INSTITUTE OF BIOORGANIC CHEMISTRY
POLISH ACADEMY OF SCIENCES IN POZNAN, POLAND

Maria Nalewaj

G-quadruplex motifs in the influenza A virus RNA genome: structural insights, ligand interaction, and potential function

The doctoral thesis has been completed
in the Department of RNA Structural Genomics

Supervisor:
Prof. dr hab. Elżbieta Kierzek
Co-supervisor:
Dr Marta Szabat

POZNAŃ, 2026

I am sincerely grateful to my supervisors, Prof. dr hab. Elżbieta Kierzek and Dr. Marta Szabat, for their support throughout my PhD journey. Their knowledge, guidance, and experience have helped me become the scientist I am today.

In addition, I would like to thank my colleagues from my Department and from the Institute of Bioorganic Chemistry for the years of collaboration and shared experiences.

Finally, I would like to express my gratitude to my closest ones – my partner, my parents, my grandma, my sisters, and friends. I am thankful for their constant encouragement, patience, and belief in me.

TABLE OF CONTENTS

LIST OF PUBLICATIONS INCLUDED IN THE DOCTORAL DISSERTATION

1. Identification and structural aspects of G-quadruplex-forming sequences from the influenza A virus genome

International Journal of Molecular Sciences, 2021, 22, 6031

Tomaszewska, M.*, Szabat, M.*, Zielińska, K., Kierzek, R.

*co-authors who equally contributed to this work

IF: 4.9 (2024)

MNiSW Points: 140

Number of citations: 24

2. Examples of structural motifs in viral genomes and approaches for RNA structure characterization

International Journal of Molecular Sciences, 2022, 23, 15917

Nalewaj, M.*, Szabat, M.*

*co-authors who equally contributed to this work

IF: 4.9 (2024)

MNiSW Points: 140

Number of citations: 9

3. The interaction of RNA G-quadruplexes from the influenza A virus vRNA with TMPyP4 and BRACO-19 ligands

PLoS One, 2025, 20(11): e0335975

Nalewaj, M., Sliwiak, J., Zielinska, K., Zmora , P., Niedziałek, D., Wieczorek, G., Kierzek, E., Szabat, M.

IF: 2.6 (2024)

MNiSW Points: 100

Number of citations: 0

- *Impact factor based on the journal website on February 16, 2026.*
- *Points of the Ministry of Science and Higher Education based on the list of point journals of the Ministry of Science and Higher Education from January 5, 2024.*
- *Number of citations based on Google Scholar on February 16, 2026.*

ABSTRACT

Influenza A virus (IAV), a cause of seasonal epidemics and occasional pandemics, is a significant research subject. Despite the high variability of its genome, viral RNA (vRNA) secondary structures remain conserved across IAV strains. Moreover, the role of vRNA structural motifs in the viral replication cycle has been previously described, making them interesting targets in anti-influenza strategies.

Among the particularly interesting nucleic acid secondary structures are G-quadruplexes (G4s). The G4 motifs are noncanonical DNA or RNA structures formed within guanine-rich regions. As a result, these motifs have significant biological functions, taking part in many cellular processes. In recent years, G-quadruplexes have been identified in viral genomes, where they play an important role in the viral replication cycle. Therefore, G-quadruplexes represent a promising target in antiviral therapies.

The first stage of my studies aimed to determine whether potential quadruplex-forming sequences (PQSs) are present within the IAV genome (A/California/04/2009(H1N1) strain). To this end, bioinformatic analyses were conducted, enabling the identification of 12 highly conserved PQSs within the IAV vRNA. In the next step of the research, we confirmed the RNA G-quadruplex formation using several spectroscopic techniques and polyacrylamide gel electrophoresis assay. Based on the obtained results, we revealed that three PQSs form stable RNA G-quadruplexes *in vitro*. Interestingly, two of these motifs were found within sequences encoding polymerase subunits and one within the hemagglutinin-coding region, suggesting their role in the viral replication cycle.

In the second part of the research, we investigated whether G4-specific ligands interact with RNA G-quadruplexes from the IAV genome. The aim of our study was to observe differences in binding modes of G4-specific ligands (*in vitro*). For this purpose, two small-molecule compounds, TMPyP4 and BRACO-19, were selected. We studied the G4/ligand interactions using reverse transcription stop assay, isothermal titration calorimetry, circular dichroism, fluorescence spectroscopy, and UV melting analysis. Importantly, differences were observed in the interaction properties between specific RNA G-quadruplexes and their ligands, with distinct binding modes identified for TMPyP4 and BRACO-19. Furthermore, we conducted biological studies in cell culture to evaluate the influence of TMPyP4 and its analogue, TMPyP2, on IAV minireplicon activity. The experiment results revealed a significant inhibition of viral replication upon the ligand addition, suggesting that selected G4s may be potential anti-influenza drug targets. To complement our studies, we employed molecular

modeling to simulate the folding of one selected RNA G-quadruplex. This method allowed us to gain additional insights into G4 topology and structure stability.

Overall, we concluded that RNA G-quadruplex structures are present within the A/California/04/2009(H1N1) genome and are highly conserved across strains. Notably, these motifs can be targeted by G4-specific small-molecule ligands. What is more, viral replication is effectively inhibited via G4 interaction with TMPyP4 and TMPyP2 compounds. All our findings suggest that selected PQS motifs from the IAV genome can serve as potential novel antiviral therapeutic targets.

STRESZCZENIE

Wirus grypy typu A, wywołujący sezonowe epidemie oraz sporadyczne pandemie, stanowi istotny obiekt badań. Pomimo dużej zmienności genomu, struktury drugorzędowe wirusowego RNA (vRNA) posiadają wysoką konserwatywność strukturalną wśród różnych szczepów wirusa. Ponadto zidentyfikowane motywy strukturalne vRNA pełnią ważne funkcje w cyklu replikacyjnym wirusa, co czyni je atrakcyjnymi celami w opracowywaniu strategii terapeutycznych skierowanych przeciwko wirusowi grypy.

Do szczególnie interesujących struktur drugorzędowych kwasów nukleinowych należą G-kwadrupleksy (G4). Motywy G4 to niekanoniczne struktury DNA lub RNA tworzące się w obrębie regionów bogatych w reszty guanozyny. W rezultacie cząsteczki te pełnią istotne funkcje biologiczne, biorąc udział w wielu procesach komórkowych. W ostatnich latach wykazano, że tworzą się one również w obrębie wirusowego RNA, gdzie odgrywają kluczową rolę w przebiegu cyklu replikacyjnego. W związku z tym G-kwadrupleksy stanowią obiecujący cel w terapiach przeciwwirusowych.

Celem pierwszego etapu moich badań było określenie, czy sekwencje potencjalnie tworzące G-kwadrupleksy (PQSs) występują w obrębie genomu wirusa grypy typu A (szczep A/California/04/2009(H1N1)). W tym celu przeprowadziliśmy analizę bioinformatyczną, która umożliwiła identyfikację dwunastu takich sekwencji w vRNA wirusa grypy typu A. Ponadto sekwencje te wykazywały silną konserwatywność w różnych szczepach. W celu potwierdzenia, czy zidentyfikowane motywy PQS tworzą G-kwadrupleksy RNA, użyliśmy technik spektroskopowych oraz elektroforezy w żelu poliakrylamidowym. Na podstawie otrzymanych wyników wykazaliśmy, że trzy motywy PQS tworzą stabilne struktury RNA G4 *in vitro*. Dwa z tych motywów występują wewnątrz sekwencji kodujących podjednostki polimerazy, a jeden wewnątrz regionu kodującego hemaglutyninę, co sugeruje ich rolę w cyklu replikacyjnym wirusa.

W drugim etapie badań sprawdziliśmy, czy ligandy specyficzne względem cząsteczek G4 oddziałują z G-kwadrupleksami RNA z genomu wirusa grypy typu A. Nasze badania koncentrowały się na zaobserwowaniu różnic w sposobach wiązania ligandów (*in vitro*). W tym celu wybraliśmy dwa związki małocząsteczkowe: TMPyP4 oraz BRACO-19. Powinowactwo ligandów do G-kwadrupleksów RNA określiliśmy, wykorzystując test zatrzymywania odwrotnej transkrypcji, kalorymetrię miareczkowania izotermicznego, dichroizm kołowy, spektroskopię fluorescencyjną, metodę topnienia UV. Co istotne, zaobserwowano różnice we właściwościach oddziaływań pomiędzy poszczególnymi G-

kwadrupleksami RNA a ligandami oraz zidentyfikowano odmienne mechanizmy wiązania dla TMPyP4 oraz BRACO-19. Ponadto przeprowadziliśmy badania biologiczne w hodowlach komórkowych w celu oceny wpływu ligandów TMPyP4 oraz jego analogu, TMPyP2, na aktywność minireplikonu wirusa grypy typu A. Wyniki eksperymentu wykazały znaczącą inhibicję replikacji wirusa w obecności ligandów, sugerując, że wybrane G-kwadrupleksy RNA mogą być potencjalnymi celami dla leków przeciwwirusowych. Aby uzupełnić nasze badania, zastosowaliśmy modelowanie molekularne do symulacji procesu tworzenia się wybranego G-kwadrupleksu RNA. Metoda to umożliwiła uzyskanie dodatkowych informacji dotyczących topologii G4 i stabilności struktury.

Podsumowując, stwierdziliśmy, że G-kwadrupleksy RNA są obecne wewnątrz genomu wirusa grypy typu A, szczepu A/California/04/2009(H1N1), gdzie wykazują wysoki poziom konserwatywności w różnych szczepach. Motywy te mogą być celem dla małocząsteczkowych ligandów specyficznych względem G-kwadrupleksów. Ponadto replikacja wirusa jest skutecznie hamowana poprzez stabilizację tych struktur z wykorzystaniem ligandów TMPyP4 oraz TMPyP2. Wszystkie uzyskane przez nas dane wskazują, że wybrane motywy G4 z genomu wirusa grypy typu A mogą być potencjalnymi celami nowych terapii przeciwwirusowych.

THE AIM OF THE PRESENTED PUBLICATIONS

The primary goal of my research was to determine whether potential quadruplex-forming sequences (PQSs) are present within the influenza A virus (IAV) genome (A/California/04/2009(H1N1) strain). Moreover, we intended to study their folding into RNA G-quadruplex structures. The research also aimed to investigate the interactions between selected RNA G4 structures from the IAV genome and G4-specific ligands. Finally, we explored the potential role of RNA G-quadruplexes, found within the polymerase subunit-coding regions, in the viral replication cycle. The aims of my research overlap with the topics of the three publications included in my dissertation.

1. The main goals of the research presented in the publication **"Identification and structural aspects of G-quadruplex-forming sequences from the influenza A virus genome"** were to identify the potential quadruplex-forming sequences (PQSs) within the influenza A/California/04/2009(H1N1) genome, determine their conservation level across IAV strains, and establish their folding into RNA G-quadruplex structures *in vitro*. In addition, we intended to obtain information on the thermal stability and structural topology of the adopted RNA G4s.
2. The review **"Examples of structural motifs in viral genomes and approaches for RNA structure characterization"** aimed to summarize the most common structural motifs found within viral genomes and to underline their importance in the viral replication cycle. We particularly focused on the motifs found in the IAV and SARS-CoV-2 genomes. Additionally, we collected and demonstrated current knowledge about approaches for RNA structure prediction and experimental validation.
3. In the research included in the publication entitled **"The interaction of RNA G-quadruplexes from the influenza A virus vRNA with TMPyP4 and BRACO-19 ligands"**, we demonstrated that RNA G-quadruplexes from the IAV genome, which were identified in our first publication, interact with G4-specific ligands, i.e., TMPyP4 and BRACO-19. Our primary objective was to observe differences in the binding mechanisms of these two ligands to RNA G4 structures. Moreover, we employed TMPyP4 and its analogue to explore the potential role of RNA G-quadruplexes located within polymerase-subunit coding regions during the IAV replication cycle.

INTRODUCTION

Although ribonucleic acid (RNA) was once considered a predominantly single-stranded molecule, it is now well established that it folds into a wide range of secondary and tertiary structures[1–3]. Consequently, numerous studies have demonstrated that RNA function is determined not only by its nucleotide sequence but also by its structural organization[4–6]. Common RNA structural motifs include duplexes, hairpins, terminal and internal loops, i-motifs, and G-quadruplexes, which perform diverse cellular functions[7,8]. At a higher level of organization, RNA tertiary structure is formed through long-range intramolecular interactions, such as pseudoknots arising from contacts between distant regions, kissing hairpin loops, and coaxial helices[7,8].

RNA structure plays a crucial role in regulating key biological processes, particularly during gene expression. Structural motifs within mRNA have been shown to modulate translation efficiency by influencing mRNA binding, codon recognition during translation initiation, the activity of initiation and elongation factors, and the folding of the synthesized proteins[9,10]. Specifically, RNA secondary structures can mediate internal ribosome entry site (IRES) recognition and restrict access of translation initiation factors. When located within the 5′ untranslated region (5′UTR), these structures can also fine-tune translation rates[11]. In addition, RNA structural elements are known to influence alternative splicing[9]. Hairpin structures involved in exon skipping and G-quadruplexes, which are proposed to interact with splicing regulatory elements, represent the most common splicing modulators[9,11].

Beyond cellular RNAs, structured RNA motifs are also widespread in viral genomes, where they play essential roles throughout the viral replication cycle[12,13]. Influenza A virus (IAV) represents a major global health concern. Its ability to cause seasonal epidemics and sporadic pandemics is associated with high morbidity and mortality. The devastating impact of the 1918–1919 Spanish influenza pandemic, which resulted in 50-100 million deaths, highlights the pandemic potential of this virus[14]. This potential is largely driven by the high genetic variability of the influenza genome. IAV variability is closely linked to viral pathogenicity, replication efficiency, growth kinetics, and multiple stages of the viral life cycle, including RNA replication, packaging, editing, and mRNA splicing. Importantly, many of these processes are regulated by the structure of viral RNA[15–18].

Influenza viral RNA (vRNA) is actively involved in viral replication and transcription through mechanisms dependent on its secondary and tertiary structures[12,19,20]. The selective packaging of all eight vRNA segments into a single virion is thought to be mediated, at least in

part, by structured RNA motifs protruding from viral ribonucleoprotein complexes (vRNPs). Within viral particles, vRNPs adopt a highly flexible double-helical conformation, with the viral polymerase bound at one end and a looped structure closing the opposite end[21]. Notably, both vRNA and viral mRNA form complex secondary structures that are highly conserved among IAV strains[12,22–25]. Furthermore, dynamic changes in vRNA secondary structure occur at different stages of the viral life cycle and contribute to the regulation of viral replication and transcription. The functional importance of many viral RNA structural motifs has been confirmed experimentally[23,26–28]. Together, these findings highlight the regulatory role of vRNA structure in influenza virus biology. Given that each vRNP acts as an independent transcription and replication unit and that viral RNA structure is critical for efficient genome packaging, vRNA represents a promising target for antiviral therapeutic strategies.

Recent advances in RNA structure investigation have further revealed the presence of genome-scale ordered RNA structure (GORS) within viral genomes[25,29–31]. GORS refers to the extensive secondary structure formed across the entire viral RNA genome, distinct from localized structural motifs found in specific regions. This global RNA architecture is highly conserved, suggesting an important role in viral biology, and has been implicated in host persistence, particularly in mammalian viruses[32]. In addition to large-scale genome organization, long-range intramolecular RNA–RNA interactions within viral genomes are critical for viral fitness. Disruption of these interactions by mutations can reduce viral infectivity[33–36]. Overall, these findings underline the central role of viral RNA structure in influenza virus biology. They also highlight vRNA as a promising target for antiviral therapeutic tools. Common antiviral strategies aimed at RNA secondary structures include small-molecule ligands, antisense oligonucleotides, small interfering RNAs (siRNAs), and catalytic nucleic acids[13,27,37–39].

1. Influenza A virus

Influenza viruses belong to the *Orthomyxoviridae* family and constitute a serious threat to public health. Orthomyxoviruses are enveloped, negative-sense, single-stranded RNA viruses. The *Orthomyxoviridae* family comprises influenza A, B, C, and D viruses, of which influenza A, B, and C infect humans. Influenza A and B viruses are responsible for seasonal influenza epidemics, but among all influenza viruses, only influenza A virus (IAV) is known to cause pandemics. A pandemic may arise when a novel IAV emerges, gains the ability to infect humans, and is efficiently transmitted between individuals. In contrast, influenza C virus

infections typically result in mild disease and are not associated with epidemics in humans. Influenza D viruses primarily infect cattle and have not been shown to cause disease in humans[40–42]. In my doctoral studies, I focused on the influenza A virus, which causes seasonal flu epidemics and pandemics. Therefore, the following paragraphs are limited to this species.

The IAV infects the upper respiratory tract and is transmitted through respiratory droplets[43]. Every year, 3-5 million people have severe flu symptoms, and more than 250,000 people die[44,45]. Therefore, over the years, the IAV has become an important research subject for scientists worldwide. There is an urgent need to understand the mechanisms governing the viral life cycle, emergence, transmission, and pandemic potential[44]. Importantly, treating IAV is especially challenging due to its high mutational potential. Two main mechanisms of genetic diversity of the IAV have previously been described namely antigenic drift and antigenic shift[46]. Antigenic drift is the accumulation of point mutations in the viral genome that alter amino acid sequences, leading to the emergence of novel viral variants. The molecular mechanism underlying such a process is the absence of proofreading-repair activity in the viral polymerase[47]. This process is responsible for seasonal IAV epidemics[46]. On the other hand, antigenic shift occurs when co-infection with various IAV strains leads to reassortment of RNA segments. By this mechanism, the virus can adapt to the new host, and new virus subtypes can emerge[47]. If the novel subtype can be efficiently transmitted among humans, the influenza pandemic can occur[46]. Due to these reasons, finding the anti-influenza drug targets within genome regions that remain conserved among IAV strains and subtypes requires particular attention. A thorough understanding of the virion structure is essential for studying the influenza A virus and, therefore, developing new antiviral approaches.

1.1. Influenza A virus genome organization

Overall, the influenza A virus genome is segmented and consists of eight negative-sense RNA segments (Figure 1). Each segment and associated proteins form viral ribonucleoprotein complexes that are enclosed within a host-derived lipid envelope. Beneath this membrane lies the matrix protein (M1), which is important for the virion's structural stability. Additionally, three proteins that are essential for the viral life cycle are embedded in the viral membrane: hemagglutinin (HA), neuraminidase (NA), and the integral ion channel (M2). Two glycoproteins, HA and NA, form the spikes on the viral surface. The IAV can be classified into antigenic subtypes based on the antigenicities of these two surface glycoproteins. Currently, 18 different HA classes (H1–H18) and 11 NA classes (N1–N11) have been identified. Moreover, each segment is numbered according to its length - the longest is segment 1, and the shortest is

segment 8. The segments are also named after the main protein they encode. Thus, segments 1–8 correspond to PB2 (polymerase basic 2), PB1 (polymerase basic 1), PA (polymerase acid), HA, NP (nucleoprotein), NA, M (matrix), and NS (non-structural), respectively[12,48]. The core of IAV consists of viral ribonucleoprotein (vRNP) complexes, which are composed of viral RNA (vRNA) segments with multiple NP molecules and the polymerase complex subunits PB2, PB1, and PA. Each vRNP complex is separate and an independent replication-transcription unit. Another protein associated with the IAV particle is the nuclear export protein (NEP, also known as NS2), which plays a role in the nuclear export of vRNP complexes[41].

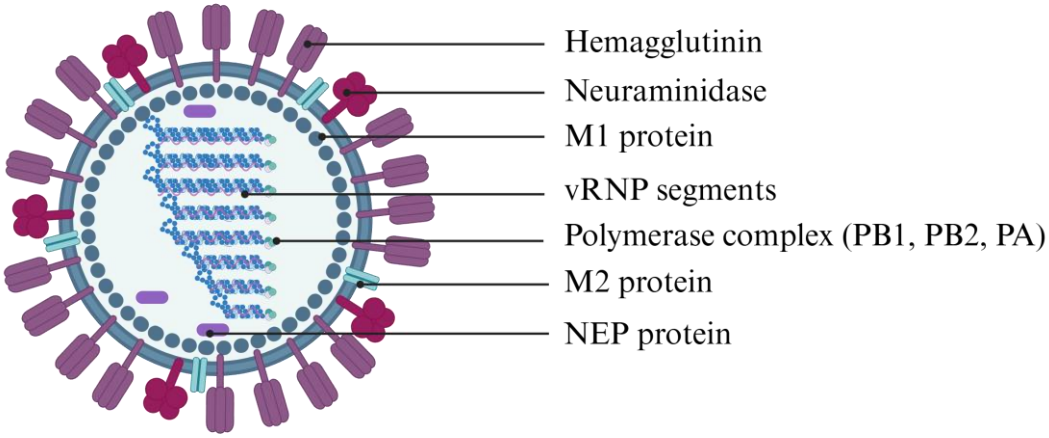


Figure 1. The structure of an influenza virus virion. Created with BioRender.com.

Importantly, all IAV segments contain untranslated regions (UTRs) at both the 3' and 5' ends, which are partially conserved across segments[41]. One of the most significant structural motifs present within IAV vRNA is the panhandle structure formed via base pairing of 5′ and 3′ ends within the promoter region[49,50]. Because this region is important for transcription and replication, it is highly conserved across all vRNA segments and IAV strains. However, several studies have shown that the panhandle structure forms only in the absence of the polymerase[12,50]. This finding suggests the dynamic nature of the vRNA structure *in vivo*. Consistent with this, studies on IAV mRNA demonstrated that secondary structures determined *in vitro* differ substantially from those observed *in vivo*. Notably, IAV mRNA appears to be largely unfolded in the cellular environment, although some structural motifs identified *in vitro* remain stable *in vivo*[22]. These observations suggest that the viral genome secondary structure should be studied not only *in vitro* but also in biologically relevant systems.

In general, the secondary and tertiary structures of IAV RNA have been confirmed to be complex, and the structure-function relationship should be analyzed. Apparently, IAV RNA secondary structure is crucial for many viral processes within the cell, including the viral

replication cycle, splicing, switching between transcription and replication, recognition by the host immune system, and vRNA packaging[12]. To date, many publications have reported on the secondary structure of the IAV genome[20,24,50,51]. Michalak et al. determined the *in vitro* secondary structure of the influenza A virus segment 5 vRNA (vRNA5) using chemical mapping[20]. The global secondary structure of vRNA5 was modeled by integrating experimental mapping data with free-energy minimization and computational analysis. Moreover, the authors found that many structural motifs are highly conserved across strains, suggesting their importance during the viral life cycle. Furthermore, based on the resultant model, the authors designed and tested short ASOs that bound accessible regions of the vRNA5 in cell culture[20]. In 2016, Lenartowicz et al. investigated the secondary structure of IAV RNA segment 8 *in vitro*[50]. To this end, *in vitro* chemical mapping, thermodynamic RNA folding predictions, comparative analysis of sequence/structure, and microarray mapping were performed. The authors found that segment 8 vRNA folds into stable, conserved secondary structures. Additionally, many of these motifs are thermodynamically favored and evolutionarily preserved, suggesting their role in the viral life cycle[50]. Importantly, in addition to examining viral RNA secondary structure *in vitro*, vRNA structure was studied *in virio* and *in cellulo*, allowing characterization of viral genomes in a biologically relevant context. An interesting finding was reported by researchers from our group, Szutkowska et al., who determined the secondary structure of segment 8 vRNA from influenza A/California/04/2009 (H1N1) in the presence of cellular and viral components[24]. The proposed segment structure largely corresponded with the previously determined *in vitro* structure. Several conserved, previously described motifs were also identified, indicating their potential role in the viral replication cycle. What is important from the clinical point of view, the segment secondary structure revealed regions accessible to novel potential therapeutics[24]. The follow-up studies showed complex folding of vRNA in cells, with partial preservation of motifs compared to the *in virio* structure[25]. Such investigations in the cellular environment are more complex and challenging; however, they provide important insight into the secondary structure of the viral genome in an infected cell.

These studies demonstrate that determining the secondary structure of IAV, including the arrangement of entire vRNA segments, is a crucial step in identifying novel antiviral targets. Moreover, discovering conserved non-canonical structures in viral genomes often reveals additional potential therapeutic targets. G-quadruplexes and their biological functions have been widely studied in various viral genomes[52–57]. Importantly, their significance is often associated with the degree of structural conservation[58,59]. Based on this, we aimed to determine

whether G-quadruplexes can form in IAV RNA and to evaluate the conservation of PQS motifs and their G-tracts.

1.2. Influenza A virus life cycle

The infection and replication of IAV are multi-step processes (Figure 2). During the viral replication cycle, not only vRNA but also complementary RNA (cRNA) and mRNA are synthesized. Because these RNA species differ in polarity and function, their secondary structures are not identical[41]. Therefore, when studying the secondary structure of the IAV genome, we should pay attention to whether a given structural motif is formed within vRNA, mRNA, or cRNA (Figure 2).

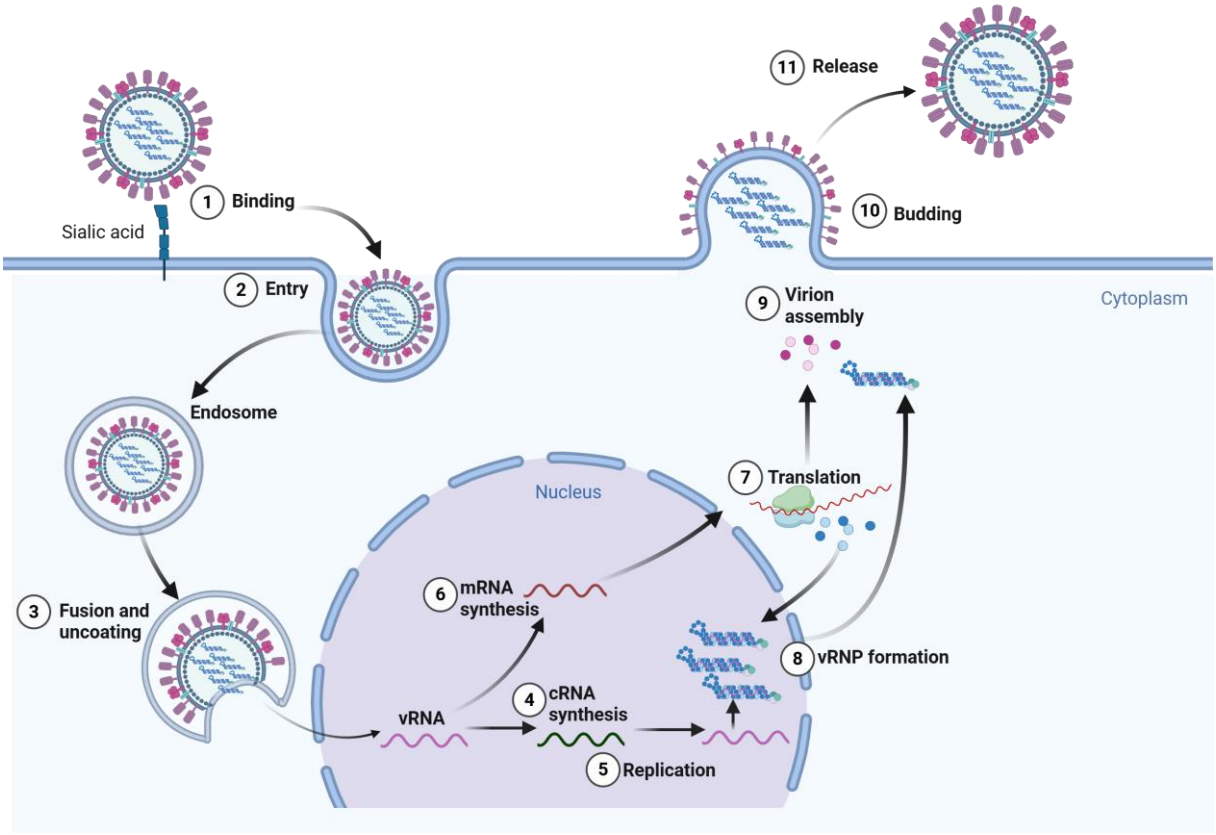


Figure 2. Influenza virus replication cycle: (1) the binding of virion to the host cell; (2) viral entry via receptor-mediated endocytosis (3) the fusion between the viral envelope and the endosomal membrane, the viral genome entry to the nucleus; (4) cRNA synthesis; (5) replication by RNA-dependent RNA polymerase complex; (6) mRNA synthesis – vRNA is transcribed into mRNA; (7) mRNA transport to the cytoplasm for translation, early viral proteins are transported back to the nucleus; (8) vRNP complex formation; (9) the assembly of new virions; (10) accumulation of viral subunits at the budding site leads to the budding of virions; (11) the release of progeny virions. Created with BioRender.com.

Overall, IAV infection of the host cells begins with viral binding to sialic acid residues on the cell membrane (Figure 2). The virus subsequently enters the cell via endocytosis and is initially localized within early endosomes. Fusion of the viral envelope with the late endosomal membrane is dependent on endosomal acidification, which induces pH-driven conformational rearrangements in hemagglutinin (HA). These structural changes expose the fusion peptide, thereby enabling viral–endosomal membrane fusion. After fusion, the viral particle is uncoated, allowing release of the viral genome into the cytosol. Once the vRNPs enter the cell, they are transported to the nucleus, and transcription begins. vRNA serves as a template for mRNA synthesis and then mRNA as a template for translation[41]. During vRNA replication, every vRNA segment acts as a separate replication unit. Complementary RNAs are synthesized, being copies of vRNA but in the positive-sense. cRNAs bind polymerase complex subunits and NP, forming the cRNP complex. Therefore, cRNAs serve as intermediates during replication as they are templates for novel vRNA molecules synthesis[12]. Finally, synthesized vRNPs are transported from the nucleus and trafficked to the plasma membrane. Viral transmembrane proteins, on the other hand, are translated into the ER membrane and transferred to the Golgi. At the last step of the viral replication cycle, the progeny virions are budded and released outside the cell[43] (Figure 2).

As previously mentioned, the secondary structure of vRNA is essential during various stages of the viral replication cycle and propagation. It has been confirmed that structural motifs have a major impact on selective genome packaging, assembly of new virions, incorporation of genome segments, and reassortment during co-infection with different viral strains[60]. Importantly, disruption of the structures involved in genome packaging reduces viral infectivity and attenuates pathogenicity in mice. This demonstrates how RNA folding directly influences viral replication[61]. RNA structural motifs formed within IAV mRNA are important functional elements during gene expression, thereby affecting replication efficiency and viral infectivity[22]. Moreover, the structures formed within viral mRNA in the proximity of splice sites are highly conserved and are likely to play regulatory roles in the processing of IAV RNA. By affecting mRNA splicing, in consequence, they have an impact on the production of proteins necessary for viral replication[62]. Also, the structure of IAV polymerase is unique – it is a heterotrimer consisting of three subunits that, when held together, form a U shape. This enzyme is highly flexible, but only when all subunits are present and bound to each other, the polymerase can catalyze the replication of the viral genome[43]. Taking into account the fact that structural motifs are widely distributed within the IAV genome and that they have a huge impact on the viral life

cycle, the secondary structure of IAV RNA represents a promising target for novel antiviral therapies[20,27,37].

1.3. Anti-IAV therapies

Due to the high genetic variability of IAV, most currently available strategies to control annual infections are only partially effective[63]. Approaches to combat IAV can be divided into two general strategies: prevention and treatment of infected individuals. The most common and effective prevention strategy is vaccination[64]. In addition to conventional influenza vaccines, RNA-based vaccines have been developed and shown high effectiveness in preventing IAV infection[65]. These include two main types: self-amplifying mRNA vaccines and non-replicating mRNA vaccines. A major advantage of RNA-based vaccines is their rapid adaptability, allowing vaccine formulation to be efficiently updated to the novel IAV variants emerging every season[65].

Despite these preventive measures, influenza infection rates remain high, largely because vaccination strategies are not fully effective and are not universally adopted. Consequently, most clinical management still relies on symptomatic treatment, while virus-specific antiviral therapies continue to be intensively investigated. One established approach to specifically treat IAV infection involves the introduction of neuraminidase inhibitors (NAIs). Neuraminidase plays a critical role in viral spread within the infected host. The NA cleaves off sialic acid residue from glycoproteins and glycolipids, enabling the release of progeny virions. Inhibition of this enzyme effectively blocks viral egress from infected cells and limits the progression of infection[64].

In recent years, increasing attention has been directed toward antiviral strategies that target viral RNA rather than viral proteins. Most RNA-targeted approaches focus on regions within vRNA and cRNA promoters or splice sites, although conserved RNA motifs located in other genomic regions are also being investigated[12,20,27,37,66–68]. These RNA-based strategies include the use of small molecules selective for specific RNA structural motifs, as well as antisense oligonucleotides (ASOs), small interfering RNAs (siRNAs), microRNAs, and catalytic nucleic acids[27,37,38,69]. The major advantages of such approaches are their high sequence selectivity and the possibility of chemical modification to enhance intracellular stability and inhibitory effects in the IAV-infected cells. The inhibitory effect of such tools within the IAV-infected cells has already been shown[20]. However, the effective design of RNA-targeted antivirals requires detailed knowledge of the secondary structures of vRNA, cRNA, and mRNA. In particular, identifying accessible regions within IAV RNA segments, such as

hairpins and internal loops, facilitates the discovery of potential therapeutic targets. This is important when these regions are involved in viral RNA–RNA or RNA–protein interactions[27,68]. For example, Soszyńska et al. combined the computational, experimental, and functional analyses to detect RNA secondary structures from the IAV genome that could be novel antiviral targets. First, the authors used a chemical mapping approach to determine the secondary structure of the IAV segment 5 (+) RNA. The obtained data were integrated into thermodynamic folding algorithms to generate a full-length RNA secondary structure model. Then, the biological significance of the RNA structures, together with conservation analysis, was evaluated to guide the design of antisense oligonucleotides targeting loops and internal regions. Importantly, the cell culture experiments revealed inhibition of IAV replication upon the addition of designed ASO[68]. Similarly, Piasecka et al. demonstrated that knowledge of viral RNA secondary structure can serve as a starting point for identifying antiviral drug targets. The authors focused on the mRNA of segment 5 of the IAV genome and used a secondary structure model in the preliminary selection of siRNA targets. As in Soszyńska et al.'s research, the targets were located within conserved motifs. Such an approach enabled authors to identify siRNAs, which inhibited IAV replication in cells[27]. Both studies have proven that viral genome secondary structures can be treated as potential antiviral targets, especially against viruses that evolve rapidly, such as the influenza A virus. Importantly, the continuous emergence of new IAV variants and subtypes leads to drug resistance, creating an ongoing need for novel antiviral strategies and combination therapies[64]. Additionally, numerous studies have demonstrated that G-quadruplex structures are widely present across viral genomes, such as SARS-CoV-2, HIV-1, Ebola virus, Zika virus, Hepatitis B and C, and IAV genomes[52,53,55–57]. All of the above motivated us to search the IAV genome for highly conserved structural motifs, specifically RNA G-quadruplexes, that could serve as potential new drug targets. In addition, we evaluated small-molecule ligands as potential inhibitors of IAV replication.

2. G-quadruplex structures

G-quadruplexes (G4s) are non-canonical RNA or DNA secondary structures that can form in guanine-rich sequences within the genomes of various organisms. Genome regions containing G-tracts, which can potentially fold into G4s, are called potential quadruplex-forming sequences (PQSs). G4 motifs are formed when four guanine residues associate through Hoogsteen hydrogen bonding to generate a planar array known as a G-tetrad (or G-quartet). Formation of a G-quadruplex structure requires the stacking of at least two G-tetrads, resulting

in a two-layer G4 stabilized by monovalent metal cations, typically K^+ or Na^+. Guanine runs in the genome, also known as G-tracts, are separated by other nucleotides, which, in G-quadruplexes, form the loops within the structure[70–72].

Regarding the G-quadruplex architecture, these structures are highly polymorphic and vary in molecularity, strand orientation, and loop geometry. However, it should be noted that only several topologies are energetically, thermodynamically, or kinetically favored and can thus be formed *in vitro*[73]. Importantly, G4s vary in the number of nucleic acid strands that fold into a G4 structure. G-quadruplexes can be intramolecular when formed by guanine residues located within a single strand or intermolecular, if guanines from two (bimolecular) or four (tetramolecular) separate strands associate with each other. Moreover, G-quadruplexes adopt various topologies depending on the strand direction concerning the 5′ to 3′ polarity of guanine runs. When all strands are in parallel orientation, the G4 is described to adopt a parallel topology (Figure 3A). On the other hand, when two strands proceed in the opposite orientation to the remaining two strands, the G4 topology is anti-parallel (Figure 3B). In addition, the hybrid (3+1) topology can be distinguished if three strands are in the parallel orientation and only one is antiparallel to them[71,73] (Figure 3C). The loops connecting the guanine tracts differ in geometry; they can be classified as edgewise (lateral), diagonal, or propeller (double chain reversal) loops. Notably, the loop type present in the particular G-quadruplex is an important factor in determining its structural topology and stability[71,73]. Besides the stand orientation and loop geometry, G-quadruplexes vary in the glycosyl bond angle of guanines, which can adopt *syn* or *anti* geometry. In parallel G4s vast majority of guanines are in the *anti* conformation, with several exceptions, whereas anti-parallel and hybrid topologies are characterized by the combination of guanines in both *syn* and *anti* geometries[73].

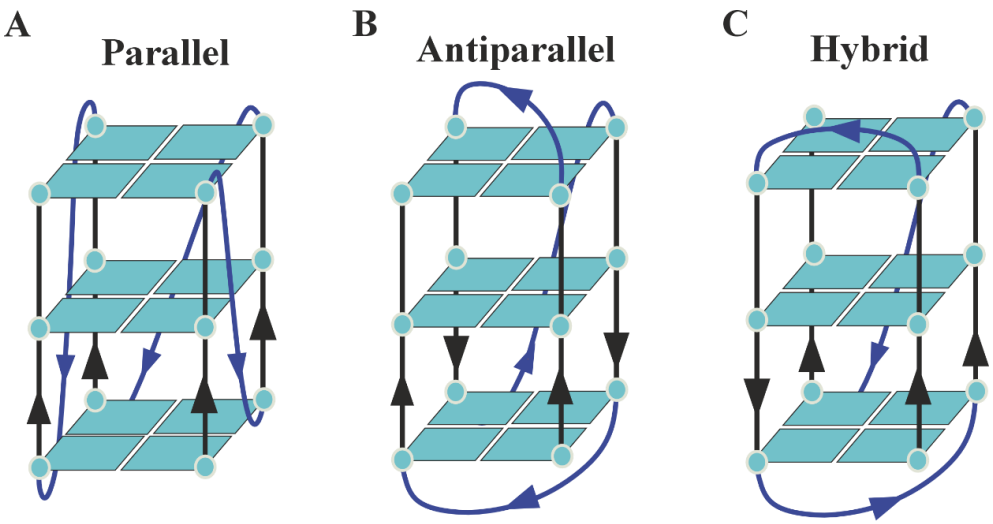


Figure 3. Schematic illustration of three G-quadruplex topologies: A) parallel, B) antiparallel, C) hybrid.

It is noteworthy that there are significant differences between G-quadruplexes formed in DNA and RNA sequences. Most RNA G4s fold into a parallel conformation, with several exceptions, e.g., in the human telomere RNA[74]. On the other hand, DNA G4s may adopt all mentioned topologies – parallel, antiparallel, or the hybrid one and switch between them[75]. Moreover, the 2′-hydroxyl group and water-mediated interactions within grooves lead to the higher thermal stability, and RNA G4s are more compact than DNA G4s. The 2′-hydroxyl group also impairs the binding of certain G4-specific ligands that typically interact with DNA G4s. The stabilizing cation binding preferences also differ between DNA and RNA G-quadruplexes. Potassium cations stabilize DNA and RNA G4s at a similar level; however, Na^+ has a significant stabilizing effect only on DNA G4s and not on RNA structures in certain cases[75].

2.1. Selected factors influencing G-quadruplex formation/stability

G-quadruplex formation and the structural stability depend on several factors that can be broadly categorized into environmental and structural features (Figure 4). The environmental factors include the type and concentration of metal cations, the solution's composition, and its pH. The structural factors contain properties of the loops (their sequence, length, and topology), chemical modifications, and the presence of ligands interacting with the G4 structure[76]. In the following section, I briefly describe several of these factors.

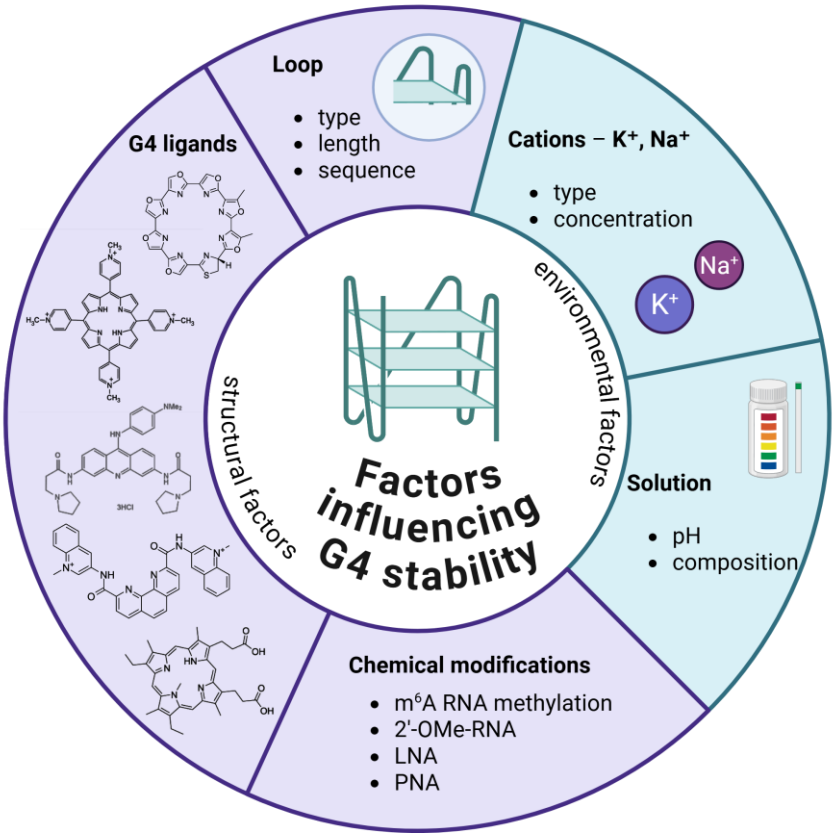


Figure 4. A schematic illustration of the key factors influencing the G4s formation and stability. Created with BioRender.com.

2.1.1. Effect of metal cations

The crucial environmental factor determining the sequence folding into G-quadruplex and the G4 stabilization is the presence and type of metal cations in the solution[77,78]. The interaction of G4 with a monovalent cation has been described to occur in the central channel of the G-tetrad. However, depending on the G4 conformation and cation type, the ions may be along the planes of the quartet or between them[79]. Most studies investigate the influence of biologically relevant cations, K^+ and Na^+, on G-quadruplex stabilization. Nevertheless, other monovalent metal cations, such as NH_4^+, Li^+, and Tl^+, and divalent cations (Mg^{2+} and Ca^{2+}) have been shown to have an impact on G4 stability[77]. Apparently, divalent cations stabilize G4 structure differently by screening backbone charges, binding to the phosphate group, influencing loop and groove interactions[79]. An important factor determining whether a given cation binds and affects G4 stability is its ionic radius. More specifically, K^+ has an optimal ionic radius to be located between the tetrads, whereas Li^+ has too small and Cs^+ too big ionic radius, making them less efficient G4 stabilizers[73].

Besides stabilizing G-quadruplexes, the type of metal cation present in solution strongly influences the adopted G4 topology. Olejko et al. investigated how various cations affect the folding of the same sequence and found that a hybrid topology is favored in the presence of K^+, whereas Na^+ yields a mixture of topologies (basket, propeller, and hybrid)[80]. In contrast, Mg^{2+} and Ca^{2+} promote the formation of parallel G4s. In addition, the cation concentration has a crucial impact on the adopted G4 structure[80]. The increasing salt concentration drastically increases the G4 thermal stability[77]. Therefore, in the biophysical experiments of my doctoral research, I used a 10 mM potassium phosphate buffer containing 50 mM KCl, because the K^+ cation was confirmed to have the most stabilizing effect during G-quadruplex formation[77].

2.1.2. Effect of the pH of the solution

The pH of the solution is another environmental factor that has an impact on G-quadruplex formation. It can both determine whether a G-quadruplex forms and affect the G-quadruplex topology[81]. Several mechanisms underlying the influence of pH on G4 formation have been proposed[82–84]. Alterations in pH may change the protonation state of some nucleobases, for example, within loops, in consequence changing G4 topology. Moreover, pH changes affect the solution's overall charge and thus affect interactions with G4-stabilizing cations. Therefore, it has been proposed that changes in pH, along with other environmental factors, influence the G4 stability, or in some cases, cause the alteration of G4 folding topology[81]. Galer et al. found that a human telomere DNA sequence forms a two-layer G-quadruplex structure at neutral pH (7.0), whereas at acidic pH (5.0), it folds into a different two-layer G-quadruplex[82]. These two forms vary in their loop arrangement, base interactions, and the behavior of the loop-end residues. Interestingly, the G4 structure can switch between these two forms depending on the pH values[82].

2.1.3. Effect of loops

G-quadruplex formation and stability are also strongly influenced by the length and composition of the loops connecting G-tetrads. In general, loop elongation destabilizes the G4 structure[76] and may increase the structural heterogeneity[85]. Both parallel and antiparallel topology G4 can be formed when more than two bases are present within every loop in the structure[76]. Additionally, various interactions involving the loop regions may alter the structure. Interactions occurring between loop bases arising from hydrogen bonding and loop bases/G-

tetrads interactions affect the G-quadruplex structure and stability. The G4 structure can also be influenced by different ligands interacting with loop regions[76].

However, not only the length, but also the sequence of the loop has an impact on the G-quadruplex stability. Research has demonstrated that the presence of thymidine, cytosine, or 1′,2′-dideoxyribose in the loop significantly stabilizes the G4 structure compared to structures with adenine within the loop[86]. Thus, it can be concluded that purine bases destabilize the structure, while pyrimidines have a stabilizing effect[86]. The loop composition also influences the G-quadruplex topology, as even small differences in the loop sequence can lead to a different G4 conformation[85]. Interestingly, the loop permutation, the change in the nucleotide order within the loop, has an impact on G-quadruplex stability and topology[87].

2.2. G-quadruplex function in the cells

The first reports concerning the fact that four guanines can interact with each other within genomes to form unique structures emerged in the 1960s, when Gellert et al. indicated this type of guanine arrangement using crystallography[88]. Then, in the 1980s, G4s were revealed to be present within human telomere sequences, making these structures a popular research subject[89]. Since that time, there have been many studies regarding the function of G-quadruplexes within the telomere sequence[90–93]. Evidently, G-quadruplexes formed within this sequence take part in telomere length regulation, protection, and heterochromatin formation. In 2018, Xiao et al. demonstrated that a region in human telomere RNA adopts an anti-parallel G4 structure[74]. The presence of G-quadruplexes in the telomere sequence controls the telomerase's access to this region (Figure 5). Therefore, this structural stabilization with G4-specific ligands inhibits telomere extension. The same ligands are capable of changing the location of the telomere protection complex, leading to cell death and thus being potential anti-cancer agents[94].

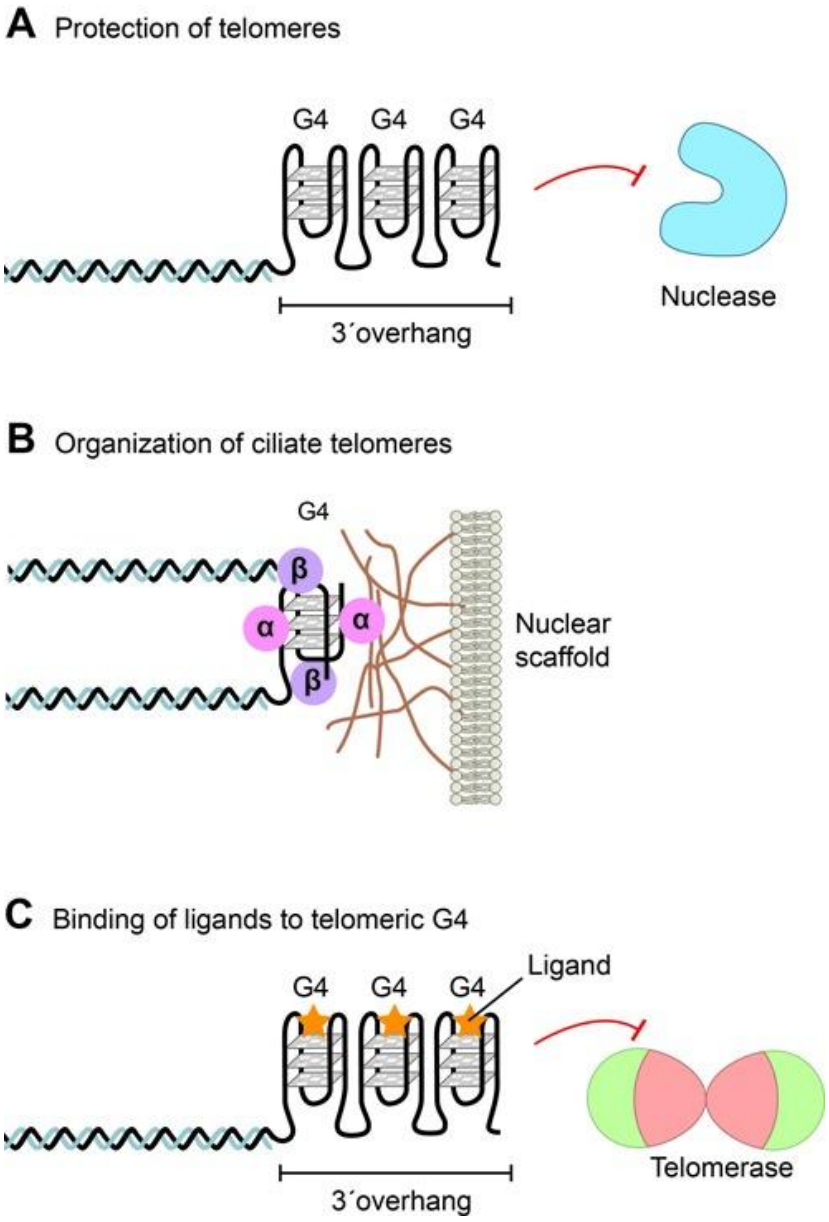


Figure 5. G-quadruplexes at telomeres. A) In human - intramolecular G4s formed at 3' overhangs protect chromosome ends from nucleases. B) In ciliates - intermolecular G4s involving two telomers are formed. The process is facilitated by the telomere-end-binding protein. C) Stabilization of G4s by specific ligands inhibits telomerase-mediated repeat synthesis[70].

Given knowledge of G-quadruplex formation within genomes, many research groups have focused on the correlation between G4 structure and function. Over the years, G4s have been confirmed to fold *in vitro*; however, the structures formed *in vitro* do not always reflect those adopted *in vivo*. Therefore, studies using biological models are needed to determine whether G-quadruplexes form in living organisms and to investigate their potential biological functions. Currently, it is known that G4s play important regulatory functions in transcription, translation, and telomere-related processes[94]. Besides telomeres, where G4s have been

extensively described, these structures form within promoter regions, DNA origins of replication, intron/exon borders, the target of immunoglobulin gene class switch recombination, and near somatic copy-number alterations. Approximately 50% of human genes contain sequences that can fold into a G-quadruplex in the proximity of their promoters, suggesting the potential function of G4 structures in gene expression[70]. During translation, the G-quadruplex fold within mRNA can act as an obstacle for proceeding ribosome, thereby regulating reaction efficiency. However, it has been shown that mRNA initially forms a stable hairpin rather than a G-quadruplex[95]. If translation begins before this hairpin can rearrange into the G4, the G-quadruplex cannot act as a ribosomal roadblock. Consequently, this leads to a disruption of translational regulation[95]. This demonstrates that secondary structures within the genome can compete with one another, each exerting a different influence on cellular processes.

Due to their abundance in the genome and their established importance in biological processes, G-quadruplex regulation has been associated with various diseases, including cancer, neurological disorders, developmental defects, and viral pathogenesis. In many cases, the mechanism underlying the disease is not correlated with the formation of G4 itself, but rather with the binding of G4 by different proteins. In cancer, it was proven that G-quadruplexes regulate the expression of cancer hallmark genes in cells[96]. Moreover, the G-rich motifs with propensity to fold into G4 structures were also identified in bacterial and viral genomes[70]. This fact encouraged us to investigate the IAV genome for PQSs and their potential biological function.

2.3. G-quadruplexes within viral genomes

To date, numerous studies have confirmed the presence of G-quadruplex structures in the genomes of different viruses, including SARS-CoV-2, HIV-1, Ebola virus, Zika virus, Hepatitis B and C, and IAV[52,53,55–57]. In many cases, those structures are highly conserved within viral genomes, suggesting their important role in regulating viral replication. Due to their apparent role in the viral life cycle, not only their identification, but also the discovery of their biological function has gained special attention.

Because G4 structures are often present in regulatory regions of viral genomes (promoters, UTRs, coding sequences), it is evident that they play crucial roles in regulating the viral life cycle, thereby affecting the progression of viral infection. They can also be involved in maintaining viral latency[97]. Furthermore, G4s are predicted to hinder gene expression in viruses, so their formation within the genomes can be directly correlated with replication rate[98]. Research on viral genomes can be challenging as various viral strains often differ in their

nucleotide sequences. This genetic variability results from a high mutation rate in viral genomes and is determined by virus- and host-related processes. Importantly, the mutation rate is higher in viruses containing RNA and single-stranded genomes[99]. The accumulation of mutations within the viral genome leads to the novel viral strains emergence[100]. In turn, those alterations affect the genome's secondary structure as they may disrupt previously existing genome secondary motifs[99]. Therefore, it is crucial to study genomes of distinct viral strains for the presence of G-quadruplexes. In general, viruses exhibit a wide range of genome types – including double-stranded (ds) DNA (HSV-1 and HSV-2), single-stranded (ss) DNA (canine parvovirus), dsRNA (rotavirus), ssRNA with positive (SARS-CoV-2) and negative (influenza viruses) sense, and viruses utilizing reverse transcriptase during their replication (HIV)[101]. Therefore, the complex studies that thoroughly examine their genomes for G-rich regions are necessary. Such research was conducted by the Brazda group, which performed studies similar to ours and investigated all newly sequenced variations of the H1N1 influenza virus to find PQS motifs[52]. The authors demonstrate that PQS location within genomes is not random and may be associated with biological processes[52].

A study on the Zika virus revealed conserved G4 structures within its genome that can bind G4-specific ligands[53]. Similar research on the Ebola virus has shown that G4 structures within its genome, when targeted with ligands, can inhibit gene expression[57]. In SARS-CoV-2, G4s are formed within both positive- and negative-sense RNA strands (+gRNA and -gRNA, respectively). Moreover, it has been confirmed that human cellular proteins bind to G4s and unfold them, thereby serving as an important regulatory factor in the viral replication cycle[55]. The phenomenon of host proteins interacting with viral G4s, influencing the viral life cycle, has been observed in many viruses[98]. It is worth mentioning that these interactions can lead either to the unfolding of a G4 structure, as seen in the case of the SARS-CoV-2 G4, or to the stabilization of a structure. An example of the second mode of action is the interaction between G4 from the HIV-1 genome and nucleolin, which results in repression of transcription[97].

In my research, I focused on studying the IAV genome, specifically the A/California/4/2009 (H1N1) strain, to identify PQSs within its RNA and investigate their potential function. Similarly, Brazda et al. analyzed the IAV genome *in vitro* to identify such structures[52]. The studies of both groups, ours and Brazda's, have shown that RNA G4 structures can form within IAV genomes, and they are often present within regions important for viral replication. Overall, the presence and function of G4s in viral genomes indicate that those structures can serve as novel potential anti-viral drug targets.

2.4. Different approaches to target G-quadruplexes in the cell

As previously described, G-quadruplexes formed within genomes have diverse functions in biological processes. Due to this fact, they are an interesting target for various compounds capable of altering these processes, often having a therapeutic outcome. Several approaches to target G4s have been developed to date[102–105]. Importantly, the choice of targeting strategy depends strongly on the desired result – whether the G-quadruplex stabilization, unfolding, or formation is required[102,106]. Some compounds can control the dynamics of G4 folding and unfolding at various time points[107]. Additionally, the G4 target topology and desired selectivity are important factors while selecting the anti-G4 strategy.

