

Instytut Chemii Bioorganicznej  
Polskiej Akademii Nauk  
w Poznaniu

Zakład Biochemii Rybonukleoprotein

mgr inż. Agnieszka Szczepańska

**Wybrane aspekty oddziaływań pomiędzy ludzką  
rybonukleazą Dicer i cząsteczkami kwasów nukleinowych**

Selected aspects of the interactions between the human Dicer ribonuclease  
and nucleic acids

Praca doktorska wykonana pod kierunkiem  
dr hab. Anny Kurzyńskiej-Kokorniak, prof. IChB PAN

Poznań, 2023

*„Wszystko, co widziałam nowego i czego się uczyłam, zachwycalo mnie.  
Było to jakby objawienie nowego świata, świata wiedzy, do którego  
nareszcie otwarto mi wolny dostęp.“*

Maria Skłodowska-Curie

Praca finansowana ze środków Narodowego Centrum Nauki, Grant SONATA BIS

nr 2016/22/E/NZ1/00422

**Pragnę wyrazić wdzięczność każdej osobie, która przyczyniła się do powstania tej pracy.**

*Szczególne podziękowania składam mojej Pani Promotor, dr hab. Annie Kurzyńskiej-Kokorniak, prof. IChB PAN za cierpliwość, zaufanie, życzliwość, wszystkie dobre słowa i bezcenne wskazówki oraz nieustające wsparcie w trudnych momentach.*

*Serdeczne podziękowania składam również Pracownikom Zakładu Biochemii Rybonukleoprotein IChB PAN za serdeczną atmosferę, a w szczególności mgr inż. Marcie Wojnickiej i dr Kindze Ciechanowskiej za wsparcie oraz niezliczone dyskusje o charakterze nie tylko naukowym.*

*Dziękuję wszystkim Pracownikom Ośrodka Genomiki Medycznej - OMICRON Uniwersytetu Jagiellońskiego za miłą atmosferę, wsparcie i podzielenie się ze mną wiedzą na temat metody irCLIP, w szczególności serdecznie dziękuję dr n. med. Pawłowi Wołkow, dr Justynie Totoń-Żurańskiej, mgr inż. Agnieszce Borys oraz mgr inż. Pawłowi Koniecznemu.*

*Dziękuję również dr Michałowi Sewerynowi za pomoc w analizie wyników sekwencjonowania NGS.*

*Serdeczne podziękowania składam także Kolegom i Koleżankom IChB PAN za ciekawe dyskusje naukowe oraz cenne wskazówki, w szczególności mgr inż. Karolinie Świtońskiej-Kurkowskiej za przyjaźń i nieustające wsparcie.*

*Serdecznie dziękuję także Rodzinie i Przyjaciółom za wsparcie, cierpliwość, a także nieustającą wiarę we mnie.*

**Dziękuję!**

## Spis treści

Dorobek naukowy autorki .....	8
Streszczenie .....	9
Abstract.....	11
Wykaz stosowanych skrótów.....	13
1. Wprowadzenie.....	16
1.1 Ludzka rybonukleaza Dicer i jej rola w procesie biogenezy miRNA.....	17
1.1.1 Biogeneza miRNA .....	17
1.1.2 Budowa ludzkiej rybonukleazy Dicer.....	18
1.1.2.1 Domena helikazy .....	19
1.1.2.2 Domena DUF283.....	20
1.1.2.3 Domeny Platforma i PAZ.....	21
1.1.2.4 Domeny RNazy III i dsRBD .....	22
1.1.3 Kanoniczne substraty ludzkiej rybonukleazy Dicer .....	22
1.2 Białka oddziałujące z pre-miRNA .....	23
1.3 Białka oddziałujące z Dicer.....	24
1.4 Rola Dicer poza szlakiem biogenezy miRNA: różnorodność substratów i funkcji Dicer.....	25
2. Cel pracy.....	27
3. Materiały i metody .....	29
3.1 Materiały .....	29
3.1.1 Oligonukleotydy .....	29
3.1.2 Bufory i roztwory barwiące .....	31
3.1.3 Materiał biologiczny .....	36
3.1.4 Żele poliakrylamidowe (PAA).....	37
3.1.5 Żele agarozowe .....	37
3.1.6 Plazmidy .....	38
3.1.7 Preparaty białkowe.....	38
3.1.8 Wzorce wielkości.....	38
3.1.9 Ważniejsze enzymy i odczynniki wykorzystane podczas badań.....	38
3.1.10 Ważniejsze komercyjnie dostępne zestawy wykorzystane w badaniach.....	39
3.2 Metody.....	40
3.2.1 Elektroforeza w żelach poliakrylamidowych w warunkach denaturujących.....	40
3.2.2 Elektroforeza w żelach poliakrylamidowych w warunkach natywnych.....	40



3.2.3	Elektroforeza w żelach typu SDS-PAGE .....	41
3.2.4	Elektroforeza w żelach agarozowych .....	41
3.2.5	Znakowanie radioizotopowe końca 5' oligonukleotydów .....	41
3.2.6	Znakowanie końca 5' oligonukleotydów barwnikami fluorescencyjnymi ..	42
3.2.7	Oczyszczanie kwasów nukleinowych w żelach PAA.....	42
3.2.8	Przygotowanie substratów dwuniciowych.....	43
3.2.9	Przygotowanie plazmidów kodujących warianty delecyjne ludzkiej rybonukleazy Dicer .....	43
3.2.10	Transformacja komórek bakteryjnych metodą szoku cieplnego (ang. <i>heat-shock</i> ) .....	45
3.2.11	Weryfikacja kolonii bakteryjnych po transformacji metodą szoku cieplnego.....	46
3.2.12	Izolacja plazmidowego DNA.....	47
3.2.13	Hodowla i transfekcja komórek HEK 293T .....	47
3.2.14	Przygotowanie lizatów komórkowych.....	48
3.2.15	Immunoprecypitacja wariantów delecyjnych hDicer oraz hDicer.....	48
3.2.16	Analiza typu Western Blot.....	49
3.2.17	Cięcia substratów przez warianty delecyjne hDicer .....	49
3.2.18	Badanie oddziaływań białko-kwas nukleinowy w teście spowolnienia migracji w żelach poliakrylamidowych (EMSA).....	50
3.2.19	Badanie aktywności wspierającej parowanie cząsteczek kwasów nukleinowych .....	51
3.2.20	Ocena ilościowo-jakościowa preparatów kwasów nukleinowych.....	52
3.2.21	Odwrotna transkrypcja oraz reakcja qPCR.....	52
3.2.22	Ligacja cząsteczek wykorzystanych w testach wiązania domeny PPC z cząsteczkami RNA.....	53
3.2.23	Produkcja preparatu 5'PM PPC hDicer w bakteryjnym systemie ekspresyjnym.....	53
3.2.24	Izolacja i oczyszczanie preparatu białkowego 5'PM PPC hDicer .....	54
3.2.25	Wymiana buforu i zagęszczenie preparatu białkowego 5'PM PPC hDicer .	54
3.2.26	Oznaczanie stężenia białka 5'PM PPC hDicer metodą Bradforda.....	55
3.2.27	Technika irCLIP (ang. <i>infrared crosslinking immunoprecipitation</i> ).....	55
3.2.27.1	Tworzenie kompleksów białko•RNA za pomocą promieni UV (ang. <i>UV crosslinking</i> ) .....	55
3.2.27.2	Przygotowanie złoża oraz lizatu komórkowego do wyławiania kompleksów białko•RNA .....	56

3.2.27.3	Defosforylacja końca 3' cząsteczek RNA związanych z białkiem oraz ligacja z adapterem wyznakowanym barwnikiem bliskiej podczerwieni.....	57
3.2.27.4	Rozdział kompleksów w żelu SDS-PAGE oraz transfer na membranę nitrocelulozową.....	58
3.2.27.5	Izolacja RNA z kompleksów białko•RNA.....	59
3.2.27.6	Odwrotna transkrypcja i ligacja fragmentów cDNA.....	60
3.2.27.7	Optymalizacja reakcji PCR oraz namnażanie fragmentów cDNA....	62
3.2.27.8	Oczyszczanie wybranych fragmentów biblioteki cDNA z żelu.....	63
3.2.28	Rozdział fragmentów bibliotek z zastosowaniem systemu TapeStation .....	64
3.2.29	Normalizacja bibliotek do sekwencjonowania NGS .....	64
3.2.30	Sekwencjonowanie z zastosowaniem sekwenatora Illumina NextSeq550..	65
3.2.31	Analiza danych pochodzących z sekwencjonowania NGS .....	65
4.	Wyniki .....	67
4.1	Badanie znaczenia domeny DUF283 dla aktywności ludzkiej rybonukleazy Dicer.....	67
4.1.1	Otrzymanie konstruktów genetycznych kodujących warianty $\Delta$ DUF(625-752) oraz $\Delta$ DUF(630-709) .....	67
4.1.2	Otrzymywanie preparatów białkowych $\Delta$ DUF(625-752) i $\Delta$ DUF(630-709) oraz hDicer .....	70
4.1.3	Badanie aktywności otrzymanych preparatów białkowych $\Delta$ DUF(625-752) oraz $\Delta$ DUF(630-709).....	73
4.1.3.1	Badanie aktywności RNazowej wariantów $\Delta$ DUF(625-752) oraz $\Delta$ DUF(630-709) w warunkach <i>in vitro</i> .....	73
4.1.3.2	Badanie aktywności RNazowej wariantów $\Delta$ DUF(625-752) oraz $\Delta$ DUF(630-709) w warunkach <i>in cellulo</i> .....	76
4.1.3.3	Badanie aktywności wspierania parowania RNA-RNA przez warianty $\Delta$ DUF(625-752) oraz $\Delta$ DUF(630-709) .....	78
4.2	Rola domeny DUF283 hDicer w wiązaniu komórkowych RNA.....	80
4.2.1	Izolacja kompleksów białko•RNA z wybranych ludzkich linii komórkowych.....	80
4.2.2	Analiza ilościowo-jakościowa przygotowanych bibliotek cDNA do sekwencjonowania NGS.....	82
4.2.3	Raport jakości uzyskanych danych pochodzących z sekwencjonowania bibliotek cDNA .....	84
4.2.4	Wstępna analiza porównawcza pul RNA wiązanych przez warianty $\Delta$ DUF(625-752), $\Delta$ DUF(630-709) oraz hDicer .....	91
4.3	Badania oddziaływań hDicer z cząsteczkami RNA i DNA przyjmującymi strukturę G-kwadruksów .....	100

5. Dyskusja .....	112
6. Podsumowanie.....	120
7. Bibliografia.....	122
8. Suplementy .....	130
9. Załączniki .....	132

## Dorobek naukowy autorki

Publikacje autorki, które dotyczą zagadnień przedstawionych w rozprawie doktorskiej:

1. **Szczepanska A.**, Wojnicka M. & Kurzynska-Kokorniak A. (2021). The Significance of the DUF283 Domain for the Activity of Human Ribonuclease Dicer. *International Journal of Molecular Sciences*, 13;22(16):8690. IF<sub>2021</sub> 6,208.
2. Koralewska N.\*, **Szczepanska A.\***, Ciechanowska K., Wojnicka M., Pokornowska M., Milewski C.M., Gudanis D., Baranowski D., Nithin C., Bujnicki M.J., Gdaniec Z., Figlerowicz M. & Kurzynska-Kokorniak A. (2021). RNA and DNA G-quadruplexes bind to human dicer and inhibit its activity. *Cellular and Molecular Life Sciences*, 78:3709–3724. IF<sub>2021</sub> 9,207.

Inne prace autorki:

1. Wojnicka M., **Szczepanska A.** & Kurzynska-Kokorniak A. (2020). Unknown Areas of Activity of Human Ribonuclease Dicer: A Putative Deoxyribonuclease Activity. *Molecules* 2020, 25(6), 1414. IF<sub>2021</sub> 4,927.
2. Pokornowska M., Milewski C.M., Ciechanowska K., **Szczepanska A.**, Wojnicka M., Radogostowicz Z., Figlerowicz M. & Kurzynska-Kokorniak A. (2020). The RNA-RNA base pairing potential of human Dicer and Ago2 proteins. *Cellular and Molecular Life Sciences*, 78:3709–3724. IF<sub>2021</sub> 9,207.
3. Ciechanowska K., **Szczepanska A.**, Szpotkowski K., Urbanowicz A. & Kurzynska-Kokorniak A. (2023). New insights into the helicase domain of human Dicer and its biochemical properties (publikacja w recenzji).

(\*) pierwszy współautor

## Streszczenie

Rybonukleazy Dicer należą do rodziny rybonukleaz III (RNaz III), będących endorybonukleazami specyficznymi wobec dwuniciowych RNA (dsRNA). Dicer znane są głównie z ich kluczowej roli w procesie biogenezy małych regulatorowych RNA (srRNA): mikroRNA (miRNA) oraz małych interferujących RNA (siRNA). W ostatnim czasie pojawia się coraz więcej doniesień literaturowych opisujących aktywności Dicer niepowiązane z jej aktywnością RNazową i produkcją srRNA. Niniejsza praca doktorska poświęcona jest rybonukleazie Dicer człowieka (hDicer). hDicer, podobnie jak inne Dicer kręgowców, jest białkiem wielodomenowym i złożona jest z N-końcowej domeny helikazowej, domeny DUF283 (ang. *domain of unknown function*), domen: Platformy i PAZ, helisy łączącej, dwóch domen RNazowych (RIIIa i RIIIb) oraz C-końcowej domeny wiążącej dsRNA (dsRBD). Domeny: Platforma i PAZ wraz z helisą łączącą są często określane jako kasetta PPC (ang. *Platform-PAZ-Connector helix*). Kasetta PPC odgrywa kluczową rolę w rozpoznawaniu i kotwiczeniu kanonicznych substratów Dicer, czyli prekursorów miRNA (pre-miRNA) i prekursorów siRNA (pre-siRNA).

Wcześniejsze badania prowadzone w Instytucie Chemii Bioorganicznej PAN wykazały, iż domena DUF283 hDicer wiąże jednoniciowe RNA i DNA, nie wiąże natomiast dwuniciowych RNA i DNA. Co ciekawe, dalsze badania wykazały, iż DUF283 może wspierać proces parowania cząsteczek kwasów nukleinowych. Kolejne analizy ujawniły, że hDicer pełnej długości również wykazuje taką aktywność, działając tym samym jak białka typu ang. *nucleic acid annealer* (białka wspierające hybrydyzację komplementarnych kwasów nukleinowych). W oparciu o wyniki przeprowadzonych badań postawiono hipotezę, że domena DUF283 jest istotna dla aktywności hDicer wspierającej parowanie cząsteczek kwasów nukleinowych. W celu weryfikacji tej hipotezy stworzono warianty delecyjne hDicer pozbawione sekwencji aminokwasowej kodującej domenę DUF283: (i) wariant  $\Delta$ DUF(630-709) oraz (ii) wariant  $\Delta$ DUF(625-752). Wariant  $\Delta$ DUF(625-752), oprócz delecji domeny DUF283, pozbawiony był także fragmentów oskrzydających tę domenę. Przeprowadzone analizy wykazały, iż w warunkach *in vitro* wariant  $\Delta$ DUF(630-709) prezentował podobną aktywność RNazową wobec użytych w badaniach pre-miRNA i pre-siRNA, co hDicer (białko typu dzikiego). Natomiast wariant  $\Delta$ DUF(625-752), wobec tych samych substratów, wykazywał zdecydowanie słabszą aktywność RNazową. W przypadku badań *in cellulo*, wariant  $\Delta$ DUF(625-752) również produkował wybrane miRNA z niższą wydajnością,

w porównaniu do wariantu  $\Delta$ DUF(630-709), czy hDicer. Co jednak ważne, oba warianty nie wykazywały aktywności wspierającej parowanie cząsteczek kwasów nukleinowych. Uzyskane dane wspierają hipotezę, że domena DUF283 jest kluczowa dla aktywności typu *nucleic acid annealer* hDicer.

Rybonukleazy Dicer mogą wiązać w komórkach nie tylko cząsteczki pre-miRNA oraz pre-siRNA, ale także mRNA, czy długie niekodujące RNA. Biorąc pod uwagę ten fakt, kolejnym celem badawczym było określenie roli domeny DUF283 hDicer w wiązaniu puli komórkowych RNA. W badaniach tych wykorzystano metodę irCLIP-seq (ang. *infrared crosslinking immunoprecipitation followed by NGS sequencing*) oraz ludzkie embrionalne komórki nerki (HEK 293T), a także pochodne linie komórkowe typu *DICER1* knock-out, produkujące: (i) wariant  $\Delta$ DUF(630-709), (ii) wariant  $\Delta$ DUF(625-752) lub (iii) hDicer typu dzikiego (tzw. *rescue control*). Przeprowadzone badania wykazały, że poszczególne warianty delecyjne hDicer wiążą inne pule RNA w komórkach HEK 293T, inna była także pula RNA wiązanych przez hDicer. Zgromadzone wyniki pozwalają sądzić, iż domena DUF283 hDicer bierze udział w rozpoznawaniu i wiązaniu określonej puli komórkowych RNA. Co ciekawe, zgromadzone dane ujawniły, że w puli RNA wiązanych przez hDicer znajdują się RNA bogate w ciągi guanozynowe (G). Analizy bioinformatyczne wykazały, iż większość z tych RNA posiada potencjał do tworzenia struktur G-kwadrupleksu. Otrzymane dane wytyczyły kolejny cel badań, który dotyczył oddziaływań hDicer z cząsteczkami RNA i DNA przyjmującymi struktury G-kwadrupleksu. Przeprowadzone badania wykazały, iż hDicer wiąże zarówno G-kwadrupleksy RNA (G4-RNA), jak i DNA (G4-DNA), a ich wiązanie odbywa się najprawdopodobniej w obrębie kasety PPC hDicer. W oparciu o przeprowadzone modelowania molekularne postawiono hipotezę, że cząsteczki pre-miRNA i cząsteczki przyjmujące strukturę G-kwadrupleksu kotwiczone są w tym samym rejonie kasety PPC hDicer. Rzeczywiście, w wyniku dodania do mieszaniny reakcyjnej G4-RNA lub G4-DNA, obserwowano inhibicję cięcia pre-miRNA przez hDicer. Zdobyta wiedza poszerza zrozumienie molekularnych podstaw oddziaływań pomiędzy rybonukleazą Dicer oraz cząsteczkami RNA i DNA przyjmującymi struktury G-kwadrupleksu.

Podsumowując, wyniki badań prowadzonych w ramach realizacji niniejszej pracy doktorskiej wspierają dotychczasowe obserwacje, że rola rybonukleaz Dicer może wykraczać daleko poza proces biogenezy srRNA.

## Abstract

Dicer ribonucleases belong to the Ribonuclease III (RNase III) family, which is a group of endoribonucleases that are specific for double-stranded RNAs (dsRNAs). Dicers are mainly known for their important role in the biogenesis of small regulatory RNAs (srRNAs), such as microRNAs (miRNAs) and small interfering RNAs (siRNAs). Recently, there has been an increasing number of reports describing Dicer activities unrelated to its RNase activity and srRNA production. This dissertation is devoted to the human Dicer ribonuclease (hDicer). hDicer, like other vertebrate Dicer proteins, is a multi-domain protein composed of an N-terminal helicase domain, a domain of unknown function 283 (DUF283), Platform, PAZ domain, a connector helix domain, two RNase III domains (RIIIa and RIIIb), and a C-terminal dsRNA-binding domain (dsRBD). The region spanning the Platform–PAZ–Connector helix domains is called the PPC cassette, and it plays a key role in the recognizing and anchoring canonical substrates of Dicer, such as miRNA precursors (pre-miRNAs) and siRNA precursors (pre-siRNAs).

Previous studies conducted at the Institute of Bioorganic Chemistry Polish Academy of Sciences have revealed that the DUF283 domain of hDicer binds single-stranded RNAs and DNAs but not the double-stranded RNA and DNA substrates. Interestingly, further studies showed that the DUF283 domain can promote base pairing of complementary nucleic acid molecules. Subsequent studies revealed that the full-length hDicer also exhibits such activity, thus acting as a nucleic acid annealer. Based on the results of these studies, it was hypothesized that the DUF283 domain is crucial for the annealing activity of hDicer. To verify this hypothesis, two variants of hDicer, lacking the amino acid sequence encoding the DUF283 domain, were created: (i) the  $\Delta$ DUF(630-709) variant and (ii) the  $\Delta$ DUF(625-752) variant. The  $\Delta$ DUF(625-752) variant, in addition to the deletion of the DUF283 domain, also lacked amino acids flanking this domain. The analyses showed that *in vitro*, the  $\Delta$ DUF(630-709) variant presented similar RNase activity towards the pre-miRNAs and pre-siRNAs used in the study, as the wild-type hDicer. In contrast, the  $\Delta$ DUF(625-752) variant showed significantly weaker RNase activity towards the same substrates. Likewise, *in cellulo* studies of the  $\Delta$ DUF(625-752) variant demonstrated that this variant produced the selected miRNAs with lower efficiency, as compared to the  $\Delta$ DUF(630-709) variant, or hDicer. However, both variants did not facilitate base pairing of complementary nucleic acids. The obtained data strongly

support the hypothesis that the DUF283 domain is crucial for the nucleic acid annealing activity of hDicer.

In the cell, Dicer ribonucleases can bind not only pre-miRNA and pre-siRNA substrates, but also other RNAs, e.g., mRNAs, or long non-coding RNAs. Given this fact, another research was conducted to determine the role of the DUF283 hDicer domain in the binding of cellular RNAs. In this study, the irCLIP-seq (infrared crosslinking immunoprecipitation followed by NGS sequencing) approach and human embryonic kidney cells (HEK 293T) were used, as well as *DICER1* knock-out cell lines producing: (i) the  $\Delta$ DUF(630-709) variant, (ii) the  $\Delta$ DUF(625-752) variant, or (iii) the wild-type hDicer (rescue control). The collected results revealed that the RNA pools bound by the  $\Delta$ DUF hDicer variants and the wild-type full-length hDicer were different. These results suggest that the DUF283 hDicer domain is involved in the recognition and binding of a specific pool of cellular RNAs. Interestingly, the collected data revealed that the pool of RNAs bound by hDicer included RNAs rich in guanosine (G) tracks. Bioinformatics analyses showed that most of these RNAs have a potential to adopt G-quadruplex structures. Consequently, another research goal was to investigate the interactions between hDicer and nucleic acid molecules adopting G-quadruplex structures. The collected results showed that hDicer binds both RNA G-quadruplexes (G4-RNA) and DNA G-quadruplexes (G4-DNA), and that this binding presumably occurs within the PPC cassette of hDicer. Based on the results of the molecular modeling, it was hypothesized that pre-miRNAs and molecules adopting the G-quadruplex structures anchor within the same region of the PPC hDicer cassette. Indeed, when G4-RNA or G4-DNA was added to the reaction mixture, inhibition of pre-miRNA processing by hDicer was observed. The knowledge gained during these studies expands our understanding of the molecular basis of interactions between Dicer ribonucleases and RNA or DNA molecules adopting G-quadruplex structures.

In conclusion, the results of the experimental research conducted in the course of this dissertation support observations that the role of Dicer ribonucleases may extend far beyond the srRNA biogenesis pathways.



## Wykaz stosowanych skrótów

Skrót	Określenie
3'PM PPC	wariant kasety PPC z mutacjami w obrębie kieszeni wiążącej 3'
5'PM PPC	wariant kasety PPC z mutacjami w obrębie kieszeni wiążącej 5'
5-LO	(ang. <i>5-lipoxygenase</i> ) 5-lipooksygenaza
aa	aminokwasy
ADAR1	(ang. <i>adenosine deaminase acting on RNA 1</i> ) deaminaza adenozyiny działająca na RNA1
Ago	(ang. <i>Argonaute protein</i> ) białko Argonaute
ATP	adenozyno-5'-trifosforan
BSA	(ang. <i>bovine serum albumin</i> ) surowicza albumina wołowa
cDNA	(ang. <i>complementary DNA</i> ) komplementarny DNA
C-koniec	koniec karboksylowy białka
Cryo-EM	(ang. <i>cryoelectron microscopy</i> ) mikroskopia krioelektronowa
DCL4	ang. <i>dicer-like 4 protein</i>
DExD/H-box	motyw strukturalny charakterystyczny dla RNA helikaz DExD/H-box
DGCR8	ang. <i>DiGeorge syndrome chromosomal region 8</i>
DHx9	ang. <i>DExH-Box Helicase 9</i>
diRNA	(ang. <i>DNA damage-induced small RNA</i> )
DNA-G4	G-kwadrupleks DNA
dNTP	trifosforany deoksyrybonukleotydów
DSBs	(ang. <i>DNA double-stranded breaks</i> ) pęknięcia podwójnej nici DNA
dsDNA	(ang. <i>double-stranded DNA</i> ) dwuniciowe DNA
dsRBD	(ang. <i>Double-stranded RNA binding domain</i> ) domena wiążąca dsRNA
dsRNA	(ang. <i>double-stranded RNA</i> ) dwuniciowe RNA
DUF283	( ang. <i>domain of unknown function 283</i> ) domena o nieznannej funkcji 283
EMSA	(ang. <i>electrophoretic mobility shift assay</i> ) metoda różnicowanej migracji w żelu poliakrylamidowym
EXP5	(ang. <i>exportin 5</i> ) eksportyna 5
G	guanozyna
G4	G-kwadrupleks
hDicer	(ang. <i>human Dicer</i> ) ludzka rybonukleaza Dicer

HEK 293T	(ang. <i>Human Embryonic Kidney 293 Cell line</i> ) linia komórkowa ludzkich komórek embrionalnych nerki nr 293
HERNA	ang. <i>helicase with RNase motif</i>
hg38	ang. <i>Genome Reference Consortium Human Build 38</i>
irCLIP	(ang. <i>infrared cross-linking immunoprecipitation</i> ) immunoprecypitacja sieciowania w podczerwieni
irCLIP-seq	(ang. <i>infrared crosslinking immunoprecipitation followed by NGS sequencing</i> ) immunoprecypitacja sieciowania w podczerwieni oraz sekwencjonowanie nowej generacji
K <sub>d</sub>	równowagowa stała wiązania
kDa	kilo Dalton
kpz	kilo par zasad
KSRP	ang. <i>KH-type splicing regulatory protein</i>
lncRNA	(ang. <i>long non-coding RNA</i> ) długie niekodujące RNA
miRNA	(ang. <i>miRNA</i> ) mikro RNA
mRNA	(ang. <i>messenger RNA</i> ) matrycowy RNA
ncRNA	(ang. <i>non-coding RNA</i> ) niekodujące RNA
N-koniec	koniec aminowy białka
NMR	(ang. <i>nuclear magnetic resonance</i> ) spektroskopia magnetycznego rezonansu jądrowego
nt	nukleotyd
PAA	poliakrylamid
PACT	(ang. <i>protein activator of protein kinase R</i> ) proteinowy aktywator kinazy proteinowej R
PAR-CLIP	ang. <i>photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation</i>
PAZ	(ang. <i>Piwi/Argonaute/Zwill</i> ) domena Piwi/Argonaute/Zwill
PCR	(ang. <i>polymerase chain reaction</i> ) reakcja łańcuchowej polimerazy
PKR	ang. <i>protein kinase R</i>
Pol II	(ang. <i>RNA polymerase II</i> ) polimeraza RNA II
pre-miRNA	(ang. <i>precursor miRNA</i> ) prekursor miRNA
pre-siRNA	(ang. <i>siRNA precursor</i> ) prekursor siRNA
pri-miRNA	(ang. <i>primary miRNA</i> ) pierwotny transkrypt miRNA
pz	par zasad
RBM3	ang. <i>RNA binding motif protein 3</i>
RHA	ang. <i>RNA helicase A</i>

RISC	ang. <i>RNA-induced silencing complex</i>
RNA-G4	G-kwadrupleks RNA
RNAi	(ang. <i>interference RNA</i> ) interferencja RNA
RNazy III	grupa rybonukleaz należących do klasy III
rpm	(ang. <i>rotations per minute</i> ) obroty na minutę
RT-qPCR	(ang. <i>reverse transcription quantitative real-time PCR</i> ) reakcja łańcuchowa polimerazy w czasie rzeczywistym
SDS	(ang. <i>sodium dodecyl sulfate</i> ) siarczan dodecyłu sodu
siRNA	(ang. <i>small interfering RNA</i> ) mały interferujący RNA
snoRNA	(ang. <i>small nucleolar RNA</i> ) mały jąderkowy RNA
srRNA	(ang. <i>small regulatory RNA</i> ) krótki regulatorowy RNA
ssDNA	(ang. <i>single-stranded DNA</i> ) jednoniciowy DNA
ssRNA	(ang. <i>single-stranded RNA</i> ) jednoniciowy RNA
TDP-43	ang. <i>TAR DNA-binding protein-43</i>
TRBP	ang. <i>TAR RNA-binding protein</i>
tRF	ang. <i>tRNA-derived fragments</i>
tRNA	(ang. <i>transfer RNA</i> ) transportujący RNA
TUT	ang. <i>Terminal Uridyl Transferase</i>
UMI	ang. <i>Unique Molecular Identifiers</i>
UTR	(ang. <i>untranslated region</i> ) rejon niepodlegający translacji
WT	(ang. <i>wild-type</i> ) białko typu dzikiego

## 1. Wprowadzenie

Historia badań nad cząsteczkami RNA regulującymi ekspresję genów sięga już 1993 roku, w którym to ukazała się praca ujawniająca, że rozwój larwalny nicienia *Caenorhabditis elegans* (*C. elegans*) zależy od inhibicji ekspresji genu *lin-14* przez krótki transkrypt genu *lin-4* [1, 2]. Ustalono wtedy, że krótkie cząsteczki RNA zawierają sekwencje komplementarne do rejonów niepodlegających translacji (UTR, ang. *untranslated region*) w mRNA (ang. *messenger RNA*) *lin-14* i uczestniczą w regulacji poziomu tego transkryptu w komórce. Mechanizm tego zjawiska pozostawał nieznany. Dopiero przełomowe wyniki badań nad procesem RNAi zespołów A. Firea i C. Mello, opublikowane w 1998 roku, pozwoliły rozwikłać zagadkę leżącą u podstaw zjawiska nazwanego przez badaczy RNAi (ang. *interference RNA*). Dzięki przeprowadzonym eksperymentom udało się pokazać, że wprowadzenie do komórek *C. elegans* dupleksu RNA o długości 742 par zasad (pz) doprowadziło do 10-krotnie wydajniejszego wyciszenia genu, w porównaniu do wprowadzenia jednoniciowej cząsteczki RNA do komórki, a do osiągnięcia takiego efektu wystarczyło kilka cząsteczek dupleksu RNA w komórce [3]. Obaj naukowcy zostali uhonorowani nagrodą Nobla w dziedzinie fizjologii i medycyny w 2006 roku za badania nad zjawiskiem interferencji RNA u nicienia *C. elegans*.

Niniejsza praca poświęcona jest ludzkiej rybonukleazie Dicer (hDicer). Pierwsze szczegółowe doniesienie literaturowe o hDicer pojawiło się w 2000 roku. W pracy Matsuda i in. po raz pierwszy opisano gen *DICER1*, ówczasie pod nazwą *HERNA* (ang. *helicase with RNase motif*). Badaczom udało się pokazać, że przewidziana sekwencja aminokwasowa produktu białkowego genu *HERNA* jest podobna do sekwencji aminokwasowej helikaz występujących u *C. elegans* (K12H4.8) oraz drożdży (C8A4.08C). Ponadto, udało się im również zidentyfikować siedem sekwencji konsensusowych typowych dla motywów strukturalnych ATP-zależnej helikazy RNA, w tym domenę wiążącą ATP, domenę DExH, domenę wiążącą RNA, a na karboksylowym końcu (C-koniec) białka Herna rozpoznali także dwa motywy rybonukleazy III (RNazy III).

Zarówno odkrycie zjawiska RNAi, jak i badania nad białkami typu Dicer, rozpoczęły prawdziwą rewolucję w badaniach nad procesem biogenezy małych regulatorowych RNA (srRNA, ang. *small regulatory RNA*): mikro RNA (miRNA, ang. *microRNA*) oraz małych interferujących RNA (siRNA, ang. *small interfering RNA*).

## 1.1 Ludzka rybonukleaza Dicer i jej rola w procesie biogenezy miRNA

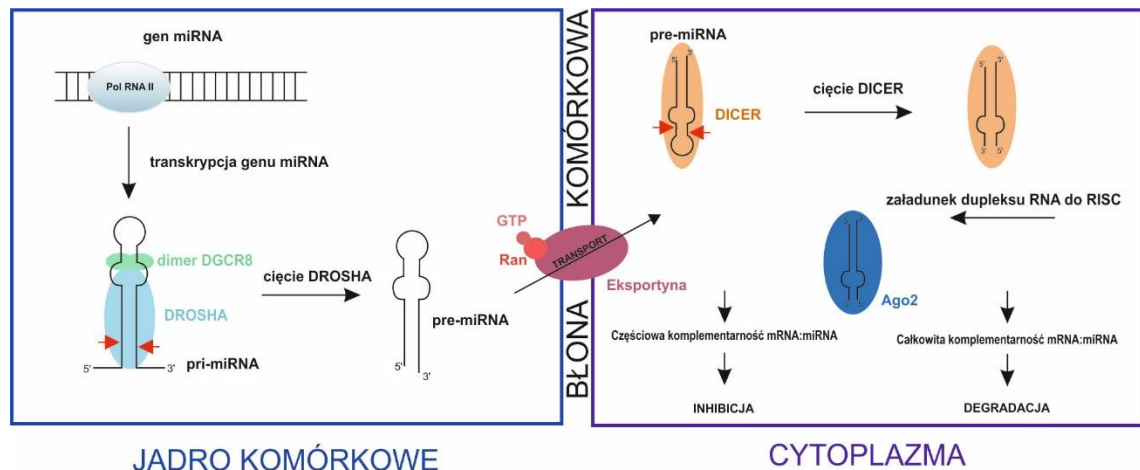
Ludzka rybonukleaza Dicer należy do klasy rybonukleaz III (RNaz III) i posiada dwie domeny RNazy III tworzące dimer, w obrębie którego znajduje się centrum katalityczne [4]. RNazy III są endorybonukleazami specyficznymi wobec dwuniciowych RNA (dsRNA, ang. *double-stranded RNA*), a hDicer znana jest głównie z kluczowej roli w procesie biogenezy miRNA [5].

### 1.1.1 Biogeneza miRNA

miRNA to grupa endogennych, niekodujących RNA o długości około 21-25 nukleotydów (nt), których obecność została potwierdzona w komórkach roślin, grzybów i zwierząt [6]. Rolą cząsteczek miRNA jest potranskrypcyjna regulacja ekspresji genów. Geny kodujące miRNA stanowią kilka procent wszystkich genów u zwierząt [7]. W przypadku człowieka szacuje się, że ponad 60% wszystkich genów kodujących białka jest regulowane przez miRNA [8]. Wspomniane geny miRNA mogą występować w obrębie egzonów oraz intronów [9] i mogą być zorganizowane w skupiska (tzw. klastry) oraz transkrybowane jako policistronowe jednostki [10].

Ścieżka kanonicznej biogenezy miRNA składa się z kilku etapów. Pierwszym z nich jest transkrypcja genów miRNA w jądrze komórkowym, w którą zaangażowana jest polimeraza RNA II (Pol II) [11]. Przeważnie większość miRNA transkrybowana jest z intronów, lecz mogą one także powstawać z egzonów, czy rejonów międzygenowych, a proces ten odbywa się niezależnie od transkrypcji genów gospodarza [12, 13]. Następnie powstający pierwotny transkrypt miRNA (pri-miRNA, ang. *primary miRNA*) jest rozpoznawany i wiązany przez kompleks białek DGCR8 (ang. *DiGeorge Syndrome Critical Region 8*) oraz rybonukleazę Drosha (tzw. kompleks Mikroprocesora). Pri-miRNA jest cięty przez rybonukleazę Drosha do pre-miRNA (ang. *precursor miRNA*) o długości około 60 nt [14]. Powstające cząsteczki pre-miRNA są transportowane do cytoplazmy przez Eksportynę 5 i czynnik Ran, który wiąże i hydrolizuje GTP [15]. W cytoplazmie cząsteczka pre-miRNA jest rozpoznawana i cięta przez rybonukleazę Dicer, która generuje dupleks miRNA/miRNA\* o długości ok. 21 pz. Dupleks ten jest następnie przekazywany do kompleksu RISC (ang. *RNA-induced silencing complex*). W kompleksie RISC jedna z nici, tzw. nić pasażerska\* (miRNA\*, ang. *passenger strand\**), jest odrzucana i ulega degradacji, a druga nić, tzw. nić wiodąca (ang. *guide strand*), pozostaje związana z kompleksem [16]. Aktywny kompleks RISC wiąże się zazwyczaj z docelowym mRNA w obrębie rejonu 3' UTR i powoduje degradację mRNA lub represję

translacji, w zależności od stopnia komplementarności pomiędzy miRNA a mRNA. W przypadku całkowitej komplementarności pomiędzy mRNA a miRNA następuje degradacja docelowego mRNA. Natomiast kiedy komplementarność ta jest tylko częściowa obserwujemy inhibicję translacji [6]. W przypadku ssaków większość mRNA ulega inhibicji translacji, a mniej niż 5% wszystkich sekwencji docelowych mRNA ulega degradacji po związaniu z kompleksem RISC [17] (**Rysunek 1**).



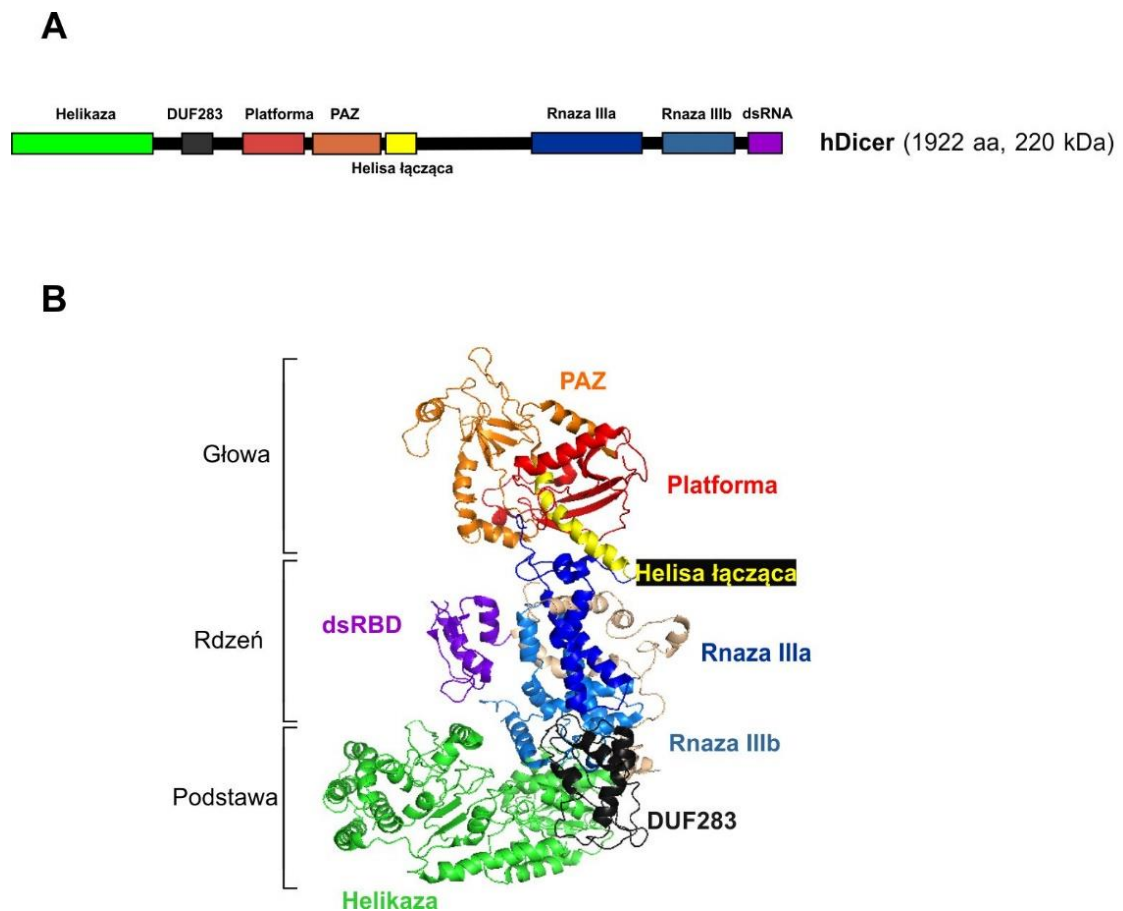
**Rysunek 1. Kanoniczny szlak biogenezy miRNA w komórce ssaczej.** Czerwonymi strzałkami zaznaczono miejsca cięcia białek Drosha oraz Dicer.

### 1.1.2 Budowa ludzkiej rybonukleazy Dicer

hDicer jest dużym (~220 kDa), wielodomenowym białkiem zbudowanym z 1922 aminokwasów (aa), kodowanym przez gen *DICER1*, który położony jest na długim ramieniu chromosomu 14. Gen *DICER1* uznawany jest za gen metabolizmu podstawowego (ang. *housekeeping gene*). Na szczególną uwagę zasługuje fakt, iż w porównaniu do innych genów tego typu, gen *DICER1* ma długi rejon 3' UTR (>4000 pz) [18]. Znanych jest wiele izoform mRNA hDicer, które generowane są m.in. w procesie alternatywnego składania pre-mRNA. Obecnie znanych jest siedem izoform mRNA, różniących się długością końców 5' oraz 3' UTR, z których powstaje pełnej długości białko hDicer [19].

hDicer zbudowana jest z następujących po sobie domen: domeny helikazowej, domeny DUF283 (ang. *domain of unknown function*), domeny Platformy i PAZ (ang. *Piwi Argonaut and Zwillie proteins*) oraz helisy łączącej, dwóch domen RNazowych (RIIIa i RIIIb) oraz C-końcowej domeny wiążącej dsRNA (dsRBD, ang. *double-stranded RNA binding domain*) [20] (**Rysunek 2**). Podobnie jak inne Dicer kręgowców, ludzka

rybonukleaza Dicer kształtem przypomina literę „L”. W przestrzennej organizacji białka można wyodrębnić tzw. głowę, rdzeń oraz podstawę białka [21, 22] (**Rysunek 2**). Rejon zwany głową tworzony jest przez domenę Platformy, PAZ oraz helisę łączącą, rdzeń enzymu stanowią domeny RNazy III, a podstawę białka współtworzą domena helikazy oraz DUF283. Badania strukturalne dowodzą, że podczas wiązania substratu dochodzi do licznych i znacznych zmian konformacyjnych obejmujących całą cząsteczkę białka, a duża elastyczność struktury enzymu umożliwia wiązanie RNA o zróżnicowanej budowie, tj. pre-miRNA oraz pre-siRNA (ang. *precursor siRNA*) [23, 24].



**Rysunek 2. Budowa ludzkiej rybonukleazy Dicer.** (A) Schematyczne przedstawienie budowy domenowej hDicer. (B) Struktura 3D białka hDicer zobrazowana w programie PyMOL na podstawie kodu 5ZAL z bazy danych PDB [25]. Rysunek na podstawie publikacji Szczepanska i in.[26].

### 1.1.2.1 Domena helikazy

Domena helikazy uważana jest za jeden z najbardziej konserwatywnych rejonów białek typu Dicer [27]. W obrębie tej domeny występuje charakterystyczny motyw DExD/H-box (ang. *DExD-H-box helicase*) odpowiedzialny za wiązanie i hydrolizę ATP.

Nazwa motywu pochodzi od tworzących go reszt aminokwasowych: kwasu asparaginowego (A), kwasu glutaminowego (E), dowolnego aminokwasu (x) oraz histydyny (H). Domena helikazy razem z domeną DUF283 tworzą tzw. podstawę białka (**Rysunek 2**), z którą to oddziałują partnerzy białkowi hDicer. Z domeną helikazową oddziałuje, między innymi białko TRBP (ang. *HIV-1 TAR RNA-binding protein*) oraz PACT (ang. *protein activator of PKR*) [28]. Obecnie wiadomo, że domena helikazy odgrywa ważną rolę w oddziaływaniach z pętlą apikalną pre-miRNA, umożliwiając tym samym dyskryminację pomiędzy pre-miRNA a pre-siRNA [29]. Dotychczas przeprowadzone analizy wykazały, że po usunięciu domeny helikazowej, hDicer nadal zachowuje swoją aktywność RNazową, a taki wariant delecyjny charakteryzuje się wyższą efektywnością cięcia kanonicznych substratów, w porównaniu do białka typu dzikiego [30]. Grupie prof. P. Svobody udało się pokazać, że domena helikazy utrzymuje Dicer w stanie „zamkniętym” [31]. Przejście do stanu „otwartego”, w którym to białko może związać i przeciąć substrat, stymulowane jest przez związanie się białka TRBP [31]. Dzięki usunięciu domeny helikazowej, białko Dicer pozostaje w „otwartej” konformacji i może wydajnie wiązać i procesować substraty pre-miRNA [31]. Inne badania wykazały również, że domena helikazy jest potrzebna do prawidłowego cięcia pre-miRNA o termodynamicznie niestabilnej strukturze [32].

### 1.1.2.2 Domena DUF283

Domena DUF283, razem z wcześniej wspomnianą domeną helikazową, tworzy podstawę białka hDicer (**Rysunek 1**). Funkcja jaką pełni ta domena wciąż pozostaje słabo poznana. Za pomocą spektroskopii magnetycznego rezonansu jądrowego (NMR, ang. *nuclear magnetic resonance*) oraz metody cryo-EM (ang. *cryoelectron microscopy*) uzyskano struktury 3D domeny DUF283 odpowiednio, roślinnego białka typu Dicer DCL4 [33] oraz hDicer [25]. Badania strukturalne ujawniły obecność motywu  $\alpha\beta\beta\alpha$  w domenie DUF283, czyli motywu charakterystycznego dla domen wiążących dsRNA. Jednakże wcześniej przeprowadzone badania w Instytucie Chemii Bioorganicznej Polskiej Akademii Nauk (IChB PAN), nie wykazały aktywności wiązania dsRNA, ani dsDNA, przez DUF283 hDicer [34]. Podobnie, badania prowadzone w grupie prof. Y.A. Yuan, nie wykazały, aby domena DUF283 DCL4 wiązała dwuniciowe substraty [35]. Wykazano natomiast, że domena DUF283 hDicer oddziałuje z jednoniciowymi RNA oraz DNA [34]. Ponadto, okazało się, że zarówno DUF283, jak pełnej długości hDicer, wspierają proces hybrydyzacji komplementarnych kwasów nukleinowych (ang. *nucleic*



*acid annealing activity*) [25, 36]. Badania zespołu prof. J. Doudny stanowią bogate źródło wiedzy na temat domeny DUF283. Dzięki tym badaniom udało się wykazać, że hDicer pozbawiona domeny DUF283 prezentuje mniejszą aktywność RNazową wobec prekursorów siRNA, w porównaniu do białka typu dzikiego [30]. Różnicy takiej nie zanotowano dla użytego w badaniach substratu typu pre-miRNA, pre-let-7 [30]. W innej pracy pokazano, że usunięcie domeny helikazowej oraz DUF283 powoduje utratę aktywności RNazowej hDicer, co sugeruje istotny wpływ obu tych domen na tę aktywność hDicer [26, 37].

W literaturze znaleźć można również informację, że domena DUF283 hDicer oddziałuje z białkiem ADAR1 (ang. *adenosine deaminase acting on RNA*), które to pozytywnie wpływa na wydajność cięcia prekursorów miRNA oraz przekazywanie miRNA do kompleksu RISC [36].

### 1.1.2.3 Domeny Platforma i PAZ

Domeny Platforma, PAZ oraz helisa łącząca tworzą tzw. kasetę PPC (ang. *PPC cassette*) i stanowią „głowę” białka (**Rysunek 2**), która odpowiada za kotwiczenie kanonicznych substratów białek typu Dicer. Kasetę PPC zawiera dwie kieszenie wiążące: kieszeń tworzoną przez domenę Platformy i PAZ, która wiąże koniec 5' substratów, tzw. kieszeń wiążącą 5' (ang. *5' binding pocket*), oraz kieszeń tworzoną przez domenę PAZ i wiążącą dwa niesparowane nukleotydy na końcu 3' substratów, tzw. kieszeń wiążącą 3' (ang. *3' binding pocket*). Ponadto, wyróżnić można jeszcze jeden motyw strukturalny w tym rejonie białka tzw. helisę łączącą, która w strukturze trzeciorzędowej białka ulokowana jest pomiędzy obiema kieszeniami wiążącymi i prawdopodobnie wspomaga proces przekazywania dupleksu miRNA lub siRNA do białka Ago [38]. Najnowsze badania grupy prof. N. Kim wskazują, że rearanżacja przestrzenna helisy łączącej jest konieczna, aby pre-miRNA prawidłowo dopasował się do doliny katalitycznej hDicer [24]. Mutacje wprowadzone w rejonie helisy łączącej lub delecja całej domeny spowodowały obniżenie aktywności RNazowej hDicer zarówno w warunkach *in vitro*, jak i *in cellulo* [24]. Rejon pomiędzy kasetą PPC i domenami RNazy III nazywany jest „linijką molekularną” (ang. *molecular ruler*), ponieważ dzięki związaniu substratu w obrębie kieszeni wiążących oraz odległości dzielącej kieszenie wiążące od domeny RNazy IIIa oraz RNazy IIIb, hDicer generuje dupleksy miRNA/siRNA o długości ~20-25 pz [23, 39]. Grupa prof. J. Doudny podzieliła białko hDicer na dwie części, pierwszą składającą się z domen DUF283 oraz PAZ (fragment DP) oraz drugą, składającą się

z domen RNazy III oraz dsRBD (fragment hDcr-C), tzw. rdzeń katalityczny. W reakcji cięcia kanonicznych substratów, fragment hDcr-C generował szereg produktów w przedziale długości od 12 do 22 nt, dopiero po dodaniu do reakcji fragmentu DP zaobserwowano pojawienie się produktów o długości odpowiadającej długości cząsteczek miRNA, tj. ok. 22 nt [40]. Warianty delecyjne białka hDicer pozbawione domeny PAZ lub kasety PPC, czy wariant pozbawiony domen: helikazy, DUF283 oraz PAZ również nie wykazywały aktywności RNazowej wobec kanonicznych substratów hDicer [41, 42].

#### 1.1.2.4 Domeny RNazy III i dsRBD

Dwie domeny RNazy III oraz domena dsRBD tworzą C-koniec hDicer (**Rysunek 2**). Centrum katalityczne białka stanowią dwie domeny RNazy III. Do zachowania swojej aktywności RNazowej hDicer, podobnie jak inne białka typu Dicer, wymaga obecności jonów dwuwartościowych. Preferowanym jonem jest jon magnezu ( $Mg^{2+}$ ), ale mogą to też być jony manganu, niklu, czy kobaltu. Rodzaj jonu dwuwartościowego może mieć wpływ na długość produktów generowanych przez Dicer [43, 44].

Domena dsRBD, podobnie jak DUF283, posiada motyw  $\alpha\beta\beta\beta\alpha$  charakterystyczny dla domen wiążących dsRNA. W przypadku dsRBD, motyw ten pozwala na wiązanie cząsteczek dsRNA niezależnie od ich sekwencji nukleotydowej, ponieważ proces wiązania substratów odbywa się głównie poprzez rozpoznanie specyficznych motywów strukturalnych w helisie A dsRNA [45]. Domena dsRBD pełni rolę domeny wspomagającej proces wiązania kanonicznych substratów hDicer [46]. Wykazano, że w przypadku usunięcia domeny dsRBD z fragmentu hDcr-C, taki wariant nie prezentował aktywności RNazowej, przypuszczalnie ze względu na nieefektywne wiązanie substratów [40].

#### 1.1.3 Kanoniczne substraty ludzkiej rybonukleazy Dicer

Kanonicznymi substratami dla hDicer są cząsteczki pre-miRNA oraz pre-siRNA, które posiadają niesparowany dwunukleotydowy koniec 3'-OH oraz grupę fosforanową na końcu 5' [47].

pre-miRNA to dość jednorodna pod względem długości i pochodzenia grupa cząsteczek RNA, które generowane są głównie przez kompleks Mikroprocesora w jądrze komórkowym, w procesie cięcia cząsteczek pri-miRNA [14]. Jest to grupa cząsteczek o długości ok. 60-70 nt, które przyjmują strukturę spinki do włosów złożoną z rejonu

trzonu (rejon dwuniciowy) oraz rejonu pętli apikalnej [48]. Niemniej jednak, struktury przyjmowane przez pre-miRNA mogą charakteryzować się większymi lub mniejszymi niesparowaniami w rejonie trzonu, w postaci pętli wewnętrznych i wybrzuszeń, czy też większą lub mniejszą pętlą apikalną, a także różnorodnością końców, co wpływa na proces dojrzewania pre-miRNA [49, 50]. Domena RNazy IIIa odpowiada za cięcie substratu pre-miRNA w obrębie ramienia 3', a domena RNazy IIIb przecina substrat w obrębie ramienia 5' [23].

Jak wspomniano powyżej, do kanonicznych substratów Dicer zaliczają się także cząsteczki pre-siRNA. Są to długie (>30 pz), w pełni sparowane dsRNA, które mogą zostać wprowadzone bezpośrednio do cytoplazmy, np. w wyniku infekcji wirusowej [51, 52]. Mogą być także wycinane z komplementarnych transkryptów przyjmujących lokalne struktury dwuniciowe [53]. Ponadto, syntetyczne pre-siRNA używane są np. w terapii genowej [54]. Ze względu na pochodzenie, cząsteczki siRNA można podzielić na egzogenne (egzo-siRNA) oraz endogenne (endo-siRNA) [55].

## 1.2 Białka oddziałujące z pre-miRNA

Cząsteczki pre-miRNA, po transporcie do cytoplazmy, mogą oddziaływać z całym szeregiem białek. Z doniesień literaturowych wiadomo, że z końcem 5' pre-miRNA oddziałuje metylotransferaza zależna od S-adenozylu-metioniny (BCDIN3D, ang. *BCDIN3 domain containing*), która odpowiedzialna jest, m.in., za dimetylację 5' monofosforanu pre-miRNA [56]. W wyniku działania tego enzymu zmniejsza się powinowactwo Dicer do pre-miRNA, co prowadzi do obniżenia poziomu produkcji dojrzałych miRNA [56]. Natomiast z końcem 3' cząsteczek pre-miRNA, które posiadają jednonukleotydowy niesparowany koniec 3', wiążą się białka z grupy TUT (ang. *Terminal Uridyl Transferase*). Białka te, dodając uracyl do końca 3' pre-miRNA prowadzą do powstania charakterystycznego dla kanonicznych substratów dwunukleotydowego niesparowanego końca 3', co promuje produkcję miRNA [57].

Z rejonem pętli apikalnej oddziałują, m. in., białka takie jak, KSRP, RBM3, TDP-43, czy Lin28A. Białko KSRP (ang. *KH-type splicing regulatory protein*) wiąże wiele cząsteczek RNA występujących w komórce, w tym pre-miRNA, rozpoznając sekwencje bogate w guaninę [58]. Prawdopodobnie wiązanie KSRP do pre-miRNA ułatwia przyłączanie się Dicer do powstałego kompleksu, co promuje proces cięcia pre-miRNA [59]. Kolejne dwa białka, RBM3 (ang. *RNA binding motif protein 3*) oraz TDP-43 (ang.

*TAR DNA-binding protein-43*), w wyniku oddziaływań z Dicer oraz pre-miRNA, również promują dojrzewanie miRNA [60, 61]. Białko Lin28A odpowiada za stabilizację struktur mRNA w procesie translacji, ale może również oddziaływać z motywami: GNGAY oraz GGAG obecnymi w części apikalnej pre-miRNA, co prowadzi do rekrutacji UTP-transferazy, a następnie do poliurydylicacji i degradacji pre-miRNA [62, 63].

### **1.3 Białka oddziałujące z Dicer**

Chociaż białko hDicer może samodzielnie procesować pre-miRNA oraz pre-siRNA, to jego aktywność może być modulowana przez inne białka, tzw. „białka partnerujące”. Najbardziej znanymi białkami oddziałującymi z hDicer są dwa strukturalnie podobne białka: TRBP oraz PACT [37, 64]. Oba białka posiadają trzy domeny dsRBD, a domena występująca na ich C-końcu odpowiada za interakcję z domeną helikazową hDicer [42, 65, 66]. Zarówno TRBP, jak i PACT wiążąc się z hDicer wpływają na jej specyficzność substratową oraz aktywność RNazową [67]. Stwierdzono, że hDicer w połączeniu z PACT wykazuje większą selektywność wobec substratów i preferencyjnie przecina pre-miRNA, a nie pre-siRNA [64]. Ta dyskryminacja substratowa jest mniej zauważalna, gdy hDicer znajduje się w kompleksie z TRBP [64]. Zauważono także, że białko TRBP promuje wiązanie pre-miRNA przez hDicer w środowisku, w którym występują cząsteczki o podobnej strukturze spinkowej [68]. Białka TRBP i PACT mogą również wpływać na wzór cięcia pre-miRNA przez hDicer, powodując powstawanie miRNA różniących się długością (tzw. izomiRów) [64, 69]. Dodatkowo, TRBP może stabilizować Dicer oraz wspomagać proces przekazywania miRNA do kompleksu RISC [70, 71]. Wydaje się, że podobną rolę może pełnić także PACT, choć jego rola w tym procesie jest słabo poznana [42].

Innymi przykładami białek współpracujących z hDicer są ADAR1 oraz RHA (ang. *RNA helicase A*). ADAR1 oddziałuje z domeną DUF283 hDicer, dzięki czemu zwiększona jest efektywność cięcia pre-miRNA. Białko to wspiera również przekazywanie miRNA do kompleksu RISC [36]. Białko RHA zaangażowane jest w bezpośrednie oddziaływanie z Dicer, TRBP, Ago2 oraz miRNA i przypuszczalnie uczestniczy w rozplataniu dupletu miRNA/miRNA\* [72]. Na uwagę zasługuje również 5-lipooksygenaza (5-LO), która oddziałując z C-końcową częścią hDicer, wpływa na długość produkowanych przez nią RNA. W wyniku oddziaływania z 5-LO, hDicer produkuje RNA o długości 10-55 nt [73, 74].

#### 1.4 Rola Dicer poza szlakiem biogenezy miRNA: różnorodność substratów i funkcji Dicer

W ostatnich latach odnotowuje się coraz więcej doniesień naukowych opisujących zaangażowanie białek Dicer w inne procesy komórkowe niż ścieżka biogenezy srRNA [75].

Do „niekanonicznych” funkcji Dicer zaliczyć można na przykład udział hDicer w naprawie pęknięć DNA, w jądrze komórkowym. Badania *in cellulo* dowodzą, że pęknięcia w obrębie dsDNA (DSBs, ang. *DNA double-stranded breaks*) indukują fosforylację Dicer, powodując jej translokację do jądra komórkowego gdzie bierze ona udział w naprawie DNA i odpowiedzi na powstałe uszkodzenia [76].

Ponadto, pojawiają się również doniesienia mówiące o udziale rybonukleaz Dicer w odpowiedzi przeciwwirusowej. Wykazano, że izoforma hDicer pozbawiona egzonu 7 oraz 8, powstająca w wyniku alternatywnego składania pre-mRNA, jest zdolna do przecinania wirusowego dsRNA i generowania siRNA w celu ochrony komórek macierzystych przed infekcją wirusową [77]. W pracy Montavon i in. również wykazano, że hDicer, dzięki obecności domeny helikazowej, może oddziaływać z białkami wiążącymi dsRNA, biorącymi udział w odpowiedzi przeciwwirusowej, przykładowo z białkiem DHx9 (ang. *DExH-Box Helicase 9*) oraz PKR (ang. *protein kinase R*) [78].

Wiele z powyżej opisanych funkcji jest niezależnych od aktywności rybonukleazowej hDicer. W publikacji Rybak-Wolf i in. wykazano, że białko hDicer może wiązać się z komórkowymi RNA w sposób „pasywny”, czyli bez przeprowadzania cięcia substratu [79]. Oddziaływania takie występują, przykładowo, pomiędzy hDicer a mRNA oraz długimi niekodującymi RNA (lncRNA, ang. *long non-coding RNA*), ale także z cząsteczkami transportującymi RNA (tRNA, ang. *transfer RNA*), czy małymi jądrerkowymi RNA (snoRNA, ang. *small nucleolar RNA*). W wyniku tych oddziaływań obserwuje się obniżenie poziomu komórkowych miRNA, co w konsekwencji prowadzi do zmian w poziomach ekspresji genów. Sugeruje się, że domeną odpowiedzialną za pasywne wiązanie substratów jest domena helikazowa hDicer [79]. Niemniej jednak, znaczenie pasywnego wiązania hDicer do komórkowych RNA pozostaje wciąż nieznanne.

W literaturze znaleźć również można publikacje mówiące o udziale hDicer w procesie biogenezy cząsteczek tRF (ang. *tRNA-derived fragments*), które powstają z tRNA [80]. Cząsteczki tRF powstają zarówno w komórkach eukariotycznych, jak

i prokariotycznych [81] pod wpływem różnych czynników, takich jak np. stres czy niedotlenienie [82, 83]. Pełnią one wiele funkcji biologicznych, działając jako cząsteczki sygnalizacyjne w odpowiedzi na stres komórkowy oraz jako regulatory ekspresji genów. Ponadto uważa się, że są one zaangażowane w metabolizm RNA, proliferację komórek, modulację odpowiedzi na uszkodzenia DNA i neurodegenerację [84]. Niektóre cząsteczki tRF działają podobnie do miRNA, a liczne badania dowiodły, że są one generowane w sposób zależny od Dicer [85, 86]. W pracy Maute i in. wykazano, że pewien tRF generowany przez hDicer moduluje proliferację i odpowiedź komórkową na uszkodzenia DNA w chłoniaku komórek B [82].

Znaleźć można również prace opisujące rolę hDicer w generowaniu srRNA o długości 20-25 nt z cząsteczek snoRNA [87]. Małe jąderkowe RNA to wysoce konserwatywna klasa RNA, która występuje u wszystkich eukariontów [88]. W pracy Ender i in. wykazano, że z cząsteczki snoRNA ACA45, w wyniku cięcia przez hDicer, może powstać funkcjonalna cząsteczka siRNA, a następnie cząsteczka taka jest wiązana przez kompleks RISC i bierze udział w potranskrypcyjnym wyciszaniu ekspresji genów [87]. Kolejnym przykładem jest snoRNA U3, które w wyniku cięcia przez hDicer staje się cząsteczką U3-miRNA i bierze udział w wyciszaniu ekspresji genu *SNX27* (ang. *sorting nexin-27*), z którego powstaje białko zaangażowane w transport wewnątrzkomórkowy [89].

Na podstawie powyższych przykładów „niekanonicznych” funkcji hDicer, można stwierdzić, że substratami Dicer są nie tylko pre-miRNA oraz pre-siRNA, ale także inne cząsteczki RNA obecne w komórkach.

## 2. Cel pracy

Wcześniejsze badania przeprowadzone w naszym zespole badawczym wykazały, że domena DUF283 hDicer wiąże jednoniciowe RNA oraz DNA, nie wiąże natomiast substratów dwuniciowych. Dalsze badania ujawniły, że zarówno sama domena DUF283, jak i pełnej długości hDicer wspierają parowanie komplementarnych cząsteczek kwasów nukleinowych. W ramach realizacji niniejszej pracy doktorskiej, w oparciu o dotychczas zgromadzone wyniki badań postawiono hipotezę, że domena DUF283 jest kluczowa dla aktywności hDicer wspierającej proces hybrydyzacji komplementarnych kwasów nukleinowych.

W celu weryfikacji postawionej hipotezy wyznaczono następujące zadania badawcze:

- stworzenie konstruktów genetycznych kodujących warianty hDicer pozbawione domeny DUF283 (warianty  $\Delta$ DUF hDicer),
- zbadanie aktywności RNazowej otrzymanych wariantów  $\Delta$ DUF hDicer wobec wybranych kanonicznych substratów hDicer, w warunkach *in vitro* oraz *in cellulo*,
- zbadanie aktywności wspierania parowania cząsteczek kwasów nukleinowych wariantów  $\Delta$ DUF hDicer.

