

Uniwersytet im. Adama Mickiewicza w Poznaniu

Wydział Biologii



Ewa Stein

**Rozpoznawanie cząsteczek RNA
przez domenę FinO białka ProQ z bakterii *Escherichia coli***

Rozprawa doktorska wykonana w Pracowni Biochemii RNA

Pod kierunkiem prof. dra hab. Mikołaja Olejniczaka

Poznań 2022

Podziękowania

Składam serdeczne podziękowania:

Mojemu promotorowi **prof. drowi hab. Mikołajowi Olejniczakowi** za umożliwienie mi udziału w intrygującym projekcie, opiekę merytoryczną i wszelką pomoc.

Dr Joannie Kwiatkowskiej za owocną współpracę naukową i pomoc w sprawdzaniu pracy, a pozanaukowo: za przyjaźń, wsparcie i rozmowy przy drożdżówce.

Byłym i obecnym **doktorantom Pracowni Biochemii RNA**, wcześniej Zakładu Biochemii, za pouczające rozmowy i miłą atmosferę pracy, w szczególności zaś **mgr Joannie Zwolenkiewicz** i **mgr Marii Mamońskiej** za pomoc w sprawdzaniu pracy.

Ponadto z całego serca dziękuję **Rodzicom** i **siostrze Kasi** za nieustanne wsparcie i wiarę we mnie.

Finansowanie

Badania opisane w rozprawie doktorskiej były finansowane ze środków:

1. Narodowego Centrum Nauki – grant OPUS: 2018/31/B/NZ1/02612 „Mechanizmy rozpoznawania RNA przez białko ProQ oraz rola tego białka w oddziaływaniach pomiędzy niekodującymi RNA a regulowanymi mRNA u bakterii *Escherichia coli*”, kierownik projektu – prof. dr hab. Mikołaj Olejniczak.
2. Narodowego Centrum Nauki – grant OPUS: 2014/15/B/NZ1/03330 „Regulacja translacji na etapie elongacji przez niekodujące RNA bakterii”, kierownik projektu – prof. dr hab. Mikołaj Olejniczak.
3. Krajowego Naukowego Ośrodka Wiodącego (KNOW) Poznańskiego Konsorcjum RNA (01/KNOW2/2014).
4. Narodowego Centrum Badań i Rozwoju – projekt „Paszport do przyszłości - Interdyscyplinarne studia doktoranckie na Wydziale Biologii UAM” (POWR.03.02.00-00-I006/17).

Wykaz artykułów powstałych podczas realizacji pracy doktorskiej

Stein Ewa M., Kwiatkowska Joanna, Basczok Maciej M., Gravel Chandra M., Berry Katherine E. i Olejniczak Mikołaj (2020). *Determinants of RNA recognition by the FinO domain of the Escherichia coli ProQ protein*. Nucleic Acids Research, 48(13), 7502–7519, doi.org/10.1093/nar/gkaa497

Wartość współczynnika IF dla czasopisma Nucleic Acids Research w 2020 r.: **16.971**.

Spis treści

PODZIĘKOWANIA	2
FINANSOWANIE	3
WYKAZ ARTYKUŁÓW POWSTAŁYCH PODCZAS REALIZACJI PRACY DOKTORSKIEJ.....	4
SPIS TREŚCI.....	5
STRESZCZENIE	8
ABSTRACT	11
WYKAZ ZASTOSOWANYCH SKRÓTÓW I SYMBOLI	13
1. WSTĘP.....	15
1.1. WPROWADZENIE.....	15
1.2. REGULATOROWE RNA U BAKTERII.....	15
1.2.1. sRNA i ncRNA powstające na drodze obróbki transkryptów	16
1.3. BIAŁKA UCZESTNICZĄCE W KONTROLI ZALEŻNEJ OD SRNA	17
1.1.1. Hfq	17
1.1.2. Białka z rodziny FinO	19
1.1.2.1. FinO	19
1.1.2.2. Inne białka z rodziny FinO	20
1.1.2.3. ProQ.....	21
1.1.2.3.1.Odkrycie.....	21
1.1.2.3.2.Budowa	22
1.1.2.3.3.Funkcja.....	24
2. CELE PRACY	27
3. MATERIAŁY	28
3.1. ODCZYNNIKI I KOLUMNY DO CHROMATOGRAFII BIAŁEK	28
3.2. ENZYMY:	29
3.3. OLIGODEOKSYRYBONUKLEOTYDY.....	29
3.3.1. Oligonukleotydy DNA użyte do reakcji wydłużania starterów.....	29
3.3.2. Oligodeoksyrybonukleotydy użyte do przygotowania konstruktyw do nadekspresji białek	33
3.3.3. Oligodeoksyrybonukleotydy użyte do mutagenyzy	34
3.4. BUFORY I ROZTWORY	34
3.5. SZCZEPY I PLAZMIDY BAKTERYJNE	36
4. METODY	37
4.1. PRZYGOTOWANIE CZĄSTECZEK RNA	37
4.1.1. Przygotowanie matrycy DNA	37

4.1.1.1.	Oczyszczanie syntetycznych oligonukleotydów.....	37
4.1.1.2.	Reakcja wydłużania starterów	37
4.1.2.	Przygotowanie cząsteczek RNA.....	38
4.1.2.1.	Transkrypcja <i>in vitro</i>	38
4.1.2.2.	Oczyszczanie cząsteczek RNA w żelu poliakrylamidowym	39
4.1.2.3.	Znakowanie cząsteczki RNA izotopem radioaktywnego fosforu ³² P.....	39
4.2.	PRZYGOTOWANIE BIAŁEK.	40
4.2.1.	Przygotowanie konstruktów do nadekspresji białek	40
4.2.1.1.	Projektowanie konstruktów	40
4.2.1.2.	Reakcja łańcuchowa polimerazy (PCR, ang. <i>polymerase chain reaction</i>)	40
4.2.1.3.	Trawienie restrykcyjne	41
4.2.1.4.	Defosforylacja plazmidu	42
4.2.1.5.	Ligacja	42
4.2.2.	Metody mikrobiologiczne.....	42
4.2.2.1.	Transformacja kompetentnych komórek bakteryjnych	42
4.2.2.2.	Selekcja kolonii przy użyciu kolonijnej PCR	43
4.2.2.3.	Hodowla bakteryjna w celu nadekspresji białka.....	43
4.2.3.	Oczyszczanie białka	44
4.2.3.1.	Liza komórek bakteryjnych	44
4.2.3.2.	Chromatografia metalopowinowactwa – pierwszy etap	44
4.2.3.3.	Chromatografia powinowactwa	45
4.2.3.4.	Trawienie proteazą TEV	46
4.2.3.5.	Chromatografia metalopowinowactwa – drugi etap	46
4.2.3.6.	Sączenie molekularne	46
4.2.3.7.	Obliczanie stężenia białka.....	47
4.2.3.8.	Przechowywanie oczyszczonych białek	47
4.3.	METODY PRACY EKSPERYMENTALNEJ:	47
4.3.1.	Wyznaczanie stałej równowagi reakcji dysocjacji (K_d) wiązania RNA do białka.....	47
4.3.2.	Kompetycja o wiązanie białka.....	48
5.	WYNIKI.....	49
5.1.	OKREŚLENIE CECH CZĄSTECZKI RNA ISTOTNYCH DLA WIĄZANIA Z N-KOŃCOWĄ DOMENĄ BIAŁKA PROQ.....	49
5.1.1.	Wyznaczenie siły oddziaływania wybranych cząsteczek RNA z białkiem ProQ i jego wyizolowaną domeną FinO	49

5.1.2.	Udział struktury terminatora transkrypcji niezależnego od białka Rho w oddziaływaniu z N-kończącą domeną białka ProQ	54
5.1.3.	Znaczenie 3'-terminalnej lokalizacji struktury spinki z sekwencją poliU dla oddziaływania RNA z N-kończącą domeną białka ProQ.....	56
5.1.4.	Znaczenie długości 3'-terminalnego ogona poliurydynowego dla wiązania przez domenę FinO białka ProQ	58
5.1.5.	Znaczenie długości regionu dwuniciowego w obrębie spinki terminatora transkrypcji dla wiązania N-końcowej domeny białka ProQ.....	60
5.1.6.	Znaczenie motywu bogatego w adenozyliny zlokalizowanego po 5'-stronie spinki terminatorowej w kompetycji pomiędzy białkami ProQ i Hfq o wiązanie cząsteczek RNA	63
5.2.	POZNANIE ZNACZENIA RESZT AMINOKWASOWYCH ZLOKALIZOWANYCH W OBRĘBIE DOMENY FINO BIAŁKA PROQ ISTOTNYCH DLA WIĄZANIA CZĄSTECZEK RNA	71
5.2.1.	Analiza zakonserwowania sekwencji aminokwasowej pośród białek zawierających domenę FinO/ProQ u wybranych gatunków enterobakterii.....	71
5.2.2.	Wyznaczenie siły oddziaływania mutantów N-końcowej domeny białka ProQ z cząsteczkami RNA.....	74
5.2.3.	Cechy cząsteczek RNA potencjalnie wpływające na odmienne sposoby wiązania do mutantów K35A i R69A domeny FinO białka ProQ.....	86
5.2.3.1.	Znaczenie sekwencji zlokalizowanych po 5' i 3' stronie spinki terminatorowej cząsteczek RNA w różnicowym wiązaniu RNA przez mutanty N-końcowej domeny białka ProQ.....	86
5.2.3.2.	Udział dolnej części spinki terminatorowej w wiązaniu mutantów domeny FinO białka ProQ.....	88
5.2.3.3.	Znaczenie długości jednoniciowego fragmentu łączącego struktury drugorzędowe w obrębie cząsteczek RNA w oddziaływaniu z mutantami N-końcowej domeny białka ProQ.....	90
6.	DYSKUSJA.....	97
7.	WNIOSKI	105
	WYKAZ RYSUNKÓW.....	106
	WYKAZ TABEL	108
	BIBLIOGRAFIA	109

Streszczenie

U bakterii małe niekodujące RNA (sRNA) biorą udział w adaptacji do zmieniających się warunków środowiska. Oddziałując z białkami, cząsteczki sRNA wpływają na translację i stabilność regulowanych cząsteczek mRNA. Jednym z najlepiej poznanych białek zaangażowanych w regulację zależną od sRNA jest białko Hfq, jednak niedawne doniesienia wskazują także na istotną rolę rodziny białek z domeną FinO w tego typu regulacji. Białko ProQ, które występuje u *Escherichia coli*, jest jednym z przedstawicieli tej rodziny. ProQ zbudowane jest z dwóch domen: N-końcowej wiążącej RNA domeny FinO (NTD) i C-końcowej domeny Tudor (CTD), które są połączone nieustrukturyzowanym regionem o długości około 50 reszt aminokwasowych. Białko ProQ wiąże liczne cząsteczki RNA w komórce bakterii *E. coli*, jednak dotąd niewiele było wiadomo na temat tego, w jaki sposób cząsteczki RNA są rozpoznawane przez to białko.

Aby wyjaśnić, w jaki sposób białko ProQ rozpoznaje cząsteczki RNA, w pierwszej części mojej pracy doktorskiej zbadalam, jakie cechy sekwencji i struktury cząsteczek RNA są niezbędne dla ich wiązania przez domenę FinO białka ProQ, a w drugiej części pracy przeanalizowałam, w jaki sposób mutacje w domenie FinO wpływają na wiązanie różnych RNA. W pierwszej części badań określiłam cechy cząsteczek RNA decydujące o ich wiązaniu do NTD. Spośród puli RNA wiązanych przez ProQ *in vivo* wybrałam sześć cząsteczek: MicA, SibA, *cspE-3'*, *malM-3'*, *hupA-5'* i *lpp-5'*, i zmierzyłam ich wiązanie do oczyszczonego białka ProQ pełnej długości oraz do jego domeny FinO (NTD). Analiza wiązania cząsteczek RNA i ich mutantów wskazała, że elementem struktury rozpoznawanym przez NTD jest terminator transkrypcji niezależny od białka Rho. Cechami struktury terminatora, które mają znaczenie dla oddziaływania z NTD, jest znajdujący się na 3' końcu ogon poliurydynowy o długości co najmniej czterech nukleotydów, a także region dwuniciowy spinki o długości co najmniej dwóch par zasad. Wprowadzenie niesparowania powyżej czwartej pary zasad regionu dwuniciowego spinki terminatorowej nie wpływa na oddziaływanie z NTD. Ponadto stwierdziłam, że motyw sekwencji bogatej w adenozyne znajdujący się po 5'-stronie spinki terminatorowej cząsteczek RNA wiązanych przez ProQ odgrywa rolę w kompetycji pomiędzy ProQ, a innym białkiem wiążącym 3' koniec RNA jakim jest białko Hfq. Zauważyłam, że substytucja adenozyne na urydyny spowodowała, że cząsteczki RNA wiązały się do NTD słabiej albo tak samo jak do RNA o naturalnej sekwencji. Wyraźnej poprawie uległo za to

wiązanie zmutowanych cząsteczek RNA do białka Hfq. Sugeruje to, że motyw bogaty w adenozyliny zapobiega wiązaniu cząsteczek RNA do białka Hfq i ułatwia ich wiązanie przez ProQ.

W drugiej części badań w obrębie N-końcowej domeny białka ProQ wskazałam aminokwasy ważne dla oddziaływania z cząsteczkami RNA. W wyniku analizy zakonserwowania ewolucyjnego reszt aminokwasowych oraz w oparciu o dostępne dane literaturowe opisujące efekty mutacji białek należących do rodziny FinO, wytypowałam dziesięć reszt aminokwasowych i sprawdziłam ich wpływ na wiązanie siedmiu cząsteczek RNA *in vitro*. Są to RNA *cspE-3'*, *cspE81-3'*, *malM-3'*, *gapA-3'*, RybB, SibA i SibB. Mutacje wprowadzone do NTD w pozycjach R58, Y70 i R80 powodują zahamowanie wiązania białka do RNA. Istotne dla oddziaływania są także pozycje K54 i R62, których mutacje znacząco osłabiają wiązanie RNA, podczas gdy najmniejszy wpływ na wiązanie miały mutacje w pozycjach R32, D41 i T65. Co ciekawe mutacje reszt aminokwasowych K35 i R69 miały odmienny wpływ na wiązanie różnych cząsteczek RNA, najmniejszy na wiązanie cząsteczki *cspE-3'*, a największy na wiązanie sRNA SibB. Może to sugerować różnice w oddziaływaniach cząsteczek RNA z domeną FinO białka ProQ. Aby wskazać, które regiony cząsteczek RNA są odpowiedzialne za zróżnicowany wpływ mutacji K35A i R69A w NTD na wiązanie RNA, zbadalam wiązanie cząsteczek RNA zawierających mutacje pozbawiające ich regionów jednoniciowych sąsiadujących ze spinką terminatorową, zmieniające sekwencję spinki terminatorowej, albo zmieniające długość linkera pomiędzy spinką terminatorową a najbliższym elementem struktury dwuniciowej po 5' stronie spinki. Skutki wprowadzenia tych mutacji w RNA jednoznacznie nie wskazywały na bezpośredni kontakt zmutowanych sekwencji RNA z resztami aminokwasowymi w pozycjach 35 i 69, dlatego jest możliwe, że substytucje K35A i R69A wywierają pośredni wpływ na wiązanie RNA, np. poprzez zmianę konformacji białka.

Podsumowując, wyniki eksperymentów przeprowadzonych w ramach mojej pracy doktorskiej pokazały, że N-końcowa domena białka ProQ rozpoznaje cząsteczki RNA poprzez wiązanie się do nich w regionie połączenia pomiędzy dolną częścią spinki terminatora transkrypcji a otaczającymi ją sekwencjami jednoniciowymi. Ponadto badania skutków mutacji wprowadzonych do domeny FinO sugerują, że największe znaczenie dla wiązania RNA mają reszty aminokwasowe znajdujące się w centralnej części strony wklęsłej domeny FinO, podczas gdy reszty aminokwasowe znajdujące się na krawędziach tej domeny lub po stronie wypukłej mają mniejsze znaczenie dla wiązania RNA lub, jak

w przypadku K35 i R69, mają zróżnicowany wpływ na wiązanie różnych cząsteczek RNA. Przedstawione wyniki pozwalają na wyjaśnienie sposobu oddziaływania białka ProQ z cząsteczkami RNA i przez to na lepsze zrozumienie udziału tego białka w procesach kontroli ekspresji genów zależnych od regulatorowych RNA u bakterii.

Abstract

Bacterial small noncoding RNAs (sRNAs) contribute to adaptation to changing environmental conditions. By interaction with specialized proteins, noncoding RNAs impact translation and target mRNA stability. One of the best known RNA binding proteins is Hfq, however recent reports point out the role of FinO-domain proteins in this kind of regulation. ProQ protein present in *Escherichia coli*, is a member of this family. ProQ consists of two globular domains, RNA-binding N-terminal FinO domain (NTD) and C-terminal Tudor domain (CTD) connected via unstructured linker of about 50 aa. ProQ binds numerous RNAs in *E. coli*, however little is known yet about its ligand recognition fashion.

In order to explain how the ProQ protein recognizes RNA molecules, in the first part of my PhD thesis I investigated which features of the sequence and structure of RNA molecules are necessary for their binding by the FinO domain of the ProQ protein, and in the second part of the thesis I examined how mutations in the domain FinO affect the binding of various RNA molecules. In the first part of my research, I established the characteristics of RNA molecules that determine their binding to NTD. From RNAs bound by ProQ *in vivo*, I selected six molecules: MicA, SibA, *cspE*-3', *malM*-3', *hupA*-5' and *lpp*-5', and I measured their binding to purified full-length ProQ protein and its FinO domain (NTD). The analysis of RNA molecules and their mutants binding indicated that the structural element recognized by NTD is the Rho-independent transcription terminator. For the binding relevant are a polyuridine tail of at least four nucleotides at the 3' end and a double-stranded hairpin region of at least two base pairs. The introduction of a mismatch above the fourth base pair of the double-stranded region of the terminator hairpin does not affect the interaction with the NTD. Furthermore, I found that the adenosine-rich sequence motif located on the 5'-side of the terminator hairpin of RNA molecules bound by ProQ plays a role in competition between ProQ and another protein binding the 3' end of RNA, which is the Hfq protein. The substitution of adenosine for uridine caused the RNA molecules to bind to NTD less or the same as to RNA of natural sequence. However, the binding of such mutant RNA molecules to the Hfq protein was clearly improved. The results suggest that the adenosine-rich motif prevents the binding of RNA molecules to the Hfq protein and facilitates their binding to ProQ.

In the second part of my research I indicated amino acids within the N-terminal domain of ProQ important for interaction with RNA molecules. Based on the analysis of the evolutionary conservation of amino acid sequence and the available literature I selected ten amino acid residues and checked their effect on the binding of seven RNAs (*cspE-3'*, *cspE81-3'*, *malM-3'*, *gapA-3'*, RybB, SibA and SibB) *in vitro*. Substitutions at positions R58, Y70 and R80 resulted in inhibition of protein binding to RNA. Significant for the interaction are also positions K54 and R62, which substitutions significantly weakened RNA binding, while mutations at positions R32, D41 and T65 had the least effect on binding. Interestingly, mutations of the K35 and R69 residues had a diverse effect on binding of different RNAs, the least on the binding of the *cspE-3'*, and the greatest on the binding to SibB sRNA. This may suggest differences in the interaction between RNAs and the FinO domain of the ProQ protein. To indicate which regions of RNA molecules are responsible for the differential effect of the K35A and R69A substitutions in NTD on RNA binding, I examined the binding of RNA molecules containing mutations that: deprived them of single-stranded regions adjacent to the terminator hairpin, changed the sequence of the terminator hairpin, or changed the length of the linker between the terminator hairpin and the nearest double-stranded element on the 5' side of the hairpin. The effects of introducing these mutations in RNA did not indicate direct contact of the mutated RNA sequences with amino acid residues at positions 35 and 69, therefore it is possible that the K35A and R69A substitutions have an indirect effect on RNA binding, e.g. by changing the conformation of the protein.

In conclusion, the results of the experiments showed that the N-terminal domain of the ProQ protein recognizes RNA molecules by binding to them within the junction between the lower part of the transcription terminator hairpin and the surrounding single-stranded sequences. In addition, studies of the effects of mutations introduced into the FinO domain suggest that amino acid residues located in the central part of the concave side of the FinO domain are crucial for RNA binding, while amino acid residues located at the edges of this domain or on the convex side are less important for RNA binding or, as in the case of K35 and R69 residues, have a diverse effect on the binding of different RNAs. The presented results allow to explain the mode of interaction of the ProQ protein with RNA molecules, and thus to better understand the role of this protein in the process of gene expression dependent on regulatory RNAs in bacteria.

Wykaz zastosowanych skrótów i symboli

% (v/v)	stężenie procentowe objętościowe
% (w/v)	stężenie procentowe objętościowo-masowe
°C	stopień Celsjusza
A	adenozyna
A	alanina
aa	aminokwasy (ang. <i>amino acids</i>)
AMP	adenozyno-5'-monofosforan
APS	nadsiarczan amonu
asRNA	antysensowny RNA (ang. <i>antisense RNA</i>)
ATP	adenozyno-5'-trifosforan
C	cytozyna
CTD	domena C-końcowa białka (ang. <i>C-terminal domain</i>)
B3H	bakteryjny system trójhybrydowy (ang. <i>bacterial three-hybrid assay</i>)
BSA	albumina surowicy bydłowej (ang. <i>bovine serum albumin</i>)
CLIP-seq	metoda immunoprecypitacji poprzedzona sieciowaniem krzyżowym i sekwencjonowaniem DNA (ang. <i>cross-linking immunoprecipitation, and DNA sequencing</i>)
CTP	cytydino-5'-trifosforan
cV	objętość kolumny (ang. <i>column volume</i>)
D	asparaginian
Da	dalton
DNA	kwask deoksyrybonukleinowy
DTT	ditiotreitól
EDTA	kwask etylenodiaminotetraoctowy
EMSA	metoda różnicowej migracji w żelu poliakrylamidowym (ang. <i>electrophoretic mobility shift assay</i>)
F	fenyloalanina
G	guanozyna
GradSeq	metoda frakcjonowania kompleksów RNA-białko w gradiencie glicerolu
GTP	guanozyno-5'-trifosforan
h	godzina
HDX	metoda spektrometrii masowej z analizą obszarów chronionych przed wymianą atomów wodoru na deuter (ang. <i>hydrogen–deuterium exchange mass spectrometry</i>)
IGR	region międzygenowy (ang. <i>intergenic region</i>)
K	lizyna
K_d	równowagowa stała dysocjacji (ang. <i>equilibrium dissociation constant</i>)
L	leucyna
M	molowe
min	minuta
mRNA	matrycowy RNA (ang. <i>messenger RNA</i>)

N	dowolny nukleotyd
ncRNA	niekodujący RNA (ang. <i>non-coding RNA</i>)
NMR	spektroskopia magnetycznego rezonansu jądrowego (ang. <i>nuclear magnetic resonance</i>)
nt	nukleotyd
NTD	N-końcowa domena (ang. <i>N-terminal domain</i>)
ORF	otwarta ramka odczytu (ang. <i>open reading frame</i>)
PNK	kinaza polinukleotydowa
PNPaza	fosforylaza polinukleotydowa
pz	pary zasad
R	arginina
RBP	białko wiążące RNA (ang. <i>RNA binding protein</i>)
RBS	miejsce wiązania rybosomu (ang. <i>ribosome binding site</i>)
RIL-seq	metoda badania oddziaływania RNA przez ligację i sekwencjonowanie (ang. <i>RNA interaction by ligation and sequencing</i>)
RNA	kwasy rybonukleinowe
rcf	relatywna siła odśrodkowa (ang. <i>relative centrifugal force</i>)
rpm	obroty na minutę (ang. <i>revolutions per minute</i>)
rRNA	rybosomalne RNA (ang. <i>ribosomal RNA</i>)
RT	temperatura pokojowa (ang. <i>room temperature</i>)
s	sekunda
S	seryna
SD	sekwencja Shine-Dalgarno
sRNA	małe regulatorowe RNA (ang. <i>small RNA</i>)
T	treonina
TBE	bufor TBE (ang. <i>Tris-Borate-EDTA</i>)
TEMED	N, N, N', N'-tetrametyloetylenodiamina
Tris	tris(hydroksymetylo)aminometan
tRNA	transferowy RNA (ang. <i>transfer RNA</i>)
U	urydyna
U	jednostka enzymatyczna (ang. <i>unit</i>)
UTP	urydyno-5'-trifosforan
UTR	region RNA niepodlegający translacji (ang. <i>untranslated region</i>)
UV	światło ultrafioletowe
V	wolt
W	wat
wt	typ dziki (ang. <i>wild type</i>)
Y	tyrozyna

1. Wstęp

1.1. Wprowadzenie

Informacja genetyczna o budowie i funkcji każdej z komórek, bez względu na to, czy stanowiącej organizm jednokomórkowy jak komórka bakteryjna, czy też budującej twór nieporównanie bardziej złożony, zapisana jest na helisie DNA. Warunkiem przetrwania organizmu jest umiejętność przystosowania do nieustannie zmieniających się warunków środowiska, do czego niezbędne jest wykształcenie wyspecjalizowanych mechanizmów adaptacyjnych. Obok DNA podstawową rolę w procesie powstawania białek pełni RNA, a hipoteza „świata RNA”, w którym zarówno nośnikiem genów, jak i katalizatorem reakcji chemicznych był RNA, opisuje poprzedzający zastany i obserwowany przez nas etap rozwoju życia. Cząsteczki RNA pełnią rozmaite funkcje biologiczne: budulcowe, matrycowe, transferowe, sygnalizacyjne, enzymatyczne, jak również adaptacyjne w postaci udziału w regulacji ekspresji genów.

Praca ta stanowi opracowanie biochemicznych analiz oddziaływania regulatorowych niekodujących cząsteczek RNA z jednym z białek wiążących RNA u bakterii *Escherichia coli* oraz próbę zrozumienia mechanizmów rozpoznawania regulowanych cząsteczek przez białko w celu lepszego poznania złożonych i fascynujących procesów adaptacyjnych komórek bakteryjnych.

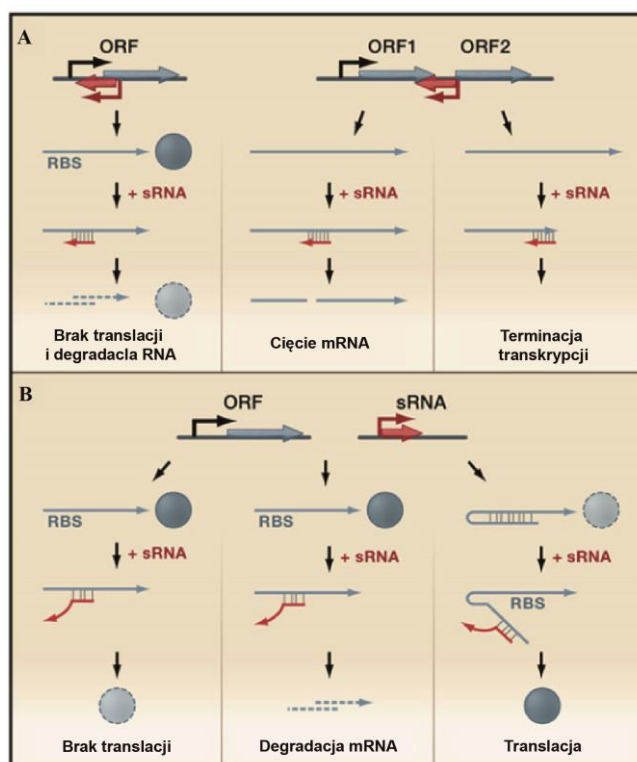
1.2. Regulatorowe RNA u bakterii

Ze względu na lokalizację informacji genetycznej stanowiącej potranskrypcyjny niekodujący RNA (ncRNA, ang. *non-coding RNA*), wyróżnia się elementy regulatorowe kodowane *in cis*, czyli zapisane na nici DNA komplementarnej do nici, na której zapisany jest regulowany gen, i *in trans*, co oznacza, że element regulatorowy powstaje z *locus* odległego od miejsca zakodowania sekwencji regulowanej. Do elementów regulatorowych *in cis* można zaliczyć ryboprzełączniki i termometry RNA, a także część sRNA (ang. *small RNAs*), natomiast większość sRNA reguluje ekspresję genów *in trans*. Regulatorowe RNA o funkcji podobnej do sRNA mogą także powstawać na drodze obróbki lub degradacji długich transkryptów, a wiele z takich cząsteczek oddziałuje z białkami na podobnej zasadzie co sRNA. Unikalną grupę regulatorowych RNA stanowią RNA CRISPR, które wspólnie z białkami z rodziny Cas tworzą swoisty system obronny przed infekcjami wirusowymi, a ryboprzełączniki i termometry RNA, które są zlokalizowane w regionach nieulegających translacji na 5'-końcu transkryptów (5'-UTR, ang. *5' untranslated region*),

modyfikują swoją konformację w odpowiedzi na zmianę stężenia metabolitu lub jonu, przy czym nie potrzebują do działania białek (Richards i Belasco 2021)

1.2.1. sRNA i ncRNA powstające na drodze obróbki transkryptów

Cząsteczki sRNA to krótkie (50-400 nt; najdłuższy znany sRNA to RNAIII ze *Staphylococcus aureus* o długości 514 nt (Novick i Geisinger 2008; Waters i Storz 2009; Storz i wsp. 2011)), niekodujące RNA zaangażowane w regulację ekspresji genów bakteryjnych, często w odpowiedzi na warunki stresowe, zmiany adaptacyjne, kontrolujące czynniki wirulencji czy *quorum sensing* (Vogel i Pappenfort 2006; Novick i Geisinger 2008; Pappenfort i Vogel 2010; Wagner i Romby 2015). W zależności od lokalizacji genu kodującego sRNA i regulowanego przez nie mRNA wyróżnia się sRNA działające *in cis* i *in trans*. sRNA działające *in cis* nazywane są także antysensownymi RNA (asRNA, ang. *antisense RNA*), ponieważ kodowane są w *locus* regulowanych mRNA, ale na antysensownej nici; z tego powodu są w całości do nich komplementarne (Rys. 1A). Z kolei sRNA działający *in trans* kodowany jest w *locus* odległym od sekwencji regulowanej, z którą dzielą tylko częściową komplementarność w regionie oddziaływania *seed* (Rys. 1B, (Storz i wsp. 2011)). Zarówno sRNA kodowane *in cis* i *in trans* poprzez związanie z regulowanym mRNA wpływają na jego degradację, umożliwienie bądź zahamowanie translacji, a czasem także terminację transkrypcji.



Rysunek 1. Lokalizacja w genomie i funkcje regulatorowe sRNA.
Opis w tekście (zmienione, na podstawie (Waters i Storz 2009)).

Początkowo sRNA odkrywano w regionach międzygenowych (IGRs, ang. *intergenic regions*), jednak wśród krótkich regulatorowych RNA zidentyfikowano cząsteczki o innym pochodzeniu (Wagner i Romby 2015). Krótkie niekodujące RNA (ncRNA, ang. *non-coding RNA*) odnaleziono w regionach niekodujących zlokalizowanych na 3'- i 5'-końcach mRNA – mogą zarówno być tam kodowane, jak i powstawać na drodze obróbki potranskrypcyjnej (Kawano i wsp. 2005; Chao i wsp. 2012; Wagner i Romby 2015; Ren i wsp. 2017), czy być wycinane z dłuższych transkryptów przez RNazę E (Dar i Sorek 2018), mogą równocześnie funkcjonować jako sRNA i kodować funkcjonalny peptyd, jak sRNA SgsR (Wadler i Vanderpool 2007) czy jak sRNA SroC powstawać podczas rozpadu operonowego mRNA (Miyakoshi i wsp. 2015). sRNA spełniają wiele funkcji, a regulacja zależna od sRNA często odbywa się przy udziale białek wiążących RNA.

1.3. Białka uczestniczące w kontroli zależnej od sRNA

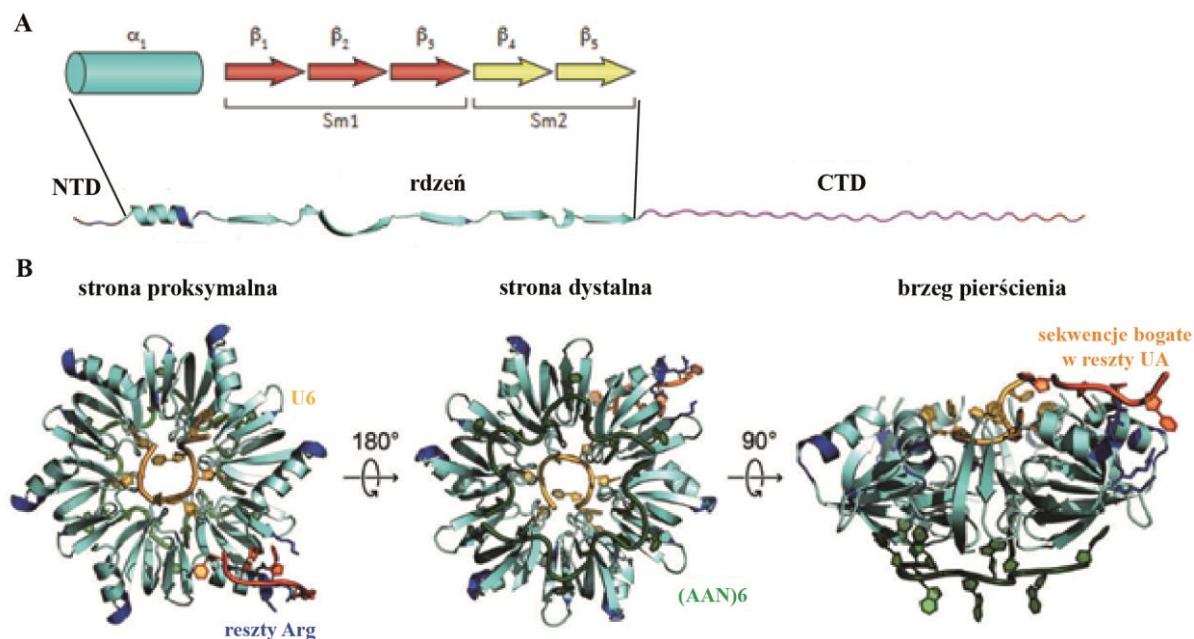
Komponent białkowy w regulacji zależnej od sRNA ułatwia odnalezienie się regulatora i regulowanego mRNA. U większości bakterii rolę tę pełni białko Hfq, które stanowi platformę wiążącą obie wchodzące w skład oddziaływania cząsteczki. Białko to wiąże ponad 70 sRNA i setki mRNA (Holmqvist i wsp. 2016; Olejniczak i Storz 2017), dlatego nazywane jest globalnym regulatorem. W 2016 roku opublikowano badania (Smirnov i wsp. 2016), w których donoszono o zidentyfikowaniu u *Salmonella enterica* drugiego globalnego regulatora sRNA – białka ProQ, należącego do rodziny białek z domeną FinO.

1.1.1. Hfq

Białko Hfq zostało opisane jako czynnik niezbędny dla replikacji RNA bakteriofaga Q β u *E. coli* (Franze de Fernandez i wsp. 1968). Jest homoheksamerem o wielkości ok. 70 kDa i kształcie pierścienia; każda z sześciu podjednostek posiada zakonserwowaną domenę Sm – N-końcowy rdzeń niosący 2 motywy sekwencyjne Sm1 i Sm2 oraz zmienny C-koniec (Rys. 2A.; (Holmqvist i Vogel 2018)). Choć domenę nazwano Sm, struktura ta bardziej niż eukariotyczną domenę Sm przypomina domenę Lsm (ang. *Sm-like*, pol. podobna do Sm; (Wilusz i Wilusz 2012)). Białko Hfq wiąże pojedynczoniowe RNA przy pomocy czterech powierzchni (Rys. 2B):

- strony proksymalnej wykazującej powinowactwo do sekwencji bogatych w urydyny, często obecne na 3'-końcu sRNA (Soper i wsp. 2010; Dimastrogiovanni i wsp. 2014);

- strony dystalnej wiążącej sekwencji zawierające powtórzenie $(AAN)_n$, gdzie A to adenina a N to dowolny nukleotyd (Mikulecky i wsp. 2004; Link i wsp. 2009; Robinson i wsp. 2014);
- zewnętrznego brzegu pierścienia, do którego wiążą się sekwencje wzbogacone w adenozyliny i urydyny (Ishikawa i wsp. 2012);
- C-końca, który odpowiada za wyparcie związanych uprzednio transkryptów (Santiago-Frangos i wsp. 2016).



Rysunek 2. Struktura i powierzchnie wiążące RNA białka Hfq.

(A) Schemat budowy podjednostkowej rdzenia białka wraz ze schematycznym przedstawieniem monomeru. (B) Strony wiążące RNA wraz z motywami przez nie rozpoznawanymi (zmienione; na podstawie (Vogel i Luisi 2011) i (Santiago-Frangos i Woodson 2018)).

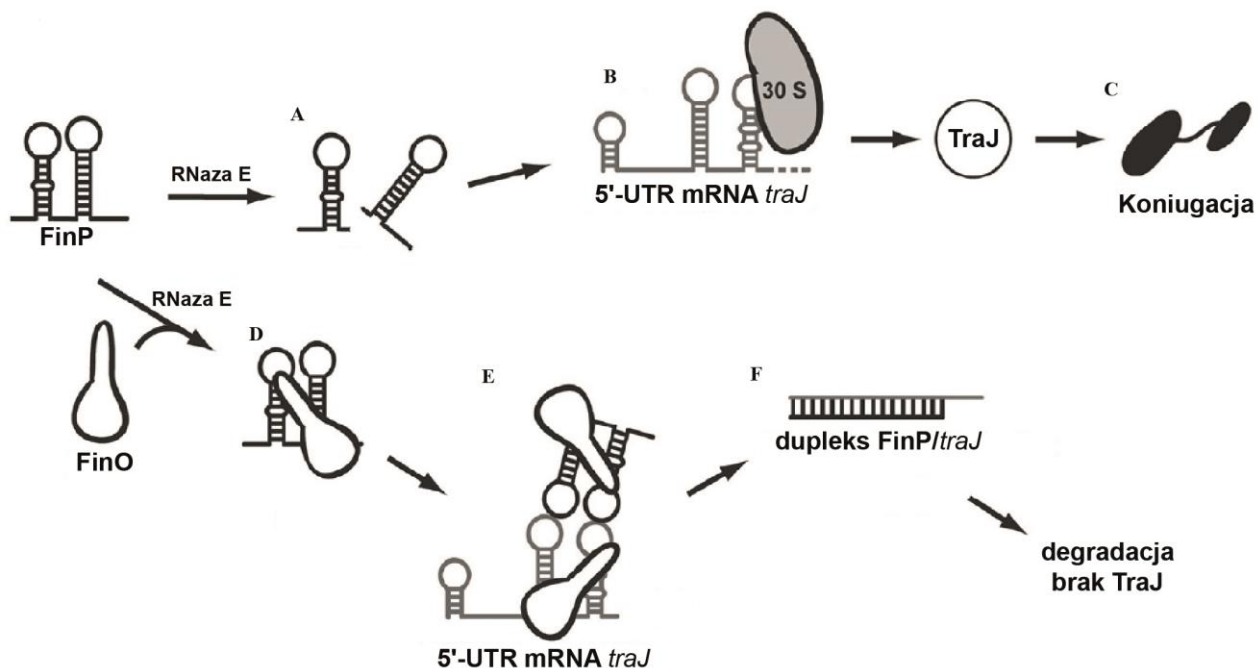
W celu odróżnienia sposobu wiązania do tego białka, cząsteczki sRNA podzielono na 2 grupy: cząsteczki klasy I wiążące Hfq poprzez strony proksymalną i brzegową, podczas gdy regulowane mRNA wiązane jest od strony dystalnej oraz sRNA klasy II, które oddziałują z białkiem stronami proksymalną i dystalną, natomiast mRNA wiązane jest przez brzeg pierścienia Hfq (Małecka i wsp. 2015; Schu i wsp. 2015; Kwiatkowska i wsp. 2018). System regulatorowy Hfq-sRNA-mRNA wpływa na zahamowanie lub aktywowanie translacji, a białko Hfq stabilizuje sRNA chroniąc je przed degradacją przez RNazę E (Vogel i Luisi 2011). Warto odnotować, że oprócz sRNA i mRNA białko to oddziałuje z rRNA (Andrade i wsp. 2018), tRNA (Lee i Feig 2008), DNA (Cech i wsp. 2016), a na liście jego partnerów białkowych są RNazaE (Morita i wsp. 2005),

polimeraza I poli(A) (Mohanty i wsp. 2004), PNPaza (Andrade i wsp. 2012), polimeraza RNA (Sukhodolets i Garges 2003) czy białko Rho (Bidnenko i Bidnenko 2018).

1.1.2. Białka z rodziny FinO

1.1.2.1. FinO

21,2 kDa białko FinO u *E. coli* kodowane jest na plazmidzie F, na którym znajduje się także operon *tra* kodujący większość białek zaangażowanych w koniugację (Frost i wsp. 1994). FinO wiąże sRNA FinP kodowane antysensownie do fragmentu 5'-UTR mRNA *traJ* i w całości do niego komplementarne. Kiedy FinO jest nieobecne, sRNA FinP zostaje zdegradowane przez RNazę E (Rys. 3A), a mRNA *traJ* ulega translacji (Rys. 3B), w wyniku której powstaje białko umożliwiające koniugację (Rys. 3C). Kiedy jednak FinO jest obecne, wiąże ono sRNA FinP i stabilizuje jego strukturę chroniąc przed degradacją zależną od RNAzy E (Rys. 3D). Następnie FinP w kompleksie z FinO wiąże komplementarne struktury w *traJ* (Rys. 3E), czym blokuje miejsce rozpoznania rybosomu i translację TraJ. Powstaje dupleks FinP/*traJ* (Rys. 3F), który podlega degradacji, zatem komórka traci zdolność koniugacji.



Rysunek 3. Schematyczne przedstawienia działania systemu FinOP regulującego translację mRNA *traJ*.

Opis w tekście (zmienione, na podstawie (Glover i wsp. 2015).

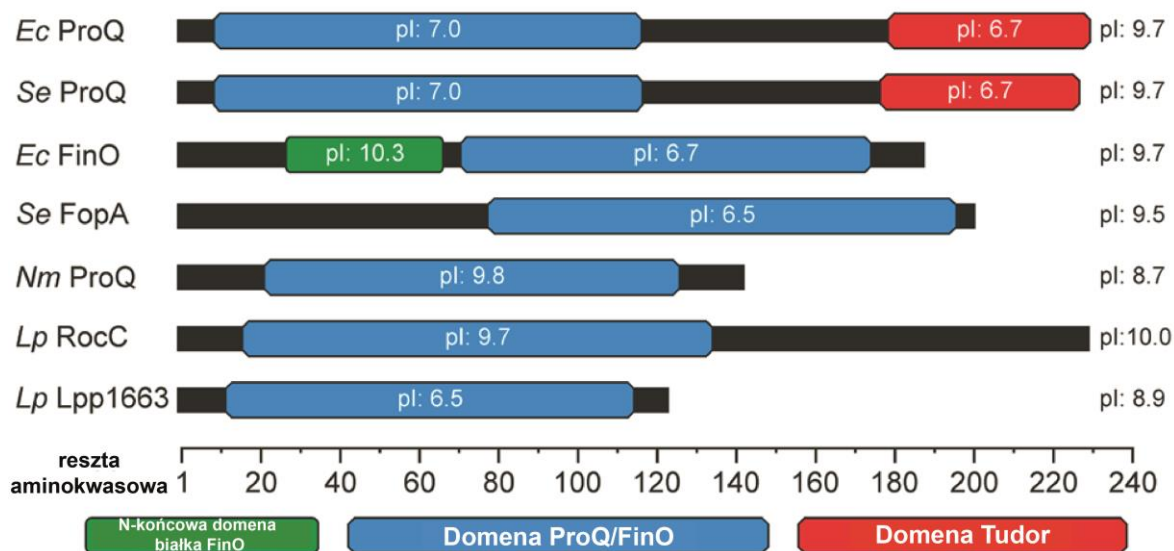
Pośród ligandów FinO u *S. enterica* (kodowanego na plazmidzie pSLT) zidentyfikowano również sRNA RepX kodowane na plazmidzie pRSF1010 (El Mouali i wsp. 2021a). Zauważono, że sposób działania sRNA RepX na mRNA *repA* może być

blizniaczo podobny do mechanizmu regulacji koniugacji opartego o dupleks *FinP-traJ*. *RepX*, podobnie jak *FinP*, kodowane jest antysensownie do mRNA *repA* i w całości komplementarne, co wskazuje na podobieństwo mechanizmu kontroli replikacji plazmidu pRSF1010 zależnej od *RepX* do zależnej od *FinP* kontroli koniugacji plazmidu pSLT (El Mouali i wsp. 2021a).

