures instructions. The eluted sample was absorbed on a floating carbon film as before for the negative-stain and the grid blotted by the Vitrobot Mark IV (FEI, Eindhoven, Netherlands) with a blot force of 12 at 4 °C and a relative humidity of 100 %. After blotting the sample was shot into liquid ethane and stored within liquid nitrogen until electron microscopy analysis. Additionally, cryonegative-stain grids were prepared (Jensen *et al.*, 2010). First, a slide coated with a supporting carbon film is inserted in the sample-containing solution. After absorption of the GraFix prepared sample (no buffer exchange is necessary) onto the carbon surface the slide is floated upside-down into a well containing the uranyl formate staining solution. The carbon film containing the stained sample is picked up by a grid. The grids rest only shortly to ensure that the stained samples are not dried completely and are frozen in liquid nitrogen and stored there until further usage. Cryo images were recorded with a 2-fold binned 4 k x 4 k Eagle camera in a Cs corrected Titan Krios (FEI Company) at a magnification of 80000 at an acceleration voltage of 80 kV and an electron dose of 20 electrons per Å<sup>2</sup>.

### **5.2.4.7 Small angle X-ray scattering (SAXS)**

To determine low resolution structures and Rg values of the hTRBP2 fragment T7 containing amino acids 157-227 in solution small angle X-ray scattering (SAXS) measurements were carried out in collaboration with Clement Blanchet from the group of Prof. Dr. Dmitri I. Svergun at the EMBL in Hamburg. Prior to SAXS measurements the hTRBP2 fragment was freshly prepared, concentrated using Millipore concentrators and the concentrations determined using Bradford and A<sub>280</sub> absorbance. SAXS-data were collected on the X33 beam line of the EMBL (DESY, Hamburg) using a 2D Photon counting Pilatus 1M-W pixel X-ray detector (Dectris, Switzerland). For measurements an automated sample changer was used. A BSA standard with different concentrations was measured at least one time before measurement of the samples for an initial calculation setup. For one measurement a final volume of 50 µl sample was used and for T7 a concentration series of 3 to 4 different concentrations was measured for each of the different buffer at 10 °C (see following table). Clement Blanchet carried out all data evaluation steps.

<b>Buffer</b>	<b>Concentration serie (mg/ml)</b>
<b>T7 (hTRBP2 aa 157-227)</b>	
500 mM NaCl, 20 mM HEPES pH 7.5, 2 mM DTT	25;12.5; 6.25; 3.13
250 mM NaCl, 20 mM NaPhosphate pH 7.5, 5% Glycerol, 2 mM DTT	6; 3; 1.5; 0.75
150 mM NaCl, 20 mM NaPhosphate pH 7.5, 5% Glycerol, 2 mM DTT	12; 6; 3; 1.5

### 5.2.4.8 Crystallization

#### 5.2.4.8.1 General crystallization setup

All proteins, protein-RNA complexes and protein-protein complexes with a concentration of 1-20 mg/ml were subjected to initial rounds of crystallization trials. For screening, a Phoenix/RE crystallization robot was used, which reliably adds 250 nl protein solution to 250 nl precipitate solution sitting drop as well as 250 nl protein solution to 125 nl precipitate solution sitting drop with a reservoir of 40  $\mu$ l precipitate solution in a 96-well format. Older screening was performed sitting drop in 24-well crystallization plates by manual mixing of 1  $\mu$ l protein solution with 1  $\mu$ l precipitate solution and a reservoir of 500  $\mu$ l precipitate solution (5.1.14). In the appendix a table of the initial crystallization screens and tested *in situ* proteolysis conditions for all proteins and complexes used in this study are listed (5.1.14; 5.2.4.8.2; see Table 4).

#### 5.2.4.8.2 *In situ* proteolysis

The addition of proteases in crystallization droplets, which is also called *in situ* proteolysis was used to induce crystallization of proteins or protein complexes that had not crystallized so far (Dong *et al.*, 2007). The proteases chymotrypsin or trypsin were mixed with the protein solution in mass ratios ranging from 1:1000 to 1:100 and incubated for 20 minutes prior to crystallization screening (see Table 4).

#### 5.2.4.8.3 Crystallization and structure determination of the second dsRNA-binding domain of hTRBP2

Prior to crystallization trials hTRBP2 comprising amino acids 157 to 366 was incubated with Chymotrypsin in a ratio of 500:1 and two different crystal forms could be obtained. Hexagonal-shaped crystals were obtained at 20 °C using the sitting drop method by mixing 250  $\mu$ l of the 5 mg/ml protein-protease mixture with 125  $\mu$ l 3.2 M Ammoniumsulfate and 0.1 M Tris, pH 8.0 after 2 month. As cryo protectant 15% glycerol in addition to the crystallization solution was used. Complete datasets were collected on the microfocus beamline ID23-2 of the ESRF (Grenoble). Data were processed in space group P3<sub>1</sub>2<sub>1</sub> with XDS and scaled using XSCALE to a final resolution of 2.28 Å (Kabsch, 2010). The phase problem was solved by molecular replacement using the crystal structure of the second dsRNA-binding domain of *H. sapiens* TRBP2 in complex with RNA (PDB-ID: 3ADL; Yang *et al.*, 2010), whereby the RNA was removed, for the search in Phaser (McCoy *et al.*, 2007). The model was improved manually using Coot (Emsley *et al.*, 2010) and refined with PHENIX (Adams *et al.*, 2010).

Diffraction quality plate-shaped crystals were obtained at 20 °C with the sitting drop method by mixing 250 nl of the 5 mg/ml protein-protease mixture with 125 nl 2 M Li<sub>2</sub>SO<sub>4</sub> and 3 % 2-Methyl-2,4-pentanediol (MPD) after 3 weeks. The crystals were flash-cooled in liquid nitrogen

## Material and Methods

without any need of further cryo protectant. Complete datasets were collected at the microfocus beamline ID23-2 of the ESRF (Grenoble). The data were processed and scaled using XDS and XCALE in the space group P422 to a final resolution of 2.95 Å (Kabsch, 2010). No reliable crystal structure could be solved by molecular replacement to date (McCoy *et al.*, 2007).

## 6 Results

### 6.1 RISC-loading complex (RLC)

In order to study the human RISC-loading complex (RLC), the aim was to reconstitute it *in vitro* using recombinant components with the ultimate goal to elucidate the molecular architecture the human RLC by means of X-ray crystallography and electron microscopy (EM). The following sections describe the heterologous production and characterization of recombinant RLC components, the *in vitro* reconstitution of the human RLC, analysis of the siRNA-binding properties of the reconstituted complex as well as structure determination attempts by means of X-ray crystallography and the electron microscopy structure of the recombinant RLC.

#### 6.1.1 Preparation of the human RISC-loading complex

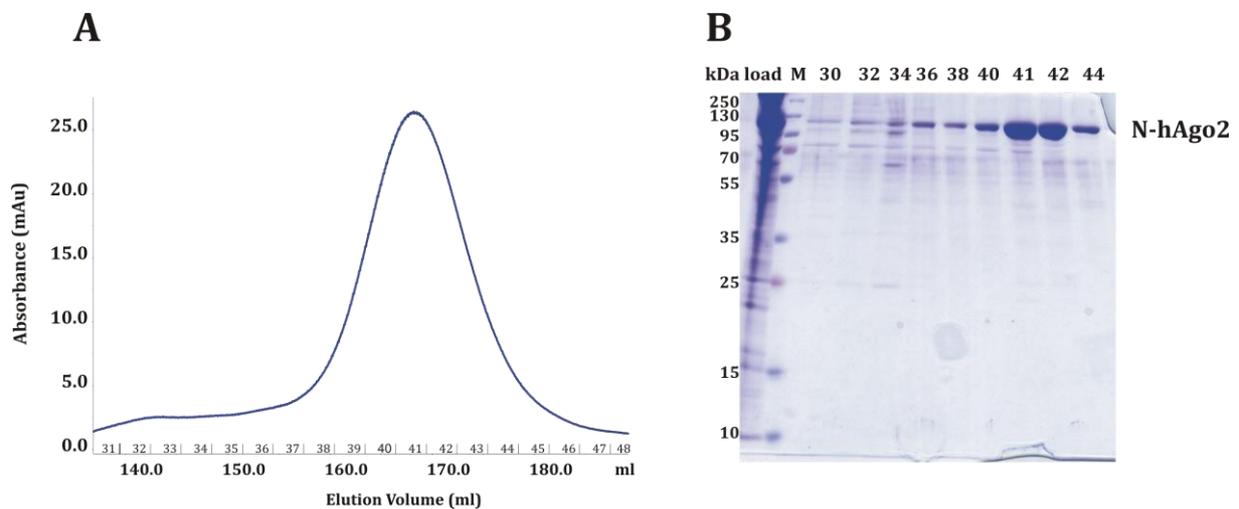
Most recombinant proteins used for structural studies were produced by overexpression in *E. coli*, which is a cheap and adaptable host. Nevertheless, many eukaryotic proteins fail to be expressed properly in *E. coli*. This may be due to the necessity of specific chaperone machineries or post-translational modifications for proper folding of these proteins that *E. coli* cannot support. The baculovirus/insect cell system is a eukaryotic expression system that allows the overexpression of such proteins (Jarvis and Summer 1989; Jarvis and Summers 1992; O'Reilly *et al.*, 1992). The large human Dicer and Ago2 proteins with sizes of 219 kDa and 97 kDa, respectively failed to be expressed or properly folded in different *E. coli* strains (own investigations and personal communication G. Meister). Initially, the aim was to reconstitute the RLC via a co-expression strategy by using the MultiBac system, which uses an engineered baculovirus DNA for later infection of insect cells and allows the simultaneous expression of multiple proteins in a single insect cell (Bieniossek *et al.*, 2008). Therefore, the genes coding for hDicer, hAgo2 and hTRBP2 were cloned into different MultiBac transfer vectors and their co-expression was tested in Sf9 and High Five insect cells (data not shown). The obtained expression rates of the human proteins were not detectable visually using SDS-PAGE or western blot. Shortly after the beginning of this study, in 2008, MacRae *et al.* successfully reconstituted the minimal human RLC using individually in Sf9 insect cell expressed and purified proteins. Additionally, MacRae *et al.* and later investigations showed that the expression of hTRBP2 in prokaryotic *E. coli* cells is sufficient for hTRBP2 function within the RLC (MacRae *et al.*, 2008; Chakravarthy *et al.*, 2010). Thus, only hAgo2 and hDicer were expressed as N-terminal hexahistidin-tagged fusion proteins in Sf9 and High five insect cells using the baculovirus

## Results

expression system from Invitrogen and hTRBP2 was expressed as a GST-fusion protein in *E. coli* BL21 (DE3) Star cells (5.1.12; 5.2.1.1).

### 6.1.1.1 Preparation of hAgo2

Human Ago2, which was expressed as a N-terminal hexahistidine tagged fusion protein (N-hAgo2) in High Five cells, can be purified in a monodisperse form as described in material and methods (5.2.2.6; 5.2.2.7; 5.2.3.5). A typical size exclusion chromatography profile of N-hAgo2 using a Superdex 200 (26/60) column and a 15% SDS-PAGE analysis of the peak fractions are shown in Figure 7. The human Ago2 protein functions as a monomeric protein within the RISCs (MacRae *et al.*, 2008) and N-hAgo2 elutes at 168 ml corresponding to an estimated molecular mass of 79 kDa. Thus, the obtained molecular mass is significantly lower than the calculated value of 100 kDa. This might be either due to a more compact folding or because of interactions with the column matrix during the gel filtration run. Additionally, the elution profile of hAgo2 did not change after cleavage of the NHis<sub>6</sub>-tag. hAgo2 is an siRNA- and miRNA-binding protein and for probable RNA-contaminations the absorptions at 260 nm and 280 nm were measured and the protein solution analyzed by urea-PAGE (5.2.3.1; 5.2.1.10). The 260 nm/280 nm value of hAgo2 was relatively high with 0.98 and on urea-PAGE small RNAs of sizes of 21-nts could be detected, indicating that hAgo2 might have become unspecifically loaded with endogenous insect cells RNAs (see 7.1.1 and Figure 54).



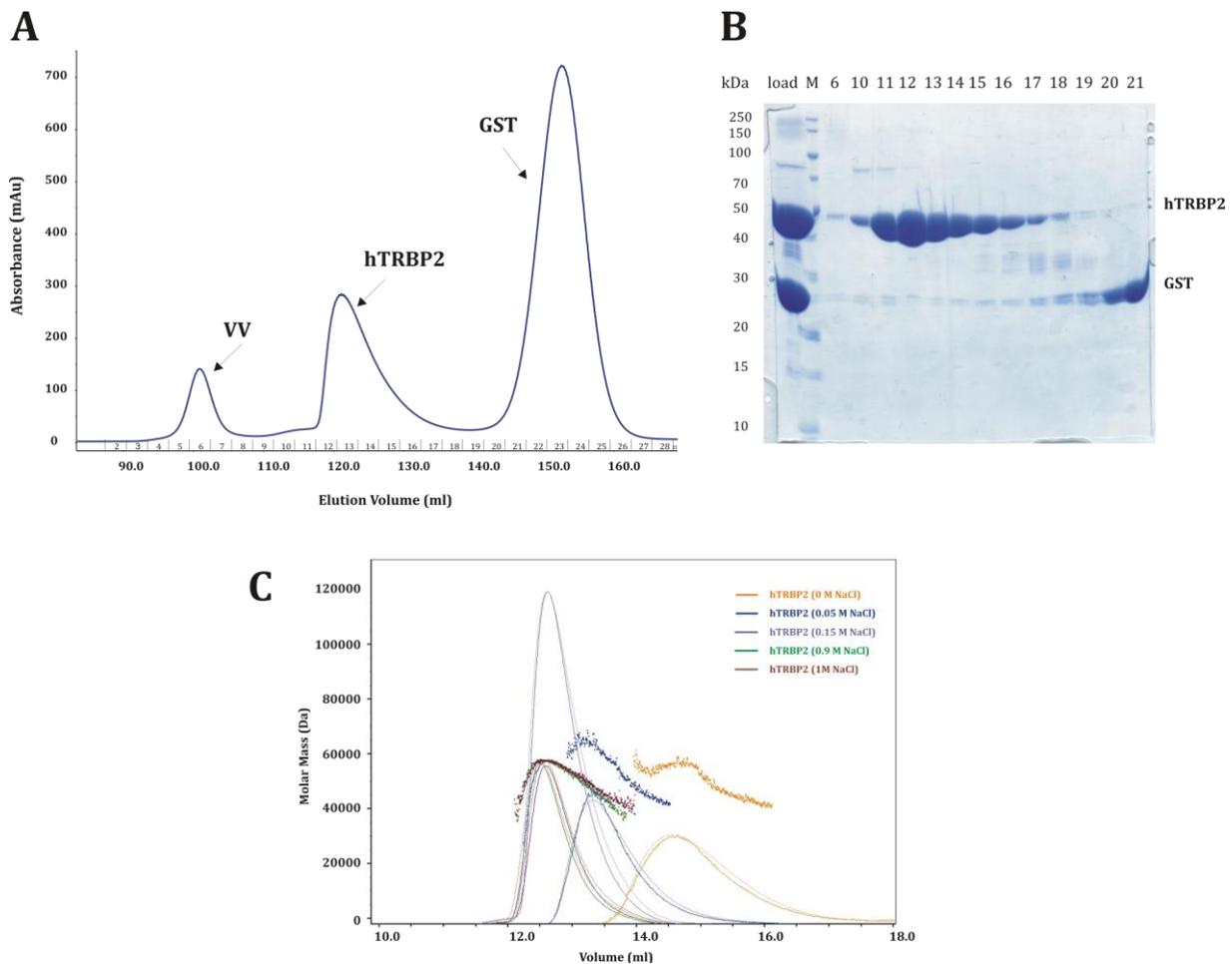
**Figure 7: Preparation of N-hAgo2.**

**A:** Superdex 200 (26/60) elution profile of full length N-hAgo2. Affinity-purified N-hAgo2 was applied on a Superdex 200 (26/60) column in buffer containing 150 mM NaCl, 20 mM HEPES pH 7.5, 5% (v/v) glycerol and 2 mM DTT. The elution profile, which was monitored by absorption at 280 nm in milli absorption units (mAu), is shown. **B:** Analysis of the peak fractions of the SEC of N-hAgo2 by SDS-PAGE with subsequent Coomassie staining. Protein samples from peak fractions were analyzed on a 15% SDS-gel, the identified N-hAgo2 (Ago2) protein is indicated on the right and a size marker is shown on the left of the gel. The numbers on top of the gel correspond to the fractions of the SEC chromatogram (M = marker).

### **6.1.1.2 Preparation of hTRBP2**

Human TRBP2, expressed as an N-terminal GST-fusion protein in *E. coli* BL21 (DE3) Star cells, can be purified in a RNA-free form using a protocol established during this work (5.2.2.5; 5.2.2.7; 5.2.3.4). A typical SEC profile of hTRBP2 using a Superdex 75 (26/60) column and a subsequent 15% SDS-PAGE analysis of the peak fractions are shown in Figure 8. TRBP2 has a theoretical molecular mass of 39 kDa, but elutes at 121 ml corresponding to an estimated molecular mass of 157 kDa. Thus, hTRBP2 has an unusual shape or forms higher oligomers. Analytical SEC of hTRBP2 revealed a molecular weight of 79 kDa corresponding to a hTRBP2-homodimer (data not shown). To determine the molecular mass of hTRBP2 more precisely SEC combined with MALS using different salt concentrations ranging from 0-1 M NaCl was carried out (Figure 8 C, 5.2.4.1). At low (0-0.025 M) NaCl concentrations hTRBP2 elutes later from SEC than at medium (0.15 M) and high (0.9-1 M) salt concentrations. This behavior can be due to a different overall shape of hTRBP2 or interactions of hTRBP2 with the column material at low salt concentrations. In all tested salt concentrations hTRBP2 is polydispers, which is already indicated by the strong tailing and the displacement of UV and light scattering. The average/mean molecular mass determined by MALS of hTRBP2 at all different concentrations is 60 kDa. At all different salt concentrations molar mass distributions varying between 66 kDa and 33 kDa were assigned. Hence, hTRBP2 forms compact homodimers and monomers at low and high salt concentrations. hTRBP2-homodimers seem to be stable (hTRBP2-homodimers exist at 1 M NaCl) and the homodimer formation seems to be independent of ionic interactions (hTRBP2-homodimers exist at 0 M NaCl).

## Results



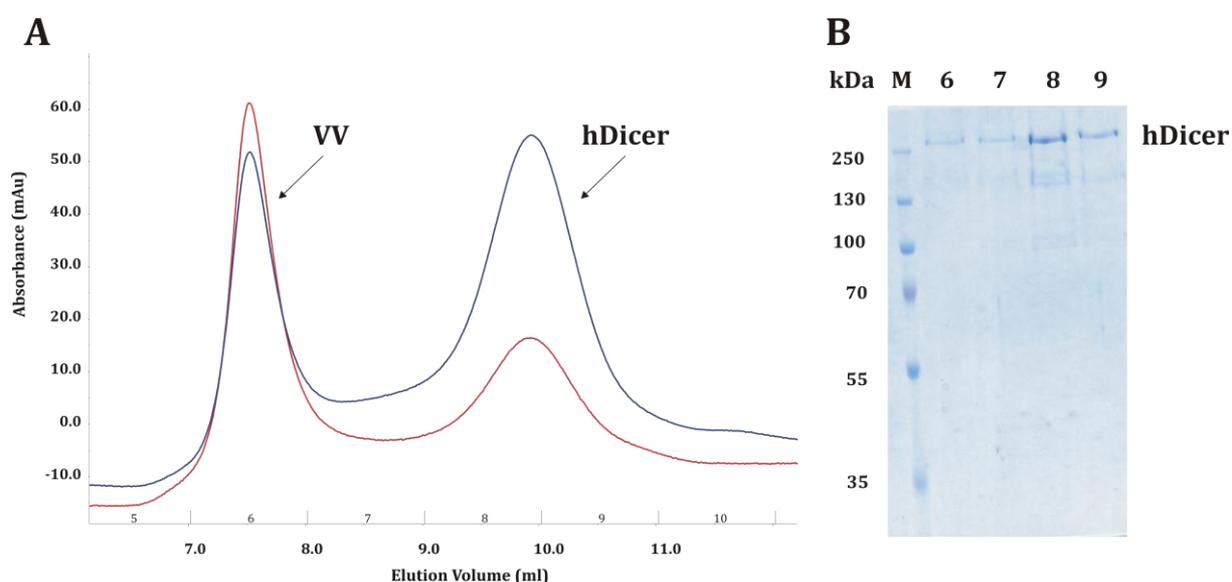
**Figure 8: Preparation of hTRBP2.**

**A:** SEC elution profile of full length hTRBP2. Affinity-purified hTRBP (incubated with PreScission protease) was applied on a Superdex S75 (26/60) column in buffer containing 150 mM NaCl, 20 mM HEPES pH 7.5, 5% (v/v) glycerol and 2 mM DTT. The elution profile, monitored by absorption at 280 nm (mAu), is shown. **B:** Analysis of the peak fractions of the SEC of full length hTRBP2 by SDS-PAGE with subsequent Coomassie staining. Protein samples from peak fractions were analyzed on a 15% SDS-gel, the identified proteins are indicated on the right and size marker is shown on the left of the gel. The numbers on top of the gel correspond to the fractions of the SEC chromatogram (M = marker; VV= void volume). **C:** SEC combined with MALS of TRBP2 at different NaCl concentrations ranging from 0-1 M NaCl. The SEC elution profiles, monitored by absorption at 280 nm (continuous lines) and by light scattering (dashed lines), and the calculated molecular mass points (points) of hTRBP2 at different concentrations are shown. hTRBP2 shows a different elution behavior at varying salt concentration, but had the same polydisperse size distribution with an average molecular weight of 65 kDa.

### 6.1.1.3 Preparation of hDicer

The expression rates of the N-terminal hexahistidine tagged hDicer protein in High Five or Sf9 insect cells using the original Dicer coding sequence were very low (data not shown). In order to increase expression rates a codon optimized hDicer gene for expression in Sf9 insect cells with an additional C-terminal Strep-tag was ordered from Geneart (10). The usage of the synthetic gene led to increased expression rates of hDicer using the MultiBac pFBDM transfer vector and the modified MultiBac-bacmid in High Five and Sf9 cells (5.1.12; 5.2.2.6). After affinity purification using a StrepTrap column the eluted protein was analyzed by SEC. The

CStrep-hDicer protein eluted in the void volume of Superdex or Superose SECs using different buffer systems, indicating that the protein is aggregated in solution and is probably not folded correctly (results not shown). In order to equip the CStrep-hDicer with an N-terminal hexahistidine tag and to perhaps produce soluble protein, the gene-optimized sequence of hDicer was cloned into a pFastBacHtb transfer vector (5.1.12, 5.2.1.1). This N-terminal His-tagged and C-terminal Strep-tagged hDicer (called N-hDicer in the following) can be expressed in High Five or Sf9 insect cells. Expression using the slower growing Sf9 cells produced less degraded protein and was therefore used for further expressions of N-hDicer (5.2.2.6). After harvesting of the Sf9 cells and cell lysis, a first affinity purification step using either a HisTrap- or a StrepTrap-column was carried out and led to already quite pure protein (5.2.2.7, 5.2.3.5). An analytical SEC using a Superose 12 (10/300) column and a subsequent SDS-PAGE analysis of the N-hDicer protein are shown in Figure 9. The N-hDicer elutes at 10 ml from a Superose 12 (10/300) column corresponding to an apparent weight of 251 kDa, which is slightly higher than the theoretical molecular weight of 223 kDa. To check hDicer for RNA-impurities the absorption at 260 nm and 280 nm of the protein solution was measured and the protein solution was analyzed by urea-PAGE (5.2.1.10). The 260 nm/280 nm value of hDicer was 0.6 and no RNA could be detected on urea-PAGE, indicating a RNA-free protein solution (results not shown).



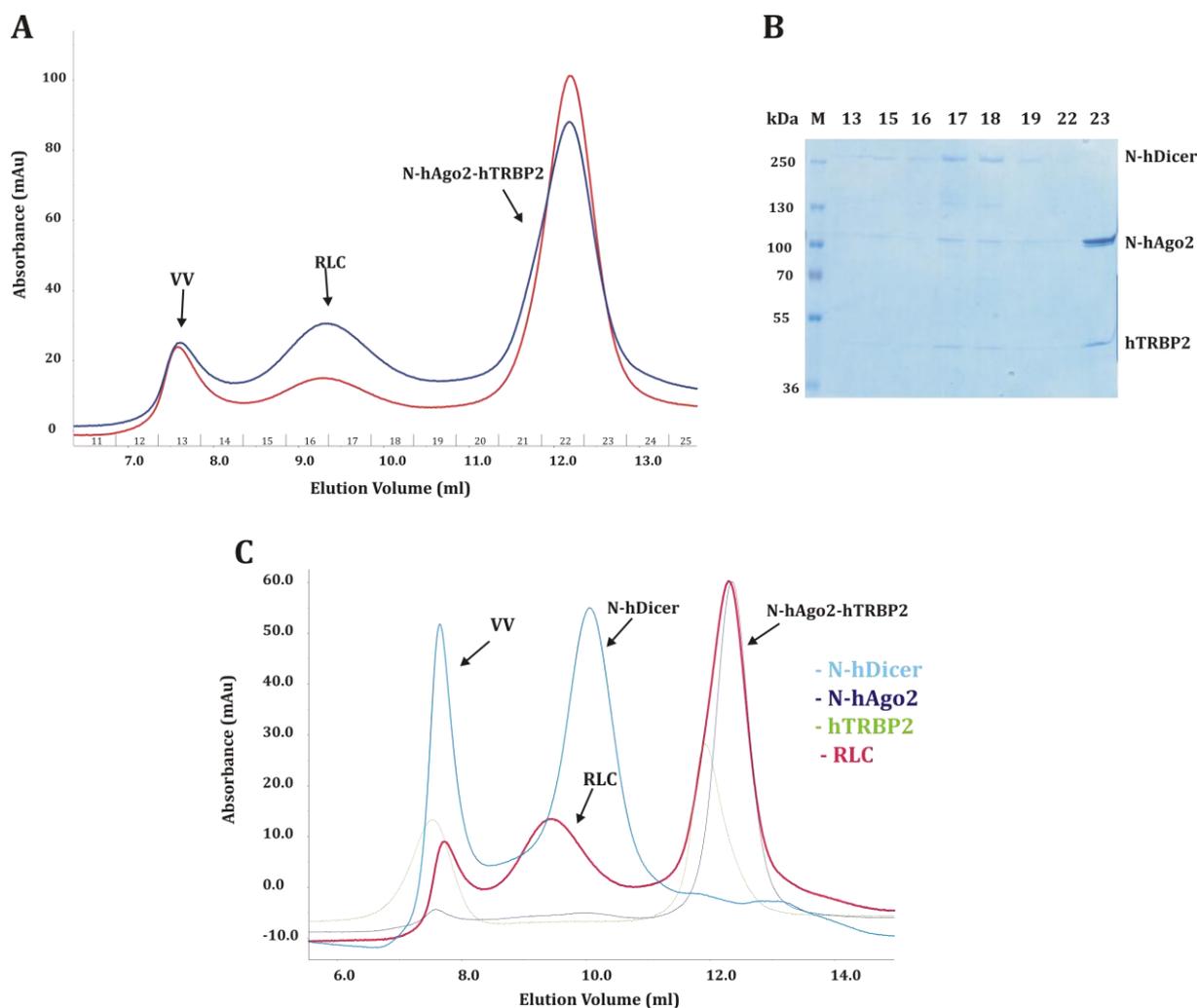
**Figure 9: Preparation of full length N-hDicer.**

**A:** Superose 12 (10/300) elution profile of full length N-hDicer. Affinity-purified N-hDicer was applied on a Superose 12 (10/300) column in buffer containing 150 mM NaCl, 20 mM HEPES pH 8.0, 5% (v/v) glycerol and 2 mM DTT. The elution profile, monitored by absorption at 280 nm (mAU), is shown. **B:** Analysis of the peak fractions of the SEC of full length N-hDicer by SDS-PAGE with subsequent Coomassie stain. Protein samples from peak fractions were analyzed on a 10% SDS-gel, the identified N-hDicer protein is indicated on the right and size marker is shown on the left of the gel. The numbers on top of the gel correspond to the fractions of the SEC chromatogram (M = marker, VV= void volume).

#### **6.1.1.4 In vitro reconstitution of a minimal human RLC**

In order to reconstitute the human RLC the affinity purified N-hDicer protein was incubated with purified hAgo2 and hTRBP2 proteins in a molar ratio of 1:3:3 for 1 hour at 4 °C and applied to a Superose 12 (10/300). A SEC elution profile and a corresponding 10% SDS-PAGE analysis are depicted in Figure 10. In the chromatogram three absorption peaks can be detected, with the first one eluting at 7.5 ml corresponding to the void volume of the Superose 12 (10/300) column where aggregated proteins elute. The second elution peak at 9.3 ml corresponds to the RLC containing N-hDicer, N-hAgo2 and hTRBP2. The third elution peak at 12-12.5 ml corresponds to unincorporated N-hAgo2 and hTRBP2 (Figure 10 B). All three proteins present in the second peak elute with a total molecular mass of about 400 kDa. This molecular mass is slightly larger than the sum of the calculated molecular masses of the individual proteins (theoretical MW = 362 kDa).

In Figure 10 C, Superose 12 (10/300) elution profiles of N-hDicer alone (in light blue), N-hAgo2 alone (in purple), hTRBP2 alone (in light green) and the reconstituted RLC (in red) are superimposed. In the elution profiles of N-hDicer, hTRBP2 and reconstituted RLC a void volume peak can be detected. These aggregated proteins can be separated from single proteins or protein complexes using a Superose 12 (10/300) column. N-hAgo2 monomers and hTRBP2-homodimers elute at approximately the same volume of 12-12.5 ml, corresponding to a size of about 70 kDa. The Dicer monomer elutes at 10 ml equivalent to an apparent molecular mass of 251 kDa. Compared with the elution of Dicer, the elution peak of the RLC is shifted to a lower volume. Taken together the SEC studies indicate that N-hDicer, N-hAgo2 and hTRBP2 spontaneously assemble into a stable complex *in vitro*.



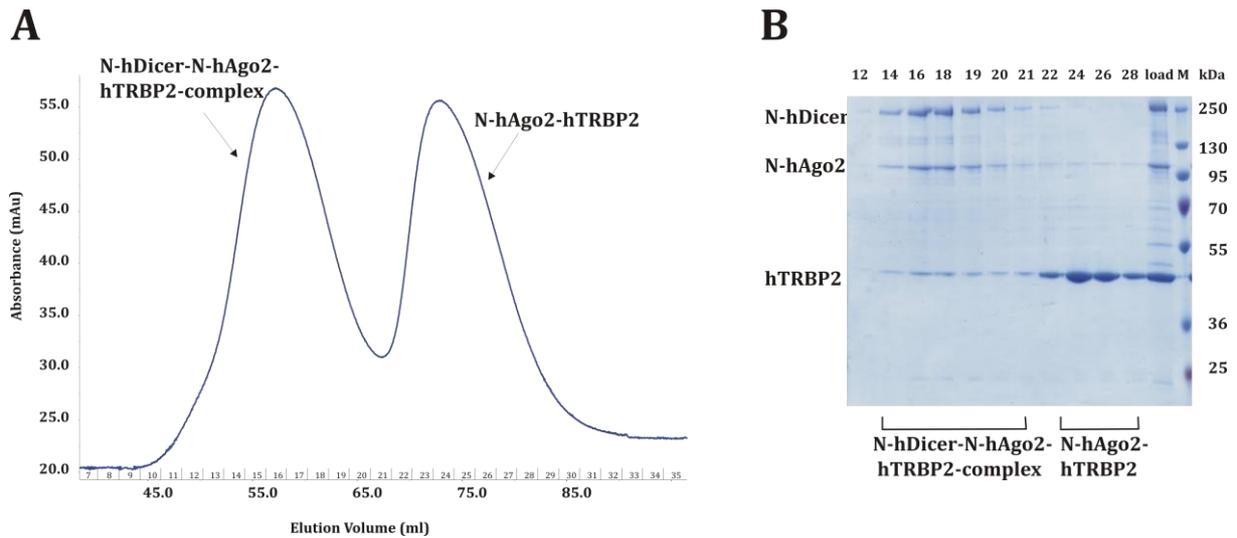
**Figure 10: *In vitro* reconstitution of the minimal human RISC-loading complex (RLC).**

**A:** SEC elution profile of full length N-hDicer-N-hAgo2-hTRBP2 complex (RLC). Affinity purified N-hDicer protein was incubated with purified N-hAgo2 and hTRBP2 proteins in a molar ratio of 1:3:3 for 1 hour at 4 °C and applied to a Superose 12 (10/300) column in a buffer containing 150 mM NaCl, 20 mM HEPES pH 8.0, 5% (v/v) glycerol and 2 mM DTT. The elution profile, monitored by absorption at 280 nm and 260 nm, is shown. **B:** SDS-PAGE analysis of the peak fractions of the SEC of RLC. Protein samples from peak fractions were analyzed on a 10% SDS-gel, the identified proteins are indicated on the right and size marker is shown on the left of the gel. **C:** SEC elution profiles of RLC, N-hDicer, N-hAgo2 and hTRBP2. Different amounts of N-hDicer (light blue), N-hAgo2 (purple), hTRBP2 (light green) and the RLC have been analyzed using a Superose 12 (10/300) SEC column.

Although, the RLC could be reconstituted by uniting separately expressed and purified proteins, the 223 kDa large N-hDicer turned out to be an unstable protein (see Figure 47) and could not reproducibly be purified in a monodisperse form. Thus, an alternative co-purification strategy to reconstitute the minimal human RLC in a preparative scale was developed. In most complexes binding partners stabilize proteins and also hDicer seems to be stabilized by hAgo2 and hTRBP2 (MacRae *et al.*, 2008; Wang *et al.*, 2009b). Thus, the N-hDicer-N-hAgo2 complex was first co-purified using a HisTrap column according to methods described in 5.2.3.5. Optionally, the N-His<sub>6</sub>-tags of N-hDicer and N-hAgo2 were removed by TEV cleavage. The affinity purified N-hDicer-N-hAgo2 complex was then incubated with purified hTRBP2 in a molar ratio

## Results

of 1:3 in HisTrap elution buffer supplemented with 10 mM DTT and 10 mM EDTA at least for 1 hour at 4 °C. This protein solution was applied to Superdex 200 (16/60) column in a buffer containing 150 mM KCl, 20 mM Imidazole pH 6.8, 3% (v/v) glycerol and 2 mM DTT. The peak fractions eluted from the SEC were analyzed using 10% SDS-PAGE. The human N-hDicer-N-hAgo2-hTRBP2 complex (RLC) elutes from a Superdex 200 (16/60) column at 57 ml corresponding to an estimated molecular weight of 400 kDa (Figure 11).



**Figure 11: *In vitro* Reconstitution of the RLC in preparative scale.**

**A:** Superdex 200 (16/60) elution profile of human RLC. Affinity-purified N-hDicer-N-hAgo2 complex was incubated with purified hTRBP2 in a ratio of 1:3 and the protein solution was applied to a Superdex 200 (16/60) column. The elution profile, which was monitored by absorption at 280 nm, is shown. The distribution of RLC and unincorporated N-hAgo2-hTRBP2 are denoted. **B:** SDS-PAGE analysis of the peak fractions of the SEC of human RLC. Protein samples from peak fractions of the SEC of the RLC were analyzed on a 10% SDS-gel, the identified proteins are indicated on the left and size marker is shown on the right of the gel. The distribution of N-hDicer-N-hAgo2-hTRBP2 complex (RLC), N-hAgo2 and hTRBP2 are denoted below the gel.

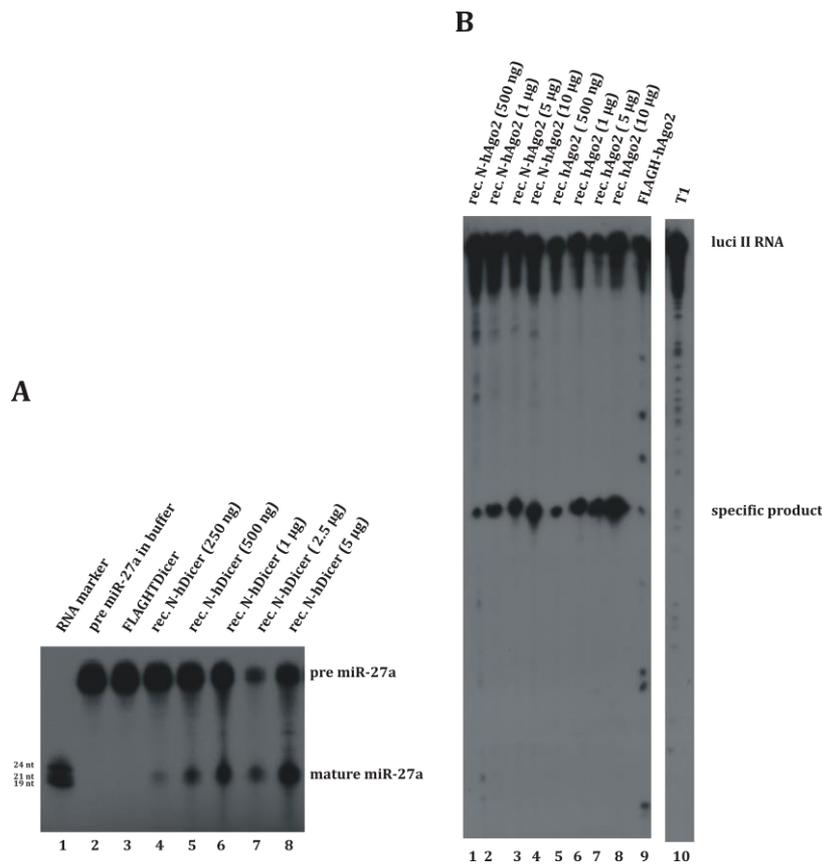
### **6.1.2 Functional analysis of the RLC proteins hDicer, hAgo2 and hTRBP2**

The human RLC composed of hDicer, hAgo2 and hTRBP2 couples the processing of pre-miRNA substrate by the hDicer to the loading and duplex unwinding of the miRNA product onto hAgo2 (4.3.1). To determine if the proteins of the reconstituted RLC are functional, the endonuclease activities of recombinant hDicer and hAgo2 were assayed and the antisense siRNA and siRNA duplex binding properties of recombinant hTRBP2 were tested. Furthermore, the antisense siRNA and siRNA duplex binding and hAgo2-loading of the reconstituted RLC were examined.

#### **6.1.2.1 Endonuclease activities of human Dicer and Argonaute2**

The human endoribonuclease Dicer cleaves pre-miRNAs and pre-siRNA (dsRNAs) substrates into 20-25 nucleotide long miRNA or siRNAs with a 2-nt overhang on the 3'-end (Bernstein *et al.*, 2001). To assay the recombinant N-hDicer for the pre-miRNA processing activity, different amounts of protein were incubated with a <sup>32</sup>P-labeled miR-27a pre-RNA and accumulation of mature miR-27a was analyzed by urea-PAGE (5.1.13; 5.2.1.11; 5.2.1.10; Figure 12). The recombinant N-hDicer is able to incorporate the miR-27a pre-RNA and endonucleolytically cleave it to produce mature miR-27a. As anticipated, mature miR-27a accumulation rises with increasing protein concentrations (see lanes 4-8 of Figure 12). This assay demonstrates that the recombinant N-hDicer is a functional endoribonuclease being able to process pre-miRNAs and that neither the N-terminal His-tag nor the C-terminal Strep-tag prevent cleavage.

Human Ago2 has been shown to endonucleolytically cleave an mRNA target when loaded with an appropriate miRNA/siRNA guide that is perfectly complementary to this mRNA target (4.2.2.2; Liu *et al.*, 2004; Meister *et al.*, 2004; Rana *et al.*, 2005). The purified recombinant N-hAgo2 and hAgo2 proteins were tested for this RISC activity (5.2.1.12). Both, the purified recombinant N-hAgo2 and hAgo2 are able to load a single stranded RNA (guide RNA) and to slice the complementary luciferase siRNA (target RNA) without the need of additional protein partners. The accumulation of sliced product increases with expanding protein concentrations (see lanes 1-4 and 5-8 from Figure 12).



**Figure 12: Endonuclease activities of hDicer and hAgo2.**

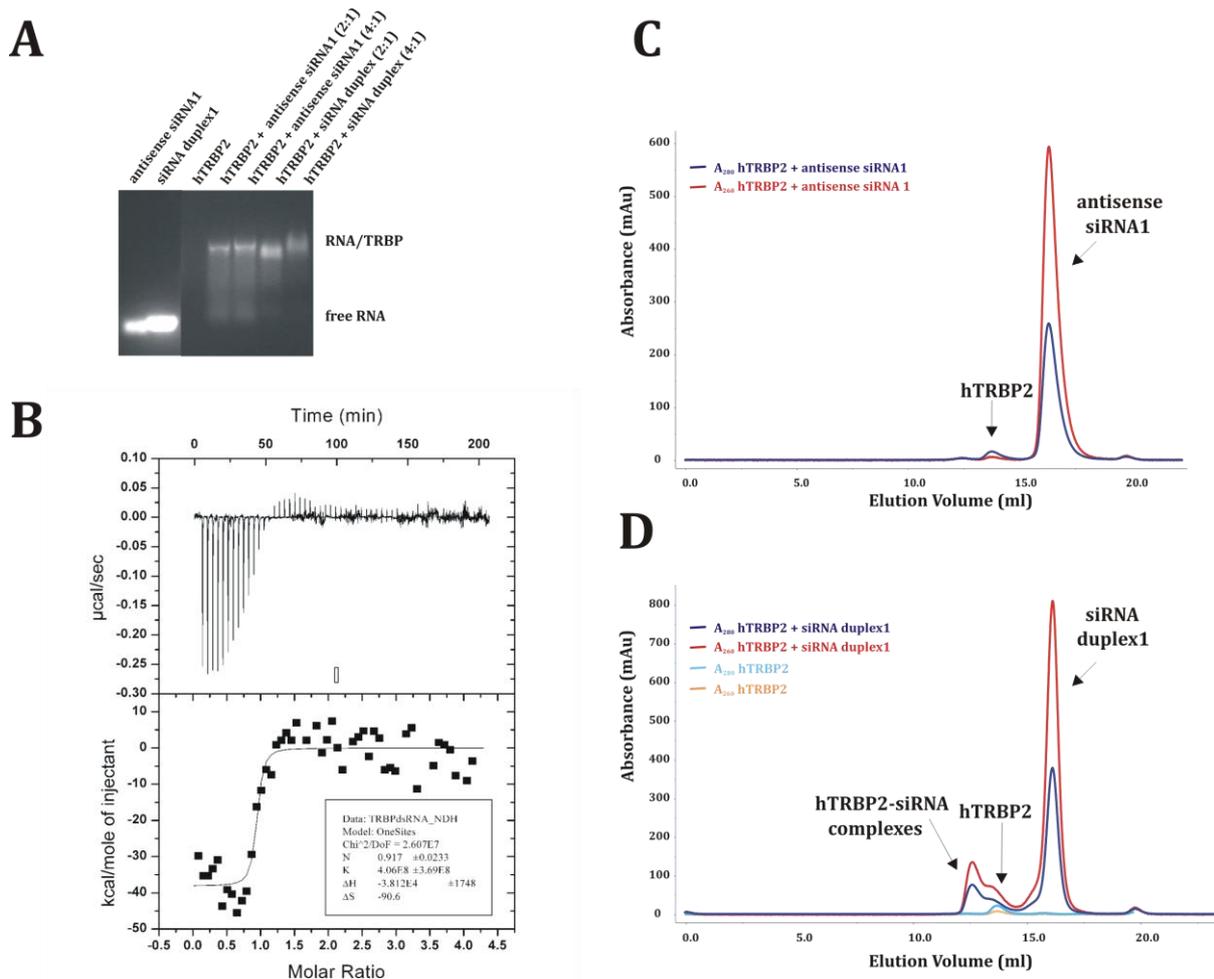
**A:** Analysis of Dicer activity. A nucleotide maker is shown on the left and the RNA substrate and product are indicated on the right. In lane 2 the <sup>32</sup>P-labeled miR-27a pre-RNA in buffer and in lane 3 <sup>32</sup>P-labeled miR-27a pre-RNA incubated with FLAGHT-Dicer immunopurified from HEK cells by A. Frohn was applied. This Dicer had no endonucleolytic cleavage activity. Within lane 4-8 the cleavage of <sup>32</sup>P-labeled miR-27a precursor RNA to mature miR-27a by different amounts of recombinant N-hDicer is shown. **B:** RISC assay analysis of recombinant N-hAgo2 and hAgo2. The cleavage of luciferase (luc) II siRNA target RNA by different amounts of recombinant human N-hAgo2 and hAgo2 are shown in lanes 1-4 and lanes 5-8, respectively. T1 refers to partial nuclease T1 digestion of the miR-21 targeted RNA substrate. In lane 9, as a positive control, FLAGHA-Ago2 immunopurified from HEK cells has been used.

### 6.1.2.2 siRNA-binding of human TRBP2

Generally, dsRBPs bind their substrates independent of the RNA sequence, while the hTRBP2 was shown to have preferences for GC-rich sequences (Gatignol *et al.*, 1991; Lee *et al.*, 1996; Saunders & Barber, 2003). Furthermore, human TRBP2 binds small siRNAs with higher affinity than long dsRNAs in a non-cooperative manner (Parker *et al.*, 2008; Chakravarthy *et al.*, 2010). To monitor the binding of recombinant hTRBP2 to antisense siRNA and siRNA duplexes *in vitro* EMSAs, analytical SECs and ITC-experiments were performed (5.2.1.9; 5.2.3.7; 5.2.4.3). For EMSAs antisense siRNA1 and siRNA duplex1 were incubated with hTRBP2 in molar ratios of 1:2 and 1:4 and analyzed on agarose gels (5.2.1.9). In the presence of the antisense siRNA1 or siRNA duplex1 well-defined band shifts could be observed, representing the hTRBP2 antisense siRNA1 complexes or hTRBP2 siRNA duplex1 complexes, respectively (Figure 13 A). The binding

to the siRNA duplex1 by hTRBP2 seems to be stronger than the binding to the single stranded antisense siRNA1. The whole duplex siRNA is shifted forming a prominent band upon addition of different amounts of hTRBP2, whereby the single stranded RNA shift experiments result in smeary bands and some RNA is not shifted upon addition of hTRBP2 indicating a weak binding. Analytical SEC runs of hTRBP2 incubated with antisense siRNA1 or siRNA duplex1 revealed that no stable complex formation of TRBP2 with an antisense siRNA can be detected (Figure 13 B) and that hTRBP2 forms larger complexes upon siRNA duplex1 binding (Figure 13 C). In the SEC elution profile of hTRBP2 incubated with siRNA duplex1 two overlapping elution peaks and a single RNA peak were observed. Analysis by SDS-PAGE and urea-PAGE revealed that both overlapping peaks contain hTRBP2-siRNA duplex1 complexes (data not shown). The larger elution peak at 12.5 ml corresponds to a size of about 160 kDa and the smaller elution peak at 13.7 ml corresponds to a size of circa 89 kDa. When separated and rerun on SEC these two different complexes elute at the original elution volume, indicating that they are stable (data not shown). By addition of the calculated molecular masses of hTRBP2 (39 kDa) and the siRNA duplex (21 kDa), one can hypothesize that the roughly 160 kDa complex may contain two siRNA duplexes and three hTRBP2 molecules (159 kDa) or one siRNA duplex and four hTRBP2 molecules (177 kDa) and that the 89 kDa complex may consist of one siRNA duplex and a hTRBP2-homodimer (99 kDa). Thus, SEC indicated that various hTRBP2-RNA complexes can be formed, but which of those might be physiological/functional will be discussed later (see 7.1.3). Initial ITC-experiments of hTRBP2 with siRNA duplex1 revealed a high affinity of hTRBP2 to siRNA duplex1 with a  $K_d$  of  $0.246 \pm 0.223$  nM (5.2.4.3, Figure 13 B). Further ITC experiments to improve data quality were not carried out as in the mean time Yamashita *et al.* published detailed ITC measurements of hTRBP2 with a similar siRNA duplex and assigning a  $K_d$  of 0.24 nM (Yamashita *et al.*, 2011, and see 7.1.3). Additionally, these experiments showed that this binding event is an enthalpy driven and not entropy driven process ( $-\Delta H$  and  $-\Delta S$  values). Thus, most likely hTRBP2 interacts with the siRNA phosphate backbone through ionic interactions or hydrogen bonds (see discussion 7.1.3).

