

Université de Montréal

**Régulation de l'activité de Dicer par TRBP dans la  
biogénèse des micro-ARN**

par Jonathan Bouvette

Département de biochimie et médecine moléculaire  
Faculté de médecine

Thèse présentée  
en vue de l'obtention du grade de Doctorat  
en biochimie et médecine moléculaire

31 Juillet 2018

© Jonathan Bouvette, 2018

## Résumé

Les microARN (miARN) sont de petits ARN non codants ayant un rôle dans la régulation post-transcriptionnelle des gènes. Ils font partie du complexe de répression induit par les miARN (*RNA-induced silencing complex*; RISC) où ils servent de guide pour cibler des ARN messagers (ARNm) par complémentarité de séquence et inhibent leur traduction. Les miARN ont le potentiel de réguler la traduction de pratiquement toutes les protéines cellulaires, ce qui leur confère un rôle central dans plusieurs processus cellulaires, dont le développement de l'organisme.

Afin d'être fonctionnels, les miARN doivent passer par une voie de biogénèse hautement régulée où ils sont d'abord transcrits, puis clivés en deux temps par les endonucléases Drosha et Dicer et, finalement, chargés sur Argonaute (AGO) pour former le complexe RISC. Cette thèse s'intéresse à l'étape de clivage par Dicer et sa régulation par le cofacteur TRBP, protéine pouvant à la fois lier Dicer et son substrat.

Le premier objectif de cette thèse est de développer un nouveau protocole pour l'expression et la purification de la protéine Dicer humaine afin de faciliter sa caractérisation *in vitro* en améliorant les rendements limités des méthodes existantes. Nous avons optimisé l'expression de Dicer dans un système d'expression basé sur des cellules humaines HEK293-6E cultivées en suspension. Ensuite, nous avons développé un protocole de purification en 3 étapes permettant la purification de Dicer en une seule journée, où chacune des étapes a été soigneusement optimisée pour minimiser les pertes et maximiser la stabilité de Dicer. Ensuite, par des études biophysiques de SEC-MALS et microscopie électronique à coloration négative, nous avons confirmé que Dicer était bien repliée sous forme monomérique. Finalement, nous avons fait la caractérisation biochimique de Dicer en mesurant sa liaison à un pre-miARN par études de retard sur gel et des études de cinétique à l'état-stationnaire. Ces expériences de cinétique nous ont permis de déterminer de nouvelles valeurs de  $k_{cat}$  et  $K_M$  qui concordent avec un modèle impliquant un changement de conformation du complexe Dicer-substrat avant la réaction de clivage.

Le deuxième objectif est de caractériser l'effet de TRBP sur l'activité de Dicer. Pour ce faire, nous avons testé quelques pre-miARN en présence de TRBP et montré que cette dernière

pouvait être un activateur ou un inhibiteur de Dicer. En effet, TRBP active Dicer à faible concentration de substrat alors qu'elle inhibe Dicer à forte concentration de substrat. Nous avons ensuite caractérisé ce mécanisme enzymatique de manière détaillée par la combinaison d'expériences de liaison et de cinétique à l'état-stationnaire. Notre modèle final montre qu'une faible concentration de TRBP est nécessaire afin d'obtenir un effet d'activation à faible concentrations de substrat. Cela suggère que TRBP active Dicer sans modifier sa grande spécificité pour son substrat. Ces nouveaux résultats suggèrent que TRBP a évolué de manière à pouvoir activer efficacement Dicer sans modifier la spécificité de l'enzyme.

**Mots-clés:** Dicer, miARN, TRBP, régulation, purification, enzyme, cinétique.

## Abstract

Micro-RNAs (miRNA) are small non-coding RNAs that act as post-transcriptional gene regulators. They are part of the RNA-induced silencing complex (RISC) where they serve as guides to target mRNAs by sequence complementarity and inhibit their translation. Together, miRNAs can regulate the translation of virtually every protein in the cell. This confers miRNAs a central role in various cellular processes such as the development of organisms.

To be functional, miRNA must undergo maturation through a highly regulated pathway where they are first transcribed, then cleaved in two steps by endonucleases Drosha and Dicer and, finally, loaded onto the Argonaute protein (AGO) to form the RISC complex. This thesis focuses on the Dicer cleavage step and its regulation by the TRBP cofactor, which is a protein that can simultaneously bind Dicer and its pre-miRNA substrate.

The first objective of the thesis is to develop a new protocol for the expression and purification of human Dicer with the intent to improve the yields of existing methods and facilitate its *in vitro* characterization. We optimized Dicer expression in a human cell expression system that uses an engineered HEK93-6E cell line grown in suspension. Then, we developed a 3-step purification protocol allowing for Dicer purification in a single day. Moreover, these steps were carefully optimized to minimize losses and maximize Dicer stability. The purified Dicer was then analyzed using SEC-MALS and negative-stain EM to confirm that it was a well-folded monomer. Finally, we biochemically characterized Dicer by measuring its binding to a pre-miRNA substrate by EMSA and steady-state kinetic studies. Through these experiments, we calculated new values of  $k_{cat}$  and  $K_M$  that are consistent with a mechanism that involves a conformational change of the Dicer-substrate complex prior to the cleavage step.

The second objective was to characterize the effect of TRBP on Dicer activity. To do this, we first tested different pre-miRNA substrates with Dicer in presence of TRBP and showed that TRBP was both an activator and inhibitor of Dicer. Indeed, TRBP is an activator at low substrate concentration and an inhibitor at higher concentration. Next, we characterized the detailed mechanism of the effect of TRBP by using a combination of binding and steady-state kinetic experiments. Our final model shows that low TRBP concentration is necessary for activation of Dicer activity at low substrate concentration. Our model suggests that TRBP exerts

its activation function without reducing Dicer's specificity for its substrate. These new results suggest that TRBP evolved to efficiently activate Dicer without modifying the specificity of the reaction.

**Keywords:** Dicer, miRNA, TRBP, regulation, purification, enzyme, kinetics.

# Table des matières

Résumé .....	ii
Abstract.....	iv
Table des matières.....	vi
Liste des tableaux.....	xi
Liste des figures .....	xii
Liste des sigles .....	xiv
Liste des abréviations .....	xvi
Remerciements.....	xvii
1. Introduction.....	19
1.1 Biogénèse des microARN.....	19
1.1.1. Historique .....	19
1.1.2. Transcription.....	21
1.1.3. Clivage par Drosha.....	23
1.1.4. Exportation du pre-miARN vers le cytoplasme.....	25
1.1.5. Clivage par Dicer .....	26
1.1.6. Chargement des miARN sur Argonaute.....	26
1.1.7. Inhibition de la traduction par le complexe RISC .....	27
1.1.8. Autres voies de biogénèse .....	29
1.2. Régulation de la biogénèse des miARN .....	29
1.2.1. TRBP et PACT .....	29
1.2.2. Lin28 .....	30
1.2.3. hnRNP A1 et KSRP .....	33
1.2.4. Autres régulateurs et efforts récents.....	33
1.3. La famille let-7 de miARN.....	34
1.4. Dicer.....	35
1.4.1. Domaines .....	36
1.4.2. Reconnaissance de l'ARN par Dicer.....	39

1.4.3.	Autres rôles de Dicer et localisation cellulaire .....	41
1.5.	TRBP.....	42
1.5.1.	Domaines.....	42
1.5.2.	Fonction de TRBP dans la régulation de Dicer .....	43
1.5.3.	Interaction de TRBP avec les pre-miARN .....	44
1.5.4.	Interaction de TRBP avec Dicer .....	46
1.5.5.	Dimérisation de TRBP .....	47
1.5.6.	TRBP influence les produits de clivage de Dicer .....	48
1.6.	Études enzymatiques à l'état stationnaire .....	48
1.6.1.	Les travaux de Michaelis et Menten .....	49
1.6.2.	Approximation de l'état stationnaire.....	51
1.6.3.	Activation et inhibition enzymatique .....	52
1.6.4.	Dérivation des équations cinétiques.....	53
1.6.5.	Conditions expérimentales à l'état stationnaire .....	54
1.7.	Objectifs de la thèse.....	55
2.	Production de grandes quantités de Dicer humaine par transfection de cellules HEK293-EBNA1 en suspension.....	58
2.1.	Mise en contexte.....	58
2.2.	Abstract .....	61
2.3.	Background .....	63
2.4.	Results.....	65
2.4.1.	Transient transfection of recombinant human Dicer.....	65
2.4.2.	Purification of recombinant human Dicer .....	67
2.4.3.	Conformational characterization of the purified Dicer .....	70
2.4.4.	Binding and cleavage of pre-let-7a-1 by Dicer.....	73
2.5.	Discussion .....	76
2.6.	Conclusion.....	78
2.7.	Methods.....	79
2.7.1.	Plasmids.....	79
2.7.2.	Cell culture and transfection.....	80

2.7.3.	Optimization of transfection conditions .....	81
2.7.4.	Dicer purification .....	82
2.7.5.	Synthesis and purification of pre-let-7a-1 .....	83
2.7.6.	SEC-MALS/RI.....	84
2.7.7.	Negative stain TEM .....	84
2.7.8.	Dicer binding assay .....	85
2.7.9.	Dicer cleavage assay .....	86
2.8.	References .....	89
2.9.	Supporting figures .....	95
3.	Le double effet de TRBP sur le clivage des pre-miARN par Dicer .....	97
3.1.	Mise en contexte.....	97
3.2.	Abstract .....	99
3.3.	Introduction .....	99
3.4.	Material and methods.....	103
3.4.1.	Plasmids.....	103
3.4.2.	Synthesis and purification of pre-let-7 RNAs .....	103
3.4.3.	Protein expression and purification.....	104
3.4.4.	Steady-state kinetic studies in the absence of TRBP .....	105
3.4.5.	Effect of TRBP on the possible production of isoforms of let-7-5p.....	106
3.4.6.	Effect of TRBP on the relative cleavage efficiency of different pre-let-7 substrates by Dicer .....	106
3.4.7.	Evaluation of a potential pre-steady-state burst.....	107
3.4.8.	Steady-state kinetic studies in the presence of TRBP .....	107
3.4.9.	Binding assays for $K_d$ determination.....	108
3.5.	Results.....	109
3.5.1.	Binding studies of pre-let-7 to Dicer.....	109
3.5.2.	Steady-state kinetic studies of pre-let-7 cleavage by Dicer .....	109
3.5.3.	General effect of TRBP on pre-let-7 cleavage .....	112
3.5.4.	No evidence for a pre-steady-state burst .....	115



3.5.5.	Building a kinetic model for pre-miRNA cleavage by Dicer in the presence of TRBP .....	115
3.5.6.	Derivation of the kinetic equations .....	118
3.5.7.	Fitting the T12 subsection of the kinetic model .....	120
3.5.8.	Fitting the T3 subsection of the kinetic model .....	122
3.5.9.	Fitting the overall kinetic model.....	123
3.6.	Discussion .....	127
3.6.1.	Similar thermodynamic and kinetic properties of pre-let-7 substrates for Dicer .. .....	127
3.6.2.	A dual role for TRBP in regulating pre-let-7 cleavage by Dicer.....	128
3.6.3.	The four phases of the kinetic model for TRBP modification .....	131
3.6.4.	Dicer cleavage under physiological substrate concentrations .....	133
3.6.5.	The kinetic model resolves previous and apparently conflicting observations .... .....	133
3.7.	Conclusion.....	134
3.8.	References .....	135
3.9.	Supplementary information.....	141
4.	Discussion .....	146
4.1.	Purification de Dicer .....	146
4.2.	Coopérativité de liaison entre Dicer et ses substrats.....	148
4.3.	Futur des études structurales de Dicer .....	150
4.4.	Modèle de modification de Dicer par TRBP.....	150
4.4.1.	Implication biologique des paramètres cinétiques .....	150
4.4.2.	Méthodes alternatives pour les études cinétiques .....	152
4.4.3.	Validation <i>in vivo</i> du modèle d'activation de Dicer par TRBP .....	153
4.4.4.	Étude cinétique du complexe Dicer-TRBP .....	154
4.4.5.	Étude plus approfondie de l'étape limitante de l'activité de Dicer.....	155
4.5.	Étude d'autres régulateurs de Dicer.....	156
4.5.1.	PACT.....	156
4.5.2.	Lin28 .....	157

4.5.3.	Pr55 <sup>gag</sup> .....	158
4.5.4.	Découverte de nouveaux régulateurs de Dicer .....	158
4.6.	Conclusion générale.....	159
5.	Bibliographie.....	i

## Liste des tableaux

Tableau I.	Summary of binding and kinetic constants of Dicer for the pre-let-7 family. ....	109
Tableau II.	Binding and kinetic constants of the effect of TRBP on Dicer cleavage of pre-let-7g. <sup>a</sup> .....	126

## Liste des figures

Figure 1.1. Modèle de l'expression et de la régulation du développement de <i>C.elegans</i> .....	20
Figure 1.2. Voie de biosynthèse des miARN. ....	22
Figure 1.3. Structure de Drosha. ....	24
Figure 1.4. Modèle de l'inhibition de la traduction par les miARN. ....	28
Figure 1.5. Organisation des domaines de PACT et TRBP. ....	30
Figure 1.6. Rôle de Lin28 dans la biogénèse de let-7. ....	32
Figure 1.7. Modèle du rôle antagoniste de hnRNP A1 et KSRP dans la maturation de let-7a-1. .....	33
Figure 1.8. Structure de la protéine Dicer humaine. ....	37
Figure 1.9. Reconnaissance de l'ARN par Dicer. ....	40
Figure 1.10. Interaction de TRBP avec un duplex d'ARN de 23 pb. ....	44
Figure 1.11. Interaction entre TRBP et Dicer. ....	46
Figure 1.12. Résumé des observations de Michaelis-Menten ....	49
Figure 1.13. Concentration des espèces réactionnelles au cours d'une réaction. ....	51
Figure 2.1. Transfection of small suspension culture with Dicer-expressing plasmid. ....	67
Figure 2.2. Dicer expression and purification. ....	69
Figure 2.3. Conformational characterization of purified Dicer. ....	72
Figure 2.4. Binding studies of Dicer to pre-let-7a-1. ....	74
Figure 2.5. Steady-state kinetics for cleavage of pre-let-7a-1 by Dicer.....	75
Figure 2.6. SEC-MALS analysis of purified WT Dicer stored in sucrose/DDM-containing storage buffer for 6 months at -80°C.....	95
Figure 2.7. Preliminary kinetics studies of pre-let-7a-1 cleavage by Dicer under non-steady-state conditions.....	96
Figure 3.1. Thermodynamic and kinetic studies of pre-let-7 cleavage by Dicer.....	111
Figure 3.2. Effect of TRBP on Dicer cleavage of pre-let-7 RNAs. ....	114
Figure 3.3. Steady-state kinetic model of TRBP modification on Dicer cleavage activity. ...	117
Figure 3.4. Thermodynamic and steady-state kinetic experiments with T12.....	122
Figure 3.5. Steady-state kinetic experiments with T3.....	123
Figure 3.6. Thermodynamic and steady-state kinetic experiments with TRBP. ....	125

Figure 3.7. Interpretation of the parametrized steady-state kinetic model of TRBP modification on Dicer cleavage activity. ....	130
Figure 3.8. Primary and secondary structures of the twelve pre-let-7 family members. ....	141
Figure 3.9. Effect of TRBP on the formation of potential isoforms of let-7-5p. ....	142
Figure 3.10. Effect of TRBP on Dicer cleavage of different pre-let-7 substrates as monitored by denaturing gel electrophoresis. ....	143
Figure 3.11. Evaluation of a potential pre-steady-state burst. ....	144
Figure 3.12. Binding of T3 to pre-let-7g. ....	145
Figure 4.1. Représentation des différents types de coopérativité. ....	149

## Liste des sigles

ARN : Acide ribonucléique

ARNr : ARN ribosomal

ATM : Protéine mutée dans le syndrome d'ataxie télangiectasie / *Ataxia telangiectasia mutated protein*

ATP : Adénosine triphosphate

DDM : *n*-dodécyl- $\beta$ -D-maltoside

DGCR8 : Région critique numéro 8 du syndrome de DiGeorge / *DiGeorge critical region 8*

DUF283 : Domaine de fonction inconnue 283 / *Domain of unknown function 283*

ERK : Kinase de réponse aux signaux extracellulaires / *Extracellular signal-regulated kinase*

dsRBD : Domaine de liaison à l'ARN double-brin / *Double-stranded RNA-binding domain*

GTP : Guanosine triphosphate

HEK : Cellules embryonnaires rénales humaines / *Human embryonic kidney cells*

KSRP: Protéine de régulation de l'épissage à homologie K / *K-homology splicing regulatory protein*

LINE: Long élément nucléaire intercalé / *Long interspersed nuclear element*

MAP : Protéine activée par les substances mitogènes / *Mitogen-activated protein*

MIR : Répétitions intercalées de mammifères / *Mammalian interspersed repeat*

PABP : Protéine de liaison à la queue de poly-A / *Poly-A binding protein*

PACT : Protéine activatrice de la kinase induite par les interférons / *Protein activator of the interferon-induced protein kinase*

PAZ : Piwi, Argonaute et Zwiille

PCR : Réaction de polymérisation en chaîne

RISC : Complexe d'inhibition par l'ARN / *RNA-induced silencing complex*

RIG : Gène inductible à l'acide rétinoïque / *Retinoic acid-induced gene*

RT-PCR : Transcription inverse suivie d'une réaction de polymérisation en chaîne

SEC : Chromatographie sur tamis moléculaire / *Size-exclusion chromatography*

SEC-MALS : Chromatographie sur tamis moléculaire couplée à la diffraction lumineuse à angles multiples / *Size-exclusion chromatography coupled to multiple-angle light scattering*

SINE : Court élément nucléaire intercalé / *Short-interspersed nuclear element*

TRBP : Protéine de liaison à l'ARN TAR. / *TAR-RNA binding protein*

TREMD : Dynamique moléculaire par remplacement de répliques de température / *Temperature replica-exchange molecular dynamics*

TU : Unité de transcription / *Transcription unit*

UTR : Région non-transcrite / *Untranslated region*

VIH : Virus de l'immunodéficience humaine

ZBD : Domaine de liaison au zinc / *Zinc-binding domain*

## Liste des abréviations

AGO : Argonaute

Alu : *Arthrobacter luteus*

AKT : Protéine kinase 3

ARNi : interférence par l'ARN

ARNm : ARN messenger

ARNt : ARN de transfert

cryo-EM : microscopie électronique à l'état cryogénisé / *Cryo-electron microscopy*

DISL2 : Ribonuclease du syndrome de Perlman

GW182 : Protéine riche en séquence glycine et tryptophanes

hnRNP : Protéine hétéronucléaire de liaison à l'ARN / *Heteronuclear ribonucleoprotein*

Hsp : Protéine de réponse au choc thermique / *Heat-shock protein*

kb : Kilobase

kDa : Kilo-Dalton

lncARN : long ARN non codant

μM : micro-molaire

Loqs : Protéine Loquacious

miARN : microARN

nM : nano-molaire

nt : Nucléotide

pM :pico-molaire

pri-miARN. : microARN primaire

pol II : ARN polymérase II

RNase : Ribonuclease

TUT4 : Poly-uridyl transferase 4



## Remerciements

Lors de mon arrivée à l'université de Montréal en 2007 pour mon baccalauréat en biochimie, jamais je n'aurais cru que mon séjour allait s'étendre jusqu'en 2018. Onze années plus tard, je suis maintenant un biochimiste s'appêtant à partir à l'étranger pour un stage postdoctoral.

Évidemment, mon parcours a été agrémenté de hauts et de bas. Les bas ont été particulièrement difficile jusqu'au point où j'avais décidé d'abandonner le doctorat. Cependant, j'ai finalement décidé de rester et terminer. Cela est dû en grande partie à l'atmosphère du laboratoire et du département qui me donnait toujours envie d'aller à l'université.

En effet, j'aimerais d'abord remercier l'équipe initiale et les anciens membres des laboratoires Legault/Omichinski avec qui j'ai partagé des tonnes de bons moments. Allant de l'escalade au canoé/camping en passant par les bières, festivals ou simplement par les pauses café, j'ai vécu beaucoup de moments agréables avec vous. Ensuite, il y a l'équipe présente du laboratoire, avec qui les journées de travaux sont agréables et à qui je souhaite beaucoup de succès dans leurs projets. Il y a aussi les membres des autres laboratoires du département, avec qui j'ai pu partager des bons moments et qui sont toujours disponibles pour offrir leur expertise. J'aimerais aussi remercier les employés du département pour leur soutien ainsi que l'équipe du laboratoire d'enseignement, avec qui il m'a été donné de travailler durant une année. Finalement, j'aimerais remercier Jim et Pascale pour m'avoir accueilli dans leur laboratoire et de m'avoir permis de me développer en tant que scientifique.

Je dois aussi remercier les amis de l'extérieure de l'université, plus particulièrement la famille de Shakti rock gym avec qui j'ai versé plusieurs gouttes de sueurs et fait des beaux voyages d'escalade. Vous m'avez permis de me défouler après mes mauvaises journées au labo. Parlant sport, j'aimerais aussi remercier les radicaux libre, club de course du département pour les sorties sur la montagne à tous les moments de l'année.

J'aimerais remercier ma famille qui est toujours présente pour moi. Merci plus particulièrement à mes parents, de chez qui je suis parti en 2012 pour vivre dans un condo et chez qui je suis revenu, avec tous mes meubles et boîtes, en janvier 2018 après la vente du condo. Finalement, merci à Amanda pour m'avoir accompagné durant les 2.5 dernières années.

Son départ en Caroline du nord pour un postdoctorat m'a grandement motivé à terminer mon doctorat et je la remercie grandement de m'avoir attendu malgré le fait que ça a pris plus de temps que prévu pour terminer ma thèse.

En conclusion, même si le doctorat est rempli d'embuches, je vais retenir de très bons souvenirs de mon passage au département de biochimie, et ce, grâce à tous les gens qui m'ont accompagné durant mon parcours correspondant, à ce jour, à un tiers de ma vie.

# 1. Introduction

## 1.1 Biogénèse des microARN

Les microARN (miARN) forment une classe d'ARN non-codants ayant généralement une longueur de 22 nucléotides (nt). Ils sont responsables de moduler la traduction d'ARN messagers (ARNm) pour lesquels ils ont une complémentarité de séquence. De ce fait, ils sont des régulateurs post-transcriptionnels de l'expression des gènes et se retrouvent au cœur de la majorité des processus cellulaires. Depuis leur découverte, les miARN ont été identifiés comme régulateurs du métabolisme, du développement, de la prolifération et de la différenciation cellulaire. Les miARN ont aussi été associés à plusieurs pathologies comme le cancer, la maladie d'Alzheimer et la maladie de Parkinson.

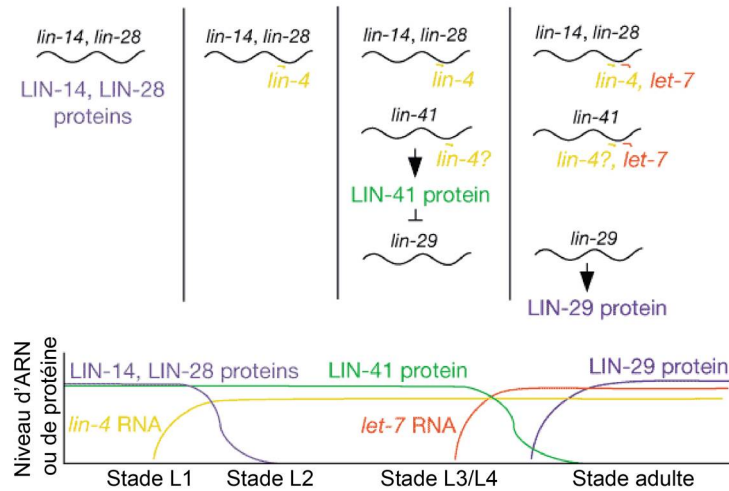
Les miARN sont produits par les cellules par un processus à plusieurs étapes appelé la biogénèse de miARN (Figure 1.2). Ces étapes nécessitent la coordination de plusieurs enzymes et cofacteurs dans différents compartiments cellulaires. Il a été démontré que ces étapes peuvent aussi être hautement régulées par des effecteurs tiers. Les travaux de cette thèse portent sur l'étude des protéines pouvant réguler la biogénèse des microARN en s'intéressant plus précisément aux protéines Dicer et TRBP.

### 1.1.1. Historique

La découverte des miARN date de 1993. Des travaux chez *Caenorhabditis elegans* avaient identifié que le gène *lin-4* induit une baisse des niveaux de la protéine LIN-14 dans le premier stade de développement larvaire (1). À ce moment, les chercheurs ont tenté d'identifier la protéine encodée par *lin-4*. Ils ont finalement découvert que le gène *lin-4* n'encode aucune protéine mais plutôt deux ARN de 22 et 61 nt. Ces ARN présentent une complémentarité de séquence pour une région répétée dans le 3'-UTR de l'ARNm de LIN-14. Ils ont donc postulé que *lin-4* régulait la traduction de LIN-14 par des interactions anti-sens (2).

Après cette découverte, il aura fallu 7 ans pour découvrir un second miARN. La découverte de *let-7* a aussi été faite chez *C. elegans* (3). *let-7* a aussi une complémentarité de

séquence pour la région 3'-UTR de gènes impliqués dans le développement du nématode (Figure 1.1).



**Figure 1.1. Modèle de l'expression et de la régulation du développement de *C.elegans*.**

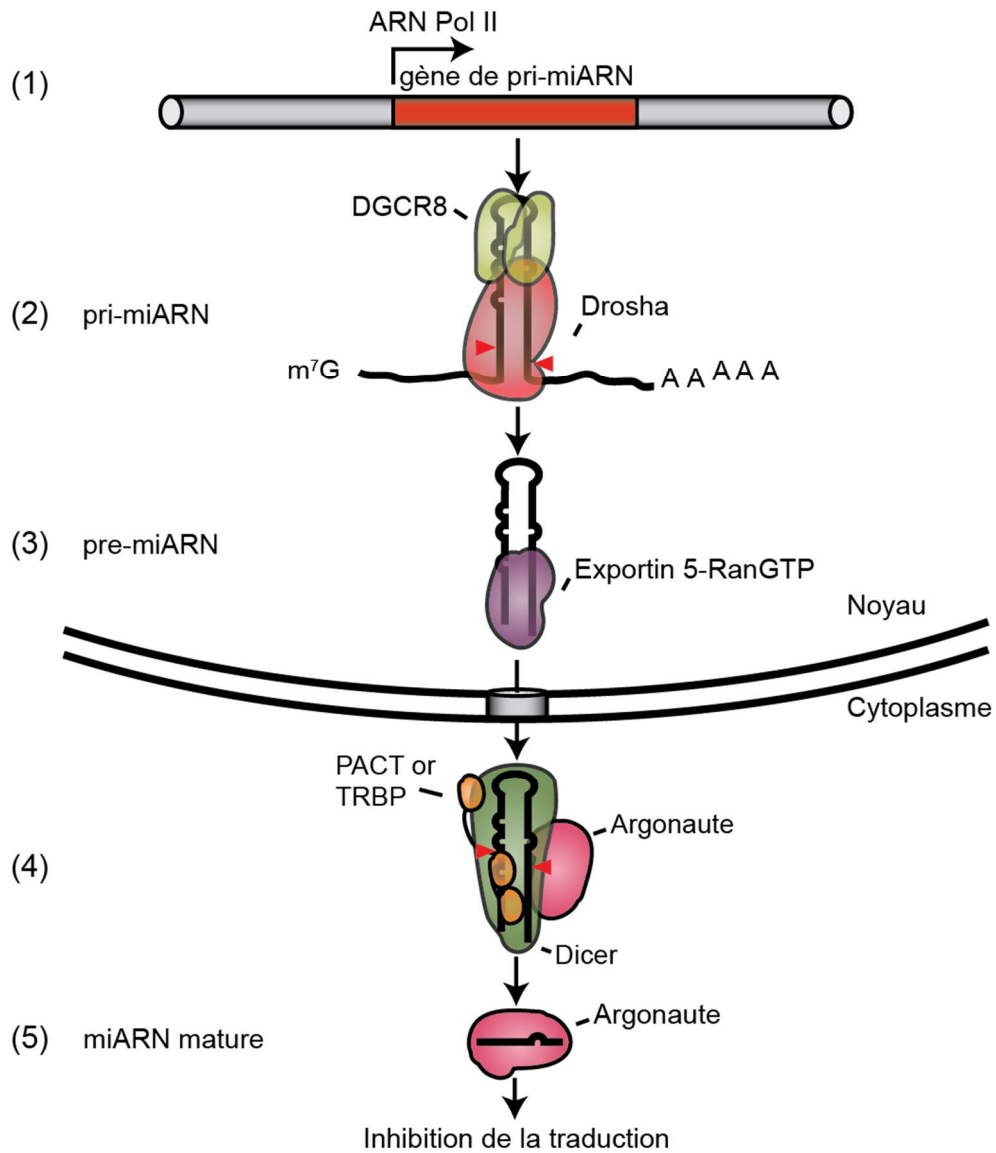
Les miARN lin-4 et let-7 régulent les niveaux des protéines LIN-14, LIN-28, LIN-29 et induisent la différenciation cellulaire. Adapté de (3).

Dans les années suivantes, les miARN ont été découverts chez plusieurs espèces allant de la drosophile aux vertébrés et jusqu'à l'humain (4-6). Aujourd'hui, des milliers de miARNs uniques ont été identifiés (7). On sait maintenant aussi que la grande majorité des ARNm peut être ciblée par un ou plusieurs miARN (8). De plus, des mutations au niveau des miARN ainsi qu'une dérégulation de leur expression sont observées dans plusieurs maladies telles le cancer, la maladie d'Alzheimer et la maladie de Parkinson (9-12). Même certains virus encodent des miARN dans leur génome (13). Donc, la découverte des miARN a vraiment changé la manière de percevoir le contrôle de l'expression des gènes. Par le passé, les modèles prédominants plaçaient ce contrôle au niveau de la transcription. Les miARN ajoutent un niveau de contrôle supplémentaire important, puisque ce mécanisme de régulation offre une réponse beaucoup plus rapide que l'activation et l'inactivation de gènes. Cela permet à la cellule de s'adapter rapidement à un changement d'environnement et de bien contrôler les étapes de différenciation cellulaire lors du développement de l'organisme (14).

## 1.1.2. Transcription

Comme la plupart des gènes codant pour des protéines, les miARN sont majoritairement transcrits par l'ARN polymérase II (pol II) et possèdent une queue de poly-A ainsi qu'une coiffe de 5'-7-méthylguanosine en 5' servant à protéger l'ARN de la dégradation par les exonucléases (15, 16). Cet ARN est appelé transcrit primaire ou pri-miARN (Figure 1.2). Il est possible de retrouver plus d'un miARN dans une unité de transcription (TU). Ces polycistrons (*clusters*) sont donc transcrits ensemble et dépendent du même promoteur. Le premier polycistron identifié est celui encodant miR-23a~27a~24-2 (16). Des études par puce à ADN (*microarray*) suivi d'un positionnement des miARN sur le génome a montré une corrélation entre la distance de deux miARN et leur niveau d'expression (17). Ces études ont également montré que les miARN situés à moins de 50 kilobases (kb) les uns des autres avaient un niveau d'expression similaire alors que la corrélation chutait drastiquement à de plus grandes distances. Cette distance est en accord avec la longueur moyenne d'un gène transcrit par pol II (55 kb). Cela suggère que ces miARN font partie de la même TU.

En plus de retrouver plusieurs miARN sur un seul gène, les miARN sont, la plupart du temps, transcrits avec un gène codant pour une protéine ou un long ARN non-codant (lncARN) (18). Plus précisément, 40 % des miARN sont situés dans des introns de gènes codant pour des protéines, 10 % dans les introns de lncARN et 14 % dans les exons de lncARN. Les polycistrons de miARN sont généralement situés dans le même intron (19). La transcription des miARN peut donc être couplée à celle d'une protéine et d'autres miARN montrant que la transcription d'un seul gène peut avoir un impact à plusieurs niveaux différents sur la cellule. Cependant, certains miARN introniques possèdent leur propre promoteur indépendant et peuvent donc être transcrits indépendamment de la protéine encodée par le gène (20). Il y a aussi une petite population de miARN est transcrite par l'ARN polymérase III (21). Il a été montré que les séquences Alu en amont des miARN contiennent des promoteurs pour l'ARN polymérase III. La même étude a aussi proposé la présence de miARN transcrits en aval de certains ARNt et des répétitions intercalées de mammifères (MIR; *Mammalian-wide interspersed repeats*), mais cette hypothèse n'a toujours pas été confirmée.



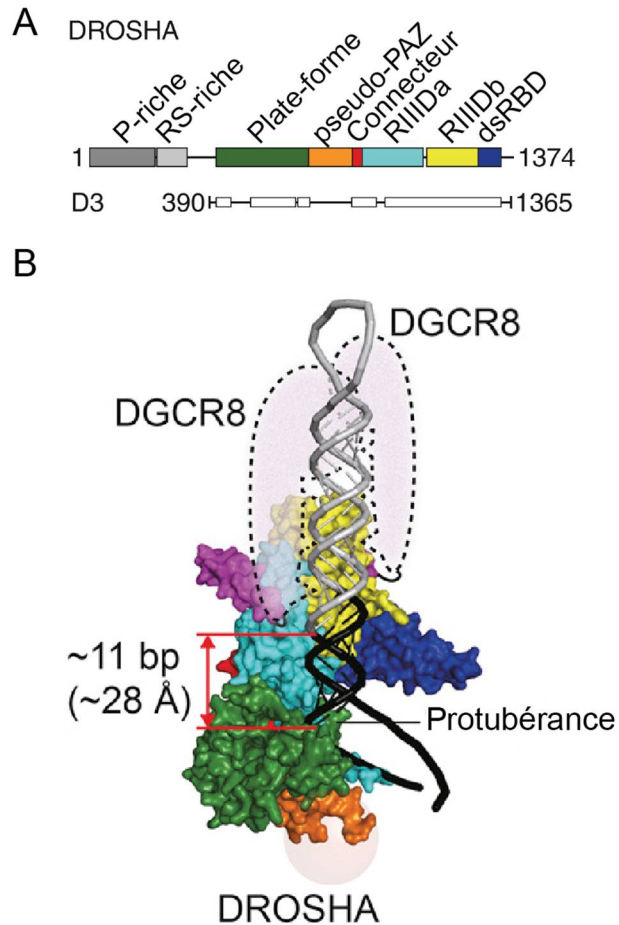
**Figure 1.2. Voie de biosynthèse des miARN.** La voie est divisée en 5 étapes (1) Transcription [Section 1.1.2] (2) Clivage par Drosha [Section 1.1.3] (3) Exportation au cytoplasme [Section 1.1.4] (4) Clivage par Dicer [Section 1.1.5] (5) Chargement sur Argonaute [Section 1.1.6].

### 1.1.3. Clivage par Drosha

Après la transcription, le pri-miARN doit passer par deux étapes de maturation. La première étape a lieu dans le noyau et implique un complexe minimal composé de Drosha et DGCR8. Drosha est l'enzyme responsable du clivage du pri-miARN en miARN précurseur (pre-miARN) d'environ 60 à 80 nt (22). DGCR8 est une protéine pouvant lier la boucle du pri-miARN et permettre le bon positionnement de Drosha. Le complexe est nommé microprocesseur.

Drosha est une nucléase à ARN (ARNase) contenant deux domaines catalytiques de type RNase III (Figure 1.3). Ces domaines utilisent le magnésium et une molécule d'eau afin de cliver le lien phosphodiester de manière irréversible, laissant des extrémités 5'-phosphate et 3'-OH (23). Drosha contient aussi un motif de liaison à l'ARN double brin (dsRBD) lui permettant de lier l'ARN. Depuis peu, la structure de Drosha a montré la présence d'un domaine de liaison au zinc (ZBD) dans le domaine plate-forme pouvant aussi interagir avec l'ARN (24). Finalement, Drosha possède aussi un domaine s'apparentant à un domaine PAZ des protéines Piwi, Argonaute et Zwille et permettant aussi la liaison de l'ARN .

Le positionnement précis de Drosha sur l'ARN est crucial car il délimite l'extrémité 5' du miARN mature. S'il y a clivage au mauvais endroit, la séquence du miARN mature pourrait faire en sorte que des ARNm différents soient reconnus. Plusieurs études ont donc visé à déterminer ce qui donne sa spécificité à Drosha. Le premier élément important est la base de l'hélice du pri-miARN. Drosha peut reconnaître cette extrémité d'hélice ainsi que des résidus spécifiques sur les deux brins basaux (25, 26). Cela permet le positionnement des sites catalytiques de Drosha à environ 11 pb de distance de la base de l'hélice (26, 27). Cependant, ce positionnement n'est pas toujours parfait lorsque les expériences sont réalisées seulement en présence de Drosha et d'un pri-miARN (25).



**Figure 1.3. Structure de Drosha.** **A** Arrangement des domaines de Drosha. **B** La structure cristallographique de la construction D3 de Drosha sur laquelle un pri-miARN a été modélisé et le positionnement attendu de DGCR8 a été dessiné en rose. Adapté de (24).

Des études *in vivo* et *in vitro* ont confirmé que DGCR8 est indispensable au clivage des pri-miARN par Drosha (28). DGCR8 contient deux domaines de liaison à l'ARN de type dsRBD liés de manière assez rigides pouvant lier la tige et la boucle du pri-miARN (29). En plus, DGCR8 possède un domaine Rhed pouvant interagir à la fois avec l'ARN et avec lui-même (30). En effet DGCR8 peut homo-dimériser en présence d'un groupement hème (31). Le groupement hème peut modifier la manière par laquelle DGCR8 interagit avec l'ARN et ainsi modifier le positionnement de Drosha (32, 33). Finalement, Le domaine C-terminal de DGCR8 est composé d'une hélice alpha qui interagit avec les domaines catalytiques de Drosha (24, 25).



Ensemble, Drosha et DGCR8 peuvent se positionner précisément sur le pri-miARN et à 11 bp et 22 pb respectivement à partir des deux extrémités de la tige d'ARN, où DGCR8 lie préférentiellement la région apicale près de la boucle terminale tandis que Drosha lie la région basale (25, 32, 33).

#### **1.1.4. Exportation du pre-miARN vers le cytoplasme**

Les pre-miARN doivent être exportés vers le cytoplasme afin de terminer leur maturation et exercer leur fonction. En 2003, Yin *et al.* ont déterminé que c'est la protéine exportin-5 qui est chargée de rôle (34). Exportin-5 peut lier les pre-miRNA en reconnaissant leur tige et leur extrémité 3'. Le complexe adopte une forme ressemblant à un gant de baseball (35, 36). La structure de l'ARN exporté peut être assez variable puisque celui-ci doit contenir au moins 16 paires de bases et entre 1 et 8 nt non-appariés en 3' (36). Le complexe peut être exporté vers le cytoplasme s'il est accompagné de la protéine de liaison au GTP nucléaire Ran-GTP (37, 38). Après la translocation, l'hydrolyse du GTP permet au complexe d'être relâché et au pre-miARN d'être libéré. L'exportation des pre-miARN est, en soit, un mécanisme de régulation car des expériences de surexpression d'exportin-5 montrent une augmentation des miARN dans les cellules (39). Cependant, la suppression d'exportin-5 dans les cellules n'inhibe pas complètement la maturation des miARN. Cela suggère l'existence de mécanismes d'exportation alternatifs (40).

Exportin-5 peut être régulée par, entre autres, la voie MAP/ERK. Une activation de cette voie peut entraîner la rétention des pre-miARN au noyau et globalement inhiber la maturation des miARN (41). L'exportation des pre-miARN est cependant activée par la voie ATM-AKT en réponse aux dommages à l'ADN (42). Ces résultats montrent que l'exportation des pre-miARN est un processus hautement régulé et dépendant des stimuli extérieurs. Finalement, des mutations d'exportin-5 trouvées dans les cellules cancéreuses ont montré une dérégulation de la maturation de miARN. Par exemple, une mutation induisant une troncature de la partie C-terminale d'exportin-5 dans certaines lignées cellulaires provenant de cancers colorectaux présente une déficience à lier les pre-miARN et Ran-GTP. Cela résulte en une inhibition globale des niveaux de miARN (43).

### 1.1.5. Clivage par Dicer

Une fois dans le cytoplasme, les pre-miARN sont clivés en un duplex de 22 nt par l'endonucléase Dicer. Cette découverte a été faite simultanément dans plusieurs organismes, dont la drosophile et *C. elegans* (44-46). Certains organismes possèdent plus d'un homologue de Dicer. C'est le cas, par exemple, de la drosophile et d'*Arabidopsis thaliana*, où les différents homologues reconnaissent des substrats légèrement différents (47, 48). Le clonage et la purification de Dicer provenant de l'humain et de la drosophile a été très bénéfique à la caractérisation initiale de l'activité de Dicer (49, 50). Il a été montré, entre autres, que Dicer clive l'ARN double brin en présence de magnésium et que l'enzyme humaine, comparativement à Dicer-2 de la drosophile, fonctionne indépendamment de la présence d'ATP (50, 51). Des mutations de Dicer ont été retrouvées dans plusieurs types de tumeurs et celles-ci affectent l'équilibre des miARN dans les cellules (52, 53). Les mutations affectent surtout les résidus du site actif, inhibant l'activité de Dicer mais certaines mutations induisent des troncations du domaine hélicase (54, 55). Ces mutations modifient les miARN formés et, par conséquent, la reconnaissance des substrats par Dicer (56, 57). Dicer sera décrit en plus de détail dans la section 1.4 de cette thèse.

### 1.1.6. Chargement des miARN sur Argonaute

Le duplex de 22 nt généré par Dicer est ensuite chargé sur la protéine Argonaute (AGO) (58). Le chargement est essentiel à la fonction des miARN. Le complexe formé par argonaute et un miARN constitue la base du complexe RISC (*RNA-induced silencing complex*). Le mécanisme de formation du complexe RISC commence par le transfert du duplex de Dicer à AGO pour former le complexe pre-RISC. Un des deux brins du duplex sera retiré tandis que l'autre, représentant le miARN mature, restera lié à AGO.

Afin d'effectuer le transfert du miARN sur AGO, cette dernière doit former un complexe ternaire avec Dicer par l'entremise de TRBP (*TAR-RNA binding protein*) (59-61). Ce complexe est nommé complexe de chargement du RISC (RLC; *RISC loading complex*). TRBP lie l'extrémité la plus stable du duplex et AGO lie le brin 5' de l'autre extrémité. TRBP a donc la fonction de coordonner le chargement du duplex et aider à la détermination du brin qui demeure sur AGO (62-64). Le chargement du duplex sur AGO est aussi un processus dépendant de l'ATP

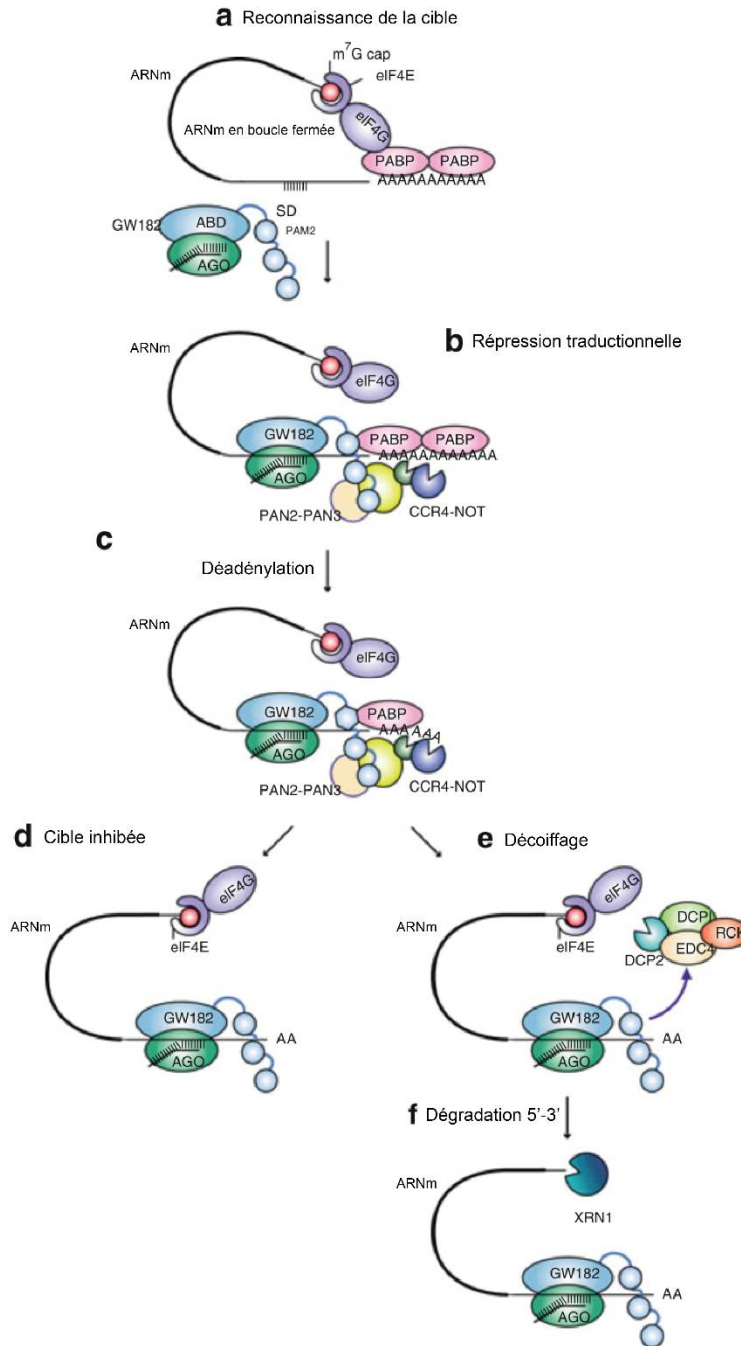
lors duquel les chaperones Hsp70 et Hsp90 promeuvent un changement de conformation de AGO (65-67).

Une fois le transfert complété et le pre-RISC formé, AGO doit se débarrasser du brin complémentaire au miARN. Ce mécanisme est indépendant de l'ATP (65, 66). Il a été proposé que le modèle passif de dissociation des brins du duplex soit lié à la tension qu'AGO accumule en s'ouvrant pour laisser la place au duplex de miARN. La relâche de la tension accumulée éjecterait le second brin (68). Il a aussi été déterminé que l'endonucléase C3PO peut cliver le brin complémentaire et faciliter sa dissociation (69).

### **1.1.7. Inhibition de la traduction par le complexe RISC**

Les miARN chargés sur AGO peuvent reconnaître des ARNm par complémentarité de séquence. La complémentarité doit être parfaite au niveau de la séquence «*seed*», composée des nucléotides 2 à 7 du miARN, afin d'observer une inhibition de la traduction du miARN (70). Chez l'humain, il existe 4 isoformes d'AGO. Parmi celles-ci, seul AGO2 a un domaine catalytique fonctionnel et est capable de cliver les ARN messagers (71). Ce clivage peut cependant se produire seulement si la complémentarité de séquence entre le miARN et sa cible est parfaite (72). Puisque c'est très rarement le cas, les ARNm cibles ne sont généralement pas clivés (6). C'est pourquoi plusieurs autres mécanismes peuvent engendrer l'inhibition de la traduction par les miARN.

Le complexe RISC mature est seulement fonctionnel si les protéines de la famille GW182 y sont associées (73, 74). En effet, les protéines GW182 recrutent les facteurs de déstabilisation et dégradation des ARNm (75). Les complexes recrutés sont surtout liés à la déadénylation de ARNm mais GW182 peut aussi promouvoir le décoiffage en 5' de l'ARNm (Figure 1.4) (73, 74).



**Figure 1.4. Modèle de l'inhibition de la traduction par les miARN.** a. Le complexe RISC composé d'un miARN, d'AGO et GW182 reconnaît un ARNm en boucle fermée sur sa région 3'-UTR. b et c. GW182 interagit avec les PABP pour réprimer la traduction. Elle recrute le complexe de déadénylation. d. Les ARNm peuvent être conservés dans un stade

de répression. **e et f.** La machinerie de décoiffage est recrutée pour ensuite envoyer l'ARNm vers la dégradation. Adapté de (76).

### **1.1.8. Autres voies de biogénèse**

Certains miARN n'ont pas besoin d'être clivés par Drosha. C'est le cas d'une classe spécifique de miARN introniques découverts chez la drosophile. Les groupes de Bartel et Lai ont montré que certains petits introns contenaient seulement l'équivalent de la séquence d'un pre-miARN flanqué directement de sites d'épissage (77, 78). Leurs études ont montré que ces pre-miARN, appelés mirtrons, sont produits par la machinerie d'épissage indépendamment de Drosha et qu'ils empruntent ensuite la voie canonique de maturation. Peu après cette découverte chez les invertébrés, une autre étude a mis en lumière la présence de mirtrons chez les mammifères (79). La position des mirtrons et les miARN provenant de mirtrons sont différents entre les espèces suggérant l'apparition indépendante de ce système durant l'évolution.

