Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Sci.*, 2024, **15**, 19520

All publication charges for this article have been paid for by the Royal Society of Chemistry

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Cyclic peptides targeting the SARS-CoV-2 programmed ribosomal frameshifting RNA from a multiplexed phage display library†

Jacob A. Iannuzzelli,‡¶[a] Rachel Bonn,§¶[bc] Andrew S. Hong,¶[a] Abhijith Saseendran Anitha,[ad] Jermaine L. Jenkins,[bc] Joseph E. Wedekind *[bc] and Rudi Fasan *[ad]

RNA provides the genetic blueprint for many pathogenic viruses, including SARS-CoV-2. The propensity of RNA to fold into specific tertiary structures enables the biomolecular recognition of cavities and crevices suited for the binding of drug-like molecules. Despite increasing interest in RNA as a target for chemical biology and therapeutic applications, the development of molecules that recognize RNA with high affinity and specificity represents a significant challenge. Here, we report a strategy for the discovery and selection of RNA-targeted macrocyclic peptides derived from combinatorial libraries of peptide macrocycles displayed by bacteriophages. Specifically, a platform for phage display of macrocyclic organo-peptide hybrids (MOrPH-PhD) was combined with a diverse set of non-canonical amino acid-based cyclization modules to produce large libraries of 10^7 structurally diverse, genetically encoded peptide macrocycles. These libraries were panned against the −1 programmed ribosomal frameshifting stimulatory sequence (FSS) RNA pseudoknot of SARS-CoV-2, which revealed specific macrocyclic peptide sequences that bind this essential motif with high affinity and selectivity. Peptide binding localizes to the FSS dimerization loop based on chemical modification analysis and binding assays and the cyclic peptides show specificity toward the target RNA over unrelated RNA pseudoknots. This work introduces a novel system for the generation and high-throughput screening of topologically diverse cyclopeptide scaffolds (multiplexed MOrPH-PhD), and it provides a blueprint for the exploration and evolution of genetically encoded macrocyclic peptides that target specific RNAs.

Received 19th June 2024
Accepted 3rd October 2024

DOI: 10.1039/d4sc04026k

rsc.li/chemical-science

Introduction

RNA has emerged as a promising therapeutic target because of its key regulatory roles in numerous biological processes.[1–4] Although traditionally viewed as a passive carrier of genetic information, whole-genome sequencing indicates that many transcripts comprise noncoding RNAs (ncRNAs)[5] that regulate transcription, RNA processing, translation, and innate immunity.[6–9] RNA also plays a pivotal role in controlling viral replication and bacterial homeostasis, providing new targets for antimicrobial development.[10–13] The ability of RNA to adopt tertiary structures with deep grooves and concave surfaces as well as motifs for protein binding renders them susceptible to recognition by drug-like molecules.[14–20]

Macrocyclic peptides represent an attractive class of bioactive agents capable of inhibiting biomolecular interactions.[21–27] Their modest size (800–3000 Da) and conformational rigidity impart advantages over their linear counterparts, including increased target affinity,[28,29] enhanced proteolytic stability[30,31] and cell permeability.[32–36] These beneficial characteristics make them ideal candidates for targeting intracellular complexes such as RNA-protein interactions. For example, representative work from our groups and others achieved disruption of the essential HIV Tat–TAR binding interaction through use of a focused group of cyclic peptides that bind the viral TAR RNA major-groove bulge with affinity and specificity.[37,38] Albeit, cyclic peptides that target specific RNA molecules have been identified thus far *via* low throughput, structure-guided-design approaches, which are laborious and limited in terms of the sequence space that can be explored.[39,40] Accordingly, high-

[a]*Department of Chemistry, University of Rochester, Rochester, NY 14627, USA*

[b]*Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, NY, 14642, USA. E-mail: joseph.wedekind@rochester.edu*

[c]*Center for RNA Biology, University of Rochester School of Medicine and Dentistry, Rochester, NY, 14642, USA*

[d]*Department of Chemistry & Biochemistry, The University of Texas at Dallas, Richardson, TX 75080, USA. E-mail: Rudi.Fasan@UTDallas.edu*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4sc04026k

‡ Present address: Merck & Co., Inc., Rahway, NJ 07065, USA.

¶ These authors contributed equally to this work.

§ Present address: Renaissance School of Medicine, Stony Brook University, Stony Brook, NY 11794, USA.

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

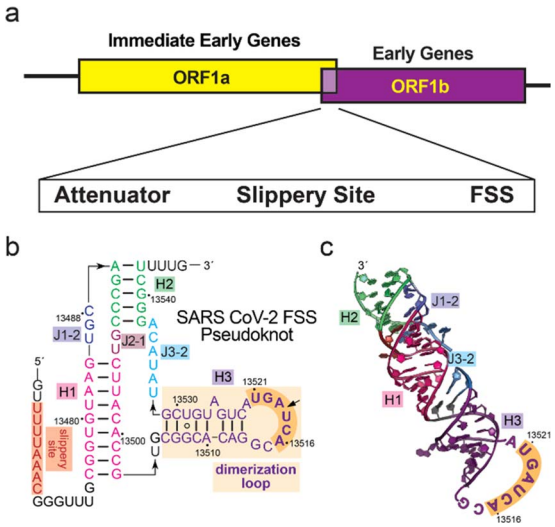


Fig. 1 Overview of the −1 programmed ribosomal frameshifting stimulatory sequence (FSS) of SARS-CoV-2. (a) Genomic organization of the SARS-CoV-2 open reading frames (ORFs) whose expression is controlled by the FSS element. (b) Secondary structure of the FSS pseudoknot (PK) of SARS-CoV-2 with the dimerization loop hairpin (DL) boxed (orange). The PK sequence used for this study began at C13476 and ended at U13543. (c) Crystal structure of the FSS PK variant (PDB 7mky). The GNRA tetra loop in the original structure was replaced with the actual dimerization loop sequence.

throughput platforms capable of probing structurally and functionally diverse libraries of macrocyclic peptides against specific RNA targets are highly desirable.

Recently, some of us introduced an innovative platform to discover bioactive cyclic peptides that combines the production of genetically encoded macrocyclic organo-peptide hybrids (MOrPHs) using non-canonical amino acids with M13 bacteriophage display (*i.e.*, MOrPH-PhD).[41] This system allows for the exploration of large libraries up to 10^9 macrocyclic peptides against various target proteins and protein–protein interactions (PPIs). In this system, cyclization of a ribosomally derived peptide is achieved *via* a spontaneous, post-translational cross-linking reaction between a genetically encoded electrophilic unnatural amino acid (eUAA) and a proximal cysteine residue, leading to a side-chain-to-side-chain linked macrocyclic peptide.[42,43] The eUAA cyclization module is incorporated into the precursor polypeptide *via* amber stop codon (UAG) suppression using an engineered aminoacyl tRNA synthetase/$tRNA^{CUA}$ pair derived from *Methanococcus jannaschii*. In the MOrPH-PhD system, the cyclic peptide is fused to the N-terminus of the pIII phage coat protein. *Via* helper phage-assisted assembly, mature M13 phage particles display the MOrPH library on their surface, establishing a physical linkage between genotype and phenotype.[41]

Building upon this work, we were interested in exploring the utility of this platform for the discovery of cyclic peptides capable of targeting a RNA molecule, using the −1 programmed ribosomal frameshifting (−1 PRF) RNA of SARS-CoV-2 as a model target (Fig. 1). This RNA element plays a key role in the viral replication of SARS-CoV-2,[44] which folds as a conserved three-stem H-type pseudoknot (PK) that functions as part of a 'frameshift-stimulating sequence' (FSS).[45] We chose to target the FSS PK because it interacts with the host ribosome, leading

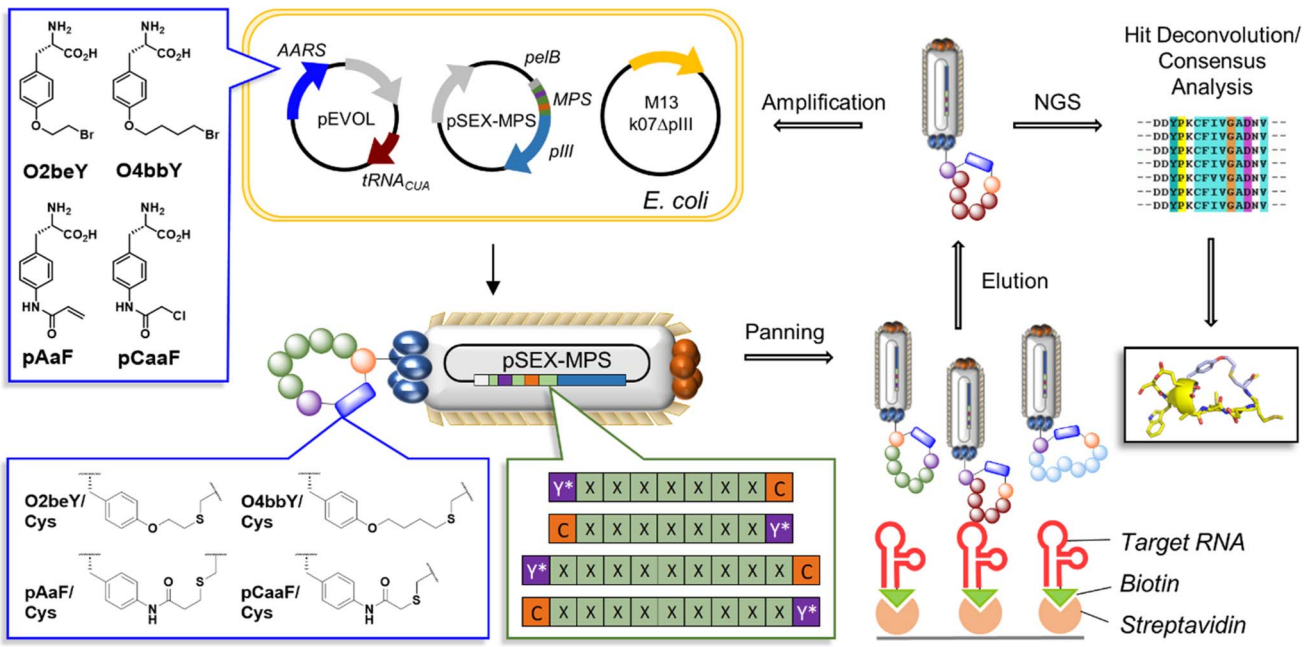


Fig. 2 Multiplexed MOrPH-PhD system for the selection of SARS-CoV-2 FSS-targeting cyclic peptides. MOrPH-PhD libraries were diversified through cyclization of 11-mer ($i/i \pm 10$) and 9-mer ($i/i \pm 8$) peptide sequences using four different eUAAs and two different orientations for the eUAA/Cys linkage (X = randomized amino acid position, C = cysteine residue, Y* = eUAA). See Fig. S1† for further information about the library design.

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

to a −1 frameshift[46] and the translation of two distinct viral polyproteins (encoded by ORF1a and ORF1ab) essential for the SARS-CoV-2 lifecycle.[47] The −1 PRF FSS system is structurally and functionally conserved within the betacoronavirus genus, and mutations to this sequence are detrimental to viral propagation.[10] Therefore, the FSS is an attractive target for antiviral therapy, as evident in recent studies showing that chemical targeting of this motif can interfere with the FSS-mediated frameshifting mechanism[48,49] and reduce SARS-CoV-2 replication in cells.[11,48] Despite progress in the development of platforms for high-throughput screening of genetically-encoded[41,50–55] or phage-encoded[56–60] macrocyclic peptide libraries, there are no reports of their successful application to target RNA to our knowledge. While the use of a phage-encoded library led to the discovery of a bicyclic peptide with sub-micromolar affinity for a DNA G-quadruplex,[61] the application of a phage-display library of cyclic peptides constrained by a disulfide linkage against BIV TAR RNA resulted in compounds with no specificity against the target RNA.[62]

Here, we report the discovery and characterization of RNA targeting cyclic peptides using a "multiplexed MOrPH-PhD" system, in which phage-displayed MOrPH libraries are diversified by means of different eUAA cyclization modules (Fig. 2). This method led to the selection of multiple macrocyclic peptides that specifically and selectively target the −1 PRF RNA of SARS-CoV-2 (Fig. 2). Details of our multiplexed MOrPH-PhD approach are presented along with binding analysis to the FSS PK and localization using chemical mapping. These findings have broader implications for the identification and development of new antiviral and antibacterial molecules.

Results

Multiplexing of the MOrPH-PhD library enhances diversity yielding a library of >50 million unique peptides

We first sought to design a set of MOrPH-PhD libraries capable of targeting the FSS pseudoknot (PK) of SARS-CoV-2. To this end, for randomization of the peptide sequence we chose to use restricted amino acid alphabets that comprise amino acid residues found to be statistically more prevalent at protein-RNA interfaces (*i.e.*, Lys, Arg, His, Phe, Tyr, Leu, Asn, Gln).[63–65] Accordingly, the macrocycle peptide sequences were randomized using different patterns of the following degenerate codons: (MRW = R, N, Q, K, H, S), (YWT = H, L, F, Y) and (HWS = N, Q, H, I, L, K, M, F, Y) (Fig. S1†). Peptide libraries were further diversified by varying the ring size (9-mer *vs.* 11-mer) and orientation of the eUAA/Cys linkage (*e.g.*, $i/i+8$ *vs.* $i/i-8$, where the ith position indicates the eUAA relative to the invariant cysteine nucleophile). Previous studies demonstrated that the orientation of the thioether linkage can drastically impact the target binding properties of these peptides.[43,66] Because of these structural variations, the overall library comprises ∼13.6 million unique peptide sequences (Fig. 2).

While O-2-bromoethyl tyrosine (O2beY) was the initial eUAA for cyclization, this choice was extended subsequently to a broader suite of eUAAs based on recent work demonstrating additional methodologies to establish thioether crosslinks.[43] Variation of the thioether crosslink was shown to impact the functional properties of bioactive cyclic peptides. Accordingly, an additional element of structural diversification was incorporated into our libraries through the multiplexed integration of our MOrPH-PhD libraries with different eUAAs as cyclization modules, namely O2beY, O-4-bromobutyl tyrosine (O4bbY), *p*-acrylamido phenylalanine (*p*AaF) and *p*-chloroacetamido phenylalanine (*p*CaaF). Altogether, the resulting library comprises 54.5 million unique macrocyclic peptides.

Panning against the target RNA reveals consensus peptides

Our initial experiments determined that phage particles — produced in bacterial cell cultures — contain significant amounts of RNase, which led to rapid degradation of the target RNA (Fig. S2†). This phenomenon highlights a fundamental challenge when applying high-throughput techniques to RNA targets. We found the addition of a protein-based RNase inhibitor resulted in significant reduction in RNA degradation with the majority of the full-length FSS PK remaining intact after overnight incubation with the phage solution (Fig. S2†). Based on these results, we included RNase inhibitor in the phage solution prior to panning against the target RNA. Furthermore, we also envisioned that the full-length FSS PK, which comprises 68 nucleotides, may be susceptible to chemical degradation, particularly within less structured joining regions. In this context, the dimerization loop (DL) hairpin, which comprises 26 conserved nucleotides that play a key role in the −1 PRF mechanism of action, is a shorter structural motif contained within the FSS and would be less susceptible to degradation than the full-length FSS PK (Fig. 2).[67] Accordingly, in addition to the full-length FSS PK, MOrPH-PhD libraries were screened in parallel against a more compact motif corresponding to the DL hairpin of the FSS.

Based on these considerations, selection experiments were carried out by panning the four MOrPH-PhD libraries, each constructed using a different eUAA, against both the DL and the full-length FSS biotinylated RNA immobilized on streptavidin-coated magnetic beads. All libraries were subjected to three rounds of affinity-based selection and amplification (Fig. 3A and B). Notably, the MOrPH-PhD libraries show differential levels of post-selection recovery depending on the nature of the eUAA module (Fig. 3A and B), likely reflecting the differential ability of these cyclization modules to generate cyclic peptides capable of interacting with the target RNA. In particular, MOrPH libraries cyclized *via* O4bbY or *p*CaaF showed substantially higher levels of phage recovery during the second and third rounds of affinity selection, relative to the MOrPH libraries cyclized *via* O2beY or *p*AaF (Fig. 3A and B). Based on these results, the library members recovered after the second and third round of selection from O4bbY- and *p*CaaF-based libraries against the DL and FSS PK targets were analyzed *via* next-generation sequencing (NGS). In addition, libraries recovered from the second round and third round of selection against the DL for *p*AaF- and O2beY-containing libraries were similarly subjected to NGS.

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

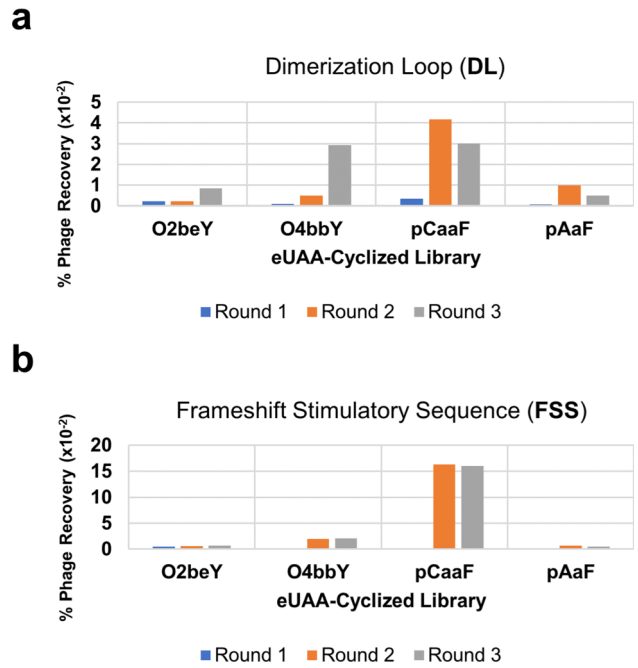


Fig. 3 Phage recovery after each round of panning. Percent phage recovery of individual eUAA-containing libraries through three rounds of affinity selection and amplification against the (a) DL and (b) FSS.

Consistent with the trend observed from phage recovery analyses, NGS showed enrichment of peptide sequences primarily in the O4bbY- and *p*CaaF-based libraries. By contrast, sequencing results from O2beY- and *p*AaF-containing libraries failed to yield peptide sequences with significant enrichment. Accordingly, macrocyclic peptides selected for further characterization were chosen solely from the O4bbY- and *p*CaaF-based cyclic peptide libraries. Interestingly, enriched macrocyclic peptides ('hits') were obtained from all four O4bbY-based libraries (*i.e.*, for both the 9-mer and 11-mer libraries and for both linkage orientations) with distinct consensus sequences associated with each library (Fig. 4A). By contrast, sequence 'hits' derived from the *p*CaaF-containing libraries were found solely in the Cys/*p*CaaF orientation (*i.e.*, $i/i-8$ or $i/i-10$), indicating a clear advantage of the Cys/*p*CaaF connectivity over the *p*CaaF/Cys connectivity for binding to the target RNAs (Fig. 4A). Overall, these results evidence the impact of different cyclization modules and linkage orientations on phage recovery and enrichment, highlighting the distinct advantage of the present multiplexed MOrPH system for exploring a larger macrocycle space than possible using a single eUAA module and thus increasing chances of success in hit identification against a desired target.

Analysis of the sequences selected from the 9-mer ($i/i-8$) *p*CaaF-based library showed a single, highly enriched peptide (*i.e.*, *p*CaaF(i − 8)-m1, called **FSS-m1**) (Fig. 4A) with markedly higher abundance over the other top-scoring sequences (19% *vs.* <0.6%). On the other hand, the 11-mer ($i/i-10$) *p*CaaF-based library yielded two highly enriched peptides, one of which represented as much as 37% of the sequence reads (*i.e.*, *p*CaaF($i-10$)-m1, called **FSS-m2**). In addition, several other sequences among the top-scoring ones show a high level of homology to **FSS-m2** (Fig. 4A). Notably, **FSS-m1** and **FSS-m2**, the top hits from the 9-mer and 11-mer libraries, respectively, share nearly identical motifs (–KQL(R/H)–).

Sequencing results from the 9-mer ($i/i+8$) O4bbY-linked library (Fig. 4A) revealed four peptides with significant abundance within the library (0.4–4.9%). Interestingly, the top sequences identified from this library show a high level of homology, displaying two distinct consensus groups of cyclic peptides. The top enriched peptide (*i.e.*, O4bbY($i+8$)-m1, called **FSS-m3**) shares a –KKYR– motif with the less abundant (0.4%) O4bbY($i+8$)-m3 peptide (called **FSS-m5**). The second and third top enriched peptides share a similar –LQ(Q/N)– motif (Fig. 4A). Furthermore, a consensus was observed among the sequences identified from the 11-mer ($i/i-10$) O4bbY-based library with the top sequence (*i.e.*, O4bbY($i-10$)-m1, called **FSS-m8**) sharing a similar –KRH– motif with the second most highly enriched peptide. In addition, a preference for aromatic amino acid residues (Phe/Tyr) is observed at the $i+1$ position relative to the preinstalled cysteine at the *i*th position, with the exception of the third top enriched peptide, which contains an amphipathic Arg residue at the $i+1$ position (Fig. 4A).

In general, panning experiments against the DL resulted in more substantial enrichment and stronger consensus sequences among hits compared to those targeting the FSS PK. For most of the libraries, sequence hits were identified from affinity selection experiments against the DL, although the same sequences were also found among those selected from panning experiments against the FSS PK. As an exception, the opposite result (*i.e.*, stronger enrichment and consensus for hits selected against the FSS PK *vs.* DL) was observed for the *p*CaaF-based libraries. Notably, sequence hits identified from the FSS PK binding experiments included **FSS-m2**, which showed the largest degree of enrichment (and strongest consensus) among hits from all libraries (Fig. 4A). The FSS PK panning experiments notably yielded strong hits from only one sub library (*i.e.*, that comprising 11-mer cyclic peptides constrained by a *p*CaaF-based thioether linkage in the Cys/*p*CaaF orientation). These results indicate that — among the 16 different macrocyclic peptide scaffolds tested — this cyclopeptide topology is optimal and nearly uniquely suited for interaction with the FSS PK — a result consistent with the distinctly higher phage recovery observed for the *p*CaaF-based libraries during panning against this target (Fig. 3B).

Solid-phase synthesis of selected cyclic peptides

Based on the sequencing results and consensus motifs derived above, we selected eight representative cyclic peptides for synthesis and further characterization derived from the O4bbY- and *p*CaaF-based MOrPH libraries. To this end, solid-phase peptide synthesis (SPPS) protocols were developed to prepare peptides cyclized *via* Cys/*p*CaaF, O4bbY/Cys and Cys/O4bbY linkages (Scheme 1A–C). Briefly, the Cys/*p*CaaF-cyclized peptides were synthesized *via* the use of *N*′-allyloxycarbonyl (Alloc)-protected *p*-aminophenylalanine (*p*AmF) and Acm-

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

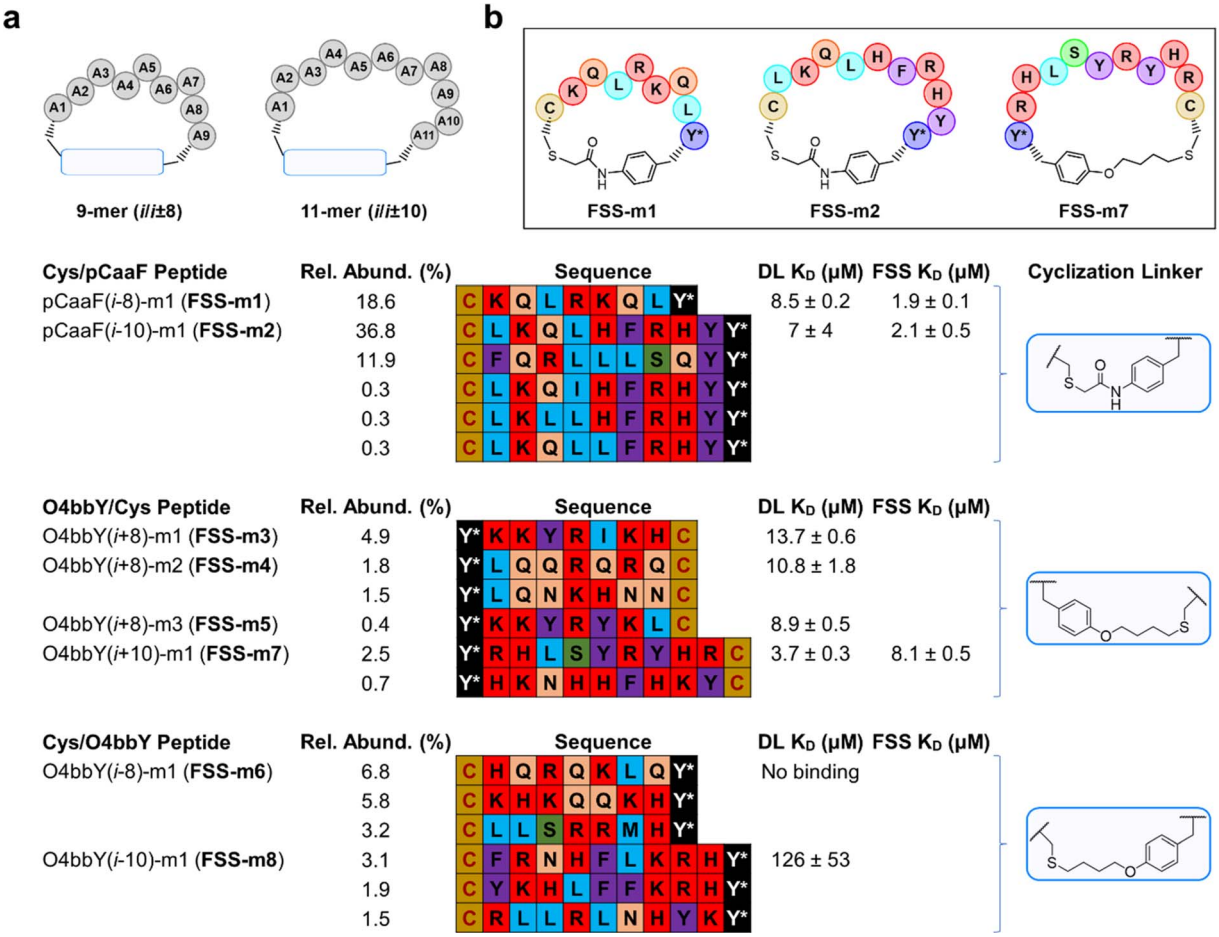


Fig. 4 Macrocyclic peptides identified from deconvoluted MOrPH-PhD libraries. (a) Top enriched cyclic peptides were identified from $i/i \pm 8$ and $i/i \pm 10$ libraries. Cyclic peptides identified from sequenced libraries panned against the FSS comprise the *p*CaaF $i/i - 10$ 11-mer library. All other peptides were identified from NGS deconvoluted libraries panned against the DL RNA from SARS-CoV-2. (b) Cyclic peptides with the highest binding affinity for the DL RNA.

