

Cite this: *Chem. Sci.*, 2023, 14, 10561

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

## Selective thiazoline peptide cyclisation compatible with mRNA display and efficient synthesis†

Minglong Liu,<sup>ab</sup> Richard Morewood,<sup>c</sup> Ryoji Yoshisada,<sup>ab</sup> Mirte N. Pascha,<sup>d</sup> Antonius J. P. Hopstaken,<sup>ab</sup> Eliza Tarcoveanu,<sup>c</sup> David A. Poole, III,<sup>ab</sup> Cornelis A. M. de Haan,<sup>d</sup> Christoph Nitsche<sup>ib</sup>\*<sup>c</sup> and Seino A. K. Jongkees<sup>id</sup>\*<sup>ab</sup>

Peptide display technologies are a powerful method for discovery of new bioactive sequences, but linear sequences are often very unstable in a biological setting. Macrocyclisation of such peptides is beneficial for target affinity, selectivity, stability, and cell permeability. However, macrocyclisation of a linear hit is unreliable and requires extensive structural knowledge. Genetically encoding macrocyclisation during the discovery process is a better approach, and so there is a need for diverse cyclisation options that can be deployed in the context of peptide display techniques such as mRNA display. In this work we show that *meta*-cyanopyridylalanine (*mCNP*) can be ribosomally incorporated into peptides, forming a macrocycle in a spontaneous and selective reaction with an N-terminal cysteine generated from bypassing the initiation codon in translation. This reactive amino acid can also be easily incorporated into peptides during standard Fmoc solid phase peptide synthesis, which can otherwise be a bottleneck in transferring from peptide discovery to peptide testing and application. We demonstrate the potential of this new method by discovery of macrocyclic peptides targeting influenza haemagglutinin, and molecular dynamics simulation indicates the *mCNP* cross-link stabilises a beta sheet structure in a representative of the most abundant cluster of active hits. Cyclisation by *mCNP* is also shown to be compatible with thioether macrocyclisation at a second cysteine to form bicycles of different architectures, provided that cysteine placement reinforces selectivity, with this bicyclisation happening spontaneously and in a controlled manner during peptide translation. Our new approach generates macrocycles with a more rigid cross-link and with better control of regiochemistry when additional cysteines are present, opening these up for further exploitation in chemical modification of *in vitro* translated peptides, and so is a valuable addition to the peptide discovery toolbox.

Received 19th June 2023

Accepted 6th September 2023

DOI: 10.1039/d3sc03117a

rsc.li/chemical-science

## Introduction

Peptides are currently experiencing a surge of interest as drug candidates,<sup>1</sup> arising from their strong target engagement, selective binding, and low immunogenicity.<sup>2</sup> However, they still have major drawbacks in their pharmacokinetics, as they are often degraded quickly. Modified peptides, or peptides with non-natural building blocks, can offer solutions to these

challenges, and prime among these is the macrocyclisation of the peptide.<sup>3</sup> This removes one or both termini from access to proteases, while conformational constraint can further reduce access to endopeptidases. It also removes some of the entropic penalty to binding by holding the peptide in a (near-)active conformation, and can improve selectivity by preventing alternate conformations. Macrocyclisation is likely the most impactful single change that can be made to a peptide to address many of the problems associated with their use in a biological setting, although it is rarely enough on its own.<sup>4,5</sup>

The approach used for peptide macrocyclisation can have a large influence on biological activity.<sup>6,7</sup> Macrocyclisation is typically well tolerated if it is already present at the discovery stage of '*de novo*' peptides (those derived from random peptide libraries<sup>8</sup>), as the peptide will be selected to fit the conformational constraints imposed by the cyclisation approach. This means that changes during lead development are minimised. By contrast, macrocyclisation introduced at a later stage can be much more difficult to optimise, requiring extensive structural information and trial and error.<sup>9,10</sup> There is thus a demand for

<sup>a</sup>Department of Chemistry and Pharmaceutical Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. E-mail: S.A.K.Jongkees@vu.nl

<sup>b</sup>Amsterdam Institute of Molecular and Life Sciences (AIMMS), Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

<sup>c</sup>Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia. E-mail: christoph.nitsche@anu.edu.au

<sup>d</sup>Section Virology, Division Infectious Diseases and Immunology, Department of Biomolecular Health Sciences, Faculty of Veterinary Medicine, Utrecht University, Yalelaan 1, 3584 CL Utrecht, The Netherlands

† Electronic supplementary information (ESI) available: Preparation of oligonucleotides, aminoacylation and translation testing, mRNA-display selection protocol, peptide and amino acid synthesis, compound characterisation. See DOI: <https://doi.org/10.1039/d3sc03117a>

peptide macrocyclisation approaches that are compatible with both ribosomal translation and with peptide production techniques such as solid-phase peptide synthesis (SPPS) or bacterial expression.

The peptide discovery technique of mRNA display is particularly suited to discovery of new macrocyclic peptides as the peptide libraries are generated by *in vitro* translation.<sup>11</sup> In this, the building block pool can be easily manipulated, both by addition of new building blocks and by omission of canonical amino acids to generate vacant codons, resulting in a shift in the chemical space of the displayed peptides. A particularly convenient method for adding new building blocks is through the use of aminoacylating ribozymes called flexizymes.<sup>12</sup> These recognise only the CCA acceptor stem of tRNA, so are anticodon-agnostic. They are also promiscuous in substrate scope, having been evolved to recognise only an aromatic group in either the amino acid or activated ester. This combination of flexible *in vitro* translation by means of flexizyme with peptide enrichment against a biologically relevant target using mRNA display has been termed the random non-standard peptide integrated discovery (RaPID) system.<sup>13</sup>

Macrocyclisation in the RaPID system is at present carried out predominantly by initiating peptide translation with a chloroacetylated aromatic amino acid, although a few other methods have been developed.<sup>14</sup> This chloroacetyl moiety reacts with the first downstream cysteine to give a stable non-reducible thioether linkage (Fig. 1).<sup>15</sup> This has many favourable properties, including spontaneous and fast reaction, predictable ring size, and easy scalability in SPPS. However, with the exception of the second position, it cannot produce macrocycles that contain additional cysteine residues (which are particularly useful as reactive handles for further chemical or enzymatic modification<sup>16,17</sup> and for access to multicyclic peptides), it generates a flexible linkage, and it cannot currently be accessed during bacterial expression. Therefore, an increased diversity in macrocyclisation approaches is necessary, which will also give rise to a corresponding increase in diversity of peptide structures and as a result may further increase the scope of targets

accessible to this method. A macrocyclisation approach that works well for one target may not work well for another, and so it is desirable to have a panel of different methods that together increase the chance of success.

