One Solution for All: Searching for Universal Aptamers for Constantly Mutating Spike Proteins of SARS-CoV-2

Jiuxing Li,[a] Zijie Zhang,[a] Ryan Amini,[a] and Yingfu Li*[a]

Aptamers that can recognize the spike (S) protein of SARS-CoV-2 with high affinity and specificity are useful molecules towards the development of diagnostics and therapeutics to fight COVID-19. However, this S protein is constantly mutating, producing variants of concern (VoCs) that can significantly weaken the binding by aptamers initially engineered to recognize the S protein of the wildtype virus or a specific VoC. One strategy to overcome this problem is to develop universal aptamers that are insensitive to all or most of the naturally emerging mutations in the protein. We have recently demonstrated this concept by subjecting a pool of S protein-binding DNA aptamers for one-round parallel-SELEX experiments targeting 5 different S protein variants for binding-based sequence enrichment, followed by bioinformatic analysis of the enriched pools. This effort has led to the identification of a universal aptamer that recognizes 8 different variants of the spike protein with equally excellent affinity.

Introduction

Over two years following its discovery, SARS-CoV-2 and COVID-19 continue to be a global health problem with no apparent ending date in sight. The continuous administration of vaccines and the accompanied elimination of most or all of COVID-19 restrictions and policies in many countries provide a positive outlook on the situation, but public health policies must nonetheless remain cautious of this virus’ continuing threats in the future. Yet, regardless of when COVID-19 becomes endemic, SARS-CoV-2 has reminded the scientific community of the importance of effective diagnostic and therapeutic interventions during a pandemic. Current testing and treatment methods are plentiful, but each has its limitations. RT-qPCR, for example, is costly and time-consuming,[3,4] rapid antigen tests, on the other hand, are far less sensitive and prone to false-negative results.[5] The development of therapeutic solutions is in a worse state, as only one drug (remdesivir) has been approved by the Food and Drug Administration.[6] Without any doubt, efforts are needed to establish better or enhanced solutions to COVID-19 detection and treatment.

As a typical coronavirus, SARS-CoV-2 is defined by its four main structural proteins: envelope, membrane, nucleocapsid, and spike (S). Although each protein has the potential to serve as an antigen target, the S protein remains the popular choice for several reasons. Firstly, it is indispensable to the virus, as the receptor binding domain (RBD) on the S1 subunit is required for viral entry.[6,7] Additionally, the S protein lies on the surface of the virus, offering a target for detection without the need for lysis.[6,9] Taken altogether, the S protein reigns supreme as the most optimal antigen target in the fight against SARS-CoV-2.

One potential approach, for both COVID-19 and future pandemics, is to use aptamers as molecular recognition elements. Aptamers are short single-stranded synthetic oligonucleotides that are capable of forming well-defined three-dimensional structures to bind a target of interest.[10,11] They are isolated from a large library of random sequences by an in vitro process known as SELEX (Systematic Evolution of Ligands through Exponential Enrichment).[12,13] In this technique, sequences are progressively enriched through iterative cycles of partitioning and amplification based on their affinity and/or specificity for the target.[14] The procedure is continued until satisfactory binding occurs, and the best-performing species are subsequently sequenced and studied. Characteristically, aptamers are similar in function to antibodies, hence their alternate name “chemical antibodies”. Yet, these nucleic acid aptamers provide some advantages over their protein counterparts.[15] They are smaller, more cost-effective, more stable, easier to synthesize and modify, and less immunogenic.[16,17]

With such benefits, it comes as no surprise that multiple labs around the world have produced aptamers relevant to diagnostic and therapeutic solution to SARS-CoV-2. Thus far, more than a dozen SELEX experiments have been conducted against the S protein, and over fifty aptamers have been derived.[18–28] Most of the aptamers display excellent affinity for the S protein, with $K_d$ (dissociation constant) values predominantly ranging between 1–20 nM. Some groups, including ours, have gone as far as engineering dimeric aptamers to further increase the binding affinity.[20] By joining two aptamers that bind two epitopes of this trimeric protein, binding affinity is enhanced, and $K_d$ values drop to the pM range. Many of the published aptamers also exhibit high recognition specificity to the S protein. The inclusion of counter selection steps against unintended targets (e.g., pseudotyped lentiviruses with SARS-
CoV-1 S protein) during the SELEX protocol helps to identify aptamers with high recognition specificity.21

