

A novel approach for selecting potent peptide inhibitors of the SARS-CoV2-Mpro protease

Alexander Pisarchik (✉ a.pisarchik@alcastem.com)

Alcamena Stem Cell Therapeutics

Research Article

Keywords: peptide inhibitors, cytotoxic protein, SARS-CoV2-Mpro protease

Posted Date: April 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-396336/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A novel approach for selecting potent peptide inhibitors of the SARS-CoV2-M^{pro} protease

Alexander Pisarchik^{1*}

¹ Alcamena Stem Cell Therapeutics

1450 S. Rolling Rd., Suite 4.069, Halethorpe MD 21227

* Corresponding Author

a.pisarchik@alcastem.com

Abstract

Here, we describe a new high throughput selection technology for identifying exceedingly specific and effective peptide inhibitors. This technology incorporates the co-expression of a cytotoxic protein and a library of peptide variants inserted directly into a loop of a carrier protein. Selection is based on the cytotoxicity neutralization by a member of a peptide library binding to and inhibiting the cytotoxic protein. Our technology provides the flexibility of screening both cyclic and linear peptides. Herein, we demonstrate the power of this technology by developing selective inhibitors of the main coronavirus protease (M^{pro}) in a matter of weeks by screening libraries of cyclic and linear peptides. This technology opens up an opportunity to develop inhibitors for a wide range of previously undruggable targets.

Modern day drug discovery has focused on the development of small molecule therapeutics. While small molecules offer many advantages, such as economical manufacturing, lower complexity and better bioavailability as compared to legacy drugs, they can only target 2-5% of the proteome^{1,2}. Biologic-based drugs have a larger binding surface and therefore a higher target specificity, allowing them to access targets that are beyond the reach of small molecules. However, most biologics are large molecules that cannot cross cell membranes, which restricts their use to extracellular targets. Peptide drugs, on the other hand, have advantages of both small molecule therapeutics and biologic drugs, but do not have many of their disadvantages. Like biologic-based drugs, peptides have a large binding surface to target leading to their higher specificity and fewer off-target effects^{3,4}. Similar to small molecules, they are smaller, have lower immunogenicity^{5,6} and higher bioavailability. Recent advances in cell penetrating peptide technology have enabled peptide drugs to be designed to access intracellular targets^{7,8}. Peptide drugs can therefore achieve the level of bioavailability comparable to that of small molecule therapeutics and activity and safety of biologic-based drugs which makes them prime candidates of drug development for previously “undruggable” targets.

The development of therapeutic peptides commonly starts with a combinatorial biology approach that involves the generation of chemical or biosynthetic peptide libraries. Chemical peptide synthesis is a well-established method for developing peptide libraries^{9,10}; however, the biosynthetic approach offers many advantages. One key advantage is the library size. Biosynthetic libraries can easily contain as many as 10^9 peptides, while chemical synthesis is limited to approximately 10^4 peptides. The most commonly used biosynthetic selection methods are phage display¹¹, yeast display¹² and RNA display^{13,14}. All of these methods select peptides

that bind to the target protein most tightly. However, a major limitation to these approaches is that the best binders may not be the best inhibitors of the target protein.

One way to solve this problem is to establish a link between binding and function by screening peptides intracellularly for their ability to attenuate or inhibit cellular processes. None of the existing cell-based assays has taken full advantage of this approach. Currently, the most promising *in vivo* peptide selection method, called split-intein circular ligation of peptides and proteins (SICLOPPS), is based on protein trans-splicing. This involves self-excision of an internal protein segment (intein) resulting in a cyclized polypeptide¹⁵. Typically, such libraries are screened in *E. coli* cells using bacterial two-hybrid system. Selection relies on disruption of a targeted protein-protein interaction (not function), detected through a reporter gene expression¹⁶. False positive clones often result due to fluctuations of gene expression, mutations in the regulatory sequences and mutations in the bacterial genome. Additionally, construct design for these peptide “processing” enzymes (inteins) is complex, they mostly work in a reduced environment¹⁷, and are often slow¹⁸.

To solve this problem, we tested a new selection system based on direct inhibition of a cytotoxic protein that is not dependent on transcription of a reporter gene (Fig.1). Our peptides mimic cyclization by insertion into a protein loop, thus avoiding the need for any processing enzymes (like inteins). This gives the flexibility of screening both cyclic and linear peptides, which further increases the library size and improves the chances for identification of the optimal peptide inhibitor. As a proof-of-concept for this new approach, we performed the selections for a small pool of peptides (10^6 variants) that consisted of cyclic and linear peptide inhibitors targeting main coronavirus protease (M^{pro}). Within five weeks, we identified an inhibitor with an IC_{50} of 33 μ M, validating this screening approach.

Results

Selection System. The selection system relies on the toxicity of a particular protein to its host (Fig.1). A peptide variant is co-expressed with the cytotoxic target protein in the host cell. The host cell only survives if a peptide variant binds to the protein, and neutralizes its cytotoxicity. Using this approach, we developed inhibitors of the difficult-to-target SARS-CoV2 main protease (M^{pro})¹⁹ that have the potential to be developed into a pancoronavirus antiviral drugs, for which there is an urgent unmet medical need, due to the regular frequency of coronavirus-caused pandemics that have occurred over the last century.

Libraries. Our constructs are presented in Fig.2. The peptide libraries were inserted into ubiquitin as a carrier protein, because it is small (8.6kD), stable in *Ecoli*, and has previously been used to express proteins and peptides^{20,21}. The first peptide library was random, built with 14 degenerate codons, resulting in 1.6×10^{18} variants. The second library was based on published sequences²² recognized by M^{pro} , and contained 2×10^9 variants. These libraries were inserted into the first loop of ubiquitin, which was selected because it had been shown to tolerate insertions²³. Variant sequences with no stop codons or frame-shifts were fully integrated into the loop of the full-length ubiquitin protein and served as a model of cyclic peptides. Variants with stop codons produced linear peptides attached to the first beta strand of ubiquitin.