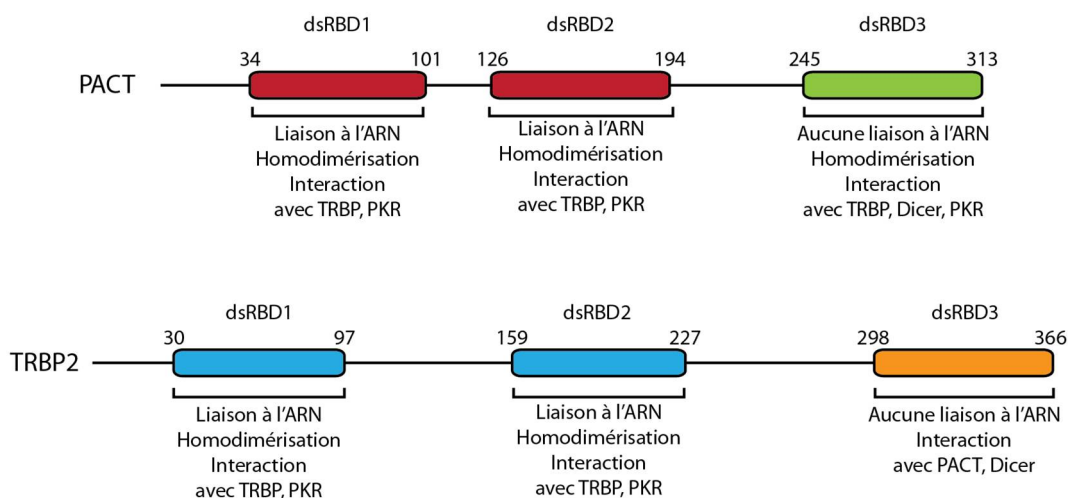
## **1.2. Régulation de la biogénèse des miARN**

Les miARN jouent un rôle important sur la régulation post-transcriptionnelle des gènes et peuvent rapidement changer la fonction d'une cellule en inhibant la traduction de certaines protéines. Il est donc extrêmement important de bien réguler la biogénèse des miARN. Plusieurs groupes de recherche ont concentré leurs efforts sur la découverte et la caractérisation de mécanismes de régulation de la maturation des miARN. Tous ces travaux ont confirmé que la biogénèse des miARN est hautement régulée et qu'il existe une très grande variété de protéines régulatrices. À partir de ces découvertes, il est maintenant supposé que le petit groupe de régulateurs découverts jusqu'à maintenant ne constitue que la partie visible de l'iceberg et que plusieurs autres régulateurs restent à être découverts. Cette section vise à résumer quelques-uns des régulateurs les mieux caractérisés ainsi que les efforts récents apportés à la découverte de nouveaux régulateurs de la biogénèse des miARN.

### **1.2.1. TRBP et PACT**

TRBP (protéine de liaison à l'ARN TAR) et PACT (Protéine activatrice de la kinase induite par les interférons) sont deux protéines hautement similaires au niveau de leurs domaines

et de leur fonction (Figure 1.5). Elles ont été caractérisées comme cofacteurs essentiels à la biogénèse des miARN (60, 61). Elles ont un rôle d'activateur de Dicer où ils aident la reconnaissance et le positionnement du pre-miARN sur Dicer. Leur rôle le plus essentiel semble cependant être au niveau du transfert du miARN mature de Dicer à Argonaute (60, 61). Il a été montré par des expériences de suppression par interférence à l'ARN (ARNi) que, sans TRBP ou PACT, le pre-miRNA peut être clivé par Dicer sans toutefois pouvoir être lié par Argonaute et pouvoir faire partie du complexe RISC. Il a aussi été montré que TRBP et PACT font partie intégrante du complexe RISC (80). La manière par laquelle ces protéines permettent le transfert du miARN de Dicer à Argonaute est toujours méconnue. TRBP sera décrit de manière plus détaillée à la section 1.5 de cette thèse.



**Figure 1.5. Organisation des domaines de PACT et TRBP.** Identification des domaines de PACT et TRBP ainsi que leurs principales fonctions.

### 1.2.2. Lin28

Lin28 est une protéine de liaison à l'ARN qui est parmi les premiers régulateurs de la biogénèse des miARN ayant été découverts. La particularité de cette protéine est sa spécificité à la famille let-7 (81). En effet, Lin28 inhibe la maturation de let-7 (82). L'équilibre Lin28/let-7 a été prouvé très important dans la coordination du développement larvaire chez *C.Elegans* (83, 84). Après cette importante découverte, Lin28 a été découverte chez plusieurs autres organismes eucaryotes. En plus d'être importante pour la maturation de let-7, Lin28 a été

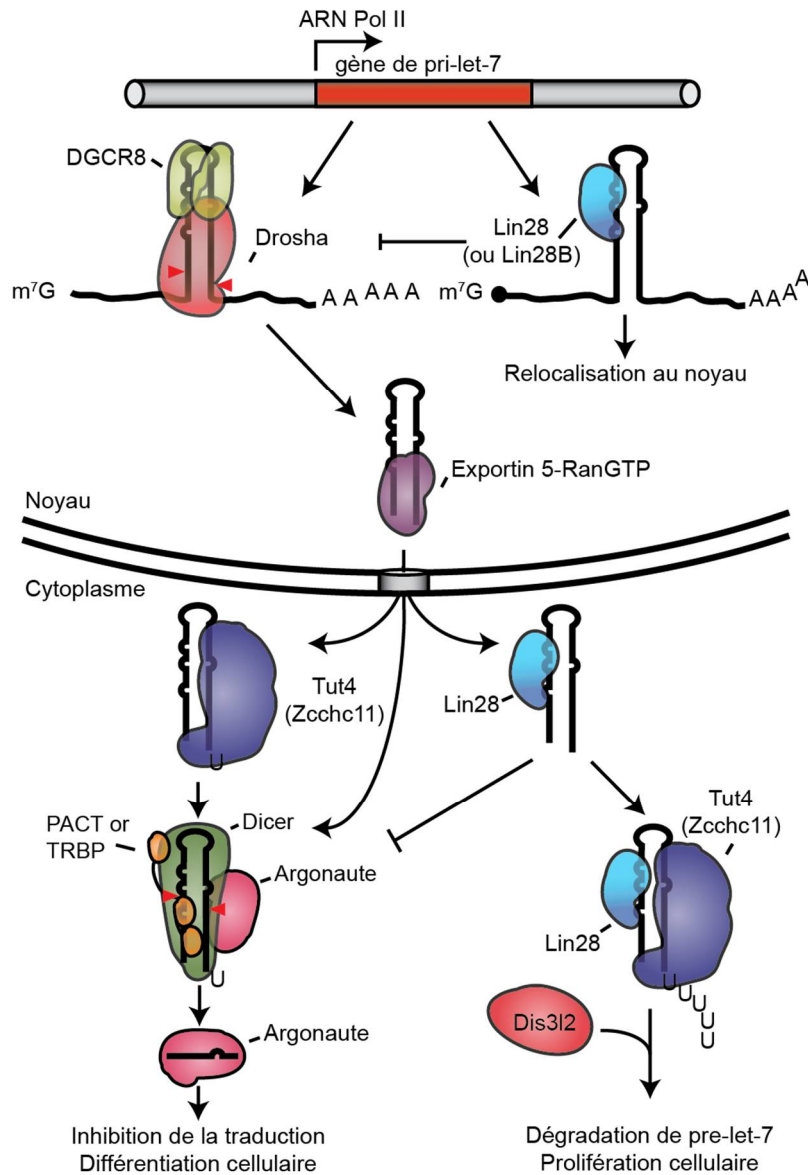
identifiée comme une des 4 protéines minimales permettant la reprogrammation d'une cellule différenciée (85, 86). Aussi, Lin28 peut réguler la traduction d'ARN messagers (87-89). Lin28 est aussi surexprimée dans la plupart des cancers, permettant aux cellules de se reproduire plus rapidement (90). La suppression de Lin28 dans des lignées cancéreuses ralentit grandement leur prolifération (82).

Lin28 possède une combinaison unique de domaines de liaison à l'ARN. Son domaine C-terminal contient une région riche en lysines et arginines ainsi que deux domaines de liaison au zinc de type CCHC hautement similaires à la protéine virale NCp7 encodée par le VIH (91). Ce domaine reconnaît et lie les guanines exposées avec une très grande affinité. Le second domaine de Lin28 est un domaine de choc hypothermique (*cold shock*). Ces domaines ont une fonction de chaperonne à ARN leur permettant de déstabiliser les structures secondaires et tertiaires de l'ARN (92). La combinaison de ces domaines permet la reconnaissance d'une région riche en guanines exposées et conservées chez 11 des 12 pri- et pre-let-7 chez l'humain et la déstabilisation des structures de cette boucle (93, 94). Seul pre-let-7a-3 ne contient pas de site de liaison à Lin28 (95).

Lin28 peut agir dans le noyau et inhiber la maturation de pri-let-7 par Drosha (84). La liaison de Lin28 à pri-let-7 induit la relocalisation du complexe dans le nucléole, le rendant inaccessible au clivage par Drosha (96). Le mécanisme menant à cette relocalisation serait favorisé par la méthylation de Lin28 sur sa lysine 135 par la méthyltransférase SET7/9 (97). La liaison de Lin28B, un homologue de Lin28 possédant une séquence de localisation nucléolaire est aussi responsable de la relocalisation (96). Les rôles de Lin28 et pri-let-7 au nucléole n'ont toujours pas été élucidés.

Lin28 peut aussi agir dans le cytoplasme en inhibant le clivage de pre-let-7 par Dicer. La déstabilisation de la boucle de pre-let-7 induit une déstabilisation du site de clivage de Dicer (94). Cette dernière a besoin que les bases situées au niveau du site de clivage soient parfaitement appariés (98). Cette déstabilisation permettrait une inhibition partielle de l'activité de Dicer. L'inhibition totale de Dicer vient du recrutement de l'enzyme TUT4, une poly-uridylyl transférase (99-101). Le modèle proposé est que la déstabilisation de la boucle de pre-let-7 par Lin28 permet à TUT4 de lier une région simple brin riche en guanines et catalyser la poly-uridylylation de pre-let-7 sur son extrémité 3'. En plus de rendre pre-let-7 un mauvais substrat de

Dicer, la queue de poly-U peut être reconnue par l'exonucléase DIS3L2 qui dégrade pre-let-7 (102, 103).

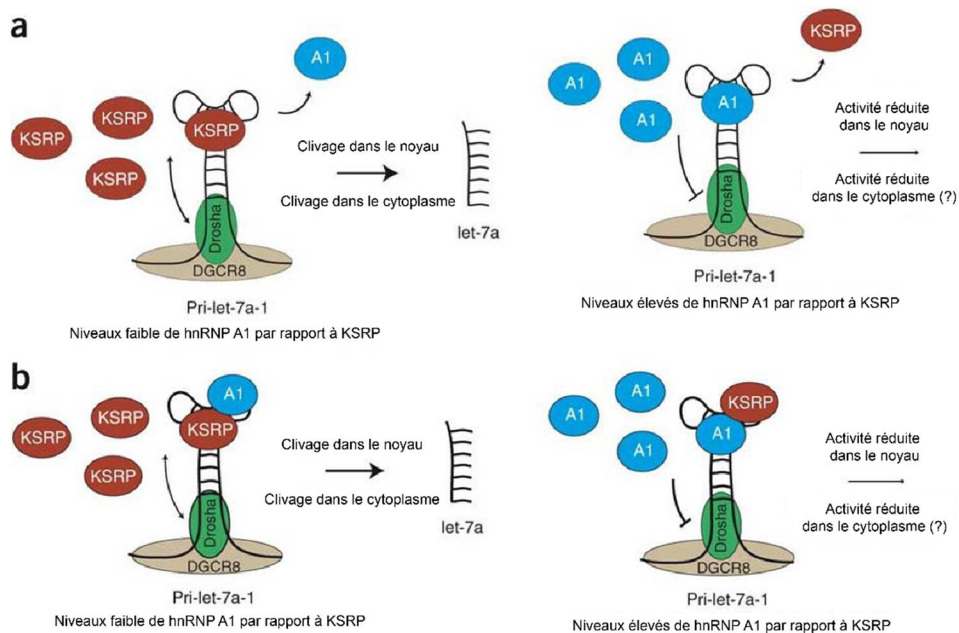


**Figure 1.6. Rôle de Lin28 dans la biogénèse de let-7.** Lin28 peut interagir avec le pri-let-7 au noyau, inhiber la maturation et relocaliser le pri-miARN vers le nucléole. Dans le cytoplasme, l'interaction de Lin28 avec pre-let-7 induit la poly-uridylation de l'ARN par TUT4 qui est ensuite dégradé par DIS3L2. TUT4 peut aussi monouridyler pre-let-7 en absence de Lin28.



### 1.2.3. hnRNP A1 et KSRP

hnRNP A1 et KSRP sont deux protéines nucléaires pouvant lier les pri-miARN. Leur fonction est antagoniste, où KSRP est un activateur du clivage de pri-let-7a-1 par Drosha tandis que hnRNP A1 est un inhibiteur (Figure 1.7) (104). Les deux protéines peuvent lier le même site contenant un motif GGG sur pre-let-7a-1 et compétitionner pour la liaison (105). Dans d'autres cas, hnRNP A1 peut aussi activer la maturation de pri-miARN, notamment pri-miR-18a, en modifiant la structure de la boucle terminale afin de la rendre plus accessible à la liaison Drosha-DGCR8 (106). KSRP peut aussi promouvoir le clivage des pre-miARN par Dicer en liant les motifs riches en guanine sur leur boucle terminale (105).



**Figure 1.7. Modèle du rôle antagoniste de hnRNP A1 et KSRP dans la maturation de let-7a-1. a.** Modèle par lequel KSRP et hnRNP A1 compétitionnent pour la liaison à pri-let-7a-1. **b.** Modèle d'occupation simultanée. Adapté de (104).

### 1.2.4. Autres régulateurs et efforts récents

La majorité des régulateurs de la maturation des miARN est encore inconnue. Des efforts récents visant à trouver ces régulateurs ont généré une très longue liste de protéines pouvant lier

les pri- et pre-miRNA. Ces régulateurs potentiels provenant d'études de criblage à large échelle demandent maintenant à être validés par des études de caractérisation *in vivo* et *in vitro*.

Une étude réalisée par notre laboratoire visait à découvrir les protéines pouvant interagir avec la boucle terminale commune à pri- et pre-let-7 (108). Les protéines ont été isolées par une méthode de purification par affinité (*pull-down*) par l'ARN suivi par la caractérisation par spectrométrie de masse. La méthode consiste à attacher le pre-miARN auquel une séquence épitope a été ajoutée (*tag*) à une résine par l'interaction de l'épitope à une protéine préalablement liée à la résine. Un lysat cellulaire est ensuite incubé avec la résine. Après plusieurs lavages, les protéines attachées à l'ARN sont éluées par le clivage de l'épitope. La fraction éluée est ensuite assujettie à une analyse par spectrométrie de masse. En répétant cette expérience avec différents membres de la famille let-7, il a été possible de générer une liste de protéines communes aux différents substrats. Parmi les protéines identifiées, certaines avaient déjà été identifiées comme régulateurs de la biogénèse des miARN et la plupart contiennent des domaines connus de liaison à l'ARN.

Une autre étude beaucoup plus étendue utilisant une méthode similaire a été réalisée par Treiber *et al.* (109). Cette étude a identifié l'interactome de 72 miARN dans 11 lignées cellulaires, ce qui a mené à la publication d'une longue liste de protéines interagissant avec la tige ou la boucle terminale des pre-miARN. Parmi les protéines identifiées, certaines sont spécifiques à un sous-groupe de miARN tandis que certaines agissent sur une très grande variété de miARN. Certaines protéines interagissant avec let-7 identifiées par notre laboratoire ont aussi été retrouvées par cette étude. Cela offre un niveau supplémentaire de validité pour les protéines ayant été identifiées par les deux études.

Les protéines identifiées par ces études de criblages constituent un bon point de départ pour la découverte de nouveaux régulateurs potentiels de la biogénèse des miARN. Il reste maintenant à tester ces candidats afin de valider leur rôle.

### **1.3. La famille let-7 de miARN**

Le miARN *lethal-7* (let-7) est très bien conservé entre les espèces. Après sa découverte chez *C. elegans*, il a été découvert chez les mammifères ainsi que dans les organismes bilatéraux

où il possède une séquence très conservée (110, 111). Bien qu'il n'y ait qu'un seul gène exprimant *let-7* chez le nématode, d'autres organismes possèdent plusieurs variants de *let-7* possédant tous la même séquence «*seed*». Ces différents *let-7* sont nommés à l'aide d'une lettre différente, allant de a à i chez l'humain. De plus, l'humain possède trois précurseurs différents donnant *let-7a* et deux donnant *let-7f*, ceux-ci différenciés par un numéro (ex. *let-7a-1*) (110). Chez l'humain, la famille *let-7* est donc composée d'un total de 12 membres dont les premiers miRNA ont des longueurs et des structures secondaires variables (Figure 3.8).

La famille *let-7* est pertinente pour l'étude de la régulation de la biogénèse des miARN car il a été montré que l'accumulation de *pri-let-7* n'est pas corrélée avec l'accumulation de *let-7* mature, suggérant la présence de mécanismes de régulation post-transcriptionnelle (82, 112). En effet, quelques mécanismes de régulation de la biogénèse de *let-7* ont été identifiés. Le premier est la régulation par *Lin28*, qui a été décrit dans la section 1.2.2. Ensuite, la poly-uridylyl transférase *TUT4* a un rôle de régulation en absence de *Lin28* où elle peut mono-uridyler certains *pre-let-7* (tous sauf *pre-let-7a-2*, *-c* et *-e*) afin d'en faire des substrats plus optimaux pour le clivage par *Dicer* [(113); voir section 1.4.2]. Finalement, les protéines *hnRNPa1* et *KSRP* peuvent réguler le clivage par *Drosha* de la famille *pre-let-7* (section 1.2.3).

L'intérêt physiologique de l'étude de la biogénèse de *let-7* provient aussi du fait, qu'en plus d'être essentiel à la régulation temporelle du développement des organismes (section 1.1.1), il a un rôle général dans la régulation de la différenciation cellulaire (112, 114, 115). *let-7* est aussi caractérisé comme suppresseur de tumeur car il peut inhiber la traduction de plusieurs oncogènes tel *Ras*, *Myc* et *HMGA2* et ralentir la prolifération cellulaire (116-118). La famille *let-7* est donc généralement sous-réglée dans les cancers et cette inhibition est souvent liée à un mauvais pronostic pour la survie des patients (9, 119, 120). Cela rend l'étude de la biogénèse de *let-7* intéressante en tant qu'avenue thérapeutique contre le cancer (121).

## 1.4. Dicer

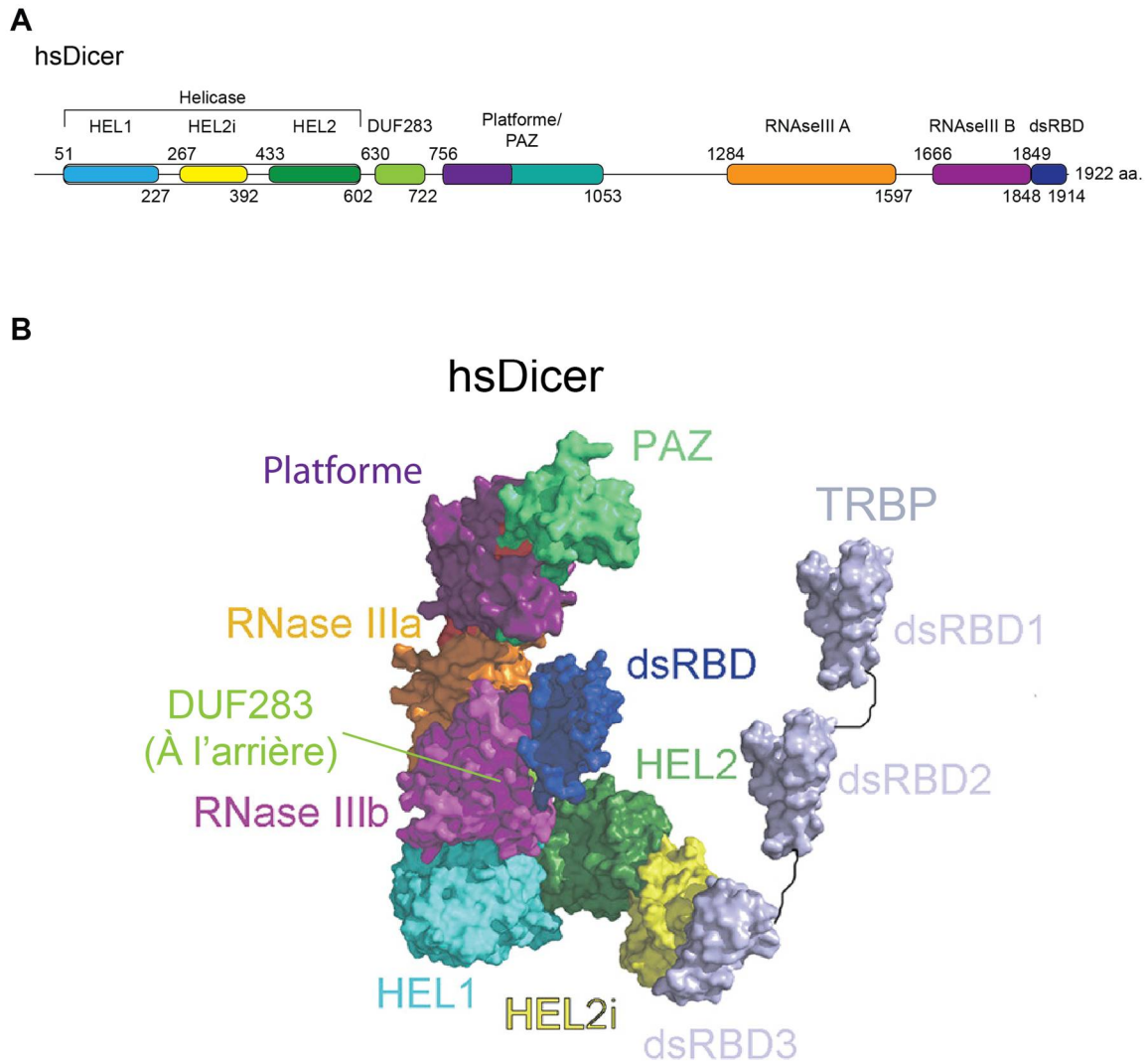
*Dicer* est responsable de la dernière étape de clivage dans la maturation des miARN. C'est une protéine à plusieurs domaines ayant tous une importance capitale pour sa fonction. Afin que le miARN mature soit fonctionnel, *Dicer* doit reconnaître le *pre-miARN* et le cliver très précisément sur les deux côtés de sa tige à 22 nt de l'extrémité 5'. Pour ce faire le substrat

doit être parfaitement positionné. De plus, l'accès au site actif de Dicer est très limité, ce qui lui confère sa grande spécificité pour les pre-miARN (122, 123).

### 1.4.1. Domaines

Dicer est une grosse protéine d'environ 220 kDa. Elle est composée de 6 domaines principaux lui permettant une reconnaissance très spécifique des miARN et le positionnement très précis du substrat. De l'extrémité N-terminale à C-terminale, Dicer possède un domaine hélicase, DUF283, PAZ, RNaseIII A, RNaseIII B et dsRBD (Figure 1.8 A).

Le premier domaine de Dicer est un domaine hélicase de type DExH-box. Ce type de domaine se replie sous forme de «U» (124). On peut le séparer en 3 sous-domaines HEL1, HEL2i et HEL2. La première section, HEL1, est un domaine de liaison à l'ATP. Ce type de domaine peut généralement lier et hydrolyser l'ATP afin d'induire un mouvement d'ouverture ou de fermeture de la pince et permettre à ce domaine de se déplacer sur l'ARN (125). Ce type de mouvement dépendant de l'ATP n'a cependant jamais été montré chez la protéine Dicer humaine mais l'hydrolyse d'ATP est nécessaire chez la protéine Dicer-2 de drosophile pour la maturation des siRNA endogènes (49-51). Le deuxième sous-domaine, HEL2i, comprend la région d'interaction à TRBP. Cette région présente une face hydrophobe au solvant permettant de recruter d'autres protéines comme TRBP et potentiellement PACT (126). Le sous-domaine HEL2 forme la partie C-terminale de l'hélicase. Son rôle est principalement de lier les acides nucléiques et il est généralement composé de plusieurs acides aminés basiques. Il n'y a pas de structure à haute résolution connue pour le domaine hélicase de Dicer, mais des analyses de séquences et des outils de prédiction de structures indiquent que sa structure serait hautement similaire à celle de RIG-I (124, 127). Cette dernière lie les ARN viraux et est impliquée dans l'immunité innée. Sa spécificité pour ces ARN lui vient d'un autre petit domaine à son extrémité N-terminale venant fermer la pince et former un anneau (125).



**Figure 1.8. Structure de la protéine Dicer humaine.** **A.** Structure primaire et organisation des domaines de Dicer. **B.** Structure tertiaire de Dicer en complexe avec TRBP (gris). Adapté de (123).

Le domaine hélicase de Dicer n'est pas essentiel à son activité catalytique. Des études de mutagenèse ont montré que la délétion complète du domaine hélicase augmente la vitesse de réaction de Dicer (57). Cette délétion cause cependant une diminution de la spécificité de Dicer de sorte que les longs ARN double brin peuvent aussi être clivés efficacement. Dicer a donc besoin de tous ces domaines pour être spécifique même si cela entraîne une perte d'activité.

Le second domaine de Dicer se nomme DUF283 (*Domain of unknown function*) car, historiquement, sa fonction n'était pas connue. Nous savons maintenant grâce à des études

structurales et biochimiques que ce domaine adopte un repliement similaire à un dsRBD, à la différence que l'hélice alpha ne se couche pas sur le feuillet beta avec le même angle (128). Ce domaine serait, par des analyses de séquence, incapable de lier les acides nucléiques. Cependant, une étude a montré que le domaine DUF283 de Dicer peut lier les ARN et, en plus, jouer un rôle de chaperone permettant de favoriser la formation d'ARN double brin (129). Ce domaine étant positionné juste entre les domaines hélicase et le domaine catalytique et pourrait venir stabiliser le pre-miRNA près du site de clivage (123).

Les domaines hélicase et DUF283 de Dicer sont très susceptibles de fonctionner de concert afin de reconnaître la boucle terminale des pre-miRNA et favoriser la formation de paires de bases en aval du site de clivage. Comme mentionné précédemment, des études de mutagenèse ont montré que le retrait du domaine hélicase lui permet de mieux cliver les ARN (57). Cela est vrai à condition que le domaine DUF283 soit présent. Le retrait du DUF283 en plus de la délétion du domaine hélicase entraîne une perte quasi-complète de l'activité de Dicer (130). Cela indique que le domaine DUF283 est très important dans le positionnement et la stabilisation du substrat près du site actif de Dicer.

Le domaine suivant est le domaine PAZ. Ce domaine est important pour la reconnaissance de la base de la tige des pre-miRNA composée des extrémités 5' et 3' (131, 132). Le domaine PAZ peut reconnaître à la fois une extrémité 5'-phosphate ainsi que 2 nt non-appariés en 3' (*2-nt 3'-overhang*). Le domaine PAZ est le point de contact avec le substrat permettant le clivage au nucléotide 22 (133). En effet, le 5'-phosphate est le point déterminant la position du site de clivage. Un ARN n'ayant pas de phosphate sera clivé avec moins de précision et produira des ARN de longueur variable (21-24 nt). L'effet est moins marqué sur le côté 3', mais la précision du clivage est affectée lorsque les ARN ne possèdent pas 2 nt non-appariés.

Après avoir bien positionné le substrat sur les domaines PAZ et hélicase, l'ARN peut être hydrolysé par deux domaines de type RNaseIII. Ce type de domaine est très répandu chez les organismes partant des bactéries jusqu'à l'humain (98). Les protéines de type RNaseIII peuvent être séparées en plusieurs classes mais elles ont tous quelques points en commun : [1] Elles forment des dimères par leur domaine RNaseIII ou possèdent deux domaines RNaseIII (46, 134) [2] Elles clivent de l'ARN double-brin (135) [3] Elles ont généralement un domaine

dsRBD en plus de leur domaine RNaseIII (136). Les domaines RNaseIII ont besoin de deux ions magnésium afin de catalyser la réaction. Chez Dicer, les deux domaines RNaseIII s'organisent en dimère interne et forment une grande face basique permettant à l'ARN de s'y lier (131, 137, 138). La catalyse se produit simultanément sur les deux côtés de la tige d'ARN, libérant la boucle et la tige.

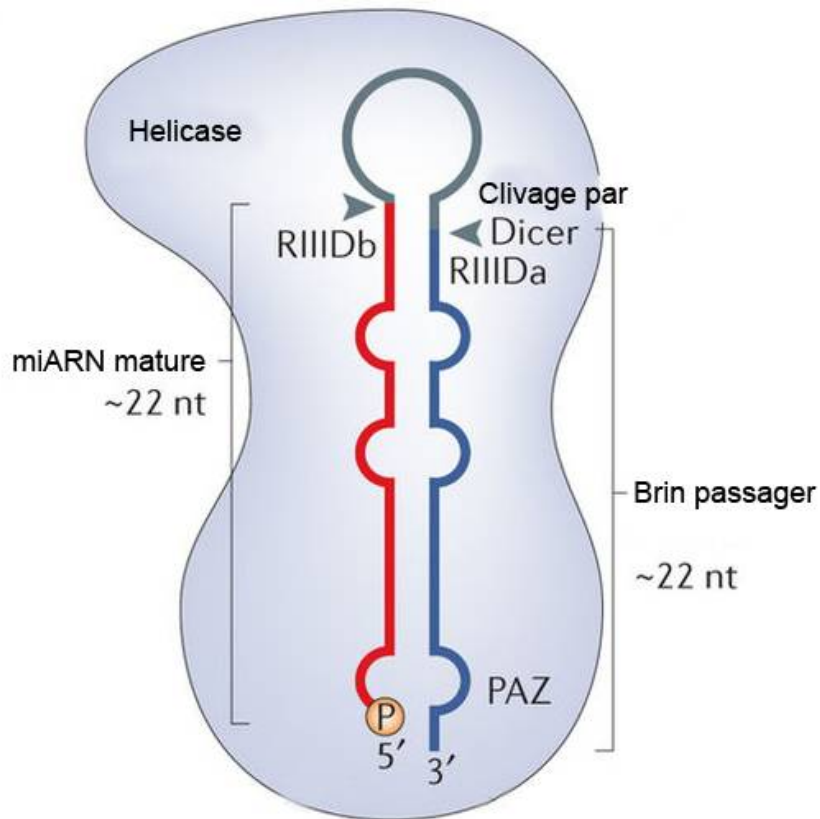
Le dernier domaine est un domaine dsRBD. Généralement, ce type de domaine lie aussi l'ARN double brin et permet de stabiliser l'ARN (139). Bien que la fonction ou le site de liaison exact du dsRBD de Dicer ne soit pas connue, quelques pistes nous indiquent que celui-ci pourrait aider le positionnement et possiblement stabiliser la structure double brin essentielle à la catalyse (123).

### **1.4.2. Reconnaissance de l'ARN par Dicer**

Plusieurs études se sont penchées sur les ARN reconnus par Dicer dans le but de trouver quels sont les éléments structuraux préférentiellement reconnus par Dicer. Ces études ont surtout été réalisées *in vitro* en présence de plusieurs différents ARN naturels et synthétiques. Même si les pre-miARN sont de bien meilleurs substrats pour Dicer comparé à un duplex d'ARN, les premières études *in vitro* ont surtout utilisé des siARN double brin ayant une parfaite complémentarité (140). Ces études ont montré l'importance d'avoir une extrémité 3' avec 2 nt simple brin. Sans changer l'affinité de Dicer pour l'ARN, cette extrémité 3' est importante pour la précision du clivage et l'homogénéité des produits générés (141). En plus de l'extrémité 3', une autre étude a montré que Dicer a aussi besoin d'un groupement phosphate en 5' pour obtenir un clivage précis et homogène (133). Cette dernière étude a montré que ces prérequis étaient les mêmes pour un pre-miARN ainsi que pour les siARN (Figure 1.9).

Ensuite, des études se sont intéressées aux mésappariements dans les tiges des pre-miARN ainsi que leur taille. Premièrement, l'absence de mésappariement semble diminuer l'activité de la Dicer humaine mais ce même effet est faible chez Dicer-1 de la drosophile (142, 143). La présence de ces mésappariements pourrait changer significativement la structure 3D des pre-miARN pour les rendre plus flexibles ou simplement changer leur structure. Une autre étude a montré que l'introduction de protubérances dans la tige du pre-miARN peut changer la position du site de clivage par Dicer (144). Cela peut changer le brin chargé sur Argonaute et

cibler des ARNm complètement différents. La longueur de la tige a aussi son importance sur la position du site de clivage et une trop longue tige résulte en une forte diminution de la vitesse de clivage, et ce, autant chez l'humain que chez la drosophile (142, 144).



**Figure 1.9. Reconnaissance de l'ARN par Dicer.** Dicer lie les extrémités 5' et 3' avec le domaine PAZ et la boucle terminale avec le domaine hélicase. ~22 nt sont mesurés à partir de l'extrémité 5'.

Finalement, la taille de la boucle terminale de l'ARN peut avoir des effets sur l'activité de Dicer. Une boucle terminale de moins de 10 nt réduit fortement l'activité de Dicer (143). Des résultats similaires ont été observés chez la drosophile avec Dicer-1 (142). Il a aussi été montré que la boucle n'a pas besoin d'être fermée mais les deux brins ont seulement besoin d'une séquence simple brin assez longue à leur extrémité pour interagir avec le domaine hélicase de Dicer (142).



Plusieurs études seront nécessaires afin d'évaluer tous les paramètres de reconnaissance des pre-miARN par Dicer. Ce qui est certain est que Dicer est capable de lier une très grande variété de substrat avec des affinités très similaires et de les cliver (143). Les différences sont surtout présentes au niveau de la taille des produits de clivage et de la vitesse de réaction.

### 1.4.3. Autres rôles de Dicer et localisation cellulaire

Un nouveau champ de recherche s'est récemment ouvert quand certains groupes ont commencé à caractériser de nouveaux rôles pour Dicer indépendamment des miRNA. En effet, une grande variété de nouveaux substrats a été attribuée à Dicer. On y retrouve des ARNt, de long et courts éléments nucléaires intercalés (LINE et SINE *Long* et *Short Interspersed Nuclear Elements*) ARNr et des ARN viraux (145). Tous ces substrats peuvent être une source de petits ARN non-codants produits par Dicer. Ces sous-domaines de recherche sont cependant nouveaux et les rôles et les mécanismes menant à ces nouveaux produits de Dicer doivent encore être étudiés.

Parmi ces nouveaux substrats, plusieurs se retrouvent au noyau. Dicer est d'abord connue pour localiser au cytoplasme, mais quelques études ont montré qu'une petite fraction de celle-ci pouvait en effet être dans les noyaux. Cette fraction correspondant à environ 5% de la quantité totale de Dicer dans les cellules murines a fait l'objet de débat dans la littérature puisqu'elle est extrêmement difficile à détecter (146-148). En effet, Dicer possède une séquence de localisation nucléaire sur son domaine dsRBD (149). Il a été proposé que cette séquence soit très peu disponible et qu'une conformation spécifique du domaine hélicase permettrait de la présenter au solvant et rendre la translocation au noyau possible. Cela est supporté par le fait que des mutants n'ayant pas de domaine hélicase entier seraient localisés au noyau (149). Aucun mécanisme n'est encore connu mais on pourrait supposer qu'une modification post-traductionnelle ou l'interaction avec une protéine soit à l'origine de ce mécanisme (148). Néanmoins, il semble effectivement y avoir présence de Dicer au noyau. Il a aussi été démontré que Dicer peut être phosphorylée et être recrutée à des sites de dommages à l'ADN. Une des hypothèses est que Dicer puisse être en mesure de produire localement des petits ARN permettant le recrutement de facteurs de réparation ou le remodelage de la chromatine (147, 148, 150).

Dicer peut aussi être localisée aux membranes dans certains types cellulaires (49). Une étude a montré que Dicer se localise au réticulum endoplasmique dans les pré-dendrites de neurones probablement pour favoriser la production locale de miARN et que TRBP peut y co-localiser de manière calcium-dépendante (151, 152).

## **1.5. TRBP**

TRBP a d'abord été découverte pour son rôle dans l'interaction avec l'ARN TAR du VIH de type 1, mais elle a très rapidement été associée à la maturation des miARN (60, 61, 153-155). Les premières études ont montré qu'elle pouvait influencer les niveaux de miARN matures mais aussi qu'elle a un rôle essentiel dans le transfert du miARN mature de Dicer vers Argonaute. TRBP fait donc partie intégrante du complexe RLC en tant que facteur essentiel à la maturation des miARN.

### **1.5.1. Domaines**

TRBP est une protéine de 266 acides aminés chez l'humain pour un poids moléculaire de 39 kDa (156). Elle a trois domaines principaux ayant chacun un repliement de type dsRBD liés par de longs peptides flexibles (Figure 1.5). Ce type de domaine est surtout connu pour lier le squelette ribose-phosphate des acides nucléiques qui sont généralement retrouvés sous forme d'une hélice de type A (139). Cependant, certains dsRBD ne possèdent pas les résidus basiques nécessaires à l'interaction avec l'ARN. Ceux-ci peuvent alors généralement interagir avec des protéines (157). C'est le cas du domaine dsRBD3 de TRBP (158)(Figure 1.5).

L'arrangement des trois domaines en tandem liés par des régions flexibles permet à TRBP de s'adapter à une très grande variété de substrat. Cela confère à TRBP l'habilité de lier pratiquement tout ARN qui forme une hélice de plus de 11 pb (159). Sa particularité provient de son dsRBD3 pouvant interagir avec des protéines (160). En pouvant lier l'ARN et les protéines, TRBP permet de rapprocher ou apporter un substrat vers une protéine (161). En effet, TRBP peut interagir, à la fois, avec Dicer et un pre-miRNA et faciliter l'interaction entre Dicer et le pre-miRNA (162). Cette interaction permet de faciliter grandement la reconnaissance et la liaison du substrat par Dicer.

## 1.5.2. Fonction de TRBP dans la régulation de Dicer

C'est en ayant la capacité de lier l'ARN et Dicer simultanément que TRBP peut influencer la maturation des miARN par Dicer. TRBP a la capacité d'influencer plus que la simple activité : elle peut aussi affecter la position à laquelle le pre-miARN sera clivé et même faciliter le clivage d'ARN n'ayant pas les caractéristiques parfaites d'un pre-miRNA (par exemple, des long ARN double-brin) (163, 164).

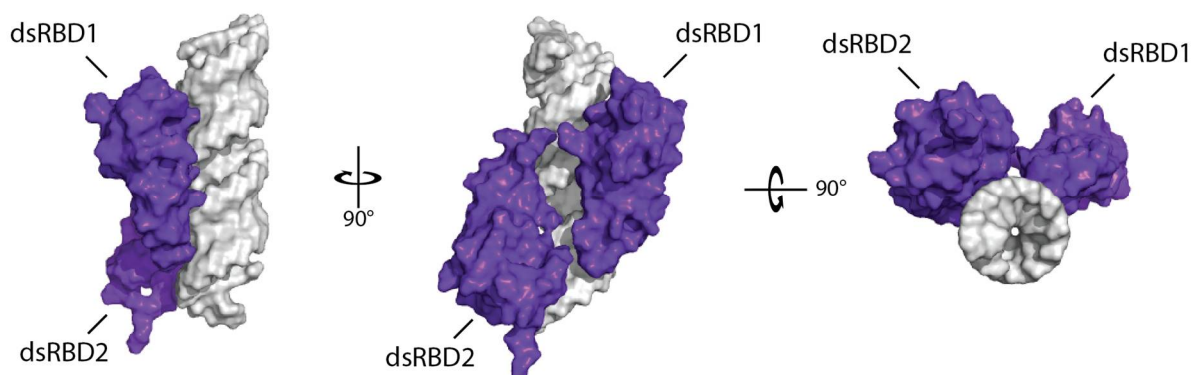
Premièrement, TRBP influence l'interaction générale de Dicer et son substrat. Il a été montré par des expériences de liaison sur filtre (*filter binding assay*) que la constante de liaison apparente de Dicer pour l'ARN était presque 1000 fois plus forte en présence de TRBP, passant d'environ 1 à 30 nM à 50 pM (140). Des cinétiques de répétition simple et multiple (*single turnover* et *multiple turnover*) ont aussi montré une augmentation générale de la vitesse de maturation des pre-miRNA (140, 163). De plus, des études d'interactions à molécule unique (*single-molecule*) ont montré que le temps d'interaction entre un ARN (ARNt ou pre-miRNA) est diminué en présence de TRBP, ce qui aide les complexes inactifs à se dissocier rapidement afin de laisser place à une re liaison qui sera peut-être active (161). Cependant il manque toujours un modèle global de l'activation de Dicer par TRBP.

Ensuite, TRBP peut aussi modifier la position à laquelle le pre-miRNA sera clivé (164-166). Des études de criblage et séquençage de miARN totaux à partir de lignées cellulaires sauvages et dans lesquelles TRBP a été réprimé ont montré que certains miRNA voyaient leur taille modifiée par plus ou moins 2 nt (164, 166). Cela peut avoir des conséquences très importantes de sorte que le brin complémentaire au miRNA puisse être chargé sur Argonaute et reconnaître différents substrats. Cela consiste en un mécanisme de contrôle permettant la formation potentielle de deux miARN différents dépendant de la présence de TRBP. Évidemment, tous les pre-miRNA ne sont pas sujet à une telle régulation. Une nouvelle étude a déterminé quels sont les prérequis pour que ce changement de taille puisse survenir (164). Il s'agit du positionnement des acides nucléiques non-appariés formant des protubérances dans la tige du pre-miRNA. Cependant, il n'y a pas encore de règles parfaitement établies pour prédire le site de clivage par rapport à la structure de la tige des pre-miARN. La prédiction du site de clivage serait une addition très utile aux outils bio-informatiques de prédiction de miRNA.

Finalement, TRBP pourrait aussi faciliter la maturation de substrats non-canoniques en miRNA. Des études *in vitro* ont montré que Dicer avait beaucoup plus de facilité à cliver un long ARN double brin en présence de TRBP (140, 163). Des études chez les différents homologues de Dicer chez la drosophile on aussi montré des résultats similaires (51, 167).

### 1.5.3. Interaction de TRBP avec les pre-miARN

Comme mentionné précédemment, TRBP interagit avec les pre-miARN par ses deux premiers domaines (dsRBD1 et dsRBD2; Figure 1.5). L'interaction d'un dsRBD avec l'ARN nécessite que l'ARN soit sous forme d'hélice A. Comme les dsRBD font contact avec trois sillons sur une même face de l'hélice d'ARN, soit sillon mineur-majeur-mineur. L'ARN doit avoir au minimum un tour complet d'hélice (11 pb) pour accommoder la liaison de ce domaine (168).



**Figure 1.10. Interaction de TRBP avec un duplex d'ARN de 23 pb.** Le dsRBD2 de TRBP lie l'extrémité la plus stable du duplex. Le dsRBD1 se positionne sur environ  $\frac{1}{4}$  de tour d'hélice plus loin sur l'ARN. Adapté à partir de PDB:5N8L (159)

La liaison des deux domaines dsRBD à l'ARN présente une affinité de l'ordre de 100 picomolaire (pM) à 10 nM nanomolaire (nM) selon les études (158, 169, 170). Ces études réalisées avec des ARN d'au moins 22 pb montrent la liaison d'une deuxième molécule de TRBP à l'ARN. Ce deuxième site présente une affinité diminuée par rapport au premier. Cela montre qu'un pre-miARN pourrait théoriquement lier deux molécules de TRBP. Cependant, la tige des pre-miARN est de longueur minimale pour permettre cette conformation (Figure 1.10).

Il ne faut cependant pas exclure que chacun des domaines de TRBP peut lier l'ARN indépendamment. Le domaine 2 aurait une affinité très légèrement plus élevée que le domaine 1 pour l'ARN et les deux domaines peuvent lier l'ARN de façon indépendante avec une affinité calculée entre 100 et 500 nM (158, 169). Cela indiquerait que TRBP puisse lier l'ARN avec un seul de ces deux domaines sur un ARN ayant une hélice plus courte.

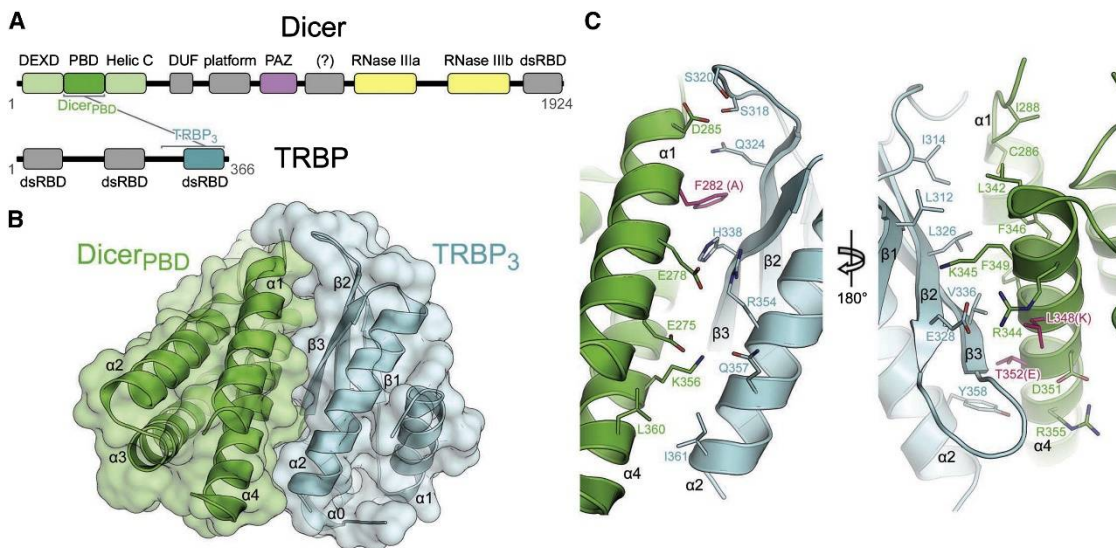
Quelques structures de TRBP et ses orthologues ont été résolues. La première structure correspond au domaine 2 de TRBP en présence d'ARN (171). Cette étude a montré que ce domaine était un dsRBD canonique et a aussi offert une comparaison de structure avec des domaines de protéines orthologue HYL1 chez *Arabidopsis thaliana* et Xlrpba de *Xenopus laevis* (139, 171). Ensuite, un autre groupe a résolu les structures des dsRBD1 et 2 séparément en absence d'ARN (169). La différence majeure entre les deux domaines est la présence d'un résidu tryptophane dans le domaine 2 qui rend le repliement du domaine plus stable que celui du domaine 1. La structure des deux domaines de TRBP liée à un ARN de 19 pb montre que les deux domaines lient l'ARN de manière antiparallèle et occupe seulement un côté de l'hélice, laissant environ 180° de la surface accessible pour, entre autres, la liaison de Dicer (159).

Aucune spécificité de séquence n'a, à ce jour, été attribuée à TRBP. Par contre, TRBP peut lier plus favorablement les séquences ayant des paires de bases stables (62, 63). Pour montrer cet effet, des ARN ayant des paires de bases plus stables ou moins stables aux deux extrémités de l'hélice ont été utilisés. Par des expériences de liaison et de réticulation (*crosslinking*), il a été montré que TRBP lie préférentiellement l'extrémité possédant l'énergie libre la plus basse. Une analyse des séquences connues de miARN montre aussi que les duplex de miARN possèdent généralement une extrémité riche en U et une autre riche en G, ce qui formerait un déséquilibre de stabilité (62). Une autre étude avec Loqs-PD, un orthologue de TRBP chez la drosophile, montre que celle-ci préfère généralement lier les extrémités des hélices d'ARN et que, en présence de TRBP, Dicer préfère l'extrémité la moins stable tandis que Loqs-PD lie l'autre extrémité (63). Il a été proposé que ce mécanisme serait à l'origine du choix du brin qui est chargé sur Argonaute.

### 1.5.4. Interaction de TRBP avec Dicer

L'interaction de TRBP avec Dicer a d'abord été proposée en 2005 par Chendrimada et Haase (60, 61) par des expériences de co-immunoprécipitation. Des résultats similaires ont aussi été montrés chez la drosophile et le nématode où des orthologues de TRBP peuvent aussi interagir avec Dicer (154, 155).

Le site d'interaction entre Dicer et TRBP a ensuite été identifié. Le domaine de TRBP impliqué dans l'interaction avec Dicer est le domaine dsRBD3 (Figure 1.11). Celui-ci interagit avec le domaine hélicase de Dicer, sur un fragment de 165 acides aminés situé entre les sous-domaines ATPase et hélicase (162). La structure cristalline de cette région d'interaction a été résolue et montre une interaction à caractère principalement hydrophobe stabilisée par quelques interactions électrostatiques (126). Le  $K_d$  de cette interaction est évalué entre 1 et 10 nM (126). Cette même région de Dicer est responsable pour l'interaction avec le domaine 3 de PACT.



**Figure 1.11. Interaction entre TRBP et Dicer.** **A.** Région d'interaction entre TRBP et Dicer sur les séquences primaires. **B.** Représentation du squelette peptidique des deux régions d'interaction à partir de la structure cristalline. **C.** Vue de l'interface d'interaction entre TRBP et Dicer. Adapté de (126).

Des études de cryo-microscopie électronique (cryo-EM) sur Dicer montrent des différences de conformation du domaine hélicase en présence de TRBP ou PACT (122). En effet, Dicer peut adopter une conformation ouverte ou fermée et la liaison de l'ARN et TRBP sur Dicer favorise la conformation ouverte de Dicer alors que l'ARN ou TRBP seuls favorise plutôt la conformation fermée. Ce changement de conformation peut expliquer en partie le rôle d'activation que TRBP a sur Dicer.

Chez la drosophile, Loqs-PD interagit avec Dicer via un domaine très différent de celui de TRBP. Au lieu d'avoir un domaine de type dsRBD en C-terminal, Loqs-PD a un domaine de 22 aa comportant un motif FDF. Ce motif peut interagir, comme TRBP, avec le domaine hélicase de Dicer-2 de la drosophile (172). Cependant, la région d'interaction sur Dicer-2 est dans le sous-domaine hel2 du domaine hélicase (167). Malgré leur différence, Loqs-PD et TRBP interagissent tous les deux avec le domaine hélicase de Dicer et peuvent activer le clivage des substrats.

### 1.5.5. Dimérisation de TRBP

En plus d'interagir avec Dicer, TRBP peut homodimériser ou hétérodimériser avec PACT, PKR et Merlin. Ces interactions se font via les différents domaines de TRBP. TRBP peut interagir avec lui-même *in vitro*. L'homodimérisation se fait via les domaines dsRBD 1 et 2 (160). Le  $K_d$  de cette interaction a été mesuré à 54  $\mu$ M par ultracentrifugation analytique (169). Il a été supposé que TRBP puisse se lier à l'ARN et ensuite multimériser (158). Il n'est cependant pas clair si TRBP peut interagir avec l'ARN et former des dimères simultanément. Des études structurales de dimères de TRBP pourraient apporter cette réponse. D'un autre côté, avec un  $K_d$  aussi élevé, il serait surprenant que cette dimérisation ait un gros effet *in vivo*.

