JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

Ribosomal synthesis of backbone-cyclic peptides compatible with *in vitro* display

Ryo Takatsuji, Koki Shinbara, Takayuki Katoh, Yuki Goto, Toby Passioura, Ryo Yajima, Yamato Komatsu, and Hiroaki Suga

J. Am. Chem. Soc., **Just Accepted Manuscript** • Publication Date (Web): 16 Jan 2019

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Ribosomal synthesis of backbone-cyclic peptides compatible with *in vitro* display

Ryo Takatsuji, Koki Shinbara, Takayuki Katoh, Yuki Goto, Toby Passioura, Ryo Yajima, Yamato Komatsu, Hiroaki Suga*

Department of Chemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

*Correspondence to: hsuga@chem.s.u-tokyo.ac.jp

Abstract

Backbone-cyclic peptides are an attractive class for therapeutic development. However, *in vitro* display technologies coupled with ribosomal synthesis are intrinsically inapplicable to such "phenotypes" due to loss of the C-terminal peptide region linking to "genotype". Here we report a methodology enabling the display of backbone-cyclic peptides. To achieve this, genetic code reprogramming was utilized to implement a rearrangement strategy involving the ribosomal incorporation of a designer initiator containing a thiazolidine-protected cysteine and 2-chloroacetoamide (ClAc) sidechain, followed by an α-thio acid and cysteine at downstream positions. Upon expression of the linear peptide, spontaneous thioester rearrangement occurs between the α-thioester and the thiol group of the cysteine, liberating the α-thio group and resulting in crosslinking to the upstream ClAc sidechain group. Then, selective deprotection of the thiazolidine-protected cysteine immediately promotes intramolecular native chemical ligation, as demonstrated for various sequences and ring sizes. In this approach, the backbone-cyclic peptides retain their C-terminal peptide regions via the sidechain thioether covalent linkage, making them compatible with *in vitro* display.

Introduction

In vitro display technologies, such as ribosome, DNA, or mRNA display enable the affinity-based screening of vast numbers of peptides (or proteins) against targets of interest[1-4]. Conventional methods are able to display peptides consisting of only the 20 proteinogenic amino acid due to limitations imposed by the natural genetic code. This limits not only the usable building blocks but also the scaffold architecture, such that only linear or cyclic forms closed by

a disulfide bond(s) between Cys sidechains are generally accessible. However, recent advances in genetic code reprogramming methodologies have enabled the expression of peptides with diverse non-proteinogenic building blocks and macrocyclic scaffolds closed by non-reducible bonds. For instance, the FIT (Flexible *In vitro* Translation) system[5,6], in which genetic code reprogramming is facilitated by flexible tRNA-acylation ribozymes (flexizymes)[5,7], has made it possible to express nonstandard peptides macrocyclized (via spontaneous thioether bond formation between an N-terminal chloroacetamide group and a downstream Cys sidechain) and containing D- and/or *N*-methyl-L-amino acids[8-12]. Integration of the FIT system with mRNA display, referred to as RaPID (Random nonstandard Peptide Integrated Discovery)(Figure 1a), allows the *in vitro* selection of thioether-macrocyclic peptide ligands against therapeutic protein targets from libraries in excess of 10^{12} unique compounds, leading to the discovery of novel and potent drug candidates[13-18].

In any of the aforementioned display technologies, the "phenotype" imparted by the peptide must be linked to its "genotype" via the peptide C-terminal region. In the case of mRNA display (including RaPID), the C-termini of individual peptides are fused to puromycin attached to the 3'-end of cognate mRNAs[3,4]. This linkage allows for iterative rounds of affinity selection against a target protein and amplification of active peptide species from the library, and is therefore indispensable. This means that current methods are able to display macrocyclic peptides consisting of only N-terminus-to-sidechain (Figure 1a) or sidechain-to-sidechain linkages, but not N-terminus-to-C-terminus (Figure 1b) linkages. To wit, backbone-cyclized peptide libraries cannot be screened using current *in vitro* display methodologies. However, some naturally occurring macrocyclic peptides, such as cyclosporine A, exhibit N-terminus-to-C-terminus

linkage[19]. This motivated us to devise a new methodology overcoming this structural limitation, and through which backbone-cyclized peptides could be displayed on mRNA. Here we report a method for ribosomal synthesis of backbone-cyclized peptides without loss of the downstream peptide region, achieving this end (Figure 1c for example).

Results and Discussion

α-Thio acid incorporation into peptide backbones

To establish the method, we envisioned two multi-reaction steps: (*S1*) a step that establishes a linkage of the downstream peptide region with a sidechain on the region to be backbone-cyclized, and (*S2*) intramolecular native chemical ligation (NCL)[20] between the N-terminal amino group and a downstream thioester that is directly introduced through ribosomal translation, leading to the formation of a backbone-cyclized scaffold (Figure 1c).

To introduce a thioester bond into the expressed peptide, we planned to reprogram a codon to assign an α-thio acid. It is well known that some α-hydroxy acids are effectively incorporated into nascent peptide chains with efficiency comparable to the corresponding α-amino acids when charged onto appropriate tRNAs[21-23]. Despite the fact that α-thio acids are analogous to α-hydroxy acids, it has been reported that ribosomal thioester bond formation using α-thio acids is very poor[24]. Therefore, we first investigated the incorporation of α-thio acids into nascent peptide chains by means of FIT (Figure S1).

We chose three α-thio acids, $^{HS}F^{Cl}$, ^{HS}V, and ^{HS}G (Figure 2a), which are α-thio analogs of *p*-chlorophenylalanine, valine, and glycine, respectively, to investigate ribosome-mediated thioester bond formation. The respective α-thio acids esterified with cyanomethyl (CME for $^{HS}F^{Cl}$) or 3,5-

dinitrobenzyl (DNB for ^{HS}V and ^{HS}G) groups were charged onto tRNA$^{AsnE2}_{CUA}$ (for simplicity, described as tRNA$_{CUA}$ below) using appropriate flexizymes (eFx for CME and dFx for DNB)[5,6]. Although ^{HS}F^{Cl} and ^{HS}G were charged onto tRNA$_{CUA}$ in satisfactory yields, ^{HS}V was not charged well, presumably due to steric hindrance of the sidechain and/or poor solubility in the reaction buffer (Figure 2b). We thus chose ^{HS}F^{Cl}-tRNA$_{CUA}$ and ^{HS}G-tRNA$_{CUA}$ for testing ribosome-mediated thioester bond formation in a FIT reaction lacking release factor 1 (RF1) and in which the UAG codon was vacant. In the translation of an mRNA (mr1-X5), the UAG codon was decoded by ^{HS}F^{Cl}-tRNA$_{CUA}$ and ^{HS}G-tRNA$_{CUA}$, yielding the model peptides (p1-X5), p1-^{S}F^{Cl}5 and p1-^{S}G5, respectively (Figure 2c). MALDI-TOF-MS analysis of the respective products showed the expected mass values (Figure 2d and Figure S2b), and MS/MS analysis confirmed their sequence identity (Supplementary Figure S2a and S2c). To the best of our knowledge, this is the first demonstration that the ribosome is able to efficiently incorporate an α-thio acid into a peptide backbone. To assess the expression level of each thioester-containing peptide, we performed tricine-SDS-PAGE analysis relying on the quantitative autoradiographic detection of [^{14}C]-Asp in the C-terminal peptide region. The concentration of p1-^{S}F^{Cl}5 was determined to be 1.78 μM, which was comparable to that of a control peptide p1-F5 containing phenylalanine (2.03 μM) and high enough for further applications, whereas the expression of p1-^{S}G5 was a magnitude lower (Figure 2e). Thus, the following experiments were performed using ^{HS}F^{Cl}.