The N-terminus of a peptide or protein is unique in its reactivity,<sup>18,19</sup> having a different pK<sub>a</sub> and having neighbouring functional groups not present in lysine, which is the other common amine in peptides and proteins. A host of methods have been developed for specific modification of the N-terminus of a protein, and these can thus serve as inspiration for functional groups that can be incorporated into amino acids for translation into peptides, and thereby afford new and more selective cyclisation reagents. We have recently reported a facile non-enzymatic approach to access peptides in bacterial *in vitro* translation that does not contain a formylated methionine, and so has a free N-terminus available for reaction.<sup>20</sup> Of the N-terminal selective functional groups, aromatic electrophilic nitriles have recently caught broad attention for their rapid formation of a thiazoline heterocycle with an N-terminal cysteine. This reaction can be efficiently exploited in peptide macrocyclisation by incorporation of *meta*-cyanopyridylalanine (*mCNP*) at the C-terminus of a peptide.<sup>21</sup> Based on this, we also recently reported a method for translation of a cyanobenzo-thiazole-containing amino acid that gave efficient peptide macrocyclisation,<sup>22</sup> while others have exploited this same reactive group together with a second cysteine-reactive moiety for stapling of peptides displayed on phage,<sup>23,24</sup> with selectivity arising from either differences in rate or reversibility of the second reaction. However, the high reactivity of cyanobenzo-thiazoles meant that extra precautions needed to be taken during translation of this moiety to prevent premature hydration. A cyanobenzothiazole-containing amino acid was also not directly amenable to SPPS, and so needed to be incorporated in its thiazoline form with the cysteine carboxylate orthogonally protected to allow a late-stage amide bond-forming macrocyclisation. This low yielding and labour-intensive scale-up is thus a hurdle to application.

In the current work we present a facile approach to peptide macrocyclisation in mRNA display (Fig. 1), overcoming all hurdles from our previous studies. By modifying the reactivity of the aromatic nitrile, in the form of a ribosomally translatable *meta*-pyridine nitrile amino acid, all side reactions can be prevented and SPPS becomes routine, while still retaining sufficiently fast macrocyclisation for application in library building in the RaPID system. Notably this new approach has selectivity over additional cysteine residues, which we show can be exploited in controlled bicyclisation during *in vitro* translation.

## Results and discussion

### Flexible *in vitro* translation of a pyridine nitrile amino acid

In order to be able to charge *mCNP* onto tRNA by flexizyme, two activated esters were synthesised with cyanomethyl ester and dinitrobenzyl ester leaving groups. These esters serve as substrates for eFx and dFx type flexizymes, respectively, with eFx and cyanomethyl esters typically being used for aromatic amino acids. Both options were synthesised here to increase the

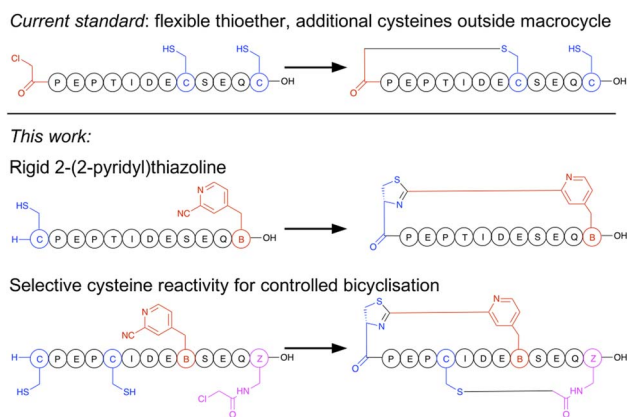
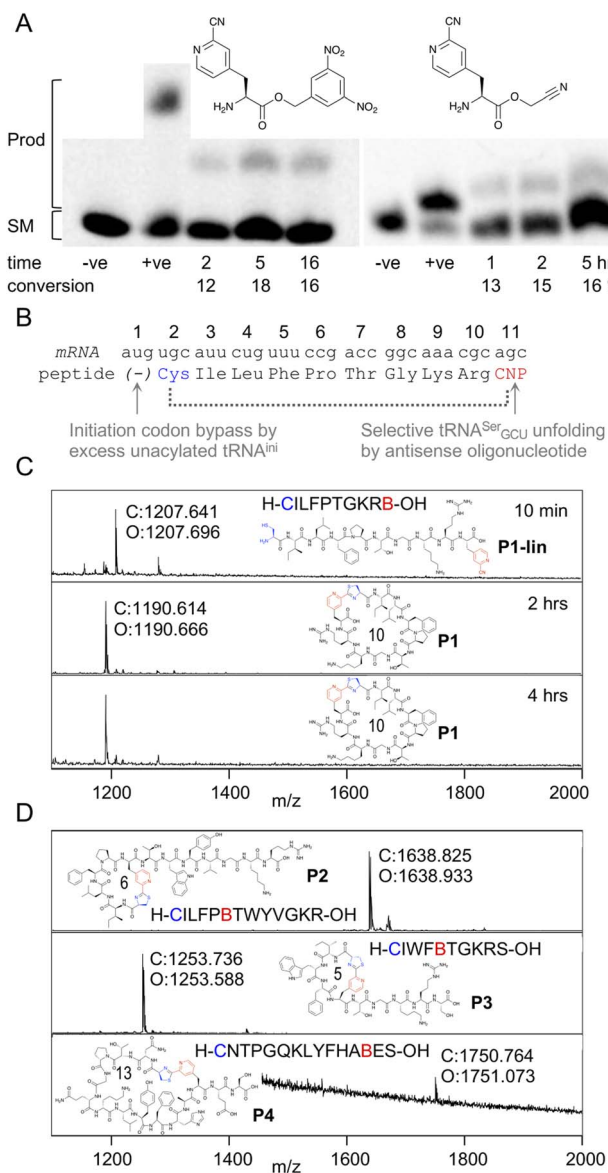


Fig. 1 Comparison of selectivity between the established approach to peptide macrocyclisation in mRNA display with that in the current work.





**Fig. 2** Aminoacylation and translation with *mCNP*. (A) Charging of activated esters shown onto a fluorescein-labeled 5-base tRNA analogue, analysed by urea PAGE. Starting material RNA and product acylated RNA are indicated to the left and time points below the gel, along with negative control '–ve' (DMSO only) and positive control '+ve' (Lys-DBE for 5 hours and Ac-Phe-CME for 2 hours) for each. Conversions are estimated by densitometry of fluorescent scan of gel. Full gels are presented in Fig. S1 and S2.† (B) Schematic representation of the template sequence used to test *mCNP* incorporation by *in vitro* translation, indicating cyclisation by dotted bracket and annotating how each relevant codon was liberated. *CNP* = *meta*-cyanopyridylalanine (C) MALDI-TOF-MS spectra after varying time points for peptide translations using tRNA acylated with *mCNP* by dFx, showing calculated (C) and observed (O) masses. Peptide linear sequence and structure is as indicated with 'B' as one-letter code for *mCNP*, with ring size in amino acids also shown. Pyridine nitrile (derived) moieties are indicated in red and 1,2-aminothiol (derived) moieties in blue. Peptide calculated and observed masses are shown by the peak. (D) MALDI-TOF-MS spectra for peptide translations with varied ring sizes and amino acid compositions, coloured as for B.

chance of success. These were prepared by alkylation of the Boc-protected amino acid with the relevant alkyl chloride before deprotection under acidic conditions. Using these activated esters, acylation onto a fluorescent 5-base tRNA mimetic substrate was assessed by urea PAGE analysis after varied incubation times (Fig. 2A).<sup>25</sup> These showed acylation occurring for both substrates, with the reaction peaking after 5 hours incubation at 18% for dFx and at 16% for eFx. However, we observed hydrolytic instability during storage with the cyanomethyl ester, and so we opted to continue with the more stable dinitrobenzyl ester for further experiments despite this being atypical for aromatic amino acids.