Struktura białka *FinO* została rozwiązana przy wykorzystaniu krystalografii rentgenowskiej (Ghetu i wsp. 2000). Domenę *FinO* zdefiniowano jako centralną domeną białka w granicach 25-186 aminokwasu. Składa się ona z 5 helis α i 2 struktur β (Olejniczak i Storz 2017), a jej kształt przyrównano do pięści prawej dłoni z wyprostowanym palcem wskazującym i kciukiem, które stanowią odpowiednio N- i C-kończącą domenę (Ghetu i wsp. 2000). Białko *FinO* dało nazwę rodzinie białek posiadających wyżej opisaną domenę określoną jako PF04352.

1.1.2.2. Inne białka z rodziny *FinO*

W skład białek z rodziny *FinO* wchodzi: *RocC* i *Lpp1663* z *Legionella pneumophila* (Attaiech i wsp. 2016), minimalne *ProQ* (opisane także jako NMB 1681) z *Neisseria meningitidis* (Chaulk i wsp. 2010; Bauriedl i wsp. 2020), czy *ProQ* i *FopA* z *S. enterica* (Smirnov i wsp. 2016; Smirnov i wsp. 2017; Gerovac i wsp. 2020). Wszystkie zawierają domenę *FinO* i różnego rodzaju rozszerzenia na N- i C-końcach (Rys. 4).



Rysunek 4. Schematyczna budowa domenowa wybranych białek z rodziny *FinO*.

Dla każdej z domen i pełnej długości białek podano teoretyczny pI; *Ec* – *Escherichia coli*; *Se* – *Salmonella enterica*; *Nm* – *Neisseria meningitidis*; *Lp* – *Legionella pneumophila* (zmienione, na podstawie (Immer i wsp. 2020).

Białko FinO i białko FopA kodowane są na plazmidach (u *E. coli* FinO kodowane jest na plazmidzie F, u *S. enterica* FinO na plazmidzie pSLT, a FopA na plazmidzie pCol1B9), natomiast ProQ u *E. coli* i *S. enterica* oraz białka znane u *N. meningitidis* i *L. pneumophila* kodowane są na chromosomie bakteryjnym. Białko RocC z *L. pneumophila* przy pomocy swojej domeny FinO rozpoznaje sRNA RocR, jego plazmidowy homolog RocRp oraz kilka wiązanych przez nie mRNA, przez co uczestniczy w regulacji procesów kompetencji (Attaiech i wsp. 2016; Durieux i wsp. 2019). Nie są znane inne cząsteczki wiązane przez drugie białko występujące u tego gatunku, Lpp1663, natomiast wskazano na preferencję wiązania przez nie sekwencji bogatych w urydyny (Immer i wsp. 2020). Dla minimalnego ProQ z *N. meningitidis* w eksperymencie wykorzystującym metodę immunoprecypitacji poprzedzoną sieciowaniem krzyżowym w świetle ultrafioletowym i sekwencjonowaniem DNA (UV-CLIP-seq; ang. *UV cross-linking immunoprecipitation, and DNA sequencing*) zidentyfikowano wiązanie z 16 sRNA i 166 mRNA (Bauriedl i wsp. 2020). Białko FopA zostało opisane jako wiążące liczne cząsteczki RNA kodowane na plazmidzie pCol1B9, jednak jego głównym ligandem jest sRNA Inc, które reguluje translację czynnika inicjacji replikacji plazmidu RepZ (Asano i wsp. 1999; Gerovac i wsp. 2020).

1.1.2.3. ProQ

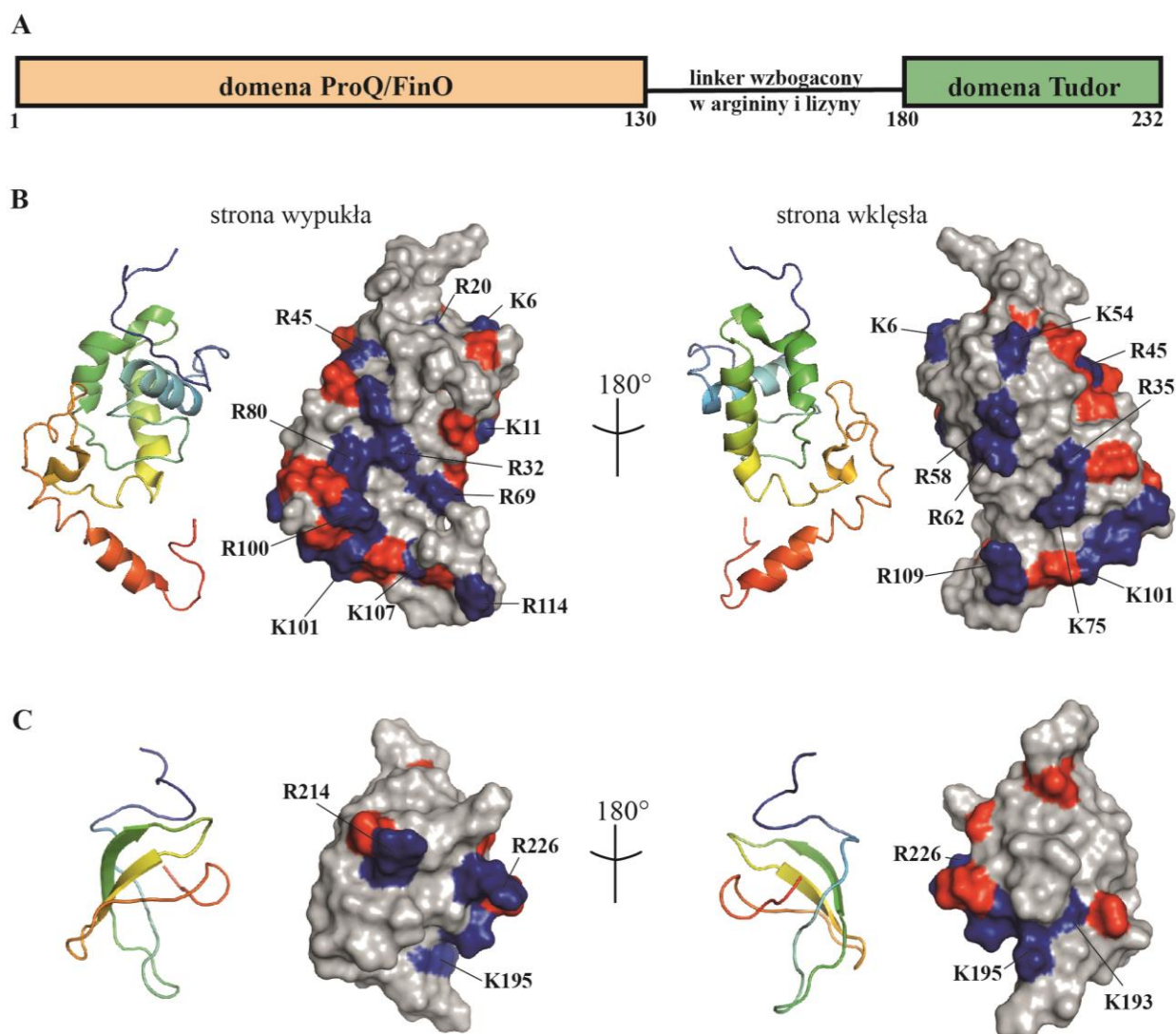
1.1.2.3.1. Odkrycie

Gen kodujący chromosomowe białko ProQ u *E. coli* odkryto w badaniu przesiewowym mutantów o zwiększonej oporności na 3,4-dehydroprolinę (Stalmach i wsp. 1983; Milner i Wood 1989). Odkryto, że fenotyp *proQ::Tn5* rzutuje na poziom symportera jonów wodorowych i osmoprotektantów ProP w zależności od fazy wzrostu i ciśnienia osmotycznego (Kunte i wsp. 1999; Chaulk i wsp. 2011), bez wpływu na transkrypcję mRNA *proP* (Milner i Wood 1989). Zauważono jednak, że u gatunków, u których występują ortologi ProQ nie zawsze znajdowano ortologi ProP, a u gatunków, u których istnieją ortologi ProP często brak ProQ (Smith i wsp. 2004; Smith i wsp. 2007), co sugerowało dodatkowe, nieznanie wówczas funkcje białka ProQ. W 2016 roku opublikowano wyniki badania frakcjonowania kompleksów RNA-białko w gradiencie glicerolu (GradSeq; (Smirnov i wsp. 2016)) u *S. enterica*, w którym obok znanych bakteryjnych białek wiążących RNA – Hfq i CsrA – zidentyfikowano białko ProQ jako oddziałujące ze znaczącą liczbą partnerów RNA. W ten sposób wskazano na rolę ProQ w regulacji ekspresji genów zależnej od sRNA.

1.1.2.3.2. Budowa

Białko ProQ zbudowane jest z 2 globularnych domen połączonych nieustrukturyzowanym linkerem (Rys. 5A), jest ono białkiem cytoplazmatycznym o masie 26 kDa (Milner i Wood 1989; Kunte i wsp. 1999; Smith i wsp. 2007). Dotychczas nie udało się uzyskać struktury krystalicznej białka pełnej długości. Początkowo strukturę N-końcowej domeny 1-130 aa wygenerowano na zasadzie homologii do krystalicznej struktury białka FinO z wykorzystaniem serwera 3D-PSSM (Smith i wsp. 2007). W 2017 roku strukturę 1-119 aa NTD opracowano z wykorzystaniem spektroskopii magnetycznego rezonansu jądrowego (PDB: 5nb9; (Gonzalez i wsp. 2017), Rys. 5B). Podobnie jak centralna domena białka FinO, N-końcowa domena białka ProQ zbudowana jest z 5 helis α i 2 struktur β , a jej powierzchnia jest pokryta dodatkowo naładowanymi resztami aminokwasowymi. Na podstawie rozmieszczenia dodatkowo naładowanych reszt aminokwasowych wyróżniono 2 powierzchnie białka: wklęsłą (ang. *concave*) oraz zlokalizowaną po przeciwnej stronie wypukłą (ang. *convex*).

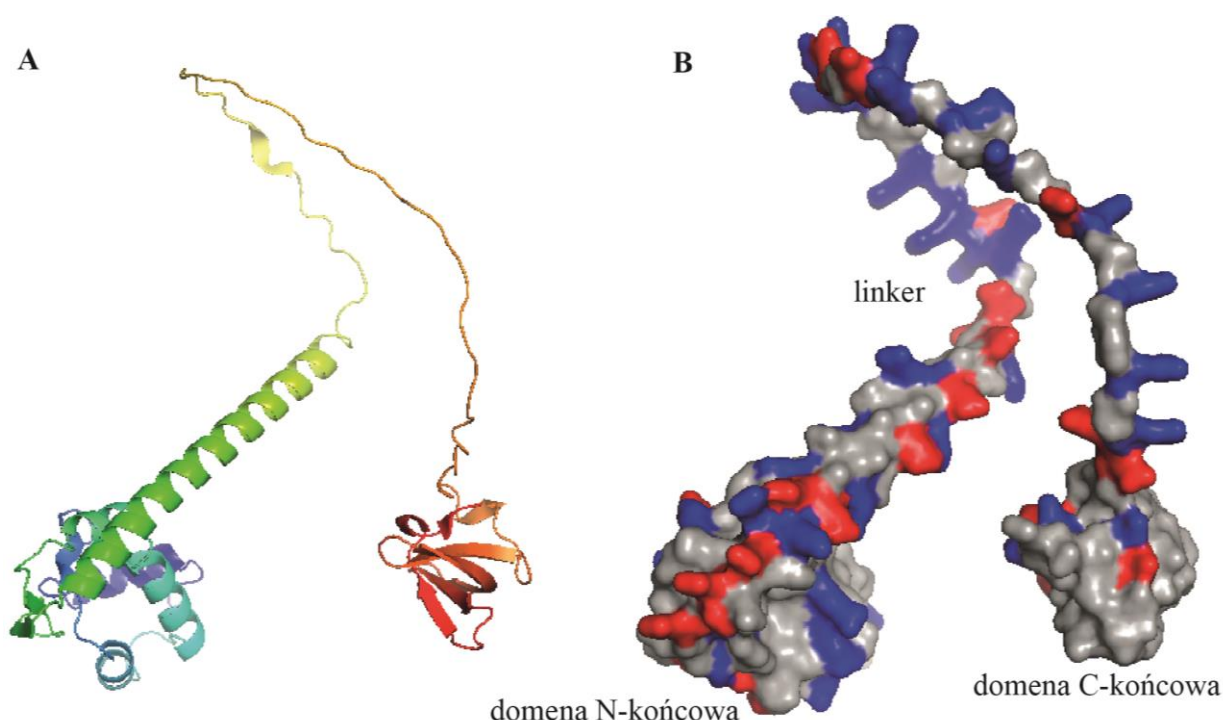
Zauważono, że C-końcowa domena białka ProQ składa się głównie ze struktur β i wykorzystano domenę SH3 do stworzenia modelu CTD (Smith i wsp. 2004). Kolejne badania strukturalne przyniosły porównanie tej struktury do domeny podobnej do Sm, pomimo nikłego podobieństwa sekwencji, a w konsekwencji CTD nazwano domeną podobną do Hfq (Smith i wsp. 2007; Chaulk i wsp. 2011). Badania z wykorzystaniem spektroskopii magnetycznego rezonansu jądrowego (PDB: 5nbb; (Gonzalez i wsp. 2017) pokazały, że trzeciorzędowa struktura CTD najbardziej przypomina domenę Tudor, która zwykle występuje u eukariotów, a mogłaby zostać nabyta na drodze horyzontalnego transferu genów (Gonzalez i wsp. 2017). Powierzchnia CTD w większości jest neutralna elektrostatycznie, choć również znajdują się na niej pojedyncze reszty aminokwasowe naładowane dodatnio, w których sąsiedztwie zlokalizowane są reszty aminokwasowe naładowane ujemnie (Rys. 5C).



Rysunek 5. Budowa białka ProQ.

(A) Schematyczne przedstawienie budowy białka ProQ u *E. coli*. (B) Model N-końcowej domeny białka ProQ uzyskany metodą spektroskopii NMR (PDB: 5nb9) ukazany z dwóch przeciwległych stron: wypukłej i wklęsłej. Modele ukazano w gradiencie kolorystycznym od niebieskiego do czerwonego, obok każdego z nich przedstawiono ich reprezentację powierzchni elektrostatycznej, reszty aminokwasowe naładowane dodatnio zaznaczono kolorem niebieskim, a te naładowane negatywnie kolorem czerwonym. Strzałkami wskazano reszty lizyny i arginynty potencjalnie uczestniczące w oddziaływaniu z RNA. (C) Model C-końcowej domeny białka ProQ uzyskany metodą spektroskopii NMR (PDB: 5nbb), przedstawiony w gradiencie kolorystycznym od niebieskiego do czerwonego. Obok przedstawiono reprezentację powierzchni elektrostatycznej, oznaczenia jak uprzednio.

Dzięki programowi AlphaFold, który w oparciu o głębokie uczenie i sztuczną inteligencję tworzy predykcje struktur trzeciorzędowych białek (Jumper i wsp. 2021), możemy poznać także predykcję pełnej długości białka ProQ z *E. coli* (PDB: AF_AFP45577F1). Linker, którym połączone są domeny N- i C-końcowa, zawiera dużo dodatkowo naładowanych reszt aminokwasowych (Rys. 6), które mogą wskazywać na rolę tej struktury w wiązaniu kwasów nukleinowych.



Rysunek 6. Model struktury pełnej długości białka ProQ z *E. coli* wygenerowany przez program AlphaFold (PDB: AF_AFP45577F1).

(A) Model białka ProQ przedstawiony w gradiencie kolorystycznym od niebieskiego do czerwonego, z uwzględnieniem elementów struktury drugorzędowej. (B) Reprezentacja powierzchni elektrostatycznej, reszty aminokwasowe naładowane dodatnio zaznaczono kolorem niebieskim, a te naładowane negatywnie kolorem czerwonym.

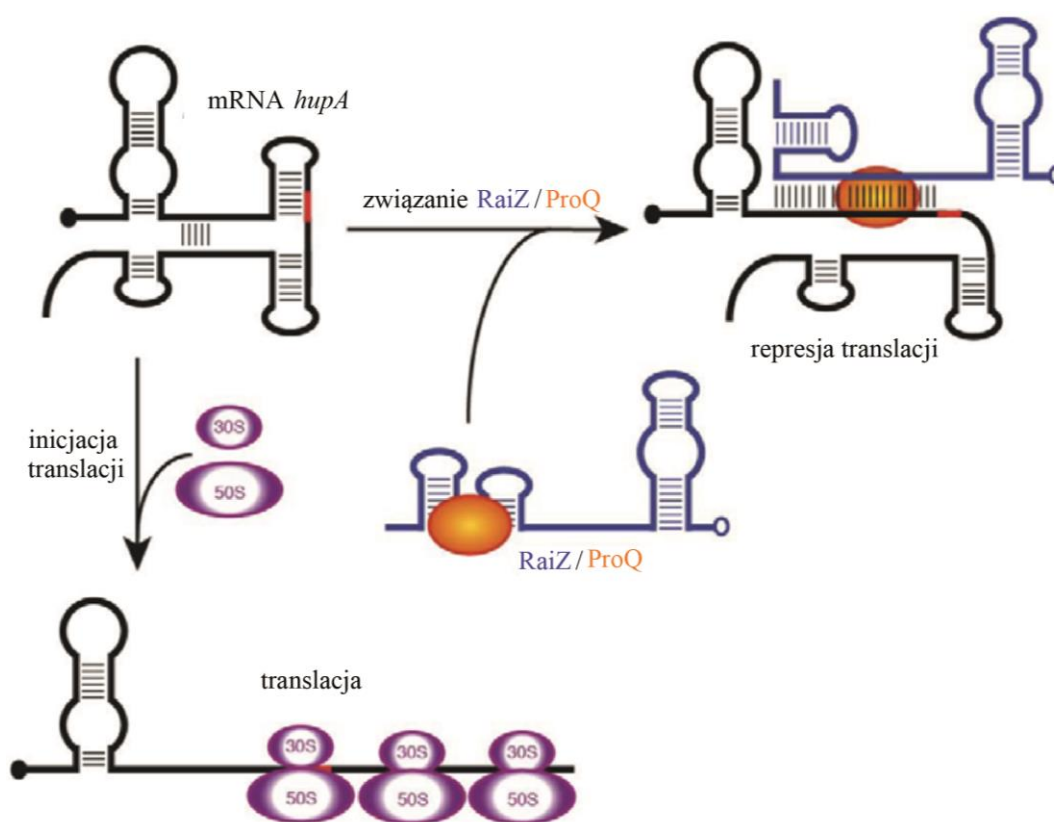
1.1.2.3.3. Funkcja

Białko ProQ uczestniczy w licznych procesach komórkowych, takich jak metabolizm DNA (Smirnov i wsp. 2016), kontrola czynników wirulencji (Westermann i wsp. 2019), kontrola czynników związanych z ruchliwością (Rizvanovic i wsp. 2021), formowanie biofilmu (Sheidy i Zielke 2013), adaptacja w warunkach stresu osmotycznego (Kerr i wsp. 2014) czy w warunkach braku dostępności zasobów (Avrani i wsp. 2017).

Zauważono, że białko ProQ wiąże cząsteczki RNA o stabilnej strukturze drugorzędowej (Smirnov i wsp. 2016). Badania CLIP-seq transkryptomów *E. coli* i *S. enterica* ujawniły, że wiązane fragmenty są dwuniciowe i wzbogacone w pary GC, w tym zidentyfikowano Rho-niezależne terminatory (Holmqvist i wsp. 2018). Analiza RIL-seq również pokazała, że białko ProQ wiąże sekwencje zawierające pary GC, po których występują urydyny (Melamed i wsp. 2020). Cząsteczki RNA wiązane przez białka z rodziny FinO – minimalne ProQ z *N. meningitidis*, RocC z *L. pneumophila* czy plazmidowe białko FinO – także zawierają motywy dwuniciowe, w tym spinki

terminatorowe (Jerome i Frost 1999; Arthur i wsp. 2011; Attaiech i wsp. 2016; Bauriedl i wsp. 2020).

Zbadany został sposób oddziaływania białka ProQ z *S. enterica* z sRNA RaiZ; białko stabilizuje związane sRNA (Smirnov i wsp. 2017). Przy braku sRNA RaiZ struktura drugorzędowa *hupA*, w którą uwikłane jest miejsce wiązania rybosomu, rozplata się, co umożliwia inicjację translacji mRNA (Rys. 7). Związanie dupletu RaiZ/ProQ nieco powyżej RBS zapobiega związaniu podjednostki rybosomu 30S i powstaniu podjednostki HU- α , która wraz z podjednostką HU- β tworzą podobne do histonu białko HU (Laine i wsp. 1980).



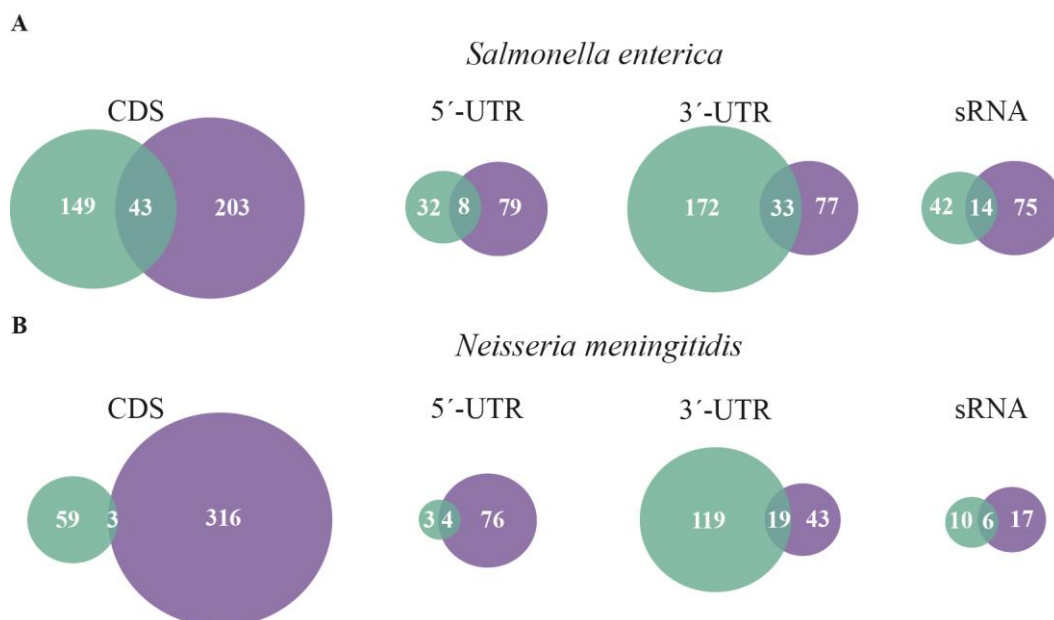
Rysunek 7. Proponowany model zależnej od białka ProQ regulacji translacji mRNA *hupA* przez sRNA RaiZ.

Opis w tekście (zmienione, na podstawie (Smirnov i wsp. 2017).

N-końcowa domena ProQ wiąże RNA tak samo dobrze, jak pełnej długości białko, co sugeruje, że NTD jest główną strukturą białka wiążącą RNA (Chaulk i wsp. 2011). Analiza obszarów chronionych przed wymianą atomów wodoru na deuter (HDX) podczas wiązania 85-nt fragmentu *cspE* 3'-UTR i 169-nt sRNA SraB do białka ProQ ujawniła najsilniejszą ochronę w obrębie NTD, choć zasugerowano także kontakty badanych RNA z linkerem i CTD (Gonzalez i wsp. 2017).

W badaniu mutantów białka ProQ opartym o bakteryjny system trójhybrydowy (B3H) aminokwasy w pozycjach K35, D41, K54, R58, Y70, L71, D82 i R80 wskazano jako ważne dla wiązania cząsteczek *cspE* i SibB (Pandey i wsp. 2020). W analizie mutantów ProQ z *S. enterica* w obrębie N-końcowej domeny zidentyfikowano ponad 20 aminokwasów, których mutacje wpłynęły na utratę funkcji białka *in vivo* (Rizvanovic i wsp. 2021). U *S. enterica* białko ProQ stabilizuje mRNA *cspE* (Holmqvist i wsp. 2018), jednak poziom tego mRNA w obecności mutantów L34Q, L57R, R80H, D82E, L83R był obniżony (Rizvanovic i wsp. 2021). W tym samym badaniu zidentyfikowano także trzy mutacje C-końcowej domeny (G216D, G185V, T200P), w których obecności poziom mRNA *cspE* był podwyższony, jednak hamowały one aktywność regulacyjną białka; mutacje nie wpływały na stabilność białka (Rizvanovic i wsp. 2021).

Zarówno białko ProQ, jak i Hfq wiążą krótkie niekodujące RNA, jednak pule wiązanych przez nie RNA są różne. Z porównania lokalizacji ligandów ProQ i Hfq u dwóch gatunków bakterii: *S. enterica* (Holmqvist i wsp. 2018) i *N. meningitidis* (Bauriedl i wsp. 2020) wynika, że większość sekwencji wiązanych przez ProQ pochodzi z regionów 3'-UTR, a w przypadku Hfq większość wiązanych sekwencji ulokowana jest w obszarze sekwencji kodujących oraz 5'-UTR (Rys. 8). Warto zauważyć, że część wiązanych RNA jest wspólnych dla obu białek, a podobne obserwacje poczyniono dla białek z *E. coli* (Melamed i wsp. 2020).



Rysunek 8. Porównanie liczby ligandów białek ProQ i Hfq u (A) *S. enterica* i (B) *N. meningitidis* w zależności od pochodzenia regulatorowego RNA..

Ligandy ProQ zidentyfikowane w eksperymencie CLIP-seq zaznaczono kolorem **turkusowym**, natomiast ligandy Hfq zaznaczono kolorem **fioletowym** (na podstawie (Bauriedl i wsp. 2020)).

2. Cele pracy

Główne cele mojej pracy scharakteryzowałam następująco:

1. Określenie cech cząsteczek RNA decydujących o ich wiązaniu przez domenę FinO białka ProQ;
2. Poznanie znaczenia reszt aminokwasowych zlokalizowanych w obrębie domeny FinO białka ProQ dla oddziaływania z różnymi cząsteczkami RNA.

Pierwsza część badań dotyczyła analizy *in vitro* oddziaływań pełnej długości białka ProQ oraz jego wyizolowanej N-końcowej domeny FinO z cząsteczkami RNA o różnych cechach strukturalnych. Aby określić cechy budowy cząsteczki RNA wpływające na wiązanie N-końcowej domeny białka ProQ, wyróżniłam następujące cele szczegółowe:

1. Wyznaczenie siły oddziaływania białka ProQ oraz jego wyizolowanej N-końcowej domeny z cząsteczkami RNA o różnej budowie;
2. Poznanie roli struktury terminatora transkrypcji niezależnego od białka Rho w oddziaływaniu z wyizolowaną N-końcową domeną białka ProQ;
3. Wyjaśnienie, czy lokalizacja struktury spinki z sekwencją poliurydynową na 3'-końcu ma znaczenie dla oddziaływania RNA z domeną FinO białka ProQ;
4. Poznanie znaczenia długości 3'-terminalnego ogona poliurydynowego dla wiązania przez N-końcową domenę białka ProQ;
5. Określenie długości regionu dwuniciowego w obrębie spinki terminatora transkrypcji niezbędnego dla wiązania N-końcowej domeny białka ProQ;
6. Wyjaśnienie, czy motyw bogaty w adenozyne zlokalizowany po 5'-stronie spinki terminatora transkrypcji wpływa na odmienne wiązanie RNA przez białka ProQ i Hfq.

W drugiej części badań zaprojektowałam zestaw mutantów domeny FinO białka ProQ zawierających substytucje pojedynczych aminokwasów i zbadalam wpływ mutacji na wiązanie różnych cząsteczek RNA *in vitro*. W tej części pracy sformułowałam następujące cele szczegółowe:

1. Wyznaczenie siły oddziaływania mutantów N-końcowej domeny białka ProQ z różnymi cząsteczkami RNA;
2. Wyjaśnienie, jakie cechy budowy cząsteczki RNA wpływają na różne sposoby wiązania cząsteczek RNA do mutantów domeny FinO białka ProQ:
 - a. Sprawdzenie, czy sekwencje zlokalizowane po 5'- i 3'-stronie terminatora transkrypcji w cząsteczkach RNA mogą wyjaśnić różnice w wiązaniu cząsteczek RNA przez mutanty N-końcowej domeny białka ProQ;
 - b. Poznanie udziału dolnej części regionu sparowanego spinki terminatorowej w wiązaniu mutantów N-końcowej domeny białka ProQ;
 - c. Określenie znaczenia długości jednoniciowego fragmentu łączącego struktury drugorzędowe w obrębie cząsteczki RNA dla oddziaływania z mutantami N-końcowej domeny białka ProQ.

3. Materiały

3.1. Odczynniki i kolumny do chromatografii białek

Odczynnik	Producent
2-betamerkaptoetanol	Sigma Aldrich
2-butanol	Stanlab
Adenozyno-5'-trifosforan (ATP)	Sigma Aldrich / ROTH
Agaroza	Prona
Akrylamid:bisakrylamid 19:1 (mieszanina)	BioShop
Akrylamid:bisakrylamid 29:1 (mieszanina)	BioShop
Ampicylina, sól sodowa	Sigma Aldrich
Bromofenol	Sigma Aldrich
Chlorek magnezu	Sigma Aldrich
Chlorek sodu	Sigma Aldrich
Comassie Brilliant Blue R250	Sigma Aldrich
Cytydino-5'-trifosforan (CTP)	Sigma Aldrich / ROTH
Ditiotreitol (DTT)	Sigma Aldrich
dNTP, mieszanina	Thermo Scientific
EDTA	Sigma Aldrich
Etanol (96%)	POCH/Stanlab
Fenol/Chloroform/Alkohol izoamyłowy (25:24:1)	Bioshop
Filtr wirówkowy Amicon Ultra-10K	Merck
Glicerol	BioShop
Guanozyno-5'-trifosforan (GTP)	Sigma Aldrich / ROTH
Imidazol	Sigma Aldrich
IPTG	BioShop
Kanamycyna, siarczan	Sigma Aldrich
Ksylencjanol	Sigma Aldrich
Kwas borowy	BioShop
Kwas octowy (99,9%)	Sigma Aldrich
Kwas solny (35-37%)	POCH/Stanlab
Laurylosiarczan sodu (SDS)	Sigma Aldrich
LB Broth Miller	BioShop
LB Broth Agar	BioShop
Marker wielkości białek PageRuler Plus	Thermo Scientific
Marker wielkości DNA Gene Ruler 100 bp	Thermo Scientific
Marker wielkości DNA Gene Ruler 1 kb	Thermo Scientific
Metanol (99,9%)	Stanlab
Midori Green	NIPPON Genetics
Mocznik	BioShop
Nadsiarczan amonu (APS)	BioShop
<i>N,N,N',N'</i> -Tetrametyloetylenodiamina (TEMED)	Bioshop
Octan sodu	Sigma Aldrich
Siarczan niklu(II), sześciowodny	Sigma Aldrich
Spermidyna	Sigma Aldrich
Surowicza albumina wołowa (BSA)	BioShop
TRIS	Bioshop
Urydino-5'-trifosforan (UTP)	Sigma Aldrich /ROTH

γ -32P-ATP (222TBq (6000Ci)/mmol 370MBq (10mCi)/ml)	Hartmann Analytics
Kolumna do chromatografii białek	
HiLoad 16/600 Superdex 16/200 pg	GE Healthcare
His Trap ff crude 5 ml	GE Healthcare/Cytvia
HiTrap Heparin 1 ml	GE Healthcare/Cytvia
Superloop 10 ml	GE Healthcare

3.2. Enzymy:

Enzym	Producent/Białko oczyszczone przez
DNaza I	Thermo Scientific
DpnI	Thermo Scientific
Fast Digest BamHI	Thermo Scientific
Kinaza polinukleotydomowa PNK z faga T4	Thermo Scientific
Ligaza DNA z faga T4	Thermo Scientific
Polimeraza DNA Dream Taq	Thermo Scientific
Polimeraza DNA Phusion	Thermo Scientific
Polimeraza RNA z faga T7	Przeze mnie/dr Joanna Kwiatkowska, Pracownia Biochemii RNA
Polimeraza TEV	dr Michał Taube, Zakład Ekspresji Genów

3.3. Oligodeoksyrybonukleotydy

3.3.1. Oligonukleotydy DNA użyte do reakcji wydłużania starterów

Nazwa	Sekwencja (5' → 3')
cspA-3'_F	TAATACGACTCACTATAGTAATCTCTGCTTAAAAGCACAGAATCTAAGATCCCTGCCATTTGGCGGGGATTTTTTT
cspA-3'_R	AAAAAAATCCCCGCCAAATGGCAGGGATCTTAGATTCTGTGCTTTTAAAGCAGAGATTAC
cspA-3'-AtoR_F	TAATACGACTCACTATAGTAATCTCTGCTTAAAAGCACAGAATCTTTTCCCTGCCATTTGGCGGGGATTTTTTT
cspA-3'- AtoR_R	AAAAAAATCCCCGCCAAATGGCAGGGAAAAAAGATTCTGTGCTTTTAAGCAGAGATTAC
cspE-3'_F	TAATACGACTCACTATAGTAAGATACGTCAGCAAGAATTCAAACCCGCTTAATC
cspE-3'_R	AAAAAAAACCCGCTGATTAAGCGGGTTTTGAATTCTTGCTGACGTATCTTAC
cspE-3'-ext+5_R	AGACCAAAAAAACCCTGATTAAGCGGGTTTTGAATTC
cspE-3'-ext+22_R	CCTCAACCGCACACTAAAGACCAAAAAAACCCTGATTAAGCGGGTTTTGAATTC
cspE-3'-6U_R	AAAAAACCCGCTGATTAAGCGGGTTTTGAATTCTTGCTGACGTATCTTAC
cspE-3'-4U_R	AAAACCCGCTGATTAAGCGGGTTTTGAATTCTTGCTGACGTATCTTAC
cspE-3'-0U_R	CCCGCTGATTAAGCGGGTTTTGAATTCTTGCTGACGTATCTTAC
cspE-3'-AtoU_F	TAATACGACTCACTATAGTAAGATACGTCAGCAAGAATTCTTTCCCGCTTAATC
cspE-3'-AtoU_R	AAAAAAAACCCGCTGATTAAGCGGGAAAAAGAATTCTTGCTGACGT

	ATCTTAC
cspE-3'-mini_F	TAATACGACTCACTATAGGAATTCGGAAGGCGCTTAATC
cspE-3'-mini_R	AAAAAAAAAGGCGCTGATTAAGCGCCTTCCGAATTCCTAT
cspE-3'-mini-6U_R	AAAAAAGGCGCTGATTAAGCGCCTTCCGAATTCCTAT
cspE-3'-mini-4U_R	AAAAGGCGCTGATTAAGCGCCTTCCGAATTCCTAT
cspE-3'-mini-5'-blunt_F	TAATACGACTCACTATAGGAAGGCGCTTAATCAGCGCCTTTTTTTT
cspE-3'-mini-5'-blunt_R	AAAAAAAAAGGCGCTGATTAAGCGCCTTCCTATAGTGAGTCG
cspE-3'-mini-tr-5'-blunt_F	TAATACGACTCACTATAGGCGCTTAATCAGCGCCTTTTTTTT
cspE-3'-mini-tr-5'-blunt_R	AAAAAAAAAGGCGCTGATTAAGCGCCTATAGTGAGTCG
cspE-3'-4XX_F	TAATACGACTCACTATAGTAAGATACGTCAGCAAGAATTC
cspE-3'-4AA_R	AAAATTTTCCCGCTGATTAAGCGGGTTTTGAATTCTTGCTGACGTA TCTTAC
cspE-3'-CG_R	AAAACCCCCCGCTGATTAAGCGGGGGGGGAATTCTTGCTGACGT ATCTTAC
cspE-3'-4GC_R	AAAAAAAGCCCGCTGATTAAGCGGGCTTTGAATTCTTGCTGACGT ATCTTAC
cspE-3'-4UA_R	AAAATTTTCCCGCTGATTAAGCGGGAAAAGAATTCTTGCTGACGTA TCTTAC
cspE-3'-4UU_R	AAAAAAAACCCGCTGATTAAGCGGGAAAAGAATTCTTGCTGACGT ATCTTAC
cspE81-3'_F	TAATACGACTCACTATAGGCCCTTCTGCTGCAAACGTAATCGCTCT GTAAGATACGTCAGCAAG
cspE81-3'_R	AAAAAAAACCCGCTGATTAAGCGGGTTTTGAATTCTTGCTGACGTA TCTTACAGAG
cspE81-3'-AtoU_R	AAAAAAAACCCGCTGATTAAGCGGGAAAAGAATTCTTGCTGACGT ATCTTACAGAG
cspE81_1nt_F	TAATACGACTCACTATAGGGAATTCTGCTGCAAACGTAATCGCTCT GTAAGATACGTCAG
cspE81_7nt_F	TAATACGACTCACTATAGGCCCTTCTGCTGCAAACGTAATCGCTCT GTAAGATACGTCAG
cspE81_7nt_R	AAAAAAAACCCGCTGATTAAGCGGGTTTTGAAAGAGTGCTGACGT ATCTTACAGAGCG
gapA-3'_F	TAATACGACTCACTATAGGACCTGATCGCTCACATCTCCAAATAAG TTGAGATGACACTGT
gapA-3'_R	AAAAAAAAGAGCGACCGAAGTCGCTCTTTTTAGATCACAGTGTC TCTCAACTTA
gapA-3'-AtoU-R	AAAAAAAAGAGCGACCGAAGTCGCTCAAAAAGATCACAGTGTC ATCTCAACTTA
gapA80_3nt_F	TAATACGACTCACTATAGGTCCTTCTCGCAGACATCTCCAAATAAG TTGAGATGACACTGTCT
gapA80_7nt_F	TAATACGACTCACTATAGGTCCTTCTCGCTCACATCTCCAAATAAG TTGAGATGACACTG
gapA80_7nt_R	AAAAAAAAGAGCGACCGAAGTCGCTCTTTTTAGAAAGACAGTGTC

	TCTCAACTTA
hupA-5'_F	TAATACGACTCACTATAGCGAAAAAAGTGGCTATCGGTGCGTGT ATGCAGGAGAGTGCTATTCTGGCATTTCGGTCGCACTCGATGCTTA GCAAGCG
hupA-5'_R	TTCATAAGTTATCCTTACAATGTGTTTATCGCTTGCTAAGCATCGA GTGCG
hupA-5'_mut_R	AAAACGTAAGTTATCCTTACGATGTGTTTATCGCTTGCTAAGCATC GAGTGCG
hupA-5'_term_R	AAAAAAGGTGCGCCAGGAGACGCACCATTTCATAAGTTATCCTTA CAATGTGTTTATCGCTTGCTAAGCATCGAGTGCG
hupA-5'_hp_R	GGTGCGCCAGGAGACGCACCATTTCATAAGTTATCCTTACAATGTGT TTATCGCTTGCTAAGCATCGAGTGCG
hupB-3'_F	TAATACGACTCACTATAGGTAAACTAAGCGTTGTCCCCAGTGGGG ATGTG
hupB-3'_R	AAAAAAGGCACATCAGTAGATGCGCCCTTGAACCTTCGTCACATCC CCACTGGGGAC
hupB-3'-RtoU_R	AAAAAAGGCACATCAGTAGATGCGCCAAAGAACTTCGTCACATCC CCACTGGGGAC
ihfA-3'_F	TAATACGACTCACTATAGCTTCGCCAAAGACGAGTAATCTGAT
ihfA-3'_R	AGAAAAAAGGCCGCAGAGCGGCCTTTTTAGTTAGATCAGATTACT CGTC
ihfA-3'-AtoU_R	AGAAAAAAGGCCGCAGAGCGGCCAAAAAAGTTAGATCAGATTAC TCGTC
infA-3'_F	TAATACGACTCACTATAGTCGCTGATTGTTTTACCGCCTGATGGGC GAAGAGAAA
infA-3'_R	ATAAAAAGGCCGGTTAAACCGACCTTTTACTCGTTCTTTCTCTTCG CCCATC
infA-3'-AtoU_R	ATAAAAAGGCCGGTTAAACCGACCAAAAAGTTAGATCAGATTAC CCCATC
lpp-5'_F	TAATACGACTCACTATAGCTACATGGAGATTAAGTCAATCTAGAG GGTATTAATAATGAAAGCTACTAAACTG
lpp-5'_R	AGTAGAACCCAGGATTACCGCGCCCAGTACCAGTTTAGTAGCTTTC ATTATTAATACCCTC
lpp-5'-mut_R	AAAAGGGCGCGGGATTACCGCGCCCAGTACCAGTTTAGTAGCTTT CATTATTAATACCCTC
lpp-5'-term_R	AAAAAAGGTGCGCCAGGAGACGCACCAGTAGAACCCAGGATTA CCGCGCCCAGTACCAGTTTAGTAGCTTTCATTATTAAT
lpp-5'-hp_R	GGTGCGCCAGGAGACGCACCAGTAGAACCCAGGATTACCGCGCCC AGTACCAGTTTAGTAGCTTTCATTATTAAT
malM-3'_F	TAATACGACTCACTATAGCTTTATCAGCAGTGTAAGGCAAGGG GTAATTACGCCCCACAGTG
malM-3'_R	AAAAAAGGTGCGCCAGGAGACGCACCAGTTGTTGCAAAATCAGC ACTGTGGGGCGTAATTA
malM-3'- ext+5_R	AGCATAAAAAAAGGTGCGCCAGGAGACGCACCAGTTGTTGCAAAA TCAGCACTGTGGGGCGTAATTA
malM-3'- ext+17_R	TCCAGGAAGGAAGCATAAAAAAAGGTGCGCCAGGAGACGCACC AGTTGTTGCAAAATCAGCACTGTGGGGCGTAATTA
malM-3- Abulge_R	AAAAAAGGTGCGCCAGGAGACGTCACCAGTTGTTGCAAAATCAG CACTGTGGGGCGTAATTA

malM-3'-AAAbulge_R	AAAAAAAGGTGCGCCAGGAGACGTTTCACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-AA_R	AAAAAAAGGTTTGGCAGGAGACTTACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-UU_R	AAAAAAAGGTAAGCCAGGAGACAAACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-4bp_R	AAAAAAAGGTGAGGACACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-3bp_R	AAAAAAAGGTAGGAACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-2bp_R	AAAAAAAGGAGGACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-1bp_R	AAAAAAAGAGGACAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-9U_R	AAAAAAAAGGTGCGCCAGGAGACGCACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-6U_R	AAAAAAGGTGCGCCAGGAGACGCACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-5U_R	AAAAAGGTGCGCCAGGAGACGCACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-4U_R	AAAAGGTGCGCCAGGAGACGCACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-3U_R	AAAGGTGCGCCAGGAGACGCACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-2U_R	AAGGTGCGCCAGGAGACGCACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-1U_R	AGGTGCGCCAGGAGACGCACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-0U_R	GGTGCGCCAGGAGACGCACCAGTTGTTGCAAATCAGCACTGTGGGGCGTAATTA
malM-3'-AtoU_R	AAAAAAAGGTGCGCCAGGAGACGCACCAGAAGAAGCAAATCAGCACTGTGGGGCGTAATTA
MicA_F	TAATACGACTCACTATAGAAAGACGCGCATTGTTATCATCATCCCTGAATTC
MicA_R	AAAAGGCCACTCGTGAGTGGCCAAAATTCATCTCTGAATTCAGGATGATGAT
model_1nt_F	TAATACGACTCACTATAGGTTTTTTCGCTAGAAATACGCTAAAACCGCTT
model_1nt_R	AAAAAAAACCCGCTGATTAAGCGGGTTTTAGCGTATTTCTACGCAAAAACC
model_2nt_F	TAATACGACTCACTATAGGTTTTTTCGCTAGAAATACGCTTAAAACCGCTT
model_2nt_R	AAAAAAAACCCGCTGATTAAGCGGGTTTTAAGCGTATTTCTACGCAAAAACC
model_3nt_F	TAATACGACTCACTATAGGTTTTTTCGCTAGAAATACGCTTTAAAACCGCTT
model_3nt_R	AAAAAAAACCCGCTGATTAAGCGGGTTTTAAGCGTATTTCTACGCAAAAACC
model_4nt_F	TAATACGACTCACTATAGGTTTTTTCGCTAGAAATACGCTTTTAAAACCGCTT

	CGCTT
model_4nt_R	AAAAAAAAACCCGCTGATTAAGCGGGTTTTAAAAGCGTATTTCTAC GCAAACC
model_5nt_F	TAATACGACTCACTATA GGTTGCGTAGAAATACGCTTTTTAAAACC CGCTT
model_5nt_R	AAAAAAAAACCCGCTGATTAAGCGGGTTTTAAAAGCGTATTTCTA CGCAAACC
model_6nt_F	TAATACGACTCACTATA GGTGGCGTAGAAATACGCTTTTTTTAAAACC CGCTT
model_6nt_R	AAAAAAAAACCCGCTGATTAAGCGGGTTTTAAAAAGCGTATTTCT ACGCACC
RybB_F	TAATACGACTCACTATA GCCACTGCTTTTCTTTGATGTCCCCATTT GTGGAGCCC
RybB_R	AAAAAACCCATCAACCTTGAACCGAAATGGCGGGGTTGATGGGCT CCACAAAATGGGG
SibA_F	TAATACGACTCACTATA GAGGGTTAGGGAGAGGTTTCCCCCTCCCC CTGGTGTTCTTAGTAAGCCTGGAAGCTAATCACTAAGAGTA
SibA_R	GGGAAAGCCTCTCCCGGAGAAGAGGGCTTTTAATAAGGAAAGGGT TATGATGAAGCACGTCATCATACTGGTGATACTCTTAGTGATTAGC T
SibB_F	TAATACGACTCACTATA GAGGGTAGAGCGGGGTTTCCCCGCCCCT GGTAGTCTTAGTAAGCGGGGAAGCTTATGACTAAGAGCACC
SibB_R	GGAAAGCCCCTCCCGAGGAAGGGGCCATAAATAAGGAAAGGGTC ATGATGAAGCTACTCATCATCGTGGTGCTCTTAGTCATAAG

Kolorem **niebieskim** zaznaczono promotor dla polimerazy RNA T7 umożliwiający transkrypcję *in vitro*. Oligodeoksyrybonukleotydy otrzymano z firmy Sigma Aldrich lub Metabion.