Strategy for sidechain-sidechain crosslinking prior to native chemical ligation

Direct thioester formation using ^{HS}F^{Cl} allowed us to examine NCL between the amino group of an N-terminal Cys residue and the thioester, (step *S2*). We previously reported a method to

generate an N-terminal Cys with a free α-amine using two recombinant enzymes, peptide deformylase and methionine aminopeptidase, allowing backbone cyclization via intramolecular NCL[25]. Unfortunately, this enzymatic method would not allow control of the "timing" of the thioester exchange reaction such that *S1* preceded *S2*. Therefore, we devised an alternative strategy, which was novel with respect to translated peptides, to chemically generate an N-terminal amino group on the N-terminal Cys1 residue. We considered that (*R*)-thiazolidine-4-carboxylic acid (Thz1), which has been sometimes used in ribosomal translation as a proline analogue[26,27], could be used as a precursor of N-terminal *N*-methyl-cysteine (^{MeN}Cys1). Although Thz1 could be deprotected by hydroxylamine or methoxyamine[28], we were concerned that such reagents would react with nucleobases to cause mutagenesis[29,30], thus making them unsuitable for application to *in vitro* display. Therefore, we chose the $NaBH_3CN$ reduction method to deprotect Thz1, which is unlikely to interfere with nucleobases[31]. By means of MALDI-TOF-MS, we indeed confirmed that Thz1-Cab2 used as an initiator reassigned by Thz1-Cab2-$tRNA^{fMet}_{CAU}$ was ribosomally incorporated and then cleanly converted to ^{MeN}Cys1 in a model peptide by the treatment with $NaBH_3CN$ (*vide infra*).

At this point, it was necessary to develop a strategy for achieving the *S1* step, and we naively considered the following scheme: when (*S*)-2-amino-4-(2-chloroacetamido)butanoic acid (Cab) was incorporated adjacent to Thz1, and when a Cys12 was incorporated at a downstream position of $^{HS}F^{Cl}$9, the chloroacetamido group on Cab2 would spontaneously react with the thiol group on Cys12 to form a thioether linkage (R1 in Figure 3). This large thioether-macrocyclic peptide could then be treated with $NaBH_3CN$ to liberate N-terminal ^{MeN}Cys1. Subsequently, the sidechain thiol group would attack the thioester bond to undergo NCL (S-S acyl migration followed by S-N acyl

migration to the N-methyl-amino group of ^{MeN}Cys1), forming a macrocyclic scaffold (R2 in Figure 3). During NCL, the thioester in the main backbone originally produced in translation would be cleaved, but the thioether linkage formed between the Cab2 and Cys12 residues prior to NCL would prevent division into two fragments, and the resulting backbone-cyclized peptide would be compatible with mRNA display (Figure 3b).

To test the above strategy, we designed an mRNA template (mr2-W8) expressing a model peptide, p2-W8, in a FIT reaction lacking Met and RF1, and in which the AUG and UAG codons were reassigned to Thz1-Cab2 and ^{HS}F^{Cl}9, respectively (Figure 3a). Following the 30 min translation reaction, 33 μg/mL chloramphenicol was added to the translation, aiming at slowing down the translation to avoid the formation of pre-matured peptide side-products. Under these conditions, expression of p2-W8 produced a peptide with a mass value consistent with loss of hydrogen chloride (Figure S3a), suggesting that the expected thioether-macrocyclic peptide, tcp2-W8-^{S}F^{Cl}9 (tcp denotes thioether-cyclic peptide), was produced (Figure 3b). Since macrocyclic peptides are not generally susceptible to MS/MS cleavage, the thioester bond in tcp2-W8-^{S}F^{Cl}9 was opened by addition of an excess amount of free Cys at pH 10, producing a linear peptide with a Cys-adduct, p2-W8-Cys (Figure 4a), as confirmed by MS (Figure 4b). MS/MS analysis of the resulting acyclic peptide suggested that the thioether bond did not form between Cab2 and the downstream Cys12 sidechain as planned (Figure 4c). Rather, the data appeared consistent with thioether bond formation between Cab2 and the α-thio group of the ^{HS}F^{Cl}9 residue (Figure 4d). This observation made us revise our original strategy (Figure 3b) to an alternative (Figure 5b).

Based on the observed MS/MS data indicating formation of the Cab2-^{S}F^{Cl}9 thioether linkage, the first event occurring after ribosomal synthesis of p2-W8 was likely to be thioester exchange

between $^{HS}F^{Cl}9$ and $_{HS}$Cys12 (the subscript HS denotes the sidechain thiol), establishing the Trp8-$_{S}$Cys12 thioester linkage between the backbone carbonyl group and the thiol sidechain group of $_{HS}$Cys12 (Rtex in Figure 5, forming p2-W8-$_{S}$C12). This event might occur rapidly prior to thioether bond formation, presumably due to fast nucleophilic attack of the primary thiol of $_{HS}$Cys12 at the secondary thioester in closer proximity than the Cab2 chloroacetamido sidechain near the N-terminus. This trans-thioesterification would liberate the secondary α-thiol of $^{HS}F^{Cl}9$, inducing subsequent thioether bond formation to generate a Cab2-$^{S}F^{Cl}9$ thioether linkage, yielding tcp2-W8-$_{S}$C12 (Figure 5b). Although the MS/MS data clearly supports this event as the major pathway, it should be noted that even if the minor (originally proposed) pathway occurs in parallel the final backbone cyclic moiety would be the same in both cases, as discussed below.

We next conducted the $NaBH_3CN$ reduction of Thz1 in tcp2-W8-$_{S}$C12. We observed an approximately 2 Da increase in mass value, indicating generation of the N-terminal ^{MeN}Cys1 (deprotected tcp2-W8-$_{S}$C12) (Figure S3a-b). The thiol group of ^{MeN}Cys1 could spontaneously undergo NCL with the Trp8-$_{S}$Cys12 thioester to generate an *N*-methyl-peptide bond yielding bcp2-W8 (bcp denotes backbone-cyclic peptide), but the mass value would not change before and after the NCL in this case (Figure 5b). Therefore, we confirmed the product in two ways. First, the product after $NaBH_3CN$ reduction was treated with 2-iodoacetamide (IAA), where if the desired NCL did not occur, we expected to see a single AA-adduct on the peptide at ^{MeN}Cys1; whereas if it occurred and gave backbone cyclized bcp2-W8, a double AA-adduct to the thiol sidechains of both ^{MeN}Cys1 and Cys12 should be observed (Figure 5c). In fact, we observed a sole peak corresponding to the double AA-adduct (bcp2-W8 + 2AA) (Figure S3c), indicative of macrocyclization. However, since we did not consider this IAA modification of bcp2-W8

conclusive, we applied a second method in which MS/MS analysis was performed on the final product (Figure 5d). The MS/MS spectrum of bcp2-W8 showed a series of fragmentations in the C-terminal liner peptide region ($b_{9'}$, $b_{10'}$, $b_{11'}$, $b_{12'}$, and $b_{13'}$) indicating thioether formation on $^{HS}F^{Cl}9$, as well as internal peptide fragments including the Trp8-^{MeN}Cys1 amide bond (i[b_A-y_B], i[b_A-y_C], i[b_A-y_E], and i[b_B-y_C]), directly indicating the expected backbone cyclization. Further, we did not observe strong peaks consistent with fragmentation of the macrocyclic peptide derived from the originally proposed pathway. We thus concluded that the revised pathway was the major process occurring in backbone cyclization of bcp2-W8 closed by the Trp8-^{MeN}Cys1 backbone amide bond (Figure 5b).