The most commonly used compounds targeting G4s are small-molecule ligands. Multiple such molecules specifically interacting with G-quadruplex structures have been described[108]. They can act either as stabilizers or disruptors of G4; some of them can do both depending on various factors, including the target sequence or desired biological outcome[107]. The small molecule compounds are mostly based on derivatives of porphyrin, pyridine, acridine, fluoroquinolone, and bisquinolinium. Moreover, they contain a characteristic polycyclic planar aromatic ring in their chemical structure. The ligands vary in their mode of interaction with G-quadruplexes; however, each of them binds to the particular G4 topology[106]. In general, small molecules that interact with G4s can block protein-RNA interactions, inhibit viral replication, or modify oncogene expression[108].

The second strategy that has gained considerable interest in recent years involves the use of antisense oligonucleotides (ASOs) either alone or conjugated to G4-specific ligands[107]. Because the ASO mechanism of action relies on complementarity to the target sequence, the major advantage of this approach is higher selectivity for a particular G4 structure formed in the cell rather than for G-quadruplex topology in general[106,109]. Antisense oligonucleotides are often developed to induce, disrupt, or promote folding of G4 structure. In many cases, the ASO-based approach is designed to influence the level of gene expression, also in viruses[106,109]. Introduction of chemical modifications into ASOs, also targeting G-quadruplexes, allows modulation of stability, affinity, pharmacokinetic properties, and cellular uptake[102]. Such modifications can be present within the ASO backbone, sugar moieties, or nucleobases. The first group includes, among others, phosphorothioate (PS), triazole, and amide modifications. Examples of sugar modifications of ASOs can be locked nucleic acid (LNA) and 2′-O-Methyl (2′-OMe) residues. The third group contains, for instance, 2-thiouridine modification of ASO[110].

In addition to antisense oligonucleotides alone, ASO-ligand conjugates are used to improve the properties of these G4-targeting compounds[104]. Although ASOs are highly specific to their target sequences, they still face several challenges. These include limited cellular uptake, insufficient structure-binding specificity, and variable biological activity. Moreover, a folded G4 structure may be inaccessible to ASO, which is designed only based on the sequence complementarity. Conjugating an ASO with a G4-specific ligand can help address these issues. Structural recognition of the ligand and sequence specificity of ASO can be combined to improve the overall properties of such tools[104]. This strategy has already been shown to selectively modulate G4 structures[104,111].

Another related approach involves peptide nucleic acids (PNAs) and PNA-conjugates to interfere with the folding of G-quadruplexes[107]. PNAs combine properties of nucleic acids and proteins. They contain a neutral backbone, allowing them to easily hybridize with RNA and DNA target sequences and to fold into a G-quadruplex structure. This way, they may interfere with G4 formation[112]. However, the introduction of modifications in PNAs is recommended to increase their cellular uptake and binding specificity[113]. The development of PNA-conjugates (small molecules, peptides, and imaging compounds) also increases their therapeutic potential[114]. Moreover, peptides themselves may be utilized to target G4s with high selectivity, as the presence of an RNA-binding domain enables them to bind and unwind G4 structures[108]. G-quadruplexes can also be targeted by L-aptamers, which are nuclease-resistant "mirror image" enantiomers of RNA molecules[108]. The development of the first G4-specific aptamer targeting RNA G4 was proposed by Chan and Kwok in 2020.[115] They developed an aptamer that forms its own G4 motif, binds to a specific G-quadruplex, and blocks its interactions with proteins[115].

Targeting approaches to detect and visualize G-quadruplex formation in the cells have also been developed. Among the most commonly used are small-molecule fluorescent probes, metal complexes binding G4s, and conjugation strategies. The last group includes, for instance, small molecules with fluorophores or peptides conjugated with ligands[116]. An interesting example of a fluorescent probe is the G-quadruplex-triggered fluorogenic hybridization probe, which consists of two parts: a fluorescent light-up part selective for G4 structure, and a DNA fragment complementary to the sequence adjacent to the G-rich region. Such a probe can detect G-quadruplex structure formed *in vitro* and in cells with high precision[117]. Additionally, G4-specific antibodies, such as BG4, can be used to visualize G-quadruplexes in cells. It detects G4 structures by forming nuclear or cytoplasmic foci, enabling the mapping of G4 distribution *in vivo*[108]. Another fluorescent small molecule that binds preferentially to G-quadruplexes is

the *o*-BMVC compound. Upon binding to the G4 structure, it exhibits a significantly longer fluorescence decay time than when binding to other genome structures. Using fluorescence lifetime imaging microscopy (FLIM) allows visualization and quantification of G-quadruplexes in cells via observation of *o*-BMVC foci[118].

Summing up, G4-targeting strategies have become a rapidly growing field in recent years. However, the delivery of the G4-specific compound, its penetration through the cell membrane, its stability within the cell, and its effect validation remain major challenges. Therefore, there is a constant need to develop next-generation G4-targeted agents with improved properties and applications in antiviral or anticancer therapies.

3. Methods to study G-quadruplex structures

Currently, various bioinformatic approaches are available to predict G-quadruplex formation within the sequence. First, the potential quadruplex-forming sequences (PQSs) must be identified within the genome. To achieve this, computational tools are employed to predict PQSs directly from RNA or DNA sequences *in silico*. These programs are typically based on algorithms that scan the input sequence for regions/fragments with characteristics indicative of PQSs[94]. Additionally, G-quadruplex folding can be investigated via molecular dynamics (MD) simulations. Molecular modeling allows investigation of the folding behavior of G-quadruplexes and their rearrangement under various conditions[73].

After bioinformatic prediction, identified PQS motifs require experimental validation to confirm if they actually fold into stable G4 structures. G-quadruplex formation can be investigated using numerous biophysical approaches. Overall, they can be divided into low- and high-resolution methods. Among the most commonly used low-resolution techniques are circular dichroism and UV melting, whereas high-resolution methods include X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. Other, more advanced experimental techniques are also available for G-quadruplex confirmation. They involve mapping by chain-extension stalling (e.g., reverse transcription stop assay), various chemical mapping methods, antibody-based mapping, and G4 imaging (e.g., immunofluorescence with G4-specific antibodies)[94]. The PQS motifs that form stable G4 structures *in vitro* can be further investigated using biological models to determine their potential functions within the cell. In the following subsections, different methods for studying G-quadruplex formation are described in detail, with particular focus on the methods I used in my doctoral research (Table 1).

Table 1. Summary of methods applied in my doctoral research, categorized by computational, biophysical, and biological methods.

Computational methods	Biophysical methods	Biological methods

• Bioinformatic tools:
  ◦ QGRS Mapper
  ◦ G4RNA Screener
• Conservation level analysis:
  ◦ WebLogo application
• Nucleotide composition comparison
• Molecular dynamics (MD) simulations
	• Spectroscopic methods:
	  ◦ circular dichroism (CD)
	  ◦ UV melting analysis
	  ◦ nuclear magnetic resonance (NMR)
	  ◦ fluorescence spectroscopy
	• Electrophoretic techniques:
	  ◦ PAGE experiment
	  ◦ Reverse transcription (RT) stop assay
	• Isothermal titration calorimetry (ITC)
		• Minireplicon assay based on luciferase expression

3.1. Bioinformatics analysis of PQSs

Bioinformatics analysis of the sequence of interest is usually the first step in studying G-quadruplex structure. Because there are specific rules that govern G4 formation, it is possible to make a preliminary assessment of whether G4 structure is formed based exclusively on its sequence. Several bioinformatic tools are currently available to search for the PQSs within the genome, including QGRS Mapper, QuadDB, Quadfinder, Greglist, or QuadBase[96]. Additionally, other tools can be used to identify PQSs; however, a common feature among the abovementioned is that they are all based on the same algorithm: (G3+N1–7G3+N1–7G3+N1–7G3+, where G refers to guanine and N is any base, including guanine). Importantly, a range of bioinformatics tools, such as G4Hunter, take into account the sequence surrounding PQS of interest[96].

Moreover, computational methods are available to assess the conservation level of sequences or their specific regions, for example, across different viral strains. To this end, the WebLogo application can be used to evaluate the conservation levels of both the sequence and G-tracts. It is also possible to analyze nucleotide residue frequency within the sequence using, for instance, the MAFFT software and the BioEdit 7.1 program[119].

Interestingly, bioinformatics analyses may reveal alterations within the nucleotide composition of G-rich sequences between different viral strains. Point mutations occurring within G-tracts raise the question of whether these changes could disrupt the formation of G-quadruplex structures in viral genomes[106].

3.2. Molecular modeling of G-quadruplexes

G-quadruplexes, like other secondary structures, can be investigated via molecular modeling[73]. Molecular modeling is a powerful tool for detailed prediction of PQS folding into G-quadruplex structures and their interactions with ligands. Such simulations allow us to acquire atomistic-resolution models, making them a valuable approach, for instance, in drug design. In molecular modeling, particularly in molecular dynamics simulations, both ion-ion and G4-ion interactions should be considered. Therefore, various restraints chosen in the molecular modeling experiment can lead to different outcome structures[120]. In general, MD simulations are based on calculations of atomic motions in the system using Newton's equations of motion. The forces between atoms and their potential energies are estimated using molecular mechanical force fields[121].

Molecular modeling is a valuable method for investigating the conformation, stability, and dynamics of G-quadruplex structures[122–124]. However, it should be emphasized that such simulations serve as a complementary approach to biophysical and biological studies, as the resulting model provides a predicted G4 conformation that may differ from the structure formed *in vitro* or *in vivo*.

3.3. Selected methods to study G4 formation

Nevertheless, existing G4 prediction tools can only identify PQSs within the input sequence and cannot confirm whether they form G-quadruplex structures *in vitro*. Therefore, the use of biophysical methods is necessary to confirm the G-quadruplex formation. The results obtained from these methods enable the identification of candidates for further investigation in biological models[75]. Biophysical techniques, known as gold standards for G-quadruplex detection, provide specific parameters characteristic of G-quadruplex formation – thermal stability and other thermodynamic parameters, molecularity, folding topology, folding kinetics, ligand interaction properties[75,125–127]. In recent years, high-throughput approaches for studying RNA G4s have received significant attention, enabling the analysis of entire transcriptomes. In general, these methods combine G4-selective chemical or ligand-based probing with RNA sequencing to map G-quadruplexes transcriptome-wide. They detect chemical footprints or signals of reverse transcriptase stalling that mark formed G4s, or enrich transcripts containing G4s using G4-selective ligands[75]. As mentioned earlier, there are many biophysical methods to determine G-quadruplex formation and its stability *in vitro*. However, in my doctoral research, I selected several commonly used techniques, and I described them in the following paragraphs.

3.3.1. Nuclear magnetic resonance (NMR) spectroscopy

The NMR method relies on the properties of atomic nuclei in which the nuclear magnetic moment (spin) interacts with a strong external magnetic field. The structure determination process is based on the analysis of magnetization transfer between protons through space mediated by carbon, nitrogen, and phosphorus atoms[128]. In our study, we employed proton nuclear magnetic resonance (^{1}H NMR), which specifically probes hydrogen-1 (H1) nuclei. The major advantage of using NMR to study G-quadruplexes is its ability to provide atomic-resolution structural information. Moreover, we can gain information about the structure, dynamics, stability, and interactions with other molecules. NMR spectroscopy measurements are performed in the solution, which can simulate cellular conditions. Importantly, proton spectra of G-quadruplexes reveal characteristic chemical shifts in the region between 10-12 ppm that are correlated with imino proton signals of guanines associated with G-tetrad formation. The number of peak signals should also correspond to the number of guanine residues, thus giving detailed information about the G4 structure. G-quadruplex-indicating signals can be easily distinguished from regions characteristic of other secondary structures, as Watson-Crick base pairing imino proton signals appear in the 12-14 ppm range[71,73]. The addition of a G4-specific ligand during the NMR experiment may provide evidence for G-quadruplex formation and its interaction with the ligand, as we should observe shifts in imino proton signals in the NMR spectra upon ligand binding.

3.3.2. UV spectroscopy

The commonly used, low-resolution spectroscopic method to study G-quadruplex formation is UV melting. In addition to confirming that DNA or RNA folds into a G4 structure, UV melting analysis provides information on the thermal stability of the G-quadruplex. Importantly, to observe the spectral curve profile characteristic of G-quadruplexes, associated with G-tetrad formation and dissociation, absorbance changes are recorded at 295 nm wavelength, where sample denaturation is demonstrated by a 50% alteration in absorbance amplitude. When the DNA or RNA molecule folds into G4, its formation is indicated by the inverted denaturation profile described as the hypochromic effect – a decrease of absorbance with increasing temperature. Based on UV melting measurements, we can obtain several thermodynamic parameters, including the melting temperature, enthalpy changes (ΔH), entropy (ΔS), and Gibbs free energy (ΔG)[71,119,129]. The melting temperature parameter provides information about the thermal stability of the structure, its resistance to thermal denaturation as

temperature increases. The obtained melting temperature is the temperature at which 50% of the structure is in the unfolded state. On the other hand, parameters such as ΔG, ΔH, and ΔS describe the thermodynamic stability of the structure[130]. However, the determination of the mentioned parameters is sometimes difficult in UV melting experiments for certain sequences. Measurement errors become substantial, preventing the extraction of reliable, quantitatively meaningful values of thermodynamic parameters. In such cases, only the melting temperature can be obtained, allowing for the assessment of the thermal stability of the structure, as demonstrated in my study.

Notably, UV melting experiments can be performed for the G-quadruplex structure alone or in the presence of various G4-targeting molecules, e.g., chemical ligands[131]. This way, the effects of the tested molecules on the G-quadruplex structure or stability and their interactions can be investigated[131]. Numerous ligands can be used in UV melting measurements, e.g., the compounds selected in my doctoral research: TMPyP4, BRACO-19, and TMPyP2.

Up to this date, various research based on UV spectroscopy has been published to confirm the formation of G-quadruplexes and/or to study the interactions of G4 with ligands[132–134]. Tluckova et al. investigated whether PQSs from human papillomavirus (HPV) fold into G4s. Among other methods, UV melting measurements were conducted to confirm G-quadruplex formation based on the melting curve profile and assess the thermal stability of PQSs[135]. In another study, ligand binding to G-quadruplexes formed within human telomeric DNA has been investigated using UV spectroscopy. Moreover, the UV melting analysis allowed the researchers to observe a different ligand binding profile depending on the G4 stabilizing cation in the solution[136].

3.3.3. Circular dichroism spectroscopy

Circular dichroism (CD) is another spectroscopic method that has attracted considerable attention in studying G-quadruplex structures due to its simplicity. The mechanism underlying CD measurements is based on the differential absorption of right- and left-handed circularly polarized light by chiral molecules. Nucleic acids that fold into G-quadruplexes usually reveal a band in the 210-300 nm wavelength range. This phenomenon is correlated with electron transitions in G-tetrads. In G4s, each topology is determined by *syn* and *anti* guanine conformation. Therefore, G-tetrads exhibit a different stacking pattern, which can be observed in the distinct shape pattern of the CD spectra[71,73]. CD profile of parallel G-quadruplexes is characterized by a positive peak around 260 nm and a negative at 240 nm. However, this only applies to tetramolecular G4 structures, as the additional positive peak around 295 nm indicates

the formation of parallel uni- or bimolecular G-quadruplex. On the contrary, the CD pattern with a positive peak at 240 nm and a negative one around 260 nm with an additional positive peak at 295 nm indicates an anti-parallel G4 topology. CD spectra of the hybrid topology of G-quadruplex exhibit a negative band at 240 nm and two positive bands at 260 and 295 nm; thus, it is difficult to distinguish from parallel uni- and bimolecular G4s[71,137].

Similarly to the case of UV meltings, CD spectra can be obtained for the G-quadruplex alone or after the addition of G4-specific compounds. Using this method, G-quadruplex interactions with ligands can be confirmed. Conformational changes in G4 resulting from this interaction are observed in the CD profile[126].

3.3.4. Fluorescence spectroscopy

An alternative spectroscopic technique to confirm that a specific sequence folds into a G-quadruplex *in vitro* involves the use of G4-specific ligands with fluorescence properties. Chemical compounds interacting with G-quadruplexes can be primarily divided into several groups depending on the effect obtained upon their addition. There are so-called "light-up" ligands, which bind to G4 and result in the fluorescent enhancement, "light-off" probes, which cause a reduction in fluorescence signal when the ligand interacts with the G4 structure, and permanent probes that bind selectively to the G4 structure; however, they do not cause any change in fluorescence signal. In general, the modes of interaction with G4 of such probes differ, as there are multiple binding sites within the structure[137]. Importantly, the result of G4-ligand interaction is visualized as the fluorescence emission spectra within the wavelength range characteristic of a particular compound.

The fluorescent ligand used in my doctoral research was thioflavin T, a compound previously shown to interact, for instance, with the telomeric G-quadruplex[138]. In this assay, fluorescence of thioflavin T is expected to increase in the presence of G4 structure. Furthermore, fluorescence spectroscopy was used to investigate the interactions between PQS motifs and two G4-specific ligands, TMPyP4 and BRACO-19. In this assay, we expected changes in the fluorescence spectrum and intensity upon binding of a selected ligand to a G-quadruplex.

3.3.5. Electrophoretic separation

Electrophoresis in native polyacrylamide gel (PAGE) can be a useful approach for detecting G-quadruplex formation as well as obtaining additional information about the folding

topology and molecularity. Based on differences in the migration of DNA or RNA in the folded state, we can determine whether the G4 adopts an intra- or intermolecular structure. Apparently, monomeric G4s migrate faster in the gel than unfolded molecules or dimeric and higher-order G-quadruplexes. This phenomenon arises from the fact that the G-quadruplex formed within one molecule is more compact than its unfolded state and therefore migrates faster through the gel. On the other hand, dimeric or tetrameric G4s involve more than one strand to form a G4; thus, the structure is more complex and does not migrate through the gel easily, resulting in lower electrophoretic mobility[139,140].

To visualize the bands indicating G4 formation and distinguish them from other structures, such as duplexes, gel staining using G4-specific dyes is necessary, e.g., thiazole orange derivatives or NMM. Thiazole orange has been described to exhibit strong fluorescence upon binding to the G4 structure[141]. Similarly, NMM has been shown to preferentially bind G-quadruplexes rather than duplexes[137]. However, this technique has certain limitations, including visualization problems (such as smearing) or difficulties with quantification[73]. Thus, the selection of an appropriate dye for G-quadruplex detection in gel-based assays is influenced by the specific G4 topology and other experimental parameters.

3.3.6. Reverse transcription (RT) stop assay

The method that allows for the identification of G-quadruplex formation within the sequence, as well as studying its interaction with G4-specific ligand, is a reverse transcription (RT) stop assay. The principle of this technique is that G-quadruplex formation has been shown to induce reverse transcriptase pausing[142]. During the RT experiment, this enzyme proceeds along the template RNA sequence until it reaches the G-quadruplex structure. As a result of reaction inhibition, an incomplete, shorter cDNA product is synthesized, which can be visualized during the PAGE experiment[143]. Moreover, G4 structure stabilization by various ligands, including TMPyP4 and BRACO-19, leads to increased enzyme stalling. Experiments on viral genomes revealed that ligand addition may inhibit viral replication, e.g., in human immunodeficiency virus (HIV)[144]. Interestingly, a similar approach can be used to study the pause sites, indicating the location of G4 in a high-throughput manner. In this variation of the experiment, instead of visualizing the product on the gel, the G-quadruplex-forming sites are analyzed by high-throughput sequencing[94].

3.3.7. Isothermal titration calorimetry (ITC)

The ITC method has been designed to investigate interactions between molecules, establish thermodynamics of these interactions, and provide insights into the forces responsible for specific binding. This technique can provide information about the enthalpy of molecule binding, binding affinity, and stoichiometry. From the ITC experiment, we can obtain several parameters, including dissociation constant (Kd), enthalpy change, Gibbs free energy change, and entropy change. The major advantage of this method is its high precision, as ITC has been employed to study G-quadruplexes and their interactions with different compounds[126,145]. In the ITC experiment, one molecule (e.g., G-quadruplex) is placed in the sample cell, while another molecule (e.g., G4-specific ligand) is added with increasing concentrations. In the moment of interaction of these two compounds, the release or absorption of heat takes place, and those changes (enthalpy) are analyzed. Released or absorbed heat reflects how much ligand is bound to the G4 structure at the moment. The peaks indicating the heat of each injection gradually decrease until they disappear. This stage is called the saturation, when all the existing binding sites are filled with ligands[146].

3.3.8. Biological studies of G-quadruplexes and their interactions with ligands

The final approach described in publication included in my doctoral thesis involved studies on G-quadruplex formation in cells, their interactions with G4-specific ligands, and the potential functions of these structures. There are various biological methods to investigate G-quadruplex folding *in cellulo.* These include the use of G4-specific antibodies and other G4-specific probes to determine the localization of G4 structures[147] and changes in gene expression levels. Additionally, genome mapping may be used to identify G4-induced pausing sites during replication[148]. However, a common feature of most biological methods is the use of G-quadruplexes-specific ligands that interact with the structure in the cell, influencing cellular processes. Importantly, in biological experiments, both the wild-type and the corresponding point-mutation variants need to be included. Mutant models often contain a point mutation within G-tracts, leading to disruption of G4 in the cell, and serve as a negative control in the experiment.

To study G-quadruplex formation within viral genomes, a minireplicon assay was developed[57]. The minireplicon system, also called a mini-genome assay, is an *in vitro* system that mimics viral genome replication and transcription machinery without using the infectious virus. Minireplicon contains essential viral RNA elements needed for these processes, but also

replaces most of the coding sequences with a reporter gene, e.g., luciferase or GFP. This enables functional analysis of RNA structures in a safer, simplified system. Such a non-infectious life cycle model was previously successfully used to investigate G-quadruplex formation within the Ebola virus genome[57]. In our case, the minireplicon system consists of five plasmids, including three plasmids with sequences encoding viral polymerase subunits (PB1, PB2, PA), one encoding viral nucleoprotein, and one encoding firefly luciferase (Figure 6). The selection of this system was justified, as the PQSs were located within those segments of the influenza A/California/04/2009 (H1N1) genome. Importantly, the bioinformatics analysis revealed the differences in nucleotide compositions of G-tracts within sequences from distinct viral strains: A/California/04/2009 (H1N1) and A/Puerto Rico/8/34 (H1N1). Therefore, we decided to create two minireplicon systems.

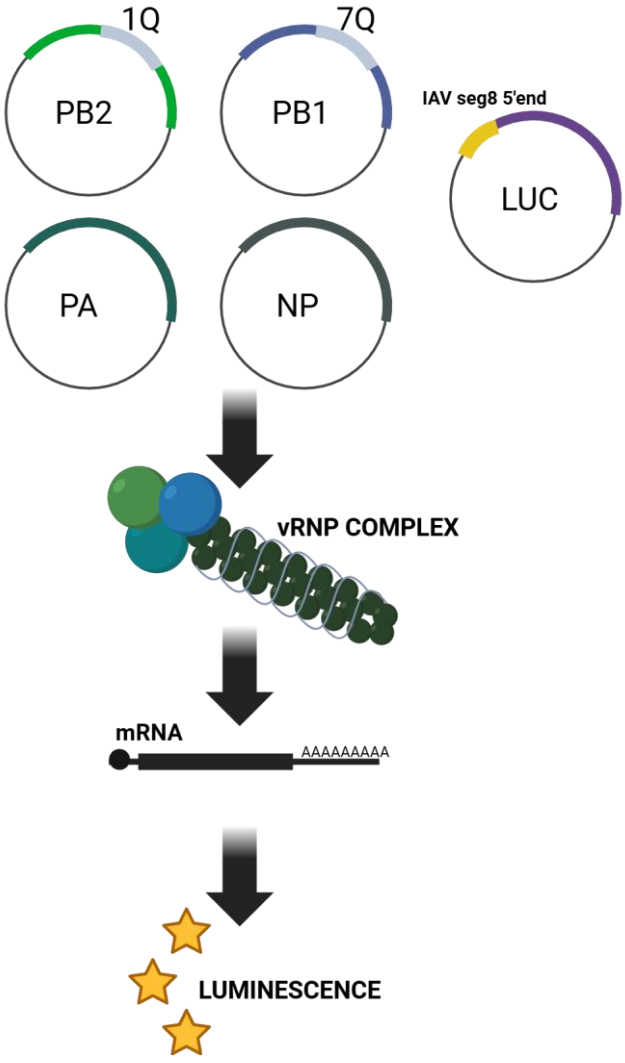


Figure 6. Scheme of the IAV minireplicon system used in the doctoral research.

4. G4-specific ligands

As previously mentioned, small molecules are commonly used as ligands that target G-quadruplex structures. Such compounds have potential applications in cancer therapy, where stabilization of G-quadruplexes by G4-specific ligands may downregulate oncogene transcription. Moreover, these ligands can also stabilize G-quadruplexes within telomeres, which has been associated with telomerase inhibition[149].

In addition to their roles in cancer research and telomere biology, G4-targeting ligands are also widely used to interact with G-quadruplexes present in viral genomes. From my perspective, this application is particularly important, as stabilizing viral G4s with small-molecule ligands has been shown to reduce viral infectivity, making these compounds promising antiviral agents[150–152]. The antiviral effect is generally attributed to the stabilization of G-quadruplex structures within viral genomes by these ligands[152].

Nevertheless, employing G4-specific ligands has both advantages and disadvantages. Their activity is largely determined by the structural features and stability of the G-quadruplexes they target. Therefore, when applied in cellular systems, potential off-target effects must be carefully considered. Many commonly used G4-binding ligands, including BRACO-19 and TMPyP4, while showing a preference for G-quadruplexes, can still interact to some extent with duplexes. Consequently, the development of more selective and specific ligands remains an important and ongoing research goal[126].

Despite the chemical diversity of G4-specific ligands, most of them share several common features. First of all, they possess large planar aromatic cores that enable their π-π stacking on terminal G-quartets. Secondly, their cationic side chains facilitate the interactions with negatively charged phosphate backbones. However, their modes of action vary: they can stack onto planar G-tetrads (BRACO-19), insert into the groove (distamycin-A), or bind to loop regions (acridine-triazole). Importantly, loop regions can adopt different conformations in response to bound ligands, thereby affecting ligand selectivity[153]. Over the years, many ligands recognizing G4 topology have been identified and applied in research[76,103,126,154]. However, in my doctoral research, I used five of them, i.e., BRACO-19, TMPyP4, TMPyP2, NMM, and thioflavin T (Figure 7). Therefore, in the following sections, I focus on them and briefly describe their properties.

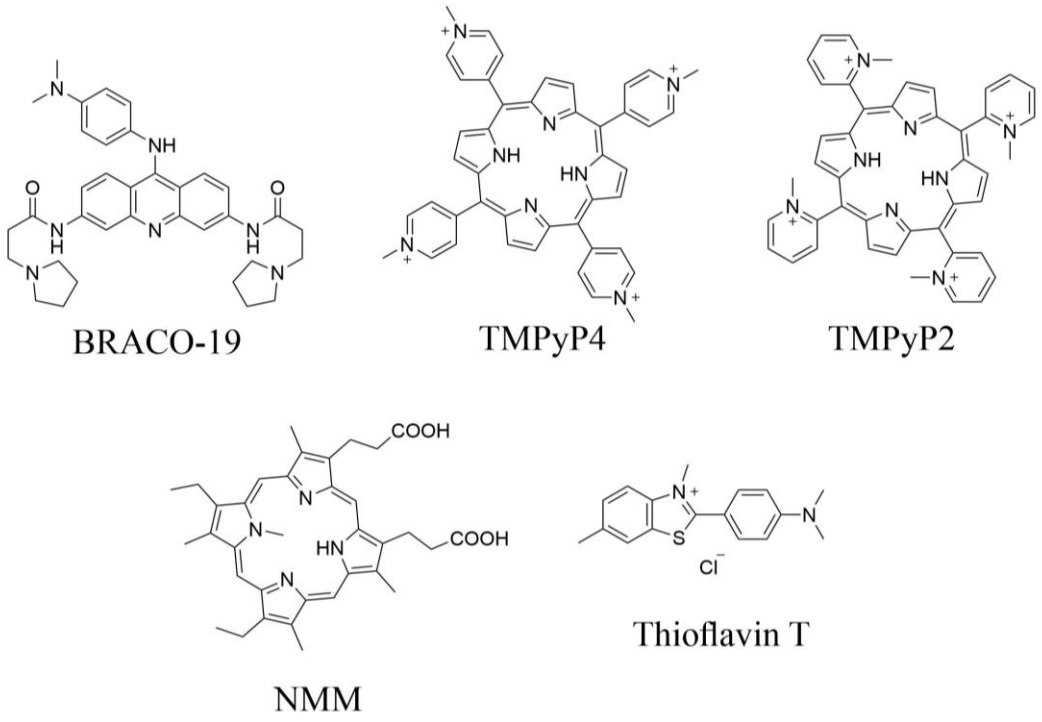


Figure 7. Examples of G4-specific ligands used in my research.

4.1. BRACO-19

BRACO-19 is a widely used, commercially available G4-targeting ligand; thus, I decided to employ it in my doctoral study. Structurally, BRACO-19 is a 3,6,9-trisubstituted acridine derivative[144] that binds G-quadruplexes with high affinity and stabilizes these structures. This stabilization can be confirmed by an increase in the G4 melting temperature[150]. BRACO-19 has also been investigated as a G4-stabilizing compound in viral genomes, where it has been shown to inhibit viral replication, supporting its use in my experiments[144]. Furthermore, BRACO-19 is a well-established stabilizer of telomeric G-quadruplexes, inhibiting telomerase activity[126]. Importantly, this ligand exhibits low cytotoxicity, as even concentrations up to 100 μM do not cause significant cell death[150]. Additionally, the BRACO-19 compound does not exhibit any fluorescence properties. One proposed mechanism of BRACO-19 interaction with G-quadruplexes involves asymmetrical stacking onto G-tetrads[155]. More recent molecular dynamics studies suggest that its binding mode depends on the G4 topology: end-stacking with parallel, bottom stacking with antiparallel, and top stacking with hybrid G4 structures[126]. For instance, studies of telomeric G-quadruplexes indicate that BRACO-19 can interact via top stacking, groove binding, or bottom intercalation[156].

BRACO-19 has been widely described to interact with G-quadruplexes in viral genomes. In 2014, Perrone et al. published an article regarding BRACO-19 binding to G4 from the HIV-1 RNA genome[144]. Importantly, this compound demonstrated antiviral activity by stabilizing G4 structures. This stabilization, in turn, inhibited reverse transcription of HIV-1 RNA, thus affecting viral replication[144]. A similar study was conducted on the G4 structures from the Herpes Simplex Virus-1 (HSV-1) genome[157]. Researchers demonstrated that BRACO-19 bound G-quadruplexes from the HSV-1 genome and altered their topology, resulting in increased thermal stability. Such stabilization resulted in the inhibition of viral DNA replication in infected cells[157]. Lv and Zhang investigated how BRACO-19 interacts with G4s from the enterovirus genome (in particular, EV-A71 virus)[158]. They confirmed that this compound stabilizes G4 structures present within the EV-A71 genome. Similar to observations in HIV-1 and HSV-1, BRACO-19 binding disrupts viral processes, which results in reduced replication and infectivity[158]. Collectively, these findings indicate that targeting viral G4 structures with BRACO-19 can produce an antiviral effect.

4.2. TMPyP4

The cationic porphyrin, tetra-(N-methyl-4-pyridyl)porphin (TMPyP4), is another ligand widely used in G-quadruplex-related studies[155,159] that I selected for my doctoral research. The TMPyP4 compound binds to G4 structures, often stabilizing them[150]. However, some studies have reported that TMPyP4 can also destabilize G-quadruplexes, although the mechanism underlying this unwinding activity remains unclear[126]. It is known that TMPyP4 primarily interacts with G-quadruplexes through π-π stacking, but its binding mode can vary. Reported binding modes include end-stacking on terminal tetrads, interactions within loop regions, and, in some cases, intercalation between G-tetrads[159]. Its binding can sometimes induce significant loop rearrangements[153]. Apparently, the mechanism of interaction can also depend on the nucleic acid type, as ITC analyses show that this ligand binds to DNA G4s via end-stacking and intercalation, whereas to RNA G4s via groove binding and intercalation[126]. The binding of TMPyP4 is associated with fluorescence quenching[155]. Importantly, at higher concentrations, this ligand shows greater toxicity than BRACO-19[150]. Similarly, as BRACO-19, the TMPyP4 compound also possesses anti-viral properties due to its specificity toward G-quadruplexes. By interacting with G4s and stabilizing them, it can, for instance, inhibit the progression of viral polymerase[160]. A major limitation of TMPyP4, however, is that it can also bind DNA duplexes to some extent[153].

TMPyP4 has been studied in the context of G-quadruplex stabilization across different viral genomes[57,160–162]. Qin et al. investigated whether this ligand targets G4 structures present within the SARS-CoV-2 genome[162]. They confirmed that TMPyP4 binds to these structures, enhancing their thermal stability. Moreover, the addition of this compound to infected cells resulted in suppressed translation of G4-containing sequences, reduced RNA replication, and, finally, decreased viral titer[162]. Research was performed on the G-quadruplex structures from the Ebola virus genome[57]. TMPyP4 treatment markedly reduced RNA-level expression of the Ebola virus gene containing G-rich sequences, due to stabilization of G4 structures by the compound. In addition, TMPyP4 inhibited the activity of the Ebola virus mini-genome in cells[57]. In 2021, Artusi et al. reported that TMPyP4 binds and stabilizes G-quadruplexes from the HSV-1 genome[160]. They clearly demonstrated this effect, showing that the ligand exhibits strong antiviral activity in infected cells. They also proposed an interesting mechanism of action for TMPyP4, involving the trapping of infectious virions inside cytoplasmic vesicles[160].

4.3. TMPyP2

TMPyP2 - 5,10,15,20-tetra-(N-methyl-2-pyridyl)porphyrin is a positional isomer of TMPyP4 confirmed to have lower affinity to G-quadruplexes, because of different peripheral cation positions[153,160]. Due to this fact, it is often used as a control probe in studies involving G4-specific ligands. In my doctoral research, I also employed this chemical compound to distinguish whether the observed biological effect is caused by the specific interaction between TMPyP4 and G-quadruplexes. TMPyP2 is a non-planar compound, in contrast to the planar structure of TMPyP4, which partially explains its weaker affinity to G4 structures. Overall, a non-planar ligand is unable to stack onto planar, flat terminal G-tetrads. Additionally, intercalative interactions require a greater distance between G-tetrads, further limiting TMPyP2 binding[163].

TMPyP2 has been commonly used in research on G-quadruplexes within viral genomes and their potential biological function[160,162]. The TMPyP2 compound is applied together with TMPyP4 to highlight its weaker binding to G4 structures. Artusi et al. applied TMPyP2 as the negative control to determine whether the observed effect after the addition of TMPyP4 was due to its specific interaction with G-quadruplexes rather than nonspecific TMPyP4 activity. The use of TMPyP2 in the experiment confirmed that the antiviral effect in the cells infected with HSV-1 is attributed with G4-targeted activity of the TMPyP4 ligand[160]. TMPyP2 was used for the same purpose in the study performed by Qin et al.[162]. The researchers used the ligand to validate TMPyP4's antiviral specificity in cells infected with SARS-CoV-2. Indeed, in viral

infection assays, TMPyP2 did not reduce the copies of viral RNA and virus titers, in contrast to TMPyP4, which significantly decreased both. This finding supports the conclusion that the effect of TMPyP4 is associated with G4 binding, as TMPyP2, which has a lower affinity for G-quadruplexes, did not lead to a comparable effect[162].

4.4. NMM

N-methyl mesoporphyrin IX (NMM), a water-soluble porphyrin, is another widely used G4-specific ligand that exhibits high selectivity for G-quadruplexes over other nucleic acid structures. Importantly, this ligand not only recognizes G4s but can also discriminate between different G4 topologies, showing a strong preference for parallel G-quadruplexes[155]. This fact was a key factor in our ligand selection, as most known RNA G-quadruplexes adopt a parallel topology[164]. Notably, the NMM ligand has been reported to possess excellent optical properties, making it a valuable G4-specific compound in biochemical studies. Its fluorescence increases markedly upon binding to DNA or RNA G-quadruplexes, allowing it to function as a "light up" probe[165,166]. In contrast, NMM addition to antiparallel G4 structures does not result in significant fluorescence, as the ligand does not interact with this topology[167]. Structurally, NMM is a non-planar and negatively charged compound, which is somewhat unconventional for a G4-binding molecule. It interacts with G-quadruplexes primarily through an end-stacking mode. Moreover, NMM promotes the formation of G-quadruplex/ligand complexes in the G-rich DNA sequences. Thus, disrupting the G-quadruplex structure results in a decrease in its fluorescence signal[168]. Importantly, NMM can induce a rearrangement of the G4 to a parallel conformation; therefore, care must be taken when interpreting experimental results[165].

4.5. Thioflavin T

Thioflavin T (ThT) is a water-soluble fluorogenic dye that exhibits high fluorescence enhancement upon binding to RNA G-quadruplexes[169]. This fact encouraged us to use this ligand in fluorescence measurement experiments. Thioflavin T may also interact with DNA G-quadruplexes. However, it is less selective towards DNA G4s since it also interacts with DNA that does not form G4 structures[169]. Molecular docking studies indicate that ThT binds G-quadruplexes through an end-stacking mode on the terminal G-tetrad. In this interaction, the benzothiazole ring is engaged in π-stacking, while the N3-methyl group is located in the central channel of G-tetrad[169].

Nevertheless, the selectivity of ThT toward G-quadruplex structures remains controversial, as numerous studies have demonstrated its affinity for other genomic secondary structures[170–172]. For instance, this ligand has been confirmed to interact with double-stranded DNA (ds-DNA) fragments that contain so-called cavities. Specifically, ThT can bind to dsDNA containing gaps, mismatches, or abasic sites, leading to fluorescence enhancement[173]. Moreover, thioflavin T has been shown to bind duplexes or triplexes, particularly those rich in GA motifs[170]. In this case, ThT can insert into the pocket formed between G-G and A-A motifs[170]. Consequently, fluorescence-based confirmation of G-quadruplex formation using ThT should be validated with an additional G4-specific ligand. Therefore, in my study, ThT assays were complemented with the NMM compound.

BRIEF DESCRIPTION OF THE PUBLICATIONS INCLUDED IN THE DOCTORAL DISSERTATION

1. Identification and structural aspects of G-quadruplex-forming sequences from the influenza A virus genome

International Journal of Molecular Sciences, 2021, 22, 6031

Tomaszewska, M.*, Szabat, M.*, Zielińska, K., Kierzek, R.

*co-authors who equally contributed to this work

It is well known that the influenza A virus remains a serious threat to public health. It causes hundreds of thousands of deaths every year, and currently available therapies are only partially effective. Additionally, this virus mutates very quickly, leading to the emergence of new variants and strains. Therefore, there is a constant need to search for novel anti-influenza targets within its genome. Moreover, the role of genome secondary structure in viral replication and other biological processes has received growing interest in recent years. Among other structural motifs, G-quadruplexes have been identified within viral genomes, including RNA viruses, and have been confirmed to play important roles during their replication cycle. Taken together, these observations motivated us to search the IAV genome for potential quadruplex-forming sequences (PQSs). The main goal of my research was to identify PQS motifs and then investigate their folding into G-quadruplex structures *in vitro* using various biophysical methods. Furthermore, we analyzed the conservation level of the identified PQSs across the IAV strains to assess the importance of these motifs in the viral life cycle. In our research, we chose the influenza A/California/04/2009(H1N1) strain, a rapidly spreading IAV H1N1 strain that caused a pandemic in 2009. The first infection was reported in California, and the strain was initially termed "swine flu"[174]. I would like to mention that most of the studies in our department on the structure of the influenza virus genome have been conducted using A/California/04/2009(H1N1) strain, which also influenced our choice of research object.

The first step of my research involved bioinformatics analysis to identify PQS motifs within the influenza A/California/04/2009(H1N1) genome. To this end, we used QGRS Mapper and G4RNA screener bioinformatics tools. During this analysis, twelve PQSs potentially folding into G4 structures were found within the IAV genome. Importantly, all of them were located within protein-coding regions of different RNA segments (Table 2).

Table 2. List of PQS motifs identified within the influenza A/California/04/2009(H1N1) genome.

Segment (protein)	PQS name	Sequence (5′-3′)
1 (PB2)	1KW	CU**GG**U**GGGG**CAGCAGCAAA**GGGG**AG
	6KW	**GG**UGCAU**GGGG**UUCAGUCGCU**GG**
2 (PB1)	2KW	UU**GG**CU**GG**ACCAU**GGG**CU**GG**CA
	7KW	**GG**UAGU**GG**UCCAUCAAUC**GGG**UUGAGCU**GGGG**
	8KW	**GG**CUGUAU**GG**A**GG**AUCUCCAGUAUAA**GGG**AAUGU**GG**
3 (PA) (PA/PA-X) (PA/PA-X)	9KW	UU**GG**A**GG**UUCCAUU**GG**UUCUCACAUAUA**GG**AA
	10KW	**GG**CAAAGA**GG**CCCAUCA**GG**CAAUCUGA**GGGG**
	3KW	AA**GG**CU**GGGG**AAGUUC**GG**U**GGG**AGACUUU**GG**UC
4 (HA)	11KW	**GG**AUGUAUAUUCUGAAAU**GGG**A**GG**CU**GG**
	4KW	UU**GG**UCAGCACUAGUA**GG**U**GG**AU**GG**
7 (M1/M2)	12KW	UC**GG**CUUUGA**GGGGG**CCUGAC**GGG**A
8 (NS1/NS2)	5KW	UC**GG**C**GG**AGCCGAUCAA**GG**AAUG**GGG**CA
G - G-rich sites, which can be involved in RNA G-quadruplex formation		

Thus, the next step was to analyze the conservation level of the identified PQSs. We first investigated nucleotide residue frequencies within the PQS motifs. Using the MAFFT software and BioEdit 7.1 program, we compared the similarity of PQSs across influenza type A and H1N1 subtype. As expected, for the majority of PQSs, the similarity was higher for strains of the same subtype. The final part of bioinformatics analysis was a PQS conservation analysis within the H1N1 subtype using the WebLogo tool. Notably, conservation levels varied among individual PQSs; however, most motifs showed 75-84% conservation. The conservation level of the G residues within the G-tracts of PQSs was also relatively high (around 70%). This high degree of conservation suggests that these motifs may play an important role in the viral life cycle. In the second part of the study, we aimed to assess the ability of selected PQS motifs to form G-quadruplex structures. To investigate this, a series of biophysical experiments employing several methods was performed and described in the publication. We applied various spectroscopic techniques, i.e., NMR, UV melting analysis, CD, fluorescence intensity measurements, and native gel electrophoresis. The combination of these complementary approaches allowed for a comprehensive analysis of PQS folding.

First, ^{1}H NMR experiments were performed to confirm the G-quadruplex formation. The measurements were conducted by dr Karolina Zielińska from the Department of Biomolecular NMR at our Institute. Overall, G-quadruplex formation is characterized by imino proton signals within the 10-12 ppm range, indicating Hoogsteen base pairing. Based on the

resultant NMR spectra, PQS oligomers were divided into three groups: PQSs that are predicted to fold into G-quadruplexes, PQSs that may adopt both G4 motifs and other secondary structures, and PQSs that do not fold into stable G4s. Moreover, the ^{1}H NMR experiment was performed for all PQS oligomers with the addition of G4-specific ligand, NMM. The observed shifts in imino proton signals upon ligand addition served as further confirmation of G-quadruplex formation. Next, UV melting analysis was performed to confirm that certain PQS motifs adopt a G4 structure and to establish their thermal stability. We analyzed the PQS melting curve profile at the 295 nm wavelength, as the over 50% change in the absorbance at this wavelength indicates G-quadruplex formation. Our goal was to observe the hypochromic effect, characterized by a decrease in absorbance with increasing temperature. However, we were unable to obtain reliable thermodynamic parameters and therefore focused on the melting profiles and changes in the melting temperature determined for the PQS variants. An inverted melting curve profile was observed for four PQS oligomers, 1KW, 3KW, 7KW, and 11KW, indicating G4 formation. Additionally, the melting temperatures of these four PQSs were determined, ranging from 28.6°C to 50.7°C. We further aimed to analyze G-quadruplex formation across all 12 PQS motifs. Circular dichroism spectroscopy was employed to confirm G-quadruplex formation and determine its topology. In CD spectra, we observed a negative peak around 240 nm and a positive peak around 265 nm for six PQS oligomers, 1KW, 3KW, 6KW, 7KW, 9KW, and 11KW. Such a CD profile was previously assigned to parallel topology G-quadruplex formation. As noted before, most RNA G4s adopt this topology[164], and our findings are consistent with these published results.

In the next step, fluorescence intensity measurements were conducted to confirm that PQS oligomers adopt G4 structures using a fluorescent dye. To this end, ThT, a ligand interacting with G-quadruplexes, was applied. We expected a significant increase in the fluorescence signal upon its binding to the G4 structure. Notably, we observed a strong fluorescence enhancement in the case of six PQS oligomers when ThT was added to the sample. This experiment demonstrated the potential of several PQS motifs, 1KW, 3KW, 5KW, 6KW, 7KW, and 11KW, to fold into RNA G-quadruplexes; however, the previously described affinity of ThT for other structural motifs besides the G4s must be taken into account. Given its lower specificity for G4s, ThT fluorescence can serve as a valuable complementary method to investigate PQS folding, but should not be the only indicator of G-quadruplex formation.

The last method used in this study was native polyacrylamide gel electrophoresis (PAGE). In this experiment, we aimed to both confirm the formation of G-quadruplexes and study their folding patterns. To this end, two G-quadruplex binding ligands, NMM and

thioflavin T, were used to stain the gels after electrophoresis. As expected, the bands visualized under UV light differed from those revealed after ligand staining, which identified only the bands corresponding to G4 formation. Moreover, differences in the electrophoretic mobility of RNA oligomers provided evidence of their various molecularity. Additionally, the appearance of multiple bands in the gel for several oligomers suggested an equilibrium between alternative conformations. Notably, the PAGE experiment confirmed our earlier observation that ThT exhibits lower specificity for G-quadruplexes compared with the NMM ligand.

In conclusion, the results from the selected biophysical methods revealed that three PQS motifs from the influenza A/California/04/2009(H1N1) genome, named 1KW, 7KW, and 11KW, can fold into RNA G-quadruplex structures *in vitro*. To our knowledge, this is the first study to show that PQS motifs are present in this IAV strain and can fold into RNA G4s. The localization of these PQSs within protein-coding regions, specifically in the PB1, PB2, and HA genes, and their high conservation across IAV strains suggest that these motifs may be functionally important during the virus life cycle. We believe that the sequences characterized in this research represent promising candidates for further investigation as novel anti-influenza drug targets.

2. Examples of structural motifs in viral genomes and approaches for RNA structure characterization

International Journal of Molecular Sciences, 2022, 23, 15917

Nalewaj, M.*, Szabat, M.*

*co-authors who equally contributed to this work

To better understand the relationship between viral secondary structure and function, and to highlight recent advances in RNA-structure methodology, we authored a review paper. In this publication, we summarized the most common structural motifs found within viral genomes, with a particular focus on IAV and SARS-CoV. Additionally, we discussed approaches for RNA structure prediction and its experimental validation.

The first part of the review covers the most common structural motifs within viral RNA and their functional aspects. Because secondary and tertiary structures play crucial roles in many viral processes, such as replication, translation, and host immune evasion, the identification and functional characterization of structural motifs have attracted considerable attention. Initially, we described the functional domains, focusing on internal ribosome entry sites (IRES), which facilitate translation initiation, and cis-acting replication elements. Although these elements can differ in structure and type, their functional roles remain conserved. Additionally, in this review, we described genome-scale ordered RNA structures (GORS), which are believed to support viral persistence and regulate the host's innate defense.

In the next chapter, we characterized other secondary motifs present in viral RNA genomes, namely pseudoknots, panhandle motifs, kissing loops, and G-quadruplexes. Pseudoknots are known to form through base-pairing between a hairpin loop and a complementary region. Their structural diversity reflects the topology-function relationship. Previous research showed that viral pseudoknots are highly diverse and dynamic, with diverse functions, e.g., in the SARS-CoV-2 life cycle. Moreover, these motifs are confirmed to have functions within the IRES of viruses. Afterwards, we focused on the panhandle structure, a highly conserved motif from the IAV genome that plays a crucial role in replication and transcription. The panhandle is formed by base-pairing of the 5′ and 3′ ends within all IAV genome segments. We discussed possible rearrangements of this region into different alternative structures upon polymerase binding. Furthermore, we described the kissing-loop interactions, the RNA tertiary structures formed by intermolecular contacts. These motifs are involved in various cellular processes, for example, RNA dimerization, cyclization of viral RNA, regulation of replication, and virion packaging. The last structural motifs discussed in

our review were G-quadruplexes, which form within guanine-rich sequences. In this part, we highlighted several studies identifying conserved G4s in viral genomes, such as SARS-CoV-2, rhinovirus, Zika virus, and IAV. In these viruses, G-quadruplexes influence key processes, including uncoating, viral genome replication, and translation. We also noted that G-quadruplex stabilization with ligands can inhibit viral replication, making them promising targets for antiviral therapies.

In the second part of the publication, we focused on selected methods widely used to investigate the formation of structural motifs within viral genomes. This section was divided into two main parts: an explanation of computational approaches for RNA structure prediction and a description of experimental techniques to investigate the formation of structural motifs.

Many RNA secondary prediction tools are based on thermodynamic calculations, phylogenetic analysis, or a combination of both. We compared several computational tools in terms of their inputs, outputs, and main features. Additionally, we mentioned that these methods are broadly classified into single- and homologous-sequence approaches. Single-sequence tools, such as RNAstructure, RNAfold, and UNAfold, use the minimum free energy principle. RNAstructure provides additional support for advanced analysis and integration with experimental data. Homologous-sequence and deep learning methods, including TurboFold and DMfold, enhance prediction accuracy and enable modeling of more complex motifs, such as pseudoknots. Computational tools, such as ScanFold and 3D structure methodology, were applied to identify conserved and functional structural motifs and to visualize higher-order elements within viral RNA genomes, including Zika virus, SARS-CoV-2, and IAV, highlighting their potential as antiviral targets.

We also emphasized recent efforts to integrate computational tools with high-throughput sequencing and chemical probing, including SHAPE-MaP, SPLASH, vPAR-CL DMS-MaPseq. Collectively, these approaches enable single-nucleotide-resolution mapping of RNA structures as well as RNA-RNA and RNA-protein interactions *in vitro* and *in vivo*. Furthermore, we discussed advances in molecular dynamics (MD) simulations as a powerful tool for probing RNA folding (including viral RNA), conformational dynamics, and structure-function relationships. For instance, the MD method was used to study the non-structural protein of SARS-CoV-2. The researchers proposed a structural model of this protein and focused on antivirals targeting its specific domains[175]. Finally, in this part of the review, we underscored the importance of combining these approaches with high-throughput and virtual screening. Such an approach can support antiviral discovery, as RNA structure prediction can

be a first step in the drug design process. Therefore, we presented several examples to illustrate how such strategies contributed to the discovery of potential RNA virus inhibitors.

However, RNA structure derived from computational predictions should always be experimentally validated. Due to this fact, the final section of this review described several experimental methods employed in RNA structure investigations. Widely used methods for determining RNA secondary structure and interactions include SHAPE (Selective 2′ Hydroxyl Acylation analyzed by Primer Extension) and its high-throughput variants, such as SHAPE-Map and DMS-MaPseq. These approaches revealed functionally important RNA structural elements in viruses: Chikungunya virus, HIV-1, Zika virus, dengue virus, SARS-CoV-2, and IAV[176–180]. Moreover, long-range interactions in the SARS-CoV-2 genome were identified using these methods. We further described the advanced interactome methods such as PARIS, SPLASH, 2CIMPL, and RING-MaP, which enable the detection of more complex intra- and intermolecular interactions in viral RNA. In addition, we highlighted cryo-electron microscopy as a powerful complementary method for revealing high-resolution viral RNA assemblies, providing insights into the viral replication mechanism.

Overall, we concluded that the integration of experimental methods with bioinformatics and molecular modeling is crucial for understanding viral RNA structure. We also highlighted that RNA function is strongly dependent on its structure; thus, detailed characterization of viral RNA structure is essential for the development of effective antiviral strategies. Despite a significant advance, viral RNA variability highlights the importance of identifying conserved structural motifs, which may serve as promising targets for antisense therapies, so-called ‘programmable antiviral agents’, and the rational design of mRNA vaccines.

3. The interaction of RNA G-quadruplexes from the influenza A virus vRNA with TMPyP4 and BRACO-19 ligands

PLoS One, 2025, 20(11): e0335975

Nalewaj, M., Sliwiak, J., Zielinska, K., Zmora , P., Niedziałek, D., Wieczorek, G., Kierzek, E., Szabat, M.

In our first research publication, "Identification and structural aspects of G-quadruplex-forming sequences from the influenza A virus genome", we discovered three PQS motifs within the influenza A/California/4/2009 (H1N1) genome that form stable RNA G-quadruplex structures *in vitro.* In recent years, studies regarding G-quadruplex interactions with many G-quadruplex-specific ligands have gained a huge interest. Moreover, G4-binders have been shown to successfully inhibit viral replication by interacting with G-quadruplexes within viral genomes[181,182]. In our previous article, PQS interactions with two G4 ligands, NMM and thioflavin T, were demonstrated. Therefore, in our research presented in this publication, we investigated the interaction of three selected PQS motifs, 1Q, 7Q, and 11Q, with other well-known G4-specific ligands, BRACO-19 and TMPyP4. It is worth noting that the 1Q, 7Q, and 11Q motifs correspond to those selected in the first publication; however, in that study, they were designated 1KW, 7KW, and 11KW, respectively. Importantly, PQS 1Q was identified within the PB2-coding region, whereas 7Q and 11Q were located within the PB1- and HA-coding regions, respectively. The goal was to investigate whether the G4s from the IAV genome interact with these small-molecule ligands, what are the binding properties of our G4s to BRACO-19 and TMPyP4, and if the G4/ligand interaction influences G-quadruplex thermal stability. Finally, we investigated whether TMPyP4 binding to G-quadruplexes formed within the viral RNA can inhibit IAV replication.

In our work, we employed several well-established methods to study the interactions of RNA G-quadruplexes with the BRACO-19 and TMPyP4 ligands. First, we applied the reverse transcription (RT) stop assay to study G4/ligand interaction and visualize the G-quadruplex stabilization in response to increasing ligand concentration. We assumed that the reverse transcriptase proceeds along the template RNA sequence until it reaches a G-quadruplex structure stabilized by a G4-specific ligand. Our experiment confirmed this hypothesis, as product synthesis was more strongly inhibited with increasing ligand concentrations. However, as expected, the addition of both ligands to the PQS motifs variants mutated in G-tracts did not cause such noteworthy inhibition of product synthesis. In this experiment, we used an additional ligand, TMPyP2, with lower affinity toward the G4 structure, which served as the control.

Furthermore, based on the resultant gels, we determined IC_{50} values, the half-maximal inhibitory concentrations of our ligands. It is noteworthy that when comparing the ligand inhibition pattern between particular G4 structures, the inhibition levels differed. It may also be confirmed by different IC_{50} values obtained by the same ligand addition to various G-quadruplexes, suggesting different binding affinity of certain ligand to G4s. Additionally, the obtained gels showed variances in the binding of different ligands to the same G4 motifs, as the IC_{50} values appeared to be slightly lower for BRACO-19 than for TMPyP4.

When analyzing ligand binding to a specific RNA G-quadruplex structure, a crucial step is determining the thermodynamic parameters of this interaction. Having that in mind, we applied the isothermal titration calorimetry (ITC) method to quantitatively characterize the interaction between BRACO-19 and TMPyP4 ligands and our three PQS motifs from the IAV genome (1Q, 7Q, and 11Q). The primary objective was to establish ligand binding affinity, stoichiometry, enthalpy, and entropy of these interactions. Additionally, as in the RT stop assay experiment, we used PQS variants mutated in G-tracts to compare thermodynamic parameters before and after G-tract disruption. All binding parameters obtained from the ITC experiment were summarized in Table 2 in the publication. The ITC analysis revealed distinct binding behaviors of both ligands toward the studied G-quadruplexes. Specifically, TMPyP4 bound to wild-type G4s with high affinity through two sets of binding sites, with the strongest affinity to the 11Q sequence. In contrast, BRACO-19 exhibited only one dominant binding mode with lower overall affinity but higher binding stoichiometry. This finding may be explained by its smaller molecular size compared with TMPyP4[183]. Importantly, the G-tract disruption inhibited the G-quadruplex formation, as we observed reduced affinity of both ligands and increased BRACO-19 stoichiometry.