3.3.2. Oligodeoksyrybonukleotydy użyte do przygotowania konstruktyw do nadekspresji białek

Nazwa	Sekwencja (5' → 3')	Cel
ProQ/NTD_1F	<u>GAAAATCTTTATTTCCAA</u> <u>TCCATGGAAAATC</u>	Insercja fragmentu kodującego sekwencję rozpoznawaną przez proteazę TEV po 5'-stronie sekwencji kodującej
ProQ_1R	GGATCC TCAGAACACCA GGTGTCTGCGC	Insercja miejsca restrykcyjnego BamHI po 3'-stronie sekwencji kodującej
NTD_1R	GTTAGCAGCCGGATCC TACTCACCAG	
ProQ/NTD_2F	GCAGCCATATGCTCGAG GATCC gGAAAATCTTT	Insercja miejsca restrykcyjnego BamHI po 5'-stronie sekwencji kodującej
ProQ_2R	GCAATACATCACGGATC CTCAGAACACCAG	Dodanie krótkiej sekwencji dodatkowej w celu ułatwienia wydajnego trawienia przez enzym BamHI
NTD_2R	GTTTAAACGGGCTTTGTTA GCAGCCGG	

GAAAATCTTTATTTCCAATCC – sekwencja kodująca peptyd rozpoznawany przez polimerazę TEV;

GGATCC – miejsce restrykcyjne BamHI;

g – dodatkowa guanina wstawiona w celu utrzymania otwartej ramki odczytu;

XXXXXX – sekwencja dodatkowa ułatwiająca efektywne trawienie enzymem BamHI;

XXXXXX – sekwencja kodująca białko ProQ/NTD.

Wszystkie oligonukleotydy otrzymano z firmy Sigma Aldrich.

3.3.3. Oligodeoksyrybonukleotydy użyte do mutagenazy

Nazwa	Sekwencja (5' → 3')
R32A_F	GGAAGGTGAAGCGGCGCCGCTGAAAATCGGTATTTTT
R32A_R	AAAAATACCGATTTTCAGCGGCGCCGCTTCACCTTCC
K35A_F	GAAGCGCGTCCGCTGGCGATCGGTATTTTTTCAGGATTT
K35A_R	AAATCCTGAAAATACCGATCGCCAGCGGACGCGCTTC
D41A_F	GAAAATCGGTATTTTCAGGCGTTGGTTCGATCGTGTTG
D41A_R	CAACACGATCGACCAACGCCTGAAAATACCGATTTTC
K54A_F	GAGCAAACGCAATTGGCGTCCGCTTTACGTCTCTACACTT
K54A_R	AAGTGTAGAGACGTAAAGCGGACGCCAATTGCGTTTTGCTC
R58A_F	GAGCAAACGCAATTGGCGTCCGCTTTACGTCTCTACACTT
R58A_R	AAGTGTAGAGACGTAAAGCGGACGCCAATTGCGTTTTGCTC
R62A_F	GCGATCCGCTTTAGCGCTCTACACTTCGAGCTGGCGTTAT
R62A_R	ATAACGCCAGCTCGAAGTGTAGAGCGCTAAAGCGGATCGC
T65A_F	GCTTTACGTCTCTACGCGTCGAGCTGGCGTTATC
T65A_R	GATAACGCCAGCTCGACGCGTAGAGACGTAAAGC
R69A_F	CTACACTTCGAGCTGGCGGTATCTTTACGGTGTAAAC
R69A_R	GTTTAAACCCGTAAAGATACGCCCCAGCTCGAAGTGTAG
Y70F_F	CACTTCGAGCTGGCGTTTCTTTACGGTGTAAACCC
Y70F_R	GGGTTTAAACCCGTAAAGAAAACGCCAGCTCGAAGTG
R80A_F	AAACCCGGCGCAACGGCGGTCGATCTTGACGGCAAC
R80A_R	GTTGCCGTCAAGATCGACCGCCGTTGCGCCGGGTTT
R80K_F	GTTAAACCCGGCGCAACGAAAGTCGATCTTGAC
R80K_R	GTCAAGATCGACTTTCGTTGCGCCGGGTTTAAAC

Na czerwono zaznaczono miejsca mutacji, wszystkie oligonukleotydy otrzymano z firmy Sigma Aldrich lub Metabion.

3.4. Bufory i roztwory

Bufory używane do pracy z kwasami nukleinowymi	
6x stężony bufor obciążający	10 mM Tris-HCl pH 7,6 0,3 % bromofenol (w/v) 0,3 % ksylencjanol (w/v) 60% glicerol (v/v) 60 mM EDTA
10x stężony TBE	89 mM Tris-HCl pH 8 55% kwas borowy (w/v) 20 mM EDTA
10x stężony bufor do transkrypcji	800 mM Tris pH 8,0 200 mM MgCl ₂ 20 mM spermidyna 100 mM NaCl

5x stężony bufor do reakcji wiązania	25 mM Tris 150 mM NaCl 1 mM MgCl ₂ 5% glicerol (v/v)
Bufory używane do pracy z białkami	
2x stężony bufor do lizy	100 mM Tris 7,5 1 M NaCl 20 % glicerol (v/v)
Bufor A	2x stężony bufor do lizy 25 mM imidazol
Bufor B	2x stężony bufor do lizy 600 mM imidazol
2x stężony bufor bez soli	100 mM Tris 7,5 20 % glicerol (v/v)
Bufor C	2x stężony bufor bez soli 100 mM M NaCl 25 mM imidazol
Bufor D	2x stężony bufor bez soli 2 M NaCl 25 mM imidazol
Bufor do sączenia molekularnego	50 mM Tris pH 7,5 300 mM NaCl 10 % glicerol (v/v) 100 mM EDTA
10x stężony bufor SDS	25 mM Tris 250 mM glicyna 0,1% SDS (w/v)
4x stężony bufor „górny” do żelu zagęszczającego	1,5 M Tris pH 6,8 0,4 % SDS (w/v)
4x stężony bufor „dolny” do żelu rozdzielającego	1,5 M Tris pH 8,8 0,4 % SDS (w/v)
2x stężony bufor Laemli	4% SDS (w/v) 20% glicerol (v/v) 10% 2-merkaptanoetanol (v/v) 0,004% błękit bromofenolowy (w/v) 0,125 M Tris HCl, pH 6,8
Pozostałe bufory	
Bufor do opłaszczania płytek 96- dołkowych roztworem BSA	10 mM Tris pH 7,5 150 mM NaCl 0,0025% BSA (v/v)
Bufory komercyjne	
Producent	
10x bufor dla polimerazy Dream Taq	Thermo Scientific
10x bufor A dla kinazy polinukleotydowej	Thermo Scientific
5x stężony bufor dla polimerazy Phusion	Thermo Scientific
10x stężony bufor dla enzymu FD BamHI	Thermo Scientific
6x stężony bufor obciążający	Thermo Scientific
Roztwory żeli	
Roztwór żelu agarozowego	1-1,2% agaroz (w/v)

		1x TBE
Roztwór denaturującego żelu poliakrylamidowego do oczyszczania DNA i RNA		10-20% roztworu akrylamid:bisakrylamid 29:1 (v/v) 8M mocznik 1x TBE
Roztwór żelu rozdzielającego		15% roztworu akrylamid:bisakrylamid 29:1 (v/v) 1x bufor „dolny”
Roztwór żelu zagęszczającego		4,5% roztworu akrylamid:bisakrylamid 29:1 (v/v) 1x bufor „górny”
Roztwór natywnego żelu poliakrylamidowego		6% roztwór akrylamid:bisakrylamid 19:1 (v/v) 1x TBE
Pozostałe roztwory		
Roztwór barwiący do żeli białkowych		0,1 % Comassie Brilliant Blue R250 (w/v) 50% metanol (v/v) 10% kwas octowy (v/v)
Roztwór odbarwiający do żeli białkowych		40% metanol (v/v) 10% kwas octowy (v/v)

3.5. Szczepy i plazmidy bakteryjne

Nazwa	Zastosowanie	Otrzymany od
<i>E. coli</i> PM1205	Matryca do amplifikacji sekwencji genu <i>proQ</i>	Uzyskany dzięki uprzejmości prof. Giseli Storz z Narodowego Instytutu Zdrowia Dziecka i Rozwoju Człowieka im. Eunice Kennedy Shriver
<i>E. coli</i> DH5 α	Do amplifikacji wektorów bakteryjnych; oporność na ampicylinę	Invitrogen (Thermo Scientific)
<i>E. coli</i> BL21 Δ hfq	Do nadekspresji oczyszczanych białek; gen <i>hfq</i> zastąpiony kasetą oporności na kanamycynę	Uzyskany dzięki uprzejmości prof. Agnieszki Szalewskiej-Pałasz z Uniwersytetu Gdańskiego
Plazmidy		
pET-15b	Pusty wektor	Novagen (Merck)
pET-15b- <i>proQ</i>	Do nadekspresji ProQ	Ta praca
pET-15b- <i>ntd</i>	Do nadekspresji NTD ProQ	Ta praca
pET-15b- <i>ntd</i> _R32A	Do nadekspresji NTD R32A	Ta praca
pET-15b- <i>ntd</i> _K35A	Do nadekspresji NTD K35A	Ta praca
pET-15b- <i>ntd</i> _D41A	Do nadekspresji NTD D41A	Ta praca
pET-15b- <i>ntd</i> _K54A	Do nadekspresji NTD K54A	Ta praca
pET-15b- <i>ntd</i> _R58A	Do nadekspresji NTD R58A	Ta praca
pET-15b- <i>ntd</i> _R62A	Do nadekspresji NTD R62A	Ta praca
pET-15b- <i>ntd</i> _T65A	Do nadekspresji NTD T65A	Ta praca
pET-15b- <i>ntd</i> _R69A	Do nadekspresji NTD R69A	Ta praca
pET-15b- <i>ntd</i> _Y70F	Do nadekspresji NTD Y70F	Ta praca
pET-15b- <i>ntd</i> _R80A	Do nadekspresji NTD R80A	Ta praca
pET-15b- <i>ntd</i> _R80K	Do nadekspresji NTD R80K	Ta praca

4. Metody

4.1. Przygotowanie cząsteczek RNA

4.1.1. Przygotowanie matryc DNA

4.1.1.1. Oczyszczanie syntetycznych oligonukleotydów

Aby mieć pewność, że pula oligonukleotydów użyta w reakcji wydłużania starterów jest jednolita i pozbawiona chemicznych pozostałości po syntezie, cząsteczki oczyszczono w żelu poliakrylamidowym. Oligodeoksyrybonukleotydy uzyskane metodą syntezy chemicznej (Sigma Aldrich lub Metabion) były zamawiane w postaci liofilizatu, który został rozpuszczony w wodzie miliporowanej do stężenia 500 μM . Następnie do 12 μl rozpuszczonych starterów dodawano 6 μl 8 M mocznika, mieszaninę denaturowano przez 2 min w 90°C, a następnie nakładano na 10% żel poliakrylamidowy. Elektroforezę prowadzono w buforze 1x TBE przy mocy 20 W przez 1 h. W celu zwizualizowania cząsteczek DNA żel przeniesiono na płytkę kwarcową TLC i naświetlano pod lampą UV przy świetle o długości fali 254 nm, które jest pochłaniane przez kwasy nukleinowe. Oligonukleotydy wycinano z żelu, zalewano 520 μl buforu do elucji i wytrząsano przy 350 rpm w 4°C przez noc. Następnego dnia bufor elucyjny przenoszono do świeżej probówki z 3 objętościami 96% etanolu i przez 1 h inkubowano w -80°C w celu strącenia DNA. Próby wirowano przez 40 min przy 14000 rpm, usuwano supernatant, osad osuszono i rozpuszczono w 50 μl wody miliporowanej. Tak przygotowane oligonukleotydy DNA posłużyły jako matryca w reakcji wydłużania starterów.

4.1.1.2. Reakcja wydłużania starterów

Reakcja wydłużania starterów oparta jest na reakcji PCR, w której matrycę stanowią nakładające się na siebie startery, w których fragment po 3'-stronie startera F oraz fragment po 5'-stronie startera R są komplementarne. Składniki reakcji podano w tabeli:

Składnik	Stężenie końcowe
Bufor 10x stężony	1x
Mix dNTP	0,2 mM
Starter F	1,5 μM
Strater R	1,5 μM
Polimeraza Dream Taq	0,05 U/ μl
H ₂ O	Uzupełniono do 200 μl

Następnie probówkę, w której wymieszano wszystkie składniki reakcji włożono do termocyklera i włączono program PCR o następujących parametrach:

Wstępna denaturacja	94°C – 2 min	
Denaturacja	94°C – 1min 15 s	} x 7 cykli
Przylączenie starterów	55°C – 2 min	
Wydłużanie starterów	72°C – 2 min	
Końcowe wydłużanie starterów	72°C – 10 min	

Po zakończeniu reakcji PCR pobrano 5 µl mieszaniny reakcyjnej w celu rozdzielenia jej na żelu agarozowym, natomiast pozostałą część połączono w stosunku 1:1 z mieszaniną fenol:chloroform:alkohol izoamylowy. Powstałą mieszaninę krótko wytrząsano do połączenia faz, a następnie krótko wirowano. Fazę wodną (górną) przeniesiono do nowej probówki typu eppendorf, do której dodano 0,1 objętości 3M roztworu octanu sodu oraz 3 objętości 96% etanolu. Próbkę inkubowano przez 1 h w -80°C, potem wirowano przez 40 min przy 14 000 rpm. Usunięto supernatant, osad osuszono w SpeedVacu i zawieszono w 50 µl wody miliporowanej.

4.1.2. Przygotowanie cząsteczek RNA

4.1.2.1. Transkrypcja *in vitro*

W wyniku reakcji transkrypcji *in vitro* na matrycy DNA przy udziale polimerazy RNA z faga T7 powstaje cząsteczka RNA. Składniki reakcji podano w tabeli:

Składnik	Stężenie końcowe/iłość
matryca DNA	5 µl
ganozyna	2 mM
bufor do transkrypcji 10x stężony	1x
MgCl ₂	18 mM
DTT	25 mM
ATP	4 mM
CTP	4 mM
GTP	4 mM
UTP	4 mM
polimeraza RNA z faga T7	4,5 µg
H ₂ O	Do 100 µl

Matrycę DNA zdenaturowano przez 5 min w 90°C, następnie pozostawiono ją przez 5 min w RT do renaturacji. Dodano pozostałe składniki mieszaniny reakcyjnej i inkubowano przez 3 h w 37°C z wytrząsaniem 350 rpm. Po tym czasie dodano 33,3 µl 100 mM EDTA, 10 µl 3 M NaOAc o pH 5,4 oraz 3 objętości (300 µl) zimnego 96% etanolu i przez 1 h inkubowano w -80°C w celu precypitacji RNA. Następnie próbkę

wirowano przy 14000 rpm przez 40 min, usunięto supernatant, a osad zawieszono w 30 μ l 8 M mocznika.

4.1.2.2. Oczyszczanie cząsteczek RNA w żelu poliakrylamidowym

Mieszaninę poddano 2 min denaturacji w 90°C, a następnie nakładano na 10-20% żel poliakrylamidowy. Przez 1 h prowadzono elektroforezę w buforze 1x TBE przy mocy 20 W. Następnie żel umieszczono na płytce kwarcowej TLC, RNA uwidaczniiano pod lampą UV przy świetle o długości fali 254 nm. Wycięte prążki zawierające RNA o pożądanej długości zalewano 520 μ l buforu do elucji i wytrząsano przy 350 rpm w 4°C przez noc. Następnie mieszaninę elucyjną przenoszono do nowej probówki zawierającej 3 objętości etanolu 96% i strącano przez godzinę w -80°C. Próbę wirowano przez 40 min przy 14000 rpm, supernatant usuwano, osad osuszono w SpeedVacu i zawieszono w 20 μ l wody miliporowanej, a następnie zmierzono stężenie RNA przy użyciu spektrofotometru NanoDrop.

4.1.2.3. Znakowanie cząsteczki RNA izotopem radioaktywnego fosforu ^{32}P

W celu uwidocznienia cząsteczki RNA na natywnym żelu poliakrylamidowym, w wyniku reakcji znakowania do wolnej grupy hydroksylowej po 5'-stronie cząsteczki RNA przy pomocy kinazy polinukleotydomowej z faga T4 dołączono radioaktywny izotop fosforu ^{32}P . Składniki reakcji podano w tabeli:

Składnik	Stężenie końcowe/ilość
bufor A 10x stężony	1x
RNA	1000 ng
kinaza polinukleotydomowa z faga T4 (PNK)	1 U
^{32}P - γ -ATP	25 μ Ci
H ₂ O	Uzupełniono do 20 μ l

Reakcję inkubowano w 37°C. Po 45 min do reakcji dodano 3 U Dnazy I w celu strawienia pozostałości matrycy DNA po transkrypcji *in vitro* i reakcję inkubowano jeszcze przez 15 min w 37°C. Następnie do reakcji dodano 20 μ l wody miliporowanej i połączono ją w stosunku 1:1 z mieszaniną fenol:chloroform:alkohol izoamylowy. Mieszaninę wytrząsano do momentu połączenia faz i krótko zwirowano. Fazę wodną (górną) przeniesiono do probówki typu eppendorf zawierającej 20 μ l 8 M mocznika. Następnie cząsteczki oczyszczano zgodnie z procedurą opisaną w podrozdziale 4.1.1.1, a wysuszony osad zawieszono w 50 μ l wody miliporowanej.

4.2. Przygotowanie białek.

4.2.1. Przygotowanie konstruktów do nadekspresji białek

4.2.1.1. Projektowanie konstruktów

Aby uzyskać konstrukty umożliwiające nadekspresję białek, sekwencje kodujące białka ProQ i jego N-kończową domenę wklonowano do wektora pET-15b (Novagen). W celu ułatwienia procesu oczyszczania białek, do konstruktów dołączono dodatkowe sekwencje zgodnie ze schematem przedstawionym na rysunku 9. Do klonowania użyto miejsca restrykcyjnego BamHI (które dodano przy pomocy starterów na 3'-końcu sekwencji kodujących), ponieważ w plazmidzie przed tą sekwencją w otwartej ramce odczytu (ORF, ang. *open reading frame*) zlokalizowany jest His₆-tag. Dodanie sekwencji kodującej His₆-tag do przygotowanych konstruktów umożliwiło oczyszczenie białek przy zastosowaniu chromatografii metalopowinowactwa. Bezpośrednio na 5'-końcu sekwencji kodujących dodano sekwencję nukleotydową kodującą 7-aminokwasowy motyw rozpoznawany przez proteazę TEV (ENLYFQ↓S). Enzym ten tnie białko po szóstym aminokwasie rozpoznanej sekwencji, zatem w procesie oczyszczania umożliwia to odcięcie dodatnio naładowanego His₆-tagu (który może wpływać na wiązanie białka z cząsteczką RNA przy udziale niespecyficznych oddziaływań), natomiast pozostawia dodatkowy aminokwas – serynę – na N-końcu białka.



Rysunek 9. Schemat przygotowania konstruktów do nadekspresji białek pełnej długości ProQ i jego N-kończowej domeny.

SD- sekwencja Shine-Dalgarno.

4.2.1.2. Reakcja łańcuchowa polimerazy (PCR, ang. *polymerase chain reaction*)

Sekwencje kodujące namnożono przy pomocy reakcji PCR na matrycy genomowego DNA komórek *E. coli* szczepu PM1205; szczep uzyskano dzięki uprzejmości pani prof. Giseli Storz z Narodowego Instytutu Zdrowia Dziecka i Rozwoju Człowieka im. Eunice Kennedy Shriver w Bethesda w Stanach Zjednoczonych. Przy pomocy starterów do sekwencji kodujących dodano sekwencję nukleotydową sekwencji rozpoznawanej przez proteazę TEV, miejsce restrykcyjne enzymu BamHI oraz krótkie sekwencje dodatkowe dla ułatwienia cięcia nukleolitycznego i dodatkowy nukleotyd

zachowujący ramkę odczytu (Materiały, podrozdział 3.3.2.). Składniki reakcji podano w tabeli:

Składnik	Stężenie końcowe/ilość
Bufor 10x stężony	1x
Mix dNTP	0,2 nM
Starter F	0,5 μ M
Starter R	0,5 μ M
Polimeraza Dream Taq	0,25 U
Matryca	100 ng
H ₂ O	Uzupełniono do 50 μ l

Probówkę, w której zmieszano wszystkie składniki reakcji umieszczono w termocyklerze i poddano reakcji o następującym profilu termicznym:

Wstępna denaturacja	95°C – 30 s	
Denaturacja	95°C – 30 s	} x 30 cykli
Przyłączanie starterów	56°C – 1 min	
Wydłużanie starterów	72°C – 1 min	
Końcowe wydłużanie starterów	72°C – 6 min	

4.2.1.3. Trawienie restrykcyjne

Zarówno plazmid pET-15b jak i produkt PCR poddano trawieniu enzymem restrykcyjnym BamHI FastDigest. Składniki reakcji podano w tabeli:

Składnik	Stężenie końcowe/ilość
Bufor FastDigest 10x stężony	1x
Produkt PCR / plazmid pET-15b	15 μ l / 100 ng
BamHI Fast Digest	1 μ l
H ₂ O	Uzupełniono do 30 μ l

Następnie próbki z reakcjami trawienia plazmidu pET-15b i produktu PCR umieszczono w termobloku i inkubowano przez 15 min w 37°C. Po tym czasie w celu inaktywacji enzymu próbki przeniesiono do 80°C na 5 min. Produkty trawienia poddano elektroforezie w 1,2 % żelu agarozowym przez godzinę przy napięciu 90 V. Prążki uwidoczniono pod lampą UV i wycięto z żelu. Następnie DNA oczyszczono z żelu przy pomocy komercyjnie dostępnego zestawu NucleoSpin Gel and PCR Clean-up (Macherey-Nagel).

4.2.1.4. Defosforylacja plazmidu

Aby zapobiec samoligacji, plazmid poddano defosforylacji przy użyciu alkalicznej fosfatazy. Składniki reakcji podano w tabeli:

Składnik	Stężenie końcowe/ilość
plazmid pET-15b	30 μ l
bufor 10x FastAP	1x
alkaliczna fosfataza Fast AP	3U

Reakcję umieszczono w termobloku, inkubowano przez 10 min w 37°C, a następnie inaktywowano enzym przez 5 min w 75°C.

4.2.1.5. Ligacja

Oczyszczony z żelu insert poddano ligacji do trawionego i zdefosforylowanego plazmidu pET-15b. Składniki reakcji podano w tabeli:

Składnik	Stężenie końcowe/ilość
bufor 10x stężony	1x
pET-15b trawiony BamHI + AP	30 ng
insert trawiony BamHI	90 ng
ligaza	1 U
H ₂ O	Do 10 μ l

Reakcję ligacji prowadzono przez noc w 16°C.

4.2.2. Metody mikrobiologiczne

4.2.2.1. Transformacja kompetentnych komórek bakteryjnych

W celu sprawdzenia skuteczności procesu ligacji, komórki kompetentne *E. coli* szczepu DH5 α transformowano produktem ligacji. Wszystkie czynności związane z pracą z komórkami kompetentnymi wykonywano w sterylnych warunkach przy ogniu palnika. Do 50 μ l porcji komórek kompetentnych dodano 1 lub 5 μ l produktu ligacji, mieszaninę inkubowano na lodzie przez 30 min, by umożliwić oblepienie komórek cząsteczkami plazmidu. Aby ułatwić wnikanie cząsteczek plazmidu do wnętrza, komórki bakteryjne poddano szokowi cieplnemu w 42°C przez 1 min i odstawiono na 2 min na lód. Następnie do mieszaniny dodano 250 μ l ciepłej pożywki LB i komórki inkubowano w 37°C przez godzinę z wytrząsaniem przy 300 rpm. Po tym czasie komórki zwirowano przez 2 min przy 3000 rcf i odpipetowano 250 μ l supernatantu. Osad bakteryjny zawieszono przez pipetowanie w resztkę pożywki i całość przeniesiono na szalkę z podłożem selekcyjnym LB z agarem suplementowanym ampicyliną o stężeniu końcowym 100 mg/ml. Plazmid pET-15 posiada gen oporności na ten antybiotyk, zatem na podłożu powinny wyrosnąć

jedynie kolonie, które go uzyskały w drodze transformacji. Następnie szalkę z posianymi komórkami bakteryjnymi inkubowano przez noc w 37°C.

4.2.2.2. Selekcja kolonii przy użyciu kolonijnej PCR

W celu selekcji kolonii niosących pusty plazmid od tych z plazmidem zawierającym konstrukt do nadekspresji białka, kolonie wyrosłe po transformacji produktem ligacji poddano reakcji kolonijnej PCR. Matrycą w tej reakcji były kolonie poddane lizie we wstępnej wydłużonej denaturacji. Składniki reakcji podano w tabeli:

Składnik	Stężenie końcowe/ilość
Bufor 10x stężony	1x
Mix dNTP	0,2 nM
Starter F	0,5 μM
Starter R	0,5 μM
Polimeraza Dream Taq	0,5 U
H ₂ O	Uzupełniono do 200 μl

Po wymieszaniu wszystkich składników reakcji i uzupełnieniu ich o fragmenty kolonii wyrosłych po ligacji, próbówki umieszczono w termobloku i poddano reakcji PCR o następujących parametrach termicznych:

Wstępna denaturacja	94°C – 4 min	
Denaturacja	94°C – 30 s	} x 25 cykli
Przyłączanie starterów	50°C – 30 s	
Wydłużanie starterów	72°C – 2 min	
Końcowe wydłużanie starterów	72°C – 7 min	

Następnie produkty reakcji PCR rozdzielano w 1,2 % żelu agarozowym. Na podstawie różnicy wielkości pustego plazmidu (5,7 mbp) i plazmidu z insertem (6,4 mbp) wybrano kilka kolonii, z których przy użyciu komercyjnie dostępnego zestawu NucleoSpin Plasmid (Macherey-Nagel) wyizolowano plazmidy. Aby upewnić się, że konstrukty są poprawne, oddano je do zsekwencjonowania do Wydziałowej Pracowni Techniki Biologii Molekularnej.

4.2.2.3. Hodowla bakteryjna w celu nadekspresji białka

Po zsekwencjonowaniu i potwierdzeniu poprawności sekwencyjnej konstruktów, przeprowadzono testową nadekspresję w małej skali hodowlanej. Do nadekspresji białek ProQ i NTD użyto szczepu *E. coli* BL21 Δ*hfq*, w którym w miejsce genu *hfq* wprowadzono kasetę oporności na kanamycynę; szczep uzyskano dzięki uprzejmości pani

prof. Agnieszki Szalewskiej-Pałasz z Katedry Genetyki Molekularnej Bakterii Uniwersytetu Gdańskiego. Zdecydowano się na nadekspresję w szczepie pozbawionym genu kodującego białko Hfq również oddziałującego z cząsteczkami RNA, aby mieć pewność, że oczyszczone preparaty ProQ będą wolne od tego białka. W tym celu stransformowano komórki kompetentne BL21 Δhfq oczyszczonym plazmidem, zgodnie z procedurą opisaną w podrozdziale 4.2.2.1. Pojedynczą kolonię przeniesiono do 10 ml pożywki LB wzbogaconej o ampicylinę w stężeniu końcowym 100 mg/ml oraz kanamycynę w stężeniu końcowym 50 mg/ml i wstawiono hodowlę na noc do 37°C z wytrząsaniem przy 160 rpm. Następnego dnia do 100 ml świeżej pożywki suplementowanej antybiotykami zaszczepiono 1 ml nocnej hodowli. Bakterie hodowano w 37°C w wytrząsaniem 160 rpm do uzyskania OD₆₀₀~0,6, po czym w celu indukowania nadekspresji dodano IPTG do końcowego stężenia 1 μM. Po 3 h zmierzono OD hodowli, pobrano próbki i przeprowadzono ich rozdział elektroforetyczny w poliakrylamidowym żelu białkowym.

Do celów oczyszczania przygotowano 3 litry hodowli analogicznie jak na testową skalę. W tym przypadku jednak po 3 h po nadekspresji hodowlę zwirowano w ultrawirówce przez 10 min przy 5500 rpm. Usunięto pożywkę, a osad zawieszono w 1x stężonym buforze do lizy. Tak przygotowaną hodowlę zamrożono w ciekłym azocie i przechowywano w -80°C.

4.2.3. Oczyszczanie białka

4.2.3.1. Liza komórek bakteryjnych

Hodowlę rozmrażano na mieszadle magnetycznym w wodzie z lodem, po czym dodano koktajl inhibitorów proteaz (Roche), by zapobiec degradacji nadprodukowanego białka przez komórkowe proteazy w trakcie lizy komórek. Lizę komórek przeprowadzono w formie sonikacji, 6 cykli po 30 s przy amplitudzie 70%. Aby zapobiec przegrzaniu mieszaniny, sonikację przeprowadzono w chłodni, a hodowla przez cały czas pozostawała w lodzie. Następnie lizat wirowano przez 30 min przy 80000 rpm, po czym supernatant przeniesiono do czystej schłodzonej zlewki.

4.2.3.2. Chromatografia metalopowinowactwa – pierwszy etap

Podczas projektowania konstruktu eksprymowane białko zaopatrzone w His₆-tag, by w procesie oczyszczania wykorzystać powinowactwo niklu do tego dodatnio naładowanego peptydu. W tym celu użyto kolumny ze złożem sefarozowym

z immobilizowanym niklem HisTrap ff crude 5 ml (GE Healthcare/Cytvia) oraz systemu do chromatografii cieczonej ÄKTA Pure (GE Healthcare).

Uzyskany po wirowaniu supernatant nałożono przy pomocy strzykawki na kolumnę HisTrap ff crude, zbierano przesącz. Kolumnę zainstalowano w systemie ÄKTA Pure i przemywano buforem A w ilości 5 objętości kolumny (cv) przy przepływie 1 ml/min. Zbierano frakcje z płukania. Następnie kolumnę przemywano gradientem liniowym z buforu A do buforu B (od 25 mM do 600 mM imidazolu) w ilości 20 cv. Gradient pozwolił wytworzyć rosnące stężenie imidazolu w buforze, a ponieważ związek ten tworzy z niklem stabilne kompleksy, możliwe było odpłukanie z kolumny wcześniej związanego białka. Na końcu przepłukano kolumnę 5 cv buforu A. Na każdym etapie zbierano frakcje przesącza, a także dzięki obecności w systemie dwóch lamp UV i detektorów, monitorowano absorbancję przy długości fali światła 280 nM i 254 nM. Z frakcji z pików oraz na każdym etapie pracy pobierano próbki do sprawdzenia na poliakrylamidowym żelu białkowym.

4.2.3.3. Chromatografia powinowactwa

Kolejnym etapem oczyszczania białek było usunięcie z preparatu pozostałości kwasów nukleinowych. W tym celu użyto kolumny HiTrap heparin 1 ml (GE Healthcare/Cytvia), która upakowana jest złożem sefarozytowym z immobilizowaną heparyną oraz systemu do chromatografii cieczonej ÄKTA Pure (GE Healthcare). Heparyna jest polisacharydem i polianionem, dlatego wiąże dodatnio naładowane białka oddziałujące z kwasami nukleinowymi. Odpłukanie związanego białka możliwe jest poprzez zwiększenie siły jonowej roztworu, do tego celu użyto buforu o wysokim stężeniu NaCl.

Po pierwszym etapie chromatografii powinowactwa preparat białkowy znajdował się w buforze o zawartości 500 mM NaCl. Przed nałożeniem go na kolumnę heparynową preparat białkowy rozcieńczono pięciokrotnie 1x stężonym buforem bez soli, do uzyskania stężenia NaCl ok. 100 mM. Następnie roztwór zagęszczono przy wykorzystaniu filtra Amicon (Sigma Aldrich) i preparat białkowy nałożono przy pomocy strzykawki na kolumnę HiTrap heparin 1 ml (GE Healthcare/Cytvia).

Kolumnę zainstalowano w porcie systemu ÄKTA Pure i przemywano 5 cv buforu C przy przepływie 0,5 ml/min. Następnie kolumnę przemywano gradientem liniowym z buforu C do buforu D (od 100 mM do 2000 mM NaCl) w ilości 20 cv. Potem

przepłukano kolumnę 5 cv buforu C, zbierano frakcje przesączu i płukania. Na tym etapie istotne również było monitorowanie absorbancji przy długości fali światła 280 nM i 254 nM. Na podstawie wykresów absorbancji dla białka i kwasów nukleinowych do kolejnych etapów oczyszczania wybrano frakcje o niskiej zawartości kwasów nukleinowych, natomiast frakcje zanieczyszczone poddano ponownie procedurze oczyszczania na kolumnie heparynowej. Z wybranych frakcji na każdym etapie pracy pobierano próbki do sprawdzenia na poliakrylamidowym żelu białkowym.

4.2.3.4. Trawienie proteazą TEV

W celu usunięcia His₆-tagu z N-końca białka ProQ, preparat trawiono proteazą TEV, którą uzyskano dzięki uprzejmości dra Michała Taube z Zakładu Fizyki Biomedycznej Uniwersytetu im. Adama Mickiewicza. Połączono frakcje wyselekcjonowane po płukaniu na kolumnie heparynowej i dodano proteazę TEV w stosunku 1 OD enzymu na 50 OD białka. Trawienie prowadzono z dodatkiem 1 mM DTT przez noc w 4°C.

4.2.3.5. Chromatografia metalopowinowactwa – drugi etap

Jako że proteaza TEV, którą zastosowano, była zaopatrzona w His₆-tag, w celu usunięcia enzymu z preparatu białka ponownie użyto kolumny ze złożem niklowym HisTrap ff crude 5 ml. Preparat białka po trawieniu proteazą TEV nałożono strzykawką na kolumnę, po czym kolumnę przemyto 40 ml buforu A, przesącz frakcjonowano po 2 ml. Następnie na kolumnę nałożono 20 ml buforu B w celu odpłukania związanej ze złożem niklowym proteazy TEV. Następnie pobrano próbki frakcji, które sprawdzono na żelu białkowym w SDS-PAGE. Frakcje zawierające białko złączono i zagęszczono na filtrze Amicon (Sigma Aldrich).

4.2.3.6. Sączenie molekularne

W ostatnim etapie oczyszczania preparat białkowy poddano sączeniu molekularnemu na kolumnie HiLoad 16/600 Superdex 200 pg, uprzednio nakładając próbkę na kolumnę Superloop.

Obie kolumny zainstalowano w systemie ÄKTA Pure. Preparat białkowy strzykawką wprowadzono do systemu na kolumnę Superloop, następnie przy przepływie 0,2 ml/min buforu do sączenia nakładano białko z kolumny Superloop na kolumnę Superdex. Po zwiększeniu przepływu do 0,5 ml/min prowadzono elucję przez 1,2 cv. Frakcje o objętości 0,5 ml zbierano po wypłynięciu 40 ml rozpuszczalnika stanowiącego

objętość pustą kolumny (V_0), próbki pobranych frakcji sprawdzono na żelu białkowym w SDS-PAGE.

4.2.3.7. Obliczanie stężenia białka

Fracje zawierające białko po sączeniu molekularnym połączono, w kuwecie kwarcowej zmierzono jego absorbcję przy 280 nm długości fali światła. Następnie na podstawie sekwencji aminokwasowej obliczono współczynnik ekstynkcji białka (<http://protcalc.sourceforge.net/>). Równania Lamberta-Beera użyto do obliczenia stężenia białka:

$$A = k \times c \times l$$

gdzie: A – absorbcja przy długości fali 280 nm, k – współczynnik ekstynkcji, c – stężenie substancji, l – grubość warstwy absorbującej (ściany kuwety).

4.2.3.8. Przechowywanie oczyszczonych białek

Aby zapobiec problemowi utraty aktywności przez częste rozmrażanie i zamrażanie, wszystkie białka przechowywano w -80°C w alikwotach po 10 μl , które zużywano podczas pojedynczego eksperymentu.

4.3. Metody pracy eksperymentalnej:

4.3.1. Wyznaczanie stałej równowagi reakcji dysocjacji (K_d) wiązania RNA do białka

Stałą równowagi reakcji dysocjacji K_d (ang. *dissociation constant*) dla tworzenia kompleksów RNA z białkiem wyznaczono przy pomocy technik różnicowej migracji kompleksów w żelu poliakrylamidowym (EMSA, ang. *electromobility shift assay*).

Każdorazowo przed rozpoczęciem eksperymentu wiązania ^{32}P -RNA inkubowano przez 2 min w 90°C w celu zdenaturowania cząsteczki. Następnie RNA przekładano na lód na 5 min, aby cząsteczka mogła się zrenaturować. Reakcje wiązania ^{32}P -RNA z białkiem ProQ lub jego N-końcową domeną przygotowano poprzez zmieszanie 10 μl 2 nM ^{32}P -RNA (stężenie końcowe w reakcji to 1 nM) z 10 μl białka przygotowanego jako seria dwukrotnych rozcieńczeń od maksymalnego stężenia (zakres użytych rozcieńczeń ze wskazaniem maksymalnej wartości jest zaznaczony ponad każdym z opisywanych żeli) do reakcji bez białka. Reakcję przygotowano w 1x stężonym buforze do wiązania (25 mM Tris pH 7,5, 150 mM NaCl, 1 mM MgCl_2 , 5% glicerol) na płytce 96-dołkowej dodatkowo opłaszczonej buforem zawierającym 0,0025% BSA, aby zapobiec niespecyficznej

agregacji białka na ściankach dołków, która zaburzyłaby jego stężenie w reakcji. Po zmieszaniu ^{32}P -RNA z białkiem reakcje inkubowano przez 30 min w RT, a następnie pipetą wielokanałową nakładano 5 μl alikwoty na 6% natywny żel poliakrylamidowy (19:1). Rozdział elektroforetyczny prowadzono przez 45 min – 1h w buforze 0,5x TBE przy mocy 8-11 W. Żele przekładano na bibułę Whatman i suszono, po czym przekładano do kasety, nakrywano ekranem fosforowym (Fuji) i pozostawiano do ekspozycji przez noc. Po tym czasie obraz sczytywano przy użyciu skanera Amersham Typhoon (GE Healthcare), a sygnał kwantyfikowano przy użyciu oprogramowania MultiGauge. Wartości K_d wyznaczono po dopasowaniu danych do równania kwadratowego w programie GraphPad:

$$Y=Y_0+\Delta Y \frac{[E]_0 + [S]_0 + K_d - \sqrt{([E]_0 + [S]_0 + K_d)^2 - 4[E]_0[S]_0}}{2[E]_0}$$

gdzie: Y oznacza frakcję związaną, Y_0 dolne plateau, ΔY to różnica między górnym a dolnym plateau, $[E]_0$ to stężenie ^{32}P -RNA, a $[S]_0$ to stężenie białka. Wartości stałej dysocjacji K_d wyznaczono na podstawie przynajmniej trzech niezależnych powtórzeń.

4.3.2. Kompetycja o wiązanie białka

Technikę EMSA zastosowane także w eksperymencie kompetycyjnym. Bezpośrednio przed rozpoczęciem eksperymentu prowadzono 2 min denaturację ^{32}P -RNA w 90°C , po czym przenoszono cząsteczkę na 5 min na lód. Następnie 10 μl 20 nM ^{32}P -RNA mieszano z 10 μl 2 nM mieszaniny fragmentów tRNA z *E. coli* (Roche) i dodawano równomolową mieszaninę 20 nM ProQ i 20 nM Hfq. Reakcję przygotowano w 1x stężonym buforze do wiązania (25 mM Tris pH 7,5, 150 mM NaCl, 1 mM MgCl_2 , 5% glicerol), końcowe stężenia składników wynosiły: 5 nM ^{32}P -RNA, 10 nM ProQ, 10 nM Hfq i 0,5 μM tRNA. Reakcje inkubowano przez 30 min w RT, po czym 5 μl alikwoty nakładano na 6% natywny żel poliakrylamidowy (19:1). Przez 1 h prowadzono rozdział elektroforetyczny w buforze 0,5x TBE przy mocy 8 W. Żele przekładano na bibułę Whatman, suszono i przekładano do kasety z ekranem fosforowym (Fuji), gdzie przez noc prowadzono ekspozycję. Obraz sczytywano przy użyciu skanera Amersham Typhoon (GE Healthcare), a sygnał kwantyfikowano przy pomocy oprogramowania MultiGauge.