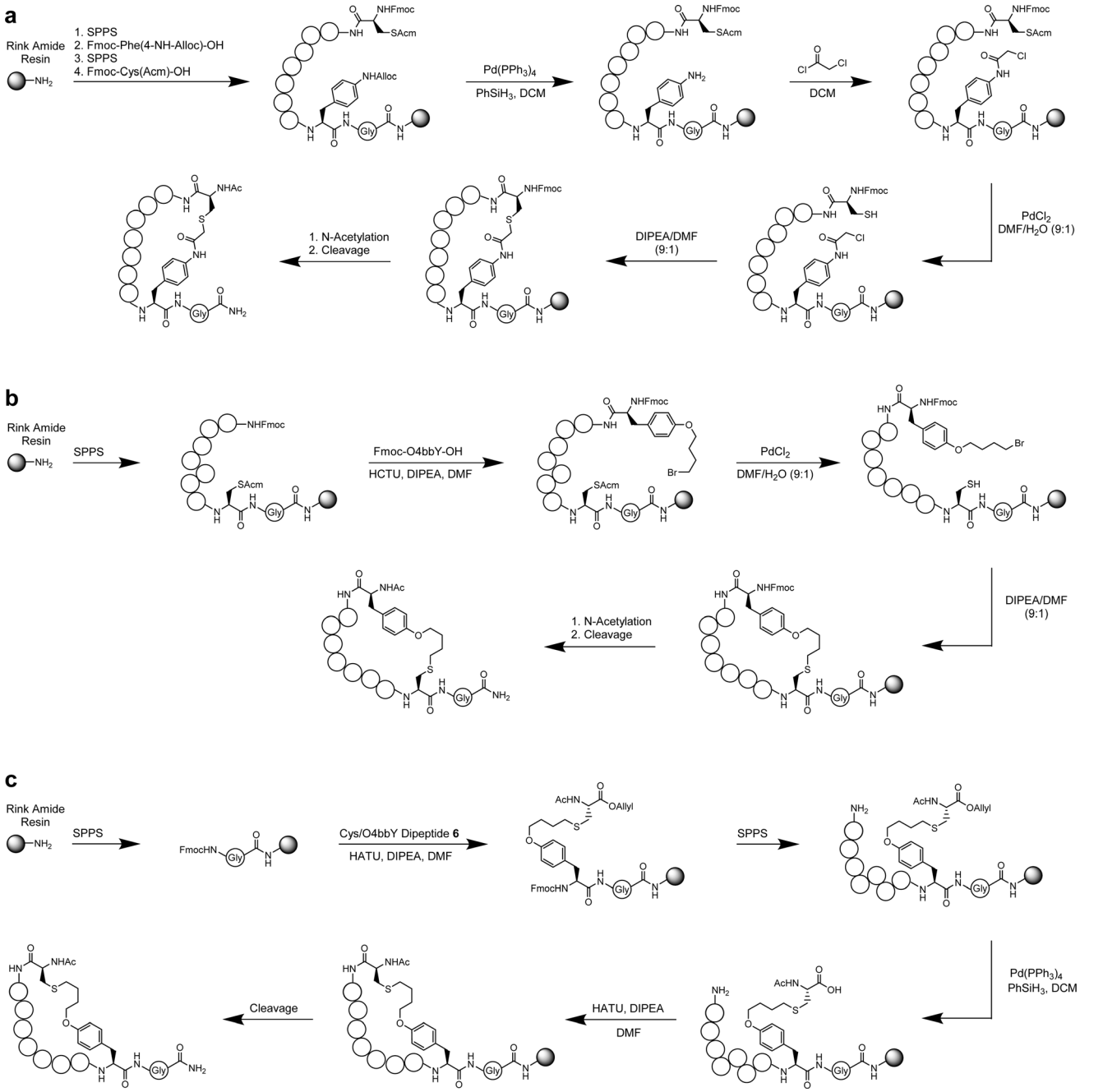
protected cysteine. After incorporation of the protected *p*AmF residue and assembly of the remainder of the peptide, the side-chain Alloc group was removed using a $Pd(PPh_3)_4$ catalyst and $PhSiH_3$ followed by acylation with chloroacetyl chloride. After peptide assembly, the cysteine residue was deprotected with a $PdCl_2$ catalyst followed by on-resin cyclization under basic conditions (Scheme 1A). The O4bbY/Cys-linked peptides were prepared *via* a different strategy using an acetamidomethyl (Acm)-protected cysteine and 9-fluorenylmethyloxycarbonyl (Fmoc)-protected O4bbY. Upon the completion of peptide assembly, cysteine deprotection and on-resin cyclization were performed in a similar manner to that described above (Scheme 1B). Synthesis of Cys/O4bbY cyclized peptides involved the use of an allyl-protected dipeptide encompassing the Cys/O4bbY linkage (Scheme S1†). Upon incorporation of the dipeptide and assembly of the remainder of the peptide, the side-chain allyl group was removed using a $Pd(PPh_3)_4$ catalyst and $PhSiH_3$ followed by cyclization *via* amide bond formation (Scheme 1C). Further details regarding methods for solid-phase synthesis and characterization of the cyclic peptides are described in the ESI.†

Cyclic peptides bind SARS-CoV-2 RNA targets

After synthesis, cyclic peptides **FSS-m1** through **FSS-m8** were characterized for DL binding affinity (apparent K_D) and binding kinetics (k_{on} and k_{off}) using surface plasmon resonance (SPR) (Table 1 for averaged binding parameters and Fig. S4† for representative binding sensorgrams and curve fits). These experiments showed that peptides **FSS-m1**, **FSS-m2**, and **FSS-m7** bind the dimerization loop with micromolar affinity (K_D = 3.6–4.6 μM). To test target specificity, the SPR experiments were repeated in the presence of a 100-fold molar excess of yeast tRNA in the SPR buffer relative to the immobilized DL RNA target. Importantly, each cyclic peptide maintained the same or only slightly weakened (2–3 fold) affinity for binding (K_D = 3.7–8.5 μM), demonstrating the specificity of each peptide for the target RNA. Peptides **FSS-m1** and **FSS-m2**, which comprise 9-mer and 11-mer cyclic peptides, share a common KQL motif, and display similar DL binding affinities (K_D values of 3.6 ± 0.2 μM and 4.6 ± 0.2 μM). In addition, each showed only a slight increase in K_D in the presence of tRNA (1.6–2.4-fold), suggesting each peptide forms a specific interaction with the DL structure. Peptide **FSS-m7**, which comprises an 11-mer cyclized *via* an

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.



Scheme 1 SPPS methods for the generation of cyclic peptides analyzed in this study. (a) SPPS of peptides comprised by a Cys/*p*CaaF linkage. (b) SPPS of peptides comprised by an O4bbY/Cys linkage. (c) SPPS of peptides comprised by a Cys/O4bbY linkage.

O4bbY/Cys linkage, maintained full affinity in the presence of tRNA, as indicated by no discernible difference in the K_D under conditions with and without tRNA (*i.e.*, K_D values of 3.7 ± 0.3 μM *vs.* 3.7 ± 1.6 μM, Table 1).

Encouraged by these results, we investigated other peptides for DL binding in the presence of tRNA. Peptides **FSS-m3**, **FSS-m4**, and **FSS-m5** (Fig. 4A) — identified from the 9-mer library containing an *i*/*i* + 8 O4bbY/Cys linkage — yielded micromolar affinity to DL (K_D values from 9–14 μM; Table 1). On the other hand, peptide **FSS-m6**, which comprises an *i*/*i* − 8 Cys/O4bbY linkage in the reverse orientation, did not exhibit binding. By contrast, peptide **FSS-m8**, which comprises an *i*/*i* − 10 Cys/O4bbY linkage, exhibited high micromolar affinity to the DL (K_D >120 μM; Table 1). This stark contrast highlights a clear advantage of the O4bbY/Cys cyclization topology over the Cys/O4bbY topology for binding to the DL.

The highest affinity cyclic peptides identified in the DL binding experiments (*i.e.*, peptides **FSS-m1**, **FSS-m2**, and **FSS-m7**) were tested next for binding to the FSS PK in the presence of tRNA (Fig. 4B). Notably, *p*CaaF-based cyclic peptides **FSS-m1** and **FSS-m2**, which share the –KQL–consensus motif, displayed three- to four-fold increases in binding affinity for the FSS PK

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Table 1 Average binding and kinetic parameters of DL binding by cyclic peptides in SPR buffer containing 100-fold molar excess yeast tRNA[c d e]

Peptide	$k_{on} \times 10^2$ (M^{-1} s^{-1})	S.E. $\times 10^2$ (M^{-1} s^{-1})	$k_{off} \times 10^{-2}$ (s^{-1})	S.E. $\times 10^{-2}$ (s^{-1})	K_D (μM)	S.E. (μM)	X^2 (RU^2)	S.E.
FSS-m1	11.3 (7.9[a])	0.05 (0.6[a])	0.96 (0.28[a])	0.01 (0.01[a])	8.5 (3.6[a])	0.2 (0.2[a])	1.9 (2.3[a])	0.05 (0.1[a])
FSS-m2	n.a.	n.a.	n.a.	n.a.	7.4[b] (4.6[a,b])	3.5[b] (0.2[a,b])	0.5[b] (12.8[a,b])	0.2[b] (2.1[a,b])
FSS-m3	17.4	0.8	2.38	0.004	13.7	0.6	222	13.2
FSS-m4	32.8	3.0	3.59	0.84	10.8	1.8	4.2	0.4
FSS-m5	31.2	0.9	2.76	0.09	8.9	0.5	100	1.3
FSS-m6	n.b.	n.b.	n.b.	n.b.	n.b.	n.b.	n.b.	n.b.
FSS-m7	44.9 (21.4[a])	3.8 (1.3[a])	1.65 (1.06[a])	0.01 (0.02[a])	3.7 (3.7[a])	0.3 (1.6[a])	10.0 (96.5[a])	0.7 (6.9[a])
FSS-m8	n.a	n.a	n.a.	n.a	126.3	52.8	2.7	1.0

[a] Peptides tested in SPR buffer lacking 100-fold molar excess yeast tRNA. [b] Equilibrium dissociation constants determined *via* steady-state analysis. [c] S.E. indicates standard error. [d] n.a. indicates not assessed. [e] n.b. indicates no detectable binding.

Table 2 Average binding and kinetic parameters of FSS PK binding by cyclic peptides **FSS-m1**, **FSS-m2**, and **FSS-m7** in buffer containing 100-fold molar excess yeast tRNA

Peptide	$k_{on} \times 10^2$ (M^{-1} s^{-1})	S.E. $\times 10^2$ (M^{-1} s^{-1})	$k_{off} \times 10^{-2}$ (s^{-1})	S.E. $\times 10^{-2}$ (s^{-1})	K_D (μM)	S.E. (μM)	X^2 (RU^2)	S.E
FSS-m1	16.3	4.8	0.31	0.10	1.9	0.01	0.03	0.01
FSS-m2	65.8	13.0	1.24	0.03	2.1	0.47	0.14	0.01
FSS-m7	10.6	0.6	0.85	0.01	8.1	0.52	92.9	0.10

compared to DL (8.5 ± 0.2 μM *vs.* 1.9 ± 0.01 μM and 7.4 ± 3.5 μM *vs.* 2.1 ± 0.5 μM; Tables 1 and 2 for averaged binding parameters and Fig. S3F and G† for representative binding sensorgrams and curve fits). It is worth noting that peptide **FSS-m2** (*p*CaaF($i-10$)-m1) was identified from the libraries panned against the FSS PK. By contrast, 11-mer cyclic peptide **FSS-m7** cyclized *via* an O4bbY/Cys linkage, produced a poorer K_D of 8.1 μM for binding to the FSS PK (Table 2 for averaged binding parameters and Fig. S4H† for a representative binding sensorgram and curve fit). This value is 2-fold worse than binding to the shorter DL, which yielded a K_D of 3.7 ± 0.3 μM (Table 1). The molecular basis for this difference in affinity is unknown at present.

To evaluate whether the cyclic peptides exhibited off-target RNA interactions with pseudoknots, we tested **FSS-m1** and **FSS-m2** for binding to riboswitch RNAs that adopt different pseudoknot folds. Specifically, we chose the type II $preQ_1$-I riboswitch, which adopts an H-type pseudoknot that binds a single $preQ_1$ (7-aminomethyl-7-deazaguanine) molecule.[68] We also tested a type I $preQ_1$-I riboswitch that folds as an H-type pseudoknot to cooperatively bind two $preQ_1$ equivalents in a single binding pocket.[69] The third riboswitch adopts an HHH pseudoknot fold that recognizes a single equivalent of guanine.[70] These experiments revealed that neither peptide showed evidence of binding to off-target pseudoknots at concentrations that elicit a binding response in the presence of FSS PK (Fig. S5† *vs.* Table 2 and Fig. S4†).

Chemical modification localizes peptide binding to the FSS dimerization loop

We next sought to identify the binding location for each of the three highest affinity cyclic peptides. Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) followed by next-generation sequencing was used to probe peptide binding to the RNA target.[71] The 26-nucleotide dimerization stem loop of the SARS-CoV-2 FSS PK was embedded in a folding cassette comprising flanking 5′ and 3′ hairpins and a downstream primer binding site (Fig. 5A);[72] two additional G–C pairs were added to the stem for stability. As an internal control to monitor non-specific binding, the 5′ hairpin was composed of the HIV-1 FSS hairpin loop. The RNA cassette was folded and bound to either **FSS-m1**, **FSS-m2** or **FSS-m7** and subsequently acylated by 2-methylnicotinic acid imidazolide (NAI).[73] Differential reactivity was quantified by deep-sequence reads measured for the bound and unbound RNA. Each peptide revealed a significant decrease in reactivity at base U13518 (Fig. 5B–D). In addition, flanking nucleotides C13517 and A13521 showed minor reductions in acylation. Whereas the 9-mer *p*CaaF($i-8$)-m1 (**FSS-m1**) showed almost no change at A13516, 11-mer cyclic peptides *p*CaaF($i-10$)-m1 (**FSS-m2**) and O4bbY($i+10$)-m1 (**FSS-m7**) revealed a slight decrease in reactivity. Reductions in acylation suggest a loss in flexibility in the presence of ligand,[74] possibly through a direct interaction or through allosteric changes due to cyclic peptide binding (Fig. 5B–D). Although the stem loop nucleotides are predicted to reside in an unpaired region, A13516 through U13521 engage in palindromic intermolecular Watson–Crick pairing *via* a kissing loop that is key for effective frameshifting and viral RNA synthesis.[67] Most notably, U13518 shows the greatest reduction in acylation in the presence of each cyclic peptide and is predicted to pair intermolecularly with A13519′.

In terms of off-target binding, our workflow allowed parallel monitoring of an upstream 5′ HIV-1 FSS element (Fig. 5). Specifically, each cyclic peptide produced acylation changes in the stem loop. Cyclic peptides **FSS-m1** and **FSS-m2** produced increases in flexibility (Fig. 5B and C), whereas **FSS-m7**

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

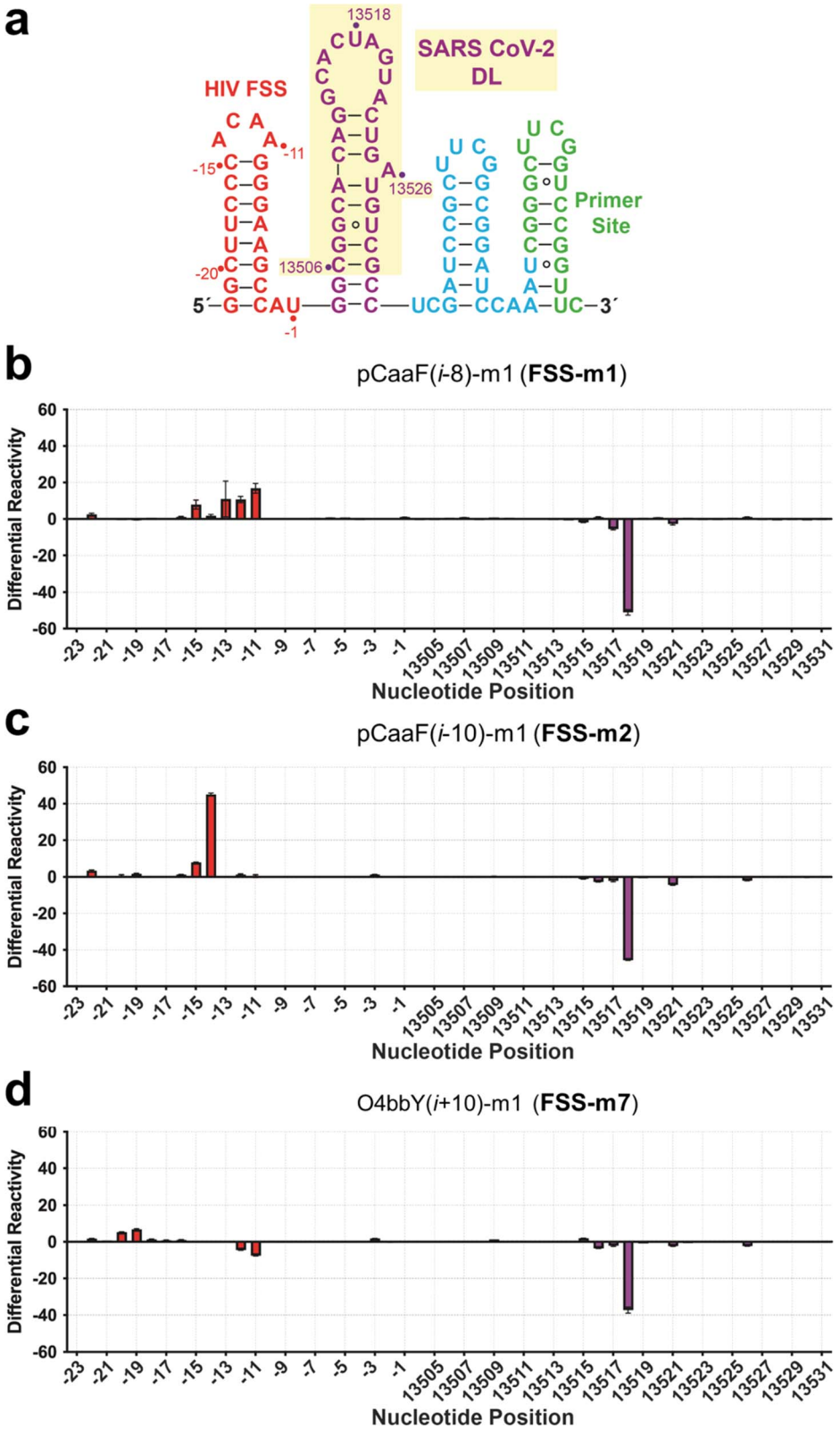


Fig. 5 SHAPE-Seq data localizes peptide binding to U13518 within the dimerization loop. (a) Secondary structure of the RNA cassette containing the stable HIV-1 FSS hairpin at the 5′ end, followed by the SARS-CoV-2 FSS dimerization loop (DL) upon its stem, a strong 3′ linker hairpin, and the reverse transcriptase primer binding site. SARS-CoV-2 numbering corresponds to reference genome NC_045512.2. (b) Differential SHAPE reactivity ($\Delta\rho$) profiles of the SARS-CoV-2 DL showing average differential acylation in the presence and absence of peptide (*i.e.*, $\rho^{+(pCaaF(i-8)\text{-}m1)} - \rho^{-(pCaaF(i-8)\text{-}m1)}$) *versus* sequence position. (c) Differential SHAPE reactivity ($\Delta\rho$) profiles of the SARS-CoV-2 DL showing average differential acylation in the presence and absence of peptide (*i.e.*, $\rho^{+(pCaaF(i-10)\text{-}m1)} - \rho^{-(pCaaF(i-10)\text{-}m1)}$) *versus* sequence position. (d) Differential SHAPE reactivity ($\Delta\rho$) profiles of the SARS-CoV-2 DL showing average differential acylation in the presence and absence of peptide (*i.e.*, $\rho^{+(O4bbY(i+10)\text{-}m1)} - \rho^{-(O4bbY(i+10)\text{-}m1)}$) *versus* sequence position. Each bar represents the average of two replicates with standard deviations shown.

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

produced minor decreases in flexibility (Fig. 5D). Such differences may reflect the different sequences of these peptides since **FSS-m1** and **FSS-m2** share a –KQL(R/H)– consensus motif, whereas **FSS-m7** is rich in positive charge but devoid of glutamine (Fig. 4). Notably, the latter peptide prompts fewer off-target changes while eliciting a strong acylation decrease on the SARS CoV-2 FSS stem loop at U13518.

Scanning mutagenesis of RNA-targeting MOrPHs reveals key amino acids

We next asked whether individual amino acids could be identified in a representative cyclic peptide that alter SARS CoV-2 dimerization loop recognition. One key advantage of the MOrPH system is that peptides can be expressed recombinantly in *E. coli*, where they undergo spontaneous and chemoselective cyclization to form the desired macrocyclic peptide.[41,43,66] Accordingly, constructs for recombinant production of cyclic peptide **FSS-m1** and alanine-scanning mutants at each position of the peptide were designed to comprise the macrocyclic peptide precursor with N-terminal Met–Gly sequence fused to an polyhistidine-tagged Mxe GyrA intein (Table S3†). Tyr was inserted at the junction between the macrocyclic peptide precursor and the intein to promote thiophenol-mediated intein cleavage after expression and isolation *via* Ni-affinity chromatography, resulting in macrocyclic peptide MG-**FSS-m1** and its alanine variants containing an additional Gly and Tyr at the N- and C-terminus, respectively (Table S6†).

After purification following intein cleavage, SPR analysis of the parent recombinant peptide MG-**FSS-m1** and its alanine variants was performed to elucidate key residues for DL binding (Fig. S7†). Responses for mutants K1A, R4A and K5A were comparable to the response for MG-**FSS-m1**, suggesting the side chains are not critical for target binding. By contrast, the response for L7A was lower than the parent peptide. Notably, the response for Q6A was substantially higher than the response for MG-**FSS-m1**, while the responses for Q2A and L3A were substantially lower, suggesting these side chains contribute to RNA affinity or cyclic peptide conformation. Interestingly, Q2 and L3 are each contained within the consensus motif (–KQL(R/H)–) shared by **FSS-m1** (*p*CaaF($i-8$)-m1) and **FSS-m2** (*p*CaaF($i-10$)-m1), suggesting this region of the motif may be important for targeting the SARS CoV-2 FSS dimerization loop. By contrast, mutation of R4A was approximately neutral, which is unexpected if the guanidinium group participates in RNA recognition.[37,75] Q6A was the only position that enhanced binding affinity relative to the parent peptide, consistent with the lack of conservation at this position in cyclic peptide **FSS-m2** (Fig. 4).

Discussion

Due to the emerging appreciation of RNA as a promising target,[4,14,49,76,77] we investigated here the potential of genetically encoded macrocyclic peptide libraries for targeting RNA, using the −1 programmed ribosomal frameshifting stimulatory sequence (FSS) RNA pseudoknot of SARS-CoV-2 as a model of a therapeutically relevant RNA molecule.[46,78] In particular, we have introduced and applied here a multiplexed version of the MOrPH-PhD platform,[41] in which libraries of phage displayed macrocyclic peptides are diversified through the use of multiple eUAAs (*i.e.*, O2beY, O4bbY, *p*AaF, and *p*CaaF) and thus produce different types of intramolecular thioether crosslinks. By exploiting the modularity of the MOrPH architecture, these cyclopeptide libraries were further diversified through variation of the ring size (9-mer *vs.* 11-mer) and orientation of the eUAA/Cys linkage (*e.g.*, $i/i+8$ *vs.* $i/i-8$, where i is the position of the eUAA with respect to the invariant cysteine residue). Finally, to favor RNA targeting, the amino acid sequences of these peptide libraries were randomized using amino acid residues that are found to be statistically more prevalent at protein-RNA interfaces, including Lys, Arg, His, Phe, Tyr, Leu, Asn, and Gln.[63–65] Altogether, the resulting RNA-targeting library contains ∼55 million unique macrocyclic peptides.

Additional elements were implemented to adapt the present phage display system for RNA targeting. As anticipated, our experiments were hindered initially by RNase degradation during library panning. This setback was overcome by use of a protein-based RNase inhibitor that allows broad application of our phage display approach against RNA targets. We also found that RNase degradation could be minimized by applying our method to the shorter 26-nucleotide dimerization loop hairpin (DL), which is substantially less complex in terms of its fold and the number of non-helical joining regions (Fig. 1).

Of the ∼55 million peptide sequences subjected to the panning procedure, we found enrichment primarily of cyclopeptide members from the O4bbY- and *p*CaaF-based libraries. Sequencing results from O2beY- and *p*AaF-containing libraries failed to yield peptide sequences with significant levels of enrichment, suggesting such cyclic peptides cannot adopt conformations suited for molecular recognition of the DL or full-length FSS PK. Moreover, panning of the same libraries against the smaller DL and the full-length FSS PK produced a larger number of hits for the former target RNA molecule. On the one hand, these results highlight the value of the present multiplexed MOrPH-PhD system toward enabling the exploration of different cyclopeptide topologies against a target of interest. On the other hand, the results suggest that the pursuit of small, well-defined motifs — such as the dimerization loop — is more tractable than more complex targets, such as the full-length FSS. This observation may be related to the inherently dynamic structural properties of the full-length FSS pseudoknot.[46,79–82] Our findings have broader implications for target choices using high-throughput ligand screening platforms.

Based on the sequencing results and consensus analysis, we chose eight representative cyclic peptides derived from the O4bbY- and *p*CaaF-based MOrPH libraries for synthesis and characterization. Binding was assessed by SPR, which resulted in the identification of three high affinity peptides for the target RNA, namely **FSS-m1**, **FSS-m2** and **FSS-m7**. Binding was established using both the DL and the full-length FSS PK. Only a small change in affinity between the two RNAs was observed, suggesting that the cyclic peptides bind structural features shared by the DL and FSS PK RNAs.

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

SHAPE-seq analysis revealed that the peptides localize to the DL sequence, which is expected to show non-canonical base-pairing properties.[78] The importance of maintaining a precise level of frameshifting is underscored by the fact that a single nucleotide mutation in the slippery sequence of the viral RNA abolishes replication.[83] Similarly, point mutants in the DL that ablate dimerization have a deleterious effect on −1 PRF. While the exact mode of cyclic peptide binding remains unknown, our present data indicated that it is localized to the apical loop region (Fig. 5). Binding of cyclic peptides within the DL may be analogous to recognition of the HIV-1 TAR UCU bulge wherein regions of non-canonical base pairing allow arginine- and glutamine-mediated readout of the major groove.[37,84,85] Similar comparisons can be made to the HCV IRES and FMN riboswitch, which recognize small molecules within a helical bulge and multi-helix junction.[17] Importantly, as demonstrated by our *in vitro* binding experiments in the presence of tRNA (Fig. S4†) and our control experiments with unrelated RNA pseudoknots (Fig. S5†), the cyclic peptides isolated through the present strategy show high specificity toward the target RNA molecule.

It is also instructive to compare these FSS targeting cyclic peptides with other compounds previously developed against this RNA target. In previous studies, the best hit isolated from the screening of ∼4000 small molecule drugs was reported to exhibit an IC_{50} of ∼20 μM in a frameshifting reporter assay,[11] while an optimized analog from an initial screen of ∼40 000 small molecules had a K_D of 60 μM against the frameshifting element RNA in an SPR assay.[48] In this context, the (low micromolar) cyclopeptides reported here constitute alternative and promising starting points for the development of cyclopeptide agents directed against this RNA target, *e.g.*, through affinity maturation *via* site-saturation and/or combinatorial mutagenesis as done previously by our groups for other bioactive cyclic peptides.[37,84] Finally, while further studies will be required to assess the activity of these compounds in cellular assays,[86] it is promising that no significant cytotoxicity was observed against mammalian cells (HEK293T) even after incubation of **FSS-m1** at 50 μM for 24 hours (Fig. S6†).

Conclusion

The emergence of RNA as a key regulator of cellular functions and human disease heightens the importance of developing strategies that can accelerate the discovery of molecules that target RNA with high affinity and specificity. While libraries of genetically encoded macrocyclic peptides have provided a valuable source of chemical agents for modulating protein–protein interactions,[41,50–60] their application toward developing RNA-targeting molecules has remained underexplored. This work demonstrates the implementation and validation of a multiplexed MOrPH phage display platform for the discovery of macrocyclic peptides that target SARS-CoV-2 RNA with specificity and low micromolar affinity. Given the importance of the SARS-CoV-2 −1 PRF FSS PK in the viral life cycle and its conservation within the betacoronavirus genus, these results represent a first step toward the development of first-in-class cyclopeptide inhibitors of the −1 PRF pathway. We further envision these compounds can be useful for the development of systems for RNA degraders.[49,77,87] More broadly, these results pave the way to the application of multiplexed MOrPH-PhD for the discovery of chemical entities directed against RNAs and other biomolecular targets.

Data availability

The data supporting this article have been included as part of the ESI.†

Author contributions

R. F. and J. E. W. conceived the project and supervised the work. J. A. I., R. B., A. S. H., A. S. A., and J. L. J. performed the experimental work. J. A. I., R. B., A. S. H., J. E. W., and R. F. wrote the manuscript, with input from the other authors.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by NIH grant R01GM134076 to R. F. and in part by NIH grants R01AI150463 & R01GM063162 to J. E. W. R. F. acknowledges endowed chair support from the Robert A. Welch Foundation (Chair, AT-0051). We thank members of the Wedekind and Fasan labs for technical support and discussions about this work. We thank Dr E. Pritchett and C. Baker (U. Rochester) for assistance with SHAPE-seq and Shelby Phelps in the Dodani Lab at UT Dallas for materials for the cytotoxicity experiments. R. B. recognizes support from NIH training grant T32 GM135134 and an E. Huntington Hooker graduate fellowship. A. S. H. is a trainee in the Medical Scientist Training Program funded by NIH grant T32 GM07356 and T32 GM152318.

