# Discovery of all-D-peptide inhibitors of SARS-CoV-2 3C-like protease

Raphael J. Eberle[1,2], Marc Sevenich[1,2,3], Ian Gering[1], Lara Scharbert[1,2], Birgit Strodel[1,2], Karoline B. Santur[1,2], Jeannine Mohrlüder[1], Mônika A. Coronado[1*] and Dieter Willbold[1,2*]

[1]Institute of Biological Information Processing: Structural Biochemistry (IBI-7), Forschungszentrum Jülich, 52425 Jülich, Germany

[2]Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf

[3]Priavoid GmbH, Düsseldorf, Germany

[*]Corresponding authors: Tel: +49-2461-619505, m.coronado@fz-juelich.de;
Tel.: +49-2461-612100, Fax: +49-2461-612023, d.willbold@fz-juelich.de

**Keywords:** COVID-19, SARS-CoV-2 3CL protease, 3CL$^{pro}$ inhibitors, D-retro-inverso peptides

## *Abstract*

During the replication process of SARS-CoV-2 the main protease of the virus (3-chymotrypsin-like protease (3CL$^{pro}$)) plays a pivotal role and is essential for the life cycle of the pathogen. Numerous studies have been conducted so far, which have confirmed 3CL$^{pro}$ as an attractive drug target to combat COVID-19. We describe a novel and efficient next generation sequencing (NGS) supported phage display selection strategy for the identification of a set of SARS-CoV-2 3CL$^{pro}$ targeting peptide ligands that inhibit the 3CL protease, in a competitive or non-competetive mode, in the low µM range. From the most efficient L-peptides obtained from the phage display, we designed all-D-peptides based on the retro-inverso (ri) principle. They had $IC_{50}$ values also in the low µM range, and in combination even in the sub-micromolar range. The inhibition modes of these D-ri peptides were the same as their respective L-peptide versions. Our results demonstrate that retro-inverso obtained all-D-peptides interact with high-affinity and inhibit the SARS-CoV-2 3CL protease, thus reinforcing their potential as therapeutic agents. The here described D-ri peptides address limitations associated with current L-peptide inhibitors and are promising lead compounds. Further optimization regarding pharmacokinetic properties will allow the development of even more potent D-peptides to be used for the prevention and treatment of COVID-19.

*Introduction*

The spread of Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), with increasing levels of infectivity and transmissibility, has strained human health and public safety worldwide.[1] By May 2022, the COVID-19 pandemic has resulted in more than 524 million confirmed cases and more than 6.2 million confirmed deaths, according to the World Health Organization (WHO).[2] To date, vaccination is considered the key strategy for ending the pandemic[3] and the worldwide vaccination campaign using clinical safe and efficient vaccines against SARS-CoV-2 have controlled the number of death, however, not yet the spread of the diseases.[4-7] So far, more than 11.8 billion vaccine doses have been administered.[2] Remdesivir, Dexamethasone, Favipiravir, Lopinavir/Ritonavir, Nirmatrelvir/Ritonavir (Main protease inhibitor) and Darunavir have been approved for emergency use to inhibit SARS-CoV-2 infection and replication.[8-12] Given a considerable limitation of direct-acting antivirals for COVID-19 and an increasing presence of SARS-CoV-2 variants (B.1.1.529, B.1.617.2, B.1.1.7, B.1.351, A.23.1, B.1.525, B.1.526 and P.1),[13] it remains a strategic priority to develop new drug candidates with minimal side effects and also targeting the new variants.

For the expression and replication of the CoV gene, proteases play essential roles involving the proteolytic processing of replicase polyproteins, which makes them attractive targets for drug development.[14,15] The main protease, also called 3CL protease ($3CL^{pro}$), is the 3-chymotrypsin-like cysteine protease that features a His41-Cys145 catalytic dyad. A glutamine (Gln) residue is commonly located at the P1 site on the protease substrate. No known human cysteine protease cleaves after Gln, thus offering potential selectivity for this viral target over human proteases.[16,17] Viral proteases have been attractive targets for oral small-molecule therapies in treating HIV and HCV.[18,19] The use of SARS-CoV-2 $3CL^{pro}$ as a drug target has several advantages towards other virus proteins: (i) its essential role in the viral replication process, (ii) its potential for

mechanistic safety and (iii) the expected lack of variant resistance challenges.[8] 3CL$^{pro}$ inhibition represents an attractive approach for a safe and orally available antiviral therapy to treat COVID-19.

Drugs are conventionally classified into two molecule weight classes: "small molecule" drugs (< 500 Da) and protein/peptide-based drugs (> 1000 Da).[20] Besides the obvious advantages of small molecules like favourable oral bioavailability and a rational design,[21] they often have low target selectivity, ultimately resulting in side effects. In comparison, peptides were often neglected as potential molecules for drug development, despite binding to their target proteins with high affinity and specificity. They are smaller than proteins and can be obtained synthetically by well-established and cost-efficient methods.[22] The disadvantages of peptide drugs are their potential immunogenicity and also low bioavailability due to degradation and short half-lives. Several approaches have been developed to approach these problems. D-enantiomeric amino acid residues effectively enhance the resistance to degradation since most proteolytic enzymes have substrate specificity for L-peptide bonds. This even allows the oral administration of D-enantiomeric peptides.[23,24] Furthermore, all-D-peptides are less immunogenic.[25]

Several studies described the development of the first all-D-peptide designed to treat Alzheimer's disease and showed it to be efficient in animal models and safe in humans.[26-28] In the case of SARS-CoV-2, three studies describe the identification of D-peptide inhibitors against SARS-CoV-2 Spike protein and 3CL$^{pro}$.[29-31] Of particular interest are all-D-peptides that are composed of D-amino acid residues. However, many peptide-based drug development strategies, like phage display selection, end up with L-peptides. So called D-retro-inverso (D-ri) peptides are composed solely of D-amino acid residues in the reverse sequence of their parental L-peptide. Such a peptide assembled in reversed sequence from D-amino acid residues will have almost the same structure, stability, and bioactivity as its parent peptide made of L-amino acid residues, but with the advantage of being more resistant to proteolytic degradation. This combination makes D-ri peptides attractive drug candidates.[32]

In the present paper, we explore the inhibition potential of L-peptides selected using phage display targeting SARS-CoV-2 $3CL^{pro}$. Subsequently, its retro-inverso versions were explored, also in combination.

## *Results*

### *Characterisation of SARS-CoV-2 $3CL^{pro}$ inhibition by L- and D-retro-inverso peptides*

All experiments were performed with recombinant SARS-CoV-2 $3CL^{pro}$_GST fusion protein (Supplementary Fig. S1). To select binding L-peptides to the $3CL^{pro}$, M13 phage display selection was performed (Fig.1A). Detailed information about the phage display process are described in supplementary text S1, Figs S1 and S2, Table S1.

To verify whether the selected peptides showed inhibition toward the target protein, they were primary tested using a final concentration of 10 μM of each peptide, separately, in an enzyme-based fluorescence assay (Fig. 1B)

With more than 60% inhibition, the peptides 3CVL-2 (SPHGWPSQSIEVQPQW), 3CVL-4 (AHEGWTWDWTPQYSWK) and 3CVL-7 (TVAPLHAHYWDVEERH) were selected for further analysis (Fig. 1C, D, E). Concerning their potential to inhibit 100% of the protease activity, determination of the $IC_{50}$ was performed using 0-140 μM (3CVL-2), 0-120 μM (3CVL-4), and 0-100 (3CVL-7).

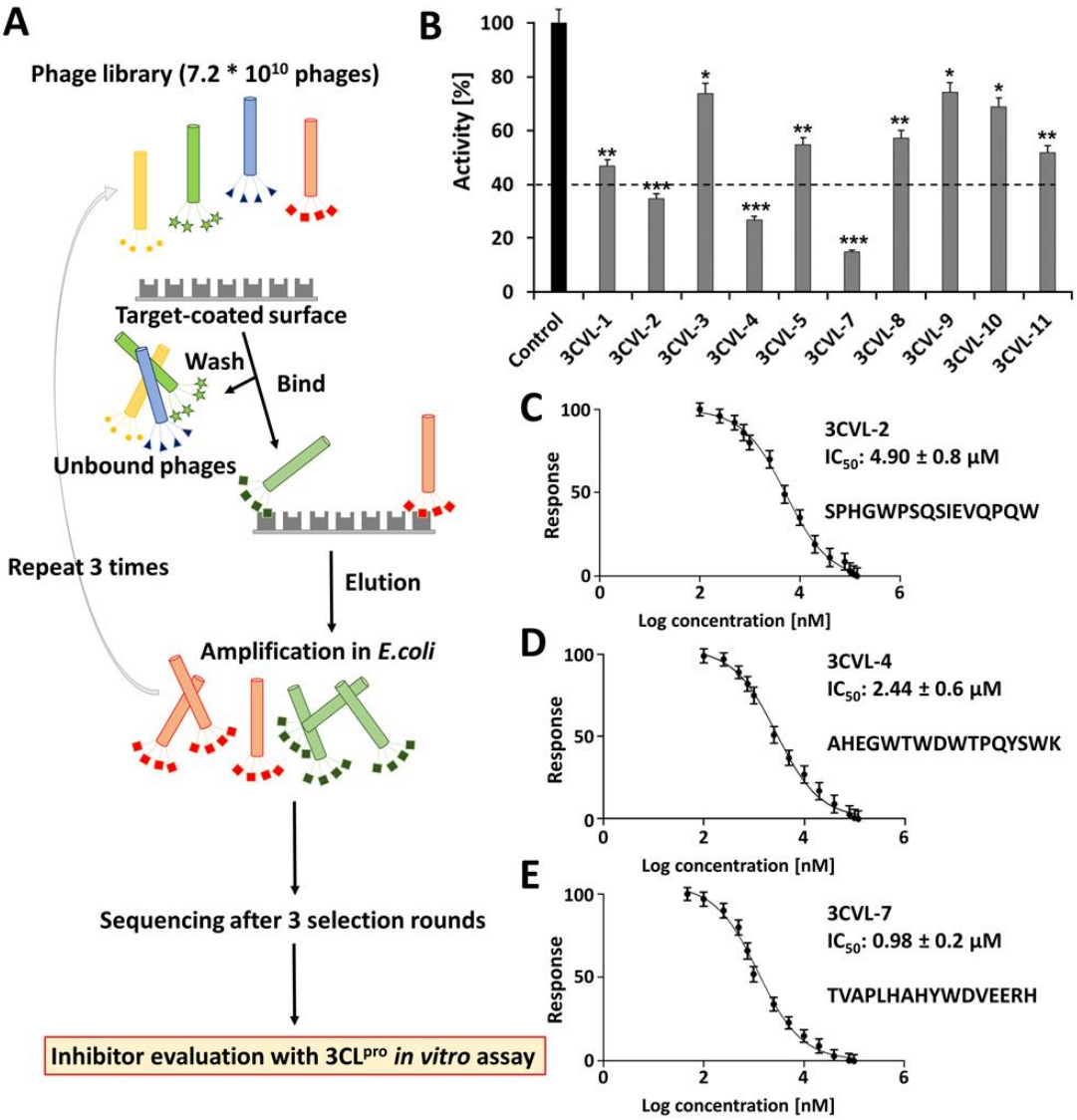


**Figure 1. Phage display Scheme and screening of phage display L-peptides against SARS-CoV-2 3CL$^{pro}$ activity. (A)** Schematic diagram of the phage display process. **(B)** The primary screen of the eleven selected peptides (10 µM) against SARS-CoV-2 3CL$^{pro}$ activity. 3CVL-2, 3CVL-4 and 3CVL-7 inhibit the virus protease activity by more than 60%. **(C–E)** Selected peptide inhibiting SARS-CoV-2 3L$^{pro}$. Dose-response curves for $IC_{50}$ values of 3CVL-2 (C), 3CVL-4 (D) and 3CVL-7 (E) were determined by nonlinear regression. Data shown are the mean ± SD from three independent measurements (n=3). Asterisks mean that the data differs from the control (0 µM inhibitor) significantly at $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (***), level according to ANOVA and Tukey's test.

