Author(s): Mironov, Vladimir; Shchugoreva, Irina A.; Artyushenko, Polina V.; Morozov, Dmitry; Borbone, Nicola; Oliviero, Giorgia; Zamay, Tatiana N.; Moryachkov, Roman V.; Kolovskaya, Olga S.; Lukyanenko, Kirill A.; Song, Yangling; Merkuleva, Iuliia A.; Zabluda, Vladimir N.; Peters, Georgy; Koroleva, Lyudmila S.; Veprintsev, Dmitry V.; Glazyrin, Yury E.; Volosnikova, Ekaterina A.; Belenkaya, Svetlana V.;

Title: Structure- and Interaction-Based Design of Anti-SARS-CoV-2 Aptamers

Year: 2022

Version: Accepted version (Final draft)

Please cite the original version:

Mironov, V., Shchugoreva, I. A., Artyushenko, P. V., Morozov, D., Borbone, N., Oliviero, G., Zamay, T. N., Moryachkov, R. V., Kolovskaya, O. S., Lukyanenko, K. A., Song, Y., Merkuleva, I. A., Zabluda, V. N., Peters, G., Koroleva, L. S., Veprintsev, D. V., Glazyrin, Y. E., Volosnikova, E. A., Belenkaya, S. V., . . . Kichkailo, A. S. (2022). Structure- and Interaction-Based Design of Anti-SARS-CoV-2 Aptamers. Chemistry : A European Journal, 28(12), Article e202104481. https://doi.org/10.1002/chem.202104481

Chemistry
A European
Journal

Chemistry Europe
European Chemical Societies Publishing

Accepted Article

Title: Structure and Interaction Based Design of Anti-SARS-CoV-2 Aptamers

Authors: Vladimir Mironov, Irina A. Shchugoreva, Polina V. Artyushenko, Dmitry Morozov, Nicola Borbone, Giorgia Oliviero, Tatiana N. Zamay, Roman V. Moryachkov, Olga S. Kolovskaya, Kirill A. Lukyanenko, Yangling Song, Iuliia A. Merkuleva, Vladimir N. Zabluda, Georgy Peters, Lyudmila S. Koroleva, Dmitry V. Veprintsev, Yury E. Glazyrin, Ekaterina A. Volosnikova, Svetlana V. Belenkaya, Tatiana I. Esina, Anastasiya A. Isaeva, Valentina S. Nesmeyanova, Daniil V. Shanshin, Anna N. Berlina, Nadezhda S. Komova, Valery A. Svetlichnyi, Vladimir N. Silnikov, Dmitriy N. Shcherbakov, Galina S. Zamay, Sergey S. Zamay, Tatyana Smolyarova, Elena P. Tikhonova, Kelvin H.-C. Chen, U-Ser Jeng, Gerolama Condorelli, Vittorio de Franciscis, Gerrit Groenhof, Chaoyong Yang, Alexander A. Moskovsky, Dmitri G. Fedorov, Felix N. Tomilin, Weihong Tan, Yuri Alexeev, Maxim V. Berezovski, and Anna S Kichkailo

WILEY-VCH

To be cited as: *Chem. Eur. J.* **2022**, e202104481

Link to VoR: https://doi.org/10.1002/chem.202104481

10.1002/chem.202104481

WILEY-VCH

RESEARCH ARTICLE

Structure and Interaction Based Design of Anti-SARS-CoV-2 Aptamers

Dr. Vladimir Mironov[+],[a,α] Irina A. Shchugoreva[+],[b,c,d] Dr. Polina V. Artyushenko[+],[b,c,d] Dr. Dmitry Morozov[+],[e] Prof. Nicola Borbone,[f] Prof. Giorgia Oliviero,[g] Prof. Tatiana N. Zamay,[d] Roman V. Moryachkov,[b,h] Dr. Olga S. Kolovskaya,[b,d] Kirill A. Lukyanenko,[b,d,i] Prof. Yanling Song,[j] Iuliia A. Merkuleva,[k] Dr. Vladimir N. Zabluda,[h] Georgy Peters,[l] Dr. Lyudmila S. Koroleva,[m] Dr. Dmitry V. Veprintsev,[b] Dr. Yury E. Glazyrin,[b,d] Dr.Ekaterina A. Volosnikova,[k] Svetlana V. Belenkaya,[k] Tatiana I. Esina,[k] Anastasiya A. Isaeva,[k] Valentina S. Nesmeyanova,[k] Daniil V. Shanshin,[k] Dr. Anna N. Berlina,[n] Dr. Nadezhda S. Komova,[n] Dr. Valery A. Svetlichnyi,[o] Prof. Vladimir N. Silnikov,[m] Dr. Dmitriy N. Shcherbakov,[k,p] Dr. Galina S. Zamay,[b,d] Dr. Sergey S. Zamay,[q] Tatyana Smolyarova,[b,i] Prof. Elena P. Tikhonova,[r] Prof. Kelvin H.-C. Chen,[s] Prof. U-Ser Jeng,[t,β] Prof. Gerolama Condorelli,[g,γ] Prof. Vittorio de Franciscis,[u] Prof. Gerrit Groenhof,[e] Prof. Chaoyong Yang,[j,v] Dr. Alexander A. Moskovsky,[a] Dr. Dmitri G. Fedorov,[w] Dr. Felix N. Tomilin,[c,h] Prof. Weihong Tan,[v,x] Dr. Yuri Alexeev,[y] Prof. Maxim V. Berezovski,*[z] and Dr. Anna S. Kichkailo*[b,d].

[a] Department of Chemistry, Lomonosov Moscow State University, Moscow, 119991 (Russia)
E-mail: vmironov@lcc.chem.msu.ru
[b] Laboratory for Digital Controlled Drugs and Theranostics, Federal Research Center KSC SB RAS, Krasnoyarsk, 660036 (Russia)
E-mail: kichkailo.as@ksc.krasn.ru
[c] School of Non-Ferrous Metals and Materials Science, Siberian Federal University, Krasnoyarsk, 660041 (Russia)
E-mail: shchugorevai@mail.ru
[d] Laboratory for Biomolecular and Medical Technologies, Krasnoyarsk State Medical University, Krasnoyarsk, 660022 (Russia)
E-mail: azamay@krasgmu.ru
[e] Nanoscience Center and Department of Chemistry, University of Jyväskylä, Jyväskylä, 40014 (Finland)
E-mail: dmitry.morozov@jyu.fi
[f] Department of Pharmacy, University of Naples Federico II, Naples, 80138 (Italy)
E-mail: nicola.borbone@unina.it
[g] Department of Molecular Medicine and Medical Biotechnologies, University of Naples Federico II, Naples, 80131 (Italy)
E-mail: giorgia.oliviero@unina.it
[h] Laboratory of Physics of Magnetic Phenomena, Kirensky Institute of Physics, 660012 (Russia)
E-mail: zvn@iph.krasn.ru
[i] School of Fundamental Biology and Biotechnology, Siberian Federal University, Krasnoyarsk, 660041 (Russia)
E-mail: kirill.lukyanenko@gmail.com
[j] Department of Chemical Biology, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, 361005 (China)
E-mail: ylsong@xmu.edu.cn
[k] State Research Center of Virology and Biotechnology VECTOR, Rospotrebnadzor, Koltsovo, 630559 (Russia)
E-mail: dnshcherbakov@gmail.com
[l] NRC Kurchatov Institute, Moscow, 117 259 (Russia)
E-mail: georgspeters@gmail.com
[m] Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 (Russia)
E-mail: v.silnikov@mail.ru
[n] Laboratory of Immunobiochemistry, A.N. Bach Institute of Biochemistry, Research Center of Biotechnology of the Russian Academy of Sciences Moscow, 119071 (Russia)
E-mail: anberlina@yandex.ru
[o] Laboratory of Advanced Materials and Technology, Tomsk State University, Tomsk, 634050 (Russia)
E-mail: v_svetlichnyi@bk.ru
[p] Altai State University, Barnaul, 656049 (Russia)
E-mail: dnshcherbakov@gmail.com
[q] Department of Molecular Electronics, Federal Research Center KSC SB RAS, Krasnoyarsk, 660036 (Russia)
E-mail: sergey-zamay@yandex.ru
[r] Department of Infectious Diseases and Epidemiology, Krasnoyarsk State Medical University, Krasnoyarsk, 660022 (Russia)
E-mail: tihonovaep@mail.ru
[s] Department of Applied Chemistry, National Pingtung University, Pingtung 900391 (Taiwan)
E-mail: kelvin0324@gmail.com
[t] National Synchrotron Radiation Research Center, Hsinchu Science Park, Hsinchu 30076 (Taiwan)
E-mail: usjeng@nsrrc.org.tw

[u] Institute of Genetic and Biomedical Research (IRGB), Research National Council (CNR), Milan, 09042 (Italy)
E-mail: defranci@unina.it
[v] Institute of Molecular Medicine, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200127 (China)
E-mail: cyyang@xmu.edu.cn
[w] Research Center for Computational Design of Advanced Functional Materials, National Institute of Advanced Industrial Science and Technology, Tsukuba, 305-8560 (Japan)
E-mail: d.g.fedorov@aist.go.jp
[x] College of Chemistry and Chemical Engineering, Hunan University, Changsha, 410082 (China)
E-mail: tan@hnu.edu.cn
[y]Computational Science Division, Argonne National Laboratory, Lemont, IL, 60439 (USA)
E-mail: yuri@alcf.anl.gov
[z]Department of Chemistry and Biomolecular Sciences, Ottawa, Ontario K1N6N5 (Canada)
E-mail: maxim.berezovski@uottawa.ca
[α] Department of Chemistry, Kyungpook National University, Daegu, 702-701 (South Korea)
E-mail: vmironov@lcc.chem.msu.ru
[β] Department of Chemical Engineering, National Tsing Hua University, Hsinchu 30013 (Taiwan)
E-mail: usjeng@nsrrc.org.tw
[γ] IRCCS Neuromed-Istituto Neurologico Mediterraneo Pozzilli, Via Atinense 18, 86077 Pozzilli (Italy)
E-mail: gecondor@unina.it

[+] These authors contributede qually to this work.
* Correspondence may be addressed to A.S.K. Tel: +7903-923-84-02; Email: azamay@krasgmu.ru, annazamay@yandex.ru, Y.A. yuri@alcf.anl.gov, and M.V.B. maxim.berezovski@uottawa.ca

Abstract

Aptamer selection against novel infections is a complicated and time-consuming approach. Synergy can be achieved by using computational methods together with experimental procedures. This study aims to develop a reliable methodology for a rational aptamer *in silico et vitro* design. The new approach combines multiple steps: (1) Molecular design, based on screening in a DNA aptamer library and directed mutagenesis to fit the protein tertiary structure; (2) 3D molecular modeling of the target; (3) Molecular docking of an aptamer with the protein; (4) Molecular dynamics (MD) simulations of the complexes; (5) Quantum-mechanical (QM) evaluation of the interactions between aptamer and target with further analysis; (6) Experimental verification at each cycle for structure and binding affinity using small-angle X-ray scattering, cytometry, and fluorescence polarization. Using a new iterative design procedure, Structure and Interaction Based Drug Design (SIBDD), a highly specific aptamer to the receptor-binding domain of the SARS-CoV-2 spike protein, was developed and validated.

The SIBDD approach enhances speed of the high-affinity aptamers development from scratch, using a target protein structure. The method could be used to improve existing aptamers for stronger binding. This approach brings to an advanced level the development of novel affinity probes, functional nucleic acids. It offers a blueprint for the straightforward design of targeting molecules for new pathogen agents and emerging variants.

Introduction

The current COVID-19 pandemic has made it clear that virus outbreaks can still be an enormous threat to society despite advances in science and technology. While the challenges in dealing with the outbreak and its profound effect on our lives demand a coordinated multidisciplinary effort, the infection itself is a molecular process that can be understood by joint efforts of natural scientists. In particular, structural biology, organic synthesis, and computational chemistry can shed light on the molecular details of host recognition (the first step in the infection mechanism) and facilitate the development of efficient means to prevent that recognition and protect the population against COVID-19. COVID-19 is caused by a coronavirus, which is a positive-sense single-stranded RNA virus. Whereas coronaviruses mostly infect animals and cause only mild respiratory infections in humans, there have been notable exceptions, including the lethal severe acute respiratory syndrome (SARS)-CoV-1 endemic in 2003, the Middle East Respiratory Syndrome (MERS) endemic in 2012,[1] and now the COVID-19 pandemic.