Translation was tested in a model peptide with sequence H-CILFPTGKRB-OH (**P1**, where B is used as a one letter code for *mCNP*). Methionine was omitted from the translation, and an excess of non-acylated initiator tRNA was used to increase the efficiency of translation start from the second codon (cysteine). The *mCNP* amino acid was charged onto the engineered<sup>26</sup> tRNA<sup>EnAsnE2</sup><sub>GCU</sub> to reprogram the serine AGC codon and native tRNA<sup>Ser</sup><sub>GCU</sub> was sequestered by binding to an antisense oligonucleotide (Fig. 2B).<sup>27</sup> Observing the product by MALDI-TOF-MS after 10 min of translation showed a peak matching the expected mass for the linear product (Fig. 2C), while after 30 min translation and a further 1.5 hour incubation only the macrocyclic product was observed. A longer incubation, for a total of 4 hours, showed no further change to this product and thus confirms that the macrocyclic peptide is stable in the *in vitro* translation reaction mixture across the range of time spans typically involved in mRNA display.

We next assessed the ability of the macrocyclisation to tolerate different ring sizes and amino acid compositions in peptides **P2** through **P4** (Fig. 2D). Rings of 5, 6, 10, and 13 amino acids were all well tolerated (including N-terminal Cys and *mCNP* in the ring size), showing only macrocyclic product after 30 min of translation and 1.5 hours of cyclisation. Collectively, these sequences cover the canonical amino acids most likely to interfere with the reaction (S, T, Y, R, K, H), and no signs were seen of side-reactions. These peptides span the size range typically used for mRNA display experiments using the RaPID system, and so indicate that the *mCNP* cyclisation is well suited to this application.

Because the *mCNP* cyclisation is selective for the 1,2-aminothiols of N-terminal cysteines,<sup>21</sup> it was anticipated to be tolerant of other cysteines in the peptide. To demonstrate the utility of this we translated a further two test peptides **P5** and **P6**. These each contained two cysteines, one at the start and another in the middle of the sequence (Fig. 3). In these we also translated two reactive groups, *mCNP* and L-N-β-chloroacetyl 2,3-diaminopropionic acid (ClAC-Dap or Cdp). This second non-canonical amino acid was incorporated by acylation onto tRNA<sup>EnAsnE2</sup><sub>CCA</sub> by dFx from a dinitrobenzyl ester<sup>28</sup> and translated at a vacant codon created by further omission of tryptophan in addition to methionine from the translation reaction solution. **P5** and **P6** differ in the relative placements of these reactive groups, to access different peptide crosslinking architectures.



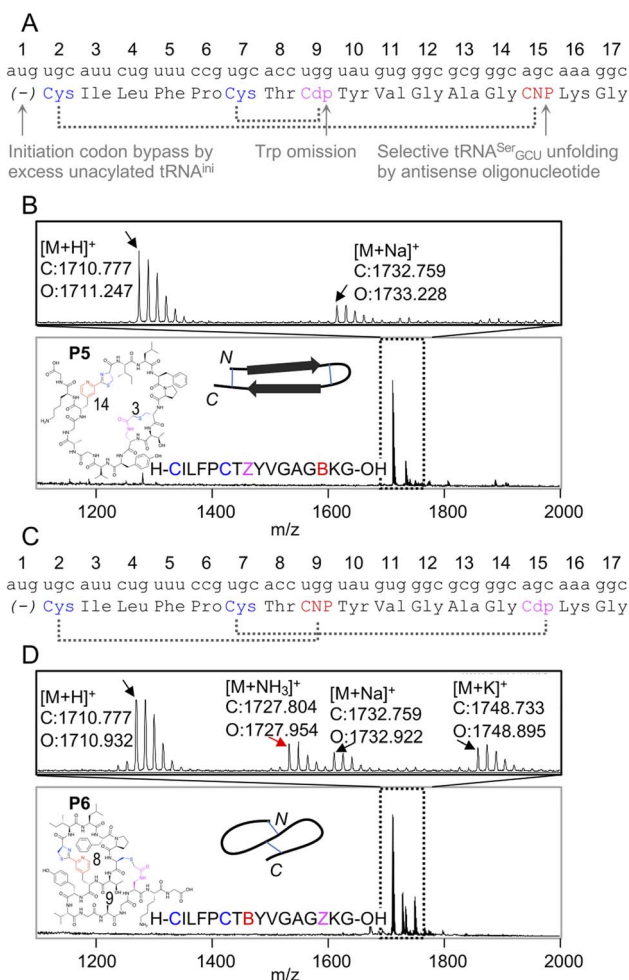


Fig. 3 Controlled bicyclisation of a test peptide with two different cysteine-reactive electrophiles; chloroacetamide and pyridine nitrile. (A) Schematic representation of the test template and reprogrammed translation, as for Fig. 2B. Cdp = L-N-β-chloroacetyl 2,3-diaminopropionic acid. (B) MALDI-TOF-MS showing peptide product matching calculated mass for bicyclisation, with expansion above showing specific peaks and their identities. Overlaid on the spectrum is the product structure, a cartoon representation of cyclisation pattern, and the linear sequence of the peptide (with 'B' as one-letter code for mCNP and 'Z' for ClAc-Dap, and ring sizes in amino acids indicated). (C) As for (A), but with alternate codon assignments for the two reactive amino acids. (D) As for (B), but showing a minor product with ammonia adduct consistent with incorrect macrocyclisation (red arrow).

In designing these test peptides, we considered two selectivity 'rules' for the two reactive groups: chloroacetamides have previously been shown to have good selectivity for the closest thiol when incorporated at the N-terminus.<sup>15</sup> and mCNP has been shown to form stable products exclusively with an N-terminal cysteine.<sup>21</sup> Of these, the chloroacetamide is the less selective, and so imposes constraints on the design. As long as the ClAc-Dap is placed so that the closest cysteine is not the N-terminal cysteine, we expect both selectivities to reinforce one another and bicyclisation to proceed with sufficient control. Thus, while these two reactive groups are not orthogonal,

careful peptide design can give spontaneous and controlled bicyclisation.