The 3CVL-2 at a final concentration of 140 µM showed 100% inhibition of the protease (Supplementary Fig. S4A) with a calculated $IC_{50}$ value of 4.90 ± 0.8 µM (Table 1, Fig. 1C). In a similar concentration range, the tested 3CVL-4 inhibited 100% of the protease

activity at a final concentration of 120 μM (Supplementary Fig. S4C), presenting an $IC_{50}$ value of 2.44 ± 0.6 μM (Table 1, Fig. 1D). The selected peptide 3CVL-7 performs even better and inhibits 100% of the recombinant SARS-CoV-2 protease activity at a final concentration of 80 μM (Supplementary Fig. S4E) with an $IC_{50}$ value of 0.98 ± 0.2 μM (Table 1, Fig.1E).

Further inhibition mode assays identified 3CVL-2 and 3CVL-7 as competitive inhibitors. These peptides compete with the substrate for the activity and/or binding sites of the protease (Supplementary Figs. S4B and S4F). Remarkably, the inhibition mode has shown that $3CL^{pro}$ can be allosterically inhibited by 3CVL-4, which discloses a non-competitive inhibition mode (Supplementary Fig. S4D).

**Table 1.** Summary of the SARS-CoV-2 $3CL^{pro}$ inhibition experiments by 3CVL-2, -4 and -7.

| **Peptide** | **$IC_{50}$ [μM]** | **Inhibition mode** |
|---|---|---|
| 3CVL-2 | 4.90 ± 0.8 | competitive |
| 3CVL-4 | 2.44 ± 0.6 | non-competitive |
| 3CVL-7 | 0.98 ± 0.2 | competitive |

To improve the poor proteolytic stability of L-enantiomeric peptides without losing affinity and binding specificity to their target protein, the peptides 3CVL-2, 3CVL-4, and 3CVL-7 were synthesized in retro-inverse D-enantiomeric (D-ri) form and designated as 3CVLri-2, 3CVLri-4, and 3CVLri-7, respectively. Glycine does not have a chiral center, that was the reason that the D-ri peptides have identical amino acid sequences, except for glycine, which was replaced by alanine. Additionally, an arginine was added to the C-termini of each D-ri peptide to increase the permeability of the peptides. The representative scheme is shown in Figs. 2A and B.

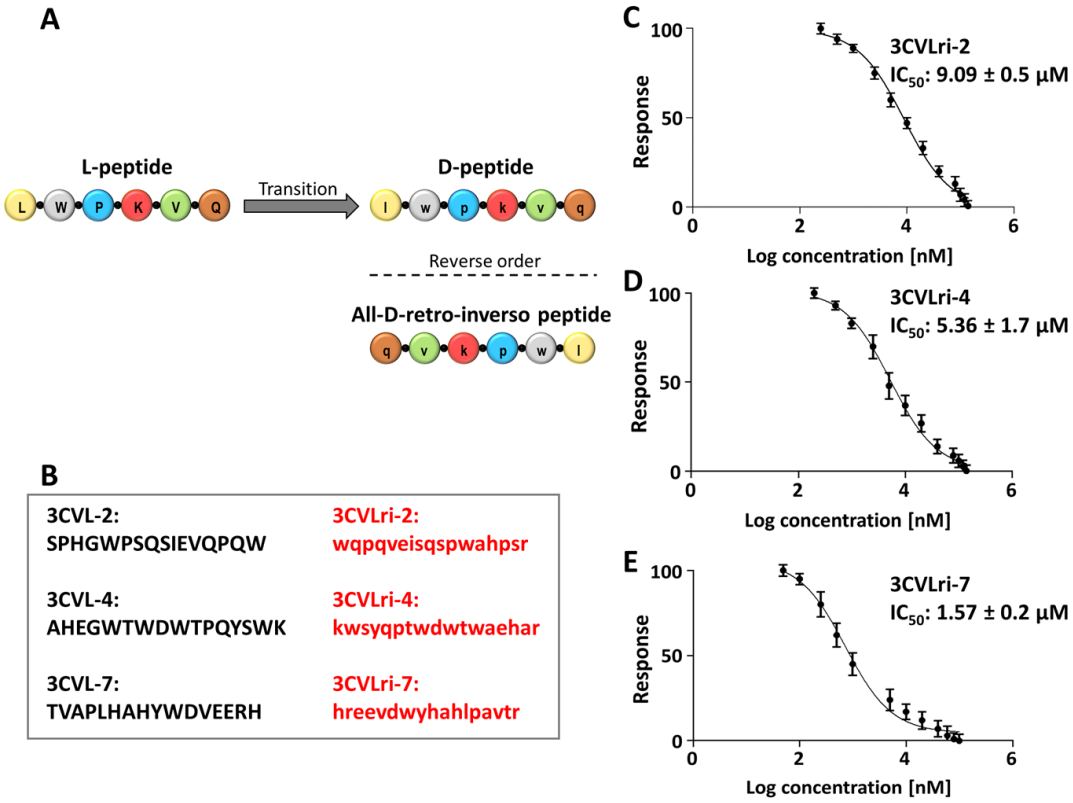


**Figure 2. Principle for obtaining D-retro-inverso peptides and inhibitory effect against SARS-CoV-2 3CL$^{pro}$ activity. (A)** Schematic diagram of D-retro-inverso (D-ri) peptides. **(B)** Sequences of 3CVLri-2, -4, -7 and their mother L-peptides. In the D-ri peptides, glycine (G) was replaced by alanine (a). **(C-E)** All-D-ri peptides inhibit SARS-CoV-2 3CL$^{pro}$ activity. Dose-response curves for $IC_{50}$ values of 3CVLri-2 (C), 3CVLri-4 (D) and 3CVLri-7 (E) were determined by nonlinear regression. Data shown are the mean ± SD from three independent measurements (n=3).

Like the L-peptides, the D-ri-peptides were tested using a concentration range of up to 140 µM (3CVLri-2 and 3CVLri-4) and up to 100 µM (3CVLri-7), which showed 100% inhibition of the protease by the corresponding used final concentration (Supplementary Fig. S5A, C, E). Based on the enzymatic assay, the $IC_{50}$ values was determined for each D-ri peptide, displaying for 3CVLri-2 (9.09 ± 0.5 µM), 3CVLri-4 (5.36 ± 1.7 µM) and 3CVLri-7 (1.57 ± 0.2 µM) an increase of the $IC_{50}$ values of about twofold compared to the L-peptides (Table 2). The mode-of-inhibition of the D-ri peptides was maintained, displaying a competitive mode of 3CLVri-2 and -7 and an allosteric mode of inhibition for 3CVLri-7 (Supplementary Figs. S5B, D and F).

**Table 2.** Summary of the SARS-CoV-2 3CL$^{pro}$ inhibition experiments by 3CVLri-2, -4 and -7.

| Peptide | $IC_{50}$ [μM] | Inhibition mode |
|---|---|---|
| 3CVLri-2 | 9.09 ± 0.5 | competitive |
| 3CVLri-4 | 5.36 ± 1.7 | non-competitive |
| 3CVLri-7 | 1.57 ± 0.2 | competitive |

***Competitive and non-competitive D-retro-inverso peptides act cooperatively to inhibit SARS-CoV-2 3CL$^{pro}$***

To presume the effect of the competitive (3CVLri-2 and 3CVLri-7) and non-competitive (3CVLri-4) D-retro-inverso peptides, a combined inhibitory assay was performed (3VLri-2+3CVLri-4 and 3CVL-7+3CVL-4). Both peptides were mixed in a 1:1 volume ratio and tested against the protease activity (Fig. 3).

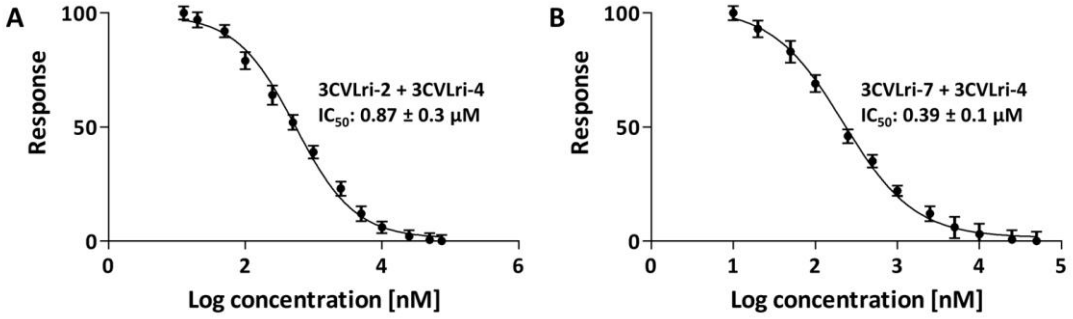


**Figure 3. SARS-CoV-2 3CL$^{pro}$ inhibition by a combination of competitive and non-competitive D-retro-inverso peptides simultaneously.** Dose-response curves for $IC_{50}$ values of the combination were determined by nonlinear regression. Data shown are the mean ± SD from three independent measurements (n=3). The corresponding normalised activity and inhibition plots of SARS-CoV-2 3CL$^{pro}$ under the influence of the combined peptides are shown in Supplementary Fig. S6. **(A)** Dose-response curve of the 3CVLri-2 + 3CVLri-4 combination. **(B)** Dose-response curve of the 3CVLri-7 + 3CVLri-4 combination.

The calculated $IC_{50}$ values for the 3CVLri-2 + 3CVLri-4 combination was 0.87 ± 0.3 μM (Fig. 3A) and for the 3CVLri-7 + 3CVLri-4 combination was 0.39 ± 0.1 μM (Fig. 3B).

