biomedicines

MDPI

Article

Phage Display-Derived Compounds Displace hACE2 from Its Complex with SARS-CoV-2 Spike Protein

Marc Sevenich [1,2], Elena Thul [3], Nils-Alexander Lakomek [1,4], Thomas Klünemann [5], Maren Schubert [6], Federico Bertoglio [6], Joop van den Heuvel [5], Patrick Petzsch [7], Jeannine Mohrlüder [1,*] and Dieter Willbold [1,4,8,*]

1 Institute of Biological Information Processing (IBI-7), Forschungszentrum Jülich, 52425 Jülich, Germany; m.sevenich@fz-juelich.de (M.S.); n.lakomek@fz-juelich.de (N.-A.L.)
2 Priavoid GmbH, 40225 Düsseldorf, Germany
3 Medizinische Fakultät, Westfälische Wilhelms Universität Münster, 48149 Münster, Germany; e_thul01@uni-muenster.de
4 Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany
5 Helmholtz Centre for Infection Research, 38124 Braunschweig, Germany; thomas.kluenemann@helmholtz-hzi.de (T.K.); joop.vandenheuvel@helmholtz-hzi.de (J.v.d.H.)
6 Institut für Biochemie, Biotechnologie und Bioinformatik—Abteilung Biotechnologie, Technische Universität Braunschweig, 38106 Braunschweig, Germany; m.bleckmann@tu-bs.de (M.S.); f.bertoglio@tu-bs.de (F.B.)
7 Biologisch-Medizinisches Forschungszentrum (BMFZ), Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany; patrick.petzsch@uni-duesseldorf.de
8 JuStruct, Forschungszentrum Jülich, 52425 Jülich, Germany
* Correspondence: j.mohrlueder@fz-juelich.de (J.M.); d.willbold@fz-juelich.de (D.W.); Tel.: +49-2461-612100 (D.W.)

check for updates

Citation: Sevenich, M.; Thul, E.; Lakomek, N.-A.; Klünemann, T.; Schubert, M.; Bertoglio, F.; van den Heuvel, J.; Petzsch, P.; Mohrlüder, J.; Willbold, D. Phage Display-Derived Compounds Displace hACE2 from Its Complex with SARS-CoV-2 Spike Protein. *Biomedicines* **2022**, *10*, 441. https://doi.org/10.3390/biomedicines10020441

Academic Editor: Dong-Yan Jin

Received: 19 January 2022
Accepted: 11 February 2022
Published: 14 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Abstract: Severe respiratory syndrome coronavirus-2 (SARS-CoV-2) is a highly contagious beta-class coronavirus. Although vaccinations have shown high efficacy, the emergence of novel variants of concern (VOCs) has already exhibited traits of immune evasion. Thus, the development of tailored antiviral medications for patients with incomplete, inefficient, or non-existent immunization, is essential. The attachment of viral surface proteins to the cell surface is the first crucial step in the viral replication cycle, which for SARS-CoV-2 is mediated by the high affinity interaction of the viral trimeric spike with the host cell surface-located human angiotensin converting enzyme-2 (hACE2). Here, we used a novel and efficient next generation sequencing (NGS) supported phage display strategy for the selection of a set of SARS-CoV-2 receptor binding domain (RBD)-targeting peptide ligands that bind to the target protein with low µM to nM dissociation constants. Compound CVRBDL-3 inhibits the SARS-CoV-2 spike protein association to hACE2 in a concentration-dependent manner for pre- as well as post-complex formation conditions. Further rational optimization yielded a CVRBDL-3 based divalent compound, which demonstrated inhibitory efficacy with an IC_{50} value of 47 nM. The obtained compounds were not only efficient for the different spike constructs from the originally isolated "wt" SARS-CoV-2, but also for B.1.1.7 mutant trimeric spike protein. Our work demonstrates that phage display-derived peptide ligands are potential fusion inhibitors of viral cell entry. Moreover, we show that rational optimization of a combination of peptide sequences is a potential strategy in the further development of therapeutics for the treatment of acute COVID-19.

Keywords: SARS-CoV-2; coronavirus; spike protein; phage display; acute COVID-19 treatment; spike-hACE2 inhibitor

1. Introduction

SARS-CoV-2 is a beta class coronavirus with the ability for human-to-human transmission that was initially discovered and characterized in Wuhan, China, at the end of 2019 [1,2]. Since then, it has challenged health care systems due to its rapid spread and

the transmission of COVID-19 throughout the world. Far more than the past coronavirus pandemics, caused by SARS-CoV in 2002 and MERS-CoV in 2012, the ongoing pandemic has claimed over 5.7 million lives with over 386 million total cases across the world [3–6].

The coronavirus replication depends on a multi-step process starting with the interaction of the viral trimeric spike protein (S) and the human angiotensin converting enzyme (hACE2) receptor that mediates uptake of the viral RNA into the host-cell cytoplasm. Each monomer of the spike protein consists of the two functional subunits S1 and S2. The S1 subunit comprises the receptor binding domain (RBD) that interacts via a defined motif sequence (RBM) with an N-terminally located helical structure of the hACE2. In accordance with previous observations from SARS-CoV, the SARS-CoV-2 RBD has been found to exist in two conformational states, from which the 'up' state initiates interaction with hACE2 and the 'down' state's main function is the shielding of receptor binding relevant proportions from immune reactions [7]. The S2 subunit plays a crucial role in the membrane fusion process. While the S1-hACE2 interaction allows viral attachment to the host cell surface, the S2 site is cleaved by the human endoprotease TMPRSS2. Together with cooperative actions of other tissue specific endogenous proteases, proteolytic cleavage results in irreversible conformational changes of the S protein that eventually results in cell membrane fusion and endocytic host-cell uptake [8].

Although the viral S proteins of SARS-CoV and SARS-CoV-2 share ~73% amino acid sequence identity, SARS-CoV-2 shows significantly enhanced transmission efficiency when compared to SARS-CoV [9,10]. Structural studies suggest that alterations at the furin cleavage site, located at the boundary between S1 and S2 subunits, may result in reduced stability of the S protein, which facilitates the conformational adaptation that is required for receptor binding [11]. In addition, SARS-CoV-2 specific amino acid residues identified within the RBM were shown to enhance the binding affinity of the SARS CoV-2 spike protein and hACE2 as compared to SARS-CoV [12–14].

Since its first appearance in 2019, the virus has undergone numerous mutational events, resulting in variants of concern (VOCs) with enhanced fitness regarding their transmissibility and disease severity [15,16]. A large proportion of these mutations cluster in the spike protein, influencing its binding affinity towards hACE2 [13,17–20]. To date, the most widespread mutants such as B.1.617.2 (δ-variant), P.1 (γ-variant) and B.1.1.529 (o-variant) have widely replaced the originally identified SARS-CoV-2 virus due to their enhanced fitness [21–25]. Interestingly, most of the circulating variants that have enhanced infectivity as compared to the ancestral strain carry a D614G substitution in the S1 domain of the spike protein. Several indications suggest that the mutation may be directly linked to the enhanced transmissibility [21,24].

Although mRNA-based vaccines in particular turned out as a successful strategy for controlling the pandemic spread of SARS-CoV-2, vaccinations have already been shown to be partially less effective for certain VOCs [19,25]. The selection pressure created by different immunization strategies in combination with naturally occurring spontaneous mutations consequently leads to the enforcement of variants with increased immune evasive traits, making adaptation of vaccines as well as repetitive renewal of immunization mandatory. This situation, which resembles the status quo of the influenza virus [26], requires the development of antiviral medication, especially for cases where either no sufficient immunity is present or already existing immunity is circumvented by novel viral strains.

Hence, various research efforts have either focused on the development of novel drug candidates for a causative treatment such as Paxlovid (Pfizer), repurposing approved compounds such as Molnupiravir (Merck), or even screening components from natural sources [27–33]. However, there are no generally proven effective antiviral treatment options for SARS-CoV-2 infections. In this study, we present a novel phage display based strategy for the development of SARS-CoV-2 RBD-targeting peptide ligands. The resulting lead compounds show high binding affinity for the RBD in the nanomolar scale and binding selectivity for the SARS-CoV-2 RBD, as well as for the B.1.1.7 mutant (α-variant). Moreover,

when analyzed in binding experiments with the isolated proteins, one of the selected peptide sequences reduced the SARS-CoV-2 spike protein binding towards the cellular receptor hACE2 when added before complex formation on one side, but was also able to displace the spike protein from the already formed complex. In addition, a first optimization of the most promising lead compound resulted in an even more efficient compound.

2. Materials and Methods

2.1. Phage Display Selection

In phage display, exogenous peptides are presented on phage particles as fusion constructs with coat proteins. Consecutive rounds of biopanning and amplification increase the fraction of phages presenting strong target binders, which can be identified by sequencing of the variable portion of the genome [34].

For biopanning, the commercially available M13-bacteriophage library TriCo-16 (Creative Biolabs, Shirley, USA) was used. The library had a capacity of 2.6×10^{10} pIII fused 16-mer peptide variants. The biotinylated SARS-CoV-2 RBD target protein (positions 319 to 541) was purchased as lyophilized powder from Acrobiosystems (# QHD43416, USA) with a purity of >95%. The protein carries a C-terminal His-Tag for Ni-NTA purification followed by an Avitag sequence (SPD-C82E9, [35]) for BirA catalyzed transfer of biotin. The protein was expressed in HEK293 cells and had a Mw of 28.2 kDa with additional 5 to 8 kDa due to glycosylation. For surface immobilization, the target protein was incubated on a streptavidin-functionalized polystyrene 96-well plate (High capacity, Thermo Fisher Scientific, Waltham, MA, USA) for 1 h at RT using 70 pmol/well of target protein in 25 mM HEPES, pH 7.4, 150 mM NaCl, 100 µM EDTA (selection buffer). Non-coupled streptavidin was quenched by 10 min incubation with 100 µL of 10 µM biotin. Subsequent surface blocking was performed for 1 h by incubation with either 1% (*w*/*v*) BSA or nonfat milk powder in selection buffer. For the first selection round, 10 µL of the phage library (6×10^{11} phages/well) was diluted in 90 µL selection buffer with 0.2% (*w*/*v*) BSA or milk powder and incubated for 20 min. The subsequent washing steps with 200 µL washing buffer (0.05% (*v*/*v*) Tween-20 and either 0.2% (*w*/*v*) bovine serum albumin (BSA) or milk powder in selection buffer) were gradually increased from 5 to 15 repetitions with the number of selection rounds. Phage elution was performed using 10 min incubation with 100 µL of 0.2 M glycine-HCl (pH 2.2), followed by neutralization in 25 µL of 1 M Tris-HCl (pH 9.1). Eluted phages were amplified for 4.5 h in 20 mL of LB-cultured *E. coli* K12 ER2783 with a starting OD_{600} of 0.1. The pelleted cells were discarded after 20 min centrifugation at 5000× *g* and the phages were precipitated by addition of 7 mL 20% (*w*/*v*) PEG-8000, 2.5 M NaCl followed by overnight (ON) incubation at 4 °C. Phages were pelleted by centrifugation at 5000× *g* for 1 h at 4 °C and resuspended in 1 mL PBS (pH 7.4). After additional centrifugation at 10,000 rpm for 10 min, the supernatant was mixed with 200 µL of 20% (*w*/*v*) PEG-8000, 2.5 M NaCl and incubated for 1 h at 4 °C. The phages were pelleted in a last purification step for 1 h at 5000 rpm and the pellet was suspended in 100 µL selection buffer. Phage concentration of the purified phages was determined by serial dilution of the phage suspension and incubation with *E. coli* K12 ER2783 on IPTG/X Gal-LB titer plates ON at 37 °C. The concentration of input phages was calculated based on the serial dilution and a concentration of 6×10^{12} phage forming units ml^{-1} (PFU) was used as the input for all selection rounds and controls. The selection was performed for three consecutive rounds on the target (target selection = TS). Additionally, input phages resulting from selection rounds on the target were applied to a parallel selection (direct control = DC) on a surface without target, which was otherwise treated identically to the TS surface. As a second control, a consecutive selection was performed exclusively without target on otherwise identically treated surfaces (empty selection = ES).

2.2. Enrichment ELISA

Enrichment-ELISA is a method to identify enrichment of target binding phages during phage display selection. All steps were carried out at RT. A total of 20 pmol/well

of the biotinylated SARS-CoV-2 RBD target protein was immobilized on a streptavidin-functionalized polystyrene 96-well plate (High capacity, Thermo Fisher, USA). Both the target-immobilized and target-free surfaces were quenched with biotin as described previously. The surfaces were blocked with 1% (*w*/*v*) BSA in selection buffer for 1 h and washed with washing buffer (0.2% (*w*/*v*) BSA, 0.05% Tween-20 in selection buffer). A total of 2.5×10^{11} phages from the TS input samples or library were diluted in 100 µL washing buffer and incubated on target and control surfaces for 1 h. After washing with 200 µL washing buffer, the anti-M13 gpVIII monoclonal HRP-antibody (Sino Biologicals, Beijing, CHN) was diluted to a final concentration of 0.4 mg mL^{-1} in washing buffer and 100 µL per well was incubated for 1 h. After six-fold washing with washing buffer, 100 µL of 3,3′,5,5′-3,3,5,5-Tetramethylbenzidin (Sigma-Aldrich, St. Louis, MO, USA) substrate solution was incubated for 20 min and stopped by addition of an equal volume of 2 M H_2SO_4. The signal intensity was quantified by absorption measurement in a Fluorostar optima platereader at 450 nm (BMG Labtech, Ortenberg, GE; $n = 3$).

2.3. ssDNA Purification and Next Generation Sequencing (NGS) of Phage Samples

The ssDNA of the phage suspensions resulting from each phage display selection round was purified by phage precipitation, as described before, followed by resuspending in 200 µL of a 10:1 mixture of 3 M sodium acetate pH 5.2:1 × TE-buffer and addition of 500 µL of 99% ethanol. ssDNA was precipitated at 14,000 rpm for 10 min at 4 °C and the supernatant was discarded. After pellet purification by the addition of 250 µL of 70% EtOH and repeated centrifugation at 10,000 rpm, the pellet was dried and the DNA concentration was determined by signal quantification at A_{260}. For next generation sequencing, a polymerase chain reaction (PCR) was used to add adapter sequences to both the 3′ and 5′ end of the amplification product. Success of the amplification was determined by agarose gel separation with ethidium bromide staining. Amplicon next generation sequencing was performed with a MiSeq system (Illumina, San Diego, CA, USA).