The observed NCL efficiency might depend on the reactivity of the thioester, dictated by the choice of amino acid residue at the adjacent upstream position[32]. To verify if other residues could affect ligation efficiency, we substituted Trp8 in p2-W8 with Gly8, Lys8, or Val8 (p2-G8, p2-K8, and p2-V8). MALDI-TOF-MS analysis of the respective peptides showed clean expression of the full length of peptide via spontaneous thioether bond formation (Figure S4a, S4d and S4g), and the subsequent deprotection of Thz1 to ^{MeN}Cys1 also yielded products with the expected MS values (Figure S4b, S4e and 4h). The IAA treatment of bcp2-G8 and bcp2-K8 demonstrated MS values consistent with their double AA-adduct (Figure S4c and S4f), indicating the completion of backbone macrocyclization. However, IAA treatment of bcp2-V8 showed two peaks, consistent with formation of both the double and single AA-adducts, suggesting that backbone macrocyclization was not complete (Figure S4i). We speculated that because of the steric hindrance of Val8, the intramolecular NCL in tcp2-V8 was slower than tcp2-W8, tcp2-G8, and tcp2-K8 under acidic conditions with $NaBH_3CN$. In order to accelerate the cyclization, we added

an additional step to neutralize the reaction buffer after $NaBH_3CN$ reduction. This resulted in observation of a single peak consistent with formation of the double AA-adduct of bcp2-V8 (Figure S4j), indicating that the complete backbone cyclization for bcp2-V8 was also achieved.

Next, we wondered if the thioester exchange between $^{HS}F^{Cl}9$ and Cys12 followed by backbone-cyclization was dependent on the linking sequence, Ile10-Gly11. We also wondered if the length of this linking sequence was of importance. We thus designed two sequences, in one of which the Ile10-Gly11 sequence was replaced with Ile10-Pro11 (p3-W8) and a second in which a Ser residue (Ser11) was inserted to give Gly11-Ser12-Cys13 (p4-W8) (Figure 6a). Both peptides were ribosomally synthesized and cyclized by the procedure outlined above. Mass spectra of the product gave peaks corresponding to the doubly AA-modified bcp3-W8 and bcp4-W8, indicating generation of the desired backbone-cyclized peptides closed by the Trp8-^{MeN}Cys1 bond (Figure 6b and 6c). This suggests that the linking sequence composition and length between $^{HS}F^{Cl}$ and the downstream Cys residue are not critical.

We further confirmed that this backbone-cyclization methodology could be applied to other sequences. We designed two linear peptide sequences, p5-L11 and p6-Q14, forming 11- and 14-residue rings closed by L11-^{MeN}Cys1 and Q14-^{MeN}Cys1 backbone amide bonds, respectively (bcp5-L11 and bcp5-Q14, Figure 6d). MS analysis demonstrated that treatment of the expressed p5-L11 and p6-Q14 peptides with $NaBH_3CN$, neutralization, and then IAA coupling yielded the respective backbone cyclic peptides with double AA-modification (Figure 6e and 6f). Thus, since the macrocyclization process proceeded smoothly in all constructs irrespective of sequence (three distinct sequences were tested) or ring size (8, 11, and 14 residue-rings were tested) even at

distinct ligation residues (Trp, Lys, Val, Leu, and Gln were tested), this strategy of backbone macrocyclization appears very likely to be applicable to a wide range of sequences and ring sizes.

To examine if this backbone-macrocyclization methodology could be compatible with the mRNA display, two experiments were performed. First, we designed a mRNA template mr7, expressing p7-L10 (Figure 7a), with UAG stop codon where ribosome would stall in the absence of RF1. After the translation of p7-L10, 1.2 μM puromycin (Pu) was added to trap the C-terminus of p7-L10-tRNA instead of chloramphenicol. After the chemical transformation using the protocol established in this report, the product was analyzed by LC/ESI-MS. We observed a single major peak in the LC diagram (Figure 7b), corresponding to an ionized mass (*m/z*) of 1168.47 (Figure 7c). This mass value was consistent with the calculated mass of $[M–2H]^{2-}$ (*m/z* = 1168.42), indicating that bcp7-L10-Pu was formed. In addition, MALDI-TOF analysis of the IAA-treated bcp-L10-Pu allowed us to detect the MS values consistent with their double AA-adduct (Figure 7d), indicating the completion of backbone macrocyclization. Second, we performed the analysis of mRNA under the conditions of incubation with 1M $NaBH_3CN$, 1M NaOAc at pH 4.8 for 21 hours, giving no damage of the mRNA template (Supplementary Figure S5). These results firmly show that the technology developed in this work enables display of backbone-macrocyclic peptides on their cognate puromycin-mRNA conjugates

Conclusion

NCL has been applied to the synthesis of various peptides and proteins[20,33]. This fundamental chemical concept is also found in a naturally occurring genetically encoded system, so-called inteins[34]. Intein systems have been utilized for the expression of backbone-cyclic peptides,

referred to as split-intein circuit ligation of peptides and proteins (SICLOPPS)[35]. Although the SICLOPPS method combined with reverse two-hybrid or other in-cell screening systems allows the identification of *de novo* backbone-cyclic peptide ligands against protein targets of interest[36-38], its library size is limited to 10^9 or less, resulting in relatively rare discovery of bioactive peptides with very high potencies. By contrast, *in vitro* display techniques (including RaPID), which are compatible with libraries of 10^{12} compounds or greater, have more reliably yielded potent macrocyclic peptide ligands against a wide range of targets including both intra- or extra-cellular proteins, and exhibiting K_D (or IC_{50}) values in the low nM range or below. Unfortunately, because such display methods rely on the linkage of peptide C-terminus to cognate mRNA template, backbone-cyclization has been incompatible to the method; *i.e.* the displayable scaffold of macrocyclic peptides on mRNA consists of N-terminus-to-sidechain or sidechain-to-sidechain linkages.

Here we have devised a new methodology compatible with *in vitro* display of backbone-cyclic peptides. Several key technical advances have been made to realize this. First, incorporation of an α-thio acid, $^{HS}F^{Cl}$, into the peptide main chain was achieved by genetic code reprogramming. Although the classic study using a puromycin analog where the α-amino group is substituted with α-thio group suggests the ribosome capability for the formation of thioester bond[39], this system is not elongatable after thioester bond formation. To the best of knowledge, there is no report in literature that an α-thio acid is ribosomally incorporated and elongated into peptide chain. Thus, this work represents the first example of such events. Second, the sidechain thiol of a Cys residue embedded at a position downstream of the $^{HS}F^{Cl}$-thioester bond was spontaneously exchanged with the ribosomally formed thioester, yielding a new thioester bond

between the main chain and Cys sidechain. This rearrangement was initially not planned, but during the course of our experiments we found that it proceeded rapidly and efficiently prior to thioether bond formation as described below. Third, the liberated thiol group of $^{HS}F^{Cl}$ attacks the upstream Cab residue bearing a chloroacetamido sidechain, forming a thioether-cyclic peptide. This Cab residue was introduced at the N-terminus together with Thz, a protected *N*-methyl-Cys (^{MeN}Cys), as a Thz-Cab residue by reprogramming of the initiation of translation. Finally, deprotection of the Thz group under mild reducing conditions converted it to ^{MeN}Cys, the thiol sidechain of which induces backbone-cyclization via NCL. Despite the complexity of the above reactions, each step of the process worked with acceptable efficiency to produce backbone-cyclic peptides without loss of the original downstream C-terminal peptide region, meaning that the methodology should be compatible with *in vitro* display methods (Figure 5c).