One application of SARS-CoV-2 S protein aptamers is within the therapeutic pipeline. The binding of aptamers with the S protein can prevent the S protein-angiotensin converting enzyme 2 (ACE2) interaction and disrupt virus entry, which has been successfully demonstrated in several studies.18,20–22 Sun et al. reported the first therapeutic aptamer named CoV2-6 against SARS-CoV-2,22 which was further minimized and engineered into a circular bivalent aptamer cb-CoV2-6 C3 for improved stability in serum (stable for 12 hours), increased binding affinity ($K_d = 0.13$ nM), and excellent ability to inhibit viral infection towards human cells ($IC_{50} = 0.42$ nM). Independently, two aptamers named Aptamer-120 and nCoV-S1-Apt124 were also developed by Zu et al. and Huang et al. to neutralize SARS-CoV-2 spike pseudotyped virus and prevent host cell interactions. Notably, Mayer et al. identified an aptamer for the S protein that did not block the RBD-ACE2 interaction but was still able to prevent the infection of a pseudovirus against S protein that did not block the RBD-ACE2 interaction but was able to prevent the infection of a pseudovirus against SARS-CoV-2 2022. In addition, S protein binding aptamers have been conjugated with gold nanoparticles or gold nanostars to increase the efficacy for suppressing SARS-CoV-2 infection.18,31 It is important to note that aptamers must possess excellent stability against nucleases in bodily fluids, in order to be useful as therapeutics. Circularization of aptamers as described above is one option,22 while chemical modification is another. 2′-fluoro-arabinonucleic acid (FANA)-based aptamer ($K_d = 10–20$ nM)22 and 2′-fluoro protected RNA aptamer ($K_d = 18$ nM)23 for the RBD of the S protein have been reported by DeStefano et al. and Kjems et al., respectively, both of which were shown to be effective towards inhibiting the infection of pseudotyped viruses. Thus far, however, most therapeutic aptamers have only been tested in vitro. Future follow-up in vivo studies on the reported aptamers will be imperative to validate their therapeutic values.

More prevalent is the application of the S protein binding “aptasensors”, which are aptamer-based biosensors. Aptasensors can be further broken down into several types based on their signal transduction mechanisms (e.g., electrochemical sensors, optical sensors, lateral flow assays, aptamer-linked immobilized sorbent assays, etc.). Considerable efforts have been made to create a diverse array of SARS-CoV-2 aptasensors based on the S protein binding aptamers. Our McMaster team, for example, recently developed the “CoV-eChip” – a rapid sensor for SARS-CoV-2 diagnosis using human saliva. The sensor is anchored by a dimeric aptamer, and functions through the principle of electrochemical impedance spectroscopy.29 It provides a rapid, single-step detection method that can be conveniently paired with simple instrumentation. Elsewhere, groups such as Huang et al.33 and Lewis et al.34 have reported optical aptasensors for detecting the S protein in biometrics such as urine, blood, and saliva. Most of these aptasensors still need to be optimized for commercialization, considering their expensive instrumentation and intricate sample handling.

However, the utility of aptamers for the S protein have been complicated by the constant mutations occurring to this protein observed with many variants of concern (VoCs) (Table 1). VoCs are defined by several amino acid mutations in the S protein. These changes are known to improve the binding of the virus to human ACE2, giving rise to more infectious and antibody-resistant strains.35,36 Alpha (B.1.1.7), Beta (B.1.351), and Gamma (P.1) variants were first to enter the scene, followed up by the more concerning Delta (B.1.617.2) and Omicron (B.1.1.529 and B.1.1.529.2) strains (Figure 1).35 Constant S protein mutations (Table 1 and Figure 1) present a significant challenge to vaccines, diagnostics, and therapeutics.