To determine the effect of ligand binding on the thermal stability of RNA G-quadruplexes, UV melting measurements were performed. Analysis of the melting curve profile showed that all PQS motifs formed RNA G4 structures both in the absence and presence of ligands. Ligand addition resulted in only minor changes in melting temperatures, causing only slight destabilization in a few complexes. Overall, all ligands (TMPyP4, TMPyP2, and BRACO-19) did not significantly affect the thermal stability of selected RNA G-quadruplexes under the tested conditions.

CD spectroscopy was utilized to examine whether ligand addition alters the RNA G-quadruplex topology. Nevertheless, CD measurements revealed that all three selected RNAs (1Q, 7Q, and 11Q) adopt a parallel topology, and titrations with ligands (TMPyP4 and BRACO-19) do not influence the overall G4 structure, causing only minor changes observed in the

graphs. This observation indicated that both ligands interact with RNA G-quadruplexes without disrupting their folding topology. Moreover, the second CD experiment was performed with the corresponding RNA sequences from the genome of the second IAV strain, A/Puerto Rico/8/34 (PR8). Because the bioinformatics analysis revealed differences in the nucleotide composition of selected G-tracts from these two IAV strain genomes (Figure 8 in the publication), we aimed to establish whether it influences the adopted G4 structure. CD analysis of RNA motifs from the PR8 strain demonstrated their ability to adopt RNA G4 structures, but with spectral deviations. This result might indicate structural heterogeneity in the RNA G-quadruplexes formed or an equilibrium between various conformations.

To further verify direct interactions between selected RNA G4s and G4-specific ligands, we employed fluorescence spectroscopy. The binding of TMPyP4 to all three RNA motifs (1Q, 7Q, and 11Q) induced changes in the emission spectra, including the peak splitting and redistribution of the fluorescence intensity. In contrast, binding of BRACO-19 to these RNAs resulted mainly in fluorescence quenching. Overall, these findings showed that TMPyP4 exhibits stronger interaction with RNA G4 structures than BRACO-19, as its fluorescence was more significantly altered upon binding to RNA G-quadruplexes.

Moreover, to complement the experimental data, we performed molecular modeling of the 1Q G-quadruplexes. This part of our study aimed to elucidate the folding behavior of this sequence and to identify structural features that can influence ligand binding. The structure obtained from molecular simulation confirmed that the 1Q motif forms a stable G-quadruplex under physiological conditions, while its loop regions remain flexible. Based on this experiment, we concluded that the 1Q motif adopts a bimolecular, four-tetrad G-quadruplex structure, which is highly dynamic. The obtained structural model partially explains our experimental observations and the challenges associated with obtaining high-resolution structural data.

Finally, the IAV minireplicon system was used to investigate whether ligand binding (TMPyP4 and its low-affinity analog, TMPyP2) to G-quadruplex structures affects minireplicon activity. We chose this system in our research because two G4-forming sequences (1Q and 7Q) are located within viral polymerase subunit-coding regions. As the minireplicon system enables testing the activity of the vRNP complex, its decreased activity upon ligand addition may reflect the viral replication inhibition. In this experiment, two minireplicon systems based on different IAV strains, A/California/04/2009 (Cal2009) and A/Puerto Rico/8/34 (PR8), were applied to assess if strain-specific nucleotide differences influence G-quadruplex formation in cells. The analysis revealed that both ligands significantly inhibited

luciferase expression in the Cal2009 minireplicon system, indicating an effective suppression of vRNP activity. On the contrary, PR8 minireplicon showed minimal or no inhibition of the vRNP activity upon ligand addition. We assumed that this result is correlated with the sequence differences in the G-rich motifs of these two IAV strains. Overall, our findings demonstrated that G4-specific ligands can modulate IAV replication in cells. However, their effects strongly depend on the viral genome sequence and the functional motifs formed within viral RNA.

The results described in this study demonstrated that RNA G-quadruplexes are present within the IAV genome and can interact with G4-specific ligands, TMPyP4 and BRACO-19. Our biophysical analyses, including ITC, UV melting, CD, and fluorescence spectroscopy, showed that both ligands bind to RNA G4s with different binding modes. While ligand binding caused only minor changes in thermal stability, it effectively inhibited reverse transcription *in vitro*. Additionally, experiments using two minireplicon systems showed that TMPyP4 suppresses viral polymerase activity in the Cal2009 system but not in the PR8 system. This observation highlights strain-dependent differences, likely arising from variations in the nucleotide composition of G-rich sequences. Overall, our findings indicated that RNA G-quadruplex motifs in IAV may be selectively targeted by chemical ligands, supporting their potential as novel anti-influenza targets.

SUMMARY

Influenza A virus, a human pathogen causing respiratory infections, represents a serious threat to global public health. Its high genomic variability makes the currently available therapies and preventive strategies only partially effective. Therefore, there is a constant need to develop new anti-influenza approaches. In recent years, viral genomes have been confirmed to adopt various structural motifs that have been reported to play significant roles during viral replication cycles. This finding motivated researchers to search viral genomes, including IAV RNA, for structural motifs that could be potential novel antiviral drug targets. Previously, non-canonical secondary structures known as G-quadruplexes were discovered within the genomes of several viruses. These highly conserved motifs play important roles in the viral replication cycle, making them promising targets for antiviral therapies.

In our research, we investigated the influenza A/California/04/2009(H1N1) genome for potential quadruplex-forming sequences (PQS). Bioinformatics analysis identified twelve highly conserved PQSs within the IAV genome, suggesting that G-quadruplex formation in these regions may be important in the influenza life cycle. In the next stage of our research, we investigated the propensity of these sequences to form RNA G-quadruplex structures using various biophysical methods. Several spectroscopic methods, including NMR spectroscopy, UV melting analysis, CD analysis, and fluorescence intensity measurements, were applied. Based on the obtained results, we confirmed that three PQSs adopt stable RNA G4s *in vitro*. Additionally, UV melting analysis enabled the determination of G4 melting temperatures, thereby providing insight into the thermal stability of the investigated G-quadruplexes. CD analysis, in turn, revealed that the proposed G4s adopt a parallel topology. Subsequently, the PAGE experiment was performed to study RNA G-quadruplex formation and to characterize their folding patterns. The resultant gels provided evidence of structural heterogeneity of the formed structures and indicated the presence of an equilibrium between different conformations.

Finally, after confirming RNA G-quadruplex formation using biophysical techniques, three promising candidates were selected for further investigation with G4-specific ligands. In particular, two well-known small-molecule ligands, TMPyP4 and BRACO-19, were used to assess their interactions and binding properties with G4s from the IAV genome. We applied several techniques to study these interactions, i.e., reverse transcription (RT) stop assay, isothermal titration calorimetry, UV melting analysis, CD spectroscopy, and fluorescence intensity measurements. In addition, molecular modeling of one investigated IAV G4 was

performed to assess and visualize the adopted RNA G4 structure. Our biophysical analyses demonstrated that both utilized ligands interact with chosen RNA G4s through distinct binding modes. Therefore, our study provides valuable insights into the interaction between RNA G4 and the selected compounds, contributing to the future development of G4-targeting strategies.

Furthermore, we aimed to investigate how TMPyP4 binding to RNA G4 structures present within the IAV genome affects the viral replication cycle. For this purpose, we conducted a minireplicon-based experiment using two IAV strains, A/California/04/2009 and A/Puerto Rico/8/34. Interestingly, TMPyP4 inhibited IAV polymerase activity only in the A/California/04/2009 system, indicating strain-dependent ligand activity likely driven by differences in the nucleotide composition of G-tracts within PQSs.

In summary, my doctoral research demonstrates that PQSs are present in the influenza A/California/04/2009(H1N1) genome, and some of these PQSs can adopt stable RNA G-quadruplex structures. Moreover, those selected RNA G4s can be targeted by small-molecule G4-specific ligands. Targeting G-quadruplexes within the coding regions of polymerase complex proteins suppresses viral polymerase activity. Taking all the data into account, we concluded that G-quadruplex structures from the IAV RNA may serve as novel potential anti-influenza drug targets.

LIST OF ABBREVIATIONS

ASOs – antisense oligonucleotides

Cal2009 – influenza A/California/04/2009

CD – circular dichroism

cRNA – complementary RNA

DNA – deoxyribonucleic acid

G4 – G-quadruplex

GORS – genome-scale ordered RNA structure

HA – hemagglutinin

HIV – human immunodeficiency virus

IAV – influenza A virus

IRES – internal ribosome entry sites

ITC – isothermal titration calorimetry

MD – molecular dynamics

mRNA – messenger RNA

NA – neuraminidase

NMM – N-methylmesoporphyrin IX

NMR – nuclear magnetic resonance

NP – nucleoprotein

PA – polymerase acid

PAGE – polyacrylamide gel electrophoresis

PB1 – polymerase basic 1

PB2 – polymerase basic 2

PNA – peptide nucleic acid

PQSs – potential quadruplex-forming sequences

PR8 – influenza A/Puerto Rico/8/34

RNA – ribonucleic acid

RT – revered transcription

SARS-CoV-2 – severe acute respiratory syndrome coronavirus 2

SHAPE – Selective 2′ Hydroxyl Acylation analyzed by Primer Extension

siRNAs – small interfering RNAs

ThT – thioflavin T

TMPyP2 – 5,10,15,20-tetra-(N-methyl-2-pyridyl)porphyrin

TMPyP4 – tetra-(N-methyl-4-pyridyl)porphin

UTR – untranslated region

vRNA – viral ribonucleic acid

vRNA5 – segment 5 vRNA

vRNPs – viral ribonucleoprotein complexes

REFERENCES

1. Le, S. Y., Zhang, K. & Maizel, J. V. RNA molecules with structure dependent functions are uniquely folded. *Nucleic Acids Res.* **30**, 3574–3582 (2002).

2. Ganser, L. R., Kelly, M. L., Herschlag, D. & Al-Hashimi, H. M. The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.* **20**, 474–489 (2019).

3. Jaafar, Z. A. & Kieft, J. S. Viral RNA structure-based strategies to manipulate translation. *Nat. Rev. Microbiol.* **17**, 110–123 (2019).

4. Mortimer, S. A., Kidwell, M. A. & Doudna, J. A. Insights into RNA structure and function from genome-wide studies. *Nat. Rev. Genet.* **15**, 469–479 (2014).

5. Sanchez de Groot, N. *et al.* RNA structure drives interaction with proteins. *Nat. Commun.* **10**, 1–13 (2019).

6. Bose, R., Saleem, I. & Mustoe, A. M. Causes, functions, and therapeutic possibilities of RNA secondary structure ensembles and alternative states. *Cell Chem. Biol.* **31**, 17–35 (2024).

7. 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).

8. Moore, P. B. Structural motifs in RNA. *Annu. Rev. Biochem.* **68**, 287–300 (1999).

9. Jacobs, E., Mills, J. D. & Janitz, M. The role of RNA structure in posttranscriptional regulation of gene expression. *J. Genet. Genomics* **39**, 535–543 (2012).

10. Hershey, J. W. B., Sonenberg, N. & Mathews, M. B. Principles of translational control: An overview. *Cold Spring Harb. Perspect. Biol.* **4**, 1–10 (2012).

11. Georgakopoulos-Soares, I., Parada, G. E. & Hemberg, M. Secondary structures in RNA synthesis, splicing and translation. *Comput. Struct. Biotechnol. J.* **20**, 2871–2884 (2022).

12. Ferhadian, D. *et al.* Structural and functional motifs in influenza virus RNAs. *Front. Microbiol.* **9**, 1–11 (2018).

13. Szczesniak, I., Baliga-Gil, A., Jarmolowicz, A., Soszynska-Jozwiak, M. & Kierzek, E. Structural and functional RNA motifs of SARS-CoV-2 and influenza A virus as a target of viral inhibitors. *Int. J. Mol. Sci.* **24**, (2023).

14. Short, K. R., Kedzierska, K. & van de Sandt, C. E. Back to the future: Lessons learned from the 1918 influenza pandemic. *Front. Cell. Infect. Microbiol.* **8**, 1–19 (2018).

15. Boerneke, M. A., Ehrhardt, J. E. & Weeks, K. M. Physical and functional analysis of viral RNA genomes by SHAPE. *Annu. Rev. Virol.* **6**, 93–117 (2019).

16. Gupta, A. & Bansal, M. RNA-mediated translation regulation in viral genomes: Computational advances in the recognition of sequences and structures. *Brief. Bioinform.* **21**, 1151–1163 (2019).

17. Ye, L. *et al.* RNA structures and their role in selective genome packaging. *Viruses* **13**, (2021).

18. Cross, S. T., Michalski, D., Miller, M. R. & Wilusz, J. RNA regulatory processes in RNA virus biology. *Wiley Interdiscip. Rev. RNA* **10**, 1–24 (2019).

19. Fodor, E. & Velthuis, A. J. W. T. Structure and function of the influenza virus transcription and replication machinery. *Cold Spring Harb. Perspect. Med.* **10**, 1–14 (2020).

20. 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**, 1–16 (2019).

21. Nakano, M. *et al.* Ultrastructure of influenza virus ribonucleoprotein complexes during viral RNA synthesis. *Commun. Biol.* **4**, 1–10 (2021).

22. 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).

23. Michalak, P. *et al.* Conserved structural motifs of two distant IAV subtypes in genomic segment 5 RNA. *Viruses* **13**, 1–14 (2021).

24. Szutkowska, B. *et al.* Secondary structure of influenza A virus genomic segment 8 RNA folded in a cellular environment. *Int. J. Mol. Sci.* **23**, (2022).

25. Mirska, B. *et al.* In vivo secondary structural analysis of influenza A virus genomic RNA. *Cell. Mol. Life Sci.* **80**, 1–24 (2023).

26. 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).

27. 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**, 627–642 (2020).

28. Takizawa, N. & Kawaguchi, R. K. Comprehensive in virio structure probing analysis of the influenza A virus identifies functional RNA structures involved in viral genome replication. *Comput. Struct. Biotechnol. J.* **21**, 5259–5272 (2023).

29. Fremin, B. J., Bhatt, A. S. & Kyrpides, N. C. Identification of over ten thousand candidate structured RNAs in viruses and phages. *Comput. Struct. Biotechnol. J.* **21**, 5630–5639 (2023).

30. Simmonds, P. Pervasive RNA secondary structure in the genomes of SARS-CoV-2 and other coronaviruses. *MBio* **11**, 1–15 (2020).

31. Priore, S. F., Moss, W. N. & Turner, D. H. Influenza B virus has global ordered RNA structure in (+) and (-) strands but relatively less stable predicted RNA folding free energy than allowed by the encoded protein sequence. *BMC Res. Notes* **6**, (2013).

32. Davis, M., Sagan, S. M., Pezacki, J. P., Evans, D. J. & Simmonds, P. Bioinformatic and physical characterizations of genome-scale ordered RNA structure in mammalian RNA viruses. *J. Virol.* **82**, 11824–11836 (2008).

33. Li, P. *et al.* Integrative analysis of Zika virus genome RNA structure reveals critical determinants of viral infectivity. *Cell Host Microbe* **24**, 875-886.e5 (2018).

34. Alvarez, D. E., Lodeiro, M. F., Ludueña, S. J., Pietrasanta, L. I. & Gamarnik, A. V. Long-range RNA-RNA interactions circularize the dengue virus genome. *J. Virol.* **79**, 6631–6643 (2005).

35. Yang, S. L. *et al.* A conserved long-range RNA interaction in SARS-CoV-2 recruits ADAR1 to enhance virus proliferation. *Nat. Commun.* **16**, 1–15 (2025).

36. 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).

37. Lenartowicz, E. *et al.* Antisense oligonucleotides targeting influenza A segment 8 genomic RNA inhibit viral replication. *Nucleic Acid Ther.* **26**, 277–285 (2016).

38. Czapik, T., Piasecka, J., Kierzek, R. & Kierzek, E. Structural variants and modifications of hammerhead ribozymes targeting influenza A virus conserved structural motifs. *Mol. Ther. - Nucleic Acids* **29**, 64–74 (2022).

39. Lee, M. K. *et al.* A novel small-molecule binds to the influenza A virus RNA promoter and inhibits viral replication. *Chem. Commun.* **50**, 368–370 (2014).

40. Piekarowicz, A. *Podstawy Wirusologii Molekularnej*. (PWN, 2004).

41. Krammer, F. & Palese, P. Orthomyxoviridae: The Viruses and Their Replication. in *Fields Virology: Emerging Viruses* 596–649 (Wolters Kluwer, 2021).

42. Centers for Disease Control and Prevention. Types of Influenza Viruses. *Centers for Disease Control and Prevention* https://www.cdc.gov/flu/about/viruses-types.html (2025).

43. Carter, T. & Iqbal, M. The influenza A virus replication cycle: A comprehensive review. *Viruses* **16**, 1–53 (2024).

44. Eisfeld, A. J., Neumann, G. & Kawaoka, Y. Influenza A virus isolation, culture and identification. *Nat. Protoc.* **9**, 2663–2681 (2014).

45. World Health Organization. Influenza (seasonal). *World Health Organization* https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal) (2025).

46. Paules, C. I. & Fauci, A. S. Influenza vaccines: Good, but we can do better. *J. Infect. Dis.* **219**, S1–S4 (2019).

47. Mostafa, A., Abdelwhab, E. M., Mettenleiter, T. C. & Pleschka, S. Zoonotic potential of influenza A viruses: A comprehensive overview. *Viruses* **10**, 1–38 (2018).

48. Nayak, D. P., Balogun, R. A., Yamada, H., Zhou, Z. H. & Barman, S. Influenza virus morphogenesis and budding. *Virus Res.* **143**, 147–161 (2009).

49. Kesy, J. *et al.* A short chemically modified dsRNA-binding PNA (dbPNA) inhibits influenza viral replication by targeting viral RNA panhandle structure. *Bioconjug. Chem.* **30**, 931–943 (2019).

50. Lenartowicz, E. *et al.* Self-folding of naked segment 8 genomic RNA of influenza A virus. *PLoS One* **11**, 1–21 (2016).

51. 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).

52. Brázda, V. *et al.* G-quadruplexes in H1N1 influenza genomes. *BMC Genomics* **22**, 1–11 (2021).

53. Fleming, A. M., Ding, Y., Alenko, A. & Burrows, C. J. Zika virus genomic RNA possesses conserved G-quadruplexes characteristic of the Flaviviridae family. *ACS Infect. Dis.* **2**, 674–681 (2016).

54. Wang, S. R. *et al.* A highly conserved G-rich consensus sequence in hepatitis C virus core gene represents a new anti–hepatitis C target. *Sci. Adv.* **2**, (2016).

55. Bezzi, G., Piga, E. J., Binolfi, A. & Armas, P. CNBP binds and unfolds in vitro G-quadruplexes formed in the SARS-CoV-2 positive and negative genome strands. *Int. J. Mol. Sci.* **22**, 1–22 (2021).

56. Ruggiero, E., Zanin, I., Terreri, M. & Richter, S. N. G-quadruplex targeting in the fight against viruses: An update. *Int. J. Mol. Sci.* **22**, (2021).

57. Wang, S. R. *et al.* Chemical targeting of a G-quadruplex RNA in the Ebola virus L gene. *Cell Chem. Biol.* **23**, 1113–1122 (2016).

58. Lavezzo, E. *et al.* G-quadruplex forming sequences in the genome of all known human viruses: A comprehensive guide. *PLoS Comput. Biol.* **14**, 1–20 (2018).

59. Li, Z. *et al.* G-quadruplexes in genomes of viruses infecting eukaryotes or prokaryotes are under different selection pressures from hosts. *J. Genet. Genomics* **49**, 20–29 (2022).

60. Dadonaite, B. *et al.* The structure of the influenza A virus genome. *Nat. Microbiol.* **4**, 1781–1789 (2019).

61. Yang, R. *et al.* Mapping of the influenza A virus genome RNA structure and interactions reveals essential elements of viral replication. *Cell Rep.* **43**, 113833 (2024).

62. Dela-Moss, L. I., Moss, W. N. & Turner, D. H. Identification of conserved RNA secondary structures at influenza B and C splice sites reveals similarities and differences between influenza A, B, and C. *BMC Res. Notes* **7**, (2014).

63. 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).

64. Takemoto, Y. Therapeutic approach for seasonal influenza and pandemic. in *Influenza - Therapeutics and Challenges* 133–150 (2018).

65. Scorza, F. B. & Pardi, N. New kids on the block: RNA-based influenza virus vaccines. *Vaccines* **6**, 1–15 (2018).

66. Hagey, R. J. *et al.* Programmable antivirals targeting critical conserved viral RNA secondary structures from influenza A virus and SARS-CoV-2. *Nat. Med.* **28**, 1944–1955 (2022).

67. Ambike, S. *et al.* Targeting genomic SARS-CoV-2 RNA with siRNAs allows efficient inhibition of viral replication and spread. *Nucleic Acids Res.* **50**, 333–349 (2022).

68. Soszynska-Jozwiak, M. *et al.* Influenza virus segment 5 (+)RNA - Secondary structure and new targets for antiviral strategies. *Sci. Rep.* **7**, 1–15 (2017).

69. Trobaugh, D. W. & Klimstra, W. B. MicroRNA regulation of RNA virus replication and pathogenesis. *Trends Mol. Med.* **23**, 80–93 (2017).

70. Rhodes, D. & Lipps, H. J. Survey and summary G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* **43**, 8627–8637 (2015).

71. Małgowska, M., Gudanis, D., Teubert, A., Dominiak, G. & Gdaniec, Z. How to study G-quadruplex structures. *Biotechnologia* **93**, 381–390 (2012).

72. Abiri, A. *et al.* Unlocking G-quadruplexes as antiviral targets. *Pharmacol. Rev.* **73**, 897–923 (2021).

73. Largy, E., Mergny, J.-L. & Gabelica, V. Role of alkali metal ions in G-quadruplex nucleic acid structure and stability. in *The Alkali Metal Ions: Their Role for Life* vol. 16 204–231 (2016).

74. Xiao, C. Da, Shibata, T., Yamamoto, Y. & Xu, Y. An intramolecular antiparallel G-quadruplex formed by human telomere RNA. *Chem. Commun.* **54**, 3944–3946 (2018).

75. Lyu, K., Chow, E. Y. C., Mou, X., Chan, T. F. & Kwok, C. K. RNA G-quadruplexes (rG4s): Genomics and biological functions. *Nucleic Acids Res.* **49**, 5426–5450 (2021).

76. Hashimoto, Y., Shil, S., Tsuruta, M., Kawauchi, K. & Miyoshi, D. Three- and four-stranded nucleic acid structures and their ligands. *RSC Chem. Biol.* **6**, 466–491 (2025).

77. Cucchiarini, A., Kledus, F., Luo, Y., Brázda, V. & Mergny, J. L. Quadruplexes with a grain of salt: Influence of cation type and concentration on DNA G4 stability. *Eur. Biophys. J.* **54**, 589–599 (2025).

78. Kim, S., Kim, D. & Kim, Y. Metal–stabilized G–quadruplexes: Biological insights and sensing applications. *BMB Rep.* **58**, 397–405 (2025).

79. Bhattacharyya, D., Arachchilage, G. M. & Basu, S. Metal cations in G-quadruplex folding and stability. *Front. Chem.* **4**, 1–14 (2016).

80. Olejko, L., Dutta, A., Shahsavar, K. & Bald, I. Influence of different salts on the G-quadruplex structure formed from the reversed human telomeric DNA sequence. *Int. J. Mol. Sci.* **23**, (2022).

81. Nishio, M., Tsukakoshi, K. & Ikebukuro, K. G-quadruplex: Flexible conformational changes by cations, pH, crowding and its applications to biosensing. *Biosens. Bioelectron.* **178**, 113030 (2021).

82. Galer, P., Wang, B., Plavec, J. & Šket, P. Unveiling the structural mechanism of a G-quadruplex pH–driven switch. *Biochimie* **214**, 73–82 (2023).

83. Navarro, A., Benabou, S., Eritja, R. & Gargallo, R. Influence of pH and a porphyrin ligand on the stability of a G-quadruplex structure within a duplex segment near the promoter region of the SMARCA4 gene. *Int. J. Biol. Macromol.* **159**, 383–393 (2020).

84. Tuntiwechapikul, W., Taka, T., Béthencourt, M., Makonkawkeyoon, L. & Lee, T. R. The influence of pH on the G-quadruplex binding selectivity of perylene derivatives. *Bioorganic Med. Chem. Lett.* **16**, 4120–4126 (2006).

85. Li, J. *et al.* Effects of length and loop composition on structural diversity and similarity of (G3TG3NMG3TG3) G-quadruplexes. *Molecules* **25**, (2020).

86. Rachwal, P. A., Brown, T. & Fox, K. R. Sequence effects of single base loops in intramolecular quadruplex DNA. *FEBS Lett.* **581**, 1657–1660 (2007).

87. Cheng, M. *et al.* Loop permutation affects the topology and stability of G-quadruplexes. *Nucleic Acids Res.* **46**, 9264–9275 (2018).

88. Gellert, M., Lipsett, M. N. & Davies, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U. S. A.* **48**, 2013–2018 (1962).

89. Sundquist, W. I. & Klug, A. Telomeric DNA dimerizes by formation of guanine tetrads between

hairpin loops. *Nature* **342**, 825–829 (1989).

90. Sun, D. *et al.* Inhibition of human telomerase by a G-quadruplex-interactive compound. *J. Med. Chem.* **40**, 2113–2116 (1997).

91. Read, M. *et al.* Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 4844–4849 (2001).

92. Neidle, S. Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS J.* **277**, 1118–1125 (2010).

93. Chen, L., Dickerhoff, J., Sakai, S. & Yang, D. DNA G-quadruplex in human telomeres and oncogene promoters: Structures, functions, and small molecule targeting. *Acc. Chem. Res.* **55**, 2628–2646 (2022).

94. Varshney, D., Spiegel, J., Zyner, K., Tannahill, D. & Balasubramanian, S. The regulation and functions of DNA and RNA G-quadruplexes. *Nat. Rev. Mol. Cell Biol.* **21**, 459–474 (2020).

95. Endoh, T. & Sugimoto, N. Conformational dynamics of the RNA G-quadruplex and its effect on translation efficiency. *Molecules* **24**, 1–14 (2019).

96. Cammas, A. & Millevoi, S. RNA G-quadruplexes: Emerging mechanisms in disease. *Nucleic Acids Res.* **45**, 1584–1595 (2017).

97. Ruggiero, E. & Richter, S. N. G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Res.* **46**, 3270–3283 (2018).

98. Cui, H. & Zhang, L. G-quadruplexes are present in human coronaviruses including SARS-CoV-2. *Front. Microbiol.* **11**, 1–8 (2020).

99. Sanjuán, R. & Domingo-Calap, P. Mechanisms of viral mutation. *Cell. Mol. Life Sci.* **73**, 4433–4448 (2016).

100. Grenfell, B. T. *et al.* Unifying the epidemiological and evolutionary dynamics of pathogens. *Science (80-.).* **303**, 327–332 (2004).

101. Goździcka-Józefiak, A. *Wirusologia*. (PWN, 2019).

102. Rouleau, S., Jodoin, R., Garant, J.-M. & Perreault, J.-P. RNA G-quadruplexes as key motifs of the transcriptome. in *Catalytically Active Nucleic Acids. Advances in Biochemical Engineering/Biotechnology* vol. 170 1–20 (2017).

103. Kosiol, N., Juranek, S., Brossart, P., Heine, A. & Paeschke, K. G-quadruplexes: A promising target for cancer therapy. *Mol. Cancer* **20**, 1–18 (2021).

104. Berner, A. *et al.* G4-Ligand-conjugated oligonucleotides mediate selective binding and stabilization of individual G4 DNA structures. *J. Am. Chem. Soc.* **146**, 6926–6935 (2024).

105. Zhai, L. Y. *et al.* Targeting the RNA G-quadruplex and protein interactome for antiviral therapy. *J. Med. Chem.* **65**, 10161–10182 (2022).

106. Pathak, R. G-Quadruplexes in the viral genome: Unlocking targets for therapeutic interventions and antiviral strategies. *Viruses* **15**, (2023).

107. Fracchioni, G., Vailati, S., Grazioli, M. & Pirota, V. Structural unfolding of G-quadruplexes: From small molecules to antisense strategies. *Molecules* **29**, (2024).

108. Wu, T. Y., Lau, H. L., Santoso, R. J. & Kwok, C. K. RNA G-quadruplex structure targeting and imaging: Recent advances and future directions. *Rna* **31**, 1053–1080 (2025).

109. Rouleau, S. G., Beaudoin, J. D., Bisaillon, M. & Perreault, J. P. Small antisense oligonucleotides against G-quadruplexes: Specific mRNA translational switches. *Nucleic Acids Res.* **43**, 595–606 (2015).

110. Sharma, V. K., Sharma, R. K. & Singh, S. K. Antisense oligonucleotides: Modifications and clinical trials. *Medchemcomm* **5**, 1454–1471 (2014).

111. Chen, Z. *et al.* Synthesis and evaluation of antisense oligonucleotides prodrug with G-quadruplex assembly and lysosome escape capabilities for oncotherapy. *Bioorg. Chem.* **148**, 107475 (2024).

112. Muangkaew, P. & Vilaivan, T. Modulation of DNA and RNA by PNA. *Bioorganic Med. Chem. Lett.* **30**, (2020).

113. Lusvarghi, S. *et al.* Loop and backbone modifications of PNA improve G-quadruplex binding selectivity. *J. Am. Chem. Soc.* **131**, 18415 (2009).

114. Krpetic, Z., Miah, A. & Arrigo, R. Targeting mRNA G-quadruplexes with peptide nucleic acid conjugates to regulate gene expression in vivo. **2025**, (2025).

115. Chan, C. & Kwok, C. K. Specific binding of a D-RNA G-quadruplex structure with an L-RNA aptamer. *Angew. Chemie* **132**, 5331–5335 (2020).

116. Esposito, D., Locatelli, A. & Morigi, R. Molecular tools for precision targeting and detection of G-quadruplex structures. *Molecules* **30**, (2025).

117. Chen, S. Bin *et al.* Visualization of NRAS RNA G-quadruplex structures in cells with an engineered fluorogenic hybridization probe. *J. Am. Chem. Soc.* **138**, 10382–10385 (2016).

118. Tseng, T. Y., Chu, I. Te, Lin, S. J., Li, J. & Chang, T. C. Binding of small molecules to G-quadruplex DNA in cells revealed by fluorescence lifetime imaging microscopy of o-BMVC foci. *Molecules* **24**, 1–13 (2019).

119. Tomaszewska, M., Szabat, M., Zielińska, K. & Kierzek, R. Identification and structural aspects of G-quadruplex-forming sequences from the influenza A virus genome. *Int. J. Mol. Sci.* **22**, 1–19 (2021).

120. Lemkul, J. A. Same fold, different properties: Polarizable molecular dynamics simulations of telomeric and TERRA G-quadruplexes. *Nucleic Acids Res.* **48**, 561–575 (2020).

121. Ortiz De Luzuriaga, I., Lopez, X. & Gil, A. Learning to model G-quadruplexes: Current methods and perspectives. *Annu. Rev. Biophys.* **50**, 209–243 (2021).

122. Li, M. H., Zhou, Y. H., Luo, Q. & Li, Z. S. The 3D structures of G-quadruplexes of HIV-1 integrase inhibitors: Molecular dynamics simulations in aqueous solution and in the gas phase. *J. Mol. Model.* **16**, 645–657 (2010).

123. Tsvetkov, V. B. Modeling possible G-quadruplexes and i-motifs at DNA–DNA contact sites: Strategy, classification, and examples. *Int. J. Mol. Sci.* **26**, 1–30 (2025).

124. Fogolari, F. *et al.* Molecular models for intrastrand DNA G-quadruplexes. *BMC Struct. Biol.* **9**, 64 (2009).

125. Bryan, T. M. & Baumann, P. G-quadruplexes: From guanine gels to chemotherapeutics. *Mol.*

Biotechnol. **49**, 198–208 (2011).

126. Santos, T., Salgado, G. F., Cabrita, E. J. & Cruz, C. G-quadruplexes and their ligands: Biophysical methods to unravel G-quadruplex/ligand interactions. *Pharmaceuticals* **14**, (2021).

127. You, J. *et al.* Effects of monovalent cations on folding kinetics of G-quadruplexes. *Biosci. Rep.* **37**, 1–12 (2017).

128. da Silva, M. W. NMR methods for studying quadruplex nucleic acids. *Methods* **43**, 264–277 (2007).

129. Mergny, J. L., Phan, A. T. & Lacroix, L. Following G-quartet formation by UV-spectroscopy. *FEBS Lett.* **435**, 74–78 (1998).

130. Mergny, J. L. & Lacroix, L. Analysis of thermal melting curves. *Oligonucleotides* **13**, 515–537 (2003).

131. Giunta, D., Roviello, G. N. & Solinas, M. Design and synthesis of new benzoxazole derivatives and study of their interactions with a G4 model by UV spectroscopy and in silico methods. *J. Mol. Liq.* **410**, 125491 (2024).

132. Nowak-Karnowska, J., Głuszyńska, A., Kosman, J. & Dembska, A. Fluorescence turn-off ligand for parallel G-quadruplexes. *Molecules* **29**, (2024).

133. Głuszyńska, A., Juskowiak, B. & Rubiś, B. Binding study of the fluorescent carbazole derivative with human telomeric G-quadruplexes. *Molecules* **23**, 1–22 (2018).

134. Zhang, Y., Chen, J., Ju, H. & Zhou, J. Thermal denaturation profile: A straightforward signature to characterize parallel G-quadruplexes. *Biochimie* **157**, 22–25 (2019).

135. Tluckova, K. *et al.* Human papillomavirus G-quadruplexes. *Biochemistry* **52**, 7207–7216 (2013).

136. Bai, L. P., Hagihara, M., Nakatani, K. & Jiang, Z. H. Recognition of chelerythrine to human telomeric DNA and RNA G-quadruplexes. *Sci. Rep.* **4**, 15–17 (2014).

137. Chaires, J. B. & Graves, D. *Quadruplex Nucleic Acids*. (Springer, 2013).

138. Mohanty, J. *et al.* Thioflavin T as an efficient inducer and selective fluorescent sensor for the human telomeric G-quadruplex DNA. *J. Am. Chem. Soc.* **135**, 367–376 (2013).

139. Schonhoft, J. D. *et al.* Direct experimental evidence for quadruplex-quadruplex interaction within the human ILPR. *Nucleic Acids Res.* **37**, 3310–3320 (2009).

140. Piekna-Przybylska, D., Sullivan, M. A., Sharma, G. & Bambara, R. A. U3 region in the HIV-1 genome adopts a G-quadruplex structure in its RNA and DNA sequence. *Biochemistry* **53**, 2581–2593 (2014).

141. Lubitz, I., Zikich, D. & Kotlyar, A. Specific high-affinity binding of thiazole orange to triplex and G-quadruplex DNA. *Biochemistry* **49**, 3567–3574 (2010).

142. Kharel, P., Becker, G., Tsvetkov, V. & Ivanov, P. Properties and biological impact of RNA G-quadruplexes: From order to turmoil and back. *Nucleic Acids Res.* **48**, 12534–12555 (2020).

143. Hagihara, M., Yoneda, K., Yabuuchi, H., Okuno, Y. & Nakatani, K. A reverse transcriptase stop assay revealed diverse quadruplex formations in UTRs in mRNA. *Bioorganic Med. Chem. Lett.* **20**, 2350–2353 (2010).

144. Perrone, R. *et al.* Anti-HIV-1 activity of the G-quadruplex ligand BRACO-19. *J. Antimicrob. Chemother.* **69**, 3248–3258 (2014).

145. Timmer, C. M. *et al.* An isothermal titration and differential scanning calorimetry study of the G-quadruplex DNA-insulin interaction. *J. Phys. Chem. B* **118**, 1784–1790 (2014).

146. Pagano, B., Mattia, C. A. & Giancola, C. Applications of isothermal titration calorimetry in biophysical studies of G-quadruplexes. *Int. J. Mol. Sci.* **10**, 2935–2957 (2009).

147. Mendoza, O., Bourdoncle, A., Boulé, J. B., Brosh, R. M. & Mergny, J. L. G-quadruplexes and helicases. *Nucleic Acids Res.* **44**, 1989–2006 (2016).

148. Spiegel, J., Adhikari, S. & Balasubramanian, S. The structure and function of DNA G-quadruplexes. *Trends Chem.* **2**, 123–136 (2020).

149. Awadasseid, A., Ma, X., Wu, Y. & Zhang, W. G-quadruplex stabilization via small-molecules as a potential anti-cancer strategy. *Biomed. Pharmacother.* **139**, 111550 (2021).

150. Majee, P. *et al.* Inhibition of Zika virus replication by G-quadruplex-binding ligands. *Mol. Ther. - Nucleic Acids* **23**, 691–701 (2021).

151. Majee, P. *et al.* Identification and characterization of two conserved G-quadruplex forming motifs in the Nipah virus genome and their interaction with G-quadruplex specific ligands. *Sci. Rep.* **10**, 1–12 (2020).

152. Molnár, O. R., Végh, A., Somkuti, J. & Smeller, L. Characterization of a G-quadruplex from hepatitis B virus and its stabilization by binding TMPyP4, BRACO19 and PhenDC3. *Sci. Rep.* **11**, 1–11 (2021).

153. Haider, S. M., Neidle, S. & Parkinson, G. N. A structural analysis of G-quadruplex/ligand interactions. *Biochimie* **93**, 1239–1251 (2011).

154. Chilka, P., Desai, N. & Datta, B. Small molecule fluorescent probes for G-quadruplex visualization as potential cancer theranostic agents. *Molecules* **24**, (2019).

155. Vummidi, B. R., Alzeer, J. & Luedtke, N. W. Fluorescent probes for G-quadruplex structures. *ChemBioChem* **14**, 540–558 (2013).

156. Sun, Z. Y., Wang, X. N., Cheng, S. Q., Su, X. X. & Ou, T. M. Developing novel G-quadruplex ligands: From interaction with nucleic acids to interfering with nucleic acid–protein interaction. *Molecules* **24**, 1–29 (2019).

157. Artusi, S. *et al.* The Herpes Simplex Virus-1 genome contains multiple clusters of repeated G-quadruplex: Implications for the antiviral activity of a G-quadruplex ligand. *Antiviral Res.* **118**, 123–131 (2015).

158. Lv, L. & Zhang, L. Characterization of G-quadruplexes in enterovirus A71 genome and their interaction with G-quadruplex ligands. *Microbiol. Spectr.* **10**, 1–17 (2022).

159. Martino, L., Pagano, B., Fotticchia, I., Neidle, S. & Giancola, C. Shedding light on the interaction between TMPyP4 and human telomeric quadruplexes. *J. Phys. Chem. B* **113**, 14779–14786 (2009).

160. Artusi, S. *et al.* Antiviral activity of the G-quadruplex ligand TMPyP4 against herpes simplex virus-1. *Viruses* **13**, (2021).

161. Cervenak, M., Molnár, O. R., Horváth, P. & Smeller, L. Stabilization of G-quadruplex structures of the SARS-CoV-2 genome by TMPyP4, BRACO19, and PhenDC3. *Int. J. Mol. Sci.* **25**, (2024).

162. Qin, G. *et al.* RNA G-quadruplex formed in SARS-CoV-2 used for COVID-19 treatment in animal models. *Cell Discov.* **8**, (2022).

163. Le, V. H., Nagesh, N. & Lewis, E. A. Bcl-2 promoter sequence G-quadruplex interactions with three planar and non-planar cationic porphyrins: TMPyP4, TMPyP3, and TMPyP2. *PLoS One* **8**, 1–9 (2013).

164. Lyons, S. M., Gudanis, D., Coyne, S. M., Gdaniec, Z. & Ivanov, P. Identification of functional tetramolecular RNA G-quadruplexes derived from transfer RNAs. *Nat. Commun.* **8**, (2017).

165. Yett, A., Lin, L. Y., Beseiso, D., Miao, J. & Yatsunyk, L. A. N-methyl mesoporphyrin IX as a highly selective light-up probe for G-quadruplex DNA. *J. Porphyr. Phthalocyanines* **23**, 1195–1215 (2019).

166. Basu, P. *et al.* RNA G-quadruplex formation in biologically important transcribed regions: Can two-tetrad intramolecular RNA quadruplexes be formed? *Nucleic Acids Res.* **52**, 13224–13242 (2024).

167. Kreig, A. *et al.* G-quadruplex formation in double strand DNA probed by NMM and CV fluorescence. *Nucleic Acids Res.* **43**, 7961–7970 (2015).

168. Li, Y., Liu, W., Zhu, Y. & Diao, L. A fluorescence method for homogeneous detection of influenza A DNA sequence based on guanine-quadruplex-N-methylmesoporphyrin IX complex and assistance-DNA inhibition. *J. Med. Virol.* **91**, 979–985 (2019).

169. Xu, S. *et al.* Thioflavin T as an efficient fluorescence sensor for selective recognition of RNA G-quadruplexes. *Sci. Rep.* **6**, 1–9 (2016).

170. Liu, S., Peng, P., Wang, H., Shi, L. & Li, T. Thioflavin T binds dimeric parallel-stranded GA-containing non-G-quadruplex DNAs: A general approach to lighting up double-stranded scaffolds. *Nucleic Acids Res.* **45**, 12080–12089 (2017).

171. Hanczyc, P., Rajchel-Mieldzioć, P., Feng, B. & Fita, P. Identification of Thioflavin T Binding Modes to DNA: A Structure-Specific Molecular Probe for Lasing Applications. *J. Phys. Chem. Lett.* **12**, 5436–5442 (2021).

172. Lages, A., Proud, C. G., Holloway, J. W. & Vorechovsky, I. Thioflavin T monitoring of guanine quadruplex formation in the rs689-dependent INS intron 1. *Mol. Ther. - Nucleic Acids* **16**, 770–777 (2019).

173. Liu, L., Shao, Y., Peng, J., Liu, H. & Zhang, L. Selective recognition of ds-DNA cavities by a molecular rotor: Switched fluorescence of thioflavin T. *Mol. Biosyst.* **9**, 2512–2519 (2013).

174. Xiao, Y. *et al.* Deep sequencing of 2009 influenza A/H1N1 virus isolated from volunteer human challenge study participants and natural infections. *Virology* **534**, 96–107 (2019).

175. Selvaraj, C. *et al.* Structure-based virtual screening and molecular dynamics simulation of SARS-CoV-2 Guanine-N7 methyltransferase (nsp14) for identifying antiviral inhibitors against COVID-19. *J. Biomol. Struct. Dyn.* **39**, 4582–4593 (2021).

176. Madden, E. A. *et al.* Using SHAPE-MaP to model RNA secondary structure and identify 3′UTR variation in Chikungunya virus. *J. Virol.* **94**, 1–22 (2020).

177. Tomezsko, P. J. *et al.* Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature* **582**, 438–442 (2020).

178. Wang, X. W., Liu, C. X., Chen, L. L. & Zhang, Q. C. RNA structure probing uncovers RNA structure-dependent biological functions. *Nat. Chem. Biol.* **17**, 755–766 (2021).

179. Yang, S. L. *et al.* Comprehensive mapping of SARS-CoV-2 interactions in vivo reveals functional virus-host interactions. *Nat. Commun.* **12**, (2021).

180. Le Sage, V. *et al.* Mapping of influenza virus RNA-RNA interactions reveals a flexible network. *Cell Rep.* **31**, 1–8 (2020).

181. Xu, J., Huang, H. & Zhou, X. G-quadruplexes in neurobiology and virology: Functional roles and potential therapeutic approaches. *JACS Au* **1**, 2146–2161 (2021).

182. Holoubek, J. *et al.* Guanine quadruplexes in the RNA genome of the tick-borne encephalitis virus: Their role as a new antiviral target and in virus biology. *Nucleic Acids Res.* **50**, 4574–4600 (2022).

183. Sullivan, H. J. *et al.* Binding of telomestatin, TMPyP4, BSU6037, and BRACO19 to a telomeric G-quadruplex-duplex hybrid probed by all-atom molecular dynamics simulations with explicit solvent. *ACS Omega* **3**, 14788–14806 (2018).

APPENDIX 1

Publications included in the doctoral dissertation

International Journal of
Molecular Sciences

MDPI

Article

Identification and Structural Aspects of G-Quadruplex-Forming Sequences from the Influenza A Virus Genome

Maria Tomaszewska [1,†], Marta Szabat [1,*,†], Karolina Zielińska [2] and Ryszard Kierzek [1,*]

[1] Department of Structural Chemistry and Biology of Nucleic Acids, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland; mtomaszewska@ibch.poznan.pl
[2] Department of Biomolecular NMR, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland; kczajczynska@ibch.poznan.pl
* Correspondence: szabat@ibch.poznan.pl (M.S.); rkierzek@ibch.poznan.pl (R.K.); Tel.: +48-61-852-8503 (M.S. & R.K.)
† These authors contributed equally to this work as co-first authors.

Abstract: Influenza A virus (IAV) causes seasonal epidemics and sporadic pandemics, therefore is an important research subject for scientists around the world. Despite the high variability of its genome, the structure of viral RNA (vRNA) possesses features that remain constant between strains and are biologically important for virus replication. Therefore, conserved structural motifs of vRNA can represent a novel therapeutic target. Here, we focused on the presence of G-rich sequences within the influenza A/California/07/2009(H1N1) genome and their ability to form RNA G-quadruplex structures (G4s). We identified 12 potential quadruplex-forming sequences (PQS) and determined their conservation among the IAV strains using bioinformatics tools. Then we examined the propensity of PQS to fold into G4s by various biophysical methods. Our results revealed that six PQS oligomers could form RNA G-quadruplexes. However, three of them were confirmed to adopt G4 structures by all utilized methods. Moreover, we showed that these PQS motifs are present within segments encoding polymerase complex proteins indicating their possible role in the virus biology.

Keywords: Influenza A virus; potential quadruplex-forming sequence; RNA G-quadruplexes

Citation: Tomaszewska, M.; Szabat, M.; Zielińska, K.; Kierzek, R. Identification and Structural Aspects of G-Quadruplex-Forming Sequences from the Influenza A Virus Genome. *Int. J. Mol. Sci.* **2021**, *22*, 6031. https://doi.org/10.3390/ijms22116031

Academic Editor: Nitin Saksena

Received: 22 April 2021
Accepted: 31 May 2021
Published: 2 June 2021

1. Introduction

Influenza virus, due to its pandemic potential, became an interesting subject in various biological research projects. There are three types of influenza virus: A, B and C. However, the influenza A virus (IAV) is gaining a lot of attention because of its responsibility for most of the pandemic outbreaks [1], the ability to infect fowl, humans and other mammals, and the highest zoonotic potential allowing the virus to cross the barrier from animals to humans [2]. This is possible due to the reassortment (antigenic shift) of the viral RNA (vRNA) segments during the co-infection with different IAV strains [3]. Besides this ability, there are other IAV traits that make it such a serious threat to public health, i.e., easy transmission, fast adaptation to the particular host immune system, and unceasing evolutionary dynamics correlated with appearing point mutations in the viral genome (antigenic drift). Accumulation of these mutations can eventually lead to the emergence of the virus with novel antigenic properties that can evade the host immune response [4]. Therefore, it is important to discover novel therapeutic strategies and anti-influenza drugs with a new mode of action.

The IAV negative-sense RNA genome consists of eight segments encoding 11 proteins: three polymerase complex proteins: polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2) and polymerase acidic protein (PA), hemagglutinin (HA), nucleoprotein (NP), neuraminidase (NA), two matrix proteins (M1 and M2) and non-structural proteins (NS1, NS2 and PB1-F2) [5]. Interestingly, each of the vRNA segments with nucleoprotein and polymerase subunits forms a viral ribonucleoprotein (vRNP) complex being a basic

element during genome replication and transcription [6]. Overall, vRNP structure is related to its functions. It was reported that the viral replication, RNA packaging, mRNA splicing regulation or recognition by the host immune system are controlled by the RNA structure [7,8]. Within the vRNP complex, NP molecules bind vRNA with high affinity and serve as the regulator of the nuclear export and import of vRNP. However, it was previously confirmed that vRNA can escape from complex with NP and potentially fold into the secondary structure in a dynamic manner [9]. Recently, there has been increasing research indicating the crucial role of viral genome secondary structure in the replication and other biological processes [7], not only in the case of the influenza virus [10,11], but also regarding other viruses. For instance, human immunodeficiency virus (HIV) vRNA structure can regulate genomic RNA transcription and gene expression that affect the cellular innate immunity [12].

Importantly, there is growing interest in the unique non-canonical structures called G-quadruplexes (G4s) [13,14]. They can be formed within guanine-rich (G-rich) sequences called potential quadruplex-forming sequences (PQS). The G-quadruplex involves G-tetrads composed of a planar array of four guanine residues associated through Hoogsteen hydrogen bonds. The G-tracts are connected with each other by different kinds of loop indicating the G-quadruplex folding topology. Depending on the sequence, length of the loops and strand orientation, the G4 structures can be parallel, antiparallel or hybrid type (3 + 1) [15]. Moreover, the number of molecules involved in the G4s formation can influence the folding topology. G-quadruplex structures can be intramolecular meaning that one strand folds within itself to form the G-tetrads or intermolecular consisting of two or more strands [16]. Previously published studies revealed the presence of G4 structures in the genomes of RNA viruses, e.g., Zaire ebolavirus (EBOV) [17], Zika virus [18] and more recently also in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [19] as well as in the IAV [20]. Apparently, these structures seem to be highly conserved in the viral genomes, indicating their important role in the viral life cycle, for instance during replication. These findings encouraged us to investigate the IAV genome for the presence of the unique G-rich sequences.

In the current study, both bioinformatics analysis and biophysical methods allowed us to show that conserved PQS motifs occur within the influenza A/California/07/2009(H1N1) genome and indeed have the ability to fold into G-quadruplexes. Based on these results and taking under consideration the G4s localization within the vRNA segments, it can be assumed that these structures can play an important role in the regulation of different steps of the viral life cycle.

2. Results

2.1. Identification of Potential Quadruplex-Forming Sequences (PQS) Motifs in the Influenza A Virus (IAV) Genome

In this work, we inspected the IAV genome for the presence of the unique G-rich sequences. For this purpose, the sequence of the influenza A/California/07/2009(H1N1) genome was retrieved from the NCBI (National Center for Biotechnology Information) database (RefSeq assembly accession: GCF_001343785.1). There are several bioinformatics tools allowing to search for G-rich regions in RNA sequences prone to fold into G-quadruplex structures [21]. Herein, G4RNA screener (accessed on 15 December 2020) (http://scottgroup.med.usherbrooke.ca/G4RNA_screener) and the Quadruplex-forming G-Rich Sequences (QGRS) Mapper (accessed on 15 December 2020) (http://bioinformatics.ramapo.edu/QGRS/analyze.php) were used to identify PQS motifs across the whole IAV genome. Our bioinformatics analysis revealed the presence of twelve PQS candidates that can potentially adopt RNA G4 structures. The PQS motif characteristics are summarized in Table 1. Interestingly, all PQS candidates contain three or more of at least two continuous guanosine residues (Gs) within a sequence and a half of them contains four (1KW, 3KW, 6KW, 7KW and 10KW) or five (12KW) continuous Gs (Table 1). Therefore, it is plausible that these sequences have the potential to form RNA G-quadruplex structures.

Table 1. Characteristics of potential quadruplex-forming sequences (PQS) motifs from vRNA of the influenza A/California/07/2009(H1N1).

Segment (Protein)	Segment Length (nt)	PQS Name	PQS Location within Segment	Sequence (5′-3′)
1 (PB2)	2341	1KW 6KW	436–460 1018–1040	CUGGUGGGGCAGCAGCAAAGGGGAG GGUGCAUGGGGUUCAGUCGCUGG
2 (PB1)	2341	2KW 7KW 8KW	360–381 2097–2128 2224–2259	UUGGCUGGACCAUGGGCUGGCA GGUAGUGGUCCAUCAAUCGGGUUGAGCUGGGG GGCUGUAUGGAGGAUCUCCAGUAUAAGGGAAUGUGG
3 (PA)	2233	9KW 10KW 3KW	495–526 1381–1412 1532–1564	UUGGAGGUUCCAUUGGUUCUCACAUAUAGGAA GGCAAAGAGGCCCAUCAGGCAAUCUGAGGGG AAGGCUGGGGAAGUUCGGUGGGAGACUUUGGUC
4 (HA)	1779	11KW 4KW	807–834 1133–1157	GGAUGUAUAUUCUGAAAUGGGAGGCUGG UUGGUCAGCACUAGUAGGUGGAUGG
7 (M1) (M2)	1027	12KW	935–959	UCGGCUUUGAGGGGGCCUGACGGGA
8 (NS1) (NS2)	890	5KW	751–778	UCGGCGGAGCCGAUCAAGGAAUGGGGCA

G—G-rich sites which can be involved in RNA G-quadruplex formation.

2.2. Identification of Nucleotide Residues Frequency in PQS Motifs

All PQS motifs found were located within the protein-coding regions of the IAV genome. Obtained results encouraged us to ask the question of whether these PQS candidates are conserved among the IAV strains. To answer this question, we first used the IRD (Influenza Research Database) database (accessed on 15 December 2020) (www.fludb.org) to retrieve complete genomic sequences of the eight IAV segments. Subsequently, the multiple sequence alignments were carried out for each segment separately, and the frequency of nucleotide residues at each position was established using MAFFT software and BioEdit 7.1 program, respectively. The percentage of the PQS motifs similarity across the IAV strains was calculated and is presented in Table 2.

Table 2. Summary of the PQS motifs similarity within the influenza virus type A and H1N1 subtype strains.

PQS Name	Influenza Type A		H1N1 Subtype	
	Number of Sequences	Similarity [%]	Number of Sequences	Similarity [%]
1KW	20,593	86.1	8793	93.0
2KW	20,430	96.0	8597	96.4
3KW	20,069	93.3	8743	93.0
4KW	20,534	71.0	9081	86.0
5KW	12,902	95.1	5569	96.7
6KW	20,593	86.8	8793	92.4
7KW	20,430	96.1	8597	93.4
8KW	20,430	97.5	8597	97.8
9KW	20,069	86.9	8743	97.4
10KW	20,069	80.0	8743	92.3
11KW	20,534	84.1	9081	95.0
12KW	11,392	95.4	4767	99.6

Number of sequences deposited in the Influenza Research Database (December 2020).

This comparative analysis showed that similarity of PQS candidates ranged from 71.0% to 97.5% and from 86.0% to 99.6% for the influenza type A and the H1N1 subtype, respectively (Table 2). As expected, the higher percentage of similarity was obtained for the strains of the H1N1 subtype compared to all the IAV strains. Interestingly, the only exception was observed in the case of 7KW variant, for which the result was opposite (similarity was higher within the influenza type A than within the H1N1 subtype, 96.1 vs. 93.4%, respectively).

Finally, the conservation level of PQS motifs within the H1N1 subtype was estimated using WebLogo application [22,23]. More specific, all PQS motifs were searched in aligned sequences and LOGO representation was generated (Table 3). The number of aligned

sequences used for this analysis varied for each vRNA segments and was presented above in Table 2. Based on our results we observed that the conservation levels were different depending on PQS motifs and ranged from 44% for 4KW to 92% for 12KW (Table 3). However, most of PQS motifs showed the conservation levels between 75% and 83.4% (1KW-3KW, 5KW, 7KW–9KW, Table 3). This indicates that some of the sequences are highly conserved, e.g., 12KW (92%, Table 3), whereas 4KW (44%, Table 3) is significantly less conserved within the analyzed sequences of H1N1 subtype. Interestingly, when we take under consideration the conservation pattern of nucleotides, we can conclude that G residues show relatively high conservation within the G-tracts of PQS motifs. The average conservation level in this position (G-rich regions) was 69.8%. Obtained data suggest that these regions can be important in genome organization of the IAV.

Noticeably, the formation of G4 structures by PQS candidates cannot be unequivocally assumed without experimental validation. Therefore, we chemically synthesized RNA oligomers and used different types of spectroscopic techniques (including nuclear magnetic resonance (NMR), ultraviolet–visible (UV-Vis) melting, circular dichroism (CD) and fluorescence assays) as well as native gel electrophoresis to examine the structure of PQS motifs identified in our bioinformatics analysis

2.3. Characterization of PQS Motifs by Nuclear Magnetic Resonance (^{1}H NMR)

NMR spectroscopy is a powerful tool to study nucleic acids secondary structures including G-quadruplex formation [24,25]. By recording ^{1}H NMR spectra of DNA or RNA samples, we can gain information whether the molecule forms quadruplex structure under appropriate solution conditions. The core of the quadruplexes are G-tetrads in which guanosine residues are associated through Hoogsteen hydrogen bonds. Signals from the imino protons involved in this type of bond are usually in the area of 10–12 ppm, whereas those from 12–15 ppm are characteristic for Watson–Crick base pairs. The imino proton region of NMR spectra between 10 and 15 ppm provides information about the nature of hydrogen bonds between DNA or RNA base pairs [15].