This methodology leverages the full power of genetic code reprogramming assisted by flexizymes. When combined with further nonproteinogenic amino acid incorporation, such as D- and *N*-methyl-L-amino acids, which is certainly possible, RaPID display of backbone-cyclized pseudo-natural peptides will be achievable through *in vitro* selection against protein targets of interest. This opens a critical door to the discovery of new therapeutic modalities and diagnostic molecules based on such pseudo-natural peptides.

Methods

Additional experimental procedures are provided in the Supplementary Information.

In vitro translation using the FIT system

The reconstituted cell-free translation system used in this study contained all the necessary components for translation except for RF1 (and methionine, if necessary). The concentration of cysteine (0.05 mM) was decreased compared to that of the other 19 proteinogenic amino acids (0.5 mM) to reduce the possibility of reaction of produced thioester-containing peptide with cysteine. The composition of the FIT reaction was as follows; 50 mM HEPES-KOH (pH 7.6), 12 mM magnesium acetate, 100 mM potassium acetate, 2 mM spermidine, 20 mM creatine phosphate, 2 mM DTT, 2 mM ATP, 2 mM GTP, 1 mM CTP, 1 mM UTP, 0.1 mM 10-formyl-5,6,7,8-tetrahydrofolic acid, 0.5 mM 19 proteinogenic amino acids other than cysteine, and 0.05 mM cysteine and 1.5 mg/mL E. coli total tRNA along with 0.73 μM AlaRS, 0.03 μM ArgRS, 0.38 μM AsnRS, 0.13 μM AspRS, 0.02 μM CysRS, 0.06 μM GlnRS, 0.23 μM GluRS, 0.09 μM GlyRS, 0.02 μM HisRS, 0.4 μM IleRS, 0.04 μM LeuRS, 0.11 μM LysRS, 0.03 μM MetRS, 0.68 μM PheRS, 0.16 μM ProRS, 0.04 μM SerRS, 0.09 μM ThrRS, 0.03 μM TrpRS, 0.02 μM TyrRS, 0.02 μM ValRS, 0.6 μM MTF, 2.7 μM IF1, 0.4 μM IF2, 1.5 μM IF3, 0.26 μM EF-G, 10 μM EF-Tu, 10 μM EF-Ts, 0.25 μM RF2, 0.17 μM RF3, 0.5 μM RRF, 0.1 μM T7 RNA polymerase, 4 μg/mL creatine kinase, 3μg/mL myokinase, 0.1 μM pyrophosphatase, 0.1 μM nucleotide-diphosphatase kinase, 1.2 μM ribosome and 400 nM DNA templates.

In vitro translation was carried out using the FIT system in the presence of 100 μM $^{HS}F^{Cl}$-$tRNA_{CUA}$ (and 100 μM Thz-Cab-$tRNA^{fMet}_{CAU}$, if necessary). Translation reactions were started by adding the acyl tRNA(s) to the FIT system at 37˚C, and translation time was set to 30 min. To promote thioether bond formation, additional incubation was performed for 90 min in the presence of 33 μg/mL chloramphenicol except for in the case of expression of p1-X5 and p7-L10.

MALDI-TOF-MS and MS/MS analysis

After *in vitro* translation, the solutions were desalted using an SPE C-tip (Nikkyo Technos) and eluted with 80% acetonitrile, 0.5% acetic acid solution 50% saturated with the matrix (*R*)-cyano-4-hydroxycinnamic acid (Bruker). MALDI-TOF-MS and MS/MS measurements were performed under linear positive mode using an Autoflex II (Bruker) or ultrafleXtreme (Bruker) with external calibration (Peptide Calibration Standard II, Bruker).

Tricine-SDS-PAGE analysis for translated peptides

For autoradiography analysis, the translation reactions (2.5 µl) were performed in the presence of 50 µM [^{14}C]-Asp instead of 500 µM cold Asp, and terminated by adding an equal volume of 2×Tricine-SDS-PAGE loading buffer (900 mM Tris-HCl (pH 8.45), 8% (w/v) SDS, 30% (v/v) glycerol). The translation products were analyzed by 15% tricine-SDS-PAGE (150 V constant for 40 min.) and autoradiography was performed using an FLA-5100 (Fujifilm Life Science) or a Typhoon FLA 7000 (GE Healthcare). The amount of peptide product was quantified based on the band intensities of known amounts of [^{14}C]-Asp as standards. Values reported are the average of three independent reactions, and error bars represent the standard deviation.

N-terminal ^{MeN}Cys liberation of expressed peptides possessing an *N*-terminal Thz residue

2.5 µL of the ^{MeN}Cys liberation buffer (1 M $NaBH_3CN$, 1 M NaOAc (pH 4.8)) was added to 2.5 µL of translation solution. After incubation at 25˚C for 21 hours, N-terminal Thz was cleanly deprotected to N-terminal ^{MeN}Cys.

Thioester bond opening in thioether-cyclic peptides by the addition of excess cysteine

To 5μL of peptide solution, 5μL of the thioester-opening buffer (800 mM Na_2CO_3-$NaHCO_3$ (pH 10.0), 20 mM cysteine, 40 mM TCEP-HCl) was added. The reaction was completed in 1 hour at 42 ˚C.

Free thiol detection in peptides with IAA

5 μL of IAA buffer (100 mM IAA, 1 M HEPES-NaOH (pH 8.0)) was added to 5μL of peptide solution. The reaction was completed in 1 hour at 25 ˚C.

Backbone macrocyclization acceleration in neutral condition

To 5μL of peptide solution, 6 μL of 1 M HEPES-NaOH (pH 8.0) and 1 μL of 400 mM TCEP-HCl were added and incubated at 25℃ for 3 hours to promote backbone macrocyclization.

Supporting Information

The supporting information is available free of charge on the ACS publications website, URL: http://pubs.acs.org.

Supporting figures, supporting materials and methods, and supporting references.

Acknowledgements

This work was partially supported by CREST for Molecular Technologies, JST to H.S., and KAKENHI (16H06444 and 26220204 to H.S.; 26560429 to T.K.; 16H01131 and 17H04762 to Y.G.) from the Japan Society for the Promotion of Science.

References

(1) Mattheakis, L. C.; Bhatt, R. R.; Dower, W. J., An in-vitro polysome display system for identifying ligands from very large peptide libraries. *Proc. Natl. Acad. Sci. USA* **1994,** *91*, 9022-9026.

(2) Odegrip, R.; Coomber, D.; Eldridge, B.; Hederer, R.; Kuhlman, P. A.; Ullman, C.; FitzGerald, K.; McGregor, D., CIS display: In vitro selection of peptides from libraries of protein-DNA complexes. *Proc. Natl. Acad. Sci. USA* **2004,** *101*, 2806-2810.

(3) Nemoto, N.; Miyamoto-Sato, E.; Husimi, Y.; Yanagawa, H., In vitro virus: Bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett.* **1997,** *414*, 405-408.

(4) Roberts, R. W.; Szostak, J. W., RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. USA* **1997,** *94*, 12297-12302.

(5) Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H., A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods* **2006,** *3*, 357-359.

(6) Goto, Y.; Katoh, T.; Suga, H., Flexizymes for genetic code reprogramming. *Nat. Protoc.* **2011**, *6*, 779-790.

(7) Xiao, H.; Murakami, H.; Suga, H.; Ferré-D'Amaré, A. R., Structural basis of specific tRNA aminoacylation by a small in vitro selected ribozyme. *Nature* **2008,** *454*, 358-

361.