Kolejnym celem badawczym było ustalenie znaczenia domeny DUF283 hDicer dla wiązania puli komórkowych RNA przez wspomnianą rybonukleazę.

Osiągnięcie tego celu wymagało realizacji kolejnych zadań badawczych:

- izolacji komórkowych kompleksów hDicer•RNA oraz  $\Delta$ DUF hDicer•RNA z wybranych linii komórkowych,
- identyfikacji pul RNA wiązanych przez hDicer oraz warianty  $\Delta$ DUF hDicer,
- analizy danych, w tym analizy różnicowej zbiorów danych dotyczących pul RNA wiązanych przez hDicer oraz warianty  $\Delta$ DUF hDicer.

Wyniki przeprowadzonych badań ujawniły, że wśród licznych RNA wiązanych przez hDicer znajdują się cząsteczki RNA bogate w ciągi reszt guanozynowych, potencjalnie przyjmujące struktury G-kwadrupleksu. Wyniki naszych wcześniejszych badań, prowadzonych w warunkach *in vitro*, także wskazywały, że hDicer może oddziaływać z cząsteczkami przyjmującymi struktury G-kwadrupleksu. Zgromadzone dane zrodziły kolejne pytania i cele badawcze zmierzające do ustalenia charakteru oddziaływań

pomiędzy hDicer i cząsteczkami kwasów nukleinowych przyjmującymi struktury G-kwadrupleksu.

Osiągnięcie tego celu wymagało realizacji następujących zadań:

- zbadania, czy domeny hDicer rozpoznające i kotwiczące kanoniczne substraty hDicer (rejon tzw. kasety PPC), mogą także oddziaływać z RNA i DNA przyjmującymi struktury G-kwadrupleksu,
- scharakteryzowania oddziaływań pomiędzy domenami hDicer wiążącymi G-kwadrupleksy RNA i DNA, a wybranymi cząsteczkami kwasów nukleinowych przyjmującymi struktury G-kwadrupleksu,
- zaproponowania, w oparciu o uzyskane wyniki, a także wcześniej przeprowadzone badania, potencjalnego biologicznego znaczenia oddziaływań pomiędzy hDicer a cząsteczkami przyjmującymi struktury G-kwadrupleksu.



### 3. Materiały i metody

#### 3.1 Materiały

##### 3.1.1 Oligonukleotydy

Stosowane w badaniach cząsteczki RNA oraz DNA zostały zakupione w firmach Genomed S.A. oraz FutureSynthesis Sp. z o.o. Cząsteczki RNA i DNA G-kwadrupleksów zostały zsyntetyzowane przez prof. dr hab. Ryszarda Kierzka (Zakład Chemii i Biologii Kwasów Nukleinowych IChB PAN). Sekwencje wszystkich oligomerów przedstawiono w **Tabeli 1** oraz **Tabeli 2**.

**Tabela 1.** Sekwencje oligonukleotydów wykorzystywanych w badaniach.

Nazwa	Sekwencja 5'-> 3'	Długość
RNA_sense*	GUGCAUUGUAGUUGCAUUGCAUGUUCUGG UCA	32 nt
RNA_ov	ACCAGAACAUGCAAUGCAACUACAAUGCA CAU	32 nt
pre-mir-16-1	UAGCAGCACGUA <del>AAU</del> AUUGGCGUUAAGAU UCU <del>AAAAU</del> UAUCUCCAGUAUUAACUGUGC UGCUGAA	65 nt
pre-mir-21	AGCUUAUCAGACUGAUGUUGACUGUUGAA UCUCAUGGCAACACCAGUCGAUGGGCUGU	58 nt
RNA21	<u>UCGAAGUAUUCGCGUACGUG</u>	21 nt
RNA50	GGUUGAACUAUUUCGUGUAUCUGGAA <u>CA</u> <u>CGUACGCGGAAUACUUCGAUU</u>	50 nt
f_ΔDUF(625-752)	ATTCCAGAGTGT <del>TTG</del> AGGGATAGTTATCCC AGACCTG	37 nt
r_ΔDUF(625-752)	TCGTGGACCACCATCGTCAGGCCTCAACAC	30 nt
f_ΔDUF(630-709)	GACCATTTGATGCCAGTTGGGAAAGAGACT	30 nt
r_ΔDUF(630-709)	GGCCGTGTTGATTGTGACTCGTGGACCAC	29 nt
P1-DICER	ATGAAAAGCCCTGCTTTGC	20 nt
P3-DICER	TATCAGACTGTCGTGCCG	18 nt
P4-DICER	TGGAAAGAAGATACACAGC	19 nt
P6-DICER	ACTCAACTTTAGAAGGCGG	19 nt
P8-DICER	CTGCGGATTTTAGATACCC	19 nt

P10-DICER	AGCAACACAGATAAAATGGG	19 nt
P12-DICER	TGGGAGATGCGATTTTGG	18 nt
P13-DICER	GGTCCATCTCTCTTTTGT	20 nt
r_PAZ	ACTGAGTGGGGTAAGATCAG	20 nt
f_miR-21-5p	GCTTATCAGACTGATGTTGAAA	22 nt
f_miR-16-1-5p	GCACGTAAATATTGGCGAA	19 nt
TER10	GGGUUAGGGU	10 nt
TER22	AGGGUUAGGGUUAGGGUUAGGG	22 nt
TEL22	AGGGTTAGGGTTAGGGTTAGGG	22 nt
G4U4G4	GGGGUUUUGGGG	12 nt
G4T4G4	GGGGTTTTGGGG	12 nt
LIN32	GUGCAUUGUAGUUGCAUUGCAUGUUC UGGUCA	32 nt
adapter	/5Phos/NNAACNNNAGATCGGAAGAGCGTC GTGGATCCTGAACCGC	41 nt
starter P3 Solexa	CAAGCAGAAGACGGCATAACGAGATCGGTCT CGGCATTCTGCTGAACCGCTCTTCCGATCT	61 nt
starter P5 Solexa	AATGATACGGCGACCACCGAGATCTACT CTTCCCTACACGACGCTCTTCCGATCT	58 nt

\*Cząsteczka RNA\_sense jest komplementarna do cząsteczki RNA\_ov i razem tworzą cząsteczkę dsRNA\_ov, która posiada dwa niesparowane nukleotydy na końcu 3'. Komplementarna sekwencja pomiędzy cząsteczkami RNA21 a RNA50 została podkreślona.

**Tabela 2.** Sekwencje indeksów wykorzystanych w protokole irCLIP.

Numer indeksu_##	Sekwencja indeksu	Przypisana próba
irCLIP_ddRT_22	NNNN,GCTCT_0,WWW	WT hDicer (1)
irCLIP_ddRT_44	NNNN,CTGTA_0,WWW	WT hDicer (2)
irCLIP_ddRT_43	NNNN,AGTTA_0,WWW	WT hDicer (3)
irCLIP_ddRT_25	NNNN,CTGAT_0,WWW	WT hDicer (IgG)
irCLIP_ddRT_12	NNNN,TCCAC_0,WWW	hDicer (1)
irCLIP_ddRT_13	NNNN,CCGGA_0,WWW	hDicer (2)
irCLIP_ddRT_1	NNNN,ACACC_0,WWW	hDicer (3)
irCLIP_ddRT_14	NNNN,AGGCA_0,WWW	hDicer (IgG)
irCLIP_ddRT_24	NNNN,TGTAT_0,WWW	ΔDUF(625-752) (1)

irCLIP_ddRT_3	NNNN,CAACA_0,WWW	ΔDUF(625-752) (2)
irCLIP_ddRT_4	NNNN,CGATA_0,WWW	ΔDUF(625-752) (3)
irCLIP_ddRT_33	NNNN,CACAG_0,WWW	ΔDUF(625-752) (IgG)
irCLIP_ddRT_23	NNNN,AGACT_0,WWW	ΔDUF(630-709) (1)
irCLIP_ddRT_7	NNNN,TTGCG_0,WWW	ΔDUF(630-709) (2)
irCLIP_ddRT_50	NNNN,TGAAA_0,WWW	ΔDUF(630-709) (3)
irCLIP_ddRT_32	NNNN,TCGAG_0,WWW	ΔDUF(630-709) (IgG)

(##) numer indeksu

(x) numer powtórzenia biologicznego

(IgG) kontrola izolacji IgG

### **3.1.2 Bufory i roztwory barwiące**

#### Bufor do reakcji cięcia substratów przez hDicer i jej warianty delecyjne

2,5 mM MgCl<sub>2</sub>

50 mM NaCl

20 mM Tris-HCl, pH 7.5

#### 5-krotnie stężony bufor do badania oddziaływań białko-kwas nukleinowy

250 mM NaCl

100 mM Tris-HCl, pH 7.5

#### Bufor TBS

50 mM Tris-HCl, pH 7.5

150 mM NaCl

#### Bufor dla białka hDicer i jej wariantów delecyjnych

20 mM Tris-HCl, pH 7.5

50 mM NaCl

10% glicerol

0,25% Triton X-100

Bufor do elektroforezy białek 10×SDS-PAGE

30,2 g Tris

144,2 g glicyna

10 g SDS

dopełniono wodą do objętości 1 L

Bufor do elektrotransferu półsuchego

2,91 g Tris

1,46 g glicyna

20% metanol

10% SDS

dopełniono wodą do objętości 1 L

Bufor I do nakładania prób na żel poliakrylamidowy (LB I, ang. *loading buffer I*)

0,05% błękit bromofenolowy

0,05% ksylencyjanol

w 7 M wodnym roztworze mocznika

Bufor II do nakładania prób na żel poliakrylamidowy (LB II, ang. *loading buffer II*)

0,05% błękit bromofenolowy

0,05% ksylencyjanol

w 50% wodnym roztworze glicerolu

Bufor III do nakładania prób na żel typu SDS-PAGE 4-krotnie stężony (4×LB III, ang. *loading buffer III*)

250 mM Tris-HCl, pH 6.8

20% v/v β-merkaptioetanol

8% SDS

0,15% błękit bromofenolowy

40% glicerol

Bufor do elektroforezy w żelach poliakrylamidowych w warunkach natywnych i denaturujących

1 x TBE (Roth)

Bufor do lizy komórek PROT S-100

30 mM HEPES, pH 7.4

100 mM KCl

5 mM MgCl<sub>2</sub>

10% glicerol

0,5 mM DTT

0,2% NP40

1 x Complete EDTA-free Protease Inhibitor Coctail (Roche)

Roztwór *n*-metylomezoporfiryny IX (NMM) do wybarwiania w żelach PAA kwasów nukleinowych o strukturze G-kwadrupleksów

8,61 mM NMM w 0,2 M HCl

20 mM Tris-HCl, pH 7.4

100 mM KCl

1 mM EDTA

Bufor SYBR Gold<sup>®</sup> do wybarwiania kwasów nukleinowych w żelach PAA

1 x SYBR Gold<sup>®</sup> (Thermo Fisher Scientific)

w 1 x TBE (Roth)

Płynna pożywka LB

10 g pepton

5 g ekstraktu drożdżowego

5 g NaCl

dopełniono wodą do objętości 1 L

### Stała pożywka LB

1 g peptonu

0,5 g ekstraktu drożdżowego

0,5 g NaCl

1,5 g agaru

dopełniono wodą do 100 ml

### 2-krotnie stężony bufor podstawowy (BP)

50 mM Tris

600 mM NaCl

10% glicerol

0,2% Triton X-100

### Bufer do lizy osadu bakteryjnego

1x bufor podstawowy

0,02 M imidazol

10% Cell Lytic B Lysis Reagent (Sigma)

1 tabletka inhibitora proteaz, cOmplete™, EDTA-free Protease Inhibitor Cocktail (Roche)

1% lizozym

### Bufer dla białka 5'PM PPC hDicer

25 mM Tris

300 mM NaCl

0,1% Triton X-100

5% glicerol

Bufor do lizy komórek w protokole irCLIP

50 mM Tris-HCl, pH 7.4

100 mM NaCl

1% Igepal CA-630 (Sigma)

0,1% SDS

0,5% deoksyholan sodu

Bufor do płukania z wysoką solą I

50 mM Tris-HCl, pH 7.4

50 mM NaCl

1 mM EDTA

1% Ipegal CA-630 (Sigma)

0,1% SDS

0,5% deoksyholan sodu

Bufor do płukania PNK

20 mM Tris-HCl, pH 7.4

10 mM MgCl<sub>2</sub>

0,2% Tween-20 (Sigma)

5 x bufor PNK, pH 6.5

350 mM Tris-HCl, pH 6.5

50 mM MgCl<sub>2</sub>

5 mM DTT

4-krotnie stężony bufor do ligacji

200 mM Tris-HCl, pH 7.4

40 mM MgCl<sub>2</sub>

4 mM DTT

#### Bufor PK + SDS

10 mM Tris-HCl, pH 7.4

100 mM NaCl

1 mM EDTA

0,2% SDS

#### Bufor NT2

50 mM Tris-HCl, pH 7.4

150 mM NaCl

1 mM MgCl<sub>2</sub>

0,0005% Ipegal CA-630 (Sigma)

#### Bufor do płukania z wysoką solą II

15 mM Tris-HCl, pH 7.4

120 mM NaCl

25 mM KCl

5 mM EDTA

1% Triton X-100 (Sigma)

0,001% SDS

1% deoksycholalan sodu

#### Bufor do elucji cząsteczek z żelu w protokole irCLIP

500 mM NaCl

1 mM NaCl

0,05% SDS

### **3.1.3 Materiał biologiczny**

W badaniach wykorzystano linię komórkową typu dzikiego HEK 293T oraz HEK 293T NoDice, czyli komórki typu knock-out z delecją w obrębie sekwencji kodującej domenę helikazową w genie *DICER1*. Komórki zostały uzyskane dzięki uprzejmości prof. B. R. Cullena [19]. Ponadto, w badaniach wykorzystano także chemikompetentne komórki



bakteryjne *Escherichia coli* (*E. coli*) Subcloning Efficiency™ DH5α oraz One Shot™ BL21 Star™ (DE3) (Invitrogen).

### **3.1.4 Żele poliakrylamidowe (PAA)**

Żele analityczne do rozdziału elektroforetycznego kwasów nukleinowych w warunkach denaturujących

10% lub 15% akrylamid/*N,N'*-bisakrylamid w stosunku 29:1, 7 M mocznik w buforze 1×TBE. Katalizatory polimeryzacji: 1000 μl 10% nadsiarczanu amonu, 100 μl TEMEDu na 100 ml żelu.

Żele preparatywne do rozdziału elektroforetycznego kwasów nukleinowych w warunkach denaturujących

8% akrylamid/*N,N'*-bisakrylamid w stosunku 29:1, 7 M mocznik w buforze 1×TBE. Katalizatory polimeryzacji: 500 μl 10% nadsiarczanu amonu, 50 μl TEMEDu na 50 ml żelu.

Żele do badania oddziaływań RNA-RNA w warunkach natywnych

10% akrylamid/*N,N'*-bisakrylamid w stosunku 29:1, 5% glicerol w buforze 1×TBE. Katalizatory polimeryzacji: 500 μl 10% nadsiarczanu amonu, 50 μl TEMEDu na 50 ml żelu.

Żele do badania oddziaływań białko-RNA w warunkach natywnych

5% lub 10% akrylamid/*N,N'*-bisakrylamid w stosunku 29:1, 5% glicerol w buforze 1×TBE. Katalizatory polimeryzacji: 500 μl 10% nadsiarczanu amonu, 50 μl TEMEDu na 50 ml żelu.

### **3.1.5 Żele agarozowe**

Do analiz wykorzystano żele agarozowe o stężeniu 0,5-2%. Do izolacji kwasów nukleinowych z żelu agarozowego wykorzystano agarozę o niskiej temperaturze topnienia firmy Biotools-B&M Labs. S. A. rozpuszczoną w buforze 1×TAE (Merck). W pozostałych przypadkach wykorzystano agarozę firmy Syngen, rozpuszczoną w buforze 1×TBE (Roth). W celu wizualizacji kwasów nukleinowych, do żelu agarozowego, dodawano barwnik Gel View (Novazym), zgodnie z instrukcją dostarczoną przez producenta.

### 3.1.6 Plazmidy

W badaniach wykorzystano wektor SureVector (Agilent) niosący sekwencję cDNA hDicer×3Flag. Plazmid ten został stworzony przez mgr inż. Martę Wojnicką (Zakład Biochemii Rybonukleoprotein IChB PAN) [41]. Do stworzenia wariantów 5'PM PPC oraz 3'PM PPC, wykorzystanych w analizach dotyczących oddziaływań kieszeni wiążących kasety PPC hDicer z G-kwadruksami RNA oraz DNA wykorzystano plazmid pcDNA3.1-FLAG-Dicer-5'-pocket mutant (#51368, Addgene) oraz plazmid pcDNA3.1-FLAG-Dicer-3'-pocket mutant (#51369, Addgene). Oba plazmidy zawierają mutacje reszt aminokwasowych, które są istotne dla wiązania kanonicznych substratów hDicer [90].

### 3.1.7 Preparaty białkowe

W badaniach wykorzystano preparat domeny PPC hDicer (PPC) otrzymany w bakteryjnym systemie ekspresyjnym *E.coli* OneShot™ BL21 Star™, który został otrzymany przez dr Kingę Ciechanowską (Zakład Biochemii Rybonukleoprotein IChB PAN) oraz preparat 3'PM PPC hDicer (3'PM PPC), który został otrzymany przez mgr inż. Martę Wojnicką (Zakład Biochemii Rybonukleoprotein IChB PAN) [90].

### 3.1.8 Wzorce wielkości

W badaniach wykorzystano następujące wzorce wielkości:

- Dla analiz w żelach agarozowych wykorzystano marker wielkości Perfect Plus™ 1 kb DNA ladder (EURx).
- Dla analiz w żelach typu SDS-PAGE wykorzystano marker wielkości Perfect™ Tricolor Protein Ladder (EURx).
- Dla analiz w analitycznych żelach poliakrylamidowych przygotowano marker wielkości poprzez alkaliczną hydrolizę (AH) wybranych cząsteczek RNA.

### 3.1.9 Ważniejsze enzymy i odczynniki wykorzystane podczas badań

*Gotaq® DNA Polymerase* (Promega)

*Gel view* (Novazym)

*Herculase II fusion DNA Polymerase* (Agilent)

*RNase I* (Thermo Fisher Scientific)

*Turbo™ DNase* (Invitrogen)

*T4 Polynucleotide Kinase (PNK)* (New England BioLabs)

*T4 RNA Ligase I* (New England BioLabs)

*Proteinase K* (Thermo Fisher Scientific)

*RNase H* (New England BioLabs)

*RNase A* (Ambion)

*CircLigaseII* (Epicentre)

*Phusion HF Master Mix* (Thermo Fisher Scientific)

*PageBlue™ Protein Staining Solution* (Thermo Fisher Scientific)

Bufor TBE (Roth)

Modyfikowany bufor TEA (Merck)

*SYBR™ Gold Nucleic Acid Gel Stain* (Thermo Fisher Scientific)

Agarozy: Syngen Daily Agarose (Syngen) lub agarozą o niskiej temperaturze topnienia (Biotools-B&M Labs. S. A.)

*Rnase Inhibitor Murine* (Biolabs)

*Dharmafact kb* (Dhermacon)

*cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail* (Roche)

### **3.1.10 Ważniejsze komercyjnie dostępne zestawy wykorzystane w badaniach**

*SuperScript IV Reverse Transcriptase* (Invitrogen)

*Mir-X miRNA First-Strand Synthesis Kit* (Takara Bio)

*SureVector next-generation cloning kit* (Agilent)

*iTaq Universal SYBR Green Supermix* (Bio-Rad)

*SuperSignal™ West Pico PLUS Chemiluminescent Substrate* (Thermo Fisher Scientific)

*Zyppy Plasmid Miniprep Kit* (Zymo Research)

*GenElute™ Endotoxin-free Plasmid Midiprep Kit* (SIGMA)

*High Sensitivity D1000 ScreenTape® System* (Agilent)

*QuantiFlor® ONE dsDNA* (Promega)

*Syngen Gel Mini Kit* (Syngen)

## **3.2 Metody**

### **3.2.1 Elektroforeza w żelach poliakrylamidowych w warunkach denaturujących**

Rozdział elektroforetyczny kwasów nukleinowych w warunkach denaturujących prowadzono w żelach poliakrylamidowych (PAA) zawierających 7 M mocznik. Do celów preparatywnych stosowano 8% żele PAA o wymiarach 19×18,5 cm, a do celów analitycznych 15% żele PAA o wymiarach 19×18,5 cm lub 39×33 cm. Analizowane próby przed nałożeniem na żel denaturowano. W tym celu próby mieszano z równą objętością (w stosunku do objętości próby) buforu LB I oraz inkubowano przez 3 min w 95°C, a następnie przez 2 min na lodzie. Rozdział prowadzono w buforze 1×TBE (Roth). Przed nałożeniem prób, żel poddawano preelektroforezie:

- dla żeli o wymiarach 19×18,5 cm były to parametry 450 V, 10 mA, 25 W, 15 min,
- dla żeli o wymiarach 39×33 cm były to parametry 1500 V, 10 mA, 50 W, 15 min.

Natomiast elektroforezę prowadzono w następujących warunkach:

- dla żeli o wymiarach 19×18,5 cm były to parametry 450 V, 40 mA, 25 W, 1-2 h,
- dla żeli o wymiarach 39×33 cm były to parametry 1500 V, 40 mA, 50 W, 1,5-2 h.

Cząsteczki RNA oraz DNA znakowane izotopem  $\gamma(^{32}\text{P})\text{ATP}$  wizualizowano na kliszach rentgenowskich (żele preparatywne) lub przy użyciu skanera FLA-5100 (Fujifilm) (żele analityczne). Do analizy uzyskanych danych wykorzystano program Multi Gauge (Fujifilm).

### **3.2.2 Elektroforeza w żelach poliakrylamidowych w warunkach natywnych**

Rozdział elektroforetyczny kwasów nukleinowych oraz kompleksów białko•kwas nukleinowego prowadzono w żelach poliakrylamidowych w warunkach natywnych. W tym celu stosowano żele o wymiarach 19×18,5 cm. Analizowane preparaty przed nałożeniem na żel mieszano z buforem LB II, do końcowego stężenia glicerolu w próbce ok. 15%. Rozdział prowadzono w buforze 1×TBE (Roth). Przed nałożeniem prób, żel poddawano preelektroforezie (warunki rozdziału: 120-180 V, 30 min), a następnie elektroforezie właściwej (warunki rozdziału: 120-180 V, 6 do 10 h w 4°C). Kwasy

nukleinowe znakowane izotopem  $\gamma(^{32}\text{P})\text{ATP}$  wizualizowano przy użyciu skanera FLA-5100 (Fujifilm), a kwasy nukleinowe niewyznakowane izotopem  $\gamma(^{32}\text{P})\text{ATP}$  wizualizowano przy użyciu tego samego skanera, po wcześniejszym wybarwieniu żelu w roztworze barwiącym: SYBR Gold (wizualizacja całkowitego RNA przy długość fali 473 nm) lub NMM (wizualizacja RNA przyjmujących strukturę G-kwadrupleksów o topologii równoległej przy długości fali 532 nm). Do analizy uzyskanych danych wykorzystywano oprogramowanie Multi Gauge (Fujifilm).

### **3.2.3 Elektroforeza w żelach typu SDS-PAGE**

Rozdział elektroforetyczny białek w warunkach denaturujących prowadzono w żelach poliakrylamidowych zawierających SDS. Do celów analitycznych stosowano 8 lub 10% żełe PAA o wymiarach 8,5×8 cm. Analizowane próby (w objętości 15  $\mu\text{l}$ ) przed nałożeniem na żel denaturowano. W tym celu próby mieszano z 5  $\mu\text{l}$  buforu 4×LB III, do końcowego stężenia 1×LB III, oraz inkubowano przez 5 min w 95°C, a następnie przez 2 min na lodzie. Rozdział prowadzono w buforze do elektroforezy białek 1×SDS-PAGE przez 30 minut przy napięciu 80 V dla żelu zagęszczającego, a następnie 50-90 min przy napięciu 195 V dla żelu rozdzielającego. Wynik rozdziału udokumentowano z wykorzystaniem urządzenia UVITEC Alliance Q9 Advanced.

### **3.2.4 Elektroforeza w żelach agarozowych**

Rozdział elektroforetyczny prowadzono w 0,5-2% żelach agarozowych. Analizowane próby, przed nałożeniem na żel, mieszano z buforem do obciążania prób 6×Loading buffer BLUE (EURx), do stężenia końcowego buforu 1×. W ten sposób przygotowane próby rozdzielano pod napięciem 100-120 V przez 0,5-1 h w buforze 1×TBE (Roth) lub 1×TAE (Merck). Wynik rozdziału udokumentowano z wykorzystaniem urządzenia UVITEC Alliance Q9 Advanced.

### **3.2.5 Znakowanie radioizotopowe końca 5' oligonukleotydów**

Cząsteczki kwasów nukleinowych znakowano radioizotopowo przez fosforylację ich końców 5' izotopem  $\gamma(^{32}\text{P})\text{ATP}$  (Hartmann). Skład i warunki reakcji podano w **Tabeli 3**. Przed przygotowaniem mieszaniny reakcyjnej cząsteczkę denaturowano poprzez inkubację w 95°C przez 3 min, a następnie umieszczenie na lodzie. Po dodaniu pozostałych składników mieszaniny, próbę inkubowano w 37°C przez 15 min. Po zakończonej inkubacji próbę zmieszano z jedną objętością 7 M mocznika (w stosunku do objętości próby) i ponownie inkubowano w 95°C przez 3 min. Produkty reakcji

rozdzielano elektroforetycznie w 8% preparatywnym żelu PAA w warunkach denaturujących (Metody 3.2.1) i oczyszczono zgodnie z procedurą opisaną w punkcie Metody: 3.2.7. Cząsteczki o długości powyżej 20 nt oczyszczano także na kolumnach NucAway™, zgodnie z protokołem dostarczonym przez producenta (Ambion).

**Tabela 3.** Skład mieszaniny i warunki reakcji fosforylacji końca 5' oligonukleotydów.

Składnik	Ilość na 1 próbę
10×bufor A	1 µl
RNA lub DNA (10 µM)	1 µl
[ $\gamma^{32}\text{P}$ ] ATP (10 µCi/µl)	1 µl
Kinaza polinukleotydowa T4 (10 U/µl)	1 µl
woda	do 10 µl
inkubacja: 37°C, 15 min	
inaktywacja enzymu: 90°C, 5 min	

### 3.2.6 Znakowanie końca 5' oligonukleotydów barwnikami fluorescencyjnymi

Cząsteczki przyjmujące strukturę G-kwadrupleksu z dołączonym barwnikiem o-BMVC zostały przygotowane przez dr Dorotę Gudanis oraz dr. Daniela Baranowskiego (Zakład Biomolekularnego NMR IChB PAN). C6-amino zmodyfikowany oligonukleotyd został rozpuszczony w roztworze fosforanu sodu (0,1 M pH 7.0) i wymieszany w stosunku molowym 1:5 z nadmiarem estru o-BMVC-C<sub>3</sub>-NHS zawieszonym w DMSO. Mieszaninę inkubowano w 37°C w ciemności przez noc. Cząsteczkę o-BMVC-C<sub>3</sub> oddzielono od wolnego o-BMVC-C<sub>3</sub>-NHS i odsolono z wykorzystaniem metody strącania za pomocą 2% mieszaniny NaClO<sub>4</sub>/aceton. Czystość i jednorodność wyznakowanej cząsteczki o-BMVC-C<sub>3</sub> sprawdzono poprzez rozdział cząsteczek w 15% żelu denaturującym w buforze 1xTBE (Roth) (Metody 3.2.1). Cząsteczki znakowane barwnikiem o-BMVC wizualizowano przy użyciu skanera FLA-5100 (Fujifilm) przy długości fali 532 nm.

### 3.2.7 Oczyszczanie kwasów nukleinowych w żelach PAA

Cząsteczki kwasów nukleinowych rozdzielano w 8% preparatywnym żelu PAA zgodnie z procedurą opisaną w punkcie: Metody 3.2.1, następnie wizualizowano na kliszach rentgenowskich i wycinano z żelu. Wycięte fragmenty żelu zawierające cząsteczki kwasów nukleinowych o pożądanej długości rozdrabniano w próbówce typu Eppendorf, dodawano 400 µl 0,3 M octanu sodu (pH 5.5) i inkubowano w 4°C przez 16 h z wytrząsaniem 500 rpm. Następnie całość mieszaniny przeniesiono do próbówki Spin-

X Costar (Sigma) zawierającej filtr celulozowy o średnicy porów 0,22  $\mu\text{m}$  i krótko wirowano (16000 $\times$ g, 4°C, 1 min), aby odseparować fragmenty żelu od supernatantu. Przefiltrowany supernatant przeniesiono do nowej probówki typu Eppendorf i dodano 900  $\mu\text{l}$  zimnego 96% etanolu, a następnie wymrażano przez noc w -20°C. Po zakończeniu inkubacji mieszaninę wirowano (16000 $\times$ g, 4°C, 40 min), roztwór dekantowano, a osad przemyto 200  $\mu\text{l}$  zimnego 70% etanolu (4°C). Próby powtórnie wirowano (16000 $\times$ g, 4°C, 30 min), po usunięciu etanolu osad suszono i rozpuszczano w wodzie wolnej od RNaz i DNaz. Stężenie materiału znakowanego radioizotopowo doprowadzono do stężenia końcowego 10 000 cpm/ $\mu\text{l}$ . Homogenność materiału po oczyszczeniu sprawdzano poprzez elektroforezę w 15% PAA w warunkach denaturujących (Metody: **3.2.1**). Analizowane cząsteczki wizualizowano przy użyciu skanera FLA-5100 (Fujifilm).

### **3.2.8 Przygotowanie substratów dwuniciowych**

W celu przygotowania dwuniciowych substratów do reakcji cięcia, niewyznakowana nić (RNA\_sense) była mieszana z wyznakowaną radioizotopowo, komplementarną nicią (RNA\_ov) w stosunku 1:1. Hybrydyzacja była prowadzona w buforze do reakcji cięcia substratów poprzez podgrzanie mieszaniny do 95°C i powolne schłodzenie (spadek temperatury -1°C/min) do temperatury 21°C.

Następnie powstałe kompleksy były rozdzielane poprzez elektroforezę w 10% natywnym żelu PAA (Metody **3.2.2**) i wycinane z żelu w celu pozbycia się niezwiązanej frakcji substratów (Metody **3.2.7**). Analizowane cząsteczki wizualizowano przy użyciu skanera FLA-5100 (Fujifilm).

### **3.2.9 Przygotowanie plazmidów kodujących warianty delecyjne ludzkiej rybonukleazy Dicer**

Plazmidy kodujące warianty  $\Delta\text{DUF}(625-752)$  oraz  $\Delta\text{DUF}(630-709)$  hDicer zostały otrzymane z wykorzystaniem metody PCR. Wszystkie startery zostały zaprojektowane na podstawie sekwencji cDNA transkryptu 2 hDicer (NM\_030621.4). Matrycą w reakcji PCR był plazmid kodujący pełnej długości hDicer ze znacznikiem 3 $\times$ Flag na C-końcu białka, stworzony za pomocą komercyjnego systemu SureVector firmy Agilent (plazmid został otrzymany w Zakładzie Biochemii Rybonukleoprotein IChB PAN [41]). W reakcji namnażania wykorzystano startery, których sekwencje są podane w punkcie Materiały: **3.1.1** w **Tabeli 1** oraz polimerazę Herculase II Fusion DNA (Agilent). Pełen opis składników oraz warunki reakcji PCR podano w **Tabeli 4** oraz **Tabeli 5**. Otrzymane

produkt PCR (fragment plazmidu kodujący sekwencję białka hDicer z delecją 384 nt, kodujących domenę DUF283 (625-752 aa) wariantu  $\Delta$ DUF(625-752) oraz plazmid z delecją 240 nt, kodujących domenę DUF283 (630-709 aa) wariantu  $\Delta$ DUF(630-709)) był wycięty z żelu i oczyszczony za pomocą zestawu Syngen Gel Mini Kit (Syngen), a następnie fosforylowany na końcu 5' poprzez kinazę polinukleotydową T4 (PNK) (Thermo Fisfer Scientific, pełny skład mieszaniny przedstawiono w **Tabeli 6**), oraz 100 ng fosforylowanego produktu połączono za pomocą ligazy DNA T4 (Thermo Fisfer Scientific, pełny skład mieszaniny przedstawiono w **Tabeli 7**). Całość mieszaniny po ligacji wykorzystano do transformacji chemikompetentnych komórek bakteryjnych *E. coli* szczep DH5 $\alpha$  (Metody: **3.2.10**).

**Tabela 4.** Skład mieszaniny do reakcji PCR w celu otrzymania wariantów białka hDicer.

Składnik	Ilość na 1 próbę
matryca: plazmid hDicer $\times$ 3Flag (5 ng/ $\mu$ l)	10 $\mu$ l
starter f_ $\Delta$ DUF(625-752) / f_ $\Delta$ DUF(630-709) (10 $\mu$ M)	2,5 $\mu$ l
starter r_ $\Delta$ DUF(625-752) / r_ $\Delta$ DUF(630-709) (10 $\mu$ M)	2,5 $\mu$ l
mix dNTP	5 $\mu$ l
Polimeraza Herculase II Fusion DNA	1 $\mu$ l
Bufor 5 $\times$ Herculase II Reaction Buffer	10 $\mu$ l
DMSO	1 $\mu$ l
woda	do 50 $\mu$ l

**Tabela 5.** Warunki reakcji PCR w celu otrzymania wariantów białka hDicer.

Etap	Temperatura	Czas	Ilość cykli
Denaturacja wstępna	95°C	120 s	1
Denaturacja	95°C	20 s	
Przyłączanie starterów	68°C	20 s	30
Synteza	72°C	6 min	
Synteza końcowa	72°C	8 min	1
Przechowywanie	25°C	$\infty$	



**Tabela 6.** Skład mieszaniny do fosforylacji produktu PCR.

Składnik	Ilość na 1 próbę
matryca: produkt PCR	30 $\mu$ l (1,5 $\mu$ g)
Bufor 10 $\times$ stężony A	5 $\mu$ l
ATP (10 mM)	5 $\mu$ l
kinaza polinukleotydowa T4	2 $\mu$ l
woda	do 50 $\mu$ l
inkubacja 37°C, 25 min	
inaktywacja enzymu 75°C, 10 min	

**Tabela 7.** Skład mieszaniny do ligacji fosforylowanego produktu PCR.

Składnik	Ilość na 1 próbę
matryca: fosforylowany produkt PCR (10 ng/ $\mu$ l)	10 $\mu$ l
Bufor do ligacji 10 $\times$ stężony	2 $\mu$ l
Ligaza T4 DNA	1 $\mu$ l
woda	do 20 $\mu$ l
inkubacja 15°C, 18 h	
inaktywacja enzymu 70°C, 10 min	

### 3.2.10 Transformacja komórek bakteryjnych metodą szoku cieplnego (ang. *heat-shock*)

Do 50  $\mu$ l zawiesiny kompetentnych komórek *E. coli* (szczep DH5 $\alpha$  lub OneShot™ BL21 Star™ (Thermo Fisher Scientific)) dodano 10 ng plazmidowego DNA w objętości 1  $\mu$ l, całość inkubowano 30 min na lodzie. Następnie, w celu wprowadzenia plazmidu do komórek bakteryjnych, zawiesinę inkubowano przez 20 s w temperaturze 42°C, po czym schłodzono w lodzie. Do komórek dodano 900  $\mu$ l sterylnej pożywki LB oraz inkubowano w 37°C (1 h, 225 rpm). Po tym czasie komórki bakteryjne wysiano na stałe podłoże LB zawierające antybiotyk selekcyjny, ampicylinę o stężeniu 100  $\mu$ g/ml. Selekcja komórek bakteryjnych posiadających pożądany plazmid była możliwa dzięki obecności w wektorze SureVector fragmentu zapewniającego oporność na ampicylinę – gen AmpR, którego produkt pozwala na hydrolizę pierścienia  $\beta$ -laktamowego ampicyliny. Szalki ze stałym podłożem LB inkubowano przez noc w 37°C.

### 3.2.11 Weryfikacja kolonii bakteryjnych po transformacji metodą szoku cieplnego

W celu sprawdzenia czy kolonie bakteryjne posiadają pożądany plazmid, zastosowano metodę kolonijnej reakcji PCR, tzw. kolonijny PCR. W tym celu wybrano kilkanaście kolonii bakteryjnych, każdą z nich przeniesiono za pomocą sterylnej końcówki do pipety do osobnej próbówki do reakcji PCR zawierającej 5  $\mu$ l wody wolnej od DNaz i RNaz. Tą samą końcówką przeniesiono również część przygotowanej zawiesiny bakteryjnej na nowe podłoże stałe LB z ampicyliną (100  $\mu$ g/ml) i inkubowano przez noc w 37°C. Przygotowaną zawiesinę komórek bakteryjnych inkubowano przez 10 min w 98°C, w celu lizy komórek bakteryjnych. Po tym czasie dodano resztę składników mieszaniny reakcyjnej PCR, cały skład podano w **Tabeli 8**, a warunki reakcji w **Tabeli 9**. Uzyskane produkty reakcji PCR analizowano w 1% żelu agarozowym z dodatkiem GelView w buforze 1 $\times$ TBE (Roth) (Metody 3.2.4). Wynik rozdziału udokumentowano z wykorzystaniem urządzenia UVITEC Alliance Q9 Advanced. Następnie, kolonie bakteryjne zawierające pożądany plazmid, namnażano w płynnej pożywce LB z ampicyliną (100  $\mu$ g/ml) i izolowano plazmidowe DNA, zgodnie z punktem Metody: 3.2.12.

**Tabela 8.** Skład mieszaniny do reakcji PCR z wybranych kolonii bakteryjnych.

<b>Składnik</b>	<b>Ilość na 1 próbę</b>
matryca: lizat komórek bakteryjnych	5 $\mu$ l
starter P4-DICER (10 $\mu$ M)	1,25 $\mu$ l
starter r_PAZ (10 $\mu$ M)	1,25 $\mu$ l
Bufor 5 $\times$ GoTaq Green	5 $\mu$ l
Polimeraza GoTaq <sup>®</sup> DNA	0,05 $\mu$ l
dNTP	2,5 $\mu$ l
woda	do 25 $\mu$ l

**Tabela 9.** Warunki reakcji PCR z wybranych kolonii bakteryjnych.

<b>Etap</b>	<b>Temperatura</b>	<b>Czas</b>	<b>Ilość cykli</b>
Denaturacja wstępna	95°C	180 s	1
Denaturacja	95°C	30 s	
Przyłączanie starterów	52°C	30 s	30
Synteza	72°C	90 s	
Synteza końcowa	72°C	5 min	1
Przechowywanie	4°C	∞	

### **3.2.12 Izolacja plazmidowego DNA**

Płynną pożywkę LB z ampicyliną (100 µg/ml) o objętości 2 ml, zaszczerpiono pojedynczą kolonią pobraną z wcześniej przygotowanego stałego podłoża LB (Metody **3.2.11**) i hodowano przez noc w 37°C z wytrząsaniem 225 rpm. Materiał izolowano za pomocą zestawu Zyppy Plasmid Miniprep Kit (Zymo Research), zgodnie z protokołem dostarczonym przez producenta.

W przypadku izolacji plazmidu w większej skali, płynną pożywkę z ampicyliną (100 µg/ml) o objętości 2 ml zaszczerpiono pojedynczą kolonią i hodowano w 37°C z wytrząsaniem 225 rpm przez 4 godziny. Następnie 150 µl tej hodowli przeniesiono do świeżej pożywki LB z ampicyliną (100 µg/ml) o objętości 250 ml i hodowano przez noc w 37°C z wytrząsaniem 225 rpm. Materiał izolowano za pomocą zestawu GenElute™ Endotoxin-free Plasmid Midiprep Kit (SIGMA), zgodnie z protokołem dostarczonym przez producenta.

Stężenie oraz czystość wyizolowanych plazmidów analizowano za pomocą spektrofotometru NanoDrop®1000 (Thermo Fisher Scientific). Przygotowane w ten sposób plazmidy były sekwencjonowane metodą Sangera w firmie Genomed (sekwencje starterów P1-P13-DICER użytych do sekwencjonowania zawarte są w punkcie: Materiały **3.1.1** w **Tabeli 1**). Próby do sekwencjonowania zostały przygotowane zgodnie z wytycznymi firmy Genomed.

### **3.2.13 Hodowla i transfekcja komórek HEK 293T**

Adherentne komórki HEK 293T NoDice lub HEK 293T typu dzikiego hodowano w pożywce Dulbecco's Modified Eagle Medium (DMEM) (Gibco) suplementowanej 10% płodową surowicą bydlęcą (FBS) (EURx), penicyliną (100 U/mL), streptomycyną

(100 µg/mL) i pirogronianem sodu (1 mM). W celu posiania komórek, komórki najpierw odczepiano od podłoża stosując trypsynę (0,05%) (Thermo Fisher Scientific), a następnie poddawano sedymentacji poprzez krótkie wirowanie (200×g, 3 min). Komórki wysiewano w ilości  $1,5 \times 10^6$  komórek na butelkę hodowlaną o powierzchni 25 mm<sup>2</sup>. Przy ~70% konfluencji, komórki transfekowano odczynnikiem DharmaFECT kb DNA (Dharmacon™), zgodnie z protokołem producenta. Po 24 h inkubacji, komórki przenoszono do nowej butelki hodowlanej o powierzchni 175 mm<sup>2</sup>. Po 72 h od momentu transfekcji, komórki ponownie odczepiono od podłoża stosując trypsynę (0,05%) i poddano sedymentacji poprzez wirowanie (200×g, 3 min). Otrzymany osad komórkowy przemyto 5 ml 1×PBS (EURx) oraz ponownie poddano go sedymentacji poprzez wirowanie (200×g, 3 min). Po usunięciu roztworu PBS, osad komórkowy był przechowywany w temperaturze -80°C do dalszych analiz.

#### **3.2.14 Przygotowanie lizatów komórkowych**

Komórki przygotowane zgodnie z punktem: Metody 3.2.13 zawieszono w buforze do lizy komórek PROT S-100 (Materiały 3.1.2). Przygotowaną w ten sposób zawiesinę komórek homogenizowano przy użyciu strzykawki i igły o wymiarach 0.9×40 mm. Otrzymany lizat poddano wirowaniu (16 000×g, 5 min, 4°C), a następnie uzyskany supernatant przeniesiono do nowej probówki typu Eppendorf. Końcowy supernatant stanowił frakcję cytoplazmatyczną, która posłużyła do dalszych badań.

#### **3.2.15 Immunoprecypitacja wariantów delecyjnych hDicer oraz hDicer**

Przygotowany wcześniej lizat komórkowy (Metody 3.2.14) inkubowano przez noc w 4°C na wytrząsarce rolkowej ze złożem ANTI-FLAG<sup>®</sup>M2 Affinity Gel (Merck) w celu wyizolowania białka. Złoże przygotowano zgodnie z protokołem dostarczonym przez producenta. Po inkubacji złoże przemyto 5-krotnie roztworem TBS. Następnie, złoże ze związanym białkiem inkubowano z peptydem 3×Flag o stężeniu końcowym 100 µg/mL w obecności buforu dla białka Dicer i jej wariantów (Materiały 3.1.2) w celu uwolnienia białka ze złoża. Stężenie białka było ustalane w oparciu o krzywą standardową dla białka BSA zgodnie z metodą Bradforda (Metody 3.2.3). Wyizolowane białko było zagęszczane na filtrze typu Amicon<sup>®</sup> (Merck) z punktem odcięcia 100 kDa, a następnie przechowywane w wyżej wymienionym buforze, który wzbogacono w 40% glicerol w -20°C.

### 3.2.16 Analiza typu Western Blot

Uzyskane frakcje białkowe były separowane w 8% żelu typu SDS-PAGE zgodnie z punktem: Metody 3.2.3. Następnie za pomocą elektrotransferu półsuchego, białko przeniesiono na membranę PVDF (Thermo Fisher Scientific). Bibuły Whatmann namoczono w 1×buforze do transferu półsuchego (Materiały 3.1.2) na około 2 min. Następnie, w dolnej kasetce (stanowiącej anodę) aparatu do półsuchej metody transferu (Trans-Blot Turbo, BioRad), umieszczono kolejno: warstwę bibuły, membranę PVDF, żel, warstwę bibuły i zamknięto górną kasetką (stanowiącą katodę). Tak przygotowaną membranę na czas transferu (90 minut) umieszczono w aparacie z ustawieniami stałego natężenia prądu wynoszącym 100 mA. Po zakończonym transferze membranę blokowano w 5% roztworze mleka.

Następnie przeprowadzono detekcję białek za pomocą przeciwciał. Dla białka hDicer i jej delecyjnych wariantów, membrany inkubowano z pierwszorzędowym, monoklonalnym mysim przeciwciałem anti-Dicer skierowanym na C-koniec białka w stosunku 1:300 (Santa Cruz Biotechnology) przez 15 min. W przypadku  $\beta$ -aktyny membranę inkubowano z pierwszorzędowym, króliczym przeciwciałem anti- $\beta$ -aktyna w stosunku 1:1100 (Cell Signaling Technology) przez 15 min. Następnie membrany inkubowano z drugorzędowymi przeciwciałami skierowanymi na IgG mysie oraz królicze (Jackson ImmunoResearch Laboratory), połączonymi z peroksydazą chrzanową (HRP) w stosunku 1:5000 przez 20 min. Obrazownie wykonano za pomocą metody chemiluminescencji, w tym celu wykorzystano zestaw SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific). Analizy typu Western Blot przeprowadzono na urządzeniu SNAP i.d.® 2.0 (Merck), zgodnie z protokołem producenta. Wynik analizy udokumentowano z wykorzystaniem urządzenia UVITEC Alliance Q9 Advanced.

### 3.2.17 Cięcia substratów przez warianty delecyjne hDicer

Reakcje cięcia były przygotowane w 10  $\mu$ l reakcji i zawierały bufor do reakcji cięcia (Materiały 3.1.2), cząsteczkę wyznakowaną na końcu 5' z użyciem [ $\gamma^{32}$ P]-ATP (10 000 cpm/ $\mu$ l, ok. 5 nM) oraz 18 nM odpowiedniego białka (skład reakcji opisano w **Tabeli 10**). Przygotowano także szereg reakcji kontrolnych, kontrolę ujemną (K-), która zawierała cząsteczkę zawieszoną w buforze reakcyjnym; a także reakcję kontrolną z dodatkiem EDTA, który chelatuje jony magnezu, o końcowym stężeniu 25 mM. Reakcje były inkubowane w 37°C przez 10, 30, 60 lub 120 minut z dodatkiem komercyjnego inhibitora RNaz (Eurx). Po zakończeniu inkubacji, w celu zatrzymania

reakcji cięcia, do mieszaniny dodano 10 µl buforu LB I (Materiały 3.1.2). Całość mieszaniny inkubowano w 95°C przez 3 min, a następnie 2 min na lodzie. Próby rozdzielano w 15% denaturującym żelu PAA (Metody 3.2.1). Analizowane cząsteczki wizualizowano przy użyciu skanera FLA-5100 (Fujifilm). Dla tego eksperymentu wykonano przynajmniej 3 powtórzenia techniczne.

**Tabela 10.** Skład mieszaniny do reakcji cięcia substratów.

<b>Składnik</b>	<b>Ilość na 1 próbę</b>
wyznakowany [ $\gamma^{32}\text{P}$ ] substrat (10 000 cpm/µl, 5 nM)	1 µl
5×bufor do reakcji cięcia substratów	2 µl
Mg <sub>2</sub> Cl (25 mM)	1 µl
inhibitor RNaz RNaseOUT (40 u/µl) (Invitrogen) (40 u/µl)	0,4 µl
wybrany preparat białkowy	x µl
EDTA (0,5 M)	1 µl
woda	do 10 µl
inkubacja 37°C	
inaktywacja enzymu poprzez dodanie 10 µl buforu LB I	

### **3.2.18 Badanie oddziaływań białko-kwas nukleinowy w teście spowolnienia migracji w żelach poliakrylamidowych (EMSA)**

Wyznakowaną z użyciem [ $\gamma^{32}\text{P}$ ]-ATP cząsteczkę (10 000 cpm/µl, ok. 5 nM) inkubowano z białkiem w obecności buforu do badania oddziaływań białko-kwas nukleinowy (Materiały 3.1.2) o stężeniu końcowym 1× (skład reakcji podano w Tabeli 11). W ten sposób przygotowane próby inkubowano na lodzie przez określony czas, który został uwzględniony w opisach rysunków. Po zakończonej inkubacji próby rozdzielano w 5% lub 10% natywnym żelu PAA (Metody 3.2.2). Analizowane kompleksy wizualizowano przy użyciu skanera FLA-5100 (Fujifilm). Dla tego eksperymentu wykonano przynajmniej 3 powtórzenia techniczne.

**Tabela 11.** Skład mieszaniny do reakcji badania oddziaływań białko-kwas nukleinowy w teście spowolnienia migracji w żelach poliakrylamidowych.

Składnik	Ilość na 1 próbę
wyznakowany [ $\gamma^{32}\text{P}$ ] substrat (10 000 cpm/ $\mu\text{l}$ )	1 $\mu\text{l}$
5×bufor badania oddziaływań białko-kwas nukleinowy	2 $\mu\text{l}$
białko	x $\mu\text{l}$
LB II	5 $\mu\text{l}$
woda	do 20 $\mu\text{l}$

inkubacja w temperaturze 4°C

### 3.2.19 Badanie aktywności wspierającej parowanie cząsteczek kwasów nukleinowych

Wyznakowaną z użyciem [ $\gamma^{32}\text{P}$ ]-ATP cząsteczkę RNA21 (10 000 cpm/ $\mu\text{l}$ , ok. 5 nM), komplementarną do niej cząsteczkę RNA50 oraz  $\Delta\text{DUF}(625-752)$ ,  $\Delta\text{DUF}(630-709)$  lub pełnej długości hDicer inkubowano przez 30 min w temperaturze 37°C. Reakcje prowadzono w rosnącym stężeniu 1.88, 3.75, 7.5, 15, 30 nM każdego z białek. Skład mieszaniny reakcyjnej przedstawiono w **Tabeli 12**. Po zakończonej inkubacji do reakcji dodawano 4% SDS do końcowego stężenia 0,2% oraz 50% glicerol do końcowego stężenia 10%. Próby rozdzielano w 10% natywnym żelu PAA (Metody: 3.2.2). Wyniki wizualizowano przy użyciu skanera FLA-5100 (Fujifilm). Dla tego eksperymentu wykonano przynajmniej 3 powtórzenia techniczne.

**Tabela 12.** Skład mieszaniny do reakcji badania aktywności wspierającej parowanie kwasów nukleinowych.

Składnik	Ilość na 1 próbę
wyznakowana [ $\gamma^{32}\text{P}$ ] cząsteczka RNA21 (10 000 cpm/ $\mu\text{l}$ )	1 $\mu\text{l}$
Niewyznakowana cząsteczka RNA50 (50 fmoli/ $\mu\text{l}$ )	1 $\mu\text{l}$
5×bufor badania oddziaływań białko-kwas nukleinowy	2 $\mu\text{l}$
białko	x $\mu\text{l}$
woda	do 10 $\mu\text{l}$

inkubacja w temperaturze 37°C, 30 min

### **3.2.20 Ocena ilościowo-jakościowa preparatów kwasów nukleinowych**

Stężenie oraz jakość kwasów nukleinowych oceniano spektrofotometrycznie poprzez pomiar absorbancji światła przy użyciu spektrofotometru NanoDrop®1000 (Thermo Fisher Scientific). Przed rozpoczęciem pomiarów urządzenie było kalibrowane rozpuszczalnikiem, buforem lub wodą, w którym była rozpuszczona badana próba. Do analizy użyto 1 µl próby. Otrzymane wyniki posłużyły do oceny preparatów i ich wykorzystania w kolejnych etapach badań. Dodatkowo jakość preparatów sprawdzano także w 1-2% żelach agarozowych z dodatkiem GelView (Metody 3.2.4) lub 10% analitycznych żelach PAA (Metody 3.2.1), które następnie wybarwiono odczynnikiem SYBR™ Gold (Thermo Fisher Scientific), zgodnie z instrukcją dostarczoną przez producenta. Wynik rozdziału udokumentowano z wykorzystaniem urządzenia UVITEC Alliance Q9 Advanced lub wizualizowano przy użyciu skanera FLA-5100 (Fujifilm) przy długości fali 473 nm.

Do oceny stężenia bibliotek cDNA przygotowanych do sekwencjonowania NGS wykorzystano zestaw do pomiaru fluorymetrycznego QuantiFlor® ONE dsDNA (Promega). Próby do analiz przygotowano zgodnie z protokołem producenta i analizowano na urządzeniu Quantus™ Fluorometr (Promega).

### **3.2.21 Odwrotna transkrypcja oraz reakcja qPCR**

Fracje całkowitego RNA były izolowane za pomocą odczynnika TRIzol (Invitrogen), zgodnie z protokołem producenta. Do reakcji odwrotnej transkrypcji użyto 8 µg całkowitego RNA. Reakcję prowadzono z wykorzystaniem zestawu Mir-X™ miRNA First-Strand Synthesis Kit (Takara Bio), zgodnie z protokołem producenta. Po zakończonej reakcji, próby rozcieńczono wodą wolną od RNaz i DNaz do końcowego stężenia 4 ng/µl. Do reakcji qPCR wykorzystano komercyjny zestaw iTaq™ Universal SYBR® Green Supermix (Bio-Rad). Do analizy poziomu miRNA wykorzystano 8 ng uzyskanego cDNA. Amplifikacja była prowadzona z wykorzystaniem urządzenia Bio-Rad CFX96™, które wyznaczało parametr  $C_q$ . Do określenia relatywnego poziomu ekspresji badanych miRNA wykorzystano metodę  $2^{-\Delta\Delta C_t}$  [91]. Normalizacja była prowadzona w oparciu o jądrowy gen U6. Wykonano również reakcje kontrolne bez dodatku matrycy cDNA. Wyniki uzyskane z trzech niezależnych powtórzeń biologicznych przedstawiono jako średnią wartość  $C_q$  z odchyleniem standardowym. Do analizy statystycznej wykorzystano jednokierunkowy test ANNOVA z testem



wielokrotnych porównań Dunnetta, dostępny w programie GraphPad Prism 6 (GraphPad Software). Sekwencje użytych starterów przedstawiono w **Tabeli 1** (Materiały **3.1.1**).

### **3.2.22 Ligacja cząsteczek wykorzystanych w testach wiązania domeny PPC z cząsteczkami RNA**

Do reakcji znakowania opisanej w punkcie: Metody **3.2.5** wykorzystano 50 pmoli cząsteczki, a po zakończonej reakcji znakowania, oczyszczono ją na kolumnie NucAway™, zgodnie z protokołem dostarczonym przez producenta (Ambion). Następnie do cząsteczki zawieszanej w wodzie dodawano roztwór KCl do końcowego stężenia 50 mM i cząsteczkę denaturowano w 95°C przez 3 min, a następnie schładzano do temperatury pokojowej, aby wspierać tworzenie się struktury G-kwadrupleksu. Kolejnym etapem była ligacja cząsteczek za pomocą ligazy T4 RNA (Thermo Fisher Scientific, pełny skład mieszaniny przedstawiono w **Tabeli 13**). Następnie, cząsteczki ponownie oczyszczono za pomocą kolumniek NucAway™ i rozcieńczono do końcowego stężenia 10 000 cpm/μl.

**Tabela 13.** Skład mieszaniny do reakcji ligacji wyznakowanych radioizotopowo cząsteczek RNA.

<b>Składnik</b>	<b>Ilość na 1 próbę</b>
wyznakowany [ $\gamma^{32}\text{P}$ ] substrat (10 pmol/μl)	5 μl
10×bufor do ligacji	2 μl
Ligaza T4 RNA (10 U/μL)	1 μl
DMSO	2 μl
inhibitor RNaz RNaseOUT (40 u/μl) (Invitrogen)	0,5 μl
woda	do 20 μl

inkubacja 17°C przez 10 h

### **3.2.23 Produkcja preparatu 5'PM PPC hDicer w bakteryjnym systemie ekspresyjnym**

Z podłoża stałego LB, na którym znajdowały się kolonie bakteryjne *E. coli* OneShot™ BL21 Star™ zawierające plazmid z wprowadzoną sekwencją wariantu 5'PM PPC hDicer, pobrano pojedynczą kolonię i zawieszono w 5 ml płynnej pożywki LB z ampicyliną (100 μg/ml). Hodowlę prowadzono przez noc (12 – 16 h) w 37°C z wytrząsaniem (225 rpm). Kolejnego dnia zwiększono skalę hodowli do 1:20, tj. do każdego 2,5 ml hodowli nocnej dodano 47,5 ml świeżej płynnej pożywki LB z ampicyliną i ponownie hodowano przez

noc w 37°C z wytrząsaniem (225 rpm). Następnego dnia rano zwiększono skalę hodowli do 1:100, tj. do kolby z 495 ml świeżej płynnej pożywki LB z ampicyliną dodano 5 ml hodowli nocnej i inkubowano w 37°C z wytrząsaniem (225 rpm). Co godzinę sprawdzano w jakiej fazie wzrostu znajdują się bakterie poprzez pomiar gęstości optycznej pobranej próbki hodowli przy długości fali 600 nm (OD600) z wykorzystaniem spektrofotometru NanoDrop®1000 (Thermo Fisher Scientific). Hodowlę prowadzono do momentu uzyskania przez bakterie logarytmicznej fazy wzrostu, tj. do momentu osiągnięcia wartości OD600 ok. 0,8. Po osiągnięciu logarytmicznej fazy wzrostu, hodowlę schłodzono w lodzie przez 30 minut, a następnie indukowano ekspresję genu polimerazy T7 dodając IPTG do końcowego stężenia 0,4 mM. Ekspresję prowadzono przez 18 h w 18°C z wytrząsaniem (225 rpm). Kolejnego dnia hodowlę wirowano (5 000×g, 20 min, 4°C), supernatant usunięto, a osad przechowywano w -80°C do dalszych analiz.

#### **3.2.24 Izolacja i oczyszczanie preparatu białkowego 5'PM PPC hDicer**

Uzyskany osad bakteryjny (Metody 3.2.23) rozpuszczono w 40 ml buforu lizującego (Materiały 3.1.2) i inkubowano w lodzie przez 30 minut. Zawiesinę sonikowano przez 20 minut w 4°C (2 sekundy; 9,9 sekundy przerwy; amplituda 80%), a następnie wirowano (11 000×g, 30 min, 4°C). Po wirowaniu lizatu komórkowego zebrano supernatant zawierający frakcję białek rozpuszczalnych. Następnie, przygotowano kolumnę chromatograficzną do oczyszczania białka metodą chromatografii powinowactwa. W tym celu, w kolumnie umieszczono 5 ml złoża niklowego Ni Sepharose® High Performance (Cytiva). Następnie upakowane złożo równoważono 5V (1V odpowiada jednej objętości złoża, tj. 5ml) 0,5 M roztworu imidazolu w buforze podstawowym 1x (BP 1x) (Materiały 3.1.2), a następnie 10V 0,02 M imidazolu w BP 1x. Po zrównoważeniu, na kolumnę nałożono supernatant zawierający frakcję białek rozpuszczalnych. Następnie złożo przepłukiwano 5V kolejno: 0.02 M, 0.05 M, 0.1 M roztworami imidazolu w BP 1x; i 10V kolejno: 0.2 M, 0.3 M, 0.5 M i 1 M roztworami imidazolu w BP 1x. Przemycanie kolumny roztworami było wspomagane pompą perystaltyczną Masterflex® L/S®. Zebrane frakcje poddano analizie metodą elektroforezy typu SDS-PAGE (Suplement S1) (Metody 3.2.3).

#### **3.2.25 Wymiana buforu i zagęszczenie preparatu białkowego 5'PM PPC hDicer**

Na podstawie wyniku analizy frakcji rozdzielanych metodą elektroforezy typu SDS-PAGE (Metody 3.2.24), do analiz aktywności wybierano frakcje zawierające białko o pożądanym stopniu homogenności (>95%). W celu wymiany buforu na bufor BP 1x

bez imidazolu oraz zagęszczenia otrzymanego preparatu białkowego wykorzystano filtry typu Amicon<sup>®</sup> (Merck) o punkcie odcięcia 10 kDa. Procedurę wymiany buforu przeprowadzono zgodnie z zaleceniami producenta, a zagęszczone białko przechowywano w buforze białkowym 5'PM PPC hDicer (Materiały: **3.1.2**), który wzbogacono w 40% glicerol i przechowywano w -20°C.

### **3.2.26 Oznaczanie stężenia białka 5'PM PPC hDicer metodą Bradforda**

Przygotowano rozcieńczenia BSA: 1 µg/µl; 0.75 µg/µl, 0.5 µg/µl, 0.25 µg/µl, 0.1 µg/µl w objętości 50 µl, w buforze BP 1x. Do wszystkich prób przygotowanych w objętości 50 µl, tj. rozcieńczeń BSA, kontroli bez białka oraz badanych prób z białkiem o nieznanym stężeniu, dodano 2,5 ml odczynnika Bradforda 1x (Bio-Rad). Następnie próby inkubowano w ciemności przez 5 minut. Absorbancję prób mierzono przy długości fali λ595 nm z wykorzystaniem spektrofotometru NanoDrop<sup>®</sup>1000 (Thermo Fisher Scientific). Na podstawie uzyskanych wyników dla znanych stężeń BSA wyznaczono krzywą wzorcową zależności absorbancji od stężenia białka i na tej podstawie wyznaczono stężenie badanego preparatu białkowego.

### **3.2.27 Technika irCLIP (ang. *infrared crosslinking immunoprecipitation*)**

Badania z wykorzystaniem techniki irCLIP zostały przeprowadzone podczas stażu doktorantki, który odbył się w ramach współpracy z Ośrodkiem Genomiki Medycznej - OMICRON na Uniwersytecie Jagiellońskim w Krakowie kierowanym przez dr Pawła Wołkova. Wszystkie etapy badań zostały przeprowadzone zgodnie z protokołem opracowanym przez zespół prof. J. Ule [92].

Adherentne komórki HEK 293T typu dzikiego oraz HEK 293T NoDice przygotowano zgodnie z procedurą opisaną w punkcie Metody: **3.2.13**. Następnie komórki wysiewano w ilości  $1 \times 10^6$  komórek na szalkę hodowlaną o powierzchni 100 mm<sup>2</sup>. Przy stanie ~70% konfluencji, komórki transfekowano wybranym plazmidem z wykorzystaniem odczynnika DharmaFECT kb DNA, zgodnie z protokołem producenta. Po 24 h inkubacji, pożywkę wymieniono na świeżą, a hodowlę prowadzono kolejne 48h.

#### **3.2.27.1 Tworzenie kompleksów białko•RNA za pomocą promieni UV (ang. *UV crosslinking*)**

Zgodnie z powyższym opisem, po 72 h hodowli komórek, szalkę wyjęto z inkubatora i umieszczono na lodzie. Następnie usunięto pożywkę, a do komórek dodano 6 ml zimnego roztworu 1×PBS (EURx). W ten sposób przygotowaną szalkę umieszczono na

lodzie w urządzeniu UV Stratalinker 2400 (Stratagene) i poddano sieciowaniu przy długości fali 254 nm i energii 150 mJ/cm<sup>2</sup>. Następnie komórki odklejono od podłoża za pomocą skrobaka do komórek i przeniesiono do probówek typu Eppendorf po 2 ml. Komórki poddano sedymentacji poprzez wirowanie w temperaturze 4°C (514×g, 3 min). Po usunięciu supernatantu, osady komórkowe zamrożono w suchym lodzie i przechowywano w -80°C. Łącznie uzyskano 3 osady komórkowe z jednego powtórzenia biologicznego. Do dalszych analiz wykorzystano 2 osady (jeden do izolacji kompleksów białko•RNA, a drugi jako kontrolę specyficzności izolacji). Dla każdego badanego układu przygotowano 3 powtórzenia biologiczne.