Table 1. List of SARS-CoV-2 variants of concern (VoCs).

| Lineage     | WHO label    | Earliest documented samples | Mutations in the S protein                                                                 |
|------------|--------------|----------------------------|------------------------------------------------------------------------------------------|
| B.1.1.7    | Alpha        | United Kingdom (September 2020) | △F69–70, △A144, N501Y, A570D, D614G, D614G, T716I, S982A, D1118H                        |
| B.1.351    | Beta         | South Africa (May 2020)       | L18F, D80A, D215G, K417N, E484K, N501Y, D614G, A701V                                   |
| P.1        | Gamma        | Brazil (November 2020)        | L18F, T20N, P26S, D138Y, R190S, K417N, K417T, E484K, N501Y, D614G, H655Y, T10271I, V1176F |
| B.1.617.2  | Delta        | India (October 2020)          | T19R, △I156–157, R158G, L452R, T478K, D614G, P681R, D950N                           |
| B.1.1.529  | Omicron      | South Africa (November 2021)  | A67V, A69–70, T95I, A142–144, Y145D, S211L, L212F, D214E, G239D, S371L, S373P, S375F, R144K, K417T, N400K, G446S, S477N, T478K, E484A, Q489K, G496S, Q498R, N501Y, S505H, T547K, D614G, H655Y, N679K, P681H, N764K, K769R, D796Y, N856K, Q954H, N969K, L981F |
| B.1.1.529.2| Omicron BA.2 | Multiple Countries (November 2021) | T19L, V24–26, A275, T95I, G142D, V213G, C339D, S371F, S373P, S375F, S376F, T736A, D405N, R408S, K417N, N440K, G446S, S477N, T478K, E484A, Q489K, G496S, Q498R, N501Y, S505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K, L981F |
Selection of Universal Aptamers

Most of the SELEX experiments conducted to date in search of S protein binding aptamers were targeted to the wildtype (Wuhan) strain of SARS-CoV-2 (Table 2). That is not to say that these aptamers do not bind to the S protein of other variants, but they are at risk of binding with reduced affinity, given that there are more and more mutations in the S protein as the virus continues to evolve (Table 1 and Figure 1). It is therefore extremely important to develop "universal" aptamers with which to build biosensors and therapeutic agents targeting emerging VoCs. Universal aptamers are defined as aptamers that show equally high affinity for the S proteins from all or most of the current and future VoCs. For this reason, selection methods that can generate universal aptamers are very desirable. Ideally, such methods should be able to derive aptamers in a short period of time taking into consideration the rapid mutations of the S protein.

To be precise, a universal aptamer is an aptamer that binds with a specific protein target but is insensitive to some key mutations of the target, which makes the aptamer able to recognize all (or most) of the variants of the same protein family regardless of amino acid mutations. Universal aptamers capable of recognizing the same S protein target from diverse VoCs would greatly simplify the diagnostic and therapeutic processes that employ these aptamers.

Isolation of universal aptamers for the SARS-CoV-2 S protein is possible in theory due to the following reasons. On the one hand, the DNA library used for aptamer selection is highly diverse in sequence options (generally containing at least $10^{14}$ unique sequences, graphically illustrated with different shapes in Figure 2A). On the other hand, even though there are many dominant mutations in the S protein that closely relate to immune evasion (Figure 2B), regions do exist on this protein that are less tolerable to mutation,[37] and these more conserved regions may provide binding sites for universal aptamers. Furthermore, there appear to be many aptamer sequence solutions for the S protein, as demonstrated by successful isolation of diverse aptamer sequences for this protein by the aptamer community. Most of the reported aptamers show high binding affinity with $K_d$ values in the nanomolar range despite their disparate sequences and secondary structures.

