

INSTYTUT CHEMII BIOORGANICZNEJ
POLSKIEJ AKADEMII NAUK

Barbara Szutkowska

**STRUKTURA DRUGORZĘDOWA vRNA WIRUSA
GRYPY TYPU A W WARUNKACH BIOLOGICZNYCH**

Praca doktorska została wykonana
w Zakładzie Genomiki Strukturalnej RNA

Promotor: prof. dr hab. Elżbieta Kierzek

POZNAŃ 2021

Spis treści

Cel pracy	9
Streszczenie	12
Abstract	15
Wstęp literaturowy	18
1. Wirus grypy typu A	18
1.1. Budowa wirusa grypy typu A	18
1.2. Budowa kompleksu rybonukleinowego vRNP	22
1.3. Cykl replikacyjny wirusa grypy	24
1.4. Białka niestrukturalne kodowane przez genom wirusa grypy i ich wpływ na cykl replikacyjny wirusa.....	29
1.5. Struktury drugorzędowe RNA wirusa grypy	32
1.5.1. Znane motywy strukturalne vRNA wirusa grypy typu A.....	35
Struktura promotora wirusowej polimerazy RNA-zależnej	36
Konserwatywne motywy strukturalne vRNA.....	36
Interakcje pomiędzy segmentami vRNA.....	42
1.5.2. Konserwatywne motywy strukturalne RNA(+) wirusa grypy typu A	44
1.6. Metody badań struktury drugorzędowej RNA.....	46
1.6.1. Metody badań struktury RNA z zastosowaniem odczynników mapujących	46
1.6.2 Metody badań struktury RNA z zastosowaniem odczynników mapujących i elektroforezy kapilarnej.....	47
1.6.3. Metody odczytu mapowania RNA oparte o zatrzymanie odwrotnej transkryptazy z zastosowaniem sekwencjonowania nowej generacji.....	48
1.7. Planowanie eksperymentu mapowania struktury RNA w warunkach biologicznych	54
1.8. Strategie przeciwwirusowe.....	56

1.8.1. Zależne od struktury drugorzędowej RNA metody inhibicji nakierowane na wirusowe RNA	57
Wyniki i dyskusja.....	61
1. Struktura drugorzędowa vRNA8 w lizacie komórek MDCK	61
1.1. Otrzymanie matrycy do transkrypcji segmentu 8 vRNA oraz transkrypcja <i>in vitro</i>	61
1.2. Namnożenie wirusa A/California/04/2009 wykorzystywanego do eksperymentów komórkowych.....	62
1.3. Infekcja komórek i przygotowanie lizatów komórkowych	63
1.4. Sprawdzenie obecności vRNA oraz białek wirusowych w lizacie komórkowym.....	64
1.5. Mapowanie struktury drugorzędowej vRNA8 w lizacie komórkowym....	65
1.6. Wyniki mapowania chemicznego struktury drugorzędowej vRNA8 w lizacie komórkowym	66
1.7. Struktura drugorzędowa vRNA8 w lizacie komórkowym	69
1.7.1. Prawdopodobieństwo parowania zasad	76
1.7.2. Analiza zachowawczości parowania zasad dla typu A wirusa grypy w strukturze drugorzędowej vRNA8 w lizacie komórkowym	78
1.7.3. Występowanie motywów strukturalnych vRNA8 w lizacie komórkowym a profil wiązania białka NP	82
2. Wyniki odczytu mapowania chemicznego vRNA w warunkach <i>in virio</i> i <i>in cellulo</i> metodą profilowania mutacji MaP	85
2.1. Optymalizacja metody odczytu profilu mutacji (MaP).....	85
2.1.1. Reakcja na matrycy RNA <i>in vitro</i>	85
2.1.2. Próbną reakcją typu SHAPE-MaP na matrycy RNA <i>in cellulo</i>	89
2.2. Wyniki mapowania struktury drugorzędowej wszystkich segmentów vRNA wirusa grypy szczepu A/California/04/2009 w warunkach <i>in virio</i> i <i>in cellulo</i>	90

2.2.1. Przygotowanie matrycy RNA z eksperymentu mapowania chemicznego <i>in virio</i>	91
2.2.2. Przygotowanie matrycy RNA z eksperymentu mapowania chemicznego <i>in cellulo</i> z zainfekowanych wirusem grypy komórek A549	91
2.2.3. Biblioteki <i>in virio</i> : Przygotowanie matrycy, reakcja tagmentacji i amplifikacja bibliotek fragmentów do sekwencjonowania	92
2.2.4. Biblioteki <i>in cellulo</i> : Przygotowanie matrycy, reakcja tagmentacji i amplifikacja bibliotek fragmentów do sekwencjonowania	97
2.2.5. Przygotowanie biblioteki do sekwencjonowania NGS dla próby kontrolnej, zawierającej RNA o znanej strukturze drugorzędowej (kontrola <i>in cellulo</i>)	99
2.2.6. Biblioteki do sekwencjonowania NGS- analiza ilościowa i jakościowa bibliotek	99
2.2.7. Wyniki sekwencjonowania	108
2.3. Struktura drugorzędowa vRNA w warunkach <i>in virio</i> oraz <i>in cellulo</i>	118
2.3.1. Segmenty 1, 2, 3, 4 oraz 6 vRNA	119
2.3.2. Segment 5 vRNA	127
2.3.3. Segment 7 vRNA	132
2.3.4. Segment 8 vRNA	136
2.4. Porównanie struktury MFE vRNA8 tworzącej się w różnych warunkach biologicznych	140
2.5. Występowanie motywów strukturalnych vRNA o wysokim prawdopodobieństwie <i>in virio</i> a profil wiązania białka NP	141
2.6. Analiza korelacji wyników mapowania struktur drugorzędowych vRNA w warunkach <i>in virio</i> oraz <i>in cellulo</i>	145
Podsumowanie	149
Materiały i metody	156
1. Materiały.....	156

1.1. Sekwencje stosowanych oligonukleotydów	156
1.2. Komercyjnie dostępne zestawy	157
1.3. Enzymy	158
1.4. Barwniki fluorescencyjne	158
1.5. Wzorce wielkości	159
1.6. Plazmidy	159
1.7. Odczynniki dostępne komercyjnie	159
1.8. Bufory i roztwory	160
1.9. Przeciwciała.....	162
1.10. Linie komórkowe.....	163
1.11. Szczep wirusa grypy	163
1.12. Media i roztwory wykorzystywane w eksperymentach komórkowych	163
2. Metody.....	165
2.1. Synteza oligonukleotydów	165
2.2. Odblokowanie oligonukleotydów po syntezie	165
2.3. Znakowanie końca 5' oligonukleotydów barwnikami fluorescencyjnymi	166
2.4. Oczyszczanie oligonukleotydów na denaturującym żelu poliakrylamidowym.....	166
2.5. Synteza odczynników chemicznych do reakcji mapowania SHAPE.....	166
2.6. Strącanie kwasów nukleinowych	167
2.7. Elektroforeza kwasów nukleinowych w żelu agarozowym	167
2.8. Oczyszczanie kwasów nukleinowych z zastosowaniem kulek magnetycznych typu SPRI	167
2.9. Synteza matrycy DNA do przygotowania drabinek sekwencyjnych oraz do transkrypcji vRNA8 szczepu A/California/04/2009 <i>in vitro</i>	168

2.10. Transkrypcja segmentu 8 vRNA szczepu A/California/04/2009 <i>in vitro</i>	169
2.11. Fałdowanie vRNA8 w warunkach <i>in vitro</i>	169
2.12. Fałdowanie vRNA8 w lizacie komórek MDCK	170
2.13. Mapowanie chemiczne	170
2.14. Przygotowanie drabinek sekwencyjnych	173
2.15. Rozdział produktów reakcji odwrotnej transkrypcji i drabinek sekwencyjnych z zastosowaniem elektroforezy kapilarnej	173
2.16. Analiza bioinformatyczna rozdziałów z zastosowaniem programu ShapeFinder	174
2.17. Predykcja struktury drugorzędowej vRNA8 mapowanej w lizacie komórkowym	175
2.18. Predykcja struktury drugorzędowej vRNA mapowanych w warunkach <i>in virio</i> oraz <i>in cellulo</i>	176
2.19. Hodowla linii komórkowych MDCK oraz A549	177
2.20. Propagacja wirusa w linii komórkowej MDCK	177
2.21. Test immunofluorescencji	178
2.22. Zagęszczanie roztworu wirusa z zastosowaniem poduszki sacharozowej	179
2.23. Infekcja wirusem grypy linii komórkowej A549	179
2.24. Infekcja wirusem grypy linii komórkowej MDCK	179
2.25. Przygotowanie lizatu komórkowo-wirusowego	179
2.26. Izolacja RNA po eksperymencie mapowania <i>in virio</i> oraz <i>in cellulo</i> ...	181
2.27. Reakcja enzymatycznej degradacji DNA w izolacie całkowitego RNA komórkowego i oczyszczanie	182
2.28. Pomiar stężenia kwasów nukleinowych	182
2.29. Analiza obecności białek wirusowych z zastosowaniem metody Western-Blot	183

2.30. Analiza obecności białek w lizatach komórkowo-wirusowych	184
2.31. Analiza vRNA z zastosowaniem metody ilościowego real-time PCR w czasie rzeczywistym	184
2.32. Reakcja odwrotnej transkrypcji indukującej mutacje w obecności jonów Mn^{2+} (MaP)	186
2.33. Reakcja syntezy drugiej nici.....	189
2.34. Reakcja PCR z zastosowaniem specyficznych starterów na wszystkie segmenty vRNA	190
2.35. Sekwencjonowanie produktów PCR metodą Sanger	191
2.36. Analiza wyników sekwencjonowania metodą Sanger	191
2.37. Przygotowanie bibliotek DNA do wysokoprzepustowego sekwencjonowania nowej generacji (NGS)	191
2.38. Rozdział fragmentów bibliotek z zastosowaniem bioanalyzera.....	193
2.39. Normalizacja bibliotek do sekwencjonowania NGS	193
2.40. Sekwencjonowanie z zastosowaniem sekwencjatora Illumina NextSeq550	194
2.41. Analiza wyników sekwencjonowania NGS	194
Dodatek	196
Bibliografia	204

Jestem jedną z tych osób, które uważają, że nauka kryje w sobie wielkie piękno. Naukowiec w swym laboratorium jest nie tylko technikiem – jest także dzieckiem stojącym twarzą w twarz ze zjawiskami natury, które budzą jego zachwyt niczym bajka.

- Maria Skłodowska-Curie

Dedykuję tę pracę Wszystkim, którzy byli inspiracją i przyczynili się do jej powstania.

Cel pracy

Istnieją cztery typy wirusa grypy: A, B, C i D, z czego trzy pierwsze mają zdolność do zakażenia organizmu ludzkiego¹². Szczepy A i B cyrkulują jako sezonowe szczepy wirusa, a zakażenie powoduje chorobę górnych dróg oddechowych o ostrym przebiegu, zwaną grypą. Zakażenie wirusem typu C u ludzi wywołuje chorobę o łagodnym przebiegu. Natomiast odnotowano wiele przypadków epidemii grypy wywołanych przez typy A i B wirusa, a pandemiczny potencjał ma jedynie wirus grypy typu A (IAV)³. Jednym z przykładów pandemii wirusa IAV była wywołana przez szczep H1N1 grypa zwana „hiszpanką”, która w latach 1918-19 dotknęła aż 1/3 ówczesnej populacji ludzkiej⁴⁻⁶. Niedawną pandemię wywołał w 2009 roku szczep A/California/04/2009 (H1N1). Według szacunków WHO, nawet 1 miliard ludzi rocznie choruje na grypę, z czego u 3-5 milionów obserwuje się ciężki przebieg choroby, a nawet 290-650 tysięcy zakażeń kończy się zgonem⁷.

Wirus IAV należy do rodziny RNA wirusów, które ze względu na potencjał pandemiczny uznawane są za jedne z najbardziej niebezpiecznych na świecie. Genom wirusa grypy składa się z 8 jednoniciowych cząsteczek RNA (vRNA) o ujemnej polarności¹. Oznacza to, że w trakcie cyklu replikacyjnego wirusa cząsteczki vRNA przepisywane są na dwie komplementarne cząsteczki – cRNA oraz mRNA⁸. Cząsteczki cRNA są matrycą do powstawania nowych cząsteczek vRNA, natomiast cząsteczki mRNA służą jako matryca do powstawania białek wirusowych. W czasie cyklu replikacyjnego cząsteczki wirusowego RNA wchodzi w liczne interakcje typu RNA-RNA oraz RNA-białko, które mają wpływ na strukturę cząsteczki. Cząsteczki genomowego RNA (vRNA) w dojrzałym wirionie tworzą kompleks rybonukleoproteinowy (vRNP), na który, oprócz RNA, składają się liczne kopie białka nukleoproteiny (NP) oraz kompleks wirusowej polimerazy⁹.

Funkcjonalność cząsteczki RNA warunkowana jest przez jej strukturę drugorzędową, na którą składają się motywy strukturalne. Nie inaczej jest w przypadku RNA wirusa grypy. Przez naszą grupę zostały przeprowadzone liczne badania struktur drugorzędowych vRNA w warunkach *in vitro*, które zostały wzbogacone w sekwencyjno-strukturalną analizę bioinformatyczną¹⁰⁻¹⁴. Badania te pokazały, że część motywów strukturalnych RNA IAV jest zakonserwowana

ewolucyjnie i jest obecna w strukturach drugorzędowych odległych genetycznie podtypów wirusa grypy typu A. Oznacza to, że pewne motywy strukturalne w cząsteczkach RNA mają znaczenie funkcjonalne, a zaburzenie w obrębie ich budowy może przyczynić się do obniżenia wirulencji wirusa, bądź do jego całkowitej niezdolności do namnażania. Zaburzenia funkcjonalności motywów strukturalnych RNA możliwe jest dzięki zastosowaniu inhibitorów naczelnianych na strukturę RNA – co pokazały skuteczne strategie antysensowe^{10,12}. Ze względu na istnienie konserwatywnych motywów strukturalnych w RNA wirusa grypy, najlepszym celem uniwersalnej dla wielu szczepów inhibicji mogą być właśnie te motywy..

Badanie struktury RNA wirusa grypy w warunkach *in vitro* stanowi podstawowy model do określenia struktury drugorzędowej tej cząsteczki, jednakże nie uwzględnia oddziaływań typu RNA-białko, które mogą zachodzić w komórce i mają wpływ na ostateczną strukturę. Dlatego istnieje potrzeba przeprowadzenia badań w warunkach naturalnych lub jak najbardziej do nich zbliżonych. Badanie wirusowych RNA w warunkach *in vivo* stanowi wyzwanie metodologiczne i wymaga zastosowania bardziej wszechstronnego podejścia w trakcie planowania eksperymentów i interpretacji wyników. Co istotne, dotychczas nie została określona struktura vRNA wirusa grypy typu A w warunkach komórkowych. W niniejszej rozprawie doktorskiej postanowiono zbadać strukturę drugorzędową vRNA, która tworzy się w trakcie oddziaływań z biomolekułami w różnych warunkach biologicznych: w lizacie komórkowo-wirusowym, w wirionie (*in virio*) oraz w żywej komórce (*in cellulo*). Jako obiekt badań wybrano pandemiczny szczep grypy A/California/04/2009 (H1N1).

W niniejszej pracy zostały wyznaczone następujące cele szczegółowe:

1. Opracowanie modelu do badania struktury drugorzędowej RNA w obecności innych cząsteczek biologicznych *in-vivo-like* poprzez stworzenie lizatu z zakażonych wirusem grypy komórek MDCK (Madin Darby Canine Kidney). Zbadanie struktury drugorzędowej segmentu 8 vRNA (vRNA8) w lizacie komórkowo-wirusowym i porównanie ze strukturą vRNA8 *in vitro*. Określenie podobieństwa strukturalnego oraz wpływu składników obecnych w lizacie na strukturę drugorzędową vRNA8. Zbadanie zachowawczości motywów strukturalnych vRNA8.

2. Opracowanie metody mapowania struktury drugorzędowej vRNA w wirionie oraz metody przygotowania biblioteki fragmentów do sekwencjonowania nowej generacji (NGS). Określenie struktur drugorzędowych segmentów vRNA *in virio*.
3. Opracowanie metody mapowania struktury drugorzędowej vRNA w ludzkich komórkach płucnych A549. Opracowanie metody przygotowania biblioteki fragmentów do sekwencjonowania NGS. Zaproponowanie struktur drugorzędowych vRNA *in cellulo*.
4. Porównanie uzyskanych struktur drugorzędowych vRNA ze strukturami drugorzędowymi vRNA zaproponowanymi na podstawie badań w różnych warunkach dla innych szczepów wirusa grypy. Określenie konserwatywnych motywów strukturalnych vRNA, które mogą stanowić cel dla uniwersalnej dla wielu szczepów IAV inhibicji namnażania wirusa.
5. Porównanie struktur drugorzędowych vRNA określonych na podstawie badań w różnych układach biologicznych.

Streszczenie

Na dzień dzisiejszy scharakteryzowano cztery typy wirusa grypy - A, B, C i D. Spośród nich za najbardziej zagrażające człowiekowi uznaje się zoonotyczne szczepy wirusa grypy typu A (IAV), które mają duży potencjał do wywołania epidemii, a także pandemii. U podstaw wirulentności wirusa grypy leżą zjawiska skoku oraz dryfu genetycznego, które przyczyniają się do szybkiej ewolucji genetycznej. Szybka ewolucja jest możliwa dzięki niewielkiemu (13 kbp) genomowi, na który składają się jednoniciowe cząsteczki RNA (vRNA) zwane również segmentami. W przypadku szczepów wirusa typu A jest to 8 jednoniciowych segmentów vRNA, które wraz z kompleksem wirusowej polimerazy oraz wieloma kopiami stabilizującego je białka NP (nukleoproteina) tworzą kompleks rybonukleoproteinowy (RNP). Składające się na genom wirusa cząsteczki vRNA mają ujemną polarność. Oznacza to, że w czasie cyklu replikacyjnego wirusa są one przepisywane na komplementarne cząsteczki cRNA, które są matrycą do powstawania nowych cząsteczek vRNA. Cząsteczki vRNA są również matrycą do transkrypcji cząsteczek mRNA ulegających translacji na białka wirusowe.

Funkcja cząsteczek RNA jest warunkowana ich strukturą drugorzędową oraz możliwością tworzenia oddziaływań z innymi biomolekułami, takimi jak oddziaływanie typu RNA-RNA czy RNA-białko. W przypadku vRNA wirusa grypy opisano szereg motywów strukturalnych pełniących funkcje na różnych etapach cyklu życiowego wirusa, których zaburzenie przyczyniało się do obniżenia namnażania wirusa. Chociaż na dzień dzisiejszy dysponujemy wiedzą dotyczącą fałdowania się wybranych segmentów vRNA wirusa grypy typu A w warunkach *in vitro* oraz opublikowano dane mapowania *in virio* dla niektórych szczepów, to nadal brakuje pełnej wiedzy dotyczącej formowania się struktur drugorzędowych vRNA w warunkach naturalnych.

Celem badań opisanych w niniejszej rozprawie doktorskiej było określenie struktur drugorzędowych vRNA w warunkach biologicznych: w lizacie komórkowo-wirusowym (*in vivo-like*), a także w zainfekowanych wirusem grypy komórkach (*in cellulo*) oraz w wirionie (*in virio*). Do badań wykorzystano szczep

A/California/04/2009 (H1N1), który był odpowiedzialny za pandemię w latach 2009-2010.

W pierwszej kolejności określona została struktura segmentu 8 vRNA (vRNA8) w lizacie zainfekowanych wirusem grypy komórek MDCK (Madin Darby Canine Kidney). Do przewidywania struktur drugorzędowych wykorzystano program RNAStructure i dane eksperymentalne z mapowań chemicznych (DMS, 1M7). Wygenerowano strukturę globalną, a także strukturę lokalną, w której zastosowano dystans parowania zasad. Umożliwiło to określenie wpływu ograniczenia dystansu parowania zasad w przewidywaniu na generowaną strukturę drugorzędową oraz wykrycie motywów strukturalnych, które były przewidziane niezależnie od zastosowanych parametrów. Następnie obliczona została konserwatywność parowania zasad (na podstawie analizy sekwencyjno-strukturalnej wszystkich szczepów grypy typu A) w obrębie całej struktury vRNA8. Na podstawie analiz zostały wyłonione motywy strukturalne, które mają wysoką konserwatywność strukturalną wśród innych szczepów IAV.

W drugiej części niniejszej rozprawy doktorskiej zostały opisane badania skupiające się na poznaniu struktur drugorzędowych vRNA w warunkach *in virio* oraz *in cellulo*. W pierwszej kolejności zostały opracowane metody mapowania chemicznego vRNA w wirionie oraz w zainfekowanych wirusem grypy ludzkich komórkach płucnych (A549). Istotnym etapem było również opracowanie metody przygotowania bibliotek do sekwencjonowania NGS. Wyniki sekwencjonowania umożliwiły obliczenie reaktywności nukleotydów względem odczynników chemicznych (NAI, DMS) dla pełnej długości wszystkich segmentów vRNA w warunkach *in virio*, a także pełnej długości segmentów vRNA5, 7 i 8 w warunkach *in cellulo*. Uzyskane wyniki reaktywności dla trzech powtórzeń biologicznych wskazały na bardzo wysoką korelację, co świadczyło o powtarzalności wyników. Na podstawie wyników sekwencjonowania wyłoniono motywy strukturalne wspólne dla struktur drugorzędowych w warunkach *in virio* oraz *in cellulo*. Uzyskane wyniki porównano również do znanych struktur drugorzędowych vRNA innych szczepów wirusa grypy uzyskanych w badaniach *in virio* oraz *in vitro*, co umożliwiło wyłonienie lokalnych motywów strukturalnych wspólnych dla wielu szczepów IAV. Badania pokazały

również istnienie nowych motywów strukturalnych dla poszczególnych segmentów vRNA.

Następnie porównano struktury drugorzędowe vRNA8 *in virio*, *in cellulo* ze strukturą drugorzędową uzyskaną w lizacie komórkowo-wirusowym. Porównanie struktur drugorzędowych wykazało, że część motywów była wspólna dla wszystkich trzech struktur vRNA8. Natomiast część motywów przewidzianych w vRNA8 w lizacie komórkowo-wirusowym była unikalna albo dla struktury drugorzędowej *in virio* albo struktury *in cellulo*. To wykazało, że struktura w lizacie jest strukturą pośrednią pomiędzy obiema strukturami.

Struktury drugorzędowe vRNA5, 7 oraz 8 o najniższej energii swobodnej (MFE) w warunkach *in virio* oraz *in cellulo* różnią się od siebie w ujęciu globalnym. Najwięcej różnic zaobserwowano w parowaniu zasad pomiędzy odległymi nukleotydami. Zmiany te sugerują dużą dynamikę struktur RNA w środowisku biologicznym. Jednocześnie część motywów strukturalnych, szczególnie tych lokalnych, została zachowana, co prawdopodobnie ma znaczenie funkcjonalne.

Opisane w niniejszej rozprawie doktorskiej badania pozwoliły na bardzo szeroką analizę struktur drugorzędowych vRNA wirusa grypy typu A w różnych warunkach biologicznych. Po raz pierwszy określono struktury vRNA w warunkach *in cellulo*, a także porównano je ze strukturami vRNA w innych warunkach biologicznych. Badania przyczyniły się do lepszego poznania biologii wirusa od strony genomowego RNA. Opisane zachowawcze motywy strukturalne vRNA mogą stanowić idealny cel dla uniwersalnej, niezależnej od szczepu, inhibicji wirusa grypy typu A na poziomie RNA.

Abstract

As of today, four types of Influenza viruses were characterized – A, B, C and D. Due to epidemic and pandemic potential, zoonotic Influenza A virus (IAV) strains are considered the most human-threatening. The phenomena of the antigenic shift and drift underlie the virulence of the IAV and contribute to rapid genetic evolution. Fast evolution is supported by a small (13 kbp) genome, which consists of single-stranded RNAs (vRNA) called segments. In the case of the Influenza A virus, the genome consists of the 8 single-stranded vRNAs, which along with the viral polymerase complex and multiple copies of stabilizing NP protein (nucleoprotein) form ribonucleoprotein complex (RNP). The genomic vRNAs are of negative-sense. This means that during the viral replication cycle, the vRNAs are replicated into complementary RNAs (cRNA), which are the template for the synthesis of new vRNAs. The vRNAs are also templates for the transcription of mRNAs which are translated into viral proteins.

RNA function depends on its secondary structure and the possibility of interactions with biomolecules such as RNA-RNA or RNA-protein interactions. In the case of influenza virus vRNA, a number of structural motifs have been described that perform function at different stages of the IAV lifecycle. The disturbance of these motifs contributed to the reduction of viral proliferation. Although we have knowledge of the *in vitro* folding of selected IAV vRNA segments and *in virio* mapping data for some strains have been published, we still lack the full knowledge of the vRNAs secondary structure folding in natural environment.

The main objectives of the research described in this doctoral thesis was to determine the vRNAs secondary structures in the biological environment: in cellular lysates (*in vivo-like*), in IAV-infected cells (*in cellulo*), and in virion (*in virio*). The pandemic strain A/California/04/2009 (H1N1) responsible for pandemics in the years 2009-2010 was used for the research.

In the first place, the vRNA8 structure in the lysate of IAV-infected MDCK (Madin Darby Canine Kidney) cells was determined. The RNAstructure program and experimental data from chemical mapping (DMS, 1M7) were used for the secondary structure prediction. The global structure, as well as local structure with maximum pairing distance, were generated during the prediction. This enabled determination of

the impact of limited base-pairing distance on the secondary structure prediction and identification of structural motifs forming regardless of the parameters used. Next, the base-pairing conservation (based on structural-sequence analysis of all IAV strains) of vRNA8 structure was calculated. The analysis revealed structural motifs with high structural conservation among different IAV strains.

The research focused on the secondary structure of vRNAs *in virio* and *in cellulo* was described in the second part of this doctoral thesis. In the first place, the methods of chemical probing of vRNAs in virion as well as IAV-infected human lung cells (A549) were established. Another essential step was the development of NGS library preparation protocols. The sequencing results enabled the calculation of single-nucleotide reactivities to chemical reagents (NAI, DMS) for the full-length of all vRNA segments *in virio* as well as the full-length vRNA5, 7, and 8 *in cellulo*. The results obtained for three independent biological replicates showed a very high correlation, which proved the reproducibility of the results. Based on the sequencing results, common RNA structural motifs were revealed for *in virio* and *in cellulo* structures. The obtained results were also compared to the known *in virio* and *in vitro* vRNA secondary structures of other IAV strains, which allowed to indicate local structural motifs common to many IAV strains. The research revealed new structural motifs in certain vRNA segments as well.

Next, the vRNA8 secondary structures determined *in virio* and *in cellulo* were compared with the secondary structure obtained in cellular lysate. The comparison of the secondary structures showed that some of the motifs were common to all three analyzed vRNA8 structures. Importantly, some of the vRNA8 motifs predicted in cell lysates were unique to either the *in virio* secondary structure or *in cellulo* structure. This indicated that the structure in lysate was an intermediate structure between both structures.

The MFE (Minimum Free Energy) secondary structures of vRNA5, 7 and 8 in *in virio* and *in cellulo* conditions differ in global folding. Most differences were observed in the long-range base-pairing between distant nucleotides. These changes suggest high structural dynamics of RNA in the biological environment. Still, some structural motifs, especially local, were preserved probably due to their functional significance.

The research described in this doctoral thesis enables a very wide analysis of IAV vRNA secondary structures in the different biological environments. For the first time, vRNA structures were determined under *in cellulo* conditions and also compared with vRNA structures under other biological conditions. This research contributed to a better understanding of viral biology from the IAV genomic RNA point of view. Described conserved vRNA structural motifs may be an ideal target for universal, strain-independent inhibition of Influenza A virus at the RNA level.

Wstęp literaturowy

1. Wirus grypy typu A

Wirus grypy należy do rodziny wirusów *Orthomyxoviridae*. Na dzień dzisiejszy scharakteryzowane zostały 4 typy wirusa grypy: A, B, C¹ oraz D². Wirus grypy typu C ma zdolność do infekcji człowieka, jednakże wywoływane przez niego choroby układu oddechowego charakteryzują się łagodnym przebiegiem³. Wirus grypy typu D ma zdolność do zakażeń bydła oraz świń, dotychczas nie potwierdzono zakażeń wśród ludzi¹⁵. Szczepy typu A i B występują jako sezonowe szczepy wirusa grypy wywołujące dotkliwe choroby układu oddechowego, dosyć często o ostrym przebiegu³. Budowa genetyczna wirusa grypy sprawia, że jest on podatny na mutacje prowadzące do powstawania nowych szczepów. Nowopowstałe szczepy mogą mieć potencjał wywołania epidemii, niekiedy o charakterze globalnym – pandemii. Epidemii mogą wywołać zarówno szczepy typu A jak i B, natomiast dotychczasowe pandemie wywołane były jedynie przez wirusy typu A³. Za najniebezpieczniejsze uznaje się szczepy zoonotyczne, które mogą przenosić się ze zwierząt na ludzi. Szacuje się, że ok. 1 miliard ludzi rocznie choruje na grypę, z czego 3-5 milionów stanowią ciężkie przypadki zachorowań, w których 290-650 tysięcy kończy się śmiercią⁷. W historii współczesnej odnotowano kilka przypadków pandemii wywołanych przez wirusa grypy typu A podtypów H1N1, H2N2 oraz H3N2^{4,5,16}. Pierwsza z nich była tzw. „grypa hiszpanka” (podtyp H1N1), która w ciągu jednego roku, na przełomie lat 1918-19, objęła swoim zasięgiem wszystkie kontynenty. Wywołujący ją szczep grypy charakteryzował się wyjątkowo wysoką wirulencją i szacuje się, że dotknął aż 1/3 ówczesnej populacji ludzkiej, doprowadzając do śmierci wielu milionów ludzi⁶. Natomiast, ostatnia pandemia wirusa grypy nastąpiła w latach 2009-10 i była wywołana zmutowanym szczepem A/California/04/2009 (podtyp H1N1) nazywanym „grypą świńską”.

1.1. Budowa wirusa grypy typu A

Wirus grypy typu A ma stosunkowo niewielki (13kbp) segmentowany genom na który składa się 8 jednoniciowych cząsteczek vRNA (ang. viral RNA) o ujemnej polarności¹. Wszystkie segmenty vRNA mają konserwatywne sekwencje UTR (ang. untranslated region) na 3' oraz 5' końcach, które otaczają otwartą ramkę odczytu¹⁷. Ze

względu na to, że wirus grypy należy do rodziny wirusów RNA o ujemnej polarności, to mówiąc o ramce odczytu vRNA należy pamiętać, że jest ona w orientacji antysensowej w stosunku do sensowej cząsteczki komplementarnego (+)RNA. Oprócz wspólnej dla wszystkich vRNA sekwencji UTR, każdy segment vRNA posiada specyficzną dla siebie sekwencję niekodującego rejonu UTR¹⁸. Sekwencje te są również zachowawcze w obrębie segmentów vRNA danego szczepu wirusa. Segmenty vRNA są ponumerowane w zależności od długości – najdłuższy oznaczany jest jako segment 1, najkrótszy zaś jako segment 8. Przez długi czas uważano, że genom wirusa grypy koduje 11 podstawowych białek wirusowych¹. Najnowsze badania wskazują jednak na to, że genom wirusa zawiera informacje genetyczną o znacznie większej ilości białek, pełniących wiele, istotnych z punktu widzenia infekcyjności wirusa^{19,20}. Potwierdzone eksperymentalnie białka i izoformy białek wirusa grypy typu A zostały zebrane **Tabeli 1**. Długości vRNA poszczególnych szczepów różnią się, dlatego wartości podane w tabeli odnoszą się do szczepu A/California/04/2009 (H1N1), na podstawie bazy danych Uniprot (www.uniprot.org).

Tabela 1 Zastawienie segmentów genomowego RNA wirusa grypy typu A oraz kodowane przez nie główne białka.

Segment	Długość vRNA (nt)	Kodowane białko i długość (AA)	Funkcje kodowanych białek
Segment 1	2341	PB2 (759)	inicjacja transkrypcji; podkradanie struktury kapu (ang. Cap-snatching); replikacja ²¹ ; regulacja aktywności kompleksu vRNP; inhibicja mitochondrialnego białka sygnalizacyjnego MAVS ²²
Segment 2	2341	PB1 (757), PB1-F2 (87, 90) PB1-N40 (717)	podjednostka RdRp, elongacja RNA; aktywność endonukleazy ¹ ; czynnik wirulencji; funkcja proapoptyczna w komórkach poprzez oddziaływanie z białkami błony mitochondrium – VDAC-1 oraz ANT-3 ²³ ; białko regulujące balans w poziomie ekspresji białek PB1 i PB1-F2 ²⁴
Segment 3	2233	PA (716) PA-X (252) PA-155 (562) PA-182 (535)	podjednostka RdRp; aktywność endonukleazy, odcinanie struktury kapu ²⁵ ; hamowanie translacji białek gospodarza, inhibicja odpowiedzi immunologicznej ^{19,24,26,27} funkcja nieznana ^{24,28} funkcja nieznana ^{24,28}
Segment 4	1778	HA (566)	białko powierzchniowe i główny antygen wirusa; odpowiedzialne za fuzję wirionu z błoną komórkową; rozpoznaje receptory na powierzchni infekowanej komórki rozpoznając reszty kwasu sjałowego tworzącego wiązania typu α -2,3/2,6-sjałowe; ¹
Segment 5	1565	NP (498)	białko wiążące vRNA i cRNA w kompleksie rybonukleoproteinowym (RNP) ^{29,9} ; posiada sekwencje

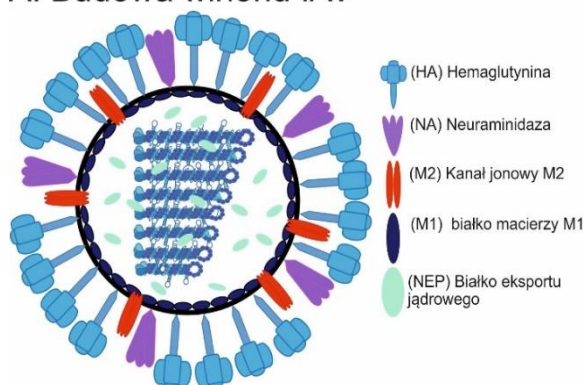
			lokalizacji jądrowej i bierze udział w imporcie vRNP do wnętrza jądra ³⁰
Segment 6	1413	NA (469)	glikoproteina powierzchniowa; aktywność sialidazy; odpowiedzialna za uwalnianie potomnych wirionów ²⁵
Segment 7	1027	M1 (248),	białko wyściełające wnętrze wirionu; bierze udział w stabilizacji kompleksów vRNP ³¹ ; odpowiedzialne za hamowanie cyklu replikacyjnego komórki gospodarza poprzez interakcję z czynnikiem inicjacji replikacji SLD5 ³² ;
		M2 (97)	transbłonowe białko budulcowe - wiroporyna o funkcji kanału jonowego zależnego od pH (IAV); na etapie późnego endosomu, pod wpływem zmiany pH, odpowiedzialny za odłączenie białek M1 od kompleksów vRNP i ich uwolnienie z wnętrza wirionu do cytoplazmy; stabilizuje pH w błonie <i>Golgi</i> zapobiegając cięciu proteolitycznemu białka H0 i tym samym przedwczesnemu powstaniu aktywnej formy tego białka ³³
		M42 (99)	Opisane dla segmentu vRNA7 białko niestrukturalne M-42 jest izoformą białka M2, które najprawdopodobniej pełni taką samą funkcję co białko M2 ^{26,24,34} .
Segment 8	890	NS1 (219), NEP/NS2 (121) NS3 (194)	białko niestrukturalne; jeden z głównych czynników wirulencji wirusa grypy; inhibitor ekspresji genów komórki gospodarza; inhibitor szlaków powstawania odpowiedzi antywirusowej, min. interferonu (IFN) oraz RIG-I ³⁵ ; selektywnie pobudza translację białek wirusowych; reguluje proces replikacji vRNA ³⁶ wspomaga transport vRNP z wnętrza jądra komórkowego; zapobiega akumulacji produktów replikacji i transkrypcji (cRNA, mRNA, vRNA) ³⁷ ; wspomaga uwalnianie wirionów potomnych poprzez stymulację komórkowej F1F0 ATPazy ³⁸ funkcja nieznaną ^{39,24,26}

Wirion wirusa grypy (**Ryc. 1A**) charakteryzuje się wysokim pleomorfizmem, jednakże po wyjściu z komórki najczęściej przybiera kształt sferyczny o średnicy 80-120 nm³. Otoczka wirusa składa się z warstwy lipidowej inkrustowanej glikoproteinami antygenowymi – hemaglutyniną (ang. Hemagglutinin, HA) i neuraminidazą (ang. Neuraminidase, NA), które występują na powierzchni wirionu w stosunku 4:1¹. Wyróżnić można 18 różnych typów białka HA oraz 11 typów białka NA i na ich podstawie rozróżnia się poszczególne subtypy wirusa³.

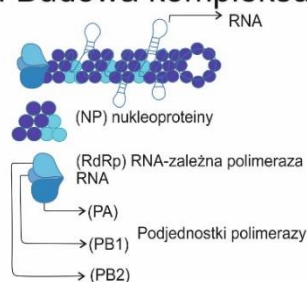
Hemaglutynina jest białkiem o dużym powinowactwie do kwasu sjałowego (kwas N-acetylneuraminowy) (ang. Sialic Acid, SA) obecnego na powierzchni komórek wielu organizmów i odpowiada za przyłączenie się cząsteczki wirusowej do infekowanej komórki. Powinowactwo do określonych komórek jest podyktowane typem wiązania, które tworzy kwas sjałowy obecny w glikoproteinach powierzchniowych komórek gospodarza i zależy również od subtypu wirusa. W przypadku komórek nabłonka górnych dróg oddechowych u ludzi, białko HA

rozpoznaje i wiąże wiele reszt kwasu sjałowego tworzącego wiązania α -2,6-sjałowe, tak więc subtypy wirusa infekujące organizmy ludzkie wykazują większą preferencję do tego właśnie typu wiązania¹. W przypadku szczepów wirusa grypy infekujących ptaki preferowane jest wiązanie do reszt kwasu sjałowego tworzącego wiązanie α -2,3-sjałowe, które występuje licznie w komórkach nabłonka jelit ptasich¹. Szczepy grypy atakującej ptaki mają również zdolność do infekcji komórek nabłonka ludzkich dolnych dróg oddechowych, które są bogate w kwas sjałowy tworzący wiązania α -2,3-sjałowe¹.

A. Budowa wirionu IAV



B. Budowa kompleksu vRNP



Ryc. 1A. Budowa wirionu IAV. Wirion wirusa grypy jest sferyczny, opłaszczony warstwą dwufosfolipidową, pochodzącą z błony komórki gospodarza (Samji, 2009). Na powierzchni występują białka HA/NA w stosunku 4:1 oraz transbłonowy kanał jonowy M2. Wewnętrzna strona wirionu pokryta jest białkiem macierzowym M1. W środku wirionu występuje białko niestrukturalne 2 (NEP/NS2) oraz 8 cząsteczek kompleksu rybonukleoproteinowego vRNP. **B.** Budowa kompleksu vRNP. Kompleks vRNP składa się z cząsteczki vRNA oraz związanych z nią czterech typów białek – wielu kopii nukleoproteiny (NP) oraz kompleksu wirusowej polimerazy RNA (RdRp) składającej się z trzech podjednostek PB1 i PB2 oraz podjednostki PA.

Drugim białkiem wirusa grypy, które charakteryzuje się dużym powinowactwem do kwasu sjałowego jest neuraminidaza. Białko to katalizuje cięcie wiązania kwasu sjałowego na ostatnim etapie infekcji, podczas uwalniania potomnych wirionów z komórki gospodarza. Odpowiedzialne jest za cięcie reszt kwasu sjałowego z receptorów obecnych na powierzchni zainfekowanej komórki⁴⁰. Cięcie reszt kwasu sjałowego zapobiega agregowaniu cząsteczek wirusowych, a także ponownemu wiązaniu cząsteczek wirusa do umierającej już komórki gospodarza⁴⁰.

Pozostałe budulcowe białka wirusowe pełnią równie istotne funkcje w cyklu życiowym wirusa grypy. Kanał jonowy zbudowany z białka M2 obecny w otoczce wirusowej (ang. M2 ion channel), pełni ważne, zależne od zmiany pH, funkcje w

czasie infekcji⁴¹. Wewnątrz wirionu jest wyścielane białkiem macierzowym M1 (ang. Matrix 1 protein), które oprócz funkcji budulcowej, zaangażowane jest w opłaszczanie genomu wirusa w czasie opuszczania zainfekowanej komórki⁴².

W środku wirionu znajduje się genom wirusa w formie kompleksu rybonukleoproteinowego RNP (*Ryc. 1B*), na który składają się wirusowe RNA (ang. Viral RNA, vRNA) wraz z cząsteczkami białka nukleoproteiny (ang. Nucleoprotein, NP) oraz trzema podjednostkami RNA-zależnej wirusowej polimerazy RNA (ang. RNA dependent RNA polymerase, RdRp)⁹. Kompleks polimerazy tworzą trzy podjednostki – PB1, PB2 (ang. Polymerase basic protein 1 and 2) oraz PA (ang. Polymerase Acidic protein), z których każda pełni istotną rolę w trakcie replikacji wirusa⁸. Dokładnie funkcje pełnione przez te białka zostały opisane w *Tabeli 1*.

Oprócz białek budulcowych, wirus grypy wytwarza również białka niestrukturalne pełniące istotne funkcje w czasie cyklu replikacyjnego, między innymi regulując replikację i transkrypcję wirusowego RNA. Dotychczas zidentyfikowano białka niestrukturalne kodowane przez segmenty vRNA2 (białka PB1-N40, PB1-F2), vRNA3 (białka PA-X, PA-N155, PA-N182), vRNA7 (białko M42) oraz vRNA8 (białka NS1, NS2 (NEP) oraz NS3)^{24,26}.

1.2. Budowa kompleksu rybonukleinowego vRNP

Jak już wspomniano, na kompleks vRNP, oprócz cząsteczki wirusowego RNA, składają się trzy podjednostki wirusowej polimerazy RNA-zależnej – PB1, PB2 oraz PA, a także wielu kopii białka NP. Dotychczas zaproponowanych zostało kilka modeli budowy kompleksu rybonukleinowego (vRNP). Jeden z modeli budowy vRNP zakładał, że vRNA jest nawinięte na rdzeń zbudowany z białek NP w częstotliwości jednego białka NP na 24 nukleotydy vRNA⁹. Według tego schematu, wiązanie NP-vRNA jest niezależne od sekwencji, a do oddziaływania dochodzi między białkiem a resztami fosforanowymi RNA. Niemniej jednak, natura wiązania białka NP do vRNA w tym kompleksie sprawia, że RNA ma zdolność do oddziaływań z innymi cząsteczkami. Co więcej, RNA w kompleksie może również tworzyć struktury wyższego rzędu.

Wysokorozdzielcze analizy kompleksów za pomocą mikroskopii elektronowej cryo-EM wskazały złożoność oddziaływań NP-RNA oraz NP-NP w kompleksie⁴³⁻⁴⁵.

Wyłoniony model strukturalny wskazał, że cząsteczka podzielona jest na trzy fragmenty: rejon sparowanych końców 3' oraz 5' cząsteczki vRNA związanych z kompleksem polimerazy, rejon centralny przybierający strukturę helikalną oraz rejon drugiego końca przybierający formę^{44,45}.

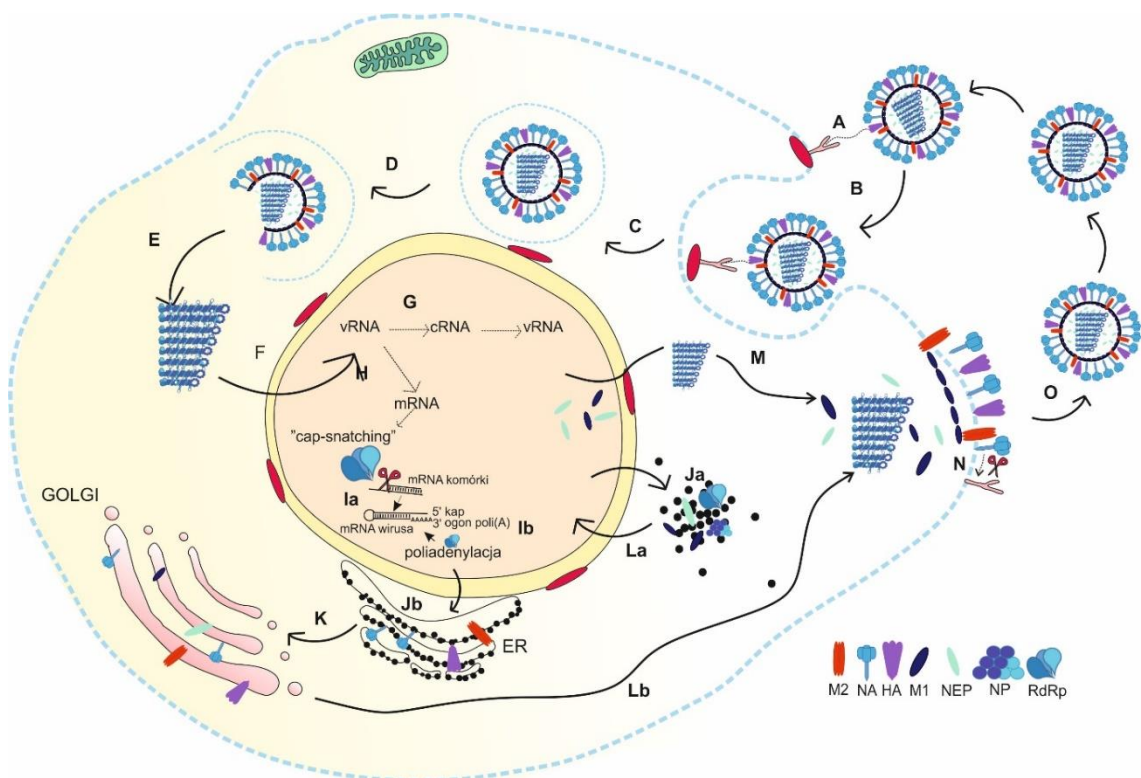
Wraz z rozwojem metod sekwencjonowania nowej generacji (ang. Next Generation Sequencing, NGS) otworzyła się możliwość badania oddziaływań RNA w kompleksie białkowym bezpośrednio w warunkach komórkowych *in cellulo*. Tak więc, drugi zaproponowany model, który uzyskano na podstawie badań vRNP za pomocą metody immunofotoszyczenia (HITS-CLIP) sprzężonego z NGS, potwierdził częściowe wypętlanie się vRNA poza obręb kompleksu vRNP⁴⁶. Ponadto, badania przeprowadzone przez *Lee et al.* potwierdziły, że model wiązania RNA w kompleksie vRNP dla różnych szczepów wirusa typu A – A/California/07/2009 oraz A/WSN/1933 jest odmienny. Kolejne eksperymenty potwierdziły, że wiązanie vRNA-NP jest niezależne od sekwencji vRNA, jednakże rejony wykazujące niski poziom wiązania, w szczególności w segmentach 1, 2 oraz 8 mają duże znaczenie w czasie cyklu replikacyjnego wirusa⁴⁷.

1.3. Cykl replikacyjny wirusa grypy

Cykl replikacyjny wirusa rozpoczyna się od wniknięcia do komórki gospodarza (**Ryc. 2**). W zainfekowanym organizmie dochodzi do rozwijania się infekcji górnych dróg oddechowych, która może przekształcić się w infekcję dolnych dróg oddechowych wywołując zagrażające życiu zapalenie płuc. Do zarażenia wirusem może dojść na trzy sposoby: 1) drogą kropelkową, poprzez bezpośredni kontakt z osobą zarażoną, 2) poprzez kontakt ze skażoną powierzchnią oraz 3) przez wdychanie powietrza, w którym znajdują się cząsteczki wirusa⁴⁸. Cykl replikacyjny wirusa grypy można podzielić na etapy: 1) rozpoznanie i wniknięcie do wnętrza komórki gospodarza; 2) transkrypcja i replikacja genomu wirusa oraz translacja białek wirusowych; 3) złożenie nowopowstałych cząsteczek vRNP i ich opłaszczenie, a następnie opuszczenie komórki gospodarza przez potomne wiriony.

Wnikanie wirionu do wnętrza komórki

Infekcja rozpoczyna się od rozpoznania receptorów kwasu sialowego na powierzchni komórki gospodarza. Dochodzi do wiązania α -2,3 lub α -2,6-sialowego pomiędzy cząsteczką wirionu a terminalną grupą kwasu sialowego obecnego w



Ryc. 2 Uproszczony schemat przedstawiający najważniejsze etapy cyklu replikacyjnego wirusa grypy typu A

receptorze powierzchniowym (**Ryc. 2A**). Proces ten jest katalizowany przez białko HA. Podczas tego etapu infekcji dochodzi do utworzenia wielu interakcji HA-SA, co znacznie wpływa na stabilność utworzonych wiązań⁴⁹. Białko HA jest homotrimerem, a każda podjednostka zbudowana jest z dwóch polipeptydów połączonych wiązaniem disiarczkowym: podjednostka HA1 posiadająca domenę wiążącą receptor (ang. Receptor Binding Site, RBS) specjalizującą się w rozpoznawaniu kwasu sialowego oraz podjednostka HA2, w skład której wchodzi peptyd fuzyjny⁸. Jednym z procesów decydujących o infekcyjności wirusa jest aktywacja prototypu białka HA. W procesie cięcia proteolitycznego, białko HA ulega aktywacji z nieaktywnej formy (HA0) do formy aktywnej, zdolnej do fuzji z błoną komórkową, składającej się z dwóch podjednostek H1 oraz H2²⁵. W zależności od momentu cyklu, do aktywacji białka HA dochodzi podczas opuszczania potomnego wirionu bądź przed endocytozą wirionu do wnętrza komórki. Proces cięcia jest katalizowany przez ludzkie proteazy transbłonowe HAT (ang. Human Airway Trypsin-like Protease) oraz TMPRSS2 (ang. Transmembrane protease serine S1 member 2), z czego tylko ta pierwsza ma zdolność do cięcia zarówno podczas wejścia jak i wyjścia wirionu z komórki, z kolei aktywność TMPRSS2 ograniczona jest do cięcia nowopowstałych białek H0⁵⁰. W przypadku szczepów grypy ptasiej białko HA ulega aktywacji w obrębie aparatu Golgiego, a za cięcie odpowiedzialna jest furyna²⁵.

Wnikanie cząstki wirusowej odbywa się na drodze endocytozy (**Ryc. 2B**). W środku pęcherzyka endocytotycznego jest środowisko o kwaśnym odczynie pH (5-6). Dzięki temu dochodzi do zależnej od pH zmiany konformacyjnej białka HA i wyeksponowany zostaje peptyd fuzyjny. Peptyd fuzyjny wchodzi następnie w interakcję z błoną infekowanej komórki inicjując tym samym fuzję cząsteczki wirusowej^{8,51} (**Ryc. 2C**). Kwaśne środowisko endosomu umożliwia oddzielenie kompleksów vRNP od stabilizujących je białek M1 (**Ryc. 2D**), a następnie ich uwolnienie do wnętrza cytoplazmy. Proces ten katalizowany jest przez kanał jonowy M2⁵² (**Ryc. 2E**).

Replikacja genomu oraz transkrypcja wirusowych mRNA

Cząsteczki vRNP są następnie transportowane do jądra komórkowego (**Ryc. 2F**), w którym następuje transkrypcja i replikacja genomu wirusa. Transport do jądra

odbywa się z wykorzystaniem jądrowego szlaku transportowego składającego się z wzajemnie kooperujących białek importyn- α oraz $-\beta$, a proces ten zapoczątkowuje rozpoznanie przez importynę- α sekwencji lokalizacji jądrowej (ang. Nuclear Localization Signal, NLS) obecnej licznie w białku NP^{8,25}.

Genom wirusa grypy to jednoniciowy RNA o ujemnej polarności. Oznacza to, że proces replikacji genomu w pierwszej kolejności wymaga przepisania vRNA w pełnej długości komplementarną cząsteczkę RNA (ang. complementary RNA, cRNA), która następnie posłuży jako matryca do syntezy potomnych cząsteczek vRNA (**Ryc. 2G**)⁸. Synteza cRNA katalizowana przez wirusową polimerazę RNA RdRp obecną w kompleksie vRNP i jest niezależna od starterów²⁵. Inicjacja transkrypcji vRNA jest możliwa dzięki strukturze promotora znajdującej się na sparowanych 3' i 5' końcach cząsteczek vRNA⁵³. W wyniku transkrypcji vRNA powstaje również drugi rodzaj cząsteczek – matrycowy RNA (ang. Messenger RNA, mRNA), który stanowi matrycę do syntezy białek wirusowych. Proces ten jest również katalizowany przez wirusową polimerazę (**Ryc. 2H**). Wobec tego, że RdRp nie ma zdolności do wytworzenia struktury kapu, wykorzystuje ona struktury kapu dostępne w komórkowych mRNA, tzw. cap-snatching (**Ryc. 2Ia**)⁸. Proces ten jest możliwy dzięki budowie wirusowej polimerazy, która składa się z trzech białek: PB1, PB2 oraz PA. Każde białko katalizuje inny proces: białko PB1 ma domenę wiążącą vRNA, białko PB2 ma domenę wiążącą strukturę kapu, z kolei białko PA ma domenę o właściwościach endonukleazy^{8,54}. RdRp jest również odpowiedzialna za syntezę ogonów poli(A) na 3' końcu cząsteczki mRNA, a badania wykazały, że sygnał do poliadenylacji znajduje się na matrycy vRNA i ma formę ogonów poli(U) (**Ryc. 2Ib**)⁵⁵.

Translacja białek wirusowych

Po wytworzeniu dojrzałych cząsteczek wirusowych mRNA, są one transportowane do cytoplazmy, gdzie następuje ich translacja przy użyciu maszynerii dostępnej w komórce gospodarza (**Ryc. 2J**). Białka HA, NA oraz M2 ulegają translacji katalizowanej przez rybosomy obecne w retikulum endoplazmatycznym (ER) (**Ryc. 2Jb**), natomiast białka PB1, PB2, PA, NP, NS1, NEP oraz M1 ulegają translacji w rybosomach cytoplazmatycznych (**Ryc. 2Ja**)²⁵. Białka wirusowe ulegają następnie obróbce potranslacyjnej w aparacie Golgiego (**Ryc. 2K**) Białko NP oraz trzy

podjednostki polimerazy wirusowej posiadające sekwencje lokalizacji jądrowej (ang. Nuclear Localization Signal, NLS) są następnie kierowane z powrotem do jądra komórkowego, gdzie uczestniczą w procesie transkrypcji oraz replikacji (**Ryc. 2La**)²⁵. Białko NS1 jest kierowane do jądra komórkowego, gdzie uczestniczy w inhibicji szlaków sygnalizacyjnych interferonu²⁵. Białka NEP oraz M2 również są kierowane do środka jądra komórkowego i uważa się, że mogą one uczestniczyć w eksporcie kompleksów vRNP z jądra do cytoplazmy komórkowej²⁵.

Białka HA, NA oraz M2 kierowane są w stronę błony komórkowej (**Ryc. 2Lb**), gdzie nastąpi pakowanie nowopowstałej cząsteczki wirusa^{1,56} (**Ryc. 2Lb**). Wyściełające wewnętrzną stronę wirionu białko M1 układa się po jej wewnętrznej stronie błony komórkowej i pełni funkcję w końcowym etapie składania cząsteczki wirusowej podczas jej zamykania i ostatecznego opuszczenia komórki⁸.

Łączenie wszystkich elementów w całość, czyli o procesie pakowania wirionu potomnego

Po utworzeniu potomnych cząsteczek vRNP w jądrze komórkowym, są one transportowane do cytoplazmy w kierunku błony komórkowej, gdzie następuje składanie potomnej cząsteczki wirusa (**Ryc. 2M**). W procesie transportu kompleksów vRNP uczestniczy komórkowe białko pełniące funkcje transporterowe Rab11, które wchodzi w interakcję z białkiem PB2²⁵. Specyficzna budowa genomu wirusa wymaga skomplikowanych mechanizmów koordynujących właściwe upakowanie cząsteczek potomnych. Dotychczas zaproponowano dwa modele mechanizmu pakowania cząsteczek vRNP w wirion potomny: 1) losowe łączenie się vRNP i przypadkowe powstawanie wirulentnych cząsteczek potomnych po spotkaniu kopii wszystkich vRNP oraz 2) selektywne i nieprzypadkowe łączenie się cząsteczek vRNP, które jest kierowane przez odpowiednie sekwencje sygnalizacyjne⁵⁷. Obecna wiedza pozwoliła potwierdzić, że pakowanie cząsteczek vRNP jest procesem nieprzypadkowym i całkowicie kontrolowanym przez składowe wirusa. Badania wykazały, że cząsteczki vRNA na końcach 3' i 5' posiadają unikalne dla każdego segmentu sekwencje sygnałowe odpowiedzialne za oddziaływania typu RNA-RNA między segmentami⁵⁸⁻⁶⁰. Co więcej, udowodniono, że sygnały pakowania wirionu zlokalizowane są nie tylko

w obrębie rejonów 3' i 5' UTR cząsteczek, ale również w obrębie sekwencji kodujących¹⁸.

Sekwencje pełniące rolę sygnałów pakowania i składania wirionu potomnego są specyficzne dla danego szczepu wirusa IAV¹⁸. Prawidłowe upakowanie wirionu potomnego jest również zależne od sekwencji aminokwasowej białka NP⁵⁹. Obraz uzyskany za pomocą wysokorozdzielczej mikroskopii elektronowej ukazał również sposób w jaki porządkują się vRNP we wnętrzu wirionu⁶¹.

Za proces opuszczenia wirionów z komórki odpowiedzialne jest białko NA⁴⁹. NA katalizuje hydrolizę wiązania glikozydowego występującego w obrębie kwasu sjałowego, uniemożliwiając utworzenie wiązania pomiędzy kwasem sjałowym a białkiem HA (**Ryc. 2N**), tym samym białko NA umożliwia uwolnienie wirionu potomnego, zapobiegając agregacji nowopowstałych cząsteczek na powierzchni komórki⁴⁰ (**Ryc. 2O**).

Konsekwencje zaburzenia procesu pakowania wirionu potomnego

W pełni infekcyjna cząsteczka wirusa zawiera komplet wszystkich segmentów vRNA o pełnej długości¹⁸. Niekiedy dochodzi do zaburzeń cyklu replikacyjnego, a także procesu pakowania wirionu potomnego, w wyniku czego powstają nieinfekcyjne cząsteczki wirusa. Cząsteczki te nazywane są defektywnymi cząsteczkami interferującymi (ang. Defective Interfering Particles, DIP)⁶². Cząsteczki DIP charakteryzują się brakiem jednego lub kilku segmentów, bądź delecją w obrębie cząsteczek vRNA. Chociaż cząsteczki defektywne mają zdolność do wnikania do wnętrza komórki, to w wyniku infekcji nie powstaje infekcyjna cząsteczka potomna⁶³.

Powstanie w pełni infekcyjnej cząsteczki wirusa wymaga właściwego upakowania kompleksów vRNP do wirionu potomnego. Jak już wspomniano, pierwsze teorie dotyczące wzoru pakowania się cząsteczek vRNP do wirionu potomnego zakładały, że proces ten zachodzi na zasadzie losowego i niespecyficznego łączenia się vRNP⁶⁴. Obecnie uznawany model według którego cząsteczki vRNP wirusa grypy typów A i B łączą się na zasadzie wzoru „1+7”. Model ten zakłada centralnie umieszczoną cząsteczkę vRNP otoczoną 7 cząsteczkami vRNP i został potwierdzony za pomocą wysokorozdzielczej mikroskopii elektronowej⁶⁵. W przypadku wirusa grypy typów A i B musi być to więc 8 różnoimiennych cząsteczek vRNP. Model upakowania wirionu

potomnego „1+7” jest taki sam w przypadku szczepów wirusa grypy typów C i D, które posiadają jedynie 7 segmentów vRNA⁶⁶. Upakowanie dodatkowej, ósmej cząsteczki vRNP nie przyczynia się do obniżenia wirulentności wirusa. To wskazuje na to, że ten sposób upakowania jest najbardziej optymalny dla wirusa i nie zależy od ilości segmentów składających się na jego genom⁶⁶.

W badaniach, w których indukowano mutacje różnych segmentach vRNA, wykazano, że nawet jednonukleotydowa mutacja synonimiczna w obrębie sekwencji kodujących odpowiedzialnych za upakowanie w pełni wirulentnych cząsteczek wirusa, powoduje zaburzenie tego procesu⁵⁹. Mutacje w obrębie sekwencji aminokwasowej białka NP również wpływają na powstawanie niewirulentnych form wirusa⁵⁹. Dowiedziono również, że zaburzenie procesu ubikwitynacji białka M2 doprowadza do powstawania niekompletnych cząsteczek wirusa⁶⁷. Badanie zaburzenia pakowania wirionu potomnego mogą przyczynić się do opracowania nowych metod inhibicji wirusa na poziomie białek.

Pakowanie wirionu potomnego jest kierowane także na poziomie RNA. Proces jest koordynowany przez sygnały pakowania, które ulokowane są w obrębie sekwencji vRNA. Sygnały pakowania są to specyficzne dla poszczególnych segmentów RNA sekwencje, które koordynują uporządkowanie i pakowanie się kompleksów vRNP do wirionu. Potwierdzenie istotności sekwencji w rejonach stanowiących sygnały pakowania było możliwe dzięki wprowadzaniu mutacji w obrębie tych sekwencji prowadzących do powstawania nieaktywnych wirionów potomnych^{68,69}. Porównanie wyników badań dotyczących sygnałów pakowania dla różnych segmentów IAV wskazały, że wykazują one pewien stopień konserwatywności¹⁸. Jednakże, badania sugerują również, że część sygnałów pakowania jest specyficzna dla danego szczepu IAV^{18,70}.

1.4. Białka niestrukturalne kodowane przez genom wirusa grypy i ich wpływ na cykl replikacyjny wirusa

Białka niestrukturalne kodowane przez segment 8 vRNA

Najmniejszy z segmentów, segment 8 vRNA, koduje dwa białka niestrukturalne – NS1 (ang. Non-structural protein1, NS1), oraz NEP/NS2 (ang. Non-structural protein 2, NS2; znane również jako NEP ang. Nuclear Export Protein), które koordynują

kluczowe dla cyklu życiowego wirusa procesy (**Tabela 1**). Białko NS1 jest białkiem zapewniającym wirulencję wirusa i wpływa bezpośrednio oraz pośrednio na odpowiedzi immunologiczne zainfekowanych komórek. Lista szlaków na które ma wpływ NS1 jest bardzo długa, a wszystkie interakcje mają na celu jak najwydajniejszy proces namnażania wirusa w komórce⁷¹. Aktywność białka NS1 ma miejsce zarówno w jądrze komórkowym jak i cytoplazmie. Wyróżnić można 5 głównych procesów, na które ma wpływ NS1: wyciszanie komórkowych RNA, aktywność polimerazy wirusowej, procesowanie i eksport komórkowych mRNA, antagonizm interferonu i antywirusowej odpowiedzi komórkowej, a także regulacja translacji wirusowych mRNA^{36,72,73}. Niedawne doniesienie potwierdziło również wpływ tego białka na proces pakowania potomnego wirionu⁷³. Wielozadaniowość białka NS1 wynika z obecności domeny wiążącej cząsteczki RNA (ang. RNA-binding domain), jak i domeny wiążącej białka (ang. Effector domain)⁷⁴. Domena wiążąca RNA umożliwia białku NS1 interakcję z cząsteczkami RNA wchodzącymi w skład kompleksu splicosomu takimi jak U6 i U6atac, a tym samym wywołuje inhibicję splicingu ssaczych pre-mRNA⁷⁵. Wykazano także, że jest zaangażowana w regulację splicingu wirusowych mRNA segmentu 7 i 8, w wyniku którego powstają białka M2 oraz NEP⁷⁶. Ponadto, NS1 wiąże się do traktów poli(A) na 3' końcu mRNA białka NS2, a także innych mRNA w tym komórkowych, w efekcie zaburzając ich eksport jądrowy⁷⁷.

Oprócz pobudzania regulacji ekspresji własnych genów, u podłoża wirulencji wirusów, w tym także wirusa grypy, leży zdolność do osłabiania odpowiedzi immunologicznej zainfekowanej komórki. Białko NS1 upośledza zdolność komórki do obrony przeciwwirusowej poprzez hamowanie procesów dojrzewania i translacji mRNA interferonu (IFN), a także poprzez zapobieganie aktywacji białek IFN-zależnych^{71,78}. NS1 wpływa na ekspresję IFN poprzez oddziaływanie z czynnikami regulacji ekspresji IFN- α/β - czynnikiem IRF-3 (ang. Interferon regulatory factor 3) oraz czynnikiem NF- κ B^{79,80}. Z kolei przykładem oddziaływania z innymi białkami odpowiedzi antywirusowej jest interakcja z indukowanym interferonem białkiem obrony antywirusowej PKR (ang. RNA-dependent protein kinase)^{81,82}, a także białkiem RIG-I (ang. Retinoic acid-inducible gene I)³⁵. Białko RIG-I jest aktywowane przez wirusowe dsRNA, pośrednio ingeruje w szlak sygnalizacyjny IFN1, dlatego też

stanowi jeden z najistotniejszych czynników immunologicznej odpowiedzi komórkowej w czasie infekcji wirusowej. Białko NS1 wpływa także na inhibicję szlaku sygnalizacyjnego OAS/RNase L, który bierze udział min. w degradacji wirusowych vRNA i mRNA⁸³.

Drugie białko niestrukturalne, nazywane białkiem transportu jądrowego (NEP), odpowiedzialne jest za komunikację z importynami, przez które transportowane są kompleksy vRNP z jądra komórkowego, a także wspomaga uwalnianie wirionów potomnych poprzez stymulację komórkowej F1F0 ATPazy³⁸. Białko NEP zapobiega akumulacji vRNA, cRNA oraz mRNA na terenie jądra³⁸. Bierze również udział w procesie opuszczania potomnych cząsteczek wirusa^{37,38}.

Stosunkowo niedawno odkryto również trzecie białko niestrukturalne (ang. Non-structural protein 3, NS3) powstające w wyniku alternatywnego splicingu mRNA segmentu 8³⁹. Dotychczas nie poznano dokładnej funkcji pełnionej przez to białko^{24,26,39}.

Białka niestrukturalne kodowane przez segment 2 vRNA

Białka niestrukturalne kodowane przez segment vRNA2 to pełniące funkcje proapoptyczne białko PB1-F2⁸⁴ oraz białko PB1-N40 (*Tabela 1*) balansujące ekspresję białek PB1 oraz PB1-^{26,85}. Białko PB1-F2 bierze udział w wielopoziomowej inhibicji szlaków sygnalizacyjnych będących odpowiedzią immunologiczną zainfekowanego organizmu^{86,87}. Białko to uznawane jest za ważny czynnik wirulencji, gdyż ingeruje w szlaki sygnalizacyjne, zależnego od mitochondrium, powstawania IFN1 oraz szlaków aktywowanych przez IFN1⁸⁸. Co więcej, wykazuje również właściwości proapoptyczne⁸⁴. Dowiedziono, że PB1-F2 pełniło kluczową rolę w patogenności szczepu grypy A/H1N1 z 1918 roku, który wywołał pandemię znaną jako „grypa hiszpanka”, a u podstaw wyjątkowej patogenności tego szczepu leżała zdolność do supresji helikazy DDX3⁸⁹. Poprzez blokowanie transbłonowego mitochondrialnego białka MAVS (ang. Mitochondrial antiviral-signaling protein), PB1-F2 ingeruje w szlak sygnalizacyjny prowadzący do produkcji IFN1⁸⁸. Niektóre właściwości tego białka i jego oddziaływanie na poszczególne ścieżki antyzapalnej odpowiedzi komórkowej, są specyficzne dla danego podtypu wirusa. Na przykład, w przypadku podtypu H7N9, białko PB1-F2 ma zdolność do agregacji na błonie mitochondrium i

ingerencji w oddziaływania TRIM31-MAVS, a w konsekwencji do agregacji i degradacji białka MAVS⁹⁰. Białko PB1-F2 blokuje również powstanie IFN1 poprzez wpływ na ścieżkę sygnalizacyjną kinazy I- κ B/NF- κ B⁹¹. W odpowiedzi na infekcję wirusową, komórka makrofaga rozpoznaje agregujące w macierzy zewnątrzkomórkowej białko PB1-F2, co z kolei wpływa na produkcję czynnika zapalnego NLRP3 oraz produkcję interleukin – IL-1 β oraz IL-18. Dowiedziono, że obecne w komórce białko PB1-F2 jest odpowiedzialne za supresję ekspresji NLRP3 i nadprodukcję IL-1 β oraz IL-18, a mechanizm ten polega na inhibicji oddziaływań białka NLRP3-MAVS⁹².

Białko PB1-N40 (718AA) jest izoformą białka PB1 (718AA), które stanowi niepełnej długości białko PB1. Badania wskazują, że wpływa ono na poziom ekspresji białek wirusowych PB1 oraz PB2-F2 poprzez oddziaływanie z wirusową polimerazą⁸⁵.

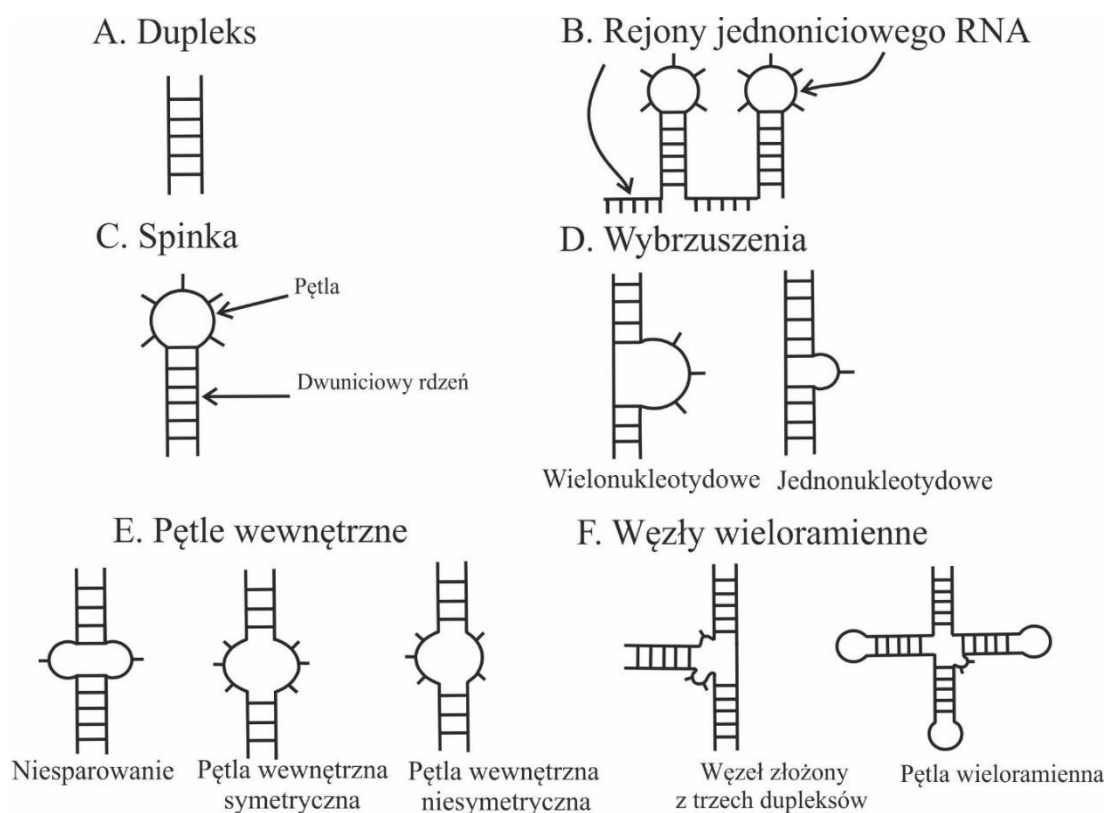
Białka niestrukturalne kodowane przez segment 3 vRNA

Białka niestrukturalne kodowane przez segment 3 vRNA również mają wpływ na infekcyjność wirusa. Najlepiej poznanym białkiem jest PA-X (*Tabela 1*), które odpowiedzialne jest za cięcie komórkowych transkryptów mRNA, wpływa na replikację genomu wirusa, patogenność i odpowiedź immunologiczną komórki^{19,24,26,27,86,93}. Mniej poznanymi białkami kodowanymi przez vRNA3 są izoformy białka PA – PA-N155 oraz PA-N182 (*Tabela 1*)^{24,26,28,86,94}. Dotychczas nie określono dokładnej funkcji pełnionej przez te białka, chociaż badania przeprowadzone na myszach wykazały, że mutanty wirusa nie posiadające tych izoform charakteryzowały się mniejszą zjadliwością²⁸.

1.5. Struktury drugorzędowe RNA wirusa grypy

RNA są cząsteczkami biologicznymi, które klasyfikuje się ze względu na pełnione przez nie funkcje. Oprócz RNA kodujących białka, wyróżnić można cały szereg RNA niekodujących (ang. non-coding RNAs, ncRNAs). Do niekodujących RNA zaliczane są RNA pełniące funkcje regulatorowe na wszystkich etapach ekspresji genów⁹⁵. RNA tworzą skomplikowaną sieć interakcyjną z wieloma typami cząsteczek takimi jak DNA, RNA, białkami oraz różnego rodzaju metabolitami⁹⁶.

U podstawy funkcjonalności RNA jest jego budowa. RNA przyjmuje budowę drugorzędową, pełną oddziaływań o charakterze lokalnym jak i dalszego zasięgu w obrębie której występują zarówno dwuniciowe rejony tzn. duplekty (ang. duplex) (*Ryc. 3A*) jak i niesparowane rejony jednoniciowe (ang. single stranded RNA, ssRNA) (*Ryc. 3B*). Podstawową jednostką struktury drugorzędowej są motywy strukturalne, na które składają się rejony sparowanych nukleotydów, a także miejsc niesparowanych tworzących różnego rodzaju pętle (ang. loops), wybrzuszenia (ang. bulges) lub dłuższe fragmenty jednoniciowego RNA⁹⁷. Rejony te są często istotnym miejscem oddziaływań RNA z innymi cząsteczkami⁹⁷.



Ryc. 3 Przykłady motywów strukturalnych występujących w strukturze drugorzędowej

Podstawowym motywem strukturalnym RNA jest motyw helisy. W helikalnych motywach RNA często występują niesparowania nukleotydów skutkujące powstawaniem wybruszeń takich jak jedno- lub wielonukleotydydowe wybrzuszenia jednostronne (ang. bulge loops) (*Ryc. 3D*). W przypadku, kiedy dochodzi do wybruszeń dwustronnych tworzą one pętle wewnętrzne (ang. internal loops). Pętle wewnętrzne mogą mieć charakter pojedynczych niesparowań (ang. mismatches), jak i wielonukleotydydowych niesparowań o charakterze symetrycznym i niesymetrycznym

(**Ryc. 3E**). W obrębie helisy często występują tandemowe niesparowania (ang. tandem mismatches) np. typu G-U oraz G-A. Jednym z najbardziej stabilnych termodynamicznie motywów helikalnych jest motyw spinki zbudowanej z dupleksu (ang. hairpin stem) zakończonego pętlą (ang. hairpin loop) (**Ryc. 3C**). Dupleksy mogą się łączyć tworząc węzły składające się z trzech lub czterech dupleksów (ang. three/four stem junctions) tworząc pętle wieloramienne (ang. multibranch loops) (**Ryc. 3F**).

Struktura wyższego rzędu powstaje w wyniku oddziaływania cząsteczki RNA z innymi cząsteczkami biologicznymi takimi jak inne RNA, DNA, białka czy metabolity^{98,99,96,100}. Funkcjonalność RNA definiowana jest właśnie jego strukturą drugorzędową i motywami strukturalnymi w jej obrębie, a także zdolnością do oddziaływań z innymi cząsteczkami. W wyniku takich oddziaływań RNA przybiera struktury trzecio- i czwartorzędowe, będące w istocie podstawą do powstawania w pełni funkcjonalnej, aktywnej biologicznie cząsteczki¹⁰¹. Należy także nadmienić, że aby w pełni zrozumieć funkcjonalność cząsteczki RNA, motywy strukturalne należy analizować nie tylko pod kątem struktury drugorzędowej, ale także w formie trójwymiarowej biorąc pod uwagę ułożenie konformacyjne rdzenia cząsteczki, występujące w jej obrębie wiązania wodorowe oraz oddziaływania warstwowe⁹⁷. Co więcej, RNA oprócz oddziaływań intermolekularnych, ma również zdolność do oddziaływań intramolekularnych bliskiego i dalekiego zasięgu w obrębie swojej cząsteczki. Tak więc, oprócz oddziaływań z innymi cząsteczkami biologicznymi, lokalne motywy strukturalne mają zdolność do tworzenia oddziaływań wyższego rzędu w obrębie tej samej cząsteczki RNA. Przykładem elementów struktury trzeciorzędowej jest struktura pseudowęzła (ang. pseudoknot), która powstaje poprzez oddziaływanie nukleotydów pętli struktury typu spinka z fragmentami jednoniciowymi spoza tej pętli¹⁰². Co istotne, motyw pseudowęzła jest powszechny dla wirusów RNA i pełni rolę w trakcie procesów związanych z replikacją genomu oraz translacją białek wirusowych^{103,104}.

Struktura RNA zależna jest nie tylko od sekwencji, ale także od zdolności parowania zasad w jej obrębie. Co istotne, struktura RNA bezpośrednio determinuje pełnione przez nie funkcje. Tym samym zachowawczość pewnych motywów strukturalnych wskazuje na pełnienie przez nie funkcji biologicznych.

1.5.1. Znane motywy strukturalne vRNA wirusa grypy typu A

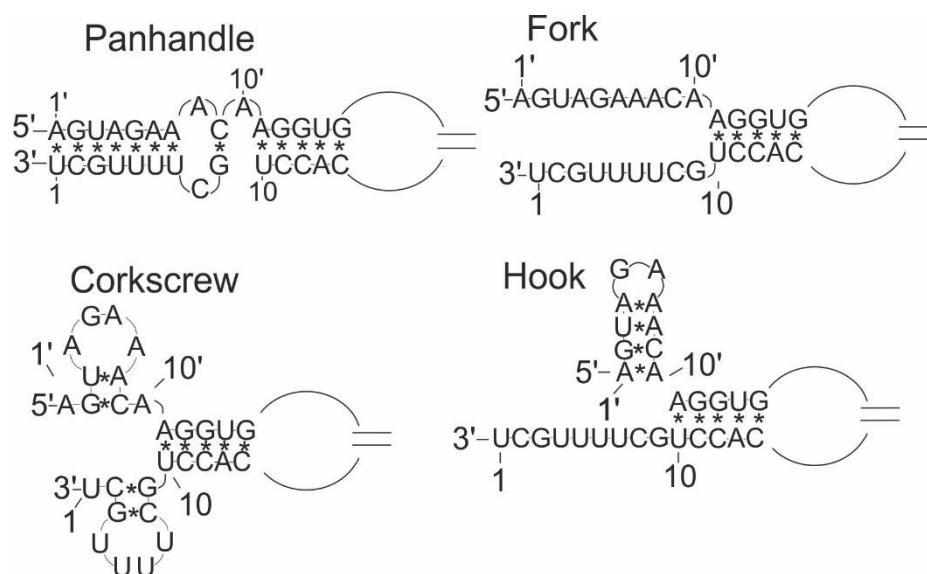
RNA wirusa grypy dzielimy na trzy typy cząsteczek – genomowy vRNA, a także powstające w czasie cyklu replikacyjnego wirusa, komplementarny do vRNA – cRNA oraz cząsteczki mRNA¹. Struktura RNA odgrywa znaczącą rolę w czasie procesów takich jak transkrypcja i replikacja genomu wirusowego, składanie kompleksów vRNP, translacji białek wirusowych, a także podczas procesu splicingu wirusowego mRNA¹⁷. W obrębie każdej z tych cząsteczek wyróżnić można zachowawcze, wspólne dla wielu szczepów, motywy strukturalne.

Na zachowawczość motywów strukturalnych RNA ma wpływ konserwatywność na poziomie sekwencji. Badania porównawcze sekwencji poszczególnych segmentów vRNA w obrębie grypy typu A wykazały istnienie sekwencji konserwatywnych¹⁰⁵. Co istotne, zachowawczość sekwencji jest nie tylko różna dla każdego segmentu vRNA, ale jest charakterystyczna dla danego podtypu wirusa. W przypadku szczepów podtypu A/H1N1 najmniejszą konserwatywność sekwencyjną zaobserwowano w przypadku segmentów vRNA4 i vRNA6. Z kolei najwięcej rejonów konserwatywnych zidentyfikowano dla segmentów vRNA7, vRNA8 i vRNA5. Co ciekawe, w przypadku podtypu A/H7N7 segment vRNA7 okazał się najbardziej zmienny, a najbardziej konserwatywny sekwencyjnie okazał się segment vRNA6.

Zidentyfikowano też konserwatywne motywy strukturalne zarówno w nici (+) RNA jak i (-) RNA, o zmiennej sekwencji, zachowującej jednak strukturę danego motywu. Takie mutacje kompensacyjne (mutacje dwupunktowe, które zachowują parowanie), oraz konsensusowe mutacje (mutacje jednopunktowe, które zachowują parowanie) są potwierdzeniem istnienia danego motywu strukturalnego RNA i jego prawdopodobnej istotności dla cyklu replikacyjnego wirusa. Konserwatywne motywy strukturalne pełniące funkcje w czasie cyklu replikacyjnego wirusa stanowią dobry target do celowanej inhibicji na poziomie RNA.

Struktura promotora wirusowej polimerazy RNA-zależnej

Jednym z najlepiej poznanych motywów strukturalnych RNA wirusa grypy jest motyw promotora rozpoznawanego przez polimerazę RdRp. Struktura promotora łączy ze sobą końce 3' i 5' vRNA związanego z białkami w kompleksie vRNP. Rejon ten pełni istotne funkcje w czasie transkrypcji oraz replikacji genomu wirusa¹⁷. Struktura RNA promotora od wielu lat jest tematem dysput naukowych i dotychczas zaproponowano kilka modeli struktury drugorzędowej: *panhandle*, *fork*, *corkscrew* oraz *hook* (**Ryc. 4**).^{53,106–109} Badania polegające na wprowadzeniu mutacji w formie



Ryc. 4 Proponowane struktury promotora vRNA rozpoznawanego przez kompleks wirusowej polimerazy RNA-zależnej RdRp: A. Struktura panhandle, B. Struktura corkscrew, C. Struktura fork, D. Struktura hook. Na podstawie rysunku Ferhadian et al. (Frontiers in Microbiology, 2018).

addycji nukleotydu U, skutkującej powstaniem sparowania pomiędzy nukleotydem A10' i U potwierdziły, że niesparowany nukleotyd A10' (**Ryc. 4**) stanowi istotny element strukturalny mający wpływ na proces transkrypcji¹⁰⁷. Wydaje się, że struktura końców 5' oraz 3' vRNA może formować się w inny sposób w wolnej od białek cząsteczce vRNA, a inaczej w kompleksie vRNP¹¹⁰. W czasie cyklu replikacji struktura promotora przechodzi przez zmiany konformacyjne, które umożliwiają przyłączenie się polimerazy vRdRp^{111,112}.

