Chemistry–A European Journal
Research Article
doi.org/10.1002/chem.202203923
Chemistry Europe
European Chemical Societies Publishing
www.chemeurj.org

An Efficient, Site-Selective and Spontaneous Peptide Macrocyclisation During in vitro Translation

Minglong Liu,[a] Ryoji Yoshisada,[a] Avand Amedi,[b] Antonius J. P. Hopstaken,[a] Mirte N. Pascha,[c] Cornelis A. M. de Haan,[c] Daan P. Geerke,[a] David A. Poole, III,[a] and Seino A. K. Jongkees*[a, b]

Abstract: Macrocyclisation provides a means of stabilising the conformation of peptides, often resulting in improved stability, selectivity, affinity, and cell permeability. In this work, a new approach to peptide macrocyclisation is reported, using a cyanobenzothiazole-containing amino acid that can be incorporated into peptides by both in vitro translation and solid phase peptide synthesis, meaning it should be applicable to peptide discovery by mRNA display. This cyclisation proceeds rapidly, with minimal by-products, is selective over other amino acids including non *N*-terminal cysteines, and is compatible with further peptide elaboration exploiting such an additional cysteine in bicyclisation and derivatisation reactions. Molecular dynamics simulations show that the new cyclisation group is likely to influence the peptide conformation as compared to previous thioether-based approaches, through rigidity and intramolecular aromatic interactions, illustrating their complementarity.

Introduction

Macrocyclic peptides are a promising class of compound for drug development, with their large interaction surface able to bind in a similar mode to antibodies while the conformational constraints both reduce the entropic cost of binding and improve protease stability.[1] Peptide display technologies, such as phage or mRNA display, offer a means to discover new macrocyclic peptides,[2] but these need macrocyclisation chemistry that is compatible with both the display platform and subsequent scale up by solid-phase peptide synthesis (SPPS).[3] Current cyclisations offer at best limited regioselectivity and are predominantly cysteine-based, meaning that peptides with multiple cysteines either give a mix of isomers or give the smallest macrocycle (further discussed below). Cysteine is uniquely convenient in its reactivity among the canonical amino acids, and so offers a convenient handle for further modification of (mRNA-displayed) peptides.[4–7] For this reason, we sought to develop a new approach for macrocyclisation that retains the convenience of spontaneous and efficient ring formation, but offers improved control. Our ideal reaction would give no DNA/RNA damage, be water compatible, spontaneous, efficient, reasonably fast (e.g. complete within 30–60 min), form a stable product, and give controlled reaction with multiple cysteines to form the largest macrocycle, thereby leaving further cysteine(s) in the macrocycle available for other reactions to further diversify the peptides.

Existing approaches for macrocyclisation in mRNA display (Figure 1) include: disulfide linkages[8] (which are not selective, and are prone to reductive opening inside cells), thioether formation[9] (which is intolerant of additional cysteine residues inside the macrocycle), 'CLIPS' reactions with benzylic halides[10] (which do not have selectivity among cysteine residues), 'click' copper-catalysed azide-alkyne cycloaddition reactions[11] (which are not spontaneous, varied in efficiency, can cause oxidation,[12] and not regioselective among alkynes/azides), and amine cross-linking by disuccinimidyl glutarate[13] (which restricts lysine placement and is difficult to control). None of these meet the criteria defined above as an ideal means for cyclisation.

In contrast, *N*-terminal selective reagents are a promising approach for controlled reactivity in peptides.[14] Among these, cyanobenzothiazoles (abbreviated here as CBTz rather than CBT to avoid confusion with 4-chlorobenzyl thioesters used in genetic code reprogramming[15]) have been reported to give fast and efficient modification of specifically the *N*-terminal

[a] *Dr. M. Liu, R. Yoshisada, A. J. P. Hopstaken, Dr. D. P. Geerke, Dr. D. A. Poole, III, Dr. S. A. K. Jongkees*
Chemistry and Pharmaceutical Sciences and Amsterdam Institute of Molecular and Life Sciences (AIMMS)
Vrije Universiteit Amsterdam
Amsterdam 1081 HV (the Netherlands)
E-mail: s.a.k.jongkees@vu.nl
Homepage: https://sakjongkees.wixsite.com/website

[b] *A. Amedi, Dr. S. A. K. Jongkees*
Department Chemical Biology and Drug Discovery and Utrecht Institute for Pharmaceutical Sciences (UIPS)
Utrecht University
Utrecht 3584 CG (the Netherlands)

[c] *M. N. Pascha, Dr. C. A. M. de Haan*
Section Virology
Division of Infectious Diseases and Immunology
Department of Biomolecular Health Sciences
Faculty of Veterinary Medicine
Utrecht University
Utrecht 3584 CL (the Netherlands)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202203923

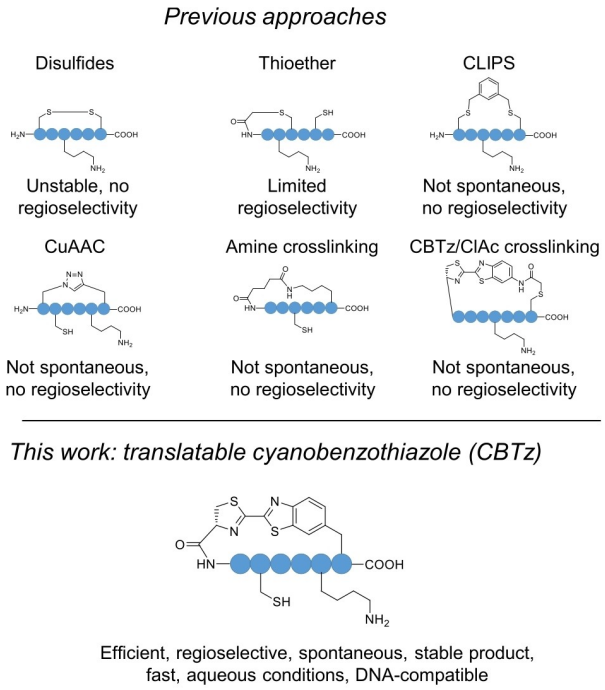


Figure 1. Examples of previous approaches for macrocyclisation of displayed peptide libraries (upper), along with limitations of each, compared with the approach developed in the current work (lower). Blue circles represent amino acids, explicit sidechains indicate selectivity over lysine and cysteine residues. CBTz, cyanobenzothiazole; ClAc, chloroacetamide.

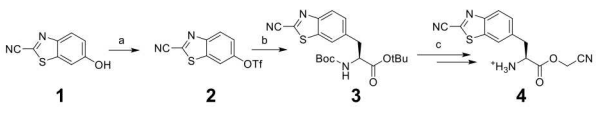
cysteine in proteins[16] and peptides,[17] are DNA-compatible,[18] and the resulting product is compatible with a biological setting.[19] The aromatic nature and rigidity of the CBTz group may also have a substantially different effect on peptide dynamics than a highly flexible thioether group, as has been seen with cysteine cross-linking agents where different linkers can give access to different conformations and result in different bioactivity.[20] As further evidence of their utility, a report was published during the preparation of this manuscript demonstrating the cross-linking of two cysteines in phage display, one *N*-terminal, using a bifunctional CBTz-based reagent.[21] However, this crosslinking approach is not spontaneous and not selective among downstream cysteines, and so precludes the further cysteine-based peptide elaboration that is one of our goals here. A cyanopyridylalanine-based system for cyclisation inside cells has also recently been reported,[22] but has not been demonstrated in the context of in vitro translation of peptide libraries, where the aminoacyl-tRNA synthetase would need to be overexpressed and purified. In this work, we investigate the use of a CBTz-containing amino acid in peptide macrocyclisation during reprogrammed in vitro translation to facilitate alternate cyclisation patterns, different conformations, and further modification on cysteine residues, all with the long term goal of building more advanced libraries in mRNA display.