2.4. Analysis of Next Generation Obtained Sequences

Variable DNA sequences resulting from next generation sequencing (NGS) were transcribed into peptide sequences, as described previously [36]. The transcribed sequences were filtered based on their relative frequency increase within the TS (library < TS1 < TS2 < TS3), the relative frequency of a sequence observed in one TS compared to its abundancy within the respective DC (TS2 > DC2; TS3 > DC3) and the relative frequency of one sequence observed in one TS round compared to its abundancy within the respective ES round (TS1 > ES1; TS2 > ES2; TS3 > ES3). Sequences of the last TS round that passed the filter were ranked corresponding to their enrichment from library to TS3 (TS3/library = enrichment factor) and their relative frequencies in TS3 compared to ES3 (TS3/ES3 = empty score). Filtered sequences were used as inputs for the Hammock clustering software [37]. The input FASTA-file included all filtered sequences as ranked by their empty score. The Hammock clustering software was performed in full mode with sequences ranked according to their input position (-R input). Cluster motif logos were generated from initial clusters after greedy clustering using the WebLogo 3.4 application.

2.5. CVRBDL Compounds

CVRBDL peptides were purchased with C-terminal amidation from CASLO (Copenhagen, DK) as lyophilized chloride salt powder with a purity of >95%.

2.6. SARS-CoV-2 Spike Protein and hACE2 Sample Preparation

Expression of SARS-CoV-2 spike protein constructs was either performed using High Five insect cells (SARS-CoV-2 RBD, SARS-CoV-2 S1S2 monomer and trimer, SARS-CoV-2 B.1.1.7 S1S2 trimer, SARS-CoV (2002) trimer, and SARS-CoV-2 "closed" S1S2 trimer) or HEK293-6E cells (hACE-2) via transient gene expression [38,39]. The High Five (Hi5) insect cell line (officially called BTI-Tn-5B1-4) was isolated by the Boyce Thompson Institute

for Plant Research (Ithaca, NY, USA); the cell line can be acquired from Thermo Fisher Scientific (USA). The HEK293-6E cell line was licensed from National Research Council (NRC), Biotechnological Research Institute (BRI, CAN). Recombinant protein genes for the CoV constructs (Figure A1) were synthesized by Genscript (Piscataway Township, NJ, USA) or Thermo Fisher Scientific (Waltham, MA, USA). Protein samples were purified using HisExcel columns (Cytiva, Marlborough, MA, USA) for 6xHis tagged proteins, StrepTrap HP columns (Cytiva, MA, USA) for TwinStrep tagged proteins (all trimeric proteins with fold-on sequences were Strep and His-tagged and purified via StrepTrap HP columns) or protein A columns (Thermo Fisher Scientific, USA) for Fc-tagged proteins (see Table A1 for complete list of all constructs and modifications). A C-terminal introduction of a T4 bacteriophage fold-on sequence (with exception of the SARS-CoV construct) was used for the induction of S1S2 stable trimers [40]. All purifications were performed on Äkta Start or Äkta Pure systems (Cytiva, USA) according to the manufacturer's protocol. Depending on the protein size, size exclusion chromatography (SEC) was performed as a final polishing step using either a Superdex 200 or Superose 6 column (Cytiva, USA) in 20 mM Tris (pH 8.0), 150 mM NaCl as running buffer. The SARS-CoV-2 nuclear magnetic resonance (NMR) sample was additionally purified via 10/300 Superdex 200 Increase SEC (Cytiva, USA) in 20 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) (pH 6.5), 100 mM NaCl. The oligomerization state and purity of each construct was verified by SEC—multiangle light scattering (SEC-MALS).

2.7. NMR Sample Preparation

SARS-CoV-2 RBD protein was expressed in HighFive insect cells as described previously. The NMR sample was prepared using 0.85 mM (25.7 mg/mL) SARS-CoV-2 RBD, dissolved in 20 mM MES (pH 6.5), 100mM NaCl sample buffer. After recording the reference NMR spectrum of the RBD domain alone (see below), CVRBDL-3 was added to a final concentration of 1 mM, resulting in a molar stoichiometry of 1.1:1 for the peptide vs. RBD domain.

2.8. NMR Spectroscopy

Two dimensional ^{1}H-^{15}N HMQC (heteronuclear multiple quantum coherence) spectra of the isolated RBD domain (natural abundance) were recorded on a Bruker Avance Neo 1.2 GHz spectrometer, equipped with a 3 mm HCN cryoprobe. Spectral dimensions were 16.02 ppm (^{1}H) × 32.89 ppm (^{15}N); 1024 and 27 complex points were recorded in the direct (^{1}H) and indirect (^{15}N) dimension, resulting in acquisition times of 53.2 ms and 6.8 ms in the direct and indirect dimension, respectively. To compensate for the low sensitivity of the natural abundance sample (^{15}N natural abundance of 0.37%), 8192 scans were recorded for each increment in the indirect dimension. The recovery delay between experiments was set to 1 s, resulting in a total experimental time of 5 d 15 h per spectrum. The experimental temperature was 20 °C. Spectra of the RBD domain in the presence of the peptide, as well as for the isolated peptide, were recorded back-to-back, under identical conditions as for the reference spectrum of the RBD domain alone. Spectra were processed using the Bruker Topsin 4 software (Billerica, MA, USA) and analyzed using the CCPN Analysis 2.3 software (Collaborative Computing Project for NMR; [41]).

2.9. Surface Plasmon Resonance (SPR) Experiments

SPR experiments were performed with a Biacore T200 device (Cytiva, USA) using 10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% (*v*/*v*) Tween-20 as assay buffer. SPR experiments were performed at a flow rate of 30 µL/min and RT if not otherwise stated.

2.10. SPR Affinity Measurements

Affinity screening for CVRBDL peptides −1 to −5 was performed by immobilization of the biotinylated SARS-CoV-2 RBD selection target (Acrobiosystems, Newark, NY, USA) on a streptavidin-functionalized sensor surface (SA sensor, Cytiva, USA). CVRBDL peptides

were injected for 300 s in a concentration range of 11.1 to 0.01 µM, followed by 600 s of dissociation phase detection. For affinity measurements on SARS-CoV-2 and COVID (2002) RBD, S1S2 monomer and S1S2 trimer constructs, the proteins were immobilized on a polycarboxylate sensor surface using EDC/sulfo-NHS chemistry at pH 5.5 (HC200M, Xantec, GE). Injections were performed for 60 s association and 100 s dissociation using CVRBDL-3 and -3_3 concentration ranges of 5 to 0.04 µM. For both setups, regeneration was performed using 2 × 30 s injections of 20 mM glycine-HCl (pH 3.5). Data evaluation was performed using Biacore T200 Evaluation Software v3.2 (Cytiva, USA).

2.11. SPR Inhibition Assay

The SPR inhibition assay was performed by immobilization of hACE-fc on a Protein A functionalized sensor surface by injecting 6 µg/mL protein for 45 s at 10 µL/min (Protein A, Cytiva, USA). Next, 50 nM of the SARS-CoV-2 S1S2 monomer was injected for 300 s with or without compound concentrations in the range of 100 µM to 6.4 nM and 5 µM to 20 nM for CVRBDL-3 and -3_3, respectively. All injections including compound were referenced with the corresponding injections of compound only on the immobilized hACE2, while the S1S2 injection without compound (before and after the concentration row) was referenced with a buffer only injection. For IC_{50} validation, the responses at injection times of t = 750 s were plotted against the $\log_{10}$ of the used compound concentration. Data fitting was performed using a Boltzmann fit model (OriginPro 2020, OriginLab, Northampton, MA, USA): $y = (A1 - A2)/(1 + e\hat{}((x - x_0)/dx)) + A2$ with A1 = initial value, A2 = final value, x_0 = inflection point = IC_{50}, and dx = time constant.

2.12. SPR Displacement Assay

The displacement assay was performed by hACE2-fc immobilization as described in the previous section. Next, 50 nM SARS-CoV-2 wt S1S2 was injected for 300 s at 30 µL/min, followed by a dissociation recording of 1200 s. During dissociation, different concentrations of CVRBDL-3 and -3_3 were injected in the range of 2 µM to 0.63 µM for 300 s. These injections were referenced with compound-only injections using the same concentration on the immobilized hACE-fc. Data evaluation was performed by plotting the response at t = 1050 s against the corresponding compound concentration. Displacement effects were fitted using a mono exponential decay model (OriginPro 2020, OriginLab, Northampton, MA, USA): $y = y_0\ A\ e\hat{}(-x/t)$ with y_0 = offset, A = amplitude, t = time constant, and $t\ 1/2 = DC_{50} = t \cdot \ln(2)$.

3. Results

3.1. Phage Display Selection on the SARS-CoV-2 RBD Target Protein

For identification of RBD binding peptide sequences, M13 phage display selection was performed on the biotinylated SARS-CoV-2 RBD target protein in three consecutive selection rounds (TS = target selection). The control selections on "empty" surfaces (same surfaces but without RBD) were either derived from selection rounds on the target (DC = direct control) or from a selection starting from the original library that was exclusively performed without target protein (ES = empty selection; Figure 1A).

During the selection process, the success of target-related sequence enrichment was verified by determination of the PFU concentration in phage elution suspensions after biopanning (Figure 1B). Here, a 270-fold increase of the output titer was observed for the TS3 as compared to TS1, while an output titer reduction by a factor of ~5 was observed when no target was present during biopanning (ES3 compared to ES1). Likewise, the titer of TS derived samples decreased when a surface without target was used for the subsequent selection round (DC). In addition, the enrichment ELISA with the input titer phage samples showed a ~5-fold increase in end-point signal on the selection target, while a low signal was identified when no selection target was present (Figure 1C).

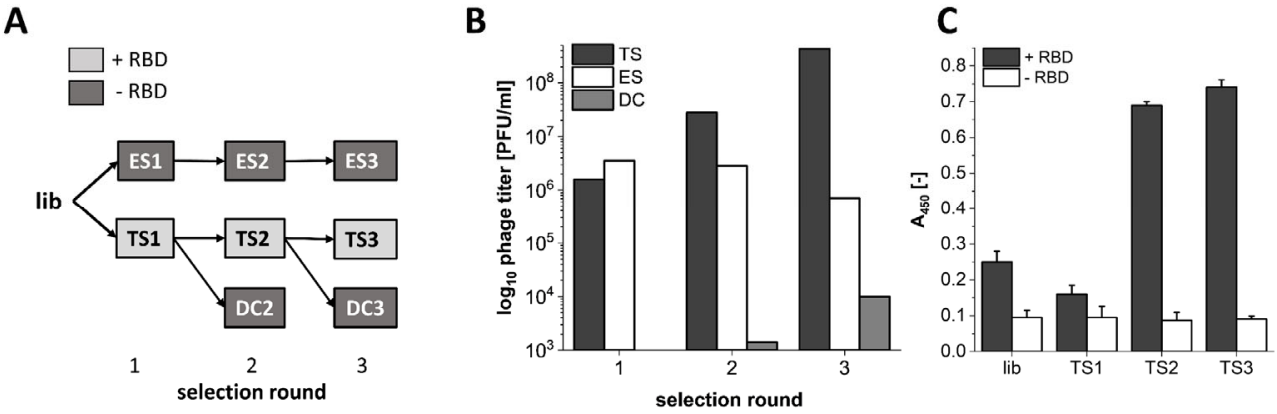

Figure 1. M13 phage display selection on SARS-CoV-2 RBD in three consecutive selection rounds. (**A**) Selection scheme: lib = naïve library; TS = target selection on RBD; DC = direct control; ES = empty selection. Samples are derived from each other in the direction of the connecting arrow starting with the naïve phage display library (lib). For TS, the selection target SARS-CoV-2 RBD was immobilized on a streptavidin-functionalized surface, while no target protein was immobilized for ES and DC. (**B**) Output phage titer in PFU/mL for selection samples of TS, ES and DC as measured past phage elution after biopanning. (**C**) Enrichment ELISA with input titer samples on the immobilized SARS-CoV-2 RBD selection target protein. Phage concentration was adjusted based on A_{265} UV-vis measurements to a total input of 2.5×10^{10} phages for all samples.

3.2. Analysis of Sequences Derived from NGS Sequencing

NGS sequencing of phage-isolated ssDNA yielded about 500,000 to 600,000 sequence reads per sample. In order to identify sequences that were enriched during phage display selection due to specific binding to the target protein, the TS3 sequences were filtered according to their enrichment on the target during the selection (lib < TS1, TS2 < TS3), their frequency in the target selection as compared to selections missing the target protein (TS1 > ES1, TS2 > ES2, TS3 > ES3), and the relative frequency decrease when a target-selection-derived sequence was subjected to a surface without target (TS2 > DC2, TS3 > DC3). In total, 90,000 unique sequences were identified that passed the filtering process. Analysis of the sequence enrichment (enrichment score [TS3/lib]) revealed that 18 unique sequences were enriched by a factor of 1000 or more. More than 50% of the total number of all sequence reads that passed the filter are represented by these 18 variants (Figure 2A). Additionally, 14 of the 18 sequences were identified with an ES of at least 1000 (empty score [TS3/ES3]), which underlines the target specificity of these sequences.

Next, the sequences were clustered according to their sequence homology in three consecutive rounds of cluster alignment and extension using Hammock clustering software (37). Clustering resulted in 16 clusters containing 57,712 unique and 383,503 total sequences, where ~96% of total sequence reads that passed the filter were assigned to five main clusters. The clusters were named CVRBDL-1 to -5, corresponding to the descending order of their total number of assigned sequence reads. The largest proportion by far, with over 70% of total sequence reads, was assigned to CVRBDL-1, followed by 10, 7, 6 and 3% for CVRBDL-2 to -5 (Figure 2B). Based on the sequence alignments, motifs were generated for each of the main clusters (Figure 2C), revealing five highly conserved sequence motifs, including amino acid residue variations, which can be used for potential sequence optimization steps. The amino acid residues with the highest occurrence at each position determined the peptide sequence of the peptides together with three C-terminally introduced arginines for enhanced solubility and C-terminal amidation: CVRBDL-1: MLDEWGWENYPSKFWHRRR-NH2; CVRBDL-2: WTYMKDPLSYWGGYYRRR-NH2; CVRBDL-3: HDFWWEYDKKNDTWTRRRR-NH2; CVRBDL-4: EVWGINFYPNRVPEKV RRR-NH2; CVRBDL-5: IGINQVAQPWWYNKSDRRR-NH2.