Important for characterisation of the product is that both cyclisations are able to be independently monitored in MALDI-TOF-MS through the loss of masses corresponding to the Cl and NH<sub>3</sub> leaving groups (from ClAc-Dap and mCNP, respectively). Undesired reaction of the N-terminal cysteine with the internal ClAc-Dap would then be expected to retain the NH<sub>3</sub> leaving group, with the internal cysteine unable to form a stable adduct to mCNP.

We observed only the mass corresponding to the bicyclic product for **P5**, demonstrating the formation of a large macrocycle with a further staple as could be used in stabilisation of a beta sheet structure.<sup>29</sup> In **P6**, these reactive groups were reversed in an otherwise identical background to instead give a 'figure 8' type structure. Notably, this spacing of *i*; *i* + 7 or *i* + 8 could be expected to stabilise a helical fold, and again for **P6** the major product was the correctly bicycled peptide. In this second case a minor peak was also observed that corresponds to the mass of an ammonia adduct, and thus indicates alternative cyclisation, but such a minor side product is not anticipated to be a problem in applying this method for bicyclic library construction.

These two peptides thus demonstrate that the two cysteine-based reactions can be used together for controlled peptide bicyclisation following *in vitro* translation despite not being fully orthogonal. Placement of the chloroacetamide C-terminal to both cysteines was expected to maximise selectivity of this less-selective pair, allowing it to exploit differences in reactivity from cysteine proximity. We observed that in **P5** the selectivity was better, and so this indicates that selectivity of this reaction benefits from being in a smaller macrocycle (3 amino acids in **P5** vs. 9 in **P6**). While placement of the reactive groups is constrained by these factors, and so this method cannot access all possible bicyclisation architectures, we see no reason all other positions could not be randomised to generate bicyclic peptide libraries in future.

In the setting of reprogrammed *in vitro* translation, the mCNP macrocyclisation reaction thus proved to be robust, selective, and efficient, consistent with previous literature on cyclisation following SPSPS<sup>21</sup> and bacterial translation using an orthogonal aaRS enzyme.<sup>30</sup> Its complementarity to other thiol-based cyclisations also allows spontaneous ribosomal generation of peptides with more complex bicyclic architectures, provided that cysteine placement is carefully considered.

### Selection of thiazoline-macrocyclised peptides targeting influenza haemagglutinin

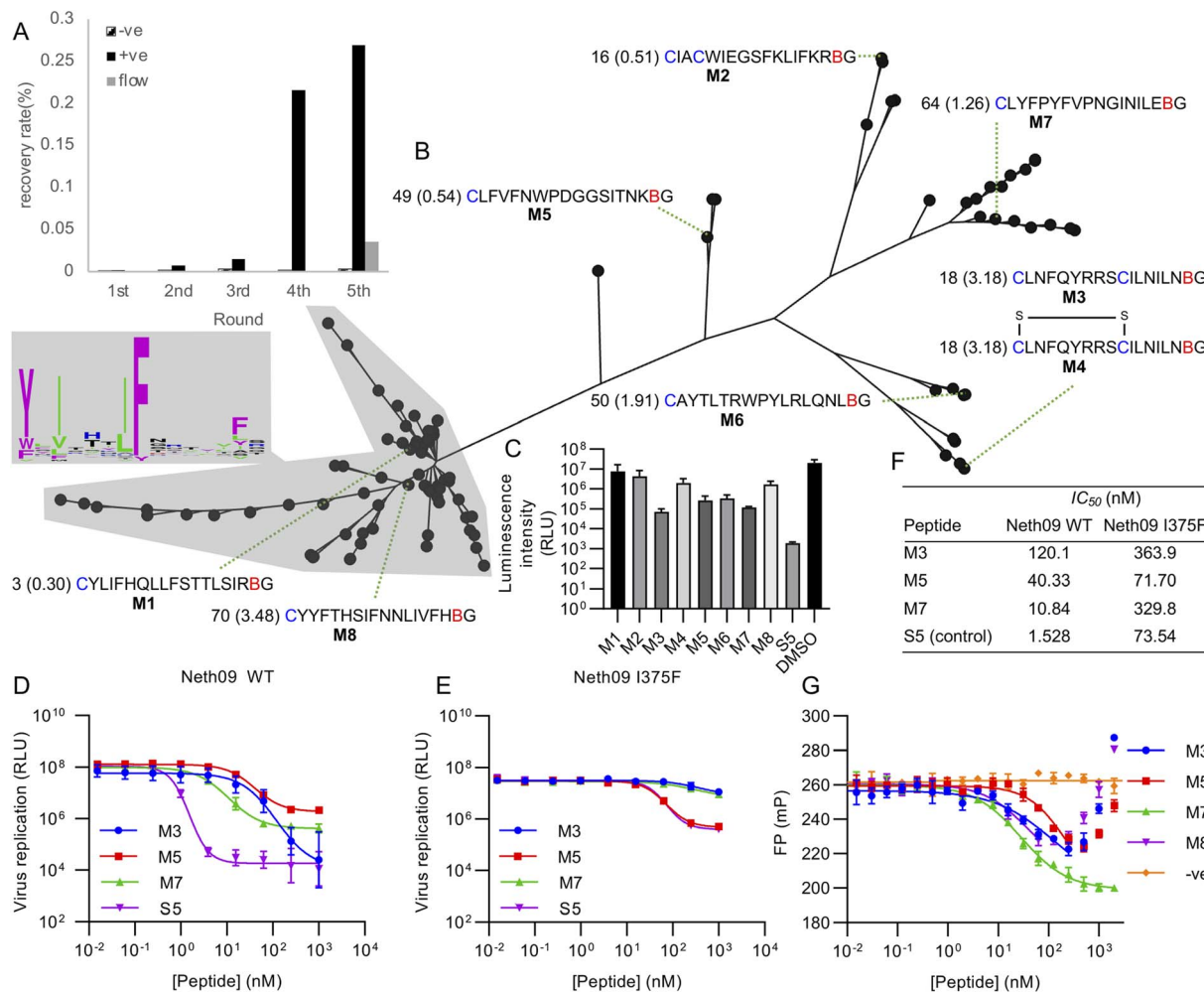
With this validation of incorporation by *in vitro* translation and subsequent macrocyclisation, we next investigated the application of mCNP macrocyclisation in mRNA display selection. For this we took as model target the haemagglutinin (HA) protein from influenza A virus, responsible for entry into host cells through its receptor binding and fusion activity and a potential target for new treatments. It is currently largely targeted by antibodies and hence smaller modalities that



capture the same type of interaction are advantageous. We have previously carried out a RAPID selection campaign against this target using thioether macrocyclisation,<sup>31</sup> and we hypothesised that a new macrocyclisation approach would give rise to a different pool of sequences, potentially also binding at different sites.