Results and Discussion

Initially we considered that amide bond formation between a sidechain carboxylate in an appropriately protected aspartate and 6-amino-2-cyanobenzothiazole would provide facile access to a CBTz-containing amino acid. The amine in this latter compound is poorly nucleophilic and previously reported methods[23,24] using isobutyl chloroformate activation proved difficult to implement in our hands, but activation with $POCl_3$ in pyridine[25] was effective and reliable in furnishing the product (Scheme S1). Incorporation of this amino acid into a test peptide derived from SPPS (Scheme S2) showed cyclisation driven by the rate-limiting TCEP deprotection of an *N*-terminal Cys-S*t*Bu (Figure S1). However, testing the 'charging' of a model tRNA-like microhelix with this amino acid by the aminoacylating ribozyme 'dFx' type flexizyme[15] showed no conversion (Figure S2). Thus, while the cyclisation reaction itself appears promising, its incorporation into peptides in this form during in vitro translation is a barrier to implementation of this reaction and so a synthetically more complex amino acid would be needed.

To improve aminoacylation yield, we sought an amino acid that could be charged by the more reactive 'eFx' type flexizyme.[26] This typically charges cyanomethyl esters of aromatic amino acids, and so we considered making a *C*-linked amino acid derivative of CBTz (Scheme 1). Synthesis of this proceed from the cheaper starting material 6-hydroxy-2-cyanobenzothiazole (**1**) through its triflate derivative **2**, which was subsequently used in a Negishi coupling with Boc/O*t*Bu protected iodoalanine to afford the protected amino acid **3**. Deprotection of both the amine and acid by trifluoroacetic acid (TFA), reprotection of the amine with Boc anhydride, and treatment with triethylamine in chloroacetonitrile afforded the activated ester, which was immediately deprotected by TFA to afford 6 mg of activated amino acid **4** in 6% overall isolated yield. While this was a small quantity, there was sufficient material for approximately 800 typical acylation reactions, including all experiments performed in this text, and so the synthesis was not further optimised. Pleasingly, charging of tRNA with this amino acid proceeded smoothly to give 44% acylated yield after one hour of reaction (Figure S2), while longer reaction times gave decreasing yields, presumably due to ester hydrolysis.

Subsequent testing of translation showed incorporation of this *C*-linked CBTz amino acid into a test peptide **P1** by the engineered[27] $tRNA^{EnGluE2}_{GCU}$ at an AGC serine codon (liberated by omission of the canonical amino acid). Translation was



Scheme 1. Synthesis of a C-linked cyanobenzothiazole amino acid. a) Tf_2O, pyridine, 94%; b) 1. Zinc dust, iodine, Boc-Ala(iodo)-O*t*Bu; 2. $Pd_2(PPh_3)_2Cl_2$, Sphos., **1**, 63% over 2 steps; c) 1. TFA; 2. Boc_2O, $NaHCO_3$; 3. chloroacetonitrile, Et_3N; 4. TFA, 10% over 4 steps.

carried out with suppression of the initiation codon[28] by excess $tRNA^{fMet}_{CAU}$ to give translation starting from the second codon (cysteine). This directly afforded the cyclised product, but with several notable side-product peaks (Figure 2). Suppression of the serine codon appeared to be incomplete, with byproducts featuring either residual serine incorporation or arginine mis-incorporation being present. A substantial peak from CBTz hydration was also detected, likely at the nitrile and in competition with cyclisation. Noteworthy is that no product was observed from cysteine or dithiothreitol adducts, despite both being present in the translation reaction mixture. This selectivity may reflect a substrate-discriminating effect of the ribosome, with cysteine-modified CBTz not being efficiently translated, or a kinetic effect, where the free cysteine only reacting slowly while the translated product rapidly cyclises due to high local concentration of the *N*-terminus.

The presence of a peptide with serine incorporation despite the omission of serine from the translation reaction mixture was presumed to derive from residual serine, either already attached to tRNA or carried over in the active site of SerRS. This could in principle be mitigated by SerRS omission, but we see no way to carry this out in the context of a commercial translation kit. Arginine incorporation may be arising through either wobble base pairing of an arginine tRNA during translation or background mis-acylation of serine tRNA in the absence of the correct substrate. To prevent serine incorporation, translation was attempted with the AGC codon instead liberated by addition of an antisense oligonucleotide that binds to a portion of $tRNA_{AGC}$ to suppress its use in translation through unfolding.[29] This approach also gives expansion, rather than replacement, in the genetic code by allowing retention of serine at other codons that use a different tRNA. This approach indeed removed the serine by-product, and pleasingly also fortuitously removed the arginine side-product (supporting the hypothesis of mis-acylation of $tRNA_{AGC}$ by ArgRS).

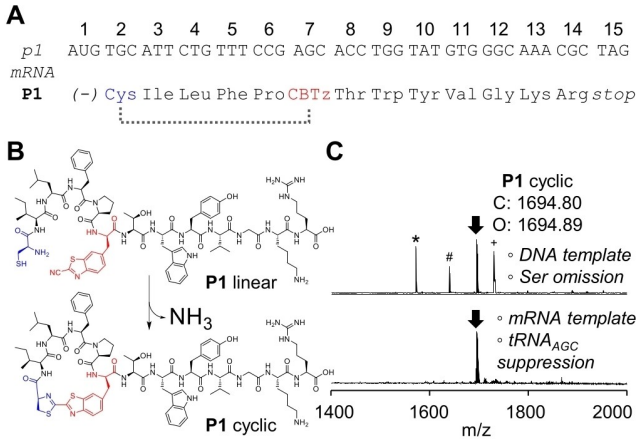


Figure 2. Translation and cyclisation testing for C-linked CBTz amino acid **3**. A.) mRNA and peptide sequences, showing the codons reprogrammed, and with the intended resulting cyclisation as a dotted line. Cyclising amino acids are coloured for emphasis. B.) Structure of the initial translated peptide and cyclised final product, coloured as in A. C.) MALDI-TOF spectra of the first test conditions (upper) and optimised conditions (lower). An arrow indicates the expected mass in each case with (C)alculated and (O)bserved masses as indicated; (–), vacant codon; *, serine incorporation at AGC; #, arginine incorporation at AGC; +, hydration rather than cyclisation at CBTz.

The hydration side-reaction was considered to be possible during both acylation and translation, but LC-MS testing of amino acid stability under acylation conditions showed only activated ester hydrolysis (data not shown). Given that the acylation and translation are carried out at the same pH, we thus concluded that this side-reaction is likely to be catalysed by some component of the translation reaction mixture. Testing of translation in a reaction mixture either omitting DTT or with a lower concentration of free cysteine did not resolve this issue (Figure S3). As an alternate solution, we instead sought to minimise the time that the CBTz amino acid spent in the translation solution before incorporation into a stable macrocyclised peptide product. To achieve this we used mRNA rather than DNA as template, to bypass any delay from initiation of the transcription step, and this largely removed the problem (still with trace hydration detectable in some cases). With these changes, reprogrammed translation proceeded efficiently and with only a minor hydration peak remaining, and so we considered this approach ready for implementation.