## Results



**Figure 13: hTRBP2 siRNA interaction.**

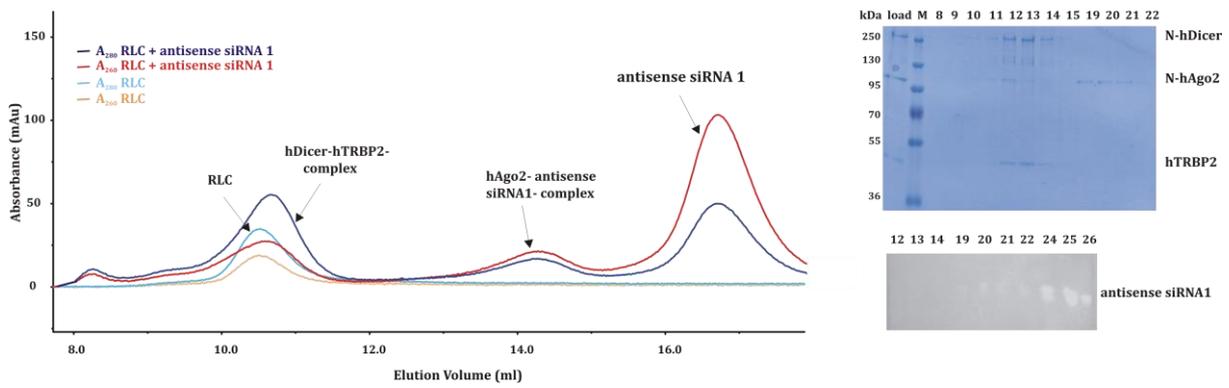
**A:** Electrophoretic mobility shift assay of hTRBP2 with antisense siRNA1 and siRNA duplex1. When incubated with hTRBP2 the antisense siRNA1 and the siRNA duplex1 are shifted to a distinct band, whereby the antisense siRNA1 interaction with hTRBP2 leads to smears on the gel and is more weakly bound by hTRBP2. **B:** Isothermal titration calorimetry of hTRBP2 with siRNA duplex1. The upper panel shows raw data of heat effect (in  $\mu\text{cal}\cdot\text{s}^{-1}$ ) of injections of siRNA duplex1 into hTRBP2 solution. The lower panel shows the fitted binding isotherms. hTRBP2 bind the siRNA duplex with a high affinity ( $K_d = 0.246 \pm 0.223 \mu\text{M}$ ) in an enthalpy and not entropy driven process ( $-\Delta H$  and  $-\Delta S$ ). **C:** Superdex 200 (10/300) elution profile of hTRBP2 incubated with antisense siRNA1. No hTRBP2-antisense siRNA complex is formed. **D:** Superdex 200 (10/300) elution profiles of hTRBP2 and hTRBP2 incubated with siRNA duplex1. Upon addition of siRNA duplexes hTRBP2 forms higher oligomers.

### 6.1.2.3 RISC-loading ability of the *in vitro* reconstituted RLC

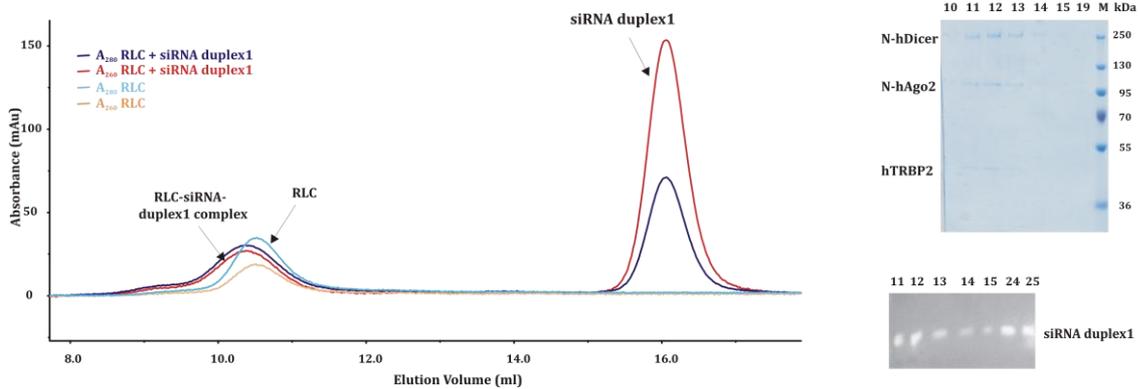
The RISC containing hDicer, hAgo2 and hTRBP2 was renamed RLC because of its ability to specifically load a miRNA guide onto hAgo2 (Maniataki & Mourelatos, 2005). In order to test if the *in vitro* reconstituted RLC is able to load hAgo2, analytical SEC runs using the recombinant RLC incubated with different siRNAs were performed (5.2.3.7, Figure 14 A-C). Prior to SEC the *in vitro* reconstituted RLC was incubated with antisense siRNA1 (Figure 14 A), siRNA duplex1 (Figure 14 B) or siRNA duplex2 (Figure 14 C) in a molar ratio of 1:3 on ice for 1 hour and separated on a SEC. Peak fractions of the SEC run were analyzed by SDS-PAGE and urea-PAGE (5.2.1.10; 5.2.3.2). When incubated with antisense siRNA1 the RLC disassembles into the hDicer-hTRBP2 complex and the hAgo2-antisense siRNA1 complex (Figure 14 A). Thus, within the RLC, hAgo2 gets loaded with the antisense siRNA1 (guide siRNA) forming a mature RISC *in vitro* and the RLC disassembles upon this loading. The RLC incubated with the siRNA duplex1, which contains the same siRNA guide, binds this siRNA duplex (as indicated by an increase in 260/280 ratio) and shift to a higher volume on SEC (Figure 14 B). This increase in size to a molecular mass of 436.5 kDa (molecular mass of RLC alone determined by SEC is 400 kDa) can be explained by binding of a single or two siRNA duplex molecules to the RLC (MW of 419 kDa or 440 kDa, respectively). The RLC incubated with the siRNA duplex2 shifts to a higher volume on SEC corresponding to a size of 426 kDa and also releases the hAgo2-siRNA duplex2 complex (Figure 14 C). Some hAgo2 molecules are loaded with the siRNA duplex1 and this hAgo2-RNA complex disassemble from the RLC and some hAgo2 molecules stay within the complex. Probably the higher molecular weight of the RLC can be explained by the structural rearrangements, structural flexibility and dynamics. In summary, the *in vitro* reconstituted RLC is able to load hAgo2 with a guide siRNA (Figure 14 A) and a siRNA duplex (Figure 14 C) and the RLC disassembles upon this loading. The hAgo2 loading is Mg<sup>2+</sup>-independent and seems to be sequence dependent (loads siRNA duplex2, but not siRNA duplex1). The siRNA duplex1 has a lower affinity than the siRNA duplex2 to hAgo2 (personal communication Jochen Deckert, Roche Kulmbach). Thus, if hAgo2 is not loaded with the siRNA duplex due to a lower affinity within the RLC, the RLC binds this siRNA, but does not disassemble.

## Results

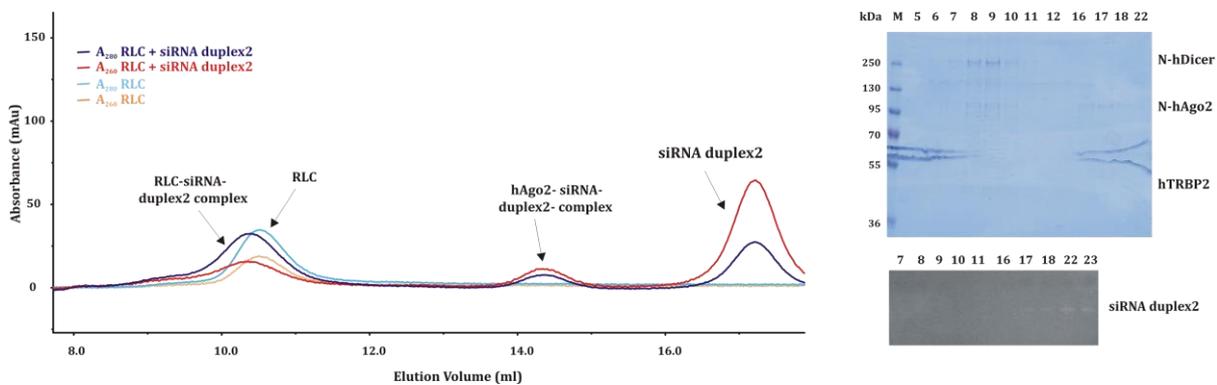
### A



### B



### C



**Figure 14: Analytical SEC, SDS-PAGE and urea-PAGE analysis of recombinant human RLC with different siRNAs.**

The recombinant RLC was incubated with antisense siRNA1 (A), siRNA duplex1 (B) or siRNA duplex2 (C) in a molar ratio of 1:3 for 1 hour on ice and applied to a Superdex 200 (10/300) column in buffer containing 150 mM KCL, 20 mM Imidazole pH 6.8, 2 % glycerol and 1 mM DTT. The elution profiles of the SECs were monitored by absorption at 260 nm as well as 280 nm and are depicted in the left part of the picture. The analytical S200 (10/300) elution profile of the RLC alone is overlaid with the elution profile of the RLC incubated with siRNA. In the upper right the SDS-PAGE analysis and within the lower right the urea-PAGE analysis of the peak fractions of the SEC run of the RLC with the siRNA are shown. **A:** Disassembly of the human RLC upon antisense siRNA1 binding into N-hDicer-hTRBP2 and N-hAgo2-antisense siRNA1 complexes. **B:** Binding of the RLC to siRNA duplex1. **C:** Partial disassembly of the human RLC upon siRNA duplex2 binding into N-hDicer-hTRBP2 and N-hAgo2-siRNA duplex2 complexes.

It has been controversy discussed if the RLC needs ATP for some rearrangement events to facilitate loading of a siRNA or miRNA onto hAgo2 (Iwasaki *et al.*, 2010;Yoda *et al.*, 2010). To test if our reconstituted RLC hydrolyses ATP to load hAgo2, photometric ATPase assays were performed (5.2.4.4). Six different single and double stranded siRNAs as well as two independent RLC preparations were used to test for ATPase activity of the RLC. Using this setup no ATP hydrolysis and thus no ATPase activity of the recombinant RLC could be detected (data not shown). Hence, one can conclude that ATP is not needed for binding of these RNAs or loading of at least some siRNAs (antisense siRNA1 and siRNA duplex2) onto hAgo2 (see Figure 14).

### **6.1.3 MALS and thermal shift assays of the minimal human RLC**

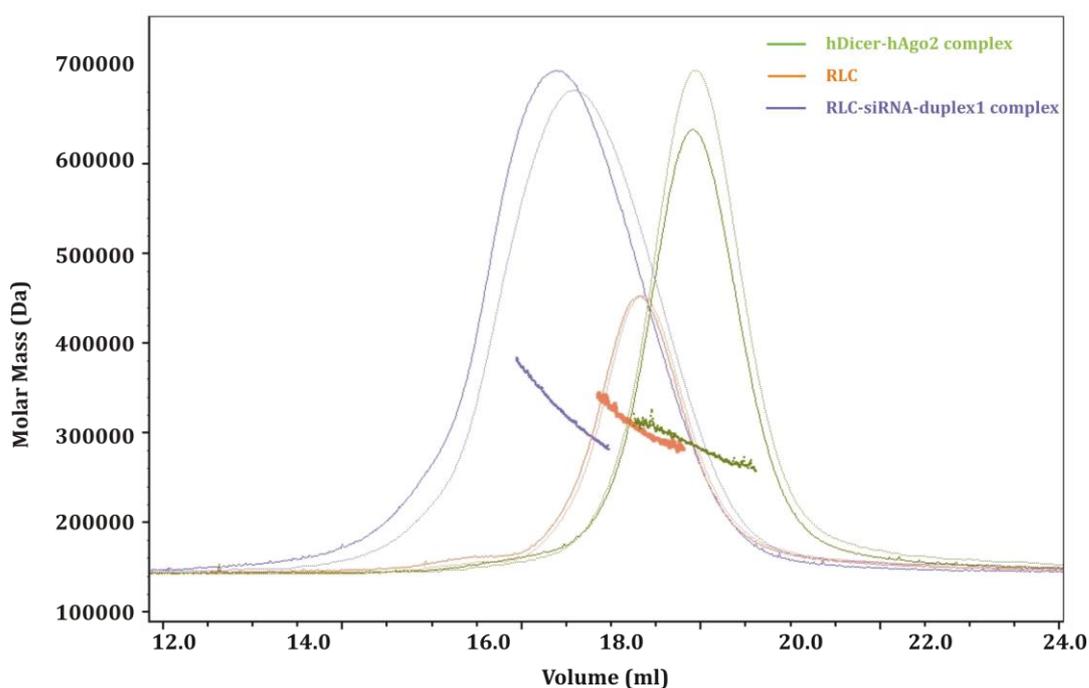
Key factors that affect the probability to obtain a 3D structure of macromolecular complexes by means of crystallography or electron microscopy are the homogeneity (chemical and conformational), the stability and the solubility of the complex. To analyze and optimize these factors SEC combined with MALS analysis and several thermal shift assays of the RLC and the RLC-siRNA duplex1 complex were carried out and are described below.

#### **6.1.3.1 MALS analysis of the RLC**

To determine the correct molecular mass, the dispersity and aggregation of the minimal RLC and the RLC-siRNA duplex1 complex in solution, SEC combined with MALS analysis were carried out in a buffer containing 150 mM KCl, 20 mM Imidazole pH 6.8, 3 % glycerol and 2 mM DTT (5.2.4.1). The results indicate that the *in vitro* reconstituted RLC is a polydisperse protein complex with an average molecular mass of 305 kDa and a broad size distribution of particles ranging from 275 kDa to 360 kDa (Figure 15). These molecular masses are about the sum of the molecular masses of the individual proteins within the RLC (362 kDa) or N-hDicer-hTRBP2 (262 kDa) and N-hDicer-N-hAgo2 (323 kDa) subcomplexes. Using SEC, the RLC elutes at a volume contributing to a size of about 398 kDa (see section 6.1.1), assuming a molecular ratio of 1:1:2 of hDicer, hAgo2 and hTRBP2 in the RLC. On the contrary, MALS data suggest a 1:1:1 composition of hDicer, hAgo2 and hTRBP2 within the RLC. However, the average determined MW of 305 kDa with the high polydispersity ( $\pm 50$  kDa) is indicative for an imperfect incorporation of hTRBP2 or hAgo2. Additionally, this data give rise to the possibility that hTRBP2 can be incorporated into the RLC as a monomer as well as a dimer. The average molecular mass of the N-hDicer-N-hAgo2-complex determined by SEC combined with MALS is 285 kDa, which is slightly smaller than the sum of the molecular masses of the two proteins (315 kDa). This complex is also very polydisperse with molecular masses ranging from 250-320 kDa (Figure 15). Using SEC, the RLC-siRNA duplex1 complex elution peak is clearly shifted to a higher molecular mass compared to the elution of the RLC alone (Figure 14B, Figure 15). Additionally, the broad elution peak of the RLC-siRNA duplex1 indicates the polydispersity of that complex and SEC combined with MALS

## Results

confirmed the hypothesis. The average molecular weight of this protein-RNA complex was determined to be 324 kDa with sizes varying from 266–380 kDa (see Figure 15). This determined molecular mass range indicates that this complex can be composed of N-hDicer, N-hAgo2, hTRBP2 and siRNA duplex1 in a ratio of 1:1:1:1 (362 kDa). The polydispersity ( $\pm 60$  kDa) indicates a non-stoichiometric incorporation of individual components and thus the coexistence of different complexes and subcomplexes in solution. Furthermore one has to consider that the method of MALS is not precise for protein-RNA complexes, because the native extinction coefficient cannot exactly be determined *in silico* and thus an accurate composition of such a complex is not demanding.



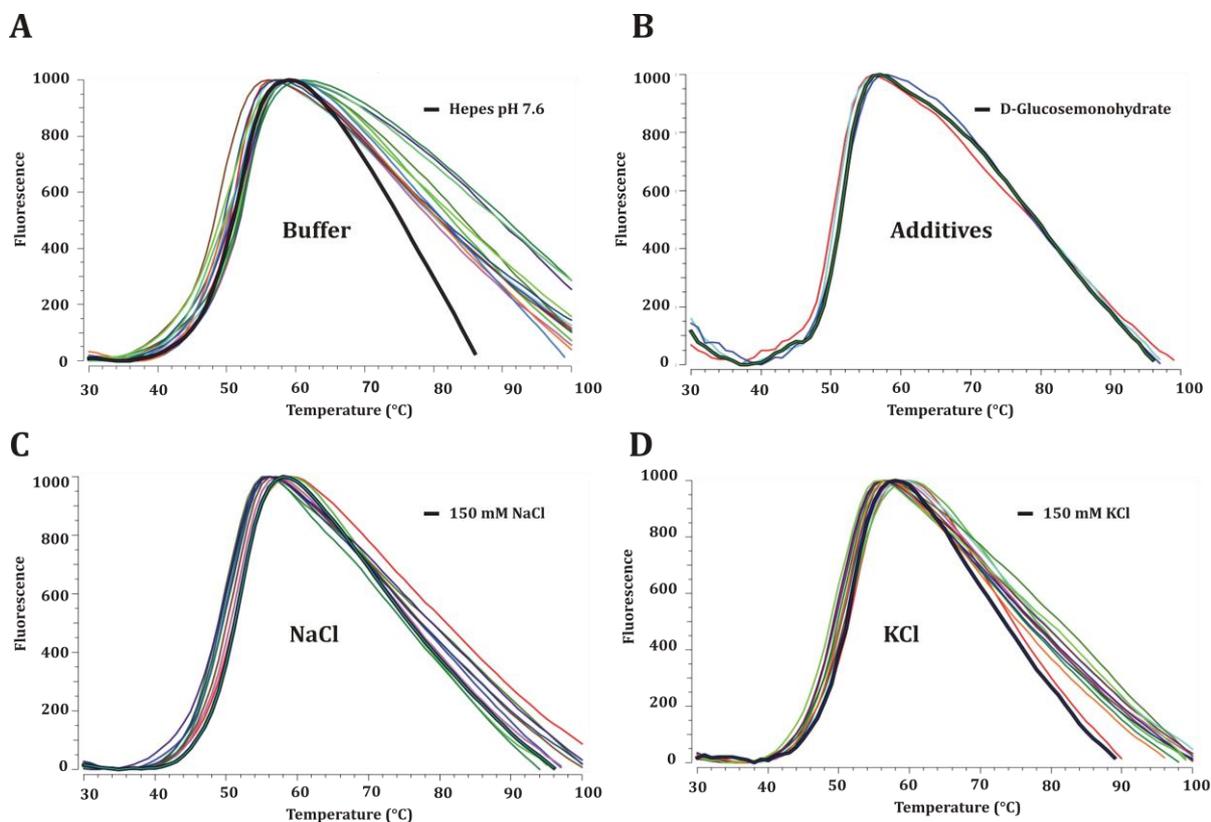
**Figure 15: SEC combined with MALS of the *in vitro* reconstituted RLC-siRNA duplex1 complex, the RLC and the N-hDicer-N-hAgo2 complex.**

The SEC elution profiles, monitored by absorption at 280 nm (continuous lines), by light scattering (thin lines), and the calculated molecular mass points (points) of RLC-siRNA duplex1 complex (purple), of the RLC (orange) and the N-hDicer-N-hAgo2 complex (green). The three different complexes are polydisperse with average molecular weights of 324 kDa (RLC-siRNA duplex1 complex), 305 kDa (RLC) and 285 kDa (N-hDicer-N-hAgo2 complex).

### **6.1.3.2 Fluorescence-based thermal shift assays of the RLC**

One method used to determine the influence of manifold factors such as pH, ionic strength, additives and precipitants on stability of a protein is the fluorescence-based thermal-shift assay (5.2.4.2). To test the influence of different buffers and pH's on the stability of the RLC, the pHAT screen (Emerson) was used for the thermal shift assays as described in 5.2.4.2 (Figure 16 A). These studies revealed that the buffer already used (HEPES, pH 7.6) seemed to be one of the optimal buffers (Figure 16 A) as well as some others like Imidazole, Phosphate, HEPES and Bis-

Tris buffers at pH between 6.6 and 7.6, the RLC is quite stable, indicated by the relatively high melting temperatures of about 51 °C. As expected, buffers with extreme pH like MES pH 4.6 or Tris pH 9.4 led to destabilization of the RLC (curves not shown).



**Figure 16: Fluorescence-based thermal shift assays of the RLC.**

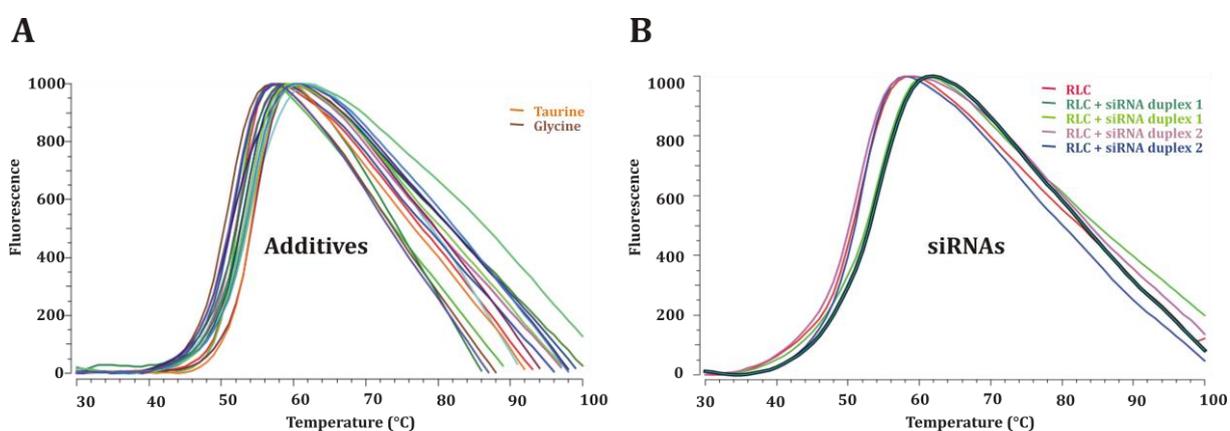
**A:** Different melting curves of the RLC using various buffers at different pH-values are shown. **B:** Different melting curves of the RLC upon addition of various additives during thermal shift assay are depicted. **C:** Melting curves of the RLC in NaCl buffers containing 25 to 500 mM NaCl. The highest melting temperature can be detected in 150 mM NaCl. **D:** The RLC is stable in KCl buffers containing 25 to 500 mM KCl, whereby the highest melting temperature can be detected in 150 mM KCl.

Fluorescence-based thermal shift assays using the additive screen (Hampton, USA) or manual screens, testing the influence of different additives and salt concentrations on the RLC (in HEPES buffer) stability, were performed (Figure 16 B to D). These assays identified CsCl, Glucose and Trehalose as stabilizing agents for the RLC (Figure 16 B). A second assay was performed using the additive screen and the RLC prepared in Imidazole buffer. Here, the addition of substances like glycerol, D-galactose, MnCl<sub>2</sub>, CaCl<sub>2</sub>, Taurine, Glycine or PEG led to higher melting temperatures and thus stabilization of the RLC (Figure 17 A). Moreover the RLC is stable at salt concentrations varying from 25 to 500 mM NaCl or KCl (Figure 16 C and D). Finally, thermal shift assays of the RLC using additives like nucleotides, reduction agents and siRNAs were performed. Most strikingly, the addition of siRNA duplex1 increased the melting temperature of the RLC to about 53°C and therefore seems to influence RLC stability (see Figure

## Results

17 B). Other siRNAs, nucleotides and reducing agents did not show any influence on RLC stability (see Figure 17, some curves are not shown).

Taken together, different buffers and additives increase RLC stability. Slightly acidic buffers at pH between 6.6 and 7.0, sugars, 2<sup>+</sup>-cations and the siRNA duplex1 seem to stabilize the RLC. These results obtained during thermal shift assays were used to optimize the RLC purification procedure by varying the purification buffers and the data quality of the EM images by testing different GraFix gradient buffers (see 6.1.5.1). Here, especially the Imidazole buffer at pH 6.8 was used for the RLC SEC and subjected as GraFix buffer helped to optimize EM image quality (Figure 11, Figure 15, 6.1.5.1). The RLC-siRNA duplex1 complex was shown to be stable during SEC, subjected to crystallization trials and used for EM (6.1.4; 6.1.5).



**Figure 17: Thermal shift assays of the RLC-siRNA duplex1 complex.**

**A:** Melting curves of the RLC and the RLC incubated either with siRNA duplex1 or siRNA duplex2 (5.1.13).  
**B:** Different melting curves of the RLC-siRNA duplex1 complex using various additives are depicted.

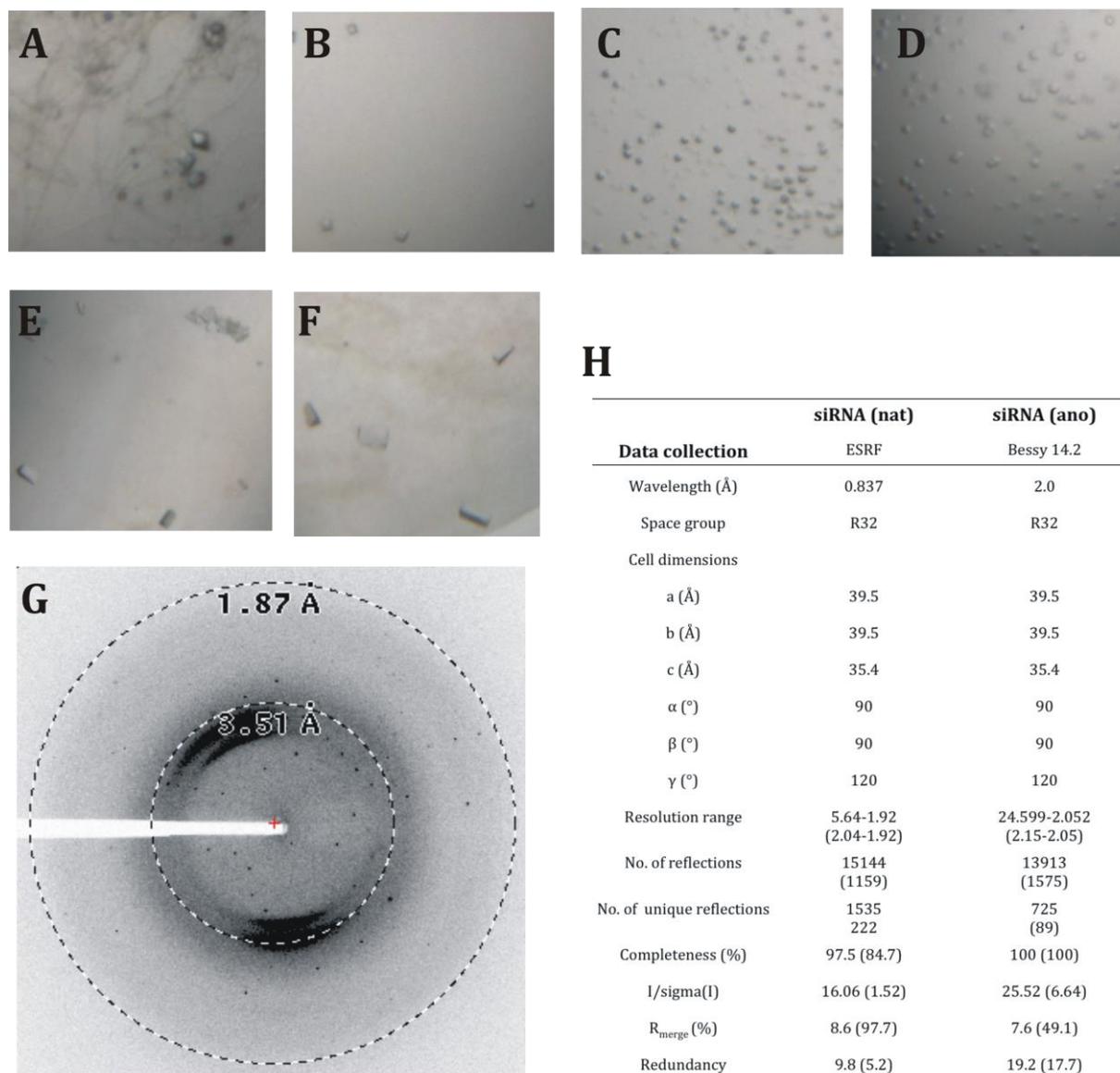
### 6.1.4 Crystallization trials of the RLC

One of the major goals was the 3D structure determination of the human RLC comprising hDicer, hAgo2 and hTRBP2. In the following section, attempts to crystallize the prepared RLCs, the RLC-siRNA duplex1 complexes, sub-complexes and single proteins are described (see 5.2.4.8.1 and Table 4). At first, the *in vitro* reconstituted RLC complex consisting of N-hDicer, N-hAgo2 and hTRBP2 was prepared as described in 6.1.1 and used for crystallizations trials at different concentrations and in varying buffers as listed in Table 4. 960 conditions were screened at 4 °C and 20 °C, but no crystals or pre-crystalline formations could be obtained. Additionally, the full length RLC consisting of hDicer, hAgo2 and hTRBP2 was used for similar crystallization trials, but this complex also did not form any crystals. The N-hDicer-hTRBP2-complex was also subjected to crystallization trials, but only salt crystals could be obtained (Figure 18 A). Addition of Chymotrypsin to this complex prior to crystallization trials led to several crystals (Figure 18 B). These crystals did not form in precipitation buffer and gel filtration buffer without the hDicer-hTRBP2 complex, but diffracting images clearly identified

these crystals as salt crystals. Using N-hAgo2 for crystallization trials, initial pre-crystalline precipitates could be obtained (Figure 18 C and D), but optimization attempts including seeding did not produce any protein crystals.

In contrast, diffracting crystals could be obtained by mixing 4 mg/ml hTRBP2 and siRNA duplex1 in a molar ratio of (1:1.1) with a reservoir containing 3.5 M Ammoniumsulfate and 0-10% MPD or 0.1 M citric acid at 20 °C (Figure 18 E and F). The crystals of around 50 x 50 x 10  $\mu\text{m}^3$  were grown in a reproducible fashion and flash frozen in reservoir solution with 10-15% glycerol for cryoprotection. Diffraction images revealed that the crystals, which diffracted to around 2 Å, only contained RNA, (Figure 18 G). Most likely, the high concentration of Ammoniumsulfate within the crystallization condition disrupted the hTRBP2-siRNA duplex1 complex, mediated siRNA duplex1 crystallization and hTRBP2 precipitation (brown precipitate in the background of Figure 18 F). However, several datasets of the RNA crystals were collected at ESRF (Grenoble). The data were processed using XDS and R32 was determined as space group with unit cell parameter of  $a=b= 39.5 \text{ \AA}$ ,  $c= 35.4 \text{ \AA}$  and  $\alpha=\beta= 90^\circ$ ,  $\gamma= 120^\circ$  (see Figure 18 H). Calculation of the Matthews coefficient (Matthews, 1968) revealed that the asymmetric unit volume of the unit cells is too small to harbor the siRNA duplex1 (21 kDa). Urea-PAGE of washed and dissolved crystals showed that the whole siRNA duplex1 is present in the crystals and SDS-PAGE analysis could not detect any protein (5.2.1.10; 5.2.3.2). The analysis of the data revealed crystallographic problems, including pseudotranslation and twinning. Pseudotranslation changes the statistical properties of intensities and might result in a wrong space group and can cause misleading results. Additionally, it is hard to find a perfect RNA model in the PDB for molecular replacement, because RNA is highly flexible (for review on RNA dynamics see Dethoff *et al.*, 2012). One attempt to solve the phase problem was to get anomalous signal of the phosphates in the siRNA backbone by collecting data at a wavelength of 2 Å (see Figure 18 H, siRNA (ano) data statistics), but the anomalous signal of the siRNA (ano) data was very weak. To date, it was not possible to solve the structure of the crystallized RNA.

## Results



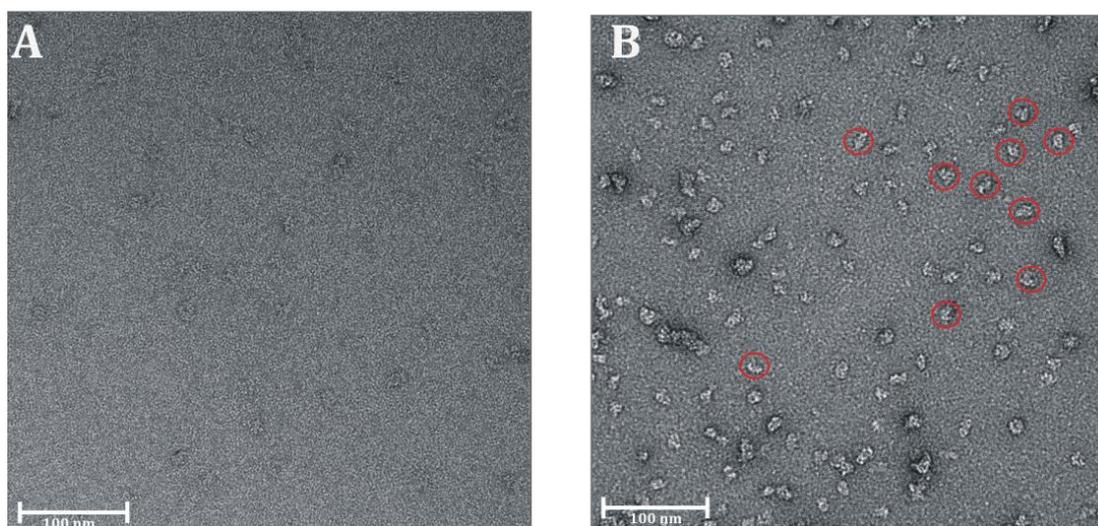
**Figure 18: Crystallization trials of the RLC-subcomplexes and individual proteins, diffraction image and data statistics of siRNA crystals.**

**A:** Salt crystals and fungal growths in a condition containing 1.7 mg/ml N-hDicer-hTRBP2 complex and 25% PEG 1000 and 0.1 M Na-HEPES pH 7.5 at 20 °C. **B:** Salt crystals in a condition containing 2 mg/ml N-hDicer-hTRBP2 complex incubated with Chymotrypsin in a molar ratio of 1000:1 and a reservoir containing 10% (w/v) PEG 4000; 0.1 M M-buffer 7.5; 20% (w/v) glycerol; 0.1 M M-NPS at 20 °C. **C:** Spherulites of N-hAgo2 obtained in 0.1 M Mg<sub>2</sub>Ac, 0.1 M MES pH 6.5 and 30% (w/v) MPD at 20°C. **D:** Spherulites obtained in a condition containing 3 mg/ml N-hAgo2, 0.1 M Mg<sub>2</sub>Cl, 0.1 M MES pH 6.5 and 30%PEG 400 at 20 °C. **E:** RNA crystals obtained by mixing 4 mg/ml hTRBP2 and siRNA duplex1 in a molar ratio of (1:1.1) with a reservoir of 3.2 M Ammoniumsulfate and 0.1 M citric acid at 20 °C. **F:** RNA crystals obtained by mixing 4 mg/ml hTRBP2 and siRNA duplex1 in a molar ratio of (1:1.1) with a reservoir of 3,01 M Ammoniumsulfate and 1.9% (w/v) MPD. **G:** Diffraction image of RNA crystals grown in conditions like E and F, prior to collection the crystals were cryoprotected in a solution containing reservoir and 10-15% glycerol. **H:** Data collection statistics of best diffracting crystals, siRNA (nat) means native dataset and siRNA (ano) an anomalous dataset.

### 6.1.5 EM-analysis of the RLC

Since no crystals could be obtained, to solve high resolution structures, EM analysis of the human RLC, the RLC-siRNA duplex1 complex, subcomplexes and Dicer were performed in collaboration with David Haselbach in Prof. Dr. Holger Stark's 3D electronmicroscopy group at the Max-Planck-Institute for Biophysical Chemistry (Göttingen).

The *in vitro* reconstituted RLC (see 6.1.1.4) was used to prepare negative-stain grids for electron microscopy (5.2.4.6.1). Initial images revealed a heterogeneity within the sample in form of particles of varied sizes and shapes (Figure 19 A). Especially, the contrast of particles compared to the background is really low (Figure 19 A). In order to overcome the heterogeneity of the sample and enhance the contrast of particles to background, the *in vitro* reconstituted RLC was applied to a GraFix gradient before preparing the EM grids (5.2.4.5). The Glutaraldehyde fixed complexes were negative-stained as before with uranyl formate and the particles analyzed by EM (5.2.4.6.1). The EM micrograph of the fixed RLC showed great improvement in image quality compared to the unfixated sample (compare Figure 19 A and B). Many particles of about the size of the RLC monomer (362 kDa) can be detected and the contrast of the single-particles to background is increased dramatically when compared to the unfixated complex particles (Figure 19 A and B). There are still particles of different sizes indicating a slightly to high concentration of complex or cross-linker during the GraFix preparation (RLC oligomers and some aggregates). When performing GraFix gradients the SEC step can be circumvented for RLC preparation. Thus, SEC was omitted for RLC preparations used for EM analysis.



**Figure 19: Effect of GraFix treatment of the recombinant RLC.**

**A:** EM micrographs of negative-stained particles of the RLC. Heterogeneous particles are present and the contrast of particle to background is low (low signal-to-noise ratio). **B:** EM micrographs of negative-stained particles of the RLC after GraFix treatment. There are still particles of different sizes (some aggregates), but the contrast of particles to background is increased (higher signal-to-noise ratio). Some single-particles of the RLC are marked with red circles.

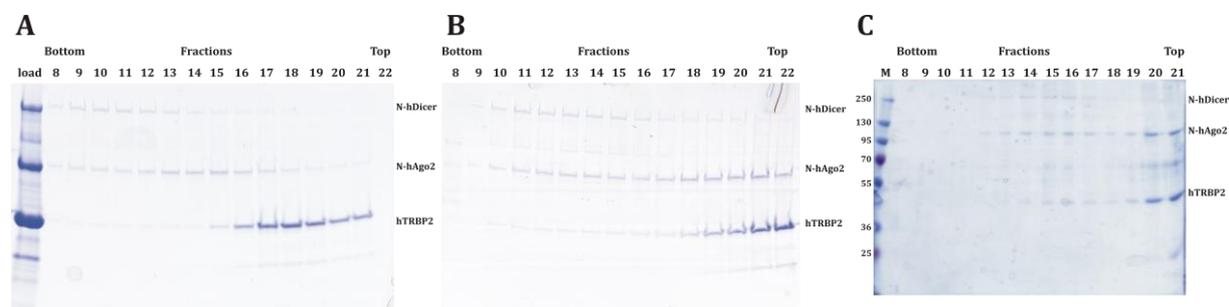
### **6.1.5.1 Optimization of EM image quality**

The heterogeneity with respect to size of the RLC particles can either stem from the sample itself or be due to GraFix and grid preparation. MALS analysis had revealed that the RLC itself is heterogeneous, thus forms different oligomers (and aggregates) and also might exist in different conformations due to its conformational flexibility (6.1.3.1). The heterogeneity in size was anticipated to be overcome by the GraFix procedure, whereby the gradient conditions (gradient buffers, gradient concentrations, centrifugation speed and time, centrifugation temperature), the fixation agent (fixation agent type and concentration), the fractionation of the gradients (fraction size, mixing of the gradients) and the stopping of the crosslinking reaction by different amino acids (glycine, aspartate) may have an influence. Within the following section several attempts to improve the EM image quality of the RLC are described.

Sample quality itself was optimized first. The purification buffers for the RLC preparations were varied based on the results observed from fluorescence-based thermal shift assays prior to EM analysis (6.1.3.2). Here, especially the change of buffers resulted in improvement and the addition of several cations or reducing agents did not show much improvement. One of the best negative-stain EM datasets of the RLC could be recorded, when reducing the pH from pH 8.0 to pH 7.0 during cell lysis and for the first affinity purification of the N-hDicer-N-hAgo2 complex and changing the gel filtration buffer from HEPES pH 7.5 to Imidazole pH 6.8 during SEC of the RLC. This SEC-purified RLC was applied to a 10-30% sucrose gradient containing 150 mM KCl, 20 mM Imidazole pH 6.8 and 2 mM DTT and 0.05% Glutaraldehyde as fixation agent and centrifuged at 38000 rpm for 18 hours at 4 °C (5.2.4.5). This setup did not work for all preparations and often resulted in samples that were still very heterogeneous in size. As mentioned above, the addition of siRNA duplex1 to the RLC resulted in increased stability of the complex during fluorescence-based thermal shift assays (6.1.3.2), therefore we aimed to solve the three dimensional structure of the RLC-siRNA duplex1 complex as well. By comparison of the two structures we aimed, to detect structural rearrangements upon siRNA duplex1 binding to the RLC and thus get insights into the siRNA-transfer mechanism within the RLC. Thus, for the following optimization steps the RLC or the RLC-siRNA duplex1 complexes were used.

To analyze the behavior of the RLC during gradient centrifugation, gradient centrifugation of the RLC and the RLC-siRNA duplex1 complexes without any fixation agent were carried out as before and the fractions analyzed by SDS-PAGE (see Figure 20). The SDS-PAGE analysis of the gradient fractions of the RLC shows that the RLC components and especially N-hDicer appears in all fractions of the gradient forming various oligomers (Figure 20). Generally fractions 13 to 16 containing mainly RLC monomers were used for EM analysis. Within these fractions a low amount of hTRBP2 can be detected compared to N-hDicer and N-hAgo2 and therefore it seems that hTRBP2 is not stoichiometrically incorporated into the RLC or disassembles from the RLC (Figure 20). One possible explanation is that RNA contaminations within the purification of the

N-hDicer-N-hAgo2 complex, the hTRBP2 protein and so the RLC preparation might prevent proper binding of hTRBP2 to the N-hDicer-N-hAgo2 complex. In order to exclude RNA impurities, additional Heparin-column purification step for N-hDicer-N-hAgo2 complex and hTRBP2 purifications were included and the RLC prepared this way used for gradient centrifugation. SDS-PAGE analysis of these gradient fractions revealed that hTRBP2 is not incorporated to a higher extent as with previous preparations. An alternate preparation strategy to ensure proper incorporation of hTRBP2 into the RLC was the addition of the N-hDicer-N-hAgo2 cell lysate to a GSH-Sepharose column with GST-TRBP2 immobilized to it. The eluted GST-hTRBP2-N-hDicer-N-hAgo2 complex was incubated with PreScission protease over night to cleave the N-terminal GST-tag from hTRBP2 and subsequently applied this solution to SEC and the appropriate fractions to GraFix preparation (5.2.4.5). Although, hTRBP2 seems to be properly incorporated into the RLC using this preparation strategy, the resulting EM micrographs showed very heterogeneous particles and did not improve data quality (data not shown). Based on the assumption that the inefficient incorporation of hTRBP2 is due to RNA impurities, RNase A was added to the *in vitro* reconstituted RLC prior to GraFix preparation. SDS-PAGE analysis showed that the addition of RNase A to the *in vitro* reconstituted RLC prior to GraFix preparation indeed resulted in a proper incorporation of hTRBP2 into the RLC (Figure 20 C).



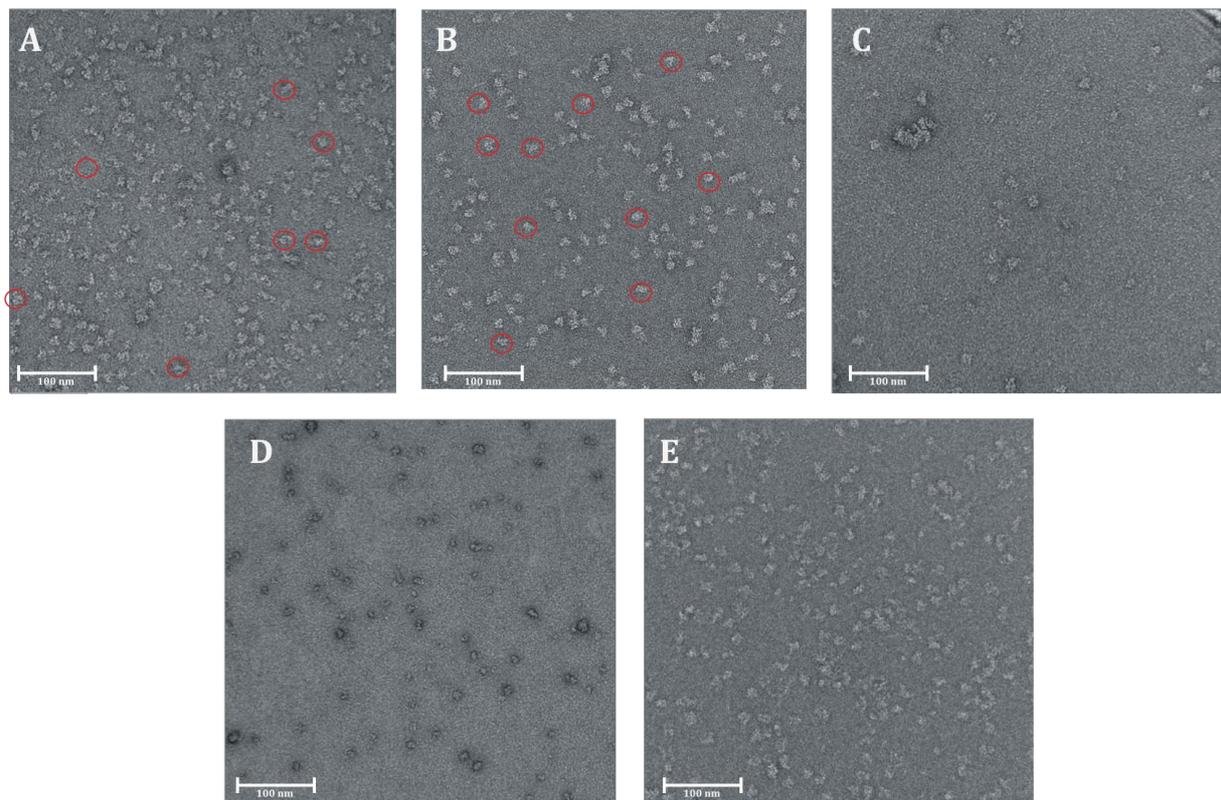
**Figure 20: SDS-PAGE analysis of RLC complex fractions after gradient centrifugation.**

**A and B:** SDS-PAGE analysis of fractions after gradient centrifugation of the RLC (A) and RLC-siRNA duplex1 complex (B). No stoichiometric incorporation of TRBP2 into the RLC (fractions 13-16). A different migration behavior of the RLC components and better incorporation of hTRBP2 into the RLC can be detected upon addition of the siRNA duplex1. **C:** SDS-PAGE analysis of fractions from gradient centrifugation of the RLC that was preincubated with RNase A prior to gradient centrifugation. A Stoichiometric incorporation of hTRBP2 into the RLC can be detected.