***Binding affinities of selected 3CL$^{pro}$ L- and D-peptide inhibitors***

The interaction kinetics of 3CVL-2, -4 and -7 with 3CL$^{pro}$ and its D-ri form was determined using surface plasmon resonance (SPR) experiments. 3CL$^{pro}$ was immobilised via covalent primary amino group coupling, and 3CVL and 3CVLri peptides were injected as analytes. The Supplementary Figs. S7 and S8 show the SPR sensorgrams for 3CVL peptides at assay concentrations of 50, 25, 12.5, 6.25, 3.12, 1.56, 0.78, and 0.39 μM on 3CL$^{pro}$ immobilised on the CM5 sensor chip. The fit for the competitive inhibitors was made using a heterogeneous ligand model in which the two binding sites on the dimer are assumed to behave differently, as described for the 3CL$^{pro}$ dimer active sites.[33-35] The fit for the non-competitive inhibitor was made using a 1:1 stoichiometric kinetic fitting.

Heterogenous ligand fitting revealed dissociation constants from the low μM to low nM range for the higher affine binding site. Global 1:1 fitting showed dissociation constants in the low μM range (Supplementary Figs. S7, S8 and Table 3).

**Table 3.** $K_D$ values determined by SPR experiments.

| **Heterogenous ligand fitting** | | |
|---|---|---|
| **Peptide** | **$K_{D1}$** | **$K_{D2}$** |
| 3CVL-2 | 0.5 nM | 6.5 μM |
| 3CVL-7 | 1.8 μM | 7.8 μM |
| 3CVLri-2 | 46.8 μM | 1.7 μM |
| 3CVLri-7 | 2.4 nM | 35 μM |
| **1:1 ligand fitting** | | |
| **Peptide** | **$K_D$** | |
| 3CVL-4 | 5.1 μM | |
| 3CVLri-4 | 25.2 μM | |

The binding kinetics parameters of the association rate $k_a$, dissociation rate $k_d$, and dissociation constant $K_D$ ($k_d/k_a$) for the interaction of 3CL$^{pro}$ with peptides 3CVL-2, -4, -7, and 3CVLri-2, -4, -7 are shown in Supplementary Figs. S7 and S8.

*24 h stability and promiscuous assays of 3CVLri peptides*

Peptide-based inhibitors targeting active proteolytic proteases can lose their inhibitory effect over time. Based on our focus, the retro-inverso peptides were tested regarding their stability over 24 h. The results demonstrated a constant inhibition of SARS-CoV-2 $3CL^{pro}$ over time, showing that the peptides are not prone to the 3CL protease digestion (Supplementary Fig. S9).

Furthermore, a detergent-based control was carried out to exclude peptide inhibitors that possibly act as an aggregator of $3CL^{pro}$, so-called "promiscuous" inhibitors. The experiment was performed by adding 0.001%, 0.01% and 0.1% of Triton X-100 detergent to the reaction. Supposed that a molecule exhibits significant inhibition of $3CL^{pro}$, but this is diminished by detergent, it may act as an aggregation-based inhibitor. This was not observed for the tested 3CVLri peptides (Supplementary Fig. S10).

*Metabolic stability of 3CVLri peptides*

To investigate the 3CVLri-2, 3CVLri-4 and 3CVLri-7 peptides resistance against enzymatic degradation, the peptides were incubated in media simulating the gastrointestinal tract, blood and liver, and the unmetabolised peptides were quantified by RP-HPLC.

In simulated gastric fluid (SGF), the resistance of 3CVLri-2, 3CVLri-4 and 3CVLri-7 was monitored over 8 hours. The results showed that the 3CVLri-4 and -7 peptides are relatively stable (≥80 ±3%) within 8 hours. In contrast, 3CVLri-2 was remarkably metabolized (~ 75 ±2%) (Fig. 4A). The resistance of the D-ri peptides was also investigated using a simulated intestinal fluid (SIF). The peptides 3CVLri-2 and 3CVLri-4 remained rather stable over the selected time (≥90 ±1.5%), as well as 3CVLri-7 (≥85 ±5%) (Fig. 4B). In human plasma, the stability of 3CVLri-2 and 3CVLri-7 was monitored for 48 hours, with approximately 5% being metabolised (≥95 ±4.5%). In contrast, the 3CVLri-4 showed about 21 ± 5.6% of degradation within 48 hours (Fig. 4C). Furthermore, human liver microsomes were monitored for 24 hours, and 3CVLri-2 and 3CVLri-7 were about 5 ±1.1%

metabolized. However, 3CVLri-7 was almost 50% metabolised (≥45 ±3.7%) in the same period (Fig. 4D). The corresponding HPLC profiles of the stability tests are shown in Supplementary Figs. S11 to S14.

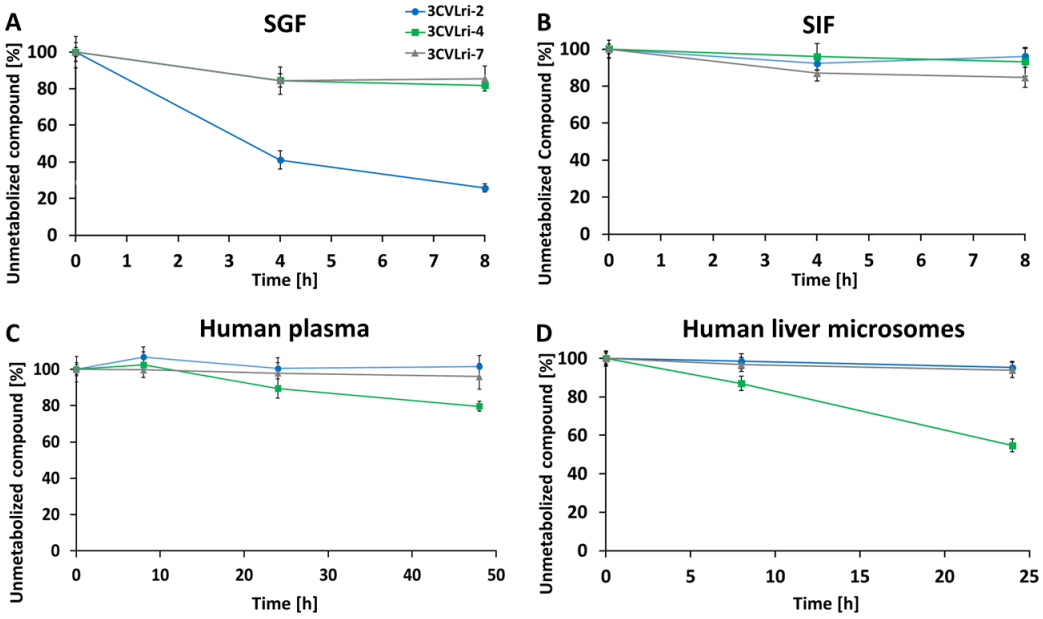


**Figure 4.** Stability of 3CVLri-2, 3CVLri-4 and 3CVLri-7 peptides incubated in SGF, SIF, human plasma and human liver microsomes. **(A)** SGF, **(B)** SIF, **(C)** human plasma and **(D)** human liver microsomes. Unmetabolised peptides were quantified by P-HPLC. The peak areas of the unmetabolised peptides after different incubation times were normalised to the peptides' peak areas after direct extraction from the media. Data are presented as mean ± SD (n = 3).

### *Structural properties of 3CVL and 3CVLri peptides and their similar interaction mode with SARS-CoV-2 3CL$^{pro}$*

Mode-of-inhibition studies indicated that 3CVL-2/3CVLri-2 and 3CVL-7/3CVLri-7 peptides follow a competitive inhibition mode, and 3CVL-4/3CVLri-4 peptides a non-competitive. Protease-peptide docking and subsequent molecular dynamics (MD) simulations were performed to investigate the possible binding interface between 3CL$^{pro}$-L and D-ri-peptides complexes.

The secondary structure of the 3D models of the L-peptides (3CVL-2, -4 and -7) and their D-ri counterparts corroborates with CD spectroscopy results. During the one μs MD simulations of the unbound peptides, they mainly contained a random coiled secondary structure. However, the L- and the D-ri-3CVL-4 peptides also more frequently formed an ◎-helix (Supplementary Fig. 15A-C).

The atomic coordinates of SARS-CoV-2 $3CL^{pro}$ (PDB entry: 6M2N) and ten representative structures per peptide from the peptide simulations were submitted for docking using AutoDock Vina.[36]

The competitive peptides 3CVL-2/3CVLri-2 and 3CVL-7/3CVLri-7 were docked into the active site and substrate-binding site of the protease (Supplementary Table S4). The non-competitive peptides 3CVL-4/3CVLri-4 were docked in a potential allosteric binding site (Supplementary Table S4). Previous studies have indicated that only one active site of the dimer is functional.[33,35] A 200 ns control MD simulation of the dimer without ligands was performed to identify a potentially less stable protomer or a collapsing of the binding sites. Based on the volume of the active sites, we did not identify any differences between both protomers (Supplementary Fig. S16). Therefore, we docked the competitive peptides into the active site of chain A. After docking, 100 ns MD simulations of selected $3CL^{pro}$/peptide complexes were conducted to further assess their stabilities. The selection of these complexes was made based on the calculated binding free energy and the distance to the critical catalytic dyad residues (Cys145 and His41). Per peptide, three different binding configurations were subjected to the MD simulations where as before $3CL^{pro}$ was simulated as a dimer. The flexibility of the peptides in complex with the protease was monitored by calculating the RMSD (root mean square deviation) and RMSF (root mean square fluctuations) (Supplementary Figs. S17 and S18). Representative structures for all three simulations of each L- and D-ri-peptide are shown in Supplementary Figs. S19 and S20. To investigate the $3CL^{pro}$ peptide interactions in more detail, we elongated one of the three MD simulations per peptide to 200 ns and conducted these additional simulations in triplicate. The selection of the complex per peptide for

extended simulations was based on the results regarding the RMSD, RMSF and the binding free energies during the initial 100 ns MD.

In order to identify the most popular binding modes sampled during the simulations, we applied conformational clustering to the 3 x 100 ns simulations but considered only the non-flexible/interacting residues (RMSF <= 5 Å) of the peptides (Supplementary Fig. S21). Interactions between the residues of the peptides and the protease were analyzed by calculating the minimum distance between them for the first three clusters per complex (Table 4).