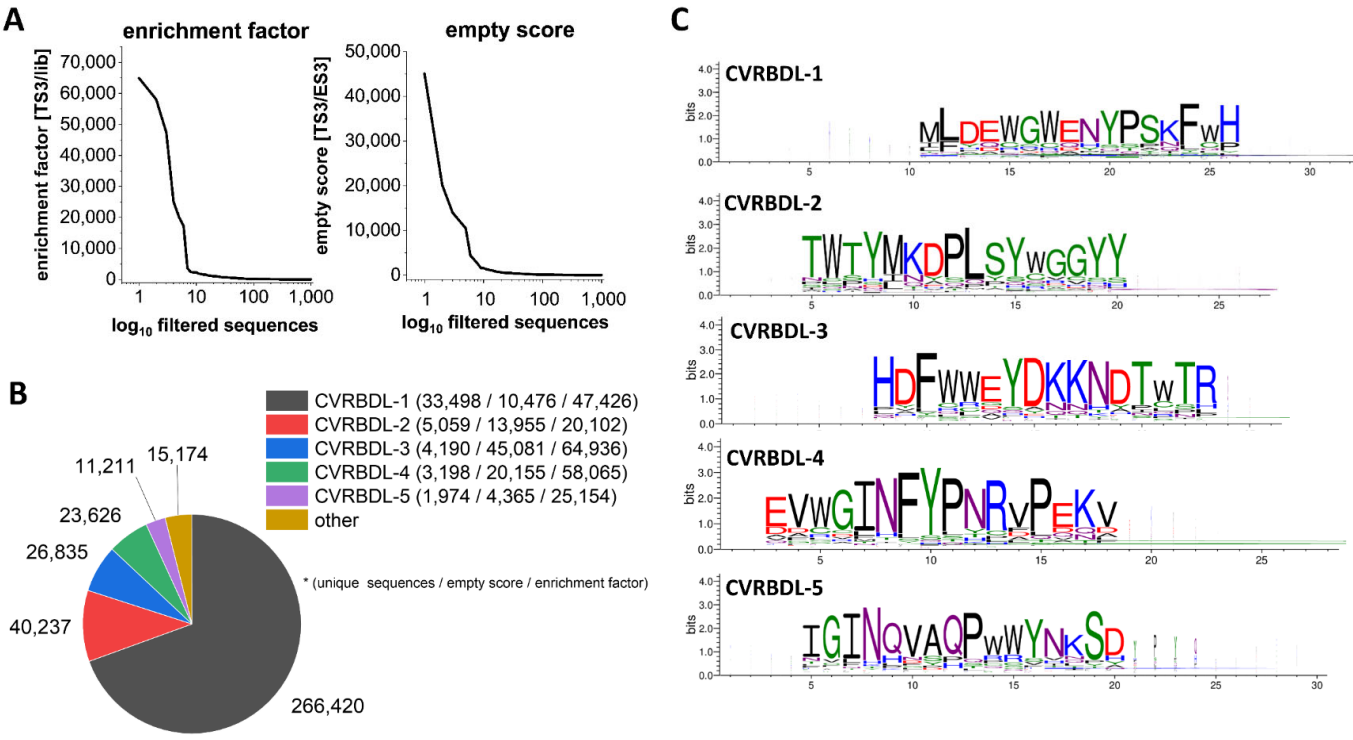

Figure 2. NGS sequence filtering and clustering reveals five main sequence motifs. (**A**) Evaluation of the enrichment factors (TS3/lib) and empty scores (TS3/ES3) of the first 1000 sequences as ranked by their enrichment factor or empty score, respectively. Sequences were filtered according to their enrichment on the target (lib > TS1 > TS2 > TS3), their frequency in the target selection as compared to the empty selection (TS1 > ES1; TS2 > ES2; TS3 > ES3) as well as their frequency decrease when the target selection was subjected to a surface without target (TS2 > DC2; TS3 > DC3). (**B**) Cluster distribution of the filtered sequences after three consecutive rounds of cluster extension and merging. The pie chart depicts the total number of sequences that were assigned to one cluster including multiple occurrences of identical sequences. The legend shows the number of unique sequences that were assigned to each cluster. Additionally, for each main sequence of the cluster, the specific empty score and enrichment factor is listed. (**C**) Conserved cluster motifs resulting from cluster assignments as shown in (**B**) derived by Hammock clustering software.

3.3. Affinity of CVRBDL Peptides towards the SARS-CoV-2 RBD

To verify whether the compounds showed affinity towards the selection target, the biotinylated SARS-CoV-2 RBD, as already used for phage display selection, was immobilized on a streptavidin-derivatized sensor surface and the compounds CVRBDL-1 to -5 were injected as analytes in a multicycle SPR experiment, which was subsequently evaluated by steady-state analysis.

All of the tested peptides showed affinity towards the immobilized SARS-CoV-2 RBD in the low µM range as determined by steady-state analysis. As typical for peptides interacting with a folded binding partner, a transient binding behavior was detected for most of the compounds (Figure 3). However, the transient binding proportion CVRBDL-1 and -2 showed a partially slower complex decay, which can be recognized by a non-mono-exponential complex decay during the dissociation phase. This suggests that multiple binding sites for these peptides may exist, which would be a reasonable explanation for these sequences to be assigned as main sequences of the most prominent clusters as shown previously (Figure 2). Interestingly, the lowest affinity constant and therefore highest complex stability was identified for CVRDBL-3 with a K_D of 1.3 µM, showing a ~70-fold higher affinity than the least stable interaction with CVRBDL-5.

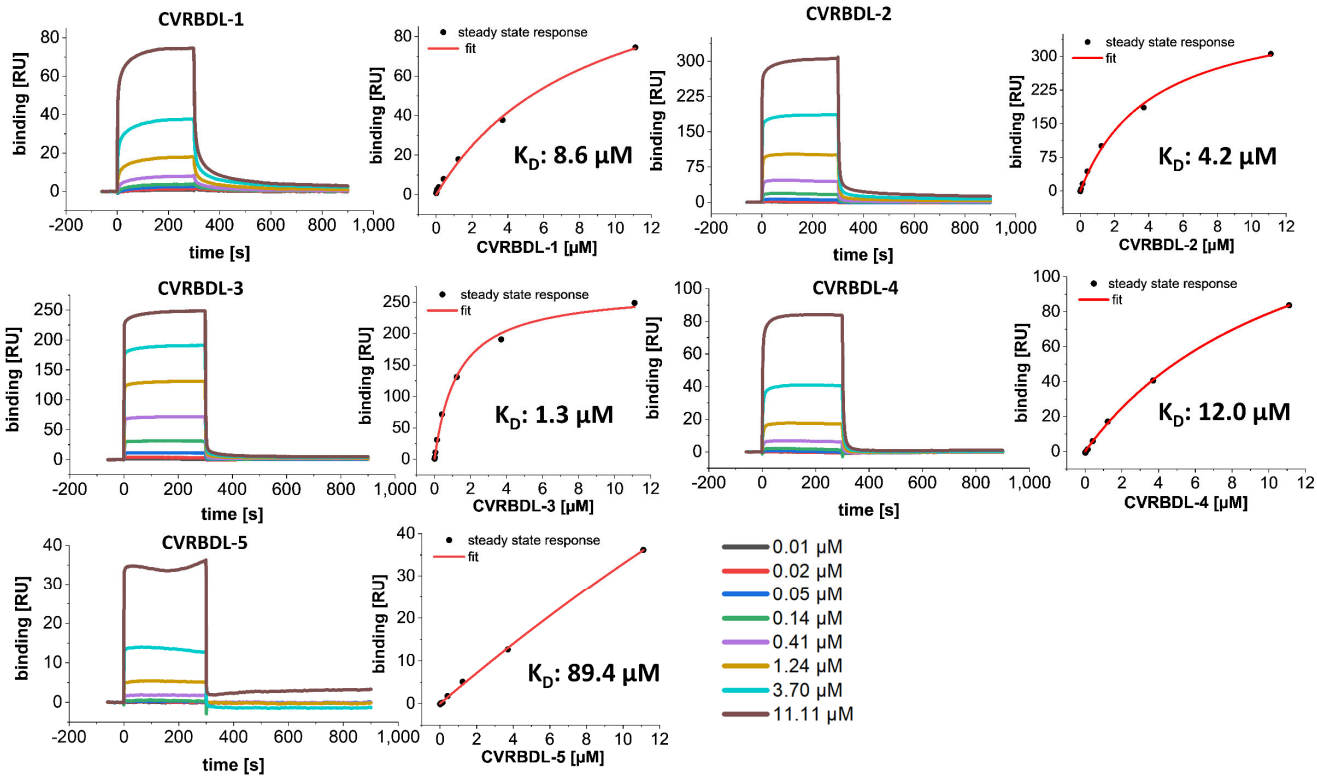

Figure 3. SPR affinity measurement of compounds CVDRBDL-1 to -5 on the immobilized SARS-CoV-2 RBD selection target. The C-terminally biotinylated SARS-CoV-2 RBD was immobilized on a streptavidin-derivatized sensor surface in a T200 SPR device (SA, Cytiva, USA). The compounds were injected for 300 s with 30 µL/min in a range of 0.01 to 11.11 µM in 10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% (*v*/*v*) Tween-20. Surface regeneration in between cycles was performed using 30 s injections of an acidic regeneration cocktail (pH 5.0), (0.15 M oxalic acid, 0.15 M H_3PO_4, 0.15 M formic acid, 0.15 M malonic acid). The K_D was determined by steady state fit as follows: CVRBDL-1: 8.6 µM ± 1.3 µM; CVRBDL-2: 4.2 µM ± 0.38 µM; CVRBDL-3: 1.3 µM ± 0.12 µM; CVRBDL-4: 12.0 µM ± 0.87 µM; CVRBDL-5: 89.4 µM ± 32.0 µM. K_D evaluation was performed using the Biacore T200 data evaluation software v. 3.2 with offset set to zero.

Further, binding properties of the peptide CVRBDL-3 to the RBD domain were investigated by solution NMR spectroscopy. Two-dimensional ^{1}H-^{15}N HMQC (Heteronuclear Multiple Quantum Coherence) spectra of an 0.85 mM RBD domain sample (unlabeled, natural abundance) were recorded using a Bruker AvanceNeo 1.2 GHz spectrometer (Figure 4), as described in the Material and Methods Section. Figure 4A shows the ^{1}H-^{15}N correlation spectrum of the RBD domain. Roughly, every resonance in the spectrum corresponds to one amino acid residue of the protein sequence, with the nitrogen resonance correlated to the proton resonance of the backbone amide group. Therefore, the HMQC spectrum serves as a "fingerprint" of the protein. Because of the natural abundance of the ^{15}N isotope of only 0.37%, the "NMR active" concentration of the 0.85 mM RBD sample corresponds to about 3 µM. Therefore, presumably only the most intense (most dynamic) residues of the RBD domain are visible in the spectra, while the more rigid regions stay below the detection threshold and thus remain invisible.

Next, the CVRBDL-3 peptide was added to the RBD domain at a molar ratio of about 1.1:1 (peptide: RBD domain). The resulting spectrum showed decreased intensities for several resonances of the RBD domain (Figure 4B, red spectrum), compared to the reference spectrum of the unbound RBD (black). A possible explanation is that residues involved in the binding become more rigidified upon peptide interaction, leading to decreased intensities, as a result of faster transverse relaxation. A second reference spectrum of the peptide only was recorded (blue, Figure 4B). Interestingly, the resonances of the peptide

(natural abundance, 1.1 higher concentration than the RBD domain) disappear in the spectra of the RBD plus peptide sample (Figure 4B, red spectrum), clearly indicating binding. Presumably, the peptide binds to the more rigid (invisible) region of the RBD and therefore becomes invisible as well (as it assumes the same NMR relaxation properties as the invisible part of the RBD).

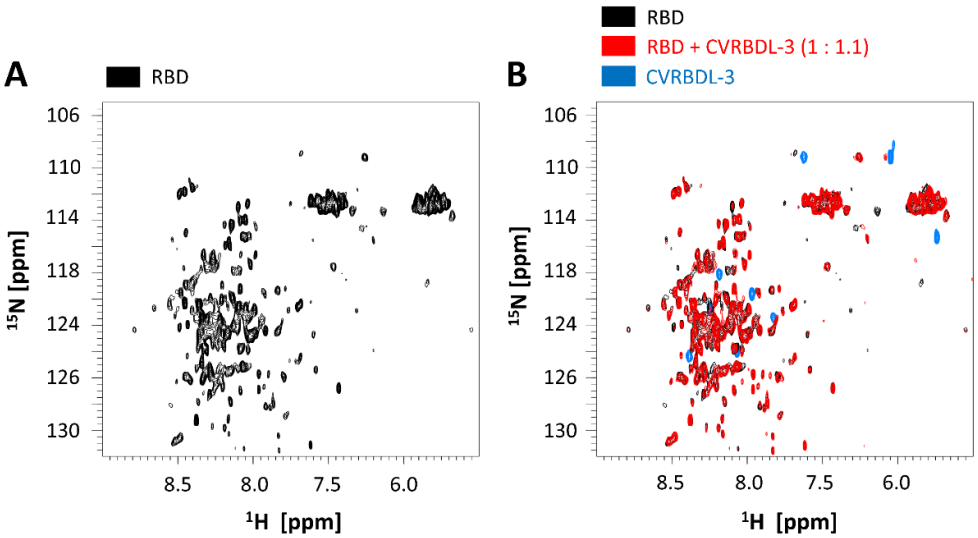

Figure 4. Two-dimensional ^{1}H-^{15}N HMQC spectra (at 1.2 GHz proton frequency) of the RBD domain (natural abundance) in absence (**A**) and presence (**B**) of the peptide CVRBDL-3. (**A**) Reference spectrum of the RBD alone. (**B**) The spectrum of the RBD domain alone (black) is overlaid with the spectra of the RBD in presence of the peptide (red), as well as the isolated peptide (blue).

3.4. Inhibition of the SARS-CoV-2 S1S2 Spike Protein Interaction with hACE2

Phage display selection was combined with NGS to identify peptides that target the SARS-CoV-2 RBD with medium to low µM affinities. To further verify whether the CVRBDL peptides are able to interfere with the association of the spike protein towards hACE2, SPR-based inhibition assays were performed. hACE-fc was immobilized on a protein A-derivatized sensor surface and the monomeric SARS-CoV-2 S1S2 protein (ectodomain of the spike protein) was injected as an analyte together with or without a 100-fold molar excess of one of the CVRBDL peptides.