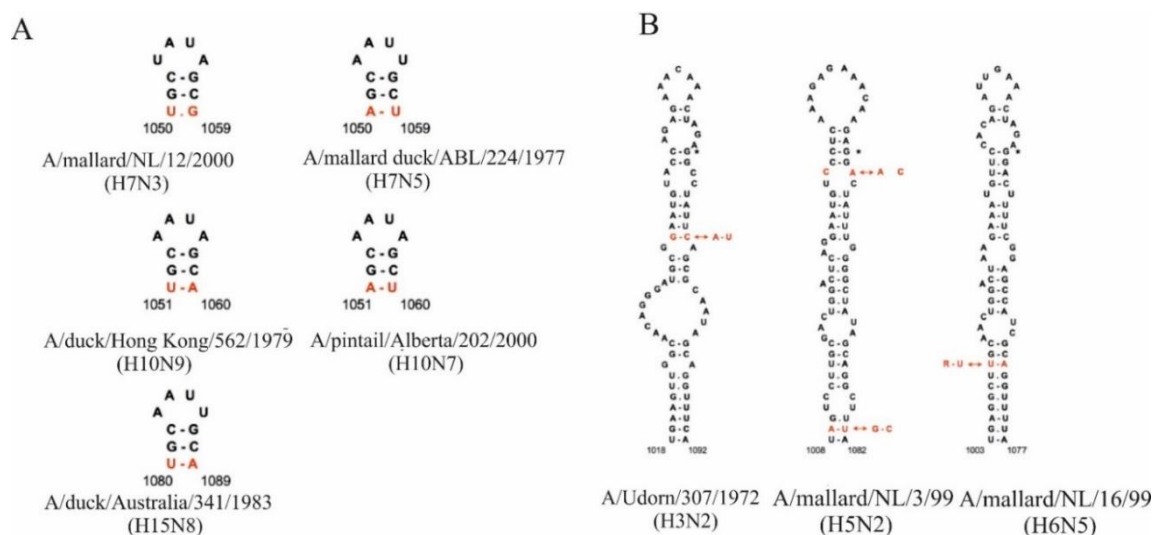
Konserwatywne motywy strukturalne vRNA

Dotychczas zaproponowano modele struktury drugorzędowej, opracowanych na podstawie przewidywań bioinformatycznych oraz badań *in vitro*, kilku segmentów

vRNA.. Modele te stanowią podstawę do opracowania struktury poszczególnych segmentów vRNA w środowisku biologicznym. Na dzień dzisiejszy najlepiej poznanymi cząsteczkami są segment 4, 5, 7 oraz 8 vRNA.

Segment 4 vRNA

Gulyaev et al. przeprowadził badania bioinformatyczne dotyczące zachowawczości motywów strukturalnych vRNA na podstawie kilku tysięcy sekwencji dostępnych w bazie danych¹¹³. Badania wykazały, że segment kodujący białko HA, nawet w obrębie jednego podtypu, posiada nie tylko różny stopień zgodności sekwencyjnej, a także prawdopodobieństwa parowania zasad w obrębie struktury RNA¹¹³. Autorzy wskazali, że zachowawczość większości motywów jest specyficzna dla danego podtypu wirusa, np. spinka formująca się w rejonie 794-839 nt jest charakterystyczna dla podtypu H3N2. Wyjątek stanowią motywy występujące w sekwencji kodującej miejsce cięcia nieaktywnego białka HA do aktywnej formy H1 oraz H2, które są wspólne dla szczepów reprezentujących kilka podtypów wirusa. Pierwszym z nich jest krótka spinka występująca w podtypie H7 (rejon 1050-59 nt), H10 (rejon 1051-60 nt) oraz H15 (rejon 1080-89 nt) (**Ryc. 5A**). Drugim natomiast, jest homologiczny motyw wspólny dla podtypów H3 (rejon 1018-92 nt), H5 (rejon 1008-82 nt) oraz H6 (rejon 1003-1077 nt) (**Ryc. 5B**). Badania polegające na wprowadzeniu

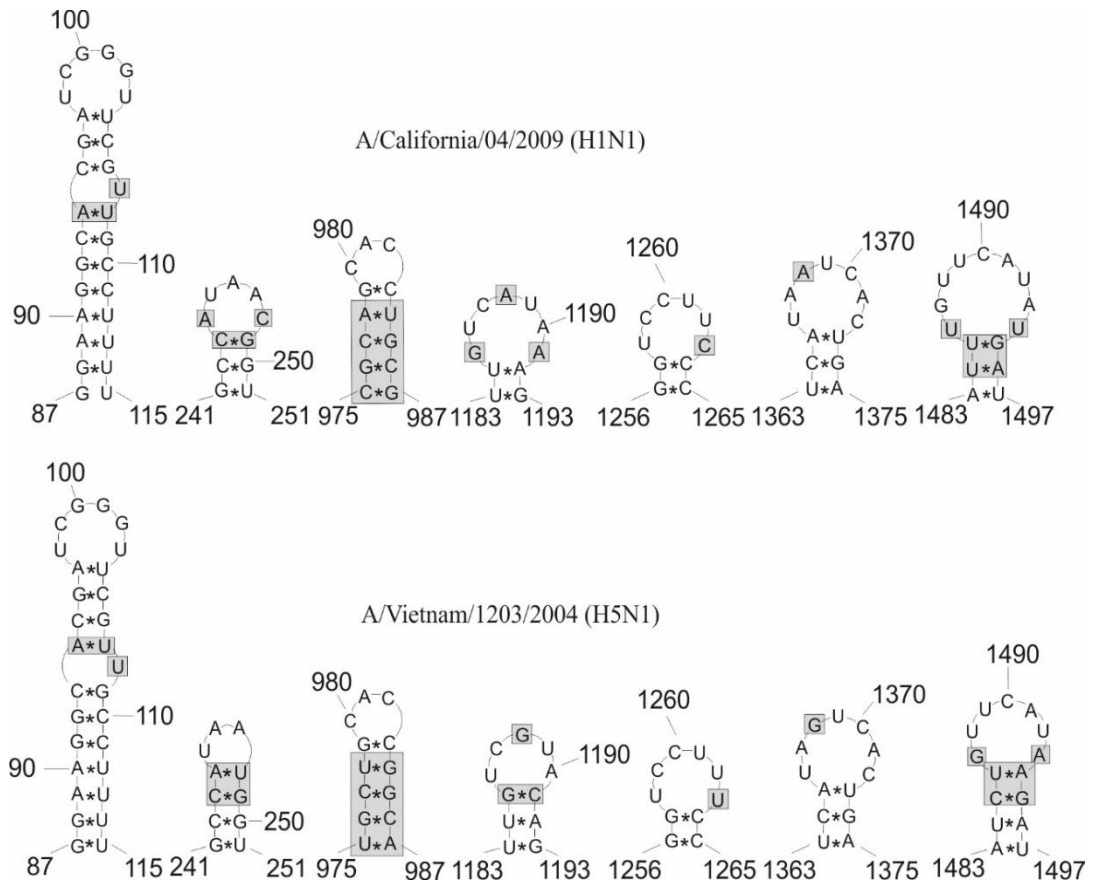


Ryc. 5 A. Spinka formująca się w segmencie 4 vRNA w rejonie sekwencji kodującej miejsce cięcia białka HA do aktywnej formy H1 i H2 przewidziana jako konserwatywna dla wielu podtypów wirusa grypy typu A. **B.** Motyw homologiczny przewidziany dla podtypów H3, H5 i H6, formujący się w miejscu cięcia białka HA.

mutacji synonimicznych w obrębie tych struktur skutkowały obniżeniem miana wirusa, tym samym wskazując na istotność tych rejonów w cyklu życiowym wirusa. Dotychczas nie potwierdzono jednak istnienia motywów, które byłyby wspólne dla wszystkich podtypów wirusa. Ze względu na niższą konserwatywność strukturalną, segment vRNA4 stanowi mniej efektywny target do celowanej inhibicji wirusa IAV.

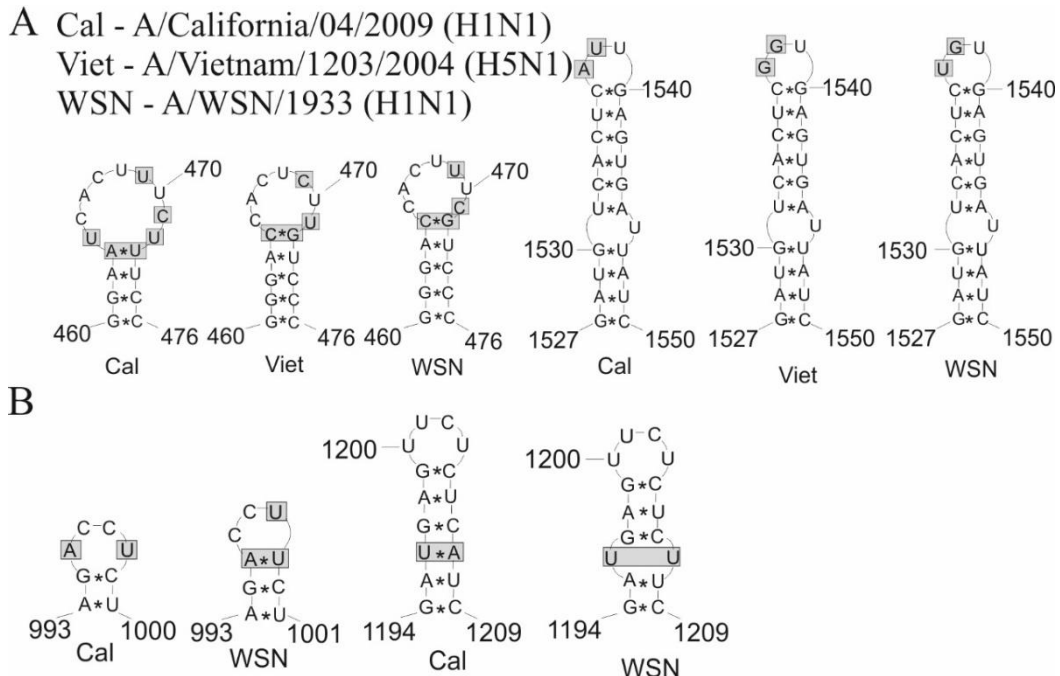
Segment 5 vRNA

Przeprowadzone przez naszą grupę, badania struktury vRNA5 pandemicznego szczepu A/Vietnam/1203/2004 (H5N1) w warunkach *in vitro* wykazały, że cząsteczka ta charakteryzuje się dużym stopniem ustrukturyzowania tworząc liczne motywy strukturalne zgrupowane w trzy główne domeny¹². Obliczenia bioinformatyczne na podstawie 18 500 sekwencji dostępnych w bazie danych wskazały, że struktura vRNA5 posiada 87% konserwatywności wśród szczepów grypy typu A¹². Nasze



Ryc. 6 Motywy strukturalne zidentyfikowane jako wspólne w cząsteczce vRNA5 szczepów A/California/04/2009 oraz A/Vietnam/1203/2004 w warunkach *in vitro*. Motywy strukturalne są zachowawcze pomimo różnic na poziomie sekwencji (zaznaczone kolorem szarym).

badania potwierdziły, że struktura drugorzędowa vRNA5 szczepu A/California/04/2009 (H1N1) (zgodność sekwencji pomiędzy powyższymi szczepami wynosi 83,3%) określona na podstawie badań eksperymentalnych *in vitro* posiada tę samą architekturę w ujęciu globalnym¹¹. Wspólne dla obu szczepów motywy strukturalne typu spinki zostały zidentyfikowane w rejonach 87-115 nt, 975-987 nt, 1256-1265 nt, 1363-1375 nt oraz 1527-1550 nt (**Ryc. 6** Motywy strukturalne zidentyfikowane jako wspólne w cząsteczce vRNA5 szczepów A/California/04/2009 oraz A/Vietnam/1203/2004 w warunkach *in vitro*. Motywy strukturalne są zachowawcze pomimo różnic na poziomie sekwencji (zaznaczone kolorem szarym).). Oprócz tego, zidentyfikowano również motywy strukturalne, w obrębie których zaobserwowano niewielkie różnice w parowaniu zasad. Są to motywy występujące w rejonach 241-251 nt, 460-476 nt, 1183-1193 nt oraz 1483-1497 nt (**Ryc. 6 i 7A**). Dwa motywy strukturalne występujące w rejonach 5' (87-115 nt) oraz 3' (1483-1497 nt) zostały zidentyfikowane jako terminalne sygnały pakowania wirionu potomnego¹¹⁴. Najnowsze badania struktury w warunkach biologicznych *in virio* oraz *ex virio* szczepu A/WSN/33 (H1N1) wykazały, że część motywów strukturalnych, przewidzianych w naszych badaniach *in vitro*, występuje również w środowisku

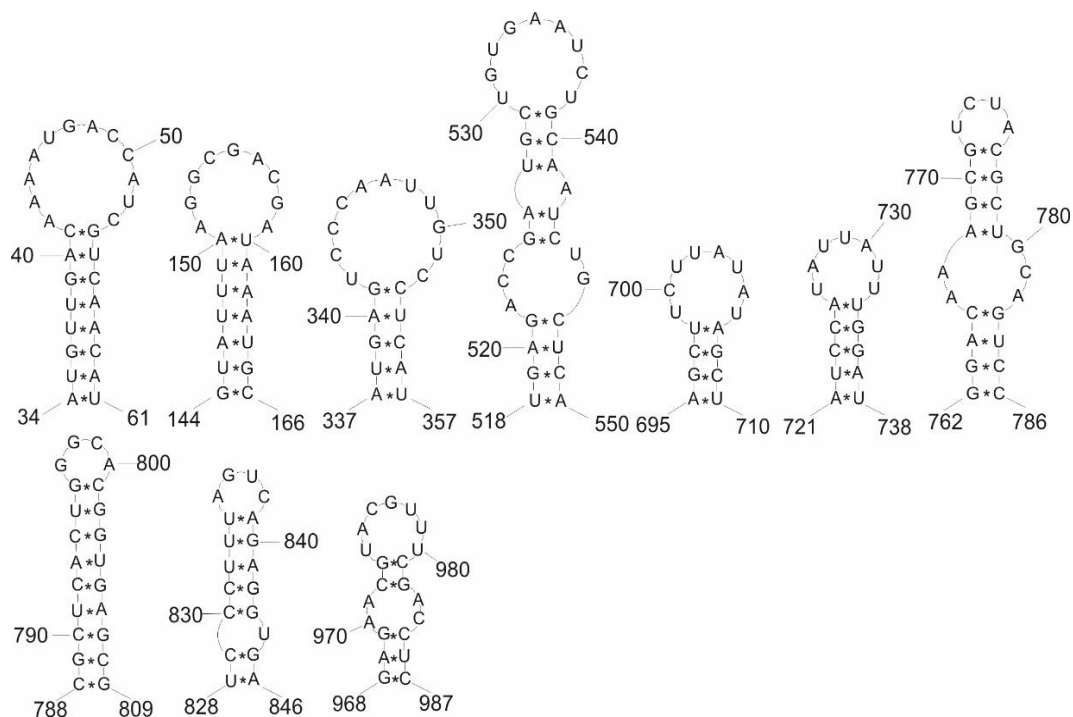


Ryc. 7 A. Motywy strukturalne vRNA5 wspólne dla wszystkich trzech analizowanych szczepów. **B.** Motywy strukturalne wspólne dla szczepów podtypu H1N1. Na szaro zaznaczone różnice w sekwencji i/lub parowaniu zasad.

biologicznym¹¹⁵. W strukturze vRNA5 szczepu A/WSN/33 zidentyfikowano cztery motywy strukturalne obecnie w strukturze vRNA5 szczepu A/California/04/2009 w rejonach 460-476 nt, 993-1001 nt, 1194-1209 nt, 1527-1550 nt (**Ryc. 7A,B**). Autorzy badań potwierdzili również występowanie w rejonie 79-154 nt struktury pseudowęzła, który został wcześniej zaproponowany dla vRNA5¹¹⁶. Warto zaznaczyć, że występowanie motywu pseudowęzła nie wyklucza formowania się spinki w rejonie 87-115 nt, którą powstaje się w rejonie uznanym za istotną w czasie pakowania wirionu.

Segment 7 vRNA

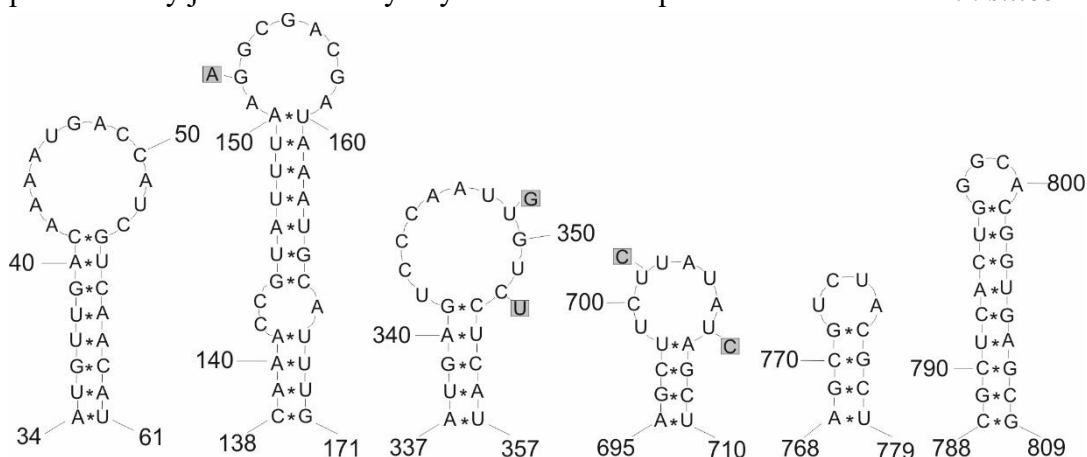
Struktura drugorzędowa segmentu 7 vRNA szczepu A/Vietnam/1203/2004 (H5N1) została zaproponowana przez naszą grupę¹³. Predykcję struktury przeprowadzono w oparciu o dane eksperymentalne pochodzące z mapowania chemicznego vRNA7 w warunkach *in vitro* oraz sprzężone z nimi obliczenia konserwatywności parowania zasad, na podstawie dostępnych w bazie danych, sekwencji wielu szczepów wirusa grypy typu A. Najbardziej konserwatywne motywy strukturalne zidentyfikowano w rejonach 34-61 nt, 144-166 nt, 337-357 nt, 518-550



Ryc. 8 Motywy strukturalne vRNA7 szczepu A/Vietnam/1203/2004 (H5N1) o najwyższym stopniu konserwatywności przewidziane na podstawie badań *in vitro*

nt, 695-710 nt, 721-737 nt, 762-786 nt, 788-809 nt, 828-846 nt oraz 966-1006 nt (**Ryc. 8**).

Opublikowane później struktury szczepu IAV - A/WSN/33 w warunkach *in virio* oraz *ex virio* potwierdziły występowanie motywów strukturalnych typu spinka w rejonach: 36-59 nt, 138-171 nt, 337-357 nt, 788-809 nt. W strukturze *in virio* potwierdzono również motywy typu spinka w rejonach 696-709 nt oraz 768-779 nt. Wspólne motywy strukturalne dla szczepów A/Vietnam/1203/2004 oraz A/WSN/33 zostały zaprezentowane na **Ryc. 9**. Występowanie tych motywów strukturalnych w innym szczepie w warunkach *in virio* potwierdziło ich istotność w cyklu replikacyjnym wirusa. Co więcej, motyw spinki w rejonie 144-166 nt został przewidziany jako konserwatywny dla wielu szczepów IAV w badaniach *in silico*¹¹⁷.

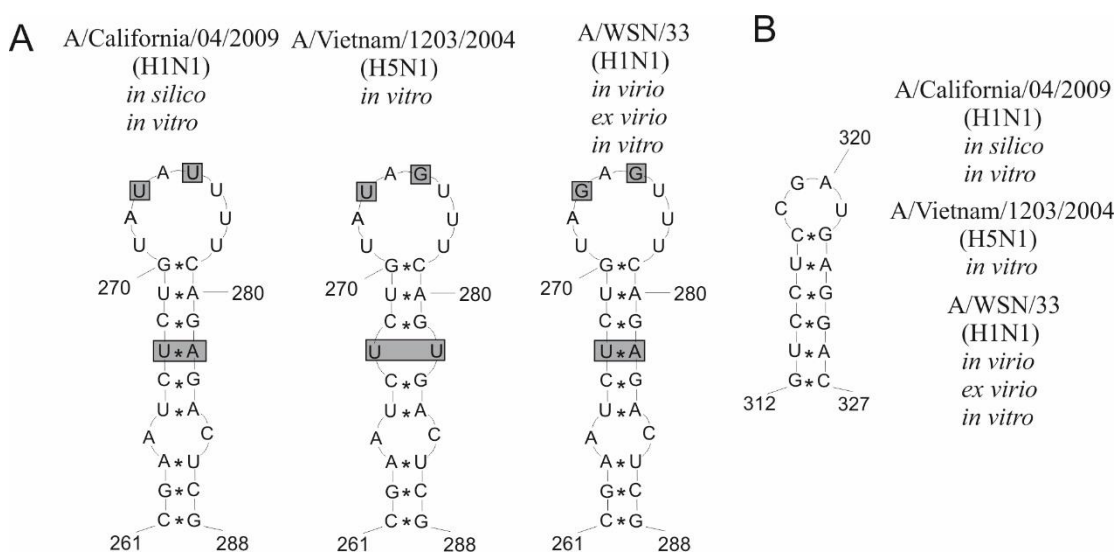


Ryc. 9 Motywy strukturalne przewidziane w strukturze vRNA7 szczepów A/Vietnam/1203/2004 (H5N1) oraz A/WSN/1933 (H1N1). Przedstawione motywy strukturalne z sekwencją dla szczepu A/Vietnam/1203/2004. Zmiany kompensacyjne w szczepie A/WSN/1933 zaznaczono kolorem szarym.

Segment 8 vRNA

Badania skupiające się na poznaniu struktury vRNA8 różnych szczepów wirusa grypy są prowadzone w naszej grupie od wielu lat. Dotychczas opublikowano strukturę vRNA8 szczepu A/Vietnam/1203/2004 *in vitro*, strukturę A/California/04/2009 modelowaną na podstawie danych eksperymentalnych szczepu A/Vietnam/1203/2004 oraz przewidywań bioinformatycznych, a także opracowano uniwersalną dla szczepów IAV strukturę vRNA8^{10,14}. Przewidziana struktura vRNA8 szczepu A/California/04/2009 została również zweryfikowana eksperymentalnie w warunkach *in vitro* (Marta Soszyńska-Jóźwiak et al., manuskrypt w recenzji). Nasze badania pozwoliły zaproponować konserwatywne motywy vRNA8, które z dużym

prawdopodobieństwem występują dla typu A wirusa grypy. Motywy o wysokiej konserwatywności to spinki w rejonach: 261-288 oraz 312-327 nt (**Ryc. 10**) oraz helisy w rejonach: 696-701/775-780 nt, 704-713/758-767 nt oraz 736-740/744-748 nt. Motywy typu spinka w rejonach 261-288 nt oraz 312-327 nt zostały potwierdzone również w strukturze *in vitro*, *ex virio* oraz *in virio* szczepu A/WSN/33¹¹⁵. Pozostałe motywy strukturalne, wskazane w publikacji *Lenartowicz et al.*¹⁴ jako konserwatywne, nie zostały potwierdzone w badaniach *Dadonaite et al.*¹¹⁵. Jednakże, wpływ na to mogło mieć zastosowanie odmiennych parametrów fałdowania cząsteczki vRNA, w której ograniczona została możliwość parowania się nukleotydów powyżej 150 nt.

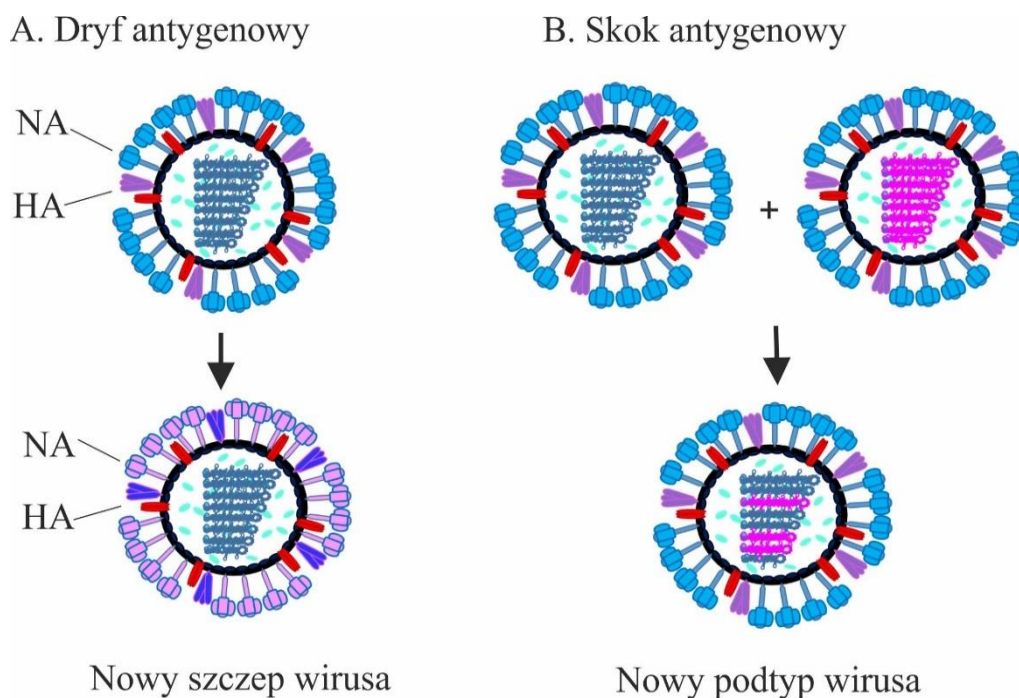


Ryc. 10 Konserwatywne motywy strukturalne występujące w strukturze vRNA8. **A.** Wspólny dla różnych szczepów wirusa grypy typu A motyw w rejonie 261-288 nt z zaznaczonymi zmianami sekwencji zachowującymi strukturę drugorzędową. **B.** Motyw strukturalny w rejonie 312-320 nt opisany dla szczepów wirusa grypy A/California/04/2009, A/WSN/33 oraz A/Vietnam/1203/2004 wykazuje 100% zachowawczości sekwencyjnej.

Interakcje pomiędzy segmentami vRNA

Nie tylko sygnały pakowania, ale także interakcje pomiędzy poszczególnymi cząsteczkami vRNA w kompleksie vRNP, stanowią obiekt intensywne badań. Rozwój wysokorozdzielczych metod fotozszywania sprzężonych z wysokoprzepustowym sekwencjonowaniem NGS umożliwił badanie tych zależności^{115,118}. Badania wykazały istnienie regionów typu *hotspots*, co w dosłownym tłumaczeniu oznacza ‘gorące miejsca’, w których występują intensywne oddziaływania typu RNA-RNA. Jednym z takich miejsc jest zidentyfikowany w centralnym regionie vRNA5 szczepu A/WSN/33 rejon oddziaływań z wieloma

segmentami i wieloma miejscami w obrębie tych segmentów¹¹⁸. Autorzy pracy wskazali, że spośród wszystkich segmentów, vRNA5 wykazuje najwięcej oddziaływań między segmentami. Najwięcej oddziaływań zidentyfikowano pomiędzy zlokalizowanym w rejonie centralnym cząsteczki ‘gorącym’ rejonem vRNA5 a rejonem 5’ vRNA6. Tak samo jak sygnały pakowania, również interakcje między segmentami mają różny stopień konserwatywności i mogą się różnić między szczepami. Zaobserwowane różnice wynikają ze zmian na poziomie sekwencji nukleotydowej, które mogą przyczyniać się do zmian w strukturze drugorzędowej vRNA, co z kolei przekłada się na oddziaływania typu RNA-RNA. Porównanie sieci interakcyjnej RNA-RNA w kompleksie vRNP blisko spokrewnionych (96% podobieństwa sekwencji), szczepów wirusa grypy typu A wykazała, że w istocie część tych interakcji jest wspólna¹¹⁵. Na poparcie swojej tezy, autorzy badań wykonali analizę sieci interakcyjnej innego podtypu wirusa IAV – A/Udorn/72 (H3N2), a także reasortanty wirusowe zawierające 6 segmentów A/PuertoRico/8/1934 oraz 2 segmenty A/Udorn/72/H3N2. Badania wskazały dopasowanie sieci interakcyjnej pomiędzy



Ryc. 11 Źródła zmienności genetycznej wirusa grypy **A.** Zjawisko dryfu antygenowego (ang. antigenic drift) występuje kiedy dochodzi do niewielkich mutacji w obrębie genów kodujących białka NA i HA, których kumulacja prowadzi ostatecznie do powstania nowych typów tych białek. W konsekwencji powstaje nowy podtyp wirusa. **B.** Skok antygenowy (ang. antigenic shift) polega na wymianie jednego lub więcej segmentów vRNA między niespokrewnionymi szczepami wirusa grypy, co prowadzi do powstania nowego szczepu wirusa grypy. Zjawisko skoku antygenowego ma miejsce w przypadku jednoczesnej infekcji komórki dwoma szczepami grypy pochodzącymi od różnych gospodarzy (szczepy infekujące ludzi, świnie lub ptaki).

segmentami vRNA z zachowaniem miejsc oddziaływań vRNA-vRNA pochodzących z macierzystego szczepu wirusa. Tym samym, badania potwierdziły, że nowe warianty wirusa grypy są w stanie dopasować nową sieć interakcji vRNA-vRNA z zachowaniem części oddziaływań i ma to duże znaczenie dla zjawiska skoku antygenowego oraz dryftu antygenowego (**Ryc. 11A, B**)¹¹⁹.

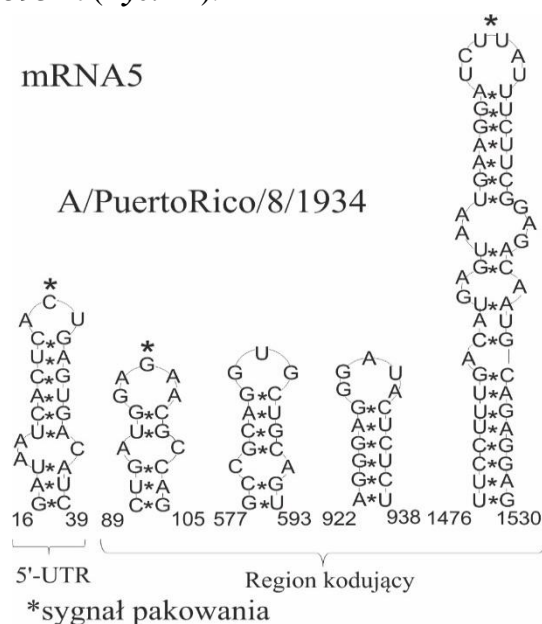
1.5.2. Konserwatywne motywy strukturalne RNA(+) wirusa grypy typu A

Większość poznanych dotychczas motywów strukturalnych RNA(+) wirusa grypy dotyczy badań *in vitro* oraz predykcji bioinformatycznych przeprowadzone przez Moss *et al.* badanie globalnego fałdowania się struktury drugorzędowej vRNA oraz cRNA wykazały, że vRNA wykazuje mniejszy stopień ustrukturyzowania¹²⁰. Ponadto, autorzy pracy wykazali, że cząsteczki RNA(+) wykazują wyższy stopień konserwatywności strukturalnej niż RNA(-). Co istotne, spośród wszystkich segmentów, 8RNA(+) wykazuje najwyższy stopień konserwatywności strukturalnej, a w obrębie samej cząsteczki 8RNA(+) najwyższą konserwatywnością wykazywały się motywy strukturalne w rejonach: 21-180 nt, 181-300 nt oraz 371-690 nt. Warto wspomnieć, że przewidziany w badaniach motyw mRNA5 w rejonie 1051-1171 nt został niedługo później potwierdzony eksperymentalnie przez naszą grupę¹²¹. Badania inhibicji celowanej na spinki występujące w obrębie tego motywu, przeprowadzone na szczepie A/California/04/2009 w linii komórkowej MDCK, wykazały obniżenie miana wirusa¹²². Wynik ten wskazuje na istotność tego motywu w przypadku szczepów IAV.

Predykcja bioinformatyczna struktur segmentów cRNA1, cRNA5, cRNA7 oraz cRNA8 wskazało, że struktury te w ujęciu globalnym różnią się pomiędzy szczepami grypy tzw. „ludzkiej”, „świńskiej” oraz „ptasiej”¹²³. Należy nadmienić, że temperatura otoczenia wpływa na fałdowanie cząsteczki RNA. Autorzy wskazują, że różnorodność tych struktur w ujęciu globalnym może wynikać właśnie z przystosowania do infekowanego organizmu, min. innej temperatury panującej w obrębie infekowanej tkanki w przypadku np. ptaków i ludzi.

Gulyaev *et al.*, przeprowadził analizę porównawczą struktur *in silico* różnych szczepów wirusa grypy, w wyniku której określone zostały konserwatywne motywy strukturalne segmentu 5 mRNA¹¹⁶. Autorzy pracy wykazali wysoką zachowawczość

strukturalna motywu w rejonie niekodującym 5'-UTR (16-39 nt) oraz 5 motywów w rejonie kodującym cząsteczki mRNA5: 89-105 nt, 577-593 nt, 922-938 nt oraz 1476-1530 nt **Ryc. 12** oznaczono motywy przewidziane w rejonach 3' i 5' zidentyfikowanym jako sygnały pakowania¹²⁴. Struktura globalna cząsteczki mRNA5 *in vitro* szczepu A/Vietnam/1203/2004 została zaproponowana przez naszą grupę¹²¹. Badania w oparciu o dane eksperymentalne uzupełnione w analizę sekwencyjno-strukturalną wyłoniły motywy konserwatywne dla wielu szczepów IAV. Co istotne, w naszej strukturze zostały potwierdzone dwa motywy w regionach 16-39 nt oraz 577-593 nt (**Ryc. 12**).



Ryc. 12 Konserwatywne motywy strukturalne mRNA5 wytypowane na podstawie analizy sekwencyjno-strukturalnej wielu szczepów IAV. Sekwencja szczepu A/PuertoRico/8/1934.

Przewidywania *in silico*, w oparciu o dane termodynamiczne oraz dostępne w bazie danych sekwencje IAV, pozwoliły na identyfikację potencjalnie konserwatywnych motywów strukturalnych mRNA segmentów 7 i 8¹²⁰. Zgodnie z przewidywaniami, konserwatywne motywy mRNA zlokalizowane są blisko lub w obrębie miejsc splicingowych w obrębie tych segmentów. Znaczenie tych motywów strukturalnych w cyklu replikacyjnym wirusa potwierdzono poprzez indukowaną mutację sekwencji mRNA mającą na celu zaburzenie powstawania tych struktur¹²⁵. Tym samym potwierdzono duże znaczenie przewidywań bioinformatycznych w identyfikacji potencjalnie funkcjonalnych motywów strukturalnych.

Niedawno opublikowana praca dotycząca struktur mRNA wirusa grypy rzuciła nowe światło na fałdowanie się tych struktur w warunkach *in vivo*¹²⁶. Autorzy pracy doszli do wniosku, że struktury mRNA w warunkach komórkowych są mniej ustrukturyzowane niż w warunkach *in vitro*. Badania wskazały, że struktura wirusowego mRNA wykazuje zachowawczość lokalnych motywów strukturalnych przewidzianych dla struktur *in vitro*. Największą zachowawczość w parowaniu zasad pomiędzy strukturami *in vitro* oraz *in vivo* zaobserwowano dla segmentów 7 i 8 mRNA. Co ciekawe, autorzy pracy wykazali istnienie struktur typu pseudowęzła w strukturach *in vitro* segmentów 1,2,3,4,5 oraz 8, jednakże ich obecność w strukturze *in vivo* nie została przez nich potwierdzona¹²⁶.

1.6. Metody badań struktury drugorzędowej RNA

Metody badań struktur drugorzędowych RNA opierają się w głównej mierze na metodach eksperymentalnych i uzupełniających je predykcjach bioinformatycznych. Istnieje kilka rodzajów metod służących do eksperymentalnego określania struktur drugorzędowych RNA. Możemy wyróżnić metody fizyczne (Cryo-EM, NMR), enzymatyczne (mapowanie enzymami np.: T1, S1, V1 lub RNazą H w obecności DNA) oraz chemiczne¹²⁷⁻¹³⁰. Idealne podejście eksperymentalne powinno łączyć te metody w dążeniu do określenia najbardziej prawdopodobnej struktury drugorzędowej. Niestety, podejście to jest zarezerwowane dla struktur w układzie *in vitro*, dosyć często również z ograniczeniem długości RNA. W przypadku określania struktury drugorzędowej RNA w warunkach biologicznych podejścia eksperymentalne są ograniczone ze względu na złożoność danego układu. Występujące licznie interakcje RNA z cząsteczkami biologicznymi obecnymi w warunkach *in vivo*, bądź *in cellulo* wymagają metod spełniających dodatkowe kryteria np.: przenikalność przez błony komórkowe czy stosunkowo krótkiego czasu reakcji. W związku z tym, że w niniejszej pracy skupiono się na określeniu struktur RNA w warunkach biologicznych, poniższy opis ograniczony został do metod chemicznego mapowania struktur.

1.6.1. Metody badań struktury RNA z zastosowaniem odczynników mapujących

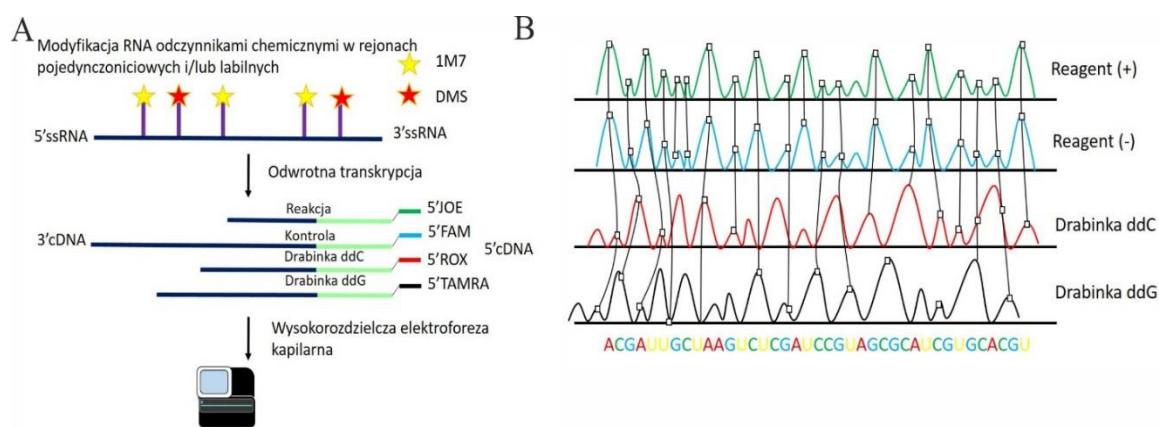
Podstawową metodą określania struktury drugorzędowej RNA jest zastosowanie odczynników mapujących, które mają zdolność do reagowania z wolną grupą 2'-OH

rybozy dowolnego nukleotydu w rejonach jednoniciowych i/lub labilnych strukturalnie (odczynniki z grupy SHAPE, ang. Selective 2'-hydroxyl acylation analyzed by primer extension)¹³¹. Drugim rodzajem odczynników chemicznych jest grupa bezpośrednio oddziałująca z zasadą azotową określonych nukleotydów w miejscach jednoniciowego RNA. Wśród specyficznie oddziałujących odczynników wyróżnia się: DMS (N1-adenozyna, N3-cytydina), DEPC (N7-adenozyna), CMCT (N3-uracyl, N1-guanozyna) czy ketoksal (N1 oraz N2-guanozyna)^{132,133}. Obecnie znane są dwie metody umożliwiające odczyt mapowania chemicznego z zastosowaniem reakcji odwrotnej transkrypcji – metoda odczytu stopów odwrotnej transkryptazy (ang. RT-stop count) oraz metoda odczytu profilu mutacji (ang. MaP, mutational profile). Obydwie z nich bazują na niezdolności do rozpoznania zmodyfikowanego nukleotydu przez enzym RT, jednakże zarówno reakcja RT jak i sposób odczytu mapowania różni się w zależności od wybranej metody.

1.6.2 Metody badań struktury RNA z zastosowaniem odczynników mapujących i elektroforezy kapilarnej

W przypadku pierwszej metody, enzym odwrotna transkryptaza zatrzymuje wydłużanie nici cDNA na matrycy RNA na jeden nukleotyd przed zmodyfikowanym nukleotydem. Następnie, określana jest reaktywność poszczególnych nukleotydów na podstawie wyliczeń zatrzymań enzymu RT w czasie reakcji odwrotnej transkrypcji i normalizacja zatrzymań poprzez odcięcie tła wynikającego z naturalnie powstających zatrzymań, najczęściej w miejscach tworzących trwale termodynamicznie struktury. W wyniku reakcji powstaje wiele fragmentów cDNA, których końce kumulują się na 1 nt przed modyfikacją. Następnie, odczyt zatrzymań RT możliwy jest z zastosowaniem rozdziału fragmentów przy użyciu elektroforezy kapilarnej (**Ryc. 13**). Elektroforeza kapilarna obu reakcji, zarówno, zmodyfikowanej chemicznie, jak i jej kontrola, przeprowadzana jest w jednej kapilarze. W celu rozróżnienia poszczególnych próbek do reakcji odwrotnej transkrypcji stosuje się znakowane fluorescencyjnie startery. Prawidłowa identyfikacja mapowań i odpowiednie dopasowanie do sekwencji RNA, możliwe jest dzięki zastosowaniu drabinek sekwencyjnych (drabinki dideoksy) (**Ryc. 13 A,B**). Drabinki sekwencyjne są produktami reakcji sekwencjonowania DNA typu Sanger na matrycy DNA do transkrypcji RNA *in vitro* badanej cząsteczki. Do reakcji stosuje się wybrane

dideoksynukleotydy (dd) np. kombinacja ddG i ddA oraz startery znakowane fluorescencyjnie. Wyniki rozdzielów są następnie analizowane za pomocą programu, np. ShapeFinder¹³⁴ albo QuShape¹³⁵.



Ryc. 13 Schemat analizy odczytu mapowań chemicznych metodą bazującą na zatrzymaniu odwrotnej transkrypcji, sprzężoną z elektroforezą kapilarną. **A.** Modyfikowany RNA jest poddany reakcji RT z zastosowaniem starterów znakowanych fluorescencyjnie. Dla każdej próbki (reakcja, kontrola, drabinka dideoksynukleotydowej) stosowany jest inny fluorofor. Rozdział kapilarny zachodzi w pojedynczej kapilarze. **B.** Dzięki zastosowaniu różnych znaczników fluorescencyjnych możliwe jest rozróżnienie poszczególnych próbek i ich analiza w programie ShapeFinder.

1.6.3. Metody odczytu mapowania RNA oparte o zatrzymania odwrotnej transkryptyzy z zastosowaniem sekwencjonowania nowej generacji

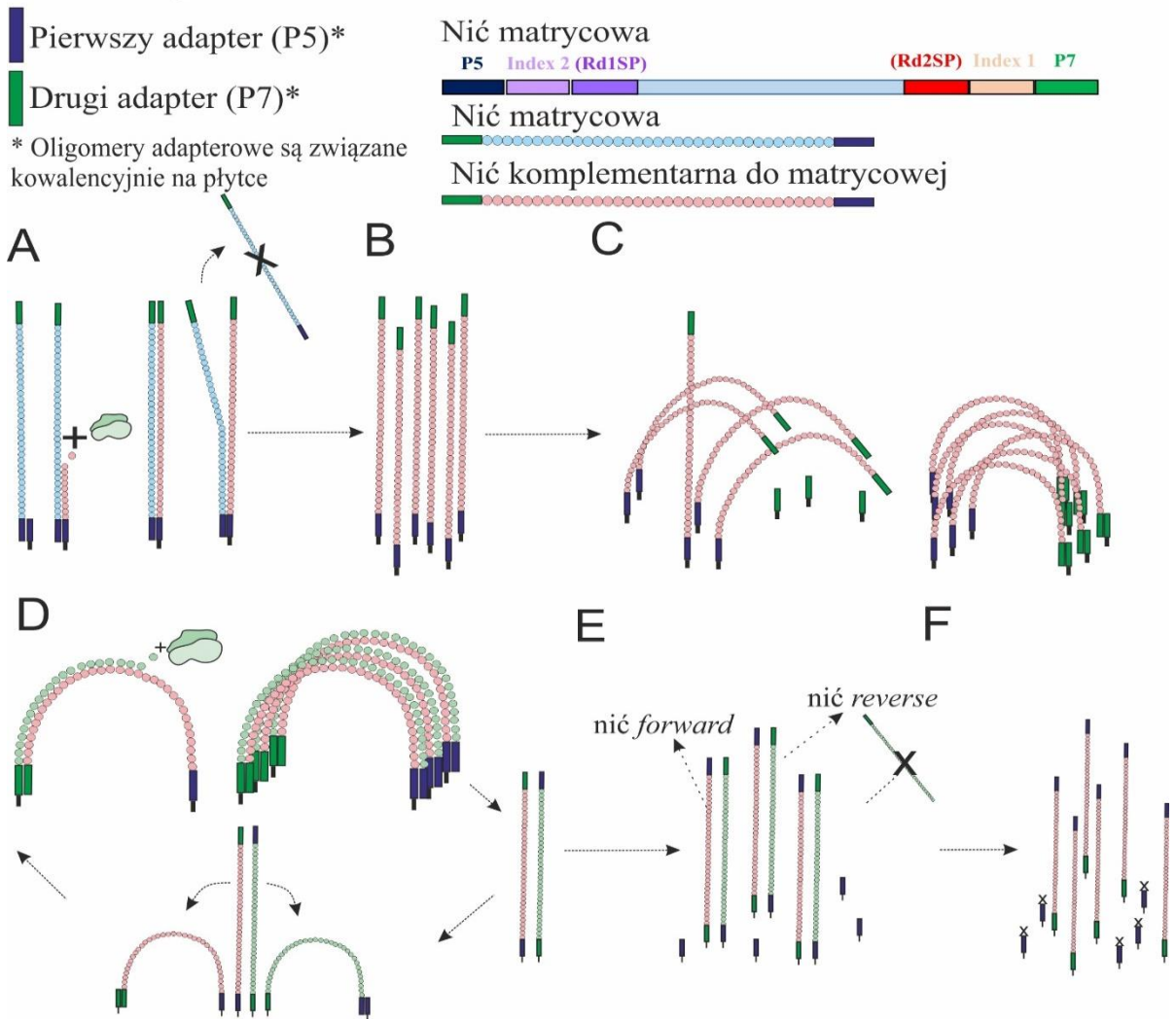
Sekwencjonowanie Nowej Generacji (NGS)

Sekwencjonowanie Nowej Generacji (ang. Next Generation Sequencing, NGS), nazywane również sekwencjonowaniem wysokoprzepustowym, jest technologią umożliwiającą równoczesne sekwencjonowanie zróżnicowanego zbioru fragmentów DNA z użyciem minimalnej ilości materiału¹³⁶. Tym samym jest rewolucyjną technologią umożliwiającą sekwencjonowanie od pojedynczych cząsteczek do całych genomów w stosunkowo krótkim czasie. Wraz z rozwojem technologii, opracowane zostały metody określania struktury drugorzędowej RNA sprzężone z sekwencjonowaniem NGS^{137–143}. Opracowane metody bazują na sekwencjonowaniu z wykorzystaniem platformy Illumina, chociaż pojawiają się również doniesienia o zastosowaniu technologii Nanopore^{144–146}. W niniejszej pracy wykorzystano metodę sekwencjonowania NGS na platformie Illumina NextSeq550, dlatego też poniższy opis dotyczy tej technologii.

Etapem poprzedzającym sekwencjonowanie NGS jest generacja klastrów. Etap ten polega na namnożeniu pojedynczych (unikalnych) fragmentów DNA w formie homogenych zbiorów (klastrów). Tak więc każdy klaster stanowi namnożony zbiór pojedynczej cząsteczki DNA. Schemat generowania klastrów został przedstawiony na schemacie **Ryc. 14**.

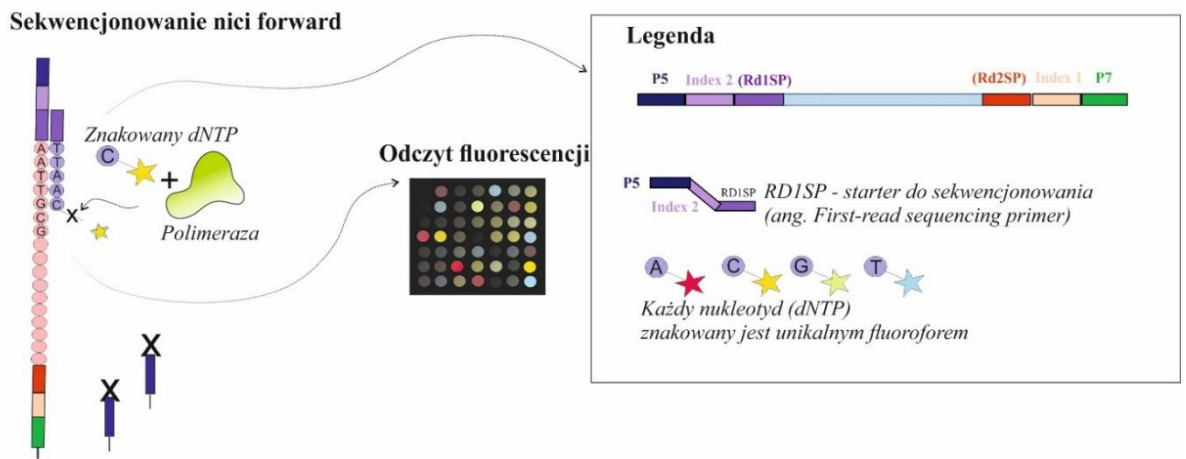
Po zakończeniu etapu generowania klastrów następuje właściwy etap sekwencjonowania. Sekwencjonowanie Illumina jest typem sekwencjonowania poprzez syntezę (ang. Sequencing by synthesis, SBS). W każdym cyklu sekwencjonowania dochodzi do generowania komplementarnej nici DNA poprzez przyłączenie znakowanego fluorescencyjnie dNTP (**Ryc. 15**). Każdy z dNTP posiada inny znacznik fluorescencyjny. Po jego przyłączeniu dochodzi do emisji fluorescencji i jej odczytu. Każdy cykl sekwencjonowania obejmuje przyłączenie jednego nukleotydu. Tak więc długość sekwencjonowania zależy od ilości cykli, np. 150 cykli oznacza sekwencjonowanie o długości 150 nt. W przypadku systemu Illumina można wykonać sekwencjonowanie jednego końca cząsteczki (ang. Single-end Sequencing **Ryc. 15**) lub sekwencjonowanie obu końców cząsteczki (ang. Paired-end Sequencing, **Ryc. 16**).

Generacja klastrów



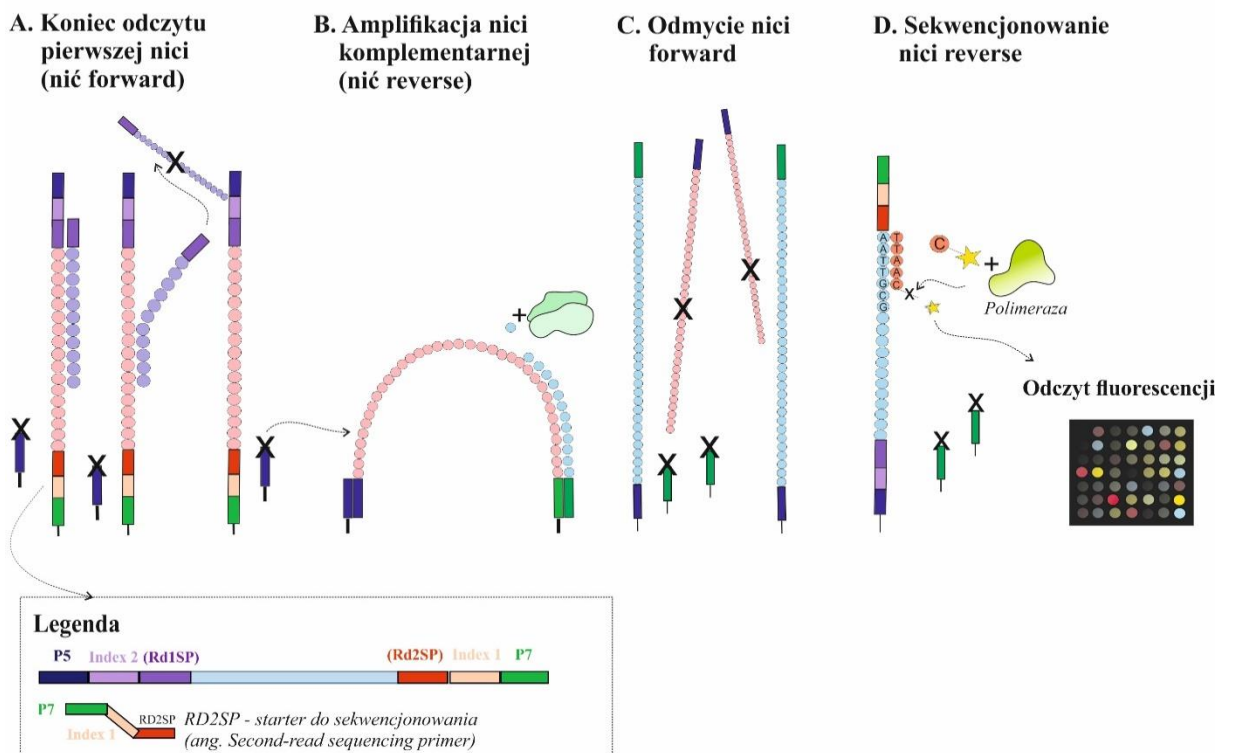
Ryc. 14 Schemat przedstawiający etap generacji klastrów. Generacja klastrów jest etapem poprzedzającym sekwencjonowanie Illumina. W przypadku systemu NextSeq550 zarówno generacja klastrów jak i sekwencjonowanie zachodzi w tym samym urządzeniu. **A.** Po przyłączeniu nici matrycowej do adapterów P5 na płycie do sekwencjonowania następuje synteza nici komplementarnej. Po syntezie nowej nici przeprowadzana jest denaturacja i odmycie nici matrycowej. **B.** Na płycie pozostają tylko komplementarne nici. **C.** Powstanie mostków (ang. Bridges) poprzez wygięcie nici ssDNA i połączenie jej drugiego końca z drugim adapterem na płycie (P7). **D.** Po utworzeniu mostków następuje cykl amplifikacji mostkowej (ang. Bridge Amplification). Cykl ten obejmuje: syntezę komplementarnej nici DNA, denaturację nici i powstanie nowego mostka. Cykl amplifikacji mostkowej zachodzi wielokrotnie, w wyniku czego powstają klastry. Każdy klaster stanowi homogeny zbiór namnożonej, unikalnej cząsteczki DNA. **E.** Po zakończeniu ostatniego cyklu generowania klastrów dochodzi do denaturacji i odłączenia nici *reverse* od oligomeru związanego na płycie. **F.** Na płycie pozostają tylko nici *forward*. Wolne oligomery adapterowe zostają zablokowane. Schemat przygotowany na podstawie animacji firmy Illumina (*dostępna pod adresem: www.youtube.com/watch?v=fCd6B5HRaZ8*).

Sekwencjonowanie typu single-end



Ryc. 15 Synteza nici komplementarnej z zastosowaniem znakowanych dNTP. Podczas sekwencjonowania dochodzi do syntezy komplementarnej nici DNA z zastosowaniem znakowanych fluorescencyjnie dNTP. W każdym cyklu przyłączany jest jeden nukleotyd. Po przyłączeniu nowego nukleotydu dochodzi do emisji fluorescencji, a następnie jej odczyt. Sekwencjonowanie typu single-end ogranicza się do sekwencjonowania jednego końca cząsteczki DNA. W czasie sekwencjonowania 3' koniec adaptera P5 (kolor granatowy) jest zablokowany. Schemat przygotowano na podstawie animacji firmy Illumina

Sekwencjonowanie typu paired-end



Ryc. 16 A. Koniec odczytu pierwszej nici (nić forward). Po zakończeniu cyklu sekwencjonowania pierwszej nici dochodzi do denaturacji i odmycia wydłużonej nici **B. Amplifikacja nici komplementarnej (nić reverse).** Po odblokowaniu 3' końca adaptera P5 (kolor granatowy), następuje cykl amplifikacji mostkowej w celu stworzenia komplementarnej nici, która będzie nicią matrycową do sekwencjonowania drugiego końca. **C. Odmycie nici forward.** Po zakończeniu amplifikacji następuje denaturacja nici forward i zablokowanie adaptera P7 (kolor zielony). **D. Sekwencjonowanie nici reverse.** Następuje sekwencjonowanie drugiej nici (reverse). Sekwencjonowanie zachodzi tak samo jak w przypadku nici forward. W schemacie nie uwzględniono etapu odczytu sekwencji indeksującej. Schemat przygotowano na podstawie animacji firmy Illumina

Metody odczytu mapowań chemicznych z zastosowaniem zatrzymań odwrotnej transkryptazy

Jedną z pierwszych metod odczytu zatrzymań RT w warunkach biologicznych sprzężonych z NGS jest metoda Structure-seq¹³⁹. Pierwszym etapem eksperymentu jest mapowanie RNA w linii komórkowej z zastosowaniem odczynnika chemicznego, a następnie izolacja całkowitego RNA. Następnie przeprowadzana jest reakcja odwrotnej transkrypcji ze starterem typu losowy heksamer, który na 5' końcu posiada znaną sekwencję adapterową. Kolejnym etapem jest dołączenie adaptera na 3' końcu cząsteczki. W tym celu przeprowadza się reakcję ligacji z zastosowaniem enzymu CircLigase. Dzięki dołączeniu znanych sekwencji na obu końcach cząsteczki, możliwe jest przeprowadzenie reakcji PCR z zastosowaniem starterów komplementarnych do sekwencji adapterowych posiadających dodatkowo sekwencje starterów (wraz z indeksami) potrzebnymi do sekwencjonowania na platformie Illumina. Poszczególne reakcje, za wyjątkiem reakcji RT, są oczyszczane z zastosowaniem elektroforezy: w przypadku ligacji jest to denaturujący żel poliakrylamidowy (PAGE), natomiast reakcja PCR jest oczyszczana za pomocą elektroforezy na żelu agarozowym. Etapy oczyszczania mają na celu pozbycie się niezwiązanych w reakcji starterów, substratu ligacji, produktów pośrednich typu primer-dimer, a także selekcję fragmentów o określonej długości. Konstrukcja biblioteki do sekwencjonowania metodą Structure-seq okazała się być bardzo czasochłonna i charakteryzowała się niską wydajnością wynikającą, między innymi, z niedopracowanej metody ligacji oraz niewydajnej elucji z żelów preparatywnych.

Niedługo potem, autorzy opublikowali ulepszoną metodę – Structure-seq2, w której zoptymalizowana została reakcja ligacji, dodany został dodatkowy etap oczyszczania z zastosowaniem PAGE po reakcji RT, a końcowy etap oczyszczania biblioteki za pomocą elektroforezy agarozowej został zastąpiony elektroforezą PAGE¹³⁸. Ulepszona reakcja ligacji bazuje na zastąpieniu jednoniciowego adaptera, adapterem strukturalnym typu spinka oraz zmianie enzymu na T4 ligazę, co znacznie poprawiło wydajność reakcji¹⁴⁷. Autorzy testowali również efektywność konstrukcji biblioteki z zastosowaniem biotynylowanych dCTP. Biotynylowane dCTP są dodatkowo dodawane do mieszaniny dNTP wchodzących w skład reakcji odwrotnej transkrypcji. Dzięki temu cząsteczki cDNA można oczyścić za pomocą kulek

streptawidynowych, które wykorzystują zjawisko silnego wiązania biotyny ze streptawidyną. Oczyszczanie na kulkach ma na celu eliminację czasochłonnego i mało wydajnego etapu oczyszczania na żelach poliakrylamidowych. Podstawowym problemem w zastosowaniu obydwóch metod jest ich ograniczenie do transkryptów, które występują w dużej ilości w mieszaninie totalnego RNA. Problem ten można obejść z zastosowaniem dodatkowego etapu oczyszczania, np. poprzez usunięcie frakcji rybosomalnego RNA (rRNA) lub zagęszczenie frakcji mRNA za pomocą sond poli-T. Natomiast, w przypadku rzadkich transkryptów wymagane jest zastosowanie specyficznych starterów do reakcji odwrotnej transkrypcji.

Metodą bazującą na specyficznych starterach do reakcji RT jest Celowany Structure-seq (ang. Targeted Structure-seq)¹⁴⁸. W metodzie tej starter typu losowy heksamer został zastąpiony zestawem specyficznych starterów również zawierających sekwencję adaptera na 5' końcu. Oprócz tego, autorzy metody wykorzystali analogiczną reakcję ligacji strukturalnego adaptera z zastosowaniem enzymu T4 ligazy. Dodatkowo, pojedyncza reakcja PCR z długimi starterami została podzielona na 2 etapy: w pierwszym powielana jest sekwencja między adapterami dodanymi na początkowych etapach konstrukcji biblioteki, a następnie przeprowadzana jest końcowa reakcja PCR z adapterami i sekwencjami indeksów potrzebnych do sekwencjonowania na platformie Illumina. Atutem tej metody jest również zastąpienie czasochłonnych etapów oczyszczania na żelach, wydajnym i szybkim oczyszczaniem na kulkach magnetycznych typu SPRI (ang. Solid-Phase Reversible Immobilization). Metoda celowanego Structure-seq ma zastosowanie w przypadku rzadkich transkryptów i dużych cząsteczek. Jednakże, zastosowanie specyficznych starterów może skutkować niskim pokryciem sekwencjonowania fragmentów pomiędzy poszczególnymi starterami.

Metody bazujące na odczycie mapowania oparte o zatrzymanie odwrotnej transkrypcji, sprzężone z sekwencjonowaniem NGS wymagają przeprowadzenia optymalizacji konstrukcji biblioteki na wielu etapach, począwszy od poszczególnych reakcji, skończywszy na wydajnym oczyszczaniu produktów. Co więcej, brak możliwości konstrukcji biblioteki za pomocą dostępnych komercyjnie zestawów kompatybilnych z nowoczesnymi platformami do sekwencjonowania może sprawiać dużą trudność w przypadku mniej doświadczonych laboratoriów. Tym samym ich

użycie zostało mocno ograniczone. W związku z tym, protokoły te zostały dosyć szybko wyparte przez bardziej zoptymalizowane metody bazujące na profilu mutacji.

Metody odczytu mapowania RNA oparte o analizę profilu mutacji wprowadzonych przez odwrotną transkryptazę

Drugą metodą umożliwiającą odczyt mapowań chemicznych RNA jest metoda bazująca na profilu mutacji (ang. Mutational Profile, MaP). Metoda MaP wykorzystuje niezdolność enzymu RT do prawidłowego odczytu zmodyfikowanego nukleotydu i bazuje na modyfikacji reakcji RT. Modyfikacja reakcji RT polega na zastosowaniu buforu reakcyjnego wzbogaconego w jony dwuwartościowe np. Mn^{2+} . Dzięki temu, podczas odwrotnej transkrypcji enzym nie zatrzymuje się i w nowo powstającej nici cDNA, w miejscu komplementarnym do modyfikowanego nukleotydu, wprowadza mutację. Następnie porównanie liczby mutacji pomiędzy próbą reakcyjną (mapowaną odczynnikiem chemicznym) a próbą kontrolną (nietraktowaną odczynnikiem) umożliwia obliczenie reaktywności poszczególnych nukleotydów. Można wyróżnić dwie metody odczytu informacji strukturalnej RNA opartej o profil mutacji, a ich nazwy zależne są od użytego odczynnika modyfikującego – SHAPE-MaP oraz DMS-MaPseq^{140,143,149}. Głównym atutem metod MaP jest ich szerokie zastosowanie do zarówno krótkich jak i długich transkryptów o różnej ilości występowania w komórce. Co więcej, wyeliminowana jest potrzeba przeprowadzania reakcji ligacji, ponieważ w czasie reakcji RT powstają cDNA o pełnej długości o znanej sekwencji. Po drugie, metoda częściowo bazuje na komercyjnie dostępnych zestawach, a tym samym nie wymaga przeprowadzenia czasochłonnej optymalizacji. Ponadto, nie wymaga stosowania etapów oczyszczania opartych o elektroforezę, które mogą być zastąpione przez oczyszczanie na kolumnach lub za pomocą kulek magnetycznych typu SPRI.

1.7. Planowanie eksperymentu mapowania struktury RNA w warunkach biologicznych

Mapowanie struktury RNA w warunkach biologicznych stanowi duże wyzwanie metodologiczne. Eksperyment wymaga optymalizacji odpowiednich warunków takich jak czas, dobór i stężenie odczynnika chemicznego, zachowanie odpowiedniego pH w trakcie mapowania chemicznego oraz efektywne zatrzymanie reakcji. Ma to na celu uzyskanie wiarygodnych wyników mapowania, a także zapobiegnięcie dalszego

działania odczynnika mapującego w trakcie izolacji RNA. W czasie mapowania chemicznego *in cellulo* ważne jest również to, aby w czasie eksperymentu komórki nie wykazywały efektu cytotoksycznego.

Wybór odczynnika mapującego podyktowany jest zdolnością jego wnikania do wnętrza komórki przy stosunkowo niskiej cytotoksyczności, dobrą rozpuszczalnością nawet w wysokich stężeniach, stosunkowo krótkim czasem działania (czas półtrwania) oraz możliwością zatrzymania reakcji mapowania. Wśród odczynników, które mają zdolność do przenikania przez błony komórkowe wyróżnia się zarówno odczynniki z grupy SHAPE jak i DMS. W ciągu ostatnich lat pojawiło się wiele publikacji eksperymentalnych, w których dyskutowano efektywność poszczególnych odczynników w warunkach *in cellulo*^{140,143,148,150–152}. Wyniki badań dosyć często przedstawiały odmienne spostrzeżenia dotyczące efektywności działania poszczególnych odczynników.

Eksperymenty mapowania struktur RNA zarówno w hodowli komórkowej jak i bakteryjnej wykazały efektywne działanie odczynnika NAI oraz FAI, z dużą przewagą efektywności tego pierwszego¹⁵¹. Autorzy pracy jednocześnie wykluczyli zdolność mapowania struktur w warunkach komórkowych z zastosowaniem NMIA. Badania przeprowadzone przez *Lee et al.* miały na celu przetestowanie zdolności mapowania RNA komórkowego z zastosowaniem odczynników FAI, NAI, NAI-N₃ oraz 1M7¹⁵². Autorzy wskazali, że stopień modyfikacji, wynikający ze stosunku reaktywności próbki traktowanej odczynnikiem do próbki traktowanej DMSO, z zastosowaniem odczynnika 1M7 jest znacznie niższy niż w przypadku pozostałych odczynników z grupy SHAPE. Autorzy pracy wskazali, że prawdopodobną przyczyną jest krótki czas półtrwania (~17 s, 37°C), niska rozpuszczalność odczynnika w większych stężeniach (10-krotne zwiększenie stężenia nie wpłynęło na stopień modyfikacji), a także przypuszczalnie niewielka zdolność 1M7 do przenikania przez błony i ściany komórkowe. Kolejne badania mapowania struktur RNA *in cellulo* z zastosowaniem odczynników 1M7, 1M6, NMIA, 5NIA oraz NAI wykazały, że wszystkie te odczynniki spełniają kryteria zdolności mapowania RNA w komórkach¹⁵⁰. Autorzy wskazali jednak, że każdy z tych odczynników wymaga innego podejścia w trakcie analizy reaktywności. Niedawno ukazała się również praca, w której z powodzeniem przeprowadzono mapowanie chemiczne *in cellulo* z wykorzystaniem odczynników

NMIA, 1M7 oraz NAI¹⁵³. Biorąc pod uwagę powyższe, wybór odczynnika 1M7 oraz NMIA wydaje się być mocno dyskusyjny, ze względu na odmienne wnioski przedstawione w niezależnych badaniach. W **Tabeli 2** znajduje się podsumowanie poszczególnych odczynników.

Tabela 2 Podsumowanie informacji dotyczących poszczególnych odczynników mapujących oraz ich potencjalne zastosowanie w badaniach *in vitro* oraz *in cellulo*.

Nazwa/pełna nazwa	Czas półtrwania (W 37°C), pH=8	Uwagi
1M7 (1-Methyl-7-nitroisatoic anhydride)	~17 s	Niezależne badania mapowania struktur RNA wykazały zróżnicowany pogląd odnośnie możliwości użycia tego odczynnika w warunkach <i>in vivo</i> ^{143,150-152}
1M6 (1-Methyl-6-nitroisatoic anhydride)	~31 s	Nie zaleca się stosowania w warunkach <i>in vivo</i> ¹⁵⁰
5NIA (5-nitroisatoic anhydride)	~100 s	Szybki czas działania; Nie wymaga dodatkowego etapu zatrzymywania reakcji. Wykazuje nadreaktywność w przypadku adenozyiny i wymaga dodatkowej normalizacji danych ¹⁵⁰
NAI (2-Methylnicotinic acid imidazolide) NAI-N ₃ (2-methylnicotinic acid imidazolide-azide)	~33 min	Skuteczny do mapowania struktur <i>in cellulo</i> , ale wymaga dodatkowego etapu zatrzymania reakcji ^{143,152}
NMIA (N-methylisatoic anhydride)	~260 s	Zróżnicowany pogląd dotyczący efektywności odczynnika w warunkach <i>in cellulo</i> ^{150,151,153} .
FAI (2-methyl-3-furoic acid imidazolide)	~73 min	Wykazuje niższy poziom reaktywności <i>in cellulo</i> niż NAI ^{150,152}
DMS (Dimethyl sulfide)	~30 min	Mapuje przy A i C; Wymaga dodatkowego etapu zatrzymania reakcji ^{140,148} . Skuteczny w mapowaniu chemicznym <i>in cellulo</i> ¹⁴⁰ .

1.8. Strategie przeciwwirusowe

Najpopularniejszą strategią antywirusową jest stosowanie profilaktyki w postaci szczepień. Co roku opracowywana jest szczepionka, na którą składają się 4 wyselekcjonowane szczepy wirusów: dwa szczepy wirusa typu A (podtyp H1N1 oraz H3N2) oraz dwa szczepy wirusa typu B. Selekcja szczepów wykonywana jest na podstawie predykcji najwyższego prawdopodobieństwa występowania w danym sezonie grypowym^{154,155}. Co istotne, produkcja szczepionki przeciwwirusowej jest czasowo- i pracochłonnym procesem¹⁵⁶. Z tego powodu skład szczepionki jest opracowywany na kilka miesięcy przed wystąpieniem sezonu grypowego. Tak więc, może się zdarzyć, że opracowana szczepionka może być nieefektywna. Taka sytuacja miała miejsce w 2009 roku kiedy światem zawładnęła pandemia wirusa grypy typu A i dość szybko dotknęła ponad 88 milionów ludzi¹⁵⁷. Rozwiązaniem może być prężnie

rozwijająca się nanotechnologia, która umożliwiła opracowanie szczepionki mRNA, a badania dotyczące tej strategii zostały przez nas zreferowane w pracy przeglądowej¹⁵⁸.

Drugą strategią antywirusową stosowaną w przypadku grypy są leki przeciwwirusowe pełniące funkcję inhibitorów białek wirusowych. Jednym z pierwszych zatwierdzonych do użytku leków były inhibitory kanału jonowego M2 – amantadyna oraz rymantadyna, chociaż ich stosowanie charakteryzuje się częstym wywoływaniem skutków ubocznych¹⁵⁹. Problemem stały się również pojawiające się szczepy odporne na działanie tych^{160–162}. Kolejnym typem leków są inhibitory neuraminidazy – oseltamiwir, zanamiwir, peramiwir i laninamiwir^{162–164}. Niestety, również i w tym przypadku zaobserwowano szczepy odporne na działanie leków z tej^{165–168}. Obecnie pokładana jest nadzieja w nowej klasie inhibitorów wirusowej polimerazy RNA do której zaliczany jest baloksawir marboksyl^{169,170}. Badania wskazały również ich skuteczność w stosowaniu w przypadku szczepów opornych na leki z grupy inhibitorów neuraminidazy¹⁷¹.

1.8.1. Zależne od struktury drugorzędowej RNA metody inhibicji nakierowane na wirusowe RNA

Jak już wspomniano wcześniej, w obrębie drugorzędowej cząsteczek vRNA jak i mRNA oraz cRNA znajdują się konserwatywne motywy strukturalne, które są niezbędne do prawidłowego powstawania wirulentnej cząsteczki potomnej wirusa. Badania mutacyjne potwierdziły wpływ zmian konformacji cząsteczki na osłabienie wirulencji, tym samym wskazały na potencjalne zastosowanie metod inhibicji wirusa nakierowanych na konserwatywne motywy RNA¹²⁵. Idealnym rozwiązaniem, które umożliwiłoby ominięcie oporności wirusów na leki, jest zaprojektowanie uniwersalnych inhibitorów nakierowanych właśnie na konserwatywne motywy RNA. Chociaż na dzień dzisiejszy nie zarejestrowano żadnego leku przeciwko wirusowi grypy opartego o technologię RNA, to opracowane metody mogą być w przyszłości wykorzystane w leczeniu lekoopornych szczepów wirusa. Poniżej zostaną pokrótce omówione metody inhibicji RNA o potencjalnym znaczeniu terapeutycznym w leczeniu zakażeń wirusem grypy. Inhibicja namnażania wirusa grypy typu A w oparciu o strukturę RNA są od wielu lat obiektem badań prowadzonych w naszym zakładzie.

Metoda inhibicji namnażania wirusa grypy z wykorzystaniem siRNA i antysensowych oligonukleotydów (ASO)

Metoda inhibicji oparta o krótkie interferujące RNA (ang. small interfering RNA, siRNA) wykorzystuje naturalnie występujące zjawisko wyciszania genów - interferencji RNA (ang. RNA interference, RNAi)^{172,173}. Technologia ta jest uniwersalna dla wszystkich typów wirusów, zarówno DNA jak i RNA, dlatego też znalazła zastosowanie w przypadku wirusa grypy¹⁷⁴. Przeprowadzone w naszym zakładzie badania wpływu siRNA na namnażanie wirusa szczepu A/California/04/2009 wykazały skuteczność inhibicji na poziomie 85%¹⁷⁵. Najwyższą efektywność zaobserwowano dla cząsteczek siRNA naceLOWanych na rejony konserwatywne cząsteczki (+)RNA5. Z kolei inne badania wyciszania ekspresji genów białek M1, M2, NS1 oraz NEP z użyciem siRNA wykazały również addycyjny efekt inhibicji namnażania wirusa przy równoczesnym stosowaniu siRNA naceLOWanych na więcej niż jedną cząsteczkę wirusowego mRNA¹⁷⁶. Należy wspomnieć, że metoda inhibicji oparta na technologii siRNA jest podatna na efekt tzw. off-target, który polega na niespecyficznym wyciszaniu genów w komórce gospodarza. Dlatego również ważny jest rozwój metod projektowania cząsteczek siRNA minimalizujący ryzyko wystąpienia efektu off-target¹⁷⁷.

Drugą metodą inhibicji namnażania wirusa grypy naceLOWaną na strukturę drugorzędową RNA są antysensowe oligonukleotydy (ang. Antisense oligonucleotides, ASO). Są to krótkie pojedynczoniowe oligonukleotydy, które łączą się komplementarnie z docelową sekwencją RNA. W przypadku wirusów RNA, efektywność ASO objawia się w zahamowaniu namnażania wirusa w komórce gospodarza. Podczas projektowania potencjalnie efektywnych ASO należy wziąć pod uwagę strukturę RNA, w celu określenia dostępności danego fragmentu RNA, tym samym możliwości przyłączenia danego ASO w tym miejscu. Badania przeprowadzone w naszym zakładzie wykazały skuteczność tak zaprojektowanych ASO w inhibicji namnażania wirusa^{10,12,122}. Wykazano inhibicję namnażania wirusa grypy na poziomie 80-96,2% (dla 5 z 12 zaprojektowanych cząsteczek) z zastosowaniem ASO nakierowanych na segment 8 vRNA¹⁰. Kolejne badania inhibicji wirusa z ASO nakierowanymi na vRNA5 wykazały zahamowanie namnażania wirusa o 88%¹². Co istotne, ten poziom inhibicji uzyskano celując na jednoniciowy rejon 887-

889 nt, który jest obecny w cząsteczce vRNA5 szczepów - A/Vietnam/1203/2004 jak i A/California/04/2009¹². Równie wysoką efektywność uzyskano celując w cząsteczkę mRNA 5 segmentu, co wskazuje na to, że zarówno vRNA jak i mRNA stanowią dobry target dla strategii antysensowej z zastosowaniem ASO¹²².

Stosowanie potencjalnych terapeutyków w postaci niemodyfikowanych ASO jest obarczone ryzykiem ich szybkiej degradacji w środowisku biologicznym. W tym celu stosuje się modyfikacje, które znacznie wpływają na ich stabilność, między innymi 2'-O-metylowane nukleotydy oraz modyfikacja LNA (ang. Locked Nucleic Acid) znacznie wpływają na stabilizację ASO w środowisku biologicznym, chroniąc przed działaniem nukleaz, a także stabilizują wiązanie ASO do docelowej cząsteczki RNA¹⁷⁸. Nasze badania potwierdziły wzmocnienie działania ASO poprzez zastosowanie modyfikowanych nukleotydów typu LNA^{12,10}. Podobnie jak w przypadku siRNA, również ASO są podatne na wywoływanie efektu niespecyficznego zahamowania. Wykazano jednak, że stosowanie modyfikacji LNA wpływa na zmniejszenie zjawiska *off-target*¹⁷⁹.

Metoda inhibicji namnażania wirusa grypy z zastosowaniem technologii CRISPR/Cas

Technologia CRISPR/Cas (ang. Clustered regularly interspaced short palindromic repeat, CRISPR) jest uznawana za przełomową na polu inżynierii genetycznej i na chwilę obecną opisano liczne ortologi białka Cas¹⁸⁰. Początkowo uważano, że jedynym zastosowaniem technologii jest manipulacja DNA. Z czasem odkryto ortologi białka Cas regulowane przez krótkie cząsteczki crRNA (ang. CRISPR RNA) umożliwiające bezpośrednią ingerencję w cząsteczkę RNA z zastosowaniem kierującej cząsteczki crRNA¹⁸¹. Grupę białek celujących w RNA opisuje się jako system CRISPR/Cas typu IV klasy 2, a w obrębie tej grupy dzieli się je dodatkowo ze względu na mechanizm działania. Do tej grupy zaliczane są białka Cas13, które mają duży potencjał do wykorzystania w terapii przeciwko wirusom¹⁸²⁻¹⁸⁵. Technologia CRISPR/Cas posiada pewne ograniczenia, które przekładają się na efektywność działania w warunkach biologicznych. Między innymi, badania aktywności ortologu białka Cas13 wskazały, że na efektywność działania tego białka bezpośredni wpływ ma struktura drugorzędowa RNA¹⁸². Wynika to z faktu, że w warunkach

biologicznych cząsteczka targetowego RNA wchodzi w wiele interakcji typu RNA-RNA czy RNA-białko, więc nie zawsze jest dostępna dla crRNA.

Bezpośrednie cięcie nici RNA, zarówno mRNA jak i cRNA, z zastosowaniem ortologu PspCas13b (ortolog białka pochodzący od bakterii *Prevotella sp. P5-125*, PsPCas13b) wykazało inhibicję namnażania wirusa grypy w rejonach konserwatywnych sekwencji RNA¹⁸⁴. Tym samym udowodniono, że białko to wykazuje aktywność zarówno na terenie jądra (cRNA) jak i cytoplazmy (mRNA). Autorzy badań również doszli do konkluzji, że zaprojektowanie efektywnych crRNA wymaga poznania struktury drugorzędowej RNA. Zastosowanie innego ortologu białka Cas13d wykazało efektywną inhibicję wirusa grypy¹⁸⁶. Autorzy badań skupili się na zaprojektowaniu cząsteczek crRNA nacelowanych na konserwatywne sekwencje IAV. Wykazali 72% inhibicji w przypadku kierowanej inhibicji zaprojektowanej na segment 6 vRNA.

Badania inhibicji wirusa grypy z zastosowaniem Cas13a zostały niedawno po raz pierwszy przeprowadzone na modelu zwierzęcym *in vivo*¹⁸⁷. Wyniki badań wskazały na efektywne obniżenie ilości wirusa w zainfekowanej tkance płucnej myszy. Targetem dla zaprojektowanych cząsteczek crRNA były rejony konserwatywne występujące w segmentach 1 i 2 wirusa grypy. Co istotne, efektywność działania systemu była testowana na zakażonym modelu zwierzęcym, a więc w momencie pełnego rozwinięcia infekcji.

Podobnie jak ASO i siRNA, również i metoda CRISPR/Cas obarczona jest pewnym ryzykiem efektu niespecyficznego działania. Tak więc, równie istotne dla tej metody jest prowadzenie badań dotyczących efektu *off-target* w komórkach ssaczych. Warto wspomnieć, że badanie aktywności Cas13a w komórkach ssaczych nie wykazało niespecyficznego cięcia RNA komórkowego¹⁸³. To jednoznacznie wskazuje na obiecującą przyszłość technologii CRISPR/Cas w leczeniu infekcji wywołanych wirusami układu oddechowego.