The structural proteins of SARS viruses are the following: the spike (S) protein, the envelope (E) protein, the membrane (M),

and the nucleocapsid (N) proteins.[2] The SARS- CoV-2 S-protein is important in virus infection because it mediates the viral entry into the host cells by binding to a host receptor via the receptor-binding domain (RBD). Therefore, the S-protein is considered the primary target for neutralizing antibodies, antiviral attachment, entry inhibitors, and vaccine development.[1]

For the SARS-CoV-2 S-protein pseudovirus system, it is known that: (I) human angiotensin-converting enzyme 2 (hACE2) is the receptor for SARS-CoV-2 binding,[3] (II) SARS-CoV-2 enters 293/hACE2 cells mainly through endocytosis, (III) PIKfyve, TPC2, and cathepsin L are critical for the entry, and (IV) the S-protein in SARS-CoV-2 is less stable than in SARS-CoV-1.[4] The second co-receptor promoting SARS-CoV-2 entry into host cells is the tyrosine-protein kinase receptor UFO (AXL), which specifically interacts with the N-terminal domain of SARS-CoV-2 spike.[5]

Monoclonal and polyclonal antibodies, developed previously against the RBD in SARS-CoV, demonstrate only a limited binding to SARS-CoV-2.[3,4,6] Poor bio-availability, high cost, and a long time needed to develop and manufacture monoclonal antibodies, raises the necessity of a new methodology for rapid design of drugs, capable of binding selectively to a desirable epitope of a protein target with a known sequence.

On the other hand, nucleic acid oligonucleotides can also form multiple three-dimensional (3D) structures, some of which can fit a given protein target. Aptamers are artificial single-stranded RNA or DNA molecules consisting of 15–100 nucleotides. They can be designed to bind with high affinity and specificity to target proteins.[7,8] The mechanism of their interaction with the target is similar to that of antibodies. It can be characterized as hydrophobic, electrostatic, hydrogen bonding, and van der Waals interactions, as well as base stacking and shape complementarity.[9]

Aptamers have a high potential in diagnostics, therapeutics, and drug delivery systems.[10–16] However, they also could be used as antiviral drugs urgently needed during outbreaks of infections. Thus, efficient methods for rapidly screening sequences capable of blocking viruses binding to host cells should be developed beforehand.

Nowadays, aptamers are designed primarily through the SELEX procedure (Systematic Evolution of Ligands by Exponential Enrichment).[7,8] SELEX is used to generate target-specific aptamers,[17] whereas high-throughput sequencing technologies with the corresponding bioinformatics tools are used to deconvolute the results of SELEX pipelines.[18] Several SELEX variants have been suggested for aptamer optimization.[19–23] Recently, alternative approaches for designing aptamers that rely on *in silico* analyses have been proposed.[24–29] A combination of docking and classical molecular dynamics (MD) simulations are often used to sample possible aptamer–target binding poses.[30–33] Several computational techniques were used to predict the binding energies and interactions in the aptamer-virus complexes: interaction analysis,[34] docking,[35] and MD-based free energy calculations.[36] These methods have been proven to be a crucial part of the most modern pipelines for drug design.

Here, we present a new methodology for aptamer's rational design based on their structure and interaction with the target. This new methodology, called SIBDD (Structure and Interaction Based Drug Design), can be considered as an extension of SBDD (Structure-Based Drug Design),[37] enhanced with the feedback from classical molecular dynamics simulations combined with quantum-chemical interaction energies evaluation. SIBDD is a powerful combination of computational screening simulations and directed mutagenesis with experimental validation at each cycle to obtain oligonucleotides with a high affinity and selectivity.

SIBDD methodology is generally designed for cases with limited information about the target. For example, it can be used for rapid response to novel, dangerous infections when the access to the samples is minimal. While it can take several months to get the first experimental structure of the target, initial rounds of SIBDD can be performed based on computational folding results

10.1002/chem.202104481

obtained from the target's primary sequence. This initial set of aptamers could be further improved when more experimental data become available. Finally, the aptamers can be adjusted afterward for various pathogen mutations.

This work demonstrated how the SIBDD approach could be applied to produce an aptamer binding to the RBD of the SARS-CoV-2 spike protein. Several RBD-binding aptamers have been designed, and their affinity has been experimentally validated and confirmed.

The pipeline for aptamers selection

The preliminary step in the SIBDD methodology (Figure 1) is computational aptamer library design, which can be found in Github repository (https://github.com/MolecularDesign/Aptamer_library_16nt). The structures of aptamer candidates and the target (RBD of SARS-CoV-2 spike protein) are determined by molecular modeling. Next step is a selection of aptamers from the created library by molecular docking to target protein. The sequence with the best affinity predicted by docking is selected for the iterative improvement in the pipeline (Figure 1). The electrostatic potential of the protein is computed and used for the docking in each iteration of candidate structures to the RBD. Next, molecular dynamics (MD) simulations are performed on the aptamer-protein complexes constructed in docking. A clustering of the MD trajectories provides a set of the most populated molecular conformations in the dynamics. In these conformations, the aptamer-protein interactions are analyzed in detail with quantum mechanical (QM) calculations based on the fragmentation approach called fragment molecular orbitals (FMO) method. The fragmentation of proteins and aptamers into smaller parts, amino acid residues, and nucleotides, respectively, makes it possible to obtain a detailed map of nucleotide-residue interactions with QM calculations by using pair interaction electrostatic decomposition analysis (PIEDA). FMO structure optimization of the protein-aptamer complexes are performed with the DFTB method, while PIEDA calculations are performed at the DFTB and RI-MP2 levels of theory. The obtained interaction energies are not expected to be quantitatively precise but sufficiently indicate which nucleotides bind strongly and to which residues, making possible to use such information in the following steps of SIBDD. They can be viewed as a qualitative measure to guide the aptamer selection process.

After *in silico* modeling, experimental validation of the most promising candidate structures binding affinities towards target are performed. The most effective aptamers will be advanced for the further selection rounds. As a result, the best binding aptamer is revealed by experimental and theoretical methods. In this work, the synthesized aptamers were tested against recombinant RBD protein using flow cytometric analysis. Fluorescence polarization assay experiments demonstrated the binding kinetics of the aptamers to RBD in comparison with random sequences. In addition, small-angle X-ray scattering (SAXS) experiments verified whether an aptamer binds to the RBD with a 1:1 ratio and forms a stable complex.

10.1002/chem.202104481

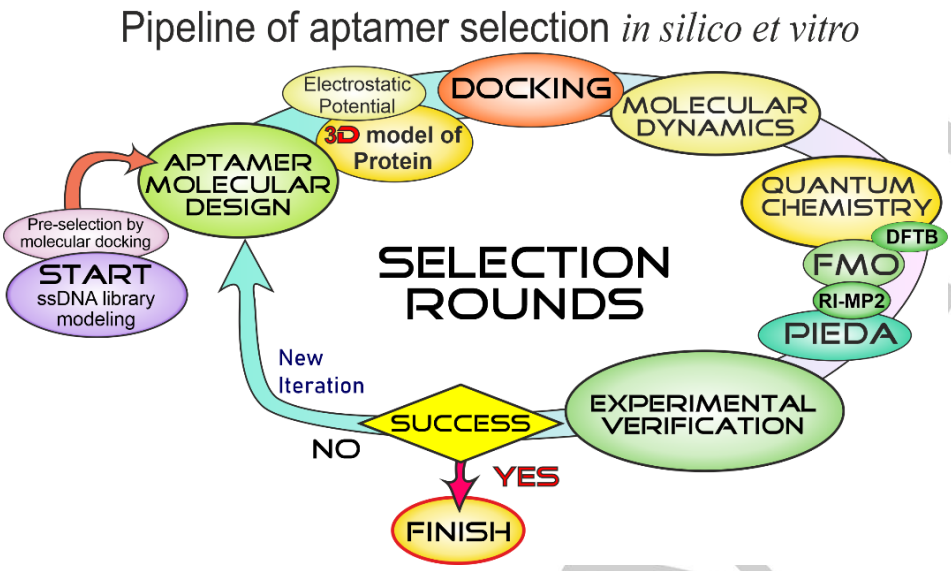


Figure 1. Structure and Interaction Based Drug Design (SIBDD) is an iterative procedure consisting of the following steps: several rounds of aptamer molecular design (screening of a DNA library and directed mutagenesis for fitting with the protein tertiary structure); 3-D modeling of the protein target and calculating its electrostatic potential; molecular docking of aptamers to the target; molecular dynamics simulations of the complexes; quantum-mechanical analysis of nucleotide-residue interactions using methods such as DFTB, FMO, RI-MP2, and PIEDA; and finally experiment verification of binding

After completing the whole procedure (Figure 1), a decision is made as to whether a desired level of binding is achieved. If it is not, then, based on the feedback from the stability of aptamer-protein complexes in MD and nucleotide-residue interactions, combined with the experimental validation, a new modified set of aptamers can be made and the whole procedure repeated.

The SIBDD protocol in this study was applied to a search for an aptamer with a high binding affinity towards the RBD of SARS-CoV-2 S-protein. Nevertheless, the proposed protocol is general and can be applied to a wider range of ligands and protein targets. It should be noted that the individual steps of the SIBDD approach are very well parallelizable, making it suitable for use within continuous high-performance workflows in the future.

Results and Discussion

Structure of the SARS-CoV-2 spike receptor-binding domain

RBD is a flexible region on the top of the coronavirus spike that permits virus anchoring to the human ACE2 receptor on the cell surface.[38] The structure of the SARS- CoV-2 spike protein is shown in Figure 2a. This protein is a trimer, comprising three identical monomers (Figure 2b). Each monomer contains an RBD (Figure 2c) and can bind to human cells.

The RBD model used in the present work was taken from the PDB entry 6M0J of the RBD complex with the ACE2 human receptor.[39] The most important residues for binding human ACE2 are reported[40] to be Asn439, Asn501, Gln493, Gly485, and Phe486. The RBD part of the spike protein has a charge of +2e, binding to the ACE2 protein with a charge of 25e. The electrostatic potential of RBD was averaged over 70 snapshots taken at 10 ns intervals from a 700 ns MD trajectory. The potential plotted on a surface (Figure 2 e.1-e.3) shows multiple charged regions. The outer side of the RBD domain (Figure 2 e.3) is mostly positively charged ("outer" refers to facing solvent and possible host cells). The opposite "inner" RBD side (Figure 2 e.1) facing the viral protein has two charged regions of opposite signs. The "top" (Figure 2 e.2) part of the RBD domain, which can interact with the ACE2 receptor, is only slightly negatively charged. Positively charged regions on the RBD surface are primary targets for binding

10.1002/chem.202104481

aptamers containing negatively charged phosphate groups in the double-stranded DNA helix.

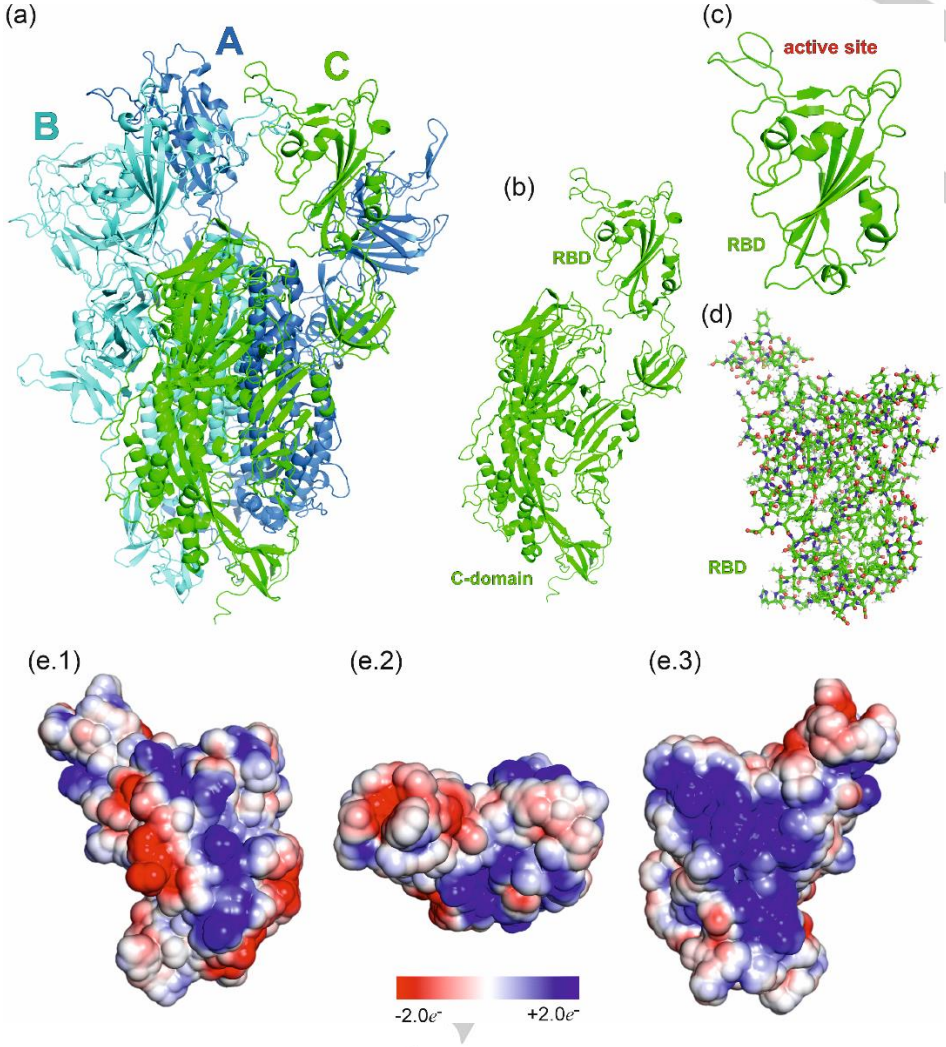


Figure 2. Structure and properties of SARS-CoV-2 spike protein RBD. (a) SARS-CoV-2 spike protein trimer, monomers are colored in blue, cyan, and green; (b) receptor-binding domain (RBD) position within the spike protein monomer; (c) active site of an RBD; and (d) an atomic model of an RBD, where C atoms are green, N atoms are blue, O atoms are red, S atoms are yellow, and H atoms are grey. Electrostatic potential of RBD is computed with APBS and plotted on a surface surrounding the protein: (e.1) inner view (i.e., facing the virus side), (e.2) top view, and (e.3) outer view (i.e., facing a host cell). Blue means positive potential value; red is negative. The total charge of RBD is +2; thus, the positive potential prevails, especially its side which is exposed to solution