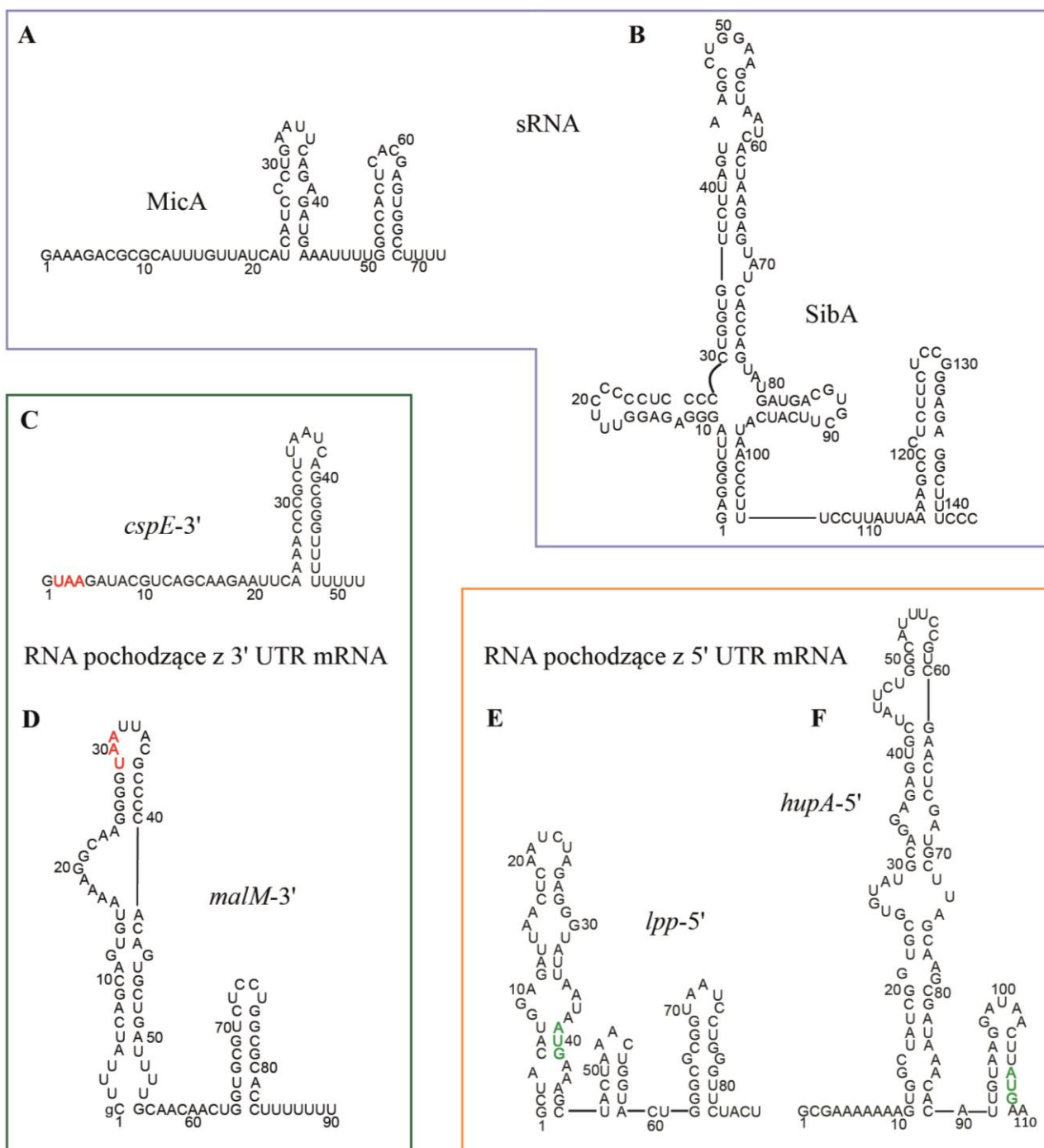
5. Wyniki

5.1. Określenie cech cząsteczki RNA istotnych dla wiązania z N-końcową domeną białka ProQ

5.1.1. Wyznaczenie siły oddziaływania wybranych cząsteczek RNA z białkiem ProQ i jego wyizolowaną domeną FinO

Do zbadania, jakie cechy cząsteczki RNA determinują wiązanie do białka ProQ, spośród ligandów o potwierdzonym oddziaływaniu z tym białkiem wybrałam 6 strukturalnie zróżnicowanych cząsteczek o różnym pochodzeniu (Rys 10.). Podane cząsteczki zostały zidentyfikowane jako ligandy białka ProQ u bakterii *E. coli* w badaniach CLIP-seq i RIL-seq (Holmqvist i wsp. 2018; Melamed i wsp. 2020). Dwie spośród wybranych cząsteczek – MicA i SibA – należą do grupy sRNA, krótkich niekodujących cząsteczek RNA, pozostałe cząsteczki pochodzą z regionów niekodujących mRNA: *cspE-3'* i *malM-3'* z regionu 3'-UTR, natomiast *hupA-5'* i *lpp-5'* z regionu 5'-UTR.

MicA jest kodowanym w *trans* sRNA o długości 72 nukleotydów, które wiąże zarówno białko ProQ i białko Hfq (Melamed i wsp. 2020). MicA negatywnie reguluje ekspresję wielu genów, m.in. *ompA* kodującego białko A błony zewnętrznej (Udekwa i wsp. 2005) czy *phoP* kodującego część dwuskładnikowego szlaku sygnałowego zależnego od jonów magnezu (Coornaert i wsp. 2010). sRNA SibA jest kodowany w *cis* jako jeden z komponentów, obok białka IbsA, systemu toksyna-antytoksyna i ma długość 144 nt (Fozo i wsp. 2008). Fragment *malM-3'* o długości 91 nt powstaje jako produkt obróbki potranskrypcyjnej z regionu 3'-UTR mRNA *malM*, natomiast 52-nt RNA *cspE-3'* kodowany jest jako niezależny transkrypt w obrębie regionu 3'-UTR mRNA *cspE* (Melamed i wsp. 2020). Fragment regionu 5'-UTR mRNA *lpp* zidentyfikowano jako komponent dwóch spośród najliczniejszych chimer związanych z białkiem ProQ (Melamed i wsp. 2020), natomiast homologiczny region 5'-UTR mRNA *hupA* opisano jako fragment regulowany przez sRNA RaiZ, który jest zależny od ProQ (Smirnov i wsp. 2017).

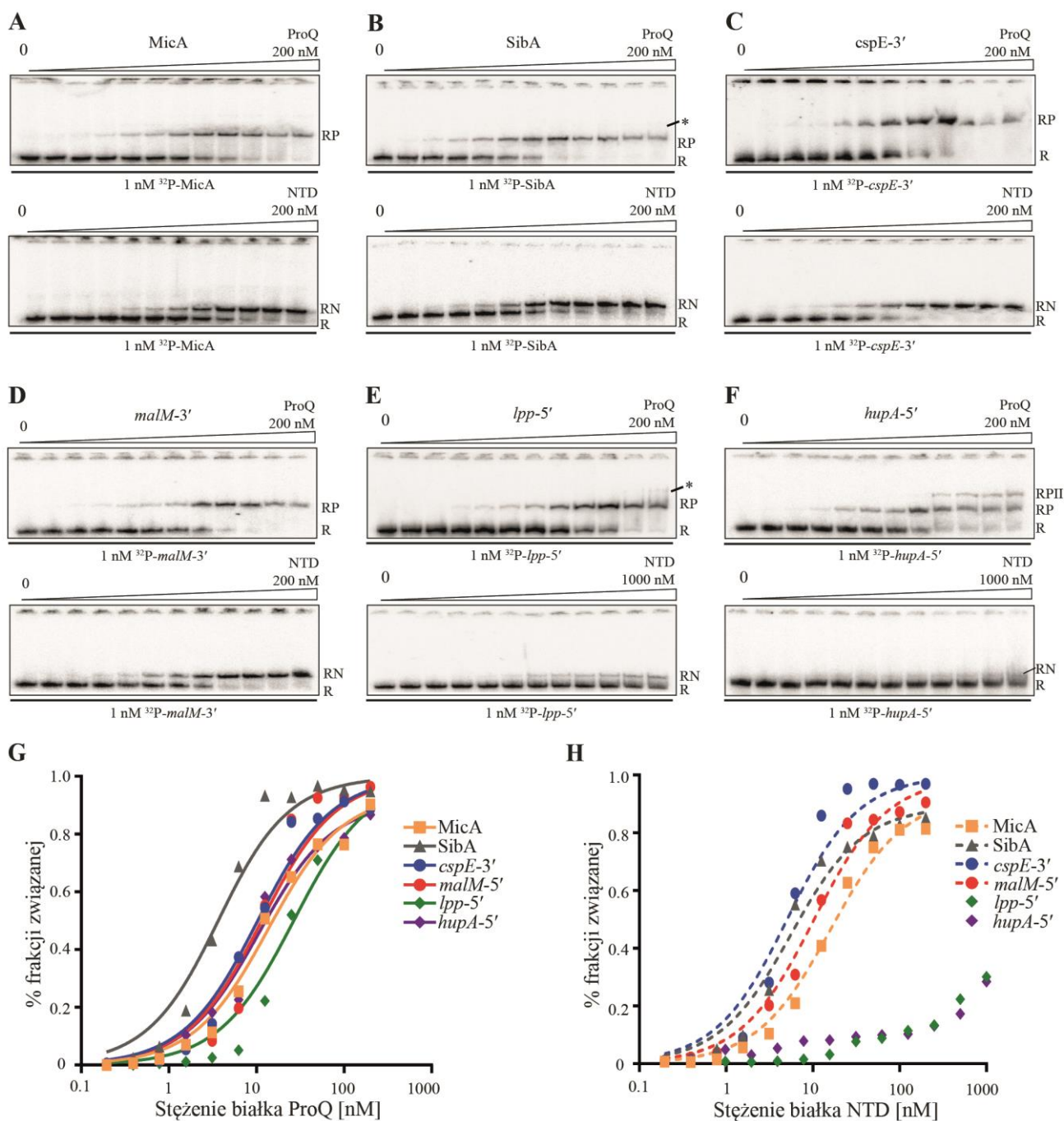


Rysunek 10. Częsteczki RNA o zróżnicowanym pochodzeniu i budowie użyte w eksperymencie wyznaczenia stałej dysocjacji w równowadze.

Struktury drugorzędowe cząsteczek sRNA (A) MicA (Udekwu i wsp. 2005) i (B) SibA (Smirnov i wsp. 2016) zaprezentowano zgodnie z referencjami. Struktury drugorzędowe cząsteczek pochodzących z regionów 3'-UTR mRNA (C) *cspE-3'* i (D) *malM-3'* oraz pochodzących z regionów 5'-UTR mRNA (E) *lpp-5'* i (F) *hupA-5'* zostały przygotowane przy użyciu oprogramowania RNAstructure. Kodony START (AUG) oznaczono kolorem **zielonym** a STOP (UAA) kolorem **czerwonym**. Małą literą „g” oznaczono guanozynę dodaną na początku sekwencji cząsteczki *malM-3'* w celu umożliwienia jej transkrypcji *in vitro* przez polimerazę z faga T7.

Zmierzyłam wiązanie wymienionych 6 cząsteczek RNA do oczyszczonego białka ProQ i jego 130-aa N-końcowej wyizolowanej domeny (NTD, ang. *N-terminal domain*) (Chaulk i wsp. 2011). Dla każdego z oddziaływań wyznaczyłam stałą równowagową reakcji dysocjacji K_d (Rys. 11, Tabela 1). Białko pełnej długości wiązało każdy z testowanych RNA z podobnym, silnym powinowactwem. Wartości K_d mieszczą się w zakresie od $2,4 \pm 0,85$ nM dla wiązania z SibA, do $21 \pm 5,9$ nM dla oddziaływania z *lpp-5'*. W wyniku wiązania z cząsteczkami MicA, SibA, *cspE-3'*, *malM-3'* i *lpp-5'* na żelu poliakrylamidowym widoczny jest pojedynczy kompleks RNA-białko. Natomiast oddziaływanie z cząsteczką *hupA-5'* uwidocznilo również dodatkowe kompleksy wyższego rzędu.

W badaniach *in vitro* dowiedziono, że 1-130 aa domena FinO białka ProQ z wysokim powinowactwem wiąże pochodną RNA FinP, naturalnego ligandu białka FinO (Chaulk i wsp. 2011). Powstało pytanie, czy wiązanie naturalnych ligandów ProQ do wyizolowanej domeny w jakiś sposób różni się od wiązania do białka pełnej długości? Wszystkie badane RNA tworzyły pojedynczy kompleks z NTD, jednak wartości K_d tych oddziaływań różnią się. Dla cząsteczek sRNA – MicA i SibA – stałe równowagowe dysocjacji dla wiązania z NTD są podobne jak w przypadku wiązania do pełnej długości ProQ (z K_d 21 ± 10 nM dla wiązania NTD-MicA i $15 \pm 5,1$ nM dla wiązania ProQ-MicA oraz wartościami $4,7 \pm 0,89$ nM dla reakcji NTD-SibA i $2,4 \pm 0,85$ nM dla ProQ-SibA) (Rys. 11, Tabela 1). RNA pochodzące z końców transkryptów mRNA również wiążą się do wyizolowanej domeny białka ProQ z dużym powinowactwem: dla *malM-3'* wartości K_d dla wiązania z ProQ i NTD są zbliżone (NTD-*malM-3'* - $11 \pm 4,9$ nM, ProQ-*malM-3'* - $12 \pm 4,7$ nM), *cspE-3'* było wiązane przez NTD dwukrotnie silniej ($4,2 \pm 1,3$ nM) niż przez białko pełnej długości ($13 \pm 4,5$ nM). Natomiast cząsteczki RNA pochodzące z początków transkryptów – *hupA-5'* i *lpp-5'* – do NTD wiążą się dużo słabiej niż do białka pełnej długości: przy zakresie maksymalnego stężenia białka do 1 μ M reakcje nie osiągały wysycenia (ze średnią frakcją związaną poniżej 30%), co przełożyło się na wartości K_d niższe co najmniej 50- lub 100-krotnie w porównaniu do wiązania z białkiem pełnej długości.



Rysunek 11. Wiązanie badanych RNA do białka ProQ z *E. coli* i jego N-końcowej domeny (NTD).

Wiązanie 1 nM ³²P-RNA (A) MicA, (B) SibA, (C) *cspE-3'*, (D) *malM-3'*, (E) *lpp-5'* i (F) *hupA-5'* do białek ProQ i NTD monitorowano przy wykorzystaniu techniki EMSA, zakres stężeń białek użytych w eksperymencie wskazano nad każdym żelem. Niezwiązane ³²P-RNA oznaczono symbolem R, RP to kompleks RNA-ProQ, natomiast RN to kompleks RNA-NTD. Kompleks wyższego rzędu z reakcji z *hupA-5'* oznaczono RPII, natomiast * wskazuje dodatkowy kompleks SibA-ProQ, który nie został skwantyfikowany z uwagi na niski stosunek sygnału do tła. Dane zależności frakcji związanej RNA od stężenia (G) ProQ i (H) NTD zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 1.

Tabela 1. Wiązanie cząsteczek RNA i ich mutantów z wydłużonymi lub zmienionymi końcami 3' do pełnej długości białka ProQ i jego N-końcowej domeny (NTD).

³² P-RNA	K_d [nM] (% średniej max. frakcji związanej)	
	ProQ	NTD
MicA	15 ± 5,1 (94%)	21 ± 10 (82%)
SibA	2,4 ± 0,85 (97%)	4,7 ± 0,89 (84%)
<i>cspE-3'</i>	13 ± 4,5 (82%)	4,2 ± 1,3 ^a (96%)
<i>malM-3'</i>	12 ± 4,7 (90%)	8,2 ± 1,8 ^a (85%)
<i>lpp-5'</i>	21 ± 5,9 ^a (93%)	> 1000 (28%)
<i>hupA-5'</i>	9,3 ± 4,1 ^b (86%)	> 1000 (19%)
<i>lpp-5'-mut</i>	3 ± 1,3 (94%)	1,3 ± 0,83 (83%)
<i>lpp-5'-term</i>	3,6 ± 1,3 (91%)	5,1 ± 1 (85%)
<i>lpp-5'-hp</i>	20 ± 2,8 ^b (78%)	n. o.
<i>hupA-5'-mut</i>	14 ± 8 ^b (91%)	16 ± 4,7 (64%)
<i>hupA-5'-term</i>	10 ± 5 (91%)	8,5 ± 0,96 (82%)
<i>hupA-5'-hp</i>	18 ± 1,9 ^b (80%)	n. o.
<i>cspE-3'-ext+5</i>	31 ± 14 ^a (87%)	17 ± 7,5 (89%)
<i>cspE-3'-ext+22</i>	17 ± 8,7 (83%)	n. o.
<i>malM-3'-ext+5</i>	33 ± 4,2 (82%)	n. o.
<i>malM-3'-ext+17</i>	8,4 ± 2,6 ^b (79%)	n. o.

Wartości K_d uzyskano przez dopasowanie danych wykresu zależności frakcji związanego RNA od stężenia białka do równania kwadratowego w programie GraphPad. Średnie wartości K_d wraz z odchyleniem standardowym obliczono na podstawie co najmniej 3 powtórzeń z niezależnych eksperymentów.

^a – w dopasowaniu danych do równania zastosowano wartość 100% jako wartość końcową maksymalną;

^b – w dopasowaniu danych do równania zastosowano sumę wszystkich widocznych kompleksów z białkiem;

n.o. – nie obserwowano wiązania w zakresie stężeń do 1 μ M.

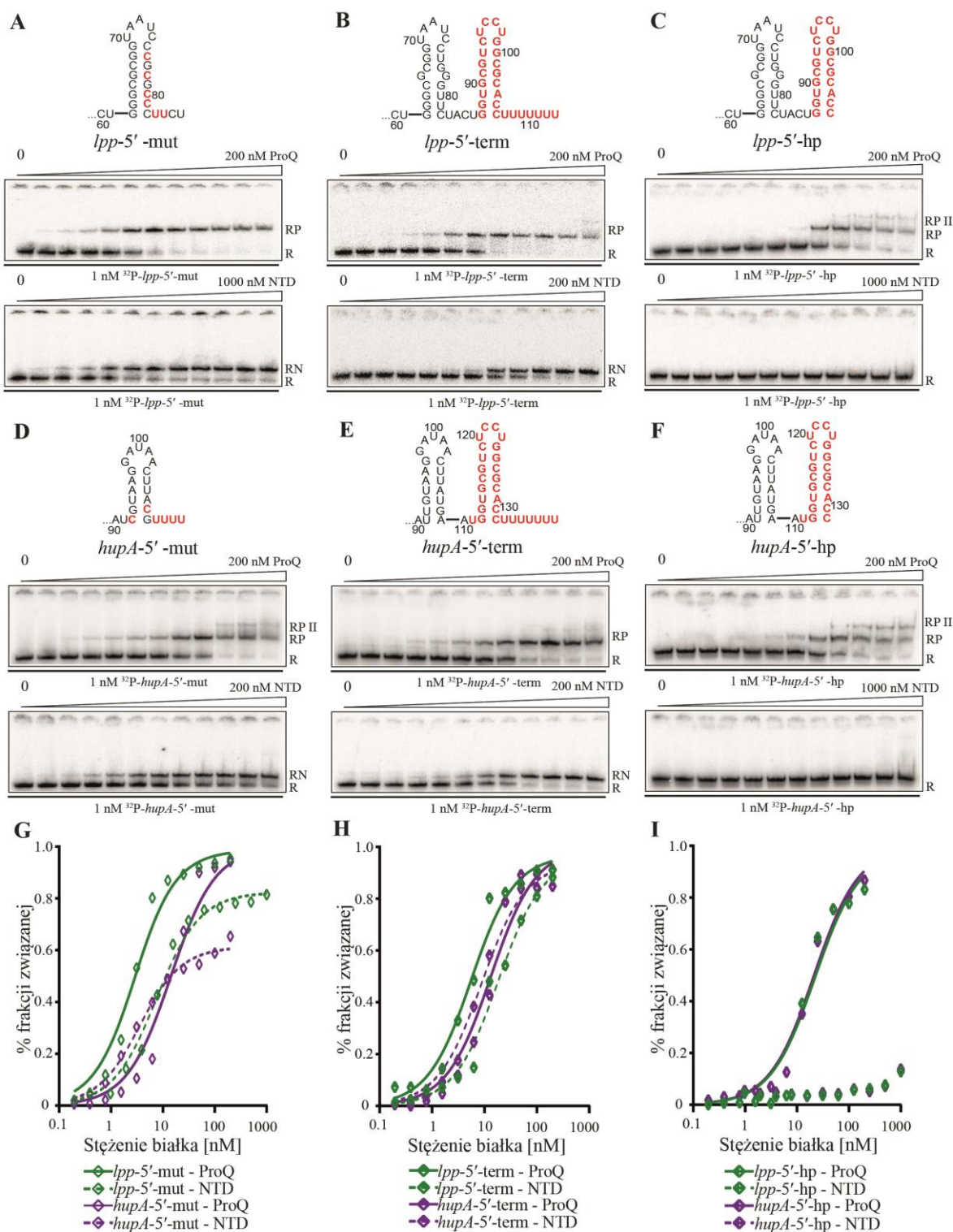
Poczynione obserwacje zachęciły mnie do ponownej analizy właściwości struktur badanych cząsteczek RNA w poszukiwaniu cech odróżniających cząsteczki o obniżonym wiązaniu do NTD. Co ciekawe cząsteczkami słabo wiążącymi domenę N-końcową białka ProQ były obie cząsteczki pochodzące z regionów 5'-UTR mRNA. Cząsteczki te nie posiadają Rho-niezależnego terminatora transkrypcji, którym jest znajdująca się na 3'-

końcu struktura typu spinki do włosów zakończona ogonem poliurydynowym, dlatego w kolejnym kroku postanowiłam sprawdzić, czy właśnie ta struktura jest niezbędna, by cząsteczka RNA została rozpoznana przez wyizolowaną N-końcową domenę białka ProQ.

5.1.2. Udział struktury terminatora transkrypcji niezależnego od białka Rho w oddziaływaniu z N-końcową domeną białka ProQ

Aby sprawdzić, czy obecność struktury o cechach Rho-niezależnego terminatora transkrypcji poprawi wiązanie cząsteczek *lpp-5'* i *hupA-5'* do NTD, do znajdujących się na ich 3'-końcach niestabilnych struktur przypominających spinki końcowe wprowadziłam mutacje mające na celu upodobnienie ich do struktury terminatora transkrypcji (Rys. 12, Tabela 1). Aby to osiągnąć do obu cząsteczek RNA wprowadziłam mutacje punktowe w celu ustabilizowania struktury dwuniciowej typu spinki do włosów (głównie poprzez zamianę par G-U w pary G-C) oraz dodałam krótki czteronukleotydowy ogon poliU. Następnie zmierzyłam wiązanie RNA *lpp-5'-mut* i *hupA-5'-mut* do białka ProQ pełnej długości i jego N-końcowej domeny. Wprowadzone zmiany znacznie poprawiły wiązanie tych cząsteczek RNA do wyizolowanej N-końcowej domeny ProQ, ponieważ wartość K_d dla ich oddziaływania z NTD obniżyła się do poziomu podobnego jak dla białka ProQ pełnej długości. Wartości K_d wynosiły odpowiednio dla *lpp-5'-mut* z ProQ: $3 \pm 1,3$ nM, z NTD: $1,3 \pm 0,83$ nM; dla *hupA-5'-mut* z ProQ: 14 ± 8 nM, z NTD: $16 \pm 4,7$ nM, podczas gdy średnia frakcja związana dla tych reakcji była niższa niż dla reakcji z ProQ o ok. 10% dla *lpp-5'-mut* i o prawie 30% dla *hupA-5'-mut*.

Następnie postanowiłam sprawdzić, czy dodanie na 3'-końcu tych cząsteczek naturalnego Rho-niezależnego terminatora pochodzącego z innej cząsteczki RNA również przywróci wiązanie cząsteczek RNA do NTD. W tym celu na 3' końcach cząsteczek *lpp-5'* i *hupA-5'* dodałam strukturę terminatora pochodzącego z cząsteczki *malM-3'*, dzięki czemu powstały RNA *lpp-5'-term* i *hupA-5'-term* (Rys. 12, Tabela 1). Zabieg ten poskutkował pełnym przywróceniem wiązania do NTD z wartościami K_d wynoszącymi $5,1 \pm 1$ nM dla *lpp-5'-term* i $8,5 \pm 0,96$ nM dla *hupA-5'-term*. Cząsteczki *lpp-5'-term* i *hupA-5'-term* wiązały się do NTD z powinowactwem zbliżonym do cząsteczek *lpp-5'-mut* i *hupA-5'-mut*, jednak przy wyższym o ok. 20% poziomie wysycenia reakcji dla *hupA-5'-mut*. W przypadku oddziaływania z pełnej długości białkiem ProQ zaobserwowałam ok. siedmiokrotnie silniejsze wiązanie cząsteczek *lpp-5'-mut* i *lpp-5'-term* w porównaniu do oddziaływania z cząsteczką *lpp-5'*, podczas gdy mutanty *hupA-5'* wiązały się podobnie jak cząsteczka o naturalnej sekwencji.



Rysunek 12. Dodanie struktury spinki z ogonem poliU na 3'-końcu *lpp-5'* i *hupA-5'* powoduje ich silne wiązanie przez NTD.

Wiązanie 1 nM ³²P-RNA (A) *lpp-5'*-mut, (B) *lpp-5'*-term, (C) *lpp-5'*-hp, (D) *hupA-5'*-mut, (E) *hupA-5'*-term i (F) *hupA-5'*-hp do białek ProQ i NTD monitorowano przy wykorzystaniu techniki EMSA, zakres stężeń białek użytych w eksperymencie wskazano nad każdym żelem. Na sekwencjach kolorem **czeryonym** zaznaczono mutacje punktowe oraz wprowadzone sekwencje. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia ProQ lub NTD zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad, na wykresach (G), (H), (I) zaprezentowane dane dla par mutantów *lpp-5'* i *hupA-5'* o analogicznych zmianach. Procent frakcji związanej oraz wartości K_d dla reakcji, w których zaobserwowano kompleksy wyższych rzędów, obliczono po zsumowaniu

uwidocznionych kompleksów. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 1.

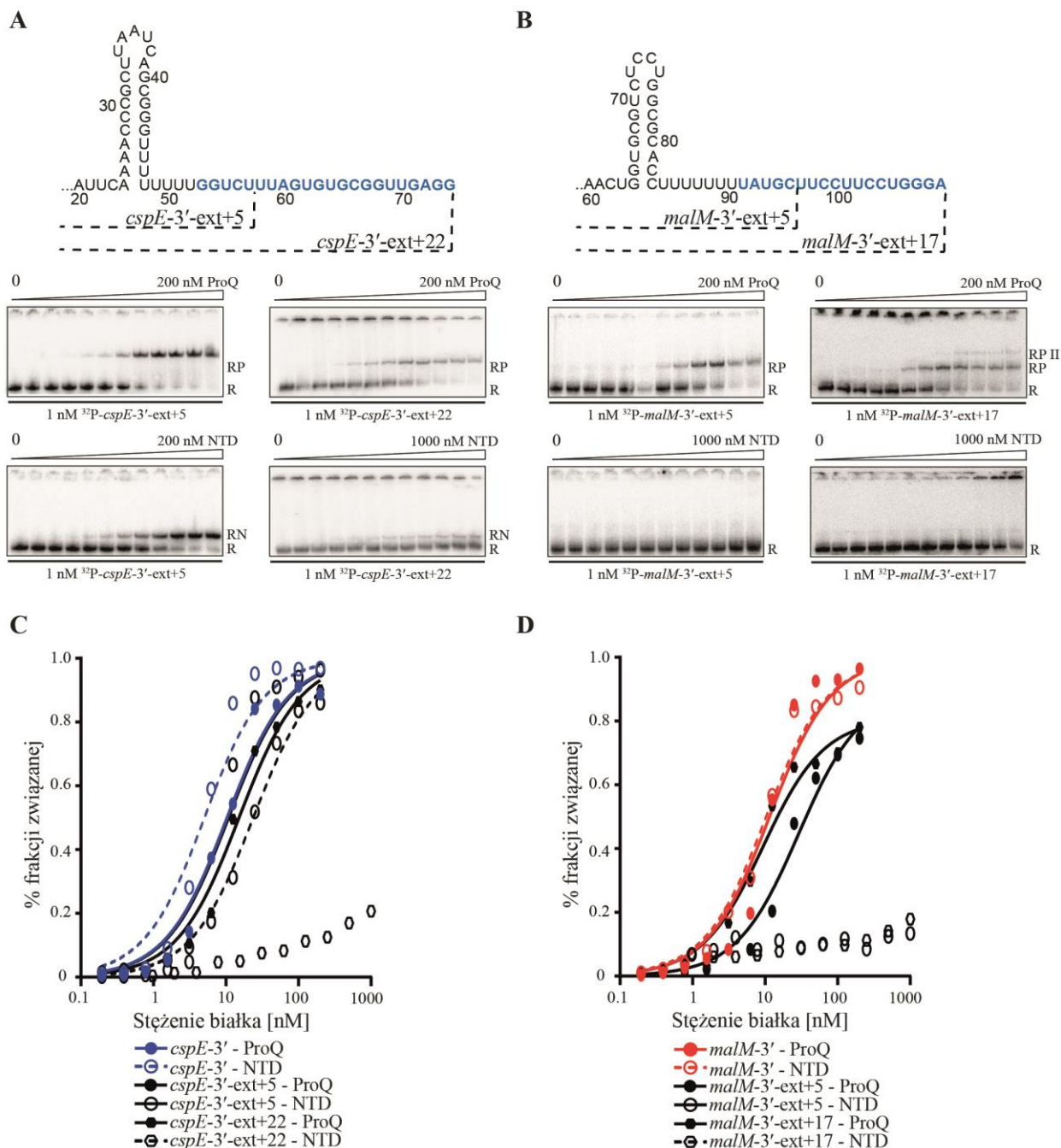
W kolejnym kroku z 3' końców cząsteczek *lpp-5'-term* i *hupA-5'-term* usunęłam naturalnie występującą w strukturze terminatora sekwencję 7 urydyn, co spowodowało, że cząsteczki te kończyły się strukturą w pełni dwuniciową. Wiązanie tak powstałych cząsteczek *lpp-5'-hp* i *hupA-5'-hp* do NTD w zakresie stężeń do 1 μ M było praktycznie niewykrywalne, podczas gdy ich wiązanie do pełnej długości białka ProQ było porównywalne z wiązaniem cząsteczek RNA o naturalnej sekwencji, z wartościami K_d $20 \pm 2,8$ nM dla *lpp-5'-hp* i $18 \pm 1,9$ nM dla *hupA-5'-hp*. Oznacza to, że nie tylko struktura spinki, ale także jednoniciowa sekwencja poliU na 3'-końcu RNA jest niezbędna do wiązania przez NTD.

Przeprowadzone analizy wskazują, że struktura złożona ze spinki oraz jednoniciowej sekwencji poliurydynowej jest niezbędna dla wiązania RNA przez N-końcową domenę białka ProQ. Sugeruje to, że naturalnie występujące w cząsteczkach RNA struktury terminatorów transkrypcji niezależnych od białka Rho są miejscami wiązania tej domeny.

5.1.3. Znaczenie 3'-terminalnej lokalizacji struktury spinki z sekwencją poliU dla oddziaływania RNA z N-końcową domeną białka ProQ

Aby sprawdzić, czy lokalizacja Rho-niezależnego terminatora transkrypcji na końcu 3' cząsteczki RNA jest istotna dla rozpoznania tej cząsteczki przez NTD, wiązaniu do obu badanych białek poddałam cząsteczki *cspE-3'* i *malM-3'* wydłużone powyżej naturalnie występujących 3'-końców (Rys. 13, Tabela 1). Przygotowałam po dwa warianty obu wydłużonych cząsteczek. Wydłużenie cząsteczki *cspE-3'* o 5 nt (*cspE-3'-ext+5*) nieznacznie osłabiło wiązanie do NTD (przy K_d $17 \pm 7,5$ nM), jednak wprowadzenie takiej samej zmiany do cząsteczki *malM-3'* (*malM-3'-ext+5*) spowodowało, że w zastosowanym zakresie stężeń białka do 1 μ M, właściwie nie wykryto kompleksu z NTD. Dalsze wydłużanie cząsteczek RNA nie poprawiło ich oddziaływania z NTD, przy jednoczesnym silnym wiązaniu wszystkich wydłużonych wariantów cząsteczek do pełnej długości białka ProQ (jedynie dla wersji wydłużonych o 5 nukleotydów wartości K_d dla wiązania z ProQ były dwukrotnie niższe w porównaniu do wiązania cząsteczek o naturalnej sekwencji).

Ten wynik wskazuje, że dla silnej interakcji z N-kończową domeną białka ProQ istotna jest lokalizacja struktury Rho-niezależnego terminatora transkrypcji na samym 3'-końcu cząsteczki RNA.



Rysunek 13. Wydłużenie sekwencji RNA poza strukturę terminatora transkrypcji wpływa negatywnie na wiązanie do N-kończowej domeny białka ProQ.

Sekwencje cząsteczek *cspE-3'* i *malM-3'* wydłużono, by powstały cząsteczki (A) *cspE-3'-ext+5* i *cspE-3'-ext+22* oraz (B) *malM-3'-ext+5* i *malM-3'-ext+17*, nukleotydy, o które wydłużono cząsteczki, zaznaczono niebieskim kolorem. Wiązanie 1 nM ^{32}P -RNA do białek ProQ i NTD monitorowano przy wykorzystaniu techniki EMSA, zakres stężeń białek użytych w eksperymencie wskazano nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia ProQ i NTD zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad, na wykresach (C) i (D) zaprezentowane dane odpowiednio dla cząsteczek *cspE-3'* i *malM-3'* i ich mutantów. Procent frakcji

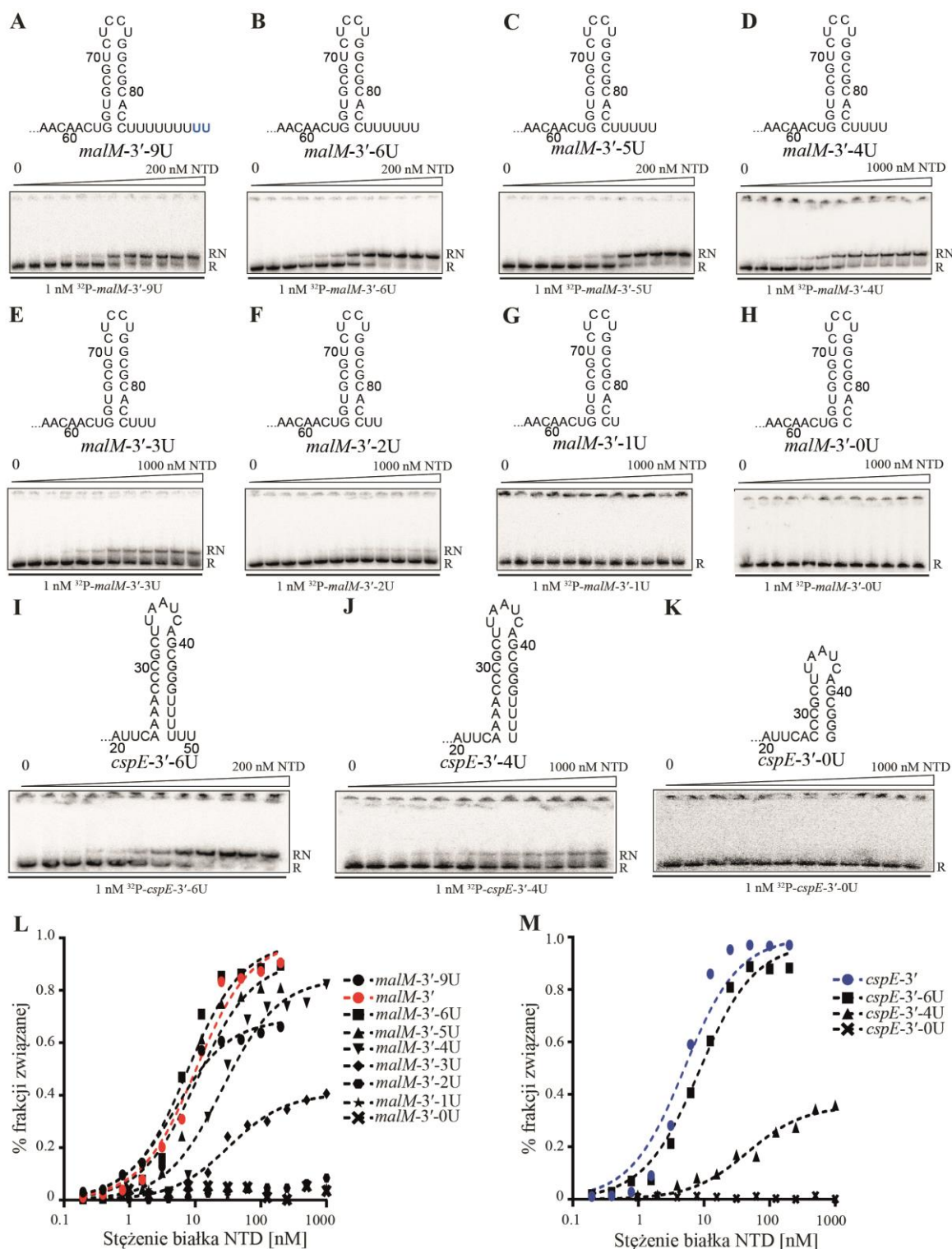
związanej oraz wartości K_d dla reakcji, w których zaobserwowano kompleksy wyższych rzędów, obliczono po zsumowaniu uwidocznionych kompleksów. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 1.

5.1.4. Znaczenie długości 3'-terminalnego ogona poliurydynowego dla wiązania przez domenę FinO białka ProQ

W celu sprawdzenia, jaka długość ogona poliU na 3'-końcu cząsteczki RNA jest niezbędna dla oddziaływania z N-końcową domeną białka ProQ, zaprojektowałam warianty cząsteczek *malM-3'* i *cspE-3'* z różną liczbą urydyn (Rys. 14, Tabela 2). Naturalnie w sekwencji cząsteczki *malM-3'* występuje ogon poliurydynowy o długości 7 nt, przygotowałam zatem wariant *malM-3'* o sekwencji ogona poliU wydłużonej do 9 nt oraz serię cząsteczek o sekwencjach skróconych o 1, 2, 3, 4, 5, 6 lub 7 urydyn. Dla porównania efektu na innej cząsteczce RNA, zaprojektowano wersje *cspE-3'*, który natywnie posiada 8 urydyn w ogonie poliU, obcięte o 2, 4 i 8 urydyn.

Wydłużenie ogona poliU o 2 urydyny nie wpłynęło na poprawę wiązania cząsteczki *malM-3'-9U* do NTD, przy K_d $5,1 \pm 2,9$ nM siła oddziaływania jest taka sama jak w przypadku cząsteczki o naturalnej sekwencji ($8,2 \pm 1,8$ nM). Z kolei dla wariantów o skróconej sekwencji ogona poliurydynowego obserwowano pełen wachlarz efektów. Cząsteczki pozbawione 1 lub 2 urydyn (*malM-3'-6U* i *malM-3'-5U*) wiązały się do NTD podobnie jak cząsteczka o naturalnej sekwencji, kiedy usunięto kolejną urydynę, cząsteczka *malM-3'-4U* wiązała się dwukrotnie słabiej w porównaniu do *malM-3'* przy K_d 27 ± 11 nM. Po usunięciu czterech urydyn z naturalnej sekwencji, reakcja z RNA *malM-3'-3U* zawierającym tylko 3 urydyny w ogonie poliU nie uzyskiwała wysycenia w zakresie stężeń białka do $1 \mu\text{M}$ przy 30% średniej frakcji związanej, natomiast wiązanie mutantów o jeszcze mniejszej liczbie urydyn na 3'-końcu było w badanym zakresie stężeń niewykrywalne.

W przypadku mutantów cząsteczki *cspE-3'* mutant *cspE-3'-6U* pozbawiony 2 końcowych urydyn wiązał się do NTD przy K_d $6,8 \pm 1,4$ nM podobnie jak cząsteczka o naturalnej sekwencji ($4,2 \pm 1,3$ nM). Usunięcie kolejnych 2 urydyn spowodowało niemal dwudziestokrotne osłabienie wiązania cząsteczki *cspE-3'-4U* w porównaniu do *cspE-3'*, natomiast wiązanie wariantu *cspE-3'-0U* pozbawionego ogona poliU w zakresie stężeń do $1 \mu\text{M}$ nie było obserwowane.



Rysunek 14. Skrócenie ogona poliU cząsteczek RNA *malM-3'* i *cspE-3'* prowadzi do zaniku wiązania do N-końcowej domeny białka ProQ.

Wiązanie 1 nM ^{32}P -RNA (A) *malM-3'-9U*, (B) *malM-3'-6U*, (C) *malM-3'-5U*, (D) *malM-3'-4U*, (E) *malM-3'-3U*, (F) *malM-3'-2U*, (G) *malM-3'-1U*, (H) *malM-3'-0U* (I) *cspE-3'-6U*, (J) *cspE-3'-4U* i (K) *cspE-3'-0U* do białka NTD monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej mutantów (L) *malM-3'* i (M) *cspE-3'* od stężenia NTD zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 2.

Tabela 2. Wiązanie wariantów cząsteczek *malM-3'* i *cspE-3'* oraz ich mutantów z różną liczbą urydyn w ogonie poliU do N-końcowej domeny białka ProQ.

³² P-RNA	<i>K_d</i> [nM] (% średniej max. frakcji związanej)
	NTD
<i>malM-3'-9U</i>	5,1 ± 2,9 (70%)
<i>malM-3' (7U, wt)</i>	8,2 ± 1,8 ^a (85%)
<i>malM-3'-6U</i>	7,8 ± 1,2 (89%)
<i>malM-3'-5U</i>	9,7 ± 1,5 (82%)
<i>malM-3'-4U</i>	27 ± 11 (74%)
<i>malM-3'-3U</i>	> 1000 (30 %)
<i>malM-3'-2U</i>	n. o.
<i>malM-3'-1U</i>	n. o.
<i>cspE-3' (8U, wt)</i>	4,2 ± 1,3 ^a (96%)
<i>cspE-3'-6U</i>	6,8 ± 1,4 (95%)
<i>cspE-3'-4U</i>	79 ± 29 (34%)
<i>cspE-3'-0U</i>	n. o.

Wartości *K_d* uzyskano przez dopasowanie danych wykresu zależności frakcji związanego RNA od stężenia białka do równania kwadratowego w programie GraphPad. Średnie wartości *K_d* wraz z odchyleniem standardowym obliczono na podstawie co najmniej 3 powtórzeń z niezależnych eksperymentów.

^a – dane takie same jak w tabeli 1;

n.o. – nie obserwowano wiązania w zakresie stężeń do 1 μM.

Opisane wyniki sugerują, że długość co najmniej 4 nt 3'-końcowego ogona poliurydynowego jest niezbędna dla silnego wiązania z N-końcową domeną białka ProQ. Fakt, że wariant *cspE-3'-4U* z 4 pozostającymi urydynami wiąże się gorzej niż analogiczna wersja *malM-3'* może być wytłumaczony tym, że w cząsteczce *cspE-3'-4U* wszystkie urydyny są sparowane z adenozydami i tworzą trzon struktury spinki, co może wpływać na ich dostępność dla białka.

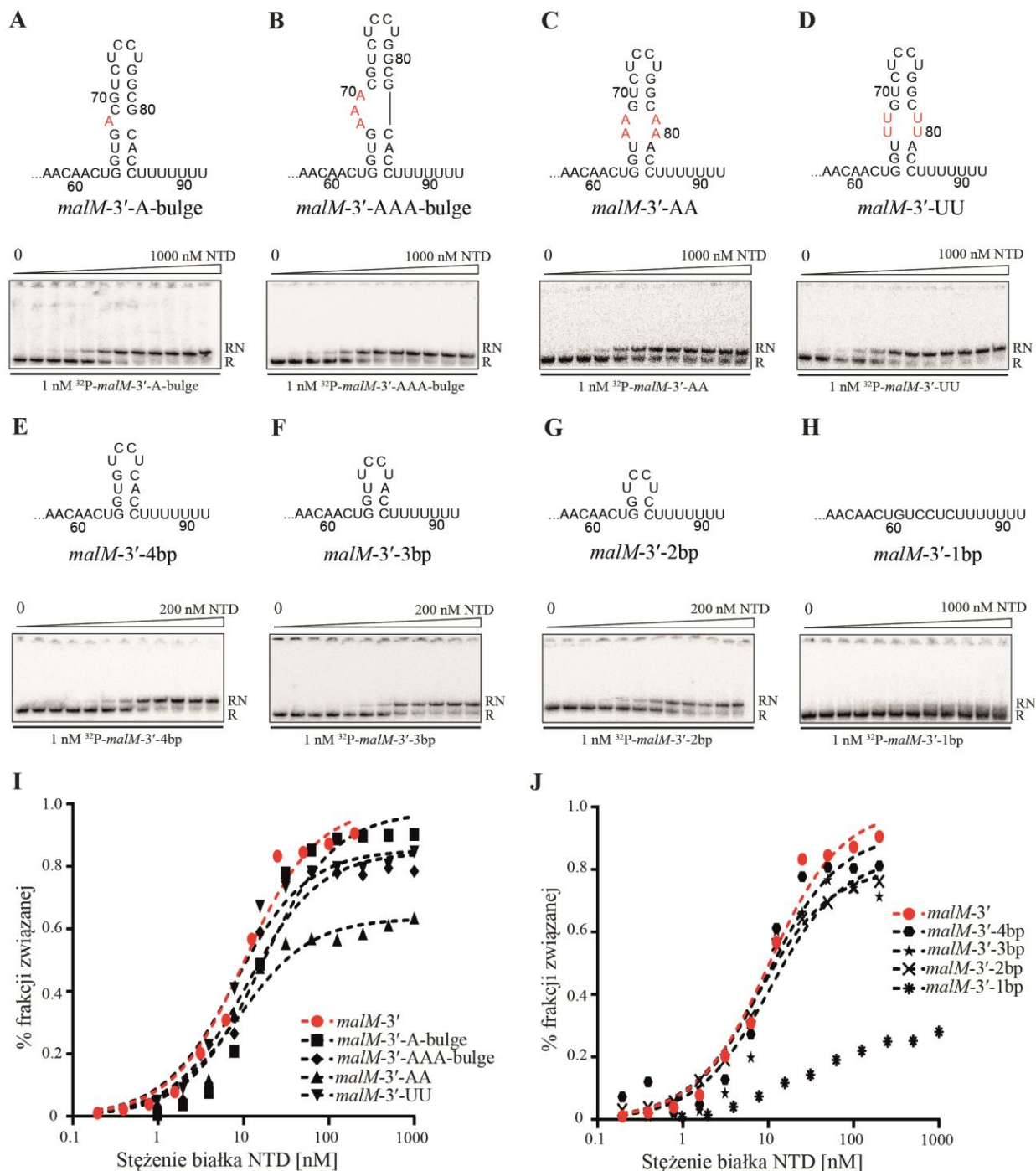
5.1.5. Znaczenie długości regionu dwuniciowego w obrębie spinki terminatora transkrypcji dla wiązania N-końcowej domeny białka ProQ

Aby sprawdzić, czy struktura trzonu spinki terminatorowej wpływa na rozpoznanie cząsteczki RNA przez N-końcową domenę białka ProQ, do cząsteczki *malM-3'* wprowadzono mutacje powodujące powstanie niesparowań w środkowej części trzonu spinki terminatorowej (Rys. 15, Tabela 3). Przygotowano cząsteczki, w których powyżej czwartej pary trzonu od dołu wprowadzono 1- lub 3-nukleotydowe wybrzuszenie, i takie,

gdzie powyżej trzeciej pary trzonu utworzono pętlę wewnętrzną zbudowaną z 4 adenozyń lub urydyn. Zastosowane zmiany nie miały znaczącego wpływu na wiązanie RNA do białka, ponieważ wartości K_d dla tych reakcji były podobne do wartości K_d dla wiązania z cząsteczką o naturalnej sekwencji. W przypadku wiązania z RNA *malM-3'-AA* zaobserwowano jednak obniżenie maksymalnej frakcji związanej o ok. 20%.