A new mRNA display library was designed that contained standard sites for T7 RNA polymerase recognition and ribosome binding, an encoding region for a random peptide of sequence MCX<sub>15</sub>MGGAGAS (NNK codons for X), and finally a puromycin oligonucleotide annealing site to allow mRNA display using the

TRAP variant.<sup>32</sup> For this initial validation of the method, we opted for a relatively simple monocyclic library. Translation of this library was carried out similar to above, with the initiating methionine codon bypassed by addition of an excess of uncharged synthetic initiator tRNA and omission of methionine to give translation start from the second codon (a fixed N-terminal Cys), while the C-terminal methionine codon was recoded to *mCNP* charged onto tRNA<sup>EnAsnE2</sup><sub>CAU</sub> by dFx. Additional cysteine and methionine codons can occur within the random regions, but the former was shown above to be compatible with macrocyclisation while the latter would



**Fig. 4** Selection of *mCNP*-macrocyclised peptides targeting influenza haemagglutinin (subtype H1, isolate A/Netherlands/602/2009 – 'Neth09'). (A) Enrichment of library across selection rounds, plotting recovered DNA as percentage of input library. Positive ('+ve') indicates recovery with HA target immobilised on magnetic beads, negative ('-ve') indicates direct recovery with the Streptactin-coated magnetic beads used for target immobilisation, and 'flow' indicates washing under continuous flow conditions. (B) Phylogenetic tree representation for alignment of high throughput sequencing results from round 5 shown in panel (A), with sequence logo showing conservation by letter height in the main cluster highlighted. Sequences chosen for synthesis are named **M1** through **M8**, with the sequences preceded by the abundance rank as well as the difference in recovery by flow washing vs. batch washing in parentheses (high number indicates higher sequence count following flow washing). (C) Screen for peptide inhibition activity at 1  $\mu$ M peptide with H1N1pdm09 virus (A/Netherlands/602/2009) in a luciferase reporter assay. S5 is a positive control thioether-macrocyclised peptide inhibitor from previous work.<sup>31</sup> (D) Inhibition of infection in the same assay from active peptides derived from this work, along with positive control peptide S5. (E) As for C, but with an I375F resistance mutant. (F) Calculated IC<sub>50</sub> values from panels (D) and (E). (G) Fluorescence polarisation competition assay with fluorescein-tagged S5 as probe and H1 as target. An unrelated SARS-CoV-2 spike-binding peptide S1b3inL1 was used as negative control ('-ve').<sup>34</sup> Data are shown as mean  $\pm$  SD for biological replicates ( $n = 2$  for panel (C), 3 for panels (D), (E) and (G)).



potentially give multiple regioisomers which could in principle be deconvoluted after hit identification if needed and should therefore not be detrimental to library enrichment.

Using this library, five rounds of selection were carried out with clear enrichment across these (Fig. 4A). The enriched library from the fourth round was then also subjected to a further round using continuous flow rather than batch washing, to attempt to drive further enrichment of sequences with a low off-rate.<sup>33</sup> This showed a large drop in recovery, but still above background. Negative control recovery with the same Strep-Tactin beads used for target immobilisation remained consistently low. Enriched libraries were then sequenced in high-throughput to identify candidate hits.

Analysis of the sequencing results revealed one main sequence family with a central YXIXXXIF motif (Fig. 4B), as well as several other smaller and less well conserved sequence families. We did not find any sequences with clear homology to hits from our previous work,<sup>31</sup> suggesting that the alternate cyclisation developed here influences the pool of candidates to adopt different interactions with the target. From these sequence families we selected one representative member each (named **M1** through **M8**), covering sequences that were both more and less strongly enriched in the final continuous-flow round to determine if this late-stage application of flow-based enrichment was helpful in finding more potent hits. We also chose two sequences containing additional cysteine residues (**M2** and **M3/M4**). These were then synthesised on solid phase by standard Fmoc SPPS as previously described,<sup>21</sup> and purified by HPLC. Unlike our previously reported cyanobenzothiazole-based macrocyclisation, this pyridine nitrile building block is stable to piperidine and does not need additional protection, making it much more convenient in scale-up. For one of the peptides with two cysteines we isolated two forms, one correctly cyclised (**M3**) and one forming a disulfide and retaining the *m*CNP residue (**M4**). As this could potentially be a relevant form present in the selection, we decided to test both rather than reduce the disulfide and force macrocyclisation.

All synthesised peptides were tested for inhibition using a cell-based luciferase infection assay with H1 (Fig. 4C). Screening at a single concentration showed clear infection inhibition for **M3**, **M5**, **M6**, and **M7**, while a similar assay with a mutant (I375F) resistant against our previous hits showed most potent activity with **M5** (Fig. S3†). Notably, we found minimal to no activity with either **M1** or **M8**, which were representatives of the main cluster of hits. Further profiling of the most promising hits **M3**, **M5** and **M7** showed the most potent hit to be **M7** (IC<sub>50</sub> of 10 nM; Fig. 4D and F), while against the resistance mutant **M5** showed essentially unchanged activity (IC<sub>50</sub> 50 nM WT vs. 70 nM I375F, Fig. 4E and F). This activity clearly demonstrates that our test selection was successful. **M4**, the disulfide variant of **M3**, did not show inhibition and so indicates that the correctly cyclised form was likely present during the selection and that the disulfide formation is a result of the different conditions in SPPS macrocyclisation vs. after *in vitro* translation. We did not observe any pattern in activity correlating with increased recovery in the continuous flow selection round, but we did not test binding

kinetics directly and hence it remains possible that this would show a clearer effect in the off-rate.

To further assess the influence of our new cyclisation approach on peptide binding to the target, we used a competitive fluorescence polarisation assay to determine if the new hits identified here are likely to be binding at the same site as our previous hits (Fig. 4G). We also tested one of the non-active peptides from the main sequence family to determine if these might be inactive because of binding in a different location, and further included as a negative control our recently reported SARS-CoV-2 spike protein binding peptide.<sup>34</sup> Polarisation was seen to decrease for all new hits in a concentration-dependent manner, but eventually increased again for all but **M7**. We attribute this increase to non-specific aggregation effects at higher concentrations. These results clearly show that peptides **M3**, **M5**, and **M7** all compete with fluorescein-labelled **S5** for binding to H1 protein, and that peptide **M8** also showed competition despite not being active. Hydrogen-deuterium exchange footprinting was previously used to show that all of our earlier hits bind to helix A in the stem of HA and thereby prevent conformational changes leading to fusion of viral and endosomal membranes,<sup>31</sup> and this pattern appears to have been maintained in this new library. That the inactive peptide **M8** still shows competition suggests it is binding in a partially overlapping site that does not have the same influence on HA conformational dynamics, and thus does not lead to inhibition. This likely also explains why **M5** and **M7** plateau at less than full inhibition. It remains unclear why this site is so strongly dominant in our *in vitro* peptide selection experiments when the sialic-acid binding 'head' domain is dominant for antibodies,<sup>35</sup> but the high conservation of this region makes it promising for the development of broadly-active macrocyclic peptides.<sup>36</sup>