Peptide evolution. Both peptide libraries were cloned in the pUbi-Mpro construct (Fig. 2) and taken through five rounds of selection in *Ecoli*. We screened 1 million clones at each round. To weed out false positives that may result from frame-shifts, deletions of M^{pro} and somatic mutations libraries were re-cloned into the original vector (pUbi-Mpro) after each round of selection. The fifth round of selection generated several sequences that were significantly overrepresented in the population. We chose the 11 most abundant peptides for further testing.

Seven of these peptides were linear, 4 were fully integrated in the loop of the carrier protein and, therefore, cyclic (Supplementary Table 1). The most abundant peptides were synthesized in a linear form and tested in an *in vitro* M^{pro} activity assay. Out of 11 peptides tested, 7 did not have any effect on protease activity (false positives). The other 4 peptides inhibited the M^{pro} with IC₅₀ ranging from 100 μ M to 1.2 mM (Supplementary Table 1). The two best peptides were M1 (RQGLDEDLHRW) and M5 (TANAFLS). Their IC₅₀ was 249 and 101 μ M, respectively (Table 1). Peptide M1 originated from the random library and peptide M5 originated from the library based on the published sequences that are recognized by M^{pro}. To be consistent with the structure in the original screen, we also synthesized peptide M5 in a cyclic form (peptide M5c) fused to a custom cell penetration sequence to improve its stability and intracellular transport. Cyclization improved IC₅₀ of M5 peptide significantly from 101 to 33 μ M (Table 1).

Following first 5 rounds of selection, peptides M1 and M5 were further mutagenized by PCR with degenerate primers, cloned in pUbi-Mpro-CAT vector and selected for 5 more rounds. pUbi-Mpro-CAT was identical to pUbi-Mpro except that M^{pro} coding sequence was fused to the *chloramphenicol acetyl transferase* (CAT) gene. This modification made selection more stringent since cell survival now required not only the presence of the peptide inhibitor but also intact CAT protein which prevented bacteria from deleting M^{pro}. In this round of selection there were no false positives, showing that the addition of the CAT improved screening effectiveness. The largest fraction of the resulting sequences was made of peptides M74 and M78 (Table 1). These peptide variants were synthesized and tested in the *in vitro* assay (Table 1). They did not show any improvement in the standard conditions (IC₅₀ was 195 and 244 μ M, respectively). However, under denaturing conditions (heat shock at 53°C) which we used to model partial protein unfolding, the IC₅₀ of peptide M74 improved 10-fold (21.2 μ M, Table 1).

Discussion

A significant disadvantage of current display technologies (eg. Phage display, RNA display, yeast display) is the lack of a connection between binding and function. That means that a peptide binding to the target protein may not necessarily inhibit its enzymatic activity or disrupt a protein-protein interaction. We solved this problem by developing a screen based on the cytotoxicity of the target protein. Other *in vivo* selection methods have relied on the toxicity of an enzyme's (target protein) substrate²⁴, products of the enzymatic reaction²⁵, a particular intermediate²⁶, or resistance to inhibitors²⁷. Our selection approach is the first to capitalize on the cytotoxicity of the target protein itself.

The problem of protein toxicity is widespread in the field of protein expression. Usually, it is a problem that has to be minimized. Our technology leverages this “problem” by screening peptides for their ability to neutralizing the cytotoxicity of the target protein. It involves co-expression of the cytotoxic target protein and a library of peptide variants. Host cells survive only when a particular peptide variant inhibits the cytotoxic protein (Fig. 1).

To demonstrate the power of this technology, we targeted the coronavirus M^{pro} protease, that is highly conserved among various coronaviruses and plays a pivotal role in their life cycles²⁸. Mutations in M^{pro} are often lethal to the virus which is why drugs targeting the M^{pro} enzyme have the potential to significantly reduce the risk of mutation-mediated drug resistance and display broad-spectrum antiviral activity¹⁹. To date, no M^{pro} targeted antiviral has been developed. Repurposing of antiviral drugs for other viruses^{29,30} have not proved effective. Drug development approaches based on converting peptides into peptidomimetics are very challenging, because side chain modifications often abolish inhibitor activity³¹ and result in off-

target effects³² and toxicity³³. Our approach avoids these pitfalls because our peptides are selected intracellularly biasing libraries towards candidates with lower toxicity, and higher stability.

Two best peptides (M1 and M5) generated by our selection system showed inhibitory activity in the low micromolar range in an *in vitro* assay (Table 1), demonstrating the utility of our selection approach. Peptide M1 was selected from the random library and peptide M5 from the candidate-based library. This observation demonstrates that our approach can identify inhibitors without prior knowledge of their ligands and can improve the inhibitory activity of known ligands. It is also important to note that peptide M5 is fully integrated in the first loop of the carrier protein (ubiquitin) which gives it a cyclic structure. Consistent with this observation, when peptide M5 was cyclized, its IC₅₀ improved significantly from 101 to 33 μ M (Table 1) which confirms that our technology is useful for screening both linear and cyclic peptides.

Additional rounds of evolution (rounds 6-10) identified peptides M74 and M78. Interestingly, these peptides were not as potent under native conditions (IC₅₀ were at 238 and 195 μ M respectively (Table 1); however, under denaturing conditions (heat shock at 53°C) the IC₅₀ of the peptide 74 improved 10-fold. We speculate that M74 may inhibit M^{Pro} during protein folding. This type of inhibitors was previously identified for HIV protease³⁴ and chicken egg lysozyme³⁵. It is believed these inhibitors interfere with protein folding by binding to certain protein sequences called local elementary structures (LES). These local elementary structures are believed to be part of protein's hydrophobic core which can only be exposed during partial protein denaturation. Not surprisingly, M74 was the only peptide with a stretch of hydrophobic amino acids (LAVVAL) at its C terminus, which have the potential to interact with the hydrophobic core of partially unfolded protease.

As for peptide (M78), it had 2 cysteines which are highly reactive groups that have the potential to form disulphide bonds not only with the cysteine in the active site of M^{pro} protease but also with each-other in a matter of minutes³⁶. Therefore, we suspect that the life of this peptide inhibitor is very short and we did not use the best conditions to detect its activity

A weakness in our study is that we only screened a small fraction of all available peptides in our library (1 million clones at each stage of selection). Despite this shortcoming, we were able to rapidly identify (after a few weeks of screening) potent peptide inhibitors with low μM activity from both our random and candidate base peptide libraries (Table 1), demonstrating the functional utility of our approach.