Thus, the existence of a double-stranded region in the conformation of aptamers may be crucial for their binding to the spike RBD, and the length of the duplex tract can be tuned to enhance the binding. Both positively charged outer and inner parts of the RBD can be a target to the aptamer stem helix. Another potential binding target is the top of the RBD. It is formed with mostly neutral polar amino acids such as Asn, Gln, Ser, Thr, and Tyr. In particular, Asn and Gln are well known as hydrogen-bonding partners[41] of nucleobases in protein complexes. A non-duplex part of the aptamer can be tuned to bind to the top side of RBD. Thus, the electrostatic potential map provided two valuable hints for aptamer design. It should have a double-stranded part

10.1002/chem.202104481

binding the outer or inner RBD and a non-duplex domain binding the top of the RBD.

Aptamer design steps

Creating an aptamer library

In this study, aptamer molecular modeling is performed using a novel *in silico* selection approach without any preliminary experiments. In general, *in silico* aptamer modeling can be divided into two stages:[25,42] (I) searching oligonucleotide sequences, which form the part of the aptamer responsible for selective binding, and (II) constructing a structure of the aptamer that maintains the optimal conformation of that binding part. Considering the binding interface of the RBD (Figure 2 e.2) as the primary target for the aptamer, a hairpin was chosen as the initial shape of the aptamer for docking with the concave upper region of the RBD protein. To obtain the initial library of aptamer candidates, a 16mer sequence 5'-GGAATT NNNN AATTCC-3' was designed to have a central variable "NNNN" loop of four nucleotides and two constant self-complementary regions at the 5' and 3' ends (Figure 3a). All possible combinations of the DNA nucleotides for the variable region have been considered in this study, resulting in the initial aptamer library of 4^4 = 256 entries. The double-stranded stem part (5'-GGAATT-... and… -AATTCC-3') improves the aptamer affinity to positively charged RBD parts and maintains the aptamer structure.

The molecular docking for all 256 models in the 16mer aptamer library (Figure 3a) was performed following the *in silico* selection scheme in Figure 1. For every aptamer, ten binding sites were predicted. The top side of the RBD is mostly exposed to the receptor (Figure 2 e.2), and is presumed to be responsible for recognizing the ACE2 protein by the virus. The molecule variant that was predicted to have the most binding sites located in the upper part of the RBD (Figure 3c, Figure S1a) was selected. That aptamer, having the AGTC sequence in the variable region (Figure 3b), is denoted as Apt16. It was used as the starting sequence for further modifications to improve the binding affinity to the key amino acid residues of the RBD reported by Lu et. al.[40]

Simulations of Apt16-RBD complex

The procedure described below for Apt16 was also applied to the other aptamers. As mentioned above, Apt16 was selected as an initial sequence, and its binding sites with the protein were predicted by docking with RBD (Figure 3c, Figure S1a). An MD simulation of the RBD/Apt16 complex was performed to refine the binding pose of this complex further. These MD simulations were done with the GROMACS software package for at least 200 ns. After MD simulation, the conformers clustering showed that Apt16 mostly binds to the top part of RBD (Figure 3 d,e).

From all clusters of conformations, the two with the highest populations in the MD simulations were chosen for further analysis. It was found that Apt16(1) (Figure 3d) mainly binds RBD with its 3'-ending nucleotides. For Apt16(2) conformation (Figure 3e), the nucleotides located both in the loop and in the stem part participate in the binding.

Next, a more detailed semi-quantitative analysis of residue-nucleotide interaction energies was performed. The MD structures of RBD/aptamer complexes were optimized at the semi-empirical density functional-based tight-binding (DFTB)[43] level. The FMO fragmentation approach was used for the optimized structures to get the interaction energies, which were evaluated with the non-empirical RI-MP2/6-31G(d,p) method. In all these simulations, the solvent environment was described by the polarizable continuum model (PCM).

The total interaction enthalpies, E_{total}, between RBD and aptamer (Table 1) were obtained as a sum of residue-nucleotide pair interaction energies (PIE) at both DFTB and RI-MP2 levels of theory. PIE values allow us to compare the partial contribution to the binding energy from a single nucleotide in the aptamer and estimate the effect of point mutations. Thus, they can be used to systematically improve the aptamer sequence for better binding to the target molecule. It should be noted that because we did not

10.1002/chem.202104481

attempt to account for the entropy contributions into PIE, E_{total} only includes binding enthalpies. However, entropy may play a significant role in the binding, thus making experimental checks important after each SIBDD cycle.

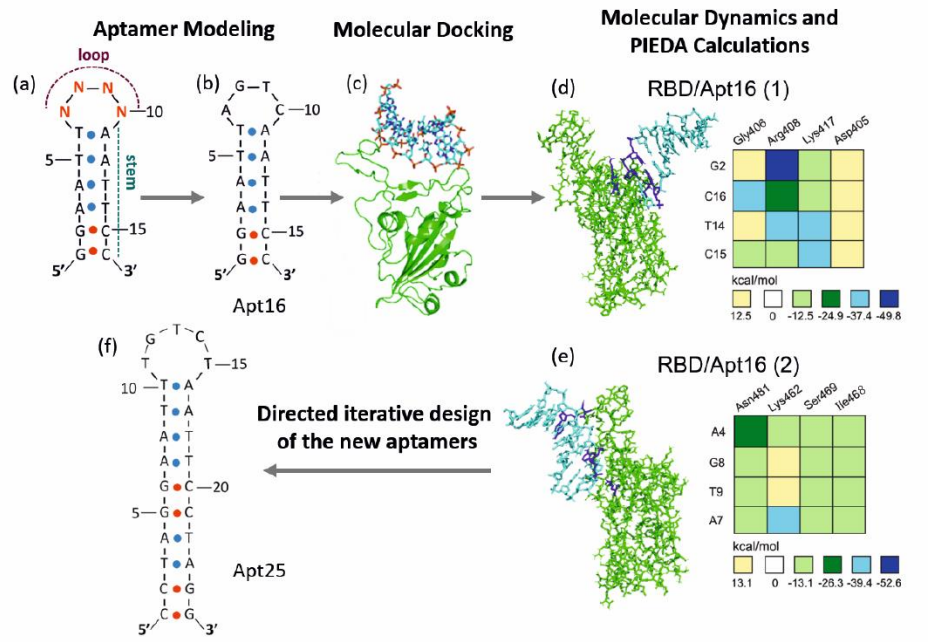


Figure 3. Molecular modeling and interaction results. (a) The secondary structure of initial aptamers, (b) aptamer Apt16 and (f) aptamer Apt25. (c) Apt16/RBD complex obtained from molecular docking. Molecular structure and QM nucleotide-residue interactions for two isomers of RBD/Apt16 complexes, (d) Apt16(1) and (e) Apt16(2). RBD is in green, and aptamers are cyan; the main interacting amino acids and nucleotides are in dark blue. The principal QM interactions are shown as 2-D maps, with nucleotides (G2, etc.) and residues (Gly406, etc.) labeling the heat map; dark blue, light blue, green, and yellow indicate attraction in the decreasing order.

For example, for the RBD/Apt16(1) and RBD/Apt16(2) complexes, the values of E_{total} are -153.0 kcal mol^{-1} and -91.6 kcal mol^{-1}, respectively, which indicates that Apt16(1) is more strongly bound to RBD compared to Apt16 (2).

The color maps in Figure 3 d, e illustrate the pairs of nucleotides and amino acids whose contribution to the binding energy is the largest. The binding in the RBD/Apt16(1) complex is largely driven by G2, C15, T14, and C16 stem nucleotides, whereas in RBD/Apt16(2) the binding is mostly due to A7, G8, and T9 loop nucleotides. Participation of the duplex base pairs in the binding is not desirable because this region can be used to enhance an aptamer later with functional groups such as fluorescent labels. Thus, the strategy to improve the binding affinity of Apt16 to RBD was to modify the aptamer to increase interactions involving the loop nucleotides rather than nucleotides of the duplex stem.

Table 1. Total interaction energy E_{total} (kcal mol^{-1}) and the number of hydrogen bonds N_{HB} between aptamers and RBD. Energies were computed using FMO2/RI-MP2 method. FMO2/DFTB3 energies are provided for comparison.

Aptamer[a]	N_{HB}	E_{total}	
		FMO2/RI-MP2	FMO2/DFTB3
Apt16(1)	12	-153.0	-100.9
Apt16 (2)	9	-91.6	-65.1
Apt25 (1)	11	-159.2	-107.8
Apt25 (2)	12	-162.7	-112.9
Apt27 (1)	14	-200.1	-132.8
Apt27 (2)	17	-217.4	-157.9
Apt31 (1)	13	-221.6	-167.9
Apt31 (2)	9	-152.2	-108.3

[a] Different conformer are labeled in parentheses.

Iterative design of aptamers for binding with RBD

Experimental flow cytometry results (Figure 6a) showed for Apt16 a weaker binding than the random sequence of 40 nucleotides. Thus, to improve the aptamer binding affinity and specificity to the SARS-CoV-2 RBD domain, the initial sequence of Apt16 was modified. Based on the results from the MD simulations and quantum chemistry calculations discussed above, both the duplex and loop regions of the Apt16 aptamer were extended. First, to strengthen the complementary duplex structure, the stem regions were elongated from six to ten base pairs on both the 3' and 5' ends, resulting in the new hairpin-forming template 5'-CCTAGGAATT-loop-AATTCCTAGG-3'.

Secondly, in RBD/Apt16 complexes, the interactions with the protein were mainly due to guanines and cytosines and, to a much lesser extent, to adenines. The reports[44,45] indicate that the asparagine and the glutamate, exposed on the RBD surface, preferentially bind thymine to the other bases. We replaced the adenine with thymine in the loop to increase the binding affinity and added one additional thymine at the loop 3' end. As a result of these modifications, a new aptamer Apt25 5'-CCTAGGAATT-TGTCT-AATTCCTAGG-3' (Figure 3f) was created, and it went through the *in silico* aptamer selection scheme in Figure 3, in the same way as Apt16.

Comparing the calculation results for RBD/Apt16 and RBD/Apt25 revealed that the aptamer structure modifications improve the binding. Table 1 shows that the total interaction energy E_{total} for RBD/Apt25 complexes is more attractive than for the two RBD/Apt16 complexes. The duplex segment of Apt25 was stable during MD simulations. In the RBD/Apt25 complexes, strong electrostatic interactions between negatively charged phosphate groups of the aptamer duplex backbone and the positively charged regions of RBD were observed in FMO PIEDA analysis. Compared to RBD/Apt16, the number of strongly interacting residue-nucleotide pairs in RBD/Apt25 is larger, including interactions with thymines of Apt25 (the maps are shown in Figure 4 a.2). The two most representative structures of RBD/Apt25 in MD simulations are shown in Figure 4 (a.3, a.5).

The total number of hydrogen bonds is almost the same as in the RBD/Apt16 complexes, but in the RBD/Apt25 clusters, they are formed mainly by T9, T11, T13, T15 thymines located both in the loop and in the stem parts. A detailed summary of the hydrogen bonds is given in SI (Table S1).