Hence, we first performed ^{1}H NMR experiments to justify the selection of RNA G4-forming motifs. Our NMR analysis of PQS candidates provided direct evidence for the formation of RNA G4 structures by several oligomers. Based on the results obtained, the model PQS oligomers were divided into three groups depending on their predicted structure. The first group includes sequences with the highest propensity to fold into G-quadruplex structures, including 1KW, 7KW and 11KW variants (Figure 1a). Obtained ^{1}H NMR spectra for these PQS displayed the imino proton peaks between 10 and 12 ppm, which are characteristic for Hoogsteen base pairs. The second group contains three PQS oligomers (3KW, 6KW and 9KW) that might be in dynamic equilibrium between two different structures, including RNA G-quadruplex and hairpin. Their ^{1}H NMR spectra demonstrated the signals within the range of 10–12 ppm, suggesting Hoogsteen hydrogen bonds involved in G4 structure formation (Figure 1b). Additionally, the presence of imino resonances found between 12 and 14 ppm supports the Watson–Crick base pairing and possible hairpin structure formation (Figure 1b). These results can indicate that 3KW, 6KW and 9KW variants have lower propensity to form RNA G4 structures in comparison with PQS oligomers assigned to the first group.

On the other hand, the third group of PQS oligomers does not form stable RNA G4 structures. It can be indicated that 2KW, 4KW, 5KW, 8KW, 10KW and 12KW variants adopt a hairpin structure. The structure of these PQS oligomers was predicted using RNAstructure and is shown in Supplementary Materials (Figure S1).

Table 3. PQS motifs presented as LOGO sequences, their conservation and variability levels.

PQS Name	LOGO Sequence	PQS Length (nt)	Conserved Nucleotides	Conservation (%)	Variable Nucleotides	Variability (%)
1KW		25	20	80.0	5	20.0
2KW		22	18	81.8	4	18.2
3KW		33	26	78.8	7	21.2
4KW		25	11	44.0	14	56.0
5KW		28	21	75.0	7	25.0
6KW		23	16	69.6	7	30.4
7KW		32	24	75.0	8	25.0
8KW		36	30	83.4	6	16.6

Table 3. *Cont.*

PQS Name	LOGO Sequence	PQS Length (nt)	Conserved Nucleotides	Conservation (%)	Variable Nucleotides	Variability (%)
9KW		32	26	81.3	6	18.7
10KW		31	20	64.5	11	35.5
11KW		28	18	64.3	10	35.7
12KW		25	23	92.0	2	8.0

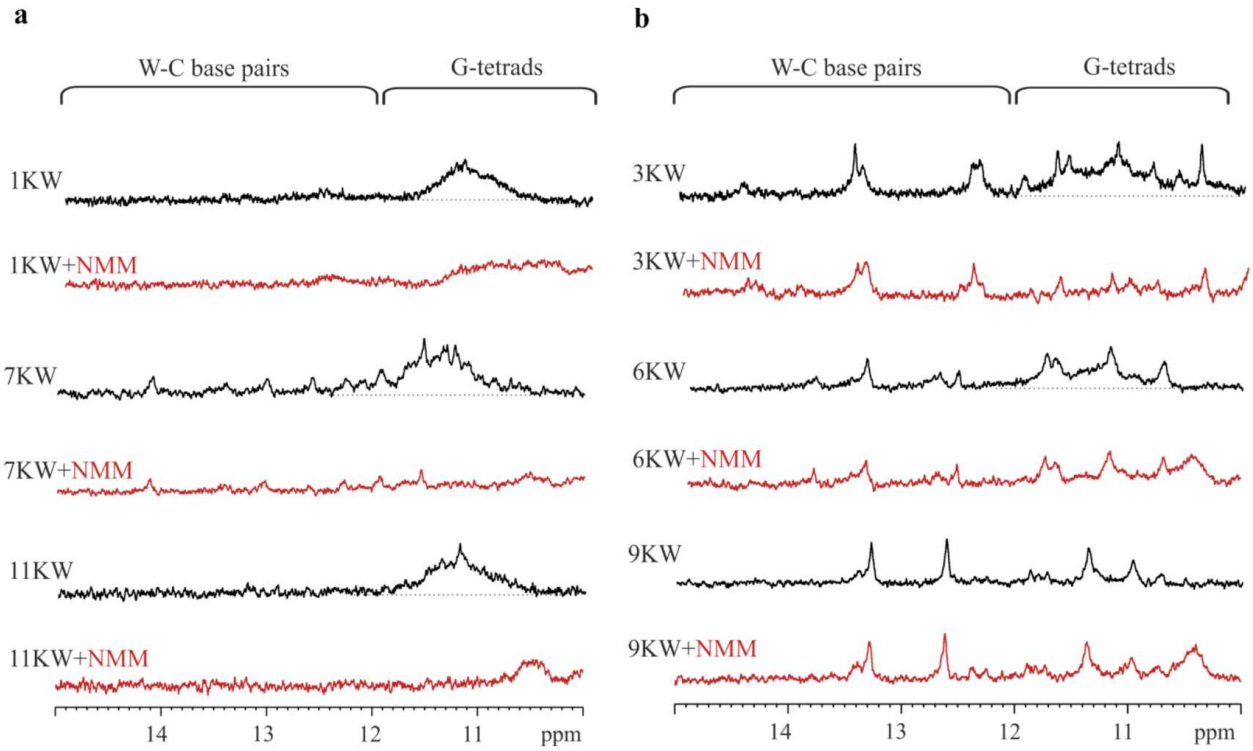


Figure 1. The imino region of the nuclear magnetic resonance (^{1}H NMR) spectra of PQS oligomers. (**a**) ^{1}H NMR spectra for 1KW, 7KW and 11KW variants (first group), (**b**) ^{1}H NMR spectra for 3KW, 6KW and 9KW variants (second group).

Additionally, we performed NMR experiments with the addition of N-methyl mesoporphyrin IX (NMM) that is a dye selectively binding to G-quadruplexes. Experiments were conducted for PQS oligomers with the addition of NMM in a 1:1 molar ratio. In general, the recorded ^{1}H NMR spectra for the first and the second group of PQS displayed broad and shifted signals between 10–12 ppm, indicating binding of NMM to the G4 structure. It is an additional confirmation of RNA G-quadruplex formation by few PQS candidates.

2.4. Determination of Thermal Melting Profiles of PQS Motifs

UV spectroscopy is a fast and simple technique to confirm G4 structure formation and evaluation of its thermal stability. Typically, the G-quadruplex unfolding is associated with a decrease in absorbance at 295 nm wavelength, at which the thermal denaturation of the G4 structure is manifested in an over 50% change in the absorbance amplitude [26,27]. Recording the profile of the melting curve at this wavelength results in hypochromic effect (decrease in absorbance with increasing temperature) [15,26,27]. Therefore, despite the fact that we performed measurements at two wavelengths: 260 and 295 nm, we decided to study the RNA G-quadruplex formation by analysis of the melting profile at 295 nm.

As a result, we were not able to obtain reliable thermodynamic parameters and decided to focus on the melting profiles and changes in the melting temperature determined for model PQS variants. The obtained results demonstrated that the melting curve profiles of four PQS oligomers (1KW, 3KW, 7KW and 11KW, Figure 2) were inverted, indicating RNA G4 formation. Unfortunately, we did not observe the hypochromic effect at 295 nm for remaining PQS variants, which suggests that these structures have lower thermal stability or an equilibrium between different folding forms exists under experimental conditions. Based on the thermal denaturation measurements we calculated the melting temperature values (T_m) that are presented in Figure 2.

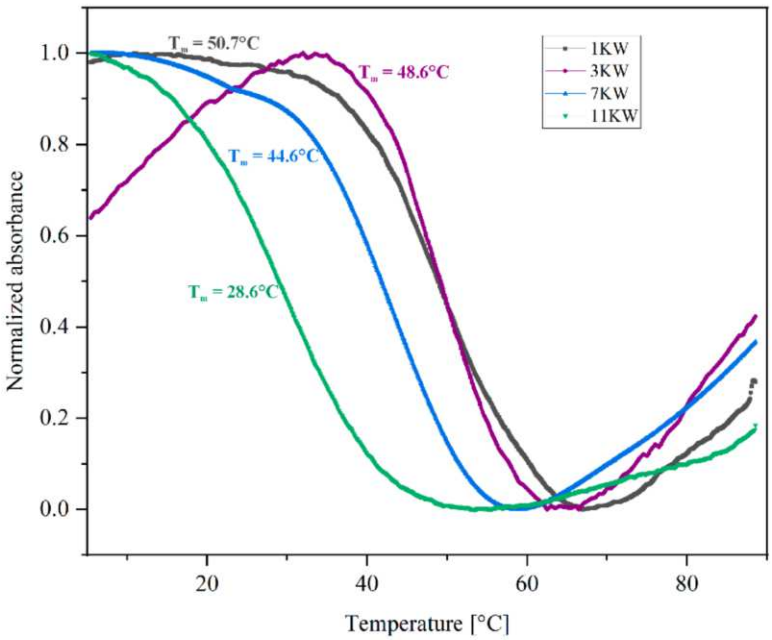


Figure 2. Representative ultraviolet (UV) melting curves of model PQS oligomers with T_m values obtained by monitoring the melting profile at 295 nm.

Based on the data obtained we were able to determine the T_m values for only four PQS oligomers (Figure 2). Individually, the melting analysis revealed that the most stable variant was 1KW (T_m = 50.7 °C, Figure 2), whereas variant 11KW was characterized by the lowest value of T_m parameter (28.6 °C, Figure 2). The 3KW and 7KW variants were less stable in comparison to 1KW and their melting temperatures were 48.6 °C and 44.6 °C, respectively (Figure 2). These results can indicate that the changes in thermal stability of studied PQS oligomers are correlated with the number of G-tracts in a strand. The 1KW oligomer contains two tracts of four continuous Gs and one tract of two continuous Gs, whereas the least stable 11KW variant has one tract of three continuous Gs and three tracts with two continuous Gs.

2.5. Circular Dichroism Spectra of PQS Oligomers

Circular dichroism (CD) spectroscopy is a simple and relatively fast biophysical method to characterize the G-quadruplex folding topology and assess its structural geometry. It is based on a phenomenon that chiral molecules absorb left-handed and right-handed circularly polarized light differently [28]. Generally, CD spectra of nucleic acids display characteristic peaks in the spectral range from 200 nm to 320 nm and are sensitive to base-stacking interactions. Due to the fact that G4 structures can adopt three topologies characterized by differences in tetrad stacking, we can observe a unique CD signature for a given topology [15,29]. Therefore, CD spectroscopy was employed herein to evaluate the ability of PQS motifs to form RNA G4s. The resultant CD spectra for these PQS oligomers displayed the profile typical for a parallel G-quadruplex structure (Figure 3).

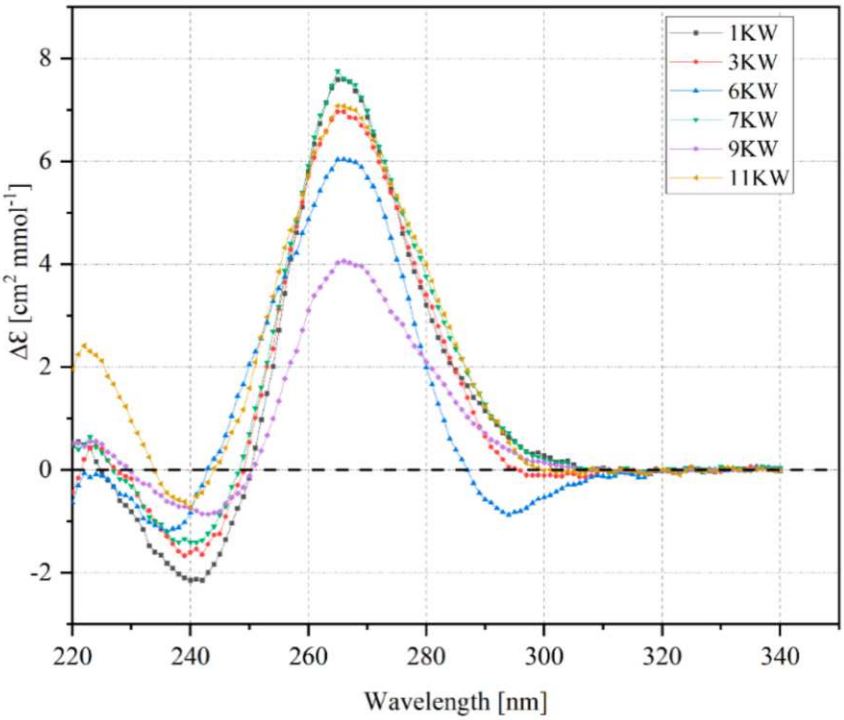


Figure 3. Circular dichroism (CD) spectra of model PQS oligomers recorded at 10 °C.

One negative band around 240 nm and one positive band near 265 nm were observed for studied PQS variants (Figure 3). Surprisingly, we observed an additional negative peak around 295 nm in case of 6KW variant that is difficult to interpret at this moment. Nevertheless, results obtained stay in accordance with previously published data regarding the formation of parallel-stranded G-quadruplexes from the viral genomes [17–19]. Interestingly, most of the RNA G4s adopt a parallel strand orientation, however, the first evidence of an antiparallel RNA G-quadruplex formed by human telomere RNA was recently described by Bao and Xu [30].

2.6. Thioflavin T Fluorescence Intensity Measurements

To further confirm that PQS motifs adopt a G-quadruplex structure, thioflavin T (ThT) fluorescence intensity measurements were performed. In 2013 ThT was applied as a specific ligand towards G-quadruplex folds [31]. Because the fluorescence enhancement results from the specific binding of ThT to G4 structure, the ThT probe was utilized for selective detection of RNA G-quadruplexes by many research groups [18,32,33]. Hence, in our studies we expected to provide additional RNA G-quadruplex confirmation by conducting ThT fluorescence assays. More specifically, we used two oligomers (2KW and 4KW from the IAV genome that failed to adopt a RNA G-quadruplex in previous experiments) as negative controls and a sequence from the Ebola virus genome examined earlier by Wang et al. [17] as a positive control, named 17KW 5′r(GGGGUCAUAUGGGAGGGAUUGAAGG). In addition, the buffer and ThT alone were used as negative controls. ThT was directly added to oligonucleotide solutions and the fluorescence measurements were performed on a microplate reader.

As a result, the fluorescence emission spectra of ThT in the presence of various RNA oligomers at wavelengths ranging from 448 nm to 700 nm and the fluorescence emission of ThT at 482 nm are presented in Figure 4a,b, respectively. Based on the results, the maximum fluorescence peak was recorded around 490 nm in all cases (Figure 4a). In addition, large increments in the fluorescence intensity for all of the RNA oligomers from the first group (1KW, 7KW and 11KW), two oligomers from the second group (3KW and 6KW) and positive control (17KW) compared with negative control probes (2KW and 4KW) were observed (Figure 4). Interestingly, we observed the fluorescence enhancement of ThT also

in the presence of the 5KW variant which was assigned to third group (non-G-quadruplex forming sequence) based on the NMR analysis. This can suggest that 5KW oligomer may adopt the G4 structure under certain conditions. As expected, 9KW variant was the only exception showing no significant increase in ThT fluorescence indicating that this oligomer does not form a G-quadruplex structure, which is consistent with the UV melting analysis.

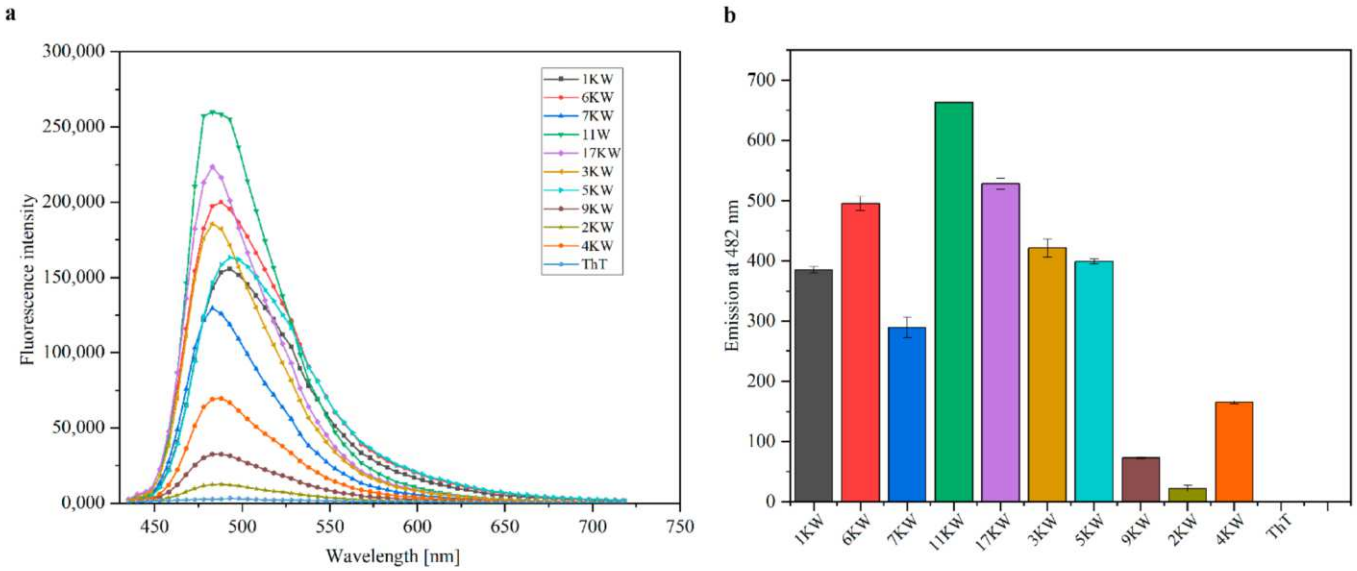


Figure 4. Thioflavin T fluorescence assays for PQS oligomers. (**a**) Fluorescence emission spectra of ThT in the presence of model PQS oligomers, fluorescence of ThT alone is also presented, (**b**) fluorescence emission of ThT at 482 nm for model PQS oligomers compared to ThT alone. The dotted line is a threshold for ThT enhancement supporting G4 formation.

Comparing the fluorescence intensity at 482 nm for tested PQS variants we found a significant increase in ThT fluorescence upon binding to almost all of them (1KW, 7KW, 11KW, 3KW, 6KW and 5KW, Figure 4b). Obtained data showed that these oligomers yielded substantially higher fluorescence enhancement (except 9KW variant) in reference to the 2KW and 4KW selected as non-G-quadruplex forming oligomers. According to these results, we observed that the fluorescence enhancement at 482 nm was 728-, 564- and 537-fold for 11KW, 17KW and 6KW, respectively, compared with the signal of ThT alone (Figure 4b) supporting RNA G-quadruplex formation. Moreover, the substantial increase in the ThT fluorescence emission compared to ThT alone was recorded in the presence of three RNA oligomers (3KW, 5KW and 7KW) and was 466-, 435- and 323-fold, respectively (Figure 4b). These findings indicate that studied PQS variants adopt RNA G-quadruplex structures. In contrast, the 9KW variant did not induce a significant fluorescence increase (only a 78-fold enhancement, Figure 4b), which suggests adopting structure other than G4s. It is worth noting that ThT exhibited higher fluorescence enhancement in the presence of 11KW variant than for our positive control, 17KW (728-fold vs. 564-fold, Figure 4b). This indicates that ThT ligand has strong binding affinity to 11KW variant and supports assumption that this PQS folds into the RNA G4 structure. Taking under consideration the above data and the previously described results [17,18,31,34], it can be concluded that the ThT fluorescence signal may be used to predict G-quadruplex formation and depends on several factors, e.g., nucleotide sequence or available binding sites.

2.7. Native Gel Electrophoresis

Finally, native polyacrylamide gel electrophoresis (PAGE) was performed to confirm the RNA G4s formation. DNA or RNA migrates at different rates in an electric field depending on its secondary structure. Moreover, based on the electrophoretic mobility, we can obtain detailed information about the conformation changes that determine both

folding topology and molecularity in G4 structure. Therefore, we used this method to study RNA G-quadruplex folding patterns.

Due to the fact that NMM is a highly selective quadruplex-specific ligand with a preference to bind to the parallel G4s [35], we applied NMM staining to confirm RNA G-quadruplex formation. Additionally, as we used ThT probe in our fluorescence investigations, we decided to employ ThT to stain the gel and determine the RNA G4s. In PAGE experiments the following RNA oligomers were used: PQS oligomers including two control probes, i.e., 17KW and its variant with mutation in G-rich region named 18KW, the hammerhead ribozyme named HHO as a negative control and three oligo(dT) of various lengths as the mobility markers (named dT and marked on the gel as T35, T45 and T55). PQS oligomers assigned to the third group (2KW, 4KW, 8KW, 10KW and 12KW) were excluded from the PAGE analysis, because no bands on the gels were observed for these variants after ligand staining. The only exception was 5KW oligomer which is suspected to fold into less stable G4 structure. The resultant gels are presented in Figure 5.

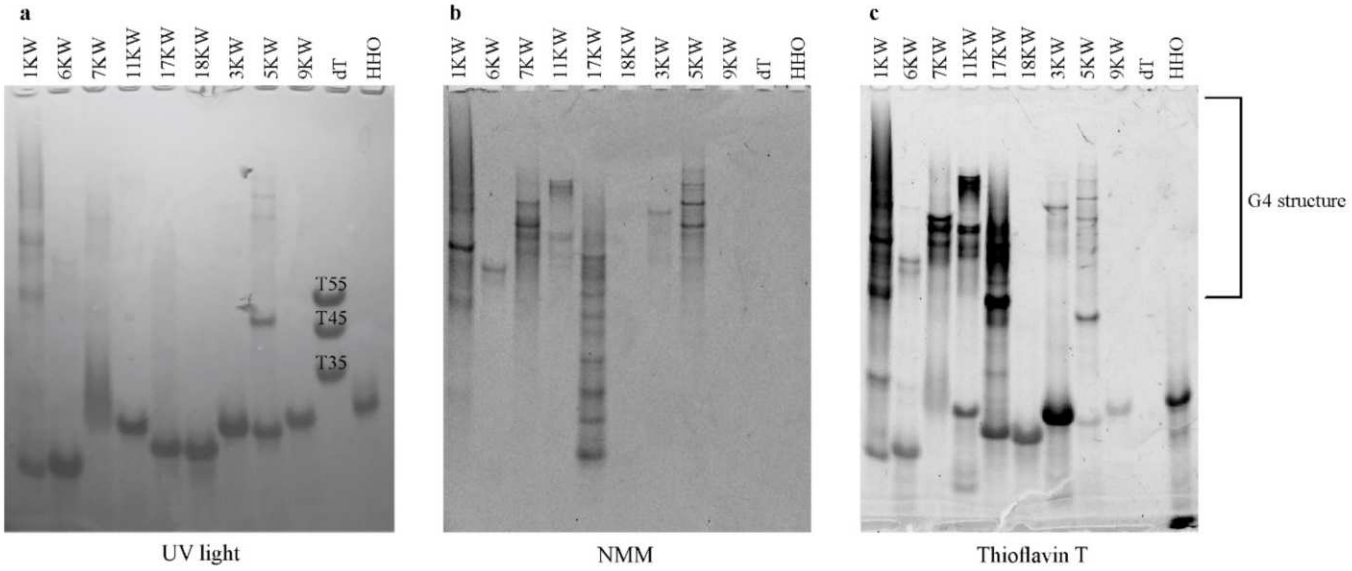


Figure 5. Visualization of model PQS oligomers migration on a native gel. (**a**) The gel under the UV light, (**b**) The gel after staining with G4 specific ligand, N-methyl mesoporphyrin IX (NMM), (**c**) The gel after staining with G4 specific ligand, thioflavin T. On the gels single-stranded dT is a mobility marker and hammerhead ribozyme (HHO) is a negative control.

The gels were first visualized under the UV light (Figure 5a). As expected, we observed evident bands in the lines of all oligomers, which confirm that the samples were properly prepared and loaded into the wells (Figure 5a). The analysis of the gel stained with NMM dye revealed the bands for 1KW, 3KW, 5KW, 6KW, 7KW, 11KW variants and control 17KW, whereas there were no bands in the line of 9KW, 18KW, HHO and dT. This phenomenon indicates that tested PQS motifs from influenza A virus (except 9KW) adopt RNA G-quadruplex folds. Moreover, these oligomers demonstrate different electrophoretic mobility, which indicates differences in their folding molecularity. The number and position of bands in the lines of 1KW, 5KW, 7KW, 11KW and 17KW showed that these structures are probably composed of more than one strand, suggesting bimolecular or tetramolecular G-quadruplexes. Furthermore, as presented in Figure 5b, we noted that for the 1KW, 5KW, 7KW, 11KW and 17KW variants appeared few bands corresponding probably to higher-order G4 structures or suggesting that the equilibrium between different conformations exists. In contrast, only one band in the line of 3KW and 6KW was observed (Figure 5b). This result shows that these PQS motifs can form bimolecular G-quadruplex.

According to the gel stained with ThT dye presented in Figure 5c, we noted that the same bands as in the case of NMM staining were observed, however, there are also many additional bands visible. The bottom panel of the gel demonstrated similar bands to those

visualized under the UV light (Figure 5a vs. Figure 5c). In this regard, we concluded that ThT ligand is less specific towards G-quadruplexes in comparison to NMM dye. These findings stay in accordance with research showing that ThT can also bind to other non-canonical structures, such as duplexes or triplexes as well as to certain trinucleotide repeats forming mismatches [34,36]. Nonetheless, ThT staining assay served as additional confirmation that most of the selected PQS oligomers fold into RNA G4s.

In summary, all the results obtained by using different spectroscopic methods and electrophoretic techniques, allowed us to discover that three PQS motifs from the influenza A/California/07/2009(H1N1) genome adopt RNA G-quadruplex structures. Nonetheless, these G4s can differ in their stability and also folding topology depending on the experimental conditions.

3. Discussion

Influenza virus continuously remains a serious threat for public health. The number of annual deaths worldwide caused by infection with this pathogen reaches above 600,000 cases. Due to the high viral genetic variability, new vaccines need to be developed every year and their effectiveness rarely exceeds 60% [37]. Therefore, development of new therapeutic strategies and the discovery of potential drug targets in vRNA secondary structure are important topics.

Based on the literature we know that the secondary structure of vRNA is highly conserved among viral strains and biologically important for virus life cycle [7]. The viral replication, RNA packaging or its recognition by the host immune system are controlled by the RNA structure. It has been discovered that several short-range vRNA secondary structural motifs are likely to play an important role in virus replication [38]. Moreover, the long-range interactions formed by the partially complementary 5′ and 3′ ends of vRNA are crucial for the regulation of viral transcription and replication [39].

In virus particles and in cells, the vRNA segments are encapsidated as individual RNP complexes. The molecular model of interaction between the vRNA and NP revealed that RNA can loop out from this complex and interact with other vRNA segments or host and virus factors [8,40,41]. Furthermore, in 2018 Williams et al. showed that some regions of vRNA interact more weakly with NP, allowing them to adopt secondary or tertiary RNA structures [41]. Interestingly, Lee and co-workers employed high-throughput sequencing of RNA isolated by the crosslinking immunoprecipitation (HITS-CLIP) methodology and revealed a distinct pattern of NP-association with vRNA [9]. Certain regions of vRNA exhibit strong NP association (peaks), whereas other regions are depleted of NP or can associate and disassociate from NP (non-peaks). Interestingly, they observed a preference in nucleotide composition for NP binding sites, which are relatively guanine-rich and uracil-poor compared to the overall genome-wide nucleotide content. It suggests that vRNA nucleotide content can be an important factor for the degree of NP binding. Moreover, it can be concluded that some regions of influenza vRNA segments not bound by NP would form functional secondary structure and be involved in RNA–RNA interactions [9].

Furthermore, taking into account described by Lee et al. NP-vRNA association profiles (including nucleotide content of NP peaks) in the A/California/07/2009 strain, we can suggest that PQS motifs may occur in G-rich regions that are NP binding positions on the vRNP. However, NP association does not exclude RNA–RNA interactions between viral RNA segments, as reported by Le Sage et al. [42]. Due to the fact that vRNA secondary structure is partially unwound by binding NP, we can postulate that this phenomenon could have an influence on the G4 structure formation in virion or infected cells. On the other hand, it is possible that G4-specific ligand binding to vRNA G-quadruplex can partially inhibit NP association with G-rich regions or induce conformational changes in the RNA structure imposed by NP.

Interestingly, previous research revealed that potential G4 sequences can be found in Zika virus and Zaire ebolavirus (EBOV) genomes and their potential biological functions were studied [17,18]. Burrows and co-workers determined that the formation of G4s by

the conserved Zika virus sequence causes polymerase stalling [18]. On the other hand, Wang et al. reported that a highly conserved G-rich sequence present in the EBOV L gene can fold into RNA G4 structure, stabilization of which by the specific ligand leads to repression of this gene [17]. A similar approach has been applied against hepatitis C virus (HCV) [43]. The authors identified a highly conserved G-rich sequence located within the negative RNA strand and revealed that stabilization of formed G4s by Phen-DC3 ligand causes inhibition of HCV viral replication [43].

In this present study, we used bioinformatical and biophysical methods to search the influenza A/California/07/2009(H1N1) genome for PQS motifs and confirm the G4 structure formation. There are many PQS searching bioinformatics tools reviewed by Lombardi et al. [44]. Here, we used two algorithms: G4RNA screener [45] and QGRS mapper [46] to identify and select PQS sites within the influenza A/California/07/2009(H1N1) genome. We found out that 12 PQS motifs are present in the few segments of the IAV genome. It should be noted that one sequence identified in our study, 12KW, was previously found also by Brázda et al. [20], while other potential quadruplex-forming sequences selected by our research groups varied. A possible explanation for this observation is that we identified PQS from influenza A virus vRNA segments using QGRS mapper and G4RNA screener, while Brázda et al. utilized in this purpose G4hunter Web tool. Moreover, in our study we focused on the influenza A/California/07/2009(H1N1) genome analysis and Brázda et al. analyzed G4-EA-H1N1 genomes. The discrepancy can be also related to the fact that different algorithms were used and the settings or parameters selected for searching varied between our and Mergny's groups. It had an impact on the number of sequences retrieved from the databases as well. In addition, the algorithms used do not consider possible substitutions of G residue by A residue in quadruplex tetrads that have been reported by Kocman and Plavec [47], stable G4 structures with long loops described by Guédin et al. [48], or G-quadruplex with unique features. Therefore, after bioinformatics analysis, the G4 formation by PQS candidates must be confirmed by experimental validation.

Here, we also analyzed the similarity of PQS motifs across the IAV genomic sequences retrieved from the IRD database. Additionally, the PQS conservation analysis based on LOGO sequences showed that G residues within G-tracts, which could be involved in G4 folding, were conserved. Recently, it has been reported that G4-prone sequences from the genomes of RNA viruses show high level of conservation among different genotypes, suggesting their crucial role in the virus life cycle [17,18,43]. For instance, the G-quadruplex formation can inhibit RNA synthesis by the RdRp in HCV negative RNA strand [43].

More recently, some papers described bioinformatics analyses of viral genomes (including human coronaviruses, [49,50] and influenza virus, [20]) to determine the presence of G-rich regions. In 2020, Bartas et al. conducted a systematic and comprehensive analyses to study the occurrence of PQS sites and inverted repeats (IRs) loci in genomes of all known *Nidovirales* [50]. The authors discovered that the G-quadruplex-forming sequences are very rare and unequally distributed in *Nidovirales* genomes, in contrast to IRs. However, location of PQS sites before 3′ UTR, inside 5′ UTR and before mRNA suggests their important regulatory role [50]. Another study concerning G-quadruplex in H1N1 influenza genomes was performed by Brázda et al. [20]. The comparison of 77 H1N1 influenza genomes provided information on G-quadruplex occurrence, localization and variance. The results showed that influenza genomes contain several highly conserved PQS, which can be potential therapeutic targets [20]. Interestingly, in silico analysis of PQS sites in the genome of all the known human-infecting viruses was reported by Lavezzo and coworkers [49]. Comprehensive analysis revealed that PQS sites presence and location are characteristic for each virus class and family. Furthermore, the researchers noticed that PQS motifs are mainly present in the ssRNA viral genomes that are more suitable to G4 structures' formation as they do not require unfolding from a complementary sequence [49]. Moreover, data indicate that most of the virus classes could regulate their genome functions by G4-mediated mechanism [49].

Our studies focused on the identification of PQS motifs and biophysical characterization of RNA G4s formation. We combined several experimental methods that are commonly used for studying the G-quadruplexes [17,18,26,27,33]. A similar investigation was performed for human coronaviruses [51]. All seven analyzed coronavirus genomes have been demonstrated to contain G4 sequences, and SARS-CoV and SARS-CoV-2 harbor the most conserved sequences [51]. This observation suggests that G-quadruplexes can be important elements in the genomes of human coronaviruses. Formation of G4 structures was experimentally confirmed using fluorescence ThT assay and CD spectroscopy [51], which were also used in our studies.

The growing interest in the G-quadruplex structures led to the discovery of many advantageous properties of these unique forms. Development of novel biosensors based on the G4s with untypical folding features or improved binding affinity reviewed by Xi et al. [14] is one of such examples. The other study presented a fluorescence method for detection of the IAV DNA sequence based on the G4-NMM complex and assistance-DNA [52]. The quadruplex-based functional DNA was designed and used for the homogeneous detection of influenza A DNA sequence [52]. Nevertheless, the current coronavirus disease 2019 (COVID-19) outbreak encouraged many scientists to search for new antiviral drugs and develop novel strategies against viral infections. One example can be the study concerning RNA G-quadruplex in SARS-CoV-2 as a potential therapeutic target [53]. The researchers revealed that PQS in SARS-CoV-2 can form stable RNA G4 structures in live cells. Furthermore, they presented that the SARS-CoV-2 N protein levels both in vitro and in vivo were reduced by a specific targeting compound—PDP (pyridostatin derivative) [53].

In conclusion, a combination of bioinformatics tools, spectroscopic techniques and gel electrophoresis method allowed us to achieve a global view of RNA G-quadruplex formation within the sequences of the IAV genome. Results obtained from our experiments are summarized in Table 4. Based on these data we were able to confirm three PQS motifs (1KW, 7KW and 11KW) to fold into G4s.

Table 4. Summary of obtained results with different methods for selected PQS oligomers.

PQS Name	^{1}H NMR Spectra	UV Melting Profile at 295nm	CD Spectra	Thioflavin T Fluorescence Enhancement	Native Polyacrylamide Gel Electrophoresis	G4 Structure Folding
1KW	√	√	√	√	√	yes
7KW	√	√	√	√	√	yes
11KW	√	√	√	√	√	yes
3KW	√ -	√ -	√	√	√	probably
6KW	√ -	√ -	√ -	√	√	probably
9KW	√ -	√ -	√	√ -	√ -	no

To the best of our knowledge, the current investigations provide the first evidence that conserved PQS motifs are present in the influenza A/California/07/2009(H1N1) genome and have a high propensity to fold into RNA G-quadruplex structures. Comparing our investigations with an already published study on the genotype 4 reassortant Eurasian avian-like H1N1 virus (*G4*-EA-H1N1) and the PQS occurrence in *G4*-EA-H1N1 genomes [20], we focused on different approaches, but came to the same conclusion that the IAV genome contains specific G-rich sequences, which can adopt a G-quadruplex fold. Further studies are necessary to confirm the G4s formation in live cells and to obtain a better understanding of the potential role in virus biology, especially taking into the account the fact that the PQS motifs were found within PB1, PB2 and PA segments. Nevertheless, the current study provides information concerning RNA G4 structure that could be explored for use as a novel therapeutic target in the future.

4. Materials and Methods

4.1. Identification of PQS Motifs in the IAV Genome

The influenza A/California/07/2009(H1N1) genome was obtained from the NCBI database (RefSeq assembly accession: GCF_001343785.1). To identify the PQS motifs, the genome was analyzed using two different tools for G-quadruplex prediction, i.e., G4RNA screener web interface (accessed on 15 December 2020) (http://scottgroup.med.usherbrooke.ca/G4RNA_screener, v.0.3, Scott Group Bioinformatics, Sherbrooke, QC, Canada) and the Quadruplex forming G-Rich Sequences (QGRS) Mapper (accessed on 15 December 2020) (http://bioinformatics.ramapo.edu/QGRS/analyze.php, Ramapo College of New Jersey, Mahwah, NJ, USA).

4.2. Sequence Alignment and Frequency of Nucleotide Residues Analysis

The nucleotide sequences for each segment of the IAV were downloaded from the Influenza Research Database (IRD) (accessed on 15 December 2020) (www.fludb.org). Two types of comparative analysis of sequence data were conducted with the help of MAFFT software (version 7, Research Institute for Microbial Diseases, Suita, Osaka, Japan) and the Bioedit 7.1 program (version 7.1, Thomas A. Hall, Vista, CA, USA). The multiple sequence alignments using MAFFT were performed for all IAV genomic sequences deposited in the IRD database and separately for all those of IAV H1N1 subtypes. Next, the aligned sequences were loaded into the Bioedit 7.1 program for the frequency of nucleotide residues analysis. The data obtained were used to examine the PQS motifs similarity across both the IAV subtypes and particularly the IAV H1N1 subtype. Then PQS motifs were searched in aligned segment sequences, LOGO representations were constructed using WebLogo application (accessed on 15 December 2020) (https://weblogo.berkeley.edu/logo.cgi, version 2.8.2, Department of Plant and Microbial Biology, University of Berkley, Berkley, CA, USA) and the conservation levels were determined.

4.3. Synthesis and Purification of Oligonucleotides

RNA oligonucleotides were synthesized on a MerMade12 (BioAutomation, Irving, TX, USA) synthesizer using standard phosphoramidite chemistry on solid support [54]. All oligonucleotides were purified by polyacrylamide gel electrophoresis. The details of deprotection and purification of oligonucleotides have been described previously [55].

4.4. NMR Spectroscopy

^{1}H NMR spectra were acquired on a Bruker AVANCE III 700 MHz spectrometer equipped with a QCI CryoProbe (Bruker, Billerica, MA, USA). The 3 mm thin wall tubes were used with the final sample volume of 200 µL H_2O/D_2O (9:1, v/v). RNA oligonucleotides were dissolved in a 10 mM potassium phosphate buffer (pH 6.8) containing 50 mM KCl and 0.1 mM EDTA (ethylenediaminetetraacetic acid). The experiments were conducted at 25 °C using RNA oligomers with or without the addition of N-methyl mesoporphyrin IX (NMM, Frontier Scientific, Logan, UT, USA) in a 1:1 molar ratio. The samples were annealed by heating at 90 °C for 5 min and then slowly cooled to room temperature. Water suppression was achieved using excitation sculpting. The spectra were processed and prepared with TopSpin 3.0 Bruker Software (version 3.0, Bruker, Billerica, MA, USA).

4.5. Ultraviolet (UV) Melting Measurements

RNA oligonucleotides were dissolved in the same buffer as used for NMR studies. Oligonucleotide single strand concentrations were calculated based on both, absorbance measured above 80 °C and extinction coefficients, which were approximated by a nearest-neighbor model using the Integrated DNA Technologies calculator on the website https://www.idtdna.com/calc/analyzer (Integrated DNA Technologies, Inc., Coralville, IA, USA). Samples were denatured for 5 min at 90 °C and then slowly cooled to room temperature. Measurements were performed in triplicate using quartz cuvettes of 0.1 cm path length with a sample volume of 30 µL. Absorbance versus temperature curves were obtained

by the UV melting method at 260 nm and 295 nm wavelengths in the temperature range 5–90 °C with a heating rate of 0.2 °C/min on a Jasco V-650 spectrophotometer (Jasco Deutschland GmbH, Pfungstadt, Germany) equipped with a thermoprogrammer. Thermal denaturation curves were analyzed and the melting temperatures (T_m) of model PQS oligomers were determined. The T_m values were obtained as an arithmetic average from three repeats.

4.6. Circular Dichroism Spectroscopy

CD spectra were recorded on a Jasco J-815 spectropolarimeter (Jasco Deutschland GmbH, Pfungstadt, Germany) using 1.5 mL quartz cuvettes with a 5 mm path length and the sample volume of 1300 µL. RNA oligonucleotides were dissolved in the same buffer as used for NMR studies, to achieve a sample concentration of 14 µM. All samples were denatured for 5 min at 90 °C and then slowly cooled to room temperature overnight before data collection. Measurements were collected at 10 °C in the 220–340 nm wavelength range with a 1 nm data interval. CD curves were established as an average of three CD measurements. The buffer spectrum was subtracted from the sample spectra. CD spectra were expressed as the difference in the molar absorption ($\Delta\varepsilon$, in units of cm^2 $mmol^{-1}$) of the right-handed and left-handed circularly polarized light and normalized for plotting and comparative purposes using Origin Pro 9.8 software (version Origin Pro 2021 (9.8), Northampton, MA, USA).

4.7. Fluorescence Intensity Measurements

RNA oligonucleotides at concentration of 1 µM were dissolved in the same buffer as used for NMR studies, to obtain a final volume of 150 µL. Samples were denatured for 5 min at 90 °C and then cooled down to room temperature overnight. After folding, the sample solutions were mixed with thioflavin T (ThT, Sigma Aldrich, Saint Louis, MO, USA) at concentration of 0.5 µM. Next, the samples were centrifuged and loaded into a 96-well plate (Greiner Bio-one, Kremsmünster, Austria) in three repeats. The fluorescence measurements were carried out with CLARIOstar Plus multimode plate reader (BMGLabtech, Ortenberg, Germany) at room temperature. ThT fluorescence emission spectra were acquired over the range of 448 to 718 nm. The excitation wavelength was 425 nm. The fluorescence intensity of ThT was measured at 482 nm with excitation at 425 nm. The data analysis was performed using the Origin Pro 9.8 software (version Origin Pro 2021 (9.8), Northampton, MA, USA).

4.8. Native Polyacrylamide Electrophoresis

Native PAGE was conducted to monitor the mobility of RNA oligomers and confirm the RNA G-quadruplex formation. RNA oligonucleotides at concentration of 0.4 µg/mL (absorbance of 0.1 at 260 nm) were dissolved in the same buffer as used for NMR studies, to obtain final volume of 4 µL. Samples were denatured for 5 min at 90 °C and then slowly cooled to room temperature. After folding, 2 µL of 35% glycerol solution was added. Samples were centrifuged and loaded into a 20% native polyacrylamide gel prepared in 1X TBE (Tris-Borate-EDTA) buffer, pH 7.0. The electrophoresis was performed at 150 V for 3 h in a cold room using XCell SureLock Mini-Cell Electrophoresis System (ThermoFisher Scientific, Waltham, MA, USA). For analysis the gel was stained in a solution of 0.1 mg ml^{-1} NMM in 1X TBE for 10 min or in solution of 0.5 µM ThT in 1X TBE for 15 min. After staining, the gels were destained for 10 min in 1X TBE. The resultant gels were visualized with the help of Fujifilm Phosphorimager (Fujifilm Holdings Corporation, Tokio, Japan) and analyzed using MultiGauge Fujifilm software (version 3.0, Fujifilm Holdings Corporation, Tokio, Japan).

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms22116031/s1.

Author Contributions: Conceptualization, M.T. and M.S.; methodology, M.T. and M.S.; investigation, M.T., M.S., K.Z. and R.K.; data curation, M.T. and M.S.; writing—original draft preparation, M.T. and M.S.; writing—review and editing, M.T., M.S., K.Z. and R.K.; visualization, M.T., M.S. and K.Z.; supervision, M.S. and R.K.; funding acquisition, M.S. and R.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Science Centre, grant numbers: UMO-2019/35/D/NZ6/01479 (M.S.) and UMO-2017/25/B/NZ1/02269 (R.K.). The APC was funded by National Science Centre, grant number: UMO-2019/35/D/NZ6/01479 (M.S.).

Data Availability Statement: The data used to support the findings of this study are available upon request to the authors.

Acknowledgments: We would like to thank Paweł Zmora for scientific support and helpful suggestions. The authors also thank Jan Sadoch for technical support during fluorescence intensity measurements.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

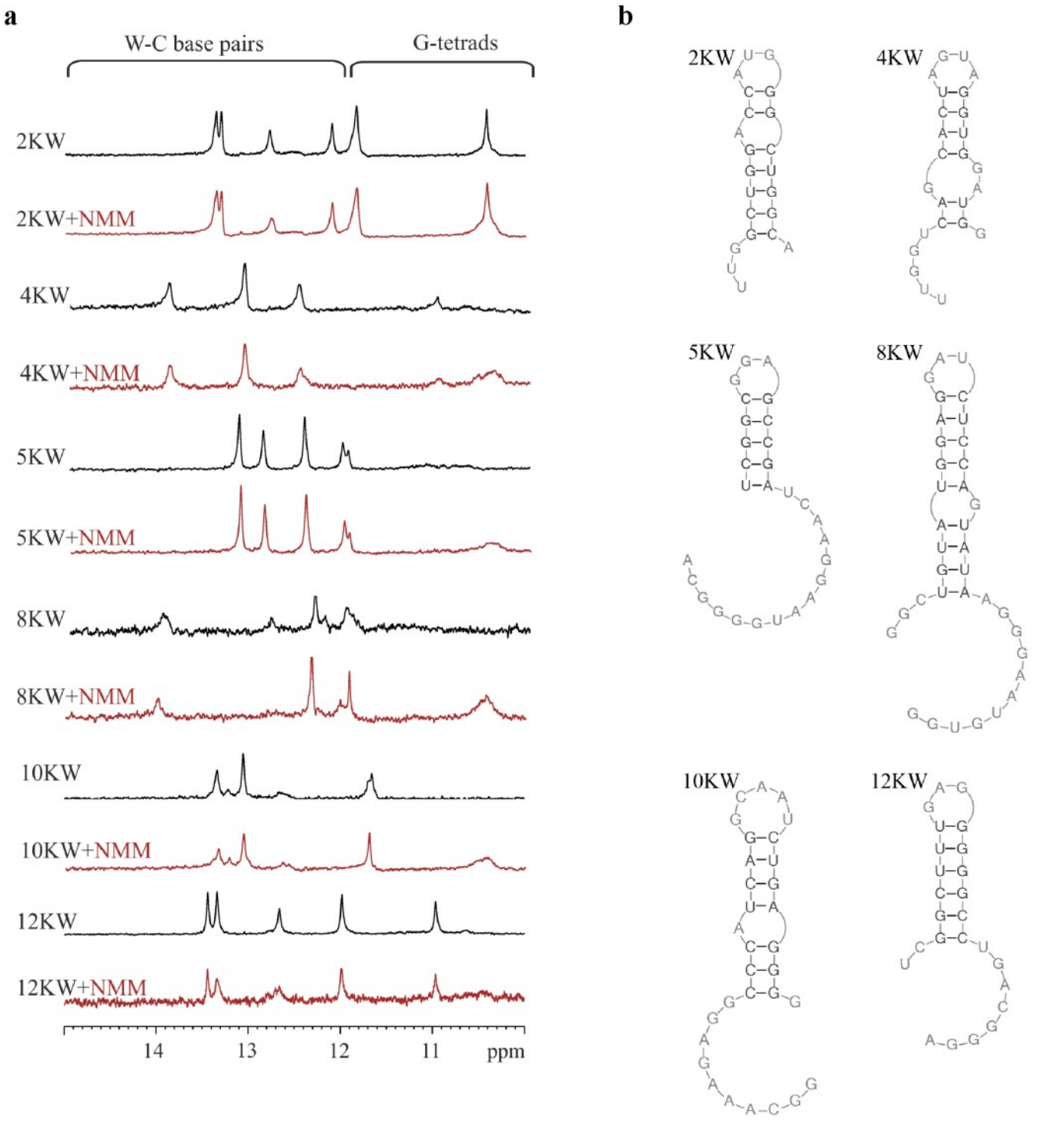
1. Takemoto, Y. Therapeutic approach for seasonal influenza and pandemic. In *Influenza-Therapeutics and Challenges*; Saxena, S.K., Ed.; IntechOpen: London, UK, 2018; Volume 1, pp. 133–150.
2. Mostafa, A.; Abdelwhab, E.M.; Mettenleiter, T.C.; Pleschka, S. Zoonotic potential of influenza A viruses: A comprehensive overview. *Viruses* **2018**, *10*, 497. [CrossRef]
3. Blümel, J.; Burger, R.; Drosten, C.; Gröner, A.; Gürtler, L.; Heiden, M.; Hildebrandt, M.; Jansen, B.; Klamm, H.; Montag-Lessing, T.; et al. Influenza Virus. *Transfus. Med. Hemother.* **2009**, *36*, 32–39.
4. Shao, W.; Li, X.; Goraya, M.U.; Wang, S.; Chen, J.L. Evolution of influenza A virus by mutation and re-assortment. *Int. J. Mol. Sci.* **2017**, *18*, 1650. [CrossRef]
5. Samji, T. Influenza A: Understanding the viral life cycle. *Yale J. Biol. Med.* **2009**, *82*, 153–159.
6. Noda, T.; Kawaoka, Y. Structure of influenza virus ribonucleoprotein complexes and their packaging into virions. *Rev. Med. Virol.* **2010**, *20*, 380–391. [CrossRef]
7. Ferhadian, D.; Contrant, M.; Printz-Schweigert, A.; Smyth, R.P.; Paillart, J.C.; Marquet, R. Structural and functional motifs in influenza virus RNAs. *Front. Microbiol.* **2018**, *9*, 1–11. [CrossRef]
8. Zheng, W.; Tao, Y.J. Structure and assembly of the influenza A virus ribonucleoprotein complex. *FEBS Lett.* **2013**, *587*, 1206–1214. [CrossRef]
9. Lee, N.; Le Sage, V.; Nanni, A.V.; Snyder, D.J.; Cooper, V.S.; Lakdawala, S.S. Genome-wide analysis of influenza viral RNA and nucleoprotein association. *Nucleic Acids Res.* **2017**, *45*, 8968–8977. [CrossRef]
10. Simon, L.M.; Morandi, E.; Luganini, A.; Gribaudo, G.; Martinez-Sobrido, L.; Turner, D.H.; Oliviero, S.; Incarnato, D. In vivo analysis of influenza A mRNA secondary structures identifies critical regulatory motifs. *Sour. Nucleic Acids Res.* **2019**, *47*, 7003–7017. [CrossRef]
11. Szabat, M.; Lorent, D.; Czapik, T.; Tomaszewska, M.; Kierzek, E.; Kierzek, R. RNA secondary structure as a first step for rational design of the oligonucleotides towards inhibition of influenza a virus replication. *Pathogens* **2020**, *9*, 925. [CrossRef]
12. Kao, S.Y.; Calman, A.F.; Luciw, P.A.; Peterlin, B.M. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature* **1987**, *330*, 489–493. [CrossRef] [PubMed]
13. Cammas, A.; Millevoi, S. RNA G-quadruplexes: Emerging mechanisms in disease. *Nucleic Acids Res.* **2017**, *45*, 1584–1595. [CrossRef]
14. Xi, H.; Juhas, M.; Zhang, Y. G-quadruplex based biosensor: A potential tool for SARS-CoV-2 detection. *Biosens. Bioelectron.* **2020**, *167*, 112494. [CrossRef] [PubMed]
15. Małgowska, M.; Gudanis, D.; Teubert, A.; Dominiak, G.; Gdaniec, Z. How to study G-quadruplex structures. *Biotechnologia* **2012**, *93*, 381–390. [CrossRef]
16. Kharel, P.; Becker, G.; Tsvetkov, V.; Ivanov, P. Properties and biological impact of RNA G-quadruplexes: From order to turmoil and back. *Nucleic Acids Res.* **2020**, *48*, 12534–12555. [CrossRef] [PubMed]
17. Wang, S.R.; Zhang, Q.Y.; Wang, J.Q.; Ge, X.Y.; Song, Y.Y.; Wang, Y.F.; Li, X.D.; Fu, B.S.; Xu, G.H.; Shu, B.; et al. Chemical targeting of a G-quadruplex RNA in the Ebola virus L gene. *Cell Chem. Biol.* **2016**, *23*, 1113–1122. [CrossRef]
18. Fleming, A.M.; Ding, Y.; Alenko, A.; Burrows, C.J. Zika virus genomic RNA possesses conserved G-quadruplexes characteristic of the Flaviviridae family. *ACS Infect. Dis.* **2016**, *2*, 674–681. [CrossRef] [PubMed]
19. Ji, D.; Juhas, M.; Tsang, C.M.; Kwok, C.K.; Li, Y.; Zhang, Y. Discovery of G-quadruplex-forming sequences in SARS-CoV-2. *Brief. Bioinform.* **2020**, *22*, 1150–1160. [CrossRef] [PubMed]

20. Brázda, V.; Porubiaková, O.; Cantara, A.; Bohálová, N.; Coufal, J.; Bartas, M.; Fojta, M.; Mergny, J.L. G-quadruplexes in H1N1 influenza genomes. *BMC Genom.* **2021**, *22*, 77. [CrossRef] [PubMed]
21. Puig Lombardi, E.; Londoño-Vallejo, A. A guide to computational methods for G-quadruplex prediction. *Nucleic Acids Res.* **2020**, *48*, 1–15. [CrossRef]
22. Stephens, R.M.; Schneider, T.D. Sequence logos: A new way to display consensus sequences. *Nucleic Acids Res.* **1990**, *18*, 6097–6100.
23. Crooks, G.; Hon, G.; Chandonia, J.; Brenner, S. WebLogo: A sequence logo generator. *Genome Res.* **2004**, *14*, 1188–1190. [CrossRef]
24. Barnwal, R.P.; Yang, F.; Varani, G. Applications of NMR to structure determination of RNAs large and small. *Arch. Biochem. Biophys.* **2017**, *628*, 42–56. [CrossRef]
25. da Silva, M.W. NMR methods for studying quadruplex nucleic acids. *Methods* **2007**, *43*, 264–277. [CrossRef] [PubMed]
26. Rachwal, P.A.; Fox, K.R. Quadruplex melting. *Methods* **2007**, *43*, 291–301. [CrossRef] [PubMed]
27. Mergny, J.L.; Lacroix, L. UV melting of G-quadruplexes. *Curr. Protoc. Nucleic Acid Chem.* **2009**, *37*, 1–15. [CrossRef]
28. Kypr, J.; Kejnovská, I.; Bednářová, K.; Vorlíčková, M. Circular dichroism spectroscopy of nucleic acids. *Compr. Chiroptical Spectrosc.* **2012**, *2*, 575–586.
29. del Villar-Guerra, R.; Trent, J.O.; Chaires, J.B. G-quadruplex secondary structure from circular dichroism spectroscopy. *Angew. Chem.* **2018**, *57*, 7171–7175. [CrossRef]
30. Bao, H.L.; Xu, Y. Investigation of higher-order RNA G-quadruplex structures in vitro and in living cells by 19 F NMR spectroscopy. *Nat. Protoc.* **2018**, *13*, 652–665. [CrossRef]
31. Mohanty, J.; Barooah, N.; Dhamodharan, V.; Harikrishna, S.; Pradeepkumar, P.I.; Bhasikuttan, A.C. Thioflavin T as an efficient inducer and selective fluorescent sensor for the human telomeric G-quadruplex DNA. *J. Am. Chem. Soc.* **2013**, *135*, 367–376. [CrossRef]
32. Xu, S.; Li, Q.; Xiang, J.; Yang, Q.; Sun, H.; Guan, A.; Wang, L.; Liu, Y.; Yu, L.; Shi, Y.; et al. Thioflavin T as an efficient fluorescence sensor for selective recognition of RNA G-quadruplexes. *Sci. Rep.* **2016**, *6*, 24793. [CrossRef]
33. Jing, S.; Liu, Q.; Jin, Y.; Li, B. Dimeric G-quadruplex: An effective nucleic acid scaffold for lighting up thioflavin T. *Anal. Chem.* **2021**, *93*, 1333–1341. [CrossRef]
34. De La Faverie, A.R.; Guédin, A.; Bedrat, A.; Yatsunyk, L.A.; Mergny, J.L. Thioflavin T as a fluorescence light-up probe for G4 formation. *Nucleic Acids Res.* **2014**, *42*, e65. [CrossRef]
35. Smith, J.S.; Johnson, F.B. Isolation of G-quadruplex DNA using NMM-sepharose affinity chromatography. *Methods Mol. Biol.* **2010**, *608*, 207–221.
36. Liu, S.; Peng, P.; Wang, H.; Shi, L.; Li, T. Thioflavin T binds dimeric parallel-stranded GA-containing non-G-quadruplex DNAs: A general approach to lighting up double-stranded scaffolds. *Nucleic Acids Res.* **2017**, *45*, 12080–12089. [CrossRef]
37. Paules, C.I.; Fauci, A.S. Influenza vaccines: Good, but we can do better. *J. Infect. Dis.* **2019**, *219*, S1–S4. [CrossRef] [PubMed]
38. Gilbertson, B.; Zheng, T.; Gerber, M.; Printz-Schweigert, A.; Ong, C.; Marquet, R.; Isel, C.; Rockman, S.; Brown, L. Influenza NA and PB1 gene segments interact during the formation of viral progeny: Localization of the binding region within the PB1 gene. *Viruses* **2016**, *8*, 238. [CrossRef] [PubMed]
39. Gultyaev, A.P.; Fouchier, R.A.M.; Olsthoorn, R.C.L. Influenza virus RNA structure: Unique and common features. *Int. Rev. Immunol.* **2010**, *29*, 533–556. [CrossRef]
40. Arranz, R.; Coloma, R.; Chichón, F.J.; Conesa, J.J.; Carrascosa, J.L.; Valpuesta, J.; Ortín, J.; Martín-Benito, J. The Structure of Native Influenza Virion Ribonucleoproteins. *Science* **2012**, *338*, 1634–1638. [CrossRef]
41. Williams, G.D.; Townsend, D.; Wylie, K.M.; Kim, P.J.; Amarasinghe, G.K.; Kutluay, S.B.; Boon, A.C.M. Nucleotide resolution mapping of influenza A virus nucleoprotein-RNA interactions reveals RNA features required for replication. *Nat. Commun.* **2018**, *9*, 465. [CrossRef] [PubMed]
42. Le Sage, V.; Kanarek, J.P.; Snyder, D.J.; Cooper, V.S.; Lakdawala, S.S.; Lee, N. Mapping of Influenza Virus RNA-RNA Interactions Reveals a Flexible Network. *Cell Rep.* **2020**, *31*, 107823. [CrossRef] [PubMed]
43. Jaubert, C.; Bedrat, A.; Bartolucci, L.; Di Primo, C.; Ventura, M.; Mergny, J.L.; Amrane, S.; Andreola, M.L. RNA synthesis is modulated by G-quadruplex formation in Hepatitis C virus negative RNA strand. *Sci. Rep.* **2018**, *8*, 8120. [CrossRef] [PubMed]
44. Lombardi, E.P.; Londoño-Vallejo, A.; Nicolas, A. Relationship between G-quadruplex sequence composition in viruses and their hosts. *Molecules* **2019**, *24*, 1942. [CrossRef] [PubMed]
45. Garant, J.M.; Perreault, J.P.; Scott, M.S. Motif independent identification of potential RNA G-quadruplexes by G4RNA screener. *Bioinformatics* **2017**, *33*, 3532–3537. [CrossRef]
46. Kikin, O.; D'Antonio, L.; Bagga, P.S. QGRS Mapper: A web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.* **2006**, *34*, 676–682. [CrossRef]
47. Kocman, V.; Plavec, J. Tetrahelical structural family adopted by AGCGA-rich regulatory DNA regions. *Nat. Commun.* **2017**, *8*, 15355. [CrossRef]
48. Guédin, A.; Gros, J.; Alberti, P.; Mergny, J.L. How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res.* **2010**, *38*, 7858–7868. [CrossRef]
49. Lavezzo, E.; Berselli, M.; Frasson, I.; Perrone, R.; Palù, G.; Brazzale, A.R.; Richter, S.N.; Toppo, S. G-quadruplex forming sequences in the genome of all known human viruses: A comprehensive guide. *PLoS Comput. Biol.* **2018**, *14*, e1006675. [CrossRef] [PubMed]

50. Bartas, M.; Brázda, V.; Bohálová, N.; Cantara, A.; Volná, A.; Stachurová, T.; Malachová, K.; Jagelská, E.B.; Porubiaková, O.; Červeň, J.; et al. In-Depth Bioinformatic Analyses of Nidovirales Including Human SARS-CoV-2, SARS-CoV, MERS-CoV Viruses Suggest Important Roles of Non-canonical Nucleic Acid Structures in Their Lifecycles. *Front. Microbiol.* **2020**, *11*, 1–16. [CrossRef]
51. Cui, H.; Zhang, L. G-quadruplexes are present in human coronaviruses including SARS-CoV-2. *Front. Microbiol.* **2020**, *11*, 1–8. [CrossRef]
52. Li, Y.; Liu, W.; Zhu, Y.; Diao, L. A fluorescence method for homogeneous detection of influenza A DNA sequence based on guanine-quadruplex-N-methylmesoporphyrin IX complex and assistance-DNA inhibition. *J. Med. Virol.* **2019**, *91*, 979–985. [CrossRef] [PubMed]
53. Zhao, C.; Qin, G.; Niu, J.; Wang, Z.; Wang, C.; Ren, J.; Qu, X. Targeting RNA G-quadruplex in SARS-CoV-2: A promising therapeutic target for COVID-19? *Angew. Chem. Int. Ed.* **2021**, *60*, 432–438. [CrossRef] [PubMed]
54. Kierzek, E.; Kierzek, R. The thermodynamic stability of RNA duplexes and hairpins containing N6-alkyladenosines and 2-methylthio-N6-alkyladenosines. *Nucleic Acids Res.* **2003**, *31*, 4472–4480. [CrossRef] [PubMed]
55. Xia, T.; SantaLucia, J.; Burkard, M.E.; Kierzek, R.; Schroeder, S.J.; Jiao, X.; Cox, C.; Turner, D.H. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **1998**, *37*, 14719–14735. [CrossRef] [PubMed]

SUPPLEMENTARY DATA

Identification and structural aspects of G-quadruplex-forming sequences from the influenza A virus genome



S1 Fig. A) The imino region of the nuclear magnetic resonance (^{1}H NMR) spectra for 2KW, 4KW, 5KW, 8KW, 10KW, and 12KW variants (third group). B) The structure of PQS oligomers predicted using RNAstructure.