Uzyskane wyniki mogą sugerować, że zmiany wprowadzone powyżej trzeciej i czwartej pary regionu dwuniciowego zlokalizowane są poza obszarem rozpoznawanym przez NTD.

Dodatkowo w celu sprawdzenia, czy długość trzonu struktury spinki wpływa na oddziaływanie z cząsteczką RNA, zaprojektowałam mutanty *malM-3'* o trzonie skróconym od góry do 4, 3, 2 oraz 1 pary zasad (Rys. 15, Tabela 3). Mutanty *malM-3'* o 4, 3 i 2 parach zasad w trzonie spinki terminatorowej wiązały się do NTD z siłą podobną jak cząsteczka o naturalnej sekwencji, przy K_d o wartościach w zakresie $6,4 - 8,7 \pm 2,4$ nM. Natomiast cząsteczka *malM-3'-1bp* w zakresie stężeń białka do $1 \mu\text{M}$ wiązała się bardzo słabo, ze średnią frakcją związaną na poziomie 23%. W modelu tej cząsteczki przewidzianym w programie RNAstructure pojedyncze sparowanie nie zapewnia stabilności struktury dwuniciowej, więc bardzo prawdopodobne jest, że cząsteczka pozostaje liniowa. Zatem uzyskany wynik jest zgodny z dotychczasowymi obserwacjami mówiącymi o tym, że białko ProQ preferuje wiązanie do mocno ustrukturyzowanych cząsteczek RNA (Smirnov i wsp. 2016).



Rysunek 15. Dolna część spinki terminatorowej uczestniczy w wiązaniu cząsteczki *malM-3'*.

Wiązanie 1 nM ^{32}P -RNA (A) *malM-3'-A-bulge*, (B) *malM-3'-AAA-bulge*, (C) *malM-3'-AA*, (D) *malM-3'-UU*, (E) *malM-3'-3U*, (F) *malM-3'-2U* (G) *malM-3'-1U* i (H) *malM-3'-0U* do białka NTD monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. (I) Dane zależności frakcji związanej RNA od stężenia NTD zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 3. Na sekwencjach kolorem **czernym** zaznaczono wprowadzone mutacje.

Tabela 3. Wiązanie cząsteczek RNA i ich mutantów z wydłużonymi lub zmienionymi końcami 3' do pełnej długości białka ProQ i jego N-końcowej domeny.

³² P-RNA	K_d [nM] (% średniej max. frakcji związanej)
	NTD
<i>malM-3'</i>	12 ± 4,8 ^a
<i>malM-3'-AAA</i> bulge	9,0 ± 2,5 (77%)
<i>malM-3'-A</i> bulge	26 ± 8 (86%)
<i>malM-3'-AA</i>	6,6 ± 1,4 (65%)
<i>malM-3'-UU</i>	8,7 ± 2,1 (77%)
<i>malM-3'-4bp</i>	8,8 ± 2,4 (81%)
<i>malM-3'-3bp</i>	8,7 ± 2,4 (73%)
<i>malM-3'-2bp</i>	6,4 ± 2 (72%)
<i>malM-3'-1bp</i>	> 1000 (23%)

Wartości K_d uzyskano przez dopasowanie danych wykresu zależności frakcji związanego RNA od stężenia białka do równania kwadratowego w programie GraphPad. Średnie wartości K_d wraz z odchyleniem standardowym obliczono na podstawie co najmniej 3 powtórzeń z niezależnych eksperymentów.

^a – dane takie same jak w tabeli 1;

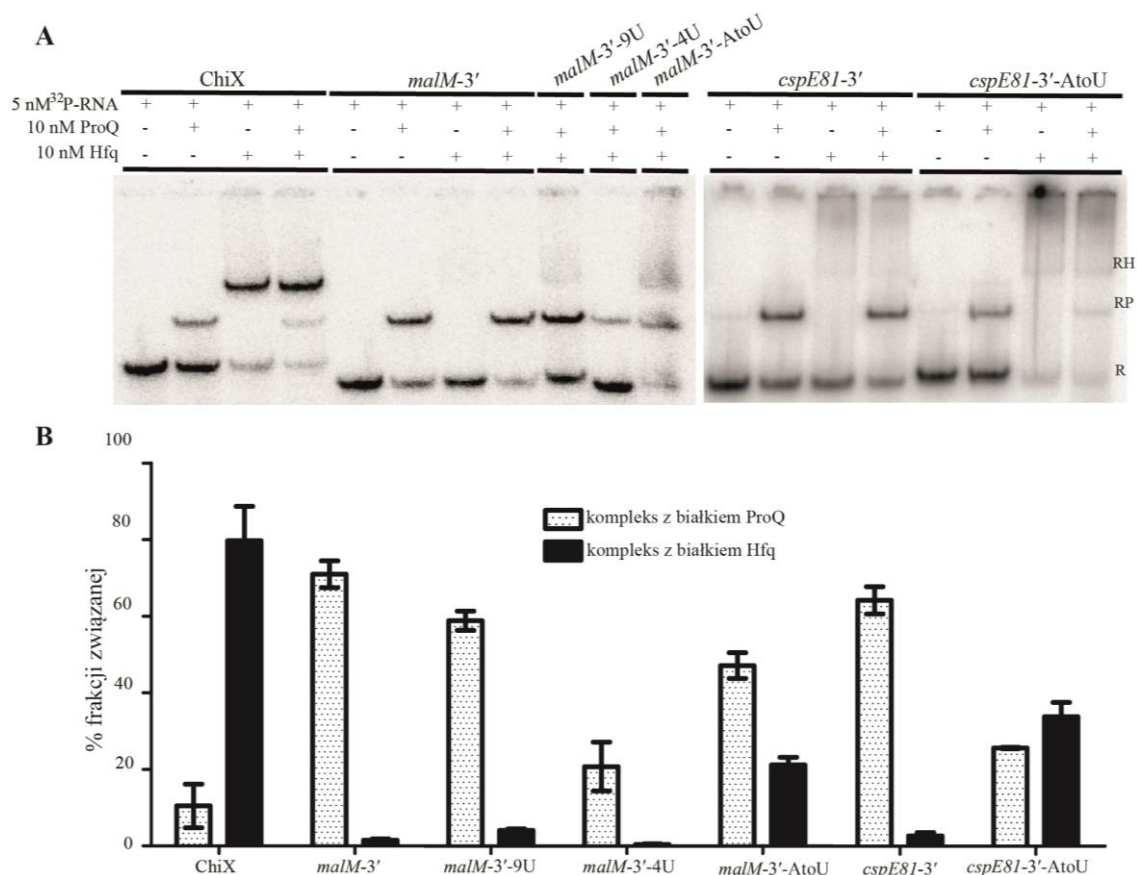
n.o. – nie obserwowano wiązania w zakresie stężeń do 1 μM.

5.1.6. Znaczenie motywu bogatego w adenozyne zlokalizowanego po 5'-stronie spinki terminatorowej w kompetycji pomiędzy białkami ProQ i Hfq o wiązanie cząsteczek RNA

Cechy cząsteczki RNA rozpoznawane przez N-końcową domenę białka ProQ przypominają znane motywy rozpoznawane przez inne białko opiekuńcze wiążące liczne cząsteczki RNA, białko Hfq, które także wiąże spinki zakończone ogonem poliU na 3' końcu (Ishikawa i wsp. 2012). Zbadano jednak, że oba białka wiążą różne pule RNA, które tylko częściowo nakładają się (Holmqvist i wsp. 2018). Jeżeli zarówno ProQ, jak i Hfq rozpoznają podobne motywy strukturalne, w jaki sposób odróżniają swoje ligandy? Czy istnieje jakaś cecha, która wpływa na rozpoznanie RNA tylko przez jedno z białek? Próbę odpowiedzi na to pytanie podjął mgr Maciej Basczok z Pracowni Biochemii RNA, który w oparciu o sekwencje najczęściej wiązanych ligandów białek ProQ i Hfq stworzył logo sekwencji 10 nukleotydów przylegających od 5' strony do trzonu spinki terminatorowej (Stein i wsp. 2020). Z jego analizy wynika, że w bezpośrednim sąsiedztwie trzonu spinki terminatorowej u ligandów ProQ znajduje się nadreprezentacja adenozyne, szczególnie

w czterech pozycjach najbliższych regionu sparowanego, natomiast w przypadku puli ligandów Hfq w odpowiadających im pozycjach przeważają urydyny.

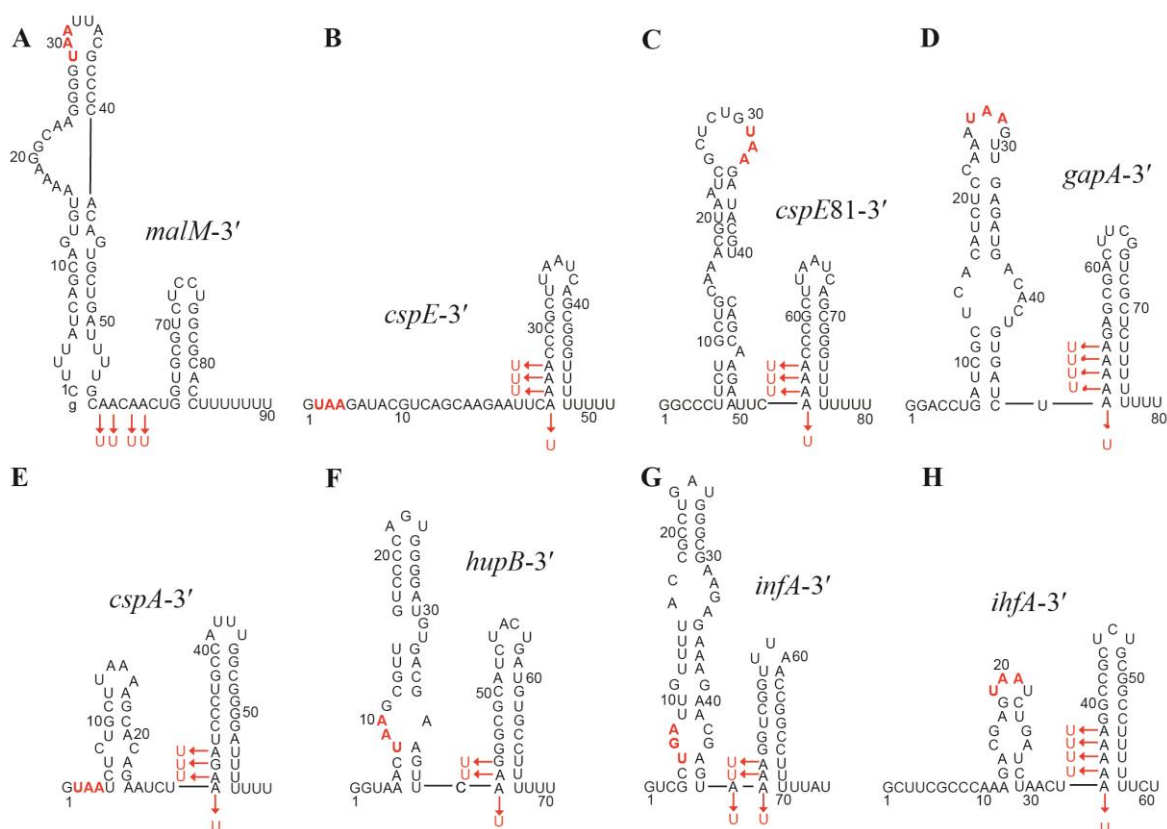
Aby porównać wiązanie RNA zmutowanych w obrębie motywu bogatego w adenozyne, przeprowadziłam wiązanie cząsteczki *malM-3'* i jej zmutowanych wersji *malM-3'-9U*, *malM-3'-4U* i *malM-3'-AtoU* oraz pary *cspE81-3'* i *cspE81-3'-AtoU* w warunkach kompetycji białek ProQ i Hfq (Rys. 16A, C; 17). Cząsteczka *cspE81-3'* została użyta uprzednio w badaniach *in vivo* wykorzystujących bakteryjny system trójhybrydowy (Pandey i wsp. 2020). Jako reakcję kontrolną zmierzyłam wiązanie z cząsteczką ChiX, która jest naturalnym ligandem Hfq. W tym eksperymencie mierzyłam wiązanie znakowanego RNA przez białka Hfq i pełnej długości ProQ w jednakowych stężeniach, a po 30 min przeprowadziłam rozdział reakcji na żelu poliakrylamidowym. W warunkach eksperymentalnych, tak jak się spodziewałam, niemal cała pula znakowanej cząsteczki ChiX została związana przez białko Hfq. W warunkach kompetycji cząsteczka *malM-3'*, przeciwnie niż ChiX, w znacznej większości wiązała się z białkiem ProQ. Jako że Hfq wiąże w większości RNA o co najmniej ośmionukleotydowym ogonie poliU (Otaka i wsp. 2011), wiązaniu w kompetycji poddałam także wersję *malM-3'-9U* z wydłużonym do 9 nt ogonem poliU. Chociaż frakcja tego RNA związana z Hfq była dwukrotnie większa niż w przypadku *malM-3'* o naturalnej 7 nt długości ogona poliU, to nadal większość RNA związana była z ProQ. W przypadku mutantu *malM-3'-4U* o zmniejszonej liczbie urydyn w ogonie poliU, zaobserwowałam obniżone wiązanie do obu białek, jednak ze znaczną przewagą frakcji z ProQ. Dla wiązania *malM-3'-AtoU* zaobserwowałam rosnącą frakcję związaną z Hfq, która stanowiła jedną trzecią całości frakcji związanej w reakcji i 14-krotnie wyższą niż w przypadku cząsteczki o naturalnej sekwencji. Podobną zależność zauważyłam dla wiązania cząsteczek *cspE81-3'* i *cspE81-3'-AtoU*; w warunkach kompetycji prawie 70% *cspE81-3'* wiązane było do ProQ, podczas gdy powstała znikoma ilość kompleksu z Hfq, natomiast w przypadku reakcji z mutantem *cspE81-3'-AtoU* powstało prawie 40% kompleksu z Hfq i ok. 30% kompleksu z ProQ. Przedstawione wyniki sugerują, że sekwencja bogata w adenozyne zlokalizowana po 5'-stronie spinki terminatorowej RNA w warunkach kompetycji decyduje o rozpoznaniu cząsteczki przez ProQ a nie przez Hfq.



Rysunek 16. Kompetycja o wiązanie RNA przez białka ProQ i Hfq.

(A) Analiza na żelu poliakrylamidowym wiązanie ³²P znakowanego 5 nM RNA z 10 nM pełnej długości białkiem ProQ i 10 nM białkiem Hfq po 30 min inkubacji. Niezwiązane ³²P-RNA oznaczono symbolem R, RP to kompleks RNA-ProQ, natomiast RH to kompleks RNA-Hfq. (B) Wykres słupkowy pokazujący procent frakcji związanej do każdego z białek w reakcji. Przedstawione wartości stanowią średnią z co najmniej 3 niezależnych powtórzeń.

Aby sprawdzić, czy motyw bogaty w adenozyne zlokalizowany po 5' stronie trzonu spinki terminatorowej RNA wiązanych przez białko ProQ jest uniwersalny dla ligandów białka ProQ, zaprojektowano szereg cząsteczek, w których ten motyw zmutowano i adenozyne zastąpiono urydynami (Rys. 17). Następnie pary cząsteczek w wersji o sekwencji naturalnej i zmienionej poddano wiązaniu w równowadze do białek NTD ProQ i Hfq. Eksperymenty wiązania do NTD zostały wykonane przez mnie, natomiast eksperymenty wiązania RNA do Hfq zostały przeprowadzone przez dr Joannę Kwiatkowską metodą retencji kompleksów na podwójnych filtrach.



Rysunek 17. Mutacje wprowadzone do sekwencji cząsteczek RNA w celu przetestowania roli motywu bogatego w adenozyne zlokalizowanego po 5'-stronie trzonu spinki terminatorowej.

Struktury drugorzędowe cząsteczek (A) *malM-3'*, (B) *cspE-3'*, (C) *cspE81-3'*, (D) *gapA-3'*, (E) *cspA-3'*, (F) *hupB-3'*, (G) *infA-3'* i (H) *ihfA-3'* zostały przygotowane przy użyciu oprogramowania RNAstructure. Mutacje wprowadzone w obrębie motywu bogatego w adenozyne zaznaczono kolorem czerwonym, małą literą „g” oznaczono guanozynę dodaną na początku sekwencji cząsteczki *malM-3'* w celu umożliwienia jej transkrypcji *in vitro* przez polimerazę z faga T7.

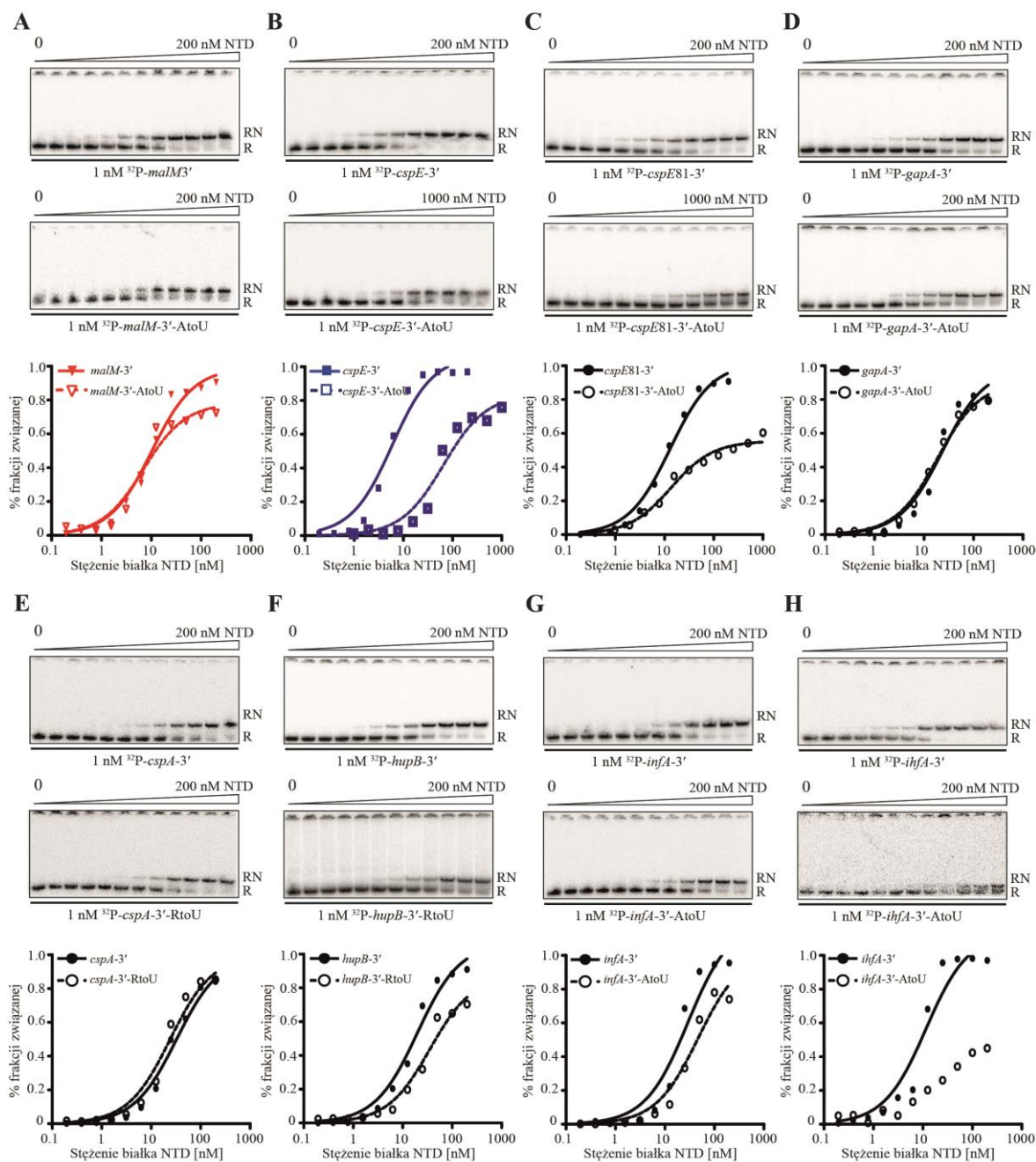
Spośród ligandów ProQ silnie wiązanych *in vivo* (Holmqvist i wsp. 2018; Melamed i wsp. 2020) wybrano 8 cząsteczek o dobrze zdefiniowanych motywach bogatych w adenozyne, mutacje wprowadzono w taki sposób, aby nie wpływać na zmianę predykcji struktury według oprogramowania RNAstructure. Oprócz cząsteczek *malM-3'* i *cspE81-3'* użytych w eksperymencie kompetycyjnym, do badania wybrano także *gapA-3'*, *cspA-3'*, *hupB-3'*, *infA-3'* i *ihfA-3'* (Rys. 17). Zbadano również wiązanie dwóch wariantów cząsteczki *cspE-3'* – *cspE-3'* o długości 52-nt i *cspE81-3'* 81-nt. Oba występują *in vivo* (Holmqvist i wsp. 2018; Melamed i wsp. 2020). Następnie zamieniono na urydyny reszty adenozynowe, oraz rzadziej guanozynowe (w przypadku cząsteczek *cspA-3'* i *hupB-3'*), które tworzyły pary zasad z resztami ogona poliurydynowego.

W przypadku wiązania do N-końcowej domeny białka ProQ substytucje urydynowe w badanym motywie nie miały wpływu na wiązanie lub osłabiły wiązanie RNA (Rys. 18, Tabela 4). Dla cząsteczek *malM-3'*, *cspE81-3'*, *gapA-3'*, *cspA-3'*, *hupB-3'*

i *infA-3'* mutacja motywu bogatego w adenozyne nie miała wpływu na wiązanie lub je nieznacznie (maksymalnie dwukrotnie) osłabiła. Większy efekt obserwowałam dla cząsteczek *cspE-3'* i *ihfA-3'*, u których mutacja motywu obniżyła K_d przynajmniej 10-krotnie. Warto także zauważyć, że niemal we wszystkich przypadkach (oprócz *gapA-3'* i *cspA-3'*) dla reakcji z mutantami zanotowałam co najmniej 10% spadek średniej frakcji związanej, a dla pary RNA *ihfA-3'* i *ihfA-3'*-AtoU spadek wynosił aż 50%.

Przeciwny efekt mutacji obserwowano dla wiązania RNA z białkiem Hfq (Rys. 19, Tabela 4) – substytucje poprawiły wiązanie wszystkich badanych cząsteczek. Dla większości z nich wartości K_d spadły ponad 10-krotnie – największy efekt zanotowano dla *ihfA-3'*-AtoU i *cspA-3'*-RtoU. Dla pary *gapA-3'* różnica była 8-krotna, natomiast najmniejszy wzrost siły wiązania (3,5-krotny) obserwowano dla zmutowanego wariantu *cspE-3'*.

W tym miejscu warto przyjrzeć się różnicy w wiązaniu do cząsteczek *cspE-3'*, *cspE81-3'* i ich zmutowanych wersji. Białko Hfq wiąże cząsteczkę *cspE-3'* z K_d $5,9 \pm 2,5$ nM, wersję *cspE-3'*-AtoU z K_d $1,7 \pm 0,88$ nM, czyli ponad trzykrotnie mocniej. Z kolei wariant 81-nt z wydłużonym fragmentem z 5'-strony białko Hfq wiąże prawie 8 razy słabiej niż wersję 52-nt (K_d 47 ± 11 nM), natomiast RNA *cspE81-3'*-AtoU oddziałuje z białkiem Hfq 18 razy mocniej w porównaniu do *cspE81-3'*, co daje podobną wartość K_d jak wiązanie *cspE-3'*-AtoU ($2,6 \pm 1,2$ nM). 52-nt *cspE-3'* po 5'-stronie spinki terminatorowej posiada ponad 20-nt element sekwencji jednoniciowej, natomiast w dłuższej 81-nt cząsteczce *cspE81-3'* fragment ten jest uwikłany w rozległą strukturę drugorzędową. Sekwencja jednoniciowa może stanowić dodatkowe miejsce wiązania dla białka Hfq, stąd efekt substytucji adenozyne do urydyn w cząsteczce *cspE-3'* może być po części maskowany.



Rysunek 18. Mutacja w obszarze motywu bogatego w adenyzy zlokalizowanego po 5'-stronie trzonu spinki terminatorowej nie wpływa lub osłabia wiązanie do N-końcowej domeny białka ProQ.

Wiązanie 1 nM 32 P-RNA (A) *malM-3'* i *malM-3'-AtoU*, (B) *cspE-3'* i *cspE-3'-AtoU*, (C) *cspE81-3'* i *cspE81-3'-AtoU*, (D) *gapA-3'* i *gapA-3'-AtoU*, (E) *cspA-3'* i *cspA-3'-RtoU*, (F) *hupB-3'* i *hupB-3'-RtoU*, (G) *infA-3'* i *infA-3'-AtoU* i (H) *ihfA-3'* i *ihfA-3'-AtoU* do białka NTD monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia NTD zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 4.

Tabela 4. Wiązanie cząsteczek RNA i ich mutantów ze zmutowanym regionem bogatym w adenozyne do N-końcowej domeny białka ProQ i Hfq.

³² P-RNA	K_d [nM] (% średniej max. frakcji związanej)	
	NTD ^a	Hfq ^b
<i>malM-3'</i>	12 ± 4,8 ^c	22 ± 7
<i>malM-3'-AtoU</i>	5,0 ± 1,8 (68%)	2,2 ± 0,8
<i>cspE-3'</i>	4,5 ± 1,5 ^c	5,9 ± 2,5
<i>cspE-3'-AtoU</i>	100 ± 51 (69%)	1,7 ± 0,88
<i>cspE81-3'</i>	7,7 ± 1,8 (86%)	47 ± 11
<i>cspE81-3'-AtoU</i>	17 ± 4,6 (57%)	2,6 ± 1,2
<i>cspA-3'</i>	33 ± 10 ^d (87%)	49 ± 15 ^d
<i>cspA-3'-RtoU</i>	42 ± 18 ^d (83%)	1,3 ± 0,66
<i>hupB-3'</i>	18 ± 1,7 (90%)	30 ± 16 ^d
<i>hupB-3'-RtoU</i>	29 ± 7,2 (79%)	1,9 ± 0,75
<i>infA-3'</i>	19 ± 10 (95%)	28 ± 12 ^d
<i>infA-3'-AtoU</i>	32 ± 13 (75%)	1,2 ± 0,55
<i>ihfA-3'</i>	10 ± 1,5 (98%)	64 ± 15
<i>ihfA-3'-AtoU</i>	5 ± 0,9 (46%)	0,67 ± 0,35
<i>gapA-3'</i>	22 ± 8,8 (84%)	11 ± 0,37
<i>gapA-3'-AtoU</i>	19 ± 8,5 (79%)	90 ± 24

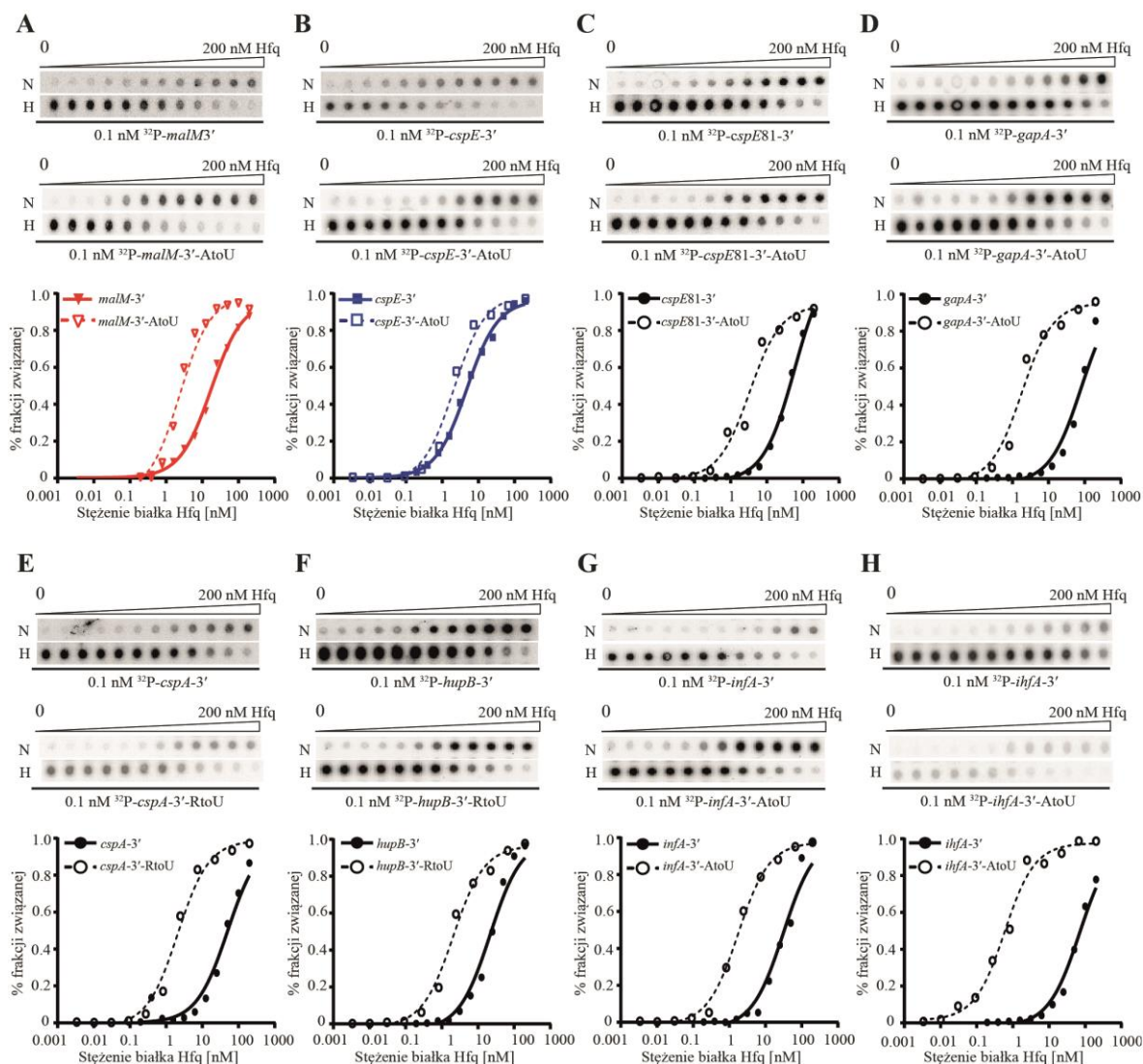
Wartości K_d uzyskano przez dopasowanie danych wykresu zależności frakcji związanego RNA od stężenia białka do równania kwadratowego w programie GraphPad. Średnie wartości K_d wraz z odchyleniem standardowym obliczono na podstawie co najmniej 3 powtórzeń z niezależnych eksperymentów.

^a – wiązanie RNA do NTD mierzone było przy pomocy EMSY;

^b – wiązanie RNA do Hfq mierzone było przy użyciu metody retencji na podwójnych filtrach, eksperyment przeprowadzony przez dr Joannę Kwiatkowską;

^c – dane takie same jak w tabeli 1;

^d - w fitowaniu zastosowano wartość 100% jako wartość końcową maksymalną.



Rysunek 19. Mutacja w obszarze motywu bogatego w adenozyne zlokalizowanego po 5'-stronie trzonu spinki terminatorowej poprawia wiązanie do białka Hfq.

Wiązanie 0,1 nM ³²P-RNA (A) *malM-3'* i *malM-3'-AtoU*, (B) *cspE-3'* i *cspE-3'-AtoU*, (C) *cspE81-3'* i *cspE81-3'-AtoU*, (D) *gapA-3'* i *gapA-3'-AtoU*, (E) *cspA-3'* i *cspA-3'-RtoU*, (F) *hupB-3'* i *hupB-3'-RtoU*, (G) *infA-3'* i *infA-3'-AtoU* i (H) *ihfA-3'* i *ihfA-3'-AtoU* do białka NTD monitorowano przy wykorzystaniu metody retencji na podwójnych filtrach, w zakresie stężeń podanym nad każdym żelem. Literą N oznaczono kompleks z białkiem na membranie nitrocelulozowej, natomiast literą H oznaczono wolne RNA zatrzymane na membranie Hybond-N⁺. Dane zależności frakcji związanej RNA od stężenia Hfq zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d zostały przedstawione w tabeli 4 [Eksperyment wykonany przez dr Joannę Kwiatkowską].

Przedstawione dane świadczą o tym, że substytucja adenozyne zlokalizowanych po 5'-stronie spinki terminatorowej w różny sposób wpływa na wiązanie białek Hfq i wyizolowanej N-końcowej domeny białka ProQ. Podczas gdy ma niewielki negatywny efekt lub nie ma wpływu na wiązanie z NTD, to wyraźnie poprawia wiązanie ligandów ProQ do białka Hfq. Powstaje zatem pytanie, czy obecność adenozyne po 5'-stronie spinki

terminatorowej cząsteczki RNA nie jest ewolucyjnie związane ze specjalizacją grupy cząsteczek RNA do preferencyjnego wiązania białka innego niż Hfq?

5.2. Poznanie znaczenia reszt aminokwasowych zlokalizowanych w obrębie domeny

FinO białka ProQ istotnych dla wiązania cząsteczek RNA

5.2.1. Analiza zakonserwowania sekwencji aminokwasowej wśród białek zawierających domenę FinO/ProQ u wybranych gatunków enterobakterii

W celu wyznaczenia cech N-końcowej domeny białka ProQ istotnych dla wiązania cząsteczek RNA, zdecydowałam się na przygotowanie mutantów białka zawierających substytucje pojedynczych aminokwasów. Aby wybrać kodony do mutagenazy, przeprowadziłam analizę opublikowanych badań *in vivo* dotyczących wpływu substytucji w obrębie badanej domeny ProQ i jej homologów na wiązanie RNA i funkcję białka (Ghetu i wsp. 2002; Chaulk i wsp. 2011; Attaiech i wsp. 2016) oraz przeprowadziłam analizę stopnia zakonserwowania sekwencji wśród białek o znanej strukturze należących do rodziny białek FinO występujących u enterobakterii (Rys. 20). Na tej podstawie wybrałam 10 reszt aminokwasowych, które potencjalnie mogłyby mieć udział w oddziaływaniu z RNA.

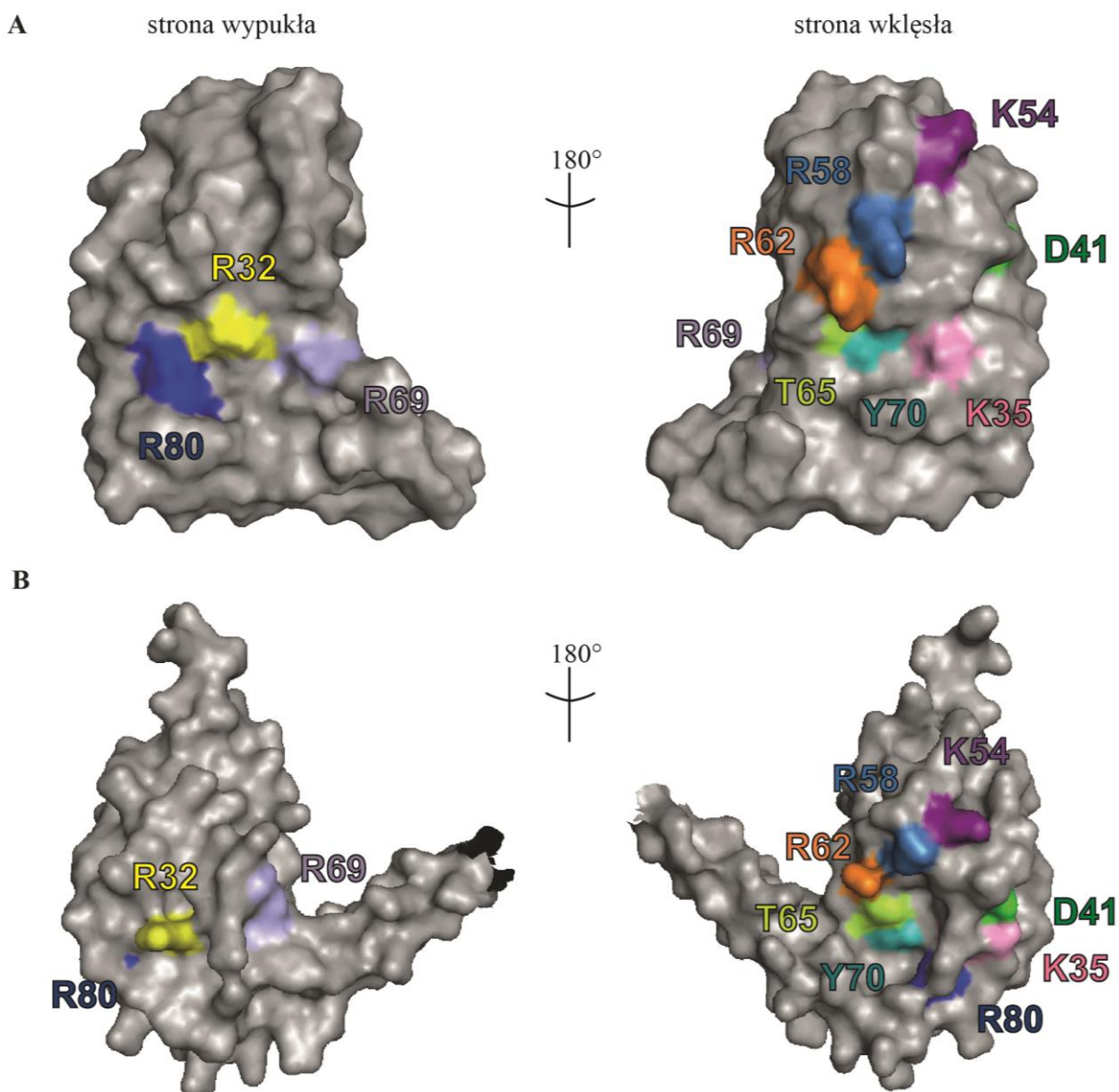
				R32	K35		D41			K54	R58	R62	T65																										
EcProQ	7	-LNSSKEVIAFLAERFP	HC	FSAEGEAR	PLK	I	GIFQDLVDRVA	--	GEMNLS	KTQLRS	ALR	LY	TSS	67																									
NmProQ	16	QTMRSKKKQTE	MIADHIY	GKYD	VFKR	FKPLALGIDQDLIAALPQ	----	YDAAL	IARV	L	NH	CRR	84																										
LpRocC	21	SKRARS	DALLWLAANF	PEAF	DNSLR	IRPLKIGIMSDILQHA	EKA	EQVGV	SKSL	REAV	VL	FTRR	84																										
EcFinO	69	NLP	TLDEAVNTL	KPWWPGLF	--	DG	TPRL	LACGIR	DDVLE	DVA	QR	NI	PLSH	KKLR	RALK	AITRS	130																						
SeFopA	80	TREY	THECV	EKIKAL	FPHLR	AEGGG	F	IPLK	I	GINND	ISAF	LA	EHPE	TEL	TM	DEWL	CAV	S	CITSR	143																			
lpp1663	8	TKKDKLQ	VIDW	LIEN	F	PN	AFF	FKGN	V	KPLK	I	GIF	DD	LID	F	YER	LD	TP	FS	SKSL	REAL	SY	SAS	72															
		.	:				*	**	.	:			..																										
				R69	Y70		R80																																
EcProQ	68	WRYLYG	-VKPGAT	RV	DL	DGNPC	GELDE	QHVE	HARK	QLE	EAKAR	VQ	QRAEQ	QAKKRE	AAA	TAGE	130																						
NmProQ	75	PRYLKA	-LAR	GGK	RFDL	NNR	FKGE	VTPEEQ	AIAQ	NHP	FV	----	QAL	QOQ	SAQ	-AAA	ETLSV	130																					
LpRocC	85	LDYLAC	-LKARE	VR	IDLH	GNP	VAEV	TEEE	AEN	SMK	IKK	RVE	KSV	KNARK	Q	VNAK	---AA	NHSY	144																				
EcFinO	131	ESYLCA	-MKAGAC	RYD	TEGY	VTEH	ISQEE	EVY	AAER	LDK	IRR	QNR	IKAE	LQ	AVL	-----	183																						
SeFopA	144	RVYLQ	RTAV	AGV	PRY	GLD	GHP	KGQ	VS	DEA	QSAG	R	LAT	LE	QK	WLR	MQ	AQ	Q	-ENIS	-----	197																	
lpp1663	73	PAYL	SC	-Q	KP	D	TAR	V	I	Y	GN	E	V	V	V	T	PE	Q	A	Y	A	Y	Q	R	Y	Q	E	R	Y	G	N	K	S	Q	D	L	K	-----	121
		**		*	.	.	:	..	*			:																											

Rysunek 20. Porównanie sekwencji aminokwasowych domen FinO sześciu białek należących do rodziny białek FinO o znanych strukturach i/lub ligandach.

U góry zaznaczono pozycje aminokwasowe w N-końcowej domenie białka ProQ wybrane do mutagenazy; EcProQ – N-końcowa domena białka ProQ z *E. coli*; NmProQ – domena FinO białka NMB1681 z *N. meningitidis*; LpRocC – domena FinO białka RocC z *L. pneumophila*; SeFopA – domena FinO białka FopA z *S. enterica*; lpp1663 – domena FinO białka Lpp1663 z *L. pneumophila*.

W zestawie pozycji wybranych do substytucji ze względu na silne zakonserwowanie ewolucyjne uwzględniłam tyrozynę 70 i argininę 80. Łańcuchy boczne tych aminokwasów mają potencjał do tworzenia wiązań wodorowych (wolna grupa hydroksylowa tyrozyny) lub oddziaływań elektrostatycznych (dodatnio naładowana arginina). Ponadto mutacje skutkujące substytucjami w tych pozycjach powodowały zahamowanie wiązania RNA *in vivo* (Pandey i wsp. 2020). Lizynę 35, asparaginian 41, lizynę 54 i argininę 58 wybrałam ze względu na znany efekt osłabienia wiązania *in vivo* (Pandey i wsp. 2020). W przypadku argininy 32 i treoniny 65 analiza filogenetyczna pokazała, że pozycje te wśród znanych białek z rodziny FinO są zwykle zastępowane przez aminokwasy o zbliżonych właściwościach (arginina 32 często była zastępowana lizyną, a treonina 65 seryną lub cysteiną), co może sugerować, że wybrane aminokwasy pełnią ważne funkcje. Argininy 62 i 69 zlokalizowane są w ewolucyjnie zmiennych pozycjach, ale stosunkowo blisko pozycji o znanym wpływie na wiązanie RNA. Ponadto oba aminokwasy niosą ładunek dodatni, co również może mieć znaczenie w powstawaniu oddziaływań elektrostatycznych między białkiem a ujemnie naładowanym kwasem nukleinowym.