TRBP et PACT peuvent interagir ensemble via leurs 3 domaines, quand les domaines dsRBD1 et dsRBD2 interagissent entre eux de manière interchangeable et les domaines dsRBD3 interagissent ensemble (160). Ces deux protéines sont extrêmement similaires et peuvent toutes les deux activer Dicer (163). Leur différence provient de l'écart entre chacun des domaines et de la capacité de leur dsRBD3 à homodimériser. En effet, PACT peut homodimériser via le dsRBD3 alors que TRBP ne peut pas (160, 173). La différence provient du fait que TRBP contient seulement la moitié des résidus essentiels à la dimérisation, ce qui l'empêche d'interagir

avec lui-même tandis que PACT a la totalité des résidus, ce qui lui permet d'interagir avec TRBP ou avec lui-même (173). Cette interaction est importante car elle empêche ces protéines d'aller exercer leurs fonctions lorsqu'elles s'associent ensemble (174).

### **1.5.6. TRBP influence les produits de clivage de Dicer**

Dès le début de la caractérisation des miARN, il a été observé que la taille des miARN matures peut parfois varier. Ces miARN variables sont appelés isomiR (175). Ces variations sont importantes car elles peuvent influencer le brin qui sera chargé sur argonaute en changeant, entre autres, la stabilité des paires de base à l'extrémité du miARN (176, 177). TRBP peut influencer le site de clivage privilégié par Dicer et promouvoir la production d'isomiR. Les premiers résultats en ce sens proviennent d'études *in vitro* comparant les activités de Dicer seule ou en présence de TRBP (165). Ce rôle de TRBP a été confirmé *in vivo* où des expériences de séquençage ont montré que l'absence de TRBP changeait l'équilibre des isomiR dans les cellules HeLa (166). Finalement, une étude a tenté de trouver le facteur permettant à un substrat de voir son site de clivage modifié par la présence de TRBP. Chez la drosophile, Loqs-PB peut modifier le site de clivage de Dicer si l'ARN dépend de la position de nucléotides non-appariés sur la tige des pre-miARN (164). Cependant il n'y a pas de règle générale concernant la position de ces nucléotides non-appariés et le résultat qui en découle car c'est unique à chaque substrat. Chez l'humain, il n'est pas encore clair si ce changement de site de clivage par TRBP est dicté par des changements sur la tige de l'ARN.

## **1.6. Études enzymatiques à l'état stationnaire**

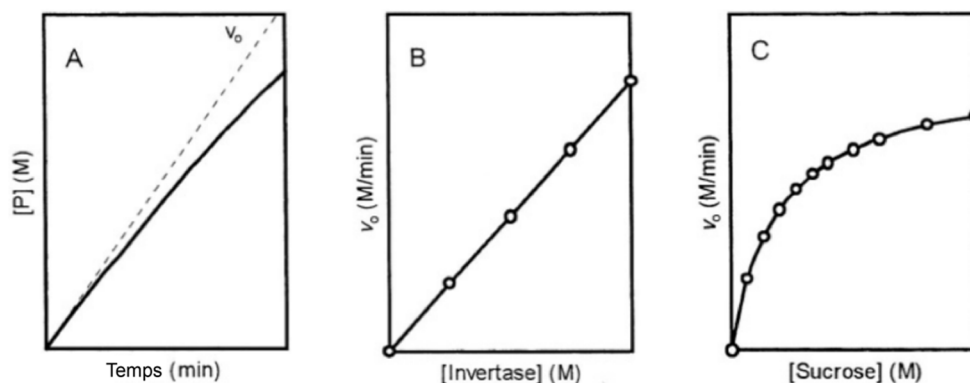
Les études cinétiques offrent une bonne façon de comprendre le mécanisme d'une enzyme afin de pouvoir émettre des hypothèses sur sa régulation et comparer son activité pour différents substrats. Puisque Dicer peut cliver plusieurs substrats différents et être régulée par d'autres protéines, elle devient donc très intéressante à étudier par des études enzymologiques. Une bonne façon d'étudier les activateurs et inhibiteurs enzymatiques est par des expériences de cinétique à l'état stationnaire. À partir des résultats, il est ensuite possible de dériver le modèle enzymatique prenant en compte les interactions entre les différentes molécules et leurs



effets sur la catalyse. Cette section vise à résumer la théorie nécessaire à ce type d'étude enzymologique.

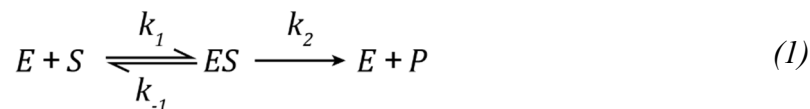
### 1.6.1. Les travaux de Michaelis et Menten

Les travaux de Henri, Michaelis et Menten ont permis d'établir les règles de base pour l'étude des cinétiques enzymatiques (178, 179). Puisque que la vitesse d'une réaction varie en fonction du temps dû à la déplétion du substrat lors de sa transformation, ils ont établi que la vitesse doit être mesurée au début de la réaction, soit dans la phase linéaire. Cette donnée de vitesse est définie comme la vitesse initiale de réaction (Figure 1.12 A).



**Figure 1.12. Résumé des observations de Michaelis-Menten** (A) Apparition des produits en fonction du temps dans une réaction enzymatique où  $v_0$  est la vitesse initiale. (B) Relation linéaire entre la vitesse initiale et la concentration d'enzyme utilisée. (C) Relation non-linéaire entre la concentration de substrat et la vitesse initiale. Provient de (180).

Ils ont ensuite montré que la vitesse initiale de réaction varie de façon linéaire avec la concentration d'enzyme, tant que la concentration d'enzyme ( $[E]_t$ ) dans la réaction est beaucoup plus petite que la concentration de substrat ( $[S]_t$ ; Figure 1.12 B). Finalement, ils ont déterminé la relation non-linéaire entre la vitesse initiale de la réaction et la concentration de substrat (Figure 1.12 C). Ils ont ensuite présenté une réaction simple comme celle-ci,



où la réaction est séparée par un équilibre rapide de liaison suivi d'une étape lente irréversible, où  $k_{-1} \gg k_2$ . Cette restriction indique que la dissociation du substrat se fait plus rapidement que sa transformation. La vitesse initiale de cette réaction est,

$$v_0 = k_2[ES] \quad (2)$$

L'équilibre réactionnel,  $K_S$  est donc,

$$K_S = \frac{k_{-1}}{k_1} = \frac{[E][S]}{[ES]} = K_d \quad (3)$$

où,  $K_S$  est égal à la constante de dissociation du complexe [ES]. Après réarrangement,

$$[ES] = \frac{[E][S]}{K_S} \quad (4)$$

Il faut ensuite appliquer la loi de la conservation de masse pour exprimer cet équilibre en fonction de  $[E]_t$  et  $[S]_t$ :

$$\begin{aligned} [E]_t &= [E] + [ES] \\ [S]_t &= [S] + [ES] \approx [S] \end{aligned} \quad (5)$$

Après substitution dans (5), on obtient,

$$[ES] = \frac{([E]_t - [ES])[S]_t}{K_S} \quad \text{ou} \quad [ES] = \frac{[E]_t[S]_t}{K_S + [S]_t} \quad (6)$$

Finalement, on peut substituer (5) dans (3) et obtenir :

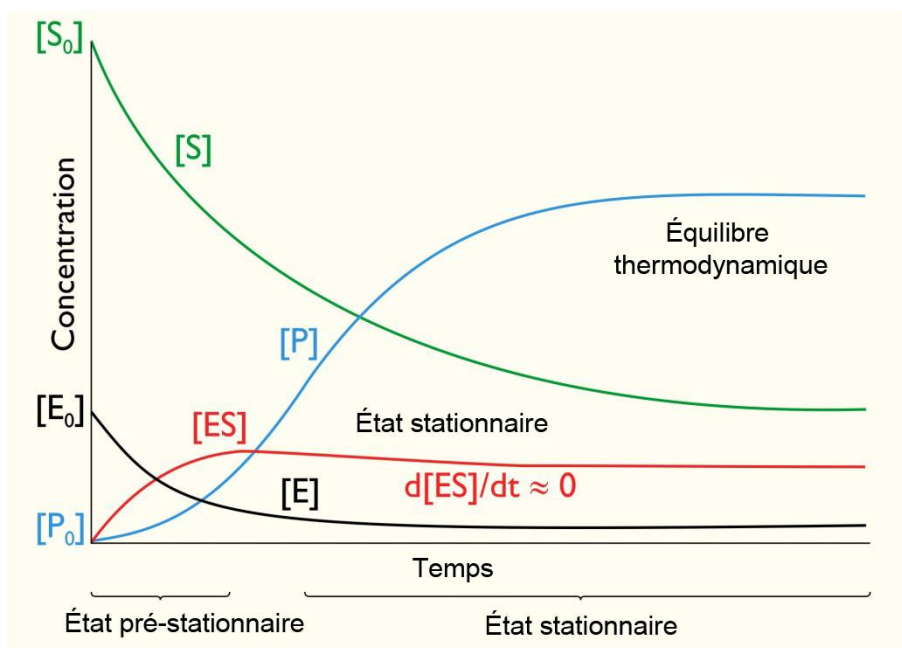
$$v_0 = \frac{k_2[E]_t[S]_t}{K_S + [S]_t} \quad (7)$$

Cette équation exprime l'effet de la concentration de substrat sur la vitesse initiale dans le cas d'une réaction irréversible à un substrat où la réaction est limitée par la formation du complexe [ES] c'est-à-dire où  $k_2 \gg k_{-1}$ . Évidemment, la plupart des enzymes ont des mécanismes bien plus complexes. Certaines sont réversibles, d'autres ont plus d'un substrat et leur étape limitante n'est pas toujours au niveau de la formation de [ES]. Ces travaux ont cependant permis de mettre un système simple en contexte et ont permis la proposition de nouveaux modèles plus généraux qui peuvent s'appliquer à une plus grande variété d'enzymes.

## 1.6.2. Approximation de l'état stationnaire

L'approximation de l'état stationnaire a été décrite par Briggs et Haldane en 1925 (181). Leurs travaux ont permis d'étendre la portée de l'équation de Michaelis-Menten. La particularité

de ces enzymes est que l'apparition des produits au début de la réaction ne varie plus de façon linéaire. Par contre, il est encore possible d'observer une phase linéaire dans l'apparition des produits, un peu après le début de la réaction. C'est exactement ce que Briggs et Haldane ont observé et ils ont supposé que, pour une cette courte période durant la réaction, il y avait un moment de pseudo-équilibre où  $[ES]$  était constante (Figure 1.13). Cette phase linéaire a été appelée « état stationnaire » tandis que la phase précédant celle-ci a été nommée « état pré-stationnaire ».



**Figure 1.13. Concentration des espèces réactionnelles au cours d'une réaction.** Formation du complexe  $[ES]$  au début de la réaction lors de l'état pré-stationnaire suivi de l'état stationnaire où  $[ES]$  reste constant durant la réaction. Adapté de (<http://elte.prompt.hu/sites/default/files/tananyagok/IntroductionToPracticalBiochemistry/ch09s02.html>)

Les vitesses initiales mesurées à l'état stationnaire montrent exactement le même type de comportement que ce qui a été décrit par Henri, Michaelis et Menten et peuvent être décrits par l'équation (7). La différence est que  $k_{-1} \approx k_2$  et  $K_S$  est maintenant décrit comme :

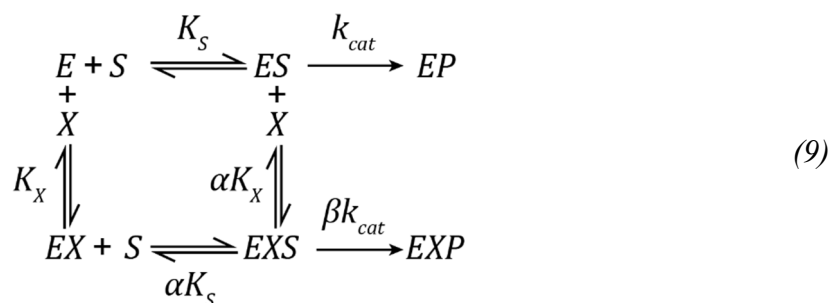
$$K_S = \frac{k_2 + k_{-1}}{k_1} \quad (8)$$

Cette formulation permet donc de s'adapter à la plupart des enzymes comme Dicer ou autres enzymes complexes.

### 1.6.3. Activation et inhibition enzymatique

Les modificateurs enzymatiques peuvent fonctionner soit en tant qu'activateur ou inhibiteur. Les modificateurs les plus connus et décrits dans les livres généraux de biochimie sont les modificateurs linéaires. Ces modificateurs sont soit des inhibiteurs pouvant complètement inhiber l'activité enzymatique ou des activateurs essentiels, sans lesquels l'enzyme ne peut pas fonctionner. Cependant, il y a une catégorie de modificateurs n'ayant pas la capacité d'inhiber complètement l'activité de l'enzyme, c'est-à-dire que le complexe enzyme-inhibiteur (EI) est encore fonctionnel. Il y a aussi les activateurs non-essentiels, sans lesquels l'enzyme peut encore fonctionner. Cette catégorie de modificateurs n'a pas un effet linéaire et forme la catégorie des modificateurs hyperboliques (180).

Comme mentionné plus haut, un modificateur est hyperbolique lorsque le complexe enzyme-modificateur (EX) et l'enzyme seule (E) peuvent catalyser la réaction. Dans son schéma le plus général, les modificateurs hyperboliques peuvent être représentés comme :

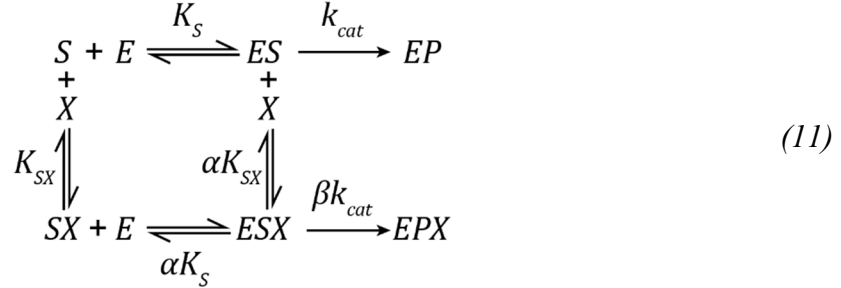


Où  $\alpha$  représente l'effet de X sur la liaison de S et cet effet est le même indépendamment de l'ordre de liaison puisqu'il est assumé que ces complexes sont à l'état stationnaire. La constante  $\beta$  est l'effet de X sur la vitesse de réaction à saturation de S. En général, on parle d'activation si  $0 < \alpha < 1$  et d'inhibition lorsque  $\alpha > 1$  et il n'y a pas d'effet si  $\alpha = 1$ . Dans le cas de  $\beta$ , on observe de l'inhibition lorsque  $0 < \beta < 1$  et activation lorsque  $\beta > 1$  et, encore une fois, aucun effet lorsque  $\beta = 1$ . On peut cependant avoir un mécanisme hybride où, par exemple,  $0 < \alpha < 1$  et  $0 < \beta < 1$ . Ce genre de modificateur peut être un activateur ou un inhibiteur dépendamment de la concentration de S (182).

Il y a aussi des modificateurs qui lient le substrat au lieu de l'enzyme où, S devient SX :



On peut représenter un schéma enzymatique simple pour ce type de modificateur :



La particularité de ce système est que, lors de sa caractérisation expérimentale, il est généralement impossible d'utiliser des concentrations de modificateur de façon à ce que  $[S] \gg [X]$  soit toujours valide. On ne peut donc pas assumer que  $[S] \approx [S]_t$  et  $[X] \approx [X]_t$ . Il faut donc exprimer  $[S]$  et  $[X]$  par la loi de conservation de masse où  $[S]_t = [S] + [SX]$ .

#### 1.6.4. Dérivation des équations cinétiques

La manière la plus simple d'obtenir les équations cinétiques en termes des différentes constantes d'équilibre est de commencer par définir l'équation de la vitesse de réaction des complexes pouvant se transformer en produit. En prenant par exemple le schéma général de modification enzymatique présenté en (9) :

$$v = k_{cat}[ES] + \beta k_{cat}[ESX] \quad (12)$$

Ensuite, on définit l'équation de conservation de masse correspondant à  $[E]_t$  :

$$[E]_t = [E] + [ES] + [EX] + [ESX] \quad (13)$$

Puis, on divise (12) par (13) :

$$\frac{v}{[E]_t} = \frac{k_{cat}[ES] + \beta k_{cat}[ESX]}{[E] + [ES] + [EX] + [ESX]} \quad (14)$$

Où les différents complexes sont ensuite exprimés en fonction de  $[S]$  et  $[X]$  :

$$\frac{v}{[E]_t} = \frac{k_{cat} \frac{[E][S]}{K_S} + \beta k_{cat} \frac{[E][S][X]}{\alpha K_S K_{SX}}}{[E] + \frac{[E][S]}{K_S} + \frac{[E][X]}{K_{SX}} + \frac{[E][S][X]}{\alpha K_S K_{SX}}} = \frac{k_{cat} \frac{[S]}{K_S} + \beta k_{cat} \frac{[S][X]}{\alpha K_S K_{SX}}}{1 + \frac{[S]}{K_S} + \frac{[X]}{K_{SX}} + \frac{[S][X]}{\alpha K_S K_{SX}}} \quad (15)$$

Cette équation peut être réarrangée sous une forme ressemblant à l'équation de Michaelis-Menten (7) :

$$\frac{v}{[E]_t} = \frac{k_{cat} \left(1 + \beta \frac{[X]}{\alpha K_X}\right) [S]}{K_S \left(1 + \frac{[X]}{K_X}\right) + [S] \left(1 + \frac{[X]}{\alpha K_X}\right)} \quad (15)$$

Le même type de procédure peut être appliqué à plusieurs schémas enzymatiques. Le chapitre 3 de cette thèse traite de l'effet de TRBP sur Dicer et les modèles cinétiques ont été dérivés de cette manière.

### 1.6.5. Conditions expérimentales à l'état stationnaire

L'étude d'enzyme à l'état stationnaire afin de dériver les différents paramètres cinétiques et établir des modèles enzymatiques pour des modificateurs est possible mais il faut d'abord s'assurer que les conditions expérimentales respectent les règles générales permettant d'assumer l'état stationnaire. Cette section vise à lister brièvement ces règles de base :

- 1)  $[S] \gg [E]$ . Cette condition permet d'assumer  $[S] \approx [S]_t$  et simplifie grandement les équations cinétiques. Dans le cas d'une enzyme rapide ayant un  $k_{cat} > 30 \text{ min}^{-1}$ , cette condition est très facile à respecter car la concentration d'enzyme nécessaire à l'obtention d'un clivage significatif ( $\sim 10\%$  du substrat total) en un temps raisonnable ( $< 3\text{h}$ ) est très faible. Cependant, pour des enzymes plus lentes comme les RNase III, cela peut être difficile d'observer une apparition significative de produits dans un temps raisonnable à toutes les concentrations de substrat testées. Il est cependant impératif d'avoir une concentration de substrat au moins  $20\times$  supérieure à la concentration d'enzyme. Une stratégie impliquant d'utiliser une quantité différente d'enzyme à chaque concentration de substrat et de présenter les résultats de vitesse initiales normalisés en fonction de la concentration d'enzyme sous forme  $(v_0/[E]_t)$ . Cette normalisation est possible car la vitesse de réaction varie de façon linéaire avec la concentration d'enzyme, tel que discuté dans la section 1.6.1.
- 2) Mesurer  $v_0$  lorsque moins de 10% de  $[S]_t$  est transformé. Ce paramètre est important car l'hydrolyse du substrat suit généralement une relation exponentielle inverse et ralentit lorsque la concentration de substrat restante diminue. De plus, dans le cas où la réaction inverse ou la liaison des produits à l'enzyme est possible, il est possible d'observer un

ralentissement de la réaction dû à un phénomène d'inhibition par les produits. Il est donc impératif de garder la concentration de produits au minimum.

- 3) S'assurer d'être à l'état stationnaire. Bien que ce soit généralement assez difficile d'observer l'état pré-stationnaire des enzymes puisque c'est une étape rapide, il est possible que certains points de temps observés au début d'une réaction ne soient pas totalement dans la phase linéaire de l'état stationnaire. L'état pré-stationnaire est généralement observé avant que les molécules d'enzymes présentes dans la réaction aient effectué leur premier revirement « *turnover* ». La meilleure manière de s'assurer de l'état-stationnaire est donc d'utiliser les points de temps se retrouvant après quelques revirements du lot d'enzyme. Le nombre de revirements correspond à  $[P]/[E]_t$ . Théoriquement, un seul revirement est nécessaire afin d'être à l'état stationnaire. Cependant, il est préférable de mesurer la vitesse de réaction avec les points correspondants à plus de cinq revirements lorsque c'est possible.
- 4)  $[X] \gg [E]$ . Lors de l'étude de modificateurs enzymatiques, il est important aussi pour le modificateur d'être en haute concentration par rapport à l'enzyme, et cela, pour les mêmes raisons que celles mentionnées en 1). Cette condition est extrêmement difficile à respecter lorsque la constante de dissociation du modificateur est du même ordre ou plus petite que les autres constantes d'équilibre du système. Ce cas s'appelle modification à liaison solide « *tight-binding inhibition or activation* ». Dans le cas où il est impossible d'avoir au moins  $[X] \geq 10 \times [E]$ , une correction supplémentaire doit être appliquée car  $[X]_t = [X] + [XE] \neq [X]$ .

## 1.7. Objectifs de la thèse

L'objectif global de notre laboratoire est d'étudier les mécanismes de régulation de la maturation des miARN. Au cours des dernières années, nous avons publié deux articles sur la reconnaissance des pre-miARN par Lin28 (93, 94). Ce projet nous a permis d'acquérir les connaissances de base sur les miARN, leur régulation et quelques méthodes essentielles à leur étude. La suite logique à ces deux premières études était d'inclure Drosha et Dicer afin de mesurer l'effet de régulateurs comme Lin28 sur leur activité. Nous avons donc décidé de se concentrer d'abord sur Dicer et l'effet de TRBP, son principal cofacteur, sur son activité.

Nous avons déjà utilisé de la Dicer de source commerciale pour un des résultats publiés avec Lin28. Cependant, ces sources de Dicer sont, non seulement, dispendieuses mais on ne connaît pas très bien leur pureté et leur concentration. Cela rend difficile des études enzymatiques à large échelle. Nous avons donc décidé de purifier Dicer au laboratoire et, puisque les protocoles existants utilisant un système d'expression par baculovirus en cellules de drosophile étaient très limités en termes de rendement, nous avons choisi de tenter d'améliorer les rendements de purification en utilisant un système d'expression basé sur les cellules humaines en suspension et de proposer un protocole de purification permettant une grande pureté avec un rendement maximal. Nous croyons donc que l'utilisation d'un système de surexpression en cellule humaine permettra l'expression d'une plus grande quantité de Dicer, augmentant ainsi le matériel de départ, et permettra, à terme, de réduire la taille des cultures utilisées pour produire quelques milligrammes de Dicer recombinante. De plus, nous croyons qu'il est possible de maximiser la récupération et la stabilité de Dicer à partir des cellules en prenant soin d'optimiser chacune des étapes de purification. Finalement, nous voulons montrer que la Dicer finale est fonctionnelle et homogène par des études de liaison et d'activité avec des pre-miARN et une caractérisation biophysique par SEC-MALS et microscopie électronique à coloration négative.

En un deuxième temps, nous voulons étudier l'effet de protéines régulatrices sur l'activité de Dicer. Au lieu de tester directement Lin28, il est important de caractériser l'effet de TRBP sur Dicer. Puisque TRBP est essentielle à la maturation des miARN, il est fort probable que celle-ci soit très souvent présente dans un contexte physiologique. Il devient donc important de connaître le rôle de TRBP sur Dicer et d'étudier les autres régulateurs en sa présence. Des résultats obtenus lors des expériences de cinétique suite à la purification de Dicer nous permettent de supposer que Dicer ne serait pas une enzyme très rapide puisqu'elle a un  $K_M$  élevé et un faible  $k_{cat}$ , seule dans la cellule. Dicer aurait donc besoin d'une protéine activatrice afin d'augmenter son activité et réduire son  $K_M$  et nous croyons que TRBP est chargée de ce rôle. Afin de tester cette hypothèse, nous allons d'abord confirmer les résultats sur le  $K_M$  obtenus dans la première étude avec d'autres substrats. Ensuite, nous proposerons un modèle cinétique de l'effet de TRBP sur Dicer. Ce modèle sera forgé à partir des informations connues sur TRBP et sera ensuite validé par plusieurs expériences de liaison et de cinétiques à l'état-stationnaire.



Le modèle final représentera, pour la première fois, l'effet de TRBP sur la vitesse de réaction de Dicer en fonction de la concentration de substrat et offrira un regard nouveau sur l'activation de Dicer par TRBP.

## **2. Production de grandes quantités de Dicer humaine par transfection de cellules HEK293-EBNA1 en suspension.**

### **2.1. Mise en contexte**

Les travaux nécessaires à ce manuscrit ont été effectués entre l'automne 2014 à l'automne 2017. Au moment de commencer ces travaux, encore aucune structure à haute résolution n'était publiée pour Dicer, tout organisme confondu. Cependant, les premières structures à basse résolution par cryo-EM venaient d'être publiées, montrant la forme générale et permettant le positionnement des domaines dans l'espace. Plusieurs études ont permis de connaître les substrats optimaux de Dicer, la fonction générale de chacun des domaines de Dicer, ainsi que le rôle général d'action de Dicer par TRBP. La plupart des études publiées produisaient Dicer à partir d'un baculovirus dans des cellules de drosophile. Ce protocole permet effectivement la production de Dicer recombinante mais les rendements de productions sont faibles (49, 50, 60, 183-185).

Cette première étude avait donc pour but d'améliorer les rendements de purification de Dicer dans le but de faciliter les études structurales et biochimiques subséquentes. En ayant une plus grande quantité de Dicer, nous pourrions donc tenter des essais de cristallisation avec plus de facilité et permettre l'utilisation d'autres méthodes requérant une plus grande quantité de matériel sans avoir à produire plusieurs dizaines de litres de culture. Nous avons donc décidé de repartir à zéro et utiliser un système d'expression en cellules humaines dans le but d'augmenter les niveaux l'expression initiaux. Ensuite, nous avons soigneusement élaboré un protocole de purification permettant de maximiser le rendement, la pureté et la stabilité de Dicer.

Pour ces travaux, j'ai (JB) développé, réalisé et analysé la plupart des expériences : le clonage, la culture cellulaire, l'optimisation des conditions de transfection, la purification de Dicer et pre-let-7a-1, les tests de liaison et d'activité ainsi que le SEC-MALS. Dursun Nizam Korkut (DNK) a purifié Dicer et fait l'acquisition de donnée de SEC-MALS (Figure 2.6). Aurélien Fouillen (AF) a réalisé et analysé les expériences de TEM. Soumiya Amellah a contribué techniquement aux travaux de culture cellulaire, de purification de Dicer et aux tests

de liaison avec JB (Figure 2.2c-g et Figure 2.4a). Antonio Nanci, Yves Durocher, Pascale Legault (PL), James G. Omichinski ont porté conseils. JB et PL ont rédigé le manuscrit. Tous les auteurs ont lu et approuvé la version finale.

***High-yield production of human Dicer by transfection of  
human HEK293-EBNA1 cells grown in suspension***

Jonathan Bouvette<sup>1</sup>, Dursun Nizam Korkut<sup>1</sup>, Aurélien Fouillen<sup>1,2</sup>, Soumiya Amellah<sup>1</sup>, Antonio Nanci<sup>1,2</sup>, Yves Durocher<sup>1,3</sup>, James G. Omichinski<sup>1</sup> and Pascale Legault<sup>1</sup>

<sup>1</sup>Département de Biochimie et Médecine Moléculaire, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, QC, Canada,

<sup>2</sup>Département de Stomatologie, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, QC, Canada, H3C 3J7

<sup>3</sup>Human Health Therapeutics Research Centre, National Research Council Canada, Montréal, QC, Canada, H4P 2R2

**Corresponding author:** [pascale.legault@umontreal.ca](mailto:pascale.legault@umontreal.ca)

**Email addresses for all authors in order of appearance:** [jonathan.bouvette@umontreal.ca](mailto:jonathan.bouvette@umontreal.ca), [dursun.korkut@umontreal.ca](mailto:dursun.korkut@umontreal.ca), [aurelien.fouillen@umontreal.ca](mailto:aurelien.fouillen@umontreal.ca), [s.amellah@umontreal.ca](mailto:s.amellah@umontreal.ca), [antonio.nanci@umontreal.ca](mailto:antonio.nanci@umontreal.ca), [yves.durocher@cnrc-nrc.gc.ca](mailto:yves.durocher@cnrc-nrc.gc.ca), [jg.omichinski@umontreal.ca](mailto:jg.omichinski@umontreal.ca), [pascale.legault@umontreal.ca](mailto:pascale.legault@umontreal.ca).

## 2.2. Abstract

**Background:** Dicer is a 219-kDa protein that plays key roles in gene regulation, particularly as the ribonuclease III enzyme responsible for cleaving precursor miRNA substrates. Its enzymatic activity is highly regulated by protein factors, and this regulation can impact on the levels of miRNAs and modulate the behavior of a cell. To better understand the underlying mechanisms of regulation, detailed enzymatic and structural characterization of Dicer are needed. However, these types of studies generally require several milligrams of recombinant protein, and efficient preparation of such quantities of pure human Dicer remains a challenge. To prepare large quantities of human Dicer, we have optimized transfection in HEK293-6E cells grown in suspension and streamlined a purification procedure.

**Results:** Transfection conditions were first optimized to achieve expression levels between 10 and 18 mg of recombinant Dicer per liter of culture. A three-step purification protocol was then developed that yields 4-9 mg of purified Dicer per liter of culture in a single day. From SEC-MALS/RI analysis and negative stain TEM, we confirmed that the purified protein is monomerically pure ( $\geq 98\%$ ) and folds with the characteristic L-shape geometry. Using an electrophoretic mobility shift assay, a dissociation constant ( $K_d$ ) of 5 nM was measured for Dicer binding to pre-let-7a-1, in agreement with previous reports. However, when probing the cleavage activity of Dicer for pre-let-7a-1, we measured  $k_{cat}$  ( $7.2 \pm 0.5 \text{ min}^{-1}$ ) and  $K_M$  ( $1.2 \pm 0.3 \mu\text{M}$ ) values that are much higher than previously reported due to experimental conditions that better respect steady-state assumption.

**Conclusions:** The expression and purification protocols described here provide high yields of monomerically pure and active human Dicer. Cleavage studies of a pre-let-7 substrate with this purified Dicer reveal higher  $k_{cat}$  and  $K_M$  values than previously reported and support the current

view that conformational changes are associated with substrate binding. Large quantities of highly pure Dicer will be valuable for future biochemical, biophysical and structural investigations of this key protein of the miRNA pathway.

**Keywords:** human HEK293-EBNA1 cells, mammalian cell suspension culture, Dicer expression, Dicer purification, pre-miRNA, pre-let-7, Dicer binding, Dicer cleavage assay.

## 2.3. Background

Dicer is a multi-functional protein that plays a critical role in regulating several fundamental cellular processes, including RNA interference, genome integrity, development, and antiviral immunity (for recent reviews see [1, 2]). It is a type III endoribonuclease (RNase) that contributes to RNA interference by processing specific hairpin structures and long double-stranded RNAs into microRNAs (miRNAs) and small interfering RNAs (siRNAs), respectively. In the miRNA pathway, Dicer is responsible for cleaving the precursor-miRNA (pre-miRNA) into a ~22-nucleotide (nt) duplex, from which one strand will be loaded into the RNA-induced silencing complex (RISC) to repress mRNA translation. Currently, it is thought that Dicer produces almost all miRNAs in human cells (~ 2000), which together control over 60% of protein-coding genes [3-7]. The pre-miRNA processing activity of Dicer is known to be highly regulated, and this likely contributes to precisely control miRNA levels and thereby determine cell behavior and cell fate. The cleavage activity is regulated by protein cofactors, such as TRBP and PACT, as well as several RNA-binding proteins known to act either as activators or inhibitors (reviewed in [8]). Moreover, high-throughput studies have identified additional non-coding RNAs that are likely processed by Dicer [9] as well as several pre-miRNA binding proteins that may regulate its cleavage activity [10-12]. At this time, there is a need to integrate the available data into a coherent and detailed mechanistic understanding of Dicer's activities and regulation via *in vitro* biochemical, biophysical and structural studies. However, for such research investigations to be carried out effectively, large quantities of highly pure and active recombinant proteins are needed.

Although bacterial overexpression systems generally provide a simple and fast method to obtain significant amount of recombinant proteins, it has not proven practical for human Dicer

(219 kDa). Eukaryotic expression systems offer an important alternative for expression of such a large eukaryotic protein because they allow for proper folding and post-translational modifications (reviewed in [13]). Over the past 15 years, purification of recombinant human Dicer following expression in insect cells (Sf9) infected by baculovirus has been achieved with yields up to 0.5-1 mg/L culture [14-16]. This has allowed for the *in vitro* characterization of human Dicer's enzymatic activity [14, 15, 17-25] and provided the first descriptions of its three-dimensional structure by cryo-electron microscopy (cryo-EM) at 20-30 Å-resolution [26-28]. Subsequently, the production method in insect cells was improved by systematic optimization of the overexpression and purification steps to yield milligram amounts (3-4 mg/L culture) of highly pure drosophila melanogaster Dicer-2 (dmDicer-2) [29], which was used for its structure determination by cryo-EM at 7-Å resolution. More recently, expression in HEK293 cells grown in suspension was reported as part of a cryo-EM study of human Dicer that allowed structural reconstruction at 4.4-Å resolution [30]. This study provided unprecedented details into Dicer's domain organization as well as its interaction with TRBP and a pre-miRNA substrate. However, the optimization of Dicer production and the yields obtained were not reported. Therefore, it is possible that large-scale expression from mammalian cells grown in suspension could be optimized to provide pure protein at higher yields than currently reported. Such a procedure could be useful for future biochemical and structural investigations as there is still no X-ray or cryo-EM structure of human Dicer at atomic resolution and there are many questions that remain about its mechanism and regulation by protein factors.

Here, we describe an efficient and fast procedure to obtain milligram quantities of pure, monomeric and stable recombinant human Dicer. To overcome the challenge of Dicer over-expression, we optimized transfection and expression in an HEK293-EBNA1 cell line. This



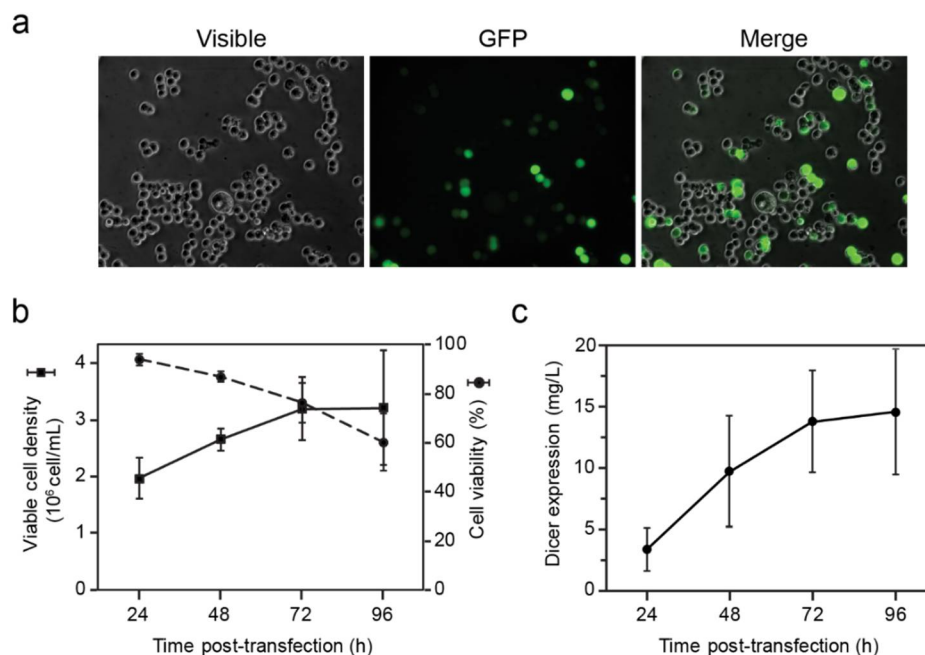
allowed for expression of up to 9 milligrams of Dicer per liter of culture. A purification protocol was designed to minimize sample handling as well as to provide maximal recovery, purity and enzymatic activity. Although we noticed that human Dicer forms multimers at high concentration, our final three-step purification yields highly-pure monomeric Dicer. The structure and activity of the purified Dicer were assessed by size-exclusion chromatography coupled to multi-angle light-scattering and refractive index (SEC-MALS/RI), negative stain transmission electron microscopy (TEM), binding assay and steady-state kinetics. Our kinetic analysis for Dicer cleavage of the pre-let-7a-1 substrate were performed under strict steady state conditions and reveals  $k_{cat}$  and  $K_M$  values that are much higher than previously reported. Overall, this new method will facilitate future biochemical, biophysical and structural characterization of Dicer, including its interaction with RNA substrates and protein regulators.

## **2.4. Results**

### **2.4.1. Transient transfection of recombinant human Dicer**

To achieve high-level expression of recombinant human Dicer, we relied on transient transfection of a suspension-grown HEK-293 cell line expressing the Epstein-Barr virus antigen-1 (HEK293-EBNA1 or 293-6E) with a pTT5 expression vector carrying a CMV promoter and the Epstein-Barr virus origin of replication *oriP* [31-33]. This mammalian expression system has been engineered for maximal expression of recombinant His-tagged proteins in serum-free media [34]. Linear polyethylenimine (PEI) was selected as an efficient and cost-effective transfection reagent. For optimization of transfection conditions, small-scale suspension cultures were grown in serum-free media, either as 2-mL cultures in 6-well plates or 20-mL cultures in 20-mL disposable Erlenmeyer flasks. In initial experiments, we found that

optimal transfection efficiency could be reached when transfecting cells that were diluted to  $0.8 \times 10^6$  cells/mL 24 h in advance using 1 mg of Dicer-expressing pTT5 plasmid with 2 mg linear PEI per L of culture. To estimate the transfection efficiency, we replaced 5% of the total transfected plasmid by a GFP-expressing plasmid and used fluorescence microscopy to count the percentage of GFP-expressing cells. Transfection efficiencies of 30-40% were typically reached at 48 hours post transfection (hpt) (Figure 2.1a). The cell cultures were monitored up to 96 hpt to assess cell viability and viable cell density as well as the intracellular expression of recombinant Dicer. It was found that cell viability decreases below 75% beyond 72 hpt, whereas viable cell density is maintained at  $3.2 \times 10^6$  cell/mL and Dicer expression levels are not significantly changed. Similar results were obtained with large-scale transfections of up to 1 L. Thus, the optimum time for collecting the cells for protein purification was determined to be 72 hpt, taking into consideration that additional time may lead to protein modifications that would adversely affect its activity. At 72 hpt, volumetric expression of the intracellular Dicer protein was  $14 \pm 4$  mg per L of culture on average (Figure 2.1c).



### **Figure 2.1. Transfection of small suspension culture with Dicer-expressing plasmid.**

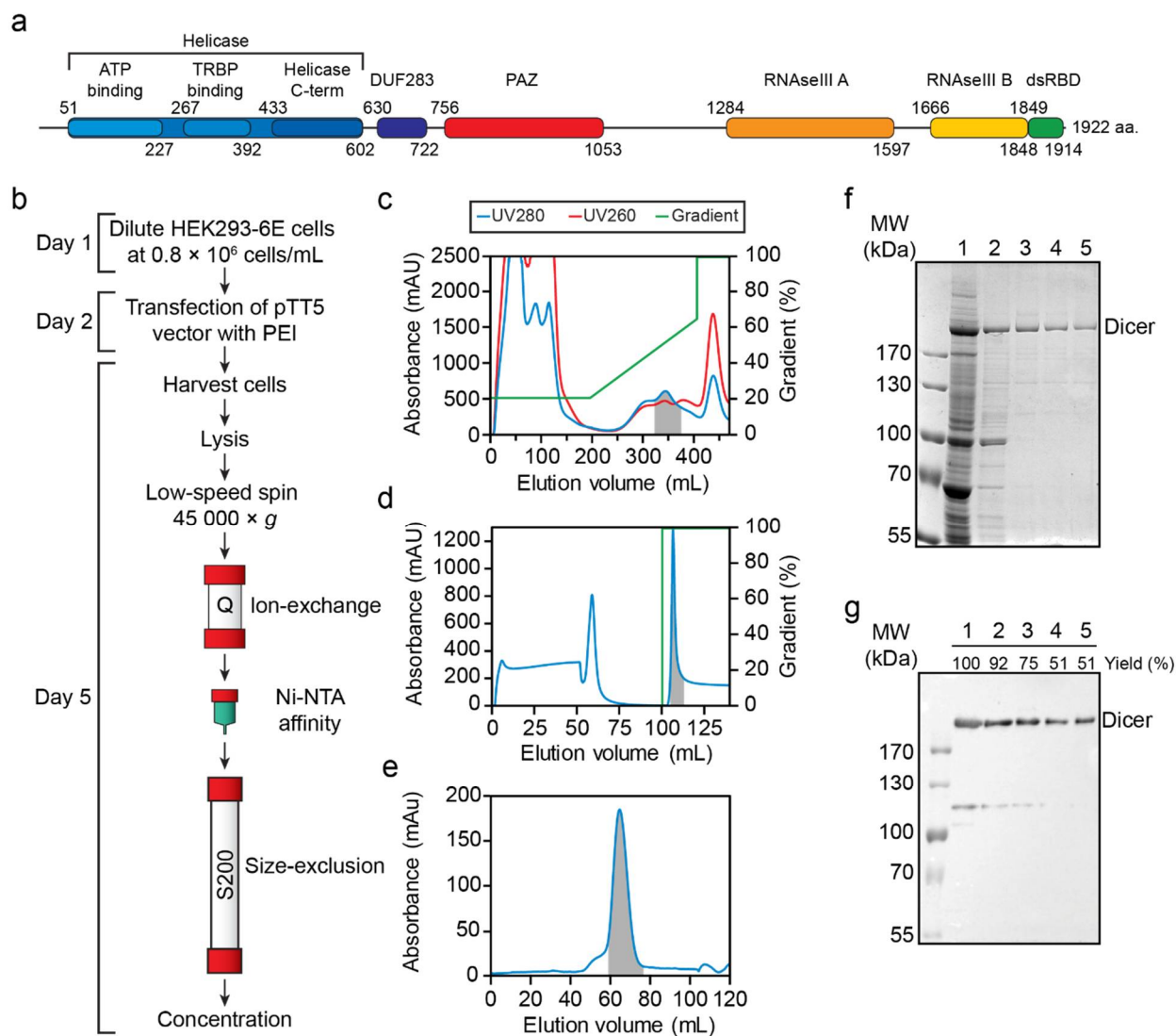
293-6E cells were transfected with PEI:pTT5-DNA complexes at a 1:2 mass ratio and monitored after transfection. **(a)** Transfection efficiency at 48 hpt. A total of  $36 \pm 7\%$  of the cells are expressing GFP ( $n = 4$ ). **(b)** Cell density and viability ( $n = 4$ ). **(c)** Expression of cytoplasmic Dicer. Dicer quantification was derived from Western blot analysis using a standard curve of pure recombinant Dicer ( $n = 4$ ). Transfections in (b) and (c) were performed in 20 mL-cultures grown in 125-mL Erlenmeyer flasks and the results are means  $\pm$  standard deviation values of replicates ( $n$ ) from independent transfections experiments.

### **2.4.2. Purification of recombinant human Dicer**

An efficient multi-step purification protocol was developed to maximize purification yields (Figure 2.2b). While optimizing this protocol, we noticed that Dicer precipitated either at high concentration or when in contact with metallic surfaces such as stainless steel needles, FPLC titanium alloy pump heads and the non-reactive Hastelloy sample cell of the isothermal titration calorimeter. Thus, the final purification protocol was optimized to avoid protein contact with metallic surfaces and minimize sample handling that could lead to protein aggregation and precipitation.

Large-scale expression of His-tagged Dicer (400 mL to 1 L) was typically achieved in a 2.8-L glass Fernbach flask and harvested 72 hpt. For cell lysis, the washed cell pellet was resuspended in a low salt buffer containing 0.1% NP-40, allowing for lysis of the plasma membrane while leaving the nuclei intact. Nuclei and cell debris were then removed by centrifugation and filtering of the resulting supernatant. The first step of purification consisted of an anion-exchange chromatography (Q Sepharose Fast Flow), which helped remove protein and nucleic acid contaminants (Figure 2.2c). The addition of 10% glycerol in the elution buffers was found

critical to purify Dicer free of contamination with Hsp70 (~70-kDa intense band in lane 1 of Figure 2.2f), which is one of the most abundant protein in the HEK-293 cell line [35]. The Dicer-containing fractions were pooled and directly loaded on a HisTrap HP column for purification by immobilized metal affinity chromatography (IMAC; Figure 2.2d). This step, which was carried out in the presence of 0.5 M NaCl and in the absence of glycerol, allowed for the separation of Dicer from the remaining contaminant proteins (Figure 2.2f). The Dicer fractions were then loaded on a size-exclusion column (SEC) to isolate the major monomeric population from multimeric contaminants. When this monomeric population of Dicer was concentrated by ultrafiltration under standard conditions (e.g. 50 mM Tris pH 8.2, 10 mM NaCl/KCl 24:1, 0.5 mM MgCl<sub>2</sub> and 0.5 mM TCEP), it resulted in a significant loss of protein via precipitation. In addition, when Dicer was concentrated prior to SEC purification, substantial amounts of Dicer multimers were detected in the SEC elution profile. This led us to identify conditions that would allow for maximum protein recovery following concentration by ultrafiltration. We found that addition of sucrose and the non-ionic detergent *n*-dodecyl- $\beta$ -D-maltoside (DDM) helped stabilize the monomeric form. Thus, these reagents were added to the mobile phase during SEC (Figure 2.2ef), and fractions containing monomeric Dicer could be readily concentrated at up to 22.5  $\mu$ M (5 mg/mL) without precipitation. Analysis of the purified protein by SDS-PAGE gels stained with Coomassie Blue shows greater than 90% purity (Figure 2.2f), whereas Western blot analysis indicate that ~50% of the protein is recovered from the cytoplasmic fraction (Figure 2.2g). Yields of 4-9 mg of human Dicer per liter of culture were typically obtained from large-scale preparations.



**Figure 2.2. Dicer expression and purification.** (a) Domain architecture of Dicer. Domains were positioned using InterPro [55] and refined using available structural data [30, 56-58]. (b) Flowchart of the Dicer expression and purification protocol. (Day 1) Cells are passaged at  $0.8 \times 10^6$  cells/mL and incubated 24 h before transfection. (Day 2) The transfection mix is prepared in culture media using 10% of the final transfection volume with 10  $\mu$ g/mL of pTT5-Dicer plasmid and 20  $\mu$ g/mL of linear PEI. After vigorous vortexing and a 15-min incubation at room temperature, the transfection mix is added to the cell culture, which is incubated for

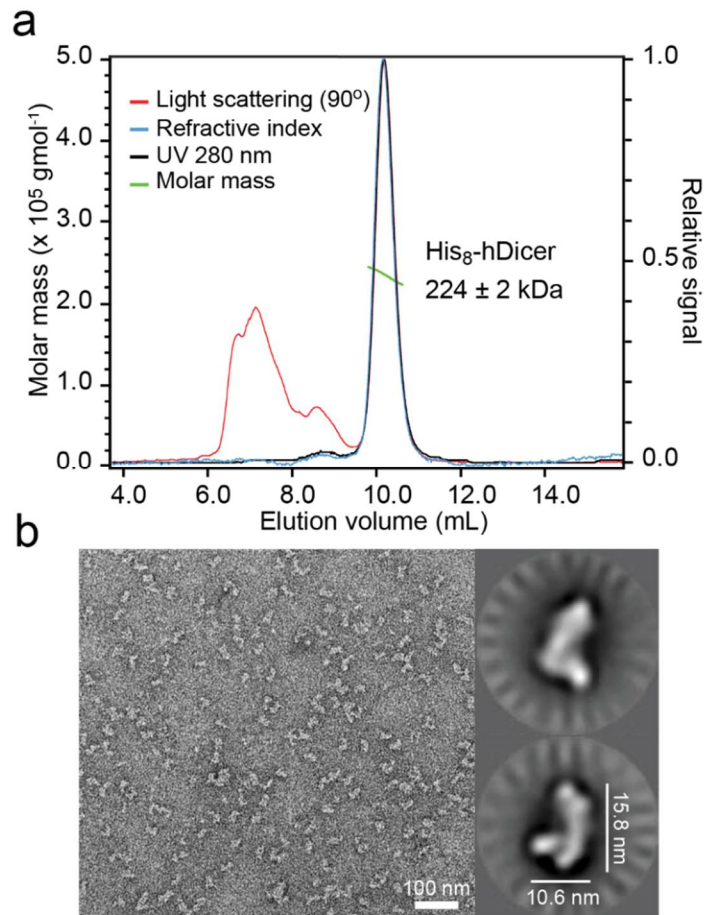
72 h (37°C, 5% CO<sub>2</sub>) with shaking (100 RPM). (Day 5) Cells are harvested by centrifugation at 200 × g and rinsed 3 times in cold PBS. After lysis, the cytoplasmic fraction is clarified by centrifugation, filtered and loaded on a 60-mL Q Sepharose Fast Flow column for ion-exchange chromatography. Fractions containing Dicer are pooled and loaded directly on a 5-mL HisTrap HP column for purification by immobilized metal affinity chromatography (IMAC). The Dicer-containing fractions are loaded directly on a 120-mL Superdex 200 column for purification by size-exclusion, and the fractions containing homogeneous Dicer are concentrated and stored at -80°C. (c-e) Typical chromatograms from the (c) ion-exchange, (d) affinity and (e) size-exclusion purifications, showing the UV absorbance at 280 nM and 260 nM along with the gradient trace. The selected fractions are highlighted by the grey area. (f-g) SDS-PAGE summary of the purification viewed by (f) Coomassie stain and (g) Western blot. Each lane is loaded proportionally to reflect the yield at each step (*Lane 1*: clarified lysate; *Lane 2*: ion-exchange fraction pool; *Lane 3*: affinity fraction pool; *Lane 4*: size-exclusion fraction pool; and *Lane 5*: concentrated protein). Yields were quantified from Western blot analysis, with loaded quantities of Dicer being in the linear range of detection.