Wyniki i dyskusja

1. Struktura drugorzędowa vRNA8 w lizacie komórek MDCK

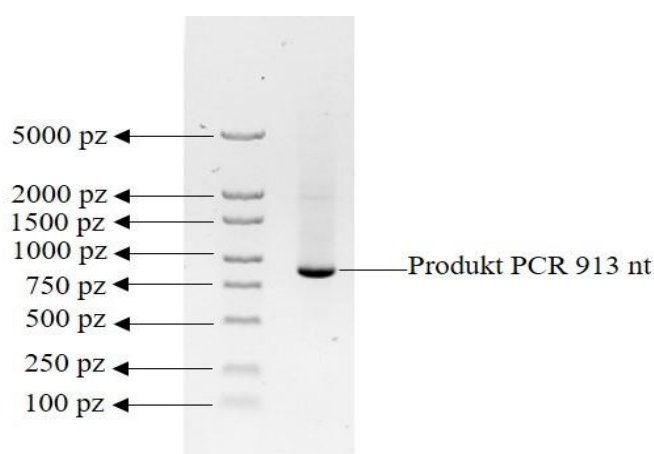
Interpretacja struktur RNA w warunkach biologicznych stanowi wyzwanie metodologiczne. Wynika to z faktu, że struktury RNA w warunkach biologicznych charakteryzują się dużą labilnością, a dana cząsteczka RNA może przyjmować wiele alternatywnych struktur. Struktura RNA w warunkach *in vivo* może różnić się od struktury przewidzianej *in vitro*, chociaż część motywów strukturalnych może pozostać zachowana. Badanie cząsteczki RNA, zarówno, w warunkach *in vitro*, jak i w warunkach środowiska aktywnego biologicznie, jest ograniczone. Przewidywanie struktury w warunkach *in vitro* nie daje informacji odnośnie jej fałdowania w czasie oddziaływań z białkami oraz kwasami nukleinowymi. Z kolei badanie struktury RNA w warunkach *in vivo* jest narażone na wiele zmiennych losowych, między innymi, zależy od momentu cyklu komórkowego czy czasu wykonania eksperymentu. Rozwiązaniem tego problemu jest stworzenie warunków mapowania *in-vivo-like* przypominających te panujące w komórce¹⁸⁸. W celu zbadania różnic między strukturą *in vitro* oraz tą otrzymaną w warunkach *in-vivo-like* wybrana została dobrze poznana cząsteczka segmentu 8 vRNA szczepu A/California/04/2009 (H1N1). Warunki przypominające te panujące w komórce zostały uzyskane poprzez stworzenie lizatu z zakażonych wirusem grypy komórek MDCK. Uzyskana struktura cząsteczki vRNA8 została następnie porównana do struktury vRNA8 uzyskanej w warunkach aktywnych biologicznie *in cellulo* oraz w wirionie *in virio*. Wyniki mapowania struktur *in cellulo* oraz *in virio* zostały opisane w drugiej części niniejszego rozdziału. Porównanie uzyskanej struktury w różnych warunkach mapowania miało na celu określenie czy warunki *in-vivo-like* mogą stanowić dobry model do badania RNA w warunkach biologicznych.

1.1. Otrzymanie matrycy do transkrypcji segmentu 8 vRNA oraz transkrypcja *in vitro*

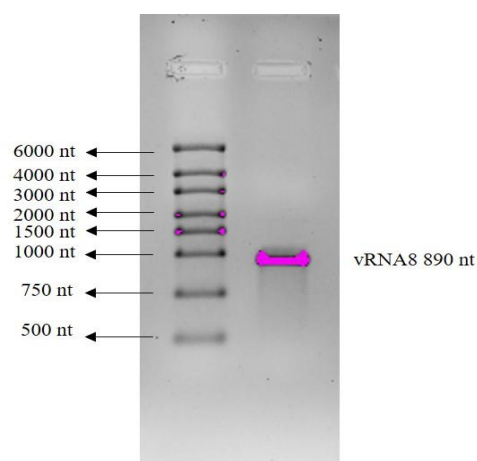
W celu otrzymania produktu DNA do transkrypcji *in vitro* vRNA8 wykonano reakcję PCR. Matrycę do reakcji stanowił plazmid pPolNS (otrzymany od prof. Baek Kim, Uniwersytet Emory, Atlanta, Georgia, USA). Na **Ryc. 18** widoczny jest produkt reakcji PCR rozdzielony na 1% żelu agarozowym. Widoczny produkt PCR (długość

913 pz) składa się matrycy vRNA8 o długości 890 pz wydłużonej o 23 pz sekwencji promotora dla polimerazy T7 RNA. Dokładny opis reakcji PCR oraz jej profil temperaturowy został zawarty w rozdziale *Materiały i metody*.

Na otrzymanej matrycy przeprowadzono następnie transkrypcję *in vitro* zgodnie z protokołem opisanym w rozdziale *Materiały i metody*. Uzyskany produkt reakcji transkrypcji rozdzielono na 1% żelu agarozowym (**Ryc. 17**) Wynik rozdziału wskazuje na prawidłową długość RNA bez produktów ubocznych.



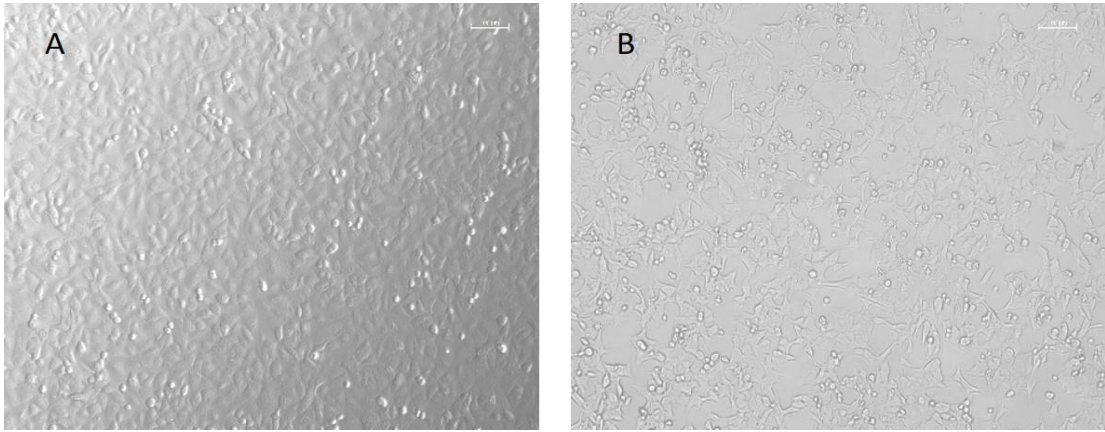
Ryc. 17 Rozdział elektroforetyczny na 1% żelu agarozowym produktu reakcji PCR matrycy do transkrypcji *in vitro* vRNA8. Jako standard długości zastosowano Perfect Plus MWQ DNA ladder (EurX). Na ścieżkę nałożono 300 ng oczyszczonego produktu PCR



Ryc. 18 Rozdział elektroforetyczny transkryptu *in vitro* vRNA8. Na żel nałożono 1 pmol produktu. Jako standard długości zastosowano RiboRuler High Range RNA Ladder (ThermoFisher).

1.2. Namnożenie wirusa A/California/04/2009 wykorzystywanego do eksperymentów komórkowych

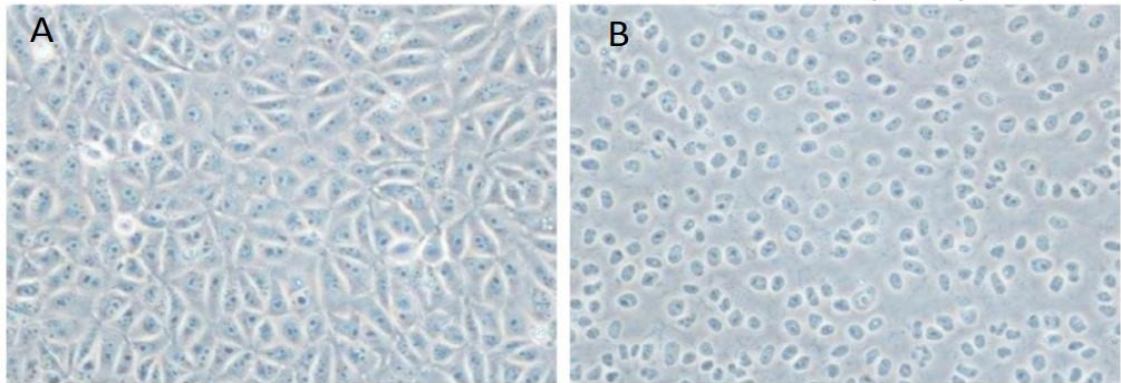
Propagację wirusa grypy typu A szczepu A/California/04/2009 przeprowadzono w linii komórkowej MDCK. Na **Ryc. 19** przedstawione są zdjęcia komórek MDCK wykonane bezpośrednio po infekcji (**Ryc. 19A**) oraz po ok. 48 godzinach po infekcji kiedy zaobserwowano widoczny efekt cytotoksyczny (**Ryc. 19B**). Efekt cytotoksyczny wskazuje na infekcję komórek. Infekcja komórek została potwierdzona za pomocą testu immunofluorescencji (IFA) według protokołu opisanego w rozdziale *Materiały i metody*.



Ryc. 19 **A** Zdjęcie komórek MDCK bezpośrednio po infekcji wirusa i dodaniu medium post-infekcyjnego. **B** Zdjęcie komórek po około 48 godzinach inkubacji w medium post-infekcyjnym. Widoczna zmiana zagęszczenia oraz kształtu komórek wskazuje na efekt cytotoksyczny będący wynikiem postępującej infekcji komórek. Zdjęcia spod mikroskopu wykonane aparatem Leica DFC450 C.

1.3. Infekcja komórek i przygotowanie lizatów komórkowych

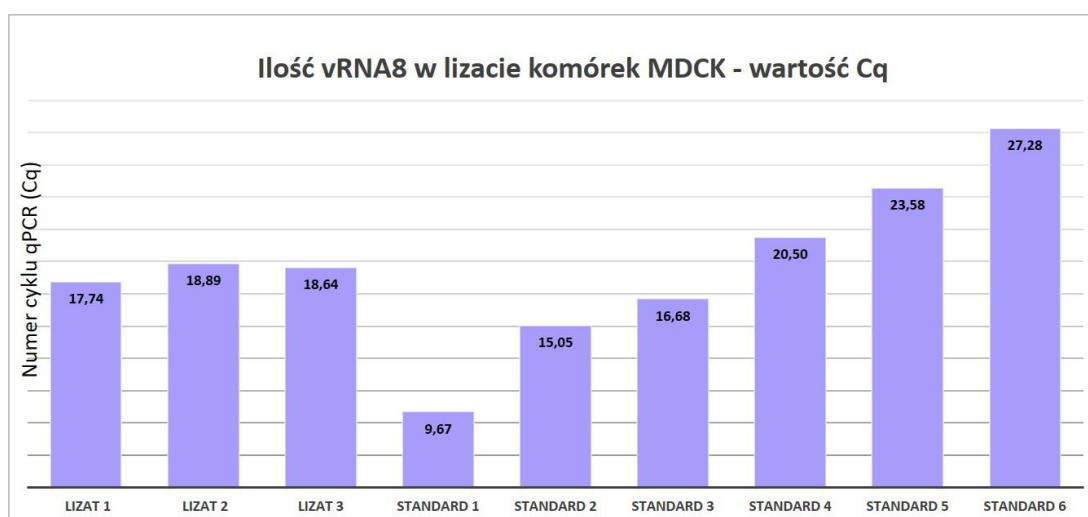
Po 24 godzinach od infekcji, komórki MDCK poddano lizie komórkowej z zastosowaniem buforu lizującego¹⁸⁹. Efekt lizy był kontrolowany za pomocą mikroskopu, poprzez obserwację morfologii komórek. Na **Ryc. 20** przedstawiony jest przykładowy obraz efektu lizy komórkowej przed (**Ryc. 20A**) oraz po inkubacji z buforem lizującym (**Ryc. 20B**)¹⁸⁹. Lizę komórkową prowadzono do momentu uzyskania obrazu mikroskopowego świadczącego o całkowitej lizie komórek.



Ryc. 20 **A** Morfologia komórek MDCK-London przed wykonaniem lizy komórkowej. **B** Komórki MDCK-London po 5 minutach inkubacji z zastosowaniem buforu lizującego. Zdjęcia pochodzą z publikacji Shatzkes et al. (Scientific Reports, 2014)

1.4. Sprawdzenie obecności vRNA oraz białek wirusowych w lizacie komórkowym

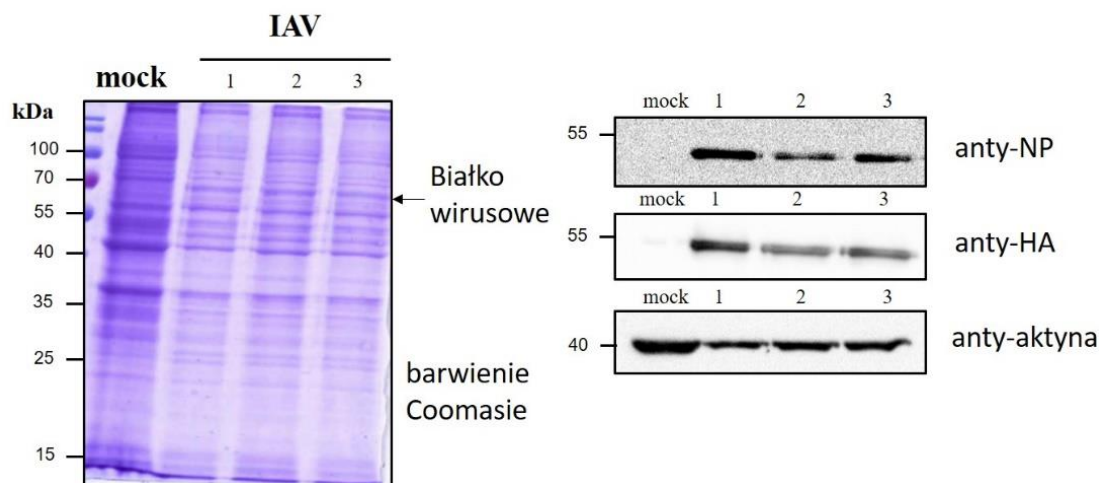
Przed przystąpieniem do właściwego eksperymentu mapowania, wykonana została reakcja PCR w czasie rzeczywistym. Umożliwiło to potwierdzenie obecności RNA wirusowego we frakcji całkowitego RNA. Reakcję real-time PCR przeprowadzono zgodnie z protokołem opisanym w sekcji *Materiały i metody* z zastosowaniem standardów umożliwiających określenie bezwzględnej ilości vRNA8 w badanej próbce RNA. Następnie przeprowadzona została analiza w programie Bio-Rad CFX Maestro (*Materiały i metody*). Wyniki reakcji real-time PCR wskazały na obecność vRNA8 w izolacie całkowitego RNA w zainfekowanych komórkach MDCK (*Ryc. 21*).



Ryc. 21 Na wykresie przedstawiona została wartość Cq (cykl qPCR), w którym został wykryty segment vRNA8 w mieszaninie całkowitego RNA w próbce. Na wykresie uwzględniono także wartości Cq dla kolejnych rozcieńczeń standardów ilościowych vRNA8.

Barwienie białek obecnych w lizacie komórkowym wskazało na obecność białek wirusowych (*Ryc. 22*; ścieżki 1-3) oraz ich brak w próbce kontrolnej (*Ryc. 22*; ścieżka mock). Następnie wykonany został Western Blot z przeciwciałami nakierowanymi na białka wirusowe – anty-NP oraz anty-HA, które potwierdziły ich obecność w lizacie komórkowym. Jako kontrolę zastosowano przeciwciało nakierowane na aktynę, która była obecna we wszystkich badanych próbkach – infekowanych oraz nieinfekowanych wirusem. Barwienie Coomassie Blue, a także

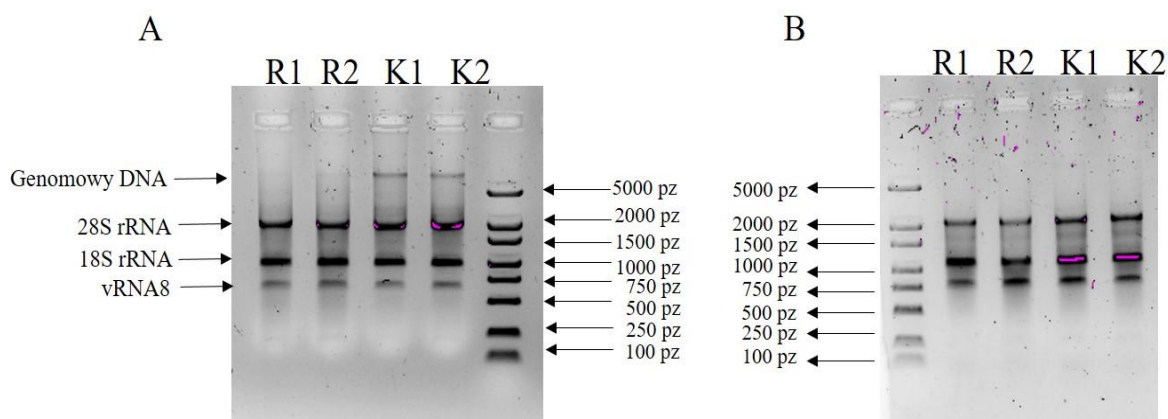
eksperymenty Western Blot zostały wykonane przez dr Pawła Zmorę (*Materiały i Metody*).



Ryc. 22 Barwienie białek obecnych w lizacie komórkowym z zastosowaniem Coomassie Blue wykazało obecność białek wirusowych w lizacie infekowanym wirusem (zaznaczone strzałką). Poniżej wynik Western Blot, który potwierdził obecność białek wirusowych z zastosowaniem przeciwciał nakierowanych na NP i HA. Jako kontrolę zastosowano przeciwciało nakierowane na aktynę. Barwienie Coomassie Blue, a także analizę Western Blot przeprowadzono dla trzech niezależnych lizatów wirusowo-komórkowych.

1.5. Mapowanie struktury drugorzędowej vRNA8 w lizacie komórkowym

Przed przystąpieniem do reakcji mapowania chemicznego, zdenaturowany transkrypt vRNA8 dodano do lizatu komórkowego i inkubowano w celu swobodnego fałdowania cząsteczki w obecności składników komórkowych i wirusowych. Po inkubacji wykonano mapowanie chemiczne z zastosowaniem odczynników 1M7 oraz DMS. Po reakcji mapowania, całkowity RNA oczyszczano na kolumnie RNeasy MinElute (QIAGEN). Następnie przeprowadzono reakcję enzymatycznej degradacji DNA enzymem DNazą I. W celu kontroli jakości RNA po reakcji mapowania (**Ryc. 23A**), a także po reakcji z DNazą I (**Ryc. 23B**), wykonano ponowny rozdział elektroforetyczny. Widoczne na żelu wyraźne prążki rRNA komórkowego świadczą o dobrej jakości RNA. Widoczny jest również transkrypt vRNA8. Zanieczyszczenie genomowym DNA (**Ryc. 23A**) zostało usunięte po reakcji enzymatycznej degradacji DNA (**Ryc. 23B**). Szczegółowy opis eksperymentów został zawarty w rozdziale *Materiały i Metody*.

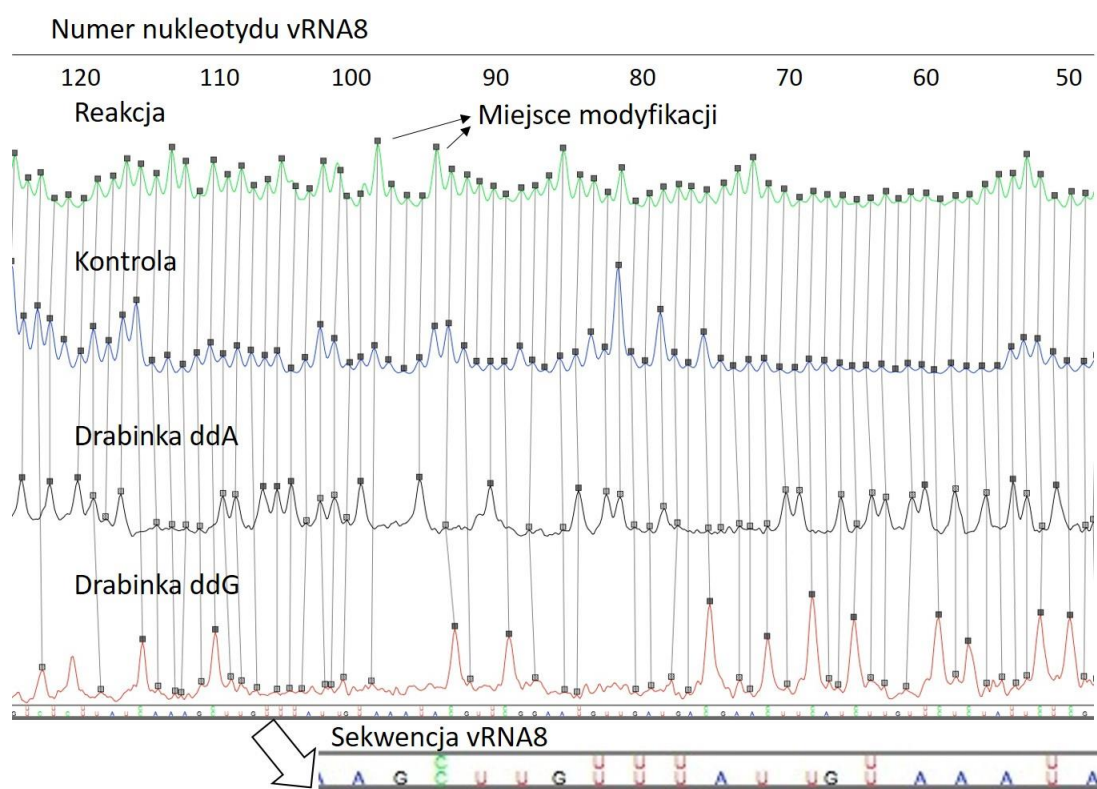


Ryc. 23 **A.** Rozdział elektroforetyczny izolatu całkowitego RNA z lizatu komórkowego wzbogaconego we frakcję transkryptu vRNA8 *in vitro*. Widoczne są prążki 28S oraz 18S rRNA oraz vRNA8 (dł. 890 nt), a także zanieczyszczenie genomowym DNA, **B.** Powtórzenie rozdziału elektroforetycznego próbek po przeprowadzeniu reakcji trawienia genomowego DNA enzymem DNazą I. Rozdział wskazuje na wydajne usunięcie genomowego DNA. Jako standard długości zastosowano drabinkę Perfect Plus MWQ (EurX). R1, R2 – niezależne powtórzenia mapowania dla próbki traktowanej odczynnikami chemicznymi DMS, K1, K2 – próbki kontrolne.

1.6. Wyniki mapowania chemicznego struktury drugorzędowej vRNA8 w lizacie komórkowym

Odczyt mapowań chemicznych RNA jest możliwy dzięki zastosowaniu reakcji odwrotnej transkrypcji (RT). Metoda odczytu mapowań chemicznych z zastosowaniem odwrotnej transkryptazy nazywana jest także reakcją wydłużania startera. Można ją zastosować do reagentów metody SHAPE (ang. Selective 2'-OH hydroxyl acylation analyzed by primer extension) i ma zastosowanie również w przypadku odczynników selektywnych dla zasad nukleotydowych np. DMS, CMCT i ketoksal. Rodzaje odczynników stosowanych do mapowania chemicznego zostały szczegółowo opisane w **Tabeli 2** w rozdziale *Wstęp literaturowy*. Enzym odwrotna transkryptaza nie rozpoznaje zmodyfikowanego chemicznie nukleotydu, co prowadzi do zatrzymania wydłużania komplementarnej nici cDNA na jeden nukleotyd przed zmodyfikowanym nukleotydem w nici RNA. W wyniku tego w czasie reakcji RT dochodzi do powstawania mieszaniny produktów cDNA o różnej długości. Odczyt zatrzymań reakcji RT jest możliwy dzięki rozdzieleniu produktów reakcji za pomocą elektroforezy kapilarnej. Obraz rozdziału wskazuje na miejsca kumulacji fragmentów o tej samej długości, które są interpretowane jako zatrzymania odwrotnej transkryptazy.

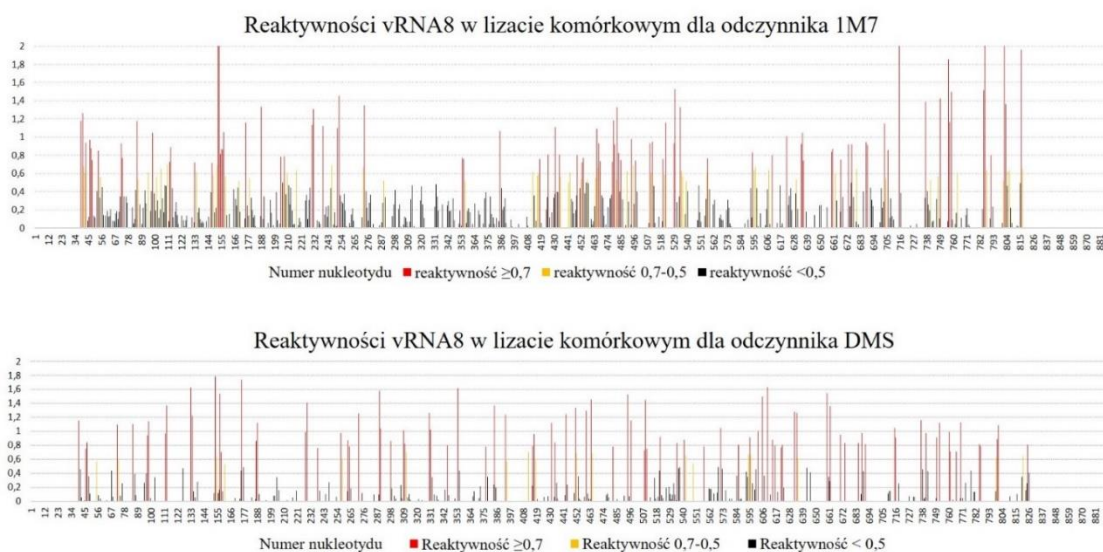
Po reakcji mapowania chemicznego całkowitego RNA wzbogaconego w transkrypt vRNA8 *in vitro* przeprowadzono reakcję RT. W celu pokrycia całej cząsteczki vRNA8 (długość 890 nt) zaprojektowano trzy startery w odległościach ok. 300 nt. Reakcję dla każdego startera (1_8, 2_8, 3_8, sekwencje starterów są wykazane w **Tabeli 8** (*Materiały i metody*) wykonano w minimum 3 powtórzeniach technicznych i biologicznych. Po precypitacji z zastosowaniem etanolu, próbki reakcji mapowania oraz kontroli mapowania łączono wraz z drabinkami sekwencyjnymi. Połączoną mieszaninę reakcji, kontroli oraz drabinek rozdzielano z zastosowaniem elektroforezy kapilarnej (Pracownia Technik Biologii Molekularnej UAM). Wyniki rozdziału analizowano w programie ShapeFinder (**Ryc. 24**)¹³⁴. Szczegółowy opis eksperymentów oraz analizy bioinformatycznej został zawarty w rozdziale *Materiały i Metody*.



Ryc. 24 Analiza rozdziału elektroforetycznego produktów odwrotnej transkrypcji po mapowaniu odczynnikami chemicznymi oraz drabinek sekwencyjnych ddA i ddG w programie ShapeFinder. Na rysunku widoczne są cztery kanały rozróżniane na podstawie użytych do poszczególnych reakcji fluoroforów. Zaznaczono również przykładowe miejsca mapowania chemicznego w próbce reakcyjnej.

Po przeprowadzeniu analizy wyników w programie ShapeFinder obliczone zostały reaktywności dla poszczególnych nukleotydów z zastosowaniem etapu normalizacji 2-8% (*Materiały i metody*)¹⁹⁰. Wyniki poszczególnych powtórzeń

połączono w celu wyliczenia średniej reaktywności dla wszystkich nukleotydów w cząsteczce vRNA8. Wyniki mapowania cząsteczki z zastosowaniem odczynnika DMS wykazały silne reaktywności ($\geq 0,7$) 89 nukleotydów (20,48% wszystkich A i C w sekwencji vRNA8), średnie reaktywności (0,7-0,5) 18 nukleotydów. Mapowanie odczynnikiem 1M7 wykazało reaktywność dla 56% wszystkich nukleotydów, z czego wysoką reaktywność miało 87 nukleotydów ($\geq 0,7$), natomiast średnią - 55 nukleotydów (0,7-0,5) (**Ryc. 25**). Rejony reaktywne względem odczynników mapujących występowały w vRNA8 w sposób zróżnicowany. Zaobserwowano miejsca o widocznie zwiększonej częstotliwości mapowań oraz miejsca o obniżonej reaktywności. Największą reaktywność nukleotydów zaobserwowano w rejonach: 39-48 nt, 147-158 nt, 448-458 nt, 464-470 nt, 476-485 nt oraz 801-806 nt. Niski poziom mapowania zaobserwowano w rejonach: 57-70 nt, 115-132 nt, 135-145 nt, 159-174 nt, 180-203 nt, 254-271 nt, 275-353 nt, 357-384 nt, 387-411 nt, 559-591 nt, 637-658 nt, 717-735 nt, 759-784 nt.



Ryc. 25 Średnia reaktywność nukleotydów vRNA8 w lizacie komórkowym określona na podstawie odczytu mapowania chemicznego odczynnikami 1M7 oraz DMS obliczona na podstawie powtórzeń biologicznych i technicznych. Kolorem czerwonym zaznaczono reaktywności o wartościach $\geq 0,7$, kolorem pomarańczowym reaktywności o wartościach w zakresie 0,7 – 0,5, na biało zaznaczone są reaktywności poniżej tych wartości.

Opublikowane niedawno wyniki mapowania struktury vRNA szczepu A/WSN/33 (H1N1) wykazały różnice w poziomie reaktywności (SHAPE) pomiędzy cząsteczkami mapowanymi w warunkach *in virio*, *ex virio* oraz *in vitro*¹¹⁵. Cząsteczki biologiczne, zarówno białka jak i kwasy nukleinowe, obecne w lizacie komórkowym

mogą swobodnie oddziaływać z cząsteczką vRNA. W efekcie można zaobserwować odmienny profil reaktywności w porównaniu do profilu uzyskanego dla vRNA8 w warunkach *in vitro*. W przypadku vRNA8 mapowanego w lizacie zaobserwowano zwiększone reaktywności nukleotydów dla obu odczynników chemicznych (DMS oraz 1M7).

1.7. Struktura drugorzędowa vRNA8 w lizacie komórkowym

W warunkach biologicznych z większym prawdopodobieństwem przewidywane są motywy o charakterze lokalnym. Podejście to nie jest pozbawione wad, gdyż utracona zostaje informacja o interakcjach pomiędzy oddalonymi od siebie nukleotydami. Z kolei fałdowanie struktury w ujęciu globalnym może również powodować powstawanie nieidealnego modelu cząsteczki, ponieważ (tak jak struktura lokalna) nie uwzględnia występujących międzycząsteczkowych interakcji typu RNA-RNA czy RNA-białko. W przypadku struktur mapowanych *in vivo* miejsca o widocznie niskim profilu reaktywności mogą być interpretowane w sposób dwojaki. Po pierwsze mogą być to miejsca, w których tworzą się pary zasad. Po drugie mogą być to miejsca oddziaływać z innymi cząsteczkami biologicznymi takimi jak inne RNA czy białka, wchodzące w interakcje z badaną cząsteczką w momencie mapowania chemicznego. Wobec tego interpretację niniejszej struktury postanowiliśmy przeprowadzać na kilku płaszczyznach.

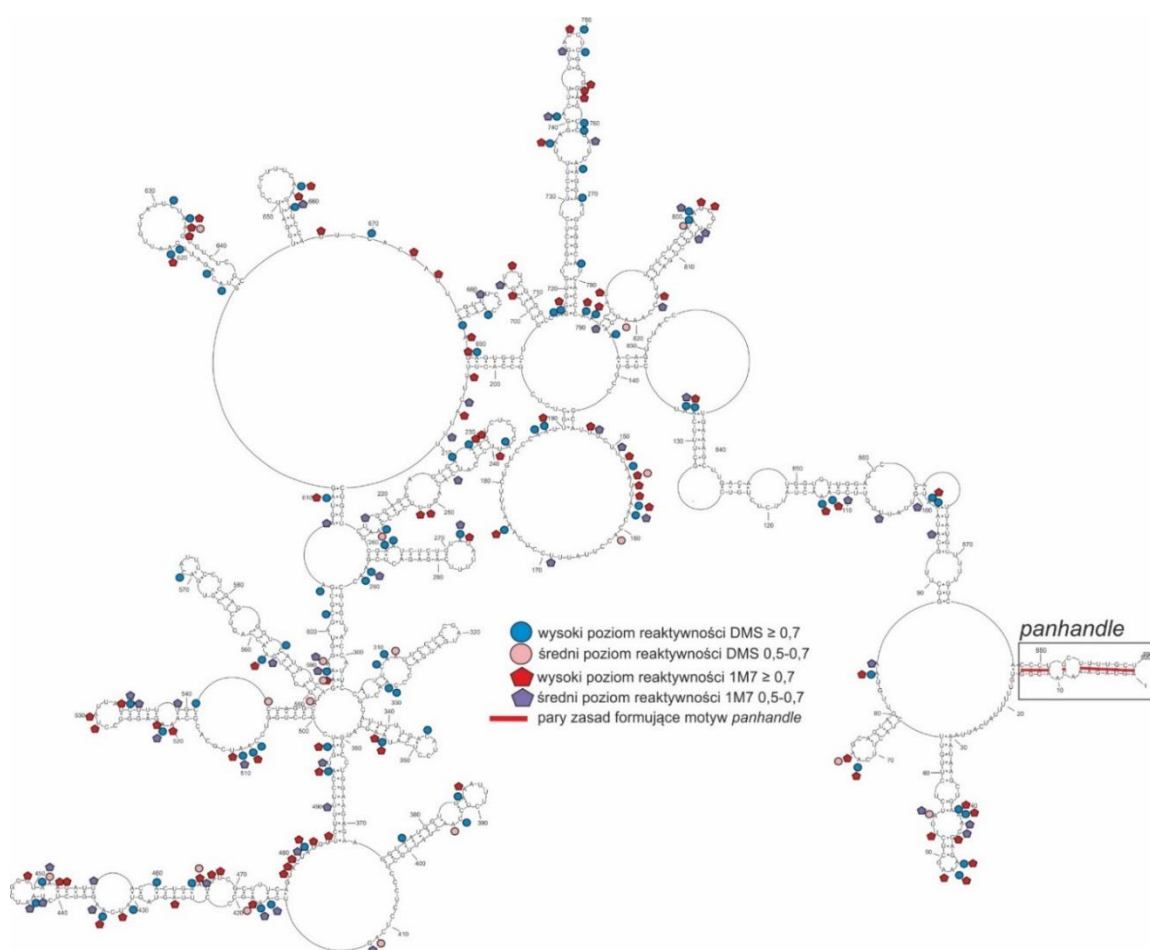
Struktura drugorzędowa vRNA8 w ujęciu globalnym

W pierwszej kolejności wygenerowano strukturę w ujęciu globalnym poprzez swobodne fałdowanie cząsteczki RNA bez ograniczenia odległości parowania zasad. Następnie zastosowano ograniczenie odległości parowania zasad w celu określenia lokalnych motywów strukturalnych. W wyniku predykcji w programie RNAStructure otrzymano zbiór wielu struktur o różnych wartościach energii swobodnej, a uzyskane struktury różniły się niewielką wartością tej energii. W przypadku przewidywań struktur krótkich RNA najczęściej opisuje się strukturę o najmniejszej wartości energii swobodnej (ang. Minimum Free Energy, MFE). W czasie predykcji struktury MFE wykorzystywane są parametry termodynamiczne Modelu Najbliższego Sąsiedztwa (ang. nearest neighbor)¹⁹¹. Przewidywania struktur długich RNA, w szczególności w warunkach biologicznych, wymaga zastosowania rozszerzonego podejścia w

interpretacji wyników eksperymentalnych. Wobec tego, w czasie analizy motywów strukturalnych w obrębie tych struktur nie skupiono się jedynie na strukturze o najniższej energii swobodnej, ale porównano ją również do struktur MEA (ang. Maximum Expected Accuracy structure, MEA structure)¹⁹². W trakcie predykcji struktury MEA oprócz danych termodynamicznych wykorzystywane są również dane dotyczące prawdopodobieństwa parowania zasad (ppz). Prawdopodobieństwo to jest obliczone na podstawie wszystkich, uzyskanych na podstawie danych eksperymentalnych, struktur. Takie samo podejście w interpretacji wyników zastosowano w przypadku wielu struktur określanych w warunkach *in vivo*, w tym struktur mRNA oraz vRNA wirusa grypy^{126,115}.

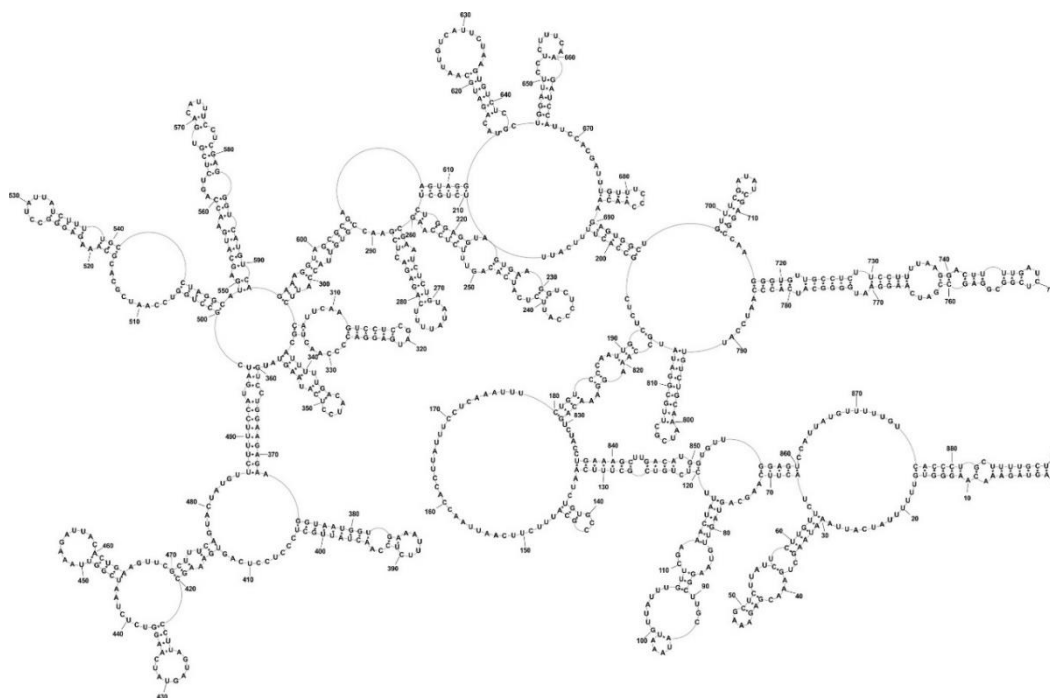
Średnią wartość reaktywności nukleotydów, obliczoną z powtórzeń eksperymentalnych, wprowadzono do programu RNAstructure (wersja 6.2). W przypadku mapowania odczynnikiem SHAPE (1M7) wprowadzono reaktywności dla każdego nukleotydu, za wyjątkiem tych których reaktywności nie udało się odczytać. Podczas rozdziału kapilarnego dochodzi do związanego z metodą ograniczenia odczytu obu końców cząsteczki. W tym wypadku nie udało się uzyskać odczytów w zakresie 1-38 nt i 818-890 nt (dla odczynnika 1M7) oraz 1-13 nt i 827-890 nt (dla odczynnika DMS). Modyfikacje DMS wprowadzono jako osobny plik z numerami nukleotydów o wysokiej reaktywności ($\geq 0,7$) (w module *chemical modification constraints*). W przypadku struktury fałdowanej w ujęciu globalnym wprowadzono dodatkowo pary zasad pomiędzy nukleotydami 5' i 3' końców, które tworzą strukturę *panhandle*, będącą promotorem wirusowej polimerazy RdRp. Formowanie się struktury *panhandle* w kompleksie vRNP zostało potwierdzone eksperymentalnie^{46,118}.

W ujęciu globalnym w strukturze drugorzędowej vRNA8 formują się zarówno wyraźne domeny silnie ustrukturyzowane, w obrębie których występują liczne motywy typu ‘spinka’ jak i długie rejonu jednoniciowe (**Ryc. 26**). Na rysunku zaznaczona została strukturę *panhandle* (rejon 1-11/881-890 nt). Zidentyfikowanych zostało 17 motywów strukturalnych typu ‘spinka’. Wyniki mapowania korelują ze strukturą drugorzędową. Miejsca o zwiększonej reaktywności przypadają na rejonu jednoniciowe takie pętle, rejonu jednoniciowe w obrębie spinek, a także miejsca pojedynczych niesparowań. Najwyższą reaktywność nukleotydów zaobserwowano w rejonie 146-190 nt, w którym formuje się długa pętla.



Ryc. 26 Struktura drugorzędowa vRNA8 szczepu A/California/04/2009 określona na podstawie mapowania chemicznego w lizacie komórek MDCK infekowanych wirusem grypy. Struktura o najniższej wartości energii swobodnej (MFE). Struktura została wygenerowana na podstawie danych eksperymentalnych. Czerwoną kreską zostały zaznaczone pary zasad użyte dodatkowo podczas fałdowania RNA w programie RNAstructure. Na rycinie zaznaczone zostały silne oraz średnie mapowania chemiczne nukleotydów odczynnikami DMS i 1M7.

Uzyskaną strukturę drugorzędową o najniższej energii swobodnej porównano do struktury MEA (**Ryc. 27**). Obie struktury były do siebie bardzo podobne i większość motywów strukturalnych przewidzianych w strukturze MFE była również obecna w strukturze MEA. Największe różnice występują w rejonach: 76-116 nt, 137-145 nt, 180-193 nt oraz 421-469 nt. W obrębie dwóch motywów strukturalnych w rejonach 305-335 nt oraz 645-666 nt zauważono lekką różnicę w parowaniu zasad.



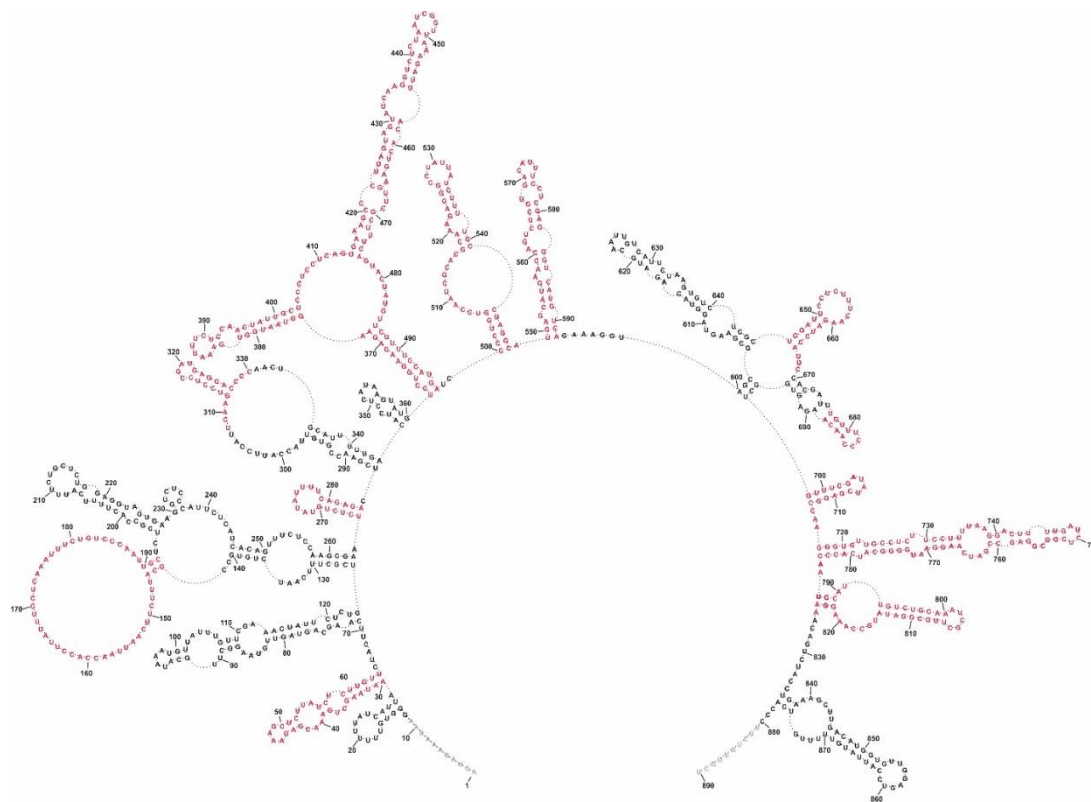
Ryc. 27 Struktura globalna MEA cząsteczki vRNA8 wygenerowana na podstawie danych eksperymentalnych i dotyczących prawdopodobieństwa parowania zasad.

Struktura drugorzędowa vRNA8 w ujęciu lokalnym

Duże cząsteczki RNA fałdując się mogą tworzyć w obrębie struktury wiele oddziaływań dalekiego zasięgu. W przypadku cząsteczek RNA, które nie tylko wchodzi w interakcję typu RNA-RNA, ale i również RNA-białko, określenie oddziaływań dalekiego zasięgu za pomocą standardowo używanych metod mapowania chemicznego jest bardzo trudne i wymaga zastosowania dodatkowych metod biofizycznych. Pomimo pręźnie rozwijających się metod wysokorozdzielczych min. mikroskopii Cryo-EM, które umożliwiają badanie kompleksów rybonukleoproteinowych *in virio*, metody te wymagają materiału w postaci oczyszczonego kompleksu vRNP¹⁹³. W związku z tym nadal istnieje ograniczenie w zastosowaniu tej metody w badaniach w warunkach biologicznych *in cellulo*. Alternatywą są metody oparte o zjawisko fotozszywania (ang. crosslinking), z

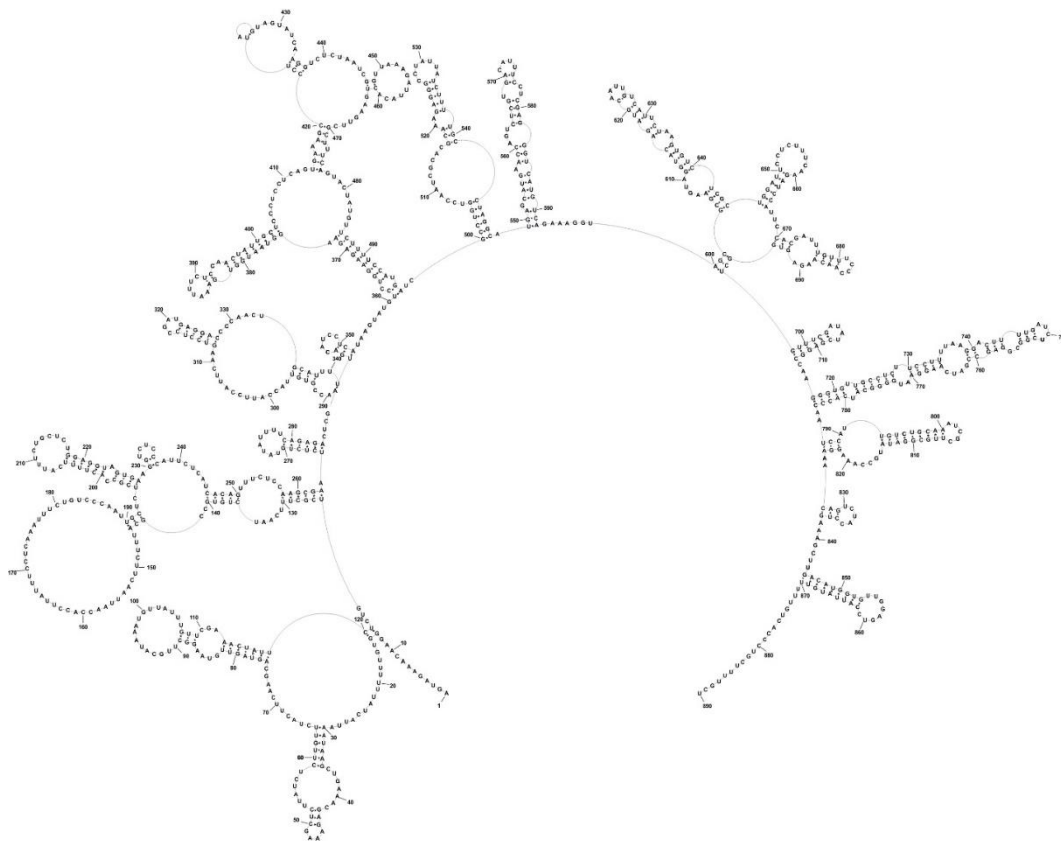
zastosowaniem odczynników takich jak psoralen lub 4-tiourydyna, które umożliwiają bezpośrednią identyfikację interakcji RNA-RNA, a także RNA-białko^{194,195}. Co istotne, metody fotozszywania znalazły zastosowanie w przypadku badania interakcji RNA-RNA wirusa grypy^{115,118}. Pojawiło się również wiele publikacji dotyczących oddziaływań wirusowego RNA z białkami, między innymi z białkiem NP w kompleksie vRNP^{196,197,47,198}. Metody fotozszywania są dobrym uzupełnieniem danych dotyczących struktury drugorzędowej, ponieważ dają informację o rejonach dwuniciowych w obrębie cząsteczki, a także miejsc oddziaływań pomiędzy różnymi cząsteczkami RNA oraz RNA-białko. W przypadku struktur RNA, które określane są w warunkach biologicznych warto zastosować kilka podejść podczas modelowania struktury drugorzędowej. W literaturze stosowane jest podejście modelowania struktur z zastosowaniem dystansu parowania zasad, tym samym umożliwia to identyfikację motywów strukturalnych o charakterze lokalnym. Dla długich RNA często stosuje się dłuższy dystans parowania zasad, np. struktura RNA całego genomu SARS CoV-2 fałdowana została z zastosowaniem dystansu parowania równym 500 nt¹⁹⁹. Taką samą wartość dystansu zastosowano w przypadku modelowania struktur wirusowych mRNA (IAV)¹²⁶. W przypadku struktur mRNA, które nie wchodzą w skład kompleksu rybonukleoproteinowego, oddziaływania dalekiego zasięgu pomiędzy nukleotydami są bardziej prawdopodobne. W przypadku struktur cząsteczek krótszych, wchodzących w liczne interakcje, stosuje się mniejszy dystans parowania zasad, np. cząsteczki vRNA wirusa grypy typu A *in virio* fałdowano z zastosowaniem dystansu 150 nt¹¹⁵. Trzeba również pamiętać o zwykłych ograniczeniach programowych w przewidywaniu struktury drugorzędowej RNA wynikających z dostępnych danych i możliwości bioinformatycznych. Poza tym, wprowadzenie danych eksperymentalnych do programu takiego jak RNAstructure powoduje wygenerowanie wielu struktur drugorzędowych, które często bardzo nieznacznie różnią się między sobą wartością energii swobodnej. Ma to miejsce zwłaszcza w przypadku długich RNA. Kiedy istnieje niewielka różnica pomiędzy strukturą o najniższej energii i kolejnymi przewidywanymi strukturami, wybór właściwej struktury jest bardzo trudny. Dlatego też w niniejszej pracy wygenerowano strukturę MEA dla vRNA8 w warunkach lizatu komórkowego.

Dodatkowo, w celu wygenerowania struktury drugorzędowej uwzględniając lokalny charakter parowania się zasad zastosowano maksymalny dystans parowania zasad równy 150 nt. Podczas modelowania struktury wykluczono rejon 3' i 5' cząsteczki, które parują się między sobą tworząc motyw *panhandle*. W tym celu wprowadzono dodatkowe ograniczenie fałdowania poprzez wymuszenie w tym miejscu rejonu jednoniciowego. Uzyskana struktura (**Ryc. 28**) wykazała zachowanie 12 z 17 struktur typu spinka, przewidzianych w globalnej strukturze vRNA8 w rejonach: 30-64 nt, 265-284 nt, 312-327 nt, 374-403 nt, 414-476 nt, 500-547 nt, 549-592 nt, 645-666 nt, 677-688 nt, 698-713 nt, 717-784 oraz 787-824 nt. Struktury formujące się pomiędzy nukleotydami oddalonymi od siebie o więcej niż 150 nt nie powstały ze względu na wprowadzone ograniczenia.



Ryc. 28 Struktura drugorzędowa vRNA8 w lizacie komórkowym uzyskana z zastosowaniem dystansu parowania zasad wynoszącym 150 nt. Uzyskana struktura wykazała zachowanie większości struktur typu spinka oraz parowania zasad w rejonach helikalnych, przewidzianych w globalnej strukturze vRNA8 (zaznaczone na czerwono). Rejon motywu panhandle został wykluczony z przewidywania struktury.

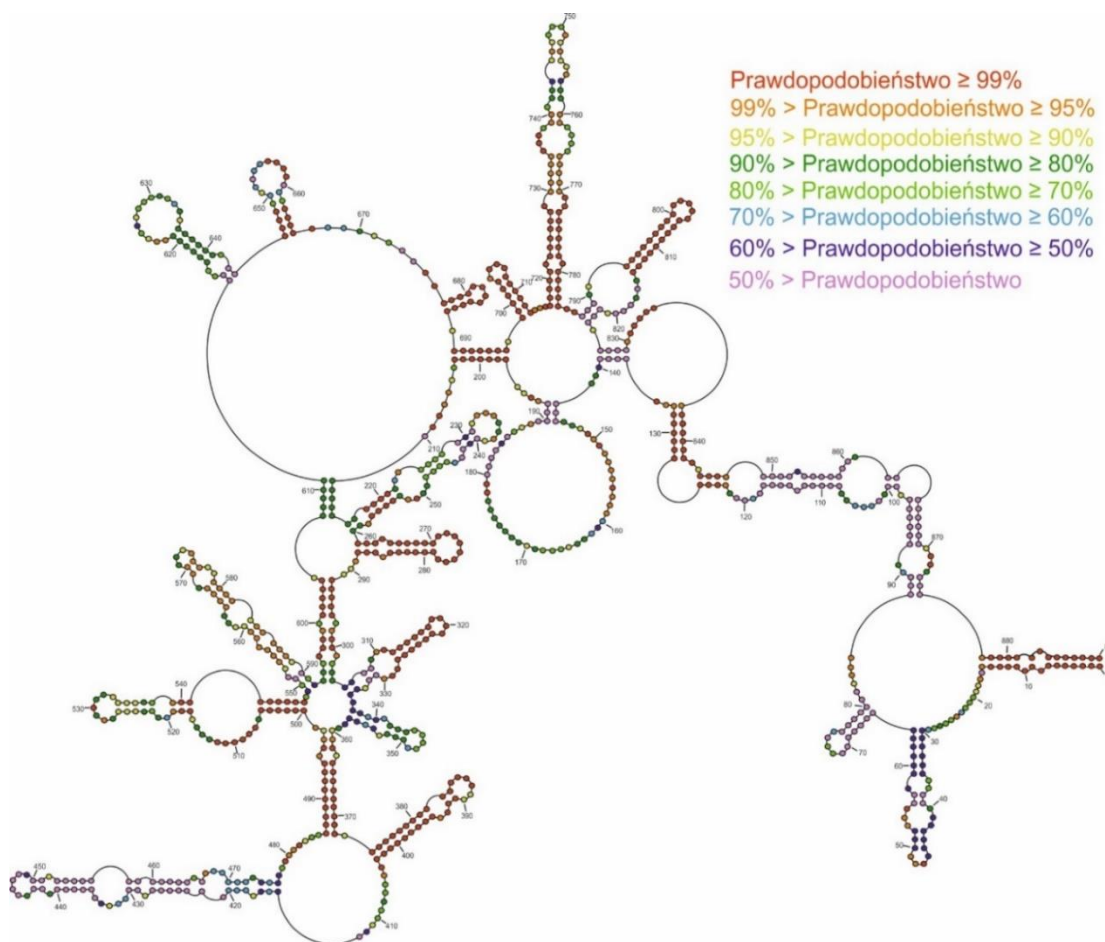
Następnie wygenerowano strukturę MEA, która wykorzystuje dane dotyczące prawdopodobieństwa parowania zasad stosując dystans parowania zasad wynoszący 150 nt (**Ryc. 29**). Porównanie ze strukturą lokalną (**Ryc. 28**) wykazało istnienie różnic. Chociaż większość motywów strukturalnych została zachowanych, zaobserwowano różnice w rejonie 5' końca (13-124 nt) oraz 3' końca (826-845 nt) cząsteczki. Różnice w obrębie motywów strukturalnych zaobserwowano w spinkach 345-360 nt oraz 414-476 nt.



Ryc. 29 Struktura MEA segment 8 vRNA wygenerowana z zastosowaniem dystansu parowania zasad 150 nt. Rejon motywu panhandle został wykluczony z przewidywania struktury.

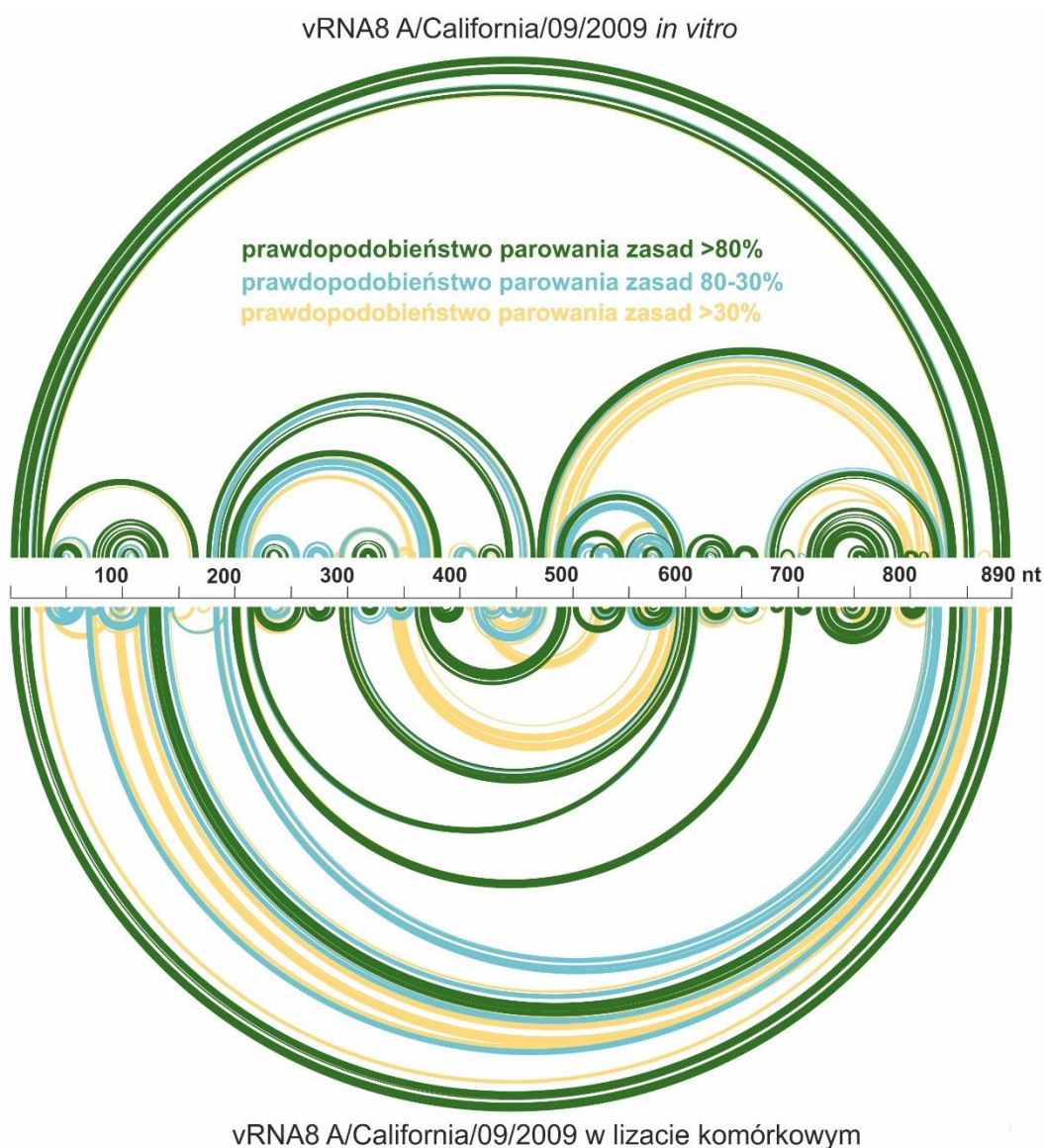
1.7.1. Prawdopodobieństwo parowania zasad

Uzyskane w programie RNAStructure prawdopodobieństwo parowania zasad, a także prawdopodobieństwa występowania nukleotydów w rejonach jednoniciowych, w strukturze MFE vRNA8 naniesiono na strukturę globalną (**Ryc. 30**). Najwyższe prawdopodobieństwo (>90%) zaobserwowano w 10 motywach strukturalnych w rejonach 261-288 nt, 310-329 nt, 374-404 nt, 500-547 nt, 553-589 nt, 645-666 nt, 677-687 nt, 698-713 nt, 717-784 nt and 792-814 nt. Najwyższe prawdopodobieństwo parowania zasad dalekiego zasięgu zaobserwowano dla nukleotydów w rejonach 123-126/845-848 nt, 127-133/842-836 nt, 197-203/696-690 nt, 292-304/605-593 nt (niesparowane 297/600 i 301/596 nt) i 360-372/598-486 nt (niesparowane 363/495 nt).



Ryc. 30 Prawdopodobieństwo występowania nukleotydów w rejonach jednoniciowych oraz tworzenia przez nie par zasad w obrębie struktury drugorzędowej vRNA8 w lizacie komórkowym. Obliczenia wykonano w programie RNAStructure 6.2 z wykorzystaniem modułu Partition Function. Kolory wskazują na procent prawdopodobieństwa

Następnie porównano wyniki prawdopodobieństwa parowania zasad pomiędzy strukturą vRNA8 uzyskaną w lizacie komórkowym oraz strukturą *in vitro* (Marta Soszyńska-Józwiak et. al. manuskrypt w recenzji). Wizualizację ppz dla obu struktur wykonano w programie IGV (**Ryc. 31**). Program umożliwia wizualizację prawdopodobieństw w określonym zakresie: >80%, 80-30 % oraz < 30%. Porównanie wykazało wspólne motywy strukturalne, a najwyższe prawdopodobieństwo (>80%) uzyskano w rejonach: 312-327 nt, 500-547 nt, 645-666 nt oraz 797-809 nt.



Ryc. 31 Wykresy prawdopodobieństwa na podstawie danych eksperymentalnych struktur vRNA8 fałdowanych w warunkach *in vitro* oraz w lizacie komórkowym. Łukami zaznaczone są parowania pomiędzy poszczególnymi nukleotydami, natomiast kolory wskazują prawdopodobieństwo ich tworzenia. Obliczenia wykonano w programie RNAstructure 6.2. Wizualizacja w programie IGV

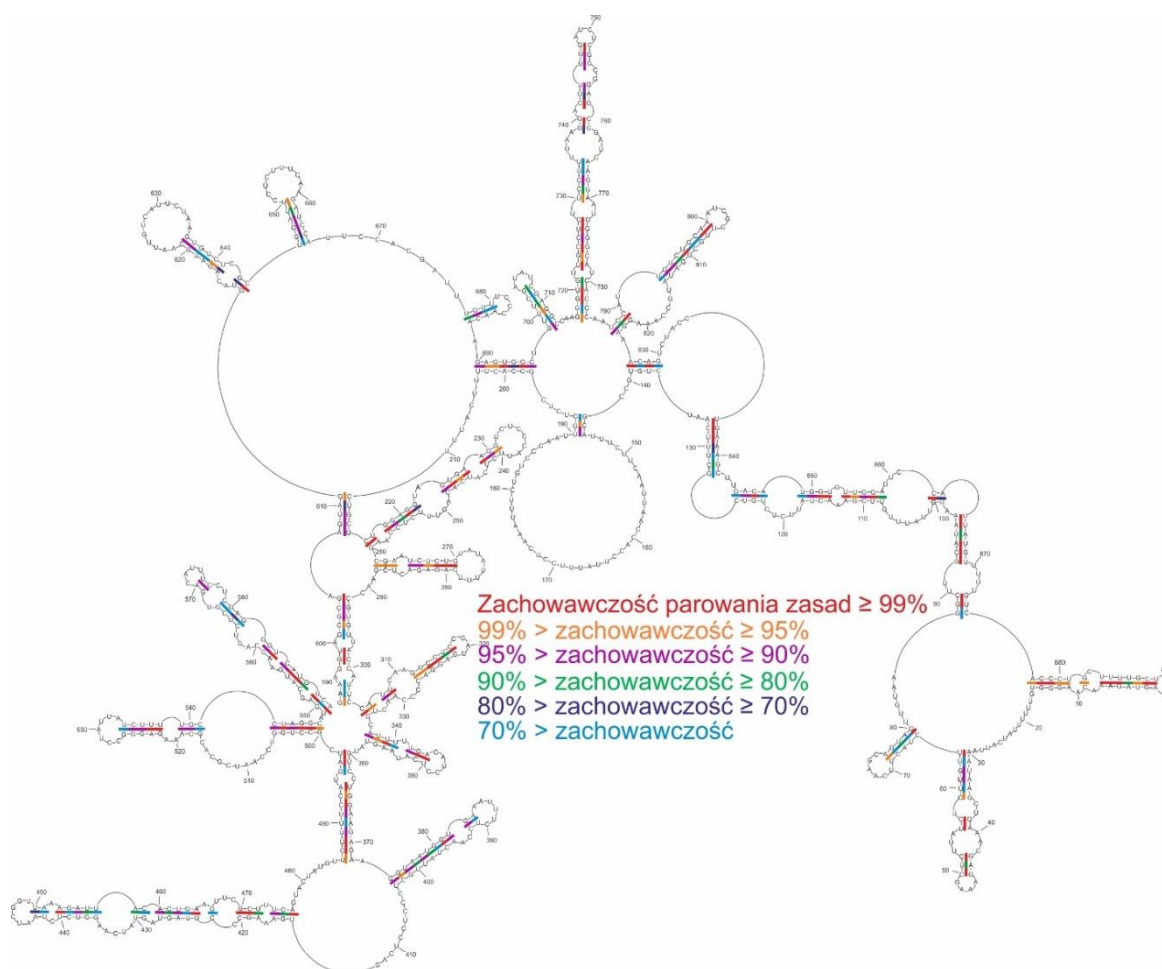
1.7.2. Analiza zachowawczości parowania zasad dla typu A wirusa grypy w strukturze drugorzędowej vRNA8 w lizacie komórkowym

Zachowawczość parowania zasad dla typu A wirusa grypy obliczona została na podstawie globalnej oraz lokalnej struktury vRNA8 (**Ryc. 26, Ryc. 28**) w odniesieniu do danych sekwencyjnych wszystkich dostępnych w bazie danych (Influenza Research Database) sekwencji szczepów wirusa grypy typu A. Obliczenia zachowawczości parowania zasad przeprowadził dr Walter Moss oraz mgr Jake Peterson (Departament Biochemii, Biofizyki oraz Biologii Molekularnej Roy J. Carver, Uniwersytet Stanowy Iowa). Analiza zachowawczości parowania zasad informuje z jakim prawdopodobieństwem dana para może powstać w obrębie struktur vRNA innych szczepów wirusa grypy. Podczas obliczeń brane są pod uwagę wszystkie możliwe pary zasad (również pary G-U) oraz niesparowania (brak danej pary zasad). Wysoka zachowawczość danej pary zasad w danym motywie strukturalnym wskazuje na jego ewolucyjną konserwatywność, a tym samym wskazuje na potencjalną funkcjonalność danego motywu w czasie cyklu replikacyjnego wirusa.

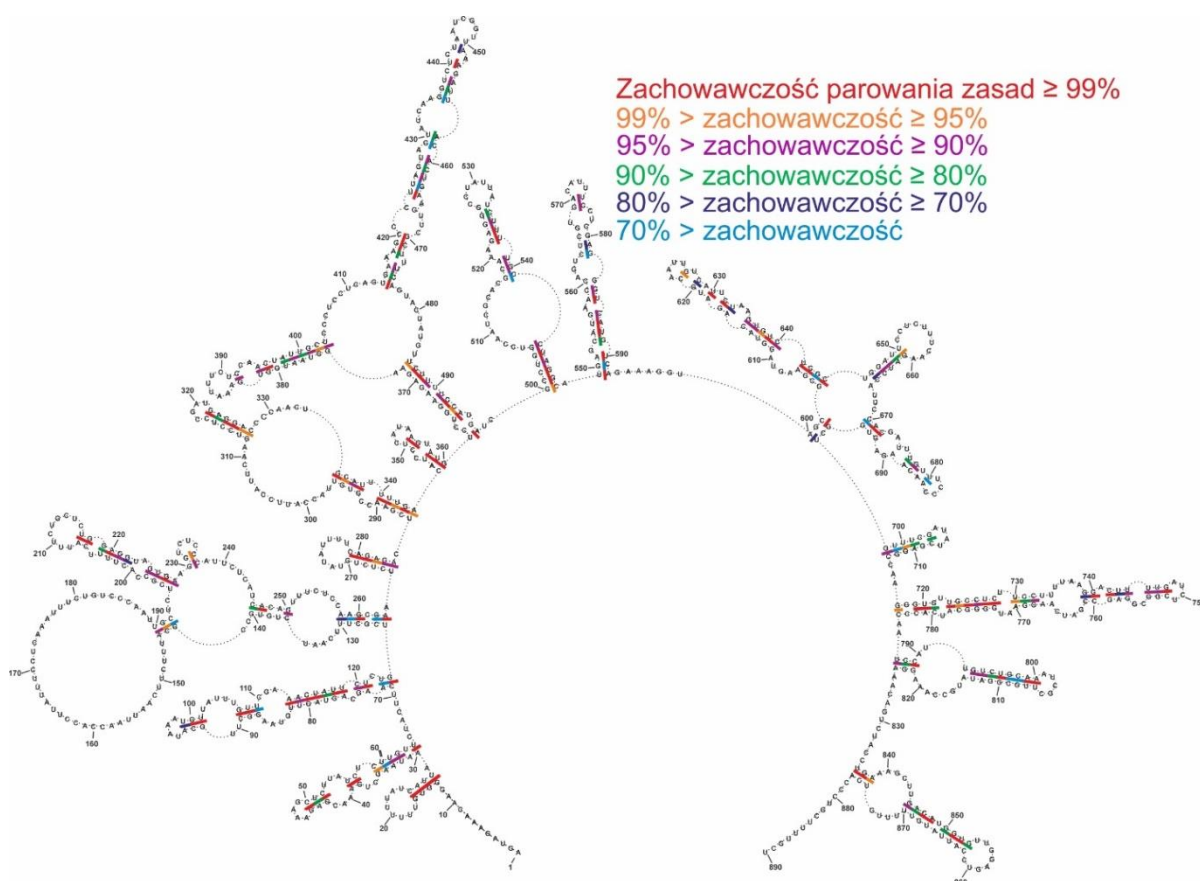
Struktura vRNA8 fałdowana w ujęciu globalnym wykazała średnią zachowawczość parowania zasad równą 86,38%, natomiast w ujęciu lokalnym 86,45%. Zachowawczość parowania zasad w obrębie struktury globalnej oraz lokalnej została zaprezentowana na **Ryc. 32 i Ryc. 33**. Badania zachowawczości parowania zasad wskazały wysoką konserwatywność w parowaniu typu dalekiego zasięgu w rejonach: 292-299/605-598 nt (91,87%), 197-203/696-690 nt (92,58%), 88-126/875-845 nt (średnia 90,85%; 95,74% dla rejonu 88-116/875-850 nt) oraz 360-372/498-486 nt (87,37%). W przypadku motywów strukturalnych zachowawczość parowania zasad na poziomie ~90% zaobserwowano w obu strukturach w rejonach: 30-64 nt (89,05%), 265-284 nt (97,03%), 312-327 nt (96,09%), 500-547 nt (89,40%) oraz 717-784 nt (88,87%). Motyw w rejonie 216-260 nt obecny w strukturze lokalnej miał 94,99% zachowawczości. Dwa motywy o wysokiej zachowawczości (92,14% w rejonie 71-124 nt oraz 90,37% w rejonie 836-877 nt) były przewidziane w strukturze lokalnej. Chociaż część motywów w strukturze vRNA8 w lizacie wykazywało wysoki poziom prawdopodobieństwa parowania zasad (**Ryc. 14**), to analiza zachowawczości tych

motywów wskazała niski poziom ich konserwatywności wśród innych szczepów IAV: 549-592 (77,81%), 645-666 (81,50%), 787-824 (84,91%).

Analiza zachowawczości parowania zasad w strukturze *in vitro* wskazała na wysoką zachowawczość parowania w rejonie formującym długą helisę 696-780 nt. W przypadku struktury vRNA8 w lizacie zaobserwowano różnice w parowaniu zasad w obrębie tego rejonu. Zmiany strukturalne mogą wynikać z oddziaływań tego regionu ze składnikami obecnymi w lizacie komórkowym. Co ciekawe, badania profilu wiązania białka NP do cząsteczki vRNA8 wskazały na wyplatanie się cząsteczki w tym rejonie poza obszar kompleksu vRNP⁴⁶.



Ryc. 32 Zachowawczość parowania zasad w strukturze globalnej vRNA8 przewidzianej na podstawie mapowania chemicznego w lizacie komórkowym.



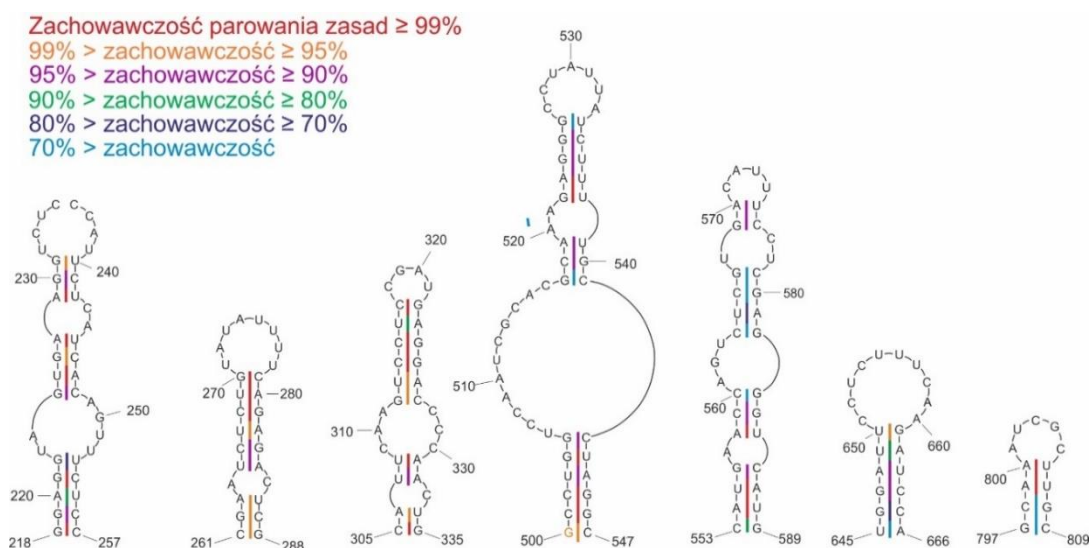
Ryc. 33 Zachowawczość parowania zasad w strukturze lokalnej vRNA8 przewidzianej na podstawie mapowania chemicznego w lizacie komórkowym.

Konserwatywne motywy strukturalne vRNA8

Struktury drugorzędowe vRNA8 *in vitro* szczepów A/Vietnam/1203/2004 (H5N1)¹⁰ oraz A/California/04/2009 (H1N1) (Marta Soszyńska-Jóźwiak et. al., manuskrypt w recenzji) zostały zaproponowane przez naszą grupę. Na podstawie danych eksperymentalnych z mapowania szczepu A/Vietnam/1203/2004 oraz analiz sekwencyjno-strukturalnych, przeprowadzonych na podstawie wszystkich dostępnych w bazie danych sekwencji IAV, zaproponowano również uniwersalną strukturę vRNA8 dla wirusa grypy typu A (struktura *in silico*)¹⁴. Nasze badania wskazały, że motywy strukturalne przewidziane w badaniach *in silico* były zachowane w strukturach *in vitro* A/California/04/2009 (Marta Soszyńska-Jóźwiak et. al., manuskrypt w recenzji). Porównanie struktury drugorzędowej vRNA8 uzyskanej w lizacie komórkowym ze strukturą uniwersalną vRNA8 również wykazało zachowanie części motywów strukturalnych przewidzianych *in silico*. Niedawno opublikowane zostały struktury drugorzędowej vRNA szczepu A/WSN/1933 (H1N1) w warunkach

in vitro, *in virio* oraz *ex virio*¹¹⁵. Autorzy ww. pracy byli zainteresowani bardziej lokalnym charakterem powstawania motywów strukturalnych vRNA w obrębie struktur powstających w warunkach biologicznych i zastosowali ograniczenie parowania zasad do 150 nt. Zastosowanie tych samych warunków fałdowania umożliwiło porównanie struktury w lizacie ze strukturą vRNA8 szczepu A/WSN/33 uzyskaną w warunkach *in virio*.

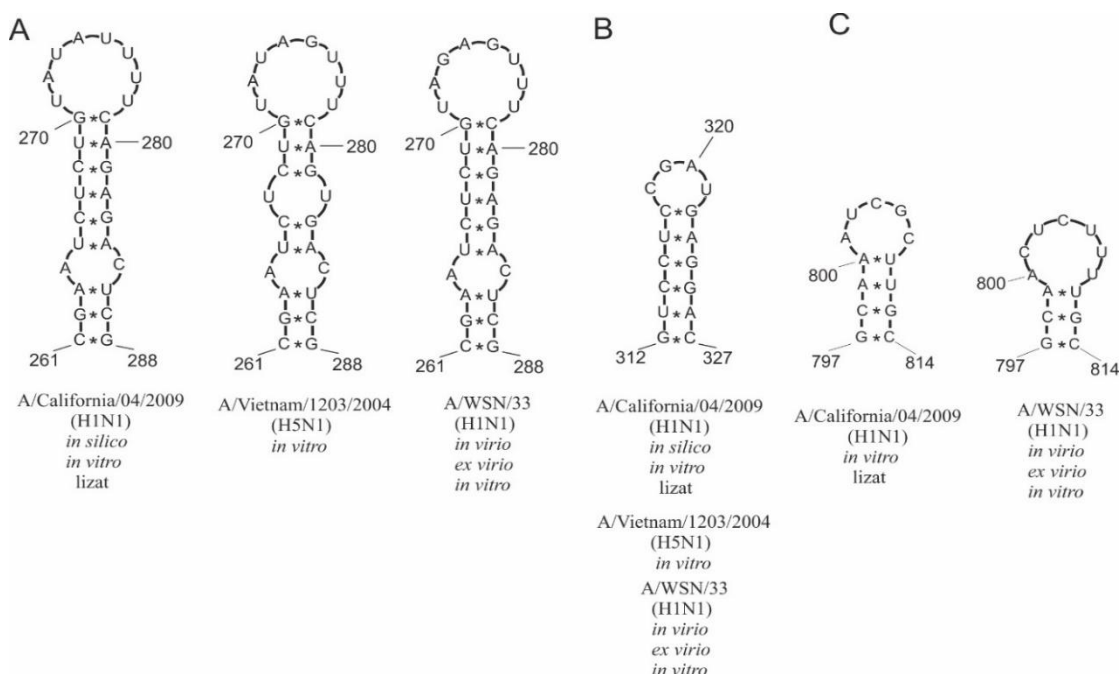
Analiza zachowawczości motywów strukturalnych w strukturach *in vitro* oraz w lizacie wykazała zarówno różnice jak i motywy wspólne. Różnice pomiędzy strukturami *in vitro* oraz strukturą vRNA8 w lizacie mogą wynikać z większej labilności struktury w środowisku biologicznym. Na **Ryc. 34** przedstawiono motywy strukturalne vRNA8 przewidziane w strukturze *in vitro* oraz w lizacie komórkowym wraz z uwzględnieniem zachowawczości parowania zasad. Część motywów wspólnych dla obu struktur wykazuje niski poziom zachowawczości parowania zasad.



Ryc. 34 Motywy strukturalne wspólne dla vRNA8 *in vitro* oraz w lizacie komórkowym. Kolorami zaznaczono zachowawczość parowania zasad w obrębie tych motywów strukturalnych.

Część motywów strukturalnych przewidzianych dla struktury *in vitro* oraz w lizacie została przewidziana w strukturze szczepu A/Vietnam/1203/2004 oraz w strukturze uniwersalnej vRNA8. Wspólne motywy strukturalne, wraz z różnicami wynikającymi z odmienności sekwencji, zostały przedstawione na **Ryc. 35**. Motyw w rejonie 261-288 nt wykazywał niewielkie różnice w parowaniu zasad, ale był zachowawczy dla wszystkich trzech analizowanych szczepów IAV (**Ryc. 35A**).

Motyw w rejonie 312-327 nt był zachowawczy pod względem sekwencji i struktury dla wszystkich trzech szczepów (**Ryc. 35B**). Jedna spinka formująca się w rejonie 797-814 nt została przewidziana w strukturze szczepów tego samego podtypu wirusa (H1N1) (**Ryc. 35C**). Zachowawczość powyższych motywów wskazuje na ich potencjalną funkcjonalność w czasie cyklu replikacyjnego wirusa, bądź w czasie pakowania wirionu potomnego. Pierwsze publikacje dotyczące sygnałów pakowania wskazywały, że są one ulokowane w odległości 50-150 nt na obu końcach cząsteczki vRNA²⁰⁰. Badania przeprowadzone przez *Dadonaite et al.*, wskazały, że sygnały pakowania mogą być ulokowane w dalszej odległości od końców cząsteczki¹¹⁵. To może sugerować, że zidentyfikowane motywy strukturalne mogą w istocie pełnić funkcje sygnałów pakowania.



Ryc. 35 Konserwatywne motywy strukturalne vRNA8 przewidziane dla różnych szczepów IAV w różnych warunkach. A. Motyw 262-288 nt wykazuje delikatne zmiany w parowaniu zasad wynikające z różnic na poziomie sekwencji. B. Motyw 312-327 nt jest wspólny dla wszystkich analizowanych szczepów i struktur. C. Motyw w rejonie 797-814 nt jest wspólny dla szczepów podtypu H1N1.

