

Uniwersytet Warszawski



Adam Mamot

Numer albumu: 345894

Synteza podwójnie znakowanego oraz kolistego RNA: narzędzia do bioobrazowania i terapii genowej

Praca doktorska

Praca wykonana pod kierunkiem

Prof. Jacka Jemielitego

w Zakładzie Biofizyki

Instytutu Fizyki Doświadczalnej

Wydziału Fizyki Uniwersytetu Warszawskiego

Warszawa 2023

Streszczenie: W ramach niniejszej rozprawy zaprezentowano wyniki wieloletnich badań nad otrzymywaniem i zastosowaniem modyfikowanych kwasów rybonukleinowych (RNA). Celem badań było opracowanie możliwie prostych i skutecznych metod znakowania oraz cyklizacji RNA. Otrzymane tymi metodami cząsteczki (analogi RNA) miały posłużyć jako narzędzia do obrazowania procesów dostarczania, ekspresji i rozpadu RNA w układach biologicznych oraz jako nośniki terapeutycznej informacji genetycznej. W tym celu, zastosowano oddolną strategię badawczą oraz szereg technik eksperymentalnych (metod syntezy organicznej i biochemicznej, analiz chromatograficznych, elektroforetycznych, spektrometrycznych i spektroskopowych, metod transkrypcji i translacji *in vitro*) z wykorzystaniem mało- oraz wielkocząsteczkowych modeli RNA. Na podstawie wstępnych wyników badań opracowano metodę bezpośredniej, chemicznej modyfikacji struktury końca 3' RNA oraz udoskonalono metodę chemoenzymatycznej modyfikacji struktury końca 5' RNA. W dalszym etapie badań, metody modyfikacji końców 5' oraz 3' RNA zostały połączone, w wyniku czego opracowano metody otrzymywania podwójnie znakowanego oraz kolistego RNA (circRNA). Podwójnie znakowane RNA zostały wykorzystane jako sondy molekularne, których zastosowanie zademonstrowano na przykładzie badań dynamiki strukturalnej wybranych RNA czy aktywności enzymów nukleolitycznych. Wraz ze współpracownikami, zbadano biologiczne właściwości podwójnie znakowanego informacyjnego RNA (mRNA) oraz wykonano obrazowanie dostarczania, lokalizacji i ekspresji mRNA w komórkach oraz w *Danio* pręgowanym (*Danio rerio*). Ponadto (wraz ze współpracownikami), otrzymano i zbadano właściwości biologiczne szeregu analogów circRNA. Wyniki tych badań wskazują na pozytywny wpływ modyfikacji circRNA (np. endocyklicznej struktury kapu) na poziomy produkowanego białka oraz wydłużone czasy półtrwania (względem liniowych form RNA). Z tego powodu, otrzymane opracowanymi metodami analogi circRNA wykazują potencjał jako nośniki terapeutycznej informacji genetycznej.

Projekt został zrealizowany w ramach programów TEAM (TEAM POIR.04.04.00-00-20A2/16) oraz START (START 044.2022) Fundacji na rzecz Nauki Polskiej współfinansowanego przez Unię Europejską z Europejskiego Funduszu Rozwoju Regionalnego.

Składam serdeczne podziękowania za wsparcie finansowe w ramach projektu „DESCARTES” współfinansowanego ze środków Unii Europejskiej w ramach Programu Operacyjnego Wiedza Edukacja Rozwój (POWR.03.02.00-00-I013/17-00).

Abstract: Herein are presented the results of several years of research on preparation and application of modified ribonucleic acids (RNA). The aim of the research was to develop simple and effective methods of labelling and circularization of RNA. The molecules obtained by these methods (RNA analogues) were to serve as tools for imaging of RNA delivery, expression and decay in biological systems, and as carriers of therapeutic genetic information. To this end, a bottom-up research strategy and a range of experimental techniques (organic and biochemical synthesis methods, chromatographic, electrophoretic, spectrometric and spectroscopic analyses, in vitro transcription and translation methods) were employed, using both low and high molecular weight RNA models. As a result of the preliminary research, a direct chemical method modification of the RNA 3' end was developed, as well as a chemoenzymatic method of the RNA 5' end modification was refined. During the further stage of research, both methods of RNA modification (for the 5' and 3' end structures) were combined, which led to the development of methods for obtaining dually labelled RNA and circular RNA (circRNA). The dually labelled RNAs were used as molecular probes, that was demonstrated in studies of structural dynamics of RNA, or the activity of several nucleolytic enzymes. Together with collaborators, the biological properties of dually labelled messenger RNA (mRNA) were investigated, as well as imaging of mRNA delivery, localization, and expression in cells and in zebrafish (*Danio rerio*). In addition (together with collaborators), a series of circRNA analogues were obtained and their biological properties were studied. The results of these studies indicate positive effects of circRNA modifications (e.g. endocyclic cap structure) on the levels of the produced proteins, and prolonged half-life (relative to linear forms of RNA). Therefore, the circRNA analogues, obtained by the developed methods, show potential as carriers of therapeutic genetic information.

This project was carried out within the *TEAM POIR.04.04.00-00- 20A2/16* and *START 044.2022* programmes of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund.

Financial support from the „*DESCARTES*” project co-financed by the European Union under the „Program Operacyjny Wiedza Edukacja Rozwój” (*POWR.03.02.00-00-I013/17-00*) is gratefully acknowledged.

Składam najserdeczniejsze podziękowania:

Prof. Jackowi Jemielitemu – za możliwość rozwoju, mentorstwo i cenne nauki na przestrzeni ostatnich dziesięciu lat;

Dr. hab. Joannie Kowalskiej – za wsparcie merytoryczne, cierpliwość i możliwości rozwoju na przestrzeni ostatnich dziesięciu lat;

Dr. Pawłowi Sikorskiemu – za lata nauk i współpracy na najwyższym poziomie;

Dr. Małgorzacie Wąsińskiej-Kałwie – za najlepszą dynamikę współpracy, koleżeństwo, cierpliwość, ogromny wysiłek włożony w realizację wspólnego projektu oraz cenne uwagi dotyczące rozprawy;

Dr. Aleksandrze Siekierskiej – za owocną współpracę nad bioobrazowaniem mRNA w układach *in vivo*;

Dr. Karolowi Czubakowi – za obiecującą współpracę nad translacją analogów kolistego RNA;

Dr. Karolinie Drażkowskiej – za lekcje technik biochemicznych oraz cenną, krytyczną perspektywę (nie tylko w kwestii rozprawy);

Dr. Mikołajowi Chromińskiemu – za lekcje i opiekę w laboratorium oraz za cenne uwagi dotyczące rozprawy;

Koleżankom i kolegom z Laboratorium Chemii Bioorganicznej – za wspólną pracę, naukę, zabawę i przygody;

Dr. hab. Piotrowi Wasylczykowi, prof. UW – za bezcenne lekcje prezentacji wyników badań naukowych;

Żonie Paulinie, najbliższej rodzinie i przyjaciółom – za ocean miłości i wsparcia.

Spis treści:

Spis skrótów.....	8
Część I: Stan wiedzy i techniki.....	10
1. RNA jako biocząsteczka.....	11
1. 1. Struktura i właściwości RNA.....	11
1. 2 Wybrane funkcje i właściwości biologiczne RNA	18
2. RNA jako narzędzie.....	33
2.1 Terapie genowe	33
2.2 Bioobrazowanie i sondy molekularne RNA.....	45
2.3 Metody modyfikacji kodującego RNA	51
Część II: Cel i wyniki badań.....	63
Cel badań	64
1. Modyfikacja końca 3' RNA	66
1. 1. Założenia i strategia badań.....	66
1. 2. Przebieg aminowania redukcyjnego w zależności od doboru <i>N</i> -nukleofila	69
1. 3. Modyfikacja RNA za pomocą analogów EDA	75
2. Podwójne znakowanie RNA.....	88
2. 1. Kotranskrypcyjna modyfikacja końca 5' i reakcja SPAAC	88
2. 2. Sprzężenie reakcji SPAAC i metody PORA.....	93
2. 3. Efekt „kotwiczenia”	97
3. Sondy molekularne RNA.....	101
3. 1. Podwójnie znakowane RNA jako sondy molekularne.....	101
3. 2. Badania wykonane ze współpracownikami: Bioobrazowanie mRNA	111
4. Cyklizacja RNA	118
4. 1. Cel i metoda otrzymywania analogów circRNA.....	118
4. 2. Metodologia badań analogów circRNA.....	119
4. 3. Badania wykonane ze współpracownikami: optymalizacja cyklizacji RNA.....	124
4. 4. Badania wykonane ze współpracownikami: właściwości biologiczne analogów circRNA	130
5. Podsumowanie i wnioski	134
Część III: Procedury doświadczalne	138
Sprzęt i odczynniki	139
Syntezy chemiczne.....	140
Karbaminian (2-bromoetylo) tert-butyłowy	140
Karbaminian (N-propargiloaminoetylo) tert-butyłowy.....	140

Dichlorowodorek N-propargiloetylenodiaminy (PEDA).....	140
2-azydoetyloamina	141
Karbaminian [N-(2-azydoetylo)aminoetylo] tert-butyłowy	141
Dichlorowodorek N-(2-azydoetylo)etylenodiaminy (AEEDA).....	141
Biot-EDA	142
Cy3-EDA oraz Cy5-EDA.....	142
FAM-EDA.....	143
pHrodo-EDA	143
DBCO-EDA	144
GMP-dial.....	144
pU ₃	145
N ₃ -GDP	145
N ₃ -m ⁷ GDP	146
N ₃ -m ⁷ GpppAG	146
EDA-PEG ₅ -m ⁷ GpppAG	147
EDA-m ⁷ GpppG	148
N ₃ -AG.....	149
EDA-AG.....	150
EDA-PEG ₅ -AG	150
Analizy chemiczne.....	151
Reakcje między GMP-dial a N-nukleofilami.....	151
Optymalizacja protokołu PORA w jednym naczyniu	151
Transkrypcja <i>in vitro</i>	153
Ogólna procedura	153
Znakowanie RNA	153
Znakowanie końca 3' mRNA	153
Znakowanie końca 5' mRNA (Cy5).....	154
Podwójne znakowanie mRNA (5' Cy5 i 3' Cy3)	154
Izolacja znakowanego mRNA.....	154
Cyklizacja RNA	154
Ogólna procedura cyklizacji chemicznej (PORA).....	154
Ogólna procedura cyklizacji enzymatycznej (ligaza RNA I T4)	155
Trawienie RNazą R	156
Test endonukleolityczny z RNazą H.....	156
Elektroforeza circRNA na żelu poliakrylamidowym (PAGE).....	156

Sondy molekularne RNA	156
Sondy FRET	156
Sonda pH	158
Translacja <i>in vitro</i>	158
Translacja znakowanego mRNA	158
Translacja kolistego RNA	158
Bibliografia	160
Lista publikacji i zgłoszeń patentowych autora	171

Spis skrótów:

AA – akrylamid
AcOH – kwas octowy
AEEDA – N-(2-azydoetylo)etylenodiamina
APS – nadsiarczan amonu
ATP – 5'-trifosforan adenozyne
circRNA – kolisty kwas rybonukleinowy
CRISPR – ang. *clustered regularly interspaced short palindromic repeats*, system obrony organizmów prokariotycznych (bakterii i archeonów) przed egzogennymi elementami genetycznym
CTP – 5'-trifosforan cytydyny
CuAAC – reakcja azydkowo-alkinowej cykloaddycji katalizowanej jonami miedzi
Cy3, Cy5 – cyjanina trzecia lub piąta (barwniki fluorescencyjne)
DBCO – dibenzocyklooktyn
DETA – dietylenotriamina
DMF – N,N-dimetyloformamid
DMSO – dimetylosulfotlenek
DNA – kwas deoksyrybonukleinowy
dsDNA, dsRNA – dwuniciowy kwas deoksy- lub rybonukleinowy
EDA – etylenodiamina
EMSA – ang. *electrophoretic mobility shift assay*, metoda elektroforetyczna
EtBr – bromek etydyny
EtOH – etanol
FAM – fluoresceina (barwnik fluorescencyjny)
FISH – ang. *fluorescent in situ hybridization*, metoda detekcji kwasów nukleinowych
FRET – rezonansowy transfer energii Förstera
GMP – 5'-monofosforan guanozyny
GTP – 5'-trifosforan guanozyny
HA – n-heksyloamina
HAA – octan n-heksyloaminy
iEDDA – ang. *inverse electron demand Diels–Alder reaction*, reakcja Dielsa-Aldera z odwróconym zapotrzebowaniem elektronowym
iPrOH – 2-propanol, izopropanol
IRES – ang. *internal ribosomal entry site*, sekwencja RNA promująca wiązanie czynników inicjujących translację
IVT – transkrypcja in vitro
lncRNA – długi niekodujący kwas rybonukleinowy
LNPs – ang. *lipid nanoparticles*, nanocząstki lipidowe
MBAA – N,N'-metylenobisakrylamid
MeCN – acetonitryl
MeOH – metanol
MFE – ang. *minimum free energy*, minimum potencjału Gibbsa
miRNA – mikro kwas rybonukleinowy
mRNA – informacyjny kwas rybonukleinowy
MS – spektrometria mas
ncRNA – niekodujący kwas rybonukleinowy
NHS – N-hydroksysukcynoimid
NMP – 5'- monofosforan rybonukleozydu

NTP – 5'-trifosforan rybonukleozydu
ORF – ang. *open reading frame*, sekwencja otwartej ramki odczytu (od kodonu start do kodonu stop)
PAGE – elektroforeza na żelu poliakrylamidowym
PEDA – N-(propargilo)etylenodiamina
PIE – ang. *permuted intorn exon system*, metoda cyklizacji RNA
poliA – RNA lub jego fragment, złożony z sekwencji poliadeniny
PORA – ang. *periodate oxidation and reductive amination*, metoda modyfikacji końca 3' RNA
RNA – kwas rybonukleinowy
RP-HPLC – wysokosprawna chromatografia cieczowa w układzie faz odwróconych
rRNA – rybosomalny kwas rybonukleinowy
SAM – S-adenozylometionina
siRNA – krótki interferujący kwas rybonukleinowy
snoRNA – mały jądrowy kwas rybonukleinowy
snRNA – mały jądrowy kwas rybonukleinowy
SPAAC – reakcja azydowo-alkinowej cykloaddycji katalizowanej naprężeniem pierścienia
ssDNA, ssRNA – jednoniciowy kwas deoksy- lub rybonukleinowy
TEA – trietyloamina
TEAA – octan trietyloaminy
TEMED – N,N,N',N'-tetrametyloetylenodiamina
THPTA – Tris(benzyltriazolylmethyl)amina
tRNA – transferowy kwas rybonukleinowy
TSS – ang. *transcription start site*, miejsce rozpoczęcia transkrypcji
UTP – 5'-trifosforan urydyny
UTR – ang. *untranslated region*, sekwencja mRNA nieulegająca translacji

Część I: Stan wiedzy i techniki

1. RNA jako biocząsteczka

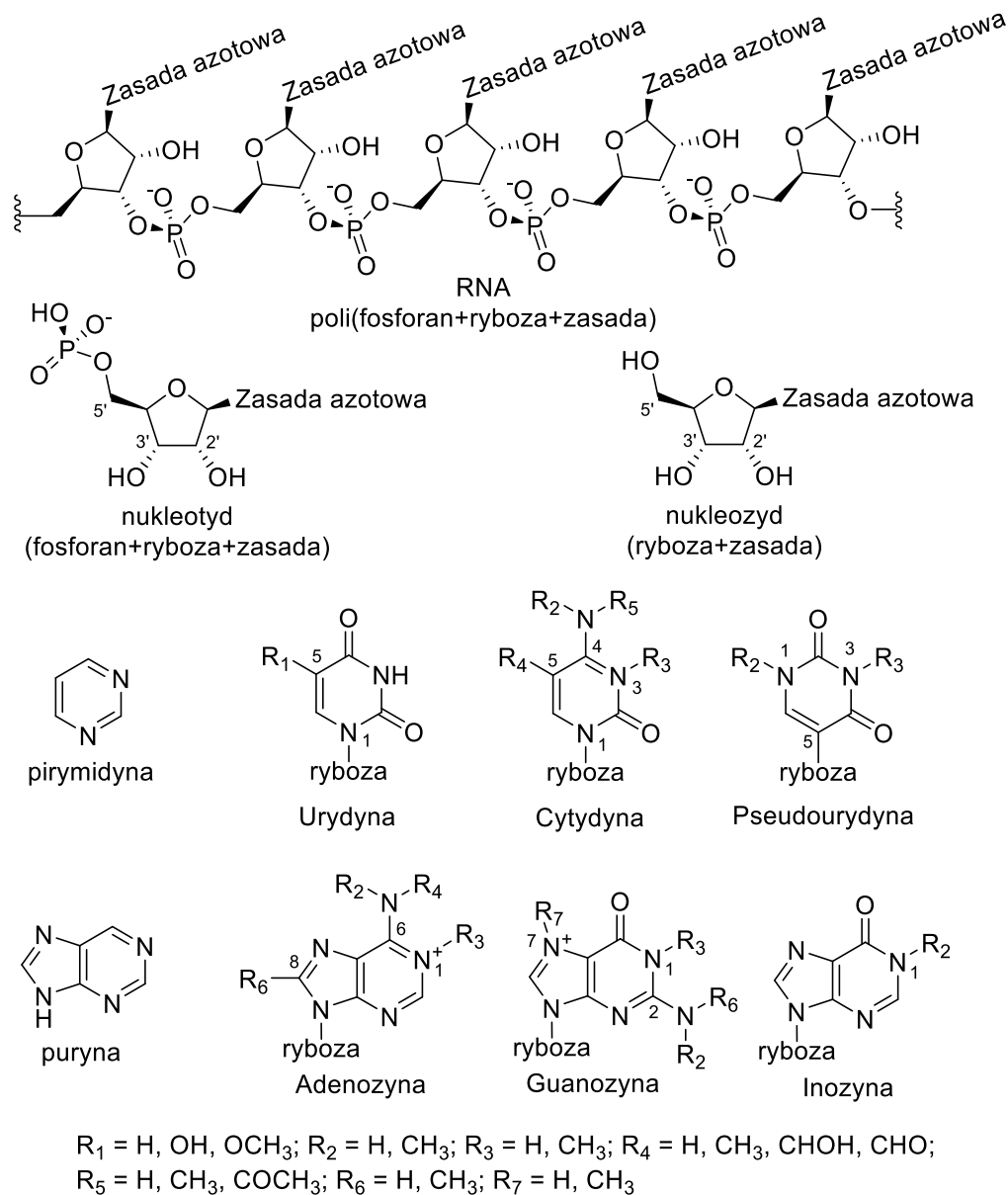
1. 1. Struktura i właściwości RNA

RNA jest naturalnie występującą substancją oraz cennym narzędziem wykorzystywanym w badaniach i terapii. Właściwości, funkcje i zastosowania RNA wynikają z jego struktury cząsteczkowej. RNA jest biopolimerem z rodziny kwasów nukleinowych. Biopolimery są to duże cząsteczki o budowie łańcuchowej złożone z podjednostek (monomerów) naturalnego pochodzenia. Kwas nukleinowy to biopolimer o budowie liniowej którego monomerami są nukleotydy. Najczęściej są to polimery liniowe (Rys. 1.1.1), ale również mogą one przybierać strukturę rozgałęzioną lub kolistą. Nukleotydy składają się z trzech elementów: reszty cukrowej, fosforanowej oraz zasady azotowej. Cukier i fosforan, połączone wiązaniem fosfodiesterowym, tworzą łańcuch główny, a zasady azotowe stanowią reszty boczne, połączone z cukrem wiązaniem N-glikozydowym. Każdy z trzech elementów nukleotydu ma specyficzne właściwości, które konstytuują strukturę przestrzenną i właściwości biopolimeru jako całości.

Struktura chemiczna reszty cukrowej definiuje rodzaj kwasu nukleinowego: ryboza znajduje się w kwasie rybonukleinowym czyli RNA, natomiast deoksyryboza – w kwasie deoksyrybonukleinowym czyli DNA. Reszty te różnicuje podstawnik przy atomie węgla 2', w RNA jest to OH, w DNA jest to H, zatem jedynie jeden atom tlenu na nukleotyd. Obecność grupy 2' hydroksylowej rybozy umożliwia tworzenie wiązania wodorowego, co jest istotne z perspektywy hydrolizy wiązania fosfodiesterowego. Ponadto, pierścień rybozy może przyjąć wiele konformacji (dwie główne to konformacja N i S), co ma znaczenie z perspektywy kinetyki hydrolizy łańcucha RNA (Rys 1.1.2).^[1]

Kolejnym elementem jest reszta fosforanowa. Jedną z najistotniejszych jej cech jest jonizowalność: w warunkach fizjologicznych (pH cytozolu ~7.2, wewnątrz komórki i organelli w zakresie 4–8),^[2] w jest ujemnie naładowana ($pK_a \sim 1$). Dzięki temu kwasy nukleinowe są rozpuszczalne w wodzie, a atomy tlenu mają silnie nukleofilowy charakter. Kolejną jej istotną cechą jest obecność atomu fosforu na piątym stopniu utlenienia (P(V)) o charakterystycznej reaktywności. Atom P(V) ma charakter elektrofilowy oraz indukowalną reaktywność (np. pod wpływem koordynacji kationów metali przez atomy tlenu grupy fosforanowej), co uznaje się powód „sukcesu” grupy fosforanowej w naturalnych procesach biochemicznych.^[3] Przykładowo, reaktywność P(V) jest kluczowa w syntezie i hydrolizie wysokoenergetycznych bezwodników fosforanowych obecnych w trifosforanach rybonukleozydów (NTP). Dzięki temu NTP są zarówno materiałem jak i

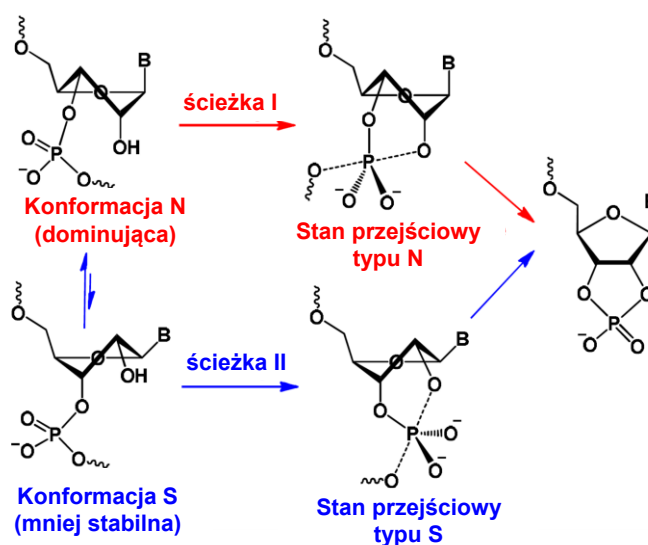
paliwem wykorzystywanym w naturalnym procesie syntezy łańcucha RNA czyli transkrypcji (omówiono w Rozdziale 1.2.1). Fosforan to grupa kluczowa także dla hydrolizy łańcucha głównego RNA, podczas której reaguje z 2' OH pod wpływem konformacji rybozy czy interakcji z jonami metali dwuwartościowych.^[1, 3]



Rysunek 1.1.1: Struktura chemiczna RNA. Struktury zasad azotowych nukleozydów na podstawie ^[4]. Dla kanonicznych zasad azotowych (występujących w adenozyinie, guanozyinie, cytydynie i urydynie) podstawniki R₁–R₇ = H.

Ostatnim elementem są zasady azotowe, które należą do grupy heterocyklicznych związków aromatycznych opartych na architekturze puryny i pirymidyny (Rys. 1.1.1). W ramach naturalnie występujących cząsteczek RNA wyróżniamy szereg zasad azotowych, które historycznie dzielone są na „kanoniczne” (Adenia, Uracyl, Guanina, Cytozyna) oraz „niekanoniczne”.^[4, 5] Nazwa „zasady azotowe” pochodzi do zawartych w nich atomów

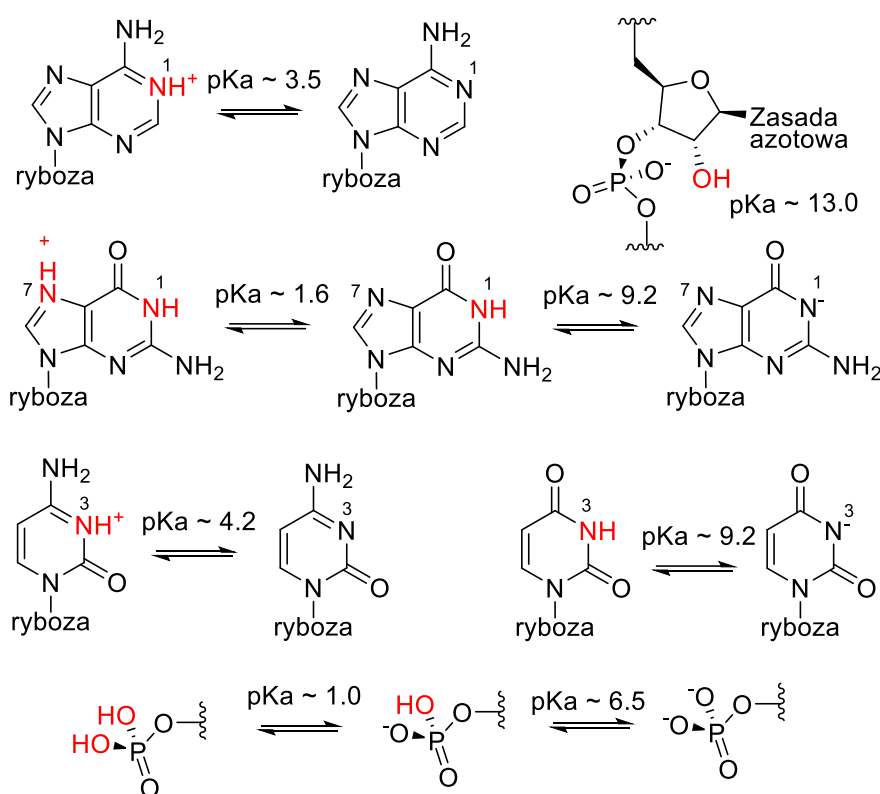
azotu (endo- i egzocykliczne) oraz możliwości ich protonowania (tj. ich zdolności do koordynacji kationu wodoru) w kwaśnym środowisku. W warunkach fizjologicznych kanoniczne zasady azotowe są obojętne elektrycznie (A i C) lub obdarzone częściowym ładunkiem dodatnim (G i U) (Rys.1.1.3).^[6] W ogólności, zasady azotowe występujące w naturalnych RNA posiadają jonizowalne grupy, zawierające zarówno atomy azotu jak i tlenu, o pK_a w bardzo szerokim zakresie (co najmniej 2–12).^[4] Powoduje to, że dana zasada azotowa, w zależności od warunków oraz struktury, może być zarówno obojętna elektrycznie, naładowana dodatnio lub ujemnie. Ponadto, wyżej wspomniane atomy azotu i tlenu stanowią platformę do tworzenia wiązań wodorowych, które są istotne między innymi w przenoszeniu informacji genetycznej (Rozdział 1.2.1 i 1.2.2). W przeciwieństwie do składników łańcucha głównego, zasady azotowe są hydrofobowe, co nadaje kwasom nukleinowym charakter amfifilowy.



Rysunek 1.1.2: Hydroliza wiązania fosfodiesterowego RNA z udziałem grupy 2' OH: wpływ konformacji pierścienia rybozy na kinetykę i stany równowagowe reakcji. B oznacza zasadę azotową. Na podstawie ^[1].

Struktura przestrzenna oraz jej dynamika mają bezpośrednie przełożenie na właściwości i funkcje biologiczne RNA.^[7, 8] Struktura i właściwości chemiczne nukleotydów powodują, że struktura przestrzenna RNA jest skomplikowana i zasadniczo nigdy nie przybiera formy kłębka statystycznego.^[5, 9] Dzieje się tak, ponieważ monomery RNA są wysoce „towarzyskie”: mogą brać udział w szeregu rozmaitych oddziaływań (wewnątrz- i międzycząsteczkowych) za pomocą każdej ze swoich trzech części.^[5] Hydrofobowe zasady azotowe ciasno przylegają do siebie za sprawą oddziaływań van der Waalsa, w szczególności nakładania orbitali π (ang. *stacking*), a polarny łańcuch ulega

solwatacji. Jednocześnie łańcuch stanowi 2/3 masy polimeru i jest jednoimiennie naładowany, co nadaje ograniczenia elektrostatyczne i konformacyjne całej cząsteczce RNA (Rys.1.1.4).^[5]



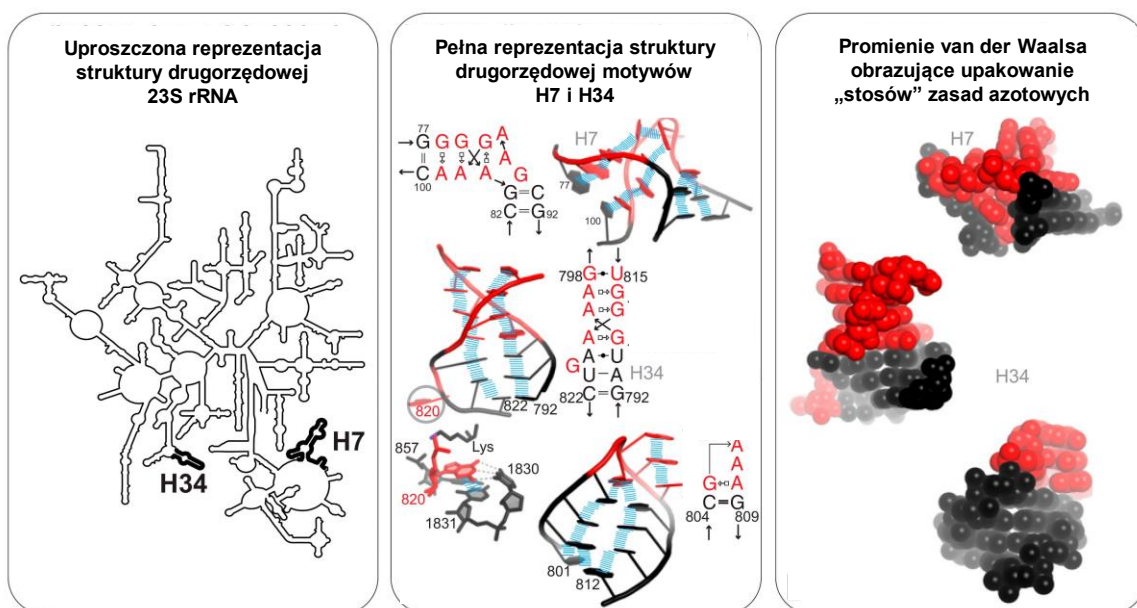
Rysunek 1.1.3: Jonizacja elementów składowych kanonicznych rybonukleotydów. Na podstawie ^[6].

Oddziaływania hydrofobowe, w szczególności *stacking* są głównym czynnikiem promującym heliakalną strukturę drugorzędową RNA. Innymi słowy, charakterystyczna heliakalna struktura kwasów nukleinowych bardziej przypomina „stos monet” niż „skręconą drabinkę”.^[5] Jednym z dowodów na to jest struktura krystaliczna poliadeniny (A₁₁), której cząsteczki tworzą podwójną helisę w pH 7.^[10] W mniejszym stopniu, struktura drugorzędowa RNA jest również promowana przez tworzenie wiązań wodorowych między parami zasad azotowych. Często, podkreśla się rolę kanonicznych par zasad Watsona-Cricka (pary A:T, A:U oraz G:C), co w wypadku RNA jest uproszczeniem graniczącym z trywializacją. W rzeczywistości, parowanie W-C reprezentuje zaledwie jeden z dwunastu typów interakcji nukleotydów i zasad azotowych RNA.^[11] Typy owych interakcji można klasyfikować pod kątem wzajemnej orientacji wiązań N-glikozydowych, oddziałujących krawędzi oraz lokalnej orientacji nici. (Rys.1.1.5) Różnorodność interakcji nukleotydów, połączona z wpływem posttranskrypcyjnych modyfikacji (modyfikacji chemicznych którym podlegają nukleotydy po procesie transkrypcji) oraz ich charakterem

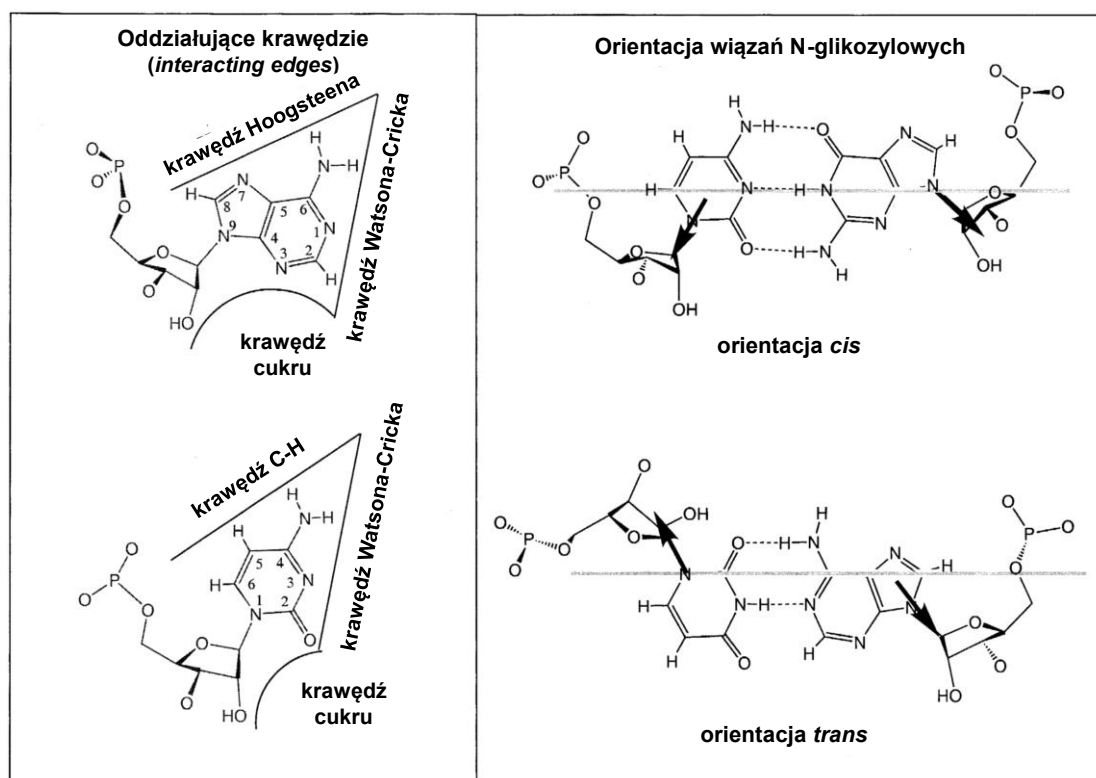
równowagowym, ma kluczowe znaczenie z perspektywy tworzenia specyficznych motywów strukturalnych i funkcji biologicznych RNA.^[12]

Struktura przestrzenna RNA jest dynamiczna i może się zmienić (spontanicznie lub pod wpływem bodźców) w skali od nanosekund do milisekund.^[8] Z tego powodu lepiej mówić o populacjach konformerów RNA i o tworzonej przez nie krajobrazie termodynamicznym niż o (pojedynczej) strukturze RNA (Rys.1.1.6). Ta perspektywa pozwala lepiej zrozumieć i precyzyjniej przewidywać właściwości RNA w układach eksperymentalnych.^[7, 8]

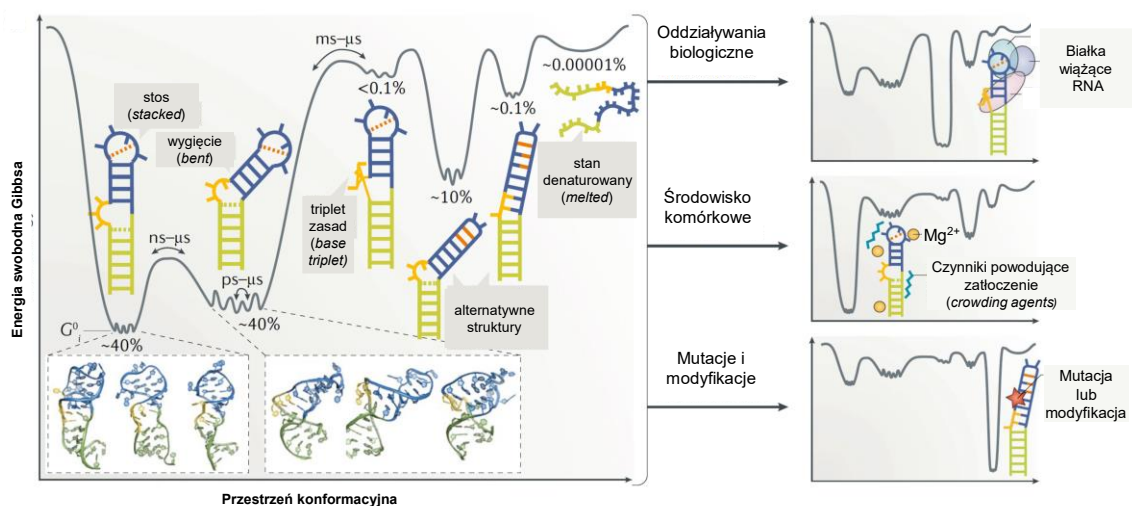
Model struktury przestrzennej danej sekwencji RNA można stworzyć za pomocą zautomatyzowanych metod obliczeniowych.^[13, 14] Bynajmniej, ich zastosowanie nie jest tożsame ze stworzeniem „rzeczywistego” modelu danej cząsteczki RNA.^[5, 15] Głównymi przyczynami tego problemu wydają się być: poziom skomplikowania mechanizmów odpowiedzialnych za tworzenie struktury przestrzennej RNA oraz niedostateczny stan wiedzy na ów temat. Przykładowo, zadaniem wielu z wyżej wymienionych narzędzi jest znalezienie struktury drugorzędowej danej sekwencji RNA o jak najniższej energii swobodnej Gibbsa (struktura MFE, ang. *minimum free energy*). W tym celu można wykorzystać algorytmy najbliższego sąsiada oraz następujące założenie (parametr wejściowy): dla dowolnego nukleotydu, prawdopodobieństwo oddziaływania jednego z drugim jest jednakowe/stałe.^[5] Innymi słowy, każdy nukleotyd może oddziaływać z każdym, bez „dyskryminacji”. W ten sposób dochodzi do faworyzowania motywów strukturalnych (np. spinek i pętli) o niższej energii Gibbsa, często podtrzymywane przez „kanoniczne” pary W-C. To może być przyczyną, dla której zaobserwowano pozytywną korelację między dokładnością modeli MFE a odsetkiem nukleotydów tworzących pary W-C, wyznaczonym na podstawie struktur krystalograficznych RNA (Rys.1.1.7).^[5] Aktualnie, najlepszym sposobem rzetelnego zastosowania modeli uzyskanych metodami obliczeniowymi jest ich konfrontacja z danymi eksperymentalnymi.



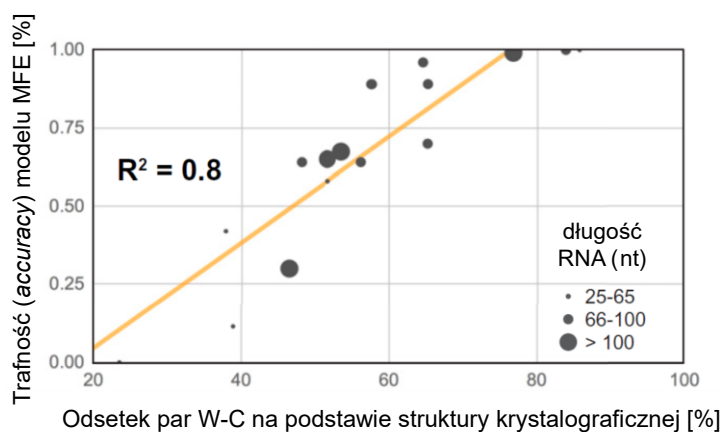
Rysunek 1.1.4: Sposoby reprezentacji struktury przestrzennej RNA na przykładzie 23S rRNA z *E. coli* (2904 nt). Na czerwono zaznaczono elementy struktury, które nie zostały uwzględnione w uproszczonej reprezentacji struktury drugorzędowej. Na podstawie publikacji w trybie wolnego dostępu [5].



Rysunek 1.1.5: Elementy wykorzystywane do precyzyjnego opisu struktury drugorzędowej RNA. Na podstawie [11].



Rysunek 1.1.6: Dynamika struktury przestrzennej RNA z perspektywy krajobrazu konformacyjnego. Lokalne minima potencjału Gibbsa odpowiadają różnym strukturom przestrzennym, które charakteryzują poszczególne populacje cząsteczek RNA o danej sekwencji. Zmiana środowiska powoduje zmianę krajobrazu. Na podstawie [8].



Rysunek 1.1.7: Korelacja między dokładnością modelu MFE a odsetkiem par W-C wyznaczona z wykorzystaniem znanych struktur krystalograficznych RNA o różnych rozmiarach. Każdy punkt (rozmiar kropki zależy od długości RNA) reprezentuje jedną strukturę krystalograficzną o znanym udziale par W-C oraz zgodność tej struktury z jej modelem MFE (odsetek prawidłowo wyznaczonych par zasad). Na podstawie publikacji w trybie wolnego dostępu [5].

1. 2 Wybrane funkcje i właściwości biologiczne RNA

Ogól biologicznych funkcji RNA

RNA jest jednym z kluczowych czynników stanowiących centralny dogmat biologii molekularnej, opisujący przepływ informacji genetycznej w organizmach. Informacja genetyczna przekazywana jest między kwasami nukleinowymi na drodze transkrypcji ($\text{DNA} \rightarrow \text{RNA}$), odwrotnej transkrypcji ($\text{RNA} \rightarrow \text{DNA}$) lub replikacji ($\text{DNA} \rightarrow \text{DNA}$ lub $\text{RNA} \rightarrow \text{RNA}$). Ponadto, informacja genetyczna może ulec translacji ($\text{RNA} \rightarrow \text{białko}$) i jednoczesnej degeneracji (za sprawą degeneracji kodu genetycznego). Owa degeneracja powoduje, iż w wyniku translacji informacja genetyczna zostaje bezpowrotnie zredukowana do struktury białka. Innymi słowy, struktura i właściwości RNA są *kluczowe* z perspektywy płynności informacji genetycznej. Jest to jedna z głównych przesłanek mówiących o cząsteczkach RNA jako „przodkach” wszystkich współczesnych organizmów. Kolejnymi przesłankami są: różnorodność funkcjonalna oraz konserwatyzm strukturalny naturalnie występujących RNA.^[16]

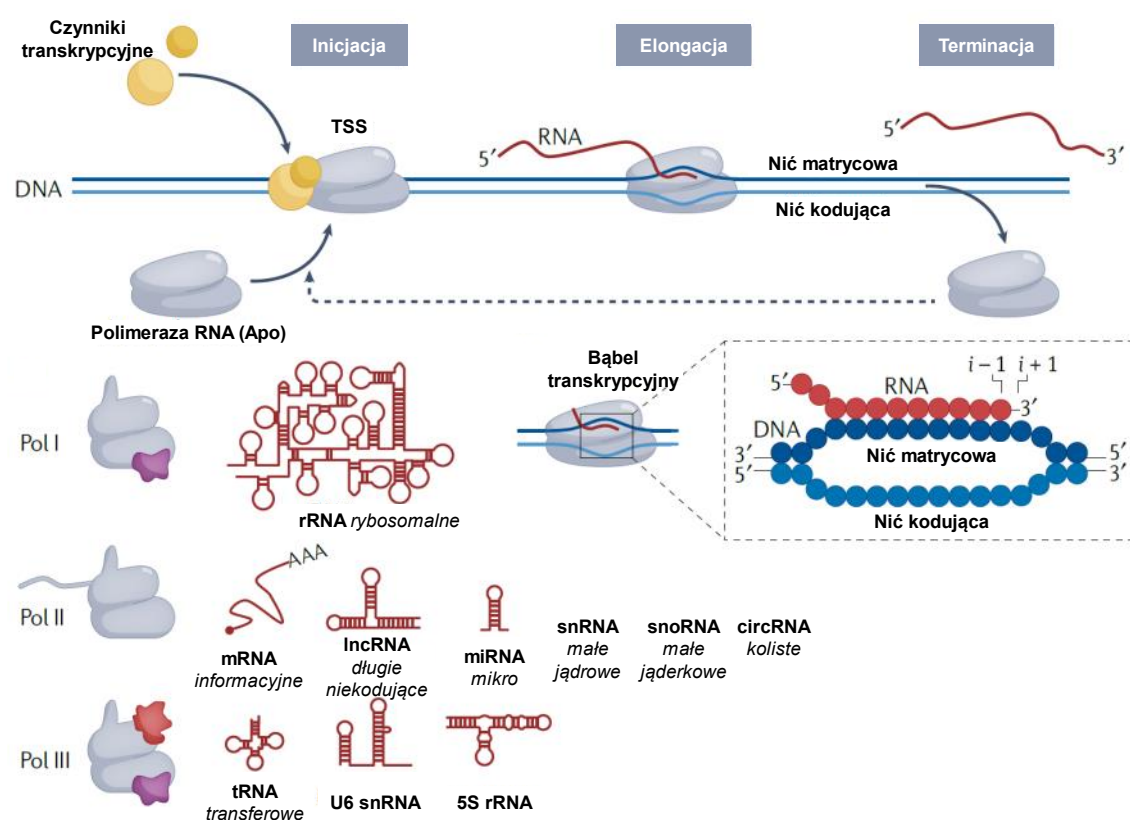
Poza przenoszeniem informacji genetycznej, cząsteczki RNA mają zdolność do wiązania małych- i wielkocząsteczkowych ligandów (aptamery czy RNA wiążące białka), katalizy reakcji chemicznych (rybozomy), a nawet do autoreplikacji. W strukturze i sekwencji, a także mechanizmach obróbki RNA rybosomalnego (rRNA), transferowego (tRNA), informacyjnego (mRNA), kolistego (circRNA) oraz różnego typu RNA niekodującego występują cechy wspólne dla eukariontów, prokariotów i archeonów. Poniżej zwięźle opisano wybrane elementy biologii RNA.

DNA-zależna transkrypcja RNA

Los każdej z endogennych (tj. naturalnego pochodzenia) cząsteczek RNA rozgrywa się w czterech głównych aktach: transkrypcji, dojrzewaniu, spełnieniu roli oraz degradacji. Są one ze sobą sprzężone oraz regulowane w skomplikowany sposób. Pierwszym z nich jest transkrypcja katalizowana przez polimerazy RNA, w szczególności te wykorzystujące informację genetyczną zawartą w DNA.

Transkrypcja katalizowana przez polimerazy RNA, w szczególności te przepisujące sekwencje zakodowane w DNA. Polimerazy RNA zależne od DNA wykazują szereg wspólnych cech mechanistycznych. Jest to grupa zawierająca zróżnicowane enzymy, przy czym zarówno jednocząsteczkowa polimeraza bakteriofaga T7 jak i eukariotyczne kompleksy polimeraz RNA (typu I, II oraz III) dokonują syntezy RNA w takich samych trzech etapach: inicjacji, elongacji i terminacji (Rys. 1.2.1).^[17-19] Pod nieobecność

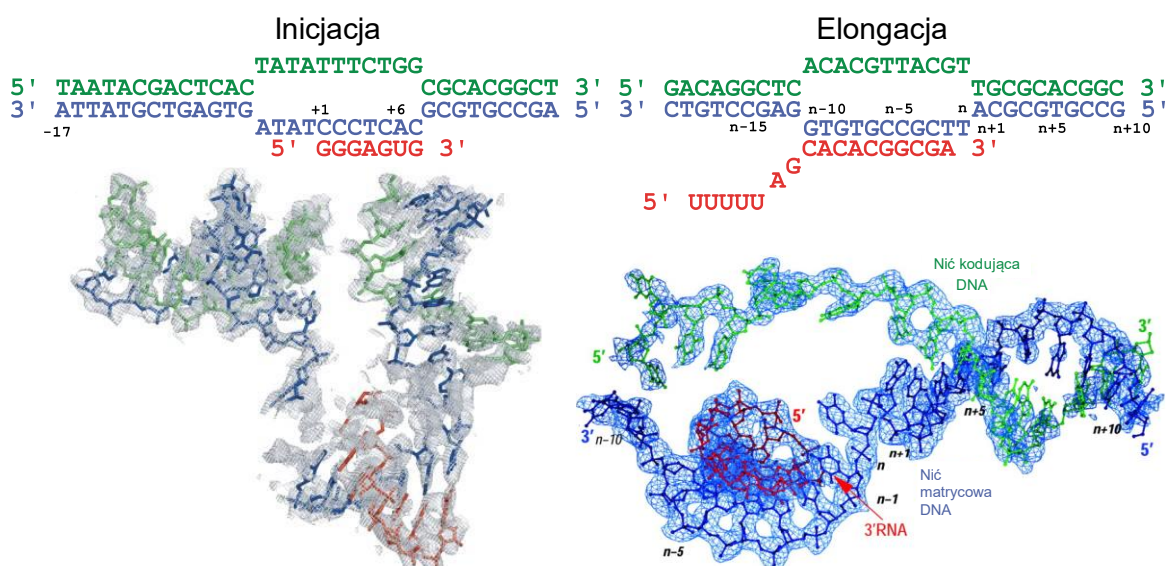
inicjatora, produktem transkrypcji jest RNA (transkrypt) o zadanej sekwencji, które zwierają grupę 5' trifosforanową na pierwszym nukleotydzie (na końcu 5') oraz 2',3'-diol rybozy nukleotydu znajdującego się na ostatnim nukleotydzie (na końcu 3') transkryptu. Z drugiej strony, różnice w mechanizmach transkrypcji objawiają się przede wszystkim w sferze jej regulacji. Co do zasady, regulacja transkrypcji może zachodzić za sprawą elementów *cis*, znajdujących się w ramach cząsteczki zawierającej dany gen (np. kasetę TATA czy promotor transkrypcji) lub elementów *trans* (np. białkowe czynniki transkrypcyjne). W eksperymentalnej i przemysłowej syntezie RNA najczęściej stosowaną polimerazą RNA jest ta pochodząca z bakteriofaga T7, zatem mechanizm transkrypcji katalizowany przez ten enzym posłuży jako przykład (Rys. 1.2.2).



Rysunek 1.2.1: Schemat procesu DNA-zależnej transkrypcji RNA u eukariontów. TSS: miejsce rozpoczęcia transkrypcji (ang. *transcription start site*). Na podstawie [18].

Transkrypcja rozpoczyna się od inicjacji, w trakcie której dochodzi do stworzenia bąbla transkrypcyjnego (Rys. 1.2.2) na drodze: wiązania DNA przez polimerazę w obszarze sekwencji promotora (23 pz, pozycje od -17 do +6), rozplecenia dwuniciowego DNA i syntezy krótkich fragmentów RNA (8-12 nt). W genomie bakteriofaga T7 obecnych jest siedemnaście różnych sekwencji wyznaczających miejsca rozpoczęcia transkrypcji (TSS, pozycje od -1 do +1).^[20, 21] W celu przejścia do fazy elongacji, kompleks bąbla

transkrypcyjnego musi przejść szereg drastycznych zmian konformacyjnych.^[19, 22] Jest to krytyczny moment, od którego zależy, czy synteza RNA zajdzie dalej czy dojdzie do przedwczesnej dysocjacji kompleksu i uwolnienia produktów przerwanej transkrypcji. Jego powodzenie oraz końcowa wydajność transkrypcji zależą od wielu czynników, takich jak sekwencja promotora oraz jej kontekst (elementy cis przed pozycją -17 oraz za +6).^[23, 24]

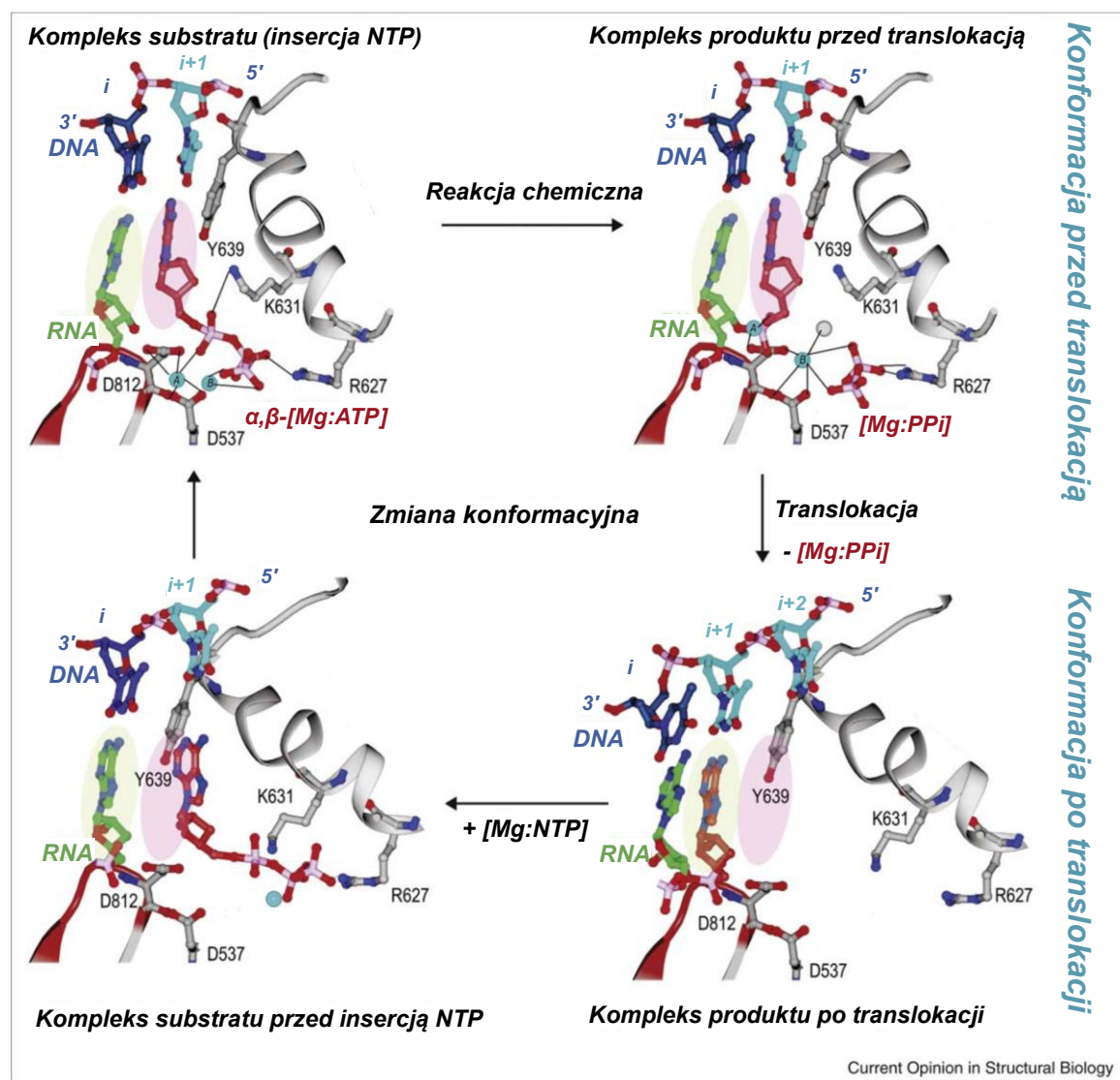


Rysunek 1.2.2: Struktura krystalograficzna bąbla transkrypcyjnego (polimeraza RNA T7) na etapach inicjacji oraz elongacji transkrypcji. Na podstawie ^[22].

W trakcie elongacji bąbel przybiera konformację optymalną dla translokacji polimerazy wzdłuż matrycowej nici DNA, między innymi dzięki obecności kanału wyjściowego dla nowopowstałej nici RNA. Szacowana prędkość syntezy łańcucha RNA na tym etapie może osiągać ponad 150 nukleotydów na sekundę.^[25] Ruch polimerazy jest napędzany hydrolizą bezwodnika fosforanowego i następczej eliminacji pirofosforanu magnezu z centrum aktywnego enzymu (Rys. 1.2.3). W trakcie elongacji dochodzi do symetrycznego splatania i rozplatania nici DNA, co minimalizuje straty energetyczne związane ze zmianami konformacyjnymi.

Do terminacji transkrypcji dochodzi za sprawą sekwencji terminatorowych. Są nimi sekwencje palindromowe, których transkrypcja prowadzi do stworzenia struktury spinki RNA. Spinka utworzona na końcu 3' RNA destabilizuje kompleks bąbla i promuje jego dysocjację.^[26] Naturalne sekwencje terminatorów bakteriofaga T7 nie gwarantują wydajnego kończenia transkrypcji. Można ją zwiększyć poprzez mutacje powodujące stabilizację struktury spinki RNA.^[17, 20] Alternatywną formą terminacji transkrypcji, dostępną wyłącznie w układach *in vitro*, jest spadek polimerazy z DNA (ang. *runoff*

transcription). To techniczne rozwiązanie również nie jest doskonałe ze względu na zwiększoną heterogenność sekwencji końców 3' wyprodukowanych transkryptów oraz powstawanie immunogennych produktów ubocznych (dwuniciowych RNA, dsRNA).^[27, 28]



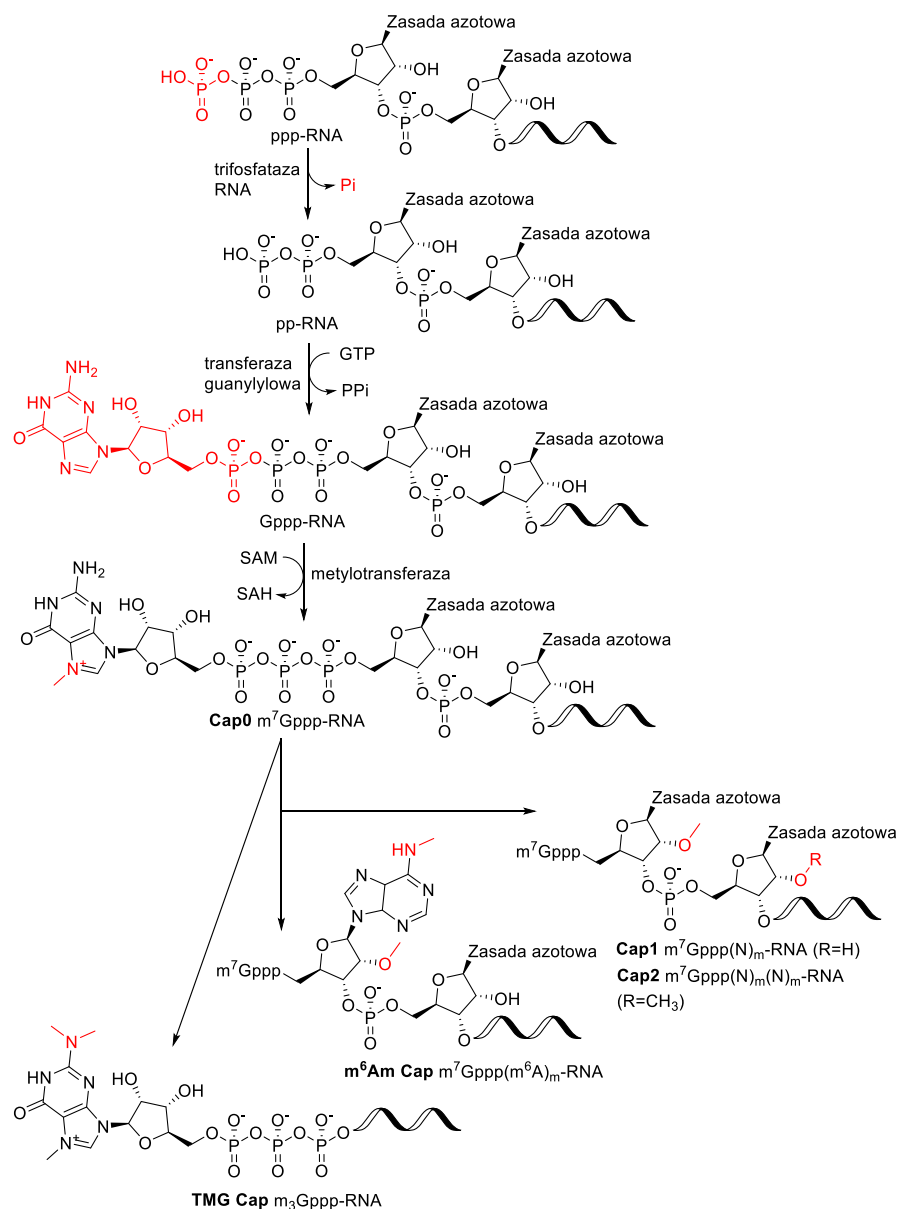
Rysunek 1.2.3: Centrum aktywne polimerazy RNA T7 w trakcie zmian konformacyjnych towarzyszących translokacji. Kationy magnezu oznaczono literami A i B. Na podstawie publikacji^[19].

Dojrzewanie informacyjnego RNA

Dojrzewanie RNA jest kluczowym etapem z perspektywy funkcjonowania transkryptomu. W jego trakcie dochodzi do uściślenia sekwencji, struktury chemicznej oraz wiązania produktów transkrypcji z czynnikami *trans*, odpowiedzialnymi za dalsze losy danej cząsteczki RNA (tj. funkcja, degradacja, transport i lokalizacja). W przypadku produktów polimerazy RNA typu II, proces dojrzewania zachodzi kotranskrypcyjnie (w

trakcie transkrypcji) jak i posttranskrypcyjnie (po jej zakończeniu). Dalej omówiono trzy główne etapy dojrzewania eukariotycznego mRNA (kaping, splicing i poliadenylacja).

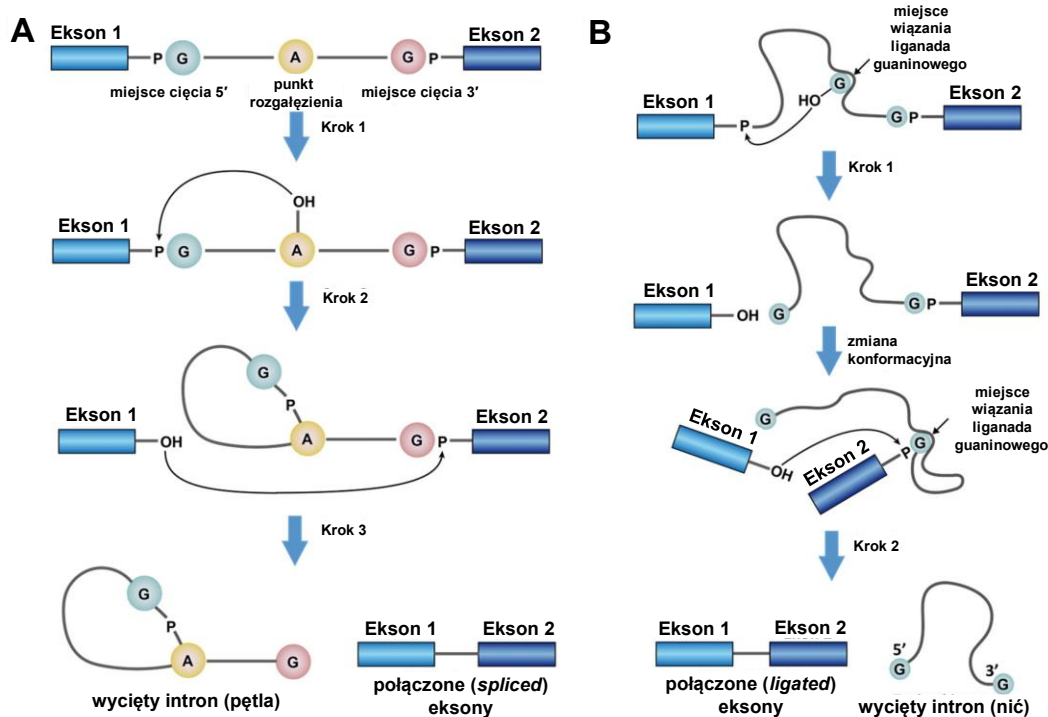
Kapowanie (ang. *capping*) jest procesem zachodzącym natychmiast po zakończeniu inicjacji transkrypcji. W wyniku kapowania dochodzi do przekształcenia grupy 5' trifosforanowej (ppp-RNA) w strukturę kapu (inaczej czapeczki, ang. *cap*), która składa się z reszty N7-metyloguanozyny (m^7G) połączonej mostkiem 5'-5' trifosforanowym z pierwszym nukleotydem transkryptu (Rys. 1.2.4). Podstawową ścieżką syntezy struktury kapu jest transfer grupy guanylowej z udziałem GTP oraz następcza metylacja w pozycji N7. Jej produktem jest struktura „Cap 0” (m^7Gppp -RNA). Koniec 5' RNA może ulegać dalszym modyfikacjom, w szczególności metylacji grup 2' OH (Cap 1, Cap 2), hipermetylacji guanozyny (TMG cap) czy metylacji pozycji N6 sąsiadującej adenozyiny (m^6A).^[29-31] Metylacja oraz inne posttranskrypcyjne modyfikacje chemiczne RNA mają kluczowy wpływ z perspektywy regulacji jego biologicznej funkcji (przykładowo translacji dla mRNA, rRNA i tRNA; splicingu lub regulacji dla snRNA, lncRNA i miRNA), jego degradacji czy immunogenności.^[30-33]



Rysunek 1.2.4: Schemat procesu kapiungu eukariotycznych RNA. SAM: S-adenozylometionina, SAH: S-adenozylhomocysteina. Na podstawie [30, 32].

Równolegle do fazy elongacji transkrypcji zachodzi splicing (czytaj *splajsing*), którego zasadniczym celem jest osiągnięcie docelowej sekwencji mRNA z jego prekursora (pre-mRNA). Splicing może również być etapem biosyntezy innych typów RNA, między innymi kolistego RNA (circRNA). Można wyróżnić dwa mechanizmy splicingu: katalizowany przez spliceosom (czytaj *splajsosom*) oraz splicing autokatalityczny. W obu wypadkach splicing prowadzi do połączenia sekwencji eksonowych (ulegających ekspresji) w wyniku dwóch, regioselektywnych reakcji transestryfikacji i eliminacji sekwencji intronowych (niekodujących, wewnątrz genu; Rys. 1.2.5). W przypadku większości eukariotycznych pre-mRNA, splicing zachodzi z udziałem spliceosomu, czyli

kompleksu złożonego z białek i snRNA (mały jądrowy RNA) o funkcji katalitycznej.^[34, 35] W tym celu grupa 2' OH adenozy, znajdującej się w punkcie rozgałęzienia, prowadzi atak nukleofilowy na grupę 3',5'-fosfodiesterową wyznaczoną przez miejsce cięcia 5'. W wyniku transestryfikacji dochodzi do syntezy wiązania 2',5'-fosfodiesterowego (zawartego w strukturze pętli intronu) oraz uwolnienia grupy 3' OH na ostatnim nukleotydzie eksonu. Uwolniona grupa 3' OH eksonu prowadzi kolejny atak nukleofilowy, tym razem na grupę 3',5'-fosfodiesterową wyznaczoną przez miejsce cięcia 3'. W wyniku drugiej transestryfikacji dochodzi do połączenia sekwencji eksonowych za pomocą wiązania 3',5'-fosfodiesterowego i eliminacji intronu o strukturze pętli (inaczej *lassa*, ang. *lariat*). Wyeliminowane pętle RNA mogą być wykorzystane, między innymi jako prekursor regulatorowych miRNA lub circRNA. W splicingu autokatalitycznym intron jest rybozymem wykorzystującym ligand (ganozynę lub jej fosforany). Ligand pełni funkcję donora grupy hydroksylowej (3' OH), która wchodzi w pierwszą reakcję transestryfikacji. Ostatecznie, autokatalityczny splicing prowadzi do połączenia eksonów i eliminacji intronu w formie liniowej (inaczej nici).



Rysunek 1.2.5: Porównanie splicingu A) katalitycznego (z udziałem kompleksu spliceosomu) oraz B) autokatalitycznego (z udziałem rybozymu i liganda guaninowego). Na podstawie publikacji w trybie wolnego dostępu^[35].