### **2.4.3. Conformational characterization of the purified Dicer**

Size-exclusion chromatography coupled to multi-angle light-scattering and refractive index (SEC-MALS/RI) is a powerful tool to characterize purified biomolecules in solution; it allows accurate molecular weight determination as well as the determination of the oligomeric state in solution. In the SEC-MALS/RI experiments performed here, SEC served only as a separation step, whereas the LS and RI detectors were used to calculate the molar mass of Dicer after purification and concentration in a sucrose/DDM-free buffer (Figure 2.3a). Although a small

percentage of multimeric protein appears to be present in the void volume of the SEC (elution volume  $\sim 8$  mL), the purified Dicer elutes predominantly as a single peak that corresponds to  $\geq 98\%$  of the total protein in the injected sample. The derived molecular weight for this peak based on light-scattering data is  $224 \pm 2$  kDa, which essentially matches the theoretical molecular weight of the His-tagged monomeric protein (221 kDa). The polydispersity index of  $1.02 \pm 0.01$  over the entire eluted peak indicates that the main form of purified Dicer is monodisperse, i.e. homogeneous with respect to molar mass. Moreover, SEC-MALS/RI analysis of a purified Dicer sample stored for 6 months at  $-80^\circ\text{C}$  in sucrose/DDM-containing buffer shows that the purified protein remains almost exclusively monomeric ( $\geq 94\%$ ), indicating that the protein is intact after long-term storage (Figure 2.6).

Negative stain transmission electron microscopy (TEM) is an efficient tool to evaluate the quality of a protein sample, providing information on the homogeneity and the three-dimensional shape of the biomolecule. TEM micrographs of negative stain preparation of the purified Dicer using uranyl formate illustrate particles of roughly equal size throughout the field of view, providing additional evidence of the homogeneity of the purified sample (Figure 2.3b). Model-free 2D class averages allowed us to observe the characteristic L-shape architecture of Dicer, with two main unique subpopulations representing different forms of the enzyme, in agreement with previous studies [27, 30, 36]. Taken together, the SEC-MALS/RI and negative stain TEM studies validate the monomeric L-shape architecture of the purified Dicer and indicate that the conformational properties of the enzyme are preserved after long-term storage.



**Figure 2.3. Conformational characterization of purified Dicer.** (a) SEC-MALS/RI analysis of purified WT Dicer stored in sucrose/DDM-free storage buffer. The relative signal of light scattering (red), refractive index (blue), and UV absorbance at 280 nm (black) are represented by solid line. The molar mass distribution is shown with green dots. A small amount of aggregated protein can be seen mainly from the light scattering signal. Peak integration of the UV absorbance trace shows that the main peak contains  $\geq 98\%$  of the eluted protein. The molar mass calculated from MALS analysis was normalized against BSA to give an average molecular weight of  $212 \pm 2$  kDa with a polydispersity index of  $1.02 \pm 0.01$  over the entire eluted peak. (b) Negative stain TEM. *Right panel* Typical grid imaging of negatively-stained WT Dicer. *Left panels* Main 2D class averages of WT Dicer.

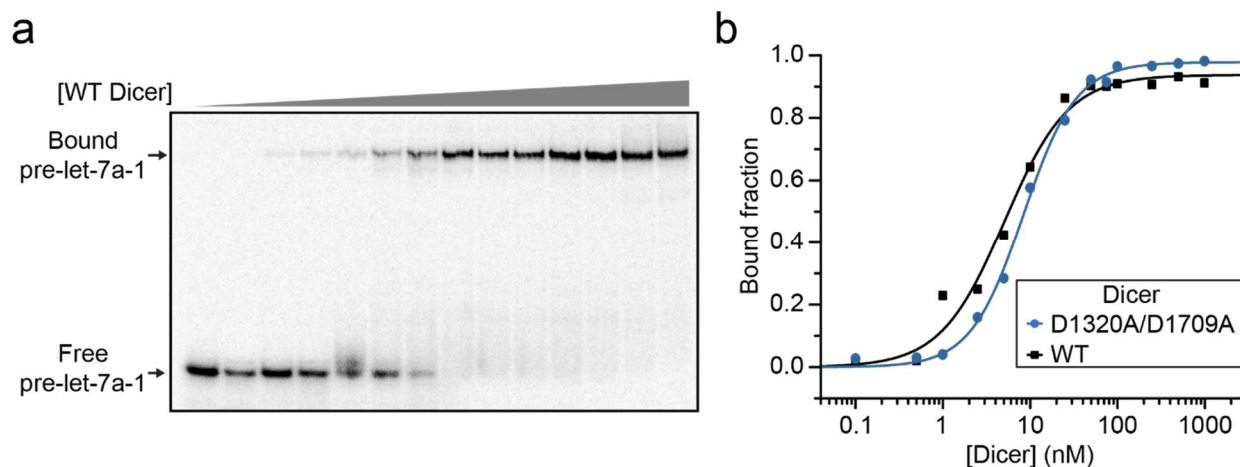


#### 2.4.4. Binding and cleavage of pre-let-7a-1 by Dicer

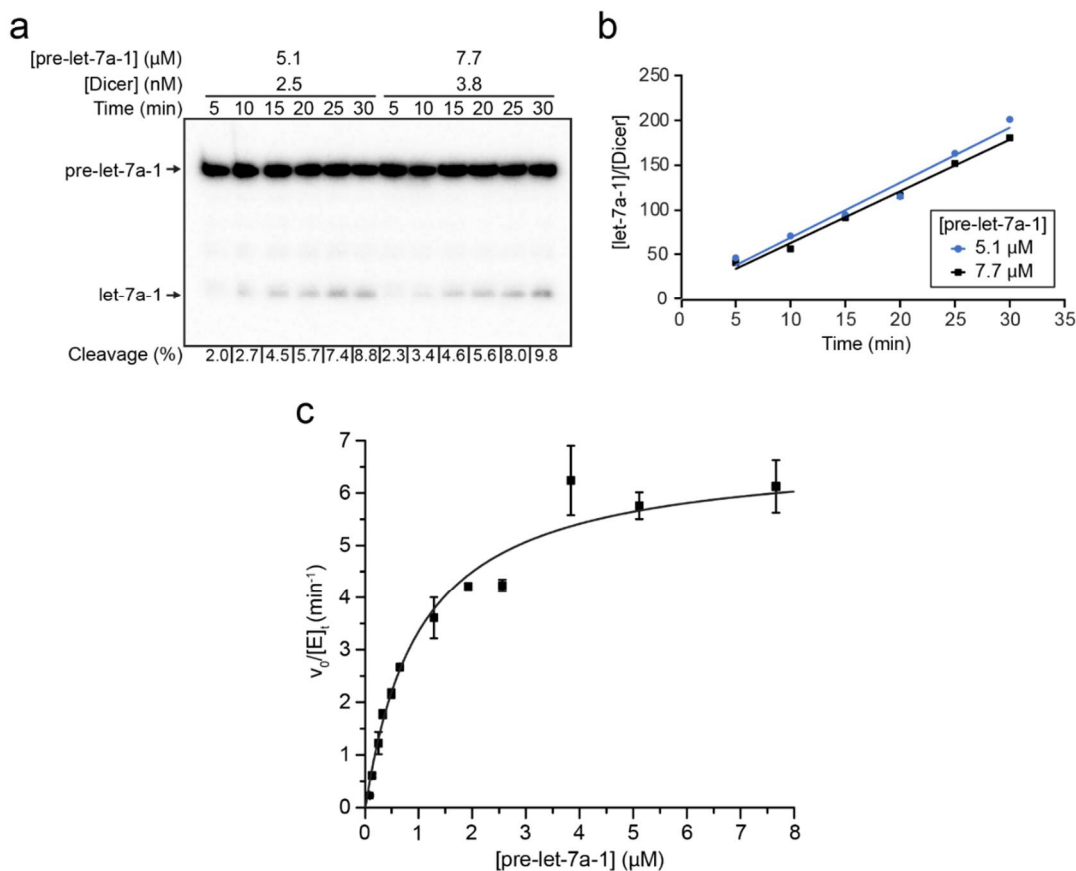
Dicer is known to bind pre-miRNA substrates and then cleave at two specific phosphodiester bonds to yield miRNA-5p/3p species. To evaluate the RNA binding and cleavage activity of the purified recombinant Dicer, we performed binding and kinetic studies using the pre-let-7a-1 substrate, one of the most commonly used pre-miRNA substrate for Dicer characterization. Binding studies were conducted via electrophoretic mobility shift assay (EMSA) using  $^{32}\text{P}$ -labeled pre-let-7a-1 and increasing concentration of Dicer, either wild-type (WT) Dicer supplemented with 0.5 mM EDTA to prevent cleavage or with the catalytically-inactive D1320/1709A variant (Figure 2.4). When fitting the binding data to the Hill equation, dissociation constants ( $K_d$ ) of  $5 \pm 1$  nM (WT Dicer) and  $9 \pm 1$  nM (D1320A/D1709A Dicer) were obtained with Hill coefficient of  $1.3 \pm 0.2$  and  $1.4 \pm 0.3$ , respectively.

To characterize the cleavage activity of the purified Dicer enzyme, we determined the steady-state kinetic parameters  $k_{cat}$  and  $K_M$  under multiple turnover conditions. To ensure that the initial cleavage rates were measured under steady state conditions, the enzyme concentration ( $[\text{E}]$ ; 0.35 nM to 3.75 nM) was varied along with the initial substrate concentration ( $[\text{S}]$ ; 0.08  $\mu\text{M}$  to 7.68  $\mu\text{M}$ ) to maintain  $[\text{S}]/[\text{E}] \geq 200$ , and time points were collected after at least one turnover of the enzyme pool at up to 10% substrate cleavage, as previously reported for steady-state kinetic studies of *E. coli* RNase III [37]. For each substrate concentration, the cleavage reaction of  $^{32}\text{P}$ -labeled pre-let-7a-1 was monitored by denaturing gel electrophoresis to quantify the percentage of cleavage (Figure 2.5a). From this data, the concentration of product normalized against enzyme concentration ( $[\text{let-7a-1}]/[\text{Dicer}]$ ) was plotted as a function of time, and the slope of the resulting time course was fitted by linear regression to derive the turnover frequency, expressed as  $v_o/[\text{E}]_t$  (Figure 2.5b). By plotting the dependence of  $v_o/[\text{E}]_t$  versus  $[\text{S}]$

and fitting to the Michaelis-Menten equation, a  $k_{cat}$  value of  $7.2 \pm 0.5 \text{ min}^{-1}$  and a  $K_M$  value of  $1.2 \pm 0.3 \text{ }\mu\text{M}$  were determined for cleavage of pre-let-7a-1 by Dicer (Figure 2.5c).



**Figure 2.4. Binding studies of Dicer to pre-let-7a-1.** (a) Typical EMSA performed with 10 pM of 5'-[ $^{32}\text{P}$ ]-labeled pre-let-7a-1 and increasing concentrations of the D1320A/D1709A variant Dicer (0, 0.1, 0.5, 1, 2.5, 5, 10, 25, 50, 75, 100, 250, 500 and 1000 nM). (b) Typical binding curves of D1320A/D1709A Dicer and WT Dicer to pre-let-7a-1. The data shown here were fitted to the Hill equation to obtain  $K_d$  values of  $5.1 \pm 0.6 \text{ nM}$  (WT Dicer) and  $8.4 \pm 0.4 \text{ nM}$  (D1320A/D1709A Dicer) with Hill coefficient of  $1.2 \pm 0.1$  and  $1.4 \pm 0.1$ , respectively. Averages from at least three independent experiments yield  $K_d$  values of  $5 \pm 1 \text{ nM}$  (WT Dicer) and  $9 \pm 1 \text{ nM}$  (D1320A/D1709A Dicer) with Hill coefficient of  $1.3 \pm 0.2$  and  $1.4 \pm 0.3$ , respectively.



**Figure 2.5. Steady-state kinetics for cleavage of pre-let-7a-1 by Dicer.** (a) Typical Dicer cleavage reaction of 5'-[ $^{32}\text{P}$ ]-labeled pre-let-7a-1 with Dicer analyzed by denaturing gel electrophoresis. Reactions were carried for 30 min in conditions that allowed less than 10% of total substrate cleavage. (b) Normalized product formation ( $[\text{let-7a-1}]/[\text{Dicer}]$ ) as a function of time for cleavage data shown in (a). The turnover frequency, expressed as  $v_o/[E]_i$ , was calculated by linear regression of the slope ( $5.8 \pm 0.3 \text{ min}^{-1}$  and  $6.1 \pm 0.5 \text{ min}^{-1}$  for pre-let-7a-1 substrate concentration of 5.1  $\mu\text{M}$  and 7.7  $\mu\text{M}$ , respectively). (c) Typical steady-state kinetic analysis showing the dependence of the turnover frequency on substrate concentration (0.08 to 7.68  $\mu\text{M}$ ). The data shown here were fitted to the Michaelis-Menten equation to yield  $k_{cat}$  of  $6.8 \pm 0.8 \text{ min}^{-1}$  and  $K_M$  of  $1.0 \pm 0.2 \mu\text{M}$ . Averages from two independent experiments yield  $k_{cat}$  of  $7.2 \pm 0.5 \text{ min}^{-1}$  and  $K_M$  of  $1.2 \pm 0.3 \mu\text{M}$ .

## 2.5. Discussion

By optimizing transfection conditions, we have achieved high-level expression of recombinant human Dicer using a 293-6E cell expression system. In addition, a fast and efficient purification procedure was developed that yields up to 9 mg of pure, monomeric and stable recombinant human Dicer per liter of culture. It may be possible to obtain higher yields by further optimizing transfection conditions as it should be possible to increase the transfection efficiency (30-40%) by about two-fold [31, 38, 39]. The reported purification protocol was designed to prevent protein aggregation and precipitation, since we observed that Dicer tends to form multimers and precipitate at high-concentration or when it contacts metallic surfaces. Nevertheless, the purified protein was shown to be monomerically pure and highly active.

Although affinity purification is often considered the best choice for the first purification step of a recombinant protein from a crude mixture, we selected an alternate approach to minimize sample handling that could lead to aggregation and precipitation, such as protein dialysis, desalting, and concentration. Following cell lysis in low-salt buffer, the cytoplasmic fraction was first loaded on an anion-exchange column. The large volume and high salt concentration of the elution fractions from this first purification step is compatible with direct loading onto the HisTrap affinity column. Subsequently, the small elution volume from the second purification step allowed for a single injection on the preparative size-exclusion column without any additional manipulations. Overall, this strategy allows for multi-step purification of milligram amounts of recombinant Dicer in a fast and efficient manner. In fact, the entire purification procedure can be conveniently completed in a single work day.

It is important to note that the last step of purification can be performed either in the presence or absence of reagents (sucrose and DDM) that were found to prevent aggregation and

precipitation at high protein concentration. These reagents are not required if the concentration of the protein eluted from the size-exclusion column, generally between 0.5-1  $\mu\text{M}$ , is sufficient for subsequent applications. However, if higher protein concentrations are required and/or long-term storage is desirable, these reagents should be included in the buffer used for size-exclusion chromatography and final protein concentration.

In binding studies, we determined  $K_d$  values of  $5 \pm 1$  nM (WT Dicer) and  $9 \pm 1$  nM (D1320A/D1709A Dicer) between Dicer and pre-let-7a-1. These values are similar to previously reported  $K_d$  values of 1.8-39 nM obtained for the binding of WT Dicer to pre-let-7a-1 under similar conditions [18, 20, 40, 41]. Moreover, our studies confirm observations with other RNaseIII enzymes that the amino acid substitutions that disrupt magnesium binding in the RNaseIIIa and RNaseIIIb domains do not affect substrate binding to Dicer [42].

From kinetic studies of pre-let-7a-1 cleavage by Dicer, we obtained a  $k_{cat}$  value of  $7.2 \pm 0.5$   $\text{min}^{-1}$  and a  $K_M$  value of  $1.2 \pm 0.3$   $\mu\text{M}$ . These steady-state kinetic parameters are much higher than those derived from previously cleavage studies of the same pre-miRNA substrate (25-375 nM) by Dicer (5 nM):  $k_{cat} = 0.45$   $\text{min}^{-1}$  and  $K_M = 20$ -30 nM [20]. Notably, we found similar values ( $k_{cat} = 0.72$   $\text{min}^{-1}$  and  $K_M = 52$  nM  $\pm$  7 nM) when performing kinetic experiments under similar conditions (5 nM Dicer and 2.5-1000 nM substrate; Figure 2.7). However, the resulting curve did not fit a hyperbolic model, which is understandable given that the following condition,  $[\text{S}] \gg [\text{E}]$ , was not respected for every substrate concentration and, thus, did not agree with the steady-state assumption [43]. By selecting experimental conditions that ensured the validity of the steady-state assumption ( $[\text{S}]/[\text{E}] \geq 200$ ), much higher  $k_{cat}$  and  $K_M$  values were obtained.

These updated steady-state kinetic values provide insights into the kinetic mechanism of pre-miRNA cleavage by Dicer. Given the  $k_{cat}/K_M$  of  $\sim 10^5 \text{ M}^{-1} \text{ sec}^{-1}$  and the very low  $k_{cat}$  value of  $\sim 0.1 \text{ sec}^{-1}$ , the rate-limiting step may be product release as observed with other RNase III enzymes [37, 44] or possibly the slow formation and/or conversion of the enzyme-substrate complex into a catalytically productive conformation. Dicer is a complex and dynamic enzyme that is known to interact with several cellular proteins [45], and there is increasing evidence that conformational changes may be important for its cleavage activity. Free Dicer is known to adopt different conformations, as confirmed here from negative stain TEM studies, and binding of pre-miRNA or double-stranded RNA substrates is linked to structural rearrangements [18, 23, 27, 36]. Furthermore, single-molecule fluorescence studies have provided evidence for Dicer adopting two different binding modes with a pre-let-7a-1 substrate [46]. More recently, cryo-EM reconstructions of a pre-miRNA/Dicer/TRBP complex captured two distinct conformations of a pre-dicing state, from which transition to the dicing state would require repositioning of the pre-miRNA in the enzyme processing center [30]. Thus, additional studies are needed to better define the detailed kinetic mechanism of Dicer and characterize the role that conformational changes play along the cleavage reaction pathway.

## 2.6. Conclusion

We have achieved large-scale expression of recombinant human Dicer in a human cell line. The expression in 293-6E cells coupled with our streamlined purification protocol yields up to 9 mg of protein per L of cell culture, which is higher than with previously reported protocols. The purification scheme was carefully optimized to minimize aggregation and maximize monomeric purity and stability of the final sample. To achieve these objectives, we minimized sample manipulation which, in turn, reduced the duration of the purification to a single work day. The

purified enzyme adopts the typical L-shape architecture and binds a pre-let-7a-1 substrate in the low nM range, as previously observed. Steady-state kinetic studies yielded higher  $k_{cat}$  and  $K_M$  values than previously reported due to the use of experimental conditions that respected the steady-state assumption. Together, these studies provide an efficient method for high-yield production of human Dicer as well as revised kinetic parameters for Dicer cleavage that should prove useful for future structural and mechanistic studies of Dicer aimed at better understanding how the catalytic production of miRNAs is achieved and regulated.

## 2.7. Methods

### 2.7.1. Plasmids

For construction of the human Dicer expression vector pTT5-Dicer, a PCR fragment encoding Dicer was first generated from plasmid pFRT/TO/FLAG/HA-DEST DICER [provided as a gift from Thomas Tuschl (Addgene plasmid # 19881)][47] and inserted between the NotI and HindIII sites of the pTT5SH8Q1 vector to allow for expression of Dicer with a C-terminal StrepTagII/His<sub>8</sub> tag [34]. The expression vector for the catalytically-inactive Dicer D1320A/D1709A variant was prepared from pTT5-Dicer using the Stratagene QuikChangeII site-directed mutagenesis method. Plasmids were amplified in *Escherichia coli* (*E. coli*) DH5 $\alpha$  cells (Invitrogen) grown in LB medium supplemented with 100 mg/L ampicillin and purified using the QIAGEN Plasmid Giga Kit. Plasmids were stored at 4°C in water at a concentration of 1 mg/mL.

The pUC19-HH-pre-let-7a-1-HDV vector used for *in vitro* transcription of pre-let-7a-1 was constructed using a pUC19 vector, which allows for synthesis of pre-let-7a-1 flanked by a 5' Hammerhead ribozyme (HH) and a 3' Hepatitis delta virus (HDV) ribozyme. Plasmids were

amplified in *E. coli* DH5 $\alpha$  cells grown in LB medium supplemented with 100 mg/L ampicillin and purified using the QIAGEN Plasmid Maxi Kit. Plasmids were linearized with HindIII (New England Biolabs) and stored as is at -20°C at a concentration of 1 mg/mL. All plasmids were verified by DNA sequencing.

### **2.7.2. Cell culture and transfection**

HEK-293-EBNA1-6E (293-6E) cells were grown in suspension in Freestyle 17 (F17) medium (Invitrogen) supplemented with 4 mM glutamine, 0.1% pluronic acid and 25  $\mu$ g/mL G418 at 37°C with 5% CO<sub>2</sub>. Tryptone N1 was added to some of the cell cultures at a weight/volume ratio of 0.5%, but was found to have minimal effect on cultures grown 72 hpt. Cultures were maintained under 2.0 x 10<sup>6</sup> cells/mL in Erlenmeyer flasks shaken at 100 RPM in a humidified incubator at 37°C with 5% CO<sub>2</sub>.

Cell cultures were diluted to 0.8 x 10<sup>6</sup> cells/mL 24 h before transfection. The transfection mix was prepared in 10% of the final culture volume in F17 medium with 10  $\mu$ g/mL of DNA plasmid and 20  $\mu$ g/mL of linear PEI 25 kDa (Polysciences Inc.). The mix was agitated by 3 vortexing cycles of 3 sec, incubated 15 min at room temperature and incorporated in the culture [32], which was incubated for 72-96 h (37°C, 5% CO<sub>2</sub>) with shaking (100 RPM).

For optimization of transfection conditions, small-scale cultures were used: 2-mL cultures were grown in 6-well plates (Starstedt #83.1839.500) and 20-mL cultures were grown in 125-mL disposable Erlenmeyer flasks (Corning #431143). For Dicer purification, large-scale cultures of 150 mL to 1 L were grown in thoroughly rinsed and sterilized 1-L glass Erlenmeyer (Corning #49851L) or 2.8-L Fernbach culture flasks (Corning #44262XL). The incubation time after transfection was fixed to 72 h for large-scale cultures.



At harvest time, cells were first counted and then pelleted by centrifugation at 200×g and washed 3 times with at least 5-pellet volumes of cold PBS. The washed cell pellet was either used immediately or frozen in liquid nitrogen and stored at -80°C for later use.

### **2.7.3. Optimization of transfection conditions**

To optimize transfection conditions, small-scale transfections were monitored over a 4-day period by withdrawing a 190-μL aliquot from the culture every 24 h to measure cell density, cell viability and recombinant Dicer expression. From this aliquot, a 10-μL volume was mixed with an equal volume of trypan blue and loaded on a hemocytometer slide to determine the cell density and viability using a Nikon Eclipse TS100 microscope. The remaining 180 μL of cell culture was pelleted, resuspended in 60 μL of lysis buffer (20 mM Tris pH 7.4, 50 mM NaCl, 10% Glycerol, 1 mM EDTA, 1 mM EGTA, 0.1% NP-40 and 0.5 mM TCEP [tris(2-carboxyethyl)phosphine]) and spun for 1 min at 16,000 ×g. The supernatant containing the cytoplasmic fraction was mixed with one volume of 2X Laemmli buffer and stored at -20°C. All samples collected during transfection were heated at 95°C for 2 min, separated on a 7.5% SDS-PAGE and blotted on a nitrocellulose membrane (Bio-Rad) for Western-blot analysis. Dicer was detected using a mouse anti-His tag primary antibody (Medimabs #MM-0165-P), an HRP-conjugated anti-mouse secondary antibody (Cedarlane #115-035-044) and the Western Lightning Plus-ECL substrate (Perkin-Elmer #NEL103E001EA). A ChemiDoc MP (Bio-Rad) Imaging system was used for chemiluminescence detection, and band intensities were quantified using the ImageLab software version 5.2 (Bio-Rad).

Transfection efficiency was monitored by substituting 5% of the total plasmid transfected by a pTT5SH8Q1-GFPq plasmid [48]. After 48 h, live cells were diluted with one volume of Freestyle 17 media supplemented with 20 mM HEPES pH 7.5, and a thin layer of diluted cells

was carefully deposited on a Lab-TEK-II chamber slide. The slide was visualized by both phase contrast microscopy and by fluorescence microscopy on a Nikon TE2000U microscope (20X magnification). The ratio of cells expressing GFP over the total cells of at least 200 cells was taken as the transfection efficiency.

#### **2.7.4. Dicer purification**

Dicer-transfected cell pellets were resuspended in 5-pellet volumes of lysis buffer supplemented with one tablet of Roche Complete EDTA-free and incubated on ice for 10 min. The lysate was clarified by centrifugation at 45 000  $\times g$  for 20 min at 4°C. The supernatant was then filtered on a 0.22- $\mu m$  pore size polyethersulfone (PES) membrane and loaded onto a 60-mL Q Sepharose Fast Flow media packed in a XK 26/20 column (GE Healthcare) equilibrated with Qseph-A buffer (20 mM Tris pH 7.4, 10% glycerol and 0.5 mM TCEP). After loading, the column was washed with 2.5 column volume (CV) of 20% Qseph-B buffer (20 mM Tris pH 7.4, 1 M NaCl, 10% glycerol and 0.5 mM TCEP) and Dicer was eluted using a linear gradient (from 20% to 70% over 4 CV) of Qseph-B. The Dicer-containing fractions were combined and loaded on a 5-mL HisTrap High Performance column (GE Healthcare) equilibrated with His-A buffer (20 mM Tris pH 7.4, 500 mM NaCl, 25 mM imidazole and 0.5 mM TCEP). After loading, the column was washed with 12 CV of His-A buffer and Dicer was eluted with a single isocratic step of 100% His-B buffer (20 mM Tris pH 7.4, 500 mM NaCl, 250 mM imidazole and 0.5 mM TCEP). The fractions containing Dicer were combined and loaded on a preparative Superdex 200 16/600 column (GE Healthcare) pre-equilibrated in storage buffer, either sucrose/DDM-free storage buffer (50 mM Tris pH 8.2, 10 mM NaCl/KCl 24:1, 0.5 mM MgCl<sub>2</sub>, and 0.5 mM TCEP) or storage buffer supplemented with 5% Sucrose and 0.3 mM DDM (*n*-dodecyl- $\beta$ -D-maltoside). When needed, fractions containing monomeric Dicer were pooled and concentrated on a 7-mL

Apollo concentrator 150-kDa MWCO (Orbital Biosciences). The concentrated Dicer was aliquoted, flash frozen in liquid nitrogen and stored at -80°C for future use. Purity and yield were assessed by 7.5% Coomassie-stained SDS-PAGE and Western blotting of the different fractions collected during purification.

### **2.7.5. Synthesis and purification of pre-let-7a-1**

The *in vitro* transcription of human pre-let-7a-1 with a 5'-HH ribozyme tag and a 3'-HDV ribozyme tag was carried out in 40 mM HEPES pH 8.0, 20 mM MgCl<sub>2</sub>, 50 mM DTT, 1 mM spermidine, 0.1% triton X-100, 4 mM of each NTP, 30 µg/mL of linearized pUC19-HH-pre-let-7a-1-HDV plasmid, 0.5 U/mL of RNasin® Plus RNase inhibitor and 30 µg/mL of His-tagged T7 polymerase prepared in house [49]. Since both ribozyme tags were cleaved off during transcription, the pre-let-7a-1 RNA was generated with homogeneous ends directly from the transcription reaction. The RNA was then purified by denaturing gel electrophoresis, as described previously [50]. To convert the purified RNA into an optimal Dicer substrate with 5'-phosphate and 3'-OH ends [51], a phosphorylation reaction was carried out with T4 polynucleotide kinase (NEB #M0201S) according to the manufacturer protocol with either cold ATP or ATP-γ-<sup>32</sup>P. The <sup>32</sup>P-labeled pre-let-7a-1 was further purified by denaturing gel electrophoresis, extracted from the gel by crush and soak, ethanol precipitated and resuspended in TE pH 7.5 (10 mM Tris pH 7.5 and 1 mM EDTA) [50]. The cold phosphorylated RNA was further purified by HPLC at 65°C on a DNAPac PA-100 9x250 mm equilibrated with 30% DNAPac-B buffer (12.5 mM Tris pH 7.4, 5 M urea and 0.5 M NaClO<sub>4</sub>) and 70% DNAPac-A buffer (12.5 mM Tris pH 7.4 and 5 M urea). The sample was loaded and eluted with a linear gradient (30% to 70%) of DNAPac-B buffer. The fraction containing pure pre-let-7a-1 was exchanged into TE pH 7.5 and concentrated using a 3-kDa MWCO Amicon® Ultra-4

Centrifugal Filter Unit (Millipore). Both the  $^{32}\text{P}$ -labeled and cold pre-let-7a-1 samples were stored at  $-20^{\circ}\text{C}$ .

### **2.7.6. SEC-MALS/RI**

Human WT Dicer final purification product was analyzed by SEC-MALS/RI using an AKTA micro FPLC (GE Healthcare) coupled to a multiple-angle light scattering and refractive index system (MALS/RI; Wyatt Dawn HELEOS II and OptiLab T-rEX). A volume of 100  $\mu\text{L}$  (75-80  $\mu\text{g}$ ) of purified Dicer was loaded on a Superdex 200 Increase 10/300 (GE Healthcare) column equilibrated in sucrose/DDM-supplemented storage buffer at a flowrate of 0.3-0.5 mL/min. The SEC-MALS/RI data was collected and analysed with ASTRA chromatography software package, version 6.1.6.5 (Wyatt Technology). The MALS detector was normalized using the monomeric and dimeric signals of bovine serum albumin (BSA; Fisher scientific BP1605). For the Dicer sample purified in DDM-free buffer, the molar mass was determined by the dual detection method (RI/LS signals) implemented in the conjugated analysis mode of the ASTRA software, using a refractive index increment ( $dn/dc$ ) of 0.190 mL/g [52]. A similar method was used for the Dicer sample purified in DDM-containing buffer, however the dual detection method used the UV and LS signals, because the RI signal may be perturbed by DDM-Dicer interactions.

### **2.7.7. Negative stain TEM**

A 3- $\mu\text{L}$  drop of 50 nM human WT Dicer in 20 mM HEPES, pH 7.5, 150 mM KCl, 3 mM EDTA, 1 mM DTT, and 2.5% glycerol was applied onto a carbon-coated copper grid previously glow-discharged using an ELMO glow discharge system (Cordouan Technologies, France). After 1 min, excess liquid was blotted and stained for 1 min with 1.5% freshly prepared uranyl formate

(Electron Microscopy Sciences, PA). Samples were imaged using a FEI Tecnai T12 (Eindhoven, The Netherlands) Transmission Electron Microscope (TEM) equipped with a LaB6 filament and operated at an acceleration voltage of 120 kV. Micrographs were collected at defocus between 1-3  $\mu\text{m}$  on a FEI Eagle 4k x 4k CCD camera at a magnification of  $\sim 67,000\times$ . From these micrographs, 125 000 particles were extracted and 2D align using xmipp from Scipion [53]. The particles were classified in a minimum of 2 classes and the 2D volumes were measured using ImageJ [54].

### **2.7.8. Dicer binding assay**

Binding assays were carried using either WT Dicer or the catalytically-inactive Dicer variant (D1320A/D1709A). Dicer D1320A/D1709A was first diluted in EMSA buffer (100 mM Tris pH 7.6, 100 mM NaCl and 20% glycerol), whereas WT Dicer was diluted in EMSA buffer supplemented with 1 mM EDTA. In parallel,  $^{32}\text{P}$ -labeled pre-let-7a-1 was refolded by heating at  $95^\circ\text{C}$  for 2 min and snap-cooling at  $4^\circ\text{C}$  for a minimum of 5 min. The RNA was then diluted in 0.1% NP-40 and 4 mM DTT to a concentration of 4 pM. The binding reactions were initiated by adding 10  $\mu\text{L}$  of the RNA sample to 10  $\mu\text{L}$  of the protein samples. Standard binding reactions contained 50 mM Tris pH 7.6, 50 mM NaCl, 10% glycerol, 0.05% NP-40, 2 mM DTT, and 2 pM  $^{32}\text{P}$ -labeled RNA (0.5 mM EDTA was added for reactions with WT Dicer) with protein concentrations varying between 0.01X and 100X of the estimated  $K_d$  value. Reactions were incubated for 30 min at  $4^\circ\text{C}$  and loaded on a 4%-15% gradient polyacrylamide gel (37.5:1 acrylamide:bis-acrylamide) in Tris-Glycine buffer (25 mM Tris-Base and 200 mM glycine), which was run for 2 h at 200 V in the cold room ( $4^\circ\text{C}$ ). Gels were dried, exposed overnight to a storage phosphor screen (Bio-Rad) and visualized with a Personal Molecular Imager (Bio-Rad). Band intensities for the bound (B) and unbound (U) RNA were quantified using the ImageLab

software version 5.2 (Bio-Rad). The fraction of bound RNA [ $F = B/(B+U)$ ] was plotted against protein concentration, and the binding data were fitted to the Hill equation with the OriginPro 8 software (OriginLab). The reported  $K_d$  value and its errors represent the average and standard deviation from two independent experiments.

### **2.7.9. Dicer cleavage assay**

For *in vitro* Dicer cleavage assays, the RNA and Dicer solutions were prepared separately at twice their final desired concentration in 30  $\mu$ L of the cleavage reaction buffer (50 mM HEPES pH 7.4, 50 mM NaCl, 5 mM MgCl<sub>2</sub> and 0.05% NP-40). The RNA solutions, containing 80 pM of <sup>32</sup>P-labeled pre-let-7a-1 and cold RNA at concentrations ranging from 0.1X to 10X the estimated  $K_M$ , were heated at 95°C for 2 min and snap-cooled at 4°C for a minimum of 5 min to refold the RNA. The concentration of the Dicer solution was adjusted for each pre-let-7a-1 substrate concentration to allow measurement of the initial velocity of the reaction (5-10% substrate cleavage in 30 min) under multiple turnover conditions (substrate/Dicer molar ratio  $\geq$  200). Both RNA and protein solutions were pre-heated for 5 min at 37°C, and cleavage reactions were initiated by adding 30  $\mu$ L of the protein solution to 30  $\mu$ L of the RNA solution. At specific time points, a 5- $\mu$ L aliquot of the cleavage reaction was taken, immediately mixed with 5  $\mu$ L of stop buffer (95 % formamide and 50 mM EDTA). For each reaction, 6 time points were taken at less than 10% reaction completion and over at least 2 full turnovers of the enzyme pool to ascertain the possible use of the steady-state assumption. Samples were analyzed by denaturing gel electrophoresis [20% acrylamide:bis-acrylamide (19:1) /7 M urea gel]. The amounts of substrate (S) and product (P) were quantified from radioactive bands as for the binding assay, and the fraction of cleaved product [ $F = P / (S + P)$ ] was plotted against time. The resulting time courses were fitted by linear regression and the slope was taken as the initial

velocity ( $v_o$ ). Subsequently, the dependence of  $v_o/[E]_t$  on  $[S]$  was plotted and fitted to the Michaelis-Menten equation (written as  $v_o/[E]_t = k_{cat}[S]/(K_M+[S])$ ) with the OriginPro 8 software (OriginLab) to derive  $k_{cat}$  and  $K_M$  values. The quality of the fit was obtained from the square of the correlation coefficient ( $R^2$ ), with  $R^2 \geq 0.97$  and by making sure that the distribution of residuals was random around the regression line. The values of  $k_{cat}$  and  $K_M$  are reported with experimental errors taken from the average and standard deviations from two independent experiments.

### **List of abbreviations**

BSA: bovine serum albumin; CCD: charge-coupled device; CMV: cytomegalovirus; CV: column volume; DDM: *n*-dodecyl- $\beta$ -D-maltoside; DTT: dithiothreitol; EM: electron microscopy; ES: enzyme-substrate; EBNA1: Epstein-Barr nuclear antigen 1; EMSA: electrophoretic mobility shift assay; FPLC: fast protein liquid chromatography; GFP: green fluorescent protein; HDV: Hepatitis delta virus; HEK293-6E: EBNA1 human embryonic kidney cell line; HH: hammerhead; Hsp70: heat-shock protein 70; LS: light-scattering; MWCO: molecular weight cut-off; NP-40: nonyl phenoxypolyethoxyethanol; PACT: protein activator of the interferon-induced protein kinase; PAGE: polyacrylamide gel electrophoresis; PES: polyethersulfone; PBS: phosphate buffer saline; PEI: polyethylenimine; SDS: sodium dodecyl sulfate; RI: refractive index; SEC-MALS: size-exclusion chromatography and multiple-angle light scattering; RISC: RNA-induced silencing complex; RPM: revolution per minute; TCEP: tris(2-carboxyethyl)phosphine; TEM: transmission electron microscopy; TRBP: TAR-RNA binding protein; WT: wild-type.

## **Acknowledgements**

We thankfully acknowledge Alexandre Desjardins and Louise Cournoyer for technical help with cell culture and transfection, Geneviève Di Tomasso and Pierre Dagenais for technical assistance with protein and RNA purification, Aleksandr Sverzhinsky and Sébastien Truche for technical assistance with SEC-MALS/RI data collection.

## **Funding**

The research was funded by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to PL and the Canada Foundation for Innovation (CFI) for SEC-MALS/RI.

## **Availability of data and materials**

All data supporting our findings are displayed in the submitted manuscript.

## **Authors' contributions**

JB developed, performed and analyzed most of the experiments: the cloning, cell culture and transfection conditions, Dicer purification, pre-let-7a-1 synthesis and purification, Dicer binding and processing assays as well as one SEC-MALS/RI experiment (Figure 2.6). DNK performed Dicer purification and SEC-MALS/RI data collection and analysis (Figure 2.3a). AF performed all negative stain TEM experiments and analysis (Figure 2.3b). SA provided technical assistance to JB for cell culture, Dicer purification and Dicer binding assays (Figure 2.2c-g and Figure 2.4a). AN, YD, PL and JGO provided guidance. JB and PL wrote the manuscript. All authors read and approved the final manuscript.

## **Ethics approval and consent to participate**

Not applicable.



## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## 2.8. References

1. Song MS, Rossi JJ. Molecular mechanisms of Dicer: endonuclease and enzymatic activity. *Biochem J.* 2017, 474(10):1603-1618.
2. Foulkes WD, Priest JR, Duchaine TF. DICER1: mutations, microRNAs and mechanisms. *Nat Rev Cancer.* 2014, 14(10):662-672.
3. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 2014, 42(Database issue):D68-73.
4. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 2009, 19(1):92-105.
5. Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature.* 2010, 465(7298):584-589.
6. Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon GJ, Lawson ND *et al.* A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science.* 2010, 328(5986):1694-1698.
7. Havens MA, Reich AA, Duelli DM, Hastings ML. Biogenesis of mammalian microRNAs by a non-canonical processing pathway. *Nucleic Acids Res.* 2012, 40(10):4626-4640.
8. Ha M, Kim VN. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol.* 2014, 15(8):509-524.
9. Li Z, Ender C, Meister G, Moore PS, Chang Y, John B. Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. *Nucleic Acids Res.* 2012, 40(14):6787-6799.
10. Di Tomasso G, Miller Jenkins LM, Legault P. ARiBo pull-down for riboproteomic studies based on label-free quantitative mass spectrometry. *RNA.* 2016.

11. Treiber T, Treiber N, Plessmann U, Harlander S, Daiss JL, Eichner N, Lehmann G, Schall K, Urlaub H, Meister G. A compendium of RNA-binding proteins that regulate microRNA biogenesis. *Mol Cell*. 2017, 66(2):270-284 e213.
12. Nussbacher JK, Yeo GW. Systematic discovery of RNA binding proteins that regulate microRNA levels. *Mol Cell*. 2018, 69(6):1005-1016 e1007.
13. Geisse S, Gram H, Kleuser B, Kocher HP. Eukaryotic expression systems: a comparison. *Protein Expr Purif*. 1996, 8(3):271-282.
14. Provost P, Dishart D, Doucet J, Frendewey D, Samuelsson B, Radmark O. Ribonuclease activity and RNA binding of recombinant human Dicer. *EMBO J*. 2002, 21(21):5864-5874.
15. Zhang H, Kolb FA, Brondani V, Billy E, Filipowicz W. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J*. 2002, 21(21):5875-5885.
16. Myers JW, Jones JT, Meyer T, Ferrell JE, Jr. Recombinant Dicer efficiently converts large dsRNAs into siRNAs suitable for gene silencing. *Nat Biotechnol*. 2003, 21(3):324-328.
17. Vermeulen A, Behlen L, Reynolds A, Wolfson A, Marshall WS, Karpilow J, Khvorova A. The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA*. 2005, 11(5):674-682.
18. Ma E, MacRae IJ, Kirsch JF, Doudna JA. Autoinhibition of human dicer by its internal helicase domain. *J Mol Biol*. 2008, 380(1):237-243.
19. Soifer HS, Sano M, Sakurai K, Chomchan P, Saetrom P, Sherman MA, Collingwood MA, Behlke MA, Rossi JJ. A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. *Nucleic Acids Res*. 2008, 36(20):6511-6522.
20. Chakravarthy S, Sternberg SH, Kellenberger CA, Doudna JA. Substrate-specific kinetics of Dicer-catalyzed RNA processing. *J Mol Biol*. 2010, 404(3):392-402.
21. Feng Y, Zhang X, Graves P, Zeng Y. A comprehensive analysis of precursor microRNA cleavage by human Dicer. *RNA*. 2012, 18(11):2083-2092.
22. Lee HY, Zhou K, Smith AM, Noland CL, Doudna JA. Differential roles of human Dicer-binding proteins TRBP and PACT in small RNA processing. *Nucleic Acids Res*. 2013, 41(13):6568-6576.

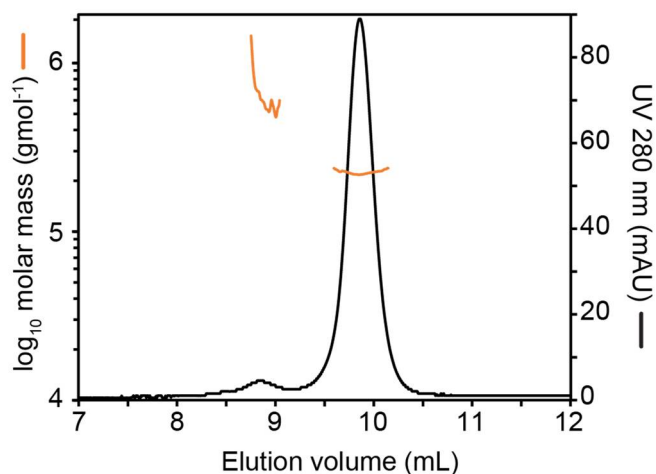
23. Liu Z, Wang J, Li G, Wang HW. Structure of precursor microRNA's terminal loop regulates human Dicer's dicing activity by switching DExH/D domain. *Protein Cell*. 2015, 6(3):185-193.
24. Kurzynska-Kokorniak A, Pokornowska M, Koralewska N, Hoffmann W, Bienkowska-Szewczyk K, Figlerowicz M. Revealing a new activity of the human Dicer DUF283 domain in vitro. *Sci Rep*. 2016, 6:23989.
25. MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA. In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci U S A*. 2008, 105(2):512-517.
26. Lau PW, Potter CS, Carragher B, MacRae IJ. Structure of the human Dicer-TRBP complex by electron microscopy. *Structure*. 2009, 17(10):1326-1332.
27. Taylor DW, Ma E, Shigematsu H, Cianfrocco MA, Noland CL, Nagayama K, Nogales E, Doudna JA, Wang HW. Substrate-specific structural rearrangements of human Dicer. *Nat Struct Mol Biol*. 2013, 20(6):662-670.
28. Wang H-W, Noland C, Siridechadilok B, Taylor DW, Ma E, Felderer K, Doudna JA, Nogales E. Structural insights into RNA processing by the human RISC-loading complex. *Nature Structural & Molecular Biology*. 2009, 16:1148.
29. Sinha NK, Bass BL. Overexpression and purification of Dicer and accessory proteins for biochemical and structural studies. *Methods*. 2017.
30. Liu Z, Wang J, Cheng H, Ke X, Sun L, Zhang QC, Wang HW. Cryo-EM structure of human Dicer and its complexes with a pre-miRNA substrate. *Cell*. 2018, 173(5):1191-1203.
31. Durocher Y, Perret S, Kamen A. High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res*. 2002, 30(2):E9.
32. Tom R, Bisson L, Durocher Y. Transfection of HEK293-EBNA1 cells in suspension with linear PEI for production of recombinant proteins. *CSH Protoc*. 2008, 2008:pdb.prot4977.
33. Tom R, Bisson L, Durocher Y. Culture of HEK293-EBNA1 cells for production of recombinant proteins. *CSH Protoc*. 2008, 2008:pdb.prot4976.

34. Cass B, Pham PL, Kamen A, Durocher Y. Purification of recombinant proteins from mammalian cell culture using a generic double-affinity chromatography scheme. *Protein Expr Purif.* 2005, 40(1):77-85.
35. Schirle M, Heurtier MA, Kuster B. Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. *Mol Cell Proteomics.* 2003, 2(12):1297-1305.
36. Lau PW, Guiley KZ, De N, Potter CS, Carragher B, MacRae IJ. The molecular architecture of human Dicer. *Nat Struct Mol Biol.* 2012, 19(4):436-440.
37. Campbell FE, Jr., Cassano AG, Anderson VE, Harris ME. Pre-steady-state and stopped-flow fluorescence analysis of Escherichia coli ribonuclease III: insights into mechanism and conformational changes associated with binding and catalysis. *J Mol Biol.* 2002, 317(1):21-40.
38. Fukumoto Y, Obata Y, Ishibashi K, Tamura N, Kikuchi I, Aoyama K, Hattori Y, Tsuda K, Nakayama Y, Yamaguchi N. Cost-effective gene transfection by DNA compaction at pH 4.0 using acidified, long shelf-life polyethylenimine. *Cytotechnology.* 2010, 62(1):73-82.
39. Delafosse L, Xu P, Durocher Y. Comparative study of polyethylenimines for transient gene expression in mammalian HEK293 and CHO cells. *J Biotechnol.* 2016, 227:103-111.
40. Lee HY, Doudna JA. TRBP alters human precursor microRNA processing in vitro. *RNA.* 2012, 18(11):2012-2019.
41. Ma E, Zhou K, Kidwell MA, Doudna JA. Coordinated activities of human dicer domains in regulatory RNA processing. *J Mol Biol.* 2012, 422(4):466-476.
42. Li H, Nicholson AW. Defining the enzyme binding domain of a ribonuclease III processing signal. Ethylation interference and hydroxyl radical footprinting using catalytically inactive RNase III mutants. *EMBO J.* 1996, 15(6):1421-1433.
43. Leskovac V: *Comprehensive Enzyme Kinetics*: Springer US; 2007.
44. Nicholson AW. Ribonuclease III mechanisms of double-stranded RNA cleavage. *Wiley Interdiscip Rev RNA.* 2014, 5(1):31-48.
45. Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. *Cell.* 2004, 118(1):57-68.

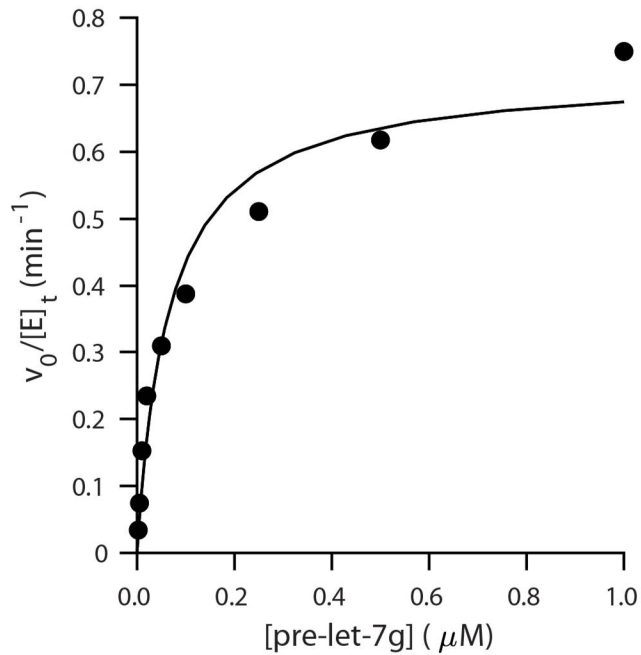
46. Fareh M, Yeom KH, Haagsma AC, Chauhan S, Heo I, Joo C. TRBP ensures efficient Dicer processing of precursor microRNA in RNA-crowded environments. *Nat Commun.* 2016, 7:13694.
47. Landthaler M, Gaidatzis D, Rothballer A, Chen PY, Soll SJ, Dinic L, Ojo T, Hafner M, Zavolan M, Tuschl T. Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA.* 2008, 14(12):2580-2596.
48. Lindner HA, Fotouhi-Ardakani N, Lytvyn V, Lachance P, Sulea T, Menard R. The papain-like protease from the severe acute respiratory syndrome coronavirus is a deubiquitinating enzyme. *J Virol.* 2005, 79(24):15199-15208.
49. Salvail-Lacoste A, Di Tomasso G, Piette BL, Legault P. Affinity purification of T7 RNA transcripts with homogeneous ends using ARiBo and CRISPR tags. *RNA.* 2013, 19(7):1003-1014.
50. Bouchard P, Lacroix-Labonte J, Desjardins G, Lampron P, Lisi V, Lemieux S, Major F, Legault P. Role of SLV in SLI substrate recognition by the *Neurospora* VS ribozyme. *RNA.* 2008, 14(4):736-748.
51. Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN. Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature.* 2011, 475(7355):201-205.
52. Zhao H, Brown PH, Schuck P. On the distribution of protein refractive index increments. *Biophys J.* 2011, 100(9):2309-2317.
53. de la Rosa-Trevin JM, Quintana A, Del Cano L, Zaldivar A, Foche I, Gutierrez J, Gomez-Blanco J, Burguet-Castell J, Cuenca-Alba J, Abrishami V *et al.* Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *J Struct Biol.* 2016, 195(1):93-99.
54. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods.* 2012, 9(7):671-675.
55. Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztanyi Z, El-Gebali S, Fraser M *et al.* InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Res.* 2017, 45(D1):D190-D199.