International Journal of
Molecular Sciences

MDPI

Review

Examples of Structural Motifs in Viral Genomes and Approaches for RNA Structure Characterization

Maria Nalewaj † and Marta Szabat *,†

Department of RNA Structural Genomics, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

* Correspondence: szabat@ibch.poznan.pl; Tel.: +48-61-852-85-03

† These authors contributed equally to this work as co-first authors.

Abstract: The relationship between conserved structural motifs and their biological function in the virus replication cycle is the interest of many researchers around the world. RNA structure is closely related to RNA function. Therefore, technological progress in high-throughput approaches for RNA structure analysis and the development of new ones are very important. In this mini review, we discuss a few perspectives on the structural elements of viral genomes and some methods used for RNA structure prediction and characterization. Based on the recent literature, we describe several examples of studies concerning the viral genomes, especially severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and influenza A virus (IAV). Herein, we emphasize that a better understanding of viral genome architecture allows for the discovery of the structure-function relationship, and as a result, the discovery of new potential antiviral therapeutics.

Keywords: viral RNA; secondary and tertiary structures; structure prediction; high-throughput methods; cryo-EM; viruses

Citation: Nalewaj, M.; Szabat, M. Examples of Structural Motifs in Viral Genomes and Approaches for RNA Structure Characterization. *Int. J. Mol. Sci.* **2022**, *23*, 15917. https://doi.org/10.3390/ijms232415917

Academic Editor: Encarnacion Martinez-Salas

Received: 19 October 2022
Accepted: 13 December 2022
Published: 14 December 2022

1. Introduction

It is known that RNA function is closely related with its secondary and higher-order structure. RNA molecules play important roles in diverse catalytic and regulatory biological processes and can serve as disease biomarkers and therapeutic targets. Therefore, it is of great importance to understand the RNA folding properties and broaden our structure-function characterization capabilities. In RNA viruses, its genetic material is used throughout the viral replication cycle, highlighting its crucial role in the virus biology. Questions have arisen on how to best understand the RNA structure features from a biological function perspective, however, the determination of the RNA architecture is still challenging.

Many efforts have been made to study the structure, composition and organization of viral genomes. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) contains enveloped, non-segmented, positive-strand genomic RNA with a 5′ cap and 3′ poly-A tail that allow it to serve as an mRNA during viral protein translation. At the 5′ end, long ORF1ab poly-protein is located followed by four structural (spike, membrane, envelope, nucleocapsid) and several nonstructural proteins. In addition, coronaviruses with a positive-strand ssRNA have the largest RNA genome [1–3]. On the contrary, the influenza A virus (IAV) genome is a negative-strand, segmented, enveloped RNA. Eight segments encode at least eleven viral proteins and viral RNA (vRNA) is transcribed into positive-sense complementary RNA (cRNA) and messenger RNA (mRNA). The vRNA with three polymerase subunits and nucleoprotein forms viral ribonucleoprotein (vRNP) [4]. Based on the literature data, it is known that all RNA viruses characterize by a high mutation rate, which allows them to quickly adapt to the environment and evolve rapidly [3].

Therefore, it is crucial to study RNA secondary structures from IAV and SARS-CoV-2 that are conserved among viral strains [4,5].

Gaining the information on RNA architecture has a great contribution to reveal its biological function within the virus and can be useful for designing various antiviral strategies. However, vRNA dynamics and their propensity to continuous mutations are the main obstacles in designing efficient tools. Therefore, discovering the highly conserved structural elements within the genome and the development of novel methodologies are crucial to RNA structure and function characterization. Based on the available literature, we observed technological progress in the field of structural biology, especially in approaches for high-throughput structure determination. Next generation sequencing (NGS), molecular dynamics (MD) simulation and other emerging methods (including cryo-electron microscopy, cryo-EM) offer scientists the possibility to know the viral genome structural organization at increasingly high resolutions and native conditions.

This mini review summarizes a few perspectives on the viral genome structure and some approaches utilized for structural characterization. Overall, this mini review can be divided into two main parts: (i) the first concerning some viral RNA structural motifs and their potential function, and (ii) the second describing several methods to vRNA folding determination. We will discuss some recent structural studies addressing different human viruses, especially SARS-CoV-2 and IAV. Herein, we will emphasize that a better understanding of viral genome architecture allows for the discovery of structure-function relationships.

2. Examples of Structural Motifs within RNA Virus Genomes

2.1. Internal Ribosome Entry Sites (IRES) and Genome-Scale Ordered RNA Structure (GORS)

The polymorphic nature and conformational flexibility of RNA facilitates their formation of a variety of structures. Many cellular processes are often regulated by RNA secondary and tertiary structures. The ability of the RNAs to fold into stable three-dimensional (3D) structures is important to RNA transcription, translation and splicing [6]. In viruses, RNA structure contributes to viral replication, evasion of host immune response, hijacking of host cell system, protein synthesis and packaging of genetic material into virions [7]. Therefore, a comprehensive understanding of RNA structures can provide important insights into its biological function in the viral life cycle. Figure 1 presents the scheme of the relationship between RNA sequence, structure, and function, which are enormously complex in the living cell.

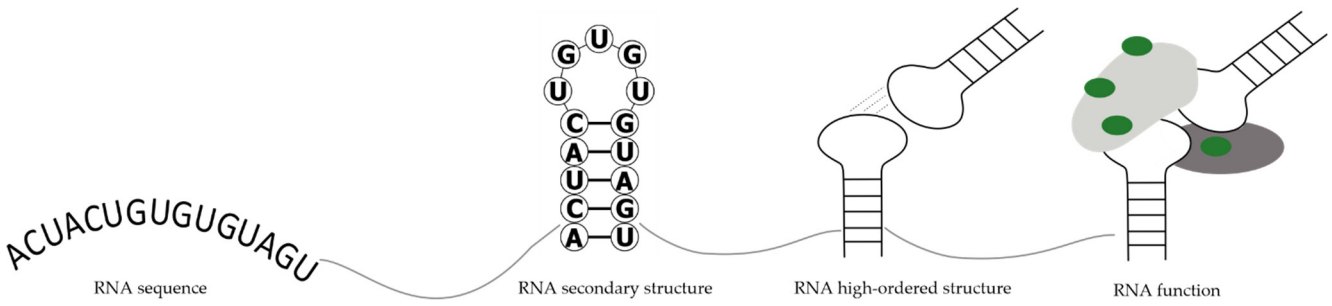


Figure 1. Scheme of the relationship between RNA sequence, structure and function.

In general, RNA secondary structures consist of single-stranded and double-stranded regions, while the 3D arrangement of RNA includes helical duplexes (Figure 2a), triplexes (Figure 2b) and other components that are connected by different types of interaction. The result of these two structural levels is RNA structurome [8,9]. Interestingly, structural RNA motifs in viruses are often highly conserved among different strains and can establish complex networks of RNA-RNA interactions to shape genome architecture and control viral processes. For instance, viral RNA (vRNA) 3D structures can interact with host factors and viral proteins that have an impact on translation machinery [10]. It

has been found that several short-range vRNA secondary structural motifs are likely to play important roles in virus replication [11–13]. Moreover, the long-range interactions formed by the partially complementary 5′ and 3′ ends of vRNA are important for the regulation of viral transcription and replication [14]. Although vRNA structure and its function are reasonably well defined and a few (mostly in vitro and bioinformatics) studies addressing the vRNA and mRNA were performed, structural studies of the vRNA, mRNA and cRNA are still in their infancy.

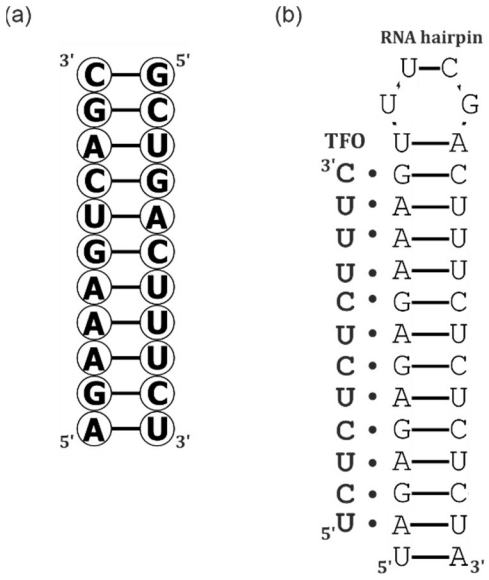


Figure 2. Examples of RNA secondary structures: (**a**) duplex and (**b**) triplex.

From the literature we know that within the viral RNA genome so-called functional RNA domains can be distinguished. Among them are the internal ribosome entry sites (IRES) playing an important role in translation initiation. They can be divided into packed IRESs that are highly conserved, extended IRESs with compact regions, flexible IRESs and inducible IRESs [15]. Various types of IRES present within the viral genome perform the same function despite different sequence and secondary RNA structure. Structural organization of IRES elements play an essential role for their activity, i.e., recruiting ribosomal subunits, initiation factors and/or RNA-protein interactions [16]. It has been described that Dicistroviridae viruses possess the IRES elements within the intergenic region of viral RNA. The architecture of the intergenic region IRESes has been well studied and found that it is highly conserved across various species of dicistroviruses [17]. Another functional RNA motifs from viral genome are cis-acting replication elements located mainly at the 3′ and 5′UTRs, but also within the coding region. These elements are involved in RNA replication and translation [18,19]. Evans and co-workers investigated these motifs in enteroviruses as conserved structural elements and determined the regions of their structure essential for activity. They also studied the influence of the genomic location of cis-acting replication elements on its function [20].

In addition to the presence of local structural motifs within the viral RNA genomes, there is an extensive RNA structure across the entire genome called genome-scale ordered RNA structure (GORS). It has been suggested that GORS is associated with viral persistence or regulates the host innate defense, potentially by shielding viral RNA from recognition in the cell [21,22]. More recently, Simmonds analyzed the coronaviruses genomes including SARS-CoV-2 for the presence of GORS and also discussed its biological purpose in the genome [22]. In other study, the authors investigated the occurrence of extensive RNA structure in the genomes of among others the Flaviviridae and Picornaviridae using large-scale thermodynamic prediction [23]. Furthermore, McFadden et al. showed that GORS can play the functional role in the genome of murine norovirus both in vitro and in

vivo. They revealed that GORS potentially contributes to viral persistence in vivo [24]. As described in this paragraph, there are several specific genome secondary structures present within viral genomes that play functional role in virus biology.

2.2. Pseudoknots, Panhandle Motif, Kissing-Loop Motif and G-Quadruplex Structure

To this day, a lot of research has been carried out to study the secondary and tertiary RNA structures of different viruses [25–28]. Viral RNA pseudoknots are known structural elements with diverse functions in virus biology (Figure 3a). As defined, a pseudoknot is formed upon base-pairing in the loop of a hairpin with a single-stranded region of complementary nucleotides elsewhere in the RNA chain. This motif is overrepresented in functional and regulatory elements of viruses and examples of pseudoknots in viral genomes are well characterized. The most common is a hairpin type (H-type) pseudoknot consisting of two helix stems (S1 and S2) and two or three single-stranded loops (L1–L3). Naturally occurring RNA pseudoknots usually forming interhelix loop contains no more than one nucleotide, and the base-paired stems tend to stack coaxially to form a quasi-continuous helix [29]. Some pseudoknots are formed upon base-pairing of single-stranded structure elements, such as bulge, interior and multibranch loops with complementary regions. It is known that the pseudoknot topology may result in many different structures that suggests the relationship between its structure and function.

For instance, Woodside and co-workers described structural dynamics of pseudoknot from SARS-CoV-2 genome showing conformers which have distinct topologies with different stability and stem-loop structures [30]. The authors found one conformer threading 5′ end and the other with unthreaded 5′ end, based on the structures observed using cryo-EM and simulations [30]. Trinity et al. investigated the structural landscape of SARS-CoV-2 ribosomal frameshifting pseudoknot using a hierarchical folding approach. They identified similarities of this motif between the novel coronavirus SARS-CoV-2, the original SARS-CoV, and the coronavirus responsible for Middle East Respiratory Syndrome [31]. Another example of pseudoknot within the viral RNA is motif found in IRES element upstream of the start codon in Senecavirus A genome [32]. This pseudoknot is predicted to be composed of two stem structures named pseudoknot stem I and II (PKS-I and -II). The authors studied the impact of motif mutations in PKS-I on viral IRES activity and on the rescue of recombinant virus. They found that mutations in PKS-I do not significantly impair the IRES activity and do not abolish luciferase expression in vitro. However, PKS-I-disrupting motifs inhibited virus rescue from mutated cDNA clones [32]. A recent study concerning the dicistrovirus 5′ IRES structural analysis has revealed the presence of pseudoknot motif relevant for proper folding and ribosome recruitment [33]. The dicistrovirus intergenic IRES adopts an overlapping triple pseudoknot RNA structure consisting of pseudoknots I, II and III (PK-I, PK-II and PK-III). This structure binds to the core of the ribosome and can regulate translation initiation [34].

Despite the flexibility and dynamics of RNA molecule, there are structural motifs that remain conserved across the strains and can serve as a potential target for antiviral therapies. One of the widely described structural motifs is a panhandle structure present in the IAV genome (Figure 3b) [4]. All eight influenza genome segments contain this motif arranged by partial base-pairing of the highly conserved 13 nucleotides at the 5′ end and 12 at the 3′ end. The panhandle region plays an important role in viral biological processes such as replication, transcription or packaging. Therefore, it became a target for antiviral compounds, e.g., small molecules. However, because it is a duplex structure, the panhandle is not easily accessible to antisense oligonucleotides or other ligands. Kesy et al. showed that panhandle motif can be bound by the dsRNA-binding PNA conjugated with neamine, what can result in reduction of viral replication [35]. Apparently, a promoter region folds into panhandle structure only in the absence of the polymerase complex, as it was recently discovered by vRNA promoter crystal structure analysis. Interestingly, it has been proposed that the viral polymerase-bound promoter region can form alternative folds. After the polymerase binding to the viral RNA termini, this region forms a

secondary structure so-called 'corkscrew' conformation (Figure 3c). In this structural motif, the very ends of the 5′ and 3′ termini adopt intra-termini hairpin loops. The corkscrew formation has been described by various researchers to be critical element for viral polymerase activity[36–39]. Notably, the polymerase complex binding to the promoter region causes structural rearrangement and the 'hook' conformation in the 5′ end proximal region can be formed. The compact stem-loop (hook) structure is formed by two canonical base-pairs of the 5′ end, flanked at both sides by noncanonical A-A base pairs, leaving the 3′ end single stranded [40]. Additionally, Fodor et al. proposed the formation of the 'fork' model for the transcription initiation. In this structural motif base pairs are formed between nucleotide 13 at the 3′ end and 11 at 5′ end, while the viral RNA ends remain single stranded [41]. These findings show that the viral RNA promoter structure plays an important role during the viral life cycle, mainly in the transcription process.

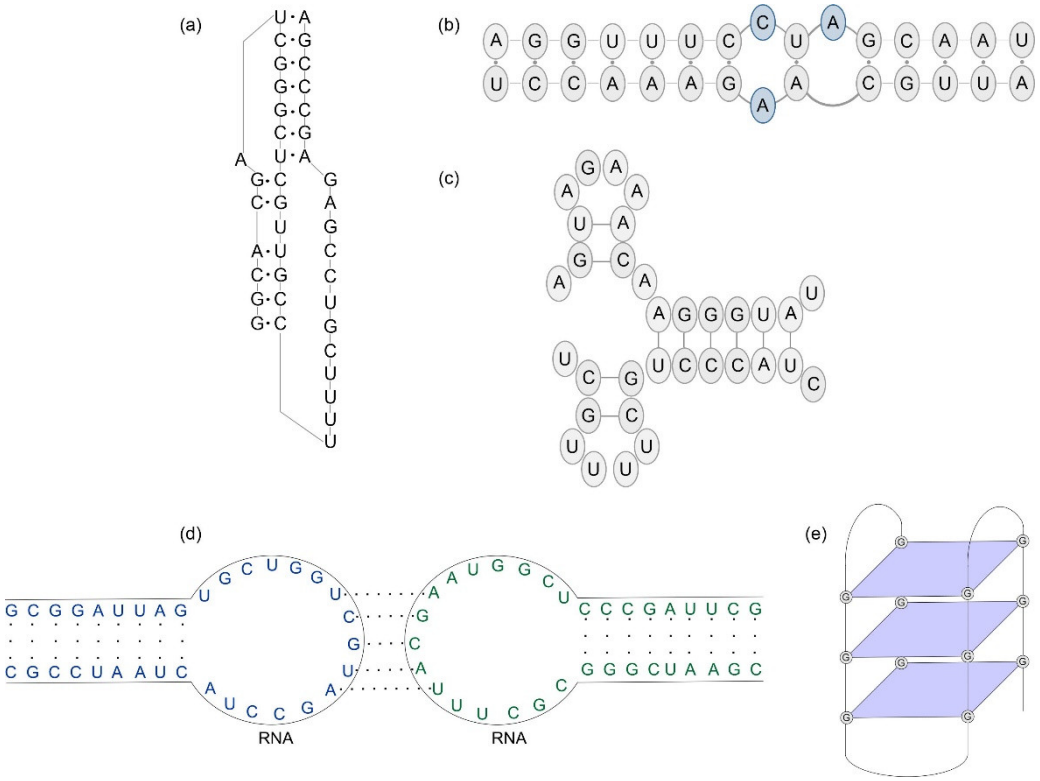


Figure 3. Examples of structural motifs found in the viral genomes, (**a**) pseudoknot, (**b**) panhandle, (**c**) corkscrew conformation, (**d**) kissing-loops, and (**e**) G-quadruplex.

It is worthy of note that the interacting RNA sequences can fold into stem-loops by tertiary interactions known as a 'kissing-loop' mechanism (Figure 3d). Many RNA-RNA complexes can be stabilized by this type of interaction. For instance, in 2017 Gamache et al. proposed a structural model of a symmetrical kissing complex [42]. In this tertiary structure occur two intermolecular interactions of complementary sequences (7 nucleotides) at the apices of two stem-loops. The authors suggest that the Ty1 RNA dimer adopts two forms, in which the kissing-loop interactions are involved. Based on these findings, this kissing complex may initiate dimerization of the retrotransposon Ty1 and packaging into virus-like particles [42]. Another example can be the study concerning the kissing-loop interaction between 5′ and 3′ ends of tick-borne Langat virus genome conducted by Pletnev and co-workers [43]. This investigation showed that destabilization of kissing-loop contacts results in the observed attenuation of virus growth, but does not affect RNA translation or its stability. The researchers suggest that the kissing-loop interaction between the genomic ends can facilitate viral RNA cyclization [43]. Moreover, since the sequences at the 5′ and 3′ loops remain invariant within the tick-borne flaviviruses, thus the

kissing-loop interactions are proposed as a common mechanism in replication regulation for the group of tick-borne encephalitis virus. Furthermore, research on a functional sequence-specific interaction of the viral RNAs revealed a direct interaction between segments: polymerase-basic 1 (PB1) (nucleotides 289–309 in positive sense numbering) and non-structural (NS) (257–277) of an avian H5N2 virus [44]. Structure modeling suggested that these regions fold into hairpins, which are involved in the RNA–RNA contact through kissing-loop interactions. The application of mutagenesis confirmed that through the formation of sequence-specific interaction network, the abovementioned regions play a functional role in virion packaging. More recently, Le Sage and colleagues established a methodology termed 2CIMPL, i.e., dual crosslinking, immunoprecipitation, and proximity ligation to study the genome-wide intersegmental RNA-RNA interactions in the influenza virions [45]. The authors showed that the interactions between viral RNA segments are flexible and that individual segment can coordinate interaction with multiple other segments. Notably, they revealed that the RNA interactions of eight segments occur at multiple sites along each segment and concentrate at specific regions (hotspots) within the viral genome [45].

Among the number of structural motifs present in the viral RNA genomes, it is worth mentioning the G-quadruplex structure formation (Figure 3e). In recent years, G-quadruplexes (G4s) became the subjects of interest in many research groups [46–50]. These structures formed within guanine-rich (G-rich) regions consist of G-tetrads composed of planar array of four guanine residues. Many reported studies concerning the G4 structures showed that they can regulate different biological processes of various organisms. For instance, Basu and colleagues demonstrated that G4 structure can regulate the maturation of human pre-microRNA 92b [51]. The formation of above-structure lead unwinding of the canonical stem-loop structure of pre-miRNA 92b [51]. Moreover, potential G4-forming sequences (PQS) have been found in the viral RNA genomes and their potential biological functions have been studied. In the recent report, Ji et al. discovered the PQS within the RNA genome of SARS-CoV-2 located in the open reading frames of the ORF1 ab, spike, ORF3a, membrane and nucleocapsid genes [50]. Two of the PQS at positions 13385 and 24268 were confirmed to fold into RNA G4 structure by different spectroscopic methods. In addition, molecular dynamics analysis suggested the guanine bases deviating from the G-quartet planes, indicate that the G4 structure was distorted by the viral helicase nsp13. The researchers proposed targeting viral helicase and G-quadruplex structures as an attractive approach for the inhibition of SARS-CoV-2 replication [50]. Other examples are studies reported by Kowalski and co-workers, showing the presence of G4-forming sequence in the viral genome [52]. Using the combination of bioinformatics and biophysical analysis, the authors revealed the occurrence of G-rich sequences in the rhinovirus A, B and C clades, and showed that four of them are highly conserved. Furthermore, they demonstrated that pyridostatin could promote the conformational changes of G-quadruplex structure. The effect of this small molecule on the in vivo uncoating of rhinovirus has been studied. Based on the results, it has been revealed that targeting G4 structure by pyridostatin specifically inhibits the uncoating of the rhinovirus and such an approach can be a promising antiviral tool [52]. In addition to the reported papers concerning G4 structures in the viral genomes, Majee et al. reported the presence of conserved PQS motifs by analyzing the genomic Zika virus (ZIKV) sequences using a bioinformatics tool [46]. Biophysical and biochemical experiments with known G4 structure-binding ligands (N,N′-(9-(4-(dimethylamino)phenylamino)acridine-3,6-diyl)bis(3-(pyrrolidin-1-yl)propanamide) hydrochloride, BRACO-19 and 5,10,15,20-tetrakis-(N-methyl-4-pyridyl)porphine, TMPyP4) confirmed the G-quadruplex formation. Moreover, the researchers demonstrated significant inhibition of infectious ZIKV yield upon the addition of above compounds to the cell culture. Overall, they suggested that binding and stabilization of G4 structure by ligands can potentially affect virus growth, genome replication and viral protein production [46]. Similar investigations were conducted by Tomaszewska et al. presenting the G4 in the IAV genome [48]. They found the PQS motifs within segments of

the IAV and evaluated their conservation among different IAV strains by bioinformatics analysis. In addition, they confirmed the propensity of these PQSs to adopt G4 folds by several biophysical methods. Taking under consideration the G-rich regions localization within the IAV RNA segments and obtained results, the authors assumed that G4 structure may play an important role in the virus life cycle [48]. As one can notice, vRNA folds into various secondary motifs, e.g., pseudoknots, kissing-loop interactions, and G-quadruplexes. Formed structures can also vary depending on viral life cycle stage, such as in the case of influenza promoter region.

3. Selected Approaches to Study Viral RNA Structure

3.1. Several Computational Tools for vRNA Structure Prediction

Viral RNA secondary and tertiary structures can be determined using the combination of computational and experimental studies. However, the flexible nature of vRNA structure makes its characterization more complex. In addition to the structural dynamics of vRNA, the length of this molecule can also be limitation, because long vRNAs can form both local or long-range interactions. Therefore, the structural analysis of such RNAs requires highly robust and reproducible methods.

To this day, many bioinformatics tools and experimental techniques have been developed for studying the RNA folding. RNA structure prediction is a mainstream approach for the structural analysis of RNA. Computational methods are based on thermodynamic calculation of minimum folding energies, phylogenetic analysis based on stochastic context-free grammars (SCFGs) and methods that combine thermodynamic and covariance detection approach. Overall, popular approaches for predicting RNA folding can be divided mainly into two categories: single-sequence and homologous-sequence methods (Table 1). The first applies the algorithms using the minimum free energy principle. Representative examples of these programs are RNAstructure [53], RNAfold [54,55], UNAfold (older Mfold) [56,57] or ScanFold [58] software. The more detailed RNAstructure includes several algorithms for RNA secondary structure prediction and analysis. This software package utilizes thermodynamics and the set of nearest neighbor parameters to predict base pairing probabilities, bimolecular structures and secondary structure common to two unaligned sequences. Moreover, RNAstructure is also useful for siRNA design or can use a number of different types of experiment mapping data (nuclear magnetic resonance, NMR, chemical/enzymatic mapping) to constrain or restrain prediction of the structure [53].

Table 1. Comparison of programs for RNA structure prediction.

Name	Alignment Method	Input Information	Output Information
RNAstructure	Single-sequence	Sequence and experimental data	Lowest free energy structure, base pair probabilities, siRNA design
RNAfold	Single-sequence	ssRNA sequence	RNA secondary structures
UNAfold	Single-sequence	One or two single-stranded sequences	Folding, hybridization, melting pathways simulations
ScanFold	Single-sequence	Input sequence	Functional RNA structures prediction
TurboFold	Homologous-sequence	Set of homologous RNA sequences	Secondary structures predictions for multiple RNA sequence alignments, base pairing probabilities
DMfold	Similar-sequence	Similar RNA in the known structures	RNA secondary structure with pseudoknots
CONTRAfold	Single-sequence	RNA sequence and stochastic context-free grammars (SCFGs)	Predicted RNA secondary structure
2dRNA	Single-sequence	Sequence information of a target RNA	Contact map that includes pseudoknot information

MXfold2	Single-sequence	RNA sequence and thermodynamic parameters	RNA secondary structure with thermodynamic integration
RNAz	Multiple-sequence	Processed multiple sequence alignment	Noncoding RNA detection
SilentMutations (SIM)	Multiple-sequence	Two interacting single-stranded RNAs	Effect of point mutations on the secondary structures of two interacting RNAs, long-range RNA-RNA interactions

The second type of software to predict the RNA structure is based on the determination of base pairs conserved among large number of homologous sequences. An example of such a program is TurboFold for predicting secondary structures for multiple RNA sequence alignments [59]. Recently, Wang et al. proposed a novel method of prediction RNA secondary structure with pseudoknots using deep learning and improved base pair maximization principle called DMfold [60]. The DMfold tool can predict the RNA structure by deep learning similar RNA in the known structures, that uses the similar sequence instead of the homologs. Very recently, similar approach with some modifications was proposed by Xiao and co-workers to obtain a length-dependent deep learning model [61]. This RNA secondary structure prediction method with pseudoknots termed as 2dRNA (a coupled deep learning neural network for RNA) needs only the sequence information of a target RNA as an input (without the sequence alignment). The output of this method is a contact map that includes pseudoknot information [61,62]. Furthermore, Sato et. al. proposed a deep neural network-based method (MXfold2 algorithm) to predict RNA secondary structure with thermodynamic integration [63].

In the literature there are several papers describing prediction and analysis of viral RNA structural motifs using computational approaches [64–66]. One of such approaches is ScanFold, which has been successfully applied in structural analysis of the Zika virus [58], human herpesvirus [67] and most recently SARS-CoV-2 virus [68]. Andrews et. al. used the ScanFold tool to predict functional RNA structures and to characterize them. Expanded analyses included all sequenced RefSeq genomes allowing genome-wide comparison between human herpesviruses. All resulting data are archived and available on a public resource—the RNAStructuromeDB [67]. Very recently, interesting studies concerning RNA genome conservation and secondary structure prediction in SARS-CoV-2 and SARS-related viruses were presented by Das and co-workers [65]. The researchers determined the conservation of genome regions across SARS-CoV-2 sequences available since the COVID-19 outbreak. Moreover, they predicted additional stable structural motifs across the viral genome and provided the conserved regions in SARS-CoV-2 as potential targets for antiviral strategies. Additionally, detailed secondary structure models for the extended 5′ and 3′ UTRs, and frameshifting stimulation element were predicted from RNAz and CONTRAfold 2.0 and presented in above paper [65]. In 2018, Marz with colleagues developed a tool named SilentMutations (SIM) to analyze long-range RNA interactions within viral genomes and structured RNAs [69]. The SIM tool simulates the effect of multiple point mutations on the secondary structure, and needs only two interacting single-stranded RNAs as an input. The authors applied the SIM approach and experimentally confirmed the IAV and hepatitis C virus interactions and presented potential double mutants for in vitro characterization [69].

Advances in the field of bioinformatics and computational techniques enable the exploration of large and complex RNA structures. It is feasible to detect a long-range base pairing or intermolecular interactions, that are common in vRNAs and during vRNAs-host RNA interactions [70]. RNA 3D structure methodology offers a helpful approach to build the large 3D models of linear RNAs. For instance, the RNAcomposer system was utilized to generate 3D models of the Dengue virus RNA fragment [71]. This computer modeling used the secondary structure of RNA predicted by RNAstructure software and resultant data from the chemical mapping (selective 2′-hydroxyl acylation analyzed by primer extension, SHAPE). The researchers provided a new insight into the tertiary

interactions within Dengue virus RNA structures and revealed that some motifs might function as autonomous structural element. This year, Antczak group published a paper presenting a novel webserver, named RNAspider, for topology examination and for analysis of entangled structural elements in RNA tertiary structures [72]. This tool allows automatic identification, classification and visualization of primary and higher-ordered entanglements in RNA 3D structure. The authors focused on two fragments of viral RNA genomes: UTRs from the SARS-CoV-2 genome and exoribonuclease resistant RNA from the flavivirus molecule. They used RNAspider to generate RNA 3D assembled models that present examples of entangled structure [72]. All the data presented above indicate that computational methods can be used both to predict the vRNA structure and to estimate their conservation level in the genome.

A range of RNA structure prediction tools and available resources of data have provided insight into the folding pathway of RNA molecules, including viral RNAs. Furthermore, advances in the next generation sequencing (NGS) and in the high-throughput sequencing have greatly enhanced the RNA structure determination and facilitated the large-scale structural analysis. The NGS technology is a powerful tool for massive scale RNA analysis through cDNA sequencing, also providing information on strand orientation (RNA-seq). Additionally, NGS approaches can often provide information on structure with nucleotide or near-nucleotide resolution. A typical workflow of NGS method consists of RNA fragmentation, reverse transcription into cDNA, and then library preparation according to the NGS methodology of choice. For instance, by the addition of adapters containing specific sequences interacting with the NGS platform, i.e., the surface of the flow cell—Illumina or the beads—Ion Torrent. After sequencing, bioinformatics analysis is conducted to determine the reactivity of individual nucleotides [73]. Overall, a rapid development of strategies based on the genome sequencing can facilitate investigating the virus spread and transmission, understanding the viral evolution or discovering the new antiviral drugs and molecular diagnostic tests. Recently, various HT-based technologies in combination with chemical probing have been successfully used for viral RNA structure analysis [26,45,74,75].

In order to analyze the IAV RNA structure at the single-nucleotide resolution, Dadonaite et al. used selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling method (SHAPE-MaP) [75]. This technique probes the conformational flexibility (base pairing) of each nucleotide. The researchers carried out the RNA structure analysis in both in virio and ex virio conditions. For the in virio experiments, the vRNA was probed directly inside the influenza A/WSN/1933 (H1N1) virions (in the context of vRNPs), whereas, for the ex virio, the eight RNA segments of IAV were transcribed using T7 RNA polymerase or naked viral RNA was purified from deproteinated virus particles. Based on the results, they found that viral RNA segments acquire distinct conformations and can form both intra- and intermolecular RNA-RNA interactions. Moreover, the authors used sequencing of psoralen-crosslinked, ligated and selected hybrids (SPLASH) approach for the structural analysis of interactions occurring in virio. The SPLASH method allowed them to identify the intersegment RNA interactions in the IAV genome [75]. An interesting paper describing the mapping of virus RNA-RNA interactions and RNA secondary structure within the virions was published by Zhou and Routh [26]. They developed an approach based on the NGS platform combined with the viral photoactivatable ribonucleoside crosslinking (vPAR-CL) to study the genomic viral RNA. It has been shown that vPAR-CL methodology allows for reliable identification of RNA-capsid interactions. The capsid binding sites were detected by crosslink specific uridine to cytidine transitions in NGS data. Furthermore, authors applied dimethyl-sulfate mutational profiling with sequencing (DMS-MaPseq) to study the RNA secondary structure in vivo. Based on obtained results they generated a high-resolution profile of single-stranded genomic RNA in virions and revealed that the RNA-protein binding sites favored double-stranded RNA regions [26].

In addition to the computational structure prediction tools described above, as well as HT- sequencing approaches, molecular dynamics (MD) is currently intensively expanding field. Molecular docking is a crucial technique for generating the hypotheses about structure-function relationships in RNA molecules. By computational simulation, it is possible to investigate various biomolecular processes, including conformational dynamics, ligand binding, molecules folding, and revealing the atom positions at very fine temporal resolutions. Hence, MD simulation in molecular biology or virology is an extremely helpful way to understand the molecular architecture of viral RNAs. In 2020, Selvaraj et. Al. performed structure-based virtual screening and MD simulation of SARS-CoV-2 guanine-N7 methyltransferase (nsp14) as a drug target [76]. The authors proposed a 3D model structure of SARS-CoV-2 nsp14 that contains both the guanine-N7 methyltransferase and exoribonuclease domains. Then, they focused on identification of antiviral compounds blocking guanine-N7 methyltransferase (N7-Mtase). It was found that new inhibitors can occupy and interact with the N7-Mtase binding site, which makes them potential antiviral tool against COVID-19 [76]. Among other examples of the use of MD simulations, there is a publication by Hu and colleagues [77]. Integrating MD dynamics and molecular mechanics generalized born surface area (MM-GBSA) calculation method allowed the researchers to find three natural compounds that can inhibit influenza virus RNA polymerase. The authors identified compounds which could bind tightly within polymerase acidic protein-polymerase basic protein 1 (PA-PB1) subunits and studied the nature and conformations stability of the protein-ligand complexes [77]. They suggested that the influenza virus PA-PB1 subunits interaction can be inhibited by compounds from the natural resource. The utility of molecular dynamics simulation in order to analyze vRNA structure has also been shown in the research of Anusuya et al. from 2016 [78]. In general, the Gromiha group used computational fragment-based approaches and molecular dynamics to identify dengue viral RNA-dependent RNA polymerase inhibitor. In more detail, they validated the stability of the dengue polymerase-lead complex, its binding mode, and its interaction pattern. The formation of this complex can keep the polymerase in closed conformation and results in viral replication inhibition [78]. The abovementioned examples of investigations concerning the structure-based virtual screening and MD simulations of viral RNAs underline how important is the study of the structure-function relationship for the drug discovery and development of antiviral therapies.

Overall, the improvement of existing methodology and development of effective protocols for HT techniques will enable us to investigate RNA secondary structure, as well as genome- or transcriptome-wide RNA-RNA interactions. Additionally, integrating the sequencing data obtained from HT methods with MD simulation results can provide important insights into the viral assembly process and be very helpful for understanding the infection process and pathogenesis, as well as the virus maturation in host cells.

The characterization of vRNA structural elements is a first step for rational design of antiviral agents. Nowadays, there is an urgency to identify and develop drugs against the highly pathogenic viruses. An effective and quick approach is a high-throughput screening (HTS) allowing to test existing therapeutics or identify new drug candidates. HTS identifies compounds (for example small molecules) starting from calculating the energy-minimized structures, followed by molecular docking of these structures against specific targets, for example, viral proteins. Crucial in the HTS method is performing biochemical and biological tests using a validated target representing a disease state. A complementary method is a computational screening named virtual HTS (vHTS), which computes the binding affinity of the target to the library compounds. This approach can be divided into two groups: i) ligand-based methods that are based on the similarity between the targets and compounds, and ii) receptor-based methods that rely on the complementarity between the compounds of interest and the target binding sites [79–81].

Several examples of high-throughput activity assays against different viral proteins have been described and validated [82–87]. Yang with colleagues used over 6,000 compounds in a library including approved drugs, drug candidates in clinical trials and

pharmacologically active compounds, and performed high-throughput drug screening to discover new antiviral inhibitors against SARS-CoV-2 papain-like protease (PLpro) [85]. They identified four potential compounds, determined their half-maximal inhibitory concentration values and examined antiviral activities in cell-based assay. Furthermore, the crystal structure of SARS-CoV-2 papain-like protease with identified compound was presented and a unique binding mode was revealed. The solved crystal structure exposed that the interaction within the complex was stabilized through hydrophobic interactions, π-stacking interactions and hydrogen bonds [85]. In another study, a replicon-based HTS approach was subjected for the screening of ZIKV replication inhibitors [88]. This replicon system was engineered and successfully used to screen small-molecule compounds library and to identify two potential anti-ZIKV inhibitors (imidazonaphthyridine and riminophenazine). By the combination of in vitro biochemical and cellular assays, as well as in silico analysis, the researchers examined the antiviral activity of these candidates and revealed that they impair ZIKV replication, probably by interacting with the viral RdRp. Resulting data showed that replicon-based HTS approach represents useful and reliable tool for drug discovery against ZIKV infection [88]. Interestingly, Rong with collaborators discovered a novel anti-influenza virus agent (compound CBS1194) by a retroviral pseudotype-based HTS [84]. The obtained results indicated that this inhibitor can specifically target group 2 hemagglutinins and significantly inhibit virus entry during the IAV infection. Furthermore, to better understand the mechanism of action of CBS1194, biological assays and molecular docking simulations were performed. The results revealed that this compound interferes with HA-mediated membrane fusion by blocking the exposure of the fusion peptide. CBS1194 compound directly binds to a pocket near the fusion peptide and blocks the conformational change within the HA [84]. Although, the HTS strategies allow the identification and validation of agents, which target different factors involved in numerous diseases, they have some limitations and should always be supplemented by biophysical and biological studies of RNA structure. To conclude, nowadays high-throughput techniques gained popularity due to their ability to analyze the structure of a large number of sequences at the same time. By the combination of the HT approach with next generation sequencing or molecular dynamics, researchers can gain a better insight into vRNA secondary structures.

3.2. Several Experimental Methods for vRNA Structure Determination

As mentioned above, RNA structure prediction should be complemented by experimental validation. To improve the accuracy of computational modeling and MD simulations, several approaches for experimental RNA structure determination can be proposed [7,27,89,90]. Currently, the SHAPE method is a commonly used technique. Generally, SHAPE technology uses hydroxyl-selective electrophilic reagents reacting with ribose 2′-hydroxyl group of flexible nucleotides, which causes formation of covalent 2′-O-adduct. Then, the reverse transcriptase during reverse transcription reaction reads out the reactivity of the particular nucleotides. The synthesized cDNA is analyzed by a sequencing gel electrophoresis [7]. Over the past several years, HT methods based on chemical probing have gained popularity, because they enable studying the structure across entire viral RNA genomes by massively parallel sequencing. Technologies based on the chemical probing coupled with mutational profiling (MaP) (for example SHAPE-MaP) and next-generation sequencing (for example dimethyl sulfate mutational profiling with sequencing, DMS-MaPseq) allow for the investigation of RNA structure and its intermolecular interactions with other molecules in a native context. The SHAPE-MaP strategy requires the usage of an enzyme that introduce a change in the chemically modified site, thus internal mutations in cDNA are detected. The implementation of various alternative reagents and adduct detection methods in SHAPE probing, enables RNA folding analysis in the infected or living cells [7,91].

For instance, Heise and co-workers applied the SHAPE-MaP approach to investigate RNA secondary structure of the Chikungunya Virus (CHIKV) genome and 3′UTR variants

of CHIKV [92]. Based on the SHAPE reactivity data, they analyzed base-pairing probabilities for the whole viral genome and presented RNA secondary structure model with highly specific structured regions. Additionally, the researchers also studied the flexibility and folding of CHIKV 3′UTR variants. These structural elements varied in their ability to replicate in mosquito cells. The authors revealed novel specific RNA structures in these 3′UTR variants and suggested their functional importance in viral replication [92]. In 2021, Tomezsko et al. applied DMS-MaPseq to determine the structure of HIV-1 RNA in the cells and suggest its role in splicing [93]. In this study, the authors described a new algorithm called detection of RNA folding ensembles using expectation-maximization (DREEM) and used it to directly cluster the experimental data from DMS-MaPseq (as an input for DREEM). This strategy revealed the alternative conformations that are formed by the same RNA sequences that indicates the heterogeneity of RNA structure across the entire HIV-1 genome. The role of RNA structure in HIV-1 splicing was also examined and obtained findings showed that conformational changes of RNA at critical splice sites can regulate the ration of transcript isoforms [93].

When looking for functional RNA elements within the viral genomes, different high-throughput and interactome strategies can be utilized. SHAPE mapping can be combined with crosslinking and ligation-based RNA-probing strategies. The usage of icSHAPE (in vivo click SHAPE) and PARIS (psoralen analysis of RNA interactions and structures) or SHAPE-MaP and SPLASH allowed for the discovery of secondary structure integration and function in Zika virus and dengue virus [8]. To assess the tertiary structure of SARS-CoV-2, Wan with collaborators performed comprehensive mapping of RNA interactions in vivo [94]. The characterization of RNA folding and functional structure elements along the SARS-CoV-2 genome was conducted using new methods, such as SHAPE-MaP, coupling RNA structure probing with Nanopore sequencing (abbreviated as PORE-cupine) and SPLASH. The authors obtained secondary structure models of the whole SARS-CoV-2 genome. The results showed that SARS-CoV-2 genome is highly structured inside the cells and contains many regions that are involved in intramolecular long-range interactions. Moreover, the study of subgenomic RNA (sgRNA) structure found unique conformations which are different from these in the full-length RNA genome. In addition, individual sgRNAs folding can vary despite sharing the same sequences [94].

An interesting approach, i.e., dual crosslinking, immunoprecipitation and proximity ligation (abbreviated as 2CIMPL) to identify RNA-RNA interactions of influenza virus was proposed [45]. Using this method, multiple interactions between all viral segments were presented. The analysis revealed that some intersegmental junctions are concentrated at the specific regions of IAV genome. This observation indicates that individual segments can coordinate various interactions with other segments. Furthermore, viral nucleoprotein binding to the genome does not prevent RNA-RNA interactions. All obtained results suggest that IAV genome-wide intersegmental interaction networks are complex, flexible and adjustable to sequence variations [45]. Another example of vRNA structure study is comprehensive interrogation of dengue virus RNA genome presented by Dethoff et al. [95]. The researchers first performed SHAPE-MaP analysis to generate the secondary structure of the entire DENV2 RNA genome, both in virion and ex virion. They identified some regions with well-determined structures. Then, they used the RNA interaction groups by mutational profiling (RING-MaP) approach to determine DENV2 RNA tertiary structures. The resulting data revealed numerous RNA regions with potential high-order interactions. Additionally, the analysis of structure-disrupting mutants showed the importance of these complex tertiary structure elements for global genome architecture and viral replication [95].

As it was emphasized before, numerous methodologies have been developed to characterize the architecture of viral RNA genomes. To provide complementary information to these methods, cryo-electron microscopy (cryo-EM) can be used to solve the high-resolution structures of viral assemblies. Technological progress in cryo-EM field gives scientists the opportunity of a better understanding the virus behaviors including structure

organization of its genome. The EM approach includes many imaging techniques, such as cryo-electron tomography, single-particle analysis or microcrystal electron diffraction. In cryo-EM strategies, preparing samples is usually based on their immobilization in vitreous ice via rapid freezing of liquid solution that allows to keep near native cellular environment and protect from radiation damage [96,97]. The high-resolution EM analysis requires small sample amounts and can be performed in very little time. Therefore, it is commonly used in structural biology or virology.

Very recently, Zhang et al. presented the structural and functional results obtained from chemical mapping experiments, cryo-EM and biological assays (in vitro and in cellulo) [98]. They solved tertiary structure of the highly conserved frameshift stimulation element from the SARS-CoV-2 RNA genome. Based on this structure, the researchers developed antisense oligonucleotides (ASO) targeting the structural motifs of SARS-CoV-2 2 frameshift stimulation element. They demonstrated that virus replication might be impaired through modulating this element function (guided by its structure) with ASO tools. Applying cryo-EM and ASO strategy provides information how coronavirus frameshifting might be regulated and inhibited efficiently [98]. An interesting article concerning cryo-EM analysis of the Ebola virus nucleoprotein–RNA complex was written by Wolf and co-workers [99]. Researchers demonstrated near-atomic resolution structure of a recombinant NP-RNA complex expressed in mammalian cells, which was determined by the single-particle cryo-EM. Atomic model of RNA-bound NP and structures of RNA-free monomeric NP model were compared, because the transition between these two states requires conformational changes within both RNA and NP. It was suggested that coordination of nucleoprotein with RNA is sequence independent, and the transition from RNA-free state to a bound state is a dynamic process. Moreover, cryo-EM structure of the Ebola virus NP–RNA complex revealed the role of the N terminus in NP subunit oligomerization, and multiple hydrophobic and electrostatic interactions essential for viral genome encapsidation [99]. In 2021, Shan et al. used cryo-EM as the major approach to investigate the structural features of mumps virus nucleoprotein-RNA complex [100]. They found the occurrence of different assembly forms of nucleocapsid including two rings with 13 and 14 protomers, one stacked-ring filament and two nucleocapsids with distinct helical pitches. Furthermore, the resulting data indicate the conformational flexibility upon binding nucleoprotein to RNA. Structural analysis suggests that the C-terminal tail of viral nucleoprotein can regulate nucleocapsids assembly, and the C-terminal arm is speculated to be important for the transition of mumps virus nucleocapsids between the dense and hyperdense states. Overall, the application of cryo-EM approach facilitated the resolution of the high-resolution structure of different mumps virus nucleocapsids and was very helpful during increasing the knowledge of molecular mechanism for its structural plasticity [100]. To sum up, there is a constant development in experimental methods to study vRNA secondary structure. Moreover, the combination of bioinformatics tools and experimental validation is the base of NGS, MD, and HTS methods that are currently used in most of the research.

4. Future Perspectives

RNA function is closely related to its structure, hence, there is a continual necessity to investigate the genome structure and organization, both in vitro and in a physiological environment. Nowadays, there are numerous methods to study secondary and tertiary structures of different genomes including viruses conserved structural RNA motifs can be involved in various processes crucial for the viral replication cycle, including genome assembly and viral packaging. Therefore, accurate structural analysis of these elements should be based on an appropriate choice of modeling and experimental methods, and proper optimization and data interpretation. Although advances in RNA structure analysis provide information regarding RNA secondary and tertiary structure, much remains to be discovered and additional work is needed to evaluate the full spectrum of vRNA structure and functions in the cell.

Overall, vRNA dynamics and its propensity to continuous mutations are the main obstacles to development of efficient antiviral strategies. Therefore, finding well-described conserved structural motifs is crucial for the alternative treatment of different viral diseases in humans. vRNA can be considered a promising target for novel therapeutics. Michalak et al. suggested that structured regions of the IAV genome that are important for virus biology, can serve as a target for antisense oligonucleotides to inhibit IAV proliferation in cells [101].

Additionally, the identification of conserved structural motifs within the IAV and SARS-CoV-2 genomes can result in the development of so-called 'programmable antiviral agents'. Various pandemic IAV strains were analyzed by SHAPE-chemical mapping and conserved motifs were discovered. Targeting critical conserved vRNA structures, by for instance LNA-modified oligonucleotides results in potent antiviral activity both in vitro and in vivo [102]. This type of 'programmable therapy' is extremely useful in the case of vaccine-resistant viral strains and during the period when vaccines are being developed. The knowledge of the mRNA secondary structure may also be useful in the development of mRNA vaccines, e.g., to design the vaccine against SARS-CoV-2, the structure of S-protein antigen encoded by mRNA-1273 was solved [103].

To sum up, the combination of extensive previous work and future efforts aimed at RNA structure determination and biological function identification will not only improve our basic scientific knowledge, but can also clarify the nature of the relationship between RNA structure and its functions, and further, could lead to discovering novel targets within viral genomes for efficient antiviral strategies.

Author Contributions: M.N. and M.S. performed the literature review and wrote the manuscript, M.S. supervised the process. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the following National Science Center grants: UMO-2019/35/D/NZ6/01479 (M.S.).

Conflicts of Interest: The authors declare that they have no competing interest.