Methods

Constructs. All genes were codon optimized, synthesized as gBlocks by IDT and cloned into the pBAD-HisA plasmid (Thermo Fisher Scientific). pBAD-HisA plasmid was amplified with primers P33 and P34 (Supplementary Table 2) introducing *HindIII* and *XhoI* restriction sites. Each PCR reaction (20 μ l) contained 20 ng of DNA template and 50 pmoles of each primer mixed with 10 μ l of Pfu Ultra II Hotstart 2x Mastermix (Agilent). The PCR reaction (20 μ L) was initially heated at 95°C for 2.5 min followed by 30 cycles of denaturation at 94°C for 15 sec., annealing at 55°C for 15 sec. and extension at 72°C for 6 min. Following amplification, the PCR fragment was gel-purified by the QIAGEN gel-band purification kit and digested with *HindIII* and *XhoI* restriction enzymes.

The pUbi-Mpro constructs (Fig.2.) use an arabinose-inducible promoter, which is also an operon (Fig.2) to express M^{pro}, ubiquitin, or ubiquitin-CAT. A Shine-Dalgarno sequence is inserted between M^{pro} and Ubiquitin to ensures the expression of both genes. The ubiquitin gene was synthesized by IDT and amplified by primers T227 and T228 (Supplementary Table 2). The amplified fragment was cut with *HindIII* and *PacI* restriction enzymes. PCR conditions are described above, with the exception, that extension was done for 1 minute. The M^{pro}-GST fusion was amplified with primers T229 and P108 (Supplementary Table 2) and cut with *PacI* and *XhoI* restriction sites as described above. Following amplification, both PCR fragments were gel-purified using the QIAGEN gel-band purification kit. Then, both fragments were ligated with pBAD backbone (cut with *HindIII* and *XhoI*). The ligation reaction (20 μ l) contained 2 μ l of 10x ligation buffer, 100 ng of each fragment and 1 μ l of T4 DNA ligase (NEB cat# M0202S). The reaction mix was incubated for 1 h at room temperature. Ligated fragments were transformed in 10G chemically competent cells (Lucigen) according to manufacturer's protocol. Transformed

cells were plated on LB plates containing 50 µg/mL ampicillin and incubated overnight at 37°C. The insert was confirmed with colony PCR. This involved resuspending a colony in 20 µl of sterile 0.9% sodium chloride solution. One µl of this solution was transferred to the PCR tube and amplified with Taq polymerase (New England Biolabs, cat# M0482S) and 30 pM of the flanking primers. Each PCR reaction (20 µL) was initially heated to 95°C for 2.5 min followed by 30 cycles of denaturation at 94°C for 20 sec., annealed at 55°C for 20 sec., and extended at 72°C for 1 min. Amplification products were visualized by agarose electrophoresis. Clones with the correct inserts were inoculated in culture tubes containing 5 mL of LB with the appropriate antibiotic and incubated overnight at 37°C. Constructs were then purified using the Monarch Plasmid miniprep kit (NEB) and sequenced.

In the pUbi-M^{pro}-CAT construct, ubiquitin and M^{pro}-CAT were constructed by extension PCR. First, the ubiquitin gene was amplified with primers T227 and T336 and with M^{pro} primers T337 and T333 (Supplementary Table 2). The *CAT* gene was synthesized by IDT and amplified with primers T334 and T335 (Supplementary Table 2). All fragments were gel-purified with the QIAGEN gel-band purification kit. Then all fragments were mixed together, amplified with primers T227 and T335, gel-purified with QIAGEN gel-band purification kit and cut with *HindIII* and *XhoI* restriction enzymes. Finally, the construct was ligated with pBAD backbone (also cut with *HindIII* and *XhoI*), tested and sequenced as described above.

Construction of Peptide libraries. Random and candidate libraries of the M^{pro}-inhibitor peptides were cloned into the first loop of ubiquitin (in the pUbi-M^{pro} constructs), which was shown previously to be tolerant to insertions and deletions²³.

The random library, was built with 14 NNK codons and amplified as two fragments which were united by ligation. The first fragment was amplified with flanking forward primer P23 (Supplementary Table 2) and the reverse primer T232 (Supplementary Table 3). The second fragment was amplified with primer T233 (Supplementary Table 3) and the reverse primer P24 (Supplementary Table 2). The PCR reaction (20 μ L) was initially heated at 95°C for 2.5 min followed by 30 cycles of denaturation at 94°C for 15 sec., annealing at 55°C for 15 sec. and extension at 72°C for 40 sec. Following amplification, PCR fragments were gel-purified by the QIAGEN gel-band purification kit and mixed and ligated with T4 DNA ligase. The ligation reaction contained 20 μ l of 10x ligation buffer, 100 ng of fragment mix, 0.5 μ l of 100 mM ATP, 1 μ l of T4 DNA ligase (NEB cat# M0202S) and 1 μ l of T4 polynucleotide kinase. The reaction mix was incubated at room temperature and used as a template for PCR with flanking primers P276 and P277 (Supplementary Table 2) using program described above. The PCR fragment was gel-purified by the QIAGEN gel-band purification kit and cut with *KasI* and *XbaI* restriction enzymes, purified by QIAGEN kit and ligated with pUbi_Mpro also digested by *KasI* and *XbaI*. Ligated fragments were transformed in 10G electrocompetent cells according to the recommendations of Bio-Rad. Transformed cells were plated on LB plates containing 50 μ g/mL ampicillin and incubated overnight at 37°C. Individual clones were sequenced verified by GeneWiz.

The M^{Pro} candidate library was based on sequences recognized by M^{Pro}²². These sequences were mutagenized by degenerate synthetic oligonucleotides. Library size was controlled by targeting mutations to one position in each codon with only first or the second codon position being changed. Library construction was done as described above with the following differences: the first fragment was amplified with flanking forward primer P23

(Supplementary Table 2) and one of the reverse primers (Supplementary Table 3, primers T234 through T255). The right fragment was amplified with primer T233 (Supplementary Table 2) and one of the forward primers (Supplementary Table 3, primers T256 through T277). Following amplification, PCR fragments were gel-purified, mixed and ligated with T4 DNA ligase. The full-length PCR band was amplified with the flanking primers P276 and P277 (Supplementary Table 2), digested with *KasI* and *XbaI* restriction enzymes and ligated with pUbi-Mpro as described above.