1.7.3. Występowanie motywów strukturalnych vRNA8 w lizacie komórkowym a profil wiązania białka NP

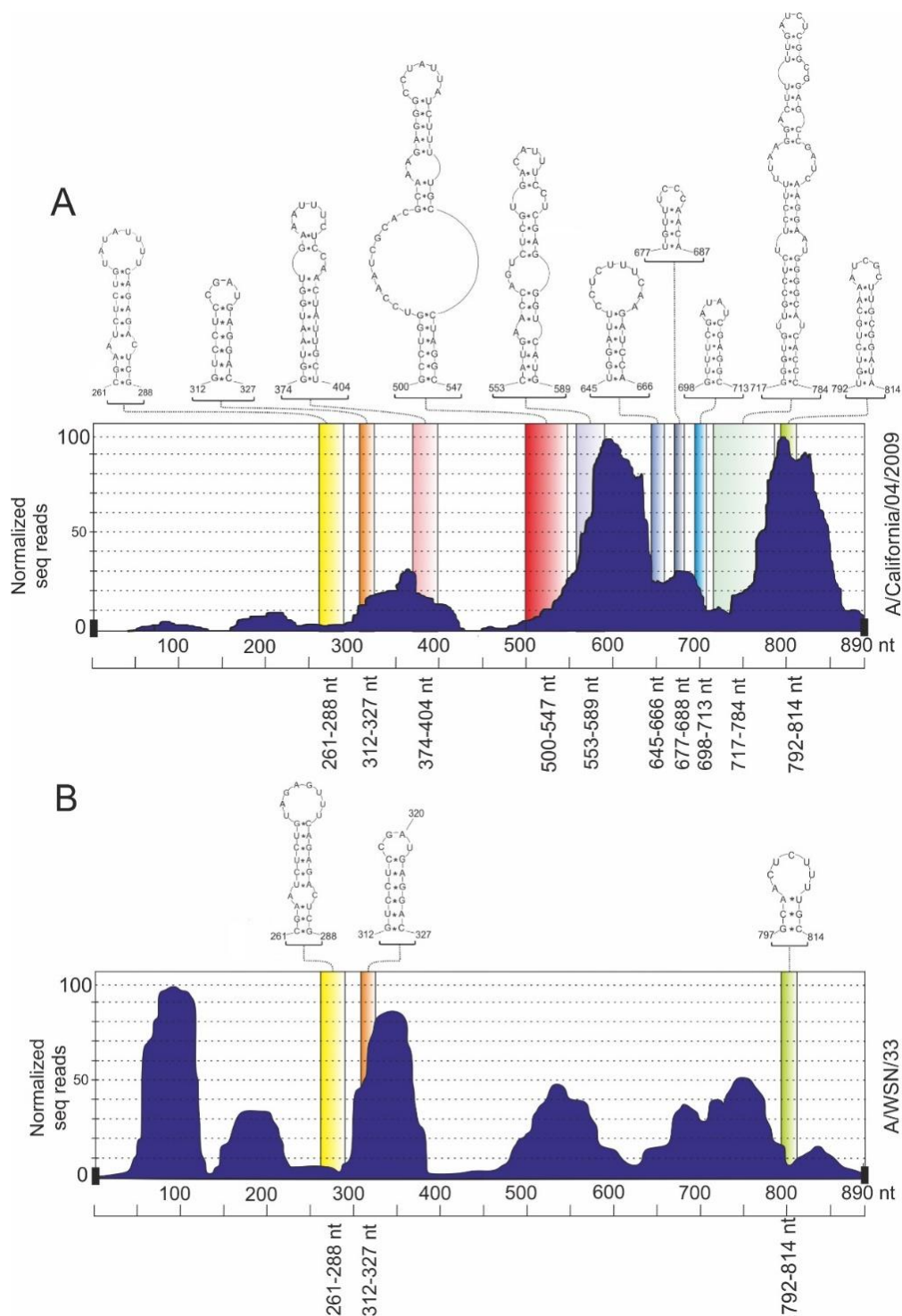
Oprócz podobieństw strukturalnych pomiędzy strukturami drugorzędowymi vRNA8 *in vitro* oraz w lizacie, zaobserwować można również wiele różnic (**Ryc. 31**). Może to być wynikiem różnych oddziaływań cząsteczki vRNA8 w lizacie

komórkowym. Co istotne, oprócz składników komórkowych, w mieszaninie lizatu obecne były również białka wirusowe, z którymi cząsteczka vRNA8 mogła swobodnie oddziaływać. Niedawno zostały opublikowane wyniki badań dotyczące profilu wiązania białka NP do cząsteczki vRNA w kompleksie vRNP dla dwóch szczepów IAV⁴⁶. Porównanie profilu wiązania białka NP dla dwóch odległych ewolucyjnie szczepów wirusa – A/California/04/2009 oraz A/WSN/1933 wykazało, że oprócz wielu podobieństw, jest on w dużej mierze unikalny dla szczepu. To sugeruje również pewne różnice na poziomie struktury drugorzędowej RNA między tymi szczepami.

W przypadku szczepu A/California/07/2009 najniższy poziom wiązania białka zidentyfikowano w rejonach 1-295 nt, 400-520 nt, 695-720 nt oraz 855-890 nt. Średni poziom wiązania białka zaobserwowano w rejonach 295-400 nt 520-555 nt, 630-695 nt oraz 720-775 nt. Rejony o niskim poziomie wiązania białka mogą być istotnym miejscem oddziaływań między różnymi segmentami vRNA. Z kolei najwyższy poziom wiązania białka zaobserwowano w rejonach 555-630 nt oraz 750-855 nt. Co ciekawe, motywy strukturalne przewidziane z bardzo wysokim prawdopodobieństwem ($\geq 90\%$) były w rejonach zarówno o niskim, jak i wysokim poziomie wiązania białka (Ryc. 20A). Motywy o wysokiej konserwatywności (261-288 nt, 312-327 nt oraz 500-547 nt) były przewidziane w miejscu słabo wiążanym przez białko NP. Z kolei dwa motywy (553-589 nt oraz 792-814 nt) przewidziane w miejscu bogatym w wiązania z białkiem NP miały zdecydowanie niższy poziom konserwatywności (77,8% i 81,78%).

W publikacji *Dadonaite et al.*¹¹⁵ również porównano występowanie motywów strukturalnych vRNA z profilem wiązania białka. Autorzy pracy doszli do konkluzji, iż występowanie motywów strukturalnych vRNA jest niezależne od profilu wiązania białka NP. Zaznaczyli, że fałdowanie się motywów strukturalnych w tych rejonach jest zależne od sekwencji RNA. Nie stwierdzili korelacji pomiędzy poziomem wiązania białka a ich występowaniem w strukturze. Nasze obserwacje również skłaniają się do tego wniosku, gdyż zaobserwowane motywy strukturalne posiadające wysokie prawdopodobieństwo parowania zasad ($>90\%$) występowały w obrębie cząsteczki niezależnie od poziomu wiązania, obecnego w lizacie komórkowym, białka NP. (Ryc. 36A). Co istotne, obecność wspólnych, dla obu analizowanych szczepów, motywów strukturalnych była niezależna od wiązania

białka NP. Na **Ryc. 36B** zaznaczono motywy szczepu A/WSN/1933, które zostały zidentyfikowane również w przypadku szczepu A/California/04/2009.



Ryc. 36 Profil wiązania białka NP szczepu **A.** A/California/04/2009 oraz **B.** A/WSN/1933. Na Ryc. A zaznaczono motywy strukturalne o wysokim prawdopodobieństwie par zasad w strukturze drugorzędowej vRNA8 w lizacie. Na Ryc. B zaznaczono motywy wspólne dla obu szczepów. Profil wiązania białka został określony za pomocą metody HITS-CLIP Lee et al. (NAR, 2017). Oś Y – poziomy wiązania białka NP znormalizowany wobec najwyższego pików. Oś X – numer nukleotydu w sekwencji vRNA8 od 5'-3' końca.

2. Wyniki odczytu mapowania chemicznego vRNA w warunkach *in virio* i *in cellulo* metodą profilowania mutacji MaP

Metoda odczytu mapowań chemicznych oparta o profil mutacji MaP wykorzystuje modyfikowaną reakcję odwrotnej transkrypcji. Reakcja odwrotnej transkrypcji przeprowadzana jest w obecności jonów Mn^{2+} , dzięki czemu w trakcie reakcji RT indukowane jest wprowadzanie mutacji lub delecji w miejscu komplementarnym do modyfikowanego chemicznie nukleotydu. Miejsca mutacji kumulują się w wydłużonej nici cDNA w miejscu komplementarnym do zmodyfikowanego nukleotydu w RNA. Analiza porównawcza profilu mutacji dla próby mapowanej odczynnikiem chemicznym i próby nietraktowanej odczynnikiem chemicznym umożliwia obliczenie reaktywności poszczególnych nukleotydów RNA. Szczegółowy opis metody, a także jej porównanie do metody odczytu za pomocą zatrzymań odwrotnej transkryptazy został zawarty w rozdziale *Wstęp literaturowy*.

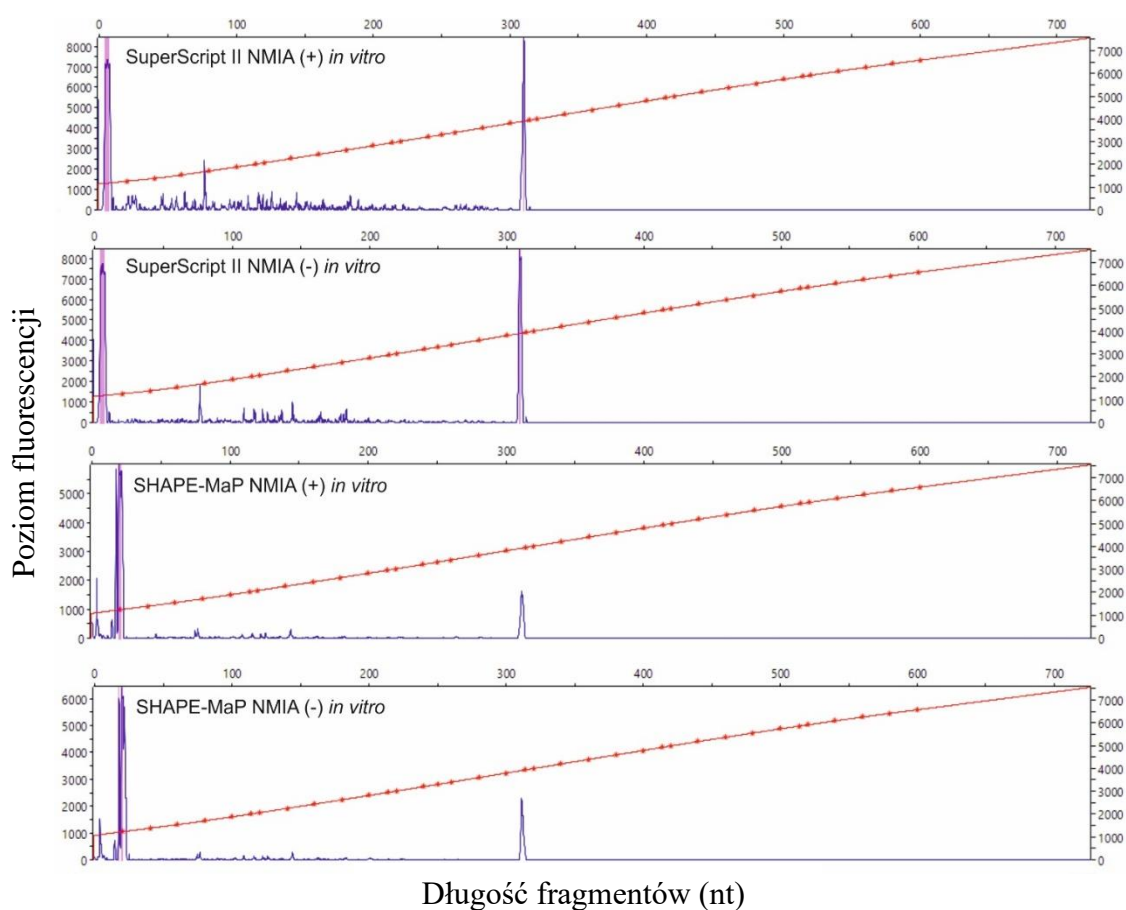
2.1. Optymalizacja metody odczytu profilu mutacji (MaP)

2.1.1. Reakcja na matrycy RNA in vitro

Przed przystąpieniem do eksperymentów w warunkach *in cellulo* oraz *in virio* wykonano optymalizację metody Map w warunkach *in vitro* oraz *in cellulo* zgodnie z procedurą opisaną w rozdziale *Materiały i metody*. Eksperymenty *in vitro* przeprowadzono na cząsteczce segmentu 8 vRNA szczepu A/California/04/2009. Po reakcji transkrypcji *in vitro* oraz potwierdzeniu integralności RNA przystąpiono do reakcji fałdowania RNA, a następnie mapowania z zastosowaniem odczynnika NMIA (reagent SHAPE). Po przeprowadzeniu reakcji z zastosowaniem enzymu SuperScript II, zarówno kontrolnej w warunkach optymalnych zgodnie z protokołem producenta (odczyt w wyniku zatrzymań RT), a także reakcji SHAPE-MaP, próby rozdzielono z zastosowaniem elektroforezy kapilarnej. Uzyskane wyniki były wizualizowane w programie PeakScanner (wersja 1.0) (**Ryc. 37**).

W przypadku reakcji odwrotnej transkrypcji w warunkach standardowych, na matrycy traktowanej odczynnikiem chemicznym, dochodzi do zatrzymań odwrotnej transkryptazy. Zatrzymania te kumulują się w nici cDNA na jeden nukleotyd przed modyfikowanym chemicznie nukleotydem w komplementarnej nici RNA (**Ryc. 37**).

Należy wziąć pod uwagę, że w przypadku próby kontrolnej (nietraktowanej odczynnikiem chemicznym) również dochodzi do naturalnych zatrzymań enzymu, min. w miejscach skomplikowanych pod względem struktury. Jednakże obserwuje się ich mniej niż w przypadku próbki mapowanej chemicznie. Tak więc, spodziewano się uzyskać zróżnicowany obraz rozdziału dla obu próbek, ze znaczną kumulacją zatrzymań enzymu w próbce traktowanej odczynnikiem chemicznym. Wyniki rozdziału prób po reakcji odwrotnej transkrypcji z zastosowaniem standardowego protokołu producenta potwierdziły kumulację zatrzymań odwrotnej transkryptazy w próbce reakcyjnej w porównaniu z próbką kontrolną (**Ryc. 37**). Tym samym potwierdzono właściwe warunki reakcji mapowania oraz reakcji RT.



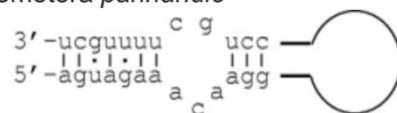
Ryc. 37 Wizualizacja wyników rozdziału elektroforezy kapilarnej w programie Peak Scanner (wersja 1.0). Nieprzereagowany starter widoczny jest na początku rozdziału.

W czasie reakcji RT metodą MaP nie dochodzi do kumulacji zatrzymań enzymu, a tym samym, w wyniku reakcji powstaje pełnej długości produkt cDNA. W tym przypadku spodziewamy się uzyskania pełnego produktu cDNA, zarówno dla próbki kontrolnej jak i reakcyjnej. Uzyskany rozdział kapilarny wskazał, że rozdział

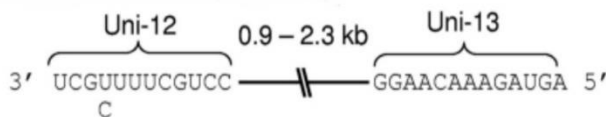
fragmentów obu próbek jest bardzo podobny, z minimalnie mniejszą ilością produktu końcowego (311 nt) w przypadku próbki reakcyjnej (**Ryc. 37**). Wyniki te wskazały na poprawność przeprowadzenia reakcji RT typu MaP. Warto zaznaczyć, że wyniki kontrolnej reakcji RT (odczyt zatrzymań RT) również wykazały powstanie pełnej długości produktu końcowego. Wyniki pokazały również, że reakcja odwrotnej transkrypcji typu MaP jest o wiele mniej wydajna, niż w przypadku standardowej reakcji odwrotnej transkrypcji. Mniejsza wydajność reakcji typu MaP została zasygnalizowana przez twórców metody^{149,143}.

Docelowy eksperyment w warunkach biologicznych wymaga doboru odpowiedniego odczynnika mapującego. Zostało to opisane w rozdziale *Wstęp literaturowy*. W planowanych eksperymentach *in virio* oraz *in cellulo* postanowiono wykorzystać odczynnik NAI (SHAPE) oraz DMS. Przygotowanie biblioteki do sekwencjonowania NGS metodą SHAPE-MaP oraz DMS-MaPseq narzuca konieczność otrzymania dwuniciowego produktu DNA^{143,140}. Tak więc, kolejnym etapem było wykonanie reakcji PCR na produkcie cDNA reakcji odwrotnej transkrypcji. Do testów *in vitro* przygotowano vRNA8 zmapowany odczynnikiem DMS oraz NAI. Następnie przeprowadzono kontrolną reakcję RT oraz reakcję RT

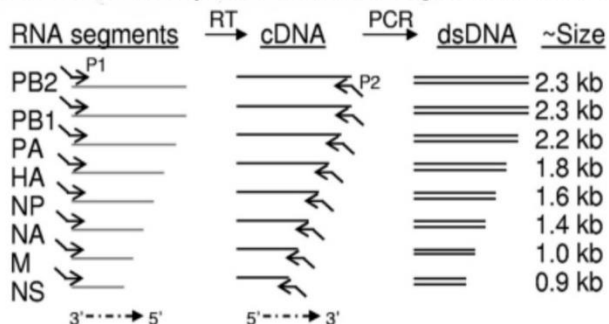
A. Struktura promotora *panhandle*



B. Genomowy RNA wirusowy (vRNA)



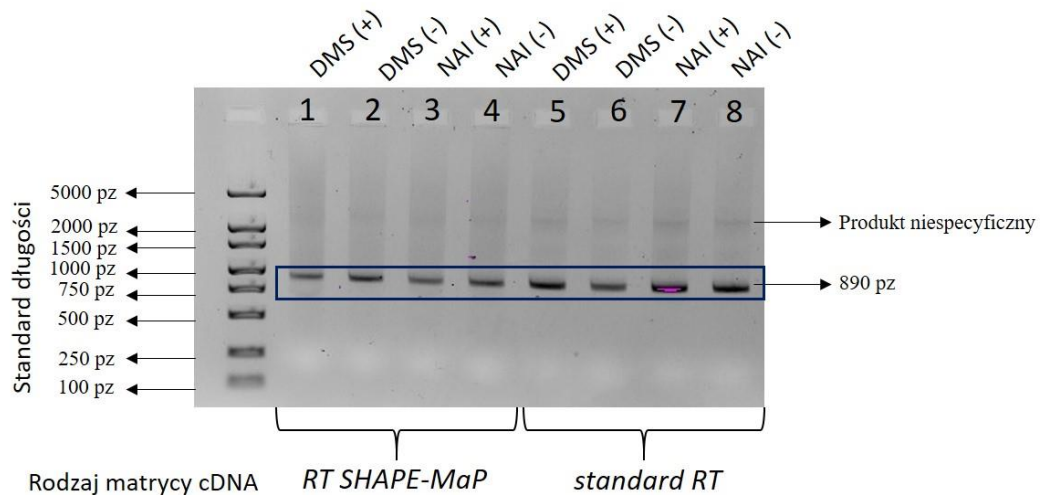
C. Schemat reakcji RT-PCR wielu segmentów vRNA



Ryc. 38 A. Struktura promotora *panhandle* B. Rejony docelowe dla uniwersalnych starterów komplementarnych do 3'końca (Uni-12) i 5' końca (Uni-13) cząsteczki vRNA C. Schemat reakcji RT-PCR, w której możliwa jest amplifikacja dla wszystkich segmentów vRNA. Rysunek na podstawie Zou et al., Journal of Virology, 2009.

typu SHAPE-MaP z zastosowaniem startera Uni-12 komplementarnego do 3' końca cząsteczki vRNA8. Co istotne, starter ten umożliwia przepisanie na cDNA wszystkich segmentów vRNA w jednej reakcji, ponieważ jest komplementarny do sekwencji 3' końca wszystkich segmentów vRNA w rejonie promotora (**Ryc. 38A**). Starter Uni-12 jest uniwersalny dla wielu szczepów wirusa grypy typu A (**Ryc. 38B**)^{201,202}. Zastosowanie sekwencji starterów Uni-12 oraz Uni-13 w RT umożliwia przeprowadzenie namnażania wszystkich segmentów w jednej, następczej, reakcji PCR (**Ryc. 38C**) Na tej podstawie została opracowana reakcja RT-PCR namnażania wszystkich segmentów w jednej reakcji PCR, wykorzystana w trakcie uzyskania danych z mapowań dla wszystkich segmentów vRNA. Reakcja PCR jest opisana w dalszej części pracy.

Produkt cDNA segmentu 8 posłużył jako matryca do reakcji PCR z zastosowaniem specyficznych starterów (NS-F i NS-R, Tabela 1.1, *Materiały i metody*). Produkty reakcji rozdzielono na 1% żelu agarozowym (**Ryc. 39**). Rozdział elektroforetyczny wskazał uzyskanie właściwych produktów dla wszystkich matryc cDNA (**Ryc. 39**).

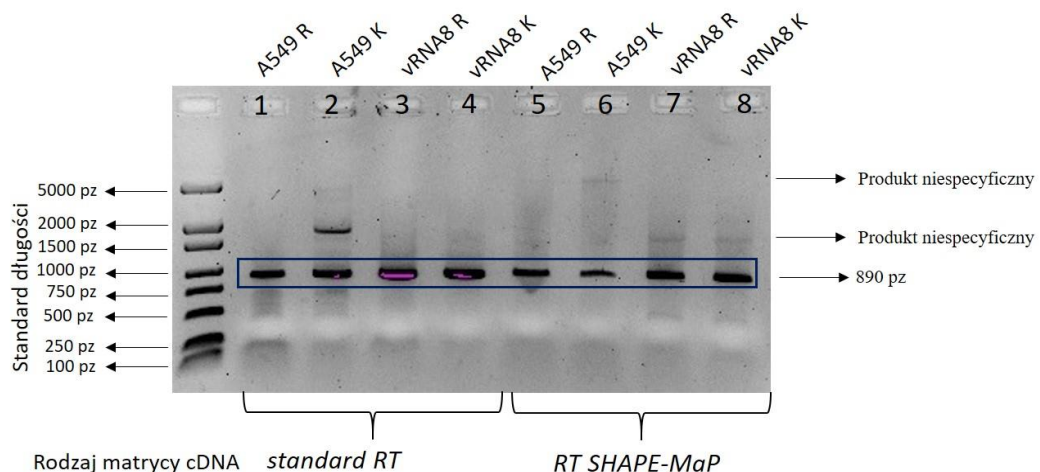


Ryc. 39 Rozdział produktów reakcji PCR dla próbek *in vitro*, z zastosowaniem starterów specyficznych do cDNA segmentu 8. Na każdą ścieżkę nałożono 100 ng produktu PCR. Ścieżki 1-4 – produkty PCR na matrycy cDNA po reakcji RT typu SHAPE-MaP. Ścieżki 5-8 – produkty PCR na matrycy cDNA z kontrolnej (standardowej) reakcji RT wg. protokołu producenta SuperScript II. Produkty PCR na matrycy cDNA po mapowaniu odczynnikami chemicznymi (DMS(+), NAI(+)) oraz ich reakcji kontrolnych (DMS(-), NAI(-)) mają właściwą długość (890 pz, zaznaczone w ramce). W przypadku reakcji PCR na matrycy cDNA z reakcji kontrolnej RT są widoczne produkty niespecyficzne (ścieżki 5-8). Jako standard długości użyto drabinkę Perfect Plus Molecular Weight Quantitative Ladder (Eurx).³⁸

2.1.2. Próbną reakcją typu SHAPE-MaP na matrycy RNA *in cellulo*

W przypadku reakcji *in vitro* istnieje możliwość wykorzystania dużej ilości czystego materiału RNA do reakcji. Reakcje w warunkach biologicznych utrudniają duże zagęszczenie RNA komórkowego przy stosunkowo niewielkiej ilości wirusowego RNA. Rozwiązaniem jest zastosowanie specyficznego startera do reakcji odwrotnej transkrypcji. Próbną reakcją RT typu SHAPE-MaP wykonano z zastosowaniem startera Uni-12. Jako matrycę wykorzystano całkowity RNA wyizolowany z zainfekowanych wirusem grypy komórek A549 mapowanych odczynnikami NAI. Matrycę *in cellulo* przygotowano zgodnie z procedurą opisaną w rozdziałach dotyczących infekcji, mapowania komórek A549, izolacji całkowitego RNA oraz oczyszczania RNA z zastosowaniem enzymu DNazy I (*Materiały i metody*). Jako kontrolę zastosowano cząsteczkę vRNA8 *in vitro*, a do reakcji RT również zastosowano starter Uni-12.

Przeprowadzono próbę standardowej reakcji RT oraz reakcji RT typu SHAPE-MaP. Produkty cDNA posłużyły jako matryce do reakcji PCR z zastosowaniem starterów specyficznych dla 8 segmentu (NS-F i NS-R, Tabela 1.1, *Materiały i metody*). Produkty PCR rozdzielono na 1% żelu agarozowym. Uzyskano właściwe produkty o długości 890 nt dla wszystkich reakcji. Oprócz tego

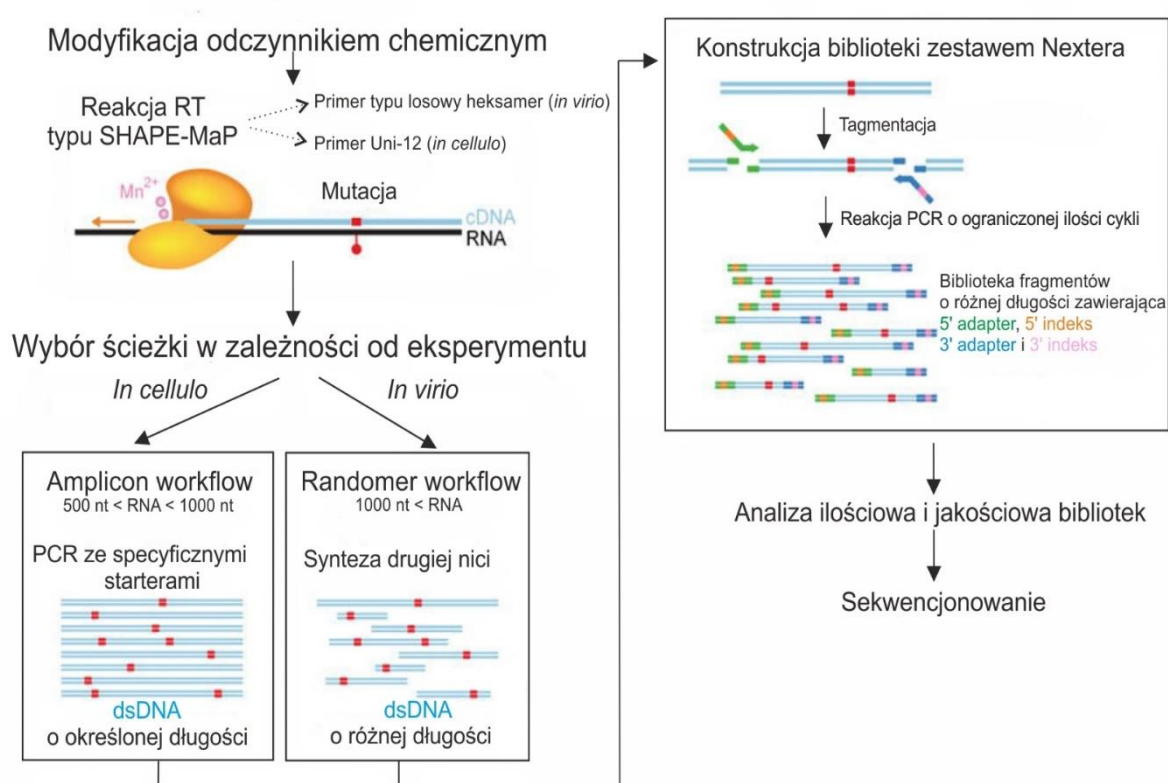


Ryc. 40 Rozdział produktów PCR dla próbek *in cellulo*. W reakcji zastosowano specyficzne startery dla segmentu 8. Ścieżki 1-4 produkty PCR na matrycy cDNA z kontrolnej reakcji RT. Ścieżki 5-8 – produkty PCR na matrycy cDNA z reakcji RT-SHAPE-MaP. Właściwy produkt reakcji o długości 890 pz zaznaczono w ramce. Na żelu widoczne są produkty niespecyficzne reakcji (ścieżka 1,2,6,7 i 8). Jako standard długości użyto drabinkę Perfect Plus Molecular Weight Quantitative Ladder (Eurx).

zaobserwowano produkty niespecyficzne na ścieżkach 2,6,7 i 8 (**Ryc. 40**). Uzyskane wyniki reakcji PCR na matrycy pochodzącej z materiału całkowitego RNA wykazały na zagęszczenie frakcji wirusowego RNA poprzez zastosowanie uniwersalnego startera Uni-12.

2.2. Wyniki mapowania struktury drugorzędowej wszystkich segmentów vRNA wirusa grypy szczepu A/California/04/2009 w warunkach *in virio* i *in cellulo*

W eksperymentach wykonanych w ramach mapowania struktury drugorzędowej wszystkich segmentów vRNA wirusa grypy szczepu A/California/04/2009 wykorzystano dwa protokoły opracowane w ramach metody SHAPE-MaP¹³¹. Na **Ryc. 41** zaprezentowany został schemat poszczególnych etapów przygotowania biblioteki fragmentów do sekwencjonowania metodą SHAPE-MaP. W ramach metody SHAPE-MaP etapy przygotowania biblioteki fragmentów (ścieżki) różnią się w zależności od pochodzenia matrycy RNA. W niniejszej pracy, w zależności od eksperymentu,



Ryc. 41 Schemat przygotowania biblioteki typu SHAPE-MaP. Poszczególne kroki eksperymentalne różnią się w zależności od pochodzenia matrycy. Ostatni krok konstrukcji biblioteki z zastosowaniem zestawu Nextera (Illumina) jest wspólny.

zastosowano dwie ścieżki: *Amplicon workflow (in cellulo)* oraz *Randomer workflow (in virio)*.

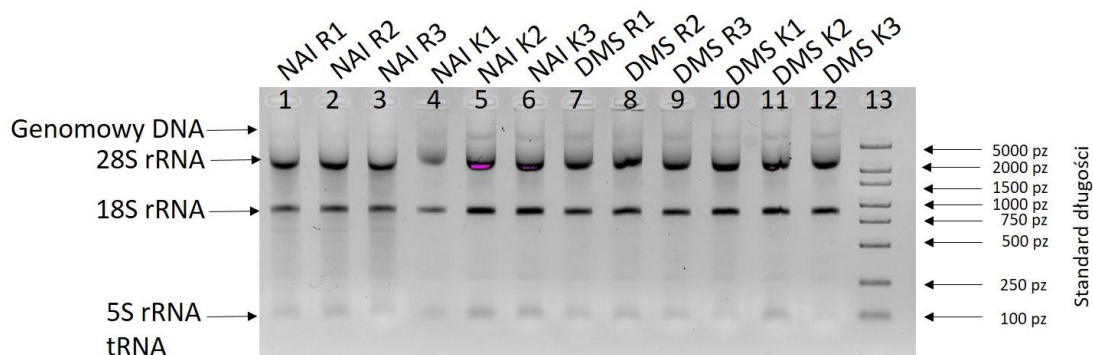
2.2.1. Przygotowanie matrycy RNA z eksperymentu mapowania chemicznego *in virio*

Jako matrycę do eksperymentów mapowania *in virio* wykorzystano zagęszczony roztwór wirusa grypy typu A szczepu A/California/04/2009. Roztwór wirusa pochodził z supernatantów zakażonych wirusem komórek MDCK. Po zagęszczeniu wirusów z supernatantów na poduszce sacharozowej wykonano mapowanie struktury drugorzędowej odczynnikami chemicznymi – DMS i NAI, a następnie izolację RNA. Eksperyment mapowania chemicznego wykonano w trzech niezależnych powtórzeniach. Szczegółowy opis eksperymentów zawarty jest w rozdziale *Materiały i metody*. Po uzyskaniu czystego izolatu RNA wirusowego mapowanego odczynnikami chemicznymi przystąpiono do przygotowania biblioteki do sekwencjonowania NGS.

2.2.2. Przygotowanie matrycy RNA z eksperymentu mapowania chemicznego *in cellulo* z zainfekowanych wirusem grypy komórek A549

Komórki A549 zostały zakażone wirusem A/California/04/2009 na 24 godziny. Czas ten odpowiada około 2 cyklom replikacyjnym wirusa. Po tym czasie wykonano eksperyment mapowania chemicznego z zastosowaniem dwóch odczynników chemicznych o potwierdzonej zdolności do przenikania przez błony komórkowe: DMS oraz NAI. Następnie przeprowadzono izolację całkowitego RNA. W celu potwierdzenia integralności RNA wykonano rozdział elektroforetyczny na żelu agarozowym (**Ryc. 42**). Widoczne na żelu wyraźne prążki 18S i 28S rRNA komórkowego świadczą o dobrej jakości wyizolowanego RNA. W próbkach pochodzących z mapowania odczynnikami NAI (**Ryc. 42**; ścieżki 1-4) widoczne jest lekkie smużenie RNA na żelu, które jest spowodowane zasoleniem próbki RNA. Na żelu widoczny jest również prążek na wysokości powyżej 5000 pz, który stanowi frakcję genomowego DNA. W celu oczyszczenia całkowitego RNA z frakcji DNA genomowego wykonano reakcję enzymatycznej degradacji DNA enzymem DNazą I. Oczyszczenie RNA po reakcji enzymatycznej umożliwiło również pozbycie się

resztek soli pozostałych po izolacji RNA. Czystość RNA potwierdzono poprzez pomiar na spektrofotometrze NanoDrop.



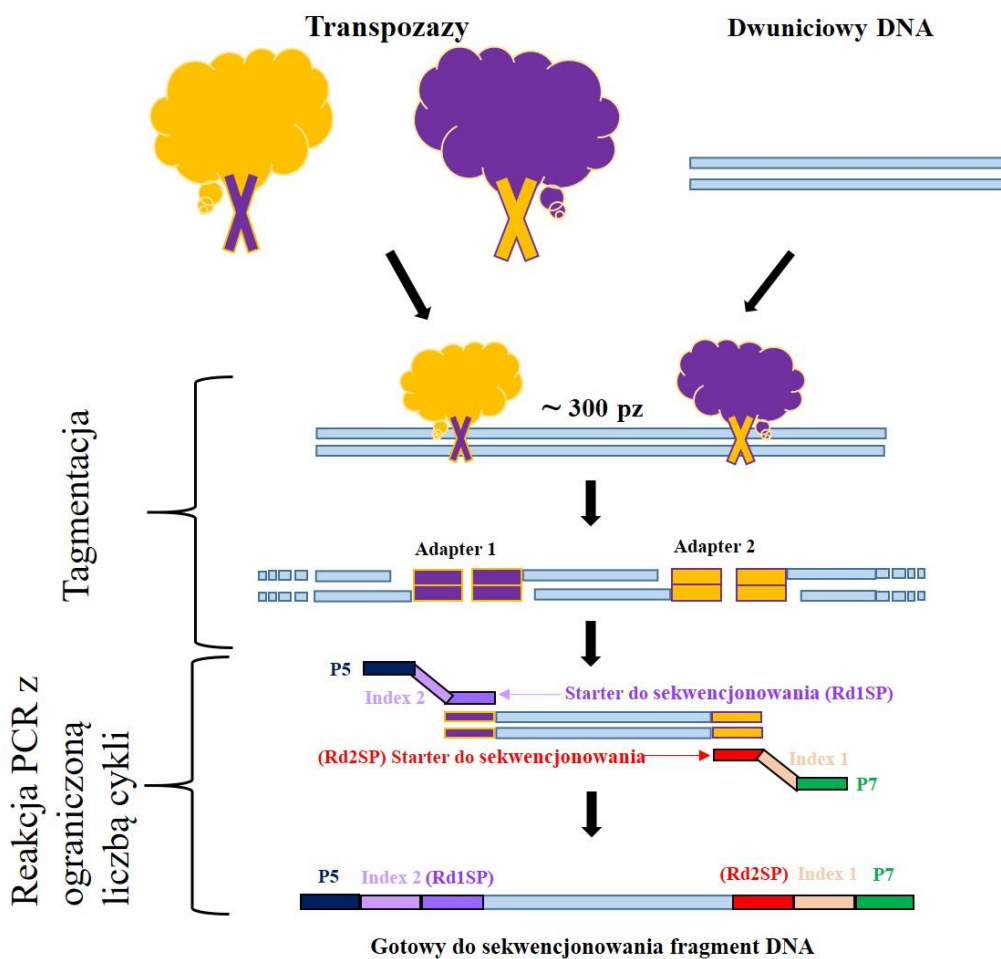
Ryc. 42 Rozdział elektroforetyczny całkowitego RNA wyizolowanego z komórek A549 na 1% żelu agarozowym. Na każdą ścieżkę nałożono 500 ng RNA. Ścieżki 1-6 RNA wyizolowany z komórek po mapowaniu odczynnikiem chemicznym NAI, Ścieżki 7-12 – RNA wyizolowany z komórek po mapowaniu odczynnikami DMS. R1, R2, R3 – próbki po reakcji z dodaniem odpowiedniego odczynnika chemicznego z trzech powtórzeń biologicznych; K1, K2, K2 – próbki kontrolne nietraktowane odczynnikami chemicznymi z trzech powtórzeń biologicznych. Ścieżka 13 – standard długości MWQ (EurX). Na ścieżkach 1-12 widoczna jest frakcja krótkich RNA komórkowych tj. 5S rRNA oraz tRNA, a także długich rRNA – 18S i 28S. Zaznaczono również widoczne na żelu zanieczyszczenie genomowym DNA.

2.2.3. Biblioteki *in virio*: Przygotowanie matrycy, reakcja tagmentacji i amplifikacja bibliotek fragmentów do sekwencjonowania

Po izolacji wirusowego RNA zmierzono stężenie za pomocą fluorymetru Qubit. W zależności od próbki, uzyskano ilość RNA w przedziale między 50 a 90 ng. Materiał RNA posłużył jako matryca do reakcji odwrotnej transkrypcji. W następnej kolejności matrycę cDNA poddano reakcji syntezy drugiej nici, której produkt posłużył do dalszych reakcji. Poniżej zostały opisane kolejne kroki przygotowania biblioteki.

Protokół przygotowania biblioteki opracowano na podstawie opublikowanej metody zgodnie ze ścieżką "randomer workflow" dla matryc powyżej 1000 nt (**Ryc. 41**)¹³¹. Całość oczyszczonego materiału RNA przeznaczono do reakcji RT typu MaP. Matryca RNA pochodząca z zagęszczonego roztworu wirusa nie była zanieczyszczona RNA komórkowym, dlatego też możliwe było zastosowanie startera typu losowy heksamer, z pomocą którego w czasie reakcji przepisana została całość RNA w próbce (*Materiały i metody*). Następnie przeprowadzono reakcję syntezy drugiej nici DNA (*Materiały i metody*). Stężenie oczyszczonych produktów syntezy drugiej nici

zmierzono za pomocą fluorymetru i wykonano rozcieńczenia matrycy DNA do stężeń 0,2 ng/μl. Do konstrukcji biblioteki pobrano 1 ng matrycy. Bibliotekę fragmentów DNA do sekwencjonowania przygotowano za pomocą zestawu Nextera XT DNA (Illumina) zgodnie z protokołem producenta. Pierwszym etapem konstrukcji biblioteki była reakcja tagmentacji (**Ryc. 43**). Reakcja tagmentacji wykorzystuje modyfikowany kompleks enzymów - transpozazy, który przeprowadza cięcie cząsteczki DNA z jednoczesnym dołączeniem uniwersalnych adapterów (ang. Tag) na obu końcach dwuniciowego DNA. Powodzenie reakcji tagmentacji zależy od kilku czynników. Ważne jest dokładne określenie stężenia próbki, ponieważ użycie ilości mniejszej niż



Ryc. 43 Schemat przygotowania biblioteki do sekwencjonowania wysokoprzepustowego z zastosowaniem zestawu Nextera XT DNA Kit (Illumina). W czasie reakcji tagmentacji matryca wejściowa DNA jest w pierwszym etapie poddana fragmentacji oraz dodaniu sekwencji adapterowych. Następnie, przeprowadzana jest reakcja PCR z ograniczoną ilością cykli, w której do sekwencji adapterowych przyłączane są startery posiadające unikalne sekwencje indeksujące (Indeks 1; Indeks 2) oraz adaptory komplementarne do oligonukleotydowych sond kowalencyjnie związanych na płycie sekwencyjnej (P5 i P7). Schemat przygotowany na podstawie protokołu Illumina.

1 ng matrycy prowadzi do zwiększonej tagmentacji, a tym samym fragmentacji, próbki. Z kolei użycie większej ilości matrycy prowadzi do zmniejszonej efektywności reakcji. Istotna jest również czystość próbki. Matryca DNA musi być wolna od zanieczyszczeń, które mogą powodować inhibicję reakcji. Do inhibicji reakcji przyczyniają się chelatory (np. EDTA), sole, białka związane z matrycowym DNA, proteiny, detergenty czy fenol²⁰³. Co więcej, aktywność enzymatyczna transpozaz jest również uzależniona od pH próbki²⁰³. Oprócz czystości próbki znaczenie ma również długość matrycy. Enzym wymaga matrycy o długości co najmniej 300 pz. Należy również pamiętać, że enzym nie ma możliwości przyłączenia adapterów bezpośrednio na końcach cząsteczki. W związku z tym dochodzi do częściowego utracenia danych na końcach cząsteczki. W trakcie planowania eksperymentów zadbano o spełnienie wszystkich wymienionych warunków.

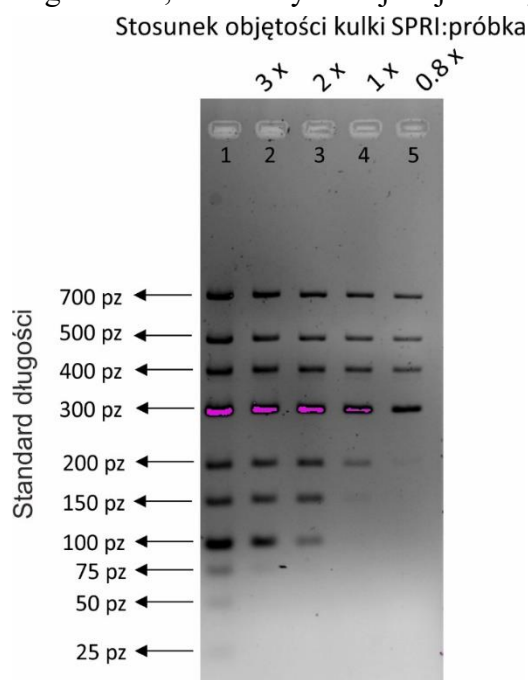
Po neutralizacji reakcji tagmentacji przystąpiono do ostatniego etapu konstrukcji biblioteki – reakcji PCR. Podczas reakcji PCR wykorzystywane są startery posiadające unikalną sekwencję indeksującą próbkę, sekwencje adapterowe komplementarne do startera do sekwencjonowania oraz sekwencje adapterowe komplementarne do oligomerów znajdujących się na płytce do sekwencjonowania (**Ryc. 14-16**, *Wstęp literaturowy*). Do każdej próbki dodano unikalne kombinacje starterów forward (i5) oraz reverse (i7). Kombinacje starterów użytych dla poszczególnych próbek zostały wyszczególnione w Tabeli 1 w rozdziale *Dodatek*.

Reakcję PCR przeprowadzono w 12 cyklach. W czasie przygotowania biblioteki ilość cykli w czasie reakcji PCR musi być znacznie mniejsza niż w przypadku standardowej reakcji amplifikacji. Po pierwsze, do sekwencjonowania wymagana jest stosunkowo niewielka ilość materiału DNA. Po drugie, zwiększenie ilości cykli może wpłynąć na generowanie większej ilości błędów przez polimerazę, a także może dojść do produkcji dużej ilości produktów niespecyficznych, takich jak dimery starterów (ang. primer-dimer). Powstawanie krótkich, niespecyficznych produktów jest szczególnie niekorzystne, ponieważ nawet niewielka ilość tych produktów w bibliotece fragmentów powoduje ich preferencyjne namnażanie na etapie generacji klastrów na płytce do sekwencjonowania (**Ryc. 14**, *Wstęp literaturowy*). Przekłada się to na ilość użytecznych odczytów sekwencjonowania, znacznie zmniejszając pokrycie docelowych cząsteczek.

Oprócz ograniczenia ilości cykli PCR, istotne jest wydajne oczyszczenie biblioteki fragmentów. W trakcie oczyszczania mieszaniny fragmentów można oczyścić próbkę z frakcji zarówno krótkich jak i długich fragmentów. Usunięcie długich fragmentów ma również istotne znaczenie w przypadku sekwencjonowania, ponieważ zbyt długie fragmenty DNA mają tendencję do „przeskakiwania” pomiędzy klastrami generowanymi na płycie do sekwencjonowania co przekłada się na obniżenie jakości sekwencjonowania. Oczyszczanie biblioteki do sekwencjonowania polega na usunięciu produktów niespecyficznego reakcji, a także selekcji fragmentów ze względu na ich długość (ang. size selection). Selekcję fragmentów można wykonać na dwa sposoby. Pierwszy z nich to oczyszczanie produktu na poliakrylamidowym żelu preparatywnym. Jednakże, metoda ta jest niewydajna i prowadzi do dużych strat oczyszczanego materiału. Drugą metodą jest wykorzystanie kulek magnetycznych typu SPRI. Metoda ta została zastosowana w trakcie eksperymentów opisanych w niniejszej pracy.

Kulki magnetyczne typu SPRI umożliwiają selektywne oczyszczanie mieszaniny fragmentów kwasów nukleinowych o różnej długości poprzez usunięcie krótkich i długich fragmentów. Efekt ten wywołany jest poprzez zastosowanie różnej objętości kulek magnetycznych w stosunku do objętości mieszaniny. Kulki magnetyczne SPRI zawieszane są w roztworze NaCl/PEG, którego składniki stwarzają warunki zagęszczenia. Modulowanie zagęszczenia mieszaniny stosunkiem objętości mieszaniny kulek w odniesieniu do próby sprawia, że im większa objętość kulek tym dłuższe fragmenty przyciągane są do powierzchni kulek, np. stosunek kulek wynoszący 1,8 x w stosunku do próby sprawia, że z próby oczyszczane są fragmenty poniżej 100 nt, z kolei stosunek 0,9 x powoduje odcięcie produktów poniżej 300 nt. Oczyszczanie na kulkach SPRI zachodzi w trzech etapach: 1) Zmieszanie próbki z wybraną objętością kulek, inkubacja i związanie kulek do magnezu na statywie magnetycznym, 2) Usunięcie supernatantu i przemycie kulek etanolem o odpowiednim stężeniu, 3) Elucja oczyszczonych fragmentów poprzez zmieszanie kulek z wodą, inkubacja, związanie kulek do magnezu na statywie magnetycznym i zebranie oczyszczonego produktu rozpuszczonego w wodzie. Przed oczyszczeniem bibliotek fragmentów wykonano próbne oczyszczanie na kulkach Ampure XP beads (Beckman Coulter) z zastosowaniem różnych objętości kulek w stosunku do

oczyszczanej próbki (**Ryc. 44**) zgodnie z opisem zawartym w rozdziale *Materiały i metody*. Testowe oczyszczanie wykonano na standardzie o fragmentach długości w zakresie od 25 do 700 pz. Optymalizacja oczyszczania mieszaniny fragmentów o różnych długościach z zastosowaniem kulek SPRI wykazała bardzo precyzyjne oczyszczenie próbki. Użycie różnej ilości kulek w stosunku do próbki spowodowało usunięcie z mieszaniny fragmentów w różnym zakresie. Zgodnie z przewidywaniami, użycie mniejszej objętości kulek skutkowało usunięciem z mieszaniny dłuższych fragmentów, bez utraty frakcji najdłuższych fragmentów.



Ryc. 44 Rozdział elektroforetyczny na 3% żelu agarozowym próbek po oczyszczeniu na kulkach SPRI. Ścieżka 1 – nieoczyszczony standard długości GeneRuler Low Range DNA Ladder (ThermoFisher). Ścieżki 2-5 – rozdział standardu długości po oczyszczeniu z różną objętością kulek SPRI (Ampure XP beads, Beckman Coulter).

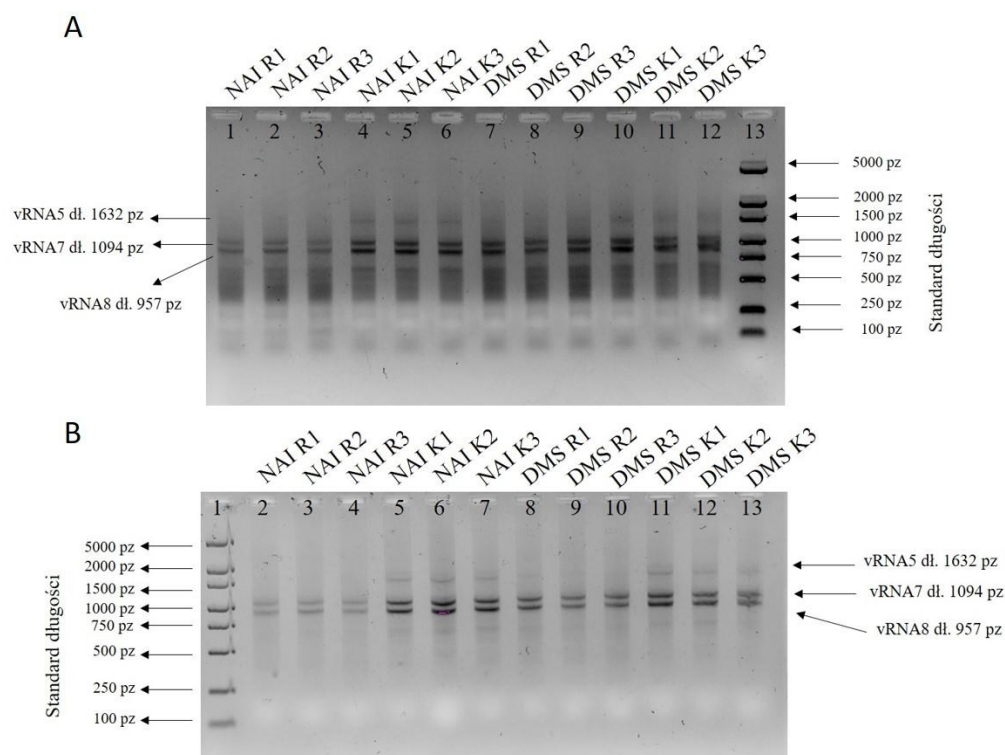
W przypadku eksperymentu mapowania vRNA w wirionie wirusa grypy długości fragmentów były w zakresie od 890 pz (8 segment) do 2341 pz (1 segment). Zgodnie z protokołem producenta zestawu Illumina XT DNA dla matryc DNA o początkowej długości powyżej 500 pz należy zastosować 0,6 x objętości kulek SPRI. Użycie stosunku objętości 0,6 x skutkowało usunięciem fragmentów o długościach poniżej 300 pz. Stężenie uzyskanych bibliotek do sekwencjonowania było niewielkie i oscylowało w granicach ok 20-100 ng dla całości próbki. Dlatego też analiza ilościowa i jakościowa próbek wymagała zastosowania instrumentów umożliwiających ocenę jakościową produktów DNA o niskim stężeniu.

2.2.4. Biblioteki *in cellulo*: Przygotowanie matrycy, reakcja tagmentacji i amplifikacja bibliotek fragmentów do sekwencjonowania

Protokół przygotowania biblioteki opracowano na podstawie opublikowanej metody zgodnie ze ścieżką "Amplicon workflow" (**Ryc. 41**)¹³¹. Do reakcji RT typu MaP pobrano 3 µg całkowitego RNA. Matryca całkowitego RNA składa się w ok. 99% z RNA komórkowego. Natomiast frakcja wirusowego RNA stanowiła mniej niż 1% w mieszaninie RNA, co zostało potwierdzone przez nas w reakcji real-time PCR. W związku z tym, do reakcji RT wykorzystano starter Uni-12 o sekwencji komplementarnej do 3' końca rejonu *panhandle* wszystkich segmentów vRNA (**Ryc. 38**)²⁰². W celu uzyskania dwuniciowego DNA wykonano reakcję amplifikacji PCR z zastosowaniem starterów HFA i HRA komplementarnych do 3' i 5' końca cząsteczki cDNA (rejon *panhandle*). Startery te posiadają sekwencję starterów Uni-12 oraz Uni-13 (**Ryc. 38B**) wydłużonych o sekwencję dla tagmentazy, dzięki czemu enzym ma możliwość przyłączenia się na początku badanej sekwencji. Miało to na celu umożliwienie odczytu obu końców cząsteczki w trakcie sekwencjonowania. Dokładny opis wszystkich reakcji został zawarty w rozdziale *Materiały i metody*. Po zmierzeniu stężeń produktów PCR za pomocą fluorymetru Qubit, sprawdzono ich jakość za pomocą rozdziału elektroforetycznego na 1% żelu agarozowym (**Ryc. 45A,B**). W wyniku reakcji spodziewano się uzyskać 8 produktów reakcji PCR, odpowiadających ośmiu segmentom vRNA. Jednakże, rozdziel elektroforetyczny (**Ryc. 45A,B**) wykazał obecność jedynie trzech produktów specyficznych, których długości odpowiadały cząsteczkom vRNA5, vRNA7 i vRNA8 wydłużonym o sekwencję starterów. Na żelu zaobserwowano również dużą frakcję produktów niespecyficznych widocznych jako smuga poniżej długości 750 pz (**Ryc. 45A,B**).

W celu usunięcia frakcji niespecyficznych produktów reakcji amplifikacji wykonano oczyszczanie z zastosowaniem kulek SPRI w stosunku objętości wynoszącym 0,5 x. Oczyszczone produkty PCR ponownie rozdzielono z zastosowaniem elektroforezy na 1% żelu agarozowym (**Ryc. 45B**). Rozdział oczyszczonych produktów PCR wskazał na powodzenie takiego oczyszczania. Brak widocznych produktów reakcji dla segmentów vRNA1, vRNA2, vRNA3, vRNA4 oraz vRNA6 mógł wynikać z ich mniejszego udziału procentowego w produkcie reakcji RT. Drugim powodem mogło być preferencyjne namnażanie segmentów 5, 7 i

8 w komórce, wynikające ze specyfiki cyklu replikacyjnego wirusa. Niedawno opublikowana praca z zakresu kinetyki namnażania RNA wirusa grypy wskazała, że w istocie jest ona zróżnicowana dla każdego typu cząsteczki (vRNA, cRNA i mRNA), a także inna dla poszczególnych segmentów RNA²⁰⁴.



Ryc. 45 Rozdział elektroforetyczny na 1% żelu agarozowym produktów reakcji PCR z zastosowaniem uniwersalnych starterów HFA i HRA. **A.** Rozdział produktów PCR przed oczyszczaniem. **B.** Rozdział produktów PCR po oczyszczaniu na kulkach SPRI. Na każdą ścieżkę nałożono 200 ng. Na ścieżkach 1-6 – produkty PCR na matrycy z eksperymentu mapowania NAI; Ścieżki 7-12 – produkty PCR na matrycy z eksperymentu mapowania DMS. R1, R2, R3 – matryca z trzech powtórzeń biologicznych reakcji mapowania dla odpowiedniego odczynnika chemicznego; K1, K2, K3 – matryca z trzech powtórzeń biologicznych dla reakcji kontrolnej, nietraktowanej odczynnikiem chemicznym. Ścieżka 13 – standard długości MWQ (EurX). Na żelu widoczne są trzy produkty reakcji PCR odpowiadające długościom segmentów vRNA5, vRNA 7 i vRNA8 wydłużonych o sekwencje starterów.

Uzyskany produkt PCR posłużył jako matryca do przygotowania biblioteki do sekwencjonowania NGS. Kolejne etapy przygotowania biblioteki, za wyjątkiem etapu oczyszczania z użyciem kulek SPRI, zostały wykonane tak samo jak w przypadku bibliotek z eksperymentu *in virio*. W przypadku eksperymentu *in virio* zastosowanie losowego startera skutkowało uzyskaniem produktów cDNA, a następnie DNA, o różnej długości. W przypadku bibliotek *in cellulo* matrycę do biblioteki stanowiły

długie produkty PCR, których fragmentacja na etapie tagmentacji mogła zachodzić nieefektywnie. Z tego względu zastosowano oczyszczanie produktów biblioteki poprzez selekcję fragmentów, polegającą na usunięciu krótkich oraz długich fragmentów w mieszaninie (*Materiały i metody*).

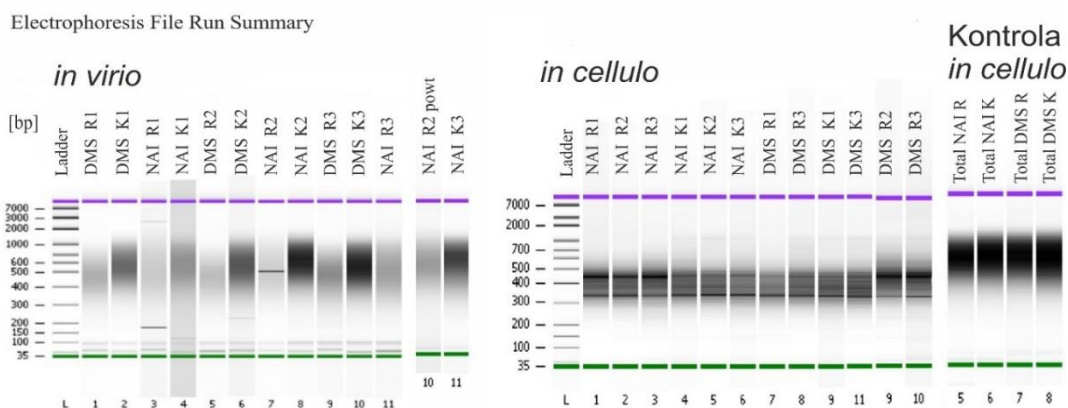
2.2.5. Przygotowanie biblioteki do sekwencjonowania NGS dla próby kontrolnej, zawierającej RNA o znanej strukturze drugorzędowej (kontrola *in cellulo*)

W celu sprawdzenia prawidłowości reakcji mapowania RNA, pod kątem późniejszego odczytu reaktywności nukleotydów badanych struktur, przygotowano próbę kontrolną o znanej strukturze drugorzędowej – wybrano 18S rRNA. Wykorzystano całkowity RNA komórkowy po reakcji mapowania chemicznego odczynnikami NAI oraz DMS. Do przygotowania biblioteki do sekwencjonowania wykorzystano protokół opracowany dla próbek *in virio*. Tak więc, w pierwszej kolejności przepisano 1 µg całkowitego RNA na cDNA w reakcji odwrotnej transkrypcji typu MaP z wykorzystaniem startera typu losowy heksamer. Następnie wykonano reakcję syntezy drugiej nici i uzyskany produkt rozcieńczono. Rozcieńczony produkt syntezy drugiej nici posłużył jako matryca do przygotowania biblioteki. Gotowe biblioteki oczyszczano z zastosowaniem kulek SPRI w stosunku objętości wynoszącym 0,6 x.

2.2.6. Biblioteki do sekwencjonowania NGS- analiza ilościowa i jakościowa bibliotek

Gotowe biblioteki do sekwencjonowania zostały sprawdzone pod względem ilościowym i jakościowym. Stężenie próbek zmierzono za pomocą fluorymetru Qubit. Biblioteki fragmentów zostały następnie sprawdzone pod względem jakościowym z zastosowaniem bioanalyzera (Agilent Bioanalyzer 2100) (*Materiały i metody*). Bioanalyzer jest urządzeniem do przeprowadzenia elektroforezy kapilarnej wysokiej rozdzielczości w technologii typu lab-on-a-chip. Wyniki elektroforezy umożliwiają ocenę jakościową próbki tj. ocenę rozkładu fragmentów w próbce wskazując ewentualne zanieczyszczenie próbki produktami ubocznymi reakcji PCR, na przykład primer-dimer. Sekwencjonowanie NGS wymaga znormalizowanego rozkładu fragmentów, w zależności od wybranej platformy sekwencyjnej odpowiedni zakres długości mieści się w wartościach pomiędzy 300-1500 nt. Na **Ryc. 46** zaprezentowany

został obraz rozdziału elektroforetycznego gotowych bibliotek z eksperymentu mapowania *in virio*, *in cellulo* oraz kontroli *in cellulo*. Właściwy rozdział fragmentów powinien się znajdować w zakresie od 300 do 1000 pz. Fragmenty ≥ 150 pz są produktami ubocznymi reakcji PCR typu primer-dimer lub niezwiązanymi starterami.

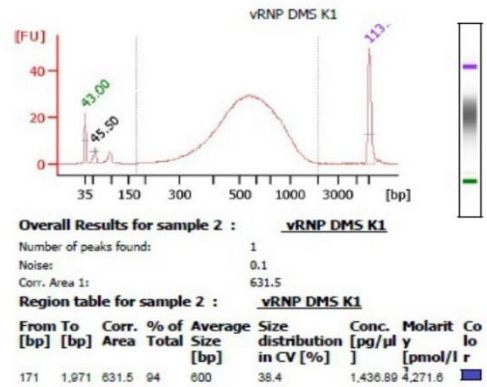
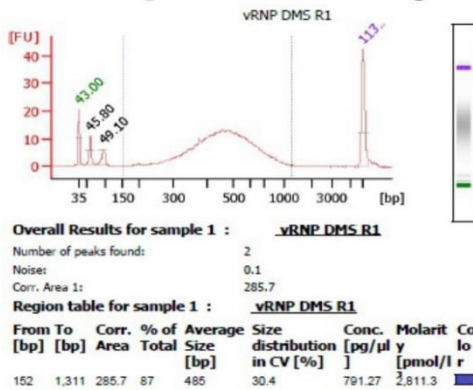


Ryc. 46 Obraz rozdziału elektroforetycznego gotowych bibliotek do sekwencjonowania *in virio* oraz *in cellulo* oraz kontroli *in cellulo*. Rozdział wykonano z zastosowaniem chipu High Sensitivity. Po lewej stronie widoczny jest rozdział standardu długości w zakresie 35-7000 pz. Do każdej próbki dodawane są markery długości – krótki o dł. 35 pz (lower marker) oraz długi o dł. 10380 pz (upper marker). Odpowiednie przypisanie pików do markerów długości gwarantuje prawidłowy odczyt długości fragmentów w próbce. Na każdej ścieżce widoczne są prawidłowo przypisane markery. Na chipie High Sensitivity można jednorazowo rozdzielić maksymalnie 11 próbek, dlatego próbka NAI K3 została rozdzielona na osobnym chipie. Ze względu na pojawienie się podczas rozdziału artefaktu, dodatkowo powtórzono rozdzielanie próbki NAI R2.

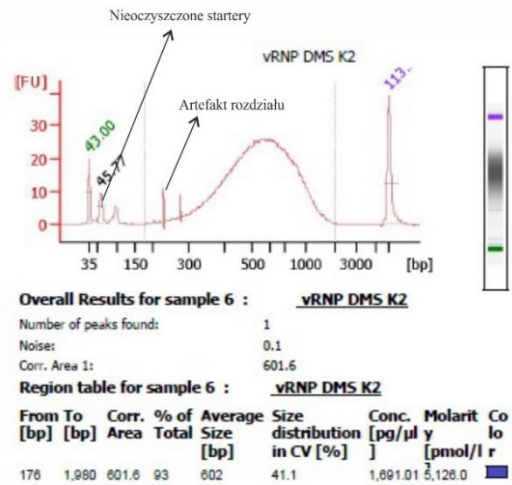
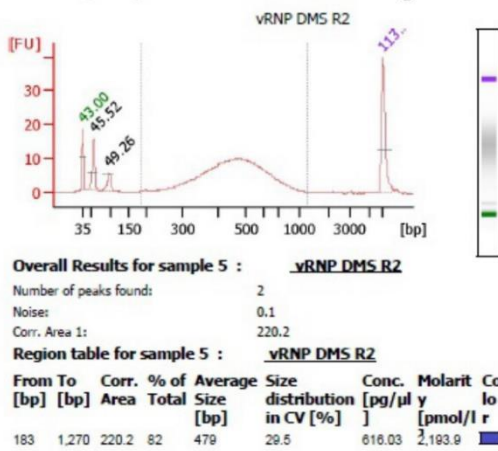
Poniżej przedstawione są wykresy rozdziału elektroforetycznego badanych próbek. Wyznaczono średnią długość fragmentów w próbce wraz ze stężeniem molarnym próbki. Było to niezbędne do prawidłowego obliczenia stężeń próbek w nM na etapie łączenia bibliotek fragmentów do sekwencjonowania (etap normalizacji bibliotek). Na **Ryc. 47-50** przedstawione są wykresy dla gotowych bibliotek. Wykresy wskazują na uzyskanie prawidłowych bibliotek do sekwencjonowania NGS dla próbek z eksperymentu *in virio*, których zakresy długości mieszczą się w zakresie 300-1000 pz. Obraz z bioanalyzera świadczy o prawidłowej fragmentacji biblioteki fragmentów. Oprócz markerów górnego i dolnego widoczne są piki świadczące o zanieczyszczeniu starterami. Zdecydowano, że niewielka ilość zanieczyszczenia starterami nie wpłynie znacznie na jakość sekwencjonowania. W przypadku wykresów dla próbek *in virio* (vRNP DMS K2, NAI R1 oraz NAI K1) widoczne są artefakty rozdziału elektroforetycznego. Wykresy wskazują na większe stężenie próbek kontrolnych (vRNP DMS K1, K2 i K3 oraz NAI K1, K2 i K3). Wynika to z

faktu, że na żel nałożono 1 μ l nierozcieńczonej próbki gotowej biblioteki, a stężenie próbek kontrolnych były średnio 2x wyższe niż dla próbek reakcyjnych (pomiar Qubit).

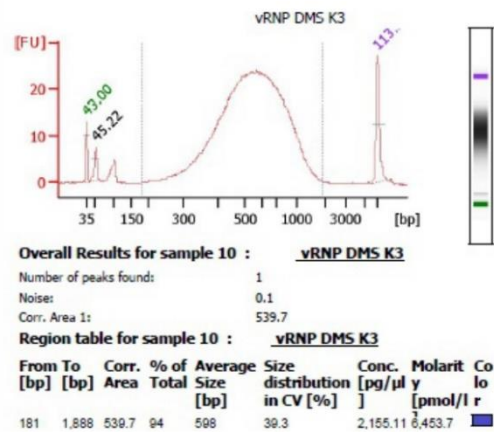
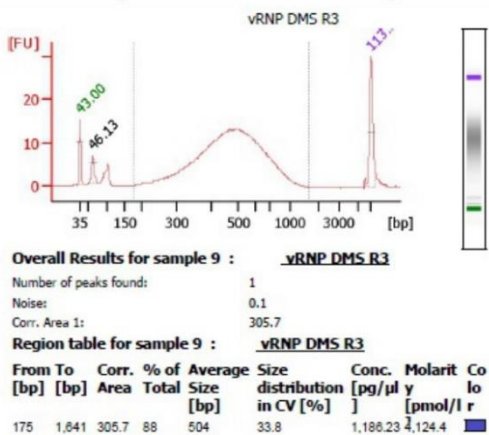
Pierwsze powtórzenie biologiczne



Drugie powtórzenie biologiczne

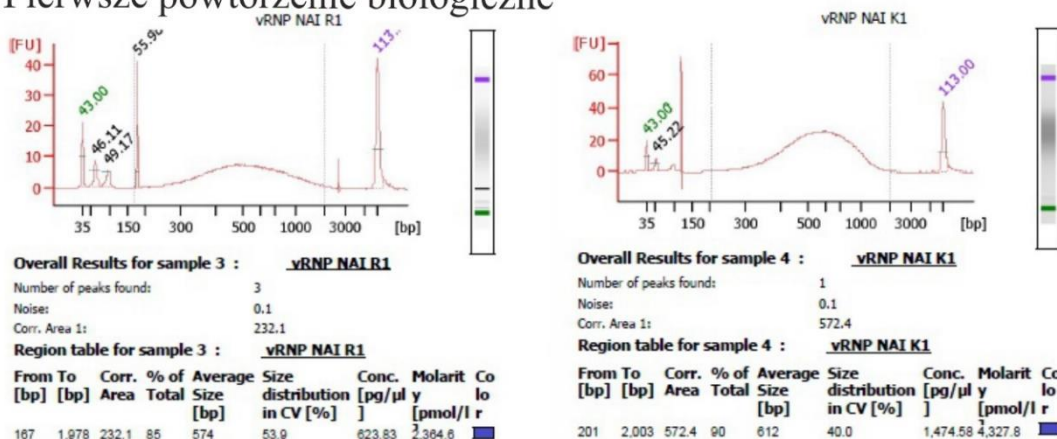


Trzecie powtórzenie biologiczne

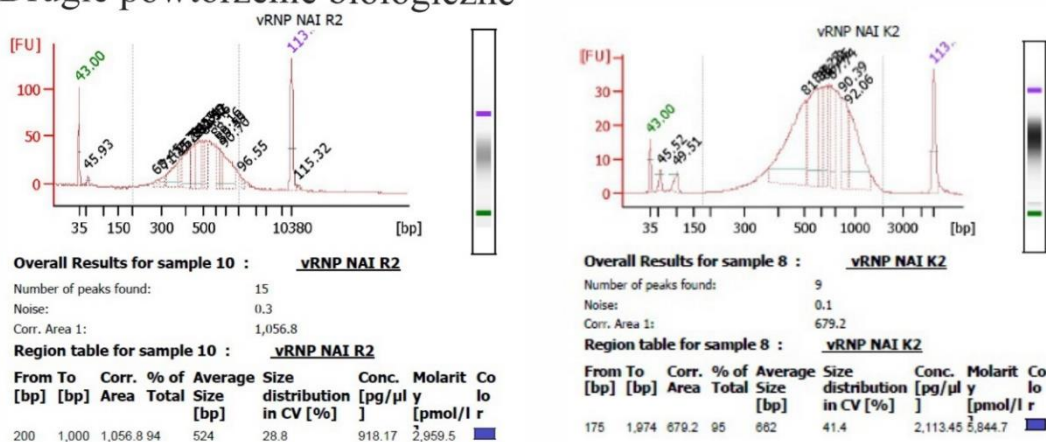


Ryc. 1 Wykresy dla bibliotek fragmentów wygenerowanych na matrycy RNA pochodzącej z mapowania *in vivo* odczynnikami DMS. R oznacza próbkę reakcyjną (traktowaną odczynnikami DMS), K oznacza próbkę kontrolną (nieotraktowaną odczynnikami DMS). Po lewej stronie (oś Y) wyznaczony jest poziom fluorescencji [FU], na dole (oś X) zaznaczona jest długość rozdzielanych fragmentów. Pod wykresami znajduje się tabela ze stężeniem próbki w wybranym regionie. Wykresy wygenerowane przez Agilent Bioanalyzer 2100 na podstawie rozdziału elektroforetycznego próbek.

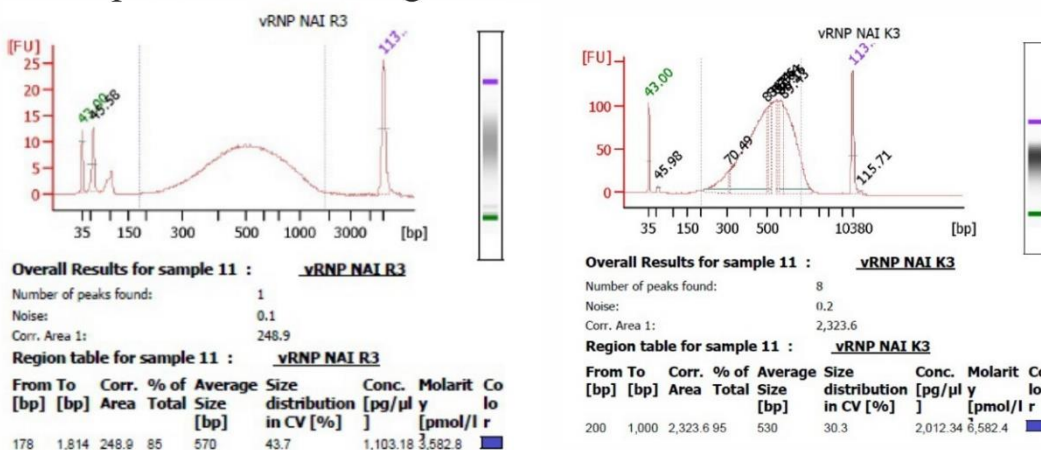
Pierwsze powtórzenie biologiczne



Drugie powtórzenie biologiczne

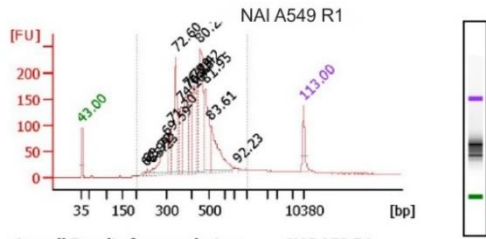


Trzecie powtórzenie biologiczne



Ryc. 48 Wykresy dla bibliotek fragmentów wygenerowanych na matrycy RNA pochodzącej z mapowania *in virio* odczynnikami NAI. R oznacza próbkę reakcyjną (traktowaną odczynnikami NAI), K oznacza próbkę kontrolną (nietraktowaną odczynnikami NAI). Po lewej stronie (oś Y) wyznaczony jest poziom fluorescencji [FU], na dole (oś X) zaznaczona jest długość rozdzielanych fragmentów. Pod wykresami znajduje się tabela ze stężeniem próbki w wybranym regionie. Wykresy wygenerowane przez Agilent Bioanalyzer 2100 na podstawie rozdziału elektroforetycznego próbek.

Pierwsze powtórzenie biologiczne

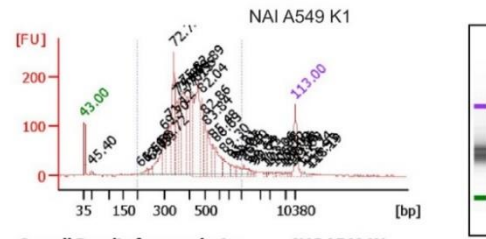


Overall Results for sample 1 : NAI A549 R1

Number of peaks found: 14
 Noise: 0.3
 Corr. Area 1: 3,391.2

Region table for sample 1 : NAI A549 R1

From [bp]	To [bp]	Corr. Area [bp]	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit y [pmol/l]	Co lo r
200	1,000	3,391.2	94	422	24.3	3,641.43	14,047.8	■



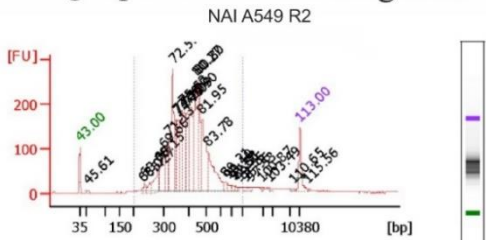
Overall Results for sample 4 : NAI A549 K1

Number of peaks found: 37
 Noise: 0.8
 Corr. Area 1: 3,375.8

Region table for sample 4 : NAI A549 K1

From [bp]	To [bp]	Corr. Area [bp]	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit y [pmol/l]	Co lo r
200	1,000	3,375.8	94	424	27.0	2,742.03	10,642.1	■

Drugie powtórzenie biologiczne

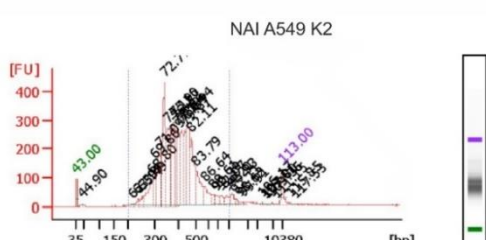


Overall Results for sample 2 : NAI A549 R2

Number of peaks found: 29
 Noise: 0.8
 Corr. Area 1: 3,763.6

Region table for sample 2 : NAI A549 R2

From [bp]	To [bp]	Corr. Area [bp]	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit y [pmol/l]	Co lo r
200	1,000	3,763.6	95	412	24.4	3,056.15	12,051.3	■



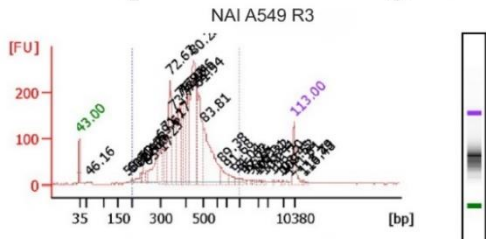
Overall Results for sample 5 : NAI A549 K2

Number of peaks found: 34
 Noise: 1.1
 Corr. Area 1: 5,367.0

Region table for sample 5 : NAI A549 K2

From [bp]	To [bp]	Corr. Area [bp]	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit y [pmol/l]	Co lo r
200	1,000	5,367.0	95	417	27.1	4,609.65	18,130.2	■

Trzecie powtórzenie biologiczne

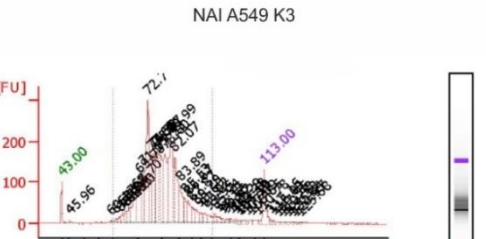


Overall Results for sample 3 : NAI A549 R3

Number of peaks found: 33
 Noise: 1.2
 Corr. Area 1: 4,000.2

Region table for sample 3 : NAI A549 R3

From [bp]	To [bp]	Corr. Area [bp]	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit y [pmol/l]	Co lo r
200	1,000	4,000.2	96	422	25.4	3,385.34	13,223.1	■



Overall Results for sample 6 : NAI A549 K3

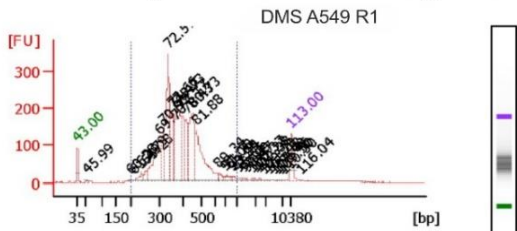
Number of peaks found: 45
 Noise: 1.3
 Corr. Area 1: 3,742.3

Region table for sample 6 : NAI A549 K3

From [bp]	To [bp]	Corr. Area [bp]	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit y [pmol/l]	Co lo r
200	1,000	3,742.2	91	417	28.3	3,062.32	12,195.3	■

Ryc. 49 Wykresy dla bibliotek fragmentów wygenerowanych na matrycy RNA pochodzącej z mapowania *in cellulo* odczynnikiem NAI. R oznacza próbkę reakcyjną (traktowaną odczynnikiem NAI), K oznacza próbkę kontrolną (nietraktowaną odczynnikiem NAI). Po lewej stronie (oś Y) wyznaczony jest poziom fluorescencji [FU], na dole (oś X) zaznaczona jest długość rozdzielanych fragmentów. Pod wykresami znajduje się tabela ze stężeniem próbki w wybranym regionie. Wykresy wygenerowane przez Agilent Bioanalyzer 2100 na podstawie rozdziału elektroforetycznego próbek.

Pierwsze powtórzenie biologiczne

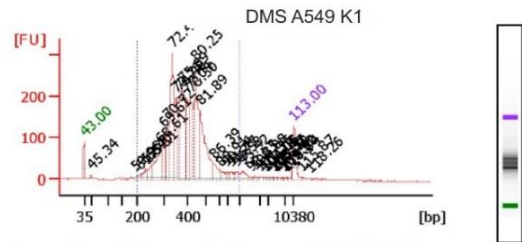


Overall Results for sample 7 : DMS A549 R1

Number of peaks found: 33
 Noise: 0.7
 Corr. Area 1: 3,850.0

Region table for sample 7 : DMS A549 R1

From [bp]	To [bp]	Corr. Area	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit [pmol/l]	Co lo
200	1,000	3,850.0	93	401	26.4	3,581.72	14,604.1	



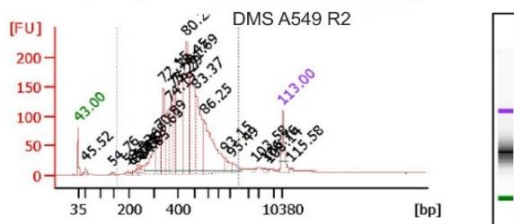
Overall Results for sample 9 : DMS A549 K1

Number of peaks found: 35
 Noise: 0.8
 Corr. Area 1: 3,959.7

Region table for sample 9 : DMS A549 K1

From [bp]	To [bp]	Corr. Area	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit [pmol/l]	Co lo
200	1,000	3,959.7	93	404	26.7	3,782.94	15,340.9	

Drugie powtórzenie biologiczne

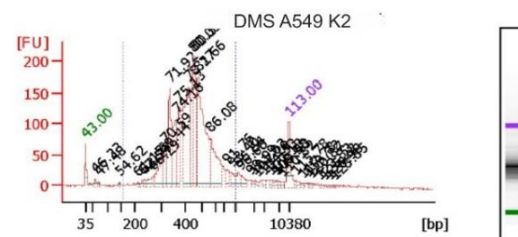


Overall Results for sample 9 : dms r2 a549 powt

Number of peaks found: 24
 Noise: 0.3
 Corr. Area 1: 3,397.5

Region table for sample 9 : dms r2 a549 powt

From [bp]	To [bp]	Corr. Area	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit [pmol/l]	Co lo
160	1,494	3,397.5	95	470	34.6	3,660.16	13,293.5	



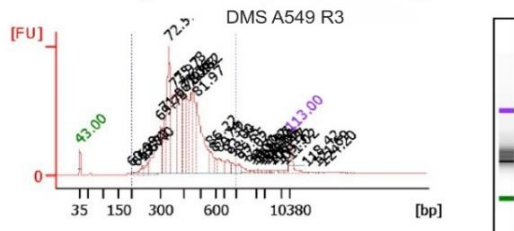
Overall Results for sample 10 : dms k2 a549 powt

Number of peaks found: 37
 Noise: 0.9
 Corr. Area 1: 3,013.8

Region table for sample 10 : dms k2 a549 powt

From [bp]	To [bp]	Corr. Area	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit [pmol/l]	Co lo
151	1,000	3,013.8	89	449	27.4	2,857.53	10,595.2	

Trzecie powtórzenie biologiczne

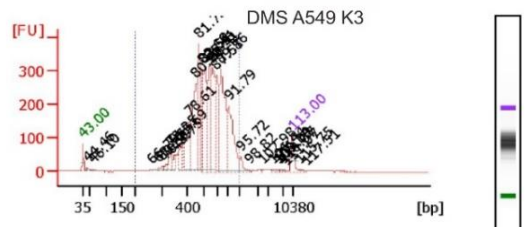


Overall Results for sample 8 : DMS A549 R3

Number of peaks found: 34
 Noise: 0.8
 Corr. Area 1: 6,484.8

Region table for sample 8 : DMS A549 R3

From [bp]	To [bp]	Corr. Area	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit [pmol/l]	Co lo
200	1,000	6,484.8	94	419	28.5	4,498.45	17,787.6	



Overall Results for sample 10 : DMS A549 K3

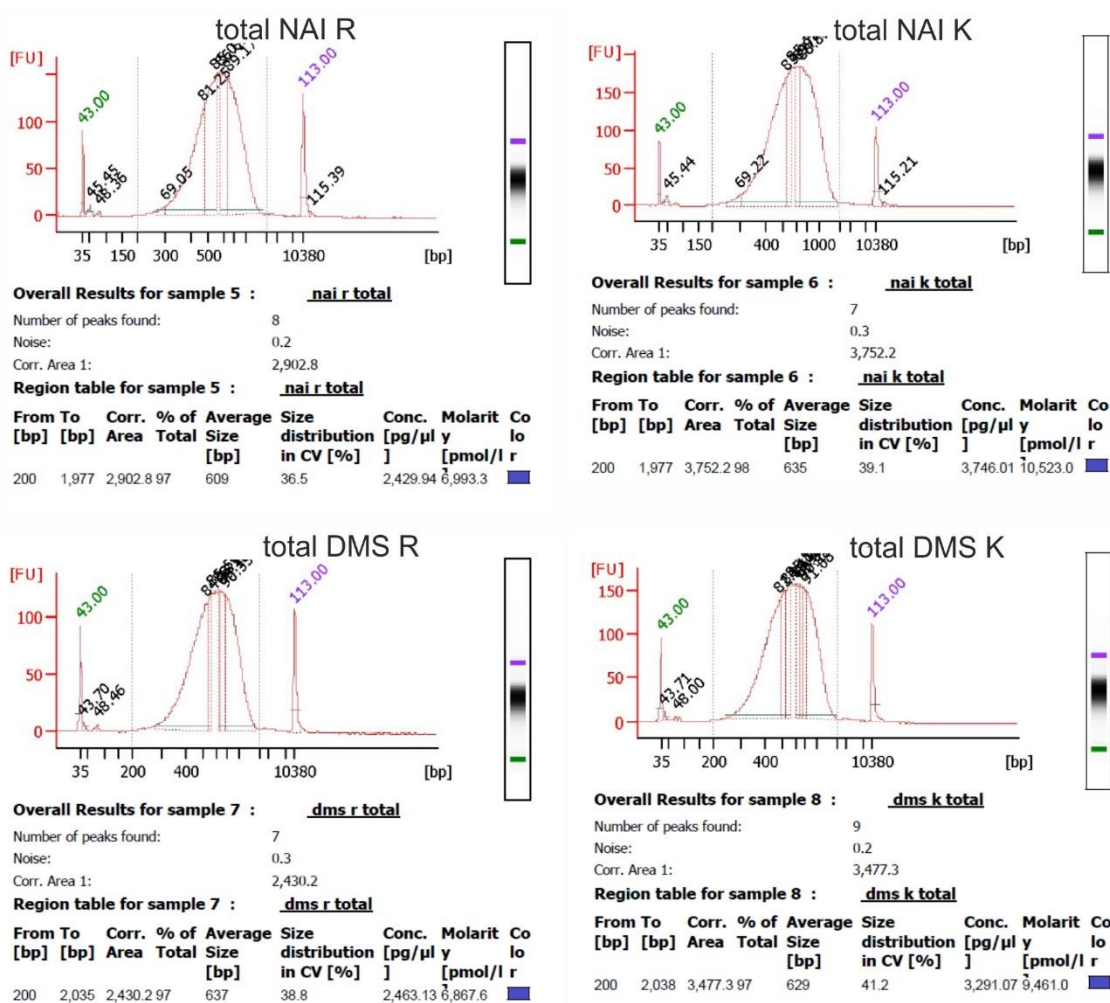
Number of peaks found: 29
 Noise: 1.0
 Corr. Area 1: 5,007.1

Region table for sample 10 : NAI A549 K2

From [bp]	To [bp]	Corr. Area	% of Total Area	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/μl]	Molarit [pmol/l]	Co lo
200	1,000	5,007.1	95	551	23.8	4,791.53	14,219.3	

Ryc. 50 Wykresy dla bibliotek fragmentów wygenerowanych na matrycy RNA pochodzącej z mapowania *in cellulo* odczynnikami DMS. R oznacza próbkę reakcyjną (traktowaną odczynnikami DMS), K oznacza próbkę kontrolną (nieotraktowaną odczynnikami DMS). Po lewej stronie (oś Y) wyznaczony jest poziom fluorescencji [FU], na dole (oś X) zaznaczona jest długość rozdzielanych fragmentów. Pod wykresami znajduje się tabela ze stężeniem próbki w wybranym regionie. Wykresy wygenerowane przez Agilent Bioanalyzer 2100 na podstawie rozdzielania elektroforetycznego próbek.