### Structural effects of thiazoline macrocyclisation

Peptides with *m*CNP macrocyclisation that were enriched for binding to H1 clearly differed from our previous hits against this same target. This difference presumably arises from the unique conformational constraints imposed by the macrocyclisation, and so we sought to visualise what influence this cross-link has on peptides of this size. We carried out molecular dynamics (MD) simulation of peptide **M7**, as well as an analogue with a crosslink derived from *para*-substituted pyridine nitrile (*p*CNP) and a linear version (with acetylated N-terminus and *m*CNP substituted to phenylalanine). Input structures for the linear version derived from the iTasser server<sup>37</sup> showed a strong propensity for beta sheet across the entire length of the peptide, with a turn at Pro9 (Fig. S4†). MD simulations showed this to be stable but with the two termini more dynamic, as would be expected for a short linear peptide (Fig. 5A). Introduction of a crosslink from *m*CNP decreased this terminal flexibility, reinforcing the beta sheet across residues Phe4 through Leu15 (Fig. 5B). The remaining amino acids form an unusual bulged turn structure that appears to be stabilised by hydrophobic interactions of the thiazoline ring with the Phe4 and Leu15 sidechains (Fig. S5†). Changing the crosslinker to *para* rather than *meta* substituted pyridine nitrile had



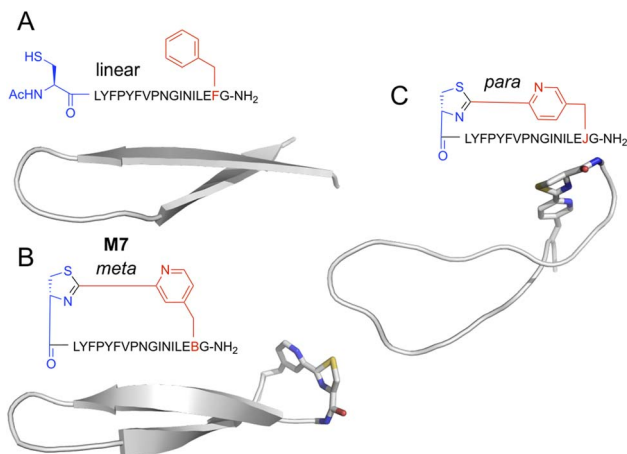


Fig. 5 Representative structures for top-ranked clusters from MD simulations of alternately-cyclised versions of peptide M7. Backbone is shown as cartoon representation, and non-canonical residues as stick. Atoms are coloured by type (grey, carbon; blue, nitrogen; red, oxygen; yellow, sulfur). (A) Linearised version, with the N-terminus acetylated and the pyridine nitrile replaced with phenylalanine. (B) *Meta*-substituted cross-link, as used in experimental work here. (C) *Para*-substituted cross-link, as previously reported by Iskandar *et al.*<sup>40</sup> Peptide sequences are as indicated semi-structurally, with 'B' as one-letter code for *mCNP* and 'J' for *pCNP*.

a dramatic effect, disrupting the beta sheet and driving the peptide to a more flexible loop conformation (Fig. 5C), with several hydrophobic sidechains forming a small core. These models indicate that the rigidity of the pyridine-thiazoline crosslink has a strong influence on the conformation of the peptide backbone, it introduces a unique functional group that can engage in stabilising sidechain interactions, and differences in the regiochemistry of the crosslink can lead to large scale differences in peptide conformation. We thus conclude that *mCNP* as a cyclisation moiety can give unique conformational influences on a macrocyclic peptide library, even in a relatively large ring.

### Comparison to alternate approaches

The most common cyclisation approach currently in use for reprogrammed mRNA display is by thioether formation from an N-terminal chloroacetamide to a C-terminal cysteine.<sup>15</sup> That method suffers from limited regioselectivity if more than one cysteine is present (typically the first in the sequence will react, although that balance can be influenced by choice of initiating amino acid<sup>38</sup>), and also gives a flexible linkage. It is hence incompatible with additional cysteine residues within the macrocycle, which are valuable handles for further bio-conjugation or bicyclisation. The more rigid thiazoline that results from the method we describe here, which also has less hydrogen bond donors from the loss of the N-terminal amide NH, has been shown to give improved permeability in a parallel artificial membrane permeability assay (PAMPA) as compared to analogous thioether or amide cyclised peptides.<sup>39</sup> We consider these two cyclisation approaches to be complementary, giving access to pools of peptides with different

conformational influences and so potentially different hits, as demonstrated by our selections against influenza HA. The ability to produce *mCNP*-cyclised peptides by bacterial culture<sup>30</sup> is a further advantage of our current method here, taking advantage of an aminoacyl tRNA synthetase that can efficiently charge both *p*- and *mCNP* and has been shown to give in-cell peptide macrocyclisation.

Recently published work from the Bowers group showed a related cyclisation using a *para*-substituted pyridine nitrile amino acid.<sup>40</sup> This was incorporated by the promiscuous aminoacyl-tRNA synthetase *p*-CNF-RS (*p*-cyanophenylalanine specific aminoacyl-tRNA synthetase), which relies on stop codon suppression and is thus less flexible in reprogramming. This nonsense suppression approach risks peptide truncations, and so necessitates an additional purification step using a C-terminal affinity tag that increases the complexity of the workflow. That approach also relied on the enzymes peptide deformylase and methionine aminopeptidase to expose the N-terminal cysteine. These enzymes have their own substrate biases<sup>41,42</sup> and so are less general than the initiation codon bypass approach we use here. Importantly, the enzyme used in that work, *p*-CNF-RS, was unable to acylate the *meta*-substituted isomer, *mCNP*, described in this work. We again consider both approaches complementary as the *meta*- and *para*-substitution patterns give access to useful differences in peptide conformation, as illustrated by our MD simulations here.

## Conclusions

In this work we have adapted the 'thiazoline click' reaction of N-terminal cysteine with *meta*-cyanopyridylalanine for application to *in vitro* translation and subsequently in genetically recoded mRNA-display based peptide selection. We show that this new approach is overall complementary to several existing macrocyclisation options and thereby increases the size of the macrocyclisation toolbox, potentially offering advantages in increased rigidity of the macrocyclisation cross-link, with positive influence on cell permeability; free choice in the genetic recoding for more flexible adjustment of the peptide chemical space; and compatibility with additional cysteine residues, including in spontaneous bicycle formation.

## Data availability

Raw sequencing data have been deposited in the DataverseNL database with <https://doi.org/10.34894/GAZQX6>.

## Author contributions

ML, DP, CN and SJ designed the research. DP, CH, CN, and SJ supervised the research. ML, RM, RY, MP, AH and ET carried out experimental work. ML, RM, RY, DP, CH, CN and SJ wrote the manuscript.