With an efficiently translating and cyclising CBTz amino acid ready, we explored the scope and selectivity of the reaction. The initial test translation template used for optimisation gave a peptide macrocycle with 4 additional amino acids spanning the cysteine and CBTz, and contains potential alternate nucleophiles in lysine, threonine, and tyrosine (which give no detectable side-reactions, expected as an ammonium adduct). A larger macrocycle **P2**, with 8 additional amino acids also gave efficient translation and cyclisation (Figure 3A), while a control peptide **P3** with an *N*-terminal alanine gave only the hydrated product and no cyclisation (Figure 3B).

To increase the flexibility in application of this new cyclisation amino acid, we showed that it can also be used in a sidechain-to-sidechain cyclisation (Figure 3C) by incorporation into peptide **P4** together with a second non-canonical amino acid *N*-ε-cysteinyl lysine. This was synthesised as a dinitrobenzyl-activated ester (Scheme S3) and charged onto $tRNA^{EnGluE2}_{CAU}$ by dFx for AUG codon reprogramming by methionine omission, together with initiation reprogramming by *N*-acetyl phenylalanine on $tRNA^{fMet}_{CAU}$ as a convenient initiating amino acid. While this showed an unusual second band at much higher apparent weight during aminoacylation (Figure S2), MALDI analysis after in vitro translation showed only the correct product. Again, in peptide **P4** two cysteines were present, and this gave a product of a single mass showing loss of ammonia and so indicative of reaction with an *N*-terminal cysteine (in this case on the lysine sidechain) and not the internal cysteine.

We next explored the potential of our approach for orthogonal cysteine-based bicyclisation. This was assessed by translation of peptide **P5** with an *N*-terminal cysteine, a *C*-terminal CBTz, a further internal cysteine, and an additional cysteine-reactive group in a *N*-γ-chloroacetylated diaminopropionic acid (Figure 3D). With the CBTz reaction selective for the *N*-terminus and the reaction of an *N*-terminal chloroacetamide having previously been shown to be selective for the closest cysteine,[9] as well as the large difference in reaction rates

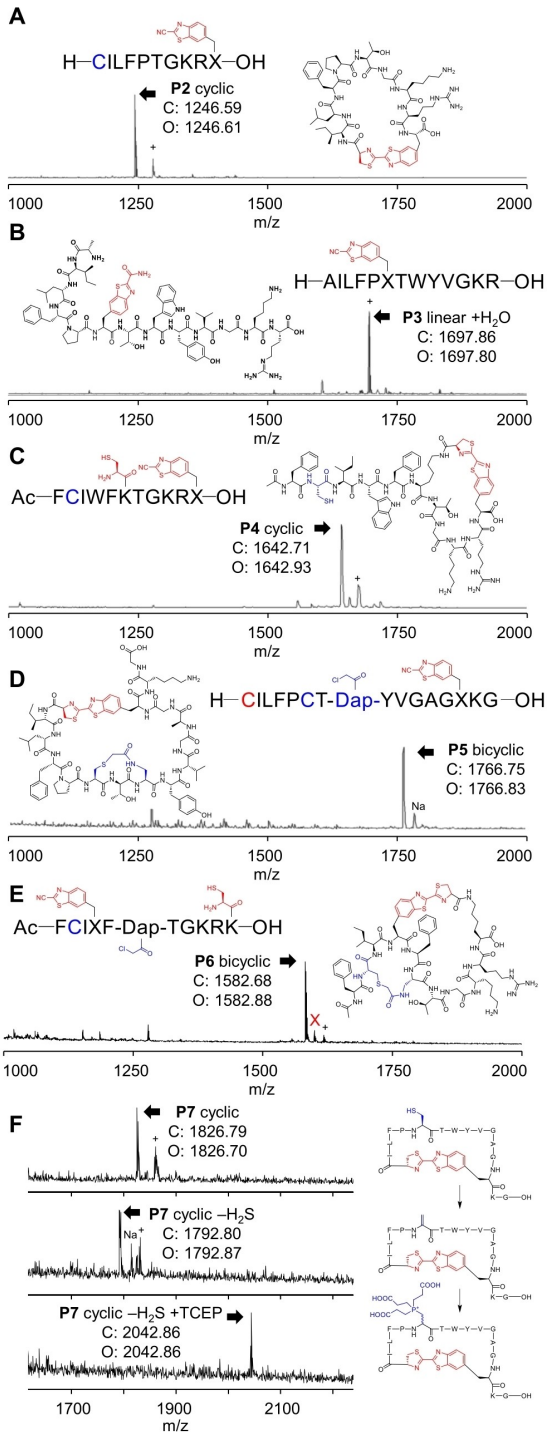


Figure 3. Testing the scope and selectivity of CBTz cyclisation. A.) Translation of a larger macrocycle. B.) Control reaction with no C-terminal cysteine. C.) Sidechain to sidechain cyclisation using a cysteine-modified lysine. D.) Head-to-sidechain and sidechain-to-sidechain bicyclisation with orthogonal cysteines. E.) Bicyclisation with two orthogonal sidechain-to-sidechain cysteine reactions. F.) modification of a secondary cysteine within the macrocycle by elimination and conjugate addition. An arrow indicates the expected mass in each case with (C)alculated and (O)bserved masses as indicated; sequence representations show the initial translated peptides and expanded structures the final products; +, hydration rather than cyclisation at CBTz; Na, sodium adduct; X, ammonium adduct indicating incorrect cyclisation.

(0.02[30] vs. ~10[31] $M^{-1}s^{-1}$ at neutral pH for an intermolecular reaction), we considered it plausible that each would proceed with sufficient selectivity to give a single defined product. Importantly, each possible side reaction is expected to afford products distinguished by their unique mass, based on loss of chloride and ammonium leaving groups, together with hydration of the CBTz if unable to cyclise. Pleasingly, only a single mass peak was observed, which indicates a cleanly bicyclised product. A further test of this concept was applied in peptide **P6**, where we combine two sidechain-to-sidechain cyclisations (chloroacetamide with thiol and CBTz with *N*-terminal Cys). In this test, where the chloroacetamide is roughly equally spaced between the two cysteines and the CBTz cyclisation needs to form the larger macrocycle, we see the first sign of incorrect cyclisation, observed as a small peak at a mass corresponding to an ammonium adduct of the expected product mass (Figure 3E). We interpret this as arising from chloroacetamide cyclisation as the undesired regioisomer followed by reversible addition of the only remaining thiol into the nitrile, which then is unable to undergo the irreversible second step that occurs with an *N*-terminal cysteine. That this does not hydrate suggests it remains a stable thiol adduct and may reflect the effect of high local concentration from an intramolecular reaction on the equilibrium constant. However, while this product is detectable, it is a small peak that arises in a'worst case' geometry and so should not be a hurdle to implementation of this method.

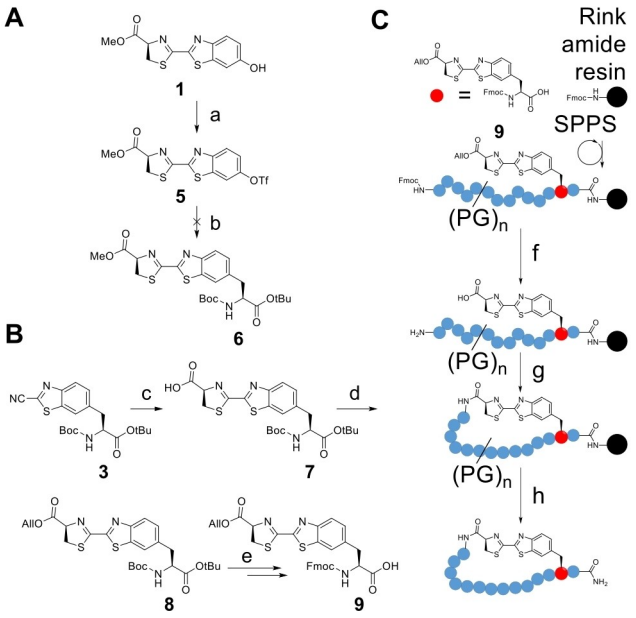
In a final test of compatibility we also investigated the use of a cysteine inside the macrocycle for peptide diversification by intermolecular reaction. Peptide **P7**, with 12 additional amino acids including a second cysteine in the macrocycle, again showed translation and cyclisation to afford a single major product. This additional cysteine was then shown to be able to be converted to dehydroalanine and subsequently undergo conjugate addition by a phosphine nucleophile using our previously reported approach[32] (Figure 3F). While signals were noisier for **P7** than other peptides in this work, peaks clearly match expected masses and mass changes across several replicates.