Istotnym aspektem regulacji splicingu jest splicing alternatywny, w wyniku którego powstają izoformy mRNA: cząsteczki mRNA pochodzące z transkrypcji jednego genu oraz różniące się sekwencją kodującą (kolejnością eksonów). Większość (>95%) ludzkich

mRNA ulega alternatywnemu splicingowi, co jest kolejnym czynnikiem pogłębiającym różnorodność i dynamikę transkryptomu oraz powiązanego z nim proteomu (sumie wszystkich białek w komórce w danym momencie). Do regulacji alternatywnego splicingu wykorzystywane są lncRNA (długie niekodujące RNA).^[36]

Splicing z udziałem spliceosomu jak i splicing autokatalityczny nie wymagają zużycia energii pochodzącej z hydrolizy bezwodników fosforanowych (tj. NTP). To można tłumaczyć tym, że opiera się on na „energetycznie obojętnych” reakcjach transestryfikacji oraz wzroście entropii towarzyszącej eliminacji intronowego RNA (innymi słowy, splicing mógłby być napędzany siłami entropowymi).^[37]

Jeszcze przed terminacją transkrypcji, pre-mRNA ulega odcięciu i poliadenylacji z udziałem kompleksu CPF (ang. *cleavage and polyadenylation factor*).^[38, 39] W ten sposób otrzymuje kolejny (obok kapu na końcu 5') charakterystyczny element strukturalny: homopolimeryczną sekwencję adeniny (tzw. ogon poliA) na końcu 3'. Oba elementy są kluczowe z perspektywy translacji oraz degradacji mRNA. Odpowiedź na pytanie, który z elementów ma większe znaczenie z perspektywy regulacji ekspresji genów nie jest prosta, ani jednoznacznie sformułowana. Co do zasady, wzrost długości ogona poliA (w zakresie od 30 do ponad 250 nukleotydów) wydłuża czas półtrwania mRNA.^[39] Analogicznie do splicingu, odcięcie i poliadenylacja mogą mieć alternatywne przebiegi.^[38, 40]

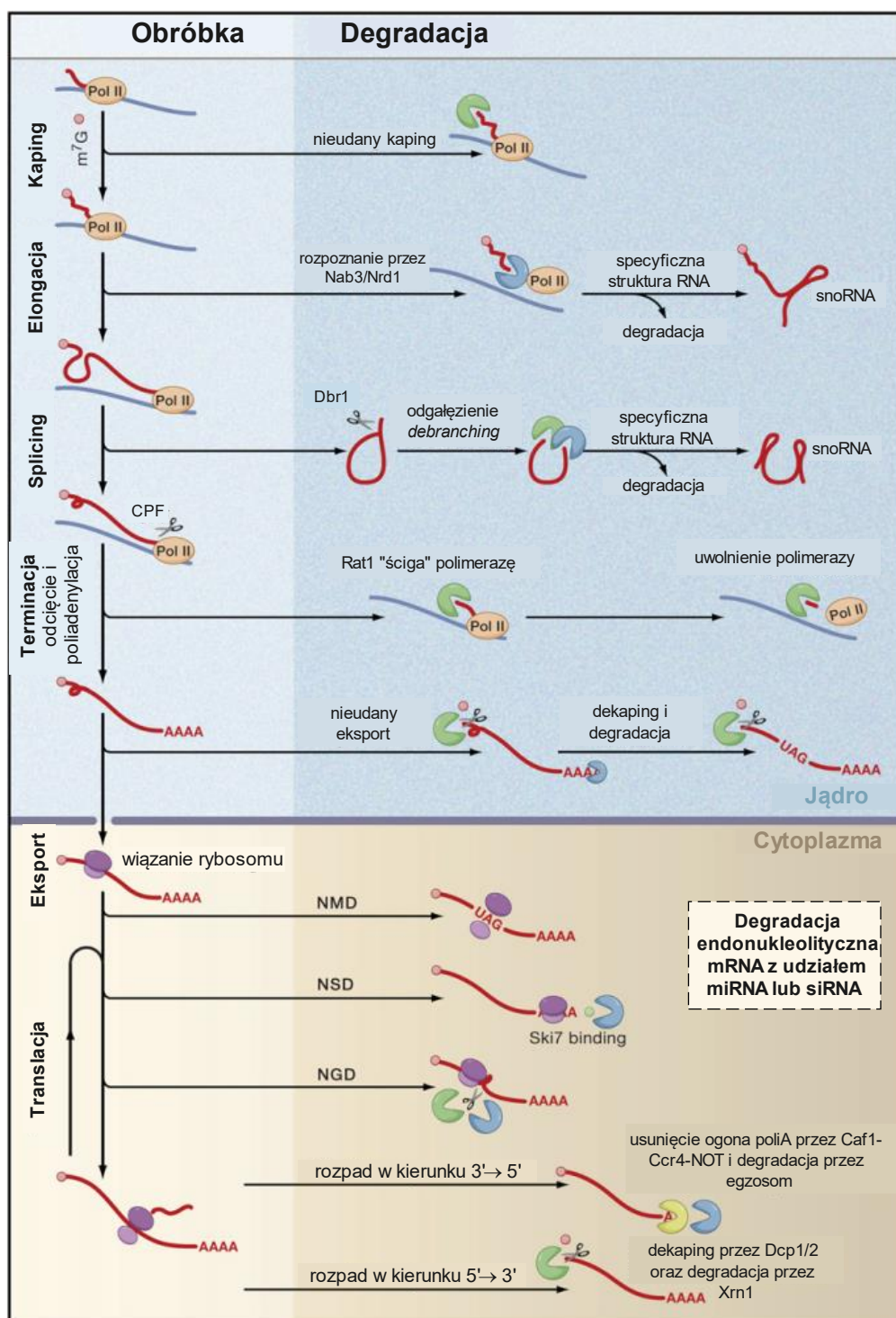
Degradacja informacyjnego RNA

Mimo wysokiej intensywności transkrypcji, RNA w komórkach nie ulega akumulacji. Dzieje się tak ze względu na istnienie złożonych, zdegenerowanych (jeden cel może zostać osiągnięty na więcej niż jeden sposób) i ściśle kontrolowanych procesów degradacji RNA.^[41] Są one zaangażowane, między innymi, w kontroli jakości i wykrywaniu błędów powstałych w trakcie obróbki (ang. *processing*) endogennych RNA czy rozpoznawaniu transkryptów pochodzenia zewnętrznego (egzogennych, np. wirusowych).^[42] W procesach degradacji RNA można wyróżnić kilka typów czynników *trans*: rybonukleazy (inaczej RNazy), helikazy, polimerazy oraz białka opiekuńcze (inaczej szaperonowe, ang. *chaperones*). Kluczową aktywność przypisuje się rybonukleazom, które klasyfikowane są ze względu na miejsce i mechanizm hydrolizy (inaczej cięcia) RNA. Wyróżnia się endonukleazy, dokonujące cięcia wewnątrz cząsteczki RNA (np. RNaza A, RNaza T1 i RNaza H), egzonukleazy działające w kierunku 5'→3' (np. Rat1 i XRN 1) oraz egzonukleazy działające w kierunku 3'→5' (np. PNPaza, RNaza R i kompleks PAN1/2).^[39, 41, 43] Działanie egzonukleaz jest ściśle powiązane z czynnikami oddziałującymi ze specyficznymi strukturami znajdującymi się końcach RNA, takimi jak enzymy i

kompleksy dekapujące (np. RppH i Dcp1/2).^[41, 44] Helikazy „rozplatają” struktury drugorzędowe RNA oraz wspomagają dysocjacje związanych z nimi makromolekularnych ligandów (np. białek), co powoduje, iż znajdują zastosowanie na wszystkich etapach procesu ekspresji genów.^[45] Polimerazy (np. kompleks TRAMP) rozbudowują koniec 3' o sekwencje poliA lub poliU, będące sygnałem i swoistym „pasem startowym” dla 3'→5' egzonukleaz, co znajduje swoją analogię w poliubikwitynacji białek.^[41, 46] Białka opiekuńcze, takie jak pierścieniowy kompleks Lsm1-7, pełnią funkcję regulatorową, poprzez selektywne promowanie oddziaływań RNA-RNA lub RNA-białko. Jednym z najbardziej charakterystycznych i uniwersalnych elementów maszynery degradacji RNA jest kompleks egzozomu, który łączy ze sobą aktywność helikazy, endonukleazy oraz 3'→5' egzonukleazy.^[41, 47]

Z perspektywy eukariotycznego mRNA, procesy degradacji zachodzą na każdym z etapów jego funkcjonowania (Rys. 1.2.6).^[41] W wyniku niepowodzenia procesu kapowania, może dojść do degradacji od końca 5' i terminacji transkrypcji za sprawą egzonukleazy Rat1. Dalej, na etapach elongacji transkrypcji i splicingu, może dojść do degradacji transkryptu i/lub syntezy prekursorów małych jąderkowych RNA (snoRNA). Jeżeli nie, pre-mRNA ulega odcięciu i poliadenylacji, a 5' egzonukleazą Rat1 „ściga” (ang. *chase*) i „torpeduje” (ang. *to torpedo*) kompleks RNA polimerazy II, co finalizuje procesy transkrypcji i dojrzewania mRNA. Dalej mRNA jest eksportowane z jądra do cytoplazmy. Niepowodzenie na tym etapie skutkuje degradacją za pomocą kompleksów dekapujących i egzonukleaz. mRNA w cytoplazmie ulega translacji, której mechanizm omówiono w dalszej części tekstu. Jeżeli mRNA zawiera błędną sekwencję, może ono ulec degradacji już w pierwszej rundzie translacji na trzech różnych ścieżkach: NMD (ang. *nonsense mediated decay*), spowodowana wystąpieniem przedwczesnego kodonu stop; NSD (ang. *non-stop mediated decay*), spowodowana brakiem kodonu stop; lub NGD (ang. *no-go mediated decay*), spowodowana przedwczesnym zatrzymaniem rybosomu. W dalszych rundach translacji również może dojść do degradacji NSD lub NDG i terminacji translacji. Równolegle, w trakcie prawidłowo zachodzącej translacji ogon poliA ulega stopniowemu skróceniu za sprawą kompleksu Caf1-Ccr4-NOT. Deadenylacja, dekapowanie oraz translacja są ze sobą sprzężone w czasie i przestrzeni, a mechanizm owego sprzężenia jest polem wieloletniej debaty.^[39] W cytoplazmie, nieaktywne translacyjnie cząsteczki mRNA są gromadzone, wraz z białkami odpowiedzialnymi za degradację mRNA, w dynamicznych strukturach (granulach) zwanych ciałkami P (ang. *processing bodies, P-bodies*).^[48] Procesy zachodzące w granulach, ich powstawanie oraz rola w regulacji

ekspresji genów również pozostają przedmiotem intensywnych badań. Ponadto, istotnym elementem posttranskrypcyjnej regulacji ekspresji genów jest endonukleolityczne cięcie mRNA z udziałem miRNA oraz siRNA.^[41, 49]



Rysunek 1.2.6: Procesy obróbki i degradacji produktów transkrypcji katalizowanej przez polimerazę RNA typu II (u Eukaryota). Skróty: CPF – kompleks odcinający i poliadenylujący pre-mRNA (ang. *cleavage and polyadenylation factor*); NMD – ścieżka spowodowana wystąpieniem przedwczesnego kodonu stop (ang. *nonsense mediated decay*); NSD – ścieżka spowodowana brakiem kodonu stop (ang. *non-stop mediated decay*); lub

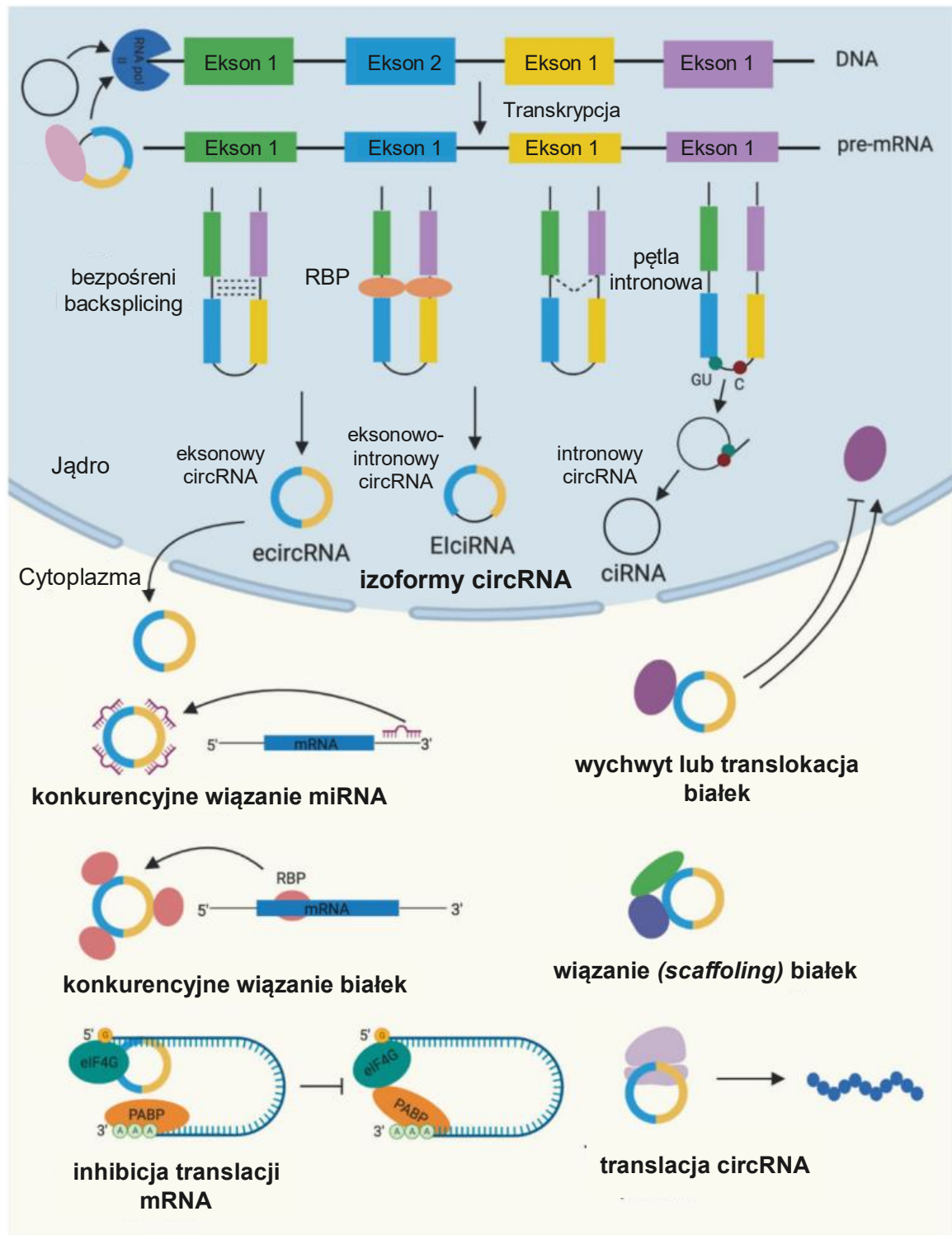
NGD – ścieżka spowodowana przedwczesnym zatrzymaniem rybosomu (ang. *no-go mediated decay*). Na podstawie publikacji w trybie wolnego dostępu ^[41].

Geneza i postulowane funkcje kolistego RNA

Koliste RNA (ang. *circular*, circRNA) różnią się od pozostałych, naturalnie występujących polimerów rybonukleotydów jedną, główną cechą — topologią. Są to cząsteczki, których końce 5' oraz 3' uległy połączeniu i w konsekwencji zanikowi. Podobnie jak w przypadku różnicy między DNA i RNA, z perspektywy cząsteczki RNA jako całości, cyklizacja jest stosunkowo niewielką zmianą (np. jest kwestią kondensacji grupy 5' fosforanowej i 3' OH), która niesie ze sobą szereg poważnych konsekwencji z perspektywy jej właściwości. Aktualnie, są one obiektem intensywnych badań.^[50-52]

Pierwsza fala badań nad circRNA przypada na okres od lat siedemdziesiątych do lat dziewięćdziesiątych ubiegłego wieku. Wtedy zidentyfikowano pierwsze cząsteczki circRNA, głównie pochodzenia wiroidowego i wirusowego, a nieliczne przypadki odnotowane u eukariontów były interpretowane jako produkty błędnego splicingu (ang. *miss-splicing*). Rozwój techniczny, w szczególności metod wysokoprzepustowego sekwencjonowania kwasów nukleinowych, zainicjował na przełomie pierwszych dekad XXI wieku kolejną falę, której częścią jest niniejsza rozprawa.

Głównym mechanizmem powstawania eukariotycznych circRNA jest alternatywny splicing, to jest splicing „wsteczny” (ang. *back-splicing*) z udziałem sekwencji eksonowych lub intronowych pre-mRNA. (Rys. 1.2.7) W procesie wstecznego splicingu mogą brać udział zarówno spliceosom jak i specyficzne sekwencje intronowe. Z tego powodu izoformy circRNA stanowią „konkurencję” dla pre-mRNA w kontekście regulacji transkryptomu i powiązanego z nim metabolizmu komórki. Z drugiej strony, cyrkularna topologia jest najczęściej prowadzana do odporności circRNA na działanie egzonukleaz, którą tłumaczy się wydłużenie czasów ich życia oraz akumulacji w komórkach. Ponadto, wykryto zróżnicowane poziomy szczególności sekwencji circRNA w zależności od typów komórek i tkanek.



Rysunek 1.2.7: Postulowane funkcje biologiczne circRNA. Na podstawie publikacji w trybie wolnego dostępu ^[53].

Lista postulowanych funkcji circRNA stale się powiększa. Co do zasady, circRNA pełnią funkcję regulacyjną poprzez wiązanie białek i kwasów nukleinowych zaangażowanych w: transkrypcję, splicing, translację i degradację mRNA (np. poprzez wiązanie miRNA), odpowiedź immunologiczną oraz inne szlaki metaboliczne i

rozwojowe. Ponadto, niektóre sekwencje circRNA mogą ulegać translacji na drodze mechanizmów niezależnych od struktury kapu (np. zależnych od sekwencji IRES, opisanych w dalszej części tekstu).

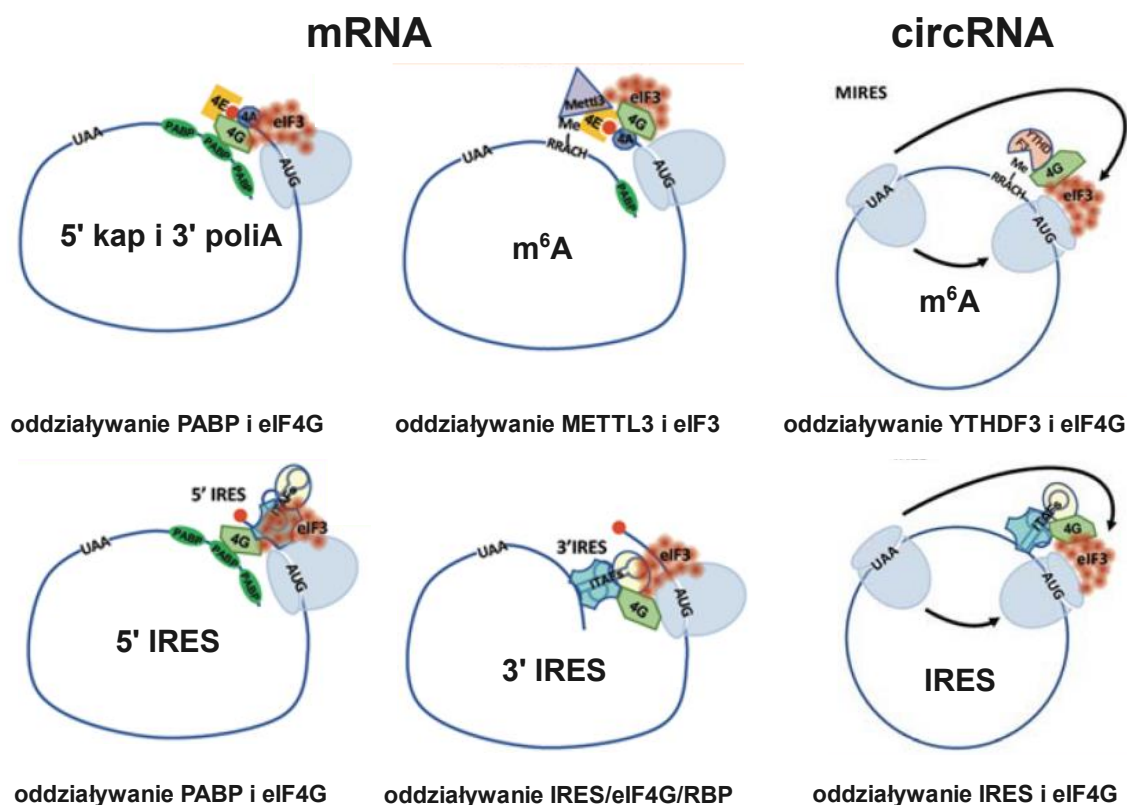
Wymienione właściwości i funkcje circRNA widziane są jako atrakcyjne z perspektywy zastosowań diagnostycznych i terapeutycznych.^[53-55] Przykładowo, ludzka ślina i krew stanowią bogate źródło circRNA, a specyfika ich ekspresji została powiązana z występowaniem szeregu stanów chorobowych. Z tego powodu próbuje się wykorzystać circRNA jako markery chorób układów: endokrynowego (np. cukrzycy), oddechowego (np. gruźlicy), pokarmowego (np. nowotworów jelita lub trzustki) czy nerwowego (np. choroby Parkinsona lub Alzheimer). Aktualnie prowadzone są również badania, w których circRNA stanowią cele terapeutyczne.^[53] Z perspektywy niniejszej rozprawy, szczególnie istotne jest potencjalne zastosowanie syntetycznych circRNA jako nośników terapeutycznej informacji genetycznej.^[55-58]

Zaletą oraz jednym z podstawowych problemów badań nad circRNA jest ich stosunkowo krótka historia. Z tego powodów wiele z postulowanych zjawisk związanych z circRNA została udokumentowana na niewielkiej liczbie niezależnych przykładów. Ponadto, nowość i innowacyjność owej dziedziny idą w parze z konkurencyjnością, która przyczynia się do powstawania presji, pośpiechu i w konsekwencji (pozornych) sprzeczności w obserwacjach. Przykładami takich sprzeczności są immunogenność circRNA na drodze aktywacji receptorów RIG-I czy wysoki odsetek fałszywie pozytywnych wyników sekwencjonowania circRNA.^[57, 59] Z tego powodu warto uzbroić się w sceptycyzm, cierpliwość oraz inwestować w udoskonalanie metod badawczych, których rozwój również stał się bardzo intensywny. Szczególnie atrakcyjne stało się udoskonalanie technik sekwencjonowania i analizy bioinformatycznej, nadekspresji, wyciszania, wykrywania, izolacji, syntezy czy uwzględnienie efektów konformacyjnych circRNA.^[52]

Translacja mRNA i circRNA u eukariontów

W wyniku translacji powstają białka (polipeptydy) na podstawie informacji genetycznej zawartej w strukturze RNA. W procesie zaangażowane są trzy główne elementy: kodujący RNA (tj. mRNA lub circRNA), kompleks rybosomu oraz tRNA. Do transferu informacji genetycznej dochodzi za pośrednictwem kontaktu struktury kodującego RNA oraz tRNA w trójnukleotydowych obszarach kodonu oraz antykodonu. Natomiast, relację sekwencji kodującego RNA i docelowego białka określa zdegenerowany i ewolucyjnie uniwersalny kod genetyczny.^[60] Największym, najbardziej

złożonym (i jednym z najstarszych ewolucyjnie) elementem maszynierii translacyjnej jest rybosom (eukariotyczny, 80S, 4200 kDa), składający się z dwóch podjednostek (małej 40S i dużej 60S), złożonych z czterech cząsteczek RNA (18S, 5S, 28S i 5.8S) i ponad pięćdziesięciu białek. Te ostatnie stanowią mniejszość masy kompleksu rybosomu. Struktura aktywnego kompleksu rybosomu (zawierająca rybosom, kodujące RNA oraz tRNA) zawiera trzy obszary, w których dochodzi do syntezy polipeptydu, wiązania i dysocjacji tRNA: E (jak „wyjście”, ang. exit), P (jak peptydylo-tRNA) i A (aminoacylo-tRNA). Z mechanistycznego punktu widzenia, w procesie translacji można wyróżnić trzy etapy: inicjację, elongację oraz terminację. Każdy z tych etapów jest wybitnie złożony, a szczegóły ich mechanizmów pozostają obiektami intensywnych badań.^[39, 50, 61, 62] W dalszej części tekstu wyróżniono proces inicjacji translacji oraz porównano jego mechanizmy dla mRNA oraz circRNA (Rys.1.2.8).



Rysunek 1.2.8: Postulowane mechanizmy inicjacji translacji mRNA oraz circRNA. RBP: białko wiążące RNA. Na podstawie publikacji w trybie wolnego dostępu ^[50].

W ogólności, kanoniczny mechanizm inicjacji kap-zależnej translacji eukariotycznego mRNA przewiduje szereg sprzężonych w czasie i przestrzeni zdarzeń: powstanie kompleksu preinicjującego 43S (PIC, złożony z 40S, inicjatorowego metionino-tRNA oraz czynników eIF2 i eIF3); wiązanie struktury kapu na końcu 5' przez czynnik

eIF4E; powstanie kompleksu inicjującego translację, złożonego z: struktur końca 5' i 3' mRNA, białek wiążących owe struktury (eIF4E oraz PABP) oraz czynników eIF4G i eIF4A; oddziaływanie PIC z kompleksem inicjującym translację za pośrednictwem IF4G i eIF4A; „odnalezienie” i związanie kodonu start (AUG) przez PIC z wykorzystaniem aktywności helikazy oraz interakcji z kompleksem inicjującym translację. Istnieje wiele, alternatywnych mechanizmów inicjacji translacji. W szczególności, wyróżnia się mechanizmy zależne od sekwencji typu IRES (ang. internal ribosomal entry site) oraz mechanizmy zależne od modyfikacji m6A.^[50, 61, 63, 64] W pierwszym przypadku, struktura przestrzenna sekwencji IRES jest zdolna do wiązania czynników ITAF (ang. *IRES trans-acting factors*), przykładowo białek związanych ze stresem komórkowym, oraz czynników promujących translację, takich jak eIF3, eIF4G oraz eIF4E.^[65] Z drugiej strony, inicjacja z udziałem m6A wymaga czynników o aktywności metylotransferaz (np. METTL3) lub białek selektywnie wiążących sekwencje zawierające m6A (np. YTHD).^[50, 63] Oba wymienione mechanizmy inicjacji niezależnej od kapu znalazły centralne miejsce w dyskusji na temat translacji circRNA.^[52, 55-57, 66, 67]

Zarówno kolistę jak i liniowe cząsteczki RNA mogą ulegać translacji niezależnej od kapu (co do zasady, obecność kapu na końcu 5' liniowego RNA może być wtórna lub zapewniać ochronę przed degradacją w kierunku 5'→3'). Co więcej, w trakcie translacji, liniowe cząsteczki kodujące ulegają niekowalencyjnej cyklizacji. Jedną z hipotez tłumaczących to zjawisko są fizyczne właściwości RNA — termodynamicznie uprzywilejowane struktury RNA zawierają końce 5' oraz 3' blisko siebie, bez konieczności udziału czynników trans.^[9, 50] Innymi słowy, niekowalencyjna cyklizacja mRNA zachodząca w trakcie translacji, która objawia się między innymi powstawaniem kolistych polisomów, miałyby być promowana przez strukturę kodującego RNA, nie zaś czynników inicjujących translację.

2. RNA jako narzędzie

2.1 Terapie genowe

O terapiach genowych

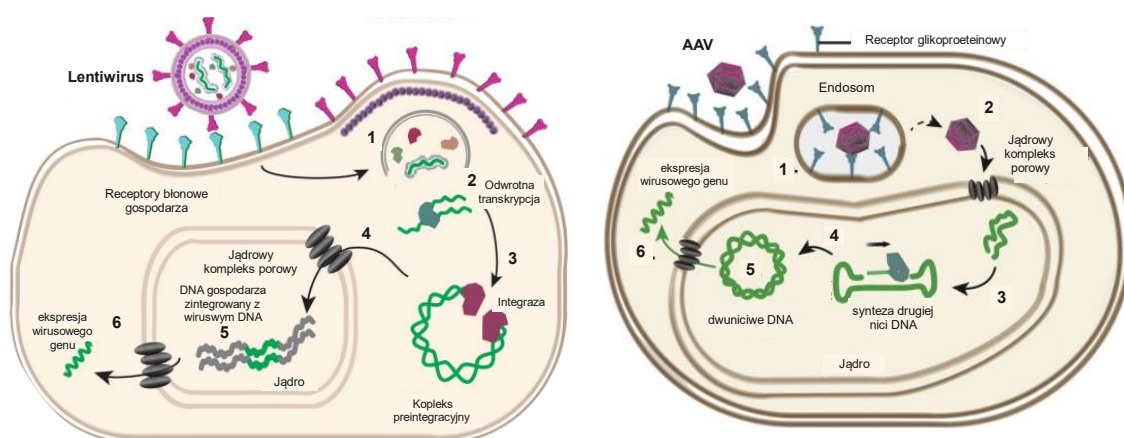
Terapie genowe są aktualnie jedną z najbardziej dynamicznych i obiecujących dziedzin z pogranicza biotechnologii i medycyny. Definicja terapii genowej nie jest jednoznaczna, przy czym oryginalnie krąży wokół traktowania informacji genetycznej jako celu terapeutycznego.^[68-72] W ramach niniejszej rozprawy zaproponowano następującą definicję: terapia genowa zakłada wykorzystanie informacji genetycznej jako celu terapeutycznego (np. edycja patologicznego genu) lub jako środka terapeutycznego (np. suplementacja genetyczna), co wymaga zastosowania kwasów nukleinowych o charakterze czynników *cis* (np. nośniki terapeutycznej informacji genetycznej) lub *trans* (np. o charakterze regulatorów ekspresji patologicznego genu). Terapie genowe można podzielić ze względu na miejsce i okoliczności ich działania (*in vivo*, *in situ* lub *ex vivo*) oraz na tryb i efekt ich działania (przejściowy lub stabilny). Terapie genowe *in vivo* zakładają wprowadzenie środka terapeutycznego do organizmu pacjenta i jego systemowe dostarczenie (np. przez krwioobieg) do docelowego organu lub tkanki (np. mózgu, wątroby czy kanału kręgowego), w którym ma zadziałać terapeutyczny kwas nukleinowy. Podobnie jest w wypadku terapii *in situ*, z tą różnicą, że w ich wypadku środek terapeutyczny jest dostarczany bezpośrednio do organu lub tkanki (np. gałki ocznej, nowotworu czy limfy). Podejście *ex vivo* zakłada pobranie komórek pacjenta (np. komórek macierzystych, nowotworowych lub układu odpornościowego), manipulację (np. edycję genomu) i hodowlę w warunkach *in vitro* i następnie wprowadzenie ich do organizmu pacjenta. Po dostarczeniu, terapeutyczny kwas nukleinowy może działać w sposób przejściowy (np. w celu suplementacji genetycznej lub regulacji ekspresji genów) lub stabilny (np. w celu rozszerzenia lub edycji genomu pacjenta).

Ponad trzydziestoletnia historia terapii genowych zawiera zarówno spektakularne sukcesy (np. leczenie pacjentów dotkniętymi chorobami neurodegeneracyjnymi, nowotworami, wrodzoną ślepotą czy zastosowanie szczepionek mRNA w pandemii COVID-19), tragiczne wydarzenia (np. śmierć Jessego Gelsingera w 2000 roku) oraz kontrowersje (np. „dzieci CRISPR” w 2018 roku).^[69, 71] Spośród garstki dopuszczonych oraz tysięcy terapii genowych w trakcie badań klinicznych można wyróżnić: terapie przeciwnowotworowe i spersonalizowane, szczepionki przeciwwzakaźne, produkcję przeciwciał, leczenie chorób wrodzonych oraz chorób rzadkich.^[70, 71, 73] Rozwój terapii

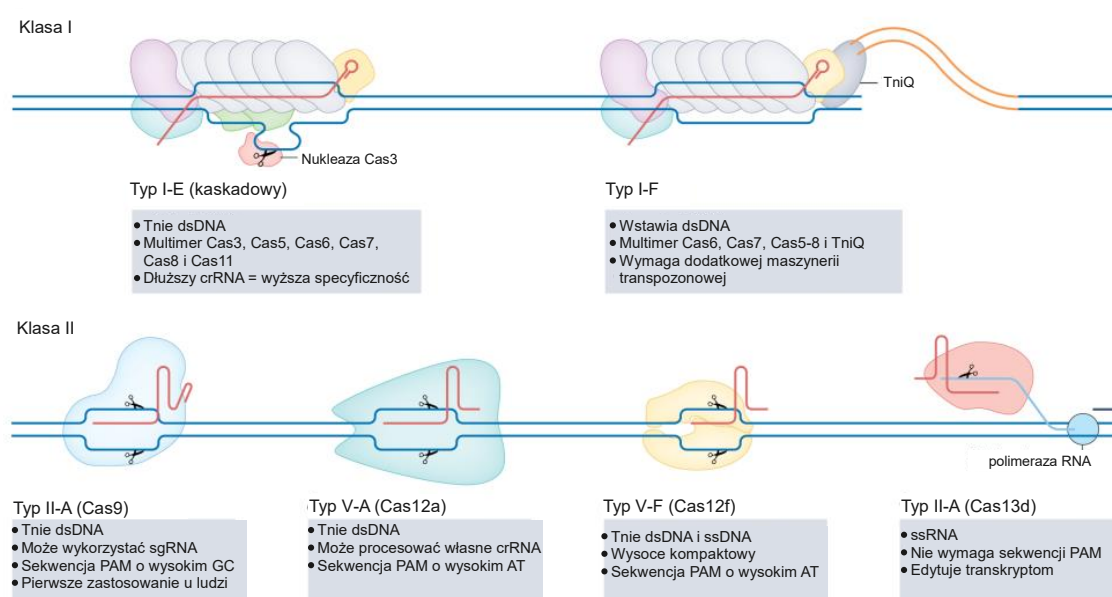
genowych (rozumianych jako „klasyczne” terapie genowe oraz zastosowanie terapeutycznych oligonukleotydów lub mRNA) wiąże się z bezprecedensowym wzrostem gospodarczym. Rozmiary współczesnego rynku wyżej wymienionych technologii można szacować na blisko 50 miliardów dolarów.^[74] Jest to, między innymi, związane z zwiększonymi kosztami badań podstawowych i klinicznych. W wypadku „klasycznych” terapii (np. z zastosowaniem małowcząsteczkowych leków), koszty badań klinicznych liczone są w dziesiątkach lub setkach milionów dolarów.^[75] W celu zwiększenia bezpieczeństwa potencjalnych pacjentów, koszty badań klinicznych innowacyjnych terapii genowych ulegają zwielokrotnieniu. Niestety, to może prowadzić do jeszcze wyższych kosztów ponoszonych przez konsumentów. Koszty wielu terapii genowych chorób rzadkich i nowotworowych sięgają dziesiątek lub setek tysięcy dolarów na pacjenta.^[71] Skrajnym przypadkiem jest Onasemnogene abeparvovec (Zolgensma), uznawany za najdroższy lek na świecie. Terapia rdzeniowego zaniku mięśni (SMA) z jego użyciem to koszt ponad dwóch milionów dolarów na pacjenta.^[76] Jednym z problemów mających potencjalnie największy wpływ na rozwój i „oblicze” terapii genowych w przyszłości jest dynamiczna równowaga między zapotrzebowaniem i innowacyjnością a dostępnością, regulacją i bezpieczeństwem.^[71]

W klasycznym rozumieniu, podstawową strategią terapii genowej jest zmiana na poziomie genomu za pomocą czynników *cis* lub *trans*. W tym celu można zastosować wektory wirusowe jako nośniki informacji genetycznej (czynniki *cis*), zawartej w formie dwu- lub jednoniciowego DNA lub RNA. (Rys. 2.1.1) W przypadku wektorów wirusowych możemy rozdzielić ich mechanizmy działania na te, które zachodzą bez integracji (wektory adenowirusowe, AdV oraz wektory wirusów związanych z adenowirusami, AVV) lub z integracją terapeutycznej informacji genetycznej z genomem gospodarza (wektory lentiwirusowe i retrowirusowe).^[77, 78] W wypadku pierwszych, informacja genetyczna w formie DNA ulega importowi do jądra i następnej translacji. W wypadku drugich, informacja genetyczna w formie RNA ulega odwrotnej transkrypcji, importowi do jądra, integracji z genomowym DNA i następnej translacji. Błędy na etapie integracji z genomowym DNA mogą prowadzić do genotoksyczności, przykładowo do aktywacji protoonkogenów czy przerwania genów supresji nowotworów.^[71] W celu precyzyjnej zmiany genomu można zastosować metodę CRISPR (ang. *clustered regularly interspaced short palindromic repeats*), nagrodzoną w 2020 roku nagrodą Nobla z chemii.^[79, 80] System CRISPR ma swoje korzenie w mechanizmach obronnych archeonów i bakterii przed infekcją wirusową. Opiera się on na działaniu kompleksu białek Cas i RNA

(Rys. 2.1.2).^[80] Białka Cas1–Cas13 pełnią w kompleksie funkcję białek wiążących, efektorowych, o aktywności helikazy lub nukleolitycznej (jak w wypadku Cas9 i Cas13, które posiadają aktywność DNazy i RNazy odpowiednio). W skład kompleksu wchodzi jedna (sgRNA ang. *single-guide RNA*) lub dwie cząsteczki RNA (crRNA:tracrRNA, ang. *CRISPR RNA (crRNA)–trans-activating CRISPR RNA (tracrRNA) complex*), których celem jest wiązanie komplementarnej sekwencji DNA lub RNA (np. celu terapeutycznego) oraz „rusztowanie” dla składników białkowych. Aktywny kompleks jest zdolny do selektywnego cięcia sekwencji wskazanego genu, po którym może dojść do edycji pojedynczych par zasad, a nawet wstawienia lub wycięcia większego obszaru DNA. W ten sposób można korygować mutacje, wyciszać lub wzmacniać ekspresję istniejących genów.

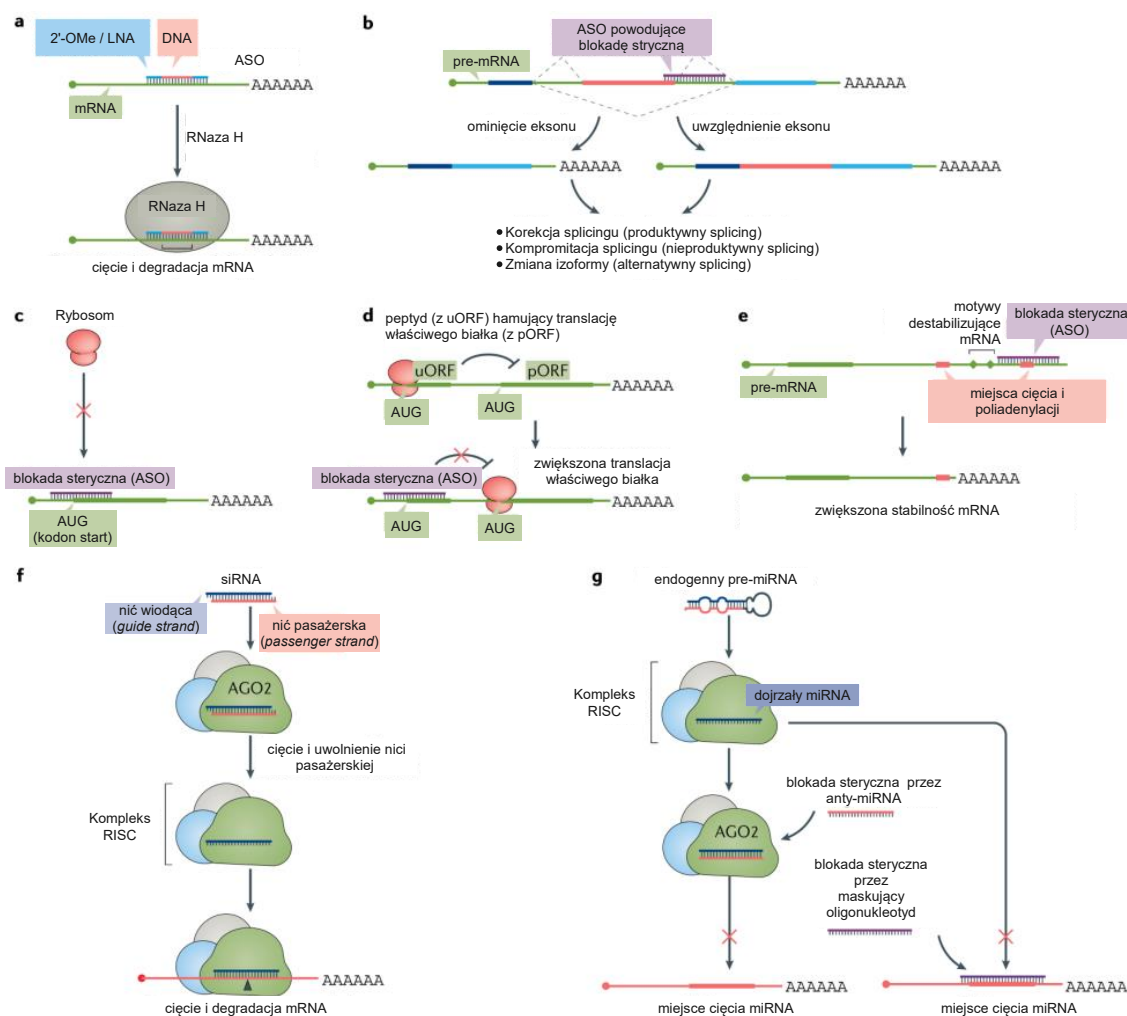


Rysunek 2.1.1: Schematy mechanizmów wprowadzania informacji genetycznej z udziałem wektorów lentiwirusowych lub AAV. 1–6 kolejne etapy uwalniania, obróbki i ekspresji wprowadzanego genu. Na podstawie publikacji w trybie wolnego dostępu ^[78].



Rysunek 2.1.2: Schematy naturalnych kompleksów CRISPR. TniQ: regulatorowe białko wiążące DNA; PAM: sekwencja DNA bezpośrednio za celem cięcia (ang. *protospacer adjacent motif*). Na podstawie publikacji w trybie wolnego dostępu^[80].

Kolejną strategią terapii genowych jest regulacja genomu lub transkryptomu za pomocą oligonukleotydów (kwasów nukleinowych o długości poniżej stu nukleotydów) o charakterze czynników *trans* i działaniu przejściowym (Rys. 2.1.3).^[70] Wyróżnia się dwa główne typy terapeutycznych oligonukleotydów: antysensowne (ASO, ang. *antisense oligonucleotides*) oraz analogi krótkiego interferującego RNA (siRNA, ang. *small interfering*). ASO mogą posłużyć do posttranskrypcyjnego wyciszania genów poprzez wiązanie docelowej sekwencji mRNA i promowanie endonukleolitycznego cięcia katalizowanego przez RNazę H. Ponadto, poprzez komplementarność i selektywne wiązanie, mogą tworzyć zawadę steryczną w wybranym rejonie kodującego RNA. W ten sposób można przeprowadzić „korektę” na etapach splicingu i poliadenylacji pre-mRNA lub hamować translację mRNA zainicjowaną na przedwczesnym kodonie start. Z drugiej strony, analogi siRNA mogą wchodzić w funkcję ich naturalnych odpowiedników (promować endonukleolityczne cięcie wskazanego mRNA) lub działać w przeciwnym kierunku, czyli stymulować translację wybranego mRNA na drodze inhibicji endonukleolitycznego cięcia z udziałem miRNA i kompleksu RISC. Ze względu na niewielki rozmiar (typowo 20-30 nt), terapeutyczne oligonukleotydy otrzymywane są na drodze syntezy chemicznej. To pozwala opatrzyć je w szereg modyfikacji w obrębie zasad azotowych, reszty cukrowej, połączenia monomerów czy wyposażyć w dodatkowe grupy funkcyjne takie jak białka i peptydy (np. przeciwciała), lipidy i steroidy (np. cholesterol) czy sacharydy (np. N-acetylogalaktozaminą, GalNAc). Modyfikacje chemiczne pozwalają zwiększyć trwałość oligonukleotydów na czas transportu i działania w organizmie, zwiększyć powinowactwo i selektywność względem danego celu czy zwiększyć wydajność i selektywność dostarczania do wybranych komórek.



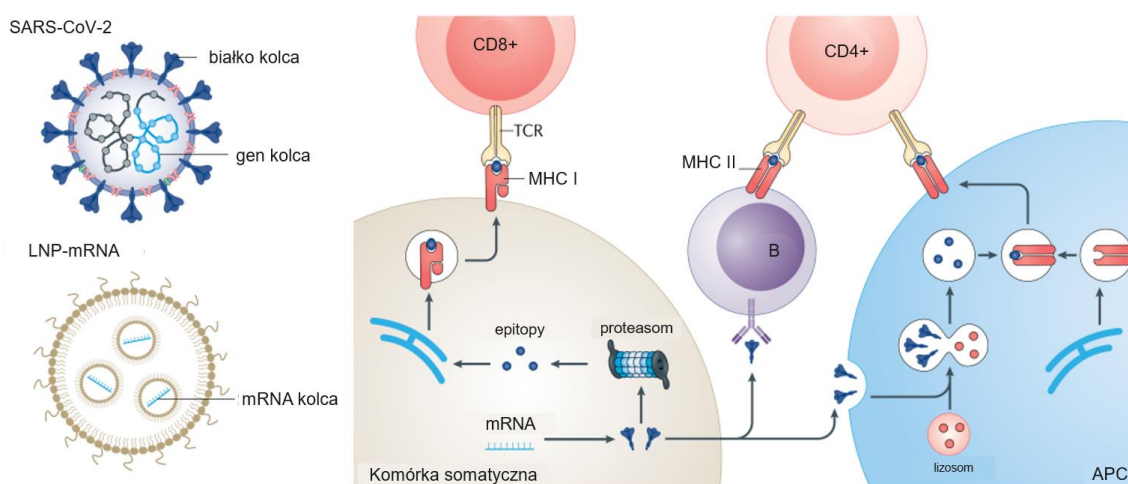
Rysunek 2.1.3: Mechanizmy posttranskrypcyjnej regulacji ekspresji genów za pomocą terapeutycznych oligonukleotydów. Na podstawie [70].

Terapeutyczne mRNA i circRNA

Z perspektywy niniejszej rozprawy, najistotniejszą strategią terapii genowych jest dostarczenie genu za pomocą kodującego RNA (tj. mRNA lub circRNA), czyli ingerencji w transkryptom za pomocą czynnika *cis* w trybie przejściowym. Z perspektywy historycznej, badania w kierunku terapeutycznego zastosowania kodującego RNA rozpoczęły się już w latach sześćdziesiątych dwudziestego wieku, wraz z odkryciem informacyjnego RNA.^[73, 81, 82] Pierwsze próby zastosowań rozpoczęły się w latach pięćdziesiątych, a w 2002 roku pierwsze badania kliniczne immunoterapii *ex vivo* z udziałem mRNA. Najbardziej dynamiczny okres rozwoju dziedziny można zaobserwować w drugiej dekadzie dwudziestego pierwszego wieku. To wtedy dochodzi do dynamicznego rozwoju technologii produkcji terapeutycznego mRNA w skali przemysłowej, wzrostu liczby patentów i badań klinicznych, pojawienia się nowych rozwiązań technologicznych (np. udoskonalonych metod dostarczania, sprzężenia z metodologią CRISPR, zastosowania

samopowielających mRNA czy circRNA), oraz zatwierdzenia i masowej produkcji szczepionek mRNA przeciw SARS-CoV-2.^[83, 84]

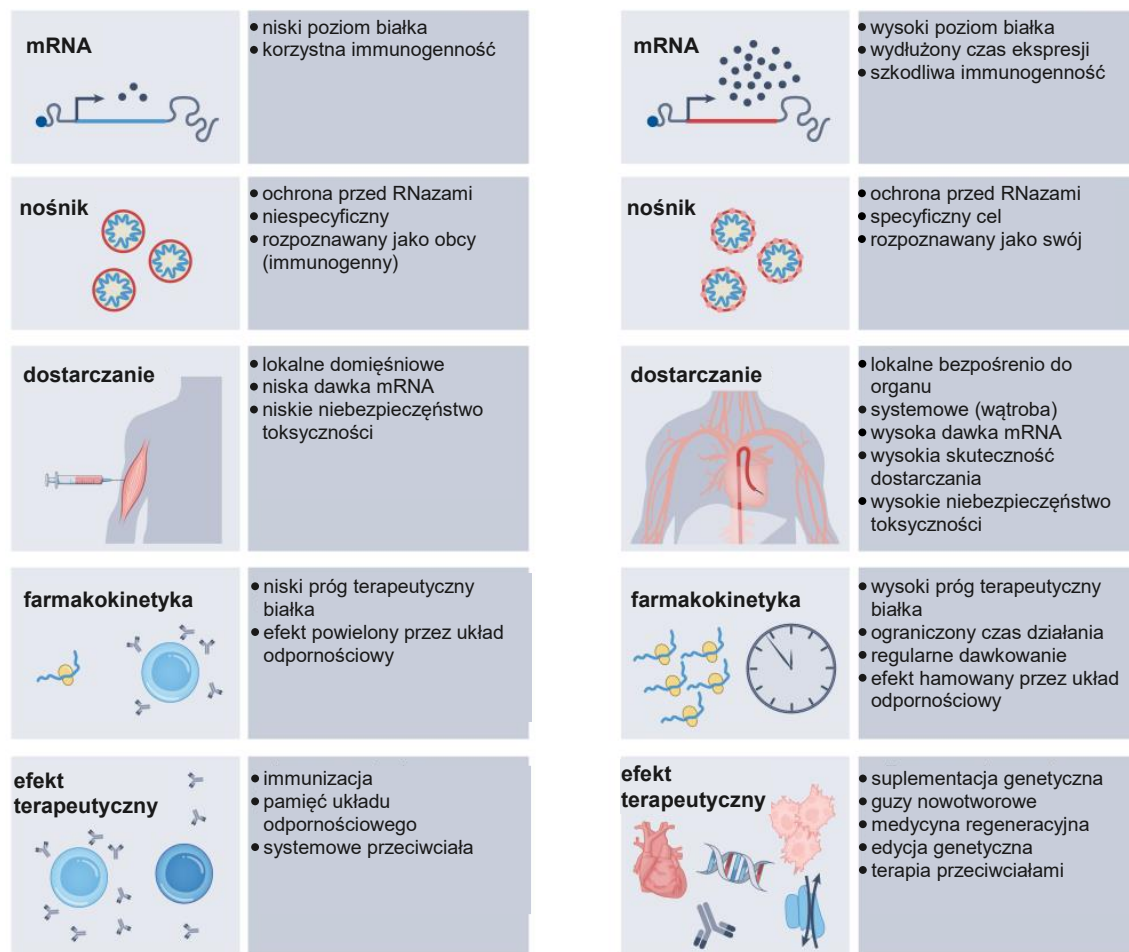
W przypadku szczepionek RNA, kodowana informacja genetyczna służy do immunizacji organizmu pacjenta: jest wychwytywana i następnie wykorzystywana do sformułowania wtórej odpowiedzi immunologicznej. Innymi słowy, efekt dostarczonej informacji genetycznej ulega zwielokrotnieniu i utrwaleniu za sprawą układu odpornościowego pacjenta. Z tego powodu w szczepionkach stosuje się stosunkowo niskie dawki RNA oraz jedno- lub kilkukrotne dawkowanie, co niesie niskie ryzyko toksyczności.^[72, 85] Przykładem są szczepionki mRNA przeciw SARS-CoV-2, kodujące wirusowe białko kolca (Rys.2.1.4).



Rysunek 2.1.4: Schemat działania immunizacji za pomocą szczepionki mRNA przeciw SARS-CoV-2. Po wstrzyknięciu domięśniowym, lipidowe nanocząstki zawierające mRNA (LNP-mRNA) są internalizowane przez komórki somatyczne (np. komórki mięśniowe) i komórki prezentujące antygen (APC, ang. *antigen presenting cells*). Ponadto, LNP-mRNA mogą koncentrować się w węzłach chłonnych, gdzie przebywają inne komórki układu odpornościowego. Antygeny (białka kolca, ang. *spike protein*) wyprodukowane w cytoplazmie są degradowane przez proteasomy, a powstałe epitopy są prezentowane głównie przez białka MHC klasy I (ang. *major histocompatibility complex*) limfocytom T CD8+. Antygeny mogą być również endocytowane przez komórki APC, gdzie są degradowane w lizosomach i dalej prezentowane limfocytom T CD4+ za pośrednictwem białek MHC klasy II. Ponadto, wydzielane antygeny mogą być internalizowane przez receptory limfocytów B i dalej przetwarzane w celu prezentacji limfocytom T CD4+. TCR: receptor komórek T. Na podstawie publikacji w trybie wolnego dostępu ^[82].

Poza zastosowaniem do masowej immunizacji, można wyróżnić szereg typów „klasycznych” terapii genowych wykorzystujących kodujące RNA: suplementacja genetyczna (np. dostarczanie genów białek będących w niedomiarze), edycja genów (poprzez sprzężenie z metodologią CRISPR), medycyna regeneracyjna (np. przeprogramowanie komórek macierzystych), immunoterapia nowotworów oraz terapia

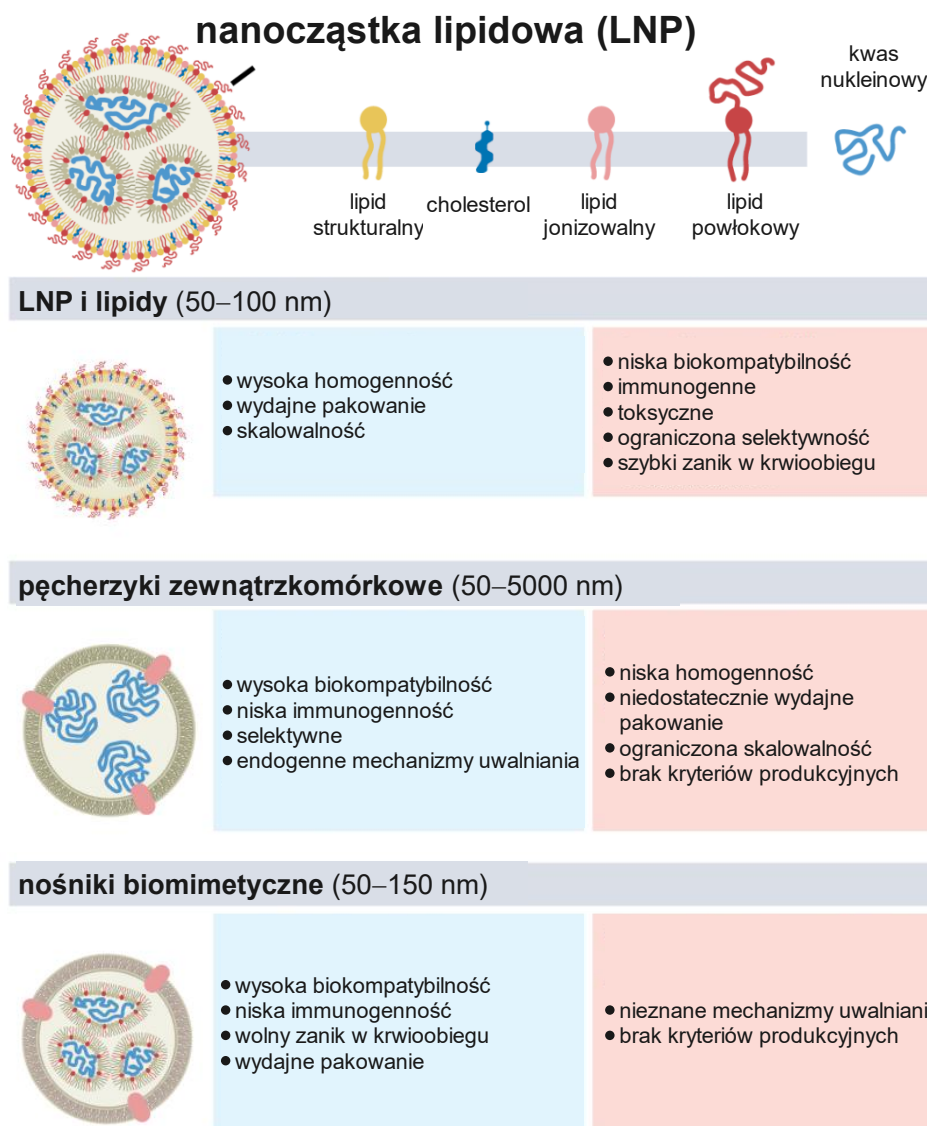
przeciwciałami (np. chorób zakaźnych) (Rys. 2.1.5).^[72, 73, 82] Osiągnięcie wyżej wymienionych efektów terapeutycznych wymagają znacznie (50–1000 krotnie) wyższych poziomów wyrażanego (produkowanego, ulegającego ekspresji) białka, wielokrotnego lub regularnego dozowania oraz metod selektywnego dostarczania terapeutycznego RNA. To zwiększa ryzyko toksyczności oraz wystąpienia efektów ubocznych.^[72]



Rysunek 2.1.5: Specyfika i zastosowanie szczepionkowych i terapeutycznych mRNA. Na podstawie publikacji w trybie wolnego dostępu^[72].

Ze względu na ujemny ładunek, kwasy nukleinowe nie penetrują błon komórkowych. Są też stale narażone na działanie nukleaz obecnych w komórkach i krwiobiegu. W celu ochrony przed degradacją oraz pokonania biologicznych barier, opracowano szereg metod dostarczania, opierających się na opakowaniu (ang. *packaging*) RNA za pomocą związków o charakterze lipidowym lub o strukturze błony lipidowej (Rys. 2.1.6).^[72, 73, 82] Opakowane RNA mogą być przechowywane i stosowane w formach zawiesin, które można wprowadzić do organizmu pacjenta (np. poprzez zastrzyk lub cewnikowanie). Po dotarciu do celu, w wyniku kontaktu opakowania z błoną komórkową, RNA jest uwalnianie do cytoplazmy na drodze zależnej od rodzaju opakowania. Dalej, RNA jest procesowane w

zależności od swojej struktury i sekwencji. Najczęściej stosowanym i najbardziej kompaktowym (50–100 nm) rodzajem opakowania są nanocząstki lipidowe (LNPs, ang. *lipid nanoparticles*). Ich formułowanie (ang. *formulation*) polega na kontrolowanym mieszanii roztworu wodnego, zawierającego RNA z roztworem alkoholowym, zawierającym mieszkankę lipidów. Mieszkanka może zawierać lipidy strukturalne, kationitowe lub jonizowane (zapewniają ładunek dodatni i wiązanie RNA), powłokowe (np. zawierające reszty glikolu polietylenowego) oraz dodatki (np. cholesterol). LNPs są udoskonalane już od sześciu dekad pod kątem minimalizacji immunogenności, zwiększania biokompatybilności, stabilności nanocząstek w roztworze, skuteczności uwalniania RNA oraz powtarzalności i skali ich produkcji. Rozwijane są również alternatywne nośniki kwasów nukleinowych. Przykładem są pęcherzyki zewnątrzkomórkowe (EVs, ang. *extracellular vesicles*). Zaliczają się do nich, naturalnie wydzielane przez komórki struktury takie jak egzosomy (50–150 nm) i ektosomy (50–500 nm). Są one, między innymi, odpowiedzialne za komunikację międzykomórkową. Są też atrakcyjnymi środkami transportu substancji terapeutycznych ze względu na wysoką biokompatybilność oraz niską immunogenność. W połączeniu ze zdolnością do ukierunkowanego dostarczania (pełna selektywność nigdy nie jest możliwa), są atrakcyjnymi nośnikami terapeutycznego RNA. Głównymi problemami związanymi z klinicznym zastosowaniem EVs jest niska wydajność i powtarzalność „pakowania”. Obiecującą alternatywą są opakowania biomimetyczne, będące hybrydą naturalnych i syntetycznych nośników. Przykładem są nośniki zbudowane z syntetycznego rdzenia (np. LNPs, metaliczne lub niemetaliczne nanocząstki) i otoczki naturalnego pochodzenia (np. EVs).

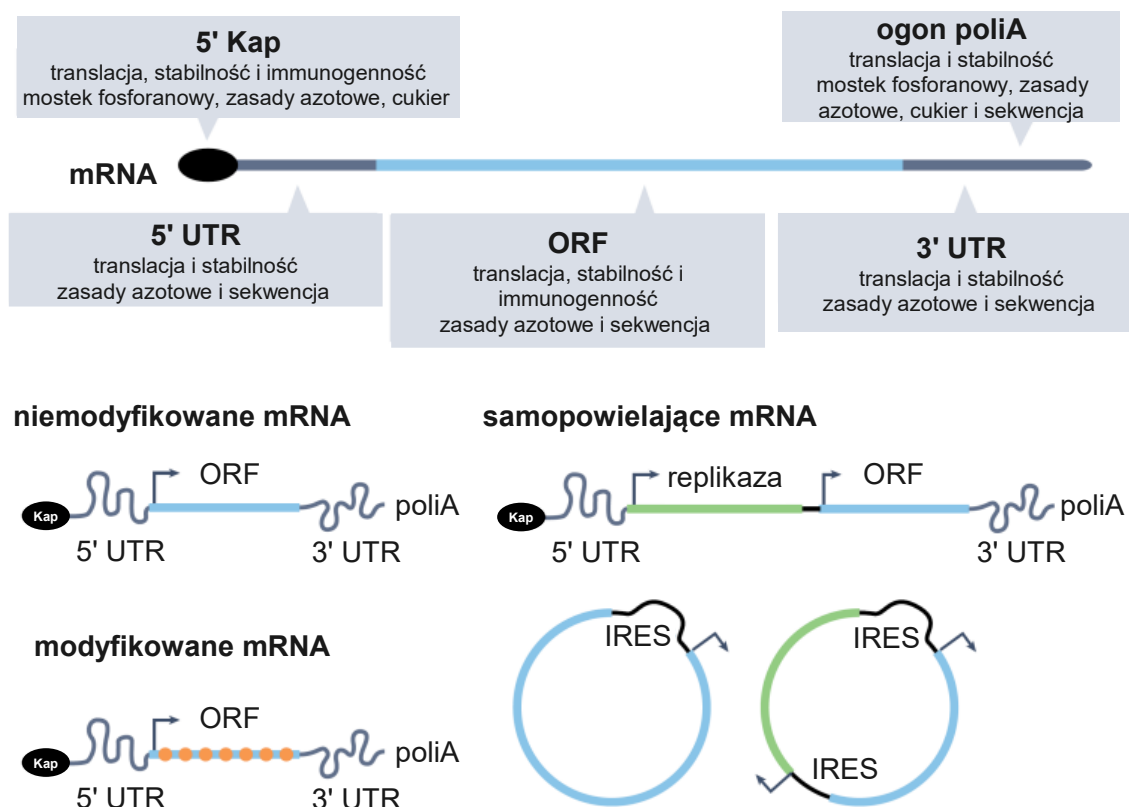


Rysunek 2.1.6: Rodzaje i właściwości nośników (opakowań) terapeutycznego RNA. Na podstawie publikacji w trybie wolnego dostępu [72].

Kolejnym istotnym aspektem terapii genowych jest projektowanie i produkcja kodujących cząsteczek RNA. Są one projektowane jest zarówno na poziomie struktury chemicznej, przestrzennej oraz sekwencji w celu zwiększenia ich stabilności, modulacji ich immunogenności, optymalizacji kinetyki i wydajności ich translacji (Rys. 2.1.7). [72, 73, 86]

W strukturze chemicznej wprowadza się modyfikacje, które obejmują zasady azotowe, reszty cukrowe i wiązania fosfodiesterowe łańcucha głównego, ogona poliA oraz kapu. [32, 73, 87, 88] W szczególności, stosowane są: modyfikacje tiofosforanowe, boranofosforanowe i metylenofosfonianowe, zmniejszające podatność na hydrolizę nukleolityczną; alkilowanie i benzylowanie nukleofilowych atomów zasad i rybozy, które zmniejszają immunogenność i mogą promować translację; niekanoniczne zasady azotowe, jak pseudourydyna i N1-metylopseudourydyna, które również wpływają na translację i

immunogenność kodującego RNA. Inną formą modyfikacji chemicznej jest cyklizacja, rozumiana jako kowalencyjne połączenie końców 5' i 3'.^[55, 57, 58] Koliste cząsteczki RNA wykazują wydłużoną ekspresję kodowanego genu, co może być korzystne z perspektywy farmakokinetyki.



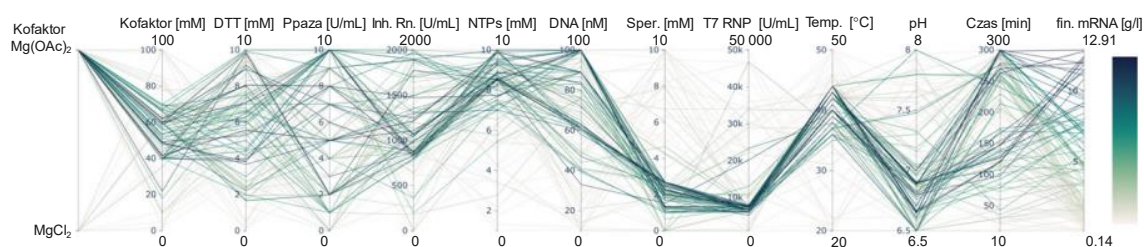
Rysunek 2.1.7: Schemat kodującego RNA jako nośnika terapeutycznej informacji genetycznej. Wyróżniono na nim elementy takie jak kap, sekwencje UTR, ORF oraz ogon poliA wraz z wpływem ich modyfikacji na właściwości mRNA. Zobrazowano również podział terapeutycznego mRNA na niemodyfikowane, modyfikowane (w tym koliste) oraz samopowielające. Na podstawie publikacji w trybie wolnego dostępu^[72].

Struktura przestrzenna oraz sekwencja kodującego RNA są parametrami nierozłącznymi oraz równie istotnymi co wyżej wymienione modyfikacje chemiczne. Z perspektywy efektu terapeutycznego kluczowa jest sekwencja otwartej ramki odczytu (ORF, ang. *open reading frame*), zawarta między kodonami start i stop włącznie. Nie tylko przekłada się na sekwencję aminokwasową białka, ale również na kinetykę jego translacji i zwijania. Ze względu na degenerację kodu genetycznego, długość sekwencji ORF oraz poziom skomplikowania efektów mających wpływ na kinetykę translacji (np. dostępność aminoacylo-tRNA w danym środowisku czy struktura przestrzenna RNA) optymalizację sekwencji ORF wykonuje się za pomocą zaawansowanych narzędzi obliczeniowych, takich jak uczenie maszynowe.^[89] Analogiczne złożony związek można zaobserwować

między właściwościami RNA (tj. stabilnością, immunogennością i wydajnością translacji) a niekodującymi rejonami jej sekwencji (UTR, ang. *untranslated region*), co dalej podkreśla nierozłączność struktury przestrzennej (całości) cząsteczki RNA i jej właściwości.^[90] Co do zasady, sekwencje 5' oraz 3' UTR stanowią rejony wiązania regulatorowych elementów *trans*, w szczególności w procesie translacji. Analogiczne funkcje i właściwości można przypisać sekwencjom typu IRES oraz ogonowi poliA, które formalnie są wyodrębniane od grupy UTR (prawdopodobnie ze względów historycznych). Mimo strukturalnej „prostoty” zależność między długością ogona poliA a ilością produkowanego białka nie jest jasna.^[39] Ponadto, prowadzenie modyfikacji ogona poliA (np. poprzez jego segmentację) na poziomie sekwencji lub struktury chemicznej również może mieć pozytywny wpływ z perspektywy zastosowań terapeutycznych.^[91-93] Z drugiej strony, sekwencje IRES charakteryzują się klasyfikacją i historią badań niezależną od terapeutycznego mRNA. Głównym elementem wspólnym jest zastosowanie sekwencji IRES przy projektowaniu circRNA jako nośnika terapeutycznej informacji genetycznej.^[55, 56, 94] Jednym z problemów zastosowania sekwencji IRES jest ich stosunkowo duży rozmiar (często powyżej 500 nt), złożona i tylko częściowo zrozumiana struktura przestrzenna i mechanizm wiązania czynników inicjujących translację. Wydajność translacji IRES-zależnej jest (zazwyczaj) niższa niż kap-zależnej. Niemniej, w odpowiednim kontekście strukturalnym i pod wpływem modyfikacji, sekwencje IRES zawarte w kolistych RNA mogą stanowić atrakcyjną alternatywę dla „klasycznych” terapeutycznych mRNA. Kolejną nowością jest zastosowanie dodatkowej sekwencji kodującej wirusową replikazą (np. białka Nsp1-4 z alfawirusa) w celu dostarczania samopowielających się cząsteczek mRNA (saRNA, ang. *self-amplifying*).^[95] Podobnie jak w wypadku circRNA, zastosowanie saRNA ma na celu wydłużenie czasu, w którym zachodzi ekspresja terapeutycznej informacji genetycznej.

Aktualnie, jedyną wydajną i skalowalną ścieżką syntezy kodującego RNA jest transkrypcja *in vitro* (IVT) i następcza chromatografia.^[84] Do IVT wykorzystane są wirusowe polimerazy DNA-zależne, takie jak polimeraza RNA SP6, T3 czy T7. Ta ostatnia jest najczęściej wybierana oraz modyfikowana na potrzeby minimalizacji produktów ubocznych (tj. dwuniciowego RNA) czy wykorzystania syntetycznych analogów substratów.^[28, 96, 97] Kolejnym elementem mieszaniny IVT jest matrycowe DNA, które można otrzymać na drodze linearyzacji plazmidu lub reakcji polimeryzacji łańcuchowej (PCR, ang. *polymerase chain reaction*).^[98] Są to metody komplementarne: plazmidowe DNA można produkować w dużej skali z udziałem hodowli bakteryjnej, natomiast za

pomocą reakcji PCR można szybko otrzymać więcej różnorodnych sekwencji kodujących. Sekwencja matrycowego DNA, w szczególności promotor i jego kontekst, są istotne z perspektywy optymalizacji produkcji RNA metodą IVT.^[24] Ostateczna ilość i jakość produktów IVT jest zależna od wielu czynników, przez co optymalizacja warunków reakcji jest trudna. Z pomocą mogą przyjść zaawansowane metody analizy danych, przykładowo optymalizacja bayesowska (Rys 2.1.8).^[84] Z jej pomocą można pokierować proces optymalizacji zarówno pod kątem czynników intuicyjnych (np. stężenia NTP, matrycowego DNA, czasu i temperatury reakcji), ale również pod kątem czynników nieintuicyjnych, jak rodzaj kofaktora (znacznie lepszy jest octan niż chlorek manganu (II)), czy wartość pH (optymalna w przedziale 6.6–6.8). Jakość produktów IVT można sprowadzić do homogenności właściwego produktu RNA (w szczególności w ramach struktury końca 5' oraz 3') oraz ilości produktów ubocznych. Homogenność można zwiększyć na drodze wydajnego kapowania i poliadenylacji, które można przeprowadzić zarówno kotranskrypcyjnie jak i posttranskrypcyjnie.^[87, 91, 99] Powstawanie immunogennych produktów ubocznych można ograniczyć za pomocą mutantów polimeraz lub czynników chaotropowych.^[97, 100] Niezależnie od jakości, produkty IVT trzeba wyizolować z mieszanin reakcyjnych w sposób wydajny i łagodny. Klasycznymi technikami jest strącanie soli RNA za pomocą soli litu lub sodu, często w połączeniu z odbiałczaniem na drodze ekstrakcji w układzie fenol/chloroform.^[101] W celu izolacji produktów wysokiej czystości w powtarzalny i skalowalny sposób stosuje się metody chromatograficzne.^[102] Do najbardziej rozpoznawanych należą chromatografia w układzie faz odwróconych (np. RP-HPLC)^[30, 103], chromatografia jonowymienna^[104], sączenie molekularne^[57, 58], chromatografia na złożach powinowactwa (np. oligo-dT)^[105] czy celulozowych^[106].



Rysunek 2.1.8: Optymalizacja bayesowska warunków transkrypcji *in vitro*. Optymalizowane warunki: rodzaj i stężeni kofaktora; stężenie ditiotreitolu (DTT), pirofosfatazy (Ppaza), inhibitora RNaz (Inh. Rn.), trifosforanów rybonukleotydów (NTPs), matrycy (DNA), spermidyny (Sper.), polimerazy RNA T7; temperatury, pH i czasu reakcji. Fin. mRNA: końcowe stężenie produktu (wydajność IVT). Na podstawie publikacji w trybie wolnego dostępu ^[84].

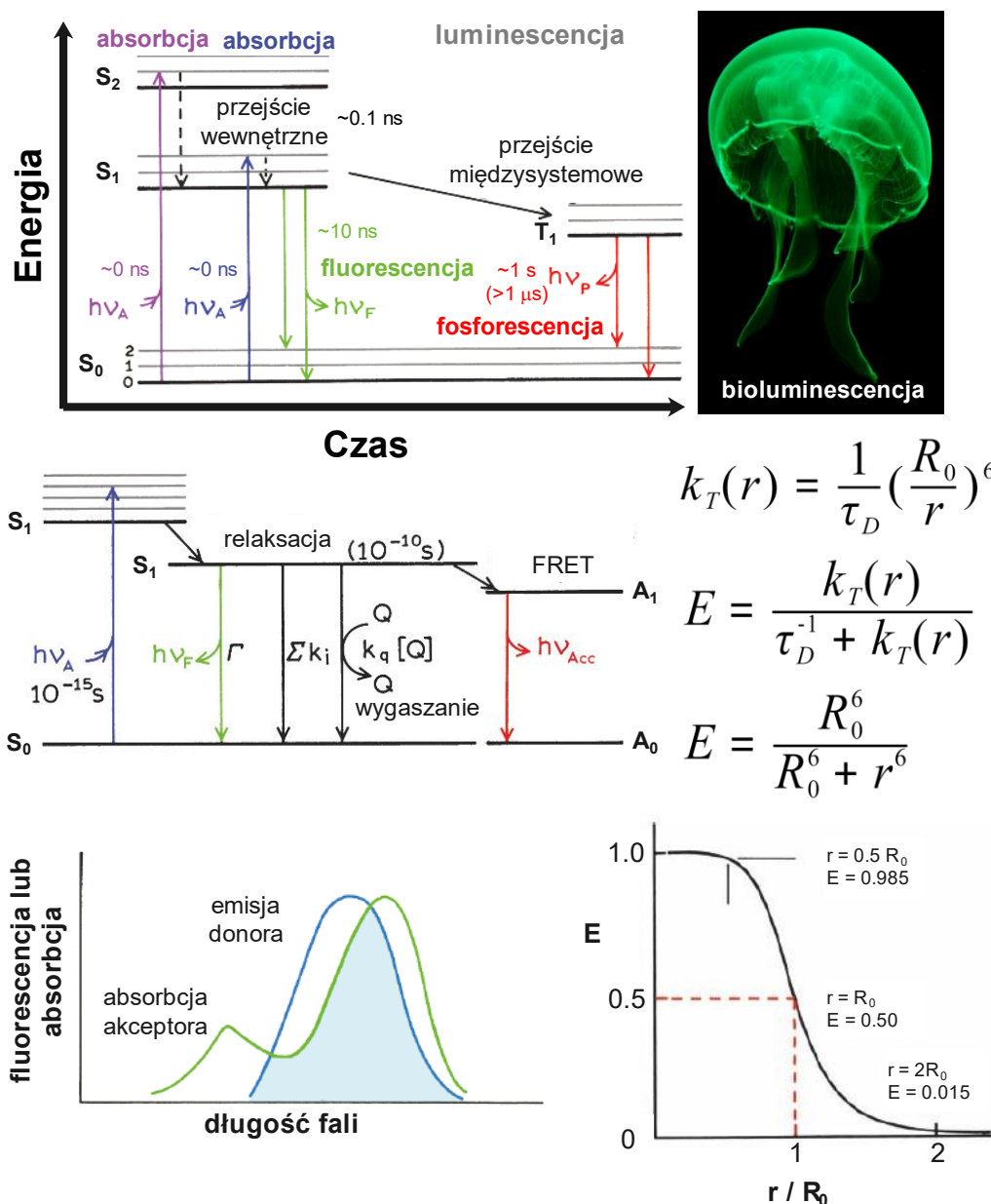
2.2 Bioobrazowanie i sondy molekularne RNA

Zjawisko fluorescencji i rezonansowego transferu energetycznego

Zjawisko podstawowe polega na elektronowej absorpcji i następczej emisji światła. Absorbowane i emitowane fotony odpowiadają różnicy energii między singletowymi stanami wzbudzonymi (S_1 i S_2) a podstawowym (S_0). Tę sytuację opisuje niżej zamieszczony diagram Jabłońskiego (Rys. 2.2.1).^[107] Ze względu na prawdopodobieństwo wzbudzenia do drugiego stanu elektronowego (S_2) oraz na wyższe stany oscylacyjne, jeszcze przed emisją zachodzi przejście wewnętrzne (z S_2 do S_1 , $\sim 10^{-12}$ s). Alternatywą jest przejście międzysystemowe, ze wzbudzonego stanu singletowego do tripletowego (z S_1 do T_1), po którym zachodzi zabronione przejście (z T_1 do S_0) i emisja fosforescencyjna. Fluorescencja i fosforescencja należą do kategorii zjawisk luminescencyjnych, zachodzących zarówno w układach eksperymentalnych jak i naturalnych. Zjawisko fluorescencji charakteryzuje się za pomocą następujących parametrów: widma fluorescencji, czasu życia fluorescencji (τ) oraz wydajności kwantowej (Q). Podstawowemu zjawisku fluorescencji towarzyszą inne, przykładowo wygaszanie czy rezonansowy transfer energii.

Widma absorpcyjne i fluorescencyjne to jedne z bardziej charakterystycznych właściwości materii. Kształty ich obwiedni, będące funkcją współczynnika absorpcji bądź wydajności kwantowej emisji, odpowiadają populacjom fotonów (o danej energii lub długości fali) zaabsorbowanych lub wyemitowanych przez dany typ cząsteczki (fluorofora) lub ich układu. Na widma mają wpływ czynniki wewnętrzne (struktura i właściwości fluorofora) oraz zewnętrzne (np. temperatura, rozpuszczalnik i inne cząsteczki). Widma absorpcji i fluorescencji są najczęściej lustrzanymi odbiciami przemieszczonymi o odległość Stokesa, co wiąże się z wzbudzaniem wyższych stanów oscylacyjnych. Każdy wzbudzony stan wraca spontanicznie do podstawowego S_0 . Średni czas, po którym odnotowujemy to zjawisko to parametr czasu życia fluorescencji τ (~ 10 ns, inaczej czas rozpadu), którego odwrotnością jest suma częstości przejść promienistych Γ (emisji, inaczej częstości rozpadu) oraz częstości przejść bez promienistych (np. na drodze wygaszania). Wydajność kwantowa Q jest stosunkiem fotonów wyemitowanych do zaabsorbowanych. Można ją przestawić jako iloczyn częstości emisji Γ i czasu życia τ ($Q = \Gamma \cdot \tau$). Wydajność kwantowa może być zaniżona przez zjawiska towarzyszące, takie jak wygaszanie. Wyróżnia się wygaszanie statyczne oraz dynamiczne. W wypadku wygaszania statycznego, fluorofor wchodzi w stabilizujące oddziaływania z cząsteczkami

wygaszacza (np. kompleksy). Wygaszanie dynamiczne polega na zderzeniach cząsteczek w stanie S1 i bezpromienistą wymianę energii z otoczeniem. Z tego powodu stopień wygaszania dynamicznego jest zależny od stężenia wygaszacza.



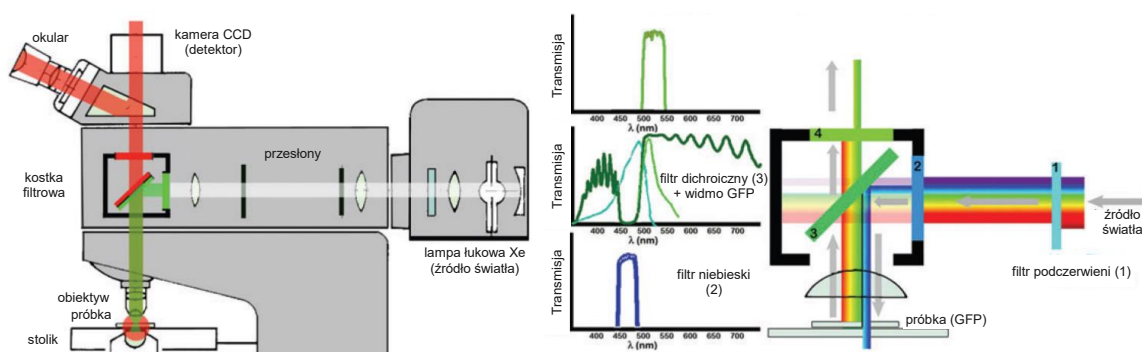
Rysunek 2.2.1: Zjawisko fluorescencji i rezonansowego transferu energetycznego Förstera (FRET). k_i – stałe wygaszania statycznego (różne mechanizmy), k_q – stała wygaszania dynamicznego, Q – wygaszcz. Na podstawie [107].