As mentioned above, the duplex function is to maintain a rigid hairpin conformation and provide electrostatic interactions with the protein. To further improve that, one of the A-T base pairs was replaced with a C-G pair to strengthen the duplex further, considering that C-G pairs are bound by three hydrogen bonds while A-T pairs only by two.[46] Because the interaction data indicated that in Apt25 the T9 nucleotide of the duplex binds to RBD (Figure 4 a.2, a.4), T9 was replaced with A9, and also the order of pairs in the central region of the duplex was changed to ensure that the specificity of the binding was due to the loop nucleotides rather than the stem. The new stem sequence was 5'- CGGATGGAAT - loop - ATTCCATCCG-3'.

On the other hand, the aptamer loop size was extended further to maximize its interaction surface with RBD and compensate for changes in the stem part. The loop was extended with an additional pair (in Apt27) and with two pairs (in Apt31) of the nucleotides. As a result, two new aptamers were created having the same duplex stem but a loop of different sizes: Apt27 (5'-stem - TTG TCTG - stem-3', Figure 4 b.1) and Apt31 (5'-stem - ACTT TGTC TGTC - stem-3', Figure 4 c.1). Aptamers Apt27 and Apt31 underwent the simulation rounds according to Figure 1. Comparative results of simulations for all aptamers, Apt16, Apt25, Apt27, and Apt31, are described in the following section.

10.1002/chem.202104481

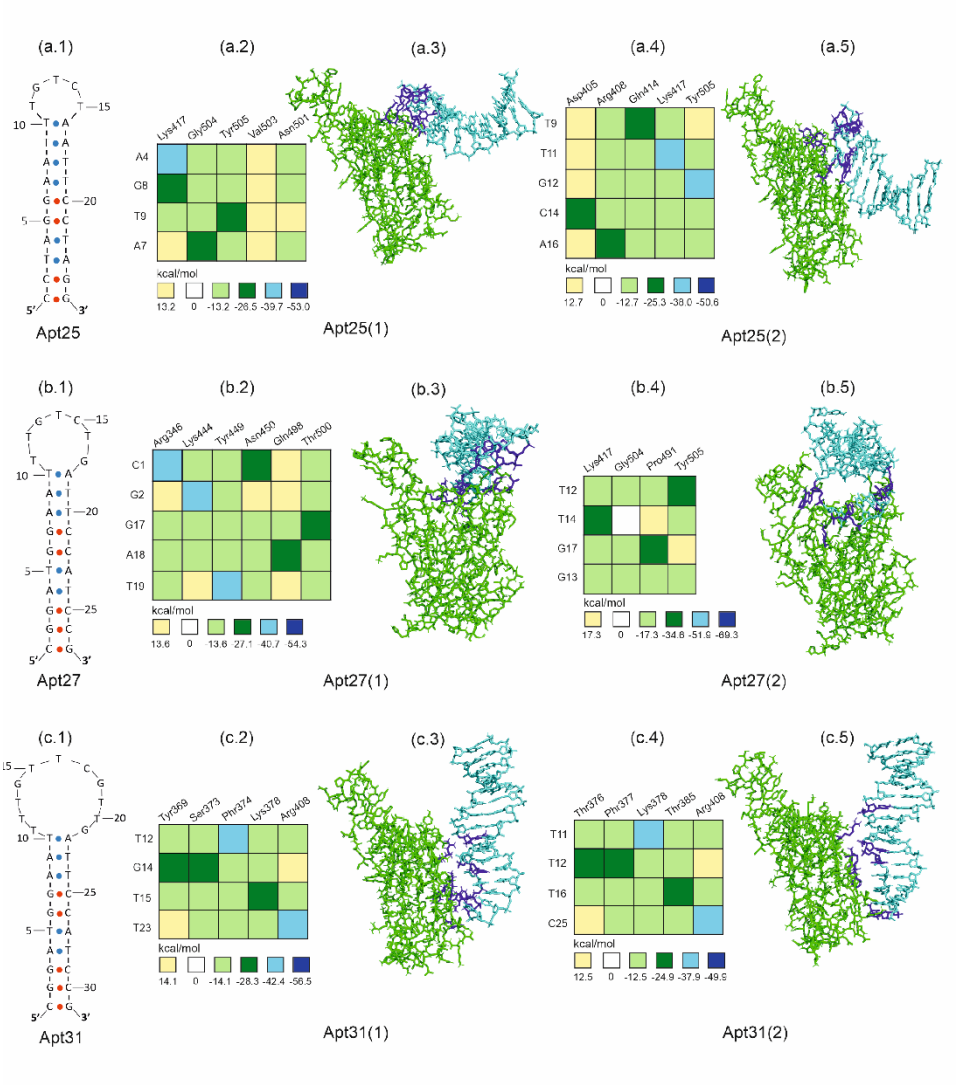


Figure 4. Molecular structure and interaction energies for RBD and aptamer complexes. RBD is shown in green, and aptamers are in cyan, except that the main interacting amino acids and nucleotides are in dark blue. (a.1), (b.1), (c.1) show the secondary structure of the Apt25, Apt27, Apt31, respectively. See also the supporting video (S.1, S.2) for details of the Apt27 and Apt31 binding. The main FMO PIEDA interactions are shown as 2-D maps, with nucleotides (A4, etc.) and residues (Lys417, etc.) labeling the heat map; dark blue, light blue, and green indicate attraction in decreasing order.

RBD/aptamer complexes: a comparative analysis

The binding of aptamers to RBD can be studied in terms of structure and interactions. For each RBD/aptamer complex, the aptamer remained bound to the RBD throughout MD simulations, adopting multiple binding sites. Apt16 and Apt27 bind primarily to the top part of the RBD, as shown in Figure 3 d,e and Figure 4 b.3, b.5, but Apt25 and Apt31 bind preferably to the sides of the RBD (Figure 4 a.3, a.5 and Figure 4 c.3, c5) driven by the electrostatic interactions between the phosphate groups of the duplex part of the aptamer and the positively charged residues of RBD.

Pairwise interaction energies between RBD residues and aptamers nucleotides have been computed to analyze the binding. The values of individual pair interactions for the primary binding contributors are shown as heat maps in Figure 4 c.2, c.4 for RBD/Apt31 clusters, and in Figure e 4 a.2, a.4, b.2, b.4 and Figure 3 d,e for other complexes. The total interaction energy E_{total} between RBD and aptamer for each complex was computed by

10.1002/chem.202104481

summing all residue-nucleotide interaction energies. The results are shown in Table 1, where more negative E_{total} values correspond to a stronger attraction. Another important feature is the number of hydrogen bonds in the RBD/aptamer complexes (Table 1). The larger the number of hydrogen bonds, the higher is aptamer binding affinity to RBD. Detailed information regarding the hydrogen bonds is presented in Supporting Information (Table S1).

One can see from Table 1 that as a result of the aptamer design, the total interaction E_{total} became stronger from –91.6 kcal mol^{-1} for Apt16(2) to –221.6 kcal mol^{-1} for Apt31(1). According to Table 1, the strongest interactions were observed for Apt27(2), showing a strong interaction of –217.4 kcal mol^{-1}. There are 17 hydrogen bonds between RBD and aptamer fragments in RBD/Apt27(2) and 13 hydrogen bonds in RBD/Apt31(1) complexes. The heat maps of interactions in Figure 4 c.2 and 4 b.4 show that in both the RBD/Apt31(1) and RBD/Apt27(2) complexes, some of the strongest interactions are due to thymines located in the loop. In the RBD/Apt27(2) complex, the largest contribution to the binding is due to pairs T12-Tyr505, T14-Lys417, and G17- Pro491, while in the RBD/Apt31(1) complex, the strongest interactions are observed for G14-Tyr369, G14-Ser373, and T15-Lys378 fragment pairs. The structures of all discussed aptamers are shown in Figure 5. In the next section, experimental studies on these aptamers are reported.

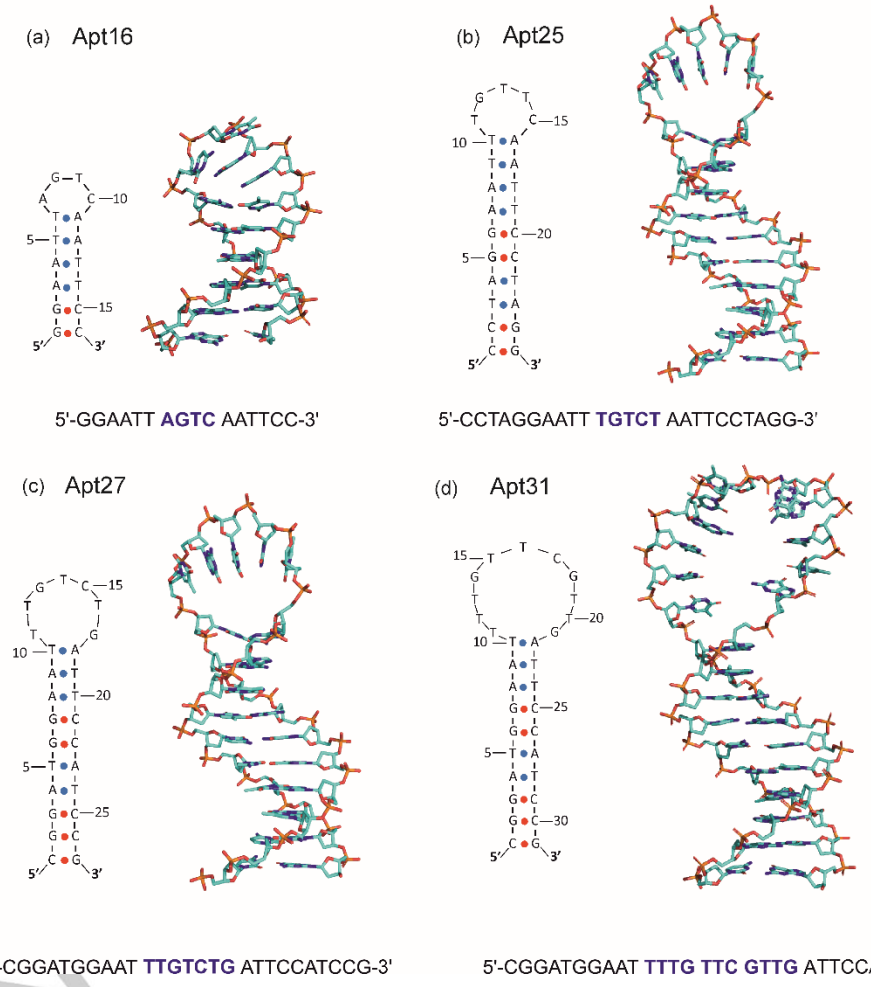


Figure 5. Secondary structures from the aptamer library and the corresponding tertiary structures optimized with FMO2-DFTB3/PCM(water). (a) Apt16, (b) Apt25, (c) Apt27, and (d) Apt31.

10.1002/chem.202104481

Experimental analyses of aptamer affinity to RBD

The experimental confirmation of aptamer binding to the spike RBD was performed by three different methods: flow cytometry, fluorescence polarization, and small-angle X-ray scattering.

Flow cytometry experiments

In silico modeling predicts different binding affinities for the aptamers with RBD. To confirm this result, flow cytometry experiments of the binding of aptamer candidates with the recombinant purified RBD peptide were performed.

Ni beads were conjugated with recombinant RBD (RBD beads) and incubated with a designed aptamer or with a randomly controlled sequence. The fluorescence intensity of RBD-coated beads bound to the oligonucleotides was then measured on a flow cytometer. RBD of SARS-CoV-2 spike protein bound with Apt31 shows a stronger fluorescence signal against RBD than the initial library and the other three candidate sequences (Figure 6 a.1, a.2).