Together, the translated peptides in Figure 3 demonstrate that this new C-linked CBTz amino acid is a flexible and convenient cyclisation reagent for peptides generated by in vitro translation. It is selective over other amino acids including cysteines not at the *N*-terminus, robustly forms macrocycles of varied size, composition, and bicyclisation architecture, and is compatible with further peptide elaboration at additional cysteines.

To ensure that macrocyclic peptide hits discovered with this amino acid can also be produced on solid phase, we tested the stability of CBTz in the presence of piperidine (Figure S4). This was particularly of concern given the formation of hydrates seen during translation (Figures 2C, 3A–F). Substantial instability was seen in this test, indicating that repeated SPPS coupling cycles with an unprotected CBTz group would likely produce unacceptable amounts of side-products. Instead of attempting to protect at the nitrile, we opted to generate the cysteine adduct of CBTz and protect this at the cysteine carboxylate with an allyl ester. Initial tests using the same starting material **1** to

prepare triflate **5** followed by the same Negishi reaction to the amino acid **6** gave only traces of product detectable by TLC and we were unable to achieve the desired amino acid (Scheme 2A). Coupling cysteine after forming the amino acid **2** proved much more efficient (54%), giving the stable compound **5**. After allyl ester installation at the sidechain acid to form **6**, protecting group removal and Fmoc installation afforded compound **7** in 19% yield over 4 steps from **2** (Scheme 2B). This can be used in routine Fmoc SPPS, and subsequently deprotected by palladium followed by removal of the final Fmoc at the *N*-terminus to allow an on-resin macrocyclisation (represented in Scheme 2C).

Using this approach, we synthesised an analogue of an anti-influenza peptide that we recently reported[33] as a model peptide (named L4 in that work) at 12.5 μmol scale in an overall 0.8% yield ('L4 CBTz'; Figure S5). In that same earlier work, peptide L4 was seen to retain some activity in its linear form as compared to the original head-to-sidechain thioether macrocycle (denoted here as 'L4 thioether' for clarity), and so we considered it should be relatively tolerant of variations in cyclisation chemistry. Our CBTz-cyclised variant of L4 was synthesised without its *N*-terminal tyrosine to compensate for the larger CBTz group affording a macrocycle of approximately equal size (57-member ring for CBTz vs. 54-member ring for thioether). To our surprise, L4 CBTz proved to be completely inactive for inhibition of H1 N1 influenza infection at up to 2.5 μM, in contrast to L4 thioether which showed the expected activity down into the low nanomolar range (Figure 4B). Further confirmation of this unexpected result came from subsequent competitive fluorescence polarisation measurements using a fluorescent version of L4 thioether as probe, where L4 thioether competed for binding while L4 CBTz did not (Figure 4C). While



Scheme 2. Fmoc SPPS using a *C*-linked cyanobenzothiazole amino acid. a) Tf_2O, pyridine, 91%; b) Zinc dust, iodine, Boc-Ala(iodo)-O*t*Bu, trace product; c) Cysteine, $NaHCO_3$, 54%; d) allyl alcohol, EDC, 71%; e) 1. TFA; 2. Fmoc-OSu, Et_3N, 50% over 2 steps; f) $Pd(PPh_3)_4$ then 20% piperidine; g) DIC/Oxyma; h) TFA/H_2O/TIPS. PG = protecting groups.

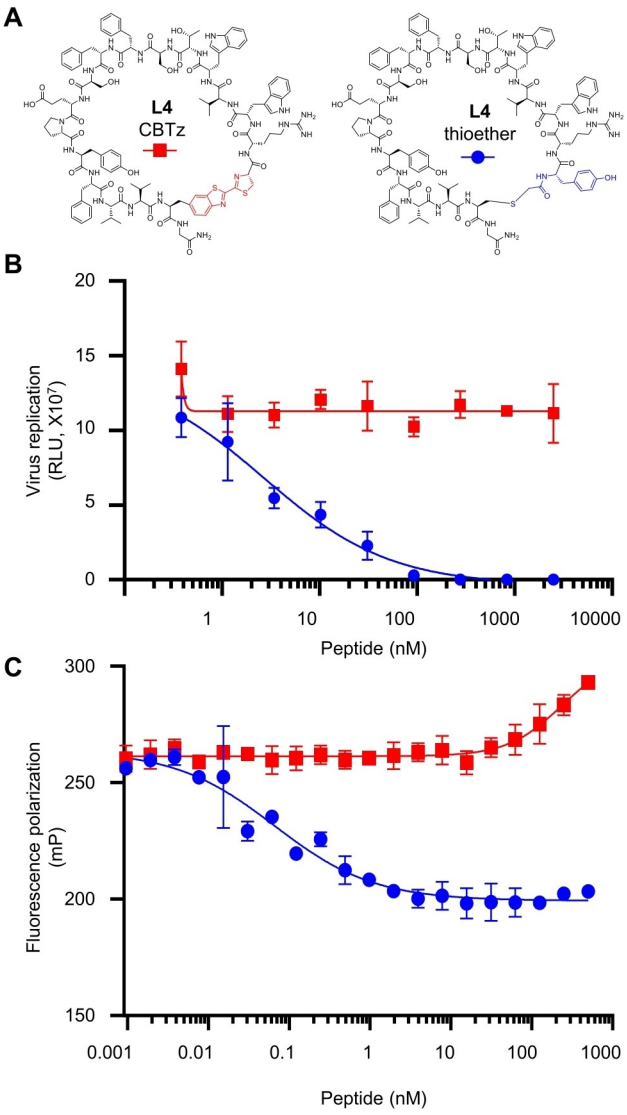


Figure 4. Solid phase synthesis and testing of a CBTz-cyclised peptide. A) Structures of the two peptides tested. B) Inhibitory activity of the two peptides in A for influenza H1 N1 infection in a luciferase reporter assay. C) Competition of the same peptides in A with a fluorescent version of L4 thioether for binding to haemagglutinin in a fluorescence polarisation assay.

initially disappointing, we now consider this a demonstration of the stark effect this change in cyclisation chemistry can have on an otherwise very similar pair of peptides.