### **3.2.27.2 Przygotowanie złoża oraz lizatu komórkowego do wyławiania kompleksów białko•RNA**

Procedurę rozpoczęto od przygotowania złoża do wyławiania kompleksów białko•RNA. W tym celu pobrano odpowiednią ilość złoża magnetycznego Dynabeads™ Protein G (100 µl na 1 eksperyment) (Invitrogen) i umieszczono w probówce typu Eppendorf. Złoże przemyto 2-krotnie buforem do lizy komórek (2×900 µl, Materiały 3.1.2). Następnie, złoże zawieszono w buforze do lizy komórek i dodano odpowiednią ilość przeciwciała anti-Dicer (sc-136979, Santa Cruz Biotechnology). Dla jednej próby było to 100 µl buforu do lizy komórek i 4 µg (20 µl) przeciwciała anti-Dicer. W ten sam sposób przygotowano również złoże, które opłaszczono przeciwciałem skierowanym na białko IgG (normal mouse IgG, sc-2025, Santa Cruz Biotechnology), które stanowiło kontrolę specyficzności immunoprecypitacji. Dla jednej kontroli było to 100 µl buforu do lizy komórek i 4 µg (10 µl) przeciwciała normal mouse IgG. Tak przygotowane mieszaniny inkubowano w temperaturze pokojowej do momentu przygotowania lizatu komórkowego, na rotatorze wertykalnym firmy Biosan.

Następnym etapem było przygotowanie buforu do lizy komórek z dodatkiem inhibitora proteaz. Na jedną próbę użyto 980 µl buforu do lizy komórek, do którego dodano 20 µl 50×stężonego inhibitora proteaz (Roche). Następnie osad komórkowy rozpuszczono w 1 ml przygotowanego buforu. Zawiesinę komórek przeniesiono do nowej probówki (dedykowanej do sonikacji) i poddano 10 cyklom sonikacji z naprzemiennym włączaniem i wyłączaniem co 30 sekund na urządzeniu Bioruptor® (Diagenode) w temperaturze 4°C. Po zakończeniu sonikacji, lizat przeniesiono do nowej probówki typu Eppendorf, z której pobierano 5 µl lizatu do pomiaru stężenia całkowitego białka oraz 15 µl do analizy typu Western blot. Do pomiaru stężenia białka wykorzystano zestaw

Coomassie Bradford Assay (Promega), a pomiar wykonano zgodnie z protokołem producenta na 96-dółkowej płytce, na urządzeniu GloMax<sup>®</sup>-Multi Detection System (Promega). Po obliczeniu stężenia całkowitego białka w próbie, próby rozcieńczano buforem do lizy komórek z dodatkiem inhibitora proteaz do końcowego stężenia 0,5 mg/ml (końcowa objętość lizatu wynosiła 1 ml). Następnie do każdego lizatu dodano 4  $\mu$ l rozcieńczonej RNazy 1 (enzym rozcieńczono do stężenia końcowego 0,1 u/ $\mu$ l, Thermo Fisher Scientific) oraz 2  $\mu$ l TURBO<sup>™</sup> DNase (Invitrogen). Próby inkubowano w temperaturze 37°C przez 3 min z wytrząsaniem 1100 rpm. Po zakończonej inkubacji, próby umieszczono na lodzie i inkubowano 3 min, po tym czasie próby wirowano w temperaturze 4°C (16 000 $\times$ g, 10 min).

Ostatnim etapem było połączenie lizatu komórkowego z wcześniej przygotowanym złożem. Przed dodaniem lizatu usunięto bufor, w którym znajdowało się złoże. Następnie, złoże zawieszono w buforze do lizy komórek i przeniesiono do nowych probówek typu Eppendorf po 100  $\mu$ l. Do tak przygotowanego złoża dodano po 1 ml lizatu komórkowego o stężeniu 0,5 mg/ml. Lizat komórkowy oraz złoże opłaskzczone wybranym przeciwciałem inkubowano na rotatorze wertykalnym w temperaturze 4°C przez noc.

### **3.2.27.3 Defosforylacja końca 3' cząsteczek RNA związanych z białkiem oraz ligacja z adapterem wyznakowanym barwnikiem bliskiej podczerwieni**

Po nocnej inkubacji próby umieszczono na statywie magnetycznym. Następnie usunięto supernatant, a złoże przemyto 2-krotnie buforem do płukania z wysoką solą I (2 $\times$ 900  $\mu$ l, Materiały 3.1.2). Przy drugim płukaniu zawiesinę inkubowano na rotatorze wertykalnym przez 1 min w 4°C i dokładnie usunięto supernatant. Dalej, złoże przemyto 2-krotnie buforem do płukania PNK (Materiały 3.1.2), przy drugim płukaniu złoże inkubowano w 4°C przez 10 min na rotatorze wertykalnym. Następnie dokładnie usunięto supernatant, a złoże zawieszono w 20  $\mu$ l mieszaniny do defosforylacji, której skład podano w **Tabeli 14**. Mieszaninę inkubowano w temperaturze 37°C przez 20 min z wytrząsaniem 1100 rpm. Dalej, do prób dodano 1 V (900  $\mu$ l) buforu do płukania PNK i umieszczono je na statywie magnetycznym. Po usunięciu supernatantu złoże przepłukano kolejno 1 V buforu do płukania z wysoką solą I oraz 1 V buforu do płukania PNK. Następnie złoże zawieszono w 100  $\mu$ l buforu do płukania PNK i przeniesiono do nowej probówki typu Eppendorf. Do złoża dodano ponownie 1 V buforu do płukania PNK i całość inkubowano w 4°C na rotatorze wertykalnym przez 10 min, a później dokładnie usunięto supernatant.

Złoże zawieszono w 20  $\mu$ l mieszaniny ligacyjnej, której skład podano w **Tabeli 15**. Próby inkubowano w 16°C z wytrząsaniem 1100 rpm przez noc.

**Tabela 14.** Skład mieszaniny i warunki reakcji defosforylacji końca 3' cząsteczek RNA związanych z białkiem.

Składnik	Ilość na 1 próbę
Bufor 5 $\times$ PNK (Materiały 3.1.2)	4 $\mu$ l
Inhibitor RNaz RNaseOUT (40 U/ $\mu$ l) (Invitrogen)	0,5 $\mu$ l
Kinaza polinukleotydowa T4 (10 U/ $\mu$ l)	0,5 $\mu$ l
H <sub>2</sub> O	do 20 $\mu$ l

inkubacja: 37°C, 20 min, wytrząsanie 1100 rpm

**Tabela 15.** Skład mieszaniny do ligacji adaptera z końcem 3' cząsteczek RNA związanych z białkiem.

Składnik	Ilość na 1 próbę
Bufor 4 $\times$ do ligacji (punkt 3.1.2)	5 $\mu$ l
Inhibitor RNaz RNaseOUT (40 U/ $\mu$ l) (Invitrogen)	0,5 $\mu$ l
T4 ligaza RNA 1 (10 U/ $\mu$ l)	1 $\mu$ l
pre-adenylowany adapter L3-IR-App (1 $\mu$ M)	2 $\mu$ l
PEG400	4 $\mu$ l
H <sub>2</sub> O	do 20 $\mu$ l

inkubacja: 16°C, przez noc z wytrząsaniem  
1100 rpm

#### 3.2.27.4 Rozdział kompleksów w żelu SDS-PAGE oraz transfer na membranę nitrocelulozową

Następny etapem była separacja kompleksów białko•RNA w żelu typu SDS-PAGE oraz ich transfer na membranę nitrocelulozową za pomocą transferu mokrego. W tym celu próby z poprzedniego etapu zawieszono w 1 V (900  $\mu$ l) buforu do płukania PNK i umieszczono na statywie magnetycznym. Po usunięciu supernatantu złoże zawieszono w 1 V (900  $\mu$ l) buforu do płukania z wysoką solą I oraz inkubowano w 4°C na rotatorze wertykalnym przez 5 min. Następnie, złoże przemyto 2-krotnie 1 V buforu do płukania PNK i dokładnie usunięto supernatant. Złoże zawieszono w 20  $\mu$ l wcześniej przygotowanego, 1 $\times$ stężonego buforu NuPAGE™ LDS Sample Buffer (Invitrogen)

w celu elucji kompleksów ze złoża magnetycznego. Na jedną próbę wykorzystano 5  $\mu$ l 4 $\times$ stężonego buforu NuPAGE™ LDS Sample Buffer, 2  $\mu$ l 10% DTT oraz 13  $\mu$ l wody. Próby inkubowano w 70°C przez 5 min z wytrząsaniem 1100 rpm. Po inkubacji, supernatant rozdzielano w żelu z gradientem stężenia 4-12% NuPAGE Bis-Tris (Invitrogen) (warunki rozdziału: 180V, 65 min). Rozdział prowadzono w obecności buforu 1 $\times$  NuPAGE™ MOPS SDS Running Buffer (Invitrogen) za pomocą urządzenia Mini Gel Tank (Thermo Fisher Scientific). Wzorcem wielkości był marker PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific). Następnie za pomocą mokrego transferu kompleksy przeniesiono na membranę nitrocelulozową. Gąbki oraz gotowy zestaw składający się z bibuły i membrany nitrocelulozowej Nitrocellulose/Filter Paper Sandwich, 0.45  $\mu$ m (Invitrogen) zanurzono na 5 min w 1 $\times$ stężonym buforze do transferu NuPAGE™ Transfer Buffer (Invitrogen). Następnie, w dolnej części modułu Mini Blot Module (Invitrogen) umieszczono kolejno: gąbkę, warstwę bibuły, żel, membranę, warstwę bibuły, gąbkę i zamknięto górnym modułem. Tak przygotowany moduł umieszczono w aparacie Mini Gel Tank (warunki rozdziału: stałe napięcie prądu 30 V, 120 min). Po zakończonym transferze membranę przepłukano sterylnym roztworem 1 $\times$ PBS (pH 7.4, Gibco), a kompleksy wizualizowano z wykorzystaniem urządzenia Odyssey Li-COR CLx Imager (Li-COR) przy długości wiązki lasera 700 oraz 800 nm.

#### **3.2.27.5 Izolacja RNA z kompleksów białko•RNA**

Po wizualizacji kompleksów, wycięto wybrany fragment membrany zawierający kompleks białko•RNA i przeniesiono go do probówki typu Eppendorf. Następnie membranę zawieszono w mieszaninie do trawienia białek opisanej w **Tabeli 16**. Całość inkubowano w temperaturze 50°C przez 90 min z wytrząsaniem 1100 rpm. Po zakończonej inkubacji supernatant przeniesiono do nowej probówki typu Eppendorf i dodano 200  $\mu$ l mieszaniny fenol-chloroform-alkohol izoamyłowy w celu wyizolowania RNA. Całość mieszaniny przeniesiono do probówki Phase Lock Gel™ (VWR), inkubowano w 30°C przez 5 min z wytrząsaniem 1100 rpm, a następnie wirowano w temperaturze pokojowej (16000 $\times$ g, 5 min). Do górnej fazy dodano 800  $\mu$ l chloroformu, zmieszano przez kilkukrotne odwracanie probówki i wirowano w 4°C (16000 $\times$ g, 5 min). Górną, wodną warstwę zawierającą RNA przeniesiono do nowej probówki typu Eppendorf, dodano mieszaninę do strącania RNA opisaną w **Tabeli 17** i inkubowano w temperaturze -20°C przez noc.

**Tabela 16.** Skład mieszaniny do trawienia białek w kompleksach białko•RNA.

Składnik	Ilość na 1 próbę
Bufor PK+SDS (Materiały 3.1.2)	190 $\mu$ l
Proteinaza K	10 $\mu$ l

inkubacja: 50°C, 90 min, z wytrząsaniem 1100 rpm

**Tabela 17.** Skład mieszaniny do strącania RNA.

Składnik	Ilość na 1 próbę
GlycoBlue™ Coprecipitant (15 mg/mL)	1 $\mu$ l
5 M NaCl	20 $\mu$ l
Alkohol etylowy	1 ml

inkubacja: -20°C, przez noc

### 3.2.27.6 Odwrotna transkrypcja i ligacja fragmentów cDNA

Po zakończonej nocnej inkubacji, próby zawierające RNA wirowano w temperaturze 4°C (16000×g, 45 min), dokładnie usunięto supernatant, a osad RNA przemyto 900  $\mu$ l 80% etanolu i ponownie wirowano w temperaturze 4°C (16000×g, 30 min). Po usunięciu supernatantu, osad zawieszono w 6  $\mu$ l wody i przeniesiono do próbki do PCR. Do każdej próby dodano starter irCLIP\_ddRT\_## (starter zawiera specyficzny numer indeksu, lista indeksów znajduje się w **Tabeli 2**, Materiały 3.1.1) oraz mieszaninę deoksynukleotydów (dNTP), całość inkubowano w 65°C przez 5 min, a następnie w temperaturze 25°C do momentu dodania następnej mieszaniny opisanej w **Tabeli 18**. Całość inkubowano kolejno w temperaturze 25°C (10 min), 50°C (10 min), 55°C (10 min), 4°C ( $\infty$ ). Po zakończonej inkubacji, znajdujące się w mieszaninie ssRNA/dsRNA oraz ssDNA trawiono mieszaniną enzymów w temperaturze 37°C przez 15 min (skład mieszaniny opisano w **Tabeli 19**). Następnie do wyizolowania cDNA przygotowano złoże MyOne C1 SA-dynabeads (7,5  $\mu$ l złoża dla 1 próby) (Invitrogen). W tym celu pobrano odpowiednią ilość złoża i przemyto je 2-krotnie buforem NT2 (Materiały 3.1.2), następnie złoże zawieszono w tym buforze i dodano po 80  $\mu$ l złoża do każdej próby zawierającej cDNA. Dalej, inkubowano próby ze złożem w temperaturze pokojowej na rotatorze wertykalnym przez 45 min. Po inkubacji usunięto supernatant, złoże przemyto 2-krotnie buforem do płukania z wysoką solą II (Materiały 3.1.2) i 3-krotnie roztworem 1×PBS (pH 7.4, Gibco). Wszystkie etapy separacji supernatantu od złoża prowadzono na



statywie magnetycznym. Po usunięciu supernatantu, złoże zawieszono w mieszaninie do elucji cDNA (skład podano w **Tabeli 20**) oraz inkubowano w temperaturach 95°C (3 min), 95-60°C (spadek temperatury -0,1°C/s), 60°C (10 min). Po zakończonej inkubacji do prób znajdujących się w termocyklerze T-100 (Bio-Rad) dodano mieszaninę do ligacji cDNA (skład podano w **Tabeli 20**) i ponownie inkubowano w temperaturze 60°C (120 min), 4°C (10 min) oraz 95°C (3 min). Po zakończonej inkubacji próby umieszczono na statywie magnetycznym, a supernatant zawierający wolne cDNA przeniesiono do nowej probówki typu Eppendorf. Pozostałe złoże zawieszono w buforze TE (185 µl) i inkubowano w termocyklerze w temperaturze 80°C (5 min) żeby odzyskać pozostałości cDNA. Po zakończonej inkubacji supernatant przeniesiono do wcześniej zebranego supernatantu zawierającego cDNA. Do mieszaniny dodano 200 µl buforu TE oraz mieszaninę do strącania cDNA (skład podano w **Tabeli 21**) i inkubowano przez noc w temperaturze -20°C.

**Tabela 18.** Skład mieszaniny do odwrotnej transkrypcji.

<b>Składnik</b>	<b>Ilość na 1 próbę</b>
starter irCLIP_ddRT_# (1 pmol/µl)	1 µl
mieszanina dNTP (10 nM)	0,5 µl
<hr/>	
inkubacja 65°C przez 5 min, 25°C (∞)	
<hr/>	
bufor 5×SSIV	2 µl
0,1 M DTT	0,5 µl
Inhibitor RNaz RNaseOUT (40 U/µl) (Invitrogen)	0,25 µl
Superscript IV Reverse Transcriptase (200 U/ µl)	0,25 µl
<hr/>	
inkubacja 25°C (10 min), 50°C (10 min), 55°C (10 min), 4°C (∞)	

**Tabela 19.** Skład mieszaniny do trawienia RNA po reakcji odwrotnej transkrypcji.

<b>Składnik</b>	<b>Ilość na 1 próbę</b>
RNaza H (5 U/µl)	0,5 µl
RNaza A (1 mg/mL)	0,5 µl
<hr/>	
inkubacja 37°C, 15 min	

**Tabela 20.** Skład mieszaniny do elucji oraz ligacji cDNA związanego ze złożem MyOne C1 SA-dynabeads.

Składnik	Ilość na 1 próbę
P3 tall (1 pmol/μl)	1 μl
MnCl <sub>2</sub> (50 mM)	0,75 μl
Betaina (5 M)	3 μl
inkubacja 95°C (3 min), 95-60°C (spadek temperatury -0,1°C/s), 60°C (10 min)	
bufor 10×CircLigase II	1,5 μl
CircLigase II ssDNA Ligase (100 U/μl)	0,75 μl
inkubacja 60°C (120 min), 4°C (10 min), 95°C (3 min)	

**Tabela 21.** Skład mieszaniny do strącania cDNA.

Składnik	Ilość na 1 próbę
GlycoBlue™ Coprecipitant (15 mg/mL)	1 μl
3 M octan sodu (pH 5.5)	40 μl
Alkohol etylowy (100%)	1 ml
inkubacja: -20°C, przez noc	

### 3.2.27.7 Optymalizacja reakcji PCR oraz namnażanie fragmentów cDNA

Po zakończonej nocnej inkubacji, próby zawierające cDNA wirowano w temperaturze 4°C (16000×g, 45 min), usunięto dokładnie supernatant, a osad cDNA przemyto 500 μl 80% etanolu i ponownie wirowano w temperaturze 4°C (16000×g, 30 min). Po usunięciu supernatantu, osad zawieszono w 14 μl wody. Następnie, przygotowano mieszaninę do reakcji optymalizacji PCR (skład opisano w **Tabeli 22**). Optymalizację prowadzono w 4 warunkach, w każdym z nich zmieniała się ilość cykli namnażania produktu PCR. Warunki reakcji PCR opisano w **Tabeli 23**, a reakcję prowadzono dla 18, 21, 24 oraz 27 cykli namnażania produktu PCR w termocyklerze T-100 (Bio-Rad). Następnie produkty reakcji rozdzielano w 6% żelu TBE Novex™ (Invitrogen) w obecności buforu 1×TBE (Roth) (warunki rozdziału: 180 V, 25 min). Jako marker wielkości zastosowano Low Molecular Weight DNA Ladder (New England BioLabs). Na żel nałożono 5 μl każdej próby z dodatkiem 1 μl buforu obciążającego dołączonego do markera wielkości. Po zakończonym rozdziale, w celu wizualizacji produktów PCR, żel barwiono w roztworze

SYBR<sup>®</sup> Green I (Sigma) przez 5 min i dokumentowano z wykorzystaniem urządzenia Gel Dox XR Imaging System (Bio Rad). Następnie przeanalizowano otrzymane obrazy i wybrano taką ilość cykli namnażania produktów PCR, w której obserwujemy podwojenie początkowej ilości matrycy.

Kolejnym etapem było ponowne namnożenie biblioteki cDNA, która stanowiła bibliotekę do sekwencjonowania NGS, czyli tzw. preparatywny PCR. W tym celu ponownie przygotowano reakcję PCR z wcześniej opisaną mieszaniną składników (**Tabela 22**) oraz 4  $\mu$ l matrycy. W związku ze zwiększeniem ilości początkowej matrycy, od wybranej ilości cykli namnażania produktu PCR odjęto 2 cykle, aby wyrównać ilości matrycy pomiędzy reakcją optymalizacji PCR a preparatywnym PCR. Reakcję PCR przeprowadzono zgodnie z opisem warunków reakcji opisanych w **Tabeli 23**.

**Tabela 22.** Skład mieszaniny do optymalizacji reakcji PCR w protokole irCLIP.

Składnik	Ilość na 1 próbę
Matryca: cDNA	1 $\mu$ l lub 4 $\mu$ l
startery P5Solexa/P3Solexa (10 $\mu$ M)	0,5 $\mu$ l
Polimeraza Phusion High-Fidelity PCR Master	10 $\mu$ l
woda	do 20 $\mu$ l
inkubacja 60°C (120 min), 4°C (10 min), 95°C (3 min)	

**Tabela 23.** Warunki reakcji optymalizacji reakcji PCR w protokole irCLIP.

Etap	Temperatura	Czas	Ilość cykli
Denaturacja wstępna	98°C	40 s	1
Denaturacja	98°C	20 s	
Przyłączanie starterów	65°C	30 s	18-27
Synteza	72°C	45 s	
Synteza końcowa	72°C	3 min	1
Przechowywanie	25°C	$\infty$	

### 3.2.27.8 Oczyszczanie wybranych fragmentów biblioteki cDNA z żelu

Ostatnim etapem było przygotowanie biblioteki cDNA do sekwencjonowania NGS. W tym celu namnożone fragmenty cDNA w reakcji preparatywnego PCR rozdzielano

w 6% żelu TBE Novex™ zgodnie z opisaną procedurą w punkcie Metody **3.2.26.7**. Po barwieniu żelu w roztworze SYBR® Green I, żel obrazowano na transiluminatorze z światłem niebieskim i wycięto fragmenty cDNA w zakresie długości 145-400 nt. Skrawki żelu umieszczono w probówce typu Eppendorf, rozdrobniono na mniejsze fragmenty i inkubowano w obecności buforu do elucji z żelu (Materiały **3.1.2**) w temperaturze 65°C przez 120 min z wytrząsaniem 1100 rpm. Po zakończonej inkubacji, mieszaninę przeniesiono do próbki Spin-X Costar (Sigma) zawierającej filtr celulozowy o średnicy porów 0,22 µm i krótko wirowano (16000×g, 1 min, 4°C), aby odseparować fragmenty żelu od supernatantu. Następnie cały supernatant przeniesiono do nowej próbki typu Eppendorf i dodano mieszaninę do strącania cDNA (**Tabela 21**). Całość inkubowano przez noc w temperaturze -20°C. Dalej, próby wirowano (16000×g, 30 min, 4°C), usunięto dokładnie supernatant, a osad cDNA przemyto 500 µl 80% etanolu i ponownie wirowano (16000×g, 30 min, 4°C). Po usunięciu supernatantu, osad zawieszono w 15 µl wody. W ten sposób otrzymaną bibliotekę fragmentów cDNA wykorzystano do sekwencjonowania NGS.

### **3.2.28 Rozdział fragmentów bibliotek z zastosowaniem systemu TapeStation**

Rozkład długości fragmentów w bibliotece oraz jakość biblioteki cDNA sprawdzono za pomocą czułego rozdzału elektroforetycznego z zastosowaniem chipów High Sensitivity D1000 ScreenTape® System na urządzeniu Agilent 2200 TapeStation System (Agilent). Roztwór markera wielkości, a także przygotowanie prób wykonano zgodnie z protokołem producenta.

### **3.2.29 Normalizacja bibliotek do sekwencjonowania NGS**

Po określeniu stężenia każdej biblioteki cDNA (Metody **3.2.20**) oraz sprawdzeniu średniego rozkładu długości fragmentów bibliotek cDNA (Metody **3.2.28**), biblioteki połączono ze sobą w jednej próbce typu Eppendorf do końcowego stężenia 0,05 ng/µl każdej z bibliotek. Kontrolę IgG potraktowano objętościowo, tzn. z każdej biblioteki kontroli IgG dodano po 2 µl do biblioteki końcowej. Następnie ponownie wykonano pomiar stężenia przygotowanej biblioteki zgodnie z punktem: Metody **3.2.20** w trzech powtórzeniach technicznych oraz określono średnią długość produktów cDNA w bibliotece zgodnie z punktem: Metody **3.2.28**. Na podstawie wyliczonych stężeń biblioteki cDNA, obliczono średnie stężenie biblioteki, które wynosiło 0,315 ng/µl. Żeby oszacować ilość kopii cDNA wartość stężenia ng/µl należy wyrazić w nM. Wyrażone w ten sposób stężenie pozwoli na jednakowe pokrycie podczas sekwencjonowania

każdego fragmentu cDNA znajdującego się w bibliotece. W celu wyrażenia stężenia w nM należy posłużyć się poniższym wzorem:

$$\frac{\text{stężenie próbki } \left(\frac{ng}{\mu l}\right)}{660 \left(\frac{g}{mol}\right) \times \text{średnia długość produktów cDNA w bibliotece (pz)}} \times 10^6 \\ = \text{stężenie biblioteki (nM)}$$

Wartość współczynnika 660 g/mol to masa molekularna jednej pary zasad.

Stężenie molowe biblioteki wyniosło 2,973 nM  $\approx$  3 nM. Przygotowanie biblioteki zostało zrealizowane dzięki współpracy z dr Justyną Totoń-Żurańską oraz dr Pawłem Wołkowem z Ośrodka Genomiki Medycznej OMICRON na Wydziale Lekarskim Uniwersytetu Jagiellońskiego w Krakowie.

### **3.2.30 Sekwencjonowanie z zastosowaniem sekwenatora Illumina NextSeq550**

Znormalizowaną bibliotekę zgodnie z punktem: Metody **3.2.29** rozcieńczono wodą wolną od DNaz i RNaz do końcowego stężenia 2 nM i przygotowano do nałożenia na kartridż zgodnie z protokołem „NextSeq 500 and 550 Sequencing Systems Denature and Dilute Libraries Guide” (Illumina). Sekwencjonowanie bibliotek wykonano na płycie o wysokiej przepustowości (High Output Flow Cell (Illumina)). Finalne stężenie biblioteki w reakcji sekwencjonowania wynosiło 1.8 pM. Sekwencjonowanie wykonano z zastosowaniem sekwenatora Illumina NextSeq550 (Illumina).

Sekwencjonowanie NGS zostało zrealizowane dzięki współpracy z dr Michałem Sewerynem oraz mgr Marcinem Słomką z Pracowni Biobank na Wydziale Biologii i Ochrony Środowiska Uniwersytetu Łódzkiego.

### **3.2.31 Analiza danych pochodzących z sekwencjonowania NGS**

Pierwszym etapem analizy uzyskanych danych było tzw. przetwarzanie wstępne (ang. *preprocessing*), w którym to dokonano kontroli jakości surowych danych (ang. *quality control*). W tym celu wykorzystano narzędzie FastQC, które dało ogólne wskaźniki jakościowe dotyczące uzyskanych odczytów. Następnym etapem było wykorzystanie narzędzia UMI-tools w celu wyodrębnienia i uporządkowania odczytów oraz przypisania nazwy próby do poszczególnego znacznika UMI. Często ten krok jest również wykonywany w ramach demultipleksacji produktów PCR powstających w reakcji sekwencjonowania z wykorzystaniem platformy Illumina. Ostatnim krokiem było usunięcie z odczytów sekwencji adapterów, a także przefiltrowanie odczytów, tzw.

przycięcie danych (ang. *quality trimming*) pod względem ich jakości. Do tego celu wykorzystano program Cutadapt, który domyślnie wycina uniwersalne sekwencje adapterów Illumina oraz usuwa odczyty krótsze niż 12 nt. Drugi etap obejmował dopasowanie (ang. *alignment*) otrzymanych odczytów do genomu referencyjnego. W tych analizach genom referencyjny stanowił ludzki genom, wersja hg38 (*Homo\_sapiens.GRCh38*). Najpierw otrzymane odczyty zmapowano do sekwencji rRNA oraz tRNA (ang. *pre-alignment*) za pomocą narzędzia Bowtie 2, a następnie wykonano dopasowanie do genomu referencyjnego dzięki wykorzystaniu programu STAR. W celu identyfikacji miejsc wiązań krzyżowych pomiędzy białkiem a RNA, tzw. crosslinków, wykorzystano najpierw narzędzie UMI-tools w celu deduplikacji produktów PCR w oparciu o indeksy UMI, a następnie programu BEDTools, który zidentyfikował ilość oraz dokładne miejsca wystąpienia wiązań krzyżowych. Wszystkie zastosowane programy i narzędzia pozwoliły wygenerować końcowy raport jakości MultiQC. Analiza różnicowa ekspresji genów została wykonana w programie DESeq2. Analiza sekwencji genów, których transkrypty zostały związane przez badane białka w komórce, została wykonana w programie IGV służącym do analiz danych genomicznych.

Dane uzyskane w wyniku sekwencjonowania NGS bibliotek cDNA uzyskanych metodą irCLIP zostały poddane analizie i obróbce bioinformatycznej przy współpracy z bioinformatykami: dr Michałem Sewerynem (Wydział Biologii i Ochrony Środowiska, Uniwersytet Łódzki) oraz mgr Ziemowitem Radogostowiczem (Zakład Biochemii Rybonukleoprotein IChB PAN).

## 4. Wyniki

### 4.1 Badanie znaczenia domeny DUF283 dla aktywności ludzkiej rybonukleazy Dicer

Wcześniejsze badania prowadzone w Instytucie Chemii Bioorganicznej PAN wykazały, że domena DUF283 hDicer (rejon obejmujący 625-752 aa hDicer, zgodnie z sekwencją transkryptu NM\_030621.4) wiąże jednoniciowe RNA i DNA, nie wiąże natomiast dsRNA i dsDNA [34]. Ponadto wykazano, że zarówno sama domena DUF283 jak i pełnej długości białko hDicer wspierają parowanie cząsteczek RNA i DNA [34]. W ramach realizacji niniejszej pracy doktorskiej, w oparciu o wyniki przeprowadzonych badań postawiono hipotezę, że domena DUF283 hDicer jest kluczowa dla aktywności hDicer wspierającej hybrydyzację komplementarnych kwasów nukleinowych. W celu weryfikacji postawionej hipotezy otrzymano dwa warianty delecyjne hDicer: (i) wariant z delecją obejmującą rejon pomiędzy 625 a 752 aa hDicer, wariant  $\Delta$ DUF(625-752) (delecja obejmowała fragment wykorzystany w pracy „*Revealing a new activity of the human Dicer DUF283 domain in vitro*” [34]) oraz (ii) wariant z delecją obejmującą rejon pomiędzy 630 a 709 aa hDicer (rejon obejmujący domenę DUF283) [30]. Wariant  $\Delta$ DUF(630-709) powstał w oparciu o dane zawarte w pracy „*Autoinhibition of Human Dicer by its Internal Helicase Domain*” Ma E. i in. [30]. Otrzymane preparaty białkowe były badane pod kątem aktywności RNazowej oraz aktywności wspierania parowania cząsteczek kwasów nukleinowych.

#### 4.1.1 Otrzymanie konstruktów genetycznych kodujących warianty $\Delta$ DUF(625-752) oraz $\Delta$ DUF(630-709)

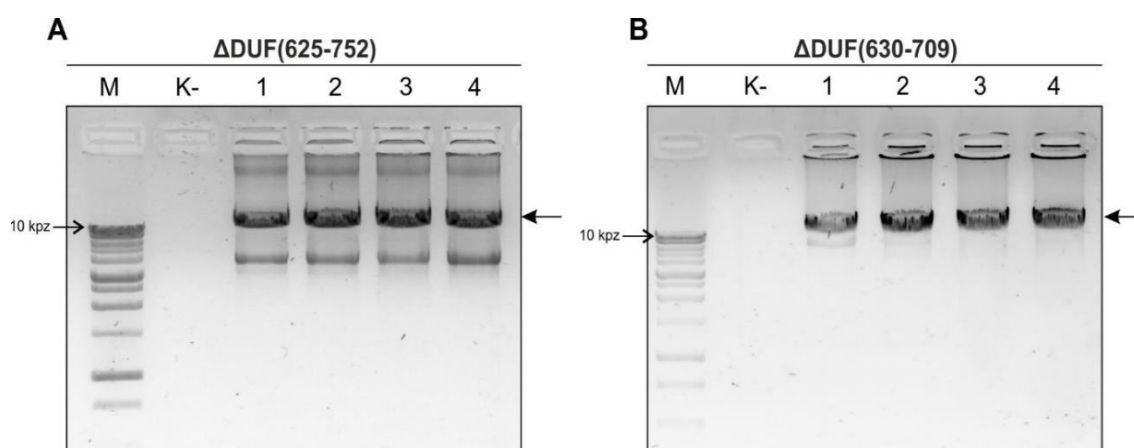
W celu otrzymania plazmidów kodujących wspomniane wyżej warianty delecyjne wykorzystano reakcję PCR (ang. *polymerase chain reaction*). Matrycą w reakcjach PCR był plazmid SureVector wt\_hDicerx3Flag (Metody 3.2.9), zawierający sekwencję kodującą hDicer typu dzikiego (NM\_030621.4), **Rysunek 3** [41]. Startery zaprojektowano w taki sposób, aby powielić cały plazmid: (i) bez rejonu kodującego aminokwasy 625-752, w przypadku wariantu  $\Delta$ DUF(625-752) lub (ii) bez rejonu kodującego aminokwasy 630-709, w przypadku wariantu  $\Delta$ DUF(630-709) (**Rysunek 3**). Listę starterów wykorzystanych w reakcjach PCR podano w rozdziale: Materiały 3.1.1 w **Tabeli 1**.



**Rysunek 3. Schemat wektora ekspresyjnego SureVector zawierającego sekwencję kodującą pełnej długości białko hDicer.** Na schemacie zaznaczono sekwencję insertu (GOI) kodującego hDicer typu dzikiego, sekwencję kodującą 3xFlag, sekwencję kodującą gen oporności na puromycynę oraz ampicylinę (PuromycinR, AmpR) oraz promotor pochodzący z cytomegalowirusa (CMV promoter). Schemat wygenerowano w programie SnapGene Viewer. Ponadto, na schemacie zaznaczono miejsca przyłączenia się starterów w reakcji PCR dla wariantu  $\Delta$ DUF(625-752) strzałkami w kolorze szarym oraz dla wariantu  $\Delta$ DUF(630-709) strzałkami w kolorze żółtym.

Produkty PCR rozdzielano w żelu agarozowym (Metody 3.2.4), **Rysunek 4**, a następnie wyizolowano DNA o długościach odpowiadających pożądanym fragmentom, tj. 10 653 pz ( $\Delta$ DUF(625-752)) oraz 10 800 pz ( $\Delta$ DUF(630-709)). Wycięte z żelu fragmenty DNA oczyszczono z wykorzystaniem zestawu Syngen Gel Mini Kit (Syngen), zgodnie z protokołem dostarczonym przez producenta.

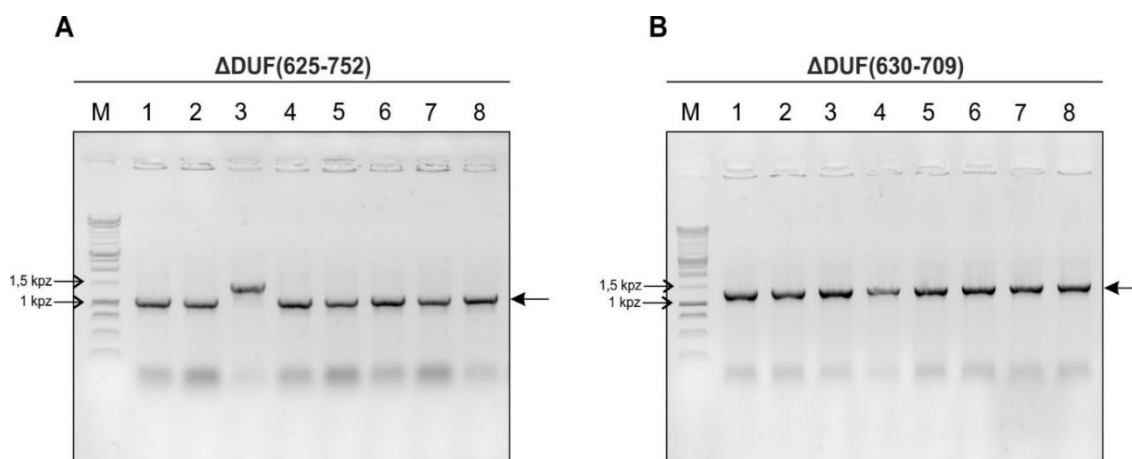




**Rysunek 4. Zdjęcia przedstawiające rozdziel elektroforetyczny produktów amplifikacji wykorzystanych do przygotowania konstruktów  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709).** Na żel nałożono 5  $\mu$ l markera Perfect Plus™ 1 kb DNA Ladder (EurX) (M), 5  $\mu$ l reakcji kontrolnej bez matrycy (K-) oraz 5  $\mu$ l produktu PCR (1-4). Strzałką wskazano pożądane produkty reakcji. (A, B) Panele przedstawiające wyniki reakcji PCR dla wariantu  $\Delta$ DUF(625-752) (A) oraz dla wariantu  $\Delta$ DUF(630-709) (B). Rozdział prowadzono w 1% żelu agarozowym.

W ten sposób przygotowane produkty fosforylowano na końcu 5' poprzez dodanie kinazy polinukleotydowej T4, a następnie końce produktów połączono za pomocą ligazy DNA T4, zgodnie z opisem: Metody 3.2.9. Całość mieszaniny ligacyjnej wykorzystano do transformacji komórek bakteryjnych *E. coli* szczep DH5 $\alpha$  (Metody: 3.2.10). Komórki bakteryjne wysiano na szalkę ze stałym podłożem LB zawierającym ampicylinę.

W celu weryfikacji czy obecne na szalce ze stałym podłożem kolonie bakteryjne posiadają pożądany plazmid wykorzystano tzw. „kolonijny PCR” (Metody 3.2.11). W reakcji wykorzystano startery: P4-DICER oraz r\_PAZ (Materiały 3.1.1, **Tabela 1**). Startery zostały zaprojektowane w taki sposób, aby zweryfikować, czy matrycą w reakcji był plazmid kodujący pełnej długości hDicer (produkt PCR o długości 1535 pz), czy też plazmid kodujący  $\Delta$ DUF(625-752) (produkt PCR o długości 1151 pz) lub  $\Delta$ DUF(630-709) (produkt PCR o długości 1294 pz). Rozdział produktów kolonijnego PCR przedstawiono na **Rysunku 5**.



**Rysunek 5. Zdjęcie przedstawiające rozdzielanie elektroforetyczne przykładowych produktów reakcji „kolonijnego PCR” (A, B).** Na żel nałożono 5  $\mu$ l markera Perfect Plus™ 1 kb DNA Ladder (EurX) (M) oraz 15  $\mu$ l produktu „kolonijnego PCR” (1-8). Strzałką wskazano pożądane produkty reakcji. Wszystkie kolonie, oprócz kolonii 3, posiadały plazmid z delecją  $\Delta$ DUF(625-752) (A). W przypadku wariantu  $\Delta$ DUF(630-709), wszystkie analizowane kolonie posiadały plazmid z delecją (B). Rozdział prowadzono w 1% żelu agarozowym.

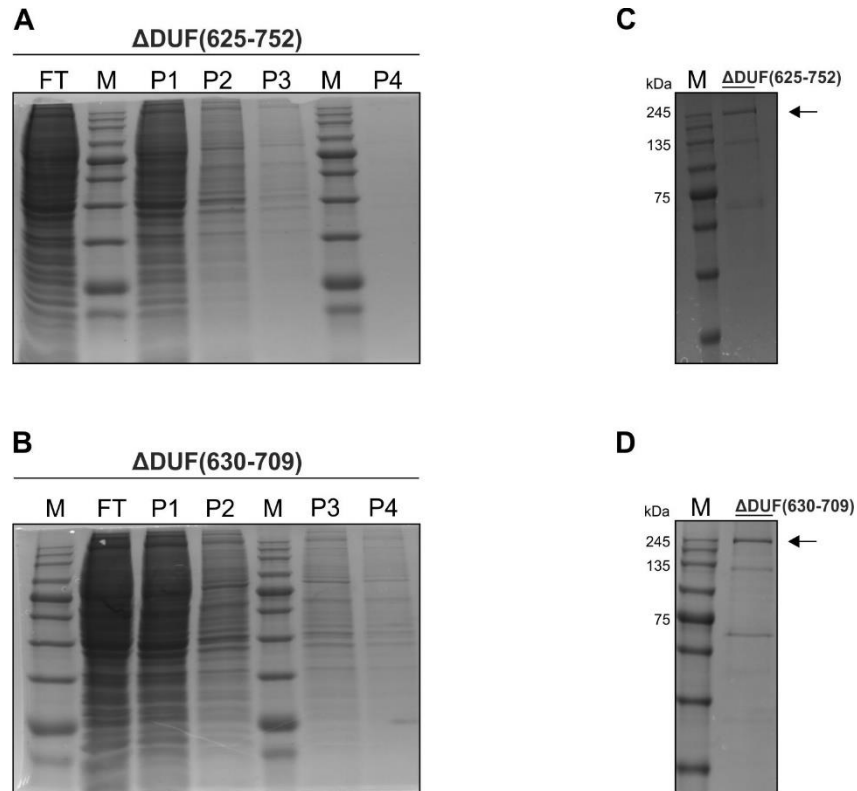
Na podstawie wyników reakcji „kolonijnego PCR” wybrano kilka kolonii bakteryjnych, z których wyizolowano plazmidowy DNA (Metody: 3.2.12). Tak otrzymane plazmidy zostały wykorzystane jako matryce w reakcji sekwencjonowania metodą Sangera (Tabela 1 w rozdziale: Materiały 3.1.1). Otrzymane rezultaty potwierdziły całkowitą zgodność sekwencji insertów w plazmidach kodujących warianty delecyjne  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709), z sekwencjami referencyjnymi, a także poprawne umiejscowienie sekwencji kodującej w ramce odczytu zgodnej z kodonem start AUG.

#### 4.1.2 Otrzymywanie preparatów białkowych $\Delta$ DUF(625-752) i $\Delta$ DUF(630-709) oraz hDicer

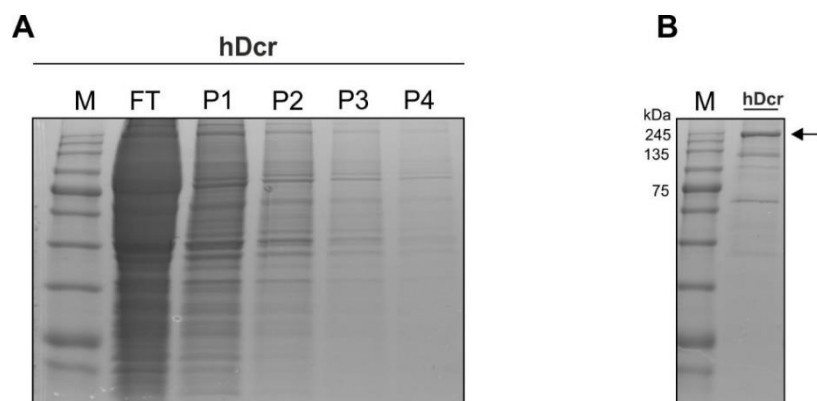
W celu otrzymania preparatów białkowych wariantów  $\Delta$ DUF(625-752),  $\Delta$ DUF(630-709) oraz hDicer (hDcr), komórki HEK 293T NoDice transfekowano odpowiednim plazmidem kodującym pożądane białko. Hodowlę komórkową oraz transfekcję prowadzono zgodnie z procedurą opisaną w rozdziale: Metody 3.2.13. Po upływie 72 h od transfekcji, komórki zbierano poprzez wirowanie (200 $\times$ g, 3 min), a otrzymany osad komórkowy był przechowywany w temperaturze -80°C do momentu dalszych analiz.

Lizę komórek HEK 293T NoDice transfekowanych odpowiednim plazmidem przeprowadzono zgodnie z opisem w punkcie: Metody 3.2.14. Wyprodukowane w komórkach warianty hDicer oraz hDcr izolowano za pomocą immunoprecypitacji

dzięki obecności znacznika 3xFlag na C-końcu białek. W tym celu wykorzystano złożę ANTI-FLAG<sup>®</sup>M2 Affinity Gel (Merck). Jest to złożę agarozowe połączone kowalencyjnie z mysim monoklonalnym przeciwciałem IgG M2, które wykazuje powinowactwo do peptydu Flag. Supernatant komórkowy nałożono na wcześniej przygotowaną zgodnie z instrukcją producenta kolumnę ze złożem agarozowym. Do uwalniania białka związanego ze złożem wykorzystano peptyd 3xFlag (Merck), zgodnie z procedurą opisaną w rozdziale: Metody 3.2.15. Wszystkie uzyskane podczas oczyszczania frakcje analizowano za pomocą elektroforezy typu SDS-PAGE w 8% żelu (Metody: 3.2.3) (**Rysunek 6A, B oraz Rysunek 7A**). Frakcja zawierająca uwolnione białko była zagęszczana na filtrze Amicon (Merck) z punktem odcięcia 100 kDa. Stężenie uzyskanego preparatu białkowego oceniono za pomocą metody Bradforda, a także za pomocą elektroforezy typu SDS-PAGE, na podstawie oceny stężenia oczyszczonego białka  $\Delta$ DUF(625-752),  $\Delta$ DUF(630-709) lub hDcr, w odniesieniu do znanych stężeń BSA. Średnie stężenie uzyskanego preparatu wynosiło  $\sim 0,1 \mu\text{g}/\mu\text{l}$  dla białka  $\Delta$ DUF(625-752),  $\sim 0,15 \mu\text{g}/\mu\text{l}$  dla białka  $\Delta$ DUF(630-709) oraz  $\sim 0,3 \mu\text{g}/\mu\text{l}$  dla hDcr. Przygotowany w ten sposób preparat białkowy był przechowywany w  $-20^{\circ}\text{C}$  w buforze zawierającym 40% glicerol. Przykładowe próbki preparatów, rozdzielone w 8% żelu typu SDS-PAGE, przedstawiono na **Rysunku 6C, D oraz Rysunku 7B**. Można zauważyć, że we wszystkich preparatach białkowych, oprócz pożądanego białka, obecne były także niewielkie ilości dwóch innych białek. W celu ich identyfikacji wykonano analizę metodą spektrometrii mas, która wykazała, że białko znajdujące się na wysokości ok. 135 kDa to białko KIF11 (ang. *Kinesin-like protein KIF11*), które bierze udział w tworzeniu wrzeciona kariokinetycznego. Natomiast drugie białko, znajdujące się na wysokości ok. 75 kDa, to metylotransferaza argininy ANM5 (ang. *Protein arginine N-methyltransferase 5*). Analizy metodą spektrometrii mas wykonano we współpracy z dr Łukaszem Marcziakiem (Pracownia Spektrometrii Mas IChB PAN).



**Rysunek 6.** Zdjęcie przedstawiające rozdzielanie elektroforetyczne frakcji otrzymanych w procesie oczyszczania preparatów  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709). (A, B) Rozdział frakcji uzyskanych w trakcie izolacji i oczyszczania białek. Na żel nałożono 5  $\mu$ l markera Perfect™ Tricolor Protein Ladder (EURx) (M) oraz po 15  $\mu$ l każdej z frakcji: (FT) frakcja po nocnej inkubacji lizatu ze złożem, (P1-P4) frakcje zebrane po każdym przemyciu złoża roztworem TBS. (C, D) Preparat białkowy po oczyszczaniu i zagęszczeniu białka na filtrze Amicon. Na żel nałożono 5  $\mu$ l markera Perfect™ Tricolor Protein Ladder (EURx) (M) oraz 15  $\mu$ l preparatu białkowego. Strzałkami zaznaczono preparat  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709). Rozdziały prowadzono w 8% żelu PAA.



**Rysunek 7.** Zdjęcie przedstawiające rozdzielanie elektroforetyczne frakcji otrzymanych w procesie oczyszczania preparatu hDcr. (A) Rozdział frakcji uzyskiwanych w trakcie izolacji i oczyszczania hDcr. Na żel nałożono 5  $\mu$ l markera Perfect™ Tricolor Protein Ladder (EURx) (M) oraz po 15  $\mu$ l każdej z frakcji: (FT) frakcja po nocnej inkubacji lizatu ze złożem, (P1-P4) frakcje zebrane po każdym przemyciu złoża roztworem TBS. (B) Preparat hDcr po oczyszczaniu i zagęszczeniu białka na filtrze Amicon. Na żel nałożono 5  $\mu$ l markera Perfect™ Tricolor Protein Ladder (EURx) (M) oraz 15  $\mu$ l preparatu hDcr. Strzałką zaznaczono preparat hDcr. Rozdziały prowadzono w 8% żelu PAA.

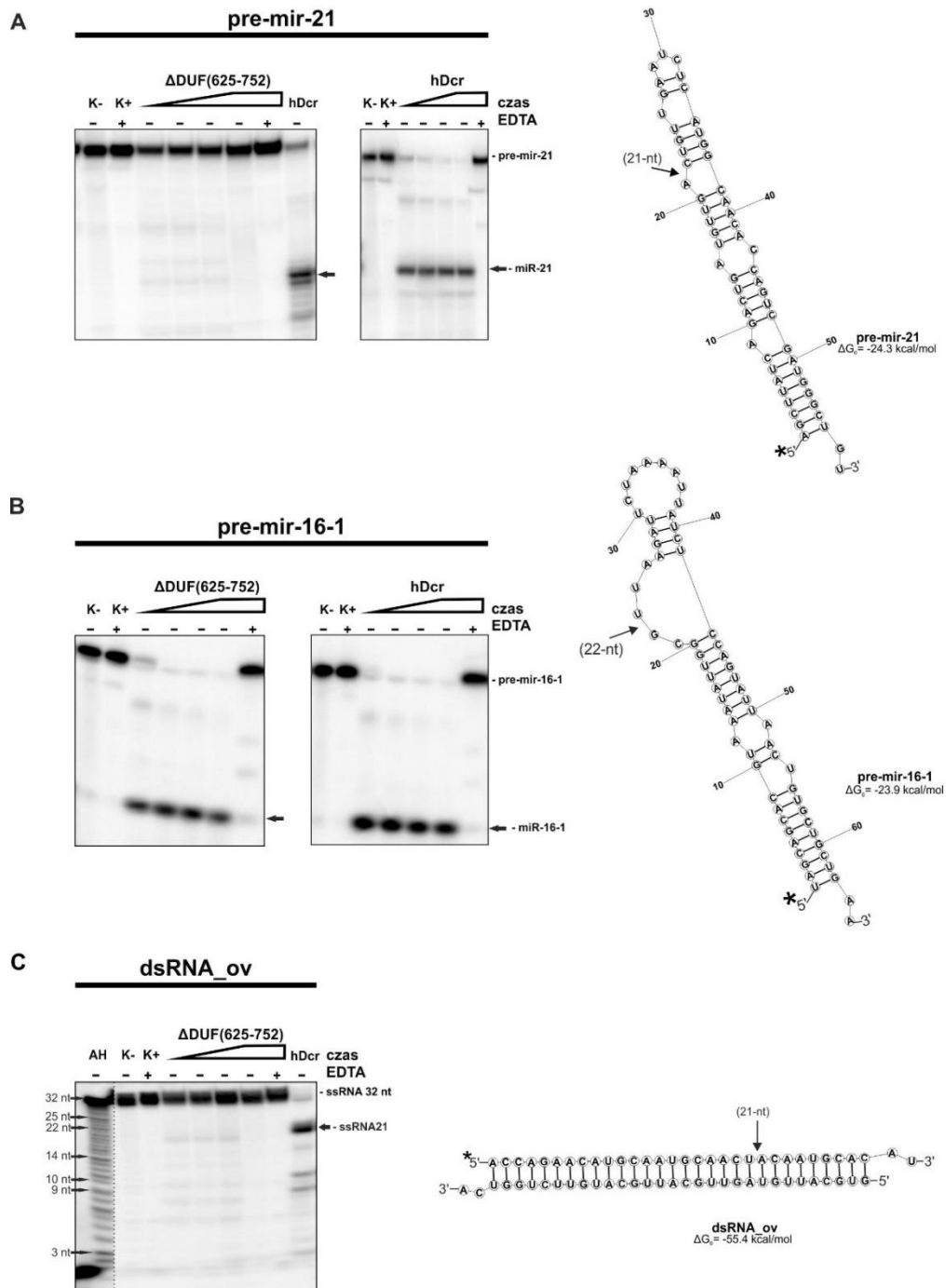
#### 4.1.3 Badanie aktywności otrzymanych preparatów białkowych $\Delta$ DUF(625-752) oraz $\Delta$ DUF(630-709)

Otrzymane preparaty  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709) zbadano pod kątem (i) aktywności RNazowej, zarówno w warunkach *in vitro*, jak i *in cellulo*, oraz (ii) aktywności wspierania parowania cząsteczek kwasów nukleinowych.

##### 4.1.3.1 Badanie aktywności RNazowej wariantów $\Delta$ DUF(625-752) oraz $\Delta$ DUF(630-709) w warunkach *in vitro*

Do badania aktywności RNazowej  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709) w warunkach *in vitro* wykorzystano kanoniczne substraty hDicer: dwie cząsteczki pre-miRNA, pre-mir-21 oraz pre-mir-16-1, a także dsRNA o długości 30 pz posiadający dwa niesparowane nukleotydy na końcach 3' dupleksu (dsRNA\_ov). Cząsteczka pre-mir-16-1 zawiera duże pętle wewnętrzne w obrębie trzonu spinki oraz 9 nt rejon pętli terminalnej. Natomiast cząsteczka pre-mir-21 przyjmuje bardziej zwartą strukturę i posiada małą, 3 nt pętlę terminalną (**Rysunek 8**). Jako kontrolę w reakcjach cięcia wykorzystano hDcr. W celu monitorowania integralności substratu w czasie trwania inkubacji przygotowano także dwie reakcje kontrolne bez białka: jedną zawierającą tylko substrat w buforze reakcyjnym (K-) oraz drugą, zawierającą substrat w buforze reakcyjnym z dodatkiem czynnika chelatującego jony  $Mg^{2+}$ , tj. EDTA (K+). Aktywność RNazowa hDicer jest zależna od jonów  $Mg^{2+}$ , więc przygotowano również kolejną reakcję kontrolną zawierającą substrat, białko oraz EDTA (+EDTA). Reakcje prowadzono w czasie (10, 30, 60, 120 min) w objętości 10  $\mu$ l (Metody 3.2.17), a produkty reakcji cięcia rozdzielano w 15% żelu PAA (**Rysunek 8** oraz **Rysunek 9**).

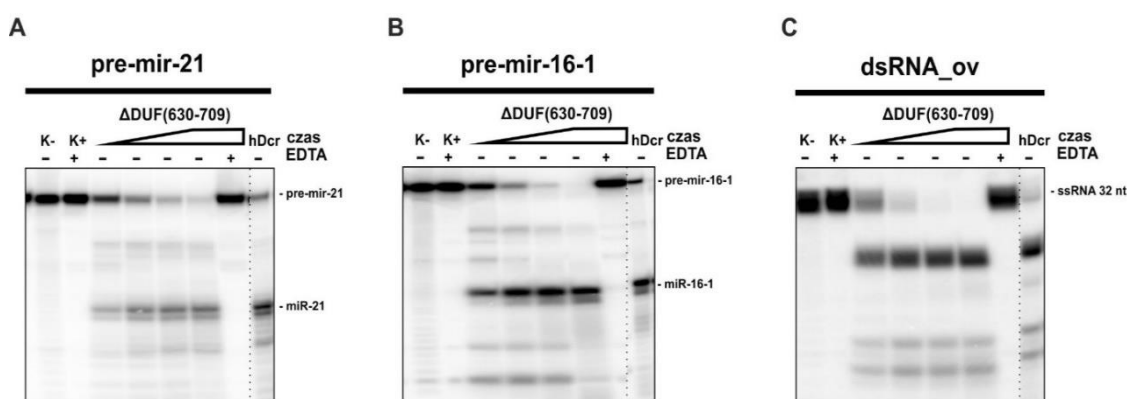
Uzyskane wyniki wykazały, że w przypadku pre-mir-21, wariant delecyjny  $\Delta$ DUF(625-752) produkował śladowe ilości produktu o długości miRNA (21 nt), w porównaniu do białka typu dzikiego, które po 120 min przecięło prawie cały substrat (**Rysunek 8A**). W przypadku pre-mir-16-1, zarówno wariant delecyjny jak i białko hDicer (hDcr) produkowały wydajnie miR-16-1-5-p, a substrat pre-mir-16-1 został całkowicie przecięty po 120 min inkubacji (**Rysunek 8B**). W reakcjach cięcia substratu dsRNA\_ov zaobserwowano produkcję śladowych ilości produktu o długości 21 nt przez wariant  $\Delta$ DUF(625-752), w przeciwieństwie do kontroli z hDcr, w której obserwowano wydajną produkcję siRNA (**Rysunek 8C**).



**Rysunek 8. Badanie aktywności RNazowej wariantu delecyjnego  $\Delta$ DUF(625-752) hDicer.** (A-C) Wyniki reakcji cięcia substratu: pre-mir-21 (A), pre-mir-16-1 (B) oraz dsRNA\_ov (C). Reakcje były prowadzone w czasie 10, 30, 60 oraz 120 min (wzrastający czas oznaczono trójkątem) i zawierały 18 nM białka  $\Delta$ DUF(625-752) lub pełnej długości białko hDicer (hDcr). (K-) kontrola bez białka zawierająca substrat w buforze reakcyjnym. (K+) kontrola bez białka zawierająca substrat w buforze reakcyjnym z dodatkiem EDTA. (+EDTA) kontrola, w której do reakcji cięcia dodano EDTA. (AH) Wzorec długości uzyskany za pomocą reakcji alkalicznej hydrolizy wyznakowanego z użyciem  $[\gamma^{32}\text{P}]\text{-ATP}$  substratu RNA o długości 32 nt. \*Gwiazdką zaznaczono wyznakowany z użyciem  $[\gamma^{32}\text{P}]\text{-ATP}$  koniec 5' substratu. Rozdziały prowadzono w 15% PAA. Struktury cząsteczek zostały wygenerowane z użyciem programu RNAstructure. Rysunek stworzony na podstawie publikacji Szczepanska i in. [26].

Podsumowując, badania aktywności RNazowej wariantu  $\Delta$ DUF(625-752) przeprowadzone w warunkach *in vitro* wykazały, że aminokwasy 625-752 hDicer (rejon domeny DUF283 oraz fragmenty oskrzydłujące tę domenę) mogą mieć znaczenie w rozpoznaniu substratów hDicer, co może wpływać na dostępność substratów dla domen RNazowych hDicer i produkcję srRNA.

W celu porównania aktywności RNazowej wariantu  $\Delta$ DUF(630-709) z wariantem  $\Delta$ DUF(625-752), wykonano serie reakcji cięcia z wykorzystaniem tych samych kanonicznych substratów, tj. pre-mir-16-1, pre-mir-21 oraz dsRNA\_ov. Wykonane analizy wykazały, że wariant  $\Delta$ DUF(630-709) hydrolizował wszystkie testowane substraty z niewiele niższą wydajnością niż hDcr, co było zgodne z wynikami grupy prof. J. Doudny, która wykazała, że wprowadzona delecja obniżyła wydajność cięcia kanonicznych substratów przez wariant  $\Delta$ DUF(630-709), w porównaniu do wydajności cięcia tych samych substratów przez białko typu dzikiego **Rysunek 9A-C** [30].



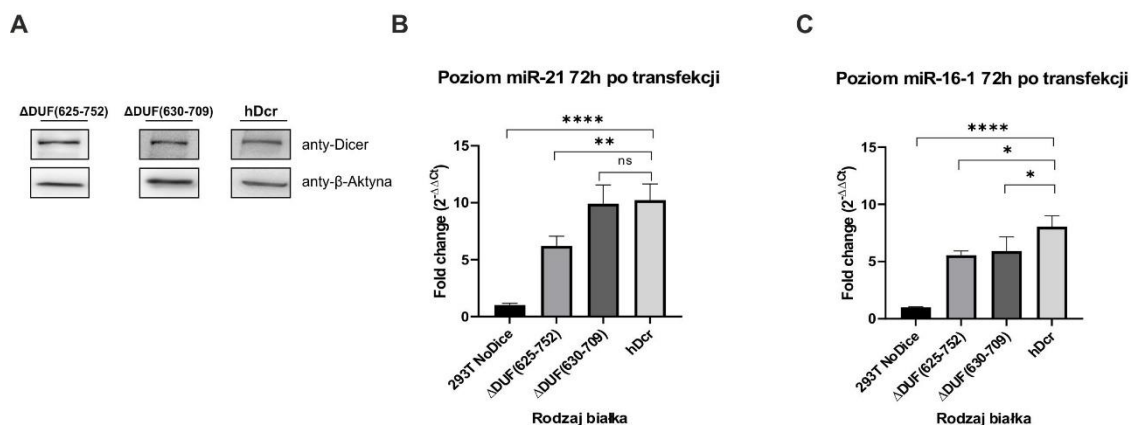
**Rysunek 9. Badanie aktywności RNazowej wariantu delecyjnego  $\Delta$ DUF(630-709) hDicer.** (A-C) Wyniki reakcji cięcia substratu: pre-mir-21 (A), pre-mir-16-1 (B) oraz dsRNA\_ov (C). Reakcje były prowadzone w czasie 10, 30, 60 oraz 120 min (wzrastający czas oznaczono trójkątem) i zawierały 18 nM białka  $\Delta$ DUF(630-709) lub pełnej długości białko hDicer (hDcr). (K-) kontrola bez białka zawierająca substrat w buforze reakcyjnym. (K+) kontrola bez białka zawierająca substrat w buforze reakcyjnym z dodatkiem EDTA. (+EDTA) kontrola, w której do reakcji cięcia dodano EDTA. Substraty posiadały wyznakowany z użyciem [ $\gamma$ <sup>32</sup>P]-ATP koniec 5'. Rozdziały prowadzono w 15% PAA. Rysunek stworzony na podstawie publikacji Szczepańska i in. [26].

#### 4.1.3.2 Badanie aktywności RNazowej wariantów $\Delta$ DUF(625-752) oraz $\Delta$ DUF(630-709) w warunkach *in cellulo*

Kolejnym etapem badań było określenie aktywności RNazowej wariantu  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709) wobec wcześniej testowanych kanonicznych substratów pre-miRNA: pre-mir-16-1 oraz pre-mir-21, w warunkach *in cellulo*. Oba badane pre-miRNA są wydajnie produkowane w komórkach HEK 293T. Komórki HEK 293T NoDice transfekowano odpowiednim plazmidem kodującym wariant  $\Delta$ DUF(625-752) lub  $\Delta$ DUF(630-709), a następnie, po upływie 72 h, zbierano osady komórkowe (Metody 3.2.13). Frakcje całkowitego RNA izolowano za pomocą odczynnika TRIzol (Invitrogen), zgodnie z protokołem producenta. Następnie, proces produkcji miRNA z ramienia 5' badano z wykorzystaniem reakcji RT-qPCR (ang. *reverse transcription quantitative real-time PCR*) [91]. Kontrolę negatywną stanowiły komórki HEK 293T NoDice potraktowane tylko odczynnikami do transfekcji, a kontrolę pozytywną stanowiły komórki HEK 293T NoDice transfekowane plazmidem kodującym pełnej długości hDcr (Metody 3.2.21).

Otrzymane wyniki pozwoliły określić względny poziom miR-21-5p oraz miR-16-1-5p w komórkach HEK 293T NoDice, po 72 h od transfekcji, do których wprowadzono plazmid kodujący:  $\Delta$ DUF(625-752),  $\Delta$ DUF(630-709) lub hDcr. Stwierdzono, że przy podobnych poziomach produkcji białek w komórkach (**Rysunek 10A**), wariant  $\Delta$ DUF(625-752) produkował oba miRNA z niższą wydajnością w porównaniu do białka typu dzikiego. Dokładniej, obniżenie poziomu miRNA wynosiło ok. 40% dla miR-21-5p i ok. 30% dla miR-16-1-5p. Natomiast w przypadku wariantu  $\Delta$ DUF(630-709) poziom produkcji miRNA obniżył się o ok. 30% dla miR-16-1-5p, a dla miR-21-5p nie odnotowano istotnych zmian w poziomie produkcji, w porównaniu do hDcr (**Rysunek 10B, C**).