Three out of five peptides (CVRBDL-2, -3 and -4) were able to interfere with association of the monomeric S1S2 spike protein to hACE2 in our experimental setup, which can be recognized by a reduction in the binding signal of the S1S2 protein to hACE2 in the steady-state phase upon compound addition (Figure 5A). Here, CVRBDL-3 showed the strongest inhibitory effect by a steady-state reduction of ~33% as compared to the setup without compound. CVRBDL-2 and -4 reduced the steady-state signal only by ~10 and 5%, respectively (Figure 5B). To verify whether the effect observed for CVRBDL-3 was concentration-dependent, compound concentrations in the range of 6.4 nM to 100 µM were applied to the inhibition setup.

When CVRBDL-3 is applied with increasing molar ratios concerning the SARS-CoV-2 S1S2 protein, a concentration-dependent inhibition of the S1S2-hACE2 complex formation can be observed (Figure 6A). Surprisingly, the compound was not able to inhibit the complex formation completely, which can be attributed either to an insufficient sterical inhibition of the hACE2 accessibility or to an inherent heterogeneity of the S1S2-hACE2 complex. The latter has already been described by us previously [42], where we identified the S1S2-hACE2 complex formation as undergoing a two-state reaction binding behavior. In this context, the compound CVRBDL-3 is able to block the primary association of the RBD with hACE2 but possibly not the secondary state transition, which is inherent to the already formed complex. However, although CVRBDL-3 is not able to completely suppress the spike protein association to hACE2, the inhibitory effect can be characterized with an

IC_{50} of 532 nM by Boltzman sigmoidal fitting, showing a maximum inhibitory saturation level of ~58% (Figure 6B).

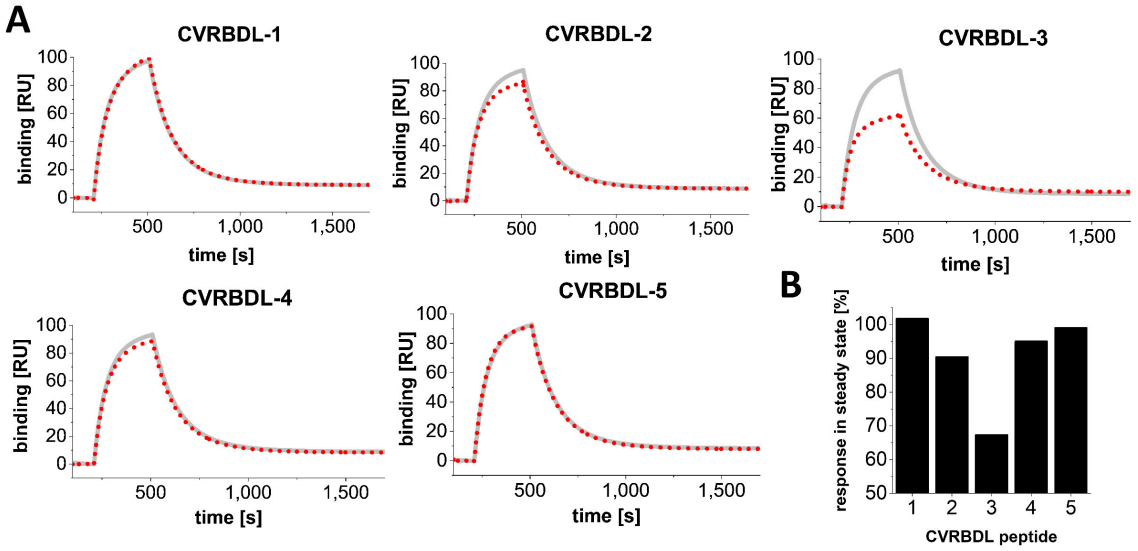

Figure 5. SPR inhibition assay of SARS-CoV-2 S1S2 spike interaction to hACE2 with and without compounds CVRBDL-1 to -5. (**A**) hACE2-fc was immobilized on a protein A/G derivatized sensor surface in a T200 SPR device (Cytiva, USA). 50 nM SARS-CoV-2 S1S2 was injected with (red dotted line) or without (solid grey line) 5 µM of compounds CVRBDL-1 to -5. Prior to each injection of S1S2 with compound, a separate injection without compound was performed. All injections with compound were referenced with compound-only injections on the hACE2 immobilized surface. Injections were performed with a flow rate of 30 µL/min in 10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% (*v*/*v*) Tween-20. (**B**) Relative steady state responses of the injections with compound as detected 270 s after injection start. CVRBDL-3 shows the strongest inhibitory effect with a steady- state reduction of ~33% as compared to SARS-CoV-2 S1S2 injection without compound.

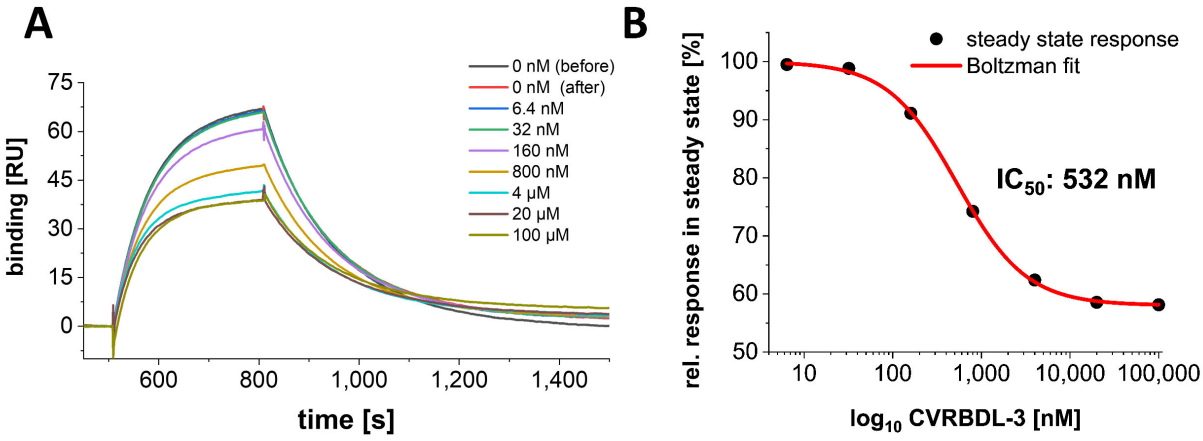

Figure 6. Concentration dependency of the inhibitory effect of CVRBDL-3 on the SARS-CoV-2 S1S2 interaction with hACE2. (**A**) hACE2-fc was immobilized on a protein A/G derivatized sensor surface in a T200 SPR device (Cytiva, USA). 50 nM SARS-CoV-2 S1S2 was injected with or without 6.4 nM to 100 µM CVRBDL-3. Injections without compound were performed prior and after the concentration row to demonstrate reproducibility of the S1S2 binding level. All injections with compound were referenced with the corresponding concentration of compound only injections on the hACE2 immobilized surface. Injections were performed with a flow rate of 30 µL/min in 10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% (*v*/*v*) Tween-20. (**B**) Relative response in steady as detected at t = 750 s. The inhibitory effect of CVRBDL-3 saturates at ~58% of the S1S2 binding level without compound. The IC_{50} was determined by a Boltzman sigmoidal fit to be 532 nM (OriginPro, OriginLab, USA).

3.5. Tail-To-Tail Construct Design: CVRBDL-3_3

In order to further increase the compound affinity towards the SARS-CoV-2 spike protein, we created a tail-to-tail construct of the compound CVRBD-3 leading to a divalent compound. This construct, named CVRBDL-3_3, consists of a lysine with two CVRBDL-3 molecules covalently linked with their C-termini to the α- and ε-amino groups of the lysine via linkers each consisting of γ-aminobutyric acid (Abu) and arginine (Figure 7). The tail-to-tail connection was expected to retain the availability of the N-termini of the CVRBDL-3 moieties, which was also exposed during phage display selection as the N-terminal extension of the M13-phage coating protein pIII.

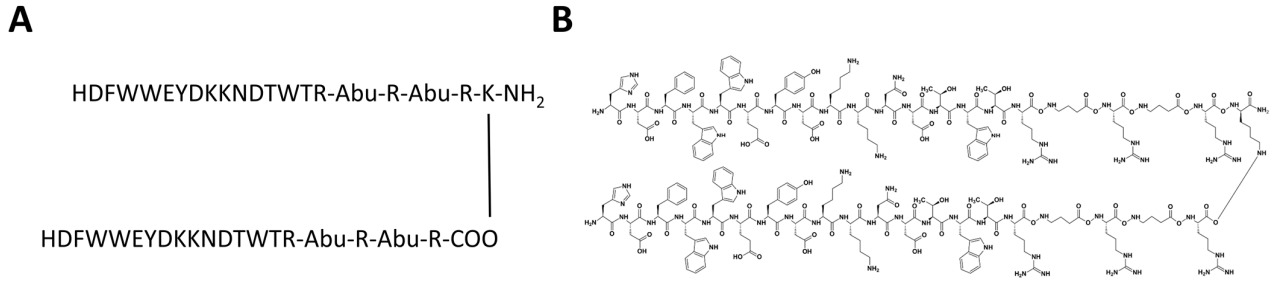

Figure 7. Amino acid single letter sequence and Natta projection of peptide construct CVRBDL-3_3. CVRBDL-3_3 was created by C-terminal attachment of two CVRBDL-3 sequences via the side chain primary amino group of a C-terminal introduced lysine. Each of the peptides carries a C-terminal [-Abu-R-Abu-R-] spacer region for increased binding flexibility and spacing of binding regions (Abu = γ-amino butyric acid). (**A**) Single letter amino acid CVRBDL-3_3 sequence. (**B**) Natta projection of the CVRBDL-3_3 sequence.

3.6. CVRBDL-3_3 Shows Increased Affinity for the CoV-2 Spike Protein Compared with CVRBDL-3

We compared the affinity of the single peptide CVRBDL-3 and the tail-to-tail combination CVRBDL-3_3 to the poly-His tagged spike protein constructs of SARS-CoV-2 RBD, S1S2 monomer and the S1S2 trimer using SPR analysis. SARS-CoV-2 RBD, S1S2 monomer and S1S2 trimer were immobilized via primary amino group coupling on a polycarboxylate sensor surface and the compounds were injected in a multicycle experiment using concentrations in the range of 0.04 to 5 μM. The subsequent evaluation was performed by steady-state analysis (Figure 8).

The interaction of CVRBDL-3 towards the SARS-CoV-2 RBD was identified to be ~two-fold more affine than for the original biotinylated selection target (K_D: 680 nM vs. 1.3 μM, respectively), whereas slightly higher affinities were determined for the S1S2 monomer and the trimeric S1S2 (K_D: 400 and 406 nM, respectively). As already observed during affinity measurement on the selection target protein, CVRBDL-3 shows again a transient binding behavior for all of the tested spike protein constructs. Additionally, the identified R_{max} was found to be lower than the theoretical R_{max} for all measurements, which indicates a 1:1 binding stoichiometry. Notably, the CVRDBL-3 binding is not influenced when the RBD is presented in the context of the monomeric and trimeric spike ectodomain. This implies that the interaction of CVRBDL-3 with the RBD binding epitope is not significantly disturbed by the RBD switch between "up-" and "down-" conformations.

CVRBDL-3_3 shows increased binding affinity as compared to the single peptide alone. The complex stability increased by a factor of ~2 towards the RBD (K_D: 327 nM) and by a factor of ~2.5 towards the monomeric and trimeric S1S2 (K_D: 162 and 154 nM, respectively). The binding kinetics changed for CVRBDL-3_3 as compared to CVRBDL-3 from a transient binding mode with fast dissociation to a more stable complex with partially slower dissociation. However, the kinetics cannot be described with a 1:1 stoichiometric Langmuir binding behavior, indicating a distinct binding mode. Moreover, the bivalent analyte model was not able to sufficiently describe the interaction's heterogeneity. A possible explanation could be that the aforementioned continuous position switching between 'up' and 'down'

conformations of the SARS-CoV-2 RBD does not influence the interaction with CVRBDL-3 on the one side, but potentially restricts the association of CVRBDL-3_3 on the other site. This interference possibly hinders the assignment of a clear kinetic model to the observed sensograms. However, since the binding kinetics of CVRBDL-3_3 appears almost identical for monomeric and trimeric S1S2, a simultaneous binding of two RBD located epitopes with an increased avidity seems implausible, leaving room for further optimization of the linker length.

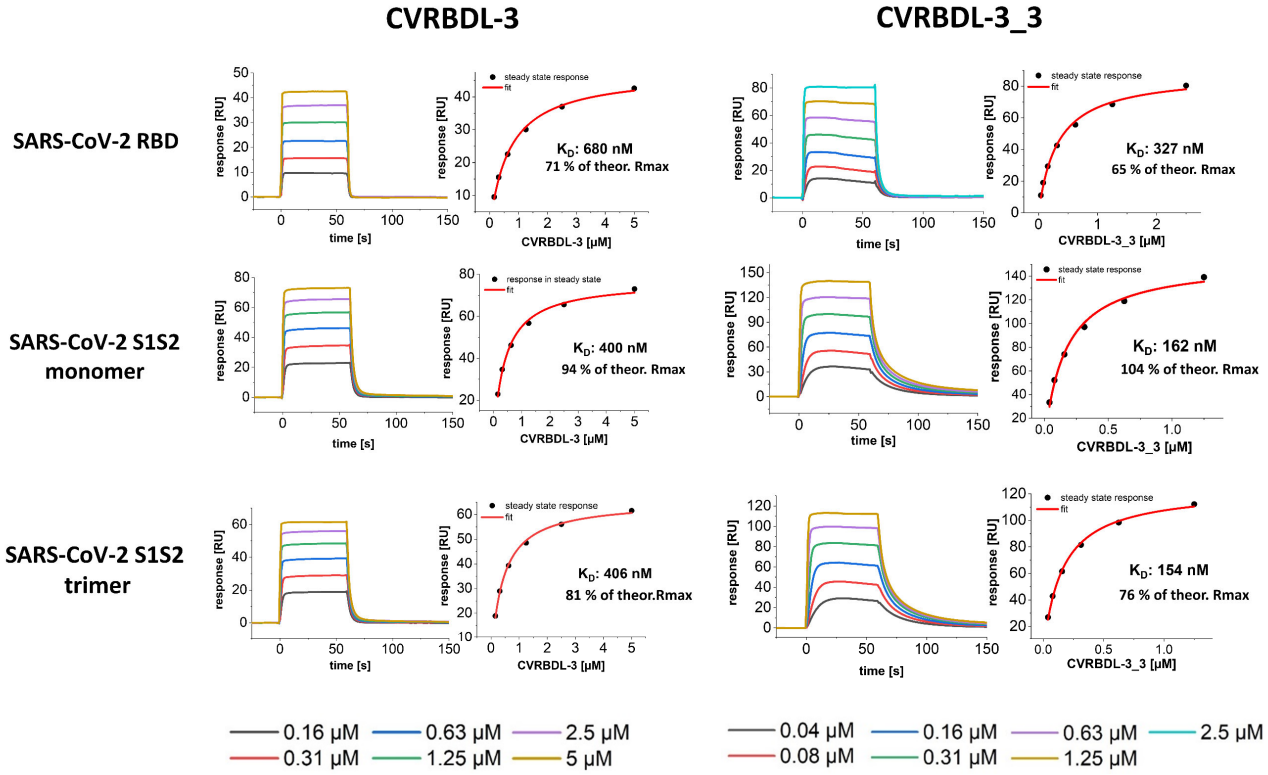

Figure 8. SPR affinity measurement of CVRBDL-3 and CVRBDL-3_3 on the SARS-CoV-2 RBD, S1S2 monomer and trimeric spike constructs. The spike constructs were immobilized on a polycarboxylate matrix by coupling via primary amino groups with sulfo-NHS ester at pH 5.5 (200 HCM, Xantec, GE). Measurements were performed in 10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% Tween-20 solution in a T200 device (Cytiva, USA). Six concentrations of each compound were injected for 60 s with a 1:1 serial dilution, with5 µM being the highest concentration for CVRBDL-3 and 2.5 µM for CVRBDL-3_3. KD evaluation was performed using Biacore T200 data evaluation software v. 3.2 with an offset of zero. The theoretical Rmax was calculated based on the molecular weight ratio of the interactants and the immobilization level after quenching.