References

1 K. V. Morris and J. S. Mattick, The rise of regulatory RNA, *Nat. Rev. Genet.*, 2014, **15**(6), 423–437, DOI: **10.1038/nrg3722**.
2 G. St Laurent, C. Wahlestedt and P. Kapranov, The Landscape of long noncoding RNA classification, *Trends Genet.*, 2015, **31**(5), 239–251, DOI: **10.1016/j.tig.2015.03.007**.
3 K. K. Ebbesen, J. Kjems and T. B. Hansen, Circular RNAs: Identification, biogenesis and function, *Biochim. Biophys. Acta*, 2016, **1859**(1), 163–168, DOI: **10.1016/j.bbagrm.2015.07.007**.
4 M. D. Disney, B. G. Dwyer and J. L. Childs-Disney, Drugging the RNA World, *Cold Spring Harbor Perspect. Biol.*, 2018, **10**(11), a034769, DOI: **10.1101/cshperspect.a034769**.
5 T. R. Cech and J. A. Steitz, The noncoding RNA revolution-trashing old rules to forge new ones, *Cell*, 2014, **157**(1), 77–94, DOI: **10.1016/j.cell.2014.03.008**.
6 P. A. Latos, F. M. Pauler, M. V. Koerner, H. B. Senergin, Q. J. Hudson, R. R. Stocsits, W. Allhoff, S. H. Stricker,

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

R. M. Klement, K. E. Warczok, *et al.*, Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing, *Science*, 2012, **338**(6113), 1469–1472, DOI: 10.1126/science.1228110.

7 C. L. Peebles, P. S. Perlman, K. L. Mecklenburg, M. L. Petrillo, J. H. Tabor, K. A. Jarrell and H. L. Cheng, A self-splicing RNA excises an intron lariat, *Cell*, 1986, **44**(2), 213–223, DOI: 10.1016/0092-8674(86)90755-5.

8 R. Barrangou, C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero and P. Horvath, CRISPR provides acquired resistance against viruses in prokaryotes, *Science*, 2007, **315**(5819), 1709–1712, DOI: 10.1126/science.1138140.

9 T. M. Schmeing, K. S. Huang, S. A. Strobel and T. A. Steitz, An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA, *Nature*, 2005, **438**(7067), 520–524, DOI: 10.1038/nature04152.

10 J. A. Kelly, M. T. Woodside and J. D. Dinman, Programmed -1 Ribosomal Frameshifting in coronaviruses: A therapeutic target, *Virology*, 2021, **554**, 75–82, DOI: 10.1016/j.virol.2020.12.010.

11 Y. Sun, L. Abriola, R. O. Niederer, S. F. Pedersen, M. M. Alfajaro, V. Silva Monteiro, C. B. Wilen, Y. C. Ho, W. V. Gilbert, Y. V. Surovtseva, *et al.*, Restriction of SARS-CoV-2 replication by targeting programmed -1 ribosomal frameshifting, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**(26), e2023051118, DOI: 10.1073/pnas.2023051118.

12 N. N. Patwardhan, L. R. Ganser, G. J. Kapral, C. S. Eubanks, J. Lee, B. Sathyamoorthy, H. M. Al-Hashimi and A. E. Hargrove, Amiloride as a new RNA-binding scaffold with activity against HIV-1 TAR, *Medchemcomm*, 2017, **8**(5), 1022–1036, DOI: 10.1039/c6md00729e.

13 T. A. Hilimire, J. M. Chamberlain, V. Anokhina, R. P. Bennett, O. Swart, J. R. Myers, J. M. Ashton, R. A. Stewart, A. L. Featherston, K. Gates, *et al.*, HIV-1 Frameshift RNA-Targeted Triazoles Inhibit Propagation of Replication-Competent and Multi-Drug-Resistant HIV in Human Cells, *ACS Chem. Biol.*, 2017, **12**(6), 1674–1682, DOI: 10.1021/acschembio.7b00052.

14 J. L. Childs-Disney, X. Yang, Q. M. R. Gibaut, Y. Tong, R. T. Batey and M. D. Disney, Targeting RNA structures with small molecules, *Nat. Rev. Drug Discovery*, 2022, **21**(10), 736–762, DOI: 10.1038/s41573-022-00521-4.

15 C. M. Connelly, M. H. Moon and J. S. Schneekloth Jr, The Emerging Role of RNA as a Therapeutic Target for Small Molecules, *Cell Chem. Biol.*, 2016, **23**(9), 1077–1090, DOI: 10.1016/j.chembiol.2016.05.021.

16 J. P. Falese, A. Donlic and A. E. Hargrove, Targeting RNA with small molecules: from fundamental principles towards the clinic, *Chem. Soc. Rev.*, 2021, **50**(4), 2224–2243, DOI: 10.1039/D0CS01261K.

17 S. S. Chavali, R. Bonn-Breach and J. E. Wedekind, Face-time with TAR: Portraits of an HIV-1 RNA with diverse modes of effector recognition relevant for drug discovery, *J. Biol. Chem.*, 2019, **294**(24), 9326–9341, DOI: 10.1074/jbc.REV119.006860.

18 J. Davila-Calderon, N. N. Patwardhan, L. Y. Chiu, A. Sugarman, Z. Cai, S. R. Penutmutchu, M. L. Li, G. Brewer, A. E. Hargrove and B. S. Tolbert, IRES-targeting small molecule inhibits enterovirus 71 replication *via* allosteric stabilization of a ternary complex, *Nat. Commun.*, 2020, **11**(1), 4775, DOI: 10.1038/s41467-020-18594-3.

19 R. J. Marcheschi, M. Tonelli, A. Kumar and S. E. Butcher, Structure of the HIV-1 frameshift site RNA bound to a small molecule inhibitor of viral replication, *ACS Chem. Biol.*, 2011, **6**(8), 857–864, DOI: 10.1021/cb200082d.

20 M. Zafferani and A. E. Hargrove, Small molecule targeting of biologically relevant RNA tertiary and quaternary structures, *Cell Chem. Biol.*, 2021, **28**(5), 594–609, DOI: 10.1016/j.chembiol.2021.03.003.

21 T. A. Hill, N. E. Shepherd, F. Diness and D. P. Fairlie, Constraining cyclic peptides to mimic protein structure motifs, *Angew Chem. Int. Ed. Engl.*, 2014, **53**(48), 13020–13041, DOI: 10.1002/anie.201401058.

22 E. Valeur, S. M. Gueret, H. Adihou, R. Gopalakrishnan, M. Lemurell, H. Waldmann, T. N. Grossmann and A. T. Plowright, New Modalities for Challenging Targets in Drug Discovery, *Angew Chem. Int. Ed. Engl.*, 2017, **56**(35), 10294–10323, DOI: 10.1002/anie.201611914.

23 T. A. F. Cardote and A. Ciulli, Cyclic and Macrocyclic Peptides as Chemical Tools To Recognise Protein Surfaces and Probe Protein-Protein Interactions, *Chemmedchem*, 2016, **11**(8), 787–794, DOI: 10.1002/cmdc.201500450.

24 M. Buyanova and D. Pei, Targeting intracellular protein-protein interactions with macrocyclic peptides, *Trends Pharmacol. Sci.*, 2022, **43**(3), 234–248, DOI: 10.1016/j.tips.2021.11.008.

25 S. Taguchi and H. Suga, Targeting of extracellular protein-protein interactions with macrocyclic peptides, *Curr. Opin. Chem. Biol.*, 2021, **62**, 82–89, DOI: 10.1016/j.cbpa.2021.02.013.

26 N. S. Robertson and D. R. Spring, Using Peptidomimetics and Constrained Peptides as Valuable Tools for Inhibiting Protein–Protein Interactions, *Molecules*, 2018, **23**(4), 959, DOI: 10.3390/molecules23040959.

27 D. M. Kruger, A. Glas, D. Bier, N. Pospiech, K. Wallraven, L. Dietrich, C. Ottmann, O. Koch, S. Hennig and T. N. Grossmann, Structure-Based Design of Non-natural Macrocyclic Peptides That Inhibit Protein-Protein Interactions, *J. Med. Chem.*, 2017, **60**(21), 8982–8988, DOI: 10.1021/acs.jmedchem.7b01221.

28 R. M. Cardoso, F. M. Brunel, S. Ferguson, M. Zwick, D. R. Burton, P. E. Dawson and I. A. Wilson, Structural basis of enhanced binding of extended and helically constrained peptide epitopes of the broadly neutralizing HIV-1 antibody 4E10, *J. Mol. Biol.*, 2007, **365**(5), 1533–1544, DOI: 10.1016/j.jmb.2006.10.088.

29 C. Ngambenjawong, J. M. B. Pineda and S. H. Pun, Engineering an Affinity-Enhanced Peptide through Optimization of Cyclization Chemistry, *Bioconjugate Chem.*, 2016, **27**(12), 2854–2862, DOI: 10.1021/acs.bioconjchem.6b00502.

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

30 T. L. Aboye, H. Ha, S. Majumder, F. Christ, Z. Debyser, A. Shekhtman, N. Neamati and J. A. Camarero, Design of a novel cyclotide-based CXCR4 antagonist with anti-human immunodeficiency virus (HIV)-1 activity, *J. Med. Chem.*, 2012, **55**(23), 10729–10734, DOI: 10.1021/jm301468k.

31 J. M. Smith, J. R. Frost and R. Fasan, Designer macrocyclic organo-peptide hybrids inhibit the interaction between p53 and HDM2/X by accommodating a functional α-helix, *Chem. Commun.*, 2014, **50**(39), 5027–5030, DOI: 10.1039/c4cc01199f.

32 L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine and S. J. Korsmeyer, Activation of Apoptosis *in Vivo* by a Hydrocarbon-Stapled BH3 Helix, *Science*, 2004, **305**(5689), 1466–1470, DOI: 10.1126/science.1099191.

33 W. M. Hewitt, S. S. F. Leung, C. R. Pye, A. R. Ponkey, M. Bednarek, M. P. Jacobson and R. S. Lokey, Cell-Permeable Cyclic Peptides from Synthetic Libraries Inspired by Natural Products, *J. Am. Chem. Soc.*, 2015, **137**(2), 715–721, DOI: 10.1021/ja508766b.

34 L. Peraro, Z. Zou, K. M. Makwana, A. E. Cummings, H. L. Ball, H. Yu, Y.-S. Lin, B. Levine and J. A. Kritzer, Diversity-Oriented Stapling Yields Intrinsically Cell-Penetrant Inducers of Autophagy, *J. Am. Chem. Soc.*, 2017, **139**(23), 7792–7802, DOI: 10.1021/jacs.7b01698.

35 A. F. L. Schneider, M. Kithil, M. C. Cardoso, M. Lehmann and C. P. R. Hackenberger, Cellular uptake of large biomolecules enabled by cell-surface-reactive cell-penetrating peptide additives, *Nat. Chem.*, 2021, **13**(6), 530–539, DOI: 10.1038/s41557-021-00661-x.

36 F. Milletti, Cell-penetrating peptides: classes, origin, and current landscape, *Drug Discovery Today*, 2012, **17**(15), 850–860, DOI: 10.1016/j.drudis.2012.03.002.

37 S. S. Chavali, S. M. Mali, R. Bonn, A. Saseendran Anitha, R. P. Bennett, H. C. Smith, R. Fasan and J. E. Wedekind, Cyclic peptides with a distinct arginine-fork motif recognize the HIV trans-activation response RNA *in vitro* and in cells, *J. Biol. Chem.*, 2021, **297**(6), 101390, DOI: 10.1016/j.jbc.2021.101390.

38 M. D. Shortridge, P. T. Wille, A. N. Jones, A. Davidson, J. Bogdanovic, E. Arts, J. Karn, J. A. Robinson and G. Varani, An ultra-high affinity ligand of HIV-1 TAR reveals the RNA structure recognized by P-TEFb, *Nucleic Acids Res.*, 2019, **47**(3), 1523–1531, DOI: 10.1093/nar/gky1197.

39 M. D. Shortridge, M. J. Walker, T. Pavelitz, Y. Chen, W. Yang and G. Varani, A Macrocyclic Peptide Ligand Binds the Oncogenic MicroRNA-21 Precursor and Suppresses Dicer Processing, *ACS Chem. Biol.*, 2017, **12**(6), 1611–1620, DOI: 10.1021/acschembio.7b00180.

40 A. K. Manna, A. Kumar, U. Ray, S. Das, G. Basu and S. Roy, A cyclic peptide mimic of an RNA recognition motif of human La protein is a potent inhibitor of hepatitis C virus, *Antiviral Res.*, 2013, **97**(3), 223–226, DOI: 10.1016/j.antiviral.2012.12.026.

41 A. E. Owens, J. A. Iannuzzelli, Y. Gu and R. Fasan, MOrPH-PhD: An Integrated Phage Display Platform for the Discovery of Functional Genetically Encoded Peptide Macrocycles, *ACS Cent. Sci.*, 2020, **6**(3), 368–381, DOI: 10.1021/acscentsci.9b00927.

42 N. Bionda, A. L. Cryan and R. Fasan, Bioinspired strategy for the ribosomal synthesis of thioether-bridged macrocyclic peptides in bacteria, *ACS Chem. Biol.*, 2014, **9**(9), 2008–2013, DOI: 10.1021/cb500311k.

43 J. A. Iannuzzelli and R. Fasan, Expanded toolbox for directing the biosynthesis of macrocyclic peptides in bacterial cells, *Chem. Sci.*, 2020, **11**(24), 6202–6208, DOI: 10.1039/d0sc01699c.

44 P. V'Kovski, A. Kratzel, S. Steiner, H. Stalder and V. Thiel, Coronavirus biology and replication: implications for SARS-CoV-2, *Nat. Rev. Microbiol.*, 2021, **19**(3), 155–170, DOI: 10.1038/s41579-020-00468-6.

45 J. A. Kelly, A. N. Olson, K. Neupane, S. Munshi, J. San Emeterio, L. Pollack, M. T. Woodside and J. D. Dinman, Structural and functional conservation of the programmed -1 ribosomal frameshift signal of SARS coronavirus 2 (SARS-CoV-2), *J. Biol. Chem.*, 2020, **295**(31), 10741–10748, DOI: 10.1074/jbc.AC120.013449.

46 P. R. Bhatt, A. Scaiola, G. Loughran, M. Leibundgut, A. Kratzel, R. Meurs, R. Dreos, K. M. O'Connor, A. McMillan, J. W. Bode, *et al.*, Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome, *Science*, 2021, **372**(6548), 1306–1313, DOI: 10.1126/science.abf3546.

47 Y. H. Wang, M. Grunewald and S. Perlman, Coronaviruses: An Updated Overview of Their Replication and Pathogenesis, *Methods Mol. Biol.*, 2020, **2203**, 1–29, DOI: 10.1007/978-1-0716-0900-2_1.

48 M. Yang, F. P. Olatunji, C. Rhodes, S. Balaratnam, K. Dunne-Dombrink, S. Seshadri, X. Liang, C. P. Jones, S. F. J. Le Grice, A. R. Ferre-D'Amare, *et al.*, Discovery of Small Molecules Targeting the Frameshifting Element RNA in SARS-CoV-2 Viral Genome, *ACS Med. Chem. Lett.*, 2023, **14**(6), 757–765, DOI: 10.1021/acsmedchemlett.3c00051.

49 H. S. Haniff, Y. Tong, X. Liu, J. L. Chen, B. M. Suresh, R. J. Andrews, J. M. Peterson, C. A. O'Leary, R. I. Benhamou, W. N. Moss, *et al.*, Targeting the SARS-CoV-2 RNA Genome with Small Molecule Binders and Ribonuclease Targeting Chimera (RIBOTAC) Degraders, *ACS Cent. Sci.*, 2020, **6**(10), 1713–1721, DOI: 10.1021/acscentsci.0c00984.

50 A. Tavassoli, SICLOPPS cyclic peptide libraries in drug discovery, *Curr. Opin. Chem. Biol.*, 2017, **38**, 30–35, DOI: 10.1016/j.cbpa.2017.02.016.

51 K. J. Hetrick, M. C. Walker and W. A. van der Donk, Development and Application of Yeast and Phage Display of Diverse Lanthipeptides, *ACS Cent. Sci.*, 2018, **4**(4), 458–467, DOI: 10.1021/acscentsci.7b00581.

52 M. Zheng, F.-J. Chen, K. Li, R. M. Reja, F. Haeffner and J. Gao, Lysine-Targeted Reversible Covalent Ligand Discovery for Proteins *via* Phage Display, *J. Am. Chem. Soc.*, 2022, **144**(34), 15885–15893, DOI: 10.1021/jacs.2c07375.

53 A. R. Horswill, S. N. Savinov and S. J. Benkovic, A systematic method for identifying small-molecule modulators of

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

protein-protein interactions, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**(44), 15591–15596, DOI: **10.1073/pnas.0406999101**.

54 S. M. Howell, S. V. Fiacco, T. T. Takahashi, F. Jalali-Yazdi, S. W. Millward, B. Hu, P. Wang and R. W. Roberts, Serum stable natural peptides designed by mRNA display, *Sci. Rep.*, 2014, **4**(1), 6008, DOI: **10.1038/srep06008**.

55 C. J. Hipolito and H. Suga, Ribosomal production and *in vitro* selection of natural product-like peptidomimetics: the FIT and RaPID systems, *Curr. Opin. Chem. Biol.*, 2012, **16**(1–2), 196–203, DOI: **10.1016/j.cbpa.2012.02.014**.

56 P. Diderich and C. Heinis, Phage selection of bicyclic peptides binding Her2, *Tetrahedron*, 2014, **70**(42), 7733–7739, DOI: **10.1016/j.tet.2014.05.106**.

57 K. Deyle, X.-D. Kong and C. Heinis, Phage Selection of Cyclic Peptides for Application in Research and Drug Development, *Acc. Chem. Res.*, 2017, **50**(8), 1866–1874, DOI: **10.1021/acs.accounts.7b00184**.

58 S. S. Kale, C. Villequey, X.-D. Kong, A. Zorzi, K. Deyle and C. Heinis, Cyclization of peptides with two chemical bridges affords large scaffold diversities, *Nat. Chem.*, 2018, **10**(7), 715–723, DOI: **10.1038/s41557-018-0042-7**.

59 S. Ng, M. R. Jafari and R. Derda, Bacteriophages and Viruses as a Support for Organic Synthesis and Combinatorial Chemistry, *ACS Chem. Biol.*, 2012, **7**(1), 123–138, DOI: **10.1021/cb200342h**.

60 T. R. Oppewal, I. D. Jansen, J. Hekelaar and C. Mayer, A Strategy to Select Macrocyclic Peptides Featuring Asymmetric Molecular Scaffolds as Cyclization Units by Phage Display, *J. Am. Chem. Soc.*, 2022, **144**(8), 3644–3652, DOI: **10.1021/jacs.1c12822**.

61 K. C. Liu, K. Roder, C. Mayer, S. Adhikari, D. J. Wales and S. Balasubramanian, Affinity-Selected Bicyclic Peptide G-Quadruplex Ligands Mimic a Protein-like Binding Mechanism, *J. Am. Chem. Soc.*, 2020, **142**(18), 8367–8373, DOI: **10.1021/jacs.0c01879**.

62 V. A. Burns, B. G. Bobay, A. Basso, J. Cavanagh and C. Melander, Targeting RNA with cysteine-constrained peptides, *Bioorg. Med. Chem. Lett.*, 2008, **18**(2), 565–567, DOI: **10.1016/j.bmcl.2007.11.096**.

63 D. M. Kruger, S. Neubacher and T. N. Grossmann, Protein-RNA interactions: structural characteristics and hotspot amino acids, *RNA*, 2018, **24**(11), 1457–1465, DOI: **10.1261/rna.066464.118**.

64 S. Jones, D. T. Daley, N. M. Luscombe, H. M. Berman and J. M. Thornton, Protein-RNA interactions: a structural analysis, *Nucleic Acids Res.*, 2001, **29**(4), 943–954, DOI: **10.1093/nar/29.4.943**.

65 Y. Yi, Y. Zhao, Y. Huang and D. Wang, A Brief Review of RNA-Protein Interaction Database Resources, *Noncoding RNAs*, 2017, **3**(1), 6, DOI: **10.3390/ncrna3010006**.

66 Y. Gu, J. A. Iannuzzelli and R. Fasan, MOrPH-PhD: A Phage Display System for the Functional Selection of Genetically Encoded Macrocyclic Peptides, *Methods Mol. Biol.*, 2022, **2371**, 261–286, DOI: **10.1007/978-1-0716-1689-5_14**.

67 D. Ishimaru, E. P. Plant, A. C. Sims, B. L. Yount Jr, B. M. Roth, N. V. Eldho, G. C. Perez-Alvarado, D. W. Armbruster, R. S. Baric, J. D. Dinman, *et al.*, RNA dimerization plays a role in ribosomal frameshifting of the SARS coronavirus, *Nucleic Acids Res.*, 2013, **41**(4), 2594–2608, DOI: **10.1093/nar/gks1361**.

68 J. L. Jenkins, J. Krucinska, R. M. McCarty, V. Bandarian and J. E. Wedekind, Comparison of a preQ$_1$ riboswitch aptamer in metabolite-bound and free states with implications for gene regulation, *J. Biol. Chem.*, 2011, **286**(28), 24626–24637, DOI: **10.1074/jbc.M111.230375**.

69 G. M. Schroeder, C. E. Cavender, M. E. Blau, J. L. Jenkins, D. H. Mathews and J. E. Wedekind, A small RNA that cooperatively senses two stacked metabolites in one pocket for gene control, *Nat. Commun.*, 2022, **13**(1), 199, DOI: **10.1038/s41467-021-27790-8**.

70 S. Hamal Dhakal, K. Kavita, S. S. S. Panchapakesan, A. Roth and R. R. Breaker, 8-oxoguanine riboswitches in bacteria detect and respond to oxidative DNA damage, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, **120**(40), e2307854120, DOI: **10.1073/pnas.2307854120**.

71 D. Loughrey, K. E. Watters, A. H. Settle and J. B. Lucks, SHAPE-Seq 2.0: systematic optimization and extension of high-throughput chemical probing of RNA secondary structure with next generation sequencing, *Nucleic Acids Res.*, 2014, **42**(21), e165, DOI: **10.1093/nar/gku909**.

72 K. A. Wilkinson, E. J. Merino and K. M. Weeks, Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution, *Nat. Protoc.*, 2006, **1**(3), 1610–1616, DOI: **10.1038/nprot.2006.249**.

73 R. C. Spitale, P. Crisalli, R. A. Flynn, E. A. Torre, E. T. Kool and H. Y. Chang, RNA SHAPE analysis in living cells, *Nat. Chem. Biol.*, 2013, **9**(1), 18–20, DOI: **10.1038/nchembio.1131**.

74 J. L. McGinnis, J. A. Dunkle, J. H. Cate and K. M. Weeks, The mechanisms of RNA SHAPE chemistry, *J. Am. Chem. Soc.*, 2012, **134**(15), 6617–6624, DOI: **10.1021/ja2104075**.

75 S. S. Chavali, C. E. Cavender, D. H. Mathews and J. E. Wedekind, Arginine Forks Are a Widespread Motif to Recognize Phosphate Backbones and Guanine Nucleobases in the RNA Major Groove, *J. Am. Chem. Soc.*, 2020, **142**(47), 19835–19839, DOI: **10.1021/jacs.0c09689**.

76 K. D. Warner, C. E. Hajdin and K. M. Weeks, Principles for targeting RNA with drug-like small molecules, *Nat. Rev. Drug Discovery*, 2018, **17**(8), 547–558, DOI: **10.1038/nrd.2018.93**.

77 Y. Tong, Y. Lee, X. Liu, J. L. Childs-Disney, B. M. Suresh, R. I. Benhamou, C. Yang, W. Li, M. G. Costales, H. S. Haniff, *et al.*, Programming inactive RNA-binding small molecules into bioactive degraders, *Nature*, 2023, **618**(7963), 169–179, DOI: **10.1038/s41586-023-06091-8**.

78 K. Zhang, I. N. Zheludev, R. J. Hagey, R. Haslecker, Y. J. Hou, R. Kretsch, G. D. Pintilie, R. Rangan, W. Kladwang, S. Li, *et al.*, Cryo-EM and antisense targeting of the 28-kDa frameshift stimulation element from the SARS-CoV-2 RNA genome, *Nat. Struct. Mol. Biol.*, 2021, **28**(9), 747–754, DOI: **10.1038/s41594-021-00653-y**.

79 N. C. Huston, H. Wan, M. S. Strine, R. de Cesaris Araujo Tavares, C. B. Wilen and A. M. Pyle, Comprehensive *in vivo* secondary structure of the SARS-CoV-2 genome reveals

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 October 2024. Downloaded on 1/31/2025 4:34:35 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

novel regulatory motifs and mechanisms, *Mol. Cell*, 2021, **81**(3), 584–598, DOI: **10.1016/j.molcel.2020.12.041**.

80 C. P. Jones and A. R. Ferre-D'Amare, Crystal structure of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) frameshifting pseudoknot, *RNA*, 2022, **28**(2), 239–249, DOI: **10.1261/rna.078825.121**.

81 C. Roman, A. Lewicka, D. Koirala, N. S. Li and J. A. Piccirilli, The SARS-CoV-2 Programmed -1 Ribosomal Frameshifting Element Crystal Structure Solved to 2.09 A Using Chaperone-Assisted RNA Crystallography, *ACS Chem. Biol.*, 2021, **16**(8), 1469–1481, DOI: **10.1021/acschembio.1c00324**.

82 T. C. T. Lan, M. F. Allan, L. E. Malsick, J. Z. Woo, C. Zhu, F. Zhang, S. Khandwala, S. S. Y. Nyeo, Y. Sun, J. U. Guo, *et al.*, Secondary structural ensembles of the SARS-CoV-2 RNA genome in infected cells, *Nat. Commun.*, 2022, **13**(1), 1128, DOI: **10.1038/s41467-022-28603-2**.

83 E. P. Plant, R. Rakauskaite, D. R. Taylor and J. D. Dinman, Achieving a golden mean: mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins, *J. Virol.*, 2010, **84**(9), 4330–4340, DOI: **10.1128/JVI.02480-09**.

84 S. S. Chavali, S. M. Mali, J. L. Jenkins, R. Fasan and J. E. Wedekind, Co-crystal structures of HIV TAR RNA bound to lab-evolved proteins show key roles for arginine relevant to the design of cyclic peptide TAR inhibitors, *J. Biol. Chem.*, 2020, **295**(49), 16470–16486, DOI: **10.1074/jbc.RA120.015444**.

85 I. A. Belashov, D. W. Crawford, C. E. Cavender, P. Dai, P. C. Beardslee, D. H. Mathews, B. L. Pentelute, B. R. McNaughton and J. E. Wedekind, Structure of HIV TAR in complex with a Lab-Evolved RRM provides insight into duplex RNA recognition and synthesis of a constrained peptide that impairs transcription, *Nucleic Acids Res.*, 2018, **46**(13), 6401–6415, DOI: **10.1093/nar/gky529**.

86 M. J. Waring, Lipophilicity in drug discovery, *Expert Opin. Drug Discovery*, 2010, **5**(3), 235–248, DOI: **10.1517/17460441003605098**.

87 M. G. Costales, Y. Matsumoto, S. P. Velagapudi and M. D. Disney, Small Molecule Targeted Recruitment of a Nuclease to RNA, *J. Am. Chem. Soc.*, 2018, **140**(22), 6741–6744, DOI: **10.1021/jacs.8b01233**.