Kodony kodujące aminokwasy w wybranych pozycjach zostały zmutowane do kodonu kodującego alaninę, oprócz tyrozyny, ponieważ poprzednie badania pokazały, że substytucja tyrozyny do fenyloalaniny wystarczyła do zaburzenia wiązania cząsteczek *cspE* i *SibA in vivo* (Pandey i wsp. 2020), i argininy 80, która została zastąpiona zarówno alaniną, jak i lizyną, aby sprawdzić, czy efekt utraty funkcji zaobserwowany *in vivo* jest związany z utratą ładunku dodatniego.



Rysunek 21. Pozycje aminokwasowe wybrane do mutagenезy zaprezentowane na strukturze N-końcowej domeny białka ProQ uzyskanej w (A) NMR ((Gonzalez i wsp. 2017); PDB: 5nb9) oraz (B) odpowiadający mu 119-nt fragment struktury uzyskany jako predykcja w programie AlphaFold.

Zgodnie ze strukturą N-końcowej domeny białka ProQ uzyskaną w spektroskopii masowej, (Gonzalez i wsp. 2017) łańcuchy boczne wszystkich wybranych aminokwasów znajdują się na powierzchni białka, dlatego nie spodziewamy się ich znaczenia dla struktury białka. Reszty R32, R69 i R80 są zlokalizowane po wypukłej (ang. *convex*) stronie białka, natomiast pozostałe reszty znajdują się po przeciwnej wklęsłej (ang. *concave*) stronie (Rys. 21).

Przygotowany zestaw aminokwasów zawiera reszty o znanym znaczeniu dla wiązania RNA *in vivo* lub położone w pobliżu takich pozycji. Niektóre z nich są silnie

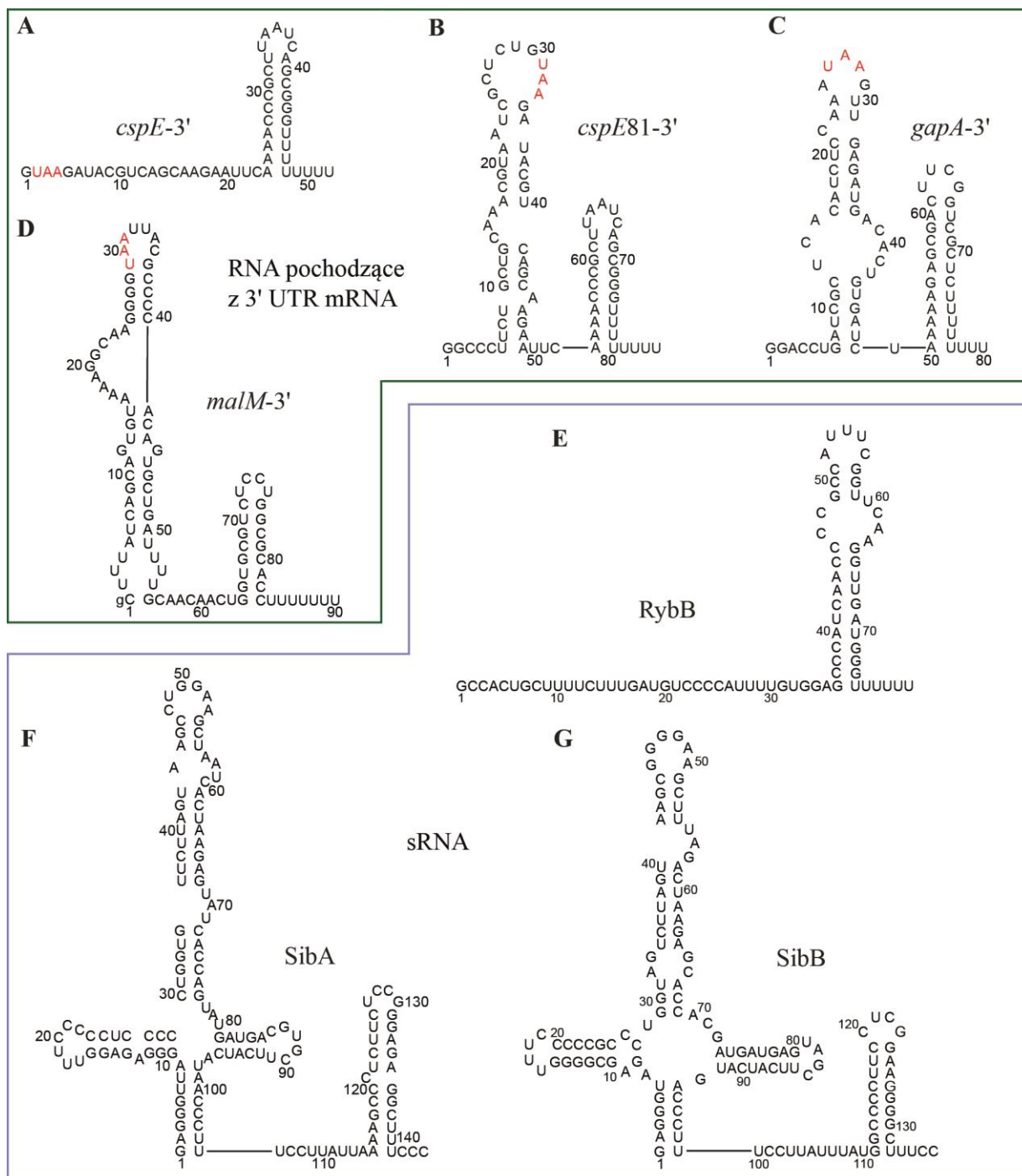
zakonserwowane ewolucyjnie, inne ułożone są w pozycjach zmiennych, zatem grupa ta zawiera pozycje o zróżnicowanych cechach, co może pomóc w identyfikacji reszt zaangażowanych w odmienne interakcje z różnymi cząsteczkami RNA.

5.2.2. Wyznaczenie siły oddziaływania mutantów N-końcowej domeny białka ProQ z cząsteczkami RNA

Aby porównać wpływ mutacji w obrębie N-końcowej domeny białka ProQ, zmierzyłam wiązanie 7 cząsteczek RNA zidentyfikowanych jako ligandy białka ProQ do 11 wyizolowanych mutantów białka. Wybrałam 4 cząsteczki RNA pochodzące z regionu 3'-UTR mRNA: *cspE-3'*, *cspE81-3'*, *gapA-3'* i *malM-3'* oraz 3 cząsteczki sRNA: RybB, SibA i SibB (Rys. 22), z których wszystkie zostały zidentyfikowane jako ligandy ProQ *in vivo* (Smirnov i wsp. 2016; Holmqvist i wsp. 2018; Melamed i wsp. 2020).

Spośród wszystkich badanych RNA cząsteczka *cspE-3'* była wiązana najsilniej przez NTD, wiązanie opisywała wartość K_d $4,2 \pm 1,3$ nM (Rys. 23, Tabela 5). Mutacje w pozycjach R32, K35, D41, T65 i R69 osłabiły wiązanie do *cspE-3'* 2-4-krotnie. Podobny poziom podwyższenia K_d obserwowałam dla substytucji K54A i R62A, z tym że dla tych białek odpowiednio o ok. 30 % i 50% obniżona była także średnia maksymalna frakcja związana. Dla wiązania z białkiem Y70F w maksymalnych stężeniach obserwowałam kompleks z *cspE-3'*, lecz w bardzo małej ilości nieprzekraczającej 30%, dlatego K_d dla tej reakcji określiłam na > 1000 nM. W przypadku białek R58A, R80A i R80K nie obserwowałam kompleksu z RNA *cspE-3'* w zakresie stężeń do 1 μ M.

Cząsteczka *cspE81-3'* była wiązana przez NTD z K_d o wartości $8,3 \pm 2,5$ nM (Rys. 24, Tabela 5). Substytucje R32A i T65A nie miały wpływu na wiązanie tego RNA, substytucje K35A, D41A i R69A osłabiały wiązanie *cspE81-3'* ponad dwukrotnie. Podobną wartość K_d obserwowałam dla substytucji K54A, jednak dla tej reakcji obniżona niemal 3-krotnie była także średnia maksymalna frakcja związana (do zaledwie 32%). W przypadku substytucji R62A i Y70F obserwowałam kompleks z *cspE81-3'* w najwyższych stężeniach białka, jednak w minimalnej ilości, dla białek R58A, R80A i R80K kompleks z RNA *cspE81-3'* w zakresie stężeń do 1 μ M nie powstawał.



Rysunek 22. Częsteczki RNA, których wiązanie do mutantów N-końcowej domeny białka ProQ zostało zmierzone w pracy.

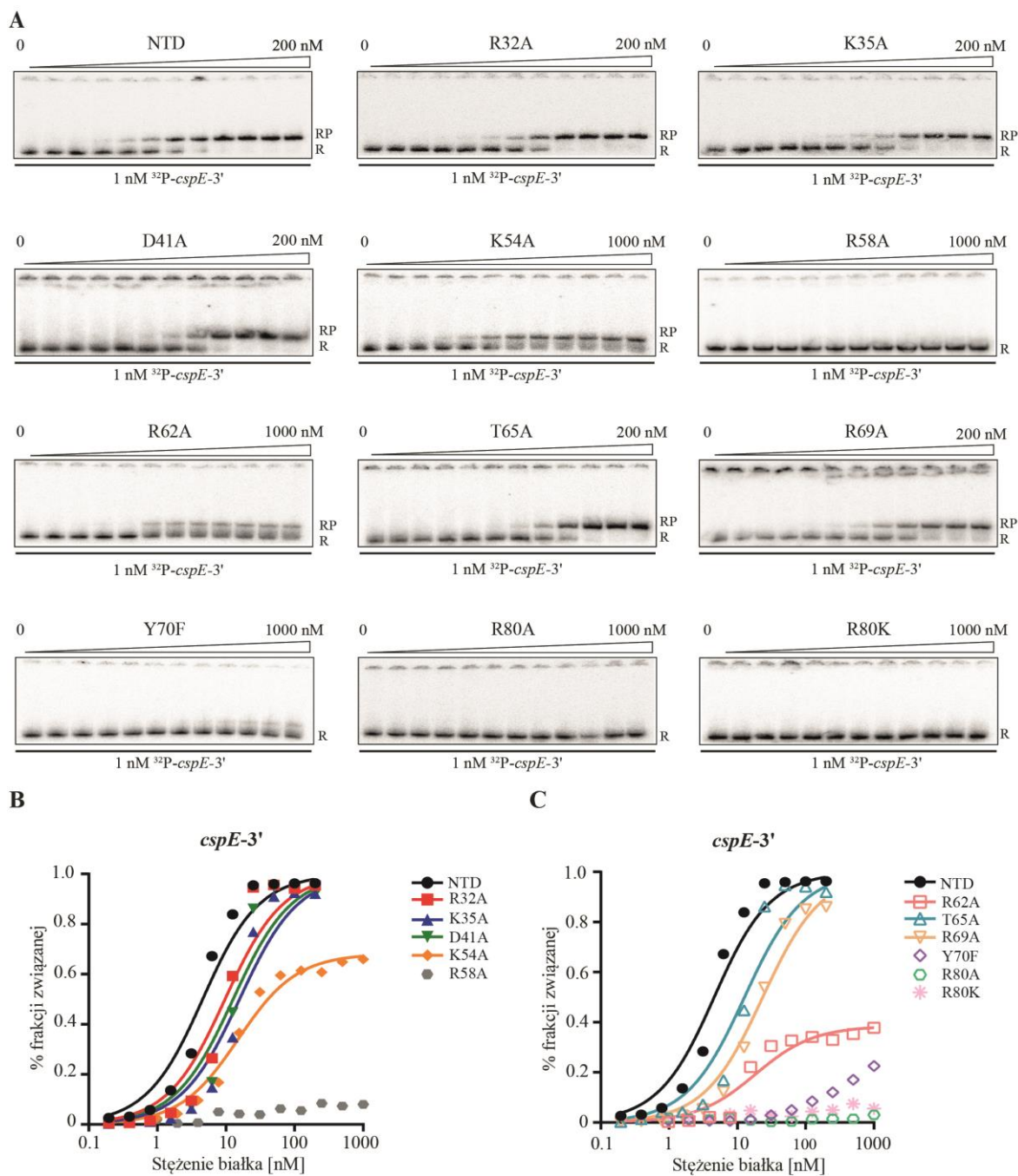
Struktury drugorzędowe cząsteczek sRNA (A) *cspE-3'*, (B) *cspE81-3'*, (C) *gapA-3'*, (D) *malM-3'* i (G) SibB zostały przygotowane przy użyciu oprogramowania RNAstructure. Drugorzędowe struktury cząsteczek (E) RybB (Balbontín i wsp. 2010) i (F) SibA (Smirnov i wsp. 2016) zaprezentowano zgodnie z referencjami. Kodon STOP (UAA) oznaczono na czerwono, małą literą „g” oznaczono guanozynę dodaną na początku sekwencji cząsteczki *malM-3'* w celu umożliwienia jej transkrypcji *in vitro* przez polimerazę z faga T7.

Tabela 5. Wiązanie cząsteczek RNA do mutantów N-końcowej domeny białka ProQ.

Mutacja	K_d [nM] (% średniej max. frakcji związanej)						
	<i>cspE-3'</i>	<i>cspE81-3'</i>	<i>gapA-3'</i>	<i>malM-3'</i>	RybB	SibA	SibB
wt NTD	4,2 ± 1,3 (96 %)	8,3 ± 2,5 (87 %)	13 ± 3 (88 %)	8,2 ± 2,8 (85 %)	26 ± 5,2 (77 %)	7,6 ± 2,6 (84 %)	9,7 ± 3,4 (74 %)
R32A	9,3 ± 0,47 (94 %)	12 ± 1,5 (83 %)	18 ± 1,3 (87 %)	13 ± 3,4 (83 %)	73 ± 24 (64 %)	21 ± 5,7 (72 %)	30 ± 7,3 (46 %)
K35A	18 ± 3,9 (88 %)	22 ± 2,6 (76 %)	45 ± 4,4 (55 %)	34 ± 15 (67 %)	78 ± 44 (37 %)	24 ± 6,5 (52 %)	n. o.
D41A	13 ± 1,7 (95 %)	19 ± 5,5 (91 %)	20 ± 2,8 (94 %)	26 ± 7,2 (87 %)	42 ± 14 (85 %)	43 ± 15 (84 %)	51 ± 14 (75 %)
K54A	18 ± 8 (59 %)	18 ± 4,5 (32 %)	>1000 (15 %)	27 ± 6,9 (33 %)	>1000 (17 %)	n. o.	n. o.
K58A	n. o.	n. o.	n. o.	n. o.	n. o.	n. o.	n. o.
R62A	19 ± 3,8 (46 %)	>1000 (17 %)	>1000 (10 %)	>1000 (25 %)	n. o.	>1000 (26 %)	n. o.
T65A	15 ± 2,1 (94 %)	15 ± 6 (86 %)	18 ± 8 (93 %)	21 ± 9,9 (90 %)	31 ± 7,4 (87 %)	16 ± 3,1 (81 %)	41 ± 7,4 (62 %)
R69A	20 ± 2,9 (88 %)	28 ± 7,5 (72 %)	140 ± 18 (44 %)	120 ± 41 (64 %)	130 ± 48 (31 %)	61 ± 18 (41 %)	n. o.
Y70F	>1000 (19 %)	>1000 (15 %)	n. o.	>1000 (16 %)	n. o.	n. o.	n. o.
R80A	n. o.	n. o.	n. o.	n. o.	n. o.	n. o.	n. o.
R80K	n. o.	n. o.	n. o.	n. o.	n. o.	n. o.	n. o.

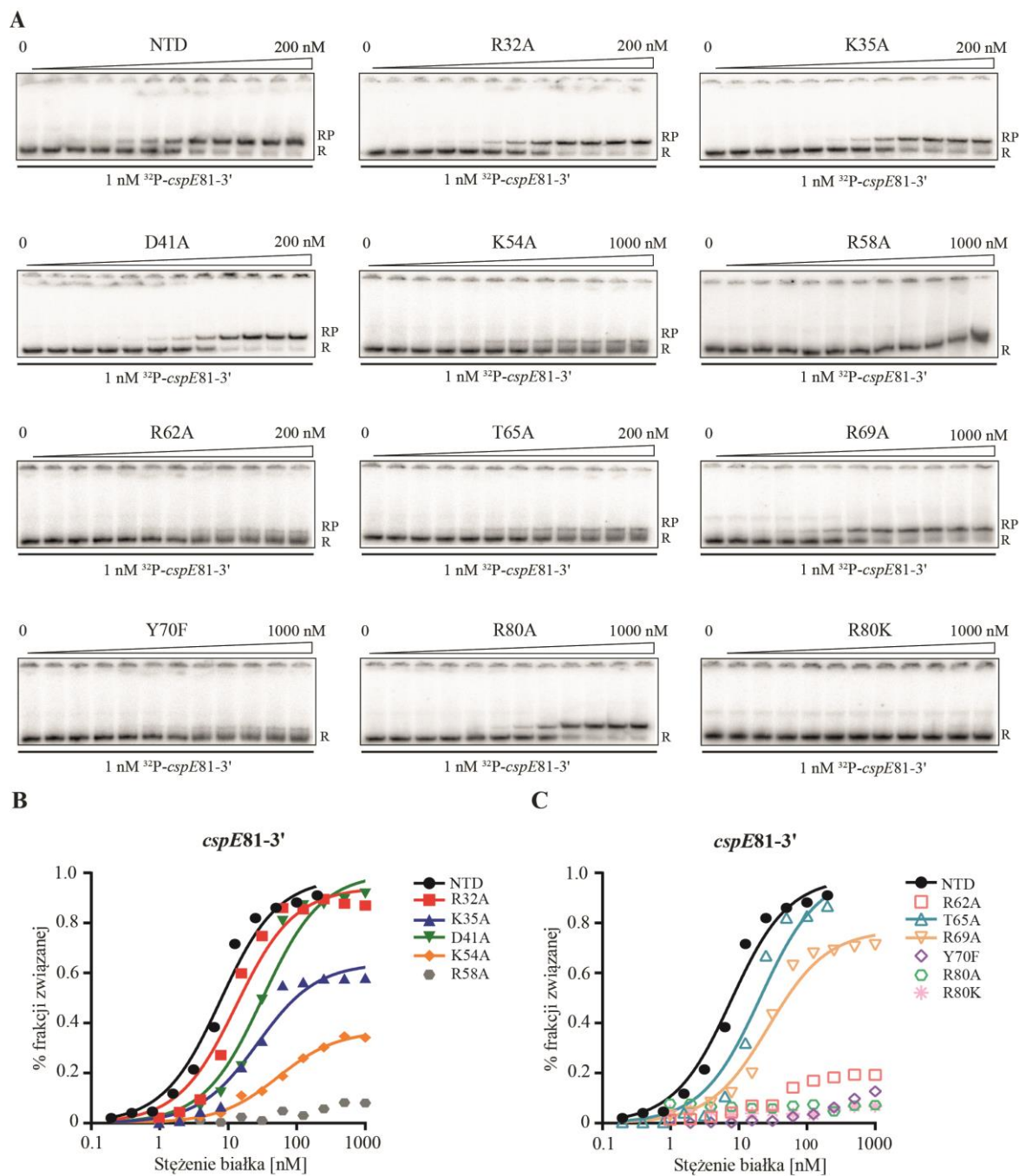
Wartości K_d uzyskano przez dopasowanie danych wykresu zależności frakcji związanego RNA od stężenia białka do równania kwadratowego w programie GraphPad. Średnie wartości K_d wraz z odchyleniem standardowym obliczono na podstawie co najmniej 3 powtórzeń z niezależnych eksperymentów.

n.o. – nie obserwowano wiązania w zakresie stężeń do 1 μ M.



Rysunek 23. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA *cspE-3'*.

(A) Wiązanie 1 nM ^{32}P -*cspE-3'* do wszystkich białek monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia białek (B) NTD, K35A, D41A, K54A i R58A oraz (C) R62A, T65A, R69A, Y70F, R80A i R80K zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 5.



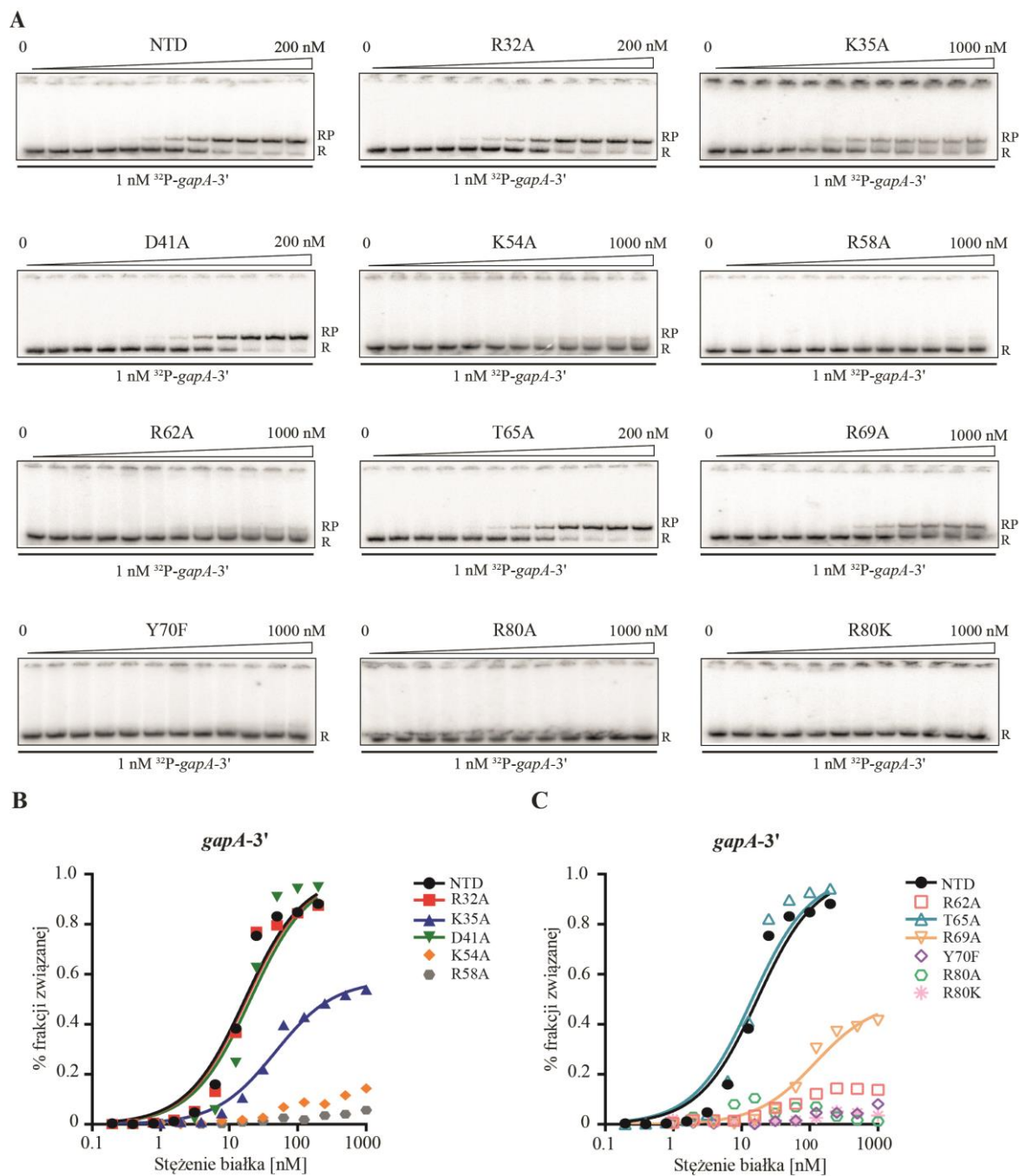
Rysunek 24. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA *cspE81-3'*.

(A) Wiązanie 1 nM ^{32}P -*cspE81-3'* do wszystkich białek monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia białek (B) NTD, K35A, D41A, K54A i R58A oraz (C) R62A, T65A, R69A, Y70F, R80A i R80K zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 5.

Wartość K_d dla wiązania cząsteczki *gapA-3'* z NTD wynosiła 13 ± 3 nM (Rys. 25, Tabela 5). Bez wpływu na wiązanie tego RNA pozostały substytucje R32A, D41A i T65A, natomiast substytucje K35A i R69A osłabiły wiązanie cząsteczki *gapA-3'* odpowiednio 3,5- i ponad 9-krotnie. Wraz z niższym K_d obserwowałam także znaczne obniżenie średniej maksymalnej frakcji związanej o 30-40%. Kompleks białko-RNA w najwyższych stężeniach widoczny był dla reakcji z białkami K54A i R62A, jednak średnia maksymalna frakcja związana nie przekraczała 15%. Dla białek R58A, Y70F, R80A i R80K nie obserwowałam kompleksu z RNA *gapA-3'* w zakresie stężeń do 1 μ M.

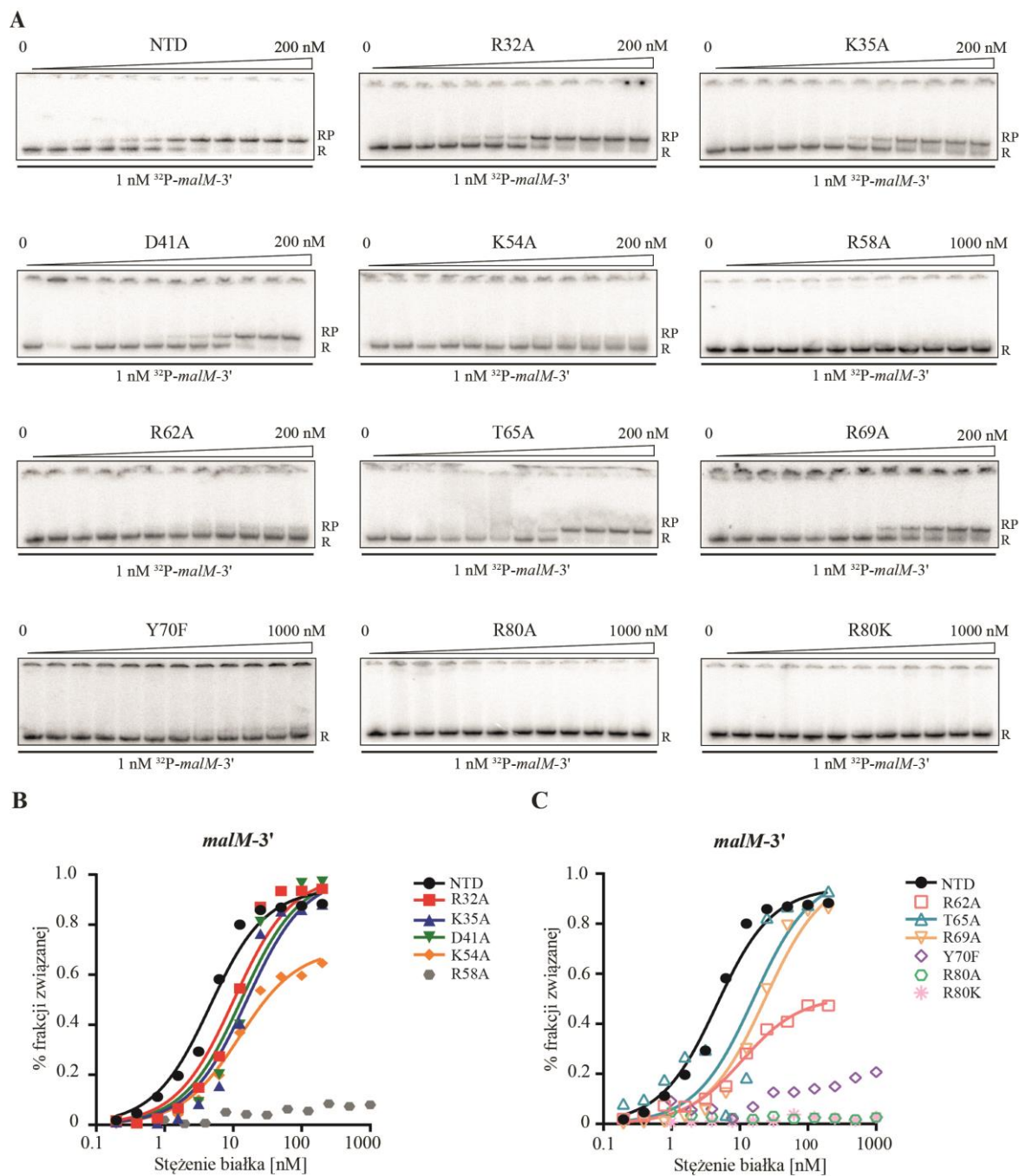
N-końcowa domena białka ProQ wiązała cząsteczkę *malM-3'* z K_d o wartości $8,2 \pm 2,8$ nM (Rys. 26, Tabela 5). Substytucja R32A nie miała wpływu na wiązanie tego RNA, substytucje D41A, i T65A osłabiały wiązanie 3-krotnie. 4-5-krotne obniżenie wartości K_d zanotowałam dla wiązania *malM-3'* z białkami K35A i R69A, jednak dla tych reakcji obserwowałam również spadek o ok. 20% średniej maksymalnej frakcji związanej, natomiast w przypadku wariantu K54A wartość K_d wskazywała na 3-krotne osłabienie wiązania z *malM-3'*, jednak średnia maksymalna frakcja związana wyniosła zaledwie 33%. Kompleks *malM-3'* w najwyższych stężeniach widoczny był dla reakcji z wariantami K54A i R62A, jednak średnia maksymalna frakcja związana nie przekraczała 25%, dlatego ich K_d określiłam jako > 1000 nM. Białka R58A, R80A i R80K nie wiązały się z cząsteczką *malM-3'* w zakresie stężeń do 1 μ M.

sRNA RybB wiązało się najslabiej do N-końcowej domeny białka ProQ spośród badanych cząsteczek, z K_d o wartości $26 \pm 5,2$ nM (Rys. 27, Tabela 5). Substytucje D41A i T65A pozostały bez wpływu na wiązanie tego RNA, substytucja R32A 3-krotnie osłabiła wiązanie do RybB, a substytucje K35A i R69A osłabiały wiązanie 3-5-krotnie przy jednoczesnym obniżeniu średniej maksymalnej frakcji związanej o 40%. W przypadku wariantu K54A obserwowałam kompleks z RNA jedynie w najwyższych stężeniach, a średnia maksymalna frakcja związana była niższa niż 20%. Dla białek R58A, R62A, Y70F, R80A i R80K nie obserwowałam kompleksu z sRNA RybB w zakresie stężeń do 1 μ M.



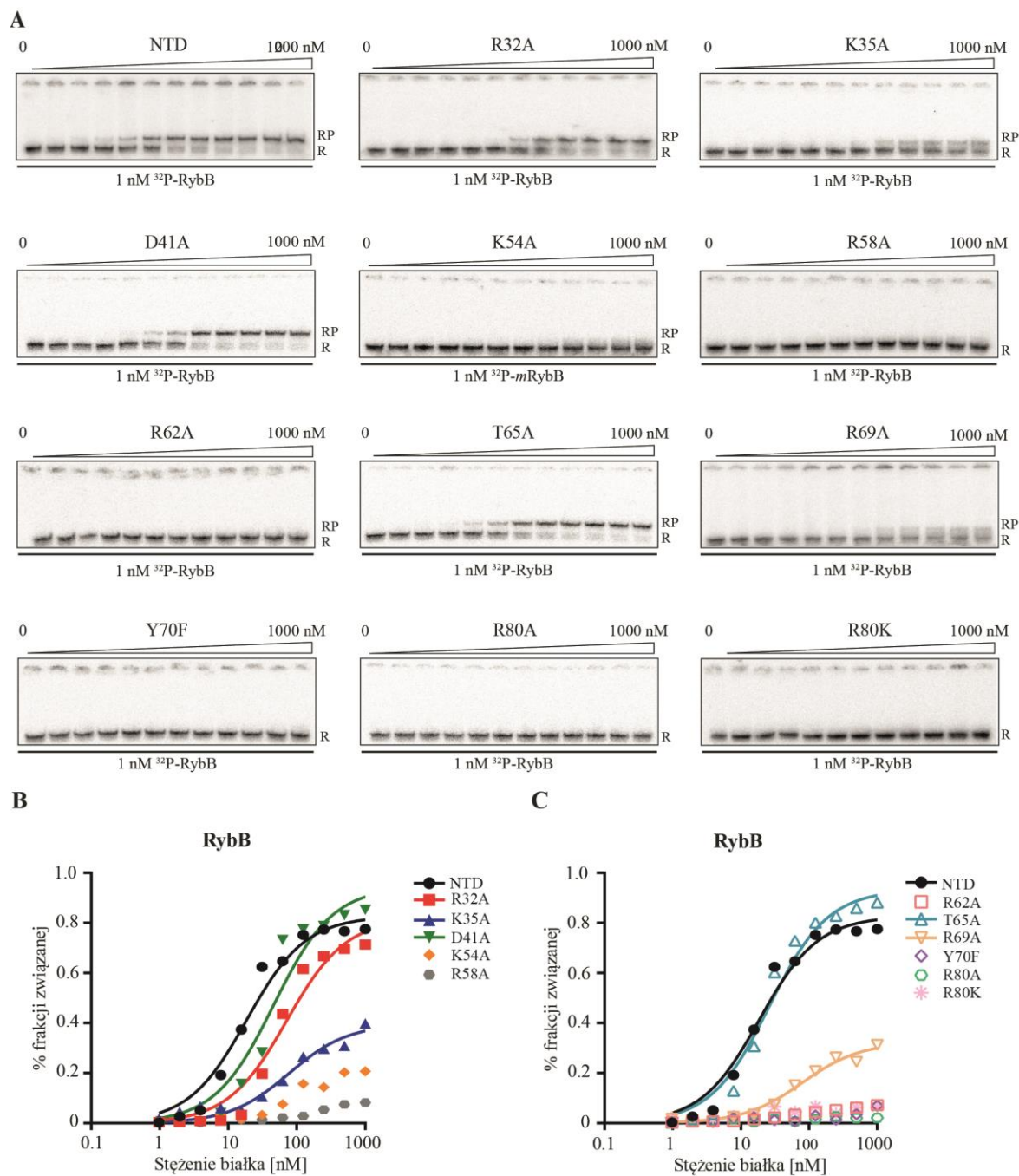
Rysunek 25. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA *gapA-3'*.

(A) Wiązanie 1 nM ^{32}P -*gapA-3'* do wszystkich białek monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia białek (B) NTD, K35A, D41A, K54A i R58A oraz (C) R62A, T65A, R69A, Y70F, R80A i R80K zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 5.



Rysunek 26. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA *malM-3'*.

(A) Wiązanie 1 nM ^{32}P -*malM-3'* do wszystkich białek monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia białek (B) NTD, K35A, D41A, K54A i R58A oraz (C) R62A, T65A, R69A, Y70F, R80A i R80K zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 5.



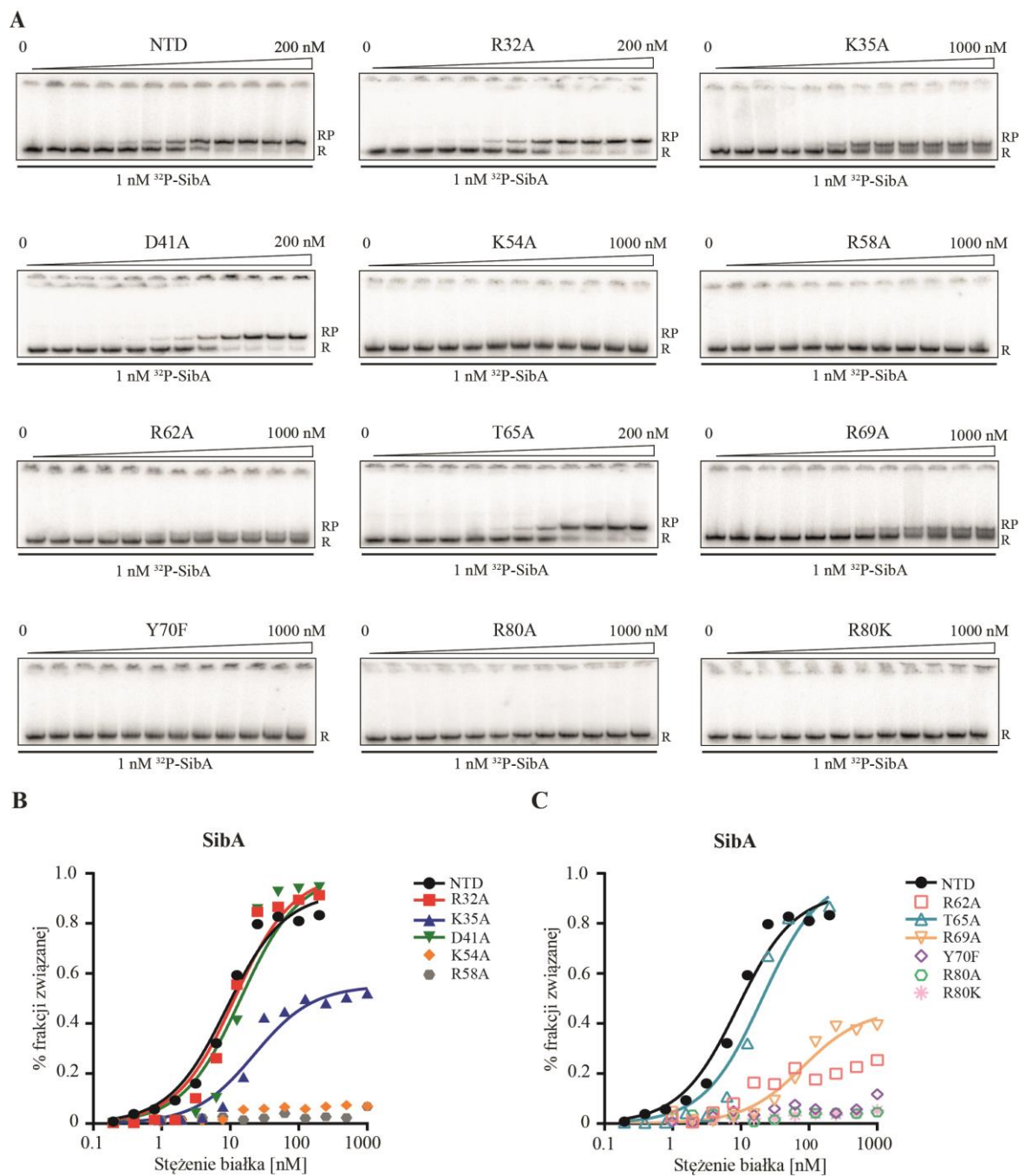
Rysunek 27. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA RybB.

(A) Wiązanie 1 nM ^{32}P -RybB do wszystkich białek monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia białek (B) NTD, K35A, D41A, K54A i R58A oraz (C) R62A, T65A, R69A, Y70F, R80A i R80K zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 5.

Kolejna z badanych cząsteczek, sRNA SibA, wiązała się do NTD z K_d o wartości $7,6 \pm 2,6$ nM (Rys. 28, Tabela 5). Mutacje w pozycjach R32A, D41A i T65A osłabiły wiązanie z SibA 2-5-krotnie, podobnie jak substytucje K35A i R69A przy jednoczesnym obniżeniu średniej maksymalnej frakcji związanej o 30-40%. W przypadku substytucji R62A zaobserwowano kompleks z sRNA SibA jedynie w najwyższych stężeniach przy średniej maksymalnej frakcji związanej 26%, natomiast białka K54A, R58A, Y70F, R80A i R80K nie wiązały się z sRNA SibA w zakresie stężeń do $1\mu\text{M}$

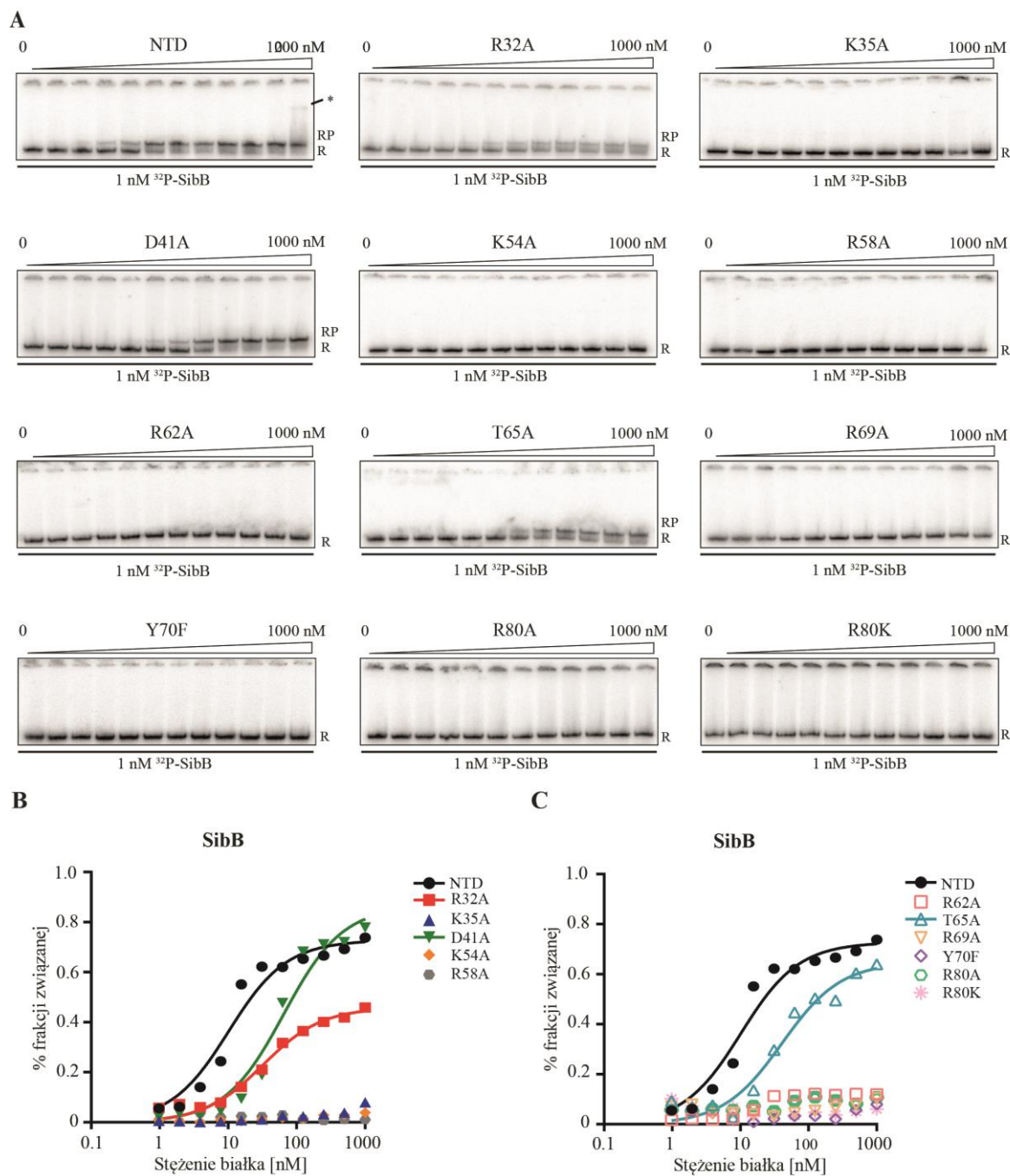
Ostatnią cząsteczką RNA, dla której zbadałam wiązanie do wszystkich zmutowanych wersji N-końcowej domeny ProQ, było sRNA SibB, które wiązało się z NTD z K_d o wartości $9,7 \pm 3,4$ nM (Rys. 29, Tabela 5). Punktowe mutacje wprowadzone na powierzchni białka najbardziej wpłynęły na wiązanie do tej cząsteczki, ponieważ aż w 8 przypadkach, dla białek K35A, K54A, R58A, R62A, R69A, Y70F, R80A i R80K nie obserwowałam kompleksu z SibB w zakresie stężeń do $1\mu\text{M}$. Kompleks białko-RNA powstał w reakcjach dla wiązania SibB z wariantami R32A, D41A i T65A przy K_d obniżonym 3-5-krotnie w stosunku do wariantu o naturalnej sekwencji; podczas gdy dla reakcji z wariantem D41A nie obserwowałam zmiany w ilości powstającego kompleksu, dla wiązania z białkiem T65A zanotowałam 10% spadek średniej maksymalnej frakcji związanej, a dla reakcji z wariantem R32A spadek tej wartości wynosił 30%.