**Table 4.** Summary of MD simulation triplicates considering non-flexible/interacting residues of peptides.

| Peptide | Binding Energy [kcal/mol] | #cluster | Representative frames first 3 clusters [%] |
|---|---|---|---|
| 3CVL-2 | 28 ± 52 | 8 | 97% |
| 3CVL-4 | 15 ± 28 | 12 | 88% |
| 3CVL-7 | 88 ± 42 | 25 | 80% |
| 3CVLri-2 | -125 ± 28 | 12 | 83% |
| 3CVLri-4 | -141 ± 38 | 10 | 99% |
| 3CVLri-7 | 91 ± 36 | 14 | 93% |

Analysis of the $3CL^{pro}$-peptide complexes provides valuable information on the interaction interface similarities between 3CVL and 3CVLri peptides. Structural representation (Fig. 5) demonstrated that 3CVL-2 extends along with S1', S1 and S3 sub pockets of the protease substrate-binding site (Fig. 5A). In contrast, the 3CVLri-2 interacts with the substrate binding site differently. However, it extends along with S1', S1, S2 and S3 (Fig. 5B), which shows one of the D-Gln residue of the 3CVLri-2 very well accommodated. The 3CVL-7 and 3CVLri-7 interaction areas comprise residues of the S1', S1, S2 and S3 subsites (Figs. 5C and D). The graphs in Fig. 5 that show the minimum distance between peptide and $3CL^{pro}$ residues reveal that all analysed peptides formed several interactions with $3CL^{pro}$ active/binding site residues. The localisation in the substrate-binding region of the 3CVL peptides and their D-ri counterparts, even when it

does not interact with the catalytic dyad (His41 and Cys145), showed the potential to block the entrance of the substrate.

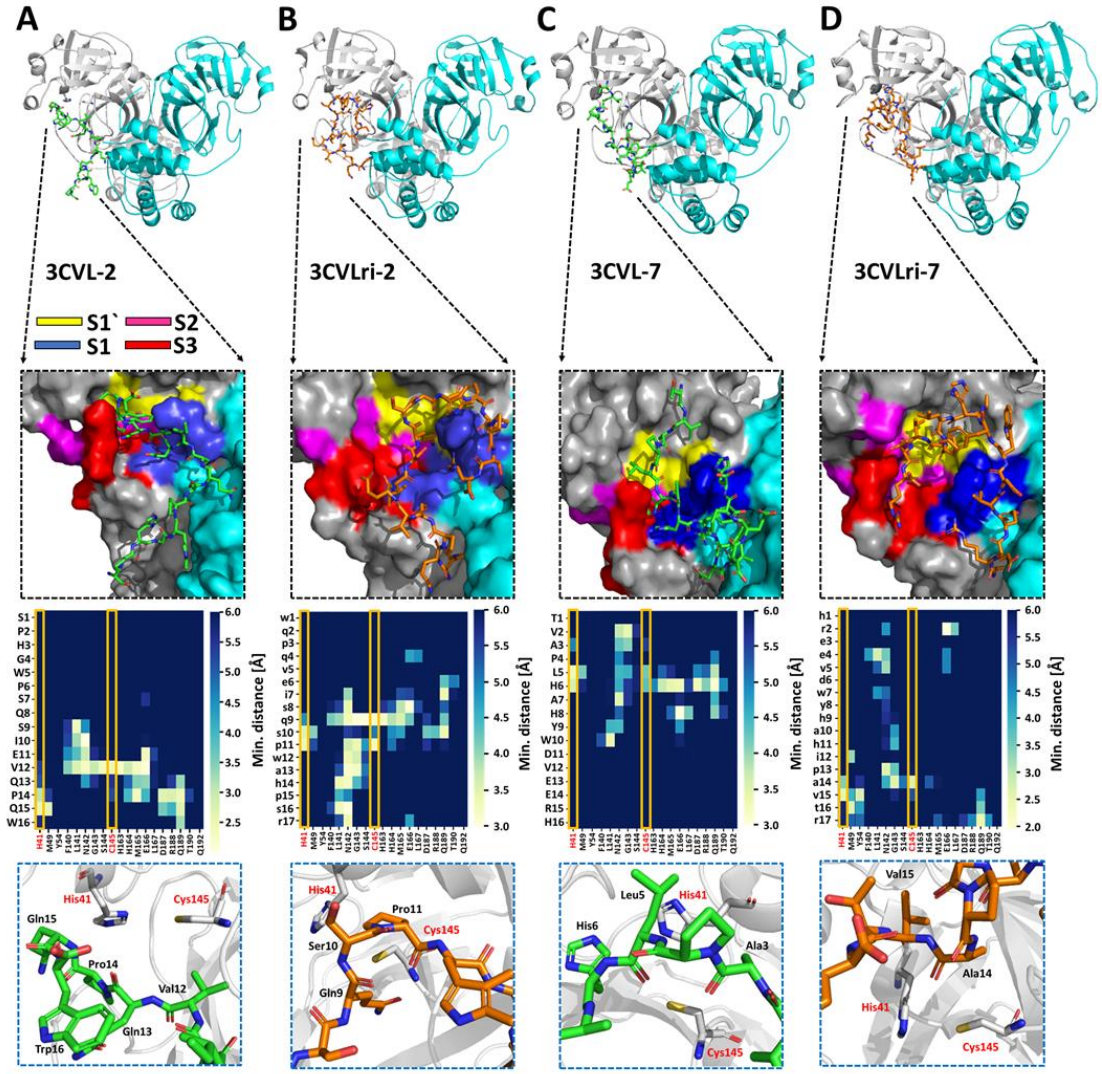


**Figure 5.** Comparison of the competitive 3CVL and 3CVLri peptide inhibitors binding position in the 3CVL$^{pro}$ active and substrate-binding site. 3CL$^{pro}$ is shown in ribbon view: chain A (grey) and chain B (turquoise). The 3CVL (green) and 3CVLri (orange) peptides are shown as sticks. The surface view demonstrates a preferred position of the peptide in the substrate binding region, and the single subsites are marked by different colours (S1': yellow; S1: blue; S2: pink; S3: red). The distance between the 3CL$^{pro}$ substrate binding site residues and the peptides show that active site residues His41 and Cys145 are blocked by the inhibitors. (**A)** 3CVL-2, **(B)** 3CVLri-2, **(C)** 3CVL-7 and **(D)** 3CVLri-7.

Contrary, the non-competitive inhibitors 3CVL-4 and 3CVLri-4 were docked in a suggested allosteric site at the interface of the protomers. A stable interaction of over 100 ns of MD simulation was observed. Structural representation of the interfaces of $3CL^{pro}$ in complex with the 3CVL-4 and 3CVLri-4 peptides are shown in Fig. 6. Analysis of the distance between residues of the allosteric site of each protomer demonstrated that several residues of both protomer chains interact with the protease in particular Lys12, 97, and 100 (Fig. 6).

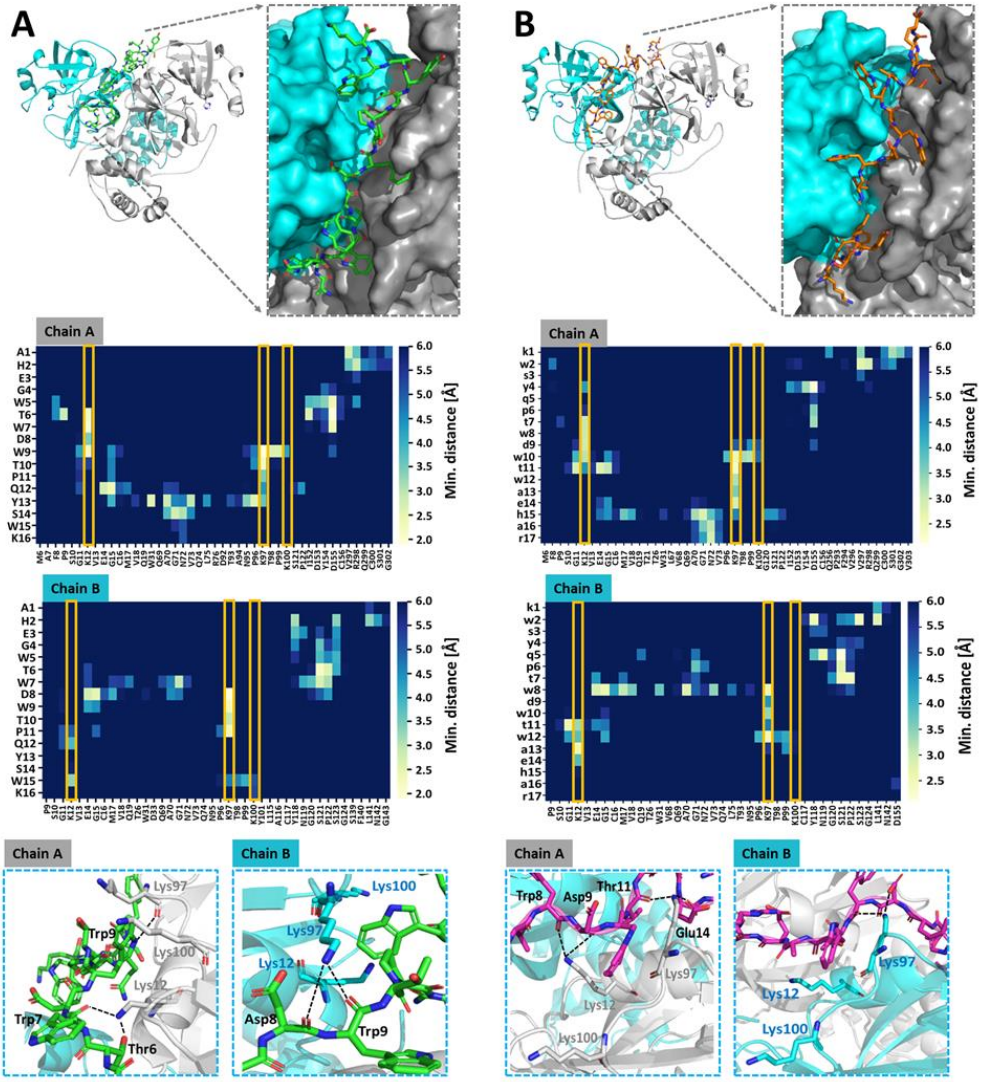


**Figure 6.** Comparison of the non-competitive 3CVL-4 and 3CVLri-4 peptide inhibitors in a supposed $3CL^{pro}$ allosteric site. $3CL^{pro}$ is shown in ribbon view and each protomer is coloured differently: chain A (grey) and chain B (turquoise). 3CVL (green) and 3CVLri (orange) peptides are shown as sticks. The surface view demonstrates a preferred position

of the peptide in the proposed allosteric site in a cleft between both protease protomers. (**A)** 3CVL-4 and **(B)** 3CVLri-4.

MD simulations of 3CVL-4 and 3CVLri-4 in a proposed allosteric site showed that this interaction induces conformational changes in the protease's active site. Determination of the volume of the active site with and without peptide in the allosteric site demonstrates a reduction of the active site overall volume (Supplementary Fig. S22).

***Discussion and conclusion***

Even upon the development of several vaccines against SARS-CoV-2, the development of suitable therapeutics against COVID-19 is important, since no causative therapy is known. Instead of inhibiting the interaction of the spike protein with the cell receptors,[37,38] we intend to target one of the most important proteases in virus replication, the 3CL protease, by D-enantiomeric peptide ligands preventing the cleavage function of $3CL^{pro}$. To this end, a phage display selection was performed.