Selection. Libraries were constructed as described above and desalted by dialysis as follows: 20 ml of Ultrapure water was poured into Petri dishes. 0.025 μ M VSWP Membrane Filters were placed on top of the water. 10 μ L drops of the DNA were pipetted on filters and incubated for 30 min. Following incubation DNA concentration was measured. Libraries were transformed into 10G strain of *E.coli* using electroporation. To make electrocompetent cells one colony from a freshly streaked plate of the *E. coli* was inoculated in 5 mL of LB and grown overnight at 37°C with shaking at 250 rpm overnight. Then 2 mL of the overnight *E. coli* culture was transferred to two 500 ml flasks containing 250 ml LB media each. Flasks were shaken at 250 rpm and incubated at 37°C until OD600 reached 0.4. Cultures were centrifuged at 4000xg at 4°C for 15 min. The supernatant was removed, cells were resuspended in 100 ml of sterile ice-cold water and centrifuged at 4000xg at 4°C for 15 min. This step was repeated a second time and then the cells were resuspended in 50 ml of ice-cold water and centrifuged at 4000xg at 4°C for 15 min. After repeating the step for the third time, cells were resuspended in 20 ml of sterile ice-cold 10% glycerol and centrifuged at 3000xg at 4°C for 15 min. Supernatant was discarded and cells were resuspended in 1 ml of 10% glycerol. Micropulser electroporator (Bio-Rad) and 0.1 mm cuvettes were used for electroporation according to the manufacturer's

recommendations. Following electroporation, 1 ml of SOC media was added to each transformation. Tubes were shaken at 37°C and 250 rpm for 1 h and added to the 500 ml flasks containing 250 ml LB media with 100 µg/mL carbenicillin and 0.04% arabinose and incubated overnight at 37°C with shaking.

Following selection, libraries were re-cloned into the original vector pUbi-M^{pro} vector to remove false positive sequences (clones that grow in the media but didn't express M^{pro} protease). For this purpose, 3 ml of cell culture was used for plasmid isolation using Qiagen Miniprep Kit. Libraries were amplified from plasmid populations using primers P276 and P277 (Supplementary Table 2). Each PCR reaction (20 µl) contained 20 ng of DNA template and 50 pmoles of each primer mixed with 10 µl of Pfu Ultra II Hotstart 2x Mastermix (Agilent). The PCR reaction (20 µL) was initially heated to 95°C for 2.5 min followed by 30 cycles of denaturation at 94°C for 15 sec., annealing at 55°C for 15 sec. and extension was at 72°C for 40 sec. Following amplification, the PCR fragment was gel-purified by the QIAGEN gel-band purification kit and used for Next Generation Sequencing (NGS) and re-cloning in the original vector and electroporated into 10G cells as described above.

Sequence Analysis. Sequences of individual clones were analyzed by Next Generation Sequencing (NGS) and Sanger Sequencing. NGS was done by submitting PCR reactions from each selection cycle to GeneWiz for Amplicon-EZ service. Libraries were also streaked on Petri dishes with LB agar containing appropriate antibiotics and incubated overnight at 37°C. Between 10 and 20 bacterial colonies were submitted for Sanger sequencing to GeneWiz. The sequences of the most abundant peptides identified by NGS and confirmed by Sanger sequencing, were selected for further testing.

Fusion Peptide Libraries. Sequences corresponding to the best two peptides (M1 and M5) were mutagenized in order to increase diversity. Two DNA fragments were amplified: left and right, which were reunited by ligation (Fig. 3) as described above for other libraries. The first fragment was amplified with the flanking forward primer P23 (Supplementary Table 2) and the reverse primers (Supplementary Table 3, primers T370 through T396 for peptide M5 or T422 through T444 for peptide M1). The second fragment was amplified with reverse primer P24 and forward primers (Supplementary Table 3, primers T343 through T369 for peptide M5 or T397 through T421 for peptide M1). Fragments were processed, inserted into pUbi-M^{Pro}-CAT and transformed into 10G cells as described above for the secondary screen. Following selection libraries were amplified using primers P276 and P277 and re-inserted into pBAD_Ubi-M^{Pro}-CAT as described above. This selection process was repeated 5 times. Selected sequences were analyzed by NGS and Sanger Sequencing as described above.

Inhibition of the M^{Pro} Protease activity by the peptides *in vitro*. Peptides were synthesized by Elim Biopharmaceuticals and purified to 95%. Inhibitory activity of these peptides on M^{Pro} was tested using 3CL Protease Kit from BPS Bioscience (Catalog # 78042-1) according to the manufacturer's recommendations. Briefly, 30 μ l of 3CL Protease enzyme solution (0.05 ng/ μ l) was mixed with 10 μ l of peptides at different concentrations and preincubated for 30 min at room temperature. Following preincubation, 10 μ l of 200 μ M 3CL Protease substrate was added and incubated for 4 h at room temperature. The fluorescence intensity was measured in a microtiter plate-reading fluorimeter with excitation at 360 nm and emission at 460 nm. The heat shock version of this protocol included incubation of the protease and peptide mixture at 53°C for 5 min followed by incubation for 4 h at room temperature. The

heat shock treatment led to a 35% loss of enzymatic activity in the positive control (enzyme without inhibitor). See Supplementary Fig.1 for details.

Acknowledgements (optional)

I thank Leon Nesti and Noreen Gervasi for their help writing and editing the manuscript.

Author Contributions

A.P. conceived the project, designed and performed the experiments, analyzed data and wrote the manuscript

Competing Interests

The author has ownership in Alcamena.