## Conflicts of interest

There are no conflicts to declare.





## Acknowledgements

CN thanks the Australian Research Council for a Discovery Project (DP230100079) and a Future Fellowship (FT220100010). SJ and RY acknowledge support from the Dutch Research Council grant number OCENW.KLEIN.248, and RY acknowledges further support from the Takenaka Scholarship Foundation.

## References

- 1 C. Morrison, Constrained peptides' time to shine?, *Nat. Rev. Drug Discovery*, 2018, **17**, 531–533.
- 2 D. R. Cary, M. Ohuchi, P. C. Reid and K. Masuya, Constrained peptides in drug discovery and development, *J. Synth. Org. Chem., Jpn.*, 2017, **75**, 1171–1178.
- 3 L. Wang, N. Wang, W. Zhang, X. Cheng, Z. Yan, G. Shao, X. Wang, R. Wang and C. Fu, Therapeutic peptides: current applications and future directions, *Signal Transduction Targeted Ther.*, 2022, **7**, 48.
- 4 A. Zorzi, K. Deyle and C. Heinis, Cyclic peptide therapeutics: past, present and future, *Curr. Opin. Chem. Biol.*, 2017, **38**, 24–29.
- 5 A. A. Vinogradov, Y. Yin and H. Suga, Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges, *J. Am. Chem. Soc.*, 2019, **141**, 4167–4181.
- 6 S. Chen, J. Morales-Sanfrutos, A. Angelini, B. Cutting and C. Heinis, Structurally diverse cyclisation linkers impose different backbone conformations in bicyclic peptides, *ChemBioChem*, 2012, **13**, 1032–1038.
- 7 S. S. Kale, C. Villequey, X.-D. D. Kong, A. Zorzi, K. Deyle and C. Heinis, Cyclization of peptides with two chemical bridges affords large scaffold diversities, *Nat. Chem.*, 2018, **10**, 715–723.
- 8 X. Li, T. W. Craven and P. M. Levine, Cyclic Peptide Screening Methods for Preclinical Drug Discovery, *J. Med. Chem.*, 2022, **65**, 11913–11926.
- 9 C. Alleyne, R. P. Amin, B. Bhatt, E. Bianchi, J. C. Blain, N. Boyer, D. Branca, M. W. Embrey, S. N. Ha, K. Jette, D. G. Johns, A. D. Kerekes, K. A. Koeplinger, D. Laplaca, N. Li, B. Murphy, P. Orth, A. Ricardo, S. Salowe, K. Seyb, A. Shahripour, J. R. Stringer, Y. Sun, R. Tracy, C. Wu, Y. Xiong, H. Youm, H. J. Zokian and T. J. Tucker, Series of Novel and Highly Potent Cyclic Peptide PCSK9 Inhibitors Derived from an mRNA Display Screen and Optimized *via* Structure-Based Design, *J. Med. Chem.*, 2020, **63**, 13796–13824.
- 10 T. J. Tucker, M. W. Embrey, C. Alleyne, R. P. Amin, A. Bass, B. Bhatt, E. Bianchi, D. Branca, T. Bueters, N. Buist, S. N. Ha, M. Hafey, H. He, J. Higgins, D. G. Johns, A. D. Kerekes, K. A. Koeplinger, J. T. Kuethe, N. Li, B. Murphy, P. Orth, S. Salowe, A. Shahripour, R. Tracy, W. Wang, C. Wu, Y. Xiong, H. J. Zokian, H. B. Wood and A. Walji, A Series of Novel, Highly Potent, and Orally Bioavailable Next-Generation Tricyclic Peptide PCSK9 Inhibitors, *J. Med. Chem.*, 2021, **64**, 16770–16800.
- 11 Y. Huang, M. M. Wiedmann and H. Suga, RNA Display Methods for the Discovery of Bioactive Macrocycles, *Chem. Rev.*, 2018, **119**, 10360–10391.
- 12 Y. Goto, T. Katoh and H. Suga, Flexizymes for genetic code reprogramming, *Nat. Protoc.*, 2011, **6**, 779–790.
- 13 Y. Goto and H. Suga, The RaPID Platform for the Discovery of Pseudo-Natural Macrocyclic Peptides, *Acc. Chem. Res.*, 2021, **54**, 3604–3617.
- 14 S. L. Richardson, K. K. Dods, N. A. Abrigo, E. S. Iqbal and M. C. Hartman, In vitro genetic code reprogramming and expansion to study protein function and discover macrocyclic peptide ligands, *Curr. Opin. Chem. Biol.*, 2018, **46**, 172–179.
- 15 K. Iwasaki, Y. Goto, T. Katoh and H. Suga, Selective thioether macrocyclization of peptides having the N-terminal 2-chloroacetyl group and competing two or three cysteine residues in translation, *Org. Biomol. Chem.*, 2012, **10**, 5783–5786.
- 16 S. A. K. Jongkees, S. Umemoto and H. Suga, Linker-free incorporation of carbohydrates into *in vitro* displayed macrocyclic peptides, *Chem. Sci.*, 2017, **8**, 1474–1481.
- 17 Y. Goto, Y. Ito, Y. Kato, S. Tsunoda and H. Suga, One-Pot Synthesis of Azoline-Containing Peptides in a Cell-free Translation System Integrated with a Posttranslational Cyclodehydratase, *Chem. Biol.*, 2014, **21**, 766–774.
- 18 C. B. Rosen and M. B. Francis, Targeting the N terminus for site-selective protein modification, *Nat. Chem. Biol.*, 2017, **13**, 697–705.
- 19 N. Asimwe, M. F. Al Mazid, D. P. Murale, Y. K. Kim and J.-S. Lee, Recent advances in protein modifications techniques for the targeting N-terminal cysteine, *Pept. Sci.*, 2021, **114**, e24235.
- 20 M. Liu, V. Thijssen and S. A. K. Jongkees, Suppression of formylation provides an alternative approach to vacant codon creation in bacterial *in vitro* translation, *Angew Chem. Int. Ed. Engl.*, 2020, **59**, 21870–21874.
- 21 C. Nitsche, H. Onagi, J.-P. Quek, G. Otting, D. Luo and T. Huber, Biocompatible Macrocyclization between Cysteine and 2-Cyanopyridine Generates Stable Peptide Inhibitors, *Org. Lett.*, 2019, **21**, 4709–4712.
- 22 M. Liu, R. Yoshisada, A. Amedi, A. J. P. P. Hopstaken, M. N. Pascha, C. A. M. de Haan, D. P. Geerke, D. A. Poole and S. A. K. Jongkees, An Efficient, Site-Selective and Spontaneous Peptide Macrocyclisation During *in vitro* Translation, *Chem.-Eur. J.*, 2023, **29**, e202203923.
- 23 J. T. Hampton, T. J. Lalonde, J. M. Tharp, Y. Kurra, Y. R. Alugubelli, C. M. Roundy, G. L. Hamer, S. Xu and W. R. Liu, Novel Regioselective Approach to Cyclize Phage-Displayed Peptides in Combination with Epitope-Directed Selection to Identify a Potent Neutralizing Macrocyclic Peptide for SARS-CoV-2, *ACS Chem. Biol.*, 2022, **17**, 2911–2922.
- 24 M. Zheng, F. Haefner and J. Gao, N-Terminal cysteine mediated backbone-side chain cyclization for chemically enhanced phage display, *Chem. Sci.*, 2022, **13**, 8349–8354.