This difference in biological activity indicates that the change in cyclisation results in the peptide either no longer being able to adopt the active conformation or that critical interacting residues are no longer available for binding. To further investigate this effect we carried out molecular dynamics (MD) simulations of both L4 thioether and L4 CBTz for 50 ns in implicit solvent at 300 K (Figure S6–8). In both productive simulations we observed a sub-population of more sheet-like structures, as well as another set of associated disordered or helix-containing structures. Circular dichroism (CD) spectra for both synthetic L4 variants showed a distinctive small positive peak at around 230 nm and were a good match to calculated

spectra for these sheet-like structures, and notably less good matches to either the helical or disordered structures (Figure 5 and S8), indicating that these sheets are likely to be representative of the dominant species in solution. These sheets align well with one another across the two cyclisation approaches in the sheet and turn regions, differing only around the cyclisation moiety. In one of the two sheet-like conformational clusters for L4 CBTz, the aromatic CBTz group is also placed similarly to the omitted tyrosine (Figure S9), validating our intended design. We also observe aromatic interactions involving the CBTz ring, particularly in edge-to-face interactions with tryptophan (Figure S10). Despite these broad similarities, the two different cyclisations induce a different angle at the *C*-terminal domain, arising from their differences in rigidity. The flexibility of each residue in the peptides was assessed through root mean square fluctuation (RMSF) analysis. Residues involved in beta-sheet formation showed lower values, indicating a more constrained nature for this region, while the turn region (Ser7-Phe9) and especially the *C*-terminal tails (Gly18) were found to fluctuate more. The beta-sheet in L4 Cbtz consists of less residues than that in L4 thioether, showing an 'unzipping' effect from the CBTz cyclisation moiety. We infer that these structural differences result in the observed reduction in activity. While specific investigation of how these two peptides interact with their target is the subject of ongoing study, it is outside the scope of this work. In generalising these observations we consider this increased rigidity and the potential for novel aromatic interactions a convincing demonstration of how CBTz cyclisation is able to provide access to a different pool of conformations when used in a library of randomised peptides, as compared to the thioether cyclisation approach, and thus highlighting their potential utility in generating macrocyclic peptide libraries.

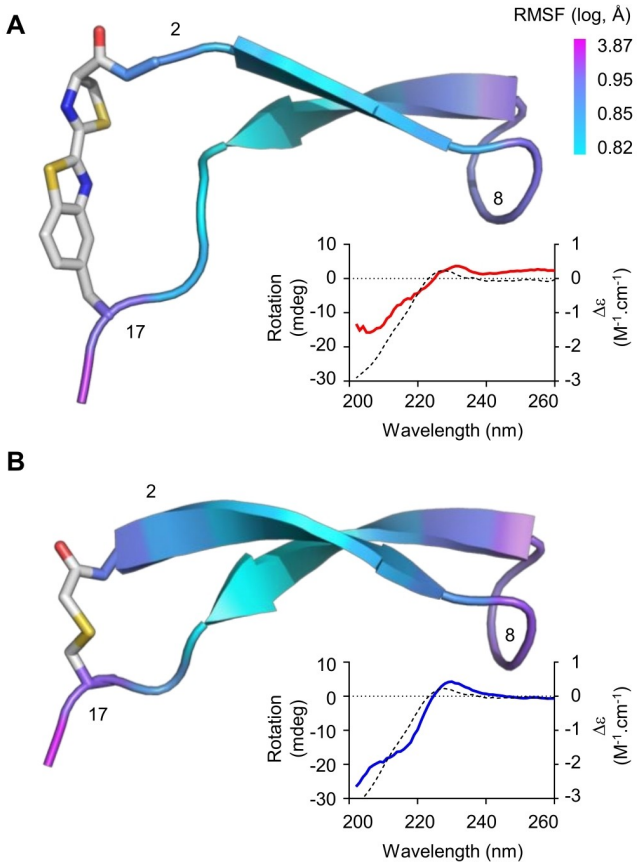


Figure 5. MD simulations of different cyclisation variants of peptide L4. A.) representative structure for L4 CBTz, shown as cartoon for the backbone (coloured by backbone RMSF with log colour scale as indicated) and stick representation for the cyclising moiety (atoms coloured carbon, grey; nitrogen, blue; oxygen, red; sulfur, yellow), with select residues numbered. Inset: measured CD spectrum for L4 CBTz in solid red (left axis) overlaid with calculated spectrum from the PDBMD2CD web server in dotted black (right axis). B.) representative structure for L4 thioether, shown as in A. Inset: measured CD spectrum for L4 thioether in solid blue (left axis) overlaid with calculated spectrum in dotted black (right axis).

Conclusion

In this work we have synthesised a new *C*-linked amino acid derivative of cyanobenzothiazole and demonstrated its efficient incorporation into peptides by reprogrammed in vitro translation. The resulting cyclisation proved to be fast, spontaneous, and selective, all while still being compatible with incorporation of another cysteine in the macrocycle and its futher exploitation in subsequent cyclisation and multi-step derivatisation reactions. We have shown that CBTz-cyclised macrocyclic peptides can also be accessed through SPPS for scale-up, and so we consider this approach suitable for use in peptide discovery by mRNA display. We anticipate that the hydrophobic and aromatic nature of the linkage and the absence of hydrogen bond donors may also aid in discovery of cell permeable peptides, while the aromatic ring system may also be able to engage in binding interactions with protein targets or conformation-stabilising intramolecular interactions. The cyanobenzothiazole cyclisation clearly influences the conformation of the peptide and its resulting bioactivity, suggesting that mRNA-displayed peptide libraries built on this cyclisation reaction may afford different hits compared to congeners from existing cyclisation reactions, even without exploiting the potential for orthogonal reactivity in library building.

Experimental Section

General information: Chemicals were purchased from Sigma Aldrich, Alfa Aesar or GLbiochem and used without further purification. HPLC-grade solvents were purchased from VWR chemicals. Unless otherwise stated, analytic LC-MS was performed on an Agilent 1260 II Infinity LC system using an Agilent poroshell-120 EC-C18 column (particle size: 2.7 µm, 100x4.6 mm) at a flow rate of 0.6 mL.min^{-1} using a linear gradient of buffer A (0.1% Formic acid in H_2O) and buffer B (0.1% Formic acid in CH_3CN) from 5–95% B over 10, 22 or 60 min, with detection by UV at 215 and 280 nm and by Agilent InfinityLab LC/MSD XT. Preparative RP-HPLC was performed on an automated preparative HPLC system equipped with a UV/VIS detector at 215/280 nm using a C18 column (particle size: 10 µm, 250x22 mm) at a flow rate of 12.5 mL.min^{-1} using a linear gradient of buffer A (0.1% TFA in CH_3CN/H2O 5:95 v/v) and buffer B (0.1% TFA in CH_3CN/H2O 95:5

v/v) from 10–70% B over 60 min. Peptide concentration was calculated through the UV absorbance at 280 nm on Nanodrop (thermo scientific). MALDI-TOF mass spectra for translated peptides were measured on Bruker ultrafleXtreme Maldi-TOF-MS and analyzed with flexAnalysis software, or for synthetic peptides on Axima CFR (Kratos analytical). U–C18 ZipTip pipette tips (ZTC18 M096) were purchased from Merck. Primers were ordered from Integrated DNA Technologies. PURexpress (E6840S) was purchased from new England biolab. ^{1}H and ^{13}C NMR spectra were recorded on Bruker 300 MHz or Bruker 500 MHz.

Preparation of flexizyme, tRNA, DNA and mRNA: DNA templates were assembled by primer extension followed by PCR as previously described.[15,34] Flexizymes (eFx and dFx), tRNAs ($tRNA^{fMet}_{CAU}$, $tRNA^{EnAsn}_{GCU}$, $tRNA^{EnGluE2}_{GCU}$, $tRNA^{EnAsn}_{CCA}$ and $tRNA^{EnAsn}_{CAU}$) and mRNAs were prepared by in vitro transcription with T7 RNA polymerase as previously described.[35] A list of primers used for each template is provided in Table S1, how these were assembled is indicated in Table S2, and resulting sequences in Table S3.