© 2024 The Author(s). Published by the Royal Society of Chemistry

Supplementary Information (SI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2024

Supporting Information

Cyclic peptides targeting the SARS-CoV-2 programmed ribosomal frameshifting RNA from a multiplexed phage display library

Jacob A. Iannuzzelli[1,4,‡], Rachel Bonn[2,3,5,‡], Andrew S. Hong[1,‡], Abhijith Saseendran Anitha[1,6], Jermaine L. Jenkins[2,3], Joseph E. Wedekind[2,3,*] and Rudi Fasan[1,6,*]

[1] Department of Chemistry, University of Rochester, Rochester, NY 14627, United States

[2] Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, NY, 14642, USA

[3] Center for RNA Biology, University of Rochester School of Medicine and Dentistry, Rochester, NY, 14642, USA

[4] Present Address: Merck & Co., Inc., Rahway, NJ 07065, USA

[5] Present Address: Renaissance School of Medicine, Stony Brook University, Stony Brook, NY 11794, USA

[6] Department of Chemistry & Biochemistry, The University of Texas at Dallas, Richardson, TX 75080, USA

[‡] These authors contributed equally to this work.

* To whom correspondence should be addressed. Joseph Wedekind, Email: joseph.wedekind@rochester.edu; Rudi Fasan, E-mail: Rudi.Fasan@UTDallas.edu

Table of Contents

Supplementary Figures……S3-S15

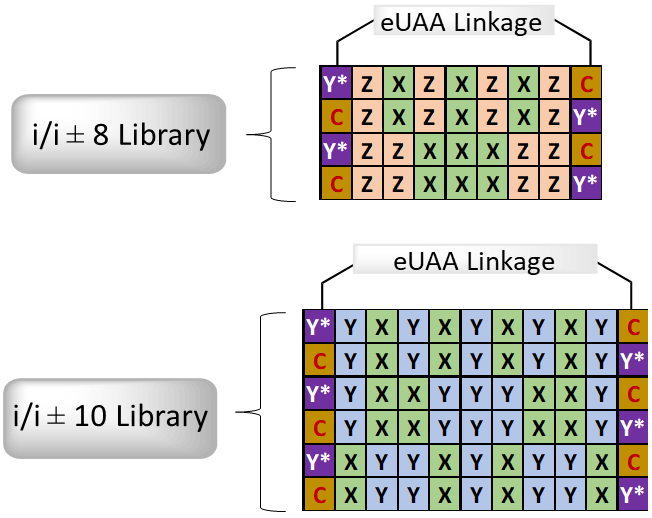
Supplementary Tables……S16-S22

Experimental Methods……S23-S46

NMR Spectra……S47-S55

References..……...……...S56

Supplemental Figures



Symbol	Codon	AAs
X	MRW	R, N, Q, H, K, S
Y	YWT	H, L, F, Y
Z	HWS	N, Q, H, I, L, K, M, F, Y

- ~6.5 X 10^7 unique DNA sequences
- ~1.4 X 10^7 unique cyclic peptides

Figure S1. Design of MOrPH-PhD Libraries. Libraries were diversified by varying the ring size, orientation of the eUAA linkage and the permutation of the degenerate codons.

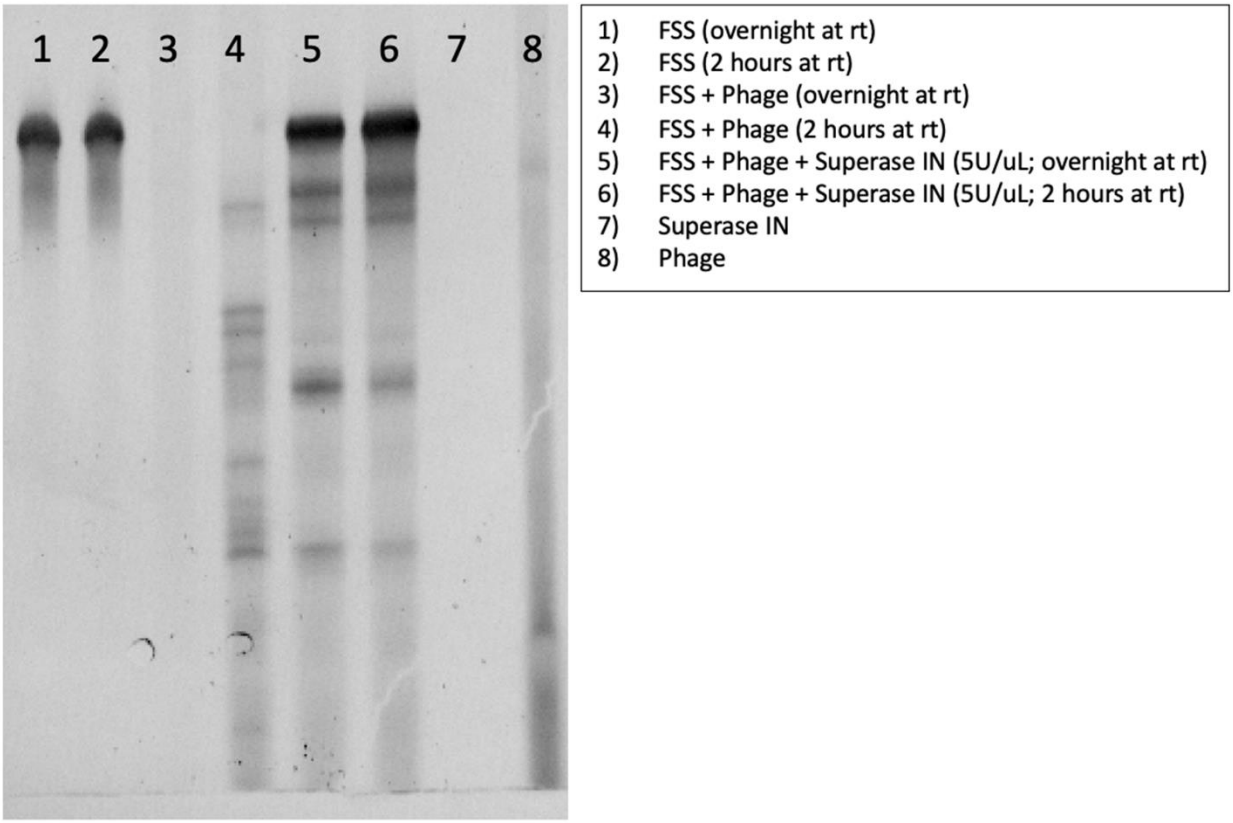


Figure S2. FSS degradation experiments. FSS PK RNA (2 mL; 250 ng/mL) (Figure 1) was incubated with 1 mL of purified phage solution used for panning experiments. Addition of a Superase IN RNase inhibitor (5 U/mL) resulted in significant reduction in FSS degradation after incubation at room temperature (rt) for 2 h (lane 6) and overnight (lane 5).

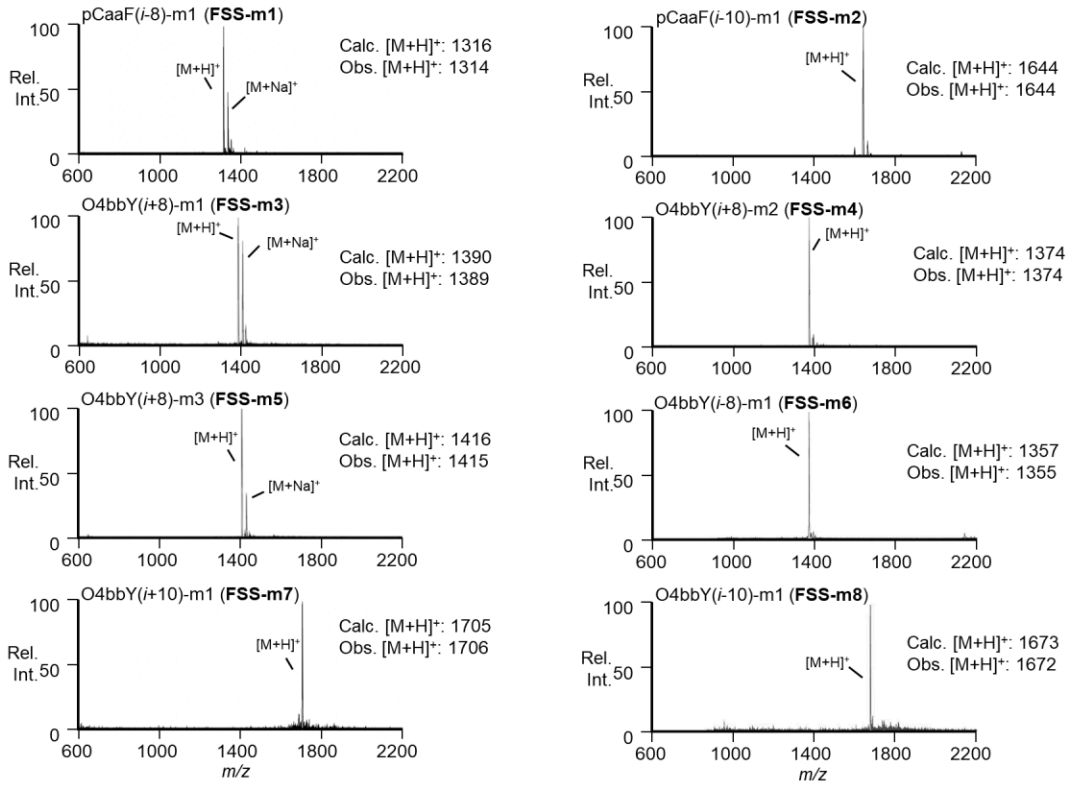


Figure S3. MALDI-TOF-MS spectra of purified synthetic cyclic peptides.

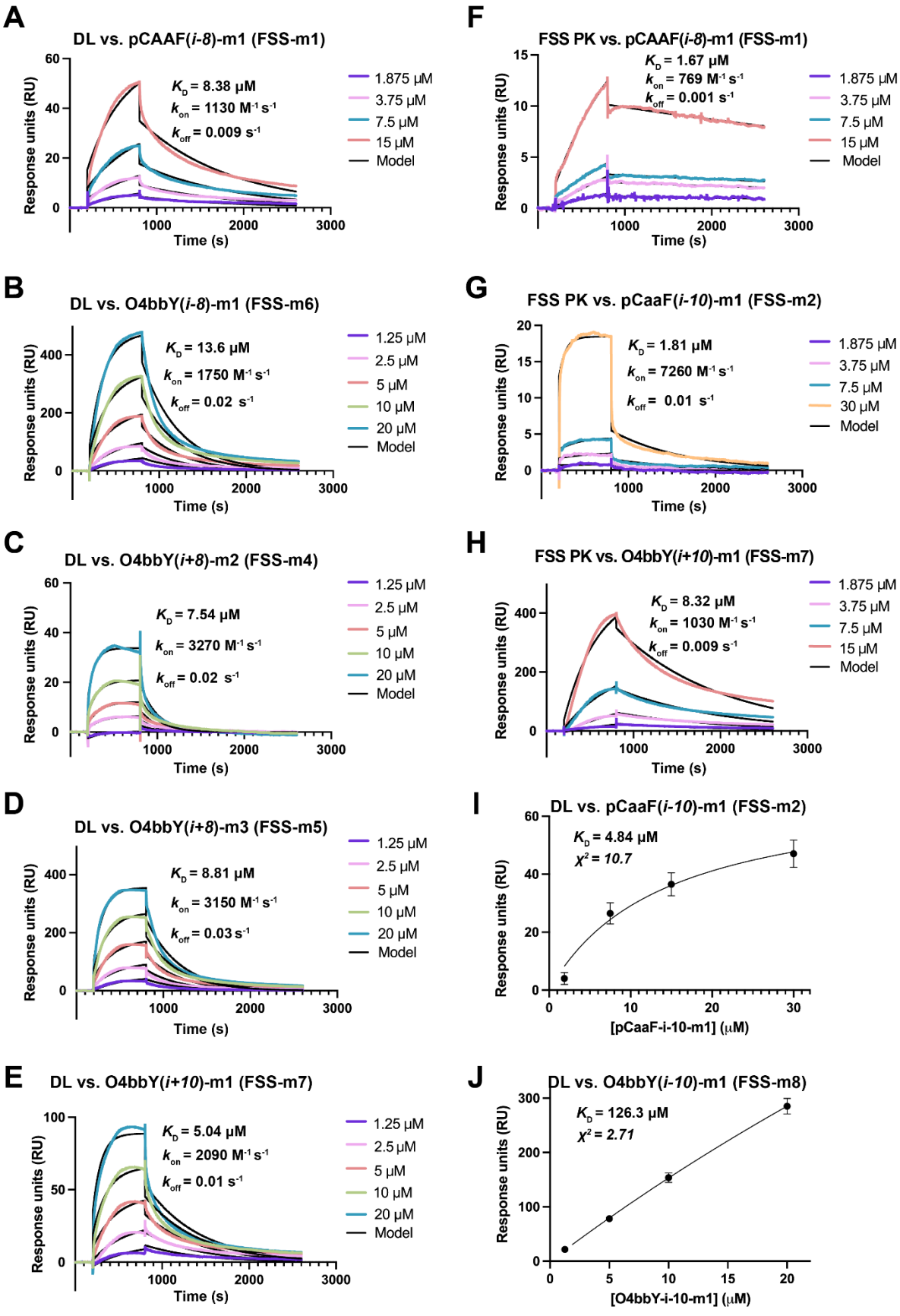

A
DL vs. pCAAF(i-8)-m1 (FSS-m1)
K_D = 8.38 μM
k_on = 1130 M^-1 s^-1
k_off = 0.009 s^-1
1.875 μM
3.75 μM
7.5 μM
15 μM
Model
Response units (RU)
Time (s)
B
DL vs. O4bbY(i-8)-m1 (FSS-m6)
K_D = 13.6 μM
k_on = 1750 M^-1 s^-1
k_off = 0.02 s^-1
1.25 μM
2.5 μM
5 μM
10 μM
20 μM
Model
C
DL vs. O4bbY(i+8)-m2 (FSS-m4)
K_D = 7.54 μM
k_on = 3270 M^-1 s^-1
k_off = 0.02 s^-1
D
DL vs. O4bbY(i+8)-m3 (FSS-m5)
K_D = 8.81 μM
k_on = 3150 M^-1 s^-1
k_off = 0.03 s^-1
E
DL vs. O4bbY(i+10)-m1 (FSS-m7)
K_D = 5.04 μM
k_on = 2090 M^-1 s^-1
k_off = 0.01 s^-1
F
FSS PK vs. pCAAF(i-8)-m1 (FSS-m1)
K_D = 1.67 μM
k_on = 769 M^-1 s^-1
k_off = 0.001 s^-1
G
FSS PK vs. pCaaF(i-10)-m1 (FSS-m2)
K_D = 1.81 μM
k_on = 7260 M^-1 s^-1
k_off = 0.01 s^-1
30 μM
H
FSS PK vs. O4bbY(i+10)-m1 (FSS-m7)
K_D = 8.32 μM
k_on = 1030 M^-1 s^-1
k_off = 0.009 s^-1
I
DL vs. pCaaF(i-10)-m1 (FSS-m2)
K_D = 4.84 μM
χ^2 = 10.7
[pCaaF-i-10-m1] (μM)
J
DL vs. O4bbY(i-10)-m1 (FSS-m8)
K_D = 126.3 μM
χ^2 = 2.71
[O4bbY-i-10-m1] (μM)

Figure S4. Representative kinetic and equilibrium binding sensorgrams and curve fits of cyclic peptide binding to the dimerization loop or full-length frameshift stimulatory sequence pseudoknot. (a) pCaaF(*i*-8)-m1 (**FSS-m1**) association with and dissociation from immobilized SARS-CoV-2 dimerization loop (DL). Here and elsewhere peptide concentrations are shown in the key; colored lines represent background subtracted data; black lines indicate the local fit to a 1:1 binding model. For this and other experiments, the average k_{on} and k_{off} rate constants and apparent K_D value from replicate runs are reported in Tables 1 and 2. (b) O4bbY(*i*+8)-m1 (**FSS-m3**) association and dissociation from immobilized DL. (c) O4bbY(*i*+8)-m2 (**FSS-m4**) association and dissociation from immobilized DL. (d) O4bbY(*i*+8)-m3 (**FSS-m5**) association and dissociation from immobilized DL. (e) O4bbY(*i*+10)-m1 (**FSS-m7**) association and dissociation from immobilized DL. (f) pCaaF(*i*-8)-m1 (**FSS-m1**) association and dissociation from immobilized SARS-CoV-2 67-mer frameshifting stimulatory sequence (FSS). (g) pCaaF(*i*-10)-m1 (**FSS-m2**) association and dissociation from immobilized SARS-CoV-2 67-mer FSS PK. (h) O4bbY(*i*+10)-m1 (**FSS-m7**) association and dissociation from immobilized SARS-CoV-2 67-mer FSS. (i) Equilibrium binding analysis of pCaaF(*i*-10)-m1 (**FSS-m2**) interacting with immobilized DL. (j) Equilibrium binding analysis of O4bbY(*i*-10)-m1 (**FSS-m8**) interacting with immobilized DL.

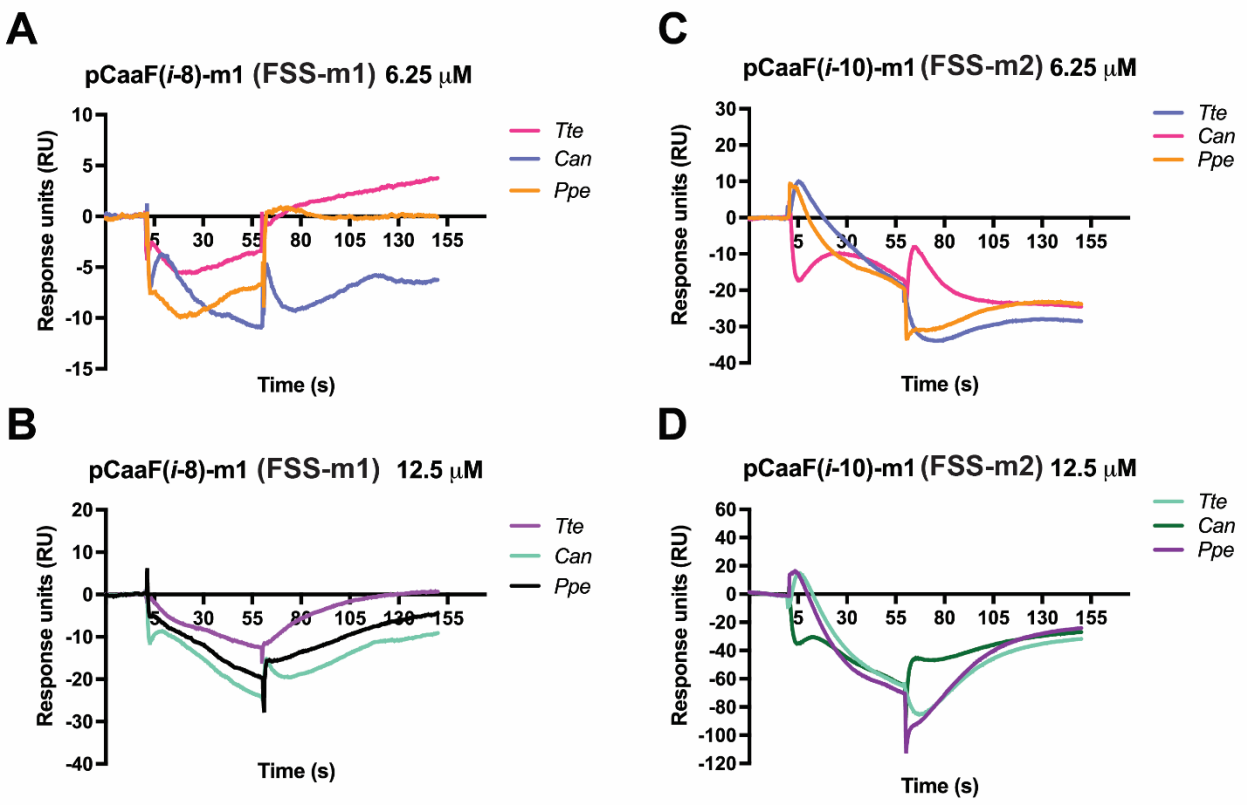


Figure S5. Representative sensorgrams for cyclic peptide binding to off-target RNAs known to adopt pseudoknot folds. (a-b) 6.25 μM or 12.5 μM pCaaF(*i*-8)-m1 (**FSS-m1**) was passed over the *Tte* (preQ$_1$-I type II), *Can* (preQ$_1$-I type I) and *Ppe* (guanine-I variant) riboswitches, which were immobilized on the chip. (c-d) 6.25 μM or 12.5 μM pCaaF(*i*-10)-m1 (**FSS-m2**) was passed over the *Tte* (preQ$_1$-I type II), *Can* (preQ$_1$-I type I) and *Ppe* (guanine-I variant) riboswitches. The absence of specific binding in the sensorgrams is apparent from the negative response units, which result from subtracting the reference cell signal from the sample cell to account for non-specific analyte interactions with the chip wherein no RNA was immobilized. In each case, there is no apparent binding as the cyclic peptide is flowed over a chip containing folded, immobilized riboswitch RNA.

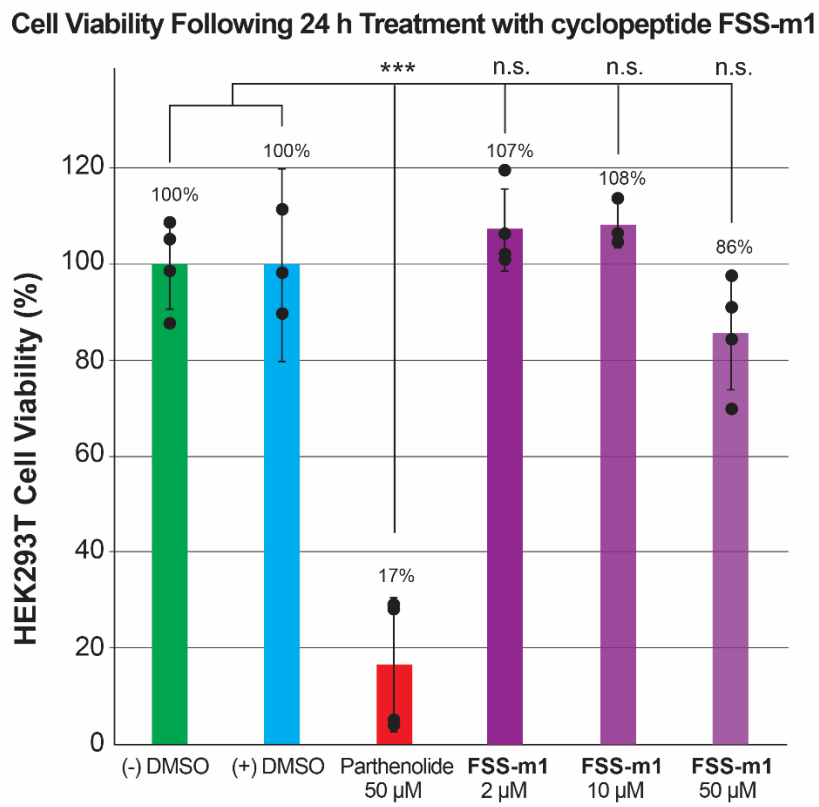
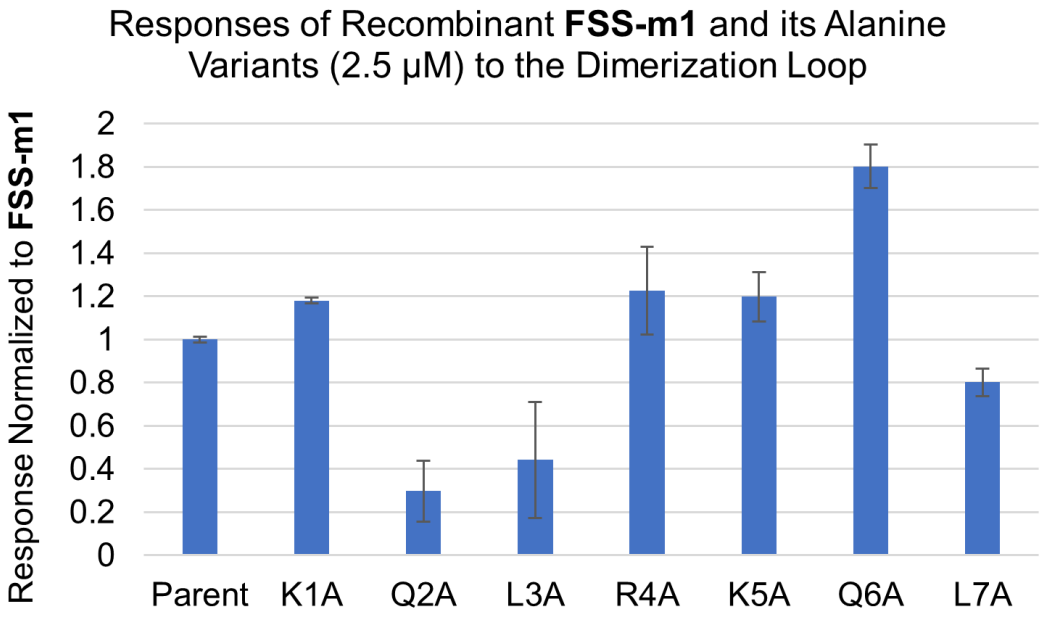
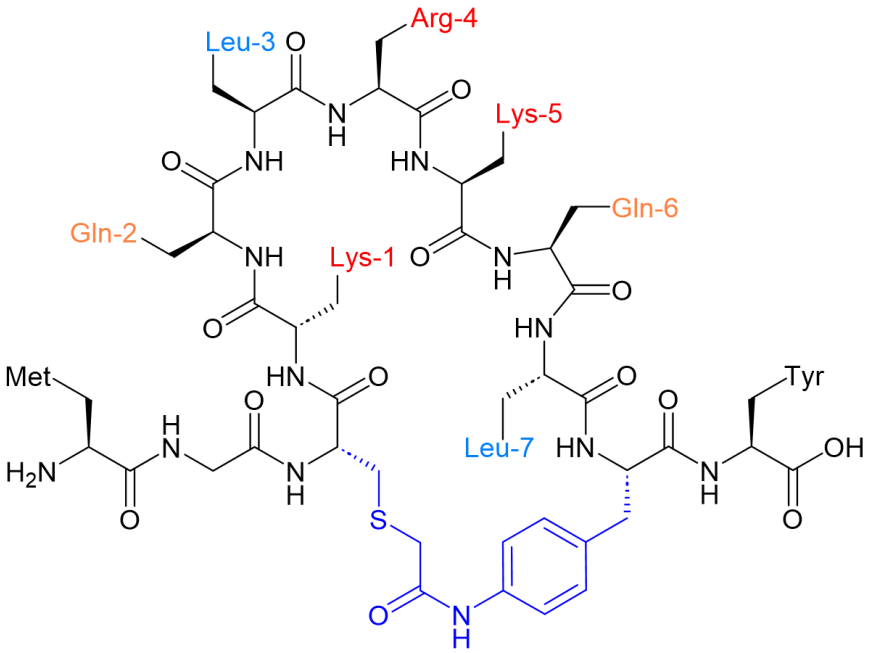


Figure S6. Viability of HEK293T cells upon treatment (24 h) with cyclopeptide **FSS-m1**. (-) DMSO = untreated cells. (+) DMSO = cells treated with vehicle only (cell culture medium +1% (v/v) DMSO). Cells were treated with parthenolide at 50 µM (positive control) or **FSS-m1** at a concentration of 2 µM, 10 µM or 50 µM in vehicle. Analyses were conducted at least in triplicate. No statistically significant difference in cell viability (n.s.) was observed for all three concentrations of **FSS-m1** compared to control cells grown in the absence (p = 0.11 for 50 µM **FSS-m1**) or presence of DMSO (p = 0.17 for 50 µM **FSS-m1**). By contrast, a statistically significant decrease in cell viability was observed for cells grown in the presence of 50 µM parthenolide ($p = 6 \times 10^{-5}$ vs. (-) DMSO cells and $p = 4 \times 10^{-4}$ vs. (+) DMSO cells).

a



b



c

	Parent	K1A	Q2A	L3A	R4A	K5A	Q6A	L7A
Avg	1	1.18	0.30	0.44	1.23	1.20	1.80	0.80
SD	0.01	0.01	0.14	0.27	0.20	0.11	0.10	0.06

Figure S7. Binding of recombinant **FSS-m1** and its alanine variants to the dimerization loop. (a) Relative response of the alanine variants of recombinant peptide **FSS-m1** (pCaaF(*i*-8)-m1) normalized to the parent peptide based on the SPR response value for each of the peptides at a concentration of 2.5 μM upon binding to the dimerization loop. Each bar represents the average of two replicates with standard deviations shown. (b) Chemical structure of the parent recombinant peptide **FSS-m1**. (c) Average and standard deviation values of the normalized SPR response for each variant.