Using a pre-structured DNA library with a 40-nucleotide (nt) random region (Figure 3A) placed in a hairpin-structured arrangement (which is often observed with many published aptamers), we successfully isolated a series of aptamers for the SARS-CoV-2 S protein.[19] The best three aptamers were MSA1 (Figure 3B, $K_d = 1.8$ nM), MSA3 (Figure 3C, $K_d = 1.9$ nM), and MSA5 (Figure 3D, $K_d = 2.7$ nM) which ranked the 1st, 3rd, and 5th respectively in the enriched library of generation 13 (G13). Nevertheless, further examination of the $K_d$ values of other aptamers in the G13 pool, including two lowly ranked aptamers, MSA50 (Figure 3E, ranked the 50th; $K_d = 10.2$ nM) and MSA439 (Figure 3F, ranked the 439th; $K_d = 36.9$ nM), showed that even these lowly ranked aptamers still possess excellent affinity for the S protein.[19]

The observations discussed above suggest that the G13 aptamer pool is sequence-diverse and may contain candidates that can universally recognize the S proteins of all VoCs. This was the foundational hypothesis examined by us in a recent study.[38] To find these aptamers quickly, we carried out five parallel one-round SELEX experiments with the G13 pool using

Table 2. Some best performing SARS-CoV-2 aptamers targeting the S protein.

| Identifier | Aptamer Name       | SELEX Target | $K_d$ [nM] | Ref. |
|------------|--------------------|--------------|-----------|-----|
| Song-2020  | CoV2-RBD-1         | RBD          | 3.1       | [18]|
| Li-2021    | MSA1               | S1           | 1.8       | [19]|
| Schmitz-2021 | SP5       | Trimeric S   | 9.2       | [21]|
| Yang-2021  | nCoV-S1-Apt3       | S1           | 0.118     | [24]|

Figure 2. (A) Schematic illustration of structural diversity of DNA libraries used for aptamer selection. (B) Mutation site summary of the SARS-CoV-2 S protein. Reproduced with permission.[37] Copyright 2021, American Chemical Society.
five different S proteins as the binding targets (Figure 4A). These variants were: the wildtype S protein plus the S proteins from B.1.1.7 (Alpha, United Kingdom, UK), B.1.351 (Beta, South Africa, SA), P.1 (Gamma, Brazil, BZ) and B.1.429 (Epsilon, California, CA) variants, all of which were commercially available at the time of the experiment. Aptamer partition was conducted using the electrophoretic mobility shift assay (EMSA), wherein each S protein was incubated with the G13 pool prior to gel separation of the S protein-DNA complexes from the unbound DNA. PCR was employed to amplify the bound DNA after elution from the gel. The amplified DNA pools were then subjected to high-throughput sequencing, followed by a comparative analysis of abundance changes of the aptamer sequences in these pools.

By comparing the ranking of the top 100 sequences in each pool after one-round selection, we found that the ranking of a particular sequence, named MSA52, was significantly increased in every pool enriched with a specific variant S protein. MSA52 was ranked the 52nd in the G13 pool selected with the wildtype S protein. However, the ranking of MSA52 increased to 24th, 19th, and 14th after one-round selection with the S protein from the UK (UKS), Brazil (BZS), South Africa (SAS), and California (CAS) variants, respectively, but remained largely unchanged (with a ranking of 46th) in the pool with the wildtype S protein (WHS).