Aminoacylation testing: Aminoacylation reactions (5 μL) were carried out with slight modifications from standard conditions.[34b,36] A mixture of 3 μL containing 125 pmol of 5b-FAM or microhelix and 125 pmol eFx in 0.25 M HEPES-KOH buffer (pH 7.5) was heated at 95 °C for 2 minutes and subsequently allowed to cool to room temperature for 5 minutes. Then, 1 μL 3 M $MgCl_2$ solution was added, followed by an incubation at room temperature for another 5 minutes. The mixture was then cooled on ice, followed by the addition of 1 μL 25 mM amino acid solution in milliq water. The reaction was left to incubate on ice for a varied amount of time (as detailed in the figures below). The reaction was quenched with 2X Loading buffer (pH 5.2) and loaded onto 20% acid PAGE gel and run for 150 minutes at 120 V in sodium acetate (pH 5.2) buffer. Gels were subsequently imaged on Azure biosystems C400 (stained by sybr green II for microhelix) and yield were determined by densitometry in ImageJ.

General procedure for aminoacylation of tRNA: Aminoacylation reactions (10 μL) were carried with the following condition[15]: A mixture of 6 μL of 250 pmol of tRNA and 250 pmol eFx (or dFx) in 0.25 M HEPES-KOH buffer (pH 7.5) was heated at 95 °C for 2 min followed by an incubation at room temperature for 5 minutes. Then, 2 μL $MgCl_2$ 3 M solution was added and the mixture was left to incubate another 5 min at room temperature. The mixture was then cooled on ice, followed by the addition of 2 μL amino acid 25 mM solution in DMSO or water. The reaction was left to incubate on ice for 2 h then quenched with 40 μL NaOAc buffer (0.3 M, pH 5.2). This was followed by the addition of 100 μL ethanol to precipitate the product. This mixture was then centrifuged at 4 °C 15000 rcf for 15 min to form a pellet of the precipitated product. The supernatant was removed and the pellet was vortexed with 40 μL NaOAc buffer (0.1 M, pH 5.2) in 70% EtOH then centrifuged for 10 minutes at 4 °C 15000 rcf, and this process was repeated for two washes. A final wash was then performed (without vortexing) using 30 μL 70% ethanol in water and followed by a 3 min centrifugation at 15000 rcf at 4 °C. Following removal of the supernatant, the pellet was left to air dry for 5 minutes and stored at −80 °C.

General procedure for translation and maldi-TOF mass test: Translation was carried out using PURExpress in vitro translation system based on an adjusted version of the manufacturer's conditions[37] at 37 °C for 10 min. Solution A of the kit was replaced with a home-made version containing 50 mM HEPES-KOH (pH 7.6), 100 mM potassium acetate, 12 mM magnesium acetate, 2 mM ATP, 2 mM GTP, 1 mM CTP, 1 mM UTP, 20 mM creatine phosphate, 2 mM spermidine, 1 mM DTT, 1.5 mg/mL E. coli total tRNA, 50 μM $tRNA^{fMet}$ (uncharged), 50 μM CBTz-$tRNA^{Englu}_{GCU}$, 0.5 mM amino acids mix (0.1 mM cysteine), 1 μM mRNA, 10 μM M5 suppressor.[29] After incubation, samples were analysed by pre-purification through C18 extraction in microtip format, washing with 4% acetonitrile in water and eluting in 80% (both with 0.1% formic acid) before spotting on MALDI plates with 70% saturated α-CHCA in the elution solvent.

Procedure for P4 reprogramming and cyclization: The translation was carried out following the general procedure without extra $tRNA^{fMet}$ and with 12.5 μM of N–Ac-Phe-$tRNA^{fMet}$, 37.5 μM of CBTz-$tRNA^{Englu}_{GCU}$ and 100 μM of Lys(Cys)-$tRNA^{EnAsn}_{CAU}$ added.

Procedure for P5 reprogramming and cysteine cyclization: The translation was carried following the general procedure except that 50 μM of Dap(ClAc)-$tRNA^{EnAsn}_{CCA}$ was also added. Following translation, an equal volume of 2X cyclization buffer (0.1 M Tris, pH 8.3, 1.4 M NaCl) was added. The reaction mixture was incubated at 37 °C for 30 min then analysed by MALDI-TOF-MS as described above.

Procedure for P6 reprogramming and cyclization: The translation was carried out following the general procedure without extra $tRNA_{fMet}$ and with 12.5 μM of *N*-Ac-Phe-$tRNA^{fMet}$, 37.5 μM of CBTz-$tRNA^{Englu}_{GCU}$, 50 μM of Dap(ClAc)-$tRNA^{EnAsn}_{CCA}$ and 50 μM of Lys(Cys)-$tRNA^{EnAsn}_{CAU}$ added. The mixture was then analysed by MALDI-TOF-MS as described above.

Procedure for P7 reprogramming and TCEP addition: After translation following the general procedure, the reaction mixture was incubated in 40 mM phosphate (pH 8.5) and 5 mM DBAA (dissolved in DMF) at 42 °C for 1.5 h. Then TCEP was added to 10 mM and the reaction mixture again incubated at 42 °C for 1.5 h. The reaction mixture was finally diluted with ultrapure water (< 5% DMF) then analysed by MALDI-TOF-MS as described above.

UV absorption of luciferin: The absorption of luciferin (CBTz reacted with cysteine) was measured on nanodrop at 255 μM in DMSO/MeOH. The extinction coefficient of luciferin at 280 nm was determined to be 13725 $M^{-1}.cm^{-1}$, with $\lambda_{max} = 305$ nm.

Piperidine stability of cyanobenzothiazole: Stability of 6-hydroxy-2-cyanobenzothiazole was tested by incubating 50 mM CBTz with 10% piperidine and 50 mM Oxyma in DMF at room temperature for 4 h. A control with 50 mM CBTz in DMF alone was carried out in parallel. After incubation, both samples were diluted to 1 mM with water and analysed by LC–MS.

Chemical synthesis of test peptide P-S4: 2-chlorotrityl chloride resin was swelled by dry DCM for 15 minutes and the solvent was then drained off from the resin. The first amino acid was coupled on the resin by using 1.0 equivalent of Fmoc-Arg(pbf)-OH dissolved in dry DCM. 5.0 equivalents of DIPEA was added and the mixture was agitated vigorously for 60 minutes. After the coupling of the first amino acid, HPLC grade methanol was added (0.8 mL per gram of resin) and mixed for 15 minutes. The resin was then filtered and washed with DCM (3x), DMF (2x), again DCM (2x) and finally with MeOH (3x). The resin was dried *in vacuo*. After the coupling of the first amino acid the peptide was further synthesized by standard Fmoc solid phase peptide synthesis on a Gyros Symphony automated peptide synthesiser. The resin was deprotected with piperidine for 30 minutes and the next amino acid (4 equivalents) were added to the resin followed by a solution of HOBT/HBTU and DIPEA in DMF (each 4 equivalents). After 30 min reaction the solution was washed from the resin. These steps were repeated until the desired peptide has formed. With a final deprotection step the peptide was removed peptide from the resin by treatment with HFIP in DCM (1:4) for 30 min. The resin was filtered and washed with another portion of HFIP in DCM (1:4). The filtrates were combined and diluted in 200 mL petroleum ether. All solvents were evaporated and the peptide was concentrated *in vacuo*.