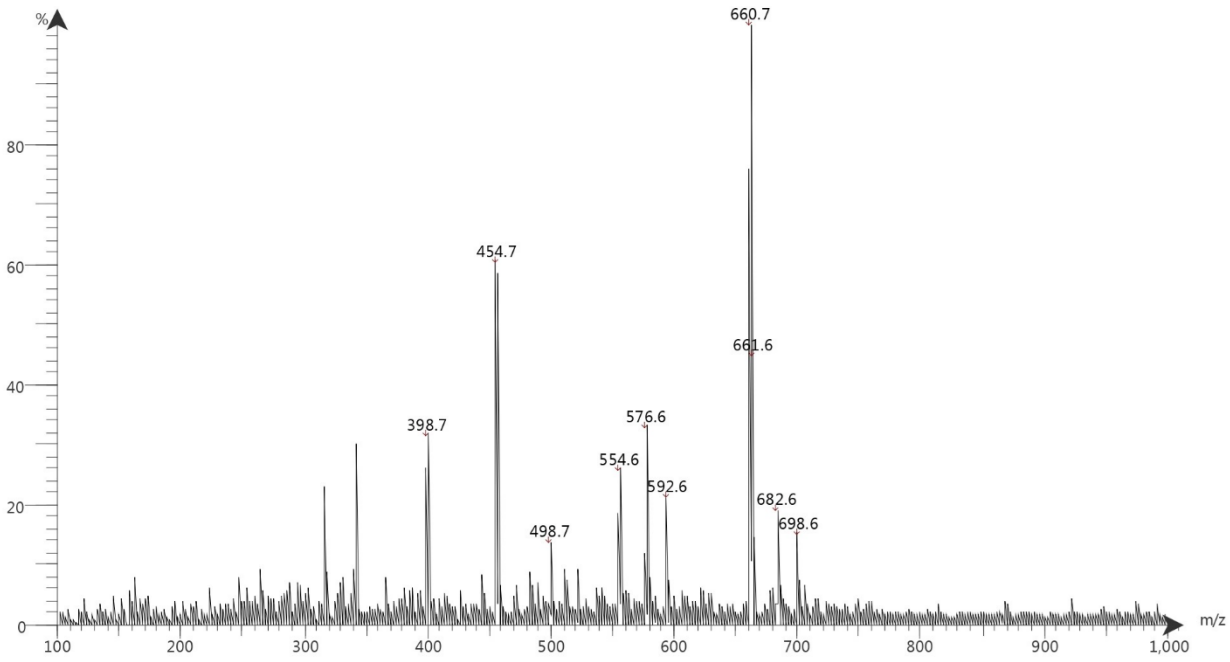
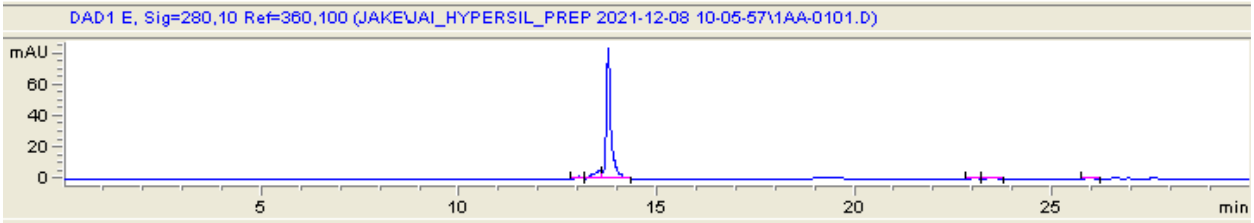


Figure S8. LC/MS spectrum for dipeptide **6**. Calc. $[M+H]^+$ *m/z* = 661.8 Da, obs. $[M+H]^+$ *m/z* = 661. 6 Da.

Figure S9. Representative high-performance liquid chromatography (HPLC) for the synthetic peptides.

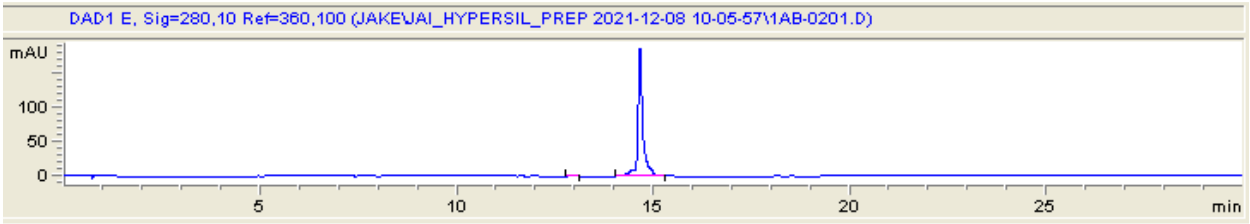
FSS-m1

C18 RP-HPLC method: Linear gradient of 5–95% Acetonitrile (0.1% TFA)/ddH_2O (0.1% TFA), *t* = 1–21 min. Purity: 90%



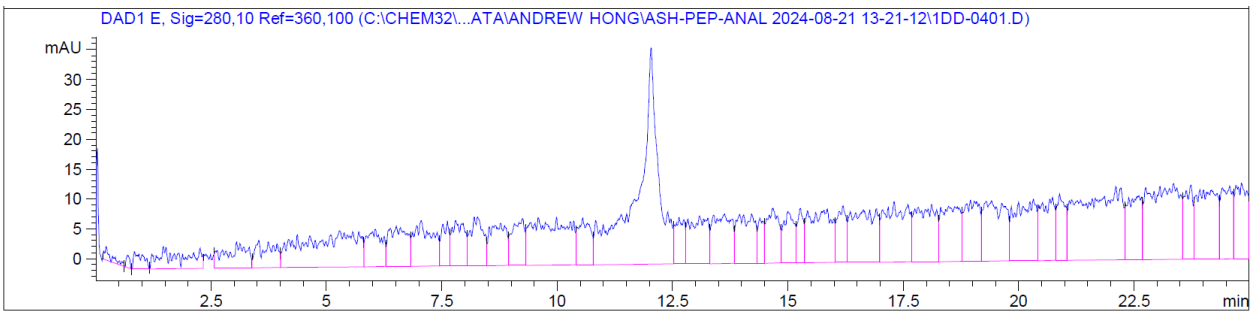
FSS-m2

C18 RP-HPLC method: Linear gradient of 5–95% Acetonitrile (0.1% TFA)/ddH_2O (0.1% TFA), *t* = 1–21 min. Purity: 92%



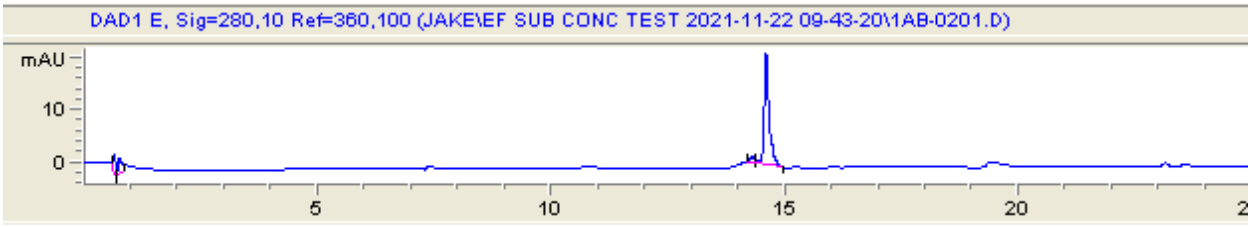
FSS-m3

C18 RP-HPLC method: Linear gradient of 5–95% Acetonitrile (0.1% TFA)/ddH_2O (0.1% TFA), *t* = 1–21 min. Purity: 90%



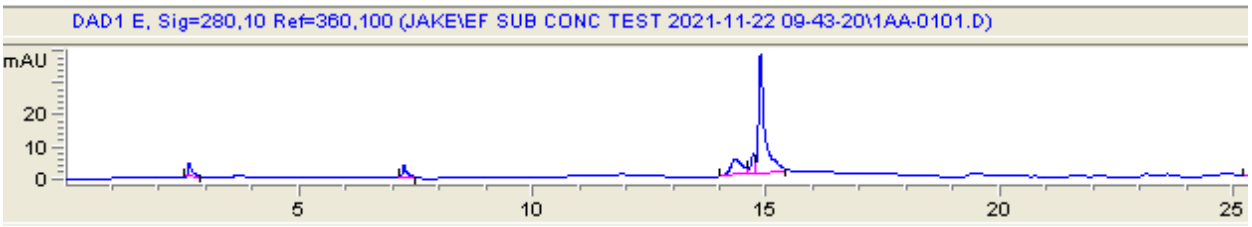
FSS-m4

C18 RP-HPLC method: Linear gradient of 5–95% Acetonitrile (0.1% TFA)/ddH_2O (0.1% TFA), *t* = 1–21 min. Purity: >90%



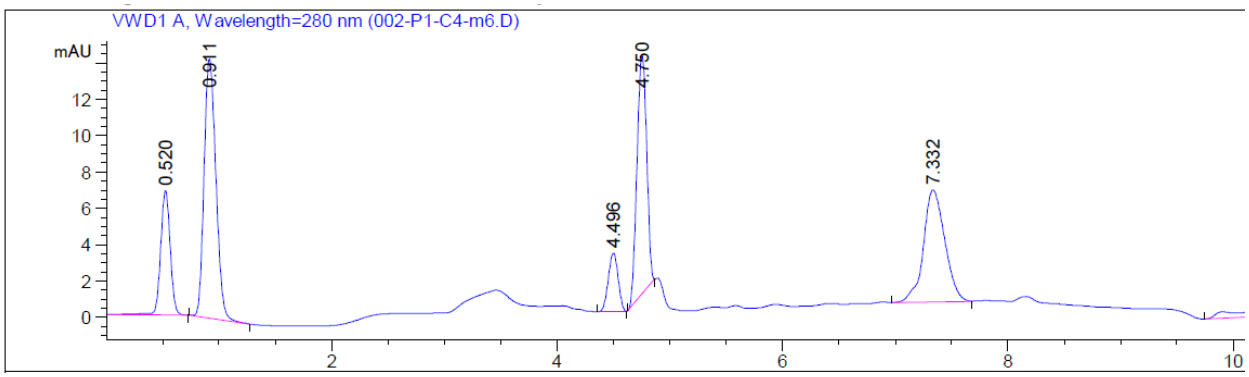
FSS-m5

C18 RP-HPLC method: Linear gradient of 5–95% Acetonitrile (0.1% TFA)/ddH_2O (0.1% TFA), *t* = 1–21 min. Purity: >80%



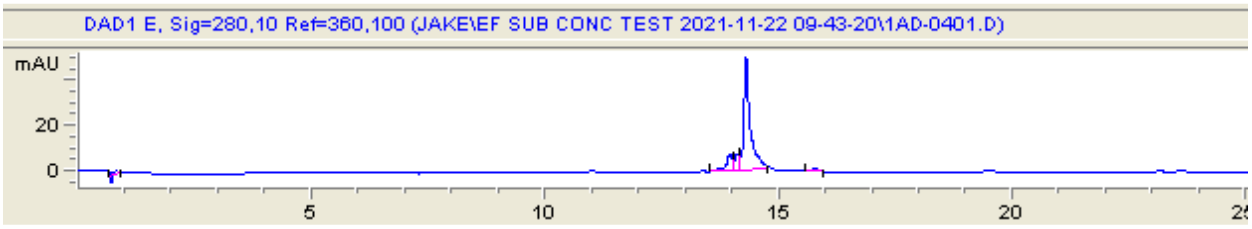
FSS-m6

C18 RP-HPLC method: Linear gradient of 5–95% Acetonitrile (0.1% formic acid)/ddH_2O (0.1% formic acid), *t* = 2–7 min. Purity: 80%



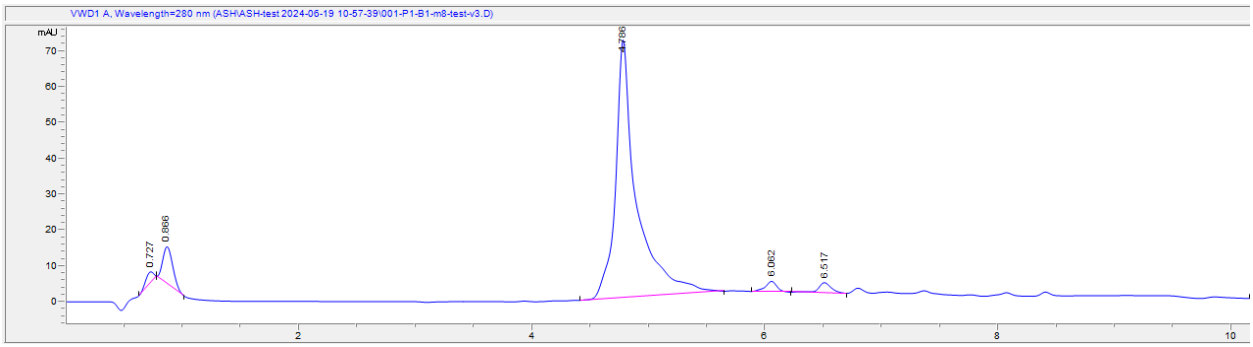
FSS-m7

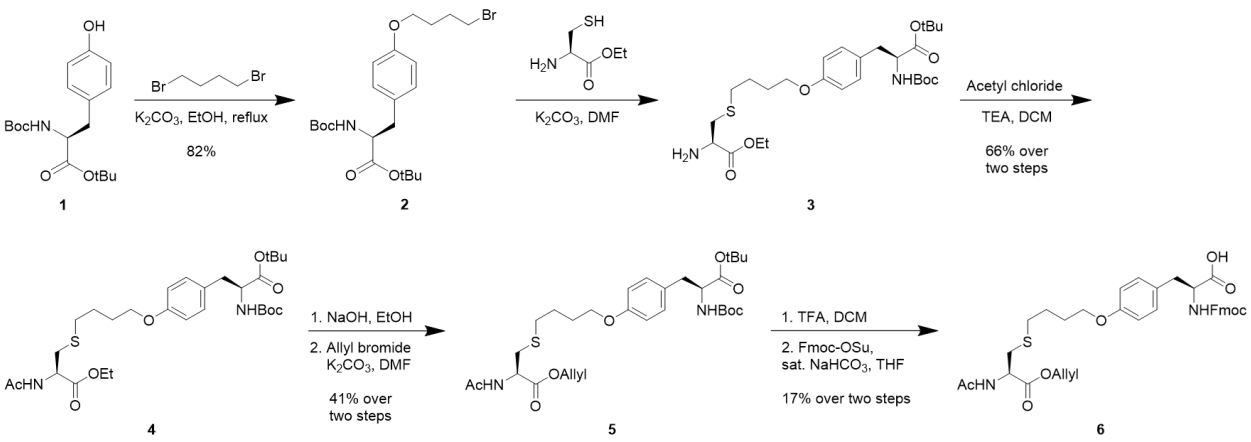
C18 RP-HPLC method: Linear gradient of 5–95% Acetonitrile (0.1% TFA)/ddH_2O (0.1% TFA), *t* = 1–21 min. Purity: 80%



FSS-m8

C18 RP-HPLC method: Linear gradient of 5–95% Acetonitrile (0.1% formic acid)/ddH_2O (0.1% formic acid), *t* = 2–7 min. Purity: >95%





Scheme S1. Synthesis of dipeptide 6.

Supplemental Tables

Table S1. Oligonucleotide sequences for the preparation of MOrPH-PhD libraries.

Primer Number	Primer Name	Primer Sequence 5' → 3'
1	pSEX81-QC Reverse	caagcccaataggaacccatgtaccgtaacac
2	pSEX81-QC Forward	gaagcggaagagcgcccaatacg
3	Library1 i/i+10 Forward	ggttcttagywtmrwywtmrwywtmrwywtmrwywt tgcggttctgcggccgctggatccaaagatatc
4	Library1 i/i-10 Forward	ggttcttgcywtmrwywtmrwywtmrwywtmrwywt tagggttctgcggccgctggatccaaagatatc
5	Library2 i/i+10 Forward	ggttcttagywtmrwmrwywtywtywtmrwmrwywt tgcggttctgcggccgctggatccaaagatatc
6	Library2 i/i-10 Forward	ggttcttgcywtmrwmrwywtywtywtmrwmrwywt tagggttctgcggccgctggatccaaagatatc
7	Library3 i/i+10 Forward	ggttcttagmrwywtywtmrwywtmrwywtywtmrw tgcggttctgcggccgctggatccaaagatatc
8	Library3 i/i-10 Forward	ggttcttgcmrwywtywtmrwywtmrwywtywtmrw tagggttctgcggccgctggatccaaagatatc
9	Library4 i/i+8 Forward	ggttcttaghwsmrwhwsmrwhwsmrwhwstgc ggttctgcggccgctggatccaaagatatc
10	Library4 i/i-8 Forward	ggttcttgchwsmrwhwsmrwhwsmrwhwstag ggttctgcggccgctggatccaaagatatc
11	Library5 i/i+8 Forward	ggttcttaghwshwsmrwmrwmrwhwshwstgc ggttctgcggccgctggatccaaagatatc
12	Library5 i/i-8 Forward	ggttcttgchwshwsmrwmrwmrwhwshwstag ggttctgcggccgctggatccaaagatatc
13	Library1 i/i+10 Reverse	agaaccgcaawrwykawrwykawrwyka wrwykawrctaagaacccgccatggccggctgagctg
14	Library1 i/i-10 Reverse	agaaccctaawrwykawrwykawrwykawrwykawrg caagaacccgccatggccggctgagctg
15	Library2 i/i+10 Reverse	agaaccgcaawrwykwykawrawrawrwykwykawr ctaagaacccgccatggccggctgagctg
16	Library2 i/i-10 Reverse	agaaccctaawrwykwykawrawrawrwykwykawrg caagaacccgccatggccggctgagctg
17	Library3 i/i+10 Reverse	agaaccgcawykawrawrwykawrwykawrawrwyk ctaagaacccgccatggccggctgagctg
18	Library3 i/i-10 Reverse	agaaccctawykawrawrwykawrwykawrawrwyk gcaagaacccgccatggccggctgagctg
19	Library4 i/i+8 Reverse	agaaccgcaswdwykswdwykswdwykswdctaagaa ccc gccatggccggctgagctg

20	Library4 i/i-8 Reverse	agaaccctaswdwykswdwykswdwykswdgcaagaaccc gccatggccggctgagctg
21	Library5 i/i+8 Reverse	agaaccgcaswdswdwykwykwykswdswdctaagaaccc gccatggccggctgagctg
22	Library5 i/i-8 Reverse	agaaccctaswdswdwykwykwykswdswdgcaagaaccc gccatggccggctgagctg

Table S2. Oligonucleotide sequences for cloning macrocyclic precursor sequences.

Primer Number	Primer Name	Primer Sequence 5' → 3'
23	FLAG-CBD_R	ctcgagcagttcgagaccgttgttacc
24	pCaaF-i-8-m1_F	agtcaggatccggttcatgcaaacagctgcgtaaacagctgtaggaattgaaggccgcaaaatcg
25	T7-terminal-ext_R	gctagttattgctcagcggtggca
26	MG-FSS-m1-Y-GyrA_F	aagcacatatgggttgcaaacagctgcgtaaacagctgtagtattgcatcacgggagat
27	MG-FSS-m1_K1A_F	tatgggttgcgcgcagctgcgtaaacag
28	MG-FSS-m1_Q2A_F	tatgggttgcaaagcgctgcgtaaacag
29	MG-FSS-m1_L3A_F	tatgggttgcaaacaggcgcgtaaacag
30	MG-FSS-m1_R4A_F	gcaaacagctggcgaaacagctgtagtatt
31	MG-FSS-m1_K5A_F	ctgcgtgcccagctgtagtattgcatc
32	MG-FSS-m1_Q6A_F	gctgcgtaaagcgctgtagtattgcatc
33	MG-FSS-m1_L7A_F	gctgcgtaaacaggcgtagtattgcatc

Table S3. Amino acid sequences of the CBD and GyrA-fused macrocycles.

Macrocycle	Amino Acid Sequence
FLAG-FSS-m1-CBD-H6	MDYKDDDDKGSGSGSCKQLRKQL(pCaaF)GIEGRKIEEGKLTNPGVSAW QVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQNNGN NGLELLEHHHHHH
MG-FSS-m1-Y-GyrA-H6	MGCKQLRKQL(pCaaF)YCITGDALVALPEGESVRIADIVPGARPNSDNAI DLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLV DVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYT VGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVY SLRVDTADHAFITNGFVSHALEHHHHHH
MG-FSS-m1_K1A-Y-GyrA-H6	MGCAQLRKQL(pCaaF)YCITGDALVALPEGESVRIADIVPGARPNSDNAI DLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLV DVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYT VGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVY SLRVDTADHAFITNGFVSHALEHHHHHH
MG-FSS-m1_Q2A-Y-GyrA-H6	MGCKALRKQL(pCaaF)YCITGDALVALPEGESVRIADIVPGARPNSDNAI DLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLV DVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYT VGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVY SLRVDTADHAFITNGFVSHALEHHHHHH
MG-FSS-m1_L3A-Y-GyrA-H6	MGCKQARKQL(pCaaF)YCITGDALVALPEGESVRIADIVPGARPNSDNAI DLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLV DVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYT VGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVY SLRVDTADHAFITNGFVSHALEHHHHHH
MG-FSS-m1_R4A-Y-GyrA-H6	MGCKQLAKQL(pCaaF)YCITGDALVALPEGESVRIADIVPGARPNSDNAI DLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLV DVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYT VGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVY SLRVDTADHAFITNGFVSHALEHHHHHH
MG-FSS-m1_K5A-Y-GyrA-H6	MGCKQLRAQL(pCaaF)YCITGDALVALPEGESVRIADIVPGARPNSDNAI DLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLV DVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYT VGVPGLVRFLEAHHRDPAQAIADELTDGRFYYAKVASVTDAGVQPVYS LRVDTADHAFITNGFVSHALEHHHHHH
MG-FSS-m1_Q6A-Y-GyrA-H6	MGCKQLRKAL(pCaaF)YCITGDALVALPEGESVRIADIVPGARPNSDNAI DLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLV DVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYT VGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVY SLRVDTADHAFITNGFVSHALEHHHHHH
MG-FSS-m1_L7A-Y-GyrA-H6	MGCKQLRKQA(pCaaF)YCITGDALVALPEGESVRIADIVPGARPNSDNAI DLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLV DVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYT

	VGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVY SLRVDTADHAFITNGFVSHALEHHHHHH

Table S4. Mass spectrometry analysis of recombinant macrocycles. The Cys and eUAA residues involved the thioether linkage are highlighted in bold and underlined.

Macrocycle	Amino Acid Sequence	Calc. $[M\text{-}Met+H]^+$ (Da)	Obs. $[M\text{-}Met+H]^+$ (Da)
MG-FSS-m1-Y	MG**C**KQLRKQL**(pCaaF)**Y	1439.10	1439.1
MG-FSS-m1_K1A-Y	MG**C**AQLRKQL**(pCaaF)**Y	1382.00	1382.8
MG-FSS-m1_Q2A-Y	MG**C**KALRKQL**(pCaaF)**Y	1382.04	1384.3
MG-FSS-m1_L3A-Y	MG**C**KQARKQL**(pCaaF)**Y	1397.02	1397.0
MG-FSS-m1_R4A-Y	MG**C**KQLAKQL**(pCaaF)**Y	1353.99	1356.8
MG-FSS-m1_K5A-Y	MG**C**KQLRAQL**(pCaaF)**Y	1382.00	1388.3
MG-FSS-m1_Q6A-Y	MG**C**KQLRKAL**(pCaaF)**Y	1382.04	1382.2
MG-FSS-m1_L7A-Y	MG**C**KQLRKQA**(pCaaF)**Y	1397.02	1397.7

Experimental Methods

Molecular Cloning and Plasmid Construction

The preparation of pEVOL-based plasmids for the expression of orthogonal AARS/tRNA^{CUA} pairs for amber stop codon suppression with O2beY, O4bbY, pAaF or pCaaF was reported previously.[1, 2] MOrPH-PhD libraries were inserted between the pelB leader sequence and pIII phage coat protein gene in a pSEX81 phagemid via a combination of splicing by overlap extension (SOE) and QuikChange PCR (Agilent Tech. Inc) using primers **1**-**22** in Table S1. In a first PCR reaction, inserts were amplified containing the MOrPH-PhD library either as a 5´ overhang (forward primers **3**-**12** with reverse primer **1**) or a 3´ overhang (reverse primers **13**-**22** with forward primer **2**) generating a ~1000 bp PCR product. Next, a second SOE PCR reaction was performed using primers from the first PCR reaction utilizing their respective overhangs for extension (e.g., PCR product from primers **1** and **3** with PCR product from primers **2** and **13**). Primers **1** and **2** were utilized to generate an ~2000 bp PCR product with the MOrPH-PhD library at its center. In a third PCR reaction, a QuikChange protocol was used to amplify the entire pSEX81 phagemid using the PCR product from the second reaction as the primer. The resulting PCR product was incubated with *Dpn*I (10 units) at 37 °C for 2 h to remove the template phagemid. The amplified phagemid was transformed into electrocompetent TOP10F' *E. coli* cells and selected on 20 cm x 20 cm 2× YT agar plates containing ampicillin (100 mg/L) and tetracycline (5 mg/L). A colony forming unit (c.f.u.) count exceeding by at least 3-fold the size of the respective DNA library was utilized for all the libraries. Colonies were then collected from the plates and the plasmid library was isolated using a plasmid midi-prep kit (Qiagen). The purified DNA libraries were confirmed via Sanger sequencing and pooled prior to phage generation.

Phage Expression and Purification

The pSEX81-based library was transformed into TOP10F' *E. coli* cells containing the pEVOL-based plasmid by electroporation. The cells were recovered with SOC media (2% w/v tryptone, 0.5% w/v yeast extract, 10 mM NaCl, 2.5 mM KCl, and 20 mM glucose), and incubated with shaking at 37 °C for 1 h. Cells were then transferred to a 125 mL Erlenmeyer flask containing 20 mL 2× YT media (1.6% w/v tryptone, 1.0% w/v yeast extract, 8.6 mM NaCl) supplemented with ampicillin (100 mg/L), chloramphenicol (34 mg/L), and tetracycline (5 mg/L). Cell cultures were grown overnight (12-16 h) at 37 °C, and then cells were recovered by centrifugation (4,000 × g). The cell pellet was diluted to an OD_{600} of 0.05 in fresh 2× YT media supplemented with ampicillin (100 mg/L), chloramphenicol (34 mg/L), tetracycline (5 mg/L) and allowed to reach an OD_{600} of 0.6. A volume equal to 10% of the final phage expression culture volume was infected with Hyperphage (Progen) at an MOI of 20. Hyperphage was allowed to infect the cells for 1 h at 37 °C with shaking (200 rpm) and then pelleted by centrifugation (4,000 × g). The pellet was resuspended in 20 mL of 2× YT supplemented with ampicillin (100 mg/L), chloramphenicol (34 mg/L), tetracycline (5 mg/L), kanamycin (30 mg/L), arabinose (0.06%), and the appropriate non-canonical amino acid (2 mM). Cultures were grown for 18 h at 30 °C with shaking (200 rpm) to express the desired eUAA-containing library. After expression, cell cultures were pelleted by centrifugation (4,000 × g). The resulting supernatant was isolated and incubated at pH 8.5 for 6 h to facilitate complete cyclization of macrocyclic peptides and then concentrated using an Amicon 30 kDa spin filter to a convenient volume (250–300 μL). The concentrated supernatant was then mixed with 1:4 (v/v) 20% polyethylene glycol buffer (20% polyethylene glycol, 2.5 M NaCl) at 4 °C and incubated overnight. The precipitated phage was pelleted by centrifugation (14,000 × g) for 30 min and resuspended in 200 μL PBS (10 mM Na_2HPO_4, 1.8 mM, KH_2PO_4, 137 mM NaCl,

2.7 mM KCl, pH 7.5). The fully resuspended phage solution was centrifuged (14,000 × *g*) for an additional 5 min to remove any insoluble cellular debris. The clarified phage solution was purified a second time and then passed through a 0.22 μm filter and stored in HEPES buffer (50 mM HEPES pH 7.0, 100 mM NaCl, 12 mM $MgCl_2$,) at 4 °C.