3.7. CVRBDL-3 and CVRBDL -3_3 Show High Affinity for the B.1.1.7 Mutant Spike Trimer and Bind the SARS-CoV (2002) and SARS-CoV-2 "Closed" RBD Trimer with Reduced Affinity

We have previously shown that the compounds CVRBDL-3 and -3_3 show high affinity with nM binding constants towards the SARS-CoV-2 RBD, S1S2 monomer and trimer. Here, we additionally verify whether the binding specificity of the compounds is limited to the "wt" of SARS-CoV-2 spike protein or can also be extended to mutant versions, where some have been found to show enhanced affinity for the cellular receptor hACE2. Therefore, we analyzed the binding affinity of the compounds towards the SARS-CoV-2 B.1.1.7 S1S2 trimeric spike mutant (α-variant) as a model for other emerging variants of the spike protein. In total, seven amino acid sequence positions differ between SARS-CoV-2 "wt" spike protein and the B.1.1.7 variant, whereas four mutations are located at intermolecular interaction sites of the subunits and one mutation, namely N501Y, is found in the RBM of the RBD [43]. This amino acid residue substitution has been shown to increase the binding affinity for the hACE2, which may be one of the reasons for the observed four to

six-fold increase in transmissibility in cooperation with other mutations that do not directly interfere with receptor binding but enhance binding e.g., by increasing RBD flexibility such as D614G [44–46].

The affinity measurement on the trimeric B.1.1.7 (Figure 9) shows clearly that the binding affinity of neither CVRBDL-3 nor -3_3 is influenced by the mutation sites of the B.1.1.7 variant. Similar binding constants as already identified for the SARS-CoV-2 "wt" S1S2 spike protein were identified (528 and 243 nM for -3 and -3_3), whereas the binding kinetics shows a clear transient binding mode for CBRBDL-3 and higher complex stability with slower complex decay for CVRBDL-3_3. Since the SARS-CoV 2002 RBD shares 73% sequence identity and high structural similarity with the SARS-CoV-2 "wt" RBD [47], we speculated that the compounds are also able to bind the COVID 2002 trimeric spike protein with high affinity.

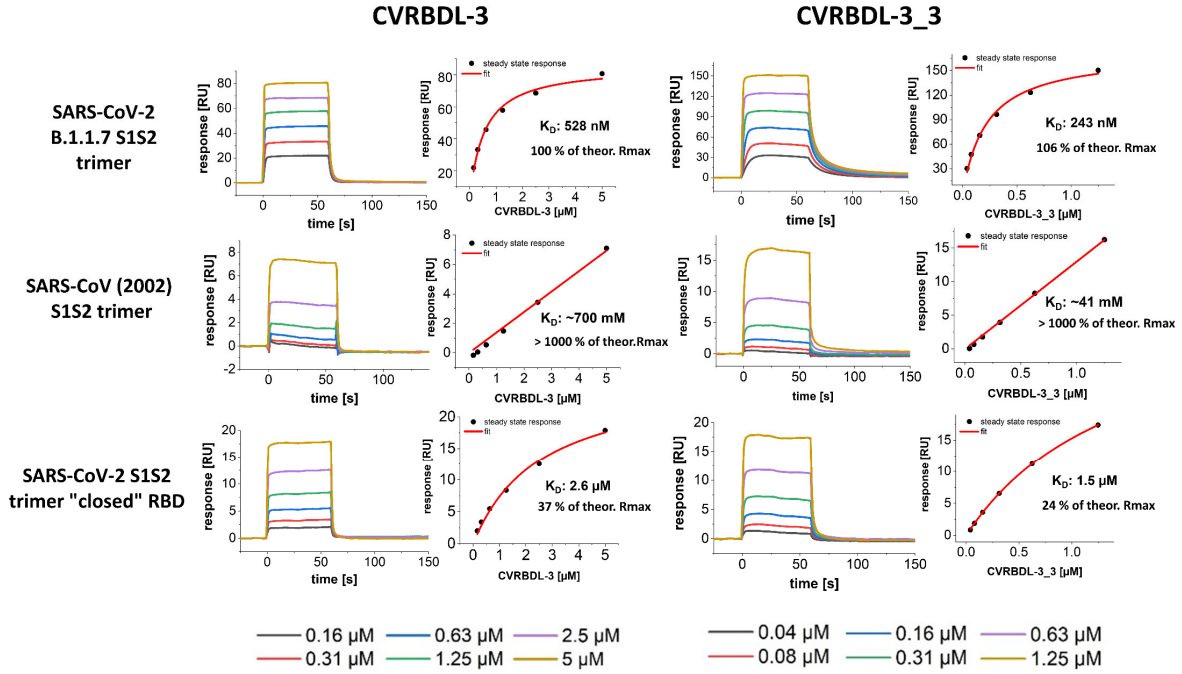

Figure 9. SPR affinity measurement of CVRBDL-3 and tail-to-tail construct CVRBDL-3_3 on the SARS-CoV-2 B.1.1.7 trimer spike, SARS-CoV (2002) trimer and the SARS-CoV-2 "closed" RBD trimeric spike. The spike constructs were immobilized on a polycarboxylate matrix by coupling via primary amino groups with sulfo-NHS ester at pH 5.5 (200 HCM, Xantec, GE). Measurements were performed in 10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% Tween-20 in a T200 device (Cytiva, USA). Six concentrations of each compound were injected for 60 s with a 1:1 serial dilution, with 5 µM being the highest concentration for CVRBDL-3 and 1.25 µM for CVRBDL-3_3. K_D evaluation was performed using Biacore T200 data evaluation software v. 3.2 with the offset set to zero. The theoretical Rmax was calculated based on the molecular weight ratio of the interactants and the immobilization level after quenching.

The affinity measurements, however, show that CVRBDL-3 and -3_3 do not have any affinity for this construct (both K_D's in high mM range). Next, we verified the binding affinity towards the "closed" SARS-CoV-2 trimer [48]. For construct design, the spike protein was manipulated by the replacement of amino acid residues at positions 383 and 985 with cysteines in order to form a disulfide bond that fixes the RBD in its "down" position. Hence, the relevant positions for interaction with hACE2 are masked in this construct. When the compounds' affinities were determined for this RBD "down" fixed S1S2 construct, significantly decreased affinities were observed. Additionally, the experimental R_{max} reached only 32 and 24% of the theoretical R_{max} for CVRBDL-3 and -3_3, respectively.

3.8. CVRBD-3_3 Efficiently Inhibits the Spike-hACE2 Complex Formation and Displaces the Spike Protein from the Pre-Formed Complex

The tail-to-tail combination of the two CVRBDL-3 peptides achieved (CVRBDL-3_3) increased binding affinities and decreased dissociation rates for the different SARS-CoV-2 constructs. Moreover, CVRBDL-3_3 retained its target specificity and the high binding affinity towards the SARS-CoV-2 B.1.1.7 mutant spike protein. Here, we compared CVRBDL-3 and CVRBDL-3_3 concerning their ability to (i) inhibit complex formation of the SARS-CoV-2 S1S2 monomer with hACE2 and (ii) displace the spike protein from the pre-formed complex, consisting of the S1S2 monomer and hACE2. Two experimental setups allowed a complete evaluation of the compound's effectiveness in either pre- or post-complex formation conditions.

We injected constant concentrations of the monomeric S1S2 spike protein with or without increasing CVRBDL-3_3 concentrations on the immobilized hACE2 receptor.

As already observed for CVRBDL-3, CVRBDL-3_3 also shows a concentration-dependent reduction of the SARS-CoV-2 S1S2 association towards hACE2 (Figure 10A). Interestingly, the inhibition saturates at 58% of the binding level without compound, which is identical with the saturation level that was previously identified for CVRBDL-3. Evaluation of the IC_{50} (Figure 10B) shows that the inhibition efficiency of complex formation increased by a factor of ~11 as compared to CVRBDL-3 (532 nM as compared to 47 nM for CVRBDL-3 and CVRBDL-3_3, respectively).

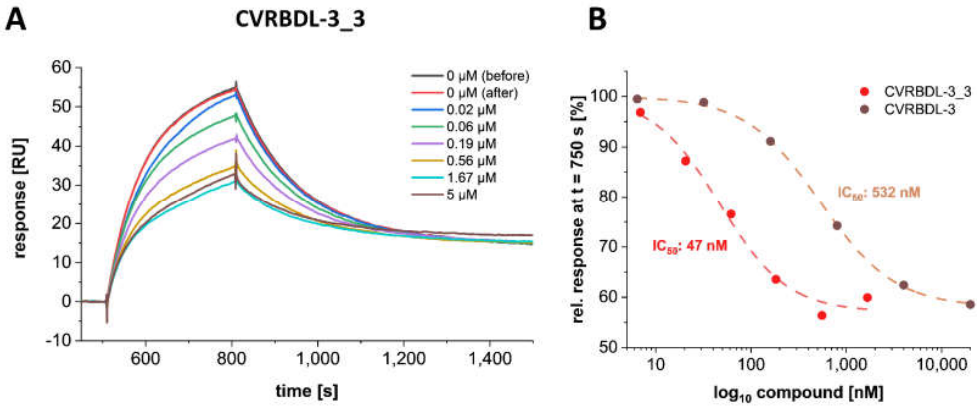

Figure 10. CVRBDL-3_3 much more efficiently inhibits the SARS-CoV-2 S1S2 spike protein interaction with hACE2 as compared to CVRBDL-3. (**A**) hACE2-fc was immobilized on a Protein A/G-derivatized sensor surface in a T200 SPR device (Cytiva, USA). 50 nM SARS-CoV-2 S1S2 was injected with or without 6.4 nM to 100 µM CVRBDL-3_3. Injections without compound were performed prior and after the concentration row to verify reproducibility of the S1S2 binding level. All injections with compound and spike protein were referenced with the corresponding concentration of compound only injections on the hACE2 immobilized surface. Injections were performed with 30 µL/min in 10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% (*v*/*v*) Tween-20. (**B**) Relative response in steady state as detected 270 s after injection start (for CVRBDL-3 see Figure 6). The inhibitory effect of CVRBDL-3 and CVRBDL-3_3 saturates at ~58% of the S1S2 binding level without compound. The IC_{50} was determined by Boltzman sigmoidal fit to be 532 nM (OriginPro, OriginLab, USA).

Next, we analyzed the effectiveness of CVRBDL-3 and -3_3 in displacing SARS-CoV-2 S1S2 monomers from the pre-formed complex with hACE2. For this assay, a constant concentration of monomeric S1S2 was injected on pre-immobilized hACE2 for all injection cycles. During dissociation, increasing concentrations of the compounds were applied to the pre-formed complex.

The sensogram course during compound injection indicates whether the compound presence accelerates the complex decay by shifting the equilibrium between hACE2 bound and unbound spike protein to the dissociated form (Figure 11A,B). For both compounds, the effect of spike protein displacement was observed by a concentration-dependent acceler-

ation of signal reduction during the compound injection phase. When the remaining signal at t = 1050 s was plotted against the injected compound concentration, the concentration dependent effect is best described with a mono-exponential decay function (Figure 11C). Here, again CVRBDL-3_3 shows an enhanced effect as compared to CVRBDL-3. Although the concentration needed for reaching 50% of the maximum displacement just slightly differs between the two compounds (DC_{50}: 510 and 417 nM for CVRBDL-3 and 3_3, respectively), the maximum displacement of CVRBDL-3_3 surpasses the one observed for CVRBDL-3 by a factor of 1.6 (y_0: 84 and 74% for CVRBDL-3 and 3_3, respectively).