Kolejnym zjawiskiem towarzyszącym fluorescencji jest rezonansowy transfer energii. Przykładem jest rezonans Förstera (FRET, ang. *Förster resonance energy transfer*), zachodzący z udziałem dalekozasięgowych oddziaływań dipolowych. Do jego zajścia potrzebne są dwa fluorofory (donor i akceptor) o nakładających się widmach emisji i absorpcji (Rys. 2.2.1). Po wzbudzeniu donora może dojść do transferu energii ze stanu S1

i wzbudzenia elektronów akceptora (z A_0 do A_1), bez udziału fotonu. Częstość transferu jest parametryzowana za pomocą stałej częstości transferu (k_T), będącej funkcją odległości między donorem i akceptorem (r). Kwantowa wydajność transferu (E) to stosunek częstości transferu do sumy częstości transferu i częstości emisji donora (k_D lub τ_D^{-1}). Częstość i wydajność transferu są silnie i odwrotnie zależne od odległości między donorem i akceptorem (k_T i $E \sim r^{-6}$). Z tego powodu wprowadzony został termin odległości Förstera (R_0), przy której transfer zachodzi z połowiczną wydajnością. Dzięki niej można wykreślić zależność wydajności transferu E od względnego dystansu między donorem a akceptorem (r/R_0). Wartości R_0 można wyznaczyć eksperymentalnie: najczęściej zawierają się w przedziale 10–60 nm, co odpowiada odległościom w skali cząsteczkowej. Dzięki temu, po odkryciu ponad sto lat temu, zjawisko FRET znalazło zastosowanie w badaniach jako „cząsteczkowa linijka”.^[108] Poza względną odległością donora i akceptora, na wydajność transferu mają wpływ: wydajność kwantowa donora, powierzchnia nakładania widm donora i akceptora oraz wzajemna orientacja tych cząsteczek. Analogicznie jak w wypadku absorpcji spolaryzowanego światła, rezonans między donorem i akceptorem jest uzależniony od orientacji ich dipolowych momentów przejścia, które są wektorami zdefiniowanymi w przestrzeni struktury cząsteczki. Transfer jest najwydajniejszy, kiedy wektory są współliniowe ($\rightarrow\rightarrow$), mniej wydajny kiedy są równoległe ($\uparrow\uparrow$), a nawet całkowicie zahamowany kiedy są prostopadłe ($\uparrow\rightarrow$).

Zjawisko fluorescencji można wykorzystać w obserwacjach mikroskopowych układów biologicznych.^[107, 109] Limit rozdzielczości klasycznego mikroskopu optycznego to połowa długości widzialnej fali świetlnej (200–400 nm), co umożliwia obserwacje w skali komórek i struktur stukomórkowych. Sytuacja ulega zmianie, gdy preparat zawiera cząsteczki zdolne do emisji światła. Analogicznie jak w wypadku teleskopii gwiazd, za pomocą mikroskopii fluorescencyjnej możliwe jest rejestrowanie obrazów o wysokim kontraście (stosunku sygnału do tła) oraz rozdzielczością na poziomie pojedynczych cząsteczek. W tym celu, preparat umieszczony na stoliku mikroskopu musi zostać oświetlony wiązką światła wzbudzającego, by następnie światło emitowane w wyniku fluorescencji próbki zostało zebrane przez obiektyw (Rys. 2.2.2). Charakterystycznym elementem mikroskopu fluorescencyjnego jest kostka filtrowa (ang. *cube filter*), która zawiera trzy filtry: prostopadłe filtry światła wzbudzającego i emitowanego (na ściankach kostki), oraz filtr dichroiczny (inaczej zwierciadło dichroiczne) na przekątnej kostki. Filtry prostopadłe przepuszczają światło o ograniczonym zakresie długości fali (np. filtr światła niebieskiego 450–500 nm lub zielonego 500–550 nm). Zwierciadło dichroiczne odbija

część pasma widzialnego (np. światło niebieskie) a pozostałą część (np. światło w zakresie od zieleni do czerwieni) przepuszcza. Ten prosty, niemniej kluczowy element układu optycznego mikroskopu pozwala na selektywne wzbudzenie preparatu i jednoczesną selekcję emitowanych fotonów (spośród wszystkich zebranych przez obiektyw). Jednym z kamieni milowych dla mikroskopii fluorescencyjnej jako narzędzia do śledzenia procesów komórkowych było odkrycie i następnie wykorzystanie fluorescencyjnych białek zaangażowanych w bioluminescencję (np. zielone białko fluorescencyjne, GFP ang. *green fluorescent protein*).



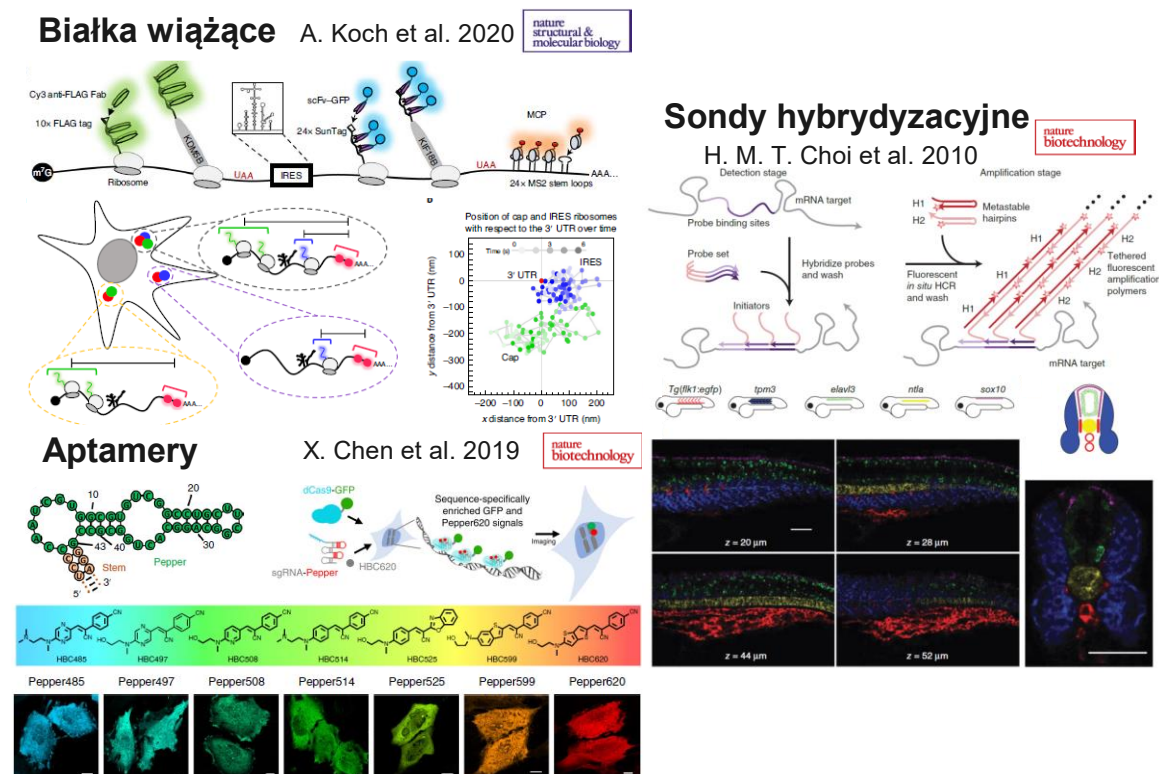
Rysunek 2.2.2: Budowa klasycznego mikroskopu fluorescencyjnego. Na podstawie [109].

Bioobrazowanie mRNA

Wiele procesów komórkowych, w których zaangażowane jest mRNA zostało lepiej zrozumianych za sprawą obrazowania metodą mikroskopii fluorescencyjnej. W tym celu, można zastosować jedną z wielu metod znakowania i detekcji mRNA (Rys. 2.2.3). W szczególności można zastosować fluorogeniczne aptamery, koniugaty białek fluorescencyjnych i wiążących mRNA, sondy hybrydizacyjne lub kowalencyjną modyfikację RNA (w tym sondy molekularne RNA). Ostatni wymieniony zbiór metod został omówiony w następnym rozdziale.

Fluorogeniczne aptamery to struktury RNA zdolne do wiązania małych cząsteczkowych fluoroforów, których wydajność kwantowa fluorescencji znacznie wzrasta w wyniku wiązania przez aptamer. Aptamery mogą być wykorzystane zarówno w syntetycznych sekwencjach mRNA jak i mogą zostać wprowadzone do sekwencji endogennego genu jako element reporterowy.^[110, 111] Odkryto oraz zsyntezowano wiele aptamerów, w tym fluorogenicznych o „warzywnych” nazwach: *broccoli*, *spinach* (emisja zielonego światła) oraz *corn* (emisja żółtego światła). Wyróżniono pracę, w której opisano otrzymanie drogą ewolucji *in vitro* (metoda SELEX) aptameru *pepper* zdolnego do silnego wiązania analogów HBC (cząsteczki zaprojektowanej na bazie fluorofora białka GFP).^[111]

Zaprezentowano również zastosowanie aptameru *pepper* do lokalizacji mRNA czy kompleksu CRISPR-Cas9.



Rysunek 2.2.3: Bioobrazowanie mRNA i znakowanie niekowalencyjne: zastosowania białek wiążących RNA, fluorogenicznych aptamerów oraz sond hybrydizacyjnych w mikroskopii fluorescencyjnej. Na podstawie [111-113].

Za złoty standard wśród wyżej wymienionych metod uznaje się zastosowanie koniugatów białek wiążących RNA i fluorescencyjnych.^[112, 114] Jednym z najstarszych jest system oparty na zastosowaniu białek powłokowych bakteriofaga MS2 wiążących fragmenty sekwencji RNA kodującej genom wirusa. Ze względu na konieczność wprowadzenia sekwencji reporterowej w ramach docelowego mRNA (najczęściej dwadzieścia cztery powtórzenia w obszarze niekodującym) oraz wyrażania transgenicznego białka reporterowego (hybrydy białka wiążącego i fluorescencyjnego), systemy tego typu znajdują zastosowanie głównie w detekcji endogennych mRNA. Wyróżnionym przykładem jest praca, w której dokonano transfekcji komórek z wykorzystaniem bicistronowego konstrukt, prowadzącej do ekspresji mRNA zawierającego strukturę kapu i ogon poliA, sekwencję IRES, sekwencję 3' UTR zawierającą spinki wiążące MS2 oraz sekwencje ORF kodujące reporterowe antygeny.^[112] Wraz z fluorescencyjnymi hybrydami białek (przeciwciał lub MS2) zarejestrowano zmiany położenia łańcucha mRNA i porównano kinetykę translacji zależnej od kapu i translacji

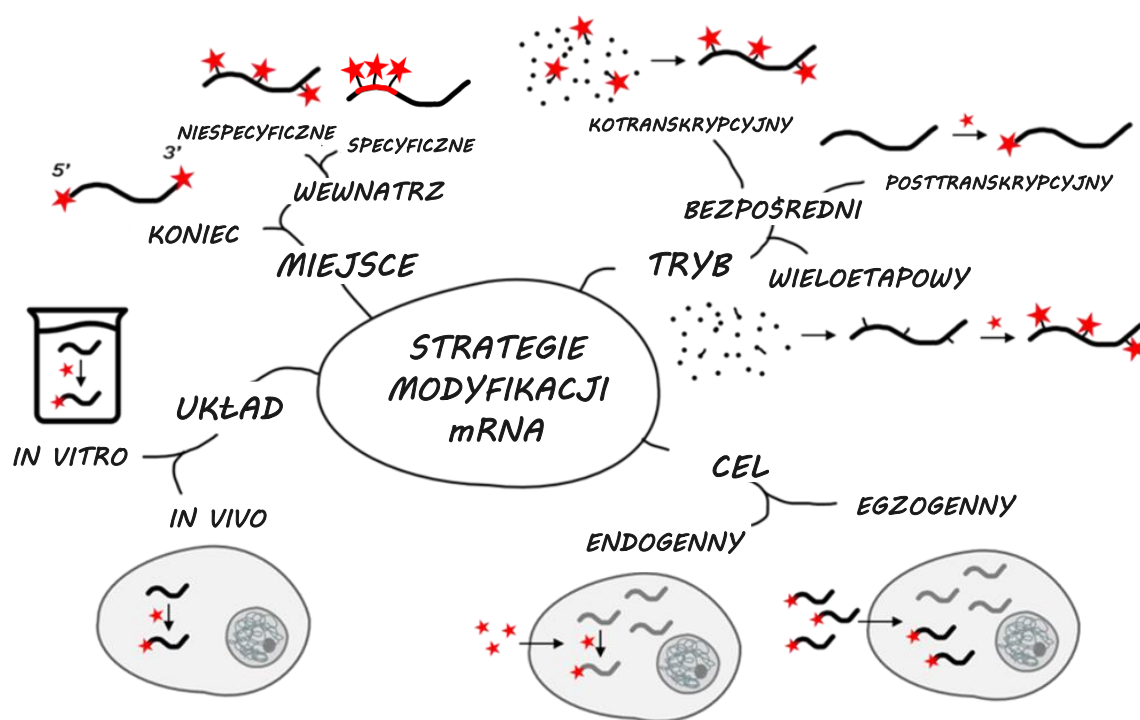
zależnej od sekwencji IRES. Co ciekawe, autorzy pracy nie zaobserwowali znacznych różnic między częstościami inicjacji translacji zależnej od kapu i IRES (częstość obu to ok. raz na 100 sekund). Zaobserwowali jednak, iż przejście ze stanu nieaktywnego translacyjnie do aktywnego (aktywacja translacji, ok. raz na 10 000 sekund) zachodzi o rząd wielkości rzadziej niż przejście odwrotne (ok. raz na 1 000 sekund). Wyjątkiem były mRNA ulegające translacji kap-zależnej (aktywowane): w ich wypadku aktywacja (równoległej) translacji IRES-zależnej zachodziła o rząd wielkości częściej (ok. raz na 1 000 sekund). Sugeruje to istnienie związku między translacją zależną od kapu oraz translacją zależną od IRES (pierwsza pozytywnie wpływa na drugą).

Sondy hybrydyzacyjne to najczęściej syntetyczne oligonukleotydy kowalencyjnie połączone z barwnikami fluorescencyjnymi (np. rodaminą lub cyjaniną). Sekwencja sondy jest projektowana tak by oddziaływała z wybraną sekwencją docelową mRNA (np. metoda FISH, ang. *fluorescent in situ hybridization*). Sondy hybrydyzacyjne znalazły zastosowania między innymi w lokalizacji endogennych mRNA czy śledzenia dostarczania syntetycznego mRNA *in vitro*.^[113, 115, 116] Wyróżniono przykładową pracę, w której wykorzystano szereg sekwencji oligonukleotydowych, które w wyniku hybrydyzacji z docelowym mRNA tworzyły „punkt zaczepienia” dla fluorescencyjnych oligonukleotydów o strukturze metastabilnej spinki.^[113] Tego typu strategia pozwoliła na wiązanie wielu fluoroforów z docelowym mRNA i lokalne zwiększenie intensywności emisji. Ze względu na mechanizm hybrydyzacji możliwe było zaprojektowanie szeregu sekwencji oligonukleotydowych, które wykorzystano do równoległego zobrazowania dystrybucji tkankowej ekspresji pięciu genów w organizmie modelowym (Danio pręgowanym, *Danio rerio*).

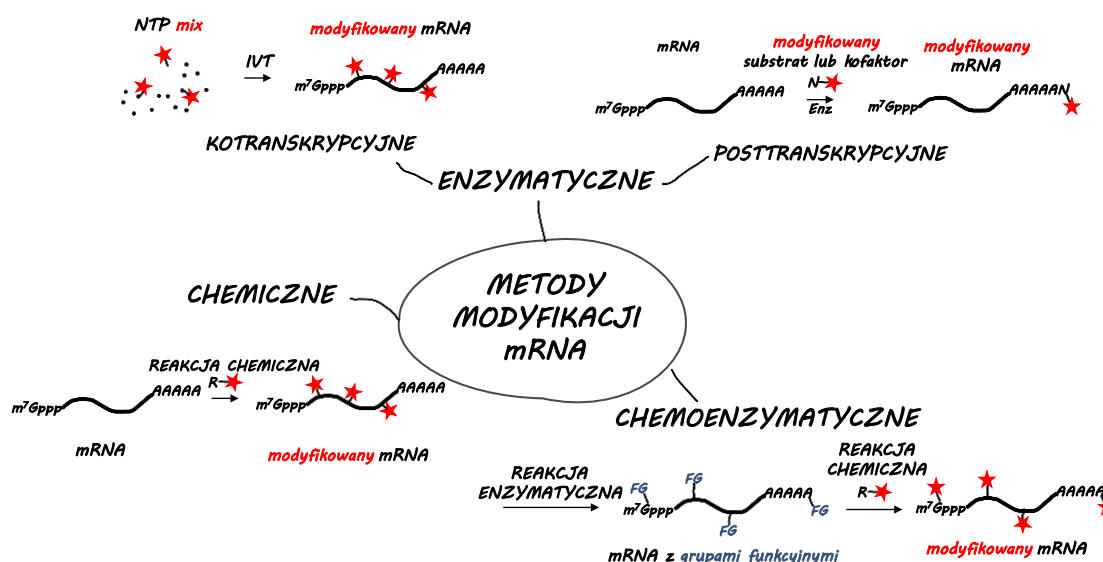
2.3 Metody modyfikacji kodującego RNA

Podział strategiczny i metodologiczny

Opublikowane metody chemicznej modyfikacji (w tym kowalencyjnego znakowania) RNA można klasyfikować według szeregu kryteriów. W ramach niniejszej rozprawy wyróżniono dwa systemy klasyfikacji: podział strategiczny oraz podział metodologiczny. Podział strategiczny uwzględnia: miejsce modyfikacji (region cząsteczki RNA); przebieg procesu modyfikacji oraz układ, w którym zachodzi; pochodzenie substratu RNA (Rys. 2.3.1). Podział metodologiczny uwzględnia trzy główne kategorie: metody enzymatyczne, chemiczne oraz chemoenzymatyczne (Rys. 2.3.2). W Tabeli 1 dokonano przeglądu prac z lat 1978–2021, w których wykorzystano wyżej wymienione metody w celu znakowania RNA lub syntezy sond molekularnych. W przeglądzie pominięto przykłady chemicznej syntezy oligonukleotydów RNA ze względu na ograniczone zastosowanie w dziedzinie modyfikacji kodującego RNA.



Rysunek 2.3.1: Strategiczny podział przykładów modyfikacji i znakowania mRNA.



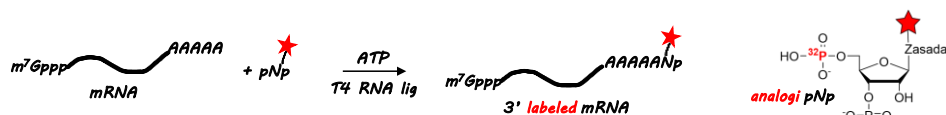
Rysunek 2.3.2: Metodologiczny podział przykładów modyfikacji i znakowania mRNA.

Cząsteczka RNA może być znakowana wewnątrz lub na końcach 5' i 3', w obrębie łańcucha, reszty cukrowej lub zasad azotowych. Wewnętrzną modyfikację można przeprowadzić zarówno w sposób zależny od sekwencji (specyficzny) lub niezależny od niej (niespecyficzny). Modyfikację można przeprowadzić bezpośrednio, w trakcie transkrypcji (kotranskrypcyjnie) lub po jej zakończeniu (posttranskrypcyjnie), lub w procesie wieloetapowym. Procedura modyfikacji może być przeprowadzona zarówno w syntetycznym układzie eksperymentalnym (*in vitro*), jak i w hodowlach komórkowych lub w organizmach (*in vivo*) na endogennym lub egzogennym RNA.

Największym zbiorem metod jest modyfikacja na drodze katalizy enzymatycznej. (Rys. 3.2.3) W ogólności, modyfikacja (podstawnik, np. znacznik fluorescencyjny) może być wprowadzona kotranskrypcyjnie lub posttranskrypcyjnie przez ligazę, polimerazę lub transferazę. W tym celu, enzym wykorzystuje syntetyczny analog substratu lub kofaktora, który zawiera pożądaną modyfikację w swojej strukturze (np. analogi nukleotydów połączone z barwnikiem fluorescencyjnym), co może prowadzić do zmniejszenia aktywności enzymu. Najczęściej wykorzystywane są enzymy wirusowe (np. bakteriofagów T4 i T7), często zmutowane w celu zwiększenia aktywności w obecności nienaturalnego substratu lub kofaktora. Ligaza RNA I T4 i z nukleotydowe substraty pNp (5',3'-difosforan nrybonukleozydu) były pierwotnie wykorzystywane do radioznakowania RNA na końcu 3'.^[117] W późniejszych pracach wykorzystano analogi pNp zawierające znaczniki (np. barwniki fluorescencyjne lub biotyne) lub bioortogonalne grupy chemiczne połączone z zasadą azotową lub 3'-fosforanem cząsteczki pNp.^[118-123] Ligaza RNA I T4 posiada również

aktywność 5' adenylazy, co zostało zastosowane do modyfikacji końca 5' RNA z wykorzystaniem analogów ATP.^[124] Modyfikacja z wykorzystaniem polimeraz zakłada zastosowanie analogów NTP modyfikowanych w obrębie zasady azotowej, rybozy lub fosforanów. Polimerazy można zastosować zarówno do modyfikacji kotranskrypcyjnej jak i posttranskrypcyjnej. W przypadku modyfikacji kontraskrypcyjnej, z udziałem DNA-zależnej polimerazy RNA (np. T7, T3 lub SP6) modyfikacja może odbyć się na etapie inicjacji (w obrębie końca 5') lub elongacji (wewnętrzna, niespecyficzna). Kotranskrypcyjna modyfikacja 5' odbywa się z udziałem inicjatora, do których zaliczamy analogi 5'-fosforanów nukleotydów (NMP), 5',3'-dinukleotydów (NpN), dinukleotydowych oraz trinukleotydowych analogów kapu (m^7GpppN oraz $m^7GpppNpN$).^[118, 122, 123, 125-131] Do modyfikacji w fazie elongacji, najczęściej stosuje się analogi zawierające modyfikacje w obrębie zasad pirymidynowych (analogi UTP i CTP).^[132-139] Alternatywnie, możliwe jest wprowadzenie modyfikacji łańcucha fosfodiesterowego (np. borano- lub trifosforanowych) z wykorzystaniem analogów ATP.^[87] Do modyfikacji posttranskrypcyjnej można zastosować enzymy o aktywności polimerazy takie jak: polimeraza poliA (PAP), polimeraza θ , polimeraza DNA (Klenow) czy syntetyczne rybozomy.^[87, 137, 140-147] W tych wypadkach modyfikacja sprowadza się do wydłużenia RNA od końca 3' z udziałem analogów NTP. Analogiczny efekt uzyskano poprzez zastosowanie analogów UTP i terminalnej urydylotransferazy (TUTazy).^[148] Innymi transferazami, wykorzystywanymi do modyfikacji RNA, są enzymy kapujące (np. VCE), metylotransferazy (np. GluTgs, Emc1, CAPAM, METTL) oraz transglikozylaza guaninowa tRNA (TGT), wykorzystujące odpowiednio analogi GTP, S-adenozylometioniny (SAM) oraz niekanoniczej zasady azotowej (preQ). Ich zastosowanie wymaga specyficznych elementów strukturalnych, takich jak 5'-trifosforan (VCE), 5' kapu (grupy $GpppN$, m^7GpppN , $m^7GpppA_{mp}N$, enzymy GluTgs, Emc1 i CAPAM) lub nukleotydów purynowych znajdujących w sekwencji specyficznie rozpoznawanej przez transferazę (METTL i TGT).^[142, 149-154] Podobnie jak w wypadku transferaz, do modyfikacji zależnej od sekwencji RNA wykorzystano katalityczne kwasy nukleinowe i analogi GTP (deoksyrybozomy) lub ATP (rybozomy).^[155-157] Katalizują one reakcję między grupą 2' OH a trifosforanem (z eliminacją pirofosforanu), której selektywność zapewnia sekwencja katalitycznego kwasu nukleinowego, komplementarna do wybranego celu RNA.

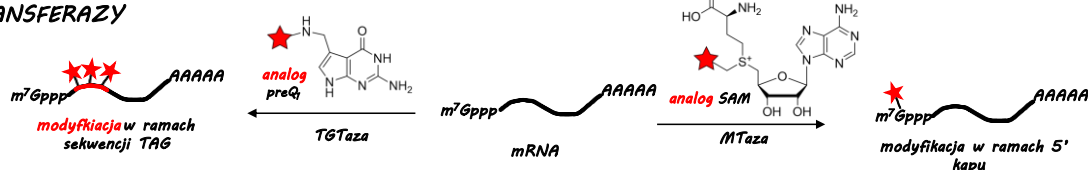
LIGAZY



POLIMERAZY



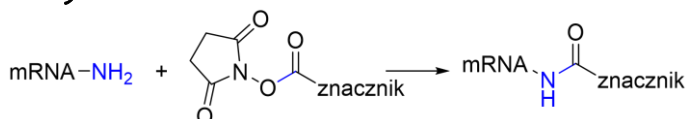
TRANSFERAZY



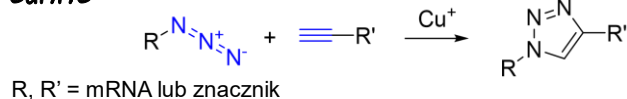
Rysunek 2.3.3: Wybrane reakcje enzymatyczne stosowane do znakowania i modyfikacji mRNA.

Wyżej wymienione metody enzymatyczne były wykorzystywane do wprowadzania nienaturalnych grup funkcyjnych, które wykorzystano w następnej reakcji chemicznej. (Rys. 3.2.4) To podejście, określane jako metody chemoenzymatyczne, pozwala na wprowadzenie grup o wysokiej i specyficznej reaktywności, przeprowadzenie wieloetapowej modyfikacji oraz na rozszerzenie wachlarza docelowych modyfikacji. W pierwszym etapie (enzymatycznym) wprowadzane były grupy takie jak aminy, azydki, alkeny i alkiny, które ulegały następczym reakcjom z estrami NHS, cykloaddycji (CuAAC, SPAAC, iEEDA), ligacji Staudingera lub reakcji Suzuki.^[118, 122, 123, 125-127, 130, 131, 136-138, 142, 151, 152, 154, 158-161] Szczególnie korzystne okazało się zastosowanie pochodnych azydkowych i reakcji SPAAC, co umożliwiło selektywną modyfikację i znakowanie mRNA *in vitro* i w komórkach.^[122, 130, 131, 137, 142, 160]

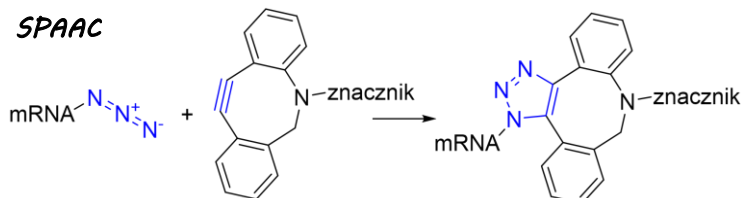
Estry NHS



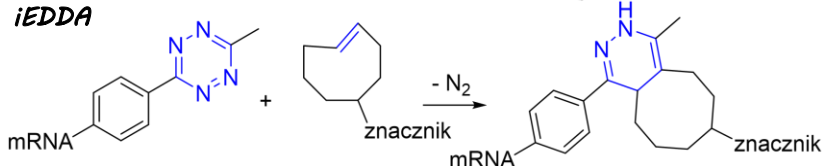
CuAAC



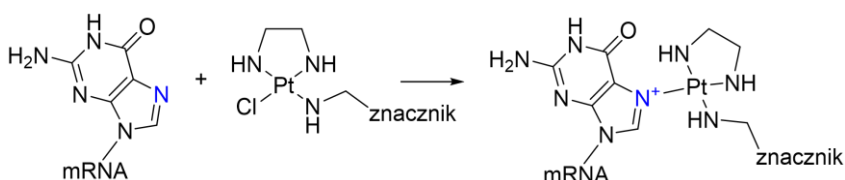
SPAAC



iEDDA



Kompleksy platyny

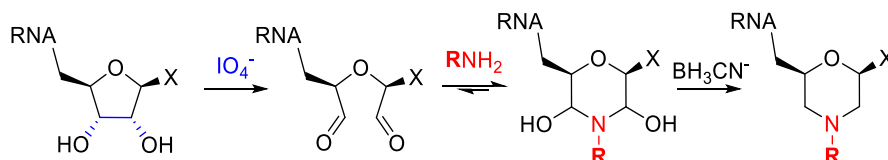


Rysunek 2.3.4: Wybrane reakcje chemiczne stosowane do znakowania i modyfikacji mRNA metodami chemoenzymatycznymi, oraz do bezpośredniej reakcji chemicznej z udziałem kompleksów platyny.

Niska wydajność, ograniczona specyficzność substratowa i wysoki koszt zwiększania skali są częstymi ograniczeniami metod opartych na reakcjach katalizowanych przez białka lub kwasy nukleinowe. Bezpośrednia chemiczna modyfikacja RNA jest wydajną alternatywą dla metod enzymatycznych. Niestety, tylko nieliczne metody chemiczne zostały zastosowane do bezpośredniej chemicznej modyfikacji RNA (nie uwzględniając syntezy modyfikowanych oligonukleotydów).^[120, 121, 162-165]

Jedną z bardziej unikalnych metod bezpośredniej chemicznej modyfikacji RNA jest oparta na reakcji oksydacyjnego rozszczepiania nadjodanem (inaczej reakcji rozszczepiania nadjodanem lub utleniania nadjodanem, Rys. 3.2.5). W trakcie reakcji, 2',3'-diol (wicynalny cis-diol) obecny w cząsteczkach RNA (w obrębie rybozy na końcu 3') jest przekształcany w acykliczną 1,5-dialdehydową pochodną, która może zostać następnie sfunkcjonalizowana na drodze reakcji z odpowiednim N-nukleofilem. W najwcześniejszych badaniach (w latach pięćdziesiątych XX w.), rozszczepianie nadjodanem zastosowano do analizy struktury końca 3' RNA naturalnego pochodzenia.^[166]

^{167]} Później zaobserwowano, że w obecności pierwszorzędowych amin i hydrazydów, 1,5-dialdehyd jest podatny na reakcję aminowania, prowadzącą do imin.^[168-170] Od tego czasu zjawisko to zostało wykorzystane do sprzęgania utlenionego końca 3' RNA z aminami lub hydrazydowymi pochodnymi barwników fluorescencyjnych, biotyny, kwasów nukleinowych, białek i żywic (przegląd prac z lat 1952–2018 w Tab. 2).



Rysunek 2.3.5: Schemat reakcji utleniającego rozszczepiania nadjodanem, następczej reakcji aminowania z N-nukleofilem i redukcji do stabilnej pochodnej morfolinowej. X: zasada azotowa.

Można wyróżnić dwa podejścia stosowane w reakcji utlenionego RNA z N-nukleofilami. Pierwsze podejście obejmuje reakcję dialdehydu z pochodną hydrazyny i następczą izolację RNA.^[168, 171-175] Powstawanie imin jest odwracalne, przez co wyizolowane produkty są niestabilne. Obejściem tego problemu jest zastosowanie aminowania redukcyjnego (tj. aminowanie w obecności czynników redukujących takich jak borowodorek sodu lub cyjanoborowodorek sodu), co prowadzi do powstania stabilnej pochodnej morfoliny.^[169, 170, 176-183] W przypadkach, w których zastosowano pierwszorzędowe aminy lub ich pochodne, wydajność redukcyjnego aminowania była znacząco wyższa niż aminowania.^[169, 177, 180, 183] Co istotne, w ramach wyżej wymienionych prac można zauważyć duże zróżnicowane warunków na etapach utleniania i aminowania: pH (5–10), czas reakcji (0.3 h–7 dni), temperatura (0–37 °C), stężenie nadjodanu (0.3 mM–1 M), długość (1–350 nt) i stężenie (5 M–100 mM) substratu RNA oraz rodzaj i stężenie N-nukleofila. Towarzyszą im zróżnicowane wydajności znakowania/modyfikacji (26–99%). Mimo, iż w niektórych pracach parametry wpływające na przebieg i wydajność reakcji były optymalizowane do pewnego stopnia, względna reaktywność różnych N-nukleofili, wpływ długości i stężenia RNA, oraz inne kluczowe czynniki nie były systematycznie badane.

Tabela 1: Przegląd prac, w których zastosowano enzymatyczne (ko- lub posttranskrypcyjne), chemiczne lub chemoenzymatyczne metody modyfikacji RNA.

Rodzaj metody	Miejsce	Etapy	Zastosowanie	Rok
chemoenz.	fosforan 5'	1) 5'-adenylacja przez lig. RNA I T4 z analogami ATP 2) Reakcja z estrem NHS	znakowanie <i>in vitro</i>	1997 ^[158]

chem.	fosforan 5'	Alkilacja odczynnikami diazowymi	modyfikacja <i>in vitro</i>	2016 ^[162]
chem.	fosforan 5'	1) kondensacja fosfoamidu 2) redukcja disiarczów 3) addycja Michaela	synteza sond molekularnych	1990 ^[163] , 2018 ^[120, 121]
chemoenz.	koniec 5'	1) IVT z analogami AMP (5'-R lub N6-R) 2) Reakcja z estrem NHS	znakowanie <i>in vitro</i>	2003 i 2008 ^[125]
chemoenz.	koniec 5'	1) IVT z 5'-nonborneno-GMP 2) reakcja iEEDA	znakowanie <i>in vitro</i>	2011 ^[126]
chemoenz.	koniec 5'	1) IVT z 5'-propargilo-GMP 2) CuAAC	synteza sond molekularnych	2013 ^[118]
chemoenz.	koniec 5'	1) IVT z dinukleotydem alkin-dUpG 2) reakcja CuAAC	znakowanie <i>in vitro</i>	2014 ^[127]
enz. kotrans.	Kap 5' (ryboza)	IVT z analogami R-m ⁷ GpppG (R: barwnik fluorescencyjny, biotyna)	znakowanie mRNA <i>in vitro</i>	2013 ^[184] , 2018 ^[129]
enz. posttrans.	kap 5' (ryboza)	Kapowanie z VCE i fluorescencyjnymi analogami GTP lub m ⁷ GTP	znakowanie mRNA <i>in vitro</i>	2016 ^[149]
enz. posttrans.	kap 5' (zasada, N2G)	1) alkilacja przez GluTgs z analogami SAM 2) reakcja CuAAC, SPAAC, iEEDA i inne	znakowanie <i>in vitro</i>	2013–2016 ^[151, 159]
enz. posttrans.	kap 5' (zasada, N7G)	1) alkilacja przez Ecm1 z analogami SAM 2) reakcja SPAAC	znakowanie mRNA <i>in vitro</i> i w komórkach	2016 ^[150]
enz. posttrans.	kap 5' (zasada, N6A)	1) alkilacja przez CAPAM z analogami SAM 2) reakcja SPAAC	znakowanie mRNA <i>in vitro</i> i w komórkach	2021 ^[142]
chemoenz.	Kap 5' (ryboza)	1) IVT z R-m ⁷ GpppG (R: NH ₂ , N ₃) 2) reakcja z estrem NHS lub SPAAC	znakowanie mRNA <i>in vitro</i> i w komórkach	2017 ^[130, 131]
chemoenz.	Kap 5' (mostek)	1) IVT z 5'-etynylo-GDP 2) reakcja CuAAC	modyfikacja mRNA <i>in vitro</i>	2017 ^[161]
enz. kotrans.	Kap 5' (ryboza)	IVT z FAM-Cap1 (trinukleotyd m ⁷ GpppNpG)	znakowanie fluorescencyjne mRNA <i>in vitro</i>	2021 ^[122, 123]
enz. kotrans.	Wewnętrzne nieselektywne (zasada)	IVT z fluorescencyjnymi analogami 5-R-UTP (R: barwnik fluorescencyjny)	znakowanie fluorescencyjne mRNA <i>in vitro</i>	2001 ^[132] , 2007 ^[133]
chem.	Wewnętrzne nieselektywne (zasada)	reakcja z udziałem reaktywnych kompleksów platyny (II)	znakowanie fluorescencyjne mRNA <i>in vitro</i>	2003 ^[164]
chem.	Wewnętrzne nieselektywne (zasada)	1) aminowanie redukcyjne w pozycji N2 guaniny z udziałem funkcjonalnych aldehydów 2) reakcja SPAAC	znakowanie <i>in vitro</i>	2020 ^[165]
enz. kotrans.	Wewnętrzne nieselektywne (zasada)	IVT z fluorescencyjnymi analogami tCTP (tricyklicznej cytydyny)	znakowanie fluorescencyjne <i>in vitro</i>	2010 ^[134]
enz. kotrans.	Wewnętrzne selektywne (zasada)	PLOR (modyfikacja metody IVT)	znakowanie <i>in vitro</i>	2015–2018 ^[135, 185]
chemoenz.	Wewnętrzne nieselektywne (zasada)	1) IVT z analogami 5-N ₃ -UTP lub 5-etynylo-UTP	znakowanie fluorescencyjne <i>in vitro</i> i w komórkach	2016 ^[136] , 2019 ^[137]

		2) reakcja CuAAC, SPAAC lub ligacja Staudingera		
chemoenz.	Wewnętrzne nieselektywne (zasada)	1) IVT z analogami 5-I-UTP 2) reakcja Suzuki	znakowanie <i>in vitro</i>	2018 ^[138]
enz. kotrans.	Wewnętrzne nieselektywne (mostek)	IVT z analogami ATP (α -S i α -BH ₃)	modyfikacja <i>in vitro</i>	2020 ^[87]
enz. kotrans.	Wewnętrzne nieselektywne (zasada)	IVT z fluorescencyjnymi analogami tCTP (tricyklicznej cytydyny)	znakowanie fluorescencyjne <i>in vitro</i>	2021 ^[139]
enz. posttrans.	wewnętrzne selektywne (zasada)	transfer zasady z TGT i fluorescencyjnymi analogami preQ	znakowanie <i>in vitro</i> i w komórkach	2015 i 2017 ^[153]
chem.	Wewnętrzne selektywne	1) reakcja cykloadycji z udziałem N6 i N1 adeniny i reaktywnego oligonukleotydu 2) reakcja CuAAC	synteza sond molekularnych	2016 i 2018 [175, 186]
chemoenz.	Wewnętrzne selektywne (zasada)	1) transfer zasady z TGT i analogami preQ 2) reakcja iEEDA lub fotoliza i aktywacja translacji	znakowanie i regulacja translacji mRNA <i>in vitro</i> i w komórkach	2018 ^[154]
enz. posttrans.	wewnętrzne selektywne (zasada, N6A)	1) alkilacja przez METTL3-14 lub METTL16 z analogami SAM 2) reakcja CuAAC	znakowanie <i>in vitro</i>	2021 ^[152]
posttrans.	wewnętrzne selektywne (ryboza)	deoksyrybozomy i analogi GTP	znakowanie i synteza sond molekularnych	2007 ^[155] , 2014 ^[156]
posttrans.	wewnętrzne selektywne (ryboza)	rybozomy i analogi GTP	znakowanie i synteza sond molekularnych	2019 i 2020 ^[157]
posttrans.	koniec 3'	elongacja z udziałem rybozomu o aktywności polimerazy i analogów NTP	znakowanie i modyfikacja <i>in vitro</i>	2018 ^[147]
enz. posttrans.	koniec 3'	ligacja przez lig. RNA I T4 z ³² P pNp	radioznakowanie RNA	1978 ^[117]
chemoenz.	Koniec 3'	1) ligacja przez lig. RNA I T4 z propargilo-pUp 2) CuAAC	synteza sond molekularnych RNA	2013 ^[118]
enz. posttrans.	koniec 3'	ligacja przez lig. RNA I T4 z fluorescencyjnymi analogami pNp	synteza sond molekularnych RNA	2017–2020 [119-121]
chemoenz.	Koniec 3'	1) ligacja przez lig. RNA I T4 z analogami p ^{R6} Ap 2) SPAAC	znakowanie mRNA <i>in vitro</i> i w komórkach; synteza sond molekularnych RNA	2021 ^[122, 123]
enz. posttrans.	koniec 3' / poliA	elongacja przez PAP z analogami ATP	radioznakowanie RNA	1993 ^[140]
chemoenz.	koniec 3' / poliA (zasada lub ryboza A)	1) elongacja przez PAP z analogami ATP 2) reakcja SPAAC, CuAAC	znakowanie mRNA <i>in vitro</i> i w komórkach	2012–2021 [137, 141, 142]
enz. posttrans.	koniec 3' / poliA (łańcuch)	elongacja przez PAP z analogami ATP (α -S i α -BH ₃)	modyfikacja i translacja mRNA	2020 ^[87]
enz. posttrans.	koniec 3'	elongacja na matrycy DNA przez polimerazę DNA I (Klenow) z α - ³² P dATP	radioznakowanie RNA	1996 ^[143] , 2016 ^[144] , 2017 ^[145]

enz. posttrans.	koniec 3'	elongacja przez polimerazę θ z analogami dUTP i ddUTP	znakowanie <i>in vitro</i>	2019 ^[146]
enz. posttrans.	koniec 3'	elongacja przez TUTazę z fluorescencyjnymi analogami UTP	znakowanie <i>in vitro</i>	2020 ^[148]
chem.	koniec 3'	1) utlenianie nadjodanem 2) aminowanie / aminowanie redukcyjne	różne	Tabela 2

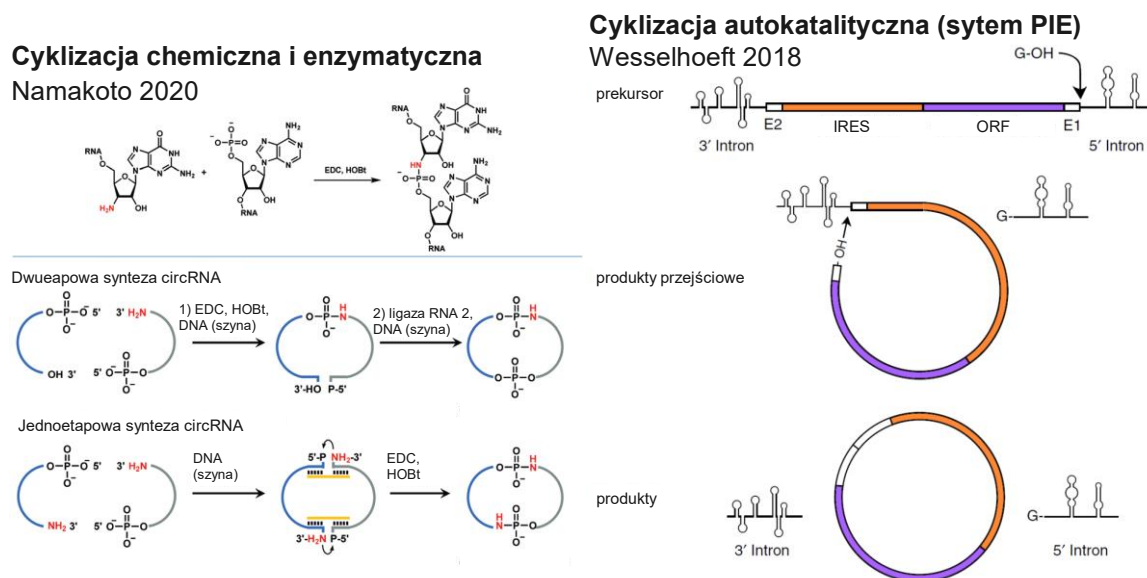
Tabela 2: Przegląd prac, w których zastosowano modyfikację rybonukleotydów lub RNA za pomocą utleniającego rozszczepiania nadjodanem i aminowania lub aminowania redukcyjnego. Wz: wydajność znakowania. *Wydajność oszacowana na podstawie zaprezentowanych w publikacji danych, niepodana bezpośrednio. **Wydajność znakowania lub podwójnego znakowania (utleniania, aminowania i następczej reakcji CuAAC).

Substrat	Warunki utleniania	Nukleofil	Warunki aminowania	Wz	Rok
Adenozyna, AMP, ATP, NpN	100 mM substrat, 100 mM NaIO ₄ , 24 h @25°C	brak	Brak	bd	1952 i 1953 ^[166]
	1 mM substrat, 1 mM NaIO ₄ , pH 5–6, 0.5 h @25°C	glicyna	33 mM dialdehyd, 200 mM nukleofil, pH 8, 16 h @25°C		1955 ^[167]
	1 mM substrat, 1 mM NaIO ₄ , pH 5–6, 0.5 h @25°C	metryloamina	1 mM dialdehyd, 10–100 mM nukleofil, 0–10 mM NaBH ₄ , pH 5–10, 1–20 h @25°C	bd	1961 ^[170]
	60 mM substrat, 100 mM NaIO ₄ , pH 5–6, 0.3 h @25°C	metryloamina	7 mM dialdehyd, 33 mM nukleofil, 90 mM NaBH ₄ , pH >8, 18 h @25°C	77%	1963 ^[169]
	20 mM substrate, 100 mM NaIO ₄ , pH 2–7, 0.5 h @4°C	hydrazyd kwasu izonikotynowego lub benzoesowego	10 mM dialdehyd, 10 mM nukleofil, pH 5–9, 2 h @25°C	70–99%	1974 ^[171]
	0.1 mmol NaIO ₄ , pH 7, 1 h @4°C	karboksylaza pirogromianowa (enzym)	0–6 mM dialdehyd, 10–20 U/ml enzym, 0–60 mM NaBH ₄ , pH 8.4, 0.6 h @25°C	bd	1976 ^[176]
	Jak w 1976 ^[176]	glicyna	50 mM dialdehyd, 140 mM nukleofil, 250 mM NaBH ₃ CN, pH 8, 2 h @37°C	bd	1985 ^[177]
		immobilizowany diaminoheksan	0.8 mM dialdehyd, 0.8 μ M nukleofil, 0–53 mM NaBH ₃ CN, pH 8, 2 h @37°C	bd	
tRNA i rRNA (70–350 nt)	5 mg/ml RNA, 100 mM NaIO ₄ , pH 5.0–6.0, 1 h @25°C	hydrazyd kwasu 2-hydroksy-3-naftowego	5 mg/ml dialdehyd, 12 mM nukleofil, pH 5, 2 h @37°C	bd	1960 ^[168]
	0.5–3 mg/ml RNA, 25 mM NaIO ₄ , pH 5.0–6.0, 1 h @25°C	hydrazyd agarozy (złoże)	0.1 μ M–0.1 mM dialdehyd, 12 mM nukleofil, pH 5, 0.2–2 h @25°C	bd	1972 ^[187]
	jak w 1960 ^[168]	fluorescencyjne hydrazydy,	~5 mg/ml dialdehyd, ~2 mM nukleofil, pH 5.6, 2 h @37°C	~60%*	1974 ^[172]

		semikarbazydy i karbohydrazony			
	~0.1 mM RNA, 1 M KIO ₄ , pH 5, 0.5 h @0°C	tiosemikarbazyd fluoresceiny	~0.5 mM dialdehyd, 10 mM nukleofil, pH 5.6, 16 h @0°C	~85%*	1994 ^[173]
	3 mM NaIO ₄ , pH 5.4, 1 h @25°C	etylenodiamina, diaminoheksan	1 mM nukleofil, pH 8.0, 5 mM NaBH ₃ CN, 1 h @37°C, dalej + 5 mM NaBH ₄ , 0.2 h @37°C	~80-99%*	1994 ^[178]
	30 µM RNA, 40 mM KIO ₄ , 1 h @25°C	hydrazyd biotyny	30 µM dialdehyd, 10 mM nukleofil, pH 5.6, 2 h @37°C, dalej + 40 mM NaBH ₄ , pH 8, 0.5 h @4°C	bd	1995 ^[179] , 2009 ^[181] , 2010 ^[182]
	50 µM RNA, 2.5 mM NaIO ₄ , pH 5, 0.9 h @0°C	tiosemikarbazyd fluoresceiny	50 µM dialdehyd, 1 mM nukleofil, pH 5.0, 16 h @0°C	bd	2000 ^[188]
oligoRNA (17-21 nt)	100 mM NaIO ₄ , pH 5, 1.5 h @25°C	hydrazyd lucifer yellow	100 mM nukleofil, pH 5.0, 4 h @25°C	~99%*	1995 ^[189]
	150 µM RNA, 0.3 mM NaIO ₄ , pH 6.0, 2 h @25°C	5'-amino RNA	150 µM dialdehyd, 0.2–0.6 mM nukleofil, 500 mM NaBH ₃ CN, pH 6.0, 168 h @25°C	~99%*	1996 ^[180]
	5 µM RNA, 30 mM NaIO ₄ , pH 6.0, 1 h @25°C	analog 5'-amino-cytydyny	25 µM dialdehyd, 1 mM nukleofil, 5 mM NaBH ₃ CN, pH 8.0, 2 h @37°C, dalej + 5 mM NaBH ₄ , 0.2 h @37°C	~50%*	2014 ^[190]
	10–200 µM RNA, 100 mM NaIO ₄ , pH 5, 0.5 h @37°C	propargiloamina	10–200 µM dialdehyd, 100 mM nukleofil, pH 7.7, 4 h @37°C, dalej + 100 mM NaBH ₃ CN, 0.5 h @37°C	~99%*, 7–88%**	2017 ^[183]
regulatorowe RNA (>300 nt)	27 µM RNA, 2.5 mM NaIO ₄ , pH 5.5, 1.5 h @25°C	hydrazyd Cy3	17 µM dialdehyd, 0.5 mM nukleofil, pH 7.0, 16 h @25°C	26–37%, 5%**	2018 ^[175]

Metody cyklizacji RNA

Szczególną formą modyfikacji jest cyklizacja, rozumiana jako kowalencyjne połączenie końców 5' i 3' RNA. Obecnie znane metody cyklizacji RNA oparte są na reakcjach chemicznych, enzymatycznych lub reakcjach autokatalitycznych, angażujących sekwencje zawarte w prekursorach circRNA (pre-circRNA, Rys. 3.2.6).



Rysunek 2.3.6: Wybrane przykłady cyklizacji kodującego RNA za pomocą metod chemicznych, enzymatycznych i autokatalitycznych.

Chemiczna cyklizacja wykorzystuje syntetyczne pre-circRNA, takie jak 3' fosorylowane RNA lub 3'-amino RNA, które ulegają kondensacji (z 5'-OH lub 5'-fosforanem odpowiednio) indukowanej bromocyjanem (BrCN) lub 1-etylo-3-(3-dimetyloaminopropylo)karbodiimidem (EDC).^[191] Chemiczna synteza kwasów nukleinowych w fazie stałej jest dogodną metodą otrzymywania modyfikowanych oligonukleotydów, jednak nie nadaje się do produkcji długich (> 200 nt) sekwencji RNA. W konsekwencji, circRNA otrzymywane metodami chemicznymi kodują stosunkowo krótkie sekwencje (oligopeptydowe epitopy, np. FLAG tag) i mają ograniczone zastosowanie w kontekście dostarczania genów.

Enzymatyczną cyklizację można przeprowadzić z pomocą ligaz, w szczególności T4 RNA ligazy I oraz II.^[58, 192] Oba enzymy wymagają 5' monofosorylowanych prekursorów. W pierwszej kolejności, prekursory są adenylowane na końcu 5', a następnie ligowane z grupą 3' OH rybozy.^[193] Ligacja może zachodzić międzycząsteczkowo lub wewnątrzcząsteczkowo, co prowadzi odpowiednio do oligomeryzacji lub cyklizacji

prekursora. Zaobserwowano, że aktywność ligaz jest silnie zależna od długości i sekwencji pre-circRNA, co jest niekorzystne z perspektywy ich zastosowania.^[56-58, 194, 195]

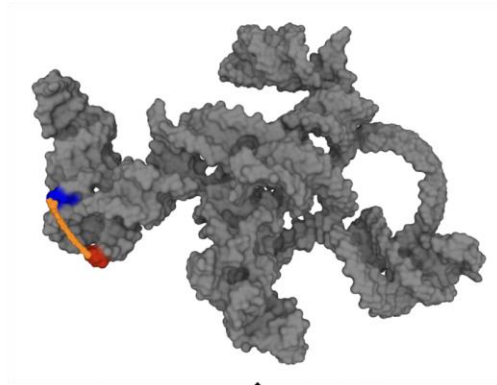
Jedne z najskuteczniejszych metod cyklizacji RNA wykorzystują aktywność katalitycznych kwasów nukleinowych.^[56-58, 195-198] Przykładem jest cyklizacja RNA *in vitro* przy użyciu systemu PIE (ang. *permuted intron-exon*) i zoptymalizowanych sekwencji intronowych zaangażowanych w autokatalityczny splicing.^[56, 57] Pre-circRNA wykorzystywane w systemie PIE składają się z sekwencji docelowego circRNA (np. zawierającej URT/IRES i ORF) oflankowanej sekwencjami przepołowionych intronów.^[196] Taki układ pozwala na zajście autokatalitycznego splicingu wstecznego (ang. *backsplicing*), który prowadzi do utworzenia circRNA i eliminacji sekwencji przepołowionych intronów. Kolejnym przykładem jest zastosowanie systemu Tornado (ang. *Twister-optimized RNA for durable overexpression*) do cyklizacji RNA *in vitro* oraz w komórkach.^[197] W tym wypadku cyklizacja zachodzi z udziałem flankujących sekwencji rybozomu *Twister*. W wyniku spontanicznej reakcji rybozymów, powstaje prekursor zawierający grupę 5'-OH (na końcu 5') oraz 2',3'-cykliczny fosforan na końcu 3', które są wykorzystywane w wewnątrzcząsteczkowej ligacji katalizowanej przez ligazę RtcB. Metody oparte na aktywności autokatalitycznych kwasów nukleinowego wymagają projektowania i optymalizacji na poziomie sekwencji, co ogranicza ich zastosowanie (nie mogą być zastosowane do cyklizacji pre-circRNA o dowolnej sekwencji). Ponadto aktywność autokatalityczna może zostać zaburzona poprzez wprowadzenie chemicznych modyfikacji (np. niekanonicznych zasad azotowych na etapie IVT).^[57]

Część II: Cel i wyniki badań

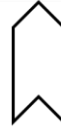
Cel badań

Podstawowym celem niniejszej pracy badawczej było opracowanie metod pozwalających na selektywną modyfikację chemicznej struktury końców 5' oraz 3' RNA. RNA jest łańcuchowym polimerem, w którym końce 5' oraz 3' są strukturalnie unikalne. Dzięki swej unikalności, miały zostać wykorzystane do regioselektywnej modyfikacji RNA. Modyfikacja końca 3' miała być osiągnięta za pomocą bezpośredniej metody chemicznej — połączeniem następczych reakcji oksydacyjnego rozszczepienia 2',3'-diolu rybozy i aminowania redukcyjnego. Modyfikacja końca 5' miała odbyć się za pomocą metody chemoenzymatycznej, kombinacji transkrypcji *in vitro* oraz chemii bioortogonalnej. Ta strategia miała pozwolić na regioselektywne i niezależne wprowadzenie znaczników (przykładowo barwniki fluorescencyjne czy hapteny) lub innych grup funkcyjnych, w obrębie struktury RNA. Swoistym rozszerzeniem koncepcji funkcjonalizacji końców 5' i 3' miała być chemoenzymatyczna cyklizacja RNA, w trakcie której sfunkcjonalizowana grupa 5' końca brałaby udział w wewnątrzcząsteczkowej reakcji aminowania redukcyjnego z utlenionym końcem 3'. Kolejnym celem badawczym było ustalenie wpływu modyfikacji (funkcjonalizacji lub cyklizacji) RNA na jego właściwości biochemiczne, biofizyczne i biologiczne. W szczególności, ustalony miał być wpływ modyfikacji na mobilność chromatograficzną i elektroforetyczną, właściwości fluorescencyjne, oraz podatność na aktywność nukleolityczną. Badania miały również objąć proces translacji i bioobrazowania modyfikowanego RNA w układach *in vitro* oraz *in vivo*. Ostatecznym celem badań było otrzymanie i zastosowanie modyfikowanego RNA jako sond molekularnych czy nośników terapeutycznej informacji genetycznej. Cele badań zostały zilustrowane na następnej stronie.

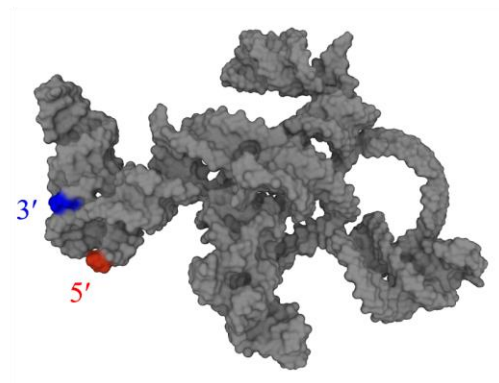
*Badania translacji
i stabilności*



Cyklizacja

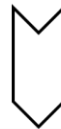


IVT

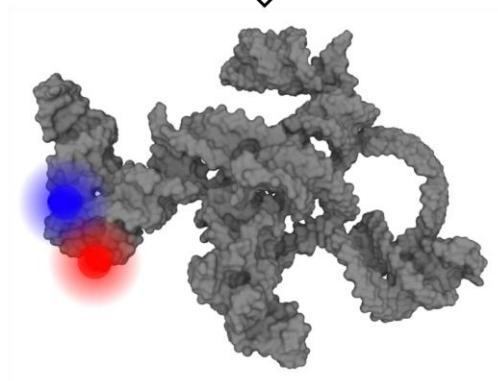


RNA

Znakowanie



*Sondy
fluorescencyjne i
bioobrazowanie*



1. Modyfikacja końca 3' RNA

1. 1. Założenia i strategia badań

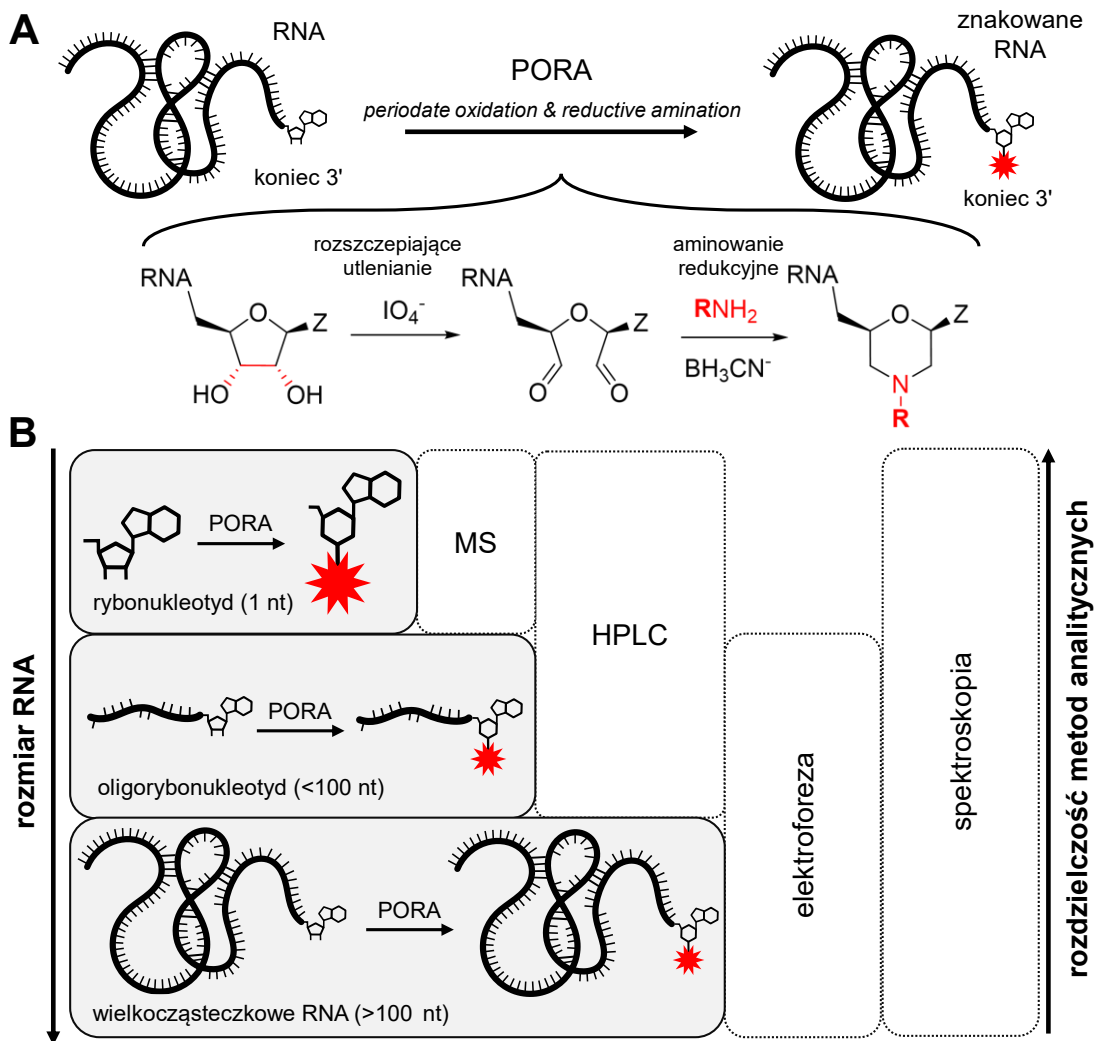
Pierwszym celem badawczym było opracowanie możliwie prostej, wydajnej oraz uniwersalnej metody pozwalającej na selektywną modyfikację struktury końca 3' wielkocząsteczkowego RNA (w szczególności mRNA). W momencie rozpoczęcia prac (rok 2018), opublikowane metody enzymatyczne i chemoenzymatyczne zostały wykluczone ze względu na ich ograniczoną wydajność (ligaza T4, polimeraza θ), brak kontroli nad umiejscowieniem (polimeraza T7) lub liczbą wprowadzanych modyfikacji (polimerazy poliA i T7). Dodatkowymi barierami były ograniczenia wynikające ze specyficzności substratowej oraz kosztów produkcji enzymów, czy z dostępnością analogów ich substratów lub kofaktorów. W świetle tego, skłoniono się ku zastosowaniu metody chemicznej – oksydacyjnego rozszczepienia 2',3'-diolu rybozy nadjodaniem i następczego aminowania redukcyjnego. W celu uproszczenia zapisu, metoda ta będzie określana mianem „PORA”, co stanowi akronim od sformułowania w języku angielskim: *periodate oxidation and reductive amination*.

Zastosowanie metody PORA miało potencjalnie ominąć wiele z wyżej wymienionych ograniczeń charakteryzujących metody enzymatyczne i chemoenzymatyczne. W szczególności, mechanizm rozszczepiającego utlenienia miał zapewnić wysoką regioselektywność i kontrolę nad liczbą wprowadzonych modyfikacji. Ponadto, w przeciwieństwie do reakcji bioortogonalnych, rozszczepiające utlenienie angażuje naturalny element struktury RNA (wycinalny 2',3'-cis diol rybozy), co pozwala przeprowadzić tę reakcję bezpośrednio na niemodyfikowanym kwasie rybonukleinowym. Stosując metody enzymatyczne, budowa i charakter podstawnika, który stanowi docelową modyfikację końca 3' RNA, jest ograniczony do struktury analogów nukleotydów. Przykładowo, analogów nukleotydów wyposażonych w barwniki fluorescencyjne, hapteny czy podstawniki bioortogonalne. Mechanizm aminowania redukcyjnego wykorzystuje cząsteczkę *N*-nukleofila, co znacząco rozszerza wachlarz potencjalnych podstawników i typów modyfikacji. W szczególności, ów *N*-nukleofil może stanowić element powierzchni nanocząstki, peptydu, białka, czy modyfikowanego kwasu nukleinowego. Kolejną zaletą metody PORA jest jej niski koszt, wynikający z zastosowania prostych odczynników chemicznych (t.j. pierwszorzędowe aminy, metanadjodany i cyjanoborowodorki).

Zarówno zalety i wady metody PORA, wynikają z jej oparcia na posttranskrypcyjnych reakcjach chemicznych. W obecności katalizatora (t.j. enzymu) reakcje zachodzą szybciej,

przez co można prowadzić je w łagodniejszych warunkach. Stosunkowo niska stabilność wielkocząsteczkowego RNA dyktuje zastosowanie możliwie łagodnych warunków. W szczególności, oznaczało to konieczność prowadzenia reakcji w roztworze wodnym o obojętnym odczynie (pH 6-7), w niskiej temperaturze (0-40 °C) i możliwie krótkim czasie. Zastosowanie PORA ma swoje korzenie w latach pięćdziesiątych dwudziestego wieku. Od tego czasu została wykorzystana do modyfikacji nukleozydów, nukleotydów, oligonukleotydów, tRNA i rRNA. Mimo tego, w pracach z lat 1952-2018 metoda PORA nie została wystandaryzowana czy wystarczająco zbadana pod kątem wpływu struktury *N*-nukleofila, czy rozmiaru i stężenia substratu RNA. Co więcej, nie znalazła zastosowania do RNA o długości przekraczającej 400 nukleotydów, w szczególności do modyfikacji mRNA (patrz Tab. 2 w rozdziale 2.3 Metody modyfikacji kodującego RNA).

Ze względu na ograniczenia dostępnych metod analitycznych, dążono do rozdzielenia problemów wynikających z przebiegu reakcji chemicznych (utleniania i redukcji) od problemów wynikających z właściwości wielkocząsteczkowego RNA. W tym celu zastosowano oddolną strategię badawczą (ang.: *bottom-up approach*). W praktyce oznaczało to rozpoczęcie badań od prób modyfikacji na małocząsteczkowych modelach końca 3' RNA i stopniowe zwiększanie ich rozmiarów (Rys. 1.1).

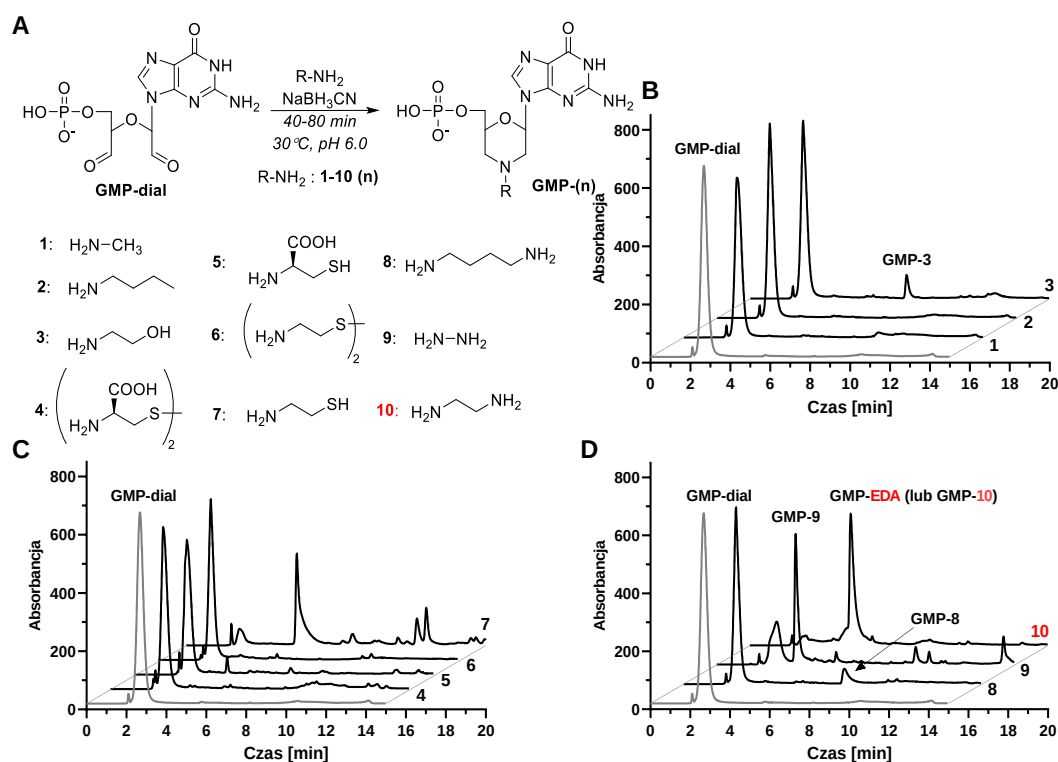


Rysunek 1.1: Strategia badań nad modyfikacją końca 3' RNA. A) Schemat metody modyfikacji na drodze rozszczepiającego utleniania i następczego aminowania redukcyjnego (PORA). B) Dobór metod analitycznych (oraz ich rozdzielczość) w zależności od rozmiaru modelowego RNA.

1. 2. Przebieg aminowania redukcyjnego w zależności od doboru *N*-nukleofila

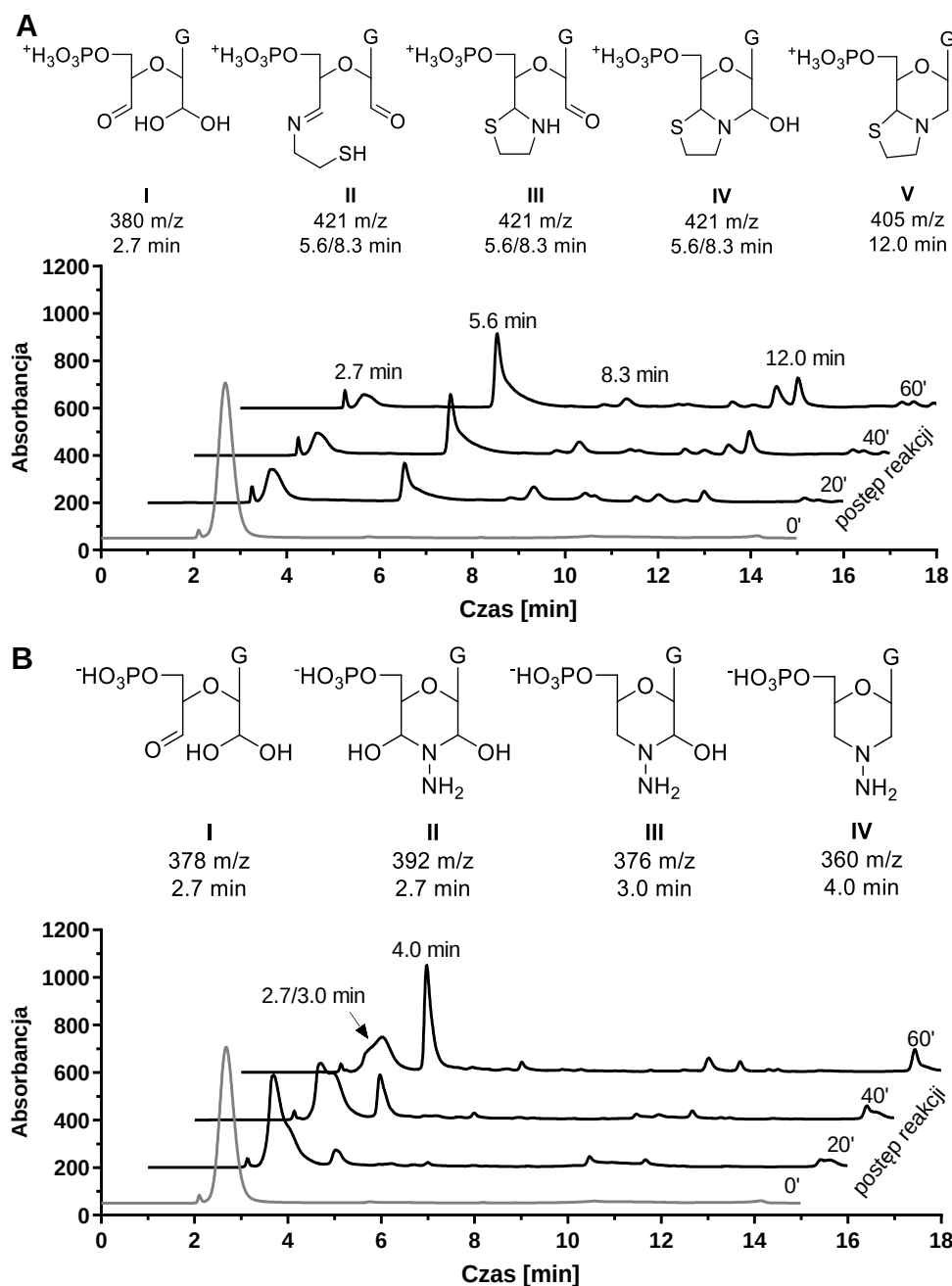
Jako pierwszy, małowczątkowy model końca 3' RNA, obrano 5'-monofosforan guanozyny (GMP). Wybór ten był podyktowany jego dostępnością (dostępny handlowo, tani), dobrą rozpuszczalnością w wodzie (lepszą niż rybonukleozydów) czy optymalnymi właściwościami chromatograficznymi (czasy retencji krótsze niż dla AMP i znacznie większa retencja niż dla UMP czy CMP). Co najistotniejsze, rozmiar cząsteczki GMP miał pozwolić na z wysokorozdzielczą analizę produktów za pomocą metod HPLC i MS, potencjalnie dając wgląd w przebieg transformacji rybozy.

W pierwszej kolejności podjęto próby otrzymania dialdehydowej pochodnej GMP (GMP-dial, Rys. 1.2.1) w wyniku rozszczepiającego utlenienia nadjodanem potasu (KIO_4). Utlenianie zachodziło szybko i ilościowo, pod warunkiem, że reakcja była prowadzona bez dostępu do światła. Wyizolowany na drodze strącania GMP-dial pozostawał stabilny w trakcie przechowywania w formie zamrożonego roztworu. Rozmrożony roztwór GMP-dial był wykorzystywany w próbach aminowania redukcyjnego. Aminowanie redukcyjne było prowadzone w łagodnych warunkach (30°C, pH 6.0, czas reakcji poniżej dwóch godzin), przy zastosowaniu niskiego stężenia GMP-dial (1 mM), różnych *N*-nukleofili (monoamin alifatycznych, diamin, tioamin, aminokwasów, etanoloaminy i hydrazyny) w dwukrotnym nadmiarze (1-2 mM, dwie grupy NH_2 na cząsteczkę GMP-dial) oraz cyjanoborowodoru sodu jako reduktora (NaBH_3CN , 10 mM). Próbkę mieszanin reakcyjnych były pobierane w punktach czasowych (co 20 min) i analizowane za pomocą HPLC i MS (Rys. 1.2.1–1.2.3).



Rysunek 1.2.1: Aminowanie redukcyjne GMP-dial w zależności od N-nukleofila. A) Schemat reakcji. B-D) Profile HPLC produktów reakcji, które zostały następnie zanalizowane za pomocą MS.