Determination of Phage Titer

10 μL aliquots of purified phage solutions were diluted serially in 10-fold dilutions with 2×YT media. 10 μL of each dilution is added to 90 μL of exponentially growing *E. coli* TOP10F' cells (OD_{600} = 0.4–0.6) in 1.7 mL microfuge tubes. The phage was allowed to infect *E. coli* cells for 1 h at 37 °C with shaking on a desktop thermoblock. 100 μL of phage infected *E. coli* was then spread on 2×YT agar plates containing ampicillin (100 mg/L) and tetracycline (5 mg/L) and grown overnight at 37 °C. The phage titer was determined by counting colony forming units.

RNA Folding

The biotinylated SARS-CoV-2 RNA was generated by chemical synthesis (Horizon Discovery) and deprotected and desalted prior to use. Approximately 100 μg of RNA was dissolved in 100 μL 0.05 M Na-HEPES, pH 7.0, and heated at 65 °C. After 3 min, the RNA was mixed with 100 μL of folding buffer warmed to 65 °C (0.05 M Na-HEPES, pH 7.0, 0.20 M NaCl, and 0.024 M $MgCl_2$) and incubated at 65 °C for 3 min. The RNA was slow cooled overnight to room temperature.

Selection of MOrPH-PhD Libraries

Streptavidin-coated magnetic beads (NEB) in 20 µl aliquots were added to two separate Eppendorf tubes and washed with HEPES buffer (3×) to remove residual storage buffer. One aliquot was incubated with 100 mL of a 1 mM solution of biotinylated DL (or FSS) in HEPES buffer for 2 h at room temperature for RNA immobilization. The beads were removed from solution via magnetic separation and rinsed with wash buffer (50 mM HEPES pH 7.0, 100 mM NaCl, 12 mM $MgCl_2$, 0.05% Tween-20) to remove residual RNA. Both aliquots of beads were then incubated with 100 mL of blocking buffer (50 mM HEPES pH 7.0, 100 mM NaCl, 12 mM $MgCl_2$, 0.5% BSA) for 2 h at room temperature, followed by rinsing with wash buffer (3 x 100 mL). Next, the phage solution was added to the aliquot lacking the immobilized RNA and incubated for 1 h at room temperature to remove library members which bind the matrix. The beads were then removed via magnetic separation to isolate the remaining phage library. The recovered phage solution was treated with SUPERaseIn RNase inhibitor (5 U/mL; Thermo Fisher) and incubated at room temperature for 15 min. The phage solution was then incubated with the immobilized RNA for 1 h at room temperature with gentle mixing. The beads were then removed from solution via magnetic separation, washed 3–5 times with wash buffer, and incubated with 100 µL of elution buffer (0.2 M glycine-HCl, pH 2.2, 1 mg/ml BSA) for 30 min at room temperature with gentle shaking. The elution solution was then neutralized with 20 µL of neutralization buffer (1 M Tris-HCl pH 9.1). 10 µl of the eluted phage solution was used to determine the titer of recovered phage. The remaining eluted phage was used to infect 2.5 mL mid-log TOP10F' cells (OD_{600} 0.4–0.6) in 2× YT for 1 h at 37 °C. This culture was then pelleted by centrifugation (4,000 × g) and resuspended in 5 mL of fresh 2× YT media supplemented with ampicillin (100 mg/L) and tetracycline (5 mg/L) and allowed to grow to saturation overnight at 37

°C. The plasmid was extracted from the overnight culture and the enriched plasmid pool was used to propagate new phage as described above. After three rounds of affinity selection and amplification, the enriched library was analyzed via next generation sequencing.

Next Generation Sequencing

Library amplicons for sequencing via an Illumina Miseq NGS platform were generated following Illumina's 16s amplicon generation protocol. In a first PCR reaction, the enriched phage libraries are amplified using primers containing Nextera-style tag sequences and locus-specific sequences for annealing to the pSEX81 phagemid. The resulting PCR product was purified via electrophoresis on a 1.2% agarose gel and the DNA band corresponding to the correct size (~200 bp) was isolated using a QIAquick gel extraction kit (Qiagen) following the manufacturer's instructions and eluted in 30 μL of ddH_2O. In a second PCR reaction, the final library amplicon was generated using the product from the first PCR as the template and Nextera-style indexing primers for the introduction of index sequences to facilitate the multiplexed analysis and deconvolution of multiples libraries in a single sequencing run. The resulting PCR product was purified in a similar manner as previously described, and the library amplicon was submitted for sequencing. Data from the sequenced libraries were aligned and organized with respect to their relative sequence count for identification of enriched members and consensus analysis.

Reagents and Analysis

Chemical reagents and solvents were purchased from AAPPTec, Acros Organics, Chem-Impex and Sigma-Aldrich. Silica gel chromatography purifications were carried out by using AMD Silica Gel 60 230-400 mesh. NMR spectra were recorded on a Brüker Avance 500 MHz

spectrometer by using solvent peaks as the reference. LC/MS analyses were performed on an Advion expression® compact mass spectrometer. MALDI-TOF spectra were acquired on a Shimadzu AXIMA Confidence MALDI-TOF spectrometer using α-cyano-4-hydroxycinnamic acid (CHCA) as the matrix. Concentrations of peptides were determined with a Thermo Scientific™ NanoDrop™ One Microvolume UV-Vis spectrophotometer.

Synthesis of Electrophilic Unnatural Amino Acids

The eUAAs O2beY, O4bbY, pAaF and pCaaF used for incorporation into MOrPH-PhD libraries were prepared as previously described.[1, 2]

Synthesis of (*S*)-2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(((allyloxy)carbonyl)-amino)phenyl)propanoic acid (N-Fmoc-N´-Alloc-*p*-amino-L-Phe-OH)

The SPPS building block for the preparation of Cys/pCaaF peptides was prepared as previously described.[3] R_f = 0.5 (1:9:0.1 methanol:DCM:acetic acid). ^{1}H NMR (400 MHz, MeOD) δ 9.20 (s, 1H), 7.79 (d, *J* = 7.8 Hz, 2H), 7.60 (d, *J* = 7.8 Hz, 2H), 7.38 (m, 4H), 7.31 (d, *J* = 7.2 Hz, 2H), 7.17 (d, *J* = 7.1 Hz, 2H), 6.00 (ddt, *J* = 17.9, 12.5, 6.3 Hz, 1H), 5.37 (d, *J* = 17.4 Hz, 1H), 5.23 (d, *J* = 10.5 Hz, 1H), 4.63 (d, *J* = 6.2 Hz, 2H), 4.41 (m, *J* = 5.4 Hz, 1H), 4.37 – 4.30 (m, 1H), 4.28 – 4.21 (m, 1H), 4.17 (t, *J* = 7.0 Hz, 1H), 3.17 (dd, *J* = 13.8, 5.5 Hz, 1H), 2.91 (dd, *J* = 14.5, 9.0 Hz, 1H); ^{13}C NMR (101 MHz, MeOD) δ 175.13, 158.36, 155.79, 145.18, 142.54, 134.28, 133.18, 130.70, 128.74, 128.16, 126.31, 126.20, 120.86, 119.91, 117.78, 67.94, 66.38, 56.80, 37.91. MS (ESI): calculated for $C_{28}H_{26}N_2O_6$: 487.19 $[M+H]^+$; found: 487.2.

Synthesis of (*S*)-2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(4-bromobutoxy)-phenyl)propanoic acid (Fmoc-O4bbY-OH)

To an ice-cold reaction mixture of O4bbY (125 mg, 0.395 mmol, 1 equivalent) dissolved in saturated sodium bicarbonate (10 mL) at pH 9, a solution of *N*-(9-fluorenylmethoxy-carbonyloxy)succinimide (120 mg, 0.356 mmol, 0.9 equivalent) in tetrahydrofuran (4 mL) was added dropwise. After the reaction was stirred overnight at room temperature, the reaction mixture was quenched with 5% HCl. The reaction mixture then was partitioned between ethyl acetate (30 mL) and water (30 mL) three times, and the combined organic layers were washed with brine (30 mL) and dried over anhydrous sodium sulfate. After solvent was removed by a rotary evaporator, the crude residue was washed with hexane (10 mL) three times. Subsequently, the remaining insoluble residue was taken up in 100 mM HCl (10 mL), flash frozen and lyophilized to afford the desired product as a white solid (108 mg, 51%). R_f = 0.4 (1:9:0.1 methanol:DCM:acetic acid). ^{1}H NMR (400 MHz, MeOD) δ 7.81 (d, *J* = 7.9 Hz, 2H), 7.60 (t, *J* = 6.2 Hz, 2H), 7.40 (t, *J* = 7.2 Hz, 2H), 7.32 (t, *J* = 6.7 Hz, 2H), 7.16 (d, *J* = 6.8 Hz, 2H), 6.81 (t, *J* = 7.6 Hz, 2H), 4.45 – 4.30 (m, 2H), 4.25 – 4.08 (m, 2H), 3.96 – 3.83 (m, 2H), 3.47 (t, *J* = 6.4 Hz, 2H), 3.18 (dd, *J* = 13.4, 5.0 Hz, 1H), 2.88 (dd, *J* = 13.9, 8.0 Hz, 1H), 1.99 (tt, *J* = 7.0 Hz, 2H), 1.86 (tt, *J* = 7.0 Hz, 2H); ^{13}C NMR (151 MHz, MeOD) δ 175.64, 159.21, 158.28, 145.32, 145.20, 142.55, 131.39, 128.75, 128.16, 126.44, 126.26, 120.89, 115.46, 68.01, 57.72, 57.19, 38.05, 34.07, 30.76, 29.00, 28.23. MS (ESI): calculated for $C_{28}H_{28}BrNO_5$: 538.12 $[M+H]^+$; found: 538.1.

Synthesis of (*S*)-2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(4-(((*R*)-2-acetamido-3-(allyloxy)-3-oxopropyl)thio)butoxy)phenyl)propanoic acid (Cys/O4bbY Dipeptide)

t-Butyl (S)-3-(4-(4-bromobutoxy)phenyl)-2-((tert-butoxycarbonyl)amino)propanoic acid (**2**). To a solution of commercially available *N*-Boc-L-tyrosine *t*-butyl ester (**1**) (2.00 g, 5.93 mmol) and anhydrous potassium carbonate (2.46 g, 17.8 mmol, 3 equivalents) in ethanol (200 proof, 20 mL) was added 1,4-dibromobutane (7.08 mL, 59.3 mmol, 10 equivalents) dropwise and stirred at reflux overnight in accordance with previously reported procedures.[2] R_f = 0.7 (1:3 ethyl acetate:hexane). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.07 (d, *J* = 8.7 Hz, 2H), 6.80 (d, *J* = 8.7 Hz, 2H), 4.96 (d, *J* = 8.2 Hz, 1H), 4.40 (d, *J* = 8.2 Hz, 1H), 3.97 (t, *J* = 6.1 Hz, 2H), 3.49 (t, *J* = 6.6 Hz, 2H), 2.98 (m, 2H), 2.11 – 2.03 (m, 2H), 1.97 – 1.89 (m, 2H), 1.42 (d, 18H). MS (ESI): calculated for $C_{22}H_{34}BrNO_5$: 494.15 $[M+Na]^+$; found: 494.2.

t-Butyl (S)-3-(4-(4-(((R)-2-amino-3-ethoxy-3-oxopropyl)thio)butoxy)phenyl)-2-((tert-butoxycarbonyl)amino)propanoic acid (**3**). A solution of **2** (1.50 g, 3.18 mmol) in dry *N*,*N*-dimethylformamide (DMF) (3 mL) was added dropwise to a solution of L-cysteine ethyl ester hydrochloride salt (884 mg, 4.76 mmol, 1.5 equivalents) with anhydrous potassium carbonate (1.76 g, 12.7 mmol, 4 equivalents) in dry DMF (5 mL) at room temperature. The mixture was stirred overnight at room temperature under argon. Following DMF removal by rotary evaporation, the mixture was diluted with ethyl acetate (30 mL) and washed with water (50 mL) twice followed by brine (30 mL). The organic layer was dried over anhydrous sodium sulfate. After solvent was removed by a rotary evaporator, crude residue of compound **3** was produced as a yellowish brown oil and used directly for the subsequent reaction. R_f = 0.3 (ethyl acetate). MS (ESI): calculated for $C_{27}H_{44}N_2O_7S$: 541.29 $[M+H]^+$; found: 540.9.

t-Butyl (S)-3-(4-(4-(((R)-2-acetamido-3-ethoxy-3-oxopropyl)thio)butoxy)phenyl)-2-((tert-butoxycarbonyl)amino)propanoic acid (**4**). Triethylamine (1330 μL, 9.53 mmol, 3 equivalents) was added to a solution of crude **3** in dichloromethane (DCM) (20 mL). After the mixture was

cooled to 0 °C, acetyl chloride (453 μL, 6.35 mmol, 2 equivalents) was added dropwise, and the reaction was stirred for three hours at room temperature under argon. Upon quenching with water, the reaction mixture was diluted with DCM (20 mL), and the reaction mixture was washed with 5% HCl (20 mL) three times, saturated sodium bicarbonate (20 mL) three times, and once with brine (20 mL). The organic layer was dried over anhydrous sodium sulfate. After solvent was removed by a rotary evaporator, the crude residue was purified by silica gel chromatography using a solvent system of ethyl acetate/hexane (3:5) to produce compound **4** as a yellowish brown oil (1.22 g, 66% over two steps). R_f = 0.5 (3:1 ethyl acetate:hexane). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.07 (d, *J* = 8.5 Hz, 2H), 6.80 (d, *J* = 8.7 Hz, 2H), 6.32 (d, *J* = 7.6 Hz, 1H), 4.97 (d, *J* = 8.7 Hz, 1H), 4.81 (m, 1H), 4.40 (m, 1H), 4.23 (m, 2H), 3.93 (t, *J* = 6.3 Hz, 2H), 3.07 – 2.91 (m, 4H), 2.59 (td, *J* = 7.2, 1.6 Hz, 2H), 2.05 (s, 3H), 1.88 – 1.82 (m, 2H), 1.78 – 1.73 (m, 2H), 1.42 (d, 18H), 1.30 (t, *J* = 7.2 Hz, 3H). MS (ESI): calculated for $C_{29}H_{46}N_2O_8S$: 605.29 $[M+Na]^+$; found: 605.0.

Allyl N-acetyl-S-(4-(4-((S)-3-(tert-butoxy)-2-((tert-butoxycarbonyl)amino)-3-oxopropyl)-phenoxy)butyl)-L-cysteine (**5**). A 2 N solution of sodium hydroxide (0.8 g) in water (10 mL) was added to a solution of **4** (1.22 g, 2.10 mmol) in ethanol (10 mL). After the reaction was stirred for 2 h at room temperature, the reaction mixture was quenched with 5% HCl. The mixture was partitioned between ethyl acetate (30 mL) and water (30 mL) three times, and the combined organic layers were washed with brine (30 mL) and dried over anhydrous sodium sulfate. Upon solvent removal via rotary evaporation, crude residue was used directly for the subsequent reaction. MS (ESI): calculated for $C_{27}H_{42}N_2O_8S$: 577.26 $[M+Na]^+$; found: 576.9. To a reaction mixture of the crude residue with anhydrous potassium carbonate (1.16 g, 8.38 mmol, 4 equivalents) in dry DMF (5 mL), allyl bromide (272 μL, 3.14 mmol, 1.5 equivalents) was added dropwise, and the reaction was stirred overnight at room temperature under argon. Upon DMF

removal via rotary evaporation, the reaction mixture was diluted with ethyl acetate (30 mL) and washed with water (30 mL) twice followed by brine (30 mL). The organic layer was dried over anhydrous sodium sulfate. After solvent was removed by rotary evaporation, the crude residue was purified by silica gel chromatography using a solvent system of ethyl acetate/hexane (2:1) to produce compound **5** as an oil (512 mg, 41% over two steps). R_f = 0.63 (3:1 ethyl acetate:hexane). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.07 (d, *J* = 8.7 Hz, 2H), 6.80 (d, *J* = 8.7 Hz, 2H), 6.41 (d, *J* = 7.6 Hz, 1H), 5.96 – 5.86 (m, 1H), 5.35 (dq, *J* = 17.1, 1.4 Hz, 1H), 5.27 (dq, *J* = 10.5, 1.2 Hz, 1H), 4.99 (d, *J* = 8.4 Hz, 1H), 4.85 (dt, *J* = 7.8, 5.2 Hz, 1H), 4.66 (dt, *J* = 5.8, 1.4 Hz, 2H), 4.40 (q, *J* = 6.1 Hz, 1H), 3.93 (t, *J* = 6.2 Hz, 2H), 3.08 – 2.95 (m, 4H), 2.59 (td, *J* = 7.2, 1.4 Hz, 2H), 2.05 (s, 3H), 1.89 – 1.81 (m, 2H), 1.75 (m, 2H), 1.42 (d, 18H). MS (ESI): calculated for $C_{30}H_{46}N_2O_8S$: 617.29 $[M+Na]^+$; found: 616.9.

(S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(4-(((R)-2-acetamido-3-(allyloxy)-3-oxopropyl)thio)butoxy)phenyl)propanoic acid (**6**). Compound **5** (512 mg, 0.861 mmol) was dissolved in a 1:1 mixture of trifluoroacetic acid (TFA) and DCM (10 mL total). After the reaction was stirred for 5 h, volatiles were removed via rotary evaporation. The sample was then suspended in diethyl ether (20 mL), and volatiles were removed via rotary evaporation. This process was performed five times to remove residual TFA, yielding crude residue that was used directly for the subsequent reaction. MS (ESI): calculated for $C_{21}H_{30}N_2O_6S$: 439.19 $[M+H]^+$; found: 438.8. To an ice-cold reaction mixture of the crude residue dissolved in saturated sodium bicarbonate (10 mL) at pH 9, a solution of *N*-(9-fluorenylmethoxycarbonyloxy)succinimide (446 mg, 1.32 mmol, 1.53 equivalents) in tetrahydrofuran (2 mL) was added dropwise. After the reaction was stirred overnight at room temperature, the reaction mixture was quenched with 5% HCl. The reaction mixture then was partitioned between ethyl acetate (30 mL) and water (30 mL)

three times, and the combined organic layers were washed with brine (30 mL) and dried over anhydrous sodium sulfate. After solvent was removed by a rotary evaporator, the crude residue was purified over silica gel chromatography using a solvent system of methanol/DCM/acetic acid (1:9:0.1) to produce compound **6** as a clear oil (95 mg, 17% over two steps). R_f = 0.19 (1:9:0.1 methanol:DCM:acetic acid). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.6 Hz, 2H), 7.56 (t, *J* = 7.4 Hz, 2H), 7.39 (t, *J* = 7.5 Hz, 2H), 7.30 (m, 2H), 7.04 (d, *J* = 8.1 Hz, 2H), 6.78 (d, *J* = 8.3 Hz, 2H), 6.51 (t, *J* = 6.6 Hz, 1H), 5.90 (m, 1H), 5.39 – 5.31 (m, 2H), 5.26 (dd, *J* = 10.4, 1.3 Hz, 1H), 4.83 (m, 1H), 4.65 (m, 3H), 4.44 (dd, *J* = 10.5, 7.1 Hz, 1H), 4.34 (dd, *J* = 10.7, 6.9 Hz, 1H), 4.20 (t, *J* = 7.2 Hz, 1H), 3.91 (t, *J* = 6.1 Hz, 2H), 3.10 (qd, *J* = 14.0, 5.6 Hz, 2H), 2.99 (qd, *J* = 13.8, 5.1 Hz, 2H), 2.56 (m, 2H), 2.04 (m, 3H), 1.81 (m, 2H), 1.72 (m, 2H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 174.61, 170.92, 170.68, 158.13, 155.89, 143.95, 143.86, 141.45, 131.36, 130.60, 127.87, 127.20, 125.20, 120.12, 119.42, 114.82, 67.30, 67.15, 66.61, 54.86, 52.25, 47.30, 37.16, 34.16, 32.54, 28.25, 26.19, 23.08. MS (ESI): calculated for $C_{36}H_{40}N_2O_8S$: 661.26 $[M+H]^+$; found: 660.7 (Figure S6; Scheme S1).

Synthesis of pCaaF-*i*/*i*-(8/10) Macrocyclic Peptides via On-Resin Cyclization

Synthesis was carried out by conventional Fmoc-solid-phase peptide synthesis on Knorr amide resin (0.4 mmol g^{-1} on a 0.1-mmol scale). Peptides were generated manually in syringes equipped with a Teflon frit. Initially, the resin was subjected to swelling for 30 min in 5 ml of a 1:1 mixture of dichloromethane/*N*,*N*-dimethylformamide (DCM/DMF). Swollen resin was then treated with 20% piperidine in DMF containing 0.05 M 1-hydroxybenzotriazole (3 cycles of 3 min, 5 min, and 3 min) to achieve deprotection of the Fmoc group. After washing with DMF (5×), the resin was loaded with four equiv. of Fmoc-Gly-OH, four equiv. of 2-(6-chloro-1H-benzotriazole-

1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU), and eight equiv. of *N*,*N*-diisopropylethylamine (DIPEA) in DMF with respect to the initial loading of the resin. The coupling reaction was allowed to proceed at room temperature for 40 min, followed by washing with DMF (5x). Fmoc deprotection was performed as previously described and, after washing with DMF (5×), the resin was loaded with 0.5 equiv. of N-Fmoc-N´-Alloc-*p*-amino-L-Phe-OH, 0.5 equiv. of HCTU, and 1.0 equiv. of DIPEA in DMF and the coupling reaction was allowed to proceed for 40 min, followed by washing with DMF (5×). The remaining deprotected N-terminal amino groups were acetylated using 5 equiv. of acetic anhydride and 10 equiv. of DIPEA for 30 min followed by washing with DMF (5×). This procedure reduces the overall loading capacity of the resin by half but facilitates on-resin peptide cyclization at the final step of synthesis. Next, Fmoc groups were deprotected and the resin was washed with DMF (5×). The resin was loaded with 2 equiv. of amino acid, 2 equiv. of HCTU, and 4 equiv. of DIPEA in DMF and the coupling reaction was allowed to proceed for 40 min. The resin was then washed with DMF (5×) and this process was repeated to attain the desired peptide sequence. Upon coupling of Fmoc-L-Cys(Acm)-OH as the final amino acid residue in the peptide chain, the resin was treated with 0.25 equiv. of $Pd(PPh_3)_4$ and 20 equiv. of $PhSiH_3$ in DCM for 1 h at room temperature with shaking to promote deprotection of the Alloc protecting group of *p*-amino-L-phenylalanine. The resin was washed with DCM (5×) and treated with 2 equiv. of chloroacetyl chloride in DCM for 1 h to acetylate *p*-amino-L-phenylalanine and washed with DCM (5×) and DMF (2×). Next, the resin was treated with 10 equiv. of $PdCl_2$ in DMF/H_2O (9:1) and incubated overnight at room temperature with shaking to promote Acm deprotection of cysteine. The resin was then washed with DMF (5×), H_2O (2×), DMF (2×), DCM (2×), DMF (2×), and treated with 0.5 M dithiothreitol (DTT) in DMF (2× 1 min) to facilitate removal of the palladium catalyst. The resin was taken up in 5 mL of a 10%

solution of DIPEA in DMF and incubated at room temperature overnight with shaking to promote cyclization between the electrophilic unnatural amino acid (pCaaF) and cysteine. Upon final Fmoc deprotection, N-terminal amino groups were acetylated using 5 equiv. of acetic anhydride and 10 equiv. of DIPEA for 30 min followed by washing with DMF (5×). The resulting macrocyclic peptide was cleaved from the resin as follows: the resin was washed with DMF (3×), DCM (3×) and dried *in vacuo*. The dried resin was taken up in 10 mL of cleavage solution comprising 95% (v/v) trifluoroacetic acid (TFA), 2.5% (v/v) triisopropylsilane (TIPS) and 2.5% (v/v) H_2O and allowed to incubate for 2 h at room temperature. The filtrate was collected and concentrated under reduced pressure and mixed with a tenfold excess volume of cold diethyl ether followed by centrifugation (4,000 ×g). Supernatant was removed, and the crude peptide precipitate was dissolved in 10 mL of a 1:1 mixture of ACN/H_2O and lyophilized. The cyclic peptide was then purified via semipreparative HPLC using a linear gradient of 5% to 60% acetonitrile in water (with 0.1% TFA) over 30 min. The correct peptide mass was confirmed via matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (MS analysis using α-cyano-4-hydroxycinnamic acid (CHCA) as the matrix (Figure S3). Subsequently, pure fractions were combined and lyophilized. Finally, for removal of TFA, peptides were resuspended in 100 mM HCl and lyophilized to yield the peptide hydrochloride salts.

Synthesis of O4bbY-*i*/*i*+(8/10) Macrocyclic Peptides

Synthesis was carried out by conventional Fmoc-solid-phase peptide synthesis on Knorr amide resin (0.4 mmol g^{-1} on a 0.1-mmol scale). Peptides were generated manually in syringes equipped with a Teflon frit. Initially, the resin was subjected to swelling for 30 min in 5 ml of a 1:1 mixture of dichloromethane/*N*,*N*-dimethylformamide (DCM/DMF). Swollen resin was then

treated with 20% piperidine in DMF containing 0.05 M 1-hydroxybenzotriazole (3 cycles of 3 min, 5 min, and 3 min) to achieve deprotection of the Fmoc group. After washing with DMF (5×), the resin was loaded with four equivalents (equiv.) of Fmoc-Gly-OH, four equiv. of 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU) and eight equiv. of *N*,*N*-diisopropylethylamine (DIPEA) in DMF with respect to the initial loading of the resin. The coupling reaction was allowed to proceed at room temperature for 40 min, followed by washing with DMF (5×). The Fmoc deprotection was performed as previously described and, after washing with DMF (5×), the resin was loaded with 0.5 equiv. of Fmoc-L-Cys(Acm)-OH, 0.5 equiv. of HCTU, and 1 equiv. of DIPEA in DMF and the coupling reaction was allowed to proceed for 40 min, followed by washing with DMF (5×). The remaining deprotected N-terminal amino groups were acetylated using 5 equiv. of acetic anhydride and 10 equiv. of DIPEA for 30 min followed by washing with DMF (5×). This procedure reduces the overall loading capacity of the resin by half and facilitates on-resin peptide cyclization at the final step of synthesis. Next, Fmoc groups were deprotected and the resin was washed with DMF (5×). The resin was loaded with 2 equiv. of amino acid, 2 equiv. of HCTU, and 4 equiv. of DIPEA in DMF and the coupling reaction was allowed to proceed for 40 min. The resin was then washed with DMF (5×) and this process was repeated to attain the desired peptide sequence. Upon coupling of Fmoc-O4bbY-OH as the final amino acid residue in the peptide chain, the resin was treated with 10 equiv. of $PdCl_2$ in DMF/H_2O (9:1) and incubated overnight at room temperature with shaking to promote Acm deprotection of cysteine. The resin was then washed with DMF (5×), H_2O (2×), DMF (2×), DCM (2×), DMF (2×), and treated with a 0.5 M solution of DTT in DMF (2× 1 min) to facilitate removal of the palladium catalyst. Next, the resin was taken up in 5 mL of a 10% solution of DIPEA in DMF and incubated at room temperature overnight with shaking to promote cyclization between the electrophilic

unnatural amino acid (O4bbY) and cysteine. Upon final Fmoc deprotection, N-terminal amino groups were acetylated using 5 equiv. of acetic anhydride and 10 equiv. of DIPEA for 30 min followed by washing with DMF (5×). The resulting macrocyclic peptide was cleaved from the resin as follows: the resin was washed with DMF (3×), DCM (3×) and dried *in vacuo*. The dried resin was taken up in 10 mL of cleavage solution comprising 95% (v/v) trifluoroacetic acid (TFA), 2.5% (v/v) triisopropylsilane (TIPS) and 2.5% (v/v) H_2O and allowed to incubate for 2 h at room temperature. The filtrate was collected and concentrated under reduced pressure and mixed with a tenfold excess volume of cold diethyl ether followed by centrifugation (4,000 ×g). Supernatant was removed, and the crude peptide precipitate was dissolved in 10 mL of a 1:1 mixture of acetonitrile (ACN)/H_2O and lyophilized. The cyclic peptide was then purified via semipreparative HPLC using a linear gradient of 5% to 60% acetonitrile in water (with 0.1% TFA) over 30 min. The correct peptide mass was confirmed via MALDI-TOF-MS analysis using CHCA as the matrix (Figure S3). Subsequently, pure fractions were combined and lyophilized. Finally, for removal of TFA, peptides were resuspended in 100 mM HCl and lyophilized to yield the peptide hydrochloride salts.