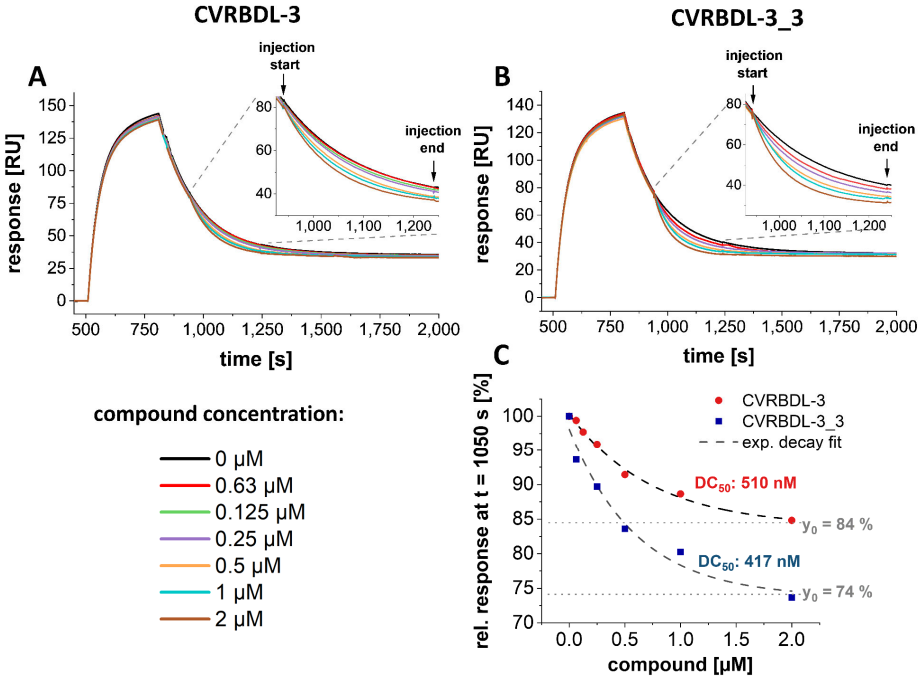

Figure 11. CVRBDL-3 and CVRBDL-3_3 displace hACE2 from the pre-formed complex with SARS-CoV-2 S1S2. (**A**,**B**) hACE2-fc was immobilized on a Protein A/G-derivatized sensor surface in a T200 SPR device (Cytiva, USA). For all samples, 50 nM SARS-CoV-2 S1S2 was injected for 300 s with 30 µL/min. After dissociation start 300 s of CVRBDL-3 (**A**) or CVRBDL-3_3 (**B**) was injected in a concentration range of 62.5 nM to 2 µM within six steps. All injections including compound were referenced with the corresponding concentration of compound only injections on the hACE2 immobilized surface. Injections were performed with 30 µL/min in 10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% (*v*/*v*) Tween-20. (**C**) The response displayed in (**A**) and (**B**) at t = 1050 s of the injection without compound was set to 100%. Relative responses of the compound injections at t = 1050 s were plotted against the corresponding compound concentration. Data fitting was performed using a mono-exponential decay fit (OriginPro, OriginLab, USA). DC_{50} [nM] = compound concentration where 50% of the maximum displacement (y_0) is reached. y_0 [%] = offset; maximum displacement.

In conclusion, both compounds show inhibitory effects when applied to pre- or post-spike-hACE2 complex formation conditions, while CVRBDL-3_3 shows enhanced inhibition as compared to CVRBDL-3.

4. Discussion

In the present study, we used phage display selection in combination with a unique and very efficient NGS-based data analysis strategy to identify ligands for the SARS-CoV-2 RBD. Filtering and clustering of the NGS-derived sequences using positive and negative control samples led to the identification of highly conserved cluster motifs that cover 96% of all sequences passing the analysis filter criteria and are represented by five distinct main sequences (Figures 1 and 2). Further affinity analysis showed all selected peptides bind the SARS-CoV-2 RBD selection target with affinities in the low µM range,

where the highest affinity was identified for sequence CVRBDL-3 with a K_D of 1.3 µM (Figure 3). The existence of a specific interaction between CVRBDL-3 and SARS-CoV-2 RBD was further verified by 2D-NMR using ^{1}H-^{15}N natural abundance correlation of the RBD (Figure 5). Here, CVRBDL-3 selectively reduced side chain resonances of the SARS-CoV-2 RBD when added at equimolar concentrations. In addition, natural resonances of the peptide were invisible when the RBD was present, suggesting a high affinity interaction between RBD and CVRBDL-3.

Further analysis in an SPR-based experiment using the isolated SARS-CoV-2 S1S2 spike and hACE2 proteins demonstrated that CVRBDL-3 inhibits the complex formation when it is injected together with the monomeric S1S2 spike protein on the immobilized hACE2 (Figure 5). CVRBDL-3 also displaced hACE2 from the spike protein when the complex was already preformed (Figure 11). These two experimental SPR set-ups resemble physiological conditions where the virus has either not attached to the cell surface or the spike protein is already attached to the cellular receptor. A potent inhibitor targeting fusion inhibition can be expected to interfere with pre- as well as post-complex formation states. Because CVRBDL-3 is efficient in both, the targeted binding site of CVRBDL-3 on the RBD surface appears to be critical for its complex formation with hACE2 and remains at least partially accessible for CVRBDL-3 once the complex with hACE2 has been formed.

The divalent version of CVRBDL-3, CVRBDL-3_3, had a two- to three-fold higher affinity for the SARS-CoV-2 spike protein (RBD, S1S2 monomer, S1S2 trimer) as compared to CVRBDL-3. The dissociation rates for the interaction of CVRBDL-3_3 with monomeric and trimeric S1S2 were substantially decreased for CVRBDL-3_3. Because there was no difference between the K_D values of CVRBDL-3_3 and monomeric and trimeric SARS-CoV-2 S1S2 spike protein (Figure 8), we conclude that the designed divalent binding mode was realized, but the increased affinity is much more likely due to the increased local concentration of the CVRBDL-3 moiety leading to decreased dissociation rates due to enhanced rebinding effects [49]. This leaves room for much more enhanced affinities by optimizing the linker length within CVRBDL-3_3 in the future. Nevertheless, CVRBDL-3_3 demonstrated increased binding affinity as compared to CVRBDL-3 in pre- as well as in post-complex formation conditions (Figures 10 and 11).

To further verify the binding specificity of CVRBDL-3 and CVRBDL-3_3, we tested their affinity towards the trimeric spike protein constructs of SARS-CoV-2 B.1.1.7, SARS-CoV (2002) and a "closed" SARS-CoV-2 spike, for which the RBD was fixed in a "down" position. CVRBDL-3 and CVRBDL-3_3 bind the B.1.1.7 mutant spike protein as efficiently as the "wt" spike trimer. This finding supports the idea that inhibitory binding sites do not necessarily overlap with the typical VOC mutational sites and thus future emerging variants may still be vulnerable to the compounds. On the other hand, the spike protein of the "old" SARS-CoV virus from the year 2002 did not show efficient binding by CVRBDL-3 and CVRBDL-3_3. This result indicates that the binding site and the observed inhibitory effect of CVRBDL-3 and CVRBDL-3_3 is specific for SARS-CoV-2 and probably not interchangeable with other SARS coronaviruses. The binding affinity, as well as the fraction of the theoretical R_{max} that is achieved towards the "closed" RBD trimeric CoV-2 constructs, is substantially decreased compared to the non-manipulated "wt" version. This result points to a decisive role of the inaccessible regions in the "closed" trimer for interaction with the compounds.

In conclusion, affinity measurements clearly show that the CVRBDL-3 and CVRBDL-3_3 compounds target a binding site of the RBD that (i) is specific to SARS-CoV-2, (ii) is conserved in a mutant version, (iii) shows reduced accessibility in the "closed" conformation, and finally (iv) is critical for the interaction with hACE2. Phage display is an appropriate strategy for the identification of potent SARS-CoV-2 fusion inhibitors. This concept is not limited to coronaviruses only, but is transferable to other viruses, opening an option for a general and straightforward identification of drug compounds against novel emerging VOCs. The presented compound selection strategy yields small peptide compounds that are not template-protein-derived and thus are less likely prone to exhibit

undesired side effects under physiological conditions. This can, for example, be expected for hACE2-template-based peptides [50,51]. The expected poor bioavailability of such L-enantiomeric peptide therapeutics may be overcome by their conversion into D-enantiomeric derivatives, for example by "retro-inverso" strategies [52] that endow all-D-enantiomeric peptide compounds with enhanced proteolytic stability [53] but conserve binding specificity and affinity. After optimization of proteolytic stability and binding affinity, future studies will first address testing of the optimized compounds in viral proliferation assays to ensure transferability of the observed effects to the cellular level.

Author Contributions: Conceived and designed the project outline: M.S. (Marc Sevenich), J.M. and D.W.; planned and performed experiments: phage display (M.S. (Marc Sevenich), E.T. and J.M.); next-generation sequencing (M.S. (Marc Sevenich) and P.P.); expression and purification of COVID and hACE-2 constructs (T.K., M.S. (Maren Schubert), F.B. and J.v.d.H.); nuclear magnetic resonance (M.S. (Marc Sevenich) and N.-A.L.); surface plasmon resonance (M.S. (Marc Sevenich)); wrote the manuscript: M.S. (Marc Sevenich) and D.W. with the help of all other authors. All authors have read and agreed to the published version of the manuscript.

Funding: N.-A.L. thanks the German Science Foundation (DFG, grant project number 433700474) for funding as well as support by the project "Virological and immunological determinants of COVID-19 pathogenesis—lessons to get prepared for future pandemics (KA1-Co-02 "COVIPA")", a grant from the Helmholtz Association's Initiative and Networking Fund.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: Computational support of the Zentrum für Informations- und Medientechnologie, especially the HPC (High Performance Computing) team at the Heinrich Heine University Düsseldorf is acknowledged.

Conflicts of Interest: M.S., J.M. and D.W. are co-inventor of patents covering the composition of matter of CVRBDL-peptides. All other authors declare no competing interest.

Appendix A

Table A1. SARS-CoV, SARS-CoV-2 spike protein and hACE2 constructs used in the study including organism used for expression, introduced tags for multimerization and purification, molecular weight (MW) and experimental usage.

Construct Name	Expression Organism	Tag	M_W Monomer [kDa]	Experimental Usage
SARS-CoV-2 RBD (Acrobiosystems)	HEK293 cells	Avitag, 6xHis-tag	28.2	Phage display screening, affinity screening
SARS-CoV-2 RBD	High Five insect cells	6x His-tag	31.8	Affinity measurement with CVRBDL-3 and CVRBDL-3_3, NMR
SARS-CoV-2 S1S2-His	High Five insect cells	6x His-tag	133.13	Inhibition and displacement analysis with CVRBDL-3 and CVRBDL-3_3
SARS-CoV (2002) S1S2-His trimer	High Five insect cells	6x His-tag	132.4	Affinity measurement with CVRBDL-3 and CVRBDL-3_3
SARS CoV-2 S1S2 trimer	High Five insect cells	6x His-tag, T4-foldon	141.9	Affinity measurement with CVRBDL-3 and CVRBDL-3_3
SARS CoV-2 B.1.1.7 S1S2	High Five insect cells	6x His-tag, T4-foldon	141.6	Affinity measurement with CVRBDL-3 and CVRBDL-3_3
hACE2	HEK293-6E cells	IgG1-Fc	95.2	Inhibition and displacement analysis with CVRBDL-3 and CVRBDL-3_3

SARS-CoV-2 S1S2-His monomer (M_w: 133.13 kDa)

```
     10         20         30         40         50         60
MAQCVNLTTR TQLPPAYTNS FTRGVYYPDK VFRSSVLHST QDLFLPFFSN VTWFHAISGT
     70         80         90        100        110        120
NGTKRFDNPV LPFNDGVYFA STEKSNIIRG WIFGTTLDSK TQSLLIVNNA TNVVIKVCEF
    130        140        150        160        170        180
QFCNDPFLGV YHKNNKSWME SEFRVYSSAN NCTFEYVSQP FLMDLEGKQG NFKNLREFVF
    190        200        210        220        230        240
KNIDGYFKIY SKHTPINLVR DLPQGFSALE PLVDLPIGIN ITRFQTLLAL HRSYLTPGDS
    250        260        270        280        290        300
SSGWTAGAAA YYVGYLQPRT FLLKYNENGT ITDAVDCALD PLSETKCTLK SFTVEKGIYQ
    310        320        330        340        350        360
TSNFRVQPTE SIVRFPNITN LCPFGEVFNA TRFASVYAWN RKRISNCVAD YSVLYNSASF
    370        380        390        400        410        420
STFKCYGVSP TKLNDLCFTN VYADSFVIRG DEVRQIAPGQ TGKIADYNYK LPDDFTGCVI
    430        440        450        460        470        480
AWNSNNLDSK VGGNYNYLYR LFRKSNLKPF ERDISTEIYQ AGSTPCNGVE GFNCYFPLQS
    490        500        510        520        530        540
YGFQPTYGVG YQPYRVVVLS FELLHAPATV CGPKKSTNLV KNKCVNFNFN GLTGTGVLTE
    550        560        570        580        590        600
SNKKFLPFQQ FGRDIDDTTD AVRDPQTLEI LDITPCSFGG VSVITPGTNT SNQVAVLYQG
    610        620        630        640        650        660
VNCTEVPVAI HADQLTPTWR VYSTGSNVFQ TRAGCLIGAE HVNNSYECDI PIGAGICASY
    670        680        690        700        710        720
QTQTNSHGSA SSVASQSIIA YTMSLGAENS VAYSNNSIAI PINFTISVTT EILPVSMTKT
    730        740        750        760        770        780
SVDCTMYICG DSTECSNLLL QYGSFCTQLN RALTGIAVEQ DKNTQEVFAQ VKQIYKTPPI
    790        800        810        820        830        840
KDFGGFNFSQ ILPDPSKPSK RSFIEDLLFN KVTLADAGFI KQYGDCLGDI AARDLICAQK
    850        860        870        880        890        900
FNGLTVLPPL LTDEMIAQYT SALLAGTITS GWTFGAGAAL QIPFAMQMAY RFNGIGVTQN
    910        920        930        940        950        960
VLYENQKLIA NQFNSAIGKI QDSLSSTASA LGKLQDVVNQ NAQALNTLVK QLSSNFGAIS
    970        980        990       1000       1010       1020
SVLNDILARL DPPEAEVQID RLITGRLQSL QTYVTQQLIR AAEIRASANL AATKMSECVL
   1030       1040       1050       1060       1070       1080
GQSKRVDFCG KGYHLMSFPQ SAPHGVVFLH VTYVPAQEKN FTTAPAICHD GKAHFPREGV
   1090       1100       1110       1120       1130       1140
FVSNGTHWFV TQRNFYEPQI ITTHNTFVSG NCDVVIGIVN NTVYDPLQPE LDSFKEELDK
   1150       1160       1170       1180       1190       1200
YFKNHTSPDV DLGDISGINA SVVNIQKEID RLNEVAKNLN ESLIDLQELG KYEQAAAHHH
                                                        HHH
```