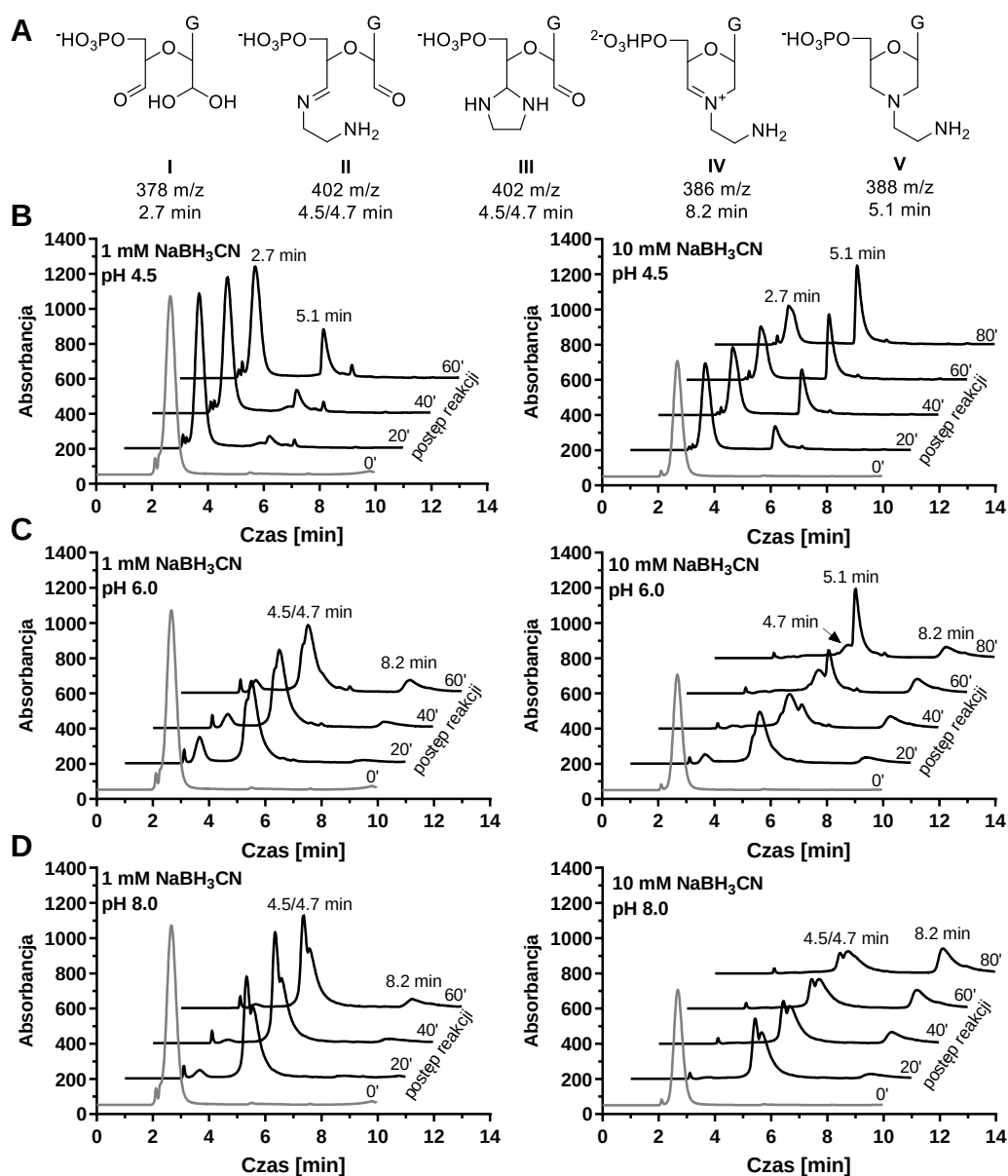
Największe różnice w przebiegach reakcji objawiały się pod wpływem zmiany nukleofila (związki 1-10 na Rys. 1.2.1). W wypadku większości z nich powstawały niewielkie lub śladowe ilości produktów aminowania redukcyjnego (Rys. 1.2.1B-C). Trzy nukleofile: hydrazyna (9), cysteamina (7) oraz etylenodiamina (10), wyróżniały się pod kątem reaktywności. Reakcja między GMP-dial a hydrazyną prowadziła do częściowej (~50% po 60 min reakcji) konwersji substratu do morfolinowej pochodnej GMP (Rys. 1.2.2B, związek IV). W reakcji z cysteaminą powstawał produkt inny niż analogiczna pochodna morfoliny. Na podstawie wyników analizy MS oraz opublikowanych badań dotyczących reakcji między aldehydami a cysteaminą lub aminokwasami^[199] uznano, iż jest to produkt zawierający pierścień tiazolidynowy (Rys. 1.2.2A, związek III). Wyłącznie w trakcie reakcji z etylenodiaminą (EDA), GMP-dial ulegał niemal całkowitej konwersji do pożądanego produktu GMP-EDA (Rys. 1.2.1D, Rys. 1.2.3 związek V).



Rysunek 1.2.2: Przebieg i analiza produktów aminowania redukcyjnego GMP-dial w obecności cysteaminy (A) lub hydrazyny (B).

Te obserwacje stanowiły podstawę do dalszych badań nad mechanizmem reakcji aminowania redukcyjnego między GMP-dial i EDA (Rys. 1.2.3). W tym celu, badano jej przebieg pod wpływem zmiennego pH (4.5, 6.0 lub 8.0) i stężenia cyjanoborowodoru (1 lub 10 mM). W pH 4.5 reakcja prowadziła selektywnie do GMP-EDA oraz ulegała przyspieszeniu przy wyższym stężeniu reduktora. Zwiększenie pH z 4.5 do 6 dodatkowo zwiększało szybkość reakcji i konwersję do GMP-EDA. W pH 8 reakcja między EDA a

GMP-dial zachodziła równie szybko co w pH 6, jednak prowadziła głównie do częściowo zredukowanej, iminowej pochodnej (Rys. 1.2.3A, związek IV). Na tej podstawie określono optymalne warunki aminowania redukcyjnego (pH 6, 10 mM NaBH₃CN) będące punktem odniesienia w późniejszych eksperymentach.



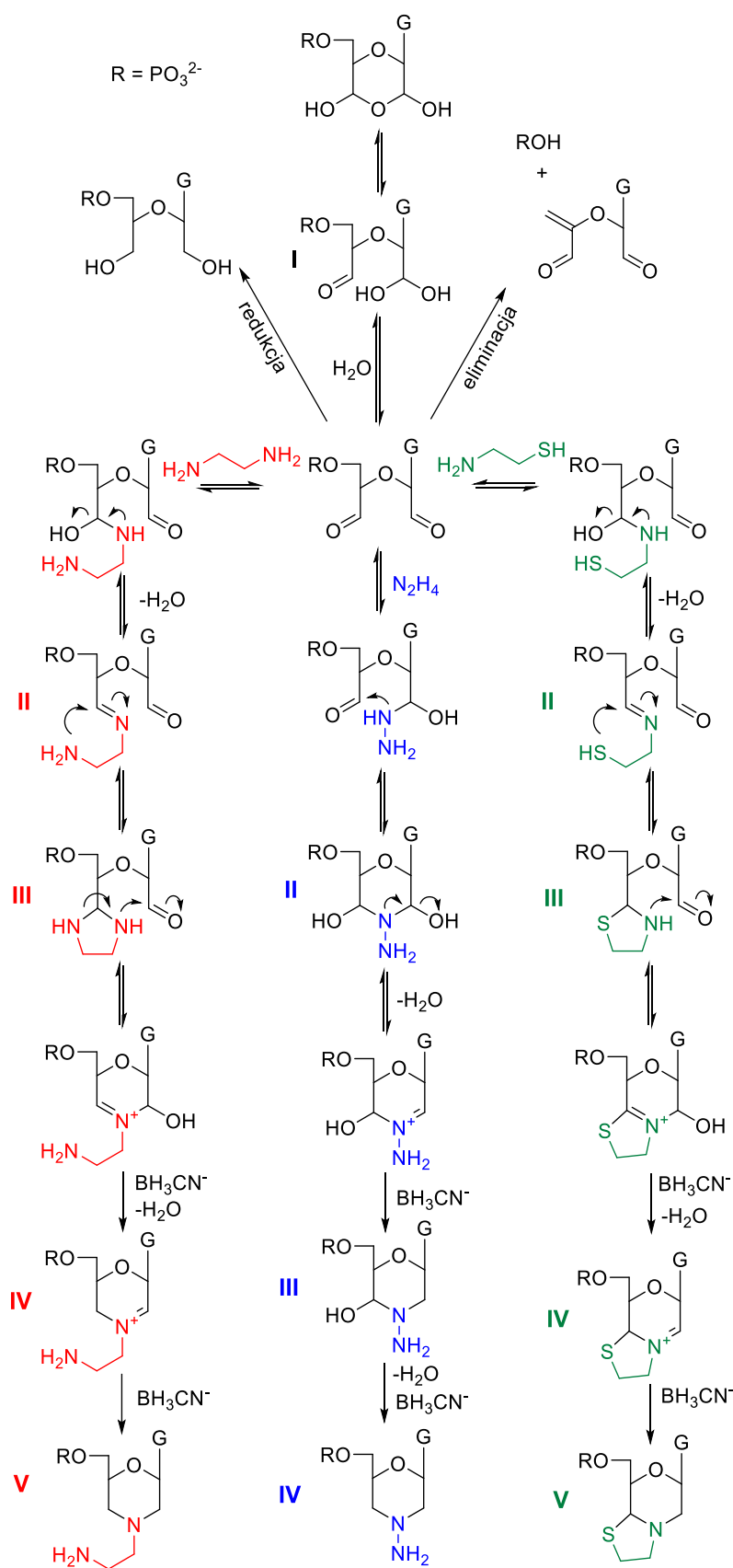
Rysunek 1.2.3: Przebieg i analiza produktów reakcji aminowania redukcyjnego między GMP-dial a etylenodiaminą. A) Struktury będące interpretacją widm masowych pików o oznaczonych czasach retencji. B-D) Profile HPLC produktów reakcji prowadzonych w zmiennym warunkach pH (4.5, 6.0 lub 8.0) oraz stężenia NaBH₃CN (1 lub 10 mM).

W trakcie reakcji biegnących przy niższym stężeniu reduktora, zaobserwowano powstawanie dwóch izomerycznych produktów pośrednich. Powstawały one w proporcjach zależnych od pH, a ich masy monoizotopowe odpowiadały pochodnej

iminowej i imidazolidynowej (Rys. 1.2.3A, związki II i III) — cząsteczkom o strukturze analogicznej do produktów reakcji z cysteaminą (Rys. 1.2.2A, związki II i III). Ponadto, w trakcie reakcji z EDA nie zauważono charakterystycznego produktu pośredniego, analogu 2,4-morfolinodiolu (t.j. związku II z Rys. 1.2.2B).

Na tej podstawie zaproponowano mechanizm reakcji wykorzystujący powstawanie alternatywnego produktu pośredniego: analogu zawierającego ugrupowanie imidazolidynowe (Rys. 1.2.3 związek III). Obecność motywu N-C-C-Nukleofil (jak w wypadku EDA, cysteaminy czy etanoloaminy) miałyby umożliwiać powstawanie pięcioczłonowego ugrupowania zawierającego atom węgla 2' lub 3' nukleotydu (Rys. 1.2.4). W związku z tym, spodziewano się wpływu grupy nukleofilowej (OH, NH₂, SH) na jego formowanie oraz na szybkość reakcji. Dodatkowo, wpływ grupy nukleofilowej byłby zależny od pH, ze względu różnice pK_a sąsiadującej grup aminowej (7.6)^[200], tiolowej (~8.2)^[201] czy hydroksylowej (~16)^[202]. W przypadku EDA, obecność dwóch nukleofilowych atomów azotu miałyby pozwalać na zajście czteroetapowej transformacji, składającej się z dwukrotnej redukcji i dehydratacji, prowadzącej do GMP-EDA. Zaproponowany mechanizm jest zgodny z wynikami badań nad reakcjami między omawianymi N-nukleofilami z aldehydami oraz między dialdehydowymi pochodnymi rybonukleozydów z tris(hydroksymetylo)aminometanem^[199].

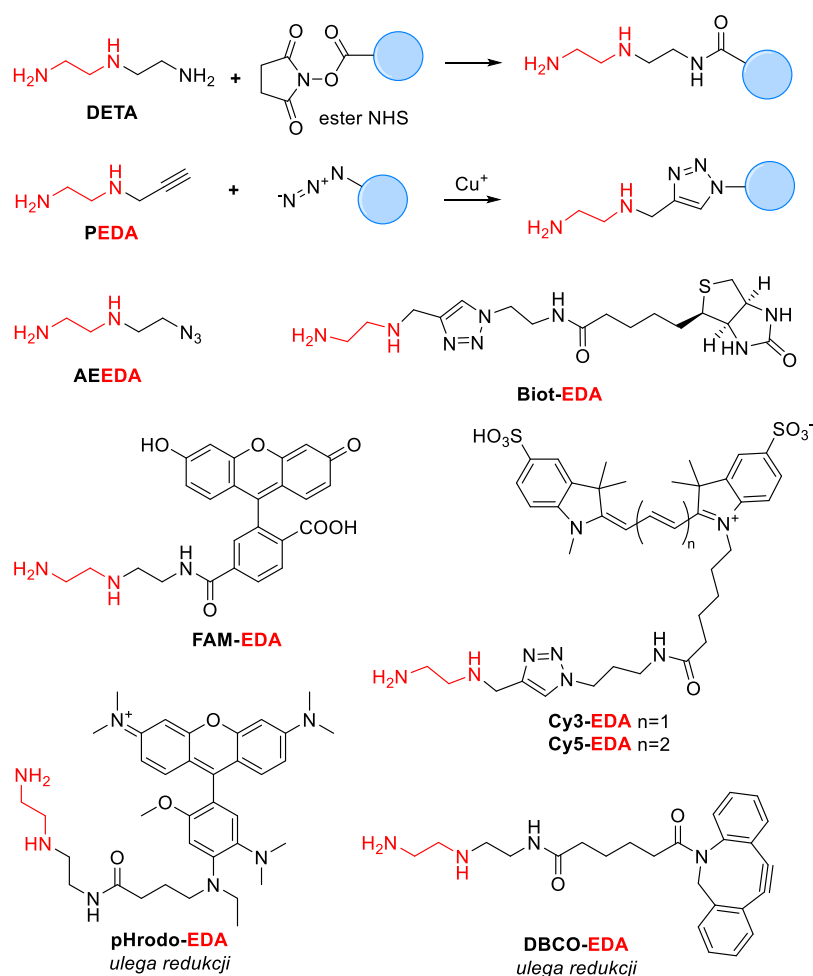
Zaproponowany mechanizm nie wykluczał możliwości, w której *N*-alkilowe pochodne EDA wykazywałyby podobną reaktywność co EDA. Ten wniosek posłużył do zaprojektowania nowego wariantu metody PORA, która miała być podstawą do osiągnięcia podstawowego celu badawczego — selektywnej modyfikacji struktury chemicznej końca 3' RNA.



Rysunek 1.2.4: Schemat reakcji, którym ulega GMP-dial w warunkach aminowania redukcyjnego. Mechanizmy reakcji w obecności trzech N-nukleofili zaproponowano na podstawie zjawisk opisanych w literaturze oraz zaobserwowanych w trakcie badań.

1. 3. Modyfikacja RNA za pomocą analogów EDA

By zweryfikować tę hipotezę, przeprowadzono syntezę szeregu pochodnych EDA. W tym celu posłużono się metodami chemii klik, w szczególności reakcją CuAAC oraz estrami NHS (Rys. 1.3.1). To pozwoliło na przeprowadzanie jednoetapowych syntez analogów EDA, wychodząc z prostych lub dostępnych handlowo prekursorów. Podstawowymi prekursorami były dietylenotriamina (DETA) i *N*-(propargilo)etylenodiamina (PEDA), oraz pochodne barwników fluorescencyjnych (Cy3, Cy5, FAM i pHrodo), biotyny czy cyklooktynu (DBCO). Przeprowadzono również syntezę *N*-(2-azydoetylo)etylenodiaminy (AEEDA), w celu poszerzenia wachlarza potencjalnych modyfikacji końca 3' RNA.

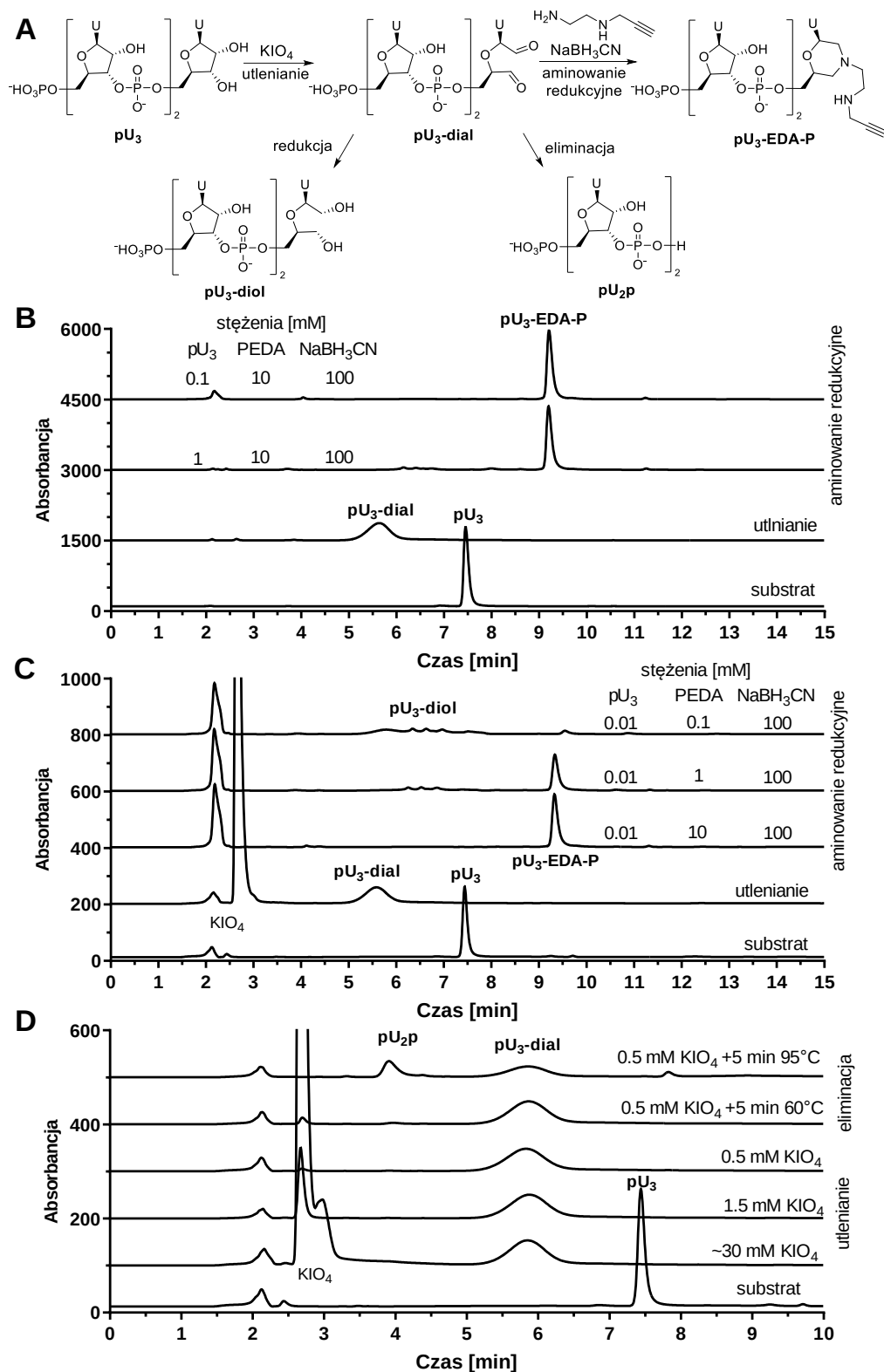


Rysunek 1.3.1: Struktury analogów EDA oraz schemat ich syntezy (reakcje CuAAC oraz estrów NHS). Analogi pHrodo-EDA oraz DBCO-EDA okazały się nieodpowiednie do przeprowadzenia aminowania redukcyjnego ze względu na ich podatność na czynniki redukujące.

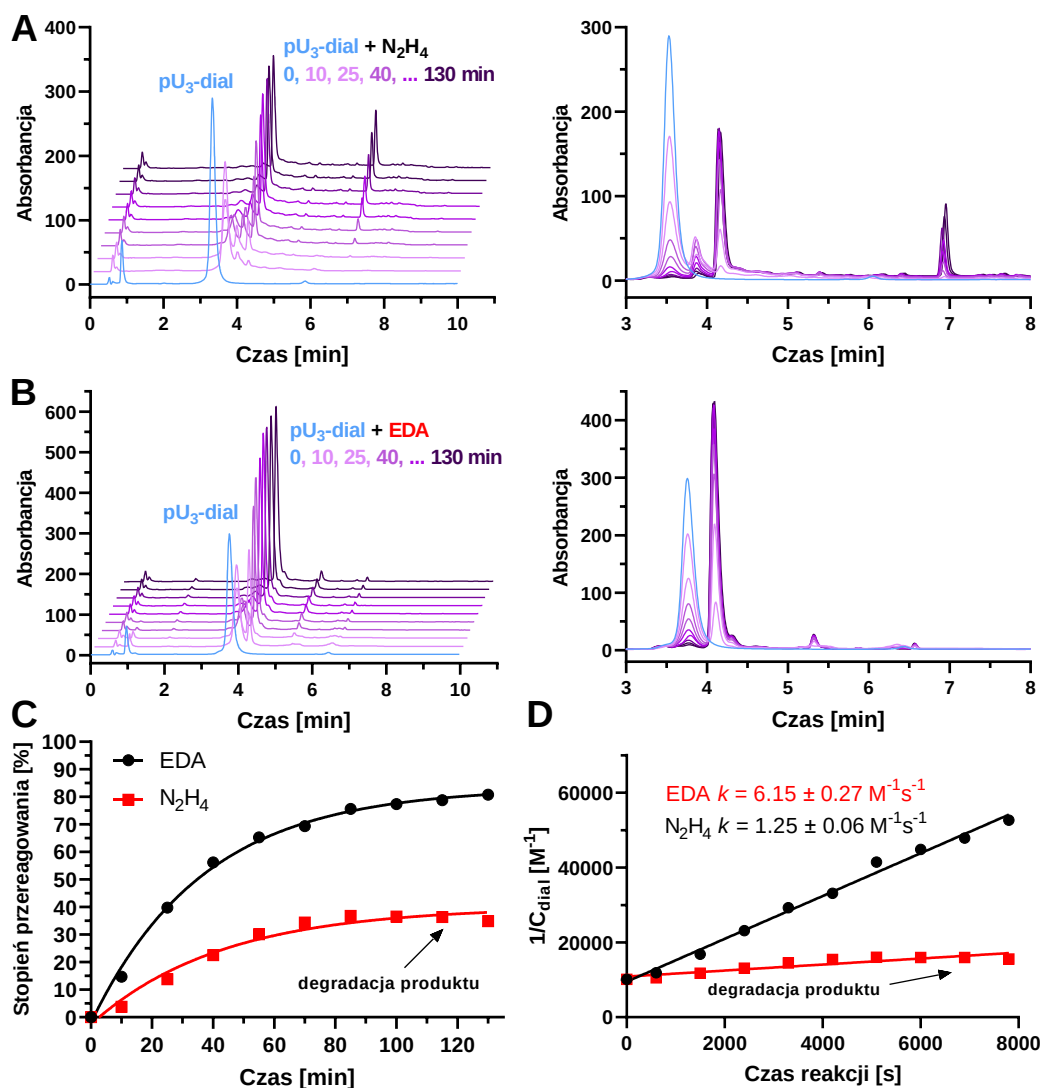
W pierwszej kolejności przeprowadzono próby PORA z wykorzystaniem 5'-monofosforanu triurydyny (pU₃) oraz PEDA (Rys. 1.3.2). Zaobserwowano, że reakcja

utleniania zachodziła szybko, jednak jej produkt (pU₃-dial) był podatny na rozkład w trakcie izolacji. Zoptymalizowano utlenianie pU₃ pod kątem czasu i stężenia KIO₄. Udało się znacząco zmniejszyć stężenie utleniacza (z 30 mM do 1 mM) przy zachowaniu krótkiego czasu reakcji (20 min) i obniżeniu temperatury (z 30°C do 25°C). W celu optymalizacji aminowania redukcyjnego badano przebiegi reakcji przy różnych stężeniach pU₃-dial (10, 100 lub 1000 μM), PEDA (0.1, 1.0 lub 10 mM) oraz NaBH₃CN (5, 50, 500 mM) (Rys. 1.3.2). W obecności wyższych stężeń PEDA (1 i 10 mM) oczekiwany produkt pU₃-EDA-P powstawał w sposób wydajny i selektywny (68-99%). Przy jej najniższym stężeniu (0.1 mM) zaobserwowano zachodzenie reakcji eliminacji (do pU₂p) oraz redukcji (do pU₃-diol) i znacznie obniżenie wydajności aminowania redukcyjnego (5-16%). Na podstawie wyników optymalizacji zaproponowano protokół PORA, będący dwuetapowym procesem odbywającym się w jednym naczyniu (Rys. 1.3.3). W pierwszej kolejności roztwór pU₃ (0.1-1 mM) był inkubowany przez 30 min w ciemni, w temperaturze 25°C i w roztworze nadjodanu (1.0-1.5 mM). Następnie do mieszaniny dodawano roztwory buforu (końcowo 100 mM KH₂PO₃ pH 6), NaBH₃CN (końcowo 20 mM) oraz analogu R-EDA (końcowo 1 mM) i kontynuowano inkubację przez dwie godziny. W ten sposób otrzymano analogi trinukleotydu (pU₃-EDA-R) zawierające grupy 3'-aminowe, -azydkowe, -alkinowe, sulfo-cyjaninę 3 (Cy3), sulfo-cyjaninę 5 (Cy5), biotynę czy karboksyfluoresceinę (FAM) z wydajnościami rzędu 75-99%.

W celu porównania reaktywności analogów R-EDA z innymi *N*-nukleofilami przeprowadzono pomiary kinetyki aminowania redukcyjnego. Za pomocą HPLC mierzono względne zmiany stężenia substratu (pU₃-dial) i jego konwersję do pożądanego produktu w czasie. Wyniki dopasowano do modelu reakcji drugiego rzędu i wyznaczono stałe szybkości reakcji (Rys. 1.3.4). Stałe szybkości reakcji dla większości zastosowanych amin (0.1-1.7 M⁻¹s⁻¹) czy hydrazyny (1.3 M⁻¹s⁻¹) były znacznie niższe niż dla EDA (6.2 M⁻¹s⁻¹) i jej pochodnych (4.5-6.0 M⁻¹s⁻¹). Wolniejsze reakcje były też mniej wydajne (5-66%). Wyjątkiem pośród analogów R-EDA była *N*-acetyloetylenodiamina (AcEDA), która okazała się jednym z najmniej reaktywnych nukleofilów (0.1 M⁻¹s⁻¹, 7%). Wyniki tych badań są zgodne z hipotezą dotyczącą mechanizmu aminowania redukcyjnego z EDA i jej analogami.



Rysunek 1.3.2: Optymalizacja metody PORA z wykorzystaniem pU₃ i PEDA. A) Schemat reakcji zaobserwowanych w trakcie procesu. B-C) Profile HPLC produktów reakcji aminowania redukcyjnego w zależności od stężenia pU₃ lub PEDA. D) Profile HPLC produktów reakcji utleniania w zależności od stężenia KIO₄ oraz reakcji eliminacji w zależności od temperatury.



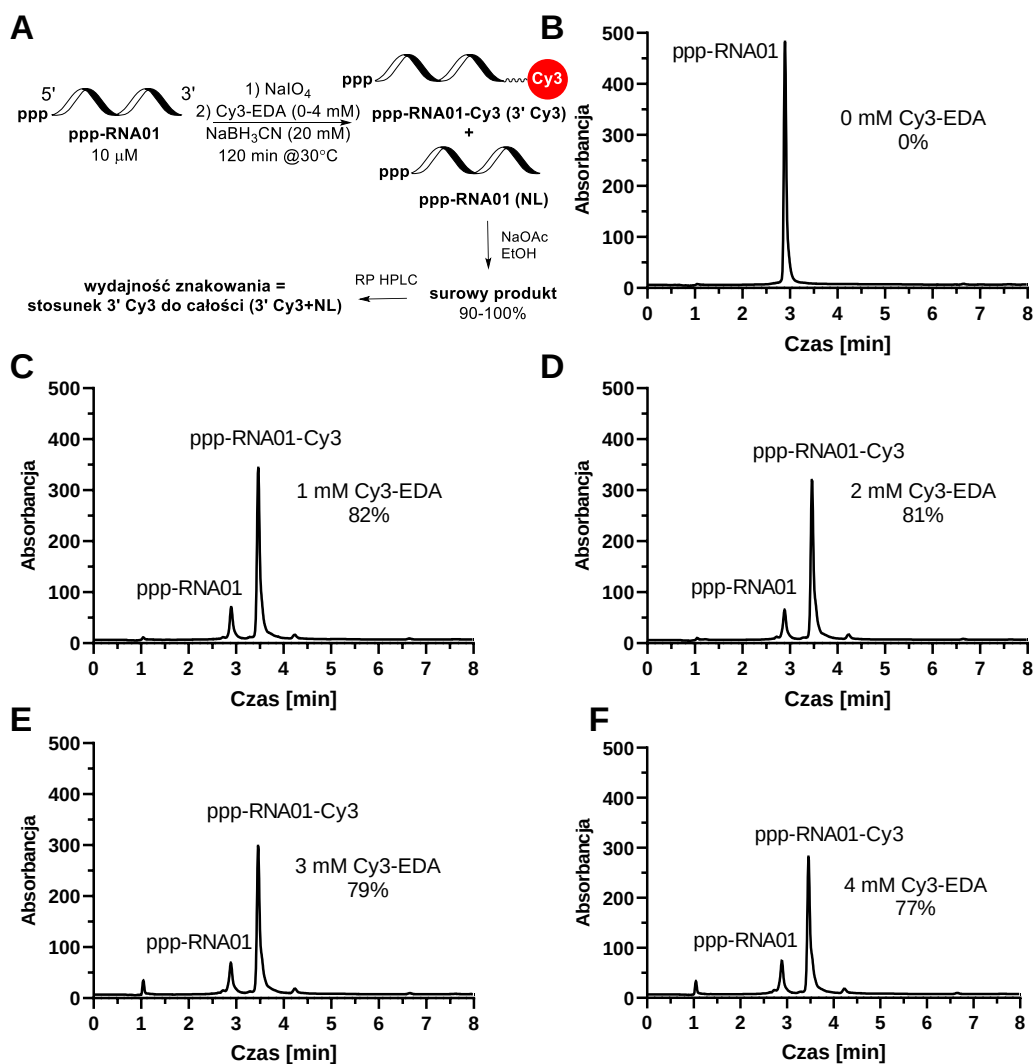
Rysunek 1.3.4: Badania kinetyki aminowania redukcyjnego za pomocą HPLC. Profile HPLC produktów reakcji z hydrazyną (A) oraz EDA (B). C) Stopień przereagowania substratu (pU₃-dial) w czasie w zależności od zastosowanego nukleofila. D) Zależność odwrotności stężenia substratu od czasu reakcji. Liniowa zależność odwzorowuje przebieg reakcji drugiego rzędu, a nachylenie dopasowanej prostej stanowi stałą szybkości reakcji.

W celu zastosowania metody PORA do modyfikacji RNA o większej masie cząsteczkowej, wykorzystano reakcję transkrypcji *in vitro* (IVT). W pierwszej kolejności przeprowadzono syntezę 5'-trifosforanu poliribonukleotydu o długości 35 nukleotydów (ppp-RNA01) na drodze IVT. Następnie, RNA poddano kolejno znakowaniu 3' Cy3, izolacji i analizie HPLC (Rys. 1.3.5). Identyfikację produktów i substratów przeprowadzono dzięki pomiarom absorbancji i fluorescencji eluatu w układzie HPLC, oraz jego analizie elektroforetycznej. Ponadto, w wyniku znakowania czas retencji RNA znacząco się wydłużał, co pozwoliło na wyznaczenie wydajności reakcji. Znakowanie ppp-

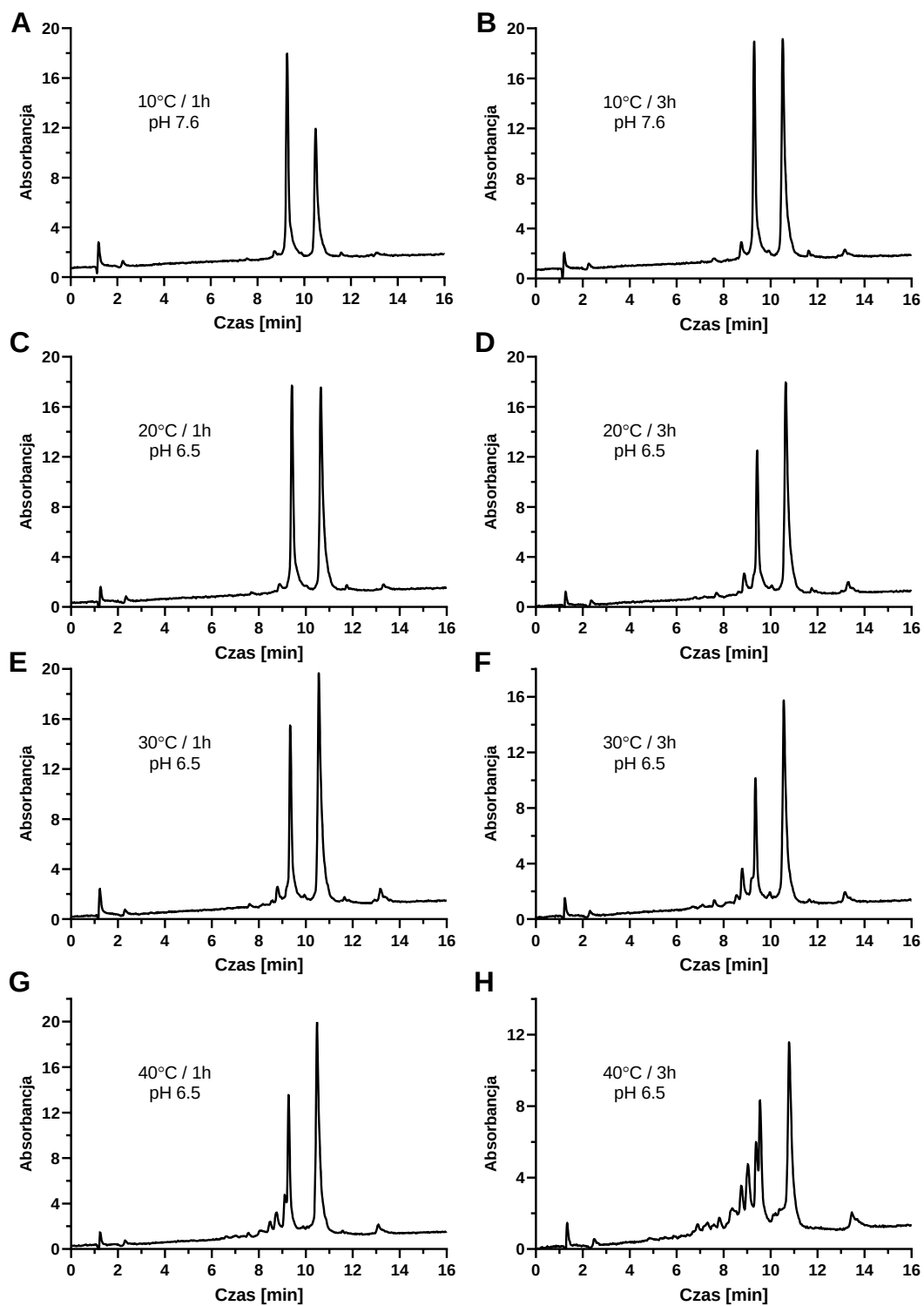
RNA01 (10 μM) zachodziło wydajnie (83%) przy zastosowaniu 1 mM Cy3-EDA, a zwiększenie stężenia barwnika (2-4 mM) miało marginalny wpływ na wydajność reakcji (79-81%). Następnie, zwieszono rozmiar modelowego RNA (z 35 do 276 nukleotydów, ppp-RNA02), który został wykorzystany do dalszej optymalizacji PORA. W tym celu, zbadano wydajność znakowania w zależności od pH (5.5, 6.0, 6.5, 7.0 lub 7.6), temperatury (10, 20, 30 lub 40 $^{\circ}\text{C}$) oraz czasu (1 lub 3 godziny) aminowania redukcyjnego (Rys. 1.3.6). W celu określenia stopnia niespecyficznego degradacji ppp-RNA02 w danych warunkach, przeprowadzono równoległe „ślepe” próby znakowania (mock), w których pominięto etap utleniania nadjodanem (Rys. 1.3.7). Wyniki wszystkich 80 prób wskazują, iż stopień degradacji i wydajność znakowania nie zależą od pH, a ich optimum odpowiada inkubacji w 20 $^{\circ}\text{C}$ przez trzy godziny lub w 30 $^{\circ}\text{C}$ przez godzinę (Rys. 1.3.8). Na tej podstawie określono, iż optymalne warunki aminowania redukcyjnego to pH 6.0, temperatura 25 $^{\circ}\text{C}$ oraz czas reakcji w zakresie 120-140 minut. Ostatecznie, wykorzystano je do modyfikacji wielkocząsteczkowego RNA (t.j. mRNA), o rozmiarach w zakresie 900-2000 nukleotydów (wybrane spośród RNA05–RNA17, Tab. 3 oraz Tab. 5 w rozdziale 4. 3. Badania wykonane ze współpracownikami: optymalizacja cyklizacji RNA). Zaskakująco, w wyniku znakowania czasy retencji wielkocząsteczkowego RNA również ulegały znacznemu wydłużeniu (Rys. 1.3.9). Jak wcześniej, to zjawisko zostało wykorzystano do badania przebiegu i wydajności znakowania. Jego wyjaśnienie zostało zaproponowane na późniejszym etapie badań oraz w dalszej części niniejszej rozprawy (rozdział 2.3. Efekt „kotwiczenia”). Konsekwencją wykorzystania wielkocząsteczkowego substratu było obniżenie jego stężenia w mieszaninie reakcyjnej (0.2-2 μM). Mimo tego, wydajności znakowania pozostawały wysokie (80-85%) oraz nie wykryto degradacji RNA spowodowanej procedurą znakowania.

Jedną z przewag metody PORA nad metodami enzymatycznymi jest to, że nie wykorzystuje ona kosztownych odczynników, w tym podatnych na utratę aktywności enzymów. By to zademonstrować, przeprowadzono serię eksperymentów mających na celu zwiększenie skali znakowania RNA. W pierwszej kolejności przeprowadzono znakowanie 3' Cy3 na modelu ppp-RNA06 (mRNA kodujące eGFP, 1065 nt) przy rosnącym stężeniu substratu (0.4-6.3 μM , 0.2-3.6 g/l, Rys. 1.3.10). Wydajność znakowania była obniżona (52-75%) po przekroczeniu stężenia ~ 3 μM (1.8 g/l). Przeprowadzono również reakcje na mniejszym modelu mRNA (ppp-RNA04, 237 nt). Ich wydajność nie zmieniała się w badanym zakresie stężeń (1-16 μM , 0.1-1.9 g/l). Stężone roztwory kwasów nukleinowych charakteryzują się zwiększoną lepkością.^[203] W przypadku badanego RNA, zjawisko to

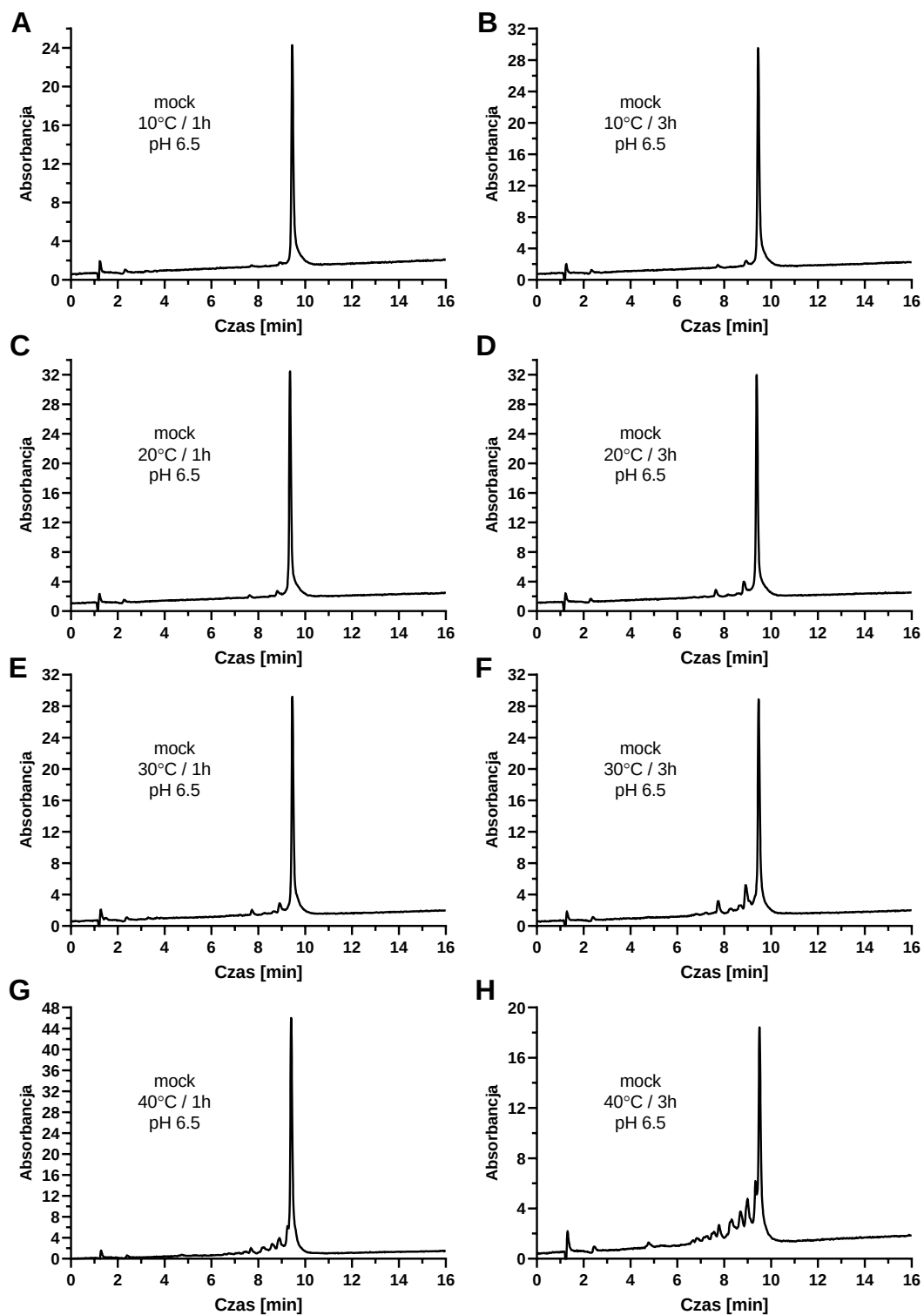
można było zaobserwować w trakcie pipetowania roztworów o stężeniu przekraczającym jeden gram na litr. Przypuszczalnie, wzrost lepkości roztworu jak i spadek wydajności znakowania mają tę samą przyczynę. Ostatecznie, przeprowadzono reakcję znakowania ppp-RNA06 w skali 500 μg (1.14 g/l). Po izolacji na drodze strącania otrzymano 376 μg (75%) materiału RNA zawierającego 77% (~290 μg) fluorescencyjnego produktu ppp-RNA06-Cy3 (Rys. 1.3.10).



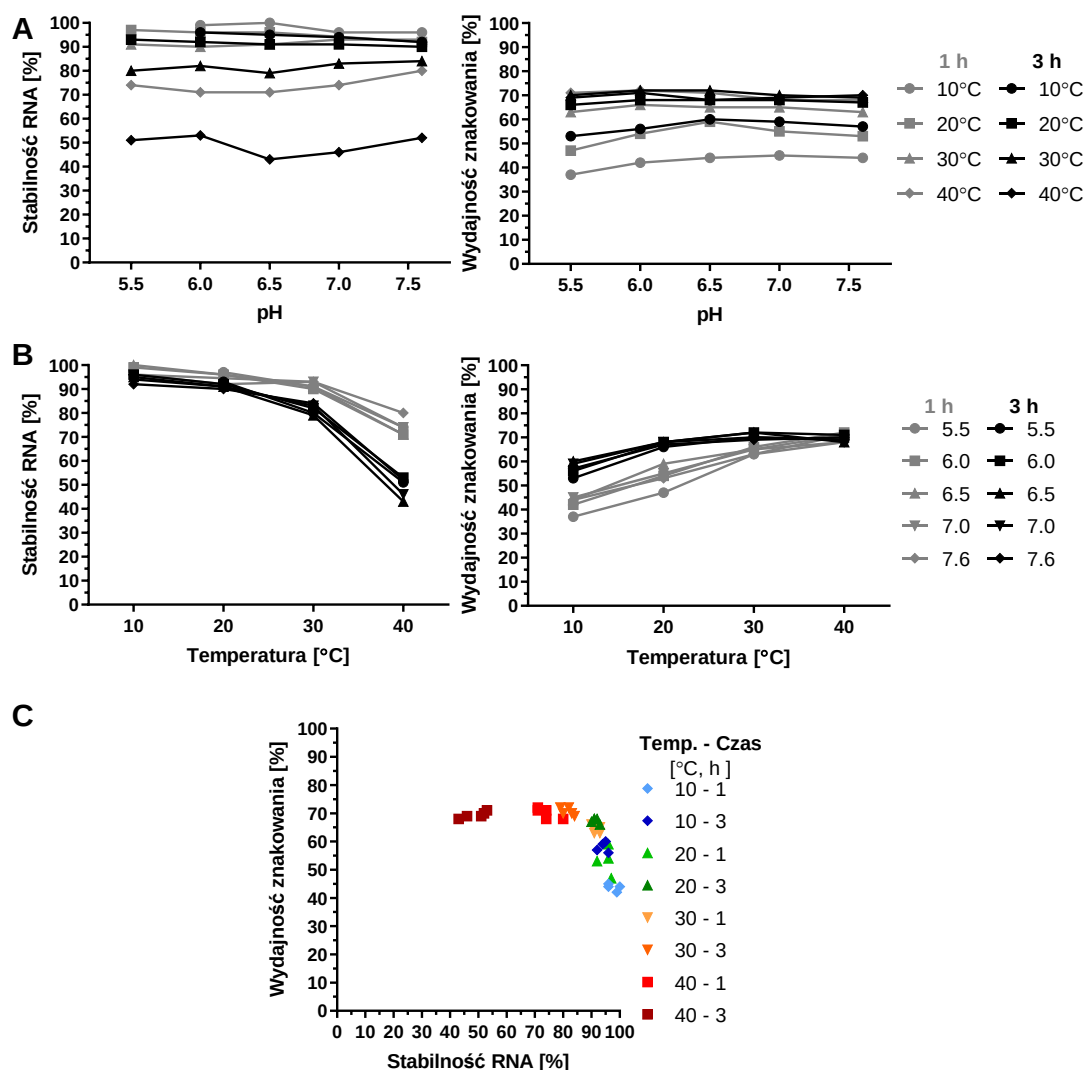
Rysunek 1.3.5: Znakowanie końca 3' oligonukleotydu RNA (ppp-RNA01, 35 nt) za pomocą Cy3-EDA . A) Schemat przebiegu eksperymentu. B-F) Profile HPLC produktów znakowania w zależności od stężenia Cy3-EDA (1–4 mM) oraz obliczone na ich podstawie wydajności znakowania.



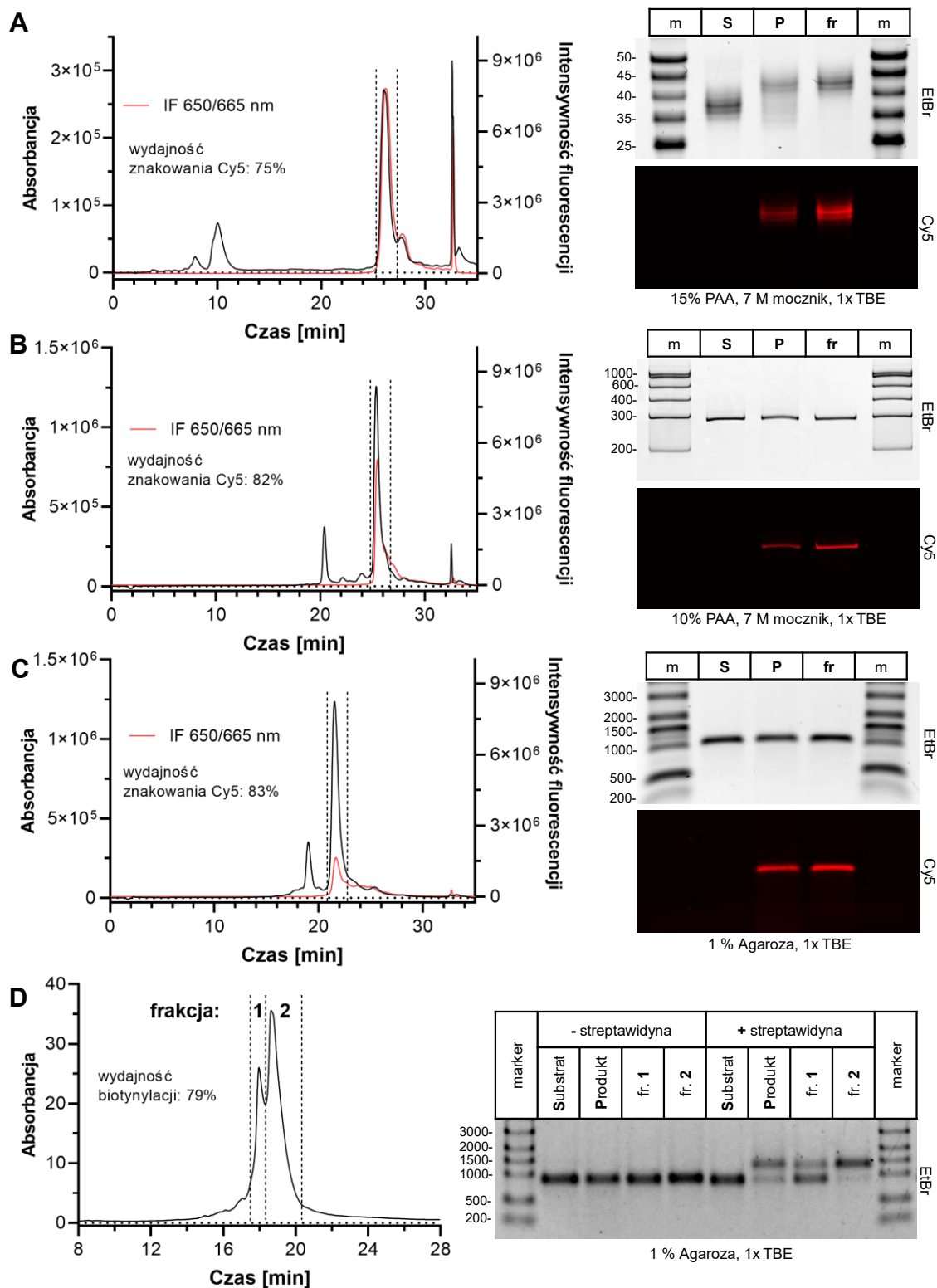
Rysunek 1.3.6: Optymalizacja znakowania RNA (RNA02, 276 nt) pod kątem pH, czasu i temperatury inkubacji. B-H) Profile HPLC produktów uzyskanych w pH 6.5.



Rysunek 1.3.7: Stabilność RNA (RNA02, 276 nt) w zależności od pH, czasu i temperatury inkubacji. B-H) Profile HPLC produktów degradacji uzyskanych w pH 6.5.

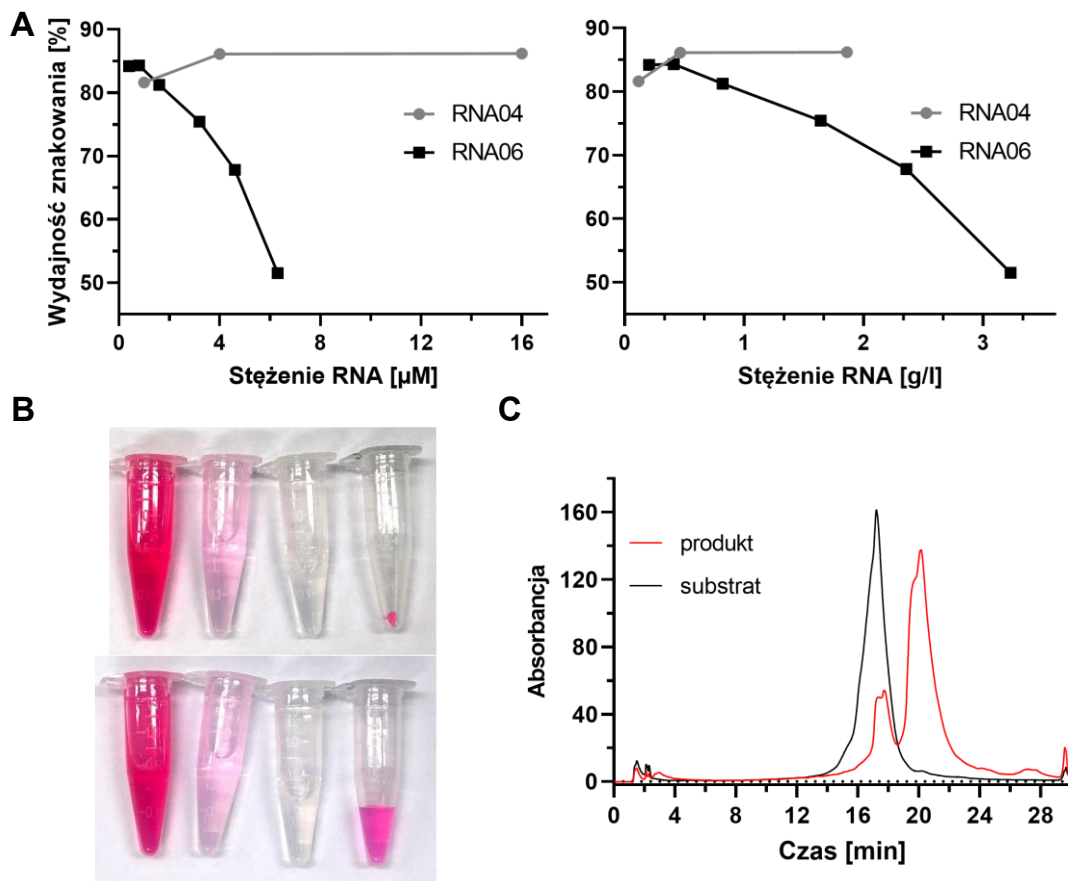


Rysunek 1.3.8: Wyniki optymalizacji znakowania RNA (RNA02, 276 nt). Stabilność RNA oraz wydajność znakowania w zależności od pH (A) lub temperatury (B). C) Mapa parametrów stabilności i wydajności dla wszystkich reakcji, pogrupowanych ze względu na czas i temperaturę inkubacji. Parametry stabilności (stosunek nieuszkodzonego RNA do całości RNA) i wydajności (stosunek znakowanego do sumy nieznakowanego i znakowanego RNA) wyznaczono na podstawie analizy chromatogramów HPLC.



Rysunek 1.3.9: Zmiana mobilności chromatograficznej RNA pod wpływem znakowania końca 3'. Profil HPLC oraz analiza elektroforetyczna produktów znakowania A) ArcaRNA01 (35 nt), B) ArcaRNA02 (276 nt) oraz C) ArcaRNA03 (1065 nt). Oznaczenia próbek naniesionych na żel: m – marker długości RNA, S – substrat RNA, P – produkty znakowania, fr – frakcje chromatograficzne (ich granice oznaczono za pomocą pionowych przerywanych linii na chromatogramach). D) Profil HPLC oraz analiza elektroforetyczna

w warunkach natywnych (test zmiany mobilność elektroforetycznej, EMSA) produktów 3' biotynylacji ArcaRNA10 (908 nt). RNA (~293 kDa) zawierające grupę biotyny wiąże się ze streptawidyną (55 kDa), przez co po dodaniu białka powstaje kompleks o zmniejszonej mobilności elektroforetycznej.



Rysunek 1.3.10: Znakowanie RNA w dużej skali. A) Wykresy zależności wydajności znakowania od stężenia ppp-RNA04 lub ppp-RNA06. B) Zdjęcia produktu znakowania w skali pół miligrama (ppp-RNA06-Cy3), po aminowaniu redukcyjnym i izolacji na drodze strącania. Namiarowy barwnik jest skutecznie usuwany w alkoholowym przesączu (pierwsze trzy próbówki od lewej) od produktu RNA (ostatnia próbówka od lewej, na górnym zdjęciu w formie stałej, poniżej w formie roztworu). C) Chromatogramy HPLC substratu i produktu znakowania ppp-RNA06 w skali pół miligrama.

Wydajności modyfikacji RNA metodą PORA z wykorzystaniem analogów EDA zostały przedstawione w Tabeli 3. Rezultaty optymalizacji doprowadziły do opracowania prostej, wydajnej oraz uniwersalnej metody modyfikacji końca 3' RNA. Prostotę zawdzięcza wykorzystywaniu dostępnych (handlowo bądź syntetycznie) odczynników, naturalnej struktury końca 3' RNA oraz izolacji poprzez strącanie. Jest wydajna i ekonomiczna ze względu na niski koszt odczynników (50–500 PLN na gram RNA), wysoką konwersję substratu w produkt (75–99%) oraz krótki czas (10–12 godzin) potrzebny na realizację całego procesu, od etapu IVT do produktu końcowego wysokiej

jakości (po HPLC). Jest wszechstronna, ponieważ toleruje szeroki zakres rozmiarów (3–2000 nt) i stężeń (0.2 μ M – 1 mM) RNA, oraz wachlarz modyfikacji, wykraczający poza zaprezentowane do tej pory barwniki fluorescencyjne, hapteny czy grupy bioortogonalne. W ten sposób zrealizowano pierwszy cel badawczy i zarazem utworzono podstawę do realizacji celów ostatecznych – podwójnego znakowania i cyklizacji RNA.

Tabela 3: Wydajności znakowania RNA określone za pomocą HPLC w zależności od warunków i rodzaju znakowania

Długość RNA [nt]	Stężenie RNA [μ M]	Warunki aminowania redukcyjnego ^a	Rodzaj oraz wydajność [%] znakowania		
			3' ^b	5' ^c	5'+3' ^d
3	100–1000	A	85–99	b.d.	b.d.
3	10–100	B	75–85	b.d.	b.d.
30–300	1–30	B	75–85	73–95	74–88
900–2000	0.2–3	B	75–85	74–88	46–56

^aWarunki aminowania redukcyjnego: A) 10 mM R-EDA, 100 mM NaBH₃CN, 500 mM KH₂PO₄ pH 6.0; B) 1 mM R-EDA, 20 mM NaBH₃CN, 100 mM KH₂PO₄ pH 6.0.

^bWydajności znakowania 3' metodą PORA. ^cWydajności znakowania 5' za pomocą reakcji SPAAC. ^dWydajności podwójnego znakowania (sprzężenie 5' SPAAC oraz 3' PORA). b.d. – brak danych. Metodologia znakowana 5' za pomocą reakcji SPAAC oraz podwójnego znakowania (5' + 3') zostały opisane w ramach rozdziału 2. Podwójne znakowanie RNA.

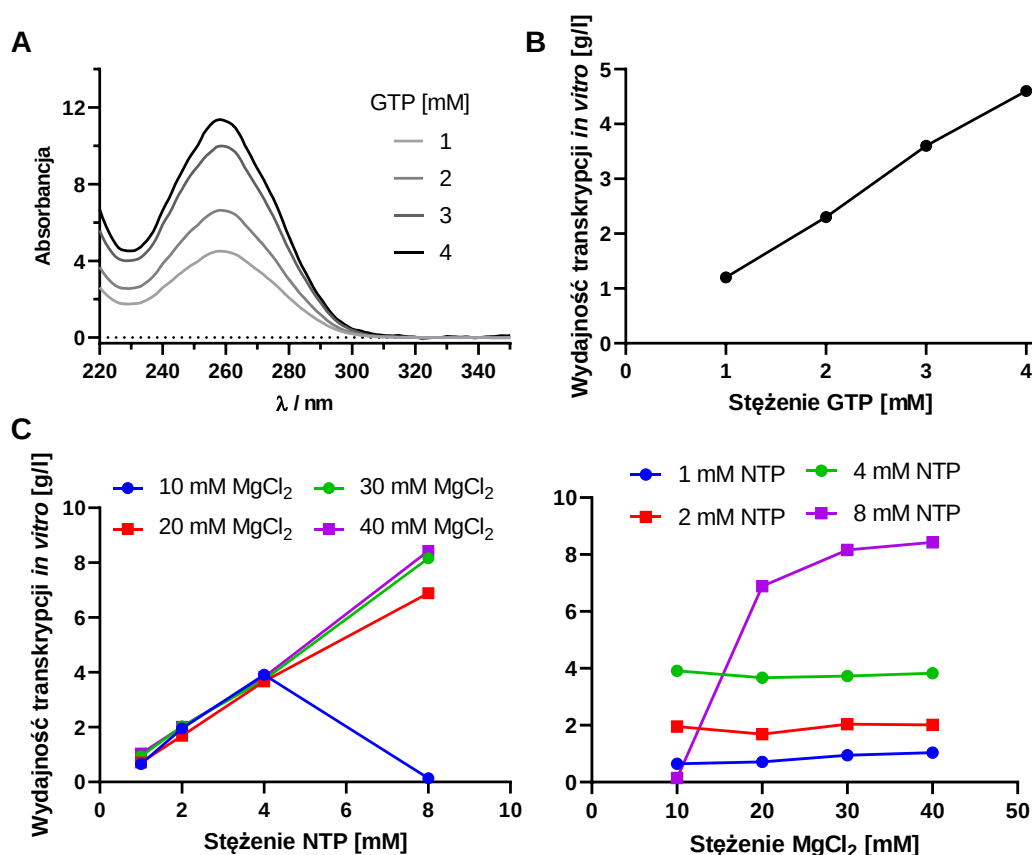
2. Podwójne znakowanie RNA

2. 1. Kotranskrypcyjna modyfikacja końca 5' i reakcja SPAAC

Dalsza realizacja niniejszego projektu badawczego wykorzystywała metodę transkrypcji *in vitro* (IVT). Zaletami IVT jest to, że można ją zastosować do syntezy RNA o (niemalże) dowolnej długości, nie wymaga stosowania zaawansowanego sprzętu (t. j. zautomatyzowany syntezer), toksycznych czy wrażliwych na wilgoć odczynników. Ponadto, jest szybka (od zmieszania odczynników do izolacji produktów w kilka godzin), oraz można ją wykorzystać do kotranskrypcyjnej modyfikacji RNA. Jednym ze sposobów na przeprowadzenie kotranskrypcyjnej modyfikacji jest zastosowanie analogów substratów polimerazy, pełniących funkcję inicjatora transkrypcji. W trakcie reakcji, polimeraza wykorzystuje analog jak pierwszy transkrybowany nukleotyd, w wyniku czego dochodzi do selektywnej modyfikacji końca 5' nowopowstałego RNA. Jedną z wad IVT jest złożoność reakcji — jej przebieg i końcowy rezultat zależą od wielu, często nieoczywiście powiązanych, czynników. Nawet przy zastosowaniu polimerazy RNA T7, czyli najdokładniej zbadanego i najczęściej stosowanego do IVT enzymu, przewidzenie ilości i jakości końcowych produktów jest trudne. Z tego powodu, optymalne warunki IVT trzeba wyznaczyć eksperymentalnie (dla danej sekwencji lub modyfikacji RNA), co wydarzyło się również w trakcie realizacji niniejszego projektu badawczego.

Zaobserwowano znane efekty, przykładowo zależność ilości produktów IVT od stężenia trifosforanów rybonukleozydów (NTP) czy chlorku magnezu(II) (Rys. 2.1.1). Wydajność IVT, mierzona jako całkowita ilość RNA wyprodukowanego na jednostkę objętości mieszaniny reakcyjnej (wyrażana w gramach na litr), zależy (najczęściej) liniowo od stężenia trifosforanu nukleozydu, będącego w niedomiarze. Ta obserwacja jest prawdziwa dla mieszanin IVT, w których obecne są dostatecznie wysokie stężenia chlorku magnezu oraz pirofosfatazy drożdżowej, która zapobiega powstawaniu pirofosforanu magnezu. Jest to trudno rozpuszczalny produkt uboczny hamujący aktywność polimerazy. W celu kotranskrypcyjnej modyfikacji 5' wprowadza się inicjator, co dalej komplikuje przebieg IVT. Inicjator konkuruje z pierwszym transkrybowanym nukleotydem (GTP lub ATP w zależności od sekwencji promotora). W celu zwiększenia wydajności wbudowywania inicjatora można zwiększyć jego stężenie. To może prowadzić do niepożądanych efektów, takich jak inhibicja polimerazy, współstrącenie z jonami magnezu czy nawet z matrycowym DNA. Te zjawiska utrudniają przeprowadzenie wydajnej transkrypcji. W praktyce, najbardziej użytecznymi substratami polimerazy (w tym

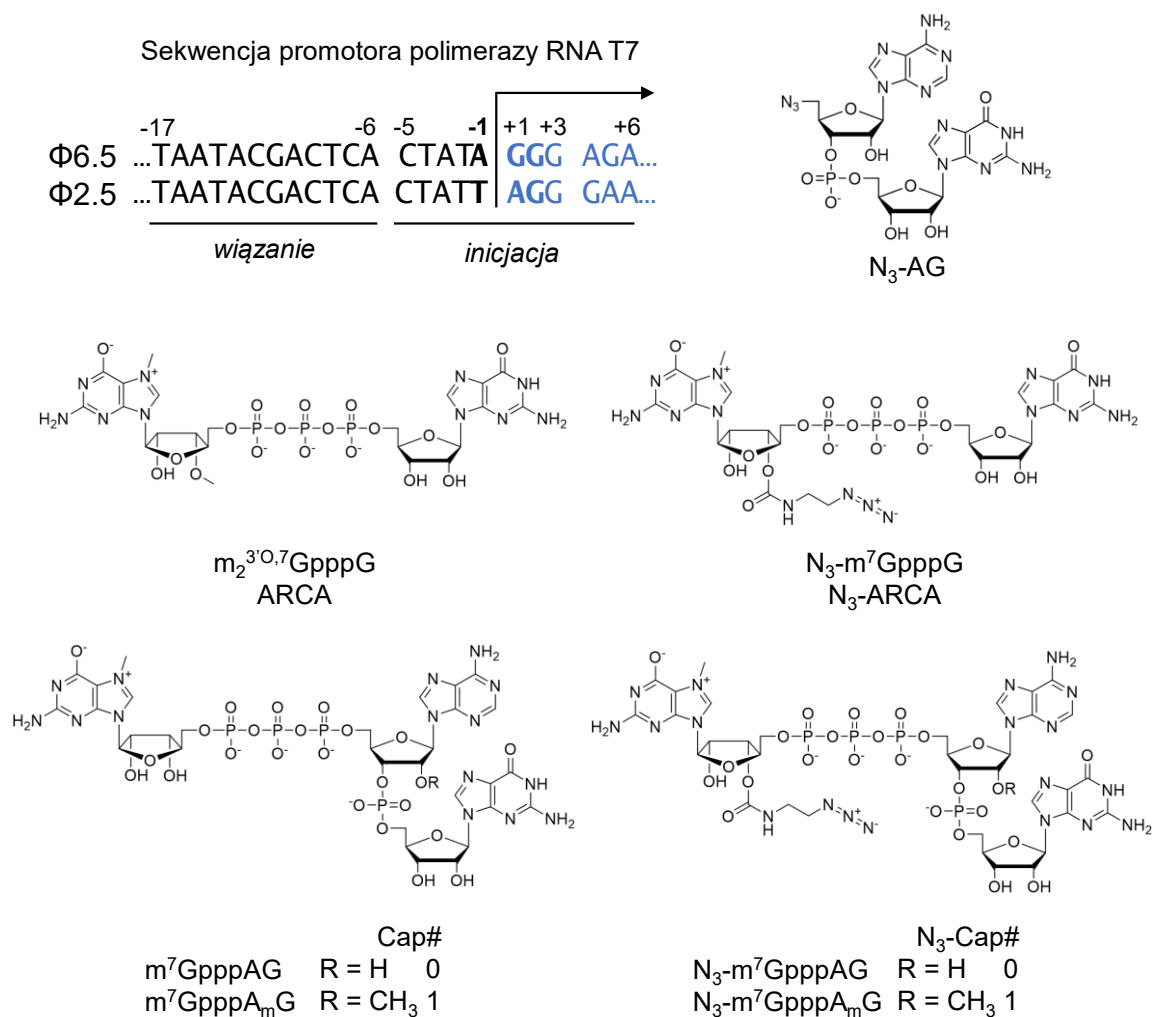
inicjatorami) są te, które nie zaburzają jej aktywności czy przebiegu transkrypcji w ogólności.



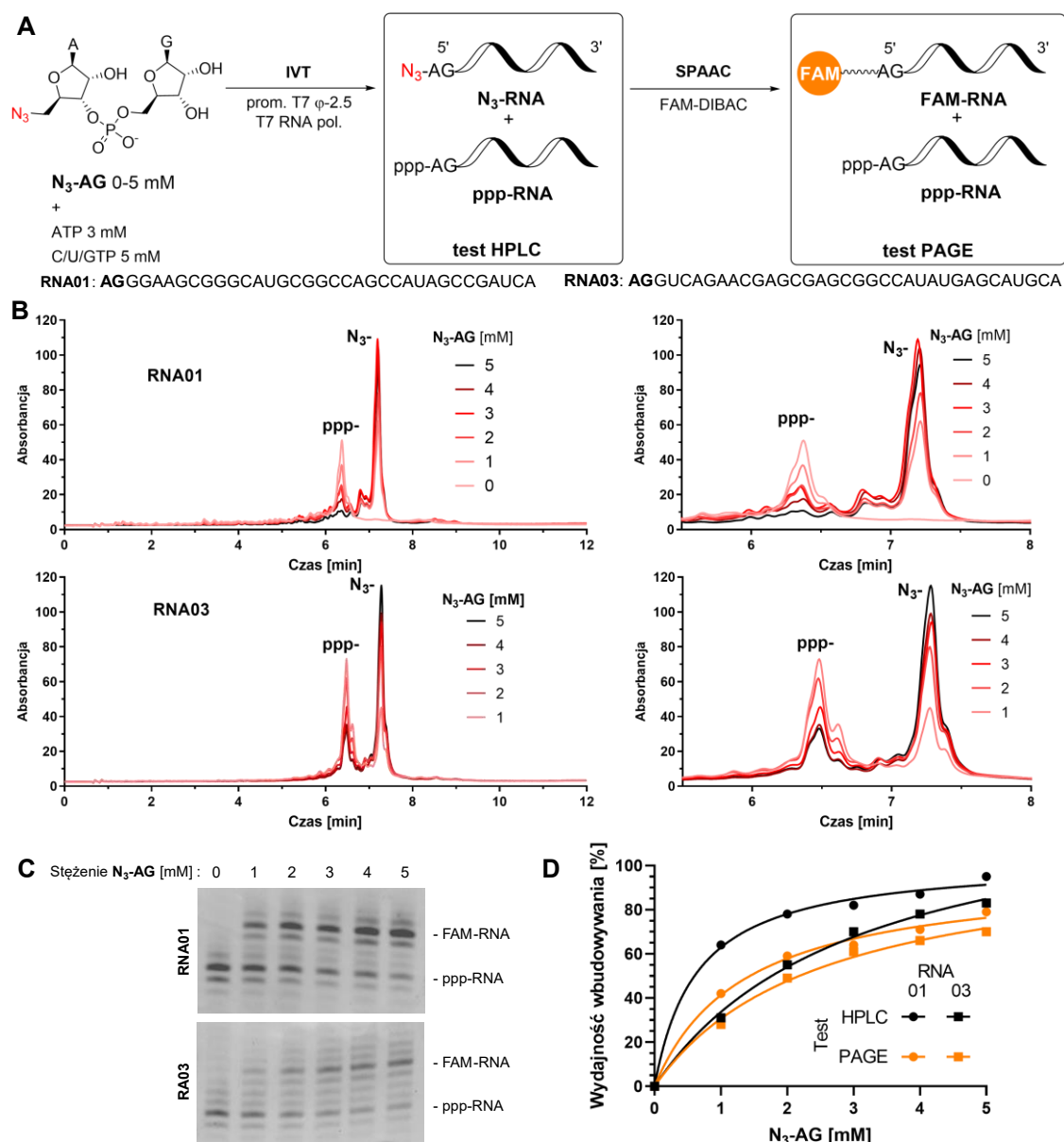
Rysunek 2.1.1: Wybrane czynniki mające wpływ na wydajność IVT. A) Widma absorpcyjne roztworów wyizolowanych produktów IVT w zależności od początkowego stężenia GTP w mieszaninie reakcyjnej (po 5 mM A/C/UTP). B) Wykres zależności wydajności IVT wyrażonej jako stosunek masy RNA wyizolowanego z mieszaniny reakcyjnej o określonej objętości. Wartość tę można określić na podstawie objętości mieszaniny reakcyjnej IVT oraz pomiarów widm absorpcyjnych wyizolowanych produktów. C) Zależność wydajności IVT od stężenia NTP lub MgCl₂. Wszystkie transkrypcje były prowadzone na matrycy RNA13 w temperaturze 37°C przez jedną (C) lub trzy godziny (A, B).

W kontekście znakowania, odpowiednimi reagentami są inicjatory zawierające grupę azydkową (Rys. 2.1.2). Grupa azydkowa jest stosunkowo mała, polarna i obojętna elektrycznie. To ułatwia jej umiejscowienie w obrębie struktury inicjatora w sposób minimalizujący ryzyko obniżenia rozpuszczalności, powstania zawady sterycznej i zaburzenia aktywności polimerazy. Przykładem takiego inicjatora jest N₃-ARCA (lub N₃-m⁷GpppG, Rys. 2.1.2), którego syntezę, właściwości i zastosowanie opublikowano już w 2017 roku.^[131] W ramach niniejszego projektu badawczego przeprowadzono syntezę nowych inicjatorów (N₃-AG, N₃-Cap0, N₃-Cap1), których strukturalnym motywem

przewodnim jest dinukleotyd AG. We wcześniejszych badaniach zaobserwowano, iż motyw AG wspomaga inicjację transkrypcji katalizowanej przez polimerazę T7.^[30] Postulowany mechanizm owego wspomagania polega na hybrydyzacji dinukleotydu AG w pozycjach -1 i +1 promotora typu $\phi 6.5$ (lub +1 i +2 dla $\phi 2.5$). W ramach niniejszego projektu, sprawdzono działanie inicjatora N₃-AG w kontekście inicjacji transkrypcji na promotorze typu $\phi 2.5$ (Rys. 2.1.3). Związek N₃-AG był wbudowywany preferencyjnie względem ATP (pierwszego transkrybowanego nukleotydu), nawet gdy był w niedomiarze. Ponadto, jego obecność zwiększała całkowitą ilość produktów transkrypcji, co sugeruje, iż N₃-AG działa jako swoisty stymulator tego procesu.