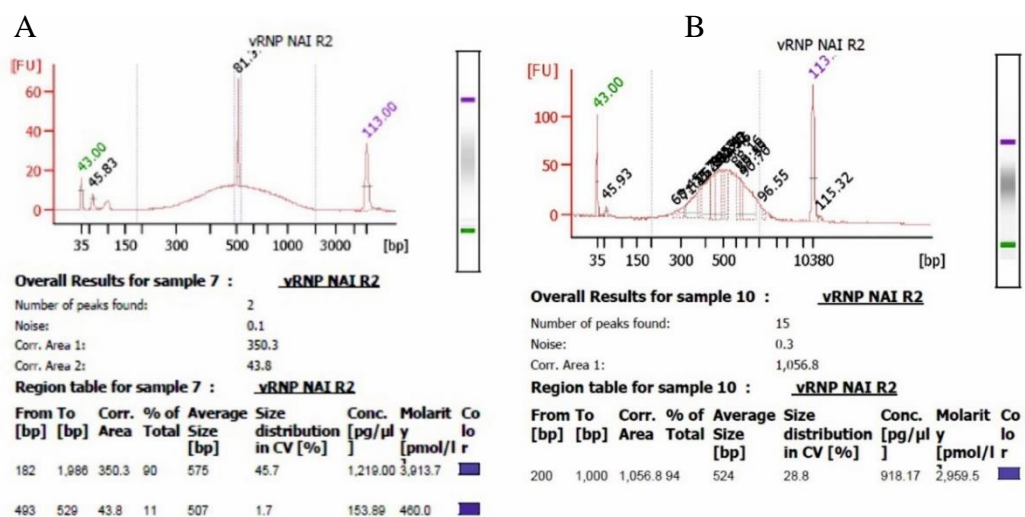
W przypadku bibliotek pochodzących z eksperymentu mapowania *in cellulo* zaobserwowano niejednorodny rozkład fragmentów (**Ryc. 49-50**). Taki rozkład jest charakterystyczny dla bibliotek przygotowywanych na matrycy produktu PCR ze starterami HFA i HRA i był zaobserwowany dla bibliotek NGS pochodzących z niezależnych eksperymentów wykonanych przez naszą grupę (dane nieopublikowane). W przypadku prób kontroli *in cellulo* uzyskano równomierny rozdział biblioteki fragmentów do sekwencjonowania NGS (**Ryc. 51**).



Ryc. 51 Wykresy dla bibliotek fragmentów wygenerowanych na matrycy RNA próby kontrolnej *in cellulo* (total) dla eksperymentu mapowania odczynnikami DMS i NAI. R oznacza próbkę reakcyjną (traktowaną odczynnikami), K oznacza próbkę kontrolną (nietraktowaną odczynnikami). Po lewej stronie (oś Y) wyznaczony jest poziom fluorescencji [FU], na dole (oś X) zaznaczona jest długość rozdzielanych fragmentów. Pod wykresami znajduje się tabela ze stężeniem próbki w wybranym regionie. Wykresy wygenerowane przez Agilent Bioanalyzer 2100 na podstawie rozdziału elektroforetycznego próbek.

W przypadku rozdziału elektroforetycznego na bioanalizatorze mogą pojawić się wspomniane wcześniej zakłócenia w rozdziale, tzn. artefakty. W celu potwierdzenia, że pojawiające się zakłócenia nie są niespecyficznymi produktami w bibliotece powtórzono rozdział próbki vRNP NAI R2 z widocznym silnym artefaktem (**Ryc. 52A**). Po lewej stronie (**Ryc. 52A**) rozdział z widocznym artefaktem na wysokości około 500 pz, który został błędnie przypisany jako pik. Po powtórzeniu rozdziału tej samej biblioteki (**Ryc. 52B**) nie zaobserwowano występowania dodatkowego pik. Potwierdziło to, że pik był wynikiem zaburzenia rozdziału elektroforetycznego.

Po potwierdzeniu poprawnego rozkładu fragmentów w bibliotece do sekwencjonowania wykonano normalizację bibliotek do sekwencjonowania. Dokładny opis tego etapu został zawarty w rozdziale *Materiały i metody*. W **Tabeli 10** (*Dodatek*) zaprezentowane jest podsumowanie informacji dotyczących zastosowanych indeksów oraz obliczenie stężeń nM. dla poszczególnych próbek do sekwencjonowania. Obliczenie stężeń w nM umożliwia połączenie równych ilości próbek w końcowej mieszaninie do sekwencjonowania. Nieprawidłowe obliczenie wartości nM może prowadzić do nadreprezentacji którejś z próbek i większe pokrycie tej próbki w trakcie sekwencjonowania. Sekwencjonowanie typu paired-end wykonano z zastosowaniem systemu Illumina NextSeq550, Długość odczytów obu końców cząsteczki ustalono na 2x150 pz.



Ryc. 52 Wykresy dla biblioteki fragmentów vRNP NAI R2. **A.** Rozdział z widocznym artefaktem na wysokości ok 500 pz, który został błędnie przypisany jako pik. **B.** Powtórzony rozdział elektroforetyczny tej samej biblioteki.

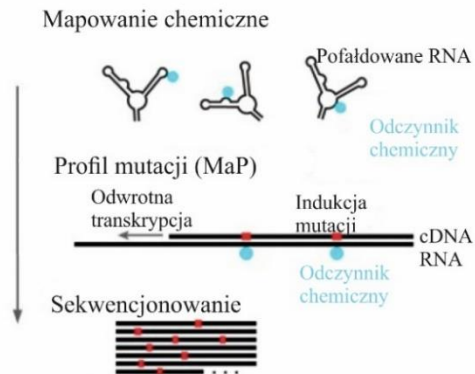
2.2.7. Wyniki sekwencjonowania

2.2.7.1 Analiza bioinformatyczna wyników sekwencjonowania

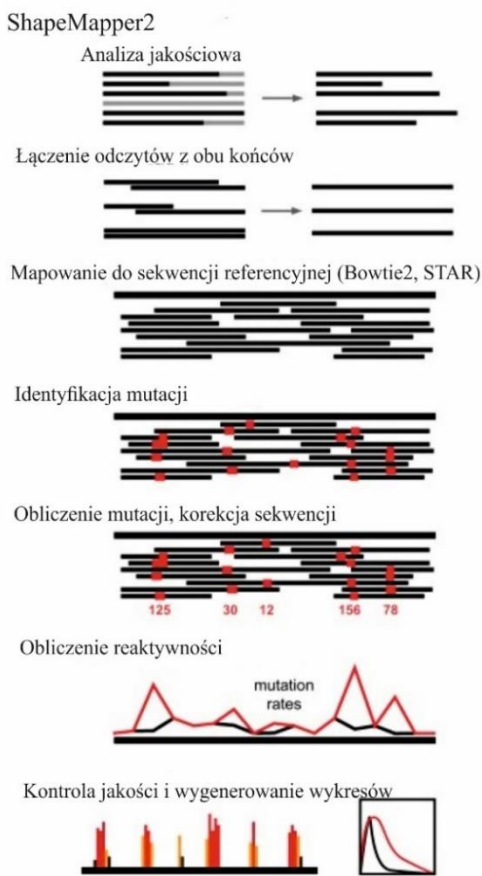
Po części eksperymentalnej, zakończonej sekwencjonowaniem nastąpił etap analizy bioinformatycznej. Analizę wyników sekwencjonowania wykonał dr. Tomasz Woźniak (Instytut Genetyki Człowieka PAN). Analiza wyników sekwencjonowania została przeprowadzona w programie ShapeMapper2²⁰⁵. Program jest dedykowany do analizy danych z sekwencjonowania metodą SHAPE-MaP oraz DMS-MaPseq. ShapeMapper2 umożliwia analizę profilu mutacji poprzez przyrównanie próby reakcyjnej (reakcja mapowania odczynnikami chemicznymi) z próbą kontrolną (nieotraktowaną odczynnikami chemicznymi). Na **Ryc. 53B** został zaprezentowany schemat analizy w programie ShapeMapper2, który obejmuje etapy takie jak: 1) Analiza jakościowa sekwencjonowania poprzez eliminację odczytów o niskiej jakości, 2) Łączenie odczytów sekwencjonowania obu końców (sekwencjonowanie typu paired-end), 3) Mapowanie do sekwencji referencyjnej wybranym algorytmem (Bowtie2, STAR), 4) Identyfikacja mutacji zarówno w próbce reakcyjnej jak i kontrolnej, 5) Przyrównanie próbki reakcyjnej i kontrolnej w celu identyfikacji miejsc o zwiększonej ilości mutacji, 6) Obliczenie reaktywności na podstawie profilu mutacji, 7) Kontrola jakości i wygenerowanie wykresów. Jako mutacje identyfikowane są zmiany takie jak jednonukleotydowe lub wielonukleotydowe niesparowania, delecje czy insercje (**Ryc. 53C,D**). Po identyfikacji miejsc o skumulowanej mutacji obliczana jest reaktywność poszczególnych nukleotydów.

Sekwenator NextSeq550 wykonuje automatyczny podział próbek na podstawie zastosowanych indeksów oraz dodatkowo usuwa sekwencje starterów w ostatecznej wersji pliku (ang. primer trimming). Uzyskane pliki w formacie FASTQ zostały poddane analizie jakościowej i ilościowej. Następnie połączono pliki dla odczytu obu końców.

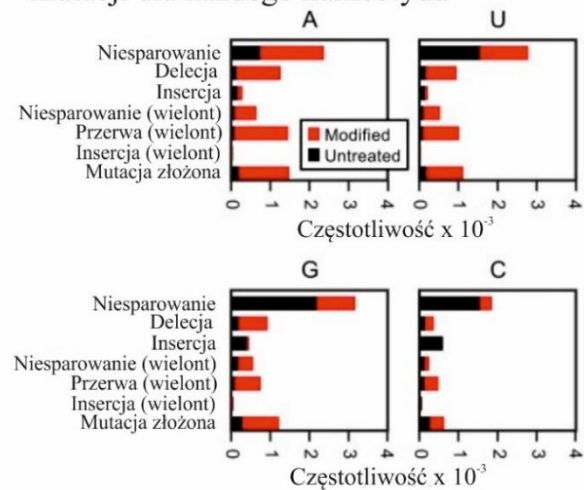
A. Część eksperymentalna



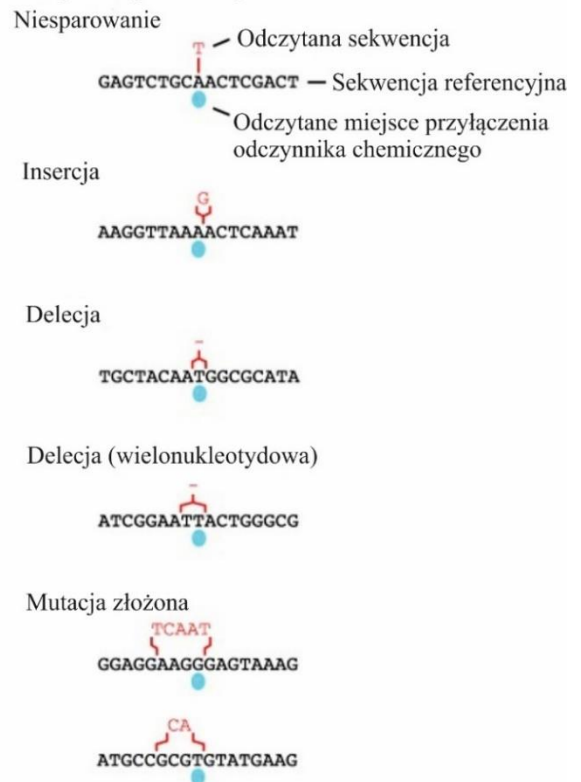
B. Analiza bioinformatyczna



C. Statystyka poszczególnych typów mutacji dla każdego nukleotydu



D. Przykłady mutacji



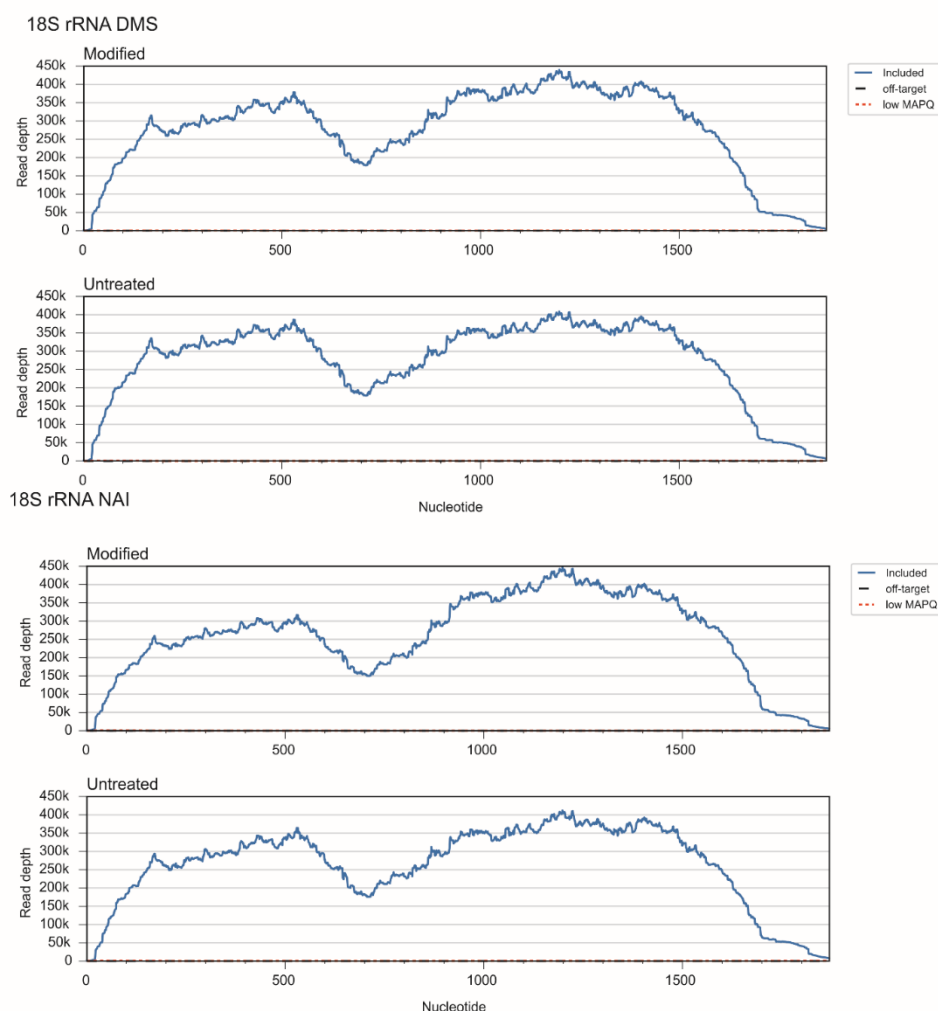
Ryc. 53 **A.** Część eksperymentalna metody detekcji modyfikacji chemicznych poprzez profilowanie mutacji MaP. **B.** Kolejne etapy analizy bioinformatycznej przeprowadzanej w programie ShapeMapper2. **C.** Typy mutacji oraz częstotliwość ich występowania dla danego nukleotydu na przykładzie analizy profilu mutacji 16S i 23S rRNA E.coli **D.** Przykłady występujących mutacji. Rysunek pochodzi z publikacji Busan S., Weeks K.M., (RNA, 2018).

2.2.7.2. Analiza wyników mapowania chemicznego RNA o znanej strukturze drugorzędowej – kontrola mapowania in cellulo

W czasie planowania eksperymentu duże znaczenie miał wybór linii komórkowej, w której przeprowadzano infekcję wirusem grypy. Wybrano ludzką linię komórkową płuc A549, co miało na celu uzyskanie informacji dotyczącej struktury vRNA w warunkach jak najbardziej zbliżonych do naturalnych. Linia komórkowa A549 jest standardowo używana w eksperymentach z wirusem grypy, co wynika między innymi z dobrej kinetyki namnażania wirusa w komórkach tej linii. Po drugie, wybór linii komórkowej podyktowany był również możliwością wybrania cząsteczki kontrolnej RNA o znanej strukturze drugorzędowej, która była mapowana w tym samym czasie co badane cząsteczki vRNA. Pozwoliło to wybrać 18S rRNA ludzki jako kontrolę pozytywną reakcji mapowania SHAPE-MaP²⁰⁶. Co istotne, dla cząsteczki 18S rRNA dostępne są również dane dotyczące struktury całego kompleksu rybonukleinoproteinowego²⁰⁷. Kontrola ta umożliwiła sprawdzenie poprawności przeprowadzonej analizy bioinformatycznej pod względem dobrania właściwych parametrów w programie ShapeMapper2. Uzyskane wyniki miały również za zadanie wskazać czy warunki eksperymentalne, takie jak stężenie odczynnika mapującego i czas reakcji, były dobrane prawidłowo.

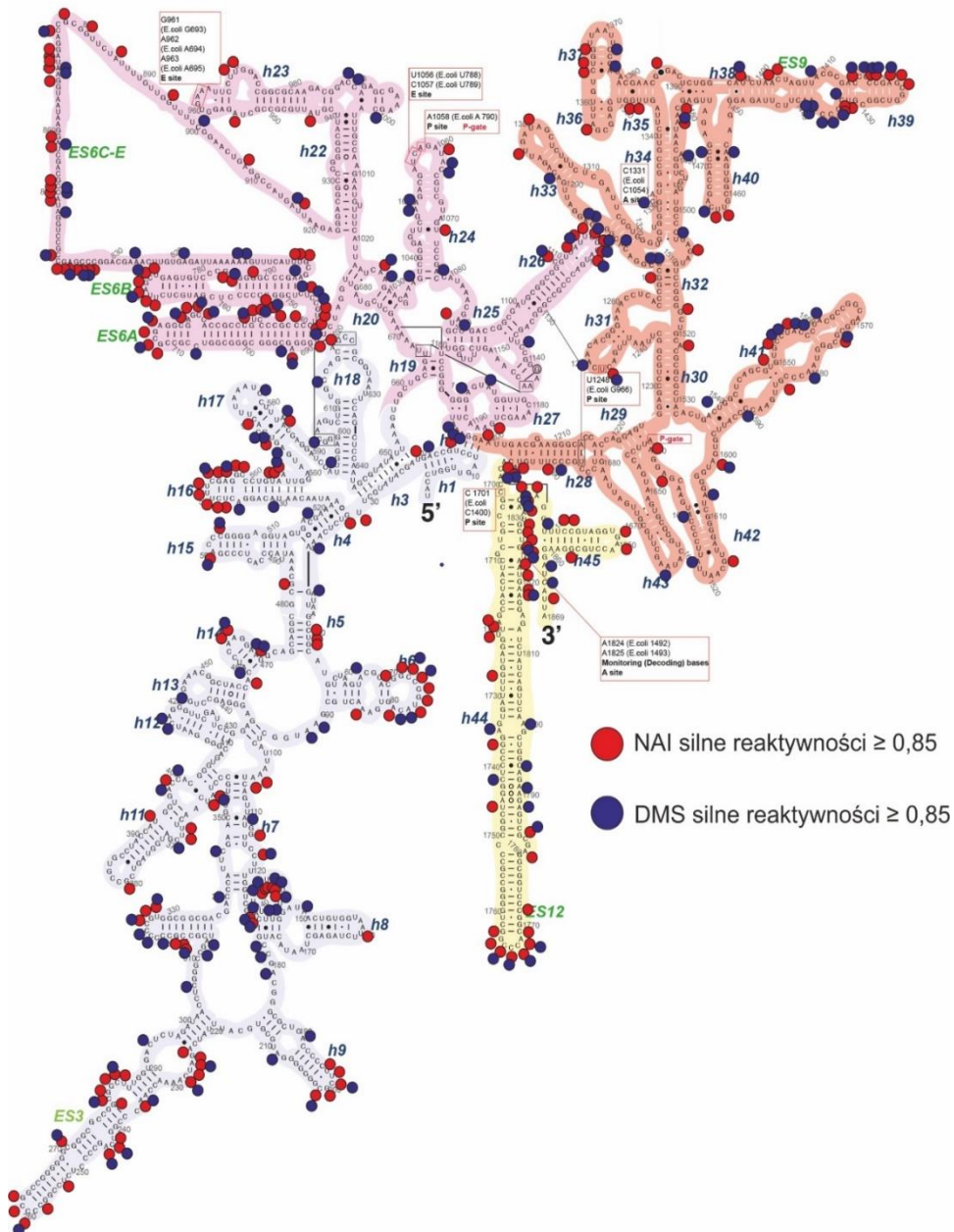
Na **Ryc. 54** zaprezentowane jest pokrycie sekwencji 18S rRNA dla prób reakcyjnych i kontrolnych mapowanych odczynnikami DMS i NAI. Po uzyskaniu informacji odnośnie pokrycia sekwencji obliczone zostały reaktywności nukleotydów. W przypadku obu odczynników chemicznych uzyskano odczyt reaktywności dla większości nukleotydów cząsteczki 18S rRNA. Nie uzyskano odczytu reaktywności jedynie dla nukleotydów 1-14. Dane mapowania chemicznego naniesiono na strukturę drugorzędową 18S rRNA (**Ryc. 55**). Wyniki mapowania korelują ze strukturą drugorzędową 18S rRNA. Silne mapowania ($\geq 0,85$) występują w rejonach jednoniciowych, bądź w obrębie par niekanonicznych (G-U, A-C). Zidentyfikowano również pewne motywy strukturalne, w których nie zaobserwowano mapowań chemicznych. W następnej kolejności zbadano dostępność tych motywów dla odczynników chemicznych w strukturze trzeciorzędowej 18S rRNA. Strukturę podjednostki 18S rybosomu ludzkiego pobrano z bazy danych PDB (Protein Data

Bank, entry: 4V6Z, chain: B2 doi: 10.2210/pdb4V6X/pdb)²⁰⁷. Wizualizacja struktury w programie Chimera (wer. 1.13.1) została przeprowadzona przez dr. Agnieszkę Ruszkowską (Zakład Chemii i Biologii Kwasów Nukleinowych, IChB PAN). Analiza wykazała, że motywy strukturalne, które były wyraźnie wyeksponowane na zewnątrz kompleksu rybosomalnego silnie mapowały się w miejscach jednoniciowych. Z kolei motywy strukturalne z nisko reaktywnymi nukleotydami zostały zidentyfikowane wewnątrz kompleksu rybosomalnego. Na **Ryc. 56** zaprezentowano przykładowe motywy strukturalne oraz ich lokalizację w cząsteczce kompleksu rybosomalnego. W pierwszym motywie strukturalnym zaobserwowano mapowania zgodne ze strukturą drugorzędową, bowiem nukleotydy o silnej reaktywności znajdowały się w obrębie regionu jednoniciowego tego motywu. Analiza miejsca występowania tego motywu

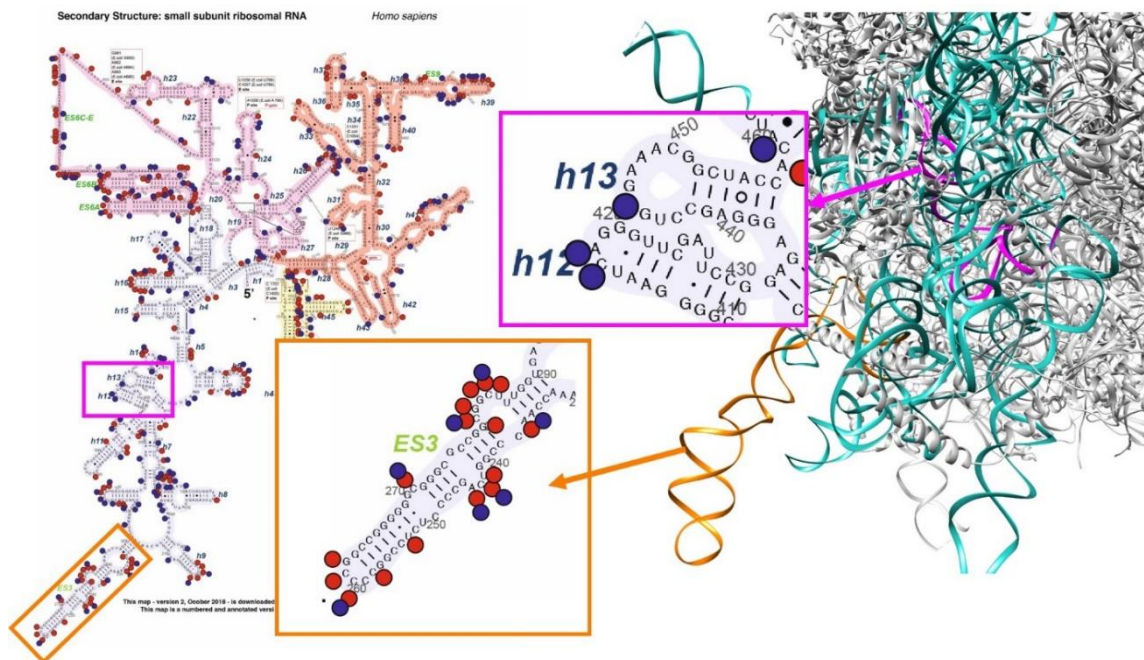


Ryc. 54 Wykresy pokrycia sekwencji 18S rRNA dla próbek reakcyjnych (modified) i kontrolnych (untreated) dla eksperymentu mapowania odczynnikami DMS i NAI. Na osi y – ilość odczytów, na osi x – numer nukleotydu w sekwencji. Wykresy zostały wygenerowane w narzędziu Tincov (pypi.org/project/tincov) (dr Tomasz Woźniak).

wskazała, że był on silnie ekspozycyjny na zewnątrz kompleksu. Drugi motyw, który zlokalizowany był wewnątrz kompleksu rybosomalnego, charakteryzował się zmniejszoną reaktywnością nukleotydów również w obrębie regionów, które zostały zidentyfikowane jako jednoniciowe w strukturze drugorzędowej. Analiza wyników mapowania cząsteczki 18S rRNA wskazała na prawidłowe przeprowadzenie eksperymentu mapowania chemicznego RNA, a także analizy bioinformatycznej.



Ryc. 55 Struktura drugorzędowa ludzkiego 18S rRNA. Na rysunku przedstawiono wyniki mapowania chemicznego odczynnikami DMS i NAI. Na strukturę naniesiono silne reaktywności ($\geq 0,85$) nukleotydów. Struktura została pobrana ze strony www.jufali.userpage.fu-berlin.de/rRNA-maps.

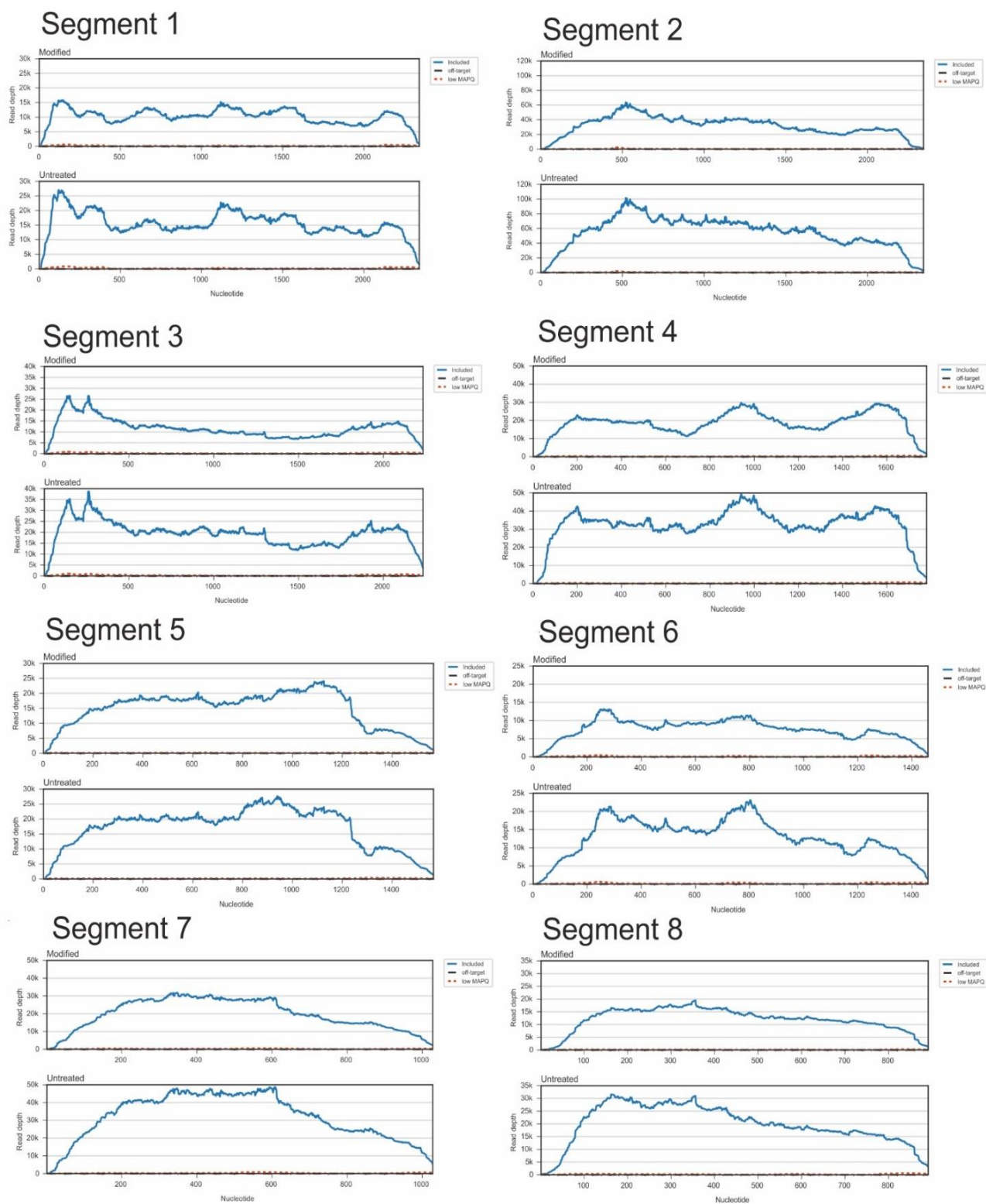


Ryc. 56 Na rysunku zaznaczono przykładowe motywy strukturalne 18S rRNA, z zaznaczeniem reaktywności nukleotydów względem zastosowanych odczynników chemicznych. Lokalizacja motywów w obrębie kompleksu wykazała, że silnie reaktywne nukleotydy występowały w obrębie jednoniciowych rejonów struktur eksponowanych poza kompleks (kolor pomarańczowy). Z kolei mniej reaktywne chemiczne motywy strukturalne były zlokalizowane w środku kompleksu (kolor różowy). Kolorem zielonym zaznaczono cząsteczkę 18S rRNA. Wizualizacja: Dr Agnieszka Ruszkowska (program: Chimera 1.13.1). Structures of the human and Drosophila 80S ribosome, Anger et al.,(2013) Nature 497. DOI: 10.2210/pdb4V6X/pdb Entry: 4V6X

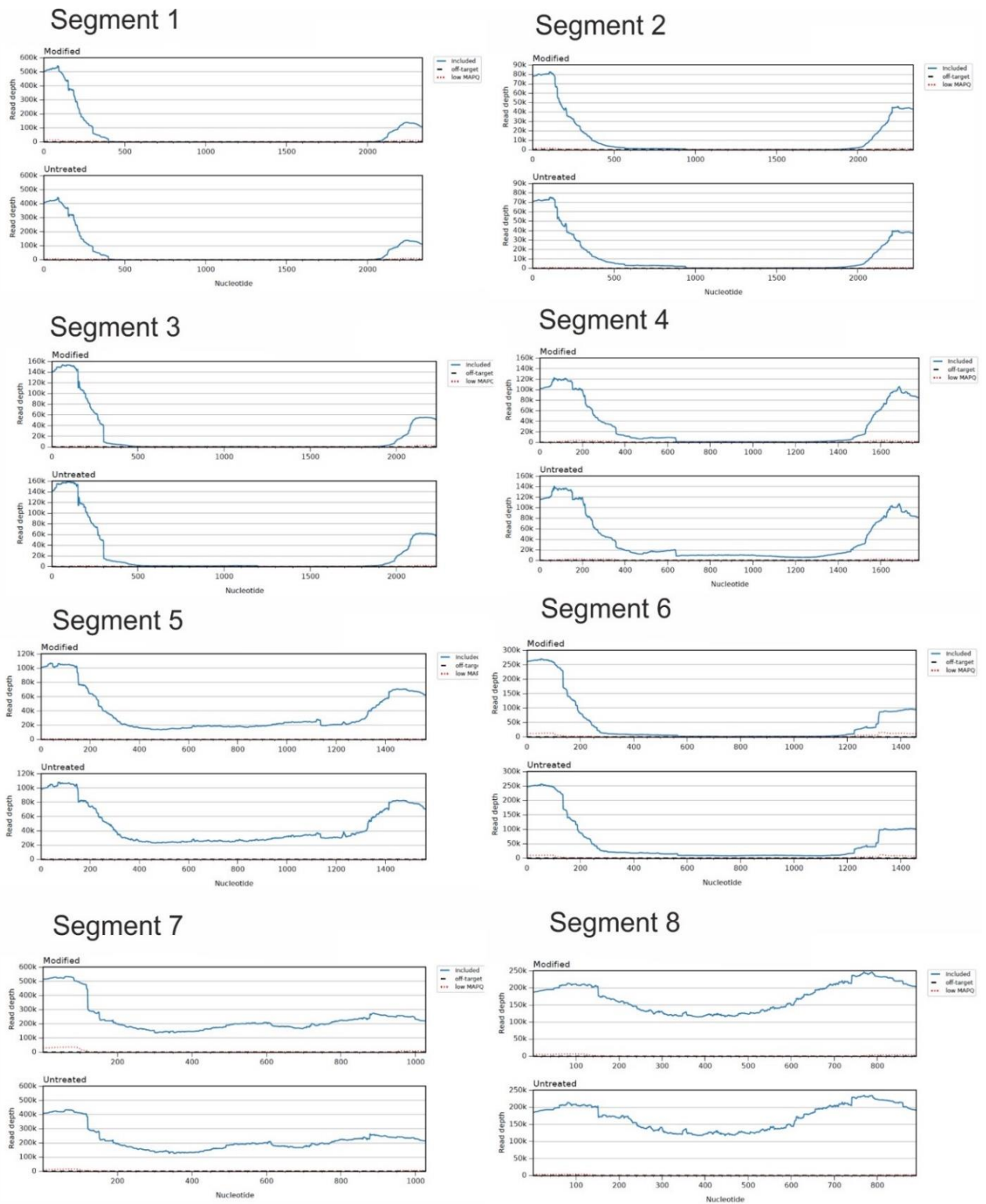
2.2.7.3. Wyniki sekwencjonowania bibliotek fragmentów z eksperymentów mapowania vRNA *in virio* oraz *in cellulo*

Pierwszym etapem analizy było sprawdzenie pokrycia sekwencji wszystkich segmentów vRNA dla każdej próbki. Po mapowaniu odczytów (ang. read mapping) do referencji uzyskano wykresy pokrycia sekwencji referencyjnych poszczególnych segmentów vRNA. Wykresy pokrycia sekwencji referencyjnych wszystkich segmentów vRNA dla przykładowej próbki reakcyjnej i kontrolnej pierwszego powtórzenia biologicznego *in virio* oraz *in cellulo* dla odczynnika NAI są zaprezentowane na **Ryc. 57-58**. W przypadku eksperymentu mapowania vRNA *in virio* uzyskano pokrycie dla wszystkich segmentów vRNA we wszystkich powtórzeniach (**Ryc. 57**). W przypadku eksperymentu mapowania w warunkach *in cellulo* uzyskano część pokrycia sekwencji referencyjnej dla segmentów 1,2,3,4 oraz 6. Dla pozostałych segmentów vRNA 5, 7 i 8 uzyskano całkowite pokrycie sekwencji referencyjnej (**Ryc. 58**). Uzyskanie nierównomiernego pokrycia sekwencji segmentów vRNA może wynikać ze zróżnicowanej kinetyki namnażania tych cząsteczek w czasie cyklu replikacyjnego wirusa²⁰⁴.

Reaktywność nukleotydów vRNA była określana poprzez porównanie profilu mutacji próbki reakcyjnej z profilem mutacji próbki kontrolnej. Na **Ryc. 60** zostały zaprezentowane wykresy wyliczonego profilu mutacji, pokrycia sekwencji referencyjnej oraz dystrybucji reaktywności na przykładzie segmentu 1 vRNA dla pierwszego powtórzenia biologicznego. Na **Ryc. 59** zaprezentowano przykładowe wykresy reaktywności oraz pokrycia dla segmentu 8 vRNA z pierwszego powtórzenia biologicznego dla odczynników NAI i DMS. Dla wszystkich segmentów uzyskano prawidłowy odczyt reaktywności dla całej cząsteczki, za wyjątkiem ok. 20-40 nukleotydów dla 5' końca cząsteczek.

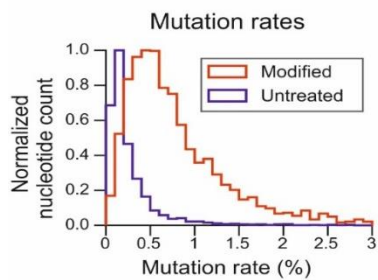


Ryc. 57 Pokrycie sekwencji referencyjnych dla segmentów 1-8 vRNA *in vitro*. Przedstawiono wykresy dla próbki reakcyjnej (modified) i kontrolnej (untreated), na przykładzie danych z pierwszego powtórzenia mapowania struktury drugorzędowej vRNA *in vitro* odczynnikiem NAI. Na osi y – ilość odczytów, na osi x – numer nukleotydu w sekwencji.



Ryc. 58 Pokrycie sekwencji referencyjnych dla segmentów 1-8 vRNA *in cellulo*. Na wykresie zostały przedstawione dane dla próbki reakcyjnej (modified) i kontrolnej (untreated), na przykładzie próbki z pierwszego powtórzenia mapowania vRNA *in cellulo* odczynnikiem NAI. Na osi y – ilość odczytów, na osi x – numer nukleotydu w sekwencji.

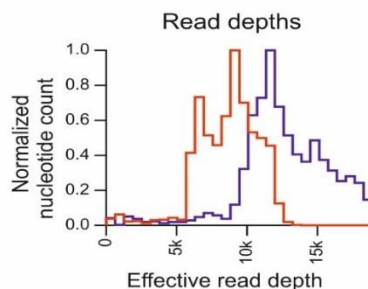
A Segment 1



Modified sample:
95th percentile rate: 2.49%
Median rate: 0.65%

Untreated sample:
95th percentile rate: 0.94%
Median rate: 0.17%

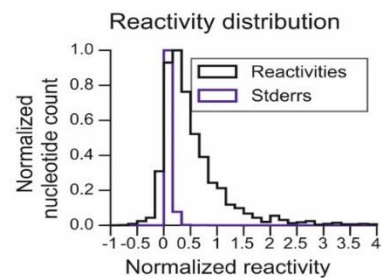
B



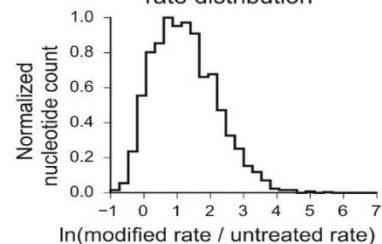
Modified sample:
Median depth: 8,909
5th percentile depth: 5,754

Untreated sample:
Median depth: 12,424
5th percentile depth: 7,573

C

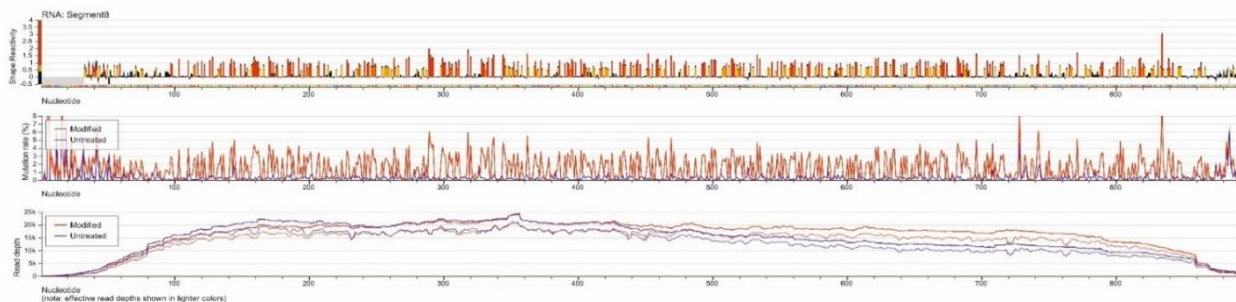


Raw background-corrected rate distribution

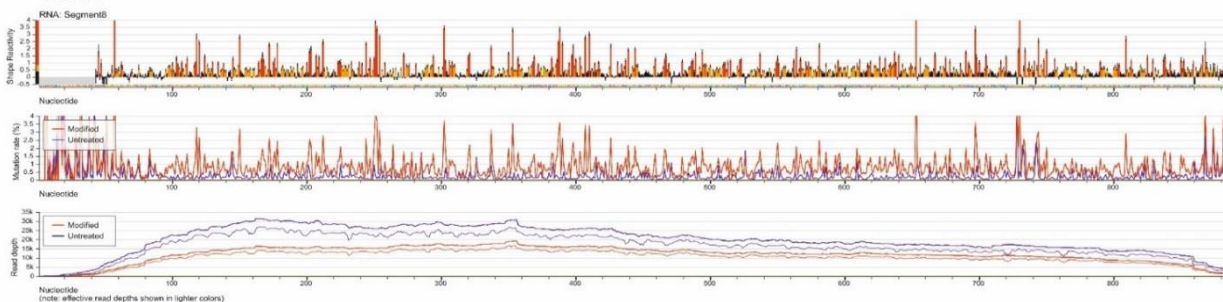


Ryc. 59 A. Przykładowa statystyka zidentyfikowanych mutacji dla próbki reakcyjnej (modified) i kontrolnej (untreated) dla segmentu 1 *in virio*. Większa ilość mutacji w próbce reakcyjnej niż kontrolnej wskazuje na prawidłowy przebieg eksperymentu i analizy bioinformatycznej. B. Ilość efektywnych odczytów dla obu próbek. C. Dystrybucja obliczonych reaktywności wraz z wyliczonym błędem standardowym.

DMS



NAI



Ryc. 60 Wykresy reaktywności nukleotydów, procentu mutacji dla poszczególnych nukleotydów oraz pokrycia sekwencji dla segmentu vRNA8 mapowanego chemicznie *in virio*. Zaprezentowane wykresy pochodzą z pierwszego powtórzenia biologicznego.

Przed przystąpieniem do obliczenia średniej reaktywności wyliczone zostały wartości korelacji Pearson'a dla poszczególnych powtórzeń biologicznych. Dla wszystkich par powtórzeń biologicznych uzyskano współczynnik korelacji w zakresie od 0,80-0,99, co wskazuje na wysoki oraz bardzo wysoki poziom korelacji. Wartości współczynnika korelacji zostały zebrane w ***Tabeli 11*** w rozdziale *Dodatek*.

Po potwierdzeniu wysokiej korelacji, wyliczono średnie wartości reaktywności dla poszczególnych nukleotydów dla obu odczynników mapujących. Podczas fałdowania struktury RNA wszystkie dane uzyskane dla odczynnika NAI zostały wprowadzone do programu RNAStructure (w wersji 6.2) jako tzw. pseudoenergie (ang. *pseudoenergy constraints*). W przypadku danych mapowania odczynnikiem DMS, do programu wprowadzone zostały silne reaktywności (wartości $\geq 0,85$). W tym celu wykorzystano moduł *chemical mapping constraints*, który jest dostosowany do tego odczynnika mapującego. Wybór progu reaktywności uznawany jako wysoki dla eksperymentu DMS-MaPseq został wybrany na podstawie opublikowanych protokołów¹⁴⁰. W czasie fałdowania zastosowano dystans parowania zasad wynoszący 150 nt. Dodatkowo, ze względu na ograniczenie dystansu, a także częściowy brak danych dla końców cząsteczki, z predykcji struktur drugorzędowych wyłączono także rejon promotora *panhandle* poprzez wymuszenie tego rejonu jako ssRNA. Przewidywanie struktury w programie RNAStructure zostało opisane w rozdziale *Materiały i metody*.

2.3. Struktura drugorzędowa vRNA w warunkach *in virio* oraz *in cellulo*

Struktury RNA w warunkach biologicznych charakteryzują się dużą labilnością, a dana cząsteczka RNA może przyjmować wiele alternatywnych struktur. Analiza struktury drugorzędowej segmentu vRNA8 w lizacie komórkowym, przeprowadzona z zastosowaniem różnych parametrów fałdowania, zarówno z ograniczeniem parowania zasad, jak i w ujęciu globalnym, wskazała na zachowawczość motywów lokalnych w każdej ze struktur. Porównanie struktur o najmniejszej energii swobodnej ze strukturami MEA wskazało również na stosunkowo niewielkie różnice pomiędzy strukturami i zachowawczość motywów lokalnych.

Wobec tego, w czasie analizy struktury drugorzędowej vRNA *in virio* i *in cellulo* skupiano się na strukturze MEA¹⁹².

W pierwszej kolejności obliczono prawdopodobieństwo par zasad (ppz) za pomocą funkcji *Partition Function* w programie RNAStructure. Jak już wspomniano wcześniej, obliczone prawdopodobieństwo daje informacje dotyczącą częstości pojawiania się określonej pary zasad we wszystkich wygenerowanych strukturach. Na podstawie obliczeń wybrano motywy strukturalne o największym prawdopodobieństwie parowania zasad (>80%). Następnie obliczono strukturę MEA poprzez wprowadzenie danych ppz do RNAStructure za pomocą funkcji *MaxExpect: Predict RNA MEA Structure*.

2.3.1. Segmenty 1, 2, 3, 4 oraz 6 vRNA

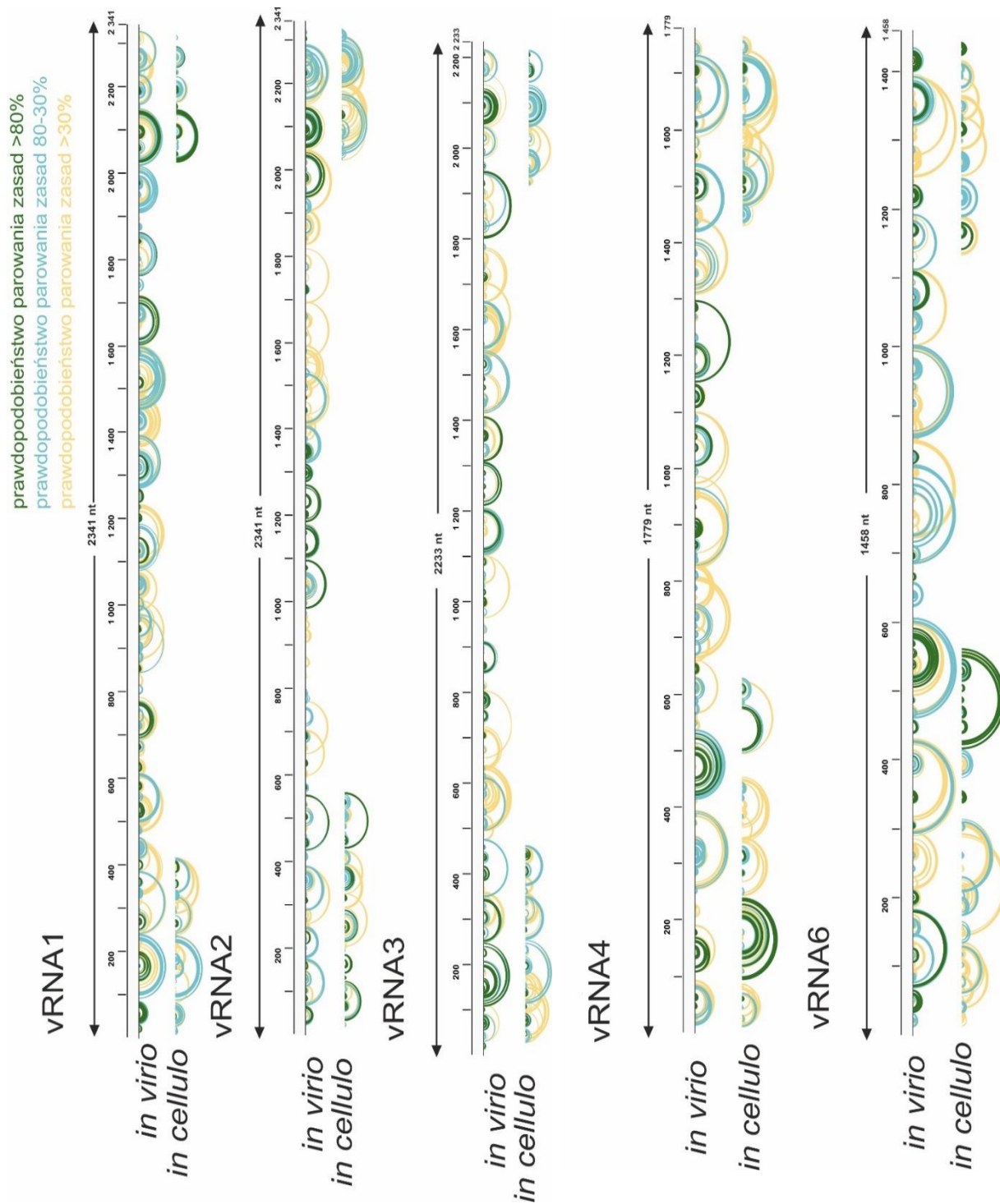
Segmenty vRNA1, vRNA2, vRNA3, vRNA4 oraz vRNA6 są najmniej poznanymi cząsteczkami vRNA. W przypadku szczepu A/California/04/2009 dotychczas nie opublikowano danych dotyczących struktur drugorzędowych dla tych cząsteczek. W niniejszej pracy po raz pierwszy opisane zostały struktury drugorzędowe tych segmentów vRNA w warunkach *in virio*, a także część danych dotyczących struktur drugorzędowych w warunkach *in cellulo* szczepu A/California/04/2009. Na **Ryc. 61** zamieszone zostały wykresy prawdopodobieństwa parowania zasad dla struktur drugorzędowych wybranych segmentów vRNA w warunkach *in virio* oraz *in cellulo*. W opisie wspólnych motywów strukturalnych wzięto pod uwagę jedynie te, których prawdopodobieństwo parowania zasad wynosiło więcej niż 80%.

Wybrane motywy strukturalne były analizowane na trzech płaszczyznach. W pierwszej kolejności wybrano motywy strukturalne vRNA *in virio* o najwyższych prawdopodobieństwach parowania zasad. Następnie porównano te motywy strukturalne z motywami vRNA obecnymi w strukturze *in cellulo*. Na ostatnim etapie analizy, dane strukturalne zostały porównane do opublikowanych niedawno struktur *in virio* oraz *ex virio* szczepu A/WSN/1933 tego samego podtypu wirusa (H1N1)¹¹⁵. Na tej podstawie określono motywy RNA wspólne dla obu szczepów.

W przypadku segmentu vRNA1 *in virio* prawdopodobieństwo ppz powyżej 80% odnotowano dla 20 motywów strukturalnych (**Ryc. 62**). Spośród nich zidentyfikowano 5 motywów, które również miały wysokie ppz (>80%) w warunkach *in cellulo*: 260-

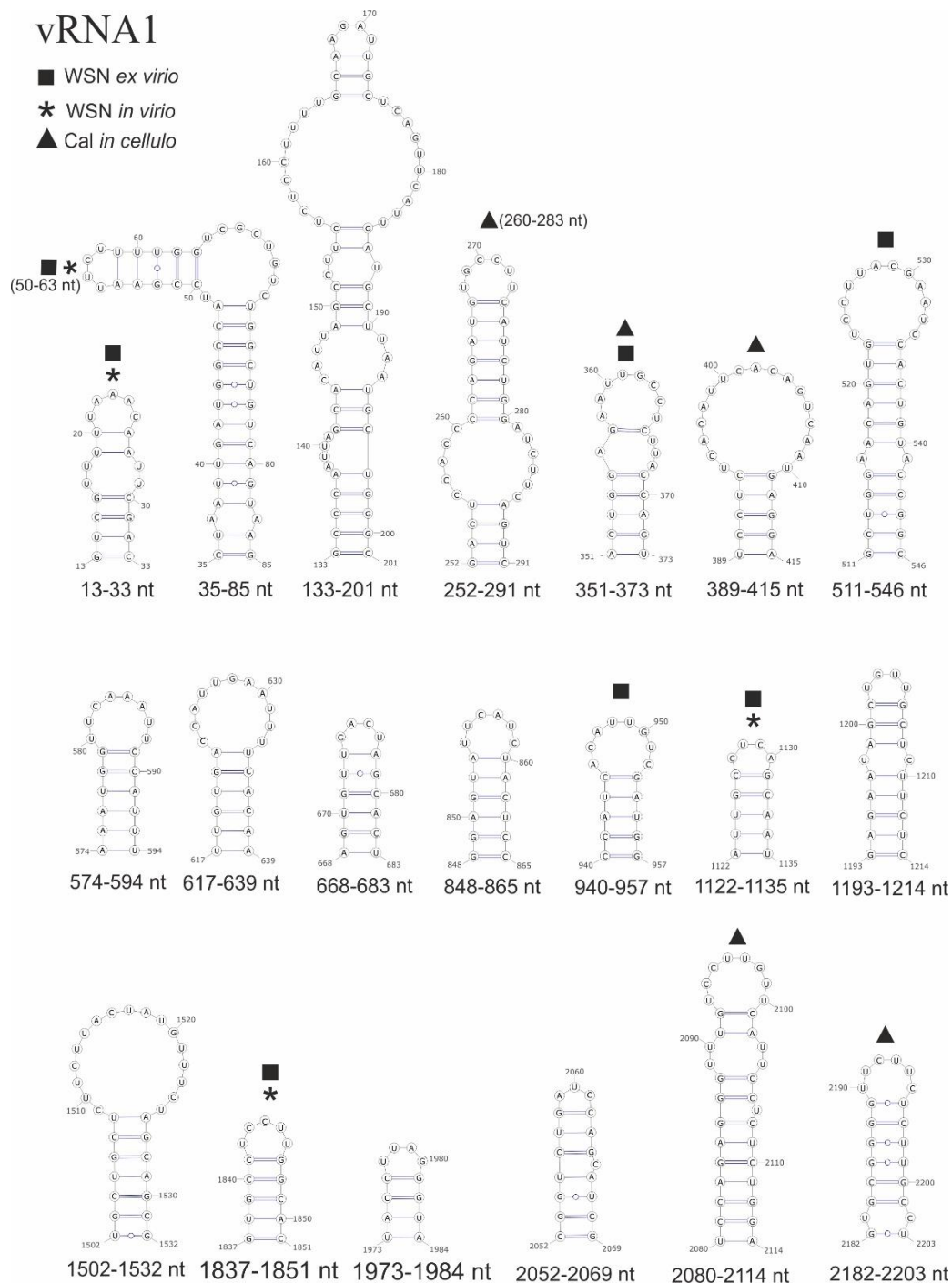
283 nt, 351-373 nt, 389-415 nt, 2080-2114 nt oraz 2182-2203 nt. Cztery motywy były obecne w strukturach *ex virio* oraz *in virio* szczepu A/WSN/1933: 13-33 nt, 50-63 nt, 1122-1135 nt oraz 1837-1851 nt (**Ryc. 62**). Natomiast motywy w rejonach 351-373 nt, 511-546 nt oraz 940-957 nt były przewidziane w strukturze *ex virio*.

W strukturze vRNA2 *in virio* zidentyfikowano 20 motywów strukturalnych z wysokim ppz, z czego 4 motywy strukturalne zostały przewidziane również w warunkach *in cellulo* (**Ryc. 63**).



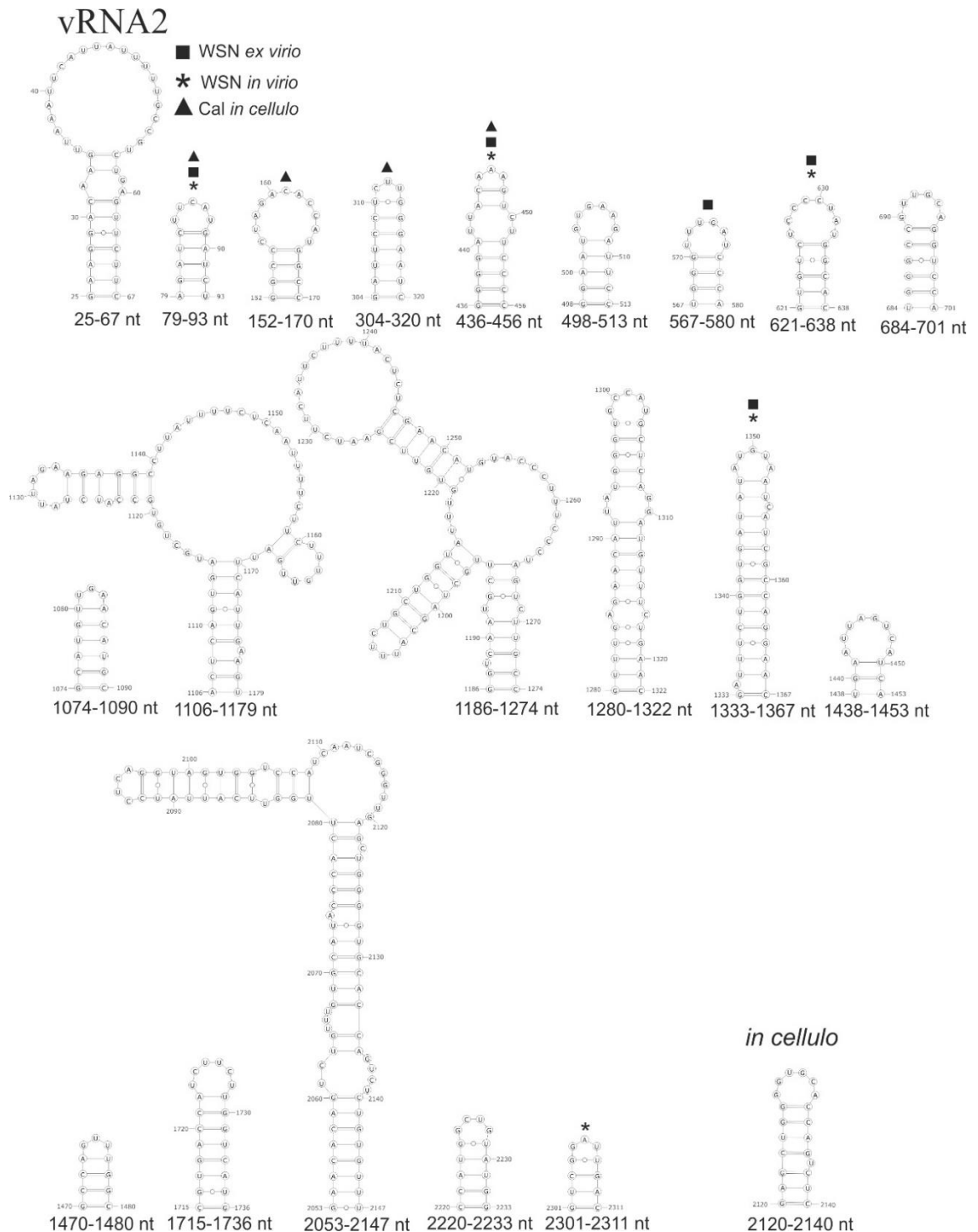
Ryc. 61 Wykresy prawdopodobieństwa parowania zasad (ppz) w obrębie struktur drugorzędowych vRNA segmentów 1,2,3,4 oraz 6. Prawdopodobieństwo obliczone na podstawie wszystkich struktur wygenerowanych w programie RNAstructure na podstawie danych eksperymentalnych. Wizualizacja wykresów została wykonana w programie IGV.

Wspólne dla obu struktur motywy strukturalne o najwyższym ppz zostały potwierdzone jedynie na 5' końcu cząsteczki w regionach: 79-93 nt, 152-170 nt, 304-320 nt oraz 436-456 nt. Co istotne, motywy strukturalne przewidziane dla vRNA2 w warunkach *in cellulo* na 3' końcu różniły się znacznie od tych przewidzianych dla



Ryc. 62 Motywy strukturalne w strukturze drugorzędowej vRNA1 określonej na podstawie mapowania *in virio*. Symbolami zaznaczono motywy zidentyfikowane w strukturze vRNA1 A/California/04/2009 (Cal) *in cellulo*, a także motywy w vRNA1 szczepu A/WSN/1933 (WSN) w warunkach *in virio* oraz *ex virio*.

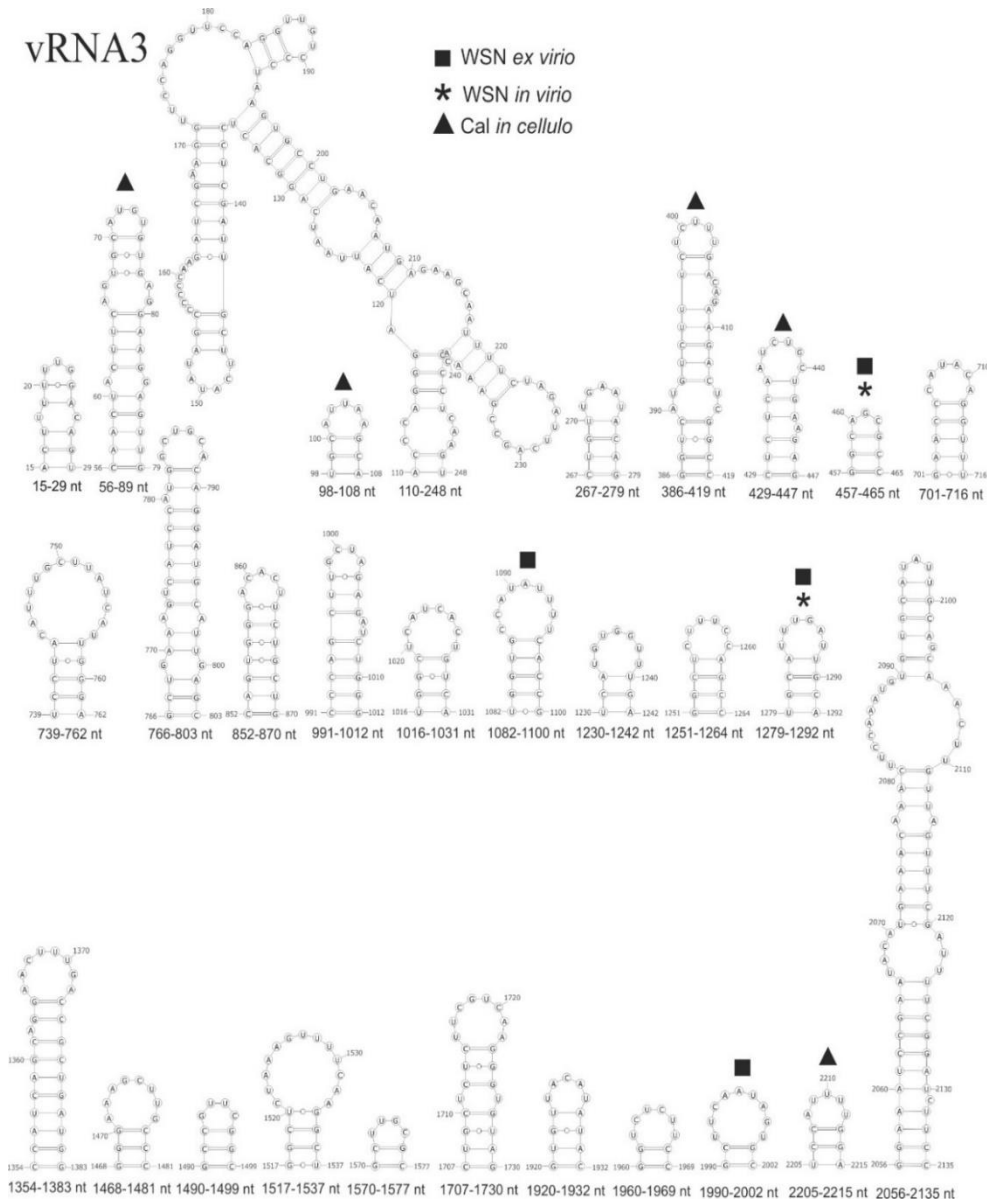
vRNA2 *in virio*. Na końcu 5' vRNA2 *in cellulo* zidentyfikowano motyw w rejonie 2120-2140 nt (>80% ppz), który był unikalny dla tej struktury (**Ryc. 63**). Wspólne dla obu struktur *in virio* motywy strukturalne znajdują się w rejonach: 79-93 nt, 436-456



Ryc. 63 Motywy strukturalne w strukturze drugorzędowej vRNA2 określonej na podstawie mapowania *in virio*. Symbolami zaznaczono motywy zidentyfikowane w strukturze vRNA2 A/California/04/2009 (Cal) *in cellulo*, a także motywy w vRNA1 szczepu A/WSN/1933 (WSN) w warunkach *in virio* oraz *ex virio*.

nt, 621-638 nt, 1333-1367 nt oraz 2301-2311 nt. Część z tych motywów przewidziano również w warunkach *ex virio*. Motyw w rejonie 567-580 nt przewidziano wcześniej jedynie w strukturze *ex virio*. Warto zaznaczyć, że motywy 79-93 nt oraz 436-456 nt są obecne w strukturach *in virio* oraz *in cellulo* szczepu A/California/04/2009, a także w strukturach *in virio* oraz *ex virio* szczepu A/WSN/1933.

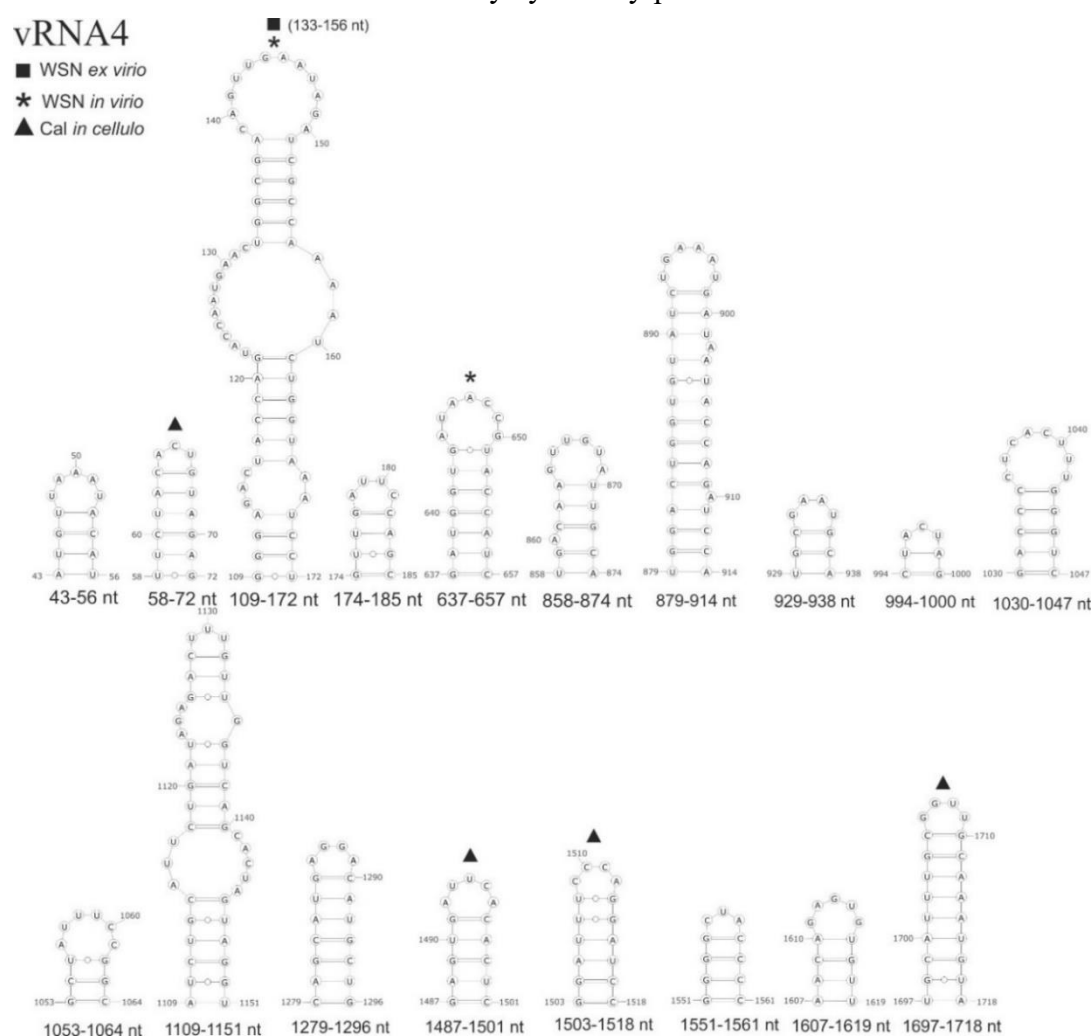
W przypadku segmentu 3 vRNA *in virio* zidentyfikowano 20 motywów o najwyższym ppz, z czego 5 przewidzianych w strukturze *in cellulo* w rejonach: 56-89



Ryc. 64 Motywy strukturalne w strukturze drugorzędowej vRNA3 określonej na podstawie mapowania *in virio*. Symbolami zaznaczono motywy zidentyfikowane w strukturze vRNA3 A/California/04/2009 (Cal) *in cellulo*, a także motywy w vRNA1 szczepu A/WSN/1933 (WSN) w warunkach *in virio* oraz *ex virio*.

nt, 98-108 nt, 386-419 nt, 429-447 nt oraz 2205-2215 nt (**Ryc. 64**). Podobnie jak w przypadku segmentu 2 vRNA, również vRNA3 *in cellulo* wykazywał większe podobieństwo na 5' końcu do struktury *in virio*. Cztery motywy strukturalne zidentyfikowano w strukturze vRNA3 szczepu A/WSN/1933 *ex virio* w rejonach: 457-465 nt, 1082-1100 nt, 1279-1292 nt oraz 1990-2002 nt. Motywy 457-465 nt oraz 1279-1292 nt zostały przewidziane również w strukturze *in virio* tego szczepu.

W strukturze drugorzędowej vRNA4 *in virio* zidentyfikowano 18 motywów o najwyższym ppz (**Ryc. 65**). Cztery motywy były przewidziane w strukturze vRNA4 *in cellulo*, z czego trzy na 3' końcu cząsteczki w rejonach: 58-72 nt, 1487-1501 nt, 1503-1518 nt oraz 1697-1718 nt. Dwa motywy zostały przewidziane w strukturze *ex virio*

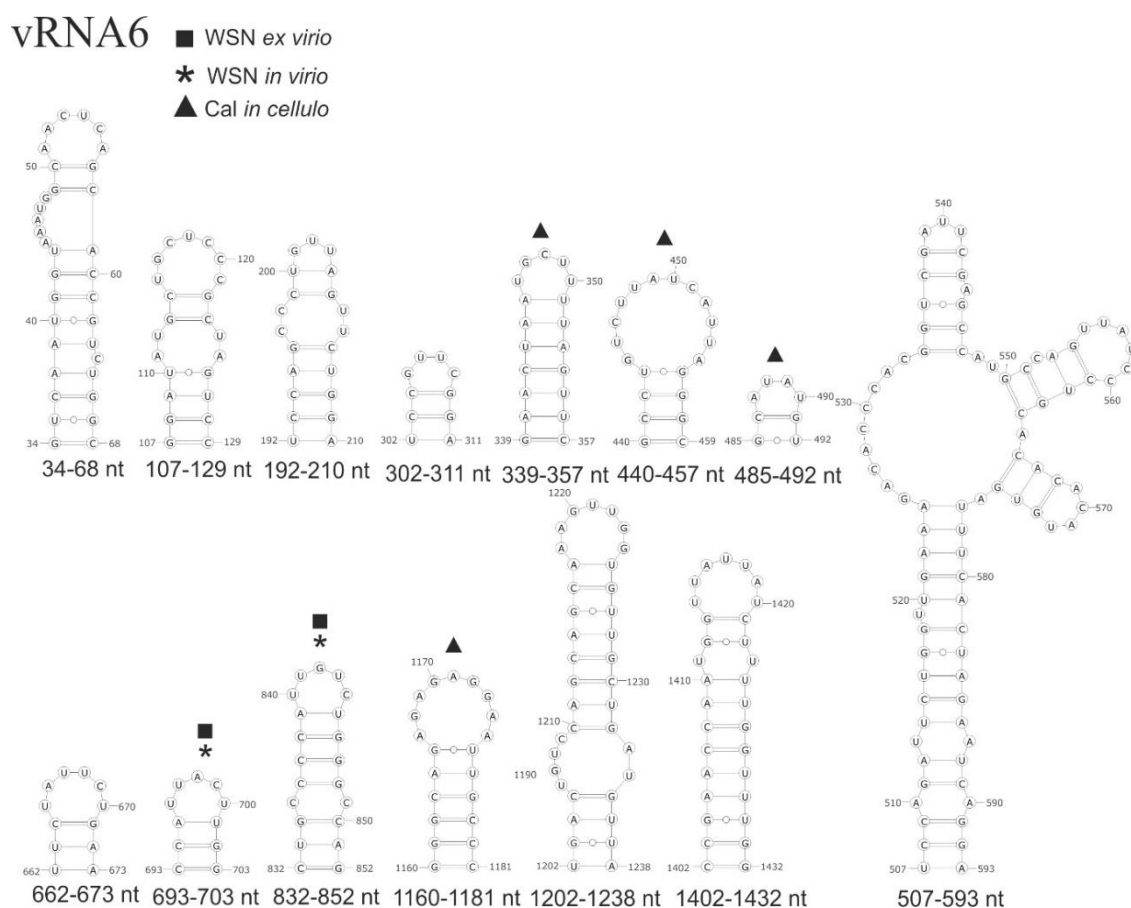


Ryc. 65 Motywy strukturalne w strukturze drugorzędowej vRNA4 określonej na podstawie mapowania *in virio*. Symbolami zaznaczono motywy zidentyfikowane w strukturze vRNA4 A/California/04/2009 (Cal) *in cellulo*, a także motywy w vRNA1 szczepu A/WSN/1933 (WSN) w warunkach *in virio* oraz *ex virio*.

szczepu A/WSN/1933: 109-172 nt oraz 637-657 nt. Pierwszy z motywów był również przewidziany dla struktury *in virio*.

W strukturze drugorzędowej vRNA6 najwyższe ppz miało 13 motywów (**Ryc. 66**). Cztery motywy w rejonach: 339-357 nt, 440-457 nt, 485-492 nt oraz 1160-1181 nt zostały przewidziane *in cellulo*. Jedynie dwa motywy były obecne w strukturach *in virio* oraz *ex virio* vRNA6 szczepu A/WSN/1933: 693-703 nt oraz 832-852 nt.

Zachowawczość motywów strukturalnych pomiędzy różnymi szczepami wirusa grypy w zależy od zmian zachowujących pary zasad na poziomie sekwencji. Mutacje kompensacyjne (dwupunktowe) i konsensusowe (jednopunktowe) dodatkowo potwierdzają istnienie danego motywu strukturalnego. Konserwatywność sekwencji w oczywisty sposób sprzyja konserwatywności motywów strukturalnych. Podobieństwo sekwencyjne pomiędzy szczepami A/California/04/2009 oraz A/WSN/1933 różni się



Ryc. 66 Motywy strukturalne w strukturze drugorzędowej vRNA6 określonej na podstawie mapowania *in virio*. Symbolami zaznaczono motywy zidentyfikowane w strukturze vRNA6 A/California/04/2009 (Cal) *in cellulo*, a także motywy w vRNA6 szczepu A/WSN/1933 (WSN) w warunkach *in virio* oraz *ex virio*.

dla poszczególnych segmentów vRNA. Poniżej w **Tabeli 3** przedstawiono % podobieństwa sekwencji pomiędzy tymi szczepami (obliczone przy pomocy narzędzia BLAST), a także podsumowanie istnienia wspólnych dla tych szczepów motywów strukturalnych przewidzianych *in virio* dla poszczególnych segmentów vRNA.

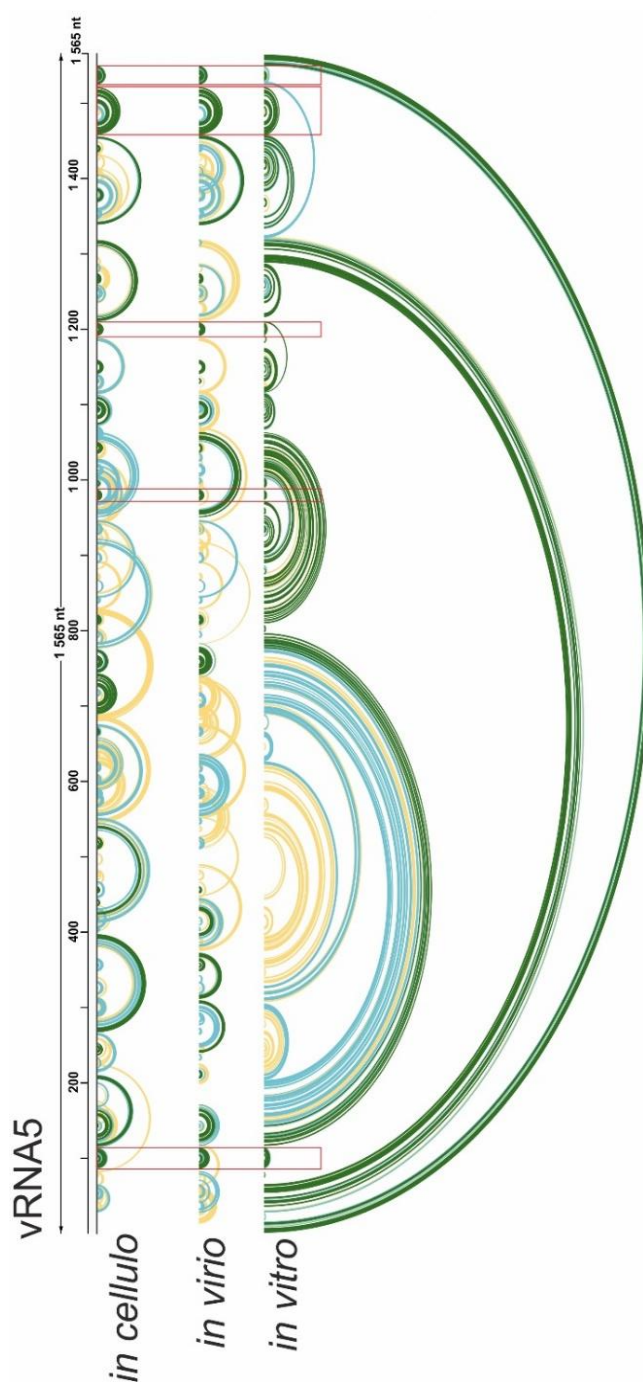
Tabela 3 Podsumowanie wspólnych motywów strukturalnych segmentów 1, 2, 3, 4 i 6 vRNA szczepów A/California/04/2009 oraz A/WSN/1933 przewidzianych w warunkach *in virio*. W tabeli uwzględniono podobieństwo sekwencji (%) pomiędzy tymi szczepami.

segment vRNA	% podobieństwa sekwencji pomiędzy A/California/04/2009 i A/WSN/1933	wspólne motywy <i>in virio</i> [nt]
vRNA1	86%	13-33 nt, 50-63 nt, 1122-1135 nt, 1837-1851 nt
vRNA2	83%	79-93 nt, 436-456 nt, 621-638 nt, 1333-1367 nt, 2301-2311 nt
vRNA3	85%	457-465 nt, 1279-1292 nt
vRNA4	80%	133-156 nt, 637-657 nt
vRNA6	81%	693-703 nt, 832-852 nt

2.3.2. Segment 5 vRNA

Segment 5 vRNA jest jedną z najlepiej poznanych cząsteczek wirusa grypy. W naszym zakładzie zdeteminowano struktury drugorzędowe vRNA5 *in vitro* dwóch szczepów: A/Vietnam/1203/2004 (H5N1) oraz A/California/04/2009 (H1N1)^{12,11}. W niniejszej pracy uzyskano dane eksperymentalne, które umożliwiły zaproponowanie struktur drugorzędowych pełnej długości vRNA5 w warunkach *in virio* oraz *in cellulo* vRNA5 szczepu A/California/04/2009. Warto zaznaczyć, że dotychczas nie opublikowano danych dotyczących struktur *in cellulo* vRNA wirusa grypy. Zaproponowane struktury drugorzędowe vRNA porównano do struktur *in virio* oraz *ex virio* szczepu A/WSN/1933¹¹⁵.