Synthesis of O4bbY-*i*/*i*-(8/10) Macrocyclic Peptides

Synthesis was carried out by conventional Fmoc-solid-phase peptide synthesis on Knorr amide resin (0.4 mmol g^{-1} on a 0.05-mmol scale). Peptides were generated manually in syringes equipped with a Teflon frit. Initially, the resin was subjected to swelling for 30 min in 6 ml of a 1:1 mixture of dichloromethane/*N*,*N*-dimethylformamide (DCM/DMF). Swollen resin was then treated with 20% piperidine in DMF containing 0.05 M 1-hydroxybenzotriazole (3 cycles of 3 min, 5 min, and 3 min) to achieve deprotection of the Fmoc group. After washing with DMF (5×), the

resin was loaded with eight equiv. of Fmoc-Gly-OH, eight equiv. of HCTU and sixteen equiv. of N, N-diisopropylethylamine (DIPEA) in DMF with respect to the initial loading of the resin. The coupling reaction was allowed to proceed at room temperature for 40 min, followed by washing with DMF (5×). The Fmoc deprotection was performed as previously described and, after washing with DMF (5×), the resin was loaded with 4 equiv. of dipeptide **6** (Scheme S1), 4 equiv. of 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluoro-phosphate (HATU) and 8 equiv. of DIPEA in DMF and the coupling reaction was allowed to proceed for 1 h, followed by washing with DMF (5×). The resin was loaded with 20 eq. of acetic anhydride and 40 eq. of DIPEA in DMF to cap any unreacted N-terminal amine groups, and the acetylation reaction was allowed to proceed for 30 min. Upon washing with DMF (5x), Fmoc groups were deprotected, and the resin was washed with DMF (5×). The resin was loaded with 4 equiv. of amino acid, 4 equiv. of HCTU, and 8 equiv. of DIPEA in DMF and the coupling reaction was allowed to proceed for 40 min. The resin was then washed with DMF (5×) and this process was repeated to attain the desired peptide sequence. After coupling the final amino acid of the desired peptide sequence, the Fmoc group was deprotected with 20% piperidine in DMF containing 0.05 M 1-hydroxybenzotriazole as previously described. Next, the resin was treated with a mixture of 20 equiv. of phenyl silane ($PhSiH_3$) and 0.35 equiv. of tetrakis(triphenylphosphine)palladium ($Pd(PPh_3)_4$) dissolved in 2.5 mL of DCM and incubated for 1 h with gentile shaking to promote the deprotection of the allyl group. The resin was then washed with DCM (7×) and DMF (4×). Peptide cyclization via amide bond formation was carried out using a mixture of 2.5 equiv. of HATU and 5 equiv. of DIPEA dissolved in 2 mL DMF and shaken overnight at room temperature. The resulting macrocyclic peptide was cleaved from the resin as follows: the resin was washed with DMF (7×), DCM (3×) and dried *in vacuo*. The dried resin was

taken up in 10 mL of cleavage solution comprising 95% (v/v) TFA, 2.5% (v/v) TIPS and 2.5% (v/v) H_2O and allowed to incubate for 2 h at room temperature. The filtrate was collected and concentrated under reduced pressure and mixed with a tenfold excess volume of cold diethyl ether followed by centrifugation (4,000 × g). Supernatant was removed, and the crude peptide precipitate was dissolved in 5 mL of a 1:1 mixture of ACN/H_2O and lyophilized. The cyclic peptide was then purified via semipreparative HPLC using a linear gradient of 10% to 55% acetonitrile in water (with 0.1% TFA) over 30 min. The correct peptide mass was confirmed via MALDI-TOF-MS analysis using CHCA as the matrix (Figure S3). Subsequently, pure fractions were combined and lyophilized. Finally, for removal of TFA, peptides were resuspended in 100 mM HCl and lyophilized to yield the peptide hydrochloride salts.

Recombinant Expression and Purification of Macrocyclic Peptides

Macrocyclic peptides were produced recombinantly in accordance with previously reported procedures.[1, 2, 4] Initially, a fusion construct for **FSS-m1** was designed in the following order: N-terminal FLAG-tag, the macrocyclic precursor sequence, Factor Xa cleavage site, chitin-binding domain (CBD) and a C-terminal His-tag. Briefly, the gene encoding the fusion construct was amplified by PCR with forward primer **24** and reverse primer **23** using a previously described pET22 template vector containing an N-terminal FLAG-tag, Factor Xa cleavage site, CBD and a C-terminal His-tag[2]. Subsequently, the PCR product was inserted between the *Bam*HI and *Xho*I restriction sites of the aforementioned construct. Later, a fusion construct for **FSS-m1** was designed to consist of the following: N-terminal Met-Gly leader, the macrocyclic precursor sequence, GyrA intein and a C-terminal His-tag. The gene encoding the GyrA-fusion construct was amplified by PCR using forward primer **26** and reverse primer **25** using a different, previously

described GyrA-containing pET22 construct as the template[4]. Subsequently, the PCR product was inserted between the *Nde*I and *Xho*I restriction sites of the same GyrA-containing vector. Plasmids for the expression of the alanine scanning variants were prepared by substituting each residue of the macrocyclic precursor sequence individually with Ala via site-directed mutagenesis using the parent GyrA-containing **FSS-m1** fusion construct as the template, forward primers **27**–**33** and reverse primer **26**. All recombinant peptides were produced in *E. coli* BL21(DE3) cells via co-transformation of the appropriate expression plasmid and pEVOL-TyrRS42, an amber suppressor plasmid that has been demonstrated previously to incorporate pCaaF upon readthrough of the amber stop codon[2]. *E. coli* BL21(DE3) cells containing both an expression plasmid and a suppressor plasmid were grown overnight in 5 mL LB media supplemented with ampicillin (100 mg/L) and chloramphenicol (34 mg/L). Overnight cultures then were used to inoculate 1 L of fresh LB media supplemented with ampicillin and chloramphenicol at 37 °C. Upon reaching an OD_{600} of 0.6, the temperature was reduced to 27 °C, and cells were induced with arabinose (0.06% m/v) and supplemented with pCaaF (2 mM). After 1 h, isopropyl β-D-1-thiogalactopyranoside (IPTG) (1 mM) was added, and cells were grown overnight. The following day, cells were harvested by centrifugation (4,000 x g), lysed via sonication, and clarified by centrifugation (10,000 x g). Peptides were purified by Ni-NTA affinity chromatography and buffer exchanged with phosphate buffered saline (PBS) three times. Subsequently, for cleavage of the GyrA intein, peptides were incubated in a solution of PBS containing 20 mM tris(2-carboxyethyl)phosphine (TCEP) and 10 mM thiophenol at pH 8.5 overnight at room temperature. Then, the cleaved peptides (i.e., MG-**FSS-m1** and its alanine variants) were purified via solid phase extraction with a step gradient of acetonitrile in water (+0.1% TFA). Upon confirmation of peptide identities by MALDI-TOF-MS,

pure fractions were combined and lyophilized. Subsequently, for removal of TFA, peptides were resuspended in 100 mM HCl and lyophilized to yield the final peptide hydrochloride salts.

Surface Plasmon Resonance

The biotinylated SARS-CoV-2 RNAs (Figure 2) were generated by chemical synthesis (Horizon Discovery) and deprotected and desalted prior to use. The RNA was folded as described above then diluted in folding buffer to create a 1 μM stock. The RNA was immobilized on CM5 chips (Cytiva) conjugated to Neutravidin to achieve ~1000 RU using an SPR buffer comprising 0.010 M HEPES pH 7.0 and 0.15 M NaCl. Similarly, a 200 nM stock of biotinylated SARS-CoV-2 dimerization loop diluted in SPR buffer was flowed over the chip surface to achieve 1700 RU. Analysis of pCaaF(*i*-10)-m1 (**FSS-m2**) proceeded in SPR buffer. Analysis of the related peptide pCaaF(*i*-8)-m1 (**FSS-m1**) and all recombinant peptides proceeded in a 2× salt SPR buffer (0.010 M HEPES pH 7.0 and 0.30 M NaCl). O4bbY/Cys peptides (i.e., *i*/*i*+8 or *i*/*i*+10) were analyzed in an SPR buffer comprising 0.010 M HEPES pH 7.0, 0.15 M NaCl and 0.05% Tween 20, while Cys/O4bbY peptides (i.e., *i*/*i*-8 or *i*/*i*-10) were analyzed in an SPR buffer comprising 0.010 M HEPES pH 7.0 and 0.15 M NaCl without Tween 20. The flow rate for each experiment was 30 μL/min. Cyclic peptides at various concentrations (1.9–30.0 μM for pCaaF family peptides and 1.3–20.0 μM for O4bbY family peptides) were injected for 60 s and allowed to dissociate for 180 s. Similarly, recombinant peptides at a concentration of 2.5 μM were injected for 60 s and allowed to dissociate for 180 s. To regenerate the RNA, 2 M NaCl was injected for 60 s for pCaaF family peptides, and 3 M guanidine HCl was injected for 60 s for O4bbY family peptides. To reduce nonspecific binding during peptide assays, a 100-fold molar excess of yeast tRNA with respect to the immobilized RNA was added to each SPR buffer and the experiments were repeated for each

peptide. Experimental data were processed using the double-referencing method[5]. The buffer-subtracted sensorgrams were fit to a 1:1 binding model using Biacore T200 analysis software to determine rate constants (k_{on} and k_{off}) and the apparent equilibrium dissociation constant (K_D)[6]. The results were plotted using Prism software (GraphPad Inc.). The K_D for the equilibrium binding measurements was determined by taking the average response from a 5 s window at equilibrium (R_{eq}) for each peptide injection versus the peptide concentration using Prism software; data were then fit to a one-site binding model. The kinetic and equilibrium experiments were repeated in triplicate with the exception of O4bbY(*i*-10)-m1 (**FSS-m8**), which was repeated in duplicate.

Biotinylated RNAs chosen for off-target binding analysis were produced by chemical synthesis and deprotected by the manufacturer followed by desalting (Horizon Discovery). Each RNA corresponded to a well-characterized riboswitch. We tested the following sequences: (i) the 33-mer class I type II preQ$_1$ riboswitch from *Thermoanaerobacter tengcongensis* (*Tte*), which binds a single equivalent of preQ$_1$ with a K_D of 2.1 ± 0.3 nM;[6] (ii) the 34-mer class I type I preQ$_1$ riboswitch from *Carnobacterium antarticum* (*Can*), which cooperatively binds two preQ$_1$ equivalents in the same pocket with a K_{D1} of 891 nM and K_{D2} of 461 nM;[7] and a 64-mer guanine-I riboswitch variant from *Paenibacillus pectinilyticus* (*Ppe*), which binds a single equivalent of guanine with a K_D of approximately 351 ± 99 nM.[8] The corresponding sequences are: (i) *Tte* riboswitch CUGGGUCGCAGUAACCCCAGUUAACAAAACCCG; (ii) *Can* riboswitch UGUGGUUCGCAACCAUCCCACAUAAAAAAACUAG; and (iii) *Ppe* riboswitch GGCCGUAUAACCUCGAUAAUUUGGUUCGGGGGCUCUACUGGGAACCUAAAUCCU AACUACGGCC. The *Tte* and *Can* strands were dissolved in 0.01 M Na-cacodylate pH 7.0; strands were folded by heating to 90 °C for 2 min. Sample tubes were rapidly cooled on ice, and an equal volume of 0.1 M Na-cacodylate containing 0.006 M $MgCl_2$ was slowly added followed

by a 30 min incubation on ice prior to warming to room temperature for SPR analysis. The *Ppe* riboswitch was folded by dissolving the RNA in 0.01 M Na-HEPES pH 7.0, followed by heating to 80 °C for 2 min, then cooling for 2 min at room temperature, followed by the slow addition of an equal volume of 0.01 M Na-HEPES pH 7.0, 0.1 M KCl, and 0.01 M $MgCl_2$. Subsequently, the riboswitch was placed on ice for at least 30 min prior to warming to room temperature for SPR. Each folded riboswitch stock was then diluted in the SPR running buffer to 300 nM prior to immobilization on CM5 chips with a neutravidin surface. The cyclic peptides (**FSS-m1** & **FSS-m2**) were analyzed in an SPR buffer comprising 0.010 M HEPES pH 7.0, 0.30 M NaCl and 0.05% Tween 20 that was passed over surfaces of each immobilized riboswitch (~800–1500 RU). The flow rate, association and dissociation conditions were the same as those used for the biotinylated SARS-CoV-2 RNAs SPR experiments (above). Cyclic peptide concentrations of 6.25 or 12.5 μM were injected and passed over the chip, which is approximately 3.1 to 6.3-fold higher than the K_D measured for **FSS-m1** and **FSS-m2** binding to the FSS-PK (Table 2) and 18–6000-fold greater than the affinity of each riboswitch for its natural ligand; 3 M guanidine HCl was injected twice for 90 s at a flow rate of 60 μL/min to regenerate the riboswitch surfaces.

Cell Viability Assay

Cell viability was determined using HEK293T cells (ATCC, CRL-3216) (a gift from Prof. Sheel Dodani) seeded in 24-well plates (Falcon), and the cells were stored in a 37 °C humidified incubator with 5% CO_2. Per manufacturer's recommendations (ATCC), HEK293T cells were plated at a density of 1 x 10^4 to 4 x 10^4 viable cells/cm^2 in 500 μL of Dulbecco's Modified Eagle Medium (DMEM) (Gibco, 11995065) supplemented with 10% fetal bovine serum (FBS) (Corning, 35011CV) and 1% penicillin/streptomycin (Gibco, 15140122). Once the cells reached a

confluency of 70–80%, the **FSS-m1** cyclopeptide was diluted from a stock solution into the aforementioned culture media and incubated with HEK293T cells to yield final concentrations of 2 μM, 10 μM and 50 μM each with 1% (v/v) dimethyl sulfoxide (DMSO). Similarly, parthenolide (PTL), a natural product which served a positive control, was diluted into the culture media from a DMSO stock solution and added to the cells at a fixed dose of 50 μM with 1% DMSO. As negative controls, culture media without cyclopeptide or PTL was incubated with HEK293T cells either in the presence or absence of DMSO (final concentration of 1%). Following 24 h of incubation, culture media was removed swiftly and discarded. Subsequently, 500 μL of thiazolyl blue tetrazolium bromide (MTT) solution (1 mg/mL) from a light-protected tube was added to the cells. As a comparator, 500 μL of the aforementioned MTT solution was added to a separate 24-well plate in parallel. Following the addition of the MTT solution, the plates were wrapped in aluminum foil and stored in the humidified incubator. Upon incubation at 37 °C for 3 h, the plates were centrifuged (3,400 x g) for 15 min, the media was removed, and 500 μL of DMSO was added to solubilize the formazan product. The absorbance of the samples was measured at a wavelength of 550 nm using a multi-well plate reader (Agilent). Analyses were conducted in triplicate for cells treated with DMSO only and 10 μM **FSS-m1**, while analyses were conducted in quadruplicate for all other cells (i.e., untreated without DMSO, 50 μM PTL treatment, 50 μM **FSS-m1** treatment and 2 μM **FSS-m1** treatment).

Selective 2′ Hydroxyl Acylation Analyzed by Primer extension (SHAPE)-Seq

For chemical-modification experiments, the dimerization loop was synthesized by T7 polymerase.[9] The transcription product contained the 26-mer dimerization loop of the SARS-CoV-2 FSS placed in a folding cassette. The 5′-end of the cassette possessed the stable HIV-1 FSS stable

hairpin, whereas the 3′-end harbored a strong hairpin followed by a unique primer binding site as described.[10] Additionally, the HIV-1 FSS serves as a control for the specificity of the peptides for the SARS-CoV-2 FSS. The 17-mer DNA primer and DNA template were ordered from IDT Inc. The transcript synthesized by T7 polymerase was:

5′-GGCTTCCCACAAGGGAAGCCATGGCGGCACAGGCACTAGTACTGAT

GTCGCCTCGATCCGCTTCGGCGGATCCAAATCGGGCTTCGGTCCGGTTC-3′. The *in vitro* transcribed product was PAGE purified, desalted[9] and stored at −20 °C in 0.5× TE buffer.

Prior to modification, 60 pmol of pure RNA was dissolved in 72 μL of 0.5× TE buffer per modification condition. The RNA was heated to 95 °C for 2 min and placed on ice 2 min. A 36 μL volume of 3× RNA folding mix (333 mM HEPES pH 8.0, 333 mM NaCl and 20 mM $MgCl_2$) was added; for the bound-state conditions, peptide was also added to a final concentration 10× the K_D of the cyclic peptide for the RNA as determined by SPR. The RNA was incubated at 24 °C for 20 min.

The 72-μL-folded-RNA sample was split in half after the addition of 36 μL of folding mix, yielding 18 μL of free-state or bound-state RNA per reaction condition. Acylation was performed with 2-methylnicotinic acid imidazolide (NAI) (Sigma-Aldrich).[11] A volume of 6 μL of 0.6 M NAI in DMSO was added to a concentration of 150 mM. Alternatively, 6 μL of DMSO was added to the control samples. The reactions proceeded 15 min at 24 °C and were terminated by addition of a stop solution comprising 90 μL water, 5 μL 4 M NaCl, 1.5 μL GlycoBlue and 2 μL of 100 mM EDTA pH 8.0. The RNA was ethanol-precipitated by addition of 350 μL of neat ethanol. Precipitated RNA was harvested by centrifugation at 18000 ×g for 40 min at 4 °C. The pellet was washed in 350 μL cold neat ethanol, and re-harvested by centrifugation at 18000 ×g for 40 min at 4 °C. The air-dried pellet was dissolved in 20 μL of 0.5× TE buffer.

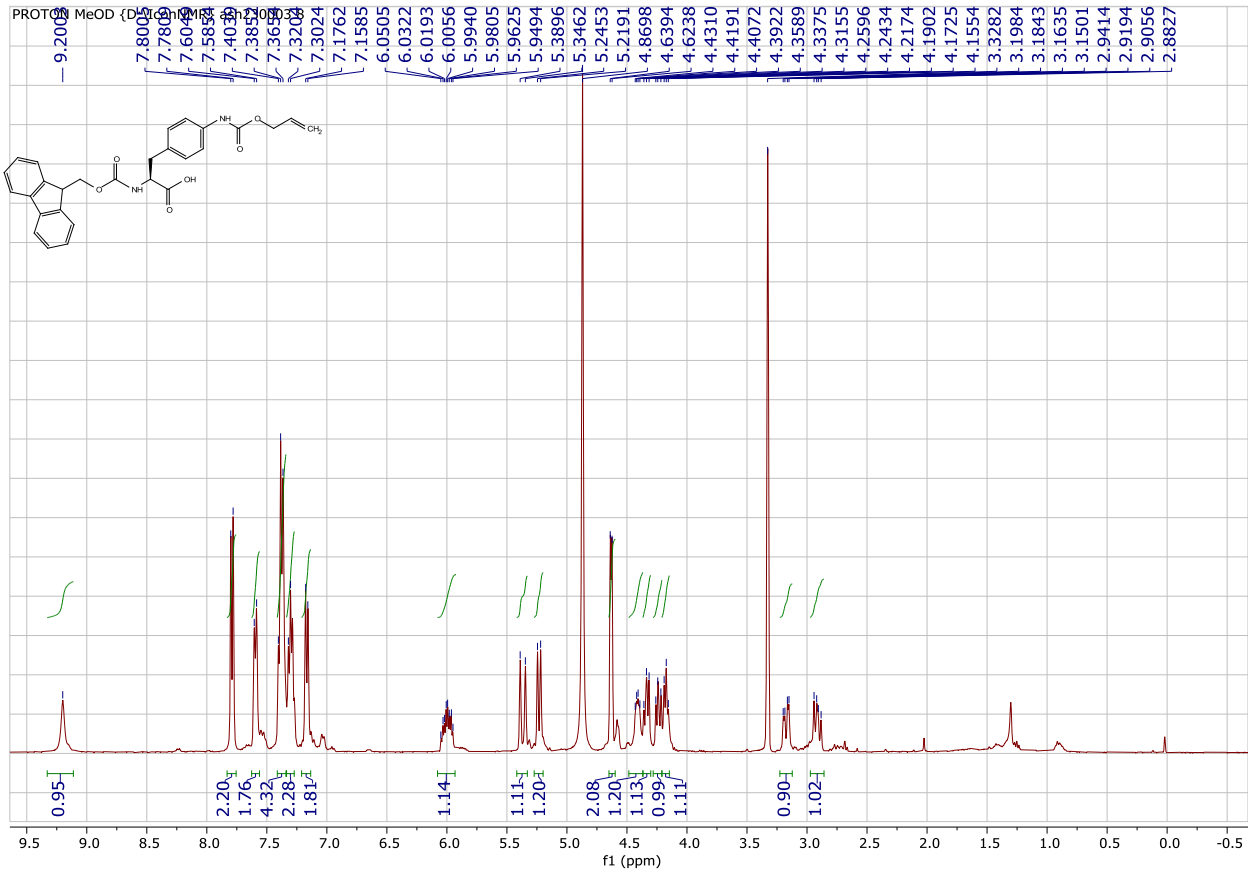
Modification of RNA was analyzed using the SHAPE-seq v2.1 workflow[12] as described[13] with an additional purification step where samples were size selected to using a 2% agarose gel on the PippinHT system (Sage Science, MA). Each sample was analyzed with SPATS version 1.9.30.[14] SHAPE reactivities were calculated using the following target sequence: GGCTTCCCACAAGGGAAGCCATGGCGGCACAGGCACTAGTACTGATGTCGCCTCGA TCCGCTTCGGCGGATCCAAATCGGGCTTCGGTCCGGTTC.

Calculation of logP values

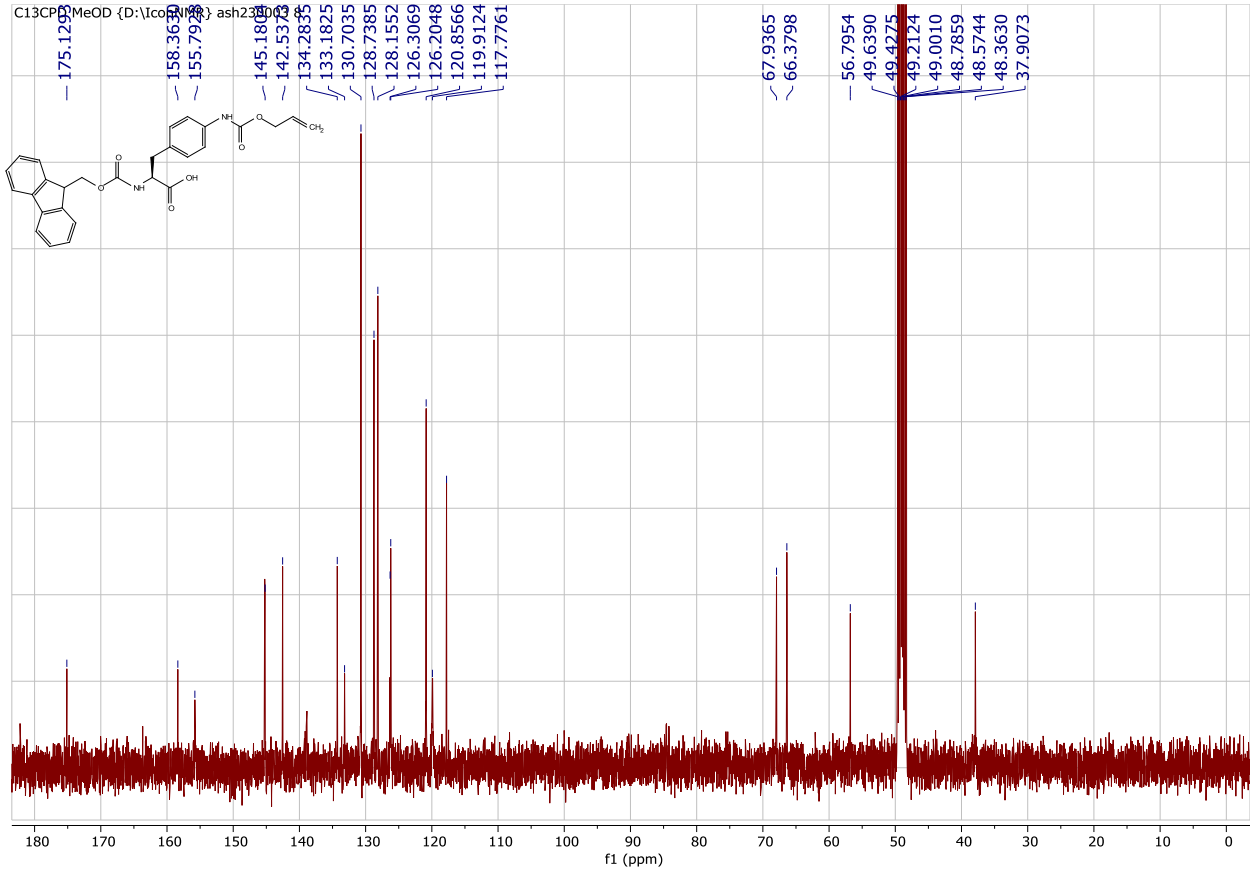
Cyclic peptides were converted into SMILES strings using ChemDraw (Revvity Signals). For each cyclic peptide, the SMILES string was copied into the program AloGPS v2.1,[15] and the logP value was calculated using the non-JAVA-enabled web tool: https://vcclab.org/web/alogps/.