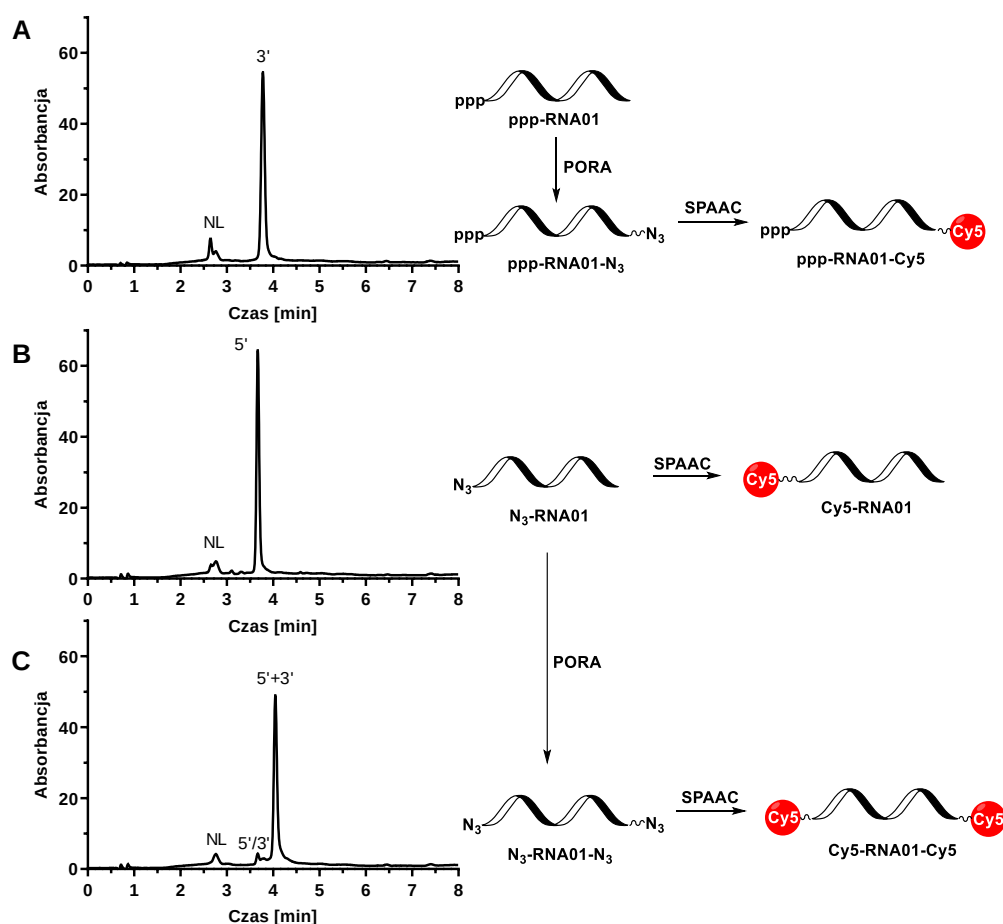
Rysunek 2.1.2: Struktury standardowych inicjatorów transkrypcji (ARCA, Cap0, Cap1), inicjatorów transkrypcji zawierających grupę azydkową (N₃-AG, N₃-ARCA, N₃-Cap0 i N₃-Cap1) oraz sekwencje promotorów $\phi 6.5$ i $\phi 2.5$ polimerazy RNA T7.



Rysunek 2.1.3: Inicjacja transkrypcji w obecności promotora ϕ 2.5 oraz inicjatora N_3 -AG. A) Schemat eksperymentu: po przeprowadzeniu IVT w obecności zmiennego stężenia N_3 -AG (0-5 mM), próbki były poddawane analizie HPLC (B) lub znakowaniu fluorescencyjnemu (reakcja SPAAC z FAM-DBCO) i następnej analizie PAGE (C). Za pomocą każdej z metod (HPLC lub PAGE) określono wydajność wbudowywania inicjatora N_3 -AG, w zależności od jego stężenia w mieszaninie reakcyjnej, dla różnych sekwencji RNA (RNA01 i RNA03) (D). Niezależnie od sekwencji, wydajność wbudowywania N_3 -AG przekracza 50%, nawet gdy jego stężenie jest niższe niż ATP (3 mM). Całkowita wydajność transkrypcji też wraść wraz ze wzrostem stężenia N_3 -AG, co jest widoczne na chromatogramach (B).

Wbudowany kotranskrypcyjny azydek można wykorzystać do posttranskrypcyjnej reakcji SPAAC. Reakcja SPAAC jest w tym celu odpowiednia, ponieważ do jej przeprowadzenia wystarczy zmieszać roztwory azydo-RNA (analog RNA zawierający

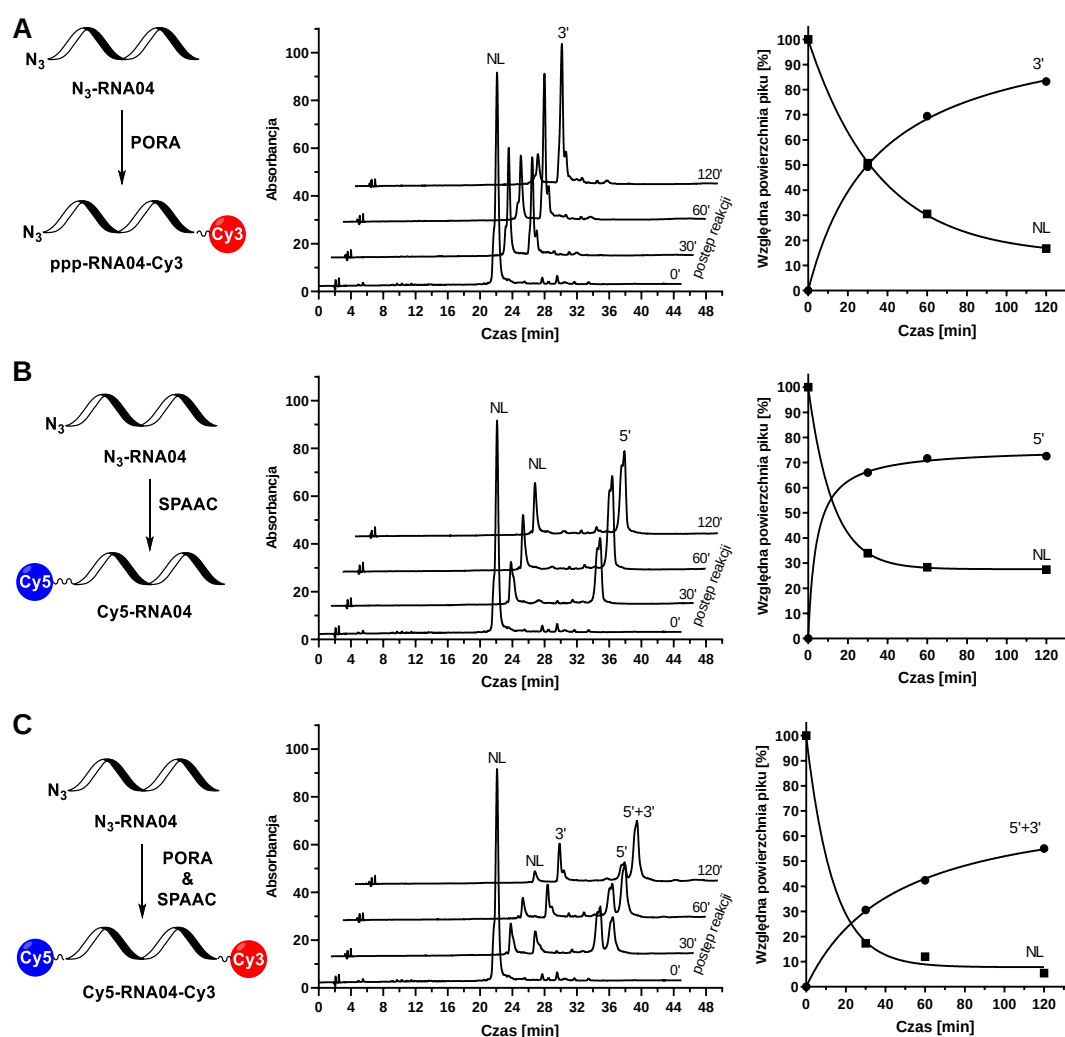
jedną lub wiele grup azydkowych) oraz pochodnej cyklooktynu. Przeprowadzono próby reakcji SPAAC na modelach N₃-RNA01, ppp-RNA01-N₃, oraz N₃-RNA01-N₃ (35 nt, Rys. 2.1.4). Znakowanie wybranych azydo-RNA zachodziło wydajnie (80-90%) już przy niskich stężeniach (0.5-2 mM) cyklooktynowej pochodnej barwnika Cy5. Obserwowana wydajność znakowania jest iloczynem wydajności kilku etapów - wprowadzania modyfikacji azydkowych oraz reakcji SPAAC. Reakcja SPAAC z udziałem N₃-RNA01-N₃ (Rys. 2.1.4C) jest przykładem podwójnego znakowania pozbawionego regioselektywności, w którym to obie grupy azydkowe (obecne na końcach 5' oraz 3' danego azydo-RNA) mogą reagować z cyklooktynową pochodną znacznika. Ta metodologia może być korzystna w szczególnych wypadkach, jednak w ramach niniejszego projektu dążono do opracowania wszechstronnej i regioselektywnej metody podwójnego znakowania. W tym celu podjęto próby sprzężenia reakcji SPAAC z metodą PORA.



Rysunek 2.1.4: Posttranskrypcyjne znakowanie azydo-RNA za pomocą reakcji SPAAC. Schematy reakcji i profile HPLC produktów znakowania końca 3' (A), końca 5' (B) oraz podwójnego znakowania (C) RNA o sekwencji RNA01 (35 nt).

2. 2. Sprzężenie reakcji SPAAC i metody PORA

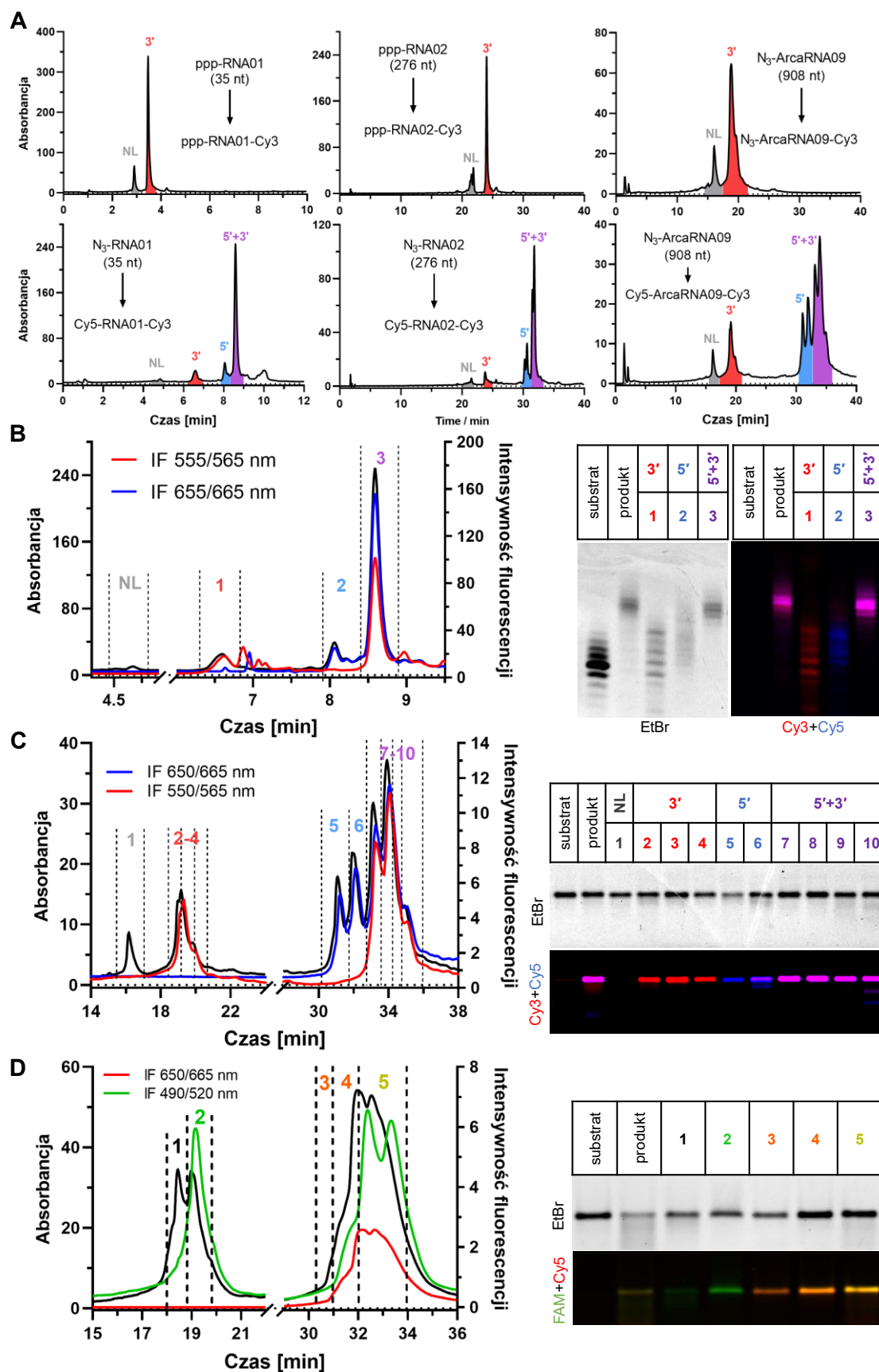
Sprzężenie reakcji SPAAC i PORA miało zapewnić regioselektywność (różne grupy funkcyjne na końcach 5' oraz 3') oraz wszechstronność (ze względu na posttranskrypcyjny dobór znaczników) metodzie podwójnego znakowania, ale również podnieść jej efektywność (poprzez zmniejszenie liczby etapów). W tym celu, reakcje SPAAC i PORA musiałyby zachodzić równolegle w jednym naczyniu. W związku z tym, przeprowadzono eksperyment, w którym N₃-RNA04 poddano równolegle znakowaniu 5' (SPAAC), znakowaniu 3' (PORA) albo podwójnemu znakowaniu (sprzężenie SPAAC i PORA, 5'+3', Rys. 2.2.1).



Rysunek 2.2.1: Sprzężenie reakcji SPAAC i PORA. Schemat reakcji, profile HPLC produktów reakcji oraz wykres względnej powierzchni pików substratu i końcowego produktu w czasie dla: znakowania 3' RNA04 za pomocą metody PORA (A), znakowania 5' za pomocą reakcji SPAAC (B) oraz podwójnego znakowania poprzez sprzężenie reakcji SPAAC i PORA (C).

Przebieg reakcji izolowanych (znakowania 5' albo 3') był podobny do przebiegu reakcji sprzężonych (równoległego znakowania 5' i 3' w jednym naczyniu). W sto dwudziestej minucie, konwersja substratu do produktu podwójnego znakowania (55%) była bliska iloczynowi wydajności znakowania 5' (73%) oraz 3' (83%). W tym wypadku, obniżona wydajność znakowania 5' (poniżej 80%) najprawdopodobniej wynikała z ograniczonej wydajności inicjacji transkrypcji z udziałem N₃-ARCA, nie zaś z obniżonej wydajności reakcji SPAAC.

Skuteczne sprzężenie reakcji SPAAC i PORA pozwoliło na otrzymanie szeregu wariantów podwójnie znakowanego RNA. Obejmowały one, między innymi pary znaczników (Cy5/Cy3, Cy5/Biotyna lub Cy5/FAM) oraz sekwencje o długościach w zakresie 30–1100 nukleotydów (Rys. 2.2.1–2.2.3). Zaobserwowano rozdział produktów podwójnego znakowania w warunkach HPLC. Rozdzielczość rozdziału była dostatecznie wysoka by wyizolować homogenne produkty przejściowe i ostateczne, oraz by mierzyć ilościowo wydajność poszczególnych reakcji. Zestawienie wydajności znakowania metodami PORA oraz SPAAC zamieszczono w Tabeli 3. Efekty zaobserwowane w trakcie eksperymentów skłoniły do sformułowania hipotezy o „efekcie kotwiczenia”, której poświęcony jest następny rozdział.



Rysunek 2.2.2: Rozdział HPLC produktów podwójnego znakowania. A) Profile HPLC produktów znakowania 3' oraz podwójnego znakowania sekwencji RNA01 (35 nt), RNA02

Rysunek 2.2.3: Znakowanie fluorescencyjne i biotynylacja RNA02 (276 nt). Chromatogramy HPLC: A) substratu (N₃-RNA02); B) produktu znakowania 5' (Cy5-RNA02); C) produktu 3' biotynylacji (N₃-RNA02-Biot) oraz D) podwójnego znakowania (Cy5-RNA02-Biot). E) analiza elektroforetyczna w warunkach natywnych (test zmiany mobilność elektroforetycznej, EMSA) wyżej wymienionych produktów. RNA (~89 kDa) zawierające grupę biotyny wiąże się ze streptawidyną (55 kDa), przez co po dodaniu białka powstaje kompleks o zmniejszonej mobilności elektroforetycznej.

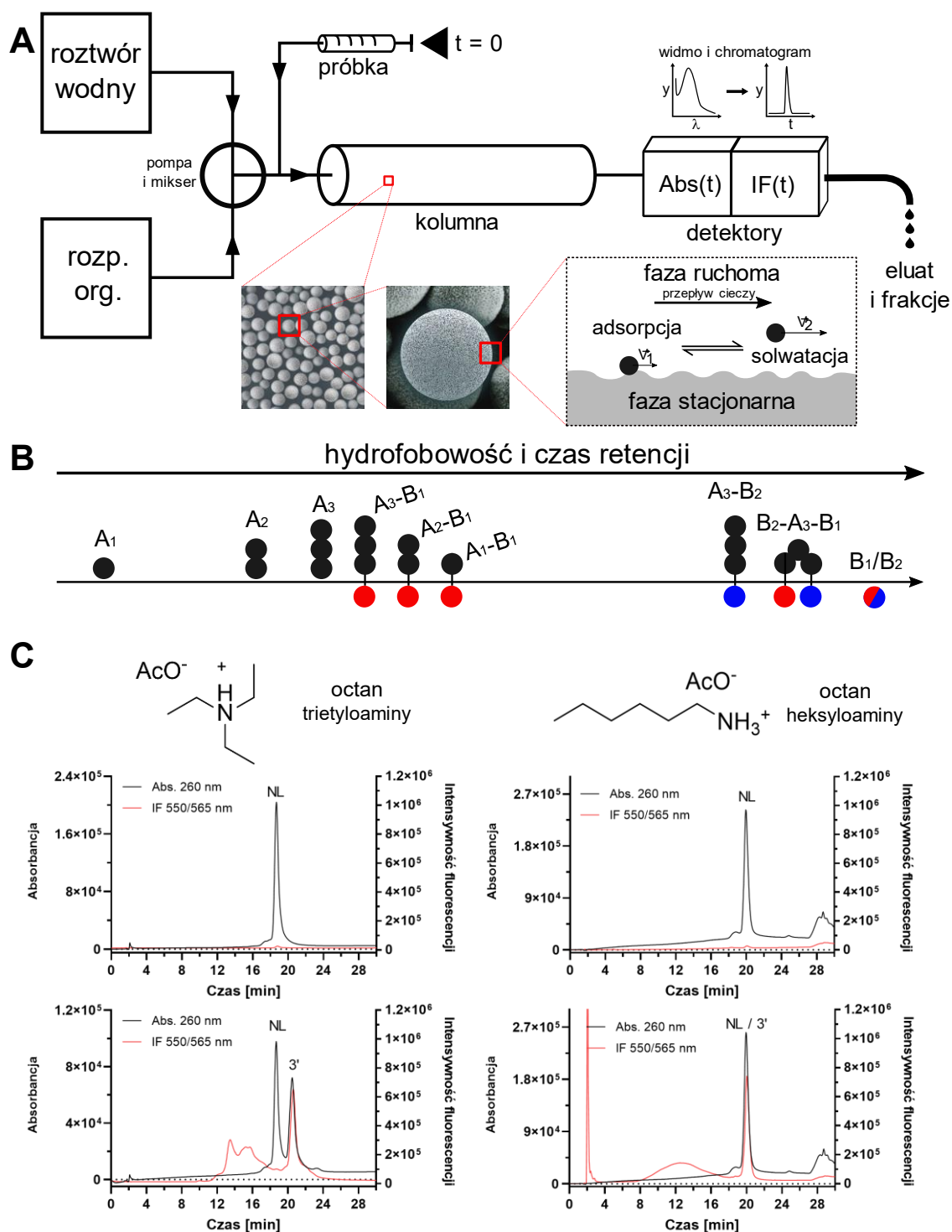
Rysunek 2.2.3: Znakowanie fluorescencyjne i biotynylacja RNA02 (276 nt). Chromatogramy HPLC: A) substratu (N₃-RNA02); B) produktu znakowania 5' (Cy5-RNA02); C) produktu 3' biotynylacji (N₃-RNA02-Biot) oraz D) podwójnego znakowania (Cy5-RNA02-Biot). E) analiza elektroforetyczna w warunkach natywnych (test zmiany mobilność elektroforetycznej, EMSA) wyżej wymienionych produktów. RNA (~89 kDa) zawierające grupę biotyny wiąże się ze streptawidyną (55 kDa), przez co po dodaniu białka powstaje kompleks o zmniejszonej mobilności elektroforetycznej.

2. 3. Efekt „kotwiczenia”

W układach RP-HPLC można zaobserwować zjawisko analogicznie jak w trakcie elektroforezy kwasów nukleinowych: zwiększenie stopnia polimeryzacji powoduje zmniejszenie szybkości migracji w układzie chromatograficznym. W konsekwencji cząsteczki RNA o większej masie molowej mają dłuższe czasy retencji. Na tej podstawie można by sformułować hipotezę, iż wydłużenie czasu retencji znakowanego RNA jest spowodowane zwiększeniem masy molowej, które towarzyszy przyłączeniu reszty barwnika. Taka hipoteza może być zgodna z efektami obserwowanymi w przypadku znakowania nukleotydów i oligonukleotydów, gdzie przyłączony barwnik stanowi duży udział w całkowitej masie cząsteczki (5–50% dla 1–20 nt). Zgodnie z tym, w wyniku znakowania wielkocząsteczkowego RNA zmiana czasu retencji powinna być marginalna (wzrost masy o ~0.1% dla 1000-2000 nt). Zaprezentowane w niniejszej rozprawie wyniki temu jednoznacznie zaprzeczają. Zaobserwowano, iż czynnikiem decydującym o różnicach w czasach retencji modyfikowanego i niemodyfikowanego RNA, nie jest stopień polimeryzacji, a hydrofobowość (oraz liczba) wprowadzonych modyfikacji. Tę obserwację lepiej wyjaśnia efekt „kotwiczenia” (Rys. 2.3.1).

Dłuższe czasy retencji w układach RP-HPLC utożsamiane są z większym parametrem hydrofobowości danej cząsteczki. W trakcie rozdzielania, cząsteczki są w równowadze między dwoma stanami — zaadsorbowane na hydrofobowej powierzchni fazy stacjonarnej lub solwatowane w hydrofilowej fazie ruchomej. Na położenie równowagi wpływają właściwości cząsteczki oraz warunki chromatografii. Hydrofobowość można parametryzować za pomocą czasu retencji (RT) danej cząsteczki. Efekt kotwiczenia może wystąpić dla cząsteczki typu A-B, składającej się z dwóch części (A i B), które w stanie wolnym i w jednakowych warunkach znacznie różnią się czasami retencji (np. $RT_A \ll RT_B$). Hipotetycznie, istniałyby warunki, w których cząsteczka A (o niskim RT) dominuje w stanie solwatowanym, a cząsteczka B (o wysokim RT) dominuje w stanie adsorbowanym. W tych warunkach, oddziaływanie hydrofobowego fragmentu (B) cząsteczki A-B z fazą stacjonarną zapobiegałoby pełnej solwatacji i uwolnieniu cząsteczki A-B do fazy ruchomej. Zgodnie z tym, im większa różnica między RT_A a RT_B , tym większa różnica między RT_A a RT_{A-B} . Jeżeli uznamy cząsteczkę A za niemodyfikowane RNA, cząsteczkę B za znacznik, a cząsteczkę A-B za znakowane RNA, to decydujący wpływ na różnicę w czasach retencji znakowanego i niemodyfikowanego RNA miałby czas retencji znacznika, nie zaś stopień polimeryzacji owego RNA.

W momencie, gdy znacząco zwiększymy hydrofobowość związku A lub zmienimy warunki chromatografii, tak by selektywnie zwiększyć jego czas retencji, to efekt kotwiczenia powinien ulec osłabieniu lub zniwelowaniu. To zjawisko zaobserwowano dla RNA za pośrednictwem zmiany przeciwjonu. W celu poprawienia jakości rozdzielów RNA za pomocą RP-HPLC, w trakcie chromatografii stosowano warunki denaturujące (temperatura powyżej 50 °C, acetonitryl w składzie fazy ruchomej) oraz pary jonowe (sole octanowe amin alifatycznych). Aminy alifatyczne, które w sprotonowanej formie są lipofilowymi kationami, wchodzą w oddziaływania zarówno z fazą stacjonarną jak i z RNA. Wysycenie nimi układu chromatograficznego zwiększa retencję (przesuwa położenie równowagi w stronę stanu zaadsorbowanego) RNA i zwiększa powtarzalność rozdzielów. Poprzez dobór pary jonowej można wpływać na retencję RNA (Rys. 2.3.B). W obecności octanu trietyloaminy elucja RNA06 (1065 nt) zachodzi w niższym stężeniu acetonitrylu (~15%) niż w obecności octanu heksyloaminy (~40%). Kiedy przeprowadzono rozdział mieszaniny RNA06 oraz RNA06-Cy3 w obecności octanu trietyloaminy zaobserwowano separację znakowanego i nieznakowanego RNA. W obecności octanu heksyloaminy owej separacji nie zaobserwowano, co jest kolejnym argumentem świadczącym o występowaniu „efektu kotwiczenia”.



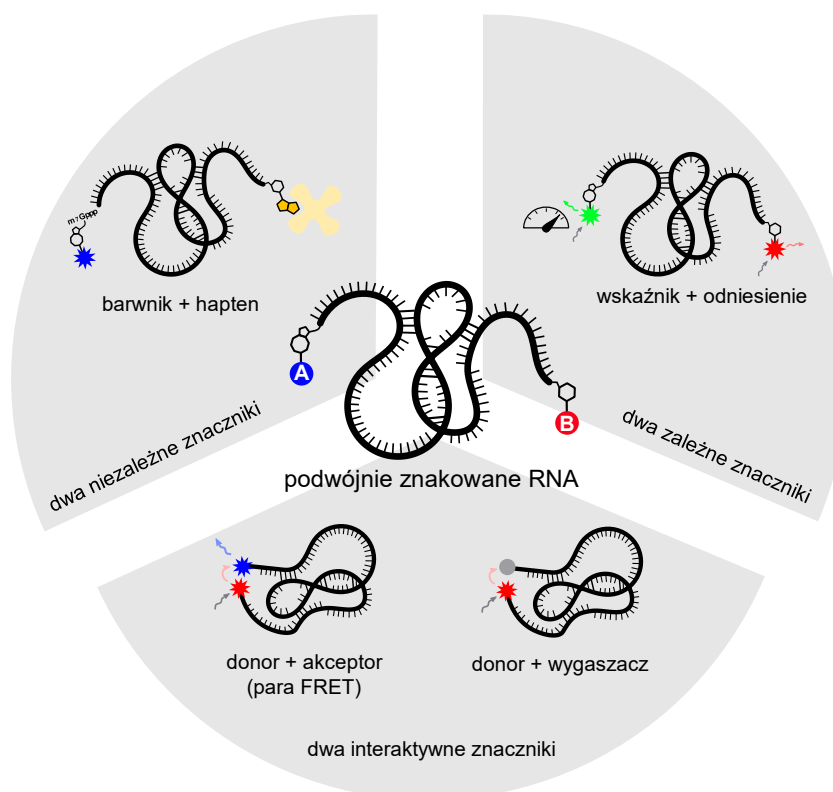
Rysunek 2.3.1: Efekt „kotwiczenia”. A) Schemat układu i mechanizmu rozdzielania HPLC. Centralnym elementem układu jest kolumna wypełniona złożem (faza stacjonarna), które jest obmywane przez ciecz (faza ruchoma) pod wysokim ciśnieniem. Substancje zawarte w próbce odwracalnie oddziałują z fazą stacjonarną i ruchomą na zasadzie adsorpcji i solwatacji, a położenie owej równowagi ma bezpośrednie przełożenie na szybkość migracji (mobilność chromatograficzną) i czas retencji. B) Zależność czasu retencji od struktury przy zastosowaniu efektu „kotwiczenia”. Związki o niskiej hydrofobowości (A_1 - A_3) odpowiadają RNA o różnych stopniach polimeryzacji. Związki o wysokiej hydrofobowości

(B₁, B₂) odpowiadają wolnym barwnikom fluorescencyjnym. Związki typu A-B odpowiadają cząsteczkom powstałym poprzez kowalencyjne połączenie RNA i barwnika. W tym układzie, czas retencji cząsteczki typu A-B jest wypadkową czasów retencji izolowanych cząsteczek A i B. C) Chromatogramy HPLC zarejestrowane dla nieznakowanego RNA06 (NL, górne chromatogramy) lub mieszaniny znakowanego (3') i nieznakowanego RNA06 (dolne chromatogramy) w układach wysyconych octanem trietyloaminy (lewy pion) lub heksyloaminy (prawy pion). W obecności octanu heksyloaminy znakowane i nieznakowane RNA mają jednakowy czas retencji.

3. Sondy molekularne RNA

3. 1. Podwójnie znakowane RNA jako sondy molekularne

W poprzednim rozdziale opisano metodologię otrzymywania dwuznakowanego RNA. W tym rozdziale zaprezentowano potencjalne zastosowania dwuznakowanego RNA jako sond molekularnych (Rys. 3.1.1).

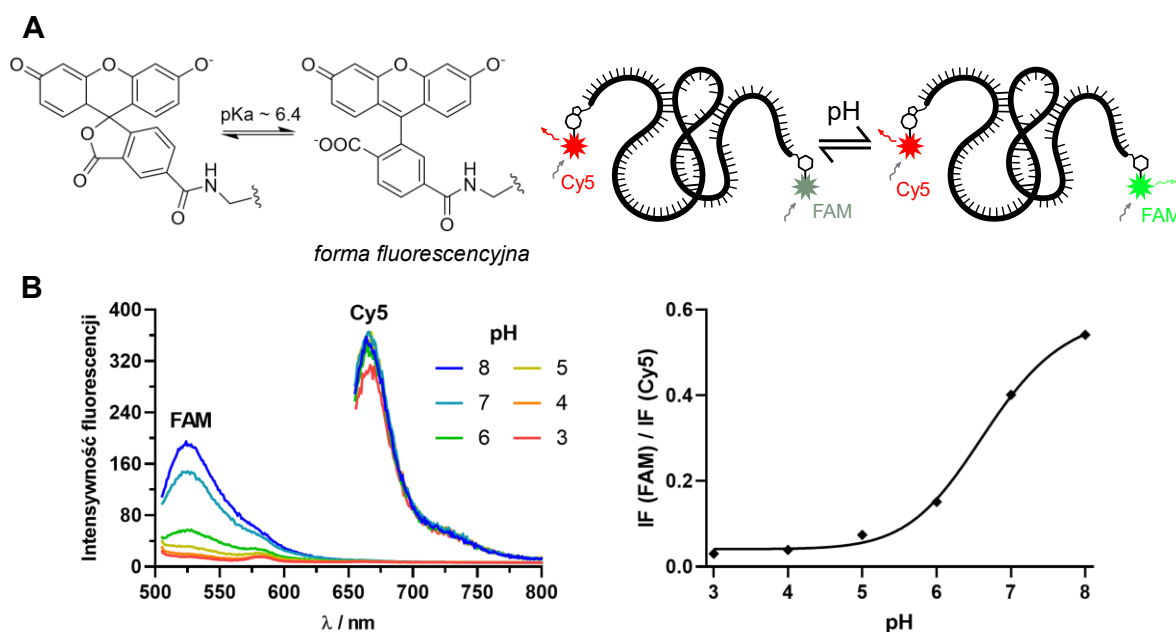


Rysunek 3.1.1: Podwójnie znakowane RNA jako sondy molekularne. Wyróżniono pary znaczników niezależnych (jak barwnik i hapten), znaczników (niebezpośrednio) zależnych (jak barwnik-wskaźnik i barwnik odniesienia) oraz znaczników interaktywnych (bezpośrednio zależnych, pary donor i akceptor/wygaszacz).

Zdolność do regioselektywnego umieszczenia dwóch modyfikacji w obrębie RNA dała możliwość tworzenia sond wyposażonych w funkcjonalne pary znaczników. Można wyróżnić pary znaczników niezależnych, zależnych niebezpośrednio i bezpośrednio (pary interaktywne). Przykładem pary niezależnych znaczników jest para barwnika umożliwiającego wykrywanie (np. barwnik fluorescencyjny) z haptenem (np. biotyną), który niezależnie wchodzi w interakcje z elementami układu eksperymentalnego (np. ze złożem ze streptawidyną). Przykładem pary znaczników zależnych niebezpośrednio jest para barwnika wskaźnikowego, czułego na zmiany środowiskowe, z barwnikiem odniesienia, który nie zmienia swoich właściwości w trakcie eksperymentu. Przykładem pary interaktywnych znaczników jest para donora i akceptora w zjawisku FRET. Ponadto,

swoistym rdzeniem sond był polimer RNA, co pozwoliło na ich zastosowanie w badaniach struktury i właściwości RNA oraz w eksperymentach, których podmiotem były cząsteczki wchodzące w interakcje z RNA. Ostatecznie, zostały one również zastosowane w badaniach obejmujących bioobrazowanie procesów dostarczania i metabolizmu mRNA, co opisano w kolejnym podrozdziale. Działanie sondy, zawierającej parę niezależnych znaczników (Cy5 i biotynę) zademonstrowano we wcześniejszej części tekstu (Rys. 2.2.3).

Przykładem sondy środowiskowej jest RNA zawierające dwa barwniki fluorescencyjne, gdzie fluorescencja pierwszego (wskaźnika) jest zależna od pH, natomiast fluorescencja drugiego (odniesienia) nie zależy od odczynu środowiska (Rys. 3.1.2). Wybrano FAM jako barwnik pH-zależny, którego zmiana w intensywności fluorescencji wynika z dysocjacji grupy karboksylowej ($pK_a \sim 6.4$), sprzężonej z układem elektronów π chromofora.^[204] Jako barwnik referencyjny wybrano Cy5. Za pomocą spektrofluorymetru zmierzono widma fluorescencji sondy Cy5-ArcaRNA06-FAM w buforowanych roztworach o pH w zakresie wewnątrzkomórkowym (3-8). Zaobserwowano stukrotną



Rysunek 3.1.2: Podwójnie znakowane RNA jako sonda środowiskowa. A) Dysocjacja fluoresceiny oraz schemat działania sondy RNA wyposażonej w barwnik-wskaźnik (FAM) oraz barwnik odniesienia (Cy5). B) Widma fluorescencyjne (FAM: wzbudzenie 490 nm, Cy5: wzbudzenie 650 nm) oraz krzywa miareczkowania wykonana dla sondy Cy5-ArcaRNA06-FAM.

zmianę intensywności fluorescencji FAM, towarzyszącą zmianie odczynu środowiska sondy. Jednocześnie, intensywność fluorescencji Cy5 pozostawała bez znaczących zmian. Stosunek intensywności fluorescencji FAM do Cy5 w zależności od pH jest krzywą

miareczkowania sondy. Jej punkt przeięcia odpowiada pH w okolicach 6.5, co odpowiada pKa wskaźnikowego barwnika. Niestety, sonda nie znalazła skutecznego zastosowania w pomiarach wewnątrzkomórkowych, ze względu na wysokie tło w kanale FAM.

Kolejnym przykładem są sondy FRET, zawierające barwniki Cy3 oraz Cy5. Fluorescencja sondy daje wgląd w odległość między końcem 5' a końcem 3' RNA. Tego typu sondy wykorzystano zarówno do badania właściwości wybranych struktur RNA jak i do eksperymentów z udziałem enzymów nukleolitycznych. Sondy o analogicznej strukturze były wykorzystywane do wyznaczenia bezwzględnych odległości między końcami 5' i 3' sekwencji mRNA i lncRNA.^[120] Autorzy wykorzystali, między innymi, zaawansowaną analizę danych spektroskopowych^[205], co pozwoliło im na wyznaczenie kwantowej wydajności rezonansowego transferu energii. Mimo zastosowania sond heterogennych, których stosunek Cy5 i Cy3 był daleki od jedności (0.2–0.3), udało im się zauważyć, że odległość między końcami 5' i 3' RNA jest zasadniczo niezależna od jego stopnia polimeryzacji czy sekwencji, oraz że waha się w przedziale 2–6 nm.

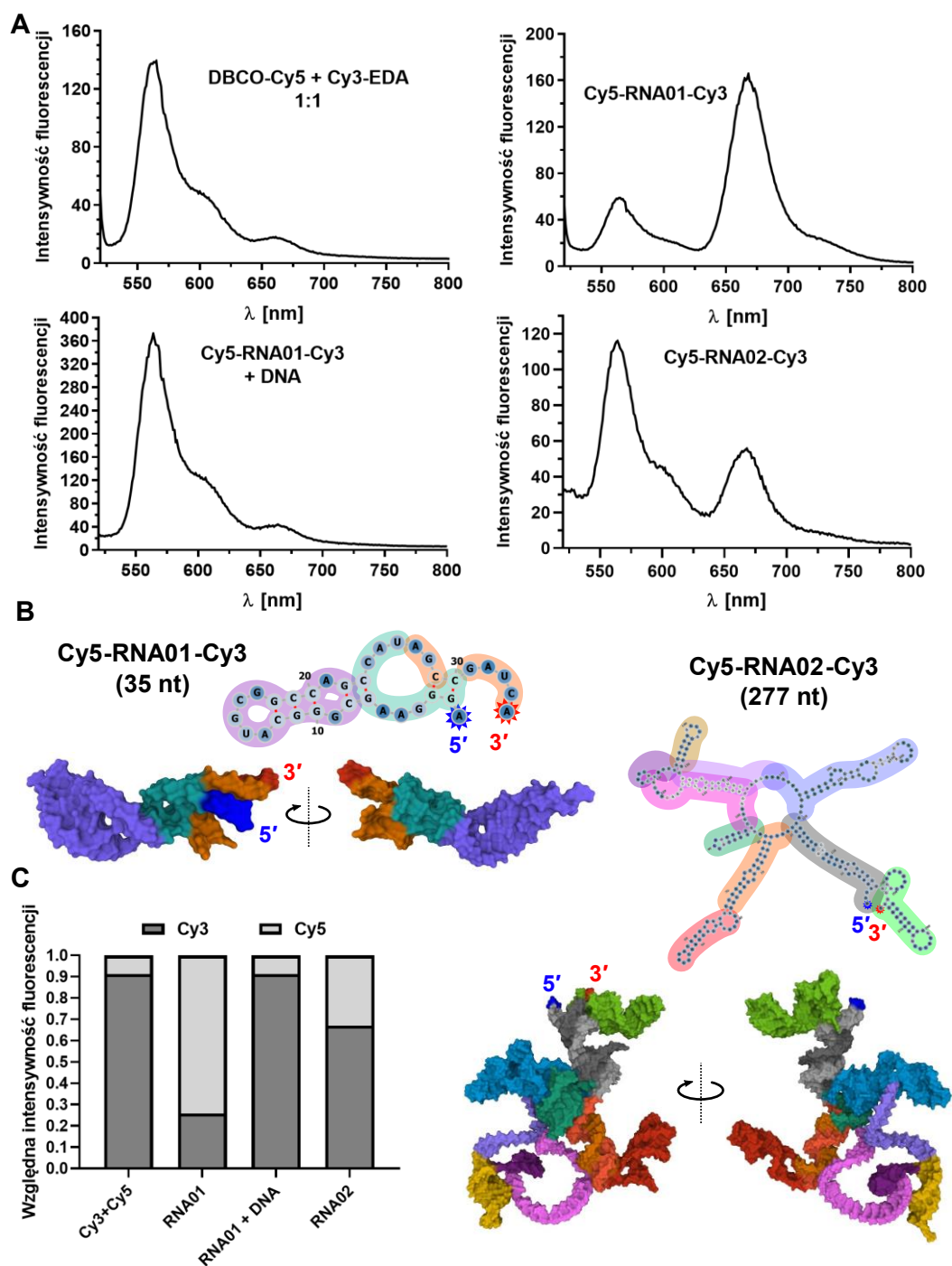
W ramach niniejszej rozprawy, za pomocą homogennych sond FRET i prostej analizy danych przedstawiono zmiany odległości między końcami RNA pod wpływem wybranych czynników, takich jak zmiana temperatury, składu roztworu czy aktywności nukleaz. Miarą owej odległości jest parametr B (jak bliskość), obliczony na podstawie wzoru:

$$B = \frac{IF_{Cy5}}{IF_{Cy3} + IF_{Cy5}}$$

gdzie IF_{Cy5} to intensywność maksimum emisji akceptora (Cy5, średnia intensywność w przedziale 660-670 nm), a IF_{Cy3} to intensywność maksimum emisji donora (Cy3, średnia intensywność w przedziale 560-570 nm), przy selektywnym wzbudzeniu donora (500 nm). Ze względu na naturę zjawiska FRET, wartość parametru B, która znajduje się w przedziale 0–1, powinna być odwrotnie skorelowana z odległością w zakresie 1–10 nm.^[205] Ponadto, podjęto próbę porównania wyników eksperymentalnych z modelami struktur drugorzędowych oraz trzeciorzędowych uzyskanych za pomocą metod obliczeniowych dostępnych w sieci (RNAfold oraz RNAcomposer).^[14, 206]

W celu zademonstrowania metodologii zastosowano przykłady pomiarów i obliczeń dla dwóch sond: Cy5-RNA01-Cy3 oraz Cy5-RNA02-Cy3 (Rys. 3.1.3). Gdy sonda Cy5-RNA01-Cy3 (35 nt) jest w stanie wolnym, końce 5' i 3' są stosunkowo blisko ($B \sim 0.67$). Po hybrydyzacji DNA na całej długości sondy końce oddalają się na odległość ponad 9 nm (orientacyjna długość 35 par zasad podwójnej helisy DNA/RNA). To odpowiada

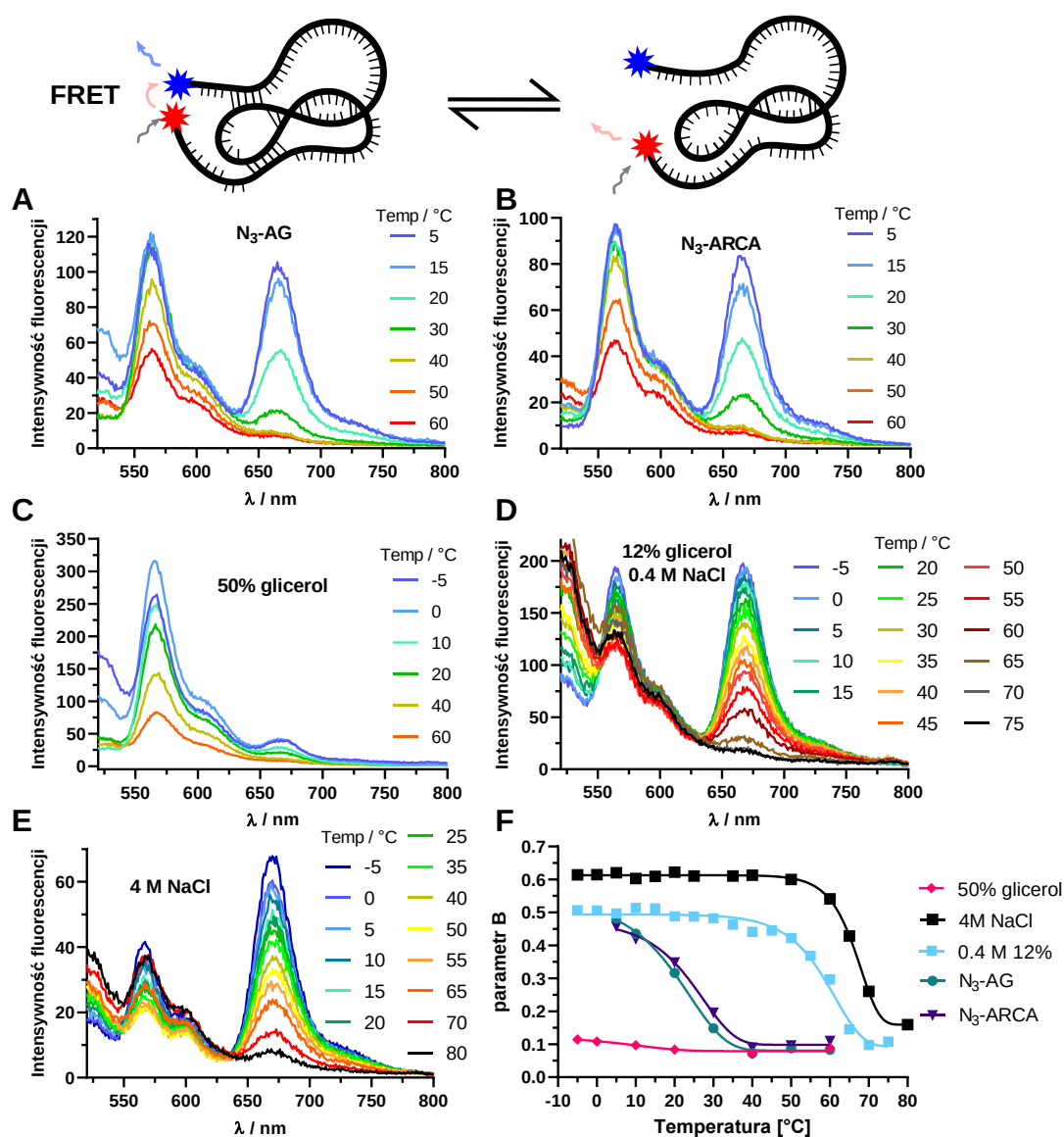
wyznaczonej wartości B mniejszej bądź równej ~ 0.1 , co również charakteryzuje widmo fluorescencyjne roztworu wolnych barwników Cy3 i Cy5. Dla sondy Cy5-RNA02-Cy3 (276 nt) w stanie wolnym wartość B to ~ 0.30 . Modele struktury drugorzędowej pozwalają lepiej przewidzieć, które elementy sekwencji mogą oddziaływać ze sobą poprzez parowanie zasad azotowych. Modele struktur trzeciorzędowych pomagają wyobrazić sobie jak poszczególne motywy strukturalne mogą rozkładać się w przestrzeni. Połączenie obu modeli pomaga zaproponować interpretacje wyników pomiarów. Sekwencja sondy Cy5-RNA01-Cy3 mogłaby tworzyć dwa motywy strukturalne: spinę oraz otwartą pętlę. Końce 5' i 3' miałyby znajdować się w motywie otwartej pętli. Sekwencja sondy Cy5-RNA02-Cy3 najpewniej tworzy więcej motywów, przez co końce 5' i 3' mogłyby znaleźć się w dwóch różnych motywach. To potencjalnie mogłoby utrudniać kontakt barwników i tłumaczyć niższą wartość parametru B. Nawet gdy zastosowanie metod obliczeniowych okazuje się pomocne w interpretacji wyników eksperymentów, należy podchodzić do nich z dystansem.



Rysunek 3.1.3: Zastosowanie sond FRET do pomiarów bliskości końców 5' i 3'. A) Widma fluorescencyjne wolnych barwników DBCO-Cy5 i Cy3-EDA, sondy Cy5-RNA01-Cy3, sondy Cy5-RNA01-Cy3 po hybrydyzacji z komplementarnym DNA, oraz sondy Cy5-RNA02-Cy3. B) Modele struktur drugorzędowych i trzeciorzędowych RNA o sekwencjach RNA01 i RNA02. C) Wartości względnych intensywności fluorescencji donora (Cy3, 565 nm) i akceptora (Cy5, 668 nm) pary FRET, które stanowią podstawę do obliczenia parametru B (bliskości końców 5' i 3' sondy).

Stosując wyżej opisaną metodologię można zbadać podstawowe właściwości danej struktury RNA i podjąć próby lepszego zrozumienia jej dynamiki. Przykładem są pomiary

temperatury topnienia (tj. denaturacji bądź „rozplecenia” struktur drugorzędowych kwasów nukleinowych) w zmiennych warunkach. Przedstawiono pomiary widm fluorescencyjnych sondy Cy5-RNA02-Cy3 lub Cy5-ArcaRNA02-Cy3 wykonane w szerokim zakresie temperatur (Rys. 3.1.4A i B). Dla obu sond, których struktury różniły się obecnością 5'-kapu, wyjściowa wartość parametru B (~ 0.30 – 0.33 w 21°C) oraz temperatura topnienia ($\sim 25^\circ\text{C}$) były podobne. W wyniku zmiany rozpuszczalnika z wody

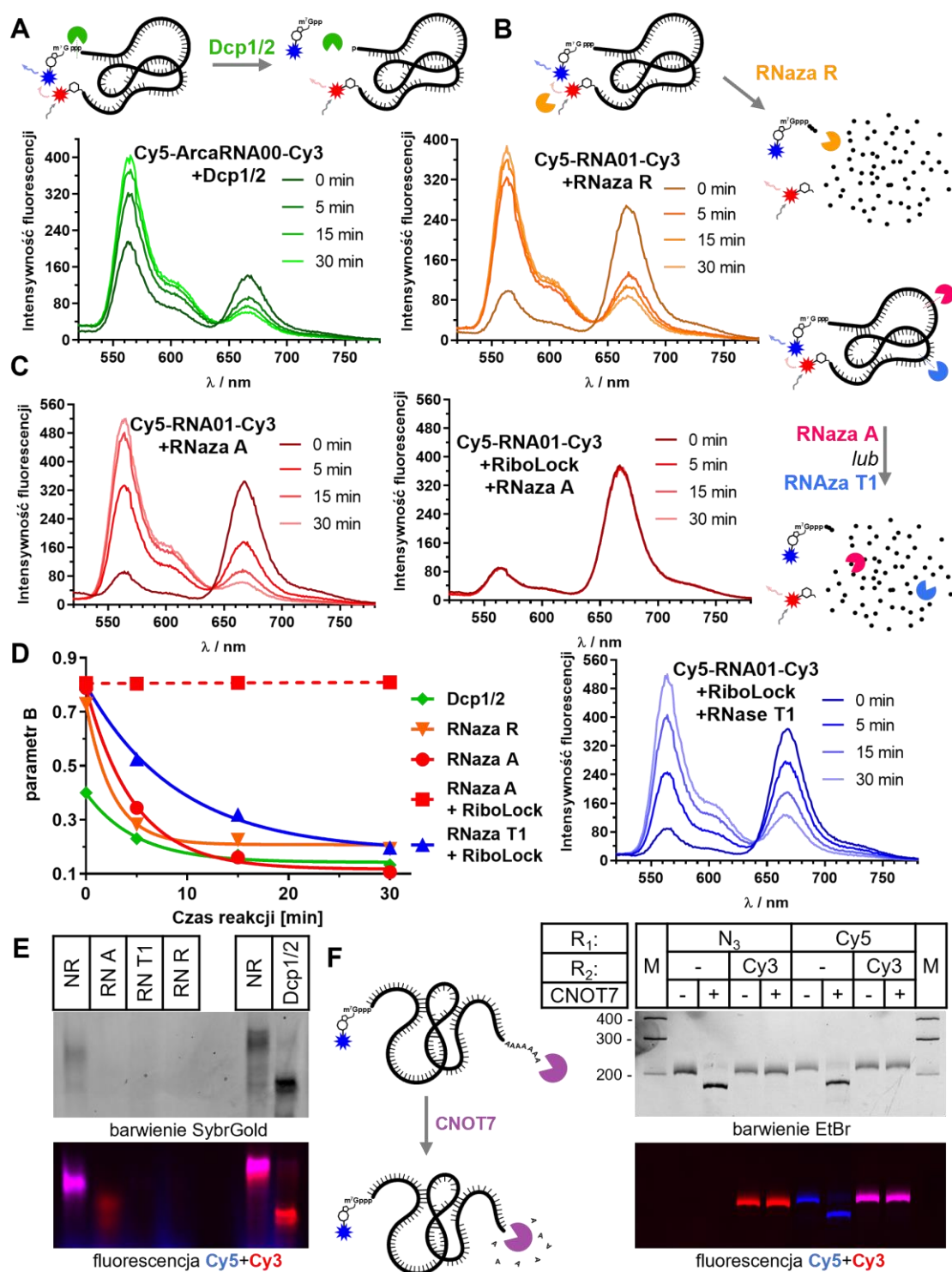


Rysunek 3.1.4: Przykładowe pomiary temperatury topnienia RNA z użyciem sond FRET. Widma fluorescencyjne sond Cy5-RNA02-Cy3 (A) oraz Cy5-ArcaRNA02-Cy3 (B) w zależności od temperatury, w buforowanych roztworach wodnych. C-E) Widma fluorescencyjne sondy Cy5-RNA02-Cy3 w zależności od temperatury i składu rozpuszczalnika. F) Wykres zmian parametru B w zależności od temperatury (krzywe topnienia) dla roztworów o różnym składzie.

na 50% glicerol wartość B drastycznie spadła (~ 0.13 , Rys. 3.1.4C). Z kolei w solance (4 M NaCl) zaobserwowano podwyższony parametr B (~ 0.6 w zakresie -8 – 50 °C) oraz temperaturę topnienia (~ 65 °C, Rys. 3.1.4D). W układzie zawierającym wysokie stężenie soli (0.4 M) oraz glicerol (12%) zanotowano stosunkowo wysoki B (~ 0.50 w 20 °C), jednak odmienny przebieg temperaturowy (Rys. 3.1.5E). W przedziale -8 – 50 °C zaobserwowano mniejszy spadek wartości B (od ~ 0.51 do ~ 0.41), a następnie większy (od ~ 0.41 do ~ 0.09), w przedziale 50 – 70 °C. Wyniki można interpretować następująco: wysokie stężenie soli obniża energię struktur RNA poprzez niwelowanie wzajemnego odpychania elektrostatycznych ładunków grup fosfodiesterowych. To powoduje jednocześnie podwyższenie i ujednolicenie temperatury topnienia wszystkich motywów w obrębie struktury sondy. Z drugiej strony, glicerol jest amfifilową cząsteczką zdolną do tworzenia wiązań wodorowych. Dzięki temu zwiększa powłokę hydratacyjną białek, stabilizując ich strukturę.^[207] W przypadku kwasów nukleinowych, hydratacja czy obecność glicerolu mają raczej efekt odwrotny, destabilizujący,^[208] co odzwierciedlają wyniki pomiarów w układach z glicerolem. Tego typu efekty mają znaczenie z perspektywy kontroli konformacji RNA, w szczególności próby kontroli dystansu między końcami 5' i 3'. To istotny problem z perspektywy cyklizacji RNA, co zostało omówione w dalszej części rozprawy (patrz rozdział 4.3 Badania wykonane ze współpracownikami: optymalizacja cyklizacji RNA).

Kolejnym przykładem zastosowania sond FRET są badania enzymów nukleolitycznych. Enzymy nukleolityczne są istotnymi narzędziami w biologii molekularnej oraz atrakcyjnymi celami terapeutycznymi. Dzielą się na egzonukleazy i endonukleazy. Pod wpływem hydrolizy sondy zmienia się jej widmo fluorescencyjne, co pozwala badać przebieg reakcji w czasie. Wybrano enzymy Dcp1/2, RNaza R, CNOT7 (egzonukleazy) oraz RNaza A, RNaza T1, RNaza H (endonukleazy). Prowadzono pomiary fluorescencji oraz analizę elektroforetyczną produktów reakcji (Rys. 3.1.5 i 3.1.6).

Dcp1/2 to egzonukleaza, która katalizuje hydrolizę trifosforanu w obrębie struktury kapu, wyniku czego zostaje uwolniony difosforan 7-metyloguanozyny (m^7GDP).^[209] W wyniku hydrolizy sondy Cy5-ArcRNA00-Cy3 uwolniony zostaje fluorescencyjny analog nukleotydu (Cy5- m^7GDP). To objawia się spadkiem wartości parametru B w czasie oraz powstaniem produktu pozbawionego fluorescencji Cy5 i o zwiększonej mobilności elektroforetycznej (Rys. 3.1.5A, D, E). Kolejnym enzymem była RNaza R. To procesywna egzonukleaza hydrolizująca ustrukturyzowane RNA w kierunku od 3' do 5' końca, w wyniku czego powstają 5' monofosforany nukleozydów.^[210] Jej aktywność prowadziła



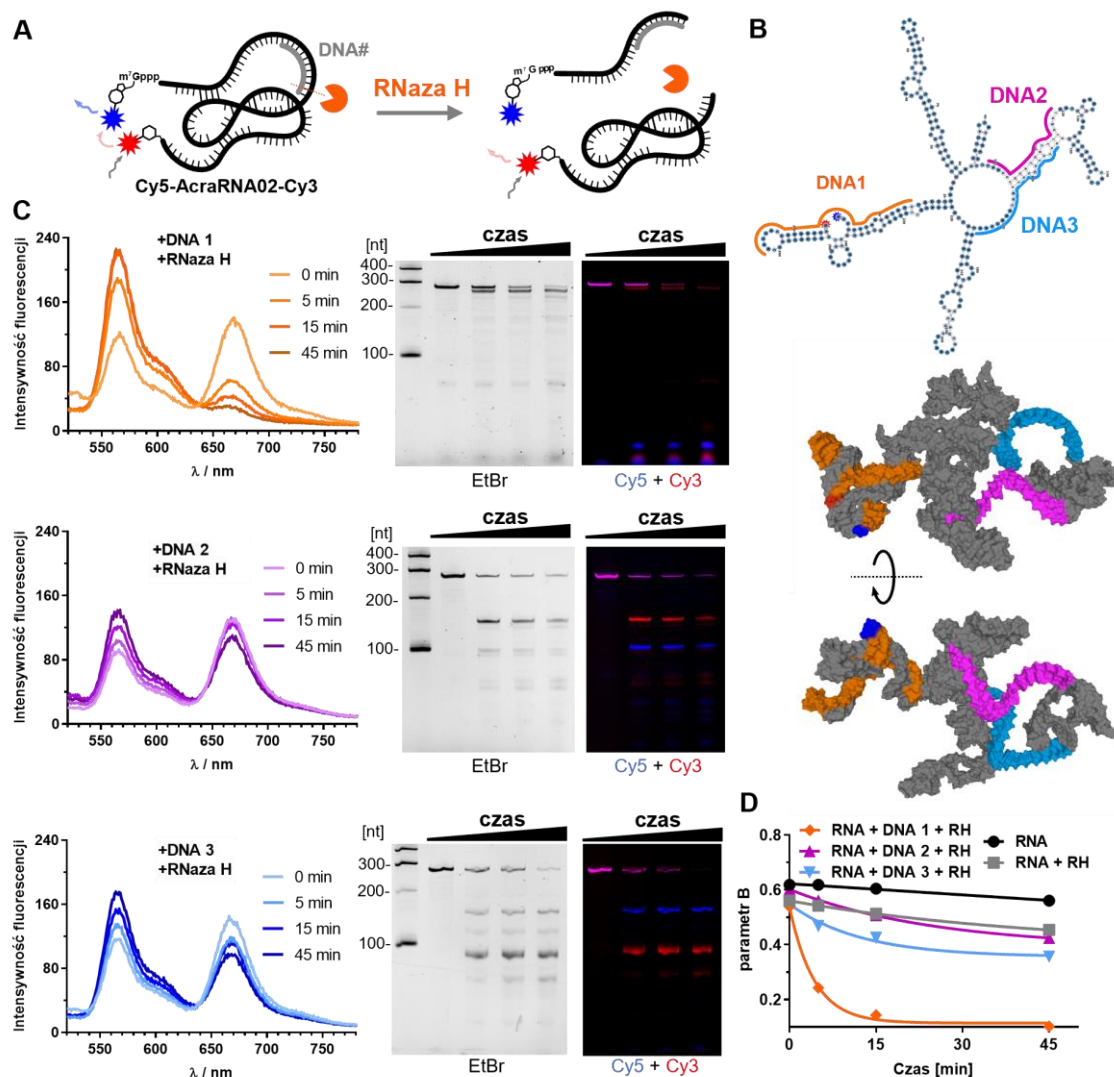
Rysunek 3.1.5: Zastosowanie sond FRET do badania enzymów nukleolitycznych. A) Schemat działania i zmiany widma fluorescencyjnego sondy Cy5-ArcaRNA00-Cy3 pod wpływem działania kompleksu dekapującego Dcp1/2. B) Schemat działania i zmiany widma fluorescencyjnego sondy Cy5-ArcaRNA01-Cy3 pod wpływem działania RNazy R. C) Schemat działania i zmiany widma fluorescencyjnego sondy Cy5-ArcaRNA01-Cy3 pod wpływem działania RNazy A lub RNazy T1, w obecności lub pod nieobecność selektywnego inhibitora RNazy A (RiboLock). D) Zmiany parametru B w czasie reakcji nukleolitycznych. E) Analiza elektroforetyczna (PAGE) substratów (NR) i produktów hydrolizy katalizowanej przez RNazę A (RN A), RNazę T1 (RN T1), RNazę R (RN R)

oraz Dcp1/2 (Dcp1/2). F) Schemat reakcji deadenylacji katalizowanej przez CNOT7 oraz analiza elektroforetyczna produktów deadenylacji sekwencji RNA04. Oznaczenia: R₁ – grupa końca 5'; R₂ – grupa końca 3'; CNOT7 – obecność lub nieobecność deadenylazy. Deadenylacja nie zachodzi w wypadku RNA modyfikowanego na końcu 3'.

również do hydrolizy sondy Cy5-RNA01-Cy3 i całkowitego zaniku fluorescencji akceptora FRET (Rys. 3.1.5B D, E). Ostatnią egzonukleazą był CNOT7, który katalizuje deadenylację od końca 3'. W celu badania enzymu konieczne jest zastosowanie sondy zawierającej sekwencję 3' poliA. Niestety, obecność sekwencji 3' poliA powoduje zanik oddziaływania FRET, przez co nie udało się zarejestrować zmian widma fluorescencji w czasie trwania deadenylacji. Ponadto, analiza elektroforetyczna produktów deadenylacji wskazuje, że w dobranych warunkach CNOT7 nie rozpoznaje RNA z modyfikacją na 3' jako substratu (Rys. 3.1.5F). To z jednej strony niekorzystne zjawisko, ponieważ ogranicza ono zastosowanie sond RNA modyfikowanych na 3' końcu. Z drugiej strony, może być ono korzystne z perspektywy projektowania inhibitorów procesu deadenylacji czy zwiększenia trwałości terapeutycznego RNA.^[87, 92]

Kolejnymi modelami były endonukleazy RNaza A i RNaza T1. Są to enzymy popularnie stosowane w biologii molekularnej, hydrolizujące wiązania fosfodiesterowe za nukleozydami pirymidynowymi (RNaza A) albo guanozyną (RNaza T1).^[211] W wyniku ich działania, sonda ulega fragmentacji, co można zaobserwować w trakcie analizy spektrofлуorymetrycznej oraz elektroforetycznej (Rys. 3.1.5C, D, E). Ponadto pokazano, iż w obecności jej selektywnego inhibitora (RiboLock), cięcie RNazy A zostaje zahamowane. Ostatnim przykładem jest RNaza H, endonukleaza tnąca RNA w heteroduplexie RNA-DNA. Cięcie zachodzi od strony 3' RNA duplexu i jest zależne od wielu czynników.^[212] Przeprowadzono pomiary dla sondy Cy5-ArcRNA02-Cy3 i trzech sekwencji DNA komplementarnych względem różnych jej obszarów (DNA1, DNA2 i DNA3, Rys. 3.1.6). Zaobserwowano, iż w trakcie reakcji z dwiema z trzech wybranych sekwencji DNA widmo fluorescencyjne nie zmienia się w czasie. Z drugiej strony reakcje hydrolizy zachodziły w obecności każdej z sekwencji DNA, o czym świadczy analiza elektroforetyczna. Wyniki zinterpretowano w sposób następujący: w trakcie wszystkich reakcji RNA ulega hydrolizie, jednak w przy zastosowaniu DNA2 i DNA3, zachowane zostają elementy struktury przestrzennej sondy odpowiedzialne za utrzymanie jej końców 5' i 3' w bliskości. Obu wypadkach, było to możliwe ze względu na to, że obszary, w których sonda ulegała hydrolizie, były oddalone od jej końców 5' i 3'. Sekwencja DNA1 jest komplementarna z sondą w obszarach zawierających końce 5' i 3'. W tym przypadku, w wyniku reakcji

dochodzi do eliminacji oligonukleotydu, zawierającego fluorescencyjną grupę końca 5', który oddysocjowuje w łagodnych warunkach (tj. bez denaturacji termicznej).



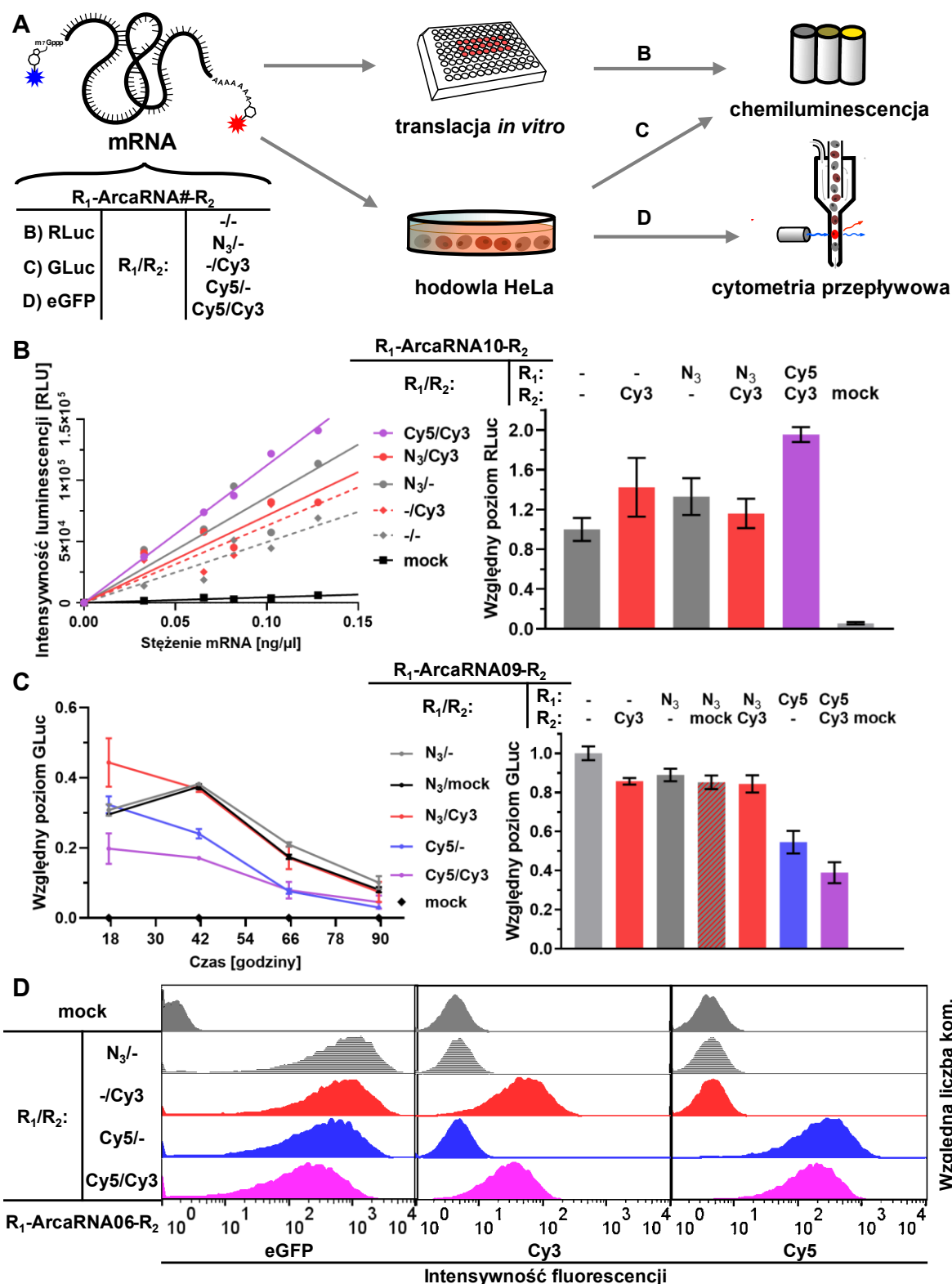
Rysunek 3.1.6: Zastosowanie sond FRET do badania aktywności RNazy H. A) Schemat działania sondy i aktywności RNazy H. B) Modele struktury sekwencji RNA02 z zaznaczonymi obszarami, względem których oligonukleotydy DNA (DNA1–3) były komplementarne. C) Widma fluorescencyjne oraz analiza elektroforetyczna produktów aktywności RNazy H. D) Zmiana parametru B w trakcie aktywności RNazy H w zależności od sekwencji oligonukleotydu DNA.

3. 2. Badania wykonane ze współpracownikami: Bioobrazowanie mRNA

Sondy molekularne zostały również zastosowane do badań w złożonych układach biologicznych. Ze względu związek z terapiami genowymi, zastosowano sondy do bioobrazowania i ekspresji mRNA. Warto zaznaczyć, że na tym etapie celem eksperymentów było zaprezentowanie możliwości sond (ang. *proof of concept*), a kompleksowe badania procesu dostarczania i ekspresji mRNA wykraczały poza ramy projektu badawczego.

W pierwszej kolejności, przeprowadzono pomiary ekspresji reporterowego genu w zależności od modyfikacji mRNA w lizatach komórkowych (wykonane przez Adama Mamota) lub w komórkach hodowanych *in vitro* (wykonane przez dr. Pawła J. Sikorskiego) (Fig. 3.2.1).^[213] Przeprowadzono syntezę modyfikowanych mRNA, kodujących wybrane białka reporterowe. Modyfikacje mRNA obejmowały obecność jednego lub dwóch barwników fluorescencyjnych w ramach końców 5' lub 3', jak i niefluorescencyjne struktury końca 5'. Następnie badano poziom reporterowego genu dla każdej z modyfikacji mRNA, dobranej pod kątem techniki i układu eksperymentalnego.

W pierwszej kolejności zbadano poziomy lucyferazy *Renilla* (RLuc) w lizatach z retikulocytów króliczych (RRL). To najprostszy z wykorzystanych, biologicznych układów, którego zalety i ograniczenia został omówione w dalszej części tekstu (patrz rozdział 5. Podsumowanie i wnioski). Poziom powstałej lucyferazy mierzono za pomocą chemiluminescencji, wynikającej z przebiegu reakcji katalizowanej przez enzym (utleniania koelenterazyny).^[214] Poziom reporterowego białka zależy zarówno od stężenia jak i rodzaju modyfikacji mRNA (Fig. 3.2.1B). Uzyskane tą metodą wyniki wskazują, że poziom białka jest statystycznie jednakowy dla mRNA modyfikowanych i niemodyfikowanych w obszarze końca 3'. Natomiast, znakowanie końca 5' zwiększało poziom białka w niewielkim stopniu.



Rysunek 3.2.1: Pomiary ekspresji reporterowego genu w zależności od modyfikacji mRNA. A) Schemat eksperymentów. B) Translacja w układzie RRL: Zależność intensywności luminescencji od stężenia i modyfikacji mRNA oraz względne poziomy RLuc w zależności od modyfikacji mRNA. C) Ekspresja GLuc w linii komórkowej HeLa: względne poziomy białka w czasie, w zależności od modyfikacji mRNA. D) Cytometria przepływowa po transfekcji komórek HeLa za pomocą modyfikowanego mRNA

kodującego eGFP. Słupki błędów odnoszą się do błędu średniej (SEM) z trzech powtórzeń eksperymentalnych. Eksperymenty z paneli C i D zostały wykonane przez dr. Pawła J. Sikorskiego.^[213]

Poziom reporterowych białek mierzono również w komórkach, do których mRNA dostarczano na drodze lipofekcji (wykonane przez dr. Pawła J. Sikorskiego.)^[213] Wykorzystano linię komórek nowotworowych szyjki macicy (HeLa) do przeprowadzenia dwóch typów eksperymentów. Pierwszy z nich wykorzystywał ekspresję mutantu lucyferazy *Gaussia* (GLuc), sekrecyjnego białka reporterowego, którego poziom można oznaczać za pomocą chemiluminescencji w próbkach medium hodowlanego (Fig. 3.2.1C). Drugi opierał się na cytometrii przepływowej i ekspresji mutantu zielonego białka fluorescencyjnego (eGFP, Fig. 3.2.1D). Wyniki eksperymentów z GLuc wskazują, iż wprowadzenie barwnika na końcu 3' mRNA nie zmienia poziomu białka. Podobnie, jak dla mRNA (N₃-ArcaRNA06-mock), którego przebieg znakowania pomijał dodatek nadjodanu na etapie utleniania (mRNA nadal było poddane następczemu aminowaniu redukcyjnemu w obecności Cy3-EDA). Natomiast, znakowanie fluorescencyjne końca 5' obniżało poziom białka o połowę, co może mieć związek z obecnością podstawników i potencjalnie zwiększoną zawadą steryczną w obszarze końca 5' i struktury kapu.