W podsumowaniu wyników badań wiązania mutantów N-końcowej domeny białka ProQ można zauważyć, że substytucje R58A, R80A i R80K praktycznie unieczynniły białko, ponieważ mutanty nie wiązały żadnego z badanych RNA. Mutacje w pozycjach D41 i T65 nie miały wpływu na wiązanie lub je nieznacznie osłabiły, podobnie substytucja R32A. Zmiana aminokwasu w pozycjach K54, R62 i Y70 znacznie osłabiła wiązanie białka do RNA, jednak w kilku przypadkach kompleks RNA-białko był widoczny w najwyższych stężeniach białka. W kontekście dalszych eksperymentów najciekawsze wydały mi się substytucje K35 i R69, ponieważ dla ich wiązania z cząsteczkami RNA obserwowałam cały wachlarz efektów: od lekkiego osłabienia dla wiązania z *cspE*-3' po zupełny zanik wiązania z SibB, co sugeruje, że aminokwasy te mogą być zaangażowane w bezpośrednie oddziaływanie z RNA, a same cząsteczki RNA mogą w różny sposób wiązać się do N-końcowej domeny białka ProQ.



Rysunek 28. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA SibA.

(A) Wiązanie 1 nM ^{32}P -SibA do wszystkich białek monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia białek (B) NTD, K35A, D41A, K54A i R58A oraz (C) R62A, T65A, R69A, Y70F, R80A i R80K zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 5.



Rysunek 29. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA SibB.

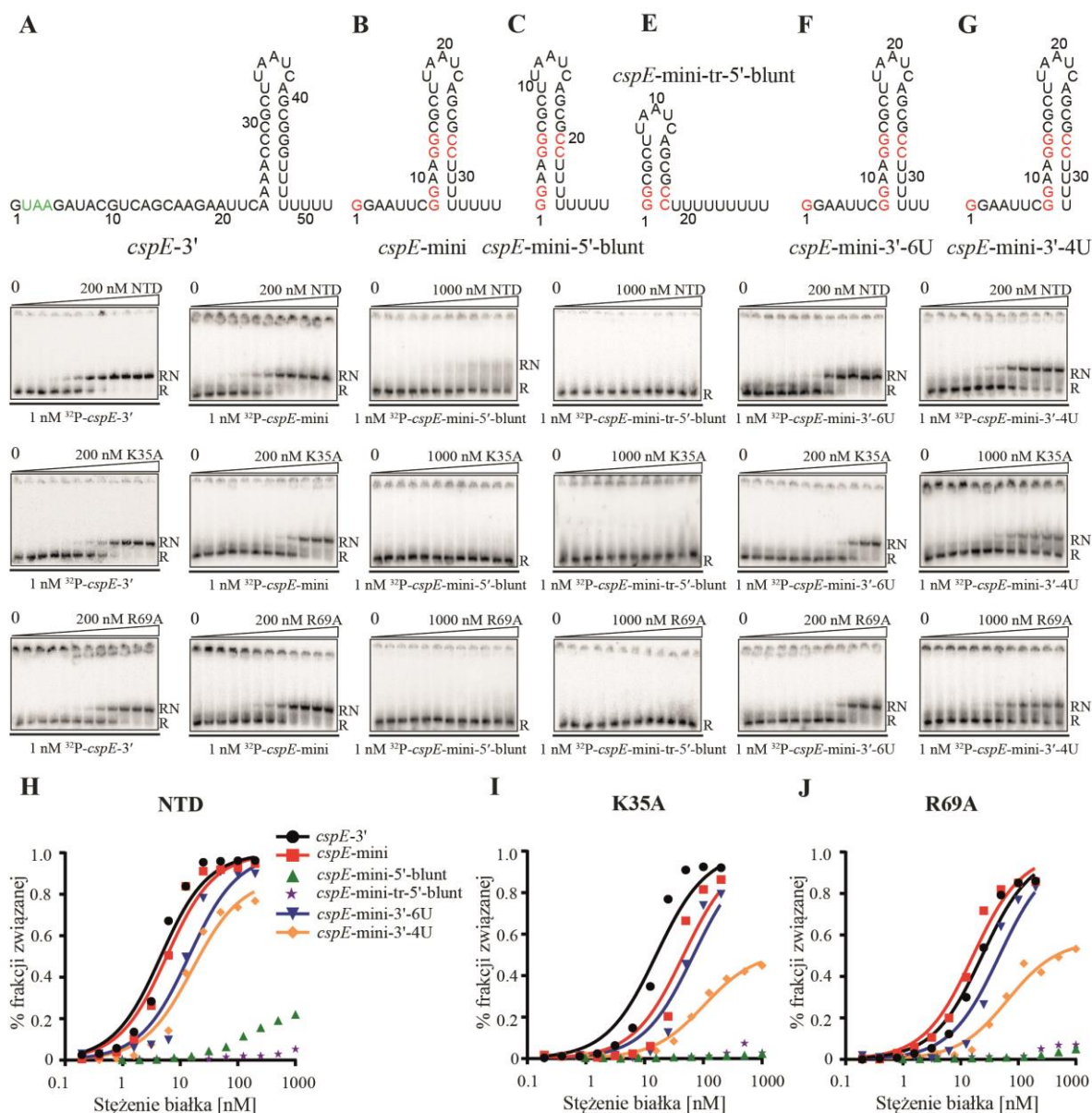
(A) Wiązanie 1 nM ^{32}P -SibB do wszystkich białek monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia białek (B) NTD, K35A, D41A, K54A i R58A oraz (C) R62A, T65A, R69A, Y70F, R80A i R80K zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 5.

5.2.3. Cechy cząsteczek RNA potencjalnie wpływające na odmienne sposoby wiązania do mutantów K35A i R69A domeny FinO białka ProQ

5.2.3.1. Znaczenie sekwencji zlokalizowanych po 5' i 3' stronie spinki terminatorowej cząsteczek RNA w różnicowym wiązaniu RNA przez mutanty N-końcowej domeny białka ProQ

W poprzedniej fazie badań zaobserwowałam, że sekwencje okalające trzon spinki terminatorowej mają istotny wpływ na rozpoznanie tej struktury przez N-kończącą domenę białka ProQ. W kontekście punktowych substytucji aminokwasowych ciekawym wydało mi się sprawdzenie, czy cechy, które wpływają na zachowanie białka zmieniają się pod wpływem mutacji. W celu sprawdzenia, czy na zróżnicowanie wiązania do mutantów o substytucjach K35A i R69A wpływają sekwencje zlokalizowane po 5' czy po 3' stronie trzonu struktury spinki terminatorowej, zaprojektowałam szereg skróconych wariantów cząsteczki *cspE-3'*. Zmierzyłam wiązanie wariantów: *cspE-3'*-mini o skróconym jednoniciowym fragmencie zlokalizowanym po 5' stronie trzonu spinki terminatorowej (z 23 do 7 nt), jego warianty ze skróconym o 2 lub 4 U ogonem poliU (*cspE-3'*-mini_6U i *cspE-3'*-mini_4U), oraz cząsteczki *cspE*-mini-5'-blunt oraz *cspE*-mini_tr-5'-blunt pozbawione sekwencji jednoniciowej na 5'-końcu (Rys. 30).

Spośród badanych RNA do NTD z wysokim powinowactwem wiąże się cząsteczka *cspE-3'*-mini oraz jej warianty pozbawione 2 i 4 urydyn na 3'-końcu (Rys. 30, Tabela 6). Wiązanie cząsteczek *cspE*-mini-5'-blunt oraz *cspE*-mini_tr-5'-blunt pozbawionych jednoniciowego fragmentu po 5'-stronie trzonu spinki w zakresie stężeń do 1 μ M białka jest w zasadzie niewidoczne. W przypadku wiązania badanych cząsteczek do białek o substytucjach K35A i R69A nie widać znaczących różnic w porównaniu wiązania *cspE-3'* z jej minimalną wersją. Skrócenie ogona poliU o 2 nt nie wpływa na siłę oddziaływania, ponowne skrócenie tej sekwencji o 2 nt, przy 4 pozostających urydynach, widocznie osłabia oddziaływanie, reakcja nie ulega wysyceniu, a średnia frakcja związana spada do ok. 50 % przy wartości ok. 80% dla białka NTD. Zaobserwowałam, że pozbawienie cząsteczki RNA jednoniciowego fragmentu po 5'-stronie trzonu spinki wpływa na zanik wiązania zarówno do zmutowanych, jak i naturalnego wariantu białka NTD.



Rysunek 30. Zarówno ogon poliurydynowy zlokalizowany na 3'-końcu cząsteczki RNA, jak i jednoniciowa sekwencja ulokowana po 5'-stronie trzonu spinki terminatorowej wpływają na wiązanie do N-końcowej domeny białka ProQ.

Struktury drugorzędowe (A) *cspE*-3', (B) *cspE*-mini, (C) *cspE*-mini-5'-blunt, (D) *cspE*-mini-tr-5'-blunt, (E) *cspE*-mini-3'-6U i (F) *cspE*-mini-3'-4U zostały przygotowane przy użyciu oprogramowania RNAstructure. Dane zależności frakcji związanej RNA od stężenia białek (G) NTD, (H) K35A i (I) R69A zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 6. Oznaczenia kompleksów są takie same jak na Rys. 11. Kolorem **czerym** oznaczono zmutowane nukleotydy.

Tabela 6. Wiązanie modelowych cząsteczek pochodzących z *cspE-3'* do N-końcowej domeny białka ProQ i mutantów K35A i R69A.

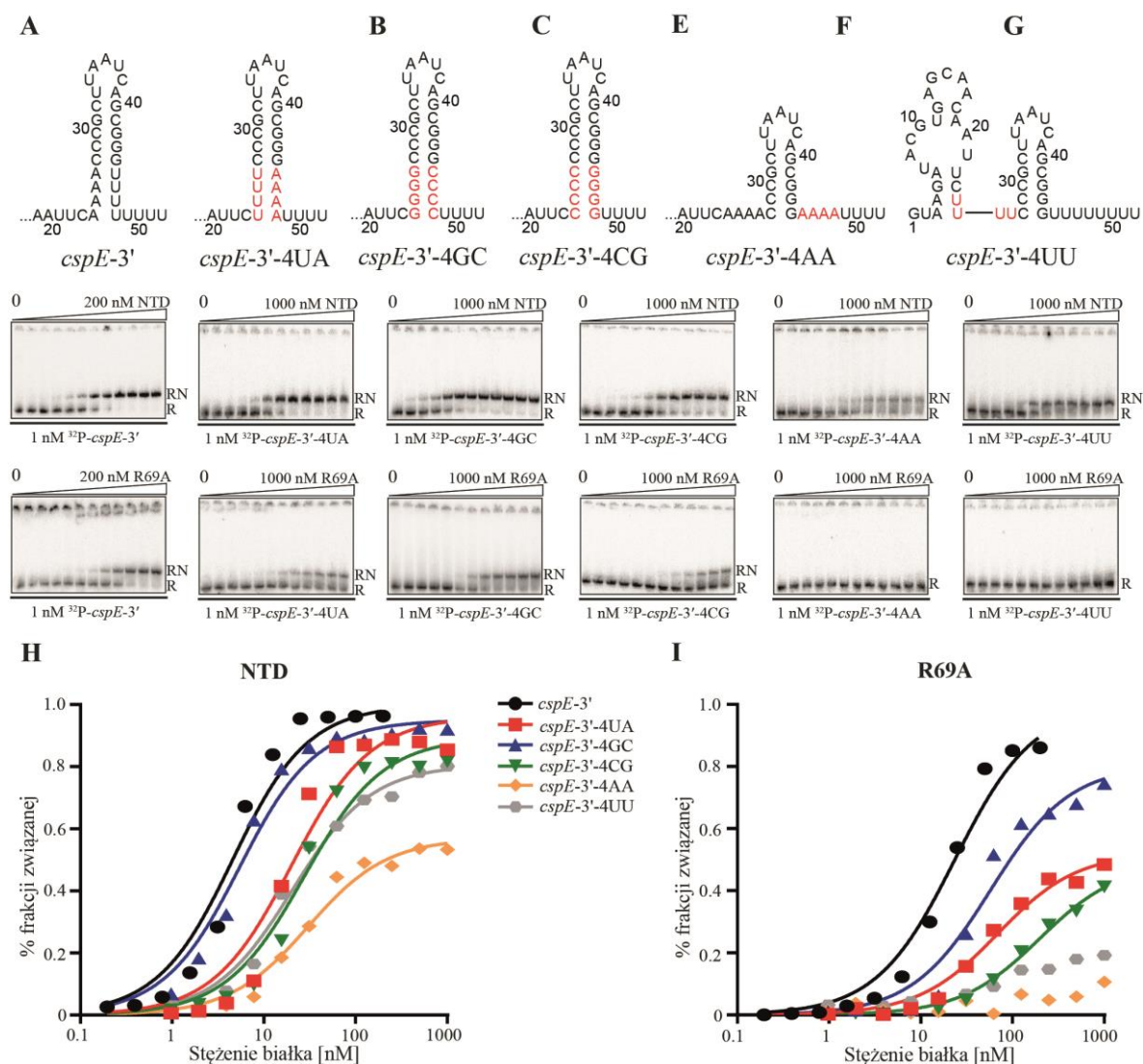
³² P-RNA	K_d [nM] (% średniej max. frakcji związanej)		
	wt NTD	K35A	R69A
<i>cspE-3'</i>	4,2 ± 1,3 (96 %)	18 ± 3,9 (88 %)	20 ± 2,9 (88 %)
<i>cspE-mini</i>	5,1 ± 1,1 (91 %)	44 ± 11 (86 %)	20 ± 7,7 (87 %)
<i>cspE-mini- 5'-blunt</i>	>1000 (25 %)	n. o.	n. o.
<i>cspE-mini_tr- 5'-blunt</i>	n. o.	n. o.	n. o.
<i>cspE-mini- 3'-6U</i>	13 ± 2,7 (91 %)	63 ± 8,4 (82 %)	46 ± 8,4 (80 %)
<i>cspE-mini- 3'-4U</i>	16 ± 1,2 (81 %)	75 ± 38 (41%)	99 ± 35 (56 %)

Wartości K_d uzyskano przez dopasowanie danych wykresu zależności frakcji związanego RNA od stężenia białka do równania kwadratowego w programie GraphPad. Średnie wartości K_d wraz z odchyleniem standardowym obliczono na podstawie co najmniej 3 powtórzeń z niezależnych eksperymentów.

n.o. – nie obserwowano wiązania w zakresie stężeń do 1 μ M.

5.2.3.2. Udział dolnej części spinki terminatorowej w wiązaniu mutantów domeny FinO białka ProQ

W celu poznania, czy struktura dolnej części trzonu spinki terminatorowej uczestniczy w wiązaniu do N-końcowej domeny białka ProQ, wprowadziłam mutacje do dwuniciowej części regionu bogatego w adenozyne i ogona poliU (Rys. 31). W pierwszym kroku 4 pary AU zamieniłam w 4 pary UA. Co ciekawe mutacja, która nie zmieniła stabilności termodynamicznej cząsteczki wpłynęła na obniżenie siły wiązania białka NTD, wartość K_d dla wiązania z cząsteczką UA spadła trzykrotnie w porównaniu do wiązania z cząsteczką o naturalnej sekwencji (Rys. 31, Tabela 7). Kolejnym krokiem było zamienienie par AU na pary GC, co nie wpłynęło na wiązanie do NTD, natomiast zamienienie kolejności wprowadzanych nukleotydów na CG osłabiło oddziaływanie. Te obserwacje mogą wskazywać na preferencję rozpoznawania przez białko w tych pozycjach nukleotydów. Jako reakcje kontrolne zastosowałam cząsteczki pozbawione sparowania w tym regionie, wersję z urydynami zamienionymi w adenozyne oraz RNA z adenozynami zmienionymi w urydyny. Oba warianty wykazują obniżone wiązanie do NTD.



Rysunek 31. Zaburzenie struktury dolnej części spinki terminatorowej RNA *cspE-3'* powoduje zwiększone osłabienie wiązania do mutanta R69A w porównaniu do wiązania z NTD.

Struktury drugorzędowe (A) *cspE-3'*, (B) *cspE-3'-4UA*, (C) *cspE-3'-4GC*, (D) *cspE-3'-4CG*, (E) *cspE-3'-4AA* i (F) *cspE-3'-4UU* zostały przygotowane przy użyciu oprogramowania RNAstructure. Dane zależności frakcji związanej RNA od stężenia do białek (G) NTD i (H) R69A zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 7. Oznaczenia kompleksów są takie same jak na Rys. 11. Kolorem czerwonym oznaczono zmutowane nukleotydy.

Tabela 7. Wiązanie modelowych cząsteczek pochodzących z *cspE-3'* ze zmienioną strukturą dolnej części trzonu spinki terminatorowej do N-końcowej domeny białka ProQ i mutantu R69A.

³² P-RNA	K_d [nM] (% średniej max. frakcji związanej)	
	wt NTD	R69A
<i>cspE-3'</i>	4,2 ± 1,3 (96 %)	20 ± 2,9 (88 %)
<i>cspE-3'-4UA</i>	15 ± 4,9 (88 %)	73 ± 14 (49 %)
<i>cspE-3'-4GC</i>	6,5 ± 1,7 (93 %)	43 ± 16 (72 %)
<i>cspE-3'-4CG</i>	27 ± 5,4 (82 %)	180 ± 30 (47 %)
<i>cspE-3'-4AA</i>	30 ± 3,3 (52 %)	n. o.
<i>cspE-3'-4UU</i>	19 ± 1,1 (68 %)	>1000 (22 %)

Wartości K_d uzyskano przez dopasowanie danych wykresu zależności frakcji związanego RNA od stężenia białka do równania kwadratowego w programie GraphPad. Średnie wartości K_d wraz z odchyleniem standardowym obliczono na podstawie co najmniej 3 powtórzeń z niezależnych eksperymentów.

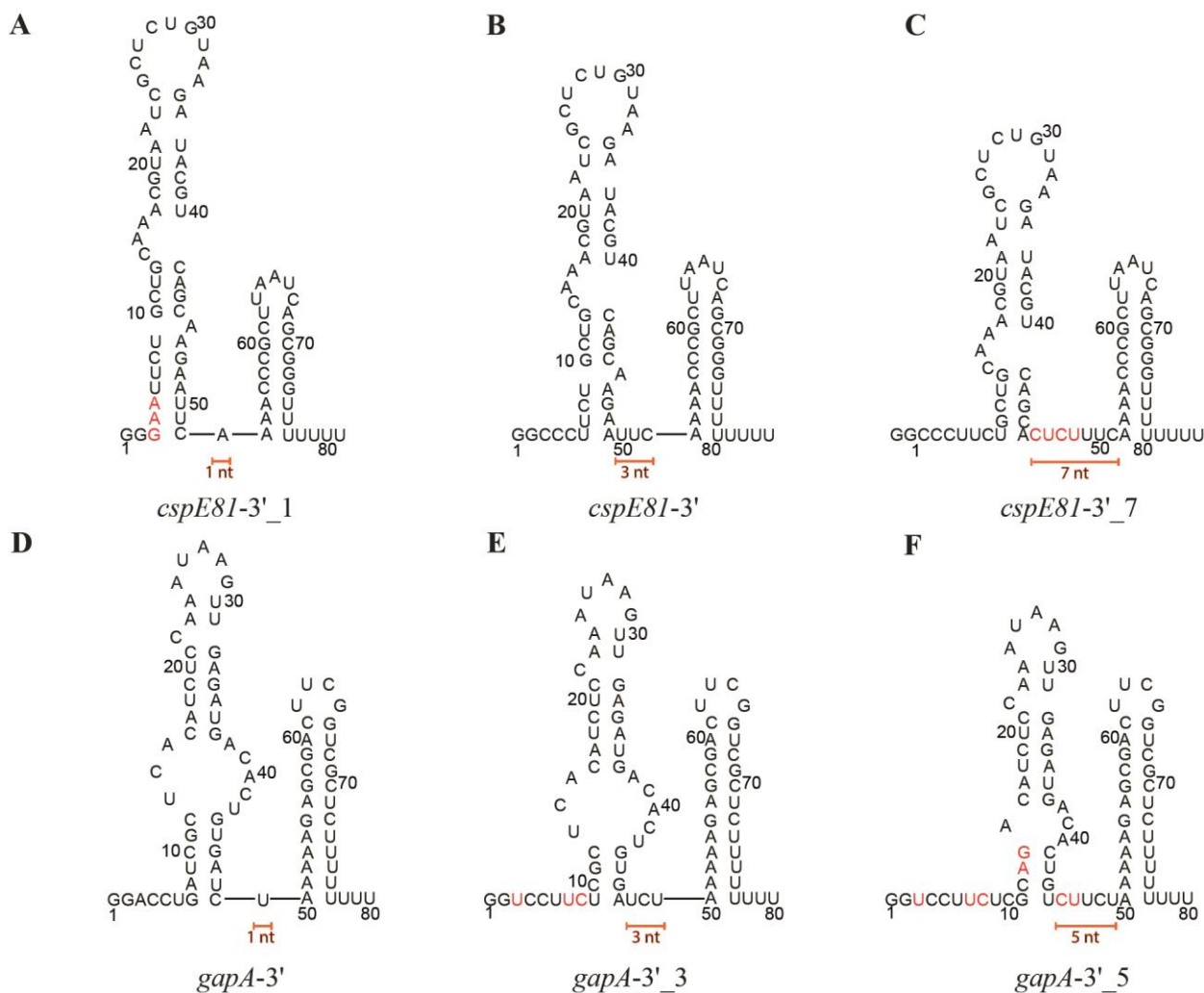
n.o. – nie obserwowano wiązania w zakresie stężeń do 1 μ M.

W przypadku wiązania do białka o substytucji R69A obserwowałam podobne różnice, z tym że ogólnie wszystkie cząsteczki słabiej wiązały się do mutantu (Rys. 31, Tabela 7). RNA *cspE-3'-4AA* i *4UU* nie wykazywały wiązania w zakresie stężeń białka 1 μ M. Hierarchia siły wiązania badanych cząsteczek dla obu białek była taka sama: AU > GC > UA > CG, co sugeruje, że dla wiązania RNA przez domenę FinO ma znaczenie geometria tej części spinki terminatorowej z preferowanym schematem par puryna-pirymidyna ponad pirymidyna-puryna. Natomiast fakt, że dla wiązania z białkiem R69A obserwowano relatywny efekt porównywalny z białkiem o naturalnej sekwencji dowodzi, że cecha ta nie jest przyczyną zróżnicowanego efektu mutacji na wiązanie RNA.

5.2.3.3. Znaczenie długości jednoniciowego fragmentu łączącego struktury drugorzędowe w obrębie cząsteczek RNA w oddziaływaniu z mutantami N-końcowej domeny białka ProQ

Aby sprawdzić czy długość jednoniciowego fragmentu łączącego struktury drugorzędowe w obrębie cząsteczki RNA ma znaczenie dla wiązania do zmutowanych wersji białka ProQ, przygotowałam zmutowane wersje cząsteczek *cspE81-3'*, *gapA-3'* oraz sztuczną cząsteczkę RNA z różną długością linkera (Rys. 32). Cząsteczka *cspE81-3'* o natywnej sekwencji posiada linker długości 3 nt, do jej sekwencji wprowadziłam

mutacje, które spowodowały skrócenie tego fragmentu do 1 nt oraz jego wydłużenie do 7 nt. W przypadku RNA *gapA-3'* natywna cząsteczka posiada tylko 1 nt linker, zatem wprowadzone mutacje poskutkowały powstaniem wariantów o sekwencjach wydłużonych do 3 i 5 nt. Wszystkie cząsteczki RNA poddałam wiązaniu do wyizolowanej N-końcowej domeny białka ProQ i jego zmutowanych wersji o substytucjach K35A i R69A.

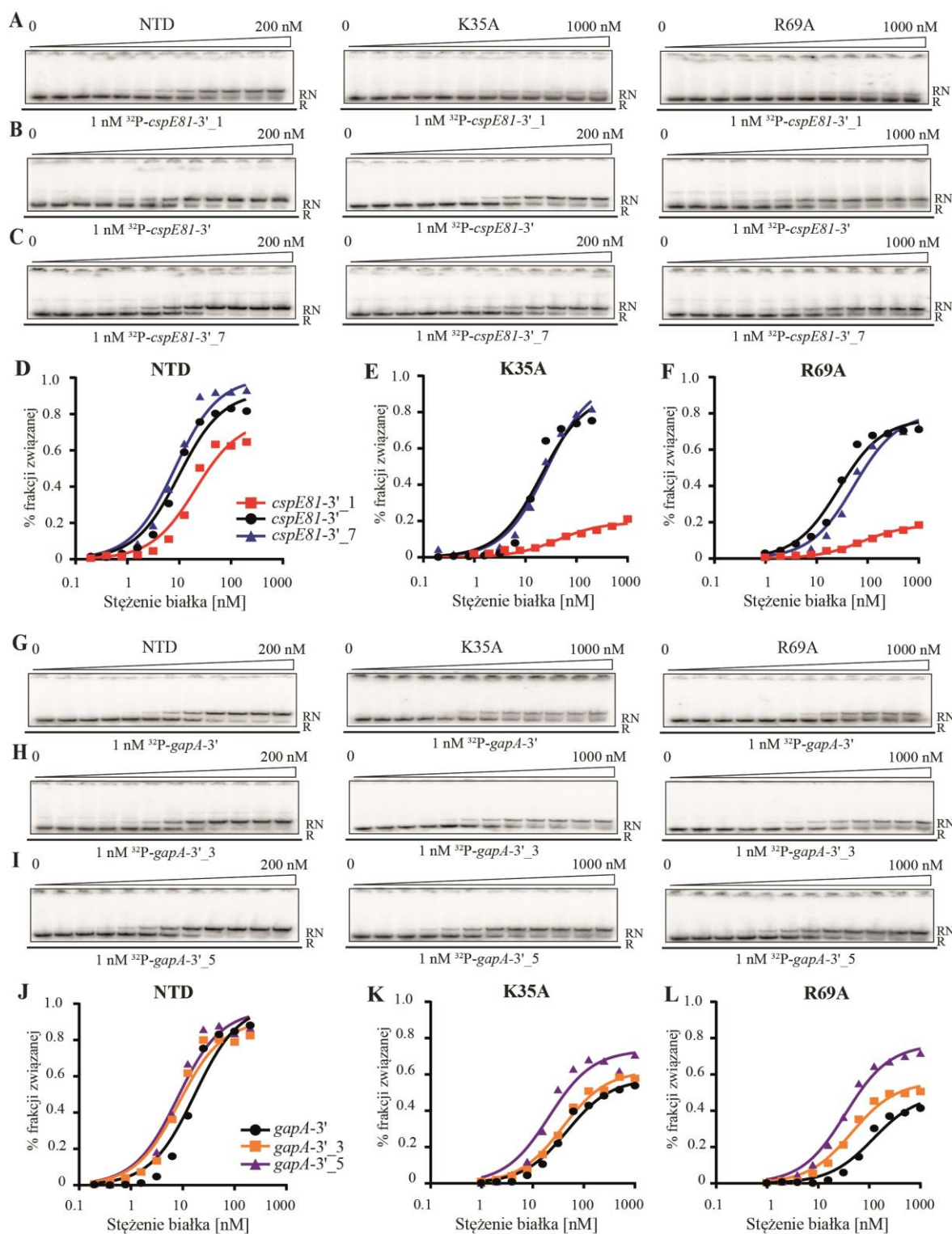


Rysunek 32. Mutanty cząsteczek *cspE81-3'* i *gapA-3'* różniące się długością linkera.

Struktury drugorzędowe cząsteczek RNA (A) *cspE81-3'_1*, (B) *cspE81-3'*, (C) *cspE81-3'_7*, (D) *gapA-3'*, (E) *gapA-3'_3* i (F) *gapA-3'_5* zostały przygotowane przy użyciu oprogramowania RNAstructure. Kolorem czerwonym oznaczono wprowadzone mutacje.

W przypadku RNA *cspE81-3'* wiązanie wszystkich wersji do białka NTD jest podobne, różnice widoczne są w przypadku wiązania ze zmutowanymi wersjami białka (Rys. 33, Tabela 8). O ile wartości K_d dla oddziaływania cząsteczek o 3 i 7 nt w linkerze są zbliżone (dla wiązania z K35A odpowiednio $22 \pm 2,6$ nM i $34 \pm 8,6$ nM), RNA o linkerze

skróconym wiąże się widocznie słabiej, ze średnią frakcją nie przekraczającą 22 %, dlatego jego K_d dla wiązania obu zmutowanych wersji określiłam jako > 1000 nM. Podobne obserwacje wypływają z eksperymentu z wariantami *gapA-3'*. W obrazie wiązania z białkiem o natywnej sekwencji różnice między oddziaływaniem z trzema wariantami różniącymi się długością linkera są niezauważalne, natomiast przy wiązaniu do białek o zmienionej sekwencji pojawia się wyraźna gradacja efektu, od silnego wiązania cząsteczki o linkerze długości 5 nt o $K_d 18 \pm 1,8$ nM dla wiązania z K35A i $37 \pm 4,4$ nM dla wiązania z R69A, do 2-4-krotnie słabsze oddziaływania z cząsteczką *gapA-3'* o naturalnej sekwencji (K_d o wartości odpowiednio $45 \pm 4,4$ nM i 140 ± 18 nM).



Rysunek 33. Skrócenie linkera w cząsteczkach RNA będących naturalnymi ligandami ProQ powoduje osłabienie wiązania do mutantów K35A i R69A.

(A) Wiązanie (A) *cspE81-3'*_1, (B) *cspE81-3'*, (C) *cspE81-3'*_7, (G) *gapA-3'*, (H) *gapA-3'*_3 i (I) *gapA-3'*_5 do białek monitorowano przy wykorzystaniu techniki EMSA, w zakresie stężeń podanym nad każdym żelem. Oznaczenia kompleksów są takie same jak na Rys. 11. Dane zależności frakcji związanej RNA od stężenia białek (D, J) NTD, (E, K) K35A i (F, L) R69A, zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 8.

Tabela 8. Wiązanie modelowych cząsteczek pochodzących z *cspE81-3'* i *gapA-3'* oraz sztucznej cząsteczki RNA z różną długością linkera do N-końcowej domeny białka ProQ i mutantów K35A i R69A.

³² P-RNA	K_d [nM] (% średniej max. frakcji związanej)		
	NTD	K35A	R69A
<i>cspE81-3'_1</i>	25 ± 8,9 (66%)	>1000 (22%)	>1000 (21%)
<i>cspE81-3'</i>	8,3 ± 2,5 (87%) ^a	22 ± 2,6 (76%) ^a	28 ± 7,5 (72%) ^a
<i>cspE81-3'_7</i>	10 ± 3,6 (87%)	34 ± 8,6 (76%)	49 ± 5,2 (74%)
<i>gapA-3'</i>	13 ± 3 (87%) ^a	45 ± 4,4 (55%) ^a	140 ± 18 (44%) ^a
<i>gapA-3'_3</i>	10 ± 2 (87%)	36 ± 3,6 (60%)	54 ± 12 (58%)
<i>gapA-3'_7</i>	6,8 ± 1,4 (91%)	18 ± 1,8 (73%)	37 ± 4,4 (71%)
model RNA_1	13 ± 1,5 (90 %)	40 ± 9,6 (70 %)	54 ± 20 (66 %)
model RNA_2	18 ± 4,6 (91 %)	47 ± 6,7 (70 %)	78 ± 12 (62 %)
model RNA_3	10 ± 1,5 (91 %)	42 ± 8,6 (73 %)	33 ± 18 (65 %)
model RNA_4	19 ± 3,9 (90 %)	56 ± 26 (73 %)	67 ± 22 (73 %)
model RNA_5	21 ± 8,6 (87 %)	60 ± 20 (79 %)	52 ± 31 (69 %)
model RNA_6	14 ± 2 (94 %)	66 ± 10 (84 %)	80 ± 13 (74 %)

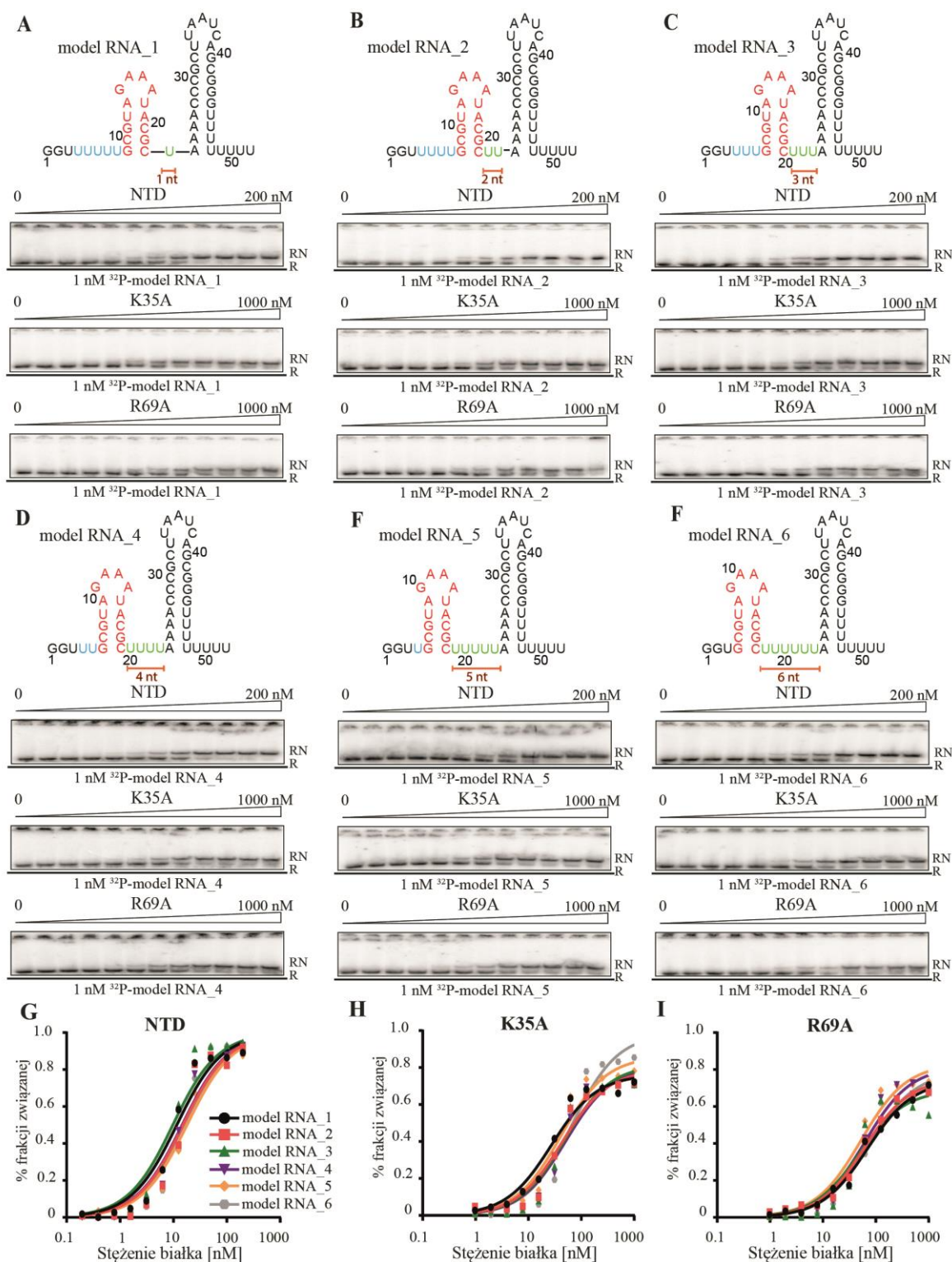
Wartości K_d uzyskano przez dopasowanie danych wykresu zależności frakcji związanego RNA od stężenia białka do równania kwadratowego w programie GraphPad. Średnie wartości K_d wraz z odchyleniem standardowym obliczono na podstawie co najmniej 3 powtórzeń z niezależnych eksperymentów.

^a – dane takie same jak w tabeli 5.

W celu sprawdzenia, czy zaobserwowana zależność odnosi się także do innych cząsteczek, w oparciu o sekwencję spinki terminatorowej cząsteczki *cspE-3'* zaprojektowałam sztuczny konstrukt RNA o zmiennej ilości urydyn (od 1 do 6 nt) w linkerze łączącym 2 struktury typu spinka do włosów (Rys. 34). W przypadku wiązania do N-końcowej domeny białka ProQ wartości K_d wszystkich reakcji mieściły się w zakresie od 10 do 21 nM (Rys. 34, tabela 8). Dla wiązania z białkiem NTD z substytucją K35A wartości K_d były 3-4-krotnie wyższe, podobnie dla wiązania z wariantem R69A,

jednak nie zaobserwowałam żadnych wyraźnych tendencji odróżniających poszczególne cząsteczki – w badanym zakresie stężeń sztuczne konstrukty modelowego RNA o zmiennej ilości urydyn w linkerze wiązały się podobnie do każdego z testowanych białek.

W oparciu o eksperyment wiązania modelowych konstruktów RNA nie można jednoznacznie potwierdzić obserwacji płynącej z wiązania zmutowanych wariantów naturalnych sekwencji. Fakt, że dłuższy linker warunkował lepsze wiązanie do białek K35A i R69A w przypadku naturalnych sekwencji mógł wynikać z tego, że konformacja zmutowanych cząsteczek ułatwiła kontakt ze zmutowanymi białkami. Natomiast cecha ta nie jest uniwersalna dla ligandów ProQ.



Rysunek 34. Zmienna ilość urydyn w linkerze sztucznej cząsteczki RNA nie ma wpływu na jej wiązanie do mutantów K35A i R69A.

Struktury drugorzędowe (A) model RNA_1, (B) model RNA_2, (C) model RNA_3, (D) model RNA_4, (E) model RNA_5 i (F) model RNA_6 zostały przygotowane przy użyciu oprogramowania RNAstructure. Dane zależności frakcji związanej RNA od stężenia białek (G) NTD, (H) K35A i (I) R69A zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad. Średnie wartości stałych K_d i maksymalnej frakcji związanej zostały przedstawione w tabeli 8. Oznaczenia kompleksów są takie same jak na Rys. 11.

6. Dyskusja

Wyniki zaprezentowane w rozprawie po raz pierwszy porównują ilościowo wiązanie domeny FinO białka ProQ do naturalnych ligandów ProQ z zastosowaniem oczyszczonych białek i cząsteczek RNA w warunkach *in vitro* oraz stanowią pierwsze porównanie znaczenia pojedynczych aminokwasów w obrębie domeny FinO białka ProQ dla wiązania różnych cząsteczek RNA *in vitro*.

NTD rozpoznaje strukturę terminatora transkrypcji niezależnego od białka Rho

Przeprowadzone analizy pozwalają na wskazanie struktury Rho-niezależnego terminatora zbudowanego ze struktury typu spinki zakończonej jednoniciowym traktem poliurydynowym, jako elementu rozpoznawanego przez domenę FinO białka ProQ. Obserwacja ta jest zgodna z wcześniejszymi doniesieniami mówiącymi o wiązaniu głównie ustrukturyzowanych cząsteczek RNA przez białko ProQ (Smirnov i wsp. 2016). Homologiczne białka pochodzące z innych gatunków bakterii także preferują wiązanie do ustrukturyzowanych RNA, domena ProQ/FinO z białka RocC u *L. pneumophila* oraz plazmidowe białko FinO wiążą Rho-niezależny terminator u swoich ligandów mocniej niż ich 5'-końce (Jerome i Frost 1999; Attaiech i wsp. 2016). Natomiast minimalne białko ProQ z *N. meningitidis* rozpoznaje nie tylko strukturę terminatora, ale również region DUS, który jest ustrukturyzowany (Chaulk i wsp. 2011; Bauriedl i wsp. 2020). Ponadto badania CLIP-seq (Holmqvist i wsp. 2018) i RIL-seq (Melamed i wsp. 2020) pokazały, że wśród ligandów białka ProQ dominują cząsteczki pochodzące z regionów 3'UTR mRNA i sRNA, które często jako element struktury zawierają Rho-niezależny terminator.

Ogon poliurydynowy jest ważny dla optymalnego oddziaływania z NTD

Przedstawione badania wskazują części struktury Rho-niezależnego terminatora, które są rozpoznawane lub bezpośrednio uczestniczą w wiązaniu do NTD: region dwuniciowy w dolnej części spinki terminatorowej i ogon poliurydynowy o długości co najmniej 4 nt. Jest to spójne z obserwacją mówiącą, że ogony ligandów ProQ są krótsze w porównaniu do ligandów Hfq (Holmqvist i wsp. 2018). Istotność dolnej części spinki terminatorowej, jak również przyległych do niej sekwencji, była także obserwowana w kontekście oddziaływań białka FinO z plazmidu F i sRNA FinP (Jerome i Frost 1999; Jerome i wsp. 1999; Arthur i wsp. 2011): związanie z białkiem chroniło dolną część spinki RNA przed trawieniem RNAzą E. Dane te sugerują, że rozpoznanie dolnej części trzonu

spinki terminatorowej wraz z ogonem poliurydynowym może być ogólną cechą białek należących do rodziny białek z domeną FinO.

Sekwencja wzbogacona w adenozyne zlokalizowana bezpośrednio po 5'-stronie spinki terminatorowej służy jako negatywna determinanta wiązania do Hfq

Mgr Maciej Basczok z Pracowni Biochemii RNA zaobserwował statystyczne wzbogacenie ligandów ProQ w adenozyne występujące po 5'-stronie spinki terminatorowej, co sugerowało rolę tego motywu jako cechy rozpoznawanej przez białko ProQ (Stein i wsp. 2020). Hipoteza została przetestowana *in vitro* przez mnie i dr Joannę Kwiatkowską, a opisany motyw służył raczej jako negatywna determinanta wiązania do białka Hfq. Obserwacje *in vitro* zostały potwierdzone *in vivo* przez zespół prof. Katherine Berry z Mount Holyoke College w South Hadley (Stein i wsp. 2020). Białko Hfq również rozpoznaje zbudowany z urydyn ogon na 3'-końcu RNA (Otaka i wsp. 2011; Sauer i Weichenrieder 2011; Ishikawa i wsp. 2012). Jednak w przypadku wiązanych przez nie sRNA z klasy I, które są dominującą grupą wiązanych przez Hfq RNA, dodatkowo do wiązania potrzebny jest motyw bogaty w urydyny znajdujący się bezpośrednio po 5' stronie spinki terminatorowej, czyli w takiej samej lokalizacji gdzie w ligandach ProQ znajduje się sekwencja bogata w adenozyne (Soper i Woodson 2008; Updegrave i wsp. 2016; Chen i Gottesman 2017). Natomiast w przypadku regulatorowych RNA klasy II dodatkowo do wiązania przez Hfq potrzebne są sekwencje bogate w adenozyne, jednak znajdujące się daleko przed spinką terminatorową (Małecka i wsp. 2015; Schu i wsp. 2015; Kwiatkowska i wsp. 2018).