NMR Spectra

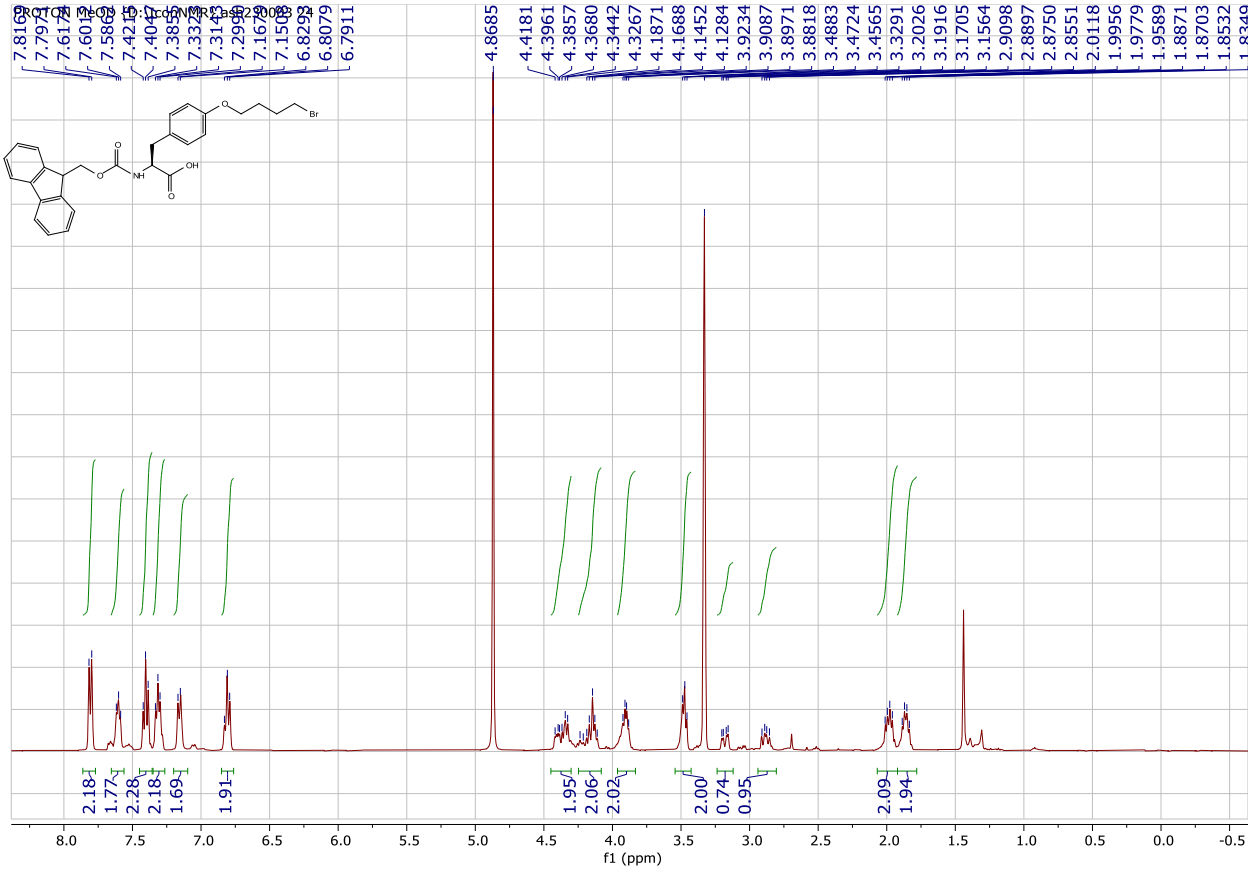
(*S*)-2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(((allyloxy)carbonyl)-amino)phenyl)propanoic acid (N-Fmoc-N´-Alloc-*p*-amino-L-Phe-OH)



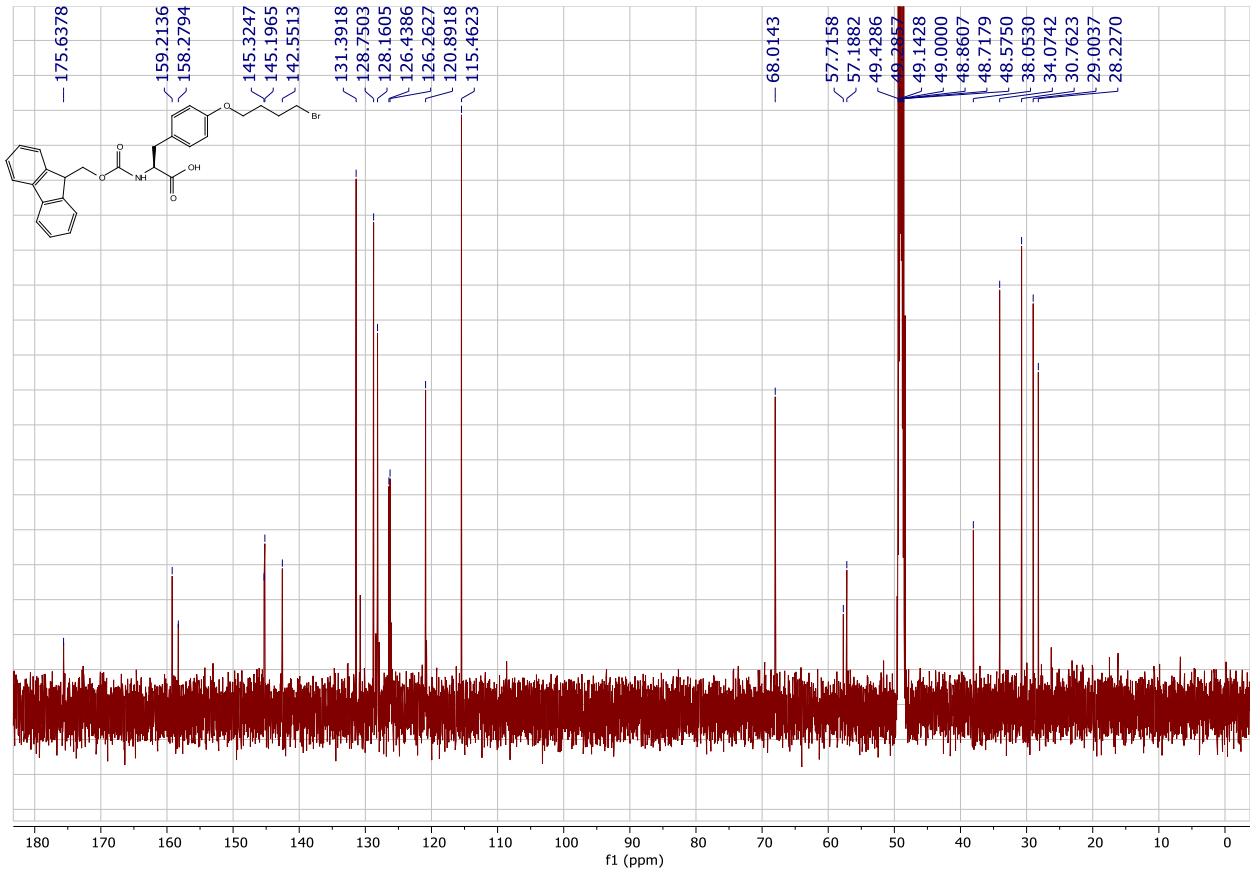
(*S*)-2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(((allyloxy)carbonyl)-amino)phenyl)propanoic acid (N-Fmoc-N´-Alloc-*p*-amino-L-Phe-OH) (cont.)



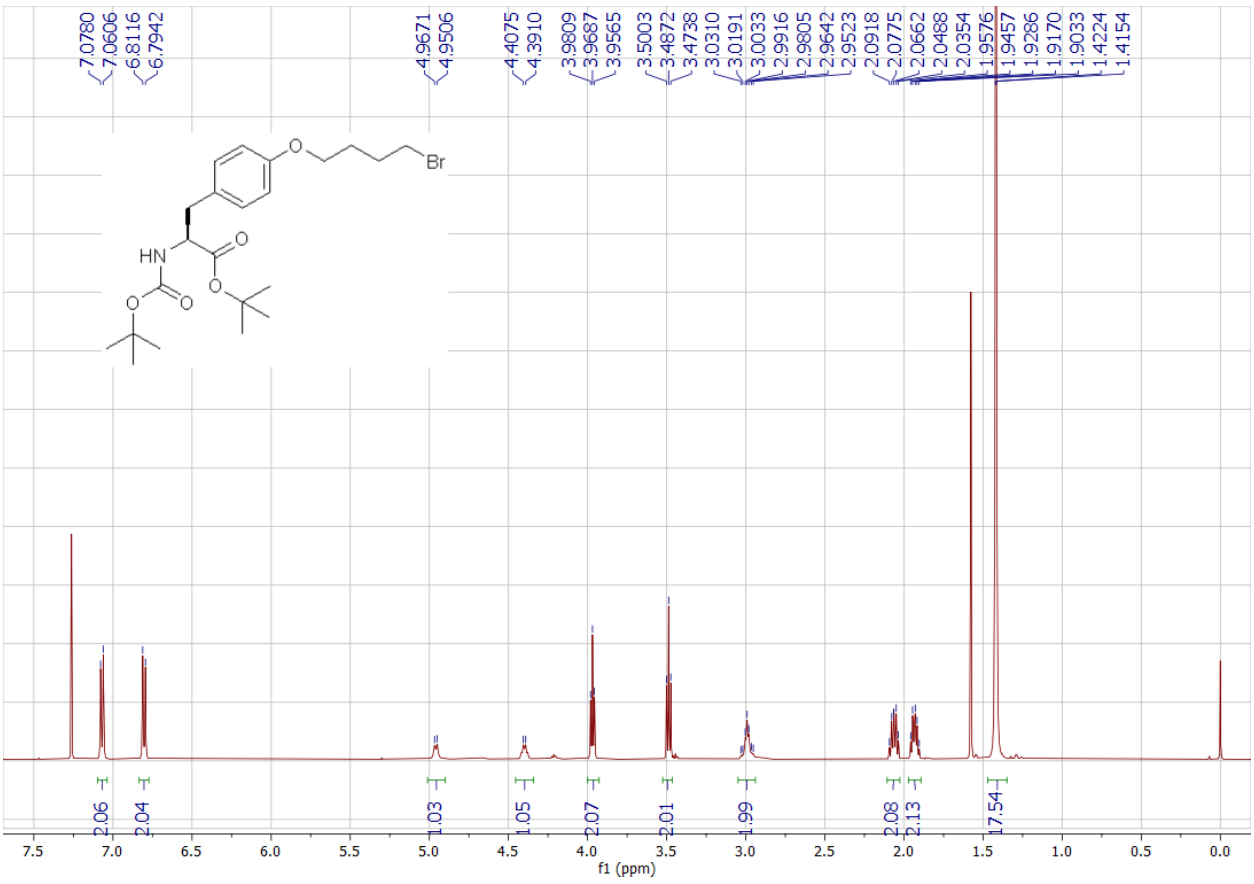
(*S*)-2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(4-bromobutoxy)-phenyl)propanoic acid (Fmoc-O4bbY-OH)



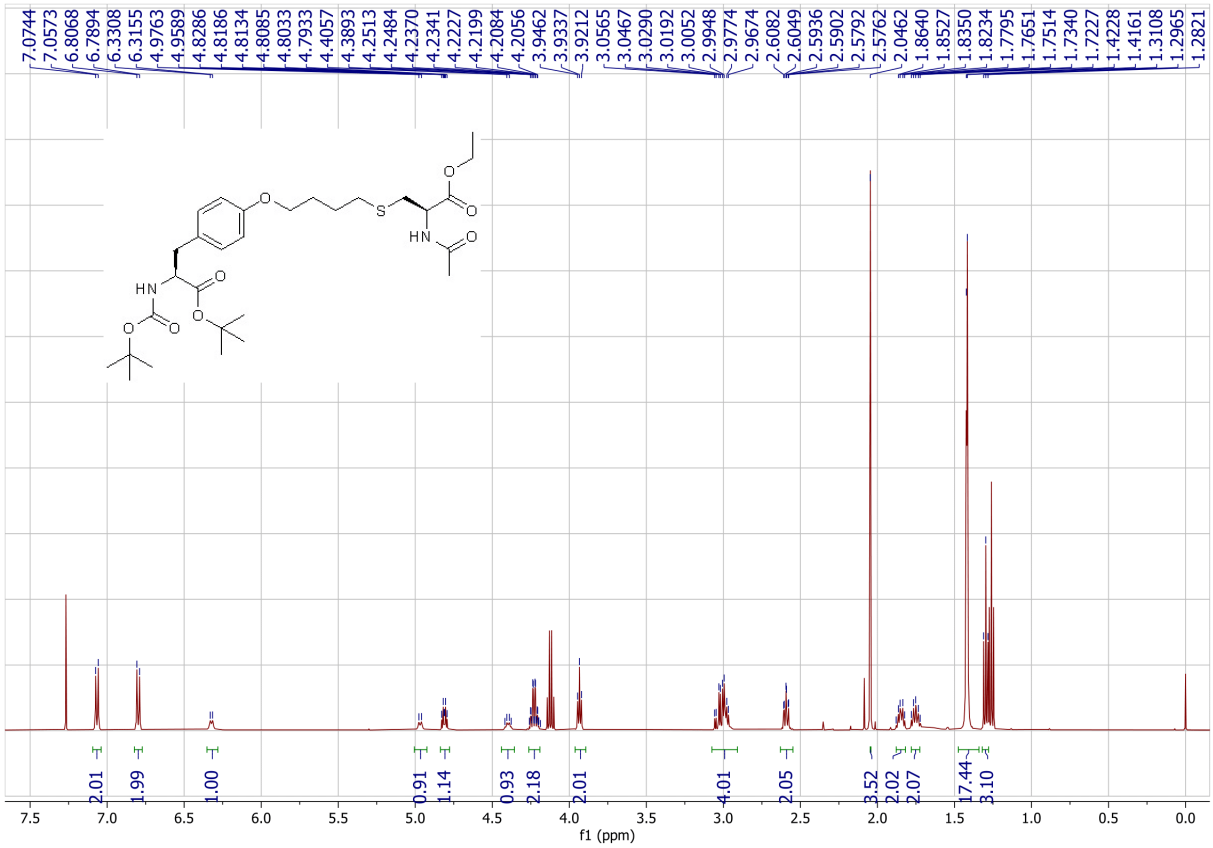
(*S*)-2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(4-bromobutoxy)-phenyl)propanoic acid (Fmoc-O4bbY-OH) (cont.)



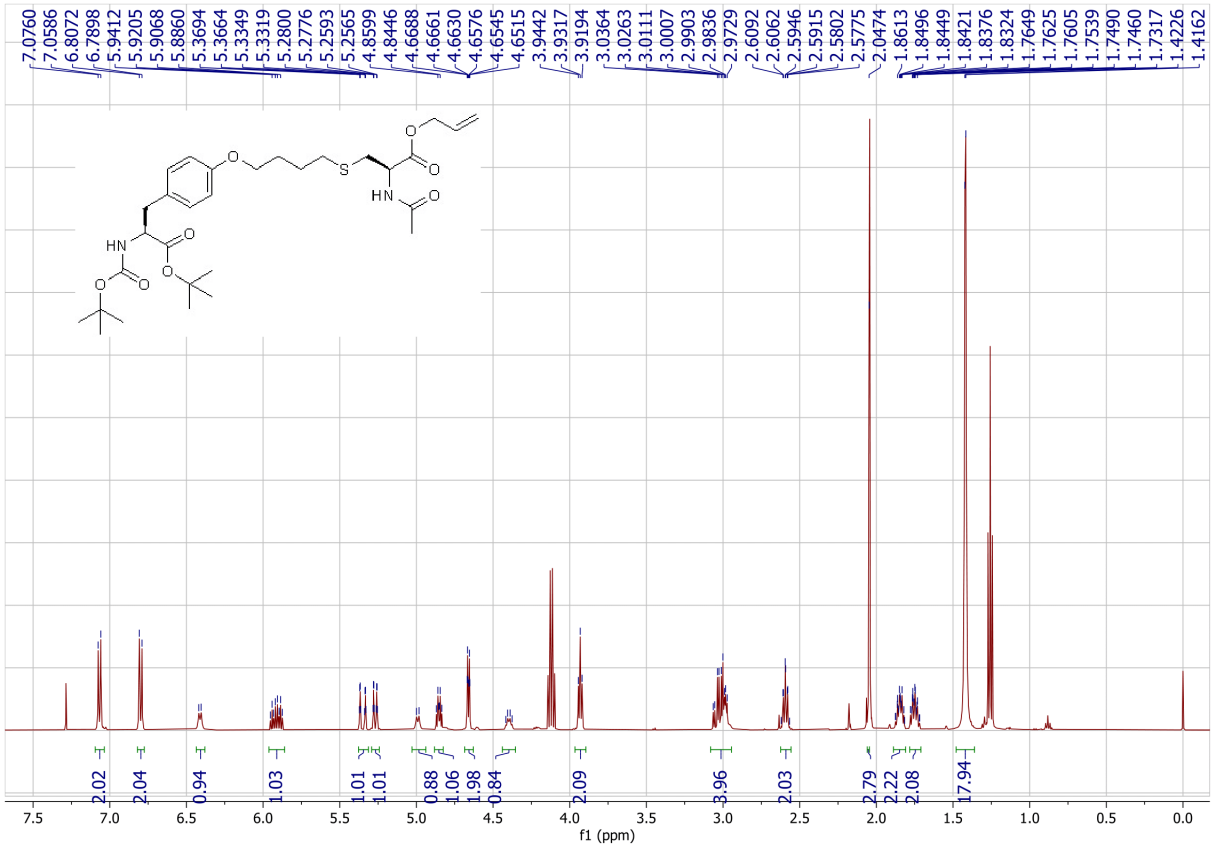
_t_-Butyl (_S_)-3-(4-(4-bromobutoxy)phenyl)-2-((_tert_-butoxycarbonyl)amino)propanoic acid (2)



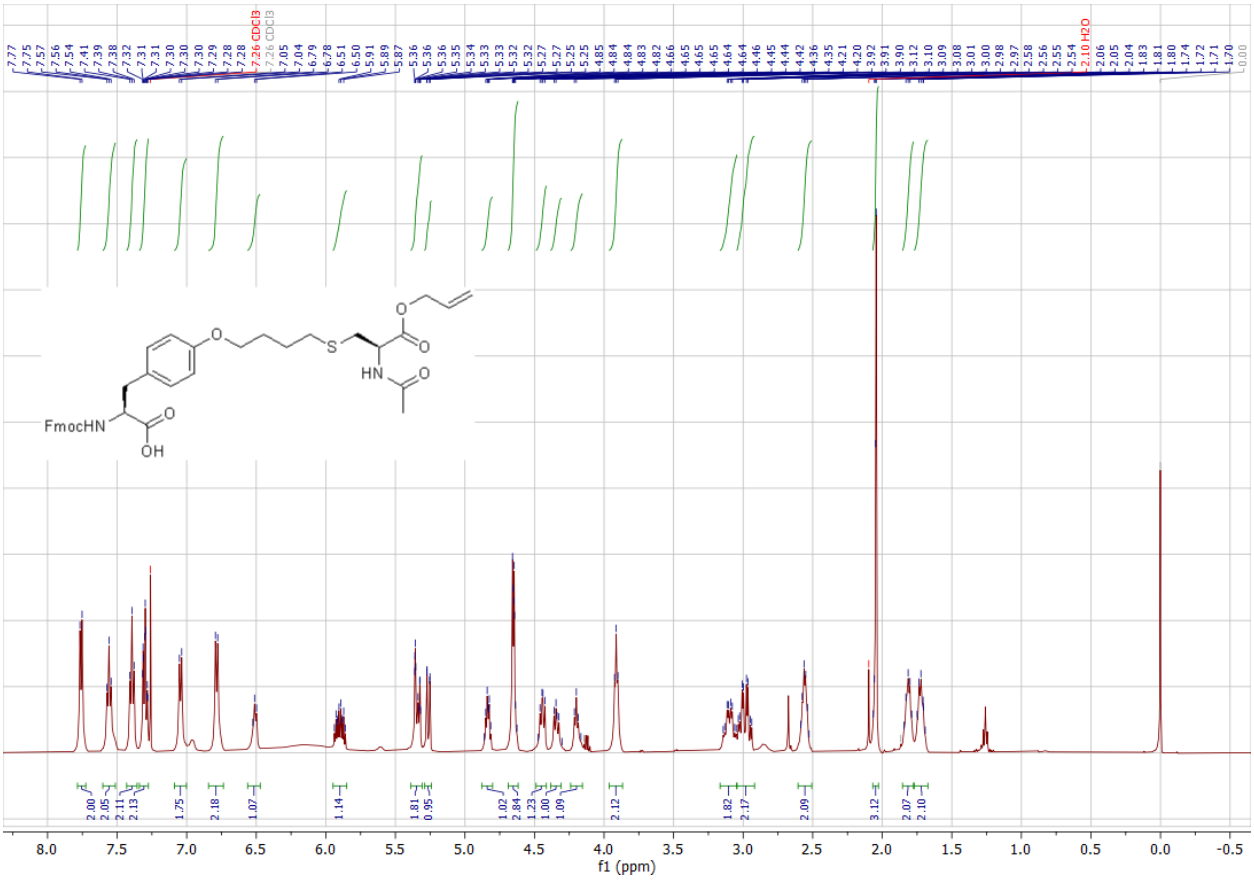
***t*-Butyl (*S*)-3-(4-(4-(((*R*)-2-acetamido-3-ethoxy-3-oxopropyl)thio)butoxy)phenyl)-2-((*tert*-butoxycarbonyl)amino)propanoic acid (4)**



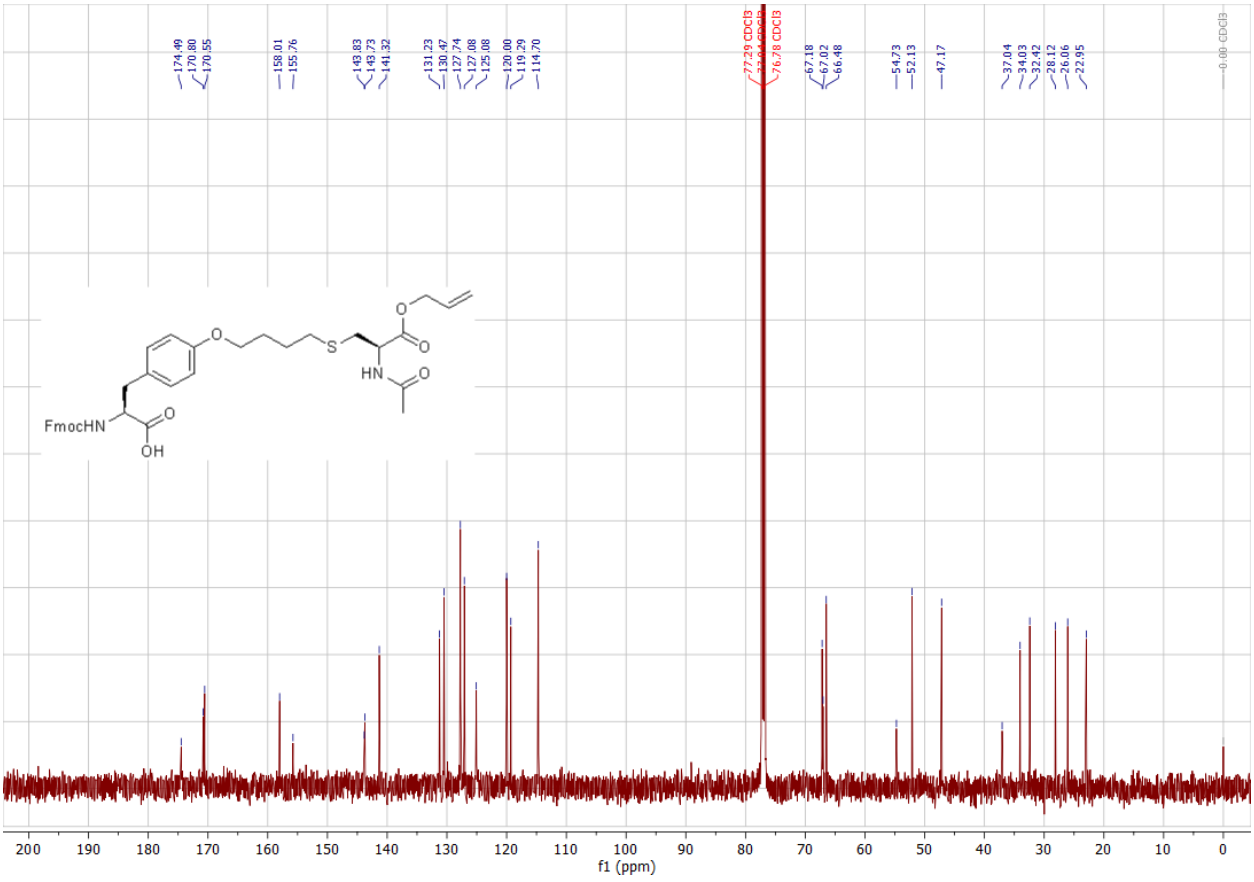
Allyl *N*-acetyl-*S*-(4-(4-((*S*)-3-(tert-butoxy)-2-((*tert*-butoxycarbonyl)amino)-3-oxopropyl)-phenoxy)butyl)-L-cysteine (5)



(*S*)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(4-(((*R*)-2-acetamido-3-(allyloxy)-3-oxopropyl)thio)butoxy)phenyl)propanoic acid (6)



(*S*)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(4-(((*R*)-2-acetamido-3-(allyloxy)-3-oxopropyl)thio)butoxy)phenyl)propanoic acid (6) (cont.)



References

(1) Bionda, N.; Cryan, A. L.; Fasan, R. Bioinspired strategy for the ribosomal synthesis of thioether-bridged macrocyclic peptides in bacteria. *ACS Chem Biol* **2014**, *9* (9), 2008-2013. DOI: 10.1021/cb500311k.

(2) Iannuzzelli, J. A.; Fasan, R. Expanded toolbox for directing the biosynthesis of macrocyclic peptides in bacterial cells. *Chem Sci* **2020**, *11* (24), 6202-6208. DOI: 10.1039/d0sc01699c.

(3) Leelasvatanakij, L.; Aldrich, J. V. A solid-phase synthetic strategy for the preparation of peptide-based affinity labels: synthesis of dynorphin A analogs. *J Pept Res* **2000**, *56* (2), 80-87. DOI: 10.1034/j.1399-3011.2000.00736.x.

(4) Owens, A. E.; Iannuzzelli, J. A.; Gu, Y.; Fasan, R. MOrPH-PhD: An Integrated Phage Display Platform for the Discovery of Functional Genetically Encoded Peptide Macrocycles. *ACS Cent Sci* **2020**, *6* (3), 368-381. DOI: 10.1021/acscentsci.9b00927.

(5) Myszka, D. G. Improving biosensor analysis. *J Mol Recognit* **1999**, *12* (5), 279-284. DOI: 10.1002/(SICI)1099-1352(199909/10)12:5<279::AID-JMR473>3.0.CO;2-3.

(6) Jenkins, J. L.; Krucinska, J.; McCarty, R. M.; Bandarian, V.; Wedekind, J. E. Comparison of a $preQ_1$ riboswitch aptamer in metabolite-bound and free states with implications for gene regulation. *J Biol Chem* **2011**, *286* (28), 24626-24637. DOI: 10.1074/jbc.M111.230375.

(7) Schroeder, G. M.; Cavender, C. E.; Blau, M. E.; Jenkins, J. L.; Mathews, D. H.; Wedekind, J. E. A small RNA that cooperatively senses two stacked metabolites in one pocket for gene control. *Nat Commun* **2022**, *13* (1), 199. DOI: 10.1038/s41467-021-27790-8.

(8) Hamal Dhakal, S.; Kavita, K.; Panchapakesan, S. S. S.; Roth, A.; Breaker, R. R. 8-oxoguanine riboswitches in bacteria detect and respond to oxidative DNA damage. *Proc Natl Acad Sci U S A* **2023**, *120* (40), e2307854120. DOI: 10.1073/pnas.2307854120.

(9) Lippa, G. M.; Liberman, J. A.; Jenkins, J. L.; Krucinska, J.; Salim, M.; Wedekind, J. E. Crystallographic analysis of small ribozymes and riboswitches. *Methods Mol Biol* **2012**, *848*, 159-184. DOI: 10.1007/978-1-61779-545-9_11.

(10) Wilkinson, K. A.; Merino, E. J.; Weeks, K. M. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat Protoc* **2006**, *1* (3), 1610-1616. DOI: 10.1038/nprot.2006.249.

(11) Spitale, R. C.; Crisalli, P.; Flynn, R. A.; Torre, E. A.; Kool, E. T.; Chang, H. Y. RNA SHAPE analysis in living cells. *Nat Chem Biol* **2013**, *9* (1), 18-20. DOI: 10.1038/nchembio.1131.

(12) Loughrey, D.; Watters, K. E.; Settle, A. H.; Lucks, J. B. SHAPE-Seq 2.0: systematic optimization and extension of high-throughput chemical probing of RNA secondary structure with next generation sequencing. *Nucleic Acids Res* **2014**, *42* (21), e165. DOI: 10.1093/nar/gku909.

(13) Schroeder, G. M.; Dutta, D.; Cavender, C. E.; Jenkins, J. L.; Pritchett, E. M.; Baker, C. D.; Ashton, J. M.; Mathews, D. H.; Wedekind, J. E. Analysis of a preQ1-I riboswitch in effector-free and bound states reveals a metabolite-programmed nucleobase-stacking spine that controls gene regulation. *Nucleic Acids Res* **2020**, *48* (14), 8146-8164. DOI: 10.1093/nar/gkaa546.

(14) Watters, K. E.; Yu, A. M.; Strobel, E. J.; Settle, A. H.; Lucks, J. B. Characterizing RNA structures in vitro and in vivo with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Methods* **2016**, *103*, 34-48. DOI: 10.1016/j.ymeth.2016.04.002.

(15) Tetko, I. V.; Tanchuk, V. Y. Application of Associative Neural Networks for Prediction of Lipophilicity in ALOGPS 2.1 Program. *Journal of Chemical Information and Computer Sciences* **2002**, *42* (5), 1136-1145. DOI: 10.1021/ci025515j.