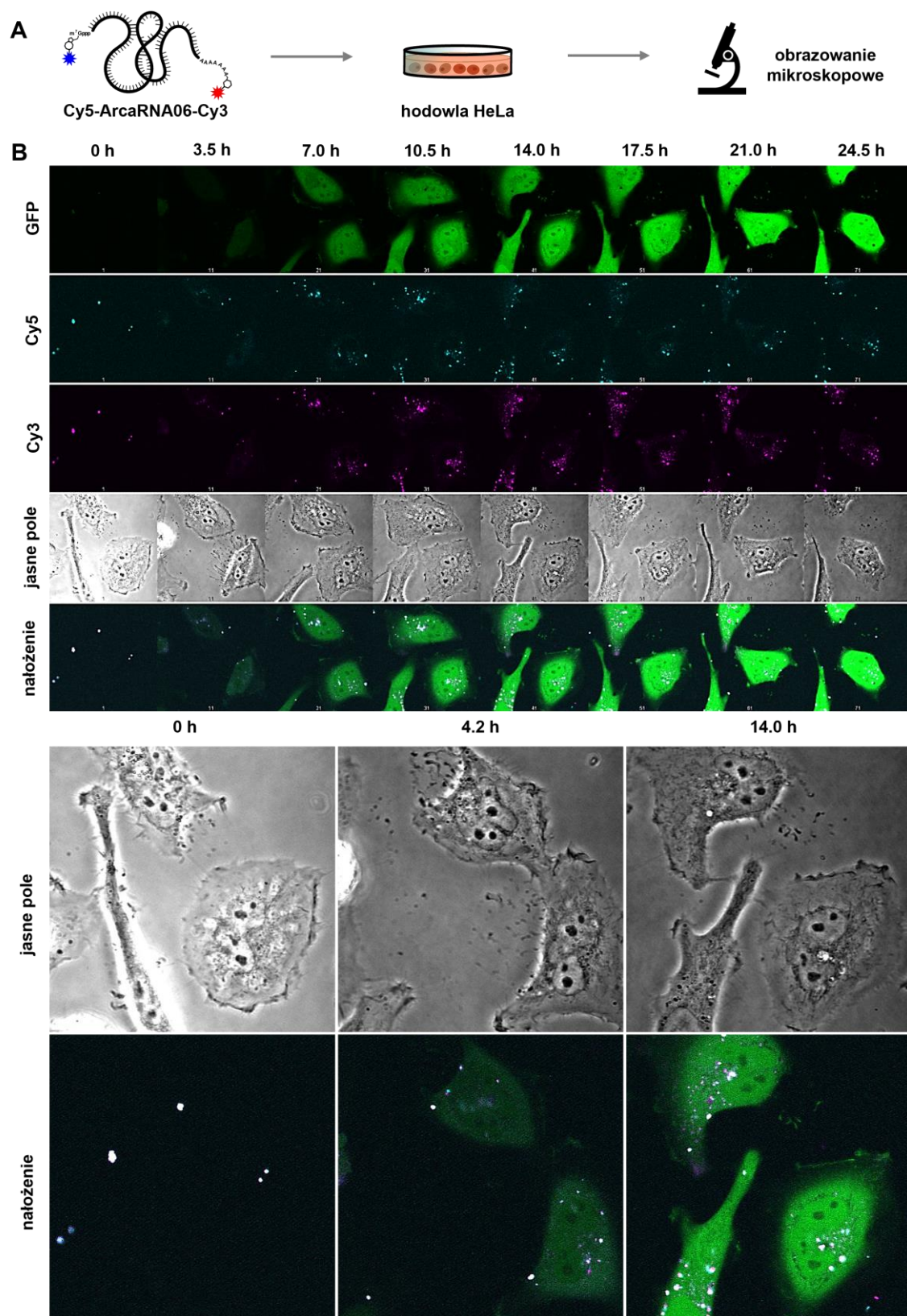
Cytometryczne pomiary poziomu eGFP są zgodne z wynikami uzyskanymi z zastosowaniem GLuc. Ponadto, pokazują rozkład liczby komórek w zależności od intensywności fluorescencji eGFP oraz znakowanego mRNA (barwników Cy3 lub Cy5). To pozwala obejść problem wynikający z nierównomiernej dystrybucji mRNA w komórkach. W wypadku uzyskanych wyników, rozkłady fluorescencji (obwiednia cytogramu) eGFP i znakowanego mRNA są podobne, co jest przejawem korelacji poziomów białka i mRNA. Udokumentowanie takiej korelacji w systemie opartym na aktywności lucyferazy wymagałoby zastosowania dwóch reporterów, co niesie techniczne konsekwencje. Z drugiej strony, zaletą systemów opartych na lucyferazach jest amplifikacja sygnału chemiluminescencyjnego, co pozwala na znacznie niższe (w tym wypadku trzydziestokrotnie niższe) zużycie mRNA. Dodatkowo, system reportera sekrecyjnego (jak GLuc) pozwala na mało inwazyjne (bez lizy czy sortowania) pomiary białka w czasie. Podsumowując, wyniki wyżej opisanych eksperymentów są pomyślne, ponieważ wskazują, iż znakowanie mRNA opracowanymi metodami nie zaburza (w sposób znaczący) procesu jego translacji. To stanowiło przesłankę by zastosować fluorescencyjne mRNA do prób bioobrazowania.

W pierwszej kolejności, do komórek wprowadzono fluorescencyjne mRNA kodujące eGFP na drodze lipofekcji, by następnie przeprowadzić obrazowanie za pomocą konfokalnej mikroskopii fluorescencyjnej (wykonane przez dr. Pawła J. Sikorskiego, Rys. 3.2.2).^[213] Obrazy były rejestrowane równolegle w kilku kanałach (fluorescencja eGFP, Cy3 oraz Cy5) w stałych odstępach czasowych (co ~21 min). Zaobserwowano występowanie mRNA w skupiskach. Jest to znane zjawisko związane z metodą dostarczania mRNA (lipofekcją) i powstawaniem endosomów.^[116] Po około trzech godzinach obserwacji ujawnia się wewnątrzkomórkowy sygnał eGFP. Bezpośrednio po rozpoczęciu transfekcji dwuznakowanym mRNA (Cy5-ArcaRNA06-Cy3), sygnały Cy3 i Cy5 w głównej mierze nakładły się. Z czasem przebiegu eksperymentu, zarejestrowano więcej izolowanych sygnałów Cy3 i Cy5. To zjawisko można interpretować jako degradację mRNA.

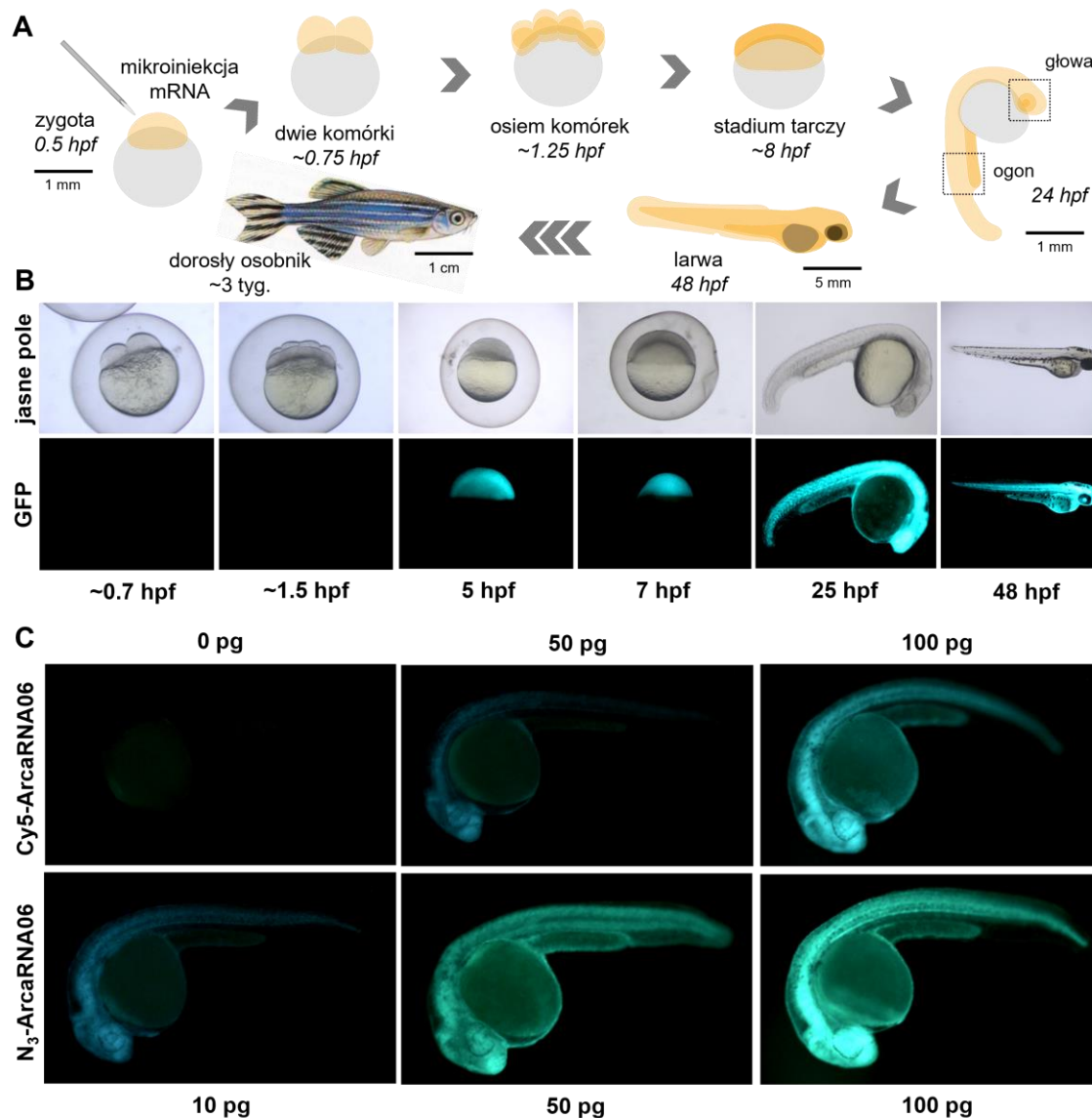
Jako model do obrazowania *in vivo* obrano danio pręgowane (*Danio rerio*, eksperymenty wykonane przez dr. Aleksandrę Siekierską).^[213] Jest to jeden z modelowych strunowców, który już od lat osiemdziesiątych dwudziestego wieku służy w badaniach genetycznych i biomedycznych.^[215] W trakcie rozwoju larwalnego komórki i tkanki zarodka ryby są przeźroczyste, co umożliwia zastosowanie mikroskopii do ich obserwacji. Wykonano mikroiniekcję roztworu mRNA (50–300 pg w 1 nl) do komórki zygotycznej. W trakcie rozwoju, mRNA ulega równomiernemu podziałowi między komórki potomne i ulega dystrybucji w całym organizmie. Na rysunku 3.2.3 przedstawiono obrazy zarodków w czasie rozwoju i ekspresji eGFP. Analogicznie jak w wypadku badań w HeLa, znakowanie mRNA na końcu 5' przekłada się na obniżony poziom reportera. Analogicznie, sygnał w kanałach od Cy3 i Cy5 był znacznie słabszy, a jego wykrycie wymagało zastosowania mikroskopii konfokalnej (Rys. 3.2.4). W trakcie rozwoju zaobserwowano zmianę rozkładu sygnałów eGFP, Cy3 i Cy5: na późniejszych etapach dochodzi do silniejszej separacji Cy3 i Cy5 od eGFP. Ponadto, przybierają one rozkłady o wzorach zależnych od struktury tkankowej zarodka.

Podsumowując, podwójnie znakowane sondy molekularne RNA i mRNA zostały skutecznie zastosowane w badaniach struktury i właściwości RNA, enzymów nukleolitycznych oraz do bioobrazowania i ekspresji mRNA. Zdolność do regioselektywnego umieszczenia dwóch modyfikacji w obrębie RNA dała możliwość wykorzystania zjawiska FRET oraz lokalizacji fragmentów polimeru w układach sztucznych i biologicznych. Określono również wpływ modyfikacji RNA na jego właściwości biochemiczne, biofizyczne i biologiczne. Wyniki opisanych do tej pory

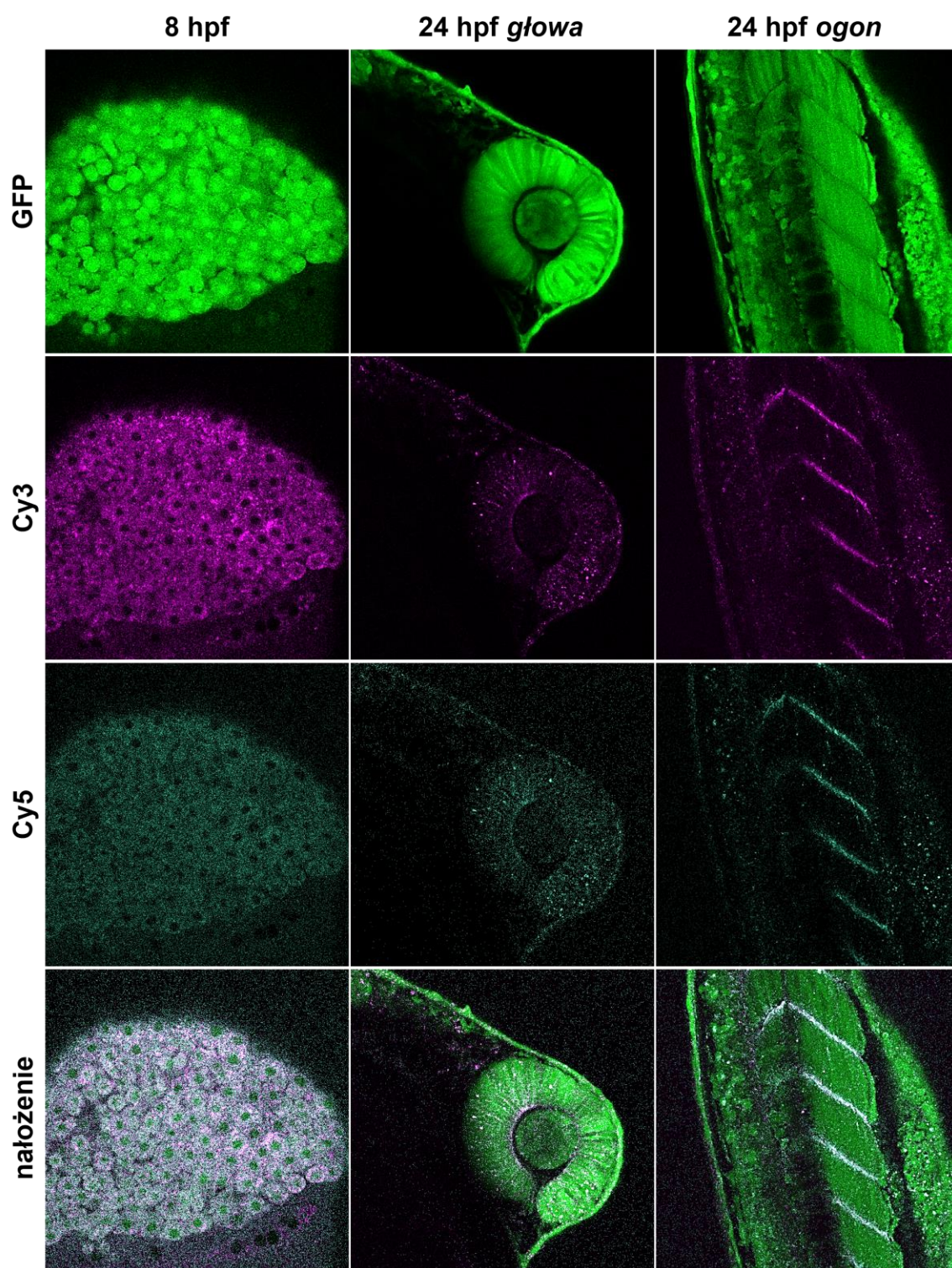
eksperymentów zostały opublikowane w formie artykułu oraz zgłoszenia patentowego.^[213] W następnych rozdziałach opisano badania nad metodą chemoenzymatycznej cyklizacji, swoistego rozszerzenia koncepcji modyfikacji końców 5' i 3' RNA.



Rysunek 3.2.2: Bioobrazowanie podwójnie znakowanego mRNA *in vitro*. A) Schemat eksperymentu. B) Obrazy mikroskopowe po transfekcji komórek HeLa za pomocą Cy5-ArcaRNA06-Cy3. Obrazy kontrolne (po transfekcji za pomocą N₃-ArcaRNA06, Cy5-ArcaRNA06 lub N₃-ArcaRNA06-Cy3) zostały opublikowane w trybie wolnego dostępu. Wszystkie obrazy są dostępne też w formie animacji. Eksperymenty zostały wykonane przez dr. Pawła J. Sikorskiego.^[213]



Rysunek 3.2.3: Mikroiniekcja modyfikowanego mRNA. A) Schemat eksperymentu oraz przebiegu rozwoju embrionalnego *Danio rerio*. B) Obrazy zarodka i jego fluorescencji po mikroiniekcji mRNA na różnych etapach rozwoju. C) Fluorescencja zarodków na etapie 24 hpf w zależności od dawki i modyfikacji mRNA. Eksperymenty zostały wykonane przez dr. Aleksandrę Siekierską.^[213]

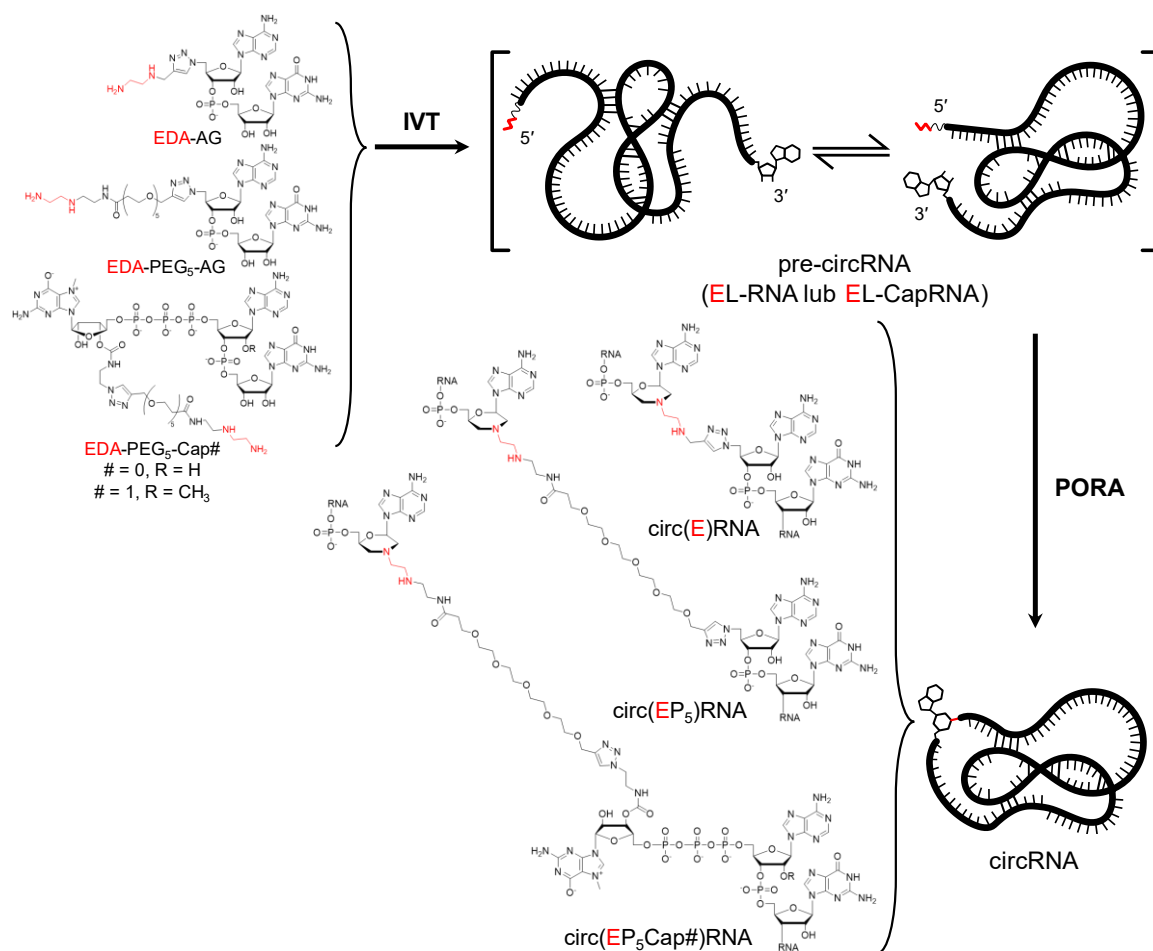


Rysunek 3.2.4: Obrazy z mikroskopu konfokalnego po mikroiniekcji Cy5-ArcRNA06-Cy3 na etapach 8 hpf oraz 24 hpf. Obrazy kontrolne (po mikroiniekcji N₃-ArcRNA06, Cy5-ArcRNA06 lub N₃-ArcRNA06-Cy3) zostały opublikowane w trybie wolnego dostępu. Eksperymenty zostały wykonane przez dr. Aleksandrę Siekierską.^[213]

4. Cyklizacja RNA

4. 1. Cel i metoda otrzymywania analogów circRNA

Chemiczna cyklizacja RNA, rozumiana jako kowalencyjne połączenie końców 5' i 3' polimeru rybonukleotydowego, stanowi istotny element tematyki terapeutycznego RNA. Wynika to z odkryć dotyczących właściwości kolistych cząsteczek RNA (circRNA), zarówno tych naturalnie występujących jak i syntetycznych. Przewaga kolistej formy RNA nad liniową wynika z różnic w strukturze chemicznej: brak elementów rozpoznawanych przez egzonukleazy (końce 5' i 3') jest korzystny pod kontem trwałości RNA, a w konsekwencji trwałości ekspresji terapeutycznej informacji genetycznej.^[52] Ponadto, opublikowano badania świadczące o wysokiej wydajności syntetycznego circRNA jako nośnika informacji genetycznej.^[55-58] Owe syntetyczne circRNA zawierały sekwencje IRES, które pozwalają wyeliminować potrzebę wprowadzania elementów struktury końca 5' i 3' mRNA (t.j. kapu oraz ogona poliA). Ograniczeniem znanych metod cyklizacji wielkocząsteczkowego RNA (ligazy, PIE) jest to, że pozwalają wyłącznie na otrzymanie circRNA o naturalnej strukturze chemicznej. Innymi słowy, ich zastosowanie nie pozwala wprowadzać nienaturalnych grup funkcyjnych, mających potencjalnie pozytywny wpływ na wydajność translacji analogów circRNA lub ich odporność na degradację. Zastosowanie IVT i PORA do chemoenzymatycznej cyklizacji miało pozwolić obejść to ograniczenie. Poprzez IVT koniec 5' liniowego prekursora circRNA miał być wyposażony w grupę EDA, elastyczne polimerowe łączniki (ang. *linkers*) oraz (ewentualnie) strukturę kapu. W celu cyklizacji wykorzystano wariant metody PORA, w trakcie której grupa 5' EDA ulegała wewnątrzcząsteczkowej reakcji aminowania redukcyjnego z utlenionym końcem 3' (Rys. 4.1.1). By zrealizować niniejszy cel, w pierwszej kolejności opracowano metodologię analizy i izolacji circRNA. Następnie przeprowadzono optymalizację cyklizacji RNA z wykorzystaniem sond FRET i metod obliczeniowych. Ostatecznie, otrzymano szereg analogów circRNA kodujących gen reporterowy GLuc oraz zbadano ich właściwości biologiczne.

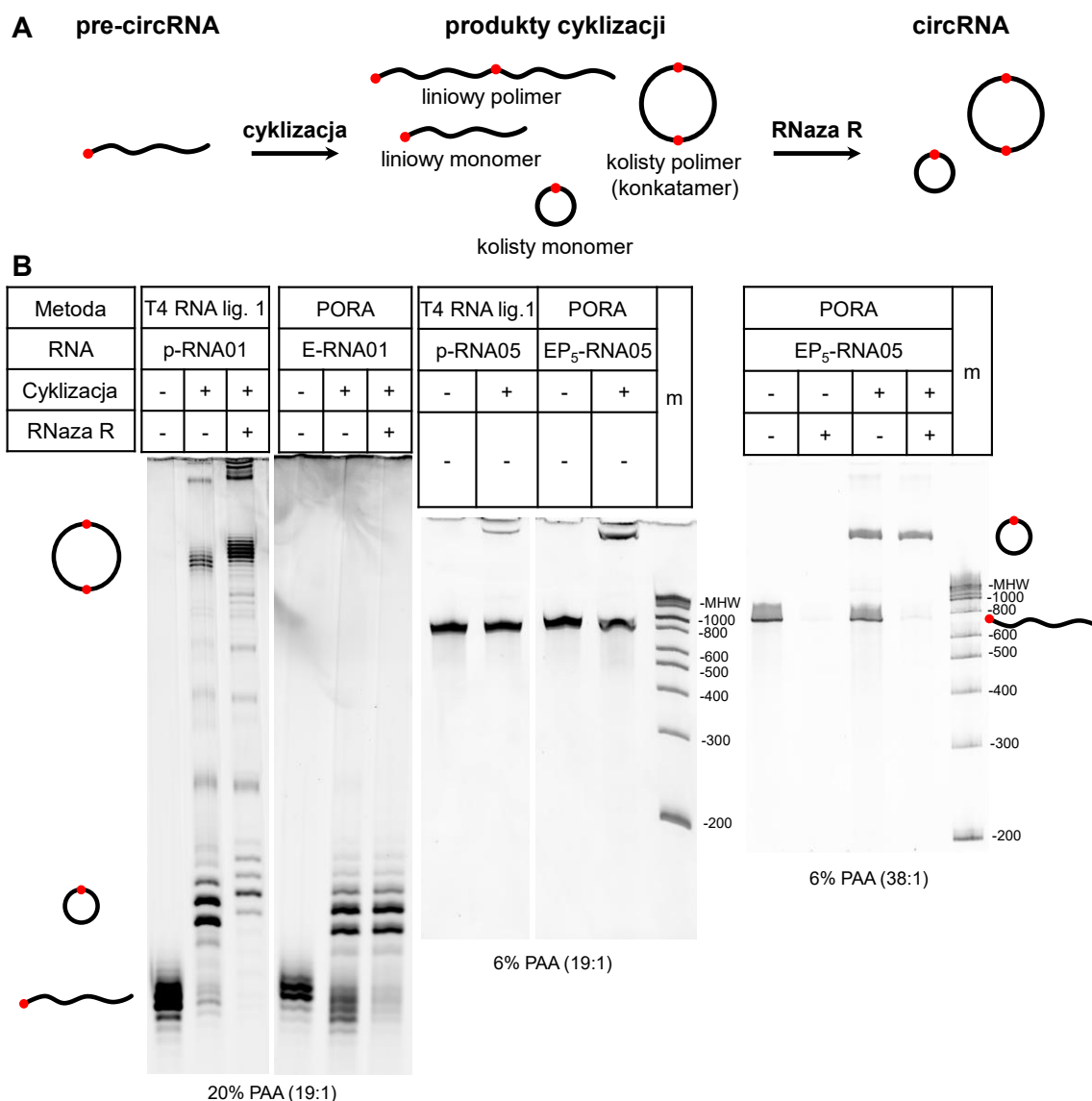


Rysunek 4.1.1: Chemoenzymatyczna cyklizacja RNA. Schemat metoda otrzymywania analogów circRNA. Struktury inicjatorów transkrypcji zawierających grupę 5'-EDA (EDA-AG, EDA-PEG₅-AG, EDA-PEG₅-Cap#) oraz odpowiadających im analogów circRNA (circ(E)RNA, circ(EP₅)RNA, circ(EP₅Cap#)RNA).

4. 2. Metodologia badań analogów circRNA

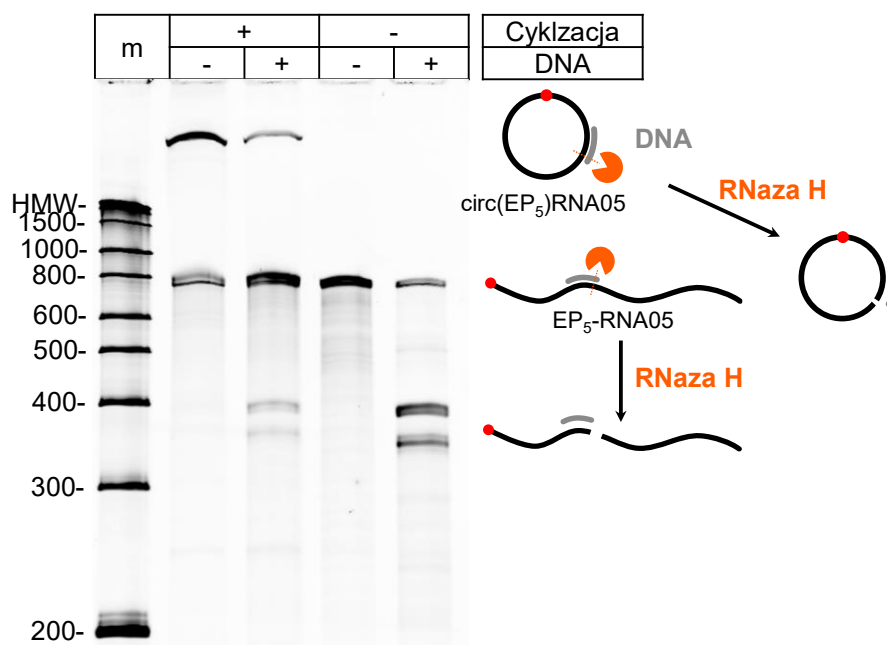
Początkowo, jednym z największych problemów było wykrywanie i mierzenie poziomów kolistego RNA w próbkach produktów reakcji cyklizacji. W związku z tym rozpoczęto od prób cyklizacji enzymatycznej (za pomocą T4 RNA ligazy 1) na jednej z modelowych sekwencji RNA (p-RNA01, 35 nt). Następnie, zaadaptowano jedną z najpowszechniejszych metod analizy mieszanin liniowych i kolistych RNA: poprzez trawienie egzonukleolityczne i elektroforezę (Rys. 4.2.1). Zastosowano RNazę R, która selektywnie usuwa z mieszaniny linowe RNA. Formy koliste są odporne na działanie egzonukleazy ze względu na brak końca 3'. W wyniku trawienia i następczej elektroforezy produktów można rozróżnić obie formy. Następnie, wykonano próby cyklizacji chemoenzymatycznej na drodze IVT i następczej chemicznej cyklizacji metodą PORA. W tym celu wykorzystano prekursorzy zawierające na końcu 5' modyfikację EDA i linker (EL-

RNA) lub modyfikację EDA, linker i analog kapu (EL-Cap#RNA). Wykonano próby cyklizacji chemoenzymatycznej dla sekwencji RNA o długości w przedziale 35-1500 nt. Jako kontrolę pozytywną wykorzystywano produkt enzymatycznej cyklizacji analogicznej sekwencji RNA (p-RNA). Jednym z pierwszych przełomowych wyników była cyklizacja sekwencji RNA05 (740 nt, Rys. 4.2.1B). Jest to przypadek, w którym chemoenzymatyczna cyklizacja wielkocząsteczkowego RNA zachodzi ponad pięciokrotnie wydajniej (33%) niż enzymatyczna (6%).



Rysunek 4.2.1: Trawienie egzonukleolityczne i następcza elektroforeza produktów cyklizacji enzymatycznej i chemoenzymatycznej. A) Schemat eksperymentu z wykorzystaniem RNazy R. B) Przykładowe żele poliakrylamidowe pokazujące produkty cyklizacji RNA o sekwencjach RNA01 (35 nt) oraz RNA05 (740 nt). Oznaczenia: Cyklizacja – informacja o tym czy próbka została poddana cyklizacji; RNaza R – informacja o tym czy próbka została poddana trawieniu RNazą R; m – marker długości RNA; HMW – mieszanina liniowych RNA o długości 1500–6000 nt.

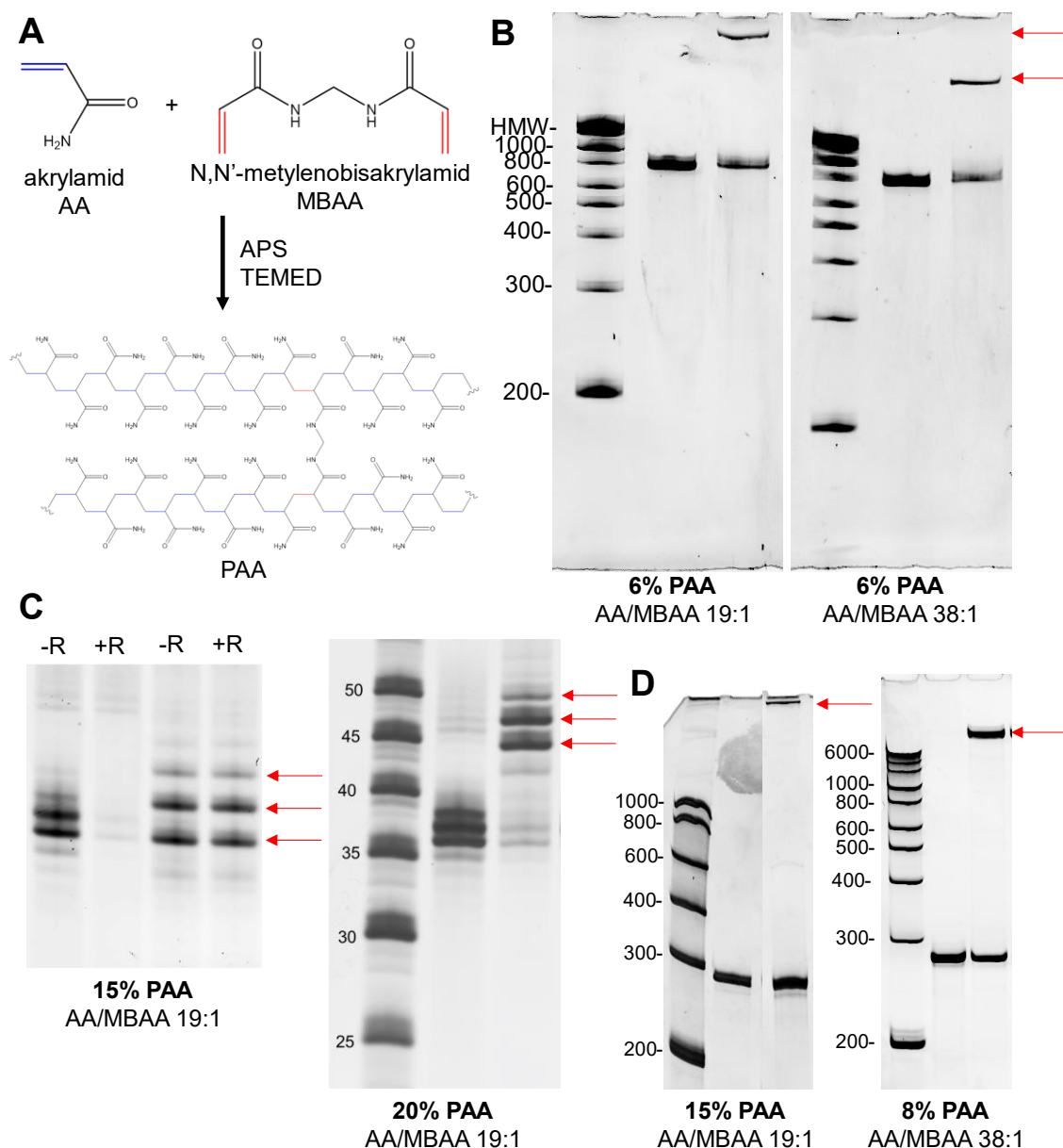
Dodatkowym sposobem weryfikacji jest wewnątrzcząsteczkowe cięcie kolistego RNA, które można przeprowadzić za pomocą RNazy H. W wyniku reakcji circRNA ulega linearyzacji co objawia się odzyskaniem mobilności elektroforetycznej na poziomie pre-circRNA. (Rys. 4.2.2) W ramach kontroli wykonano cięcie prekursora, które prowadziło do dwóch cząstek RNA o mniejszej masie.



Rysunek 4.2.2: Trawienie endonukleolityczne analogu circ(EP₅)RNA05 oraz jego liniowego prekursora (EP₅-RNA05) z wykorzystaniem RNazy H. Analiza elektroforetyczna wykonana za pomocą żelu poliakrylamidowego (6% PAA, 38:1). Oznaczenia: Cyklizacja – informacja o tym czy próbka została poddana cyklizacji; DNA – informacja o tym czy w mieszaninie reakcyjnej obecny był komplementarny oligonukleotyd DNA; m – marker długości RNA; HMW – mieszanina liniowych RNA o długości 2000–6000 nt.

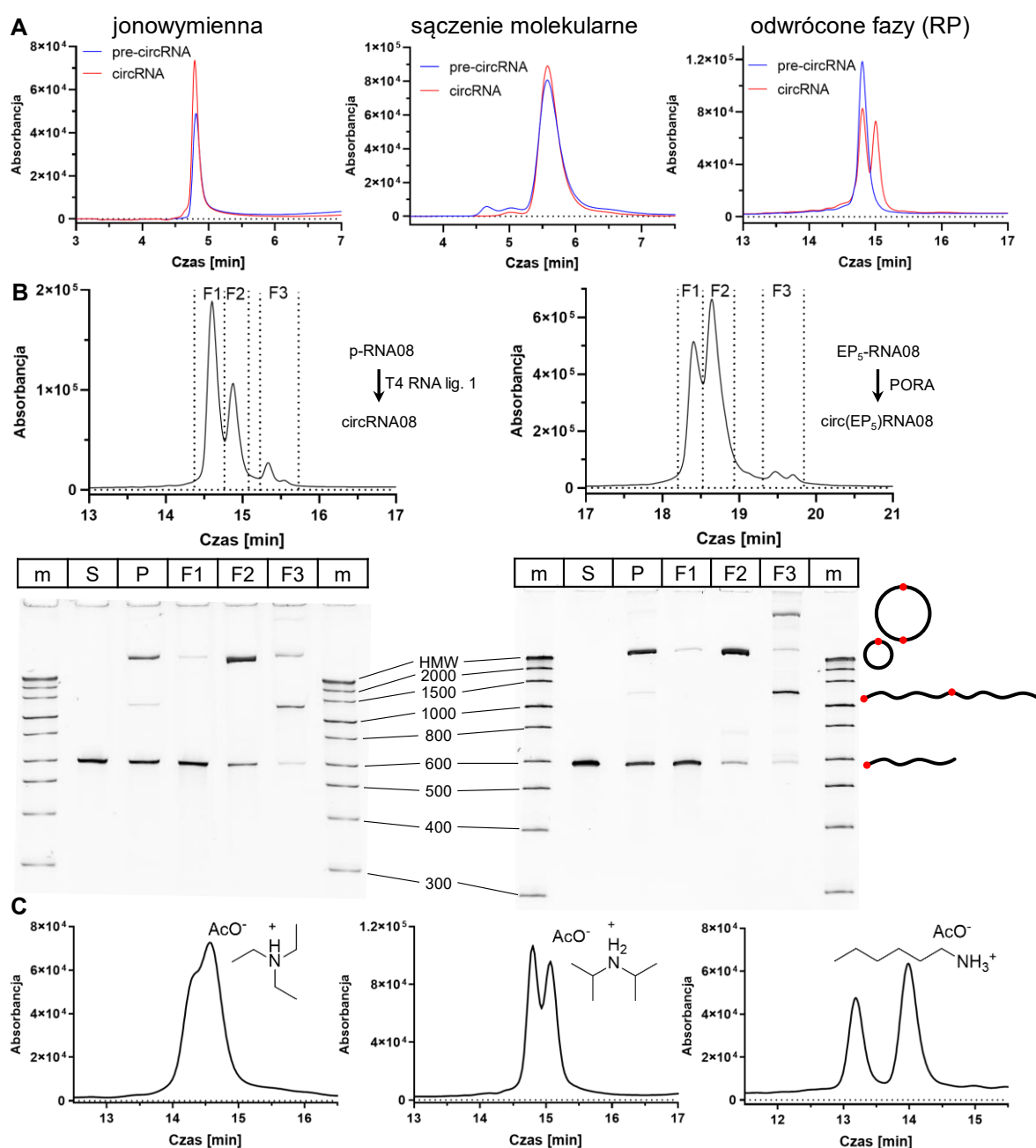
W trakcie badań wykorzystano również elektroforezę w żelu poliakrylamidowym (PAGE) w warunkach denaturujących. Elektroforeza wykorzystuje przyciąganie różnoimiennych ładunków w polu elektrycznym oraz opór ośrodka (żelu). Mobilność elektroforetyczna naładowanych polimerów różnicuje się ze względu na ich promień hydrodynamiczny, na który wpływ mają stopień polimeryzacji i konformacja cząsteczki. Jedną z opublikowanych metod analizy wielkocząsteczkowego circRNA jest elektroforeza w żelach agarozowych. Objętość żelu wypełniają splątane włókna polisacharydu, co tworzy korzystne warunki do rozdziału dużych cząsteczek, natywnych form i wielkocząsteczkowych kompleksów. Metoda wykazuje niską powtarzalność i rozdzielczość pod kontem topologii RNA.^[216] W trakcie eksperymentów zaobserwowano, iż rozdział RNA w poliakrylamidzie (PAA) wykazują wysoką rozdzielczość pod kontem

topologii. (Rys. 4.2.3) Przeprowadzono skuteczny rozdział liniowych i kolistych RNA (30-1400 nt) w żelach PAA o gęstościach w zakresie 6-20%. Różnica w mobilności elektroforetycznej liniowego i kolistego RNA była zależna od przede wszystkim od gęstości i stopnia usieciowania żelu (stosunku akrylamidu do N,N'-metylenobisakrylamidu). Ponadto, skuteczność rozdziałów wzrastała wraz ze wzrostem stopnia polimeryzacji RNA.



Rysunek 4.2.3: Elektroforeza circRNA na żelu poliakrylamidowym. A) Schemat reakcji polimeryzacji akrylamidu (AA) w obecności metylenobisakrylamidu (MBAA). Rozdział pre-circRNA i circRNA w zależności od gęstości i stopnia usieciowania (stosunku AA do MBAA) żelu dla sekwencji RNA05 (B), RNA01 (C) oraz RNA02 (D). Czerwone strzałki wskazują położenie kolistych form RNA. HMW: mieszanina liniowych RNA o długości 1500-6000 nt. Na panelu C wybrane próbki były (+R) bądź nie były (-R) trawione RNazą R.

Przeprowadzono również rozdziały chromatograficzne circRNA. (Rys. 4.2.4) Wykorzystano aparaturę układu HPLC oraz szereg typów chromatografii: jonowymienną, sączenie molekularne oraz RP. Tylko typ RP pozwalał na rozdział pod kontem topologii czy izolację circRNA na drodze frakcjonowania. Rozdzielczość rozdziału uzależniona była od składu fazy ruchomej, a najlepszą osiągnięto w układzie z octanem n-heksyloaminy. (Rys. 4.2.4C)



Rysunek 4.2.4: Chromatografia cieczowa circRNA. A) Chromatogramy HPLC próbek EP₅-RNA05 (pre-circRNA) oraz produktów jego cyklizacji (circRNA) wykonane dla różnych typów chromatografii (jonowymienną, sączenia molekularnego lub RP). B) Frakcjonowanie produktów enzymatycznej (T4 RNA lig. 1) lub chemoenzymatycznej

cyklizacji (PORA) sekwencji RNA08 (578 nt) za pomocą HPLC oraz analiza elektroforetyczna frakcji. Oznaczenia: S – substrat (EP₅-RNA05); P – produkty cyklizacji; F1–F3 – frakcje chromatograficzne; m – marker długości RNA; HMW – mieszanina liniowych RNA o długości 3000, 4000 oraz 6000 nt. C) Rozdział mieszaniny liniowego i kolistego RNA o sekwencji RNA14 (598 nt) w zależności od składu fazy ruchomej (wysycenia układu octanem trietyloaminy, diizopropylaminy lub heksyloaminy).

4. 3. Badania wykonane ze współpracownikami: optymalizacja cyklizacji RNA

Za pomocą wyżej opisanej metodologii przeprowadzono chemoenzymatyczną cyklizację wielkocząsteczkowego RNA (500-1400 nt, Tab. 4 i 5, eksperymenty wykonane przez Adama Mamota lub dr. Małgorzatę Wasińską-Kałwę). W trakcie tego procesu, rozpoznano problem: zależność wydajności cyklizacji od sekwencji RNA. W celu jego rozwiązania podjęto próbę kontroli ich struktury trzeciorzędowej. Z założenia, jest to nietrywialne zadanie. Poniżej przedstawiono studium przykładów, dokumentujących zjawiska, które mogą posłużyć do potencjalnego rozwiązania problemu w przyszłości. Mimo ograniczeń, udało się otrzymać szereg analogów wielkocząsteczkowego circRNA z wysokimi wydajnościami (30–60%), których właściwości biologiczne omówiono w następnym rozdziale.

Tabela 4: Wydajności cyklizacji wielkocząsteczkowego RNA określone na podstawie analizy PAGE. Część reakcji została wykonana lub powtórzona przez dr. Małgorzatę Wasińską-Kałwę.

Nazwa circRNA	Inicjator w IVT	Metoda cyklizacji	oligonukleotyd DNA	Wydajność cyklizacji [%]	
circRNA05	GMP	Lig. I T4	brak	5–10	
circ(EP ₅)RNA05	EDA-PEG ₅ -AG	PORA		30–45	
circRNA08	GMP	Lig. I T4		36	
circ(EP ₅)RNA08	EDA-PEG ₅ -AG	PORA		60	
circ(EP ₅ Cap1)RNA08	EDA-PEG ₅ -Cap1			53	
circ(EP ₅ Cap1)RNA11	EDA-PEG ₅ -Cap1	PORA		46	
circRNA12	GMP	Lig. I T4	brak	13	
		Lig. II T4	brak	31	
			DNA5	1–45	
circ(EP ₅ Cap1)RNA12	EDA-PEG ₅ -Cap1	PORA	brak	10	
			DNA5	50	
circRNA13	GMP	Lig. I T4	brak	9	
		Lig. II T4	brak	32	
			DNA4	61	
circ(EP ₅ Cap1)RNA13	EDA-PEG ₅ -Cap1	PORA	brak	2	
			DNA4	50	
circRNA15	GMP	Lig. I T4	brak	10	
		Lig. II T4	DNA6	46	
circ(EP ₅)RNA15	EDA-PEG ₅ -AG	PORA	brak	44	
circ(EP ₅ Cap1)RNA15	EDA-PEG ₅ -Cap1		DNA6	44	
			DNA6	41	
circRNA14	GMP	Lig. I T4	brak	40	
		Lig. II T4		84	

circ(EP ₃)RNA14	EDA-PEG ₅ -AG	PORA		52
circRNA16	GMP	Lig. I T4		5–10
circ(EP ₃)RNA16	EDA-PEG ₅ -AG	PORA		53

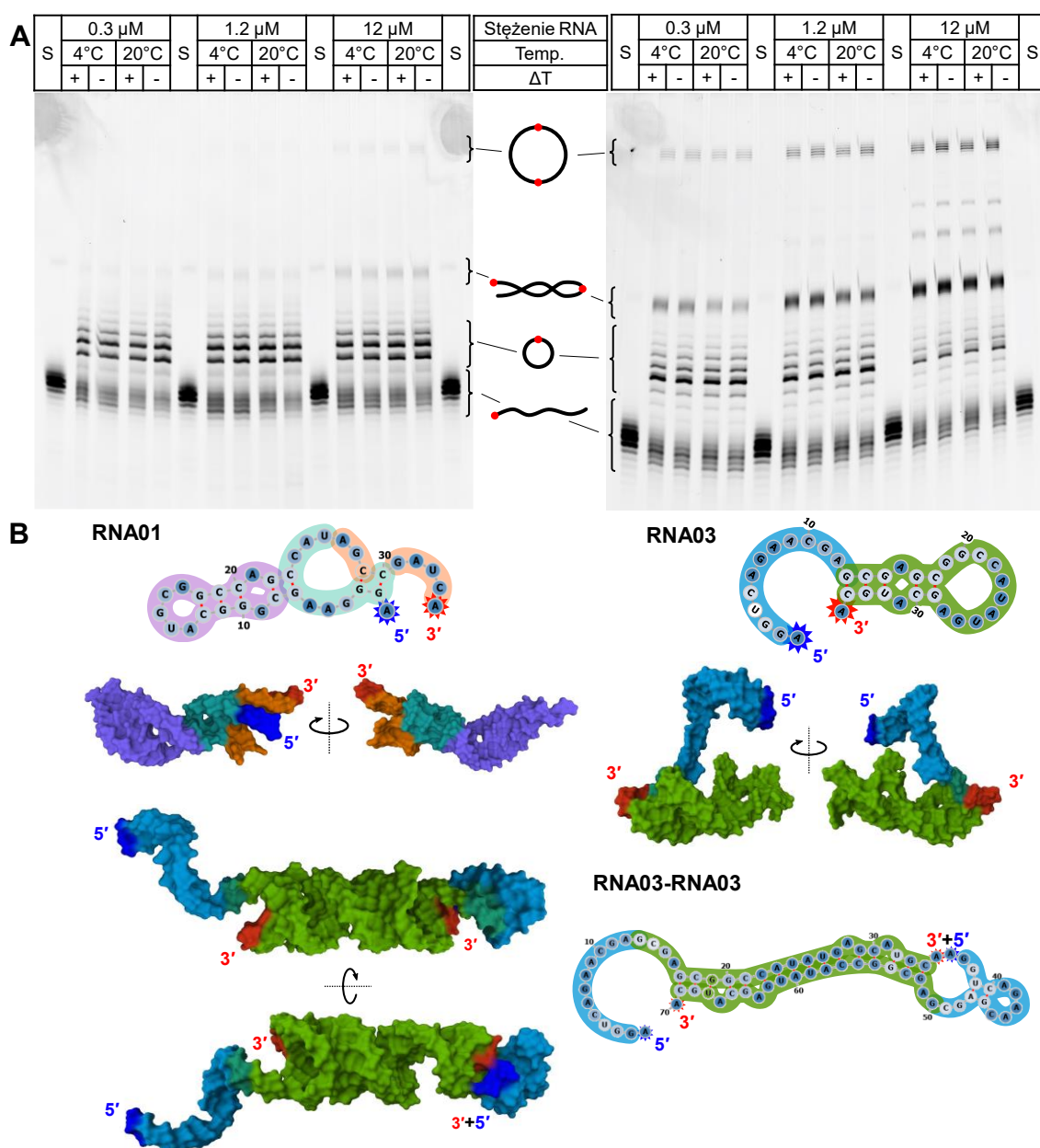
Tabela 5: Lista i opis sekwencji RNA otrzymanych na drodze IVT

Nazwa sekwencji	Elementy sekwencji RNA				Długość [nt]	E ₂₆₀ ^f
	5' UTR ^a	ORF ^b	3' UTR ^d	Ogon 3' ^e		
RNA00	GAAGAAGCGGGCAUGC GGCCAGCCAUAGCCGAUCA				35	0.438
RNA01	AGGGAAGCGGGCAUGC GGCCAGCCAUAGCCGAUCA				35	0.437
RNA02	MCS	RLuc/-	-	-	276	3.31
RNA03	AGGUCAGAACGAGCGAGCGGCCAUUAUGAGCAUGCA				35	0.446
RNA04	Kozak	V5x3	-	A60	237	2.91
RNA05	Kozak	eGFP	-	-	740	8.93
RNA06	Kozak	eGFP	HBBx2	A30	1065	12.8
RNA07	MCS	RLuc	HBBx2	A30	1354	16.4
RNA08	Kozak	Gluc	-	-	578	6.99
RNA09	Kozak	Gluc	HBBx2	-	870	10.4
RNA10	Kozak	Gluc	HBBx2	A30	908	11.0
RNA11	HBB	Gluc	-	-	626	7.58
RNA12	HBB	Gluc	HBBx2	-	918	11.0
RNA13	HBB	Gluc	HBBx2	A30	952	11.5
RNA14	RSV	Gluc	-	-	598	7.25
RNA15	EMCV	Gluc	-	-	1152	13.9
RNA16	OligoIRES	GLuc	-	-	711	9.35
RNA17	Kozak	FLuc	HBBx2	A30	2008	24.4

^aSekwencja poprzedzająca kodon start: MCS (ang. *multiple cloning site*) – sekwencja ułatwiająca klonowanie (84 nt); Kozak – krótka sekwencja promująca translację (8–12 nt); HBB (ang. *human beta-globin*) – sekwencja pochodząca z genu ludzkiej beta-globiny (50 nt); RSV – sekwencja promująca m⁶A-zależną translację (MIREs) z syncytialnego wirusa oddechowego (27 nt)^[67]; EMCV – sekwencja promująca kap-niezależną translację (IRES) z wirusa zapalenia mózgu i mięśnia sercowego świni (583 nt)^[56]; OligoIRES – sztuczna sekwencja promująca kap-niezależną translację (200 nt)^[94]. ^bOtwarta ramka odczytu czyli sekwencja kodująca białko (ang. *open reading frame*): RLuc/- – sekwencja pierwszych 62 aminokwasów lucyferazy *Renilla* (186 nt); V5x3 – potrójna sekwencja epitopowa pochodząca z bakteriofaga V5 (162 nt); eGFP – sekwencja kodująca wzmocnione białko zielone fluorescencji (720 nt); RLuc – sekwencja kodująca lucyferazę *Renilla* (936 nt); GLuc – sekwencja kodująca lucyferazę *Gaussia-Dura* (558 nt); FLuc – sekwencja kodująca lucyferazę robaczka świetojąńskiego (ang. *firefly*, 1653 nt). ^cSekwencja za kodonem stop: HBBx2 – podwojona sekwencja pochodząca z genu ludzkiej beta-globiny (270 nt); - – sekwencja krótkiego łącznika (0–20 nt). ^dSekwencja za sekwencją 3' UTR: A30 – ogon poliadeninowy (30 nt); - – sekwencja krótkiego łącznika (0–20 nt). ^fMolowy współczynnik ekstynkcji (dla długości fali 260 nm, jednostka: $\mu\text{M}^{-1}\text{cm}^{-1}$) oszacowany za pomocą bioinformatycznego narzędzia OligoCalc.^[217]

Przykład 1: chemoenzymatyczna cyklizacja krótkiego RNA (RNA01 oraz RNA03, 35 nt) w zależności od stężenia i temperatury (Rys. 4.3.1). We wszystkich reakcjach zaobserwowano powstawanie produktów ubocznych (polimerów i konkatamerów). Dla

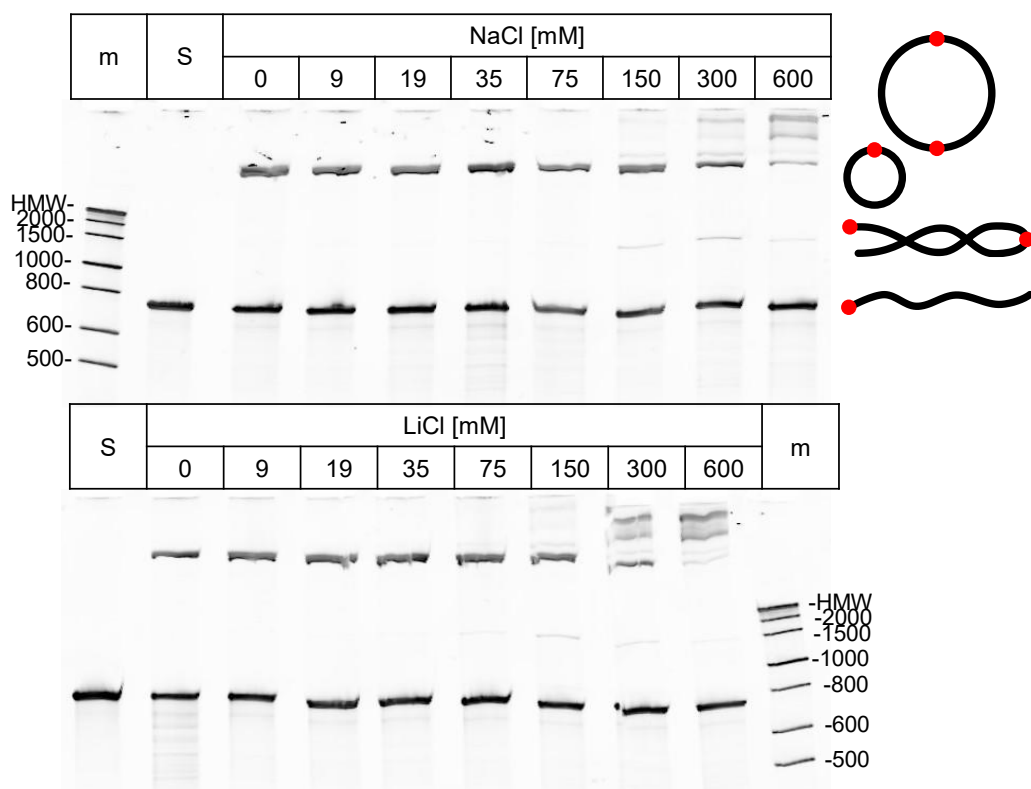
obu sekwencji warunki nie miały większego wpływu na konwersję liniowego prekursora (rzędu 50-70%). Sekwencja RNA03 ulega cyklizacji mniej selektywnie niż RNA01. W celu próby wyjaśnienia tego zjawiska wykorzystano modele struktury RNA01 oraz RNA03. W wypadku RNA01, końce 5' i 3' mogą być stosunkowo blisko siebie, co zostało omówione wcześniej (Rys. 3.1.3). Natomiast, RNA03 może tworzyć spinkę o niskiej stabilności z odseparowanym fragmentem pierwszych dziesięciu nukleotydów od końca 5'. (Rys. 4.3.1B) Model produktu ligacji RNA03 (dimeru, RNA03-RNA03) sugeruje, iż tworzy on stabilniejszą strukturę drugorzędową. To tłumaczyłoby, dlaczego w większych stężeniach RNA03 głównym produktem są dimery.



Rysunek 4.3.1: Chemoenzymatyczna cyklizacja krótkiego RNA (RNA01 oraz RNA03, 35 nt) w zależności od stężenia i temperatury. A) Analiza elektroforetyczna produktów

cyklizacji (RNA01 z lewej, RNA03 z prawej). Oznaczenia: S – substrat; Stężenie RNA – stężenie substratu w mieszaninie reakcyjnej; Temp. – temperatura inkubacji; ΔT – informacja o tym czy substrat był denaturowany termicznie przed etapem utleniania. B) Modele struktur RNA01, RNA03 oraz dimeru RNA03-RNA03.

Przykład 2: wpływ stężenia soli na cyklizację wielkocząsteczkowego RNA (RNA05, 740 nt). W trakcie pomiarów temperatury topnienia RNA za pomocą sond FRET zauważono wpływ składu rozpuszczalnika na bliskość końców 5' i 3' (Rys. 3.1.4). Ze względu na to, przeprowadzono optymalizację cyklizacji pod kątem stężenia chlorku sodu lub litu (Rys. 4.3.2). We wszystkich reakcjach konwersja substratu była podobna (40–60%). Dla każdej z soli, w stężeniu powyżej 150 mM zaobserwowano selektywne powstawanie konkatamerów. Ilość liniowego monomeru w mieszaninie nie zmieniała się znacznie, a udział liniowego dimeru był zerowy lub bardzo niski. W wyniku zwiększenia stężenia kationów metali dochodzi do maskowania ładunku reszt fosfodiesterowych, co indukuje agregację kwasów nukleinowych.^[101] W agregatach, reakcja ligacji i następnej cyklizacji RNA mogłyby zachodzić równie szybko (lub szybciej) co w roztworze, jednocześnie prowadząc do produktów o zwiększonej masie.



Rysunek 4.3.2: Produkty chemoenzymatycznej cyklizacji EP₅-RNA05 w zależności od stężenia chlorku sodu lub litu. Oznaczenia: S – substrat (EP₅-RNA05); m – marker długości RNA; HMW – mieszanina liniowych RNA o długości 3000, 4000 oraz 6000 nt.

Przykład 3: Kontrola struktury wielkocząsteczkowego RNA z wykorzystaniem komplementarnych oligonukleotydów DNA. Sekwencja RNA13 (952 nt) składała się z następujących elementów: sekwencji kodującej białko reporterowe GLuc, sekwencji niekodujących (5' i 3' UTR) ludzkiej beta-globiny (HBB) oraz sekwencji ogona 3' poliadeninowego (Rys. 4.3.3A). Ostatni element struktury (poliadenina) zamiast tworzyć wiązania wodorowe między zasadami i struktury drugorzędowe, tworzy heliakalną strukturę, opartą na oddziaływaniach hydrofobowych i wyeksponowaną w sferze solwacyjnej mRNA.^[5] To powoduje izolację końca 3' i w konsekwencji niską wydajność cyklizacji (rzędu 2%, Rys. 4.3.3C). Z wykorzystaniem modeli struktury RNA13 zaprojektowano sekwencje komplementarnych oligonukleotydów DNA (DNA1–4, Tab. 4). Poprzez komplementarność, oligonukleotyd DNA miał zadziałać jak szyna (ang. *splint*), zbliżając do siebie końce 5' i 3' RNA. Następnie, wykonano pomiary widm fluorescencyjnych z wykorzystaniem sondy FRET (Cy5-RNA13-Cy3) po jej hybrydyzacji z DNA (Rys. 4.3.3B). Na podstawie pomiarów wybrano optymalną sekwencję oligonukleotydu (DNA2). Jego zastosowanie (eksperyment wykonany przez dr. Małgorzatę Wasińską-Kałwę) doprowadziło do zwiększenia wydajności chemicznej cyklizacji ponad dwudziestokrotnie (do 50%, Rys. 4.3.3C) oraz blisko dwukrotnie w wypadku cyklizacji katalizowanej przez T4 RNA ligazę 2 (z 32 do 61%, Tab. 6). W obecności DNA2 (28 nt), cyklizacja zachodziła selektywnie. Dłuższe sekwencje (DNA3, 48 nt i DNA4, 67 nt) obniżały selektywność cyklizacji kosztem tworzenia liniowych dimerów. Krótsza sekwencja (DNA1, 10 nt) nie wpływała na przebieg reakcji (wydajność 2%). W celu wyjaśnienia zaobserwowanych efektów wykorzystano model struktury RNA13 (Rys. 4.3.3A). RNA i optymalny oligonukleotyd DNA tworzą razem strukturę heteroduplesku o długości 28 nukleotydów. Od strony końca 3' RNA i sekwencji poliA, heterodupleks sąsiaduje z odcinkiem ~20 nukleotydów adeninowych, co może być korzystne z perspektywy tworzenia heteroduplesku od strony końca 3' RNA. Z drugiej strony, od końca 5' RNA, oligonukleotyd jest w pełni komplementarny, przez co może tworzyć strukturę stabilniejszą niż wewnątrzcząsteczkowe struktury drugorzędowe RNA, które są oparte na niepełnej komplementarności. W ten sposób, w trakcie hybrydyzacji dochodzi do stworzenia trwałej struktury (dupleksu RNA-DNA o szacowanej temperaturze topnienia ~50°C) kosztem zniszczenia mniej trwałej struktury drugorzędowej (dupleksu RNA-RNA, o szacowanej temperaturze topnienia w przedziale 20–30°C) co byłoby energetycznie korzystne.

DNA6	RNA15	AGGGGGGGAGGGAGAGGGGTACCCCTATCCTTAGTCACCACCGG CCCCCT
------	-------	--

4. 4. Badania wykonane ze współpracownikami: właściwości biologiczne analogów circRNA

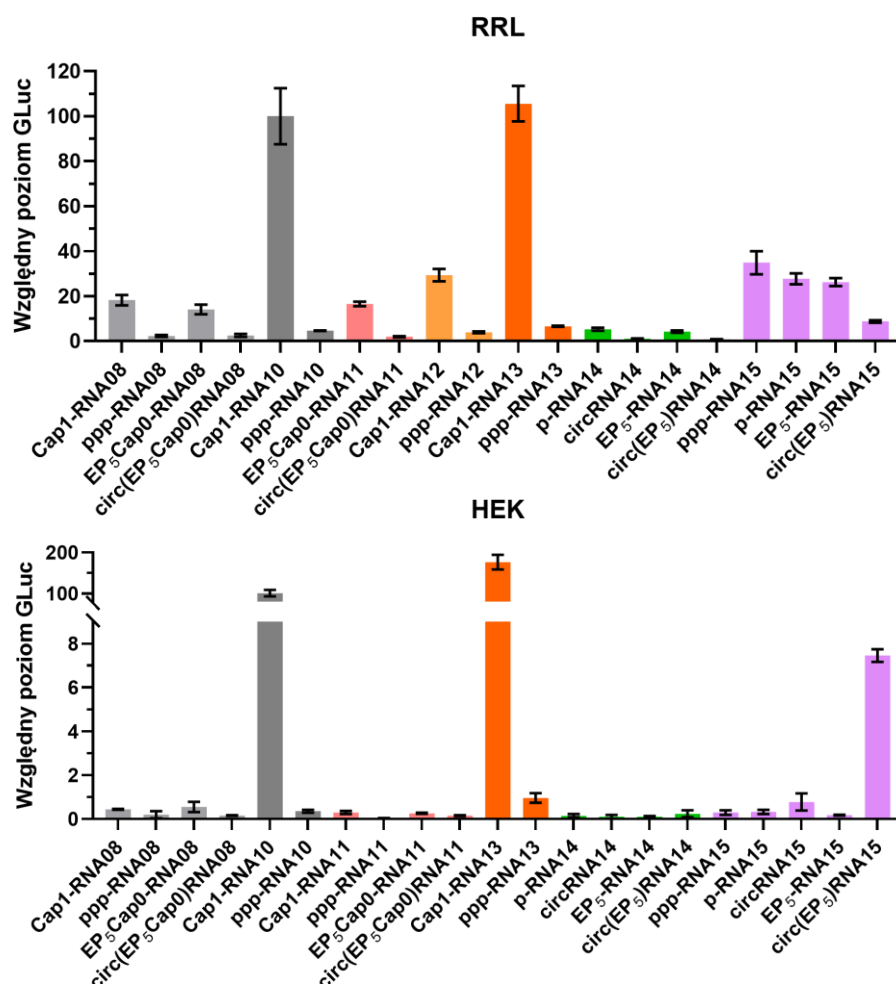
Wykonano badania ekspresji circRNA w lizatach (wykonał Adam Mamot) i liniach komórkowych z wykorzystaniem reportera GLuc (wykonał dr. Karol Czubak) (Rys. 4.4.1–3). Badano ilość produkowanego białka (całkowitą oraz w czasie) w zależności od sekwencji RNA, struktury chemicznej końca 5' (w wypadku liniowych cząsteczek) lub struktury połączenia końców 5' i 3' (w wypadku cząsteczek kolistych). Sekwencje promujące translację, takie jak EMCV IRES, oraz reporter GLuc wybrano ze względu na wcześniej opublikowane badania dotyczące ekspresji circRNA.^[56]

W pierwszej kolejności porównano ekspresję circRNA w układzie RRL oraz w komórkach linii HEK (Rys. 4.4.1). Uzyskane wyniki były rozbieżne. W RRL najwyższe poziomy białka zanotowano dla kanonicznych mRNA (Cap1RNA10 lub Cap1RNA13), a następnie dla liniowych RNA zawierających sekwencję EMCV IRES (RNA15). W RRL cyklizacja obniżała całkowity poziom białka. W komórkach HEK, poza kanonicznymi mRNA, powstawanie białka zaobserwowano wyłącznie dla analogu circ(EP₅)RNA15, zawierającego modyfikowane połączenie końca 5' i 3' oraz sekwencję EMCV IRES. Na tej podstawie zawężono badania do analogów circRNA zawierających sekwencję IRES lub elementy kanonicznego mRNA (strukturę kapu czy ogon poliA). Niestety, analogu circRNA zawierającego zarówno strukturę kapu jak i sekwencję ogona poliA (circ(EP₅Cap1)RNA13) nie udało się wyizolować za pomocą opracowanych metod. Następnie, dr. Karol Czubak wykonał eksperymenty z zastosowaniem linii komórkowych HEK, HeLa, Hep2 oraz A549 (Rys. 4.4.2). Ponownie, najwyższe poziomy białka odnotowano dla kanonicznego mRNA (Cap1-RNA13). Poziomy białka dla circRNA o sekwencji RNA12 lub RNA15 były wyższe jeżeli zawierały one strukturę kapu lub inną modyfikację połączenia końców 5' i 3' (poziom białka dla circ(EP₅Cap1)RNA > circ(EP₅)RNA > circRNA). Ponadto, cyklizacja RNA zawierających sekwencję IRES (RNA15) skutkowała znacznym zwiększeniem poziomu białka. Najwyższe poziomy wśród wszystkich circRNA oraz ich prekursorów odnotowano dla analogu circ(EP₅Cap1)RNA15, który zawierał kombinację struktury kapu oraz sekwencji IRES. Niezależnie od linii komórkowej czy ilości wyprodukowanego białka, zaobserwowano wydłużone czasy ekspresji circRNA i ich analogów. Na podstawie zależności ilości produkowanego białka od czasu oszacowano czasy półtrwania RNA (Rys. 4.4.3). Szacowane czasy półtrwania

circRNA (24-53 h) były wydłużone względem ich liniowych prekursorów (pre-circRNA, 3-9 h) czy kanonicznego mRNA (6-20 h).

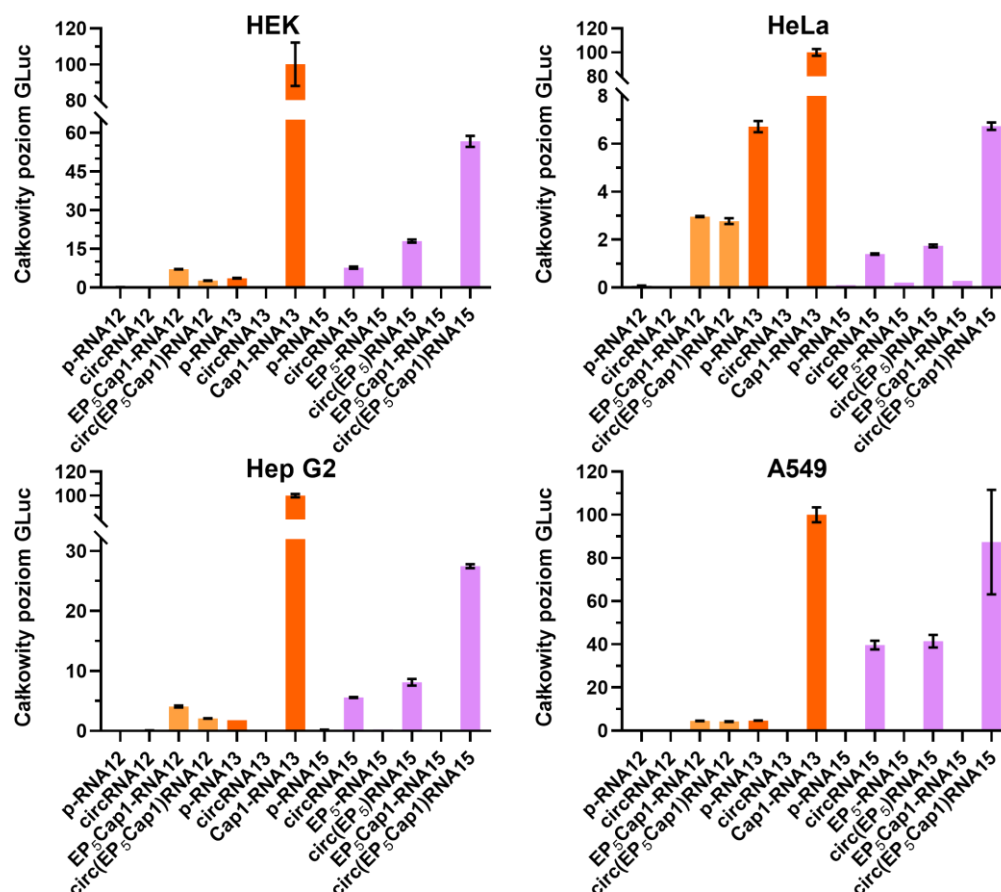
Wzrost poziomu białka w wyniku modyfikacji, stosunkowo wysoki poziom dla kombinacji struktury kapu i sekwencji IRES oraz wydłużone czasy półtrwania circRNA są obiecującymi wynikami, które zachęcają do dalszych poszukiwań atrakcyjnych analogów circRNA, uzyskanych metodą chemoenzymatycznej cyklizacji.

Sekwencja	5' UTR	CDS	3' UTR	Ogon 3'	długość [nt]
RNA08	Kozak	GLuc	-	-	578
RNA09	Kozak	GLuc	HBB	-	870
RNA10	Kozak	GLuc	HBB	A30	908
RNA11	HBB	GLuc	-	-	626
RNA12	HBB	GLuc	HBB	-	918
RNA13	HBB	GLuc	HBB	A30	952
RNA14	RSV	GLuc	-	-	598
RNA15	EMCV	GLuc	-	-	1152



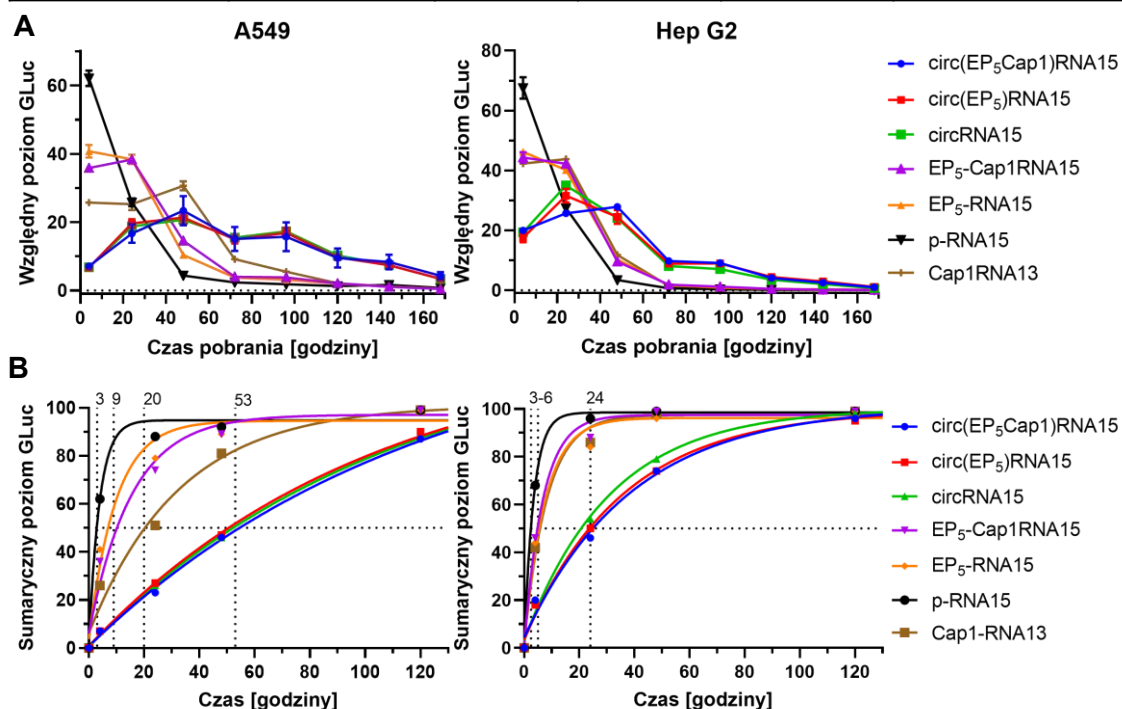
Rysunek 4.4.1: Lista sekwencji oraz wyniki ekspresji analogów circRNA w układzie RRL oraz linii komórkowej HEK. Słupki błędów oznaczają błąd średniej (SEM) z trzech powtórzeń eksperymentalnych.