The identified L-peptides showed a strong inhibitory effect in the enzyme activity assay, with $IC_{50}$ values varying between 0.9 and 9 μM. As we demonstrated experimentally, 3CVL-4 is a non-competitive inhibitor with an $IC_{50}$ value of 2.4 μM, which was maintained in its D-enantiomer retro-inverso form; however, with an increased $IC_{50}$ value of about twofold (5.4 μM). A 1:1 molar combination of the competitive and non-competitive D-ri-peptides reduced the $IC_{50}$ values remarkably. 3CVLri-2 + 3CVLri-4 combination reduces the $IC_{50}$ value to 0.9 μM and the combination of 3CVLri-7 + 3CVLri-4 to 0.4 μM. It was clearly demonstrated that a combination of competitive and non-competitive inhibitors increased the inhibitory potential of the all D-peptides.

The structural analysis of $3CL^{pro}$ in complex with the studied peptides by docking followed by molecular dynamics simulation predicted a stable conformation for each of the complexes, indicating significant intermolecular interaction. Of note, amino acid residues of the peptide, when not directly interacting with the catalytic dyad, are located close to the pocket entry, blocking the substrate entrance. The predicted structures of

3CL$^{pro}$ in complex with 3CVLri-2 indicates the occurrence of significant intermolecular interaction between the D-glutamine of the all-D-ri peptide and amino acid residues of the S1 subpocket, which explains the strong preference for D-glutamine at P1.[39]

A combination of competitive and non-competitive D-ri-peptides (3CVLri-2/3CVLri-4 and 3CVLri-7/3CVLri-4) increased the inhibitory effect against 3CL$^{pro}$.

The all-D-peptides 3CVLri-2 and -7 showed a remarkable resistance against metabolisation over 8 hours. Our study yielded all-D-peptides with a promising therapeutic potential against SARS-CoV-2 3CL$^{pro}$.

## *Materials & Methods*

### *Peptides*

All peptides were purchased from Caslo (Lyngby, Denmark) as a lyophilised powder with >98% purity. Peptides 3CVL-1 to 10 consist of 16 amino acid residues in L-configuration with amidated C-termini. The D-retro-inverso (D-ri) peptides 3CVLri-2, 3CVLri-4 and 3CVLri-7 have identical amino acid sequences, except for glycine, which was replaced by alanine in the case of 3CVLri-2 and 3CVLri-4. All amino acid residues in D-configuration and with a reversal peptide bond. Additionally, an arginine was added to the C-termini of each ri-peptide resulting in 17 amino acid residues. The C-termini of the D-ri-peptides are amidated, and the N-termini are acetylated. An analysis certificate from Caslo (Lyngby, Denmark) demonstrated the purity of each peptide (Supplementary Figs. S23 to S26 and Table S5), High-performance liquid chromatography (HPLC) analysis was performed with a C18 column. For the mass spectrometry certificate, the MALDI-TOF method was used.

### *Protein expression and purification*

SARS-CoV-2 3CL$^{pro}$ (Uniprot entry: P0DTD1, virus strain: hCoV-19/Wuhan/WIV04/2019) was cloned, expressed, and purified previously.[35]

***Phage display***

Two independent phage display selections (selection A and B), each with three rounds, were performed to obtain peptides that bind $3CL^{pro}$. Therefore, in both selections, 100 µg/mL purified $3CL^{pro}$ diluted in 100 µL 20 mM potassium phosphate buffer pH 7.5 with 150 mM sodium chloride were immobilised on an amino plate (Nunc Amino Immobilizer 96 well plates, polystyrene; Thermo Scientific, Waltham, USA). After 30 min at RT, the $3CL^{pro}$ solution was removed, and the surface was quenched with 1 M ethanolamine pH 8.5 (Cytiva; Chicago, USA) for 1 h at RT. Additionally, the surface was blocked with 200 µL 10 mg/mL BSA in 20 mM potassium phosphate buffer pH 7.5 with 150 mM sodium chloride in the first round and with 200 µL 10 mg/mL milk powder in 20 mM potassium phosphate buffer pH 7.5 with 150 mM sodium chloride in the second round to reduce unspecific binding events. In the third round, the surface was not additionally blocked to vary the selection conditions. The surface was washed 6 times with 200 µL washing buffer (20 mM potassium phosphate buffer pH 7.5 with 150 mM sodium chloride and 0.05% tween-20). During the first and second rounds, 2 mg/mL BSA or milk powder was added to the washing buffer, respectively. Then, $7.2 * 10^{10}$ phages (TriCo-16 phage display peptide library; Creative Biolabs, New York, USA) in 100 µL of the corresponding washing buffer were added and incubated for 20 min at RT with the exception of the third round of selection B. Here, the incubation time was extended to 30 min. After removing of non-binding phages, the well was washed with 200 µL of the respective washing buffer of the corresponding round. The number of washing steps was varied in each selection round (selection A: 2-5-5; selection B: 5-8-10). Elution of phages was performed by incubation with 100 µL 0.2 M Glycine-HCl pH 2.2 for 10 min at RT. The phage containing solution was then removed from the plate and neutralised by transferring it to a tube containing 25 µL 1 M Tris-HCl pH 9.1. To determine the output titer, 5 µL of eluted phages were used. Therefore, a dilution series from $10^{-2}$ to $10^{-8}$ was prepared with the eluted phages in a total volume of 100 µL LB medium. Each dilution was mixed with 100 µL *E. coli* K12 ER2738 (OD600 of 0.6) and plated with 800 µL top agar

on plates (35 x 10 mm; Sarstedt, Nümbrecht, Germany) containing LB-Agar-IPTG-XGal. After overnight incubation at 37 °C, the plaques were counted to determine the output titer.

The remaining eluted phages (120 µL) were amplified in 20 mL *E. coli* K12 ER2738, starting with an $OD_{600}$ of 0.1 for 3.5 h at 37 °C and 120 rpm. The culture was then centrifuged for 20 min at 2700 x g and 4 °C. The pellet was discarded, and the supernatant was incubated with 7 mL PEG-8000/2.5 M sodium chloride overnight at 4 °C for phage precipitation. Subsequently, the solution was centrifuged for 1 h at 4 °C, and 2700 x g. The phage containing pellet was dissolved in 1 mL 1 x PBS before another centrifugation step at 4 °C and 11000 x g for 5 min to get rid of residual bacterial components. Afterwards, the supernatant was added to 200 µL PEG-8000/2.5 M sodium chloride and incubated for 1 h on ice, followed by final centrifugation for 45 min at 2700 x g and 4 °C. The phage containing pellet was properly resuspended in 100 µL 1 x PBS. The input titer was determined by spectrophotometry[40] in 1 x PBS using a 1:10 dilution. The resulting phages were then used in the next selection round with the same phage amount as before and further used for an enrichment ELISA and prepared for next generation sequencing (NGS) analysis.

Besides the primary selection with $3CL^{pro}$ as the target (Target Selection, TS), two control selections were performed as described before[40] to enable an efficient evaluation of the selection success and to further facilitate the identification of target-binding peptides over those peptides that have accumulated, e.g. because of an affinity to the surface. The first control selection, called Empty Selection (ES), was performed in the same way as TS but without $3CL^{pro}$. During immobilisation, 20 mM potassium phosphate buffer pH 7.5 with 150 mM sodium chloride was used instead of $3CL^{pro}$. One further control, Direct Control (DC), was performed starting from the second round without $3CL^{pro}$, but in contrast to ES, the phages resulting from each round of TS were used for the following selection rounds.

*Enrichment ELISA*

An enrichment ELISA was performed to validate the success of the phage display selection. Therefore, 100 µg/mL $3CL^{pro}$ diluted in 100 µL 20 mM potassium phosphate buffer pH 7.5 with 150 mM sodium chloride was immobilised on an amino plate (Nunc Immobilizer Amino 96 well plate, polystyrene; Thermo Scientific, Waltham, USA) for 30 min at RT. Non-coated wells (20 mM potassium phosphate buffer pH 7.5 with 150 mM sodium chloride without $3CL^{pro}$) were used as a control for each selection round. The solution was removed, and the surface was quenched with 200 µL 1 M ethanolamine pH 8.5 (Cytiva; Chicago, USA) for 1 h at RT, followed by a blocking step with 200 µL blocking solution (10 mg/mL BSA in 20 mM potassium phosphate buffer pH 7.5 with 150 mM sodium chloride) for 15 min at RT. After three washing steps with 200 µL washing buffer (150 mM sodium chloride, 0.05 % Tween-20, 2 mg/mL BSA in 20 mM potassium phosphate buffer pH 7.5), the amplified and purified phages of each target selection round were diluted in washing buffer to a total amount of $5 * 10^{10}$ phages in 100 µL. The wells immobilised with $3CL^{pro}$ and non-coated wells were incubated with the corresponding phage solution for 1 h at RT. Unbound phages were removed by five washing steps with 150 µL washing buffer. The anti-M13 antibody (HRP conjugated mouse monoclonal anti-bacteriophage M13 antibody; Sino Biological, Peking, China) was diluted in washing buffer to a final concentration of 0.35 ng/µL, and 100 µL was added to each well for 1 h at RT. After 6 washing steps with 150 µL washing buffer, the supernatant was removed entirely, and 100 µL of the 3,3',5,5'-tetramethylbenzidine (TMB) solution (TMB was previously dissolved in 1 mL DMSO and diluted with 9 mL 0.05 M phosphate citrate buffer pH 5) were added into each well. After 2.5 min incubation at RT the reaction was stopped with 100 µL 2 M $H_2SO_4$. The absorption was quantified at 450 nm using a microplate reader (BMG Labtech, Ortenberg, Germany).

In addition to the samples from the target selection (TS1, TS2, TS3), the library and the wild type phage without peptides were analysed. The measurement was performed

in single determination to save protein samples. Non-coated wells that were incubated with washing buffer without phages were used as a control for the background signal.

***Sequence analysis***

Extraction and the amplification of the single-stranded phage DNA from each sample from the phage display selection were purified and prepared for NGS analysis as described previously.[40] NGS analysis was performed by the Biologisch-Medizinisches Forschungszentrum at the Heinrich Heine University Düsseldorf, Germany. The resulting sequences were further processed and evaluated with the targeted sequencing analysis tool (TSAT) and the open source software Hammock for subsequent sequence alignment.[40,41]

***SARS-CoV-2 3CL$^{pro}$ activity assay***

SARS-CoV-2 3CL$^{pro}$ activity assay was performed as described previously.[35,42-44] A fluorogenic substrate DABCYL-KTSAVLQ↓SGFRKME-EDANS (Bachem, Bubendorf, Switzerland) was used, and the assay buffer contained 20 mM Tris pH 7.2, 200 mM NaCl, 1 mM EDTA, and 1 mM TCEP. The reaction mixture was pipetted in a Corning 96-Well plate (Merck, Darmstadt, Germany) consisting of 0.5 μM 3CL$^{pro}$. The assay was initiated with the addition of the substrate at a final concentration of 50 μM. The fluorescence intensities were measured at 60 s intervals over 30 minutes using an Infinite 200 PRO plate reader (Tecan, Männedorf, Switzerland), and the temperature was set to 37 °C. The excitation and emission wavelengths were 360 and 460 nm, respectively.