56. Qin H, Chen F, Huan X, Machida S, Song J, Yuan YA. Structure of the Arabidopsis thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein-protein interaction. *RNA*. 2010, 16(3):474-481.
57. Tian Y, Simanshu DK, Ma JB, Park JE, Heo I, Kim VN, Patel DJ. A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. *Mol Cell*. 2014, 53(4):606-616.
58. Kolakofsky D, Kowalinski E, Cusack S. A structure-based model of RIG-I activation. *RNA*. 2012, 18(12):2118-2127.

## 2.9. Supporting figures



**Figure 2.6. SEC-MALS analysis of purified WT Dicer stored in sucrose/DDM-containing storage buffer for 6 months at -80°C.** Peak integration of the UV absorbance trace shows that the main peak contains  $\geq 94\%$  of the eluted protein. The molar mass calculated from MALS analysis was normalized against BSA to give an average molecular weight of  $215 \pm 3$  kDa with a polydispersity index of  $1.00 \pm 0.2$  over the entire eluted peak. The hydrodynamic radius from quasi-elastic light-scattering (QELS) is  $13.7$  nm ( $\pm 4.5\%$ ).



**Figure 2.7. Preliminary kinetics studies of pre-let-7a-1 cleavage by Dicer under non-steady-state conditions.** The turnover frequency ( $v_0/[E]_t$ ) is plotted as a function of substrate concentration (2.5 nM to 1  $\mu$ M). The kinetic studies and analysis were performed as described in Methods (Dicer cleavage assay), but the concentration of Dicer (5 nM) was kept constant for each substrate concentration. The data shown here were fitted to the Michaelis-Menten equation to provide underestimated  $k_{cat}$  ( $0.72 \pm 0.05 \text{ min}^{-1}$ ) and  $K_M$  ( $65 \pm 16 \text{ nM}$ ) values, since the reactions conditions did not allow for the steady-state assumption (see text). Averages from two independent experiments yield  $k_{cat}$  of  $0.8 \pm 0.1 \text{ min}^{-1}$  and  $K_M$  of  $66 \pm 2 \text{ nM}$ .



## **3. Le double effet de TRBP sur le clivage des pre-miARN par Dicer**

### **3.1. Mise en contexte**

Lors de la collecte de résultats de cinétique du premier article sur la purification de Dicer (Chapitre 2), nous avons calculé un  $K_M$  bien plus élevé de ce qui avait précédemment été montré. En général, très peu d'études enzymatiques à l'état stationnaire avaient été réalisées, mais cette valeur de  $K_M$  suggérait que Dicer avait besoin d'un activateur afin d'améliorer son efficacité catalytique. Nous avons donc supposé que TRBP avait un rôle à jouer dans cette activation.

Au moment de commencer ces travaux, à l'automne 2017, nous avons accès à beaucoup de données sur TRBP, soit la fonction des différents domaines de celle-ci ainsi que son interaction avec elle-même, avec Dicer et les pre-miARN. Toutes ces informations ont été utilisées pour l'élaboration initiale du modèle enzymatique d'activation de Dicer par TRBP.

Cette deuxième étude avait donc pour but de présenter un modèle global de l'action de TRBP sur l'activité de Dicer d'un point de vue de cinétique à l'état stationnaire. Cette étude présentait un défi car TRBP agit à la fois sur Dicer et son substrat. Cependant, avec la combinaison d'expériences de liaison et de cinétique à l'état stationnaire sur différentes constructions de TRBP, nous avons pu atteindre l'objectif. Au moment de présenter cette thèse, les travaux de ce projet sont encore préliminaires et bénéficieront de quelques résultats complémentaires qui seront discutés dans la section 4 de cette thèse.

La plupart des travaux de cette étude, allant du clonage jusqu'à l'analyse des résultats de cinétique ont été réalisés par moi-même. Fatima El Azzouzi, Rishov Mukhopadhyay et Soumiya Amellah ont réalisé certaines des expériences que j'ai analysées par la suite. La planification des expériences a été faite par moi-même et Pascale Legault (PL). La préparation initiale des figures et du manuscrit a été faite par moi-même et révisé par PL.

***TRBP as a dual modifier of pre-miRNA cleavage by Dicer***

Jonathan Bouvette<sup>1</sup>, Fatima El Azzouzi<sup>1</sup>, Rishov Mukhopadhyay<sup>1</sup>, Soumiya Amellah<sup>1</sup> and  
Pascale Legault<sup>1</sup>

<sup>1</sup>Département de Biochimie et Médecine Moléculaire, Université de Montréal, C.P. 6128,  
Succursale Centre-Ville, Montréal, QC, Canada,

**Corresponding author:** [pascale.legault@umontreal.ca](mailto:pascale.legault@umontreal.ca)

## 3.2. Abstract

Dicer is the enzyme responsible for the last step of miRNA maturation that allows production of ~2000 miRNAs in human cells, which regulate gene expression and profoundly affect health and diseases. Dicer specifically recognizes and cleave pre-miRNA hairpins, but it is not yet clear how its activity varies across such a large pool of natural substrates. Its partner protein TRBP has been shown to enhance pre-miRNA cleavage and promote the assembly of the RNA-induced silencing complex, however there is paradoxical evidence that TRBP can also inhibit Dicer cleavage. Here, we performed binding studies and steady-state kinetics of Dicer to compare the cleavage activity of Dicer for the twelve members of the human pre-let-7 family and thoroughly investigate the role of TRBP in the cleavage reaction. We show that, despite structural discrepancies amongst the pre-let-7 substrates, Dicer has a similar affinity and specificity towards all family members. Our kinetic results are consistent with the current view that the enzyme-substrate complex undergoes a rate-limiting conformational change before the cleavage step. We found that TRBP plays a dual role as an enzymatic modifier; it activates Dicer cleavage activity at low substrate concentrations and inhibits it at high substrate concentrations. To thoroughly describe the effect of TRBP on Dicer processing, we developed and parametrized a detailed steady-state kinetic model using pre-let-7g as the substrate. Our kinetic model provides a mechanistic view of how TRBP modulates Dicer activity, which helps to reconcile previous results and allows us to reconsider the role of TRBP in a physiological context.

## 3.3. Introduction

A detailed understanding of biological macromolecules, such as enzymes and molecular machines, is essential to make rational progress in both basic and applied biomedical research. In fact, it is often a challenge to relate available structural data with results from *in vitro* and *in vivo* functional studies and define the role of regulators. To address these issues, it often becomes necessary to develop and evaluate functional models. One such approach, global fitting of enzymatic models, is particularly well suited to describing regulation mechanisms. Here, we propose global fitting of a kinetic model for the enzymatic regulation of human Dicer by its molecular partner TRBP, which together form a key molecular machine in the maturation of microRNAs.

MicroRNAs (miRNAs) play important roles in the regulation of gene expression by recognizing mRNAs through sequence complementarity as part of the RNA-induced silencing complexes (RISC), which generally triggers translation arrest and miRNA destabilization (1). The human genome encodes around 2000 miRNAs which, together, have the potential to regulate over 60% of all mRNAs (2-6). Misexpression or misregulation in the miRNA biogenesis pathway can lead to faulty gene expression, alter the behavior of a cell and lead to development defects and diseases (reviewed in (7)). These effects have been well characterized by the study of the pre-let-7 family of miRNAs, which were initially identified as essential for the development of *C. elegans* to the adult stage and later associated with tumor suppression, metabolism and growth in human (reviewed in (8, 9)). To become functional, canonical miRNAs must go through a tightly-regulated biogenesis pathway that starts by the transcription of a primary miRNA that is cleaved into a precursor miRNA hairpin (pre-miRNA) by Drosha/DGCR8 and then cleaved into a miRNA duplex by Dicer from which one strand is loaded into Argonaute (Ago) to form the RISC. The regulation of the two cleavage steps allows for the specific temporal and spatial expression of miRNAs.

Dicer is responsible the last step of miRNA maturation. It is a large RNase III enzyme that contains a combination of different RNA-binding domains (Figure 3.1a) to recognize the pre-miRNA substrate and allow for the precise positioning of the cleavage sites (10, 11). Dicer adopts a L-shape structure where the PAZ domain is positioned at the top, the helicase domain at the base and the catalytic domains in between (12-14). This architecture allows Dicer to clamp the extremities of the pre-miRNA hairpin, the 5'-phosphate/2-nt 3'-overhang and terminal loop with its PAZ and helicase domains, respectively. Mutagenesis studies have shown that the presence of the 5'-phosphate is required for precise cleavage and that altering the length of the 3'-overhang impairs Dicer activity (10). Moreover, Dicer efficiently cleaves pre-miRNAs with terminal loops > 9 nt and elongation of the dsRNA stem leads to reduced and inaccurate processing (15-18). This confers Dicer with a high specificity for pre-miRNAs over other cellular and viral RNAs.

The pitfall of having such high specificity for pre-miRNA substrates is a loss of catalytic efficiency, as previously shown from studies of Dicer truncation mutants. Removing the helicase domain of Dicer was shown to drastically increase the rate of the cleavage reaction (19), but this came concomitantly with a loss of specificity as this truncated form of Dicer

cleaved long dsRNA even more efficiently. Similarly, Dcr-1 in drosophila lacking its helicase domain could not discriminate between canonical and mutant pre-miRNAs (16). Moreover, in a recent cryo-EM study of Dicer bound to a pre-miRNA, most particles show an inactive enzyme-substrate conformation with the substrate positioned away from the catalytic site (13). These results indicate that the catalytic site of Dicer is rather inaccessible, even for a canonical pre-miRNA substrate. Several questions remain regarding the mechanism of Dicer towards these substrates and how it can function efficiently in a cellular context where several competing RNA substrates are present. For example, it is still ambiguous as to whether Dicer prefers some pre-miRNA substrates over others as their size, sequence and structure can vary significantly. Moreover, recent steady-state kinetic studies with pre-let-7a-1 revealed a surprisingly high  $K_M$  value (1.2  $\mu\text{M}$ ), which suggests that Dicer may be an inefficient enzyme *in vivo* at nanomolar substrate concentrations. It would be interesting to know if this is a general phenomenon and it would be important to understand the mechanisms by which partner proteins, such as TRBP and PACT, and regulatory factors help modulate the efficiency of Dicer cleavage.

The TAR-RNA binding protein (TRBP) was originally identified via a yeast triple hybrid assay as an important interactor of the HIV-1 TAR RNA hairpin, and later found to play diverse roles in RNA interference (20). TRBP possess three double-stranded RNA binding domains (dsRBDs; Figure 3.2a), of which its two N-terminal dsRBDs are known to interact with the ribose-phosphate backbone of RNA double helices (17, 21, 22) (Figure 3.2a). In contrast, its C-terminal dsRBD is unable to bind nucleic acids, but there is solid biochemical and structural evidence that it interacts with the helicase domain of Dicer (23, 24). Although TRBP was initially described as being necessary for the assembly of the RISC complex, there were contradictory results about whether or not it could activate Dicer cleavage *in vivo* (25-27). TRBP was later confirmed as an activator of Dicer through *in vitro* kinetic studies as it could enhance the cleavage activity of Dicer by up to 5-fold (17, 28). Cryo-EM studies of Dicer bound to TRBP revealed that TRBP could induce a change in the conformation of the helicase domain of Dicer and potentially affect its interaction with the pre-miRNA (14). Single-molecule fluorescence studies have shown that TRBP changes the interaction dynamics of Dicer with RNAs, enhancing the specific selection of canonical pre-miRNAs over competing tRNAs (29). TRBP was also found to change the Dicer cleavage site of some pre-miRNAs, leading to the formation of miRNA isoforms or isoMiRs and thereby changing the set of target mRNAs (17, 23, 30-32).

Intriguingly, a few reports have described the ability of TRBP to inhibit the activity of Dicer for certain pre-miRNAs (17) and under certain conditions (33). At this time, it remains unclear how TRBP can activate pre-miRNA processing without impairing the high specificity of Dicer for pre-miRNAs and how it can both serve as an activator and an inhibitor of cleavage activity.

In this paper, we performed kinetic studies of Dicer cleavage to investigate its activity towards different pre-miRNA substrates as well as to provide a mechanistic understanding of TRBP as a Dicer modifier. First, we measured the binding affinity and steady-state kinetic parameters of Dicer for all twelve pre-let-7 family members in the absence of TRBP. Then, we performed several assays to investigate the effect of TRBP on the cleavage of pre-let-7 substrates by Dicer, revealing that TRBP can both act as an activator and an inhibitor of a given substrate depending on its concentration. To further investigate the role of TRBP as a dual enzyme modifier, we propose a kinetic model for cleavage of a pre-miRNA substrate by Dicer in the presence of TRBP, which incorporates current knowledge on the interaction network between Dicer, TRBP and pre-miRNAs. For global fitting of the kinetic model, we performed thermodynamic and steady-state kinetic experiments with the pre-let-7g substrate in the presence of either TRBP or TRBP fragments. Our parametrized kinetic model provides a mechanistic understanding of how TRBP modulates Dicer activity depending on substrate concentration. Moreover, it synthesizes current knowledge and helps to resolve apparently conflicting results from the literature. Interestingly, this enzymatic framework provides novel insights into the role of TRBP in a physiological context.

## 3.4. Material and methods

### 3.4.1. Plasmids

The pTT5 vectors used for expression of his-tagged human Dicer proteins (WT and D1320A/D1709A variant) and the pUC19-HH-pre-let-7-HDV vectors used for *in vitro* transcription of human pre-let-7 RNAs were obtained as described previously (34). The vector expressing full-length TRBP (pET-TEV-TRBP) was generated by PCR amplification of pDONR221-TARBP2 (DNASU #HsCD00045566) (35-37) and inserted between the BamHI and EcoRI sites of a modified pET-15b vector to allow for its expression with a His<sub>6</sub> tag and a TEV cleavage sequence at the N-terminus. The vector expressing a TRBP variant containing only the first two dsRBD (T12; residues 1-236 of TRBP) was obtained by site-directed mutagenesis through incorporation of an amber stop codon in pET-TEV-TRBP. The vector expressing the dsRBD3 of TRBP (T3; residues 293-366 of TRBP) was generated by PCR amplification from plasmid pDONR221-TARBP2 and inserted between BamHI and EcoRI sites of a modified pGEX-2T vector to allow its expression with a GST tag and a TEV cleavage sequence at the N-terminus. The sequence of all plasmid vectors was verified by DNA sequencing.

### 3.4.2. Synthesis and purification of pre-let-7 RNAs

The synthesis of all pre-let-7 substrates and 5'-end <sup>32</sup>P-labeling was carried out as described previously (34, 38). Briefly, each pre-let-7 RNA was transcribed *in vitro* with 5'-Hammerhead and 3'-HDV ribozyme tags that self-cleaved in the transcription reaction to yield an RNA with homogeneous 5' - and 3'-ends. The pre-miRNA was then purified by denaturing gel electrophoresis and phosphorylated at its 5'-end with T4 polynucleotide kinase with either unlabeled ATP or ATP- $\gamma$ -<sup>32</sup>P. Subsequently, the 5'-monophosphorylated RNAs were purified either by denaturing anion-exchange HPLC (cold RNA) or by denaturing gel electrophoresis (<sup>32</sup>P-labeled RNA) and stored in TE buffer (10 mM Tris pH 8.0 and 1 mM EDTA) at -20°C.

### 3.4.3. Protein expression and purification

The Dicer protein and the D1320A/D1709A Dicer variant were expressed from HEK293-6E cells and purified by anion-exchange chromatography, Ni-NTA affinity chromatography and size-exclusion chromatography, as described previously (34).

TRBP and T12 were expressed in *Escherichia coli* (*E. coli*) Rosetta (DE3) pLys strain (Millipore-Sigma #70956). The bacterial cultures were grown at 37°C in LB medium supplemented with 35 mg/L chloramphenicol and 100 mg/L ampicillin, induced at an OD<sub>600</sub> of 0.8 with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and then grown for 16 h at 16°C. The cells were harvested by centrifugation and resuspended in HisA buffer (20 mM Tris pH 8.0, 500 mM NaCl, 20 mM imidazole and 4 M Urea) supplemented with one tablet of Roche cOmplete protease inhibitor (Sigma # 4693159001). The cells were then lysed by French press and centrifuged at 100,000 ×g for 1 h at 4°C. The supernatant was loaded onto a 20-mL gravity flow Nickel-bound IMAC column (GE Healthcare) equilibrated with HisA buffer. The column was then washed with 10 column volumes (CV) of HisA buffer and the bound proteins eluted with HisB buffer (20 mM Tris pH 8.0, 500 mM NaCl and 300 mM imidazole). The fractions containing TRBP were pooled, supplemented with TEV protease and dialyzed (Spectra Por 4, 12-14 MWCO) overnight in TEV buffer (20 mM Tris pH 8.0, 100 mM NaCl and 5 mM DTT) at room temperature. The protein mixture was then loaded on 60-mL of SP Sepharose High Performance media packed in a XK 26/20 column (GE Healthcare) and equilibrated with SPsephA buffer (20 mM Tris pH 8.0 and 1 mM DTT). The column was then washed with 2.5 CV of SPsephA buffer, and TRBP was eluted using a linear gradient (from 0% to 100% over 4 CV) of SPsephB buffer (20 mM Tris pH 8.0, 500 mM NaCl and 1 mM DTT). Fractions containing TRBP were pooled and dialyzed in storage buffer (100 mM Tris pH 7.4, 100 mM NaCl and 20% glycerol). The dialyzed sample was then concentrated to 30-50 μM using a 10-kDa NMWL Amicon® Ultra 15-mL centrifugal filters (Millipore-Sigma #UFC901024). The final TRBP sample was aliquoted, flash frozen in liquid nitrogen and stored at -80°C for future use.

T3 was expressed in *E. coli* Rosetta2 cells (Millipore-Sigma #71402). The bacterial cultures were grown at 37°C in LB medium supplemented with 35 mg/L chloramphenicol and 100 mg/L ampicillin and induced at an OD<sub>600</sub> of 0.6 with 1 mM IPTG for 4 h at 30°C. The cells



were harvested by centrifugation and resuspended in GSTA buffer (20 mM Tris pH 8.0, 1 M NaCl and 1mM DTT) supplemented with one tablet of Roche cOmplete protease inhibitor (Sigma #4693159001). The cells were then lysed by French press and centrifuged at 100,000  $\times g$  for 1 h at 4°C. The supernatant was incubated for 1 h with 5 mL of Glutathione (GSH) Sepharose 4B resin (GE Healthcare) equilibrated with GSTA buffer. After incubation, the resin was washed with 10 CV of GSTA buffer by gravity flow. T3 was eluted with GSTB buffer (20 mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT and 20 mM reduced GST). Fractions containing T3 were supplemented with TEV protease and dialyzed (Spectra Por 4, 3500 Da MWCO) overnight in TEV buffer (20 mM Tris pH 8.0, 100 mM NaCl and 5 mM DTT) at room temperature. T3 was further purified by SP Sepharose ion-exchange chromatography, dialyzed in storage buffer, aliquoted and stored, as described above for TRBP.

#### **3.4.4. Steady-state kinetic studies in the absence of TRBP**

Steady-state kinetic studies involving Dicer and the pre-let-7 substrates were carried out as described previously (34). Briefly, RNA and proteins solutions were prepared separately at twice their final desired concentration in 30  $\mu\text{L}$  of the cleavage reaction buffer (50 mM HEPES pH 7.4, 50 mM NaCl, 5 mM  $\text{MgCl}_2$  and 0.05% NP-40). The RNA solutions, containing 80 pM of  $^{32}\text{P}$ -labeled pre-let-7g (or other pre-let-7) and the non-labeled RNA equivalent at concentrations ranging from 0.1X to 10X the estimated  $K_M$ , were heated at 95°C for 2 min and snap-cooled at 4°C for at least 5 min to refold the RNA. Both RNA and protein solutions were pre-heated for 5 min at 37°C, and cleavage reactions were initiated by adding 30  $\mu\text{L}$  of the protein solution to 30  $\mu\text{L}$  of the RNA solution. At six specific time points, a 5- $\mu\text{L}$  aliquot of the cleavage reaction was taken and immediately mixed with 5  $\mu\text{L}$  of stop buffer (95 % formamide and 50 mM EDTA). Aliquots of the cleavage reaction were analyzed by denaturing gel electrophoresis [20% acrylamide:bis-acrylamide (19:1)/7 M urea gel]. The amounts of substrate (S) and product (P) were quantified from radioactive bands using a Personal Molecular Imager (PMI) System (Bio-Rad), and the fraction of cleaved product [ $F = P / (S + P)$ ] was plotted against time. The resulting time courses were fitted by linear regression and the slope was taken as the initial velocity ( $v_0$ ). To ensure that initial cleavage rates were measured under steady-state conditions, the enzyme concentration ( $[E]$ ) was varied along with the initial substrate concentration ( $[S]$ ) to maintain  $[S]/[E] \geq 50$ , and time points were collected after at least one

turnover of the enzyme pool and with 5-10% of the substrate being cleaved within 30 min. The dependence of  $v_0/[E]_t$  on  $[S]$  was plotted and fitted to the Michaelis-Menten equation (written as  $v_0/[E]_t = k_{cat}[S]/(K_M+[S])$ ) with the OriginPro 8 software (OriginLab) to derive  $k_{cat}$  and  $K_M$  values. The values of  $k_{cat}$  and  $K_M$  are reported with standard deviations from a single fit of two independent experiments made at staggered substrate concentrations. The quality of the fit was obtained from the square of the correlation coefficient ( $R^2$ ), and in all cases  $R^2 \geq 0.97$  and the distribution of residuals was random around the regression line.

### **3.4.5. Effect of TRBP on the possible production of isoforms of let-7-5p**

To analyze the 5'-cleavage products resulting from cleavage of all human pre-let-7 by Dicer, multiple turnover reactions were set up containing 1  $\mu$ M of non-labeled pre-let-7 and 40 pM of  $^{32}$ P-labeled pre-let-7 in cleavage reaction buffer and one of the following combinations: 1) no Dicer and no TRBP; 2) 40 nM Dicer and no TRBP; and 3) 40 nM and 100 nM TRBP. Reactions were initiated as described for the steady-state kinetic studies by mixing 10  $\mu$ L each of the RNA and protein solutions. After a 15-min incubation at 37°C, the reactions were stopped by adding 60  $\mu$ L of stop buffer. The T<sub>1</sub> ladders were generated by incubating 100 fmol of  $^{32}$ P-labeled pre-let-7a-2 in a 9- $\mu$ L volume containing 20 mM sodium citrate pH 5.0, 1 mM EDTA, 7 M Urea, 0.37 mg/mL yeast tRNA with 1 U of T<sub>1</sub> nuclease (Sigma-Aldrich) for 20 min at 37°C. The hydroxyl ladders were generated by incubating 100 fmol of  $^{32}$ P-labeled pre-let-7a-2 in a 8- $\mu$ L volume containing 50 mM NaHCO<sub>3</sub>/Na<sub>2</sub>CO<sub>3</sub> pH 9.2, 1 mM EDTA, 0.25 mg/mL yeast tRNA for 5 min at 95 °C. Reactions were stopped by adding ~1X volume of stop buffer. Each sample was analyzed by loading 10  $\mu$ L (0.1 fmol)  $^{32}$ P-labeled RNA on a denaturing gel electrophoresis [20% acrylamide:bis-acrylamide (19:1) /7 M urea] gel run at 60 W for about 2.5 h. The gel was dried, exposed for 16 h to a storage phosphor screen and visualized using a PMI System (Bio-Rad).

### **3.4.6. Effect of TRBP on the relative cleavage efficiency of different pre-let-7 substrates by Dicer**

To compare the effect of TRBP on the cleavage of different pre-let-7 substrates (pre-let-7a-1, -7c and -7g), several reactions were prepared in cleavage reaction buffer at five concentrations of TRBP (0, 10, 40, 80, 160 and 1000 nM), each one using five 5 different

substrate concentrations (5, 25, 125, 625 and 3125 nM). The specific concentration of Dicer used at each substrate concentration was calculated such that  $[S]/[E] \geq 50$ , and time points were collected such that less than 10% of the substrate was cleaved within 20 min in the absence of TRBP. Each reaction was prepared in two separate tubes at twice their final concentration: one 15- $\mu$ L tube containing pre-let-7 and another tube containing both Dicer and TRBP. Both tubes were pre-incubated for 10 min at 37°C and reactions were started by adding 15  $\mu$ L from the Dicer-TRBP tube to the pre-let-7 tube. After 20 min, the reactions were quenched by adding 20  $\mu$ L of stop buffer. For each substrate concentration, the relative cleavage efficiency was calculated by dividing the fraction of cleaved substrate obtained at a specific concentration of TRBP by the fraction of cleaved substrate measured in the absence TRBP.

### **3.4.7. Evaluation of a potential pre-steady-state burst**

To evaluate the possibility for a pre-steady-state burst, cleavage assays were carried out in cleavage reaction buffer as for standard steady-state kinetic studies. Cleavage reactions in the absence of TRBP contained 5  $\mu$ M of pre-let-7g and either 6, 12, or 18 nM of Dicer, whereas those with TRBP contained 0.25  $\mu$ M pre-let-7g, 80 nM of TRBP and 0.75, 1.5 or 2.25 nM of Dicer. The cleavage reactions were incubated in a 60- $\mu$ L volume at 4°C for 30 min. Several aliquots (2  $\mu$ L) were collected at different times (20-1800 sec) and mix immediately with 10  $\mu$ L of stop buffer before transferring to ice. These aliquots were analyzed by denaturing gel electrophoresis to derive the fraction of cleaved substrate (F), and the concentration of product ( $[P] = F \times [S]$ ) was then plotted against time. From the resulting time course, the early time points that fall under the steady-state conditions ( $F < 10\%$  and  $[P] \geq [E]$ ) were fitted by linear regression; the slope was taken as the initial rate of reaction ( $v_0$ ) and the y-intercept as the amplitude of the potential burst ( $\pi$ ).

### **3.4.8. Steady-state kinetic studies in the presence of TRBP**

Steady-state kinetic studies in the presence of either TRBP or TRBP domains (T12 or T3) were carried out essentially as those performed in the absence of TRBP, but with the pre-let-7g substrate only. The concentration of TRBP proteins was kept at a  $[TRBP]/[Dicer] \geq 10$  to ensure that that the presence of Dicer does not significantly affect the concentration of free TRBP. For each set of experiments with a given TRBP protein (TRBP, T12, or T3), several

cleavage reactions were prepared at different concentrations of TRBP (For T12 and T3: 0, 20, 80, 320 and 3000 nM and for TRBP: 0, 5, 10, 20, 40, 80, 160, 320 and 3000 nM), each one using 7-9 substrate concentrations, and the dependence of  $v_0/[E]_t$  on  $[S]$  was analyzed at each TRBP concentration. Values of  $v_0/[E]_t$  were corrected by subtracting from the cleavage data the averaged slope of product degradation measured on two degradation bands on the gel, to take into account the possibility that the band of product could result from RNA degradation instead of Dicer cleavage. To correct for the minor variations in  $k_{cat}$  values that originate from the daily preparation of Dicer dilutions, each value of  $v_0/[E]_t$  was normalized on a 0 to 1 scale by dividing by the  $k_{cat}$  value obtained at 0 nM TRBP when fixing the  $K_M$  value for that experiment with the one previously determined in triplicate for pre-let-7g (Tableau I). For each of the three TRBP proteins (TRBP, T12, and T3), two independent sets of experiments were collected, and the data from both experiments were fitted simultaneously to the appropriate enzymatic model (see result section). The T12 and T3 models were first fitted independently; the T12 model to evaluate the  $\gamma$  and  $g$  parameters ( $R^2 = 0.89$ ) and the T3 model to evaluate of the  $\alpha$ ,  $a$  and  $K_{EB}$  parameters ( $R^2 = 0.90$ ). Values obtained from fitting the T12 and T3 models were then fixed to fit the global TRBP enzymatic model and evaluate the  $\beta$  and  $b$  parameters ( $R^2 = 0.93$ ). The values extracted from these fits are provided with their associated standard deviation. The data for T3 were fitted with OriginPro 8 software (OriginLab) and the data for T12 and TRBP with MATLAB R2018a.

### 3.4.9. Binding assays for $K_d$ determination

Binding assays to characterize the pre-let-7/Dicer interaction were performed with the D1320A/D1709A Dicer variant using electrophoretic mobility shift assay (EMSA), as described previously (34). Binding assays to characterize interaction between pre-let-7g and TRBP were carried out using similar procedures. Briefly, the binding reactions were initiated by adding 10  $\mu$ L of the RNA sample to 10  $\mu$ L of the protein samples, and the mixture was incubated at 37°C for 30 min. TRBP-binding reactions were prepared in TB buffer (50 mM Tris pH 7.4, 50 mM NaCl, 0.05% NP-40, 20% glycerol and 2 mM DTT), whereas Dicer-binding reactions were in DB buffer (50 mM Tris pH 7.6, 50 mM NaCl, 10% glycerol, 0.05% NP-40 and 2 mM DTT) and incubated at 4°C for 30 min. Both TRBP and Dicer binding reactions contained 2 pM  $^{32}$ P-labeled RNA and protein concentrations varying between 0.01X and 100X of the estimated  $K_d$  value. The Dicer-binding reactions were loaded on a 4%-15% gradient polyacrylamide gel (37.5:1

acrylamide:bis-acrylamide) in Tris-Glycine buffer (25 mM Tris–Base and 200 mM glycine) and run for 2 h at 200 V in the cold room. TRBP-binding reactions were loaded on a 9% acrylamide gel (37.5:1 acrylamide:bis-acrylamide) in Tris-Glycine buffer and run for 2 h at 250 V in the cold room. The protein-bound and unbound RNA fractions were quantified by phosphorimaging, the fraction of bound RNA was then plotted against protein concentration, and the binding data were fitted to the Hill equation with the OriginPro 8 software (OriginLab).

## 3.5. Results

### 3.5.1. Binding studies of pre-let-7 to Dicer

To investigate the affinity of human Dicer for different pre-miRNAs, we performed binding studies with all twelve pre-let-7 RNAs found in human. From electrophoretic mobility shift assays (EMSA) with  $^{32}\text{P}$ -end labeled RNA, the fraction of Dicer-bound RNA was estimated at several protein concentrations (Figure 3.1c). These data were fitted to the Hill equation (Figure 3.1d) to derive the dissociation constant ( $K_d$ ) and Hill coefficient ( $n$ ). From binding assays performed in triplicate, we found that the  $K_d$  values for all pre-let-7 fall within an 11-fold window in the low nanomolar range (5-60 nM) with Hill coefficients between 0.9 and 2.8 (Tableau I). These results demonstrate that Dicer has a similar affinity to all twelve pre-let-7 substrates.

### 3.5.2. Steady-state kinetic studies of pre-let-7 cleavage by Dicer

To investigate the pre-miRNA cleavage activity of Dicer, we carried out kinetic studies under steady-state conditions for all twelve human pre-let-7 substrates (Figure 3.8). This was achieved by first determining the initial velocity ( $v_0$ ) of the cleavage reaction at several different substrate concentrations ( $[\text{S}]$ ), and then fitting the plot of  $v_0/[\text{E}]_t$  as a function of  $[\text{S}]$  to derive the  $k_{cat}$  and  $K_M$  values (Figure 3.1ef). Although the  $k_{cat}$  values vary by up to 15 folds ( $1.4\text{-}21\text{ min}^{-1}$ ) and the  $K_M$  values by up to 20 folds ( $0.13\text{-}2.6\text{ }\mu\text{M}$ ), the range of observed  $k_{cat}/K_M$  values is limited to 4.4 folds ( $2.7\text{-}12\text{ min}^{-1}\mu\text{M}^{-1}$ ; Tableau I). These results indicate that the catalytic efficiency of Dicer is similar for cleavage of the different pre-let-7 substrates.

Tableau I. Summary of binding and kinetic constants of Dicer for the pre-let-7 family.

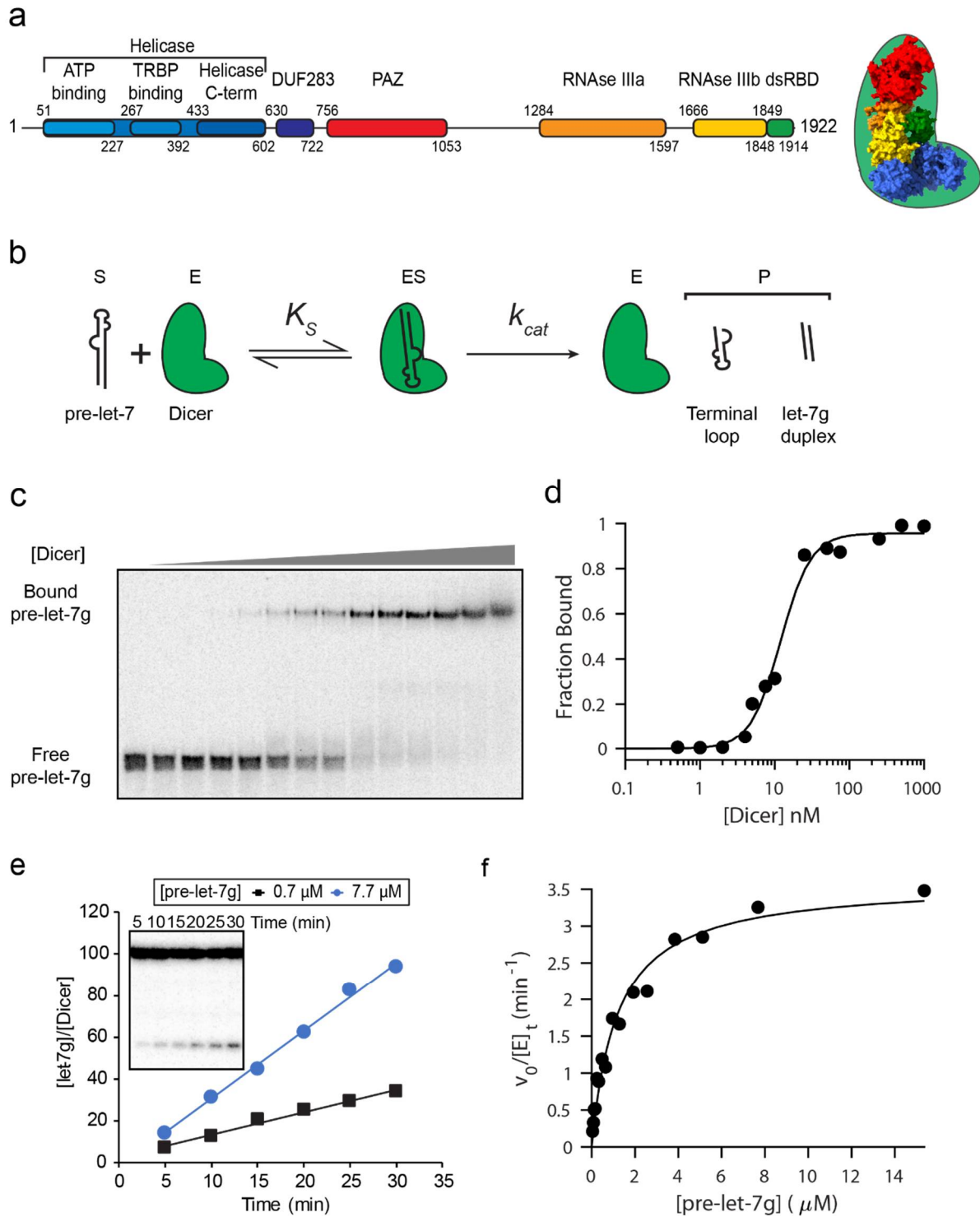
pre-let-7 substrate	$K_M$ ( $\mu\text{M}$ )	$k_{cat}$ ( $\text{min}^{-1}$ )	$k_{cat}/K_M$ ( $\text{min}^{-1}\mu\text{M}^{-1}$ )	$k_{cat}/K_M$ ratio <sup>1</sup>	$K_d$ (nM)	Hill coefficient	$K_d$ ratio <sup>2</sup>
a1 <sup>3</sup>	$1.2 \pm 0.3$	$7.2 \pm 0.5$	$6 \pm 1$	$2.2 \pm 0.6$	$5 \pm 1$	$1.3 \pm 0.2$	$1.0 \pm 0.3$
a2	$0.4 \pm 0.2$	$1.4 \pm 0.4$	$3 \pm 2$	$1.2 \pm 0.6$	$28 \pm 3$	$1.0 \pm 0.2$	$6 \pm 1$
a3	$0.90 \pm 0.06$	$4.6 \pm 0.1$	$5.1 \pm 0.4$	$1.9 \pm 0.2$	$48 \pm 19$	$1.3 \pm 0.3$	$10 \pm 5$
b	$0.50 \pm 0.05$	$3.2 \pm 0.2$	$6.3 \pm 0.8$	$2.3 \pm 0.4$	$9 \pm 1$	$1.5 \pm 0.2$	$1.8 \pm 0.4$
c	$0.27 \pm 0.03$	$3.0 \pm 0.1$	$11 \pm 1$	$4.1 \pm 0.6$	$8 \pm 6$	$1.6 \pm 0.6$	$2 \pm 1$
d	$2.6 \pm 0.6$	$21 \pm 4$	$8 \pm 3$	$3 \pm 1$	$40 \pm 3$	$1.04 \pm 0.02$	$8 \pm 2$
e	$0.13 \pm 0.01$	$1.6 \pm 0.1$	$12 \pm 1$	$4.4 \pm 0.6$	$15 \pm 1$	$2.8 \pm 0.1$	$3.0 \pm 0.7$
f1	$1.2 \pm 0.1$	$8.7 \pm 0.6$	$7.1 \pm 0.9$	$2.6 \pm 0.4$	$20 \pm 3$	$2.1 \pm 0.3$	$4 \pm 1$
f2	$0.81 \pm 0.08$	$8.4 \pm 0.4$	$10 \pm 1$	$3.8 \pm 0.6$	$17 \pm 9$	$2.2 \pm 0.5$	$3.5 \pm 2.1$
g	$1.11 \pm 0.08$	$3.3 \pm 0.2$	$3.0 \pm 0.2$	$1.1 \pm 0.1$	$16 \pm 4$	$2.4 \pm 0.8$	$3 \pm 1$
i	$1.1 \pm 0.2$	$3.2 \pm 0.3$	$2.7 \pm 0.3$	$1.0 \pm 0.2$	$17 \pm 3$	$1.7 \pm 0.4$	$4 \pm 1$
98	$0.36 \pm 0.06$	$2.1 \pm 0.2$	$6 \pm 1$	$2.1 \pm 0.5$	$19 \pm 9$	$1.5 \pm 0.2$	$4 \pm 2$

<sup>1</sup>  $k_{cat}/K_M$  ratios were calculated with the value for pre-let-7i as a reference

$$[(k_{cat}/K_M)_{pre-let-7}/(k_{cat}/K_M)_{pre-let-7i}]$$

<sup>2</sup>  $K_d$  ratios were calculated with the value for pre-let-7a-1 as a reference ( $K_{dpre-let-7}/K_{dpre-let-7a-1}$ )

<sup>3</sup> Values reported from (33)



**Figure 3.1. Thermodynamic and kinetic studies of pre-let-7 cleavage by Dicer.** (a) Schematics of the primary and tertiary structures of human Dicer using the same color code to identify the protein domains. The overlay of the three-dimensional structure of Dicer on the

cartoon representation was produced from PDB ID 5ZAL using the ChimeraX software (13, 62). **(b)** Pictogram representation of the simplified kinetic model for pre-let-7 cleavage by Dicer. **(c-d)** Binding studies of Dicer to pre-let-7g. **(c)** Typical EMSA performed with 10 pM of 5'-[<sup>32</sup>P]-labeled pre-let-7g and increasing concentrations of the D1320A/D1709A mutant Dicer (0, 0.1, 0.5, 1, 2.5, 5, 7.5, 15, 25, 50, 100, 250, 500 and 1000 nM). **(d)** Binding curve for the data shown in (c). These data were fitted to the Hill equation to obtain a  $K_d$  value of  $12.0 \pm 0.9$  nM. **(e-f)** Steady-state kinetics of cleavage of pre-let-7g by Dicer. **(e)** Plot of the normalized product formation ( $[\text{let-7g}]/[\text{Dicer}]$ ) as a function of time. The gel used to quantify the cleavage data with  $[\text{pre-let-7g}] = 0.7 \mu\text{M}$  is shown as an insert. The turnover frequency ( $v_o/[E]_i$ ) was calculated by linear regression of the slope ( $1.08 \pm 0.06 \text{ min}^{-1}$  for  $[\text{pre-let-7g}] = 0.7 \mu\text{M}$  and  $3.3 \pm 0.1 \text{ min}^{-1}$  for  $[\text{pre-let-7g}] = 7.7 \mu\text{M}$ ). **(f)** Typical steady-state kinetic analysis showing the dependence of the turnover frequency on substrate concentration (0.04 to 15.36  $\mu\text{M}$ ). The data shown here were fitted to the Michaelis-Menten equation to yield  $k_{cat}$  of  $3.6 \pm 2 \text{ min}^{-1}$  and  $K_M$  of  $1.2 \pm 0.2 \mu\text{M}$ .

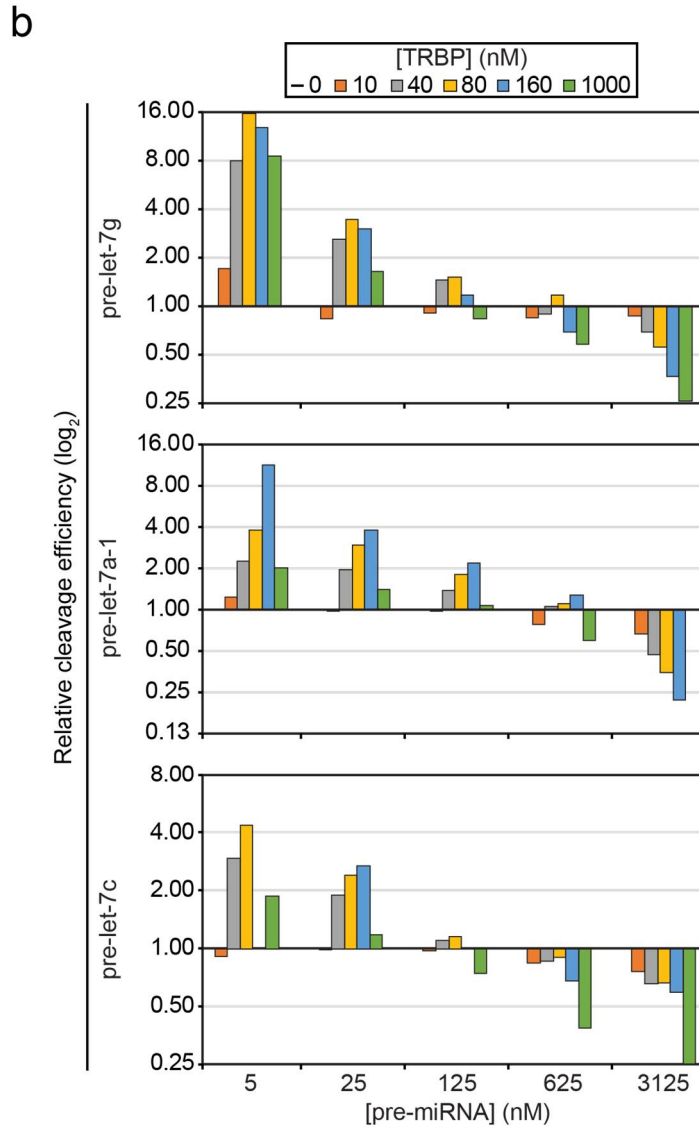
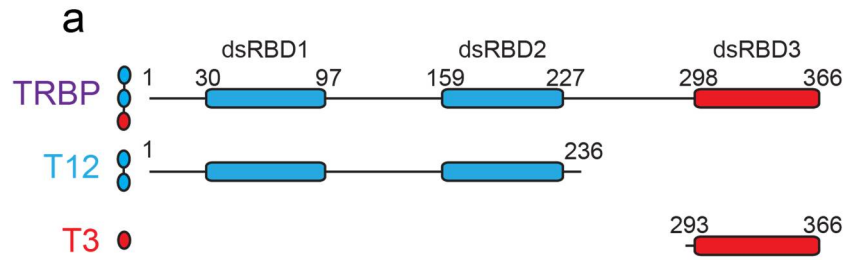
### 3.5.3. General effect of TRBP on pre-let-7 cleavage

To examine the effect of TRBP on pre-let-7 cleavage, we first investigated if the presence of TRBP could induce the formation of isoforms of let-7-5p with different lengths, since it was previously shown that TRBP can trigger the formation of isomiRs that are longer than the canonical sequence by one nucleotide (17, 23, 30, 31). In the absence of TRBP, *in vitro* cleavage of all pre-let-7 substrates by Dicer yield a single 22-nt let-7-5p product, except for pre-let-7e and pre-let-7f-1 which yield a minor 23-nt product in addition to the main 22-nt let-7-5p product (Figure 3.9). Based on predicted secondary structures (Figure 3.8), the 5'-strand cleavage sites of both pre-let-7e and 7f-1 differ from those of other pre-let-7 substrates in that they are farther away from a loop/bulge structure, and this was previously shown to induce imprecise cleavage events (18). The presence of TRBP did not change either the nature or the proportions of the detected let-7-5p products, demonstrating that for the pre-let-7 family, TRBP does not contribute to the production of isomiRs by Dicer.

Next, we investigated the effect of TRBP on the relative cleavage efficiency of pre-let-7 cleavage. We monitored the cleavage of pre-let-7 after a 20-min incubation with Dicer at varying pre-let-7 and TRBP concentrations (Figure 3.2 and Figure 3.10). For the three pre-let-7



substrates investigated, we observed that TRBP activates Dicer cleavage at low substrate concentration ( $< 125$  nM; Figure 3.2b). This TRBP activation reaches its maximum at the lowest pre-let-7 concentration investigated (5 nM), with the relative cleavage efficiency increased by 16 folds for pre-let-7g, 11 folds for pre-let-7a-1 and 4 folds for pre-let-7c. The activation by TRBP is attenuated as the substrate concentration is increased to  $\sim 125$  nM pre-let-7. Surprisingly, as the substrate concentration is further increased, TRBP becomes an inhibitor of the cleavage reaction, and a maximum inhibition with a 4-fold decrease in relative cleavage efficiency is observed for all pre-let-7 tested at the highest concentration of substrate tested (3.1  $\mu$ M). The critical substrate concentration at which TRBP activation switches to inhibition is between 125 and 625 nM for pre-let-7g, around 625 nM for pre-let-7a-1 and around 125 nM for pre-let-7c. These results indicate that both the apparent  $k_{cat}$  and  $K_M$  values for the cleavage reaction by Dicer are modulated by TRBP according to both the concentration and type of pre-miRNA. At low substrate concentration, the cleavage efficiency reaches a maximum at a TRBP concentration of 80-160 nM, whereas at higher substrate concentrations, the cleavage efficiency reaches a minimum at the highest TRBP concentration tested (1.0  $\mu$ M). Overall, these results show that TRBP is a dual enzyme modulator, which functions mainly as an activator of Dicer cleavage at low pre-let-7 concentrations, but can also function as an inhibitor at higher pre-let-7 concentrations.



**Figure 3.2. Effect of TRBP on Dicer cleavage of pre-let-7 RNAs. (a)** Domain architecture of TRBP and the T12 and T3 domain constructs used in this study. Domains were positioned from InterPro and published data (33, 42). **(b)** Histograms of the relative Dicer cleavage efficiency

of three pre-let-7 substrates on a  $\log_2$  scale. Each bar represents data obtained at a different TRBP concentration and each bar cluster gathers data at different concentrations of pre-miRNA. Note that data for 3125 nM pre-let-7a-1 at 1000 nM TRBP and for 5 nM pre-let-7c at 160 nM TRBP are not available.

#### **3.5.4. No evidence for a pre-steady-state burst**

In RNase III enzymes, a pre-steady burst is expected if product release is the rate-limiting step (39, 40). To evaluate the possibility for such a pre-steady state burst with Dicer, multiple-turnover time course experiments were performed under steady-state conditions at a single pre-let-7g substrate concentration (5  $\mu\text{M}$  for Dicer and 0.25  $\mu\text{M}$  for Dicer-TRBP), but with different enzyme concentrations ( $[E]$ , Figure 3.11). As expected, the initial cleavage rate ( $v_0$ ) determined from the slope of these time courses increase proportionally with  $[E]$  (Figure 3.11). In contrast, the amplitudes of the potential burst ( $\pi$ ) taken as the y-intercept from the linear fit appears to be independent of  $[E]$ . In the absence of TRBP,  $\pi$  values of 1 nM, 9 nM and 3 nM were obtained at Dicer concentrations of 6 nM, 12 nM and 18 nM, respectively (Figure 3.11a), whereas in the presence of TRBP,  $\pi$  values of 0.4 nM, -0.2 nM and -0.2 nM were obtained at Dicer concentration of 0.75 nM, 1.5 nM and 2.25 nM, respectively (Figure 3.11c). The lack of a correlation between  $\pi$  and  $[E]$ , indicates that there is no pre-steady-state burst, and, thus, that the rate-limiting step under multiple-turnover conditions is likely not product release, either in the absence and presence of TRBP. However, given that the observed  $\pi$  values show substantial deviation from the expected zero value, further investigations would be needed to validate these conclusions.