By comparing the ranking of the top 100 sequences in each pool after one-round selection, we found that the ranking of a particular sequence, named MSA52, was significantly increased in every pool enriched with a specific variant S protein. MSA52 was ranked the 52nd in the G13 pool selected with the wildtype S protein. However, the ranking of MSA52 increased to 24th, 19th, 17th, and 14th after one-round selection with the S protein from the UK (UKS), Brazil (BZS), South Africa (SAS), and California (CAS) variants, respectively, but remained largely unchanged (with a ranking of 46th) in the pool with the wildtype S protein (WHS). These observations indicate that MSA52 competed favorably for binding all four S protein variants with the top-ranking sequences established using WHS, pointing to the possibility of this aptamer being the universal aptamer that we were searching for.
An in-depth analysis of the MSAS2 aptamer family led to the discovery of a total of 321 members, which are named MSAS2Rx, where R refers to ranking, x is the ranking in the MSAS2 family. The ranking increases of the top 10 sequences (MSAS2R1-MSAS2R10) in the UK, B2, SA, and CA pools are calculated by comparison with their rankings in the WH pool (Figure 4B). Interestingly, all these sequences exhibit significant ranking increases in all the variant pools, many of which show ~50% ranking increases. For example, the ranking of MSAS2R1 increased from 46th in the WH pool to 14th in the CA pool, producing a ranking increase of 69.5%. The results once again confirmed that the members in this family competed well for the binding of all four variants used for selection. Taken together, the sequencing analysis identified MSAS2 as a universal aptamer candidate for the S protein.

The sequence and predicted secondary structure of MSAS2 were provided in Figure 5A.[38] MSAS2 consists of 79 nucleotides and the overall structure contains 4 pairing arms (P1, P2, P3 and P4) and 7 unpaired elements (SS12, SS23, SS34, SS41, and the loops L2, L3, L4). P1 was the pairing arm that was pre-designed in the structure of the library (Figure 3A). The 22-nt segment (purple colored) that is located from the 24th G residue in L2 to the 45th G residue in P4 was determined to be the key sequence elements for binding the spike protein. To reduce the aptamer size, a series of sequence truncations were conducted for MSAS2. The minimal 69-nt version, MSAS2-T5 (Figure 5B), displays the same binding affinity as MSAS2 for the S protein.

![Figure 5. The secondary structures of (A) the full-length MSAS2 and (B) truncated version MSAS2-T5. Reproduced with permission. Copyright 2022, Wiley-VCH.](image)

The binding affinity of MSAS2 for the spike proteins of SARS-CoV-2 and its B.1.1.7, B.1.351, P.1, B.1.429 variants were confirmed by dot blot assays. The \( K_D \) values were listed in Table 3. MSAS2 showed strong binding to these five spike proteins with \( K_D \) values in the 3.6–10.2 nM range; the affinities for the wildtype (3.6 nM), B.1.1.7 (3.8 nM), and B.1.429 (3.8 nM) spike proteins were nearly the same, the affinities were slightly weaker for the B.1.351 (8.5 nM) and P.1 (10.2 nM) S protein variants.[38]

Importantly, we were pleased to find out that MSAS2 also showed excellent affinity to three variant S proteins that were not used in the selection experiment. Specifically, we had tested the binding of MSAS2 for three newly emerged S protein variants, B.1.617.1 (Kappa), B.1.617.2 (Delta) and B.1.1.529 (Omicron), and obtained \( K_D \) values of 2.8, 3.7 and 6.2 nM, respectively, which were close to the \( K_D \) value observed for the wildtype S protein (\( K_D = 3.6 \) nM).[38] The results further confirm that MSAS2 is a universal aptamer as it is capable of recognizing all of the eight S protein variants of SARS-CoV-2 that we have tested to date. The binding affinity of MSAS2 for the S protein from the B.1.1.529.2 (Omicron BA.2) variant is currently under investigation.

In addition to the spike proteins, MSAS2 also showed universally high affinities for the pseudotyped SARS-CoV-2 lentiviruses (PVs) that were engineered to display the spike proteins of SARS-CoV-2 variants within the viral envelope. The \( K_D \) values of MSAS2 for the eight PVs expressing the spike proteins of the wildtype SARS-CoV-2 and 7 variants are listed in Table 3, which range from 18.4 to 49.0 pM.[38] Note that each PV virus is supposed to carry a comparable number of spike proteins with SARS-CoV-2, approximately 30 copies per virus.[39,40] Taken together, MSAS2 can recognize all of the tested PV variants with high binding affinities.