References

1. Hopkins, A. L. & Groom, C. R. The druggable genome. *Nat. Rev. Drug Discov.* **1**, 727–730 (2002).
2. Drews, J. Drug discovery: A historical perspective. *Science* vol. 287 1960–1964 (2000).
3. Gorr, S. U., Flory, C. M. & Schumacher, R. J. In vivo activity and low toxicity of the second-generation antimicrobial peptide DGL13K. *PLoS One* **14**, (2019).
4. Craik, D. J., Fairlie, D. P., Liras, S. & Price, D. The Future of Peptide-based Drugs. *Chem. Biol. Drug Des.* **81**, 136–147 (2013).
5. Mariani, M. *et al.* Immunogenicity of a free synthetic peptide: Carrier-conjugation enhances antibody affinity for the native protein. *Mol. Immunol.* **24**, 297–303 (1987).
6. Van Regenmortel, M. H. V. Antigenicity and immunogenicity of synthetic peptides. in *Biologicals* vol. 29 209–213 (Academic Press, 2001).
7. Walrant, A., Cardon, S., Burlina, F. & Sagan, S. Membrane Crossing and Membranotropic Activity of Cell-Penetrating Peptides: Dangerous Liaisons? *Acc. Chem. Res.* **50**, 2968–2975 (2017).
8. Dougherty, P. G., Sahni, A. & Pei, D. Understanding Cell Penetration of Cyclic Peptides. *Chemical Reviews* vol. 119 10241–10287 (2019).
9. Marasco, D., Perretta, G., Sabatella, M. & Ruvo, M. Past and Future Perspectives of Synthetic Peptide Libraries. *Curr. Protein Pept. Sci.* **9**, 447–467 (2008).
10. Lam, K. S. *et al.* A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **354**, 82–84 (1991).
11. Smith, G. P. Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science (80-.)*. **228**, 1315–1317 (1985).

12. Bowen, J. *et al.* Screening of yeast display libraries of enzymatically treated peptides to discover macrocyclic peptide ligands. *Int. J. Mol. Sci.* **22**, 1–20 (2021).
13. Nemoto, N., Miyamoto-Sato, E., Husimi, Y. & Yanagawa, H. In vitro virus: Bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett.* **414**, 405–408 (1997).
14. Roberts, R. W. & Szostak, J. W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12297–12302 (1997).
15. Tavassoli, A. & Benkovic, S. J. Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in *E. coli*. *Nat. Protoc.* **2**, 1126–1133 (2007).
16. Tavassoli, A. *et al.* Inhibition of HIV budding by a genetically selected cyclic peptide targeting the Gag-TSG101 interaction. *ACS Chem. Biol.* **3**, 757–764 (2008).
17. Bhagawati, M. *et al.* A mesophilic cysteine-less split intein for protein trans-splicing applications under oxidizing conditions. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 22164–22172 (2019).
18. Aranko, A. S., Wlodawer, A. & Iwaï, H. Nature's recipe for splitting inteins. *Protein Engineering, Design and Selection* vol. 27 263–271 (2014).
19. Goyal, B. & Goyal, D. Targeting the Dimerization of the Main Protease of Coronaviruses: A Potential Broad-Spectrum Therapeutic Strategy. *ACS Combinatorial Science* vol. 22 297–305 (2020).
20. Baker, R. T. Protein expression using ubiquitin fusion and cleavage. *Curr. Opin. Biotechnol.* **7**, 541–546 (1996).
21. Yoo, Y., Rote, K. & Rechsteiner, M. Synthesis of peptides as cloned ubiquitin extensions. *J. Biol. Chem.* **264**, 17078–17083 (1989).

22. Muramatsu, T. *et al.* SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12997–13002 (2016).
23. Ferraro, D. M., Hope, E. K. & Robertson, A. D. Site-specific reflex response of ubiquitin to loop insertions. *J. Mol. Biol.* **352**, 575–584 (2005).
24. Jiang, P. *et al.* Design and application of a novel high-throughput screening technique for 1-deoxynojirimycin. *Sci. Rep.* **5**, (2015).
25. McLoughlin, S. Y., Jackson, C., Liu, J. W. & Ollis, D. Increased expression of a bacterial phosphotriesterase in *Escherichia coli* through directed evolution. *Protein Expr. Purif.* **41**, 433–440 (2005).
26. Boersma, Y. L. *et al.* A novel genetic selection system for improved enantioselectivity of *Bacillus subtilis* lipase A. *ChemBioChem* **9**, 1110–1115 (2008).
27. Dickinson, B. C., Packer, M. S., Badran, A. H. & Liu, D. R. A system for the continuous directed evolution of proteases rapidly reveals drug-resistance mutations. *Nat. Commun.* **5**, (2014).
28. Ziebuhr, J. The coronavirus replicase. *Current Topics in Microbiology and Immunology* vol. 287 57–94 (2005).
29. Shie, J. J. *et al.* Inhibition of the severe acute respiratory syndrome 3CL protease by peptidomimetic α,β -unsaturated esters. *Bioorganic Med. Chem.* **13**, 5240–5252 (2005).
30. Jenwitheesuk, E. & Samudrala, R. Identifying inhibitors of the SARS coronavirus proteinase. *Bioorganic Med. Chem. Lett.* **13**, 3989–3992 (2003).
31. Kang, C. B. *et al.* Exploring the binding of peptidic West Nile virus NS2B-NS3 protease inhibitors by NMR. *Antiviral Res.* **97**, 137–144 (2013).
32. Yang, H. *et al.* Design of wide-spectrum inhibitors targeting coronavirus main proteases.

- PLoS Biol.* **3**, (2005).
33. 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. *Journal of Medicinal Chemistry* vol. 59 6595–6628 (2016).
 34. Broglia, R. A. *et al.* A folding inhibitor of the HIV-1 protease. *Proteins Struct. Funct. Genet.* **62**, 928–933 (2006).
 35. Caldarini, M. *et al.* Identification and characterization of folding inhibitors of hen egg lysozyme: An example of a new paradigm of drug design. *Proteins Struct. Funct. Bioinforma.* **74**, 390–399 (2009).
 36. Landgraf, B. J., Ren, G., Masuch, T., Boyd, D. & Berkmen, M. From Biology to Biotechnology: Disulfide Bond Formation in *Escherichia coli*. in *Escherichia coli - Recent Advances on Physiology, Pathogenesis and Biotechnological Applications* (InTech, 2017). doi:10.5772/67393.