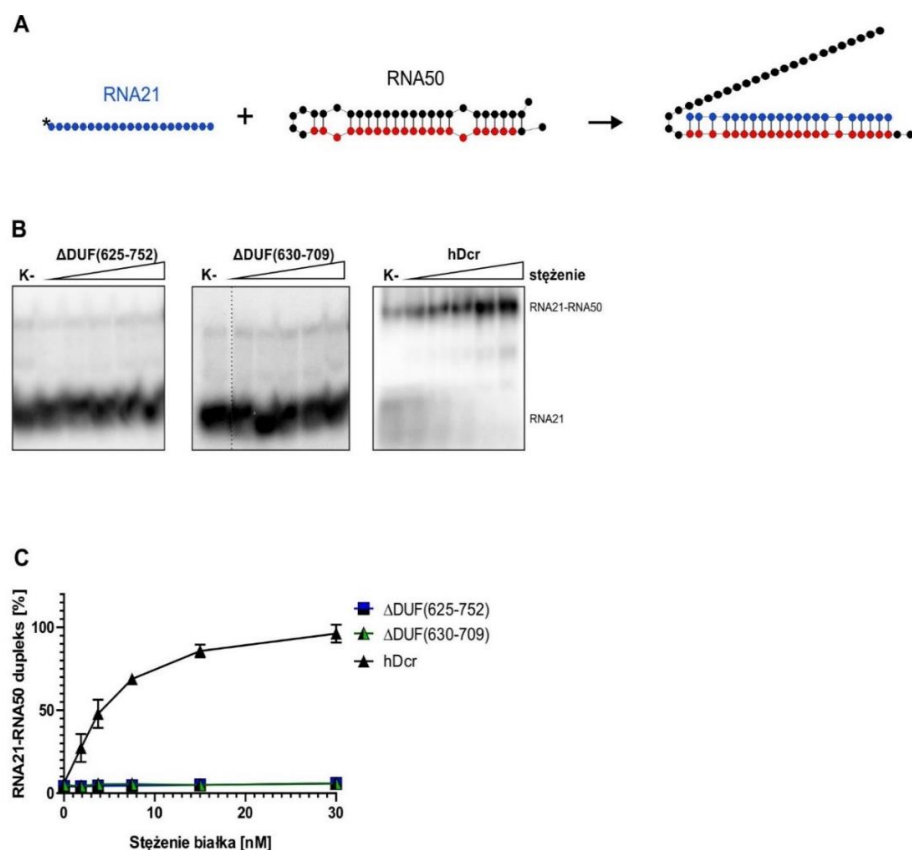


**Rysunek 10. Porównanie poziomu produkcji wybranych miRNA przez warianty delecyjne hDicer.** (A) Określenie poziomu wskazanych białek w komórkach HEK 293T NoDice za pomocą metody Western blot. (B, C) Analiza RT-qPCR względnego poziomu dojrzałych miR-21 (B) oraz miR-16-1 (C) w komórkach HEK 293T NoDice (kontrola negatywna) lub w komórkach 293T NoDice transfekowanych plazmidem kodującym wskazane białko. Poziomy miRNA normalizowano względem poziomu referencyjnego genu U6. Wyniki uzyskane z trzech niezależnych powtórzeń biologicznych przedstawiono jako średnią wartość  $Cq \pm SD$ . Dane analizowano za pomocą jednokierunkowego testu ANNOVA, a następnie za pomocą testu wielokrotnych porównań Dunetta. Uzyskane wartości p są następujące, dla miR-16-1: \*  $p < 0,0220$  \*\*\*\*  $p < 0,0001$ ; dla miR-21: \*\*  $p < 0,0079$  \*\*\*\*  $p < 0,0001$ . Rysunek stworzony na podstawie publikacji Szczepańska i in. [26].

Po przeanalizowaniu otrzymanych rezultatów stwierdzono, że oba warianty,  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709), wykazują aktywność RNazową, choć aktywność ta różniła się w zależności od wybranego substratu oraz warunków reakcji. Wariant  $\Delta$ DUF(625-752) w warunkach *in vitro* wykazywał aktywność RNazową tylko wobec substratu pre-miR-16-1. Natomiast w warunkach *in cellulo*, wariant  $\Delta$ DUF(625-752) produkował miR-16-1-5p oraz miR-21-5p, ale z niższą wydajnością w porównaniu do hDcr, odpowiednio o 30% i 40%. W przypadku wariantu  $\Delta$ DUF(630-709) odnotowano niewielkie obniżenie wydajności produkcji obu badanych miRNA w warunkach *in vitro*, natomiast w warunkach *in cellulo* zaobserwowano obniżenie wydajności produkcji miR-16-1-5p o 30%, a dla miR-21-5p nie odnotowano takich różnic, w porównaniu do białka typu dzikiego.

#### 4.1.3.3 Badanie aktywności wspierania parowania RNA-RNA przez warianty $\Delta$ DUF(625-752) oraz $\Delta$ DUF(630-709)

Biorąc pod uwagę postawioną hipotezę, iż domena DUF283 jest kluczowa dla aktywności wspierania parowania cząsteczek kwasów nukleinowych hDicer, kolejnym etapem badań było sprawdzenie jak usunięcie domeny DUF283 wpłynie na wspomnianą aktywność hDicer. W tym celu przeprowadzono badanie aktywności wspierania parowania względem kwasów nukleinowych (ang. *nucleic acids annealing activity*) wykorzystując dwie wzajemnie komplementarne cząsteczki RNA: RNA50 o długości 50 nt oraz RNA21 o długości 21 nt. Dłuższa cząsteczka RNA50 przyjmuje strukturę typu spinki do włosów i zawiera w pełni komplementarne miejsce docelowe dla cząsteczki RNA21 (**Rysunek 11A**). Ta para substratów została wykorzystana we wcześniejszych badaniach dotyczących aktywności wspierania parowania cząsteczek kwasów nukleinowych przez hDicer [93]. Przed rozpoczęciem reakcji, cząsteczkę RNA50 inkubowano w temperaturze 95°C, a następnie powoli schładzano do temperatury pokojowej. Następnie cząsteczkę RNA50 mieszało w stosunku molowym 1:1 z 5'-[ $\gamma^{32}$ P]-ATP-RNA21 i inkubowano przez 30 min z rosnącymi stężeniami (1.88, 3.75, 7.5, 15, 30 nM)  $\Delta$ DUF(625-752),  $\Delta$ DUF(630-709) lub hDcr (kontrola pozytywna) (**Rysunek 11B**). Kontrolę ujemną stanowiła reakcja bez dodatku białka. Dokładny opis procedury zawarty jest w rozdziale: Metody 3.2.19.



**Rysunek 11. Badanie aktywności wspierania parowania komplementarnych par zasad przez wariant  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709).** (A) Schematyczne przedstawienie eksperymentu. Częsteczka RNA21 (kolor niebieski) jest w pełni komplementarna do 21 nt sekwencji docelowej częsteczki RNA50 (kolor czerwony). (B) Wyniki badania aktywności wspierania parowania cząsteczek kwasów nukleinowych. Reakcje prowadzono w czasie 30 min przy wzrastających stężeniach białek  $\Delta$ DUF(625-752),  $\Delta$ DUF(630-709) lub hDcr (1.88, 3.75, 7.5, 15, 30 nM, wzrastające stężenia oznaczono trójkątem). (K-) kontrola bez białka. Rozdział prowadzono w 10% żelu PAA. (C) Graficzne przedstawienie wyników pochodzących z badania aktywności opiekuńczej białek względem kwasów nukleinowych. Słupki błędów przedstawiają odchylenie standardowe (SD) z trzech oddzielnych eksperymentów. Rysunek przygotowany na podstawie publikacji Szczepanska i in. [26].

Na podstawie uzyskanych wyników zaobserwowano, że w odróżnieniu od białka typu dzikiego, zarówno wariant  $\Delta$ DUF(625-752) jak i  $\Delta$ DUF(630-709), w zastosowanych warunkach reakcyjnych nie wspierały hybrydyzacji 21-nt RNA do komplementarnej wobec tej cząsteczki sekwencji, obecnej w rejonie trzonu RNA50 (**Rysunek 11**). Zebrane dane w znacznym stopniu potwierdzają hipotezę, że domena DUF283 jest niezbędna dla aktywności wspierającej parowanie cząsteczek kwasów nukleinowych hDicer.

Wyniki przedstawione w tej części doktoratu zostały opublikowane w: „*The Significance of the DUF283 Domain for the Activity of Human Ribonuclease Dicer*”, 2021, *International Journal of Molecular Sciences*, Szczepańska i in. (**Załącznik 1**).

## 4.2 Rola domeny DUF283 hDicer w wiązaniu komórkowych RNA

Biorąc pod uwagę wyniki powyżej opisanych badań zadano sobie pytanie jakie może być znaczenie domeny DUF283 hDicer dla wiązania komórkowych RNA przez tę rybonukleazę. W celu znalezienia odpowiedzi na to pytanie zastosowano metodę typu CLIP (ang. *crosslinking immunoprecipitation*), która umożliwia izolację kompleksów białko•RNA powstających w żywych komórkach.

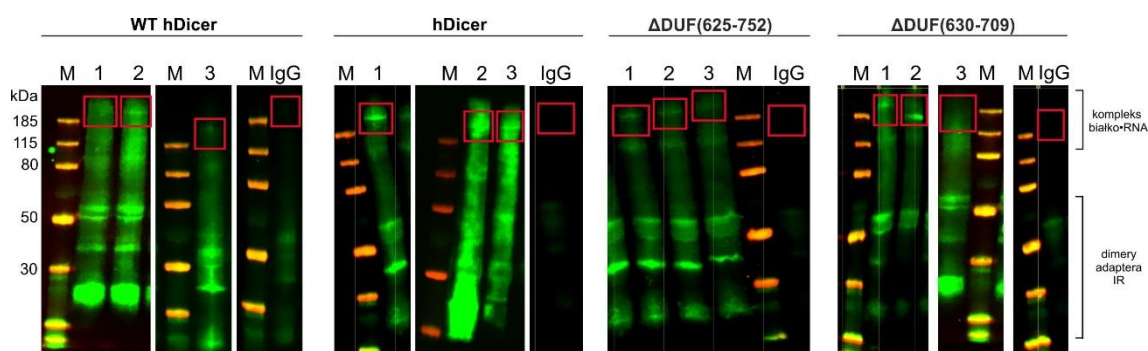
Warto wspomnieć w tym miejscu, iż badania prowadzone przez grupę prof. N. Rajewskiego, dotyczące puli komórkowych RNA wiązanych przez rybonukleazę Dicer człowieka oraz *C. elegans*, zostały wykonane z wykorzystaniem metody PAR-CLIP (ang. *photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation*) [79]. Metoda ta charakteryzuje się wysoką czułością dzięki zastosowaniu w badaniach fotoreaktywnej 4-tiourydyny (4-SU), która pozwala na dokładne określenie miejsca wiązania białka z RNA, z rozdzielczością nukleotydową. Metoda PAR-CLIP posiada również wady, które mogą wpływać na otrzymywane wyniki. Do tych wad można zaliczyć cytotoksyczność wykorzystywanej w badaniach 4-SU, czy też różny poziom jej wnikania do komórek [94, 95]. Dlatego, w ramach realizacji niniejszej pracy doktorskiej, do zidentyfikowania pul komórkowych RNA wiązanych przez hDicer i warianty  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709), postanowiono wykorzystać inną pochodną metody CLIP, dokładniej metodę irCLIP (ang. *infrared crosslinking immunoprecipitation*). W przypadku tej metody główną jej zaletą, w porównaniu do metody PAR-CLIP, jest wykorzystanie adaptera z barwnikiem bliskiej podczerwieni (IR, ang. *infrared*) do wizualizacji kompleksów białko•RNA, przez co nie ma potrzeby wprowadzania do komórek cytotoksycznej 4-SU. Dzięki zastosowaniu tego adaptera metoda irCLIP jest także metodą bardziej czułą, w porównaniu do metody PAR-CLIP [92, 96].

### 4.2.1 Izolacja kompleksów białko•RNA z wybranych ludzkich linii komórkowych

W celu zbadania znaczenia domeny DUF283 hDicer w wiązaniu komórkowych RNA wykorzystano następujące modele badawcze – linie komórkowe: (i) HEK 293T – linia typu dzikiego (WT hDicer), (ii) HEK 293T NoDice transfekowane plazmidem kodującym pełnej długości białko hDicer (tzw. *rescue control*, hDicer), (iii) HEK 293T NoDice transfekowane plazmidem kodującym  $\Delta$ DUF(625-752) oraz (iv) HEK 293T NoDice transfekowane plazmidem kodującym  $\Delta$ DUF(630-709). Charakterystyka

aktywności wariantów  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709) została przedstawiona w poprzednim rozdziale (Rozdział 4.1.3).

Po 72 h od transfekcji, komórki naświetlano światłem UV (254 nm), w celu utworzenia wiązań kowalencyjnych pomiędzy białkami, a związanymi z nimi cząsteczkami RNA (ang. *UV crosslinking*). Komórki lizowano, a następnie izolowano kompleksy białko•RNA za pomocą złoża magnetycznego opłaszczonego przeciwciałem specyficznym wobec C-końca białka hDicer (Anti-Dicer Antibody sc-136979, Santa Cruz Biotechnology) lub przeciwciałem specyficznym wobec białka IgG, które służyło jako kontrola izolacji specyficznych kompleksów białko•RNA (Normal Mouse IgG sc-2025, Santa Cruz Biotechnology). Kolejnym etapem była defosforylacja końca 3' cząsteczek RNA związanych z białkiem oraz przyłączenie do tych cząsteczek adaptera z barwnikiem IR. Odczepione od złoża kompleksy białko•RNA poddano rozdzielowi elektroforetycznemu w żelu typu SDS-PAGE, a następnie przeprowadzono ich transfer na membranę nitrocelulozową. Następnie, kompleksy białko•RNA wizualizowano z wykorzystaniem urządzenia Odyssey Li-COR CLx Imager (**Rysunek 12**).



**Rysunek 12.** Wizualizacja kompleksów hDicer•RNA,  $\Delta$ DUF(625-752)•RNA oraz  $\Delta$ DUF(630-709)•RNA wyizolowanych za pomocą metody irCLIP, po transferze na membranę nitrocelulozową. Zaznaczone czerwoną ramką fragmenty membrany wykorzystano w dalszych analizach. (M) marker wielkości PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific). IgG – kontrola specyficzności izolacji kompleksów. Kompleksy wizualizowano z wykorzystaniem aparatu Odyssey Li-COR CLx Imager dzięki zastosowaniu barwnika bliskiej podczerwieni (IR).

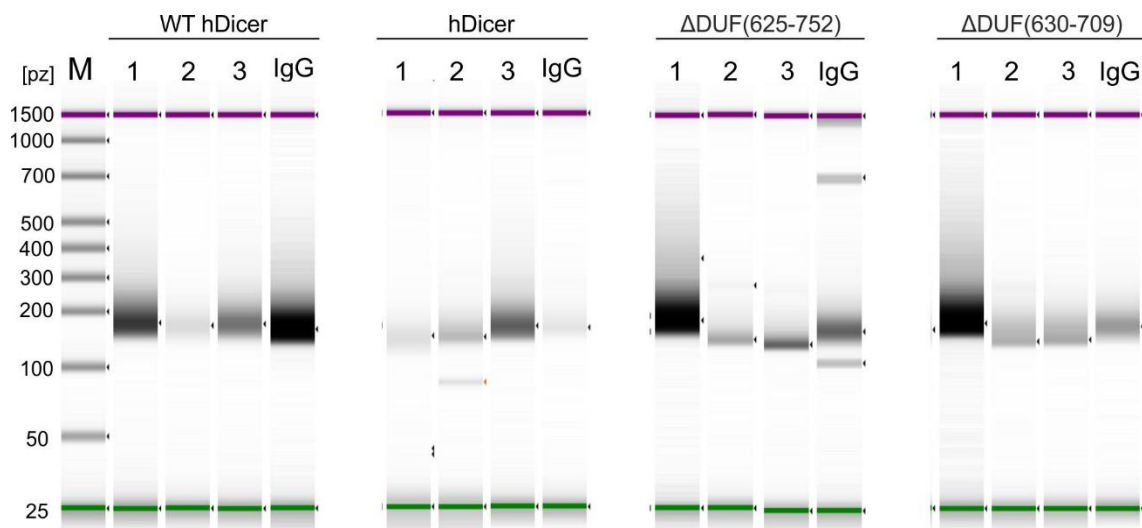
Z membrany wycięto fragmenty, w których znajdowały się pożądanе kompleksy białko•RNA (fragmenty zaznaczone czerwoną ramką na **Rysunku 12**). Kolejnym etapem było odzyskanie RNA z kompleksów białko•RNA. W tym celu wycięte fragmenty membrany inkubowano z proteinazą K, która trawi białka obecne w kompleksach, pozostawiając jedynie pojedyncze aminokwasy związane z RNA w miejscu gdzie

wytworzyło się wiązanie kowalencyjne. RNA oczyszczono metodą fenol:chloroform:alkohol izoamylowy i strącano za pomocą 96% etanolu. Oczyszczony RNA poddano reakcji odwrotnej transkrypcji, a otrzymany cDNA cyrkularyzowano i następnie amplifikowano za pomocą reakcji PCR. Tak uzyskany materiał został wykorzystany do przygotowania bibliotek cDNA do reakcji sekwencjonowania NGS. Szczegółowy opis izolacji kompleksów białko•RNA metodą irCLIP został opisany w rozdziale: Metody 3.2.27.

#### **4.2.2 Analiza ilościowo-jakościowa przygotowanych bibliotek cDNA do sekwencjonowania NGS**

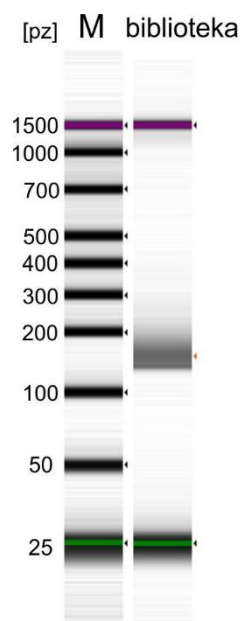
Gotowe biblioteki cDNA zostały sprawdzone pod względem ilościowym i jakościowym. Do zmierzenia stężenia otrzymanych bibliotek cDNA wykorzystano zestaw QuantiFlor® ONE dsDNA (Promega). Próby do analiz przygotowano zgodnie z protokołem producenta i analizowano z wykorzystaniem urządzenia Quantus™ Fluorometr (Promega) (Metody 3.2.20). W zależności od próbki, uzyskano stężenia w przedziale 0,15-7,5 ng/μl.

Następnie jakość bibliotek została sprawdzona dzięki wykorzystaniu bioanalizatora Agilent Bioanalyzer 2100 (Metody 3.2.28). Bioanalizator to urządzenie, które służy do przeprowadzenia wysokiej rozdzielczości elektroforezy kapilarnej typu TapeStation, a wyniki tego rozdziału umożliwiają ocenę jakościową próby tj. ocenę rozkładu długości fragmentów kwasu nukleinowego w badanej próbce, wskazując także ewentualne zanieczyszczenie próby. Na **Rysunku 13** zaprezentowano wyniki kapilarnego rozdziału elektroforetycznego gotowych bibliotek cDNA. Oczekiwana długość fragmentów dla analizowanych cDNA powinna zawierać się w przedziale 145-200 pz. Rozkład długości w poszczególnych bibliotekach jest zależny od długości związanej przez badane białko cząsteczki RNA. Ponadto, do każdej związanej cząsteczki RNA dodawany jest 131 nt adapter, który zawiera m. in. sekwencję komplementarną do sekwencji starterów P3/P5 Solexa oraz sekwencję UMI (ang. *Unique Molecular Identifiers*). Dzięki przypisaniu odpowiednich sygnałów (w postaci pików) pochodzących od analizowanej próby do pików markera długości, możliwe jest określenie długości fragmentów w próbce. Przeprowadzona analiza wykazała, że rozkład długości fragmentów cDNA w badanych próbach znajdował się w przedziale 160-400 nt.



**Rysunek 13. Obraz rozdziału elektroforetycznego typu TapeStation bibliotek cDNA.** Rozdział wykonano z użyciem chipów High Sensitivity D1000 ScreenTape®. (M) marker wielkości High Sensitivity D1000 Ladder (Agilent). Do każdej próby dodawane są 2 markery wielkości: krótki o długości 25 pz (zaznaczony na zielono) oraz długi o długości 1500 pz (zaznaczony na fioletowo).

Analiza ilościowo-jakościowa bibliotek cDNA wykazała, że każda z bibliotek może zostać wykorzystana w reakcji sekwencjonowania NGS. Kolejnym krokiem było przygotowanie jednej, tzw. znormalizowanej biblioteki cDNA (Metody 3.2.29). Końcowe stężenie każdej z bibliotek wynosiło 0,05 ng/μl. Kontrolę IgG potraktowano objętościowo, tj. z każdej biblioteki kontroli IgG dodano po 2 μl do biblioteki końcowej. Następnie znormalizowaną bibliotekę sprawdzono pod względem ilościowym i jakościowym wykorzystując wcześniej opisaną procedurę, a pomiar stężenia wykonano w trzech powtórzeniach technicznych. Średnie stężenie biblioteki wynosiło 0,315 ng/μl. Na końcu wyznaczono stężenie molowe biblioteki, które wyniosło ~3 nM. Na **Rysunku 14** zaprezentowano wyniki rozdziału elektroforetycznego znormalizowanej biblioteki fragmentów cDNA z wykorzystaniem analizy typu TapeStation. Średni rozkład długości biblioteki wynosił 160 pz, co wskazuje na oczekiwany rozkład długości fragmentów w próbce.



**Rysunek 14. Obraz rozdziału elektroforetycznego typu TapeStation znormalizowanej biblioteki.** Rozdział wykonano z użyciem chipu High Sensitivity D1000 ScreenTape®. (M) marker wielkości High Sensitivity D1000 Ladder (Agilent). Do każdej próby dodawane są 2 markery wielkości: krótki o długości 25 pz (zaznaczony na zielono) oraz długi o długości 1500 pz (zaznaczony na fioletowo).

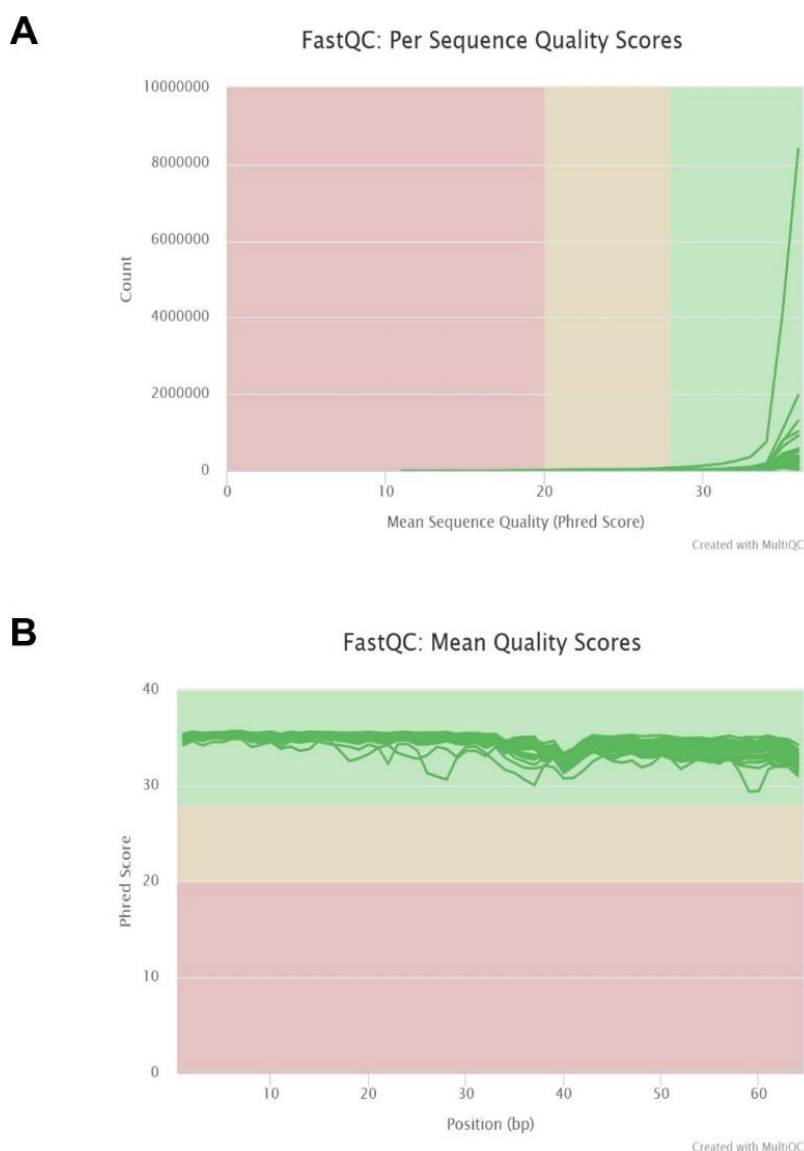
W ten sposób znormalizowana biblioteka została następnie przygotowana do sekwencjonowania NGS, zgodnie z protokołem NextSeq 500 and 550 Sequencing Systems: Denature and Dilute Libraries Guide (#15048776v16 (Illumina)). Zastosowano sekwencjonowanie o wysokiej przepustowości NextSeq 500/550 High Output Kit v2.5 (75 Cycles) (Illumina). Sekwencjonowanie przebiegło z użyciem sekwenatora Illumina NextSeq550 (Illumina) (Metody 3.2.30) i pozwoliło uzyskać odczyty 1x75 pz do formatu *.fastq*.

#### **4.2.3 Raport jakości uzyskanych danych pochodzących z sekwencjonowania bibliotek cDNA**

Surowe dane pochodzące z reakcji sekwencjonowania NGS poddano kontroli jakości, wstępnej obróbce i uporządkowaniu, zgodnie z opisem zawartym w punkcie: Metody 3.2.31. Tak przygotowane dane analizowano zgodnie z protokołem nf-core/clipseq, który opisuje w jaki sposób należy analizować dane pochodzące z sekwencjonowania CLIP-seq [97]. Dzięki zastosowaniu wspomnianego protokołu wygenerowano końcowy raport MultiQC, który stanowi podsumowanie jakości uzyskanych danych.



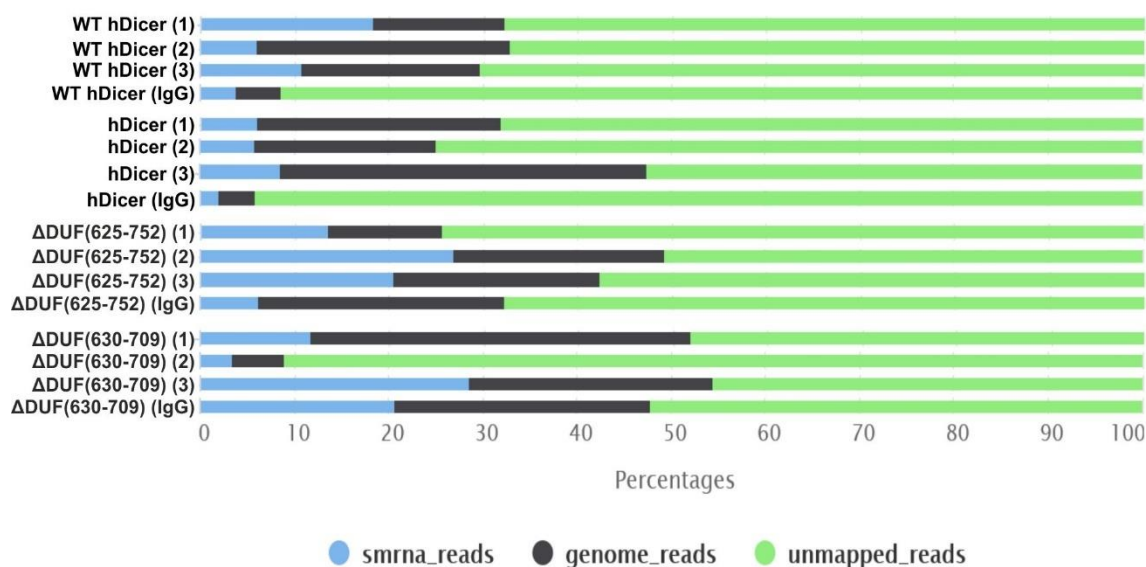
Przeprowadzona analiza wykazała bardzo dobrą jakość sekwencjonowania, na co wskazuje wartość współczynnika logarytmicznego Q, którego wartość dla wszystkich badanych prób wynosiła  $Q \geq 35$ . Wartości takie świadczą o bardzo niskim prawdopodobieństwie popełnienia błędu przez polimerazę podczas sekwencjonowania (Rysunek 15).



**Rysunek 15. Raport FastQC przedstawiający jakość otrzymanych odczytów.** (A) Wykres pokazujący średnią jakość sekwencji odczytów dla każdej próby. Oś OY (Count) określa łączną liczbę otrzymanych odczytów dla każdej próby. Oś OX (Mean Sequence Quality (Phred Score)) określa średnią jakość dla każdej próby. (B) Wykres pokazujący średnią wartość jakości dla każdej pozycji nukleotydu, który został odczytany w trakcie sekwencjonowania. Oś OY (Phred Score) na wykresie pokazuje wyniki jakości dla każdej z prób. Oś OX (Position (bp)) wskazuje pozycję nukleotydu, dla której nastąpił odczyt. Wszystkie badane próby pozycjonują się na zielonym tle ( $Q \geq 35$ ), co świadczy o bardzo dobrej jakości otrzymanych odczytów. Wykresy zostały wygenerowane przez narzędzie MultiQC.

Następnie wygenerowano ogólne podsumowanie przypisania (tzw. zmapowania) wszystkich odczytów uzyskanych dla poszczególnych prób. Odczyty te były mapowane do krótkich cząsteczek RNA o długości 18-30 nt, tzw. smrna\_reads, oraz do genomu referencyjnego człowieka (wersja hg38) tzw. genome\_reads (**Rysunek 16**). Odczyty, które nie zostały zmapowane oznaczone są na **Rysunku 16** jako odczyty niezmapowane, tzw. unmapped\_reads. Średnia liczba odczytów dla próby WT hDicer wynosiła: (i) 58 633 zmapowanych odczytów do smRNA, (ii) 115 256 odczytów zmapowanych do genomu oraz (iii) 376 340 odczytów niezmapowanych. Dla próby hDicer otrzymano średnią liczbę odczytów (i) zmapowanych do smRNA: 28 791, (ii) odczytów zmapowanych do genomu: 147 221 oraz (iii) niezmapowanych odczytów: 231 991. W przypadku próby  $\Delta$ DUF(625-752) średnia liczba odczytów wynosiła: (i) 146 326 odczytów zmapowanych do smRNA, (ii) 131 366 odczytów zmapowanych do genomu oraz (iii) 363 848 odczytów niezmapowanych. Dla próby  $\Delta$ DUF(630-709) otrzymano średnią liczbę odczytów (i) zmapowanych do smRNA: 1 128 509, (ii) odczytów zmapowanych do genomu: 1 172 368 oraz (iii) niezmapowanych odczytów: 1 972 261. Można zauważyć, że wiele z otrzymanych odczytów nie zostało przypisanych do smRNA czy referencyjnego genomu, co związane jest z faktem, że uzyskane odczyty były za krótkie lub były dopasowane z równą dokładnością do zbyt wielu miejsc w genomie referencyjnym. Dodatkowo, podczas sekwencjonowania jako kontrolę jakości i wydajności reakcji wykorzystano bibliotekę bakteriofaga PhiX.

## Mapping

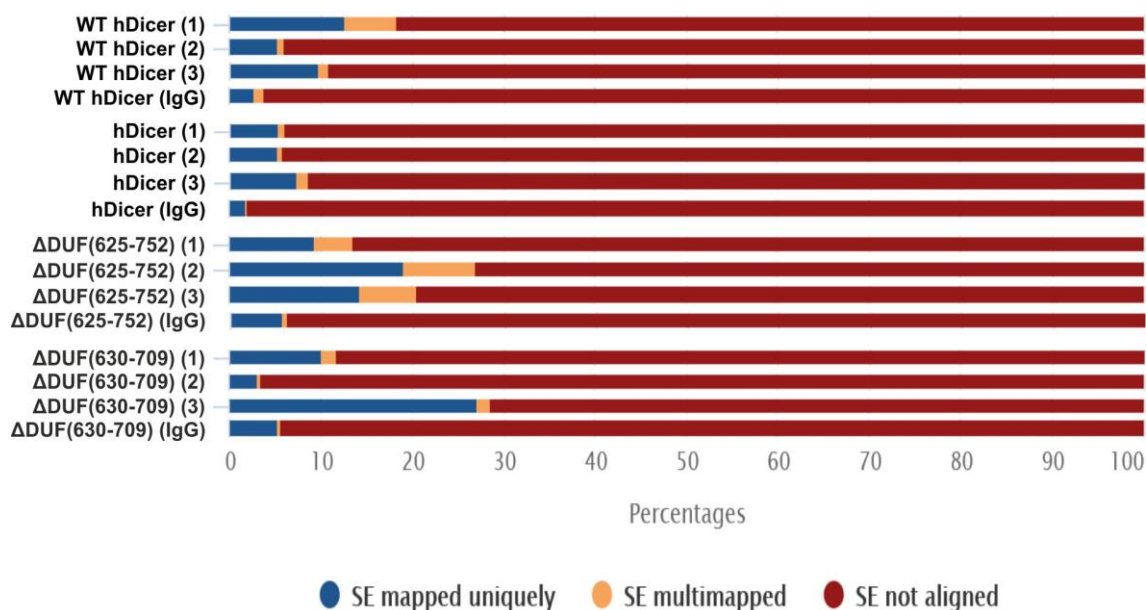


Created with MultiQC

**Rysunek 16. Procentowe przedstawienie zmapowania otrzymanych odczytów dla każdej badanej próby.** Odczyty przyrównywano do krótkich RNA (smrna\_reads, kolor niebieski) oraz do genomu człowieka hg38 (genome\_reads, kolor szary). Odczyty niezmapowane oznaczono kolorem zielonym (unmapped\_reads). Wykres został wygenerowany przez narzędzie MultiQC.

W celu przypisania otrzymanych odczytów do genomu referencyjnego człowieka (hg38) najpierw wykonano wstępne przypisanie odczytów (ang. *pre-mapping*) do sekwencji rRNA oraz tRNA. Do tej analizy wykorzystano narzędzie Bowtie 2, które pozwala szybko i wydajnie przyrównać otrzymane odczyty do krótkich sekwencji RNA zawartych w genomie referencyjnym. Wstępne mapowanie dla wszystkich badanych prób przedstawiono na wykresie, na **Rysunku 17**. Wykres ten pokazuje procent odczytów zmapowanych do referencyjnego genomu na trzy różne sposoby: (i) odczyty zmapowane jednoznacznie (ang. *SE mapped uniquely*), które mają tylko jedno dopasowanie w genomie referencyjnym; (ii) odczyty zmapowane kilkakrotnie (ang. *SE multimapped*) do referencyjnego genomu (zmapowanie do więcej niż 1 miejsca) oraz (iii) odczyty niezmapowane (ang. *SE not aligned*).

## Bowtie 2: SE Alignment Scores

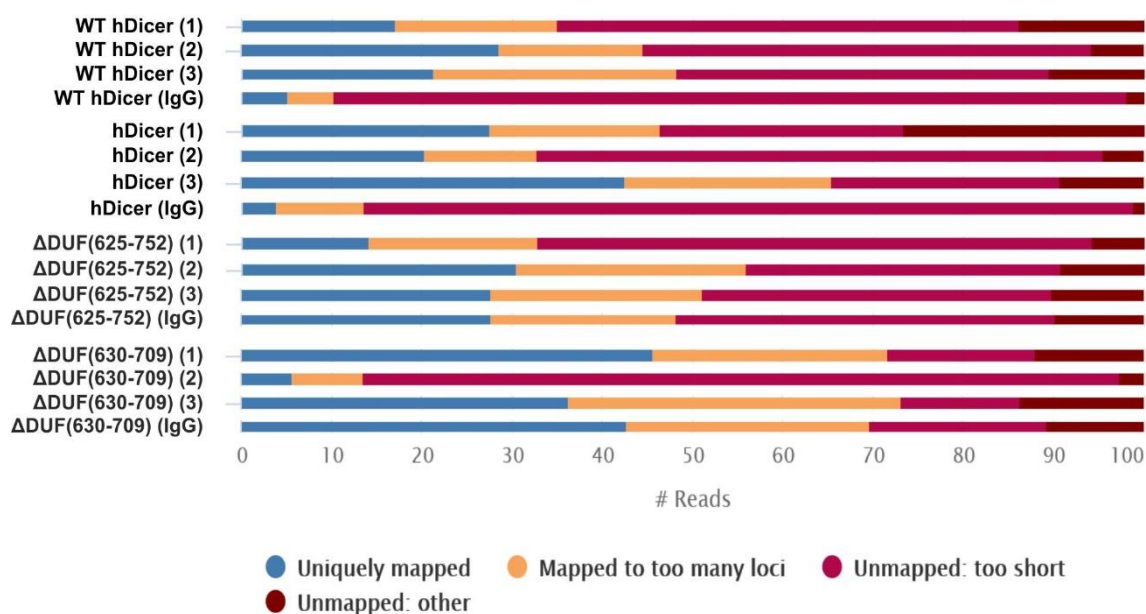


Created with MultiQC

**Rysunek 17.** Wykres obrazujący zmapowanie odczytów do sekwencji rRNA oraz tRNA znajdujących się w genomie referencyjnym człowieka (hg38), wyrażone w procentach. Kolorem niebieskim oznaczono odczyty zmapowane jednoznacznie (SE mapped uniquely), kolorem pomarańczowym oznaczono odczyty zmapowane kilkakrotnie (SE multimapped) oraz kolorem czerwonym oznaczono odczyty niezmapowane (SE not aligned) do genomu referencyjnego człowieka hg38. Wykres został wygenerowany przez narzędzie MultiQC.

Kolejnym etapem było przypisanie odczytów, niezmapowanych przez Bowtie 2, dla badanych prób do genomu referencyjnego człowieka za pomocą narzędzia STAR (ang. *mapping*), które pozwala przyrównać odczyty do poszczególnych miejsc w ludzkim genomie (**Rysunek 18**). W przypadku narzędzia STAR występują cztery rodzaje mapowania: (i) odczyty zmapowane jednoznacznie (ang. *uniquely mapped*); (ii) odczyty zmapowane do więcej niż 5 miejsc w genomie (ang. *mapped to too many loci*); (iii) odczyty krótsze niż 12 nt, które nie zostały zmapowane (ang. *unmapped: too short*) oraz (iv) odczyty niezmapowane (ang. *unmapped other*). Liczba niezmapowanych odczytów do genomu referencyjnego może wynikać z błędów różnych enzymów wykorzystanych w procesie produkcji bibliotek cDNA, czego wynikiem jest niesparowanie odczytu z referencyjnym genomem człowieka.

## STAR: Alignment Scores

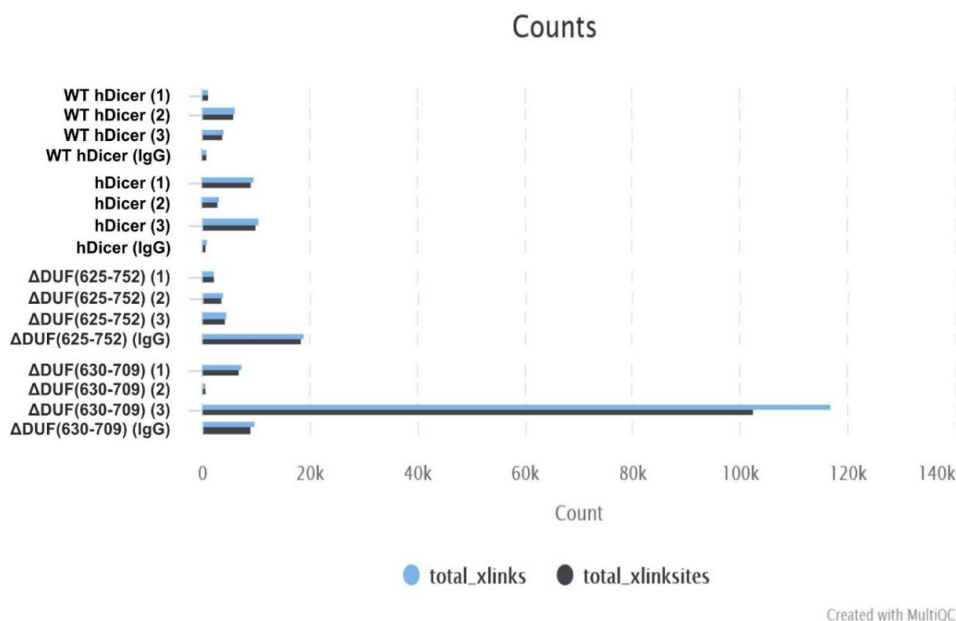


Created with MultiQC

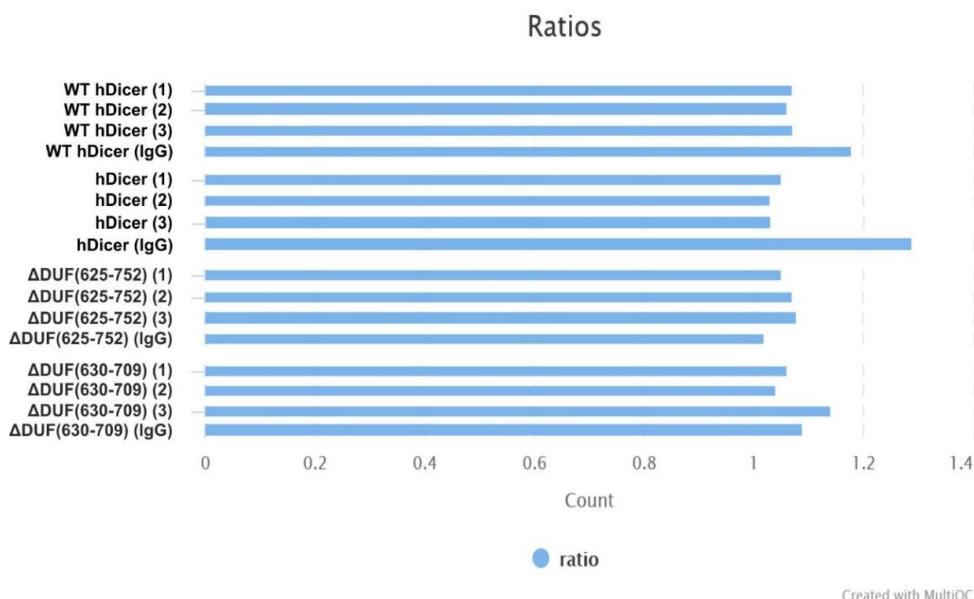
**Rysunek 18.** Wykres obrazujący zmapowanie odczytów do genomu referencyjnego człowieka (hg38), wyrażone w procentach. Kolorem niebieskim oznaczono odczyty zmapowane jednoznacznie (uniquely mapped); kolorem pomarańczowym oznaczono odczyty zmapowane do więcej niż 5 miejsc w genomie (mapped to too many loci); kolorem czerwonym oznaczono odczyty krótsze niż 12 nt, które nie zostały zmapowane (unmapped: too short) oraz kolorem brązowym oznaczono inne niezmapowane odczyty (unmapped other). Wykres został wygenerowany przez narzędzie MultiQC.

Określono również liczbę zdarzeń wiązań pomiędzy białkiem a cząsteczkami RNA (ang. *crosslink identification*). W tym celu wykorzystano dwa narzędzia: UMI-tools oraz BEDTools. Pierwsze narzędzie służy do tzw. deduplikacji odczytów (pogrupowania odczytów powstających z tego samego produktu PCR) względem sekwencji UMI przypisanej do każdej próby. Dzięki sekwencji UMI identyczne kopie, powstające w wyniku odrębnych zdarzeń wiązania białka z RNA, mogą być odróżnione od tych powstających w wyniku amplifikacji PCR pojedynczego zdarzenia wiązania. Następnie wykorzystano narzędzie BEDTools, które pozwala na identyfikację miejsc wiązania białka do RNA, tzw. crosslinku. Narzędzie to generuje dwa pliki: (i) plik o rozszerzeniu *.bed*, w którym znajdziemy informację o miejscach crosslinku, z rozdzielczością do pojedynczego nukleotydu, a wynik analizy informuje nas o liczbie wiązań w danej pozycji nukleotydowej oraz (ii) plik o rozszerzeniu *.bedgraph*, w którym znajduje się informacja o wiązaniach zidentyfikowanych w obrębie nici dodatniej oraz ujemnej genomu stanowiącego referencję. Na **Rysunku 19** przedstawiono liczbę

zidentyfikowanych miejsc wiązań oraz liczbę zidentyfikowanych miejsc crosslinku dla każdej badanej próby. Liczba zidentyfikowanych miejsc crosslinku jest różna dla każdej próby, co wynika z różnej liczby odczytów przypisanych zgodnie z sekwencją UMI. Jak można zauważyć na wykresie, różnica pomiędzy liczbą zdarzeń wiązania a liczbą miejsc crosslinku dla każdej badanej próby jest niewielka. Stosunek liczby zdarzeń wiązania do liczby miejsc crosslinku waha się w przedziale 1.02-1.29 (**Rysunek 20**).



**Rysunek 19. Liczba zdarzeń wiązań oraz miejsc crosslinku dla każdej badanej próby.** Liczba zidentyfikowanych zdarzeń wiązań oznaczona jest kolorem niebieskim (total\_xlinks) a liczba miejsc crosslinku kolorem szarym (total\_xlinksites). Wykres został wygenerowany przez narzędzie MultiQC.



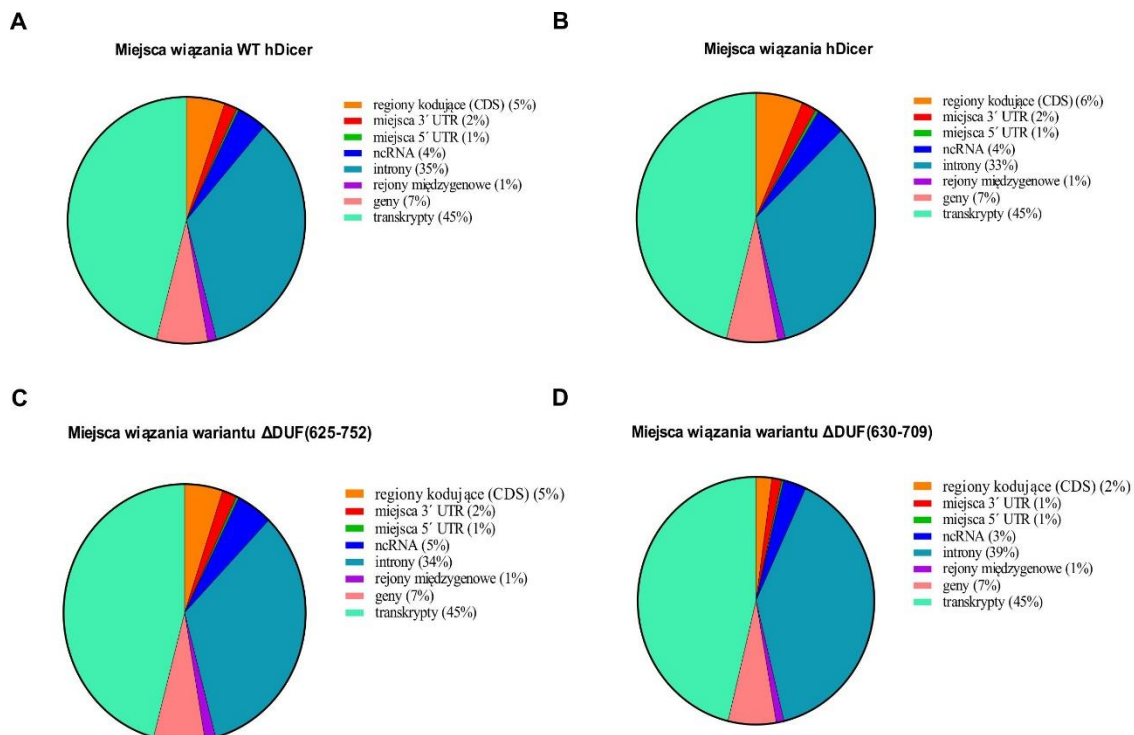
**Rysunek 20. Stosunek liczby zidentyfikowanych zdarzeń wiązań do liczby miejsc crosslinku dla każdej badanej próby.** Wykres został wygenerowany przez narzędzie MultiQC.

Na podstawie zebranego raportu o jakości uzyskanych danych pochodzących z sekwencjonowania NGS stwierdzono, że dane spełniają wszystkie wymagane kryteria i mogą zostać poddane dalszym analizom.

#### **4.2.4 Wstępna analiza porównawcza pul RNA wiązanych przez warianty $\Delta$ DUF(625-752), $\Delta$ DUF(630-709) oraz hDicer**

Jak wspomniano powyżej, w badaniach komórkowej sieci oddziaływań pomiędzy RNA a hDicer oraz RNA i wariantami delecyjnymi  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709), wykorzystano następujące układy – linie komórkowe: (i) komórki HEK 293T – linia typu dzikiego (WT hDicer), (ii) komórki HEK 293T NoDice transfekowane plazmidem kodującym pełnej długości białko hDicer (tzw. *rescue control*, hDicer), (iii) komórki HEK 293T NoDice transfekowane plazmidem kodującym  $\Delta$ DUF(625-752) oraz (iv) HEK 293T NoDice transfekowane plazmidem kodującym  $\Delta$ DUF(630-709).

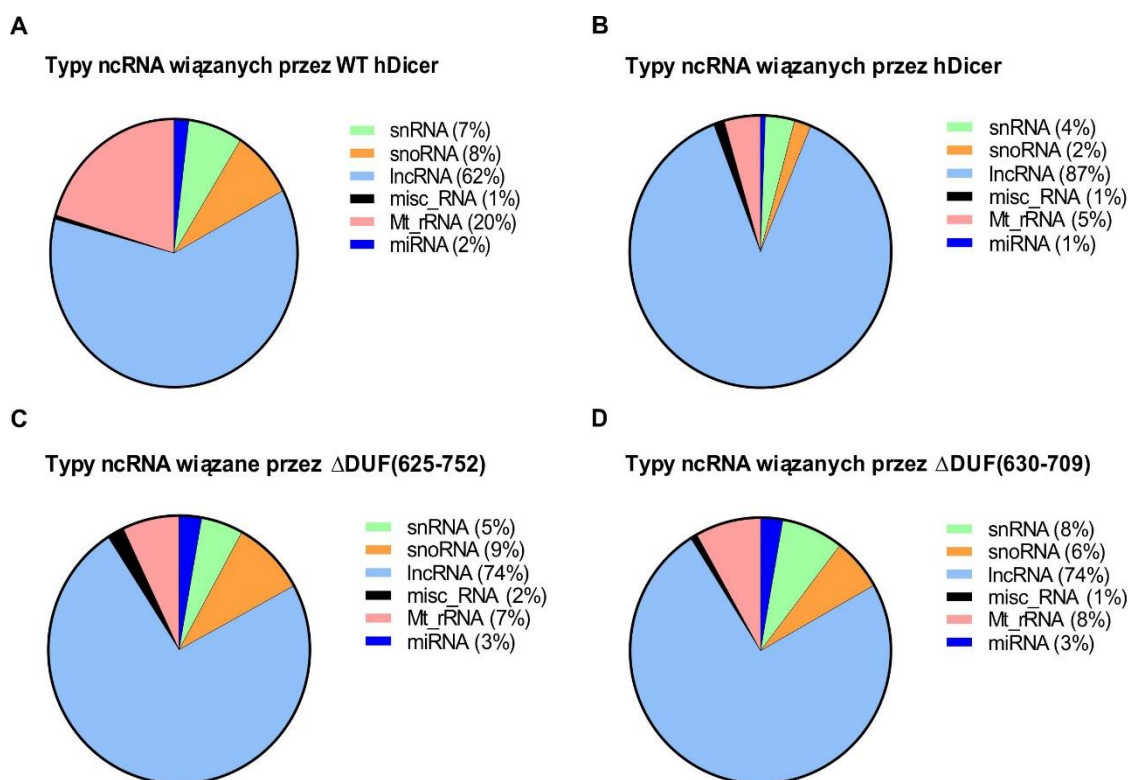
Analizę uzyskanych danych rozpoczęto od identyfikacji rejonów, z którymi wiązały się badane białka: WT hDicer, hDicer,  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709). Wykorzystując protokół *nf-core/clipseq* [97], wyznaczono średnią liczbę zidentyfikowanych miejsc wiązania, która wynosiła odpowiednio: 1 492 612 dla WT hDicer, 1 571 380 dla hDicer, 862 941 dla  $\Delta$ DUF(625-752) oraz 16 175 056 dla  $\Delta$ DUF(630-709). Ponieważ odczyty były mapowane do genomu człowieka (wersja hg 38), zatem większość miejsc wiązania się badanych białek została przypisana do szerokiego zakresu miejsc w obrębie genomu, np. fragmentów genomu ulegających transkrypcji do mRNA, czy też niekodujących RNA (ncRNA), ale także do miejsc egzonowych, czy intronowych wielu genów (**Rysunek 21**). Warto w tym miejscu podkreślić, że w przypadku wcześniejszych analiz pul RNA wiązanych przez hDicer w komórkach typu HEK 293T, z wykorzystaniem metody PAR-CLIP-seq, również otrzymano podobne rezultaty [79].



**Rysunek 21. Wykresy kołowe obrazujące typy miejsc wiązania dla badanych białek.** W nawiasach przedstawiono wartości procentowe przypisane poszczególnym typom miejsc wiązania, w odniesieniu do sumy wszystkich uzyskanych odczytów dla każdego badanego białka. Odczyty mapowano do ludzkiego genomu (wersja hg38).

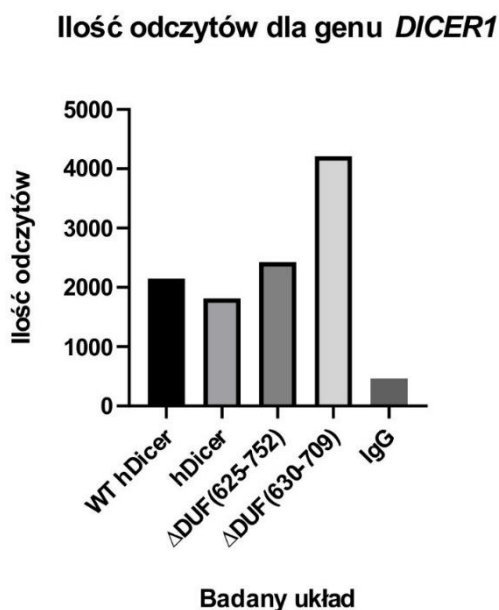
Następnie postanowiono sprawdzić z jakimi typami cząsteczek ncRNA wiązały się poszczególne białka. Jak można zauważyć na **Rysunku 22**, wszystkie badane białka najliczniej oddziaływały z lncRNA, a najmniej wiązań odnotowano dla cząsteczek misc\_RNA, czyli cząsteczek RNA, których nie przypisano do innych, znanych typów ncRNA. Ponownie, otrzymane dane odpowiadają rezultatom zaprezentowanym w publikacji Rybak-Wolf i in. [79], w której wykazano, że hDicer wiąże się licznie z mRNA oraz lncRNA.





**Rysunek 22. Wykresy kołowe obrazujące typy niekodujących RNA wiązanych przez badane białka.** W nawiasach przedstawiono wartości procentowe przypisane poszczególnym typom, ncRNA, w odniesieniu do sumy wszystkich uzyskanych odczytów dla ncRNA, dla każdego badanego białka.

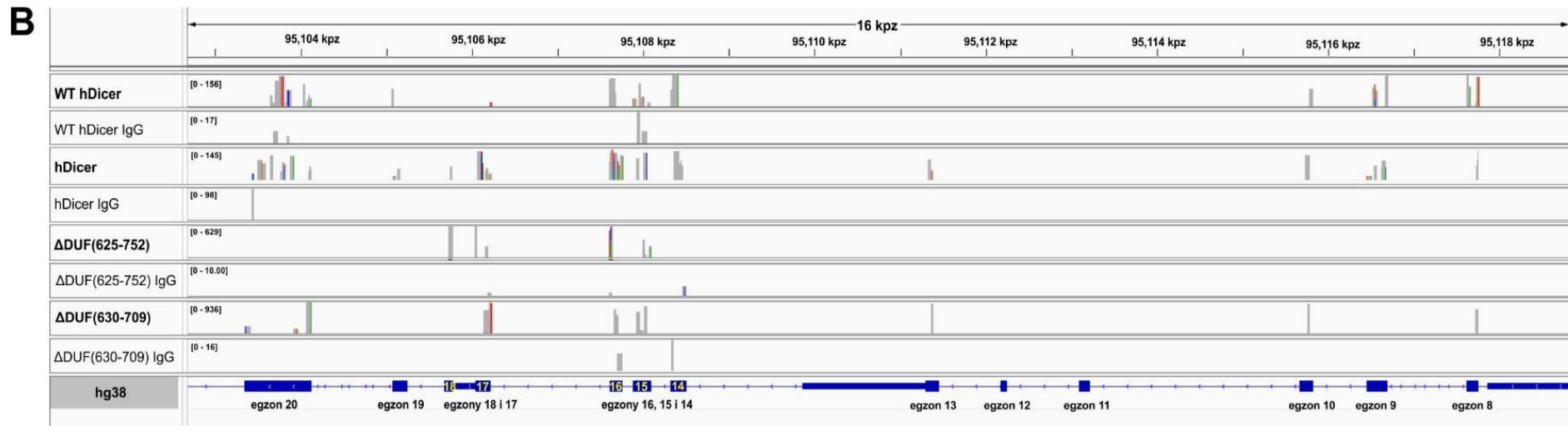
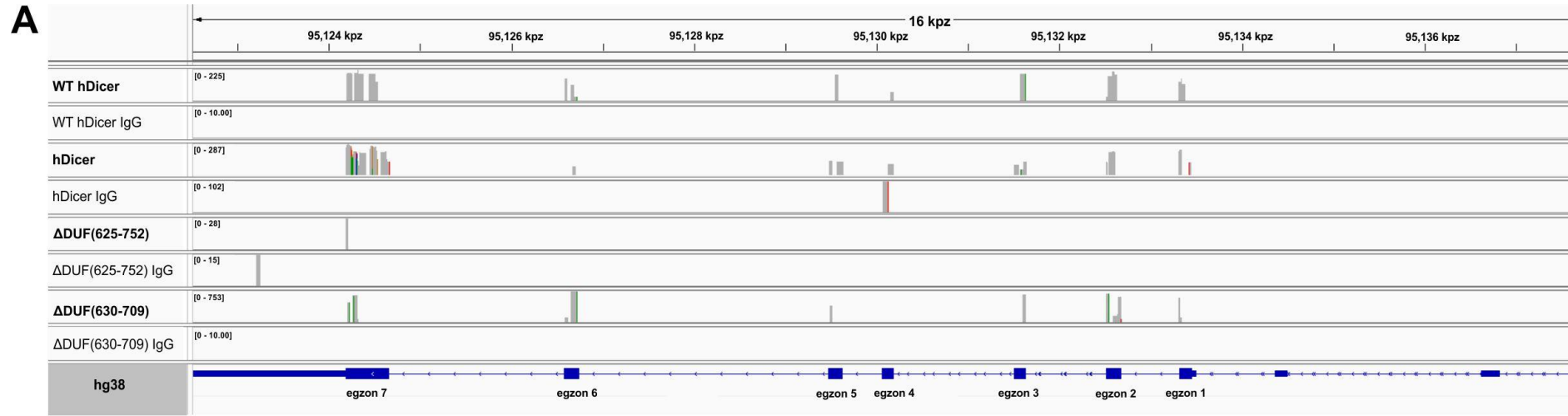
We wspomnianej publikacji odnotowano także wiele miejsc wiązania hDicer w obrębie transkryptu genu *DICER1*. Postanowiono zatem sprawdzić ile odczytów zostało zmapowanych do transkryptu genu *DICER1* w badanych układach (**Rysunek 23**). Jak można zauważyć, we wszystkich badanych układach odnotowano liczne miejsca wiązań w obrębie transkryptu genu *DICER1*: dla WT hDicer było to średnio 2174 odczytów, dla hDicer było to 1814 odczytów, dla  $\Delta$ DUF(625-752) było to 2428 odczytów, a dla  $\Delta$ DUF(630-709) było to 4214 odczytów. Dla kontroli IgG średnia liczba odczytów dla genu *DICER1* wynosiła 433 odczyty. Relatywnie niewielka liczba odczytów przypisanych do kontroli IgG świadczy o specyficzności izolacji kompleksów białko•RNA. W tym miejscu warto podkreślić, że zarówno w komórkach HEK 293T, jak i HEK 293T NoDice, obecny jest transkrypt genu *DICER1*. Jednakże w przypadku komórek HEK 293T NoDice gen *DICER1* posiada delecję w obrębie rejonu kodującego domenę helikazy [19].

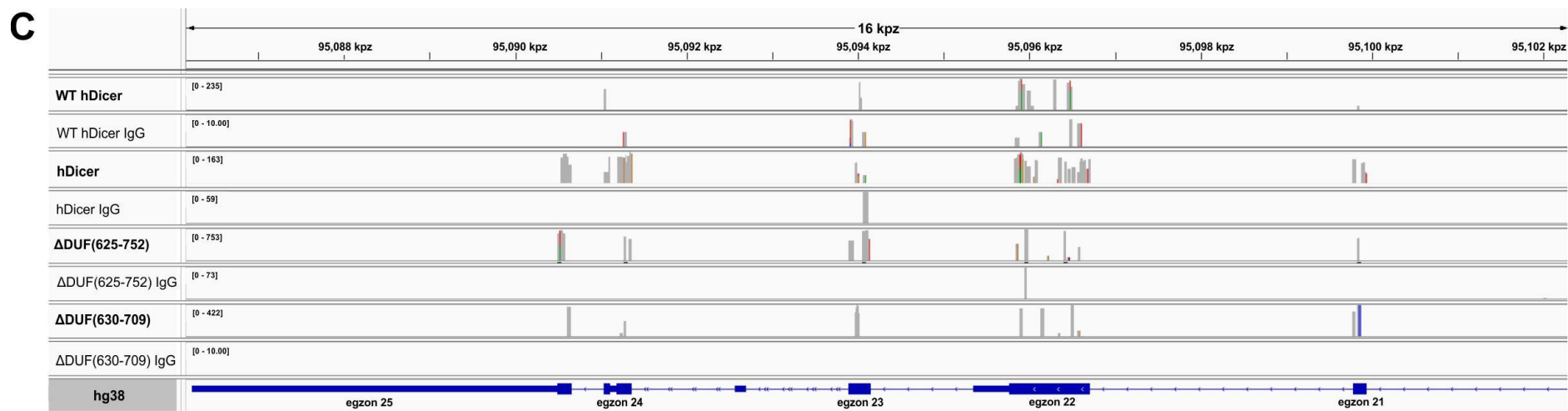


**Rysunek 23.** Wykres przedstawiający liczbę odczytów dla miejsc wiązań zmapowanych do genu *DICER1* w komórkach HEK 293T typu dzikiego oraz NoDice produkujących wskazane białka.