***SARS-CoV-2 3CL$^{pro}$ inhibition assay***

Inhibition of SARS-CoV-2 3CL$^{pro}$ activity by 3CVL-1 to 10 and 3CVLri-2, -4 and -7 was investigated using the described activity assay. 10 μM of the L-peptides was used for a preliminary screening test. For the final inhibition assays, 0.5 μM of the protein was incubated with 0-140 μM 3CVL-2, 0-120 μM 3CVL-4 and 0-100 μM 3CVL-7. The

corresponding 3CVLri peptides were incubated with the same inhibitor concentrations, except for 3CVLri-4 (0-140 μM). The mixtures were incubated for 30 minutes at RT. When the substrate with a final concentration of 50 μM was added to the mixture, the fluorescence intensities were measured at 60 s intervals over 30 minutes using an Infinite 200 PRO plate reader (Tecan, Männedorf, Switzerland). The temperature was set to 37 °C, and the excitation and emission wavelengths were 360 and 460 nm, respectively. Inhibition assays were performed as triplicates.

For the 3CVLri-2 + 3CVLri-4 and 3CVLri-4 + 3CVLri-7 combination test, a 1:1 stock solution of the molecules was prepared. 0.5 μM of the protein was incubated with 0-75 μM (3CVLri-2 + 3CVLri-4) and 0-50 μM (3CVLri-4 + 3CVLri-7) of the combined molecules. The $IC_{50}$ values were calculated by plotting the initial velocity against various concentrations of the peptides using a dose-response curve in GraphPad Prism software version 8 (San Diego, CA, USA). Data are presented as mean ± SD.

***Determination of inhibition mode***

The inhibition mode was determined using different final concentrations of the inhibitors and substrate. Briefly, 0.5 μM SARS-CoV-2 $3CL^{pro}$ was incubated with the inhibitor, in various concentrations, for 30 minutes at RT. Subsequently, the reaction was initiated by adding the corresponding concentration series of the substrate. The data were analysed using a Lineweaver-Burk plot; therefore, the reciprocal of velocity (1/V) vs the reciprocal of the substrate concentration (1/[S]) was compared[45,46]. All measurements were performed in triplicate, and data are presented as mean ± SD.

***Inhibitor stability over 24 h***

Stable inhibition of SARS-CoV-2 $3CL^{pro}$ by 3CVLri-2, -4, and -7 was followed via a 24h inhibition experiment. Briefly, 0.5 μM SARS-CoV-2 $3CL^{pro}$ was incubated with 5 μM (D-ri peptides 2 and 4) and 20 μM (D-ri peptide 7) and incubated for ½ h, 1 h, 2 h, 3 h, 4 h, 5 h and 24 h at RT. The control was performed with $3CL^{pro}$ without the peptides and

measured together after each time point. Subsequently, the reaction was initiated by the addition of the substrate. All measurements were performed in triplicate, and data are presented as mean ± SD.

**Assay to exclude D-ri peptides as promiscuous inhibitors.**

A detergent-based control assay was performed to exclude inhibitors that possibly act as promiscuous aggregators of the $3CL^{pro}$ by adding 0.001%, 0.01%, and 0.1% of Triton X-100 to the reaction[47]. Four concentrations of 3CVLri-2 (1 μM, 5 μM, 10 μM and 20μM), 3CVLri-4 (1 μM, 5 μM, 10 μM and 20μM) and 3CVLri-7 (0.25 μM, 0.5 μM, 1 μM and 5 μM) were tested. All measurements were performed in triplicate, and data are presented as mean ± SD.

***Circular Dichroism spectroscopy***

Circular Dichroism (CD) measurements were carried out with a Jasco J-1100 Spectropolarimeter (Jasco, Germany). Far-UV spectra were measured at 190 to 260 nm using 5 μM $3CL^{pro}$ in 20 mM $K_2HPO_4/KH_2PO_4$ ph 7.4 and a single peptide concentration of 30 μM in $H_2O$. The secondary structure of 3CVL-2, 3CVL-4, 3CVL-7 and the related D-ri peptides was checked. A 1 mm path length cell was used for the measurements; 15 repeat scans were obtained for each sample, and five scans were conducted to establish the respective baselines. The averaged baseline spectrum was subtracted from the averaged sample spectrum. The results are presented as molar ellipticity [θ], according to the equation (1):

$$[\theta]\lambda = \theta/(c*0.001*l*n) \quad (1)$$

where θ is the ellipticity measured at the wavelength λ (deg), c is the peptide concentration (mol/L), 0.001 is the cell path length (cm), and n is the number of amino acids.

***Surface Plasmon Resonance measurements***

Surface Plasmon Resonance (SPR) measurements with 3CVL peptides and SARS-CoV-2 $3CL^{pro}$ were performed using a Biacore 8K device (GE Healthcare, Uppsala, Sweden). Recombinant $3CL^{pro}$ was immobilised on a CM5 sensor surface (GE Healthcare, Uppsala, Sweden) via covalent coupling of primary amino groups. The protease was diluted to a final concentration of 15 μg/ml in 10 mM NaAc pH 5.5 and injected for 600 s at 10 μl/min to reach a final immobilisation level of 2700 to 3000 RU. After quenching with Ehanolamine at pH 8.5, kinetic experiments were performed as parallel injections on all 8 flow channels. 3CVL and 3CVLri peptides were diluted in 20 mM Tris pH 7.4 100 mM NaCl 1 mM EDTA 0.05 % Tween-20 in concentration range of 0.03 to 50 μM with a 1:1 dilution series and injected at 30 μl/min. Surface regeneration was performed after dissociation ended for 2 x 45 s at 30 μl/min using an acidic regeneration solution consisting of 150 mM oxalic acid, 150 mM phosphoric acid, 150 mM formic acid and 150 mM malonic acid at pH 5.0. Data evaluation was performed with Biacore Insight Evaluation Software v3.0.12 (GE Healthcare, Uppsala, Sweden) using a heterogenous ligand fit model for the competitive inhibitors and a global 1:1 kinetic fit model for the non-competitive inhibitor.

***Metabolic Stability***

Metabolic stability experiments of the 3CVLri peptides were performed as described before.[48,49] Solutions simulating gastric and intestinal fluid (SGF & SIF) were prepared according to the European Pharmacopoeia 7.0. Human plasma samples was purchased from Innovative Research (Novi, Michigan, USA), and human liver microsomes were purchased from Sekisui XenoTech (Kansas City, USA; H1000). The liver microsomes were diluted in an NADPH regenerating system (NRS). For the stability tests, 150 μM of the 3CVLri peptides were incubated in SGF, SIF, human plasma and human liver microsomes in triplicate at 37 °C with slight shaking for different periods. The peptides were extracted by precipitating the proteins with a solution containing

acetonitrile and 0.5% formic acid. Afterwards, the mixture was mixed and heated for 5 minutes at 95°C. The sample was subsequently centrifuged at 14.000 g for 10 minutes at 4°C. The extracted peptides are in the supernatant and were analysed by reversed-phase high-performance liquid chromatography (RP-HPLC). The RP-HPLC system (Agilent Technologies, Santa Clara, USA; 1260 series) consisted of an autosampler, quaternary pump, a thermostatted column compartment and a variable wavelength detector. Chromatography was performed with a C18 column (Agilent Technologies, Santa Clara, USA; ZORBAX 300SB-C18 5 μm, 4.6 × 250 mm) at 25 °C and 214 nm with a flow rate of 1 mL/min. The sample injection volume was 20 μl. Chromatograms were recorded and analysed with the Agilent software OpenLab version 2-5. Mobile phases were acetonitrile + 0.1 % Trifluoroacetic acid (TFA) (A) and water + 0.1 % TFA (B) for sample analysis from stability tests. The samples were measured with an initial isocratic step at 15% solvent A for 3 min followed by a gradient elution to 45 % A in 15 min. All measurements were performed in triplicate, and data are presented as mean ± SD.

### *Statistical analysis*

All data are expressed as the mean ± the standard deviations (SDs). The statistical significance of the mean values' differences was assessed with one-way analyses of variance (ANOVA), followed by Tukeys' multiple comparison test. Significant differences were considered at $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (***). All statistical analyses were performed with GraphPad Prism software version 8 (San Diego, CA, USA).

### *Docking and Molecular dynamics*

### *Peptide and Protein structure preparation*

The peptide models were constructed as linear chains using the python module PeptideBuilder[50]. In order to obtain D-retro-inverso (D-ri) peptides, the peptides were first generated as L-retro-inverso molecules, followed by the inversion of their stereochemical

configuration by flipping the structure files with cartesian coordinates along the x-axis. The sequences of the peptides are shown in Table 6.

**Table 6.** Sequences of 3CVL and 3CVLri peptide models.

| Peptide | Sequence |
|---|---|
| 3CVL-2 | SPHGWPSQSIEVQPQW |
| 3CVL-4 | AHEGWTWDWTPQYSWK |
| 3CVL-7 | TVAPLHAHYWDVEERH |
| 3CVLri-2 | wqpqveisqspwahpsr |
| 3CVLri-4 | kwsyqptwdwtwaehar |
| 3CVLri-7 | hreevdwyhahlpavtr |

The initial structural data of the $3CL^{pro}$ dimer was obtained from a crystal structure (PDB: 6M2N).

***Docking of the peptides against the 3CL$^{pro}$ dimer***

3CVL-2, 3CVL-7, 3CVLri-2 and 3CVLri-7 were docked against the active site of chain A of the $3CL^{pro}$ dimer (PDB: 6M2N), which can be defined by the residues His41, Met49, Tyr54, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166, Leu167, Asp187, Arg188, Gln189, Thr190 and Gln192[51].

3CVL-4 and 3CVLri-4 bind $3CL^{pro}$ allosterically, and both peptides were docked against the potential binding site residues Lys12, Lys97, Lys100, Tyr101, Lys102, Phe103, Val104 and Arg105.[35]

Autodock Tools[52,53] were used to generate a docking grid around the positions of the binding site residues of the $3CL^{pro}$ dimer. The grid was centered at x,y,z-position of -33.572, -62.854, 40.9 (active site) and -49.16, -47.323, 38.911 (allosteric binding site). After adding hydrogen atoms to the $3CL^{pro}$ dimer and the peptides, Gasteiger partial charges were computed and added to the dimer and the peptides using AutoDock Tool. The protein and peptide models were saved in the PDBQT format. The docking was then performed using AutoDock Vina,[36] which treats the ligand as flexible while keeping the receptor rigid.