Figure legends

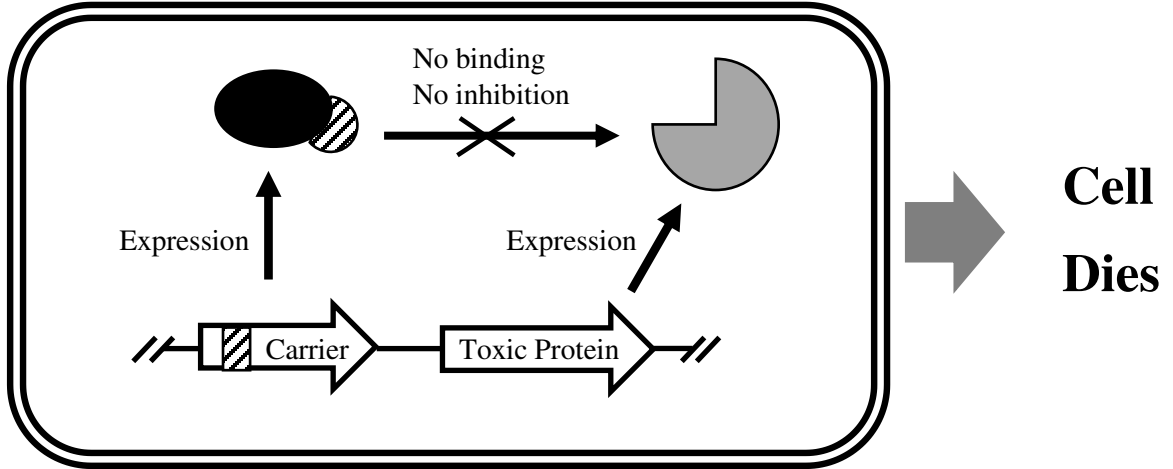
Fig. 1 Selection System. Toxic protein is co-expressed with a library of peptides. Peptide variants are inserted into a carrier protein. A) When a peptide does not inhibit the toxic protein, host cell dies. B) If a candidate peptide binds to and inhibits the target protein, its cytotoxicity is neutralized and host cells survives.

Fig. 2 Expression constructs. M^{pro} and Ubiquitin expression is controlled by the arabinose-inducible promoter. A) The ubiquitin and M^{pro} -GST fusion expression construct. B) The ubiquitin and M^{pro} -CAT (*chloramphenicol-acetyl-transferase*) fusion expression construct.

Fig. 3 Construction of the Fusion Library. DNA sequences coding for the M1 or M5 peptides are mutated the following way: A) The sequences are amplified by degenerate oligos: B). Mutated fragments are ligated using T4 DNA ligase and T4 polynucleotide kinase.

Figure 1

A. Peptides are not binding to the toxic protein



B. Peptide is binding to the toxic protein

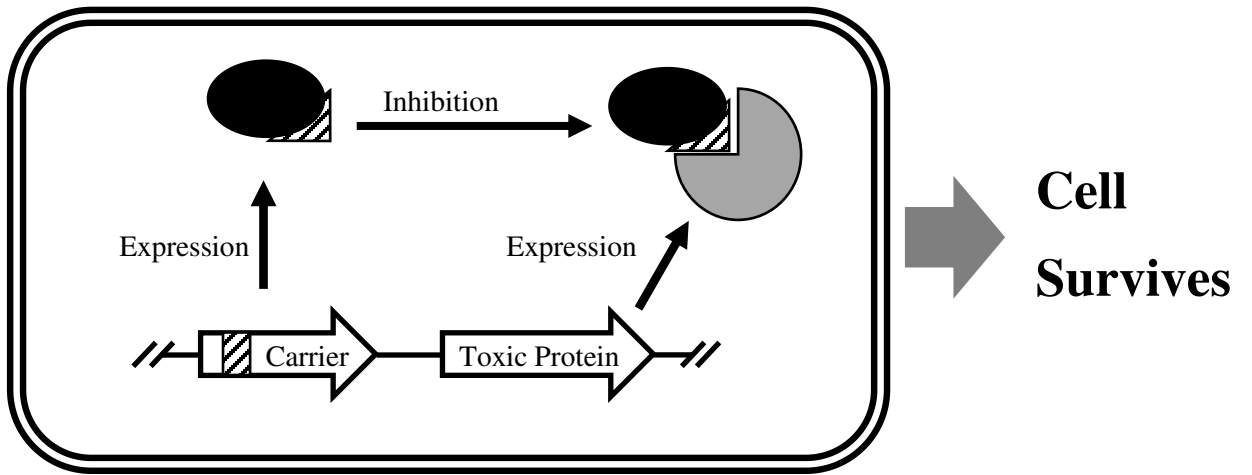
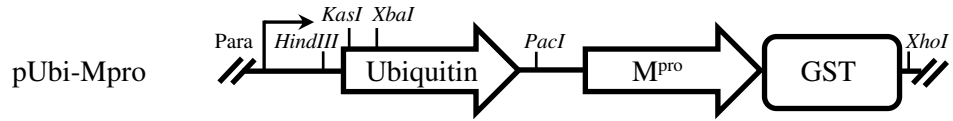


Figure 2

A)



B)