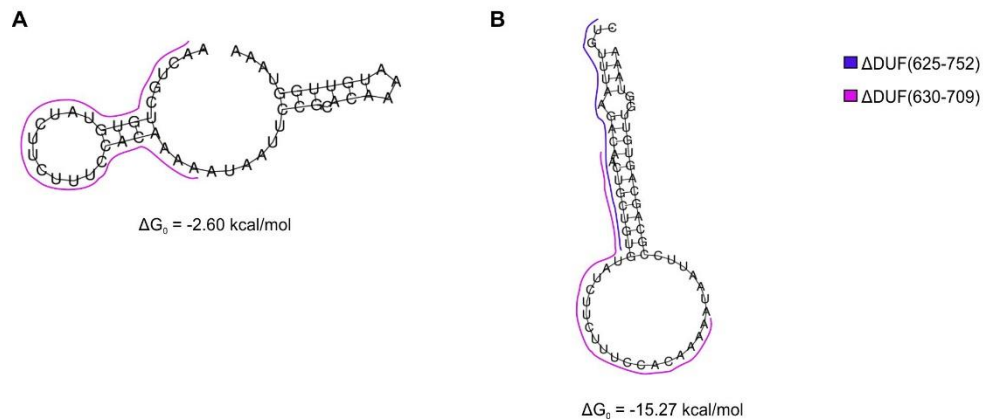
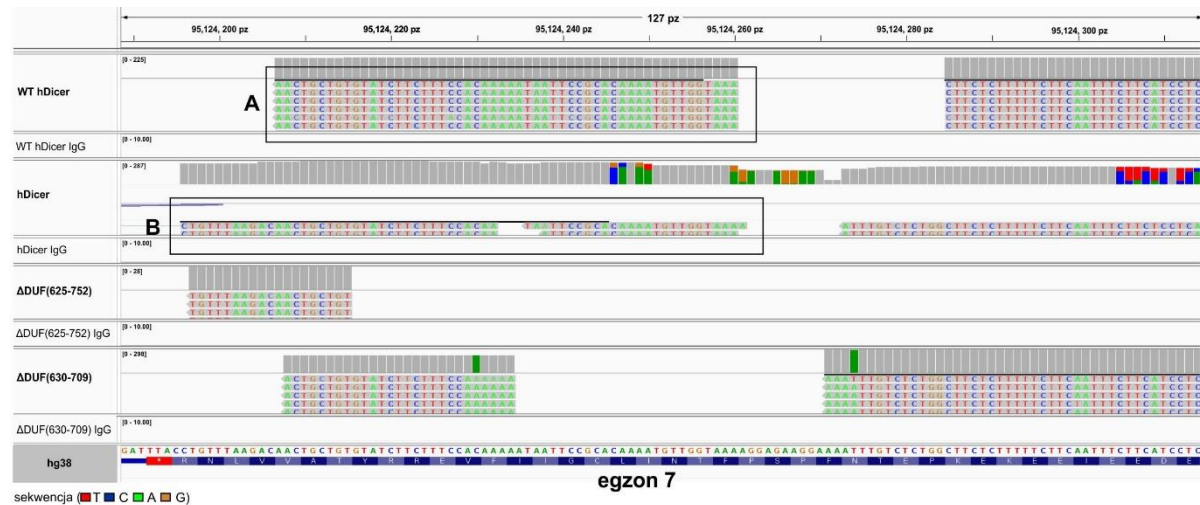
Następnie sprawdzono do jakich rejonów transkryptu *DICER1* wiązały się badane białka (**Rysunek 24**). Przeprowadzona analiza w programie IGV (ang. *Integrative Genomics Viewer*) wykazała, że wszystkie badane białka licznie wiązały się do miejsc mapujących się w obrębie genu *DICER1*. Udało się wskazać rejony, które były wiązane przez wszystkie badane białka (**Rysunek 25**), ale też takie, które były wiązane tylko przez hDicer (**Rysunek 26**). Nie zaobserwowano rejonów w obrębie transkryptu genu *DICER1*, z którymi wiązały się tylko warianty ΔDUF hDicer. W kolejnym etapie przewidziano struktury drugorzędowe zidentyfikowanych sekwencji, w tym celu wykorzystano program RNAfold [98]. Wiązanie hDicer oraz wariantów ΔDUF hDicer występowało w obrębie sekwencji, które potencjalnie mogą przyjmować struktury typu spinki do włosów (**Rysunek 25, 26**). Warto zauważyć, że warianty ΔDUF hDicer wiązały się z krótszymi fragmentami sekwencji, które tworzyły struktury typu spinki do włosów (**Rysunek 25**). W przypadku białka hDicer zaobserwowano wiązanie się z dłuższymi sekwencjami, które mogą przyjmować bardziej złożone struktury składające się z więcej niż jednej struktury typu spinki do włosów (**Rysunek 26**).

Podsumowując, dopiero analizy różnicowe prowadzone na poziomie sekwencji, z rozdzielczością do pojedynczego nukleotydu, wskazały różnice pomiędzy pulami RNA wiązаныmi przez hDicer oraz warianty  $\Delta$ DUF hDicer. W niniejszej pracy przedstawiono wyniki wstępnych analiz różnicowych pul RNA wiązanych przez hDicer i poszczególne warianty  $\Delta$ DUF hDicer. Badania te będą kontynuowane i rozszerzone o szczegółowe analizy miejsc wiązań w obrębie transkryptów innych genów.

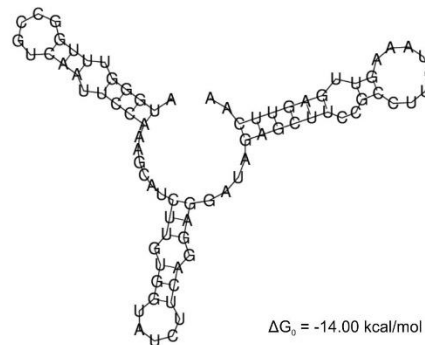
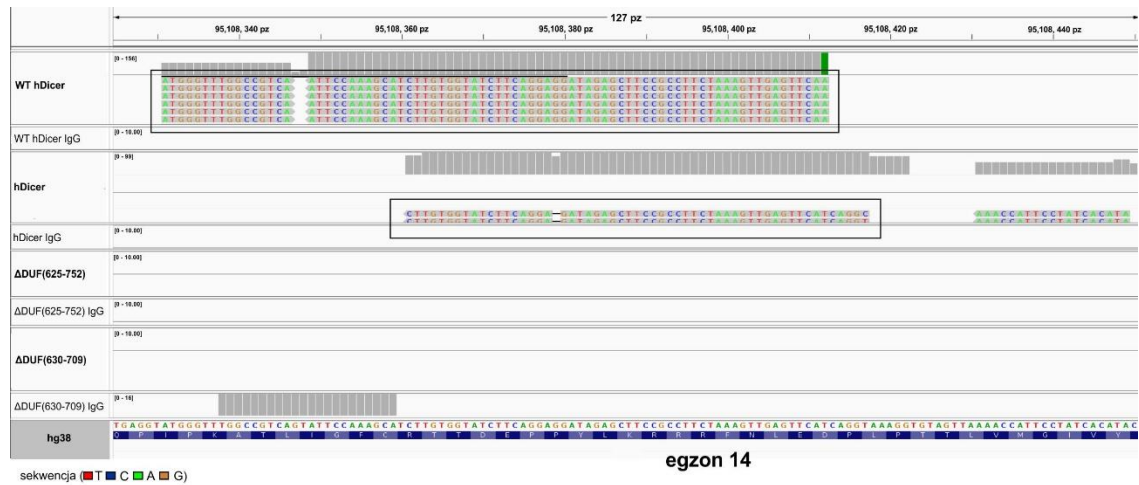




**Rysunek 24. Graficzne przedstawienie odczytów zmapowanych do egzonów genu *DICER1* (NM\_001195573.1) (A-C). Referencję stanowił genom hg38 człowieka. Analiza wykonana w programie IGV. Pokrycie poszczególnych egzonów jest przedstawione w formie skali logarytmicznej.**



**Rysunek 25. Graficzne przedstawienie odczytów zmapowanych do egzonu 7 genu *DICER1* (NM\_001195573.1) oraz przewidywane struktury drugorzędowe rejonów wiązanych przez badane białka . Na panelu **A** oraz **B** przedstawiono miejsca wiązania dla WT hDicer (**A**) oraz hDicer (**B**). Miejsca wiązań wariantów delecyjnych hDicer, w obrębie struktur wiązanych przez hDicer i przedstawionych na rysunku, oznaczono wskazanymi kolorami. Struktury zostały wygenerowane w programie RNAfold.**



Rysunek 26. Graficzne przedstawienie odczytów zmapowanych do egzonu 14 genu *DICER1* (NM\_001195573.1) oraz przewidywane struktury drugorzędowe rejonów wiązanych przez WT hDicer oraz hDicer. Struktura zostały wygenerowane w programie RNAfold.

### **4.3 Badania oddziaływań hDicer z cząsteczkami RNA i DNA przyjmującymi strukturę G-kwadrupleksów**

Badania przeprowadzone z wykorzystaniem metody irCLIP-seq pozwoliły na zidentyfikowanie pul RNA wiązanych przez hDicer w komórkach HEK 293T. Wśród tych RNA znaleziono wiele cząsteczek bogatych w ciągi guanozynowe (G). Na początku sprawdzono, czy wśród licznych komórkowych RNA bogatych w ciągi G i wiązanych przez białko hDicer, występują takie sekwencje, które mogą przyjmować struktury G-kwadrupleksu. Dzięki wykorzystaniu programu Quadron [99] udało się zidentyfikować liczne sekwencje, które potencjalnie mogą tworzyć takie struktury (**Tabela 24**). Co ciekawe, wyniki wcześniejszych badań, zaprezentowane w rozprawie doktorskiej dr N. Koralewskiej, ujawniły, iż hDicer wydajnie wiąże cząsteczki RNA przyjmujące struktury G-kwadrupleksu [A]. Biorąc pod uwagę zgromadzone dane postanowiono przyjrzeć się bliżej problemowi oddziaływań pomiędzy hDicer a cząsteczkami RNA i DNA przyjmującymi strukturę G-kwadrupleksu.



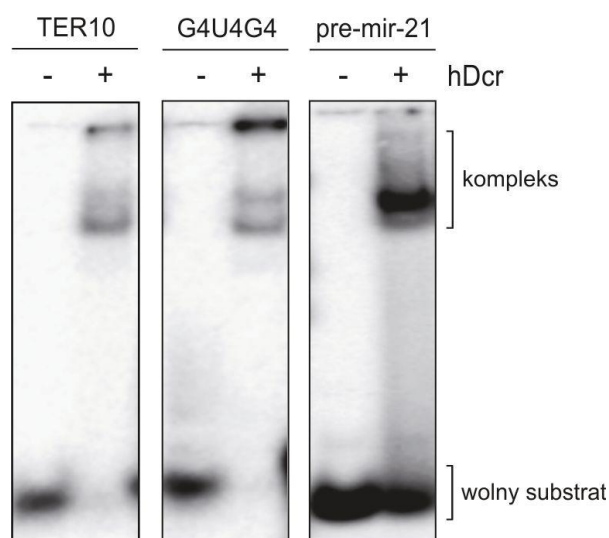
Nazwa genu	ID transkryptu	Koordynaty	Nr. chromosomu/ (orientacja nici)	Sekwencja	Lokalizacja w transkrypcie
<i>HES6</i>	ENST00000409002.3	238239932-238240025	2 (-)	GGGCTGCTG CTGGGCGGC CGCGGGGCA GCGGAGGG	5'-UTR
<i>FEM1C</i>	ENST00000274457.3	115544690-115544741	5 (-)	GGGCGGGAG GCAGCTCGG GTGGCGGGG	5'-UTR
<i>ZNF703</i>	ENST00000331569.4	37699261-37699413	8 (+)	GGGGGTGGA GGGGGCCAC CTGGGTGGC AGGGG	3'-UTR, miejsce wiązania docelowych miRNA
<i>U2AF2</i>	ENST00000308924.4	55674075-55674122	19 (+)	GGGGGAGGG TGGGGGCAG GGCTGGCTG GGGG	3'-UTR, miejsce wiązania docelowych miRNA
<i>E2F1</i>	ENST00000343380.5	33676081-33676154	20 (-)	GGGGAATCA GCAAAGGGG AGGAGGGGT GTGGGG	3'-UTR, miejsce wiązania docelowych miRNA

**Tabela 24. Przykładowe fragmenty transkryptów bogate w ciągi G, potencjalnie przyjmujące struktury G-kwadrupleksu i wiązane przez hDicer (na podstawie wyników analiz danych pochodzących z sekwencjonowania NGS bibliotek cDNA uzyskanych z izolacji kompleksów hDicer•RNA metodą irCLIP). Analizę potencjału RNA do przyjmowania struktur G-kwadrupleksu wykonano z użyciem narzędzia Quadron [99]. Na czerwono zaznaczono ciągi reszt guanozyny.**

W badaniach oddziaływań pomiędzy hDicer i RNA przyjmującymi struktury G-kwadrupleksu (RNA-G4) wykorzystywano preparat hDicer (hDcr) otrzymywany w komórkach HEK 293T NoDice (Metody 3.2.15) (Rysunek 7). Początkowo, potwierdzono zdolność hDcr do wiązania RNA-G4. W tym celu wykorzystano metodę spowolnionej migracji w żelach poliakrylamidowych (EMSA). W mieszaninach

reakcyjnych obecne były: hDcr (250 nM) i znakowane na końcach 5', z użyciem [ $\gamma^{32}$ P]-ATP, cząsteczki RNA-G4 (5 nM): TER10 [r(GGG UUA GGGU)] – fragment telomerowego RNA (tzw. TERRA) człowieka oraz G4U4G4 [r(GGG GUU UUG GGG)] – fragment TERRA orzęska *Oxytricha nova*. Wybrane RNA-G4 zostały wcześniej scharakteryzowane pod kątem przyjmowania struktur G4 [100, 101]. Szczegółowe informacje na temat oligomerów przedstawiono w rozdziale: Suplementy w **Tabeli S1**. Kontrolę ujemną stanowiła reakcja bez dodatku białka hDcr (K-). Pozytywną kontrolę reakcji wiązania stanowiła reakcja, w której wykorzystano pre-miRNA (pre-mir-21) (Metody 3.2.18).

Uzyskane wyniki potwierdziły, że hDicer wydajnie wiąże cząsteczki TER10 oraz G4U4G4 (**Rysunek 27**).



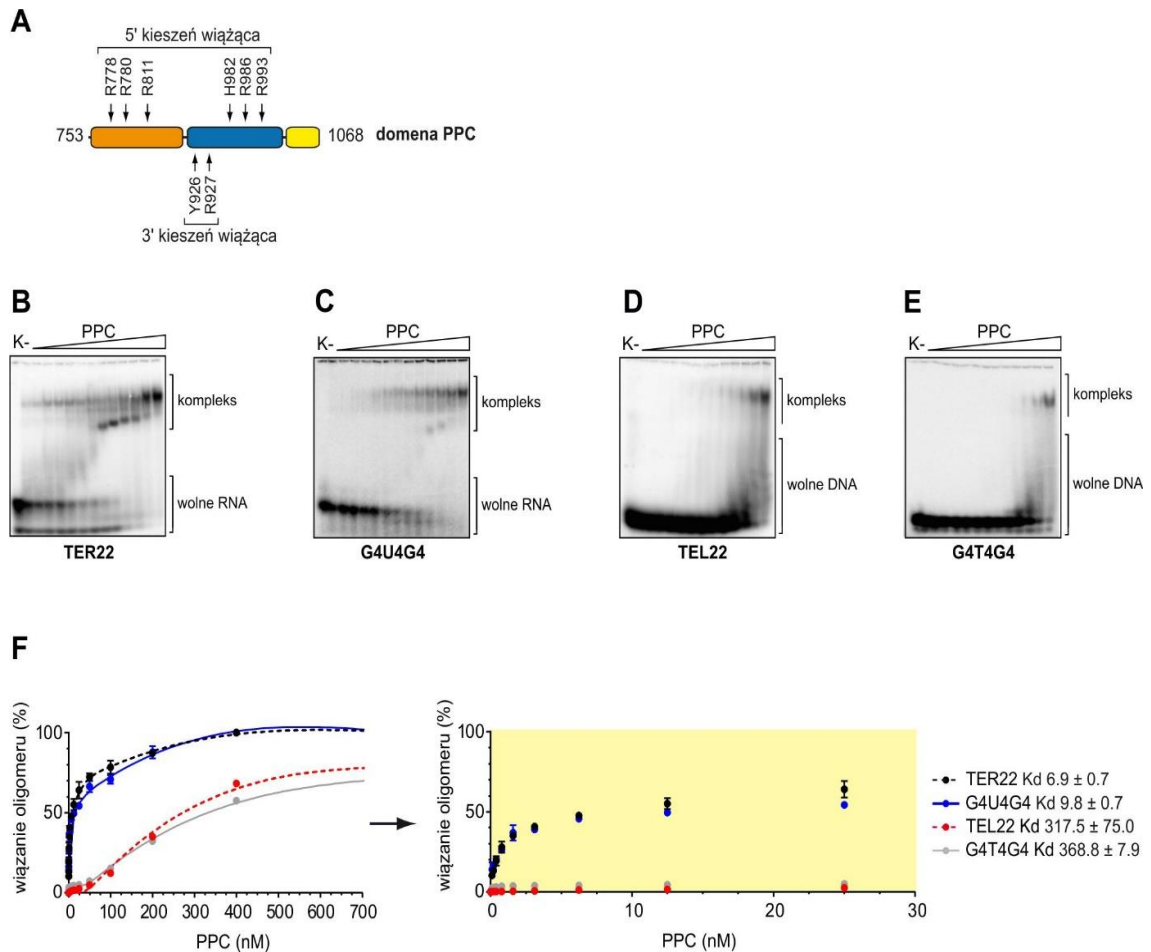
**Rysunek 27. Badanie oddziaływań pomiędzy cząsteczkami RNA przyjmującymi strukturę G-kwadrupleksu a hDcr.** Do badania wykorzystano wyznakowane, za pomocą [ $\gamma^{32}$ P]-ATP na końcu 5', cząsteczki RNA-G4 (TER10, G4U4G4) lub pre-mir-21 oraz preparat hDcr (250 nM). Znaki: (+) oraz (-) odnoszą się do obecności hDcr w mieszaninie reakcyjnej. Próby rozdzielano w 5% PAA. Rysunek przygotowano na podstawie publikacji Koralewska&Szczepanska i in. [102].

Kanoniczne substraty Dicer, tj. cząsteczki pre-miRNA i pre-siRNA, kotwiczone są w obrębie kieszeni wiążących (5' i 3') obecnych w kasecie PPC Dicer, czyli w rejonie obejmującym domeny platformy, PAZ oraz helisy łączącej (**Rysunek 28A**) [38, 90]. Biorąc pod uwagę te informacje, postanowiono sprawdzić czy kasetka PPC może również wiązać cząsteczki RNA przyjmujące strukturę G-kwadrupleksu. W badaniach

wykorzystano znakowane na końcach 5', z użyciem [ $\gamma^{32}\text{P}$ ]-ATP, cząsteczki TER22 (r[AGGG(UUAGGG)<sub>3</sub>] - dłuższa wersja cząsteczki TER10 oraz G4U4G4.

Ze względu na fakt, że sekwencje zawierające powtórzenia telomerowe są obecne nie tylko w TERRA, ale także w telomerowym DNA na końcach chromosomów, w badaniach wykorzystano także odpowiedniki DNA cząsteczek TER22 oraz G4U4G4, tj. cząsteczki: TEL22 (d[AGGG(TTAGGG)<sub>3</sub>]) oraz G4T4G4 (d(GGGGTTTTGGGG)) (Metody 3.1.1, **Tabela 1**). Wybrane DNA-G4 zostały wcześniej scharakteryzowane pod kątem przyjmowania struktur G4 [103, 104]. W reakcji wiązania zastosowano wzrastające stężenie PPC (0.1, 0.2, 0.4, 0.8, 1.6, 3.125, 6.25, 12.5, 25, 50, 100, 200, 400 nM) oraz 5 nM wskazanego 5'-[ $\gamma^{32}\text{P}$ ]-ATP-oligomeru. Reakcją kontrolną ujemną (K-) była reakcja bez białka (Metody 3.2.18).

Otrzymane rezultaty wykazały, że domena PPC wiąże wszystkie cztery testowane substraty (**Rysunek 28B-E**). W przypadku cząsteczek RNA-G4 można było zaobserwować tworzenie się stabilnych kompleksów już przy najniższym stężeniu białka PPC (ok. 0.2 nM) (**Rysunek 28B, C**). Cząsteczki DNA-G4 były wiązane z dużo niższą wydajnością (**Rysunek 28D, E**), ponieważ wiązanie następowało dopiero przy wyższym stężeniu PPC (ok. 50 nM).



**Rysunek 28. Badanie oddziaływań pomiędzy domeną PPC hDicer a cząsteczkami RNA oraz DNA przyjmującymi strukturę G-kwadupleksu. (A)** Schematyczne przedstawienie kasety PPC z zaznaczonymi resztami aminokwasowymi w obrębie kieszeni 5' i 3' kieszeni wiążącej, które oddziałują z substratami RNA. **(B, C)** Testy EMSA dla substratów RNA-G4 wyznakowanych za pomocą  $[\gamma^{32}\text{P}]$ -ATP na końcu 5': TER22 **(B)**, G4U4G4 **(C)** oraz preparatu PPC. **(D, E)** Testy EMSA dla substratów DNA-G4 wyznakowanych za pomocą  $[\gamma^{32}\text{P}]$  na końcu 5': TEL22 **(D)**, G4T4G4 **(E)** oraz preparatu PPC. (K-) to próby kontrolne bez białka. Trójkątami oznaczono rosnące stężenie preparatu PPC (0.1, 0.2, 0.4, 0.8, 1.6, 3.125, 6.25, 12.5, 25, 50, 100, 200, 400 nM). Próby rozdzielano w 10% PAA. **(F)** Graficzne przedstawienie ilościowego powinowactwa pomiędzy wykorzystanymi substratami a preparatem PPC. Rysunek przygotowano na podstawie publikacji Koralewska&Szczepanska i in. [102].

Na podstawie wyników rozdzielania elektroforetycznego kompleksów PPC•RNA-G4 oraz PPC•DNA-G4 wyznaczono równowagowe stałe dysocjacji ( $K_d$ ). W tym celu przeprowadzono analizę densytometryczną powstałych kompleksów oraz wolnego substratu, z wykorzystaniem programu Multi Gauge (Fujifilm). Wartości  $K_d$  wyliczono zgodnie ze wzorem:

$$f = \frac{A \times P_0}{P_0 + K_d}$$

f – frakcja związanego substratu

$P_0$  – stężenie molowe białka dla danej próby ( $\mu\text{M}$ )

A – maksymalne wiązanie RNA (%)

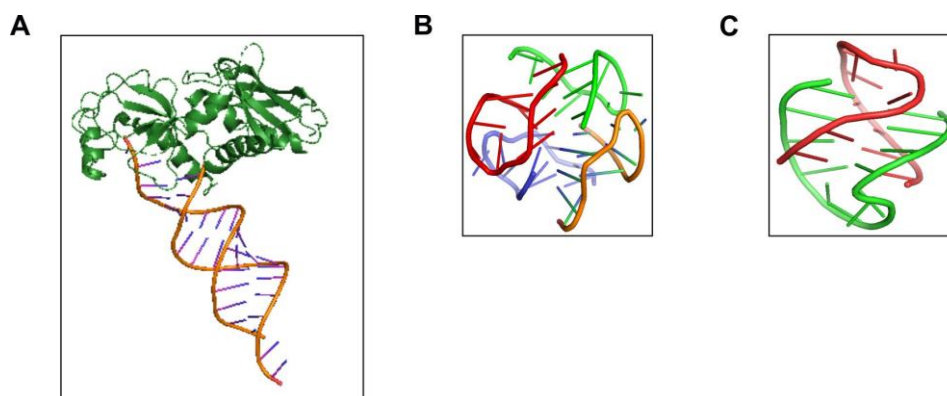
$K_d$  wyliczono dla wartości  $f = 0,5$ , czyli dla układu, w którym połowa substratu była związana z białkiem.  $P_0$  - stanowiło stężenie białka, dla którego osiągnięto 50% wiązania substratu. Dla parametru A, przyjęto 95% dla kompleksów PPC•RNA-G4 hDicer i 65% dla kompleksu PPC•DNA-G4 hDicer.

Otrzymane eksperymentalnie wartości  $K_d$  dla kompleksów PPC•RNA-G4 wyniosły:  $\sim 7$  nM dla kompleksu PPC•TER22 oraz  $\sim 10$  nM dla kompleksu PPC•G4U4G4 (**Rysunek 28F**). Dla porównania, określono też wartości stałych wiązania dla kompleksów: PPC•pre-mir-21 oraz PPC•RNA21, które wynosiły odpowiednio:  $\sim 205$  nM oraz  $\sim 46$  nM [102]. Na podstawie otrzymanych wyników można było wnioskować, że izolowana kasetka PPC hDicer wiąże cząsteczki RNA-G4 z wyższą wydajnością niż substraty pre-miRNA, czy jednoniciowe RNA o długości miRNA (21 nt). Dla kompleksów PPC•DNA-G4 wartości  $K_d$  wynosiły:  $\sim 318$  nM dla kompleksu PPC•TEL22 oraz  $\sim 369$  nM dla kompleksu PPC•G4T4G4 (**Rysunek 28F**). Uzyskane dane wykazały, że G-kwadrupleksy DNA mają znacznie niższe powinowactwo do kasety PPC niż ich odpowiedniki RNA. Dodatkowo zauważono, że cząsteczki TER22 i TEL22 były wiązane nieco wydajniej przez preparat PPC, w porównaniu do cząsteczek G4U4G4 oraz G4T4G4, tj. cząsteczka TER22 była wiązana  $\sim 1,4$  razy wydajniej niż cząsteczka G4U4G4, a cząsteczka TEL22 była wiązana  $\sim 1,2$  razy wydajniej niż cząsteczka G4T4G4.

Podsumowując, uzyskane wyniki sugerują, że w przypadku cząsteczek przyjmujących strukturę G-kwadrupleksu, zarówno rodzaj cząsteczki (RNA lub DNA), jak i jej długość, wpływają na powinowactwo danej cząsteczki RNA-G4 lub DNA-G4 do kasety PPC hDicer.

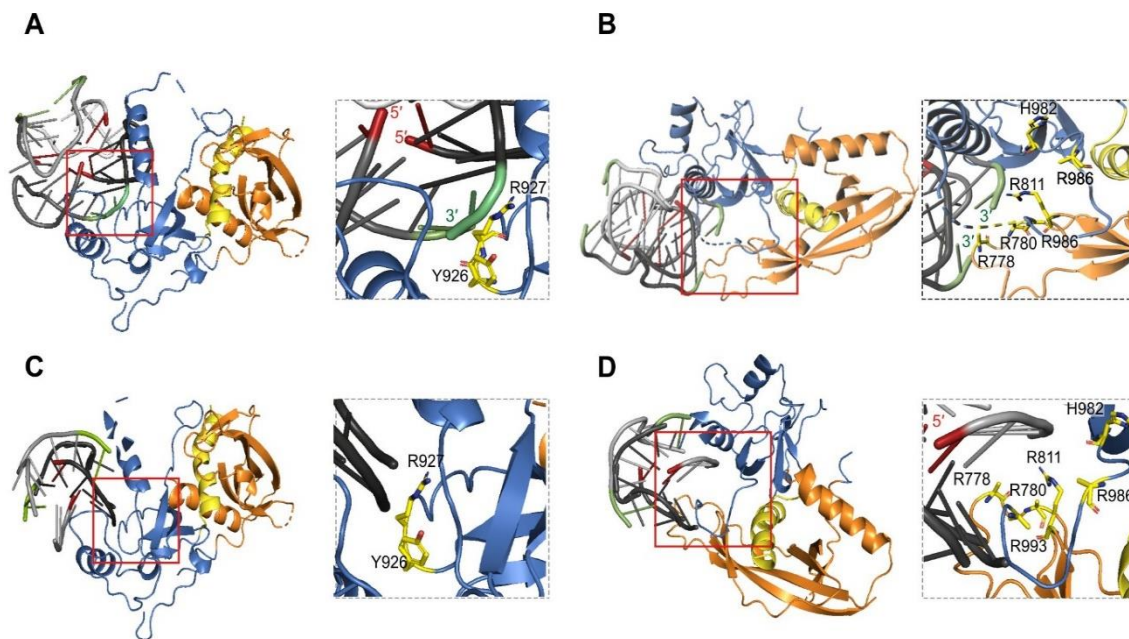
Biorąc pod uwagę fakt, że kasetka PPC hDicer posiada dwie kieszenie wiążące kanoniczne substraty, tj. kieszeń 5' i kieszeń 3', postanowiono sprawdzić czy obie te kieszenie biorą udział w wiązaniu cząsteczek RNA-G4 i DNA-G4. W tym celu wykonano modelowanie molekularne, a do analiz wykorzystano dane strukturalne dotyczące kasety PPC hDicer (kod PDB: 4NGF) [38] oraz dane strukturalne dotyczące RNA TER10, który

przyjmuje strukturę G-kwadrupleksu o typie dimeru bimolekularnego ( kod PDB: 2M18) [101], a także DNA G4T4G4, który przyjmuje strukturę G-kwadrupleksu o typie bimolekularnym (kod PDB: 1JPQ) [104]. Struktury PPC hDicer w kompleksie z 17 nt siRNA, TER10 oraz G4T4G4 przedstawiono na **Rysunku 29**.



**Rysunek 29. Struktury domeny PPC hDicer oraz cząsteczek G-kwadrupleksów wykorzystane w modelowaniu molekularnym. (A)** Struktura kasyety PPC hDicer w kompleksie z 17 nt siRNA (kod PDB: 4NGF). **(B)** Struktura cząsteczki G-kwadrupleksu RNA TER10 (kod PDB: 2M18). **(C)** Struktura cząsteczki G-kwadrupleksu DNA G4T4G4 (kod PDB: 1JPQ). Struktury zobrazowane w programie PyMOL.

W wyniku przeprowadzonego modelowania molekularnego wyłoniono osiem najlepszych modeli dla kompleksu PPC•TER10 oraz sześć najlepszych modeli dla kompleksu PPC•G4T4G4. W siedmiu z ośmiu modeli wybranych dla kompleksu PPC•TER10, cząsteczka TER10 znajdowała się w obrębie kieszeni wiążącej 3' kasyety PPC. We wszystkich siedmiu modelach koniec 3' jednej z czterech nici TER10 tworzących G-kwadrupleks, był zakotwiczony w kieszeni wiążącej 3' (**Rysunek 30A**). Tylko w jednym z ośmiu modeli koniec 3' cząsteczki TER10 znajdował się w pobliżu reszt aminokwasowych kieszeni wiążącej 5' (**Rysunek 30B**). Natomiast w przypadku modeli kompleksu PPC•G4T4G4, trzy z nich wykazały, że cząsteczka G4T4G4 była ulokowana w obrębie kieszeni wiążącej 3' kasyety PPC (**Rysunek 30C**), a pozostałe trzy modele wskazywały na wiązanie się cząsteczki w obrębie kieszeni wiążącej 5' (**Rysunek 30D**). Analizy modelowania molekularnego wykonano w ramach współpracy z dr Chandran Nithin oraz prof. Januszem Bujnickim z Laboratorium Bioinformatyki i Inżynierii Białka w Międzynarodowym Instytucie Biologii Molekularnej i Komórkowej w Warszawie.

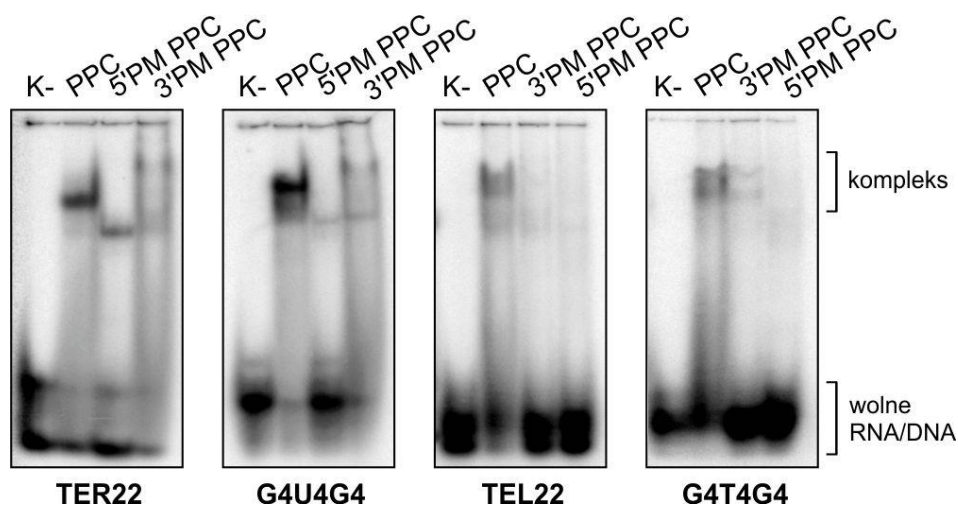


**Rysunek 30. Obrazy przedstawiające modele molekularne kompleksów RNA/DNA G-kwadrupleksów z kasetą PPC hDicer. (A, B) Modele 3D cząsteczki TER10 związanej w kieszeni wiążącej 3' (A) oraz kieszeni wiążącej 5' (B) kasety PPC hDicer. (C, D) Modele 3D cząsteczki G4T4G4 związanej w kieszeni wiążącej 3' (C) oraz kieszeni wiążącej 5' (D) kasety PPC hDicer. Końce 3' oraz 5' oligomeru oznaczono odpowiednio kolorem zielonym i czerwonym. Domena Platformy oznaczona jest kolorem pomarańczowym, domena PAZ kolorem niebieskim, a helisa łącząca kolorem żółtym. Rysunek przygotowany na podstawie publikacji Koralewska&Szczepanska i in. [102].**

W celu weryfikacji uzyskanych modeli, przeprowadzono reakcje wiązania z wykorzystaniem substratów: TER22, G4U4G4, TEL22, G4T4G4 oraz dwóch wariantów kasety PPC hDicer (Metody 3.2.24). Jeden z wariantów zawierał dwie substytucje aminokwasowe: Y926F oraz R927A w kieszeni wiążącej 3' (wariant 3'PM PPC), a drugi wariant zawierał pięć substytucji: R778A, R780A, R811A, H982A, R986A oraz R993A w kieszeni wiążącej 5' (wariant 5'PM PPC). Oba warianty hDicer z substytucjami w obrębie kasety PPC, zostały otrzymane i scharakteryzowane przez grupę prof. N. Kim [90]. Z przeprowadzonych analiz wynikało, że reszty aminokwasowe: R778, R780, R811, H982, R986, R993 (kieszeń wiążąca 5') oraz Y926 i R927 (kieszeń wiążąca 3'), są kluczowe dla oddziaływań kasety PPC z substratami RNA [38]. Zdolność wiązania RNA-G4 i DNA-G4 przez 5'PM PPC oraz 3'PM PPC badano z wykorzystaniem metody EMSA. W reakcjach wykorzystano 5 nM substratu (5'-[ $\gamma^{32}$ P]-ATP-TER22 lub 5'-[ $\gamma^{32}$ P]-ATP-G4U4G4 oraz 5'-[ $\gamma^{32}$ P]-ATP-TEL22 lub 5'-[ $\gamma^{32}$ P]-ATP-G4T4G4) oraz 500 nM białka PPC, 5'PM PPC lub 3'PM PPC. Kontrolę negatywną (K-) stanowiła reakcja bez białka, a kontrolę pozytywną reakcja z PPC typu dzikiego (Metody 3.2.18).



Wyniki przeprowadzonych eksperymentów wykazały, że oba warianty PPC (5'PM PPC oraz 3'PM PPC) wiązały RNA-G4 oraz DNA-G4 ze zdecydowanie niższą wydajnością niż PPC typu dzikiego (**Rysunek 31**). Ponadto zaobserwowano, że wprowadzone substytucje miały większy wpływ na wiązanie się PPC z TEL22 i G4T4G4, niż z cząsteczkami TER22 i G4U4G4. Warto jednakże przypomnieć, że PPC wykazała mniejsze powinowactwo do TEL22 i G4T4G4, niż do TER22 i G4U4G4 (**Rysunek 28B-E**), zatem prawdopodobnie mutacje obecne w kieszeniach wiążących PPC tylko proporcjonalnie zmniejszyły wydajność oddziaływań białko-kwas nukleinowy. Wydajniejsze wiązanie PPC z cząsteczkami RNA-G4, w porównaniu do wiązania DNA-G4, może także wynikać z obecności grupy 2'OH w rybozach RNA-G4, które potencjalnie mogą być zaangażowane w tworzenie wiązań wodorowych z resztami aminokwasowymi obecnymi w kieszeniach wiążących kasy PPC.



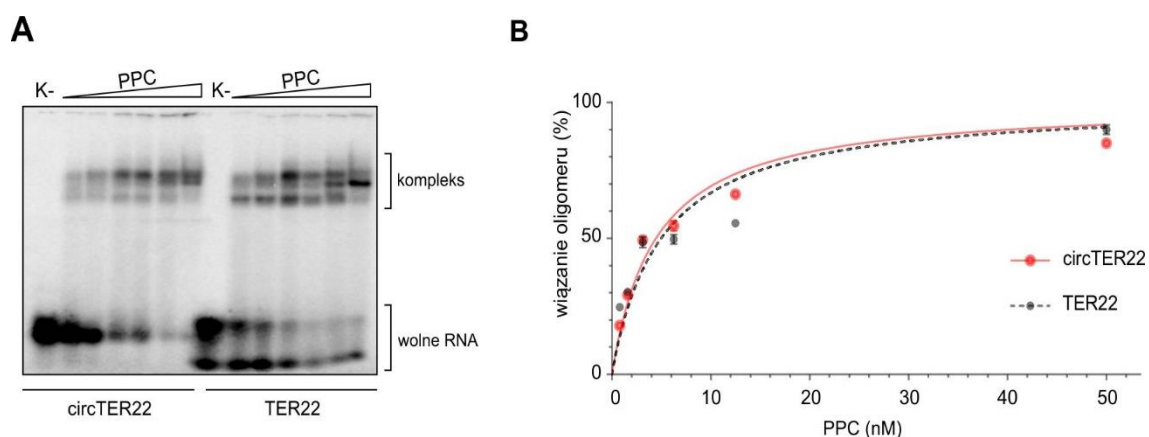
**Rysunek 31. Badanie oddziaływań pomiędzy cząsteczkami RNA oraz DNA przyjmującymi strukturę G-kwadrupeksu a wariantami kasy PPC hDicer.** W reakcjach wykorzystano 5 nM RNA-G4 (TER22, G4U4G4) lub DNA-G4 (TEL22, G4T4G4) wyznakowanych za pomocą [ $\gamma^{32}\text{P}$ ]-ATP na końcu 5' oraz 500 nM preparatu PPC, 5'PM PPC lub 3'PM PPC. (K-) to kontrola reakcji bez białka. Próby rozdzielano w 10% PAA. Rysunek przygotowany na podstawie publikacji Koralewska&Szczepanska i in. [102].

Podsumowując, zgromadzone dane potwierdziły, że obie kieszenie kasy PPC są zaangażowane w wiązanie cząsteczek RNA oraz DNA przyjmujących struktury G-kwadrupeksów.

Zaprezentowane powyżej badania prowadzone były z wykorzystaniem cząsteczek RNA-G4 lub DNA-G4 posiadających wolne końce 5' i 3'. G-kwadrupeksy posiadające wolne końce występują na końcach chromosomów oraz w sekwencji TERRA [105].



Większość struktur G-kwadrupleksowych występuje jednakże wewnątrzcząsteczkowo; przykładowo, w obrębie sekwencji promotorowych (G-kwadrupleksy DNA) lub rejonach regulatorowych mRNA (G-kwadrupleksy RNA obecne w rejonach 5' UTR i 3' UTR) [106, 107]. Biorąc pod uwagę tę okoliczność, następnym krokiem było zbadanie, czy kasetka PPC hDicer może wiązać G-kwadrupleksy nieposiadające wolnych końców 5' i 3'. W tym celu przeprowadzono reakcje wiązania, w których wykorzystano cząsteczkę TER22, której końce 5' i 3' zostały połączone ze pomocą ligazy T4 RNA (circTER22) [108] (Metody 3.2.22). W reakcjach wiązania wykorzystano wzrastające stężenie PPC (0.8, 1.6, 3.125, 6.25, 12.5, 50 nM), a ujemną kontrolę reakcji (K-) stanowiła reakcja bez białka. Równoległe prowadzono reakcje wiązania z udziałem cząsteczki TER22 posiadającej wolne końce 5' i 3'. Otrzymane rezultaty wykazały, że preparat PPC wiąże circTER22 i TER22 z podobną wydajnością (**Rysunek 32**). Zgromadzone wyniki sugerują, że kasetka PPC hDicer może wiązać zarówno G-kwadrupleksy wewnątrzcząsteczkowe, jak i te posiadające wolne końce 5' i 3'.

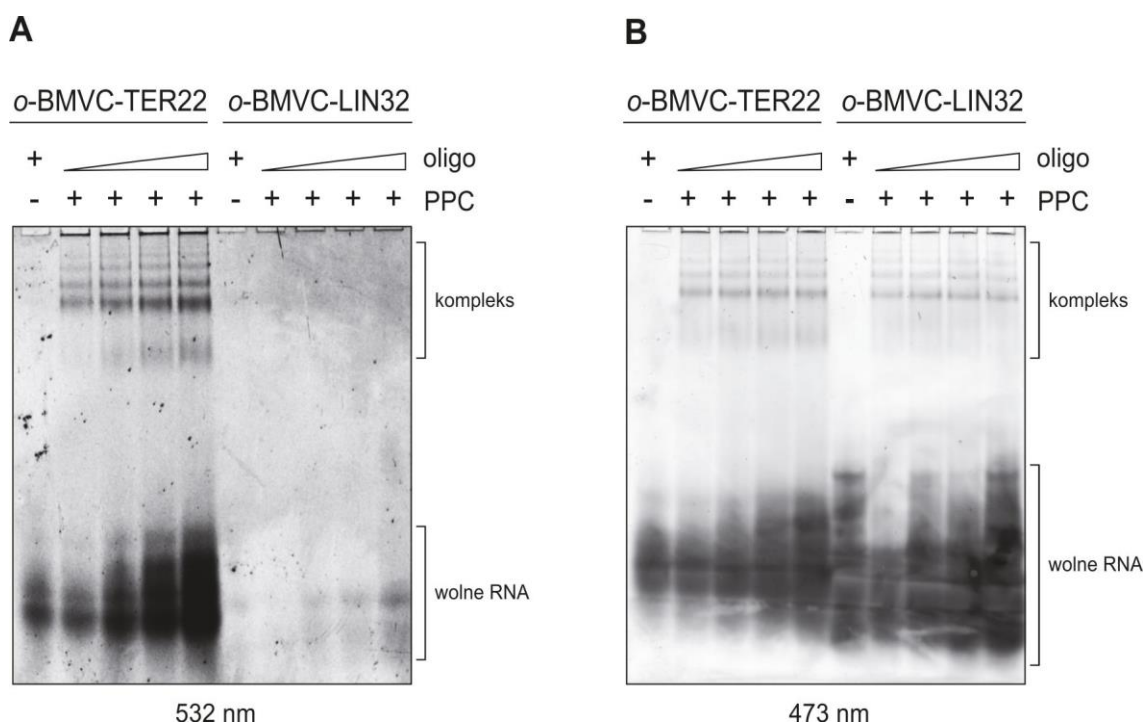


**Rysunek 32. Badanie oddziaływań pomiędzy kasetą PPC hDicer a cząsteczką TER22 z wolnymi końcami (TER22) lub z połączonymi końcami (circTER22) (A).** Trójkąty reprezentują wzrastające stężenie PPC w reakcji (0.8, 1.6, 3.125, 6.25, 12.5, 50 nM). Cząsteczki były wyznakowane za pomocą  $[\gamma^{32}\text{P}]\text{-ATP}$  na końcu 5'. (K-) to reakcja kontrolna bez białka. Próby rozdzielano w 10% PAA. **(B)** Wykres przedstawiający analizę ilościową wiązania pomiędzy preparatem PPC, TER22 lub circTER22. Rysunek przygotowany na podstawie publikacji Koralewska&Szczepanska i in. [102].

Kolejnym etapem było zbadanie, czy G-kwadrupleksy po związaniu przez kasetę PPC hDicer nadal zachowują swoją strukturę. W tym celu przeprowadzono reakcje wiązania z udziałem cząsteczki TER22 z dołączoną kowalencyjnie na jej końcu 5' pochodną 3,6-bis (1-dijodku 3,6-bis (1-metylo-2-winylo-pirydyniowym)) karbazolu (o-BMVC). O-BMVC jest fluorescencyjną sondą, która selektywnie wiąże się do G-kwadrupleksów poprzez oddziaływania z ich G-tetradami [109]. W reakcji kontrolnej

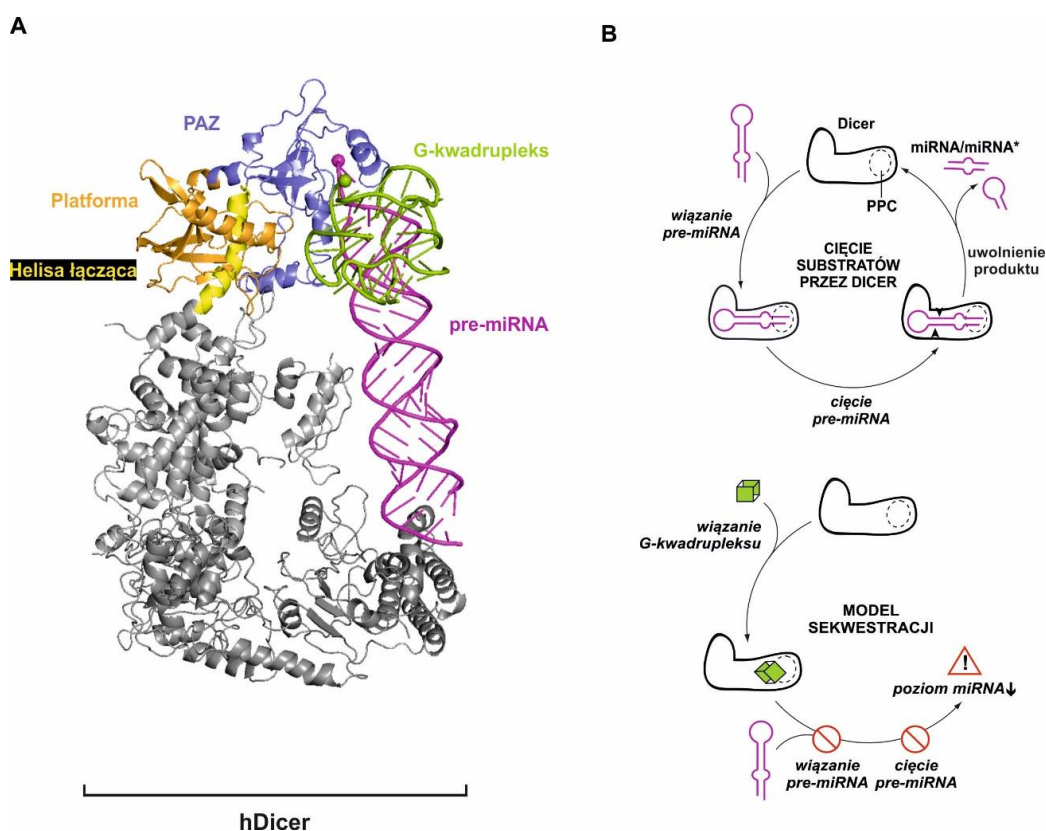
wykorzystano cząsteczkę RNA o długości 32 nt (LIN32), która nie przyjmuje struktury G-kwadrupleksu i którą wyznakowano w taki sam sposób jak cząsteczkę TER22 (**Suplement S2**) (Metody 3.2.6).

W reakcji wiązania wykorzystano 500 nM PPC oraz o-BMVC-TER22 lub o-BMVC-LIN32 o stężeniu 7.5, 15, 22.5 oraz 30  $\mu$ M. Kontrolę ujemną (K-) stanowiła reakcja bez białka. Po zakończonej elektroforezie, żel skanowano przy użyciu skanera FLA-5100 (Fujifilm) przy długości fali 532 nm, w celu wizualizacji struktur G-kwadrupleksów wyznakowanych sondą o-BMVC (**Rysunek 33A**). Następnie ten sam żel barwiono w roztworze SYBR Gold, przez 15 min w celu wizualizacji całkowitej puli RNA (**Rysunek 33B**) i ponownie skanowano przy użyciu skanera FLA-5100 (Fujifilm), tym razem przy długości fali 473 nm. Otrzymane wyniki wykazały, że TER22 po związaniu w obrębie kasety PPC hDicer zachowuje strukturę G-kwadrupleksu.



**Rysunek 33. Badanie oddziaływań pomiędzy G-kwadrupleksem RNA a kasetą PPC hDicer w celu sprawdzenia czy związana cząsteczka zachowuje strukturę G-kwadrupleksu. (A)** Testy wiązania dla PPC i TER22 lub LIN32. Oba oligomery na końcu 5' posiadały sondę o-BMVC. W reakcji wykorzystano 500 nM PPC oraz rosnące stężenia oligomeru (oligo): 7.5, 15, 22.5, 30  $\mu$ M (oznaczone trójkątem). Znaki: (+) oraz (-) odnoszą się odpowiednio do obecności, lub braku PPC lub cząsteczki RNA w mieszaninie reakcyjnej. Aby zwizualizować struktury G-kwadrupleksów, żel skanowano przy długości fali 532 nm. **(B)** Ten sam żel, co przedstawiony w panelu A, barwiono w roztworze SYBR Gold i skanowano przy długości fali 473 nm, w celu wizualizacji całkowitej puli RNA. Próby rozdzielano w 10% PAA. Rysunek przygotowany na podstawie publikacji Koralewska&Szczepanska i in. [102].

Wyniki wcześniejszych badań, zaprezentowane w rozprawie doktorskiej dr N. Koralewskiej [A], wykazały, że aktywność RNazowa Dicer może być hamowana przez krótkie cząsteczki RNA i DNA przyjmujące struktury G-kwadrupleksu. W oparciu o wyniki tych badań, a także na podstawie danych zaprezentowanych w niniejszej pracy doktorskiej, zaproponowano model sekwestracji hDicer przez cząsteczki przyjmujące struktury G-kwadrupleksów: G-kwadrupleksy wiążąc się z kasetą PPC hDicer blokują wiązania substratów pre-miRNA, a tym samym hamują powstawanie miRNA (**Rysunek 34**).



**Rysunek 34. Model sekwestracji hDicer przez cząsteczki przyjmujące strukturę G-kwadrupleksów.** (A) Model struktury hDicer w kompleksie z pre-miRNA (kod PDB: 5ZAL), z nałożoną strukturą kasety PPC w kompleksie z TER10 (struktury nałożono w programie PyMOL). Na rysunku oznaczono końce 3' pre-miRNA oraz TER10. (B) Model przedstawiający przypuszczalny mechanizm regulacji aktywności Dicer, zależny od sekwestracji enzymu. Górny panel przedstawia kanoniczny proces cięcia substratów przez Dicer, a panel dolny obrazuje model sekwestracji białka i zahamowanie aktywności RNazowej Dicer, w wyniku związania cząsteczki G-kwadrupleksu. Rysunek przygotowano na podstawie publikacji Koralewska&Szczepanska i in. [102].

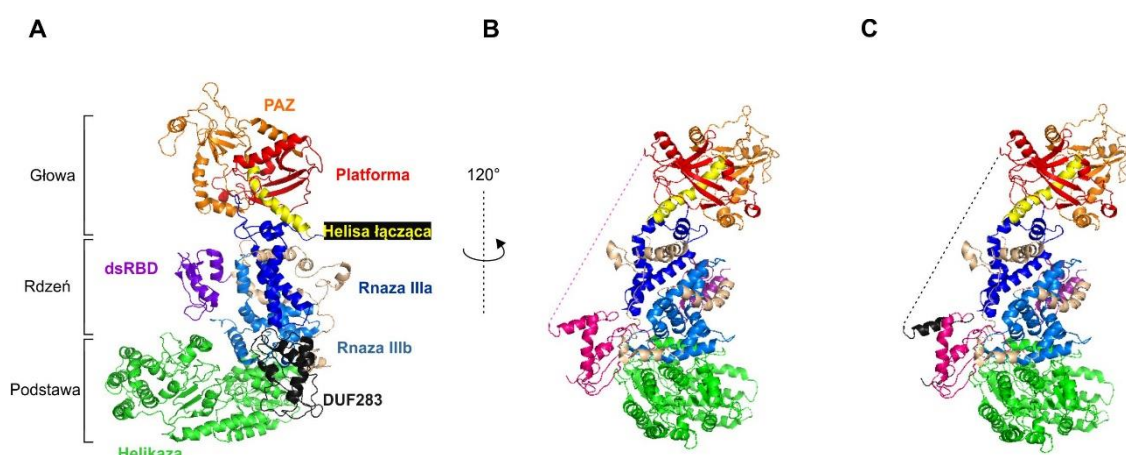
Wyniki przedstawione w tej części doktoratu zostały opublikowane w: „*RNA and DNA G-quadruplexes bind to human dicer and inhibit its activity*”, 2021, *Cellular and Molecular Life Sciences*, Koralewska&Szczepanska i in. (**Załącznik 2**).

## 5. Dyskusja

hDicer jest dużym, wielodomenowym białkiem biorącym udział w procesie produkcji miRNA w komórce. Rola jaką pełnią poszczególne domeny hDicer w procesie cięcia kanonicznych substratów, czyli cząsteczek pre-miRNA oraz pre-siRNA, została już dobrze poznana. Wyjątek stanowi domena DUF283, której znaczenie dla aktywności Dicer wciąż pozostaje niejasne. Początkowo sugerowano, że domena DUF283 może pełnić istotną funkcję w procesie cięcia pre-miRNA, ponieważ badania przeprowadzone z udziałem wariantów Dicer człowieka oraz *Drosophila melanogaster* (*D. melanogaster*) pozbawionych domeny DUF283 oraz domeny helikazy, wykazały brak aktywności RNazowej wspomnianych wariantów delecyjnych [42, 110]. Późniejsze badania *in vitro* ujawniły, że usunięcie tylko domeny DUF283 zmniejsza wydajność cięcia kanonicznych substratów Dicer, w porównaniu do wydajności cięcia białka typu dzikiego [30]. Wcześniejsze badania *in vitro* przeprowadzone w Instytucie Chemii Bioorganicznej PAN ujawniły, że DUF283 hDicer wiąże jednoniciowe RNA oraz DNA, nie oddziałuje natomiast z dwuniciowymi RNA oraz DNA [34]. Ponadto, wykazano, że domena ta może wspierać proces parowania cząsteczek kwasów nukleinowych. Aktywność taką zaobserwowano również dla pełnej długości hDicer. Na podstawie uzyskanych wyników postanowiono hipotezę, że domena DUF283 jest kluczowa dla aktywności hDicer wspierającej parowanie kwasów nukleinowych. W celu potwierdzenia tej hipotezy stworzono dwa warianty delecyjne hDicer: (i) wariant z delecją domeny DUF283 oraz fragmentów oskrzydających tę domenę (wariant z delecją pomiędzy 625 a 752 aa hDicer:  $\Delta$ DUF(625-752)), oraz (ii) wariant z delecją domeny DUF283 pomiędzy 630 a 709 aa ( $\Delta$ DUF (630-709)). Warto wspomnieć w tym miejscu, że fragment 625-752 aa hDicer został wykorzystany w naszych wcześniejszych badaniach charakteryzujących aktywności biochemiczne domeny DUF283 hDicer [34]. Natomiast wariant  $\Delta$ DUF(630-709) został scharakteryzowany pod kątem aktywności RNazowej przez grupę prof. J. Doudny [30].

Trzeciorzędowa struktura Dicer przypomina kształtem literę „L” [21, 25]. W obrębie tej struktury można wyróżnić głowę, rdzeń oraz podstawę (**Rysunek 2**). Głowę budują domeny PAZ i Platforma, domeny RNazy III stanowią rdzeń katalityczny enzymu, natomiast domena helikazy oraz DUF283 stanowią podstawę białka. Usunięcie fragmentu hDicer pomiędzy 625 a 752 aa, obejmującego nie tylko domenę DUF283, ale także część fragmentu łączącego głowę z podstawą białka, przypuszczalnie wpłynęło na

położenie domeny helikazowej w strukturze trzeciorzędowej białka (**Rysunek 35**). W przypadku wariantu  $\Delta$ DUF(630-709), ze względu na obecność wspomnianego powyżej łącznika, domena helikazy prawdopodobnie zachowała większą zdolność do przemieszczania się w odniesieniu do pozostałych domen hDicer, w odróżnieniu od wariantu  $\Delta$ DUF(625-752) (**Rysunek 35**).



**Rysunek 35. Budowa ludzkiej rybonukleazy Dicer oraz wariantów delecyjnych hDicer.** (A) Struktura 3D białka hDicer z zaznaczeniem poszczególnych domen. (B) Struktura 3D wariantu  $\Delta$ DUF(625-752), usunięty fragment został oznaczony na różowo. (C) Struktura 3D wariantu  $\Delta$ DUF(630-709), usunięty fragment został oznaczony na różowo. Struktury zobrazowano w programie PyMOL na podstawie kodu: 5ZAL z bazy danych PDB [25]. Rysunek na podstawie publikacji Szczepanska i in.[26].

Niemniej jednak liczne badania wykazały, w tym najnowsze przedstawiające model cięcia pre-miRNA przez hDicer [24, 27, 31], że domena helikazy jest ważna podczas wyboru substratów, jednakże nie bierze ona udziału w procesie ich cięcia [30, 37, 111]. Zatem, można przypuszczać, że odmienne ułożenie domeny helikazy nie powinno wpłynąć na samą aktywność RNazową obu testowanych wariantów delecyjnych hDicer. Można jednakże domniemywać, że zmiana położenia domeny helikazowej może wpłynąć na proces rozpoznawania i wiązania substratów przez hDicer.

Przeprowadzone badania *in vitro* wykazały, że wariant  $\Delta$ DUF(625-752) przecinał substrat pre-mir-16-1 tak samo wydajnie jak hDicer typu dzikiego (**Rysunek 8**). Różnice w aktywności RNazowej wariantu  $\Delta$ DUF(625-752) odnotowano dla pozostałych dwóch testowanych substratów: pre-mir-21 oraz dsRNA\_ov, dla których obserwowano powstawanie śladowych ilości produktów cięcia (**Rysunek 8**). Należy zauważyć, że substrat pre-mir-16-1 zawiera wewnętrzne pętle i wybrzuszenia w rejonie trzonu oraz 9 nt pętlę apikalną, przez co charakteryzuje się bardziej rozluźnioną strukturą

w porównaniu do pre-mir-21, który przyjmuje bardziej zwartą strukturę, z 3 nt pętlą w części apikalnej (**Rysunek 8**). Analizując przewidywane struktury drugorzędowe testowanych substratów (**Rysunek 8**) można zaobserwować, że ze względu na znaczny stopień sparowania zasad pre-mir-21, struktura tego substratu jest bardziej podobna do struktury dsRNA, niż do struktury pre-mir-16-1. Przepuszczalnie jednoniciowe rejony pre-mir-16-1 zapewniają większą elastyczność substratu i dzięki temu jego lepsze dopasowanie do doliny katalitycznej wariantu  $\Delta$ DUF(625-752). Na podstawie otrzymanych wyników można wnioskować, iż substraty: pre-mir-21 oraz dsRNA<sub>ov</sub>, w warunkach *in vitro*, nie wpasowały się odpowiednio w rejon katalityczny domen RNazowych wariantu  $\Delta$ DUF(625-752), co mogło być przyczyną niewydajnej produkcji miRNA, czy też siRNA (**Rysunek 8**). Co ciekawe, wariant  $\Delta$ DUF(630-709) przecinał wszystkie badane substraty co najmniej tak samo wydajnie jak hDicer pełnej długości (**Rysunek 9**). Otrzymane wyniki w dużym stopniu potwierdzają przypuszczenia, że odmienne położenie domeny helikazowej w wariantach  $\Delta$ DUF(625-752) i  $\Delta$ DUF(630-709) mogło wpłynąć na proces selekcji substratów w warunkach *in vitro*, co zostało odzwierciedlone w wydajności powstawania określonych produktów reakcji cięcia (**Rysunek 8 i 9**). Ponadto, można domniemywać, że nie tylko domena helikazy [29], ale również domena DUF283 może odgrywać ważną rolę w rozpoznawaniu substratów hDicer i ich dopasowaniu do doliny katalitycznej. Usunięcie obu domen hDicer: domeny helikazy oraz domeny DUF283, spowodowało brak cięcia kanonicznych substratów hDicer w warunkach *in vitro* [42]. W oparciu o otrzymane wyniki oraz dane literaturowe [24] można przypuszczać, że orientacja przestrzenna domen Dicer ma kluczowe znaczenie w procesie selekcji substratów oraz ich dopasowaniu do doliny katalitycznej, co w konsekwencji wpływa na wydajność produkcji cząsteczek miRNA oraz siRNA przez białka typu Dicer. W warunkach komórkowych białka partnerujące rybonukleazom Dicer mogą wpływać na proces selekcji substratu i przypuszczalnie przestrzenną orientację domen enzymu. Przykładami białek partnerujących hDicer są białka TRBP, PACT oraz ADAR1 [25, 28, 36, 42], przy czym z danych literaturowych wiadomo, iż z domeną DUF283 hDicer oddziałuje białko ADAR1 [36]. Badania przeprowadzone w warunkach *in cellulo* wykazały, że w komórkach HEK 293T NoDice produkujących wariant  $\Delta$ DUF(625-752) poziom miR-16-1-5p oraz miR-21-5p był obniżony o odpowiednio 40% i 30%, w porównaniu do poziomu odpowiednich miRNA powstających w komórkach HEK 293T NoDice produkujących pełnej długości hDicer (kontrola pozytywna). W przypadku wariantu  $\Delta$ DUF(630-709), poziom produkcji miR-



16-1-5p obniżył się o ok. 30%, a dla miR-21-5p nie odnotowano zmian w poziomie produkcji, w porównaniu do kontroli pozytywnej, którą stanowiło białko typu dzikiego (**Rysunek 10B, C**). Można zatem zauważyć, że w zależności od substratu pre-miRNA oraz zastosowanego układu reakcyjnego (*in vitro* vs *in cellulo*), poziom produkcji określonych miRNA przez warianty  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709) był różny. Przyczyną zaobserwowanych rozbieżności może być wpływ czynników komórkowych na strukturę cząsteczek pre-miRNA i ich dostępność dla hDicer [67, 112]. Ponadto, brak domeny DUF283 przypuszczalnie mógł spowodować zaburzenie wiązania się białek partnerujących, lub innych czynników komórkowych, do wariantów delecyjnych hDicer, co mogło także wpłynąć na poziom produkcji badanych miRNA.

Co więcej, przeprowadzone badania wykazały, że zarówno delecja 625-752 aa, jak i delecja 630-709 aa w obrębie domeny DUF283 hDicer, zniósł aktywność wspierania parowania cząsteczek kwasów nukleinowych odpowiednich wariantów delecyjnych, w badanych układach reakcyjnych (**Rysunek 11**). Otrzymane rezultaty w znacznym stopniu wspierają hipotezę, że domena DUF283 jest kluczowa dla aktywności hDicer wspierającej parowanie cząsteczek kwasów nukleinowych.

Chcąc dalej zgłębiać znaczenie domeny DUF283 dla aktywności hDicer, postanowiono zbadać rolę tej domeny w wiązaniu puli komórkowych RNA przez hDicer. W tym celu zastosowano podejście irCLIP-seq oraz następujące układy – linie komórkowe: (i) komórki HEK 293T – linia typu dzikiego (WT hDicer), (ii) komórki HEK 293T NoDice transfekowane plazmidem kodującym pełnej długości hDicer (tzw. *rescue control*, hDicer), (iii) komórki HEK 293T NoDice transfekowane plazmidem kodującym  $\Delta$ DUF(625-752) oraz (iv) HEK 293T NoDice transfekowane plazmidem kodującym  $\Delta$ DUF(630-709). Przeprowadzone analizy wykazały, że wszystkie badane białka licznie wiązały się z transkryptami komórkowymi. Na poziomie analizy porównawczej typów RNA wiązanych przez badane białka, nie odnotowano większych różnic pomiędzy poszczególnymi pulami RNA wiazanymi przez hDicer i warianty  $\Delta$ DUF hDicer (**Rysunek 21**). Dopiero analizy różnicowe prowadzone na poziomie sekwencji, z rozdzielczością do pojedynczego nukleotydu, wykazały różnice pomiędzy pulami RNA wiazanymi przez hDicer oraz warianty  $\Delta$ DUF hDicer (**Rysunek 24- 26**). Badania te będą kontynuowane i rozszerzone o szczegółowe analizy miejsc wiązań w obrębie transkryptów innych genów.

