Proceedings of the 36th European Peptide Symposium
Michal Lebl (Editor)
European Peptide Society, 2022 *https://doi.org/10.17952/36EPS.2022.291*

A Novel Selection Technology Identifies Potent Inhibitor Peptides Against 3CL Protease of SARS-Cov-2 Coronavirus

Alexander Pisarchik and Edmund Nesti
1450 S. Rolling Rd., Suite 4.069, Halethorpe MD 21227, USA

Introduction

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 other drug modalities (e.g. antibodies and nucleic acid based therapeutics), they can only target 2-5% of the proteome [1,2]. Biologic-based drugs (e.g. antibodies) have a larger binding surface and therefore a higher target specificity, allowing them to access targets that are beyond the reach of small molecules and have fewer off-target effects. 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]. Similar to small molecules, they are smaller, have lower immunogenicity [4] and higher bioavailability. Recent advances in cell penetrating peptide technology have enabled peptide drugs to be designed to access intracellular targets [5]. 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 for "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 [6]; 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 approximately10^4 peptides. The most commonly used biosynthetic selection methods are phage display [7], yeast display [8] and mRNA display [9]. 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, because binding does not always occur in the active region of the 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 polypeptide [10]. Typically, such libraries are screened in *E.coli* cells using bacterial two-hybrid system. Selection relies on disruption of a targeted protein-protein interaction, detected through a reporter gene expression. 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, generally restricted to a reduced environment [11], and are time consuming [12].

To solve this problem, we developed a new selection system based on direct inhibition of a cytotoxic protein (Figure 1). 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 constrained 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 constrained and linear peptide inhibitors targeting main coronavirus protease (3CL). Within five weeks, we identified an inhibitor with an IC_{50} of 33 μM, validating this screening approach.

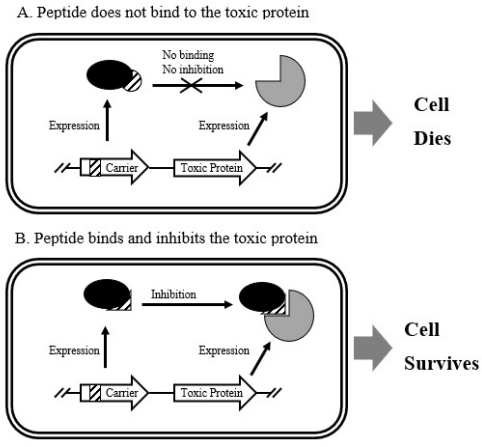

Fig. 1. Selection System. Toxic protein is co-expressed with a library of peptides. Peptide variants are inserted into a carrier protein. A) When a peptide does not inhibit the toxic protein, host cell dies. B) If a candidate peptide binds to and inhibits the target protein, its cytotoxicity is neutralized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

Results and Discussion

Selection. 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 addressed this problem by developing a selection based on the cytotoxicity of the target protein. Other *in vivo* selection methods have relied on the toxicity of an enzyme's (target protein) substrate [13], products of the enzymatic reaction [14], a particular intermediate [15], or resistance to inhibitors [16]. Our selection approach is the first to capitalize on the cytotoxicity of the target protein itself. 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 (Figure 1). Main protease (3CL) of SARS-CoV2 virus was chosen as a model protein because it plays a central role in the virus life cycle. It processes viral polyproteins and controls the replicase complex activity [17], which makes it very attractive target for drug development.

Cytotoxicity of 3CL protease was confirmed by expressing it in *E.coli*. We've built the p3CL plasmid where 3CL gene was cloned under the control of arabinose-inducible promoter of pBAD-HisA vector (Figure 2). This plasmid was transformed in 10G strain of *E.coli* and streaked on the plates with 0.4% arabinose. Strain harboring p3CL plasmid did not grow under these conditions while strains containing empty vector or mutant 3CL proteases did which confirmed our hypothesis that toxicity of this protease was caused by its enzymatic activity.

Libraries. We inserted our peptide libraries into the first loop of ubiquitin. We selected ubiquitin because of its small size (8.6kD), stability in *Ecoli*, and history of used to expressing proteins and peptides [18]. The first loop was selected for library insertion because loops are generally tolerant to insertions and deletions and this particular loop was previously used for insertions [19]. Ubiquitin was co-expressed with 3CL from the same expression construct pUbi-3CL which is shown in Figure 2. In this construct ubiquitin and 3CL genes are arranged in an operon fashion under the control of the arabinose-inducible promoter. A Shine-Dalgarno sequence is inserted between 3CL and Ubiquitin to ensure the expression of both genes.

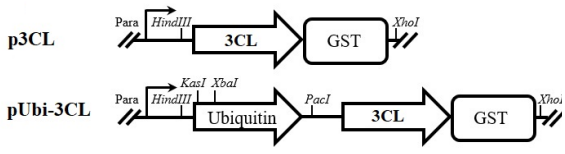

Fig. 2. Expression constructs. Expression of 3CL and Ubiquitin is controlled by the arabinose-inducible promoter in both constructs. 3CL is fused to GST for the expression purposes.

The library construction method is presented in Figure 3. The first peptide library was random, built with 14 degenerate codons, resulting in up to $1.6x10^{18}$ variants. The second library was based on published sequences [20] recognized by 3CL, and contained approximately $2x10^{9}$ variants. 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 were expected to produce linear peptides attached to the first beta strand of ubiquitin.

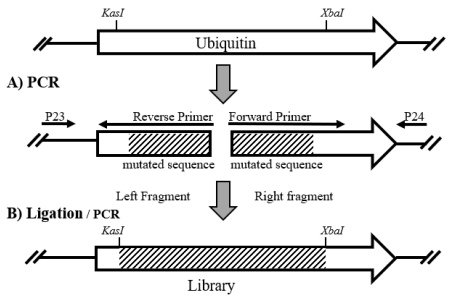

Fig. 3. Library Construction. Degenerate sequences are introduced by PCR: A) The sequences are amplified by degenerate oligos: B) Mutated fragments are phosphorylated by T4 polynucleotide kinase and ligated using T4 DNA ligase.