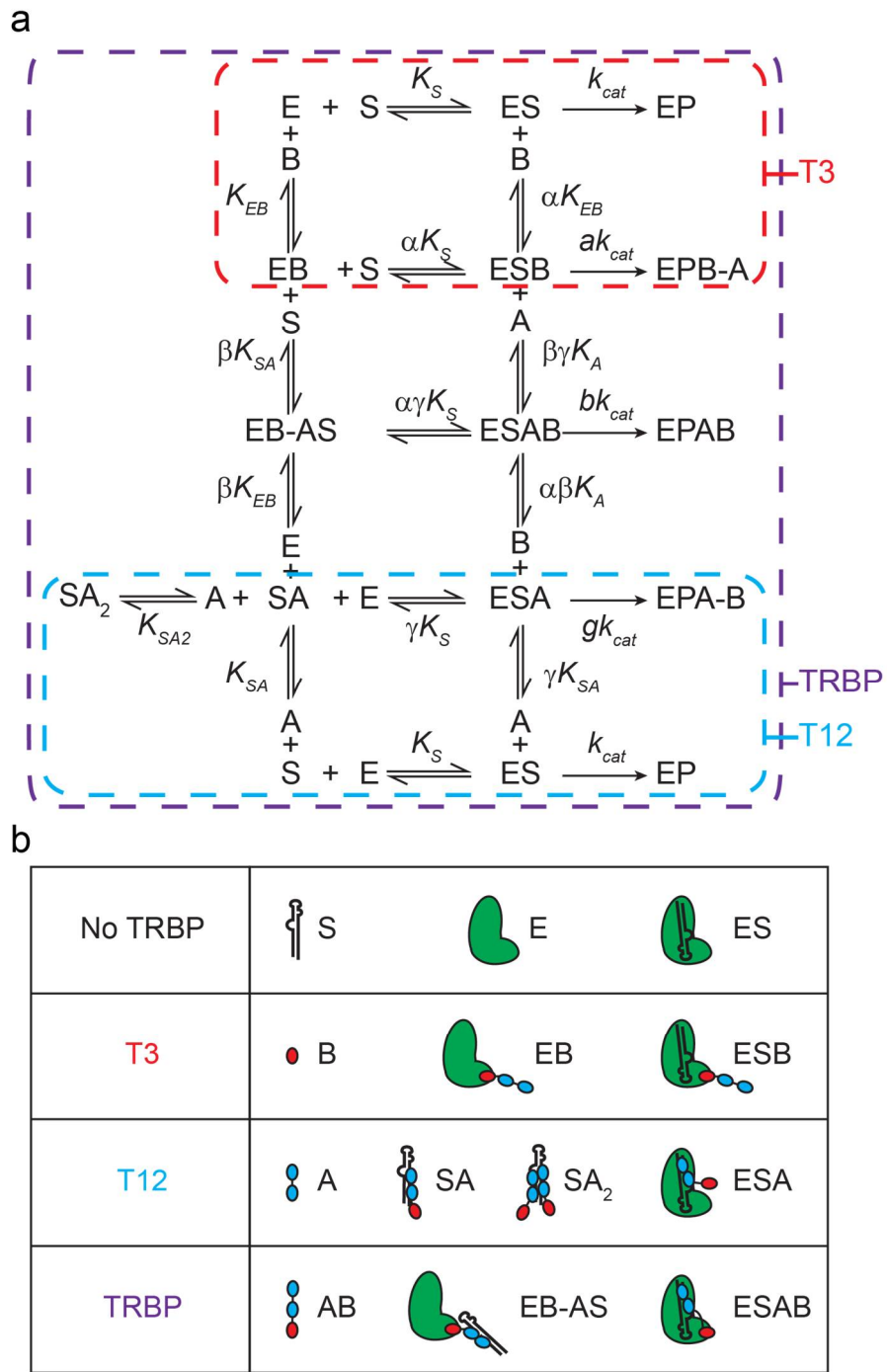
#### **3.5.5. Building a kinetic model for pre-miRNA cleavage by Dicer in the presence of TRBP**

To better understand the dual role of TRBP as both an activator and an inhibitor of the cleavage activity of pre-let-7 by Dicer, we built a kinetic model that incorporates the binding equilibria and catalysis steps of the complexes that are known to form between Dicer, TRBP and a given pre-miRNA substrate (Figure 3.3). The unique feature of TRBP as a modifier of enzymatic activity is that it can bind both the enzyme (Dicer) and its substrate (pre-miRNA).

Fortunately, these two binding activities are associated with different domains of TRBP. The region encompassing the first two dsRBD domains (T12) is known to interact with RNA (21, 33, 41), whereas the C-terminal dsRBD (T3) only interacts with proteins (23, 24, 41), such as Dicer (Figure 3.2a). To simplify the formalism and subsequent experimental investigations, TRBP is considered here as two separate but interconnected entities (A: T12 and B: T3), one for each of its two binding activities. This allowed us to individually define two subsections of the kinetic model that relate to either T12 or T3 and can be treated separately.

The T12 submodel (Figure 3.3a; blue box) takes into account that T12 ("A" in the model) binds the pre-miRNA to form the SA complex, but does not directly bind to Dicer (no direct EA interaction). Although high-resolution structures of such an SA complex are not available, the NMR structure of a related T12/dsRNA complex shows that only one face of the RNA double helix interacts with TRBP (33), leaving the other face available for Dicer binding and cleavage. Thus, the SA complex should interact with Dicer to form a catalytically-competent ESA complex. On the other hand, there is also evidence that multiple copies of TRBP can bind to a single pre-miRNA (21, 41). Thus, binding of a second molecule of T12 to a pre-miRNA substrate to form an SA<sub>2</sub> complex would likely interfere with the binding of Dicer and potentially explain the inhibition that we observed at high TRBP concentration (Figure 3.2b). For simplicity, we did not consider the formation of higher order T12/pre-miRNA complexes bound to more than two copies of T12. Also, we did not consider the possible oligomerization of T12 in the absence of RNA, since TRBP homodimerization occurs at protein concentrations that are much higher than those used in this study ( $K_d = 54 \mu\text{M}$ ) (21).

The T3 submodel (Figure 3.3a; red box) takes into account that T3 ("B" in the model) does not interact with RNA (no direct BS interaction) but binds the helicase domain of Dicer with a nanomolar  $K_d$  (23) to form the EB complex. Formation of the EB complex should be compatible with subsequent substrate binding and cleavage. The T3 domain of TRBP does not appear to form homodimers by itself (42), thus the dimerization equilibrium of T3 was not incorporated in our kinetic model.



**Figure 3.3. Steady-state kinetic model of TRBP modification on Dicer cleavage activity.** (a) Modifier kinetic model built using available data (see text). Sections of the kinetic model can be isolated by using the TRBP fragments T12 (cyan contour) and T3 (red contour). (b) Pictogram representation of the components of the kinetic model shown in (a).

The complete kinetic model of TRBP incorporates the T12 and T13 subsections, as well as adding a connecting part (Figure 3.3a; purple box), in which TRBP can simultaneously bind Dicer and its pre-miRNA substrate to form an active enzyme complex (ESAB). Altogether, this working model has six kinetic parameters ( $a$ ,  $\alpha$ ,  $b$ ,  $\beta$ ,  $g$  and  $\gamma$ ) and three binding parameters ( $K_{EB}$ ,  $K_{SA}$  and  $K_{SA2}$ ) that are unknowns. The strategy to resolve the unknowns is to first fit the T12 and T3 submodels and then fit the global TRBP kinetic model to combinations of binding and steady-state kinetics data using pre-let-7g as the substrate. This strategy makes it possible to define the unknowns by fitting a maximum of 3 parameters at once with less risk of overfitting the data.

### 3.5.6. Derivation of the kinetic equations

The expressions for the initial velocities of the kinetic model were derived from the inspection of equilibria between enzyme species for which we assumed that the system is under rapid equilibrium (43). The information behind the derivation of these types of model can be found in several enzymology textbooks (44, 45). In the equation for the kinetic model with TRBP,

$$\frac{v_0}{[E]_t} = \frac{k_{cat}([ES] + a[ESB] + \gamma[ESA] + b[ESAB])}{[E] + [ES] + [ESB] + [ESA] + [ESAB] + [EB] + [EB-AS]} \quad (1)$$

each term can be expressed in terms of  $[E]$  from their respective equilibrium.

$$[ES] = \frac{[E][S]}{K_S}$$

$$[ESA] = \frac{[E][S][A]}{\gamma K_S K_{SA}}$$

$$[EB] = \frac{[E][B]}{K_{EB}}$$

$$[ESB] = \frac{[E][S][B]}{\alpha K_S K_{EB}}$$

$$[EB-AS] = \frac{[E][S][B][A]}{\beta K_{SA} K_{EB}}$$

$$[EBAS] = \frac{[E][S][B][A]}{\alpha \gamma \beta K_S K_{SA} K_{EB}}$$

Hence, equation (1) becomes:

$$\frac{v_0}{[E]_t} = \frac{k_{cat} \left( \frac{[S]}{K_S} + a \frac{[S][B]}{\alpha K_S K_{EB}} + \gamma \frac{[S][A]}{\gamma K_S K_{SA}} + b \frac{[S][B][A]}{\alpha \gamma \beta K_S K_{SA} K_{EB}} \right)}{1 + \frac{[S]}{K_S} + \frac{[S][B]}{\alpha K_S K_{EB}} + \frac{[S][A]}{\gamma K_S K_{SA}} + \frac{[S][B][A]}{\alpha \gamma \beta K_S K_{SA} K_{EB}} + \frac{[B]}{K_{EB}} + \frac{[S][B][A]}{\beta K_{SA} K_{EB}}} \quad (2)$$

Similarly, the equations of the kinetic submodels for T12 and T3 are given by equations (3) and (4), respectively.

$$\frac{v_0}{[E]_t} = \frac{k_{cat} \left( \frac{[S]}{K_S} + \gamma \frac{[S][A]}{\gamma K_S K_{SA}} \right)}{1 + \frac{[S]}{K_S} + \frac{[S][A]}{\gamma K_S K_{SA}}} \quad (3)$$

$$\frac{v_0}{[E]_t} = \frac{k_{cat} \left( \frac{[S]}{K_S} + a \frac{[S][B]}{\alpha K_S K_{EB}} \right)}{1 + \frac{[S]}{K_S} + \frac{[S][B]}{\alpha K_S K_{EB}} + \frac{[B]}{K_{EB}}} \quad (4)$$

In the case of the T12 submodel and the TRBP model, [S] and [A] are not equal to [S]<sub>t</sub> and [A]<sub>t</sub> since they interact together and are not negligible with respect to one another. By applying the law of mass conservation for [A] and [S], we get,

$$[S]_t = [S] + [S]_b$$

$$[S]_t = [S] + [AS] + [A_2S] \quad (5)$$

$$[A]_t = [A] + [A]_b$$

$$[A]_t = [A] + [AS] + 2[A_2S] \quad (6)$$

where,

$$[AS] = \frac{[A][S]}{K_{SA}}$$

$$[A_2S] = \frac{[A]^2[S]}{K_{SA}K_{SA2}}$$

By substitution and rearrangement of equation (5), [S] can be expressed as,

$$[S] = \frac{[S]_t}{1 + \frac{[A]}{K_{SA}} + \frac{[A]^2}{K_{SA}K_{SA2}}} \quad (7)$$

Then, by substituting of equation (7) into equation (6), we get

$$[A]_t = [A] + \frac{[A]}{K_0} \frac{[S]_t}{1 + \frac{[A]}{K_0} + \frac{[A]^2}{K_0 K_1}} + 2 \frac{[A]^2}{K_0 K_1} \frac{[S]_t}{1 + \frac{[A]}{K_0} + \frac{[A]^2}{K_0 K_1}} \quad (8)$$

and then by rearranging equation (8), we can express [A] in the form of a third-degree polynomial,

$$0 = [A]^3 + a[A]^2 + b[A] - c$$

where,

$$a = K_{SA2} - [A]_t + 2[S]_t$$

$$b = K_{SA}K_{SA2} - [A]_tK_{SA2} + [S]_tK_{SA2}$$

$$c = -[A]_tK_{SA}K_{SA2}$$

The roots of the polynomial can be found through a trigonometric solution as previously described (45, 46). The discriminant of the polynomial confirmed that the system possesses three distinct real roots, and since [A] is always a positive number, only one root is relevant and [A] can be described by:

$$[A] = \frac{2}{3}\sqrt{a^2 - 3b} \cos\left(\frac{1}{3}\cos^{-1}\left(\frac{-2a^3 + 9ab - 27c}{2\sqrt{(a^2 - 3b)^3}}\right)\right) - \frac{a}{3} \quad (9)$$

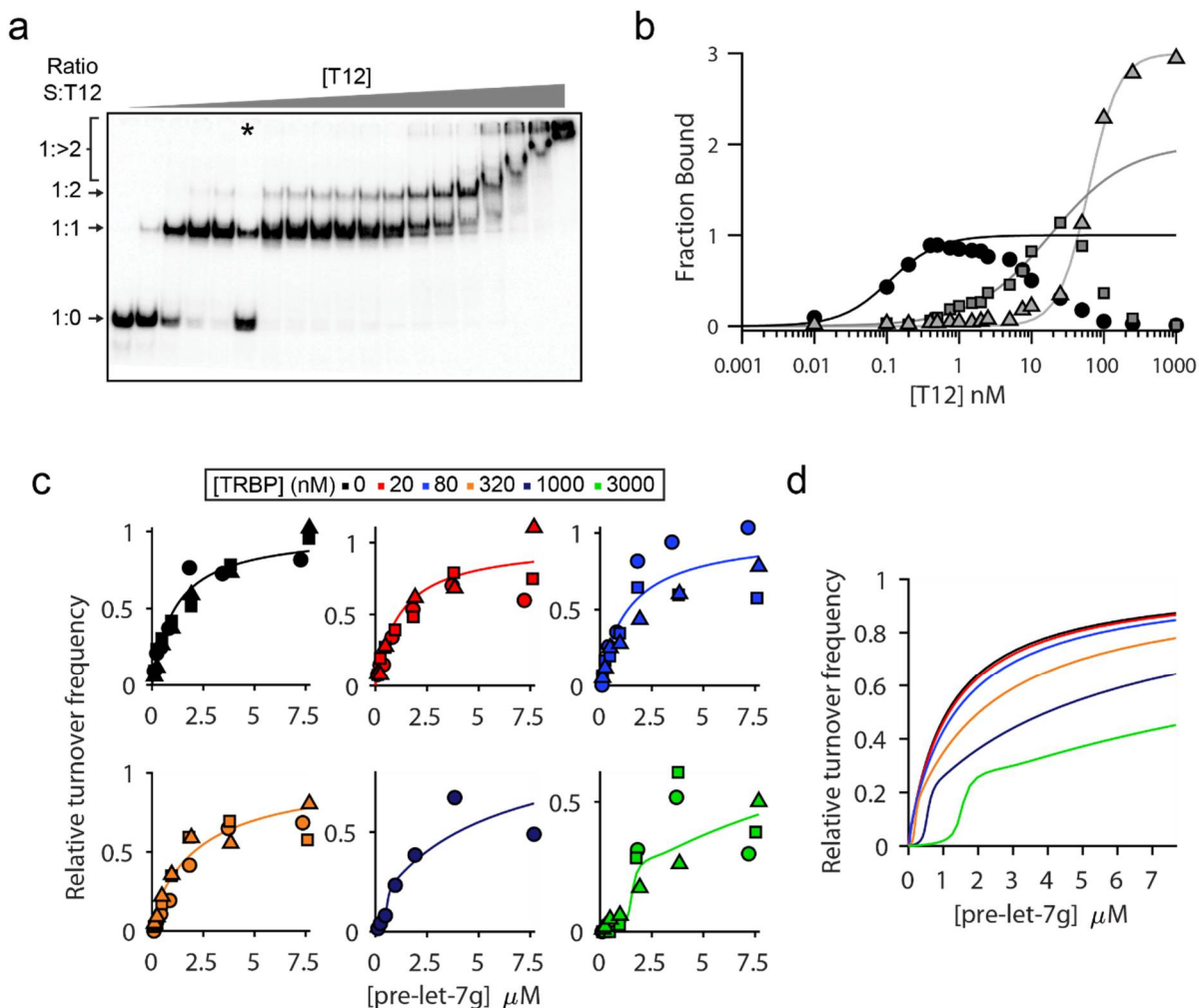
### 3.5.7. Fitting the T12 subsection of the kinetic model

To determine the values of the  $g$  and  $\gamma$  parameters (Figure 3.3a), thermodynamic and kinetic experiments were performed with T12. First, the values of  $K_{SA}$  and  $K_{SA2}$  were evaluated by EMSA. The denaturing gel of  $^{32}\text{P}$ -labeled pre-let-7g incubated with increasing concentration of T12 shows the formation of distinct complexes with different stoichiometries (Figure 3.4a). The fraction of bound RNA corresponding to the first, second and third band shifts were fitted to the Hill equation with  $B_{max}$  values of 1, 2 and 3, respectively, to derive the following dissociation constants:  $K_{SA} = 107$  pM,  $K_{SA2} = 15.8$  nM and  $K_{SA3} = 74$  nM (Figure 3.4b and Tableau I), which are similar to those previously determined by isothermal titration calorimetry (ITC) between TRBP and small interfering RNAs ( $K_{SA}$  of 0.24 nM and  $K_{SA2}$  of 13.3 nM) (21).

Steady-state kinetic experiments of pre-let-7g cleavage were performed at five different concentrations of T12 (Figure 3.4c). For each concentration, we derived [A] and [S] from the experimental values of  $K_{SA}$  and  $K_{SA2}$  using equations (9) and then (7). With these values of [A] and [S] and all five sets of data (Figure 3.4c), we could then fit equation (3) to derive the values of  $g$  and  $\gamma$  (Tableau II). The resulting fit is overall fairly good ( $R^2 = 0.89$ ), except at the highest



concentration of T12 (3  $\mu\text{M}$ ). This is likely due to the large experimental errors that are associated with quantification of low-intensity bands on autoradiographs in conditions of high [S] that lead to a very small percentage of substrate being cleaved and where a low level of substrate degradation can be detected. The overlay of fitted curves (Figure 3.4d) shows that increasing T12 concentration leads to small reductions of  $k_{\text{cat}}$  and  $K_M$  values. The  $g$  of 0.31 and  $\gamma$  of 0.20 mean that 3-fold and 5-fold maximal reduction of the apparent  $k_{\text{cat}}$  and  $K_M$  are expected, respectively, at saturating substrate concentrations and high T12 concentration (3  $\mu\text{M}$ ). Moreover, a severe inhibition of the cleavage rate is detected at [T12] = 3  $\mu\text{M}$  and low pre-let-7g concentrations (< 1  $\mu\text{M}$ ), due to formation of the SA<sub>2</sub> complex that block productive Dicer binding, in agreement with previous observations that excess T12 inhibits cleavage by Dicer (33). Overall, these data are consistent with T12 being an inhibitor of the cleavage activity of Dicer due to sequestration of the substrate in an inactive ternary complex.

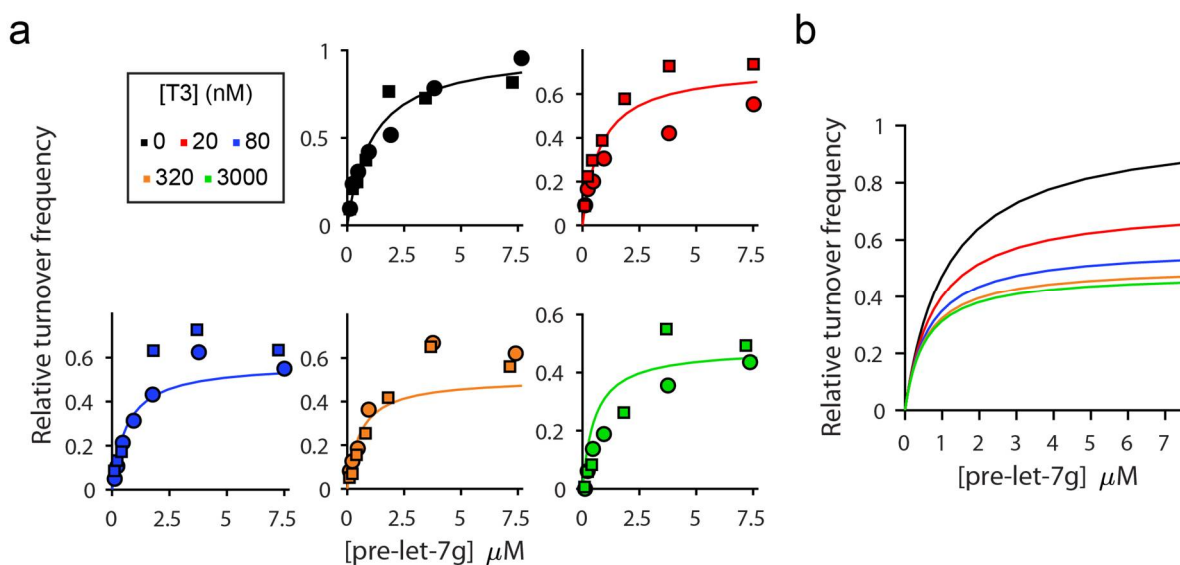


**Figure 3.4. Thermodynamic and steady-state kinetic experiments with T12.** (a) EMSA performed with 10 pM of 5'-[<sup>32</sup>P]-labeled pre-let-7g and increasing concentrations of T12 (0, 0.01, 0.2, 0.4, 0.1\*, 0.5, 0.75, 1, 1.5, 2, 2.5, 5, 7.5, 10, 25, 50, 100, 250 and 1000 nM). The asterisk indicates that the 0.1-nM sample was not loaded in the intended order of increasing concentration, but should have been loaded right after the second lane. (b) Typical plot of the fraction of T12-bound pre-let-7g against the T12 concentration. The fraction of the T12-bound pre-let-7g contained in the first (black circles), second (dark grey squares) and higher (light grey triangles) band shifts were independently fitted to the Hill equation. The following binding constants were obtained from two independent experiments:  $K_d = 107 \pm 2$  pM and  $n = 1.50 \pm 0.07$  (first shift),  $K_d = 15.8 \pm 0.2$  nM and  $n = 0.95 \pm 0.05$  (second shift) and  $K_d = 74 \pm 2$  nM and  $n = 1.73 \pm 0.03$  (higher shifts). For the fit,  $B_{max}$  values were set to one, two and three, respectively. (c) Global fit of steady-state kinetic data to the T12 section of the kinetic model. Plots of relative turnover frequency (turnover frequency/ $k_{cat}$ ) for pre-let-7g cleavage by Dicer against total substrate concentration are provided for five different concentrations of T12 (0, 20, 80, 320 and 3000 nM). The following values were fixed:  $K_S = 1113$  nM,  $K_{SA} = 107$  pM and  $K_{SA2} = 15.8$  nM. From the fit of two independent experiments represented individually with squares and circles, the values of the following two unknown were obtained:  $g = 0.31 \pm 0.04$  and  $\gamma = 0.20 \pm 0.07$ . (d) Overlay of the fitted curves shown in (c).

### 3.5.8. Fitting the T3 subsection of the kinetic model

To determine the values of  $a$ ,  $\alpha$  and  $K_{EB}$  (Figure 3.3a) the C-terminal dsRBD of TRBP (T3) was used. First, we confirmed that T3 does not interact with pre-let-7g by EMSA (21), and we did not observe any binding at  $[T3] = 1 \mu\text{M}$  (Figure 3.12T). Next, we proceeded with steady-state kinetic experiments performed in the same manner as for T12 (Figure 3.5). The kinetic model is much simpler for T3 since no correction needs to be applied to the kinetic equations given that  $T3 \gg [E]$  and  $[S] \gg [E]$  and T3 does not interact with  $[S]$ . The allosteric model used for T3 (Figure 3.3a; blue box) is expected to be one of hyperbolic modification because both forms of the enzyme, whether bound or unbound to T3, should bind and cleave the substrate (43). When fitting all five sets of data collected at different T3 concentrations (Figure 3.5a) to equation (4), the values of  $a = 0.62$  and  $\alpha = 1.1$  were obtained (Tableau II). Given these values, T3 fits the characteristics of a hyperbolic mixed, balanced inhibitor [ $0 < a < 1$  and  $\alpha \approx 1$ ; (43)],

for which the apparent  $K_M$  is independent of the T3 concentration and the apparent  $k_{cat}$  diminishes as T3 is increased until it reaches its minimum, defined as  $a \times k_{cat}$  (Figure 3.5b). The fitted value of  $K_{EB}$  (31 nM) falls within the low nanomolar range, in agreement with previous studies, where a value of 4.3 nM was obtained by ITC (23). Determination of  $K_{EB}$  by an independent method using the proteins purified for the current study is in progress as it should increase the confidence on the fitted parameters. Together, these results show that T3 is a relatively weak inhibitor that interacts with Dicer to reduce its cleavage activity.



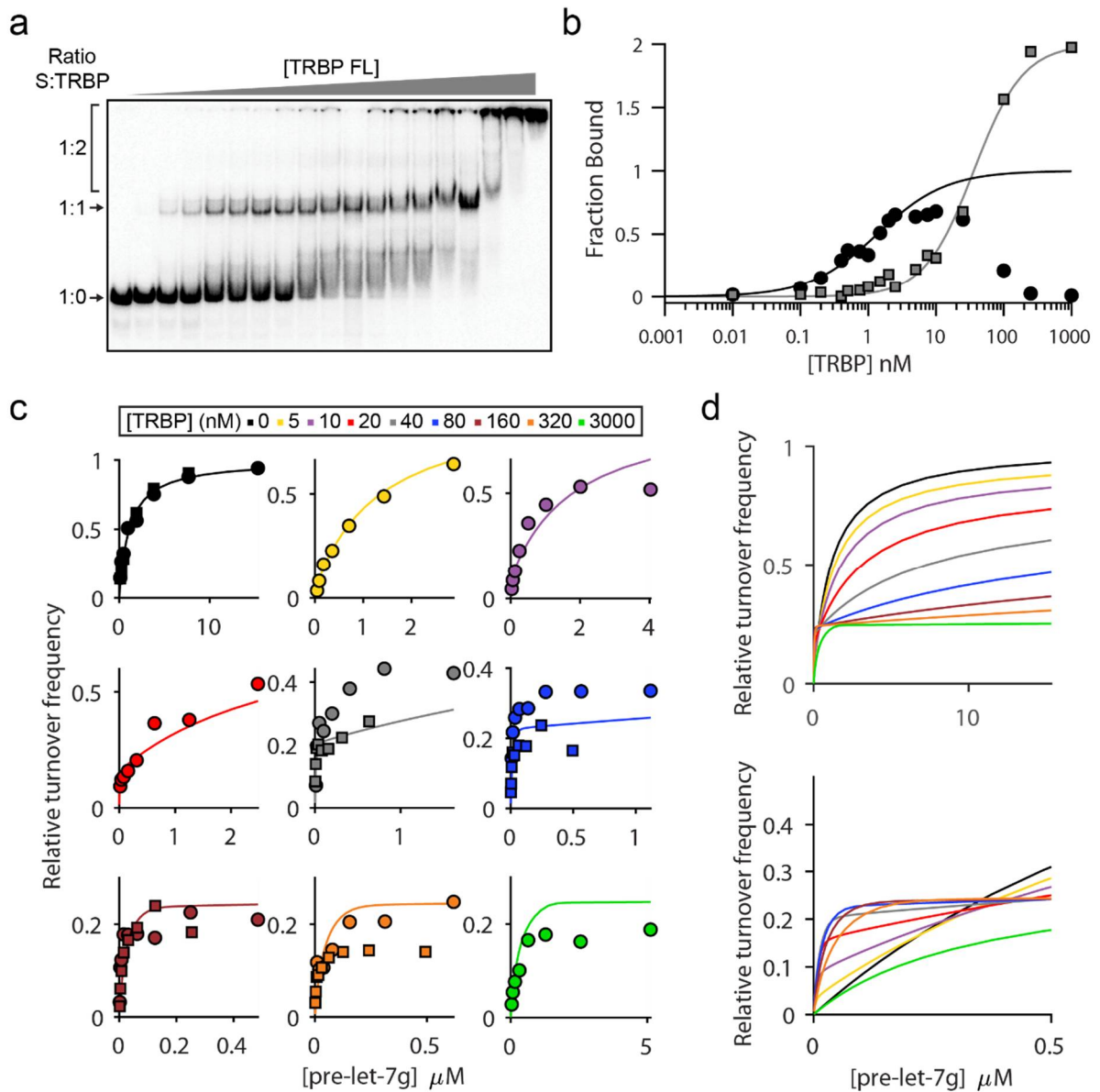
**Figure 3.5. Steady-state kinetic experiments with T3. (a)** Global fit of steady-state kinetic data to the T3 section of the kinetic model. Plots of relative turnover frequency (turnover frequency/ $k_{cat}$ ) for pre-let-7g cleavage by Dicer against total substrate concentration are provided for five different concentrations of T3 (0, 20, 80, 320 and 3000 nM). The value of  $K_S$  was fixed to 1.11  $\mu\text{M}$ . From the fit of two independent experiments represented individually with squares and circles, the values of the following three unknown were obtained:  $K_{EB} = 31 \pm 10$ ,  $a = 0.62 \pm 0.08$  and  $\alpha = 1.1 \pm 0.2$ . **(b)** Overlay of the fitted curves shown in (a).

### 3.5.9. Fitting the overall kinetic model

With several key values derived from fitting the T12 and T3 submodels, it became possible to fit the overall kinetic model using binding and kinetic data with full-length TRBP (Figure 3.6). First, the values of  $K_{SA}$  and  $K_{SA2}$  were reevaluated by EMSA using full-length

TRBP (Figure 3.6a). As with T12, increasing concentrations of TRBP lead to the formation of distinct complexes with different stoichiometries (Figure 3.6a). However, only two band shifts were observed with TRBP instead of the three seen with T12. The most shifted band is found very close to the loading well, and although this suggests formation of high order aggregates, it was considered here as representing a 1:2 RNA:TRBP complex. The fraction of bound RNA corresponding to the first and second shifts were fitted to the Hill equation to derive the following dissociation constants:  $K_{SA} = 1.4$  nM and  $K_{SA2} = 36$  nM (Figure 3.6b and Tableau I), which are respectively 13-fold and 2-fold higher the equivalent parameters obtained for T12 (Tableau II). The lower  $K_{SA}$  value with T12 compared to TRBP may result from an auto-inhibitory mechanism mediated by interactions of T12 with other regions of the protein in the context of the full-length TRBP.

Steady-state kinetic experiments of pre-let-7g cleavage were performed at nine different concentrations of TRBP (Figure 3.6cd). These data were fitted to the overall kinetic model to derive the values of  $\beta = 4.1$  and  $b = 61$  (Tableau II). Given that the values of  $\alpha$  (1.1),  $\beta$  (4.1) and  $\gamma$  (0.20) are within 5-fold of the unity, this indicates that the equilibrium interactions involving any two partners among Dicer, TRBP and pre-let-7g are not substantially affected by the presence of the other partner in terms of kinetic equilibrium. On the other hand, the relative turnover frequency of the reaction is drastically increased at low substrate concentrations, whereas it is decreased at high concentration. This defines a critical region of substrate concentration (0.3-0.4  $\mu$ M) where all the curves cross (except for the curve at 3  $\mu$ M TRBP) and where TRBP has a minimal effect on the activity of Dicer (Figure 3.6d). At high TRBP concentrations, a decrease in activity is observed for the same reason as for T12 where the binding of more than one TRBP molecule to pre-let-7g inhibits the binding and processing by Dicer. The discrepancies between the fitted curves and the experimental data at higher TRBP concentrations are the result of performing a global fit on sets of experiments collected independently on different days. In particular, the curves at 320 nM and 3000 nM TRBP would individually better fit a value of  $b$  that is lower ( $b \sim 50$ ) than the result of the global fit, but this would not change the global conclusions drawn from the fitted model. The fitted kinetic model clearly confirms our previous observation (Figure 3.2b) that TRBP is a dual modifier of the cleavage activity of Dicer; it can act as a strong activator at low substrate concentration, and as a relatively weak inhibitor at the highest concentrations investigated.



**Figure 3.6. Thermodynamic and steady-state kinetic experiments with TRBP.** (a) EMSA performed with 10 pM of 5'-[<sup>32</sup>P]-labeled pre-let-7g and increasing concentrations of TRBP (0, 0.01, 0.1, 0.2, 0.4, 0.5, 0.75, 1, 1.5, 2, 2.5, 5, 7.5, 10, 25, 50, 100, 250 and 1000 nM). (b) Typical plot of the fraction of TRBP-bound pre-let-7g against the total TRBP concentration. The fraction of the TRBP-bound pre-let-7g contained in the first (black circles) and higher (grey squares) band shifts were independently fitted to the Hill equation. The following thermodynamic values were obtained from a single experiment:  $K_d = 1.4$  nM and  $N = 0.9$  (first

shift) and  $K_d = 36$  nM and  $N = 1.2$  (higher shifts). For the fit,  $B_{max}$  values were set to one and two, respectively. **(c)** Global fit of steady-state kinetic data to the TRBP kinetic model. Plots of relative turnover frequency (turnover frequency/ $k_{cat}$ ) for pre-let-7g cleavage by Dicer against total substrate concentration are provided for different concentrations of TRBP (0, 5, 10, 20, 40, 80, 160, 320 and 3000 nM). The following values were fixed:  $K_S = 1113$  nM,  $K_{SA} = 1.1$  nM,  $K_{SA2} = 47$  nM,  $K_{EB} = 31$  nM,  $\alpha = 1.1$ ,  $a = 0.62$ ,  $\gamma = 0.20$  and  $g = 0.31$ . From the fit of two independent experiments represented individually with squares and circles, the values of the following two unknown were obtained:  $b = 61 \pm 3$  and  $\beta = 4.4 \pm 0.7$ . **(d)** Overlay of fitted curves shown in (c). The top panel represents the whole range of tested substrate concentrations (0-15  $\mu$ M), whereas the lower panel shows only the 0-0.5  $\mu$ M range.

Tableau II. Binding and kinetic constants of the effect of TRBP on Dicer cleavage of pre-let-7g.<sup>a</sup>

	$K_S^{a,b}$	$K_{SA}^b$	$K_{SA2}^b$	$K_{EB}^b$	$\alpha$	a	$\beta$	b	$\gamma$	g
T12		<b><math>107 \pm 2 \times 10^{-3}</math></b>	<b><math>15.8 \pm 0.2</math></b>	-	-	-	-	-	<b><math>0.20 \pm 0.05</math></b>	<b><math>0.31 \pm 0.03</math></b>
T3	1113	-	-	<b><math>31 \pm 10</math></b>	<b><math>1.1 \pm 0.2</math></b>	<b><math>0.62 \pm 0.08</math></b>	-	-	-	-
TRBP		<b><math>1.1 \pm 0.2</math></b>	<b><math>47 \pm 13</math></b>	31	1.1	0.62	<b><math>4.4 \pm 0.8</math></b>	<b><math>61 \pm 2</math></b>	0.21	0.33

<sup>a</sup>Except for  $K_S$  taken from Tableau I, the binding and kinetic constants were derived from the data shown in Figure 3.4, Figure 3.5 and Figure 3.6.

<sup>b</sup>Values are expressed in nM.

## 3.6. Discussion

### 3.6.1. Similar thermodynamic and kinetic properties of pre-let-7 substrates for Dicer

Although the miRNAs of the let-7 family share the same seed sequence, their precursors vary considerably in length (67 to 79 nts) and in terms of their secondary structure (Figure 3.8). In addition, these precursors belong to different groups, based on whether or not they can be mono-uridylylated at the 3'-end by TUT4 (48). Nevertheless, we found that the  $K_d$  values for all pre-let-7 fall within a narrow window (5-60 nM) and it does not appear that group (I versus II) or size of pre-miRNAs contribute to the binding affinities (Figure 3.8). Moreover, we could not identify any specific sequence or secondary structure feature that would explain the observed changes in  $K_d$  values. The Hill coefficients derived from the binding curves range between 0.9 and 2.8 (Tableau I). Although a Hill coefficient  $> 1$  generally indicates positive *allosteric* cooperativity, which may occur when more than one protein binds to a single RNA molecule, Dicer has been previously shown to bind its pre-miRNA substrate with a 1:1 ratio. The Hill coefficients  $> 1$  observed here likely reflect *configurational* cooperativity, also termed chelate cooperativity, that result from the coupling of interactions in self-assembled systems (49, 50). Configurational cooperativity is compatible with current knowledge of pre-miRNA/Dicer interactions, which involves a multidentate RNA ligand (the 5'-end, the cleavage sites and the terminal loop) and a multisite protein receptor (multiple RNA-binding domains). In addition, complex formation may be coupled to partial unfolding of the pre-miRNA terminal loop and/or changes in Dicer conformation upon RNA binding (13, 14, 51).

By examining the kinetic values for the different pre-let-7s, we found that Dicer does not discriminate between these different RNA substrates. The  $k_{cat}/K_M$  values of all pre-let-7 substrates are essentially identical ( $0.45\text{-}2.0 \times 10^5 \text{ M}^{-1} \text{ sec}^{-1}$ ) and much below the diffusion limit ( $\sim 10^7 \text{ M}^{-1} \text{ sec}^{-1}$ ). Taken together with the very low  $k_{cat}$  values of  $0.02\text{-}0.35 \text{ sec}^{-1}$ , this suggests that the rate limiting step for the cleavage of pre-miRNA by Dicer is likely not the actual cleavage step(s). The observed kinetic values are more consistent with the rate-limiting step being product release as for other RNase III enzymes (39, 52) or a conformational change that occurs before cleavage. Here, we provide experimental evidence that argues against product

release being the rate-limiting step both in the presence and absence of TRBP. Rather, our results are consistent with the current view that the conformational changes of an enzyme-substrate complex into a catalytically productive conformation is important for cleavage activity (12-14, 19, 29, 34, 53). Dicer forms a stable ES complex with its pre-miRNA substrate in a conformation that is incompatible with the cleavage reaction, as observed from recent cryo-EM structures of Dicer, where the PAZ and helicase domains respectively contact the 5'-end and terminal loop of the pre-miRNA substrate, while the RNase III active sites are located away from the pre-miRNA cleavage sites (13, 14). Cleavage occurs either through a slow conformational change towards an active form of the ES complex and/or via the dissociation and re-binding of the substrate. In previous studies, it was shown that a truncated Dicer lacking the helicase domain enhances the reaction velocity without significantly changing the  $K_d$  of ES formation (19). The helicase domain was defined as an auto-inhibitory domain that may restrict the accessibility of substrates for the active site (19), and such a mechanism was later shown to limit access of the active site to canonical pre-miRNA substrates and thereby raise the overall specificity of Dicer for those substrates (10, 16). The cost of such surveillance mechanism is low  $k_{cat}$  and high  $K_M$  values, making Dicer a particularly inefficient enzyme at substrate concentrations that are below the  $K_M$ , namely in the nanomolar range or below. Thus, it is not surprising that regulatory mechanisms have evolved to overcome the limitations of the free enzyme.

### **3.6.2. A dual role for TRBP in regulating pre-let-7 cleavage by Dicer**

One such regulatory mechanism is through the molecular partner of Dicer, TRBP, which can bind both the helicase domain of Dicer and the stem portion of the pre-miRNA substrate (21, 23, 33). Here, we confirmed that TRBP can be an activator of Dicer cleavage activity, but further defined TRBP as a dual modifier. This dual role of Dicer as an activator and an inhibitor of cleavage activity was observed for the three pre-let-7 substrates that were investigated and can likely be extrapolated to other pre-miRNA substrates. However, our results indicate that its specific modifier effect may vary from one pre-miRNA to the other as the maximum relative cleavage efficiency that we observed was different for pre-let-7g (16 fold), pre-let-7a-1 (11 fold) and pre-let-7c (4 fold). This lower value for pre-let-7c may result from the fact that its  $K_M$  is 4 times lower than pre-let-7g; at the low substrate concentrations tested its basal activity in the absence of TRBP is higher than for pre-let-7g and this could explain the limited activation effect

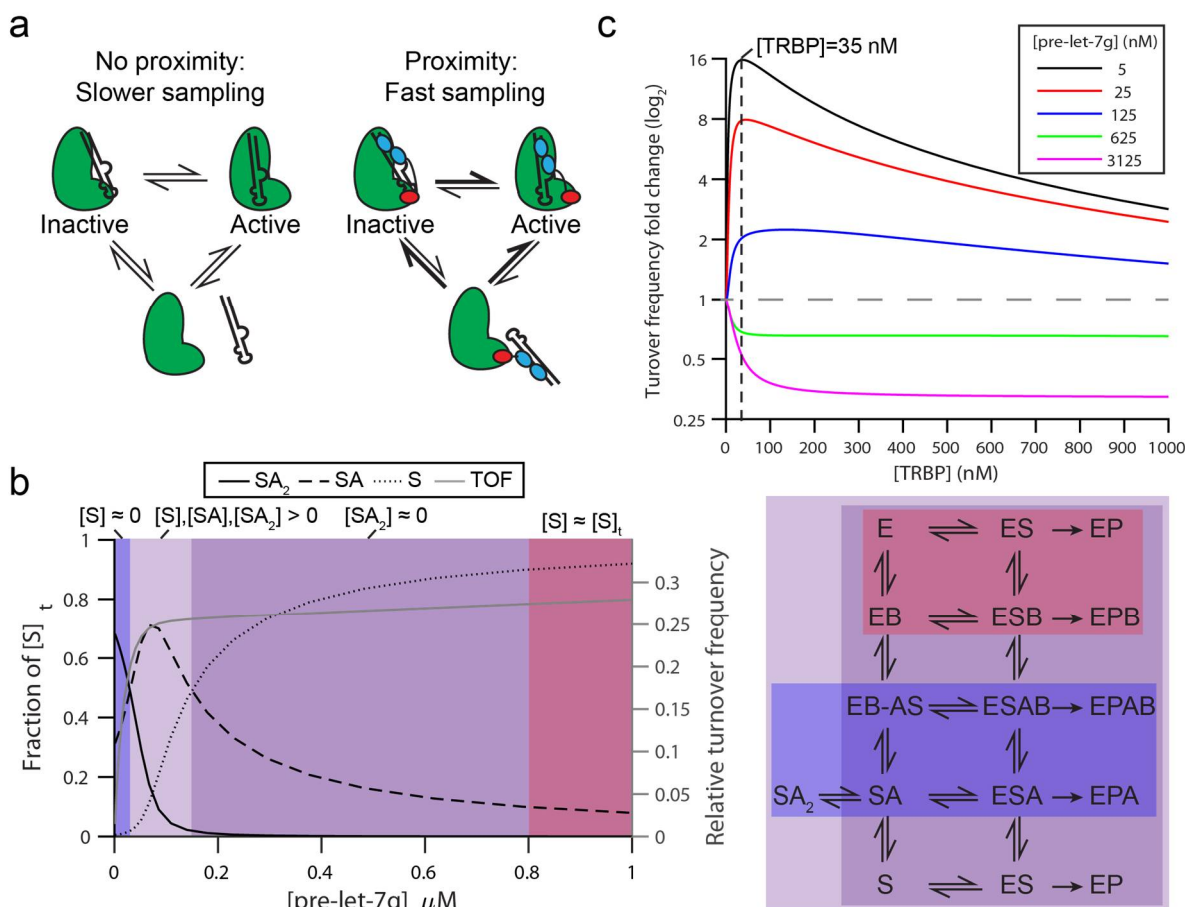


of TRBP. We show that the relative cleavage efficiency increases at low concentrations of TRBP ( $\leq 320$  nM) depending on substrate concentration. Under our experimental conditions, the critical substrate concentration at which TRBP changes from activator to inhibitor and where its presence has negligible effect on Dicer cleavage is between 125 nM and 625 nM, depending on the substrate.

The mechanism by which TRBP changes Dicer activity has then been characterized in more detail by fitting the steady-state kinetic model of TRBP modification. From this fit, the equilibrium modifier values of  $\alpha$ ,  $\beta$ ,  $\gamma$  are all relatively close to 1, indicating that the equilibrium for formation of the ES complex is not drastically affected by TRBP (43). In other words, the binding of TRBP to Dicer and its substrate does not substantially affect the interaction of Dicer with pre-let-7g.

A more pronounced effect for TRBP is observed for the apparent  $k_{cat}$  of the cleavage reaction, as revealed from analysis of the kinetic modifier values (a, b and g). The values of a and g show a slight inhibitory effect of TRBP, which could potentially result from conformational hindrance induced by TRBP binding to either the substrate or Dicer. Indeed, it was previously shown by cryo-EM studies that the binding of TRBP to the helicase domain of Dicer modifies its conformational equilibrium (14). This modification of Dicer structure may slow down the substrate docking into the active site. Moreover, it has been shown that the binding of TRBP only covers one face of the helix of the pre-miRNA, leaving the other face available for interaction with Dicer (14, 33). Again, this could slow down the positioning of the substrate into the active site by restricting active site accessibility. Moreover, if the RNA must slightly unfold to gain access to the active site, as previously suggested (13), this process would be slowed down by the interaction with the dsRBDs since these domains are known to stabilize RNA helices (54). Finally, the magnitude of the b value reveals that formation of the ternary complex ESBA leads to a drastic increase of the  $k_{cat}$  by 61-fold. Since the A (T12) and B (T3) regions of TRBP are linked together and are, on their own, inhibitors of Dicer processing, the most probable explanation for this rise in the reaction velocity is that TRBP mainly acts to bring the pre-miRNA in proximity of Dicer. In the case where the substrate needs to fully or partially dissociate from Dicer and reassociate to reach an active ES state, the substrate would remain attached to Dicer via TRBP and thus would be able to bind again much more rapidly than if the

substrate had diffused away (Figure 3.7a). Multiple partial dissociation/association cycles within a Dicer-TRBP-substrate ternary complex may promote conformational exchange to favor formation of an active ES state. In agreement with this model, the lifetime of exchange between Dicer and its substrate measured from single-molecule fluorescence was found to be significantly reduced in the presence of TRBP (29). Moreover, the stepwise formation of the inactive ES state and its transition to the active state may provide a double energy barrier to allow specific cleavage of pre-miRNA substrates while preventing cleavage of other cellular RNAs (29). Future structural studies and pre-steady-state kinetics studies should help provide more detailed insights into the conformational change that brings the ES complex in a catalytically-competent state and how binding of TRBP increases the  $k_{cat}$  of the cleavage reaction.



**Figure 3.7. Interpretation of the parametrized steady-state kinetic model of TRBP modification on Dicer cleavage activity. (a) Schematic representation of TRBP activation of**

Dicer at low substrate concentrations (see text). In absence of TRBP (left panel), binding and dissociation events occur slowly due to diffusion, increasing the time required to bind the substrate in an active conformation. When TRBP is bound to Dicer (right panel), the substrate can stay in the vicinity of Dicer for fast reassociation, accelerating sampling for an active conformation. **(b)** Simulated plot of turnover frequency (TOF) and fractions of different forms of substrates (S, SA, SA<sub>2</sub>) as a function of total pre-let-7g substrate concentration [S]<sub>t</sub> (left panel). For this simulation, the TRBP parameters in Tableau II were used together with a TRBP concentration of 80 nM. The plot illustrates four different phases of the kinetic model that depend on the total substrate concentration. In phase 1 (violet), the concentration of free substrate is negligible ([S] ≈ 0). In phase 2 (light pink), [S] > 0, [SA] > 0 and [SA<sub>2</sub>] > 0. In phase 3 (dark pink), [S] > 0, [SA] > 0 and [SA<sub>2</sub>] ≈ 0. In phase 4 (magenta), [S] = [S]<sub>T</sub>, [S] ≫ [SA] and [SA<sub>2</sub>] ≈ 0. These phases of the kinetic model can be represented by the main equilibria at play involving specific forms of the substrate (left panel), and sections of the kinetic model are highlighted using the color(s) of the phase(s) that they represent best. **(c)** Simulated plot of the TOF as a function of TRBP concentration. The relative TOF was calculated at different concentrations of pre-let-7g: 0 nM (grey dotted line), 5 nM (black), 25 nM (red), 125 nM (blue), 625 nM (green) and 3125 nM (purple). A TRBP concentration (35 nM) of maximal TOF was identified at low concentration of pre-let-7g (5 nM).

### 3.6.3. The four phases of the kinetic model for TRBP modification

A more detailed understanding of the kinetic model for TRBP modification can be obtained from the analysis of simulated plots derived from the parametrized kinetic model (Figure 3.7b). Several of these plots were drawn on the same graph to examine how the fractions of different forms of substrates that are not bound to Dicer (S, SA, SA<sub>2</sub>) vary along with the turnover frequency (TOF) as a function of total substrate concentration ([S]<sub>t</sub>) under conditions in which TRBP functions as an activator ([TRBP] = 80 nM). The resulting graph can be divided into four main phases defined with respect to the relative concentration of the various forms of the substrate (Figure 3.7c). It is important to note that the equilibrium between [E] and [EB] does not vary with substrate concentration and is constant throughout the four phases. The first phase at low [S]<sub>t</sub> represents a regime where all substrates are bound to form either SA or SA<sub>2</sub> complexes, and thus the concentration of free substrate is negligible ([S] ≈ 0). Under this regime,

the cleavage reaction progresses mainly via formation of the ESA and ESBA complexes, where ESBA contributes predominantly to the total velocity since  $b \gg g$ . This results in a sharp positive slope of the velocity dependence versus  $[S]_t$  in this phase of the graph. Notably, in the cases where  $[TRBP] \gg K_{SA_2}$  (not shown on the graph), the substrate would mostly exist in the  $SA_2$  form, the activation effect would not be observed and inhibition could occur. The second phase of the graph is observed when  $[S]_t \geq [A]_t$ , which result in  $[S] > 0$ . From that point on, a substantial part of the equilibrium will be pushed towards the formation of the ES and ESB complexes, reducing the contribution of ESBA to the reaction velocity and leading to a reduction of the slope of the relative TOF. The third phase is similar to the previous phase with the exception that  $SA_2$  stops contributing to the overall activity. The final phase is the phase where  $[S]_t$  is high enough so that it corresponds to the free  $[S]$  and that  $[A]$ ,  $[SA]$  and  $[SA_2]$  are negligible. The equilibrium of the reaction is then shifted towards  $[ES]$  and  $[ESB]$  where, if the concentration of TRBP is saturated, all the enzyme will be found as  $[EB]$  and the reaction will mainly proceed through  $[ESB]$ .