Fluorescence-based thermal shift assays of the RLC identified several sugars like trehalose, glucose, sucrose, galactose and the sugar alcohol glycerol to stabilize the complex. To improve image quality GraFix fixation of the RLC using 10-30% glycerol, sucrose, trehalose, galactose, glucose or arabinose were performed and the samples compared using negative-stain EM analysis (5.2.4.5; 5.2.4.6.1). The glycerol and sucrose gradients resulted in high quality single-particles and were used within the following preparations (EM micrographs not shown). High

## Results

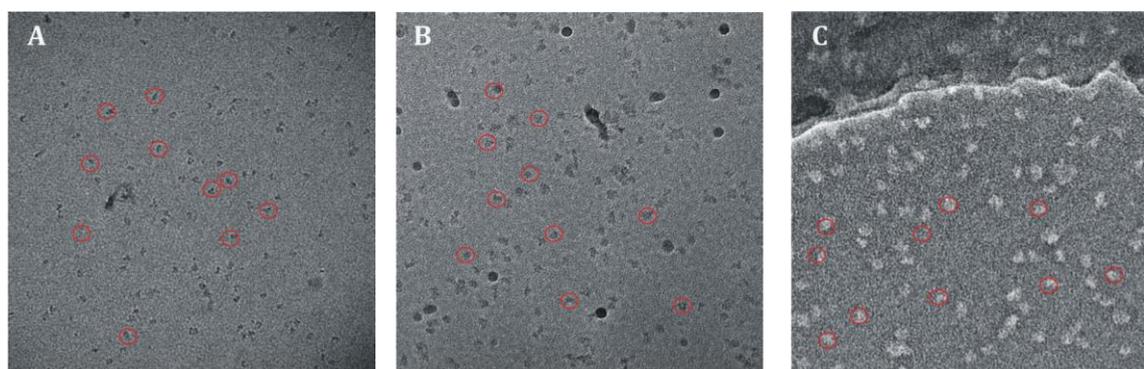
concentrations of Glutaraldehyde can lead to intermolecular fixation and thereby increasing sample heterogeneity, because intermolecular cross-linked complexes tend to form aggregates (Jensen, 2010). In order to test, which concentration of Glutaraldehyde is necessary and sufficient for crosslinking of the RLC, concentrations ranging from 0.1 to 0.001% were applied during GraFix preparation and the gradient fractions analyzed on native SDS-PAGE. A concentration of 0.01% Glutaraldehyde was found to be adequate to crosslink the RLC and was therefor used for the subsequent GraFix preparations (data not shown). Different available fixating agents like Glutaraldehyde (Figure 21 A), PMPI (Figure 21 B), SMCC (Figure 21 C), DFDNB (Figure 21 D) and EDC (Figure 21 E) were tested during GraFix preparation and the resulted samples analyzed by negative-stain EM (Figure 21 A-E). Glutaraldehyde and PMPI resulted in mainly monodisperse RLC samples, SMCC and DFDNB led to aggregation of the RLC and EDC did not crosslink the RLC components well or even caused disassembly of the RLC (see Figure 21 A-E). Taken together Glutaraldehyde and PMPI are the best fixation agents for the RLC and in the following only Glutaraldehyde and PMPI or a combination of these two was used during GraFix preparations.



**Figure 21: Effect of different cross-linker during GraFix of the recombinant RLC.**

EM micrographs of negative-stained particles of the RLC after GraFix treatment, whereby Glutaraldehyde (A), PMPI (B), SMCC (C), DFDNB (D) and EDC (E) were used as cross-linkers. Using Glutaraldehyde and PMPI many homogeneous particles are present, upon addition of SMCC and DFDNB RLCs aggregates and EDC does not crosslink the RLC. Some single-particles of the RLC are marked with red circles.

The heavy metal salt grains used during negative-stain EM analysis limit the resolution of a complex to about 20 Å. To obtain a higher resolution reconstruction of the RLC cryo electron microscopy (cryo-EM) is needed. Therefore, the sample has to be free of any cryo-protectant like glycerol or sucrose and thus after GraFix preparation the fixated complexes were applied to a gel filtration column (5.2.4.6.2). Using the GraFix protocol sometimes negative-stain images of relatively uniform particles could be recorded, but during buffer exchange for single-particle cryo-EM the RLC aggregated in most cases. This was recognized by comparing negative-stain images of preparations before and after this buffer exchange. In most cases no cryo-EM data could be recorded. However, when using a 10-30% glycerol or sucrose gradient containing 250 mM KCl, 20 mM HEPES/KOH pH 7.5, 5 mM MgCl<sub>2</sub> 5 mM DTT and 0.01% Glutaraldehyde and PMPI as fixation agents, the RLC-siRNA duplex1 complex stayed intact after buffer exchange and cryo images showed homogeneous particles (Figure 22) that were used for a 3D reconstruction (6.1.5.2). As these conditions did not work for all preparations other setups were tried to obtain better images and a high-resolution structure. Once the GraFix preparation of the RLC was carried out at -10 °C (-6.6 °C) instead of 4 °C. This variation resulted in a complex that was stable during buffer exchange and good cryo-EM images could be recorded (Figure 22 B). Because the RLC often aggregated during buffer exchange for cryo-EM, cryo negative-stain grids were prepared, whereby no buffer exchange is necessary (5.2.4.6.2, Figure 22 C). In addition to the formation of several oligomers in the RLC solution, degradation of hDicer could be detected during preparations (6.2.1; Figure 47). This phenomenon could be prevented by addition of EDTA during GraFix preparation and increased sample homogeneity, lead to reduced aggregation during buffer exchange and thus cryo-grids could be prepared and data recorded (images not shown; 5.2.4.6.2).



**Figure 22: Cryo images of the recombinant RLC.**

**A:** Cryo-EM images of the RLC siRNA duplex1 complex using 10-30 % glycerol gradient containing 250 mM KCl, 20 mM HEPES (KOH) pH 7.5, 5 mM MgCl<sub>2</sub> 5 mM DTT and Glutaraldehyde and PMPI as fixation agents at 4°C. **B:** Cryo-EM images of the RLC using GraFix preparation with Glutaraldehyde at -6.6 °C. **C:** Cryo negative-stain images of the RLC. Some single-particles of the RLC are marked with red circles.

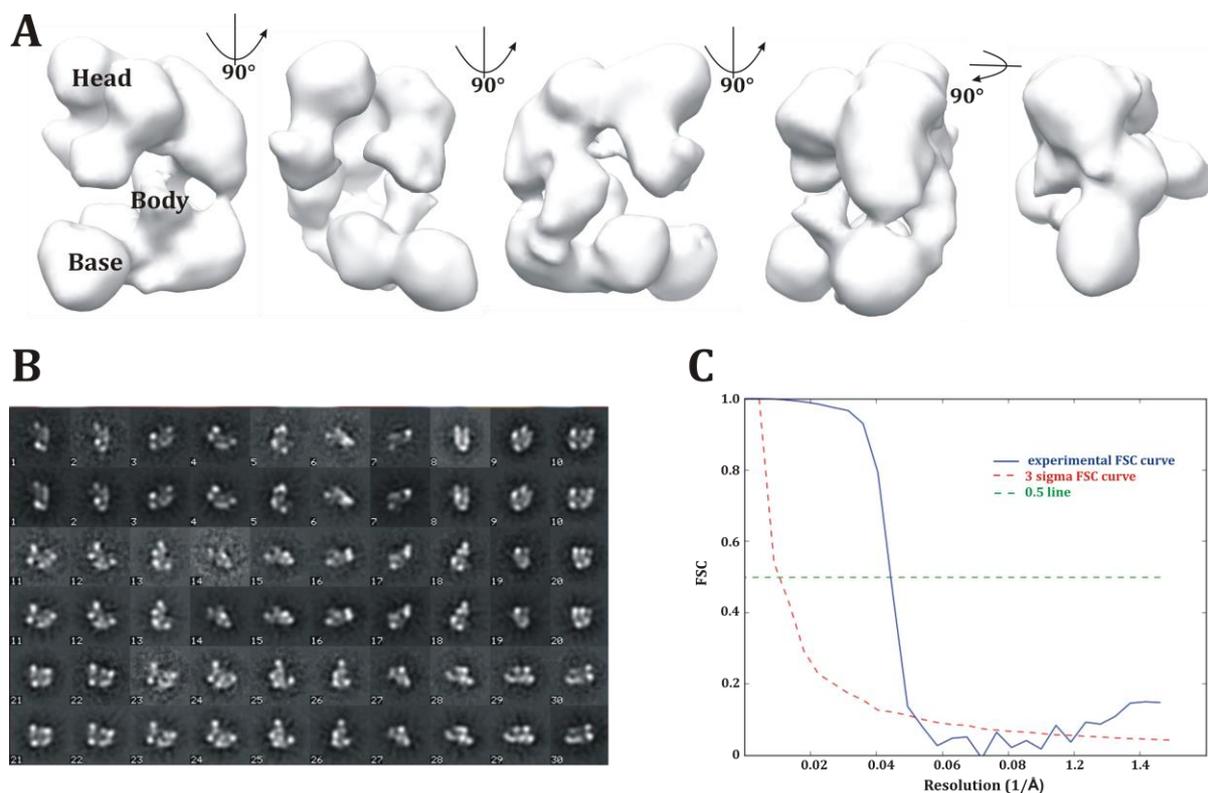
## Results

In summary, heterogeneity and aggregation of the RLC cannot be completely reduced during EM analysis using different GraFix setups (as described above). The obtained optimization results are often inconsistent and were sometimes not reproducible, which is most likely due to a different sample quality after purification. The sample quality is most likely depending on the varying expression rates of N-hDicer and N-hAgo2 using insect cells (5.2.2.6.7). Generally, best EM images could be obtained, when a RLC preparation was applied to 10-30% glycerol or sucrose gradients in a buffer containing 150-250 mM KCl, 20 mM HEPES (KOH) 7.0-7.5 or Imidazole pH 6.8, 0-5 mM MgCl<sub>2</sub>, 0-5 mM EDTA and 0-5 mM DTT fixed with 0.01% Glutaraldehyde or PMPI at 4 °C and -10 °C. During data processing, the remaining heterogeneity in size could be overcome by picking only particles of a size that correspond to RLC monomers and a possible heterogeneity of the complex itself (different shapes) is analyzed by computational methods used during EM analysis (see Figure 24).

### **6.1.5.2 Preliminary EM structure of the RLC**

Using a cryo EM dataset of the RLC-siRNA duplex1 complex, single-particles were picked manually, aligned and sorted into several class sums and an initial three-dimensional reconstruction was built by an *ab initio* reconstruction method by David Haselbach and Prof. Holger Stark. After several rounds of refinement a preliminary 3D reconstruction of the RLC could be obtained (see Figure 23 A). The head, body and base of the reconstruction are indicated in Figure 23. The RLC adopts a C-shape with an elongation on the head facing to the base (see Figure 23 A).

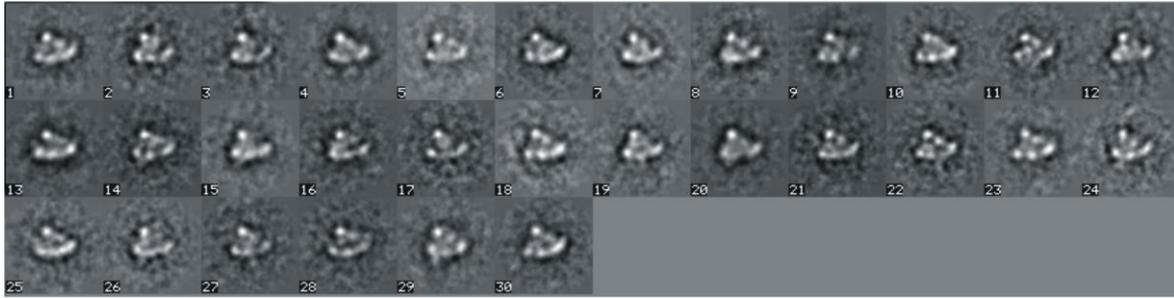
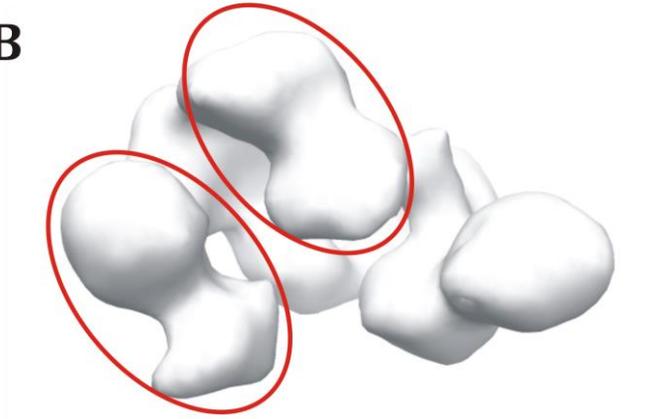
To determine the resolution of the reconstruction the Fourier Shell Correlation (FSC) is used. Therefore the data of the reconstruction are split into two sets and two independent reconstruction are calculated. The correlation between Fourier transforms of the two reconstructions in Fourier space was calculated and together with the 3-sigma ( $\sigma$ ) threshold curve plotted against the resolution (Figure 23 C). The resolution of the RLC was determined as 22.5 Å using the 3- $\sigma$  threshold criteria (Figure 23). Moreover the RLC map was back projected into 2D-images with 30 degree spacing and these projections were compared with the class averages. In Figure 23 the projections of the reconstituted RLC (raw one, three and five) are sorted to the corresponding class averages (raw two, four and six). The class averages and projections are identical, which they have to it for a correct 3D calculation. Also this is not a final proof for the correctness of the calculation.



**Figure 23: EM structure of human RLC.**

**A:** 3D reconstruction of human RLC shown in different orientations. **B:** Comparison of projections of the reconstituted RLC (first, third and fifth row) with corresponding class sums (second, fourth and sixth row). **C:** Fourier Shell Correlation (FSC) plotted against resolution for the reconstruction of the RLC (blue line). The threshold at 0.5 (green dotted line) and the 3-sigma ( $\sigma$ ) level (dotted line) are shown.

When comparing 2D-images that correspond to the RLC-siRNA complex reconstruction front, different arrangements of the arms can be detected (Figure 24). It is still an open question if this is truly flexibility within the RLC (probably caused by the siRNA duplex) or a heterogeneity artefact. To address this question extensive computational work is necessary, that has not been finished so far.

**A****B**

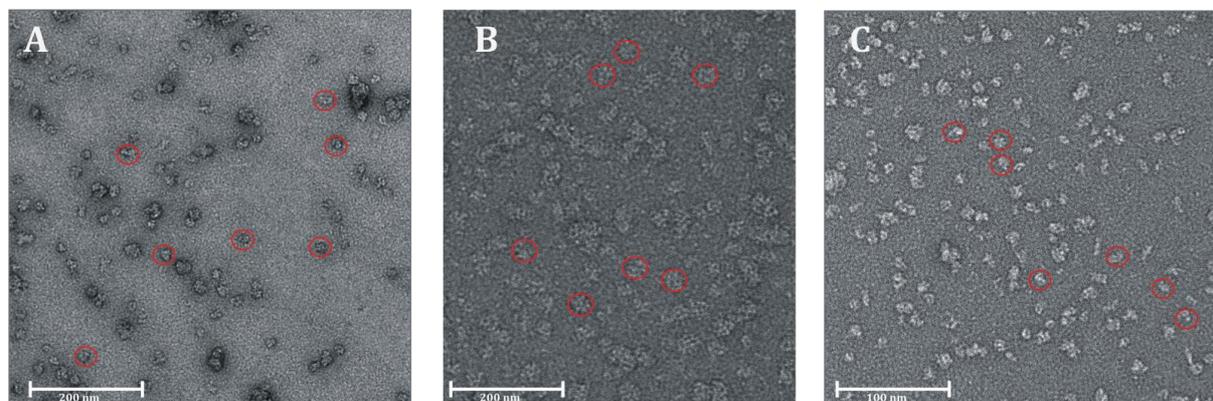
**Figure 24: Heterogeneity analysis of RLC.**

**A:** Different class sums of the front view of the RLC reveals flexible regions in the middle of the molecule. **B:** The flexible regions identified by comparing different alignments are marked by red rings in the RLC reconstruction.

### **6.1.5.3 EM analysis of RLC subcomplexes and Dicer**

To map the individual RLC components within the obtained EM structure the aim was to calculate reference density maps of the subcomplexes and hDicer alone. The complexes and Dicer (with and without his-tag) were prepared as described earlier and applied to GraFix preparations prior to EM analysis (6.1.1.3, 6.1.1.4, 6.2.1, 5.2.4.5). The EM micrographs of negative-stained particles of the N-hDicer-N-hAgo2 complex, the N-hDicer-hTRBP2 complex and N-hDicer are depicted in Figure 25. These EM micrographs revealed a high degree of heterogeneity of the samples. The heterogeneity is even higher than for the trimeric RLC (for comparison see Figure 19 and Figure 21). When analyzing micrographs of the RLC, subcomplexes and N-hDicer an increase in homogeneity upon addition of complex partners was detected. Thus, the heterogeneity of RLC is lower (the complex is more stable), than the heterogeneity of N-hDicer-N-hAgo2 or N-hDicer-hTRBP2 complexes and heterogeneity of the subcomplexes is lower than the heterogeneity of N-hDicer. We suppose that hAgo2 or hTRBP2 stabilizes hDicer within the subcomplexes and that the trimeric RLC is in an even more stable state. As such an increase in stability and homogeneity upon addition of protein binding partners to Dicer can be detected, we imagine that the trimeric RLC misses an additional

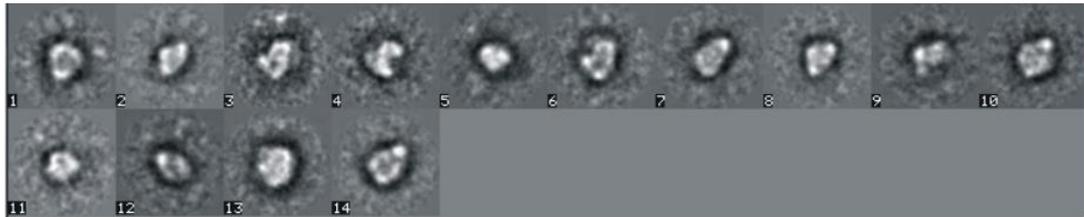
component for proper stabilization/homogeneity. To test this hypothesis, additional known RLC interaction partners like hPACT1 and Hsp90 were added to the complex prior to EM analysis. Addition of these two protein components did not remarkably improve RLC EM image quality and were not further investigated.



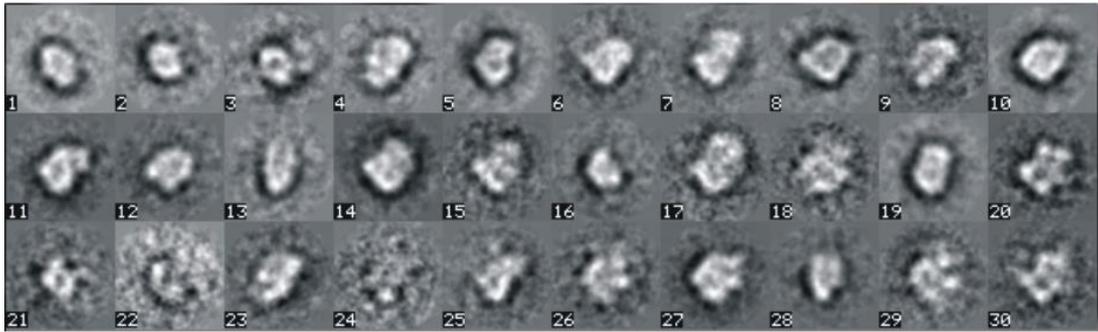
**Figure 25: EM micrographs of negative-stained particles of N-hDicer (A), the N-hDicer-N-hAgo2 complex (B) and N-hDicer-hTRBP2 complex (C).** The EM micrographs showed high degrees of heterogeneity within the samples.

However, particles of N-hDicer, N-hDicer-N-hAgo2 complex and the N-hDicer-hTRBP2 complex were picked, aligned and class sums could be calculated (Figure 26). The 2D-images look similar to the 2D-images obtained for the RLC. To date reconstructions of credible 3D-maps of hDicer or the subcomplexes has not been successful.

**A**



**B**



**C**

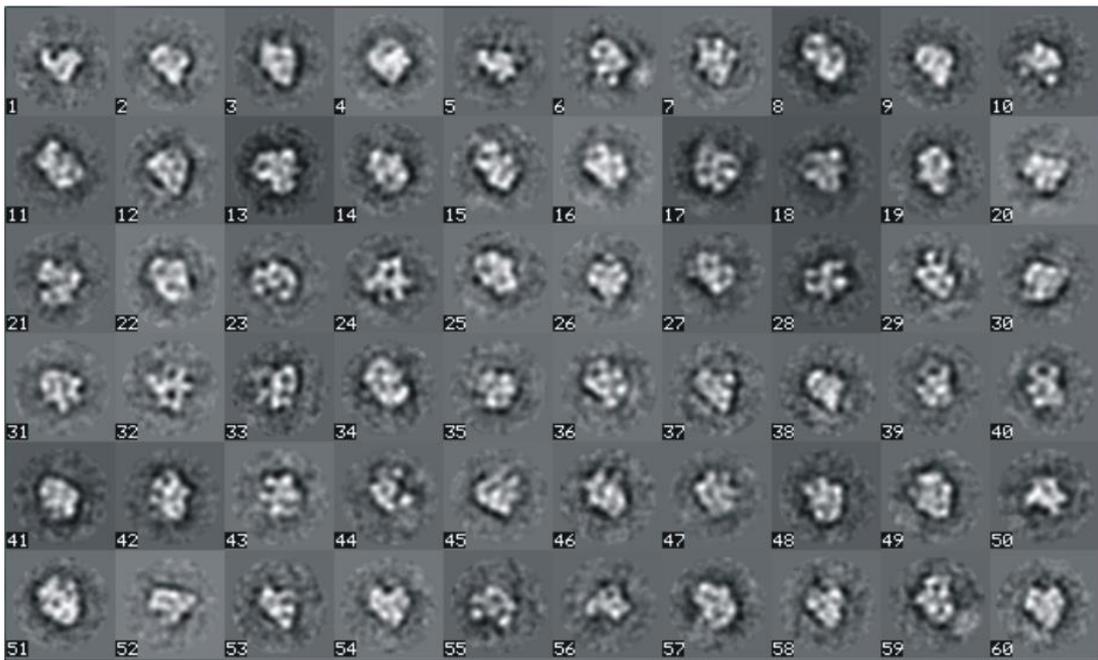


Figure 26: Class sums of N-hDicer (A), the N-hDicer-N-hAgo2 complex (B) and N-hDicer-hTRBP2 complex (C).

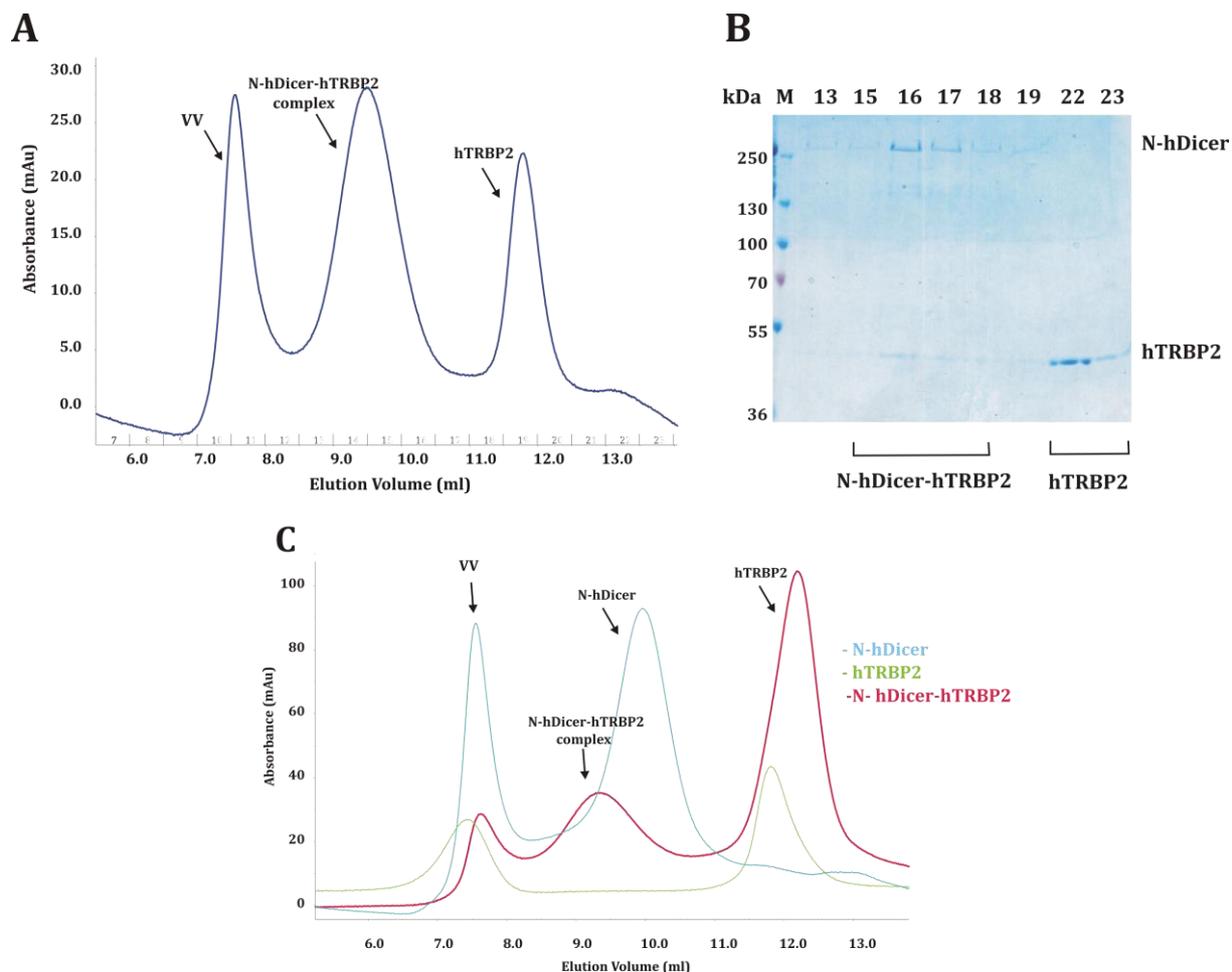
## **6.2 Dicer-TRBP2 complexes**

In order to gain insights into the structure and function of the human Dicer-TRBP2 complex, a recombinant human Dicer-TRBP2 complex was prepared for crystallization trials and EM analysis (see previous sections 6.1.4; 6.1.5.3). In the second part of this section, the preparation and crystallization trials of minimal human Dicer-TRBP2 complexes are outlined (6.2.2; 6.2.4).

### **6.2.1 Purification of full length human Dicer-TRBP2 complex**

Initially the reconstruction of the full length human Dicer-TRBP2 complex was aimed by assembling independently expressed and purified human Dicer and TRBP2 proteins. Using individually expressed and purified proteins, the human Dicer-TRBP2 complex could be reconstituted *in vitro* by incubating N-hDicer with a 2-fold excess of hTRBP2 and applying that protein solution onto a SEC to separate the complex from aggregated protein (VV) and unincorporated hTRBP2 (Figure 27 A). Peak fractions of the SEC run were analyzed by 10% SDS-PAGE indicated that N-hDicer and hTRBP2 are present in the second elution peak (Figure 27 B). Upon addition of hTRBP2 N-hDicer elutes at a 9.6 ml and thus earlier than N-hDicer alone (10 ml). N-hDicer has an apparent molecular weight of 251 kDa (6.1.1.3) and the N-hDicer-hTRBP2 complex elution profile corresponds to a molecular weight of 371 kDa. Thus an N-hDicer-hTRBP2 complex can be stably reconstituted *in vitro* (Figure 27 C).

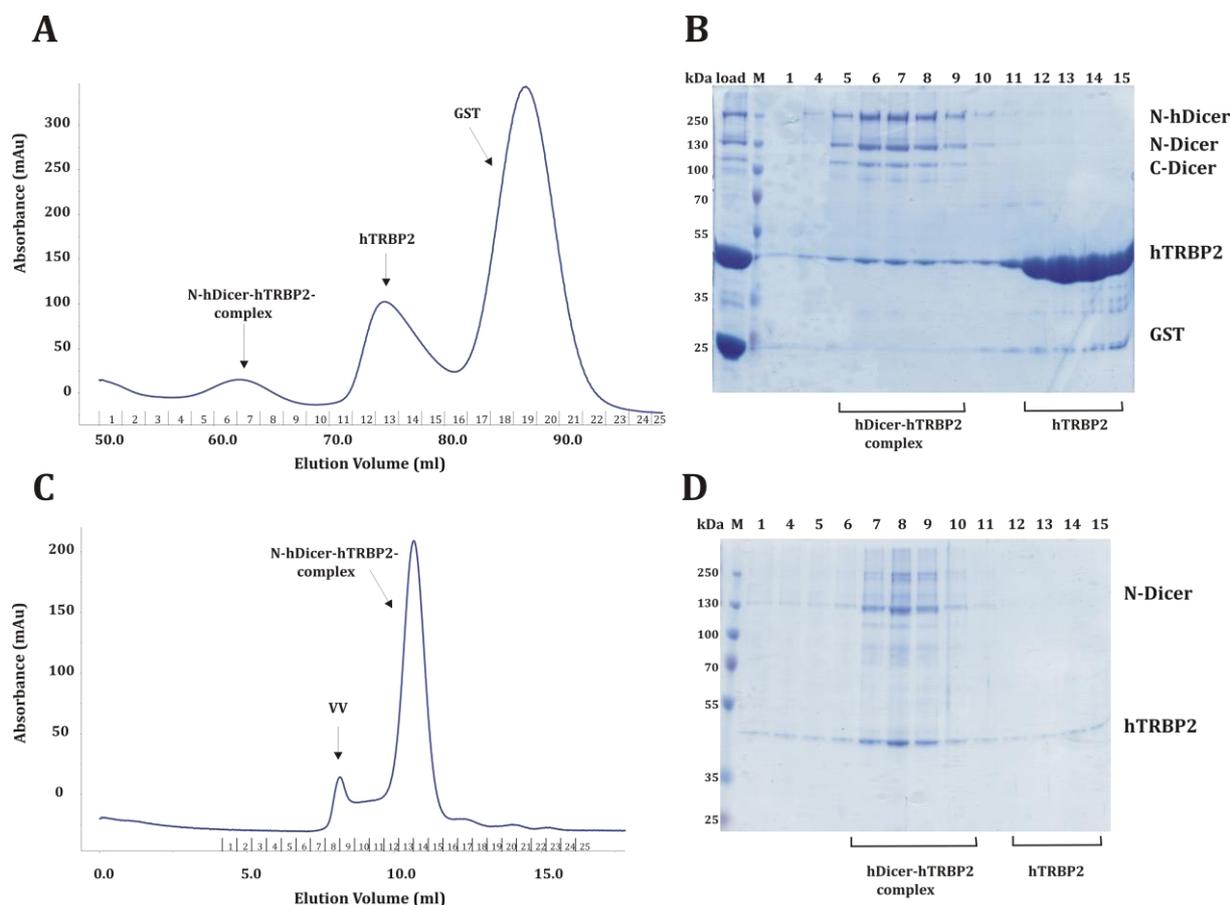
## Results



**Figure 27: *In vitro* reconstitution of human Dicer-TRBP2 complex.**

**A:** Superose 12 (10/300) elution profile of N-hDicer-hTRBP2 complex. Affinity-purified N-hDicer was incubated with purified hTRBP2 in a molar ratio of 1:2 and applied onto a Superose 12 (10/300) column in buffer containing 150 mM KCl, 20 mM HEPES pH 7.5, 5% (v/v) glycerol and 2 mM DTT. The elution profile monitored by absorption at 280 nm is shown. **B:** Analysis of the peak fractions of the SEC of full length N-hDicer-hTRBP2 complex by 10% SDS-PAGE with subsequent Coomassie staining. The identified proteins are indicated on the right and size marker is shown on the left of the gel. The distribution of N-hDicer-hTRBP2 complex (hDicer-hTRBP2) and hTRBP2 are denoted below the gel. **C:** SEC elution profiles of N-hDicer-hTRBP2 complex, N-hDicer and hTRBP2. Different amounts of N-hDicer (light blue) hTRBP2 (light green) and the N-hDicer-hTRBP2 complex were analyzed by SEC.

The hTRBP2 protein could be purified to >95% purity as described in 6.1.1.2, however the monodispersity of hDicer protein could not be reproduced in every preparation (6.1.1.3). To improve the preparation of full length N-hDicer-hTRBP2 complex, a co-purification strategy as described in 5.2.3.6 was carried out. A representative chromatogram of a SEC run and the corresponding SDS-PAGE analysis of the N-hDicer-hTRBP2 complex are shown in Figure 28. Excess of hTRBP2 and GST can be separated from the complex. The N-hDicer-hTRBP2 complex elutes in one peak at 62 ml corresponding to a molecular weight of 252 kDa, which is about the size of the sum the individual proteins (263 kDa). Apart from the full length N-hDicer and hTRBP2, additional protein bands can be identified on the SDS-gel (Figure 28 B).



**Figure 28: Preparation of human Dicer-TRBP2 complex.**

**A** and **B**: Superdex 200 (16/60) elution profile of the N-hDicer-hTRBP2 complex (hDicer-hTRBP2 complex) and subsequent 10% SDS-PAGE analysis of the peak fractions. Besides full length N-hDicer and hTRBP2 proteins, additional bands can be identified. **C** and **D**: Superdex 200 (10/300) elution profile of the Dicer-TRBP2 complex and subsequent 10% SDS-PAGE analysis of the peak fractions.

Western blot and mass spectrometry analysis identified these two additional bands as fragments of hDicer (see Figure 47). The upper band (140 kDa) is the N-terminal fragment ranging from the N-terminus to the PAZ domain of human Dicer (N-Dicer), whereas the ca. 100 kDa band is the C-terminal fragment containing the RNase III domains and the C-terminal dsRBD of human Dicer (C-Dicer). These two bands can be detected in other preparations of N-hDicer. Thus, in addition to the full length Dicer protein, cleaved fragments of hDicer coelute on SEC (Figure 28).

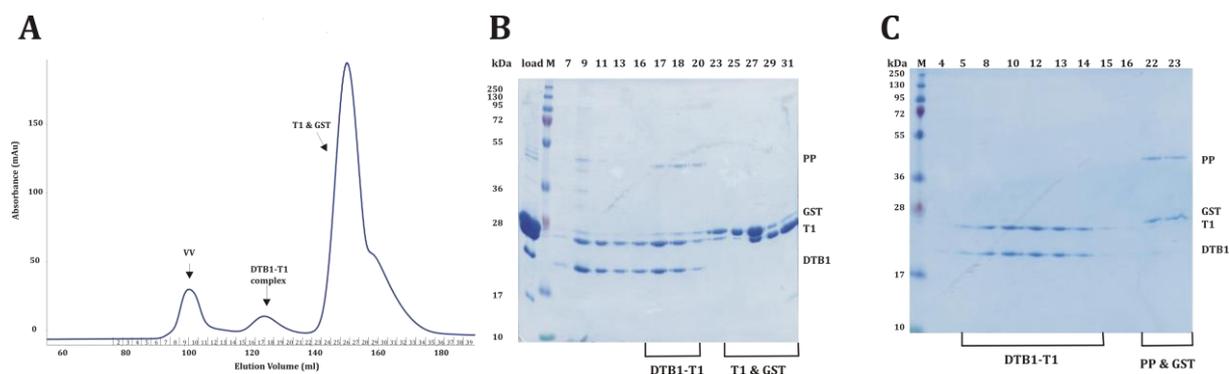
To analyze the change in amounts of degradation over time, a sample of the concentrated N-hDicer-hTRBP2 complex (from Figure 28) was left on ice for 2 days and analyzed by analytical Superdex 200 (10/300) column and subsequent SDS-PAGE analysis (Figure 28 C and D). The analysis showed that N-hDicer is degraded, but the 140 kDa the N-terminal (Helicase) fragment stayed intact. This suggests that the N-terminal part of Dicer is stabilized through hTRBP2 interaction and is therefore not degraded. The prepared N-hDicer-hTRBP2 was subjected to EM

## Results

analysis and crystallization trials directly after affinity purification or SEC (less degradation products of Dicer, see 6.1.4 and 6.1.5.3).

### **6.2.2 Preparation of minimal human Dicer-TRBP2 complexes**

It has been known for several years that the DExD-box helicase domain of human Dicer interacts with human TRBP2 (Laraki *et al.*, 2008). Not much was known about the Dicer helicase domain at the beginning of this study and in order to understand the function of the Dicer DExD-box helicase domain and its interaction with hTRBP2, a DExD-box helicase domain hDicer-hTRBP2 complex should be prepared. The preparation of a DExD-box helicase domain hDicer-hTRBP2 complex turned out to be only possible by co-expression of the DExD-box helicase domain fragment with hTRBP2. The yield of pure heterodimeric DExD-box helicase domain hDicer-hTRBP2 complex was about 0.1 mg per 5 Liter *E. coli* expression culture and therefore not suitable for crystallization trials (Diploma thesis Franka Voigt). When expressed in *E. coli* cells, the DExD-box helicase domain of hDicer aggregates and might be not properly folded as it elutes within the void volume during SEC. A co-purification of the heterodimeric DExD-box helicase domain hDicer-hTRBP2 complex was not possible (data not shown). In 2009, Daniels *et al.* were able to map the hDicer-hTRBP2 interactions sites more precisely. Residues 298-366 of hTRBP2 were identified to be responsible for hDicer binding. The hTRBP2-binding site comprises amino acids 267-431 and is located between the ATPase and helicase domains of hDicer (Larkai *et al.*, 2008; Daniels *et al.*, 2009). In order to crystallize a minimal hDicer-hTRBP2 complex, different hDicer and hTRBP2 fragments were cloned (5.1.12, 5.2.1.1). hDicer comprising amino acids 261-432 (called DTB1) or 261-439 (called DTB2) and hTRBP2 comprising aa 157-366 (called T1) or 291-366 (called T3) were expressed as a GST-fusion protein in BL21 (DE3) Star cells (5.2.2.5). For preparation of a minimal hDicer-hTRBP2 complex, the individually expressed GST-fusion protein containing *E. coli* cells were lysed and affinity purified together as described in 5.2.3.4. A typical Superdex 75 (26/60) elution profile of the DTB1-T1 complex and the corresponding SDS-PAGE analysis are shown in Figure 29 A and B. Aggregated protein, unincorporated T1 protein and most GST could be separated from the DTB1-T1 complex, but the DTB1-T1 complex still contains some GST and PreScission protease impurities (see Figure 29 B) and was therefore applied to a final GSH-Sepharose column (5.2.3.4). An SDS-PAGE analysis of such a run is depicted in Figure 29 C. The fractions of the flow-through (fractions 5-15) contained the pure DTB1-T1 complex, the fractions of the elution contained GST and PreScission protease (fractions 22 and 23) and thus the impurities could be separated from the DTB1-T1 complex. The DTB1-T3 and DTB2-T1 complexes were purified accordingly (6.2.2).



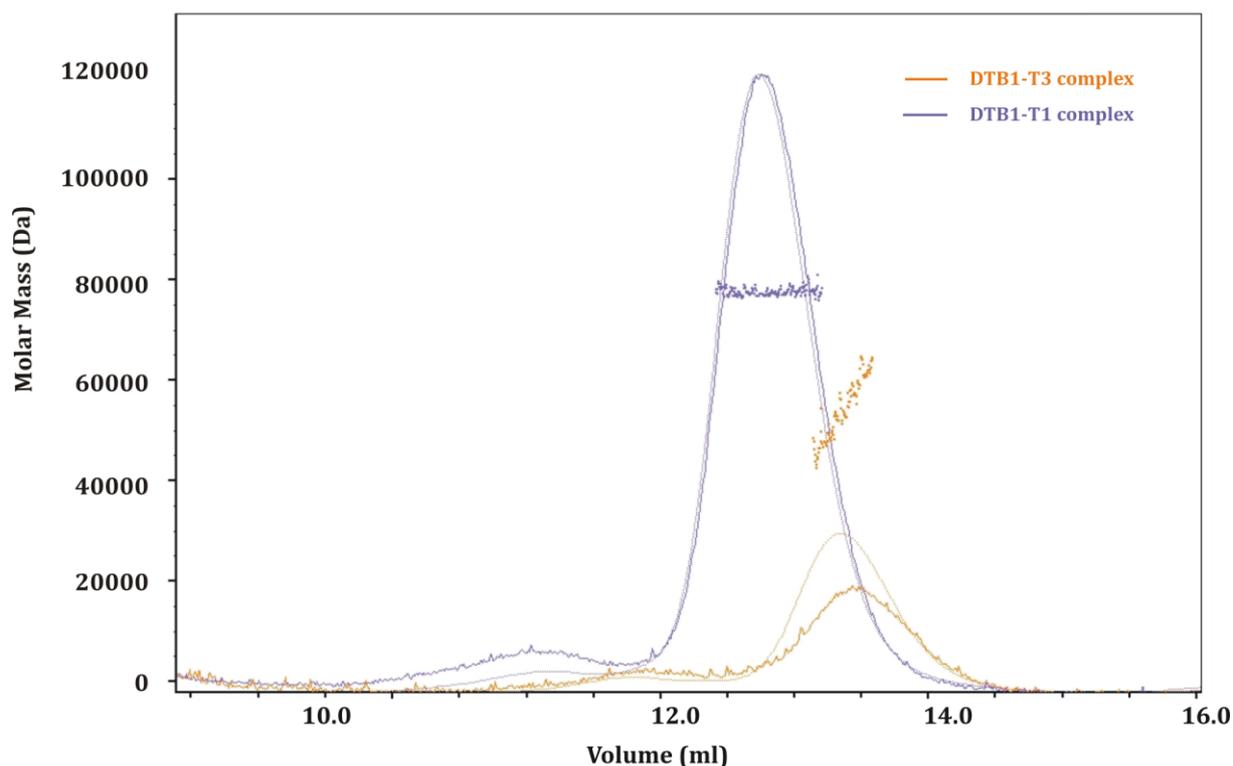
**Figure 29: Purification of DTB1-T1 complex.**

**A:** Superdex 75 (26/60) elution profile of the DTB1-T1-complex. **B:** 15% SDS-PAGE analysis of the peak fractions of the Superdex 75 (26/60) runs of the DTB1-T1-complex. **C:** 15% SDS-PAGE analysis of the peak fractions of the second GSH-Sepharose of the DTB1-T1-complex. The fractions of the flow-through (fractions 5-15) contained the pure DTB1-T1 complex, the fractions of the elution contained GST and PreScission protease (PP; M = marker, VV = void volume).

The purified DTB1-T1 complex was concentrated and subsequently analyzed by SEC combined with MALS for homogeneity and the molecular mass determination. This experiment revealed that the DTB1-T1 complex is homogeneous (Figure 30). Using the theoretical and the natural extinction coefficient molar masses of 86.2 kDa and 77.7 kDa, respectively, could be determined. These results correspond to twice the sum of the molecular masses of the individual proteins ( $MW_{DTB1} = 20.5$  kDa and  $MW_{T1} = 22.8$  kDa) meaning that the DTB1-T1 complex exists as heterotetramer in solution. A minimal amount of DTB1-T1 complex is aggregated and forms higher oligomers. As this complex is homogeneous it was used for crystallization trials (6.2.4). In contrast, the DTB1-T3 complex is polydisperse with an average molecular weight of 57.3 kDa and a broad size distribution of particles ranging from 39 kDa to 65 kDa. Hence, the DTB1-T3 proteins form heterotetramers ( $MW = 56$  kDa) and different subcomplexes.

To investigate the human Dicer interaction domain in more detail, human Dicer fragments containing amino acids 261-414 (DTB3), amino acids 261-415 (DTB4) or amino acids 264-414 (DTB5) were cloned as GST-fusion proteins and expressed in BL21 (DE3) Star *E. coli* cells (5.1.12; 5.2.1.1; 5.2.2.5). Using the co-purification strategy described above, preparation of DTB3, DTB4 or DTB5-T3 complexes was tried. Interestingly, no complex could be formed, thus the amino acids 264-414 seem to be important for human Dicer interaction with hTRBP2 *in vitro* (Figure 51).

## Results



**Figure 30: MALS of the DTB1-T1 and DTB1-T3 complexes.**

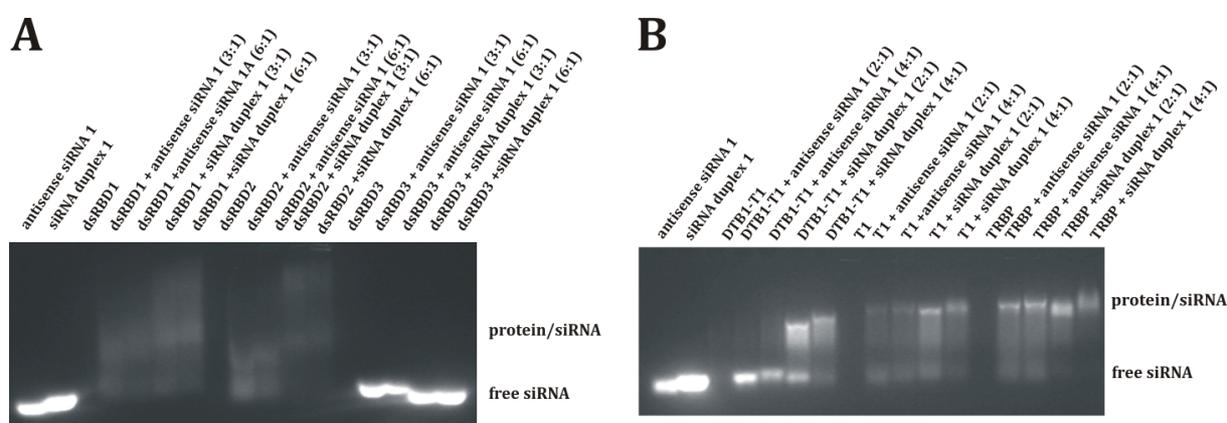
The absorption (continuous line) and the light scattering (thin line) of the DTB1-T1 (purple) and the DTB1-T3 (orange) complexes are plotted against the elution volume in ml. The molar mass distribution of the complexes is shown as purple and orange dots for the DTB1-T1 and DTB1-T3 complex, respectively.