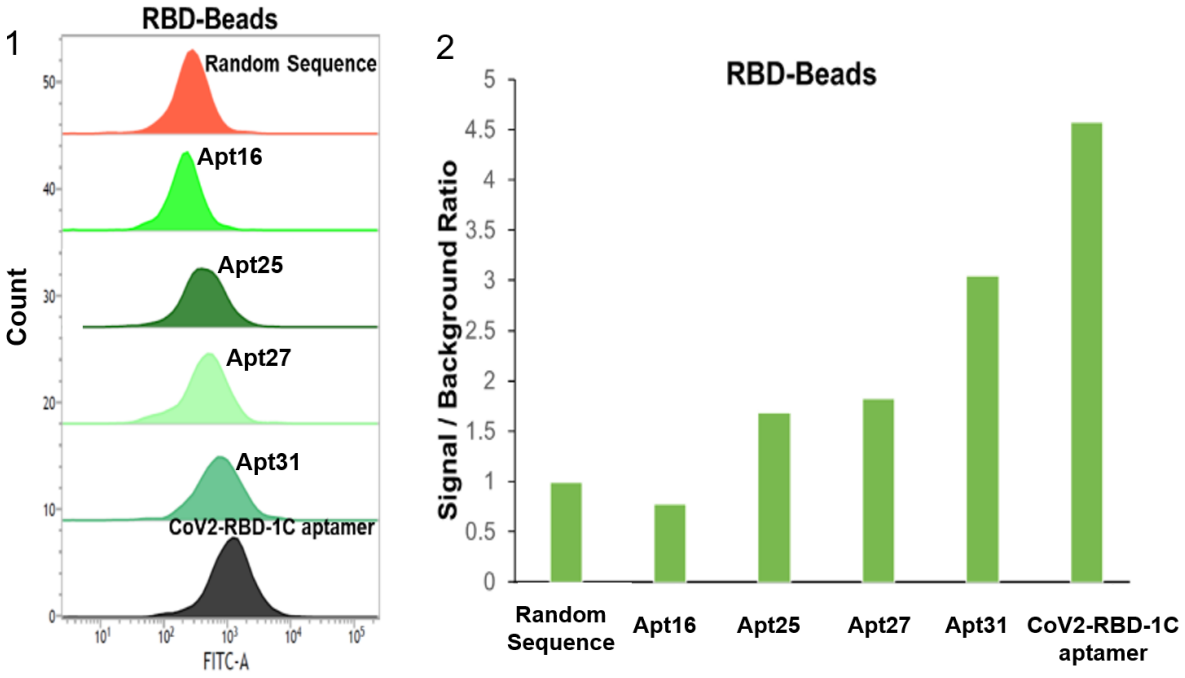
As shown in Figure 6a, the fluorescence intensity induced by all aptamers, except for Apt16, is higher than that caused by a random sequence. Moreover, the binding level of each consequent aptamer developed during the *in silico* evolution is higher than that of all others preceding it, validating the simulation-based design.

In good agreement with the interaction energies from QM calculations, the flow cytometry data confirmed that Apt31 has the strongest affinity for the RBD relative to a random DNA sequence. Furthermore, its induced fluorescence signal was comparable to that induced by RBD beads incubation with the *in vitro* selected DNA aptamer CoV2- RBD-1C [63] (Figure 6).

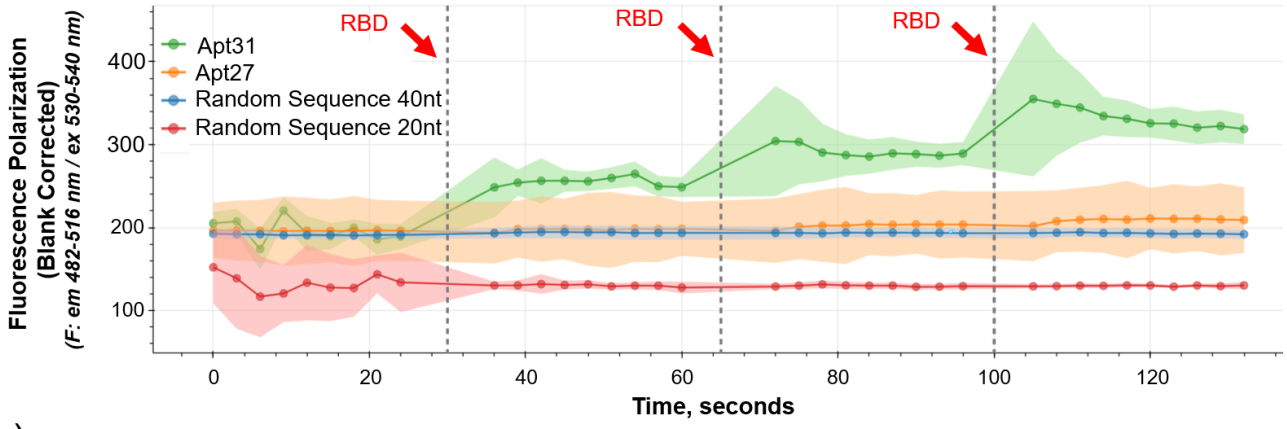
10.1002/chem.202104481

RESEARCH ARTICLE

(a)



(b)



(c)

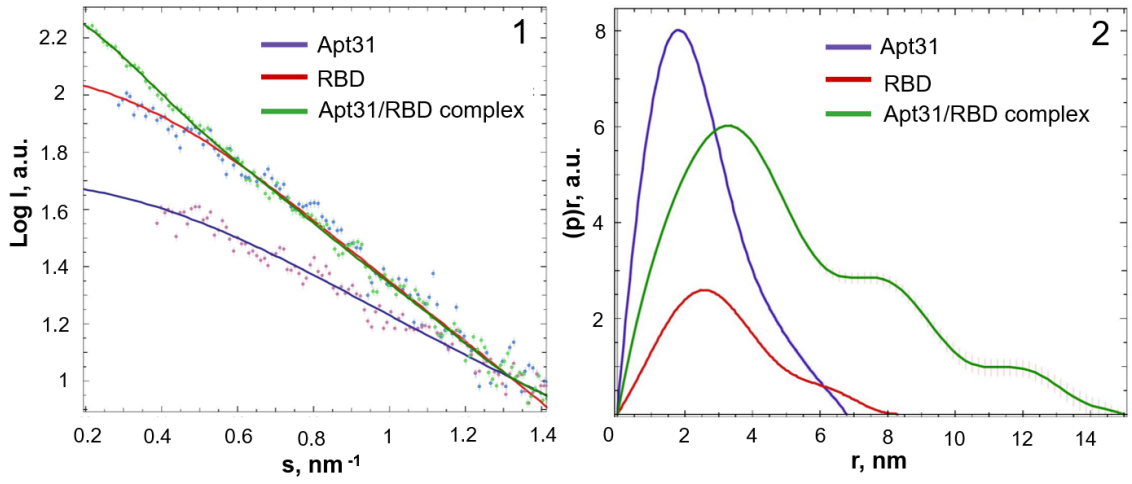


Figure 6. Experimental analyses of aptamer affinity to RBD. (a) Flow cytometric analysis of a random sequence, candidate aptamers (obtained in this work), and CoV2-RBD-1C aptamer (reported earlier (47) from an *in vitro* selection), which is shown as flow cytometric binding histograms (subfigure 1) and signal-to-background ratios of aptamers against RBD beads in the buffer (subfigure 2). (b) Fluorescence polarization analyses of aptamers binding to the recombinant RBD floating in the solution. (c) SAXS data from Apt31, RBD, and their complex. SAXS curves correspond to the aptamer, RBD, and their bound state (subfigure 1). The SAXS results for the complex indicate a larger size compared to the Apt31 and RBD molecules separately. Pair distance distribution functions p(r) of the samples in real space (subfigure 2) for Apt31 indicate a more compact prolate shape of the aptamer molecule, RBD has a larger size and a more folded form, and the Apt31-RBD complex has the largest volume.

Experimental analyses of aptamer binding to recombinant RBD measured using fluorescence polarization assays

Flow cytometry experiments confirmed binding of Apt27 and Apt31 with the recombinant RBD of the spike protein fixed on the Ni beads. To ensure that the two best candidates can bind the non-fixed RBD, an additional method, fluorescence polarization assay (FAM), was used. FAM is a highly sensitive method capable of detecting polarization changes of fluorescent molecules due to their binding with other molecules floating in the solution.

This experiment used synthetic FAM-labeled Apt27 and Apt31 and two random control sequences of 20 and 40 nucleotides in length to estimate their binding to the recombinant RBD peptide. RBD was produced based on the gene sequence encoding the 308V–542N region of the SARS-CoV-2 spike protein (GenBank MN908947). A Chinese hamster ovary (CHO) cell expression system was used to harvest this peptide. The RBD has post-translational modifications similar to the native viral RBD protein; therefore, it is a useful model for testing the binding ability of *in silico* generated candidate aptamers.

Binding kinetics measurements were performed on a Clariostar Plus microplate reader (BMG LABTECH, Germany); 10 nM of FAM-labeled aptamers Apt27, Apt31, or 10 nM control random sequences R20 (20-nucleotide long) and R40 (40-nucleotide long) were titrated with increasing concentrations of RBD peptide. As shown in Figure 6b, the fluorescence polarization of aptamer Apt31 increased from 200 to 250 mP after the first RBD injection, to 290 mP after the second injection, and to 330 mP after the third injection. Further injections did not increase fluorescence polarization, thus indicating a binding saturation (Figure 6b). The fluorescence polarization of aptamer Apt27 slightly increased after adding the RBD peptide. Conversely, the R20 and R40 control sequences did not significantly bind to RBD, thus confirming the enhancement of the aptamer candidate's binding abilities during the *in silico* selection.

Experimental analyses of aptamer binding to recombinant RBD by small-angle X- ray scattering

Using the SAXS intensity plots against the scattering angle, one can estimate the molecule sizes and form factor in the solution. The SAXS curve slope shows features of molecular shape (Figure 6c). The inverse Fourier transformation of SAXS curves reveals the pair-distance distribution in the real space of a molecule in solution.

According to the SAXS data, the RBD protein has a more folded structure and larger size than the aptamer (Dmax = 8.3 nm; the radius of gyration is 2.48 nm). Apt31 has a prolate shape with a 2 nm thickness, 6.0 ± 0.2 nm length, and radius of gyration Rg = 1.77 ± 0.07 nm. The volume of the particle is estimated as Vp = 33.91 nm-3, wich corresponds to the molecular weight MWSAXS = 8.95 (7.95-9.95) kDa. It has good accordance with the expected molecular weight of the aptamer MWexp = 9.52 kDa. The bead model was used for the molecular design of the aptamer model, and finally, the theoretical SAXS data were compared to the experimental data (Figure S2 a, b). The obtained discrepancy between molecular model and SAXS data $\chi 2$ = 1.298 shows a good coincidence in the spatial structure restoration. SAXS spectrum of an equimolar mixture of the aptamer and the RBD

protein was obtained. We observed the additional peaks with respect to individual RBD or aptamer SAXS spectra corresponding to a complex with an approximate size of 15 nm (Figure 6.2). The overall electron density volume corresponds to the molecular weight in the range of 40–52 kDa. The masses of aptamer and protein were estimated to be in ranges of 10.8-13.1 and 20.8-33.5 kDa, respectively (with 95% credibility). The SAXS dataset for Apt31 has deposited at the SASBDB[47] (https://www.sasbdb.org/draft-preview/3576/1oiq9akb5r/), accession code: SASDMU7.

Conclusions

DNA aptamers are promising ligands for the detection of SARS-CoV-2[48,49] and blocking its viral activity[50] because of their high specificity, ease of modification, and use. In this work, a novel methodology, structure- and interaction-based drug design (SIBDD), has been presented and applied to design a new 31mer DNA aptamer targeting with high affinity the RBD domain of the SARS-CoV-2 spike protein.

The selection procedure in SIBDD combines (1) of *in silico* library of aptamer structures with (2) electrostatic potential calculations to guide the docking of aptamers to RBD, (3) molecular dynamics simulations and (4) quantum chemistry-based analysis of the binding energies. The qualitative insights into the residue-nucleotide interactions obtained in this procedure are used to guide the rational modification process of the aptamers *in vitro*. The validity of the SIBDD approach has been validated with (1) flow cytometry, (2) fluorescence polarization, and (3) small-angle X-ray scattering measurements. The experimental results indicated the aptamer Apt31 as the most effective, fully supporting the SIBDD *in silico* prediction.

To the best of our knowledge, this is the first report of a systematic and rational *in silico* design of a DNA aptamer. The binding efficacy of Apt31 to the RBD has been proven using three different experimental methods. Further experimental studies are underway to assess the antiviral properties of Apt31.

Two routes can be envisaged to take advantage of the new aptamer: (1) therapeutic applications to prevent the virus from binding to human cells, and (2) diagnostic usage to detect the virus presence. Preliminary results demonstrate that the developed aptamers are promising candidates for detecting and blocking the SARS-CoV-2 virus. Furthermore, the SIBDD protocol described in this work can facilitate *in silico* the development of diagnostic and therapeutic ligands for SARS family viruses.

Given the high versatility of aptamers, the computational design of RBD-binding molecules offers a promising blueprint for the development of diagnostic and therapeutic tools for other transmissible diseases.

Methods

Molecular modeling

Molecular dynamics simulations

The X-ray structure of the complex of the SARS-Cov-2 RBD protein with human enzyme ACE2 (PDB ID: 6M0J) was used as a reference.[51] The remaining sugar residues were removed from all glycosylation sites. Protonation states of histidines and hydrogen atoms positions were assigned using the GROMACS pdb2gmx module.

The protein model was then submerged into a water box 15 nm ×15 nm ×15 nm. The negative charge of the protein complex system was neutralized with Na^+ ions. Then, Na^+ and Cl^- ions at the physiological concentration (0.15 M) were added to the system.