Na **Ryc. 67** zaprezentowany został wykres prawdopodobieństwa parowania zasad dla struktury drugorzędowej segmentu vRNA5 w warunkach *in virio*, *in cellulo* oraz *in vitro*. Na wykresie zaznaczono motywy strukturalne o wysokim ppz wspólne dla wszystkich trzech struktur w rejonach: 87-115 nt, 974-988 nt, 1193-1200 nt, 1460-1522 nt oraz 1527-1550 nt. . Wykresy ppz ukazują różnice w strukturze drugorzędowej vRNA5 w zależności od warunków eksperymentu. Należy również podkreślić, że struktura *in vitro* była przewidziana w ujęciu globalnym, bez ograniczenia w odległości parowania zasad. Struktura drugorzędowa vRNA5 w warunkach

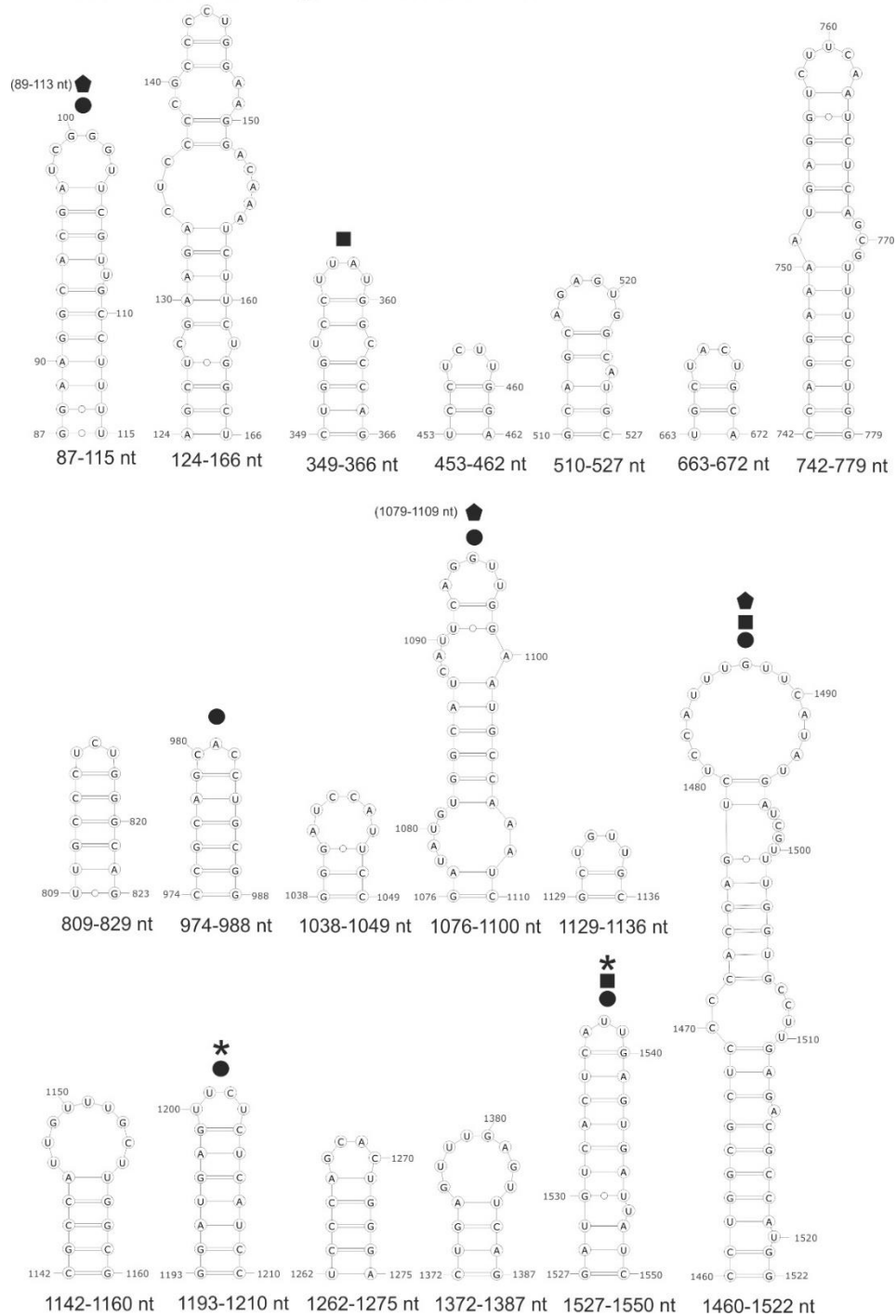


Ryc. 67 Na wykresie przedstawiono prawdopodobieństwa parowania zasad (ppz) dla struktur drugorzędowych vRNA5 uzyskanych w warunkach *in cellulo*, *in virio* oraz *in vitro* szczerpu A/California/04/2009. Na czerwono zaznaczono motywy strukturalne o najwyższym ppz wspólne dla wszystkich trzech struktur.

biologicznych jest w dużej mierze jednociowa, z zachowaniem pojedynczych motywów strukturalnych o wysokim ppz. Obie struktury drugorzędowe vRNA5 w

wspólne motywy vRNA5 *in virio* oraz *in cellulo*

- WSN *ex virio*
- ◆ Vietnam *in vitro*
- * WSN *in virio*
- California *in vitro*

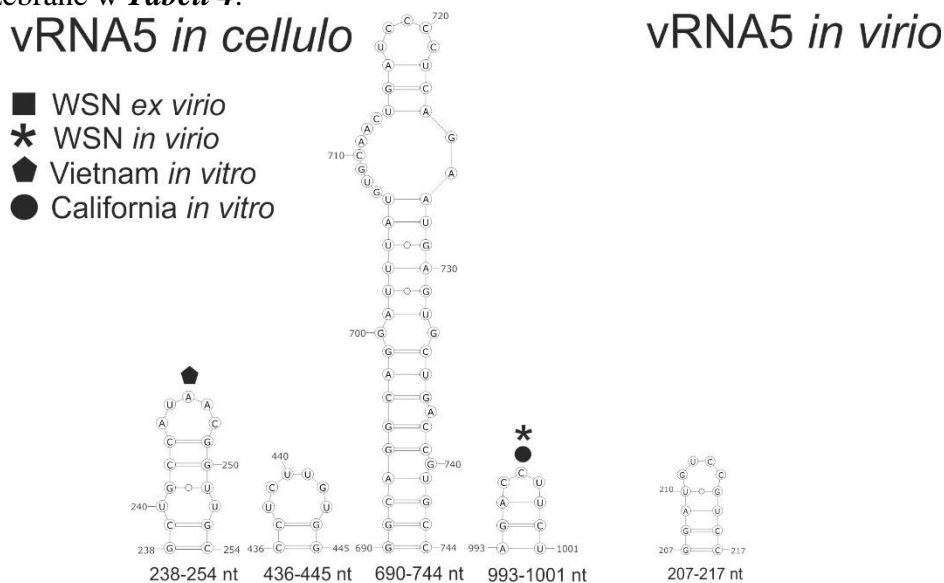


Ryc. 68 Motywy strukturalne vRNA5 szczepu A/California/04/2009 wspólne dla struktur *in virio* oraz *in cellulo*. Na rysunku zaznaczono motywy zidentyfikowane dla struktur *in virio*, *ex virio* szczepu A/WSN/1933 oraz struktur *in vitro* A/California/04/2009 oraz A/Vietnam/1203/2004.

warunkach biologicznych wykazują niższy stopień ustrukturyzowania niż *in vitro*.

Pomimo zaobserwowanych różnic, struktury vRNA5 *in virio* oraz *in cellulo* mają wiele wspólnych motywów strukturalnych o wysokim ppz. Część wspólnych motywów strukturalnych wykazywała również niższe ppz (30-80%). Na **Ryc. 68** przedstawiono motywy strukturalne wspólne dla obu struktur MEA vRNA5. Na rysunku uwzględniono motywy, które miały >80% ppz dla co najmniej jednej ze struktur (*in virio* bądź *in cellulo*).

Kilka motywów strukturalnych było unikalnych dla struktury drugorzędowej vRNA5 *in cellulo* w rejonach: 238-254 nt, 436-445 nt, 690-744 nt oraz 993-1001 nt (**Ryc. 69**). Może to wskazywać na ich powstawanie w komórkach w czasie cyklu replikacyjnego wirusa. Na **Ryc. 69** przedstawiono również jeden motyw vRNA5 w rejonie 207-217 nt, który był unikalny dla struktury *in virio*. Motyw strukturalny vRNA5 charakterystyczny dla wirionu może mieć znaczenie w pakowaniu wirionu. Wartości ppz dla poszczególnych motywów w różnych warunkach mapowania zostały zebrane w **Tabeli 4**.



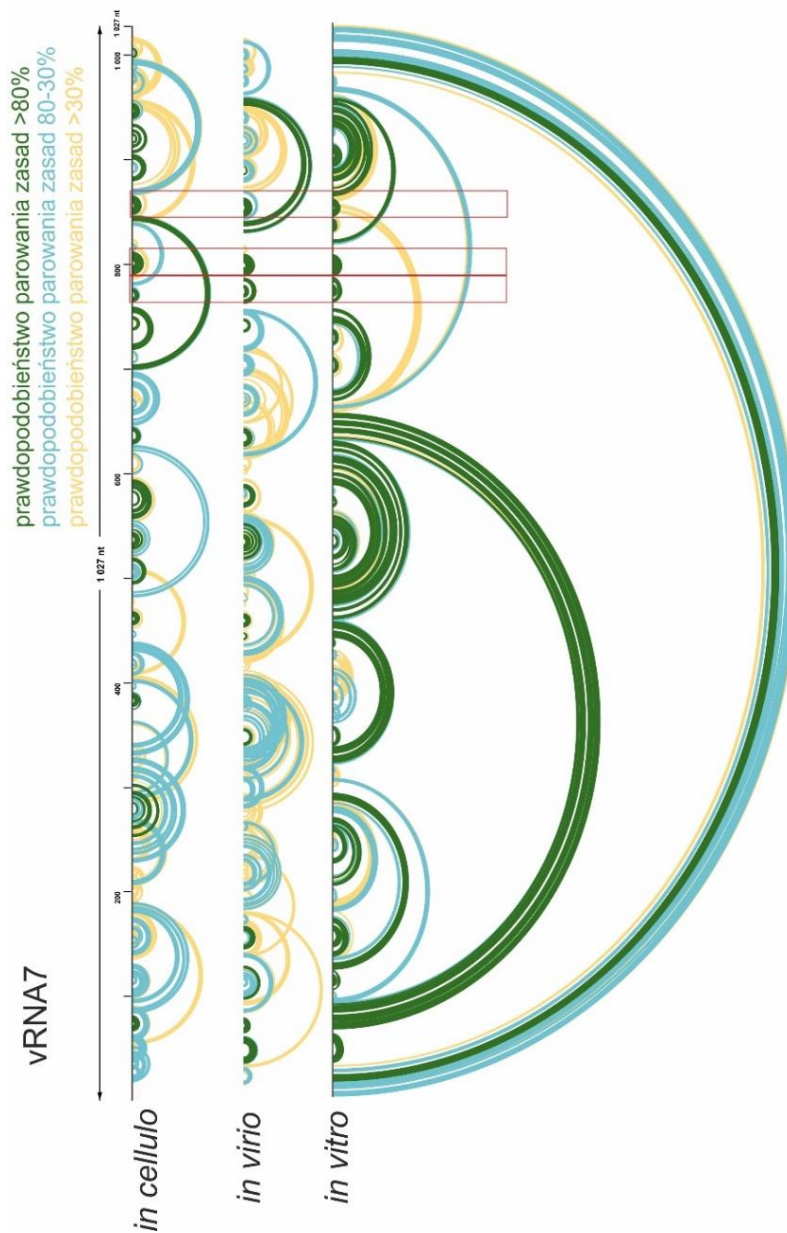
Ryc. 69 Motywy strukturalne unikalne dla struktury drugorzędowej vRNA5 w warunkach *in cellulo* oraz motyw unikalny dla struktury drugorzędowej vRNA5 *in virio*.

Tabela 4 Podsumowanie informacji dotyczących występowania poszczególnych motywów strukturalnych vRNA5 dla szczepów A/California/04/2009 (*in vitro*, *in cellulo*, *in vitro*), A/WSN/1933 (*in vitro*, *ex vitro*) oraz A/Vietnam/1203/2004 (*in vitro*). W tabeli zaznaczono % prawdopodobieństwa parowania zasad.

Nazwa szczepu	A/California/04/2009			A/WSN/1933			A/Vietnam/1203/2004	
	in virio	in cellulo	prawdopodobieństwo parowania zasad in vitro	in vitro	ex vitro	in vitro	in vitro	
86-115	>80 %	>80 %	>80 %	<10%	brak	>80% dla 89-113 nt		
124-166	>80 %; spinka 140-147 nt - 30-80%	>80%	brak	brak	brak	brak	brak	
207-218	>80%	brak	brak	brak	brak	brak	brak	
238-254	brak	>80%	brak	brak	brak	>80%		
349-366	>80%	30-80%	brak	>30-80%	>80%	brak	brak	
436-445	brak	>80%	brak	brak	brak	brak	brak	
453-462	>80%	>80%	brak	brak	brak	brak	brak	
510-527	30-80%	>80%	brak	brak	brak	brak	brak	
663-672	30-80%	>80%	brak	brak	brak	brak	brak	
690-744	brak	>80%	brak	brak	brak	brak	brak	
742-779	>80%	>80% dla 46-776 nt	brak	brak	brak	brak	brak	
810-822	>80%	>80%	brak	brak	brak	brak	brak	
974-988	>80%	>80%	>80%	<10% dla 976-986 nt	30-80%	brak	brak	
993-1001	brak	>80%	>80%	>80%	<10%	brak	brak	
1038-1049	>80%	1035-1051 nt >80%	brak	brak	brak	brak	brak	
1076-1110	>80%; spinka 1094-1099 nt 30-80%	>80%	>80%	brak	brak	>80% dla 1079-1109 nt		
1129-1136	>80%	30-80%	brak	>10%	brak	brak	brak	
1142-1160	>80%	>80%	brak	nt	brak	brak	brak	
1193-1210	>80%	>80%	>80%	>80%	30-80%	brak	brak	
1262-1275	>80%	>80%	brak	30-80%	brak	brak	brak	
1372-1387	30-80%	>80%	brak	brak	30-80%	brak	brak	
1436-1446	30-80%	>80%	brak	brak	brak	brak	brak	
1460-1522	>80%	>80%	>80%	brak	80% (zmiany w parowaniu zasad)	80% (zmiany w parowaniu zasad)	80% (zmiany w parowaniu zasad)	
1527-1550	>80%	>80%	>80%	>80%	>80%	>80%	30-80%	

2.3.3. Segment 7 vRNA

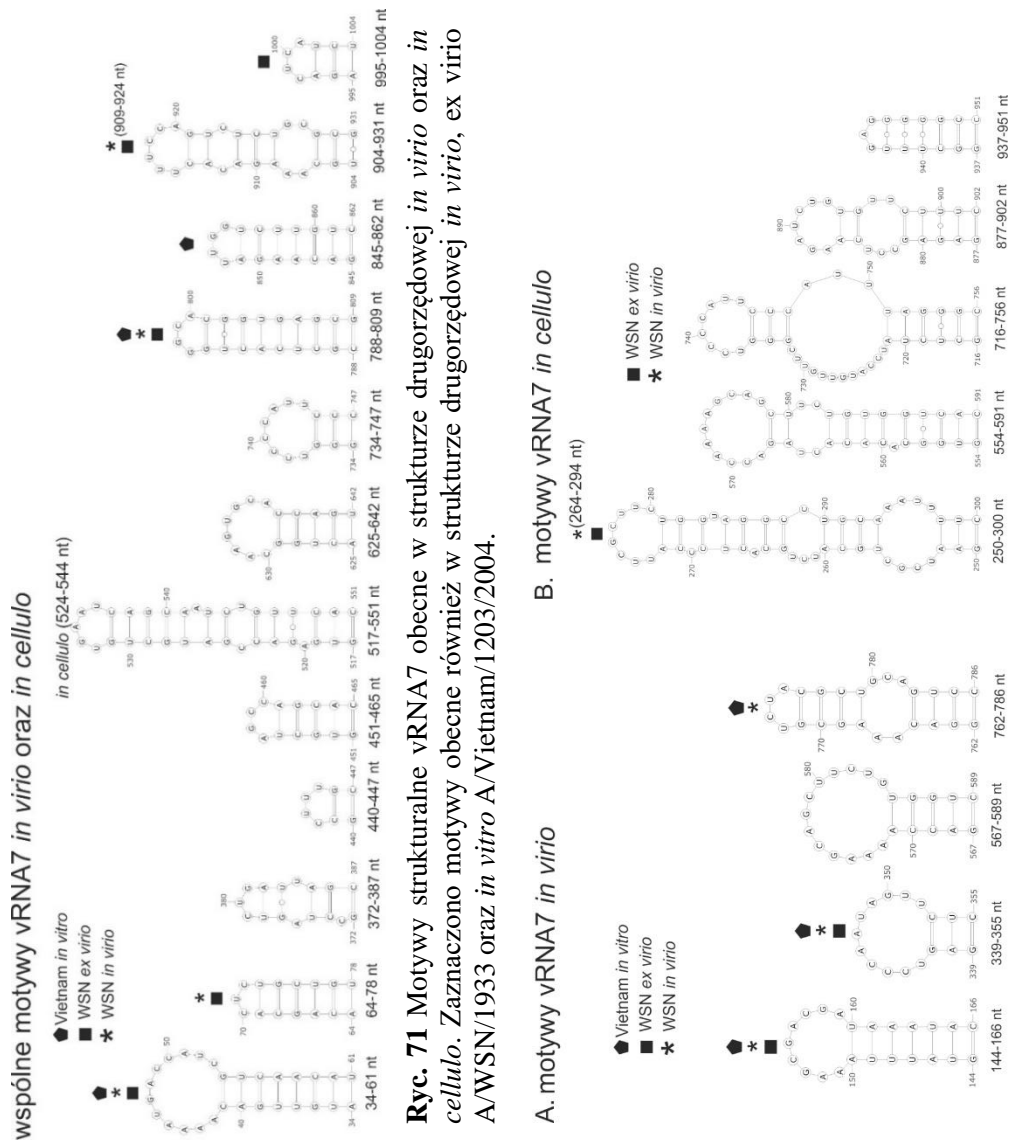
Na podstawie wyników mapowania uzyskano dane strukturalne dotyczące całej długości segmentu vRNA7 mapowanego w warunkach *in virio* oraz *in cellulo*. Struktura drugorzędowa vRNA7 dla szczepu A/California/04/2009 w warunkach *in vitro* jest nieznana. Jednakże, przez naszą grupę zaproponowana została struktura drugorzędowa vRNA7 *in vitro* szczepu A/Vietnam/1203/2004 (H5N1)¹³. Tak więc uzyskane wyniki dla struktur drugorzędowych vRNA7 *in virio* oraz *in cellulo* zostaną porównane do struktury *in vitro* tego szczepu, a także struktur *ex virio* oraz *in virio* szczepu A/WSN/1933 (H1N1)¹¹⁵. Na **Ryc. 70** zaprezentowano wykresy ppz dla



Ryc. 70 Na wykresie przedstawiono prawdopodobieństwa parowania zasad (ppz) dla struktur drugorzędowych vRNA7 uzyskanych w warunkach *in cellulo*, *in virio* szczepu A/California/04/2009 oraz *in vitro* szczepu A/Vietnam/1203/2004. Na czerwono zaznaczono motywy strukturalne o najwyższym ppz wspólne dla wszystkich trzech struktur

struktur drugorzędowych *in virio*, *in cellulo* oraz struktury drugorzędowej *in vitro* A/Vietnam/1203/2004. Struktura *in vitro* była przewidziana w ujęciu globalnym, bez wprowadzania ograniczenia parowania zasad. Tak samo jak w przypadku struktur drugorzędowych vRNA5, porównanie struktur vRNA7 *in virio*, *in cellulo* i *in vitro* wskazało na różnice między nimi z zachowaniem części wspólnych motywów strukturalnych. Podobnie jak dla vRNA5 obie struktury drugorzędowe vRNA7 w warunkach biologicznych są mniej ustrukturyzowane niż *in vitro*. Na rysunku zaznaczono motywy wspólne dla wszystkich trzech struktur vRNA7.

Zidentyfikowano 12 motywów strukturalnych, które były wspólne dla struktury drugorzędowej vRNA7 *in virio* oraz *in cellulo* (**Ryc. 71**). Dwa motywy vRNA7 w rejonach 34-81 nt oraz 788-809 nt były obecne w strukturze drugorzędowej *in virio* oraz *ex virio* A/WSN/1933 oraz *in vitro* A/Vietnam/1203/2004. Dwa motywy vRNA7



Ryc. 72 A. Motywy strukturalne unikalne dla vRNA7 *in virio* oraz *in cellulo*. Zaznaczono motywy, które zidentyfikowano w strukturach A/WSN/1933 oraz A/Vietnam/1203/2004

w rejonach 64-78 nt oraz 904-931 nt były obecne w obu zaproponowanych strukturach drugorzędowych A/WSN/1933, natomiast motyw 995-1004 nt jedynie w strukturze *ex virio* tego szczepu.

Część motywów strukturalnych vRNA7 o wysokim ppz (>80%) było unikalnych dla obu struktur. W przypadku vRNA7 *in virio* zidentyfikowano 4 motywy strukturalne nieobecne w strukturze *in cellulo*: 144-166 nt, 339-355 nt, 567-589 nt oraz 782-788 nt (**Ryc. 72A**). Co ciekawe, trzy z motywów przewidziano również w strukturze drugorzędowej *in virio* szczepu A/WSN/1933 oraz w strukturze drugorzędowej *in vitro* A/Vietnam/1203/2004. Może to wskazywać na rearanżację strukturalną w czasie cyklu replikacyjnego w komórce. Pięć motywów strukturalnych było unikalnych dla struktury drugorzędowej *in cellulo*, w rejonach: 250-300 nt, 554-591 nt, 716-756 nt, 877-902 nt oraz 937-951 nt (**Ryc. 72 B**). Fragment pierwszego motywu (w rejonie 264-294 nt) został przewidziana z dużym ppz w strukturach A/WSN/1933. Zdeterminowane w tej pracy doktorskiej nowe motywy strukturalne vRNA7 *in cellulo* mogą się tworzyć w czasie cyklu replikacyjnego wirusa i być miejscem oddziaływań vRNA7 z biomolekułami obecnymi w komórce. Podsumowanie wszystkich motywów strukturalnych vRNA7 charakterystycznych dla różnych warunków biologicznych wraz z wartościami ppz zostały zebrane w **Tabeli 5**.

Tabela 5 Podsumowanie informacji dotyczących występowania poszczególnych motywów strukturalnych vRNA7 dla szczepów A/California/04/2009 (*in virio*, *in cellulo*), A/WSN/1933 (*in virio*, *ex virio*) oraz A/Vietnam/1203/2004 (*in vitro*). W tabeli zaznaczono % prawdopodobieństwa parowania zasad.

Nazwa szczepu	A/California/04/2009		A/WSN/1933		A/Vietnam/1203/2004
	in virio	in cellulo	in virio	ex virio	
region [nt]			prawdopodobieństwo parowania zasad		in vitro
34-61	>80%	30-80%	>80%	>80%	>80%
64-78	>80%	>80%	>80%	>80%	brak
144-166	>80%	brak	>80%	>80%	>80%
250-300	brak	>80%	>80% dla 264-294 nt	>80%	brak
339-355	>80%	brak	>80%	>80%	>80%
372-387	30-80%	>80%	brak	brak	brak
440-447	>80%	30-80%	brak	brak	brak
451-465	>80%	>80%	brak	brak	brak
517-551	>80%	>80% dla 524-544 nt	brak	brak	brak
554-591	brak	>80%	brak	brak	brak
567-589	>80%	brak	brak	brak	brak
625-642	>80%	>80%	brak	brak	brak
716-756	brak	>80%	brak	brak	brak
734-747	>80%	>80%	brak	brak	>10%
762-786	>80%	brak	>80%	brak	>80%
761-774	brak	>80%	brak	brak	brak
788-809	>80%	>80%	>80%	>80%	>80%
845-862	>80%	>80%	brak	brak	>80%
877-902	brak	>80%	brak	brak	brak
904-931	30-80%	>80%	>80% dla 909-924 nt	30-80% dla 909-924 nt	brak
937-951	brak	>80%	brak	brak	brak
995-1004	30-80%	>80%	brak	30-80%	brak

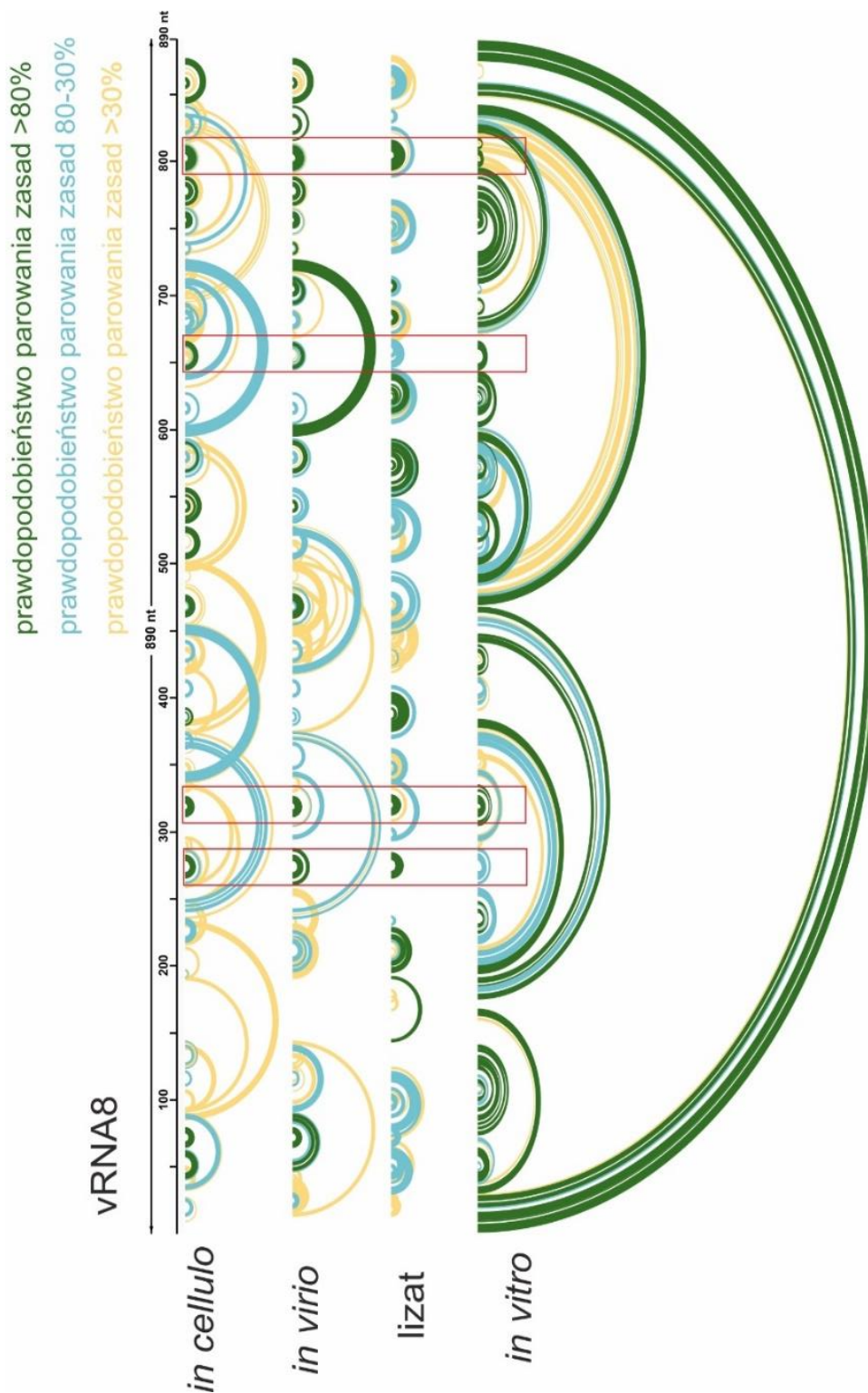
2.3.4. Segment 8 vRNA

Uzyskane dane eksperymentalne z sekwencjonowania NGS umożliwiły zaproponowanie struktury drugorzędowej pełnej długości vRNA8 szczepu A/California/04/2009 w warunkach *in virio* oraz *in cellulo*. Segment 8 vRNA jest najlepiej poznaną cząsteczką genomowego RNA wirusa grypy typu A. W pierwszej części niniejszej pracy określona została struktura drugorzędowa vRNA8 szczepu A/California/04/2009 w lizacie komórkowym. W naszej grupie zaproponowano również struktury drugorzędowe vRNA8 *in vitro* (Marta-Soszyńska Józwiak et. al., manuskrypt w recenzji), *in silico* szczepu A/California/04/2009 oraz strukturę *in vitro* szczepu A/Vietnam/1203/2004^{10,14}. Struktury drugorzędowe vRNA8 zostały również porównane do struktur vRNA8 *in vitro*, *ex virio* oraz *in virio* szczepu A/WSN/1933¹¹⁵.

Na **Ryc. 73** przedstawione zostały wykresy prawdopodobieństwa parowania zasad dla struktur drugorzędowych vRNA8 *in virio*, *in cellulo* oraz *in vitro* (przewidziana bez dystansu parowania zasad) a także w lizacie komórkowym. Podobnie jak w przypadku vRNA5 i vRNA7 wszystkie struktury drugorzędowe vRNA8 zdeterminowane w warunkach biologicznych wykazują niższy stopień ustrukturyzowania niż *in vitro*. Na rysunku zaznaczono motywy strukturalne o wysokim ppz wspólne dla wszystkich czterech struktur.

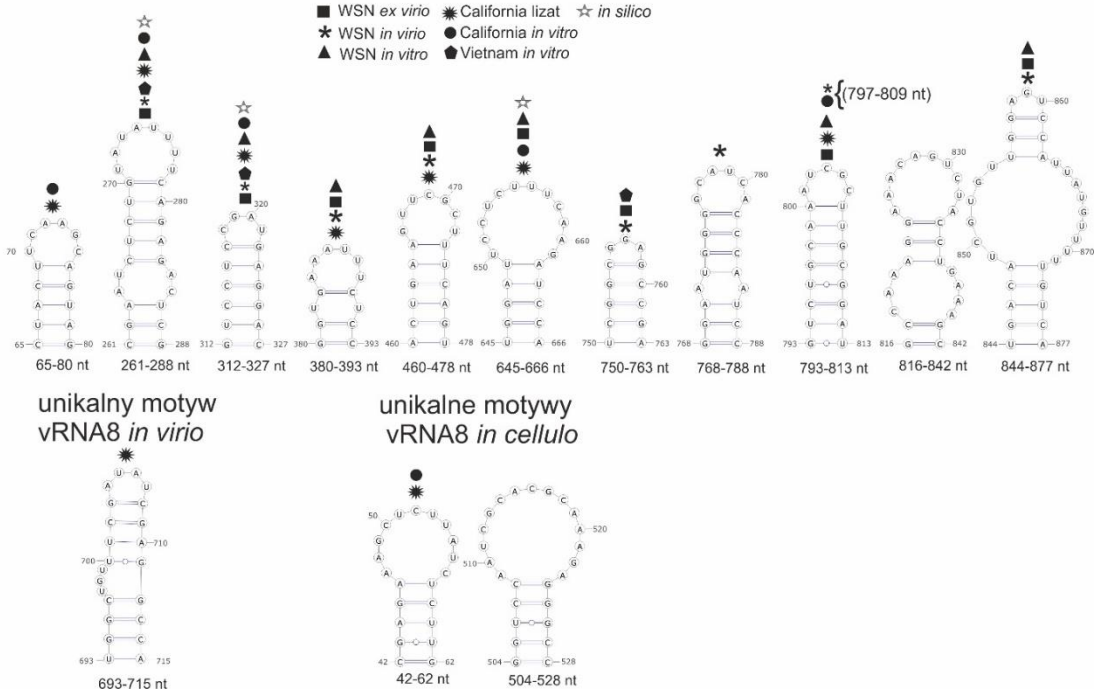
Analiza podobieństwa motywów strukturalnych vRNA8 wykazała, że 11 motywów strukturalnych było wspólnych dla struktur drugorzędowych *in cellulo* oraz *in virio* (**Ryc. 74**). Co istotne, trzy motywy strukturalne w rejonach: 261-288 nt, 312-327 nt, 794-813 nt (za wyjątkiem *in vitro* A/Vietnam/1203/2004) były wspólne dla wszystkich analizowanych struktur. Dwa pierwsze motywy zostały również przewidziane w strukturze *in silico*. To wskazuje na ich zachowawczość wśród różnych szczepów wirusa grypy typu A. Siedem motywów strukturalnych vRNA8 przewidzianych *in virio* i *in cellulo* zostało również zidentyfikowanych w warunkach lizatu komórkowego. Co ciekawe, jedna struktura spinkowa unikalna dla warunków *in virio* (693-715 nt) oraz jedna w przypadku *in cellulo* (42-62 nt) były przewidziane w strukturze vRNA8 w lizacie komórkowym (**Ryc. 74**). Jeden motyw, w rejonie 504-528 nt był obecny jedynie w strukturze *in cellulo*. W **Tabeli 6** zebrano podsumowanie

występowania motywów strukturalnych vRNA8 dla różnych warunków wraz z wartościami ppz.



Ryc. 73 Na wykresie przedstawiono prawdopodobieństwa parowania zasad (ppz) dla struktur drugorzędowych vRNA8 szczepu A/California/04/2009 uzyskanych w warunkach *in cellulo*, *in virio*, *in vitro* oraz w lizacie komórkowym. Na czerwono zaznaczono motywy wspólne dla wszystkich struktur.

wspólne motywy vRNA8 *in virio* oraz *in cellulo*



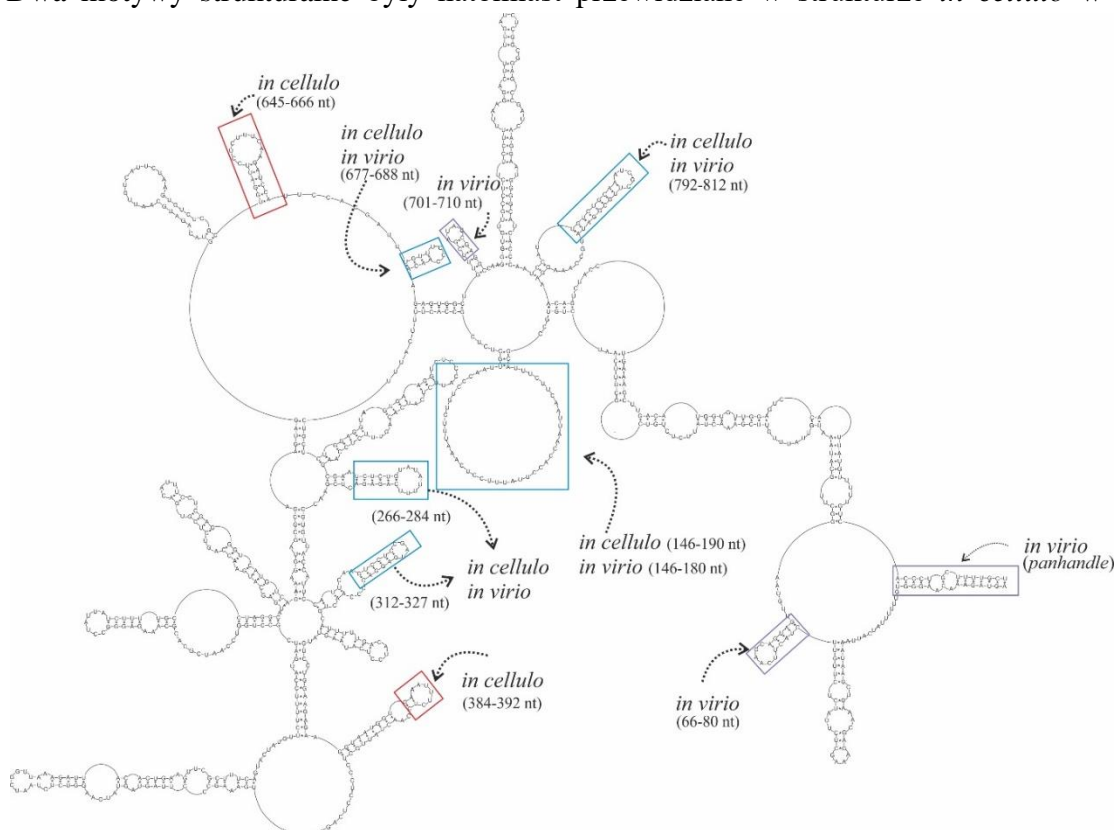
Ryc. 74 Motywy strukturalne vRNA8 przewidziane z dużym ppz dla struktur drugorzędowych *in virio* oraz *in cellulo*. Zaznaczono motywy strukturalne przewidziane w strukturach *in vitro*, *in silico* i w lizacie komórkowym dla szczepu A/California/04/2009 oraz *in vitro*, *ex virio* i *in virio* szczepu A/WSN/1933, a także *in vitro* A/Vietnam/1203/2004.

Tabela 6 Podsumowanie informacji dotyczących występowania poszczególnych motywów strukturalnych vRNA 8 dla szczepów A/California/04/2009 (*in vitro*, *in cellulo*, *in vitro*, w lizacie), A/WSN/1933 (*in vitro*, *ex virio*) oraz A/Vietnam/1203/2004 (*in vitro*). W tabeli zaznaczono % prawdopodobieństwa parowania zasad.

Nazwa szczepu	A/California/04/2009				A/WSN/1933				A/Vietnam/1204/2004	
	in virio	in cellulo	lizat	in vitro	in virio	ex virio	in vitro	in vitro	in vitro	in vitro
region [nt]										
42-62	brak	>80%	30-80%	>80%	brak	brak	brak	brak	brak	brak
65-80	>80%	>80%	30-80%	brak	brak	brak	brak	brak	brak	brak
261-288	>80%	>80%	>80%	30-80%	>80%	>80%	>80%	>80%	>80%	>80%
312-327	>80%	>80%	>80%	>80%	>80%	>80%	>80%	>80%	>80%	>80%
380-393	30-80%	>80%	>80%	brak	>80%	>80%	>80%	>80%	>80%	brak
460-478	>80%	>80%	30-80%	brak	>80%	>80%	>80%	>80%	>80%	brak
504-528	brak	>80%	brak	brak	brak	brak	brak	brak	brak	brak
645-666	>80%	>80%	30-80%	>80%	brak	>80%	>80%	>80%	>80%	brak
693-715	>80%	brak	30-80%	brak	brak	brak	brak	brak	brak	brak
750-763	>80%	>80%	brak	brak	>80%	30-80%	brak	brak	brak	>80% (736-748 nt)
768-788	>80%	>80%	brak	brak	30-80%	brak	brak	brak	brak	brak
794-813	>80%	>80%	>80%	>80% dla 797-809 nt	30-80%; >80% dla 797-809 nt	30-80%	>80%	>80%	>80%	brak
816-842	>80%	30-80%	brak	brak	brak	brak	brak	brak	brak	brak
844-877	>80%	>80%	brak	brak	>80%	>80%	>80%	>80%	>80%	brak

2.4. Porównanie struktury MFE vRNA8 tworzącej się w różnych warunkach biologicznych

W celu porównania struktur drugorzędowych vRNA8 *in virio* oraz *in cellulo* ze strukturą w lizacie komórkowym wygenerowano (RNAstructure) struktury o minimalnej wartości energii swobodnej (MFE) bez ograniczenia parowania zasad (**Ryc. 85, Ryc. 86, Dodatek**). W strukturze vRNA8 fałdowanej w lizacie komórkowym zidentyfikowano lokalne motywy strukturalne, które były obecne w strukturach *in virio* oraz *in cellulo*. Wspólne dla wszystkich trzech struktur były motywy w rejonach: 266-284 nt, 312-327 nt, 677-688 nt oraz 794-812 nt. Motywy te, za wyjątkiem (677-688 nt) były przewidziane w wszystkich strukturach MEA vRNA8 w ujęciu lokalnym, co świadczy o ich dużej stabilności termodynamicznej. We wszystkich strukturach zidentyfikowano również długi fragment jednoniciowy w rejonie 146-190 nt (146-180 nt w przypadku *in virio*). Dwa motywy strukturalne vRNA8 w lizacie komórkowym były przewidziane jedynie w strukturze *in virio* w rejonach: 66-80 nt oraz 701-710 nt. Dwa motywy strukturalne były natomiast przewidziane w strukturze *in cellulo* w

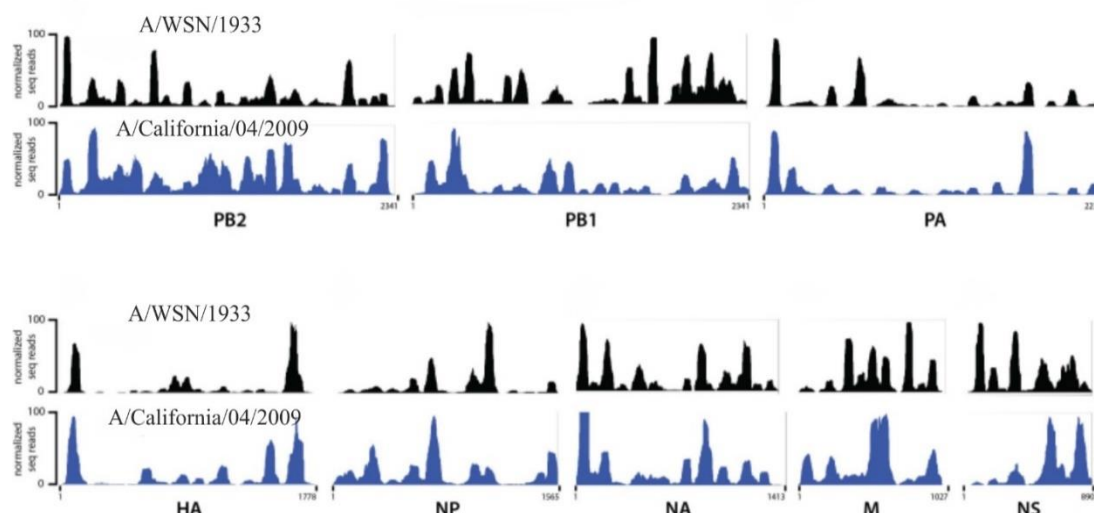


Ryc. 75 Struktura drugorzędowa o najniższej wartości energii swobodnej (MFE) vRNA8 w lizacie komórkowym. Na strukturze vRNA8 zaznaczono motywy strukturalne przewidziane również w strukturach MFE cząsteczki vRNA8 w warunkach *in virio* oraz *in cellulo*.

rejonach: 384-392 nt i 645-666 nt. Wszystkie powyższe motywy zostały zaznaczone na **Ryc. 75**. Wskazuje to na to, że struktura drugorzędowa vRNA8 w lizacie stanowiła pewną wypadkową dla obu struktur. Struktury MFE vRNA8 w różnych warunkach biologicznych różniły się znacznie w ujęciu globalnym, co wynikało przede wszystkim z różnic w parowaniu dalekiego zasięgu. Różnice te ukazują możliwość dużej labilności vRNA8 w różnych warunkach biologicznych, a także, z drugiej strony, z ograniczeń programów do przewidywania struktur. Co istotne, motyw promotora *panhandle* został przewidziany w strukturze drugorzędowej *in vitro* vRNA8. Jest to zgodne z potwierdzoną cykliczną strukturą vRNA w kompleksie vRNP. Natomiast w strukturze drugorzędowej vRNA8 *in cellulo* zaobserwowano w tym rejonie rearanżacje strukturalne, co potwierdza labilność tego motywu w czasie cyklu replikacyjnego.

2.5. Występowanie motywów strukturalnych vRNA o wysokim prawdopodobieństwie *in vitro* a profil wiązania białka NP

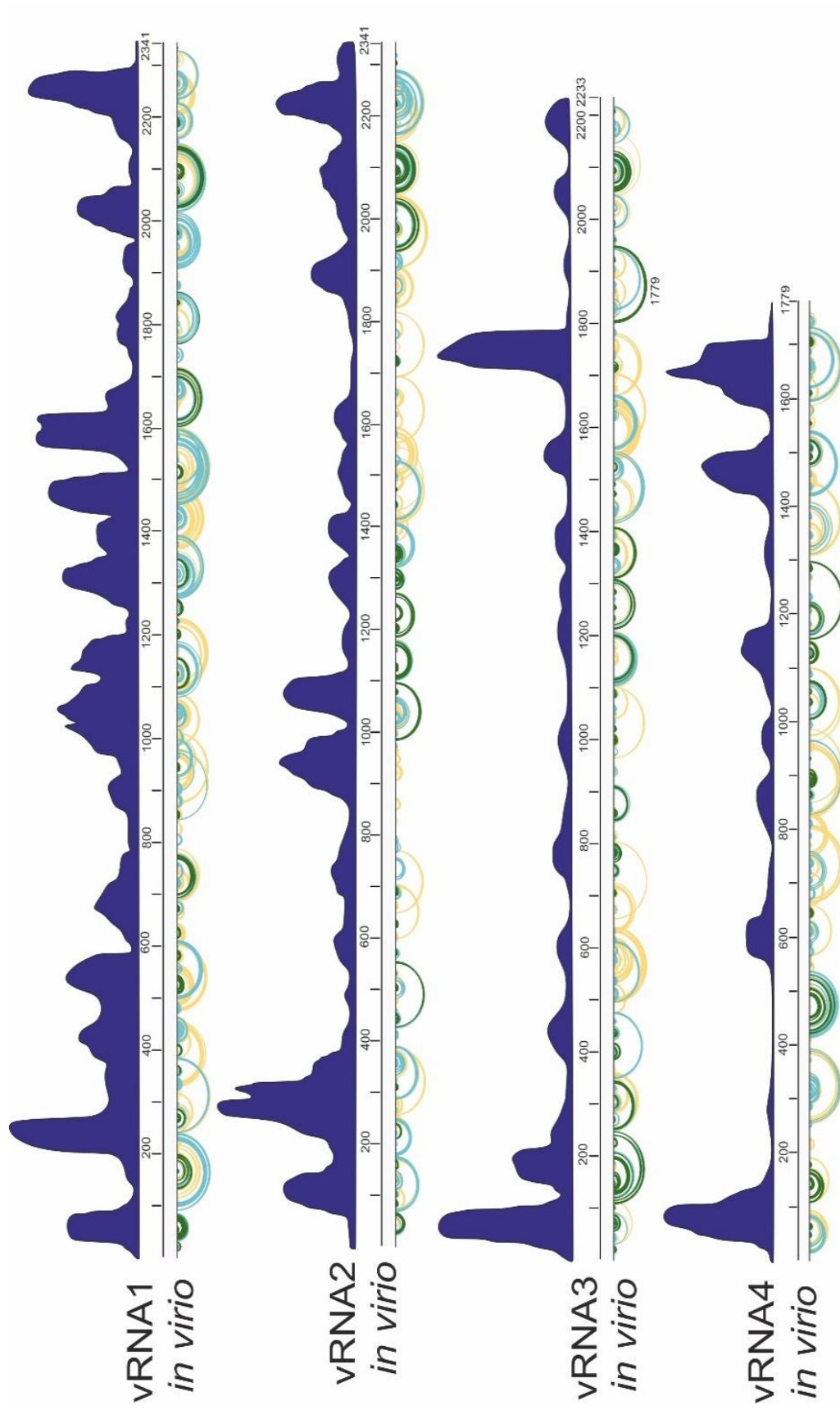
Badania dotyczące profilu wiązania białka NP do cząsteczki vRNA w kompleksie vRNP wykazały, że wiązanie to jest specyficzne dla danego szczepu IAV, chociaż pojawiają się pewne podobieństwa⁴⁶. Dwa opisane szczepy A/California/04/2009 oraz A/WSN/1933 tego samego podtypu wirusa (H1N1) wykazywały generalnie odmienny profil wiązania jednak można zauważyć rejony o podobnym zaangażowaniu w wiązanie białka NP (**Ryc. 76**). Co ciekawe, porównanie korelacji wiązania białka NP



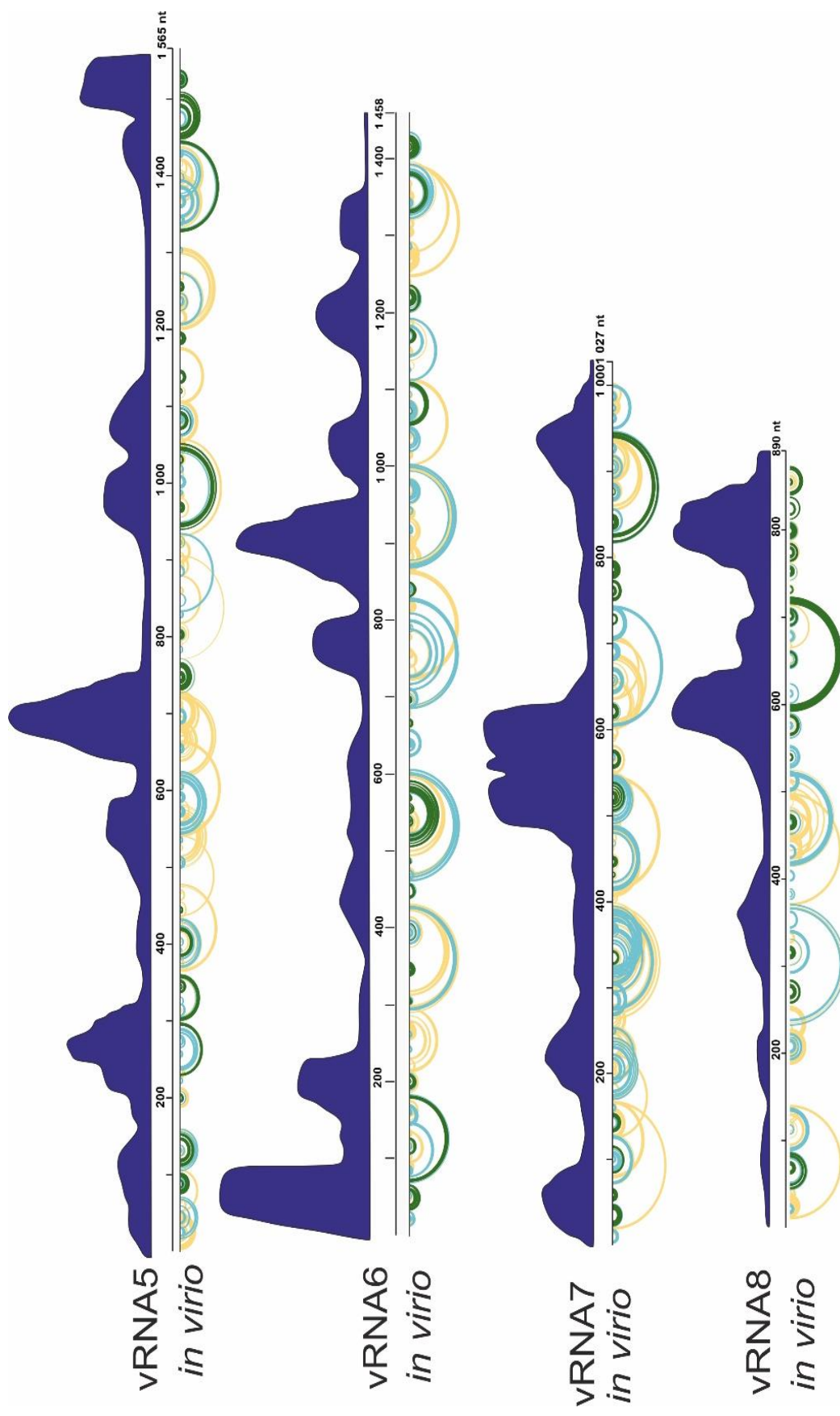
Ryc. 76 Profil wiązania białka NP dla poszczególnych segmentów vRNA dla dwóch szczepów podtypu H1N1: A/WSN/1933 oraz A/California/04/2009. Rysunek pochodzi z publikacji Lee et al. (Nucleic Acid Research, 2017).

do cząsteczki vRNA z występowaniem motywów strukturalnych w warunkach *in virio* wykazało brak zależności w przypadku szczepu A/WSN/1933¹¹⁵. Autorzy badań wskazali, że przewidziane z wysokim prawdopodobieństwem motywy strukturalne vRNA zlokalizowane były zarówno w rejonach wskazanych jako silnie, jak i słabo wiązane przez białko NP. Autorzy ww. badań doszli do wniosku, że występowanie motywów strukturalnych w większej mierze zależne jest od sekwencji vRNA, niż od wiązania białka NP w kompleksie vRNP.

Dysponując danymi dotyczącymi struktur drugorzędowych vRNA w warunkach *in virio* szczepu A/California/04/2009 oraz profilem wiązania białka w kompleksie vRNP tego szczepu, możliwe były określenie zależności pomiędzy występowaniem motywów strukturalnych a wiązaniem białka NP. Porównanie profilu wiązania białka NP z występowaniem motywów strukturalnych vRNA w kompleksie vRNP wskazało na brak zależności (**Ryc. 77-78**). Występowanie zachowawczych motywów strukturalnych, wspólnych dla obu analizowanych szczepów jest niezależne od wiązania białka NP do cząsteczki vRNA.



Ryc. 77 Porównanie profilu wiązania białka NP do cząsteczki vRNA w kompleksie vRNP dla segmentów 1-4 szczepu A/California/04/2009. Na podstawie Lee et al. (Nucleic Acid Research, 2017)



Ryc. 78 Porównanie profilu wiązania białka NP do cząsteczki vRNA w kompleksie vRNP ze strukturą drugorzędową vRNA dla segmentów 5-8 szczepu A/California/04/2009. Profil wiązania białka NP przedstawiony na podstawie Lee et al. (Nucleic Acid Research, 2017).

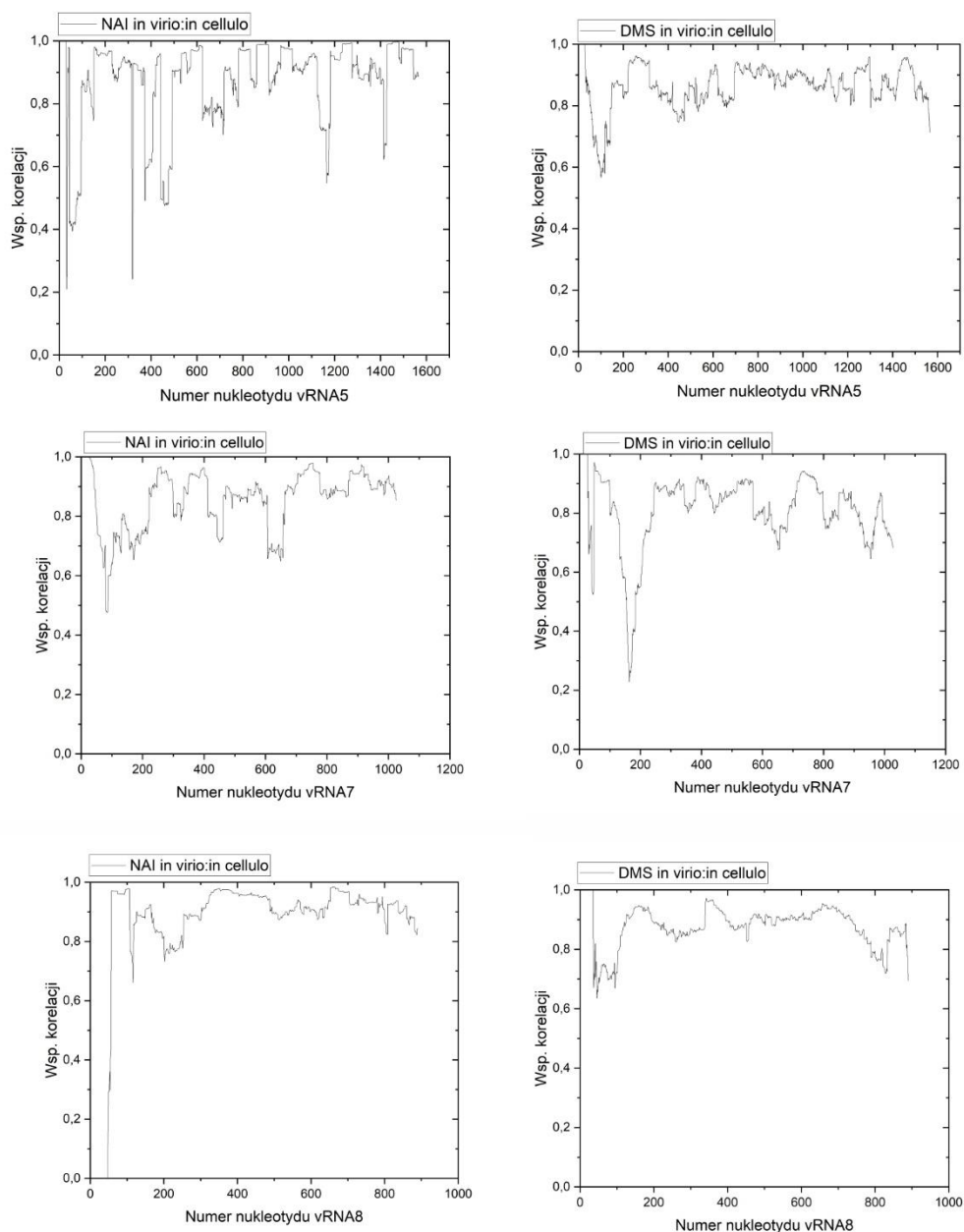
2.6. Analiza korelacji wyników mapowania struktur drugorzędowych vRNA w warunkach *in virio* oraz *in cellulo*.

W celu zbadania korelacji pomiędzy wynikami reaktywności uzyskanymi dla struktur drugorzędowych mapowanych w warunkach *in virio* oraz *in cellulo* wyliczony został współczynnik korelacji Pearson'a. Do obliczeń korelacji wybrano jedynie te segmenty vRNA dla których uzyskano reaktywności dla pełnej długości cząsteczek, tj. segmentów 5, 7 oraz 8 w obu warunkach mapowania. Wartości korelacji są wskazane w **Tabeli 7**. Obliczone współczynniki korelacji wskazały na bardzo wysoką

Tabela 7 Współczynnik korelacji Pearson'a dla reaktywności nukleotydów uzyskanych z mapowania odczynnikami chemicznymi NAI i DMS dla segmentów 5, 7 oraz 8 vRNA mapowanych w warunkach *in virio* oraz *in cellulo*.

	vRNA5	vRNA7	vRNA8
DMS	0,8142	0,79739	0,85368
NAI	0,90046	0,79705	0,90701

(>0,9) korelację reaktywności w przypadku odczynnika NAI dla segmentów vRNA5 oraz vRNA8. Natomiast wysoką wartość korelacji (0,7-0,9) uzyskano dla segmentu vRNA7 (reaktywności dla odczynnika NAI i DMS) oraz w przypadku reaktywności względem odczynnika DMS dla segmentów 5 i 8 vRNA. Wysoka wartość korelacji dla reaktywności nukleotydów (DMS i NAI) wskazuje, że vRNA w warunkach *in virio* oraz *in cellulo* może fałdować się w bardzo podobny sposób. W celu dokładnego zbadania miejsc, w których może dochodzić do zmian strukturalnych pomiędzy strukturami drugorzędowymi RNA w warunkach *in virio* oraz *in cellulo* wyliczono współczynnik korelacji kroczącej (ang. Sliding window correlation). Korelacja krocząca uwzględnienia dynamikę zmian pomiędzy obiema zmiennymi. Na **Ryc. 79** zaprezentowane są wykresy korelacji kroczącej między reaktywnościami uzyskanymi w eksperymencie mapowania struktury *in virio* oraz *in cellulo* dla segmentów vRNA5, 7 i 8. Analiza korelacji dla reaktywności wykazała, że struktury drugorzędowe w warunkach *in virio* mogą być w dużej mierze zachowane w warunkach *in cellulo*. Jednakże zaobserwowano również rejony, w których wartość korelacji spada. W rejonach o niższej wartości korelacji kroczącej analizowane struktury drugorzędowe vRNA mogą się różnić. Analizując miejsca o obniżonej korelacji brano pod uwagę te, w których współczynnik spadał poniżej wartości 0,7. W przypadku segmentu vRNA5



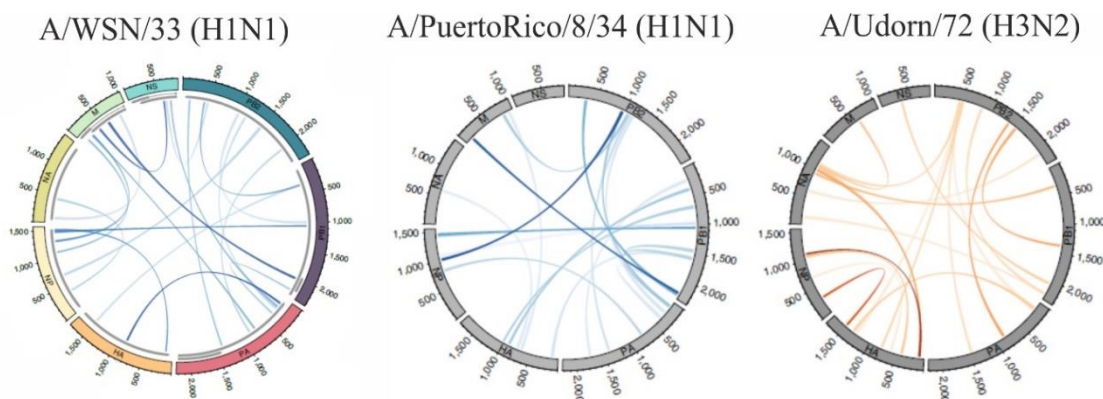
Ryc. 79 Wykresy korelacji kroczącej dla reaktywności SHAPE i DMS poszczególnych nukleotydów obliczone dla cząsteczek vRNA5, 7 i 8 mapowanych w warunkach *in virio* oraz *in cellulo*. Wielkość okna ustalono na 50.

największe obniżenie korelacji zaobserwowano w rejonach: 47-88 nt, 317-320 nt, 374-407 nt, 442-492 nt, 1163-1178 nt oraz 1415-1427 nt (NAI), a także w rejonach 90-116 nt oraz 129-136 nt (DMS). W przypadku segmentu vRNA7 spadek korelacji był widoczny w rejonach: 63-106 nt, 126-130 nt, 157-176 nt i 607- 657 nt (NAI), natomiast w przypadku odczynnika DMS w rejonach: 31-35 nt, 40-47 nt, 136-209 nt, 649-657 nt 936-957 nt. W przypadku vRNA8 korelacja dla reaktywności *in virio* oraz *in cellulo* była bardzo wysoka dla prawie całej długości cząsteczki. Obniżenie korelacji

zaobserwowano dla reaktywności dla obu odczynników jedynie w rejonach: 1-80 nt, 109-118 nt oraz 201-244 nt.

Jak wspomniano, obniżenie współczynnika korelacji wskazuje na prawdopodobną zmianę w strukturze drugorzędowej vRNA pomiędzy warunkami *in virio* oraz *in cellulo*. Różnice w strukturze mogą wynikać między innymi z interakcji między segmentami vRNA, które tworzą się w warunkach *in virio* pomiędzy kompleksami vRNP. Przeprowadzone przez *Dadonaite et al.* badanie interakcji między segmentami vRNA różnych szczepów IAV potwierdziły, że segmenty wchodzą w liczne interakcje vRNA-vRNA (**Ryc. 80**)¹¹⁵. Co istotne, autorzy pracy wskazali, że oddziaływania między segmentami vRNA nie ograniczają się jedynie do sygnałów pakowania. Wzór interakcji vRNA-vRNA jest unikalny dla poszczególnych szczepów IAV, chociaż część oddziaływań jest wspólnych dla wszystkich szczepów. Dodatkowo, niektóre segmenty vRNA mogą wchodzić w bardzo liczne interakcje w miejscach nazywanych „hotspots”. Na przykład w przypadku szczepu A/WSN/1933 cząsteczki vRNA5 i vRNA7 tworzą liczne oddziaływania między segmentami, a jednocześnie niektóre segmenty (vRNA6, vRNA8) tworzą ich bardzo niewiele (**Ryc. 80**).

Zaprezentowany na **Ryc. 79** wykres wartości korelacji reaktywności dla segmentu vRNA5 oraz vRNA7 szczepu analizowanego w niniejszej pracy (A/California/04/2009 (H1N1)) wskazuje na liczne miejscowe obniżenia wartości. Może to świadczyć o tym, że w tych miejscach dochodzi do oddziaływań vRNA-vRNA *in virio*, które nie powstają tak licznie w warunkach *in cellulo*, bądź oddziaływań w



Ryc. 80 Interakcje pomiędzy segmentami vRNA w warunkach *in virio* zidentyfikowane dla różnych szczepów IAV. Rycina pochodzi z publikacji *Dadonaite et al.* (Nature Microbiology, 2019)

warunkach *in cellulo*, które nie powstają w wirionie. W przypadku segmentu vRNA8 stwierdzono istnienie nielicznych rejonów, w których dochodzi od obniżenia wartości korelacji reaktywności. Co ciekawe, badania przeprowadzone przez *Dadonaite et al.* wskazały, że w przypadku większości analizowanych szczepów, segment vRNA8 tworzy bardzo niewiele interakcji między segmentami (**Ryc. 80**)¹¹⁵. Uzyskany dla vRNA8 (**Ryc. 79**) współczynnik korelacji kroczącej dla reaktywności nukleotydów może zatem świadczyć o mniejszej ilości oddziaływań z innymi segmentami *in virio*.

Następnie sprawdzono czy zmiany reaktywności pomiędzy warunkami *in virio* oraz *in cellulo* wpływają na zmiany w strukturze drugorzędowej we wskazanych miejscach o obniżonej korelacji. Porównano globalne struktury segmentu vRNA5, 7 i 8 o najniższej wartości energii swobodnej (MFE). Porównanie struktur wykazało, że w ujęciu globalnym struktury różnią się między sobą znacznie. Świadczy to o tym, że niewielka zmiana w reaktywności może wynikać ze znacznie większej zmiany w fałdowaniu struktury. Pomimo różnic w fałdowaniu, które dotyczyły głównie parowania dalekiego zasięgu, potwierdzono, że niektóre motywy o charakterze lokalnym były wspólne dla struktur *in virio* oraz *in cellulo*. Struktury MFE segmentów vRNA5, 7 i 8 (*in virio* oraz *in cellulo*) zostały zawarte w rozdziale *Dodatek*.

Podsumowanie

Celem badań w niniejszej pracy doktorskiej było określenie struktury drugorzędowej vRNA wirusa grypy typu A w warunkach biologicznych.. Strukturę vRNA określano w trzech układach biologicznych: w warunkach lizatu komórkowo-wirusowego, w wirionie (*in virio*) oraz w czasie infekcji wirusowej ludzkich komórek płucnych A549 (*in cellulo*).

Po raz pierwszy zostały zaproponowane struktury drugorzędowe kilku segmentów vRNA wirusa grypy typu A w warunkach komórkowych. Ponadto, po raz pierwszy przeprowadzono szerokie badania struktury drugorzędowej szczepu A/California/04/2009 (H1N1) w lizatach, w wirionie i w linii komórkowej. Dzięki temu możliwe było porównanie struktury vRNA, która może się zmieniać w zależności od otoczenia oraz na różnych etapach cyklu wirusowego. Cenne było również porównanie struktur drugorzędowych vRNA *in virio* i *in cellulo* do struktur uzyskanych w warunkach *in vitro*. Cel pracy został zrealizowany a uzyskane dane będą źródłem dalszych głębszych analiz.

Badania z wykorzystaniem mapowania chemicznego w lizacie komórkowo-wirusowym (*in vivo-like*) pozwoliły na zaproponowanie struktury drugorzędowej vRNA8 wirusa grypy typu A szczepu A/California/04/2009 (H1N1). Aby wyodrębnić charakterystyczne cechy struktury drugorzędowej vRNA8, w modelowaniu struktury drugorzędowej wykorzystano różne parametry fałdowania, co umożliwiło analizę zmian strukturalnych, na które wpływ mogły mieć właśnie zastosowane parametry. Na podstawie danych eksperymentalnych przeprowadzono fałdowanie vRNA8 w programie RNAStructure zarówno w ujęciu całościowym (globalnym), jak i w ujęciu lokalnym z zastosowaniem dystansu parowania zasad. Na podstawie danych eksperymentalnych obliczone zostało również prawdopodobieństwo parowania zasad oraz rejonów jednoniciowych w globalnej strukturze drugorzędowej vRNA8. Prawdopodobieństwo to obliczane jest na podstawie wszystkich wygenerowanych w programie RNAStructure struktur drugorzędowych. Największe prawdopodobieństwo parowania zasad (>90%) stwierdzono dla 10 motywów strukturalnych w rejonach 261-288 nt, 310-329 nt, 374-404 nt, 500-547 nt, 553-589 nt, 645-666 nt, 677-687 nt, 698-713 nt, 717-784 nt and 792-814 nt. Wszystkie te motywy istnieją również w strukturze

przewidzianej z zastosowaniem ograniczenia parowania zasad do 150 nt. Wskazuje to na trwałość termodynamiczną lokalnych motywów strukturalnych niezależnie od wybranych parametrów fałdowania i duże prawdopodobieństwo ich powstawania. Stwierdzono poza tym, że vRNA8 w lizacie komórkowo-wirusowym jest mocno ustrukturyzowane. W strukturze globalnej można wyróżnić domeny, kilka stabilnych termodynamicznie spinek, a także rejony parowania dalekiego zasięgu. Zaobserwowano również rejony o wyraźnie większej reaktywności świadczącej o labilności danego regionu.

Struktura drugorzędowa RNA zaproponowana na podstawie badań *in vitro* różni się od tej formującej się w lizacie komórkowo-wirusowym. Największe różnice zaobserwowano dla rejonów tworzących oddziaływania typu dalekiego zasięgu, a więc tworzących pomiędzy odległymi nukleotydami. Na różnice w obu strukturach mogły mieć wpływ interakcje typu RNA-RNA czy RNA-białko, które vRNA8 mogło tworzyć ze składnikami obecnymi w lizacie komórkowym. Zaobserwowano również motywy strukturalne wspólne dla obu struktur w rejonach: 218-257 nt, 261-288 nt, 305-335 nt, 500-547 nt, 553-589 nt, 645-666 nt oraz 797-809 nt. W następnej kolejności przeprowadzona została analiza zachowawczości parowania zasad na podstawie danych sekwencyjnych dla wszystkich dostępnych w bazie danych (Influenza Research Database) szczepów wirusa grypy typu A. Średnia zachowawczość parowania zasad dla struktury globalnej vRNA8 wynosiła 86,38%, a lokalnej 86,45%. Największą zachowawczość wspólnych motywów w strukturze globalnej i lokalnej vRNA8 w lizacie stwierdzono w rejonach: 30-64 nt (89,05%), 265-284 nt (97,03%), 312-327 nt (96,09%), 500-547 nt (89,40%) oraz 717-784 nt (88,87%). W przypadku struktury lokalnej zaobserwowano trzy motywy o wysokiej zachowawczości w rejonach: 216-260 nt (94,99%), 71-124 nt (92,14%) oraz 836-877 nt (90,37%).

Porównanie struktury drugorzędowej vRNA8 A/California/04/2009 (H1N1) w lizacie z zaproponowanymi strukturami drugorzędowymi vRNA8 innych szczepów wirusa wykazało istnienie zachowawczych motywów strukturalnych w różnych warunkach biologicznych. Dwa motywy w rejonach 261-288 nt oraz 312-327 nt były obecne w strukturach drugorzędowych vRNA8: *in vitro* szczepu A/Vietnam/1203/2004 (H5N1)¹⁴ oraz *in vitro*, *ex virio* oraz *in virio* szczepu

A/WSN/1933 (H1N1)¹¹⁵. Jeden motyw w rejonie 797-809 nt był przewidziany we wszystkich trzech strukturach szczepu A/WSN/1933, co może wskazywać na jego zachowawczość w obrębie podtypu H1N1.

Następnie porównano występowanie motywów strukturalnych w strukturze vRNA8 w lizacie z profilem wiązania białka NP określonego dla kompleksu vRNP8⁴⁶. Wysokoprawdopodobne (prawdopodobieństwo parowania zasad – RNAStructure) motywy strukturalne vRNA8 w lizacie były przewidziane w rejonach charakteryzujących się zarówno wysokim jak i niskim profilem wiązania z białkiem NP. To pokazało, że istnienie tych motywów strukturalnych jest generalnie niezależne od wiązania białka NP. Brak korelacji pomiędzy strukturą drugorzędową vRNA w kompleksie vRNP a wiązaniem białka NP wskazali autorzy innej pracy¹¹⁵.

W drugiej części niniejszej pracy doktorskiej skupiono się na określeniu struktur drugorzędowych wszystkich segmentów vRNA szczepu A/California/04/2009 (H1N1) w warunkach *in virio* oraz *in cellulo*. W pierwszej kolejności opracowane zostały metody przygotowania biblioteki fragmentów do sekwencjonowania NGS. Zoptymalizowane zostały dwa niezależne protokoły mapowania oraz przygotowania biblioteki NGS metodą profilowania mutacji (SHAPE-MaP) na matrycy RNA mapowanego odczynnikami chemicznymi w warunkach *in virio* oraz *in cellulo*. Jako kontrola pozytywna wyników mapowania w warunkach *in vivo* wybrana została cząsteczka o znanej strukturze drugorzędowej oraz trzeciorzędowej – 18S rRNA ludzki. Kontrolna próbka (biblioteka NGS) została przygotowana na tej samej matrycy całkowitego RNA, który został wykorzystany do przygotowania biblioteki vRNA *in cellulo*. Pozwoliło to na określenie wiarygodności uzyskanych wyników dotyczących mapowania struktur drugorzędowych vRNA *in vivo*. Analiza wyników mapowania chemicznego struktury 18S rRNA wskazała na wysoką korelację wyników mapowania ze znaną strukturą drugorzędową tej cząsteczki. Wysokie reaktywności zostały zaobserwowane dla nukleotydów w rejonach jednoniciowych. W obrębie niektórych motywów strukturalnych 18S rRNA stwierdzono niższą reaktywność nukleotydów w rejonach przewidzianych jako jednoniciowe w warunkach *in vitro*. Analiza struktury cząsteczki 18S rRNA w kompleksie rybosomalnym ujawniła, że przewidziane *in vitro* rejony jednoniciowe wykazujące niską reaktywność *in cellulo* znajdowały się w miejscu niedostępnym dla odczynników chemicznych. Z kolei wysokie reaktywności

nukleotydów zaobserwowano dla motywów eksponowanych na zewnątrz kompleksu rybosomalnego. Analiza próby kontrolnej wskazała na uzyskanie wiarygodnych wyników mapowania, a także potwierdziła właściwe dobranie warunków eksperymentalnych.

Analiza wyników sekwencjonowania pokazała wysokie pokrycie wszystkich cząsteczek vRNA w eksperymencie mapowania *in virio*. W przypadku eksperymentu *in cellulo* uzyskano częściowe pokrycie sekwencji segmentów vRNA 1, 2, 3, 4 i 6 oraz pokrycie pełnej długości sekwencji segmentów 5, 7 oraz 8. Uzyskanie zróżnicowanego pokrycia cząsteczek vRNA poszczególnych segmentów w warunkach *in cellulo* mogło być wynikiem zróżnicowanej kinetyki namnażania tych cząsteczek w czasie cyklu komórkowego. Niedawno opublikowane badania dotyczące kinetyki namnażania segmentów wirusowego RNA w istocie wykazały, że namnażanie segmentów vRNA nie ma zależności liniowej²⁰⁴.

Na podstawie wyników sekwencjonowania przeprowadzona została analiza profilu mutacji, który posłużył do obliczenia reaktywności poszczególnych nukleotydów vRNA. Przed przystąpieniem do wyliczenia średniej reaktywności poszczególnych nukleotydów vRNA z trzech niezależnych powtórzeń biologicznych wykonano test korelacji Pearson'a. Wykazano wysoki (0,8-0,9) oraz bardzo wysoki współczynnik korelacji (0,9-0,99) dla wszystkich powtórzeń biologicznych względem siebie, w zależności od segmentu vRNA oraz stosowanego odczynnika. Po potwierdzeniu powtarzalności danych wyliczona została średnia reaktywność nukleotydów dla wszystkich cząsteczek vRNA dla obu odczynników chemicznych.

Z użyciem programu RNAStructure zaproponowano struktury drugorzędowe MEA vRNA wszystkich segmentów w warunkach *in virio* oraz częściowe struktury drugorzędowe segmentów 1, 2, 3, 4, 6 i pełnej długości struktury drugorzędowe vRNA segmentu 5, 7 oraz 8 w warunkach *in cellulo*. Następnie przeprowadzono analizy strukturalne skupiające się na wyłonieniu wspólnych motywów strukturalnych dla obu warunków biologicznych oraz motywów unikatowych dla struktur drugorzędowych *in virio* oraz *in cellulo* danego vRNA. Uzyskane struktury drugorzędowe vRNA porównano również ze znanymi strukturami drugorzędowymi *in vitro* oraz *in virio* dla innych szczepów wirusa grypy typu A. Przeprowadzone analizy umożliwiły

zapropozowanie unikalnych dla szczepu A/California/04/2009 jak i zachowawczych dla wielu szczepów IAV motywów strukturalnych dla poszczególnych vRNA.

Przewidywanie struktury drugorzędowej RNA w istniejących programach takich jak np.: RNAStructure, wiąże się z pewnymi ograniczeniami. W przypadku predykcji struktur drugorzędowych RNA *in vivo* i *in virio* na podstawie danych z mapowań chemicznych z większym prawdopodobieństwem uzyskuje się informacje dotyczące lokalnych motywów strukturalnych. Dużo trudniej jest prawidłowo przewidzieć oddziaływania dalekiego zasięgu, mając na uwadze możliwe zasłanianie RNA przez biomolekuły i np.: obniżoną reaktywność rejonów jednoniciowych, które są uwikłane w oddziaływania. Stosowane w trakcie predykcji struktur MEA ograniczenia parowania zasad uniemożliwiły ocenę podobieństwa struktur vRNA w ujęciu całościowym (globalnym). Dlatego też w celu dokładnego zbadania różnic oraz podobieństw strukturalnych w strukturze drugorzędowej vRNA8 w różnych warunkach biologicznych wygenerowano struktury globalne vRNA8 *in virio* oraz *in cellulo* o minimalnej wartości energii swobodnej (struktury MFE). Następnie porównano uzyskane struktury ze strukturą drugorzędową vRNA8 uzyskaną w warunkach lizatu (*in vivo-like*). Porównanie struktur drugorzędowych wykazało, że część motywów strukturalnych było wspólnych dla wszystkich trzech struktur, niezależnie od zastosowanych parametrów fałdowania. W strukturze lizatu zidentyfikowano dwa motywy strukturalne wspólne ze strukturą *in cellulo* oraz dwa motywy strukturalne wspólne ze strukturą *in virio*. Wskazuje to na to, że struktura drugorzędowa vRNA8 w lizacie komórkowo-wirusowym jest strukturą pośrednią pomiędzy strukturami *in virio* i *in vivo*. Najwięcej różnic strukturalnych zaobserwowano w oddziaływaniach dalekiego zasięgu. Różnice te pokazują, że struktura drugorzędowa vRNA8 w warunkach biologicznych jest labilna, choć niektóre motywy (szczególnie te lokalne) zostają zachowane. W globalnej strukturze *in virio* oraz w lizacie wspólny był również motyw promotora *panhandle*. Obecność struktury promotora *panhandle* w kompleksie vRNP była wcześniej potwierdzona eksperymentalnie¹¹⁵.

Opublikowana niedawno praca dotycząca struktur drugorzędowych vRNA *in virio* innego szczepu IAV (A/WSN/1933)¹¹⁵ wskazała, że występowanie motywów strukturalnych o wysokim prawdopodobieństwie parowania zasad (ppz,

prawdopodobieństwo obliczone na podstawie funkcji *Partition function* w programie RNAStructure) było niezależne od profilu wiązania białka NP w kompleksie vRNP badanego szczepu⁴⁶. Porównanie struktur drugorzędowych uzyskanych w niniejszej pracy dla wybranych cząsteczek vRNA *in virio* wykazało, że motywy przewidziane z dużym ppz były przewidziane w rejonach zarówno bogatych jak i ubogich w wiązania z białkiem NP.