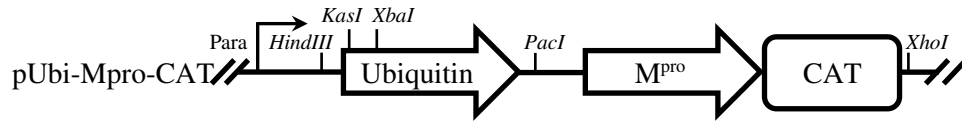


Figure 3

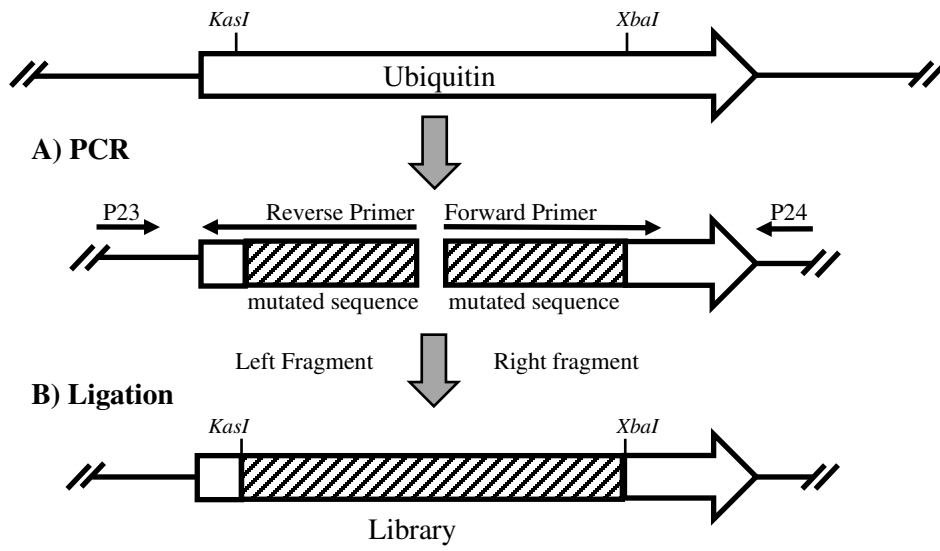


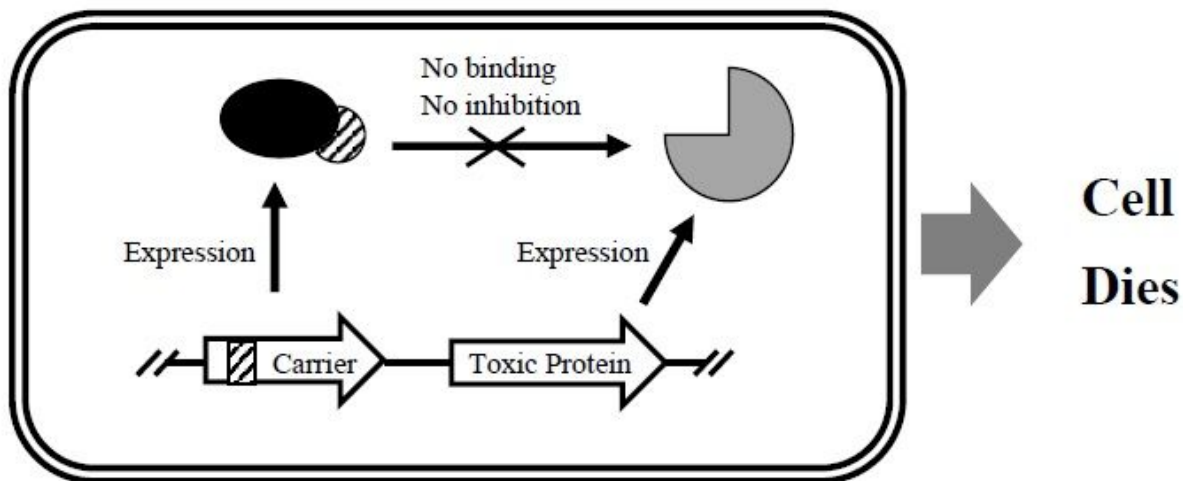
Table 1

M^{pro} Inhibition Assay of the best peptides

Peptide Name	Peptide Sequence		IC ₅₀ (μM)	
			Normal Conditions	Heated at 53C
<i>Peptides from rounds 1-5</i>				
M1	<u>GARQGLDEDLHRW</u>	linear	249.4±47.4	255.0±49.5
M5	<u>GATANAFLSGSGSRG</u>	linear	101.7±16.8	100±6.7
M5c	<u>WRRWWRRRRTANAFLS</u>	cyclic	33.9±8.3	29.1±1.3
<i>Peptides from rounds 6-10</i>				
M74	<u>GIASTMVKPILGLAVVAL</u>	linear	238.4±9.1	21.2±12.0
M78	<u>CKCMDKNPYSRYPG</u>	linear	195.1±27.2	198.5±13.0

Figures

A. Peptides are not binding to the toxic protein



B. Peptide is binding to the toxic protein

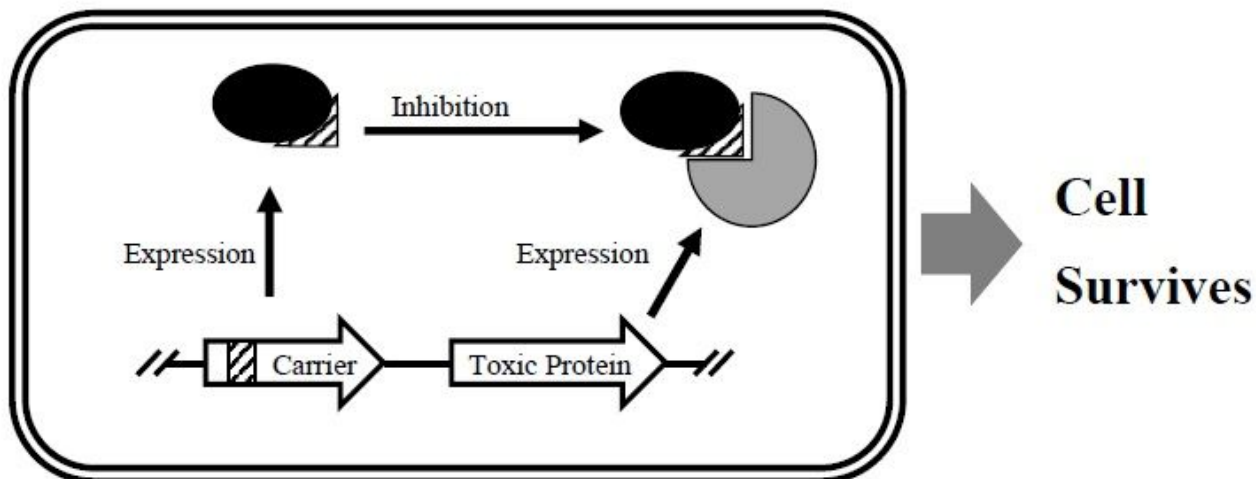


Figure 1

Selection System. Toxic protein is co-expressed with a library of peptides. Peptide variants are inserted into a carrier protein. A) When a peptide does not inhibit the toxic protein, host cell dies. B) If a candidate peptide binds to and inhibits the target protein, its cytotoxicity is neutralized and host cells survives.