Sekwencja	5' UTR	CDS	3' UTR	Ogon 3'	długość [nt]
RNA12	HBB	GLuc	HBB	-	918
RNA13	HBB	GLuc	HBB	A30	952
RNA15	EMCV	GLuc	-	-	1152



Rysunek 4.4.2: Lista sekwencji oraz wyniki ekspresji analogów circRNA w liniach komórkowych HEK, HeLa, Hep G2 oraz A549. Słupki błędów oznaczają błąd średniej (SEM) z trzech powtórzeń eksperymentalnych. Eksperymenty wykonał dr. Karol Czubak.

Sekwencja	5' UTR	CDS	3' UTR	Ogon 3'	długość [nt]
RNA13	HBB	GLuc	HBB	A30	952
RNA15	EMCV	GLuc	-	-	1152



Rysunek 4.4.3: Lista sekwencji oraz wyniki ekspresji analogów circRNA w czasie, dla linii komórkowych A549 oraz Hep G2. A) Poziom GLuc w pożywce, w danym punkcie czasowym względem całkowitego poziomu białka, uzyskanego w wyniku transfekcji danym RNA. Słupki błędów oznaczają błąd średniej (SEM) z trzech powtórzeń eksperymentalnych. B) Sumaryczny poziom GLuc uzyskany do danego momentu w czasie, znormalizowany względem całkowitego poziomu białka, uzyskanego w wyniku transfekcji danym RNA. Eksperymenty wykonał dr. Karol Czubał.

5. Podsumowanie i wnioski

Opracowano metody pozwalające na selektywną modyfikację i funkcjonalizację końców 5' oraz 3' RNA:

- Modyfikacja końca 3' została osiągnięta za pomocą metody PORA (ang. *periodate oxidation and reductive amination*) z wykorzystaniem analogów etylenodiaminy (EDA). W tym celu zbadano przebieg aminowania redukcyjnego na dwóch modelach RNA (nukleotydzie GMP oraz trinukleotydzie pU₃) pod kątem: struktury N-nukleofila, stężenia reduktora (NaBH₃CN) i pH.
- Metoda PORA została zoptymalizowana pod kontem jej prostoty, wszechstronności i wydajności. Dzięki temu możliwe jest przeprowadzenie modyfikacji końca 3' RNA na drodze prostego protokołu (w 2.5 godziny, w jednym naczyniu), w szerokim zakresie stężeń (0.2 µM – 1 mM) oraz długości (3–2000 nt) substratu RNA z wydajnościami zankowania w zakresie 75–99%. Struktura modyfikacji, wywodząca się ze struktury zastosowanego analogu EDA, może wykroczać poza zademonstrowane w pracy grupy chemiczne (tj. aminy, azydki, alkiny), hapteny (tj. biotynę) czy barwniki fluorescencyjne (tj. fluoresceiny i cyjaniny).
- Zgodnie z opisanym stanem techniki, zoptymalizowana metoda PORA jest pierwszą metodą chemiczną pozwalającą na bezpośrednią, wydajną i selektywną modyfikację końca 3' wielkocząsteczkowego RNA (powyżej 200 nt, otrzymanego na drodze transkrypcji), w szczególności mRNA.
- Modyfikację końca 5' przeprowadzono za pomocą metody chemoenzymatycznej, będącej kombinacją transkrypcji *in vitro* (IVT) oraz reakcji SPAAC. Metoda została rozwinięta poprzez syntezę analogów typu N₃-AG oraz N₃-m⁷GpppAG, których zastosowanie pozwoliło na wydajne wprowadzenie grupy azydkowej w obrębie końca 5' RNA na drodze IVT.
- Procedury modyfikacji końca 5' oraz 3' połączono, by móc przeprowadzić je równolegle, w jednym naczyniu. W wyniku tego, opracowano metodę podwójnego znakowania RNA, która pozwoliła na regioselektywne wprowadzenie różnych znaczników w obrębie końców 5' oraz 3'.
- Opracowana metoda podwójnego znakowania pozwoliła na otrzymanie szeregu sond molekularnych RNA o rozmiarach w zakresie 30–2000 nt z wydajnościami rzędu 46–88%.

- Przeprowadzono skuteczną izolację homogennych produktów znakowania oraz podwójnego znakowania RNA za pomocą chromatografii RP-HPLC oraz nieudokumentowanego wcześniej efektu „kotwiczenia”.
- Zgodnie z opisanym stanem techniki, opracowana metoda podwójnego znakowania jest bezprecedensowo wydajna i prosta w porównaniu z innymi metodami pozwalającymi na regioselektywne i ortogonalne znakowanie wielkocząsteczkowego RNA, w szczególności mRNA.

Za pomocą opracowanych metod otrzymano szereg sond molekularnych oraz zademonstrowano ich przykładowe zastosowania:

- Otrzymano sondy typu barwnik-RNA-biotyna, które zastosowano w eksperymentach demonstrujących wiązanie sądy przez streptawidynę (EMSA).
- Otrzymano sondy typu odniesienie-RNA-wskaźnik oraz zaprezentowano ich potencjalne zastosowanie jako wskaźników zmian środowiskowych (na przykładzie sondy wrażliwej na zmiany pH).
- Otrzymano sondy FRET (typu akceptor-RNA-donor), które zastosowano do badań struktury przestrzennej oraz temperatury topnienia wybranych sekwencji RNA.
- Sondy FRET zostały również zastosowane do monitorowania w czasie rzeczywistym aktywności hydrolitycznej enzymów nukleolitycznych takich jak: RNaza A, RNaza T1, RNaza R, RNaza H oraz kompleks dekapujący Dcp1/2.

Koncepcję funkcjonalizacji końców 5' i 3' rozszerzono o wariant chemoenzymatycznej cyklizacji RNA:

- Opracowano metodę cyklizacji cząsteczek RNA zawierających grupę 5'-EDA z zastosowaniem zoptymalizowanej metody PORA. W tym celu, przeprowadzono syntezę analogów typu EDA-linker-AG oraz EDA-linker-m⁷GpppAG, które wykorzystano w reakcji IVT.
- W celu izolacji oraz potwierdzenia kolistej topologii otrzymanych analogów circRNA wykorzystano metody chromatograficzne, elektroforetyczne oraz trawienie produktów cyklizacji za pomocą RNazy R lub RNazy H.
- Zgodnie z opisanym stanem techniki, opracowana metoda chemoenzymatycznej cyklizacji jest pierwszą chemoenzymatyczną metodą cyklizacji wielkocząsteczkowego RNA.

- Wraz z dr. Małgorzatą Wasińską-Kalwą zoptymalizowano warunki cyklizacji, przez co otrzymano szereg kodujących cząsteczek circRNA z wydajnościami cyklizacji rzędu 30–60%.

Wraz z dr. Pawłem J. Sikorskim, dr. Aleksandrą Siekierską oraz dr. Karolem Czubakiem przeprowadzono badania znakowego mRNA oraz analogów circRNA w układach biologicznych (*in vitro* i *in vivo*):

- Wykazano, iż modyfikacja końca 3' metodą PORA nie zaniża (w znaczącym stopniu) poziomów białek reporterowych, kodowanych w sekwencji znakowanych mRNA.
- Wykazano, iż podwójnie znakowane mRNA pozostają aktywne translacyjnie (poziom białek reporterowych zmniejszony o połowę), oraz że najadają się jako sondy molekularne do obserwacji mikroskopowych w komórkach oraz *in vivo* (na przykładzie *Danio rerio*).
- W takie badań translacji analogów circRNA zaobserwowano synergistyczny efekt struktury kapu i sekwencji EMCV IRES (elementów zaangażowanych w mechanizmy translacji zależnej i niezależnej od kapu).

Opisane w niniejszej rozprawie badania i metody obarczone są następującymi ograniczeniami:

- Zaproponowany mechanizm aminowania redukcyjnego z EDA nie w pełni potwierdzony. W tym celu należałoby przeprowadzić kompleksowe badania, przykładowo z zastosowaniem jądrowego rezonansu magnetycznego oraz zbadać reaktywność N-alkilowych pochodnych hydrazyny czy hydroksyloaminy
- Metoda PORA ma ograniczone zastosowanie do modyfikacji endogennych RNA (np. eukariotycznych mRNA), które mogą zawierać więcej niż jedna strukturę cis-diolu.
- Do modyfikacji metodą PORA można wykorzystać wyłącznie analogi EDA, które nie reagują z cyjanoborowodorkami.
- Efekty obserwowane w trakcie pomiarów widm fluorescencyjnych sond FRET (zmiana parametru B) mogą być wynikiem oddziaływań wewnątrzcząsteczkowych lub międzycząsteczkowych. W trakcie badań nie podjęto próby ich rozróżnienia czy rozdzielania. W celu ułatwienia ich interpretacji, wyniki pomiarów skonfrontowano z modelami struktur drugo- i trzeciorzędowych RNA. Pomiary z udziałem sond FRET dają wgląd w średnią odległości między barwnikami (Cy5 i Cy3), będącą wypadkową całego krajobrazu konformacji oraz oddziaływań międzycząsteczkowych zachodzących w danych warunkach. Z drugiej strony, zastosowane metody obliczeniowe pozwalają

wyłączyć jedną z wielu możliwych struktur przestrzennych danej sekwencji RNA. Ponadto, ich trafność (prawdopodobieństwo na stworzenie rzeczywistej struktury) maleje wraz ze wzrostem rozmiaru modelowanej cząsteczki RNA. W praktyce, modele struktur przestrzennych RNA są bardzo cennym narzędziem, jednak należy do nich podchodzić z dystansem.

- W trakcie badań wykonano eksperymenty polegające na translacji w lizatach komórkowych (RRL). Przykładowymi zaletami tego typu układów jest brak ograniczeń związanych z dostarczaniem mRNA do komórek czy problemów wynikających z hodowli linii komórkowych. Jedną z największych wad jest silna zależność wyników od składu mieszaniny reakcyjnej, przez co wyniki mogą mieć niską zbieżność z wynikami otrzymanymi w komórkach.

Część III: Procedury doświadczalne

Sprzęt i odczynniki

Stężenia masowe kwasów nukleinowych [ng/μl] określano mierząc absorbancję przy 260 nm za pomocą aparatu NanoDrop 2000c (Thermo, 40 ng/μl odpowiada $A_{260} = 1$ dla długości drogi optycznej = 1 mm). Stężenia molowe kwasów nukleinowych określano na podstawie absorbancji roztworu oraz współczynnika ekstynkcji obliczonego dla każdej z sekwencji RNA na podstawie opublikowanego protokołu.^[217]

Po elektroforezie na żelach agarozowych lub poliakrylamidowych, wykonywano wizualizację fluorescencji za pomocą aparatu Typhoon FLA 9500 (GE Healthcare). Następnie przeprowadzono barwienie kwasów nukleinowych za pomocą bromku etydyny (Roth) lub SybrGold (Invitrogen). Po barwieniu, fluorescencja interkalatora była wizualizowana za pomocą aparatu Typhoon FLA 9500 lub GelDoc XR+ (BioRad).

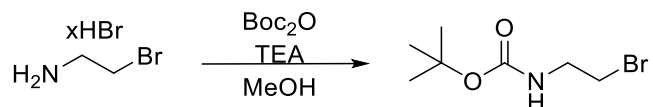
Wszystkie rozdziały HPLC przeprowadzono na modułowych, niskociśnieniowych gradientowych aparatach HPLC produkcji Shimadzu (seria LC-40/LC-20) lub Agilent (seria 1260/1290 Infinity) z termostatowanym uchwytem kolumny lub piecem kolumnowym, detektorem DAD i detektorem fluorescencyjnym. Faza ruchoma składała się z nieautoklawowanej wody MQ, acetonitrylu klasy HPLC (J.T.Baker), wysokiej jakości octanu amonu (J.T.Baker), trietyloaminy (Bioultra, Sigma), n-heksyloaminy klasy HPLC (Ottokemi) oraz lodowatego kwasu octowego klasy HPLC (J.T.Baker). Rozpuszczalniki o pożądanym składzie, stężeniu i pH filtrowano za pomocą filtra 0.45 μm i przechowywano w ciemności w temperaturze 4°C przez dłuższy czas. Chromatogramy rejestrowano z detekcją absorbancji przy 260 nm (ref. 360 nm, szerokość pasma 10 nm), widma absorbancji (220-700 nm) oraz fluorescencji zainteresowania (Cy3: 550/565, Cy5: 650/665, FAM: 490/520).

Wysokorozdzielcze widma masowe (HRMS) otrzymanych związków zarejestrowane zostały w Środowiskowym Laboratorium Spektrometrii Mas Instytutu Biochemii i Biofizyki PAN, na aparacie Micromass QToF 1 MS, w trybie jonów ujemnych. Widma NMR wykonano odpowiednio na spektrometrze BRUKER AVANCE III HD (1H: 500.24, 31P: 202.49, 13C: 125.80 MHz) na Wydziale Fizyki UW.

Do chromatografii na złożu anionitowym DEAE-Sephadex A-25 (forma wodorowęglanowa), zastosowano liniowy gradient buforu trietyloamonowego w wodzie (TEAB, pH 7.3) oraz aparat Biotage Selekt System. Bufor przygotowywano przepuszczając gazowy CO₂ przez roztwór wodny trietyloaminy, do uzyskania odpowiedniego pH. Skład frakcji wykazujących absorbancję przy 260 nm został dodatkowo zbadany za pomocą HPLC i MS. Frakcje zawierające pożądane produkty połączono i odparowano do sucha, dodając porcjami etanolu w celu usunięcia nadmiaru buforu trietyloamonowego.

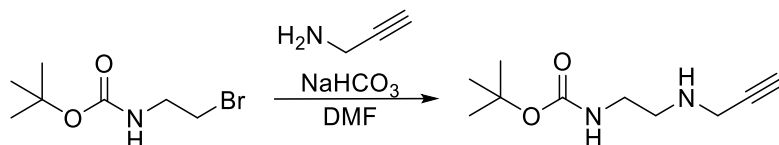
Syntezy chemiczne

Karbaminian (2-bromoetylo) tert-butylo



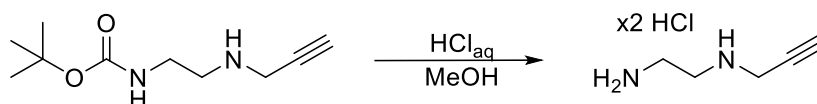
Syntezę przeprowadzono na podstawie opublikowanej metody.^[218] Do roztworu bromowodoru 2-bromoetyloaminy (1.02 g, 4.96 mmol) w metanolu (70 ml) dodano trietyloaminę (1.38 ml, 9.92 mmol) w temperaturze 0/4°C. Następnie, dodano roztwór diwęglanu di-tert-butyłu (1.09 g, 4.99 mmol) w metanolu (20 ml). Po 15 min mieszania w temperaturze 0/4°C usunięto łaźnię chłodzącą, a roztwór mieszano przez 2 h w temperaturze pokojowej. Roztwór rozcieńczono mieszaniną woda/dichlorometan (65 ml/65 ml), przeniesiono do rozdzielacza i przemyto and dichlorometanem (65 ml). Połączone frakcje organiczne wysuszone nad Na₂SO₄, przefiltrowano i zatężono pod zmniejszonym ciśnieniem. Otrzymano produkt w formie bezbarwnego oleju (1.11 g, 4.95 mmol, ~100%). ¹H NMR (500 MHz, CDCl₃) δ [ppm]: 4.94 (br s, 1H, CONH), 3.54 (t, *J* = 5.8 Hz, 2H, CH₂CH₂), 3.45 (t, *J* = 5.8 Hz, 2H, CH₂CH₂), 1.45 (s, 9H, tBut)

Karbaminian (N-propargiloaminoetylo) tert-butylo



Do karbaminianu (2-bromoetylo) tert-butylo (495 mg, 2.21 mmol) dodano wodorowęglan sodu (200 mg, 2.38 mmol), DMF (2.0 ml), oraz propargiloaminę (849 µl, 13.25 mmol). Zawiesina była refluksowana w 60°C przez 4.5 h. Przebieg reakcji był monitorowany na TLC (n-heksan/2-propanol 1:1, barwienie ninhydriną, RF~0.4). Zawiesina została rozcieńczona mieszaniną nasyconego węglanu sodu i dichlorometanu (20 ml / 30 ml), przeniesiona do rozdzielacza i przemyta dichlorometanem (2 x 20 ml). Połączone fazy organiczne zostały osuszone nad Na₂SO₄, przefiltrowane i zatężone pod zmniejszonym ciśnieniem. Otrzymany rudy olej został rozdzielony za pomocą chromatografii FLASH (dryload, 12g silicagel cartridge) w schodkowym gradiencie 2-propanolu w n-heksanie. Frakcje zawierające pożądany produkt połączono i zatężono pod zmniejszonym ciśnieniem. Otrzymano produkt w formie żółtego oleju (374 mg, 1.88 mmol, 86%). ¹H NMR (500 MHz, CDCl₃) δ [ppm]: 3.42 (d, *J* = 2.4 Hz, 2H, CH₂CCH), 3.24 (q, *J* = 5.7 Hz, 2H, CH₂CH₂), 2.81 (t, *J* = 5.7 Hz, 2H, CH₂CH₂), 2.22 (t, *J* = 2.4 Hz, 1H, CCH), 1.44 (s, 9H, tBut).

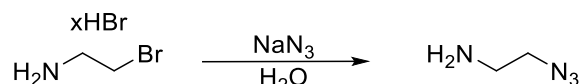
Dichlorowodorek N-propargiloetylenodiaminy (PEDA)



Do roztworu karbaminianu (N-propargiloaminoetylo) tert-butylo (374 mg, 1.88 mmol) w metanolu (20 ml) wkroplono kwas solny (4 ml, ~37% wt.). Po 30 min inkubacji w temperaturze pokojowej dodano etanol i odparowano roztwór pod zmniejszonym ciśnieniem. Następnie dodawano porcjami bezwodny etanol do momentu wytrącenia się

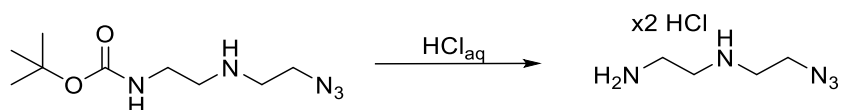
osadu. Następnie mieszanina została schłodzona, osad przefiltrowano i przemyto minimalną objętością zimnego bezwodnego etanolu (zużyto całkowicie 100 ml bezwodnego etanolu). Osad rozpuszczono w wodzie i zliofilizowano by otrzymać produkt w formie kremowego proszku (156 mg, 0.912 mmol, 48%). ^1H NMR (500 MHz, D_2O) δ [ppm]: 4.08 (d, $J = 2.6$ Hz, 2H, CH_2CCH), 3.58 (td, $J = 7.0$, 1.8 Hz, 2H, CH_2CH_2), 3.46 (td, $J = 7.0$, 1.9 Hz, 2H, CH_2CH_2), 3.10 (t, $J = 2.6$ Hz, 1H, CCH). ^{13}C NMR (126 MHz, D_2O) σ [ppm]: 79.01 (CCH), 43.11 (CH_2CH_2), 36.98 (CH_2CCH), 35.37 (CH_2CH_2).

2-azydoetyloamina



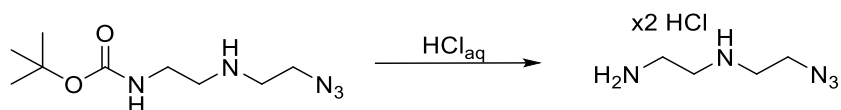
Syntezę przeprowadzono na podstawie opublikowanej metody.^[219] Do roztworu azydku sodu (8.14 g, 148 mmol) w wodzie (40 ml) dodano bromowodorek 2-bromoetyloaminy (10.09 g, 49.3 mmol) i refluksowano w 70°C przez 16 h. Mieszaninę schłodzono i dodano roztwór wodorotlenku potasu (14 g) w wodzie (10 ml), a następnie dichlorometan (50 ml). Po mieszanii w temperaturze pokojowej przez 30 min, zawiesina została przeniesiona do rozdzielacza i przemyta dichlorometanem (5 x 50 ml). Połączone fazy organiczne zostały osuszone Na_2SO_4 , przefiltrowane i ostrożnie zatężone pod zmniejszonym ciśnieniem. (w zakresie ciśnień 600-50 mbar w 30°C, do momentu ustalenia się masy produktu). Otrzymano produkt w formie bezbarwnego oleju (3.78 g, 43.9 mmol, $d = 1.04$ g/ml, 89%).

Karbaminian [N-(2-azydoetylo)aminoetylo] tert-butyłow



Do zawiesiny składającej się z 2-azydoetyloaminy (1.63 ml, 19.73 mmol), wodorowęglanu sodu (0.91 g, 6.58 mmol) i DMF (8 ml) wdroplono roztwór karbaminianu (2-bromoetylo) tert-butyłow (1.47 g, 6.58 mmol) w DMF (2 ml) w ciągu 1 h w temperaturze 70°C. Przebieg reakcji był monitorowany na TLC (n-heksan/2-propanol 1:1, barwienie ninhydriną, $\text{RF} \sim 0.4$). Zawiesina została rozcieńczona mieszaniną nasyconego węglanu sodu i dichlorometanu (40 ml / 50 ml), przeniesiona do rozdzielacza i przemyta dichlorometanem (3 x 50 ml). Połączone fazy organiczne zostały osuszone Na_2SO_4 , przefiltrowane i zatężone pod zmniejszonym ciśnieniem. Otrzymany bezbarwny olej został rozdzielony za pomocą chromatografii FLASH (dryload, 12g silicagel cartridge) w schodkowym gradiencie 2-propanolu w n-heksanie. Frakcje zawierające pożądany produkt połączono i zatężono pod zmniejszonym ciśnieniem. Otrzymano produkt w formie bezbarwnego oleju (940 mg, 4.14 mmol, 63%).

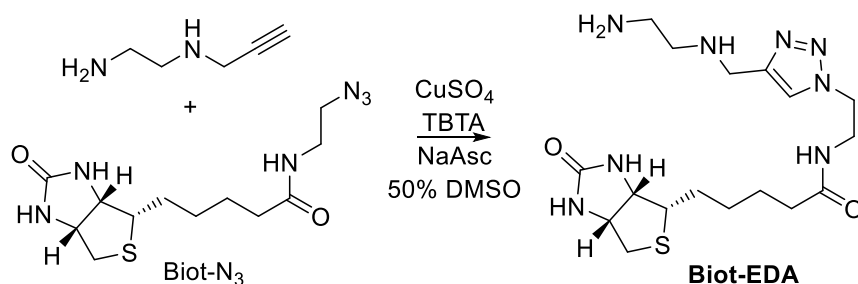
Dichlorowodorek N-(2-azydoetylo)etylenodiaminy (AEEDA)



Do karbaminianu [N-(2-azydoetylo)aminoetylo] tert-butyłow (374 mg, 1.88 mmol) wdroplono rozcieńczony kwas solny (5 ml, ~10% wt.). Po 1h inkubacji w temperaturze pokojowej mieszaninę rozcieńczono wodą (20 ml), przeniesiono do rozdzielacza i

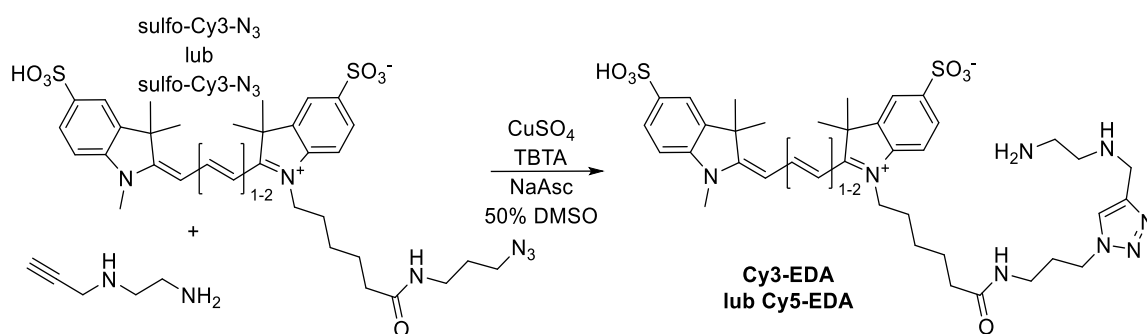
przemyto dichlorometanem (3 x 25 ml). Do fazy wodnej dodano etanol i odparowano roztwór pod zmniejszonym ciśnieniem. Następnie dodawano porcjami bezwodny acetonitryl do momentu wytrącenia się osadu. Następnie mieszanina została schłodzona, osad przefiltrowano i przemyto minimalną objętością zimnego bezwodnego etanolu. Osad osuszono pod zmniejszonym ciśnieniem by otrzymać produkt w formie białego proszku (70 mg, 0.53 mmol, 13%). ^1H NMR (500 MHz, D_2O) σ [ppm]: 3.82 (t, J = 5.5 Hz, 2H), 3.49 (m, 2H), 3.43 (m, 2H), 3.35 (t, J = 5.5 Hz, 2H). ^{13}C NMR (126 MHz, D_2O) σ [ppm]: 49.61, 49.48, 46.84, 38.01.

Biot-EDA



Syntezę Biot- N_3 przeprowadzono na podstawie opublikowanej metody.^[219] Do odważonych reagentów Biot- N_3 (9.81 mg, 31.4 μmol), PEDAx2HCl (8.27 mg, 48.4 μmol) i askorbinianu sodu (146 mg, 234 μmol) dodano mieszaninę 50% DMSO (1.26 ml) oraz roztwór kompleksu CuSO_4 -TBTA (340 μl , 9.4/10 mM w 50% DMSO). Po mieszanii przez 105 min w temperaturze pokojowej, do roztworu dodano EDTA (400 μl 0.5 M) i wodę (6 ml). Po filtracji za pomocą filtru strzykawkowego przeprowadzono rozdział na HPLC: kolumna SUPELCOSIL LC-18-T 250x4.6 mm 5 μm , A – 50 mM NH_4OAc pH 5.9, B – 50 mM NH_4OAc pH 5.9 / MeOH 1:1, 1.3 ml/min @ 22°C, gradient: 0-100% B w 30 min (RT = 15 min). Po połączeniu i trzykrotnej liofilizacji frakcji otrzymano produkt Biot-EDA w formie soli octanowej (4.5 mg, 8.48 μmol , M = 530.7 g/mol) z wydajnością 29%. MS ESI(+): m/z = 411.4 (Teor. $[\text{M}+\text{H}]^+$: m/z = 411.2).

Cy3-EDA oraz Cy5-EDA

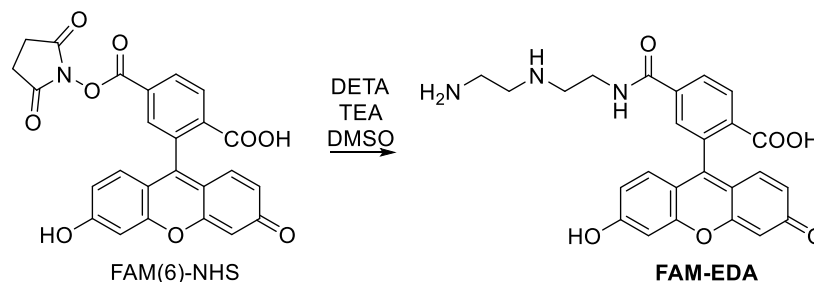


Do sulfo-Cy3- N_3 (2.22 mg, 3.01 μmol , Lumiprobe) dodano askorbinian sodu (5.19 mg, 26.2 μmol) oraz PEDAx2HCl (1.13 mg, 6.63 μmol) i rozpuszczono w 50% DMSO (270 μl). Następnie dodano roztwór kompleksu CuSO_4 -TBTA (32 μl , 9.4/10 mM w 50% DMSO) i inkubowano przez 60 min w temperaturze pokojowej (22°C). Następnie dodano wodę (2.7 ml) i roztwór EDTA (6 μl , 0.5 M, pH 8.0) i przeprowadzono rozdział HPLC kolumna Gemini® NX-C18 5 μm 110 Å 250 x 10 mm; układ A: 100 mM TEAA pH 7.0 B: 75% MeCN, program: 0-27% B w 40 min, 100% B przez 10 min, 5 ml/min @25°C (RT

= 35 min). Po połączeniu i trzykrotnej liofilizacji frakcji, oraz rozpuszczeniu w 50% DMSO (180 μ l) otrzymano produkt w formie roztworu soli octanu trietyloaminy (1.66 μ mol, 9.2 mM, $A_{550} = 1492$, $\epsilon_{548} = 162 \text{ mM}^{-1}\text{cm}^{-1}$) z wydajnością 55%. MS ESI(-): $m/z = 795.6$ (Teor. $[M-H]^-$: $m/z = 796.3$) ESI(+): $m/z = 797.4$ (Teor. $[M-H]^+$: $m/z = 797.3$).

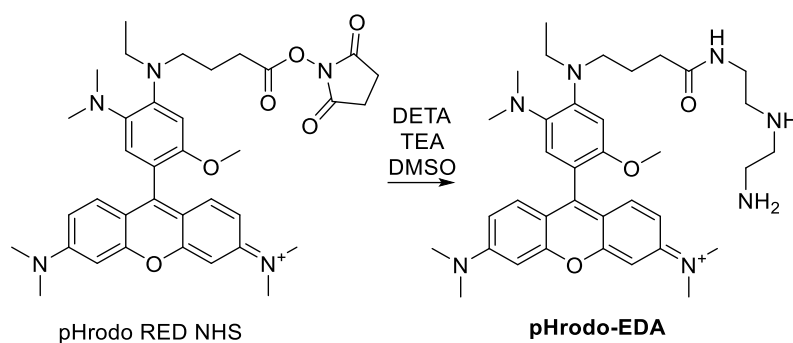
Syntezę Cy5-EDA przeprowadzono w analogiczny sposób wychodząc z sulfo-Cy5-N₃.

FAM-EDA



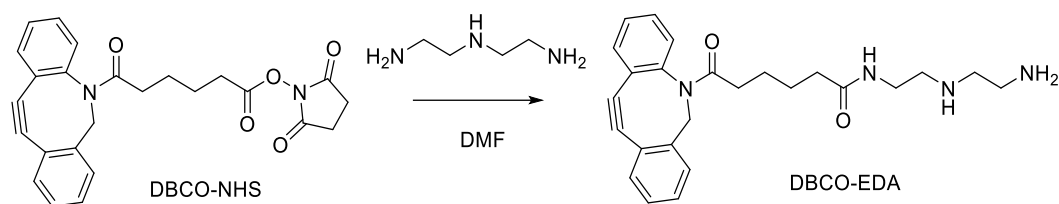
Do roztworu estru NHS 6-karboksyfluoresceiny (4.5 mg, 9.5 μ mol, ChemGenes) w DMSO (200 μ l) dodano trietyloaminę (10 μ l, bioultra), dietylenotriaminę (30 μ l) i inkubowano w 22 °C przez 60 min. Następnie dodano roztwór etanolu (1 ml, 80%) i odparowano pod zmniejszonym ciśnieniem. Czynność powtórzono dwukrotnie. Dodano roztwór octanu amonu (1.8 ml, 0.5 M, pH 5.9) i przeprowadzono rozdział za pomocą HPLC: kolumna Gemini® 5 μ m NX-C18 110 Å, 250x10 mm; układ A: 50 mM NH₄OAc pH 5.9 B: MeOH; gradient 0-50% B w 30 min, 5.0 ml/min @22°C (RT=27 min). Po trzykrotnej liofilizacji produkt w formie soli octanowo-amonowej (2.8 μ mol, $E_{490} = 83 \text{ mM}^{-1}\text{cm}^{-1}$, 30%) rozpuszczono w 60% DMSO otrzymując roztwór o stężeniu 10 mM. MS ESI(-): $m/z = 460.5$ (Teor. $[M-H]^-$: $m/z = 460.2$).

pHrodo-EDA



Do roztworu pHrodo RED NHS ester (1 mg, 2.0 μ mol, Thermo) w DMSO (200 μ l) dodano trietyloaminę (10 μ l, bioultra), dietylenotriaminę (20 μ l) i inkubowano w 22 °C przez 60 min w ciemności. Następnie dodano roztwór etanolu (1 ml, 80%) i odparowano pod zmniejszonym ciśnieniem. Czynność wykonano dwukrotnie. Dodano roztwór octanu trietyloaminy (1.8 ml, 0.1 M, pH 7.0) i przeprowadzono rozdział za pomocą HPLC: kolumna Gemini® 5 μ m NX-C18 110 Å, 250x10 mm; układ A: 50 mM AA pH 5.9 B: MeCN; gradient 0-100% B w 60 min, 5.0 ml/min @22°C (RT=26 min, MS). Po dwukrotnej liofilizacji i rozpuszczaniu w wodzie (100 μ l) otrzymano roztwór produktu (2.8 μ mol, $\epsilon_{560} = 65.0 \text{ mM}^{-1}\text{cm}^{-1}$), o stężeniu 12.6 mM ($A_{560} = 82$, $A_{260} = 30$, 1 mm, pH 7.0) z wydajnością 63%. MS ESI(+): $m/z = 630.4$ (Teor. $[M+H]^+$: $m/z = 630.4$).

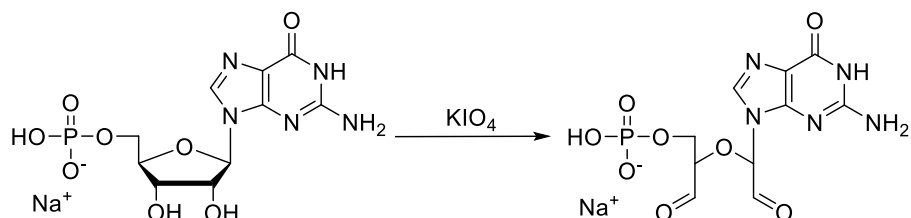
DBCO-EDA



DBCO-NHS (30.8 mg, 71.6 μmol , Lumiprobe) rozpuszczono w DMF (50 μl) i dodano dietylenotriaminę (38.7 μl , 358 μmol , 5 eq, Sigma). Po 20 min worteksowania w 37°C przeprowadzono analizę HPLC (próbka 4 μl R + 4 μl 50% AcOH + 2 ml 50 mM TEAA pH 7). Zaobserwowano powstanie produktów ubocznych, więc obniżono temperaturę do 20°C. Po godzinie zakończono reakcję poprzez dodanie 80 μl 50% AcOH oraz 1 ml 50 mM TEAA pH 7 (pH 6-7 ostatecznie) i zamrożenie w -20°C.

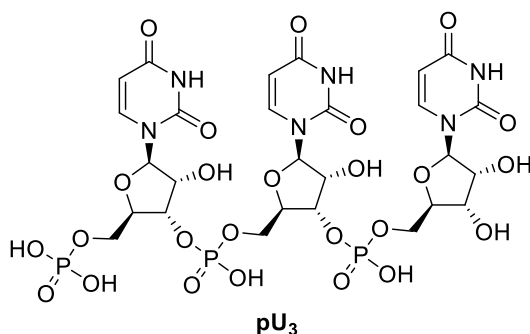
DBCO-NHS (32.1 mg, 74.6 μmol , Lumiprobe) rozpuszczono w DMF (100 μl) i dodano dietylenotriaminę (40.3 μl , 358 μmol , 5 eq, Sigma). Po godzinie worteksowania w 20°C przeprowadzono analizę HPLC. Pożądany produkt powstał selektywnie i wydajnie. Zakończono reakcję, produkty obu syntez połączono, a następnie rozdzielono za pomocą HPLC: kolumna Gemini 5 μm NX-C18 110 Å 250 x 10 mm, 5 ml/min @23°C, A: 50 mM TEAA pH 7, B: MeCN. Gradient: 0-41% B w 25 min, 41-100% B w 10 min, 100% B przez 5 min. Po dwukrotnej liofilizacji otrzymano produkt w formie brązowego proszku (13.4 mg, 25 μmol , $M_{+2\text{AcOH}}=540$ g/mol, 17%). MS ESI(+): $m/z = 419.3$ (Teor. $[\text{M}+\text{H}]^+$: $m/z = 419.5$).

GMP-dial



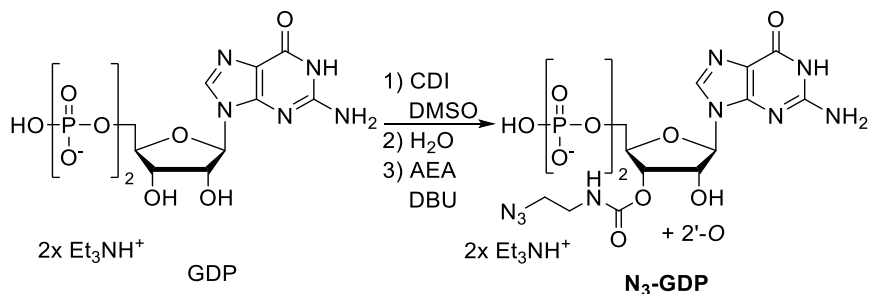
Dostępna handlowo sól sodowa 5' monofosforanu guanozyny została przekształcona w sól trietyloamoniową zgodnie z opublikowaną metodą.^[131] Sól trietyloamoniową 5' monofosforanu guanozyny (257 mg, 0,52 mmol, 6200 mOD₂₆₀) rozpuszczono w buforowanym roztworze octanu sodu (10.2 ml, 0.3 M pH 5.4). Następnie, stały nadjodan potasu (155 mg, 674 μmol , 1.25 eq) i mieszaninę mieszano przez godzinę w ciemności w temperaturze pokojowej. Następnie, do kolby wlewo aceton (250 ml). Otrzymany biały osad przefiltrowano i przemyto świeżym acetonem (150 ml) i wysuszono pod zmniejszonym ciśnieniem, otrzymując biały stały produkt (350 mg, 0.52 mmol, 6200 mOD₂₆₀, HPLC₂₆₀=99%, wydajność ~100%).

pU₃



Synteza została przeprowadzona na podstawie opublikowanej metody^[30] z wykorzystaniem syntetyzera AKTA Oligopilot plus 10 i złoża 5'-O-DMT-2'-O-TBDMS-rU 3'-Icaa Primer Support 5G ribo U 300 (170 mg, 50.7 μ mol, 298 μ mol/g, GE Healthcare). W trakcie sprzęgania złoże w kolumnie zostało przemyte roztworem 3'-O-amidofosforanu 5'-O-DMT-2'-O-TBDMS urydyny (ChemGenes) lub amidofosforanu biscyjanoetylowego (ChemGenes) w acetonitrylu (0.6 ml, 0.2 M, 2.4 eq) wraz z roztworem 5-(benzylthio)-1*H*-tetrazolu w acetonitrylu (0.30 M) w czasie 15 min. Jako odczynnik do detrytylacji wykorzystano roztwór kwasu dichlorooctowego w toluenie (3% v/v), jako utleniacza wykorzystano roztwór jodu w pirydynie (0.05 M), *N*-metylimidazol w acetonitrylu (20% v/v) jako Cap A oraz mieszaninę bezwodnika octowego (40% v/v) i pirydyny (40% v/v) w acetonitrylu jako Cap B. Po ostatnim cyklu syntetycznym produkt RNA na złożu inkubowano w roztworze dietyloaminy w acetonitrylu w celu usunięcia grup 2-cyjanoetylowych. Złoże zostało przemyte acetonitrylem i wysuszone argonem. W celu odcięcia i odbezpieczenia produktu, złoże inkubowano w AMA (3 ml 40% wt metyloaminy i 3 ml 30% wt wody amoniakalnej) przez godzinę w 40°C. Uzyskany roztwór odparowano a produkt rozpuszczono w DMAO (0.220 ml). Grupy TBDMS usunięto za pomocą trihydrofluorku trietyloaminy (250 μ l, 65°C, 3 h). Po ostudzeniu roztwór rozcieńczono roztworem wodorowęglanu sodu (20 ml, 0.25 M). Produkt został wydzielony za pomocą chromatografii jonowymiennej na złożu DEAE Sephadex (elucja gradientowa 0–1.2 M TEAB). Po odparowaniu frakcji otrzymano produkt w formie soli trietyloamoniowej (29 mg, 21.0 μ mol, 630 mOD₂₆₀, HPLC₂₆₀=99%, wydajność 57%). MS ESI(-): *m/z* = 636.3 (Teor. [M-H]⁻: *m/z* = 636.1).

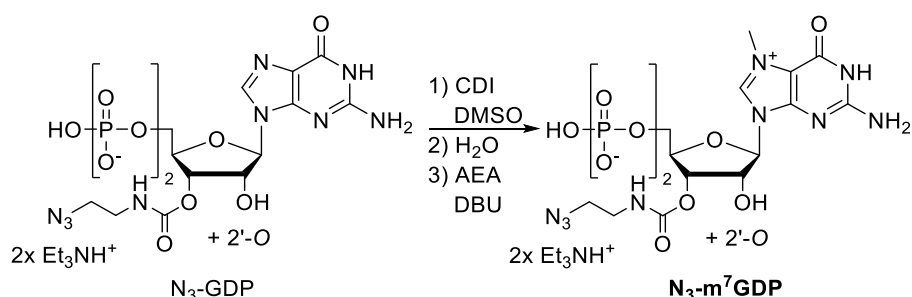
N₃-GDP



Reakcje przeprowadzono na podstawie wcześniej opublikowanej metody.^[131] Do roztworu GDP(TEAH)₂ (96 mg, 150 μ mol) w DMSO (1.5 mL) dodano 1,1-karbonylodiimidazol (CDI, 121 mg, 0.74 mmol, 5 eq.). Mieszaninę reakcyjną mieszano w temperaturze pokojowej i monitorowano za pomocą RP-HPLC. Po pięciu godzinach reakcji, nadmiar CDI zhydrolizowano wodą (20 μ L), dodano 2-azidoetyloaminę (96 mg, 93 μ L, 1.12 mmol,

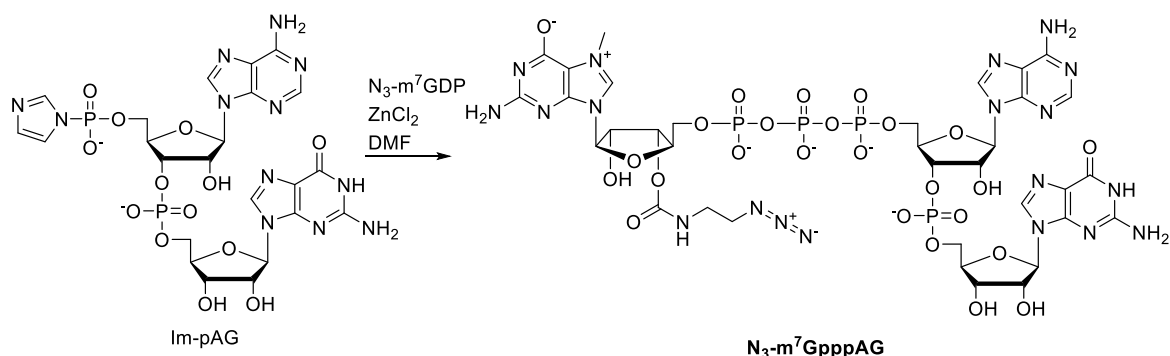
7.5 eq.), a następnie DBU (113 mg, 111 μ L, 0.74 mmol). Powstałą mieszaninę mieszano przez 3 h w 25°C. Postęp reakcji monitorowano za pomocą RP-HPLC. Kwaśną hydrolizę prowadzono w trzech etapach: przez 5 h w 22°C, przez noc w -20°C i następnego dnia 1.5 h w 22°C. Przeprowadzono rozdział na aparacie FLASH z użyciem złoża DEAE Sephadex (kolumna M, 40 ml/min) i rozdzielono w układzie 0-1.2 M TEAB. Po analizie frakcji chromatograficznych (HPLC) i połączeniu frakcji zawierających czysty produkt otrzymałem 620 ml roztworu o $A_{260} = 2.55$ (pH 7.0). Po odparowaniu i liofilizacji otrzymałem 105 mg (106 μ mol, 1.49 OD, 94% HPLC) soli trietyloamoniowej N_3 -GDP z wydajnością 71%. Analiza HPLC: kolumna Gemini® 5 μ m NX-C18 110 Å, 150x4.6 mm; układ A: 50 mM NH_4OAc pH 5.9 B: MeOH; 0-100% B w 30 min, 1.0 ml/min @22°C.

N_3 -m⁷GDP



Do zawiesiny N_3 -GDP (mieszanina izomerów) (105 mg, 106 μ mol) w DMSO (1.06 mL) dodano jodek metylu (1.58 mmol, 100 μ L, 15 eq.), a otrzymaną mieszaninę mieszano w 20°C. Postęp reakcji monitorowano za pomocą RP-HPLC. Po 3.5 h reakcję zatrzymano wodą (15 mL) i przeprowadzono ekstrakcję w układzie H₂O/ $CHCl_3$ (3x 1:1 v/v). Przeprowadzono rozdział na aparacie FLASH z użyciem złoża DEAE Sephadex (kolumna M, 40 ml/min) i rozdzielono w układzie 0-1.2 M TEAB. Po analizie frakcji chromatograficznych (HPLC) połączeniu frakcji (20-31) roztwór zamrożono na noc. Po zatkanie roztworu otrzymałem produkt (18 μ mol, 78% HPLC) w formie soli trietyloamoniowej z wydajnością 45%. Analiza HPLC: kolumna Gemini® 5 μ m NX-C18 110 Å, 150x4.6 mm; układ A: 50 mM NH_4OAc pH 5.9 B: MeOH; 0-100% B w 30 min, 1.0 ml/min @22°C.

N_3 -m⁷GpppAG

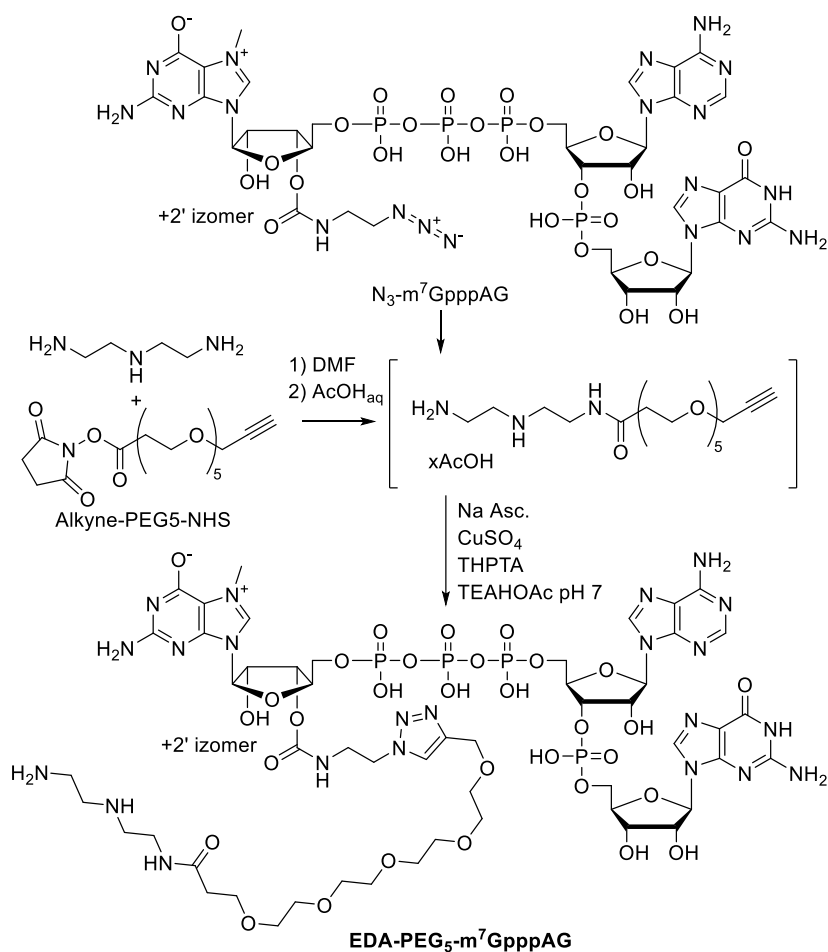


Reakcje przeprowadzono na podstawie wcześniej opublikowanej metody.^[30] Do wysuszonego Im-pAG^[30] (16.0 μ mol, 1.0 eq) dodano przygotowaną wcześniej mieszaninę wysuszonych reagentów: N_3 -m⁷GDP (17.6 μ mol, 1.1 eq) i ZnCl₂ (17.5 mg, 160 μ mol, 10 eq) w DMF (666 μ L). Otrzymaną zawiesinę worteksowano w 25°C, jednocześnie analizując

przebieg reakcji za pomocą HPLC. Po godzinie od rozpoczęcia dodano wodę (0.40 ml) oraz roztwór EDTA (0.34 ml, 0.5 M, pH 8.0) otrzymując roztwór o pH ~6, który zamrożono w -20°C na noc. Przeprowadzono rozdział na aparacie FLASH z użyciem złoża DEAE Sephadex, frakcje 8-19, 220 ml, A260 = 1.74, HPLC =85%, 325 mOD, E = 32.0 mM⁻¹cm⁻¹, 10.2 μmol, W = 64%.

Syntezę powtórzyła dr. Małgorzata Wasińska-Kałwa, a produkty oczyszczono metodą RP-HPLC z zastosowaniem liniowego gradientu (0-40%) acetonitrylu w octanie amonu (pH 5.9, 0.05 M) w czasie 60 min. Zebrane frakcje łączono i zatężano w celu usunięcia nadmiaru ACN. Produkt (mieszanina regioizomerów 2'-O/3'O) (3232 mOD₂₆₀, 55 μmol, 53%) otrzymano jako biały proszek po liofilizacji.

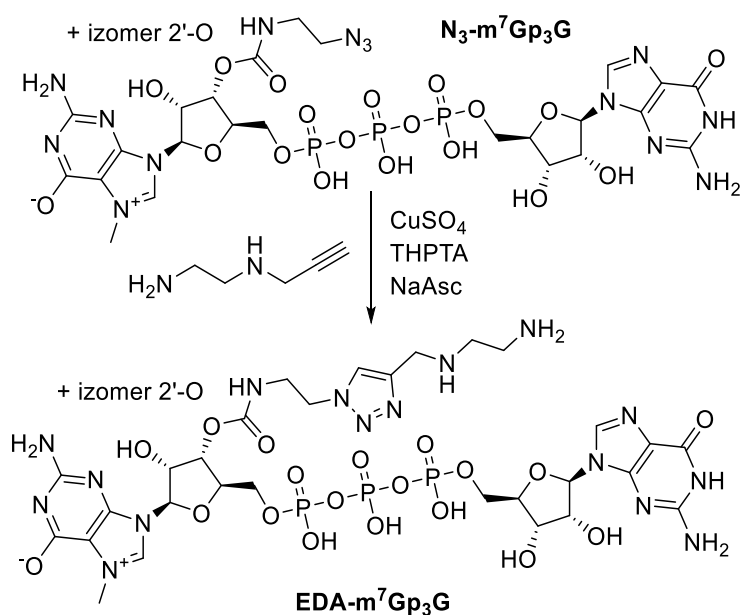
EDA-PEG₅-m⁷GpppAG



Alkyne-PEG₅-NHS (4.5 mg, 11.2 μmol, ChemCruz) rozpuszczono w DMF (70 μl) i dodano dietylenotriaminę (12 μl, 122 μmol, 11 eq, Sigma). Po godzinie worteksowania (2000 RPM) w 20°C dodano TEAHOAc (200 μl, 50 mM) i AcOH (24 μl, 50 %), pH 6-7 końcowo. Następnie dodano roztwory N₃-m⁷GpppAGxTEAH (10 μmol, 250 μl), Cu-THPTA (50 mM, 20 μl), oraz askorbinian sodu (15 mg, 76 μmol) i kontynuowano worteksowanie. Przebieg reakcji monitorowano na HPLC: kolumna Gemini 3 μm NX-C18 110 Å 150 x 4.6 mm, 1.0 ml/min @25°C, A: 50 mM NH₄OAc pH 5.9, B: 50 mM NH₄OAc pH 5.9 / MeOH 1:1 (v/v). Program: 0-100% B w 30 min. Po godzinie zwiększono temperaturę (do 25°C), zmniejszono intensywność mieszania (do 300 RPM) oraz dodano roztwór Cu-THPTA (50 mM, 10 μl) i askorbinian sodu (10 mg, 50 μmol). Po trzech

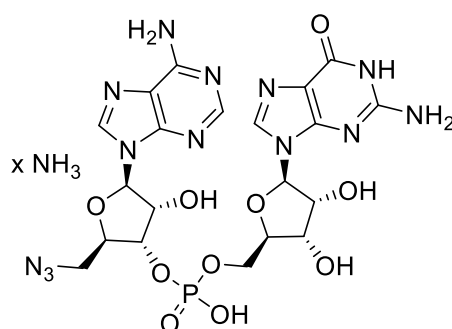
godzinach od rozpoczęcia reakcji CuAAC ponownie dodano roztwór Cu-THPTA (50 mM, 10 μ l) oraz askorbinian sodu (23 mg, 116 μ mol). Po pięciu godzinach od rozpoczęcia reakcji CuAAC mieszaninę zamrożono w -20°C na noc. Po rozmrożeniu dodano roztwory EDTA (450 μ l, 50 mM), NH_4OAc (4 ml, 50 mM), przefiltrowano mieszaninę za pomocą filtra strzykawkowego (0.45 μm). Następnie przeprowadzono rozdział na HPLC: kolumna Gemini 5 μm NX-C18 110 $\text{\AA}$ 250 x 10 mm, 5.0 ml/min @25 $^{\circ}\text{C}$, A: 50 mM NH_4OAc pH 5.9, B: MeCN. Program: 0-12% B w 35 min, 12-70% B w 5 min. Po dwukrotnej liofilizacji octanowo-amonową sól produktu rozpuszczono w wodzie (40 mM, 40 μ l, 1.6 μ mol, $\epsilon_{260} = 32.0 \text{ mM}^{-1}\text{cm}^{-1}$, HPLC 99%, W= 16%). HRMS: $m/z = 1631.42458$, 815.20983 (Teor. $[\text{M}-\text{H}]^-$: $m/z = 1631.42375$, $[\text{M}-2\text{H}]^{2-}$: $m/z = 815.20824$).

EDA- $m^7\text{GpppG}$



Do $\text{N}_3\text{-}m^7\text{Gp}_3\text{G}$ (10 mg, 8.4 μmol)^[131] dodano askorbinian sodu (10.3 mg, 52 μmol) oraz PEDAx2HCl (3.4 mg, 20 μmol) i rozpuszczono w odgazowanym roztworze octanu trietyloaminy (2.0 ml, 45 mM, pH 7). Następnie dodano roztwór kompleksu $\text{CuSO}_4\text{-THPTA}$ (10 μ l, 100/500 mM w wodzie) i inkubowano przez 2 h w temperaturze pokojowej. Następnie dodano roztwór EDTA (20 μ l, 100 mM, pH 7.0) i przeprowadzono rozdział HPLC kolumna Vydac Denali HiChrom C18 150x 10 mm 5 μm 120 $\text{\AA}$; układ A: 100 mM NH_4OAc pH 5.9 B: 30% MeCN, gradient: 0-25% B w 40 min, 4.5 ml/min @22 $^{\circ}\text{C}$. Po połączeniu i trzykrotnej liofilizacji frakcji otrzymano produkt w formie soli octanowo-amonowej (1.6 mg, 1.75 μmol , $\epsilon_{260} = 20.0 \text{ mM}^{-1}\text{cm}^{-1}$) z wydajnością 21%. MS ESI(-): 505.4 (Teor. $[\text{M}-2\text{H}]^{2-}$: $m/z = 505.1$), ESI(+): $m/z = 507.4$ (Teor. $[\text{M}+2\text{H}]^{2+}$: $m/z = 507.1$).

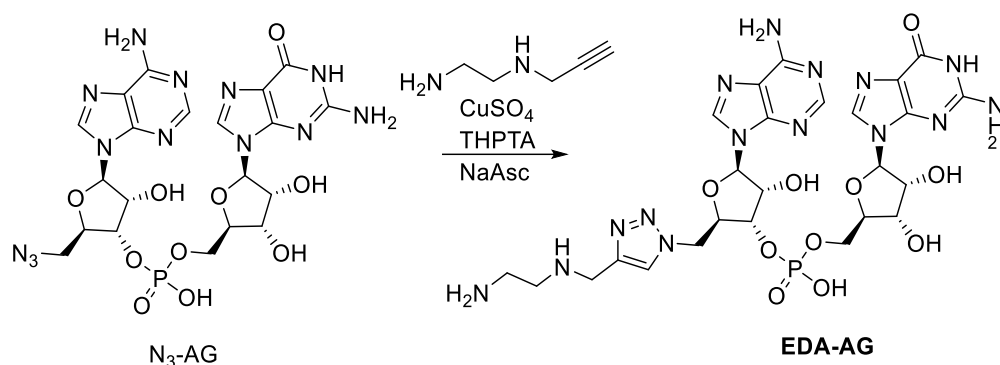
N₃-AG



Synteza została przeprowadzona na podstawie znanej metody^[220] z wykorzystaniem syntetyzera AKTA Oligopilot plus 10 i złoża 5'-O-DMT-2'-O-TBDMS-rG^{iBu} 3'-lcaa Primer Support 5G ribo G 300 (163 mg, 49.0 μmol, 300 μmol/g, GE Healthcare). W trakcie sprzęgania złoże w kolumnie zostało przemyte roztworem 3'-O-amidofosofrynu 5'-O-DMT-2'-O-TBDMS rA^{Pac} w acetonitrylu (0.6 ml, 0.2 M, 2.4 eq) wraz z roztworem 5-(benzylthio)-1*H*-tetrazolu w acetonitrylu (0.30 M) w czasie 15 min. Jako odczynnik do detrytylacji wykorzystano roztwór kwasu dichlorooctowego w toluenie (3% v/v), jako utleniacza wykorzystano roztwór jodu w pirydynie (0.05 M), *N*-methylimidazol w acetonitrylu (20% v/v) jako Cap A oraz mieszaninę bezwodnika octowego (40% v/v) i pirydyny (40% v/v) w acetonitrylu jako Cap B. Po ostatnim cyklu syntetycznym produkt RNA na złożu inkubowano w roztworze dietyloaminy w acetonitrylu w celu usunięcia grup 2-cyanoetylowych. Złoże zostało przemyte acetonitrylem i wysuszone argonem. Złoże przemywano w zamkniętym obiegu roztworem jodku trifenoksymetylofosfiny ((PhO)₃PCH₃⁺I⁻) w DMF (1.0 ml, 0.6 M) przez 15 min. Złoże zostało kolejno przemyte DMF, acetonitrylem, wysuszone argonem i przeniesione do probówki. Następnie dodano nasycony roztwór azydku sodu w DMF (1 ml) i mieszano intensywnie przez godzinę w 60°C. Złoże zostało kolejno przemyte wodą, etanolem, acetonitrylem, wysuszone argonem. W celu odcięcia I odbezpieczenia produktu, złoże inkubowano w AMA (3 ml 40% wt metyloaminy i 3 ml 30% wt wody amoniakalnej) przez godzinę w 50°C. Uzyskany roztwór odparowano a produkt rozpuszczono w wodzie. Produkt został wydzielony za pomocą chromatografii jonowymiennej na złożu DEAE Sephadex (elucja gradientowa 0–0.6 M TEAB). Po odparowaniu frakcji otrzymano produkt w formie soli trietyloamoniowej (21 mg, 20.8 μmol, 550 mOD₂₆₀, *E*₂₆₀ = 24.3 mM⁻¹cm⁻¹ HPLC₂₆₀ = 92%, wydajność 42%).

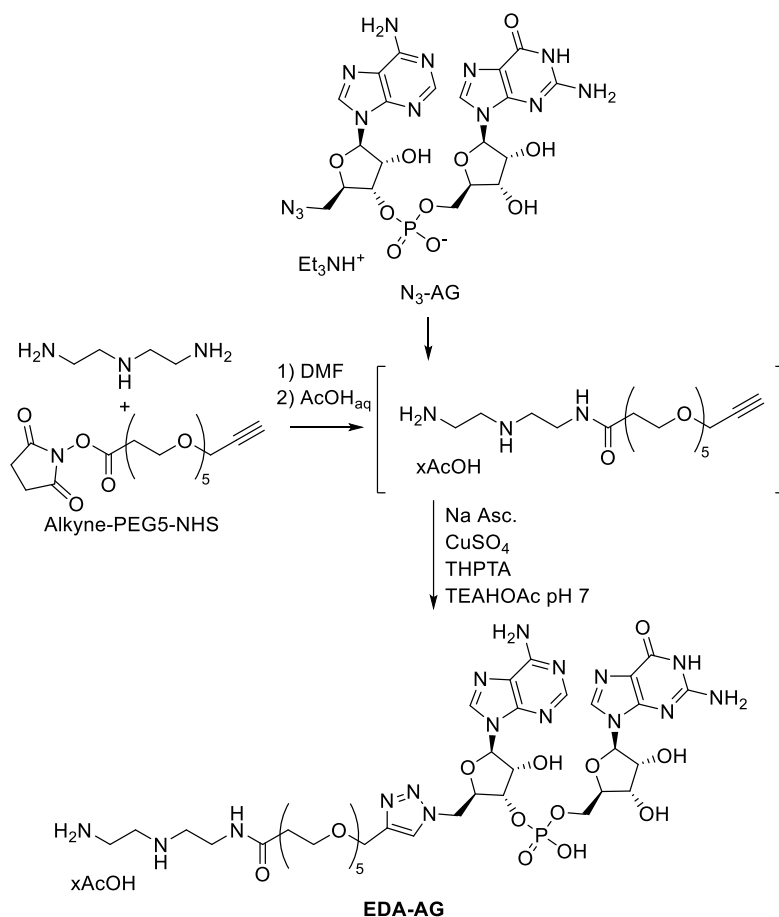
Do zastosowania w transkrypcji *in vitro* część produktu (11 mg, 11.1 μmol) została dodatkowo oczyszczona za pomocą HPLC: kolumna Vydac Denali HiChrom C18 150x 10 mm 5 μm 120 Å; rozpuszczalniki A - 50 mM NH₄OAc pH 5.9, B - 50 mM NH₄OAc pH 5.9 / MeCN 7:3 v/v, gradient: 0–25% B w 40 min, przepływ 4.5 ml/min w 25°C (RT = 25 min). Po połączeniu i trzykrotnej liofilizacji frakcji otrzymano produkt w formie soli amonowej (5.5 mg, 9.0 μmol, 220 mOD₂₆₀, *E*₂₆₀ = 24.3 mM⁻¹cm⁻¹, HPLC₂₆₀ ~ 100%, wydajność 81%). MS ESI(-): *m/z* = 636.3 (Teor. [M-H]⁻: *m/z* = 636.1).

EDA-AG



Do N₃-AG (7.4 mg, 11.5 μmol) dodano askorbinian sodu (10.3 mg, 52 μmol) oraz PEDAx2HCl (3.4 mg, 20 μmol) i rozpuszczono w odgazowanym roztworze octanu trietyloaminy (2.0 ml, 45 mM, pH 7). Następnie dodano roztwór kompleksu CuSO₄-THPTA (10 μl, 100/500 mM w wodzie) i inkubowano przez 2 h w temperaturze pokojowej. Następnie dodano roztwór EDTA (20 μl, 100 mM, pH 7.0) i przeprowadzono rozdzielanie HPLC kolumna Vydac Denali HiChrom C18 150x 10 mm 5 μm 120 Å; układ A: 100 mM NH₄OAc pH 5.9 B: 30% MeCN, gradient: 0-25% B w 40 min, 4.5 ml/min @22°C. Po połączeniu i trzykrotnej liofilizacji frakcji otrzymano produkt w formie soli octanowo-amonowej (2.8 mg, 3.50 μmol, $E_{260} = 24.3 \text{ mM}^{-1}\text{cm}^{-1}$) z wydajnością 30%. MS ESI(-): m/z = 734.4 (Teor. [M-H]⁻: m/z = 734.2).

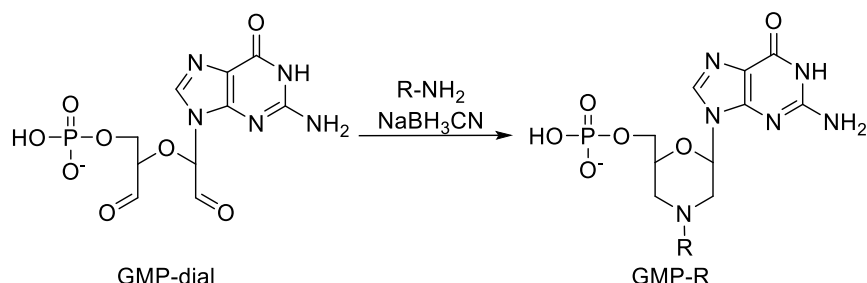
EDA-PEG₅-AG



Alkyne-PEG5-NHS ester (12.5 mg, 31.2 μmol , Sigma) rozpuszczono w DMF (200 μl) i dodano dietylenotriaminę (33.7 μl , 310 μmol , 10 eq, Sigma). Po 2.5 h worteksowania w 20°C dodano NH_4OAc (50 μl , 500 mM) i AcOH (60 μl , 50 %), pH 7-8 końcowo. Następnie dodano $\text{N}_3\text{-AGxTEAH}$ (28 μmol , 2.00 ml), Cu-THPTA (3 μmol , 50 mM, 60 μl), oraz askorbinian sodu (31 mg, 157 μmol). Przebieg reakcji monitorowano na HPLC. Po dwóch godzinach dodano TEAA (200 μl , 1.0 M), Cu-THPTA (1 μmol , 50 mM, 20 μl), oraz askorbinian sodu (114 mg, 586 μmol). Po nocy inkubacji w 20°C dodano EDTA (20 μl , 0.5 M) i przeprowadzono rozdział na HPLC: kolumna Gemini 5 μm NX-C18 110 Å 250 x 10 mm, 5 ml/min @25°C, A: 50 mM NH_4OAc pH 5.9, B: MeCN. Program: 0-41% B w 25 min, 41-100% B w 10 min, 100% B przez 5 min. Po dwukrotnej liofilizacji rozpuszczono produkt w wodzie (43 mM, 9.2 μmol , HPLC 95%, W= 33%). HRMS: m/z = 1025.39748 (Teor. $[\text{M-H}]^-$: m/z = 1025.39598).

Analizy chemiczne

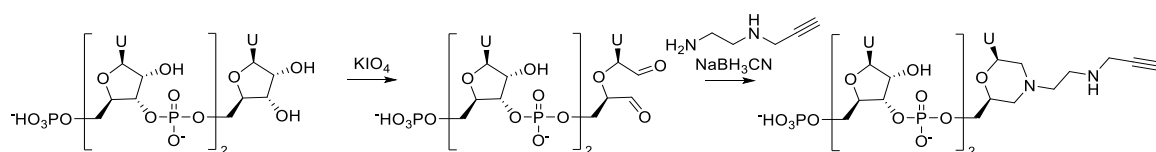
Reakcje między GMP-dial a N-nukleofilami



Ogólna procedura: Dejonizowana i odgazowana woda (150 μl) została zmieszana z buforowanym roztworem aminy (10.0 μl , 20 mM lub 40 mM w buforze fosforanu potasu 1.0 M pH 6.0) i świeżym roztworem cyjanoborowodoru sodu (20.0 μl , 100 mM w wodzie). Następnie do mieszaniny dodano roztwór GMP-dial (20.0 μl , 10 mM w wodzie) i inkubowano w temperaturze 30°C. Co 20 min od rozpoczęcia pobierano próbkę (50 μl) mieszaniny reakcyjnej, którą natychmiast analizowano za pomocą HPLC: kolumna SUPELCOSIL LC-18-T 250x4.6 mm 5 μm ; rozpuszczalniki A - 50 mM NH_4OAc pH 5.9, B - 50 mM NH_4OAc pH 5.9 / MeOH 1:1 v/v; gradient: 0-100% B w 30 min, przepływ 1.3 ml/min w 25°C.

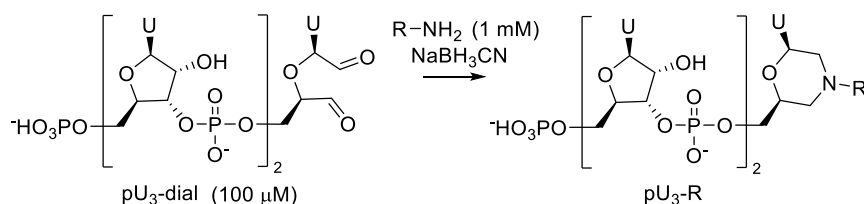
Reakcje z cysteiną i cysteaminą: Dejonizowaną i odgazowaną wodę (130 μl) zmieszano z buforowanym roztworem aminy (10.0 μl , 20 mM w buforze fosforanu potasu 1.0 M pH 6.0) i świeżym roztworem ditiotretolu (20.0 μl , 20 mM w wodzie) i inkubowano w 30°C przez 10 min. Następnie dodano świeży roztwór cyjanoborowodoru sodu (20.0 μl , 100 mM w wodzie) i roztwór GMP-dial (20.0 μl , 10 mM w wodzie) dodawano do mieszaniny i inkubowano w 30°C. Co 20 min od rozpoczęcia pobierano próbkę (50 μl) mieszaniny reakcyjnej, którą natychmiast analizowano za pomocą HPLC (warunki jak wyżej).

Optymalizacja protokołu PORA w jednym naczyniu



Do roztworu soli trietyloamoniowej pU₃ (20–2000 μM) w buforze fosforanu potasu (500 mM pH 6.0) dodano stały nadjodan potasu (1–2 mg). Mieszaninę mieszano w ciemności w temperaturze 30°C aż do całkowitego utlenienia monitorowanego za pomocą HPLC (20–50 min). Następnie roztwór odwirowano i podzielono na porcje (po 250 μl). Następnie do wybranej porcji dodano świeży, wstępnie wymieszany roztwór 2xHCl PEDA (0.2–2.0 mM) i cyjanoborowodoru sodu (200 mM) w wodzie (250 μl), a powstałą mieszaninę reakcyjną (500 μl całkowitej objętości) inkubowano w 30°C. W celu monitorowania postępu reakcji próbki mieszanin reakcyjnych (15, 150 lub 200 μl) były nastrzykiwane na HPLC: kolumna SUPELCO SIL LC-18-T 250x4.6 mm 5 μm; rozpuszczalniki A - 50 mM NH₄OAc pH 5.9, B - 50 mM NH₄OAc pH 5.9 / MeOH 1:1 v/v; gradient: 0–50% B w 15 min, przepływ 1.3 ml/min w 22°C.

Kinetyka aminowania redukcyjnego z pU₃-dial



Etap utleniania: Do wodnego roztworu soli trietyloamoniowej pU₃ (1080 μl, 167 μM) dodano świeży roztwór NaIO₄ (180 μl, 10 mM) i inkubowano w ciemności w temperaturze 25 °C przez 35 min. Otrzymany roztwór pU₃-dial przechowywano na lodzie (lub zamrażano w -20 °C na noc).

Etap aminowania redukcyjnego: pU₃-dial (105 μl, 143 μM), świeży NaBH₃CN (30 μl, 100 mM), i roztwór N-nukleofila (15 μl, 10 mM) w buforze fosforanu potasu (pH 6, 1.0 M) zmieszano i inkubowano w temperaturze 25 °C w zespole autosamplera HPLC. Analizy wykonywano w odstępach 15 min. Konfiguracja HPLC: kolumna: Clarity 1.7 μm Oligo-XT 100 Å 100 x 2.1 mm. Fazy (A) 100 mM TEAHOAc pH 7.0; (B) 200 mM TEAHOAc/MeCN 1:1 v/v; przepływ 450 μl / min; 60 °C; gradient: od 0–16.7% B w ciągu 10 min; 0% B przez 5 min.

Stałe szybkości obliczono wykonując następujące operacje: 1) ustalono linię bazową dla chromatogramów HPLC tworząc linię przecinającą chromatogram w czasie 0.2–0.4 min oraz w czasie retencji 9.5–10.0 min; 2) zintegrowany chromatogram był uszczuplony o sygnały występujące w czasie martwym (0.0–1.5 min); 3) uzyskana powierzchnia całkowita (PC_i) i powierzchnia piku produktu pU₃-R oraz powierzchnię piku produktu (PP_i) wprowadzono do wzoru na 1/C_{Si} (odwrotność stężenia substratu C_S w danym punkcie czasom i):

$$1/C_{Si} = \frac{10^4}{1 - \frac{PP_i \times PC_0}{PC_i \times PS_0}} \times \frac{L}{mol}$$

gdzie PC₀ to powierzchnia całkowita, a PS₀ to powierzchnia substratu pU₃-dial przed rozpoczęciem aminowania redukcyjnego; 4) wartości 1/C_{Si} dla każdego chromatogramu (punktu czasowego i) były wykreślane względem czasu reakcji (w sekundach). Po analizie regresji liniowej wyznaczono stałą szybkości reakcji drugiego rzędu k [M⁻¹s⁻¹] jako nachylenie dopasowanej prostej.

Transkrypcja *in vitro*

Ogólna procedura

Przygotowano mieszaniny reakcyjne w temperaturze pokojowej. Typowa mieszanina reakcyjna zawierała końcowo: Tris-HCl (40 mM, pH 7.9), MgCl₂ (31 mM), DTT (10 mM), NaCl (10 mM) i spermidyny (2 mM), PPazę (2 U/ml, Thermo), Ribolock (1 U/μl, Thermo), polimerazę RNA T7 (10 U/μl, Thermo); matrycę DNA w formie linearyzowanego plazmidu (30-60 ng/μl) lub dwuniciowego oligoDNA (1 μM), inicjator i NTP. Stężenia inicjatora i NTP opisano poniższą tabelą:

Typ RNA	Inicjator		Stężenie NTP		Wydajność [g/l]
	Typ	Stężenie	GTP*	A/U/CTP*	
ppp-RNA	-	-	5 mM	3x 5 mM	4.5–5.6
p-RNA	GMP*	20 mM	3 mM		1.7–2.8
N ₃ -RNA	N ₃ -AG, EDA-AG, EDA-PEG5-AG	12 mM	3 mM		1.7–1.8
R-ArcaRNA	ARCA, N ₃ -ARCA, EDA-ARCA	6 mM	1 mM		~1.2
R-Cap#RNA	m ⁷ GpppAG, N ₃ -m ⁷ GpppAG, EDA-PEG ₅ -m ⁷ GpppAG	12 mM	3 mM		3.2–3.8

Po dwóch godzinach inkubacji w 37°C oraz DNazowaniu (dodatek DNazy I, Thermo, końcowo 0.05 U/ μl i następcza inkubacja 15 min w 37°C) produkty reakcji wydzielono za pomocą Monarch RNA Cleanup Kit (NEB). Otrzymano RNA w formie roztworów z wydajnościami zaprezentowanymi w wyżej wymienionej tabeli.