To better understand the effect of TRBP concentration on the cleavage activity of Dicer, we calculated how the relative TOF changes with TRBP concentration from the parametrized kinetic model (Figure 3.7c). At the lowest pre-let-7g concentration (5 nM), the model predicts an increase in apparent relative TOF of over 10-fold in the 9-140 nM range of TRBP concentrations, with a maximum 15-fold increase at 35 nM TRBP. In other words, a large range of TRBP concentrations in the nanomolar range is compatible with a substantial increase in catalytic efficiency at low substrate concentration, in agreement with our initial studies performed with different pre-let-7 substrates (Figure 3.2b). At the highest pre-let-7g concentration (3125 nM), the model predicts an asymptotic decrease in apparent relative TOF with increasing TRBP concentrations with a 3-fold maximum decrease at high TRBP concentrations. Interestingly, the overall trends in relative turnover frequencies derived from the parametrized kinetic model reproduce well the experimental results that we obtained with pre-let-7g (Figure 3.2 and Figure 3.6).

### 3.6.4. Dicer cleavage under physiological substrate concentrations

Central metabolic enzymes have  $K_M$  values that roughly match the cellular concentration of their substrate, and this means that slight variations in substrate concentration have a meaningful impact on the rate of reaction (55). However, this trait was shown to differ in secondary metabolic enzymes, where some enzymes would mostly work at saturating substrate concentration. Given the  $K_M$  values of Dicer for the pre-let-7 substrates (0.1-1  $\mu\text{M}$ ), it is of interest to investigate how these values relate to cellular concentrations of substrates. Although, it is not entirely clear how the concentration of pre-miRNA substrates varies across different cell types, cellular compartments and cellular conditions, we attempted to make an estimate using available values. It was previously established that pre-miRNA substrates average around 1000 copies/cell, although a large extreme of 1-2 copies to  $\sim 50\,000$  copies per cell have been observed for pre-miRNAs in human tissues (56-58). Given this average of 1000 copies/cell and the fact that mammalian cell volumes vary from around 100 fL(57) to 10 000 fL(59), we calculated that the average concentration of each pre-miRNA substrates would range between 0.2 nM and 20 nM. In comparison with the larger  $K_M$  values for the pre-let-7 substrates, this means that Dicer would be very inefficient on its own to process these substrates in a physiological context. Interestingly, TRBP would exhibit its maximal activation effect on those pre-miRNA substrates that average  $\sim 1000$  copies/cell. In contrast, for pre-miRNA substrates with higher copy numbers that could reach higher cellular concentrations of  $\sim 1\ \mu\text{M}$ , TRBP would have an inhibitory effect. Thus, TRBP is a dual modifier that promotes the cleavage activity of Dicer for poorly-abundant pre-miRNAs while dampening Dicer processing of highly-abundant pre-miRNAs. In other words, TRBP could function as an *equalizer* by reducing the effect of substrate concentration on cleavage efficiency.

### 3.6.5. The kinetic model resolves previous and apparently conflicting observations

Although previous studies reported apparently conflicting results regarding the role of TRBP as a Dicer regulator, it is now clear that TRBP is a dual modifier that modulates its regulatory role according to its own concentration as well as the type and concentration of pre-miRNA substrates. In previous studies highlighting the importance of TRBP in transferring the

miRNA from Dicer to Ago, divergent results were obtained regarding the effect of TRBP on Dicer processing (25-27, 32). Two initial studies looked at the effect of TRBP knockdown on the levels of miR-16 but within different cell lines; using HEK-293 cells no effect was noted on miR-16 levels (26) whereas in HeLa cells, a reduction of the miR-16 level was observed (25). According to the miRmine database (60), miR-16 is expressed at higher levels in HEK-293 cells compared to HeLa cells, which is consistent with the activating effect of TRBP being more important in HeLa cells and the observed reduction of miR-16 levels by TRBP knockdown in those cells. In HEK-293 cells and in most tissues, miR-16 is abundantly expressed, most likely at near or above the critical substrate concentration where its cleavage efficiency is not increased by TRBP (60, 61). Moreover, *in vitro* studies have shown a 5-fold reduction in pre-miR-31 cleavage by Dicer upon addition of TRBP, whereas cleavage of pre-let-7a and pre-miR-21 was activated by TRBP under similar conditions (17). This is consistent with our observation that the critical concentration at which TRBP switches from an activator to an inhibitor and becomes ineffective varies for different pre-miRNA substrates.

### 3.7. Conclusion

The kinetic model presented in this study clearly establishes that TRBP acts as a dual modifier of the pre-miRNA cleavage activity of Dicer. We demonstrate for the first time that this dual effect can be observed for a specific pre-miRNA and is modulated by substrate concentration. The activation effect of TRBP may be particularly important at typical physiological concentrations of pre-miRNA substrates. In contrast, the inhibitory effect of TRBP would likely be relevant only for highly abundant substrates. Although the dual effect of TRBP was observed for all the members of the pre-let-7 family with some variations, it may also vary to different extents with other pre-miRNA substrates. Here, TRBP was shown to function as an *equalizer* by reducing the effect of substrate concentration on cleavage efficiency, and such a mechanism is likely necessary to allow the enzyme to process a wide variety of substrates without being completely saturated by those that are highly abundant. This new model for TRBP regulation helps to consolidate previous seemingly contradictory *in vitro* and *in vivo* results, and future studies should help to provide additional evidence for this model in different cellular contexts.

## Acknowledgements

We thankfully acknowledge Dr. Jurgen Sygusch for discussions regarding the kinetic model as well as Geneviève Di Tomasso and Pierre Dagenais for technical assistance with protein and RNA purifications.

## Funding

The research was funded by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to PL.

## Authors' contributions

JB developed the kinetic model and performed almost all of the experiments and analysis. SA assisted with cell culture and Dicer purification and with the data collection of EMSA studies with Dicer and the pre-let-7. PL provided guidance. JB and PL wrote the manuscript. All authors read and approved the final manuscript.

## 3.8. References

1. Winter, J., et al., *Many roads to maturity: microRNA biogenesis pathways and their regulation*. Nat Cell Biol, 2009. **11**(3): p. 228-34.
2. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs*. Genome Res, 2009. **19**(1): p. 92-105.
3. Cheloufi, S., et al., *A dicer-independent miRNA biogenesis pathway that requires Ago catalysis*. Nature, 2010. **465**(7298): p. 584-9.
4. Cifuentes, D., et al., *A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity*. Science, 2010. **328**(5986): p. 1694-8.
5. Havens, M.A., et al., *Biogenesis of mammalian microRNAs by a non-canonical processing pathway*. Nucleic Acids Res, 2012. **40**(10): p. 4626-40.
6. Kozomara, A. and S. Griffiths-Jones, *miRBase: annotating high confidence microRNAs using deep sequencing data*. Nucleic Acids Res, 2014. **42**(Database issue): p. D68-73.

7. Sayed, D. and M. Abdellatif, *MicroRNAs in development and disease*. *Physiol Rev*, 2011. **91**(3): p. 827-87.
8. Roush, S. and F.J. Slack, *The let-7 family of microRNAs*. *Trends Cell Biol*, 2008. **18**(10): p. 505-16.
9. Sun, X., et al., *The insights of Let-7 miRNAs in oncogenesis and stem cell potency*. *J Cell Mol Med*, 2016.
10. Park, J.E., et al., *Dicer recognizes the 5' end of RNA for efficient and accurate processing*. *Nature*, 2011. **475**(7355): p. 201-5.
11. Macrae, I.J., et al., *Structural basis for double-stranded RNA processing by Dicer*. *Science*, 2006. **311**(5758): p. 195-8.
12. Lau, P.W., et al., *The molecular architecture of human Dicer*. *Nat Struct Mol Biol*, 2012. **19**(4): p. 436-40.
13. Liu, Z., et al., *Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate*. *Cell*, 2018.
14. Taylor, D.W., et al., *Substrate-specific structural rearrangements of human Dicer*. *Nat Struct Mol Biol*, 2013. **20**(6): p. 662-70.
15. Feng, Y., et al., *A comprehensive analysis of precursor microRNA cleavage by human Dicer*. *RNA*, 2012. **18**(11): p. 2083-92.
16. Tsutsumi, A., et al., *Recognition of the pre-miRNA structure by Drosophila Dicer-1*. *Nat Struct Mol Biol*, 2011. **18**(10): p. 1153-8.
17. Lee, H.Y. and J.A. Doudna, *TRBP alters human precursor microRNA processing in vitro*. *RNA*, 2012. **18**(11): p. 2012-9.
18. Gu, S., et al., *The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo*. *Cell*, 2012. **151**(4): p. 900-11.
19. Ma, E., et al., *Autoinhibition of human dicer by its internal helicase domain*. *J Mol Biol*, 2008. **380**(1): p. 237-43.
20. Gatignol, A., et al., *Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR*. *Science*, 1991. **251**(5001): p. 1597-600.
21. Yamashita, S., et al., *Structures of the first and second double-stranded RNA-binding domains of human TAR RNA-binding protein*. *Protein Sci*, 2011. **20**(1): p. 118-30.



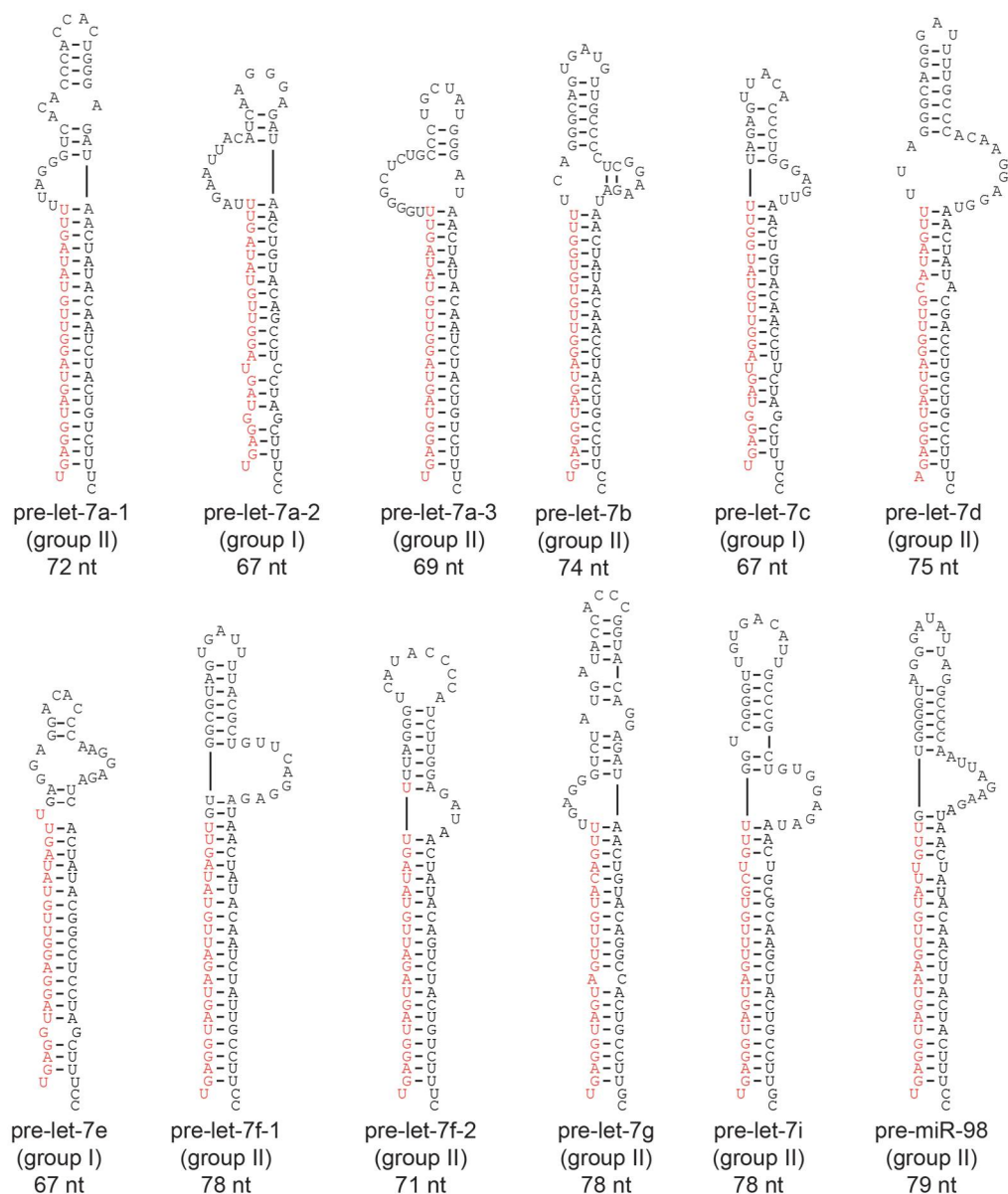
22. Ryter, J.M. and S.C. Schultz, *Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA*. EMBO J, 1998. **17**(24): p. 7505-13.
23. Wilson, R.C., et al., *Dicer-TRBP complex formation ensures accurate mammalian microRNA biogenesis*. Mol Cell, 2015. **57**(3): p. 397-407.
24. Daniels, S.M., et al., *Characterization of the TRBP domain required for dicer interaction and function in RNA interference*. BMC Mol Biol, 2009. **10**: p. 38.
25. Chendrimada, T.P., et al., *TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing*. Nature, 2005. **436**(7051): p. 740-4.
26. Haase, A.D., et al., *TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing*. EMBO Rep, 2005. **6**(10): p. 961-7.
27. De Vito, C., et al., *A TARBP2-dependent miRNA expression profile underlies cancer stem cell properties and provides candidate therapeutic reagents in Ewing sarcoma*. Cancer Cell, 2012. **21**(6): p. 807-21.
28. Chakravarthy, S., et al., *Substrate-specific kinetics of Dicer-catalyzed RNA processing*. J Mol Biol, 2010. **404**(3): p. 392-402.
29. Fareh, M., et al., *TRBP ensures efficient Dicer processing of precursor microRNA in RNA-crowded environments*. Nat Commun, 2016. **7**: p. 13694.
30. Zhu, L., S.K. Kandasamy, and R. Fukunaga, *Dicer partner protein tunes the length of miRNAs using base-mismatch in the pre-miRNA stem*. Nucleic Acids Res, 2018.
31. Fukunaga, R., et al., *Dicer partner proteins tune the length of mature miRNAs in flies and mammals*. Cell, 2012. **151**(3): p. 533-46.
32. Kim, Y., et al., *Deletion of human tarbp2 reveals cellular microRNA targets and cell-cycle function of TRBP*. Cell Rep, 2014. **9**(3): p. 1061-74.
33. Masliah, G., et al., *Structural basis of siRNA recognition by TRBP double-stranded RNA binding domains*. EMBO J, 2018. **37**(6).
34. Bouvette, J.K., D.N.; Fouillen, A.; Amellah, S.; Nanci, A.; Durocher, Y.; Omichinski, J.G.; Legault, P., *High-yield production of human Dicer by transfection of human HEK293-EBNA1 cells grown in suspension*. In preparation, 2018.

35. Cormier, C.Y., et al., *Protein Structure Initiative Material Repository: an open shared public resource of structural genomics plasmids for the biological community*. Nucleic Acids Res, 2010. **38**(Database issue): p. D743-9.
36. Cormier, C.Y., et al., *PSI: Biology-materials repository: a biologist's resource for protein expression plasmids*. J Struct Funct Genomics, 2011. **12**(2): p. 55-62.
37. Seiler, C.Y., et al., *DNASU plasmid and PSI: Biology-Materials repositories: resources to accelerate biological research*. Nucleic Acids Res, 2014. **42**(Database issue): p. D1253-60.
38. Desjardins, A., J. Bouvette, and P. Legault, *Stepwise assembly of multiple Lin28 proteins on the terminal loop of let-7 miRNA precursors*. Nucleic Acids Res, 2014. **42**(7): p. 4615-28.
39. Campbell, F.E., Jr., et al., *Pre-steady-state and stopped-flow fluorescence analysis of Escherichia coli ribonuclease III: insights into mechanism and conformational changes associated with binding and catalysis*. J Mol Biol, 2002. **317**(1): p. 21-40.
40. Comeau, M.A., D.A. Lafontaine, and S. Abou Elela, *The catalytic efficiency of yeast ribonuclease III depends on substrate specific product release rate*. Nucleic Acids Res, 2016. **44**(16): p. 7911-21.
41. Takahashi, T., et al., *Distinguishable in vitro binding mode of monomeric TRBP and dimeric PACT with siRNA*. PLoS One, 2013. **8**(5): p. e63434.
42. Laraki, G., et al., *Interactions between the double-stranded RNA-binding proteins TRBP and PACT define the Medipal domain that mediates protein-protein interactions*. RNA Biol, 2008. **5**(2): p. 92-103.
43. Baici, A., *Kinetics of Enzyme-Modifier Interactions: Selected Topics in the Theory and Diagnosis of Inhibition and Activation Mechanisms*. 2015: Springer Vienna.
44. Leskovac, V., *Comprehensive Enzyme Kinetics*. 2007: Springer US.
45. Segel, I.H., *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*. Wiley Classics Library Edition ed. 1975: Wiley-interscience publication. 957.
46. Wang, Z.X. and R.F. Jiang, *A novel two-site binding equation presented in terms of the total ligand concentration*. FEBS Lett, 1996. **392**(3): p. 245-9.

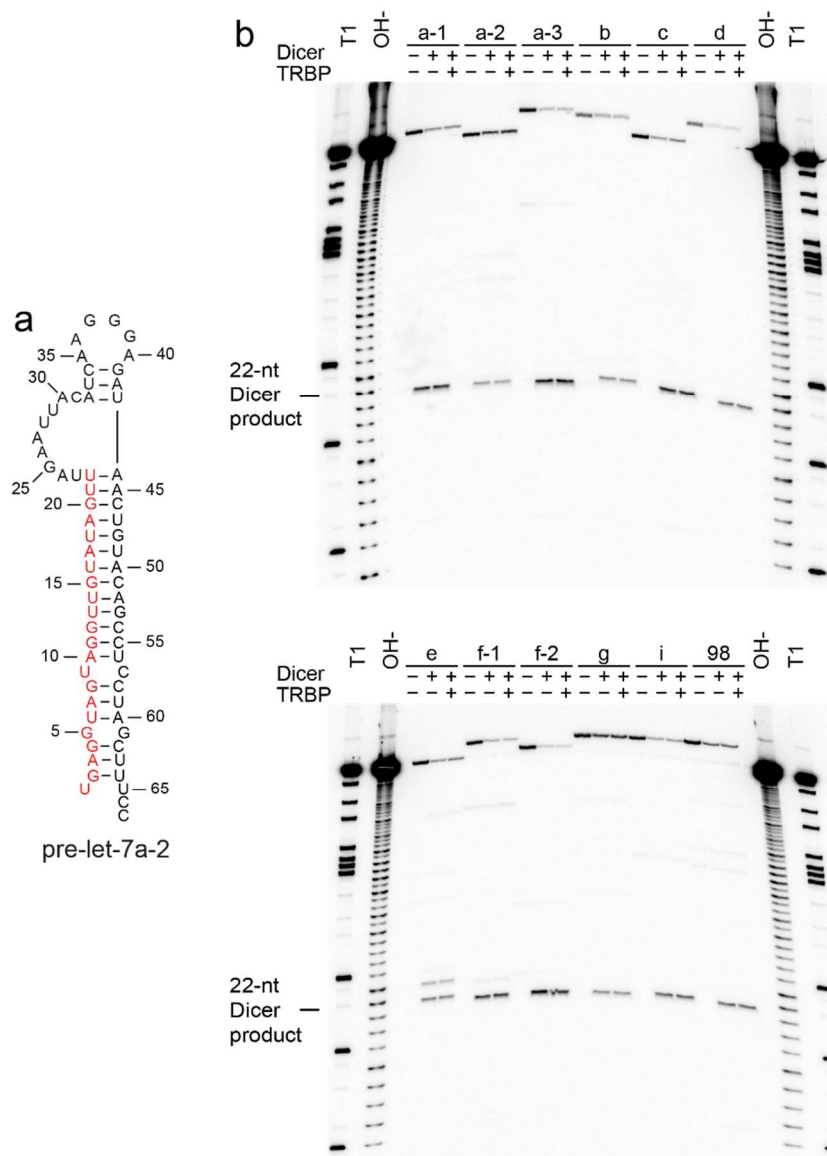
47. Zwillinger, D., *CRC Standard Mathematical Tables and Formulas, 33rd Edition*. 2018: CRC Press LLC.
48. Heo, I., et al., *Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs*. *Cell*, 2012. **151**(3): p. 521-32.
49. Hunter, C.A. and H.L. Anderson, *What is cooperativity?* *Angew Chem Int Ed Engl*, 2009. **48**(41): p. 7488-99.
50. Whitty, A., *Cooperativity and biological complexity*. *Nat Chem Biol*, 2008. **4**(8): p. 435-9.
51. Pyle, A.M., *Translocation and unwinding mechanisms of RNA and DNA helicases*. *Annu Rev Biophys*, 2008. **37**: p. 317-36.
52. Nicholson, A.W., *Ribonuclease III mechanisms of double-stranded RNA cleavage*. *Wiley Interdiscip Rev RNA*, 2014. **5**(1): p. 31-48.
53. Liu, Z., et al., *Structure of precursor microRNA's terminal loop regulates human Dicer's dicing activity by switching DExH/D domain*. *Protein Cell*, 2015. **6**(3): p. 185-93.
54. Hitti, E., A. Neunteufl, and M.F. Jantsch, *The double-stranded RNA-binding protein Xlrbpa promotes RNA strand annealing*. *Nucleic Acids Res*, 1998. **26**(19): p. 4382-8.
55. Bar-Even, A., et al., *The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters*. *Biochemistry*, 2011. **50**(21): p. 4402-10.
56. Lee, E.J., et al., *Systematic evaluation of microRNA processing patterns in tissues, cell lines, and tumors*. *RNA*, 2008. **14**(1): p. 35-42.
57. Ballas, S.K., *Erythrocyte concentration and volume are inversely related*. *Clin Chim Acta*, 1987. **164**(2): p. 243-4.
58. Zhao, L., et al., *Intracellular water-specific MR of microbead-adherent cells: the HeLa cell intracellular water exchange lifetime*. *NMR Biomed*, 2008. **21**(2): p. 159-64.
59. Andersen, B.B., H.J. Gundersen, and B. Pakkenberg, *Aging of the human cerebellum: a stereological study*. *J Comp Neurol*, 2003. **466**(3): p. 356-65.
60. Panwar, B., G.S. Omenn, and Y. Guan, *miRmine: a database of human miRNA expression profiles*. *Bioinformatics*, 2017. **33**(10): p. 1554-1560.
61. Liang, Y., et al., *Characterization of microRNA expression profiles in normal human tissues*. *BMC Genomics*, 2007. **8**: p. 166.

62. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis*. J Comput Chem, 2004. **25**(13): p. 1605-12.
63. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction*. Nucleic Acids Res, 2003. **31**(13): p. 3406-15.

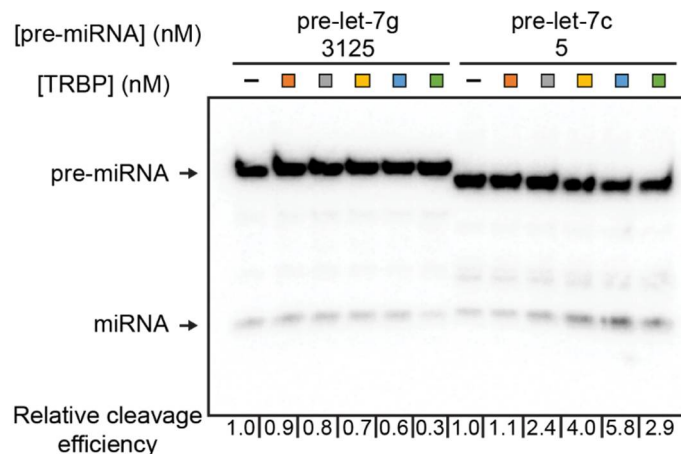
### 3.9. Supplementary information



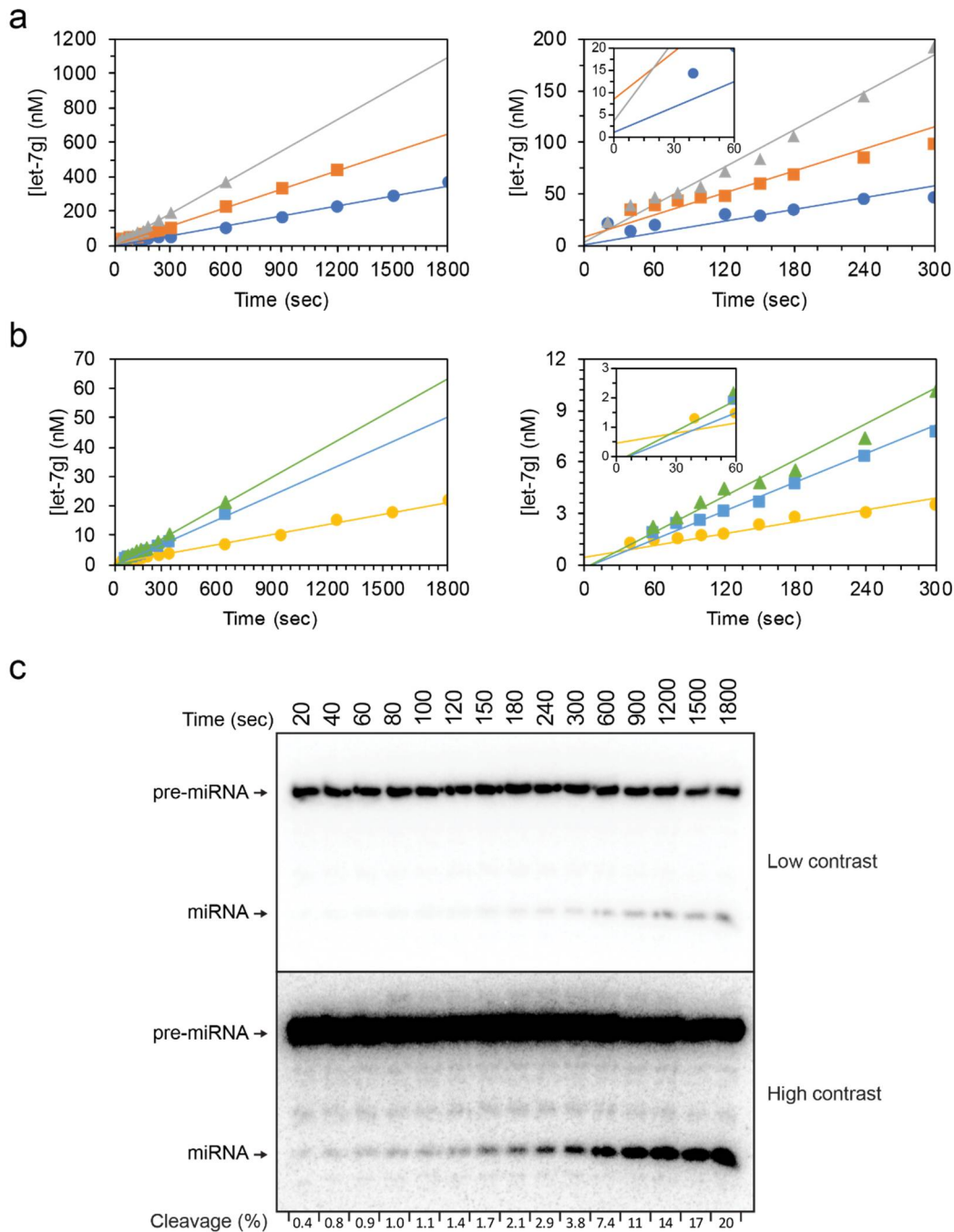
**Figure 3.8. Primary and secondary structures of the twelve pre-let-7 family members.** The sequences of the pre-let-7 RNAs incorporating the most abundant mature miRNA (in red) were obtained from miRbase (6) and secondary structures of pre-let-7 were calculated with the mfold software (63). Group II substrates are known to be mono-uridylated at their 3'-end by the TUT4 polyuridyl transferase, whereas those of group I do not get modified (48).



**Figure 3.9. Effect of TRBP on the formation of potential isoforms of let-7-5p.** (a) Secondary structure and numbering of pre-let-7a-2, used to generate the T<sub>1</sub> (T<sub>1</sub>) and hydroxyl (OH<sup>-</sup>) ladders shown in (b) and (c). (b) Polyacrylamide denaturing gels of cleavage reactions of 1 μM pre-let-7 substrates (top gel: a-1, a-2, a-3, b, c and d; bottom gel: e, f-1, f-2, g, i and 98) with (+) or without (-) 40 mM Dicer and 100 nM TRBP, as shown. Cleavage reactions were incubated for 15 min at 37°C. It is important to know that RNA cleavage reactions for the control T<sub>1</sub> and OH<sup>-</sup> ladders yield products with a 2'-3' cyclic phosphate, whereas Dicer cleavage yields a product with a 3'-OH termini. Thus, the 22-nt let-7-5p, migrates close to the same position as the 23-nt product form the OH<sup>-</sup> ladder.



**Figure 3.10. Effect of TRBP on Dicer cleavage of different pre-let-7 substrates as monitored by denaturing gel electrophoresis.** Typical autoradiograph of a denaturing polyacrylamide gel for the data reported in Figure 3.2, where cleavage reactions of different pre-let-7 substrates by Dicer were monitored at different concentrations of TRBP. For cleavage at a given concentration of a select pre-miRNA, different concentrations of TRBP were used (no color: 0 nM; orange: 10 nM; grey: 40 nM; yellow: 80 nM. blue: 160 nM and green: 1000 nM) and incubated with Dicer for 20 min. The relative Dicer cleavage efficiency was taken as the ratio of the fraction of substrate cleaved at a given TRBP concentration over the fraction of substrate cleaved under the same conditions but in the absence of TRBP.



**Figure 3.11. Evaluation of a potential pre-steady-state burst.** (a)(b) Plots of the concentration of products over the time course of Dicer cleavage reactions containing in (a) 5  $\mu$ M pre-let-7g with 6 nM (blue circles), 12 nM (orange squares) or 18 nM (grey triangles) Dicer and in (b) 250 nM pre-let-7g and 80 nM of TRBP with 0.75 nM (yellow circles), 1.5 nM (light-blue squares) or 2.25 nM (green triangles) Dicer. Full plots are shown on the left and partial



plots of the shorter time points are shown on the right with inserts to show the y-intercept. The plots were fitted by linear regression to obtain the initial rate of reaction ( $v_0$ ) from the slope and the amplitude of the potential burst ( $\pi$ ) from the y-intercept. From data shown in **(a)**, the following values were obtained:  $v_0 = 0.19 \text{ nM s}^{-1}$  and  $\pi = 1 \text{ nM}$  (6 nM Dicer),  $v_0 = 0.36 \text{ nM s}^{-1}$  and  $\pi = 9 \text{ nM}$  (12 nM Dicer) and  $v_0 = 0.61 \text{ nM s}^{-1}$  and  $\pi = 3 \text{ nM}$  (18 nM Dicer). From data shown in **(b)**, the following values were obtained:  $v_0 = 12 \text{ pM s}^{-1}$  and  $\pi = 0.4 \text{ nM}$  (0.75 nM Dicer),  $v_0 = 28 \text{ pM s}^{-1}$  and  $\pi = -0.2 \text{ nM}$  (1.5 nM Dicer) and  $v_0 = 35 \text{ pM s}^{-1}$  and  $\pi = -0.2$  (2.25 nM Dicer). **(c)** Example of an autoradiograph from the denaturing gel used to monitor a single time course. Here, the cleavage conditions are as described for the grey triangles (18 nM Dicer) in (a). Top and bottom panels show the same autoradiograph at different contrast ratios of signal intensity.



**Figure 3.12. Binding of T3 to pre-let-7g.** EMSA performed with 10 pM of 5'-[ $^{32}\text{P}$ ]-labeled pre-let-7g only (control) or after incubation with 1  $\mu\text{M}$  of either T3 or T12.

## 4. Discussion

### 4.1. Purification de Dicer

Bien que nous ayons atteint notre objectif d'améliorer la purification de Dicer, le développement de la méthode s'est avéré un plus gros défi qu'anticipé. Après avoir optimisé les conditions d'expression, le rendement était initialement assez bon pour poursuivre l'optimisation des premières étapes de purification. Cependant, la purification a été assez difficile car il a fallu optimiser au maximum le rendement de chaque étape tout en visant la meilleure pureté finale.

Au début, nous avons eu beaucoup de problèmes de pureté. La purification par affinité comme première étape de purification ne générait pas un échantillon aussi pur que ce que l'on peut généralement obtenir avec des lysats bactériens. Beaucoup d'impuretés étaient co-éluées avec Dicer, dont les chaperonnes Hsp90 et Hsp70. Ces protéines sont les deux plus abondantes dans un lysat de cellules HEK293 et sont connues pour s'associer à Dicer et à Argonaute (65, 186). Il a été déterminé que ces protéines s'associent aux protéines surtout par interactions hydrophobes (187, 188). Avec cette information, j'ai donc tenté d'ajouter des agents connus pour déstabiliser ces interactions. Il s'est avéré que l'ajout de glycérol soit suffisant pour retirer la majorité de ces chaperonnes de la fraction de Dicer. Cela a constitué un moment clé lors de l'optimisation de la purification de Dicer.

L'ajout de glycérol n'a cependant pas suffi à obtenir la pureté désirée. L'ajout d'une étape de purification supplémentaire était nécessaire. Nous avons opté pour une résine échangeuse d'ion puisque cette étape nous assurerait d'éliminer la plupart des acides nucléiques en plus d'aider au retrait des autres contaminants. Ensuite, une étape de filtration sur gel était de mise afin de s'assurer que l'échantillon soit homogène et monomérique.

C'est au moment de passer à la filtration sur gel que d'autres problèmes sont survenus. Nous devons concentrer Dicer avant cette étape afin de limiter le nombre d'injections. Cependant, Dicer était très instable et l'étape de concentration par micro-filtration générait énormément de pertes liées à la précipitation de la protéine. Pour éviter ce problème, nous avons

donc tenté d'interchanger les étapes de purification. L'idée était que la fraction générée par la colonne d'affinité était d'un volume assez petit pour ne pas avoir à concentrer l'échantillon. Après quelques essais, nous en sommes venus à procéder d'abord avec l'étape d'échange d'ion en présence de glycérol pour éliminer les chaperonnes, suivit ensuite de la colonne d'affinité. Le volume de la fraction contenant Dicer était assez petit pour procéder finalement à une seule injection sur la colonne de filtration sur gel.

Nous avons alors évité le problème des pertes lors de la concentration, mais il fallait encore trouver une façon de pouvoir concentrer Dicer après la filtration sur gel. Nous avons donc essayé quelques conditions inspirées d'un kit de stabilisation de protéines (kit Foldit de Hampton Research maintenant discontinué) et avons pu trouver une condition permettant un rendement optimal lors de la concentration de Dicer (189). En utilisant cette condition comme phase mobile lors de la filtration sur gel, nous avons aussi réussi à augmenter l'homogénéité de la protéine en réduisant la multimérisation de Dicer.

La condition finale nous permet de concentrer Dicer jusqu'à 5 mg/mL avant de constater des pertes liées à l'agrégation. De plus, nous pouvons conserver Dicer pendant plus d'une année à -80 °C sans observer de perte apparente de l'activité et de l'homogénéité.

Il y a cependant un point négatif à cette condition. La présence de *n*-dodécyl- $\beta$ -D-maltoside (DDM) rend difficile certaines étapes subséquentes. Ce détergent forme des micelles d'environ 70 kDa (190). Il n'est donc pas possible de le dialyser efficacement. Il peut aussi interférer avec certaines méthodes de quantification comme la coloration de Bradford ou des essais de température de fusion par détection fluorescente avec des molécules fluorescentes comme le SyPRO Orange qui sont basés sur des interactions hydrophobes (191). Heureusement, la plupart des étapes subséquentes présentées dans cette thèse nécessitent de diluer grandement l'échantillon par rapport à la protéine stock, minimisant l'effet du DDM. Il serait probablement possible de trouver une autre condition dans laquelle Dicer serait stable. Cela nécessiterait cependant des appareils de criblage permettant l'essais de nombreuses conditions en parallèle.

## 4.2. Coopérativité de liaison entre Dicer et ses substrats.

La notion de liaison coopérative allostérique est bien établie pour des modèles d'interaction où plusieurs sites de liaison d'un ligand sur un substrat existent et que la liaison d'un ligand change l'interaction subséquente des autres ligands (Figure 4.1a)(192). Dans ce modèle d'interaction, les données expérimentales sont généralement ajustées sur l'équation de Hill, présentée à l'équation 16, et le coefficient ( $n$ ) qui en découle donne une indication du nombre de sites de liaison potentiels pour des ligands. Lorsque la liaison d'un premier ligand favorise la liaison des suivants, on parle de coopérativité positive et la valeur de  $n$  est supérieure à 1. Inversement, la coopérativité est négative lorsque la liaison du premier ligand nuit à la liaison des autres et la valeur de  $n$  est donc entre 0 et 1.

$$\theta = \frac{[L]^n}{(K_d)^n + [L]^n} \quad (16)$$

Où,

$\theta$  = fraction liée

$[L]$  = concentration de ligand

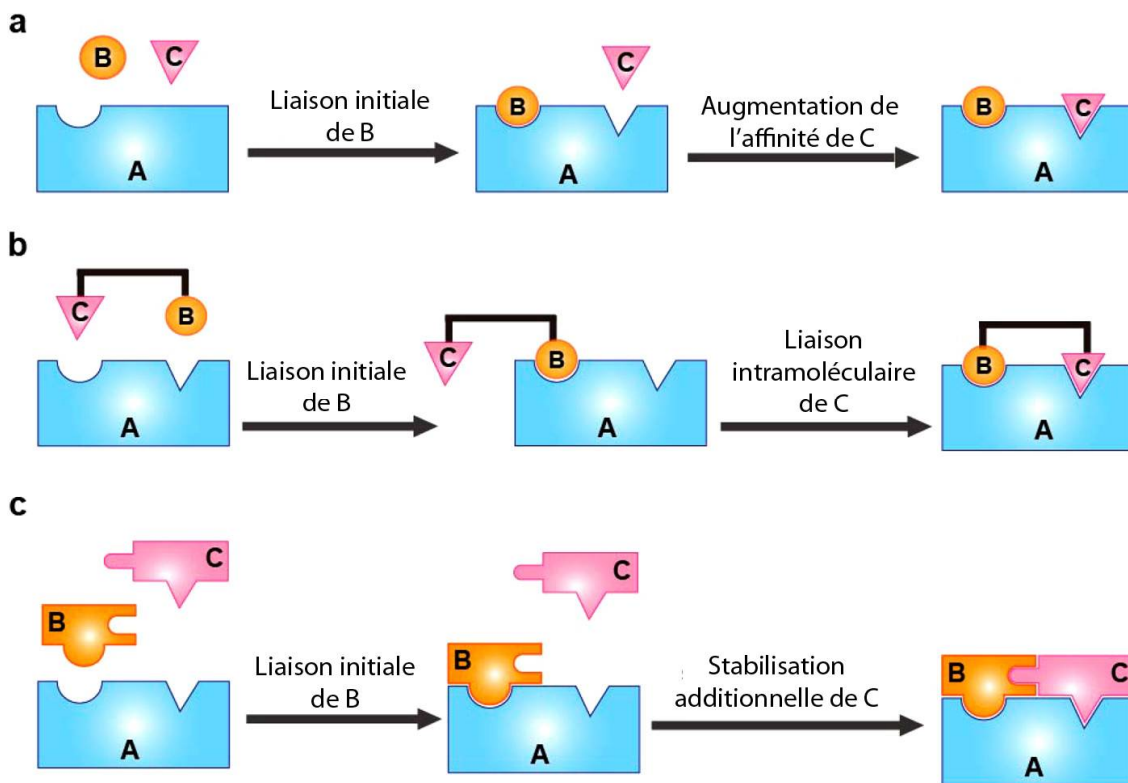
$K_d$  = constante de dissociation

$n$  = coefficient de Hill

Une interaction à stœchiométrie de 1 :1 peut aussi avoir une valeur de coopérativité lorsqu'il y a plusieurs points de contacts distincts entre les deux molécules. On parle alors d'un mécanisme de coopérativité de chélation (*Chelate cooperativity*) ou de pré-organisation [ Figure 4.1bc ; (192, 193)]. Ce type de mécanisme peut aussi être ajusté sur l'équation de Hill mais la valeur de  $n$  ne peut pas être interprétée de la même manière. Ce type d'allostérie n'aura jamais une valeur de  $n$  inférieure à 1. De plus, la valeur de  $n$  peut donner une indication du nombre de sites d'interaction seulement dans les cas extrêmes de très forte coopérativité de chélation positive. En résumé, même si, au niveau mathématique, les deux types de coopérativité peuvent être confondus, ils sont très différents au niveau moléculaire. Il existe aussi des cas plus complexe où les deux types de coopérativité sont présents dans le même système.

Dans le cas de l'interaction de Dicer avec les pre-miARN, il est très bien établi que les domaines PAZ et hélicase de Dicer interagissent avec des régions distinctes du pre-miARN et il pourrait aussi y avoir l'effet des autres domaines qui peuvent aussi interagir avec le pre-miARN (section 1.4.2). Cette multitude de sites d'interaction indique la possibilité de

coopérativité de chélation pour ce système. Nos observations (Tableau I) nous ont montré que certains pre-let-7, notamment pre-let-7e présentent une valeur de  $n > 1$ . Cela suggère que le contact d'une des régions de pre-let-7e avec Dicer facilite la liaison des autres domaines. Cela peut s'expliquer par un changement conformationnel favorable induit par ce premier contact sur une des deux molécules, ce qui a pour effet d'augmenter l'affinité d'interaction envers les sites subséquents. Un indice de tels changement de conformation provient des données structurales connues de Dicer avec certains ligands. Une première étude a déterminé que Dicer change de conformation de manière différente lorsqu'il est lié à un pre-miARN vs un duplex d'ARN (122). Une autre étude structurale a observé des changements de conformation au niveau de la boucle terminale du pre-miARN lors de l'interaction avec Dicer (123). Notre interprétation des résultats restent cependant à confirmer avec des études plus avancées pour caractériser les déterminants structuraux des pre-miARN permettant la coopérativité positive de liaison de Dicer.



**Figure 4.1. Représentation des différents types de coopérativité.** (a) Coopérativité allostérique. La liaison du ligand B à A modifie l'interaction de C sur A (b) Coopérativité de chélation ou de préorganisation. La liaison de B à A facilite la liaison de C sur A. (c)

Coopérativité de préorganisation. La liaison de B sur A permet la stabilisation additionnelle lors de la liaison de C, facilitant la formation du complexe ABC. (adapté de (193))

### **4.3. Futur des études structurales de Dicer**

L'accès à une grande quantité de Dicer et la capacité de concentrer la protéine jusqu'à 5 mg/mL ouvre la porte aux études structurales par cristallographie. En effet, les essais de cristallisation demandent généralement une grande quantité d'échantillon car il faut travailler à haute concentration et tester énormément de conditions avant de trouver celles qui permettent la formation de cristaux avec un bon potentiel de diffraction.

L'autre raison pour laquelle aucune structure cristalline de Dicer n'est publiée est probablement que la cristallisation de cette protéine est très difficile. Effectivement, avec toutes les difficultés encourues lors de l'optimisation du protocole de purification, il est fort probable que Dicer seule ne soit pas assez stable pour former des cristaux. Il serait cependant possible d'élaborer une stratégie permettant de stabiliser Dicer en formant des complexes avec différents substrats et autres protéines. Il a été montré que le domaine dsRBD 3 de TRBP permet une sélection conformationnelle du domaine hélicase (122). En réduisant le nombre de conformations possible de Dicer par l'addition de ce domaine, les chances d'obtenir un cristal seraient plus grandes. L'addition d'un substrat à Dicer pourrait aussi avoir un effet similaire.

Il y a cependant une autre protéine qui pourrait être efficace à cette stabilisation. Comme mentionné précédemment, Dicer a tendance à co-éluer avec des chaperones. Ces protéines sont généralement présentes pour stabiliser les protéines et favoriser leur bon repliement. Il serait possible d'ajouter des chaperones à Dicer et tenter de cristalliser un complexe stable. Il faudra très certainement commencer par étudier l'effet de ces chaperones sur l'activité et la stabilité de Dicer mais elles pourraient être une option viable pour aider à la cristallisation.

### **4.4. Modèle de modification de Dicer par TRBP**

#### **4.4.1. Implication biologique des paramètres cinétiques**

Il y a plusieurs facteurs qui peuvent faire varier le  $K_M$  d'une enzyme. Le plus important est la taille du substrat et le nombre d'interaction que celui-ci peut faire avec l'enzyme. Il a été

montré que les enzymes métaboliques ayant un substrat de plus de 350 Da ont un  $K_M$  moyen 60% moins élevé ( $K_M = 40 \mu\text{M}$ ) que celles ayant une masse plus petite ( $K_M = 100 \mu\text{M}$ ) (194). Dicer ne fait pas exception à la règle. Son  $K_M$  est d'environ  $1 \mu\text{M}$  et ses substrats ont une masse d'environ 20 kDa. Le  $K_M$  de Dicer se retrouve aussi parmi les 99% des enzymes ayant un  $K_M$  de plus de  $0.1 \mu\text{M}$  (194). La particularité de Dicer vient du fait que, contrairement aux enzymes métaboliques qui fonctionnent soit à saturation ou à des concentrations de substrat avoisinant leur  $K_M$ , ses substrats se retrouvent généralement en faible concentration dans les cellules. Il y a donc un besoin de mécanismes connexes afin de réduire son  $K_M$  apparent. Un de ces cofacteurs est TRBP.

Dicer n'est pas la seule enzyme à avoir un cofacteur permettant de réduire son  $K_M$ . Un des meilleurs exemples de ce type de mécanisme provient de l'addition de Coenzyme A (CoA) sur des petits substrats comme l'acétate dans la synthèse des acides gras par acetyl-CoA carboxylase (194). Cette addition de cofacteur permet d'augmenter le nombre d'interaction entre l'enzyme et le substrat pour réduire le  $K_M$  et augmenter l'efficacité enzymatique à faibles concentrations de substrat. TRBP pourrait jouer un rôle similaire auprès de Dicer mais celle-ci joue sur le  $k_{cat}$  plutôt que le  $K_M$ . Il en résulte néanmoins une amélioration de l'activité et une réduction du  $K_M$  apparent à faibles concentrations de substrat.

Un des rôles principaux des microARN est d'offrir une réponse rapide à un stimulus extérieur en inhibant rapidement la traduction de certains ARNm. Modifier l'expression des gènes au niveau de la chromatine peut prendre des heures, voire même des jours (195). Dans le cas des miARN, il est possible d'observer un effet sur l'expression d'une protéine en une heure et une réponse maximale en quelques heures (196). La vitesse de réaction de Dicer dans l'ordre du nanomolaire est sous la barre de  $1 \text{ min}^{-1}$  en absence de cofacteurs. Ce temps de l'ordre d'une réaction par minute corrèle bien avec un temps de réponse de quelques heures. Cependant, un miARN très peu exprimé sera clivé à une vitesse nettement plus lente que  $1 \text{ min}^{-1}$  en absence de cofacteurs et cela ralentira le temps de réponse. Nous avons déterminé que TRBP agissait en tant qu'activateur de l'activité de Dicer dans la gamme du nanomolaire tout juste sous la barre de  $1 \text{ min}^{-1}$  dans le cas de pre-let-7g. TRBP permettrait donc à des miARN exprimés à différents ordres de concentration (du nM au  $\mu\text{M}$ ) d'offrir des temps de réponse équivalents.

#### 4.4.2. Méthodes alternatives pour les études cinétiques

Les expériences décrivant le modèle enzymatique de TRBP ont permis de montrer que l'effet activateur de TRBP permet d'égaliser la vitesse de réaction des pre-miARN sur une grande gamme de concentrations de substrat. Cependant, les résultats des cinétiques sont loin de la perfection et l'erreur sur les résultats est grande. Cela est dû, en grande partie, au design expérimental qui nécessite beaucoup de manipulations. L'expérience a été établie ainsi pour sa simplicité. L'optimisation a été rapide et l'expérience ne nécessite aucune modification chimique au niveau du substrat. Bien qu'il y ait de la variabilité au niveau des résultats et que la régression ne soit pas parfaite, les résultats illustrent bien la tendance et la conclusion ne serait pas bien différente avec de meilleurs résultats.

Il existe des méthodes alternatives pour suivre les réactions en temps réel. Les alternatives qui mériteraient d'être testées impliquent toutes l'utilisation d'un fluorophore pour la détection de l'ARN. D'abord, il possible serait d'ajouter une modification chimique en ajoutant un fluorophore sur le pre-miARN. C'est cependant un gros défi de trouver un bon endroit pour modifier le substrat sans changer l'interaction avec Dicer puisque cette dernière interagit avec toutes les régions du substrat, et cela ne compte pas l'interaction avec TRBP. Une autre option serait d'ajouter une 2-aminopurine sur une paire de bases de la boucle terminale qui voit son environnement chimique modifié après le clivage par Dicer. En effet, la 2-aminopurine voit son signal de fluorescence diminué lorsque qu'elle localisée dans une hélice d'ARN par rapport à sa forme libre car elle participe à des interactions d'empilement (197). Finalement, une autre option serait d'ajouter un ADN complémentaire muni d'un fluorophore à un des produits de clivage afin de détecter les produits. C'est cependant une méthode indirecte qui dépend de la cinétique de l'hybridation des brins. Il faudrait s'assurer que cette molécule n'interfère pas avec Dicer. Un autre problème général de l'utilisation de fluorophore est que le simple changement d'environnement chimique dû à la liaison à Dicer ou TRBP pourrait induire un changement de signal et rendre difficile l'analyse des fluctuations liées à la catalyse comparativement aux interactions.