Dane zaprezentowane w niniejszej pracy wspierają model, w którym motyw wzbogacony w adenozyne zlokalizowany po 5'-stronie trzonu spinki terminatorowej u ligandów białka ProQ służy jako negatywna determinanta wiązania do białka Hfq. Już wcześniej zauważono, że ligandy ProQ mają krótsze ogony poliU w porównaniu do cząsteczek wiązanych przez Hfq (Holmqvist i wsp. 2018), co może mieć związek z tym, że część urydyn zaangażowana jest w wiązanie do adenozyne zlokalizowanych po przeciwnej stronie trzonu spinki terminatorowej. Z drugiej strony brak urydyn po 5' stronie spinki terminatorowej być może zapobiega wiązaniu do pierścienia Hfq, który ma powinowactwo do sekwencji wzbogaconych w AU (Updegrave i wsp. 2016). Pomijając to, który z tych mechanizmów jest ważniejszy, zgodnie z badaniami *in vitro* (Rys. 10-11) i *in vivo* (Stein i wsp. 2020) osłabia to wiązanie do Hfq, co sugeruje, że poprawne rozpoznanie

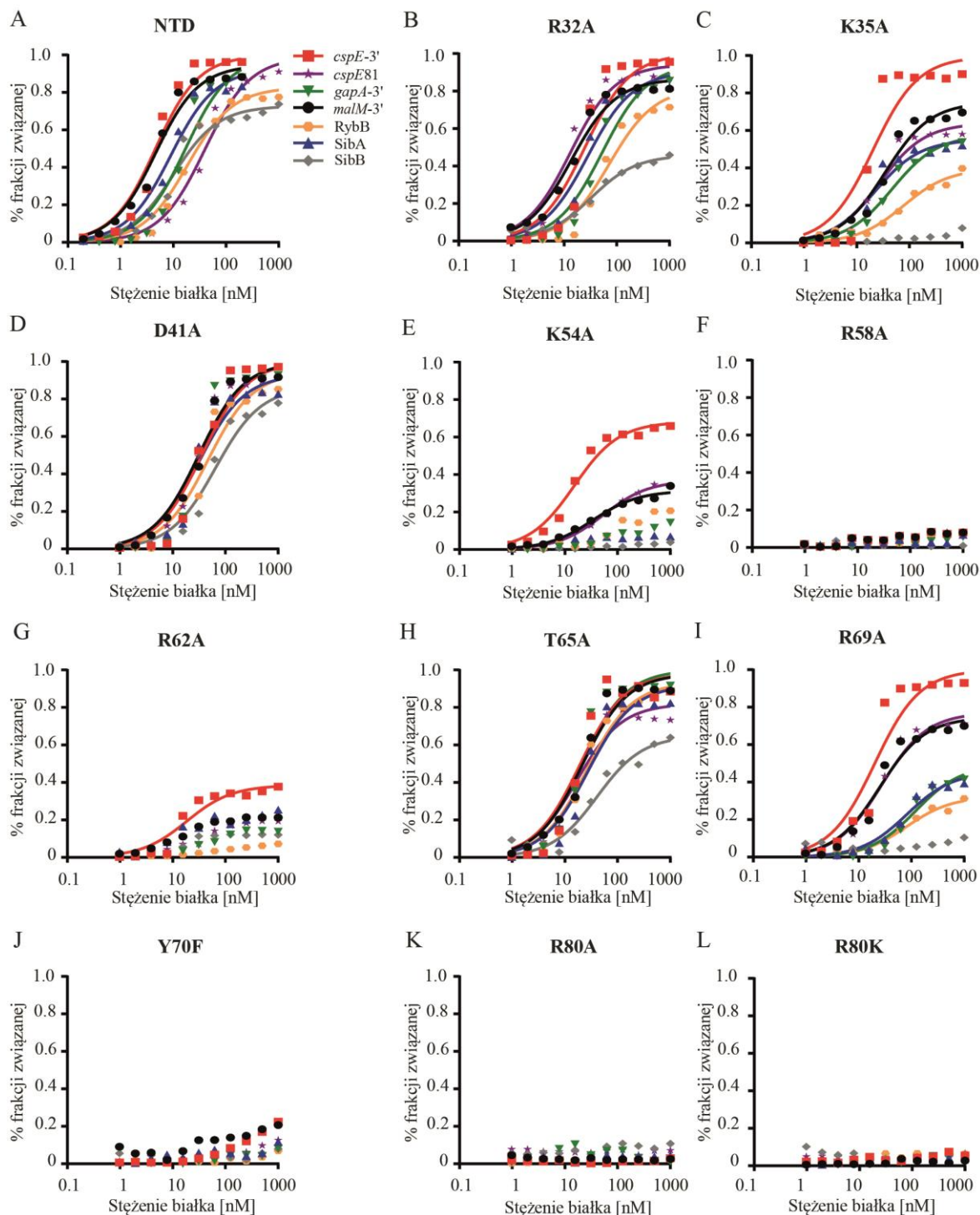
ligandu przez ProQ choć w części polega na kompetycji z Hfq o dostęp do cząsteczki RNA.

Cząsteczki RNA wiążą się do NTD w różny sposób

Wyniki wiązania *in vitro* siedmiu cząsteczek będących naturalnymi ligandami ProQ do mutantów N-końcowej domeny białka ProQ sugerują, że cząsteczki te wiążą się do NTD w różny sposób (Rys. 35). Mutacje argininy 58, tyrozyny 70 i argininy 80 miały jednakowo negatywny wpływ na wiązanie wszystkich badanych cząsteczek, natomiast substytucje lizyny 35 na wklęsłej stronie NTD oraz argininy 69 zlokalizowanej na jego wypukłej stronie wywołały szeroki zakres efektów. W obu przypadkach największe osłabienie dotyczyło wiązania 143-nt cząsteczki SibB, natomiast najmniejszy wpływ badane mutacje wykazały na wiązanie 52-nt fragmentu *cspE*-3'. Różnice w oddziaływaniu RNA z białkiem ProQ zasugerowano także wcześniej, w oparciu o różnice zakresu obszarów chronionych przed wymianą atomów wodoru na deuter (HDX) w NTD, CTD i linkerze podczas wiązania 85-nt fragmentu *cspE* i 169-nt sRNA SraB (Gonzalez i wsp. 2017).

W badaniach *in vivo* (Pandey i wsp. 2020), podobnie jak w przedstawionych badaniach *in vitro*, substytucja K35A w mutancie pozbawionym CTD również wykazała zróżnicowany wpływ na wiązanie 2 cząsteczek RNA: zahamowała wiązanie SibB, a tylko lekko osłabiła wiązanie 81-nt fragmentu *cspE*. Odpowiedź na pytanie, co powoduje tak różne sposoby wiązania cząsteczek RNA do domeny FinO ujawnione przez mutacje K35A i R69A jednak nie jest jednoznaczna. Potencjalnym wytłumaczeniem, dlaczego efekt mutacji w tych dwóch pozycjach jest widoczny jedynie dla niektórych z badanych RNA może być sytuacja, w której tylko niektóre cząsteczki RNA bezpośrednio oddziałują ze zmienionymi aminokwasami. Możliwa jest również sytuacja, w której dodatkowe kontakty z RNA w innych miejscach białka kompensują wpływ substytucji K35A i R69A. Podobna hierarchia efektów wiązania, lecz w znacznie węższym zakresie – z *cspE*-3' wiązaniem najsilniej i SibB najslabiej – może być zaobserwowana w przypadku substytucji K54A i R62A (Rys. 27). Nie można zatem wykluczyć możliwości, że obserwowana różnica efektów wiązania do białek zawierających substytucje K35A i R69A nie wynika z różnych sposobów wiązania RNA do zmienionych aminokwasów, gdyż mogą pojawić się oddziaływania kompensujące w innych obszarach białka, które to interakcje dla jednych RNA są silniejsze niż dla innych. Podczas gdy ta opcja wydaje się jednak bardziej

prawdopodobna, możliwa rola w pozycjonowaniu RNA na białku przez reszty K35 i R96 nie może zostać wykluczona w oparciu o zaprezentowane dane *in vitro*.



Rysunek 35. Cząsteczki RNA różnią się w sposobach wiązania do zmutowanych wersji N-końcowej domeny białka ProQ.

Dane zależności frakcji związanej *cspE-3'*, *cspE81-3'*, *gapA-3'*, *malM-3'*, RybB, SibA i SibB od stężenia białek (A) NTD, (B) R32A, (C) K35A, (D) D41A, (E) K54A, (F) R58A, (G) R62A, (H) T65A, (I) R69A, (J) Y70F, (K) R80A i (L) R80K zostały dopasowane do równania kwadratowego przy użyciu oprogramowania GraphPad.

Reszty K54, R58, R62, Y70 i R80 uczestniczą w oddziaływaniu z cząsteczkami RNA

Wyniki pomiarów siły wiązania *in vitro* pokazują, że kilka aminokwasów zlokalizowanych na powierzchni N-końcowej domeny białka ProQ jest istotnych dla oddziaływania z cząsteczkami RNA. Uniwersalnie zakonserwowaną resztą aminokwasową, która może być zaangażowana w wiązanie RNA, jest arginina 80. Obserwacje poczynione *in vitro* potwierdzają zgubny efekt mutacji R80A na wiązanie RNA *in vivo* (Pandey i wsp. 2020). W genetycznym teście przesiewowym opartym o B3H u *E. coli*, substytucje reszty R80 seryną, histydyną czy cysteiną także zaburzały wiązanie do RNA (Pandey i wsp. 2020). Ponadto, w dwóch niezależnych genetycznych testach przesiewowych poszukujących mutacji upośledzających działanie białka ProQ u *S. enterica*, zidentyfikowano substytucje R80 do histydyny, seryny, glicyny i alaniny (El Mouali i wsp. 2021b), lub do histydyny (Rizvanovic i wsp. 2021). W warunkach *in vivo* białko ProQ z substytucją R80A było podatne na degradację przez proteazę Lon, która trawi potencjalnie toksyczne lub niefunkcjonalne białka (El Mouali i wsp. 2021b). Interesującą kwestią pozostaje umiejscowienie łańcucha bocznego reszty R80 na powierzchni białka. Zgodnie ze strukturą N-końcowej domeny białka ProQ z *E. coli* uzyskaną na drodze spektroskopii magnetycznego rezonansu jądrowego (Gonzalez i wsp. 2017), reszta R80 zlokalizowana jest na stronie wypukłej w sąsiedztwie reszty R32 (Rys. 13). Natomiast w strukturze przewidzianej przez program AlphaFold łańcuch boczny reszty R80 skierowany jest w drugą stronę ku powierzchni wklęsłej, ulokowany w sąsiedztwie reszt Y70 i K35 (Rys. 21). W strukturze białka Lpp1663 z *L. pneumophila* homologiczna reszta R86 również skierowana jest w stronę powierzchni wklęsłej (Immer i wsp. 2020), podobnie jak w strukturach krystalicznych białek NMB1681 z *N. meningitidis*, FinO z *E. coli* (Immer i wsp. 2020) i RocC z *L. pneumophila* (Kim i wsp. 2022).

Kolejną resztą o potwierdzonym *in vitro* wpływie na wiązanie cząsteczek RNA jest tyrozyna 70. Izosteryczna substytucja do fenyloalaniny zahamowała wiązanie do RNA, co pozostaje w zgodzie z danymi *in vivo* pokazującymi zgubny efekt substytucji w tej pozycji u ProQ z *E. coli* (Pandey i wsp. 2020) lub jego homologów u innych gatunków (Attaiech i wsp. 2016; Immer i wsp. 2020; El Mouali i wsp. 2021a; El Mouali i wsp. 2021b; Rizvanovic i wsp. 2021). Sąsiadujące reszty aminokwasowe R69 i L71, albo ich odpowiedniki u homologów (Attaiech i wsp. 2016; Immer i wsp. 2020; El Mouali i wsp.

2021b; Rizvanovic i wsp. 2021; Kim i wsp. 2022), są również istotne dla wiązania, być może mają wpływ na pozycjonowanie reszty Y70. Ponadto zarówno w strukturze NTD uzyskanej na drodze spektroskopii magnetycznego rezonansu jądrowego (Gonzalez i wsp. 2017), jak i w predykcji z AlphaFold, grupa hydroksylowa tej reszty skierowana jest ku powierzchni białka (Rys. 13). W strukturze krystalicznej kompleksu białka RocC z *L. pneumophila* z cząsteczką RNA reszta Y87 homologiczna do Y70 z ProQ u *E. coli*, jak również reszta R97 z RocC homologiczna do R80 z ProQ zaangażowane są w bezpośrednie wiązanie grupy fosforanowej urydyny w pozycji 28 badanego fragmentu cząsteczki RNA RocR (Kim i wsp. 2022).

Podobnie negatywny wpływ na wiązanie badanych RNA *in vitro* miała substytucja argininy 58 alaniną, co potwierdza obserwację efektu tej mutacji w eksperymencie B3H (Pandey i wsp. 2020). Arginina w homologicznej pozycji 121 na plazmidowym białku FinO wiązała się poprzez crosslink z RNA SLII (Ghetu i wsp. 2002); boczny łańcuch R58 u NTD ProQ jest dostępny na powierzchni wklęsłej, co również przemawia za zaangażowaniem tej reszty w wiązanie cząsteczek RNA.

Silny efekt osłabienia wiązania zaobserwowano także dla białek z substytucjami w pozycjach K54 i R62, które ewolucyjnie zlokalizowane są w pozycjach zmiennych. Substytucja K54A osłabiała wiązanie RNA w eksperymencie B3H (Pandey i wsp. 2020), nie zidentyfikowano jej jednak w żadnym z genetycznych testów przesiewowych u ProQ z *E. coli* ani jego homologów z innych gatunków. Za udziałem tej reszty w oddziaływaniu z RNA może świadczyć fakt, że łańcuchy boczne homologów reszt K54 i R58 w białku RocC z *L. pneumophila*, odpowiednio K71 i R75, związane są z cytydynami w pozycjach 22 i 23 badanego fragmentu cząsteczki RNA RocR (Kim i wsp. 2022). Reszty R62 także nie zidentyfikowano w żadnym z genetycznych testów przesiewowych, jednak lizyna w homologicznej pozycji 125 na białku FinO wiązała się poprzez crosslink z RNA SLII (Ghetu i wsp. 2002), co może sugerować, że reszta R62 kontaktuje się z RNA, tym bardziej że znajduje się ona na powierzchni białka.

Warto zaznaczyć, że wszystkie reszty aminokwasowe omówione powyżej i istotne dla wiązania z cząsteczkami RNA *in vitro*, zgodnie z modelem struktury uzyskanym dzięki oprogramowaniu AlphaFold (Rys. 21 B), zlokalizowane są w bliskim sąsiedztwie po stronie wklęsłej domeny FinO białka ProQ. Można zatem stwierdzić, że tworzą one kieszeń wiążącą cząsteczkę RNA.

Udział reszt R32, D41 i T65 w wiązaniu cząsteczek RNA jest niepotwierdzony

Kilka pozostałych z analizowanych substytucji miało nieznaczny wpływ na wiązanie cząsteczek RNA. Takim przykładem jest substytucja asparaginianu 41 do alaniny, która *in vitro* nie wpływała na wiązanie RNA *gapA-3'* i RybB lub 2-5-krotnie je osłabiała w przypadku cząsteczek *cspE-3'*, *cspE81-3'*, *malM-3'*, SibA i SibB (Rys. 35). Ta sama substytucja w badaniach *in vivo* wykazywała jednoznacznie szkodliwy wpływ na wiązanie RNA w B3H (Pandey i wsp. 2020), podobnie jak substytucja tyrozyną w homologicznej pozycji D56 w białku RocC u *L. pneumophila* (Attaiech i wsp. 2016). Uzyskane wyniki mogą sugerować, że reszta D41 nie oddziałuje z RNA konkurującymi o dostęp do białka, co zaobserwowałam bezpośrednio w systemie z oczyszczonymi składnikami, natomiast zahamowanie wiązania do mutantów D41 w warunkach *in vivo* może być spowodowane kontaktem z innymi komórkowymi komponentami, których brakuje *in vitro*. W pozycji 32, gdzie w ProQ u *E. coli* arginina zlokalizowana jest na powierzchni wypukłej, ewolucyjnie u innych gatunków znajdują się lizyny lub argininy, co może sugerować udział tego aminokwasu w wiązaniu RNA. Jednak substytucja alaniną nie wpłynęła znacząco na wiązanie badanych cząsteczek w teście *in vitro* – w przypadku RNA *cspE81-3'*, *gapA-3'* i *malM-3'* wartości K_d były takie same jak dla oddziaływania z wt NTD, natomiast dla cząsteczek *cspE-3'*, RybB, SibA i SibB zanotowano jedynie 2-3-krotne osłabienie wiązania. Podobnie rzecz ma się z treoniną 65, która w modelu strukturalnym bezpośrednio przylega do zakonserwowanej T70, natomiast ewolucyjnie w tej pozycji występują treonina, seryna lub cysteina, co sugeruje jej potencjalną rolę w tworzeniu wiązania wodorowego. W przypadku 3 z 7 badanych RNA (*cspE81-3'*, *gapA-3'* i RybB) substytucja alaniną nie wpływa na wiązanie, natomiast u pozostałych (*cspE-3'*, *malM-3'*, SibA i SibB) 2-4-krotnie je osłabia. Zaobserwowane zależności mogą sugerować, że reszty R32 i T65 pełnią raczej rolę w utrzymaniu struktury białka, lecz nie uczestniczą bezpośrednio w wiązaniu RNA.

Rola CTD i linkera w wiązaniu cząsteczek RNA przez białko ProQ

Bardzo ciekawą kwestią w kontekście zaprezentowanych badań wydaje się rola domeny C-końcowej białka ProQ i bogatego w dodatnio naładowane reszty aminokwasowe linkera. Przedstawione wyniki pokazują, że obecność Rho-niezależnego terminatora na 3'-końcu cząsteczki RNA warunkuje wiązanie do NTD, natomiast wiązanie pełnej długości białka ProQ wydaje się być mniej zależne od obecności tej struktury.

Zostało to potwierdzone w przypadku wiązania fragmentów pochodzących z mRNA *lpp* i *hupA* z regionów nieulegających translacji na 5'-końcu, które do pełnej długości białka wiążą się z silnym powinowactwem, podczas gdy ich oddziaływanie z wyizolowaną domeną FinO jest silnie osłabione (Rys. 11). Pośrednio sugeruje to kontakt pozostałych części funkcjonalnych białka z cząsteczką RNA. Zdolność linkera i CTD do niespecyficznego wiązania RNA została potwierdzona (Sheidy i Zielke 2013; Gonzalez i wsp. 2017; Pandey i wsp. 2020), zauważono także różnice w oddziaływaniu z różnymi RNA. W eksperymentach *in vivo* wykorzystujących HDX i B3H wykazano, że CTD i linker wpływają na wiązanie sRNA SraB i SibB (Gonzalez, 2017, Pandey, 2020), które są dużo dłuższe niż 52-nt *cspE-3'* (SraB – 169 nt, SibB 143 – nt). Możliwe, że linker i CTD formują dodatkowe kontakty z RNA wiązanymi przez NTD lub znajdującymi się w pobliżu. Pokazano, że linker wspólnie z NTD również był niezbędny przy wiązaniu ProQ do rybosomów (Sheidy i Zielke 2013).

Zastanawiające jest także, dlaczego niektóre z białek z rodziny, takie jak FinO (Ghetu i wsp. 2002), RocC (Attaiech i wsp. 2016) czy FopA (Gerovac i wsp. 2020) rozpoznają tylko kilka RNA, podczas gdy ProQ oddziałuje z setkami cząsteczek. Ponadto u jednego gatunku bakterii może występować więcej niż jedno białko z domeną ProQ, np. u *E. coli* znaleziono dwa takie białka (ProQ i plazmidowe białko FopA), natomiast u *S. enterica* aż trzy (ProQ oraz plazmidowe białka FinO i FopA). Wydaje się, że spośród białek z tej rodziny tylko ProQ można nazwać globalnie oddziałującym białkiem opiekuńczym, pozostałe wiążą jedynie wąskie spektrum ligandów.

Zaobserwowane zależności i opisane spostrzeżenia dotyczące oddziaływań N-końcowej domeny białka ProQ z cząsteczkami RNA stanowią próbę zrozumienia procesów regulacji ekspresji genów zależnej od niekodujących RNA u bakterii. Mając na uwadze charakter prowadzonych badań, należy jednak pamiętać, że system *in vitro* nie odzwierciedla dokładnie całej złożoności procesów zachodzących jednocześnie w żywej komórce. Natomiast niewątpliwie przedstawione analizy biochemiczne uzupełniają wiedzę o mechanizmach rozpoznawania cząsteczek RNA przez białka wiążące RNA *in vivo* oraz wskazują nowe aspekty oddziaływań białko-RNA.

7. Wnioski

W toku pracy sformułowano następujące wnioski końcowe:

- Struktura Rho-niezależnego terminatora zbudowanego ze struktury typu spinki zakończonej ogonem poliurydynowym jest elementem RNA rozpoznawanym przez N-kończącą domenę FinO białka ProQ.
- Dolna część spinki terminatorowej uczestniczy w wiązaniu cząsteczek RNA przez domenę FinO białka ProQ.
- Minimalna długość ogona poliurydynowego wymagana do optymalnego związania z domeną FinO białka ProQ wynosi 4 nukleotydy.
- Motyw wzbogacony w adenozyiny występujący po 5'-stronie spinki terminatorowej cząsteczek RNA wiązanych przez białko ProQ w komórkach *E. coli* osłabia wiązanie do białka Hfq, co sugeruje kompetycję między białkami ProQ i Hfq o wiązanie do cząsteczek RNA.
- Reszty aminokwasowe zlokalizowane w kieszeni wiążącej po stronie wklęsłej (K54, R58, R62, Y70 i R80) są istotne dla wiązania domeny FinO białka ProQ z cząsteczkami RNA, podczas gdy udział reszt R32, D41 i T65 w wiązaniu nie został potwierdzony *in vitro*.
- Różna wrażliwość wiązania siedmiu badanych cząsteczek RNA na mutacje K35A i R69A sugeruje odmienne sposoby ich oddziaływania z N-kończącą domeną białka ProQ, jednak cechy budowy RNA wpływające na to pozostają nieznane.

Wykaz rysunków

Rysunek 1. Aranżacja genomowa i funkcje regulatorowe sRNA.	16
Rysunek 2. Struktura i powierzchnie wiążące RNA białka Hfq.	18
Rysunek 3. Schematyczne przedstawienia działania systemu FinOP regulującego translację mRNA <i>traJ</i>	19
Rysunek 4. Schematyczna budowa domenowa wybranych białek z rodziny FinO. ...	20
Rysunek 5. Budowa białka ProQ.	23
Rysunek 6. Model struktury pełnej długości białka ProQ z <i>E. coli</i> wygenerowany przez program AlphaFold (PDB: AF_AFP45577F1).....	24
Rysunek 7. Proponowany model zależnej od białka ProQ regulacji translacji mRNA <i>hupA</i> przez sRNA RaiZ.	25
Rysunek 8. Porównanie liczby ligandów białek ProQ i Hfq u (A) <i>S. enterica</i> i (B) <i>N. meningitidis</i> w zależności od pochodzenia regulatorowego RNA.....	26
Rysunek 9. Schemat przygotowania konstruktów do nadekspresji białek pełnej długości ProQ i jego N-końcowej domeny.	40
Rysunek 10. Częsteczki RNA o zróżnicowanym pochodzeniu i budowie użyte w eksperymencie wyznaczenia stałej dysocjacji w równowadze.	50
Rysunek 11. Wiązanie badanych RNA do białka ProQ z <i>E. coli</i> i jego N-końcowej domeny (NTD).	52
Rysunek 12. Dodanie struktury spinki z ogonem poliU na 3'-końcu <i>lpp-5'</i> i <i>hupA-5'</i> powoduje ich silne wiązanie przez NTD.....	55
Rysunek 13. Wydłużenie sekwencji RNA poza strukturę terminatora transkrypcji wpływa negatywnie na wiązanie do N-końcowej domeny białka ProQ.	57
Rysunek 14. Skrócenie ogona poliU cząsteczek RNA <i>malM-3'</i> i <i>cspE-3'</i> prowadzi do zaniku wiązania do N-końcowej domeny białka ProQ.	59
Rysunek 15. Dolna część spinki terminatorowej uczestniczy w wiązaniu cząsteczki <i>malM-3'</i>	62
Rysunek 16. Mutacje wprowadzone do sekwencji cząsteczek RNA w celu przetestowania roli motywu bogatego w adenozyliny zlokalizowanego po 5'-stronie trzonu spinki terminatorowej.	66
Rysunek 17. Mutacja w obszarze motywu bogatego w adenozyliny zlokalizowanego po 5'-stronie trzonu spinki terminatorowej nie wpływa lub osłabia wiązanie do N-końcowej domeny białka ProQ.	68
Rysunek 18. Mutacja w obszarze motywu bogatego w adenozyliny zlokalizowanego po 5'-stronie trzonu spinki terminatorowej poprawia wiązanie do białka Hfq.	70
Rysunek 19. Kompetycja o wiązanie RNA przez białka ProQ i Hfq.....	Błąd! Nie zdefiniowano zakładki.
Rysunek 20. Porównanie sekwencji aminokwasowych domen FinO sześciu białek należących do rodziny białek FinO o znanych strukturach i/lub ligandach.	71
Rysunek 21. Pozycje aminokwasowe wybrane do mutagenyzy zaprezentowane na strukturze N-końcowej domeny białka ProQ uzyskanej w (A) NMR ((Gonzalez i wsp. 2017); PDB: 5nb9) oraz (B) odpowiadający mu 119-nt fragment struktury uzyskany jako predykcja w programie AlphaFold.....	73
Rysunek 22. Częsteczki RNA , których wiązanie do mutantów N-końcowej domeny białka ProQ zostało zmierzone w pracy.....	75

Rysunek 23. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA <i>cspE</i>-3'.....	77
Rysunek 24. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA <i>cspE81</i>-3'.....	78
Rysunek 25. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA <i>gapA</i>-3'.....	80
Rysunek 26. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA <i>malM</i>-3'.....	81
Rysunek 27. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA RybB.	82
Rysunek 28. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA SibA.....	84
Rysunek 29. Znaczenie substytucji różnych aminokwasów w obrębie N-końcowej domeny białka ProQ dla wiązania RNA SibB.....	85
Rysunek 30. Zarówno ogon poliurydynowy zlokalizowany na 3'-końcu cząsteczki RNA, jak i jednoniciowa sekwencja ulokowana po 5'-stronie trzonu spinki terminatorowej wpływają na wiązanie do N-końcowej domeny białka ProQ.	87
Rysunek 31. Zaburzenie struktury dolnej części spinki terminatorowej RNA <i>cspE</i>-3' powoduje zwiększone osłabienie wiązania do mutantu R69A w porównaniu do wiązania z NTD.....	89
Rysunek 32. Mutanty cząsteczek <i>cspE81</i>-3' i <i>gapA</i>-3' różniące się długością linkera.	91
Rysunek 33. Skrócenie linkera w cząsteczkach RNA będących naturalnymi ligandami ProQ powoduje osłabienie wiązania do mutantów K35A i R69A.	93
Rysunek 34. Zmienna ilość urydyn w linkerze sztucznej cząsteczki RNA nie ma wpływu na jej wiązanie do mutantów K35A i R69A.....	96
Rysunek 35. Cząsteczki RNA różnią się w sposobach wiązania do zmutowanych wersji N-końcowej domeny białka ProQ.	100

Wykaz tabel

Tabela 1. Wiązanie cząsteczek RNA i ich mutantów z wydłużonymi lub zmienionymi końcami 3' do pełnej długości białka ProQ i jego N-końcowej domeny (NTD).	53
Tabela 2. Wiązanie wariantów cząsteczek <i>malM-3'</i> i <i>cspE-3'</i> oraz ich mutantów z różną liczbą urydyn w ogonie poliU do N-końcowej domeny białka ProQ.	60
Tabela 3. Wiązanie cząsteczek RNA i ich mutantów z wydłużonymi lub zmienionymi końcami 3' do pełnej długości białka ProQ i jego N-końcowej domeny.	63
Tabela 4. Wiązanie cząsteczek RNA i ich mutantów ze zmutowanym regionem bogatym w adenozyiny do N-końcowej domeny białka ProQ i Hfq.	69
Tabela 5. Wiązanie cząsteczek RNA do mutantów N-końcowej domeny białka ProQ. .	76
Tabela 6. Wiązanie modelowych cząsteczek pochodzących z <i>cspE-3'</i> do N-końcowej domeny białka ProQ i mutantów K35A i R69A.	88
Tabela 7. Wiązanie modelowych cząsteczek pochodzących z <i>cspE-3'</i> ze zmienioną strukturą dolnej części spinki terminatorowej do N-końcowej domeny białka ProQ i nego mutantu R69A.	90
Tabela 8. Wiązanie modelowych cząsteczek pochodzących z <i>cspE81-3'</i> i <i>gapA-3'</i> oraz sztucznej cząsteczki RNA z różną długością linkera do N-końcowej domeny białka ProQ i jego mutantów K35A i R69A.	94

Bibliografia

- Andrade JM, Dos Santos RF, Chelysheva I, Ignatova Z, Arraiano CM. 2018. The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity *EMBO Journal* **37**: e97631.
- Andrade JM, Pobre V, Matos AM, Arraiano CM. 2012. The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq *RNA* **18**: 844-855.
- Arthur DC, Edwards RA, Tsutakawa S, Tainer JA, Frost LS, Glover JNM. 2011. Mapping interactions between the RNA chaperone FinO and its RNA targets. *Nucleic Acids Research* **39**: 4450-4463.
- Asano K, Hama C, Inoue S-i, Moriwaki H, Mizobuchi K. 1999. The plasmid ColIb-P9 antisense Inc RNA controls expression of the RepZ replication protein and its positive regulator repY with different mechanisms. *The Journal of Biological Chemistry* **274**: 17924-17933.
- Attaiech L, Boughammoura A, Brochier-Armanet C, Allatif O, Peillard-Fiorente F, Edwards RA, Omar AR, Macmillan AM, Glover M, Charpentier X. 2016. Silencing of natural transformation by an RNA chaperone and a multitarget small RNA. *Proceedings of the National Academy of Sciences* **113**: 8813-8818.
- Avrani S, Bolotin E, Katz S, Hershberg R. 2017. Rapid Genetic Adaptation during the First Four Months of Survival under Resource Exhaustion. *Molecular Biology and Evolution* **34**: 1758-1769.
- Balbontín R, Fiorini F, Figueroa-Bossi N, Casadesús J, Bossi L. 2010. Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in *Salmonella enterica*. *Molecular Microbiology* **78**: 380-394.
- Bauriedl S, Gerovac M, Heidrich N, Bischler T, Barquist L, Vogel J, Schoen C. 2020. The minimal meningococcal ProQ protein has an intrinsic capacity for structure-based global RNA recognition. *Nature Communications* **11**, 2823.
- Bidnenko E, Bidnenko V. 2018. Transcription termination factor Rho and microbial phenotypic heterogeneity. *Current Genetics* **64**: 541-546.
- Cech GM, Szalewska-Pałasz A, Kubiak K, Malabirade A, Grange W, Arluison V, Węgrzyn G. 2016. The *Escherichia coli* Hfq Protein: An Unattended DNA-Transactions Regulator *Frontiers in Molecular Biosciences* **3**.
- Chao Y, Papenfort K, Reinhardt R, Sharma CM, Vogel J. 2012. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs *EMBO Journal* **31**: 4005-4019.
- Chaulk SG, Lu J, Tan K, Arthur DC, Edwards RA, Frost LS, Joachimiak A, Glover JNM. 2010. *N. meningitidis* 1681 is a member of the FinO family of RNA chaperones. *RNA Biology* **7**: 812-819.

- Chaulk SG, Smith–Friday MN, Arthur DC, Culham DE, Edwards RA, Soo P, Frost LS, Keates RAB, Glover JNM, Wood JM. 2011. ProQ Is an RNA Chaperone that Controls ProP Levels in *Escherichia coli*. *Biochemistry* **50**: 3095-3106.
- Chen J, Gottesman S. 2017. Hfq links translation repression to stress-induced mutagenesis in *E. coli*. *Genes & Development* **31**: 1382-1395.
- Coornaert A, Lu A, Mandin P, Springer M, Gottesman S, Guillier M. 2010. MicA sRNA links the PhoP regulon to cell envelope stress. *Molecular Microbiology* **76**: 467-479.
- Dar D, Sorek R. 2018. Bacterial Noncoding RNAs Excised from within Protein-Coding Transcripts *mBio* **9**: e01730-18.
- Dimastrogiovanni D, Frochlich KS, Bandyra KJ, Bruce HA, Hohensee S, Vogel J, Luisi BF. 2014. Recognition of the small regulatory RNA RydC by the bacterial Hfq protein *eLife* **3**: e05375.
- Durieux I, Ginevra C, Attaiech L, Pick K, Juan P-A, Jarraud S, Charpentier X. 2019. Diverse conjugative elements silence natural transformation in *Legionella* species *Proceedings of the National Academy of Sciences* **116**: 18613-18618.
- El Mouali Y, Gerovac M, Mineikaitė R, Vogel J. 2021a. In vivo targets of *Salmonella* FinO include a FinP-like small RNA controlling copy number of a cohabitating plasmid. *Nucleic Acids Research* **49**: 5319-5335.
- El Mouali Y, Ponath F, Scharrer V, Wenner N, Hinton JCD, Vogel J. 2021b. Scanning mutagenesis of RNA-binding protein ProQ reveals a quality control role for the Lon protease. *RNA* **27**: 1512-1527.
- Fozo EM, Hemm MR, Storz G. 2008. Small Toxic Proteins and the Antisense RNAs That Repress Them. *Microbiology and Molecular Biology Reviews* **72**: 579-589.
- Franze de Fernandez MT, Eoyang L, August JT. 1968. Factor fraction required for the synthesis of bacteriophage Qbeta-RNA *Nature* **219**: 588-590.
- Frost LS, Ippen-Ihler K, Skurray RA. 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiological Reviews* **58**: 162-210.
- Gerovac M, el Mouali Y, Kuper J, Kisker C, Barquist L, Vogel J. 2020. Global discovery of bacterial RNA-binding proteins by RNase-sensitive gradient profiles reports a new FinO domain protein. *RNA* **26**: 1448-1463.
- Ghetu AF, Arthur DC, K. KT, Mark GJN. 2002. Probing FinO-FinP RNA interactions by site-directed protein-RNA crosslinking and gelFRET. *RNA*: 816-823.
- Ghetu AF, Gubbins MJ, Frost LS, Glover JNM. 2000. Crystal structure of the bacterial conjugation repressor FinO. *Nature Structural Biology* **7**: 565-569.
- Glover JNM, Chaulk SG, Edwards RA, Arthur D, Lu J, Frost LS. 2015. The FinO family of bacterial RNA chaperones. *Plasmid* **78**: 79-87.

- Gonzalez GM, Hardwick SW, Maslen SL, Skehel JM, Holmqvist E, Vogel J, Bateman A, Luisi BF, Broadhurst RW. 2017. Structure of the Escherichia coli ProQ RNA-binding protein. *RNA* **23**: 696-711.
- Holmqvist E, Li L, Bischler T, Barquist L, Vogel J. 2018. Global Maps of ProQ Binding In Vivo Reveal Target Recognition via RNA Structure and Stability Control at mRNA 3' Ends. *Molecular Cell* **70**: 971-982.e976.
- Holmqvist E, Vogel J. 2018. RNA-binding proteins in bacteria. *Nature Reviews Microbiology* **16**: 601-615.
- Holmqvist E, Wright PR, Li L, Bischler T, Barquist L, Reinhardt R, Backofen R, Vogel J. 2016. Global sRNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. *The EMBO Journal* **35**: 991-1011.
- Immer C, Hacker C, Wöhnert J. 2020. Solution structure and RNA-binding of a minimal ProQ-homolog from Legionella pneumophila (Lpp1663). *RNA* **26**: 2031-2043.
- Ishikawa H, Otaka H, Maki K, Morita T, Aiba H. 2012. The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3' poly(U) tail. *RNA* **18**: 1062-1074.
- Jerome LJ, Frost LS. 1999. In Vitro Analysis of the Interaction between the FinO Protein and FinP Antisense RNA of F-like Conjugative Plasmids. *Journal of Biological Chemistry* **274**: 10356-10362.
- Jerome LJ, van Biesen T, Frost LS. 1999. Degradation of FinP antisense RNA from F-like plasmids: the RNA-binding protein, FinO, protects FinP from ribonuclease E. *Journal of Molecular Biology* **285**: 1457-1473.
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A i wsp. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**: 583-589.
- Kawano M, Reynolds AA, Miranda-Rios J, Storz G. 2005. Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in Escherichia coli *Nucleic Acids Research* **33**: 1040-1050.
- Kerr CH, Culham DE, Marom D, Wood JM. 2014. Salinity-dependent impacts of ProQ, Prc, and Spr deficiencies on Escherichia coli cell structure. *Journal of Bacteriology* **196**: 1286-1296.
- Kim HJ, Black M, Edwards RA, Peillard-Fiorente F, Panigrahi R, Klingler D, Eidelpes R, Zeindl R, Peng S, Su J, Omar AR, MacMillan AM, Kreutz C, Tollinger M, Charpentier X, Attaiech L, Glover JNM. 2022. Structural basis for recognition of transcriptional terminator structures by ProQ/FinO domain RNA chaperones. *Nature Communications* **13**: 7076.

- Kunte HJ, Crane RA, Culham DE, Richmond D, Wood JM. 1999. Protein ProQ Influences Osmotic Activation of Compatible Solute Transporter ProP in Escherichia coli K-12. *Journal of Bacteriology* **181**: 1537-1543.
- Kwiatkowska J, Wroblewska Z, Johnson KA, Olejniczak M. 2018. The binding of Class II sRNA MgrR to two different sites on matchmaker protein Hfq enables efficient competition for Hfq and annealing to regulated mRNAs. *RNA* **24**: 1761-1784.
- Laine B, Kmiecik D, Sautiere P, Biserte G, Cohen-Solal M. 1980. Complete amino-acid sequences of DNA-binding proteins HU-1 and HU-2 from Escherichia coli *European Journal of Biochemistry* **103**: 447-461.
- Lee T, Feig AL. 2008. The RNA binding protein Hfq interacts specifically with tRNAs *RNA* **14**: 514-523.
- Link TM, Valentin-Hansen P, Brennan RG. 2009. Structure of Escherichia coli Hfq bound to polyriboadenylate RNA *Proceedings of the National Academy of Sciences* **106**: 19292-19297.
- Małeczka EM, Stróżecka J, Sobańska D, Olejniczak M. 2015. Structure of Bacterial Regulatory RNAs Determines Their Performance in Competition for the Chaperone Protein Hfq. *Biochemistry* **54**: 1157-1170.
- Melamed S, Adams PP, Zhang A, Zhang H, Storz G. 2020. RNA-RNA Interactomes of ProQ and Hfq Reveal Overlapping and Competing Roles. *Molecular Cell* **77**: 411-425.e417.
- Mikulecky PJ, Kaw MK, Brescia CC, Takach JC, Sledjeski DD, Feig AL. 2004. Escherichia coli Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nature Structural & Molecular Biology* **11**: 1206-1214.
- Milner JL, Wood JM. 1989. Insertion proQ220::Tn5 Alters Regulation of Proline Porter II, a Transporter of Proline and Glycine Betaine in Escherichia coli. *Journal of Bacteriology* **171**: 947-951.
- Miyakoshi M, Chao Y, Vogel J. 2015. Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *The EMBO Journal* **34**: 1478-1492.
- Mohanty BK, Maples VF, Kushner SR. 2004. The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. *Molecular Microbiology* **54**: 905-920.
- Morita T, Maki K, Aiba H. 2005. RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs *Genes & Development* **19**: 2176-2186.
- Novick RP, Geisinger E. 2008. Quorum sensing in staphylococci *Annual Review of Genetics* **42**: 541-564.

- Olejniczak M, Storz G. 2017. ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers? *Molecular Microbiology* **104**: 905-915.
- Otaka H, Ishikawa H, Morita T, Aiba H. 2011. PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proceedings of the National Academy of Sciences* **108**: 13059-13064.
- Pandey S, Gravel CM, Stockert OM, Wang CD, Hegner CL, Leblanc H, Berry KE. 2020. Genetic identification of the functional surface for RNA binding by *Escherichia coli* ProQ. *Nucleic Acids Research* **48**: 4507-4520.
- Papenfert K, Vogel J. 2010. Regulatory RNA in Bacterial Pathogens. *Cell Host & Microbe* **8**: 116-127.
- Ren G-X, Guo X-P, Sun Y-C. 2017. Regulatory 3' Untranslated Regions of Bacterial mRNAs. *Frontiers in Microbiology* **8**:1276.
- Richards J, Belasco JG. 2021. Riboswitch control of bacterial RNA stability. *Molecular Microbiology* **116**: 361-365.
- Rizvanovic A, Kjellin J, Söderbom F, Holmqvist E. 2021. Saturation mutagenesis charts the functional landscape of *Salmonella* ProQ and reveals a gene regulatory function of its C-terminal domain. *Nucleic Acids Research* **49**: 9992-10006.
- Robinson KE, Orans J, Kovach AR, Link TM, Brennan RG. 2014. Mapping Hfq-RNA interaction surfaces using tryptophan fluorescence quenching *Nucleic Acids Research* **42**: 2736-2749.
- Santiago-Frangos A, Kavita K, Schu DJ, Gottesman S, Woodson SA. 2016. C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA. *Proceedings of the National Academy of Sciences* **113**: E6089-E6096.
- Santiago-Frangos A, Woodson SA. 2018. Hfq chaperone brings speed dating to bacterial sRNA. *WIREs RNA* **9**: e1475.
- Sauer E, Weichenrieder O. 2011. Structural basis for RNA 3'-end recognition by Hfq. *Proceedings of the National Academy of Sciences* **108**: 13065-13070.
- Schu DJ, Zhang A, Gottesman S, Storz G. 2015. Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition. *The EMBO Journal* **34**: 2557-2573.
- Sheidy DT, Zielke RA. 2013. Analysis and Expansion of the Role of the *Escherichia coli* Protein ProQ. *PLoS ONE* **8**: e79656.
- Smirnov A, Förstner KU, Holmqvist E, Otto A, Günster R, Becher D, Reinhardt R, Vogel J. 2016. Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proceedings of the National Academy of Sciences* **113**: 11591-11596.
- Smirnov A, Wang C, Drewry LL, Vogel J. 2017. Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA. *The EMBO Journal* **36**: 1029-1045.

- Smith MN, Crane RA, Keates RAB, Wood JM. 2004. Overexpression, Purification, and Characterization of ProQ, a Posttranslational Regulator for Osmoregulatory Transporter ProP of Escherichia coli. *Biochemistry* **43**: 12979-12989.
- Smith MN, Kwok SC, Hodges RS, Wood JM. 2007. Structural and Functional Analysis of ProQ: An Osmoregulatory Protein of Escherichia coli. *Biochemistry* **46**: 3084-3095.
- Soper T, Mandin P, Majdalani N, Gottesman S, Woodson SA. 2010. Positive regulation by small RNAs and the role of Hfq. *Proceedings of the National Academy of Sciences* **107**: 9602-9607.
- Soper TJ, Woodson SA. 2008. The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA* **14**: 1907-1917.
- Stalmach ME, Grothe S, Wood JM. 1983. Two Proline Porters in Escherichia coli K-12. *Journal of Bacteriology* **156**: 481-486.
- Stein EM, Kwiatkowska J, Basczok MM, Gravel CM, Berry KE, Olejniczak M. 2020. Determinants of RNA recognition by the FinO domain of the Escherichia coli ProQ protein. *Nucleic Acids Research* **48**: 7502-7519.
- Storz G, Vogel J, Wassarman KM. 2011. Regulation by Small RNAs in Bacteria: Expanding Frontiers. *Molecular Cell* **43**: 880-891.
- Sukhodolets MV, Garges S. 2003. Interaction of Escherichia coli RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry* **42**: 8022-8034.
- Udekwi KI, Darfeuille F, Vogel J, Reimegård J, Holmqvist E, Wagner EGH. 2005. Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes & Development* **19**: 2355-2366.
- Updegrove TB, Zhang A, Storz G. 2016. Hfq: the flexible RNA matchmaker. *Current Opinion in Microbiology* **30**: 133-138.
- Vogel J, Luisi BF. 2011. Hfq and its constellation of RNA. *Nature Reviews Microbiology* **9**: 578-589.
- Vogel J, Papenfort K. 2006. Small non-coding RNAs and the bacterial outer membrane. *Current Opinion in Microbiology* **9**: 605-611.
- Wadler CS, Vanderpool CK. 2007. A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proceedings of the National Academy of Sciences* **104**: 20454-20459.
- Wagner EGH, Romby P. 2015. Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Advances in Genetics* **90**: 133-208.
- Waters LS, Storz G. 2009. Regulatory RNAs in bacteria. *Cell* **134**: 615-628.

Westermann AJ, Venturini E, Sellin ME, Förstner KU, Hardt W-D, Vogel J. 2019. The Major RNA-Binding Protein ProQ Impacts Virulence Gene Expression in *Salmonella enterica* Serovar Typhimurium *mBio* **10**: e02504-18.

Wilusz CJ, Wilusz J. 2012. Lsm proteins and Hfq: Life at the 3' end *RNA Biology* **10**: 592-601.