Abbreviations

2CIMPL	dual crosslinking, immunoprecipitation and proximity ligation
ASO	antisense oligonucleotides
Cryo-EM	cryo electron microscopy
DMS-MaPseq	dimethyl-sulfate mutational profiling with sequencing
DREEM	detection of RNA folding ensembles using expectation-maximization
G4	g-quadruplex
GORS	genome-scale ordered RNA structure
HTS	high-throughput screening
IAV	influenza A virus
IRES	internal ribosome entry sites
MaP	mutational profiling
MD	molecular dynamics
MM-GBSA	MD dynamics and molecular mechanics generalized born surface area
NGS	next generation sequencing
NMR	nuclear magnetic resonance
NS	non-structural protein
PA	polymerase acidic protein
PARIS	psoralen analysis of RNA interactions and structures
PB1	polymerase basic 1 protein
PKS	pseudoknot stem
PQS	potential quadruplex-forming sequence
RING-MaP	RNA interaction groups by mutational profiling
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SCFGs	stochastic context-free grammars

sgRNA	subgenomic RNA
SHAPE	selective 2′-hydroxyl acylation analyzed by primer extension
SPLASH	sequencing of psoralen-crosslinked, ligated and selected hybrids
vHTS	virtual high-throughput screening
vPAR-CL	viral photoactivatable ribonucleoside crosslinking
vRNA	viral RNA
ZIKV	Zika virus

References

1. Rahimi, A.; Mirzazadeh, A.; Tavakolpour, S. Genetics and genomics of SARS-CoV-2: A review of the literature with the special focus on genetic diversity and SARS-CoV-2 genome detection. *Genomics* **2021**, *113*, 1221–1232.
2. Phan, T. Genetic diversity and evolution of SARS-CoV-2. *Infect. Genet. Evol.* **2020**, *81*, 104260.
3. ur Rehman, S.; Shafique, L.; Ihsan, A.; Qingyou, L. Evolutionary trajectory for the emergence of novel coronavirus SARS-CoV-2. *Pathogens* **2020**, *9*, 240.
4. Ferhadian, D.; Contrant, M.; Printz-Schweigert, A.; Smyth, R.P.; Paillart, J.C.; Marquet, R. Structural and functional motifs in influenza virus RNAs. *Front. Microbiol.* **2018**, *9*, 559.
5. Huston, N.C.; Wan, H.; Strine, M.S.; de Cesaris Araujo Tavares, R.; Wilen, C.B.; Pyle, A.M. Comprehensive in vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. *Mol. Cell* **2021**, *81*, 584–598.e5.
6. Jacobs, E.; Mills, J.D.; Janitz, M. The role of RNA structure in posttranscriptional regulation of gene expression. *J. Genet. Genom.* **2012**, *39*, 535–543.
7. Boerneke, M.A.; Ehrhardt, J.E.; Weeks, K.M. Physical and functional analysis of viral RNA genomes by SHAPE. *Annu. Rev. Virol.* **2019**, *6*, 93–117.
8. Wang, X.W.; Liu, C.X.; Chen, L.L.; Zhang, Q.C. RNA structure probing uncovers RNA structure-dependent biological functions. *Nat. Chem. Biol.* **2021**, *17*, 755–766.
9. Yang, X.; Yang, M.; Deng, H.; Ding, Y. New era of studying RNA secondary structure and its influence on gene regulation in plants. *Front. Plant Sci.* **2018**, *9*, 1–7.
10. Lozano, G.; Fernandez, N.; Martinez-Salas, E. Modeling three-dimensional structural motifs of viral IRES. *J. Mol. Biol.* **2016**, *428*, 767–776.
11. Gerber, M.; Isel, C.; Moules, V.; Marquet, R. Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol.* **2014**, *22*, 446–455.
12. Noton, S.L.; Fearns, R. Initiation and regulation of paramyxovirus transcription and replication. *Virology* **2015**, *479–480*, 545–554.
13. Yang, D.; Leibowitz, J.L. The structure and functions of coronavirus genomic 3′ and 5′ ends. *Virus Res.* **2015**, *206*, 120–133.
14. Gultyaev, A.P.; Fouchier, R.A.M.; Olsthoorn, R.C.L. Influenza virus RNA structure: Unique and common features. *Int. Rev. Immunol.* **2010**, *29*, 533–556.
15. Romero-López, C.; Berzal-Herranz, A. Unmasking the information encoded as structural motifs of viral RNA genomes: A potential antiviral target. *Rev. Med. Virol.* **2013**, *23*, 340–354.
16. Nakashima, N.; Uchiumi, T. Functional analysis of structural motifs in dicistroviruses. *Virus Res.* **2009**, *139*, 137–147.
17. Khong, A.; Bonderoff, J.M.; Spriggs, R.V.; Tammpere, E.; Kerr, C.H.; Jackson, T.J.; Willis, A.E.; Jan, E. Temporal regulation of distinct internal ribosome entry sites of the dicistroviridae cricket paralysis virus. *Viruses* **2016**, *8*, 25.
18. Niepmann, M.; Shalamova, L.A.; Gerresheim, G.K.; Rossbach, O. Signals involved in regulation of hepatitis C virus RNA genome translation and replication. *Front. Microbiol.* **2018**, *9*, 395.
19. Romero-López, C.; Barroso-Deljesus, A.; Berzal-Herranz, A. The chaperone-like activity of the hepatitis C virus IRES and CRE elements regulates genome dimerization. *Sci. Rep.* **2017**, *7*, 43415.
20. Goodfellow, I.G.; Kerrigan, D.; Evans, D.J. Structure and function analysis of the poliovirus cis-acting replication element (CRE). *RNA* **2003**, *9*, 124–137.
21. Stewart, H.; Bingham, R.J.; White, S.J.; Dykeman, E.C.; Zothner, C.; Tuplin, A.K.; Stockley, P.G.; Twarock, R.; Harris, M. Identification of novel RNA secondary structures within the hepatitis C virus genome reveals a cooperative involvement in genome packaging. *Sci. Rep.* **2016**, *6*, 22952.
22. Simmonds, P. Pervasive RNA secondary structure in the genomes of SARS-CoV-2 and other coronaviruses. *MBio* **2020**, *11*, e01661-20.
23. Simmonds, P.; Azim Ansari, M. Extensive C->U transition biases in the genomes of a wide range of mammalian RNA viruses; potential associations with transcriptional mutations, damage- or host-mediated editing of viral RNA. *PLoS Pathog.* **2021**, *17*, e1009596.
24. McFadden, N.; Arias, A.; Dry, I.; Bailey, D.; Witteveldt, J.; Evans, D.J.; Goodfellow, I.; Simmonds, P. Influence of genome-scale RNA structure disruption on the replication of murine norovirus—Similar replication kinetics in cell culture but attenuation of viral fitness in vivo. *Nucleic Acids Res.* **2013**, *41*, 6316–6331.
25. Soszynska-Jozwiak, M.; Ruszkowska, A.; Kierzek, R.; O'leary, C.A.; Moss, W.N.; Kierzek, E. Secondary structure of subgenomic RNA M of SARS-CoV-2. *Viruses* **2022**, *14*, 322.

26. Zhou, Y.; Routh, A. Mapping RNA-capsid interactions and RNA secondary structure within virus particles using next-generation sequencing. *Nucleic Acids Res.* **2020**, *48*, e12.
27. Tomezsko, P.; Swaminathan, H.; Rouskin, S. Viral RNA structure analysis using DMS-MaPseq. *Methods* **2020**, *183*, 68–75.
28. Li, P.; Zhou, X.; Xu, K.; Zhang, Q.C. RASP: An atlas of transcriptome-wide RNA secondary structure probing data. *Nucleic Acids Res.* **2021**, *49*, D183–D191.
29. Brierley, I.; Pennell, S.; Gilbert, R.J.C. Viral RNA pseudoknots: Versatile motifs in gene expression and replication. *Nat. Rev. Microbiol.* **2007**, *5*, 598–610.
30. Neupane, K.; Zhao, M.; Lyons, A.; Munshi, S.; Ileperuma, S.M.; Ritchie, D.B.; Hoffer, N.Q.; Narayan, A.; Woodside, M.T. Structural dynamics of single SARS-CoV-2 pseudoknot molecules reveal topologically distinct conformers. *Nat. Commun.* **2021**, *12*, 4749.
31. Trinity, L.; Lansing, L.; Jabbari, H.; Stege, U. SARS-CoV-2 ribosomal frameshifting pseudoknot: Improved secondary structure prediction and detection of inter—Viral structural similarity. *bioRxiv* **2020**, *2*, 1–17.
32. Liu, F.; Wang, N.; Wang, Q.; Shan, H. Motif mutations in pseudoknot stem I upstream of start codon in Senecavirus A genome: Impacts on activity of viral IRES and on rescue of recombinant virus. *Vet. Microbiol.* **2021**, *262*, 109223.
33. Gross, L.; Vicens, Q.; Einhorn, E.; Noireterre, A.; Schaeffer, L.; Kuhn, L.; Imler, J.L.; Eriani, G.; Meignin, C.; Martin, F. The IRES5′UTR of the dicistrovirus cricket paralysis virus is a type III IRES containing an essential pseudoknot structure. *Nucleic Acids Res.* **2017**, *45*, 8993–9004.
34. Wang, X.; Vlok, M.; Flibotte, S.; Jan, E. Resurrection of a viral internal ribosome entry site from a 700 year old ancient Northwest territories cripavirus. *Viruses* **2021**, *13*, 493.
35. Kesy, J.; Patil, K.M.; Kumar, S.R.; Shu, Z.; Yong, H.Y.; Zimmermann, L.; Ong, A.A.L.; Toh, D.F.K.; Krishna, M.S.; Yang, L.; et al. A short chemically modified dsRNA-binding PNA (dbPNA) inhibits influenza viral replication by targeting viral RNA panhandle structure. *Bioconjug. Chem.* **2019**, *30*, 931–943.
36. 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.* **2004**, *78*, 6263–6270.
37. Brinson, R.G.; Szakal, A.L.; Marino, J.P. Structural characterization of the viral and cRNA panhandle motifs from the infectious salmon anemia virus. *J. Virol.* **2011**, *85*, 13398–13408.
38. Liu, G.; Park, H.-S.; Pyo, H.-M.; Liu, Q.; Zhou, Y. Influenza A virus panhandle structure is directly involved in RIG-I activation and interferon induction. *J. Virol.* **2015**, *89*, 6067–6079.
39. Tomescu, A.I.; Robb, N.C.; Hengrung, N.; Fodor, E.; Kapanidis, A.N. Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E3335–E3342.
40. Pflug, A.; Guilligay, D.; Reich, S.; Cusack, S. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* **2014**, *516*, 355–360.
41. 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.* **1995**, *69*, 4012–4019.
42. Gamache, E.R.; Doh, J.H.; Ritz, J.; Laederach, A.; Bellaousov, S.; Mathews, D.H.; Joan Curcio, M. Structure-function model for kissing loop interactions that initiate dimerization of Ty1 RNA. *Viruses* **2017**, *9*, 93.
43. Tsetsarkin, K.A.; Liu, G.; Shen, K.; Pletnev, A.G. Kissing-loop interaction between 5′ and 3′ ends of tick-borne Langat virus genome 'bridges the gap' between mosquito-and tick-borne flaviviruses in mechanisms of viral RNA cyclization: Applications for virus attenuation and vaccine development. *Nucleic Acids Res.* **2016**, *44*, 3330–3350.
44. Gavazzi, C.; Yver, M.; Isel, C.; Smyth, R.P.; Rosa-Calatrava, M.; Lina, B.; Moulès, V.; Marquet, R. A functional sequence-specific interaction between influenza A virus genomic RNA segments. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16604–16609.
45. Le Sage, V.; Kanarek, J.P.; Snyder, D.J.; Cooper, V.S.; Lakdawala, S.S.; Lee, N. Mapping of influenza virus RNA-RNA interactions reveals a flexible network. *Cell Rep.* **2020**, *31*, 107823.
46. Majee, P.; Pattnaik, A.; Sahoo, B.R.; Shankar, U.; Pattnaik, A.K.; Kumar, A.; Nayak, D. Inhibition of Zika virus replication by G-quadruplex-binding ligands. *Mol. Ther.-Nucleic Acids* **2021**, *23*, 691–701.
47. Sarkar, S.; Armitage, B.A. Targeting a potential G-quadruplex forming sequence found in the West Nile virus genome by complementary gamma-peptide nucleic acid oligomers. *ACS Infect. Dis.* **2021**, *7*, 1445–1456.
48. Tomaszewska, M.; Szabat, M.; Zielińska, K.; Kierzek, R. Identification and structural aspects of G-quadruplex-forming sequences from the influenza A virus genome. *Int. J. Mol. Sci.* **2021**, *22*, 6031.
49. Sun, X.; Zhang, Q.; Al-Hashimi, H.M. Resolving fast and slow motions in the internal loop containing stem-loop 1 of HIV-1 that are modulated by Mg^{2+} binding: Role in the kissing-duplex structural transition. *Nucleic Acids Res.* **2007**, *35*, 1698–1713.
50. Ji, D.; Juhas, M.; Tsang, C.M.; Kwok, C.K.; Li, Y.; Zhang, Y. Discovery of G-quadruplex-forming sequences in SARS-CoV-2. *Brief. Bioinform.* **2021**, *22*, 1150–1160.
51. Mirihana Arachchilage, G.; Dassanayake, A.C.; Basu, S. A potassium ion-dependent RNA structural switch regulates human pre-miRNA 92b maturation. *Chem. Biol.* **2015**, *22*, 262–272.
52. Real-Hohn, A.; Groznica, M.; Kontaxis, G.; Zhu, R.; Chaves, O.; Vazquez, L.; Hinterdorfer, P.; Blaas, D.; Kowalski, H. Targeting G-quadruplexes in the rhinovirus genome by Pyridostatin inhibits uncoating and highlights a critical role for sodium ions. *Res. Sq.* **2021**, doi:10.21203/rs.3.rs-646190/v1.
53. Bellaousov, S.; Reuter, J.S.; Seetin, M.G.; Mathews, D.H. RNAstructure: Web servers for RNA secondary structure prediction and analysis. *Nucleic Acids Res.* **2013**, *41*, W471–W474.

54. Gruber, A.R.; Lorenz, R.; Bernhart, S.H.; Neuböck, R.; Hofacker, I.L. The Vienna RNA Websuite. *Nucleic Acids Res.* **2008**, *36*, 70–74.
55. Lorenz, R.; Bernhart, S.H.; zu Siederdissen, C.H.; Tafer, H.; Flamm, C.; Stadler, P.F.; Hofacker, I.L. ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **2011**, *6*, 26.
56. Markham, N.R.; Zuker, M. UNAFold: Software for nucleic acid folding and hybridization. In *Bioinformatics, Volume II: Structure, Function and Applications*; Keith, J.M., Ed.; Humana Press: Totowa, NJ, USA, 2008; Volume 453, pp. 3–31.
57. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **2003**, *31*, 3406–3415.
58. Andrews, R.J.; Roche, J.; Moss, W.N. ScanFold: An approach for genome-wide discovery of local RNA structural elements-applications to Zika virus and HIV. *PeerJ* **2018**, *6*, e6136.
59. Harmanci, A.O.; Sharma, G.; Mathews, D.H. TurboFold: Iterative probabilistic estimation of secondary structures for multiple RNA sequences. *BMC Bioinform.* **2011**, *12*, 108.
60. Wang, L.; Liu, Y.; Zhong, X.; Liu, H.; Lu, C.; Li, C.; Zhang, H. DMFold: A novel method to predict RNA secondary structure with pseudoknots based on deep learning and improved base pair maximization principle. *Front. Genet.* **2019**, *10*, 143.
61. Mao, K.; Wang, J.; Xiao, Y. Prediction of RNA secondary structure with pseudoknots using coupled deep neural networks. *Biophys. Rep.* **2020**, *6*, 146–154.
62. Mao, K.; Wang, J.; Xiao, Y. Length-dependent deep learning model for RNA secondary structure prediction. *Molecules* **2022**, *27*, 1030.
63. Sato, K.; Akiyama, M.; Sakakibara, Y. RNA secondary structure prediction using deep learning with thermodynamic integration. *Nat. Commun.* **2021**, *12*, 941.
64. Rouse, W.B.; Andrews, R.J.; Booher, N.J.; Wang, J.; Woodman, M.E.; Dow, E.R.; Jessop, T.C.; Moss, W.N. Prediction and analysis of functional RNA structures within the integrative genomics viewer. *NAR Genom. Bioinforma.* **2022**, *4*, lqab127.
65. Rangan, R.; Zheludev, I.N.; Hagey, R.J.; Pham, E.A.; Wayment-Steele, H.K.; Glenn, J.S.; Das, R. RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: A first look. *Rna* **2020**, *26*, 937–959.
66. Baranovskaya, I.; Sergeeva, M.; Fadeev, A.; Kadirova, R.; Ivanova, A.; Ramsay, E.; Vasin, A. Changes in RNA secondary structure affect NS1 protein expression during early stage influenza virus infection. *Virol. J.* **2019**, *16*, 162–168.
67. Andrews, R.J.; O'Leary, C.A.; Moss, W.N. A survey of RNA secondary structural propensity encoded within human herpesvirus genomes: Global comparisons and local motifs. *PeerJ* **2020**, *8*, e9882.
68. Andrews, R.J.; Peterson, J.M.; Haniff, H.S.; Chen, J.; Williams, C.; Grefe, M.; Disney, M.D.; Moss, W.N. An in silico map of the SARS-CoV-2 RNA structurome. *bioRxiv* **2020**, *3*, 1–18.
69. Desirò, D.; Hölzer, M.; Ibrahim, B.; Marz, M. SilentMutations (SIM): A tool for analyzing long-range RNA–RNA interactions in viral genomes and structured RNAs. *Virus Res.* **2019**, *260*, 135–141.
70. Smyth, R.P.; Negroni, M.; Lever, A.M.; Mak, J.; Kenyon, J.C. RNA structure—A neglected puppet master for the evolution of virus and host immunity. *Front. Immunol.* **2018**, *9*, 2097.
71. Sztuba-Solinska, J.; Teramoto, T.; Rausch, J.W.; Shapiro, B.A.; Padmanabhan, R.; Le Grice, S.F.J. Structural complexity of Dengue virus untranslated regions: Cis-acting RNA motifs and pseudoknot interactions modulating functionality of the viral genome. *Nucleic Acids Res.* **2013**, *41*, 5075–5089.
72. Luwanski, K.; Hlushchenko, V.; Popenda, M.; Zok, T.; Sarzynska, J.; Martsich, D.; Szachniuk, M.; Antczak, M. RNAspider: A webserver to analyze entanglements in RNA 3D structures. *Nucleic Acids Res.* **2022**, *50*, W663–W669.
73. Hu, T.; Chitnis, N.; Monos, D.; Dinh, A. Next-generation sequencing technologies: An overview. *Hum. Immunol.* **2021**, *82*, 801–811.
74. Rasmussen, T.B.; Boniotti, M.B.; Papetti, A.; Grasland, B.; Frossard, J.P.; Dastjerdi, A.; Hulst, M.; Hanke, D.; Pohlmann, A.; Blome, S.; et al. Full-length genome sequences of porcine epidemic diarrhoea virus strain CV777; Use of NGS to analyse genomic and sub-genomic RNAs. *PLoS ONE* **2018**, *13*, e0193682.
75. Dadonaite, B.; Gilbertson, B.; Knight, M.L.; Trifkovic, S.; Rockman, S.; Laederach, A.; Brown, L.E.; Fodor, E.; Bauer, D.L.V. The structure of the influenza A virus genome. *Nat. Microbiol.* **2019**, *4*, 1781–1789.
76. Selvaraj, C.; Dinesh, D.C.; Panwar, U.; Abhirami, R.; Boura, E.; Singh, S.K. Structure-based virtual screening and molecular dynamics simulation of SARS-CoV-2 Guanine-N7 methyltransferase (nsp14) for identifying antiviral inhibitors against COVID-19. *J. Biomol. Struct. Dyn.* **2021**, *39*, 4582–4593.
77. Jin, Z.; Wang, Y.; Yu, X.F.; Tan, Q.Q.; Liang, S.S.; Li, T.; Zhang, H.; Shaw, P.C.; Wang, J.; Hu, C. Structure-based virtual screening of influenza virus RNA polymerase inhibitors from natural compounds: Molecular dynamics simulation and MM-GBSA calculation. *Comput. Biol. Chem.* **2020**, *85*, 107241.
78. Anusuya, S.; Velmurugan, D.; Gromiha, M.M. Identification of dengue viral RNA-dependent RNA polymerase inhibitor using computational fragment-based approaches and molecular dynamics study. *J. Biomol. Struct. Dyn.* **2016**, *34*, 1512–1532.
79. Cihlova, B.; Huskova, A.; Böserle, J.; Nencka, R.; Boura, E.; Silhan, J. High-throughput fluorescent assay for inhibitor screening of proteases from RNA viruses. *Molecules* **2021**, *26*, 3792.
80. Olubiy, O.O.; Olagunju, M.; Keutmann, M.; Loschwitz, J.; Strodel, B. High throughput virtual screening to discover inhibitors of the main protease of the coronavirus SARS-CoV-2. *Molecules* **2020**, *25*, 3193.
81. Shen, L.; Niu, J.; Wang, C.; Huang, B.; Wang, W.; Zhu, N.; Deng, Y.; Wang, H.; Ye, F.; Cen, S.; et al. High-throughput screening and identification of potent broad-spectrum inhibitors of coronaviruses. *J. Virol.* **2019**, *93*, e00023-19.

82. Park, S.J.; Kim, J.; Kang, S.; Cha, H.J.; Shin, H.; Park, J.; Jang, Y.S.; Woo, J.S.; Won, C.; Min, D.H. Discovery of direct-acting antiviral agents with a graphene-based fluorescent nanosensor. *Sci. Adv.* **2020**, *6*, eaaz8201.
83. Khataniar, A.; Pathak, U.; Rajkhowa, S.; Jha, A.N. A comprehensive review of drug repurposing strategies against known drug targets of COVID-19. *Covid* **2022**, *2*, 148–167.
84. Du, R.; Cheng, H.; Cui, Q.; Peet, N.P.; Gaisina, I.N.; Rong, L. Identification of a novel inhibitor targeting influenza A virus group 2 hemagglutinins. *Antiviral Res.* **2021**, *186*, 105013.
85. Zhao, Y.; Du, X.; Duan, Y.; Pan, X.; Sun, Y.; You, T.; Han, L.; Jin, Z.; Shang, W.; Yu, J.; et al. High-throughput screening identifies established drugs as SARS-CoV-2 PLpro inhibitors. *Protein Cell* **2021**, *12*, 877–888.
86. Manikyam, H.K.; Joshi, S.K. Whole genome analysis and targeted drug discovery using computational methods and high throughput screening tools for emerged novel coronavirus (2019-nCoV). *J. Pharm. Drug Res.* **2020**, *3*, 341–361.
87. Zhang, J.W.; Wang, H.; Liu, J.; Ma, L.; Hua, R.H.; Bu, Z.G. Generation of a stable GFP-reporter Zika virus system for high-throughput screening of Zika virus inhibitors. *Virol. Sin.* **2021**, *36*, 476–489.
88. Fernandes, R.S.; de Godoy, A.S.; Santos, I.A.; Noske, G.D.; de Oliveira, K.I.Z.; Gawriljuk, V.O.; Gomes Jardim, A.C.; Oliva, G. Discovery of an imidazonaphthyridine and a riminophenazine as potent anti-Zika virus agents through a replicon-based high-throughput screening. *Virus Res.* **2021**, *299*, 198388.
89. Soszynska-Jozwiak, M.; Pszczola, M.; Piasecka, J.; Peterson, J.M.; Moss, W.N.; Taras-Goslinska, K.; Kierzek, R.; Kierzek, E. Universal and strain specific structure features of segment 8 genomic RNA of influenza A virus-application of 4-thiouridine photocrosslinking. *J. Biol. Chem.* **2021**, *297*, 101245.
90. Wacker, A.; Weigand, J.E.; Akabayov, S.R.; Altincekic, N.; Bains, J.K.; Banijamali, E.; Binas, O.; Castillo-Martinez, J.; Cetiner, E.; Ceylan, B.; et al. Secondary structure determination of conserved SARS-CoV-2 RNA elements by NMR spectroscopy. *Nucleic Acids Res.* **2020**, *48*, 12415–12435.
91. Smola, M.J.; Weeks, K.M. In-cell RNA structure probing with SHAPE-MaP. *Nat. Protoc.* **2018**, *13*, 1181–1195.
92. Madden, E.A.; Plante, K.S.; Morrison, C.R.; Kutchko, K.M.; Sanders, W.; Long, K.M.; Taft-Benz, S.; Cruz Cisneros, M.C.; White, A.M.; Sarkar, S.; et al. Using SHAPE-MaP to model RNA secondary structure and identify 3′UTR variation in Chikungunya virus. *J. Virol.* **2020**, *94*, e00701–e00720.
93. Tomezsko, P.J.; Corbin, V.D.A.; Gupta, P.; Swaminathan, H.; Glasgow, M.; Persad, S.; Edwards, M.D.; Mcintosh, L.; Papenfuss, A.T.; Emery, A.; et al. Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature* **2020**, *582*, 438–442.
94. Yang, S.L.; DeFalco, L.; Anderson, D.E.; Zhang, Y.; Aw, J.G.A.; Lim, S.Y.; Lim, X.N.; Tan, K.Y.; Zhang, T.; Chawla, T.; et al. Comprehensive mapping of SARS-CoV-2 interactions in vivo reveals functional virus-host interactions. *Nat. Commun.* **2021**, *12*, 5113.
95. Dethoff, E.A.; Boerneke, M.A.; Gokhale, N.S.; Muhire, B.M.; Martin, D.P.; Sacco, M.T.; McFadden, M.J.; Weinstein, J.B.; Messer, W.B.; Horner, S.M.; et al. Pervasive tertiary structure in the dengue virus RNA genome. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 11513–11518.
96. Luque, D.; Castón, J.R. Cryo-electron microscopy for the study of virus assembly. *Nat. Chem. Biol.* **2020**, *16*, 231–239.
97. Shepherd, D.C.; Dalvi, S.; Ghosal, D. From cells to atoms: Cryo-EM as an essential tool to investigate pathogen biology, host–pathogen interaction, and drug discovery. *Mol. Microbiol.* **2022**, *117*, 610–617.
98. Zhang, K.; Zheludev, I.N.; Hagey, R.J.; Haslecker, R.; Hou, Y.J.; Kretsch, R.; Pintilie, G.D.; Rangan, R.; Kladwang, W.; Li, S.; et al. Cryo-EM and antisense targeting of the 28-kDa frameshift stimulation element from the SARS-CoV-2 RNA genome. *Nat. Struct. Mol. Biol.* **2021**, *28*, 747–754.
99. Sugita, Y.; Matsunami, H.; Kawaoka, Y.; Noda, T.; Wolf, M. Cryo-EM structure of the Ebola virus nucleoprotein–RNA complex at 3.6 Å resolution. *Nature* **2018**, *563*, 137–140.
100. Shan, H.; Su, X.; Li, T.; Qin, Y.; Zhang, N.; Yang, L.; Ma, L.; Bai, Y.; Qi, L.; Liu, Y.; et al. Structural plasticity of mumps virus nucleocapsids with cryo-EM structures. *Commun. Biol.* **2021**, *4*, 833.
101. Michalak, P.; Soszynska-Jozwiak, M.; Biala, E.; Moss, W.N.; Kesy, J.; Szutkowska, B.; Lenartowicz, E.; Kierzek, R.; Kierzek, E. Secondary structure of the segment 5 genomic RNA of influenza A virus and its application for designing antisense oligonucleotides. *Sci. Rep.* **2019**, *9*, 3801.
102. Hagey, R.J.; Elazar, M.; Pham, E.A.; Tian, S.; Ben-Avi, L.; Bernardin-Souibgui, C.; Yee, M.F.; Moreira, F.R.; Rabinovitch, M.V.; Meganck, R.M.; et al. Programmable antivirals targeting critical conserved viral RNA secondary structures from influenza A virus and SARS-CoV-2. *Nat. Med.* **2022**, *28*, 1944–1955.
103. Diamond, M.S.; Pierson, T.C. The challenges of vaccine development against a new virus during a pandemic. *Cell Host Microbe* **2020**, *27*, 699–703.

RESEARCH ARTICLE

The interaction of RNA G-quadruplexes from the influenza A virus vRNA with TMPyP4 and BRACO-19 ligands

Maria Nalewaj[1], Joanna Sliwiak[1], Karolina Zielinska[1], Pawel Zmora[1], Dorota Niedziałek[2], Grzegorz Wieczorek[3], Elzbieta Kierzek[1], Marta Szabat[1]*

1 Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego, Poznan, Poland, **2** Ensemble3 sp. z o.o., Wólczyńska 133, 01-919 Warsaw, Warsaw, Poland, **3** Mossakowski Medical Research Institute, Polish Academy of Sciences, A. Pawińskiego 5, 02-106 Warsaw, Warsaw, Poland

* szabat@ibch.poznan.pl

Abstract

The influenza virus is an interesting research subject due to its serious threat to global public health. To date, various structural motifs from the influenza A virus (IAV) genome have been studied. Recently, RNA G-quadruplexes (G4s), noncanonical structures formed within the G-rich sequences of the IAV genome, have been reported. These motifs are suggested to be promising antiviral targets, and studying the G4 binding ligands has attracted increasing research interest. We hypothesized that RNA G4s can play a crucial role in IAV replication. This study focused on the interactions between RNA G4s and ligands, which have not been extensively studied in the influenza A virus California/4/2009 (H1N1) to date. Herein, commonly used G4-specific ligands, TMPyP4 and BRACO-19, were selected. First, we performed a reverse transcription stop assay to study the effect of both ligands on the inhibition of cDNA synthesis. Our results showed that both compounds inhibited this process in all wild-type G4 variants, with one variant exhibiting the most noticeable effect after the addition of TMPyP4. We also examined the binding affinity of TMPyP4 and BRACO-19 to IAV RNA G4s using isothermal titration calorimetry, circular dichroism, and fluorescence spectroscopy. Some differences in the binding properties of the three selected G4s were found. Furthermore, UV melting analysis was conducted to evaluate the effect of the ligands on the thermal stability of RNA G4s. To supplement our experimental approaches, we applied, in the limited range, molecular modeling to simulate the folding of 1Q G-quadruplex and provide further insights into its structural stability and topology. Finally, the influence of TMPyP4 on the IAV minireplicon activity was investigated, revealing significant inhibition of IAV replication. Overall, interactions between TMPyP4 and BRACO-19 with IAV G4s were demonstrated for the first time, suggesting that G4s can be potential anti-influenza drug targets.

OPEN ACCESS

Citation: Nalewaj M, Sliwiak J, Zielinska K, Zmora P, Niedziałek D, Wieczorek G, et al. (2025) The interaction of RNA G-quadruplexes from the influenza A virus vRNA with TMPyP4 and BRACO-19 ligands. PLoS One 20(11): e0335975. https://doi.org/10.1371/journal.pone.0335975

Editor: Hari S. Misra, NMIMS Deemed to be University - Mumbai Campus: NMIMS, INDIA

Received: May 6, 2025

Accepted: October 17, 2025

Published: November 19, 2025

Data availability statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: M.S. Grant number: UMO-2019/35/D/NZ6/01479 National Science Centre Poland M.S. Grant number: UMO-2023/49/B/

ST4/03763 National Science Centre Poland https://www.ncn.gov.pl/en No, the work funder didn't play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

1 Introduction

Influenza, commonly known as the flu, is a highly contagious disease that occurs seasonally as epidemics and occasionally as pandemics. The pandemic potential of the influenza virus is a result of its high genome variability and is correlated with its pathogenicity, replication, and growth kinetics. It is known that the secondary and tertiary structures of viral RNA control the above processes. Different RNA structural motifs are formed and used at the various stages of the viral replication cycle, highlighting their fundamental role in influenza virus biology.

Interestingly, the secondary structure of viral RNA (vRNA), including influenza A virus (IAV), is highly conserved across different viral strains, which suggests its importance during RNA replication. Furthermore, it has been shown that G-rich regions are present within the genomes of many viruses, for example, SARS-CoV-2 or HIV-1 [1,2]. These RNA regions can fold into noncanonical structures called G-quadruplexes (G4s) that are stabilized by Hoogsteen hydrogen bonds. It is assumed that they can have various important biological functions [2–5]. Therefore, understanding the vRNA structure-function relationship is essential, and research on this topic should be continued.

Recently, investigations concerning G-quadruplex binders as effective inhibitors of virus infections were published. For instance, Zou et al. used five G4 ligands, Ber, BRACO-19, 360A, PDS, and NiL, to study their interactions with the ZIKV RNA G4 structure [4]. The results revealed that PDS shows a high G4 binding affinity and thermal stability, as well as higher anti-ZIKV activity compared to the other G4 ligands. (4) Another report concerning G-quadruplexes as therapeutic targets was published by Lv and colleagues in 2022 [5]. The authors confirmed G4 structure formation within the chikungunya virus (CHIKV) genome using different biophysical techniques. Importantly, they demonstrated that BRACO-19 and TMPyP4 ligands inhibit CHIKV genome replication, showing that these compounds inhibit the production of infectious virions [5].

To date, a significant number of ligands binding G-quadruplexes have been studied [6–8]. Some are considered potential therapeutic agents [1,4,9]. Progress in the design and synthesis of new G4-interacting compounds allows for testing them as antiviral agents. Several such ligands have been identified and tested *in vitro* for their ability to inhibit viral replication [10–13]. Importantly, the biological relevance and *in vivo* therapeutic potential of these G4-targeting compounds remain areas of considerable interest.

Lately, we identified G-rich regions within the IAV A/California/4/2009 (H1N1) genome and examined their propensity to fold into G-quadruplexes using various biophysical methods [11]. Additionally, we showed that they are present within vRNA segments encoding polymerase complex proteins, indicating their possible role in the virus biology. Previously, Mergny and co-workers investigated the occurrence, localization, and variation of potential G-quadruplex-forming sequences in 77 H1N1 influenza genomes and provided a rational basis for targeting PQS motifs as a potential strategy against influenza [14].

In the current work, the primary goal was to investigate whether TMPyP4 and BRACO-19, well-characterized G4 ligands, could bind to the IAV RNA G-quadruplexes. Moreover, these compounds were used to study the differences in binding patterns of the selected G4-forming sequences and their mutated variants.

Here, we assessed whether G4-specific ligands influence the formation of vRNA G-quadruplexes using the reverse transcription stop assay. Next, we studied the interaction between the viral G4s and the selected ligands by isothermal titration calorimetry (ITC), circular dichroism (CD), and fluorescence spectroscopy. Additionally, the effect of ligand addition on the thermal stability of IAV RNA G4s was examined using UV melting analysis. To supplement our biophysical approaches, we applied, in the limited range, molecular modelling to simulate the folding of 1Q G-quadruplex and get more insights into its structural stability and topology. Finally, we investigated the effect of specific G4 ligands on virus replication using a minireplicon system.

To summarize, we report studies on the interaction between RNA G-quadruplexes from the IAV pandemic strain, A/California/4/2009, and TMPyP4 and BRACO-19 ligands. By analyzing the results obtained from various spectroscopic techniques, molecular biology methods, and biological assays, we demonstrated that selected G4-specific compounds differ in their binding properties. Moreover, our investigation confirms that G-quadruplex structures are formed within the influenza A virus genome and provides valuable data for the design of G4 stabilizers targeting viral genomes.

2 Materials and methods

2.1 Oligonucleotides and chemical compounds

Synthetic oligonucleotides were purchased from Genomed (Poland). Table 1 provides specific information about oligomer sequences and applications. The G4-specific ligands: TMPyP4 and BRACO-19 were bought from Sigma Aldrich (Germany).

2.2 Reverse transcription stop assay

A reverse transcription stop assay was performed using SuperScript™ II Reverse Transcriptase (Thermo Fisher Scientific, USA). RNA oligonucleotides (2 pmol) and appropriate primers (2 pmol) were dissolved in a 10 mM potassium phosphate buffer (pH 6.8) containing 50 mM KCl and 0.1 mM EDTA to obtain a final volume of 10 µl. Samples were denatured at 85 °C for 5 minutes and then cooled down to room temperature overnight. After folding, the sample solutions were mixed with ligands (TMPyP4 or BRACO-19) at increasing concentrations (0, 6.25, 12.5, 25, 50, and 100 µM) and incubated at room temperature for 15 min. Next, the reaction mixture without RT enzyme was added to the sample solutions, gently mixed, and incubated at 42°C for 2 min. After that, the SuperScript™ II RT enzyme was added (200 units), mixed by pipetting up and down, and incubated at 42 °C for 50 min. The reaction was inactivated by heating at 70 °C for 15 min. The products of reactions were purified using the RNA Clean & Concentrator kit according to the manufacturer's protocol (Zymo Research, Germany). Finally, electrophoresis of the products with the addition of 8 M urea was performed using 12% polyacrylamide denaturing gels, run at 140 V for 40 min in a cold room using the XCell SureLock Mini-Cell Electrophoresis System (Thermo Fisher Scientific, USA). The resultant gels were visualized using an Amersham Typhoon Biomolecular Imager and analyzed using ImageQuant™ TL 10.2 analysis software. IC_{50} values were calculated as the mean of three replicates

Table 1. Characterization of selected G-rich sequences from the IAV genome.

Sequence name	Sequence (5′-3′)	Segment (Protein)	Segment length (nt)	Sequence location within segment
1Q	CUGGUGGGGCAGCAGCAAAGGGGAG	1 (PB2)	2341	436-460
7Q	GGUAGUGGUCCAUCAAUCGGGUUGAGCUGGGG	2 (PB1)	2341	2097-2128
11Q	GGAUGUAUAUUCUGAAAUGGGAGGCUGG	4 (HA)	1779	807-834

https://doi.org/10.1371/journal.pone.0335975.t001

using the free online tool, https://mythreyaherbal.com/easy-ic50-calculator-for-herbal-drug-and-bioassay-studies/. Graphs were created using OriginPro 2021 software.

2.3 UV melting analysis

The UV melting measurements were recorded on a Jasco V-650 spectrophotometer (Jasco Deutschland GmbH, Pfungstadt, Germany) equipped with a thermoprogrammer. Single-stranded oligonucleotide concentrations were calculated based on both the absorbance measured above 80 °C and the extinction coefficients, which were estimated by the nearest-neighbor model using the Integrated DNA Technologies calculator available on the website https://www.idtdna.com/calc/analyzer (Integrated DNA Technologies, Inc., Coralville, IA, USA). Samples (final concentration of 5 μM) were dissolved in the same buffer used for the RT stop assay and denatured at 90 °C for 5 min, then slowly cooled to room temperature. Next, the pre-folded RNAs were mixed with the tested ligands at a 1:1 ratio (final concentration of 5 μM) and incubated for 30 min at room temperature. Measurements for RNAs alone and in the presence of ligands were performed in triplicate using quartz cuvettes with a 0.1 cm path length and a sample volume of 30 μl. Absorbance versus temperature curves were obtained using the UV melting method at 260 nm and 295 nm wavelengths in a temperature range of 5–90 °C, with a heating rate of 0.2 °C/min. Thermal denaturation curves were analyzed, and the melting temperature (T_m) values at 295 nm were determined as the mean of three replicates using OriginPro 2021 software. Graphs were also created using OriginPro 2021 software.

2.4 Circular dichroism spectroscopy

CD spectra were recorded on a Jasco J-815 spectropolarimeter (Jasco Deutschland GmbH, Pfungstadt, Germany) using 850 μl quartz cuvettes with a 5 mm path length and a sample volume of 600 μl. RNA oligonucleotides were dissolved in the same buffer as used for the RT stop assay to achieve a sample concentration of 6.8 nM. All samples were denatured for 5 min at 90 °C and then slowly cooled to room temperature overnight before data collection. The titration of the pre-folded RNA oligonucleotides was carried out by adding a solution of the tested ligands (TMPyP4 and BRACO-19) in 12 aliquots of 2.4 μl each, within a concentration range from 3.4 nM to 40.8 nM. Measurements were collected at 20 °C in the 210–350 nm wavelength range with a 1 nm data interval. CD measurements of RNA sequences from the PR8 genome were conducted in the same buffer as described above, at a final sample concentration of 5 μM, using 850 μl quartz cuvettes with a 5 mm path length and a sample volume of 600 μl. All samples were denatured at 90 °C for 5 min and then slowly cooled to room temperature overnight before data collection. Measurements were performed at 10 °C in the 220–340 nm wavelength range with a 1 nm data interval. All CD curves were established as an average of three CD measurements. The buffer spectrum was subtracted from the sample spectra. CD spectra were expressed as difference in molar absorption (Δε, in units of $cm^2\ mmol^{-1}$) between right- and left-handed circularly polarized light and normalized for plotting and comparative purposes using the OriginPro 2021 software.

2.5 Fluorescence measurements

Fluorescence spectra were recorded using a JASCO J-815 CD/fluorescence spectropolarimeter (Jasco Deutschland GmbH, Pfungstadt, Germany) using 850 μl quartz cuvettes with a 5 mm path length and a sample volume of 600 μl. RNA oligonucleotides were dissolved in the same buffer used for the RT stop assay to achieve a final sample concentration of 6.8 nM. All samples were denatured for 5 min at 90 °C and then slowly cooled to room temperature overnight before data collection. The titration of the pre-folded RNA oligonucleotides was carried out by adding a solution of the tested ligands (TMPyP4 and BRACO-19) in 12 aliquots of 2.4 μl each, within a concentration range from 3.4 nM to 40.8 nM. Fluorescence spectra were measured in triplicate at 20 °C with a detector sensitivity of 900 V. For the TMPyP4 ligand, measurements were recorded from 600 to 800 nm with a 2 nm step, using an excitation wavelength of 433 nm. For the BRACO-19 ligand, from 500 to 650 nm with a 2 nm step, and using an excitation wavelength of 285 nm. Graphs were created using the OriginPro 2021 software.

2.6 Isothermal titration calorimetry (ITC)

Microcalorimetric measurements were carried out to yield thermodynamic insights into the interaction between IAV RNA G4s, their mutated variants, and two ligands (TMPyP4 and BRACO-19). The ITC experiments were conducted using MicroCal PEAQ-ITC equipment at 25°C in the same buffer as used for the RT stop assay. Titration of the ligand (kept at 350−1200 μM concentration in the syringe) against RNA 1Q, 7Q, and 11Q oligomers in the cell (kept at 8−30 μM concentration determined at 260 nm) was performed. Ligand solutions were injected in 19−38 aliquots of 2 μl until saturation was observed. Raw data were analyzed using MicroCal PEAQ-ITC Analysis Software with a One Set of Sites model (for interactions with BRACO-19) or Two Sets of Sites model (for interactions between TMPyP4 and wild-type of G4s) to obtain thermodynamic parameters, such as binding stoichiometry (N), dissociation constant (K_d), and the changes in the enthalpy (ΔH) and entropy (ΔS). The ITC measurements were performed in duplicate. Data analysis and graphs were prepared using MicroCal PEAQ-ITC software, following the manufacturer's instructions [15].

2.7 Molecular modelling of the 1Q G-quadruplex

The 1Q bimolecular G-quadruplex was constructed using an in-house RNA folding program, Nucleic Acids Bender, which operates based on tools and molecular dynamics implementation from the GROMACS-2025.2 molecular modelling suite [16]. Due to the strong restraints required during G-quadruplex folding, the preparation process for molecular dynamics simulations, aiming to replicate physiological conditions, involved a gradual reduction of restraint strength by approximately 35-fold concurrently with a gradual temperature increase from 0 to 310 K in the constant volume conditions (NVT). The resulting G-quadruplex was then subjected to multi-stage molecular dynamics simulations in constant pressure (NPT). Representative G-quadruplex geometries obtained every 10 ps from approximately 300 ns of the production runs of molecular dynamics simulation were clustered using the Jarvis-Patrick method with default parameters implemented in GROMACS into 22 distinct groups. The trajectories from all MD simulations were visually examined and analysed using VMD 1.9.4a57 software [17].

2.8 Cell line culture

The human embryonic kidney cell line HEK 293T was bought from LGC Standards, Poland. The cell line was maintained in high-glucose Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific, USA) supplemented with 10% fetal bovine serum (FBS; Thermo Fisher Scientific, USA), 100 U/mL penicillin, and 100 μg/mL streptomycin. The cells were incubated at 37 °C in a humidified incubator with 5% CO_2.

2.9 Luminescent cell viability assay

The cytotoxicity of ligands was tested using Cell Titter-Glo® cell viability reagent according to the manufacturer's protocol (Promega, USA). Briefly, HEK 293T cells were seeded at 1 x 10^5 cells/ml in 96-well plates, along with the tested compounds (TMPyP4, TMPyP2, and BRACO-19 ligands). The cells were grown for 48 hours. Cell Titter-Glo® reagent was added to the cell supernatant in a 96-well plate, and after 2 minutes of mixing and 10 minutes of incubation, the solution was transferred to white plates. Finally, the luminescence measurements were conducted using the HIDEX microplate reader.

2.10 Preparation of IAV plasmids and the minireplicon construct

The minireplicon system used in this study consists of pCAGGS (A/IvPR/8/34) plasmids encoding the PB1, PB2, PA, and NP proteins of the influenza A virus (strain A/Puerto Rico/8/1934 H1N1) and the pPolI-LucRT plasmid encoding the firefly luciferase gene under the control of the influenza A segment 8 5´-end promoter (a kind gift from Prof. Stefan Poehlmann, German Primate Center – Leibniz Institute for Primate Research, Goettingen, Germany). To generate the California

2009 minireplicon, the protein-coding regions of pCAGGS plasmids (A/IvPR/8/34) for the PR8 strain were replaced with sequences encoding the same proteins from the influenza A virus strain A/California/07/2009(H1N1). The cloning method used to create these plasmids is called Gibson assembly. For this, we designed primers (see Table 1) in the SnapGene program and used GeneArt™ Gibson Assembly HiFi Master Mix (Invitrogen) for plasmid preparation. The procedure was carried out according to the manufacturer's protocol with minor modifications.

2.11 IAV minireplicon assay

This assay was performed as previously described with minor modifications.(18) HEK 293T cells were seeded in a 24-well plate at a density of 4 x 10^5 cells/ml per well and incubated overnight at 37 °C. Cultured cells were co-transfected with pCAGGS plasmids prepared as described above at various concentrations (California 2009: 150 ng of PB1, PB2, PA, and Luc, and 300 ng of NP; Puerto Rico 8: 50 ng of PB1, PB2, PA, and Luc, and 100 ng of NP) using Lipofectamine™ 2000 reagent at 80% cell confluency according to the manufacturer's protocol (Thermo Fisher Scientific, USA). Plasmids containing GFP (1 μg) or pCMV-Luc Firefly Luciferase mammalian expression reporter vector (pCMV-Luc, 200 ng) (Origene) served as control probes. Next, the TMPyP4 or TMPyP2 compounds at the final concentrations (1.25, 3.125, 6.25 or 3.125, 6.25, 12.5 μM, respectively) were added to the co-transfected cells. After 18 hours of incubation at 37 °C, the cell medium was replaced with fresh DMEM medium containing FBS and antibiotics, and the cells were incubated for an additional 30 hours at 37 °C. Next, the cells were collected into Eppendorf tubes and centrifuged at 3000 rpm for 3 minutes. Then, the cell pellet was resuspended in 250 μl of ONE-Glo™ luciferase reagent (Promega, USA) and incubated at room temperature for 15 minutes to ensure complete cell lysis. Finally, 50 μl of each sample lysate was transferred in triplicate to a 96-well white polystyrene plate, and the luminescence was measured using a HIDEX microplate reader. Mean luciferase expression and standard deviations (SDs) were calculated from luminescence measurements of six separate wells. The data analysis was performed using OriginPro 2021 and GraphPad software.

3 Results

Our previous studies identified 12 potential quadruplex-forming sequences (PQSs) within the IAV A/California/4/2009 (H1N1) genome and determined their propensity to fold into RNA G-quadruplex structures [11]. By combining bioinformatics tools and biophysical methods, we found that three PQS motifs are located within the PB1, PB2, and HA genome segments and can form stable G4s. During the influenza virus replication cycle, complementary RNA (cRNA) is synthesized from vRNA by the RNA-dependent RNA polymerase (RdRp). Additionally, in the cell nucleus, vRNA is transcribed into mRNA, which serves as a template for the translation of viral protein. Therefore, the presence of selected G-quadruplexes within the vRNA encouraged us to investigate their potential role during the viral replication cycle. We hypothesize that G-quadruplex structures form within the IAV genome and can be implicated in viral replication regulation. In this study, we selected three PQS motifs that have been previously validated to adopt G4 structures using various biophysical methods [19]. We assumed that the sequences 1Q, 7Q, and 11Q have the highest propensity to fold into G4s.

Herein, we investigated the formation and structural features of IAV RNA G-quadruplexes upon the addition of G4-specific ligands using various biophysical techniques, including RT stop assay, ITC technique, CD, UV melting, and fluorescence spectroscopy. Additionally, in the limited range, we examined the influence of the TMPyP4 compound on viral replication within the IAV minireplicon system. Fig 1 illustrates the experimental workflow used in our study. The characteristics of G-rich sequences are summarized in Table 1, whereas the oligomer sequences used in this study, along with their applications, are presented in S1 Table (Supplementary Materials).

3.1 G4-specific ligands inhibit the proceeding of reverse transcription by interacting with G4-forming sequences

Overall, a native polyacrylamide gel electrophoresis (PAGE) can provide valuable information about the binding affinity of ligands to G-quadruplex structures and their conformational preferences. Therefore, given the findings that

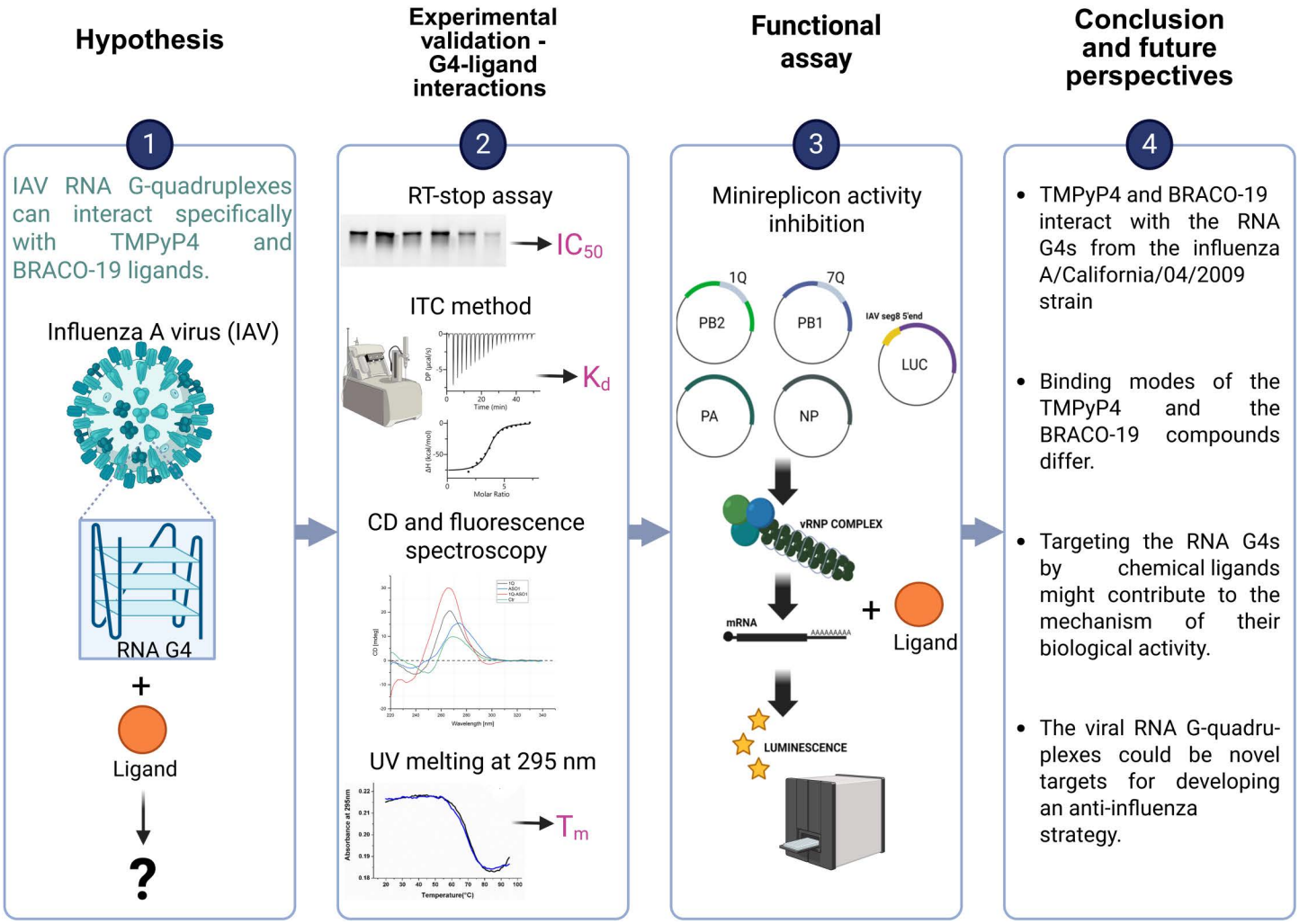


Fig 1. Scheme of the experimental design and workflow of the study.

https://doi.org/10.1371/journal.pone.0335975.g001

G-quadruplexes can induce reverse transcription pausing, [12,13] we employed an RNA-dependent DNA polymerase stop assay, called a reverse transcription (RT) stop assay. In this experiment, we used denaturing polyacrylamide gel electrophoresis to more accurately assess the quantity of the product after the RT reaction.

In the RT stop assay, the enzyme proceeds along the RNA template until it encounters a stable, ligand-interacting RNA G4 structure. Therefore, we assumed that the cDNA synthesis catalyzed by reverse transcriptase would be inhibited after adding a G4-specific ligand to the reaction. Moreover, we postulated that the synthesis of the full-length product would be more significantly arrested with an increasing ligand concentration. However, in the case of G4 variants mutated within the G-rich regions, inhibition of cDNA synthesis was expected to be less noticeable during PAGE analysis. Importantly, before the reverse transcription reaction, all RNA templates were folded in a buffer containing potassium cations required for stable G-quadruplex formation. Then, the G4 variants were incubated with G4-specific chemical compounds, ensuring ligand binding to the structure. In our experiments, we used two commercially available ligands, TMPyP4 and BRACO-19. Additionally, the TMPyP4 analog, TMPyP2, which has a known lower affinity to G-quadruplex structures, was used as a control [18–21].

Based on our results, we reported that the reverse transcription process was hindered by both G4-specific compounds for all wild-type G4 variants (1q, 7q, and 11q). We provided representative gels for three RNA G-quadruplexes with TMPyP4 and BRACO-19 ligands (Figs 2 and 3, respectively). To quantify the effect of ligands on the RT reaction, the IC_{50} values were determined as the average of three replicates. The IC_{50}, the half-maximal inhibitory concentration, is a commonly used parameter to compare drug/ligand potency. The IC_{50} values are presented in Figs 2 and 3. Because the products of RT stop reactions were cleaned as described in the Materials and Methods section, the resultant gels display only the full-length products. The full PAGE images are shown in the Supplementary Materials (S1 and S2 Figs). Moreover, the gels obtained for the G4s with the TMPyP2 compound are presented in Supplementary Materials (S3 Fig).

Overall, the gel analysis revealed some differences in the reverse transcription pausing between wild-type and mutant sequences. According to the resulting gels in Fig 2, cDNA synthesis catalyzed by SuperScript II reverse transcriptase was inhibited in all cases (1q, 7q, and 11q) after the addition of TMPyP4 ligand. We assume that ligand binding to the sequence promotes the formation of a G4 structure. Therefore, there is no significant difference in the band intensities between wild-type and mutated variants without the ligand addition. As expected, we observed that the full-length product synthesis was more significantly inhibited in the wild-type G4s compared to the mutated variants with increasing TMPyP4 concentration (Fig 2). Interestingly, when comparing the results for 1q/1qm versus 7q/7qm, remarkable differences in the level of RT inhibition were displayed after the ligand addition. Specifically, a significant change in band intensity between 1q and 7q variants is already noticeable at 25 μM TMPyP4 concentration (Fig 2A vs. 2B). Similar effects were noted for 11q and 7q, as indicated by the gel electrophoresis patterns in Fig 2B and 2C. On the other hand, product synthesis inhibition observed after 100 μM TMPyP4 addition (also to mutated variants) indicates the nonspecific reaction inhibition unrelated to G-quadruplex structure formation. This suggests that such a TMPyP4 concentration may be too high to be used as evidence of G4 structure-ligand interaction. Furthermore, we revealed different band patterns for the tested compounds by comparing the obtained results for TMPyP4 (Fig 2) and TMPyP2 (S3 Fig). The significant differences in band intensity between wild-type (1q, 7q, and 11q) and mutated variants (1qm, 7qm, and 11qm) with increasing TMPyP2 concentration were not found (S3 Fig). This confirms that this compound is less specific toward RNA G-quadruplex structures.

Our analysis of the IC_{50} parameter shows that the 1q variant with TMPyP4 ligand has the lowest value (28.8 μM, Fig 2A), whereas 7q and 11q have higher values of 55.8 μM and 45.2 μM, respectively (Fig 2B and C, respectively). This suggests that the TMPyP4 ligand can bind more tightly to the 1q motif than to 7q and 11q. Additionally, we noticed that the IC_{50} values calculated for mutated variants, 1qm and 7qm (50.8 μM and 59.7 μM, respectively, Fig 2A and B), are much higher than those for wild-type motifs. This observation may indicate the specificity of the interaction between RNA G4s and the TMPyP4 compound. However, a different effect was seen for 11q compared to 11qm, which could be due to differences in band intensity on the gel (Fig 2C). Furthermore, it is worth mentioning that in a few cases, slightly higher band intensities compared to the control (RNA template without the ligand) may result from differences in product concentration after the purification on the columns. The RT stop assay results indicate the stabilization of G4 after the ligand addition. Moreover, the different patterns of intensity on the gels may result from different stabilities of the RNA G-quadruplexes.

Considering the RT stop assay with the BRACO-19 compound, the resulting gels showed a slightly different band pattern (Fig 3) than after adding TMPyP4. We observed that BRACO-19 can influence RT processing; however, the effect appeared less noticeable based on the gel analyses. Nevertheless, the band intensity for 1q differs from the 1qm variant (Fig 3A), indicating more effective reverse transcription inhibition by BRACO-19 for the wild-type variant (1q) (Fig 3A). Moreover, we noticed less remarkable changes in band intensity for the 7q/7qm variants (Fig 3B). In the case of 11q G4, a similar effect to that observed for 1q G4 was noted, with a noticeable change appearing between 25 and 50 μM BRACO-19 concentration (Fig 3C).