### Table 3. Binding affinity of MSAS2 for SARS-CoV-2 variants.

| SARS-CoV-2 Variants | WHO label | Binding Affinity \( K_D \) [nM] | Spike protein [nM] | PV [pM] [b] |
|---------------------|-----------|--------------------------------|-------------------|------------|
| Wildtype            | SARS-CoV-2| 3.6 ± 0.4                      | 18.4 ± 1.8        |            |
| B.1.1.7             | Alpha     | 3.8 ± 0.2                      | 33.6 ± 3.2        |            |
| B.1.351             | Beta      | 8.5 ± 0.8                      | 46.4 ± 5.6        |            |
| P.1                 | Gamma     | 10.2 ± 1.4                     | 38.2 ± 2.6        |            |
| B.1.429             | Epsilon   | 3.8 ± 0.6                      | 36.5 ± 4.7        |            |
| B.1.617.1           | Kappa     | 2.8 ± 0.6                      | 49.0 ± 3.9        |            |
| B.1.617.2           | Delta     | 3.7 ± 0.4                      | 28.9 ± 2.6        |            |
| B.1.1.529           | Omicron   | 6.2 ± 0.6                      | 24.6 ± 3.6        |            |

[a] The \( K_D \) values were from reference [38].
[b] PV: pseudotyped viruses of SARS-CoV-2.

### Summary and Outlook

For the past two years, many aptamers had been isolated for the wildtype S protein or its RBD; however, these aptamers have not been comprehensively examined towards recognizing the S protein of diverse VoCs.[18–26,41] We had randomly chosen some of these aptamers to compare their relative affinity for variant S proteins of B.1.1.7, B.1.351, P.1, and B.1.617.2 in a sandwich assay.[18] We found that none of these aptamers showed consistently excellent affinity for all these variants, whether the targets were purified proteins or PVs in contrast to the universally high affinity by MSAS2 for these targets.[38] This finding suggests that MSAS2 is unique among the published S protein binding aptamers in terms of differences in epitope recognition. A likely explanation for this observation is that MSAS2 recognizes amino acids of the S protein that are conserved in the investigated variants, while other aptamers interact with some of the amino acids of the wildtype S protein that are mutated in the variant S proteins.

The discovery of MSAS2 has provided a universal high-affinity recognition element for the S protein of diverse SARS-CoV-2 VoCs, which makes it possible to develop persistently...
effective diagnostic and therapeutic solutions to the constantly evolving SARS-CoV-2 viruses. We expect, and hope, that MSAS2 may also recognize the S proteins of future VoCs. However, if this prediction does not hold true for a future VoC, we should be able to find another universal aptamer employing the same approach for the discovery of MSAS2. Given the fact that the combined parallel selection and bioinformatic approach is a rapid process (taking less than a week to complete), we believe this strategy is well suited for developing solutions when encountering another future pandemic like COVID-19.

Acknowledgements

Funding for research on COVID-19 in our laboratories have been provided by the Canadian Institutes of Health Research (CIHR), the Natural Science and Engineering Research Council of Canada (NSERC), Mitacs and McMaster University.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: COVID-19 · aptamer · spike protein · diagnostics · therapeutics

[1] P. Khan, L. M. Auff德brinck, A. E. Engelhart, ACS Synth. Biol. 2020, 9, 2861–2880.
[2] M. J. Miña, K. G. Andersen, Science 2021, 371, 126–127.
[3] V. Parvu, D. S. Gary, J. Mann, Y.-C. Lin, D. Mills, L. Cooper, J. C. Andrews, Y. C. Manabe, A. Pekosz, C. K. Cooper, E. Marra, B. S. Graham, J. S. McLellan, Science 2020, 367, 498–502.
[4] J. Dinnes, J. J. Deeks, A. Adrian, S. Berhane, C. Davenport, S. Dittrich, D. Empederador, Y. Tukwoingi, J. Cunningham, S. Beese, J. Dretzke, L. Yang, J. Wang, H. Wang, L. Yang, J. Sci. Adv. 2021, 7, eabh2848.