***Simulation setup and production run***

All MD simulations and analyses were performed with GROMACS 2021.4.[54] AMBER14SB[55] was used as a force field and the TIP3P water model[56] for explicit water simulation. The peptides and the protein-peptide complexes were centred in a dodecahedron box with 240.8 $nm^3$-370.5 $nm^3$ and 872.5 $nm^3$, solvated with water and neutralised with $Na^+$ and $Cl^-$ ions resulting in a system size of ~23,500-36,000 and ~86000 atoms, respectively.

The systems' energy was minimised using the steepest descent algorithm.[57] The systems were then equilibrated in two steps. First, a 0.1 ns simulation was performed in the *NVT* ensemble and second, a 1 ns simulation in the *NpT* ensemble at 310 K - velocity-rescaling thermostat[58] and 1.0 bar - Berendsen barostat.[59] First, 1 μs peptide MD simulations were performed. After the peptides were docked against the $3CL^{pro}$ dimer, we performed 100 ns of protein-peptide complexes with the temperature set to 310 K (Nosé-Hoover thermostat[60] and the pressure set to 1 bar (Parinello-Rahman barostat).[61] One selected simulation of each protein-peptide complex was then extended to 200 ns and ran in triplicates.

Electrostatic interactions were processed with the particle-mesh Ewald method,[62,63] combined with periodic boundary conditions and a real-space cutoff of 12 Å. The same cutoff was applied to Lennard-Jones (LJ) interactions. The leapfrog stochastic dynamics integrator was used with a time step of 2 fs to integrate the equations of motion. During the MD simulations, all bond lengths were constrained using the LINCS algorithm.[64] The MD simulations were performed in triplicates (3 x 500 ns), and coordinates were saved every 20 ps.

***Molecular Dynamics Analysis***

The peptide's root means square deviation (RMSD) indicates how flexible the ligand is within the binding site. The protein structures sampled during the MD

simulations were aligned to the MD starting structure, and then the RMSD was calculated for the ligand using *gmx rms*.

In order to investigate the effect of the allosteric peptides and the occupation of the peptides that bind to the active site, we used the VMD plugin Epock 1.0.5 to evaluate the pocket volume of both active sites of the dimer for 1000 snapshots of the whole MD simulation after fitting the protein to the starting structure.[65]

The following simulations were conducted for the second half of the MD simulations. The root mean square fluctuation (RMSF) was determined for each peptide with *gmx rmsf* to identify the flexible and rigid areas of the peptide within the protein-peptide complex. The average minimum distance between specific residues of the protein and peptide using *gmx mindist*. The binding free energy $\Delta G_{bind}$ between the protein and the peptide was computed using the MM/PBSA method as implemented in g_mmpbsa.[66] For this analysis, 500 MD snapshots were used in the second half of the simulation. Here, the binding free energy is defined as:

$$\Delta G_{bind} = \langle G_{complex} - G_{protein} - G_{ligand} \rangle$$

with $\langle\rangle$ indicating the average of over 100 snapshots. The free energy for each entity is given as:

$$G = E_{bonded} + E_{Coul} + E_{LJ} + G_{polar} + G_{nonpolar} - TS$$

where $E_{bonded}$ describes the bonded interactions, $E_{Coul}$ and $E_{LJ}$ indicate the coulomb and LJ interactions, and $G_{polar}$ and $G_{nonpolar}$ are the polar and nonpolar contributions to the solvation free energy. The TS term includes the absolute temperature, T, and the configurational entropy, S.

To identify the peptide structures with the most populated configuration or binding pose in the protein-peptide complex, the algorithm of Daura et al. was used as implemented in GROMACS (*gmx cluster*).[67]

The cutoff for clustering the peptide configurations was set to 3.5 Å, and the average structures of the 10 most populated clusters were used for docking against the dimer of the main protease. The binding poses of the peptides in the elongated protein-peptide complex simulations were clustered using a cutoff of 3.0 Å. To avoid biasing of non-binding flexible areas of the peptide, only those peptide residues with an RMSF of < 6 Å were considered for clustering. The three most prominent clusters were then analysed for interacting residues by generating a contact map.

### *Conflicts of interest*

The authors declare no competing financial interest.

### *Submission Declaration and Verification*

The authors guarantee that the manuscript describes original work, and it is not under consideration for publication elsewhere and has not been published in any medium, including electronic journals and computer databases of a public nature. All authors approved the manuscript and this submission.

### *Author Contributions*

Conceptualisation, R.J.E., J.M. and M.A.C.; methodology, R.J.E., M.S., I.G., L.S., B.S., K.B.S., J.M. and M.A.C.; validation, R.J.E., M.S., I.G., L.S., B.S., K.B.S., J.M. and M.A.C.; formal analysis, R.J.E., M.S., I.G., L.S., B.S., K.B.S., J.M. and M.A.C.; investigation, R.J.E. and M.A.C.; resources, B.S. and D.W.; writing—original draft preparation, R.J.E. and M.A.C.; writing—review and editing, R.J.E., M.S., I.G., L.S., B.S., K.B.S., J.M., M.A.C. and D.W. All authors have read and agreed to the published version of the manuscript.

### *Acknowledgments*

We would like to thank the support of the Institute of Biological Information Processing (IBI-7) Forschungszentrum Jülich, Germany. L.S. and B.S. gratefully

acknowledge the computing time granted by the JARA Vergabegremium and provided under project "3cldtp" on the JARA Partition part of the supercomputer JURECA at Forschungszentrum Jülich.

*Funding*

D.W. received funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project-ID 267205415, SFB 1208. B.S. and L.S. recieved funding from the VolkswagenStiftung for that project (project number 99761).

**References**

(1) Hu, B.; Guo, H.; Zhou, P.; Shi, Z.L. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **2021,** *19,* 141-154.

(2) WHO. 2022. World Health Organization, Coronavirus disease 2019 (COVID-19) Dashboard (Acessed 25.05.2022).

(3) Mathieu, E.; Ritchie, H.; Ortiz-Ospina, E.; Roser, M.; Hasell, J.; Appel, C.; Giattino, C.; Rodés-Guirao, L. A global database of COVID-19 vaccinations. *Nat. Hum. Behav.* **2021,** *5,* 947–953.

(4) Polack, F.P.; Thomas, S.J.; Kitchin, N.; Absalon, J.; Gurtman, A.; Lockhart, S.; Perez, J.L.; Marc, G.P.; Moreira, E.D.; Zerbini, C.; Bailey, R.; Swanson, K.A.; Gruber, W.C. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Engl. J. Med.* **2020,** *383,* 2603-2615.

(5) Baden, L.R.; El Sahly, H.M.; Essink, B.; Kotloff, K.; Frey, S.; Novak, R.; Diemert, D.; Spector, S.A.; Rouphael, N.; Creech, C.B.; McGettigan, J.; Khetan, S.; Zaks, T. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med.* **2021,** *384,* 403-416.

(6) Voysey, M.; Clemens, S.A.C.; Madhi, S.A.; Weckx, L.Y.; Folegatti, P.M.; Aley, P.K.; Bijker, E. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *The Lancet.* **2021,** *397,* 99-111.

(7) Livingston, E.H.; Malani, P.N.; Creech, C.B. The Johnson & Johnson Vaccine for COVID-19. *JAMA.* **2021,** *325,* 1575-1575.

(8) Owen, D.R.; Allerton, C.M.; Anderson, A.S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R.D.; Carlo, A.; Coffman, K.J.; Dantonio, A. (2021) An oral SARS-

CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. *Science.* **2021,** *374,* 1586-1593.

(9) Madsen, L.W. Remdesivir for the Treatment of Covid-19-Final Report. *N. Engl. J. Med.* **2020,** *338,* 1813-1826.

(10) Hosseinzadeh, M.H.; Shamshirian, A.; Ebrahimzadeh, M.A. Dexamethasone Vs. COVID-19: An Experimental Study in Line with the Preliminary Findings of a Large *Trial. Int. J. Clin. Pract.* **2020,** *75,* e13943.

(11) Nangaku, M.; Kadowaki, T.; Yotsuyanagi, H.; Ohmagari, N.; Egi, M.; Sasaki, J.; Sakamoto, T.; Hasegawa, Y.; Ogura, T.; Chiba, S.; Node, K.; Suzuki, R.; Yamaguchi, Y.; Murashima, A.; Ikeda, N.; Morishita, E.; Yuzawa, K.; Moriuchi, H.; Hayakawa, S.; Nishi, D.; Irisawa, A.; Miyamoto, T.; Suzuki, H.; Sone, H.; Fujino, Y. The Japanese Medical Science Federation COVID-19 Expert Opinion English Version. *JMA J.* **2021,** *4,* 148-162.

(12) Italian Society of Infectious and Tropical Diseases (SIMIT). Guidelines for the treatment and support management of patients with COVID-19 coronavirus infection (second edition) (Italy). 2020.

(13) Rambaut, A.; Holmes, E.C.; O'Toole, Á.; Hill, V.; McCrone, J.T.; Ruis, C.; du Plessis, L.; Pybus, O.G. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **2020,** *5,* 1403-1407.

(14) Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S.H. An overview of Severe Acute Respiratory Syndrome-Coronavirus (SARS-CoV) 3CL protease inhibitors: Peptidomimetics and small molecule chemotherapy. *J. Med. Chem.* **2016,** *59,* 6595–6628.

(15) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, LW.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. **Nature. 2020,** *582,* 289–293.

(16) Rut, W.; Groborz, K.; Zhang, L.; Sun, X.; Zmudzinski, M.; Pawlik, B.; Wang, X.; Jochmans, D.; Neyts, J.; Młynarski, W.; Hilgenfeld, R.; Drag, M. SARS-CoV-2 Mpro inhibitors and activity-based probes for patient-sample imaging. *Nat. Chem. Biol.* **2021,** *17,* 222–228.

(17) Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J.R.; Hilgenfeld, R. Coronavirus main proteinase (3CLpro) structure: Basis for design of anti-SARS drugs. *Science.* **2003,** *300,* 1763–1767.

(18) Ghosh, A.K.; Osswald, H.L.; Prato, G. Recent progress in the development of HIV-1 protease inhibitors for the treatment of HIV/AIDS. *J. Med. Chem.* **2016,** *59,* 5172–5208.

(19) Cannalire, R.; Barreca, M.L.; Manfroni, G.; Cecchetti, V. A journey around the medicinal chemistry of Hepatitis C virus inhibitors targeting NS4B: From target to preclinical drug candidates. *J. Med. Chem.* **2016,** *59,* 16–41.

(20) Craik, D.J.; Fairlie, D.P.; Liras, S.; Price, D. The future of peptide-based drugs. *Chem. Biol. Drug Des.* **2013,** *81,* 136-147.