Co ciekawe, w puli RNA wiązanych przez hDicer odnotowano liczne sekwencje bogate w ciągi G. Z wykorzystaniem programu Quadron [99], który służy do predykcji struktur G-kwadrupleksów, ustalono, że wiele ze zidentyfikowanych RNA bogatych w ciągi G ma potencjał do tworzenia struktur G-kwadrupleksu (**Tabela 24**). Wyniki wcześniejszych badań, zaprezentowane w rozprawie doktorskiej dr N. Koralewskiej [A], ujawniły także, że w warunkach *in vitro* hDicer może wiązać się z cząsteczkami przyjmującymi strukturę G-kwadrupleksu.

G-kwadrupleksy to niekanoniczne struktury tworzone przez bogate w ciągi guanin cząsteczki RNA oraz DNA, w których cztery guaniny, tzw. tetrazy G, są stabilizowane poprzez wiązania wodorowe Hoogsteena i monowalentny kation, najczęściej  $\text{Na}^+$  lub  $\text{K}^+$  [113]. Motywy zawierające G-kwadrupleksy występują w genomach i transkryptomach wielu organizmów, w tym także człowieka. Po raz pierwszy G-kwadrupleksy DNA zostały opisane w kontekście telomerowego DNA [114]. G-kwadrupleksy mogą odgrywać ważną rolę w wielu procesach komórkowych, m. in. pełnią ważną funkcję w regulacji replikacji DNA, ekspresji genów i utrzymaniu integralności telomerów [115]. Zaburzenia w tworzeniu struktur G-kwadrupleksu powiązane mogą być także z niektórymi chorobami człowieka, przykładowo z nowotworami [116], czy chorobami neurodegeneracyjnymi, np. ataksjami czy zespołem łamliwego chromosomu X [115, 117, 118]. Zaobserwowano również udział struktur G-kwadrupleksów w regulacji metabolizmu RNA, w tym ścieżki biogenezy miRNA, gdzie struktury te mogą wpływać na aktywność Dicer i inne etapy biogenezy miRNA [119-121].

Białka typu Dicer znane są głównie z ich kluczowej roli w procesie biogenezy srRNA: miRNA oraz siRNA. Pojawia się jednak coraz więcej doniesień literaturowych wskazujących, że poza tą kanoniczną rolą, rybonukleazy Dicer pełnią także szereg innych funkcji [122]. Dicer jako endorybonukleaza bierze udział w metabolizmie różnych typów RNA, przykładowo: tRNA [123], RNA powstających w trakcie pęknięć DNA (tzw. diRNA, ang. *DNA damage-induced small RNA*) [124], czy też transpozonów [125]. Ponadto u nicieni *C. elegans*, skrócona forma Dicer, generowana przez aktywowane w procesie apoptozy kaspazy, działa jako deoksyrybonukleaza inicjująca proces fragmentacji chromosomalnego DNA [126]. Jednakże aktywność rybonukleaz Dicer nie musi być powiązana z procesem hydrolizy kwasów nukleinowych. Jak wspomniano powyżej, badania prowadzone w grupie prof. N. Rajewskiego wykazały, że Dicer może



wiązać szereg komórkowych RNA, np. mRNA i lncRNA, w sposób pasywny, czyli bez przeprowadzania cięcia [79].

Wyniki badań zaprezentowanych w niniejszej rozprawie doktorskiej wskazują, że struktura G-kwadrupleksu, występująca w obrębie cząsteczki RNA, czy też DNA, nie stanowi przeszkody w wiązaniu się takiej cząsteczki do hDicer (**Rysunek 27**). G-kwadrupleksy wiązane są przypuszczalnie w obrębie kasety PPC hDicer (**Rysunek 28**). W zastosowanych warunkach reakcyjnych udało się eksperymentalnie wyznaczyć równowagowe stałe wiązania, które wynosiły dla kompleksów PPC•RNA-G4, odpowiednio: ~7 nM dla kompleksu PPC•TER22 oraz ~10 nM dla kompleksu PPC•G4U4G4 (**Rysunek 28F**). Dla kompleksów PPC•DNA-G4 wartości K<sub>d</sub> wynosiły: ~318 nM dla kompleksu PPC•TEL22 oraz ~369 nM dla kompleksu PPC•G4T4G4 (**Rysunek 28F**). Wiązanie substratów o podobnej długości, lecz nieprzyjmujących struktur G-kwadrupleksu, było mniej wydajne; obliczone wartości równowagowych stałych wiązania wynosiły odpowiednio: ~205 nM dla kompleksu PPC•pre-mir-21 oraz ~46 nM dla kompleksu PPC•RNA21 [102]. Zgromadzone dane sugerowały zatem, iż tak silne wiązanie G-kwadrupleksów do hDicer może wpływać na proces cięcia pre-miRNA. Rzeczywiście, przeprowadzone reakcje cięcia substratów pre-miRNA przez hDicer w obecności G-kwadrupleksu RNA czy DNA wykazały, że zahamowany zostaje proces cięcia kanonicznych substratów [A]. Obserwowana inhibicja aktywności RNazowej hDicer była wydajna i stabilna w czasie prowadzonych reakcji cięcia [102].

W oparciu o uzyskane wyniki zaproponowano, że związanie się hDicer do struktur G-kwadrupleksów tworzonych przez komórkowe RNA, czy też DNA, może powodować sekwestrację hDicer i zablokowanie procesu cięcia pre-miRNA (**Rysunek 34**). Podobne scenariusze zostały zaproponowane także przez inne grupy badawcze: sekwestracja Dicer poprzez wiązanie do struktur typu spinki do włosów tworzonych w obrębie endogennych transkryptów [79] lub przyjmowanych przez wirusowe RNA [127].

Na podstawie wyników modelowania molekularnego kompleksów PPC•G4-RNA oraz PPC•G4-DNA (**Rysunek 30**) oraz przeprowadzonych analiz z wariantami białka PPC posiadającymi mutacje w obrębie kieszeni wiążących (mutacje upośredzające wiązanie kanonicznych substratów hDicer) (**Rysunek 31**), zaproponowano, że cząsteczki G4-RNA i G4-DNA mogą wiązać się zarówno w obrębie kieszeni wiążącej 5' jak i kieszeni wiążącej 3' kasety PPC (**Rysunek 30**). Co więcej, wiązanie cząsteczek

przyjmujących struktury G-kwadrupleksu było niezależne od dostępności ich końców 5' i 3' (**Rysunek 32**). Wydaje się zatem, że hDicer może wiązać się do struktur G-kwadrupleksów niezależnie od tego czy występują one wewnątrzcząsteczkowo (np. w obrębie sekwencji promotorowych, czy też wewnątrz transkryptów [106]), czy też na końcach RNA lub DNA (sekwencje telomerowe).

Białka wiążące struktury G-kwadrupleksów występujące w obrębie cząsteczek mRNA są ważne dla wewnątrzkomórkowego transportu tych cząsteczek. Transport transkryptów jest szczególnie ważny w takich komórkach jak neurony, w których wzrost i różnicowanie zależy od długodystansowego transportu określonego zestawu mRNA do miejsc translacji, tj. do obszarów rosnącego aksonu czy synaps [106]. W pracy Ishiguro i in. wykazano, że białko TDP-43 (ang. *TAR DNA-binding protein of 43 kDa*) odgrywa istotną rolę w wewnątrzkomórkowym transporcie mRNA, który zawiera struktury G-kwadrupleksów, do dystalnych neurytów [128]. Dodatkowo zaobserwowano, że zaburzenie funkcji TDP-43 skutkuje zaburzeniami neurodegeneracyjnym, tj. stwardnieniem zanikowym bocznym, czy zwyrodnieniem płatów czołowych [128]. Ponadto, białka wiążące G-kwadrupleksy mogą działać jako „przełączniki” - regulatory kontrolujące dostęp miRNA do sekwencji komplementarnych bogatych w guaninę i przyjmujących struktury G-kwadrupleksu, występujących w obrębie rejonów 3' UTR mRNA. Przykładowo, w pracy Stefanovic i in. wykazano, że rejon bogaty w ciągi reszt guaniny, obecny w mRNA, z którego powstaje białko PSD-95 (ang. *postsynaptic density protein 95*), przyjmuje strukturę G-kwadrupleksu, z którą oddziałuje białko FMRP (ang. *Fragile X mental retardation protein*). Białko FMRP, poprzez wiązanie do struktury G-kwadrupleksu obecnej w mRNA PSD-95, umożliwia dostęp miR-125a i w ten sposób odwracalnie hamuje proces biosyntezy PSD-95 w neuronach [129, 130].

Biorąc pod uwagę zgromadzone dane dotyczące wiązania się hDicer do struktur G-kwadrupleksów występujących w komórkowych RNA, można przypuszczać, że rybonukleazy Dicer mogą być także bezpośrednio zaangażowane w kontrolę biosyntezy białek powstających z mRNA tworzących lokalne struktury G-kwadrupleksu. Możliwe bezpośrednie zaangażowanie hDicer w proces potranskrypcyjnej kontroli ekspresji genów było już szeroko badane w kontekście aktywności hDicer wspierającej proces parowania RNA-RNA oraz interakcji pomiędzy hDicer, związanymi z nią miRNA i docelowymi transkryptami [93]. Udało się wykazać, że w przeciwieństwie do białka Ago2, które jest głównym składnikiem kompleksu RISC, to właśnie hDicer wykazuje

zdolność wspierania parowania miRNA do komplementarnych miejsc docelowych występujących w obrębie stabilnych struktur mRNA [93]. Obecność hDicer została także potwierdzona w jądrze komórkowym [76], co razem z odkryciem, że hDicer oddziałuje z G-kwadrupeksami DNA, sugeruje, że białko to mogłoby także wiązać się do struktur G-kwadrupeksów występujących w genomowym DNA.

Badania realizowane w ramach niniejszej pracy doktorskiej wspierają hipotezę, że komórkowa rola Dicer nie ogranicza się do ścieżki biogenezy srRNA. Odkrycie, że hDicer może wiązać G-kwadrupeksy RNA i DNA, oraz że ta interakcja wpływa na aktywność RNazową Dicer, otwiera nowe drogi badań nad komórkowymi funkcjami Dicer.

## 6. Podsumowanie

Celem badań prowadzonych w ramach realizacji przedłożonej pracy doktorskiej było poszerzenie wiedzy na temat wybranych aspektów oddziaływań pomiędzy ludzką rybonukleazą Dicer i cząsteczkami kwasów nukleinowych. Prowadzone badania skupiały się w głównej mierze na pogłębieniu wiedzy na temat znaczenia domeny DUF283 dla aktywności hDicer. W badaniach aktywności RNazowej i aktywności wspierającej hybrydyzację komplementarnych cząsteczek RNA wykorzystano dwa otrzymane preparaty białkowe hDicer:  $\Delta$ DUF(625-752) oraz  $\Delta$ DUF(630-709). Pierwszy z wariantów, wariant  $\Delta$ DUF(625-752), oprócz domeny DUF283, pozbawiony był także fragmentów oskrzydlających tę domenę. Drugi wariant, wariant  $\Delta$ DUF(630-709), pozbawiony był samej domeny DUF283, tj. fragmentu hDicer pomiędzy 630 a 709 aa. Z wykorzystaniem metody irCLIP-seq (ang. *infrared crosslinking immunoprecipitation followed by NGS sequencing*), badano również jak delecja domeny DUF283 wpłynie na zdolność wiązania komórkowych RNA przez warianty  $\Delta$ DUF hDicer. W badaniach tych wykorzystano następujące układy – linie komórkowe: (i) HEK 293T – linia typu dzikiego (WT hDicer), (ii) HEK 293T NoDice transfekowane plazmidem kodującym pełnej długości białko hDicer (tzw. *rescue control*, hDicer), (iii) HEK 293T NoDice transfekowane plazmidem kodującym  $\Delta$ DUF(625-752) oraz (iv) HEK 293T NoDice transfekowane plazmidem kodującym  $\Delta$ DUF(630-709). Następnie, biorąc pod uwagę wyniki wstępnych analiz wskazujące, iż hDicer może oddziaływać w komórkach z cząsteczkami kwasów nukleinowych przyjmującymi struktury G-kwadrupleksu, badano oddziaływania pomiędzy hDicer i cząsteczkami RNA i DNA przyjmującymi struktury G-kwadrupleksu. W tym celu wykorzystano metodę spowolnionej migracji w żelach poliakrylamidowych oraz modelowanie molekularne.

Najważniejszymi rezultatami płynącymi z przeprowadzonych badań są:

- potwierdzenie, że domena DUF283 hDicer ma wpływ na proces cięcia pre-miRNA,
- wykazanie, że domena DUF283 jest niezbędna dla aktywności hDicer wspierającej parowanie cząsteczek kwasów nukleinowych,
- wykazanie różnic w pulach komórkowych RNA wiązanych przez hDicer, wariant  $\Delta$ DUF(625-752) oraz wariant  $\Delta$ DUF(630-709), co sugeruje, że domena DUF283

hDicer bierze udział w rozpoznawaniu i wiązaniu określonej puli komórkowych RNA,

- wskazanie, że rejonem hDicer odpowiedzialnym za wiązanie G-kwadrupleksów RNA oraz DNA jest przypuszczalnie kasetta PPC oraz, że w wiązanie G-kwadrupleksów przez hDicer nie wymaga dostępności końców 5' i 3' tych cząsteczek,
- zaproponowanie modelu sekwestracji hDicer przez cząsteczki przyjmujące strukturę G-kwadrupleksu.

Omówione w rozprawie doktorskiej wyniki zostały przedstawione w następujących publikacjach:

1. **Szczepanska A.**, Wojnicka M. & Kurzynska-Kokorniak A. (2021). The Significance of the DUF283 Domain for the Activity of Human Ribonuclease Dicer. *International Journal of Molecular Sciences*, 13;22(16):8690. IF<sub>2021</sub> 6,208.
2. Koralewska N.\*, **Szczepanska A.\***, Ciechanowska K., Wojnicka M., Pokornowska M., Milewski C.M., Gudanis D., Baranowski D., Nithin C., Bujnicki M.J., Gdaniec Z., Figlerowicz M. & Kurzynska-Kokorniak A. (2021). RNA and DNA G-quadruplexes bind to human dicer and inhibit its activity. *Cellular and Molecular Life Sciences*, 78:3709–3724. IF<sub>2021</sub> 9,207.

(\*) pierwszy współautor

## 7. Bibliografia

- [1] Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*. 1993;75:843-54.
- [2] Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell*. 1993;75:855-62.
- [3] Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*. 1998;391:806-11.
- [4] MacRae IJ, Doudna JA. Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr Opin Struct Biol*. 2007;17:138-45.
- [5] Provost P, Dishart D, Doucet J, Friendewey D, Samuelsson B, Radmark O. Ribonuclease activity and RNA binding of recombinant human Dicer. *EMBO J*. 2002;21:5864-74.
- [6] Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004;116:281-97.
- [7] Berezhikov E, van Tetering G, Verheul M, van de Belt J, van Laake L, Vos J, et al. Many novel mammalian microRNA candidates identified by extensive cloning and RAKE analysis. *Genome Res*. 2006;16:1289-98.
- [8] Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res*. 2009;19:92-105.
- [9] O'Brien J, Hayder H, Zayed Y, Peng C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front Endocrinol (Lausanne)*. 2018;9:402.
- [10] Kim VN. MicroRNA biogenesis: coordinated cropping and dicing. *Nat Rev Mol Cell Biol*. 2005;6:376-85.
- [11] Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, et al. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J*. 2004;23:4051-60.
- [12] Kim YK, Kim VN. Processing of intronic microRNAs. *EMBO J*. 2007;26:775-83.
- [13] de Rie D, Abugessaisa I, Alam T, Arner E, Arner P, Ashoor H, et al. An integrated expression atlas of miRNAs and their promoters in human and mouse. *Nat Biotechnol*. 2017;35:872-8.
- [14] Lee Y, Jeon K, Lee JT, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J*. 2002;21:4663-70.
- [15] Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev*. 2003;17:3011-6.
- [16] Kawamata T, Seitz H, Tomari Y. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat Struct Mol Biol*. 2009;16:953-60.
- [17] Gamazon ER, Im HK, Duan S, Lussier YA, Cox NJ, Dolan ME, et al. Exprtarget: an integrative approach to predicting human microRNA targets. *PLoS One*. 2010;5:e13534.
- [18] Martello G, Rosato A, Ferrari F, Manfrin A, Cordenonsi M, Dupont S, et al. A MicroRNA targeting dicer for metastasis control. *Cell*. 2010;141:1195-207.
- [19] Bogerd HP, Whisnant AW, Kennedy EM, Flores O, Cullen BR. Derivation and characterization of Dicer- and microRNA-deficient human cells. *RNA*. 2014;20:923-37.
- [20] Macrae IJ, Li F, Zhou K, Cande WZ, Doudna JA. Structure of Dicer and mechanistic implications for RNAi. *Cold Spring Harb Symp Quant Biol*. 2006;71:73-80.

- [21] Lau PW, Guiley KZ, De N, Potter CS, Carragher B, MacRae IJ. The molecular architecture of human Dicer. *Nat Struct Mol Biol.* 2012;19:436-40.
- [22] Svobodova E, Kubikova J, Svoboda P. Production of small RNAs by mammalian Dicer. *Pflugers Arch.* 2016;468:1089-102.
- [23] Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. *Cell.* 2004;118:57-68.
- [24] Lee YY, Lee H, Kim H, Kim VN, Roh SH. Structure of the human DICER-pre-miRNA complex in a dicing state. *Nature.* 2023;615:331-8.
- [25] Liu Z, Wang J, Cheng H, Ke X, Sun L, Zhang QC, et al. Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate. *Cell.* 2018;173:1191-203.e12.
- [26] Szczepanska A, Wojnicka M, Kurzynska-Kokorniak A. The Significance of the DUF283 Domain for the Activity of Human Ribonuclease Dicer. *Int J Mol Sci.* 2021;22.
- [27] Kidwell MA, Chan JM, Doudna JA. Evolutionarily conserved roles of the dicer helicase domain in regulating RNA interference processing. *J Biol Chem.* 2014;289:28352-62.
- [28] Ciechanowska K, Pokornowska M, Kurzynska-Kokorniak A. Genetic Insight into the Domain Structure and Functions of Dicer-Type Ribonucleases. *Int J Mol Sci.* 2021;22.
- [29] Taylor DW, Ma E, Shigematsu H, Cianfrocco MA, Noland CL, Nagayama K, et al. Substrate-specific structural rearrangements of human Dicer. *Nature structural & molecular biology.* 2013;20:662-70.
- [30] Ma E, MacRae IJ, Kirsch JF, Doudna JA. Autoinhibition of human dicer by its internal helicase domain. *J Mol Biol.* 2008;380:237-43.
- [31] Zapletal D, Taborska E, Pasulka J, Malik R, Kubicek K, Zanova M, et al. Structural and functional basis of mammalian microRNA biogenesis by Dicer. *Mol Cell.* 2022;82:4064-79 e13.
- [32] Soifer HS, Sano M, Sakurai K, Chomchan P, Saetrom P, Sherman MA, et al. A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. *Nucleic Acids Res.* 2008;36:6511-22.
- [33] Dlakic M. DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold. *Bioinformatics.* 2006;22:2711-4.
- [34] 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.
- [35] 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;16:474-81.
- [36] Ota H, Sakurai M, Gupta R, Valente L, Wulff BE, Ariyoshi K, et al. ADAR1 Forms a Complex with Dicer to Promote MicroRNA Processing and RNA-Induced Gene Silencing. *Cell.* 2013;153:575-89.
- [37] Lee Y, Hur I, Park S-Y, Kim Y-K, Suh MR, Kim VN. The role of PACT in the RNA silencing pathway. *The EMBO Journal.* 2006;25:522-32.
- [38] Tian Y, Simanshu DK, Ma JB, Park JE, Heo I, Kim VN, et al. A Phosphate-Binding Pocket within the Platform-PAZ-Connector Helix Cassette of Human Dicer. *Mol Cell.* 2014;53:606-16.

- [39] Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, et al. Structural basis for double-stranded RNA processing by Dicer. *Science*. 2006;311:195-8.
- [40] Ma E, Zhou K, Kidwell MA, Doudna JA. Coordinated activities of human dicer domains in regulatory RNA processing. *J Mol Biol*. 2012;422:466-76.
- [41] Wojnicka M, Szczepanska A, Kurzynska-Kokorniak A. Unknown Areas of Activity of Human Ribonuclease Dicer: A Putative Deoxyribonuclease Activity. *Molecules*. 2020;25.
- [42] Lee Y, Hur I, Park SY, Kim YK, Suh MR, Kim VN. The role of PACT in the RNA silencing pathway. *EMBO J*. 2006;25:522-32.
- [43] Li HL, Chelladurai BS, Zhang K, Nicholson AW. Ribonuclease III cleavage of a bacteriophage T7 processing signal. Divalent cation specificity, and specific anion effects. *Nucleic Acids Res*. 1993;21:1919-25.
- [44] Robertson HD, Webster RE, Zinder ND. Purification and properties of ribonuclease III from *Escherichia coli*. *J Biol Chem*. 1968;243:82-91.
- [45] Chang KY, Ramos A. The double-stranded RNA-binding motif, a versatile macromolecular docking platform. *FEBS J*. 2005;272:2109-17.
- [46] Wostenberg C, Lary JW, Sahu D, Acevedo R, Quarles KA, Cole JL, et al. The role of human Dicer-dsRBD in processing small regulatory RNAs. *PLoS One*. 2012;7:e51829.
- [47] Lima WF, Murray H, Nichols JG, Wu H, Sun H, Prakash TP, et al. Human Dicer binds short single-strand and double-strand RNA with high affinity and interacts with different regions of the nucleic acids. *J Biol Chem*. 2009;284:2535-48.
- [48] Lehrbach NJ, Miska EA. Regulation of pre-miRNA Processing. *Adv Exp Med Biol*. 2010;700:67-75.
- [49] Feng Y, Zhang X, Graves P, Zeng Y. A comprehensive analysis of precursor microRNA cleavage by human Dicer. *RNA*. 2012;18:2083-92.
- [50] Gu S, Jin L, Zhang Y, Huang Y, Zhang F, Valdmanis PN, et al. The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo. *Cell*. 2012;151:900-11.
- [51] Mello CC, Conte D, Jr. Revealing the world of RNA interference. *Nature*. 2004;431:338-42.
- [52] Baulcombe D. Viral suppression of systemic silencing. *Trends Microbiol*. 2002;10:306-8.
- [53] Li L, Liu Y. Diverse small non-coding RNAs in RNA interference pathways. *Methods Mol Biol*. 2011;764:169-82.
- [54] Merritt WM, Lin YG, Han LY, Kamat AA, Spannuth WA, Schmandt R, et al. Dicer, Drosha, and outcomes in patients with ovarian cancer. *N Engl J Med*. 2008;359:2641-50.
- [55] Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. *Nat Rev Genet*. 2009;10:94-108.
- [56] Xhemalce B, Robson SC, Kouzarides T. Human RNA methyltransferase BCDIN3D regulates microRNA processing. *Cell*. 2012;151:278-88.
- [57] Heo I, Ha M, Lim J, Yoon MJ, Park JE, Kwon SC, et al. Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. *Cell*. 2012;151:521-32.
- [58] Trabucchi M, Briata P, Garcia-Mayoral M, Haase AD, Filipowicz W, Ramos A, et al. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature*. 2009;459:1010-4.



- [59] Ruggiero T, Trabucchi M, De Santa F, Zupo S, Harfe BD, McManus MT, et al. LPS induces KH-type splicing regulatory protein-dependent processing of microRNA-155 precursors in macrophages. *FASEB J.* 2009;23:2898-908.
- [60] Kawahara Y, Mieda-Sato A. TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proc Natl Acad Sci U S A.* 2012;109:3347-52.
- [61] Pilotte J, Dupont-Versteegden EE, Vanderklish PW. Widespread regulation of miRNA biogenesis at the Dicer step by the cold-inducible RNA-binding protein, RBM3. *PLoS One.* 2011;6:e28446.
- [62] Newman MA, Thomson JM, Hammond SM. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA.* 2008;14:1539-49.
- [63] Viswanathan SR, Daley GQ, Gregory RI. Selective blockade of microRNA processing by Lin28. *Science.* 2008;320:97-100.
- [64] 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.
- [65] Chakravarthy S, Sternberg SH, Kellenberger CA, Doudna JA. Substrate-specific kinetics of Dicer-catalyzed RNA processing. *J Mol Biol.* 2010;404:392-402.
- [66] Daniels SM, Melendez-Pena CE, Scarborough RJ, Daher A, Christensen HS, El Far M, et al. Characterization of the TRBP domain required for dicer interaction and function in RNA interference. *BMC Mol Biol.* 2009;10:38.
- [67] Yoshida T, Asano Y, Ui-Tei K. Modulation of MicroRNA Processing by Dicer via Its Associated dsRNA Binding Proteins. *Noncoding RNA.* 2021;7.
- [68] 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.
- [69] Fukunaga R, Han BW, Hung JH, Xu J, Weng Z, Zamore PD. Dicer partner proteins tune the length of mature miRNAs in flies and mammals. *Cell.* 2012;151:533-46.
- [70] Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, et al. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature.* 2005;436:740-4.
- [71] Haase AD, Jaskiewicz L, Zhang H, Laine S, Sack R, Gatignol A, et al. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep.* 2005;6:961-7.
- [72] Robb GB, Rana TM. RNA helicase A interacts with RISC in human cells and functions in RISC loading. *Mol Cell.* 2007;26:523-37.
- [73] Dinibas-Renqvist V, Pepin G, Rakonjac M, Plante I, Ouellet DL, Hermansson A, et al. Human Dicer C-terminus functions as a 5-lipoxygenase binding domain. *Biochim Biophys Acta.* 2009;1789:99-108.
- [74] Radmark O, Werz O, Steinhilber D, Samuelsson B. 5-Lipoxygenase: regulation of expression and enzyme activity. *Trends Biochem Sci.* 2007;32:332-41.
- [75] Song MS, Rossi JJ. Molecular mechanisms of Dicer: endonuclease and enzymatic activity. *Biochem J.* 2017;474:1603-18.
- [76] Burger K, Schlackow M, Potts M, Hester S, Mohammed S, Gullerova M. Nuclear phosphorylated Dicer processes double-stranded RNA in response to DNA damage. *J Cell Biol.* 2017;216:2373-89.

- [77] Poirier EZ, Buck MD, Chakravarty P, Carvalho J, Frederico B, Cardoso A, et al. An isoform of Dicer protects mammalian stem cells against multiple RNA viruses. *Science*. 2021;373:231-6.
- [78] Montavon TC, Baldaccini M, Lefevre M, Girardi E, Chane-Woon-Ming B, Messmer M, et al. Human DICER helicase domain recruits PKR and modulates its antiviral activity. *PLoS Pathog*. 2021;17:e1009549.
- [79] Rybak-Wolf A, Jens M, Murakawa Y, Herzog M, Landthaler M, Rajewsky N. A Variety of Dicer Substrates in Human and C-elegans. *Cell*. 2014;159:1153-67.
- [80] Kazimierczyk M, Wojnicka M, Biala E, Zydowicz-Machtel P, Imiolczyk B, Ostrowski T, et al. Characteristics of Transfer RNA-Derived Fragments Expressed during Human Renal Cell Development: The Role of Dicer in tRF Biogenesis. *Int J Mol Sci*. 2022;23.
- [81] Sun W, Yang Y, Xu C, Guo J. Regulatory mechanisms of long noncoding RNAs on gene expression in cancers. *Cancer Genet*. 2017;216-217:105-10.
- [82] Maute RL, Schneider C, Sumazin P, Holmes A, Califano A, Basso K, et al. tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;110:1404-9.
- [83] Fu H, Feng J, Liu Q, Sun F, Tie Y, Zhu J, et al. Stress induces tRNA cleavage by angiogenin in mammalian cells. *FEBS Lett*. 2009;583:437-42.
- [84] Shen Y, Yu X, Zhu L, Li T, Yan Z, Guo J. Transfer RNA-derived fragments and tRNA halves: biogenesis, biological functions and their roles in diseases. *J Mol Med (Berl)*. 2018;96:1167-76.
- [85] Kumar P, Anaya J, Mudunuri SB, Dutta A. Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. *BMC Biol*. 2014;12:78.
- [86] Pederson T. Regulatory RNAs derived from transfer RNA? *RNA*. 2010;16:1865-9.
- [87] Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, et al. A human snoRNA with microRNA-like functions. *Mol Cell*. 2008;32:519-28.
- [88] Matera AG, Terns RM, Terns MP. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol*. 2007;8:209-20.
- [89] Lemus-Diaz N, Ferreira RR, Bohnsack KE, Gruber J, Bohnsack MT. The human box C/D snoRNA U3 is a miRNA source and miR-U3 regulates expression of sortin nexin 27. *Nucleic Acids Res*. 2020;48:8074-89.
- [90] Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, et al. Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature*. 2011;475:201-5.
- [91] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>( $\Delta\Delta C_T$ ) Method. *Methods*. 2001;25:402-8.
- [92] Huppertz I, Attig J, D'Ambrogio A, Easton LE, Sibley CR, Sugimoto Y, et al. iCLIP: protein-RNA interactions at nucleotide resolution. *Methods*. 2014;65:274-87.
- [93] Pokornowska M, Milewski MC, Ciechanowska K, Szczepanska A, Wojnicka M, Radogostowicz Z, et al. The RNA-RNA base pairing potential of human Dicer and Ago2 proteins. *Cell Mol Life Sci*. 2019.
- [94] Danan C, Manickavel S, Hafner M. PAR-CLIP: A Method for Transcriptome-Wide Identification of RNA Binding Protein Interaction Sites. *Methods Mol Biol*. 2022;2404:167-88.

- [95] Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*. 2010;141:129-41.
- [96] Huppertz I, Haberman N, Ule J. 'Read-through marking' reveals differential nucleotide composition of read-through and truncated cDNAs in iCLIP. *Wellcome Open Res*. 2018;3:77.
- [97] Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, et al. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol*. 2020;38:276-8.
- [98] Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL. The Vienna RNA websuite. *Nucleic Acids Res*. 2008;36:W70-4.
- [99] Sahakyan AB, Chambers VS, Marsico G, Santner T, Di Antonio M, Balasubramanian S. Machine learning model for sequence-driven DNA G-quadruplex formation. *Sci Rep*. 2017;7:14535.
- [100] Vondruskova J, Kypr J, Kejnovska I, Fialova M, Vorlickova M. Role of loops in the guanine quadruplex formation by DNA/RNA hybrid analogs of G4T4G4. *Int J Biol Macromol*. 2008;43:463-7.
- [101] Martadinata H, Phan AT. Structure of human telomeric RNA (TERRA): stacking of two G-quadruplex blocks in K(+) solution. *Biochemistry*. 2013;52:2176-83.
- [102] Koralewska N, Szczepanska A, Ciechanowska K, Wojnicka M, Pokornowska M, Milewski MC, et al. RNA and DNA G-quadruplexes bind to human dicer and inhibit its activity. *Cell Mol Life Sci*. 2021;78:3709-24.
- [103] Qi J, Shafer RH. Human telomere quadruplex: refolding and selection of individual conformers via RNA/DNA chimeric editing. *Biochemistry*. 2007;46:7599-606.
- [104] Haider S, Parkinson GN, Neidle S. Crystal structure of the potassium form of an *Oxytricha nova* G-quadruplex. *J Mol Biol*. 2002;320:189-200.
- [105] Chebly A, Ropio J, Baldasseroni L, Prochazkova-Carlotti M, Idrissi Y, Ferrer J, et al. Telomeric Repeat-Containing RNA (TERRA): A Review of the Literature and First Assessment in Cutaneous T-Cell Lymphomas. *Genes (Basel)*. 2022;13.
- [106] Fay MM, Lyons SM, Ivanov P. RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. *J Mol Biol*. 2017;429:2127-47.
- [107] Yang SY, Lejault P, Chevrier S, Boidot R, Robertson AG, Wong JMY, et al. Transcriptome-wide identification of transient RNA G-quadruplexes in human cells. *Nat Commun*. 2018;9:4730.
- [108] Kiliszek A, Blaszczyk L, Kierzek R, Rypniewski W. Stabilization of RNA hairpins using non-nucleotide linkers and circularization. *Nucleic Acids Res*. 2017;45:e92.
- [109] Tseng TY, Chien CH, Chu JF, Huang WC, Lin MY, Chang CC, et al. Fluorescent probe for visualizing guanine-quadruplex DNA by fluorescence lifetime imaging microscopy. *J Biomed Opt*. 2013;18:101309.
- [110] Ye X, Paroo Z, Liu Q. Functional anatomy of the *Drosophila* microRNA-generating enzyme. *J Biol Chem*. 2007;282:28373-8.
- [111] Welker NC, Maity TS, Ye X, Aruscavage PJ, Krauchuk AA, Liu Q, et al. Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. *Mol Cell*. 2011;41:589-99.
- [112] Michlewski G, Caceres JF. Post-transcriptional control of miRNA biogenesis. *RNA*. 2019;25:1-16.

- [113] Kwok CK, Merrick CJ. G-Quadruplexes: Prediction, Characterization, and Biological Application. *Trends Biotechnol.* 2017;35:997-1013.
- [114] Henderson E, Hardin CC, Walk SK, Tinoco I, Jr., Blackburn EH. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell.* 1987;51:899-908.
- [115] Hansel-Hertsch R, Di Antonio M, Balasubramanian S. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat Rev Mol Cell Biol.* 2017;18:279-84.
- [116] Lansdorp PM. Telomeres and disease. *EMBO J.* 2009;28:2532-40.
- [117] Di Prospero NA, Fischbeck KH. Therapeutics development for triplet repeat expansion diseases. *Nat Rev Genet.* 2005;6:756-65.
- [118] Crawford DC, Acuna JM, Sherman SL. FMR1 and the fragile X syndrome: human genome epidemiology review. *Genet Med.* 2001;3:359-71.
- [119] Mirihana Arachchilage G, Dassanayake AC, Basu S. A potassium ion-dependent RNA structural switch regulates human pre-miRNA 92b maturation. *Chem Biol.* 2015;22:262-72.
- [120] Pandey S, Agarwala P, Jayaraj GG, Gargallo R, Maiti S. The RNA Stem-Loop to G-Quadruplex Equilibrium Controls Mature MicroRNA Production inside the Cell. *Biochemistry.* 2015;54:7067-78.
- [121] Rouleau SG, Garant JM, Bolduc F, Bisailon M, Perreault JP. G-Quadruplexes influence pri-microRNA processing. *RNA Biol.* 2018;15:198-206.
- [122] Kurzynska-Kokorniak A, Koralewska N, Pokornowska M, Urbanowicz A, Tworak A, Mickiewicz A, et al. The many faces of Dicer: the complexity of the mechanisms regulating Dicer gene expression and enzyme activities. *Nucleic Acids Res.* 2015;43:4365-80.
- [123] Cole C, Sobala A, Lu C, Thatcher SR, Bowman A, Brown JW, et al. Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA.* 2009;15:2147-60.
- [124] Francia S, Micheli F, Saxena A, Tang D, de Hoon M, Anelli V, et al. Site-specific DICER and DROSHA RNA products control the DNA-damage response. *Nature.* 2012;488:231-5.
- [125] Kaneko H, Dridi S, Tarallo V, Gelfand BD, Fowler BJ, Cho WG, et al. DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. *Nature.* 2011;471:325-30.
- [126] Nakagawa A, Shi Y, Kage-Nakadai E, Mitani S, Xue D. Caspase-dependent conversion of Dicer ribonuclease into a death-promoting deoxyribonuclease. *Science.* 2010;328:327-34.
- [127] Andersson MG, Haasnoot PC, Xu N, Berenjian S, Berkhout B, Akusjarvi G. Suppression of RNA interference by adenovirus virus-associated RNA. *J Virol.* 2005;79:9556-65.
- [128] Ishiguro A, Kimura N, Watanabe Y, Watanabe S, Ishihama A. TDP-43 binds and transports G-quadruplex-containing mRNAs into neurites for local translation. *Genes Cells.* 2016;21:466-81.
- [129] Stefanovic S, Bassell GJ, Mihailescu MR. G quadruplex RNA structures in PSD-95 mRNA: potential regulators of miR-125a seed binding site accessibility. *RNA.* 2015;21:48-60.

- [130] DeMarco B, Stefanovic S, Williams A, Moss KR, Anderson BR, Bassell GJ, et al. FMRP - G-quadruplex mRNA - miR-125a interactions: Implications for miR-125a mediated translation regulation of PSD-95 mRNA. PLoS One. 2019;14:e0217275.
- [131] Zhang DH, Fujimoto T, Saxena S, Yu HQ, Miyoshi D, Sugimoto N. Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factors. Biochemistry. 2010;49:4554-63.

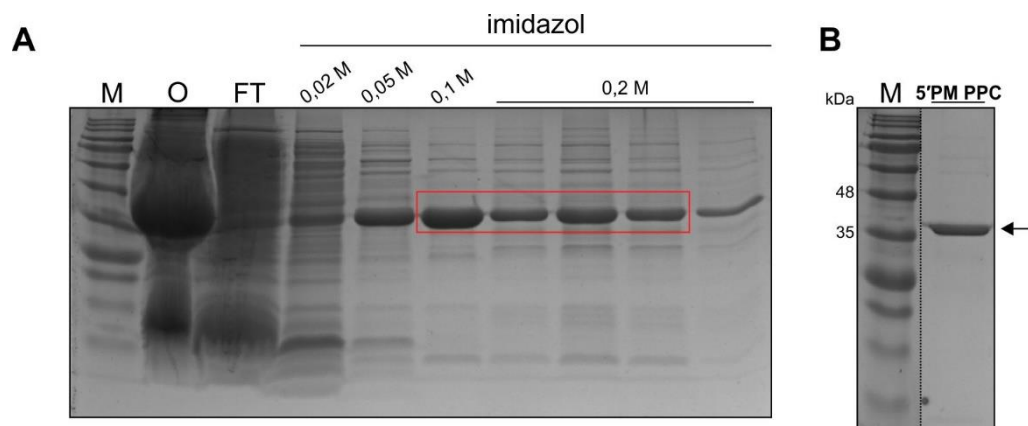
**Cytowana rozprawa doktorska:**

- A. Rozprawa doktorska dr Natalii Koralewskiej, pt. „Krótkie cząsteczki kwasów nukleinowych jako regulatory procesu uwalniania mikroRNA z ich prekursorów”.  
[https://www.ibch.poznan.pl/files/biblioteka/rozprawy\\_doktorskie/Koralewska%20Natalia%20Rozprawa%20doktorska.pdf](https://www.ibch.poznan.pl/files/biblioteka/rozprawy_doktorskie/Koralewska%20Natalia%20Rozprawa%20doktorska.pdf)

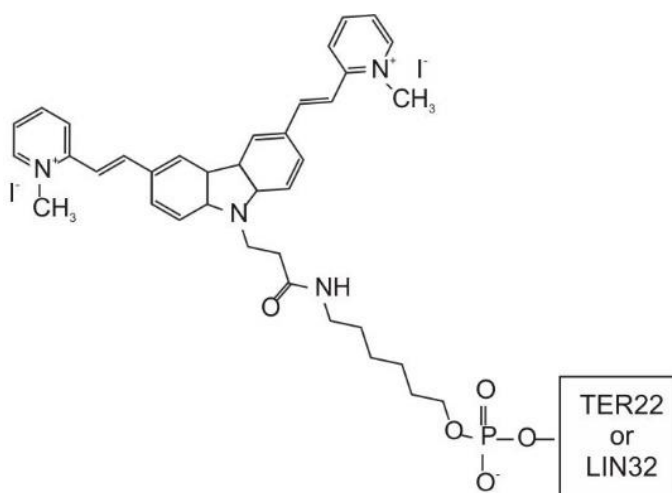
## 8. Suplementy

**Tabela S1. Tabela zawierająca podstawowe informacje dotyczące oligomerów o strukturze G-kwadrupleksów wykorzystanych w analizach.**

Nazwa	Sekwencja	Identyfikator PDB	Typ	Topologia	Referencja
TER10	r(GGGUUAGG GU)	2M18	dimer bimolekularnych G-kwadrupleksów	równoległa	[101]
TER22	r[AGGG(UUA GGG) <sub>3</sub> ]	-	monomolekularny	równoległa	[131]
G4U4G4	r(GGGUUUU GGGG)	-	bi- lub tetramolekularny	równoległa	[100]
TEL22	d[AGGG(TTA GGG) <sub>3</sub> ]	-	monomolekularny	mieszanina topologii równoległej i antyrównoległej	[103]
G4T4G4	d(GGGGTTTT GGGG)	1JPQ	bimolekularny	antyrównoległa	[104]



**Suplement S1. Zdjęcie przedstawiające rozdzielanie elektroforetyczne frakcji otrzymanych w procesie oczyszczania preparatu 5'PM PPC.** (A) Rozdział frakcji uzyskiwanych w trakcie izolacji i oczyszczania preparatu 5'PM PPC. Na żel nałożono 5  $\mu$ l markera Perfect™ Tricolor Protein Ladder (EURx) (M) oraz po 15  $\mu$ l każdej z frakcji: (O) osad komórkowy po lizie, (FT) frakcja zebrana po przepuszczeniu lizatu przez złożę, frakcje uzyskane po nałożeniu na kolumnę roztworów imidazolu o wzrastającym stężeniu. Czerwoną ramką zaznaczono frakcje wybrane do dalszych analiz. (B) Preparat białkowy po oczyszczaniu i zagęszczeniu białka na filtrze Amicon. Na żel nałożono 5  $\mu$ l markera Perfect™ Tricolor Protein Ladder (EURx) (M) oraz 1  $\mu$ l preparatu 5'PM PPC. Strzałką zaznaczono preparat 5'PM PPC. Rozdział prowadzono w 10% żelu PAA.



**Suplement S2. Schematyczne przedstawienie struktury oligomerów wyznakowanych sondą o-BMVC.** Rysunek przygotowany na podstawie publikacji Koralewska&Szczepanska i in. [102].

## 9. Załączniki

**Publikacje naukowe, które dotyczą zagadnień przedstawionych w rozprawie doktorskiej:**

1. **Szczepanska A.**, Wojnicka M. & Kurzynska-Kokorniak A. (2021). The Significance of the DUF283 Domain for the Activity of Human Ribonuclease Dicer. *International Journal of Molecular Sciences*, 13;22(16):8690. IF<sub>2021</sub> 6,208.
2. Koralewska N.\*, **Szczepanska A.\***, Ciechanowska K., Wojnicka M., Pokornowska M., Milewski C.M., Gudanis D., Baranowski D., Nithin C., Bujnicki M.J., Gdaniec Z., Figlerowicz M. & Kurzynska-Kokorniak A. (2021). RNA and DNA G-quadruplexes bind to human dicer and inhibit its activity. *Cellular and Molecular Life Sciences*, 78:3709–3724. IF<sub>2021</sub> 9,207.

(\*) pierwszy współautor





Article

# The Significance of the DUF283 Domain for the Activity of Human Ribonuclease Dicer

Agnieszka Szczepanska , Marta Wojnicka and Anna Kurzynska-Kokorniak \*

Department of Ribonucleoprotein Biochemistry, Institute of Bioorganic Chemistry Polish Academy of Sciences, 61-704 Poznan, Poland; aszczepanska@ibch.poznan.pl (A.S.); mwojnicka@ibch.poznan.pl (M.W.)

\* Correspondence: akurzyns@man.poznan.pl; Tel.: +48-61-852-8503 (ext. 1264)

**Abstract:** Dicers are multidomain proteins, usually comprising an amino-terminal putative helicase domain, a DUF283 domain (domain of unknown function), a PAZ domain, two RNase III domains (RNase IIIa and RNase IIIb) and a dsRNA-binding domain. Dicer homologs play an important role in the biogenesis of small regulatory RNAs by cleaving single-stranded precursors adopting stem-loop structures (pre-miRNAs) and double-strand RNAs into short RNA duplexes containing functional microRNAs or small interfering RNAs, respectively. Growing evidence shows that apart from the canonical role, Dicer proteins can serve a number of other functions. For example, results of our previous studies showed that human Dicer (hDicer), presumably through its DUF283 domain, can facilitate hybridization between two complementary RNAs, thus, acting as a nucleic acid annealer. Here, to test this assumption, we prepared a hDicer deletion variant lacking the amino acid residues 625-752 corresponding to the DUF283 domain. The respective 128-amino acid fragment of hDicer was earlier demonstrated to accelerate base-pairing between two complementary RNAs in vitro. We show that the  $\Delta$ DUF(625-752) hDicer variant loses the potential to facilitate RNA-RNA base pairing, which strongly proves our hypothesis about the importance of the DUF283 domain for the RNA-RNA annealing activity of hDicer. Interestingly, the in vitro biochemical characterization of the obtained deletion variant reveals that it displays different RNA cleavage properties depending on the pre-miRNA substrate.

**Keywords:** Dicer; RNase III; nucleic acid annealers; RNA-binding proteins; RNA-RNA base pairing; RNA cleavage activity



**Citation:** Szczepanska, A.; Wojnicka, M.; Kurzynska-Kokorniak, A. The Significance of the DUF283 Domain for the Activity of Human Ribonuclease Dicer. *Int. J. Mol. Sci.* **2021**, *22*, 8690. <https://doi.org/10.3390/ijms22168690>

Academic Editor: Alexandre Smirnov

Received: 5 July 2021

Accepted: 10 August 2021

Published: 13 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Dicer belongs to the ribonuclease III (RNase III) family of double stranded RNA (dsRNA)-specific endoribonucleases [1]. Dicer proteins generate short RNA duplexes containing functional microRNAs (miRNAs) or small interfering RNAs (siRNAs) [2], which are loaded into a multi-protein complex referred to as the RNA-induced silencing complex (RISC) [3]. During RISC activation, one strand of the RNA duplex (called the “passenger” strand) is released and degraded and the other strand (called the “guide” strand) remains in the complex and acts as a sequence-specific probe guiding RISC to complementary transcripts [4]. Depending on the degree of complementarity between the small RNA (miRNA or siRNA) and the targeted transcript, RISC binding results in either translational repression or mRNA cleavage and degradation [5]. The vast majority of cellular processes are controlled by miRNAs or siRNAs, e.g., developmental timing, growth control, differentiation [6], apoptosis, chromatin rearrangements [7] and even viral defense [8]. Consequently, a disruption of Dicer activity can initiate pathological processes like carcinogenesis, neurodegenerative, rheumatic or immune system disorders [1,9–11].

Human Dicer (hDicer) contains 1992 amino acid residues (220-kDa) and is one of the most structurally complex members of the RNase III family. It comprises an amino (N)-terminal helicase domain, a domain of unknown function (DUF283), Platform, Piwi-Argonaute-Zwille (PAZ) domain, a Connector helix, two RNase III domains (RNase IIIa

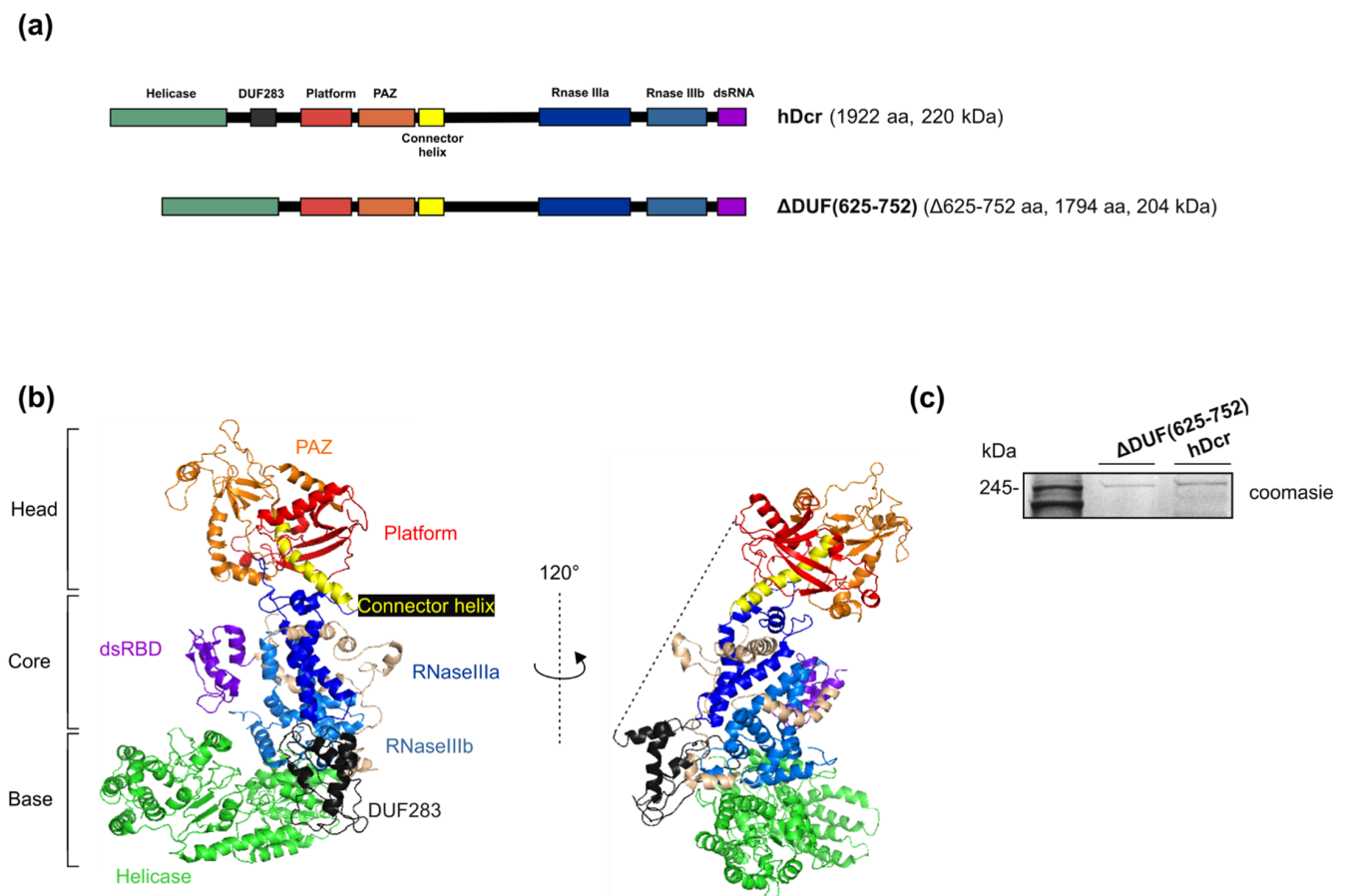
and RNase IIIb) and a dsRNA-binding domain (dsRBD) [12–14]. Extensive studies have deciphered the roles of the Dicer domains in binding and processing of its canonical substrates, i.e., single-stranded hairpin precursors of miRNAs (pre-miRNAs) and dsRNAs. The helicase domain through selective interactions with the apical loop of pre-miRNAs can discriminate between substrates [15,16]. The DUF283 domain can bind single-stranded nucleic acids in vitro [17]; thus, it is possible that this domain may be involved in interactions with the apical loop of pre-miRNA hairpins as well [18]. The Platform and PAZ domains anchor the 5' phosphate and 2-nt 3' overhang of a substrate [19]. The RNase IIIa and RNase IIIb domains form a single dsRNA-cleavage center. Each of the domains cleaves one strand within the RNA duplex structure, generating products with 2-nt 3'-overhangs [14]. Finally, the carboxy (C)-terminal dsRBD is presumed to play an auxiliary role in RNA substrate binding [20].

The function of DUF283 is still not well understood. Structural studies have revealed that DUF283 adopts a fold typical for proteins that bind dsRNA [21,22]. However, in vitro studies have shown only trace dsRNA-binding activity of the DUF283 domain from *Arabidopsis thaliana* DCL4 [21] and have not demonstrated even weak dsRNA-binding activity for DUF283 of hDicer [17]. Instead, DUF283 was found to be responsible for interacting with Dicer's protein partners in both mammals [23] and plants [21]. Additionally, results of our previous in vitro studies have revealed that the DUF283 domain of hDicer is capable of binding single-stranded nucleic acids and can accelerate base pairing between two complementary RNAs or DNAs [17]. Likewise, full-length hDicer can facilitate hybridization between short RNAs and their targets [17,24]. We hypothesize that the DUF283 domain is crucial for the RNA-RNA annealing activity of hDicer. Consequently, to test this assumption, here we prepared a hDicer deletion variant lacking the amino acid residues 625-752 corresponding to the DUF283 domain. We showed that the  $\Delta$ DUF(625-752) hDicer variant loses the potential to facilitate hybridization between two complementary RNA molecules, which strongly supports our hypothesis about the importance of the DUF283 domain for the RNA-RNA annealing activity of hDicer. Moreover, the in vitro biochemical characterization of the obtained deletion variant demonstrated that it displays different RNA cleavage properties depending on the pre-miRNA substrate. Interestingly, the in cellulo studies did not show major differences in the cleavage activity of the  $\Delta$ DUF(625-752) variant with respect to the investigated substrates, but suggested decreased cleavage potential of this variant in comparison with the wild-type protein.

## 2. Results

### 2.1. Production of the hDicer DUF283 Deletion Variant

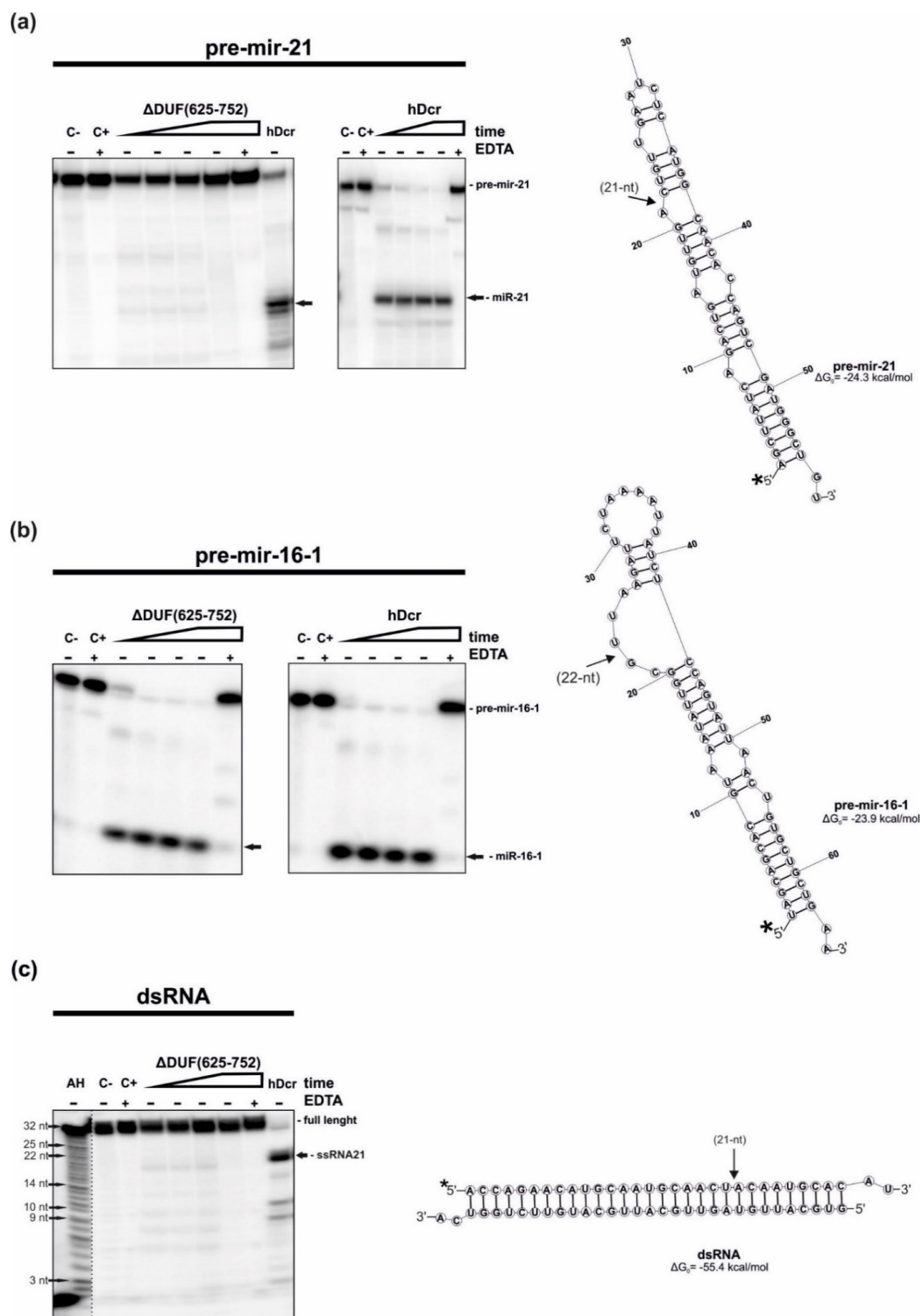
To investigate whether the DUF283 domain is indeed indispensable for hDicer RNA-RNA base pairing activity, we obtained the hDicer variant lacking the amino acid residues 625-752 corresponding to the DUF283 domain (a variant named  $\Delta$ DUF(625-752)) (Figure 1a). This 128-amino acid fragment of hDicer was earlier demonstrated to accelerate base-pairing between two complementary RNA or DNA molecules in vitro [17]. The expression plasmid encoding the hDicer  $\Delta$ DUF(625-752) variant was produced by a PCR approach using as a template the plasmid encoding the wild-type full-length hDicer fused with the 3xFlag-tag (called "hDcr") that was earlier obtained in our laboratory [25]) (Figure 1a,b) (for details please see Materials and Methods). The expression plasmid was used for the transfection of 293T NoDice cells (human cells that do not produce hDicer) [26]. A total of 72 h after transfection, the protein was isolated and purified from cell extracts by immunoprecipitation with anti-Flag antibody conjugated to agarose beads (Figure 1c).



**Figure 1.** Full-length hDicer and its mutant lacking the DUF283 domain. (a) Schematic representation of hDicer domain architecture and the DUF283 domain deletion variant:  $\Delta$ DUF(625-752). (b) The 3D structure of hDicer (PDB 5ZAL) visualized by PyMOL. The fragment removed in the  $\Delta$ DUF(625-752) variant, i.e., DUF283 and the linker (missing in 5ZAL and drafted by the dashed line), is indicated in dark grey. (c) PAGE analysis of  $\Delta$ DUF(625-752) and hDcr preparations. The C-terminally 3xFlag-tagged proteins were expressed in 293T NoDice cells, then they were purified by immunoprecipitation and analyzed by SDS-PAGE followed by Coomassie Blue Staining.

## 2.2. RNase Activity of the hDicer DUF283 Deletion Variant

First, we tested the RNase activity of the  $\Delta$ DUF(625-752) variant. The variant was assayed for the RNase activity using two pre-miRNA substrates: pre-mir-21 and pre-mir-16-1 and a 30-base pair (bp) RNA duplex having 2-nt 3'-overhanging ends (called "dsRNA"). The cleavage assays involved 18 nM of the protein and ~5 nM of either 5'-<sup>32</sup>P-labeled pre-mir-21, pre-mir-16-1 or dsRNA in which one of the two strands was 5'-<sup>32</sup>P-labeled. Thus, the reactions were performed under the low-turnover conditions, i.e., more than threefold molar excess of hDcr over a substrate was used. Two control reactions without the protein were also prepared: one containing only the substrate in the reaction buffer (C-) and the other containing the substrate in the reaction buffer with the addition of the Mg<sup>2+</sup>-chelating agent, EDTA (25 mM), (C+). As the Dicer cleavage activity is dependent on Mg<sup>2+</sup>, EDTA would abrogate this activity of Dicer. Thus, another control reactions included the substrate, protein and 25 mM EDTA (+EDTA). Yet another control contained the substrate and 18 nM hDcr. All reactions were carried out at 37 °C and then they were halted by adding 1 volume of 7 M UREA loading buffer and heating for 5 min at 95 °C. Reaction mixtures were separated on a 15% polyacrylamide gel with 7 M urea and 1xTBE (PAGE) and visualized by phosphorimaging (Figure 2). The cleavage assays were conducted in triplicate.



**Figure 2.** RNase activity assay of hDcr and  $\Delta$ DUF(625-752). (a–c) The results of the RNA-cleavage assays involving: (a) pre-mir-21, (b) pre-mir-16-1, (c) 30-bp dsRNA with 2-nt 3'-overhangs. Reaction mixtures were incubated for: 10, 30, 60, 120 min (increasing time is represented by a triangle) with 18 nM of either  $\Delta$ DUF(625-752) or hDcr. (C-) controls containing only the substrate in the reaction buffer. (C+) controls incubated without a protein but with 25 mM EDTA. (+EDTA) supplementation of the reaction buffer with 25 mM EDTA. (AH) Ladder generated by alkaline hydrolysis of 5'-<sup>32</sup>P-labeled 21-nt RNA. Dicer cleavage sites are indicated by arrows. \* The asterisk indicates the <sup>32</sup>P 5'-end label. The reproducible results were obtained using at least two batches of recombinant  $\Delta$ DUF(625-752) and hDcr.

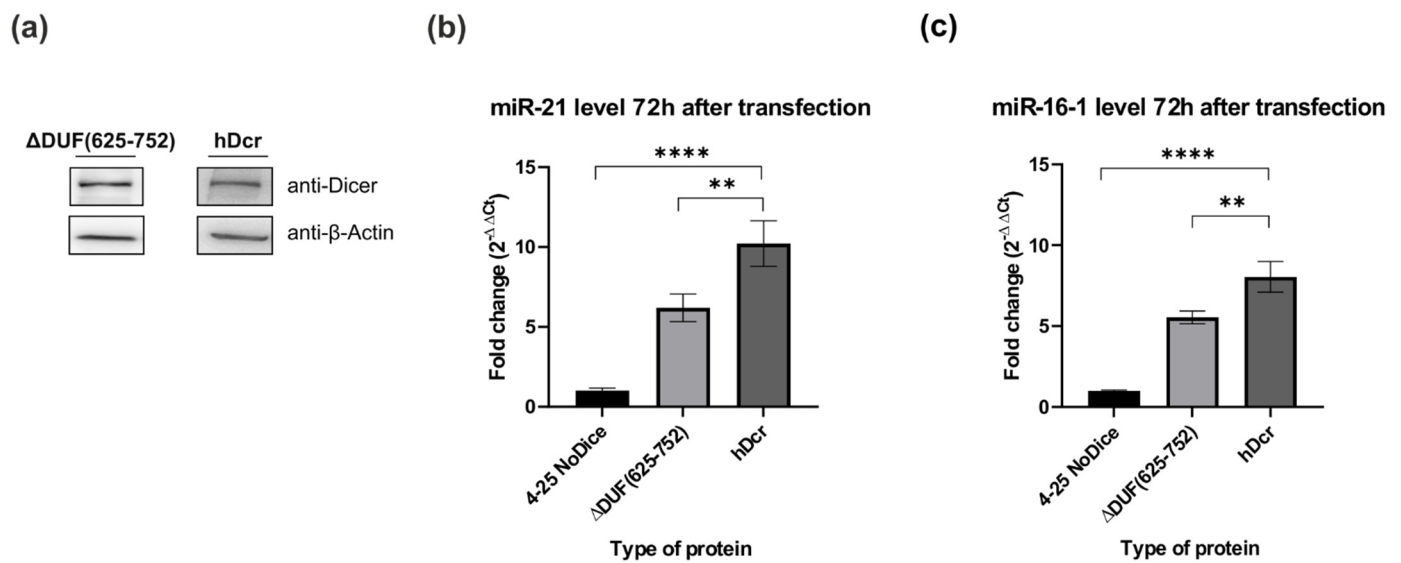
We found that in the assay with pre-mir-21,  $\Delta$ DUF(625-752) generated only traces of 21-nt miRNA products (faint bands observed at the position of miR-21 generated by hDcr in the control experiment). Under the same reaction conditions, within 2 h incubation, hDcr processed almost all of the substrate (Figure 2a). In the case of the reactions carried out with the pre-mir-16-1 substrate, both  $\Delta$ DUF(625-752) and hDcr produced efficiently miR-16-1. After 2 h incubation time, all pre-mir-16-1 was cut by both enzymes (Figure 2b). For the cleavage assay involving 30-bp dsRNA and  $\Delta$ DUF(625-752), we noticed faint bands migrating as fast as 21-nt RNA that was generated efficiently by hDcr in the control experiment (Figure 2c). Collected data revealed that the deletion of amino acids 625-752 affects hDicer's ability to cleave pre-mir-21 and 30-bp dsRNA, but not pre-mir-16-1, which suggests that this fragment of hDicer might be involved in substrate recognition and discrimination.