(8) Goto, Y.; Ohta, A.; Sako, Y.; Yamagishi, Y.; Murakami, H.; Suga, H., Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chem. Biol.* **2008,** *3*, 120-129.

(9) Goto, Y.; Murakami, H.; Suga, H., Initiating translation with D-amino acids. *RNA* **2008,** *14*, 1390-1398.

(10) Kawakami, T.; Murakami, H.; Suga, H., Messenger RNA-programmed incorporation of multiple N-methyl-amino acids into linear and cyclic peptides. *Chem. Biol.* **2008,** *15*, 32-42.

(11) Katoh, T.; Tajima, K.; Suga, H., Consecutive elongation of D-amino acids in translation. *Cell Chem. Biol.* **2017,** *24*, 46-54.

(12) Katoh, T.; Iwane, Y.; Suga, H., Logical engineering of D-arm and T-stem of tRNA that enhances D-amino acid incorporation. *Nucleic Acids Res.* **2017,** *45*, 12601-12610.

(13) Yamagishi, Y.; Shoji, I.; Miyagawa, S.; Kawakami, T.; Katoh, T.; Goto, Y.; Suga, H., Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chem. Biol.* **2011,** *18*, 1562-1570.

(14) Hayashi, Y.; Morimoto, J.; Suga, H., In vitro selection of anti-Akt2 thioether-macrocyclic peptides leading to isoform-selective inhibitors. *ACS Chem. Biol.* **2012,** *7*, 607-613.

(15) Morimoto, J.; Hayashi, Y.; Suga, H., Discovery of macrocyclic peptides armed with a mechanism-based warhead: isoform-selective inhibition of human deacetylase SIRT2.

Angew. Chem. Int. Ed. Engl. **2012,** *124*, 3479-3483.

(16) Ito, K.; Sakai, K.; Suzuki, Y.; Ozawa, N.; Hatta, T.; Natsume, T.; Matsumoto, K.; Suga, H., Artificial human Met agonists based on macrocycle scaffolds. *Nat. Commun.* **2015,** *6*, 6373.

(17) Song, X.; Lu, L.Y.; Passioura, T.; Suga, H., Macrocyclic peptide inhibitors for the protein-protein interaction of Zaire Ebola virus protein 24 and karyopherin alpha 5. *Org. Biomol. Chem.* **2017,** *15*, 5155-5160.

(18) Kawamura, A.; Munzel, M.; Kojima, T.; Yapp, C.; Bhushan, B.; Goto, Y.; Tumber, A.; Katoh, T.; King, O. N.; Passioura, T.; Walport, L. J.; Hatch, S. B.; Madden, S.; Muller, S.; Brennan, P. E.; Chowdhury, R.; Hopkinson, R. J.; Suga, H.; Schofield, C. J., Highly selective inhibition of histone demethylases by de novo macrocyclic peptides. *Nat. Commun.* **2017,** *8*, 14773.

(19) Joo, S. H., Cyclic peptides as therapeutic agents and biochemical tools. *Biomol. Ther.* **2012,** *20*, 19.

(20) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S., Synthesis of proteins by native chemical ligation. *Science* **1994,** *266*, 776-779.

(21) Fahnestock, S.; Rich, A., Ribosome-catalyzed polyester formation. *Science* **1971,** *173*, 340-343.

(22) Ohta, A.; Murakami, H.; Higashimura, E.; Suga, H., Synthesis of polyester by means of genetic code reprogramming. *Chem. Biol.* **2007,** *14*, 1315-1322.

(23) Fahnestock, S.; Rich, A., Synthesis by ribosomes of viral coat protein containing ester linkages. *Nat. New Biol.* **1971,** *229*, 8.

(24) Ellman, J. A.; Mendel, D.; Schultz, P. G., Site-specific incorporation of novel backbone structures into proteins. *Science* **1992,** *255*, 197-200.

(25) Kawakami, T.; Ohta, A.; Ohuchi, M.; Ashigai, H.; Murakami, H.; Suga, H., Diverse backbone-cyclized peptides via codon reprogramming. *Nat. Chem. Biol.* **2009,** *5*, 888.

(26) Bekhor, I. J.; Mohseni, Z.; Bavetta, L. A., Inhibition of proline-C14 incorporation into rat liver ribosomes by thiazolidine-4-carboxylic acid in a cell-free system. *Proc. Soc. Exp. Biol. Med.* **1965**, *119*, 765-769.

(27) Hartman, M. C.; Josephson, K.; Lin, C. W.; Szostak, J. W., An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides. *PLoS One* **2007**, *2*, e972.

(28) Villain, M.; Vizzavona, J.; Rose, K., Covalent capture: a new tool for the purification of synthetic and recombinant polypeptides. *Chem. Biol.* **2001,** *8*, 673-679.

(29) Stolarski, R.; Kierdaszuk, B.; Hagberg, C. E.; Shugar, D., Hydroxylamine and methoxyamine mutagenesis: displacement of the tautomeric equilibrium of the promutagen N6-methoxyadenosine by complementary base pairing. *Biochemistry* **1984,** *23*, 2906-2913.

(30) Stolarski, R.; Kierdaszuk, B.; Hagberg, C. E.; Shugar, D., Mechanism of hydroxylamine mutagenesis - tautomeric shifts and proton-exchange between the promutagen N6-methoxyadenosine and cytidine. *Biochemistry* **1987,** *26*, 4332-4337.

(31) Tsukiji, S.; Pattnaik, S. B.; Suga, H., An alcohol dehydrogenase ribozyme. *Nat. Struct. Mol. Biol.* **2003,** *10*, 713.

(32) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E., Protein synthesis by native chemical

ligation: Expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. USA* **1999,** *96*, 10068-10073.

(33) Thapa, P.; Zhang, R. Y.; Menon, V.; Bingham, J. P., Native chemical ligation: a boon to peptide chemistry. *Molecules* **2014,** *19*, 14461-14483.

(34) Shah, N. H.; Muir, T. W., Inteins: Nature's gift to protein chemists. *Chem. Sci.* **2014,** *5*, 446-461.

(35) Scott, C. P.; Abel-Santos, E.; Wall, M.; Wahnon, D. C.; Benkovic, S. J., Production of cyclic peptides and proteins in vivo. *Proc. Natl. Acad. Sci. USA* **1999,** *96*, 13638-13643.

(36) Horswill, A. R.; Savinov, S. N.; Benkovic, S. J., A systematic method for identifying small-molecule modulators of protein-protein interactions. *Proc. Natl. Acad. Sci. USA* **2004,** *101*, 15591-15596.

(37) Tavassoli, A.; Benkovic, S. J., Genetically selected cyclic-peptide inhibitors of AICAR transformylase homodimerization. *Angew. Chem. Int. Ed. Engl.* **2005,** *44*, 2760-2763.

(38) Young, T. S.; Young, D. D.; Ahmad, I.; Louis, J. M.; Benkovic, S. J.; Schultz, P. G., Evolution of cyclic peptide protease inhibitors. *Proc. Natl. Acad. Sci. USA* **2011,** *108*, 11052-11056.

(39) Gooch, J.; Hawtrey, A. O., Synthesis of thiol-containing analogues of puromycin and a study of their interaction with N-acetylphenylalanyl-transfer ribonucleic acid on ribosomes to form thioesters. *Biochem. J.* **1975,** *149*, 209-220.