(21) Mikitsh, J.L.; Chacko, A.M. Pathways for small molecule delivery to the central nervous system across the blood-brain barrier. *Perspect. Med. Chem.* **2014,** *6,* 11-24.

(22) Goodwin, D.; Simerska, P.; Toth, I. Peptides as therapeutics with enhanced bioactivity. *Curr. Med. Chem.* **2012,** *19,* 4451-4461.

(23) Jiang, N.; Leithold, L.H.; Post, J.; Ziehm, T.; Mauler, J.; Gremer, L.; Cremer, M.; Schartmann, E.; Shah, N.J.; Kutzsche, J.; Langen, K.J. Preclinical pharmacokinetic studies of the tritium labelled D-enantiomeric peptide D3 developed for the treatment of Alzheimer s disease. *PloS one.* **2015,** *10,* e0128553.

(24) Leithold, L.H.; Jiang, N.; Post, J.; Ziehm, T.; Schartmann, E.; Kutzsche, J.; Shah, N.J.; Breitkreutz, J.; Langen, K.J.; Willuweit, A.; Willbold, D. Pharmacokinetic properties of a novel D-peptide developed to be therapeutically active against toxic β-amyloid oligomers. *Pharm. Res.* **2016,** *33,* 328-336.

(25) Chong, P.; Sia, C.; Tripet, B.; James, O.; Klein, M. Comparative immunological properties of enantiomeric peptides. *Lett. Pept. Sci.* **1996,** *3,* 99-106.

(26) Zhang, T.; Gering, I.; Kutzsche, J.; Nagel-Steger, L.; Willbold, D. Toward the mode of action of the clinical stage all-D-enantiomeric peptide RD2 on Aβ42 aggregation. *ACS Chem. Neurosci.* **2019,** *10,* 4800-4809.

(27) Kutzsche, J.; Jürgens, D.; Willuweit, A.; Adermann, K.; Fuchs, C.; Simons, S.; Windisch, M.; Hümpel, M.; Rossberg, W.; Wolzt, M.; Willbold, D. Safety and pharmacokinetics of the orally available antiprionic compound PRI-002: A single and multiple ascending dose phase I study. *Alzheimer's Dement: Transl. Res. Clin. Interv.* **2020,** *6,* e12001.

(28) Schemmert, S.; Camargo, L.C.; Honold, D.; Gering, I.; Kutzsche, J.; Willuweit, A.; Willbold, D. In Vitro and In Vivo Efficacies of the Linear and the Cyclic Version of an All-d-Enantiomeric Peptide Developed for the Treatment of Alzheimer's Disease. *Int. J. Mol. Sci.* **2021,** *22,* 6553.

(29) Valiente, P.A.; Wen, H.; Nim, S.; Lee, J.; Kim, H.J.; Kim, J.; Perez-Riba, A.; Paudel, Y.P.; Hwang, I.; Kim, K.D.; Kim, S. Computational Design of Potent D-Peptide Inhibitors of SARS-CoV-2. *J. Med. Chem.* **2021,** *64,* 14955-14967.

(30) Hernández González, J.E.; Eberle, R.J.; Willbold, D.; Coronado, M.A. A computer-aided approach for the discovery of D-peptides as inhibitors of SARS-CoV-2 main protease. *Front. Mol. Biosci.* **2022,** *8,* 816166.

(31) Eberle, R.J.; Gering, I.; Tusche, M.; Ostermann, P.N.; Müller, L.; Adams, O.; Schaal, H.; Olivier, D.S.; Amaral, M.S.; Arni, R.K.; Willbold, D.; Coronado, M.A. Design of D-Amino Acids SARS-CoV-2 Main Protease Inhibitors Using the Cationic Peptide from Rattlesnake Venom as a Scaffold. *Pharmaceuticals.* **2022,** *15,* 540.

(32) Cardoso, M.H.; Cândido, E.S.; Oshiro, K.G.N.; Rezende, S.B.; Franco, O.L. in Peptide Applications in Biomedicine, Biotechnology and Bioengineering, edited by Koutsopoulos, S. (Woodhead Publishing), **2018.**

(33) Chen, H.; Wei, P.; Huang, C.; Tan, L.; Liu, Y.; Lai, L. Only one protomer is active in the dimer of SARS 3C-like proteinase. *J. Biol. Chem.* **2006,** *281,* 13894-13898.

(34) Kidera, A.; Moritsugu, K.; Ekimoto, T.; Ikeguchi, M. Allosteric regulation of 3CL protease of SARS-CoV-2 and SARS-CoV observed in the crystal structure ensemble. *J. Mol. Biol.* **2021,** *433,* p.167324.

(35) 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 3CLpro In Vitro. *Viruses.* **2021,** *13,* 873.

(36) Trott, O.; Olson, A. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimisation, and multithreading. *J. Comput. Chem.* **2010,** *31,* 455–461.

(37) Bojadzic, D.; Alcazar, O.; Chen, J.; Chuang, S.T.; Capcha, J.M.C.; Shehadeh, L.A.; Buchwald, P. Small-Molecule Inhibitors of the Coronavirus Spike: ACE2 Protein–Protein Interaction as Blockers of Viral Attachment and Entry for SARS-CoV-2. *ACS Infect. Dis.* **2021,** *6,* 1519-1534.

(38) 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.

(39) Singh, E.; Khan, R.J.; Jha, R.K.; Amera, G.M.; Jain, M.; Singh, R.P. Muthukumaran, J.; Singh, A.K. A comprehensive review on promising antiviral therapeutic candidates identified against main protease from SARS-CoV-2 through various computational methods. *J. Genet. Eng. Biotechnol.* **2020,** *18,* 1-12.

(40) Santur, K.; Reinartz, E.; Lien, Y.; Tusche, M.; Altendorf, T.; Sevenich, M.; Tamgüney, G.; Mohrlüder, J.; Willbold, D. Ligand-Induced Stabilisation of the Native Human Superoxide Dismutase 1. *ACS Chem. Neurosci.* **2021,** *12,* 2520-2528.

(41) 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.

(42) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved $\alpha$-ketoamide inhibitors. *Science.* **2020,** *368,* 409-412.

(43) Zhang, L.; Lin, D.; Kusov, Y.; Nian, Y.; Ma, Q.; Wang, J.; De Wilde, A. $\alpha$-Ketoamides as broad-spectrum inhibitors of corona-virus and enterovirus replication: Structure-based design, synthesis, and activity assessment. *J. Med. Chem.* **2020,** *63,* 4562–4578.

(44) Ma, C.; Sacco, M.D.; Hurst, B.; Townsend, J.A.; Hu, Y.; Szeto, T.; Zhang, X.; Tarbet, B.; Marty, M.T.; Chen, Y.; Wang, J. Bo-ceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res.* **2020,** *30,* 678-692.

(45) Roy, A.; Lim, L.; Srivastava, S.; Lu, Y.; Song, J. Solution conformations of Zika NS2B-NS3pro and its inhibition by natural products from edible plants. *PloS one.* **2017,** *12,* e0180632.

(46) Motulsky, H.; Christopoulos, A. Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting. Oxford University Press. **2004.**

(47) Feng, B.Y.; Shoichet, B.K. A detergent-based assay for the detection of promiscuous inhibitors. *Nat. Protoc.* **2006,** *1,* 550-553.

(48) Elfgen, A.; Santiago-Schübel, B.; Gremer, L.; Kutzsche, J.; Willbold, D. Surprisingly high stability of the Aβ oligomer eliminating all-d-enantiomeric peptide D3 in media simulating the route of orally administered drugs. *Eur. J. Pharm. Sci.* **2017,** *107,* 203-207.

(49) Elfgen, A.; Hupert, M.; Bochinsky, K.; Tusche, M.; González de San Román, M.E.; Gering, I.; Sacchi, S.; Pollegioni, L.; Huesgen, P.F.; Hartmann, R.; Santiago-Schübel, B. Metabolic resistance of the D-peptide RD2 developed for direct elimination of amyloid-β oligomers. *Sci. Rep.* **2019,** *9,* 1-13.

(50) Tien, M.Z.; Sydykova, D.K.; Meyer, A.G.; Wilke, C.O. PeptideBuilder: A simple Python library to generate model peptides. *Peer J.* **2013,** *1,* e80.

(51) Tang, B.; He, F.; Liu, D.; Fang, M.; Wu, Z.; Xu, D. AI-aided design of novel targeted covalent inhibitors against SARS-CoV-2. *bioRxiv.* **2020,** 2020.03.03.972133

(52) Goodsell, D.; Morris, G.; Olson, A. Automated docking of flexible ligands: Applications of AutoDock. *J. Mol. Recognit.* **1996,** *9,* 1–5.

(53) Santos-Martins, D.; Forli, S.; Ramos, M.; Olson, A. AutoDock4Zn: An Improved AutoDock Force Field for Small-Molecule Docking to Zinc Metalloproteins. *J. Chem. Inf. Model.* **2014,** *54,* 2371–2379.

(54) Abraham, M.; Murtola, T.; Schulz, R.; P´all, S.; Smith, J.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* **2015,** *2,* 19–25.

(55) Maier, I.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14sb: Improving the accuracy of protein side chain and backbone parameters from ff99sb. *J. Chem. Theory Comput.* **2015,** *11,* 3696–3713.

(56) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of simple potential functions for simulating liquid water. *J. Chem. Theory Comput.* **1983,** *79,* 926–935.

(57) Cauchy, M.A. Méthode généerale pour la ystemson des ystems d' éequa- tions simultanéees. *CR Hebd. Acad. Sci.* **1847,** *25,* 536–538.

(58) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007,** *126,* 014101.

(59) Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984,** *81,* 3684–3690.

(60) Nos´e, S. Molecular-dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984,** *52,* 255–268.

(61) Parrinello, M.; Rahman, A. Polymorphic transitions in single-crystals – a new molecular-dynamics method. *Mol. Phys.* **1981,** *52,* 7182–7190.

(62) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald – an N. Log(N) method for ewald sums in large systems. *J. Chem. Phys.* **1993,** *98,* 10089–10092.

(63) Essmann, U.; Perera, L.; Berkowitz, M. A smooth particle mesh ewald method. *J. Chem. Phys.* **1995,** *103,* 8577–8593.

(64) Hess, B.; Bekker, H.; Berendsen, H.; Fraaije, J. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997,** *18,* 1463–1472.

(65) Laurent, B.; Chavent, M.; Cragnolini, T.; Dahl, A.C.E.; Pasquali, S.; Derreumaux, P.; Sansom, M.S.P.; Baaden, M. Epock: rapid analysis of protein pocket dynamics. *Bioinformatics.* **2015,** *31,* 1478–1480.

(66) Kumari, R.; Kumar, R.; Lynn, A. g_mmpbsa–A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Comp. Sci.* **2014,** *54,* 1951–1962.

(67) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W.; Mark, A.E. Peptide folding: when simulation meets experiment. *Angew. Chem. Int. Ed.* **1999,** *38,* 236–240.