MD simulations were performed by using GROMACS 2019.8 software.[52] The Amber14sb[53] force field for proteins and the TIP3P model for water[54] were used. NPT MD trajectories with 200 ns length were generated at the 310 K temperature using a V-rescale thermostat with 1ps time constant[55] and at 1 bar

pressure using a Parrinello-Rahman barostat with 5ps time constant[56] and 4.5×10^{-5} bar^{-1} compressibility.

The lengths of all bonds involving hydrogen atoms were constrained to their equilibrium values, taken from the force field parameters, with the LINCS[57] method. The SETTLE algorithm was used to constrain the internal degrees of water molecules.[58] This allowed using a 2 femtoseconds time step for trajectory integration. The smooth particle mesh Ewald method[59] was used to account for periodic electrostatics with a real-space cutoff of 1.2 nm. Lennard-Jones non-bonded interactions were calculated using a cutoff of 1.2 nm with a smooth switching starting at 1.0 nm.

Molecular electrostatic potential

The structures of SARS-Cov-2 spike RBD in water for computing the potential were obtained from MD simulations by taking snapshots every 5 ns, resulting in about 70 structures. The protein atoms were taken from these snapshots and aligned by minimizing RMSD between backbone heavy-atom positions. These structures were prepared for the molecular electrostatic potential calculation using the PDB2PQR program[60] and the PARSE[61] continuum electrostatics force field.

The linearized Poisson-Boltzmann equations were solved by using a multigrid approach implemented in APBS.[62] The Na^+ and Cl^- ionic concentration of 0.15 M was used to simulate physiological conditions. The computed molecular electrostatic potential (MEP) values on a 3-D grid were then averaged over all structures and plotted on a surface representing the protein shape.

Clustering of MD trajectories

The most representative structures of aptamer/RBD complexes were obtained by the clustering analysis of MD simulations. Before clustering, the trajectories were aligned by the positions of the Cα atoms of the RBD.

Clusters of structures and their centers were computed using the quality threshold algorithm[63] implemented in VMD. The RMSD of the phosphorus atoms was used as a metric function; the cutoff was 0.7 nm. At most, five clusters were extracted for each aptamer.

Structure preparation of aptamer-RBD complexes

The RBD model was taken from the MD simulations described above. The three- dimensional models of aptamers were built using the Avogadro[64] program. Aptamer-RBD complexes were obtained by docking using HDOCK, a web server for protein-DNA/RNA docking.[35,65] For each complex, the top 10 models according to the HDOCK score were collected for further analysis.

Quantum-mechanical calculations

For accelerating QM calculations and obtaining pairwise interactions, the fragment molecular orbital (FMO) method[66] was used. In the FMO calculations, each residue within the RBD protein was represented by a fragment (189 fragments total); the fragmentation was performed using the Facio[67] program. Apt16 was divided into 16 nucleotide fragments, and other aptamers were likewise fragmented into nucleotides. Fragment calculations were performed in the fully polarizable embedding, and by calculating pairs of fragments, inter fragment interactions were incorporated, such as charge transfer, exchange-repulsion, electrostatics, dispersion, and solvent screening.[68] FMO was validated in various protein-ligand binding studies, including complexes of SARS-CoV-2.[69–72]

The geometry optimization of both aptamers and RBD/aptamer complexes was carried out using third-order density-functional tight-binding (DFTB3) combined with the two-body FMO2[73] using 3ob parameters,[43] D3(BJ) empirical dispersion,[74] and the conductor-like polarizable continuum model of solvation (C-

PCM).[73] All calculations were performed with GAMESS(US) program.[75]

Individual pair interaction energies (PIEs) can also be computed by using FMO. In the RBD-aptamer systems, they represent residue-nucleotide interaction energies. By summing the values of all pairs, the total value (E_{total}) was obtained. The PIE and total interactions can be computed at a higher level of theory than the level used for geometry optimizations. In this work, we used two methods for PIE energy calculation. First one was the DFTB3 method used for structure calculation. The second one is RI-MP2 method with the 6-31G(d,p) basis set together with the auxiliary rimp2-cc-pVDZ basis set[76] for the resolution of identity. Both methods were combined with the C-PCM[73] model for solvent. The choice of basis sets is a compromise between speed of calculation on the one hand, and accuracy of interaction energies on the other. We note that for the purpose of informing aptamer design, these interaction energies need to be qualitatively correct (i.e. trends), rather than quantitatively and small errors are tolerable. See Supporting Information Tables S2 and S3 for the estimates of BSSE errors associated with the chosen method.

Experimental details

Oligonucleotide synthesis

Oligonucleotide synthesis for fluorescence polarization assay was performed on an ASM- 800 DNA/RNA synthesizer (Biosset, Russia) on a 0.1 μM scale. Oligonucleotides were synthesized using a standard phosphoramidite 0.05 M solution in CH_3CN (ChemGenes Inc., USA).

The fluorescent dye 5(6) carboxyfluorescein was attached at the 5'-end of the oligonucleotide, which was deprotected and removed from the solid support using concentrated NH_3(aq). According to standard manufacturer protocols, purification was accomplished by using a reverse-phase cartridge (ChemGenes Inc., USA).

Synthesis and purification of recombinant RBD

Plasmid construction was performed based on the gene sequence encoding the SARS-CoV-2 Spike protein (GenBank MN908947). The codon composition of the sequence was optimized by using GeneOptimizer.[77] The resulting nucleotide was synthesized as part of the pGH vector (ltd. DNA-Synthesis, Russia).

The sequence corresponding to the region 308V-542N was used to obtain RBD. The signal sequence of tissue plasminogen activator was inserted by PCR using the following primers: TPa-F5'-GACCGCCATGTTGGCATTG-3' and Tpa-R5'-CAGCAGCACA CAGCAGAGCC CTCTCTTCAT TGCATCCATG GTGGCCCCGG GGCTAGCCTA TAGTGAG-3' on the pIPP vector matrix with the use of primers TpaRBD- F5'-TGCTGTGTGC TGCTGCTGTG TGGAGCAGTC TTCGTTTCGG CCGTGGAAAA GGGCATCTAC CAGAC-3' and RBD4-R5'-AAAAAAGTCG ACGAGGCTGA TCAGCGGTTT AAAC-3' on the pGH-S matrix. PCR products were annealed and inserted into the vector pVEAL2 AhlI and SalI to allow 6×His to be included in the protein. The CHO cell expression system was used to develop the SARS-CoV-2 Spike protein trimer and RBD. Using Lipofectamine 3000 (ThermoFisher, USA) per the manufacturer's protocol, the CHO-K1 cell line was transfected with pVEAL2-RBD or pVEAL2-TRIS vectors and plasmid pCMV (CAT) T7-SB100. The latter encodes transposase SB100, providing integration of expression cassettes into the mammalian host genome.

After three days, the antibiotic puromycin (InvivoGen, USA) was added to the culture medium at a final concentration of 10 μgmL-1 to select resistant cell clones that had received the resistance gene to this antibiotic from the pVEAL2 vector. The selection was carried out for three days; then, the polyclonal cell culture was

plated into a 96-well plate at the final concentration of one cell per well and incubated for two weeks under selective pressure. The wells were examined for the presence of colonies, and cultural fluids were screened for productivity. Clones with the highest productivity were used for protein isolation. Protein purity was determined by SDS-PAGE and visualized by Coomassie blue staining and Western blotting using rabbit anti-RBD antibodies (Sino Biological, China).

According to the manufacturer's protocol, the recombinant RBD was purified by metal chelate chromatography on a Ni NTA column (Qiagen, Germany). Additional purification was carried out by chromatography on columns connected in series with cation exchange (SP-Sepharose) and anion exchange (Q-Sepharose) sorbents, equilibrated with 20 mM Tris-HCl, pH 8.2. After the protein was applied, the columns were washed with 20 mM Tris-HCl, pH 8.2. Then, the column with the SP-sepharose sorbent, on which RBD is not adsorbed under these conditions (pH 8.2), was disconnected; and the proteins bound to the Q-sepharose sorbent were eluted in a linear NaCl concentration gradient from 0 to 1 M in 20 mM Tris-HCl, pH 8.2.

Fractions of the target protein were analyzed by denaturing electrophoresis in 15 % PAGE. The protein solution was dialyzed against PBS and subjected to sterilizing filtration through filters with a pore size of 0.22 μM. Fractions of the target protein were analyzed by denaturing electrophoresis in 15 % PAGE. The quantitative analysis of the protein content was performed by the Lowry method.

Flow cytometry binding analyses

To determine the binding performance of candidate sequences, positive RBD beads or S-protein beads were incubated with 200 nM FAM-labeled candidate sequences in 200 μL binding buffer (PBS, pH 7.4, including 136.8 mM NaCl, 10.1 mM Na_2HPO_4, 2.7 mM KCl, 1.8 mM K_2HPO_4, 0.55 mM $MgCl_2$) at 25 °C for 30 min. The beads were washed twice using the binding buffer and suspended in a 200 μL binding buffer. The fluorescence intensity of beads was measured by flow cytometry (FACSVerse, BD) by counting about 1000–3000 events.

Ni beads for His-tagged RBD or S-Protein conjugation were from GE Healthcare (USA). His-tagged-RBD of SARS- CoV-2 Spike Protein (40592-V08B) and SARS-CoV-2 Spike S1+S2 ECD-His recombinant protein (40589-V08B1) were purchased from Sino Biological Inc. (China).[78] All DNA sequences for flow cytometry experiments were synthesized by Sangon Biotech with HPLC purification (Shanghai, China).

Fluorescence polarization analyses of aptamer binding to RBD

The Clariostar Plus microplate reader (BMG LABTECH, Germany) was used for fluorescence polarization measurements. The optical system had the following settings: excitation (482±16) nm, dichroic filter LP 504 nm, emission (530±40) nm, and target polarization (P) value 200 mP (mP = P/1000). Signal time integration was 3s for each measurement. The measurements were performed at 37 °C.

Aptamers were titrated with the protein as follows: 150 μL of aptamers Apt27 and Apt31 or control random sequences R20 (20 nucleotides long) and R40 10 nM (40 nucleotides long) at a concentration of 10 nM in phosphate-buffered saline were placed into wells of flat-bottomed black 96 well plates (Grenier, Germany) in 14 replicates.

The RBD protein was injected into the wells and vortexed automatically by the instrument Clariostar Plus microplate reader. The final concentration of the protein after each ejection was 7 pM, 14 pM, 17 pM, 20 pM, and 24 pM. The kinetics of the fluorescence polarization changes was measured. Aptamer and RBD synthesis were described above.

RBD/Apt31 complex preparation for SAXS

First, 117 μL of binding buffer (PBS, pH 7.4, including 136.8 mM NaCl, 10.1 mM Na_2HPO_4, 2.7 mM KCl, 1.8 mM K_2HPO_4, 0.55 mM MgCl2) was added to 58 μL of RBD (15 mg mL^{-1}) stock solution

to prepare a 5 mg mL^{-1} protein concentration. Then 1 µL of aptamer Apt31 solution (745 nM) was mixed with 48 µL binding buffer, pH 7.4 (stock solution of Apt31). 0.7 µL of Apt31 stock solution was added to 70 µL RBD (5 mg mL^{-1}) to obtain the first complex. 1 µL of Apt31 stock solution was mixed with 105 µL RBD (5 mg mL^{-1}) to obtain the second complex. Both complexes were incubated for 15 min at ambient conditions and then stored at 4 °C before analysis.

SAXS measurements and data processing

SAXS data from RBD, Apt31, and RBD/Apt31 complex samples were obtained at the BioMUR beamline at the NRC "Kurchatov Institute".[79] The wavelength on the beamline was 0.14 nm, which corresponds to energy 8.58 keV. The sample to detector distance was 700 mm. The SAXS data were recorded by a Dectris Pilatus 1M detector during the X-ray 300s expositions of each image. The temperature of each sample during the exposition was set to 20 °C. These images were radially averaged relatively to the center of the beam. The aptamer and the protein were measured at the concentrations of 8, 4 and 2 mg/ml. For the RBD/Apt31 complex the concentrations were 5, 2.5 and 1.25 mg/ml. Dilution of the samples allowed us to define the most appropriate concentration which gives sufficient SAXS intensity with a good signal-to-noise ratio and does not already show multiple scattering effects.

The SAXS patterns were obtained in the range of scattering vectors from 0.1 to 4.5 nm^{-1} ($s = 4\pi\sin(\theta)/\lambda$, where 2θ is the scattering angle). The SAXS pattern from the buffer was also recorded and subtracted from the solution signal to get the SAXS data from the sample molecules only.