As in the case of TMPyP4, we determined the IC_{50} values for three RNA G4s and the BRACO-19 compound, as shown in Fig 3. Based on the results, the lowest IC_{50} value was obtained for the 1q variant (27.5 μM, Fig 3A). On the other hand, we found that BRACO-19 can bind to 7q and 11q with similar affinity (IC_{50} values = 52.1 μM and 46.9 μM, respectively, Fig

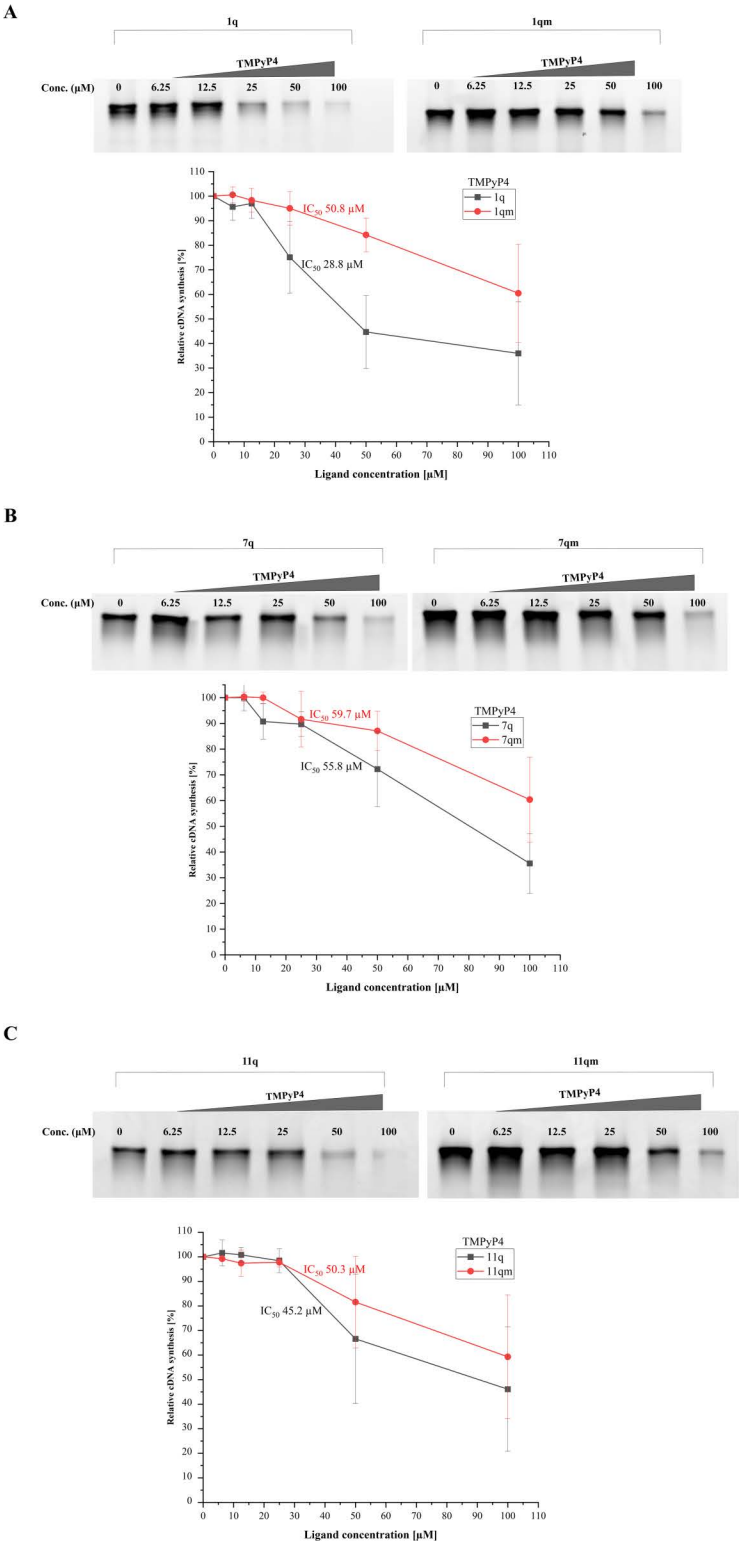


Fig 2. The results of the inhibition of cDNA synthesis catalyzed by the reverse transcriptase in the presence of TMPyP4; the resultant gels (upper) and the plots of band intensity vs. ligand concentration (bottom) for 1q/1qm, 7q/7qm, and 11q/11qm variants (panel A, panel B, and panel C, respectively). The most representative gel results for each sequence are presented here. The IC_{50} values were obtained as the average of three replicates.

https://doi.org/10.1371/journal.pone.0335975.g002

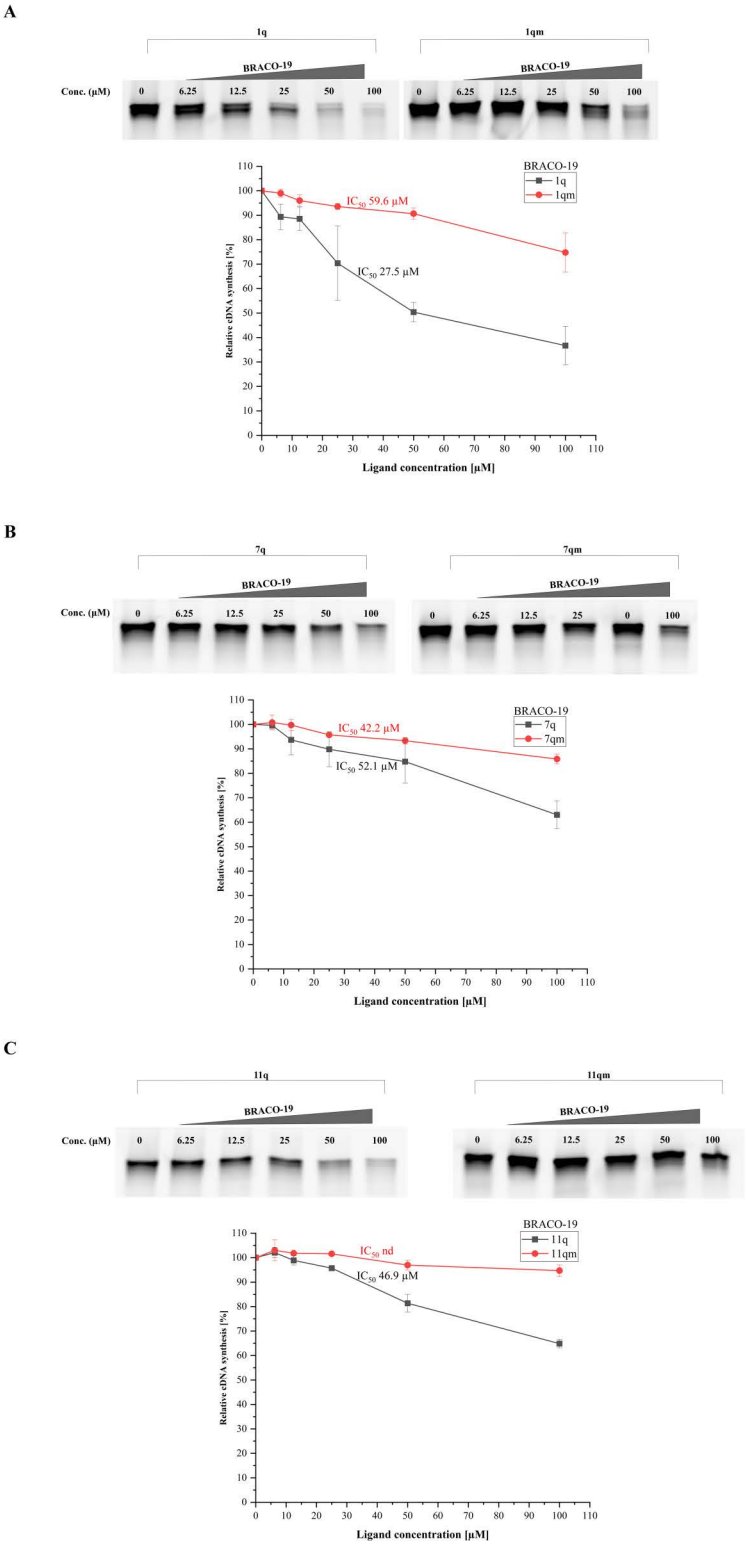


Fig 3. The results of the inhibition of cDNA synthesis catalyzed by the reverse transcriptase in the presence of BRACO-19; the resultant gels (upper) and the plots of band intensity vs. ligand concentration (bottom) for 1q/1qm, 7q/7qm, and 11q/11qm variants (panel A, panel B, and panel C, respectively). The most representative gel results for each sequence are presented here. The IC_{50} values were obtained as the average of three replicates.

https://doi.org/10.1371/journal.pone.0335975.g003

3B and C). We suggest that the BRACO-19 ligand is characterized by a higher binding affinity to RNA G-quadruplexes than the TMPyP4 compound, as indicated by its lower IC_{50} value.

Overall, the RT stop assay results suggest that RNA sequences capable of G4 formation are stabilized after the addition of tested ligands (TMPyP4 and BRACO-19). However, it is worth noting that the above compounds are not very active. The enzyme-pausing effect caused by these ligands required relatively high concentrations (Figs 2 and 3), which may indicate low selectivity of the selected compounds for IAV RNA G4 structures or moderate stabilization of the G-quadruplexes. The TMPyP2 analog did not affect cDNA synthesis (resulting gels shown in S3 Fig in Supplementary Materials). Based on this observation, we can conclude that two chemical compounds, TMPyP4 and BRACO-19, can interact with the IAV RNA G-quadruplexes; however, TMPyP2 has a low affinity toward the G4s.

In summary, reverse transcription of the RNA G4-containing sequence revealed a clear pausing site corresponding to the predicted G4 motif. The addition of a G4-specific ligand caused a dose-dependent increase in RT stalling. Moreover, control experiments using a mutated (non-G4-forming) sequence showed less significant pausing, confirming the specificity of the interaction. However, the differences observed in the gel analysis may result from several reasons, such as G4 structure formation, folding topology, molecularity, or the mechanism of ligand binding. Additionally, methodological and assay-related factors may influence the different band patterns and signal intensities observed on the gels.

3.2 Characterization of the interaction between the IAV RNA G-quadruplexes and ligands by isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) is a method used to determine thermodynamic features of biomolecular interactions. It has been proven useful in studying the folding kinetics, ligand binding specificity, and energetic aspects of G-quadruplex structure-ligand interactions [4,22,23]. Several examples of ITC applications in studies of ligand binding properties to G4s have been reported [22–24]. Hence, our investigations employed the ITC method to examine the interactions between TMPyP4 and BRACO-19 with IAV RNA G4s.

Among all studied RNA G-quadruplexes (1Q, 7Q, and 11Q), we could distinguish two sets of binding sites in the interaction with the TMPyP4 ligand. At the first set, the stoichiometry of TMPyP4 ligand binding could be approximated to ~ 3 (1Q and 11Q) or ~4 (7Q), while the second set of binding sites for the TMPyP4 with a stoichiometry of ~ 2 (Fig 4A, Table 2). TMPyP4 ligand binding is characterized by similar (within the margin of SD) parameters and mode, regardless of the G-quadruplex sequence; however, the 11Q demonstrates the highest affinity to the TMPyP4 (Table 2). The first type of binding site exhibits lower affinity to TMPyP4 with K_D in the mid-nanomolar range (~100−250 nM, Table 2), where the binding of the ligand is driven entirely by the enthalpy. The second type of TMPyP4 binding sites binds the ligand with several hundred-fold higher affinity, and this interaction was mainly driven by the entropy changes (Table 2). For the BRACO-19 ligand, the binding curves exhibited one significant inflection. Since the first inflection is inconspicuous, the model assuming the existence of one type of equivalent binding site was fitted to the obtained data (Fig 4B). Fitting titration curves of 1Q and 11Q resulted in one set of binding sites with similar K_D values of ~162 and 106 nM, respectively, and stoichiometry of ~8 and ~9, respectively (Table 2). In the case of 7Q, the interaction with BRACO-19 seemed to be more than twice as weak as the K_D is ~ 530 nM (Table 2). Interactions of all RNA G-quadruplexes with BRACO-19 were driven by enthalpy, although the entropy contribution was still significant (Table 2). The overall stoichiometry of BRACO-19 binding was almost twice as high as that of the TMPyP4 ligand, which could be explained by the fact that BRACO-19 molecular mass is more than 2-fold less than the molecular mass of TMPyP4 (Table 2).

The G-to-A mutations introduced to the G-quadruplex sequences affect the binding affinity and stoichiometry of the studied ligands. In the case of the TMPyP4 compound, the stoichiometry was reduced to ~3 (1Qm and 11Qm) or ~4 (7Qm), and this effect is limited to the low-affinity set of binding sites. Based on the analysis of the binding curves (Fig 4A), the first inflation completely disappears. We can explain this phenomenon by the fact that the mutation introduction may disrupt the folded RNA G-quadruplex structure, resulting in a loss of some additional TMPyP4 binding clefts. The affinity

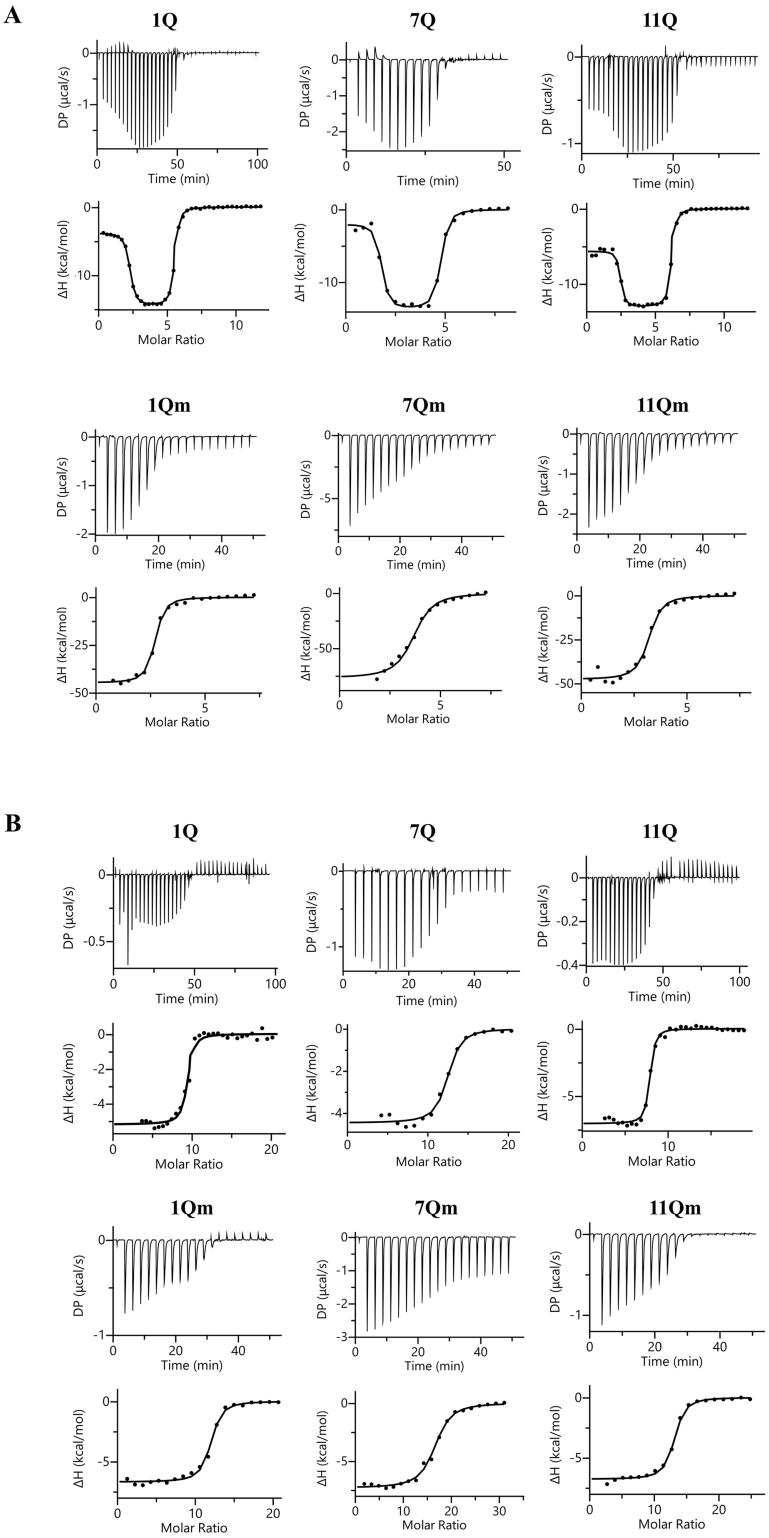


Fig 4. Representative raw data (upper panels) and integration of the peaks (bottom panels) with the best fit of the 'One Set of Sites' or 'Two Sets of Sites' model (only in the case of wild-type sequences and TMPyP4 ligand). The data were obtained after titrations of 1Q, 7Q, and 11Q G-quadruplexes and 1Qm, 7Qm, and 11Qm mutants with TMPyP4 (A) or BRACO-19 (B) ligands.

https://doi.org/10.1371/journal.pone.0335975.g004

Table 2. The binding parameters from the calorimetric titration of the wild-type and mutated variants of 1Q, 7Q, and 11Q with TMPyP4 and BRACO-19 ligands.

	Parameters	1Q	1Qm	7Q	7Qm	11Q	11Qm
TMPyP4	**N_1**	3.3 ± 0.1	2.6 ± 0.1	4.4 ± 2.0	3.7 ± 0.2	3.1 ± 0.2	3.02 ± 0.02
	K_{d1} [nM]	258 ± 84	146 ± 48	276 ± 42	699 ± 144	106 ± 2	217 ± 82
	ΔH_1 [kcal/mol]	−13.0 ± 2.2	−42.1 ± 3.8	−12.8 ± 1.3	−77.0 ± 4.3	−12.7 ± 0.6	−46.7 ± 1.4
	$-T\Delta S_1$ [kcal/mol]	4.0 ± 1.9	32.8 ± 4.0	3.8 ± 1.4	70.0 ± 2.3	3.2 ± 0.6	37.6 ± 1.6
	N_2	1.9 ± 0.4	–	2.5 ± 1.2	–	1.9 ± 0.5	–
		K_{d2} [nM]	2.8 ± 2.3		1.2 ± 0.9		0.3 ± 0.1
		ΔH_2 [kcal/mol]	−2.6 ± 1.6		−2.1 ± 0.2		−4.3 ± 1.9
		$-T\Delta S_2$ [kcal/mol]	−9.1 ± 1.0		−10.2 ± 0.3		−8.6 ± 1.7
BRACO-19	**N**	8.5 ± 1.0	12.0 ± 0.7	10.2 ± 2.6	16.2 ± 0.4	8.7 ± 1.6	12.8 ± 0.4
	K_d [nM]	162 ± 5	540 ± 213	530 ± 33	1580 ± 160	106 ± 8	584 ± 30
	ΔH [kcal/mol]	−4.8 ± 0.6	−6.8 ± 0.2	−4.4 ± 0.1	−7.4 ± 0.1	−6.2 ± 1.1	−6.7 ± 0.4
	-TΔS [kcal/mol]	−4.5 ± 0.6	−1.8 ± 0.4	−4.2 ± 0.1	−0.5 ± 0.1	−3.3 ± 1.1	−1.8 ± 0.4

https://doi.org/10.1371/journal.pone.0335975.t002

of TMPyP4 binding to the "low-affinity" set of binding sites was not affected significantly by the introduced mutations in the case of 1Q and 11Q, whereas in the case of 7Q, it was reduced by half. It is noteworthy that the binding of TMPyP4 to all mutated sequences caused large enthalpy and entropy changes, possibly due to the structurization of the unstructured sequences in the presence of a large TMPyP4 porphyrin ring. Interestingly, the described mutations also reduced the affinity for BRACO-19 by threefold (in 1Qm and 7Qm) and by fivefold (in 11Qm), while simultaneously increasing the binding stoichiometry (Fig 3B, Table 2). Moreover, these new stoichiometry values seem to be directly proportional to the length of sequences, as the N of binding is ~ 12, ~ 16, and ~13 for 1Qm, 7Qm, and 11Qm, respectively, whereas the length of sequences is 25, 32, and 28 nucleotides, respectively. This phenomenon can indicate that about two nucleotides could create the binding site for one BRACO-19 molecule in mutated sequences. The BRACO-19 binding curves for the mutated sequences are also devoid of the first vague inflection, which was observed for wild-type sequences. Mutations in G-rich regions also caused an increase in the enthalpy contribution to the binding and lowered the entropy contribution. Moreover, Libera et al. described the molecular mechanism of the interaction between BRACO-19 and human telomeric G-quadruplex [25]. The authors reported that modifying the topology of G4 may contribute to ligand interactions and binding modes [25].

In general, we noticed the differences in the binding properties between TMPyP4 and BRACO-19 compounds toward wild-type and mutated RNA G4s. It can be concluded that the cationic porphyrin is characterized by a higher binding affinity compared to the acridine derivative. This finding may result from different interaction mechanisms for these ligands, which remain unclear.

3.3 UV melting measurements of the IAV RNA G-quadruplexes

In general, UV spectroscopy is used to study the thermal stability of DNA and RNA structures, including G-quadruplexes, by monitoring changes in their absorbance as a function of temperature. Typically, the G4 structures exhibit a characteristic hypochromic effect at ~ around 295 nm due to the base stacking in the G-tetrads. To gain insight into the thermal stability of the IAV RNA G4s upon ligand binding, we carried out UV melting measurements. Although we performed measurements at two wavelengths: 260 and 295 nm, we chose to analyze the melting profile at 295 nm.

As a result, we were unable to obtain reliable thermodynamic parameters using this method; therefore, we focused on the melting profiles and changes in the melting temperature determined for 1Q, 7Q, and 11Q alone and with the addition of ligands (TMPyP4, BRACO-19, and TMPyP2). The obtained results showed that the melting curves of RNA oligomers

(1Q, 7Q, and 11Q), both in the absence and presence of ligands, were inverted (Fig 5), indicating the formation of G-quadruplexes. This observation suggests that the ligand binding does not disrupt the folding of G4 structures under the experimental conditions. Furthermore, based on the UV melting analysis, we determined the melting temperature values (T_m) for all RNA oligomers (without and with ligand addition), as shown in Fig 5.

Our results revealed only slight differences in the thermal stability of the 1Q, 7Q, and 11Q variants upon the ligand addition (Fig 5). Minor destabilization of the G-quadruplex structure was observed for 1Q in the presence of the TMPyP4 compound. The T_m values of 1Q and 1Q with TMPyP4 were 65.4 and 62.9 °C, respectively (Fig 5). A larger destabilization was noticed in the case of 1Q with BRACO-19, where the presence of this ligand caused a decrease in T_m by 6.8 °C (Figure X). For 1Q with the TMPyP2 compound, reliable results could not be obtained. According to our UV melting analysis for 7Q, the TMPyP4 presence induced slight changes in thermal stability (7Q alone T_m = 54.7 °C vs. 7Q with TMPyP4 T_m = 55.5 °C, Fig 5). A subtle difference was revealed for 7Q after the addition of BRACO-19, where this compound induced a destabilization effect, decreasing its melting temperature by 2.2 °C. Surprisingly, the highest decrease in T_m value was observed in the case of 7Q with TMPyP2, by 6.5 °C, as shown in Fig 5. By analyzing the results obtained for 11Q, we found only minor differences between thermal stability after the addition of the tested ligands. More specifically, the exact change in melting temperature of 11Q was noticed in the presence of BRACO-19 and TMPyP2 compounds; the increase by 0.9 °C was recorded (Fig 5). In the case of 11Q with TMPyP4 ligand, the difference between ligand-bound and ligand-free conditions was negligible (only 0.2 °C in T_m values). We conclude that the thermal stability of 11Q was only minimally affected by these ligands. Moreover, based on our observations, we suggest that TMPyP4, BRACO-19, and TMPyP2 compounds can influence the melting temperature of RNA G4 structures, but only within a limited range under the conditions tested.

3.4 Circular dichroism measurements of the IAV RNA G-quadruplexes upon ligand titration

Circular dichroism (CD) spectroscopy provides information on the geometry of nucleic acid structures and various types of tertiary interactions. It is a simple and relatively fast technique for evaluating the G4 folding topology and assessing its structural features. The unique CD spectral signatures help monitor conformational changes in G4s under experimental conditions. Overall, CD spectroscopy is commonly used to determine the topology of G4s, cation effect, G4/ligand interactions, and ligand-induced stabilization, being a low-resolution complement to high-resolution techniques [5,10,11,18]. Therefore, we applied CD spectroscopy as a complementary technique to check the effect of TMPyP4 and BRACO-19 ligands on the RNA G4 structure.

The CD analysis was conducted for three RNA G4s with increasing concentrations of TMPyP4 and BRACO-19 compounds. As a result, the CD spectra of 1Q, 7Q, and 11Q are presented in Figs 6 and 7. They are dominated by a strong positive band near 265 nm and a smaller negative band at 240 nm, typical for the parallel G-quadruplex structure. In general, the cationic porphyrin under tested conditions did not affect the positions of peaks, indicating that the 1Q, 7Q, and 11Q structures retained the parallel topology (Fig 6). A similar trend was noticed upon the addition of the BRACO-19 compound (Fig 7).

Importantly, we found minimal changes in the CD spectrum intensity that reflect the increase in the concentration of both ligands. However, no global RNA G-quadruplex structure differences were detected (Figs 4 and 5). We found that the G4-specific ligands in the solution slightly reduced the intensity of the CD spectra at 265 nm during the titration for 1Q, 7Q, and 11Q variants. The abovementioned findings suggest that the TMPyP4 and BRACO-19 compounds could not disturb the RNA G-quadruplex structure under experimental conditions.

3.5 Circular dichroism measurements of the RNA sequences from the PR8 genome

The CD spectroscopy was used here to evaluate the structural conformations of RNA sequences from the A/Puerto Rico/8/34 (PR8) genome. We aimed to investigate whether differences in nucleotide composition of G-rich motifs influence

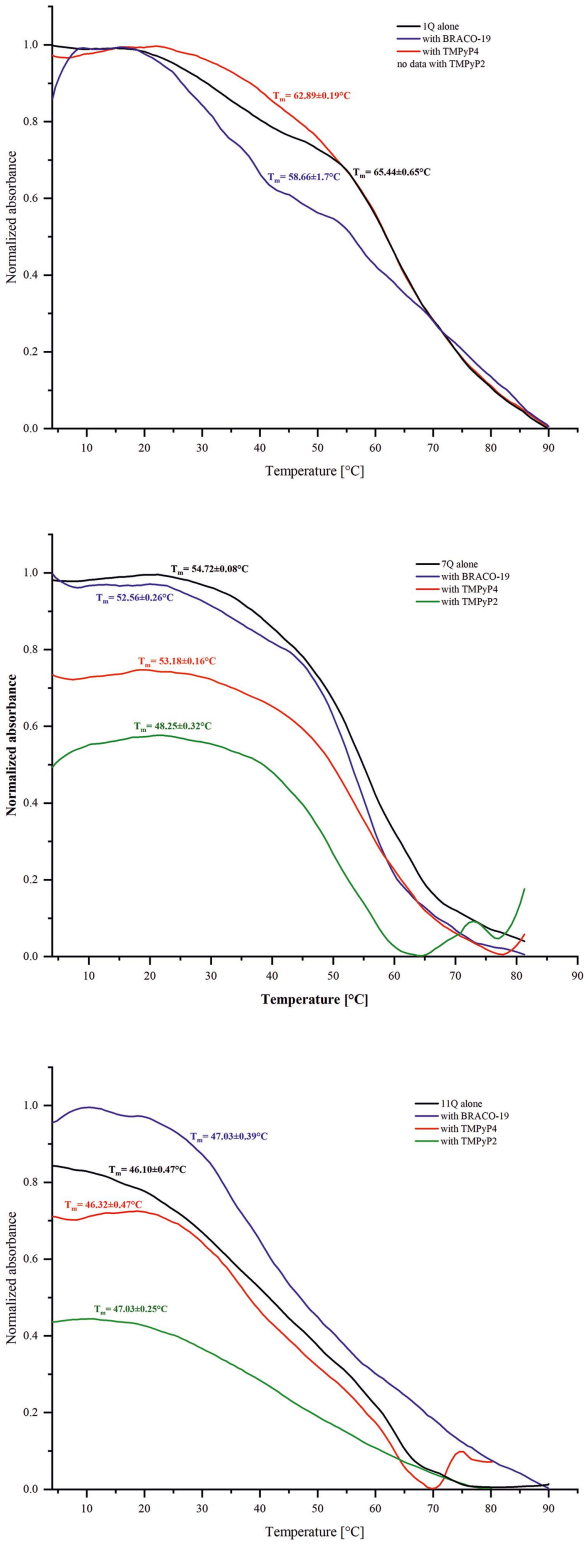


Fig 5. Representative UV melting curves of RNA oligomers in the presence or absence of ligands with T_m values obtained by monitoring the melting profile at 295 nm. The T_m values were obtained as the average of three replicates.

https://doi.org/10.1371/journal.pone.0335975.g005

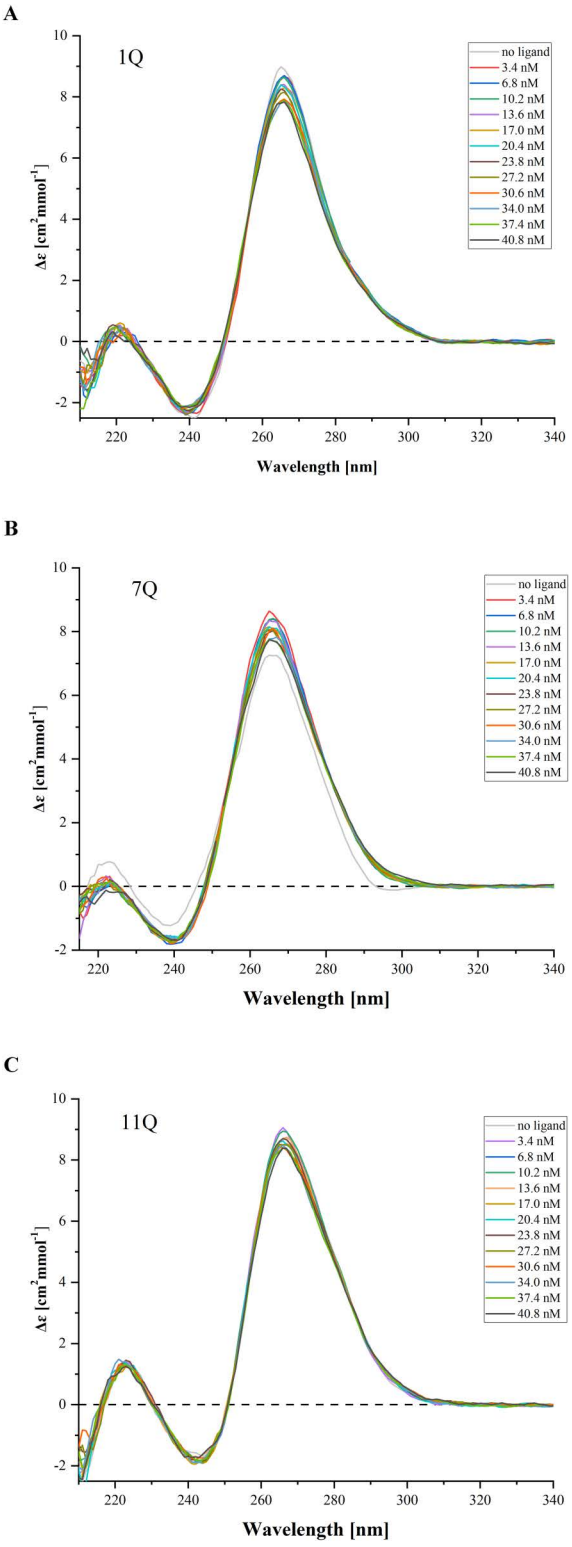


Fig 6. Circular dichroism spectra for pre-folded RNA G4s oligomers: 1Q (A), 7Q (B), and 11Q (C) with TMPyP4 increasing concentrations.

https://doi.org/10.1371/journal.pone.0335975.g006

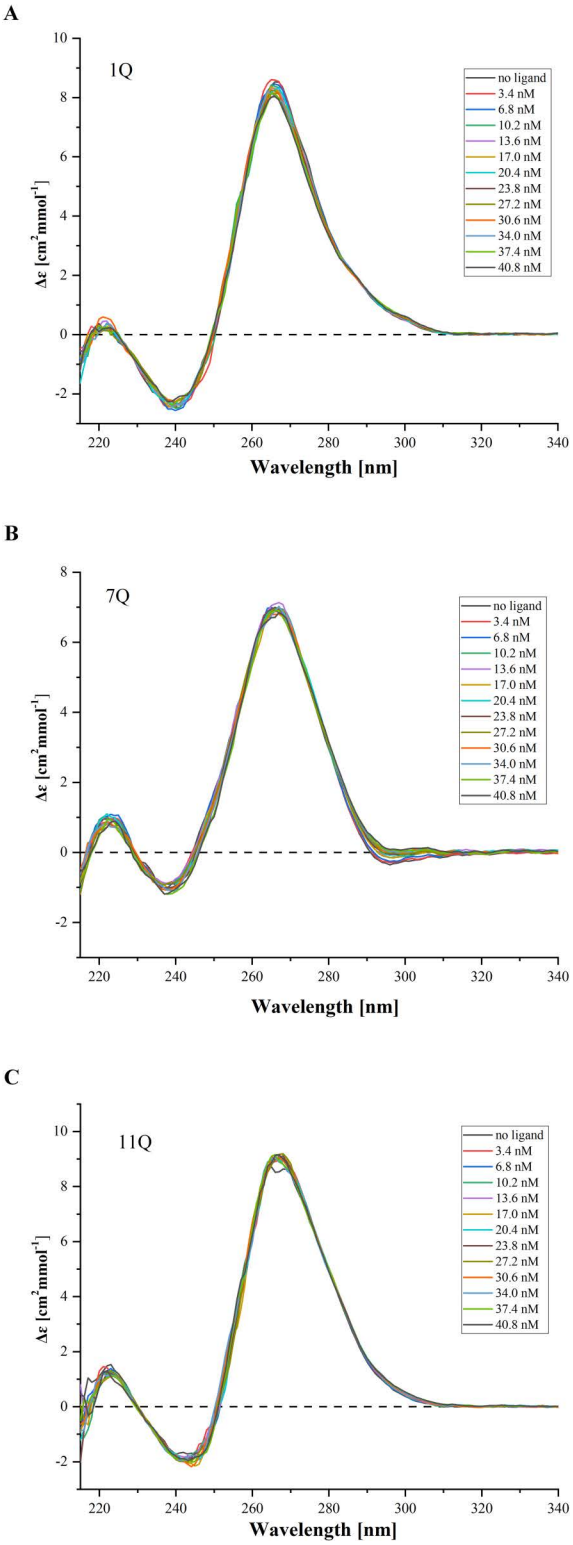


Fig 7. Circular dichroism spectra for pre-folded RNA G4s oligomers: 1Q (A), 7Q (B), and 11Q (C) with BRACO-19 increasing concentrations.

https://doi.org/10.1371/journal.pone.0335975.g007

the geometry of RNA. Fig 8 presents three G-rich sequences from the IAV Cal2009 (1Q, 7Q, and 11Q) and PR8 (1qPR8, 7qPR8, and 11qPR8) genomes with marked mutations in guanine tracts.

We obtained the CD profiles for three RNA oligomers, named 1qPR8, 7qPR8, and 11qPR8, which are shown in Fig 9. In general, the resultant CD profiles for these RNAs have similar features. Namely, the CD spectrum of 1qPR8 displays two negative bands near 235 and 295 nm, and a single maximum at 270 nm (Fig 9). For 7qPR8, we observed a broad positive band between 245 and 285 nm, with a maximum near 265 nm (Fig 9). Additionally, two minima were also found near 235 and 295 nm (7qPR8, Fig 9). In the case of 11qPR8, a single positive peak from 250 to 290 nm with a maximum at 268 nm was noticed (Fig 9). We also observed two negative bands, one with a maximum near 240 nm and the second near 295 nm (Fig 9). It has been reported in the literature that the typical CD profile of parallel G-quadruplex displays a strong, positive band near 265 nm and a negative peak around 245 nm [26–28].

According to our CD analysis, we suggest that 1qPR8, 7qPR8, and 11qPR8 can fold into G4 structures. However, we noted a slight deviation from the expected spectral features. Specifically, in the resulting CD spectra, a negative peak near 295 nm appeared (Fig 9). This result may indicate the existence of higher-order G4 structures or an equilibrium between different conformations, such as A-form RNA or hairpin. Additionally, this finding highlights the structural heterogeneity and conformational flexibility of the analyzed RNAs. Furthermore, these effects can be due to various factors, not only the nucleotide composition of the G-rich sequence but also the RNA concentration used in the CD experiment, or the presence of metal cations in the solution.

Interestingly, we used RNAStructure 6.1 Software for RNA secondary structure prediction for sequences from both Cal2009 (1Q, 7Q, and 11Q) and PR8 (1qPR8, 7qPR8, and 11qPR8) strains. All predicted RNA structures are shown in S4 Fig in the Supplementary Materials. The differences observed in RNA folding patterns confirm the structural diversity among the tested sequences. It can be concluded that the nucleotide composition of G-rich sequences influences the G-quadruplex folding behaviour.

3.6 Fluorescence spectroscopy analysis of the IAV RNA G-quadruplex-ligand interactions

To gain further insight into the interaction between RNA G4s and selected ligands, we applied fluorescence spectroscopy. We expected changes in the fluorescence spectrum and intensity when a selected ligand binds to a G-quadruplex structure. Herein, the fluorescence titration spectra of pre-folded 1Q, 7Q, and 11Q G4s (keeping RNA concentration constant) with increasing concentrations of TMPyP4 and BRACO-19 ligands were recorded. The fluorescence emission spectra for

Sequence 1

1Q	5' CUGGUGGGGCAGCAGCAAAGGGGAG
1qPR8	5' CUGGUGGAGCAGCAGCAAAGGGAAG

Sequence 2

7Q	5' GGUAGUGGUCCAUCAAUCGGGUUGAGCUGGGG
7qPR8	5' GGUAGUGGUCCAUCAAUCGGGUUGAGCUGCGG

Sequence 3

11Q	5' GGAUGUAUAUUCUGAAAUGGGAGGCUGG
11qPR8	5' GGGUGUAUAUUCUGGAAAGGGAGACUGC

Fig 8. The differences in nucleotide composition of G-rich sequences from the Cal2009 and PR8 strains used in our study. Red color indicates the point mutations; the G-tracts are underlined.

https://doi.org/10.1371/journal.pone.0335975.g008

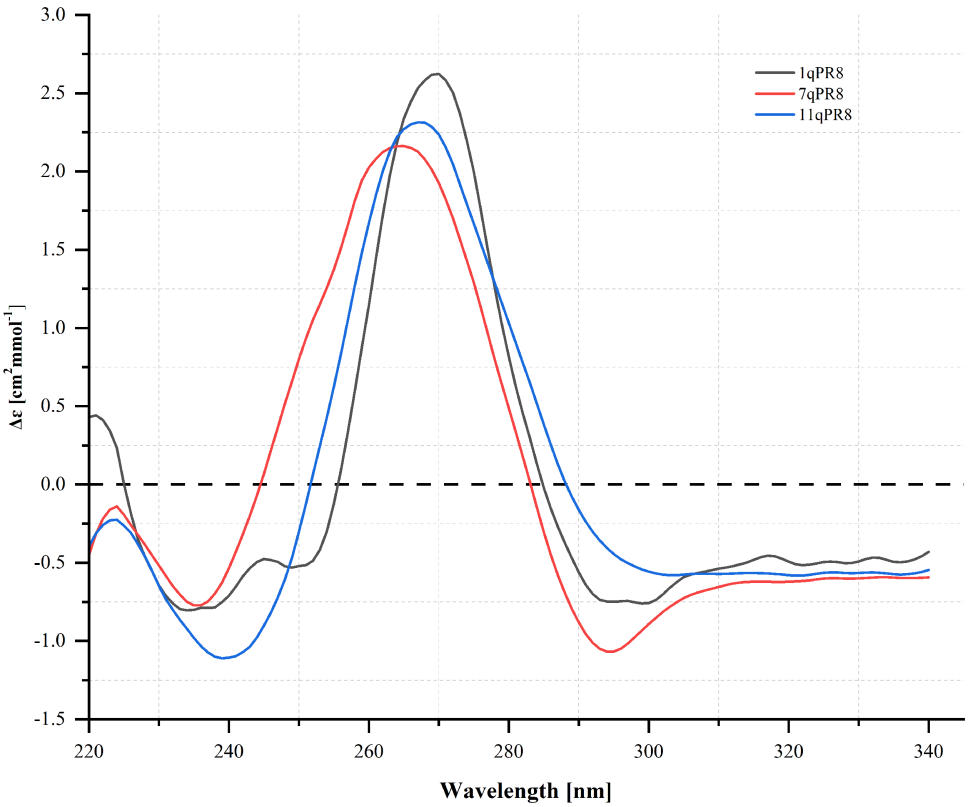


Fig 9. Circular dichroism spectra of 1qPR8, 7qPR8, and 11qPR8 oligomers.

https://doi.org/10.1371/journal.pone.0335975.g009

the TMPyP4 compound, measured at wavelengths ranging from 600 to 800 nm with excitation at 433 nm, are presented in Fig 10. Overall, the measurements were continued until the changes in TMPyP4 fluorescence intensity became negligible. We observed a decrease in the distance between peaks at emission maxima with increasing concentrations of TMPyP4 up to six-fold excess to the RNA G4s. Consequently, the fluorescence spectrum of TMPyP4 alone features a single broad emission peak centered near 700 nm (Fig 10, light green dashed line). In contrast, the fluorescence spectra changed after the addition of RNA G4 oligomers, and two emission maxima at 660 nm and near 730 nm appeared (Fig 10). The same trend was observed in all cases (1Q, 7Q, and 11Q), which suggests the TMPyP4 compound interacts with the G-quadruplexes present in the solution. In other words, different shapes of recorded spectra confirm that RNA G4s can interact with the cationic porphyrin under experimental conditions.

As mentioned earlier, we also investigated the effects of the three G4s on the BRACO-19 fluorescence spectrum. The fluorescence spectra of the BRACO-19 compound with 1Q, 7Q, and 11Q G4s were recorded in the range from 500 to 650 nm with excitation set at 285 nm. Similar to the TMPyP4 ligand, the measurements were conducted until the changes in BRACO-19 fluorescence intensity were insignificantly small. The results for all RNA G4s variants with increasing concentrations of acridine compound are presented in Fig 11. We observed a different trend for BRACO-19 upon the addition of 1Q, 7Q, and 11Q G4s compared to the TMPyP4 ligand. The overall shape of the recorded spectra remained unchanged when RNA G4s were present in the solution. In all cases, the spectral profile of BRACO-19 fluorescence is characterized by a single positive peak with a maximum near 570 nm (Fig 11). However, comparing the fluorescence intensity of the BRACO-19 ligand *vs.* the addition of G-quadruplexes, one can notice its decrease after the addition of RNA G4s (Fig 11). Surprisingly, the most significant change in fluorescence intensity was found in the case of the 11Q variant (BRACO-19

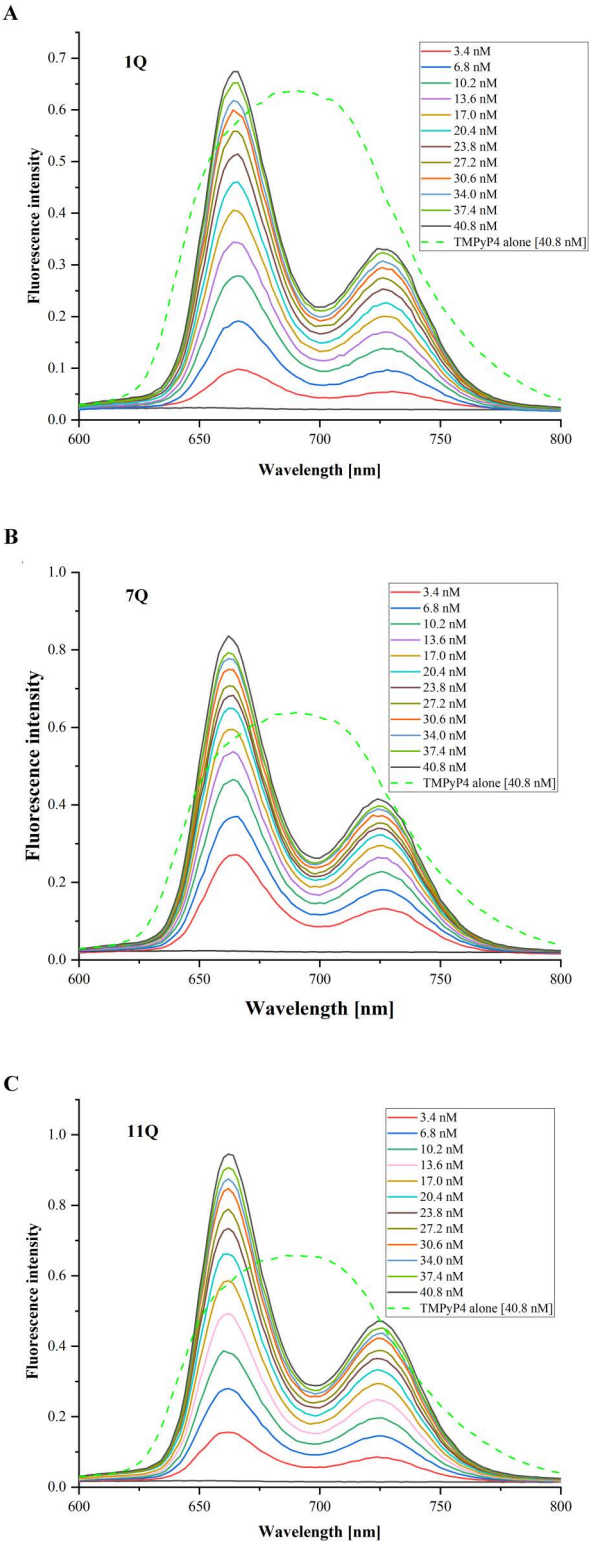


Fig 10. The fluorescence titration spectra of the TMPyP4 ligand in the absence and presence of pre-folded RNA G4s oligomers: 1Q (A), 7Q (B), and 11Q (C). The spectrum of TMPyP4 alone in the buffer is indicated as a light green dashed line.

https://doi.org/10.1371/journal.pone.0335975.g010

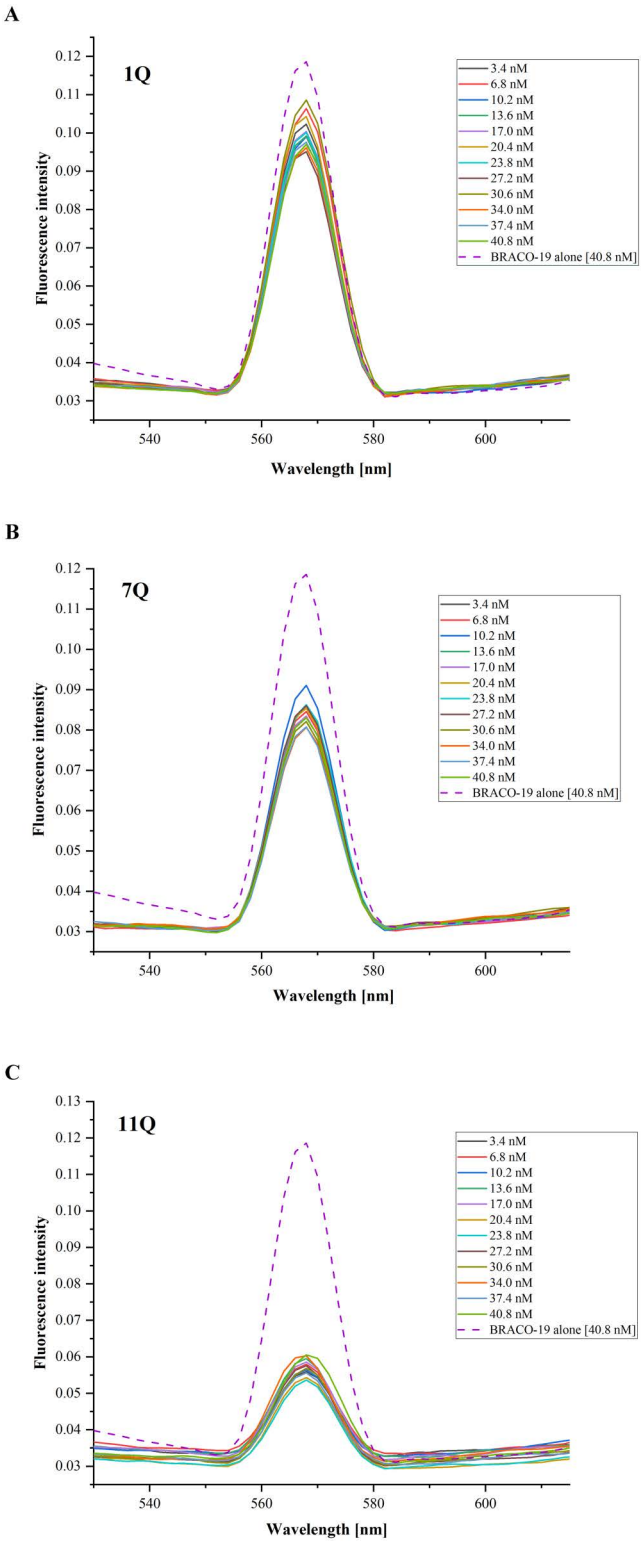


Fig 11. The fluorescence titration spectra of the BRACO-19 ligand in the absence and presence of pre-folded RNA G4s oligomers: 1Q (A), 7Q (B), and 11Q (C). The spectrum of BRACO-19 alone in the buffer is indicated as a purple dashed line.

https://doi.org/10.1371/journal.pone.0335975.g011

alone *vs.* 11Q with 40.8 nM BRACO-19, Fig 11C). In contrast, the smallest was for the 1Q variant (BRACO-19 alone versus 1Q with 40.8 nM BRACO-19, Fig 11A). Additionally, the changes in fluorescence intensity across varying concentrations of BRACO-19 were relatively small (Fig 11). This observation may suggest that the presence of RNA G4s in the solution did not significantly affect the fluorescence of the acridine derivative.

3.7 Modeling of 1Q G-quadruplex

To supplement our biophysical and experimental approaches, we applied molecular modeling to simulate the folding behaviour of 1Q G-quadruplex, providing further insights into its structural stability and topology. Given the 1Q RNA sequence, the program generated its two identical elongated RNA strands with stacking base interactions positioned according to imposed distance restraints to rigidify the initial structure (Fig 12A). Subsequently, guanine (G) quartets, expected to form adjacent pairs via intramolecular Hoogsteen hydrogen bonding, were defined. Following π–π stabilization of the four adjacent stacking guanines by additional distance restraints, RNA molecules were folded according to distance restraints between guanines involved in the intramolecular Hoogsteen interactions, resulting in the formation of two halves of a four-layered bimolecular G-quadruplex (Fig 12B). Next, these folded RNA molecules were brought together using an additional set of distance restraints between guanines involved in the intermolecular Hoogsteen interactions to form a complete, bimolecular, four-tetrad G-quadruplex model (Fig 12C).

Subsequently, three potassium ions were inserted between the stacked guanine tetrads to stabilize the structure. The system was then placed in a simulation box filled with an explicitly represented 150 mM aqueous NaCl solution (Fig 13). The prepared G-quadruplex was subjected to a multiple-step minimization and equilibration protocol designed to gradually remove the imposed restraints. Initially, only the solvent atoms were relaxed, followed by a stepwise loosening of the G-quadruplex loops. In the final stage, all restraints, including those on the core formed by the four guanine tetrads, were fully released, and the system was further optimized. Due to the strong restraints required during G-quadruplex folding, the preparation process for molecular dynamics simulations, aiming to replicate physiological conditions, involved a gradual reduction of restraint strength by approximately 35-fold concurrently with a gradual temperature increase from 0 to 310 K in the constant volume conditions (NVT). The resulting G-quadruplex was then subjected to multi-stage molecular

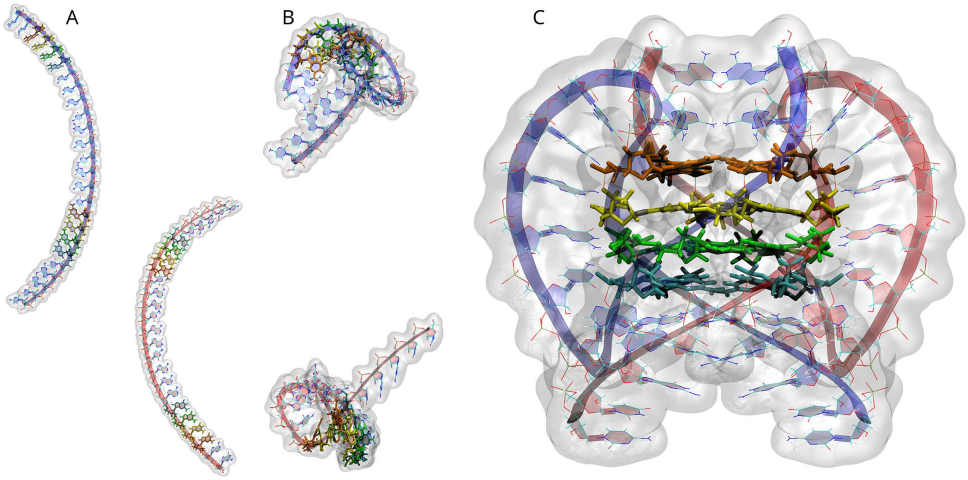


Fig 12. Three main stages of building the 1Q model. A) Generating RNA strands; B) Intramolecular folding; C) Intermolecular folding. The RNA molecules are shown using the new ribbon representation and highlighted with a white, semi-transparent surface. The two strands forming the G-quadruplex are coloured blue and red. The four layers of tetrads are shown in liquorice representation and coloured red, yellow, green, and blue.

https://doi.org/10.1371/journal.pone.0335975.g012

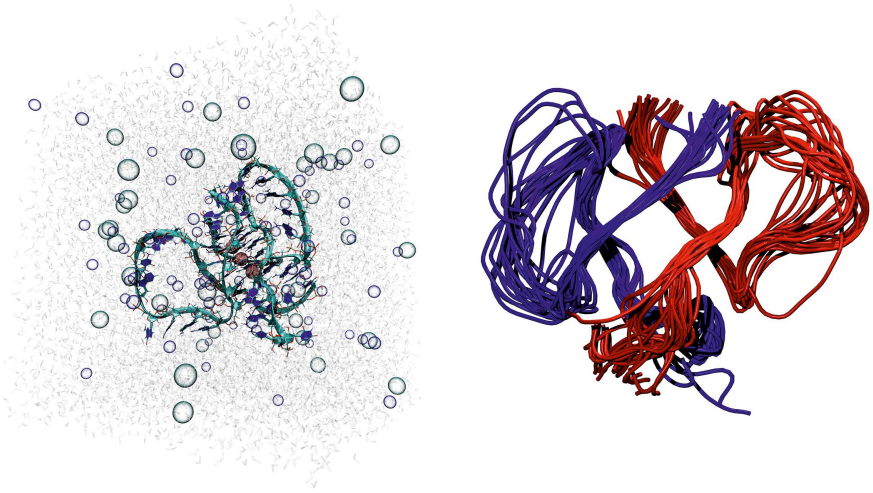


Fig 13. Simulation box with 1Q bimolecular G-quadruplex in the new ribbon representation, explicit water solution with ions (represented as transparent bubbles) during molecular dynamics simulation. The pink spheres are potassium ions in the channel of the core between the stacked G-tetrads (left panel). Representative structures of the 1Q G-quadruplex obtained during 300 ns of molecular dynamics simulations, superimposed using the phosphate groups of the core (consisting of 16 guanines) as the fitting centres. The two strands forming the G-quadruplex are shown in a tube representation of the RNA backbone and are colored blue and red. The core of the molecule remains stable, while the loops do not reach equilibrium during the entire simulation (right panel).

https://doi.org/10.1371/journal.pone.0335975.g013

dynamics simulations in constant pressure (NPT). Representative G-quadruplex geometries obtained every 10 ps from approximately 300 ns of the production runs of molecular dynamics simulation were clustered into 22 distinct groups, as shown in the right panel of Fig 13.

Molecular dynamic simulations indicate that the bimolecular G-quadruplex remains stable, i.e., its four stacked tetrads persistently maintain their structure under physiological conditions, whereas the long loops do not reach equilibrium, as throughout the simulation, no loop adopts the same conformation twice. The results presented in the right panel of Fig 13 suggest that although 1Q RNA strands can form a G-quadruplex that remains stable under physiological conditions, their excessive flexibility, due to the dynamic behaviour of the long loops, may cause significant challenges for crystallization or NMR analysis of this bimolecular RNA system.

3.8 Biological studies of the influence of G4-specific ligand on the IAV minireplicon system activity

Our previous studies have shown that identified G-rich motifs (1Q, 7Q, and 11Q) are located within segments encoding polymerase complex proteins (PB1 and PB2) and within the hemagglutinin (HA) segment [11]. Considering the localization of G4s within the vRNA segments, it can be assumed that these structures may play essential roles in regulating different steps of the viral life cycle. Recently, a non-infectious life cycle modeling system has been developed using reverse genetic methods [29,30]. This minireplicon system (also known as a minigenome) has been used in studies concerning the chemical targeting of a G-quadruplex RNA in the Ebola virus L gene [9]. Therefore, in our research, we aimed to examine the influence of G4-specific ligand binding on luciferase expression using the genomic vRNA analog system (Fig 14). This system relies on a viral mini-genome, which controls the expression of firefly luciferase (Luc). Within this minireplicon system, the expression of the viral proteins (PB2, PB1, PA, and NP) and vRNA-like transcript results in the *in vitro* reconstitution of active vRNP complexes. Those vRNP complexes generate the corresponding mRNA, which controls the expression of firefly luciferase with chemiluminescence activity. The luciferase is encoded under the influenza A virus segment 8 5´end promoter in the plasmid. Thus, its expression is only possible if the vRNP complex is reconstituted correctly.