### ***6.2.3 siRNA-binding properties of DTB1-T1 complex compared to hTRBP2 and hTRBP2 fragments***

The siRNA-binding abilities of the minimal human Dicer-TRBP2 complexes, TRBP2 and TRBP2 domains were analyzed by EMSAs and SEC (5.2.1.9, 5.2.3.7). Human TRBP2 contains three domains, two dsRBDs for RNA-binding and a Medipal domain for protein-protein interaction (4.2.3). EMSAs showed that hTRBP2s dsRBD1 and dsRBD2 are able to bind antisense siRNA and siRNA duplexes, whereby the Medipal domain (dsRBD3) is not able to bind RNA (Figure 31 A). Different oligomers seem to be bound upon addition of RNA and the affinity to duplexes seems to be higher than the affinity to single stranded RNA (nearly all free dsRNAs are bound). The dsRBD2 seems to have a higher affinity to the RNA (more free RNA bound, less smearing) than the dsRBD1. Combination of the two dsRBD1 and dsRBD2 in full length hTRBP2 protein results in affinity comparable to dsRBD2 (Figure 31 A and B). Here again, the affinity to siRNA duplexes seem to be higher than to single stranded siRNA (all free dsRNA incorporated). The DTB1-T1 complex containing dsRBD2 and dsRBD3 of hTRBP2 is expected to bind single stranded and siRNA duplexes, but the DTB1-T1 complex can only incorporate siRNA duplexes (Figure 31 A). In contrast, the free T1 fragment is able to bind single stranded and double

stranded siRNA (Figure 31 B). Hence, the DTB1 seem to prevent binding of single stranded siRNA by T1 and the hTRBP2-hDicer-interaction could influence TRBP2 siRNA selection.

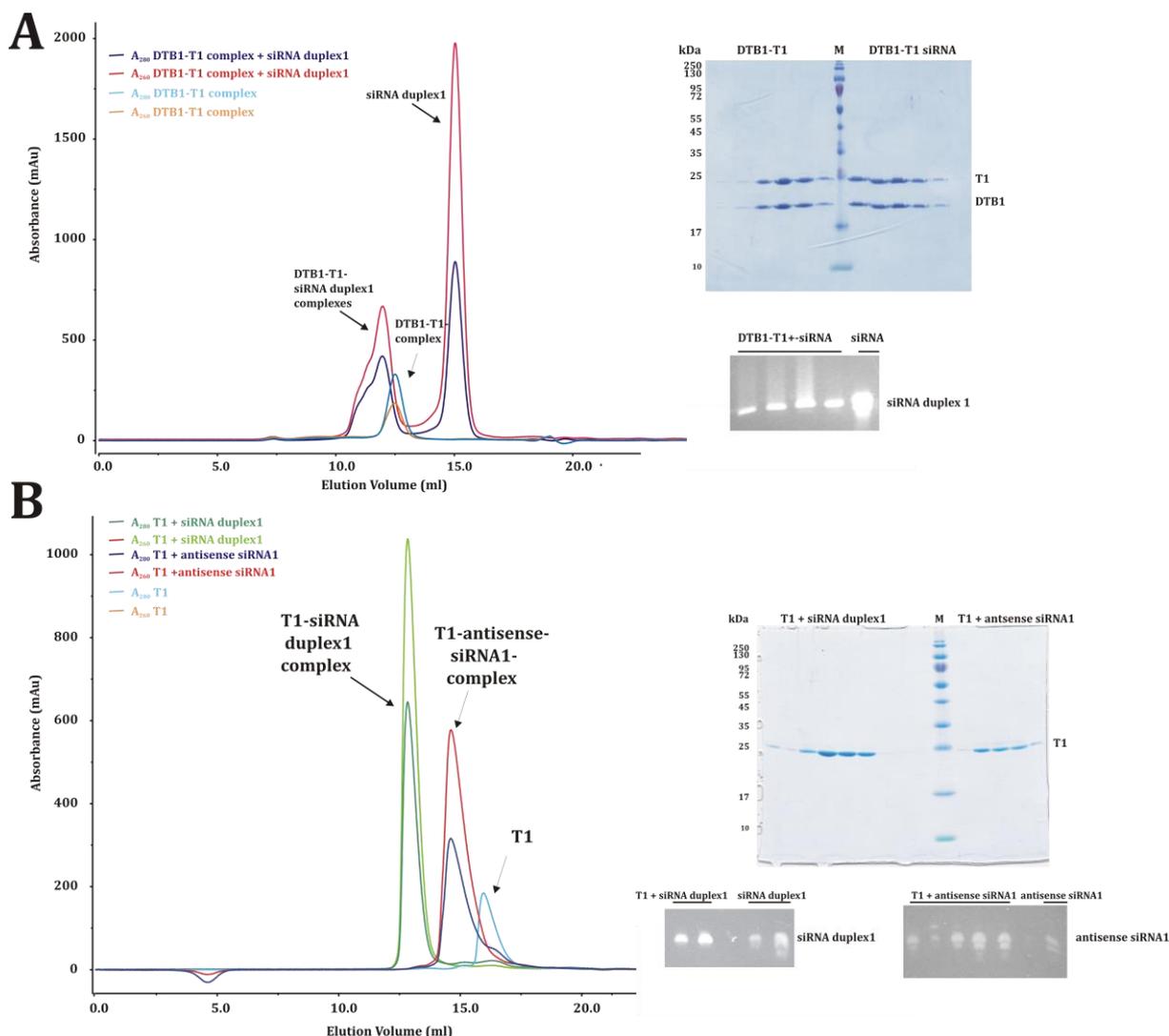
Two overlapping peaks corresponding to DTB1-T1-siRNA duplex1 complexes elute during SEC, thus different oligomers of DTB1-T1 complex could be identified, when incubated with siRNA duplex1 (Figure 32 A). The DTB1-T1-siRNA-duplex 1-complexes with a size of about 200 kDa were used for crystallization trials (6.2.4). T1 elutes with an apparent molecular weight of 33 kDa and upon addition of antisense siRNA1, the peak shifts to a molecular weight of 50.1 kDa, thus a T1 homodimer seem to bind one single stranded RNA molecule ( $MW_{\text{theoretical}} = 56.6$  kDa). Upon addition of siRNA duplexes, T1 forms complexes with a molecular weight of 120 kDa, thus forming higher oligomers (Figure 32 B).



**Figure 31: EMSAs of DTB1-T1 complex, hTRBP2 and hTRBP2 fragments.**

**A:** EMSAs of hTRBP2 dsRBDs with antisense siRNA1 and siRNA duplex2. Human TRBP2s dsRBD1 and dsRBD2 are capable to recognize and bind single stranded siRNA and siRNA duplexes, whereby the dsRBD2 binds siRNA duplexes with a higher affinity than the dsRBD1 (less amount of free RNA). The Medipal domain (dsRBD3) cannot interact with the single stranded siRNA or siRNA duplexes. The dsRBD1 containing aa 22-99 of hTRBP2, the dsRBD2 containing aa 157-366 of hTRBP2 and the dsRBD3 containing aa 291-366 of hTRBP2 are also named T4, T7 and T3, respectively, in this work. **B:** EMSA's of DTB1-T1 complex, T1 and hTRBP2 with antisense siRNA1 and siRNA duplex1. Human TRBP2 and T1 are capable to recognize and bind single stranded siRNA and siRNA duplexes, whereby the siRNA duplex binding seems stronger (less free RNA). The DTB1-T1 complex interacts with siRNA duplexes.

## Results

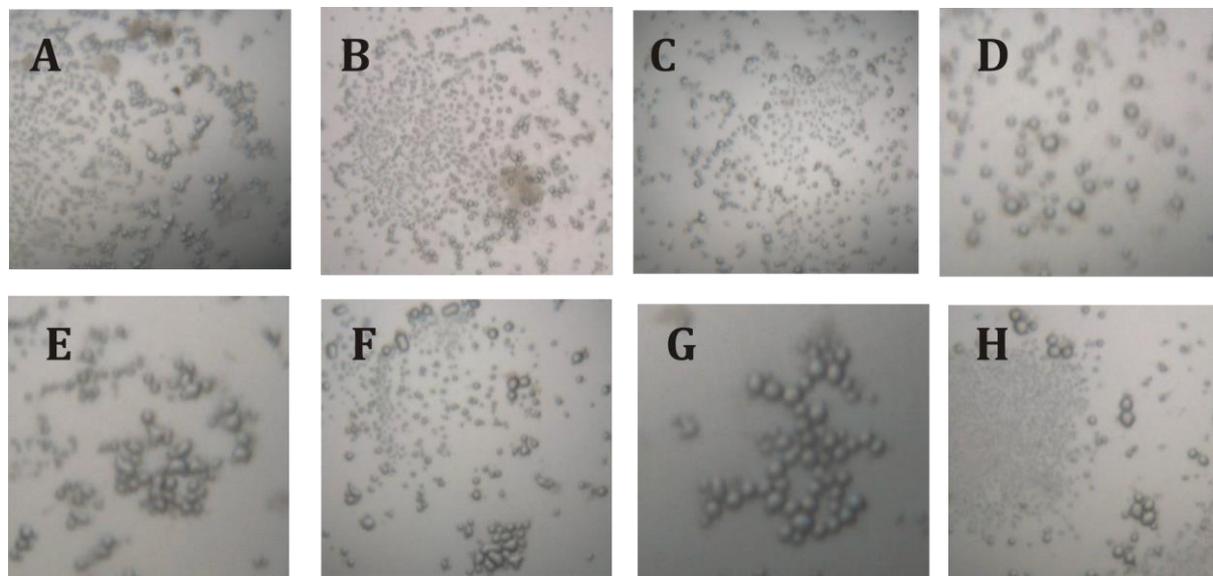


**Figure 32: Analytical SEC of DTB1-T1 and T1.** **A:** Superdex 200 (10/300) elution profiles of DTB1-T1 and DTB1-T1-siRNA duplex1 complexes with subsequent SDS- and urea-PAGE analysis. The absorbance at 280 nm and 260 nm in mAu is shown (light blue = 280 nm absorption of DTB1-T1 complex; orange = 260 nm absorption of DTB1-T1 complex; blue = 280 nm absorption of DTB1-T1-siRNA duplex1 complex; red = 260 nm absorption of DTB1-T1-siRNA duplex1 complex). The corresponding complexes and the free siRNA are indicated within the chromatogram. 15% SDS-PAGE and urea-PAGE analysis of the peak fractions of the SECs of DTB1-T1 and DTB1-T1-siRNA duplex1 complexes are shown in the upper right part and the lower right part of the figure, respectively. **B:** Superdex 200 (10/300) elution profiles of T1, T1-antisense siRNA1 complex and T1-siRNA duplex1 complex with subsequent SDS- and urea-PAGE analysis. The absorbance at 280 nm and 260 nm in mAu is shown (light blue = 280 nm absorption of T1, orange = 260 nm absorption of T1, blue = 280 nm absorption of T1-antisense-siRNA-1-complex; red = 260 nm absorption of T1-antisense-siRNA-1-complex; purple = 280 nm absorption of T1-siRNA duplex1-complex, yellow = 260 nm absorption of T1-siRNA duplex1-complex). 15% SDS-PAGE and urea-PAGE analysis of the peak fractions of the SECs of DTB1-T1 and DTB1-T1-siRNA duplex1 complexes are shown in the upper right part and the lower right part of the figure, respectively.

### 6.2.4 Crystallization trials of minimal human Dicer-TRBP2 complexes

The purified minimal human Dicer-TRBP2 complexes were subjected to initial crystallization trials as described in 5.2.4.8. In a condition containing 0.1 M HEPES/NaOH pH 7.0, 0.1 M KCl and 15% polyethyleneglycol monomethyl ether 5000 (PEG MME 5000), DTB1-T1 complex spherulites could be obtained (Figure 33). These spherulites could be reproduced when mixing 4-4.7 mg/ml DTB1-T1 complex with a reservoir containing 13.3-26.7% PEG MME 5000 or PEG

MME 2000, 0.1 M HEPES pH 6.5-8.5 and 0-0.3% KCl at 20°C (Figure 33). Extensive optimization including variations in the type and the concentration of PEG, the kind and pH of buffers, changes of the cations, additive screening and microseeding using the spherulites did not lead to crystal formation.



**Figure 33: Spherulites of the DTB1-T1 and DTB2-T1 complexes.**

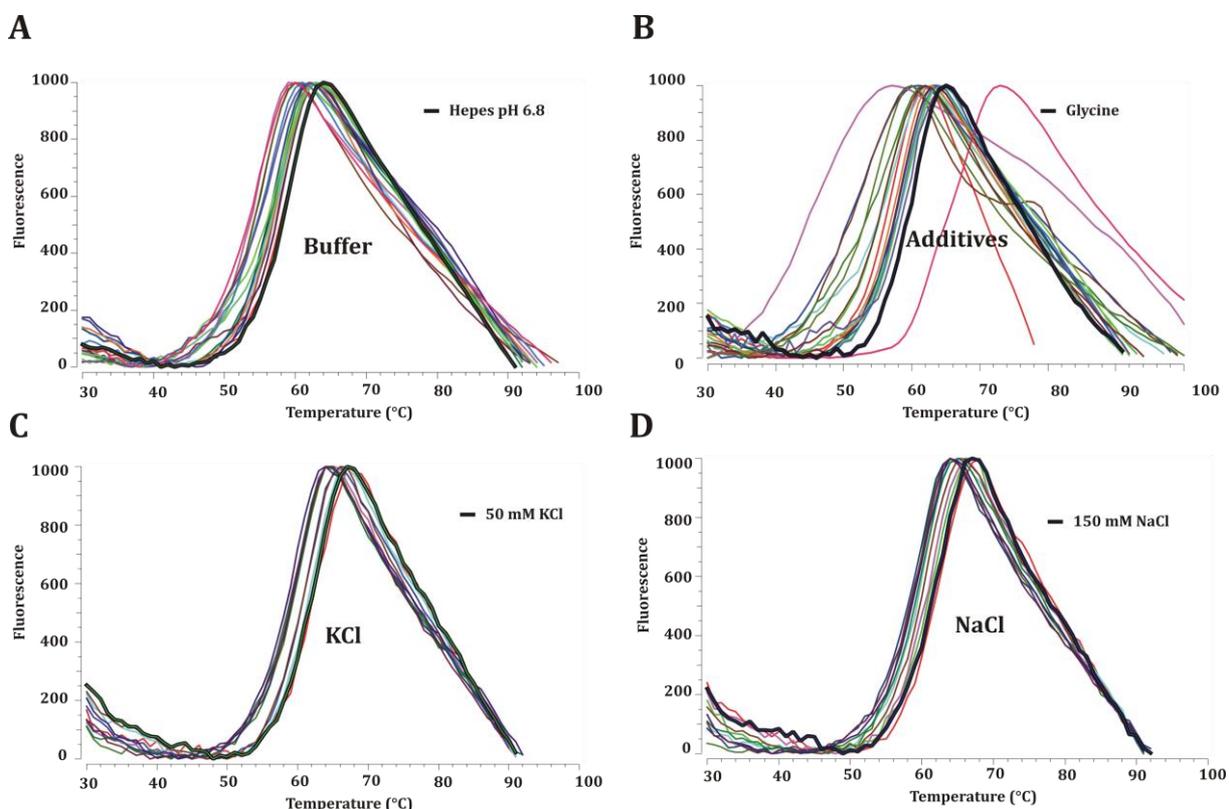
**A:** Spherulites obtained, when mixing 4.68 mg/ml DTB1-T1 complex with a reservoir containing 0.3 M KCl, 13.3% PEGMME 5000 and 0.1 M HEPES pH 7.17 at 20°C. **B:** Spherulites obtained, when mixing 4 mg/ml DTB1-T1 complex with a reservoir of 13.3% PEG MME 5000 and 0.1 M HEPES pH 6.5 in a ratio of 1:1 at 20 °C. **C:** Spherulites of the DTB1-T1 complex in a crystallization condition containing 0.1 M KCl, 15% PEG MME 5000 and 0.1 M HEPES pH 7.0 at 20°C. **D:** Spherulites obtained, when mixing 4 mg/ml DTB1-T1 complex with a reservoir of 0.1 M KCl, 15% PEG MME 5000 and 0.1 M HEPES pH 7.0 in a ratio of 2:1 at 20 °C. **E:** Spherulites of the DTB2-T1 complex in a crystallization condition containing 13.3% PEG MME 5000 and 0.1 M HEPES, pH 6.5 at 20°C. **F:** Spherulites obtained, when mixing 3.87 mg/ml DTB2-T1 complex with a reservoir of 0.12 M KCl, 13.3% PEG MME 5000 and 0.1 M HEPES, pH 6.5 in a ratio of 1:1 at 20 °C. **G:** Spherulites obtained, when mixing 3.87 mg/ml DTB2-T1 complex with a reservoir of 0.18 M KCl, 13.3% PEG MME 5000 and 0.1 M HEPES, pH 6.5 in a ratio of 1:1 at 20 °C. **H:** Spherulites obtained, when mixing 3.87 mg/ml DTB2-T1 complex with a reservoir of 0.06 M KCl, 13.3% PEG MME 5000 and 0.1 M HEPES, pH 7.17 in a ratio of 1:1 at 20 °C.

In order to optimize the DTB1-T1 complex spherulites and to optimize the stability of the complex and thus the purification yield, fluorescence-based thermal shift assays were performed (5.2.4.2). These studies revealed that the buffer used for crystallization (HEPES pH 7.5) seemed to be one of the optimal buffers (Figure 34 A). In the buffers HEPES, Tris and Phosphate at pHs between 6.6 and 7.4 the DTB1-T1 complex is most stable, which is indicated by relatively high melting temperatures of about 60 °C. In contrast, in buffers with slightly basic pH (for example phosphate pH 8.2) the DTB1-T1 complex has only a melting temperature of 53 °C and is less stable (Figure 34 A). To optimize the DTB1-T1 spherulites, the HEPES buffer of the crystallization condition was exchanged to Tris, but this only lead to precipitates. Additives have been shown to induce crystallization, thus fluorescence-based thermal shift assays using the available additive screen were performed. LiCl, CaCl<sub>2</sub> and Glycine increased DTB1-T1 complex's melting temperature to 56 °C, so these compounds were added to the initial crystallization

## Results

condition. None of these optimizations of the initial crystallization condition led to crystal formation.

Since salt concentration in which a complex is dissolved has a high influence on its stability, thermal shift assays with NaCl and KCl concentrations varying from 20 to 500 mM were performed. The DTB1-T1 is stable at salt concentrations from 20 to 500 mM NaCl or KCl. The highest melting temperature of 62 °C and 63°C could be obtained in 50 mM KCl and 350 mM NaCl, respectively (Figure 34 C and D). To optimize protein yield, the purification buffers were adapted based on these results, but no significant increase in yield could be detected. Besides, no influence in crystallization behavior could be observed, when changing the purification buffers for the DTB1-T1 complex.

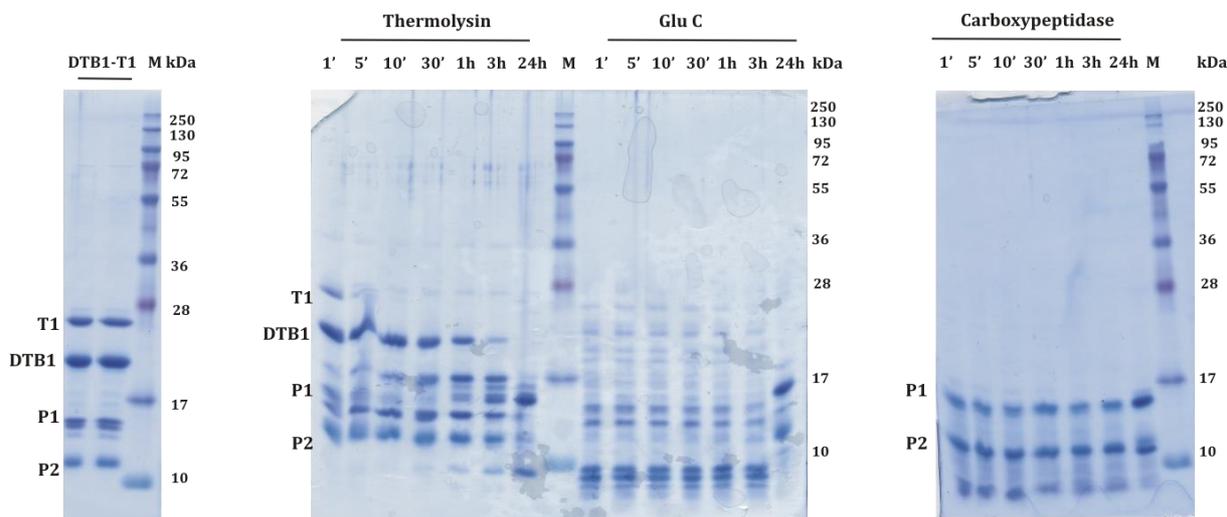


**Figure 34: Fluorescence-based thermal shift assays of the DTB1-T1 complex.**

**A:** Different melting curves of the DTB1-T1 complex using various buffers at different pH-values are shown. **B:** Different melting curves of the DTB1-T1 complex upon addition of various additives during thermal shift assays are depicted. **C:** Melting curves of the DTB1-T1 complex in 25 to 500 mM KCl. The highest melting temperature can be detected in 50 mM KCl. **D:** The DTB1-T1 complex is stable in 25 to 500 mM NaCl. The highest melting temperature is observed when using 350 mM NaCl.

Analyzing the DTB1-T1 complex two or three days after purification (left on ice), two prominent degradation bands can be detected (Figure 35). As these bands could correspond to a stable complex, limited proteolysis of the DTB1-T1 complex was performed using different proteases (5.2.3.9). To identify the protein fragments, the DTB1-T1 complex was incubated with a protease in a molar ratio of 100:1 and the reaction was analyzed by SDS-PAGE at different time points (Figure 35). Proteolysis with Thermolysin, GluC and Carboxypeptidase resulted in two

stable degradation products (Figure 35, Figure 49). It was hypothesized that crystal formation is prevented in the crystallization condition by flexible loops regions, *in situ* proteolysis using GluC and Carboxypeptidase was included in the crystallization trials. However, no crystals could be obtained.



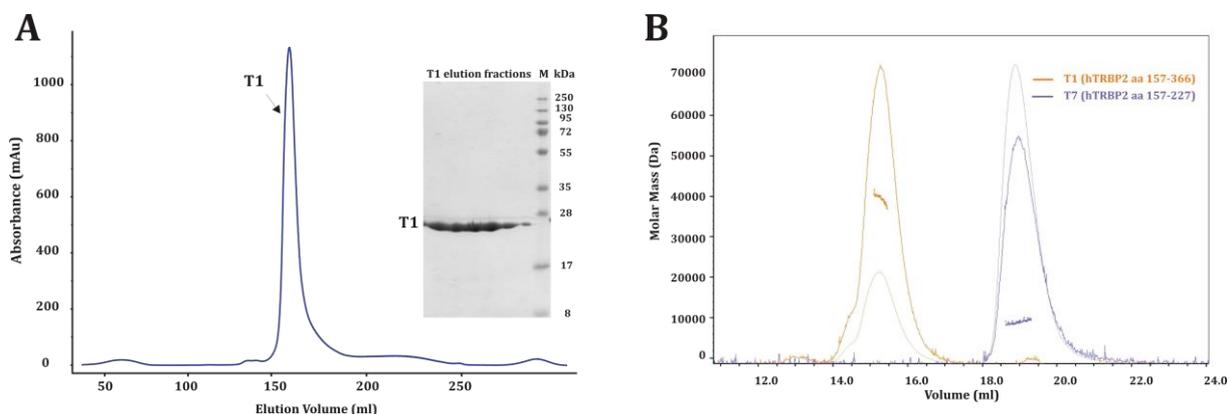
**Figure 35: Limited proteolysis of the DTB1-T1 complex.**

In the first gel, the DTB1-T1 complex was applied three days after purification and two prominent degradation products could be obtained (P1 and P2). The DTB1-T1 complex was incubated with Thermolysin, GluC or Carboxypeptidase in a molar ratio of 100:1. Samples were taken after different time points, the reaction stopped with addition of 2% SDS and applied to 15% SDS-gels. All three proteases produced stable products (P1 and P2) of the DTB1-T1 complex.

In addition to the spherulites, some precipitate could be detected in the initial crystallization condition. Most likely one component of the DTB1-T1 complex forms the spherulites, whereas the second precipitates and thereby hinders crystallization. It could only be tested how purified T1 alone behaves within the crystallization condition, because DTB1 could not be purified as single protein. When T1 was used for crystallization in 13.3-26.7% PEG MME 5000 or PEG MME 2000, 0.1 M HEPES/NaOH pH 6.5-8.5 and 0-0.3 M KCl at 20°C no spherulites could be observed. Finally, the similar DTB2-T1 complex was subjected to a crystallization setup in 13.3-26.7% PEG MME 5000 or PEG MME 2000, 0.1 M HEPES/NaOH pH 6.5-8.5 and 0-0.3% KCl at 20°C. Using this complex at slightly lower concentration than the DTB1-T1 complex, spherulites could be obtained (Figure 33). For the DTB2-T1 complex extensive crystallization trials have not been carried out yet. To date all attempts to obtain diffracting crystals from a minimal human Dicer-TRBP complex failed.

### 6.3 Preparation and crystallization of hTRBP2-domains

Human TRBP2 is a dsRBP containing two dsRBDs for both RNA-binding and homodimerization and a Medipal-domain for interaction with other RNAi components (St. Johnston *et al.*, 1992; Barber *et al.*, 2009; Daviet *et al.*, 2000; Erard *et al.*, 1998; Gagnol *et al.*, 1991; Laraki *et al.*, 2008). Various structures of individual dsRBDs from plants and mammals have been solved, but structures of several dsRBDs or whole dsRBPs are rare. In addition, structures of Medipal domains or type B dsRBDs have not been solved so far. Such structures will give important information about the interaction surface for binding to protein partners. As crystallization of full length human TRBP2 in the absence or presence of RNA (6.1.4), several hTRBP2 fragments were designed (5.1.12). To investigate the structure of the dsRBD3, the hTRBP2 fragment T1 (aa157-366) containing dsRBD2 and dsRBD3, T2 (aa 226-366) and T3 (aa 291-366), each containing the dsRBD3 were cloned, expressed in *E. coli* BL21 (DE3) Star cells and purified (5.1.12; 5.2.1.1; 5.2.2.5; 5.2.3.4). The purification of the T2 and the T3 fragments are shown in Figure 52. A SEC with a corresponding SDS-PAGE analysis of the 22.7 kDa large T1 is shown in Figure 36. SEC combined with MALS reveals that T1 has a molecular weight of 38 ( $\pm$  4) kDa and thus is supposed to exist as homodimer in solution (Figure 36 B).

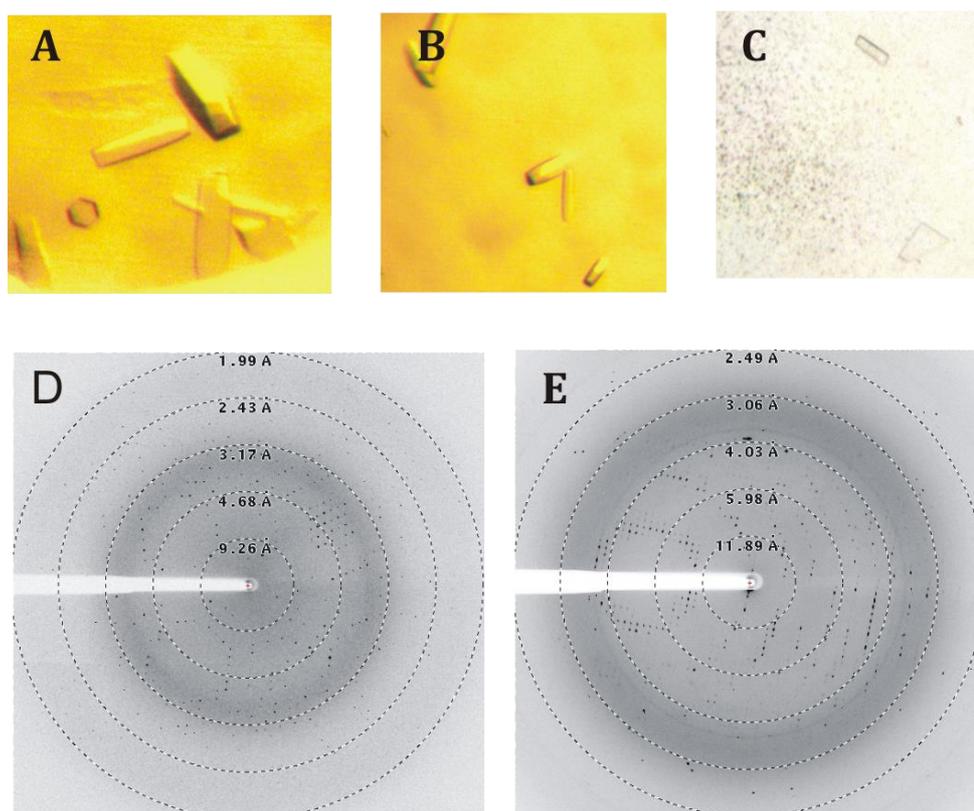


**Figure 36: SEC and MALS of hTRBP2 fragments.**

**A:** Superdex 75 (26/60) elution profile of T1 (hTRBP2 aa 157-366) monitored by absorption at 280 nm is shown. A subsequent 15% SDS-PAGE analysis (Coomassie stained) is shown next to the chromatogram. The identified protein is indicated on the left and size marker is shown on the right of the gel (M= Marker). **B:** SEC combined with MALS of T1 and T7 (hTRBP2 aa 157-227). The absorption (continuous line) and the light scattering (dashed line) of the T1 (orange) and the T7 (purple) are plotted against the elution volume in ml. The molar mass distribution of the complexes is shown as orange and purple dots for the T1 and T7, respectively.

It could be shown in section 6.2.3 that the T1-fragment is able to bind guide and duplex siRNA and that the T3 fragment does not bind RNA (Figure 31, Figure 32). Furthermore T1 and T3 are able to form complexes with hDicer's TRBP-binding domain and thus seem to be properly folded and fully functional (6.2.2). The functionality was also emphasized by pull down assays, revealing that T3 containing only the Medipal-domain is able to form a RLC complex (5.2.3.10;

Figure 53). In order to obtain structural information of functional hTRBP2 fragments T1, T2, T3, the T1-antisense-siRNA1-complex and the T1-siRNA-duplex-1 complex were subjected to initial crystallization trials (5.2.4.8.1; Table 4). However no pre-crystalline precipitate or protein crystals could be obtained. The reason might be the existence of flexible loops hindering crystallization processes. The method of *in situ* proteolysis can cleave flexible loops and thus can lead to reduction of the conformational heterogeneity and promote crystallization of proteins that had previously failed to crystallize (Dong *et al.*, 2007). Prior to crystallization trials T1, T2, T3 or the T1-siRNA complexes were incubated with Chymotrypsin in a ratio of 500:1 (see 5.2.4.8.2 and Table 4). Hexagonal-shaped crystals of T1 could be obtained after 2 month at 4 °C and 20 °C in a condition containing 3.2 M Ammoniumsulfate and 0.1 M Tris, pH 8.0 (Figure 37 A and B). Crystals were cryo protected in a solution containing ionic conditions as the reservoir with 10-15% glycerol. A typical diffraction image is shown in Figure 37 D. Plate-shaped crystals of T1 were achieved after 3 weeks at 20°C in a condition containing 2 M Li<sub>2</sub>SO<sub>4</sub> and 3 % 2-Methyl-2,4-pentanediol (Figure 37 C). The crystals were flash-cooled in liquid nitrogen without any need of further cryoprotectant. A representative diffraction image is shown in Figure 37 E.



**Figure 37: Crystallization and diffraction images of T1 crystals.**

**A:** Crystals grown in a condition containing T1 and 3.2 M Ammoniumsulfate and 0.1 M Tris, pH 8.0 after 2 month at 20 °C. **B:** Crystals grown in a condition containing T1 and 3.2 M Ammoniumsulfate and 0.1 M Tris, pH 8.0 after 2 month at 4 °C. **C:** Plate-shaped crystals of T1 in a condition containing 2 M Li<sub>2</sub>SO<sub>4</sub> and 3% MPD after 3 weeks at 20°C. **D:** Diffraction image of a single crystal from B collected at beamline ID23-2 (ESRF, Grenoble). **E:** Diffraction image of a single crystal from C collected at beamline ID23-2 (ESRF, Grenoble).

### 6.3.1 Determination of the crystal structure of the second dsRBD of hTRBP2

Complete datasets of both crystal-forms were collected on the microfocus beamline ID23-2 of the ESRF (Grenoble), processed with XDS and scaled using XSCALE (Kabsch, 2010). The corresponding data statistics are depicted in Table 1. Data of hexagonal-shaped crystals were processed in hexagonal space group  $P3_121$  and scaled to a final resolution of 2.28 Å. In contrast, the plate shaped crystals belonged to the tetragonal space group  $P4_22_12$  and diffracted to a final resolution of 2.95 Å (Table 1).

**Table 1: Data collection and processing of T1 crystals and plate-shaped crystals.**

	T1 (hexagonal crystals)	T1 (crystal-plates)
<b>Data collection</b>	ESRF	ESRF
Wavelength (Å)	0.8726	0.8726
<b>Space group</b>	$P3_121$	$P4_22_12$
<b>Cell dimensions</b>		
a, b, c (Å)	70.20, 70.20, 122.91	89.10, 89.10, 82.58
$\alpha, \beta, \gamma$ (°)	90, 90, 120	90, 90, 90
<b>Resolution range</b>	43.22-2.28 (2.38-2.28)	17.00-2.95 (3.05-2.95)
<b>No. of reflections</b>	100086 (11976)	57989 (5516)
<b>No. of unique reflections</b>	16490 (1936)	7284 (679)
<b>Completeness (%)</b>	99.9 (100)	98.0 (99.4)
<b>I/sigma(I)</b>	13.67 (3.34)	16.87 (2.95)
<b>R<sub>merge</sub> (%)</b>	11.2 (64.6)	13.4 (76.7)
<b>Redundancy</b>	6.07 (6.19)	7.96 (8.12)

$R_{\text{sym}}(I) = \sum hkl \sum i |I_i(hkl) - \langle I(hkl) \rangle| / \sum hkl \sum i I_i(hkl)$  † for  $n$  independent reflections and  $i$  observations of a given reflection;  $\langle I(hkl) \rangle$  – average intensity of the  $i$  observations

For both crystal types, the phase problem could be solved by molecular replacement in PHASER (McCoy *et al.*, 2007) using the crystal structure of the second dsRBD of hTRBP2 in complex with RNA, whereby the RNA was removed, as search model (PDB-ID: 3ADL; Yang *et al.*, 2010). For the hexagonal crystals with space group  $P3_121$  four molecules were identified in the asymmetric unit (Figure 38). Manual model building and refinement were performed with Coot and PHENIX until the crystallographic R and  $R_{\text{free}}$  factors converged. The final model was refined at a resolution of 2.28 Å to  $R_{\text{work}}$  and  $R_{\text{free}}$  values of 19.85% and 24.09%, respectively, retaining good stereochemistry. The refinement statistics are provided in Table 2 and the structure is described in 6.3.2.

**Table 2: Refinement statistics of dsRBD2 of hTRBP2**

T1	
Refinement	PHENIX
Resolution range (Å)	43.223-2.285
Reflections (working/test)	16456
R <sub>work</sub> /R <sub>free</sub> (%)	19.45/24.09
Number of atoms	2209
Protein	2159
Solvent	60
B-factor (Å <sup>2</sup> )	39.135
B-factor Protein (Å <sup>2</sup> )	34.138
B-factor Solvent (Å <sup>2</sup> )	44.127
Deviations from ideals (r.m.s.d.)	
Bond distances (Å)	0.009
Bond angles (°)	1.32
Dihedrals (°)	17.026
Ramachandran plot	
favoured regions (%)	98.48
allowed regions (%)	1.52
outlier regions (%)	0

R =  $\sum hkl |F_{obs} - F_{calc}| / \sum hkl |F_{obs}|$ ; R<sub>work</sub> =  $\sum hkl \notin T$ ; R<sub>free</sub> =  $\sum hkl \in T$ ; T = test set

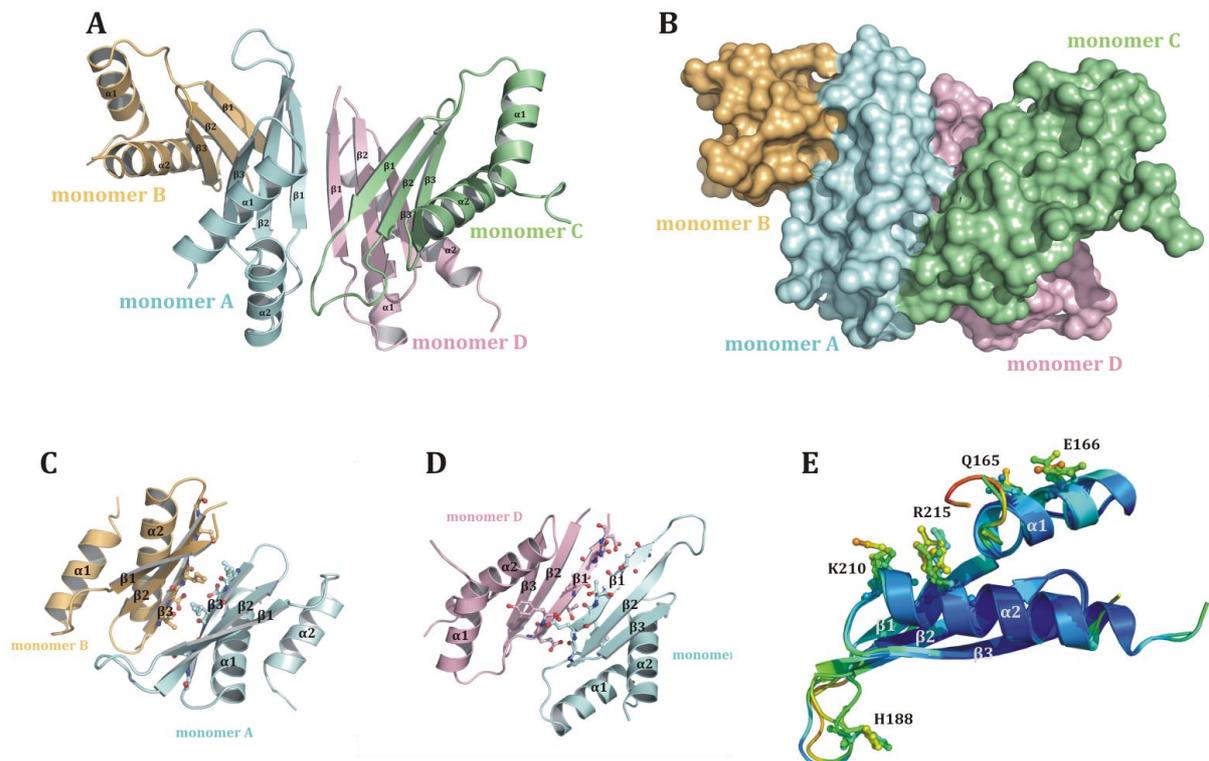
The low resolution of about 3 Å of the T1 tetragonal crystals made molecular replacement more complicated. Using either the structure of the dsRBD2 (solved and refined in this work) or the structure of the dsRBD2 that has been solved by others as a search model (PDB-ID: 3ADL), it was possible to find three monomers of dsRBD2 of hTRBP2 in the asymmetric unit when using the space group P4<sub>2</sub>2<sub>1</sub>2. The B-factors calculated by PHASER during molecular replacement of each particular solution suggested high level of thermal disorder of the second and the third molecule. This could be a result of *in situ* proteolysis and conformational heterogeneity of different molecules occupying the asymmetric unit. Due to high value of solvent content (76 %) of that solution and high R-factors of 43.48% and 39.24% for R and R<sub>free</sub>, respectively an additional molecular replacement search was performed but no reliable position of the fourth dsRBD2 molecule could be found. Molecular replacements with truncated models did not led to a reliable structure to date.

### 6.3.2 Overall structure of the second dsRBD of human TRBP2

As described above the crystal structure of the second dsRBD of hTRBP2 comprising amino acids 157-226 could be solved and refined. The asymmetric unit of the hexagonal-shaped crystals contains four dsRBD2 monomers and is shown in Figure 38 A and B. Each monomer consists of a typical  $\alpha$ - $\beta$ - $\beta$ - $\beta$ - $\alpha$  fold, in which the two  $\alpha$ -helices pack against one side of the three-stranded antiparallel  $\beta$ -sheet, a common feature of dsRBDs (see 4.2.3; Figure 38). Interactions between the different monomers in the crystallographic homotetramer are mediated by  $\beta$ 1 and  $\beta$ 3 strands of the different monomeric units (Figure 38 C and D). Superposition of the monomers in the unit cell reveals that each monomer has a similar overall structure (r.m.s.d. of 0.459-0.721

## Results

Å) with conformational differences in the loop region between the  $\beta$ 1- and  $\beta$ 2-strands (Figure 38 E). This loop region of monomer B (residues 184-190) and D (residues 185-189) could not be built into the electron density map and is most likely disordered. In contrast, these loop regions of monomers A and C are involved in crystal packing and are not exposed to the solvent. The missing loop regions of monomer B and D might be a result of protease cleavage or might be due to flexibility in this region as indicated by coloring of each monomer according to B-factors from monomer A and C (Figure 38 E). Apart from the N- and C-termini, the B-factors of the loop region between  $\beta$ 1- and  $\beta$ 2-strands are the highest ones of monomer A and C (Figure 38 E). Interestingly, H188 that is important for dsRNA-binding by hTRBP2's dsRBD2 is located in this loop. Additionally residues Gln165, Glu166, Lys210, Lys214 and Arg215 of dsRBD2 are important for dsRNA-binding (Yamashita *et al.*, 2011; see 7.2.1). The orientation of these residues among the different monomers varies and the B-factors of these residues are relatively high indicating a high thermal mobility (Figure 38 E).

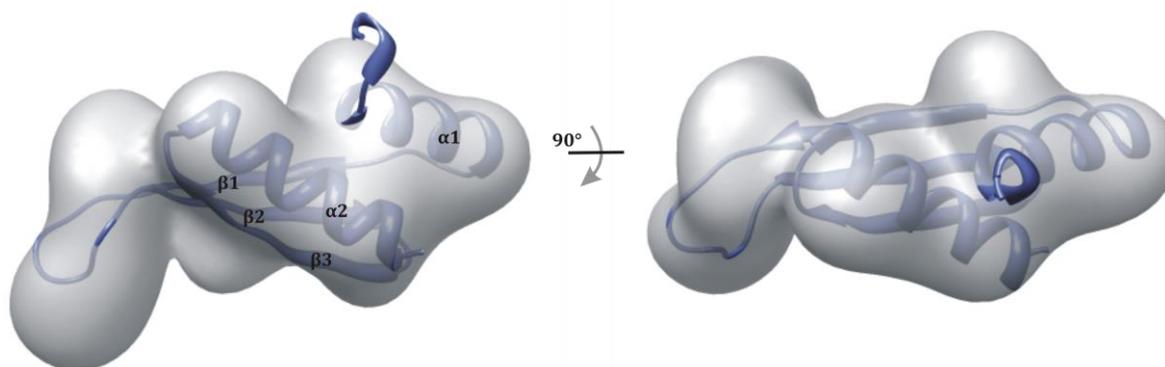


**Figure 38: Overall structure of the dsRBD2 of hTRBP2.**

**A:** Crystallographic tetramer of hTRBP2's dsRBD2 is shown in cartoon representation. The  $\beta$ 1- and  $\beta$ 3-strands of the molecules mediate the assembly of a two dimer in the crystal. **B:** Surface representation of the asymmetric unit of hTRBP2's dsRBD2, each monomer is shown in light blue, yellow, light-green or pink. **C:** Residues of the  $\beta$ 3-sheet of the dsRBD2 of hTRBP2 can interact with each other and mediate dimerization. **D:** Residues of the  $\beta$ 1-sheets pack against each other and mediate dimerization. **E:** Superposition of the different monomers of the crystallographic tetramer colored according to B-factors. Each monomer has a typical  $\alpha$ - $\beta$ - $\beta$ - $\alpha$ -dsRBD-fold. Apart from the N- and C-termini, the regions/residues that are important for RNA-binding are the most flexible ones and are shown in ball-and stick representation.

### 6.3.3 Crystallization artifact or dimerization interface?

In general, dsRNA-binding proteins have the tendency to dimerize and hTRBP2 was shown to form homodimers and heterodimers with other proteins (4.2.3). The crystal structure of hTRBP2 dsRBD2 reveals two possible dimerization interfaces for the interaction of two dsRBDs with each other (see discussion 7.2.2; Figure 38 C and D). In order to investigate the behavior of hTRBP2's dsRBD2 in solution, an hTRBP2 construct containing amino acids 157-226 (T7) was cloned, expressed and purified (5.1.12; 5.2.1.1; 5.2.2.5; 5.2.3.4). Since SEC was not accurate enough to determine the correct molecular mass of such a small protein (9 kDa), MALS and SAXS measurements were carried out (5.2.4.1; 5.2.4.7). All SAXS data analysis steps were carried out by Clement Blanchet from the group of Dmitri Svergun (EMBL, Hamburg). SEC combined with MALS revealed that T7 is monodispers and has a molecular mass of 9.3 ( $\pm$  0.5) kDa (Figure 36). SAXS measurements confirmed that T7 exist as monomer in solution having a radius of gyration ( $R_g$ ) of 1.3 and a molar mass of 9 kDa (Figure 39). Docking of the crystal structure of the second dsRBD of hTRBP2 into the SAXS model uncovers that the  $\alpha$ 1-helix and the loop region between the  $\beta$ 1- and the  $\beta$ 2-strands seem to be highly flexible in solution. This was assumed, because there is extra free density at these regions that could be the result of flexibility (Figure 39; see 7.2.2).



**Figure 39: Docking of the crystal structures of the second dsRBD of hTRBP2 into a calculated SAXS damfilt model.** The SAXS damfilt model reveals that the dsRBD2 exists as monomer in solution. Furthermore, the regions important for RNA-binding are highly flexible allowing the binding of various RNA substrates.

Results

## **7 Discussion**

### **7.1 The human RISC-loading complex**

The major goal of this thesis, was the structural analysis of the minimal human RLC comprising Dicer, Ago2 and TRBP2. In the first parts of this section, the purification and functionality of the individual RLC proteins are discussed. Thereafter, the results of preparation and loading ability of the RLC are examined, followed by the comparison of the reconstructed RLC EM map with other EM structures of the RLC and its subcomplexes. Finally, a model for RLC function is discussed.