Figure Legends

Figure 1| Compatibility of macrocyclization with *in vitro* display. (a) The N-terminal ClAc group reacts with the sidechain of a downstream Cys, forming a thioether macrocyclic peptide, which retains the fusion of downstream peptide region (shown in cyan) to Pu-mRNA. Pu indicates puromycin. This system is compatible with *in vitro* display techniques. (b) Backbone macrocyclization via native chemical ligation (NCL), producing detachment of the downstream peptide region from Pu-mRNA. This system is incompatible with *in vitro* display techniques. (c) A new strategy to display backbone-cyclized peptides via an *N*-methyl-peptide bond, which is compatible with *in vitro* display technologies. Peptides are linked to mRNA by two steps (*S1* and *S2*). Details are discussed in this work.

Figure 2| Ribosome-mediated thioester bond formation. (a) Chemical structure of the α-thio acids used in this study. (b) Quantification of microhelix RNA thioacylation by acid-PAGE analysis. (c) The mRNA sequence (mr1) encoding the p1-X5 peptide. X represents an α-thio acid of choice as shown in (a). (d) MALDI-TOF-MS of the thioester-containing peptide P1-^{s}F$^{p\text{-}Cl}$ expressed from mr1 in the FIT system. (e) Quantification of peptides expressed from mr1 in the FIT system. The expression levels (μM in the reaction mixture) of peptides were quantified based on the band intensities of known amounts of [^{14}C]-Asp as standards. Lane 1, p1-F5 suppressed with phenylalanine F-tRNA$_{CUA}$; p1-^{S}F^{Cl}5 suppressed with ^{HS}F^{Cl}-tRNA$_{CUA}$; p1-^{S}G5 suppressed with ^{HS}G-tRNA$_{CUA}$.

Figure 3| The initial strategy for linking upstream and downstream sequences to form backbone-cyclized peptides compatible with *in vitro* display. (a) The mRNA sequence (mr2-

W8) encoding the p2-W8 peptide. R1 represents the first reaction that spontaneously forms a thioether bond between Cab2 and Cys12. R2 represents the second reaction (NCL forming an amide backbone) after deprotection of Thz1. (b) Schematic view of the entire process via R1 and R2. In this process, the downstream sequence of the peptide would not detach from the backbone-cyclized peptide, and would thus be compatible with *in vitro* display. "tcp" and "bcp" stand for thioether-cyclic peptide and backbone-cyclic peptide, respectively. The C-terminal Ser residue is shown as "S". Pu-mRNA shows the potential fusion site at the C-terminal of the peptide.

Figure 4| Analysis of tcp2-W8 by MALDI-TOF MS/MS. (a) Schematic view of cleavage of the thioester in tcp2-W8 by the addition of Cys via thioester exchange followed by intramolecular amidation. (b) MALDI-TOF-MS of p2-W8-Cys**.** (c) MS/MS analysis of p2-W8-Cys. Fragmentation data was inconsistent with the predicted structure of p2-W8-Cys as shown in the right fragmentation profile. (d) The revised structure of p2-W8-Cys, the fragmentation profile of which was consistent with the observed data.

Figure. 5| Revised pathway of backbone-cyclized peptide synthesis. (a) The mRNA sequence (mr2-W8) encoding the p2-W8 peptide. Rtex represents spontaneous thioester exchange between $^{HS}F^{Cl}9$ and Cys12 prior to R1 and R2. R1 forms a thioether bond between Cab2 and $^{HS}F^{Cl}9$, and R2 forms a backbone-cyclized peptide, bcp2-W8, via NCL after deprotection of Thz1. (b) Schematic view of the entire process via Rtex, R1, and R2. In this process, the downstream peptide region would not detach from the backbone-cyclized peptide, and would thus be compatible with *in vitro* display. (c) Treatment of bcp2-W8-$_S$C12 and bcp2-W8 with 2-

iodoacetoamide (IAA). The initial peptide intermediate should be modified with a single AA, whereas the latter backbone cyclized peptide should be modified with double AA (see Figure S3 for the MS data). (d) MALDI-TOF MS/MS analysis of bcp2-W8. "i[b_X-y_X]" labels indicate the peaks corresponding to internal peptide fragments resulting from two fragmentations inside the macrocyclic moiety. The internal peptide fragments containing the newly generated Trp8-^{MeN}Cys1 amide bond are shown in bold.

Figure 6| Backbone-cyclization of various peptides compatible with *in vitro* display methods. (a) Two peptides, p3-W8 and p4-W8, with two or three amino acids between the $^{HS}F^{Cl}$9 and the downstream Cys. (b) MS data of tcp3-W8 (top), bcp3-W8 (middle) after $NaBH_3CN$ treatment, and bcp3-W8 2IAA adduct (bottom). (c) MS data of tcp4-W8 (top), bcp4-W8 (middle) after $NaBH_3CN$ treatment, and bcp4-W8 2IAA adduct (bottom). (d) Two peptides, p5-L11 and p6-Q14, with different amino acid sequences between the Cab2 and the $^{HS}F^{Cl}$9. (e) MS data of tcp5-L11 (top), bcp5-L11 (middle) after $NaBH_3CN$ treatment, and bcp5-L11 2IAA adduct (bottom). (f) MS data of tcp6-Q14 (top), bcp6-Q14 (middle) after $NaBH_3CN$ treatment, and bcp6-Q14 2IAA adduct (bottom).

Figure 7| Conjugation of a model backbone-cyclic peptide with a puromycin.
(a) The mRNA sequence (mr7) encoding p7-L10 conjugated with a puromycin at the C-terminus followed by backbone cyclization to yield bcp7-L10. (b) Mass chromatogram at *m/z* = 1168.42 obtained from LC/ESI-MS analysis of p7-L10. (c) Mass spectrum of p7-L10 isolated at

the retention of 11.51 min. (d) MALDI-TOF MS of IAA-treated bcp-L10-Pu. In the right panel, the region of 2450–2500 *m/z* was expanded to show their isotope peaks.

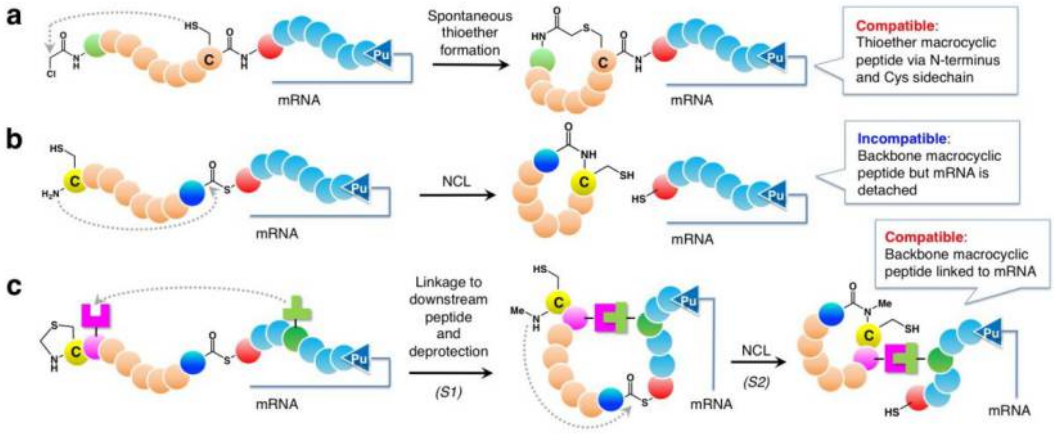
a
Spontaneous thioether formation
mRNA
mRNA
Compatible:
Thioether macrocyclic peptide via N-terminus and Cys sidechain
b
NCL
mRNA
mRNA
Incompatible:
Backbone macrocyclic peptide but mRNA is detached
c
Linkage to downstream peptide and deprotection
(S1)
mRNA
NCL
(S2)
mRNA
Compatible:
Backbone macrocyclic peptide linked to mRNA