Na ostatnim etapie analiz postanowiono zbadać różnice pomiędzy uzyskanymi danymi reaktywności nukleotydów (względem reagentów SHAPE i DMS) vRNA *in virio* i *in cellulo*. W tym celu wykorzystano narzędzia statystyczne. W pierwszej kolejności zbadano współczynnik korelacji Pearson'a pomiędzy danymi reaktywności dla cząsteczek vRNA5, 7 i 8 uzyskanymi w warunkach *in virio* oraz *in cellulo*. Współczynniki korelacji wskazały na wysoką (0,7-0,9) oraz bardzo wysoką (>0,9) korelację pomiędzy wartościami reaktywności. Wskazuje to na duże podobieństwo między uzyskanymi danymi. W celu zbadania miejsc, w których wartości reaktywności różnią się między sobą wyznaczono dodatkowo współczynnik korelacji kroczącej. Wyniki wskazały rejony, w których z dużym prawdopodobieństwem struktura może być zachowana, a także rejony w których dochodzi do zmian reaktywności nukleotydów. Najwięcej zmian zaobserwowano dla segmentu vRNA5 i vRNA7, a najmniej w przypadku vRNA8. Na podstawie badań oddziaływań między segmentami vRNA w kompleksach vRNP stwierdzono, że cząsteczki te tworzą bardzo rozbudowaną sieć interakcyjną¹¹⁵. Co ciekawe, wyniki badań wskazały, że segment vRNA5 tworzy liczne interakcje między-segmentalne, tymczasem segment vRNA8 posiada ich zdecydowanie mniej. Obniżenie wartości współczynnika korelacji kroczącej pomiędzy danymi *in virio* oraz *in cellulo* mogą więc wskazywać, że w tych miejscach dochodzi do oddziaływań vRNA-vRNA. Z drugiej strony mogą to być to zmiany spowodowane oddziaływaniem vRNA, lub rearanżacją struktury podczas cyklu wirusowego już w komórce. Potwierdzenie tych obserwacji jest możliwe po wyznaczeniu dokładnych danych dotyczących interakcji między segmentami vRNA badanego w niniejszej pracy szczepu A/California/04/2009. W celu zbadania zależności pomiędzy obniżeniem wartości korelacji kroczącej a zmianami w strukturach drugorzędowych (*in virio* oraz *in cellulo*) porównano struktury globalne o najniższej energii swobodnej (MFE). Porównanie tych struktur wykazało, że lokalne

zmiany w wartościach reaktywności nukleotydów przekładają się na większe zmiany w strukturze w ujęciu globalnym. Różnice były najbardziej widoczne w parowaniu odległych rejonów co powodowało tworzenie nowych domen. Pomimo tych różnic, w strukturach drugorzędowych *in virio* oraz *in cellulo* danego segmentu (vRNA5, 7 i 8) stwierdzono wspólne motywy strukturalne.

Chociaż struktury drugorzędowe RNA wirusa grypy od lat stanowią obiekt badań wielu grup naukowych na całym świecie, to dotychczas nie zaproponowano żadnej struktury w warunkach *in cellulo*. Wiedza dotycząca tych struktur w ujęciu biologicznym była od dawna oczekiwana przez środowisko naukowe, a opisane w niniejszej pracy doktorskiej badania po raz pierwszy prezentują szerokie analizy dotyczące dynamiki struktur drugorzędowych vRNA w środowisku biologicznym. Rozszerzenie analiz o struktury vRNA8 w lizacie oraz o struktury *in virio* wszystkich segmentów badanego szczepu umożliwiły porównanie fałdowania vRNA w różnych warunkach środowiska. Uzyskane dane są całkowicie nową wiedzą dotyczącą biologii wirusa grypy i ukazują one istotność struktury vRNA w cyklu życiowym wirusa. Ze względu na silną zależność między strukturą a funkcją pełnioną przez RNA, poznanie motywów strukturalnych wspólnych dla wielu szczepów wirusa grypy typu A umożliwi opracowanie nowych inhibitorów namnażania wirusa.

Materialy i metody

1. Materiały

1.1. Sekwencje stosowanych oligonukleotydów

Tabela 8 Sekwencje stosowanych oligonukleotydów

Nazwa	Sekwencja 5'-3'
Startery do odwrotnej transkrypcji do rozdziału na elektroforezie kapilarnej	
1_8	*AGCAAAAGCAGGGTGACAAAAC
2_8	*ACATGACCCTCGAGGAAATGTC
3_8	*CTTGAATGGAATGGTAACACGGTTC
	* znakowane wybranym barwnikiem fluorescencyjnym - FAM, JOE, TAMRA, ROX
Startery do odwrotnej transkrypcji komplementarne do vRNA8	
7VRT	GGCCGTCATGGTGGCGAATAGCCGAGATCGCGCAGAGACT
8VRT	GGCCGTCATGGTGGCGAATAATCGTGGAATGGATCTTGA
Starter do odwrotnej transkrypcji uniwersalny dla wszystkich segmentów vRNA	
Uni-12	AGCAAAAGCAGG
Startery do reakcji PCR dla poszczególnych segmentów vRNA A/California/04/2009	
PB2-F	AGCGAAAGCAGGTCAATTAT
PB2-R	AGTAGAAACAAGGTCGTTTTTAAAC
PB1-F	AGCGAAAGCAGGCAAACCAT
PB1-R	AGTAGAAACAAGGCATTTTTTCATG
PA-F	AGCGAAAGCAGGTAATGATCC
PA-R	AGTAGAAACAAGGTAATTTTTTGG
HA-F	AGCAAAAGCAGGGGAA
HA-R	AGTAGAAACAAGGGTGTTTT
NP-R	AGCAAAAGCAGGGTAGATAATC
NP.-F	AGTAGAAACAAGGGTATTTTTTC
NA-R	AGCAAAAGCAGGAGTTTAAAATG
NA-F	AGTAGAAACAAGGAGTTTTTTGAAC
M-F	AGCAAAAGCAGGTAGATATTG
M-R	AGTAGAAACAAGGTAGTTTTTTAC
NS-F	AGCAAAAGCAGGGTGACAAA
NS-R	AGTAGAAACAAGGGTGTTTT
Startery do PCR w czasie rzeczywistym	
VTAG	GGCCGTCATGGTGGCGAAT
8VQR	TGTCAGAAAGGTAGCGCGAA
Startery do reakcji PCR - matryca do transkrypcji <i>in vitro</i> vRNA8	
RC8	AGCAAAAGCAGGGTGACAAAACATAATG
FC8	GCGTAATACGACTCACTATAGGGAGTAGAAACAAGGGTGTTTTTATC
Startery do reakcji PCR - matryca do biblioteki NGS, amplifikacja wszystkich segmentów vRNA	
HFA**	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGCAAAAGCAGG
HRA**	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGTAGAAACAAGG
	**na czerwono zaznaczono sekwencje adapterowe dla transpozazy

Tabela 9 Sekwencje starterów z indeksami Illumina

Index	sekwencja 5'-3'
Index 1 [i7] adapter	CAAGCAGAAGACGGCATACGAGAT[i7]GTCTCGTGGGCTCGG
Index 2 [i5] adapter	AATGATACGGCGACCACCGAGATCTACAC[i5]TCGTCCGTCAGCGTC
i7 adapter	sekwencja 5'-3'
N701	TAAGGCGA
N702	CGTACTAG
N703	AGGCAGAA
N704	TCCTGAGC
N705	GGACTCCT
N706	TAGGCATG
i5 adapter	sekwencja 5'-3'
S502	ATAGAGAG
S503	AGAGGATA
S504	TCTACTCT
S517	TCTTACGC

1.2. Komercyjnie dostępne zestawy

Zestaw do transkrypcji in vitro

- AmpliScribe™ T7-Flash™ Transcription Kit (Lucigen)

Zestaw do oczyszczania RNA in vitro

Rneasy MinElute Cleanup Kit (QIAGEN)

Zestaw do oczyszczania RNA po reakcjach enzymatycznych

- Rneasy Mini Kit (QIAGEN)

Zestaw do oczyszczania po reakcji odwrotnej transkrypcji

- illustra MicroSpin G-25 Columns (GE Healthcare)

Zestaw do oczyszczania po reakcji syntezy drugiej nici

- PureLink PCR Micro Kit (ThermoFisher)

Zestaw do oczyszczania produktów PCR

- PCR Mini Kit (Syngen)

Zestaw do rozdziału kwasów nukleinowych na bioanalyzerze

- High Sensitivity DNA kit (Agilent)

Zestawy do mierzenia stężeń na fluorymetrze QUBIT

- HS dsDNA Assay Kit, BR dsDNA Assay Kit, HS RNA Assay Kit, BR RNA Assay Kit, RNA IQ Assay (ThermoFisher)

Zestaw indeksów do bibliotek NGS

- Nextera XT Index Kit 24 indexes (Illumina)

Zestaw do przygotowania bibliotek NGS

– Nextera XT DNA Library Preparation Kit (Illumina)

Kulki magnetyczne typu SPRI do oczyszczania kwasów nukleinowych

– Ampure RNAClean XP Beads, Ampure XP Beads (Beckman)

Zestaw do reakcji odwrotnej transkrypcji

– Super Script® III Reverse Transcriptase (Invitrogen)

Zestaw do reakcji odwrotnej transkrypcji typu SHAPE-MaP

– Super Script® II Reverse Transcriptase (Invitrogen)

Zestaw do syntezy drugiej nici

- NEBNext® Ultra Non-Directional RNA Second Strand Synthesis Module (New England Biolabs)

Zestaw do sekwencjonowania drabinek DNA

-USB®Thermo Sequenace Cycle Sequencing Kit (Affymetrix)

Zestaw do reakcji PCR w czasie rzeczywistym

- 5x HOT FIREPol® EvaGreen qPCR Plus (No ROX) (Solis Biodyne)

Zestaw membran nitrocelulozowych wykorzystywanych w metodzie Western Blot

- Trans-Blot Turbo RTA Mini 0.2 µm nitrocellulose transfer kit (BioRad)

Zestaw do detekcji białek

Westar Sun (Cyanagen)

1.3. Enzymy

-Polimeraza Q5 High-Fidelity DNA (New England Biolabs)

-Polimeraza Phusion High-Fidelity PCR Master 2x Mix with HF Buffer (New England Biolabs)

-Rekombinowany inhibitor Rnaz - Rnasin® (Promega)

-Enzym DNaza I – Ambion™ DNase I (RNase-Free) (ThermoFisher)

1.4. Barwniki fluorescencyjne

Startery DNA używane do reakcji odwrotnej transkrypcji w odczycie mapowań chemicznych RNA z zastosowaniem elektroforezy kapilarnej były znakowane na 5' końcu następującymi barwnikami fluorescencyjnymi:

5-FAM (5-karboksyfluoresceina),

5-ROX (5-karboksy-x-rodamina),

6-JOE (6-karboksy-4', 5'-dichloro-2'7'-dimetoksyfluoresceina),

6-TAMRA (6-karboksytetrametylorodamina) (AnaSpec.)

1.5. Wzorce wielkości

Wzorzec wielkości RNA - RiboRuler High Range RNA Ladder (ThermoFisher)

Wzorzec wielkości DNA – Perfect Plus MWQ DNA Ladder (EurX)

Wzorzec wielkości DNA – Low Molecular Weight DNA Ladder (NEB)

Wzorzec wielkości białek – PageRuler Prestained Protein Ladder (ThermoFisher)

1.6. Plazmidy

Podczas badań wykorzystany został plazmid pPolNS (Invitrogen) z wklonowanym insertem DNA segmentu 8 vRNA, który posłużył do przygotowania matrycy DNA do transkrypcji vRNA8 *in vitro*.

1.7. Odczynniki dostępne komercyjnie

- Bezwodny roztwór dimetylosulfotlenku (DMSO) 99.9%; roztwór dimetylosiarczku (DMS) 99%; EDTA (SIGMA); -N,N,N',N'-Tetrametyloetylenodiamina (TEMED); MgCl₂; NaCl; tetraboran sodu; nadsiarzan amonu (APS); błękit bromofenolowy; surowicza albumina wołowa (BSA); penicylina-streptomycyna; trypsyna traktowana TPCK; chlorek manganu, roztwór trypsyna/EDTA do pasażu komórek, Triton X-100, Trizol™ LS Reagent, chloroform:alkohol izoamylowy 24:1 (SIGMA)
- GlycoBlue™ Coprecipitant, GeneRuler Low Range DNA ladder (ThermoFisher)
- Simply Safe, dNTP (EurX)
- Gel Loading Dye Purple 6x; Random Primer Mix (New England Biolabs)
- Absolutny alkohol etylowy 99.9%, 2-propanol (POCH)
- 2M roztwór chlorku potasu (KCl) (Ambion)
- Lipofectamine 2000; 1M roztwór Tris pH 8.0 (Invitrogen)
- Kwas borowy (Chempur)
- Agaroza, 40% roztwór akrylamid-bisakrylamid 29:1; octan sodu; mocznik; PBS 10x pH 7.4 (BioShop)
- Dullbecco's Modified Eagle's Medium - DMEM; Reduced serum medium - Opti-MEM, serum płodowe cieląt FBS; penicylina-streptomycyna-glutamina 100x (Gibco)
- bufor EB (10 mM Tris-HCl) (QIAGEN)

1.8. Bufory i roztwory

Bufor do znakowania barwnikami fluorescencyjnymi

0.1 M tetraboran sodu pH 8,5

Bufor do fałdowania RNA 2x

100 mM HEPES pH 7,5

600 mM NaCl

10 mM MgCl₂

Bufor do elucji kwasów nukleinowych z żelu poliakrylamidowego

0,3 M Octan sodu pH 5,5

Bufor 1x TBE

100 mM Tris-HCl pH 8.3

100 mM kwas borowy

1 mM EDTA

Roztwór żelu poliakrylamidowego 12%

12% roztwór akrylamid-bisakrylamid

1x TBE

8,3 M Mocznik

0,07% APS

0,04% TEMED

Roztwór obciążający do rozdzielania kwasów nukleinowych w żelu poliakrylamidowym w warunkach denaturujących

8,3 M Mocznik

1x TBE

0,03% błękit bromofenolowy

0,03% oranż metylowy

Pre-bufor do odwrotnej transkrypcji SHAPE-MaP 5x

250 mM Tris-HCl pH 8.0

375 mM KCl

50 mM DTT

2,5 mM dNTP

Bufor do odwrotnej transkrypcji SHAPE-MaP 2,5x

Pre-bufor 5x stężony zmieszany z 30mM roztworem MnCl₂ w stosunku 1:1

Bufor TNE

50 mM Tris-HCl

100 mM NaCl

0,1 mM EDTA

Roztwór 20% sacharozy do zagęszczania wirusa grypy

20% roztwór sacharozy w buforze TNE

Bufor do zatrzymywania reakcji z DMS

50 mM Tris-HCl pH 7,5

100 mM NaCl

3 mM MgCl₂

40 mM β-merkaptioetanol

Roztwory wykorzystywane do barwienia białek z Coomassie Blue

Roztwór utrwalający

50% etanol

Kwas octowy

do 500 ml H₂O

Roztwór do barwienia

50% metanol

kw. octowy

0,25 g Coomassie Brilliant Blue R-250 (Serva)

do 500 ml H₂O

Roztwór odbarwiający

5% metanol

kw. octowy

do 500 ml H₂O

Bufory i roztwory stosowane w reakcji Western-Blot

Bufor lizujący 2x stężony

60 mM Tris pH 6,8

10% glicerol

2% SDS

5% β-merkaptioetanol

0,1% błękit bromofenolowy
1 mM EDTA

Roztwór 12% żelu poliakrylamidowego rozdzielającego

12% roztwór bis-poliakrylamidu

0,25 M Tris-HCl pH 8.8

0,1% SDS (w/o)

0,1% APS

0,01% TEMED (SIGMA)

do 7,5 ml H₂O

Roztwór 5% żelu poliakrylamidowego zagęszczającego

5% roztwór bis-poliakrylamidu

0,125 M Tris-HCl pH 6,8

0,1% SDS (w/o)

0,1% APS

0,01% TEMED

do 5 ml H₂O

Bufor do elektroforezy

250 mM Tris

1,92 M glicyna

1% SDS (w/o)

do 500 ml H₂O

Bufor PBS-T

Roztwór 1000 ml niesterylnego buforu PBS z dodatkiem 1 ml Tween 20

5% roztwór do blokowania

2,5 g mleka w proszku zmieszanego z 50 ml buforu PBS-T

Bufor do transferu

200 ml (5x) buforu do transferu zmieszano z 200 ml EtOH (96%) i uzupełniono do 1000 ml wodą

1.9. Przeciwciała

Przeciwciała używane do testu immunofluorescencji

- Przeciwciało pierwszorzędowe – mouse anti-influenza primary antibody targeting nucleoprotein (NP) (MAB8257, Merck)
- Przeciwciało drugorzędowe – FITC-conjugated secondary rabbit anti-mouse IgG antibody (AP160F, Merck)

Przeciwciała używane do znakowania białek w metodzie Western-Blot

- Przeciwciało pierwszorzędowe na białka NA i HA – anti-capsid goat α -Flu (AB1074, Merck)
- Przeciwciała drugorzędowe – mouse anti-goat, HRP α -goat (AB186P), goat anti-rabbit, HRP (AP3074P, Merck)
- Przeciwciało pierwszorzędowe na białko aktyne – actin monoclonal antibody (ACTN05(C4)), Invitrogen)

1.10. Linie komórkowe

Propagacja wirusa grypy oraz eksperymenty dotyczące przygotowania lizatów komórkowych wykonane zostały na linii komórkowej Madin-Darby Canine Kidney (MDCK) (Sigma-Aldrich, ATCC®CCL-34). Eksperymenty mapowania odczynnikami chemicznymi vRNA przeprowadzone zostały na linii komórkowej adenocarcinomic human alveolar basal epithelial cells (A549) (Sigma-Aldrich, ATCC®CCL-185TM).

1.11.Szczep wirusa grypy

Wszystkie eksperymenty wykonano z zastosowaniem wirusa grypy A/H1N1 szczepu A/California/04/2009, który otrzymano dzięki uprzejmości prof. Luisa Martinez’a-Sobrido z Texas Biomedical Research Institute, San Antonio, USA.

1.12. Media i roztwory wykorzystywane w eksperymentach komórkowych

Bufor do lizy komórkowej CL

10 mM Tris-HCl pH 7.4

150 mM NaCl

0,25% Igepal

Bufor do mapowania chemicznego RNA in virio

10 mM Tris-HCl pH 7.4

100 mM NaCl

0,1 mM EDTA

Medium hodowlane

1x Dulbecco's Modified Eagle Medium (DMEM) 4,5 g/l D-glukozy, 110 mg/l pirogronianu sodu

10% Serum płodowe cieląt (FBS)

2 mM Glutamina

100 U/ml Penicylina

100 µg/ml Streptomycyna

Medium infekcyjne

1x PBS

0,3% Surowicza albumina wołowa (BSA)

100 U/ml Penicylina

100 µg/ml Streptomycyna

100 mM Ca²⁺/Mg²⁺

Medium post-infekcyjne

1x DMEM

0,3% Surowicza albumina wołowa (BSA)

100 U/ml Penicylina

100 µg/ml Streptomycyna

2 mM Glutamina

Trypsyna traktowana TPCK:

1 µg/ml - Komórki MDCK

0,5 µg/ml - Komórki A549

Medium post-infekcyjne z roztworem mikrocelulozy (Avicel)

Standardowe medium post-infekcyjne wymieszane z 3% roztworem Avicelu w stosunku 2:1

Roztwór utrwalający komórki

0,5% Triton X-100

4% Formaldehyd

PBS do 50 ml

Bufor blokujący

2% BSA

do 50 ml PBS

Roztwór przeciwciała I-rzędowego

Przeciwciało anti-NP (1:1000)

do 1 ml bufor blokujący

Roztwór przeciwciała II-rzędowego

Przeciwciało FITC (1:150)

DAPI (1:1000)

do 1 ml bufor blokujący

2. Metody

2.1. Synteza oligonukleotydów

Wszystkie oligonukleotydy (DNA) uwzględnione w **Tabeli 8** zostały syntetyzowane metodą amidofosforynową na podłożu stałym przez prof. dr hab. Ryszarda Kierzka, kierownika Zakładu Chemii i Biologii Kwasów Nukleinowych IChB PAN.

2.2. Odblokowanie oligonukleotydów po syntezie

W celu odblokowania i zdjęcia oligonukleotydów z podłoża po syntezie, do podłoża silikażelu dodano 1,5 ml roztworu 32% amoniaku i inkubowano w 55°C. Po 12 godzinach inkubacji próbówki schłodzono w -20°C. Znad podłoża zebrano roztwór amoniaku z oligomerem i przeniesiono do nowej próbówki. Podłoże przemyto dodatkowo dwukrotnie wodą. Miało to na celu wypłukanie pozostałości oligomeru. Następnie roztwór wysuszono w z zastosowaniem koncentratora próżniowego (Speedvac). Osad oligomeru rozpuszczono w wodzie i przechowywano w -20°C do czasu oczyszczenia.

Oligomery z grupą C6-aminolinkerową na 5'-końcu (w celu późniejszego przyłączenia fluoroforu) były dodatkowo inkubowane przez 3 godziny z dodatkiem 0,9 ml 80% kwasu octowego. Osad oligomeru rozpuszczono w 0,2 ml wody i strącono w 10 ml roztworu 1% nadchloranu sodu w acetonie. Uzyskany roztwór zwirowano przez 15 minut 1700 g. Przed rozpuszczeniem w wodzie, osad przepłukano dwa razy acetonem i wysuszono próżniowo. Odblokowany oligomer przechowywano w -20°C do czasu znakowania barwnikami fluorescencyjnymi.

2.3. Znakowanie końca 5' oligonukleotydów barwnikami fluorescencyjnymi

Po określeniu stężenia oligomerów z grupą C6-aminolinkerową na 5'- końcu pobrano 300 µg oligomeru i wysuszono w koncentratorze próżniowym. Przed reakcją znakowania osad oligomeru rozpuszczono w 11 µl H₂O. Do roztworu oligomeru dodano 75 µl buforu do znakowania oraz 14 µl barwnika rozpuszczonego w DMSO*. Reakcję znakowania inkubowano przez 18 godzin w termobloku w temperaturze pokojowej (RT) z wytrząsaniem (300 rpm). Mieszaninę reakcyjną strącano w etanolu według standardowej procedury.

*Mieszaninę barwnika fluorescencyjnego przygotowywano poprzez rozpuszczenie określonej ilości barwnika w 140 µl bezwodnego roztworu DMSO. W zależności od stosowanego barwnika stosowano inne stężenia: 2 mg barwnika w przypadku 5-FAM i 6-TAMRA, z kolei w przypadku 6-JOE i 5-ROX było to 2,5 mg.

2.4. Oczyszczanie oligonukleotydów na denaturującym żelu poliakrylamidowym

Oligonukleotydy DNA bezpośrednio po odblokowaniu, bądź po reakcji znakowania barwnikami fluorescencyjnymi, oczyszczono na preparatywnym 12% żelu poliakrylamidowym. Do roztworu oczyszczanego oligomeru dodano obciążacz denaturujący z mocznikiem i barwnikami (cyjanol ksylenowy, błękit bromofenolowy, oranż G). Produkty syntezy wizualizowano z zastosowaniem lampy UV i wycinano z żelu przy pomocy skalpela. W celu elucji oczyszczonego oligomeru, skrawki żelu inkubowano przez noc w buforze elucyjnym w 4°C. W sumie wykonano 3 elucje z żelu, każdorazowo dodając nową porcję buforu elucyjnego. Roztwór po elucji strącano w etanolu według standardowej procedury.

2.5. Synteza odczynników chemicznych do reakcji mapowania SHAPE

Odczynnik 1M7 do reakcji mapowania chemicznego SHAPE został otrzymany według opublikowanej procedury chemicznej²⁰⁸. Odczynnik NAI został otrzymany zgodnie z procedurą opublikowaną przez *Spitale et al.*¹⁵¹. W celu zapobiegnięcia degradacji odczynników chemicznych zostały one zabezpieczone

parafilmem i przechowywane w małych porcjach w -20°C. Syntezy chemiczne odczynników zostały wykonane przez prof. dr hab. Ryszarda Kierzka, kierownika Zakładu Chemii i Biologii Kwasów Nukleinowych IChB PAN.

2.6. Strącanie kwasów nukleinowych

W celu wytrącenia kwasów nukleinowych z roztworu wodnego stosowano mieszaninę składającą się z 0,3 M octanu sodu stanowiącego 1/10 roztworu wodnego kwasu nukleinowego oraz 2,5 objętości (dla RNA) albo 3 objętości (dla DNA) absolutnego alkoholu etylowego. Mieszaninę inkubowano przez noc w -20°C. Po tym czasie mieszaninę wirowano przez 30 minut 18 000 g. Po ściągnięciu supernatantu, osad przemywano 70% alkoholem etylowym i wirowano 10 minut 18 000 g. Procedurę przemywania wykonywano dwa razy. Osad suszono w cieplarni w 37°C i rozpuszczono w wodzie.

2.7. Elektroforeza kwasów nukleinowych w żelu agarozowym

Elektroforezę kwasów nukleinowych (DNA oraz RNA) prowadzono na 1% żelu agarozowym w buforze 1x TBE. Jako odczynnik interkalujący stosowano SimplySafe, który dodawany był do wystudzonego roztworu agarozy bezpośrednio przed jej wylaniem do aparatu do elektroforezy.

2.8. Oczyszczanie kwasów nukleinowych z zastosowaniem kulek magnetycznych typu SPRI

Optymalizację oczyszczania mieszaniny fragmentów DNA z krótkich fragmentów przeprowadzono z zastosowaniem standardu długości DNA GeneRuler Low Range DNA Ladder (ThermoFisher). Do każdej reakcji oczyszczania stosowano 5 µl standardu długości oraz 45 µl wody. Gotowa mieszanina (50 µl) była zmieszana z różną objętością kulek: 3 x (150 µl), 1,8 x (90 µl), 1 x (50 µl) oraz 0,8 x (40 µl) i inkubowana przez 10 minut. Po tym czasie próbki umieszczono na statywie magnetycznym i przemyto dwukrotnie 80% EtOH w objętości 200 µl. Probówki z etanolem inkubowano przez 30 sekund przed usunięciem. Po usunięciu resztek etanolu, kulki suszono ok. 5 minut aż do widocznego odparowania resztek etanolu. Probówki zdjęto ze statywu magnetycznego i dodano 15 µl wody w celu elucji. Produkty oczyszczania

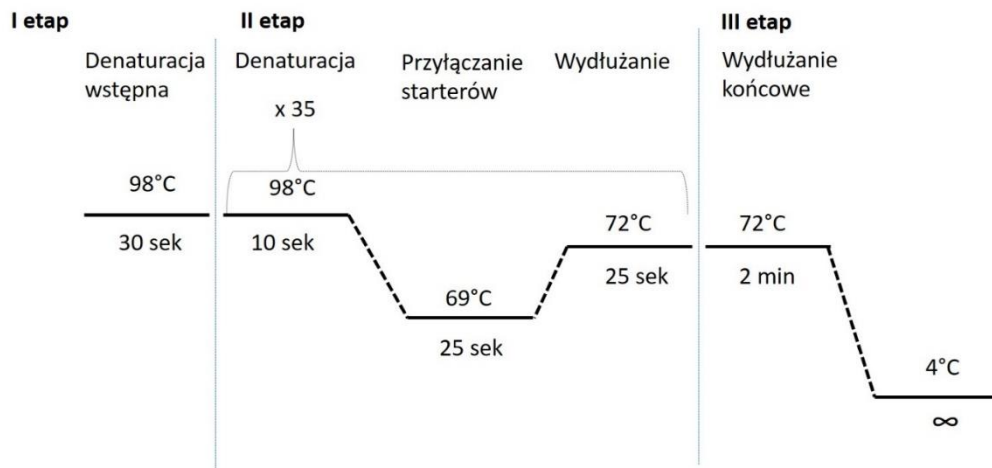
rozdzielono na 3% żelu agarozowym. Ostateczne objętości kulek, jakie zostały zastosowane zostały podane w opisie wyników eksperymentów.

W przypadku oczyszczania mieszaniny DNA z fragmentów krótkich oraz długich stosowano protokół do selekcji fragmentów (ang. size selection). Selekcję fragmentów wykonano poprzez zastosowanie dwóch objętości kulek. Po inkubacji mieszaniny fragmentów z pierwszą objętością kulek, supernatant ściągnięto i przeniesiono do nowej probówki. Następnie dodano drugą objętość kulek i ponownie inkubowano. Etapy przemycia kulek oraz elucji były takie same jak w przypadku standardowego oczyszczania.

2.9. Synteza matrycy DNA do przygotowania drabinek sekwencyjnych oraz do transkrypcji vRNA8 szczepu A/California/04/2009 *in vitro*

W celu otrzymania matrycy DNA do transkrypcji vRNA8 *in vitro* przeprowadzono reakcję PCR. Jako matrycę wykorzystano plazmid pPolNS (otrzymany od prof. Baek Kim, Uniwersytet Emory, Atlanta, Georgia, USA) z wklonowanym insertem DNA cząsteczki vRNA8. Reakcję przeprowadzono przy użyciu polimerazy Q5 High-Fidelity (NEB). Startery były komplementarne do obu końców cząsteczki vRNA8. Starter forward (FC8) był wydłużony o sekwencję promotora dla polimerazy T7 RNA niezbędną do przeprowadzenia reakcji transkrypcji *in vitro*²⁰⁹. Reakcję PCR przeprowadzono w objętości 50 µl na które składały się: 1x stężony bufor Q5, 10 mM dNTP, 15 µM startera FC8, 15 µM startera RC8, 1U polimerazy Q5 oraz 150 ng matrycy pPolNS. Stężenie oraz czystość produktu PCR zmierzono z zastosowaniem spektrometru NanoDrop One.

Profil temperaturowy reakcji:



Analizę długości produktu przeprowadzono na 1% żelu agarozowym z zastosowaniem wzorca długości DNA. Produkt PCR oczyszczono na kolumnie Syngen PCR Mini Kit zgodnie z protokołem producenta. Elucję wykonano w 80 μ l wody.

2.10. Transkrypcja segmentu 8 vRNA szczepu A/California/04/2009 *in vitro*

Transkrypcję *in vitro* segmentu 8 vRNA (890 nt) przeprowadzono na produkcie PCR posiadającym promotor dla polimerazy T7 RNA. Reakcję transkrypcji wykonano według protokołu producenta komercyjnie dostępnego zestawu AmpliScribe T7-Flash Transcription Kit (Lucigen) w 50 μ l. Do reakcji użyto 1 μ g matrycy. Mieszaninę reakcyjną inkubowano przez 90 minut w 37°C. Po tym czasie przeprowadzono enzymatyczną degradację matrycy DNA z zastosowaniem enzymu DNazy I (15 minut, 37°C). Produkt transkrypcji oczyszczono za pomocą komercyjnie dostępnego zestawu RNeasy MinElute (QIAGEN). Stężenie i jakość RNA określono za pomocą spektrometru NanoDrop One. Uzyskaną wartość współczynnika λ_{260} stosowano do przeliczenia stężenia uzyskanego RNA na wartość μ M według wzoru:

$$c[\mu M] = \frac{\lambda_{A260} * 10^6}{\epsilon}$$

gdzie: c – stężenie RNA w μ M, λ_{A260} – współczynnik absorpcji ssRNA przy długości fali $\lambda_{260\text{nm}} = 1$ dla ssRNA = 40 μ g/ml, ϵ – współczynnik ekstynkcji dla cząsteczki vRNA8

W celu potwierdzenia integralności RNA, produkt transkrypcji *in vitro* rozdzielono na 1% żelu agarozowym. Jako standard długości zastosowano RNA RiboRuler High Range RNA Ladder. Próbkę RNA oraz standard długości zdenaturowano w mieszaninie obciążającej (3 minuty w 70°C) i schładzano przez (3 minuty w 4°C) przed nałożeniem na żel.

2.11. Fałdowanie vRNA8 w warunkach *in vitro*

Fałdowanie RNA *in vitro* przeprowadzono według następującej procedury. 4,5 μ l roztworu RNA (2pmol) w H₂O podgrzewano w termobloku przez 5 minut w

80°C. Po wyłączeniu termobloku monitorowano temperaturę za pomocą termometru. Po osiągnięciu temperatury 50°C dodawano 4,5 µl 2x stężonego buforu do fałdowania RNA i inkubowano do momentu osiągnięcia przez próbkę 37°C.

2.12. Fałdowanie vRNA8 w lizacie komórek MDCK

Fałdowanie vRNA8 przeprowadzono w lizacie komórek MDCK, które były zainfekowane wirusem grypy. Każda reakcja fałdowania składała się z 180 µl lizatu wzbogaconego w 5 pmol transkryptu vRNA8. Przed dodaniem do lizatu, RNA było denaturowane przez 5 minut w 80°C. Do reakcji fałdowania dodawano 50U inhibitora Rnaz (Promega). Fałdowanie przeprowadzano w czasie 30 minut w 37°C.

2.13. Mapowanie chemiczne

Mapowanie vRNA8 w warunkach in vitro odczynnikiem DMS

RNA po fałdowaniu w warunkach *in vitro* było mapowane odczynnikiem DMS. Stężenie końcowe odczynnika zostało wyznaczone eksperymentalnie i wynosiło 0,25%. Mapowanie przeprowadzano w objętości 10 µl na którą składało się 9 µl RNA rozpuszczonego w wodzie oraz 1 µl 10x stężonego odczynnika mapującego DMS rozpuszczonego w alkoholu etylowym. Do próbki kontrolnej dodano 1 µl etanolu. Reakcję inkubowano przez 15 minut w 37°C, a następnie zatrzymano poprzez dodanie β-merkaptioetanolu do ostatecznego stężenia 0.04M i schłodzono na lodzie. Natychmiastowo po zatrzymaniu reakcji, RNA wytrącano według standardowej procedury strącania kwasów nukleinowych.

Mapowanie vRNA8 w warunkach in vitro odczynnikami NAI oraz NMIA

Reakcję mapowania vRNA8 *in vitro* przeprowadzono w 10 µl, poprzez dodanie 1 µl 10x stężonego odczynnika SHAPE (NAI lub NMIA). Do próby kontrolnej dodawano 1 µl DMSO. W przypadku NAI końcowe stężenie odczynnika wynosiło 50 mM. Reakcję inkubowano przez 15 minut w 37°C, a następnie zatrzymano poprzez dodanie 1 µl stężonego (1M) DTT i schłodzono na lodzie. W przypadku mapowania z odczynnikiem NMIA stosowano stężenie ostateczne wynoszące 4

mM. Reakcje inkubowano przez 40 minut w 37°C. RNA wytrącano natychmiast po reakcji według standardowej procedury strącania kwasów nukleinowych.

Mapowanie lizatów komórkowych odczynnikiem DMS

Po reakcji fałdowania RNA w lizacie komórkowym, przeprowadzono reakcję mapowania siarczkiem dimetylu – DMS. Do 180 µl lizatu dodano 20 µl 10x stężonego odczynnika DMS rozcieńczonego w absolutnym alkoholu etylowym, tak aby ostateczne stężenie odczynnika DMS wynosiło 0,4%. Próbę kontrolną traktowano analogicznie, z zastosowaniem czystego alkoholu etylowego. Ostateczne stężenie odczynnika DMS zostało określone empirycznie w czasie wykonania eksperymentu optymalizacyjnego z zastosowaniem stężeń DMS w zakresie od 0,25%-0,6%. Reakcję inkubowano przez 5 minut w RT, a następnie zatrzymano poprzez dodanie β-merkaptoetanolu do ostatecznego stężenia w mieszaninie wynoszącego 0.04M.

Mapowanie lizatów komórkowych odczynnikiem 1M7

Po reakcji fałdowania RNA w lizacie komórkowym, przeprowadzono reakcję mapowania chemicznego z zastosowaniem odczynnika 1M7 (ang. 1-Methyl-7-nitroisatoic anhydride). Do 180 µl lizatu dodawano 20 µl 10x stężonego odczynnika 1M7 rozpuszczonego w bezwodnym DMSO. Do kontroli dodano 20 µl bezwodnego DMSO. Reakcję inkubowano przez 5 minut. Ostateczne stężenie 1M7 zostało wyznaczone eksperymentalnie w czasie eksperymentu wstępnego z zastosowaniem stężeń ostatecznych w zakresie od 10-25 mM. Ostatecznie wybrano stężenie 25 mM.

Reakcja enzymatycznej degradacji DNA we frakcji całkowitego RNA po reakcji mapowania w lizacie komórkowym

Od razu po wykonaniu reakcji mapowania natychmiastowo przystąpiono do oczyszczania z zastosowaniem zestawu RNeasy MinElute (QIAGEN). Następnie Izolat RNA był poddany trawieniu z pozostałości genomowego DNA za pomocą enzymu DNazy I (Ambion). Reakcję przeprowadzano w objętości 40 µl poprzez zmieszanie całkowitego RNA (~ 8 µg) z 1x stężonym buforem oraz 8U DNazy I. Reakcję inkubowano przez 30 minut w 37°C, a następnie inaktywowano enzym poprzez inkubację w 75°C przez 10 minut w obecności 2,5 mM EDTA. RNA

wytrącano z mieszaniny reakcyjnej według standardowej procedury. W celu kontroli czystości i integralności RNA, 1 µl izolatu nakładano rozdzielano elektroforetycznie na 1% żelu agarozowym.

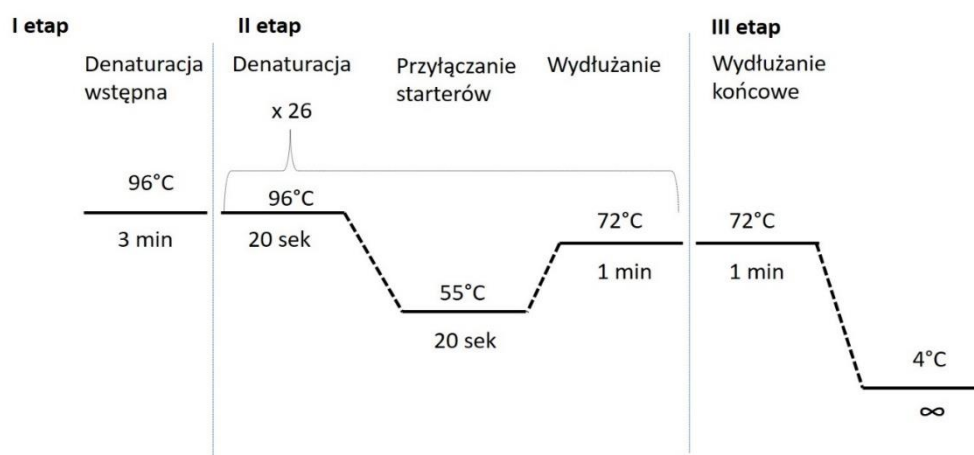
Mapowanie chemiczne RNA metodą SHAPE (ang. Selective 2'-OH hydroxyl acylation analyzed by primer extension). Odwrotna transkrypcja z użyciem starterów znakowanych fluorescencyjnie

W celu przepisania RNA na komplementarny cDNA przeprowadzono reakcję odwrotnej transkrypcji z zastosowaniem enzymu odwrotnej transkryptazy (SuperScript III). Dla próby reakcyjnej mapowanej odczynnikiem chemicznym zastosowano starter znakowany barwnikiem 6-JOE, natomiast do próby kontrolnej zastosowano starter znakowany 5-FAM. W celu odczytu mapowań dla całej cząsteczki vRNA8 zaprojektowano 3 startery (1_8, 2_8 i 3_8) w odstępach wynoszących ok. 300 nt. Reakcję przeprowadzono dwuetapowo. Pierwszy etap reakcji składał się z 24 µl mieszaniny na którą składały się: matryca RNA, 1x stężony bufor FS, 3 µM starter 6-JOE (dla reakcji) lub 5-FAM (dla kontroli) oraz 15 mM roztwór dNTP (deoksyrybonukleotydów). Mieszaninę inkubowano w 90 °C przez 3 minuty, 55°C przez 10 minut oraz schłodzono na lodzie przez 3 minuty. W drugim etapie reakcji dodano 6 µl mieszaniny składającej się z 1x stężonego buforu FS, 150 mM DTT, 60 U inhibitora RNaz oraz 300 U enzymu SuperScript IV. Reakcję inkubowano przez 20 minut w 55°C i inaktywowano przez 10 minut w 80°C. Następnie reakcje oczyszczono poprzez wytrącanie kwasów nukleinowych w etanolu według standardowej procedury.

2.14. Przygotowanie drabinek sekwencyjnych

W celu identyfikacji zatrzymań odwrotnej transkryptazy przygotowano drabinki sekwencyjne dla każdego startera. Matrycę dla drabinek stanowił produkt PCR vRNA8 szczepu A/California/04/2009. Reakcję sekwencjonowania przeprowadzono z zastosowaniem komercyjnie dostępnego zestawu USB®Thermo Sequenase Cycle Sequencing Kit. Mieszanka reakcyjna składała się z 1x stężonego buforu, 5 μ M jednego ze starterów: 1_8, 2_8, 3_8 znakowanego barwnikiem 5-ROX lub 6-TAMRA, 4 μ l dideoksy mix ddATP lub ddGTP, 1,8 z 2U enzymu sekwenazy oraz 150 ng matrycy PCR, a całość uzupełniono wodą do objętości 8,2 μ l.

Profil temperaturowy reakcji:



Produkt sekwencjonowania był następnie strącany na lodzie poprzez dodanie 3,2 μ l mieszaniny strącającej (0,4 M octan sodu, 13 mM EDTA, 0,05 mg/ml glikogenu i 20 μ l etanolu). Mieszaninę wirowano według standardowej procedury do strącania kwasów nukleinowych, przemyto 2x 70% etanolem, wysuszono i rozpuszczono w 20 μ l H₂O.

2.15. Rozdział produktów reakcji odwrotnej transkrypcji i drabinek sekwencyjnych z zastosowaniem elektroforezy kapilarnej

Produkty odwrotnej transkrypcji zarówno dla próby reakcyjnej, jak i kontrolnej, połączono wraz z drabinkami sekwencyjnymi. Na jedną reakcję mapowania dodano po 5 μ l każdej drabinki: ddA uzyskanej dla startera 6-TAMRA oraz ddG dla startera 5-ROX. Całość wysuszono w suszarce próżniowej i rozpuszczano w 10 μ l wody. Rozdziały prób z czterema barwnikami

fluorescencyjnymi były przeprowadzane w Pracowni Techniki Biologii Molekularnej UAM.

Rozdział czterech różnych barwników fluorescencyjnych w jednej kapilarze wiąże się ze zróżnicowaną migracją poszczególnych fragmentów DNA. W celu umożliwienia integracji wszystkich kanałów względem siebie na etapie analizy bioinformatycznej, przeprowadzona była dodatkowa reakcja sekwencjonowania, analogiczna do tej opisanej w punkcie *Przygotowanie drabinek sekwencyjnych*, z tą różnicą, że dla każdego startera 5-ROX oraz 6-TAMRA stosowano mix z tym samym dideoksynukleotydem - ddC. Ustawienia parametru przesunięcia pików dla odpowiednich kanałów w celu integracji wszystkich odczytów dla poszczególnych barwników (ang. *Mobility Shift*) przeprowadzono zgodnie z procedurą opisaną w Vasa et al.¹³⁴. Kalibrację migracji barwników wykonano w zakładce *Mobility shift* → *Cubic* → *New* → *Create by sliding traces*. Poszczególne odczyty z kanałów były następnie dopasowywane względem siebie. Kalibrację migracji wykonano dla każdego startera oddzielnie. Zapisany algorytm przesunięcia był implementowany na początkowym etapie analizy rozdzielów kapilarnych.

2.16. Analiza bioinformatyczna rozdzielów z zastosowaniem programu ShapeFinder

Analizę produktów odwrotnej transkrypcji i identyfikacji zatrzymań odwrotnej transkryptazy przeprowadzano z zastosowaniem programu ShapeFinder¹³⁴. Po dopasowaniu odczytu pików do linii bazowej (*Fitted Baseline Adjust*) i ustawienia szerokości okna (*window width*), wprowadzano korekcję migracji barwników poprzez implementację, wybranej w zależności od analizowanego fragmentu vRNA, kalibracji *Mobility Shift* → *Cubic*. Kolejnym krokiem było zawężenie rejonu analizy poprzez odcięcie fragmentów o niskim sygnale oraz fragmentu, w którym było widoczne podbicie sygnału fluorescencji pochodzącego od startera użytego do reakcji (*Signal Decay Correction*). W przypadku użycia czterech barwników fluorescencyjnych może dojść do osłabienia sygnału fluorescencji w jednym z kanałów. W takim przypadku dokonywano korekcji wysokości sygnałów wykonywano w zakładce *Scale Factor*, poprzez obniżanie lub podwyższanie ich wysokości w zakresie od 0.1 do 1, gdzie 1 to maksymalna wartość podniesienia

piku, którego wysokość spada wraz z wprowadzoną wartością parametru. Ostatnim etapem jest wprowadzenie sekwencji analizowanej cząsteczki RNA (sekwencja vRNA8 (-) A/California/04/2009), a także wskazanie drabinek ddNTP (ddA – TAMRA, ddG – ROX). Dane te są wprowadzane w zakładce *Align and Integrate*. Po implementacji wszystkich parametrów przystępowano do właściwej analizy rozdziałów. Ewentualne niedopasowania pików względem sekwencji referencyjnej przeprowadzano poprzez edycję manualną, do momentu uzyskania pełnego dopasowania produktów rozdziału do fragmentu sekwencji vRNA8. Ostatnim etapem analizy było wygenerowanie pliku z reaktywnościami dla poszczególnych nukleotydów.

Analiza i normalizacja danych w programie Excel

Plik z reaktywnościami wygenerowany w programie ShapeFinder był następnie poddany normalizacji z zastosowaniem metody normalizacji 2-8%¹⁹⁰. Metoda ta polega na wybraniu 10% najwyższych reaktywności, z których następnie wyklucza się 2% najwyższych reaktywności, a dla pozostałych 8% obliczana jest średnia. Następnie reaktywności dla wszystkich nukleotydów są normalizowane poprzez podzielenie ich wartości przez uzyskaną średnią. Znormalizowane reaktywności większe lub równe 0,7 uznawane były jako wysokie, te mieszczące się w zakresie 0,7-0,5 jako średnie, natomiast te poniżej 0,5 jako słabe.

2.17. Predykcja struktury drugorzędowej vRNA8 mapowanej w lizacie komórkowym

Predykcję struktury drugorzędowej przeprowadzono w programie RNAStructure w wersji 6.2. W pierwszej kolejności przewidziana była struktura MFE. W tym celu wprowadzono dwa pliki zawierające dane eksperymentalne. Pierwszy plik (rozszerzenie .shape) zawierał uśrednione (dla wszystkich powtórzeń biologicznych i technicznych) reaktywności pochodzące z eksperymentu mapowania odczynnikami 1M7. Dla nukleotydów o nieznannej wartości mapowania wprowadzono wartość -999, która jest interpretowana przez program jako brak danych. Ustawienia parametru *Slope* wynosiło 1.8 (kcal/mol), natomiast parametru *Intercept* wynosiło -0.6 (kcal/mol). Wielkość okna ustawiona

została na 15. W drugim pliku (rozszerzenie .con) wprowadzono dane mapowania DMS w postaci numerów nukleotydów o wysokiej reaktywności ($\geq 0,7$). W przypadku struktur generowanych bez zastosowania dystansu parowania zasad (algorytm *Maximum Pairing Distance*) do pliku .con wprowadzono także pary zasad dla motywu *panhandle*. W przypadku struktury fałdowanej z zastosowaniem dystansu parowania zasad do 150 nt, rejon *panhandle* był wyłączony z obszaru analizy. Wyłączenie rejonu *panhandle* z obszaru analizy wykonano poprzez wprowadzenie do pliku .con wymuszenia rejonu 1-13 oraz 878-890 jako ssRNA.

W celu obliczenia prawdopodobieństwa parowania zasad wykorzystano algorytmu *Partition Function*. Do obliczenia wykorzystano te same dane, które posłużyły do wygenerowania struktury. Do wygenerowania struktury MEA wykorzystano obliczone prawdopodobieństwa parowania zasad (plik z rozszerzeniem .pfs). W tym celu wykorzystano algorytm *MaxExpect: Predict RNA MEA structure*.

2.18. Predykcja struktury drugorzędowej vRNA mapowanych w warunkach *in virio* oraz *in cellulo*

Predykcję struktury drugorzędowej przeprowadzono w programie RNAStructure w wersji 6.2. Do przewidywania struktury MFE wprowadzono dwa pliki. Pierwszy z nich zawierał (rozszerzenie .shape) uśrednione dla wszystkich powtórzeń biologicznych reaktywności pochodzących z eksperymentu mapowania odczynnikiem NAI. Dla nukleotydów o nieznannej wartości mapowania wprowadzono wartość -999, która jest interpretowana przez program jako brak danych. Ustawienia parametru *Slope* wynosiło 1.8 (kcal/mol), a parametru *Intercept* wynosiło -0.6 (kcal/mol). Wielkość okna ustawiono na 15. W drugim pliku (rozszerzenie .con) wprowadzono informacje odnośnie mapowania DMS w postaci numerów nukleotydów o silnym mapowaniu chemicznym ($\geq 0,85$). Fałdowanie struktur vRNA wykonywano z zastosowaniem dystansu parowania zasad (algorytm *Maximum pairing distance*). Rejon *panhandle* był wyłączony z obszaru analizy. Wyłączenie rejonu *panhandle* z obszaru analizy wykonano poprzez wprowadzenie do pliku .con wymuszenia rejonu promotora (*panhandle*) jako ssRNA.

Prawdopodobieństwo parowania zasad obliczono na podstawie tych samych danych eksperymentalnych (algorytm *Partition function*). Uzyskany plik z wartościami prawdopodobieństwa wprowadzono do RNAStructure (algorytm *MaxExpect: Predict RNA MEA structure*) w celu wygenerowania struktury MEA.

Wizualizacja struktury drugorzędowej RNA

W programie RNAStructure został wygenerowany plik zawierający informację o wyliczonej strukturze drugorzędowej. Plik w formacie dotbracket (.db) zawierający sekwencję wraz z informacją o strukturze był wizualizowany w programie PseudoViewer (wer. 2.5) oraz VARNA^{210–212}.

2.19. Hodowla linii komórkowych MDCK oraz A549

Linia komórkowa MDCK (Merck) oraz A549 (ATCC) były przechowywane w ciekłym azocie. Do mrożenia komórek stosowano standardowe medium hodowlane suplementowane krioprotektantem (5% DMSO). Probówki z roztworem komórek bezpośrednio po rozmrożeniu odwirowano z pozostałości DMSO (3 minuty 1000 rpm). Odwirowany roztwór komórek zawieszono w medium suplementowanym 30% FBS i wysiano na szalkę hodowlaną. Po 24 godzinach medium zmieniono na standardowe medium hodowlane. Standardową hodowlę komórek prowadzono w inkubatorze w temperaturze 37°C z 95% wilgotnością powietrza i 5% CO₂. Przed wykonaniem pasażu, monowarstwę komórek przemywano dwukrotnie 1x PBS, a następnie dodawano roztwór trypsyny i inkubowano przez 3-5 minut w 37°C. Po tym czasie trypsyna była inaktywowana pożywką hodowlaną, a część komórek wysiewano ponownie na nową szalkę hodowlaną. Linia komórkowa była pasażowana na 10 cm² szalkach hodowlanych, w rozcieńczeniach w zakresie od 1:2 do 1:10.

2.20. Propagacja wirusa w linii komórkowej MDCK

Propagację wirusa grypy szczepu A/California/04/2009 prowadzono w linii komórkowej MDCK. Dzień przed infekcją komórki były wysiane w ilości 3x10⁶ na szalkę 10 cm². Do infekcji stosowano stężenie wirusa wynoszące MOI 0.01 (ang. Multiplicity of Infection), co oznacza, że 1 cząsteczka wirusa przypadała na ok. 100 komórek. Przed dodaniem rozcieńczenia wirusa, monowarstwę komórek

przemyto dwukrotnie 1x PBS. Następnie do komórek dodano 5 ml rozcieńczenia wirusa w medium infekcyjnym i inkubowano w RT przez 1 godzinę na platformie kołyszącej. Po tym czasie usunięto medium infekcyjne, dodano medium post-infekcyjne z trypsyną infekcyjną (TPCK) i szalki umieszczono w inkubatorze w temperaturze 33°C. Po zaobserwowaniu efektu cytotoksycznego (48-72 godziny) zbierano supernatant, który wirowano 3 minuty w 1000 rpm. Supernatant zawierający wirusa porcjowano i mrożono w -80°C.

2.21. Test immunofluorescencji

W celu określenia stężenia wirusa wykonano test immunofluorescencji FFA (ang. Focus Forming Assay). Roztwór wirusa po propagacji był testowany w trzech niezależnych powtórzeniach. Na 24 godziny przed infekcją, komórki wysiano na płytkę 96-dołkową w stężeniu $2 \cdot 10^4$ /dołek. W dniu infekcji, przygotowano seryjne 10-krotne rozcieńczenia roztworu wirusa. Następnie monowarstwę komórek przemyto dwukrotnie 1x roztworem PBS i dodano po 50 µl rozcieńczenia wirusa na dołek. Po godzinnej inkubacji w RT usunięto medium infekcyjne i dodano medium post-infekcyjne z dodatkiem 3% roztworu mikrocelulozy (Avicel)*. Po 16 godzinach inkubacji w 33°C na każdy dołek dodano 100 µl roztworu do utrwalania komórek i inkubowano przez 20 minut w RT. Po tym czasie usunięto roztwór, komórki przemyto 2-3 razy roztworem PBSu, dodano 100 µl buforu blokującego i płytki umieszczono w 37°C na 30 minut. Następnie wykonano barwienie przeciwciałami nakierowanymi na białka wirusa grypy. W tym celu dodano 45 µl 1000-krotnego rozcieńczenia przeciwciała I-rzędowego i inkubowano przez 1-2 godziny w 37°C. Po tym czasie komórki ponownie przemyto 3-4 razy roztworem PBS i dodano po 45 µl roztworu przeciwciała II-rzędowego z dodatkiem DAPI (barwienie jąder komórkowych). Płytki inkubowano w 37°C przez 30-60 minut. Po tym czasie prowadzono obserwację pod mikroskopem fluorescencyjnym. Liczenie miana wirusa rozpoczęto od dołka, który posiadał ilość skupisk wirusa możliwych do policzenia (np. w 4 dołku, gdzie wirus był rozcieńczony 1000 razy). Następnie liczono średnią z trzech dołków i przeliczano miano wirusa ze wzoru:

$$FFU[ml] = (Ilość\ skupisk * 20) * rozcieńczenie\ wirusa$$

gdzie: FFU (ang. Focus Forming Unit) – ilość cząsteczek wirusa w jednostce na 1 ml

* Zastosowanie roztworu mikrocelulozy umożliwia obserwację infekcyjnych cząsteczek wirusa. Jest to możliwe dzięki wytworzeniu zęszczenia środowiska wokół komórek, przez co wirus nie jest w stanie migrować do dalszych komórek, a tym samym ich infekować. W efekcie infekuje on komórki leżące blisko siebie i tworzy skupiska zainfekowanych komórek. Dzięki temu można określić zdolność danego wirusa do efektywnej infekcji.

2.22. Zagęszczanie roztworu wirusa z zastosowaniem poduszki sacharozowej

Zagęszczenie roztworu wirusa wykonano poprzez nałożenie 1400 μ l supernatantu wirusowego na 140 μ l buforu do zagęszczania roztworu wirusa i wirowanie przez 6 godzin w 14 000 g (16 900g) w 4°C. Po tym czasie delikatnie usunięto roztwór z nad osadu, który następnie zawieszono w 180 μ l buforu do mapowania *in virio*.

2.23. Infekcja wirusem grypy linii komórkowej A549

Na 24 godziny przed infekcją wirusem grypy komórki wysiano na płytce 6-dolkowej w stężeniu $0,5 \cdot 10^6$ /dolek. Po osiągnięciu 80-90% konfluencji, komórki były infekowane wirusem w stężeniu MOI wynoszącym 0.01 i inkubowane przez 24 godziny w 33°C.

2.24. Infekcja wirusem grypy linii komórkowej MDCK

Komórki MDCK były wysiewane w ilości $3 \cdot 10^6$ /10cm² szalkę na 24 godziny przed infekcją. Infekcję wirusem przeprowadzano z MOI wynoszącym 0.01 i komórki inkubowano przez 24 godziny przed wykonaniem lizy.

2.25. Przygotowanie lizatu komórkowo-wirusowego

W celu przygotowania lizatu komórkowo-wirusowego wykonano lizę komórkową z zastosowaniem buforu lizującego według procedury Shatzkes et al.¹⁸⁹. Przed dodaniem buforu lizującego monowarstwę komórek dwukrotnie przemyto 1x roztworem PBS. Następnie na szalkę dodano 6 ml buforu lizującego

i inkubowano przez 5-10 minut, aż do czasu pełnej lizy komórkowej widocznej pod mikroskopem. Lizaty porcjowano i zamrożono w -80 °C.

Mapowanie chemiczne struktury RNA in virio

Protokół mapowania *in virio* opracowano na podstawie publikacji Dadonaite et al.¹¹⁵. Do protokołu wprowadzono zmiany.

Mapowanie odczynnikiem DMS in virio

Do 180 µl buforu do mapowania dodano 20 µl 4% rozcieńczenia DMS w etanolu do ostatecznego stężenia 0,4%. Do próbki kontrolnej dodano 20 µl czystego etanolu. Po 5 minutach inkubacji w 37°C. W celu zatrzymania reakcji mapowania dodano 20 µl 1M DTT i inkubowano przez dodatkowe 5 minut w 37°C, po czym przystąpiono do izolacji RNA,

Mapowanie odczynnikiem NAI in virio

Przed reakcją mapowania sporządzono 1 M roztwór odczynnika NAI w bezwodnym roztworze DMSO. Reakcję mapowania przeprowadzono w objętości 200 µl na którą składało się 180 µl roztworu wirusa w buforze do mapowania oraz 20 µl 1 M NAI. Do próbki kontrolnej dodano 20 µl DMSO. Po 10 minutach inkubacji w 37°C, reakcję zatrzymano przez dodanie 20 µl 1M DTT i próbki inkubowano przez 5 minut w 37°C, po czym przystąpiono do izolacji RNA.

Mapowanie chemiczne in cellulo

Protokół mapowania chemicznego komórek opracowano na podstawie dostępnych publikacji^{115,131,150}. Przed modyfikacją chemiczną komórki trypsynizowano z 0,5 ml trypsyny w 37°C przez 3-5 minut. Następnie dodano 1 ml medium hodowlanego, a komórki zebrano i zwirowano przez 5 minut 3000g. Osad komórkowy przemyto 1 ml roztworu PBS i ponownie odwirowano. Osad komórkowy zawieszono w 180 µl roztworu 1x PBS.

Mapowanie zainfekowanej wirusem grypy linii komórkowej A549 odczynnikiem DMS

Przed reakcją mapowania chemicznego przygotowano 10x stężony roztwór DMS w etanolu. Następnie do 180 µl roztworu komórek dodano 20 µl

rozcieńczenia DMS do ostatecznego stężenia 0,3%. Do reakcji kontrolnej dodano 20 µl etanolu. Reakcję zmieszano i inkubowano 5 minut w RT. Po tym czasie dodano 1 ml roztworu zatrzymującego reakcję i komórki odwirowano 3 minuty 3000 g. Przemycie roztworem zatrzymującym reakcję powtórzono trzykrotnie. Do osadu komórek dodano 675 µl roztworu 1x PBS i izolowano całkowity RNA z zastosowaniem Trizolu LS.

Mapowanie zainfekowanej wirusem grypy linii komórkowej A549 odczynnikiem NAI

Do 180 µl zawiesiny komórkowej dodano 20 µl 1M roztworu NAI w bezwodnym DMSO, do ostatecznego stężenia 100 mM. Do kontroli dodano 20 µl DMSO. Reakcję zmieszano i umieszczono w termobloku na 8 minut w 37°C. W celu zatrzymania reakcji, do zawiesiny komórek dodano 30 µl 1M DTT i wirowano 3 minuty 3000 g. Osad ponownie przemyto roztworem 160 mM DTT w PBS, zwirowano 3 minuty 3000 g. Osad zawieszono w 675 µl 1x PBS i całość reakcji wyizolowano według standardowej procedury odczynnikiem Trizol LS.

2.26. Izolacja RNA po eksperymencie mapowania *in virio* oraz *in cellulo*

Po przeprowadzeniu reakcji mapowania chemicznego *in virio* oraz *in cellulo* natychmiastowo przystąpiono do izolacji RNA. W tym celu zastosowano odczynnik Trizol LS, który umożliwia izolację RNA z fazy wodnej. Do 225 µl reakcji dodano 675 µl Trizolu LS, wymieszano i inkubowano przez 5 minut w RT. Następnie dodano 180 µl roztworu chloroform:alkohol izoamylowy, wymieszano i inkubowano przez 3 minuty w RT. Po widocznym rozdzieleniu faz, mieszaninę wirowano 12 000 g przez 15 minut w 4°C. Po delikatnym ściągnięciu fazy wodnej, dodano 450 µl izopropanolu i 0,5 µl glikogenu, wymieszano i inkubowano przez 10 minut w RT. Po tym czasie próby ponownie odwirowano (10 minut, 4°C, 12 000g), supernatant usunięto, a osad przemyto 2 razy 70% etanolem, każdorazowo wirując próby przez 5 minut (4°C, 12 000 g). Osad wysuszono w cieplarni, rozpuszczono w wodzie i zmierzono stężenie. Jakość izolatu sprawdzano poprzez rozdział elektroforetyczny na 1% żelu agarozowym z zastosowaniem standardu wielkości.

2.27. Reakcja enzymatycznej degradacji DNA w izolacie całkowitego RNA komórkowego i oczyszczanie

Reakcję enzymatycznego oczyszczania całkowitego RNA z zanieczyszczenia genomowym DNA przeprowadzono z zastosowaniem enzymu DNazy I. Do mieszaniny reakcyjnej o końcowej objętości 40 μ l dodano: 34 μ l RNA, 4 μ l 10x stężonego buforu oraz 4 U enzymu DNazy I. Reakcję inkubowano przez 30 minut w 37°C, następnie dodano EDTA do końcowego stężenia 2,5 mM i inaktywowano enzym poprzez inkubację przez 10 minut w 75°C. Reakcję oczyszczano na kolumnach RNeasy Mini Kit według protokołu producenta.

2.28. Pomiar stężenia kwasów nukleinowych

Pomiar ilościowo-jakościowy kwasów nukleinowych

W celu sprawdzenia czystości i stężenia kwasów nukleinowych stosowano pomiar spektrofotometryczny UV/Vis z zastosowaniem spektrofotometru NanoDrop One. Pomiar ten umożliwia również wykrycie kontaminacji próbki białkami, fenolem lub resztkami soli (np. izotiocyanian guanidyny), które mogą pozostać w próbce po izolacji. Jako czyste próbki uważane są te, które posiadają współczynnik absorbancji A260/A280 równy \sim 1.8 dla DNA lub 2.0 dla RNA, a także współczynnik A260/A230 mieszczący się w wartościach pomiędzy 1.8-2.0. Próbki, które wykazywały odstępstwa od tych wartości, przed wykonaniem reakcji enzymatycznych były dodatkowo wytrącane w etanolu.

Pomiar ilościowy kwasów nukleinowych z zastosowaniem fluorymetru Qubit

W przypadku próbek, które mają bardzo niskie stężenie kwasów nukleinowych – np. próbki po izolacji RNA z zagęszczonego roztworu wirusa oraz próbek których stężenie musiało być bardzo dokładnie oszacowane (np. bibliotek do sekwencjonowania NGS) stężenia były określane za pomocą fluorymetru Qubit. Pomiar z zastosowaniem fluorymetru ma przewagę nad standardową procedurą spektrofotometryczną, ponieważ nie uwzględnia w pomiarze widma pochodzącego od kontaminacji reakcji innymi kwasami nukleinowymi np. starterami użytymi do reakcji, bądź trójfosforanami nukleotydów. Pomiar wykonywany był według procedury producenta z zastosowaniem zestawów do

pomiaru RNA, DNA o różnych zakresach wykrywanego stężenia. Jakość RNA po izolacji z komórek sprawdzano dodatkowo z zastosowaniem zestawu RNA IQ, dzięki któremu można było określić poziom integracji RNA w próbce w skali RIN (ang. RNA Integrity) równej od 1-10, gdzie 1 oznacza próbę zdegradowaną, a 10 próbę niezdegradowaną. Do eksperymentów stosowano RNA, którego pomiar RIN mieścił się w zakresie od 8 do 10.

2.29. Analiza obecności białek wirusowych z zastosowaniem metody Western-Blot

Lizat z komórek MDCK (przygotowany z zastosowaniem buforu CL) zmieszano z buforem lizującym 2x stężonym w objętości 1:1 i inkubowano przez 5 minut w 95°C w celu denaturacji białek. Próbkę rozdzielono na elektroforezie kapilarnej na 12% żelu PAA z dodatkową warstwą 5% żelu zagęszczającego. Elektroforezę prowadzono przez 15 minut (80V) dla żelu zagęszczającego, a następnie 50-60 min (180V) dla żelu rozdzielającego. Kolejnym etapem był transfer białek z żelu na membranę nitrocelulozową. Bibułki wraz z membraną namaczano przez 2-3 minuty w 1 x buforze do transferu (100 ml). Następnie, w dolnej kasetce (stanowiącej anodę) aparatu do pół-suchej metody blottingu (Trans-Blot Turbo, BioRad), umieszczono kolejno: warstwę bibuły, membranę, żel, warstwę bibuły i zamknięto górną kasetką (stanowiącą katodę). Tak przygotowaną membranę na czas transferu (7 minut) umieszczono w aparacie z ustawieniami natężenia prądu wynoszącą 1.3A i różnicą potencjału - 25V.

Następnie przeprowadzono znakowanie białek przeciwciałami. Wszystkie inkubacje przeprowadzono w RT z wytrząsaniem. W tym celu membranę inkubowano przez 1 h w 10 ml 5% mleka PBS-T. Po inkubacji dodano 10 µl przeciwciał pierwszorzędowych (1:1000) i inkubowano przez 1 h. Membranę przepłukano trzykrotnie roztworem PBS-T, każdorazowo inkubując ją 5 minut. Do membrany dodano 10 ml 5% mleka w PBS-T oraz 1 µl przeciwciał drugorzędowych (1:10000) i inkubowano 1h. Membranę ponownie trzykrotnie przemyto roztworem PBS-T. Detekcję białek wykonano z zastosowaniem komercyjnie dostępnego zestawu Westar Sun. W skrócie, na membranę równomiernie nakropiono 2 ml roztworu mieszaniny Luminol Enhancer oraz

mieszaniny Peroxide (zmieszane w stosunku 1:1) i inkubowano przez 2 minuty. Następnie, wykonano wizualizację membrany i zdjęcie za pomocą aparatu ChemiDoc XRS+ System (BioRad). Czas ekspozycji ustalano w zależności od stopnia wybarwienia białek. W sumie przygotowano trzy żele oraz trzy membrany, które posłużyły do wizualizacji białek za pomocą przeciwciał nakierowanych na białko HA, NP oraz aktyne.

2.30. Analiza obecności białek w lizatach komórkowo-wirusowych

W celu potwierdzenia obecności białek w lizatach komórkowych infekowanych wirusem przeprowadzono barwienie z barwnikiem błękitem kumasyny (Coomasie blue). Próbkę lizatów, a także rozdział elektroforetyczny opisane zostały w punkcie dotyczącym metody Western-Blot. Żel z rozdzielonymi białkami utrwalono w mieszaninie buforu utrwalającego przez 15-30 minut w RT. Następnie żel inkubowano z buforem barwiącym z błękitem kumasyny przez 45-60 minut. Ostatni etap, czyli inkubacja w mieszaninę odbarwiającej, był przeprowadzany przez noc.

2.31. Analiza vRNA z zastosowaniem metody ilościowego real-time PCR w czasie rzeczywistym

W celu określenia powodzenia infekcji komórek stosowano reakcję ilościowego PCR w czasie rzeczywistym. Standardowo wykonywano reakcje ilościowe w celu określenia stężenia 8 segmentu vRNA A/California/04/2009 w komórkach.

Przygotowanie matrycy do reakcji PCR w czasie rzeczywistym

500 ng całkowitego RNA przeznaczono do reakcji enzymatycznej degradacji DNA przy użyciu enzymu DNazy I. Do mieszaniny reakcyjnej składającej się z RNA w objętości 8,5 µl dodano 1 µl 10x stężonego buforu oraz 2U enzymu DNazy I. Reakcję inkubowano przez 30 minut w 37°C. Następnie do reakcji dodano EDTA do końcowego stężenia 2,5 mM i reakcję inaktywowano przez 10 minut w 75°C.

Do reakcji odwrotnej transkrypcji pobierano 1 µl RNA po reakcji z DNazą i przeprowadzano dwuetapową reakcję odwrotnej transkrypcji. Pierwszy etap

reakcji (5 μ l) składał się z mieszaniny buforu FS (1x stężony), 2 μ M startera (7VRT dla segmentu 7 lub 8VRT dla segmentu 8 vRNA), oraz 1 μ l RNA po reakcji enzymatycznej degradacji DNA. Reakcję inkubowano w 90°C przez 3 minuty, 55°C przez 10 minut, a następnie schłodzono na lodzie przez 3 minuty. Do reakcji dodano 5 μ l mieszaniny składającej się z 10 U inhibitora Rnaz (RNasin), 12 mM dNTP, 1x bufor FS, 60 mM DTT oraz 50U enzymu SuperScript III. Reakcję inkubowano w 55°C przez 50 minut, a następnie inaktywowano enzym w 70°C przez 15 minut.

Przygotowanie standardów do reakcji PCR w czasie rzeczywistym

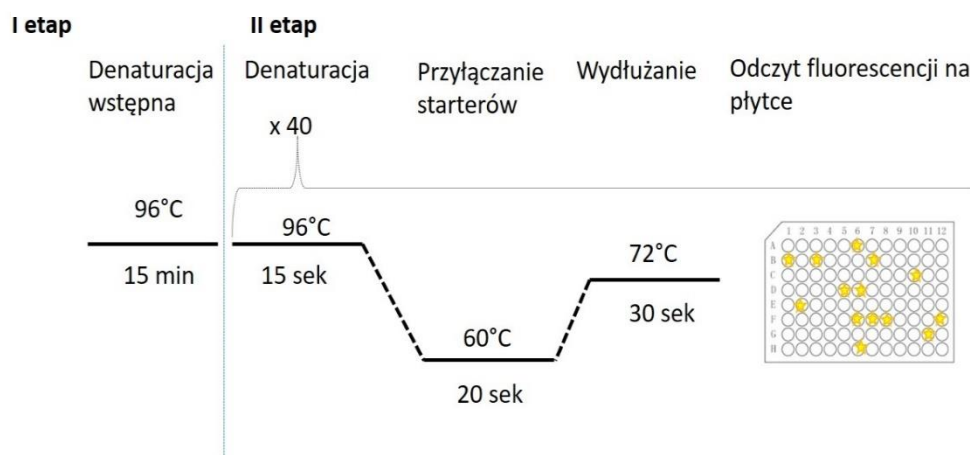
W celu przygotowania standardów do reakcji PCR w czasie rzeczywistym, przeprowadzono reakcję transkrypcji *in vitro* vRNA8 na matrycy pUC19, a następnie reakcję odwrotnej transkrypcji z enzymem SuperScript III analogicznie do tej opisaną w punkcie 2.45. Reakcję RT wykonano w objętości 50 μ l. Do reakcji dodano 185 ng RNA. Zgodnie z tym przyjęto, że nierozcieńczony standard ma stężenie 3,7 ng/ μ l. Następnie wykonano 10-krotne rozcieńczenia standardów i przeliczono wartości na femtogramy. Wartości standardów w femtogramach zostały wprowadzone na płytkę w programie Bio-Rad CFX Maestro.

Reakcja PCR w czasie rzeczywistym

Startery do reakcji PCR w czasie rzeczywistym specyficzne dla cDNA segmentów vRNA7 i vRNA8 A/California/04/2009 zaprojektowano na podstawie dostępnych publikacji^{213,214}. Reakcję PCR wykonano w trzech powtórzeniach dla każdej próbki. Całkowita objętość reakcji wynosiła 10 μ l na którą składały się: 1 μ l matrycy cDNA, 5 μ M starter VTAG, 5 μ M starter 7VQR (dla 7 segmentu vRNA) lub 8VQR (dla 8 segmentu vRNA) oraz 1x Master Mix Hot FIREPol®EvaGreen® qPCR.

Reakcję przeprowadzano w termocyklerze CFX96TM Real-Time System (Bio-Rad). Odczyt fluorescencji miał miejsce po każdym cyklu.

Profil temperaturowy reakcji:



Analiza danych z reakcji PCR w czasie rzeczywistym

Pierwszy etap analizy danych z reakcji PCR w czasie rzeczywistym przeprowadzano w programie Bio-Rad CFX Maestro. Wygenerowany plik w formacie .pcrd analizowano pod względem jakościowym. Zgodnie z protokołem producenta, wartość odchylenia standardowego parametru C_q tj. liczby kopii/ μ l pomiędzy trzema powtórzeniami dla każdej próbki powinna wynosić maksymalnie 1,69. W przypadku, kiedy wartość ta jest wyższa należy odrzucić próbkę, której wartość C_q znacząco odbiega od pozostałych 2, tak, aby obniżyć wartość odchylenia standardowego. W tym celu wykonywana była ręczna edycja polegająca na wykluczeniu pojedynczych powtórzeń z analizy próbek. Następny etap analizy obejmował przeliczenie wartości $SQ Mean$ [fm/ μ l] na wartości w ng/ μ l dla każdej z próbek, dzięki czemu możliwe było przeliczenie ilości procentowej vRNA w próbce całkowitego RNA.

2.32. Reakcja odwrotnej transkrypcji indukującej mutacje w obecności jonów Mn^{2+} (MaP)

Reakcję odwrotnej transkrypcji typu SHAPE-MaP opracowano na podstawie opublikowanych protokołów^{131,143}.

Optymalizacja metody MaP na matrycy in vitro oraz in cellulo

Optymalizację metody MaP przeprowadzono z zastosowaniem matrycy *in vitro* oraz *in cellulo*. Do eksperymentów *in vitro* wykorzystano produkt transkrypcji vRNA8 (A/California/04/2009). Po potwierdzeniu integralności RNA transkrybowanego *in vitro*, przeprowadzono procedurę fałdowania i reakcję mapowania odczynnikami NMIA. Następnie przeprowadzono reakcję odwrotnej transkrypcji z zastosowaniem enzymu odwrotnej transkryptazy (SuperScript II). W celu potwierdzenia efektywności reakcji MaP zdecydowano się na wykonanie kontrolnej reakcji z enzymem SuperScript II zgodnie z protokołem producenta. Reakcje odwrotnej transkrypcji w eksperymencie *in vitro* wykonano na RNA zmapowanym odczynnikami NMIA oraz kontroli (nieotraktowanej odczynnikami). W przypadku próby z warunków *in cellulo* zastosowano RNA wyizolowany z infekowanych wirusem grypy komórek A549, które poddano mapowaniu chemicznemu odczynnikami NAI.

Standardową reakcję odwrotnej transkrypcji przeprowadzono w trzech etapach. W pierwszym etapie mieszaninę 9 μ l RNA (2 pmol), 1 μ l dNTP (10 mM) oraz 1 μ l startera 3.8 FAM (2 pmol) inkubowano przez 5 minut w 65°C i schłodzono na lodzie. Na drugim etapie dodano 4 μ l buforu FS (5x), 2 μ l DTT (0.1 M) oraz 1 μ l inhibitora RNaz (10 U) i inkubowano 2 minuty w 42°C. W ostatnim etapie dodano 1 μ l enzymu SuperScript II (200 U) i reakcję inkubowano 50 minut w 42°C i 15 minut w 70°C w celu inaktywacji enzymu.

Reakcję MaP przeprowadzono z zastosowaniem startera 3.8 FAM oraz vRNA8 otrzymanym *in vitro*. Reakcja była przeprowadzana w trzech etapach. W pierwszym etapie 10 μ l RNA (2 pmol) zmieszano z 1 μ l startera 3.8 FAM (2 pmol) i inkubowano 5 minut w 65°C. Po schłodzeniu na lodzie dodano 8 μ l 2.5 x stężonego buforu MaP i inkubowano 2 minuty w 42°C. Po dodaniu 1 μ l enzymu SuperScript II i inkubowano 3 godziny w 42°C i 15 minut w 70°C.

Reakcje odwrotnej transkrypcji zostały rozdzielone z zastosowaniem elektroforezy kapilarnej. Ze względu na zastosowanie pojedynczego fluoroforu (FAM), każda próbka była rozdzielana w osobnej kapilarze (Pracownia Technik

Biologii Molekularnej UAM). Próbki analizowano w programie PeakScanner (wer. 1.0).

Reakcje odwrotnej transkrypcji dla próbek z warunków *in cellulo* przeprowadzono analogicznie, za wyjątkiem zastosowanego startera oraz matrycy. Do reakcji dodano 1 µl startera Uni-12 (10 pmol) oraz matrycę całkowitego RNA (1000 ng). Następnie przeprowadzono reakcję PCR z zastosowaniem starterów HFA i HRA według protokołu opisanego w punkcie 0 (str. 189). Produkty PCR rozdzielano za pomocą elektroforezy na 1% żelu agarozowym.

Mapowanie vRNA in virio - reakcja odwrotnej transkrypcji typu MaP

W przypadku mapowania *in virio*, a więc czystego izolatu vRNA wirusa grypy, do reakcji zastosowano starter typu „losowy heksamer” (mieszanina losowych heksamerów). Reakcję wykonano w 3 etapach. W etapie pierwszym zmieszano 10 µl RNA (50-100 ng) z 1 µl startera (60 pmol) typu losowy heksamer (NEB). Mieszaninę inkubowano przez 5 minut w 65°C i dodano 2,5x stężony bufor MaP oraz 10 U inhibitora RNaz. Po 2 minutach inkubacji w 25°C dodano 200U enzymu SuperScript II. Reakcję inkubowano przez 25°C przez 10 minut, następnie przez 3 godziny w 42°C, po czym inaktywowano enzym w 70°C przez 15 minut. Produkty reakcji odwrotnej transkrypcji były następnie oczyszczane na kolumnie MicroSpin G-25 (GE Healthcare) według protokołu producenta. Następnie wykonano reakcję syntezy drugiej nici (str. 189).

Mapowanie vRNA in cellulo - reakcja odwrotnej transkrypcji typu MaP

W przypadku mapowania RNA w zainfekowanych wirusem grypy komórkach A549 zastosowano specyficzny starter Uni-12, który jest komplementarny do 3' końca wszystkich vRNA – rejonu *panhandle* wirusa grypy typu A²⁰². Reakcję RT przeprowadzono trzyetapowo. W pierwszym etapie do 9 µl RNA (~3 µg) dodano 1 µl startera Uni-12 (10 µM) i mieszaninę inkubowano przez 5 minut w 65°C. Po schłodzeniu próbki na lodzie, dodano 8 µl 2.5x buforu MaP oraz 10U inhibitora RNaz. Mieszaninę inkubowano przez 2 minuty w 42°C, po czym dodano 200 U enzymu SuperScript II. Reakcję inkubowano przez 3 godziny w 42°C oraz 15 minut w 70°C w celu inaktywacji enzymu. Reakcję oczyszczono na kolumnie

MicroSpin G-25. Produkt reakcji posłużył jako matryca do reakcji PCR z zastosowaniem uniwersalnych starterów (str. 189).

2.33. Reakcja syntezy drugiej nici

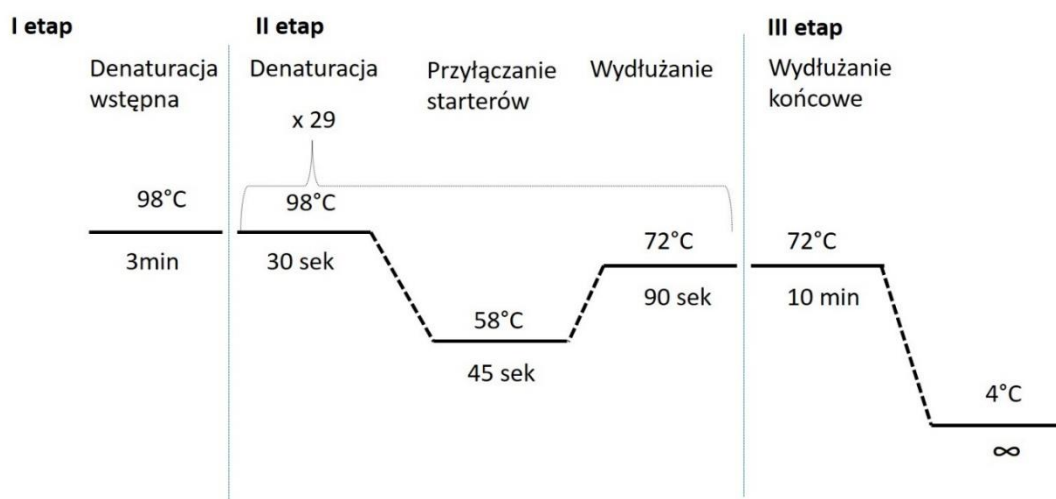
W przypadku reakcji RT z zastosowaniem losowego 6-meru uzyskano mieszaninę fragmentów cDNA o różnej długości. W celu przygotowania dwuniciowej matrycy do biblioteki do sekwencjonowania NGS przeprowadzono reakcję syntezy drugiej nici z zastosowaniem zestawu NEBNext® Ultra II Non-Directional RNA Second Strand Synthesis (NEB). Reakcję przygotowano w końcowej objętości 80 µl według protokołu producenta poprzez zmieszanie całej reakcji RT (20 µl) z 8 µl NEBNext Second Strand Synthesis Reaction Buffer, 4 µl enzymu NEBNext Second Strand Synthesis Enzyme Mix. Reakcję inkubowano w termocyklerze w 16 °C przez 2,5 h. Produkty reakcji oczyszczano z zastosowaniem kolumn PureLink PCR Micro.

Reakcja PCR z zastosowaniem uniwersalnych starterów dla wszystkich segmentów vRNA

W wyniku reakcji odwrotnej transkrypcji z zastosowaniem uniwersalnego startera Uni-12 uzyskano mieszaninę fragmentów cDNA odpowiadającą 8 cząsteczkom (segmentom) wirusowego RNA. Kolejnym etapem przygotowania matrycy do reakcji sekwencjonowania było przeprowadzenie reakcji PCR z uniwersalnymi dla wszystkich segmentów starterami HFA oraz HRA na podstawie protokołu opublikowanego przez *Illumina inc.*²¹⁵. Startery te były komplementarne do rejonów odpowiadających promotorowi wirusowej RNA-zależnej polimerazy RNA (RdRp) na 3' oraz 5' końcu każdego z segmentów vRNA²⁰¹. Oprócz sekwencji komplementarnej, na końcach starterów znajdują się dodatkowe sekwencje potrzebne na pierwszym etapie przygotowania biblioteki do sekwencjonowania. Szczegółowy opis jest zaprezentowany w rozdziale *Wstęp literaturowy*.

Reakcję PCR przygotowywano w objętości 50 μ l poprzez zmieszanie 10 μ l cDNA (1/5 oczyszczonej reakcji RT), 125 mM MgCl₂, 10 μ M startera HFA, 10 μ M startera HRA oraz 1x stężonego Phusion Master Mix.

Profil temperaturowy reakcji:



Podczas reakcji z zastosowaniem starterów HFA/HRA często powstają niespecyficzne produkty PCR. Jest to wynikiem tego, że startery są komplementarne na krótkim odcinku (12 i 13 nt) do końców cząsteczek vRNA. W celu oczyszczenia produktu PCR z powstałych niespecyficzności zastosowano selektywnie odcinające krótsze fragmenty kulki magnetyczne Ampure XP Beads. W celu optymalizacji oczyszczania zastosowano różne objętości kulek w stosunku do produktu PCR. Najlepsze rezultaty uzyskano przy zastosowaniu kulek w stosunku 0,5 objętości do objętości próbki. Po zmieszaniu 25 μ l kulek Ampure XP z 50 μ l reakcji PCR, dokładnie wymieszano całą reakcję i następnie inkubowano przez 5 minut na stole. Po usunięciu supernatantu, kulki przemyto dwa razy 85% etanolem, suszono 3 minuty i eluowano w 25 μ l H₂O przez 5 minut. Wydajność reakcji oczyszczania sprawdzono na 1% żelu agarozowym.