SARS-CoV (2002) S1S2-His trimer without foldon sequence (M_w monomer: 132.4 kDa)

```
     10         20         30         40         50         60
MGSDLDRCTT FDDVQAPNYT QHTSSMRGVY YPDEIFRSDT LYLTQDLFLP FYSNVTGFHT
     70         80         90        100        110        120
INHTFDNPVI PFKDGIYFAA TEKSNVVRGW VFGSTMNNKS QSVIIINNST NVVIRACNFE
    130        140        150        160        170        180
LCDNPFFAVS KPMGTQTHTM IFDNAFNCTF EYISDAFSLD VSEKSGNFKH LREFVFKNKD
    190        200        210        220        230        240
GFLYVYKGYQ PIDVVRDLPS GFNTLKPIFK LPLGINITNF RAILTAFSPA QDTWGTSAAA
    250        260        270        280        290        300
YFVGYLKPTT FMLKYDENGT ITDAVDCSQN PLAELKCSVK SFEIDKGIYQ TSNFRVVPSG
    310        320        330        340        350        360
DVVRFPNITN LCPFGEVFNA TKFPSVYAWE RKKISNCVAD YSVLYNSTFF STFKCYGVSA
    370        380        390        400        410        420
TKLNDLCFSN VYADSFVVKG DDVRQIAPGQ TGVIADYNYK LPDDFMGCVL AWNTRNIDAT
    430        440        450        460        470        480
STGNYNYKYR YLRHGKLRPF ERDISNVPFS PDGKPCTPPA LNCYWPLNDY GFYTTTGIGY
    490        500        510        520        530        540
QPYRVVVLSF ELLNAPATVC GPKLSTDLIK NQCVNFNFNG LTGTGVLTPS SKRFQPFQQF
    550        560        570        580        590        600
GRDVSDFTDS VRDPKTSEIL DISPCSFGGV SVITPGTNAS SEVAVLYQDV NCTDVSTAIH
    610        620        630        640        650        660
ADQLTPAWRI YSTGNNVFQT QAGCLIGAEH VDTSYECDIP IGAGICASYH TVSLLRSTSQ
    670        680        690        700        710        720
KSIVAYTMSL GADSSIAYSN NTIAIPTNFS ISITTEVMPV SMAKTSVDCN MYICGDSTEC
    730        740        750        760        770        780
ANLLLQYGSF CTQLNRALSG IAAEQDRNTR EVFAQVKQMY KTPTLKYFGG FNFSQILPDP
    790        800        810        820        830        840
LKPTKRSFIE DLLFNKVTLA DAGFMKQYGE CLGDINARDL ICAQKFNGLT VLPPLLTDDM
    850        860        870        880        890        900
IAAYTAALVS GTATAGWTFG AGAALQIPFA MQMAYRFNGI GVTQNVLYEN QKQIANQFNK
    910        920        930        940        950        960
AISQIQESLT TTSTALGKLQ DVVNQNAQAL NTLVKQLSSN FGAISSVLND ILSRLDKVEA
    970        980        990       1000       1010       1020
EVQIDRLITG RLQSLQTYVT QQLIRAAEIR ASANLAATKM SECVLGQSKR VDFCGKGYHL
   1030       1040       1050       1060       1070       1080
MSFPQAAPHG VVFLHVTYVP SQERNFTTAP AICHEGKAYF PREGVFVFNG TSWFITQRNF
   1090       1100       1110       1120       1130       1140
FSPQIITTDN TFVSGNCDVV IGIINNTVYD PLQPELDSFK EELDKYFKNH TSPDVDLGDI
          1150       1160       1170       1180       1190
       SGINASVVNI QKEIDRLNEV AKNLNESLID LQELGKYEQY IKWPAAAHHH HHH
```

Figure A1. *Cont*.

SARS-CoV-2 RBD-His (M_w: 31.8 kDa)

```
          10         20         30         40         50         60
MARVQPTESI VRFPNITNLC PFGEVFNATR FASVYAWNRK RISNCVADYS VLYNSASFST
          70         80         90        100        110        120
FKCYGVSPTK LNDLCFTNVY ADSFVIRGDE VRQIAPGQTG KIADYNYKLP DDFTGCVIAW
         130        140        150        160        170        180
NSNNLDSKVG GNYNYLYRLF RKSNLKPFER DISTEIYQAG STPCNGVEGF NCYFPLQSYG
                  190        200        210        220        230        240
FQPTNGVGYQ PYRVVVLSFE LLHAPATVCG PKKSTNLVKN KCVNFNFNGL TGTGVLTESN
                           250        260        270        280
KKFLPFQQFG RDIADTTDAV RDPQTLEILD ITPCSAAAHH HHHH
```

SARS CoV-2 S1S2 trimer with foldon sequence (M_w monomer: 141.9 kDa)

```
          10         20         30         40         50         60
MAQCVNLTTR TQLPPAYTNS FTRGVYYPDK VFRSSVLHST QDLFLPFFSN VTWFHAIHVS
          70         80         90        100        110        120
GTNGTKRFDN PVLPFNDGVY FASTEKSNII RGWIFGTTLD SKTQSLLIVN NATNVVIKVC
         130        140        150        160        170        180
EFQFCNDPFL GVYYHKNNKS WMESEFRVYS SANNCTFEYV SQPFLMDLEG KQGNFKNLRE
         190        200        210        220        230        240
FVFKNIDGYF KIYSKHTPIN LVRDLPQGFS ALEPLVDLPI GINITRFQTL LALHRSYLTP
                  250        260        270        280        290        300
GDSSSGWTAG AAAYYVGYLQ PRTFLLKYNE NGTITDAVDC ALDPLSETKC TLKSFTVEKG
                  310        320        330        340        350        360
IYQTSNFRVQ PTESIVRFPN ITNLCPFGEV FNATRFASVY AWNRKRISNC VADYSVLYNS
         370        380        390        400        410        420
ASFSTFKCYG VSPTKLNDLC FTNVYADSFV IRGDEVRQIA PGQTGKIADY NYKLPDDFTG
                  430        440        450        460        470        480
CVIAWNSNNL DSKVGGNYNY LYRLFRKSNL KPFERDISTE IYQAGSTPCN GVEGFNCYFP
                  490        500        510        520        530        540
LQSYGFQPTN GVGYQPYRVV VLSFELLHAP ATVCGPKKST NLVKNKCVNF NFNGLTGTGV
                  550        560        570        580        590        600
LTESNKKFLP FQQFGRDIAD TTDAVRDPQT LEILDITPCS FGGVSVITPG TNTSNQVAVL
                  610        620        630        640        650        660
YQDVNCTEVP VAIHADQLTP TWRVYSTGSN VFQTRAGCLI GAEHVNNSYE CDIPIGAGIC
         670        680        690        700        710        720
ASYQTQTNSP GSASSVASQS IIAYTMSLGA ENSVAYSNNS IAIPTNFTIS VTTEILPVSM
                  730        740        750        760        770        780
TKTSVDCTMY ICGDSTECSN LLLQYGSFCT QLNRALTGIA VEQDKNTQEV FAQVKQIYKT
         790        800        810        820        830        840
PPIKDFGGFN FSQILPDPSK PSKRSFIEDL LFNKVTLADA GFIKQYGDCL GDIAARDLIC
                  850        860        870        880        890        900
AQKFNGLTVL PPLLTDEMIA QYTSALLAGT ITSGWTFGAG AALQIPFAMQ MAYRFNGIGV
                  910        920        930        940        950        960
TQNVLYENQK LIANQFNSAI GKIQDSLSST ASALGKLQDV VNQNAQALNT LVKQLSSNFG
         970        980        990       1000       1010       1020
AISSVLNDIL SRLDPPEAEV QIDRLITGRL QSLQTYVTQQ LIRAAEIRAS ANLAATKMSE
                 1030       1040       1050       1060       1070       1080
CVLGQSKRVD FCGKGYHLMS FPQSAPHGVV FLHVTYVPAQ EKNFTTAPAI CHDGKAHFPR
                 1090       1100       1110       1120       1130       1140
EGVFVSNGTH WFVTQRNFYE PQIITTDNTF VSGNCDVVIG IVNNTVYDPL QPELDSFKEE
                 1150       1160       1170       1180       1190       1200
LDKYFKNHTS PDVDLGDISG INASVVNIQK EIDRLNEVAK NLNESLIDLQ ELGKYEQAAA
                 1210       1220       1230       1240       1250       1260
LVPRGSPGSG YIPEAPRDGQ AYVRKDGEWV LLSTFLVENL YFQGGSAWSH PQFEKGGGSG
                                    1270       1280
                          GGSGGSAWSH PQFEKLEHHH HHHHH*
```

Figure A1. *Cont.*

SARS CoV-2 B.1.1.7 S1S2 trimer with foldon sequence (M_w monomer: 141.6 kDa)

```
    10         20         30         40         50         60
MAQCVNLTTR TQLPPAYTNS FTRGVYYPDK VFRSSVLHST QDLFLPFFSN VTWFHAISGT
    70         80         90        100        110        120
NGTKRFDNPV LPFNDGVYFA STEKSNIIRG WIFGTTLDSK TQSLLIVNNA TNVVIKVCEF
   130        140        150        160        170        180
QFCNDPFLGV YHKNNKSWME SEFRVYSSAN NCTFEYVSQP FLMDLEGKQG NFKNLREFVF
   190        200        210        220        230        240
KNIDGYFKIY SKHTPINLVR DLPQGFSALE PLVDLPIGIN ITRFQTLLAL HRSYLTPGDS
   250        260        270        280        290        300
SSGWTAGAAA YYVGYLQPRT FLLKYNENGT ITDAVDCALD PLSETKCTLK SFTVEKGIYQ
   310        320        330        340        350        360
TSNFRVQPTE SIVRFPNITN LCPFGEVFNA TRFASVYAWN RKRISNCVAD YSVLYNSASF
   370        380        390        400        410        420
STFKCYGVSP TKLNDLCFTN VYADSFVIRG DEVRQIAPGQ TGKIADYNYK LPDDFTGCVI
   430        440        450        460        470        480
AWNSNNLDSK VGGNYNYLYR LFRKSNLKPF ERDISTEIYQ AGSTPCNGVE GFNCYFPLQS
   490        500        510        520        530        540
YGFQPTYGVG YQPYRVVVLS FELLHAPATV CGPKKSTNLV KNKCVNFNFN GLTGTGVLTE
   550        560        570        580        590        600
SNKKFLPFQQ FGRDIDDTTD AVRDPQTLEI LDITPCSFGG VSVITPGTNT SNQVAVLYQG
   610        620        630        640        650        660
 VNCTEVPVAI HADQLTPTWR VYSTGSNVFQ TRAGCLIGAE HVNNSYECDI PIGAGICASY
   670        680        690        700        710        720
QTQTNSHGSA SSVASQSIIA YTMSLGAENS VAYSNNSIAI PINFTISVTT EILPVSMTKT
   730        740        750        760        770        780
SVDCTMYICG DSTECSNLLL QYGSFCTQLN RALTGIAVEQ DKNTQEVFAQ VKQIYKTPPI
   790        800        810        820        830        840
KDFGGFNFSQ ILPDPSKPSK RSFIEDLLFN KVTLADAGFI KQYGDCLGDI AARDLICAQK
   850        860        870        880        890        900
FNGLTVLPPL LTDEMIAQYT SALLAGTITS GWTFGAGAAL QIPFAMQMAY RFNGIGVTQN
   910        920        930        940        950        960
VLYENQKLIA NQFNSAIGKI QDSLSSTASA LGKLQDVVNQ NAQALNTLVK QLSSNFGAIS
   970        980        990       1000       1010       1020
SVLNDILARL DPPEAEVQID RLITGRLQSL QTYVTQQLIR AAEIRASANL AATKMSECVL
  1030       1040       1050       1060       1070       1080
GQSKRVDFCG KGYHLMSFPQ SAPHGVVFLH VTYVPAQEKN FTTAPAICHD GKAHFPREGV
  1090       1100       1110       1120       1130       1140
FVSNGTHWFV TQRNFYEPQI ITTHNTFVSG NCDVVIGIVN NTVYDPLQPE LDSFKEELDK
  1150       1160       1170       1180       1190       1200
YFKNHTSPDV DLGDISGINA SVVNIQKEID RLNEVAKNLN ESLIDLQELG KYEQAAALVP
  1210       1220       1230       1240       1250       1260
RGSPGSGYIP EAPRDGQAYV RKDGEWVLLS TFLVENLYFQ GGSAWSHPQF EKGGGSGGGS
                     1270       1280
                GGSAWSHPQF EKLEHHHHHH HH
```

hACE2-fc (M_w: 95.2 kDa)

```
        10         20         30         40         50         60
MQSTIEEQAK TFLDKFNHEA EDLFYQSSLA SWNYNTNITE ENVQNMNNAG DKWSAFLKEQ
    70         80         90        100        110        120
STLAQMYPLQ EIQNLTVKLQ LQALQQNGSS VLSEDKSKRL NTILNTMSTI YSTGKVCNPD
   130        140        150        160        170        180
NPQECLLLEP GLNEIMANSL DYNERLWAWE SWRSEVGKQL RPLYEEYVVL KNEMARANHY
   190        200        210        220        230        240
EDYGDYWRGD YEVNGVDGYD YSRGQLIEDV EHTFEEIKPL YEHLHAYVRA KLMNAYPSYI
   250        260        270        280        290        300
SPIGCLPAHL LGDMWGRFWT NLYSLTVPFG QKPNIDVTDA MVDQAWDAQR IFKEAEKFFV
   310        320        330        340        350        360
SVGLPNMTQG FWENSMLTDP GNVQKAVCHP TAWDLGKGDF RILMCTKVTM DDFLTAHHEM
   370        380        390        400        410        420
GHIQYDMAYA AQPFLLRNGA NEGFHEAVGE IMSLSAATPK HLKSIGLLSP DFQEDNETEI
   430        440        450        460        470        480
NFLLKQALTI VGTLPFTYML EKWRWMVFKG EIPKDQWMKK WWEMKREIVG VVEPVPHDET
   490        500        510        520        530        540
YCDPASLFHV SNDYSFIRYY TRTLYQFQFQ EALCQAAKHE GPLHKCDISN STEAGQKLFN
   550        560        570        580        590        600
MLRLGKSEPW TLALENVVGA KNMNVRPLLN YFEPLFTWLK DQNKNSFVGW STDWSPYADA
   610        620        630        640        650        660
AASDKTHTCP PCPAPELLGG PSVFLFPPKP KDTLMISRTP EVTCVVVDVS HEDPEVKFNW
   670        680        690        700        710        720
YVDGVEVHNA KTKPREEQYN STYRVVSVLT VLHQDWLNGK EYKCKVSNKA LPAPIEKTIS
   730        740        750        760        770        780
KAKGQPREPQ VYTLPPSRDE LTKNQVSLTC LVKGFYPSDI AVEWESNGQP ENNYKTTPPV
           790        800        810        820        830
    LDSDGSFFLY SKLTVDKSRW QQGNVFSCSV MHEALHNHYT QKSLSLSPGK
```

Figure A1. Single letter amino acid sequence of the SARS-CoV spike protein and hACE2 constructs used for SPR experiments. T4 fold-on sequences are marked in yellow, Poly-His sequences are marked in green and fc-tags are marked in turquoise.