Since the signal-to-noise ratio of the recorded SAXS curves on wide angles and the set of concentrations used for the experiment were not optimal, only the scattering vector range $0.2 < s < 1.4$ nm^{-1} was used. Data obtained from SAXS measurements at the concentrations 4 mg/ml for the aptamer and protein separately and 2.5 mg/ml for the complex yielded the monodispersity of the solution for each sample, which allows the calculation of the structure parameters for the aptamer, protein, and their complex. To reconstruct the spatial structure of the Apt31 aptamer, additional size-exclusion-chromatography (SEC) SAXS measurements were performed at the TPS 13A BioSAXS beamline at NSRRC, Taiwan.[80] The sample was prepared in Tris buffer solution and exposed by X-rays of 15 keV at 22°C with an online HPLC system. The initial concentration was 13.4 mg/ml; HPLC column provided a high dilution of the sample during the measurements. The X-Ray beam size at the sample position was 320x260 µm². Sample to detector distance was 1210.06 mm. Total 128 data frames were recorded by the detector Eiger X 9M, each 2s exposure time per frame. All of the frames recorded with a constant Rg value over the HPLC (high-performance liquid chromatography) sample flow were radially averaged into the one-dimensional scattering intensities, two of them were merged into resulting SAXS curve. SAXS measurements, data acquisition, and processing were performed according to standard pipelines.[81–83] The reduction and analysis of experimental SAXS data (buffer subtraction, averaging, merging, gyration radius calculation, Guinier and Kratky analyses, pair-distance distribution function *p(r)* construction, bead modeling, molecular model validation, MD and bead model alignment) were made using the program suite ATSAS.[84] The bead model was constructed using the program DAMMIN[85] according to the calculated *p(r)* function. Evaluation of the molecular model was performed by a simulated annealing process using the program CRYSOL.[86]

Acknowledgements

The authors are grateful to JCSS Joint Super Computer Center of the Russian Academy of Sciences – Branch of Federal State Institution "Scientific Research Institute for System Analysis of the Russian Academy of Sciences" for providing supercomputers for computer simulations. The authors thank the RSC Group (www.rscgroup.ru) and personally Mr. Oleg Gorbachev for the constant support and establishment of "The Good Hope Net Project" (www.thegoodhope.net) multifunctional non-profit anti-CoVID research project. The authors also thank the Helicon

Company (www.helicon.ru) and personally Olesya Kucenko, Alexander Kolobov, Leonid Klimov for instrumental support and help with conducting fluorescence polarization assays, which were performed on a demo instrument Clariostar Plus microplate reader (BMG LABTECH, Germany). We thank Dr. Yong- Zhen Zhang for providing the genome sequence of 2019-nCoV and Dr. Xinquan Wang for providing the crystal structure of the binding domain of the SARS-2 Spike protein. The authors are grateful to Aptamerlab LCC financial support (www.aptamerlab.com). Y.A.'s work at Argonne National Laboratory was supported by the U.S. Department of Energy, Office of Science, under contract DE-AC02-06CH11357. The work of D.M. and G.G. has been done as part of the BioExcel CoE (www.bioexcel.eu), a project funded by the European Union contracts H2020-INFRAEDI- 02-2018-823830 and H2020-EINFRA-2015-1-675728. D.M. and G.G. also thank the CSC-IT center in Espoo, Finland, as well as PRACE for awarding access to resource Curie-Rome based in France at GENCI. V.M. thanks Russian Foundation for Basic Research (project number 19-03-00043). A.B.'s and N.K.'s work was supported by the Ministry of Science and Higher Education of Russian Federation (state assignment of the Research Center of Biotechnology RAS). V. deF. G.C., N.B and G.O. are grateful to FISR2020 _00177 Shield, Italian Ministry of Education and Research, for funding. GC is grateful to the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement: cONCReTE 872391; PRISAR2 872860. Use of the 13A BioSAXS beamtime at the Taiwan Photon Source is acknowledged. The work of M.V.B was funded by the Canadian Institutes of Health Research grant OV1-170353. SAXS measurements and PIEDA analyses were funded by the Russian Science Foundation (project № 21-73-20240 for A.S.K.).

Author Contributions: V.M., I.A.Shch., P.V.A., D.M., D.G.F., Y.A., F.N.T., A.A.M., G.G., N.B., G.O., G.C., V.deF., T.N.Z., and A.S.K. performed the computational design of the aptamers. L.S.K. and V.N.S. synthesized oligonucleotides. I.A.M., D.N.Sh., E.A.V., S.V.B., T.I.E., A.A.I., V.S.N., and D.V.Sh. synthesized and purified proteins. Y.S., C.Y., and W.T. performed flow cytometry analyses. O.S.K., K.A.L., Y.E.G., V.A.S., and A.S.K. performed fluorescence polarization analyses. R.V.M., V.N.Z., G.P., A.N.B., N.S.K., K.H.C.C, and U.J. performed small-angle X- ray measurements. D.V.V. performed mathematical analyses. E.P.T. provided consultations on virology. D.G.F., G.S.Z., A.S.K., F.N.T, T.N.Z., S.S.Z., and D.V.V. contributed to the overall design of the project. A.S.K., W.T., F.N.T., Y.A., D.G.F., A.A.M, E.P.T., M.V.B., V.deF., G.C., D.N.Sh, and V.A.S. provided supervising in different fields. V.M., I.A.S., P.V.A., D.M., G.G., N.B., G.O., R.V.M., D.V.V., G.C., V.deF., T.N.Z., A.S.K., V.N.S., I.A.M., D.N.Sh., Ch.Y., and K.A.L. wrote the manuscript.

Keywords: *aptamers, fragment molecular orbitals method, molecular dynamics, SARS-CoV-2, SAXS*

References

[1] S. F. Ahmed, A. A. Quadeer, M. R. McKay, *Viruses* **2020**, *12*, 254.

[2] A. R. Fehr, S. Perlman, in *Coronaviruses Methods Protoc.*, Springer New York, **2015**, pp. 1–23.

[3] W. Tai, L. He, X. Zhang, J. Pu, D. Voronin, S. Jiang, Y. Zhou, L. Du, *Cell. Mol. Immunol.* **2020**, *17*, 613–620.

[4] X. Ou, Y. Liu, X. Lei, P. Li, D. Mi, L. Ren, L. Guo, R. Guo, T. Chen, J. Hu, Z. Xiang, Z. Mu, X. Chen, J. Chen, K. Hu, Q. Jin, J. Wang, Z. Qian, *Nat. Commun.* **2020**, *11*, 1620.

[5] S. Wang, Z. Qiu, Y. Hou, X. Deng, W. Xu, T. Zheng, P. Wu, S. Xie, W. Bian, C. Zhang, Z. Sun, K. Liu, C. Shan, A. Lin, S. Jiang, Y. Xie, Q. Zhou, L. Lu, J. Huang, X. Li, *Cell Res.* **2021**, *31*, 126–140.

[6] D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham, J. S. McLellan, *bioRxiv Prepr. Serv. Biol.* **2020**, *367*, 1260–1263.

[7] A. D. Ellington, J. W. Szostak, *Nature* **1990**, *346*, 818–822.

[8] C. Tuerk, L. Gold, *Science* **1990**, *249*, 505–510.

[9] S. M. Nimjee, C. P. Rusconi, B. A. Sullenger, *Annu. Rev. Med.* **2005**, *56*, 555–583.

[10] S. Catuogno, C. L. Esposito, V. de Franciscis, *Pharmaceuticals* **2016**, *9*, 69.

[11] G. Sharma, S. Sharma, P. Sharma, D. Chandola, S. Dang, S. Gupta, R. Gabrani, *J. Appl. Microbiol.* **2016**, *121*, 309–319.

[12] A. Dhiman, P. Kalra, V. Bansal, J. G. Bruno, T. K. Sharma, *Sensors Actuators, B Chem.* **2017**, *246*, 535–553.

[13] J. Zhou, J. Rossi, *Nat. Rev. Drug Discov.* **2017**, *16*, 181–202.

[14] F. Nici, G. Oliviero, A. P. Falanga, S. D'Errico, M. Marzano, D. Musumeci, D. Montesarchio, S. Noppen, C. Pannecouque, G. Piccialli, N. Borbone, *Org. Biomol. Chem.* **2018**, *16*, 2349–2355.

[15] M. Scuotto, E. Rivieccio, A. Varone, D. Corda, M. Bucci, V. Vellecco, G. Cirino, A. Virgilio, V. Esposito, A. Galeone, N. Borbone, M. Varra, L. Mayol, *Nucleic Acids Res.* **2015**, *43*, 7702–7716.

[16] M. Terracciano, L. De Stefano, N. Borbone, J. Politi, G. Oliviero, F. Nici, M. Casalino, G. Piccialli, P. Dardano, M. Varra, I. Rea, *RSC Adv.* **2016**, *6*, 86762–86769.

[17] A. B. Kinghorn, L. A. Fraser, S. Lang, S. C. C. Shiu, J. A. Tanner, *Int. J. Mol. Sci.* **2017**, *18*, 2516.

[18] V.-T. Nguyen, B. H. Lee, S. H. Kim, M. B. Gu, *Biotechnol. J.* **2016**, *11*, 843–849.

[19] M. Blind, M. Blank, *Mol. Ther. - Nucleic Acids* **2015**, *4*, e223.

[20] K. Sefah, D. Shangguan, X. Xiong, M. B. O, W. Tan, *Nature Protocols* **2010**, *5*, 1169–1185.

[21] S. Gao, X. Zheng, B. Jiao, L. Wang, *Anal. Bioanal. Chem.* **2016**, *408*, 4567–4573.

[22] M. Yüce, N. Ullah, H. Budak, *Analyst* **2015**, *140*, 5379–5399.

[23] P. Kalra, A. Dhiman, W. C. Cho, J. G. Bruno, T. K. Sharma, *Front. Mol. Biosci.* **2018**, *5*, 41.

[24] M. Ashrafuzzaman, C.-Y. Tseng, J. Kapty, J. R. Mercer, J. A. Tuszynski, *Nucleic Acid Ther.* **2013**, *23*, 418–426.

[25] D. S. Shcherbinin, O. V. Gnedenko, S. A. Khmeleva, S. A. Usanov, A. A. Gilep, A. V. Yantsevich, T. V. Shkel, I. V. Yushkevich, S. P. Radko, A. S. Ivanov, A. V. Veselovsky, A. I. Archakov, *J. Struct. Biol.* **2015**, *191*, 112–119.

[26] D. A. Belinskaia, P. V. Avdonin, P. P. Avdonin, R. O. Jenkins, N. V. Goncharov, *Comput. Biol. Chem.* **2019**, *80*, 452–462.

[27] D. R. Bell, J. K. Weber, W. Yin, T. Huynh, W. Duan, R. Zhou, *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 8486–8493.

[28] S. Soon, N. A. Nordin, *Mater. Today Proc.* **2019**, *16*, 2096–2100.

[29] R. Ahirwar, S. Nahar, S. Aggarwal, S. Ramachandran, S. Maiti, P. Nahar, *Sci. Rep.* **2016**, *6*, 1–11.

[30] Z. Gong, Y. Zhao, C. Chen, Y. Xiao, *J. Biomol. Struct. Dyn.* **2011**, *29*, 403–416.

[31] K. L. Rhinehardt, G. Srinivas, R. V. Mohan, *J. Phys. Chem. B* **2015**, *119*, 6571–6583.

[32] Y. Biniuri, B. Albada, I. Willner, *J. Phys. Chem. B* **2018**, *122*, 9102–9109.

[33] J. Jeevanandam, K. X. Tan, M. K. Danquah, H. Guo, A. Turgeson, *Biotechnol. J.* **2020**, *15*, 1900368.

[34] Y. Alexeev, M. P. Mazanetz, O. Ichihara, D. G. Fedorov, *Curr. Top. Med. Chem.* **2012**, *12*, 2013–2033.

[35] Y. Yan, H. Tao, J. He, S. Y. Huang, *Nat. Protoc.* **2020**, *15*, 1829–1852.

[36] V. Gapsys, B. L. De Groot, *J. Chem. Theory Comput.* **2017**, *13*, 6275–6289.

[37] P. Śledź, A. Caflisch, *Curr. Opin. Struct. Biol.* **2018**, *48*, 93–102.

[38] Y. Chen, Y. Guo, Y. Pan, Z. J. Zhao, *Biochem. Biophys. Res. Commun.* **2020**, *525*, 135–140.

[39] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, X. Wang, *Nature* **2020**, *581*, 215–220.

[40] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Y. Bi, X. Ma, F. Zhan, L. Wang, T. Hu, H. Zhou, Z. Hu, W. Zhou, L. Zhao, J. Chen, Y. Meng, J. Wang, Y. Lin, J. Yuan, Z. Xie, J. Ma, W. J. Liu, D. Wang, W. Xu, E. C. Holmes, G. F. Gao, G. Wu, W. Chen, W. Shi, W. Tan, *Lancet* **2020**, *395*, 565–574.