2.34. Reakcja PCR z zastosowaniem specyficznych starterów na wszystkie segmenty vRNA

W celu potwierdzenia obecności wszystkich segmentów vRNA przeprowadzono 8 reakcji PCR na matrycach cDNA z reakcji RT, ze starterami specyficznymi dla każdego segmentu vRNA. Reakcje PCR wykonano analogicznie do tej opisanej w punkcie 0. W każdej reakcji zastosowano parę

specyficznych dla każdego segmentu starterów. Sekwencje użytych starterów znajdują się w Tabeli 1. Po zmierzeniu stężenia, na 1% żel agarozowy nałożono ok. 150 ng każdego z produktów PCR.

2.35. Sekwencjonowanie produktów PCR metodą Sanger

Sekwencjonowanie produktów PCR metodą Sanger zlecane było Pracowni Techniki Biologii Molekularnej UAM. Standardowo do sekwencjonowania przygotowywano matrycę o stężeniu wynoszącym 5 ng/ μ l produktu na każde 100 pz. Do reakcji sekwencjonowania stosowano starter o stężeniu 10 μ M.

2.36. Analiza wyników sekwencjonowania metodą Sanger

Pliki w formacie ab.1 z wynikami sekwencjonowania typu Sanger były wizualizowane w programie BioEdit (wersja 7.0.9.0). W programie dokonywano ewentualnej korekcji odczytanej sekwencji poprzez wstawianie nukleotydu A-C-G-T w miejscu błędnie odczytanym jako N. Następnie uzyskaną sekwencję w formacie FASTA wprowadzano do programu GeneDoc, gdzie dokonywano przyrównania uzyskanych wyników sekwencjonowania względem sekwencji referencyjnej vRNA8.

2.37. Przygotowanie bibliotek DNA do wysokoprzepustowego sekwencjonowania nowej generacji (NGS)

Biblioteka fragmentów do sekwencjonowania NGS przygotowywana była za pomocą komercyjnie dostępnego zestawu Nextera XT DNA Library Prep. Wszystkie odczynniki były przechowywane i rozmrażane zgodnie z instrukcjami producenta.

Reakcja tagmentacji

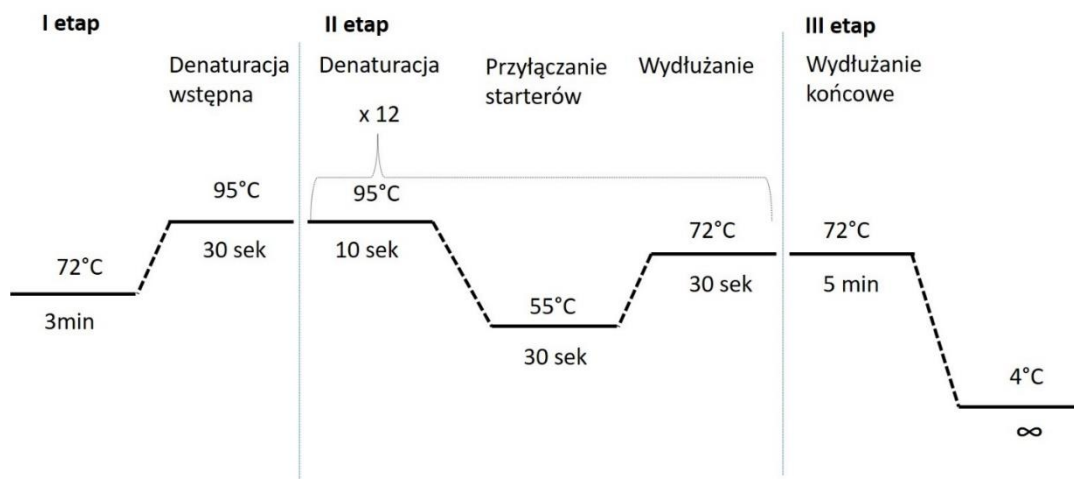
Reakcja enzymatycznej fragmentacji i dołączenia adapterów tzw. tagmentacji (ang. tagmentation) przeprowadzana była według protokołu producenta. W pierwszej kolejności przygotowano rozcieńczenie matrycy DNA do 0,2 ng/ μ l. Stężenie matrycy było kontrolowane poprzez pomiar na Qubit z zastosowaniem kitu HS QUBIT dsDNA. Do reakcji tagmentacji przygotowano 1 ng DNA, do którego dodano 10 μ l buforu TD (Tagment DNA Buffer) oraz 5 μ l mieszaniny enzymu ATM (Amplicon Tagment Mix). Całość delikatnie wymieszano pipetą,

zwirowano 280 g 1 minutę i umieszczono w termocyklerze na 5 minut w 55°C. Po schłodzeniu reakcji do 10°C, dodano 5 µl buforu inaktywującego enzym NT (Neutralize Tagment Buffer), zwirowano 280 g 1 minutę i inkubowano przez 5 minut w RT.

Reakcja PCR z dodaniem unikalnych indeksów

Następnym etapem przygotowania biblioteki była reakcja amplifikacji bibliotek wraz dodaniem specyficznych starterów do sekwencjonowania Illumina. W tym celu przeprowadzono selekcję kombinacji indeksów i5 oraz i7, które zostały zastosowane do każdej z próbek. Lista indeksów wraz z sekwencjami, które zostały wykorzystane dla poszczególnych próbek jest wyszczególniona w Tabeli 1 w rozdziale *Dodatek*. Reakcję amplifikacji bibliotek przeprowadzono poprzez dodanie do mieszaniny z poprzedniego etapu po 5 µl indeksu i5, 5 µl indeksu i7 oraz 15 µl mieszaniny enzymatycznej NPM (Nextera PCR Master Mix). Reakcję wymieszano delikatnie pipetą, zwirowano 1 minutę w 280 g i umieszczono w termocyklerze.

Profil temperaturowy reakcji:



Gotowe biblioteki zostały oczyszczone z zastosowaniem kulek Ampure XP beads. W przypadku bibliotek pochodzących z eksperymentu *in virio* zastosowano objętość kulek Ampure XP Beads w stosunku objętości wynoszącym 0,6 x. Objętość mieszaniny elucyjnej wynosiła 52,5 µl. Biblioteki pochodzące z eksperymentu *in cellulo* poddano selekcji fragmentów poprzez usunięcie frakcji krótkich fragmentów stosując objętość kulek 0,6x oraz długich fragmentów

stosując objętość kulek 0,15x. Elucję wykonano w 32,5 µl. Stężenie bibliotek mierzono za pomocą fluorymetru Qubit.

2.38. Rozdział fragmentów bibliotek z zastosowaniem bioanalyzera

Rozkład fragmentów oraz jakość biblioteki była testowana za pomocą czułego rozdzału elektroforetycznego z zastosowaniem chipów High Sensitivity DNA na bioanalyzerze Agilent 2100 Bioanalyzer. Roztwór żelu z barwnikiem użyty do rozdzału elektroforetycznego, a także przygotowanie chipu do rozdzału wykonano zgodnie z protokołem producenta. Na dołek nakładano 1 µl nierozcieńczonego produktu biblioteki.

2.39. Normalizacja bibliotek do sekwencjonowania NGS

Przygotowanie puli bibliotek fragmentów do sekwencjonowania wymaga obliczenia ich stężeń molowych. Do tych obliczeń potrzebny jest pomiar stężenia za pomocą fluorymetru QUBIT, a także określenie średniej długości fragmentów w próbce za pomocą rozdzału na bioanalyzerze.

Masa cząsteczkowa bibliotek (Library Molecular Weight) obliczana jest ze wzoru:

$$\text{Średnia długość DNA [pz]} \times 650 \text{ [g/mol na pz]} = \text{masa cząsteczkowa biblioteki} \\ \text{[g/mol=ng/nmol]}$$

Stężenie molowe obliczane jest ze wzoru:

$$\text{Stężenie próbki [ng/µl]} \times \text{masa cząsteczkowa biblioteki [ng/nmol]} = \text{stężenie} \\ \text{biblioteki [nmol/µl]}$$

Po obliczeniu stężeń bibliotek wyrażonych w nM przygotowane zostały rozcieńczenia bibliotek buforem elucyjnym EB (QIAGEN) do stężenia końcowego 10 nM, a następnie przygotowano mieszaninę bibliotek o ostatecznym stężeniu 2 nM. Następnie wykonano denaturację bibliotek. W tym celu mieszaninę bibliotek zmieszano z jedną objętością 0,2 M NaOH (10 µl), inkubowano przez 5 minut w RT i dodano 10 µl 200 mM Tris-HCl (pH 7,0). Kolejnym etapem było dodanie 970 µl zimnego (0°C) buforu hybrydizacyjnego HT1 (Illumina). Ostatecznie otrzymano mieszaninę zdenaturowanych bibliotek o ostatecznym stężeniu 20 pM.

2.40. Sekwencjonowanie z zastosowaniem sekwenatora Illumina NextSeq550

Sekwencjonowanie bibliotek wykonano na płytce o średniej przepustowości (Mid Output Flow Cell). Finalne stężenie bibliotek na tego płytce wynosi 1.5 pM. W celu wykonania ostatecznego stężenia bibliotek zmieszano mieszaninę 20 pM bibliotek z buforem HT1 w następującej proporcji: 97 μ l bibliotek z 1203 μ l buforu HT1. Całość (1300 μ l) nałożono na kartridż z odczynnikami do sekwencjonowania. Sekwencjonowanie typu paired-end (sekwencjonowanie z dwóch końców nici DNA) wykonano z zastosowaniem sekwenatora Illumina NextSeq550 oraz z odczynnikami pozwalającymi na uzyskanie długości odczytów 2x150 pz. Sekwencjonowanie zostało wykonane przez dr. Luizę Handschuh oraz dr. Jana Podkowińskiego (Pracownia Genomiki Instytutu Chemii Bioorganicznej PAN).

2.41. Analiza wyników sekwencjonowania NGS

Sekwenator NextSeq550 wykonuje automatyczny podział próbek na podstawie zastosowanych indeksów (ang. demultiplexing) oraz dodatkowo usuwa sekwencje adapterów w ostatecznej wersji pliku (ang. trimming). Sekwencjonowanie typu paired-end generuje osobne pliki odczytu dla obu końców cząsteczki, które na początkowym etapie analizy połączono w jeden plik. Następnie przeprowadzona została analiza jakościowa i ilościowa odczytów. Po mapowaniu wyników sekwencjonowania do sekwencji referencyjnych uzyskano wykresy pokrycia dla poszczególnych segmentów vRNA. Jako kontrolę dla warunków *in cellulo* wykorzystano cząsteczkę RNA o znanej strukturze drugorzędowej - 18S rRNA, który ulega wysokiej ekspresji w komórkach. Cząsteczka ta była sekwencjonowana jednocześnie z próbkami wirusowymi.

Analizę wyników sekwencjonowania przeprowadzono przy zastosowaniu narzędzia ShapeMapper2¹⁴. Program służy do analizy danych z sekwencjonowania metodą SHAPE-MaP oraz DMS-MaPseq. ShapeMapper2 wykonuje analizę profilu mutacji poprzez przyrównanie próby reakcyjnej (wywodzącej się z RNA mapowanego odczynnikiem chemicznym) do próby kontrolnej (wywodzącej się z RNA nietraktowanego odczynnikiem chemicznym).

Identyfikacja miejsc o skumulowanej liczbie mutacji posłużyła do obliczenia reaktywności poszczególnych nukleotydów. Jako mutacje identyfikowane były zmiany takie jak jednonukleotydowe lub wielonukleotydowe niesparowania, delecje czy insercje, a także wielonukleotydowe przerwy w odczytanej sekwencji dla próby reakcyjnej w stosunku do próby kontrolnej. W celu detekcji ustalono minimalne pokrycie badanej cząsteczki na 1000 odczytów. Całościową analizę bioinformatyczną wykonał dr Tomasz Woźniak (Instytut Genetyki Człowieka PAN). Analiza została przeprowadzona z zastosowaniem standardowych parametrów programu ShapeMapper2.

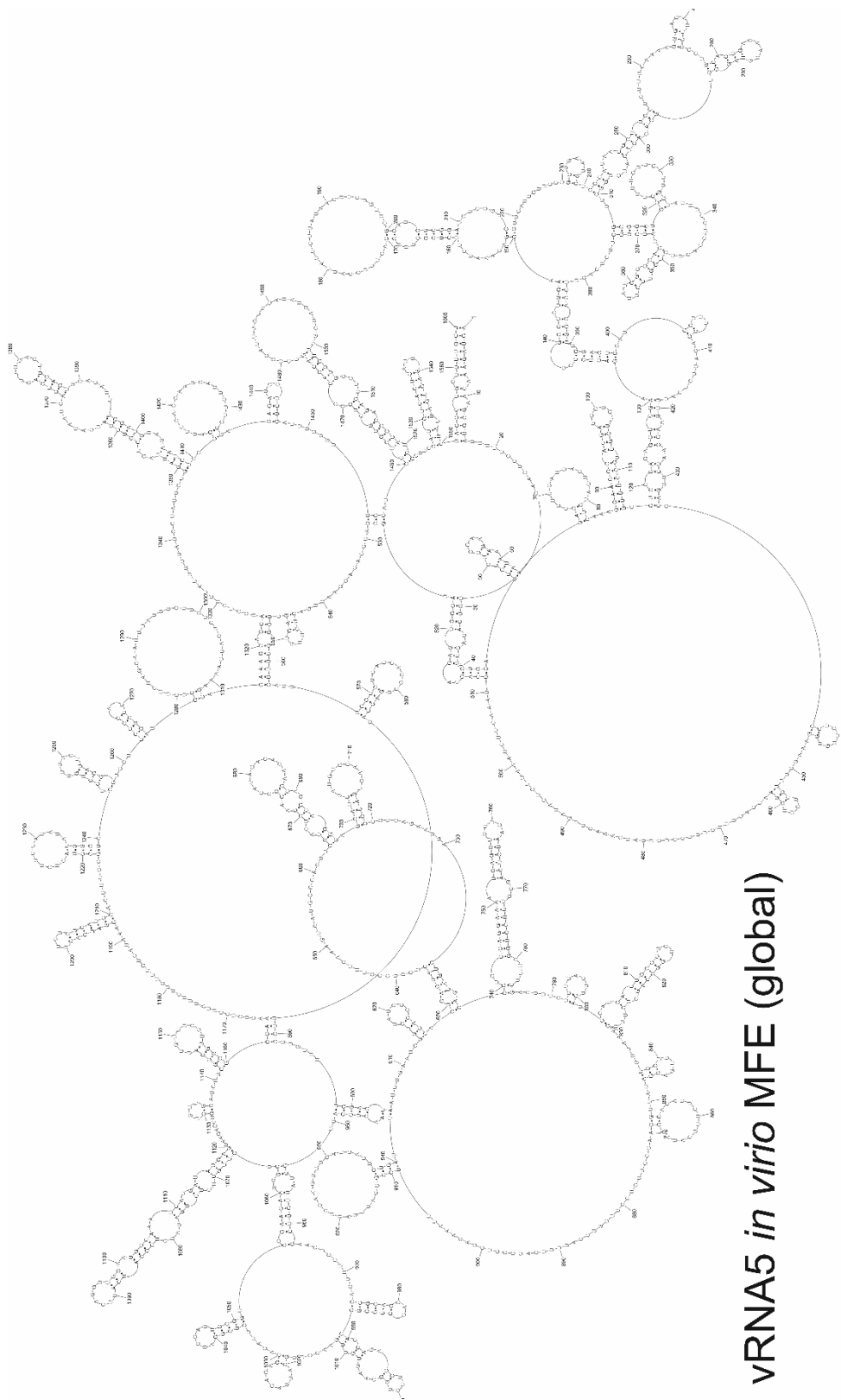
Dodatek

Tabela 10 Podsumowanie informacji dotyczących poszczególnych bibliotek fragmentów, które zostały przeznaczone do sekwencjonowania. Każda próbka posiada zestaw unikalnych indeksów, na podstawie których była identyfikowana. Do obliczenia stężenia biblioteki w nM niezbędny był pomiar stężenia (QUBIT) oraz identyfikacja średniej długości fragmentów (Bioanlizator).

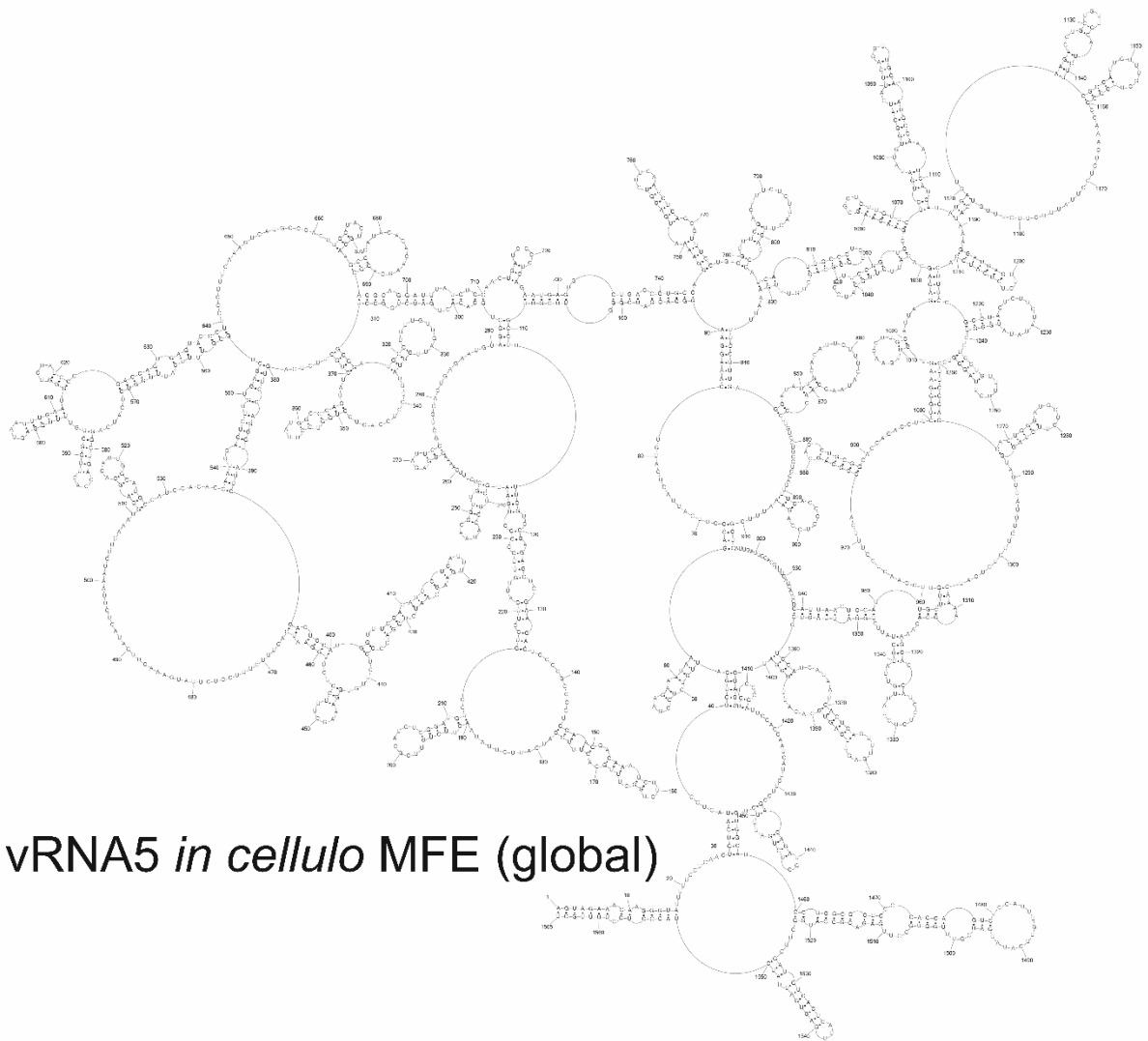
Nazwa próbki-A1:129	indeks i7	Sekwencja indeksu i7	indeks i5	Sekwencja indeksu i5	Stężenie [ng/μl] (Qubit)	Średnia długość fragmentów [bp] (Bioanlizator)	Masa molowa biblioteki [g/mol]	stężenie [nM]
<i>in vitro</i> DMS R1	N701	TAAGGCGA	S502	ATAGAGAG	0,984	485	315250	3,12
<i>in vitro</i> DMS K1	N702	CGTACTAG	S502	ATAGAGAG	2,26	600	390000	5,79
<i>in vitro</i> DMS R2	N705	GGACTCCT	S502	ATAGAGAG	0,746	479	311350	2,40
<i>in vitro</i> DMS K2	N706	TAGGCATG	S502	ATAGAGAG	1,8	602	391300	4,60
<i>in vitro</i> DMS R3	N703	AGGCAGAA	S503	AGAGGATA	1,26	504	327600	3,85
<i>in vitro</i> DMS K3	N704	TCCTGAGC	S503	AGAGGATA	2,48	598	388700	6,38
<i>in vitro</i> NAI R1	N703	AGGCAGAA	S502	ATAGAGAG	0,738	574	373100	1,98
<i>in vitro</i> NAI K1	N704	TCCTGAGC	S502	ATAGAGAG	1,63	612	397800	4,10
<i>in vitro</i> NAI R2	N701	TAAGGCGA	S503	AGAGGATA	1,25	553	359450	3,48
<i>in vitro</i> NAI K2	N702	CGTACTAG	S503	AGAGGATA	2,28	662	430300	5,30
<i>in vitro</i> NAI R3	N705	GGACTCCT	S503	AGAGGATA	1,21	570	370500	3,27
<i>in vitro</i> NAI K3	N706	TAGGCATG	S503	AGAGGATA	2,46	561	364650	6,75
<i>in cellulo</i> NAI R1	N701	TAAGGCGA	S504	TCTACTCT	4,54	414	269100	16,87
<i>in cellulo</i> NAI R2	N702	CGTACTAG	S504	TCTACTCT	4,72	412	267800	17,63
<i>in cellulo</i> NAI R3	N703	AGGCAGAA	S504	TCTACTCT	5,26	422	274300	19,18
<i>in cellulo</i> NAI K1	N704	TCCTGAGC	S504	TCTACTCT	4,74	436	283400	16,73
<i>in cellulo</i> NAI K2	N705	GGACTCCT	S504	TCTACTCT	7	434	282100	24,81
<i>in cellulo</i> NAI K3	N706	TAGGCATG	S504	TCTACTCT	4,68	430	279500	16,74
<i>in cellulo</i> DMS R1	N701	TAAGGCGA	S517	TCTTACGC	4,72	401	260650	18,11
<i>in cellulo</i> DMS R2	N702	CGTACTAG	S517	TCTTACGC	5,48	470	305500	17,94
<i>in cellulo</i> DMS R3	N703	AGGCAGAA	S517	TCTTACGC	8,08	434	282100	28,64
<i>in cellulo</i> DMS K1	N704	TCCTGAGC	S517	TCTTACGC	5,18	416	270400	19,16
<i>in cellulo</i> DMS K2	N705	GGACTCCT	S517	TCTTACGC	4,92	449	291850	16,86
<i>in cellulo</i> DMS K3	N706	TAGGCATG	S517	TCTTACGC	6,82	426	276900	24,63
total NAI R	N708	CAGAGAGG	S502	ATAGAGAG	3,02	609	395850	7,63
total NAI K	N708	CAGAGAGG	S503	AGAGGATA	4,46	635	412750	10,81
total DMS R	N709	GCTACGCT	S502	ATAGAGAG	2,82	637	414050	6,81
total DMS K	N709	GCTACGCT	S503	AGAGGATA	3,84	629	408850	9,39

Tabela 11 Współczynnik korelacji Pearson’a dla reaktywności poszczególnych nukleotydów vRNA z mapowania SHAPE oraz DMS dla wszystkich trzech powtórzeń biologicznych. Współczynnik korelacji został obliczony pomiędzy wszystkimi trzema powtórzeniami biologicznymi (rep1, rep2 oraz rep3). Dla prób z warunków *in cellulo* dla segmentów 1,2,3,4 i 6 obliczono współczynnik korelacji dla wyników uzyskanych dla fragmentów vRNA użytych w predykcji struktury drugorzędowej.

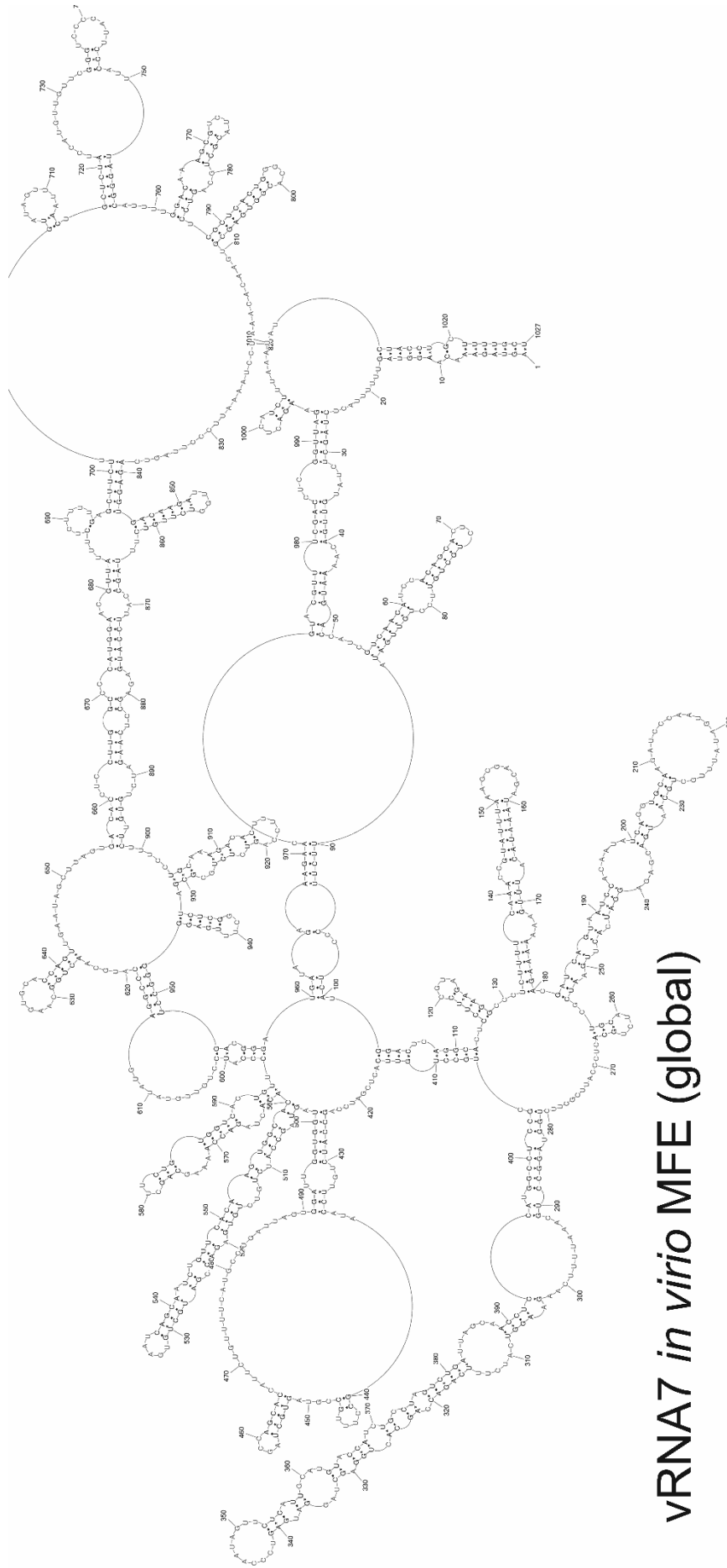
	DMS			NAI		
<i>in virio</i>	rep1:rep 2	rep 2:rep 3	rep 1:rep 3	rep1:rep 2	rep 2:rep 3	rep 1:rep 3
vRNA1	0,9797	0,9814	0,9810	0,9746	0,9384	0,9473
vRNA2	0,9852	0,9871	0,9893	0,9328	0,9491	0,9750
vRNA3	0,9760	0,9814	0,9794	0,9141	0,9394	0,9481
VRNA4	0,9833	0,9834	0,9871	0,9386	0,9627	0,9577
VRNA5	0,9820	0,9866	0,9848	0,9588	0,9678	0,9805
vRNA6	0,9663	0,9732	0,9740	0,8660	0,8925	0,9049
vRNA7	0,9806	0,9879	0,9858	0,8474	0,8870	0,9442
vRNA8	0,9835	0,9841	0,9884	0,8903	0,9429	0,9395
	DMS			NAI		
<i>in cellulo</i>	rep1:rep 2	rep 2:rep 3	rep 1:rep 3	rep1:rep 2	rep 2:rep 3	rep 1:rep 3
vRNA1	0,9540	0,9365	0,9445	0,9812	0,9815	0,9850
vRNA2	0,9200	0,8063	0,8028	0,9842	0,9829	0,9843
vRNA3	0,9554	0,8148	0,8150	0,9580	0,9635	0,9624
VRNA4	0,9461	0,9249	0,9031	0,9426	0,96752	0,9659
VRNA5	0,9691	0,9685	0,9688	0,9919	0,9874	0,9876
vRNA6	0,9291	0,9136	0,8805	0,9082	0,9109	0,9414
vRNA7	0,9823	0,9830	0,9891	0,9967	0,9941	0,9950
vRNA8	0,9893	0,9870	0,9911	0,9967	0,9956	0,9962



Ryc. 81 Struktura drugorzędowa vRNA5 uzyskana na podstawie danych eksperymentalnych z mapowania chemicznego RNA w warunkach *in vitro*. Struktura o minimalnej wartości energii swobodnej (MFE).

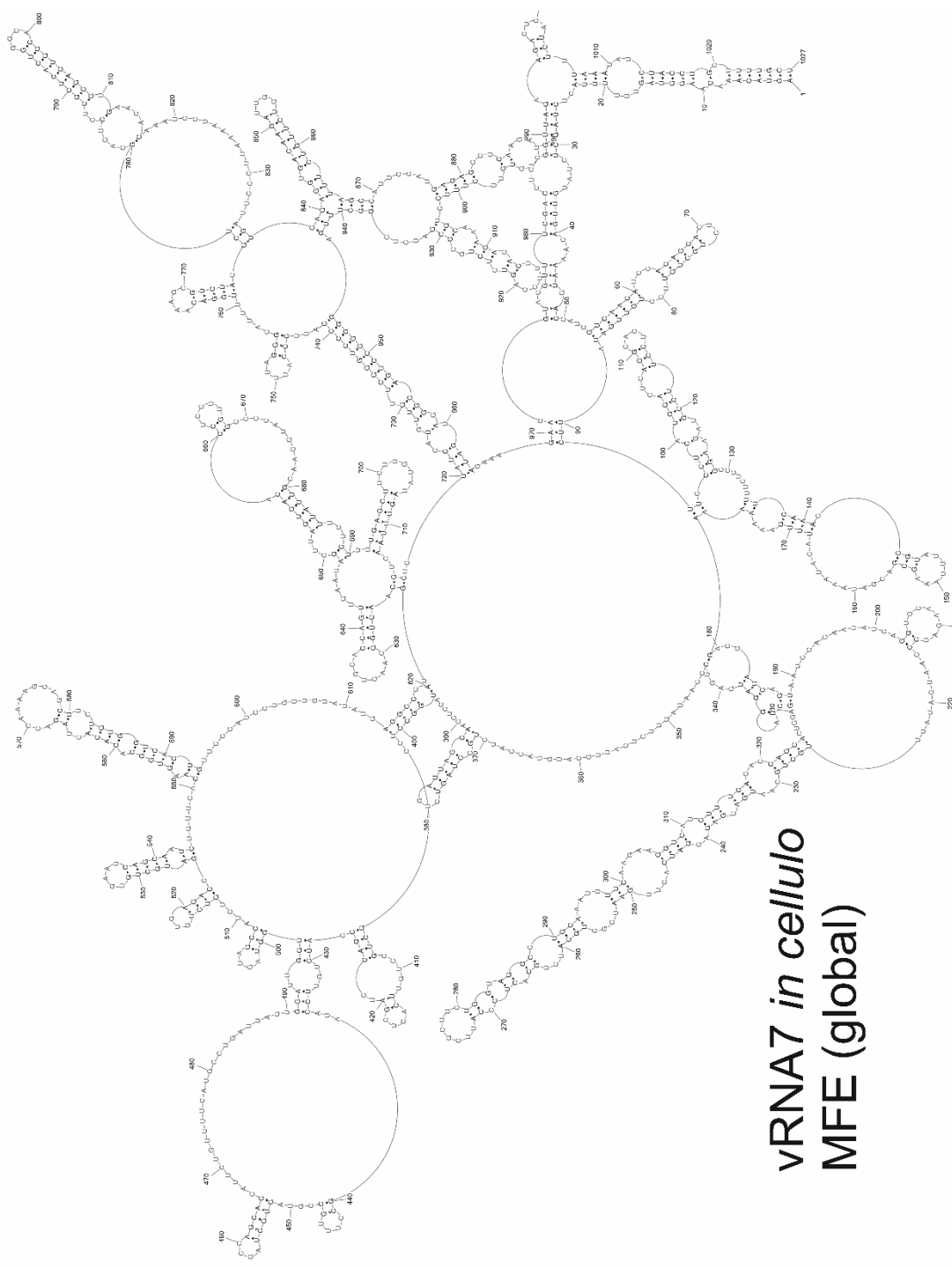


Ryc. 82 Struktura drugorzędowa vRNA5 uzyskana na podstawie danych eksperymentalnych z mapowania chemicznego RNA w warunkach *in cellulo*. Struktura o minimalnej wartości energii swobodnej (MFE).

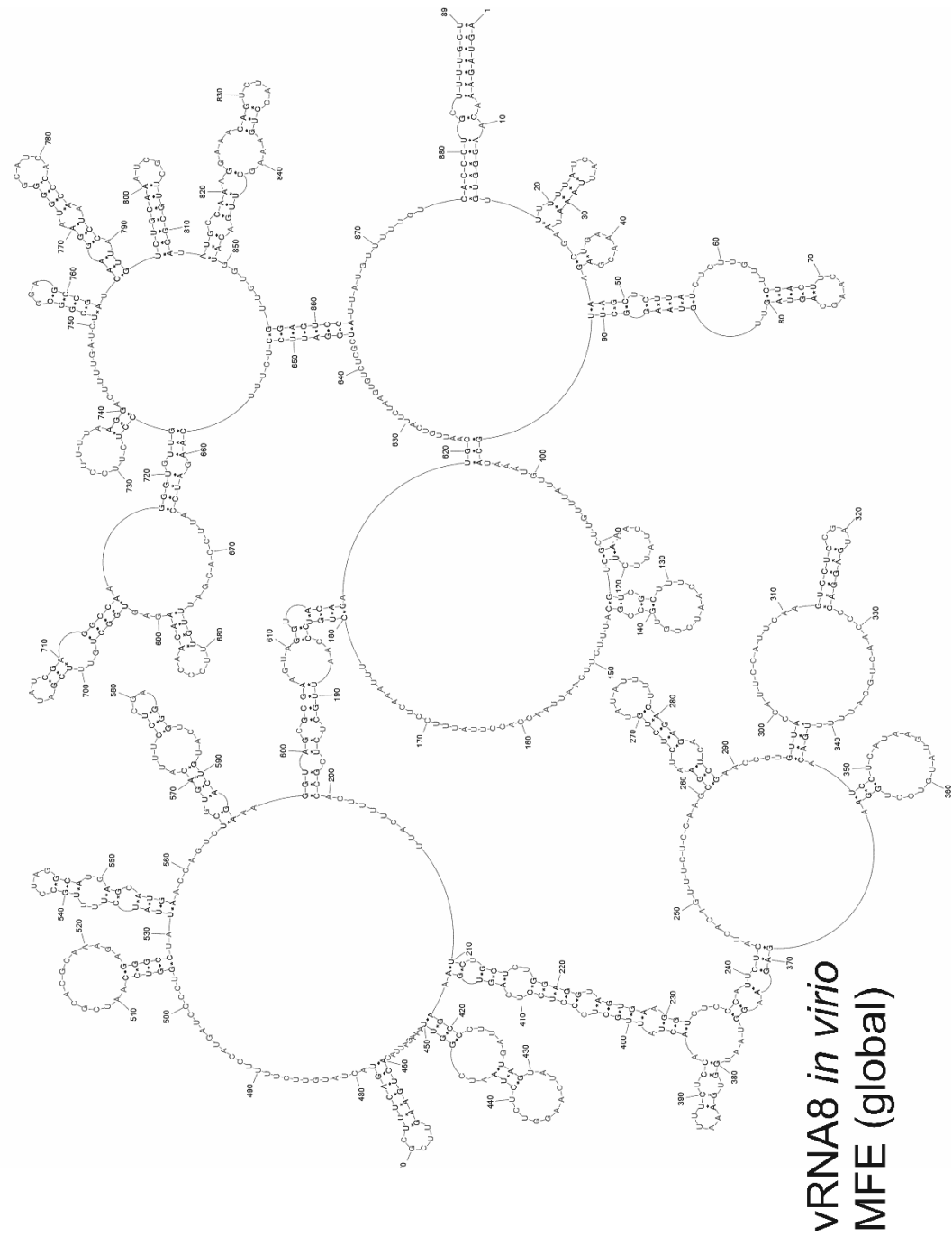


vRNA7 *in vitro* MFE (global)

Ryc. 83 Struktura drugorzędowa vRNA7 uzyskana na podstawie danych eksperymentalnych z mapowania chemicznego RNA w warunkach *in vitro*. Struktura o minimalnej wartości energii swobodnej (MFE).

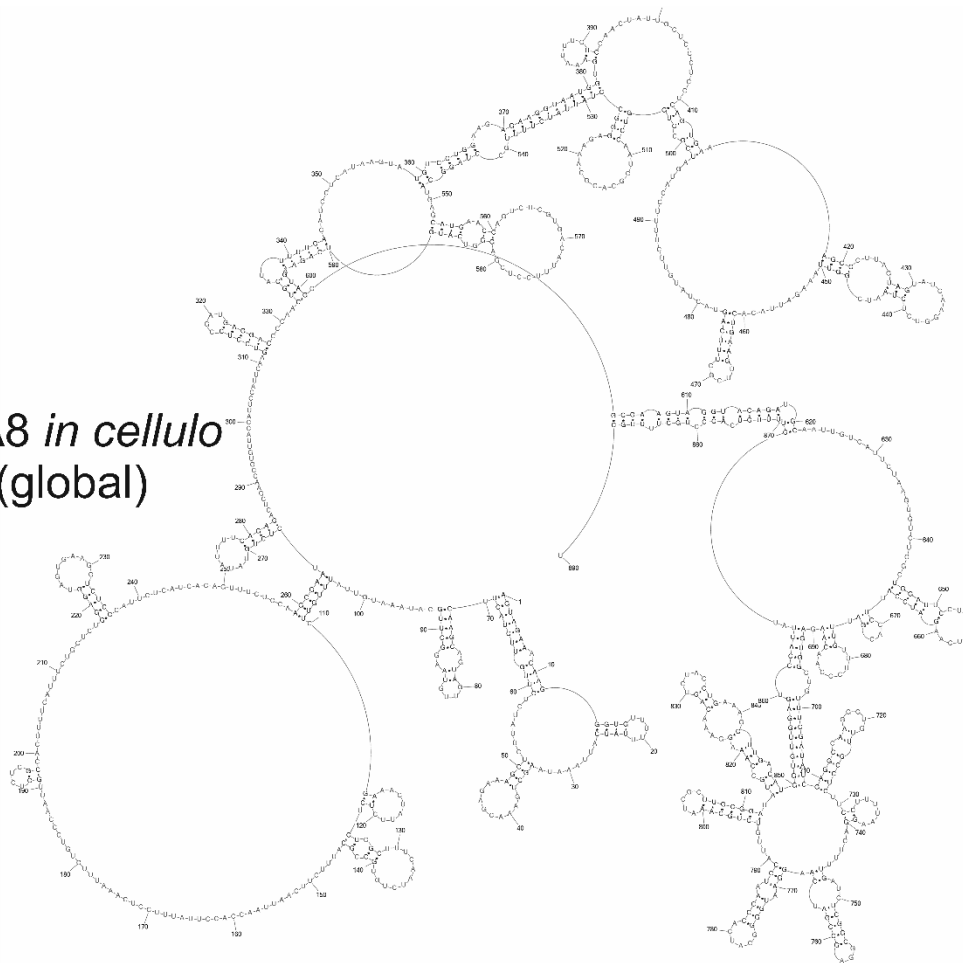


Ryc. 84 Struktura drugorzędowa vRNA7 uzyskana na podstawie danych eksperymentalnych z mapowania chemicznego RNA w warunkach *in cellulose*. Struktura o minimalnej wartości energii swobodnej (MFE).



Ryc. 85 Struktura drugorzędowa vRNA8 uzyskana na podstawie danych eksperymentalnych z mapowania chemicznego RNA w warunkach *in virio*. Struktura o minimalnej wartości energii swobodnej (MFE).

vRNA8 *in cellulo*
MFE (global)



Ryc. 86 Struktura drugorzędowa vRNA7 uzyskana na podstawie danych eksperymentalnych z mapowania chemicznego RNA w warunkach *in virio*. Struktura o minimalnej wartości energii swobodnej (MFE).

Bibliografia

1. Bouvier, N. M. & Palese, P. The biology of influenza viruses. *Vaccine* **26**, 49–53 (2008).
2. Su, S., Fu, X., Li, G., Kerlin, F. & Veit, M. Novel Influenza D virus: Epidemiology, pathology, evolution and biological characteristics. *Virulence* **8**, 1580–1591 (2017).
3. Louten, J. Influenza Viruses In:Essential Human Virology. *Essent. Hum. Virol.* 171–191 (2016) doi:10.1016/B978-0-12-800947-5.00010-7.
4. Monto, A. S. & Fukuda, K. Lessons from influenza pandemics of the last 100 years. *Clin. Infect. Dis.* **70**, 951–957 (2020).
5. Saunders-Hastings, P. R. & Krewski, D. Reviewing the history of pandemic influenza: Understanding patterns of emergence and transmission. *Pathogens* **5**, (2016).
6. Gavrilova, N. S. & Gavrilov, L. A. Patterns of mortality during pandemic: An example of Spanish flu pandemic of 1918. *Popul. Econ.* **4**, 56–64 (2020).
7. World Health Organization. *Global influenza strategy*.
8. Samji, T. Influenza A: Understanding the viral life cycle. *Yale J. Biol. Med.* **82**, 153–159 (2009).
9. Eisfeld, A. J., Neumann, G. & Kawaoka, Y. At the centre: Influenza A virus ribonucleoproteins. *Nat. Rev. Microbiol.* **13**, 28–41 (2015).
10. Lenartowicz, E. *et al.* Antisense Oligonucleotides Targeting Influenza A Segment 8 Genomic RNA Inhibit Viral Replication. *Nucleic Acid Ther.* **26**, 277–285 (2016).
11. Michalak, P. *et al.* Conserved Structural Motifs of Two Distant IAV Subtypes in. *Viruses* **13**, 1–14 (2021).
12. Michalak, P. *et al.* Secondary structure of the segment 5 genomic RNA of influenza A virus and its application for designing antisense oligonucleotides. *Sci. Rep.* **9**, 3801 (2019).

13. Ruszkowska, A., Lenartowicz, E., Moss, W. N., Kierzek, R. & Kierzek, E. Secondary structure model of the naked segment 7 influenza A virus genomic RNA. *Biochem. J.* **473**, 4327–4348 (2016).
14. Lenartowicz, E. *et al.* Self-folding of naked segment 8 genomic RNA of influenza a virus. *PLoS One* **11**, 1–21 (2016).
15. Asha, K. & Kumar, B. Emerging Influenza D Virus Threat: What We Know so Far! *J. Clin. Med.* **8**, 192 (2019).
16. Potter, C. W. A history of influenza. *J. Appl. Microbiol.* **91**, 572–579 (2001).
17. Gultyaev, A. P., Fouchier, R. A. M. & Olsthoorn, R. C. L. Influenza virus RNA structure: Unique and common features. *Int. Rev. Immunol.* **29**, 533–556 (2010).
18. Li, X., Gu, M., Zheng, Q., Gao, R. & Liu, X. Packaging signal of influenza A virus. *Viol. J.* **18**, 1–10 (2021).
19. Hayashi, T., MacDonald, L. A. & Takimoto, T. Influenza A Virus Protein PA-X Contributes to Viral Growth and Suppression of the Host Antiviral and Immune Responses. *J. Virol.* **89**, 6442–6452 (2015).
20. Gong, Y. N., Chen, G. W., Chen, C. J., Kuo, R. L. & Shih, S. R. Computational analysis and mapping of novel open reading frames in influenza A viruses. *PLoS One* **9**, 1–22 (2014).
21. Deng, T., Sharps, J. L. & Brownlee, G. G. Role of the influenza heterotrimeric RNA polymerase complex in the initiation of replication. *J. Gen. Virol.* **87**, 3373–3377 (2006).
22. Graef, K. M. *et al.* The PB2 Subunit of the Influenza Virus RNA Polymerase Affects Virulence by Interacting with the Mitochondrial Antiviral Signaling Protein and Inhibiting Expression of Beta Interferon. *J. Virol.* **84**, 8433–8445 (2010).
23. Hai, R. *et al.* PB1-F2 Expression by the 2009 Pandemic H1N1 Influenza Virus Has Minimal Impact on Virulence in Animal Models. *J. Virol.* **84**, 4442–4450 (2010).

24. Vasin, A. V. *et al.* Molecular mechanisms enhancing the proteome of influenza A viruses: An overview of recently discovered proteins. *Virus Res.* **185**, 53–63 (2014).
25. Dou, D., Revol, R., Östbye, H., Wang, H. & Daniels, R. Influenza A virus cell entry, replication, virion assembly and movement. *Front. Immunol.* **9**, 1–17 (2018).
26. Hao, W., Wang, L. & Li, S. Roles of the non-structural proteins of influenza A virus. *Pathogens* **9**, 1–19 (2020).
27. Jagger, B. W. *et al.* An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* **337**, 199–204 (2012).
28. Muramoto, Y., Noda, T., Kawakami, E., Akkina, R. & Kawaoka, Y. Identification of Novel Influenza A Virus Proteins Translated from PA mRNA. *J. Virol.* **87**, 2455–2462 (2013).
29. Chaimayo, C., Hayashi, T., Underwood, A., Hodges, E. & Takimoto, T. Selective incorporation of vRNP into influenza A virions determined by its specific interaction with M1 protein. *Virology* **505**, 23–32 (2017).
30. Luo, W. *et al.* Phospholipid scramblase 1 interacts with influenza A virus NP, impairing its nuclear import and thereby suppressing virus replication. *PLoS Pathog.* **14**, e1006851 (2018).
31. Kordyukova, L. V, Shtykova, E. V, Baratova, L. A., Svergun, D. I. & Batishchev, O. V. Matrix proteins of enveloped viruses: a case study of Influenza A virus M1 protein. *J. Biomol. Struct. Dyn.* **37**, 671–690 (2019).
32. Zhu, L. *et al.* Influenza virus matrix protein M1 interacts with SLD5 to block host cell cycle. *Cell. Microbiol.* **21**, e13038 (2019).
33. To, J. & Torres, J. Viroporins in the Influenza Virus. *Cells* **8**, 654 (2019).
34. Wise, H. M. *et al.* Identification of a Novel Splice Variant Form of the Influenza A Virus M2 Ion Channel with an Antigenically Distinct Ectodomain. *PLoS Pathog.* **8**, (2012).

35. Rückle, A. *et al.* The NS1 Protein of Influenza A Virus Blocks RIG-I-Mediated Activation of the Noncanonical NF- κ B Pathway and p52/RelB-Dependent Gene Expression in Lung Epithelial Cells. *J. Virol.* **86**, 10211–10217 (2012).
36. Marc, D. Influenza virus non-structural protein NS1: Interferon antagonism and beyond. *J. Gen. Virol.* **95**, 2594–2611 (2014).
37. Robb, N. C., Smith, M., Vreede, F. T. & Fodor, E. NS2/NEP protein regulates transcription and replication of the influenza virus RNA genome. *J. Gen. Virol.* **90**, 1398–1407 (2009).
38. Paterson, D. & Fodor, E. Emerging Roles for the Influenza A Virus Nuclear Export Protein (NEP). *PLoS Pathog.* **8**, (2012).
39. Selman, M., Dankar, S. K., Forbes, N. E., Jia, J. J. & Brown, E. G. Adaptive mutation in influenza A virus non-structural gene is linked to host switching and induces a novel protein by alternative splicing. *Emerg. Microbes Infect.* **1**, 0 (2012).
40. McAuley, J. L., Gilbertson, B. P., Trifkovic, S., Brown, L. E. & McKimm-Breschkin, J. L. Influenza virus neuraminidase structure and functions. *Front. Microbiol.* **10**, 26–32 (2019).
41. Alvarado-Facundo, E. *et al.* Influenza Virus M2 Protein Ion Channel Activity Helps To Maintain Pandemic 2009 H1N1 Virus Hemagglutinin Fusion Competence during Transport to the Cell Surface. *J. Virol.* **89**, 1975–1985 (2015).
42. Bui, M., Wills, E. G., Helenius, A. & Whittaker, G. R. Role of the Influenza Virus M1 Protein in Nuclear Export of Viral Ribonucleoproteins. *J. Virol.* **74**, 1781–1786 (2000).
43. Chang, S. *et al.* Cryo-EM Structure of Influenza Virus RNA Polymerase Complex at 4.3Å Resolution. *Mol. Cell* **57**, 925–935 (2015).
44. Moeller, A., Kirchdoerfer, R. N., Potter, C. S., Carragher, B. & Wilson, I. A. Organization of the influenza virus replication machinery. *Science* **338**, 1631–1634 (2012).

45. Arranz, R. *et al.* The structure of native influenza virion ribonucleoproteins. *Science (80-.)*. **338**, 1634–1637 (2012).
46. Lee, N. *et al.* Genome-wide analysis of influenza viral RNA and nucleoprotein association. *Nucleic Acids Res.* **45**, 8968–8977 (2017).
47. Williams, G. D. *et al.* Nucleotide resolution mapping of influenza A virus nucleoprotein-RNA interactions reveals RNA features required for replication. *Nat. Commun.* **9**, 1–12 (2018).
48. Killingley, B. & Nguyen-Van-Tam, J. Routes of influenza transmission. *Influenza Other Respi. Viruses* **7**, 42–51 (2013).
49. Kosik, I. & Yewdell, J. W. Influenza hemagglutinin and neuraminidase: Yin–yang proteins coevolving to thwart immunity. *Viruses* **11**, (2019).
50. Böttcher-Friebertshäuser, E. *et al.* Cleavage of Influenza Virus Hemagglutinin by Airway Proteases TMPRSS2 and HAT Differs in Subcellular Localization and Susceptibility to Protease Inhibitors. *J. Virol.* **84**, 5605–5614 (2010).
51. Huang, Q. *et al.* Early steps of the conformational change of influenza virus hemagglutinin to a fusion active state: Stability and energetics of the hemagglutinin. *Biochimica et Biophysica Acta - Biomembranes* vol. 1614 3–13 (Elsevier, 2003).
52. Cady Sarah D., Luo Wenbin, Hu Fanghao, H. M. Structure and Function of the Influenza A M2 Proton Channel. *Biochemistry* **48**, 7356–7364 (2009).
53. Crow, M., Deng, T., Addley, M. & Brownlee, G. G. Mutational Analysis of the Influenza Virus cRNA Promoter and Identification of Nucleotides Critical for Replication. *J. Virol.* **78**, 6263–6270 (2004).
54. De Vlugt, C., Sikora, D. & Pelchat, M. Insight into influenza: A virus cap-snatching. *Viruses* **10**, (2018).
55. Poon, L. L. M., Pritlove, D. C., Fodor, E. & Brownlee, G. G. Direct Evidence that the Poly(A) Tail of Influenza A Virus mRNA Is Synthesized by Reiterative Copying of a U Track in the Virion RNA Template. *J. Virol.* **73**, 3473–3476 (1999).

56. Mora, R., Rodriguez-Boulan, E., Palese, P. & García-Sastre, A. Apical Budding of a Recombinant Influenza A Virus Expressing a Hemagglutinin Protein with a Basolateral Localization Signal. *J. Virol.* **76**, 3544–3553 (2002).
57. Noda, T. & Kawaoka, Y. Packaging of influenza virus genome: Robustness of selection. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 8797–8798 (2012).
58. Muramoto, Y. *et al.* Hierarchy among Viral RNA (vRNA) Segments in Their Role in vRNA Incorporation into Influenza A Virions. *J. Virol.* **80**, 2318–2325 (2006).
59. Bolte, Hardin, Rosu, Miruna E., Hagelauer, Elena, Garcia-Sastre, Adolfo, Schwemmle, M. Packaging of the Influenza Virus Genome Is Governed by a Plastic Network of RNA- and Nucleoprotein-Mediated Interactions. *J. Virol.* **93**, 1–11 (2019).
60. Nakatsu, S. *et al.* Complete and Incomplete Genome Packaging of Influenza A and B Viruses. *MBio* **7**, e01248-16 (2016).
61. Fournier, E. *et al.* A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Res.* **40**, 2197–2209 (2012).
62. Ziegler, C. M. & Botten, J. W. Defective Interfering Particles of Negative-Strand RNA Viruses. *Trends Microbiol.* **28**, 554–565 (2020).
63. Brooke, C. B. Population Diversity and Collective Interactions during Influenza Virus Infection. *J. Virol.* **91**, 1–13 (2017).
64. Enami, M., Sharma, G., Benham, C. & Palese, P. An influenza virus containing nine different RNA segments. *Virology* **185**, 291–298 (1991).
65. Noda, T. *et al.* Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* **439**, 490–492 (2006).
66. Nakatsu, S. *et al.* Influenza C and D Viruses Package Eight Organized Ribonuclein Complexes. *J. Virol.* **92**, 1–9 (2018).
67. Su, W.-C., Yu, W.-Y., Huang, S.-H. & Lai, M. M. C. Ubiquitination of the Cytoplasmic Domain of Influenza A Virus M2 Protein is Crucial for Production

- of Infectious Virus Particles. *J. Virol.* **92**, JVI.01972-17 (2017).
68. Marsh, G. A. *et al.* Specific Residues of the Influenza A Virus Hemagglutinin Viral RNA Are Important for Efficient Packaging into Budding Virions □. **81**, 9727–9736 (2007).
 69. Ozawa, M. *et al.* Nucleotide Sequence Requirements at the 5' End of the Influenza A Virus M RNA Segment for Efficient Virus Replication □. **83**, 3384–3388 (2009).
 70. Zhao, L. *et al.* New Insights into the Nonconserved Noncoding Region of the Subtype-Determinant Hemagglutinin and Neuraminidase Segments of Influenza A Viruses. *J. Virol.* **88**, 11493–11503 (2014).
 71. Hale, B. G., Randall, R. E., Ortin, J. & Jackson, D. The multifunctional NS1 protein of influenza A viruses. *J. Gen. Virol.* **89**, 2359–2376 (2008).
 72. Munir, M., Zohari, S., Belák, S. & Berg, M. Double-stranded RNA-induced activation of activating protein-1 promoter is differentially regulated by the non-structural protein 1 of avian influenza A viruses. *Viral Immunol.* **25**, 79–85 (2012).
 73. Sha, T. W. *et al.* Influenza A virus NS1 optimises virus infectivity by enhancing genome packaging in a dsRNA-binding dependent manner. *Virol. J.* **17**, 107 (2020).
 74. Trapp, S. *et al.* Major contribution of the RNA-binding domain of NS1 in the pathogenicity and replication potential of an avian H7N1 influenza virus in chickens. *Virol. J.* **15**, 1–12 (2018).
 75. Wang, W. & Krug, R. M. U6atac snRNA, the highly divergent counterpart of U6 snRNA, is the specific target that mediates inhibition of AT-AC splicing by the influenza virus NS1 protein. *Rna* **4**, 55–64 (1998).
 76. Garaigorta, U. & Ortín, J. Mutation analysis of a recombinant NS replicon shows that influenza virus NS1 protein blocks the splicing and nucleocytoplasmic transport of its own viral mRNA. *Nucleic Acids Res.* **35**, 4573–4582 (2007).

77. Qiu, Y. & Krug, R. M. The influenza virus NS1 protein is a poly(A)-binding protein that inhibits nuclear export of mRNAs containing poly(A). *J. Virol.* **68**, 2425–2432 (1994).
78. Krug, R. M. Functions of the influenza A virus NS1 protein in antiviral defense. *Curr. Opin. Virol.* **12**, 1–6 (2015).
79. Talon, J. *et al.* Activation of Interferon Regulatory Factor 3 Is Inhibited by the Influenza A Virus NS1 Protein. *J. Virol.* **74**, 7989–7996 (2000).
80. Wang, X. *et al.* Influenza A Virus NS1 Protein Prevents Activation of NF- κ B and Induction of Alpha/Beta Interferon. *J. Virol.* **74**, 11566–11573 (2000).
81. Dauber, B. & Wolff, T. Activation of the antiviral kinase PKR and viral countermeasures. *Viruses* **1**, 523–544 (2009).
82. Schierhorn, K. L. *et al.* Influenza A Virus Virulence Depends on Two Amino Acids in the N-Terminal Domain of Its NS1 Protein To Facilitate Inhibition of the RNA-Dependent Protein Kinase PKR. *J. Virol.* **91**, 1–18 (2017).
83. Drappier, M. & Michiels, T. Inhibition of the OAS/RNase L pathway by viruses. *Curr. Opin. Virol.* **15**, 19–26 (2015).
84. Varga, Z. T. & Palese, P. The influenza A virus protein PB1-F2: killing two birds with one stone? *Virulence* **2**, 542–546 (2011).
85. Wise, H. M. *et al.* A Complicated Message: Identification of a Novel PB1-Related Protein Translated from Influenza A Virus Segment 2 mRNA. *J. Virol.* **83**, 8021 LP – 8031 (2009).
86. Klemm, C., Boergeling, Y., Ludwig, S. & Ehrhardt, C. Immunomodulatory Nonstructural Proteins of Influenza A Viruses. *Trends Microbiol.* **26**, 624–636 (2018).
87. Zamarin, D., Ortigoza, M. B. & Palese, P. Influenza A virus PB1-F2 protein contributes to viral pathogenesis in mice. *J. Virol.* **80**, 7976–7983 (2006).
88. Varga, Z. T., Grant, A., Manicassamy, B. & Palese, P. Influenza virus protein PB1-F2 inhibits the induction of type I interferon by binding to MAVS and

- decreasing mitochondrial membrane potential. *J. Virol.* **86**, 8359–8366 (2012).
89. Park, E. *et al.* Co-degradation of interferon signaling factor DDX3 by PB1-F2 as a basis for high virulence of 1918 pandemic influenza. *EMBO J.* **38**, 1–13 (2019).
 90. Cheung, P.-H. H. *et al.* Virus subtype-specific suppression of MAVS aggregation and activation by PB1-F2 protein of influenza A (H7N9) virus. *PLoS Pathog.* **16**, e1008611 (2020).
 91. Leymarie, O. *et al.* Influenza virus protein PB1-F2 interacts with CALCOCO2 (NDP52) to modulate innate immune response. *J. Gen. Virol.* **98**, 1196–1208 (2017).
 92. Cheung, P.-H. H. *et al.* PB1-F2 protein of highly pathogenic influenza A (H7N9) virus selectively suppresses RNA-induced NLRP3 inflammasome activation through inhibition of MAVS-NLRP3 interaction. *J. Leukoc. Biol.* (2020) doi:10.1002/JLB.4AB0420-694R.
 93. Oishi, K., Yamayoshi, S. & Kawaoka, Y. Mapping of a Region of the PA-X Protein of Influenza A Virus That Is Important for Its Shutoff Activity. *J. Virol.* **89**, 8661–8665 (2015).
 94. Wang, Q. *et al.* Host Interaction Analysis of PA-N155 and PA-N182 in Chicken Cells Reveals an Essential Role of UBA52 for Replication of H5N1 Avian Influenza Virus. *Front. Microbiol.* **9**, 936 (2018).
 95. Zhang, P., Wu, W., Chen, Q. & Chen, M. Non-Coding RNAs and their Integrated Networks. *J. Integr. Bioinform.* **16**, 1–12 (2019).
 96. Dai, X., Zhang, S. & Zaleta-Rivera, K. RNA: interactions drive functionalities. *Mol. Biol. Rep.* **47**, 1413–1434 (2020).
 97. Hendrix, D. K., Brenner, S. E. & Holbrook, S. R. RNA structural motifs: Building blocks of a modular biomolecule. *Q. Rev. Biophys.* **38**, 221–243 (2005).
 98. Sanchez de Groot, N. *et al.* RNA structure drives interaction with proteins. *Nat. Commun.* **10**, 1–13 (2019).

99. Guil, S. & Esteller, M. RNA-RNA interactions in gene regulation: The coding and noncoding players. *Trends Biochem. Sci.* **40**, 248–256 (2015).
100. Schmidt, N. *et al.* A direct RNA-protein interaction atlas of the SARS-CoV-2 RNA in infected human cells. *bioRxiv* 2020.07.15.204404 (2020) doi:10.1101/2020.07.15.204404.
101. Laing, C., Jung, S., Iqbal, A. & Schlick, T. Tertiary Motifs Revealed in Analyses of Higher-Order RNA Junctions. *J. Mol. Biol.* **393**, 67–82 (2009).
102. Rietveld, K., Bosch, L. & W.A.Pleij, C. A new principle of RNA folding based on pseudoknotting. *Nucleic Acids Res.* **13**, 1717–1731 (1985).
103. Brierley, I., Pennell, S. & Gilbert, R. J. C. Viral RNA pseudoknots: Versatile motifs in gene expression and replication. *Nat. Rev. Microbiol.* **5**, 598–610 (2007).
104. Staple, D. W. & Butcher, S. E. Pseudoknots: RNA structures with diverse functions. *PLoS Biol.* **3**, 0956–0959 (2005).
105. Maximova, V., Kirilov, K., Markov, S. & Ivanov, I. Comparative genomic studies of influenza A viruses performed on BlueGene P supercomputer: Part 1. conservative nucleotide sequences in influenza a virus genomes revealed by multiple sequence alignment. *Biotechnol. Biotechnol. Equip.* **25**, 2672–2678 (2011).
106. Fodor, E., Pritlove, D. C. & Brownlee, G. G. Characterization of the RNA-fork model of virion RNA in the initiation of transcription in influenza A virus. *J. Virol.* **69**, 4012–4019 (1995).
107. Flick, R., Neumann, G., Hoffmann, E., Neumeier, E. & Hobom, G. Promoter elements in the influenza vRNA terminal structure. *Rna* **2**, 1046–1057 (1996).
108. Flick, R. & Hobom, G. Interaction of influenza virus polymerase with viral RNA in the ‘corkscrew’ conformation. *J. Gen. Virol.* **80**, 2565–2572 (1999).
109. Hsu, M., Parvin, J. D., Gupta, S., Krystal, M. & Palese, P. Terminal Panhandle. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 8140–8144 (1987).

110. Ferhadian, D. *et al.* Structural and functional motifs in influenza virus RNAs. *Front. Microbiol.* **9**, 1–11 (2018).
111. Pflug, A., Lukarska, M., Resa-Infante, P., Reich, S. & Cusack, S. Structural insights into RNA synthesis by the influenza virus transcription-replication machine. *Virus Res.* **234**, 103–117 (2017).
112. Robb, N. C., Te Velthuis, A. J. W., Fodor, E. & Kapanidis, A. N. Real-time analysis of single influenza virus replication complexes reveals large promoter-dependent differences in initiation dynamics. *Nucleic Acids Res.* **47**, 6466–6477 (2019).
113. Gulyaev, A. P. *et al.* Subtype-specific structural constraints in the evolution of influenza A virus hemagglutinin genes. *Sci. Rep.* **6**, 38892 (2016).
114. Gavazzi, C. *et al.* An in vitro network of intermolecular interactions between viral RNA segments of an avian H5N2 influenza A virus: Comparison with a human H3N2 virus. *Nucleic Acids Res.* **41**, 1241–1254 (2013).
115. Dadonaite, B. *et al.* The structure of the influenza A virus genome. *Nat. Microbiol.* **4**, 1781–1789 (2019).
116. Gulyaev, A. P. file:///C:/Users/Ela%20Lenartowicy/Desktop/rn.-11-942. pd. *et al.* RNA structural constraints in the evolution of the influenza A virus genome NP segment. *RNA Biol.* **11**, 942–952 (2014).
117. Kobayashi, Y. *et al.* Computational and molecular analysis of conserved influenza A virus RNA secondary structures involved in infectious virion production. *RNA Biol.* **13**, 883–894 (2016).
118. Le Sage, V. *et al.* Mapping of Influenza Virus RNA-RNA Interactions Reveals a Flexible Network. *Cell Rep.* **31**, 107823 (2020).
119. Kim, H., Webster, R. G. & Webby, R. J. Influenza Virus: Dealing with a Drifting and Shifting Pathogen. *Viral Immunol.* **31**, 174–183 (2018).
120. Moss, W. N., Priore, S. F. & Turner, D. H. Identification of potential conserved RNA secondary structure throughout influenza A coding regions. *Rna* **17**, 991–1011 (2011).

121. Soszynska-Jozwiak, M., Michalak, P., Moss, W. N., Kierzek, R. & Kierzek, E. A conserved secondary structural element in the coding region of the influenza A virus nucleoprotein (NP) mRNA is important for the regulation of viral proliferation. *PLoS One* **10**, 1–16 (2015).
122. Soszynska-Jozwiak, M. *et al.* Influenza virus segment 5 (+)RNA - Secondary structure and new targets for antiviral strategies. *Sci. Rep.* **7**, 1–15 (2017).
123. Priore, S. F., Moss, W. N. & Turner, D. H. Influenza A virus coding regions exhibit host-specific global ordered RNA structure. *PLoS One* **7**, 2–10 (2012).
124. Ozawa, M. *et al.* Contributions of Two Nuclear Localization Signals of Influenza A Virus Nucleoprotein to Viral Replication. *J. Virol.* **81**, 30–41 (2007).
125. Jiang, T., Nogales, A., Baker, S. F., Martinez-Sobrido, L. & Turner, D. H. Mutations designed by ensemble defect to misfold conserved RNA structures of influenza A segments 7 and 8 affect splicing and attenuate viral replication in cell culture. *PLoS One* **11**, 1–16 (2016).
126. Simon, L. M. *et al.* In vivo analysis of influenza A mRNA secondary structures identifies critical regulatory motifs. *Nucleic Acids Res.* **47**, 7003–7017 (2019).
127. Barnwal R.P., Yang F., Varani G. Applications of NMR to structure determination of RNAs large and small. *Arch. Biochem. Biophysics* **628**, 42–56 (2017).
128. Garmann, R. F. *et al.* Visualizing the global secondary structure of a viral RNA genome with cryo-electron microscopy. *Rna* **21**, 877–886 (2015).
129. Kauffmann, A. D., Campagna, R. J., Bartels, C. B. & Childs-Disney, J. L. Improvement of RNA secondary structure prediction using RNase H cleavage and randomized oligonucleotides. *Nucleic Acids Res.* **37**, 1–11 (2009).
130. Weeks, K. M. Advances in RNA structure analysis by chemical probing. *Curr. Opin. Struct. Biol.* **20**, 295–304 (2010).
131. Smola, M. J., Rice, G. M., Busan, S., Siegfried, N. A. & Weeks, K. M. Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling

- (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nat. Protoc.* **10**, 1643–1669 (2015).
132. Deigan, K. E., Li, T. W., Mathews, D. H. & Weeks, K. M. Accurate SHAPE-directed RNA structure determination a SHAPE experiment can be interpreted as a pseudo-free energy high accuracy . Free energy minimization , by using SHAPE pseudo-. *Sci. York* **2008**, 97–102 (2008).
 133. Tijerina, P., Mohr, S. & Russell, R. DMS footprinting of structured RNAs and RNA-protein complexes. *Nat. Protoc.* **2**, 2608–2623 (2007).
 134. Vasa, S. M., Guex, N., Wilkinson, K. A., Weeks, K. M. & Giddings, M. C. ShapeFinder: A software system for high-throughput quantitative analysis of nucleic acid reactivity information resolved by capillary electrophoresis. *Rna* **14**, 1979–1990 (2008).
 135. Karabiber, F., McGinnis, J. L., Favorov, O. V & Weeks, K. M. QuShape : Rapid , accurate , and best-practices quantification of nucleic acid probing information , resolved by capillary electrophoresis. 63–73 (2013) doi:10.1261/rna.036327.112.groups.
 136. Behjati, S. & Tarpey, P. S. What is next generation sequencing? *Arch. Dis. Child. Educ. Pract. Ed.* **98**, 236–238 (2013).
 137. Bevilacqua, P. C., Ritchey, L. E., Su, Z. & Assmann, S. M. Genome-Wide Analysis of RNA Secondary Structure. *Annu. Rev. Genet.* **50**, 235–266 (2016).
 138. Ritchey, L. E. *et al.* Structure-seq2: sensitive and accurate genome-wide profiling of RNA structure in vivo. *Nucleic Acids Res.* **45**, e135 (2017).
 139. Ding, Y. *et al.* In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **505**, 696–700 (2014).
 140. Zubradt, M. *et al.* DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods* (2016) doi:10.1038/nmeth.4057.
 141. Kwok, C. K., Ding, Y., Tang, Y., Assmann, S. M. & Bevilacqua, P. C. Determination of in vivo RNA structure in low-abundance transcripts. *Nat. Commun.* **4**, 1–12 (2013).

142. Lin, Y., May, G. E. & McManus, C. J. *Mod-seq: A high-throughput method for probing RNA secondary structure. Methods in Enzymology* vol. 558 (Elsevier Inc., 2015).
143. Smola, M. J. & Weeks, K. M. In-cell RNA structure probing with SHAPE-MaP. *Nat. Protoc.* (2018) doi:10.1038/nprot.2018.010.
144. Ashley Aw, J. G. *et al.* Direct RNA sequencing reveals structural differences between transcript isoforms. (2020) doi:10.1101/2020.06.11.147223.
145. Stephenson, W. *et al.* Direct detection of RNA modifications and structure using single molecule nanopore sequencing. *bioRxiv* 1–31 (2020) doi:10.1101/2020.05.31.126763.
146. Ghut, J. *et al.* Determination of isoform-specific RNA structure with nanopore long reads. *Nature Biotechnology* vol. 39 (2020).
147. Kwok, C. K., Ding, Y., Sherlock, M. E., Assmann, S. M. & Bevilacqua, P. C. A hybridization-based approach for quantitative and low-bias single-stranded DNA ligation. *Anal. Biochem.* **435**, 181–186 (2013).
148. Fang, R., Moss, W. N., Rutenberg-Schoenberg, M. & Simon, M. D. Probing Xist RNA Structure in Cells Using Targeted Structure-Seq. *PLoS Genet.* **11**, 1–29 (2015).
149. Siegfried, N. A., Busan, S., Rice, G. M., Nelson, J. A. E. & Weeks, K. M. RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **11**, 959–965 (2014).
150. Busan, S., Weidmann, C. A., Sengupta, A. & Weeks, K. M. Guidelines for SHAPE Reagent Choice and Detection Strategy for RNA Structure Probing Studies. *Biochemistry* **58**, 2655–2664 (2019).
151. Spitale, R. C. *et al.* RNA SHAPE analysis in living cells. *Nat. Chem. Biol.* **9**, 18–20 (2013).
152. Lee, B. *et al.* Comparison of SHAPE reagents for mapping RNA structures inside living cells. *Rna* **23**, 169–174 (2017).

153. Andrzejewska, A., Zawadzka, M., Gumna, J., Garfinkel, D. J. & Pachulska-Wieczorek, K. In vivo structure of the Ty1 retrotransposon RNA genome. *Nucleic Acids Res.* **49**, 2878–2893 (2021).
154. Duwe, S. Influenza viruses - antiviral therapy and resistance. *GMS Infect. Dis.* **5**, (2017).
155. Houser, K. & Subbarao, K. Influenza vaccines: Challenges and solutions. *Cell Host and Microbe* (2015) doi:10.1016/j.chom.2015.02.012.
156. Plotkin, S., Robinson, J. M., Cunningham, G., Iqbal, R. & Larsen, S. The complexity and cost of vaccine manufacturing – An overview. *Vaccine* **35**, 4064–4071 (2017).
157. States, U. CDC Estimates of 2009 H1N1 Influenza Cases, Hospitalizations and Deaths in the United States, April – December 12, 2009. *Emerg. Infect. Dis.* **2009**, 1–7 (2010).
158. Wieczorek, K., Szutkowska, B. & Kierzek, E. Anti-influenza strategies based on nanoparticle applications. *Pathogens* **9**, 1–24 (2020).
159. Jefferson, T., Demicheli, V., Di Pietrantonj, C. & Rivetti, D. Amantadine and rimantadine for influenza A in adults. *Cochrane Database Syst. Rev.* (2006) doi:10.1002/14651858.CD001169.pub3.
160. Belshe, R.B., Burk, B., Newman, F., Cerrutic, R.L., Sim, I. S. Resistance of influenza A virus to amantadine and rimantadine: results of the one decade of surveillance. *Juornal Infect. Dis.* **159**, 430–5 (1989).
161. Saito, R. *et al.* Frequency of amantadine-resistant influenza A viruses during two seasons featuring cocirculation of H1N1 and H3N2. *J. Clin. Microbiol.* **41**, 2164–2165 (2003).
162. Stiver, G. The treatment of influenza with antiviral drugs. *Cmaj* **168**, 49–57 (2003).
163. Hata, A., Akashi-Ueda, R., Takamatsu, K. & Matsumura, T. Safety and efficacy of peramivir for influenza treatment. *Drug Des. Devel. Ther.* **8**, 2017–2038 (2014).

164. Bassetti, M., Castaldo, N. & Carnelutti, A. Neuraminidase inhibitors as a strategy for influenza treatment: pros, cons and future perspectives. *Expert Opin. Pharmacother.* **20**, 1711–1718 (2019).
165. Cheng, P. K. C. *et al.* Oseltamivir-and amantadine-resistant influenza viruses A (H1N1). *Emerg. Infect. Dis.* **15**, 966–968 (2009).
166. Zaraket, H. *et al.* Genetic makeup of amantadine-resistant and oseltamivir-resistant human influenza A/H1N1 viruses. *J. Clin. Microbiol.* **48**, 1085–1092 (2010).
167. Mishin, V. P., Hayden, F. G. & Gubareva, L. V. Susceptibilities of antiviral-resistant influenza viruses to novel neuraminidase inhibitors. *Antimicrob. Agents Chemother.* **49**, 4515–4520 (2005).
168. Fang, Q. & Wang, D. Advanced researches on the inhibition of influenza virus by Favipiravir and Baloxavir. *Biosaf. Heal.* **2**, 64–70 (2020).
169. O’Hanlon, R. & Shaw, M. L. Baloxavir marboxil: the new influenza drug on the market. *Curr. Opin. Virol.* **35**, 14–18 (2019).
170. Heo, Y. A. Baloxavir: First Global Approval. *Drugs* **78**, 693–697 (2018).
171. Koszalka, P., Tilmanis, D., Roe, M., Vijaykrishna, D. & Hurt, A. C. Baloxavir marboxil susceptibility of influenza viruses from the Asia-Pacific, 2012–2018. *Antiviral Res.* **164**, 91–96 (2019).
172. Xu, W., Jiang, X., Huang, L., Sciences, B. & Sar, H. K. RNA interference technology. *Compr. Biotechnol.* **5**, 560–575 (2019).
173. Szabat, M. *et al.* Rna secondary structure as a first step for rational design of the oligonucleotides towards inhibition of influenza a virus replication. *Pathogens* **9**, 1–19 (2020).
174. Barik, S. siRNA for influenza therapy. *Viruses* **2**, 1448–1457 (2010).
175. Piasecka, J. *et al.* RNA Secondary Structure Motifs of the Influenza A Virus as Targets for siRNA-Mediated RNA Interference. *Mol. Ther. - Nucleic Acids* **19**, (2020).

176. McMillen, C. M. *et al.* Inhibition of influenza A virus matrix and nonstructural gene expression using RNA interference. *Virology* **497**, 171–184 (2016).
177. Zhang, D. & Lu, J. In Silico design of siRNAs targeting existing and future respiratory viruses with VirusSi. *bioRxiv* (2020) doi:10.1101/2020.08.13.250076.
178. Kierzek, E. *et al.* The influence of locked nucleic acid residues on the thermodynamic properties of 2'-O-methyl RNA/RNA heteroduplexes. *Nucleic Acids Res.* **33**, 5082–5093 (2005).
179. Yoshida, T. *et al.* Evaluation of off-target effects of gapmer antisense oligonucleotides using human cells. *Genes to Cells* **24**, 827–835 (2019).
180. Gasiunas, G. *et al.* A catalogue of biochemically diverse CRISPR-Cas9 orthologs. *Nat. Commun.* **11**, (2020).
181. Wang, F. *et al.* Advances in CRISPR-Cas systems for RNA targeting, tracking and editing. *Biotechnol. Adv.* **37**, 708–729 (2019).
182. Abudayyeh, O. O. *et al.* C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* (80-.). (2016) doi:10.1101/054742.
183. Abudayyeh, O. O. *et al.* RNA targeting with CRISPR-Cas13. *Nature* (2017) doi:10.1038/nature24049.
184. Freije, C. A. *et al.* Programmable Inhibition and Detection of RNA Viruses Using Cas13. *Mol. Cell* **76**, 826-837.e11 (2019).
185. Shmakov, S. *et al.* Discovery and functional characterization of diverse Class 2 CRISPR-Cas systems HHS Public Access Graphical abstract. *Mol Cell* **60**, 385–397 (2015).
186. Abbott, T. R. *et al.* Development of CRISPR as an Antiviral Strategy to Combat SARS-CoV-2 and Influenza. *Cell* **181**, 865-876.e12 (2020).
187. Blanchard, E. L. *et al.* Treatment of influenza and SARS-CoV-2 infections via mRNA-encoded Cas13a in rodents. *Nat. Biotechnol.* (2021)

doi:10.1038/s41587-021-00822-w.

188. Leamy, K. A., Assmann, S. M., Mathews, D. H. & Bevilacqua, P. C. Bridging the gap between in vitro and in vivo RNA folding. *Q. Rev. Biophys.* **49**, e10 (2016).
189. Shatzkes, K., Teferedegne, B. & Murata, H. A simple, inexpensive method for preparing cell lysates suitable for downstream reverse transcription quantitative PCR. *Sci. Rep.* **4**, 1–7 (2014).
190. Low, J. T. & Weeks, K. M. SHAPE-directed RNA secondary structure prediction. *Methods* **52**, 150–158 (2010).
191. Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940 (1999).
192. Hajiaghayi, M., Condon, A. & Hoos, H. H. Analysis of energy-based algorithms for RNA secondary structure prediction. *BMC Bioinformatics* **13**, (2012).
193. Riedel, C. *et al.* Cryo EM structure of the rabies virus ribonucleoprotein complex. *Sci. Rep.* **9**, 9639 (2019).
194. Harris, M. E. & Christian, E. L. RNA Crosslinking Methods. *Methods Enzymol.* **468**, 127–146 (2009).
195. McHugh, C. A., Russell, P. & Guttman, M. Methods for comprehensive experimental identification of RNA-protein interactions. *Genome Biol.* **15**, 1–10 (2014).
196. Selzer, L., Su, Z., Pintilie, G. D., Chiu, W. & Kirkegaard, K. Full-length three-dimensional structure of the influenza A virus M1 protein and its organization into a matrix layer. *PLoS Biol.* **18**, 1–26 (2020).
197. Le Sage, V. *et al.* Non-uniform and non-random binding of nucleoprotein to influenza A and B viral RNA. *Viruses* **10**, 1–11 (2018).
198. Labaronne, A. *et al.* Binding of RNA by the nucleoproteins of influenza viruses A and B. *Viruses* **8**, 1–14 (2016).

199. Huston, N. C. *et al.* Comprehensive in vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. *Mol. Cell* **81**, 584-598.e5 (2021).
200. Hutchinson, E. C., von Kirchbach, J. C., Gog, J. R. & Digard, P. Genome packaging in influenza A virus. *J. Gen. Virol.* **91**, 313–328 (2010).
201. Hoffmann, E., Stech, J., Guan, Y., Webster, R. G. & Perez, D. R. Universal primer set for the full-length amplification of all influenza A viruses. *Arch. Virol.* **146**, 2275–2289 (2001).
202. Zhou, B. *et al.* Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses. *J. Virol.* **83**, 10309–10313 (2009).
203. Illumina. Nextera XT Library Prep: Tips and Troubleshooting. *Guide* (2015).
204. Phan, T. *et al.* Segment-specific kinetics of mRNA, cRNA and vRNA accumulation during influenza infection. *J. Virol.* (2021) doi:10.1128/JVI.02102-20.
205. Busan, S. & Weeks, K. M. Accurate detection of chemical modifications in RNA by mutational profiling (MaP) with ShapeMapper 2. 143–148 (2018) doi:10.1261/rna.061945.117.
206. Petrov, A. S. *et al.* Secondary structures of rRNAs from all three domains of life. *PLoS One* **9**, 1–6 (2014).
207. Anger, A. M. *et al.* Structures of the human and Drosophila 80S ribosome. *Nature* **497**, 80–85 (2013).
208. Turner, R. & Shefer, K. Safer one-pot synthesis of the ‘ SHAPE ’ reagent 1-methyl-7-nitroisatoic anhydride (1m7). 1857–1863 (2013) doi:10.1261/rna.042374.113.4.
209. Sousa, R. T7 RNA Polymerase. *Encycl. Biol. Chem.* **4**, 355–359 (2013).
210. Han, K., Lee, Y. & Kim, W. PseudoViewer: automatic visualization of RNA pseudoknots. *Bioinformatics* **18**, S321–S328 (2002).

211. Darty, K., Alain, D. & Yann, P. VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* **25**, 1974–1975 (2009).
212. Byun, Y. & Han, K. PseudoViewer: Web application and web service for visualizing RNA pseudoknots and secondary structures. *Nucleic Acids Res.* **34**, 416–422 (2006).
213. Kawakami, E. *et al.* Strand-specific real-time RT-PCR for distinguishing influenza vRNA, cRNA, and mRNA. *J. Virol. Methods* **173**, 1–6 (2011).
214. Vester, D. *et al.* Real-time RT-qPCR assay for the analysis of human influenza A virus transcription and replication dynamics. *J. Virol. Methods* **168**, 63–71 (2010).
215. Illumina, I. Sequencing the rapidly evolving influenza A virus on the MiSeq system. 1–4 (2015).