In addition, we performed the RNA cleavage assays involving the hDicer DUF283 deletion variant lacking the region spanning 630-709 amino acid residues ( $\Delta$ DUF(630-709)). This variant was earlier produced and characterized by Ma E. and colleagues [27]. We found that under the same reaction conditions as those for  $\Delta$ DUF(625-752), the  $\Delta$ DUF(630-709) variant processed pre-mir-21, pre-mir-16-1 and 30-bp dsRNA as efficiently as the wild-type enzyme, hDcr (Figure S1).

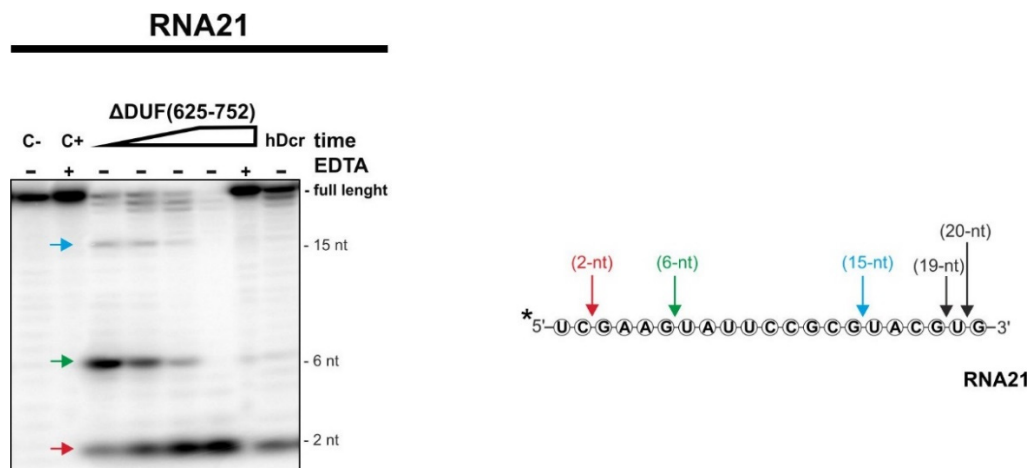
The aforementioned experiments were carried out under the *in vitro* conditions, consequently we asked the question about the ability of the  $\Delta$ DUF(625-752) variant to process pre-mir-21 and pre-mir-16-1 under *in cellulo* conditions. By applying the RT-qPCR approach, we assessed the relative level of miR-21-5p and miR-16-1-5p produced in 293T NoDice cells expressing the  $\Delta$ DUF(625-752) variant. As a negative control, we used 293T NoDice cells treated only with a transfection reagent and as a positive control, we used 293T NoDice cells transfected with the plasmid expressing the wild-type full-length hDicer, hDcr. All cells were harvested 72 h after transfection. We found that, under similar protein expression levels (Figure 3a), both miR-21-5p and miR-16-1-5p were less abundant in 293T NoDice cells expressing  $\Delta$ DUF(625-752), than in 293T NoDice cells expressing hDcr (Figure 3b,c). Precisely, a reduction in miRNA level was by about 40% (for miR-21-5p) and by about 30% (for miR-16-1-5p) compared to the respective positive control reactions. Thus, we found that miR-21-5p and miR-16-1-5p were produced with different efficiency, depending on the used system: *in vitro* (Figure 2a,b) or *in cellulo* (Figure 3b,c).

hDicer, apart from the canonical substrates, can bind and cut miRNA-size RNA molecules [24,25,28–30]; therefore, we assume that Dicer proteins might be involved in the miRNA turnover [25]. Accordingly, we performed cleavage assays with a 21-nt RNA substrate, called "RNA21". We found that  $\Delta$ DUF(625-752) generated 15-nt, 6-nt and 2-nt RNA products (Figure 4), as it was earlier observed for hDcr and the hDicer variants lacking the PAZ domain [25]. With incubation time, we observed disappearance of 15-nt and 6-nt RNA products and accumulation of 2-nt RNAs. Based on these results, we conclude that 2-nt RNAs were cut off from 15-nt and/or 6-nt RNAs. Consequently, 15-nt and 6-nt RNAs can be defined as primary cleavage products, while 2-nt RNAs, as secondary cleavage products. We also noticed that, under the adequate reaction conditions, processing of RNA21 was much more efficient in the case of  $\Delta$ DUF(625-752), compared to hDcr.





**Figure 3.** Comparison of miRNA levels produced in NoDice cells expressing  $\Delta$ DUF(625-752) and hDcr. (a) Western blot analysis of  $\Delta$ DUF(625-752) and hDcr expressed in 293T NoDice cells. Cells were harvested and analyzed by Western blotting with anti-Dicer and anti- $\beta$ -Actin antibodies 72 h after transfection with the respective expression plasmids. (b,c) Results of the RT-qPCR analysis of the level of endogenous, mature miR-21 (b) and miR-16-1 (c) in the 293T NoDice cells (a negative control), or 293T NoDice cells expressing either  $\Delta$ DUF(625-752) or hDcr. 72 h after transfection, cells were harvested and total RNA was isolated. miRNA levels were determined by RT-qPCR and normalized to the expression of U6 snRNA, as a reference gene. The values are averaged from three biological replicates as mean  $\pm$  S.D. Data were analyzed by one-way ANNOVA test followed by Dunnett's multiple comparisons test. The obtained  $p$  values are as follows, for miR-16-1: \*\*  $p < 0.0021$ , \*\*\*\*  $p < 0.0001$ ; for miR-21: \*\*  $p < 0.0040$ , \*\*\*\*  $p < 0.0001$ .

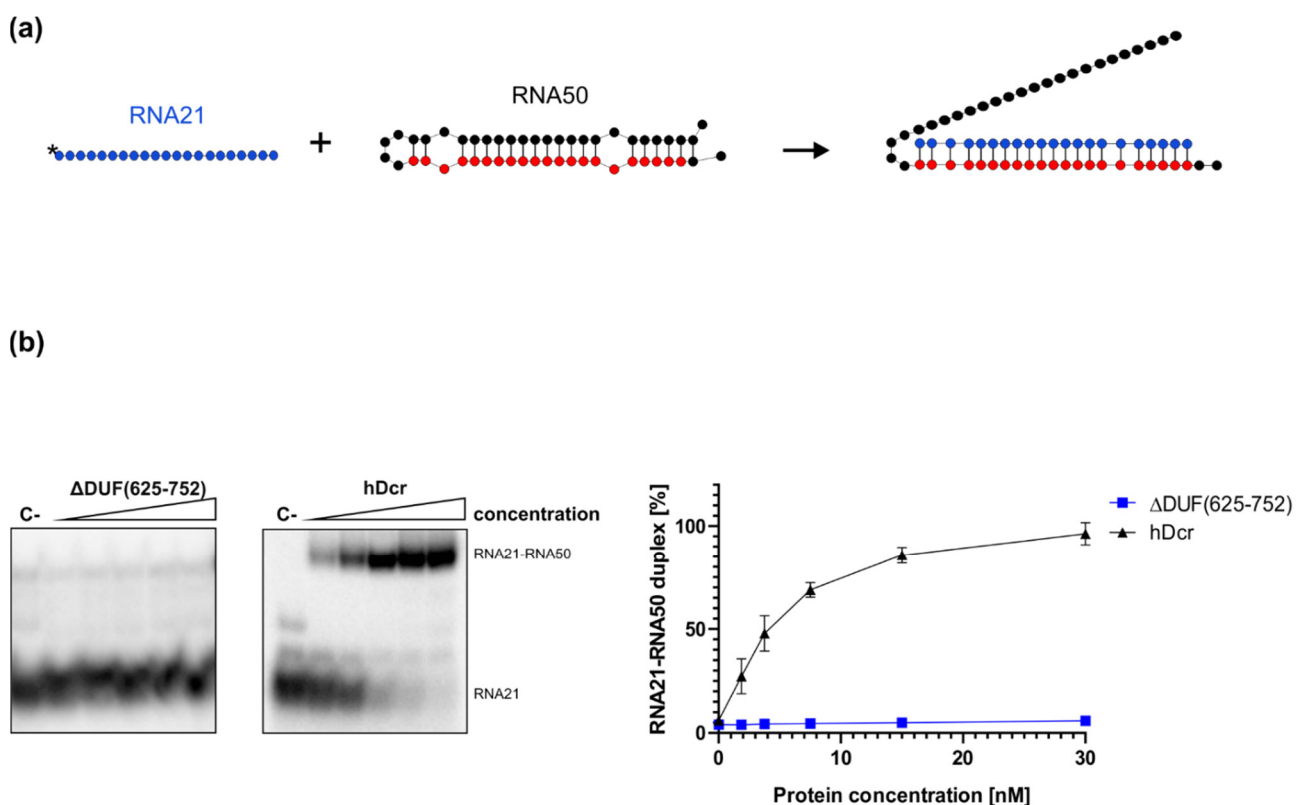


**Figure 4.** The results of the RNA-cleavage assays involving 21-nt RNA. Reaction mixtures were incubated for 10, 30, 60, 120 min for  $\Delta$ DUF(625-752) (increasing time is represented by a triangle) with 18 nM of protein. (C-) control containing only the substrate in the reaction buffer. (C+) controls incubated without a protein but with 25 mM EDTA. Additional control reaction included the substrate and 18 nM hDicer protein (hDcr) incubated for 120 min. (+EDTA) supplementation of the reaction buffer with 25 mM EDTA. The asterisk indicates the  $^{32}P$  5'-end label. The reproducible results were obtained using at least two batches of recombinant  $\Delta$ DUF(625-752) and hDcr.

### 2.3. The RNA-RNA Base Pairing Potential of the hDicer DUF283 Domain Deletion Variant

To investigate how removing the DUF283 domain from hDicer influences its RNA-RNA base pairing activity, we carried out the annealing assay involving a pair of complementary RNAs: 21-nt RNA (RNA21) and a 50-nt RNA (called "RNA50") adopting a hairpin structure and containing the fully complementary RNA21 target site, schematically

described in Figure 5a. This pair of complementary RNAs was used in our earlier annealing assays [24]. Before the reaction, RNA50 was incubated at 95 °C for 5 min and then slowly cooled to room temperature to ensure proper folding. Then, RNA50 was mixed in annealing buffer with 5'-<sup>32</sup>P-labeled RNA21 at a molar ratio of approximately 1:1 and incubated for 30 min with increasing amounts of  $\Delta$ DUF(625-752) (1.88, 3.75, 7.5, 15, 30 nM) at 37 °C (Figure 5b). The same set of reactions was carried out for hDcr. Spontaneous annealing was determined by excluding the enzyme in the assay mixture (control reaction (C-)). Based on the results obtained from three independent experiments, for each reaction we calculated the percentage ratios between the fraction containing the RNA21-RNA50 duplex and the free RNA21 fraction. The average percentage content of the RNA21-RNA50 duplex was plotted against the protein concentration (Figure 5b, right panel). The collected results demonstrated that, in contrast to hDcr,  $\Delta$ DUF(625-752) did not support base pairing between RNA21 and RNA50 under the applied reaction conditions. Neither we observed the RNA-RNA base pairing activity for the  $\Delta$ DUF(630-709) variant (Figure S2). Altogether, collected data strongly support our hypothesis that the DUF283 domain is indispensable for hDicer RNA-RNA annealing activity.



**Figure 5.** Annealing activity assay of  $\Delta$ DUF(625-752) and hDcr. (a) Schematic representation of the experiment. RNA21 (in blue) is fully complementary to the 21-nt target sequence (in red) of RNA50. (b) Native PAGE gel showing the results of annealing reactions carried out for 30 min and involving 1.88, 3.75, 7.5, 15, 30 nM of either  $\Delta$ DUF(625-752) or hDcr. (C-) a control incubated without the protein. The reactions were resolved in buffer containing SDS at final concentration of 0.2%. Graph presentation of the results obtained from three independent annealing assays. Error bars represent S.D. from three separate experiments.

### 3. Discussion

Initially, it was suggested that the DUF283 domain is critical for pre-miRNA processing, because the *in vitro* cleavage activity is lost in human and *Drosophila melanogaster* Dicer variants lacking both DUF283 and the helicase domain [31,32]. However, subsequent *in vitro* studies demonstrated that the deletion of only DUF283 decreases the cleavage of dsRNAs by hDicer, without affecting the cleavage of pre-miRNA substrates [27]. The

mentioned DUF283 deletion variant, produced by Ma E. and colleagues, lacks the region spanning 630-709 amino acid residues (the  $\Delta$ DUF(630-709) variant), (Figure S1) [27]. The DUF283 deletion variant produced in our laboratory lacks the region between amino acids 625 and 752, including a linker joining the DUF283 domain to Platform (the  $\Delta$ DUF(625-752) variant), (Figure 1a,b). Although removing this linker might affect the spatial organization of hDicer domains, we found that the  $\Delta$ DUF(625-752) variant is successfully expressed in NoDice cells (Figure 1c). After purification, this protein retained its integrity and activity up to 12 months when stored at  $-20^{\circ}\text{C}$ . The three-dimensional structure of Dicer resembles the letter L [18,33]. Within this structure, the head, core and base can be distinguished. The head constitutes the PAZ and the Platform domains, the RNase III domains are in the core, whereas the helicase domain forms the base (Figure 1b). We can deduce that removing the linker, that joins the base to the head, precludes a proper positioning of the base (i.e., the helicase domain) in relation to the head (i.e., PAZ and Platform). Consequently, we assume that the orientation of the helicase domain differs between the  $\Delta$ DUF(625-752) variant and the wild-type hDicer or the  $\Delta$ DUF(630-709) variant. However, the helicase domain was shown to be dispensable for processing of both pre-miRNA and dsRNA substrates by hDicer [27]. Thus, mispositioning of the helicase domain should not abolish the cleavage activity of the  $\Delta$ DUF(625-752) variant. Indeed, we found that, under the in vitro conditions, the  $\Delta$ DUF(625-752) variant processed pre-mir-16-1 as efficiently as the wild-type enzyme (Figure 2b). In contrast to the case of pre-mir-16-1, the deletion of amino acids 625-752 compromised hDicer ability to cleave pre-mir-21 and 30-bp dsRNA in vitro (Figure 2a,c). What are the differences between pre-mir-16-1 and pre-mir-21 substrates? We can notice that pre-mir-16-1 contains large internal loops and bulges and has relaxed terminal loop region, whereas pre-mir-21 adopts more compact structure, with the small terminal loop (Figure 2). Therefore, taking into account the content of paired nucleotides, the structure of pre-mir-21 is more similar to the structure of the dsRNA substrate, rather than to the structure of pre-mir-16-1. Single-stranded regions of pre-mir-16-1 provide the greater flexibility of this substrate, compared to the other two substrates. Accordingly, we assume that pre-mir-16-1, due its flexibility, may better fit into the cleavage center of the  $\Delta$ DUF(625-752) variant. In contrast to  $\Delta$ DUF(625-752), the  $\Delta$ DUF(630-709) variant processed all tested canonical substrates at least as efficiently as the wild-type enzyme (Figure S1). These data further support our hypothesis that the spatial orientation of the domains differs between  $\Delta$ DUF(625-752) and  $\Delta$ DUF(630-709) variants. Additionally, the collected data suggest that not only the helicase domain [15], but the DUF283 domain may as well play an important role in substrate recognition and discrimination. Importantly, as mentioned above, the removal of both the helicase and the DUF283 domains was shown to abolish in vitro processing of pre-miRNA substrates by the respectively truncated hDicer variant [31]. Collectively, based on the literature data and results of the cleavage assays generated for  $\Delta$ DUF(625-752) and  $\Delta$ DUF(630-709), we hypothesize that factors influencing the spatial orientation of hDicer domains may impact the enzyme's substrate specificity and its cleavage properties.

Nevertheless, in contrast to the in vitro RNase cleavage assays (Figure 2), we found that in NoDice cells expressing  $\Delta$ DUF(625-752), the levels of miR-21-5p and miR-16-1-5p were similarly decreased (by about 40% and 30%, respectively) compared to the control cell lines producing hDcr (Figure 3). These results indicate the possible involvement of cellular factors, interacting with pre-miRNAs, in the studied process.

We also conclude that the RNase III cleavage center of the  $\Delta$ DUF(625-752) variant is accessible for miRNA/siRNA-size RNAs (Figure 4). The collected data showed that  $\Delta$ DUF(625-752) generates the same cleavage pattern for 21-nt RNA as the wild-type enzyme and the PAZ deletion variant, which was reported in our previous studies [25]. Accordingly, taking into account the results of this and earlier studies [25], we deduce that the cleavage of miRNA-size RNAs by hDicer neither involves DUF283 nor PAZ domains. As we proposed earlier, we believe that Dicer might be involved in the miRNA turnover by promoting degradation of passenger strands that are discarded from the RISC complex [25].



Finally, we found that the deletion of the DUF283 domain (either the fragment spanning amino acids 625-752 or the fragment spanning amino acids 630-709) abolishes hDicer RNA-RNA annealing activity in vitro (Figure 5) and (Figure S2). Altogether the presented data provide further insight into the role of the DUF283 domain for the activity of human ribonuclease Dicer. The next challenge will be to translate all these results into the in vivo activity of Dicer-type proteins.

## 4. Materials and Methods

### 4.1. Oligonucleotides

DNA primers and RNA oligonucleotides were purchased from Genomed (Warsaw, Poland) and FutureSynthesis (Poznan, Poland), respectively. Sequences of all oligonucleotides used in this study are listed in Table 1.

**Table 1.** Sequences of oligonucleotides used in the experiments.

Name	Sequence (5' → 3')
RNA_sense *	GUGCAUUGUAGUUGCAUUGCAUGUUCUGUCA
RNA_ov	ACCAGAACAUGCAAUGCAACUACAAUGCACAU
pre-mir-16-1	UAGCAGCACGUAUUUAUUGGCGUUUAGAUUCUAAAAUUUUCUCCAGUAUUUACUGUGCUGCUGAA
pre-mir-21	AGCUUUAUCAGACUGAUGUUGACUGUUGAAUCUCAUGGCAACACCAGUCGAUGGGCUGU
RNA21	<u>UCGAAGUAUUCGCGUACGUG</u>
RNA50	GGUUGAACUAUUUCGUGUAUCUGGAAAC <u>ACGUACGCGGAAUACUUCGAUU</u>
f_ΔDUF(625-752)	ATTCCAGAGTGTTTGAGGGATAGTTATCCAGACCTG
r_ΔDUF(625-752)	TCGTGGACCACCATCGTCAGGCTCAACAC
f_ΔDUF(630-709)	GACCATTGATGCCAGTTGGGAAAGAGACT
r_ΔDUF(630-709)	GGCCGTGTTGATTGTGACTCGTGGACCAC
f.miR-21-5p	GCTTATCAGACTGATGTTGAAA
f.miR-16-1-5p	GCACGTAAATATTGGCGAA

\* RNA\_sense serves as a complementary strand to RNA\_ov to form dsRNA. The complementary sequences between RNA21 and RNA50 are underlined.

### 4.2. <sup>32</sup>P Labeling of Oligonucleotides

The 5'-end labeling of oligonucleotides was performed using 10 pmol of RNA, 1 μL γ-<sup>32</sup>P-ATP (Hartman Analytic GmbH, Braunschweig, Germany) and 10 U T4 polynucleotide kinase (Thermo Fisher Scientific, Waltham, MA, USA) in a final volume of 10 μL. The composition and reaction conditions were in accordance with the manufacturer's instructions. Before preparing the reaction mixture, the RNA was denatured by heating to 95 °C and rapidly cooling on ice. The radiolabeled oligonucleotides were PAGE-purified in 8% denaturing polyacrylamide gels and resuspended in water to a final concentration of approximately 10,000 cpm/μL. The homogeneity of the material after gel purification was checked by electrophoresis in 15% PAA under denaturing conditions. Gels were exposed to a phosphorimager plate, which was subsequently scanned with FLA-5100 Fluorescent Image Analyzer (Fujifilm, Minato, Tokyo, Japan) to visualize the bands.

### 4.3. Preparation of dsRNA

To prepare a dsRNA substrate, non-labeled strand (RNA\_sense) was hybridized, at a molar ratio of approximately 1:1, with <sup>32</sup>P-labeled complementary strand (RNA\_ov) in buffer containing 50 mM NaCl, 2.5 mM MgCl<sub>2</sub> and 20 mM Tris-HCl, pH 7.5, by heating up to 95 °C and then slowly cooling down to room temperature. Next, the reaction mixtures were analyzed on 10% native polyacrylamide gels to check whether the double-stranded complexes were free of single-stranded species.

#### 4.4. Preparation of the $\Delta$ DUF(625-752) and $\Delta$ DUF(630-709) Expression Plasmids

The expression plasmids encoding the  $\Delta$ DUF(625-752) and  $\Delta$ DUF(630-709) variants were prepared using PCR amplification. All primers were designed based on the cDNA encoding transcript variant 2 of human *DICER1* (NM\_030621.4). A template in the PCR amplification was the plasmid (a derivative of SureVector expression vector (Agilent, Santa Clara, CA, USA)) encoding the full-length wild-type hDicer fused at its C-terminus with the 3xFlag peptide, that was earlier prepared in our laboratory [25]. In the reaction we used the following primers: (i) for  $\Delta$ DUF(625-752), f\_ $\Delta$ DUF(625-752) (forward) and r\_ $\Delta$ DUF(625-752) (reverse); (ii) for  $\Delta$ DUF(630-709), f\_ $\Delta$ DUF(630-709) (forward) and r\_ $\Delta$ DUF(630-709) (reverse). The PCR amplification involved PfuUltra II Fusion polymerase (Agilent). The obtained PCR product: a cDNA of the plasmid including the sequence of the full-length wild-type hDicer missing the fragment encoding the DUF283 domain (either the 384-nt sequence corresponding to amino acids 625-752 of hDicer or the 240-nt sequence corresponding to amino acids 630-709 of hDicer), was phosphorylated by T4 polynucleotide kinase (Thermo Fisher Scientific, Waltham, MA, USA) and ligated by T4 DNA Ligase (Thermo Fisher Scientific). The prepared constructs were sequenced by Sanger method (Genomed, Warsaw, Poland). The results of sequencing reactions are presented in Figure S3.

#### 4.5. Cell Culture and Transfection

293T NoDice cells [26] were cultured in DMEM (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 10% FBS (Gibco), Penicillin-Streptomycin (100 U/mL of penicillin and 100  $\mu$ g/mL of streptomycin, Gibco) and 1 mM Sodium Pyruvate (Gibco), as described in Bogerd et al. [26]. Transfection was carried out by DharmaFECT kb DNA Transfection Reagent (Dharmacon, Lafayette, CO, USA), according to the manufacturer's instructions. 293T NoDice cell line was kindly provided by Prof. Bryan R. Cullen.

#### 4.6. Immunoprecipitation

Cells were harvested 72 h after transfection by centrifugation at 1500 rpm for 3 min and suspended in lysis buffer (30 mM Hepes pH 7.4, 100 mM KCl, 5 mM MgCl<sub>2</sub>, 10% glycerol, 0.5 mM DTT and 0.2% Tergitol) containing 1  $\times$  protease inhibitor mix without EDTA (Roche, Warsaw, Poland) and broken by passing through a 0.9 mm  $\times$  40 mm needle. Lysates were centrifuged at 13,000 rpm for 5 min at 4  $^{\circ}$ C. The supernatant was incubated overnight on a rotator with ANTI-FLAG<sup>®</sup>M2 Affinity Gel (Merck, Darmstadt, Germany) that was pre-washed with TBS buffer. After incubation, the beads were washed five times with TBS buffer. Isolated protein was eluted with 100  $\mu$ g/mL 3XFlag Peptide (Sigma, Kawasaki, Japan). Purified protein was suspended in the buffer (20 mM Tris pH 7.5, 50 mM NaCl, 10% glycerol and 0.25% Triton X-100). Protein concentration was estimated by the Bradford assay (BioRad, Irvine, CA, USA), relative to a BSA standard curve and on SDS-PAGE with BSA as a standard. Proteins were concentrated using Amicon filters (Merck) in the respective buffer (as described above) enriched with 40% glycerol and stored at  $-20^{\circ}$ C.

#### 4.7. Western Blot Analysis

Obtained proteins were separated on 8% SDS-PAGE and electro-transferred onto a PVDF membrane (Thermo Fisher Scientific, Waltham, MA, USA). For hDicer and hDicer deletion variant, the blots were probed with a mouse monoclonal primary anti-Dicer antibody mapping at the C-terminus of hDicer (1:300, Santa Cruz Biotechnology, Dallas, TX, USA), for  $\beta$ -Actin, the blots were probed with a rabbit monoclonal primary anti- $\beta$ -Actin antibody (1:1100, Cell Signaling Technology, Danvers, MS, USA) and subsequently with HRP-conjugated secondary antibody, anti-mouse or anti-rabbit (1:5000, Jackson ImmunoResearch Laboratories, Inc., Cambridgeshire, UK). The immunoreactions were detected using SuperSignal<sup>™</sup>West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific, Waltham, MA, USA).

#### 4.8. The RNA Cleavage Assay

The cleavage assay was performed in 10  $\mu$ L reactions containing 50 mM NaCl, 2.5 mM  $MgCl_2$  and 20 mM Tris-HCl, pH 7.5, 5'- $^{32}P$ -labeled substrate (10,000 cpm, approximately 5 nM) and 18 nM of the protein (hDcr or  $\Delta DUF(625-752)$ , or  $\Delta DUF(630-709)$ ). In addition, a reaction mixture without the protein was prepared as a control. In controls including EDTA, the reaction buffer was supplemented with the chelating agent to the final concentration of 25 mM. Reactions were carried out at 37 °C for 10, 30, 60 and 120 min with the addition of the commercial RNase-inhibitor cocktail (NEB, Ipswich, USA). The reactions were stopped by the addition of 1 volume of 7 M urea loading buffer and heating for 5 min at 95 °C. Samples were separated on a 15% denaturing polyacrylamide gel in 1  $\times$  TBE running buffer.

#### 4.9. Reverse Transcription and Quantitative PCR

The RNA was extracted using a standard TRIzol protocol (Invitrogen, Waltham, MA, USA). The RNA quantity and quality were measured using NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA). For reverse transcription (RT), 8  $\mu$ g of total RNA was used. RT was performed by using Mir-X<sup>TM</sup> miRNA First-Strand Synthesis Kit (Takara Bio, Canada, USA) in accordance with the manufacturer's instructions. RT-qPCR was performed using iTaq<sup>TM</sup> Universal SYBR<sup>®</sup> Green Supermix (Bio-Rad, Irvine, CA, USA). For miRNA analysis, 8 ng of total cDNA was used. Amplification was performed using the Bio-Rad CFX96<sup>TM</sup> system (Bio-Rad) and the software determined  $C_t$  thresholds. The relative expression level of miRNA was evaluated by the  $2^{-\Delta\Delta C_t}$  method [34]. miRNA levels were normalized to U6 small nuclear RNA as a reference gene. Control reactions lacking template were performed to verify clean backgrounds in all samples. The results from three independent experiments were presented as mean  $\pm$  S.D. One-way ANNOVA test followed by Dunnett's multiple comparisons test was used for statistical evaluation. Statistical analyses were performed using GraphPad Prism 6 Software (GraphPad Software, San Diego, CA, USA). The primer sequences used for qRT-PCR are listed in Table 1.

#### 4.10. Annealing Assay

The reactions were carried out in 10  $\mu$ L volumes. As substrates, we used the same RNA pair as in our previous annealing assays [24]. Precisely, each reaction set contained 10,000 cpm (approximately 5 nM) of the 5'-end  $^{32}P$ -labeled RNA21 and 5 nM of long complementary RNA (RNA50). The molecules were mixed in annealing buffer (50 mM NaCl, 20 mM Tris-HCl pH 7.5) and incubated for 30 min at 37 °C with dilutions of hDcr or its  $\Delta DUF(625-752)$ , or  $\Delta DUF(630-709)$  variant (1.88, 3.75, 7.5, 15, 30 nM).

#### 4.11. Gel Imaging and Analysis

The data were collected using a Fujifilm FLA-5100 Fluorescent Image Analyzer (Fujifilm, Minato, Tokyo, Japan). The amounts of  $^{32}P$ -labeled substrates and products were determined from the intensity of the respective bands in the gels measured by MultiGauge 3.0 software (Fujifilm, Minato, Tokyo, Japan). The diagrams from annealing experiments were created using GraphPad Prism 6 Software (GraphPad Software, San Diego, CA, USA). In the case of all diagrams, error bars represent S.D. values calculated based on three independent experiments.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/ijms22168690/s1>.

**Author Contributions:** A.S. and M.W. planned and performed experiments; all authors analyzed the data and interpreted the results; A.K.-K. coordinated the research, supervised and provided advice; A.S. and A.K.-K. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Science Centre, Poland (NCN, grant SONATA BIS 2016/22/E/NZ1/00422 to A.K.-K.). Funding for the open access charge was provided by the National Science Centre, Poland (SONATA BIS 2016/22/E/NZ1/00422 to A.K.-K.).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We would like to thank Bryan R. Cullen for providing Dicer-deficient (NoDice) 293T human cell lines.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Tili, E.; Michaille, J.-J.; Costinean, S.; Croce, C.M. MicroRNAs, the immune system and rheumatic disease. *Nat. Clin. Pract. Rheumatol.* **2008**, *4*, 534–541. [[CrossRef](#)]
2. Song, M.-S.; Rossi, J.J. Molecular mechanisms of Dicer: Endonuclease and enzymatic activity. *Biochem. J.* **2017**, *474*, 1603–1618. [[CrossRef](#)]
3. Rybak-Wolf, A.; Jens, M.; Murakawa, Y.; Herzog, M.; Landthaler, M.; Rajewsky, N. A Variety of Dicer Substrates in Human and *C. elegans*. *Cell* **2014**, *159*, 1153–1167. [[CrossRef](#)]
4. Bartel, D.P. MicroRNAs: Target Recognition and Regulatory Functions. *Cell* **2009**, *136*, 215–233. [[CrossRef](#)]
5. Valencia-Sanchez, M.A.; Liu, J.; Hannon, G.J.; Parker, R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev.* **2006**, *20*, 515–524. [[CrossRef](#)]
6. Grishok, A.; Pasquinelli, A.E.; Conte, D.; Li, N.; Parrish, S.; Ha, I.; Baillie, D.L.; Fire, A.; Ruvkun, G.; Mello, C.C. Genes and Mechanisms Related to RNA Interference Regulate Expression of the Small Temporal RNAs that Control *C. elegans* Developmental Timing. *Cell* **2001**, *106*, 23–34. [[CrossRef](#)]
7. Volpe, T.A.; Kidner, C.; Hall, I.M.; Teng, G.; Grewal, S.I.S.; Martienssen, R.A. Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi. *Science* **2002**, *297*, 1833–1837. [[CrossRef](#)] [[PubMed](#)]
8. Kurzynska-Kokorniak, A.; Jackowski, P.; Figlerowicz, M.; Figlerowicz, M. Human- and virus-encoded microRNAs as potential targets of antiviral therapy. *Mini Rev. Med. Chem.* **2009**, *9*, 927–937. [[CrossRef](#)] [[PubMed](#)]
9. Calin, G.; Croce, C.M. MicroRNA Signatures in Human Cancers. *Nat. Rev. Cancer* **2006**, *6*, 857–866. [[CrossRef](#)]
10. Esquela-Kerscher, A.; Slack, F. Oncomirs—microRNAs with a role in cancer. *Nat. Rev. Cancer* **2006**, *6*, 259–269. [[CrossRef](#)] [[PubMed](#)]
11. Hebert, S.; de Strooper, B. Alterations of the microRNA network cause neurodegenerative disease. *Trends Neurosci.* **2009**, *32*, 199–206. [[CrossRef](#)] [[PubMed](#)]
12. Macrae, I.; Li, F.; Zhou, K.; Cande, W.; Doudna, J. Structure of Dicer and Mechanistic Implications for RNAi. *Cold Spring Harb. Symp. Quant. Biol.* **2006**, *71*, 73–80. [[CrossRef](#)] [[PubMed](#)]
13. MacRae, I.J.; Doudna, J.A. Ribonuclease revisited: Structural insights into ribonuclease III family enzymes. *Curr. Opin. Struct. Biol.* **2007**, *17*, 138–145. [[CrossRef](#)] [[PubMed](#)]
14. Zhang, H.; Kolb, F.A.; Jaskiewicz, L.; Westhof, E.; Filipowicz, W. Single Processing Center Models for Human Dicer and Bacterial RNase III. *Cell* **2004**, *118*, 57–68. [[CrossRef](#)]
15. Taylor, D.W.; Ma, E.; Shigematsu, H.; Cianfrocco, M.A.; Noland, C.L.; Nagayama, K.; Nogales, E.; Doudna, J.A.; Wang, H.-W. Substrate-specific structural rearrangements of human Dicer. *Nat. Struct. Mol. Biol.* **2013**, *20*, 662–670. [[CrossRef](#)]
16. Ciecchanowska, K.; Pokornowska, M.; Kurzynska-Kokorniak, A. Genetic Insight into the Domain Structure and Functions of Dicer-Type Ribonucleases. *Int. J. Mol. Sci.* **2021**, *22*, 616. [[CrossRef](#)]
17. 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. [[CrossRef](#)]
18. Liu, Z.; Wang, J.; Cheng, H.; Ke, X.; Sun, L.; Zhang, Q.C.; Wang, H.-W. Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate. *Cell* **2018**, *173*, 1191–1203.e12. [[CrossRef](#)]
19. Tian, Y.; Simanshu, D.; Ma, J.-B.; Park, J.-E.; Heo, I.; Kim, V.N.; Patel, D.J. A Phosphate-Binding Pocket within the Platform-PAZ-Connector Helix Cassette of Human Dicer. *Mol. Cell* **2014**, *53*, 606–616. [[CrossRef](#)]
20. Wostenberg, C.; Lary, J.W.; Sahu, D.; Acevedo, R.; Quarles, K.A.; Cole, J.L.; Showalter, S.A. The role of human Dicer-dsRBD in processing small regulatory RNAs. *PLoS ONE* **2012**, *7*, e51829. [[CrossRef](#)]
21. Qin, H.; Chen, F.; Huan, X.; Machida, S.; Song, J.; Yuan, Y.A. Structure of the Arabidopsis thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein-protein interaction. *RNA* **2010**, *16*, 474–481. [[CrossRef](#)]
22. Dlakić, M. DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold. *Bioinformatics* **2006**, *22*, 2711–2714. [[CrossRef](#)]
23. Ota, H.; Sakurai, M.; Gupta, R.; Valente, L.; Wulff, B.-E.; Ariyoshi, K.; Iizasa, H.; Davuluri, R.V.; Nishikura, K. ADAR1 Forms a Complex with Dicer to Promote MicroRNA Processing and RNA-Induced Gene Silencing. *Cell* **2013**, *153*, 575–589. [[CrossRef](#)]
24. Pokornowska, M.; Milewski, M.C.; Ciecchanowska, K.; Szczepańska, A.; Wojnicka, M.; Radogostowicz, Z.; Figlerowicz, M.; Kurzynska-Kokorniak, A. The RNA–RNA base pairing potential of human Dicer and Ago2 proteins. *Cell. Mol. Life Sci.* **2019**, *77*, 3231–3244. [[CrossRef](#)]

25. Wojnicka, M.; Szczepanska, A.; Kurzynska-Kokorniak, A. Unknown Areas of Activity of Human Ribonuclease Dicer: A Putative Deoxyribonuclease Activity. *Molecules* **2020**, *25*, 1414. [[CrossRef](#)]
26. Bogerd, H.P.; Whisnant, A.W.; Kennedy, E.M.; Flores, O.; Cullen, B.R. Derivation and characterization of Dicer- and microRNA-deficient human cells. *RNA* **2014**, *20*, 923–937. [[CrossRef](#)] [[PubMed](#)]
27. Ma, E.; MacRae, I.J.; Kirsch, J.F.; Doudna, J.A. Autoinhibition of Human Dicer by Its Internal Helicase Domain. *J. Mol. Biol.* **2008**, *380*, 237–243. [[CrossRef](#)] [[PubMed](#)]
28. Koralewska, N.; Hoffmann, W.; Pokornowska, M.; Milewski, M.; Bienkowska-Szewczyk, K.; Figlerowicz, M.; Lipinska, A.; Kurzynska-Kokorniak, A. How short RNAs impact the human ribonuclease Dicer activity: Putative regulatory feedback-loops and other RNA-mediated mechanisms controlling microRNA processing. *Acta Biochim. Pol.* **2017**, *63*, 773–783. [[CrossRef](#)] [[PubMed](#)]
29. Kurzyńska-Kokorniak, A.; Koralewska, N.; Tyczewska, A.; Twardowski, T.; Figlerowicz, M. A New Short Oligonucleotide-Based Strategy for the Precursor-Specific Regulation of microRNA Processing by Dicer. *PLoS ONE* **2013**, *8*, e77703. [[CrossRef](#)] [[PubMed](#)]
30. Tyczewska, A.; Kurzynska-Kokorniak, A.; Koralewska, N.; Szopa, A.; Kietrys, A.M.; Wrzesinski, J.; Twardowski, T.; Figlerowicz, M. Selection of RNA Oligonucleotides That Can Modulate Human Dicer Activity In Vitro. *Nucleic Acid Ther.* **2011**, *21*, 333–346. [[CrossRef](#)]
31. Ye, X.; Paroo, Z.; Liu, Q. Functional Anatomy of the Drosophila MicroRNA-generating Enzyme. *J. Biol. Chem.* **2007**, *282*, 28373–28378. [[CrossRef](#)] [[PubMed](#)]
32. Lee, Y.; Hur, I.; Park, S.-Y.; Kim, Y.K.; Suh, M.R.; Kim, V.N. The role of PACT in the RNA silencing pathway. *EMBO J.* **2006**, *25*, 522–532. [[CrossRef](#)] [[PubMed](#)]
33. Lau, P.-W.; Guiley, K.Z.; De, N.; Potter, C.S.; Carragher, B.; Macrae, I.J. The molecular architecture of human Dicer. *Nat. Struct. Mol. Biol.* **2012**, *19*, 436–440. [[CrossRef](#)]
34. Livak, K.J.; Schmittgen, T.D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the  $2^{-\Delta\Delta CT}$  Method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]





# RNA and DNA G-quadruplexes bind to human dicer and inhibit its activity

Natalia Koralewska<sup>1</sup> · Agnieszka Szczepanska<sup>2</sup> · Kinga Ciechanowska<sup>2</sup> · Marta Wojnicka<sup>2</sup> · Maria Pokornowska<sup>2</sup> · Marek C. Milewski<sup>1</sup> · Dorota Gudanis<sup>3</sup> · Daniel Baranowski<sup>3</sup> · Chandran Nithin<sup>4</sup> · Janusz M. Bujnicki<sup>4,5</sup> · Zofia Gdaniec<sup>3</sup> · Marek Figlerowicz<sup>1</sup> · Anna Kurzynska-Kokorniak<sup>2</sup>

Received: 6 July 2020 / Revised: 27 January 2021 / Accepted: 19 February 2021 / Published online: 17 March 2021  
© The Author(s) 2021

## Abstract

Guanine (G)-rich single-stranded nucleic acids can adopt G-quadruplex structures. Accumulating evidence indicates that G-quadruplexes serve important regulatory roles in fundamental biological processes such as DNA replication, transcription, and translation, while aberrant G-quadruplex formation is linked to genome instability and cancer. Understanding the biological functions played by G-quadruplexes requires detailed knowledge of their protein interactome. Here, we report that both RNA and DNA G-quadruplexes are bound by human Dicer *in vitro*. Using *in vitro* binding assays, mutation studies, and computational modeling we demonstrate that G-quadruplexes can interact with the Platform–PAZ–Connector helix cassette of Dicer, the region responsible for anchoring microRNA precursors (pre-miRNAs). Consequently, we show that G-quadruplexes efficiently and stably inhibit the cleavage of pre-miRNA by Dicer. Our data highlight the potential of human Dicer for binding of G-quadruplexes and allow us to propose a G-quadruplex-driven sequestration mechanism of Dicer regulation.

**Keywords** MiRNA biogenesis · PAZ domain · Dicer PPC cassette · Dicer inhibition · Ribonucleoprotein complexes · Regulation of enzyme activity

---

Natalia Koralewska and Agnieszka Szczepanska contributed equally.

✉ Anna Kurzynska-Kokorniak  
akurzyns@man.poznan.pl

- <sup>1</sup> Department of Molecular and Systems Biology, Institute of Bioorganic Chemistry Polish Academy of Sciences, 61-704 Poznan, Poland
- <sup>2</sup> Department of Ribonucleoprotein Biochemistry, Institute of Bioorganic Chemistry Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland
- <sup>3</sup> Department of Biomolecular NMR, Institute of Bioorganic Chemistry Polish Academy of Sciences, 61-704 Poznan, Poland
- <sup>4</sup> Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, 02-109 Warsaw, Poland
- <sup>5</sup> Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland

## Introduction

Dicer belongs to the ribonuclease III (RNase III) family of double-stranded RNA (dsRNA)-specific endoribonucleases that are essential for the maturation and decay of coding and noncoding RNAs in both prokaryotes and eukaryotes [1]. Proteins classified into this family share a unique fold (RNase III domain) that dimerizes to form a catalytic center where the hydrolysis of phosphodiester bonds occurs, leaving a characteristic dsRNA product with a 5' phosphate and a 2-nucleotide (nt) overhang with a hydroxyl group at the 3' end (reviewed in [1]). Human Dicer (hDicer), consisting of 1992 amino acids (220 kDa), is one of the most structurally complex members of the RNase III family. It comprises an amino (N)-terminal putative helicase domain, a domain of unknown function (DUF283), Platform, Piwi–Argonaute–Zwille (PAZ) domain, a Connector helix, two RNase III domains (RNase IIIa and RNase IIIb), and a dsRNA-binding domain (dsRBD) [2, 3].

Extensive studies have deciphered the roles of the Dicer domains in binding and processing of its canonical substrates, i.e., dsRNAs and single-stranded hairpin precursors

of microRNAs (pre-miRNAs). The putative helicase domain selectively interacts with the apical loop of pre-miRNA and helps to discriminate between substrates [4]. The DUF283 domain has been implicated in the binding of single-stranded nucleic acids [5], and may therefore be involved in interactions with the apical loop of pre-miRNA hairpins [6]. Two adjacent domains, Platform and PAZ, anchor the 5' phosphate and 2-nt 3' overhang of a substrate [3]. The RNase IIIa and RNase IIIb domains form a single dsRNA-cleavage center [7]. Finally, the C-terminal dsRBD is presumed to play an auxiliary role in RNA binding [8].

G-quadruplexes are noncanonical structures formed by guanine-rich DNA and RNA molecules. They are organized in stacks of two or more G-quartets, in which four guanines are associated through Hoogsteen base pairing and stabilized by a monovalent cation, such as Na<sup>+</sup> or K<sup>+</sup> (for reviews see [9]). G-quadruplex-forming motifs occur in the genomes and transcriptomes of many species, including humans. Biologically relevant G-quadruplexes were first discovered in eukaryotic telomeres [10]. Since then, multiple studies have shown the importance of G-quadruplexes in the regulation of DNA replication, gene expression, and telomere maintenance, and have linked G-quadruplexes to several human diseases (for review see [11]). The formation of G-quadruplexes has also been reported in the regulation of RNA metabolism, including the miRNA pathway, where such structures may influence Dicer activity and other steps of miRNA biogenesis [12–14].

Previously, we characterized the inhibitory potential of structurally diverse short RNA molecules, 12- to 60-nt in length, able to influence processing of pre-miRNA by hDicer either by binding to the enzyme or by base pairing with the substrate (pre-miRNA) [15–17]. We showed that 12-mers are too short to bind to hDicer efficiently, and consequently, to act as competitive inhibitors. However, if these 12-mers base pair with the apical region of a pre-miRNA hairpin, they preclude the cleavage by hDicer, thereby acting as pre-miRNA-specific inhibitors [16]. Interestingly, in contrast to the other tested 12-mers, one guanine-rich molecule bound to hDicer and inhibited cleavage of all tested pre-miRNAs [16]. However, our previous studies did not explain the mechanism of this phenomenon.

Here, we demonstrate that a hDicer-binding guanine-rich 12-mer can form a G-quadruplex structure. This finding inspired us to investigate possible interactions between hDicer and nucleic acids adopting G-quadruplex structures. In our studies, we used a full-length hDicer, as well as the hDicer fragment responsible for canonical substrate binding, encompassing the Platform and PAZ domains. We also analyzed several structurally well-characterized short telomeric RNA and DNA G-quadruplexes from human and the ciliate *Oxytricha nova*. Using molecular modeling and mutation studies, we demonstrated that the Platform and PAZ domains

can bind both RNA and DNA G-quadruplexes. We showed that binding of RNA and DNA G-quadruplexes by hDicer precludes pre-miRNA cleavage, which suggests the existence of yet another mechanism regulating Dicer activity and miRNA biogenesis in vivo.

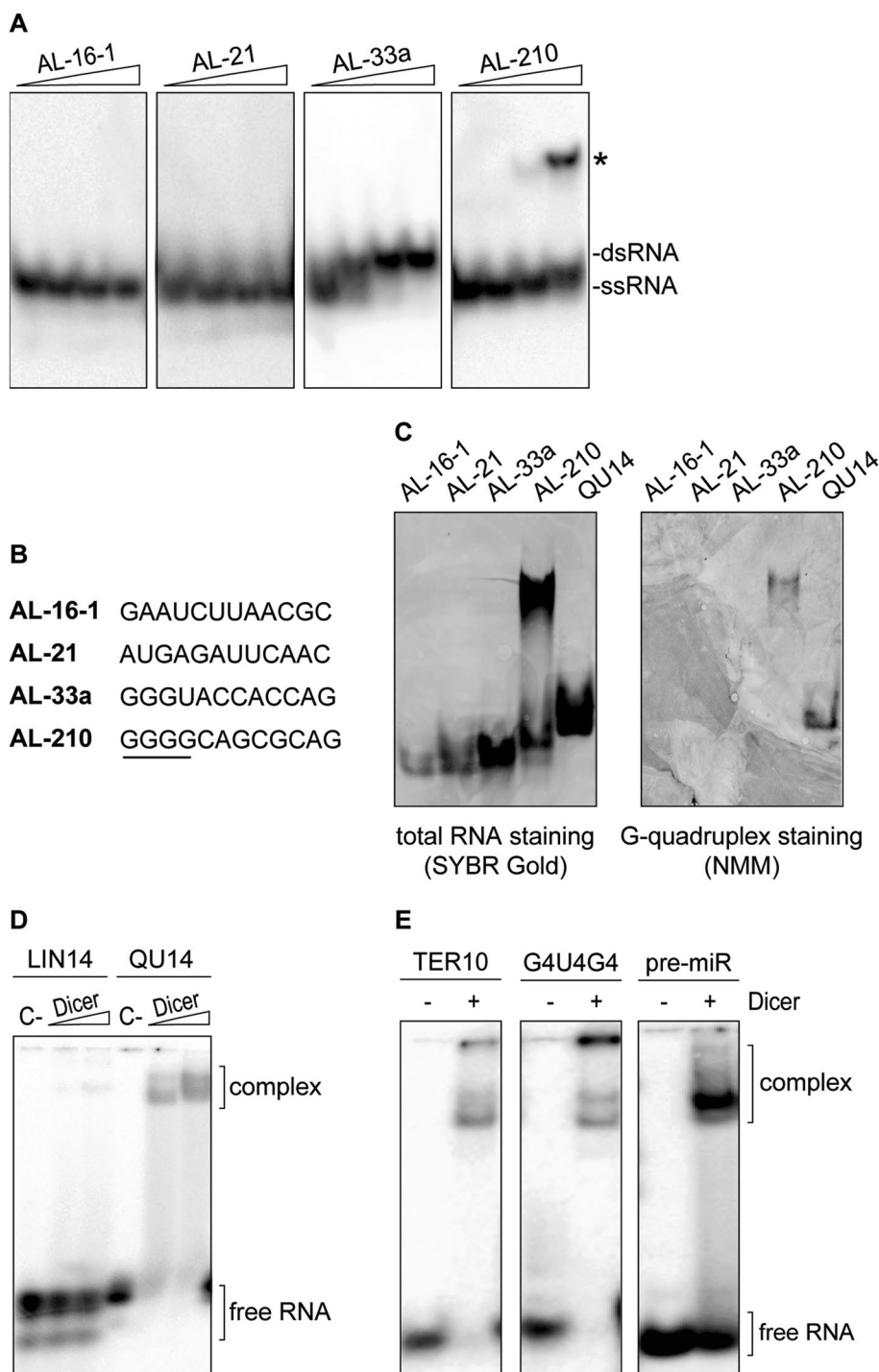
## Results

### hDicer binds RNA G-quadruplexes

The 12-nt guanine-rich RNA, found to bind to hDicer, was perfectly complementary to the apical loop of pre-mir-210, and consequently, it was named “AL-210” [16]. We hypothesized that the structure adopted by AL-210 determined its ability to bind to hDicer and to inhibit its cleavage activity. Accordingly, we used Fold and bifold algorithms provided by the RNA structure web server [18] to predict the lowest free energy secondary structures for AL-210 and three other 12-mers used in our previous studies: AL-16-1, AL-21, AL-33a (i.e., the oligomers designed to target the apical loops of pre-mir-16-1, pre-mir-21, pre-mir-33a, respectively). The results indicated that AL-33a and AL-210 can form stable homodimers (Supplementary Fig. S1). AL-16-1 and AL-21 were predicted to be monomeric. Since the structure adopted by RNA may depend on the RNA concentration, each 12-nt RNA was assayed by polyacrylamide gel electrophoresis (PAGE), under nondenaturing conditions, at four different concentrations (0.01, 0.1, 1, and 10 μM). Under the applied conditions AL-16-1 and AL-21 always migrated in the gel as a single conformer, whereas AL-33a and AL-210, depending on the concentration, migrated as faster- or slower-moving conformers. The faster-moving conformers seemed to represent monomers (single-stranded RNA, ssRNA) (Fig. 1a). At concentrations equal to or greater than 0.1 μM, AL-33a migrated as the slower-moving conformer, presumably corresponding to a homodimer (dsRNA) form. Interestingly, at a concentration equal to or greater than 1 μM, AL-210 migrated even slower than a putative dsRNA form (Fig. 1a and Supplementary Fig. S1).

When analyzing the primary structure of the 12-mers (Fig. 1b), we noticed that a striking feature of AL-210 is a stretch of four guanines at the 5' end. Such a sequence motif is commonly found in nucleic acids adopting G-quadruplex structures. To test the possibility that the slower-migrating conformer of AL-210 represents a G-quadruplex form, we performed native PAGE followed by G-quadruplex-specific staining with the *N*-methyl mesoporphyrin IX (NMM) fluorescent probe [19]. To promote and stabilize a potential G-quadruplex structure, oligomers were incubated in buffer containing 100 mM KCl before the electrophoresis. As a positive control, we used a 14-nt RNA (called “QU14”) known to adopt a G-quadruplex architecture [20]. Under the

**Fig. 1** hDicer binds RNA G-quadruplexes. **a** Native PAGE analysis of the mixtures of unlabeled and 5'-<sup>32</sup>P-labeled RNA 12-mers (AL). *Triangles* represent increasing amounts of a given oligomer (0.01, 0.1, 1, 10 μM). Bands corresponding to single- and double-stranded forms of the oligomers are indicated. **b** Comparison of the sequences of RNA 12-mers used in the study. A guanine tract in the sequence of AL-210 is underlined. **c** Native PAGE analysis of RNA 12-mers (AL) and a control 14-nt G-quadruplex (QU14) (100 pmol each). Gels were treated with nucleic acid stain SYBR Gold (*left*) or G-quadruplex-specific dye, NMM (*right*). **d** EMSAs with hDicer and the 5'-<sup>32</sup>P-labeled 14-nt RNA not adopting a G-quadruplex structure (LIN14), or the 5'-<sup>32</sup>P-labeled RNA G-quadruplex (QU14). *C-* a control sample with no protein. *Triangles* represent increasing amounts of hDicer (100, 250 nM). **e** EMSAs with hDicer (250 nM) and the 5'-<sup>32</sup>P-labeled RNA G-quadruplex (TER10 or G4U4G4) or a control pre-miRNA. ± refers to Dicer presence in the reaction mixture



applied conditions, QU14, AL-16-1, AL-21, and AL-33a each migrated as a single conformer (Fig. 1c, left panel), whereas AL-210 migrated as faster- and slower-moving conformers, as demonstrated above (Fig. 1a). Staining with NMM showed that the G-quadruplex-specific dye was bound only by QU14 and the slower-moving AL-210 conformer, but not by AL-16-1, AL-21, AL-33a, or fast-moving AL-210

conformers (Fig. 1c, right panel). Together, these results suggest that AL-210 adopts a G-quadruplex structure and, presumably, in such a form, it could bind to hDicer.

We hypothesized that hDicer could also bind other RNAs containing G-quadruplex structures, including QU14. We performed electrophoretic mobility shift assays (EMSAs) involving hDicer and 5'-<sup>32</sup>P-labeled QU14 or a control 14-nt



RNA (called “LIN14”) that did not adopt a G-quadruplex structure. QU14 was bound efficiently by hDicer, while only residual binding of LIN14 was detected (Fig. 1d). Weak binding of LIN14 agrees with our previous observations that hDicer does not interact efficiently with RNAs shorter than 20-nt [17]. Next, we tested whether hDicer would also bind other short RNAs (<20-nt) with G-quadruplex structures. Accordingly, we performed EMSAs with hDicer and 5′-<sup>32</sup>P-labeled oligonucleotides with sequences corresponding to the well-characterized telomeric repeat-containing RNAs (TERRAs) of human [21] or ciliate *Oxytricha nova* [22] origin, i.e., r(GGGUUAGGGU) called “TER10” and r(GGGGUUUUGGGG) called “G4U4G4”, respectively (detailed information about guanine-rich oligomers used in the studies is presented in Table S1). Both TER10 and G4U4G4 were bound by hDicer. A control binding reaction contained the canonical hDicer substrate, 58-nt pre-miRNA (Fig. 1e). These results confirm that hDicer is able to bind RNAs adopting a G-quadruplex structure.

### The PPC cassette of hDicer binds RNA and DNA G-quadruplexes

According to the current knowledge about the mechanism of Dicer action, the initial recognition and anchoring of the substrate occur within the region spanning the Platform–PAZ–Connector helix cassette, called “PPC” [3]. The PPC contains two adjacent pockets, a 2-nt 3′-overhang-binding pocket (3′-pocket) within the PAZ domain and a phosphate-binding pocket (5′-pocket) within the Platform and PAZ domains (Fig. 2a) [3, 23]. To test whether PPC can also bind RNA G-quadruplexes, we produced PPC in a bacterial expression system (Supplementary Fig. S2A) and then carried out EMSAs with PPC and either 5′-<sup>32</sup>P-labeled r[AGGG(UUAGGG)<sub>3</sub>] called “TER22”, which is an extended version of TER10, or G4U4G4 (Table S1). PPC formed stable complexes with both RNA oligomers, with a  $K_d$  value of ~7 nM for the PPC and TER22 complex (Fig. 2b), and ~10 nM for the PPC and G4U4G4 complex (Fig. 2c). For comparison, we also determined binding affinities for PPC and 5′-<sup>32</sup>P-labeled 21-nt RNA (called “LIN21”), 58-nt pre-mir-21, and 19-bp RNA duplexes having 2-nt 3′ overhanging ends (called “dsRNA\_OV”) (Supplementary Fig. S2B). The  $K_d$  values calculated for PPC and RNA substrates were ~46 nM for LIN21, ~205 nM for pre-mir-21 and ~50 nM for dsRNA\_OV (Supplementary Fig. S2C). The results indicate that the PPC cassette of hDicer binds RNA G-quadruplexes with higher affinity than it binds a pre-miRNA substrate, a miRNA-size RNA, or a miRNA-like duplex.

Given the fact that telomeric repeat-containing sequences are present not only in telomeric RNA but also within telomeric DNA at the chromosome ends, we asked whether

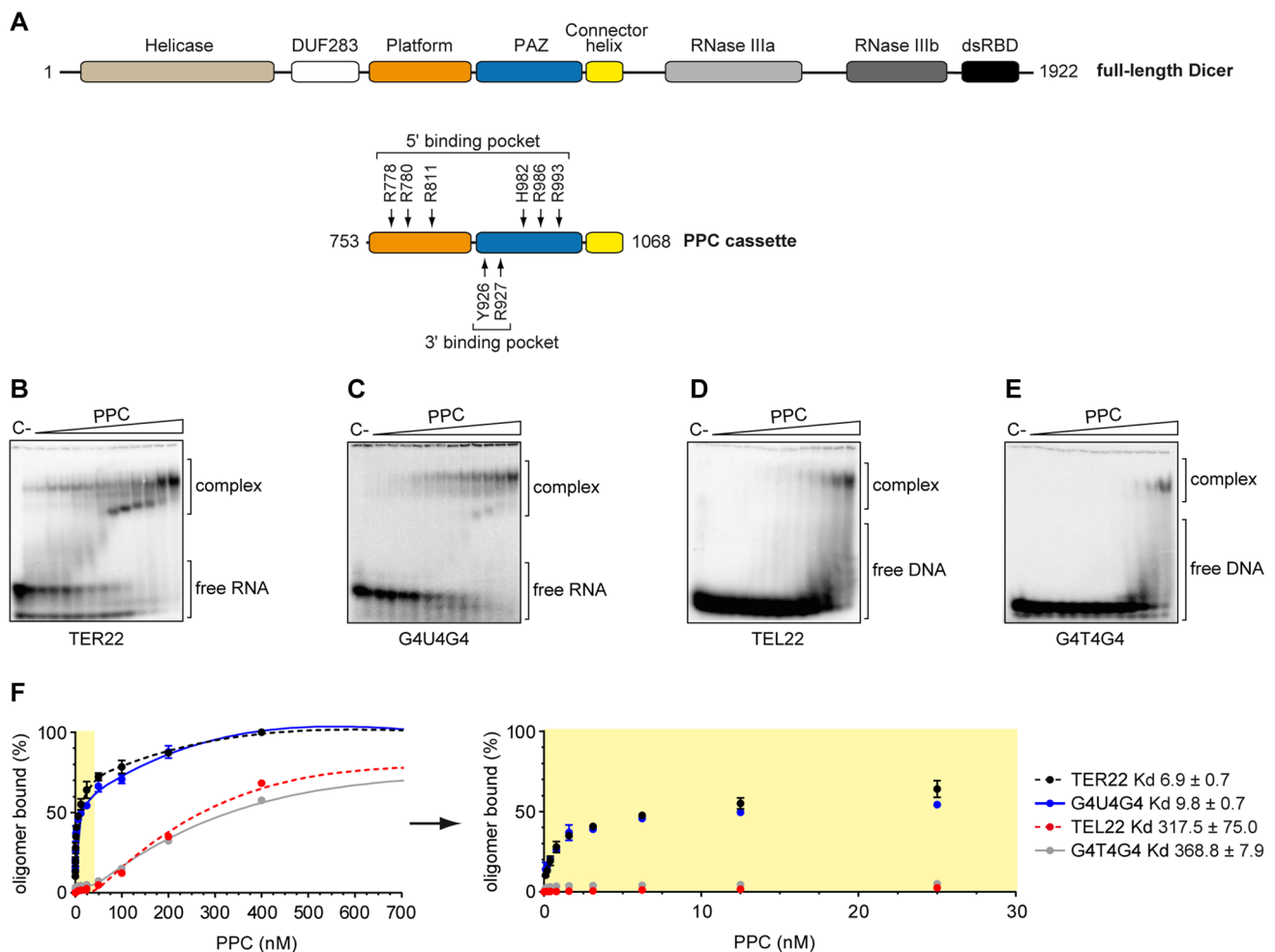
PPC could interact with DNA G-quadruplexes as well. Accordingly, we used 5′-<sup>32</sup>P-labeled DNA oligomers with sequences identical to TER22 and G4U4G4, i.e., “TEL22” and “G4T4G4”, respectively (Fig. 2d, e). The  $K_d$  value was ~318 nM for the PPC and TEL22 complex and ~369 nM for the PPC and G4T4G4 complex. These data indicated that PPC bound DNA G-quadruplexes much weaker than their RNA G-quadruplex counterparts. Additionally, TER22 and TEL22 were bound by PPC slightly better (~1.4 and ~1.2 times) than G4U4G4 and G4T4G4, respectively. Together, these results indicate that both the type and length of nucleic acid adopting a G-quadruplex structure determine its binding affinity to the PPC cassette of hDicer.

### Both the 3′-pocket and the 5′-pocket of hDicer PPC cassette are involved in the binding of G-quadruplexes

To test whether oligonucleotides adopting G-quadruplex structures could, like canonical Dicer substrates, be bound in the 3′-pocket and/or 5′-pocket of the PPC cassette, we performed molecular docking and modeling. For this we used the structure of hDicer PPC cassette (PDB entry 4NGF) [3] and the G-quadruplex structures of the 10-nt human TERRA (TER10) (PDB entry 2M18) [21] or *O. nova* G4T4G4 (PDB entry 1JPQ) [22]. TER10 formed a dimer of bimolecular G-quadruplexes [21], whereas G4T4G4 appeared as a bimolecular G-quadruplex [22].

We inspected docked poses of both TER10 and G4T4G4 with hDicer protein for clashes with pre-miRNA binding (for detailed information see “Materials and methods” section). We selected eight best models (docked poses) for the TER10-PPC complex and six best models for the G4T4G4-PPC complex. In seven out of the eight models selected for the TER10-PPC complex, TER10 was positioned within the 3′-pocket of the PPC cassette. In all these models, the 3′ end of one out of the four RNA strands was anchored in the 3′-pocket (Fig. 3a). In one of the eight models, the 3′ end of TER10 was located in the vicinity of amino acid residues lining up the 5′-pocket (Fig. 3b). Three docked poses obtained for the G4T4G4-PPC complex showed that the G4T4G4 quadruplex bound within the 3′-pocket (Fig. 3c), and the other three poses revealed G4T4G4 located within the 5′-pocket of the PPC cassette (Fig. 3d).