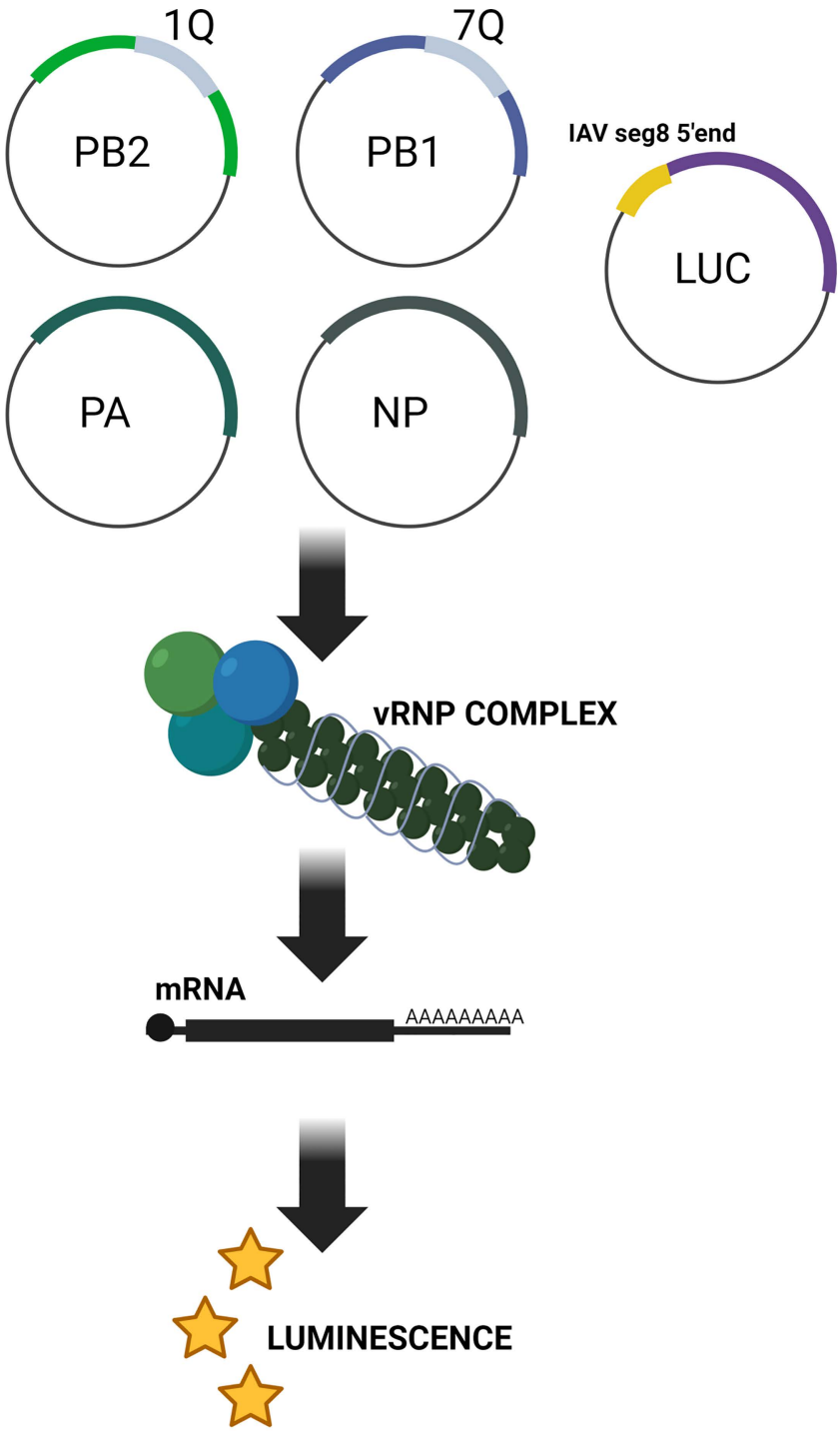


Fig 14. Scheme of the IAV minireplicon system used in this study.

https://doi.org/10.1371/journal.pone.0335975.g014

A clear in our study, we used the IAV minireplicon system, which enables an accurate and rapid test of the activity of the vRNP complex based on its ability to replicate a virus-like RNA encoding the luciferase protein. Moreover, we assumed that adding the TMPyP4 ligand should, to some extent, inhibit minireplicon activity in the cells. Additionally, we decided to use another ligand in biological experiments, the analog TMPyP2, which is reported to have a lower affinity for the G-quadruplex and is therefore often used as a negative control compound in TMPyP4 activity studies [18–21]. Both compounds were confirmed not to have a toxic effect on the selected cell line using a cytotoxicity assay. The results of cytotoxicity of TMPyP4 and TMPyP2 compounds in HEK 293T cell culture are presented in S2 Table in Supplementary Materials.

Notably, based on our previous studies, we found the difference in nucleotide composition of G-rich sequences (1Q and 7Q) between A/California/04/2009 (Cal2009) and A/Puerto Rico/8/34 (PR8) genomes, as indicated earlier in Fig 8 (*3.5 Circular dichroism measurements of the RNA sequences from the PR8 genome section*). We reported that PQS regions from segments 1 and 2 of PR8 differ from the corresponding PQS regions in the Cal2009 genome. The 1Q sequence from PR8 contains A residues instead of G residues within two different G-tracts in the Cal2009 genome. In contrast, the 7Q sequence from PR8 has a C residue instead of a G residue in the G-rich region, compared with the sequence from Cal2009. From this observation, it may be suggested that differences in nucleotide composition can affect the conformational changes of G-quadruplex structures in the PR8 strain genome. This finding motivated us to use the PR8 minireplicon system in our investigation.

By analyzing the luciferase luminescence signals, we conclude that both ligands markedly inhibited minireplicon activity compared to the negative control (TMPyP4 IC_{50} = 4.04 μM; TMPyP2 IC_{50} = 4.32 μM). Additionally, the observed effect appears to be dose-independent, as all selected ligand concentrations caused minireplicon activity inhibition (Fig 15). During the experimental optimization, we decided to select three concentrations for each ligand. We agreed not to use the lowest TMPyP2 concentration (1.25 μM) as no ligand effect was observed after its addition. Conversely, the effect on the minireplicon system after the addition of the highest TMPyP4 concentration (12.5 μM) did not differ from the effect caused by half of that concentration. Moreover, our findings suggest a strain-dependent inhibitory effect of the ligands. In the case of PR8 minireplicon, ligands do not show such a strong repressive effect compared to the negative control, but instead seem to have no significant influence on minireplicon activity (Fig 15). Interestingly, TMPyP2 at a concentration of 6.25 μM has an inhibitory effect on PR8 minireplicon system activity (Fig 15). On the other hand, the TMPyP4 ligand at the lowest concentration appears to have a stimulatory effect on the PR8 minireplicon, as its addition increased the luminescence

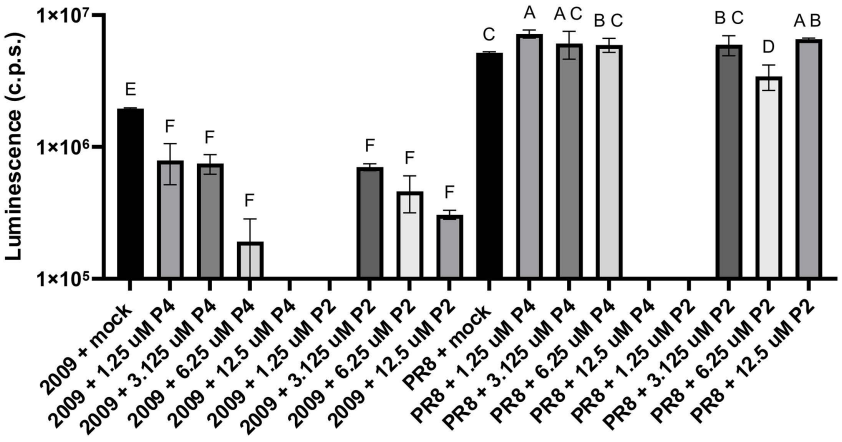


Fig 15. The inhibition of Cal2009 and PR8 minireplicon activity in HEK 293T cells after TMPyP4 or TMPyP2 treatment. All data are presented as the means from two biological replicates, each subjected to three technical replicates. Error bars reflect the SD.

https://doi.org/10.1371/journal.pone.0335975.g015

signal. At this moment, the mechanism of this stimulation is unknown, but we suspect that it may be a nonspecific effect. Further structural and bioinformatic investigations are necessary, starting with RNA sequence analysis, examining its interaction with the ligand, and exploring the possible function of RNA G4s. The fact that we observe an inhibitory effect only in the case of the Cal2009 minireplicon highlights the importance of our selected PQS motifs that were identified in the vRNA of IAV A/California/04/2009 strain and confirmed to contain mutations within the IAV A/Puerto Rico/8/34 vRNA. It is also worth noting that these two minireplicon systems have different activity levels. Even though we used three times higher plasmid concentrations for the Cal2009 minireplicon, the activity of the PR8 minireplicon remained noticeably higher.

Although we assessed binding affinity using the ITC method and examined the effect of ligands on thermal stability and folding of RNA G4s, a clear correlation with the observed inhibitory effects in the cellular environment was not consistently evident across the tested ligands. The relationship between G4 thermal stability, folding dynamics, and ligand accessibility is complex and may not be directly correlated. The minireplicon assay revealed a significant inhibition of IAV replication. However, the structural features of RNA G4/ligand interactions observed in our biophysical studies do not necessarily predict the ligand's effect in a cellular context, such as in the minireplicon system. Several factors may explain this lack of direct correlation. For example, the complex and dynamic nature of RNA G4s in a cellular environment, where folding kinetics, RNA–protein interactions, and structural accessibility can significantly influence ligand binding and functional outcomes. Additionally, ligand specificity and potential off-target effects should also be considered. G4-specific ligands may bind to other RNA or DNA structures or even interact with proteins, affecting replication independently of G4 binding. Therefore, effects observed on minireplicon activity may not strictly reflect G4 stabilization.

4 Discussion

The influenza virus, which causes the common flu, has a high epidemic and pandemic potential, resulting from its high genome variability. It is known that viral biological processes, such as replication and pathogenicity, are controlled by the secondary and tertiary structures of viral RNA. RNA is utilized throughout the replication cycle, highlighting its fundamental role in IAV biology.

The secondary structure of vRNA, including IAV vRNA, is highly conserved among different strains, suggesting its importance for the viral life cycle. Additionally, it has been shown that RNA structural motifs, such as G-quadruplexes, are present within the viral genomes and can have important biological functions [1,31–33]. Therefore, we recommend further research into the correlation between vRNA structure and function. In our previous paper, we identified and described structural aspects of G-quadruplexes from the IAV genome [11].

A significant number of G4-specific ligands have been studied, some of which show potential as antiviral therapeutic agents [34,35]. For example, TMPyP4 and BRACO-19 compounds stabilize G-quadruplex formation and could be inhibitors of viral replication [10,18]. Given the findings that the presence of G-quadruplexes can induce reverse transcription pausing, [12,13] here, we tested both TMPyP4 and BRACO-19 ligands on the RNA G4s from the IAV genome. However, due to the known lack of site specificity of these ligands, we primarily treated them as indicators of the presence and functional relevance of G-quadruplexes in the IAV replication, suggesting that these sequences may serve as targets for anti-viral drug development.

In our current study, we first examined the effect of TMPyP4 and BRACO-19 ligands on the inhibition of cDNA synthesis. The obtained results revealed some differences between the reverse transcription pausing after the addition of both tested compounds. The changes in the intensity patterns for 1q, 7q, and 11q on the gels may result from the different stability of the RNA G-quadruplexes. The relatively weak binding or effect observed in the RT stop assay suggests that viral RNA G4s may have structural features that hinder effective ligand recognition or stabilization, which represents an important finding. Additionally, our UV melting experiments support this, showing that the thermal stability (T_m) of the RNA G-quadruplexes shows only minor changes with ligand addition, with some cases showing a slight decrease, and others minimal increases. These subtle shifts indicate that, under our experimental conditions, the ligands do not strongly

stabilize the RNA G4s. Notably, previous studies on TMPyP4 and BRACO-19 ligands in the context of polymerase pausing have yielded varying results. This reflects differences in G4 folding topology, sequence context, or assay conditions. For example, Chashchina et al. observed that TMPyP4 and BRACO-19 induced minimal or no detectable polymerase pausing in native promoter sequences of selected oncogenes [36]. They noted the pause sites induced by BRACO-19 in the *MYC* promoter at high ligand concentrations. Based on this result, the authors suggest that the tested compounds exhibit low selectivity for G4 structures or provide insignificant stabilization of G4s [36]. Another example is the study by Zhou and co-workers, who investigated a small molecule targeting the RNA G-quadruplex structure in a negative-sense RNA virus [9]. The authors showed that the TMPyP4 inhibits RNA synthesis at the G4 site. The RNA synthesis pausing was gradually promoted at G4 sites upon treatment with increasing concentrations of the TMPyP4 in the presence of potassium cations [9]. Moreover, Richter and co-workers showed that the TMPyP4 ligand interacts with Herpes Simplex Virus-1 G4s and inhibits polymerase progression *in vitro* [18]. A strong pausing effect was observed, and with the increasing concentrations of TMPyP4, one/two additional bands at the stop site suggested an increased stabilization of G4 structure. In contrast, the TMPyP2 compound did not show any effects different from the untreated oligonucleotide [18].

Furthermore, a combination of the ITC technique, CD measurements, UV melting, and fluorescence spectroscopy allowed us to characterize the RNA G-quadruplex-ligand interactions. Our biophysical analyses provided information about the binding properties of TMPyP4 and BRACO-19 ligands to three RNA G4s from the IAV genome. They revealed some interaction changes that may be caused by different binding mechanisms of the ligand. The structural features of the ligand and its target (in this case, G-quadruplexes) can explain the changes in the affinity and specificity among TMPyP4 and BRACO-19 compounds binding to the same RNA G4s. Several investigations concerning the characterization of the ligand-G4 interactions were reported [37–39]. For instance, Zhang and co-workers analyzed the binding ability of ligands, including BRACO-19 and TMPyP4, to G4s from the enterovirus A71 (EV-A71) genome [37]. The authors found that G4-specific ligands stabilize EV-A71 G-quadruplex structures with high affinity. However, they observed different affinity and selectivity of such interactions [37]. Our results are consistent with literature reports, suggesting that various ligands and their structural rigidity can affect the behavior of G-quadruplex structures.

The main difference between TMPyP4-G-quadruplex and BRACO-19-G-quadruplex interactions lies in their binding modes. TMPyP4 binds externally, mainly through stacking on base pairs and loop regions, and avoids direct stacking on the central G-quartet. For instance, crystallographic studies by Neidle and co-workers showed that intercalation is clearly not the mode of interaction of TMPyP4 with the G4 structures [40]. Notably, Han *et al.* reported that TMPyP4 appears to bind to parallel G4s through external stacking at the ends rather than through intercalation between the G-tetrads [20]. The authors concluded that intercalative binding, although less favored than external binding, may occur, though it remains undetectable using the employed methods, such as the photocleavage assay [20].

In contrast, BRACO-19 binds via an intercalative mode, sandwiched between G-quartets, and also interacts with the grooves through its side chains. This binding pattern makes BRACO-19 a more specific and selective ligand for the G-quadruplex structure [41]. Recently, Si *et al.* reported experimental and molecular docking results showing that BRACO-19 binds with the residues of the DNA G4. The π -moiety of BRACO-19 interacts with the G-tetrad through π-π stacking and the side chain interaction with the groove of DNA G4 [42].

Furthermore, the folding topology and molecularity of 1Q, 7Q, and 11Q structures may also influence their interaction with G4-specific compounds. Based on our CD analysis, we identified the parallel-stranded topology for these three G4s (1Q, 7Q, and 11Q). However, G-quadruplexes as polymorphic structures can differ in their loop regions, the number of stacking guanines within the strand, and their arrangements. Although the binding mechanisms of TMPyP4 and BRACO-19 ligands to RNA G4s are still unclear, results from multiple biophysical methods presented above suggest that these compounds can influence the 1Q, 7Q, and 11Q G-quadruplex structure formation. Taken together, our data emphasize that the interactions between selected ligands and viral RNA G4s may be more complex and less predictable than previously thought, highlighting the need for case-by-case structural and functional validation.

Recent studies have explored G-quadruplex binders as potential antiviral agents, e.g., a paper relating to G-quadruplex as a therapeutic target was described by Lv and colleagues in 2022 [5]. The authors confirmed G4 structure formation within the chikungunya virus (CHIKV) genome using different biophysical techniques. Importantly, they determined the inhibition of CHIKV genome replication by BRACO-19 and TMPyP4 ligands. They also demonstrated that the production of infectious virus was inhibited by these compounds [5]. Our previous studies showed that 1Q and 7Q sequences are present in the segments encoding polymerase complex proteins, whereas 11Q is found within the hemagglutinin (HA) segment [11]. Considering the above findings, we examined the effect of TMPyP4 and TMPyP2 compounds on the IAV minireplicon system activity. The non-infectious life cycle modeling system has been developed using reverse genetic methods [29,30]. It was previously employed, for example, in the chemical targeting of a G4 RNA in the Ebola virus L gene [9].

Here, we assumed that G-quadruplex structures could play an important role in the regulation of the viral replication cycle; therefore, we used the minireplicon system. Importantly, the IAV A/California/04/2009 strain causes dangerous pandemic outbreaks in contrast to the IAV A/Puerto Rico/8/34, a lab-adapted strain, which serves as a model for seasonal influenza viruses. An important point is that we found a difference in the nucleotide composition of G-rich sequences between these two strains within the G-quadruplex-forming regions. Therefore, the use of the minireplicon system based on Cal2009 and PR8 for *in vitro* studies is justified. We found that the compounds we utilized affect the viral polymerase activity in the IAV minireplicon reporter assay, suggesting that they can interact with the G-quadruplexes formed within the vRNA of polymerase complex protein segments.

Additionally, we observed differences in IAV replication inhibition among IAV strains, with the inhibitory effect being more significant for the Cal2009 strain. Therefore, whether this naturally occurring subtle difference, specifically the single nucleotide substitution in the G-rich region within the IAV genome, can be related to viral pathogenicity remains an open question. We hypothesize that this single nucleotide substitution in the PR8 genome could affect G-quadruplex formation, leading to different gene regulation patterns. The G4-specific ligand is not an efficient inhibitor of IAV replication, as the compound has no binding target.

In summary, our paper demonstrates the molecular and structural basis of how TMPyP4 and BRACO-19 bind to selected IAV RNA G4s from the A/California/04/2009 strain. This is the first study to indicate that these structures are formed within this pandemic influenza A virus strain and can interact with G4-specific ligands. Our investigation has proven that both compounds can interact with RNA G-quadruplexes under experimental conditions; however, their binding modes differ. Moreover, the findings reported here show that TMPyP4 and BRACO-19 inhibit reverse transcription reactions *in vitro*. Finally, we demonstrated that TMPyP4 effectively inhibits IAV minireplicon activity in a cellular environment. Collectively, our study suggests that targeting viral RNA G4s by chemical compounds might contribute to the mechanism of their biological activity. Therefore, by employing chemical ligands, we have shown that the PQS regions within the IAV genome represent potential target sites for the development of new, sequence-specific therapeutics. Moreover, we suggest that RNA G4s may play a potentially important role in viral replication, a finding that requires further detailed biological study. Overall, the viral RNA G-quadruplexes could be novel targets for developing an anti-influenza strategy.

Supporting information

S1 Table. List of RNA and DNA oligomers used in this work.
(DOCX)

S2 Table. Cytotoxicity of TMPyP4 and TMPyP2 compounds in HEK 293T cell culture.
(DOCX)

S1 Fig. The full gel images of the inhibition of cDNA synthesis catalyzed by the reverse transcriptase in the presence of TMPyP4 ligand.
(TIF)

S2 Fig. The full gel images of the inhibition of cDNA synthesis catalyzed by the reverse transcriptase in the presence of BRACO-19 ligand.
(TIF)

S3 Fig. The resultant gels from the cDNA synthesis catalyzed by the reverse transcriptase in the presence of TMPyP2 for 1q/1qm, 7q/7qm, and 11q/11qm variants.
(TIF)

S4 Fig. RNA secondary structures for sequences from both Cal2009 (1Q, 7Q, and 11Q) and PR8 (1qPR8, 7qPR8, and 11qPR8) strains. RNA secondary structures were predicted and generated using RNAStructure 6.1 software.
(TIF)

Acknowledgments

We would like to thank Dagny Lorent for her scientific support and helpful suggestions. Dorota Niedzialek gratefully acknowledgethes the "ENSEMBLE3 – Center of Excellence for Nanophononics, Advanced Materials, and Novel Crystal Growth-Based Technologies" (the European Union Horizon 2020 research and innovation program Teaming for Excellence) for technical assistance. We gratefully acknowledge the Polish high-performance computing infrastructure PLGrid (HPC Center: ACK Cyfronet AGH) for providing computer facilities.

Author contributions

Conceptualization: Maria Nalewaj, Pawel Zmora, Marta Szabat.

Data curation: Maria Nalewaj, Marta Szabat.

Funding acquisition: Marta Szabat.

Investigation: Maria Nalewaj, Joanna Sliwiak, Karolina Zielinska, Dorota Niedziałek, Grzegorz Wieczorek, Marta Szabat.

Methodology: Maria Nalewaj, Joanna Sliwiak, Karolina Zielinska, Pawel Zmora, Marta Szabat.

Supervision: Elzbieta Kierzek, Marta Szabat.

Writing – original draft: Maria Nalewaj, Joanna Sliwiak, Karolina Zielinska, Marta Szabat.

Writing – review & editing: Maria Nalewaj, Pawel Zmora, Dorota Niedziałek, Grzegorz Wieczorek, Elzbieta Kierzek, Marta Szabat.

References

1. Qin G, Zhao C, Liu Y, Zhang C, Yang G, Yang J, et al. RNA G-quadruplex formed in SARS-CoV-2 used for COVID-19 treatment in animal models. Cell Discov. 2022;8(1):86. https://doi.org/10.1038/s41421-022-00450-x PMID: 36068208
2. Perrone R, Butovskaya E, Daelemans D, Palù G, Pannecouque C, Richter SN. Anti-HIV-1 activity of the G-quadruplex ligand BRACO-19. J Antimicrob Chemother. 2014;69(12):3248–58. https://doi.org/10.1093/jac/dku280 PMID: 25103489
3. Liu G, Du W, Sang X, Tong Q, Wang Y, Chen G, et al. RNA G-quadruplex in TMPRSS2 reduces SARS-CoV-2 infection. Nat Commun. 2022;13(1):1444. https://doi.org/10.1038/s41467-022-29135-5 PMID: 35301316
4. Zou M, Li J-Y, Zhang M-J, Li J-H, Huang J-T, You P-D, et al. G-quadruplex binder pyridostatin as an effective multi-target ZIKV inhibitor. Int J Biol Macromol. 2021;190:178–88. https://doi.org/10.1016/j.ijbiomac.2021.08.121 PMID: 34461156
5. Lv L, Cui H, Chen Z, Zhou Y, Zhang L. G-quadruplex ligands inhibit chikungunya virus replication. J Med Virol. 2022;94(6):2519–27. https://doi.org/10.1002/jmv.27622
6. Ruggiero E, Zanin I, Terreri M, Richter SN. G-Quadruplex Targeting in the Fight against Viruses: An Update. Int J Mol Sci. 2021;22(20):10984. https://doi.org/10.3390/ijms222010984
7. Sun Z-Y, Wang X-N, Cheng S-Q, Su X-X, Ou T-M. Developing Novel G-Quadruplex Ligands: from Interaction with Nucleic Acids to Interfering with Nucleic Acid·Protein Interaction. Molecules. 2019;24(3):396. https://doi.org/10.3390/molecules24030396 PMID: 30678288

8. Berner A, Das RN, Bhuma N, Golebiewska J, Abrahamsson A, Andréasson M, et al. G4-Ligand-Conjugated Oligonucleotides Mediate Selective Binding and Stabilization of Individual G4 DNA Structures. J Am Chem Soc. 2024;146(10):6926–35. https://doi.org/10.1021/jacs.3c14408
9. Wang S-R, Zhang Q-Y, Wang J-Q, Ge X-Y, Song Y-Y, Wang Y-F, et al. Chemical Targeting of a G-Quadruplex RNA in the Ebola Virus L Gene. Cell Chem Biol. 2016;23(9):1113–22. https://doi.org/10.1016/j.chembiol.2016.07.019 PMID: 27617851
10. Majee P, Pattnaik A, Sahoo BR, Shankar U, Pattnaik AK, Kumar A, et al. Inhibition of Zika virus replication by G-quadruplex-binding ligands. Mol Ther Nucleic Acids. 2021;23:691–701. https://doi.org/10.1016/j.omtn.2020.12.030 PMID: 33575115
11. Tomaszewska M, Szabat M, Zielińska K, Kierzek R. Identification and Structural Aspects of G-Quadruplex-Forming Sequences from the Influenza A Virus Genome. Int J Mol Sci. 2021;22(11):6031. https://doi.org/10.3390/ijms22116031 PMID: 34199658
12. Butovskaya E, Soldà P, Scalabrin M, Nadai M, Richter SN. HIV-1 Nucleocapsid Protein Unfolds Stable RNA G-Quadruplexes in the Viral Genome and Is Inhibited by G-Quadruplex Ligands. ACS Infect Dis. 2019;5(12):2127–35. https://doi.org/10.1021/acsinfecdis.9b00272 PMID: 31646863
13. Katsuda Y, Sato S, Inoue M, Tsugawa H, Kamura T, Kida T, et al. Small molecule-based detection of non-canonical RNA G-quadruplex structures that modulate protein translation. Nucleic Acids Res. 2022;50(14):8143–53. https://doi.org/10.1093/nar/gkac580
14. Brázda V, Porubiaková O, Cantara A, Bohálová N, Coufal J, Bartas M, et al. G-quadruplexes in H1N1 influenza genomes. BMC Genomics. 2021;22(1):77. https://doi.org/10.1186/s12864-021-07377-9 PMID: 33485319
15. Malvern. Manual: MicroCal PEAQ-ITC Analysis Software user manual (English) MAN0576. 2015. Available: www.malvern.com
16. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013;29(7):845–54. https://doi.org/10.1093/bioinformatics/btt055
17. Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. 1996.
18. Artusi S, Ruggiero E, Nadai M, Tosoni B, Perrone R, Ferino A, et al. Antiviral Activity of the G-Quadruplex Ligand TMPyP4 against Herpes Simplex Virus-1. Viruses. 2021;13(2):196. https://doi.org/10.3390/v13020196 PMID: 33525505
19. Le VH, Nagesh N, Lewis EA. Bcl-2 promoter sequence G-quadruplex interactions with three planar and non-planar cationic porphyrins: TMPyP4, TMPyP3, and TMPyP2. PLoS One. 2013;8(8):e72462. https://doi.org/10.1371/journal.pone.0072462 PMID: 23977303
20. Han H, Langley DR, Rangan A, Hurley LH. Selective interactions of cationic porphyrins with G-quadruplex structures. J Am Chem Soc. 2001;123(37):8902–13. https://doi.org/10.1021/ja002179j PMID: 11552797
21. Boschi E, Davis S, Taylor S, Butterworth A, Chirayath LA, Purohit V, et al. Interaction of a Cationic Porphyrin and Its Metal Derivatives with G-Quadruplex DNA. J Phys Chem B. 2016;120(50):12807–19. https://doi.org/10.1021/acs.jpcb.6b09827 PMID: 27936741
22. Majee P, Kumar Mishra S, Pandya N, Shankar U, Pasadi S, Muniyappa K, et al. Identification and characterization of two conserved G-quadruplex forming motifs in the Nipah virus genome and their interaction with G-quadruplex specific ligands. Sci Rep. 2020;10(1):1477. https://doi.org/10.1038/s41598-020-58406-8 PMID: 32001794
23. Oliva R, Mukherjee S, Manisegaran M, Campanile M, Del Vecchio P, Petraccone L, et al. Binding Properties of RNA Quadruplex of SARS-CoV-2 to Berberine Compared to Telomeric DNA Quadruplex. Int J Mol Sci. 2022;23(10):5690. https://doi.org/10.3390/ijms23105690 PMID: 35628500
24. Mishra SK, Shankar U, Jain N, Sikri K, Tyagi JS, Sharma TK, et al. Characterization of G-Quadruplex Motifs in espB, espK, and cyp51 Genes of Mycobacterium tuberculosis as Potential Drug Targets. Mol Ther Nucleic Acids. 2019;16:698–706. https://doi.org/10.1016/j.omtn.2019.04.022 PMID: 31128421
25. Libera V, Fasolato C, Ripanti F, Catalini S, Bertini L, Petrillo C, et al. Molecular mechanisms behind BRACO19 and human telomeric G-quadruplex interaction. Spectrochimica Acta Part A Mol Biomol Spectrosc. 2024;322:124684. https://doi.org/10.1016/j.saa.2024.124684
26. Del Villar-Guerra R, Trent JO, Chaires JB. G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. Angew Chem Int Ed Engl. 2018;57(24):7171–5. https://doi.org/10.1002/anie.201709184 PMID: 29076232
27. von Hacht A, Seifert O, Menger M, Schütze T, Arora A, Konthur Z, et al. Identification and characterization of RNA guanine-quadruplex binding proteins. Nucleic Acids Res. 2014;42(10):6630–44. https://doi.org/10.1093/nar/gku290 PMID: 24771345
28. Diggins L, Ross D, Bhanot S, Corallo R, Daley R, Patel K, et al. CD spectra reveal the state of G-quadruplexes and i-motifs in repeated and other DNA sequences. Biophys Rep. 2025;5(1):100187. https://doi.org/10.1016/j.bpr.2024.100187
29. Watt A, Moukambi F, Banadyga L, Groseth A, Callison J, Herwig A, et al. A Novel Life Cycle Modeling System for Ebola Virus Shows a Genome Length-Dependent Role of VP24 in Virus Infectivity. J Virol. 2014;88(18):10511–24. https://doi.org/10.1128/jvi.01272-14
30. Uebelhoer LS, Albariño CG, McMullan LK, Chakrabarti AK, Vincent JP, Nichol ST, et al. High-throughput, luciferase-based reverse genetics systems for identifying inhibitors of Marburg and Ebola viruses. Antiviral Res. 2014;106:86–94. https://doi.org/10.1016/j.antiviral.2014.03.018 PMID: 24713118
31. Chung W-C, Ravichandran S, Park D, Lee GM, Kim Y-E, Choi Y, et al. G-quadruplexes formed by Varicella-Zoster virus reiteration sequences suppress expression of glycoprotein C and regulate viral cell-to-cell spread. PLoS Pathog. 2023;19(1):e1011095. https://doi.org/10.1371/journal.ppat.1011095 PMID: 36630443
32. Fang P, Xie C, Pan T, Cheng T, Chen W, Xia S, et al. Unfolding of an RNA G-quadruplex motif in the negative strand genome of porcine reproductive and respiratory syndrome virus by host and viral helicases to promote viral replication. Nucleic Acids Res. 2023;51(19):10752–67. https://doi.org/10.1093/nar/gkad759 PMID: 37739415

33. Wang J, Huang H, Zhao K, Teng Y, Zhao L, Xu Z, et al. G-quadruplex in hepatitis B virus pregenomic RNA promotes its translation. J Biol Chem. 2023;299(9):105151. https://doi.org/10.1016/j.jbc.2023.105151 PMID: 37567479

34. Abiri A, Lavigne M, Rezaei M, Nikzad S, Zare P, Mergny J-L, et al. Unlocking G-Quadruplexes as Antiviral Targets. Pharmacol Rev. 2021;73(3):897–923. https://doi.org/10.1124/pharmrev.120.000230 PMID: 34045305

35. Ruggiero E, Richter SN. G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. Nucleic Acids Res. 2018;46(7):3270–83. https://doi.org/10.1093/nar/gky187 PMID: 29554280

36. Chashchina GV, Tevonyan LL, Beniaminov AD, Kaluzhny DN. Taq-Polymerase Stop Assay to Determine Target Selectivity of G4 Ligands in Native Promoter Sequences of MYC, TERT, and KIT Oncogenes. Pharmaceuticals (Basel). 2023;16(4):544. https://doi.org/10.3390/ph16040544 PMID: 37111301

37. Lv L, Zhang L. Characterization of G-Quadruplexes in Enterovirus A71 Genome and Their Interaction with G-Quadruplex Ligands. Microbiol Spectr. 2022;10(3):e0046022. https://doi.org/10.1128/spectrum.00460-22 PMID: 35446122

38. Shitikov E, Bespiatykh D, Malakhova M, Bespyatykh J, Bodoev I, Vedekhina T, et al. Genome-Wide Transcriptional Response of Mycobacterium smegmatis MC2155 to G-Quadruplex Ligands BRACO-19 and TMPyP4. Front Microbiol. 2022;13:817024. https://doi.org/10.3389/fmicb.2022.817024 PMID: 35308348

39. Molnár OR, Végh A, Somkuti J, Smeller L. Characterization of a G-quadruplex from hepatitis B virus and its stabilization by binding TMPyP4, BRACO19 and PhenDC3. Sci Rep. 2021;11(1). https://doi.org/10.1038/s41598-021-02689-y

40. Parkinson GN, Ghosh R, Neidle S. Structural basis for binding of porphyrin to human telomeres. Biochemistry. 2007;46(9):2390–7. https://doi.org/10.1021/bi062244n PMID: 17274602

41. Sullivan H-J, Readmond C, Radicella C, Persad V, Fasano TJ, Wu C. Binding of Telomestatin, TMPyP4, BSU6037, and BRACO19 to a Telomeric G-Quadruplex-Duplex Hybrid Probed by All-Atom Molecular Dynamics Simulations with Explicit Solvent. ACS Omega. 2018;3(11):14788–806. https://doi.org/10.1021/acsomega.8b01574 PMID: 30555989

42. Si MK, Pramanik SK, Ganguly B. Tuning the ring strain effect in acridine derivatives on binding affinity with G-quadruplex-DNA: A computational and experimental study. Int J Biol Macromol. 2019;124:1177–85. https://doi.org/10.1016/j.ijbiomac.2018.12.019 PMID: 30521912

SUPPLEMENTARY DATA

The interaction of RNA G-quadruplexes from the influenza A virus vRNA with TMPyP4 and BRACO-19 ligands

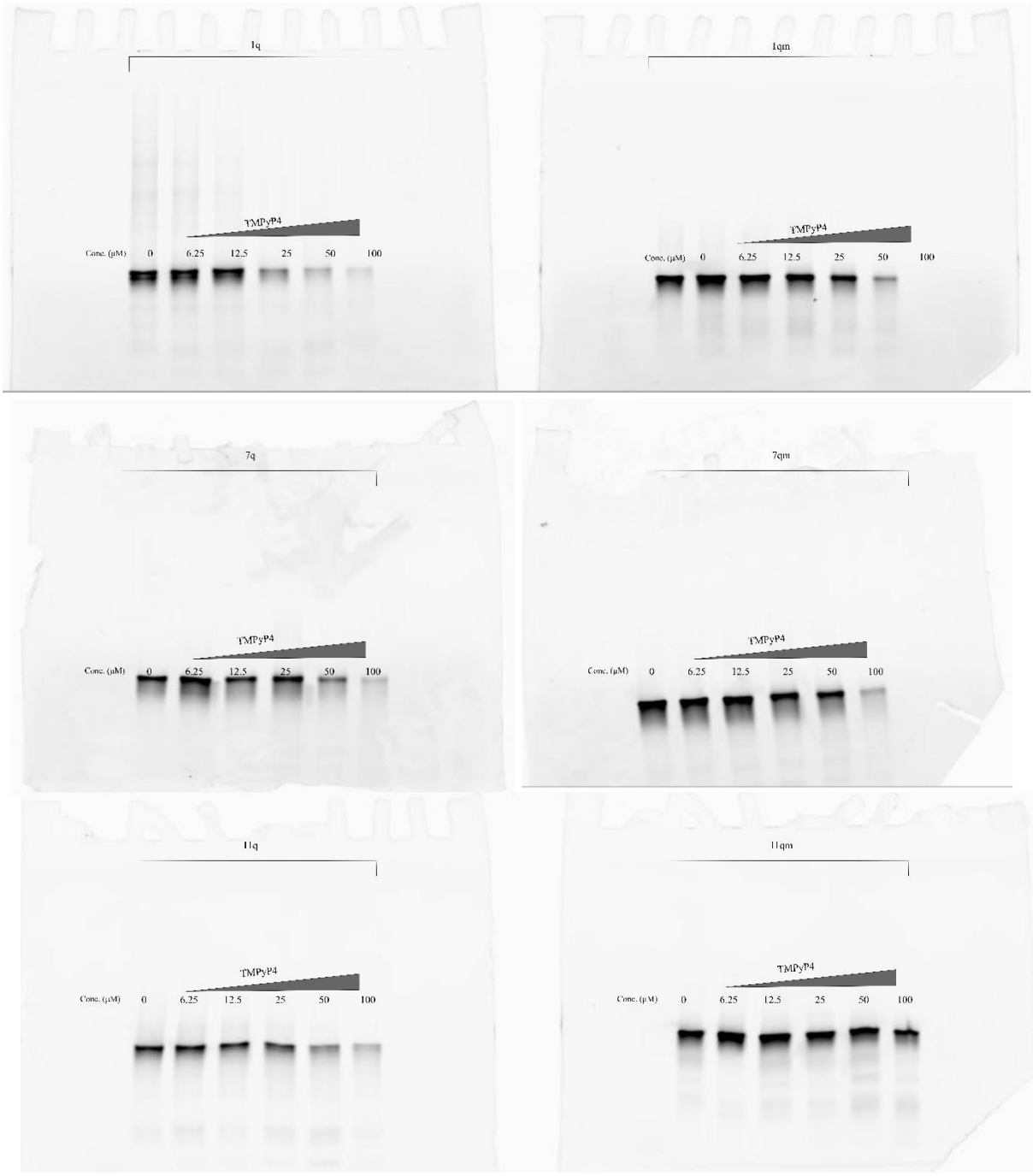
Table S1. List of RNA and DNA oligomers used in this work.

		Name	Sequence (5′-3′)	Application
RNA	wild-type	1q	CUGGUGGGGCAGCAGCAAAGGGGAGCCGCUUGUACCGA	RT stop assay
		7q	GGUAGUGGUCCAUCAAUCGGGUUGAGCUGGGGCCGCUUGUACCGA	
		11q	GGAUGUAUAUUCUGAAAUGGGAGGCUGGCCGCUUGUACCGA	
	mutant	1qm	CUGGUGGAGCAGCAGCAAAAGGAAGCCGCUUGUACCGA	
		7qm	GGUAGUGGUCCAUCAAUCGGAUUGAGCUGAGGCCGCUUGUACCGA	
		11qm	GGAUGUAUAUUCUGAAAUGGAAGACUUGCCGCUUGUACCGA	
DNA		primer 1	FAM -TCGGTACAAGCGG	
RNA		1Q	CUGGUGGGGCAGCAGCAAAGGGGAG	ITC method
		7Q	GGUAGUGGUCCAUCAAUCGGGUUGAGCUGGGG	
		11Q	GGAUGUAUAUUCUGAAAUGGGAGGCUGG	
		1Qm	CUGGUGGAGCAGCAGCAAAAGGAAG	
		7Qm	GGUAGUGGUCCAUCAAUCGGAUUGAGCUGAGG	
		11Qm	GGAUGUAUAUUCUGAAAUGGAAGACUUG	
DNA		For-PB1 insert	ACCATGGATGTCAATCCGACTCTAC	minireplicon system preparation
		Rev-PB1 insert	CAATGGTGGAACAGATCTTCATGATCTC	
		For-PB1 vector	GAAGATCTGTTCCACCATTGAAGAACTCAG	
		Rev-PB1 vector	GTCGGATTGACATCCATGGTGGTAC	
		For-PB2 insert	CACCATGGAGAGAATAAAAGAACTGAGAGA	
		Rev-PB2 insert	GCTCGAGCTAATTGATGGCCATCCGAA	
		For-PB2 vector	GGCCATCAATTAGCTCGAGCTAGCAGAT	
		Rev-PB2 vector	CTTTTATTCTCTCCATGGTGATGGGTACCATGCA	
		For-PA insert	GGTACCACCATGGAAGACTTTGTGCGAC	
		Rev-PA insert	CTAGCTCGAGCTACTTCAGTGCATGTGTGAG	

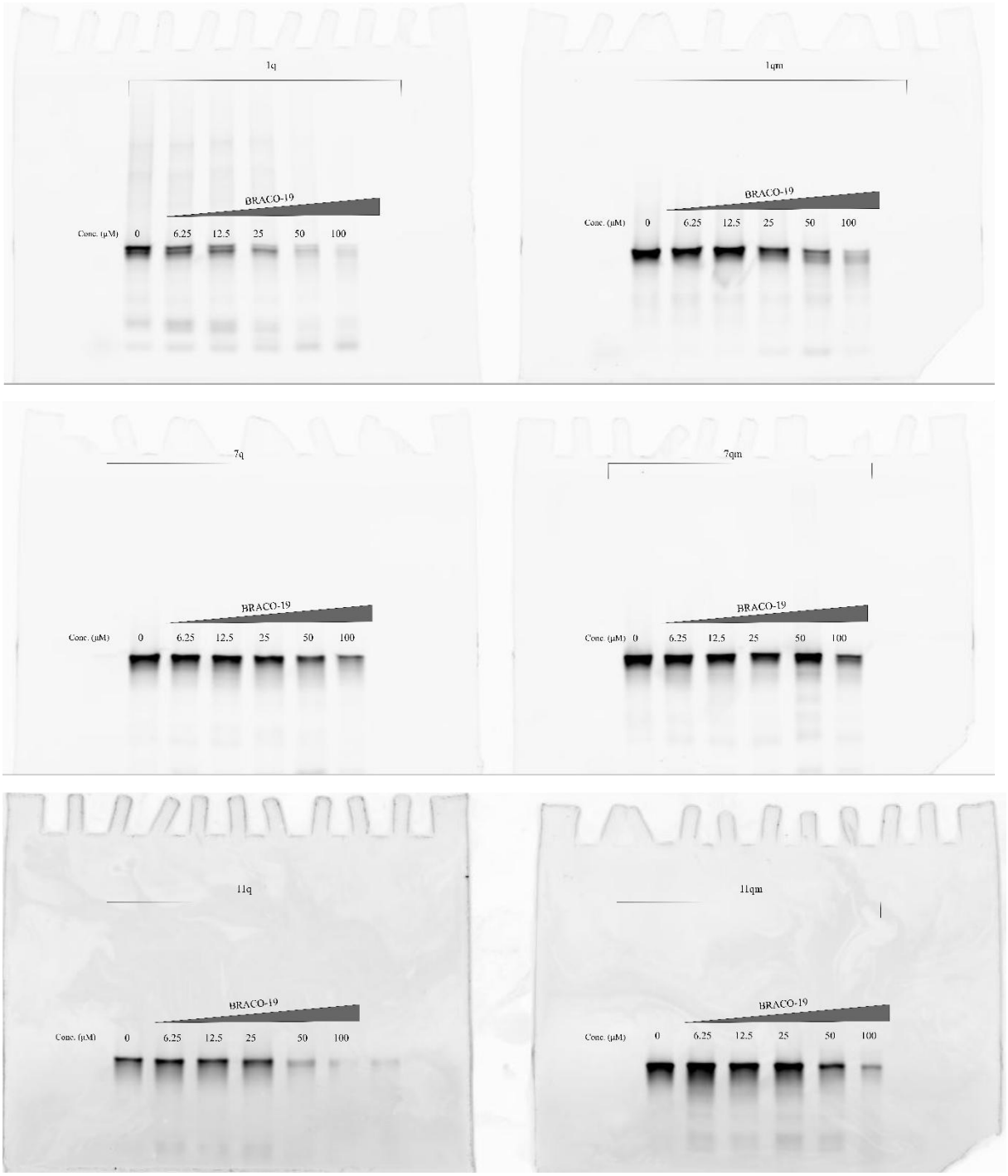
		For-PA vector	ACTGAAGTAGCTCGAGCTAGCAGATCTTTTT	
		Rev-PA vector	AAGTCTTCCATGGTGGTACCATGCATCGATG	
		For-NP insert	CATGGTACCACCATGGCGTCTCAAGGCA	
		Rev-NP insert	CTAGCTCGAGTCAACTGTCATACTCCTCTGCA	
		For-NP vector	TGACAGTTGACTCGAGCTAGCAGATCTTTTTCC	
		Rev-NP vector	GACGCCATGGTGGTACCATGCATCGATG	
Red color indicates point mutations in G-rich region. FAM – fluorescein modification at the 5′-end				

Table S2. Cytotoxicity of TMPyP4 and TMPyP2 compounds in HEK 293T cell culture.

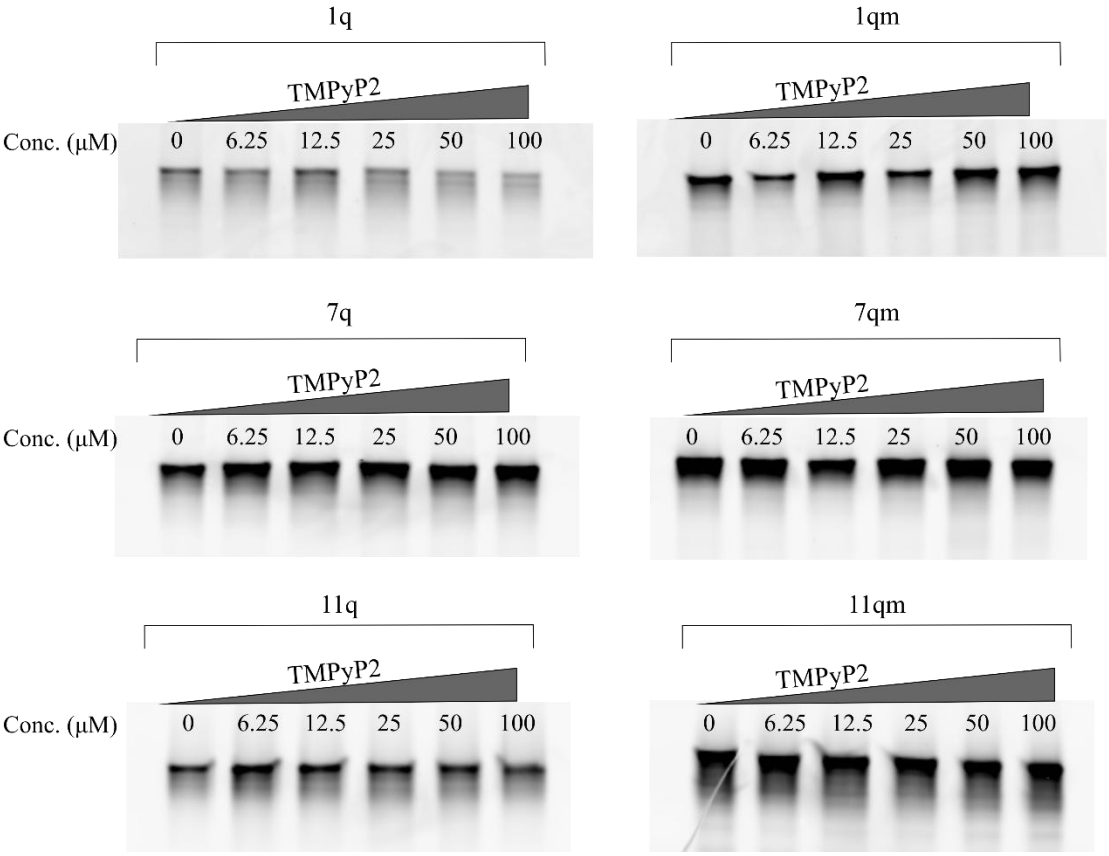
Compound concentration [μM]	**Cellular viability** (% of cells alive within 48 h)	**Standard deviation**
TMPyP4		
1.25	104.54	± 6.8
3.125	95.8	±6.86
6.25	89.01	±4.05
7.5	83.24	± 4.61
TMPyP2		
3.125	98.50	± 5.25
6.25	89.68	± 6.64
12.5	78.38	± 4.64



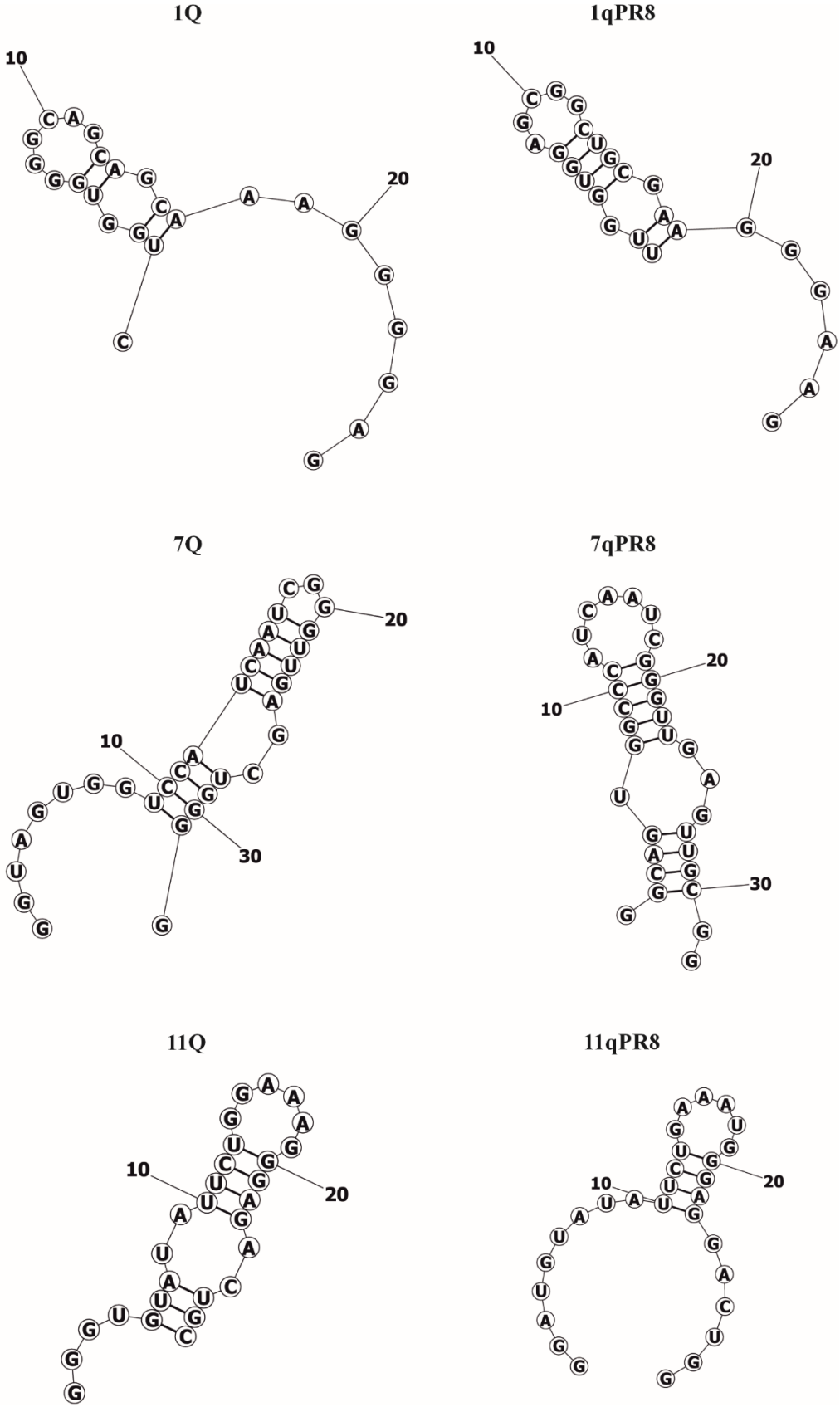
S1 Fig. The full gel images of the inhibition of cDNA synthesis catalyzed by the reverse transcriptase in the presence of TMPyP4 ligand.



S2 Fig. The full gel images of the inhibition of cDNA synthesis catalyzed by the reverse transcriptase in the presence of BRACO-19 ligand.



S3 Fig. The resultant gels from the cDNA synthesis catalyzed by the reverse transcriptase in the presence of TMPyP2 for 1q/1qm, 7q/7qm, and 11q/11qm variants.



S4 Fig. RNA secondary structures for sequences from both Cal2009 (1Q, 7Q, and 11Q) and PR8 (1qPR8, 7qPR8, and 11qPR8) strains.
RNA secondary structures were predicted and generated using RNAStructure 6.1 software.

APPENDIX 2

Statements determining the candidate's contribution

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

Poznań, 22.02.2026

CONTRIBUTION STATEMENT

Publication title: *Identification and structural aspects of G-quadruplex-forming sequences from the influenza A virus genome*
Authors: Tomaszewska, M.*, Szabat, M.*§, Zielińska, K., Kierzek, R.§
Journal: International Journal of Molecular Sciences
Release year: 2021

* co-authors who equally contributed to this work
§corresponding author

I declare that my contribution to the above-mentioned scientific work included the conceptualization of the experiments, conducting the experimental work, analysis and interpretation of the results, preparation of the figures, and drafting and editing the manuscript.

Maria Nalewaj

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

22.02.2026

CORRESPONDING AUTHORS STATEMENT ABOUT THE PHD CANDIDATE CONTRIBUTION

Publication title: *Identification and structural aspects of G-quadruplex-forming sequences from the influenza A virus genome*
Authors: Tomaszewska, M.*, Szabat, M.*§, Zielińska, K., Kierzek, R.§
Journal: International Journal of Molecular Sciences
Release year: 2021

* co-authors who equally contributed to this work
§corresponding author

We declare that the contribution of Maria Nalewaj to the above-mentioned scientific work included the conceptualization of the experiments, conducting the experimental work, analysis and interpretation of the results, preparation of the figures, and drafting and editing the manuscript.

Marta Szabat Ryszard Kierzek

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

Poznań, 22.02.2026

CONTRIBUTION STATEMENT

Publication title: *Examples of structural motifs in viral genomes and approaches for RNA structure characterization*
Authors: Nalewaj M.*, Szabat, M.*§
Journal: International Journal of Molecular Sciences
Release year: 2022

* co-authors who equally contributed to this work
§corresponding author

I declare that my contribution to the above-mentioned scientific work included performing the literature review, preparation of the figures, and drafting and editing the manuscript.

Maria Nalewaj

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

22.02.2026

CORRESPONDING AUTHOR STATEMENT ABOUT THE PHD CANDIDATE CONTRIBUTION

Publication title: *Examples of structural motifs in viral genomes and approaches for RNA structure characterization*
Authors: Nalewaj M.*, Szabat, M.*§
Journal: International Journal of Molecular Sciences
Release year: 2022

* co-authors who equally contributed to this work
§corresponding author

I declare that the contribution of Maria Nalewaj to the above-mentioned scientific work included performing the literature review, preparation of the figures, and drafting and editing the manuscript.

Marta Szabat

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

Poznań, 22.02.2026

CONTRIBUTION STATEMENT

Publication title: *The interaction of RNA G-quadruplexes from the influenza A virus vRNA with TMPyP4 and BRACO-19 ligands*
Authors: **Nalewaj, M.**, Sliwiak, J., Zielinska, K., Zmora , P., Niedziałek, D., Wieczorek, G., Kierzek, E., Szabat, M. §
Journal: PLoS One International Journal of Molecular Sciences,
Release year: 2025

§corresponding author

I declare that my contribution to the above-mentioned scientific work included the conceptualization of the experiments, conducting the experimental work, analysis and interpretation of the results, preparation of the figures, and drafting and editing the manuscript.

Maria Nalewaj

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

22.02.2026

CORRESPONDING AUTHOR STATEMENT ABOUT THE PHD CANDIDATE CONTRIBUTION

Publication title: *The interaction of RNA G-quadruplexes from the influenza A virus vRNA with TMPyP4 and BRACO-19 ligands*
Authors: **Nalewaj, M.**, Sliwiak, J., Zielinska, K., Zmora , P., Niedziałek, D., Wieczorek, G., Kierzek, E., Szabat, M. §
Journal: PLoS One International Journal of Molecular Sciences,
Release year: 2025

§corresponding author

I declare that the contribution of Maria Nalewaj to the above-mentioned scientific work included the conceptualization of the experiments, conducting the experimental work, analysis and interpretation of the results, preparation of the figures, and drafting and editing the manuscript.

Marta Szabat