- 25 T. Fujino, T. Kondo, H. Suga and H. Murakami, Exploring of minimal RNA substrate of flexizymes, *ChemBioChem*, 2019, **20**, 1959–1965.
- 26 T. Katoh, Y. Iwane and H. Suga, tRNA engineering for manipulating genetic code, *RNA Biol.*, 2018, **15**, 453–460.
- 27 Z. Cui, Y. Wu, S. Mureev and K. Alexandrov, Oligonucleotide-mediated tRNA sequestration enables one-pot sense codon reassignment *in vitro*, *Nucleic Acids Res.*, 2018, **46**, 6387–6400.
- 28 Y. Sako, Y. Goto, H. Murakami and H. Suga, Ribosomal synthesis of peptidase-resistant peptides closed by a nonreducible inter-side-chain bond, *ACS Chem. Biol.*, 2008, **3**, 241–249.
- 29 M. Wendt, R. Bellavita, A. Gerber, N. L. Efrém, T. van Ramshorst, N. M. Pearce, P. R. J. Davey, I. Everard, M. Vazquez-Chantada, E. Chiarparin, P. Grieco, S. Hennig and T. N. Grossmann, Bicyclic  $\beta$ -Sheet Mimetics that Target the Transcriptional Coactivator  $\beta$ -Catenin and Inhibit Wnt Signaling, *Angew. Chem., Int. Ed.*, 2021, **60**, 13937–13944.
- 30 E. H. Abdelkader, H. Qianzhu, J. George, R. L. Frkic, C. J. Jackson, C. Nitsche, G. Otting and T. Huber, Genetic Encoding of Cyanopyridylalanine for In-Cell Protein Macrocyclization by the Nitrile–Aminothiol Click Reaction, *Angew. Chem., Int. Ed.*, 2022, **61**, e202114154.
- 31 M. N. Pascha, V. Thijssen, J. E. Egido, M. W. Linthorst, J. H. Van Lanen, D. A. A. Van Dongen, A. J. P. P. Hopstaken, F. J. M. Van Kuppeveld, J. Snijder, C. A. M. De Haan and S. A. K. Jongkees, Inhibition of H1 and H5 Influenza A Virus Entry by Diverse Macrocyclic Peptides Targeting the Hemagglutinin Stem Region, *ACS Chem. Biol.*, 2022, **17**, 2425–2436.
- 32 T. Ishizawa, T. Kawakami, P. C. Reid and H. Murakami, TRAP display: A high-speed selection method for the generation of functional polypeptides, *J. Am. Chem. Soc.*, 2013, **135**, 5433–5440.
- 33 W. E. Evenson, W. Z. S. Lin, K. Pang, A. T. Czaja, F. Jalali-Yazdi, T. T. Takahashi, N. Malmstadt and R. W. Roberts, Enabling Flow-Based Kinetic Off-Rate Selections Using a Microfluidic Enrichment Device, *Anal. Chem.*, 2020, **92**, 10218–10222.
- 34 V. Thijssen, D. L. Hurdiss, O. J. Debski-Antoniak, M. A. Spence, C. Franck, A. Norman, A. Aggarwal, N. J. Mokiem, D. A. A. van Dongen, S. W. Vermeir, M. Liu, W. Li, M. Chatziandreou, T. Donselaar, W. Du, I. Drulyte, B. J. Bosch, J. Snijder, S. G. Turville, R. J. Payne, C. J. Jackson, F. J. M. van Kuppeveld and S. A. K. Jongkees, A broad-spectrum macrocyclic peptide inhibitor of the SARS-CoV-2 spike protein, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, e2303292120.
- 35 D. Angeletti and J. W. Yewdell, Is it possible to develop a “universal” influenza virus vaccine?: Outflanking antibody immunodominance on the road to universal influenza vaccination, *Cold Spring Harbor Perspect. Biol.*, 2018, **10**, a028852.
- 36 R. U. Kadam, J. Juraszek, B. Brandenburg, C. Buyck, W. B. G. Schepens, B. Kesteley, B. Stoops, R. J. Vreeken, J. Vermond, W. Goutier, C. Tang, R. Vogels, R. H. E. Friesen, J. Goudsmit, M. J. P. Van Dongen and I. A. Wilson, Potent peptidic fusion inhibitors of influenza virus, *Science*, 2017, **358**, 496–502.
- 37 J. Yang and Y. Zhang, I-TASSER server: New development for protein structure and function predictions, *Nucleic Acids Res.*, 2015, **43**, W174–W181.
- 38 R. Okuma, T. Kuwahara, T. Yoshikane, M. Watanabe, P. Dranchak, J. Inglese, S. Shuto, Y. Goto and H. Suga, A Macrocyclic Peptide Library with a Structurally Constrained Cyclopropane-containing Building Block Leads to Thiol-independent Inhibitors of Phosphoglycerate Mutase, *Chem.-Asian J.*, 2020, **15**, 2631–2636.
- 39 T. Tamura, M. Inoue, Y. Yoshimitsu, I. Hashimoto, N. Ohashi, K. Tsumura, K. Suzuki, T. Watanabe and T. Hohsaka, Chemical Synthesis and Cell-Free Expression of Thiazoline Ring-Bridged Cyclic Peptides and Their Properties on Biomembrane Permeability, *Bull. Chem. Soc. Jpn.*, 2022, **95**, 359–366.
- 40 S. E. Iskandar, J. M. Pelton, E. T. Wick, D. L. Bolhuis, A. S. Baldwin, M. J. Emanuele, N. G. Brown and A. A. Bowers, Enabling Genetic Code Expansion and Peptide Macrocylation in mRNA Display via a Promiscuous Orthogonal Aminoacyl-tRNA Synthetase, *J. Am. Chem. Soc.*, 2022, **145**, 1512–1517.
- 41 Q. Xiao, F. Zhang, B. A. Nacev, J. O. Liu and D. Pei, Protein N-terminal processing: Substrate specificity of *Escherichia coli* and human methionine aminopeptidases, *Biochemistry*, 2010, **49**, 5588–5599.
- 42 Y. J. Hu, Y. Wei, Y. Zhou, P. T. R. Rajagopalan and D. Pei, Determination of substrate specificity for peptide deformylase through the screening of a combinatorial peptide library, *Biochemistry*, 1999, **38**, 643–650.