After the synthesis of the peptide a last amino acid containing the CBTz group (**S6**) (10 mg, 0.020 mmol) was coupled on the peptide (60 mg, 0.026 mmol, 1.3 equiv.) using DIPEA (14 μL, 0.080 mmol, 4.0 equiv.) and BOP coupling reagent (12 mg, 0.025 mmol, 1.3 equiv.). After coupling of this amino acid all amino acids except the cysteine were deprotected using a solution of TFA/TIPS/H_2O (95 : 2.5 : 2.5). The deprotected peptide was purified by HPLC and purity was assessed by analytical HPLC. After purification TCEP was added to deprotect the cysteine to allow the peptide to cyclize. Peptide masses are given in Table S4.

Chemical synthesis of L4 CBTz: The peptide was synthesized by standard Fmoc solid phase peptide synthesis using a Gyros protein technologies PurePepChorus automated peptide synthesiser. Fmoc-amino acids were coupled for 15 min at 55 °C on TG XV RAM resin (Rapp polymere) with 10 equiv. of DIC and 5 equiv. of Oxyma, capped with acetic anhydride and pyridine for 5 min, deprotected with 20 % piperidine, 0.1 M oxyma in DMF at 80 °C for 1.5 min. CBTz amino acid coupling was carried out with 1 equiv. of compound **5**, 2 equiv. of DIC and 2 equiv. of Oxyma at room temperature overnight. OAlloc deprotection was carried under nitrogen flow with 3 equiv. of of $Pd(PPh_3)_4$ in a mixed solvent ($CHCl_3$/CH_3COOH/NMM = 37/2/1, 15 mL/g of resin) for 8 h. The resin was washed with DMF, DCM and 0.5 % HOBt in DMF. Final Fmoc deprotection was carried out as above. On resin cyclization was carried out with 6 equiv. of DIC and 6 equiv. of oxyma in DMF/DCM = 1/1. The resin was washed with DMF, DCM and dried under vacuum, cleaved off with TFA/H_2O/TIPS (92.5/5/2.5) for 3 h. The crude peptide mixture was analyzed by LC-MS and purified by RP-HPLC. The peptide concentration was quantified by the extinction coefficient at 280 nm on nanodrop. Calculated $[M+3H]^{3+} = 774.3$, observed 774.1. Final yield was 460 μg of L4 CBTz peptide (0.8 % based on 12.5 μmol scale starting resin).

Virus neutralisation assays: Virus neutralization was assessed by a luciferase reporter assay as previously described.[33] HeLa R19 cells were seeded in 96 well plates (10^4 cells/well). After 24 h cells were transfected with the pHH-Gluc luciferase reporter plasmid using FuGENE transfection reagent (Promega) according to manufacturer's instructions. The pHH-Gluc plasmid contains the gaussian luciferase gene, flanked by 3′ and 5′ untranslated regions of the IAV NP genome segment, under control of the RNA polymerase I promoter in negative sense orientation. Generation of luciferase-encoding mRNAs and thus of luciferase protein is only observed after infection of transfected cells with IAV. After a further 24 h incubation, virus with peptides was added. Peptide dilutions were prepared in triplicates in Opti-MEM medium (Gibco) in separate plates, mixed with an equal volume of diluted virus, and incubated for 10 minutes at RT before the mixture was added to the cells. At 16 h post infection, the Renilla luciferase assay system (Promega) was used according to the manufacturer's instructions to assay samples of the cell supernatant for luciferase activity. Luminescence was measured in relative light units (RLU) using Promega GloMax Explorer.

Fluorescence polarization: Fluorescence polarization competition assays were performed in black, low-volume, nonbinding 384 well microplates "784900" (Greiner Bio-One, the Netherlands) using a BMG Labtech (Germany) PHERAstar FS microplate reader. Assay buffer consisted of 10 mM Tris.HCl pH 7.5, 150 mM NaCl, 1 % DMSO and 0.01 % Tween-20. L4 variant peptides were titrated in triplicate to 6 nM H1 ectodomain by performing a 1 : 1 serial dilution. After 10 min of incubation at room temperature, FAM-labeled L4[33] was added 1 : 1 to each well to reach a final concentration of 3 nM H1 ectodomain and 5 nM L4FAM. Recombinant soluble trimeric H1 ectodomain was generated and purified as described previously.[38] Measurements were performed after another 60 min incubation at room temperature (Excitation: 485 nm, EmissionA: 520 nm, EmissionB: 520 nm). Data were fit using GraphPad Prism 8.4.3 software to the following equation:

$$Y = bottom + \frac{top - bottom}{1 + \left(\frac{IC_{50}}{X}\right)^{hillslope}}$$

Molecular dynamics simulations of L4 thioether and L4 CBTz: For parameterisation, models of the cyclisation regions were constructed to feature the cyclisation-adjacent residues appropriate to each cyclisation methodology (Figures S6 and S7). These initial models were then minimised with DFT at a B3LYP/def2-SVP level of theory using Gaussian 16, revision D.[39] The resulting DFT densities were used for RESP fitting to produce models of the cyclisation motifs with antechamber following standard procedures.[40]

Five predicted structures of H-YRWVWTSFFSEPYFVVAG-OH were prepared with ColabFold.[41] Each model was modified to thioether cyclic peptide (cyclo[YRWVWTSFFSEPYFVVC]G-NH_2), and CBTz cyclic peptide (cyclo[CRWVWTSFFSEPYFVV-Cbtz]G-NH_2). Molecular dynamics simulations were carried out using the CUDA-enabled Amber20 software.[42] Models were firstly energy minimised in implicit water, then heated up to 300 K over 10 psec. The productive simulation was carried out at 300 K over 50 ns. The trajectories of five models were (*k*-means) conformationally clustered to produce the top 10 representative structures. Detailed conditions can be found in the input files in Tables S5 and S6. Cluster populations can be found in Tables S7 and S8. The backbone root mean square fluctuation (RMSF) relative to the average coordinates was analyzed using frames between 200 to 2000 (5 ns to 50 ns). The obtained RMSF values (in Å) was applied to the residue colours, where fluctuating residues were colored pink and rigid residues colored light blue. Per-residue RMSF values are given in Table S9.

Circular dichroism measurements of L4 thioether and L4 CBTz: Circular dichroism was measured by Jasco J-1500 Circular Dichroism Spectrometer. The peptides were diluted to 7.5 μM in 66 % MeCN at pH 6 (PBS-T, 1/40). The measurements were carried out over the wavelength between 190 nm and 350 nm with an interval of 0.5 nm. The plot was obtained as the average of the three repeated measurement on the same sample.

CD spectrum simulations were carried out using PDBMD2CD web server.[43] In the cases of both L4 thioether and L4 CBTz, the representative structures for beta sheet clusters were found to correspond qualitatively to the experimental CD spectra.