Figure 1

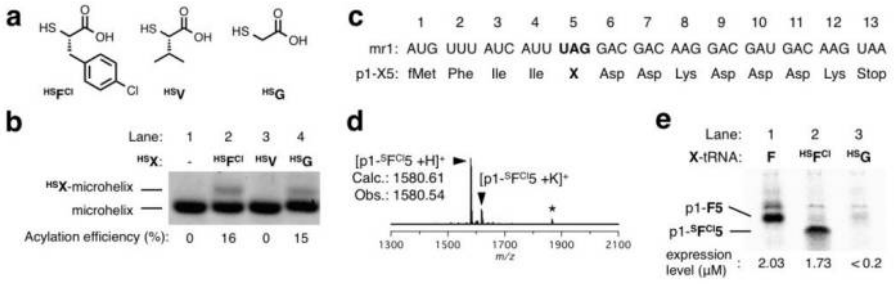
a
HS
OH
O
Cl
HSFCl
HSV
HSG
b
Lane: 1 2 3 4
HSX: - HSFCl HSV HSG
HSX-microhelix
microhelix
Acylation efficiency (%): 0 16 0 15
c
1 2 3 4 5 6 7 8 9 10 11 12 13
mr1: AUG UUU AUC AUU UAG GAC GAC AAG GAC GAU GAC AAG UAA
p1-X5: fMet Phe Ile Ile X Asp Asp Lys Asp Asp Asp Lys Stop
d
[p1-5FCl5 +H]+
Calc.: 1580.61
Obs.: 1580.54
[p1-5FCl5 +K]+
1300 1500 1700 1900 2100
m/z
e
Lane: 1 2 3
X-tRNA: F HSFCl HSG
p1-F5
p1-5FCl5
expression level (μM): 2.03 1.73 <0.2

Figure 2

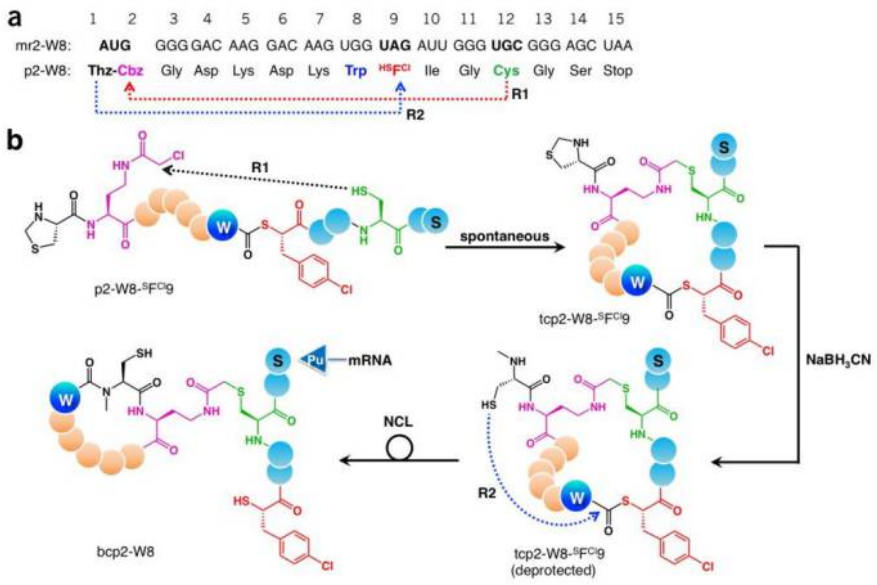
a
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
mr2-W8: AUG GGG GAC AAG GAC AAG UGG UAG AUU GGG UGC GGG AGC UAA
p2-W8: Thz-Cbz Gly Asp Lys Asp Lys Trp HSFCl Ile Gly Cys Gly Ser Stop
R1
R2
b
R1
p2-W8-SFCl9
spontaneous
tcp2-W8-SFCl9
NaBH3CN
mRNA
NCL
bcp2-W8
tcp2-W8-SFCl9
(deprotected)
R2

Figure 3

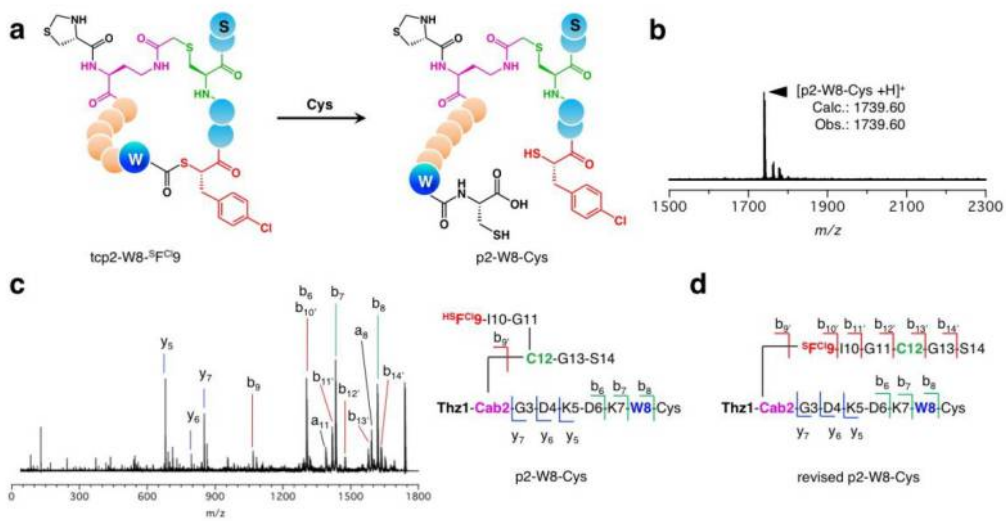
a
Cys
tcp2-W8-SFCl9
p2-W8-Cys
b
[p2-W8-Cys +H]+
Calc.: 1739.60
Obs.: 1739.60
1500
1700
1900
2100
2300
m/z
c
0
300
600
900
1200
1500
1800
m/z
HSFCl9-I10-G11
C12-G13-S14
Thz1-Cab2-G3-D4-K5-D6-K7-W8-Cys
p2-W8-Cys
d
SFCl9-I10-G11-C12-G13-S14
Thz1-Cab2-G3-D4-K5-D6-K7-W8-Cys
revised p2-W8-Cys