[41] J. Kondo, E. Westhof, *Nucleic Acids Res.* **2011**, *39*, 8628–8637.

[42] A. A. Buglak, A. V. Samokhvalov, A. V. Zherdev, B. B. Dzantiev, *Int. J. Mol. Sci.* **2020**, *21*, 8420.

[43] M. Gaus, X. Lu, M. Elstner, Q. Cui, *J. Chem. Theory Comput.* **2014**, *10*, 1518–1537.

[44] M. M. Hoffman, M. A. Khrapov, J. C. Cox, J. Yao, L. Tong, A. D. Ellington, *Nucleic Acids Res.* **2004**, *32*, D174.

[45] D. Lejeune, N. Delsaux, B. Charloteaux, A. Thomas, R. Brasseur, *Proteins Struct. Funct. Genet.* **2005**, *61*, 258–271.

[46] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737–738.

[47] A. G. Kikhney, C. R. Borges, D. S. Molodenskiy, C. M. Jeffries, D. I. Svergun, *Protein Sci.* **2020**, *29*, 66–75.

[48] G. Yang, Z. Li, I. Mohammed, L. Zhao, W. Wei, H. Xiao, W. Guo, Y. Zhao, F. Qu, Y. Huang, *Signal Transduct. Target. Ther. 2021 61* **2021**, *6*, 1–4.

[49] A. Idili, C. Parolo, R. Alvarez-Diduk, A. Merkoçi, *ACS Sensors* **2021**, *6*, 3093–3101.

[50] M. Sun, D. S. Liu, X. Wei, S. Wan, M. Huang, T. Song, Y. Lu, X. Weng, Z. Lin, P. H. Chen, P. Y. Song, P. C. Yang, *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 10266–10272.

[51] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, X. Wang, *Nat.* **2020**, *581*, 215–220.

[52] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindah, *SoftwareX* **2015**, *1–2*, 19–25.

[53] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

[54] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.

[55] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 014101.

[56] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182–7190.

[57] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463–1472.

[58] S. Miyamoto, P. A. Kollman, *J. Comput. Chem.* **1992**, *13*, 952–962.

[59] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577–8593.

[60] T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker, *Nucleic Acids Res.* **2004**, *32*, W665.

[61] D. Sitkoff, K. A. Sharp, B. Honig, *J. Phys. Chem.* **1994**, *98*, 1978–1988.

[62] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D. W. Gohara, T. Dolinsky, R. Konecny, D. R. Koes, J. E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G. W. Wei, M. J. Holst, J. A. McCammon, N. A. Baker, *Protein Sci.* **2018**, *27*, 112–128.

[63] L. J. Heyer, S. Kruglyak, S. Yooseph, *Genome Res.* **1999**, *9*, 1106–1115.

[64] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeerschd, E. Zurek, G. R. Hutchison, *J. Cheminform.* **2012**, *4*, 17.

[65] Y. Yan, D. Zhang, P. Zhou, B. Li, S.-Y. Huang, *Nucleic Acids Res.* **2017**, *45*, 365–373.

[66] D. G. Fedorov, *WIREs Comput. Mol. Sci.* **2017**, *7*, e1322.

[67] M. Suenaga, *J. Comput. Chem. Japan* **2005**, *4*, 25–32.

[68] D. G. Fedorov, *J. Chem. Theory Comput.* **2019**, *15*, 5404–5416.

[69] R. Hatada, K. Okuwaki, Y. Mochizuki, Y. Handa, K. Fukuzawa, Y. Komeiji, Y. Okiyama, S. Tanaka, *J. Chem. Inf. Model.* **2020**, *60*, 3593–3602.

[70] H. Lim, A. Baek, J. Kim, M. S. Kim, J. Liu, K. Y. Nam, J. H. Yoon, K. T. No, *Sci. Rep.* **2020**, *10*, 16862.

[71] A. Acharya, R. Agarwal, M. B. Baker, J. Baudry, D. Bhowmik, S. Boehm, K. G. Byler, S. Y. Chen, L. Coates, C. J. Cooper, O. Demerdash, I. Daidone, J. D. Eblen, S. Ellingson, S. Forli, J. Glaser, J. C. Gumbart, J. Gunnels, O. Hernandez, S. Irle, D. W. Kneller, A. Kovalevsky, J. Larkin, T. J. Lawrence, S. Legrand, S. H. Liu, J. C. Mitchell, G. Park, J. M. Parks, A. Pavlova, L. Petridis, D. Poole, L. Pouchard, A. Ramanathan, D. M. Rogers, D. Santos-Martins, A. Scheinberg, A. Sedova, Y. Shen, J. C. Smith, M. D. Smith, C. Soto, A. Tsaris, M. Thavappiragasam, A. F. Tillack, J. V. Vermaas, V. Q. Vuong, J. Yin, S. Yoo, M. Zahran, L. Zanetti-Polzi, *J. Chem. Inf. Model.* **2020**, *60*, 5832–5852.

[72] K. Akisawa, R. Hatada, K. Okuwaki, Y. Mochizuki, K. Fukuzawa, Y. Komeiji, S. Tanaka, *RSC Adv.* **2021**, *11*, 3272–3279.

[73] Y. Nishimoto, D. G. Fedorov, *Phys. Chem. Chem. Phys.* **2016**, *18*, 22047–22061.

[74] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.*, **2011,** 32, 1456-1465

[75] G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez, T. Harville, S. Irle, J. Ivanic, K. Kowalski, S. S. Leang, H. Li, W. Li, J. J. Lutz, I. Magoulas, J. Mato, V. Mironov, H. Nakata, B. Q. Pham, P. Piecuch, D. Poole, S. R. Pruitt, A. P. Rendell, L. B. Roskop, K. Ruedenberg, T. Sattasathuchana, M. W. Schmidt, J. Shen, L. Slipchenko, M. Sosonkina, V. Sundriyal, A. Tiwari, J. L. Galvez Vallejo, B. Westheimer, M. Włoch, P. Xu, F. Zahariev, M. S. Gordon, *J. Chem. Phys.* **2020**, *152*, 154102.

[76] F. Weigend, A. Köhn, C. Hättig, *J. Chem. Phys.* **2002**, *116*, 3175-3183.

[77] D. Raab, M. Graf, F. Notka, T. Schödl, R. Wagner, *Syst. Synth. Biol.* **2010**, *4*, 215–225.

[78] Y. Song, J. Song, X. Wei, M. Huang, M. Sun, L. Zhu, B. Lin, H. Shen, Z. Zhu, C. Yang, *Anal. Chem.* **2020**, *92*, 9895–9900.

[79] G. S. Peters, O. A. Zakharchenko, P. V. Konarev, Y. V. Karmazikov, M. A. Smirnov, A. V. Zabelin, E. H. Mukhamedzhanov, A. A. Veligzhanin, A. E. Blagov, M. V. Kovalchuk, *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **2019**, *945*, 162616.

[80] Din GoaLiu, C. H. Chang, M. H. Lee, C. Y. Liu, C. F. Chang, L. C. Chiang, C. S. Hwang, J. C. Huang, A. Sheng, C. K. Kuan, Y. Q. Yeh, C. J. Su, K. F. Liao, W. R. Wu, O. Shih, U. S. Jeng, *AIP Conf. Proc.* **2019**, *2054*, 060021.

[81] J. K. Raynes, L. Day, M. A. Augustin, J. A. Carver, *J. Dairy Sci.* **2015**, *98*, 2172–2182.

[82] C. E. Blanchet, D. I. Svergun, *Annu. Rev. Phys. Chem.* **2013**, *64*, 37–54.

[83] P. C. Chen, P. Masiewicz, V. Rybin, D. Svergun, J. Hennig, *ACS Comb. Sci.* **2018**, *20*, 197–202.

[84] K. Manalastas-Cantos, P. V. Konarev, N. R. Hajizadeh, A. G. Kikhney, M. V. Petoukhov, D. S. Molodenskiy, A. Panjkovich, H. D. T. Mertens, A. Gruzinov, C. Borges, C. M. Jeffries, D. I. Svergun, D. Franke, *J. Appl. Crystallogr.* **2021**, *54*, 343-355.

[85] D. I. Svergun, *Biophys. J.* **1999**, *76*, 2879–2886.

[86] D. Svergun, C. Barberato, M. H. Koch, *J. Appl. Crystallogr.* **1995**, *28*, 768–773.

10.1002/chem.202104481

WILEY-VCH

RESEARCH ARTICLE

Table of Contents graphical abstract

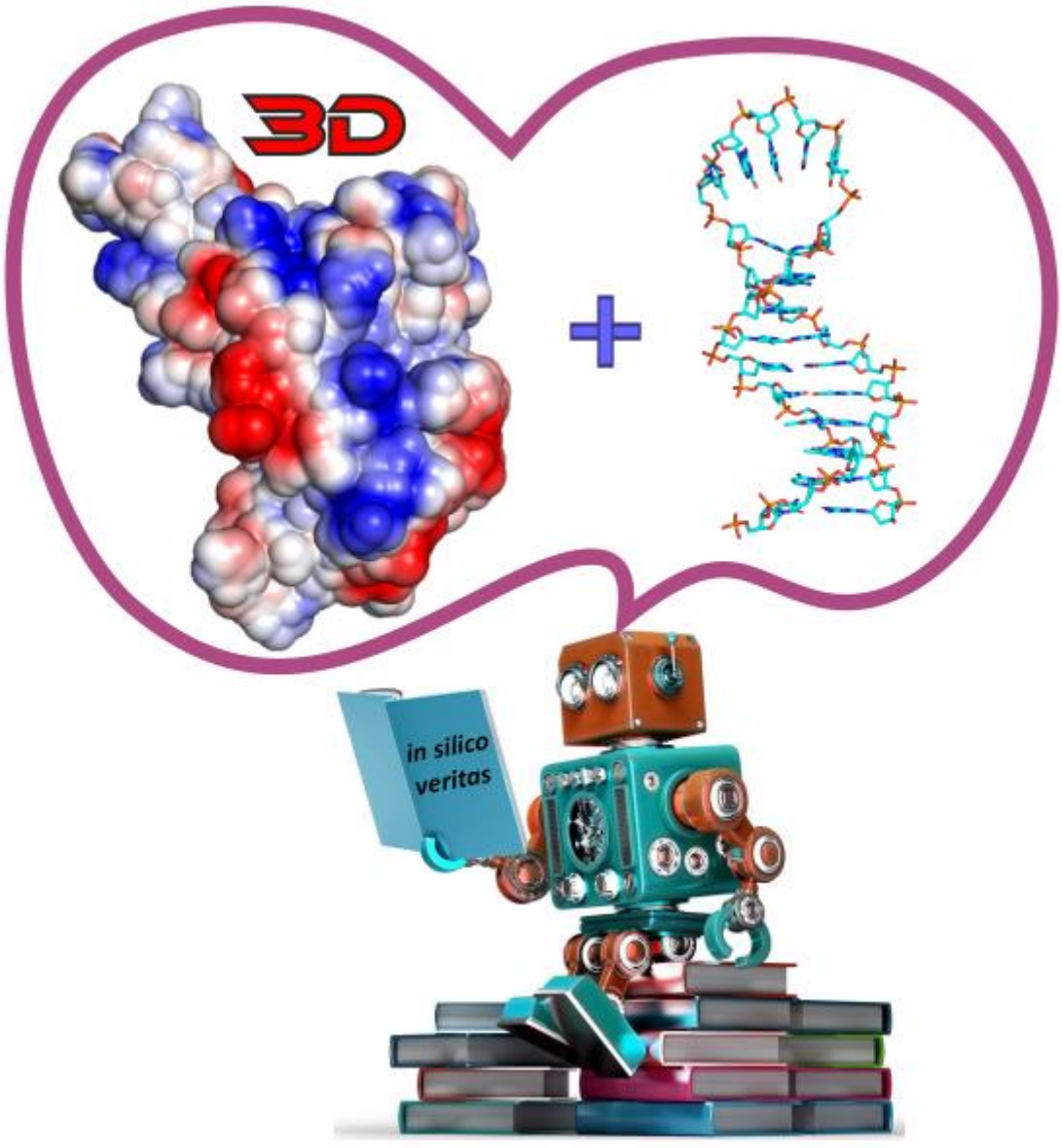


Table of Contents text

Applying computational methods together with experimental procedures can speed up an aptamer design. Using a new iterative design procedure, Structure and Interaction Based Drug Design (SIBDD), a highly specific aptamer to the receptor-binding domain of the SARS-CoV-2 spike protein, was designed and validated. The approach combines molecular dynamics with quantum chemistry to estimate binding affinity *in silico* with *in vitro* verification.