Les méthodes basées sur la fluorescence demandent donc énormément d'optimisation. Ces méthodes comportent cependant leur lot d'avantages. Les lectures sont faites en temps réel donc il y a beaucoup plus de points de temps mesurés. De plus, une fois la réaction lancée, il



n'est pas nécessaire d'aller prélever des aliquotes et de les charger sur un gel. Cela réduit donc considérablement le nombre de manipulations et les risques d'erreurs expérimentales associées à ces étapes et accélère la procédure. Il serait aussi possible de paralléliser les réactions en utilisant des plaques de 96 puits. Finalement, avec un appareil de flux arrêté, il deviendrait possible de réaliser des expériences à l'état pré-stationnaire afin de dériver plus de paramètres cinétiques.

#### **4.4.3. Validation *in vivo* du modèle d'activation de Dicer par TRBP**

Un modèle enzymatique généré par des expériences *in vitro* bénéficie toujours d'une confirmation par des études *in vivo*. Le modèle proposé par les expériences cinétiques est que la présence de TRBP égalise la vitesse de réaction des différents substrats dans la gamme de concentrations physiologiques de substrat. En absence de TRBP, on s'attendrait à voir une forte diminution de l'activité à basse concentration de substrat (< 100 nM), tandis que l'effet irait de modéré à marginal à plus haute concentration (entre 100 et 600 nM). Le but serait donc de tenter de montrer l'effet de la présence de TRBP sur la maturation de pre-miARN exprimés à des niveaux différents à partir de cellules.

Il est très difficile de calculer la concentration exacte d'un pre-miARN dans la cellule car, même si on arrive à calculer la concentration moyenne, il est possible qu'il y ait des endroits où la concentration locale soit plus élevée ou moins élevée. Cependant, il n'est pas nécessaire de calculer les concentrations absolues pour avoir une idée de l'activité de Dicer. On peut mesurer les niveaux relatifs des miARN et des pre-miARN par q-PCR. La meilleure stratégie serait d'utiliser 3 lignées cellulaires, une exprimant TRBP à son niveau constitutif, une où l'expression de TRBP a été supprimée par siARN ou par délétion CRISPR/Cas9 et une où TRBP est surexprimée. Ensuite, il faudrait exprimer des quantités croissantes d'un pre-miARN dans chacune de ces lignées. Cela pourrait se faire par la transfection d'un plasmide encodant un pre-miARN dont l'expression est sous le contrôle d'un promoteur inductible (198). Une autre alternative serait de comparer des pre-miARN endogènes dont les niveaux d'expression sont différents. En effet, il est possible de savoir quels miARN sont plus abondants dans différentes lignées cellulaires en regardant dans les bases de données comme miRmine (199). L'idée générale de cette expérience est donc de vérifier les ratios miARN: pre-miARN à différents

niveaux d'expression en présence et en absence de TRBP. L'idée d'utiliser un gène rapporteur contenant un site de liaison du miARN choisi ne serait pas une bonne option car TRBP a aussi un rôle dans la formation du complexe RISC. Les résultats dérivés d'une telle méthode seraient donc biaisés. Il est donc important d'obtenir une mesure directe des substrats et produits de réaction.

#### 4.4.4. Étude cinétique du complexe Dicer-TRBP

Le modèle final de l'effet de TRBP sur l'activité de Dicer est très complexe et il serait bénéfique de tenter de le simplifier en isolant d'autres parties du modèle. Nous avons déjà utilisé les deux parties de TRBP de manière isolée mais il y aurait possiblement une manière d'isoler la partie centrale du modèle correspondant à la formation du complexe ternaire composé de Dicer, TRBP et le pre-miARN. L'idée serait de stabiliser le complexe Dicer-TRBP.

Afin de stabiliser ce complexe, on pourrait tenter de lier Dicer et TRBP de manière covalente à l'aide de la réticulation («*Crosslink*»). Il est effectivement possible de cibler un site particulier de réticulation en ajoutant, de manière générale, une cystéine sur une étiquette fusionnée à la protéine d'intérêt (200). Cependant Dicer et TRBP contiennent respectivement 48 et 11 cystéines. Bien que les cystéines ne se retrouvent généralement pas sur la surface des protéines, il serait difficile de cibler la formation d'un complexe covalent localisé précisément au site d'interaction canonique de cas précis de Dicer et TRBP (201).

La meilleure façon d'obtenir le complexe Dicer-TRBP serait de produire une Dicer chimérique possédant les domaines 1 et 2 de TRBP à son extrémité N-terminale. L'enzyme résultante aurait la capacité de recruter le pre-miARN de manière similaire à TRBP et posséderait la totalité des domaines de Dicer. Cette chimère aurait théoriquement une efficacité catalytique nettement supérieure à Dicer et on s'attendrait, selon les données du modèle cinétique, à obtenir un  $K_M$  similaire avec un  $k_{cat}$  beaucoup plus élevé. Si ce genre de chimère fonctionne avec un activité telle qu'attendue et qu'il est possible de purifier la protéine de manière similaire à Dicer, celle-ci aurait un potentiel commercial. Dicer est déjà disponible commercialement en kit permettant de préparer des siARN pour l'inhibition de gènes dans les cellules. Une enzyme plus efficace pourrait améliorer la production de ces siRNA.

#### 4.4.5. Étude plus approfondie de l'étape limitante de l'activité de Dicer

Nos études cinétiques de Dicer ont montré que l'étape limitante semble être située avant la catalyse sans être la dissociation du complexe entre Dicer et son substrat puisque  $K_M > K_d$ . L'hypothèse, appuyée par le plus récent modèle structural de cryo-EM, est que Dicer peut lier le substrat dans une conformation où le substrat n'est pas positionné dans le site actif (123). Il devient donc intéressant d'étudier la manière par laquelle Dicer et son substrat passent à une conformation active car cela pourrait constituer l'étape limitante de l'enzyme.

Il est possible d'évaluer la constante d'équilibre de ce changement de conformation en trouvant  $k_{on}$  et  $k_{off}$  de la liaison entre le Dicer et son substrat. Cela peut être réalisé par plusieurs méthodes différentes telles la résonance plasmonique de surface (202), des expériences de cinétiques de liaison à flux arrêté à l'aide de fluorophore (135) ou encore par des expériences à molécules uniques aussi à l'aide de fluorophores (203). Sans répéter les problèmes possibles liés à l'utilisation de fluorophores, la difficulté principale est de différencier la formation du complexe inactif à celle du complexe actif où le substrat est bien positionné dans le site actif.

Maintenant que des structures à résolution atomique sont disponibles, il serait possible d'étudier ce genre de transition conformationnelle par des approches *in silico* de dynamique moléculaire. Le problème principal avec Dicer et un pre-miARN est le nombre élevé d'atomes qui rend les temps de calculs très longs. Cependant, les avancées technologiques permettent maintenant de supporter ce genre de calculs et le système Dicer-TRBP est en deçà des limitations qui se situent à environ 100 000 atomes avec des infrastructures informatiques raisonnables (204). Certaines méthodes permettent aussi de traiter les atomes par groupes, réduisant ainsi le nombre de trajectoires à calculer. Ces méthodes à grain grossier (*coarse-grained*) permettent une réduction très efficace des temps de calculs mais ce genre de simplification réduit aussi le degré de flexibilité des molécules simulées (205). Le temps de simulation de ces techniques est limité à l'ordre de la femtoseconde pour des simulations incluant tous les atomes à quelques millisecondes pour les modèles à grain grossiers (204). Des simulations dépassant ces temps rendent les calculs très longs. Cependant, la plupart des processus biologiques se produisent dans des ordres de grandeur dépassant la seconde. Afin d'explorer un plus gros espace conformationnel avec des temps de calculs raisonnable, des méthodes comme le TREMD (*temperature replica-exchange molecular dynamics*) ont été mises

au point (206). Le TREMD consiste à lancer plusieurs simulations en parallèle à différentes températures et, à certains moments au cours de la simulation, les modèles générés sont échangés avec ces des autres températures, permettant de sortir des possibles puits énergétiques. Cette méthode est moins appropriée pour les études cinétiques puisque la simulation change de condition au cours du temps mais permet d'explorer un vaste espace conformationnel. La TREMD permettrait probablement d'observer un état où le pre-miARN est parfaitement positionné dans le site actif. Il y a donc une grande variété d'options pour l'approche *in silico* de la cinétique de Dicer.

## 4.5. Étude d'autres régulateurs de Dicer

Bien que les travaux de cette thèse s'intéressent à l'effet de TRBP sur Dicer, la plupart des expériences mises au point peuvent être utilisées pour l'étude de régulateurs différents.

### 4.5.1. PACT

Tel que présenté à la section 1.2.1, PACT est hautement similaire à TRBP (Figure 1.5). À priori, ces deux protéines pourraient avoir des fonctions redondantes dans la régulation de la biogénèse des miARNs. Cependant, il existe des différences permettant de supposer que PACT n'aurait pas une activité activatrice aussi importante que TRBP.

La différence majeure entre les deux protéines provient des résidus R353 et R354 de TRBP qui correspondent à H300 et N301 chez PACT. Ces acides aminés chez PACT permettent son homodimérisation par le dsRBD 3 (160, 173). La présence d'arginines chez TRBP réduit grandement l'affinité de cette interaction (169). Des études de mutagenèse ont confirmé le remplacement des résidus 353 et 354 de TRBP par des acides aminés polaires augmente la stabilité des dimères. De plus, l'interface de dimérisation est aussi utilisée pour l'interaction avec le domaine hélicase de Dicer (126). L'analyse de la structure entre TRBP et Dicer montre que l'arginine 354 forme un pont salin avec Dicer alors que cette interaction est absente dans le cas de PACT. Cela suggère qu'en plus d'avoir potentiellement une affinité plus faible pour Dicer, l'homodimérisation de PACT peut compétitionner pour la liaison à la même interface. Il a donc été proposé que TRBP ait évolué de manière à lier Dicer tandis que PACT a plutôt évolué pour d'autres fonctions et n'aurait pas un rôle tout à fait redondant dans les cellules (173, 207).

En effet, TRBP et PACT agissent aussi dans la réponse antivirale en régulant la kinase PKR (174). Ces informations indiquent donc que PACT aurait un plus faible effet sur la maturation des miARN par rapport à TRBP.

#### 4.5.2. Lin28

La première protéine qui serait intéressante à étudier est Lin28. Nous avons déjà décrit l'interaction de Lin28 avec pre-let-7g dans le passé (93, 94). Cependant, les études *in vitro* de l'effet de Lin28 sur l'activité de Dicer sont minimales. Dans une publication précédente, j'avais montré que l'activité de Dicer est partiellement inhibée par Lin28 lorsqu'on augmente la concentration de cette dernière (94). Cette expérience avait montré une inhibition maximale de 50 à 75 % de l'activité de Dicer, selon les substrats testés. Cette inhibition partielle même à très haute concentration de Lin28 suggère que Dicer est capable de cliver le complexe Lin28/pre-miARN avec une activité moindre.

En revenant sur cette expérience aujourd'hui, je la reprendrais de manière différente. Premièrement, la concentration de substrat utilisée était de 500 nM, ce qui est probablement à la limite supérieure de la gamme de concentrations physiologiques. À la lumière des connaissances acquises dans cette thèse, il serait préférable de répéter cette expérience à plusieurs concentrations différentes de substrat. De plus, il est possible que l'inhibition de Dicer par Lin28 suive un modèle enzymatique similaire à T12. Lin28 semble seulement lier le substrat et aucune interaction directe avec Dicer n'a été observée. Tout comme T12, il y a plus d'un site de liaison pour Lin28 sur pre-let-7.

Il serait aussi intéressant d'étudier l'effet de Lin28 sur Dicer en présence de TRBP. Nous avons déjà supposé que le haut  $K_M$  de Dicer la rende possiblement inefficace dans la gamme de concentrations physiologiques de substrat. Nous avons ensuite montré que l'effet activateur de TRBP se retrouve précisément dans cette gamme de concentrations. Lin28 présente un très fort pouvoir inhibiteur de la maturation de la famille let-7 par plusieurs études *in vivo* (81, 82). Nos résultats préliminaires montrant seulement une inhibition partielle de Lin28 ne semblent pas concorder avec cet effet. Cependant, il est possible qu'en plus d'inhiber Dicer, Lin28 inhibe aussi l'effet de TRBP. Si c'est le cas, il est probable de mitiger une très grande partie de l'activité de Dicer et annulant l'effet de TRBP et en inhibant partiellement l'activité de Dicer, la rendant

quasi nulle dans un contexte physiologique surtout pour les pre-miARN ayant un faible niveau d'expression.

### **4.5.3. Pr55<sup>gag</sup>**

Lors de notre première étude sur la liaison de Lin28 à let-7, nous avons fait un parallèle entre le domaine C-terminal de Lin28 et le fragment NCp7 de la polyprotéine Pr55<sup>gag</sup> du VIH-I (section 1.2.2) et nous avons montré que ces deux domaines lient pre-let-7g avec une affinité similaire. Nous avons donc émis l'hypothèse que Pr55<sup>gag</sup> pourrait lier pre-let-7g et possiblement réguler l'activité de Dicer.

Cette hypothèse a donné lieu à un projet exploratoire réalisé par moi-même et un stagiaire d'été. Le but du projet était de purifier Pr55<sup>gag</sup> afin de valider sa liaison aux précurseurs de la famille let-7 et tester son effet sur le clivage par Dicer. Nos résultats préliminaires nous montrent que Pr55<sup>gag</sup> peut effectivement lier pre-let-7 et semble avoir un effet marqué sur l'activité de Dicer.

Ces résultats sont toujours dans un état préliminaire mais cela peut constituer la base d'un projet à part entière. Il serait intéressant de montrer que le VIH-I puisse influencer directement la maturation des miARN et, possiblement, modifier le comportement d'une cellule en faveur de la réplication et de l'infection virale.

### **4.5.4. Découverte de nouveaux régulateurs de Dicer**

Avec le développement des technologies de criblage à haut débit tels le séquençage de nouvelle génération et la spectrométrie de masse, quelques études se sont intéressées à l'interactome de la maturation des miARN. Ces études, telles que mentionnées en section 1.2.4, ont identifié plusieurs protéines ayant possiblement la capacité de réguler l'activité de Dicer.

Afin de valider ces cibles, il est possible de les purifier et de tester leur effet sur Dicer de manière similaire à ce que nous avons fait avec TRBP en Figure 3.2. Cette expérience constitue une manière rapide de tester l'effet d'une protéine à différentes concentrations de substrat et de la protéine d'intérêt. Si ce test s'avère positif, il sera alors envisageable d'étudier le mécanisme plus en profondeur de manière similaire à ce qui a été présenté dans cette thèse.

La partie la plus difficile est de choisir quelles protéines de ces listes semblent les plus intéressantes à tester en premier afin de maximiser les chances de trouver une protéine régulatrice. Voici une liste des points principaux qui me semblent les plus importants à considérer : (1) Choisir une protéine étant déjà liée à la maturation des miARN en vérifiant les bases de données et la littérature scientifique; (2) Les protéines ayant des domaines de liaison à l'ARN connus sont plus susceptibles de pouvoir interagir avec le pre-miARN avec un avantage supplémentaire pour les protéines pouvant participer à des interactions protéine-protéine; (3) Choisir une protéine ayant été identifiée par plus d'une étude de criblage.

## 4.6. Conclusion générale

Les travaux dans cette thèse ont permis, en premier lieu, de proposer une nouvelle méthode rapide pour purifier une grande quantité de Dicer recombinante pure, active et monomérique, facilitant grandement l'entreprise d'études *in vitro* de Dicer. Cela nous a permis de réaliser des études enzymatiques de Dicer dans le clivage des pre-miARN. Ces expériences nous ont permis de montrer que Dicer avait un  $K_M$  et un  $k_{cat}$  bien plus élevé que précédemment suggéré et qu'un activateur comme TRBP sera nécessaire à augmenter l'activité de Dicer dans un contexte physiologique.

Ensuite, nous avons montré que TRBP est un activateur très efficace de l'activité de Dicer à faible concentration de substrat, mais un inhibiteur à plus haute concentration. Ce rôle inhibiteur serait cependant rarement observé *in vivo* puisque les pre-miARN seraient, dans la majorité des cas, présents dans la gamme de concentration où TRBP est activateur. Ce modèle offre, pour la première fois, un regard sur le rôle de TRBP en fonction de la concentration de substrat. Ce modèle pourra maintenant être validé par des études *in vivo*. Ces travaux constituent des bases solides pour l'étude ultérieure des autres régulateurs de Dicer en présence et en absence de TRBP.

Finalement, ces études apportent un nouveau regard sur l'activité de Dicer et sa régulation par TRBP en ayant une meilleure conscience de l'effet de la concentration de substrat sur la réaction observée. Le modèle proposé est en accord avec les données publiées précédemment et offre une explication à des données, en apparence, contradictoires. De manière

plus générale, cette thèse permet de mieux comprendre la maturation des miARN et facilitera l'étude ultérieure des autres mécanismes de régulation de Dicer.



## 5. Bibliographie

1. Ambros, V. and H.R. Horvitz, *The lin-14 locus of Caenorhabditis elegans controls the time of expression of specific postembryonic developmental events*. Genes Dev, 1987. **1**(4): p. 398-414.
2. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14*. Cell, 1993. **75**(5): p. 843-54.
3. Reinhart, B.J., et al., *The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans*. Nature, 2000. **403**(6772): p. 901-6.
4. Xu, P., et al., *The Drosophila microRNA Mir-14 suppresses cell death and is required for normal fat metabolism*. Curr Biol, 2003. **13**(9): p. 790-5.
5. Abrahante, J.E., et al., *The Caenorhabditis elegans hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs*. Dev Cell, 2003. **4**(5): p. 625-37.
6. Lagos-Quintana, M., et al., *Identification of novel genes coding for small expressed RNAs*. Science, 2001. **294**(5543): p. 853-8.
7. Kozomara, A. and S. Griffiths-Jones, *miRBase: annotating high confidence microRNAs using deep sequencing data*. Nucleic Acids Res, 2014. **42**(Database issue): p. D68-73.
8. Backes, C., et al., *miRPathDB: a new dictionary on microRNAs and target pathways*. Nucleic Acids Res, 2017. **45**(D1): p. D90-D96.
9. Dong, Q., et al., *MicroRNA let-7a inhibits proliferation of human prostate cancer cells in vitro and in vivo by targeting E2F2 and CCND2*. PLoS One, 2010. **5**(4): p. e10147.
10. Mi, S., et al., *MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia*. Proc Natl Acad Sci U S A, 2007. **104**(50): p. 19971-6.
11. Noh, H., et al., *Prediction of miRNA-mRNA associations in Alzheimer's disease mice using network topology*. BMC Genomics, 2014. **15**: p. 644.
12. Gui, Y., et al., *Altered microRNA profiles in cerebrospinal fluid exosome in Parkinson disease and Alzheimer disease*. Oncotarget, 2015. **6**(35): p. 37043-53.

13. Tuddenham, L., et al., *Small RNA deep sequencing identifies microRNAs and other small noncoding RNAs from human herpesvirus 6B*. J Virol, 2012. **86**(3): p. 1638-49.
14. Ebert, M.S. and P.A. Sharp, *Roles for microRNAs in conferring robustness to biological processes*. Cell, 2012. **149**(3): p. 515-24.
15. Cai, X., C.H. Hagedorn, and B.R. Cullen, *Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs*. RNA, 2004. **10**(12): p. 1957-66.
16. Lee, Y., et al., *MicroRNA genes are transcribed by RNA polymerase II*. EMBO J, 2004. **23**(20): p. 4051-60.
17. Baskerville, S. and D.P. Bartel, *Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes*. RNA, 2005. **11**(3): p. 241-7.
18. Rodriguez, A., et al., *Identification of mammalian microRNA host genes and transcription units*. Genome Res, 2004. **14**(10A): p. 1902-10.
19. Kim, Y.K. and V.N. Kim, *Processing of intronic microRNAs*. EMBO J, 2007. **26**(3): p. 775-83.
20. Monteys, A.M., et al., *Structure and activity of putative intronic miRNA promoters*. RNA, 2010. **16**(3): p. 495-505.
21. Borchert, G.M., W. Lanier, and B.L. Davidson, *RNA polymerase III transcribes human microRNAs*. Nat Struct Mol Biol, 2006. **13**(12): p. 1097-101.
22. Lee, Y., et al., *The nuclear RNase III Drosha initiates microRNA processing*. Nature, 2003. **425**(6956): p. 415-9.
23. Ji, X., *The mechanism of RNase III action: how dicer dices*. Curr Top Microbiol Immunol, 2008. **320**: p. 99-116.
24. Kwon, S.C., et al., *Structure of Human DROSHA*. Cell, 2016. **164**(1-2): p. 81-90.
25. Nguyen, T.A., et al., *Functional Anatomy of the Human Microprocessor*. Cell, 2015. **161**(6): p. 1374-87.
26. Han, J., et al., *Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex*. Cell, 2006. **125**(5): p. 887-901.
27. Zeng, Y., R. Yi, and B.R. Cullen, *Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha*. EMBO J, 2005. **24**(1): p. 138-48.

28. Gregory, R.I., et al., *The Microprocessor complex mediates the genesis of microRNAs*. Nature, 2004. **432**(7014): p. 235-40.
29. Sohn, S.Y., et al., *Crystal structure of human DGCR8 core*. Nat Struct Mol Biol, 2007. **14**(9): p. 847-53.
30. Faller, M., et al., *DGCR8 recognizes primary transcripts of microRNAs through highly cooperative binding and formation of higher-order structures*. RNA, 2010. **16**(8): p. 1570-83.
31. Faller, M., et al., *Heme is involved in microRNA processing*. Nat Struct Mol Biol, 2007. **14**(1): p. 23-9.
32. Partin, A.C., et al., *Heme enables proper positioning of Drosha and DGCR8 on primary microRNAs*. Nat Commun, 2017. **8**(1): p. 1737.
33. Quick-Cleveland, J., et al., *The DGCR8 RNA-binding heme domain recognizes primary microRNAs by clamping the hairpin*. Cell Rep, 2014. **7**(6): p. 1994-2005.
34. Yi, R., et al., *Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs*. Genes Dev, 2003. **17**(24): p. 3011-6.
35. Zeng, Y. and B.R. Cullen, *Structural requirements for pre-microRNA binding and nuclear export by Exportin 5*. Nucleic Acids Res, 2004. **32**(16): p. 4776-85.
36. Okada, C., et al., *A high-resolution structure of the pre-microRNA nuclear export machinery*. Science, 2009. **326**(5957): p. 1275-9.
37. Bohnsack, M.T., K. Czaplinski, and D. Gorlich, *Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs*. RNA, 2004. **10**(2): p. 185-91.
38. Lund, E., et al., *Nuclear export of microRNA precursors*. Science, 2004. **303**(5654): p. 95-8.
39. Yi, R., et al., *Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs*. RNA, 2005. **11**(2): p. 220-6.
40. Kim, Y.K., B. Kim, and V.N. Kim, *Re-evaluation of the roles of DROSHA, Exportin-5, and DICER in microRNA biogenesis*. Proc Natl Acad Sci U S A, 2016. **113**(13): p. E1881-9.
41. Sun, H.L., et al., *ERK activation globally downregulates miRNAs through phosphorylating Exportin-5*. Cancer Cell, 2016. **30**(5): p. 723-736.

42. Wan, G., et al., *DNA-damage-induced nuclear export of precursor microRNAs is regulated by the ATM-AKT pathway*. Cell Rep, 2013. **3**(6): p. 2100-12.
43. Melo, S.A., et al., *A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells*. Cancer Cell, 2010. **18**(4): p. 303-15.
44. Ketting, R.F., et al., *Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans*. Genes Dev, 2001. **15**(20): p. 2654-9.
45. Hutvagner, G., et al., *A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA*. Science, 2001. **293**(5531): p. 834-8.
46. Bernstein, E., et al., *Role for a bidentate ribonuclease in the initiation step of RNA interference*. Nature, 2001. **409**(6818): p. 363-6.
47. Xie, Z., et al., *Genetic and functional diversification of small RNA pathways in plants*. PLoS Biol, 2004. **2**(5): p. E104.
48. Lee, Y.S., et al., *Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways*. Cell, 2004. **117**(1): p. 69-81.
49. Provost, P., et al., *Ribonuclease activity and RNA binding of recombinant human Dicer*. EMBO J, 2002. **21**(21): p. 5864-74.
50. Zhang, H., et al., *Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP*. EMBO J, 2002. **21**(21): p. 5875-85.
51. Sinha, N.K., et al., *Drosophila dicer-2 cleavage is mediated by helicase- and dsRNA termini-dependent states that are modulated by Loquacious-PD*. Mol Cell, 2015. **58**(3): p. 406-17.
52. Jakymiw, A., et al., *Overexpression of dicer as a result of reduced let-7 MicroRNA levels contributes to increased cell proliferation of oral cancer cells*. Genes Chromosomes Cancer, 2010. **49**(6): p. 549-59.
53. Karube, Y., et al., *Reduced expression of Dicer associated with poor prognosis in lung cancer patients*. Cancer Sci, 2005. **96**(2): p. 111-5.
54. Heravi-Moussavi, A., et al., *Recurrent somatic DICER1 mutations in nonepithelial ovarian cancers*. N Engl J Med, 2012. **366**(3): p. 234-42.
55. Wu, M.K., et al., *Biallelic DICER1 mutations occur in Wilms tumours*. J Pathol, 2013. **230**(2): p. 154-64.

56. Gurtan, A.M., et al., *In vivo structure-function analysis of human Dicer reveals directional processing of precursor miRNAs*. RNA, 2012. **18**(6): p. 1116-22.
57. Ma, E., et al., *Autoinhibition of human dicer by its internal helicase domain*. J Mol Biol, 2008. **380**(1): p. 237-43.
58. Hammond, S.M., et al., *Argonaute2, a link between genetic and biochemical analyses of RNAi*. Science, 2001. **293**(5532): p. 1146-50.
59. Tomari, Y., et al., *RISC assembly defects in the Drosophila RNAi mutant armitage*. Cell, 2004. **116**(6): p. 831-41.
60. Chendrimada, T.P., et al., *TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing*. Nature, 2005. **436**(7051): p. 740-4.
61. Haase, A.D., et al., *TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing*. EMBO Rep, 2005. **6**(10): p. 961-7.
62. Gredell, J.A., et al., *Recognition of siRNA asymmetry by TAR RNA binding protein*. Biochemistry, 2010. **49**(14): p. 3148-55.
63. Tants, J.N., et al., *Molecular basis for asymmetry sensing of siRNAs by the Drosophila Loqs-PD/Dcr-2 complex in RNA interference*. Nucleic Acids Res, 2017. **45**(21): p. 12536-12550.
64. Pham, J.W., et al., *A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila*. Cell, 2004. **117**(1): p. 83-94.
65. Iwasaki, S., et al., *Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes*. Mol Cell, 2010. **39**(2): p. 292-9.
66. Yoda, M., et al., *ATP-dependent human RISC assembly pathways*. Nat Struct Mol Biol, 2010. **17**(1): p. 17-23.
67. Kawamata, T., H. Seitz, and Y. Tomari, *Structural determinants of miRNAs for RISC loading and slicer-independent unwinding*. Nat Struct Mol Biol, 2009. **16**(9): p. 953-60.
68. Kawamata, T. and Y. Tomari, *Making RISC*. Trends Biochem Sci, 2010. **35**(7): p. 368-76.
69. Liu, Y., et al., *C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation*. Science, 2009. **325**(5941): p. 750-3.
70. Brennecke, J., et al., *Principles of microRNA-target recognition*. PLoS Biol, 2005. **3**(3): p. e85.

71. Meister, G., et al., *Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs*. Mol Cell, 2004. **15**(2): p. 185-97.
72. Elbashir, S.M., et al., *Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate*. EMBO J, 2001. **20**(23): p. 6877-88.
73. Rehwinkel, J., et al., *A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing*. RNA, 2005. **11**(11): p. 1640-7.
74. Behm-Ansmant, I., et al., *mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes*. Genes Dev, 2006. **20**(14): p. 1885-98.
75. Eulalio, A., E. Huntzinger, and E. Izaurralde, *GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay*. Nat Struct Mol Biol, 2008. **15**(4): p. 346-53.
76. Braun, J.E., E. Huntzinger, and E. Izaurralde, *The Role of GW182 Proteins in miRNA-Mediated Gene Silencing*, in *Ten Years of Progress in GW/P Body Research*, E.K.L. Chan and M.J. Fritzler, Editors. 2013, Springer New York: New York, NY. p. 147-163.
77. Ruby, J.G., C.H. Jan, and D.P. Bartel, *Intronic microRNA precursors that bypass Drosha processing*. Nature, 2007. **448**(7149): p. 83-6.
78. Okamura, K., et al., *The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila*. Cell, 2007. **130**(1): p. 89-100.
79. Berezikov, E., et al., *Mammalian mirtron genes*. Mol Cell, 2007. **28**(2): p. 328-36.
80. MacRae, I.J., et al., *In vitro reconstitution of the human RISC-loading complex*. Proc Natl Acad Sci U S A, 2008. **105**(2): p. 512-7.
81. Newman, M.A., J.M. Thomson, and S.M. Hammond, *Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing*. RNA, 2008. **14**(8): p. 1539-49.
82. Viswanathan, S.R., G.Q. Daley, and R.I. Gregory, *Selective blockade of microRNA processing by Lin28*. Science, 2008. **320**(5872): p. 97-100.
83. Moss, E.G., R.C. Lee, and V. Ambros, *The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA*. Cell, 1997. **88**(5): p. 637-46.

84. Van Wynsberghe, P.M., et al., *LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in Caenorhabditis elegans*. Nat Struct Mol Biol, 2011. **18**(3): p. 302-8.
85. Yu, J., et al., *Induced pluripotent stem cell lines derived from human somatic cells*. Science, 2007. **318**(5858): p. 1917-20.
86. Nakagawa, M., et al., *Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts*. Nat Biotechnol, 2008. **26**(1): p. 101-6.
87. Xu, B. and Y. Huang, *Histone H2a mRNA interacts with Lin28 and contains a Lin28-dependent posttranscriptional regulatory element*. Nucleic Acids Res, 2009. **37**(13): p. 4256-63.
88. Jin, J., et al., *Evidence that Lin28 stimulates translation by recruiting RNA helicase A to polysomes*. Nucleic Acids Res, 2011. **39**(9): p. 3724-34.
89. Lei, X.X., et al., *Determinants of mRNA recognition and translation regulation by Lin28*. Nucleic Acids Res, 2012. **40**(8): p. 3574-84.
90. West, J.A., et al., *A role for Lin28 in primordial germ-cell development and germ-cell malignancy*. Nature, 2009. **460**(7257): p. 909-13.
91. Piskounova, E., et al., *Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28*. J Biol Chem, 2008. **283**(31): p. 21310-4.
92. Landsman, D., *RNP-1, an RNA-binding motif is conserved in the DNA-binding cold shock domain*. Nucleic Acids Res, 1992. **20**(11): p. 2861-4.
93. Desjardins, A., et al., *Importance of the NCp7-like domain in the recognition of pre-let-7g by the pluripotency factor Lin28*. Nucleic Acids Res, 2012. **40**(4): p. 1767-77.
94. Desjardins, A., J. Bouvette, and P. Legault, *Stepwise assembly of multiple Lin28 proteins on the terminal loop of let-7 miRNA precursors*. Nucleic Acids Res, 2014. **42**(7): p. 4615-28.
95. Triboulet, R., M. Pirouz, and R.I. Gregory, *A Single Let-7 MicroRNA Bypasses LIN28-Mediated Repression*. Cell Rep, 2015. **13**(2): p. 260-6.
96. Piskounova, E., et al., *Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms*. Cell, 2011. **147**(5): p. 1066-79.

97. Kim, S.K., et al., *SET7/9 methylation of the pluripotency factor LIN28A is a nucleolar localization mechanism that blocks let-7 biogenesis in human ESCs*. Cell Stem Cell, 2014. **15**(6): p. 735-49.
98. Robertson, H.D. and J.J. Dunn, *Ribonucleic acid processing activity of Escherichia coli ribonuclease III*. J Biol Chem, 1975. **250**(8): p. 3050-6.
99. Heo, I., et al., *Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA*. Mol Cell, 2008. **32**(2): p. 276-84.
100. Hagan, J.P., E. Piskounova, and R.I. Gregory, *Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells*. Nat Struct Mol Biol, 2009.
101. Heo, I., et al., *TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation*. Cell, 2009. **138**(4): p. 696-708.
102. Ustianenko, D., et al., *Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs*. RNA, 2013.
103. Chang, H.M., et al., *A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway*. Nature, 2013. **497**(7448): p. 244-8.
104. Michlewski, G. and J.F. Caceres, *Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis*. Nat Struct Mol Biol, 2010. **17**(8): p. 1011-8.
105. Nicastrò, G., et al., *Noncanonical G recognition mediates KSRP regulation of let-7 biogenesis*. Nat Struct Mol Biol, 2012.
106. Guil, S. and J.F. Caceres, *The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a*. Nat Struct Mol Biol, 2007. **14**(7): p. 591-6.
107. Trabucchi, M., et al., *The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs*. Nature, 2009. **459**(7249): p. 1010-4.
108. Di Tomasso, G., L.M. Miller Jenkins, and P. Legault, *ARiBo pull-down for riboproteomic studies based on label-free quantitative mass spectrometry*. RNA, 2016.
109. Treiber, T., et al., *A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis*. Mol Cell, 2017. **66**(2): p. 270-284 e13.
110. Roush, S. and F.J. Slack, *The let-7 family of microRNAs*. Trends Cell Biol, 2008. **18**(10): p. 505-16.
111. Pasquinelli, A.E., et al., *Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA*. Nature, 2000. **408**(6808): p. 86-9.



112. Wulczyn, F.G., et al., *Post-transcriptional regulation of the let-7 microRNA during neural cell specification*. *FASEB J*, 2007. **21**(2): p. 415-26.
113. Heo, I., et al., *Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs*. *Cell*, 2012. **151**(3): p. 521-32.
114. Johnson, C.D., et al., *The let-7 microRNA represses cell proliferation pathways in human cells*. *Cancer Res*, 2007. **67**(16): p. 7713-22.
115. Swaminathan, S., et al., *Differential regulation of the Let-7 family of microRNAs in CD4+ T cells alters IL-10 expression*. *J Immunol*, 2012. **188**(12): p. 6238-46.
116. Johnson, S.M., et al., *RAS is regulated by the let-7 microRNA family*. *Cell*, 2005. **120**(5): p. 635-47.
117. Lee, Y.S. and A. Dutta, *The tumor suppressor microRNA let-7 represses the HMGA2 oncogene*. *Genes Dev*, 2007. **21**(9): p. 1025-30.
118. Sampson, V.B., et al., *MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells*. *Cancer Res*, 2007. **67**(20): p. 9762-70.
119. Yu, F., et al., *let-7 regulates self renewal and tumorigenicity of breast cancer cells*. *Cell*, 2007. **131**(6): p. 1109-23.
120. Shell, S., et al., *Let-7 expression defines two differentiation stages of cancer*. *Proc Natl Acad Sci U S A*, 2007. **104**(27): p. 11400-5.
121. Barh, D., et al., *MicroRNA let-7: an emerging next-generation cancer therapeutic*. *Curr Oncol*, 2010. **17**(1): p. 70-80.
122. Taylor, D.W., et al., *Substrate-specific structural rearrangements of human Dicer*. *Nat Struct Mol Biol*, 2013. **20**(6): p. 662-70.
123. Liu, Z., et al., *Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate*. *Cell*, 2018.
124. Luo, D., et al., *Structural insights into RNA recognition by RIG-I*. *Cell*, 2011. **147**(2): p. 409-22.
125. Myong, S., et al., *Cytosolic viral sensor RIG-I is a 5'-triphosphate-dependent translocase on double-stranded RNA*. *Science*, 2009. **323**(5917): p. 1070-4.
126. Wilson, R.C., et al., *Dicer-TRBP complex formation ensures accurate mammalian microRNA biogenesis*. *Mol Cell*, 2015. **57**(3): p. 397-407.

127. Kolakofsky, D., E. Kowalinski, and S. Cusack, *A structure-based model of RIG-I activation*. RNA, 2012. **18**(12): p. 2118-27.
128. Qin, H., et al., *Structure of the Arabidopsis thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein-protein interaction*. RNA, 2010. **16**(3): p. 474-81.
129. Kurzynska-Kokorniak, A., et al., *Revealing a new activity of the human Dicer DUF283 domain in vitro*. Sci Rep, 2016. **6**: p. 23989.
130. Lee, Y., et al., *The role of PACT in the RNA silencing pathway*. EMBO J, 2006. **25**(3): p. 522-32.
131. Macrae, I.J., et al., *Structural basis for double-stranded RNA processing by Dicer*. Science, 2006. **311**(5758): p. 195-8.
132. Tian, Y., et al., *A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer*. Mol Cell, 2014. **53**(4): p. 606-16.
133. Park, J.E., et al., *Dicer recognizes the 5' end of RNA for efficient and accurate processing*. Nature, 2011. **475**(7355): p. 201-5.
134. Gan, J., et al., *Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III*. Cell, 2006. **124**(2): p. 355-66.
135. Campbell, F.E., Jr., et al., *Pre-steady-state and stopped-flow fluorescence analysis of Escherichia coli ribonuclease III: insights into mechanism and conformational changes associated with binding and catalysis*. J Mol Biol, 2002. **317**(1): p. 21-40.
136. Liang, Y.H., et al., *Structure of a eukaryotic RNase III postcleavage complex reveals a double-ruler mechanism for substrate selection*. Mol Cell, 2014. **54**(3): p. 431-44.
137. MacRae, I.J., K. Zhou, and J.A. Doudna, *Structural determinants of RNA recognition and cleavage by Dicer*. Nat Struct Mol Biol, 2007. **14**(10): p. 934-40.
138. Sinha, N.K., et al., *Dicer uses distinct modules for recognizing dsRNA termini*. Science, 2018. **359**(6373): p. 329-334.
139. Ryter, J.M. and S.C. Schultz, *Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA*. EMBO J, 1998. **17**(24): p. 7505-13.
140. Chakravarthy, S., et al., *Substrate-specific kinetics of Dicer-catalyzed RNA processing*. J Mol Biol, 2010. **404**(3): p. 392-402.

141. Vermeulen, A., et al., *The contributions of dsRNA structure to Dicer specificity and efficiency*. RNA, 2005. **11**(5): p. 674-82.
142. Tsutsumi, A., et al., *Recognition of the pre-miRNA structure by Drosophila Dicer-1*. Nat Struct Mol Biol, 2011. **18**(10): p. 1153-8.
143. Feng, Y., et al., *A comprehensive analysis of precursor microRNA cleavage by human Dicer*. RNA, 2012. **18**(11): p. 2083-92.
144. Gu, S., et al., *The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo*. Cell, 2012. **151**(4): p. 900-11.
145. Song, M.S. and J.J. Rossi, *Molecular mechanisms of Dicer: endonuclease and enzymatic activity*. Biochem J, 2017. **474**(10): p. 1603-1618.
146. Much, C., et al., *Endogenous Mouse Dicer Is an Exclusively Cytoplasmic Protein*. PLoS Genet, 2016. **12**(6): p. e1006095.
147. Burger, K., et al., *Nuclear phosphorylated Dicer processes double-stranded RNA in response to DNA damage*. J Cell Biol, 2017.
148. Burger, K. and M. Gullerova, *Nuclear re-localization of Dicer in primary mouse embryonic fibroblast nuclei following DNA damage*. PLoS Genet, 2018. **14**(2): p. e1007151.
149. Doyle, M., et al., *The double-stranded RNA binding domain of human Dicer functions as a nuclear localization signal*. RNA, 2013. **19**(9): p. 1238-52.
150. Hawley, B.R., et al., *The emerging role of RNAs in DNA damage repair*. Cell Death Differ, 2017. **24**(4): p. 580-587.
151. Antoniou, A., et al., *The dynamic recruitment of TRBP to neuronal membranes mediates dendritogenesis during development*. EMBO Rep, 2018. **19**(3).
152. Sambandan, S., et al., *Activity-dependent spatially localized miRNA maturation in neuronal dendrites*. Science, 2017. **355**(6325): p. 634-637.
153. Gatignol, A., et al., *Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR*. Science, 1991. **251**(5001): p. 1597-600.
154. Tabara, H., et al., *The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans*. Cell, 2002. **109**(7): p. 861-71.
155. Liu, Q., et al., *R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway*. Science, 2003. **301**(5641): p. 1921-5.

156. Gatignol, A., C. Buckler, and K.T. Jeang, *Relatedness of an RNA-binding motif in human immunodeficiency virus type 1 TAR RNA-binding protein TRBP to human PI/dsI kinase and Drosophila staufer*. Mol Cell Biol, 1993. **13**(4): p. 2193-202.
157. Hitti, E.G., et al., *Oligomerization activity of a double-stranded RNA-binding domain*. FEBS Lett, 2004. **574**(1-3): p. 25-30.
158. Takahashi, T., et al., *Distinguishable in vitro binding mode of monomeric TRBP and dimeric PACT with siRNA*. PLoS One, 2013. **8**(5): p. e63434.
159. Masliah, G., et al., *Structural basis of siRNA recognition by TRBP double-stranded RNA binding domains*. EMBO J, 2018. **37**(6).
160. Laraki, G., et al., *Interactions between the double-stranded RNA-binding proteins TRBP and PACT define the Medipal domain that mediates protein-protein interactions*. RNA Biol, 2008. **5**(2): p. 92-103.
161. Fareh, M., et al., *TRBP ensures efficient Dicer processing of precursor microRNA in RNA-crowded environments*. Nat Commun, 2016. **7**: p. 13694.
162. Daniels, S.M., et al., *Characterization of the TRBP domain required for dicer interaction and function in RNA interference*. BMC Mol Biol, 2009. **10**: p. 38.
163. Lee, H.Y., et al., *Differential roles of human Dicer-binding proteins TRBP and PACT in small RNA processing*. Nucleic Acids Res, 2013. **41**(13): p. 6568-76.
164. Zhu, L., S.K. Kandasamy, and R. Fukunaga, *Dicer partner protein tunes the length of miRNAs using base-mismatch in the pre-miRNA stem*. Nucleic Acids Res, 2018.
165. Lee, H.Y. and J.A. Doudna, *TRBP alters human precursor microRNA processing in vitro*. RNA, 2012. **18**(11): p. 2012-9.
166. Kim, Y., et al., *Deletion of human tarbp2 reveals cellular microRNA targets and cell-cycle function of TRBP*. Cell Rep, 2014. **9**(3): p. 1061-74.
167. Trettin, K.D., et al., *Loquacious-PD facilitates Drosophila Dicer-2 cleavage through interactions with the helicase domain and dsRNA*. Proc Natl Acad Sci U S A, 2017. **114**(38): p. E7939-E7948.
168. Gleghorn, M.L. and L.E. Maquat, *'Black sheep' that don't leave the double-stranded RNA-binding domain fold*. Trends Biochem Sci, 2014. **39**(7): p. 328-40.
169. Yamashita, S., et al., *Structures of the first and second double-stranded RNA-binding domains of human TAR RNA-binding protein*. Protein Sci, 2011. **20**(1): p. 118-30.

170. Benoit, M.P., et al., *The RNA-binding region of human TRBP interacts with microRNA precursors through two independent domains*. Nucleic Acids Res, 2013. **41**(7): p. 4241-52.
171. Yang, S.W., et al., *Structure of Arabidopsis HYPONASTIC LEAVES1 and its molecular implications for miRNA processing*. Structure, 2010. **18**(5): p. 594-605.
172. Hartig, J.V. and K. Forstemann, *Loqs-PD and R2D2 define independent pathways for RISC generation in Drosophila*. Nucleic Acids Res, 2011. **39**(9): p. 3836-51.
173. Heyam, A., et al., *Conserved asymmetry underpins homodimerization of Dicer-associated double-stranded RNA-binding proteins*. Nucleic Acids Res, 2017. **45**(21): p. 12577-12584.
174. Singh, M., et al., *Stress-induced phosphorylation of PACT reduces its interaction with TRBP and leads to PKR activation*. Biochemistry, 2011. **50**(21): p. 4550-60.
175. Fernandez-Valverde, S.L., R.J. Taft, and J.S. Mattick, *Dynamic isomiR regulation in Drosophila development*. RNA, 2010. **16**(10): p. 1881-8.
176. Czech, B., et al., *Hierarchical rules for Argonaute loading in Drosophila*. Mol Cell, 2009. **36**(3): p. 445-56.
177. Okamura, K., N. Liu, and E.C. Lai, *Distinct mechanisms for microRNA strand selection by Drosophila Argonautes*. Mol Cell, 2009. **36**(3): p. 431-44.
178. Michaelis, L., et al., *The original Michaelis constant: translation of the 1913 Michaelis-Menten paper*. Biochemistry, 2011. **50**(39): p. 8264-9.
179. Henri, V., *Lois générales de l'action des diastases*. 1903: Librairie Scientifique A. Hermann.
180. Leskovac, V., *Comprehensive Enzyme Kinetics*. 2007: Springer US.
181. Briggs, G.E. and J.B. Haldane, *A Note on the Kinetics of Enzyme Action*. Biochem J, 1925. **19**(2): p. 338-9.
182. Baici, A., *Kinetics of Enzyme-Modifier Interactions: Selected Topics in the Theory and Diagnosis of Inhibition and Activation Mechanisms*. 2015: Springer Vienna.
183. Myers, J.W., et al., *Recombinant Dicer efficiently converts large dsRNAs into siRNAs suitable for gene silencing*. Nat Biotechnol, 2003. **21**(3): p. 324-8.
184. Ye, X. and Q. Liu, *Expression, purification, and analysis of recombinant Drosophila Dicer-1 and Dicer-2 enzymes*. Methods Mol Biol, 2008. **442**: p. 11-27.

185. Sinha, N.K. and B.L. Bass, *Overexpression and purification of Dicer and accessory proteins for biochemical and structural studies*. Methods, 2017.
186. Schirle, M., M.A. Heurtier, and B. Kuster, *Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry*. Mol Cell Proteomics, 2003. **2**(12): p. 1297-305.
187. Zhu, X., et al., *Structural analysis of substrate binding by the molecular chaperone DnaK*. Science, 1996. **272**(5268): p. 1606-14.
188. Karagoz, G.E., et al., *Hsp90-Tau complex reveals molecular basis for specificity in chaperone action*. Cell, 2014. **156**(5): p. 963-74.
189. Armstrong, N., A. de Lencastre, and E. Gouaux, *A new protein folding screen: application to the ligand binding domains of a glutamate and kainate receptor and to lysozyme and carbonic anhydrase*. Protein Sci, 1999. **8**(7): p. 1475-83.
190. Strop, P. and A.T. Brunger, *Refractive index-based determination of detergent concentration and its application to the study of membrane proteins*. Protein Sci, 2005. **14**(8): p. 2207-11.
191. Niesen, F.H., H. Berglund, and M. Vedadi, *The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability*. Nat Protoc, 2007. **2**(9): p. 2212-21.
192. Hunter, C.A. and H.L. Anderson, *What is cooperativity?* Angew Chem Int Ed Engl, 2009. **48**(41): p. 7488-99.
193. Li, Y., et al., *Cooperativity Principles in Self-Assembled Nanomedicine*. Chem Rev, 2018. **118**(11): p. 5359-5391.
194. Bar-Even, A., et al., *The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters*. Biochemistry, 2011. **50**(21): p. 4402-10.
195. Voss, T.C. and G.L. Hager, *Visualizing chromatin dynamics in intact cells*. Biochim Biophys Acta, 2008. **1783**(11): p. 2044-51.
196. Lu, L., et al., *Time Series miRNA-mRNA integrated analysis reveals critical miRNAs and targets in macrophage polarization*. Sci Rep, 2016. **6**: p. 37446.
197. Jean, J.M. and K.B. Hall, *2-Aminopurine fluorescence quenching and lifetimes: role of base stacking*. Proc Natl Acad Sci U S A, 2001. **98**(1): p. 37-41.

198. Gossen, M., et al., *Transcriptional activation by tetracyclines in mammalian cells*. Science, 1995. **268**(5218): p. 1766-9.
199. Panwar, B., G.S. Omenn, and Y. Guan, *miRmine: a database of human miRNA expression profiles*. Bioinformatics, 2017. **33**(10): p. 1554-1560.
200. Trahan, C. and M. Oeffinger, *Targeted cross-linking-mass spectrometry determines vicinal interactomes within heterogeneous RNP complexes*. Nucleic Acids Res, 2016. **44**(3): p. 1354-69.
201. Burns, A., P. Olszowy, and P. Ciborowski, *2 - Biomolecules*, in *Proteomic Profiling and Analytical Chemistry (Second Edition)*, P. Ciborowski and J. Silberring, Editors. 2016, Elsevier: Boston. p. 7-24.
202. Fivash, M., E.M. Towler, and R.J. Fisher, *BIAcore for macromolecular interaction*. Curr Opin Biotechnol, 1998. **9**(1): p. 97-101.
203. Luo, Y., J.A. North, and M.G. Poirier, *Single molecule fluorescence methodologies for investigating transcription factor binding kinetics to nucleosomes and DNA*. Methods, 2014. **70**(2-3): p. 108-18.
204. Hospital, A., et al., *Molecular dynamics simulations: advances and applications*. Adv Appl Bioinform Chem, 2015. **8**: p. 37-47.
205. Orozco, M., et al., *Coarse-grained representation of protein flexibility. Foundations, successes, and shortcomings*. Adv Protein Chem Struct Biol, 2011. **85**: p. 183-215.
206. Sugita, Y. and Y. Okamoto, *Replica-exchange molecular dynamics method for protein folding*. Chemical Physics Letters, 1999. **314**(1): p. 141-151.
207. Pullagura, S.R.N., et al., *Functional Redundancy of DICER Co-factors TARBP2 and PRKRA During Murine Embryogenesis Does Not Involve miRNA Biogenesis*. Genetics, 2018.