#### **7.1.1 Purification and functionality of hAgo2**

The human Ago2 has been shown to function as a monomeric protein (4.2.2; Figure 4). In this work, sufficient amounts of monomeric N-hAgo2 could be obtained from baculovirus-infected High Five insect cell extracts by affinity purification and subsequent SEC (6.1.1.1; Figure 7). This procedure involved the addition of 1-2 M LiCl during cell lysis to disrupt protein-nucleotide interactions. Though a relative high 260 nm/280 nm value of 0.98 and RNA contaminations detected by urea-PAGE indicate that N-hAgo2 is unspecifically loaded with endogenous RNAs from insect cells (see Figure 54). Thus, the addition of 1-2 M LiCl during cell lysis did not lead to RNA-free hAgo2. Completely RNA-free Ago2 protein could also not be produced when RNase A was added during cell lysis or during later purification steps. Therefore it seems that the endogenous RNA is bound tightly, most likely in a pocket inaccessible to RNase A. Generally, under low salt concentration (0-100 mM) RNase A cleaves single and double stranded RNA, but at salt concentrations of over 300 mM NaCl RNase A specifically cleaves single stranded RNA (Ausubel *et al.*, 1994). Hence, most likely the specificity of the RNase A for dsRNA was reduced under the salt concentrations used within this study (>300 mM KCl under cell lysis conditions or >150 mM before SEC). Other groups also observed binding of endogenous RNA to hAgo2, when hAgo2 was isolated from baculovirus-infected Sf9 insect cells (MacRae *et al.*, 2008; Schirle *et al.*, 2012; Elkayam *et al.*, 2012). They further investigated the nature of the bound RNAs by small RNA cloning and deep sequencing techniques. The vast majority of bound RNA has a length of 19-23-nts and belongs to siRNA and miRNA Sf9 RNA species. Therefore, human Ago2 is able to bind siRNAs and miRNAs from other species. Furthermore, the stability of Ago2 is increased, when it is bound to RNA (Rivas *et al.*, 2005; Elkayam *et al.*, 2012). Based on these results, one can suppose that the preparations of N-hAgo2 in this work contained both free N-hAgo2 and N-hAgo2 bound to endogenous RNA. However, the purified N-hAgo2 and hAgo2

## Discussion

proteins were shown to be functional as they are endonucleolytically active and can interact with N-hDicer to form RLCs (see sections 6.1.1.1, 6.1.2.1, Figure 10-12).

Recently, the crystal structure of full length human Ago2 was solved (Schirle *et al.*, 2012; Elkayam *et al.*, 2012). Several factors might have abolished crystallization of the hAgo2 protein that was used in this study. First of all, there seem to be two types of hAgo2 (RNA-free and RNA-bound) in the preparations. Second, the RNA-free hAgo2 is less stable as it seems more flexible. Third, the hAgo2 expressed in High Five insect cells could be differently posttranscriptionally modified than hAgo2 from Sf9 cells (see Figure 54). Finally, the addition of 0.12 M phenol combined with a basic pH of 9.0 was shown to be critical for crystallization (Schirle *et al.*, 2012; Elkayam *et al.*, 2012). Such high pH's in combination with this high amounts of the organic compound phenol were never added to the crystallization conditions during this work.

### **7.1.2 Purification and endonuclease activity of the hDicer**

Human Dicer expression in insect cells using the native gene coding sequence resulted in low expression rates. In order to increase expression rates, a hDicer gene optimized for expression in Sf9 insect cells with an additional C-terminal Strep-tag was ordered from GenScript (5.2.1.1, 10). Indeed, expression rates increased, but this 223 kDa large multidomain protein could not be purified in a reproducible manner and often resulted in polydisperse N-hDicer protein and low amounts of functional monomeric protein (Zhang *et al.*, 2004; 6.1.1.3; 6.1.5.3). In addition, the hDicer protein contains 48 cysteine residues that could form disulfide bonds, which may result in different oligomers. Although, the purification of N-hDicer was carried out under reducing conditions, gradient centrifugations (GraFix preparations) and subsequent negative-stain EM analysis revealed that hDicer still forms various oligomers and tends to aggregate (6.1.5.3). This could not be overcome to date. Other groups working on human Dicer detected, that it runs abnormally fast through SEC and elutes at 440 kDa, which would correspond to a homodimer (MacRae *et al.*, 2008). Additionally, on negative-stain micrographs different oligomers of hDicer could be detected, thus it seems that sample heterogeneity is a general problem during hDicer purification (Wang *et al.*, 2009b; Lau *et al.*, 2009; Lau *et al.*, 2012a). Moreover, the purified recombinant Dicer showed a prominent degradation pattern (6.2.1; Figure 47). These degradation products co-eluted with N-hDicer on SEC and were able to bind hTRBP2 and hAgo2. Western blot and mass spectrometry analyses revealed that the 140 kDa fragments contain the N-terminal domains ranging from the N-terminus to the PAZ domain of human Dicer, whereas the ca. 100 kDa fragments contain the RNase III domains and the C-terminal dsRBD of hDicer. This observation is in good agreement with experiments from other groups, which could reconstitute an active and functional human Dicer enzyme by co-expression of an N-terminal and a C-terminal fragment of human Dicer (Noland *et al.*, 2011; Ma *et al.*, 2012). It was mentioned that the prominent degradation bands result from a protease sensitive loop in the

RNase III domains, but an exact location (sequence) of this region has not been published so far (Lau *et al.*, 2012b). Interestingly, purified endogenous human Dicer shows the same prominent degradation pattern (personal communication Anne Frohn). It seemed that the cleavage does not alter the function and structure of hDicer (Noland *et al.*, 2011; Ma *et al.*, 2012; Lau *et al.*, 2012b). As higher sample heterogeneity in negative-stain micrographs was observed when using Dicer samples with higher degree of degradation, the addition of protease inhibitors and EDTA was tried to reduce degradation and mostly obtain the full length protein. The purified N-hDicer was endonucleolytically active and interacts with the other RLC components hAgo2 and hTRBP2, thus, is functional.

### **7.1.3 Preparation and RNA-binding abilities of the dsRBP hTRBP2**

Human TRBP2 could be expressed and purified in from *E. coli* BL21 (DE3) Star cells in meaningful amounts (6.1.1.2). SEC combined with MALS revealed that hTRBP2 monomers and homodimers stay in equilibrium at salt concentrations varying from 0-1 M NaCl (6.1.1.2). It has been shown that hTRBP2 exists as homodimer in solution *in vivo* and *in vitro* using yeast two hybrid screening in most cases (Cosentino, *et al.*, 1995; Duarte *et al.*, 2000; Gatignol *et al.*, 1996; Daher *et al.*, 2001; MacRae *et al.*, 2008). It seems that the homodimerization is mediated by strong interactions that cannot be easily disrupted by high salt concentrations (Figure 8). Later investigations showed that hTRBP2 homodimer only dissociates at really low concentrations (1  $\mu$ M TRBP2) into monomers (Yamashita *et al.*, 2011). SEC analysis revealed that recombinant hTRBP2 forms larger complexes upon addition of siRNA duplex1 (see Figure 13). These observations are in agreement with other groups finding that hTRBP2 forms larger complexes containing at least one siRNA duplex and up to six MBP-hTRBP2 molecules when performing native gel shift assays (Gredell *et al.*, 2010). To date it is not known which of these complexes are functionally relevant. ITC experiments revealed that the siRNA duplex binding by hTRBP2 is an enthalpy driven and not entropy driven process (Figure 13). Therefore most likely, hTRBP2 interacts with the siRNA phosphate backbone through ionic interactions or hydrogen bonds. This hypothesis was structurally and experimentally confirmed later as two groups were able to crystallize dsRBD1 of hTRBP1 (PDB-ID: 3LLH) and dsRBD2 of hTRBP2 in complex with a 10-nt GC-duplex (PDB-ID: 3ADL). They could show that these dsRBDs interact with the siRNA phosphate backbone via salt and hydrogen bridges (Figure 44; Yang *et al.*, 2010; Yamashita *et al.*, 2011). Moreover, the hTRBP2 siRNA duplex interactions were analyzed in detail using ITC by Yamashita *et al.*, 2011. They obtained a  $K_d$  of 0.24 nM for hTRBP2-siRNA duplex interaction, which is about the same value that could be obtained in this study. In further agreement with their results is, that hTRBP2's first and second dsRBDs were shown to be able to interact with siRNA duplexes, whereas the third dsRBD cannot interact with siRNA duplexes, although no  $K_d$  values for the single domains were obtained in this study (see Figure 13). Another group also

## Discussion

obtained a similar  $K_d$  value of 0.29 nM for hTRBP2 binding to a 19-nt duplex (Chakravarthy *et al.*, 2010). As the  $K_d$  value of the recombinant hTRBP2 to the siRNA duplex used in this study is similar to those obtained by other groups, a fully functional binding to dsRNA of hTRBP2 was assumed (6.1.2.2; Yamashita *et al.*, 2011; Chakravarthy *et al.*, 2010). As described above, based on SEC and combined SEC/MALS experiments of recombinant hTRBP2, it was assumed that hTRBP2 exists in equilibrium between homodimer and monomer under various salt concentrations (see Figure 13). From EMSA and SEC experiments it is not clear if hTRBP2 binds siRNA duplexes as a monomer or a dimer. In 2011, Yamashita *et al.* performed ITC measurements with hTRBP2 and a siRNA and assigned TRBP2 binding to siRNA as monomers. Under the conditions used during ITC measurements, only about 1  $\mu$ M TRBP2 solution was present within the cell chamber and the authors defined that at this concentration TRBP2 exists as a monomer (Yamashita *et al.*, 2011). TRBP2, when present in a higher concentration exists as a homodimer and also at these concentrations binds siRNAs. So it is still unclear if TRBP2 binds siRNAs as a monomer or a dimer. The ITC measurements of Yamashita *et al.* 2011 observed a stoichiometry of 1:1, so it could not be detected if an hTRBP2 monomer binds one siRNA duplex or if an hTRBP2 homodimer binds two siRNA duplexes. As hTRBP2 exists in equilibrium between homodimer and monomer in solution both scenarios are possible. SEC revealed that hTRBP2 forms different oligomers with molecular masses of 90-180 kDa upon addition of siRNA duplexes (Figure 13). This would exclude binding of one hTRBP2 monomer to one siRNA duplex (60 kDa) and indicate that hTRBP2 homodimer or higher oligomers bind to siRNA duplexes. To fully address this question, additional ITC measurements would be necessary.

### **7.1.4 Purification and homogeneity analysis of RLC**

By combining the individually purified N-hDicer, N-hAgo2 and hTRBP2 proteins followed by subsequent SEC the RLC could be reconstituted *in vitro* (Figure 10; 6.1.1.4). As N-hDicer is very unstable and heterogeneous as single protein and binding partners were known to stabilize Dicer, a co-purification strategy was employed. This resulted in higher amount of purified RLC (0.3 mg vs. 2 mg), which was used for biochemical and structural analysis (Figure 11; 6.1.1.4; 6.1.5.3; MacRae *et al.*, 2008; Wang *et al.*, 2009b). Using SEC combined with MALS and GraFix combined with negative-stain EM analysis it became clear that the *in vitro* reconstituted RLC is heterogeneous in size (6.1.3.1; 6.1.5). Additionally, the different methods to analyze the size of the RLC resulted in different observed molar masses of the RLC. Using SEC and SEC/MALS-analysis molar masses of 400 kDa and 324 ( $\pm$  60) kDa, respectively could be determined. Those two results give rise to a stoichiometry of 1:1:1 or 1:1:2 for hDicer, hAgo2 and hTRBP2 in the RLC. One could conclude that hTRBP2 can bind as a monomer or a dimer. The molecular mass of 324 kDa and the very high polydispersity ( $\pm$  60 kDa) of the RLC indicate a non-stoichiometric incorporation of individual components. Thus, a coexistence of different complexes and

subcomplexes in solution was assumed, which was confirmed by gradient centrifugation and negative-stain EM analysis (Figure 19, Figure 20). SEC experiments of the complex performed by other groups revealed a molecular mass of about 500 kDa and subsequent mass spectrometry analysis uncovered a molecular mass of 368 kDa for the RLC. With this precise method it could be determined that the RLC consists of hDicer, hAgo2 and hTRBP2 at a 1:1:1 ratio and that the RLC assembly requires hTRBP2 dissociation (MacRae *et al.*, 2008). The observations made in this study can also be due to the limitations of the used methods to determine this relative small size difference. Furthermore, a diffusion of hTRBP2 from the RLC can be detected when analyzing the RLC by gradient centrifugation and subsequent SDS-PAGE analysis (6.1.5.1). This disassembly might be induced by the low concentrations of the RLC used during gradient centrifugation or by unspecific RNA contaminations. RNA contaminations seemed to be a problem, as addition of RNase A to the complex prior to centrifugation abolished disassembly (Figure 20). A disassembly of the RLC at low concentration was also observed by others (Wang *et al.*, 2009b).

In addition to the problem to obtain a stable RLC, it was recently discovered, that hTRBP2 can be phosphorylated at 4 Serine residues by the mitogen-activated protein kinase (MAPK) Erk. Phosphorylation of hTRBP2 leads to increased stabilization of hDicer-hTRBP2 complex, which in turn enhances miRNA production (Paroo *et al.*, 2009). In contrast, when comparing the influence of wild type hTRBP2 and a phospho-mimic mutant of hTRBP2 on recombinant hDicer, no difference in cleavage activities could be detected (Chakravarthy *et al.*, 2010). Nevertheless, phosphorylation of hTRBP2 could enhance Dicer cleavage activity by other mechanisms that are not known so far. These experiments raise the possibility that phosphorylated hTRBP2 leads to hDicer-hTRBP2 stabilization. Human TRBP2 expressed in bacteria is not phosphorylated, whereas hTRBP2 expressed in insect cells exhibits phosphorylation (Paroo *et al.*, 2009). However, it has been stated that insect cells phosphorylate hTRBP2 irregularly (Chakravarthy *et al.*, 2010). In order to enhance binding of hTRBP2 to hDicer and the RLC a phospho-mimic mutant of hTRBP2 should be produced in the future, as bacterial produced hTRBP2 seems to be incorporated in sub-stoichiometric amounts into the RLC (see Figure 20). This could also lead to an increased RLC sample homogeneity.

The heterogeneity of the RLC made it hard to solve the structure of this complex by means of X-ray crystallography and EM (6.1.4; 6.1.5). In order to improve RLC sample quality, fluorescence-based thermal shift assays were performed and the conditions of the best results applied to purification and GraFix preparations to increase homogeneity (6.1.3.2). Some of the stabilizing agents detected during fluorescence-based thermal shift assays improved RLC sample quality, but the results were inconsistent and often not reproducible (6.1.5). The diversity of purifications of the N-hAgo2-N-hDicer complex revealed unequal yields of this complex independent of used buffers. Thus it seems that the expression of N-hAgo2 and N-hDicer seems

## Discussion

to be highly sensitive to subtle changes in the transfection, temperature, virus and the expression time, within High Five and Sf9 insect cells. Furthermore, it was shown that these proteins of the miRNA pathway could be regulated by posttranslational modifications. Human Dicer might be phosphorylated at Tyr664 and Ser1016, whereas the consequences of these modifications on stability and conformation are not known and a detailed analysis has not been carried out so far (Rush *et al.*, 2005; Dephoure *et al.*, 2008). For the human Ago2 protein, various modifications have been detected. One is the hydroxylation of Pro700, which stabilizes Ago2 (Qi *et al.*, 2008). On the other hand, polyubiquitination of hAgo2 marks the protein for degradation by the proteasome and thus has a destabilizing effect (Rybak *et al.*, 2009). The first confirmation of a suspected phosphorylation of Ago2 came from the work of Zeng *et al.*, 2008. They showed that Ser387 is phosphorylated *in vivo* and that this modification is enhanced under stress and leads to P-body localization (Zeng *et al.*, 2008). In 2010, more putative phosphorylation sites in hAgo2 were detected, which have an influence on hAgo2s stability and RNA-binding ability (Rüdel *et al.*, 2010). Since insect cells are able to post-translationally modify proteins, it is possible that these modifications occur, but lead to inhomogeneous populations as detected for hTRBP2 (Chakravarthy *et al.*, 2010). In addition, overexpression of proteins is stressful for the cells and this could induce post-translational modifications that might mark the overexpressed proteins for degradation. Thus, different modification patterns of the overexpressed N-hDicer and N-hAgo2 proteins could be the reason for the different yields and these modifications can have influences on the protein stability itself. The post-translational modifications for N-hDicer was not analyzed in detail, only phosphorylation of hAgo2 was analyzed and a different phosphorylation pattern than the one observed in HEK 294 cells could be detected (Figure 54). Especially, Tyr529 phosphorylation could be detected, which was shown to influence RNA-binding of hAgo2 and may be important for the results obtained during antisense siRNA and siRNA duplex binding assays of the RLC (see Figure 14; Rüdel *et al.*, 2010). Finally, it could be that higher oligomeric RLC complexes are functional, because native endogen RISCs are larger in size.

### **7.1.5 siRNA-loading ability of the RLC**

The human RLC comprising hDicer, hAgo2 and hTRBP2 was shown to be able to dice long RNA precursors, load the substrate duplexes onto hAgo2, unwind these duplexes and mediate target RNA slicing (Gregory *et al.*, 2005; Maniataki & Mourelatos, 2005; MacRae *et al.*, 2008). However, later studies revealed that human RLC assembly is uncoupled from dicing in HeLa cells (Yoda *et al.*, 2010). The endonucleolytic activities of the recombinant N-hDicer and N-hAgo2 proteins could be shown and as the RLC mostly facilitates loading of hAgo2, the *in vitro* reconstituted RLC was analyzed for siRNA-binding and N-hAgo2 loading by SEC (Figure 12; Figure 14; 6.1.2.3). It could be shown that the RLC binds siRNA duplex1 and the RLC changes

elution from SEC upon siRNA duplex1 binding. When antisense siRNA1 or siRNA duplex2 were applied to the RLC, the siRNA guide or duplex are loaded onto N-hAgo2. Subsequently, the N-hAgo2-siRNA guide or siRNA duplex2 complexes disassemble from the N-hDicer-hTRBP2 complexes (Figure 14). Thus, the *in vitro* reconstituted RLC is functional in Ago2 loading with a guide or duplex siRNA and disassembles upon this loading (Figure 14; Maniataki & Mourelatos, 2005). A different behavior between duplex1 and duplex2 can be detected, both duplexes are bound by the RLC, but only siRNA duplex2 is loaded onto hAgo2 (Figure 14). The siRNA duplex1 has a lower affinity than the siRNA duplex2 to hAgo2 (personal communication Jochen Deckert, Roche Kulmbach). The so-called nucleotide binding loop within the MID domain of hAgo2 determines that U and A nucleotides are favored at the 5'-end of the guide strand (see Figure 54, Frank *et al.*, 2010). The siRNA duplex2, containing the favored 5'-uridine in the guide strand is incorporated into Ago2 within the RLC, whereas the siRNA duplex1 containing a 5'-cytosine within the guide strand is bound by the RLC, but most likely not loaded onto hAgo2 (no disassembly of N-hDicer-hTRBP2 and N-hAgo2-siRNA duplex complexes). The siRNA duplex1 contains the antisense siRNA1, which is loaded onto N-hAgo2, whereas the duplex is not and therefore it seems that single stranded RNAs are preferably loaded onto N-hAgo2. However, a part of the purified N-hAgo2 seems to be already loaded with small RNAs from insect cells (see 7.1.1) and this could probably explain why not all N-hAgo2 molecules are loaded with antisense siRNA1 or siRNA duplex2 (Figure 14 A and C). Furthermore, mass spectrometry analysis of the purified Ago2 protein from High Five insect cells revealed that the protein is phosphorylated at Thr526 and Tyr529, which may prevent proper binding of the 5'-end of the siRNAs within the MID domain of Ago2 (see Figure 54).

Additionally, central mismatches (siRNA has only perfect complementary base pairing) and ATP were shown to promote loading in HeLa cells and since the used siRNAs do not have these structural preferences and no ATP was added, this could also explain the imperfect loading of N-hAgo2 (Yoda *et al.*, 2010; 6.1.2.3). However, the same study also revealed, that recombinant hAgo2 does not need ATP for loading and that there are differences in ATP dependence of endogenous (ATP-dependent) and recombinant RISCs (ATP-independent) (Yoda *et al.*, 2010). The ATPase assay that was performed in this study also showed that no ATP is needed for binding of the RNAs nor for loading of some of these siRNAs (antisense siRNA1 and siRNA duplex2) onto N-hAgo2. Furthermore, these findings of ATP-independent Ago2 loading are in agreement with recent observations found for *C. elegans* and *D. melanogaster* Dicer. The authors revealed that Dicer hydrolyse ATP to ADP only in the presence of long dsRNA (= substrate RNA) and not in the presence of RNA products, which were used in this work (Welker *et al.*, 2011; Cenic *et al.*, 2011). In order to test if human Dicer behaves like its homologs, ATPase assays in the presence of siRNA substrates (long dsRNAs or pre-miRNAs) have to be carried out.

## Discussion

Finally, no unwinding of the siRNA duplexes<sup>2</sup> by N-hAgo2 can be detected. Structural preferences like seed or 3'-mid mismatches were shown to facilitate unwinding, whereas the siRNA duplex<sup>2</sup> is perfectly complementary (Yoda *et al.*, 2010; 5.1.13). Generally, human Ago2 endonucleolytically cleaves perfect complementary passenger strands, which is Mg<sup>2+</sup>-dependent and no Mg<sup>2+</sup>-ions were added to the reaction, thus no strand separation can be detected (Wang *et al.*, 2009 b). A binding of the RLC could be obtained, but because the affinity to hAgo2 is low this RNA is not transferred to hAgo2. Thus the siRNA is protected, and this supports the idea that the RISCs are also important for RNA-transfer and protection. Furthermore it could be that this siRNA duplex is handed over to another RLC. However, this hypothesis has not been tested so far. In addition, it is known that in case of miRNA, loading really occurs within the miRLC and when the unwinding of a miRNA duplex is blocked, a duplex should remain in the miRLC (Liu *et al.*, 2012). Thus it could also be that the unwinding of siRNA duplex<sup>1</sup> is blocked (see above) and that this duplex stays bound to the RLC because of that. However, it was not tested which of the RLC components is bound to siRNA duplex<sup>1</sup>. Due to the low affinity of human Dicer to its substrate RNAs (700 ± 100 nM) and the high affinity of hTRBP2 to small siRNA duplexes (0.24-0.29 nM) one can exclude that the siRNA is bound to Dicer (Chakravarthy *et al.*, 2010). However, in the RLC these affinities to small siRNA could be different. The affinity of hAgo2 to siRNAs is not known, thus either hTRBP2 or hAgo2 can bind the siRNA duplex<sup>1</sup>. A binding to hTRBP2 seems more likely, because loading onto hAgo2 leads to RLC disassembly. However, the RLC-siRNA duplex<sup>1</sup> was used for further studies, because the RLC binds this siRNA, the siRNA duplex increases RLC stability and does not induce disassembly of the RLC.

### **7.1.6 In vitro reconstitution of minimal hDicer-hTRBP2 complexes**

Initially, the hDicer-hTRBP2 interaction was demonstrated by immunoprecipitation using a Dicer antibody and subsequent mass spectrometry of the interacting proteins (Chendrimada *et al.*, 2005; Haase *et al.*, 2005). Additionally, full length human Dicer-TRBP2 complexes can be reconstituted *in vitro* and were subjected to structural analysis (Lau *et al.*, 2009; Chakravarthy *et al.*, 2010; Ma *et al.*, 2012; 6.2.1). Although, the human TRBP2 stabilizes Dicer (6.1.5.3; 6.2.1), the analysis of such a big 260 kDa complex by X-ray crystallography is challenging and was not successful to date. Using yeast-to-hybrid screens the hTRBP2 binding site could be localized to amino acids 261-431, which is the domain between the DExD/H and helicase domains, of human Dicer (Daniels *et al.*, 2009). Based on these yeast-two-hybrid screens, minimal hDicer-hTRBP2 complexes were prepared *in vitro* for analysis by means of X-ray crystallography (6.2.2). With this procedure pre-crystalline formations (spherulites) could be obtained (6.2.4), but no diffracting crystals could be produced to date. However, this initial crystallization hit can be used for further optimization using the DTB2-T1 complex, where no extensive optimization was carried out yet (Figure 33). The minimal hDicer-hTRBP2 complexes are homogeneous and

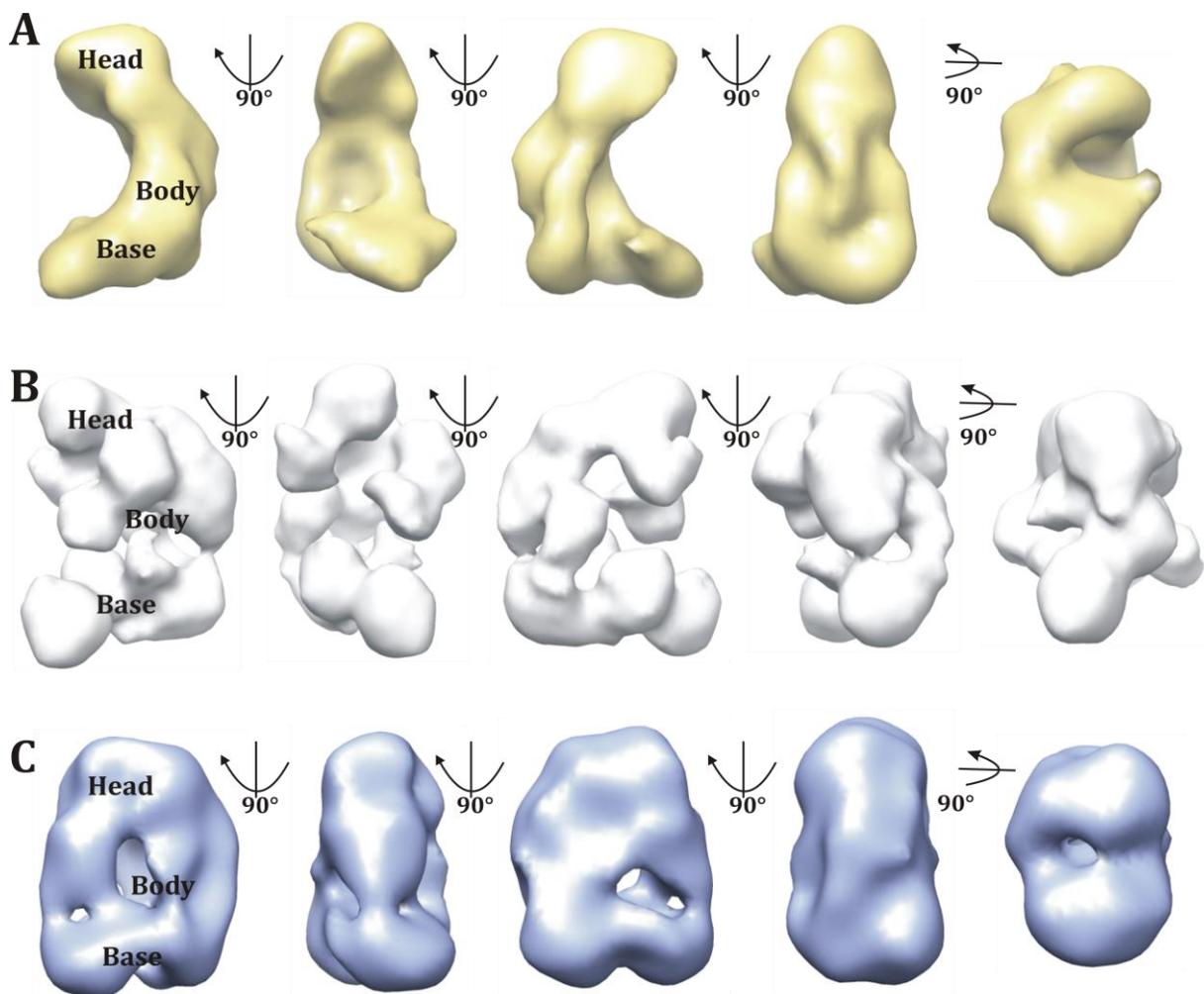
stable, but form heterotetramers (6.2.2). These heterotetramers seem not functional, because full length Dicer-TRBP complexes function as heterodimer (Lau *et al.*, 2009; Chakravarthy *et al.*, 2010; Ma *et al.*, 2012). To investigate the human Dicer interaction domain in more detail, human Dicer fragments containing amino acids 261-414 (DTB3) or amino acids 261-415 (DTB4) were used to prepare minimal hDicer-hTRBP2 complexes. Interestingly, no complex could be formed (see Figure 51), thus the amino acids 416-431 seem to be important for human Dicer interaction with hTRBP2 *in vitro*. However, an experiment showing that amino acids 416-431 of human Dicer are sufficient to bind to human TRBP2 has to be carried out in the future. The DTB1-T1 complex containing dsRBD2 and dsRBD3 of hTRBP2 was expected to bind single stranded and siRNA duplexes, but the DTB1-T1 complex can only incorporate siRNA duplexes (Figure 31 A). In contrast, the free T1 fragment is able to bind single stranded and double stranded siRNA (Figure 31 B). Hence, the DTB1 seems to prevent binding of single stranded siRNA by T1 and the hTRBP2-hDicer-interaction could influence hTRBP2 siRNA selection.

### **7.1.7 EM structure of human RLC**

The obtained 3D reconstruction of the RLC (22.6 Å resolution) was compared with another RLC reconstruction (33 Å resolution; Wang *et al.*, 2009b) and the hDicer-hTRBP2-complex reconstruction (20 Å resolution; Lau *et al.*, 2009; Lau *et al.*, 2012a) at different orientations (Figure 40). The EM reconstruction of the human Dicer-TRBP2 complex was calculated by Lau *et al.* 2009 in 2009 and was renamed as EM reconstruction of human Dicer by Lau *et al.* in 2012. Thus, it is unclear if this EM reconstruction corresponds to human Dicer alone or the Dicer-TRBP2 complex. Here this reconstruction is named as reconstruction of the hDicer-hTRBP2 complex. The calculated 3D reconstruction of the RLC obtained in this study adopts a C-shape with an elongation at the head facing to the base (Figure 40 B). In contrast, the RLC structure obtained by Wang *et al.*, 2009 has a triangular architecture. Here an agglomeration of density, which the authors identified to correspond to hAgo2, connects the head with the base of the RLC (Figure 40 C; Wang *et al.*, 2009). The obtained EM structure of the RLC of this study has a more open architecture and a connection between the head and the base branch is missing (Figure 40 B). Superposition of the two EM structures clearly shows that the two structures are similar, but the new structure has a way more open conformation (see Figure 55). Additionally, in the RLC reconstruction of this study the domains are more defined, whereas the other RLC reconstruction has a lower resolution where such details are not visible. The Dicer-TRBP2-EM reconstruction appears as a L-shaped structure that contains a vertical platform and a perpendicular base branch (Figure 40; Lau *et al.*, 2009; Lau *et al.*, 2012a). This L-shape structure most likely results from an initial L-shaped model that the authors used to reconstruct the EM map (Lau *et al.*, 2009). In contrast, the 3D-reconstruction of the RLC of this study was built by an *ab initio* reconstruction method and the other RLC structure by random conical tilt (RCT) method

## Discussion

followed by projection matching refinement (Wang *et al.*, 2009b). However, the two RLC 3D-reconstructions and the 3D-reconstruction of the hDicer-hTRBP2 complex superimpose quite well with each other. Both RLC structures are much bigger in size and the bases are oriented in a different conformation than the hDicer-hTRBP2 structure (see Figure 40). An agglomeration of density between the head and the base of the RLC structure obtained by Wang *et al.* is the main difference between this RLC structure and the hDicer-hTRBP2 structure (see also Figure 55). Comparison of the RLC structure of this study with the hDicer-hTRBP2 structure reveals an elongated base and extra density at the head region in the RLC structure (comparison of Figure 40 A and B, Figure 55).

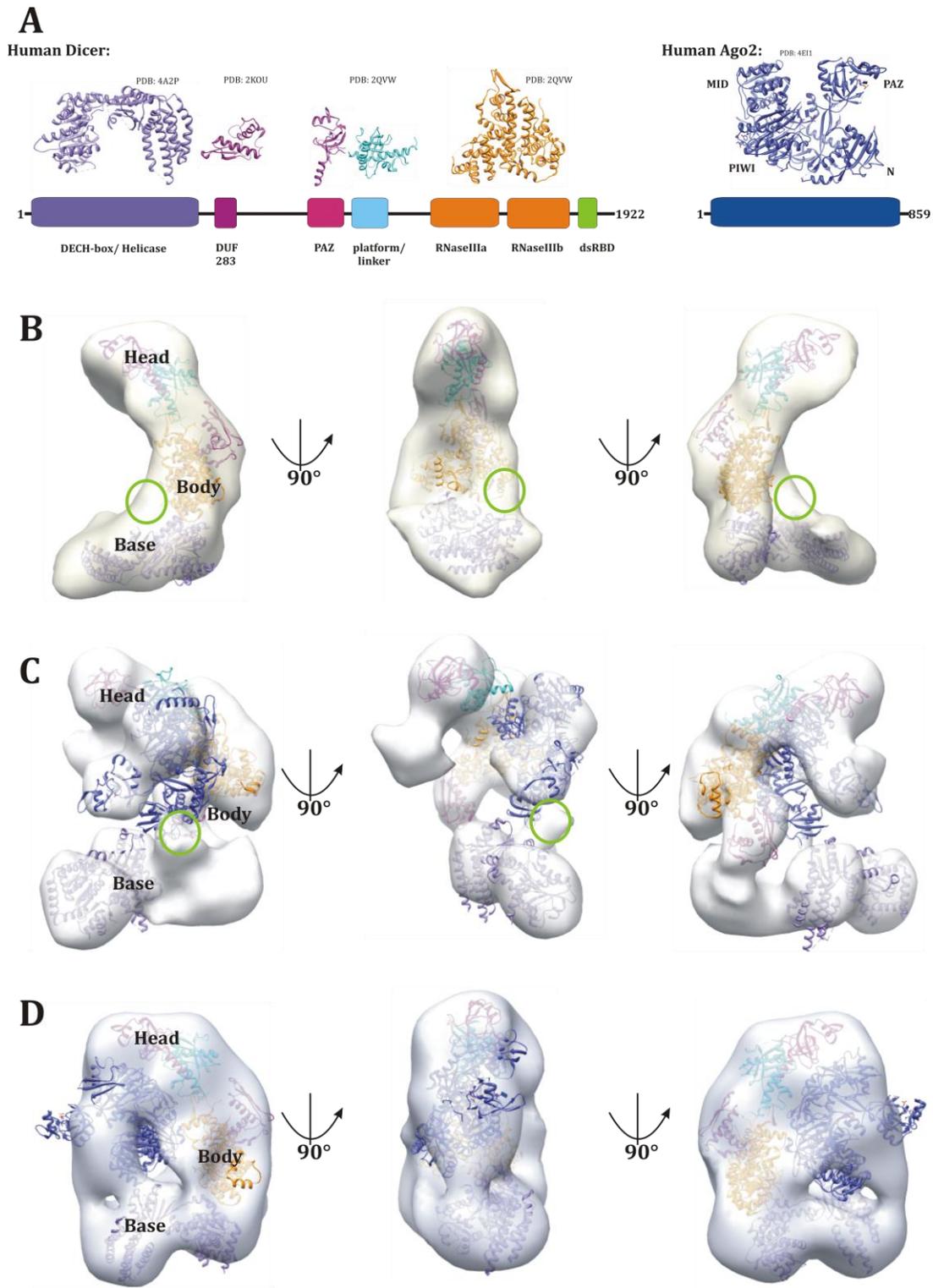


**Figure 40: Comparison of 3D reconstructions of the human Dicer-TRBP2 complex and human RLC complexes in different orientations.**

**A:** 3D reconstruction of human Dicer-TRBP2 complex (Dicer) at different orientations (EMDB-ID: 1646). **B:** 3D reconstruction of the human RLC calculated during this study at different orientations (6.1.5.2). **C:** 3D reconstruction of the human RLC at different orientations (Wang *et al.*, 2009b).

To map the different RLC components of the RLC structure, reference density maps of the subcomplexes were aimed during this study (6.1.5.3). To date no reliable 3D reconstructions of the RLC subcomplexes or hDicer could be obtained, because of the sample heterogeneity (Figure

25). However, class sums were calculated. Some of the class sums of N-hDicer (e.g. numbers 5 to 9 of Figure 26 A) and the N-hDicer-hTRBP2 complex (e.g. numbers 5, 11, 16, 19 of Figure 26) have high similarity with the class averages the previously calculated 3D structure of hDicer-hTRBP2 complex. It is hard to determine the localization of the single components of the RLC based on these class sums. Therefore, we used biochemical knowledge and the results observed by others, which calculated reference density maps or used a domain-mapping technique combined with negative-stain EM analysis, to map the different hDicer domains into the hDicer-hTRBP2-complex and RLC EM-maps (Wang *et al.*, 2009b; Lau *et al.*, 2009; Lau *et al.*, 2012a; Lau *et al.*, 2012b). Using calculations of reference maps the, the RIG-I like DExD-helicase domain of hDicer could be located within the base branch of the hDicer-hTRBP2 complex or of the previous RLC (see Figure 41 B; Wang *et al.*, 2009b; Lau *et al.*, 2009). A new mapping technique, which is called DOLORS (domain localization in electron microscopy), clearly identified the hDicer PAZ domain to be positioned in the head, that the RNase III heterodimer is positioned in the body, whereas the DUF 238 domain is located close to the RNase domains in the body of the hDicer-hTRBP2 reconstruction (Figure 41 B). These findings contradict earlier speculations about the positioning of the hDicer domains within the EM maps, where it was hypothesized that the RNase domains are located in the head and that the PAZ domain is located in the body region (see Figure 6; Wang *et al.*, 2009b; Lau *et al.*, 2009). The positioning of the catalytic core adjacent to RIG I like DExD-helicase domain is in good agreement with established biochemical properties. The close spatial arrangement of the DExD-helicase domain to the RNase domains, can easily explain an inhibition by DExD-helicase domain and activation by hTRBP2 on hDicers catalytic activity that is mediated by the RNase heterodimer (Haase *et al.*, 2005; Ma *et al.*, 2008; Chakravarthy *et al.*, 2010). Fitting of the structure of *G. intestinalis* Dicer, the *H. sapiens* RIG-I domain and the *A. thaliana* DUF-238 domain into the 3D reconstruction of hDicer-hTRBP2 complex (Lau *et al.*, 2009; Lau *et al.*, 2012a) leaves only one free space for hTRBP2, which is located adjacent to the RNase domains (Figure 41 B). However, there is no free density for the missing structural domains left and thus, the reconstruction seems too small. This observation that the reconstruction seems to small, is in agreement with another 3D reconstruction of hDicer, which had an L-shaped structure, but more density at the head and the base regions and thus would be able to fit all missing structural elements (Wang *et al.*, 2009b). Based on these results and superposition with the hDicer-hTRBP2 reconstruction, the single domains of hDicer were fitted into the new and the old RLC maps (Figure 41). Additionally, a hDicer fragment containing the RNaseIIIa, RNaseIIIb heterodimer and the dsRBD was identified to bind to the PIWI lobe of Ago (Sasaki & Shimizu, 2007; Tahbaz *et al.*, 2004). Therefore, the structure of hAgo2 was fitted into the new and the old RLC maps in an orientation that the PIWI domain of hAgo2 can interact with hDicers RNase domains. Additionally, when hAgo2 is oriented that way a bound guide strand is free to complementary bind to an mRNA target (Figure 41).



**Figure 41: Comparison of Dicer-TRBP2 and RLC reconstructions.**

**A:** Domain organization and structures of Dicer domains and structure of human Ago2. **B:** 3D reconstruction of human Dicer-TRBP2 complex (Dicer) at different orientations (EMDB-ID: 1646). **C:** 3D reconstruction of the human RLC calculated during this study at different orientations (6.1.5.2). **D:** 3D reconstruction of the human RLC at different orientations (Wang *et al.*, 2009b). B-D: Docking of the atomic models of the PAZ, the Platform and the RNase III domains of *G. intestinalis* Dicer (PDB-ID: 2QVW), the *H. sapiens* RIG-I domain (purple, PDB-ID: 4A2P), the *A. thaliana* DUF238 domain (PDB-ID: 2KOU) and the *H. sapiens* Ago2 (PDB-ID: 4E11) into the density maps. The head, platform and base of each reconstruction are indicated. A possible location of hTRBP2 is indicated by green circles.

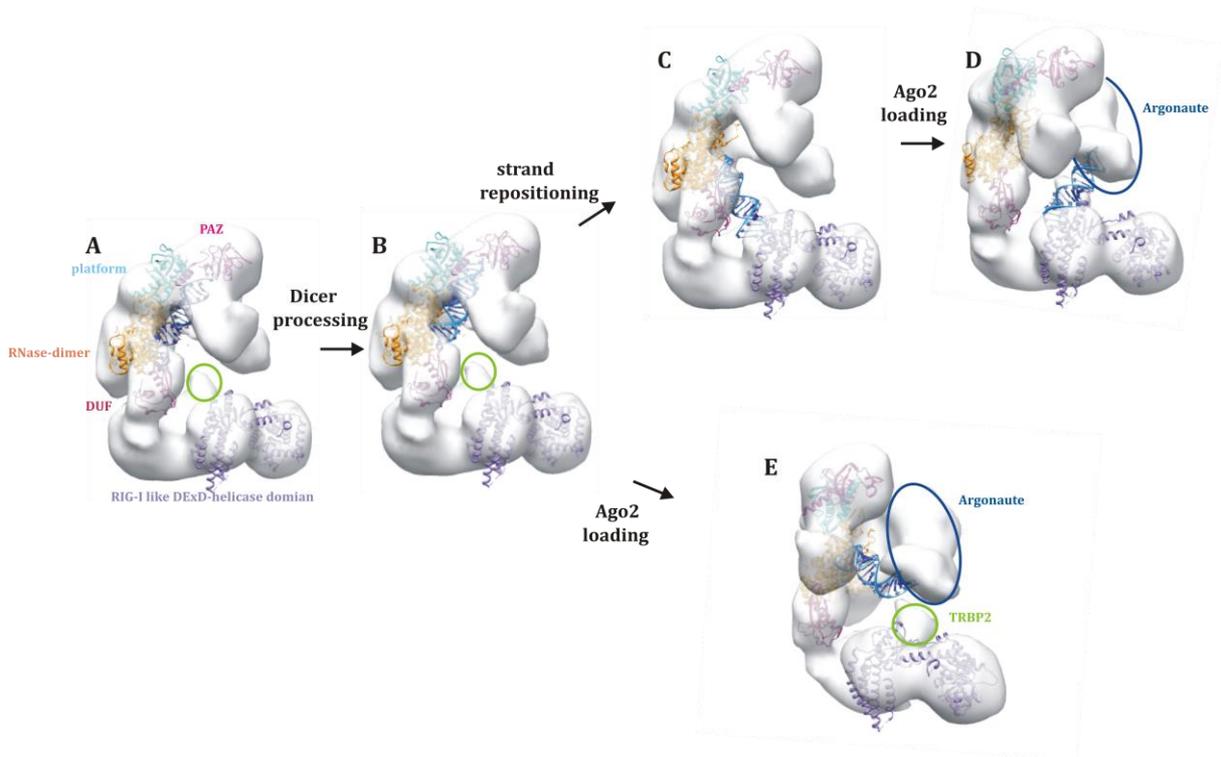
As mentioned above, the new and the old RLC reconstructions are quite similar, but the new one resembles a far more open conformation and individual domains can be separated. All structural elements can be fitted in both maps, leaving space for the missing hDicer domains and hTRBP2. The possible location of hTRBP2 in the new RLC reconstruction is indicated by a green circle (Figure 41). This density was visible in some reconstructions, but missing in others. We determined that hTRBP2 is incorporated into the RLC stoichiometrically and therefore we assumed that hTRBP2 is localized here (Figure 20; Figure 41 C).

The new positioning of Dicer domains within the reconstructions necessitates a new model for the RISC-loading processes (see also section 4.3.2; Wang *et al.*, 2009b). The known structural elements of hDicer were positioned into the new RLC reconstruction as shown in Figure 41 and hTRBP2 and hAgo2 are indicated by green and blue circles, respectively (Figure 42). Although, the dicing process seems to be uncoupled from the loading process, the RLC is able to cleave a pre-miRNA (4.2.1; 4.3.1; MacRae *et al.*, 2008). The 2-nt 3'-overhang of the pre-miRNA is bound by the PAZ domain, that is oriented to the RNaseIII heterodimer in a spatial arrangement that a about 22-nt small RNA can be produced (Figure 42 A). Thereby a small RNA having two 2-nt 3'-overhangs is produced (Figure 42 B). The small RNA-product can be directly transferred to Ago2's PAZ domain or another domain of hAgo2 (Figure 42 E). When the small RNA has to be repositioned for correct loading, the small RNA product is first transferred to the DEXD-helicase domain of hDicer and can then be loaded onto hAgo2 (Figure 42 C and D). Additionally, it seems that hAgo2 also has to undergo a repositioning of the PAZ domain in order to get loaded with a small RNA or hAgo2 binds to the RLC in a different fashion than proposed earlier (Figure 41). Using this model an involvement of hTRBP2 (marked with green ring) on hDicer processing, hDicer stability, RNA repositioning and transfer to hAgo2 is easily imaginable (Haase *et al.*, 2005; Chendrimada *et al.*, 2005; Chakravarthy *et al.*, 2010; Noland *et al.*, 2012). Moreover, it seems that the main transfer processes occur in the body of the RLC. We could detect heterogeneity or flexibility in the flexible arm regions in the head (Figure 24). The fitting identified that one arm seems to contain hDicer's PAZ domain as well as an additional unknown domain of hDicer (probably the C-terminal dsRBD) and that the second flexible arm contains hAgo2 (6.1.5.2; Figure 24). Thus these hDicer domains and hAgo2 can engage different conformations within the RLC. In order to address the question whether the observed effect is caused by conformational flexibility or is a heterogeneity artefact, extensive computational sorting is necessary that has not been finished so far.

In summary, the presented EM reconstruction of the human RLC allow to obtain a more detailed information about the structure and localization of the individual components of the RLC. The newly proposed model for the RISC-loading process is in agreement with actual biochemical data as well as structural results and presents a mechanism for the small RNA duplex transfer onto human Ago2. However, new reference density maps of the RLC

## Discussion

subcomplexes and human Dicer have to be calculated to confirm the proposed model. Additionally, flexibility or heterogeneity that could be detected in EM analysis has to be investigated in more detail in order to excise individual structures of the RLC in different conformations to fully detect and understand the single steps of the RISC-loading processes.



**Figure 42: Model for human RISC-loading processes.**

Docking of the atomic models of the PAZ (pink), the Platform (cyan) and the RNase III domains (orange) of *G. intestinalis* Dicer (PDB-ID: 2QVW), the *H. sapiens* RIG-I domain (purple, PDB-ID: 4A2P), the *A. thaliana* DUF238 domain (dark-pink, PDB-ID:2KOU) and a pre-miRNA (dark-blue, PDB-ID: 3A6P) or siRNA duplex (PDB-ID: 1R9F) into the electron density maps. **A:** Dicer processing of a pre-miRNA. The 2-nt 3'-overhang is bound to the appropriate binding pocket in PAZ domain of hDicer and the pre-miRNA oriented, that the catalytic center generates a 22-nt long product. **B:** A hDicer-product containing RLC. **C:** Repositioning of the small RNA from the RNase III domains to the DEXD-helicase domain. **D** and **E:** Human Ago2-loading either from hDicers DEXD-helicase domain (D) or directly after hDicer processing (E).