To validate these models, we generated two variants of the hDicer PPC cassette, one containing two substitutions in the 3′-pocket (Y926F/R927A variant), and the other containing six substitutions in the 5′-pocket (R778A/R780A/R811A/H982A/R986A/R993A variant) (Supplementary Fig. S3A). These changes significantly affect the binding of small interfering RNA (siRNA) by hDicer PPC cassette [3], and the corresponding full-length hDicer variants have been well characterized biochemically [23]. Using EMSAs, we



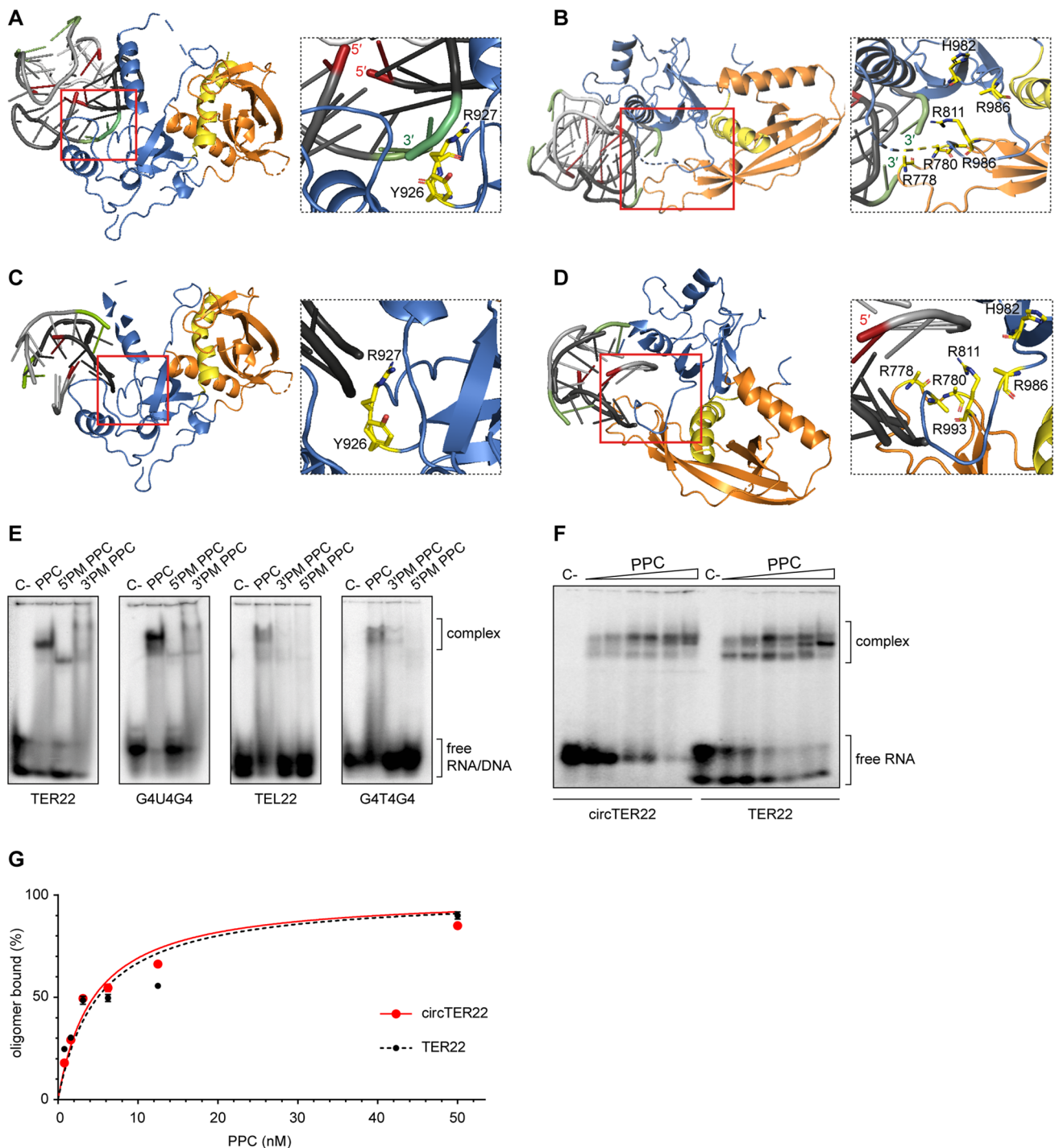
**Fig. 2** The PPC cassette of hDicer binds RNA and DNA G-quadruplexes. **a** Domain architecture of hDicer (*top*) and PPC cassette (*below*). *Arrows* indicate the RNA-interacting residues within the 5'- and 3'-pocket. **b, c** EMSAs with PPC and the 5'-<sup>32</sup>P-labeled RNA G-quadruplexes TER22 **b** and G4U4G4 **c**. *C-* a control sample with no protein. *Triangles* represent increasing amounts of PPC (0.1, 0.2, 0.4, 0.8, 1.6, 3.125, 6.25, 12.5, 25, 50, 100, 200, 400 nM). The  $K_d$  of RNA-PPC complexes was determined from the binding isotherms by curve-fitting using nonlinear regression. Error bars represent SD from three separate experiments. **d, e** EMSAs with PPC

and the 5'-<sup>32</sup>P-labeled DNA G-quadruplexes TEL22 **d** and G4T4G4 **e**. *C-* a control sample with no protein. *Triangles* represent increasing amounts of PPC (0.1, 0.2, 0.4, 0.8, 1.6, 3.125, 6.25, 12.5, 25, 50, 100, 200, 400 nM). The  $K_d$  values of DNA-PPC complexes were determined from the binding isotherms by curve fitting using nonlinear regression. Error bars represent SD from three separate experiments. **f** Quantitative analysis of the binding assay between PPC and 5'-<sup>32</sup>P-labeled RNA/DNA G-quadruplexes. Error bars represent SD from three separate experiments

examined the ability of PPC variants to bind 5'-<sup>32</sup>P-labeled TER22, G4U4G4, TEL22, and G4T4G4 oligomers. Changes in both the 3'-pocket and the 5'-pocket reduced PPC binding to all G-quadruplexes analyzed (Fig. 3e). Mutations in the PPC pockets had a greater impact on binding to the DNA G-quadruplexes than to their RNA counterparts. However, this can be explained by the fact that, under the applied reaction conditions, the wild-type protein displayed much lower maximum binding capacity for DNA G-quadruplexes than for the corresponding RNA G-quadruplexes (Fig. 2b–e), and mutations in the PPC pockets only proportionally decreased binding. The lower binding capacity of PPC to

DNA G-quadruplexes than to RNA G-quadruplexes may be due to the participation of ribose 2'-hydroxyl groups in hydrogen bonding with amino acid residues of the PPC cassette. Taken together, binding assays conducted for the wild-type protein and the 3'-pocket and the 5'-pocket variants (Fig. 3e) corresponded well with the models of PPC and G-quadruplex complexes (Fig. 3a–d), collectively indicating that both pockets of the hDicer PPC cassette are important for the binding of G-quadruplexes.

To investigate whether the free ends of the RNA adopting a G-quadruplex structure are necessary for its binding to the hDicer PPC cassette, we circularized 5'-<sup>32</sup>P-labeled



**Fig. 3** Both the 3'- and 5'-pocket of hDicer PPC cassette are involved in the binding of G-quadruplexes. **a** 3D model of TER10 forming a dimer of bimolecular G-quadruplexes bound to the 3'-pocket of hDicer PPC cassette. The 3'- and 5'-ends of the oligonucleotide are colored in green and red, respectively. The Platform is colored in orange, PAZ in blue and Connector helix in yellow. The same color coding is maintained in subpanels **b–d** of this **b** 3D model of TER10 forming a dimer of bimolecular G-quadruplexes bound to the 5'-pocket of hDicer PPC cassette. **c** 3D model of G4T4G4 forming a bimolecular G-quadruplex bound to the 3'-pocket of hDicer PPC cassette. **d** 3D model of G4T4G4 forming a bimolecular G-quadruplex bound to the 5'-pocket of hDicer PPC cassette.

plex bound to the 5'-pocket of hDicer PPC cassette. **e** EMSAs with PPC, PPC 5'- and 3'-pocket mutants (5'PM, 3'PM), and 5'-<sup>32</sup>P-labeled G-quadruplexes (TER22, G4U4G4, TEL22, G4T4G4). 500 nM protein and 10,000 cpm (approximately 5 nM) of RNA/DNA were used per lane. C- a control sample with no protein. **f** EMSAs of binding between PPC and TER22 with free ends (TER22) or TER22 with ligated ends (circTER22). Triangles represent the increasing amount of PPC (0.8, 1.6, 3.125, 6.25, 12.5, 50 nM). C- a control sample with no protein. **g** Quantitative analysis of the binding assay between PPC and TER22/circTER22 G-quadruplexes. Error bars represent SD from three separate experiments

TER22 using T4 RNA ligase, as described previously [24] (Supplementary Fig. S3B, C). In adequate protein concentrations, PPC bound this “circTER22” with similar efficiency as the TER22 RNA with free ends (Fig. 3f, g), indicating that hDicer PPC cassette binds G-quadruplexes no matter whether their 3'- or 5' ends are available.

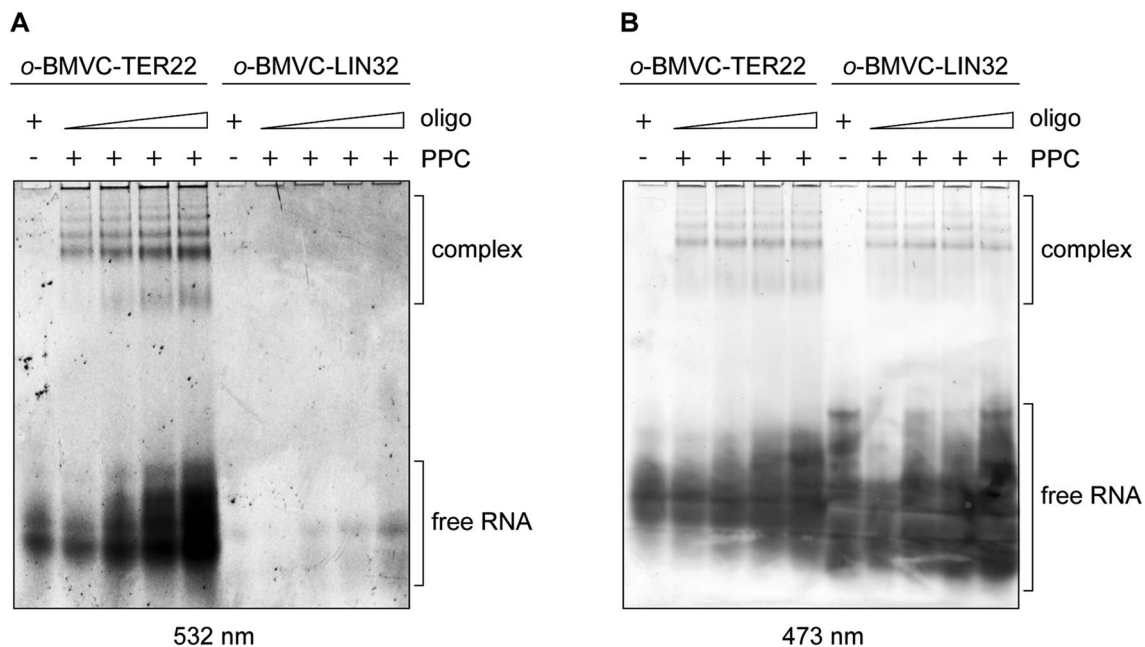
### RNA G-quadruplexes bound by the PPC cassette retain their structure

To test whether RNA G-quadruplexes retain their structures upon binding the hDicer PPC cassette, we conducted a binding assay involving TER22 with a derivative of 3,6-bis(1-methyl-2-vinyl-pyridinium) carbazole diiodide (*o*-BMVC) covalently attached to its 5'-end (Supplementary Fig. S3D). *o*-BMVC is a fluorescent light-up probe that selectively binds to G-quadruplex structures [25]. As a control, we used 32-nt RNA (called “LIN32”) that does not contain a G-quadruplex motif, labeled with *o*-BMVC in the same way as TER22. Each oligomer was incubated with PPC, and the reaction mixtures were analyzed in a polyacrylamide gel under native conditions. After electrophoresis, we exposed the gel first to 532 nm light to detect bands corresponding to the RNA species adopting G-quadruplex structures (Fig. 4a). Subsequently, we stained the gel with SYBR Gold solution and exposed it to 473 nm light to visualize the total RNA

pool (Fig. 4b). The results showed that RNA G-quadruplexes retained their structure upon binding the PPC cassette.

### RNA and DNA G-quadruplexes inhibit pre-miRNA processing by hDicer

In vitro, RNA oligonucleotides bound to hDicer inhibit the cleavage of pre-miRNAs by this enzyme [15, 16]. To investigate whether oligonucleotides containing G-quadruplex motifs can also influence the cleavage activity of hDicer, we used RNA oligomers representing human and *O. nova* TERRA (TER10, TER22, G4U4G4) and DNA oligomers representing the corresponding telomeric repeats (TEL22, G4T4G4). As well as those mentioned above, we used three RNAs of 12 to 20-nt: TER12, TER18, TER18-2A. All of the chosen oligomers are known to adopt various G-quadruplex architectures (Table S1). To assess the effect of RNA and DNA G-quadruplexes on hDicer cleavage activity, we performed a set of assays involving hDicer, 5'-<sup>32</sup>P-labeled pre-mir-21 or pre-mir-33a, and a respective oligomer. The particular pre-miRNAs were chosen as they did not interact with tested oligomers (Supplementary Fig. S4A), which ensures that the effects observed in the experiment does not stem from direct RNA–DNA or RNA–RNA binding, but rather, from protein–DNA or RNA interaction. Moreover, pre-mir-21 and pre-mir-33a represent structurally distinct substrates (Supplementary Fig. S5E).



**Fig. 4** RNA G-quadruplexes bound by the PPC cassette retain their structure. **a** EMSAs with 500 nM PPC and TER22 or LIN32, which does not adopt a G-quadruplex structure. Both RNA oligonucleotides were labeled at the 5'-end with a G-quadruplex specific probe (*o*-BMVC). ± indicates the presence or absence of PPC, or the RNA

oligomer. *Triangles* represent increasing amounts of the RNA oligomer (7.5, 15, 22.5, 30 μM). To visualize RNA adopting G-quadruplex structures, the gel was exposed to 532 nm light. **b** The same gel as presented in panel **a** exposed to 473 nm light after staining with SYBR Gold to visualize total RNA



First, we carried out cleavage assays with RNA G-quadruplexes. The efficiency of pre-miRNA (~5 nM) cleavage in reactions with individual oligomers applied at one of three concentrations (0.1, 0.5, or 2  $\mu$ M) was normalized to the cleavage efficiency in a control reaction with no oligomer added. Another set of control reactions involved in 12-nt RNA (called “LIN12”), which does not bind to hDicer, nor adopts a G-quadruplex structure. In all cases, upon the addition of an RNA G-quadruplex, we observed a dose-dependent inhibition of the cleavage of pre-mir-21 (Fig. 5a and Supplementary Fig. S4B) and pre-mir-33a (Fig. 5b and Supplementary Fig. S4C). At the lowest concentration of any of the tested RNA G-quadruplexes, the level of miRNA was reduced by at least 50% in comparison to the control reaction with no oligomer added. When the highest concentration of any of the tested RNA G-quadruplexes was applied, the cleavage of pre-miRNA was abolished by at least 80%. No decrease in the miRNA level was observed in control reactions with LIN12. We did not find any apparent correlation between the length or architecture of the RNA G-quadruplexes and the degree of inhibition they exerted.

Next, we investigated whether the cleavage of pre-miRNA (~5 nM) by hDicer can be affected by DNA counterparts of G4U4G4 and TER22, i.e., G4T4G4 and TEL22 (Fig. 5c, d and Supplementary Fig. S4D, E). At the lowest concentration (0.1  $\mu$ M), G4T4G4 inhibited cleavage of pre-mir-21 and pre-mir-33a by ~90%, while at the highest concentration (2  $\mu$ M), the inhibition reached ~96% for pre-mir-21 and 100% for pre-mir-33a. We observed a similar tendency for TEL22; it inhibited the cleavage of pre-mir-21 and pre-mir-33a by ~70% at the lowest concentration (0.1  $\mu$ M), and by ~95% at the highest concentration (2  $\mu$ M). The results are comparable with those obtained for G4U4G4 (~90% and ~100% inhibition at 0.1 and 2  $\mu$ M concentration, respectively) and TER22 (~75–80% and ~95% inhibition at 0.1 and 2  $\mu$ M concentration, respectively), which indicates no differences in the inhibition potency between RNA G quadruplexes and their DNA G quadruplex counterparts under the applied reaction conditions (Fig. 5a–d and Supplementary Fig. S4B–E).

Subsequently, we performed a time course assay to measure the level of miR-21 or miR-33a produced in reaction with either no inhibitor, a selected G-quadruplex (G4T4G4, G4U4G4, TEL22, TER22) or pre-mir-16-1 added as a competitor. The reactions were performed under the low-turnover conditions, i.e., a twofold molar excess of hDicer (10 nM) over a substrate was used. In addition, the G-quadruplexes and the competitor were in 50-fold molar excess to hDicer. As expected, in control reactions with either pre-mir-21 or pre-mir-33a and hDicer, but no other oligomer added, we observed a hyperbolic relation between the yield of miRNA and the incubation time (Fig. 5e, f and Supplementary Fig.

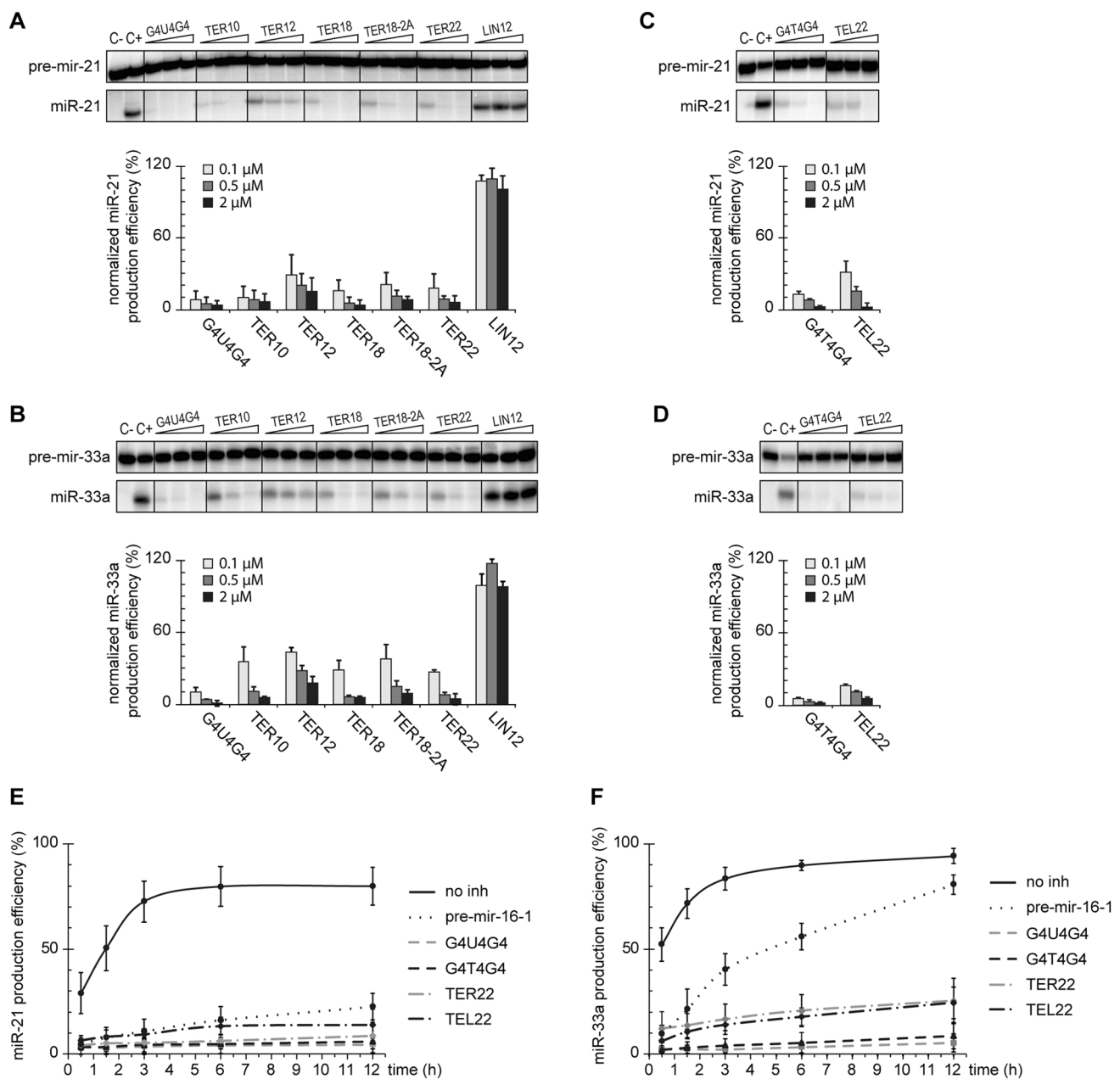
S5A–D). After 12 h, ~90% of either of the substrates was processed by the enzyme.

In reactions with pre-mir-21 as a substrate and pre-mir-16-1 as a competitor, we observed a significant inhibition of pre-mir-21 cleavage at all analyzed time points; even after 12 h only ~22% of pre-mir-21 was processed (Fig. 5e and Supplementary Fig. S5A, B). We also observed stable inhibition for the tested G-quadruplexes; the effect exerted by them was even more prominent than in the case of pre-mir-16-1: after 12 h only ~5% of pre-mir-21 was processed in reactions with G4T4G4 or G4U4G4, and ~10% in reactions with TEL22 or TER22 (Fig. 5e and Supplementary Fig. S5A, B).

In reactions with pre-mir-33a as a substrate, and pre-mir-16-1 as a competitor, initially the levels of miR-33a were much lower than in the control reactions without the inhibitor (e.g., ~10% vs ~50% after the first 30 min of the reaction) (Fig. 5f and Supplementary Fig. S5C, D) but with time the inhibition was gradually abolished; after 12 h the amount of pre-mir-33a processed in the reaction with the competitor and in the control reaction reached ~80% and ~95%, respectively. In reactions with the G-quadruplexes, we also observed low initial levels of miR-33a as in the analogous reactions with pre-mir-16-1 and the reduction of the inhibition with time (Fig. 5f and Supplementary Fig. S5C, D). The amount of pre-mir-33a cut by hDicer increased from ~5% after 30 min to ~10% after 12 h in reactions with G4T4G4 or G4U4G4, and from ~10% to ~25% in reactions with TEL22 or TER22. Despite the observed accumulation of miRNA over time, and in contrast to pre-mir-16-1, the inhibitory effect exerted by G-quadruplexes remained high even after 12 h of incubation (~75% decrease in miRNA production in comparison to the control reaction in the case of G-quadruplexes, vs ~15% in the case of pre-mir-16-1).

Altogether, these findings indicate that RNA and DNA adopting a G-quadruplex structure can affect the cleavage of pre-miRNA by hDicer. Under the low-turnover conditions (excess enzyme to substrate), RNA and DNA G-quadruplexes exerted a similar inhibition effect on the hDicer cleavage of both pre-miRNAs used (Fig. 5e, f). G4U4G4 and G4T4G4 were slightly better inhibitors of pre-mir-21 and pre-mir-33a cleavage than TER22 and TEL22 (95% vs 90% inhibition of pre-mir-21 cleavage after 12 h, and 90% vs 75% inhibition of pre-mir-33a cleavage after 12 h, respectively) (Fig. 5e, f, and Supplementary Fig. S5A–D).

Subsequently, we performed a time course assay under the high-turnover conditions using a 50-fold molar excess of a substrate to hDicer (0.5 nM), and a 100-fold molar excess of the G-quadruplexes or the competitor to the enzyme. We found that G4U4G4, G4T4G4 and TER22 retained their inhibitory potential after 12 h incubation (Supplementary Fig. S6), whereas the inhibition of either miR-21 or miR-33a production was overcome in reactions with TEL22



**Fig. 5** G-quadruplexes inhibit cleavage of pre-miRNA by hDicer in a dose-dependent manner. **a, b** Inhibition assay to assess the effect of RNA G-quadruplexes on the cleavage of pre-mir-21 (**a**) or pre-mir-33a (**b**) by hDicer. Reactions were carried out for 30 min at 37 °C under the low-turnover conditions. LIN12—a control 12-mer not adopting a G-quadruplex structure, C– a sample with no protein, nor inhibitor added, C+ a sample with hDicer, without inhibitor. Graphs show miRNA production efficiency normalized to the level of miRNA generated in C+. Error bars represent SD from three separate experiments. See also Figure S4B, C for full gel images. The reproducible results were obtained using at least two batches of recombi-

nant hDicer. **c, d** The results of the inhibition assay performed similarly as in (**a, b**) for selected DNA G-quadruplexes and pre-mir-21 (**c**) or pre-mir-33a (**d**). See also Figure S4D, E for full gel images. The reproducible results were obtained using at least two batches of recombinant hDicer. **e, f** Quantitative analysis of the time course of hDicer inhibition by G-quadruplexes in reactions with pre-mir-21 (**e**) or pre-mir-33a (**f**). Error bars represent SD from three separate experiments; see also Figure S5A–D for representative gel images. The reproducible results were obtained using at least two batches of recombinant hDicer

(Supplementary Fig. S6). Similar results as for TEL22 were obtained for pre-mir-16-1 competitor, therefore, we conclude that under the applied reaction conditions (1:2 molar ratio of

a substrate and an inhibitor), TEL22 can act as a competitive inhibitor. This conclusion is supported by the  $K_d$  values calculated for the PPC and TEL22 complex ( $K_d \sim 318$  nM),

and the PPC and pre-miRNA complex ( $K_d \sim 205$  nM). However, considering the binding affinities, we cannot explain why G4T4G4 was not outcompeted by the substrate over the incubation time. TEL22 represents the human telomeric sequence, while G4T4G4 corresponds to ciliate *O. nova* telomeric DNA. Given that the human enzyme was used in the studies, a species-specific regulatory mechanism for Dicer binding to telomeric DNA might be responsible for the effects we observed.

Under the low-turnover conditions, we did observe a difference in the effect of pre-mir-16-1 competitor on pre-mir-21 and pre-mir-33a cleavage by hDicer (Fig. 5e, f). Despite the addition of 100-fold molar excess of pre-mir-16-1 with respect to the other pre-miRNA, after 12 h incubation,  $\sim 80\%$  of pre-mir-33a (Fig. 5f) and only  $\sim 22\%$  of pre-mir-21 were processed (Fig. 5e). These results can be explained by the differences among the structures of the three pre-miRNAs. Pre-mir-21 adopts the most compact structure, with the smallest terminal loop, whereas structures of pre-mir-33a and pre-mir-16-1 contain large internal loops and bulges. In addition, pre-mir-16-1 has the most relaxed terminal loop region (Supplementary Fig. S5E). It has been reported that the pre-miRNA structure influences the efficiency of miRNA processing by Dicer [26]. In addition, the results of our previous studies have indicated that pre-miRNAs can compete for binding to hDicer [15]. Based on these data and the models generated for PPC and individual G-quadruplexes (Fig. 3a–d), we propose that a pre-miRNA and a G-quadruplex compete for binding to substrate-anchoring domains of hDicer, i.e., PAZ and Platform. Since the interactions between pre-miRNA and Dicer encompass not only the PPC region of Dicer, but other Dicer domains as well [4, 6] (Fig. 6a), we hypothesize that the competition between two pre-miRNAs for binding to Dicer is more complex, compared with the case involving a pre-miRNA and a G-quadruplex. Consequently, we deduce that under the low-turnover conditions, even despite the high excess of the inhibitor, pre-mir-33a can outcompete pre-mir-16-1 from binding to hDicer, which was not observed for pre-mir-21. However, under the high-turnover conditions, when the excess of a substrate to the enzyme was applied, and the substrate to inhibitor molar ratio was low (1:2), the degree of the miRNA production inhibition caused by pre-mir-16-1 was similar in the case of both substrates (Supplementary Fig. S6E, F).

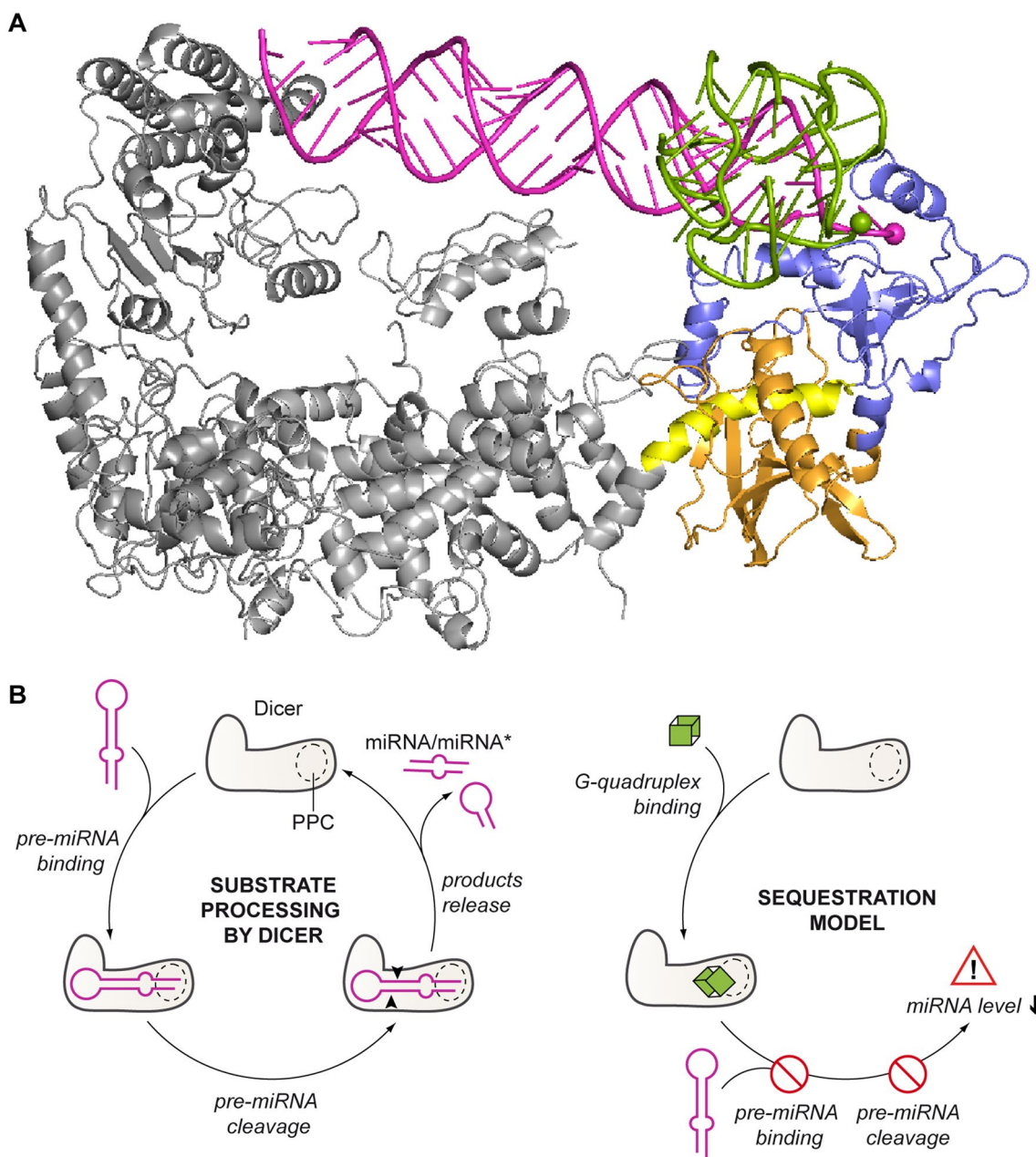
## Discussion

Dicer homologs are widely known for their important role in the biogenesis of small regulatory RNAs: miRNAs and siRNAs. However, growing evidence shows that, apart from their canonical role, Dicer proteins serve a number of

other functions (for review see [27]). As an endoribonuclease, Dicer participates in processing of diverse groups of RNA, e.g., tRNA [28], DNA-damage-induced RNA [29] or transposable elements [30]. Additionally, in *Caenorhabditis elegans*, a caspase-cleaved form of Dicer acts as a deoxyribonuclease that introduces breaks in chromosomal DNA during apoptosis [31]. Moreover, the emerging evidence points to possible cleavage-independent regulatory roles of Dicer; in *C. elegans* and human cells, Dicer binds various RNAs, including mRNAs and long noncoding RNAs, passively, i.e., without further cleavage [32]. Furthermore, hDicer can act as a nucleic acid annealer, by facilitating base pairing between two complementary RNAs or DNAs [5, 33]. Our finding that hDicer can bind RNA and DNA G-quadruplexes, and that this interaction influences Dicer ability to process its canonical substrates, opens new avenues of research on the cellular functions of Dicer.

A variety of proteins bind G-quadruplex-forming RNAs [34]. In most cases, however, it is not clear whether the binding involves the G-quadruplex structure or a guanine-rich RNA sequence. In the current study, we demonstrated that G-quadruplexes do not constitute a steric hindrance for hDicer binding (Fig. 1d, e). We also showed that the PPC cassette of hDicer binds RNA G-quadruplexes (Fig. 2b, c) with much higher affinity than it binds non-G-quadruplex RNAs (Supplementary Fig. S2B). Based on the models of the PPC and G-quadruplex complexes (Fig. 3a–d), and mutation studies (Fig. 3e), we propose that RNA and DNA molecules that adopt G-quadruplex structures can bind to either the 3'-pocket or the 5'-pocket of the hDicer PPC cassette. Although the results of our computational modeling indicated that binding of an RNA G-quadruplex by PPC may involve anchoring the 3' end of the oligonucleotide (Fig. 3a, b), we found that a 5'–3'-end-ligated G-quadruplex can be bound by PPC as well (Fig. 3f). Future work will reveal whether Dicer binds G-quadruplex-containing nucleic acids in cells and what the biological significance of this activity is.

Our results indicate that hDicer can bind both RNA and DNA G-quadruplexes. A search of the transcriptome-wide map of hDicer targets [32], using Quadron software [35], revealed that Dicer binds to several guanine-rich RNAs (Dataset S1). We can speculate about the possible implications of interactions between Dicer and an RNA G-quadruplex. For example, G-quadruplexes may act as molecular anchors that bind and sequester Dicer. A similar role has already been proposed for another type of RNA structure recognized by Dicer, i.e., RNA hairpins formed within endogenous transcripts [32] or adopted by viral RNAs [36]. Short RNAs, including miRNAs [37], piRNAs [38], and tRNA fragments [39, 40], can adopt G-quadruplex structures. The length of these RNAs ( $\sim 20$ – $30$  nt) is close to the length of the guanine-rich oligonucleotides used in the



**Fig. 6** Both pre-miRNA and G-quadruplexes are anchored within the same region of hDicer PPC cassette. **a** Superposed structures of hDicer in complex with a pre-miRNA substrate (PDB entry 5ZAL, does not include 15 nt from the apical loop of the pre-miRNA) and PPC cassette in complex with TER10. The Platform is colored in orange, PAZ in blue and Connector helix in yellow, the remaining portion of hDicer is presented in grey; pre-miRNA—in magenta, and TER10—in green. 3' ends of pre-miRNA and TER10 within the

3'-binding pocket of PAZ are marked as spheres. **b** Putative mechanism of sequestration-dependent regulation of Dicer activity by G-quadruplexes. Dicer anchors pre-miRNA ends within the PPC binding pockets and cleaves precursors to release miRNA products (left). G-quadruplexes compete with pre-miRNA for binding to Dicer. The enzyme sequestered in complex with a G-quadruplex cannot bind pre-miRNA and does not generate miRNA (right)

present study (10–22 nt). Accordingly, we hypothesize that short RNAs adopting G-quadruplex structures may sequester Dicer in vivo. Such a mechanism could be a part of autoregulatory loops, in which Dicer-processed guanine-rich miRNAs, in the form of G-quadruplexes, act in a negative feedback regulation of Dicer [17, 41]. The sequestration of

Dicer by RNA G-quadruplexes may control the pool of Dicer available for the biogenesis of miRNA. This hypothesis is supported by the results of the inhibition assay presented in Fig. 5, Supplementary Fig. S5A–D and S6. The observed RNA G-quadruplex-driven inhibition of pre-miRNA cleavage by hDicer was very efficient and stable over time. Since,



under the applied reaction conditions, the binding affinity of hDicer PPC for the RNA G-quadruplex ( $K_d \sim 7\text{--}10$  nM) (Fig. 2b, c) was much higher than the binding affinity of hDicer PPC for the pre-miRNA ( $K_d \sim 205$  nM) (Supplementary Fig. S2B), we conclude that an RNA G-quadruplex, once bound to hDicer, occupies the substrate binding pockets of PPC thereby blocking access for a pre-miRNA (Fig. 6b). This is in contrast to a pre-miRNA substrate that, once bound and cleaved by hDicer, leaves the binding pockets, making them available for another substrate molecule.

Proteins binding RNA G-quadruplexes are important for the intracellular transport of mRNA for its local translation [42]. Moreover, they act as a switch that controls the access of miRNA to its mRNA targets embedded in guanine-rich regions adopting G-quadruplex structures [43, 44]. Since Dicer binds both miRNA [32] and RNA G-quadruplexes, it may be directly involved in the control of translation of mRNAs that form local G-quadruplex structures. Possible direct involvement of Dicer in the posttranscriptional control of gene expression has been already extensively investigated in the context of interactions between miRNA-bound Dicer and complementary mRNA targets [33].

Dicer has been detected in the nucleus [45], which together with our finding that hDicer binds to DNA G-quadruplexes, leads us to propose that Dicer may interact with telomeric DNA G-quadruplexes or be involved in the transcriptional regulation of gene expression by binding to G-quadruplexes in promoter regions.

Altogether, we have demonstrated that G-quadruplexes can bind to Dicer and affect its activity. The results reported here support the notion that Dicer is a versatile protein, whose function is not limited to RNase III activity.

## Materials and methods

### Oligonucleotides

Sequences of all oligonucleotides used in this study are listed in Supplementary Table S2.

### Folding of oligonucleotides

Guanine-rich oligonucleotides were denatured in 50 mM KCl for 3 min at 90 °C, immediately transferred to 75 °C and slowly cooled down to 10 °C.

For the RNA duplexes used in the binding assays, 5'-<sup>32</sup>P-labeled oligomers were hybridized in water with 10 pmol of complementary oligomers by heating and slowly cooling the mixtures from 90 to 4 °C. The reaction mixtures were PAGE-purified in 12% native polyacrylamide gels to obtain double-stranded fractions free of single-stranded species.

### In-gel G-quadruplex staining

200 pmol of nonlabeled oligomers were folded in 20  $\mu$ L of 50 mM KCl, mixed with 20  $\mu$ L of loading buffer, divided into two equal portions (20  $\mu$ L) and separately loaded on a 12% polyacrylamide gel supplemented with 5% glycerol. The samples were arranged in duplicate so that after electrophoresis, the gel could be vertically cut into halves containing the same samples. One half was then stained in a solution of SYBR Gold (Thermo Fisher Scientific)/1 $\times$  TBE for 20 min and scanned with Fujifilm FLA-5100 Fluorescent Image Analyzer at 473 nm. The other half was stained with *N*-methyl mesoporphyrin IX (NMM; Frontier 6 Scientific), as described previously [46]. Briefly, NMM stock (5 mg/mL in 0.2 M HCl) was diluted in 20 mM Tris pH 7.6, 100 mM KCl, 1 mM EDTA to a final NMM concentration of 1  $\mu$ g/mL (NMM staining solution). Next, the gel was incubated in the NMM staining solution for 20 min. Subsequently, the gel was scanned at 532 nm.

### *o*-BMVC labeling of oligonucleotides

In-house synthesized C6-amino-modified oligo dissolved in sodium phosphate buffer (0.1 M, pH 7.0) was combined with fivefold molar excess of in-house made *o*-BMVC-C3-NHS ester in DMSO. The mixture was incubated at 37 °C in the dark overnight. The oligo-*o*-BMVC-C3 conjugate was separated from salt and free *o*-BMVC-C3-NHS by 2% NaClO<sub>4</sub>/acetone precipitation. The purity and homogeneity of the oligo-*o*-BMVC-C3 conjugate were verified by denaturing gel electrophoresis.

### Preparation of the 5'–3'-end-ligated TER22

The 5'-<sup>32</sup>P-labeled TER22 was denatured in 100 mM KCl for 3 min at 90 °C immediately transferred to 75 °C and slowly cooled down to 10 °C to adopt its native structure. The ligation reaction was carried out using T4 RNA Ligase 1 (Thermo Fisher Scientific) overnight at 17 °C. RNA was purified using NucAway Spin Columns (Thermo Fisher Scientific) according to the manufacturers protocol and resuspended in water to the final concentration of approximately 10,000 cpm/ $\mu$ L (50 nM). To verify the efficiency of TER22 circularization, RNA (10,000 cpm) was treated with 0.5 U of alkaline phosphatase (Thermo Fisher Scientific) according to the manufacturers protocol and analyzed by denaturing PAGE.

### RNA/DNA–protein binding assay

RNA/DNA–protein complex formation was analyzed using an electrophoretic mobility shift assay (EMSA). The reactions were carried out in 10 or 20- $\mu$ L volumes. PPC or PPC

variants (total protein concentration ranging from 0.1 to 400 nM (unless stated otherwise in the figure legend) were mixed with a trace amount of the 5'-<sup>32</sup>P-labeled RNA/DNA oligonucleotide (approximately 10,000 cpm) or *o*-BMVC-labeled oligonucleotide (from 7.5 to 30 μM, as indicated) and incubated in binding buffer (20 mM Tris-HCl, pH 7.4, 100 mM KCl) for 30 min (unless stated otherwise) on ice. The samples were separated in 5% or 10% nondenaturing polyacrylamide gels (as indicated) at 4 °C for approximately 10 h at 7 V/cm in 1× TBE. For radiolabeled oligomers, gels were exposed to a phosphorimager plate, which was subsequently scanned with Fujifilm FLA-5100 Fluorescent Image Analyzer to visualize the bands. For *o*-BMVC-labeled oligomers, gels were first scanned with Fujifilm FLA-5100 Fluorescent Image Analyzer at 532 nm, then stained in a solution of SYBR Gold (Thermo Fisher Scientific)/1× TBE for 20 min and scanned again at 473 nm.

### hDicer cleavage inhibition assay

Dicer cleavage inhibition assay was performed in 10 μL reactions containing 20 mM Tris-HCl, pH 7.5, 100 mM KCl, and 2.5 mM MgCl<sub>2</sub> buffer, 10 nM hDicer, 5'-<sup>32</sup>P-labeled pre-miRNA (10,000 cpm, approximately 5 nM) and the indicated oligonucleotide (0.1, 0.5, 2 μM). In addition, two control reactions were carried out: (i) a negative control (C<sup>-</sup>) with no enzyme and no inhibitor, to test the integrity of the substrate during the incubation time, and (ii) a positive control (C<sup>+</sup>) with enzyme but no inhibitor added. All samples were incubated at 37 °C for 30 min. The reactions were halted by adding 1 volume of 8 M urea loading buffer and heating for 5 min at 95 °C, and then separated in a 15% polyacrylamide gel with 7 M urea and 1× TBE.

In the time-course assay performed under the low-turnover conditions, a 10 μL reaction contained reaction buffer as described above, 10 nM hDicer, 5'-<sup>32</sup>P-labeled pre-miRNA (10,000 cpm, approximately 5 nM), and the indicated oligonucleotide (0.5 μM). Samples were incubated at 37 °C for 30 min, 1.5 h, 3 h, 6 h, or 12 h. In addition, two negative control reactions were carried out to monitor the integrity of the substrate during the incubation time: (i) a sample frozen immediately after the mixture was prepared (C<sub>0</sub>) and (ii) a sample incubated at 37 °C for 12 h (C<sub>12</sub>). After the incubation time, all of the samples were processed as described above.

For the high-turnover conditions, the time-course assay was performed similarly, except that a 10 μL reaction contained 0.5 nM hDicer, 25 nM 5'-<sup>32</sup>P-labeled pre-miRNA, and 50 nM oligonucleotide.

### Gel imaging and data analysis

For binding and cleavage inhibition assays, the data were collected using Fujifilm FLA-5100 Fluorescent Image Analyzer and quantified using MultiGauge 3.0 software (Fujifilm). Data from binding experiments were fit to a saturation isotherm using Prism 8.2.1 (GraphPad). Dissociation constants were calculated using the equation for one site, specific binding:  $Y = B_{\max} * X / (K_d + X)$  and  $B_{\max} = 100\%$  set as a constraint;  $K_d$  values together with the associated standard errors are reported in the figures. Diagrams presenting data from cleavage assays were prepared in Excel 2016. In the case of all diagrams, error bars represent SD values calculated based on three independent experiments.

### Production and purification of hDicer PPC cassette and the 3'-pocket and the 5'-pocket PPC variants

The PPC cDNA, which corresponds to the 316-amino acid (aa) sequence located between 753 and 1068 aa of hDicer, was amplified by PCR using a purchased plasmid encoding a complete *Homo sapiens* Dicer1 ribonuclease type III sequence (PubMed, NM\_030621) (GeneCopoeia). In the case of PPC variants, we used plasmids encoding hDicer 3'-pocket double mutant (Y926F, R927A), and the 5'-pocket sextuple mutant (R778A, R780A, R811A, H982A, R986A, R993A) [23] (Addgene). PCR fragments were subsequently cloned into the pMCSG7 vector (courtesy of Laboratory of Protein Engineering, Institute of Bioorganic Chemistry, Polish Academy of Sciences), which allows introduction of a 6xHis tag at the N-terminus of the protein. hDicer PPC and the PPC variants were expressed in *E. coli* strain BL21Star (Thermo Fisher Scientific), in standard Luria-Bertani (LB) medium. Gene expression was induced with 0.4 mM IPTG and bacteria were cultured for 18 h at 18 °C with shaking. The cell pellets were lysed and purified with Ni<sup>2+</sup>-Sepharose High-Performance beads (GE Healthcare) with imidazole gradient (0.02–1 M) in 25 mM HEPES (pH 8.0) (in the case of PPC), or 25 mM Tris (pH 8.0) (in the case of PPC variants), supplemented with 300 mM NaCl, 0.1% Triton X-100, and 5% glycerol. The protein purity was assessed by SDS-PAGE. Selected fractions containing homogeneous protein were concentrated using Amicon filters (Merck) in respective buffers (as described above) enriched with 40% glycerol, and stored at -20 °C.

### Docking and molecular modeling

Docking of RNA and the protein was performed following a meta-approach using different docking methods to generate the docking poses, followed by the rescoring and selection of best poses [47]. The hDicer (PDB entry 4NGF) and G-quadruplex (PDB entry 2M18) were docked using the following

methods: 3dRPC/RPDOCK [48], ClusPro [49], HADDOCK [50], HDock [51], Hex [52], PatchDock [53] and ZDOCK [54]. While 3dRPC and HADDOCK have modules to dock protein and RNA explicitly, ClusPro allows only the use of RNA as a receptor molecule. For the methods HDock, Hex, PatchDock, and ZDOCK the docking was performed both with protein as receptor and RNA as ligand and vice versa. In addition, the protein–RNA complex structures were generated with an in-house method SimRNP developed by the Bujnicki group, an extension of SimRNA [55] that enables flexible modeling of RNA structures as well as protein–RNA complexes (Michał Boniecki and J.M.B., unpublished). Additional information on modeling is available in the Supplementary Methods. The simulation trajectory is presented in the Supplementary Fig. S7, and the representative frame is provided as PDB file (Supplementary File S1). PDB files for the reported models have been deposited in Figshare and can be accessed under the following URL: <https://figshare.com/s/1dfdb3d2945d404e8065>.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00018-021-03795-w>.

**Acknowledgments** We would like to thank prof. Ryszard Kierzek for providing G-rich oligomers, Dr. Anna Urbanowicz for her technical support in RNA-protein binding studies, and Life Science Editors for editing services. This work was supported by the National Science Centre, Poland [2016/22/E/NZ1/00422 to A.K.-K., 2017/26/A/NZ1/01083 to J.M.B., 2017/01/X/ST5/00577 to D.B. and 2017/01/X/ST5/00589 to D.G.], the Polish Ministry of Science and Higher Education [KNOW program for years 2014–2018] and the IIMCB statutory funds [to J.M.B. and C.N.]. Computational analyses were performed using the resources of IIMCB, the Poznan Supercomputing and Networking Center at the Institute of Bioorganic Chemistry, Polish Academy of Sciences [grant 312 to J.M.B.], the Polish Grid Infrastructure [grant rnpmc to C.N.] and the Interdisciplinary Centre for Mathematical and Computational Modelling at the University of Warsaw [grants G73-4 and GB76-30 to C.N.]. Funding for open access charge: [2016/22/E/NZ1/00422 to A.K.-K.].

**Author contributions** N.K., A.S., K.C., M.W., M.P., M.C.M., D.G., and D.B. performed experiments; C.N. and J.M.B. performed computational analyses and modeling; all authors analyzed the data, and interpreted the results; J.M.B., Z.G., M.F., and A.K.K. supervised and provided advice; N.K. and A.K.K. wrote the manuscript with input from all other authors.

**Availability of data and material** The data that support the findings of this study are openly available in figshare at <https://doi.org/10.6084/m9.figshare.12085938>. The materials used in this study are available from the corresponding author, A.K.-K., upon reasonable request.

## Declarations

**Conflict of interest** The authors declare no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source,

provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Court DL, Gan J, Liang YH, Shaw GX, Tropea JE, Costantino N, Waugh DS, Ji X (2013) RNase III: genetics and function; structure and mechanism. *Annu Rev Genet* 47:405–431. <https://doi.org/10.1146/annurev-genet-110711-155618>
2. Lau PW, Guiley KZ, De N, Potter CS, Carragher B, MacRae IJ (2012) The molecular architecture of human Dicer. *Nat Struct Mol Biol* 19(4):436–440. <https://doi.org/10.1038/nsmb.2268>
3. Tian Y, Simanshu DK, Ma JB, Park JE, Heo I, Kim VN, Patel DJ (2014) A phosphate-binding pocket within the platform–PAZ–connector helix cassette of human dicer. *Mol Cell* 53(4):606–616. <https://doi.org/10.1016/j.molcel.2014.01.003>
4. Taylor DW, Ma E, Shigematsu H, Cianfrocco MA, Noland CL, Nagayama K, Nogales E, Doudna JA, Wang H-W (2013) Substrate-specific structural rearrangements of human Dicer. *Nat Struct Mol Biol* 20(6):662–670. <https://doi.org/10.1038/nsmb.2564>
5. Kurzynska-Kokorniak A, Pokornowska M, Koralewska N, Hoffmann W, Bienkowska-Szewczyk K, Figlerowicz M (2016) Revealing a new activity of the human Dicer DUF283 domain in vitro. *Sci Rep* 6:23989. <https://doi.org/10.1038/srep23989>
6. Liu Z, Wang J, Cheng H, Ke X, Sun L, Zhang QC, Wang H-W (2018) Cryo-EM structure of human dicer and its complexes with a pre-miRNA substrate. *Cell* 173(5):1191–1203.e1112. <https://doi.org/10.1016/j.cell.2018.03.080>
7. Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W (2004) Single processing center models for human Dicer and bacterial RNase III. *Cell* 118(1):57–68. <https://doi.org/10.1016/j.cell.2004.06.017>
8. Wostenberg C, Lary JW, Sahu D, Acevedo R, Quarles KA, Cole JL, Showalter SA (2012) The role of human Dicer-dsRBD in processing small regulatory RNAs. *PLOS One* 7(12):e51829. <https://doi.org/10.1371/journal.pone.0051829>
9. Kwok CK, Merrick CJ (2017) G-Quadruplexes: prediction, characterization, and biological application. *Trends Biotechnol* 35(10):997–1013. <https://doi.org/10.1016/j.tibtech.2017.06.012>
10. Henderson E, Hardin CC, Walk SK, Tinoco I Jr, Blackburn EH (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine–guanine base pairs. *Cell* 51(6):899–908. [https://doi.org/10.1016/0092-8674\(87\)90577-0](https://doi.org/10.1016/0092-8674(87)90577-0)
11. Hansel-Hertsch R, Di Antonio M, Balasubramanian S (2017) DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat Rev Mol Cell Biol* 18(5):279–284. <https://doi.org/10.1038/nrm.2017.3>
12. Mirihana Arachchilage G, Dassanayake AC, Basu S (2015) A potassium ion-dependent RNA structural switch regulates human pre-miRNA 92b maturation. *Chem Biol* 22(2):262–272. <https://doi.org/10.1016/j.chembiol.2014.12.013>
13. Pandey S, Agarwala P, Jayaraj GG, Gargallo R, Maiti S (2015) The RNA stem-loop to G-quadruplex equilibrium controls mature microRNA production inside the cell. *Biochemistry*



- 54(48):7067–7078. <https://doi.org/10.1021/acs.biochem.5b00574>
14. Rouleau SG, Garant JM, Bolduc F, Bisailon M, Perreault JP (2018) G-Quadruplexes influence pri-microRNA processing. *RNA Biol* 15(2):198–206. <https://doi.org/10.1080/15476286.2017.1405211>
  15. Tyczewska A, Kurzynska-Kokorniak A, Koralewska N, Szopa A, Kietrys AM, Wrzesinski J, Twardowski T, Figlerowicz M (2011) Selection of RNA oligonucleotides that can modulate human dicer activity in vitro. *Nucleic Acid Ther* 21(5):333–346. <https://doi.org/10.1089/nat.2011.0304>
  16. Kurzynska-Kokorniak A, Koralewska N, Tyczewska A, Twardowski T, Figlerowicz M (2013) A new short oligonucleotide-based strategy for the precursor-specific regulation of microRNA processing by dicer. *PLoS One* 8(10):e77703. <https://doi.org/10.1371/journal.pone.0077703>
  17. Koralewska N, Hoffmann W, Pokornowska M, Milewski M, Lipinska A, Bienkowska-Szewczyk K, Figlerowicz M, Kurzynska-Kokorniak A (2016) How short RNAs impact the human ribonuclease Dicer activity: putative regulatory feedback-loops and other RNA-mediated mechanisms controlling microRNA processing. *Acta Biochim Pol* 63(4):773–783. [https://doi.org/10.18388/abp.2016\\_1339](https://doi.org/10.18388/abp.2016_1339)
  18. Bellaousov S, Reuter JS, Seetin MG, Mathews DH (2013) RNA structure: web servers for RNA secondary structure prediction and analysis. *Nucleic Acids Res* 41(Web Server issue):W471–W474. <https://doi.org/10.1093/nar/gkt290>
  19. Nicoludis JM, Barrett SP, Mergny JL, Yatsunyk LA (2012) Interaction of human telomeric DNA with N-methyl mesoporphyrin IX. *Nucleic Acids Res* 40(12):5432–5447. <https://doi.org/10.1093/nar/gks152>
  20. Liu H, Matsugami A, Katahira M, Uesugi S (2002) A dimeric RNA quadruplex architecture comprised of two G:G(:A):G:G(:A) hexads, G:G:G: G tetrads and UUUU loops. *J Mol Biol* 322(5):955–970. [https://doi.org/10.1016/S0022-2836\(02\)00876-8](https://doi.org/10.1016/S0022-2836(02)00876-8)
  21. Martadinata H, Phan AT (2013) Structure of human telomeric RNA (TERRA): stacking of two G-quadruplex blocks in K(+) solution. *Biochemistry* 52(13):2176–2183. <https://doi.org/10.1021/bi301606u>
  22. Haider S, Parkinson GN, Neidle S (2002) Crystal structure of the potassium form of an *Oxytricha nova* G-quadruplex. *J Mol Biol* 320(2):189–200. [https://doi.org/10.1016/S0022-2836\(02\)00428-X](https://doi.org/10.1016/S0022-2836(02)00428-X)
  23. Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN (2011) Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* 475(7355):201–205. <https://doi.org/10.1038/nature10198>
  24. Kiliszek A, Blaszczyk L, Kierzek R, Rypniewski W (2017) Stabilization of RNA hairpins using non-nucleotide linkers and circularization. *Nucleic Acids Res* 45(10):e92. <https://doi.org/10.1093/nar/gkx122>
  25. Tseng TY, Chien CH, Chu JF, Huang WC, Lin MY, Chang CC, Chang TC (2013) Fluorescent probe for visualizing guanine-quadruplex DNA by fluorescence lifetime imaging microscopy. *J Biomed Opt* 18(10):101309. <https://doi.org/10.1117/1.JBO.18.10.101309>
  26. Starega-Roslan J, Koscianska E, Kozlowski P, Krzyzosiak WJ (2011) The role of the precursor structure in the biogenesis of microRNA. *Cell Mol Life Sci* 68(17):2859–2871. <https://doi.org/10.1007/s00018-011-0726-2>
  27. Kurzynska-Kokorniak A, Koralewska N, Pokornowska M, Urbanowicz A, Tworak A, Mickiewicz A, Figlerowicz M (2015) The many faces of Dicer: the complexity of the mechanisms regulating Dicer gene expression and enzyme activities. *Nucleic Acids Res* 43(9):4365–4380. <https://doi.org/10.1093/nar/gkv328>
  28. Cole C, Sobala A, Lu C, Thatcher SR, Bowman A, Brown JW, Green PJ, Barton GJ, Hutvagner G (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* 15(12):2147–2160. <https://doi.org/10.1261/rna.1738409>
  29. Francia S, Michelini F, Saxena A, Tang D, de Hoon M, Anelli V, Mione M, Carninci P, d'Adda di Fagnana F (2012) Site-specific DICER and DROSHA RNA products control the DNA-damage response. *Nature* 488(7410):231–235. <https://doi.org/10.1038/nature11179>
  30. Kaneko H, Dridi S, Tarallo V, Gelfand BD, Fowler BJ, Cho WG, Kleinman ME, Ponicsan SL, Hauswirth WW, Chiodo VA, Kariko K, Yoo JW, Lee DK, Hadzhimetovic M, Song Y, Misra S, Chaudhuri G, Buaas FW, Braun RE, Hinton DR, Zhang Q, Grossniklaus HE, Provis JM, Madigan MC, Milam AH, Justice NL, Albuquerque RJ, Blandford AD, Bogdanovich S, Hirano Y, Witta J, Fuchs E, Littman DR, Ambati BK, Rudin CM, Chong MM, Provost P, Kugel JF, Goodrich JA, Dunaief JL, Baffi JZ, Ambati J (2011) DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. *Nature* 471(7338):325–330. <https://doi.org/10.1038/nature09830>
  31. Nakagawa A, Shi Y, Kage-Nakadai E, Mitani S, Xue D (2010) Caspase-dependent conversion of Dicer ribonuclease into a death-promoting deoxyribonuclease. *Science* 328(5976):327–334. <https://doi.org/10.1126/science.1182374>
  32. Rybak-Wolf A, Jens M, Murakawa Y, Herzog M, Landthaler M, Rajewsky N (2014) A variety of dicer substrates in human and *C. elegans*. *Cell* 159(5):1153–1167. <https://doi.org/10.1016/j.cell.2014.10.040>
  33. Pokornowska M, Milewski MC, Ciechanowska K, Szczepanska A, Wojnicka M, Radogostowicz Z, Figlerowicz M, Kurzynska-Kokorniak A (2019) The RNA–RNA base pairing potential of human Dicer and Ago2 proteins. *Cell Mol Life Sci*. <https://doi.org/10.1007/s00018-019-03344-6>
  34. von Hacht A, Seifert O, Menger M, Schutze T, Arora A, Konthur Z, Neubauer P, Wagner A, Weise C, Kurreck J (2014) Identification and characterization of RNA guanine–quadruplex binding proteins. *Nucleic Acids Res* 42(10):6630–6644. <https://doi.org/10.1093/nar/gku290>
  35. Sahakyan AB, Chambers VS, Marsico G, Santner T, Di Antonio M, Balasubramanian S (2017) Machine learning model for sequence-driven DNA G-quadruplex formation. *Sci Rep* 7(1):14535. <https://doi.org/10.1038/s41598-017-14017-4>
  36. Andersson MG, Haasnoot PC, Xu N, Berenjian S, Berkhout B, Akusjarvi G (2005) Suppression of RNA interference by adenovirus virus-associated RNA. *J Virol* 79(15):9556–9565. <https://doi.org/10.1128/JVI.79.15.9556-9565.2005>
  37. Chan KL, Peng B, Umar MI, Chan CY, Sahakyan AB, Le MTN, Kwok CK (2018) Structural analysis reveals the formation and role of RNA G-quadruplex structures in human mature microRNAs. *Chem Commun (Camb)* 54(77):10878–10881. <https://doi.org/10.1039/c8cc04635b>
  38. Balaratnam S, Hettiarachchilage M, West N, Piontkivska H, Basu S (2019) A secondary structure within a human piRNA modulates its functionality. *Biochimie* 157:72–80. <https://doi.org/10.1016/j.biochi.2018.11.002>
  39. Lyons SM, Gudanis D, Coyne SM, Gdaniec Z, Ivanov P (2017) Identification of functional tetramolecular RNA G-quadruplexes derived from transfer RNAs. *Nat Commun* 8(1):1127. <https://doi.org/10.1038/s41467-017-01278-w>
  40. Jackowiak P, Hojka-Osinska A, Gasiorek K, Stelmaszczuk M, Gudanis D, Gdaniec Z, Figlerowicz M (2017) Effects of G-quadruplex topology on translational inhibition by tRNA fragments in mammalian and plant systems in vitro. *Int J Biochem Cell Biol* 92:148–154. <https://doi.org/10.1016/j.biocel.2017.10.001>

41. Durica-Mitic S, Gorke B (2019) Feedback regulation of small RNA processing by the cleavage product. *RNA Biol* 16(8):1055–1065. <https://doi.org/10.1080/15476286.2019.1612693>
42. Ishiguro A, Kimura N, Watanabe Y, Watanabe S, Ishihama A (2016) TDP-43 binds and transports G-quadruplex-containing mRNAs into neurites for local translation. *Genes Cells* 21(5):466–481. <https://doi.org/10.1111/gtc.12352>
43. Stefanovic S, Bassell GJ, Mihailescu MR (2015) G quadruplex RNA structures in PSD-95 mRNA: potential regulators of miR-125a seed binding site accessibility. *RNA* 21(1):48–60. <https://doi.org/10.1261/rna.046722.114>
44. DeMarco B, Stefanovic S, Williams A, Moss KR, Anderson BR, Bassell GJ, Mihailescu MR (2019) FMRP-G-quadruplex mRNA–miR-125a interactions: implications for miR-125a mediated translation regulation of PSD-95 mRNA. *PLoS One* 14(5):e0217275. <https://doi.org/10.1371/journal.pone.0217275>
45. Burger K, Schlackow M, Potts M, Hester S, Mohammed S, Gullerova M (2017) Nuclear phosphorylated Dicer processes double-stranded RNA in response to DNA damage. *J Cell Biol* 216(8):2373–2389. <https://doi.org/10.1083/jcb.201612131>
46. Booy EP, Meier M, Okun N, Novakowski SK, Xiong S, Steinfeld J, McKenna SA (2012) The RNA helicase RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the formation of the P1 helix template boundary. *Nucleic Acids Res* 40(9):4110–4124. <https://doi.org/10.1093/nar/gkr1306>
47. Nithin C, Ghosh P, Bujnicki JM (2018) Bioinformatics tools and benchmarks for computational docking and 3D structure prediction of RNA–protein complexes. *Genes* 9(9):432. <https://doi.org/10.1038/srep01887>
48. Huang Y, Liu S, Guo D, Li L, Xiao Y (2013) A novel protocol for three-dimensional structure prediction of RNA–protein complexes. *Sci Rep* 3:1887. <https://doi.org/10.1038/srep01887>
49. Comeau SR, Gatchell DW, Vajda S, Camacho CJ (2004) ClusPro: a fully automated algorithm for protein–protein docking. *Nucleic Acids Res* 32(suppl\_2):W96–W99. <https://doi.org/10.1093/nar/gkh354>
50. van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastiris PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ, Bonvin AMJJ (2016) The HADDOCK22 web server: user-friendly integrative modeling of biomolecular complexes. *J Mol Biol* 428(4):720–725. <https://doi.org/10.1016/j.jmb.2015.09.014>
51. Yan Y, Zhang D, Zhou P, Li B, Huang S-Y (2017) HDock: a web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res* 45(W1):W365–W373. <https://doi.org/10.1093/nar/gkx407>
52. Macindoe G, Mavridis L, Venkatraman V, Devignes M-D, Ritchie DW (2010) HexServer: an FFT-based protein docking server powered by graphics processors. *Nucleic Acids Res* 38(suppl\_2):W445–W449. <https://doi.org/10.1093/nar/gkq311>
53. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res* 33(suppl 2):W363–W367. <https://doi.org/10.1093/nar/gki481>
54. Pierce BG, Wiehe K, Hwang H, Kim B-H, Vreven T, Weng Z (2014) ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. *Bioinformatics* 30(12):1771–1773. <https://doi.org/10.1093/bioinformatics/btu097>
55. Boniecki MJ, Lach G, Dawson WK, Tomala K, Lukasz P, Soltysinski T, Rother KM, Bujnicki JM (2015) SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic Acids Res* 44(7):e63. <https://doi.org/10.1093/nar/gkv1479>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.