Selection. Both peptide libraries were cloned in the pUbi-3CL construct (Figure 2) and taken through five rounds of selection in *Ecoli*. Selection was done by inoculating libraries in liquid LB media with 0.4% arabinose and incubation overnight at 37°C with shaking. Every round of selection included 1 million clones for each library. To weed out false positives that may result from frame-shifts, deletions of 3CL and somatic mutations libraries were re-cloned into the original vector (pUbi-3CL) after each round of selection. The fifth round of selection generated several sequences that were significantly overrepresented in the population. We chose the most abundant peptides for further testing. They inhibited the 3CL protease with IC50 ranging from 100 µM to 1.2 mM. The two best peptides were M1 (RQGLDEDLHRW) and M5 (TANAFLS). Their IC_{50} was 249 and 101 µM, 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 3CL. This observation demonstrates that this approach can identify inhibitors without prior knowledge of their ligands and can improve the inhibitory activity of known ligands.

Table 1. Peptide Inhibitors of 3CL Protease.

Peptide Name	*Peptide Sequence*		*IC50 (µM)* *N=4, P<0.05*
M1	GARQGLDEDLHRW	linear	249±47
M5	GATANAFLSGSGSRG	linear	101±17
M5c	WRRWWRRRRTANAFLS	cyclic	34±8

Selected peptide sequences are underlined

It is also important to note that peptide M5 was fully integrated in the first loop of the carrier protein (ubiquitin) which gives it a cyclic structure. To be consistent with the structure in the original screen, we also synthesized peptide M5 in a cyclic form (peptide M5c, Table 1) fused to a custom cell penetration sequence (WRRWWRRRR) to improve its stability and intracellular transport. Cyclization improved IC50 of M5 peptide significantly from 101 to 33 µM (Table 1). These data demonstrate the utility of this selection approach for screening both linear and cyclic peptides.

A weakness of this study is that we were able to screen only a small fraction of all available peptide variants (1 million clones at each stage of selection). We also do not know the exact mechanism of the inhibition yet. Despite these shortcomings, we were able to rapidly identify (in a few weeks) potent peptide inhibitors with low µM activity (Table 1), validating this approach.

References

1. Hopkins, A.L., Groom, C.R. *Nat. Rev. Drug Discov*. **1**, 727-730 (2002), http://dx.doi.org/10.1038/nrd892
2. Drews, J. *Science* **287**, 1960-1964 (2000), http://dx.doi.org/10.1126/science.287.5460.1960
3. Craik, D.J., Fairlie, D.P., Liras, S., Price, D. *Chem. Biol. Drug Des.* **81**, 136-147 (2013), http://dx.doi.org/10.1111/cbdd.12055
4. Van Regenmortel, M.H.V. *Biologicals. Academic Press* **29**, 209-213 (2001), http://dx.doi.org/10.1006/biol.2001.0308
5. Dougherty, P.G., et al. *Chem. Rev.* **119**, 10241-10287 (2019), http://dx.doi.org/10.1021/acs.chemrev.9b00008

6. Marasco, D., et al. *Curr. Protein Pept. Sci.* **9**, 447-467 (2008), http://dx.doi.org/10.2174/138920308785915209
7. Smith, G.P. *Science* **228**, 1315-1317 (1985), http://dx.doi.org/10.1126/science.4001944
8. Bowen, J., et al. *Int. J. Mol. Sci*. **22**, 1-20 (2021), http://dx.doi.org/10.3390/ijms22041634
9. Roberts, R.W., Szostak, J.W. *Proc. Natl. Acad. Sci. USA* **94**, 12297-12302 (1997), http://dx.doi.org/10.1073/pnas.94.23.12297
10. Tavassoli, A., Benkovic, S. *J. Nat. Protoc*. **2**, 1126-1133 (2007), http://dx.doi.org/10.1038/nprot.2007.152
11. Bhagawati, et al. *Proc. Natl. Acad. Sci. USA* **116**, 22164-22172 (2019), http://dx.doi.org/10.1073/pnas.1909825116
12. Aranko, et al. *Protein Eng. Des. Sel*. **27**, 263-271 (2017), http://dx.doi.org/10.1093/protein/gzu028
13. Jiang, P., et al. *Sci. Rep*. **5**, 8563 (2015), http://dx.doi.org/10.1038/srep08563
14. McLoughlin, S.Y., et al. *Protein Expr. Purif*. **41**, 433-440 (2015), http://dx.doi.org/10.1016/j.pep.2005.01.012
15. Boersma, Y.L., et al. *ChemBioChem* **9**, 1110-1115 (2008), http://dx.doi.org/10.1002/cbic.200700754
16. Dickinson, et al. *Nat. Commun.* **5**, 5352 (2014), http://dx.doi.org/10.1038/ncomms6352
17. Anand, K., et al. *Science* **300**, 1763-1767 (2003), http://dx.doi.org/10.1126/science.1085658
18. Baker, R.T. *Curr. Opin. Biotechnol*. **7**, 541-546 (1996), http://dx.doi.org/10.1016/s0958-1669(96)80059-0
19. Ferraro, D.M., et al. *J. Mol. Biol*. **352**, 575-584 (2005), https://doi.org/10.1016/j.jmb.2005.07.012
20. Muramatsu, T., et al. *Proc. Natl. Acad. Sci. USA* **113**, 12997-13002 (2016), http://dx.doi.org/10.1073/pnas.1601327113