Przed cyklizacją, translacją *in vitro* i eksperymentami biologicznymi produkty transkrypcji były dodatkowo oczyszczane za pomocą HPLC. Kolumna: PolymerX 3μm RP-1 100 Å 150 x 4.1 mm (00F-4338-Z0, Phenomenex) z SecurityGuard RP-1 3.0 mm ID (KJ0-4282 i AJ0-5809, Phenomenex); rozpuszczalnik A: 100 mM octan trietyloamonowy pH 7. 0; rozpuszczalnik B: 100 mM octan trietyloaminy w 50% acetonitrylu; gradient: 15-35% B w 30 min (dla RNA o długości 30–300 nt) lub 20-35% B w 30 min (dla RNA o długości >300 nt), następnie 35-100% B w 1 min i 100% B przez 2 min; Przepływ: 1.0 ml/min @ 60°C. Zebrane frakcje dzielono na porcje (po 0.7 ml) i dodawano octan sodu (70 μl, 3.0 M, pH 5.2), glikogen (1.0 μl, 5 mg/ml, Thermo) oraz 2-propanol (0.80 ml, 99.8%) (mieszając po każdym dodaniu). Roztwór chłodzono przez 30 min w temperaturze -80°C. Pelet odwirowano (30 min, 15 000 g, 4°C), przemyto etanolem (0.80 ml, 80%), odwirowano (20 min, 15 000 g, 4°C), wysuszono pod zmniejszonym ciśnieniem i rozpuszczono w wodzie.

Znakowanie RNA

Znakowanie końca 3' mRNA

Do roztworu RNA (12 μl, stężenie w zakresie 0.1–2.0 μg/μl) dodano świeży roztwór NaIO₄ (2 μl, 10 mM) i inkubowano przez 30 min w 25°C bez dostępu do światła. Następnie, dodano bufor KH₂PO₄ (2 μl, 1 M, pH 6.0), świeży roztwór NaBH₃CN (2 μl, 200 mM) oraz analogu EDA (2 μl, 10 mM: Cy3-EDA, Cy5-EDA, FAM-EDA, Biot-EDA). Po inkubacji

przez 120 min w 25°C produkty reakcji wydzielono za pomocą Monarch RNA Cleanup Kit (NEB, odzysk 80–95%) lub poprzez strącanie w układzie NaOAc/EtOH (odzysk 75–95%, zanieczyszczenie solami fosforanowymi).

Znakowanie końca 5' mRNA (Cy5)

Do roztworu N₃-RNA (16 µl, stężenie w zakresie 0.1–1.5 µg/µl) dodano bufor KH₂PO₄ (2 µl, 1 M, pH 6.0) oraz DIBAC-sCy5 (2 µl, 20 mM, 50% DMSO, Lumiprobe). Po inkubacji przez 120 min w 25°C produkty reakcji wydzielono za pomocą Monarch RNA Cleanup Kit (NEB, odzysk 80–95%) lub poprzez strącanie w układzie NaOAc/EtOH (odzysk 75–95%, zanieczyszczenie solami fosforanowymi).

Podwójne znakowanie mRNA (5' Cy5 i 3' Cy3)

Do roztworu N₃-RNA (10 µl, stężenie w zakresie 0.1–2.0 µg/µl) dodano świeży roztwór NaIO₄ (2 µl, 10 mM) i inkubowano przez 30 min w 25°C. Następnie dodano bufor KH₂PO₄ (2 µl, 1 M, pH 6.0), świeży roztwór NaBH₃CN (2 µl, 200 mM), DIBAC-sCy5 (2 µl, 20 mM, 50% DMSO, Lumiprobe) oraz Cy3-EDA (2 µl, 10 mM, 50% DMSO). Po inkubacji przez 120 min w 25°C produkty reakcji wydzielono za pomocą Monarch RNA Cleanup Kit (NEB, odzysk 80–95%) lub poprzez strącanie w układzie NaOAc/EtOH (odzysk 75–95%, zanieczyszczenie solami fosforanowymi).

Izolacja znakowanego mRNA

Przed translacją *in vitro* i eksperymentami biologicznymi produkty znakowania były dodatkowo oczyszczane za pomocą HPLC. Kolumna: PolymerX 3µm RP-1 100 Å 150 x 4.1 mm (00F-4338-Z0, Phenomenex) z SecurityGuard RP-1 3.0 mm ID (KJ0-4282 i AJ0-5809, Phenomenex); rozpuszczalnik A: 100 mM octan trietyloamonowy pH 7.0; rozpuszczalnik B: 100 mM octan trietyloaminy w 50% acetonitrylu; gradient: 15-35% B w 30 min (dla RNA o długości 30–300 nt) lub 20-35% B w 30 min (dla RNA o długości >300 nt), następnie 35-100% B w 1 min i 100% B przez 2 min; Przepływ: 1.0 ml/min @ 60°C. Zebrane frakcje dzielono na porcje (po 0.7 ml) i dodawano octan sodu (70 µl, 3.0 M, pH 5.2), glikogen (1.0 µl, 5 mg/ml, Thermo) oraz 2-propanol (0.80 ml, 99.8%) (mieszając po każdym dodaniu). Roztwór chłodzono przez 30 min w temperaturze -80°C. Pelet odwirowano (30 min, 15 000 g, 4°C), przemyto etanolem (0.80 ml, 80%), odwirowano (20 min, 15 000 g, 4°C), wysuszono pod zmniejszonym ciśnieniem i rozpuszczono w wodzie (odzysk w przedziale 20–70%, zależny od jakości produktu i skali rozdzielania).

Cyklizacja RNA

Ogólna procedura cyklizacji chemicznej (PORA)

Rozcieńczony roztwór prekursorowego EDA-RNA (30–40 µg w 490 µl, ~140 µM) zmieszano z buforem (70 µl, 800 mM KH₂PO₄, 200 mM NaCl, pH 7.0), komplementarnym oligonukleotydem (opcjonalnie, ~11 µl, 10 µM, 1.5 eq) i inkubowano w 65°C przez 5 min. Następnie, gorący roztwór przeniesiono na blat, gdzie stygł do temperatury pokojowej. Po 10–15 min w temperaturze pokojowej dodano świeży roztwór nadjodanu sodu (70 µl, 10 mM), a mieszaninę reakcyjną natychmiast umieszczono w ciemności i inkubowano w 25°C. Po 30 min dodano świeży roztwór cyjanoborowodoru sodu (70 µl, 200 mM), a inkubację kontynuowano przez kolejne 90 min. Surowy produkt circRNA izolowano z mieszaniny reakcyjnej przy użyciu zestawu Monarch RNA cleanup kit (50 µg, NEB). Wyizolowane RNA (80-99% ilości wyjściowej) rozdzielano przy użyciu RP-HPLC:

kolumna PolymerX RP-1 3 μm 100 $\text{\AA}$, 4.1 x 150 mm; przepływ 1.0 ml/min w 60°C; rozpuszczalnik A: 100 mM HAA pH 7.0; rozpuszczalnik B: 100 mM HAA pH 7 w 75% acetonitrylu; elucja: od 55% do 60% B w 25 min, następnie od 60% do 100% B w 2 min, potem 100% B przez 2 min. Eluat HPLC zawierający pożądany produkt circRNA zatężono poprzez strącanie w układzie NaOAc/iPrOH (patrz sekcja izolacja znakowanego RNA). Wysuszony pelet zawieszano w wodzie, aby uzyskać produkt circRNA (5–10 μg , 15–25% ilości początkowej). W celu dodatkowego oczyszczenia i usunięcia liniowego RNA, produkt poddano działaniu RNazy R, jak opisano poniżej (patrz sekcja Trawienie RNazą R).

Ogólna procedura cyklizacji enzymatycznej (ligaza RNA I T4)

Rozcieńczony roztwór prekursorowego p-RNA (10–15 μg w 95 μl , ~150 μM) zmieszano z buforem (5 μl , 100 mM Tris-HCl, 500 mM NaCl, pH 7.0) i inkubowano w 65°C przez 5 min. Następnie, gorący roztwór przeniesiono na blat, gdzie stygł do temperatury pokojowej. Po 10–15 min w temperaturze pokojowej, dodano bufor (15 μl , 500 mM Tris-HCl, 100 mM MgCl_2 , 10 mM DTT, pH 7.5, NEB), PEG 8000 (15 μl , 50%, NEB), ATP (0.75 μl , 10 mM, NEB), RiboLock (3.75 μl , 40 U/ μl , Thermo) oraz ligazę RNA I T4 (15 μl , 10 U/ μl , NEB). Mieszaninę reakcyjną inkubowano w 25°C przez 150 min. Surowy produkt circRNA wyizolowano z mieszaniny reakcyjnej przy użyciu zestawu Monarch RNA cleanup kit (10 μg , NEB). Wyizolowane RNA rozdzielano przy użyciu RP-HPLC: kolumna PolymerX RP-1 3 μm 100 $\text{\AA}$, 4.1 x 150 mm; przepływ 1.0 ml/min w 60°C; rozpuszczalnik A: 100 mM HAA pH 7.0; rozpuszczalnik B: 100 mM HAA pH 7 w 75% acetonitrylu; elucja: od 55% do 60% B w 25 min, następnie od 60% do 100% B w 2 min, potem 100% B przez 2 min. Eluat HPLC zawierający pożądany produkt circRNA zatężono poprzez strącanie w układzie NaOAc/iPrOH (patrz sekcja izolacja znakowanego RNA). Wysuszony pelet zawieszano w wodzie, aby uzyskać produkt circRNA (1–4 μg , 10–30% surowego produktu). W celu dodatkowego oczyszczenia i usunięcia liniowego RNA, produkt poddano działaniu RNazy R, jak opisano poniżej (patrz sekcja Trawienie RNazą R).

Ogólna procedura cyklizacji enzymatycznej (ligaza RNA II T4)

Rozcieńczony roztwór prekursorowego p-RNA (30–40 μg w 220 μl , ~320 μM) zmieszano z komplementarnym oligonukleotydem (~11 μl , 10 μM , 1.5 eq) i inkubowano w 65°C przez 5 min. Następnie, gorący roztwór przeniesiono na blat, gdzie stygł do temperatury pokojowej. Po 10–15 min w temperaturze pokojowej, dodano bufor (35 μl , 500 mM Tris-HCl, 20 mM MgCl_2 , 10 mM DTT, 4 mM ATP, pH 7.5, NEB), PEG 8000 (70 μl , 50%, NEB), RiboLock (8.75 μl , 40 U/ μl , Thermo) oraz dodano ligazę RNA II T4 (3.5 μl , 10 U/ μl , NEB). Mieszaninę reakcyjną inkubowano w 25°C przez 150 min. Surowy produkt circRNA wyizolowano z mieszaniny reakcyjnej przy użyciu zestawu Monarch RNA cleanup kit (500 μg , NEB). Surowy produkt circRNA wyizolowano z mieszaniny reakcyjnej przy użyciu zestawu Monarch RNA cleanup kit (10 μg , NEB). Wyizolowane RNA rozdzielano przy użyciu RP-HPLC: kolumna PolymerX RP-1 3 μm 100 $\text{\AA}$, 4.1 x 150 mm; przepływ 1.0 ml/min w 60°C; rozpuszczalnik A: 100 mM HAA pH 7.0; rozpuszczalnik B: 100 mM HAA pH 7 w 75% acetonitrylu; elucja: od 55% do 60% B w 25 min, następnie od 60% do 100% B w 2 min, potem 100% B przez 2 min. Eluat HPLC zawierający pożądany produkt circRNA zatężono poprzez strącanie w układzie NaOAc/iPrOH (patrz sekcja izolacja znakowanego RNA). Wysuszony pelet zawieszano w wodzie, aby uzyskać produkt circRNA (5–10 μg , 15–25% ilości początkowej). W celu

dotatkowego oczyszczenia i usunięcia liniowego RNA, produkt poddano działaniu RNazy R, jak opisano poniżej (patrz sekcja Trawienie RNazą R).

Trawienie RNazą R

Roztwór RNA (43 μ l, 10 μ g) inkubowano w 65°C przez 5 min i następnie gwałtownie schłodzono w lodzie. Po 5 min w temperaturze 0–4°C dodawano bufor (5 μ l, 200 mM Tris-HCl, 1 M KCl, 1 mM MgCl₂, pH 7.5, ABM), RiboLock (1.25 μ l, 40 U/ μ l, Thermo) oraz RNazę R (1 μ l, 10 U/ μ l, ABM). Po 30 min inkubacji w 37°C produkty trawienia były analizowane bezpośrednio metodą PAGE lub izolowane z mieszaniny reakcyjnej przy użyciu zestawu Monarch RNA cleanup kit (10 μ g, NEB).

Test endonukleolityczny z RNazą H

Roztwór (8.5 μ l) zawierający circRNA2 (100 ng), oligonukleotyd DNA (0 lub 150 nM, CCTTCAGCTCGATGCGTTCA) oraz bufor (20 mM Tris-HCl, 10 mM KCl, 8 mM MgCl₂, 1 mM DTT, Thermo) inkubowano w 65°C przez 5 min, a następnie stopniowo schłodzono do 37°C przez 30 min. Następnie, dodawano RiboLock (0.5 μ l, 40 U/ μ l, Thermo) oraz RNazę H (1 μ l, 5 U/ μ l, Thermo). Po 30 min inkubacji w 37°C produkty trawienia analizowano bezpośrednio metodą PAGE.

Elektroforeza circRNA na żelu poliakrylamidowym (PAGE)

Żel został przygotowany poprzez: przygotowanie mieszaniny zawierającej akrylamid (AA, końcowe stężenie w zakresie 6–20%), metylenobisakrylamid (MBAA, 1:19 lub 1:38 w stosunku do AA), mocznik (7M) i TBE x1 (89 mM Tris, 89 mM kwas borowy, 2 mM EDTA); zainicjowanie polimeryzacji poprzez dodatek APS (końcowo 1%) oraz TEMED (końcowo 0.4–0.8 %); polimeryzację żelu w formie o wymiarach 17 cm x 10 cm x 1 mm przez minimum godzinę w temperaturze pokojowej. Rozdział przeprowadzono w aparacie do elektroforezy pionowej z zastosowaniem prądu o stałej mocy (15 W) w buforze TBE x1. Przed nałożeniem, próbki RNA były denaturowane termicznie (5 min w 60 °C) w buforze obciążającym (końcowo ~47% formamid, 2.5 mM EDTA, 0.05% SDS, ~0.13% błękitu bromofenolowego i cyjanolu ksylenu).

Sondy molekularne RNA

Sondy FRET

Temperatury topnienia: Roztwór sondy FRET rozcieńczono buforem (80 mM KH₂PO₄, 20 mM NaCl, pH 7.0, objętość końcowa 100 μ l) do uzyskania końcowego stężenia sondy 100 nM i przeniesiono do kwarcowej kuwety fluorescencyjnej (1 x 1 x 350 mm). Opcjonalnie dodano roztwór komplementarnego oligonukleotydu DNA (2 μ l, 10 μ M). Mieszaninę ogrzewano do temperatury od 20 do 90°C w ciągu 8 min i stopniowo schładzano do 20°C lub 4°C w ciągu 20–30 min. Widma emisyjne rejestrowano na spektrofлуorometrze Cary Eclipse (Agilent), wyposażonym w lampę ksenonową, akcesorium single cell peltier (spvf 1x0, Agilent), wzbudzenie przy 500 nm, emisja 510–800 nm, szczelina 10 mm.

RNazy R, T1 i A: Bufor reakcyjny (4 mM Tris-HCl pH 7.5, 15 mM NaCl, 0.1 mM EDTA) odgazowano pod zmniejszonym ciśnieniem. Następnie, roztwór sondy FRET rozcieńczono buforem do uzyskania końcowego stężenia sondy 100 nM. Roztwór (50 μ l) podgrzano i powoli schodzono (od 95 do 25°C w 1 h, gradient schodkowy -5°C / ~4 min), a następnie

inkubowano na lodzie w bez dostępu do światła. Po rozcieńczeniu odgazowanym buforem (150 μ l) lub buforem z inhibitorem RNaz RiboLock (Thermo, 10 μ l + 140 μ l) roztwór umieszczono w kuwecie kwarcowej (1x1x350 mm) i zarejestrowano widmo fluorescencyjne (wzbudzenie 500 nm, zakres 510–850 nm, uśrednienie z trzech widm, szczelina 10 nm). Mierzono zmiany widma emisyjnego w temperaturze 5°C. Po ustabilizowaniu układu (5–15 min) dodawano enzym w odpowiednim stężeniu:

RNazę A (Thermo): Stock 10 mg/ml, do kuwety dodano 1 μ l milionkrotnie rozcieńczonego (~10 ng/ml, H₂O) enzymu

RNazę T1 (Thermo): Stock 1000 U/ μ l, do kuwety dodano 1 μ l stukrotnie rozcieńczonego (10 U/ μ l, H₂O) enzymu

RNazę R (ABM): Stock 10 U/ μ l, do kuwety dodano 1 μ l enzymu

Następnie mierzono zmiany widma emisyjnego w temperaturze 5°C w czasie postępu reakcji.

hDcp1/2: Bufor reakcyjny (4 mM Tris-HCl pH 7.5, 15 mM NaCl, 0.1 mM EDTA) odgazowano pod zmniejszonym ciśnieniem. Następnie, roztwór sondy FRET (Cy5-ArcaRNA00-Cy3) rozcieńczono buforem do uzyskania końcowego stężenia sondy 100 nM. Roztwór (40 μ l) został podgrzany i powoli schodzony (od 95 do 25°C w 1 h, gradient schodkowy -5°C / ~4 min) a następnie inkubowany na lodzie w bez dostępu do światła. Po rozcieńczeniu odgazowanym buforem (150 μ l) dodano MgCl₂ (1 μ l, 1 M), przeniesiono do kuwety kwarcowej (1x1x350 mm) i zarejestrowano widmo fluorescencyjne (wzbudzenie 500 nm, zakres 510-850 nm, uśrednienie z trzech widm, szczelina 10 nm). Mierzono zmiany widma emisyjnego w temperaturze 5°C. Po ustabilizowaniu układu (5–15 min) dodano kompleks Dcp1/2 z *Schizosaccharomyces pombe* (10 μ l, 7 μ M).^[221] Następnie mierzono zmiany widma emisyjnego w temperaturze 5°C w czasie postępu reakcji.

RNaza H: Do roztworu sondy FRET (Cy5-ArcaRNA02-Cy3, ~100 nM, 50 μ l) w buforze (412 μ l; 4 mM Tris-HCl pH 7.5, 15 mM NaCl, 0.1 mM EDTA) dodano roztwór DNA (6.00 μ l, 1 μ M, 1.2 eq, o sekwencji TGCTCGGGGTCGTACACCTT, TCATTTGCTTGACGAGGCC lub CGTGATATCTCTCCCGTGCCTCCACAGGTA) lub wodę (6.00 μ l), podgrzano i powoli schłodzono (od 95 do 25°C w 1 h, gradient schodkowy -5°C / ~4 min). Następnie dodano odgazowaną wodę (160 μ l), bufor (24 μ l, 200 mM Tris-HCl pH 7.5, 500 mM NaCl, 100 mM MgCl₂, 10 mM DTT) i umieszczano w kuwecie kwarcowej (240 μ l, 1x1x350 mm). Mierzono zmiany widma emisyjnego w temperaturze 35°C (wzbudzenie 500 nm, zakres 510-850 nm, uśrednienie z trzech widm, szczelina 10 nm). Po ustabilizowaniu układu (2–5 min) dodano RNazę H (2.00 μ l, 0.1 mg/ml)^[212] mierzono zmiany widma emisyjnego w temperaturze w czasie postępu reakcji.

CNOT7: 60 nanogramów sondy RNA (N₃-ArcaRNA04, N₃- ArcaRNA04-Cy3, Cy5-ArcaRNA04 lub Cy5- ArcaRNA04-Cy3) rozpuszczano w buforze (10 mM tris-HCl pH 8, 5 mM MgCl₂, 10 mM DTT, 50 mM KCl) i inkubowano w obecności deadenylazy CNOT7 (0.2 mg/ml)^[92] w 37°C przez 30 min (końcowa objętość reakcji 5 μ l). Reakcja była wygaszana przez dodanie buforu obciążającego (5 μ l, 95% formamid, 5 mM EDTA, 0.1% SDS, ~0.25% błękitu bromofenolowego i cyjanolu ksylenu) i denaturację termiczną (5 min 65°C). Produkty były rozdzielane przy użyciu 8% denaturującego PAGE.

Sonda pH

Rozcieńczono próbkę sondę Cy5-ArcaRNA03-FAM w wodzie (1400 ng / 1090 µl) i przygotowano serię buforów cytrynianowo-fosforanowych (pH 3.0–8.0, 0.1–0.2 M). Zmierzono widma emisyjne dziesięciokrotnie rozcieńczonych buforów oraz roztworów sondy w różnych pH (180 µl RNA + 20 µl buforu) przy wzbudzeniu 488 nm i 639 nm w 36 °C (zakresy 505–800 nm i 655–800 nm, uśrednienie z trzech widm, napięcie 800V, filtr ex „auto”, filtr em „open”, szczelina 10 nm).

Biotynylowana sonda powinowactwa i EMSA

Sondę Cy5-RNA06-Biot (30 ng w 4.0 µl wody) inkubowano przez 5 min w 65°C i następnie chłodzono na lodzie. Następnie, dodawano streptawidynę (1.0 µl, 2.5 µM, Promega) lub wodę (1.00 µl). Po 20 min inkubacji na lodzie dodawano glicerol (2.0 µl, 80%) i przeprowadzono elektroforezę na żelu agarozowym (1.0%, 1x TBE, 80 V, 60 min). Po wizualizacji fluorescencji za pomocą aparatu Typhoon FLA 9500, przeprowadzono barwienie kwasów nukleinowych za pomocą bromku etydyyny (Sigma).

Translacja *in vitro*

Translacja znakowanego mRNA

Mieszanina reakcyjna (8 µl) zawierała lizat z retikulocytów (4 µl, RRL, Promega), mieszaninę aminokwasów bez leucyny (25 µM, Promega), mieszaninę aminokwasów bez metioniny (25 µM, Promega), octan potasu (250 mM), oraz MgCl₂ (1.25 mM). Po godzinie inkubacji w 30 °C, dodano roztwór mRNA kodującego lucyferazę *Renilla* (2 µl, 0.64, 0.51, 0.41, 0.33, 0.16 lub 0.00 ng/l) i kontynuowano inkubację mieszaniny reakcyjnej w temperaturze 30 °C przez godzinę. Reakcję zatrzymano przez zamrożenie w ciekłym azocie. Następnie, mieszaniny reakcyjne przeniesiono na płytkę 96-dółkową (białą, płaskodenną), płytkę zwirowano (1 min, 1000 rcf, 22°C) i umieszczono w czytniku płytek Synergy H1 (BioTek). Do każdego dołka dodano roztwór h-coelenterazyny (50 µl, 2 µM, NanoLight) w PBS i odczytano luminescencję (25°C, gain 240). Luminescencja została wykreślona jako funkcja stężenia mRNA i analizowana za pomocą metody regresji liniowej. Współczynnik regresji liniowej (nachylenie) dla każdego powtórzenia został normalizowany do wartości wyznaczonej dla Arca-RNA07, aby uzyskać parametr względnego poziomu białka. Wyniki z trzech niezależnych eksperymentów uśredniono podając przy tym błąd średniej (SEM).

Translacja kolistego RNA

Przygotowano mieszaninę reakcyjną: do wody (82.5 µl) dodano KOAc (137.5 µl, 1 M), MgCl₂ (27.5 µl, 25 mM), mieszaninę aminokwasów bez leucyny (13.75 µl, 1 mM, Promega), mieszaninę aminokwasów bez metioniny (13.75 µl, 1 mM, Promega) oraz Rabbit Reticulocyte Lysate (275 µl, Promega). Po godzinie inkubacji w 30°C mieszaninę podzielono na porcje (po 8.0 µl, płytka 96-dółkowa, przeźroczysta, V-dno), do których następnie dodano roztwór RNA kodującego lucyferazę *Gaussia* (2.0 µl, 1.14 nM, 1-2 ng). Mieszaniny reakcyjne inkubowano przez godzinę w 30°C, po czym zatrzymywano reakcje zamrażając je w ciekłym azocie. Seria reakcji zawierała wszystkie badane typy RNA w trzech powtórzeniach oraz kontrolę bez dodatku RNA (próba ślepa). Następnie, płytkę zwirowano (1 min, 1000 rcf, 22°C) i umieszczono w czytniku płytek Synergy H1 (BioTek). Do każdego dołka dodano roztwór h-coelenterazyny (50 µl, 2 µM, NanoLight) w PBS i odczytano luminescencję (25°C, gain 240). Wyniki z trzech powtórzeń technicznych

uśredniono i podzielono przez średnią wartość uzyskaną dla Cap1-RNA10, podając przy tym błąd średniej (SEM).

Bibliografia:

- [1] F. Guo, Z. Yue, M. Trajkovski, X. Zhou, D. Cao, Q. Li, B. Wang, X. Wen, J. Plavec, Q. Peng, Z. Xi, C. Zhou, *Journal of the American Chemical Society* **2018**, *140*, 11893-11897.
- [2] J. R. Casey, S. Grinstein, J. Orlowski, *Nature Reviews Molecular Cell Biology* **2010**, *11*, 50-61.
- [3] K. W. Knouse, D. T. Flood, J. C. Vantourout, M. A. Schmidt, I. M. McDonald, M. D. Eastgate, P. S. Baran, *ACS Central Science* **2021**, *7*, 1473-1485.
- [4] E. L. Jones, A. J. Mlotkowski, S. P. Hebert, H. B. Schlegel, C. S. Chow, *The Journal of Physical Chemistry A* **2022**, *126*, 1518-1529.
- [5] Q. Vicens, J. S. Kieft, *Proceedings of the National Academy of Sciences* **2022**, *119*.
- [6] P. Thaplyal, P. C. Bevilacqua, in *Methods in Enzymology*, Elsevier, **2014**, pp. 189-219.
- [7] E. A. Dethoff, J. Chugh, A. M. Mustoe, H. M. Al-Hashimi, *Nature* **2012**, *482*, 322-330.
- [8] L. R. Ganser, M. L. Kelly, D. Herschlag, H. M. Al-Hashimi, *Nature Reviews Molecular Cell Biology* **2019**, *20*, 474-489.
- [9] D. N. Ermolenko, D. H. Mathews, *Wiley Interdisciplinary Reviews: RNA* **2021**, *12*, e1611.
- [10] N. Safaei, A. M. Noronha, D. Rodionov, G. Kozlov, C. J. Wilds, G. M. Sheldrick, K. Gehring, *Angewandte Chemie International Edition* **2013**, *52*, 10370-10373.
- [11] N. B. Leontis, E. Westhof, *RNA* **2001**, *7*, 499-512.
- [12] S. E. Butcher, A. M. Pyle, *Accounts of Chemical Research* **2011**, *44*, 1302-1311; P. P. Seelam, P. Sharma, A. Mitra, *RNA* **2017**, *23*, 847-859.
- [13] E. Rivas, *RNA Biology* **2013**, *10*, 1185-1196; M. Popenda, M. Szachniuk, M. Antczak, K. J. Purzycka, P. Lukasiak, N. Bartol, J. Blazewicz, R. W. Adamiak, *Nucleic Acids Research* **2012**, *40*, e112-e112; B. A. Sweeney, D. Hoksza, E. P. Nawrocki, C. E. Ribas, F. Madeira, J. J. Cannone, R. Gutell, A. Maddala, C. D. Meade, L. D. Williams, A. S. Petrov, P. P. Chan, T. M. Lowe, R. D. Finn, A. I. Petrov, *Nature Communications* **2021**, *12*.
- [14] A. R. Gruber, R. Lorenz, S. H. Bernhart, R. Neuböck, I. L. Hofacker, *Nucleic Acids Research* **2008**, *36*, W70-W74.
- [15] K. J. Doshi, J. J. Cannone, C. W. Cobough, R. R. Gutell, *BMC Bioinformatics* **2004**, *5*, 105.
- [16] P. P. Amaral, M. E. Dinger, T. R. Mercer, J. S. Mattick, *Science* **2008**, *319*, 1787-1789; P. G. Higgs, N. Lehman, *Nature Reviews Genetics* **2015**, *16*, 7-17; E. Schmitt, P.-D. Coureux, R. Kazan, G. Bourgeois, C. Lazennec-Schurdevin, Y. Mechulam, *Frontiers in Microbiology* **2020**, *11*.
- [17] D. G. Calvopina-Chavez, M. A. Gardner, J. S. Griffiths, *G3 Genes|Genomes|Genetics* **2022**, *12*, jkac070.
- [18] M. Girbig, A. D. Misiaszek, C. W. Müller, *Nature Reviews Molecular Cell Biology* **2022**, *23*, 603-622.
- [19] T. A. Steitz, *Current Opinion in Structural Biology* **2009**, *19*, 683-690.
- [20] J. J. Dunn, F. W. Studier, M. Gottesman, *Journal of Molecular Biology* **1983**, *166*, 477-535.
- [21] R. Benoni, M. Culka, O. Hudeček, L. Gahurova, H. Cahová, *ACS Chemical Biology* **2020**, *15*, 1765-1772.

- [22] Y. W. Yin, T. A. Steitz, *Science* **2002**, 298, 1387-1395; K. J. Durniak, S. Bailey, T. A. Steitz, *Science* **2008**, 322, 553-557.
- [23] G.-Q. Tang, R. P. Bandwar, S. S. Patel, *Journal of Biological Chemistry* **2005**, 280, 40707-40713.
- [24] T. Conrad, I. Plumbom, M. Alcobendas, R. Vidal, S. Sauer, *Communications Biology* **2020**, 3.
- [25] V. S. Anand, S. S. Patel, *Journal of Biological Chemistry* **2006**, 281, 35677-35685; P. Thomen, P. J. Lopez, U. Bockelmann, J. Guillerez, M. Dreyfus, F. Heslot, *Biophysical Journal* **2008**, 95, 2423-2433.
- [26] K. Severinov, *Proceedings of the National Academy of Sciences* **2001**, 98, 5-7.
- [27] F. J. Triana-Alonso, M. Dabrowski, J. Wadzack, K. H. Nierhaus, *Journal of Biological Chemistry* **1995**, 270, 6298-6307; Y. Gholamalipour, A. Karunanayake Mudiyansele, C. T. Martin, *Nucleic Acids Research* **2018**, 46, 9253-9263; X. Mu, E. Greenwald, S. Ahmad, S. Hur, *Nucleic Acids Research* **2018**, 46, 5239-5249.
- [28] M. Z. Wu, H. Asahara, G. Tzertzinis, B. Roy, *RNA* **2020**, 26, 345-360.
- [29] M. Warminski, P. J. Sikorski, J. Kowalska, J. Jemielity, *Topics in Current Chemistry* **2017**, 375.
- [30] P. J. Sikorski, M. Warminski, D. Kubacka, T. Ratajczak, D. Nowis, J. Kowalska, J. Jemielity, *Nucleic acids research* **2020**, 48, 1607-1626.
- [31] A. Sugita, S. Kuruma, N. Yanagisawa, H. Ishiguro, R. Kano, Y. Ohkuma, Y. Hirose, *The Journal of Biochemistry* **2021**, 170, 203-213.
- [32] M. Warminski, P. J. Sikorski, J. Kowalska, J. Jemielity, *Topics in Current Chemistry* **2017**, 375, 16.
- [33] B. Roy, in *Methods in Molecular Biology*, Springer US, **2021**, pp. 327-356; M. Zhang, J. Song, W. Yuan, W. Zhang, Z. Sun, *Frontiers in Immunology* **2021**, 12; K. Dratzkowska, R. Tomecki, M. Warminski, N. Baran, D. Cysewski, A. Depaix, R. Kasprzyk, J. Kowalska, J. Jemielity, J. Sikorski, Pawel, *Nucleic Acids Research* **2022**, 50, 9051-9071; L. Cui, R. Ma, J. Cai, C. Guo, Z. Chen, L. Yao, Y. Wang, R. Fan, X. Wang, Y. Shi, *Signal Transduction and Targeted Therapy* **2022**, 7.
- [34] P. Morais, H. Adachi, Y.-T. Yu, *Frontiers in Genetics* **2021**, 12.
- [35] Z. Tang, J. Zhao, Z. J. Pearson, Z. V. Boskovic, J. Wang, *Molecules* **2021**, 26, 2263.
- [36] J. Ouyang, Y. Zhong, Y. Zhang, L. Yang, P. Wu, X. Hou, F. Xiong, X. Li, S. Zhang, Z. Gong, Y. He, Y. Tang, W. Zhang, B. Xiang, M. Zhou, J. Ma, Y. Li, G. Li, Z. Zeng, C. Guo, W. Xiong, *British Journal of Cancer* **2022**, 126, 1113-1124.
- [37] M. Osella, M. Caselle, *Physical Biology* **2009**, 6, 046018.
- [38] J. Nourse, S. Spada, S. Danckwardt, *Biomolecules* **2020**, 10, 915.
- [39] L. A. Passmore, J. Collier, *Nature Reviews Molecular Cell Biology* **2022**, 23, 93-106.
- [40] B. Tian, J. L. Manley, *Nature Reviews Molecular Cell Biology* **2017**, 18, 18-30.
- [41] J. Houseley, D. Tollervy, *Cell* **2009**, 136, 763-776.
- [42] E. Abernathy, B. Glaunsinger, *Virology* **2015**, 479-480, 600-608.
- [43] R. Tomecki, A. Dziembowski, *RNA* **2010**, 16, 1692-1724.
- [44] S. Kramer, A. G. McLennan, *WIREs RNA* **2019**, 10, e1511; O. Hudeček, R. Benoni, P. E. Reyes-Gutierrez, M. Culka, H. Šanderová, M. Hubálek, L. Rulíšek, J. Cvačka, L. Krásný, H. Cahová, *Nature Communications* **2020**, 11.
- [45] C. F. Bourgeois, F. Mortreux, D. Auboeuf, *Nature Reviews Molecular Cell Biology* **2016**, 17, 426-438.
- [46] T. E. Mullen, W. F. Marzluff, *Genes & Development* **2008**, 22, 50-65.
- [47] M. Schmid, T. H. Jensen, *Trends in Biochemical Sciences* **2008**, 33, 501-510.

- [48] Y. Luo, Z. Na, S. A. Slavoff, *Biochemistry* **2018**, 57, 2424-2431.
- [49] M. A. Valencia-Sanchez, J. Liu, G. J. Hannon, R. Parker, *Genes & Development* **2006**, 20, 515-524.
- [50] A.-C. Prats, F. David, L. H. Diallo, E. Roussel, F. Tatin, B. Garmy-Susini, E. Lacazette, *International Journal of Molecular Sciences* **2020**, 21, 8591.
- [51] Y. Liang, N. Liu, L. Yang, J. Tang, Y. Wang, M. Mei, *Current Genomics* **2021**, 22, 485-495.
- [52] C.-X. Liu, L.-L. Chen, *Cell* **2022**, 185, 2016-2034.
- [53] A. T. He, J. Liu, F. Li, B. B. Yang, *Signal Transduction and Targeted Therapy* **2021**, 6.
- [54] E. C. S. Lee, S. A. M. Elhassan, G. P. L. Lim, W. H. Kok, S. W. Tan, E. N. Leong, S. H. Tan, E. W. L. Chan, S. K. Bhattamisra, R. Rajendran, M. Candasamy, *Biomedicine & Pharmacotherapy* **2019**, 111, 198-208; Q. Chen, J. Li, P. Shen, H. Yuan, J. Yin, W. Ge, W. Wang, G. Chen, T. Yang, B. Xiao, Y. Miao, Z. Lu, P. Wu, K. Jiang, *Cell & Bioscience* **2022**, 12.
- [55] R. Chen, S. K. Wang, J. A. Belk, L. Amaya, Z. Li, A. Cardenas, B. T. Abe, C.-K. Chen, P. A. Wender, H. Y. Chang, *Nature Biotechnology* **2022**.
- [56] R. A. Wesselhoeft, P. S. Kowalski, D. G. Anderson, *Nature Communications* **2018**, 9.
- [57] R. A. Wesselhoeft, P. S. Kowalski, F. C. Parker-Hale, Y. Huang, N. Bisaria, D. G. Anderson, *Molecular Cell* **2019**, 74, 508-520.e504.
- [58] L. Qu, Z. Yi, Y. Shen, L. Lin, F. Chen, Y. Xu, Z. Wu, H. Tang, X. Zhang, F. Tian, C. Wang, X. Xiao, X. Dong, L. Guo, S. Lu, C. Yang, C. Tang, Y. Yang, W. Yu, J. Wang, Y. Zhou, Q. Huang, A. Yisimayi, S. Liu, W. Huang, Y. Cao, Y. Wang, Z. Zhou, X. Peng, J. Wang, X. S. Xie, W. Wei, *Cell* **2022**, 185, 1728-1744.e1716.
- [59] Y. G. Chen, M. V. Kim, X. Chen, P. J. Batista, S. Aoyama, J. E. Wilusz, A. Iwasaki, H. Y. Chang, *Molecular Cell* **2017**, 67, 228-238.e225; J. Wang, K. Liu, Y. Liu, Q. Lv, F. Zhang, H. Wang, *Oncotarget* **2017**, 8, 110914-110921.
- [60] E. V. Koonin, A. S. Novozhilov, *IUBMB Life* **2009**, 61, 99-111; D. de la Torre, J. W. Chin, *Nature Reviews Genetics* **2021**, 22, 169-184.
- [61] A. Fukao, T. Tomohiro, T. Fujiwara, *Cells* **2021**, 10, 1711.
- [62] H. F. Noller, L. Lancaster, J. Zhou, S. Mohan, *Nature Structural & Molecular Biology* **2017**, 24, 1021-1027; B. Xu, L. Liu, G. Song, *Frontiers in Molecular Biosciences* **2022**, 8; C. U. T. Hellen, *Cold Spring Harbor Perspectives in Biology* **2018**, 10, a032656.
- [63] X. Jiang, B. Liu, Z. Nie, L. Duan, Q. Xiong, Z. Jin, C. Yang, Y. Chen, *Signal Transduction and Targeted Therapy* **2021**, 6.
- [64] R. Marques, R. Lacerda, L. Romão, *Biomedicines* **2022**, 10, 1865.
- [65] A.-C. Godet, F. David, F. Hantelys, F. Tatin, E. Lacazette, B. Garmy-Susini, A.-C. Prats, *International Journal of Molecular Sciences* **2019**, 20, 924; E. Martinez-Salas, R. Francisco-Velilla, J. Fernandez-Chamorro, A. M. Embarek, *Frontiers in Microbiology* **2018**, 8; D. Koirala, Y. Shao, Y. Koldobskaya, J. R. Fuller, A. M. Watkins, S. A. Shelke, E. V. Pilipenko, R. Das, P. A. Rice, J. A. Piccirilli, *Nature Communications* **2019**, 10.
- [66] Y. Yang, X. Fan, M. Mao, X. Song, P. Wu, Y. Zhang, Y. Jin, Y. Yang, L.-L. Chen, Y. Wang, C. C. Wong, X. Xiao, Z. Wang, *Cell Research* **2017**, 27, 626-641; M. Lei, G. Zheng, Q. Ning, J. Zheng, D. Dong, *Molecular Cancer* **2020**, 19.
- [67] X. Fan, Y. Yang, C. Chen, Z. Wang, *Nature Communications* **2022**, 13.
- [68] J. S. Sherkow, P. J. Zettler, H. T. Greely, *Journal of Law and the Biosciences* **2018**, 5, 786-793.

- [69] K. Bulaklak, C. A. Gersbach, *Nature Communications* **2020**, *11*.
- [70] T. C. Roberts, R. Langer, M. J. A. Wood, *Nature Reviews Drug Discovery* **2020**, *19*, 673-694.
- [71] E. Papanikolaou, A. Bosio, *Frontiers in Genome Editing* **2021**, *3*.
- [72] E. Rohner, R. Yang, K. S. Foo, A. Goedel, K. R. Chien, *Nature Biotechnology* **2022**, *40*, 1586-1600.
- [73] S. Qin, X. Tang, Y. Chen, K. Chen, N. Fan, W. Xiao, Q. Zheng, G. Li, Y. Teng, M. Wu, X. Song, *Signal Transduction and Targeted Therapy* **2022**, *7*.
- [74] Gene Therapy Market Size, Share & Growth Report, 2030 (grandviewresearch.com); mRNA Therapeutics Market Size & Share Report, 2022-2030 (grandviewresearch.com); Global Oligonucleotide Synthesis Market Size Report, 2025 (grandviewresearch.com).
- [75] L. Martin, M. Hutchens, C. Hawkins, A. Radnov, *Nature Reviews Drug Discovery* **2017**, *16*, 381-382.
- [76] T. F. Broekhoff, C. C. G. Sweegers, E. M. Krijkamp, A. K. Mantel-Teeuwisse, H. G. M. Leufkens, W. G. Goettsch, R. A. Vreman, *Value in Health* **2021**, *24*, 759-769; M. Nuijten, *Journal of Market Access & Health Policy* **2022**, *10*; S. H. S. Chan, C. K. H. Wong, T. Wu, W. Wong, M. K. L. Yu, I. C. H. Au, G. C. F. Chan, *The European Journal of Health Economics* **2022**.
- [77] M. Charman, C. Herrmann, M. D. Weitzman, *FEBS Letters* **2019**, *593*, 3531-3550; J. E. Vargas, L. Chicaybam, R. T. Stein, A. Tanuri, A. Delgado-Cañedo, M. H. Bonamino, *Journal of Translational Medicine* **2016**, *14*.
- [78] M. E. Maes, G. Colombo, R. Schulz, S. Siegert, *Neuroscience Letters* **2019**, *707*, 134310.
- [79] S. Nidhi, U. Anand, P. Oleksak, P. Tripathi, J. A. Lal, G. Thomas, K. Kuca, V. Tripathi, *International Journal of Molecular Sciences* **2021**, *22*, 3327.
- [80] M. Chavez, X. Chen, P. B. Finn, L. S. Qi, *Nature Reviews Nephrology* **2023**, *19*, 9-22.
- [81] U. Sahin, K. Karikó, Ö. Türeci, *Nature Reviews Drug Discovery* **2014**, *13*, 759-780.
- [82] X. Hou, T. Zaks, R. Langer, Y. Dong, *Nature Reviews Materials* **2021**, *6*, 1078-1094.
- [83] C. Martin, D. Lowery, *Nature Reviews Drug Discovery* **2020**, *19*, 578-578; M. Gaviria, B. Kilic, *Nature Biotechnology* **2021**, *39*, 546-548.
- [84] S. S. Rosa, D. Nunes, L. Antunes, D. M. F. Prazeres, M. P. C. Marques, A. M. Azevedo, *Biotechnology and Bioengineering* **2022**, *119*, 3127-3139.
- [85] T. Akaishi, T. Onodera, T. Takahashi, H. Harigae, T. Ishii, *Scientific Reports* **2022**, *12*.
- [86] L. Jia, S.-B. Qian, *Accounts of Chemical Research* **2021**, *54*, 4272-4282.
- [87] D. Strzelecka, M. Smietanski, P. J. Sikorski, M. Warminski, J. Kowalska, J. Jemielity, *RNA* **2020**, *26*, 1815-1837.
- [88] M. K. Franco, K. S. Koutmou, *Biophysical Chemistry* **2022**, *285*, 106780.
- [89] H. Bae, J. Collier, *Molecular Cell* **2022**, *82*, 1467-1476; G. Hanson, J. Collier, *Nature Reviews Molecular Cell Biology* **2018**, *19*, 20-30; H. Fu, Y. Liang, X. Zhong, Z. Pan, L. Huang, H. Zhang, Y. Xu, W. Zhou, Z. Liu, *Scientific Reports* **2020**, *10*.
- [90] K. H. Asrani, J. D. Farelli, M. R. Stahley, R. L. Miller, C. J. Cheng, R. R. Subramanian, J. M. Brown, *RNA Biology* **2018**, *1-7*; K. Suknuntha, L. Tao, V. Brok-Volchanskaya, S. S. D'Souza, A. Kumar, I. Slukvin, *Stem Cell Reviews and Reports* **2018**, *14*, 525-534; N. Sultana, Y. Hadas, M. T. K. Sharkar, K. Kaur, A. Magadum, A. A. Kurian, N. Hossain, B. Alburquerque, S. Ahmed, E. Chepurko, L. Zangi, *Molecular Therapy - Methods & Clinical Development* **2020**, *17*, 622-633; P.

- Eisenhut, A. Mebrahtu, M. Moradi Barzadd, N. Thalén, G. Klanert, M. Weinguny, A. Sandegren, C. Su, D. Hatton, N. Borth, J. Rockberg, *Nucleic Acids Research* **2020**, *48*, e119-e119; K. Leppek, G. W. Byeon, W. Kladwang, H. K. Wayment-Steele, C. H. Kerr, A. F. Xu, D. S. Kim, V. V. Topkar, C. Choe, D. Rothschild, G. C. Tiu, R. Wellington-Oguri, K. Fujii, E. Sharma, A. M. Watkins, J. J. Nicol, J. Romano, B. Tunguz, E. Participants, M. Barna, R. Das, Cold Spring Harbor Laboratory, **2021**.
- [91] Z. Trepotec, J. Geiger, C. Plank, M. K. Aneja, C. Rudolph, *RNA* **2019**, *25*, 507-518.
- [92] O. Perzanowska, M. Smietanski, J. Jemielity, J. Kowalska, *Chemistry – A European Journal* **2022**, *28*.
- [93] K. Xiang, D. P. Bartel, *eLife* **2021**, *10*, e66493.
- [94] C.-K. Chen, R. Cheng, J. Demeter, J. Chen, S. Weingarten-Gabbay, L. Jiang, M. P. Snyder, J. S. Weissman, E. Segal, P. K. Jackson, H. Y. Chang, *Molecular Cell* **2021**, *81*, 4300-4318.e4313.
- [95] T. Beissert, M. Perkovic, A. Vogel, S. Erbar, K. C. Walzer, T. Hempel, S. Brill, E. Haefner, R. Becker, Ö. Türeci, U. Sahin, *Molecular Therapy* **2020**, *28*, 119-128; M. C. Ballesteros-Briones, N. Silva-Pilipich, G. Herrador-Cañete, L. Vanrell, C. Smerdou, *Current Opinion in Virology* **2020**, *44*, 145-153.
- [96] J.-C. Boulain, J. Dassa, L. Mesta, A. Savatier, N. Costa, B. H. Muller, G. L'Hostis, E. A. Stura, A. Troesch, F. Ducancel, *Protein Engineering, Design and Selection* **2013**, *26*, 725-734; D. Meyer, Kate, P. Patil, Deepak, J. Zhou, A. Zinoviev, A. Skabkin, Maxim, O. Elemento, V. Pestova, Tatyana, S.-B. Qian, R. Jaffrey, Samie, *Cell* **2015**, *163*, 999-1010; T. Egorova, E. Shuvalova, S. Mukba, A. Shuvalov, P. Kolosov, E. Alkalaeva, *International Journal of Molecular Sciences* **2021**, *22*, 5186.
- [97] A. Dousis, K. Ravichandran, E. M. Hobert, M. J. Moore, A. E. Rabideau, *Nature Biotechnology* **2022**.
- [98] J. L. Brunelle, R. Green, in *Methods in Enzymology*, Vol. 530 (Ed.: J. Lorsch), Academic Press, **2013**, pp. 101-114; J. C. Bowman, B. Azizi, T. K. Lenz, P. Roy, L. D. Williams, in *Recombinant and In Vitro RNA Synthesis: Methods and Protocols* (Ed.: G. L. Conn), Humana Press, Totowa, NJ, **2012**, pp. 19-41; W. De Mey, P. De Schrijver, D. Autaers, L. Pfitzer, B. Fant, H. Locy, A. Esprit, L. Lybaert, C. Bogaert, M. Verdonck, K. Thielemans, K. Breckpot, L. Franceschini, *Molecular Therapy - Nucleic Acids* **2022**, *29*, 943-954.
- [99] P. Arbuthnot, A. Ely, K. Bloom, *BioTechniques* **2019**, *66*, 37-39.
- [100] X. Piao, V. Yadav, E. Wang, W. Chang, L. Tau, B. E. Lindenmuth, S. X. Wang, *Molecular Therapy - Nucleic Acids* **2022**, *29*, 618-624.
- [101] S. E. Walker, J. Lorsch, in *Methods in Enzymology*, Vol. 530 (Ed.: J. Lorsch), Academic Press, **2013**, pp. 337-343.
- [102] S. S. Rosa, D. M. F. Prazeres, A. M. Azevedo, M. P. C. Marques, *Vaccine* **2021**, *39*, 2190-2200.
- [103] K. Karikó, H. Muramatsu, J. Ludwig, D. Weissman, *Nucleic Acids Research* **2011**, *39*, e142-e142.
- [104] L. E. Easton, Y. Shibata, P. J. Lukavsky, *RNA* **2010**, *16*, 647-653; S. Fekete, H. Yang, K. Wyndham, M. Lauber, *Journal of Chromatography Open* **2022**, *2*, 100031.
- [105] J. Green Mr Fau - Sambrook, J. Sambrook; B. J. Engel, B. J. Grindel, J. P. Gray, S. W. Millward, *Bioorganic & Medicinal Chemistry Letters* **2020**, *30*, 126934.
- [106] M. Baiersdörfer, G. Boros, H. Muramatsu, A. Mahiny, I. Vlatkovic, U. Sahin, K. Karikó, *Molecular Therapy - Nucleic Acids* **2019**, *15*, 26-35.

- [107] L. J. R., *Principles of Fluorescence Spectroscopy*, 3 ed., Springer New York, NY, **2006**.
- [108] J. Deal, D. J. Pleshinger, S. C. Johnson, S. J. Leavesley, T. C. Rich, *Cellular Signalling* **2020**, 75, 109769.
- [109] J. W. Lichtman, J.-A. Conchello, *Nature Methods* **2005**, 2, 910-919.
- [110] X. Li, H. Kim, J. L. Litke, J. Wu, S. R. Jaffrey, *Angewandte Chemie International Edition* **2020**, 59, 4511-4518.
- [111] X. Chen, D. Zhang, N. Su, B. Bao, X. Xie, F. Zuo, L. Yang, H. Wang, L. Jiang, Q. Lin, M. Fang, N. Li, X. Hua, Z. Chen, C. Bao, J. Xu, W. Du, L. Zhang, Y. Zhao, L. Zhu, J. Loscalzo, Y. Yang, *Nature Biotechnology* **2019**, 37, 1287-1293.
- [112] A. Koch, L. Aguilera, T. Morisaki, B. Munsky, T. J. Stasevich, *Nature Structural & Molecular Biology* **2020**, 27, 1095-1104.
- [113] H. M. T. Choi, J. Y. Chang, L. A. Trinh, J. E. Padilla, S. E. Fraser, N. A. Pierce, *Nature Biotechnology* **2010**, 28, 1208-1212.
- [114] J.-H. Spille, U. Kubitscheck, *Journal of Cell Science* **2015**, 128, 3695-3706; B. Kim, J. Seol, Y. K. Kim, J.-B. Lee, *Experimental & Molecular Medicine* **2023**.
- [115] K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, *Science* **2015**, 348, aaa6090; S. H. Rouhanifard, I. A. Mellis, M. Dunagin, S. Bayatpour, C. L. Jiang, I. Dardani, O. Symmons, B. Emert, E. Torre, A. Cote, A. Sullivan, J. A. Stamatoyannopoulos, A. Raj, *Nature Biotechnology* **2019**, 37, 84-89.
- [116] J. L. Kirschman, S. Bhosle, D. Vanover, E. L. Blanchard, K. H. Loomis, C. Zurla, K. Murray, B. C. Lam, P. J. Santangelo, *Nucleic Acids Research* **2017**, 45, e113-e113.
- [117] T. E. England, O. C. Uhlenbeck, *Nature* **1978**, 275, 560-561.
- [118] C. M. Dojahn, M. Hesse, C. Arenz, *Chemical Communications* **2013**, 49, 3128-3130.
- [119] T. Shibata, Y. Fujita, H. Ohno, Y. Suzuki, K. Hayashi, K. R. Komatsu, S. Kawasaki, K. Hidaka, S. Yonehara, H. Sugiyama, M. Endo, H. Saito, *Nature Communications* **2017**, 8, 540; E. Sonnleitner, A. Wulf, S. Campagne, X.-Y. Pei, M. T. Wolfinger, G. Forlani, K. Prindl, L. Abdou, A. Resch, F. H. T. Allain, B. F. Luisi, H. Urlaub, U. Bläsi, *Nucleic Acids Research* **2018**, 46, 1470-1485; E. L. Van Nostrand, G. A. Pratt, B. A. Yee, E. C. Wheeler, S. M. Blue, J. Mueller, S. S. Park, K. E. Garcia, C. Gelboin-Burkhart, T. B. Nguyen, I. Rabano, R. Stanton, B. Sundararaman, R. Wang, X.-D. Fu, B. R. Graveley, G. W. Yeo, *Genome Biology* **2020**, 21, 90; J. Guillén-Boixet, A. Kopach, A. S. Holehouse, S. Wittmann, M. Jahnel, R. Schlüßler, K. Kim, I. R. E. A. Trussina, J. Wang, D. Mateju, I. Poser, S. Maharana, M. Ruer-Gruß, D. Richter, X. Zhang, Y.-T. Chang, J. Guck, A. Honigsmann, J. Mahamid, A. A. Hyman, R. V. Pappu, S. Alberti, T. M. Franzmann, *Cell* **2020**, 181, 346-361.e317.
- [120] W.-J. C. Lai, M. Kayedkhordeh, E. V. Cornell, E. Farah, S. Bellaousov, R. Rietmeijer, E. Salsi, D. H. Mathews, D. N. Ermolenko, *Nature Communications* **2018**, 9, 4328.
- [121] W.-J. C. Lai, D. N. Ermolenko, *Methods* **2018**, 137, 37-48.
- [122] A. Depaix, A. Mlynarska-Cieslak, M. Warminski, P. J. Sikorski, J. Jemielity, J. Kowalska, *Chemistry – A European Journal* **2021**, 27, 12190-12197.
- [123] R. W. Tibble, A. Depaix, J. Kowalska, J. Jemielity, J. D. Gross, *Nature Chemical Biology* **2021**, 17, 615-623.
- [124] Y. Kinoshita, Y. Nishigaki K Fau - Husimi, Y. Husimi.
- [125] F. Huang, T. Wang G Fau - Coleman, N. Coleman T Fau - Li, N. Li; F. Huang, J. He, Y. Zhang, Y. Guo, *Nature Protocols* **2008**, 3, 1848-1861.

- [126] J. Schoch, S. Ameta, A. Jäschke, *Chemical Communications* **2011**, 47, 12536-12537.
- [127] A. Samanta, A. Krause, A. Jäschke, *Chemical Communications* **2014**, 50, 1313-1316.
- [128] M. Ziemniak, M. Strenkowska, J. Kowalska, J. Jemielity, *Future Medicinal Chemistry* **2013**, 5, 1141-1172.
- [129] S. Bednarek, V. Madan, P. J. Sikorski, R. Bartenschlager, J. Kowalska, J. Jemielity, *Philosophical Transactions of the Royal Society B: Biological Sciences* **2018**, 373, 20180167.
- [130] M. Warminski, P. J. Sikorski, Z. Warminska, M. Lukaszewicz, A. Kropiwnicka, J. Zuberek, E. Darzynkiewicz, J. Kowalska, J. Jemielity, *Bioconjugate chemistry* **2017**, 28, 1978-1992.
- [131] A. Mamot, P. J. Sikorski, M. Warminski, J. Kowalska, J. Jemielity, *Angewandte Chemie International Edition* **2017**, 56, 15628-15632.
- [132] B.-J. Cha, B. S. Koppetsch, W. E. Theurkauf, *Cell* **2001**, 106, 35-46.
- [133] M. D. Blower, E. Feric, K. Weis, R. Heald, *Journal of Cell Biology* **2007**, 179, 1365-1373.
- [134] G. Stengel, M. Urban, B. W. Purse, R. D. Kuchta, *Analytical Chemistry* **2010**, 82, 1082-1089.
- [135] Y. Liu, E. Holmstrom, P. Yu, K. Tan, X. Zuo, D. J. Nesbitt, R. Sousa, J. R. Stagno, Y.-X. Wang, *Nature Protocols* **2018**, 13, 987-1005.
- [136] A. A. Sawant, P. P. Mukherjee, R. K. Jangid, S. Galande, S. G. Srivatsan, *Organic & Biomolecular Chemistry* **2016**, 14, 5832-5842; A. A. Sawant, A. A. Tanpure, P. P. Mukherjee, S. Athavale, A. Kelkar, S. Galande, S. G. Srivatsan, *Nucleic Acids Research* **2016**, 44, e16-e16.
- [137] S. Croce, S. Serdjukow, T. Carell, T. Frischmuth, *ChemBioChem* **2020**, 21, 1641-1646.
- [138] M. B. Walunj, A. A. Tanpure, S. G. Srivatsan, *Nucleic Acids Research* **2018**, gky185-gky185.
- [139] T. Baladi, J. R. Nilsson, A. Gallud, E. Celauro, C. Gasse, F. Levi-Acobas, I. Sarac, M. R. Hollenstein, A. Dahlén, E. K. Esbjörner, L. M. Wilhelmsson, *Journal of the American Chemical Society* **2021**, 143, 5413-5424.
- [140] J. Lingner, W. Keller, *Nucleic Acids Research* **1993**, 21, 2917-2920.
- [141] M.-L. Winz, E. C. Linder, J. Becker, A. Jäschke, *Chemical Communications* **2018**, 54, 11781-11784; M.-L. Winz, A. Samanta, D. Benzinger, A. Jäschke, *Nucleic Acids Research* **2012**, 40, e78-e78; L. Anhäuser, S. Hüwel, T. Zobel, A. Rentmeister, *Nucleic Acids Research* **2019**, 47, e42-e42; K. J. Westerich, K. S. Chandrasekaran, T. Gross-Thebing, N. Kueck, E. Raz, A. Rentmeister, *Chemical Science* **2020**, 11, 3089-3095.
- [142] M. van Dülmen, N. Muthmann, A. Rentmeister, *Angewandte Chemie International Edition* **2021**, 60, 13280-13286.
- [143] Z. Huang, J. W. Szostak, *Nucleic Acids Research* **1996**, 24, 4360-4361.
- [144] M. A. Estrella, F.-T. Kuo, S. Bailey, *Genes & Development* **2016**, 30, 460-470.
- [145] M.-T. Le, W. K. Kasprzak, T. Kim, F. Gao, M. Y. L. Young, X. Yuan, B. A. Shapiro, J. Seog, A. E. Simon, *eLife* **2017**, 6, e22883.
- [146] C. Thomas, T. Rusanov, T. Hoang, T. Augustin, T. Kent, I. Gaspar, R. T. Pomerantz, *Nucleic Acids Research* **2019**, 47, 3272-3283.
- [147] B. Samanta, D. P. Horning, G. F. Joyce, *Nucleic Acids Research* **2018**, 46, e103-e103.
- [148] J. T. George, S. G. Srivatsan, *Chemical Communications* **2020**, 56, 12319-12322.

- [149] A. V. Domashevskiy, D. J. Rodriguez, D. Gunawardana, D. J. Goss, in *Synthetic mRNA: Production, Introduction Into Cells, and Physiological Consequences* (Ed.: R. E. Rhoads), Springer New York, New York, NY, **2016**, pp. 61-75.
- [150] J. M. Holstein, L. Anhäuser, A. Rentmeister, *Angewandte Chemie International Edition* **2016**, 55, 10899-10903.
- [151] D. Schulz, J. M. Holstein, A. Rentmeister, *Angewandte Chemie International Edition* **2013**, 52, 7874-7878.
- [152] A. Ovcharenko, F. P. Weissenboeck, A. Rentmeister, *Angewandte Chemie International Edition* **2021**, 60, 4098-4103.
- [153] S. C. Alexander, K. N. Busby, C. M. Cole, C. Y. Zhou, N. K. Devaraj, *Journal of the American Chemical Society* **2015**, 137, 12756-12759; C. Y. Zhou, S. C. Alexander, N. K. Devaraj, *Chemical Science* **2017**, 8, 7169-7173.
- [154] F. Ehret, C. Y. Zhou, S. C. Alexander, D. Zhang, N. K. Devaraj, *Molecular Pharmaceutics* **2018**, 15, 737-742; D. Zhang, C. Y. Zhou, K. N. Busby, S. C. Alexander, N. K. Devaraj, *Angewandte Chemie International Edition* **2018**, 57, 2822-2826.
- [155] D. A. Baum, S. K. Silverman, *Angewandte Chemie International Edition* **2007**, 46, 3502-3504.
- [156] L. Büttner, F. Javadi-Zarnaghi, C. Höbartner, *Journal of the American Chemical Society* **2014**, 136, 8131-8137.
- [157] M. Ghaem Maghami, C. P. M. Scheitl, C. Höbartner, *Journal of the American Chemical Society* **2019**, 141, 19546-19549; M. Ghaem Maghami, S. Dey, A.-K. Lenz, C. Höbartner, *Angewandte Chemie International Edition* **2020**, 59, 9335-9339.
- [158] Y. Kinoshita, Y. Nishigaki K Fau - Husimi, Y. Husimi, *Nucleic Acids Research* **1997**, 25, 3747.
- [159] J. M. Holstein, D. Stummer, A. Rentmeister, *Chemical Science* **2015**, 6, 1362-1369; F. Muttach, A. Rentmeister, *Angewandte Chemie International Edition* **2016**, 55, 1917-1920.
- [160] M. Holstein Josephin, L. Anhäuser, A. Rentmeister, *Angewandte Chemie International Edition* **2016**, 55, 10899-10903.
- [161] S. Walczak, A. Nowicka, D. Kubacka, K. Fac, P. Wanat, S. Mroczek, J. Kowalska, J. Jemielity, *Chemical Science* **2017**, 8, 260-267.
- [162] C. M. Gampe, M. Hollis-Symynkywicz, F. Zécéri, *Angewandte Chemie International Edition* **2016**, 55, 10283-10286.
- [163] S. S. Ghosh, P. M. Kao, A. W. McCue, H. L. Chappelle, *Bioconjugate Chemistry* **1990**, 1, 71-76.
- [164] R. J. Heetebrij, E. G. Talman, M. A. v. Velzen, R. P. M. v. Gijlswijk, S. S. Snoeijers, M. Schalk, J. Wiegant, F. v. d. Rijke, R. M. Kerkhoven, A. K. Raap, H. J. Tanke, J. Reedijk, H. J. Houthoff, *ChemBioChem* **2003**, 4, 573-583; V. Gupta, A. Cherkassky, P. Chatis, R. Joseph, A. L. Johnson, J. Broadbent, T. Erickson, J. DiMeo, *Nucleic Acids Research* **2003**, 31, e13-e13.
- [165] B. A. Bhoge, P. Mala, J. S. Kurian, V. Srinivasan, I. Saraogi, *Chemical Communications* **2020**, 56, 13832-13835.
- [166] R. Markham, J. D. Smith, *Biochemical Journal* **1952**, 52, 558-565; R. Whitfield Pr Fau - Markham, R. Markham.
- [167] D. M. Brown, M. Fried, A. R. Todd, *Journal of the Chemical Society (Resumed)* **1955**, 2206-2210.
- [168] P. C. Zamecnik, M. L. Stephenson, J. F. Scott, *Proceedings of the National Academy of Sciences of the United States of America* **1960**, 46, 811.

- [169] J. X. Khym, *Biochemistry* **1963**, 2, 344-350.
- [170] J. X. Khym, W. E. Cohn, *J Biol Chem* **1961**, 236, Pc9-10.
- [171] F. Hansske, M. Sprinzl, F. Cramer, *Bioorganic Chemistry* **1974**, 3, 367-376.
- [172] S. A. Reines, C. R. Cantor, *Nucleic acids research* **1974**, 1, 767-786.
- [173] G. Lentzen, B. Dobberstein, W. Wintermeyer, *FEBS letters* **1994**, 348, 233-238.
- [174] E. A. Griffiin Jr, Z. Qin, W. J. Michels Jr, A. M. Pyle, *Chemistry & biology* **1995**, 2, 761-770.
- [175] M. Zhao, F. D. Steffen, R. Börner, M. F. Schaffer, Roland K O. Sigel, E. Freisinger, *Nucleic Acids Research* **2018**, 46, e13-e13.
- [176] S. B. Easterbrook-Smith, J. C. Wallace, D. B. Keech, *European journal of biochemistry* **1976**, 62, 125-130.
- [177] R. Rayford, D. D. Anthony Jr, R. E. O'Neill Jr, W. C. Merrick, *Journal of Biological Chemistry* **1985**, 260, 15708-15713.
- [178] B.-K. Oh, N. R. Pace, *Nucleic acids research* **1994**, 22, 4087-4094.
- [179] U. von Ahsen, H. F. Noller, *Science* **1995**, 267, 234-237.
- [180] L. Bellon, C. Workman, J. Scherrer, N. Usman, F. Wincott, *Journal of the American Chemical Society* **1996**, 118, 3771-3772.
- [181] O. Rossbach, L.-H. Hung, S. Schreiner, I. Grishina, M. Heiner, J. Hui, A. Bindereif, *Molecular and cellular biology* **2009**, 29, 1442-1451.
- [182] D. Li, M. H. F. Meyer, D. K. Willkomm, M. Keusgen, R. K. Hartmann, *Biochimie* **2010**, 92, 772-778.
- [183] J. Yamamoto, S. Ebisuda, L. Kong, H. Yamago, S. Iwai, *Chemistry Letters* **2017**, 46, 767-770.
- [184] M. Ziemniak, M. Szabelski, M. Lukaszewicz, A. Nowicka, E. Darzynkiewicz, R. E. Rhoads, Z. Wieczorek, J. Jemielity, *RSC Advances* **2013**, 3, 20943-20958.
- [185] Y. Liu, E. Holmstrom, J. Zhang, P. Yu, J. Wang, M. A. Dyba, C. De, J. Ying, S. Lockett, D. J. Nesbitt, A. R. Ferré-D'Amaré, R. Sousa, J. R. Stagno, Y.-X. Wang, *Nature* **2015**, 522, 368-372.
- [186] D. Egloff, I. A. Oleinich, M. Zhao, S. L. B. König, R. K. O. Sigel, E. Freisinger, *ACS Chemical Biology* **2016**, 11, 2558-2567.
- [187] D. L. Robberson, N. Davidson, *Biochemistry* **1972**, 11, 533-537.
- [188] S. Busch, L. A. Kirsebom, H. Notbohm, R. K. Hartmann, *Journal of molecular biology* **2000**, 299, 941-951.
- [189] E. A. Griffin, Jr., Z. Qin, W. J. Michels, Jr., A. M. Pyle, *Chem Biol* **1995**, 2, 761-770.
- [190] N. Rublack, S. Müller, *Beilstein Journal of Organic Chemistry* **2014**, 10, 1906-1913.
- [191] N. G. Dolinnaya, M. Blumenfeld, I. N. Merenkova, T. S. Oretskaya, N. Krynetskaya, M. G. Ivanovskaya, M. Vasseur, Z. A. Shabarova, *Nucleic Acids Research* **1993**, 21, 5403-5407; O. A. Fedorova, M. B. Gottikh, T. S. Oretskaya, Z. A. Shabarova, *Nucleosides and Nucleotides* **1996**, 15, 1137-1147; K. Nakamoto, N. Abe, G. Tsuji, Y. Kimura, F. Tomoike, Y. Shimizu, H. Abe, *Chemical Communications* **2020**, 56, 6217-6220.
- [192] J. E. Rigden, M. A. Rezaian, *Virology* **1992**, 186, 201-206; C.-y. Chen, P. Sarnow, *Science* **1995**, 268, 415-417.
- [193] N. M. Nichols, S. Tabor, L. A. McReynolds, *Current Protocols in Molecular Biology* **2008**, 84, 3.15.11-13.15.14.
- [194] G. Kaufmann, T. Klein, U. Z. Littauer, *FEBS Letters* **1974**, 46, 271-275.
- [195] S. Kameda, H. Ohno, H. Saito, *Nucleic Acids Research* **2023**, 51, e24-e24.

- [196] M. Puttaraju, M. Been, *Nucleic Acids Research* **1992**, 20, 5357-5364; E. Ford, M. Ares, *Proceedings of the National Academy of Sciences* **1994**, 91, 3117-3121.
- [197] J. L. Litke, S. R. Jaffrey, *Nature Biotechnology* **2019**, 37, 667-675.
- [198] A. Kumar, N. Palmer, K. Miyasaki, E. Finburgh, Y. Xiang, A. Portell, A. Dailamy, A. Suhardjo, W. L. Chew, E. J. Kwon, P. Mali, Cold Spring Harbor Laboratory, **2022**.
- [199] R. J. Reischl, W. Bicker, T. Keller, G. Lamprecht, W. Lindner, *Analytical and bioanalytical chemistry* **2012**, 404, 1779-1787; Z.-C. Song, G.-Y. Ma, H.-L. Zhu, *RSC advances* **2015**, 5, 24824-24833; M. Kicsák, A. Mándi, S. Varga, M. Herczeg, G. Batta, A. Béneyi, A. Borbás, P. Herczegh, *Organic & biomolecular chemistry* **2018**, 16, 393-401.
- [200] National Center for Biotechnology Information, PubChem Compound Summary for CID 3301, Ethylenediamine, **2023**.
- [201] National Center for Biotechnology Information, PubChem Compound Summary for CID 6058, Cysteamine, **2023**.
- [202] National Center for Biotechnology Information, PubChem Compound Summary for CID 700, Ethanolamine, **2023**.
- [203] I. Elkin, A. K. Weight, A. M. Klibanov, *International Journal of Pharmaceutics* **2015**, 494, 66-72.
- [204] J. Han, K. Burgess, *Chemical Reviews* **2010**, 110, 2709-2728.
- [205] R. M. Clegg, in *Methods in Enzymology*, Vol. 211, Academic Press, **1992**, pp. 353-388.
- [206] M. Biesiada, K. J. Purzycka, M. Szachniuk, J. Blazewicz, R. W. Adamiak, in *RNA Structure Determination: Methods and Protocols* (Eds.: D. H. Turner, D. H. Mathews), Springer New York, New York, NY, **2016**, pp. 199-215.
- [207] V. Vagenende, M. G. S. Yap, B. L. Trout, *Biochemistry* **2009**, 48, 11084-11096.
- [208] Y. P. Bhavsar-Jog, E. Van Dornshuld, T. A. Brooks, G. S. Tschumper, R. M. Wadkins, *Biochemistry* **2014**, 53, 1586-1594; S. Hadži, J. Lah, *Biochimica et Biophysica Acta (BBA) - General Subjects* **2021**, 1865, 129774.
- [209] E. Grudzien-Nogalska, J. Jemielity, J. Kowalska, E. Darzynkiewicz, R. E. Rhoads, *RNA* **2007**, 13, 1745-1755; J. P. Wurm, I. Holdermann, J. H. Overbeck, P. H. O. Mayer, R. Sprangers, *Proceedings of the National Academy of Sciences* **2017**, 114, 6034-6039.
- [210] L.-Y. Chu, T.-J. Hsieh, B. Golzarroshan, Y.-P. Chen, S. Agrawal, H. S. Yuan, *Nucleic Acids Research* **2017**, 45, 12015-12024.
- [211] J. M. Messmore, D. N. Fuchs, R. T. Raines, *Journal of the American Chemical Society* **1995**, 117, 8057-8060; S. Loverix, A. Winqvist, R. Strömberg, J. Steyaert, *Chemistry & Biology* **2000**, 7, 651-658.
- [212] M. Nowotny, S. A. Gaidamakov, R. Ghirlando, S. M. Cerritelli, R. J. Crouch, W. Yang, *Molecular cell* **2007**, 28, 264-276.
- [213] A. Mamot, P. J. Sikorski, A. Siekierska, P. de Witte, J. Kowalska, J. Jemielity, *Nucleic Acids Research* **2022**, 50, e3-e3.
- [214] S. Bhaumik, S. S. Gambhir, *Proceedings of the National Academy of Sciences* **2002**, 99, 377-382.
- [215] T.-Y. Choi, T.-I. Choi, Y.-R. Lee, S.-K. Choe, C.-H. Kim, *Experimental & Molecular Medicine* **2021**, 1-8.
- [216] B. T. Abe, R. A. Wesselhoeft, R. Chen, D. G. Anderson, H. Y. Chang, *Molecular Cell* **2022**, 82, 1768-1777.e1763.
- [217] W. A. Kibbe, *Nucleic Acids Research* **2007**, 35, W43-W46.
- [218] M. R. T. JOHNSON, William R., (Ed.: W. I. P. Organization), **2016**.

- [219] P. Wanat, S. Walczak, B. A. Wojtczak, M. Nowakowska, J. Jemielity, J. Kowalska, *Organic letters* **2015**, 17, 3062-3065.
- [220] M. Warminski, J. Kowalska, J. Jemielity, *Organic letters* **2017**, 19, 3624-3627.
- [221] S. N. Floor, B. N. Jones, G. A. Hernandez, J. D. Gross, *Nature structural & molecular biology* **2010**, 17, 1096.

Lista publikacji i zgłoszeń patentowych autora:

1. Adam Mamot, Pawel J. Sikorski, Aleksandra Siekierska, Peter de Witte, Joanna Kowalska, Jacek Jemielity, „Ethylenediamine derivatives efficiently react with oxidized RNA 3' ends providing access to mono and dually labelled RNA probes for enzymatic assays and in vivo translation” *Nucleic Acids Research* **2022**, DOI: 10.1093/nar/gkab867.
2. Adam Mamot, Paweł Sikorski, Joanna Kowalska, Jacek Jemielity “N-(2-aminoetylo)morfolinowe analogi RNA, sposób ich otrzymywania i zastosowanie” (N-(2-aminoethyl)morpholine RNA analogs, method for the preparation and use thereof). Międzynarodowy numer publikacji: WO 2022/191725 A1, Międzynarodowy numer zgłoszenia: PCT/PL2022/050013
3. Renata Kasprzyk, Mateusz Fido, Adam Mamot, Przemysław Wanat, Mirosław Śmietanski, Michał Kopciał, Victoria Cowling, Joanna Kowalska, Jacek Jemielity, “Direct High-Throughput Screening Assay for mRNA Cap Guanine-N7 Methyltransferase Activity” *Chemistry - A European Journal* **2020**, 26, 11266.
4. Adam Mamot, Pawel J. Sikorski, Marcin Warmiński, Joanna Kowalska, Jacek Jemielity, “Azido-Functionalized 5' Cap Analogues for the Preparation of Translationally Active mRNAs Suitable for Fluorescent Labeling in Living Cells” *Angewandte Chemie International Edition* **2017**, 55, 15628.