Acknowledgements

Peptide L4 thioether was previously synthesised by Ir. V Thijssen. We thank Dr. Ir. J A W Kruijtzer of Utrecht University for assistance with initial model peptide synthesis, Dr. S A Neubacher of VU Amsterdam for assistance with CD measurements, and Prof. P E G Leonards and Dr E R Amstalden van Hove for assistance with MALDI measurements. 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. MNP and CAMH acknowledge financial support from the One Health Investment Fund from the Faculty of Veterinary Medicine of the Utrecht University. TOC figure generated in part using DALL.E 2.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: chemoselective reactions · cysteine · in vitro translation · macrocyclisation · peptides

[1] A. A. Vinogradov, Y. Yin, H. Suga, *J. Am. Chem. Soc.* **2019**, *141*, 4167–4181.
[2] H. Peacock, H. Suga, *Trends Pharmacol. Sci.* **2021**, *42*, 385–397.
[3] X. Li, T. W. Craven, P. M. Levine, *J. Med. Chem.* **2022**, *65*, 11913–11926.
[4] S. Li, R. W. Roberts, *Chem. Biol.* **2003**, *10*, 233–239.
[5] Y. V. G. Schlippe, M. C. T. Hartman, K. Josephson, J. W. Szostak, *J. Am. Chem. Soc.* **2012**, *134*, 10469–10477.
[6] S. A. K. Jongkees, S. Umemoto, H. Suga, *Chem. Sci.* **2017**, *8*, 1474–1481.
[7] G. B. Vamisetti, A. Saha, Y. J. Huang, R. Vanjari, G. Mann, J. Gutbrod, N. Ayoub, H. Suga, A. Brik, *Nat. Commun.* **2022**, *13*, 6174.
[8] J. Yamaguchi, M. Naimuddin, M. Biyani, T. Sasaki, M. Machida, T. Kubo, T. Funatsu, Y. Husimi, N. Nemoto, *Nucleic Acids Res.* **2009**, *37*, e108.
[9] K. Iwasaki, Y. Goto, T. Katoh, H. Suga, *Org. Biomol. Chem.* **2012**, *10*, 5783–5786.
[10] P. Timmerman, J. Beld, W. C. Puijk, R. H. Meloen, *ChemBioChem* **2005**, *6*, 821–824.
[11] D. E. Hacker, J. Hoinka, E. S. Iqbal, T. M. Przytycka, M. C. T. Hartman, *ACS Chem. Biol.* **2017**, *12*, 795–804.
[12] S. Li, H. Cai, J. He, H. Chen, S. Lam, T. Cai, Z. Zhu, S. J. Bark, C. Cai, *Bioconjugate Chem.* **2016**, *27*, 2315–2322.
[13] S. W. Millward, T. T. Takahashi, R. W. Roberts, *J. Am. Chem. Soc.* **2005**, *127*, 14142–14143.
[14] N. Asiimwe, M. F. Al Mazid, D. P. Murale, Y. K. Kim, J.-S. Lee, *Pept. Sci.* **2021**, *114*, e24235.
[15] Y. Goto, T. Katoh, H. Suga, *Nat. Protoc.* **2011**, *6*, 779–790.
[16] L. Cui, J. Rao, in *Site-Specific Protein Labeling Methods Protoc.*, (Eds.: A. Gautier, M. J. Hinner), Springer, **2015**, pp. 81–92.
[17] Z. Zheng, P. Chen, G. Li, Y. Zhu, Z. Shi, Y. Luo, C. Zhao, Z. Fu, X. Cui, C. Ji, F. Wang, G. Huang, G. Liang, *Chem. Sci.* **2016**, *8*, 214–222.
[18] Y. Cheng, H. Peng, W. Chen, N. Ni, B. Ke, C. Dai, B. Wang, *Chem. Eur. J.* **2013**, *19*, 4036–4042.
[19] M. Jin, G. Koçer, J. I. Paez, *ACS Appl. Mater. Interfaces* **2022**, *14*, 5017–5032.
[20] S. Chen, J. Morales-Sanfrutos, A. Angelini, B. Cutting, C. Heinis, *ChemBioChem* **2012**, *13*, 1032–1038.
[21] J. T. Hampton, T. J. Lalonde, J. M. Tharp, Y. Kurra, Y. R. Alugubelli, C. M. Roundy, G. L. Hamer, S. Xu, W. R. Liu, *ACS Chem. Biol.* **2022**, *17*, 2911–2922.
[22] E. H. Abdelkader, H. Qianzhu, J. George, R. L. Frkic, C. J. Jackson, C. Nitsche, G. Otting, T. Huber, *Angew. Chem. Int. Ed.* **2022**, *61*, e202114154.
[23] D. Ye, G. Liang, M. L. Ma, J. Rao, *Angew. Chem. Int. Ed.* **2011**, *50*, 2275–2279; *Angew. Chem.* **2011**, *123*, 2323–2327.
[24] Y. Lin, Y. Gao, Z. Ma, T. Jiang, X. Zhou, Z. Li, X. Qin, Y. Huang, L. Du, M. Li, *Bioorg. Med. Chem.* **2018**, *26*, 134–140.
[25] D. T. S. Rijkers, H. P. H. M. Adams, H. C. Hemker, G. I. Tesser, *Tetrahedron* **1995**, *51*, 11235–11250.
[26] J. Morimoto, Y. Hayashi, K. Iwasaki, H. Suga, *Acc. Chem. Res.* **2011**, *44*, 1359–1368.
[27] T. Katoh, Y. Iwane, H. Suga, *RNA Biol.* **2018**, *15*, 453–460.
[28] M. Liu, V. Thijssen, S. A. K. Jongkees, *Angew. Chem. Int. Ed. Engl.* **2020**, *59*, 21870–21874.
[29] Z. Cui, Y. Wu, S. Mureev, K. Alexandrov, *Nucleic Acids Res.* **2018**, *46*, 6387–6400.
[30] H. Lindley, *Biochem. J.* **1962**, *82*, 418–425.
[31] G. Liang, H. Ren, J. Rao, *Nat. Chem.* **2010**, *2*, 54–60.
[32] M. Liu, M. Sovrovic, H. Suga, S. Jongkees, *Org. Biomol. Chem.* **2022**, *20*, 3081–3085.
[33] M. N. Pascha, V. Thijssen, J. E. Egido, M. W. Linthorst, J. H. Van Lanen, D. A. A. Van Dongen, A. J. P. Hopstaken, F. J. M. Van Kuppeveld, J. Snijder, C. A. M. De Haan, S. A. K. Jongkees, *ACS Chem. Biol.* **2022**, *17*, 2425–2436.
[34] a) H. Murakami, A. Ohta, H. Ashigai, H. Suga, *Nat. Methods* **2006**, *3*, 357–359; b) T. Fujino, Y. Goto, H. Suga, H. Murakami, *J. Am. Chem. Soc.* **2013**, *135*, 1830–1837.
[35] N. Terasaka, G. Hayashi, T. Katoh, H. Suga, *Nat. Chem. Biol.* **2014**, *10*, 555–557.
[36] T. Fujino, T. Kondo, H. Suga, H. Murakami, *ChemBioChem* **2019**, *20*, 1959–1965.
[37] R. Maini, H. Kimura, R. Takatsuji, T. Katoh, Y. Goto, H. Suga, *J. Am. Chem. Soc.* **2019**, *141*, 20004–20008.
[38] S. Zhao, N. Schuurman, M. Tieke, B. Quist, S. Zwinkels, F. J. M. Van Kuppeveld, C. A. M. De Haan, H. Egberink, *J. Clin. Microbiol.* **2020**, *58*, e01689–20.
[39] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. a. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. a. Petersson, H. Nakatsuji, X. Li, M. Caricato, a. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, a. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. a. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. a. Keith, R. Kobayashi, J. Normand, K. Raghavachari, a. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, *Gaussian 16, Revis. D.01* **2016**, Gaussian, Inc., Wallingford CT.
[40] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graphics Modell.* **2006**, *25*, 247–260.
[41] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, *Nat. Methods* **2022**, *19*, 679–682.
[42] D. A. Case, H. M. Aktulga, K. Belfon, I. Y. Ben-Shalom, J. T. Berryman, S. R. Brozell, D. S. Cerutti, I. T. E. Cheatham, G. A. Cisneros, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, G. Giambasu, M. K. Gilson, H. Gohlke, A. W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, P. A. Kollman, *Amber 2022* **2022**, University of California, San Francisco.
[43] E. D. Drew, R. W. Janes, *Nucleic Acids Res.* **2021**, *48*, W17–W24.

Manuscript received: December 15, 2022
Accepted manuscript online: December 18, 2022
Version of record online: February 6, 2023