## 7.2 Crystal structure of dsRBD2 of human TRBP2

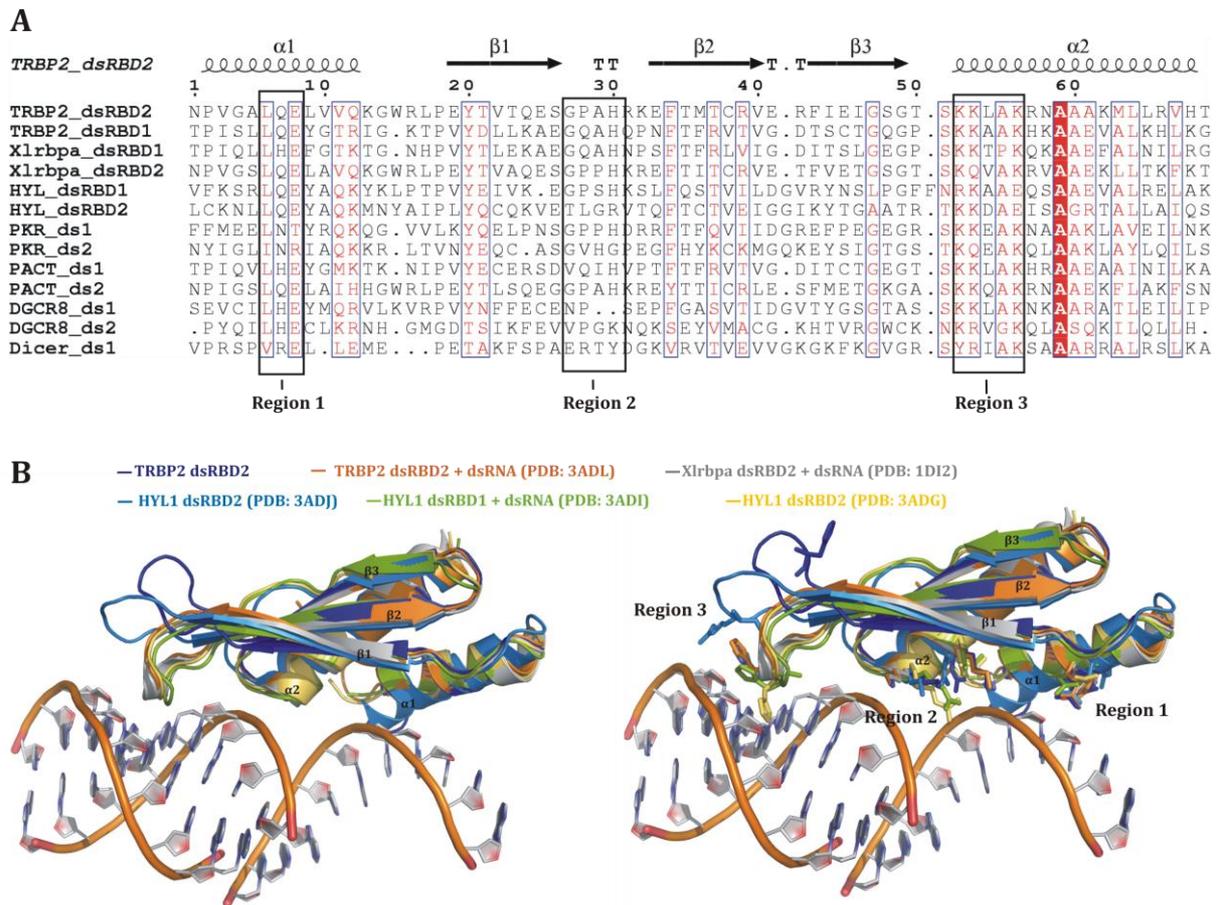
In order to investigate the structure of consecutive dsRBDs, the full length hTRBP2 protein with and without RNA was subjected to crystallization trials (6.1.4). However, this approach failed to produce protein crystals and therefore hTRBP2 fragments were prepared. These fragments either contained the second and the third dsRBDs (dsRBD2 and dsRBD3) or the dsRBD3 of hTRBP2 (6.3). Initial crystallization of these fragments in the presence or absence of single stranded or double stranded siRNAs failed (6.3). The method of *in situ* proteolysis led to diffracting crystals of the T1 fragment (aa157-366) containing dsRBD2 and dsRBD3 of hTRBP2 (6.3.1, Dong *et al.*, 2007). The crystal structure revealed that the asymmetric unit of the crystal is composed of four monomers of the dsRBD2 and does not contain the dsRBD3 of hTRBP2. Thus, the Chymotrypsin that has been added to the preparation prior to crystallization, quantitatively cleaved of the C-terminal dsRBD3 of hTRBP2 (6.3.1). In the future, other proteases that do not cleave of the C-terminus should be used. However, it is not known if the dsRBD3 of hTRBP2 will crystallize using other proteases. Unfolding experiments identified that the dsRBD2 of hTRBP2 is more thermo stable than dsRBD1 (Yamashita *et al.*, 2011). A higher thermostability of the dsRBD2 could also be the reason why this domain crystallized, while the dsRBD3 did not crystallize.

### 7.2.1 dsRBD2 of hTRBP2 has a canonical dsRBD fold that has high flexibility in the RNA-binding regions

The crystal structure of the second dsRBD (dsRBD2) of hTRBP2 comprising amino acids 157-226 could be solved at 2.28 Å resolution (6.3.1). The asymmetric unit of the hexagonal-shaped crystals is occupied by four dsRBD2 monomers, which superimpose well (6.3.2; Figure 38). Each monomer shows a canonical dsRBD ( $\alpha$ - $\beta$ - $\beta$ - $\beta$ - $\alpha$ ) fold, where the two  $\alpha$ -helices pack against one side of the three-stranded antiparallel  $\beta$ -sheet (Figure 38; Kharrat *et al.*, 1995; Bycroft *et al.*, 1995). In order to compare the structure with other dsRBDs a sequence alignment and a superposition was carried out (Figure 43). Several amino acids of diverse dsRBDs are conserved along the entire domain and the most conserved part is located in the C-terminal part of the domain (Figure 43 A). Using mutagenesis analysis, three different regions that are important for dsRNA-binding were uncovered. These regions are highly conserved among diverse dsRBDs. Region 1 was defined to contain a highly conserved Glu (E) in the  $\alpha$ 1-helix, the region 2 contains the GPxH motif in the loop between the  $\beta$ 1- and  $\beta$ 2-strands and region 3 contains the KKxAK motif in the  $\alpha$ 2-helix (Figure 43; Green & Mathews 1992; Bycroft *et al.*, 1995; Ramos *et al.*, 2000; Yang *et al.*, 2010; Ryter & Schultz, 1998; Green *et al.*, 1995; McMillan *et al.*, 1995; Patel *et al.*, 1996). Differences from that consensus can be detected in the dsRBD3 of hTRBP or hPACT. Because of those differences in amino acid sequence which results in reduced

## Discussion

RNA-binding affinity, these dsRBDs are referred to as type B dsRBDs (St Johnston *et al.*, 1992; Krovat & Jantsch, 1996; Yamashita *et al.*, 2011).

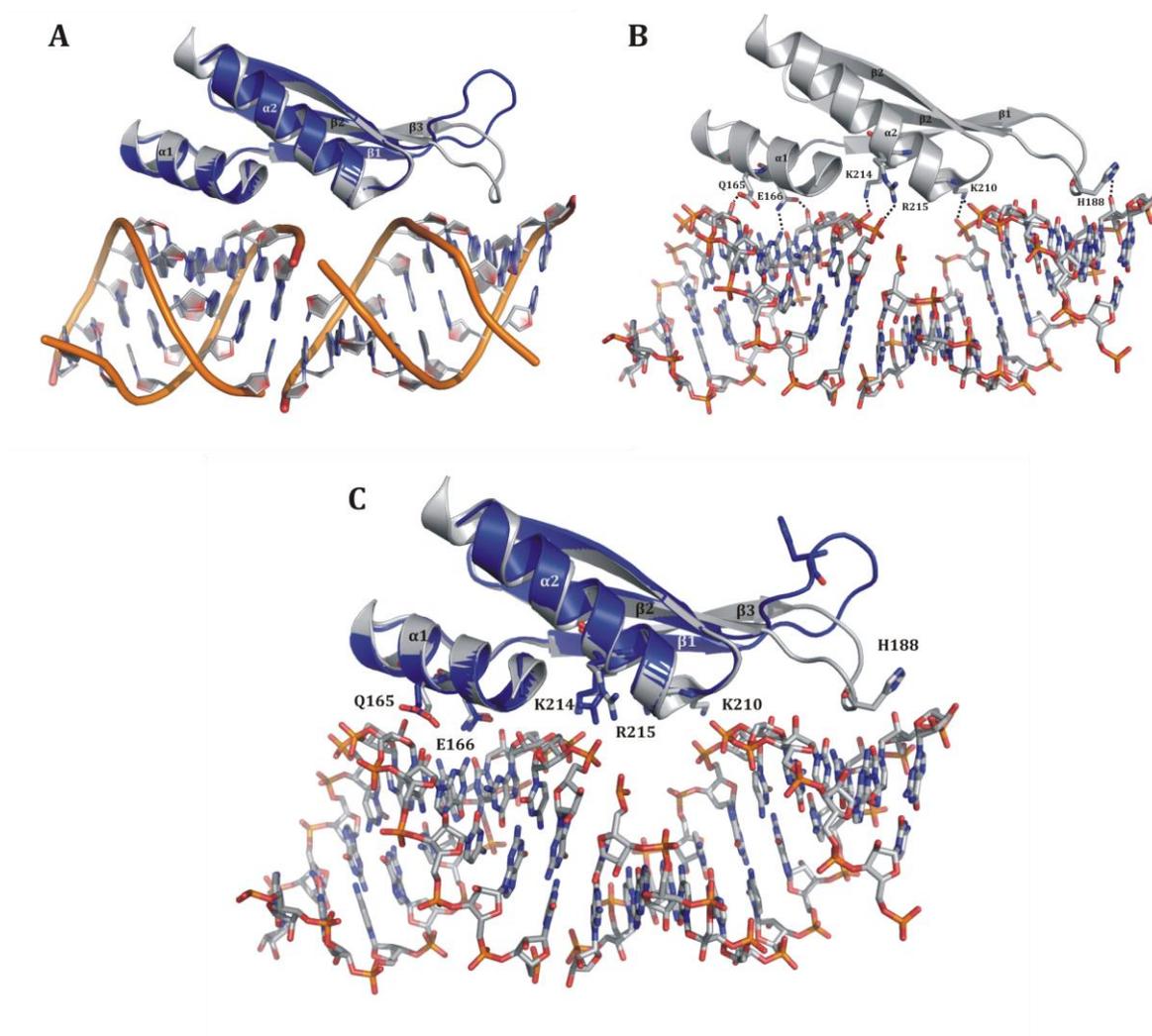


**Figure 43: Sequence alignment of various dsRBDs and superposition of the crystal structure of dsRBD2 of hTRBP2 with other dsRBDs.** **A:** Sequence alignment of different dsRBDs with the secondary structure of the dsRBD2 of hTRBP2 on top. Complete conserved residues are colored in white with a red background. Conserved amino acids are colored in red with white background. The three regions important for RNA-binding are marked. **B:** Superposition of the crystal structures of dsRBD2 of hTRBP2 with other dsRBDs. In the right part of the picture the amino acids that were shown to be important for dsRNA-binding of hTRBP2 are shown in all structures as sticks.

Apart from the importance for dsRNA-binding, other conserved hydrophobic residues were shown to be important for stabilization of the overall domain. These are mainly aliphatic, but also aromatic residues equally distributed over the domain (Krovat & Jantsch 1996; Bycroft *et al.*, 1995, Ramos *et al.*, 2000; Ryter & Schultz 1998). A superposition of several dsRBDs shows that especially the loop region between the  $\beta$ 1- and  $\beta$ 2-strands (RNA-binding region 2) can vary (Figure 43). When comparing all residues that were shown to be important for dsRNA-binding in hTRBP2, several differences in the orientation of these conserved residues can be obtained (Figure 43). Interestingly, one of the structures shown in the superposition of Figure 43 is the crystal structure of the dsRBD2 of hTRBP2 in complex with a short CG-duplex (PDB-ID: 3ADL). Single superposition of the two structures revealed that the overall structures of the dsRBD2 and the dsRBD2 when bound to RNA (3ADL) are similar (r.m.s.d. = 1.14 Å for 70 amino acids),

however especially the RNA-binding region 2, containing His188 which is important for RNA-binding, varies (Figure 44). When bound to RNA the loop faces the RNA, while in the absence of RNA the loop is oriented contrarily (Figure 44 A). In addition to this loop other residues that are important for RNA-binding have a different conformation, when or when not bound to RNA. The structure of the dsRBD2 of hTRBP2 does not allow His188 and Lys210 to interact with the RNA, because these residues are oriented away from the dsRNA (Figure 44 C). All the other residues that were shown to be important for dsRNA-binding in TRBP2 (Gln165, Glu166, Lys214 and Arg215) would be able to interact with the dsRNA, although if differently oriented.

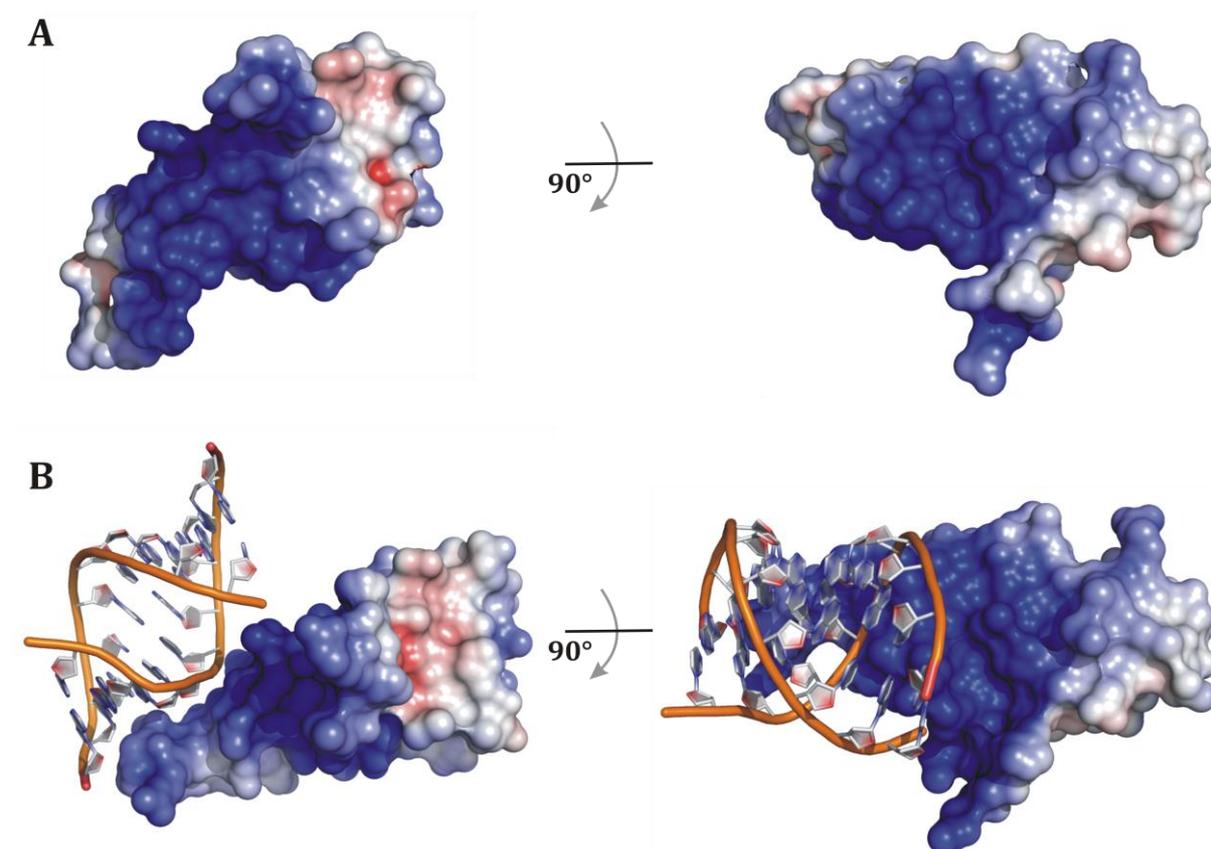
It has been shown that hTRBP2 only binds dsRNA, but not dsDNA (Gredell *et al.*, 2010). The dsRBD2 binds to the phosphate backbone and 2' OH groups of the ribose, which are not present in DNA (Figure 44). Especially, residues of the RNA-binding regions 1 and 2 interact with the 2' OH groups. For example, Glu166 makes hydrogen bonds with the 2'OH group of the RNA. Thus, hTRBP2 is able to discriminate between dsRNA and dsDNA. Furthermore, hTRBP2 was shown to preferably bind to CG-rich dsRNA sequences (Gatignol *et al.*, 1991; Lee *et al.*, 1996) and it was proposed that hTRBP2 senses the asymmetry of an siRNA and therefore must be able to sequence specific recognizing small (Gredell *et al.*, 2010). Double stranded RNA adopts an A-form RNA helix form containing a wide and flat minor groove and a tight and deep major groove (Saenger, 1984; Dickerson *et al.*, 1982; see Figure 44). The bases of the RNA are accessible in the minor groove, but not in the major groove. Residues from the RNA-binding region 1 and 2 of dsRBD2 of hTRBP2 could be able to recognize the bases of a dsRNA. The side chain of Glu165 of hTRBP2 were indeed shown to interact with the O2 atom of guanine or the N3 atom of cytosine via hydrogen bonds (Figure 44). In contrast, only hydrogen bonding of His188 to the ribose could be determined for the second dsRBD2 of hTRBP2 (Figure 44). Residues from the RNA-binding region 3 recognize the major groove of potentially bound dsRNA. Thus, hTRBP2 is able to specifically distinguish between dsRNA and dsDNA via direct interactions with the characteristic dsRNA regions. Furthermore, using residues of the RNA-binding region 1 hTRBP2 recognize CG-sequences and can thus mediate its CG-specificity for dsRNA (Yang *et al.*, 2010).



**Figure 44: Crystal structure of the second dsRBD of hTRBP2 reveals that hTRBP2 can undergo structural rearrangements upon RNA-binding.**

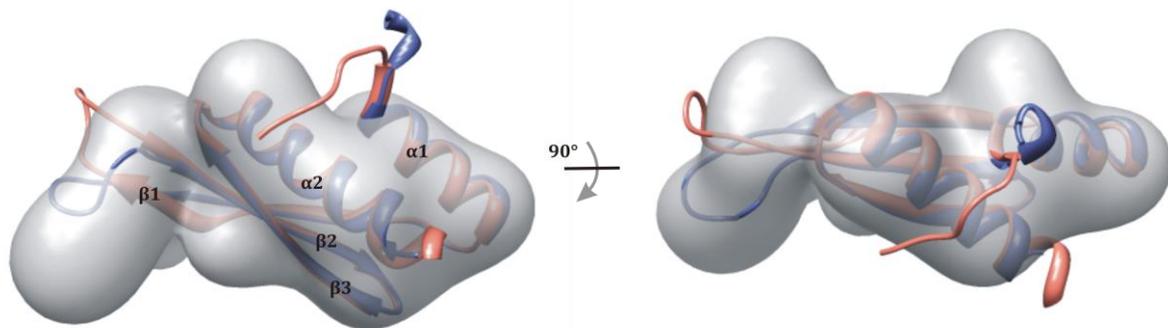
**A. B:** Comparison of the crystal structures of dsRBD2 of hTRBP2 (colored in blue) and dsRBD2 of hTRBP2 in complex with 10-nt GC-duplex (colored in grey, PDB-ID: 3ADL). Especially the loop between  $\beta 1$  and  $\beta 2$ -strands is oriented differently. **C and D:** Superposition of monomer C of dsRBD2 of hTRBP2 (colored in blue) with the dsRBD2 of hTRBP2 in complex with RNA (colored in grey, PDB-ID: 3ADL) with all residues important for RNA-binding shown in stick mode (colored according to molecule). The residues important for RNA-binding are oriented differently depending on whether RNA is bound or not.

The electrostatic surface potential of the crystal structure of the dsRBD2 and the crystal structure of the dsRBD2 of hTRBP2 in complex with a short CG-duplex are mainly similar. The overall electrostatic surface potentials are mainly basic with the most basic parts in the RNA-binding regions (Figure 45). The electrostatic surface potentials would allow the dsRNA also to bind the dsRNA in a different fashion, namely perpendicular to the molecule. If the dsRNA would bind perpendicular, then not all the residues that were shown to be important for RNA-binding could interact. Furthermore the dsRNA-interaction interface with the dsRBD2 of hTRBP is similar to the interaction sides of other dsRBD2 (Yang *et al.*, 2010). As the interaction side to dsRNA is a canonical and conserved characteristic of dsRBDs, it is most likely that the dsRNA binds to dsRBD2 of TRBP2 as shown in Figure 43-45.



**Figure 45: Comparison of the electrostatic surface potentials of TRBP2's dsRBD2.** **A:** Electrostatic surface potential of the dsRNA-binding domain of hTRBP2 in two different orientations (at  $\pm 5$  kcal/electron units, +5= blue, -5 = red) **B:** Electrostatic surface potential of the dsRNA-binding domain of hTRBP2 in complex with a 10-nt GC-RNA in two different orientations (at  $\pm 5$  kcal/electron units, +5= blue, -5 = red; PDB-ID: 3ADL). The RNA is shown in cartoon representation.

In summary, comparison of the structure of dsRBD2 of TRBP2 with the previously solved structure of the dsRBD2 in complex with a short CG-duplex reveals that some residues that are important for dsRNA-binding can undergo significant structural rearrangements upon RNA-binding (Figure 43). Additionally, the SAXS structure of the dsRBD2 uncover that the regions 1 and 2 important for RNA-binding are highly flexible in solution allowing the binding of various RNA substrates (Figure 39).



**Figure 46: The SAXS-damfilt models of the dsRBD2 of TRBP2 indicate a high flexibility of the dsRNA-binding regions 1 and 2.** Docking of the crystal structure of the dsRBD2 of hTRBP2 (colored in blue) and the crystal structure of the second dsRBD2 of hTRBP2 in complex with RNA, whereas the RNA was removed (PDB-ID: 3ADL) into a SAXS damfilt model.

### **7.2.2 The crystal structure of dsRBD2 of hTRBP2 reveals two possible dimerization interfaces**

Apart from their function in dsRNA-binding, several dsRBDs are important for protein-protein interactions. Here, the dsRBDs can bind to other dsRBDs or other domains (Huang *et al.*, 2009; Patel *et al.*, 1998; Schuldt *et al.*, 1998; Micklem *et al.*, 2000; Zhang *et al.*, 2001; Tremblay *et al.*, 2002; Gupta *et al.*, 2003; Hitti *et al.*, 2004; Haase *et al.*, 2005; Chendrimada *et al.*, 2005). Moreover, some dsRBPs were shown to form homodimers with their dsRBDs (Costinnetto *et al.*, 1995; Lararki *et al.*, 2008). The linker regions between two consecutive dsRBDs are flexible and therefore crystal structures of more than one dsRBD are rare. However, when looking at symmetry mates of already crystallized dsRBDs, prominent contacts between the  $\beta$ 1- and the  $\beta$ 3-strands of two dsRBDs can be detected. Such contacts can also be determined between monomer A and D of the dsRBD2 crystal structure (Figure 38 D). Moreover, the crystal structure of the dsRBD2-homodimer of *A.thaliana* HYL1 also uncovers this putative dimerization interface and it has been concluded that it is no crystallization artefact. However, validation by mutagenesis studies failed, because the mutations led to mainly misfolded protein (Yang *et al.*, 2010). Thus, it is still unclear if this is a functional relevant dimerization interface. The available crystal structure of the DGCR8 core containing two different dsRBDs gives the only structural insight into the interaction surface of two dsRBDs in the context of an intact protein. Here, different regions (before and after the conserved dsRBDs motif) mediate the dimerization (Yang *et al.*, 2010). This is in good agreement with the results obtained here. Although, the asymmetric unit

contains four molecules of the dsRBD2 which form two possible dimerization interfaces, this domain in solution is monomeric as determined by MALS and SAXS (Figure 36, Figure 39 and Figure 46). In contrast, both the full length hTRBP2 protein and the T1 fragment containing three and two dsRBDs, respectively exist as homodimer in solution (Figure 8, Figure 36). This led to the conclusion that the dsRBD2 seems not directly to be involved in dimerization and suggests that the loop regions between the dsRBDs of hTRBP2 are important for dimerization *in vitro*. Thus, most likely the identified dimerization interfaces are only crystallization artifacts and not functionally relevant. Additionally, some other groups, localized the dimerization of the dsRBD2 of hTRBP2 using yeast-two hybrid screens. Here, dimerization could be detected when constructs containing amino acids 152-234 or 152-224 of hTRBP2 are used (Laraki *et al.*, 2008, Daher *et al.*, 2001). As no dimerization can be detected when using amino acids 157-226, the region comprising amino acids 152-159 or 226-234 seems critical for dimerization. Further experiments using different constructs of hTRBP2 are necessary to validate this hypothesis, however the results presented here, show that not the conserved dsRBD itself mediate dimerization, but that rather the loop regions between these domains are important for such interactions.

In the future it would be important to understand how dsRBDs interact with each other and how these interactions collaborate in dsRNA-binding (affinity, specificity, cooperativity) and transfer.

Discussion

## 8 References

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D66*, 213-221
- Ausubel, F.M., et al., *Current Protocols in Molecular Biology*, vol. 1, John Wiley & Sons, Inc., Brooklyn, New York, 3.13.1, 1994-2005.
- Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P., and Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* 22, 2773–2785.
- Baek, D., Villén, J., Shin, C., Camargo, F.D., Gygi, S.P., and Bartel, D.P. (2008). The impact of microRNAs on protein output. *Nature* 455, 64–71.
- Bannwarth, S., Talakoub, L., Letourneur, F., Duarte, M., Purcell, D.F., Hiscott, J., and Gatignol, A. (2001). Organization of the human tarbp2 gene reveals two promoters that are repressed in an astrocytic cell line. *J Biol Chem* 276, 48803–48813.
- Barber, G.N. (2009). The NFAR's (nuclear factors associated with dsRNA): evolutionarily conserved members of the dsRNA-binding protein family. *RNA Biol* 6, 35–39.
- Barraud, P., Emmerth, S., Shimada, Y., Hotz, H.-R., Allain, F.H.-T., and Buhler, M. (2011). An extended dsRBD with a novel zinc-binding motif mediates nuclear retention of fission yeast Dicer. *Embo J* 30, 4223–4235.
- Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297.
- Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233.
- Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev* 20, 1885–1898.
- Bentwich I, Avniel A, Karov Y, Aharonov R, Gilad S, Barad O, Barzilai A, Einat P, Einav U, Meiri E, Sharon E, Spector Y, Bentwich Z (2005). "Identification of hundreds of conserved and nonconserved human microRNAs". *Nat. Genet.* 37 (7), 766–70.
- Berezikov, E., Chung, W.-J., Willis, J., Cuppen, E., and Lai, E.C. (2007). Mammalian mirtron genes. *Mol Cell* 28, 328–336.
- Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363–366.
- Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. (2003). Dicer is essential for mouse development. *Nat Genet* 35, 215–217.
- Betancur, J.G., and Tomari, Y. (2012). Dicer is dispensable for asymmetric RISC-loading in mammals. *Rna* 18, 24–30.
- Bieniossek C., Richmond, T.J. Berger, I (2008) MultiBac: multigene baculovirus-based eukaryotic

## References

protein complex production. *Curr Protoc Protein Sci.* 5, 5-20.

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

Bogerd, H.P., Karnowski, H.W., Cai, X., Shin, J., Pohlert, M., and Cullen, B.R. (2010). A mammalian herpesvirus uses noncanonical expression and processing mechanisms to generate viral MicroRNAs. *Mol Cell* 37, 135–142.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. *Embo J* 17, 170–180.

Bohnsack, M.T., Czaplinski, K., and Görlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *Rna* 10, 185–191.

Boland, A., Tritschler, F., Heimst Auml Dt, S., Izaurrealde, E., and Weichenrieder, O. (2010). Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. *EMBO Rep* 11, 522–527.

Borchert GM, Lanier W, Davidson BL. (2006) RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol*,13, 1097–1101.

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

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, 72, 248-254.

Brameier, M., Herwig, A., Reinhardt, R., Walter, L., and Gruber, J. (2011). Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. *Nucleic Acids Res* 39, 675–686.

Braun, J.E., Huntzinger, E., Fauser, M., and Izaurrealde, E. (2011). GW182 Proteins Directly Recruit Cytoplasmic Deadenylase Complexes to miRNA Targets. *Mol Cell* 44, 120–133.

Bryson, K., McGuffin, L.J., Marsden, R.L., Ward, J.J., Sodhi, J.S., and Jones, D.T. (2005). Protein structure prediction servers at University College London. *Nucleic Acids Res* 33, W36–W38.

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

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

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

Castanotto, D., Sakurai, K., Lingeman, R., Li, H., Shively, L., Aagaard, L., Soifer, H., Gagnol, A., Riggs, A., and Rossi, J.J. (2007). Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. *Nucleic Acids Res* 35, 5154–5164.

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

Mol Cell 42, 172–184.

Chakravarthy, S., Sternberg, S.H., Kellenberger, C., and Doudna, J.A. (2010). Substrate-Specific Kinetics of Dicer-Catalyzed RNA Processing. *J Mol Biol*, 1–11.

Chekulaeva, M., and Filipowicz, W. (2009). Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Current Opinion in Cell Biology* 21, 452–460.

Chekulaeva, M. et al. (2011). miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nature Struct. Mol. Biol.* 18, 1218–1226

Cheloufi, S., Santos, Dos, C.O., Chong, M.M.W., and Hannon, G.J. (2010). A Dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465, 584–589.

Chen, C.-Y.A., Zheng, D., Xia, Z., and Shyu, A.-B. (2009). Ago–TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. *Nat Struct Mol Biol* 16, 1160–1166.

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

Christensen HS, Daher A, Soye KJ, Frankel LB, Alexander MR, Laine S, Bannwarth S, Ong CL, Chung SW, Campbell SM, Purcell DF, Gatig- nol A: Small interfering RNAs against the TAR RNA-binding protein, TRBP, a Dicer cofactor, inhibit human immunodeficiency virus type 1 long terminal repeat expression and viral production. *J Virol* 2007, 81(10), 5121-5131.

Chu, C.-Y., and Rana, T.M. (2006). Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biology* 4, e210.

Chung, W.-J., Okamura, K., Martin, R., and Lai, E.C. (2008). Endogenous RNA Interference Provides a Somatic Defense against *Drosophila* Transposons. *Current Biology* 18, 795–802.

Cifuentes, D., Xue, H., Taylor, D.W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G.J., Lawson, N.D., et al. (2010). A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* 328, 1694–1698.

Cogoni C. and Macino G. (2000). Post-transcriptional gene silencing across kingdoms. *Curr Opin Genet Dev* 10:638–643. *Opin Genet Dev* 10, 638–643.

Cole, C., Sobala, A., Lu, C., Thatcher, S.R., Bowman, A., Brown, J.W.S., Green, P.J., Barton, G.J., and Hutvagner, G. (2009). Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *Rna* 15, 2147–2160.

Covey, S, Al-Kaff, N. Lángara, A and Turner, D (1997). Plants combat infection by gene silencing. *Nature* 385, 781-782.

Creamer, K.M., and Partridge, J.F. (2011). RITS-connecting transcription, RNA interference, and heterochromatin assembly in fission yeast. *WIREs RNA* 2, 632–646.

Cuellar, T.L., Davis, T.H., Nelson, P.T., Loeb, G.B., Harfe, B.D., Ullian, E., and McManus, M.T. (2008). Dicer loss in striatal neurons produces behavioral and neuroanatomical phenotypes in the absence of neurodegeneration. *Proc Natl Acad Sci USA* 105, 5614–5619.

Czech, B., Malone, C.D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J.A., Sachidanandam, R., et al. (2008). An endogenous small interfering RNA

## References

pathway in *Drosophila*. *Nature* *453*, 798–802.

Daniels, S.M., Melendez-Peña, C.E., Scarborough, R.J., Daher, A., Christensen, H.S., Far, M.E., Purcell, D.F.J., Lainé, S., and Gatignol, A. (2009). Characterization of the TRBP domain required for Dicer interaction and function in RNA interference. *BMC Mol Biol* *10*, 38.

Daviet, L., Erard, M., Dorin, D., Duarte, M., Vaquero, C., and Gatignol, A. (2000). Analysis of a binding difference between the two dsRNA-binding domains in TRBP reveals the modular function of a KR-helix motif. *Eur. J. Biochem.* *267*, 2419–2431.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., Snoeyink, J., Richardson, J.S., et al. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* *35*, W375–W383.

Davis, I.W., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* *32*, W615–W619.

de Carvalho, F., Gheysen, G., Kushnir, S., Van Montagu, M., Inzé, D., and Castresana, C. (1992). Suppression of beta-1,3-glucanase transgene expression in homozygous plants. *Embo J* *11*, 2595–2602.

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

Dephoure N., Zou C., Villen J., Beausoleil SA, Bakalarski C.E., Elledge S.J. and Gygi S.P. (2008). A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci*;105(31), 10762-7.

Deribe YL, Pawson T, Dikic I. (2010). Post-translational modifications in signal integration. *Nat Struct Mol Biol.* *17*(6), 666-72.

Dethoff, E.A., Chugh, J., Mustoe, A.M., and Al-Hashimi, H.M. (2012). Functional complexity and regulation through RNA dynamics. *Nature* *482*, 322–330.

Dickerson RE, Drew HR, Conner BN, Wing RM, Fratini AV, Kopka ML. (1982). The anatomy of A-, B-, and Z-DNA. *Science* *216* (4545), 475—485.

Diederichs, S., and Haber, D.A. (2007). Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* *131*, 1097–1108.

Djuranovic, S., Nahvi, A., and Green, R. (2011). A parsimonious model for gene regulation by miRNAs. *Science* *331*, 550–553.

Doench, J.G. and Sharp P.A. (2004). Specificity of microRNA target selection in translational repression. *Genes Dev* *18* (5), 504–511.

Doi, N., Zenno, S., Ueda, R., Ohki-Hamazaki, H., Ui-Tei, K., and Saigo, K. (2003). Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. *Current Biology* *13*, 41–46.

Dong, A., Xu, X., Edwards, A.M., Chang, C., Chruszcz, M., Cuff, M., Cymborowski, M., Di Leo, R., Egorova, O., Evdokimova, E., Filippova, E., Gu, J., Guthrie, J., Ignatchenko, A., Joachimiak, A., Klostermann, N., Kim, Y., Korniyenko, Y., Minor, W., Que, Q., Savchenko, A., Skarina, T., Tan, K., Yakunin, A., Yee, A., Yim, V., Zhang, R., Zheng, H., Akutsu, M., Arrowsmith, C., Avvakumov, G.V., Bochkarev, A., Dahlgren, L.G., Dhe-Paganon, S., Dimov, S., Dombrowski, L., Finerty, P., Jr., Flodin, S., Flores, A., Graslund, S., Hammerstrom, M., Herman, M.D., Hong, B.S., Hui, R., Johansson, I., Liu,

- Y., Nilsson, M., Nedyalkova, L., Nordlund, P., Nyman, T., Min, J., Ouyang, H., Park, H.W., Qi, C., Rabeh, W., Shen, L., Shen, Y., Sukumard, D., Tempel, W., Tong, Y., Tresagues, L., Vedadi, M., Walker, J.R., Weigelt, J., Welin, M., Wu, H., Xiao, T., Zeng, H. and Zhu, H. (2007) *In situ* proteolysis for protein crystallization and structure determination. *Nat Methods*, **4**, 1019-1021.
- Duarte M., Graham K., daher A., Battisti PL, Bannwarthe S., Segeal E., Jeang KT and Gatignol A. (2000). Characterization of TRBP1 and TRBP2. Stable stem-loop structure at the 5'-end of TRBP2 mRNA resembles HIV-1 TAR and is not found in its processed pseudogene. *J Biomed Sci*. 7(6), 494-506.
- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001a). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498.
- Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001b). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* **15**, 188–200.
- Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001c). Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *Embo J* **20**, 6877–6888.
- Elkayam, E., Kuhn, C.-D., Tocilj, A., Haase, A.D., Greene, E.M., Hannon, G.J., and Joshua-Tor, L. (2012). The Structure of Human Argonaute-2 in Complex with miR-20a. *Cell* **150**, 100–110.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Cryst* (2010). D66, 486-501.
- Ender, C., Krek, A., Friedländer, M.R., Beitzinger, M., Weinmann, L., Chen, W., Pfeffer, S., Rajewsky, N., and Meister, G. (2008). A human snoRNA with microRNA-like functions. *Mol Cell* **32**, 519–528.
- Erard, M., Barker, D.G., Amalric, F., Jeang, K.T., and Gatignol, A. (1998). An Arg/Lys-rich core peptide mimics TRBP binding to the HIV-1 TAR RNA upper-stem/loop. *J Mol Biol* **279**, 1085–1099.
- Ericsson UB, Hallberg BM, DeTitta GT, Dekker N, Nordlund P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem*, **357** (2), 289-298.
- Erson, A.E., and Petty, E.M. (2009). miRNAs and cancer: New research developments and potential clinical applications. *Cancer Biol. Ther.* **8**, 2317–2322.
- Eulalio, A., Behm-Ansmant, I., and Izaurralde, E. (2007). P bodies: at the crossroads of post-transcriptional pathways. *Nat Rev Mol Cell Biol* **8**, 9–22.
- Eulalio, A., Helms, S., Fritzscher, C., Fauser, M., and Izaurralde, E. (2009). A C-terminal silencing domain in GW182 is essential for miRNA function. *Rna* **15**, 1067–1077.
- Eulalio, A., Huntzinger, E., and Izaurralde, E. (2008). Getting to the Root of miRNA-Mediated Gene Silencing. *Cell* **132**, 9–14.
- Fabian, M.R., and Sonenberg, N. (2012). The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. *Nat Struct Mol Biol* **19**, 586–593.
- Fabian, M. R. et al. (2011) miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. *Nature Struct. Mol. Biol.* **18**, 1211–1217.

## References

- Fabian, M.R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J.T., Svitkin, Y.V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J.A., et al. (2009). Mammalian miRNA RISC Recruits CAF1 and PABP to Affect PABP-Dependent Deadenylation. *Mol Cell* 35, 868–880.
- Fabian MR, Sundermeier TR, Sonenberg N. (2010). Understanding how miRNAs post-transcriptionally regulate gene expression. *Prog Mol Subcell Biol.* 50,1–20.
- Fagegaltier, D., Bougé, A.-L., Berry, B., Poisot, E., Sismeiro, O., Coppée, J.-Y., Théodore, L., Voinnet, O., and Antoniewski, C. (2009). The endogenous siRNA pathway is involved in heterochromatin formation in *Drosophila*. *Proc Natl Acad Sci USA* 106, 21258–21263.
- Fairman-Williams, M.E., Guenther, U.-P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Curr Opin Struct Biol* 20, 313–324.
- Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* 9, 102–114.
- Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811.
- Flynt, A.S., Greimann, J.C., Chung, W.-J., Lima, C.D., and Lai, E.C. (2010). MicroRNA Biogenesis via Splicing and Exosome-Mediated Trimming in *Drosophila*. *Mol Cell* 38, 900–907.
- Förstemann, K., Horwich, M.D., Wee, L., Tomari, Y., and Zamore, P.D. (2007). *Drosophila* microRNAs Are Sorted into Functionally Distinct Argonaute Complexes after Production by Dicer-1. *Cell* 130, 287–297.
- Förstemann, K., Tomari, Y., Du, T., Vagin, V.V., Denli, A.M., Bratu, D.P., Klattenhoff, C., Theurkauf, W.E., and Zamore, P.D. (2005). Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biology* 3, e236.
- Frank, F., Fabian, M.R., Stepinski, J., Jemielity, J., Darzynkiewicz, E., Sonenberg, N., and Nagar, B. (2011). Structural analysis of 5'-mRNA-cap interactions with the human AGO2 MID domain. *EMBO Rep.* 12 (5), 415-20.
- Frank, F., Sonenberg, N., and Nagar, B. (2010). Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* 465, 818–822.
- Francia S, Michelini F, Saxena A, Tang D, de Hoon M, Anelli V, Mione M, Carninci P, d'Adda di Fagagna F. (2012). Site-specific Dicer and Drosha RNA products control the DANN-damage response. *Nature* 488 (7410), 231-5.
- Friedman, R.C., Farh, K.K.H., Burge, C.B., and Bartel, D.P. (2008). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19, 92–105.
- Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T., and Oshimura, M. (2004). Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nature* 431, 784–791.
- Garre, P., Pérez-Segura, P., Díaz-Rubio, E., Caldés, T., and la Hoya, de, M. (2010). Reassessing the TARBP2 mutation rate in hereditary nonpolyposis colorectal cancer. *Nat Genet* 42, 817–8–818.
- Gatignol, A., Buckler-White, A., Berkhout, B., and Jeang, K.T. (1991). Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. *Science* 251, 1597–1600.

- Gatignol, A., Duarte M., Daviet L., Chang YN. and Jeang, K.T. (1996). Sequential steps in Tat trans-activation of HIV-1 mediated through cellular DNA, RNA, and protein binding factors. *Gene Expr.* 5(4-5), 217-28.
- Gatignol A, Lainé S, Clerzius G. (2005). Dual role of TRBP in HIV replication and RNA interference: viral diversion of a cellular pathway or evasion from antiviral immunity? *Retrovirology* 2,65.
- Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. *Nat Rev Genet* 10, 94–108.
- Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L.W., Zapp, M.L., Weng, Z., et al. (2008). Endogenous siRNAs Derived from Transposons and mRNAs in *Drosophila* Somatic Cells. *Science* 320, 1077–1081.
- Giles, K.E., Ghirlando, R., and Felsenfeld, G. (2009). Maintenance of a constitutive heterochromatin domain in vertebrates by a Dicer-dependent mechanism. *Nature* 12, 94–99.
- Gill, S.C., and Hippel, P.H.V. (1989). Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data. *Analytical Biochemistry* 182, 319–326.
- Giraldez, A. J. et al. (2006). Zebrafish miR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* 312, 75–79.
- Golden, D.E., Gerbasi, V.R., and Sontheimer, E.J. (2008). An Inside Job for siRNAs. *Mol Cell* 31, 309–312.
- Goto, D.B., and Nakayama, J.-I. (2011). RNA and epigenetic silencing: Insight from fission yeast. *Development, Growth & Differentiation* 54, 129–141.
- Gouet, P. (2003). ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* 31, 3320–3323.
- Gouet, P., Courcelle, E., Stuart, D.I., and Métoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, 305–308.
- Gredell, J.A., Dittmer, M.J., Wu, M., Chan, C., and Walton, S.P. (2010). Recognition of siRNA asymmetry by TAR RNA-binding protein. *Biochemistry* 49, 3148–3155.
- Gregory, R.I., Chendrimada, T.P., Cooch, N., and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* 123, 631–640.
- Gregory, R.I., Yan, K.-P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235–240.
- Green SR, Mathews MB (1992) Two RNA-binding motifs in the double-stranded RNA- activated protein kinase, DAI. *Genes Dev* 6 (12B), 2478-2490.
- Green SR, Manche L, Mathews MB (1995) Two functionally distinct RNA-binding motifs in the regulatory domain of the protein kinase DAI. *Mol Cell Biol* 15 (1), 358-364.
- Grimm, C., Chari, A., Reuter, K., and Fischer, U. (2010). A crystallization screen based on alternative polymeric precipitants. *Acta Cryst. D* 66, 685-697
- Grimson, A., Farh, K.K.-H., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007). MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Mol Cell* 27,

## References

91–105.

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

Guang, S., Bochner, A.F., Pavelec, D.M., Burkhart, K.B., Harding, S., Lachowiec, J., and Kennedy, S. (2008). An Argonaute Transports siRNAs from the Cytoplasm to the Nucleus. *Science* *321*, 537–541.

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

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

Gupta V, Huang X, Patel RC (2003) The carboxy-terminal, M3 motifs of PACT and TRBP have opposite effects on PKR activity. *Virology* *315* (2), 283-291.

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

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

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

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

Han, J., Lee, Y., Yeom, K.-H., Nam, J.-W., Heo, I., Rhee, J.-K., Sohn, S.Y., Cho, Y., Zhang, B.-T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* *125*, 887–901.

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

Haussecker, D., Huang, Y., Lau, A., Parameswaran, P., Fire, A.Z., and Kay, M.A. (2010). Human tRNA-derived small RNAs in the global regulation of RNA silencing. *Rna* *16*, 673–695.

Hendrickson, D.G., Hogan, D.J., McCullough, H.L., Myers, J.W., Herschlag, D., Ferrell, J.E., and Brown, P.O. (2009). Concordant Regulation of Translation and mRNA Abundance for Hundreds of Targets of a Human microRNA. *PLoS Biology* *7*, e1000238.

Hitti EG, Sallacz NB, Schoft VK, Jantsch MF (2004) Oligomerization activity of a double-stranded RNA-binding domain. *FEBS Lett* *574* (1-3), 25-30.

Höck, J., and Meister, G. (2008). The Argonaute protein family. *Genome Biol.* *9*, 210.

Hu, H., Yan, Z., Xu, Y., Hu, H., Menzel, C., Zhou, Y., Chen, W., and Khaitovich, P. (2009). Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics* *10*, 413.

Huang Y, Ji L, Huang Q, Vassilyev DG, Chen X, Ma J-B (2009) Structural insights into mechanisms of the small RNA methyltransferase HEN1. *Nature* *461* (7265), 823-827.

- Hutvagner, G., and Simard, M.J. (2008). Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 9, 22–32.
- Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Bálint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834–838.
- Hutvagner G, Zamore PD (2002) A microRNA in a multiple-turnover RNAi enzyme complex. *Science* 297, 2056–2060.
- Iki, T., Yoshikawa, M., Nishikiori, M., Jaudal, M.C., Matsumoto-Yokoyama, E., Mitsuhashi, I., Meshi, T., and Ishikawa, M. (2010). In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone HSP90. *Mol Cell* 39, 282–291.
- Iwasaki, S., Kobayashi, M., Yoda, M., Sakaguchi, Y., Katsuma, S., Suzuki, T., and Tomari, Y. (2010). Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC-loading of small RNA duplexes. *Mol Cell* 39, 292–299.
- Izant, J.G., and Harold, W. (1984). Inhibition of Thymidine Kinase Gene Expression by Anti-Sense RNA: A Molecular Approach to Genetic Analysis. *Cell* 36, 1007–1015.
- Jensen, G. (2010). *Cryo-EM, Part A: sample preparation and data collection*. Preface. *Meth Enzymol* 481, xv–xvi.
- Jiang, F., Ye, X., Liu, X., Fincher, L., McKearin, D., and Liu, Q. (2005). Dicer-1 and R3D1-L catalyze microRNA maturation in *Drosophila*. *Genes Dev* 19, 1674–1679.
- Jinek, M., and Doudna, J.A. (2009). A three-dimensional view of the molecular machinery of RNA interference. *Nature* 457, 405–412.
- Johnston, D.S., Brown, N.H., Gall, J.G., and Jantsch, M. (1992). A conserved double-stranded RNA-binding domain. *Pnas* 89, 10979–10983.
- Kabsch, W. (2010). XDS. *Acta Cryst.* D66, 125–132
- Kastner B, Fischer N, Golas MM, Sander B, Dube P, Boehringer D, Hartmuth K, Deckert J, Hauer F, Wolf E, et al. (2008). GraFix: Sample preparation for single-particle electron cryomicroscopy. *Nat Methods* 5, 53–55.
- Kato, H., Goto, D.B., Martienssen, R.A., Urano, T., Furukawa, K., and Murakami, Y. (2005). RNA polymerase II is required for RNAi-dependent heterochromatin assembly. *Science* 309, 467–469
- Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* 19, 489–501.
- Kawamata, T., and Tomari, Y. (2010). Making RISC. *Trends in Biochemical Sciences* 35 (7), 368–376.
- Kawamata, T., Seitz, H., and Tomari, Y. (2009). Structural determinants of miRNAs for RISC-loading and slicer-independent unwinding. *Nat Struct Mol Biol* 16, 953–960.
- Kennerdell, J.R., and Carthew, R.W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. *Cell* 95, 1017–1026.
- Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer

## References

functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* 15, 2654–2659.