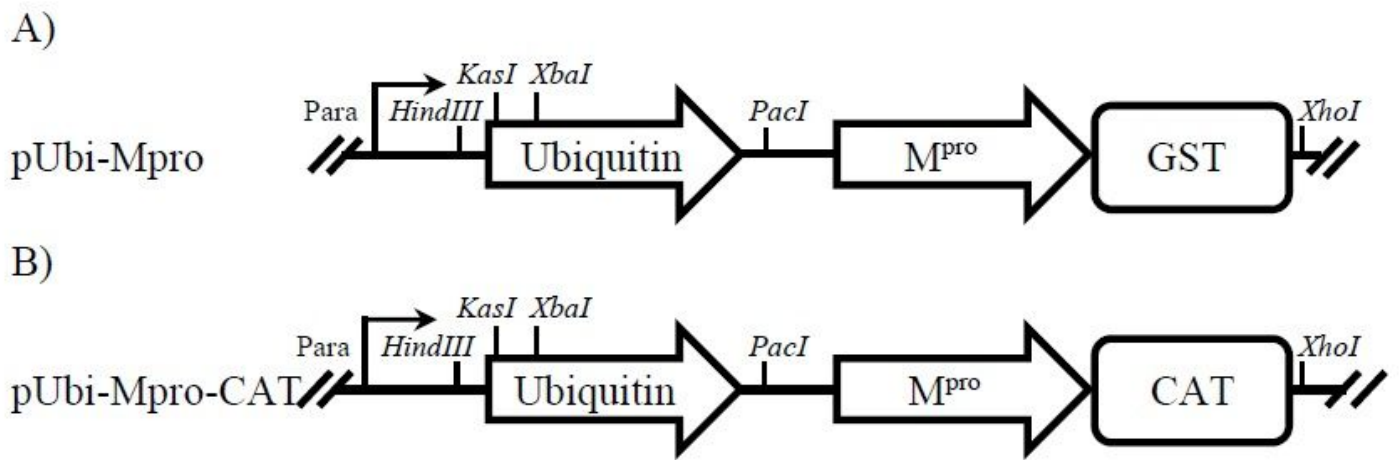


Figure 2

Expression constructs. Mpro and Ubiquitin expression is controlled by the arabinose-inducible promoter. A) The ubiquitin and Mpro-GST fusion expression construct. B) The ubiquitin and Mpro-CAT (chloramphenicol-acetyl-transferase) fusion expression construct.

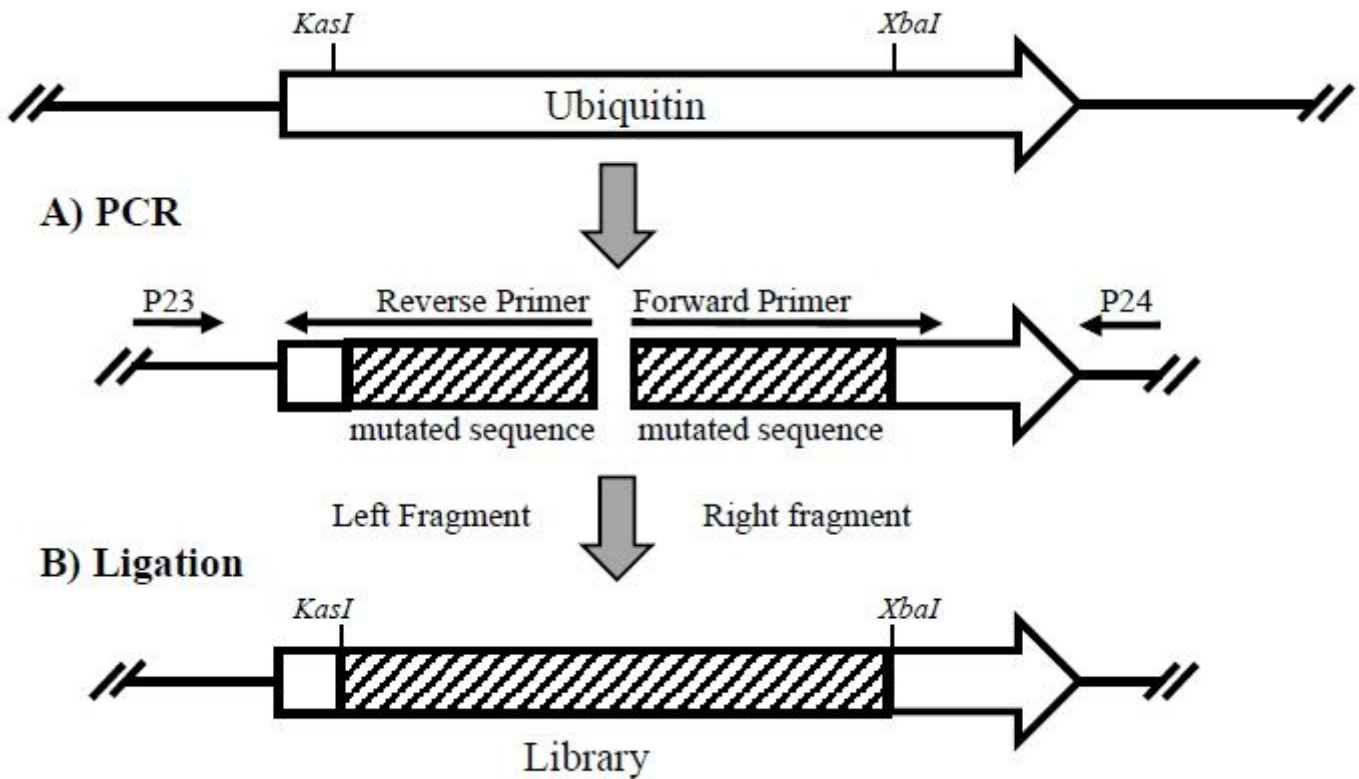


Figure 3

Construction of the Fusion Library. DNA sequences coding for the M1 or M5 peptides are mutated the following way: A) The sequences are amplified by degenerate oligos: B). Mutated fragments are ligated using T4 DNA ligase and T4 polynucleotide kinase.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [20210414SupplementaryInfofile.pdf](#)