References

1. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **2020**, *395*, 497–506. [CrossRef]
2. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K.S.; Lau, E.H.; Wong, J.Y.; et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* **2020**, *382*, 1199–1207. [CrossRef] [PubMed]
3. Rabaan, A.A.; Al-Ahmed, S.H.; Haque, S.; Sah, R.; Tiwari, R.; Malik, Y.S.; Dhama, K.; Yatoo, M.I.; Bonilla-Aldana, D.K.; Rodriguez-Morales, A.J. SARS-CoV-2, SARS-CoV, and MERS-COV: A comparative overview. *Infez Med.* **2020**, *28*, 174–184.
4. Lee, N.; Hui, D.; Wu, A.; Chan, P.; Cameron, P.; Joynt, G.M.; Ahuja, A.; Yung, M.Y.; Leung, C.; To, K.; et al. A major outbreak of severe acute respiratory syndrome in Hong Kong. *N. Engl. J. Med.* **2003**, *348*, 1986–1994. [CrossRef] [PubMed]
5. Hänel, K.; Willbold, D. SARS-CoV accessory protein 7a directly interacts with human LFA-1. *Biol. Chem.* **2007**, *388*, 1325–1332. [CrossRef] [PubMed]
6. WHO Coronavirus (COVID-19) Dashboard. Available online: https://covid19.who.int/ (accessed on 7 February 2022).
7. Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.Y.; et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, *181*, 894–904.e899. [CrossRef] [PubMed]
8. Koch, J.; Uckeley, Z.M.; Doldan, P.; Stanifer, M.; Boulant, S.; Lozach, P.Y. TMPRSS2 expression dictates the entry route used by SARS-CoV-2 to infect host cells. *Embo J.* **2021**, *40*, e107821. [CrossRef]
9. Chen, J.; Wang, R.; Wang, M.; Wei, G.W. Mutations Strengthened SARS-CoV-2 Infectivity. *J. Mol. Biol.* **2020**, *432*, 5212–5226. [CrossRef]
10. Choe, H.; Farzan, M. How SARS-CoV-2 first adapted in humans. *Science* **2021**, *372*, 466–467. [CrossRef]
11. Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221–224. [CrossRef]
12. Wan, Y.S.; Shang, J.; Graham, R.; Baric, R.S.; Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: An Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol.* **2020**, *94*, e00127-20. [CrossRef] [PubMed]
13. Letko, M.; Marzi, A.; Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.* **2020**, *5*, 562–569. [CrossRef]
14. Tang, J.W.; Tambyah, P.A.; Hui, D.S. Emergence of a new SARS-CoV-2 variant in the UK. *J. Infect.* **2021**, *82*, e27–e28. [CrossRef] [PubMed]
15. Guruprasad, L. Human SARS CoV-2 spike protein mutations. *Proteins* **2021**, *89*, 569–576. [CrossRef] [PubMed]
16. Huang, Y.; Yang, C.; Xu, X.F.; Xu, W.; Liu, S.W. Structural and functional properties of SARS-CoV-2 spike protein: Potential antivirus drug development for COVID-19. *Acta Pharmacol. Sin.* **2020**, *41*, 1141–1149. [CrossRef]
17. Yi, C.; Sun, X.; Ye, J.; Ding, L.; Liu, M.; Yang, Z.; Lu, X.; Zhang, Y.; Ma, L.; Gu, W.; et al. Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies. *Cell Mol. Immunol.* **2020**, *17*, 621–630. [CrossRef]
18. Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581*, 215–220. [CrossRef]
19. Harvey, W.T.; Carabelli, A.M.; Jackson, B.; Gupta, R.K.; Thomson, E.C.; Harrison, E.M.; Ludden, C.; Reeve, R.; Rambaut, A.; Peacock, S.J.; et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **2021**, *19*, 409–424. [CrossRef]
20. Volz, E.; Mishra, S.; Chand, M.; Barrett, J.C.; Johnson, R.; Geidelberg, L.; Hinsley, W.R.; Laydon, D.J.; Dabrera, G.; O'Toole, Á.; et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* **2021**, *593*, 266–269. [CrossRef]
21. Mlcochova, P.; Kemp, S.A.; Dhar, M.S.; Papa, G.; Meng, B.; Ferreira, I.A.T.M.; Datir, R.; Collier, D.A.; Albecka, A.; Singh, S.; et al. SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature* **2021**, *599*, 114–119. [CrossRef]
22. Karim, S.S.A.; Karim, Q.A. Omicron SARS-CoV-2 variant: A new chapter in the COVID-19 pandemic. *Lancet* **2021**, *398*, 2126–2128. [CrossRef]
23. Tagliamonte, M.S.; Mavian, C.; Zainabadi, K.; Cash, M.N.; Lednicky, J.A.; Magalis, B.R.; Riva, A.; Deschamps, M.M.; Liautaud, B.; Rouzier, V.; et al. Rapid emergence and spread of SARS-CoV-2 gamma (P.1) variant in Haiti. *Clin. Infect. Dis.* **2021**. [CrossRef] [PubMed]
24. Arora, P.; Pohlmann, S.; Hoffmann, M. Mutation D614G increases SARS-CoV-2 transmission. *Signal Transduct. Target. Ther.* **2021**, *6*, 101. [CrossRef]
25. Thomson, E.C.; Rosen, L.E.; Shepherd, J.G.; Spreafico, R.; da Silva Filipe, A.; Wojcechowskyj, J.A.; Davis, C.; Piccoli, L.; Pascall, D.J.; Dillen, J.; et al. Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell* **2021**, *184*, 1171–1187. [CrossRef] [PubMed]
26. Couch, R.B. Seasonal inactivated influenza virus vaccines. *Vaccine* **2008**, *26* (Suppl. 4), D5–D9. [CrossRef]
27. Bolarin, J.A.; Oluwatoyosi, M.A.; Orege, J.I.; Ayeni, E.A.; Ibrahim, Y.A.; Adeyemi, S.B.; Tiamiyu, B.B.; Gbadegesin, L.A.; Akinyemi, T.O.; Odoh, C.K.; et al. Therapeutic drugs for SARS-CoV-2 treatment: Current state and perspective. *Int. Immunopharmacol.* **2021**, *90*, 107228. [CrossRef]
28. Mohapatra, S.; Nath, P.; Chatterjee, M.; Das, N.; Kalita, D.; Roy, P.; Satapathi, S. Repurposing therapeutics for COVID-19: Rapid prediction of commercially available drugs through machine learning and docking. *PLoS ONE* **2020**, *15*, e0241543. [CrossRef]

29. Eberle, R.J.; Olivier, D.S.; Amaral, M.S.; Gering, I.; Willbold, D.; Arni, R.K.; Coronado, M.A. The Repurposed Drugs Suramin and Quinacrine Cooperatively Inhibit SARS-CoV-2 3CL(pro) In Vitro. *Viruses* **2021**, *13*, 873. [CrossRef]
30. Loschwitz, J.; Jäckering, A.; Keutmann, M.; Olagunju, M.; Eberle, R.J.; Coronado, M.A.; Olubiyi, O.O.; Strodel, B. Novel inhibitors of the main protease enzyme of SARS-CoV-2 identified via molecular dynamics simulation-guided in vitro assay. *Bioorg. Chem.* **2021**, *111*, 104862. [CrossRef]
31. Gordon, C.J.; Tchesnokov, E.P.; Schinazi, R.F.; Gotte, M. Molnupiravir promotes SARS-CoV-2 mutagenesis via the RNA template. *J. Biol. Chem.* **2021**, *297*, 100770. [CrossRef]
32. Mahase, E. Covid-19: Pfizer's paxlovid is 89% effective in patients at risk of serious illness, company reports. *BMJ* **2021**, *375*, n2713. [CrossRef] [PubMed]
33. Ibrahim, M.A.; Abdelrahman, A.H.; Hussien, T.A.; Badr, E.A.; Mohamed, T.A.; El-Seedi, H.R.; Pare, P.W.; Efferth, T.; Hegazy, M.E.F. In silico drug discovery of major metabolites from spices as SARS-CoV-2 main protease inhibitors. *Comput. Biol. Med.* **2020**, *126*, 104046. [CrossRef] [PubMed]
34. Barderas, R.; Benito-Pena, E. The 2018 Nobel Prize in Chemistry: Phage display of peptides and antibodies. *Anal. Bioanal. Chem.* **2019**, *411*, 2475–2479. [CrossRef] [PubMed]
35. Cronan, J.E., Jr. Biotination of proteins in vivo. A post-translational modification to label, purify, and study proteins. *J. Biol. Chem.* **1990**, *265*, 10327–10333. [CrossRef]
36. Santur, K.; Reinartz, E.; Lien, Y.; Tusche, M.; Altendorf, T.; Sevenich, M.; Tamgüney, G.; Mohrlüder, J.; Willbold, D. Ligand-Induced Stabilization of the Native Human Superoxide Dismutase 1. *ACS Chem. Neurosci.* **2021**, *12*, 2520–2528. [CrossRef]
37. Krejci, A.; Hupp, T.R.; Lexa, M.; Vojtesek, B.; Muller, P. Hammock: A hidden Markov model-based peptide clustering algorithm to identify protein-interaction consensus motifs in large datasets. *Bioinformatics* **2016**, *32*, 9–16. [CrossRef]
38. Schubert, M.; Nimtz, M.; Bertoglio, F.; Schmelz, S.; Lukat, P.; Heuvel, J.V.D. Reproducible and Easy Production of Mammalian Proteins by Transient Gene Expression in High Five Insect Cells. *Methods Mol. Biol.* **2021**, *2305*, 129–140.
39. Meyer, S.; Lorenz, C.; Başer, B.; Wördehoff, M.; Jäger, V.; Heuvel, J.V.D. Multi-Host Expression System for Recombinant Production of Challenging Proteins. *PLoS ONE* **2013**, *8*, e68674. [CrossRef]
40. Papanikolopoulou, K.; Forge, V.; Goeltz, P.; Mitraki, A. Formation of highly stable chimeric trimers by fusion of an adenovirus fiber shaft fragment with the foldon domain of bacteriophage T4 fibritin. *J. Biol. Chem.* **2004**, *279*, 8991–8998. [CrossRef]
41. Vranken, W.; Boucher, W.; Stevens, T.J.; Fogh, R.H.; Pajon, A.; Llinas, M.; Ulrich, E.L.; Markley, J.L.; Ionides, J.; Laue, E.D. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **2005**, *59*, 687–696. [CrossRef]
42. Sevenich, M.; van den Heuvel, J.; Gering, I.; Mohrlüder, J.; Willbold, D. Conversion rate to the secondary conformation state in the binding mode of SARS-CoV-2 spike protein to human ACE2 may predict infectivity efficacy of the underlying virus mutant. *bioRxiv* **2021**.
43. Ostrov, D.A. Structural Consequences of Variation in SARS-CoV-2 B.1.1.7. *J. Cell Immunol.* **2021**, *3*, 103–108. [PubMed]
44. Starr, T.N.; Greaney, A.J.; Hilton, S.K.; Ellis, D.; Crawford, K.H.; Dingens, A.S.; Navarro, M.J.; Bowen, J.E.; Tortorici, M.A.; Walls, A.C.; et al. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **2020**, *182*, 1295–1310. [CrossRef] [PubMed]
45. Khan, A.; Zia, T.; Suleman, M.; Khan, T.; Ali, S.S.; Abbasi, A.A.; Mohammad, A.; Wei, D.Q. Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data. *J. Cell Physiol.* **2021**, *236*, 7045–7057. [CrossRef] [PubMed]
46. Zhang, L.; Jackson, C.B.; Mou, H.; Ojha, A.; Peng, H.; Quinlan, B.D.; Rangarajan, E.S.; Pan, A.; Vanderheiden, A.; Suthar, M.S.; et al. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat. Commun.* **2020**, *11*, 6031. [CrossRef] [PubMed]
47. Brielle, E.S.; Schneidman-Duhovny, D.; Linial, M. The SARS-CoV-2 Exerts a Distinctive Strategy for Interacting with the ACE2 Human Receptor. *Viruses* **2020**, *12*, 497. [CrossRef] [PubMed]
48. Xiong, X.; Qu, K.; Ciazynska, K.A.; Hosmillo, M.; Carter, A.P.; Ebrahimi, S.; Ke, Z.; Scheres, S.H.; Bergamaschi, L.; Grice, G.L.; et al. A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat. Struct. Mol. Biol.* **2020**, *27*, 934–941. [CrossRef]
49. Vauquelin, G. Effects of target binding kinetics on in vivo drug efficacy: K(off), k(on) and rebinding. *Br. J. Pharmacol.* **2016**, *173*, 2319–2334. [CrossRef]
50. Larue, R.C.; Xing, E.; Kenney, A.D.; Zhang, Y.; Tuazon, J.A.; Li, J.; Yount, J.S.; Li, P.K.; Sharma, A. Rationally Designed ACE2-Derived Peptides Inhibit SARS-CoV-2. *Bioconjugate Chem.* **2021**, *32*, 215–223. [CrossRef]
51. Panda, S.K.; Gupta, P.S.S.; Biswal, S.; Ray, A.K.; Rana, M.K. ACE-2-Derived Biomimetic Peptides for the Inhibition of Spike Protein of SARS-CoV-2. *J. Proteome Res.* **2021**, *20*, 1296–1303. [CrossRef]
52. Rai, J. Peptide and protein mimetics by retro and retroinverso analogs. *Chem. Biol. Drug Des.* **2019**, *93*, 724–736. [CrossRef] [PubMed]
53. Elfgen, A.; Hupert, M.; Bochinsky, K.; Tusche, M.; González de San Román Martin, E.; Gering, I.; Sacchi, S.; Pollegioni, L.; Huesgen, P.F.; Hartmann, R.; et al. Metabolic resistance of the D-peptide RD2 developed for direct elimination of amyloid-beta oligomers. *Sci. Rep.* **2019**, *9*, 5715. [CrossRef] [PubMed]