Ketting, R.F., Haverkamp, T.H., van Luenen, H.G., and Plasterk, R.H. (1999). Mut-7 of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* 99, 133–141.

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

Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115, 209–216.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* 10, 126–139.

Kim DH, Villeneuve LM, Morris KV, Rossi JJ. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat. Struct. Mol. Biol.* 13, 793–797.

Kini, H.K., and Walton, S.P. (2009). Effect of siRNA terminal mismatches on TRBP and Dicer binding and silencing efficacy. *Febs J* 276, 6576–6585.

Kloosterman, W.P., and Plasterk, R.H.A. (2006). The diverse functions of microRNAs in animal development and disease. *Dev Cell* 11, 441–450.

Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* 293, 2269–2271.

Kok, K.H., Ng, M.-H.J., Ching, Y.-P., and Jin, D.-Y. (2007). Human TRBP and PACT Directly Interact with Each Other and Associate with Dicer to Facilitate the Production of Small Interfering RNA\*. *J Biol Chem* 282, 17649–17657.

Kolakofsky, D., Kowalinski, E., and Cusack, S. (2012). A structure-based model of RIG-I activation. *Rna* 18, 2118–2127.

Koralov, S.B., Muljo, S.A., Galler, G.R., Krek, A., Chakraborty, T., Kanellopoulou, C., Jensen, K., Cobb, B.S., Merckenschlager, M., Rajewsky, N., et al. (2008). Dicer ablation affects antibody diversity and cell survival in the B lymphocyte lineage. *Cell* 132, 860–874.

Koscianska, E., Starega-Roslan, J., and Krzyzosiak, W.J. (2011). The Role of Dicer Protein Partners in the Processing of MicroRNA Precursors. *PLoS ONE* 6, e28548.

Krovat BC and Jantsch MF (1996) Comparative mutational analysis of the double-stranded RNA-binding domains of *Xenopus laevis* RNA-binding protein A. *J Biol Chem* 271 (45), 28112-28119.

Kotaja, N., Bhattacharyya, S.N., Jaskiewicz, L., Kimmins, S., Parvinen, M., Filipowicz, W., and Sassone-Corsi, P. (2006). The chromatoid body of male germ cells: similarity with processing bodies and presence of Dicer and microRNA pathway components. *Proc Natl Acad Sci USA* 103, 2647–2652.

Kurreck J. (2009). RNA interference: from basic research to therapeutic applications. *Angew Chem Int Ed Engl*, 48(8), 1378-98.

Kuzuoglu-Ozturk, D., Huntzinger, E., Schmidt, S., and Izaurralde, E. (2012). The *Caenorhabditis elegans* GW182 protein AIN-1 interacts with PAB-1 and subunits of the PAN2-PAN3 and CCR4-NOT deadenylase complexes. *Nucleic Acids Res* 1–15.

- Kwak, P.B., and Tomari, Y. (2012). The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat Struct Mol Biol* 19, 145–151.
- Kwak, P.B., Iwasaki, S., and Tomari, Y. (2010). The microRNA pathway and cancer. *Cancer Sci* 101, 2309–2315.
- Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. *Current Biology* 14, 2162–2167.
- Laraki, G., Clerzius, G., Daher, A., Melendez-Pena, C., Daniels, S., and Gagnon, A. (2008). Interactions between the double-stranded RNA-binding proteins TRBP and PACT define the Medial domain that mediates protein-protein interactions. *RNA Biol* 5, 92–103.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948.
- Lau, P.-W., Guiley, K.Z., De, N., Potter, C.S., Carragher, B., and Macrae, I.J. (2012). The molecular architecture of human Dicer. *Nat Struct Mol Biol* 19, 436–440.
- Lau, P.-W., Potter, C.S., Carragher, B., and Macrae, I.J. (2009). Structure of the Human Dicer-TRBP Complex by Electron Microscopy. *Structure* 17, 1326–1332.
- Lau PW., Potter C.S., Carragher, B., and Macrae, I.J. (2012). DOLORS: versatile strategy for internal labeling and domain localization in electron microscopy. *Structure*. 20(12), 1995-2002
- Lawrence, R.J., and Pikaard, C.S. (2003). Transgene-induced RNA interference: a strategy for overcoming gene redundancy in polyploids to generate loss-of-function mutations. *The Plant Journal* 36, 114–121.
- Lee, J.Y., Kim, H., Ryu, C.H., Kim, J.Y., Choi, B.H., Lim, Y., Huh, P.-W., Kim, Y.-H., Lee, K.-H., Jun, T.-Y., et al. (2004a). Merlin, a tumor suppressor, interacts with transactivation-responsive RNA-binding protein and inhibits its oncogenic activity. *J Biol Chem* 279, 30265–30273.
- Lee K, Fajardo MA, Braun RE. (1996). A testis cytoplasmic RNA-binding protein that has the properties of a translational repressor. *Mol. Cell. Biol.* 16, 3023–3034.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843–854.
- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415–419.
- Lee, Y., Hur, I., Park, S.-Y., Kim, Y.-K., Suh, M.R., and Kim, V.N. (2006). The role of PACT in the RNA silencing pathway. *Embo J* 25, 522–532.
- Lee, Y., Jeon, K., Lee, J.-T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *Embo J* 21, 4663–4670.
- Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S.H., and Kim, V.N. (2004b). MicroRNA genes are transcribed by RNA polymerase II. *Embo J* 23, 4051–4060.
- Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004c). Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* 117, 69–81.

## References

- Lejeune, E., and Allshire, R.C. (2011). Common ground: small RNA programming and chromatin modifications. *Current Opinion in Cell Biology* 23, 258–265.
- Leuschner, P.J.F., Ameres, S.L., Kueng, S., and Martinez, J. (2006). Cleavage of the siRNA passenger strand during RISC assembly in human cells. *Nature Publishing Group* 7, 314–320.
- Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120, 15–20.
- Li, Y.-X., Farrell, M.J., Liu, R., Mohanty, N., and Kirby, M.L. (2000). Double-Stranded RNA Injection Produces Null Phenotypes in Zebrafish. *Developmental Biology* 217, 394–405.
- Lian, S.L., Li, S., Abadal, G.X., Pauley, B.A., Fritzler, M.J., and Chan, E.K.L. (2009). The C-terminal half of human Ago2 binds to multiple GW-rich regions of GW182 and requires GW182 to mediate silencing. *Rna* 15, 804–813.
- Lim HG, Suzuki K, Cooper DA, Kelleher AD. (2008). Promoter-targeted siRNAs induce gene silencing of simian immunodeficiency virus (SIV) infection in vitro. *Mol. Ther.* 16, 565–570
- Lippman, Z., and Martienssen, R. (2004). The role of RNA interference in heterochromatic silencing. *Nature* 431, 364–370.
- Liu, J., Rivas, F.V., Wohlschlegel, J., Yates, J.R., Parker, R., and Hannon, G.J. (2005). A role for the P-body component GW182 in microRNA function. *Nat Cell Biol* 7, 1261–1266.
- Liu, Q. (2003). R2D2, a Bridge Between the Initiation and Effector Steps of the Drosophila RNAi Pathway. *Science* 301, 1921–1925.
- Liu, X., Jin, D.-Y., McManus, M.T., and Mourelatos, Z. (2012). Precursor MicroRNA-Programmed Silencing Complex Assembly Pathways in Mammals. *Mol Cell* 46, 507–517.
- Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. *Science* 303, 95–98.
- Ma, E., Macrae, I.J., Kirsch, J.F., and Doudna, J.A. (2008). Autoinhibition of Human Dicer by Its Internal Helicase Domain. *J Mol Biol* 380, 237–243.
- Ma, J.B., Ye, K., and Patel, D.J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* 429, 318–322.
- Ma, J.-B., Yuan, Y.-R., Meister, G., Pei, Y., and Patel, T.T.D.J. (2005). Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* 434, 666–670.
- Ma, E., Zhou, K., Kidwell, M.A., and Doudna, J.A. (2012). Coordinated Activities of Human Dicer Domains in Regulatory RNA Processing. *J Mol Biol* 422, 466–476.
- McMillan NA, Carpick BW, Hollis B, Toone WM, Zamanian-Daryoush M, Williams BR (1995) Mutational analysis of the double-stranded RNA (dsRNA) binding domain of the dsRNA-activated protein kinase, PKR. *J Biol Chem* 270 (6), 2601-2606.
- Macrae, I.J., and Doudna, J.A. (2007). Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr Opin Struct Biol* 17, 138–145.
- Macrae, I.J., Ma, E., Zhou, M., Robinson, C.V., and Doudna, J.A. (2008). In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci USA* 105, 512–517.

- Macrae, I.J., Zhou, K., and Doudna, J.A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. *Nat Struct Mol Biol* 14, 934–940.
- Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural Basis for Double-Stranded RNA Processing by Dicer. *Science* 311, 195–198.
- Maniataki, E., and Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev* 19, 2979–2990.
- Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., et al. (2010). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* 39, D225–D229.
- Martienssen RA, Zaratiegui M, Goto DB (2005) RNA interference and heterochromatin in the fission yeast *Schizosaccharomyces pombe*. *Trends Genet* 21, 450–456.
- Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* 110, 563–574.
- Matranga, C., Tomari, Y., Shin, C., Bartel, D.P., and Zamore, P.D. (2005). Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* 123, 607–620.
- Matthews, B.W. (1968). Solvent content of protein crystals. *J Mol Biol* 33, 491–497.
- Matzke, M.A., Primig, M., Trnovsky, J., and Matzke, A.J. (1989). Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. *Embo J* 8, 643–649.
- McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Cryst. D* 63, 32–41.
- Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human Argonaute2 Mediates RNA Cleavage Targeted by miRNAs and siRNAs. *Mol Cell* 15, 185–197.
- Meister, G., Landthaler, M., Peters, L., Chen, P.Y., Urlaub, H., Lührmann, R., and Tuschl, T. (2005). Identification of novel argonaute-associated proteins. *Current Biology* 15, 2149–2155.
- Melo, S. A., Roperio, S., Moutinho, C., Aaltonen, L., Yamamoto, H., Calin, G. A. et al. (2009). A 646 TARBP2 mutation in human cancer impairs micro- 647 RNA processing and DICER1 function. *Nat. Genet.* 41, 365–370.
- Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A., and Matzke, A.J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *Embo J* 19, 5194–5201.
- Micklem DR, Adams J, Grunert S, St Johnston D (2000) Distinct roles of two conserved Staufen domains in oskar mRNA localization and translation. *EMBO J* 19 (6), 1366–1377.
- Miki, D. (2005). RNA Silencing of Single and Multiple Members in a Gene Family of Rice. *Plant Physiol.* 138, 1903–1913.
- Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M.C. (2005). Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes Dev* 19, 2837–2848.
- Miyoshi, T., Takeuchi, A., Siomi, H., and Siomi, M.C. (2010). A direct role for Hsp90 in pre-RISC formation in *Drosophila*. *Nat Struct Mol Biol* 1–3.

## References

- Mooij, W.T.M., Mitsiki, E., and Perrakis, A. (2009). ProteinCCD: enabling the design of protein truncation constructs for expression and crystallization experiments. *Nucleic Acids Res* 37, W402–W405.
- Morris KV, Chan SW, Jacobsen. (2004). SE, Looney DJ. Small interfering RNA-induced transcriptional gene silencing in human cells. *Science* 305,1289–1292.
- Mortensen, R.D., Serra, M., Steitz, J.A., and Vasudevan, S. (2011). Posttranscriptional activation of gene expression in *Xenopus laevis* oocytes by microRNA-protein complexes (microRNPs). *Proc Natl Acad Sci USA* 108, 8281–8286.
- Moser, J.J., Eystathioy, T., Chan, E.K.L., and Fritzler, M.J. (2007). Markers of mRNA stabilization and degradation, and RNAi within astrocytoma GW bodies. *J. Neurosci. Res.* 85, 3619–3631.
- Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci USA* 102, 12135–12140.
- Murchison, E.P., Stein, P., Xuan, Z., Pan, H., Zhang, M.Q., Schultz, R.M., and Hannon, G.J. (2007). Critical roles for Dicer in the female germline. *Genes Dev* 21, 682–693.
- Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* 2, 279–289.
- Newman et al (2005). Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. *Acta Cryst.* D61, 1426-1431
- Nilsen T.W. (2008) Endo-siRNAs: Yet another layer of complexity in RNA silencing. *Nat Struct Mol Biol* 15, 546–548.
- Noland, C.L., Ma, E., and Doudna, J.A. (2011). siRNA repositioning for guide strand selection by human Dicer complexes. *Mol Cell* 43, 110–121.
- Ohrt, T., Mutze, J., Staroske, W., Weinmann, L., Hock, J., Crell, K., Meister, G., and Schwillle, P. (2008). Fluorescence correlation spectroscopy and fluorescence cross-correlation spectroscopy reveal the cytoplasmic origination of loaded nuclear RISC in vivo in human cells. *Nucleic Acids Res* 36, 6439–6449.
- Okamura, K. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev* 18, 1655–1666.
- Okamura, K., and Lai, E.C. (2008). Endogenous small interfering RNAs in animals. *Nat Rev Mol Cell Biol* 9, 673–678.
- Okamura, K., Chung, W.-J., and Lai, E.C. (2008a). The long and short of inverted repeat genes in animals: microRNAs, mirtrons and hairpin RNAs. *Cell Cycle* 7, 2840–2845.
- Okamura, K., Chung, W.-J., Ruby, J.G., Guo, H., Bartel, D.P., and Lai, E.C. (2008b). The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* 453, 803–806.
- Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* 130, 89–100.
- Ong CL, Thorpe JC, Gorry PR, Bannwarth S, Jaworowski A, Howard JL, Chung S, Campbell S, Christensen HS, Clerzius G, Moulard AJ, Gatignol A, Purcell DF: Low TRBP levels support an innate human immunodeficiency virus type 1 resistance in astrocytes by enhancing the PKR

- antiviral response. *J Virol* 2005, 79(20),12763-12772
- Orban, T.I., and Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *Rna* 11, 459–469.
- Orom, U. A., Nielsen, F. C. & Lund, A. H. (2008) MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol. Cell* 30, 460–471.
- O'Reilly D R, Miller L K and Luckow V A (1992) Baculovirus expression vectors—A laboratory manual
- Pak J. and Fire A. (2007) Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* 315,241–244.
- Park, J.-E., Heo, I., Tian, Y., Simanshu, D.K., Chang, H., Jee, D., Patel, D.J., and Kim, V.N. (2011). Dicer recognizes the 5'-end of RNA for efficient and accurate processing. *Nature* 475, 201–205.
- Park, Y.D., Papp, I., Moscone, E.A., Iglesias, V.A., Vaucheret, H., Matzke, A.J., and Matzke, M.A. (1996). Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in methylation and gene activity. *Plant J.* 9, 183–194.
- Parker, G.S., Maity, T.S., and Bass, B.L. (2008). dsRNA-binding Properties of RDE-4 and TRBP Reflect Their Distinct Roles in RNAi. *J Mol Biol* 384, 967–979.
- Parker, J.S., Roe, S.M., and Barford, D. (2005a). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* 434, 663–666.
- Parker, J.S., Roe, S.M., and Barford, D. (2005b). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* 434, 663–666.
- Patel RC, Stanton P, Sen GC (1996) Specific mutations near the amino terminus of double-stranded RNA-dependent protein kinase (PKR) differentially affect its double-stranded RNA-binding and dimerization properties. *J Biol Chem* 271 (41):25657-25663.
- Petersen, C.P., Bordeleau, M.-E., Pelletier, J., and Sharp, P.A. (2006). Short RNAs Repress Translation after Initiation in Mammalian Cells. *Mol Cell* 21, 533–542.
- Pham, J.W., Pellino, J.L., Lee, Y.S., Carthew, R.W., and Sontheimer, E.J. (2004). A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in *Drosophila*. *Cell* 117, 83–94.
- Piao, X., Zhang, X., Wu, L., and Belasco, J.G. (2010). CCR4-NOT Deadenylation mRNA Associated with RNA-Induced Silencing Complexes in Human Cells. *Mol Cell Biol* 30, 1486–1494.
- Pillai, R.S., Artus, C.G., and Filipowicz, W. (2004). Tethering of human Ago proteins to mRNA mimics the miRNA-mediated repression of protein synthesis. *Rna* 10, 1518–1525.
- Pillai, R.S., Bhattacharyya, S.N., and Filipowicz, W. (2007). Repression of protein synthesis by miRNAs: how many mechanisms? *Trends in Cell Biology* 17, 118–126.
- Potterton, E., Briggs, P., Turkenburg, M. and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr*, 59, 1131-1137.
- Provost, P., Dishart, D., Doucet, J., Freundewey, D., Samuelsson, B., and Rådmark, O. (2002). Ribonuclease activity and RNA-binding of recombinant human Dicer. *Embo J* 21, 5864–5874.
- Qin H., Chen F., Huan X., Machida S., Song, J. Yuan YA (2010). Structure of the Arabidopsis

## References

- thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein-protein interaction. *RNA* 16, 474–481.
- Radaev, S., Li, S., and Sun, P.D. (2006). A survey of protein–protein complex crystallizations. *Acta Crystallogr D Biol Crystallogr* 62, 605–612.
- Rana, T.M. (2007). Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol* 8 (1), 23–36.
- Rand, T.A., Petersen, S., Du, F., and Wang, X. (2005). Argonaute2 Cleaves the Anti-Guide Strand of siRNA during RISC Activation. *Cell* 123, 621–629.
- Ramos A, Grünert S, Adams J, Micklem DR, Proctor MR, Freund S, Bycroft M, St Johnston D, Varani G (2000) RNA recognition by a Staufen double-stranded RNA-binding domain. *EMBO J* 19 (5), 997–100.
- Ratcliff, F. (1997). A Similarity Between Viral Defense and Gene Silencing in Plants. *Science* 276, 1558–1560.
- Reinhart BJ, Slack F.J., Basson M., Pasquinelli A.E., Rougvie A.E., Horvitz H.R., and Ruvkun G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403 (6772), 901–6.
- Ren, Y.-F., Li, G., Wu, J., Xue, Y.-F., Song, Y.-J., Lv, L., Zhang, X.-J., and Tang, K.-F. (2012). Dicer-Dependent Biogenesis of Small RNAs Derived from 7SL RNA. *PLoS ONE* 7, e40705.
- Reyes-Turcu FE and Grewal SI, (2012) Different means, same end-heterochromatin formation by RNAi and RNAi-independent RNA processing factors in fission yeast. *Curr Opin Genet Dev* 22(2), 156–63.
- Rigoutsos, I. (2009). New Tricks for Animal MicroRNAs: Targeting of Amino Acid Coding Regions at Conserved and Nonconserved Sites. *Cancer Research* 69, 3245–3248.
- Rivas, F.V., Tolia, N.H., Song, J.-J., Aragon, J.P., Liu, J., Hannon, G.J., and Joshua-Tor, L. (2005). Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol* 12, 340–349.
- Robb, G.B., Brown, K.M., Khurana, J., and Rana, T.M. (2005). Specific and potent RNAi in the nucleus of human cells. *Nat Struct Mol Biol* 12, 133–137.
- Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Res.* 14, 1902–1910.
- Rossi, J.J. (2005). Mammalian Dicer finds a partner. *Nature Publishing Group* 6, 927–929.
- Röther S. and Meister, G. (2011) Small RNAs derived from longer non-coding RNAs. *Biochimie* 93 (11), 1905–1915.
- Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-Scale Sequencing Reveals 21U-RNAs and Additional MicroRNAs and Endogenous siRNAs in *C. elegans*. *Cell* 127, 1193–1207.
- Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature* 448, 83–86.
- Rüdel, S., Wang, Y., Lenobel, R., Körner, R., Hsiao, H.-H., Urlaub, H., Patel, D., and Meister, G.

- (2010). Phosphorylation of human Argonaute proteins affects small RNA-binding. *Nucleic Acids Res.* 39 (6), 2330-2343.
- Rush J., Moritz A., Lee K.A., Guo A., Goss V.L., Spek E. J., Zhang H., Zha X.-M., Polakiewicz R.D., and Comb M.J. (2005). Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol*, 23, 94-101
- Ryter, J., and Schultz, S. (1998). Molecular basis of double-stranded RNA–protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *Embo J Vol 17*, 7505–7513.
- Saenger W (1984) Principles of nucleic acid structure. Springer-Verlag New York,
- Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2005). Processing of pre-microRNAs by the Dicer-1-Loquacious complex in *Drosophila* cells. *PLoS Biology* 3, e235.
- Saraiya, A.A., and Wang, C.C. (2008). snoRNA, a novel precursor of microRNA in *Giardia lamblia*. *PLoS Pathog* 4, e1000224.
- Sasaki T., and Shimizu (2007). Evolutionary conservation of a unique amino acid sequence in human DICER protein essential for binding to Argonaute family proteins. *Gene* 396 (2): 312-320.
- Sasaki, T., Shiohama, A., Minoshima, S., and Shimizu, N. (2003). Identification of eight members of the Argonaute family in the human genome ☆. *Genomics* 82, 323–330.
- Saunders, L.R., and Barber, G.N. (2003). The dsRNA-binding protein family: critical roles, diverse cellular functions. *The FASEB Journal* 17, 961–983.
- Saxena, S. (2003). Small RNAs with Imperfect Match to Endogenous mRNA Repress Translation: Implications for off-target activity of small inhibitory RNA in mammalian cells. *Journal of Biological Chemistry* 278, 44312–44319.
- Schirle, N.T., and Macrae, I.J. (2012). The crystal structure of human Argonaute2. *Science* 336, 1037–1040.
- Schneider, M.R. (2012). MicroRNAs as novel players in skin development, homeostasis and disease. *Br. J. Dermatol.* 166, 22–28.
- Schuldt AJ, Adams JH, Davidson CM, Micklem DR, Haseloff J, St Johnston D, Brand AH (1998) Miranda mediates asymmetric protein and RNA localization in the developing nervous system. *Genes Dev* 12 (12), 1847-1857.
- Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199–208.
- Seitz, H., Tushir, J.S., and Zamore, P.D. (2011). A 5'-uridine amplifies miRNA/miRNA\* asymmetry in *Drosophila* by promoting RNA-induced silencing complex formation. *Silence* 2, 4.
- Selbach, M. et al. Widespread changes in protein synthesis induced by microRNAs. *Nature* 455, 58–63 (2008).
- Shen, B., and Goodman, H.M. (2004). Uridine addition after microRNA- directed cleavage. *Science* 306, 997.
- Sijen, T., Vijn, I., Rebocho, A., van Blokland, R., Roelofs, D., Mol, J.N., and Kooter, J.M. (2001). Transcriptional and posttranscriptional gene silencing are mechanistically related. *Current Biology* 11, 436–440.

## References

- Sinkkonen, L., Hugenschmidt, T., Filipowicz, W., and Svoboda, P. (2010). Dicer is associated with ribosomal DNA chromatin in mammalian cells. *PLoS ONE* 5, e12175.
- Siomi, H., and Siomi, M.C. (2009). On the road to reading the RNA-interference code. *Nature* 457, 396–404.
- Smith, C.J.S., Watson, C.F., Bird, C.R., Ray, J., Schuch, W. and Grierson, D. (1990). Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants. *Mol. Gen. Genet.* 224, 447-481.
- Soifer, H.S., Sano, M., Sakurai, K., Chomchan, P., Sætrom, P.L., Sherman, M.A., Collingwood, M.A., Behlke, M.A., and Rossi, J.J. (2008). A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. *Nucleic Acids Res* 36, 6511–6522.
- Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA-binding motif in RNAi effector complexes. *Nat Struct Biol* 10, 1026-1032.
- Song, J.-J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity. *Science* 305, 1434–1437.
- Sridhar, P., Awasthi, A., Azim, A., Burma, S., Habib, S., Jain, A., Mukherjee, B., Ranjan, A., and Hasnain, S. (1994). Baculovirus vector-mediated expression of heterologous genes in insect cells. *J Biosci* 19, 603–614.
- Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41, 207–234.
- Stura E.A., Nemerow G.R., Wilson I.A (1992) Strategies in the crystallization of glycoproteins and protein complexes. *Journal of Crystal Growth* 122, 273-285
- Suzuki, K., Juelich, T., Lim, H., Ishida, T., Watanebe, T., Cooper, D.A., Rao, S., and Kelleher, A.D. (2008). Closed chromatin architecture is induced by an RNA duplex targeting the HIV-1 promoter region. *J Biol Chem* 283, 23353–23363.
- Svoboda, P., Stein, P., Hayashi, H., and Schultz, R.M. (2000). Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development* 127, 4147–4156.
- Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* 99, 123–132.
- Tahbaz, N., Kolb, F.A., Zhang, H., JARONCZYK, K., and β, W.F.T.C.H.D. (2004). Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. *EMBO Rep* 5, 189–194.
- Takeshita, D., Zenno, S., Lee, W.C., Nagata, K., Saigo, K., and Tanokura, M. (2007). Homodimeric structure and double-stranded RNA cleavage activity of the C-terminal RNase III domain of human dicer. *J Mol Biol* 374, 106–120.
- Tan, G.S., Garchow, B.G., Liu, X., Yeung, J., Morris, J.P., Cuellar, T.L., McManus, M.T., and Kiriakidou, M. (2009). Expanded RNA-binding activities of mammalian Argonaute 2. *Nucleic Acids Res* 37, 7533–7545.
- Tang, F., Kaneda, M., O'Carroll, D., Hajkova, P., Barton, S.C., Sun, Y.A., Lee, C., Tarakhovskiy, A., Lao, K., and Surani, M.A. (2007). Maternal microRNAs are essential for mouse zygotic development.

Genes Dev 21, 644–648.

Tay, Y., Zhang, J., Thomson, A. M., Lim, B. & Rigoutsos, I. (2008). MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* 455, 1124–1128

Tolia, N.H., and Joshua-Tor, L. (2007). Slicer and the argonautes. *Nat Chem Biol* 3, 36-43.

Tomari, Y. (2004). A Protein Sensor for siRNA Asymmetry. *Science* 306, 1377–1380.

Tomari Y, Zamore PD. (2005) Perspective: machines for RNAi. *Genes Dev*,19,517–529.

Tomari, Y., Du, T., and Zamore, P.D. (2007). Sorting of *Drosophila* Small Silencing RNAs. *Cell* 130, 299–308.

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

van Blokland, K, van der Geest, N, Mol, J, and Kooter, J (1994). Transgene-mediated suppression of chalcone synthase expression in *Petunia hybrida* results from an increase in RNA turnover. *Plant J.* 6, 861-877.

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

Vasudevan, S., Tong, Y. & Steitz, J. A. Switching (2007). from repression to activation: microRNAs can up-regulate translation. *Science* 318, 1931–1934.

Vaucheret H. (2006) Post-transcriptional small RNA pathways in plants: Mechanisms and regulations. *Genes Dev* 20:759–771.

Vella, M.C. (2004). The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. *Genes Dev* 18, 132–137.

Verdel, A., Vavasseur, A., Le Gorrec, M., and Touat-Todeschini, L. (2009). Common themes in siRNA-mediated epigenetic silencing pathways. *Int. J. Dev. Biol.* 53, 245–257.

Volpe, T.A. (2002). Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi. *Science* 297, 1833–1837.

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

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

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

Wang H.W., Noland C., Siridechadilok B., Taylor DW, Ma E., Felder K., Doudna JA., and Nogales E. (2009b). Structural insights into RNA processing by the human RISC-loading complex. *Nat Struct Mol Biol* 16, 1148–1153.

Wargelius, A., Ellingsen, S., and Fjose, A. (1999). Double-Stranded RNA Induces Specific Developmental Defects in Zebrafish Embryos. *Biochemical and Biophysical Research*

## References

Communications 263, 156–161.

Wassenegger, M. (2005). The Role of the RNAi Machinery in Heterochromatin Formation. *Cell* 122, 13–16.

Wassenegger, M., Heimes, S., Riedel, L., and Sanger, H.L. (1994). RNA-directed de novo methylation of genomic sequences in plants. *Cell* 76, 567–576.

Waterhouse, P.M., Graham, M.W., and Wang, M.B. (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc Natl Acad Sci USA* 95, 13959–13964.

Webb, M.R. (1992). A continuous spectrophotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems. *Pnas* 89, 4884–4887.

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

Welker, N.C., Pavelec, D.M., Nix, D.A., Duchaine, T.F., Kennedy, S., and Bass, B.L. (2010). Dicer's helicase domain is required for accumulation of some, but not all, *C. elegans* endogenous siRNAs. *Rna* 16, 893–903.

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

Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* 11, 228–234.

Wu, L., and Belasco, J.G. (2008). Let Me Count the Ways: Mechanisms of Gene Regulation by miRNAs and siRNAs. *Mol Cell* 29, 1–7.

Wu, L., Fan, J., and Belasco, J.G. (2008). Importance of Translation and Nonnucleolytic Ago Proteins for On-Target RNA Interference. *Current Biology* 18, 1327–1332.

Yamashita, S., Nagata, T., Kawazoe, M., Takemoto, C., Kigawa, T., Guntert, P., Kobayashi, N., Terada, T., Shirouzu, M., Wakiyama, M., et al. (2011). Structures of the first and second double-stranded RNA-binding domains of human TAR RNA-binding protein. *Protein Sci* 20, 118–130.

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

Yang, J.-S., and Lai, E.C. (2010). Dicer-independent, Ago2-mediated microRNA biogenesis in vertebrates. *Cell Cycle* 9, 4455–4460.

Yang, J.-S., and Lai, E.C. (2011). Alternative miRNA Biogenesis Pathways and the Interpretation of Core miRNA Pathway Mutants. *Mol Cell* 43, 892–903.

Yang, S.W., Chen, H.-Y., Yang, J., MACHIDA, S., Chua, N.-H., and YUAN, Y.A. (2010). Structure of arabidopsis HYPONASTIC LEAVES1 and its molecular implications for miRNA processing. *Structure* 18, 594–605.

Yekta S, Shih IH, Bartel DP (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science* 304: 594–596.

Ye, X., Paroo, Z., and Liu, Q. (2007). Functional anatomy of the *Drosophila* microRNA-generating

enzyme. *J Biol Chem* 282, 28373–28378.

Yeom, K.-H., Lee, Y., Han, J., Suh, M.R., and Kim, V.N. (2006). Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. *Nucleic Acids Res* 34, 4622–4629.

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

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

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

Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: Double-Stranded RNA Directs the ATP-Dependent Cleavage of mRNA at 21 to 23 Nucleotide Intervals. *Cell* 101, 25–33.

Zehir, A., Hua, L.L., Maska, E.L., Morikawa, Y., and Cserjesi, P. (2010). Dicer is required for survival of differentiating neural crest cells. *Developmental Biology* 340, 459–467.

Zeng, Y., and Cullen, B.R. (2003). Sequence requirements for micro RNA processing and function in human cells. *Rna* 9, 112–123.

Zeng, Y., and Cullen, B.R. (2004). Structural requirements for pre-microRNA-binding and nuclear export by Exportin 5. *Nucleic Acids Res* 32, 4776–4785.

Zeng, Y.Y., Yi, R.R., and Cullen, B.R.B. (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Pnas* 100, 9779–9784.

Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *Embo J* 21, 5875–5885.

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single Processing Center Models for Human Dicer and Bacterial RNase III. *Cell* 118, 57–68.

Zhang F, Romano PR, Nagamura-Inoue T, Tian B, Dever TE, Mathews MB, Ozato K, Hinnebusch AG (2001) Binding of double-stranded RNA to protein kinase PKR is required for dimerization and promotes critical autophosphorylation events in the activation loop. *J Biol Chem* 276 (27):24946-24958.

Zhong J, Peters AH, Lee K, Braun (1999). RE: A double-stranded RNA-binding protein required for activation of repressed mes- sages in mammalian germ cells. *Nat Genet*, 22(2):171-174.

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

References

## **9 Acknowledgements**

I would like to thank Prof. Dr. Ralf Ficner for the supervision of this thesis, for advice, support, trust and fruitful discussions. I also would like to thank Prof. Dr. Holger Stark and Prof. Dr. Kai Tittmann for taking part in my thesis committee, for all their support, encouragement and great ideas to promote my work throughout the years. I would also like to thank them for generous ability to use the instruments and facilities of their departments. Furthermore I am grateful to Prof. Dr. Marina Rodnina, Prof. Dr. Heinz Neumann and Dr. Jochen Hub for being members of my extended thesis committee.

Especially I would like to thank David Haselbach for his great collaboration, for his optimism, his immense imagination and for introducing me into the fascinating world of electron microscopy. I am also grateful to all the members of the group of Prof. Dr. Holger Stark for providing a kind atmosphere working there. I would also like to especially thank Dr. Ashwin Chari for helpful discussions and lots of helpful advises and Dr. Boris Busche for his help with the submission of the thesis.

I would like to thank Dr. Bernhard Schmitt and Olaf Bernhard from the Faculty of Medicine-Department of Biochemistry II in Göttingen for mass spectrometry support. I would like to thank Anne Frohn and Prof. Dr. G. Meister for their help in carrying out activity assays within their laboratory. I'm also thankful to Dr. Jochen Deckert for Roche, Kulmbach for his great collaboration and for providing us with antisense siRNAs and siRNAs duplexes.

I would like to thank the beamline staff from ESRF (Grenoble), Bessy (Berlin) and HASY-lab for their great technical support. I would like to thank Clement Blanchet for his help during data collection of SAXS data, for evaluation of the data and helpful discussions during data interpretation.

I would like to thank all former and present members of the department of structural biology for their help in the lab and a great atmosphere. I am deeply thankful to Dr. Piotr Neumann for his continuous support, his teaching in theoretical and practical aspects of crystallography, his help during the collection and processing of various data sets, always having an open door and being such a pleasant person. I'm thankful to Dr. Thomas Monecke for introducing me into several techniques and instruments used within the lab and for lots of helpful comments and suggestions. Thanks to Michael Franke and Dipl. Ing. Daniel Weinrich for carrying out several crystallization trials using the phoenix robot. Thanks to Michael Franke and Dr. Marcus Resch for their support within the lab, their help and ideas dealing with the insect cell culture and a great time in the lab. Additionally, I would like to thank Dr. Marcus Resch for his immense creativity and his great knowledge dealing with several problems. I would also love to thank, Dr. Achim Dickmanns for his continuous support throughout the years. Thanks to my

## Acknowledgements

office mates, Dr. Heide Marie Roth, Andreas Schmitt, Dr. Eike Schulz, Dr. Sohail Koshnevis and Lin-ta Hsu for helpful discussions, their encouragement, lovely coffee breaks and such a great time.

Thanks to the members of the bioanalytics department for all their help carrying out experiments and using their instruments. In particular I would like to thank Dr. Danilo Meyer, Dr. Stefan Lüdke and Astrid Sitte. I would also love to thank Stefan Lüdke for his help with data evaluation, structure refinement and his friendship.

Furthermore I would like to thank Dr. Achim Dickmanns, Dr. Thomas Monecke, Yasar Laqman Ahmed, Andreas Schmitt and David Haselbach for feedback and corrections of this thesis.

I would like to thank the Göttingen Graduate School for Neuroscience, Biophysics and Molecular Bioscience (GGNB) office, particularly Kristin Pöhlker, Christina Bach and Christin Fischer for all their assistance through the administrative work and their open ears. I am additionally deeply thankful to the GGNB for the bridging stipend that enabled me to finish this thesis.

I like to acknowledge the Deutsche Forscher Gruppe DFG 855 for funding of three years as well as the SFB 806 for funding the last month of my thesis.

I would like to thank my family and my friends for their support, love, their good advices and their encouragements.

And thanks to Matthias for his love, patience and support.

## 10 Appendix

Base sequence of *H. sapiens* Dicer optimized for expression in Sf9 insect cells, which was orderend from Geneart:

ATGAAGTCCCCGCTCTGCAGCCCCTGTCTATGGCTGGCCTGCAGCTGATGACCCCCGCTTCTCCC  
 CTATGGGCCCTTCTTCGGTCTGCCCTGGCAGCAGGAGGCTATCCACGACAACATCTACACCCCCGCA  
 AGTACCAGGTGGAGCTGCTGGAGGCTGCTCTGGACCACAACACCATCGTGTGCCTGAACACCGGTTCCG  
 GCAAGACCTTCATCGCTGTGCTGCTGACCAAGGAGCTGTCTACCAGATCCGTGGCGACTTCTCCCGTA  
 ACGGCAAGCGTACCGTGTTCCTGGTCAACTCCGCTAACCAGGTGGCCCAGCAAGTCAGCGCTGTGCGTA  
 CCCACTCCGACCTGAAGGTCGGCGAGTACTCCAACCTGGAGGTCAACGCTTCTGGACCAAGGAGAGGT  
 GGAACCAGGAGTTCACCAAGCACCAGGTGCTGATCATGACTTGTACGTGGCTCTGAACGTGCTGAAG  
 AACGGTTACCTGTCCCTGTCCGACATCAACCTGCTGGTGTTCGACGAGTGCCACCTGGCTATCCTGGAC  
 CACCCCTACCGGAGATCATGAAGCTGTGCGAGAAGTGGCCCTCCTGCCCCGATCCTGGGCCTGACC  
 GCTTCCATCCTGAACGGCAAGTGGACCCCCGAGGAGCTGGAGGAGAAGATCCAGAAGCTGGAGAAGAT  
 CCTCAAGTCCAACGCTGAGACTGCTACCGACCTGGTGGTGTGGACCGTTACACCTCCAGCCCTGCGA  
 GATCGTGGTGGACTGCGGTCCCTTACCAGCCGTTCCGGCCTGTACGAGCGTCTGCTGATGGAGCTGGA  
 GGAGGCTCTGAACCTCATCAACGACTGCAACATCTCCGTGCACTCCAAGGAGCGTGACTCCACCCTGAT  
 CTCCAAGCAGATCCTGTCCGACTGCCGTGCTGTGCTGGTCTGCTGGGCCCTTGGTGCCTGACAAGGT  
 GGCCGGCATGATGGTCCGCGAGCTGCAGAAGTACATCAAGCAGGAGCAGGAGGAGCTGCACCGCAAGT  
 TCTTGCTGTTACCGACACATTCCTCCGCAAGATCCACGCTCTGTGCGAGGAGCACTTCTCCCCGCCTC  
 CCTGGACCTGAAGTTCGTGACCCCCAAGGTCATCAAGCTGCTGGAGATCCTGCGCAAGTACAAGCCCTA  
 CGAGCGTCAGCAGTTCGAGTCCGTGGAGTGGTACAACAACCGTAACCAGGACAACCTACGTGTCTTGGT  
 CCGACTCCGAGGACGACGACGAGGACGAGGAGATCGAGGAGAAGGAGAAGCCCGAGACTAACTTCCCC  
 TCCCCCTTACCAACATCCTGTGCGGTATCATCTTCGTGGAGCGTCGTTACACCGCTGTGGTGTGAAC  
 CGCCTGATCAAGGAGGCTGGCAAGCAGGACCCCGAGCTGGCTTACATCTCCTCCAACCTCATCACCGGT  
 CACGGTATCGGCAAGAACCAGCCCCGTAACAAGCAGATGGAGGCTGAGTTCGCAAGCAGGAGGAGGT  
 CCTGAGGAAGTTCGCGCTCACGAGACTAACCTGCTGATCGCTACCTCCATCGTGGAGGAGGGCGTGGA  
 CATCCCCAAGTGCAACCTGGTGTGCGTTCGACCTCCCCACCGAGTACCGTTCCTACGTGCAGTCCAA  
 GGGTCTGCTCGTGTCTCCATCTCCAACCTACATCATGCTGGCTGACACCGACAAGATCAAGTCTTCGA  
 GGAGGACCTCAAGACCTACAAGGCCATCGAGAAGATCCTGCGTAACAAGTGTCCAAGTCCGTGGACA  
 CCGGAGAGACTGACATCGACCCCGTGATGGACGATGACGACGTGTTCCCTCCCTACGTCTGCGTCCCG  
 ACGACGGTGGTCCCGTGTGACCATCAACACCGCTATCGGTACATCAACCGTTACTGCGCTCGTCTGC  
 CCTCCGACCCCTTACCCACCTGGCTCCCAAGTGCCGTACCCGCGAGCTGCCCCGACGGCACCTTCTACTC  
 TACCCTGTACCTCCCATCAACTCCCCCTGCGTGTTCATCGTGGGTCCCCCATGTCTGCGTGC  
 CTGGCTGAGAGGGTGGTGGCTCTGATCTGCTGCGAGAAGCTGCACAAGATCGGCGAGCTGGACGACCA  
 CCTGATGCCCGTGGGCAAGGAGACTGTCAAGTACGAAGAGGAGCTGGACCTCCACGACGAGGAGGAGA  
 CTTCCGTCCCGGTCTGTCGCGTTCACCAAGCGTCGTGAGTGTACCCCAAGGCTATCCCCGAGTGCCT  
 GCGCGACTCCTACCCCGTCCCGACAGCCCTGCTACCTGTACGTGATCGGCATGGTGTGACCACTCCT  
 CTGCCCCGACGAGCTGAACTTCCGTGTCGCAAGCTGTACCCCCCGAGGACACCACCCGTTGCTTCGGT  
 ATCCTGACCGCCAAGCCATCCCCCAGATCCCCACTTCCCCGTGTACACCCGTTCCGGCGAGGTCACCA  
 TCTCCATCGAGCTGAAGAAATCCGGTTCATGCTGTCCCTGCAGATGCTGGAGCTGATCACCCGCTGC  
 ACCAGTACATCTTCTCCACATCCTGCGTCTGGAGAAGCCTGCTCTGGAGTTCAAGCCCACCGACGCTG  
 ACTCCGCTTACTGCGTGTGCCCCTGAACGTGGTCAACGACTCCTCCACCCTGGACATCGACTTCAAGT  
 TCATGGAGGACATCGAGAAGTCCGAGGCTCGTATCGGTATCCCTCCACCAAGTACACCAAGGAGACTC  
 CCTTCGTTTTCAAGCTGGAGGACTACCAGGACGCTGTGATCATCCCCGTTACCGTAACTTCGACCAGC  
 CTCACCGTTTCTACGTGGCTGACGTGTACACCGACCTGACCCCCCTGTCCAAGTTCCTTCCCCGAGTA  
 CGAGACTTTCGCGGAGTACTACAAGACCAAGTACAACCTCGACCTGACCAACCTGAACCAGCCTCTGCT  
 GGACGTGGACCACACCTCCTCCCGTCTGAACCTCCTGACCCCTCGTCACTCAACCAGAAGGGCAAGGC  
 TCTGCCCTGTCTCCGCTGAGAAGCGCAAGGCTAAGTGGGAGTCTCTGCAGAACAAGCAGATCCTCGT  
 GCCCCGAGCTGTGCGCTATCCACCCATCCCCGCTTCCCTGTGGCGCAAGGCTGTGTGCTGCCCTCCATC  
 CTGTACCGTCTGCACTGCCTGCTGACCGCTGAGGAGCTGCGTGTGCTCAGACCGCTTCCGACGCTGGTGT  
 GGTGTCCGTTCCCTGCCCGCTGACTTCCGTTACCCCAACCTGGACTTCCGTTGGAAGAAGTCCATCGAC

TCCAAGTCCTTCATCTCCATCTCTAACTCCTCCAGCGCTGAGAACGACAACACTACTGCAAGCACTCCACC  
 ATCGTGCCTGAGAACGCTGCTCACCAGGGTGCTAACCGTACCTCCTCCCTGGAGAACCACGACCAGATG  
 TCCGTGAAGTGGCGTACCCTGCTGTCCGAGTCCCCCGCAAGCTGCACGTGGAGGTGTCCGCTGACCTG  
 ACCGCTATCAACGGCCTGTCTACAACCAGAACCTGGCTAACGGTTCCTACGACCTGGCCAACCGTGAC  
 TTCTGCCAGGGCAACCAGCTGAACTACTACAAGCAGGAGATCCCCGTCCAGCCCACCACCTCCTACTCC  
 ATCCAGAACCTGTACTCCTACGAGAACCAGCCACAGCCCTCCGACGAGTGCACCTGTCTGTCCAACAAG  
 TACCTGGACGGCAACGCTAACAAGTCCACCTCCGACGGTTCCTCCGTCATGGCTGTGATGCCCGGCACC  
 ACCGACACCATCCAGGTCTGAAGGGTCGTATGGACTCCGAGCAGTCCCCCTCCATCGGTTACTCCTCC  
 CGTACCCTGGGTCCCAACCCCGGCCTGATCCTGCAGGCCCTGACCCTGTCCAACGTTCTGACGGTTTCA  
 ACCTGGAGCGCCTGGAGATGCTGGGCGACTCCTTCTGAAGCACGCTATCACCACCTACCTGTTCTGCA  
 CCTACCCCGACGCTCACGAGGGTGTCTGTCTACATGCGTTCGAAGAAGGTGTCCAACCTGCAACCTGT  
 ACCGCTGGGCAAGAAGAAGGGCCTGCCTTCCCGTATGGTGGTGTCCATCTTCGACCCCCCGTGAAC  
 GGCTGCCCCCTGGCTACGTCGTGAACCAGGACAAGTCCAACACCGACAAGTGGGAGAAGGACGAGATG  
 ACCAAGGACTGCATGCTGGCCAACGAAAGCTGGACGAGGACTACGAAGAAGAGGACGAAGAGGAAGA  
 ATCCCTGATGTGGCGTGTCCCAAGGAGGAGGCGGACTACGAGGACGACTTCTGGAGTACGACCAGG  
 AGCACATCCGTTTCATCGACAACATGCTGATGGGTTCGGTGTCTTCGTGAAGAAGATCTCCCTGTCCC  
 CCTTCTCCACCACCGACTCCGCCTACGAGTGAAGATGCCAAGAAGTCCCTCCCTGGGTTCATGCCCT  
 TCTTAGCGACTTCGAGGACTTCGACTACTCCTCCTGGGACGCTATGTGCTACCTGGACCCCTCCAAGG  
 CTGTGAGGAGGACGATTTTCGTCGTCGGTTTCTGGAACCCCTCCGAGGAGAACTGCGGCGTGCATACCG  
 GCAAGCAGTCCATCTCTTACGATCTCCACACTGAGCAGTGCATCGCTGACAAGTCTATCGCTGACTGCG  
 TGGAGGCTCTGCTGGGCTGTACCTGACCTCCTGCGGAGAGAGGGCTGCCAGCTGTTCTGTGCTCCC  
 TGGGACTGAAGGTGCTGCCCGTGTCAAGCGCACCGACCGGAGAAAGGCTCTGTGCCCCACCCGCGAGA  
 ACTTCAACTCCCAGCAGAAGAACCTGTCTGTCTCCTGCGCTGCTGCTTCCGTGGCTTCTCCAGGTCCT  
 CCGTGTGAAGGACTCCGAGTACGGTTGCCTGAAGATCCCCCTCGTTGCATGTTGACACCCTGACG  
 CTGACAAGACCCTGAACCACCTCATCTCCGGTTTCGAGAAGTTCGAGAAGAAGATCAACTACCGCTTCA  
 AGAACAAGGCTTACCTCCTGCAAGCCTTACCCACGCCTCTTACCACTACAACACTATCACCAGCTGCT  
 ACCAGGCTCTGGAGTTCCTCGGCGACGCCATCCTGGACTACCTGATCACTAAGCACCTGTACGAGGACC  
 CCCGTCAGCACTCCCCGGTGTCTGACCGACCTCCGTTCCGCTCTGGTCAACAACACTATCTTCGTTTC  
 CCTGGCTGTGAAGTACGACTACCACAAGTACTTCAAGGCTGTCTCCCCGAGCTGTTCCACGTGATCGA  
 CGACTTCGTGCAGTTCAGCTGGAGAAGAACGAGATGCAGGGCATGGACTCTGAGCTGCGTCGTTCCG  
 AGGAGGACGAGGAAAAGGAGGAGGACATCGAGGTCCCCAAGGCTATGGGTGACATCTTCGAGTCCCTG  
 GCCGGTGTATCTACATGGACTCCGGAATGTCCCTGGAGACTGTGTGGCAGGTCTACTACCCCATGATG  
 CGTCTCTGATCGAGAAGTTCCTCCGCTAACGTGCCCCGTTCCCTGTGCGTGAGCTGCTGGAGATGGAG  
 CCCGAGACTGCTAAGTTCCTCCCGCTGAGAGGACCTACGACGGCAAGGTCCGCGTGACCGTGGAGGTC  
 GTGGGAAAGGGCAAGTTCAGGGTGTGGTTCCTACCGTATCGCTAAGTCCGCTGCTGCTCGTCGT  
 GCTCTGCGTTCCTGAAGGCTAACCAGCCCCAGGTGCCAACTCCGCTTGGTCCCACCCCAAGTTCGAG  
 AAGTAA

**Table 3: DNA-oligonucleotides used in this study.**

The sequence of the DNA-oligonucleotides is shown in 5'-3'-direction and the melting temperatures given in the table are according to the manufacturer.

Primer	Sequence 5'→3'	T <sub>m</sub> [°C]
<b><i>H.s. Argonaute2</i></b>		
Ago2_iso_f	ATGTAAGTCCGGGAGCCGGCC	63.1
Ago2_iso_r	TCAAGCAAAGTACATGGTGC	60.3
Ago2_Bam_f	CGCGGATCCACCATGTAAGTCCGGGAGCCGGCC	>75.0
Ago_Hind_r	CCCAAGCTTTCAGCAAAGTACATGGTGCAGAGT	73.8
Ago2_601_f	CTGCTCTAACCTCTTGGCGGG	65.8
Ago2_1261_f	CGTACTGGGCGGGTGTGCAG	69.6
<b><i>H.s. TRBP2</i></b>		
TRBP2_iso-f	ATGAGTGAAGAGGAGCAAGGCTCC	64.4
TRBP2_iso-r	TCACTTGCTGCCTGCCATGATC	62.1

TRBP2_Xho_f	CCGCTCGAGACCATGAGTGAAGAGGAGCAA	72.1
T22_Bam_fw	CGCGGATCCATGCTGGCCGCAAC	71.3
T157_Bam_fw	CGCGGATCCGAGTGCAACCCGTTGG	72.7
T226_Bam-fw	CGCGGATCCCACACGGTGCCTCTGG	72.8
T291_Bam_fw	CGCGGATCCGGCCCTGCCTGCTGC	74.7
T99_Xho_rev	CCGCTCGAGTCAGCTCCGCCCTTTGAG	72.6
T227_Xho_rev	CCGCTCGAGTCACGTGTGCACTCGAAG	71.0
T366_Xho_rev	CCGCTCGAGTCACTTGCTGCCTGCCATG	72.4
<b>H.s. Dicer original sequence</b>		
Dicer_iso_f	ATGAAAAGCCCTGCTTTGCAACCC	62.7
Dicer_iso_r	TCAGCTATTGGGAACCTGAGG	59.8
Dcr_164_r	CAGAGCTGCTTCAAGCAGTTCAAC	62.7
Dcr_297_f	GGTCAACTCTGCAAACCAGTTGCTC	66.4
Dcr_601_f	CGCATTTTGGGACTAACTGCTTCC	62.7
Dcr_1174_f	CGACAGCAGTTTAAAAGCGTTGAGTGG	66.5
Dcr_1391_f	CTGGCAAACAAGATCCAGAGCTGGC	66.3
Dcr_2201_f	CCAGTGTTCAGGAAGACCAGTTCC	68.0
Dcr_3002_f	CACCTCGACATTTGAATCAGAAGGGG	64.8
Dcr_3821_f	GGATGGATTCTGAGCAGAGCCCTTC	66.3
Dcr_4641_f	CTTGACACTGAGCAGTGTATTGC	62.7
Dicer_5401_f	TCTGACAGCTGACACTTGTGAGC	62.7
Dcr_5680_f	GGTCGAAGTTACAGGATTGCCAAATCTGC	66.7
Dcr_Sall_f	TCTGTGCACCATGAAAAGCCCTGCTTTGCAACC	70.7
Dcr_Not_r	TTGCGGCCGCTCAGCTATTGGGAACCTG	71.0
Dcr_BglII_f	GGAAGATCTATGAAAAGCCCTGCTTTGCAACCC	69.5
<b>H.s. Dicer optimized sequence</b>		
D1_Bam_fw	CGCGGATCCATGAAGTCCCCGCTCTG	72.6
D_XbaI_rev	TCTAGATTACTTCTCGAACTGGGGG	63
D261_Bam_fw	CGCGGATCCACCGACCGTTCGGGCC	74.5
D264_Bam_fw	CGCGGATCCTCCGGCCTGTACGAGCG	84.2
D414_Xho_rev	CCGCTCGAGTCAGTCGGACCAAGACACGTAGTTG	82.1
D415_Xho_rev	CCGCTCGAGTCAGGAGTCGGACCAAGACACGT	83.4
D432_Xho_rev	CCGCTCGAGTCAGTTAGTCTCGGGCTTCTC	72.2
D439_Xho_rev	CCGCTCGAGTCAGTTGGTGAAGGGGGAG	72.2
SeqD276_fw	GGCAAGCGTACCGTGTTCCTG	63.7
SeqD366_rev	GCCGACCTTCAGGTCGGAG	63.1
SeqD972_fw	ATGGTCCGCGAGCTGCAGAAG	63.7
SeqD1683_fw	ATGCTGGCTGACACCGACAAG	61.8
SeqD2322_fw	ATGGTGTGACCACTCCTCTG	61.8
SeqD3024_fw	GGCAAGGCTCTGCCCCTG	62.8
SeqD3594_fw	GGCAACCAGCTGAACTACTACAAG	62.7
SeqD4203_fw	ATGACCAAGGACTGCATGCTGG	62.1
SeqD4758_fw	GGACTGAAGGTGCTGCCCCG	63.1
SeqD5343_fw	ATGCAGGGCATGGACTCTGAG	61.8
<b>Plasmids</b>		
pGEX_f	GCTGGCAAGCCACGTTTGGT	72.4
pGEX_r	CGTCTCCGGGAGCTGCATGT	74.4
M13_f	GTTTTCCAGTCACGAC	52.8
M13_r	CAGGAAACAGCTATGAC	50.4
pFastBac_fw	GTCCGAAACCATGTCGTAC	56.7
pFastBac_rev	CCTCTACAAATGTGGTATGG	55.3
T7_fw	TAATACGACTCACTATAGGG	50.8
T7_rev	GCTAGTTATTGCTCSGCGG	60.1

**Table 4: Initial crystallization conditions of the protein-protein complexes, protein-RNA complexes, proteins and protein fragments mentioned in this thesis are listed:**

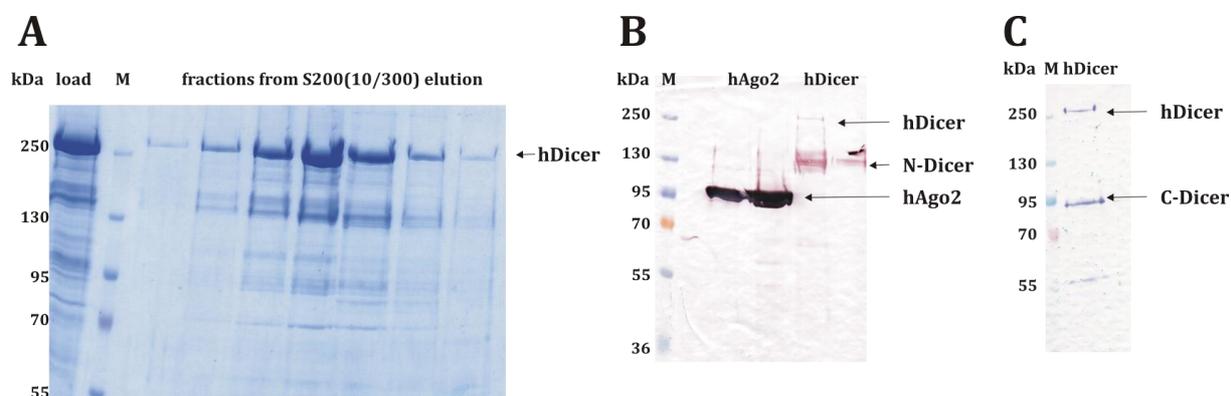
<b>Protein/complex</b>	<b>Buffer</b>	<b>Protease</b>	<b>Crystallization Screens</b>	<b>Temp.</b>
1.5 mg/ml <b>RLC</b> (N-hDicer-N-hAgo2-hTRBP2-complex)	150 mM NaCl, 20mM HEPES 7.8	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG; ProPlex; PGA; JB Nuc-Pro HTS; Ammoniumsulfate;	20 °C and 4°C
1.0 mg/ml <b>RLC</b> (N-hDicer-N-hAgo2-hTRBP2-complex)	150 mM NaCl, 20mM HEPES 7.8, 5 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); ProPlex; PGA;	20°C and 4°C
2 mg/ml <b>RLC</b> (CStrep-hDicer-hAgo2-hTRBP2-complex)	150 mM NaCl, 20mM HEPES 7.8, 5 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JB Nuc-Pro HTS; Ammoniumsulfate; ProPlex; PGA	20°C and 4°C
1.5 mg/ml <b>RLC</b> (N-hDicer-N-hAgo2-hTRBP2-complex)	150 mM NaCl, 20 mM Imidazol pH 6.8, 2% (v/v) (v/v) Glycerol, 1 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG ; JB Nuc-Pro HTS; ProPlex; PGA; Ammoniumsulfate; Natrinx	20°C and 4°C
2.4 mg/ml <b>RLC-siRNA duplex1-complex</b> (NHis <sub>6</sub> -Cstrep-hDicer-Nhis <sub>6</sub> -hAgo2-hTRBP2-siRNA duplex1-complex)	150 mM KCl, 20mM HEPES 7.5, 5% (v/v) Glycerol, 2 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG ; JB Nuc-Pro HTS; ProPlex; PGA	4 °C
1.8 mg/ml <b>RLC-siRNA duplex1-complex</b> (Nhis <sub>6</sub> -Cstrep-hDicer-Nhis <sub>6</sub> -hAgo2-hTRBP2-siRNA duplex1-complex)	150 mM KCl, 10 mM HEPES 7.8, 2% (v/v) Glycerol, 2 mM DTT	none	JB Nuc-Pro HTS; Morpheus, Natrinx	4°C and 20°C
0.5 mg/ml <b>RLC-siRNA duplex1-complex</b> (N-hDicer-N-hAgo2-hTRBP2-siRNA duplex1-complex -GraFix prepared)	150 mM KCl, 20 mM HEPES pH 7.5, 5% (v/v) Glycerol, 2 mM MgCl <sub>2</sub> , 5 mM DTT	none	ProPlex; JB Nuc-Pro HTS; Morpheus	4°C and 20°C
1.3 mg/ml <b>RLC</b> (CStrep-hDicer-hAgo2-hTRBP2) + <b>siRNA duplex1-complex</b> (1:1.3)	200 mM NaCl, 20 mM HEPES pH 7.5, 5% (v/v) Glycerol, 2 mM MgCl <sub>2</sub> , 5 mM DTT	none	ProPlex; JB Nuc-Pro HTS	4°C and 20°C
1.7 mg/ml <b>N-hDicer-hTRBP2 complex</b>	150 mM KCL, 20 mM HEPES 7.5, 5% (v/v) Glycerol, 2 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG ; JB Nuc-Pro HTS; ProPlex; PGA; Ammoniumsulfate;	4°C and 20°C
2 mg/ml <b>N-hDicer-hTRBP2 complex</b>	150 mM KCL, 20 mM HEPES 7.5, 5% (v/v) Glycerol, 2 mM DTT	Chymo-trypsin (1:1000)	MIDAS, ProPlex; JCSG; Morpheus	20°C
1.2 mg/ml <b>N-hDicer-hTRBP2 complex</b>	150 mM KCL, 20mM HEPES 7.5, 5% (v/v) Glycerol, 2 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG ; JB Nuc-Pro HTS; ProPlex; PGA; Ammoniumsulfate;	20°C
1 mg/ml <b>CStrep-hDicer-hTRBP2 complex</b>	150 mM KCL, 20mM HEPES 7.5, 5% (v/v) Glycerol, 2 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG ; JB Nuc-Pro HTS; ProPlex; PGA; Ammoniumsulfate;	4°C
3 mg/ml <b>N-hAgo2</b>	150 mM NaCl, 20mM HEPES 7.5, 5% (v/v) Glycerol, 2 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); Manual screening: Footprint Screen 1-3; Crystal Screen 1&2; Crystal Screen lite 1&2;	20°C

			Crystal Screen PEG/ION; Magic Screen 1-3; Structure	
6 mg/ml <b>N-hAgo2</b>	150 mM NaCl, 20mM HEPES 7.5, 5% (v/v) Glycerol, 2 mM DTT	none	Manual screening: Footprint Screen 1-3; Crystal Screen 1&2	20°C
15 mg/ml, 10 mg/ml, 7 mg/ml and 5 mg/ml <b>hTRBP2</b>	150mM NaCl, 20 mM HEPES pH 7.5, 5% (v/v) Glycerol, 2 mM DTT	none	Manual screening: Footprint Screen 1-3	20°
7 mg/ml <b>hTRBP2</b>	150mM NaCl, 20 mM HEPES pH 7.5, 5% (v/v) Glycerol, 2 mM DTT	none	Manual screening: Footprint Screen 1-3; Crystal Screen 1&2; Crystal Screen lite 1&2; Crystal Screen PEG/ION; Magic Screen 1-3; Structure Screen	20° and 4°C
7.5 mg/ml <b>hTRBP2</b>	150mM NaCl, 20 mM HEPES pH 7.5, 5% (v/v) Glycerol, 2 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG; ProPlex;	20° and 4°C
4 mg/ml <b>hTRBP2</b>	150 mM NaCl, 20 mM HEPES, 2 % (v/v) Glycerol	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG; ProPlex;	20° and 4°C
5 mg/ml <b>hTRBP2 + siRNA-duplex 1 (1:1.1)</b>	150 mM NaCl, 20 mM HEPES 7.5, 2% (v/v) Glycerol, 2 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG; ProPlex;	20° and 4°C
4 mg/ml <b>hTRBP2-siRNA + duplex 1 (1:1.1)</b>	150mM NaCl, 20 mM HEPES 7.5, 2 % (v/v) Glycerol, 2 mM DTT	none	Ammoniumsulfate;	20°C and 4°C
4 mg/ml, 5 mg/ml and 6 mg/ml <b>hTRBP2 + siRNA-duplex 1</b> (each in ration 1:1.1; 1:1.2 and 1:1,3)	50 mM NaCl, 20 mM Tris, 7.5	none	JB Nuc-Pro HTS	20°C and 4°C
7.35 mg/ml <b>hTRBP2-siRNA-duplex 1-complex</b>	100 mM NaCl, 20 mM Tris pH 7.5, 2 mM MgCl <sub>2</sub>	none	JB Screens (1,2,4,5); JB Sscreens (6,7,8,10); JCSG; ProPlex; PGA; JB Nuc-Pro HTS; Ammoniumsulfate;	20°C and 4°C
3.8 mg/ml <b>hTRBP2-siRNA-duplex 1-complex</b>	150 mM KCl, 10 mM HEPES 7.8, 2% (v/v) Glycerol, 2 mM DTT	none	JB Nuc-Pro HTS; MIDAS MD; MIDAS-MD diluted; Natrrix HT; Morpheus, JCSG	20°C and 4°C
5 mg/ml <b>T1</b> (hTRBP2 aa 157-366)	150 mM NaCl, 20 mM Tris pH 7.5, 2 % (v/v) Glycerol, 1 mM DTT	none	Manual screening: Footprint Screen 1-3; Structure Screens 1-2; Crystal Screens lite 1-2; Magic Screens 1-3; JB Screens 1-10; Crystal Screen Natrrix	20°C
5.5 mg/ml <b>T1</b> (hTRBP2 aa 157-366)	150 mM NaCl, 20 mM Tris pH, 2% Glycerin, 2 mM DTT	Chymo-trypsin (1:500)	ProPlex; Nuc-Pro HTS; JCSG; Ammoniumsulfate; Morpheus; Natrrix HT; JB_1-2-4-5; JB_6-7-8-10; PGA	4°C and 20°C
2.6 mg/ml <b>T1+siRNA-duplex 1-complex</b> (hTRBP2 aa 157-366-siRNA duplex1-complex)	150 mM KCl, 10 mM Imidazol pH 7.0, 2 mM MgCl <sub>2</sub> , 1 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG; ProPlex; PGA; JB Nuc-Pro HTS; Ammoniumsulfate; Natrrix, Morpheus	20°C
5 mg/ml <b>T2</b> (hTRBP2 aa 226-366)	150 mM NaCl, 20 mM Hepes pH 7.0, 2 % (v/v) Glycerol, 2 mM EDTA;1 mM DTT	none	Manual screening: Footprint Screen 1-3	20°C
2.5 mg/ml <b>T2</b> (hTRBP2	150 mM NaCl, 20 mM	none	Manual screening: Footprint	20°C

Appendix

aa 226-366)	HEPES pH 7.0, 2 % (v/v) Glycerol, 2 mM EDTA;1 mM DTT		Screen 1-3; Structure Screens 1-2; Crystal Screens 1-2; Crystal Screens lite 1-2; Magic Screens 1-3; JB Screens 1-10; Crystal Screen Matrix	
5.0 mg/ml <b>T2</b> (hTRBP2 aa 226-366) + 0.1 M Lys/Cys	150 mM NaCl, 20 mM HEPES pH 7.0, 2 % (v/v) Glycerol, 2 mM EDTA;1 mM DTT	none	Manual screening: Footprint Screen 1-3; Structure Screens 1-2; Crystal Screens 1-2; Crystal Screens lite 1-2; Magic Screens 1-3; JB Screens 6, 9, 10	20°C
5.0 mg/ml <b>T2</b> (hTRBP2 aa 226-366)	150 mM NaCl, 20 mM Tris 7.5, 2% Glycerin, 2 mM DTT	Chymo- trypsin (1:500)	Ammoniumsulfate	20°C
5.0 mg/ml <b>T3</b> (hTRBP2 aa 291-366)	150 mM NaCl, 20 mM HEPES pH 7.0, 2 % (v/v) Glycerol, 2 mM EDTA;1 mM DTT	none	Manual screening: Footprint Screen 1-3; Structure Screens 1-2; Crystal Screens 1-2; Crystal Screens lite 1-2; Magic Screens 1-3; JB Screens 1-10; Crystal Screen Matrix	20°C
2.5 mg/ml <b>T3</b> (hTRBP2 aa 291-366)	150 mM NaCl, 20 mM HEPES pH 7.0, 2 % (v/v) Glycerol, 2 mM EDTA;1 mM DTT	none	Manual screening: Footprint Screen 1-3; Structure Screens 1-2; Crystal Screens 1-2; Crystal Screens lite 1-2; Magic Screens 1-3; Crystal Screen Matrix; JB Screens 1-10	20°C
4 mg/ml <b>T3</b> (hTRBP2 aa 291-366)	150 mM NaCl, 20 mM Tris 7.5, 2% Glycerin, 2 mM DTT	Chymo- trypsin (1:500)	Ammoniumsulfate	20°C
4 mg/ml <b>DTB1-T1 complex</b> (hDicer aa 261-432- hTRBP2 aa 157-366 -complex)	150 mM NaCl, 20 mM Tris pH 7.5, 2 % (v/v) Glycerol, 1 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG; ProPlex; PGA; JB Nuc-Pro HTS; Ammoniumsulfate;	20°
2 mg/ml and 3 mg/ml <b>DTB1-T1 complex</b> (hDicer aa 261-432- hTRBP2 aa 157-366 -complex)	200 mM KCL, 20mM HEPES 7.5, 5% (v/v) Glycerol, 2 mM DTT	none	MIDAS Screen; MIDAS Screen 1:1 diluted	20° and 4°
1.2 mg/ml <b>DTB1-T1 complex</b> (hDicer aa 261-432- hTRBP2 aa 157-366 -complex)	150 mM NaCl, 20 mM Tris pH, 2% Glycerin, 2 mM DTT	Chymo- trypsin (1:100)	ProPlex; Nuc-Pro HTS; JCSG; Ammoniumsulfate; Morpheus; Natrix HT	20°C
2 mg/ml <b>DTB1-T1 complex</b> (hDicer aa 261-432- hTRBP2 aa 157-366 -complex)	50 mM KCl, 20 mM HEPES 7.5, 2% Glycerin 2 mM DTT	Carboxy- peptidase A (1:100)	ProPlex; Morpheus, JCSG; JB_1-2-4-5	20°C
2.5 mg/ml <b>DTB1-T1 complex</b> (hDicer aa 261-432- hTRBP2 aa 157-366 -complex)	200 mM KCl, 20 mM HEPES pH 7.5, 5% Glycerin, 2 mM DTT	Thermo- lysin (1:100)	ProPlex; Morpheus,; JB_1-2-4-5; Ammoniumsulfate;	20 °C
2.2 mg/ml <b>DTB1-T1-siRNA duplex1-complex</b> (hDicer aa 261-432- hTRBP2 aa 157-366 - siRNA duplex1 complex)	150 mM KCl, 10 mM Imidazol pH 7.0, 2 mM MgCl <sub>2</sub> , 1 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); ProPlex; JB Nuc-Pro HTS; Morpheus; PGA	20°C
3.38 mg/ml <b>DTB1-T3 complex</b> (hDicer aa 261-432- hTRBP2 aa 291-366 -complex)	150mM NaCl, 20 mM Tris 7.5, 2% (v/v) Glycerol, 2 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG; ProPlex; JB Nuc-Pro HTS; Ammoniumsulfate; PGA	20°C and 4°C

3.38 mg/ml <b>DTB1-T3 complex</b> (hDicer aa 261-432- hTRBP2 aa 291-366 -complex)	150mM NaCl, 20 mM Tris 7.5, 2% (v/v) Glycerol, 2 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG; ProPlex; JB Nuc-Pro HTS; Ammoniumsulfate; PGA	20°C
2.5 mg/ml <b>DTB1-T3 complex</b> (hDicer aa 261-432- hTRBP2 aa 291-366 -complex)	150 mM NaCl, 20 mM HEPES pH 7.5, 3% Glycerin, 2 mM DTT	Carboxy-peptidase A (1:100)	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG; MIDAS; MIDAS diluted;; JB Nuc-Pro HTS; PGA; ProPlex; Morpheus; Natrx HT; Ammoniumsulfate	20 °C
3.87 mg/ml <b>DTB2-T1 complex</b> (hDicer aa 261-439- hTRBP2 aa 157-366 -complex)	150 mM NaCl, 20 mM Tris pH 7.5, 2 % (v/v) Glycerol, 1 mM DTT	none	JB Screens (1,2,4,5); JB Screens (6,7,8,10); JCSG; ProPlex; JB Nuc-Pro HTS; Ammoniumsulfate; PGA	20°



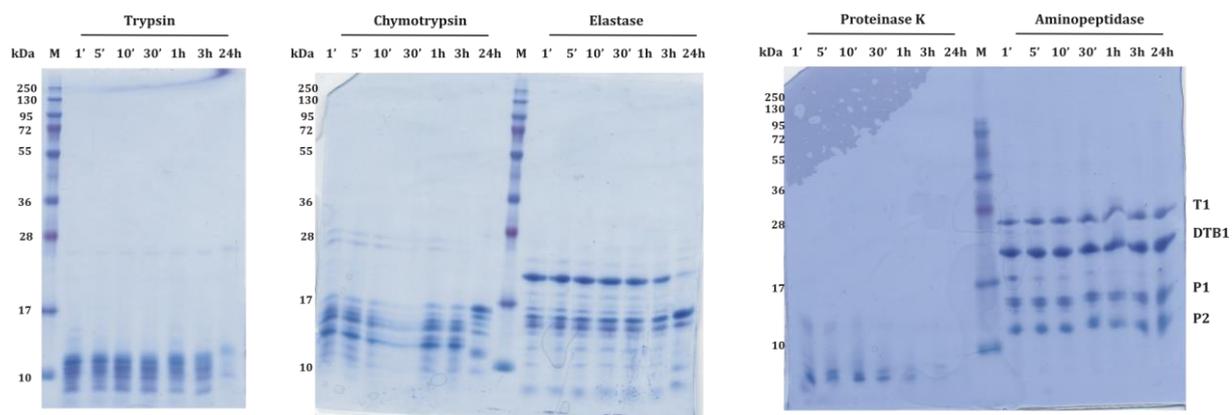
**Figure 47: Degradation analysis of human Dicer.**

**A:** 6% SDS-PAGE analysis of N-hDicer Superdex 200 (10/300) elution fractions. The full length N-hDicer (hDicer) protein is denoted on the right of the gel, size marker in kDa in shown in the left (M= marker). **B:** Anti-His Western blot of N-hAgo2 (hAgo2) and N-hDicer (hDicer). The identified proteins are denoted on the right of the gel, whereby N-Dicer is an N-terminal fragment of human Dicer of about 130 kDa. The size marker in kDa in shown in the left (M= marker). **C:** Anti-Strep western blot of N-hDicer (hDicer). The identified proteins are denoted on the right of the gel, whereby C-Dicer is a C-terminal fragment of hDicer of about 100 kDa. The size marker in kDa in shown in the left (M= marker).

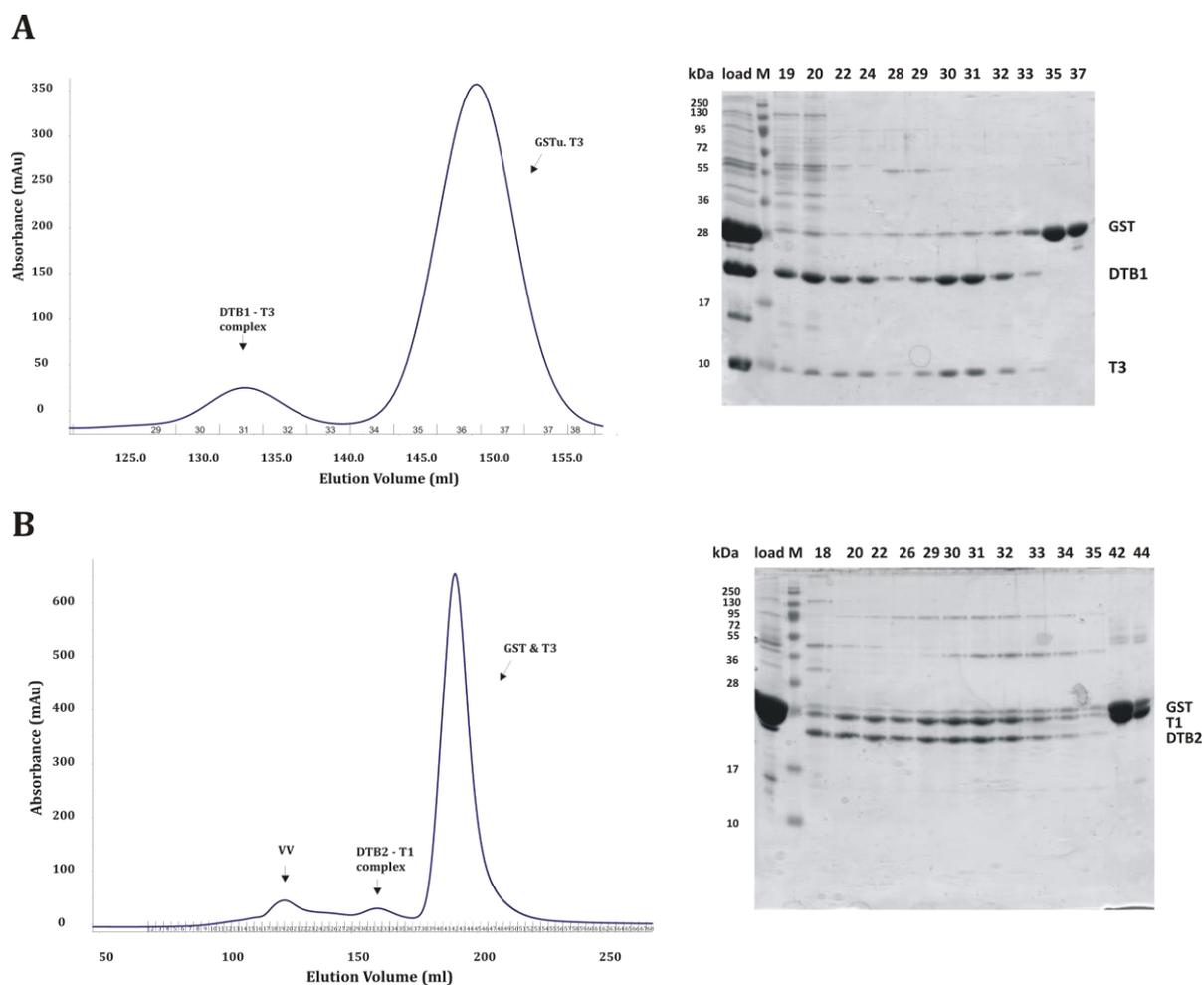
# Appendix



**Figure 48: Comparison of back-projections of the reconstituted RLC to corresponding 2D class sums.**

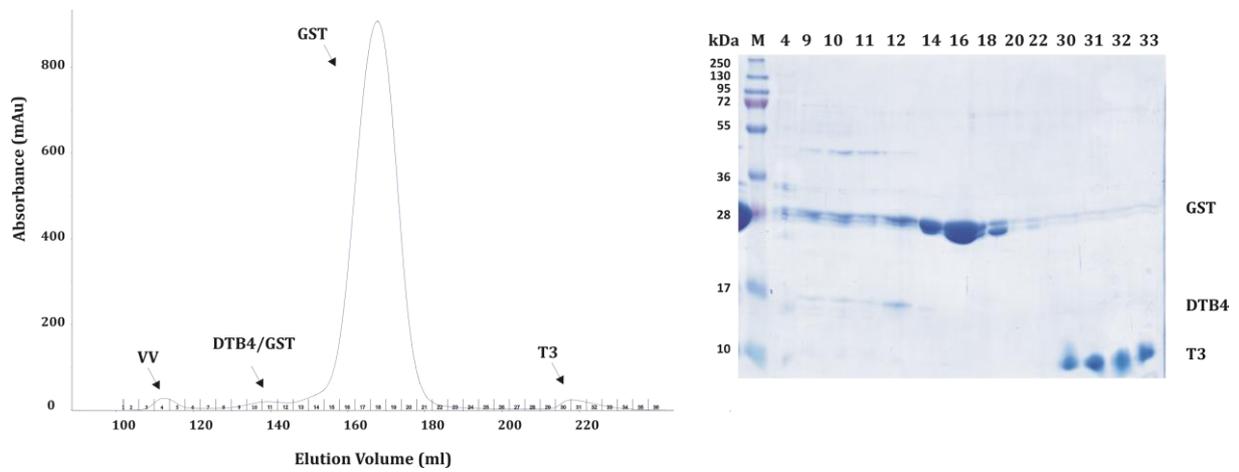


**Figure 49: Limited proteolysis of DTB1-T1 complex.** Limited proteolysis of the DTB1-T1 complex with Trypsin, Chymotrypsin, Elastase, Proteinase K and Aminopeptidase in a molar ratio of 100:1.

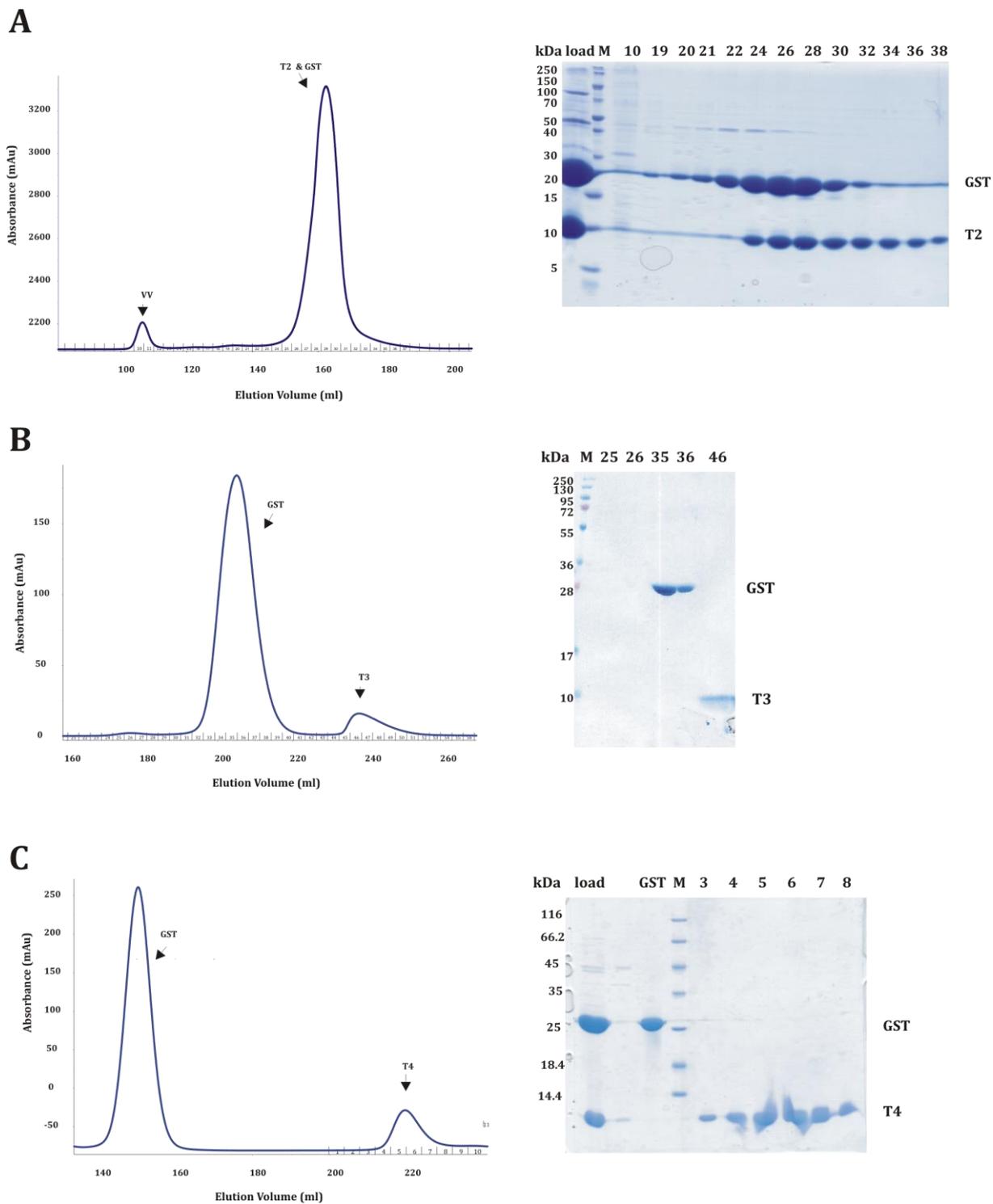


**Figure 50: Purification of DTB1-T3 (A) and DTB2-T1 complexes (B).** A: Superdex 75 (26/60) elution profile of the DTB1-T3-complex and 15% SDS-PAGE analysis of the peak fractions of the SEC run. B: Superdex 75 (26/60) elution profile of the DTB2-T1-complex and 15% SDS-PAGE analysis of the peak fractions of the SEC run (M = marker, VV = void volume).

## Appendix

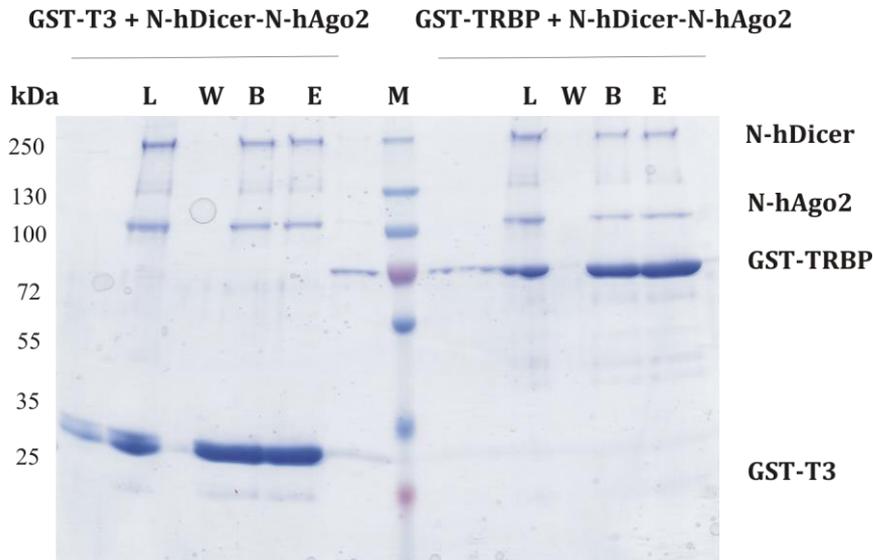


**Figure 51: SEC with subsequent SDS-PAGE analysis of the DTB4-T3 proteins.** Superdex 75 (26/60) elution profile of the DTB4-T1 proteins with subsequent 15% SDS-PAGE analysis of the peak fractions. Numbers on top of the gel correspond to the fractions of the SEC run (M = marker, VV = void volume).

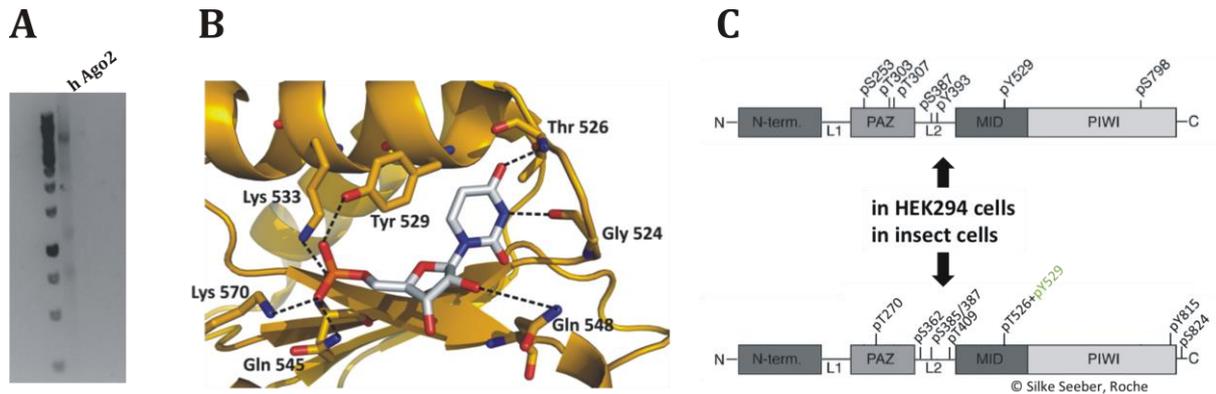


**Figure 52: Purification of hTRBP2 fragments T2 (A), T3 (B) and T4 (C).**

Superdex 75 (26/60) elution profile of the T2 (A), T3 (B) and T4 (C) proteins with subsequent 15% SDS-PAGE analysis of the peak fractions. Numbers on top of the gels correspond to the fractions of the SEC runs (M = marker, VV = void volume).

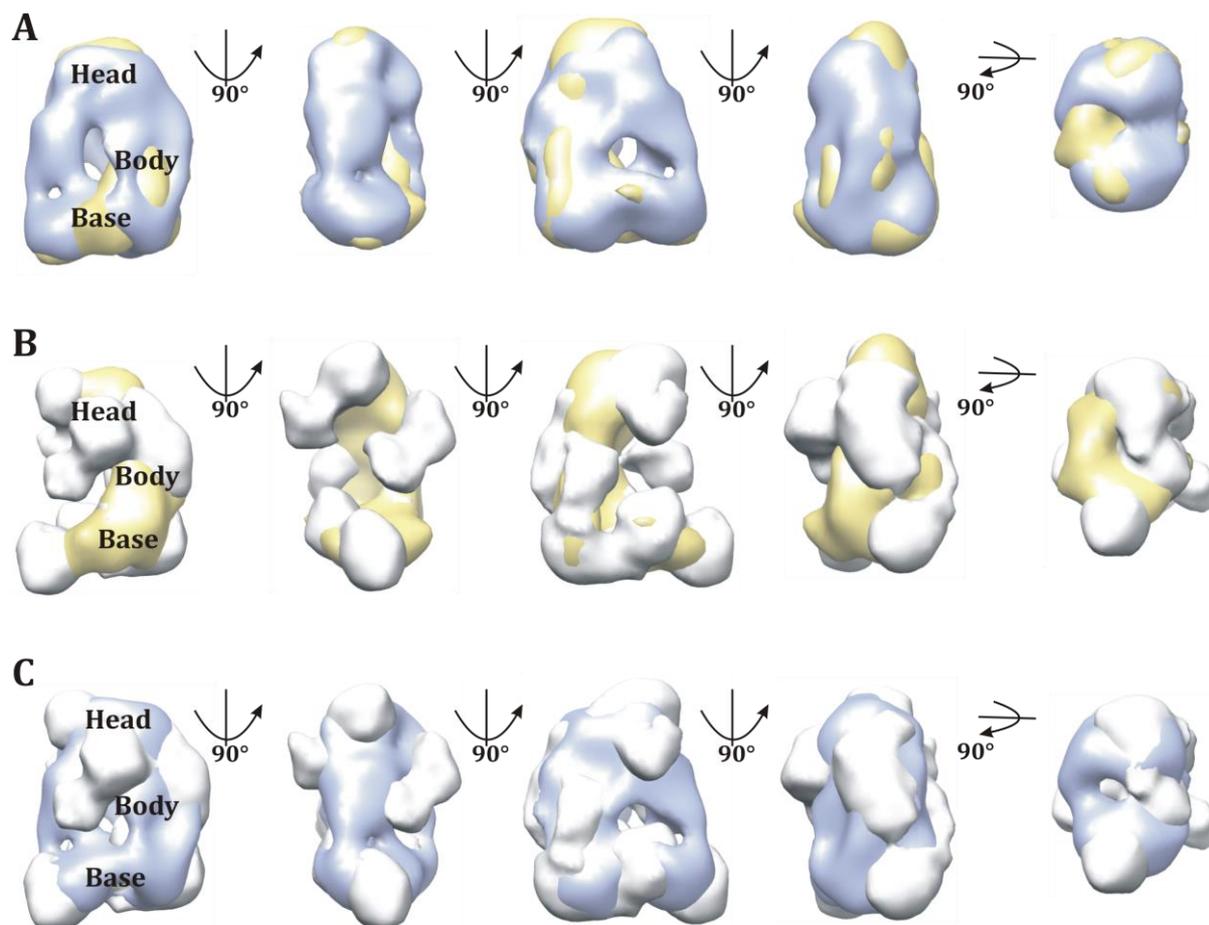


**Figure 53: Pull-down assay of GST-hTRBP2 and GST-T3-fragment with N-hDicer-N-hAgo2 complex.** (L= load, W= wash, B= beats before elution, E= elution, M= Marker).



**Figure 54: Analysis of human Ago2.**

**A:** Urea-PAGE of hAgo2 purified from baculovirus infected High Five cells (left lane = marker, right lane = hAgo2). **B:** Specific 5'-nt binding pocket of hAgo2 MID domains with bound UTP (modified from Frank *et al.*, 2011). The residues, which are important for nucleotide binding are shown as sticks. **C:** Comparison of postranslational phosphorylations of hAgo2 purified either from HEK or High Five cells. (modified from Rüdél *et al.*, 2010)



**Figure 55: Superposition of different reconstructed RLC complexes and the hDicer-hTRBP2 complex.** The RLC obtained by Wang *et al.*, 2009 is colored in purple, the Dicer-TRBP2-reconstruction in yellow and the RLC of this study in light grey. (modified from Wang *et al.*, 2009b, and Lau *et al.*, 2009, Lau *et al.*, 2012a). **A:** Superposition of RLC obtained by Wang *et al.*, 2009 with hDicer-hTRBP2 complex. **B:** Superposition of RLC obtained in this study with hDicer-hTRBP2 complex. **C:** Superposition of RLC obtained by Wang *et al.*, 2009 with RLC obtained in this study.



## 11 Curriculum Vitae

### Stephanie Schell

Georg-August-Universität, Göttingen  
 Department of Structural Biology  
 Justus-von-Liebig Weg 11  
 Göttingen, Germany

Phone: +49 551 3914074; +49 551 3914190

E-mail: sshell1@gwdg.de

### Personal Details

Date of birth	June 16 <sup>th</sup> 1984
Place of birth	Nordhausen
Nationality	German

### Education

04/2008-03/2013	PhD student, Georg-August Universität Göttingen, department of structural biology, Structural characterization of the minimal human RISC-loading complex. Thesis committee: Prof. Dr. Ralf Ficner, Prof. Dr. Holger Stark, Prof. Dr. Kai Tittmann
01/2008	Degree: Diploma
01/2007-09/2007	Diploma thesis, Georg-August Universität Göttingen, department of structural biology: „Präparation, Kristallisation und Strukturanalyse der Poly(A)-spezifischen Ribonuklease PARN“
09/2004-07/2005	Erasmus student at the University of Birmingham, UK
07/2004	Degree: Pre-Diploma
10/2002-01/2008	Diploma study: Biology at the Georg-August Universität, Göttingen
06/2002	Abitur
08/1994-06/2002	Fridrich-Schiller Gymnasium, Bleicherode
08/1990-06/1994	Adolf-Diesterweg Grundschule, Niedergebra

### Publications

Monecke, M.\*, Schell, S.\*, Dickmanns, A., Ficner, R., **2008**. Crystal Structure of the RRM Domain of Poly(A)-Specific Ribonuclease Reveals a Novel m<sup>7</sup>G-Cap-Binding Mode. *Journal of Molecular Biology*, 382 (4):827-834.

\* contributed equally to this work