Figure 4

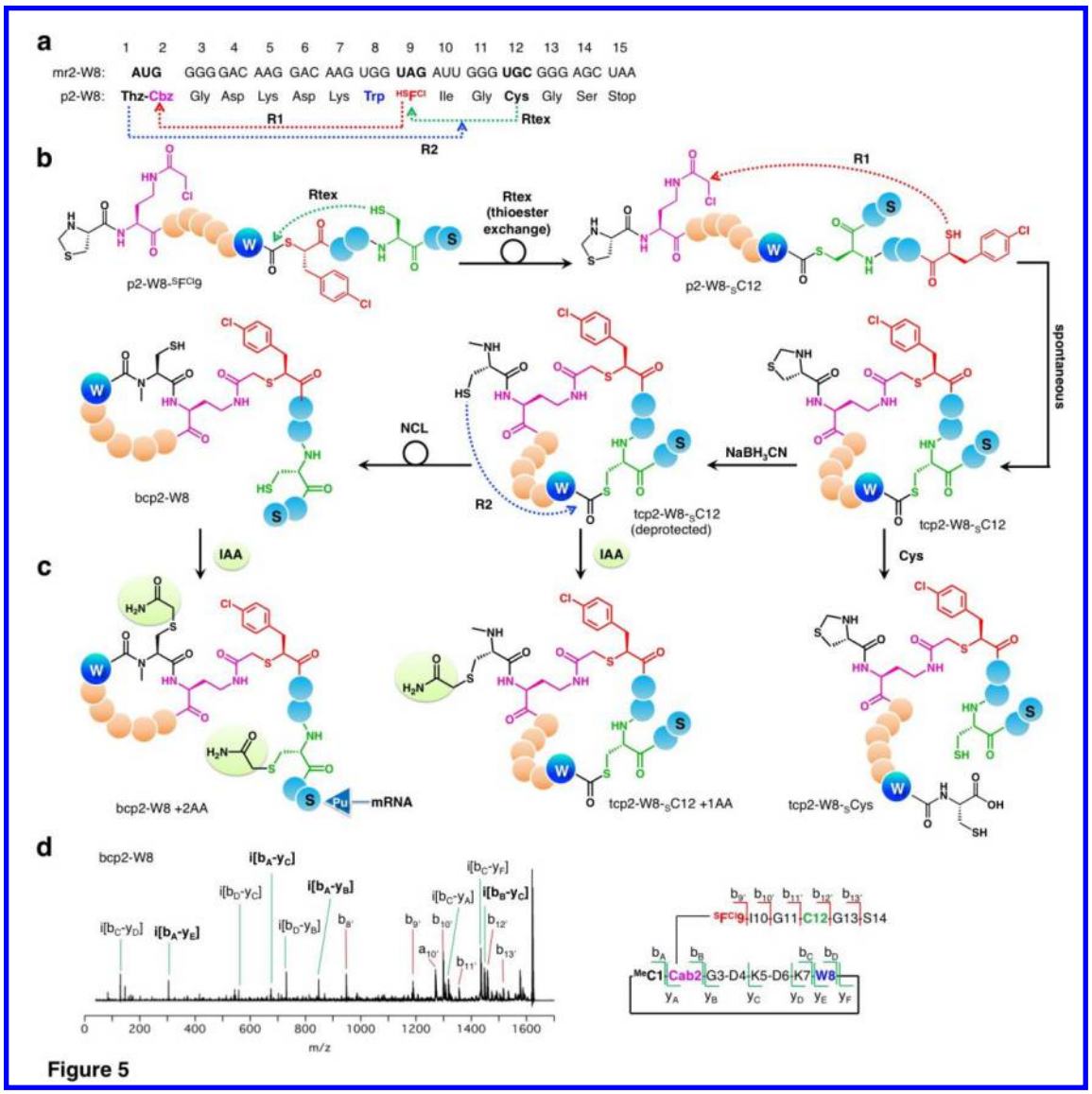
a
mr2-W8: AUG GGG GAC AAG GAC AAG UGG UAG AUU GGG UGC GGG AGC UAA
p2-W8: Thz-Cbz Gly Asp Lys Asp Lys Trp HSFCl Ile Gly Cys Gly Ser Stop
R1
R2
Rtex
b
p2-W8-SFClg
Rtex
Rtex (thioester exchange)
p2-W8-SC12
R1
spontaneous
NaBH3CN
tcp2-W8-SC12
tcp2-W8-SC12 (deprotected)
R2
NCL
bcp2-W8
IAA
IAA
Cys
c
bcp2-W8 +2AA
mRNA
tcp2-W8-SC12 +1AA
tcp2-W8-SCys
d
bcp2-W8
m/z
SFClg I10 G11 C12 G13 S14
MeC1 Cab2 G3 D4 K5 D6 K7 W8

Figure 5

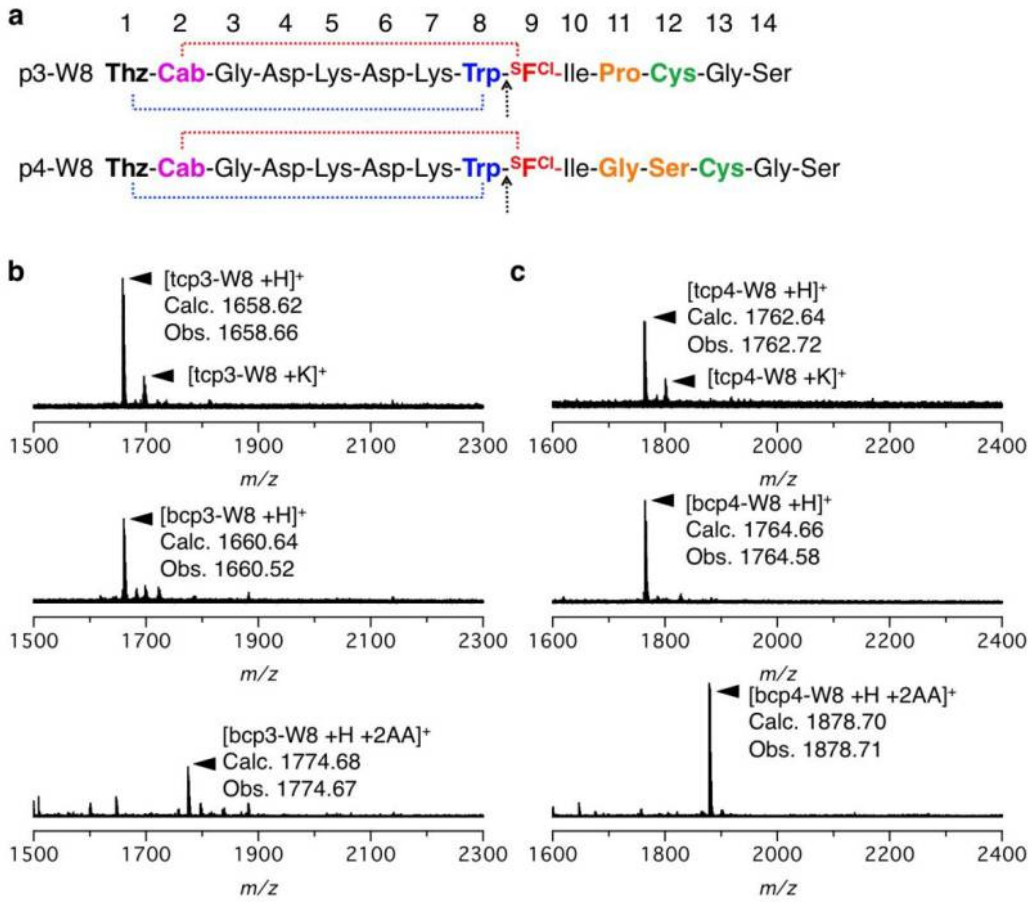
a
1 2 3 4 5 6 7 8 9 10 11 12 13 14
p3-W8 Thz-Cab-Gly-Asp-Lys-Asp-Lys-Trp-SFCl-Ile-Pro-Cys-Gly-Ser
p4-W8 Thz-Cab-Gly-Asp-Lys-Asp-Lys-Trp-SFCl-Ile-Gly-Ser-Cys-Gly-Ser
b
[tcp3-W8 +H]+
Calc. 1658.62
Obs. 1658.66
[tcp3-W8 +K]+
[bcp3-W8 +H]+
Calc. 1660.64
Obs. 1660.52
[bcp3-W8 +H +2AA]+
Calc. 1774.68
Obs. 1774.67
c
[tcp4-W8 +H]+
Calc. 1762.64
Obs. 1762.72
[tcp4-W8 +K]+
[bcp4-W8 +H]+
Calc. 1764.66
Obs. 1764.58
[bcp4-W8 +H +2AA]+
Calc. 1878.70
Obs. 1878.71
1500 1700 1900 2100 2300
1600 1800 2000 2200 2400
m/z

Figure 6

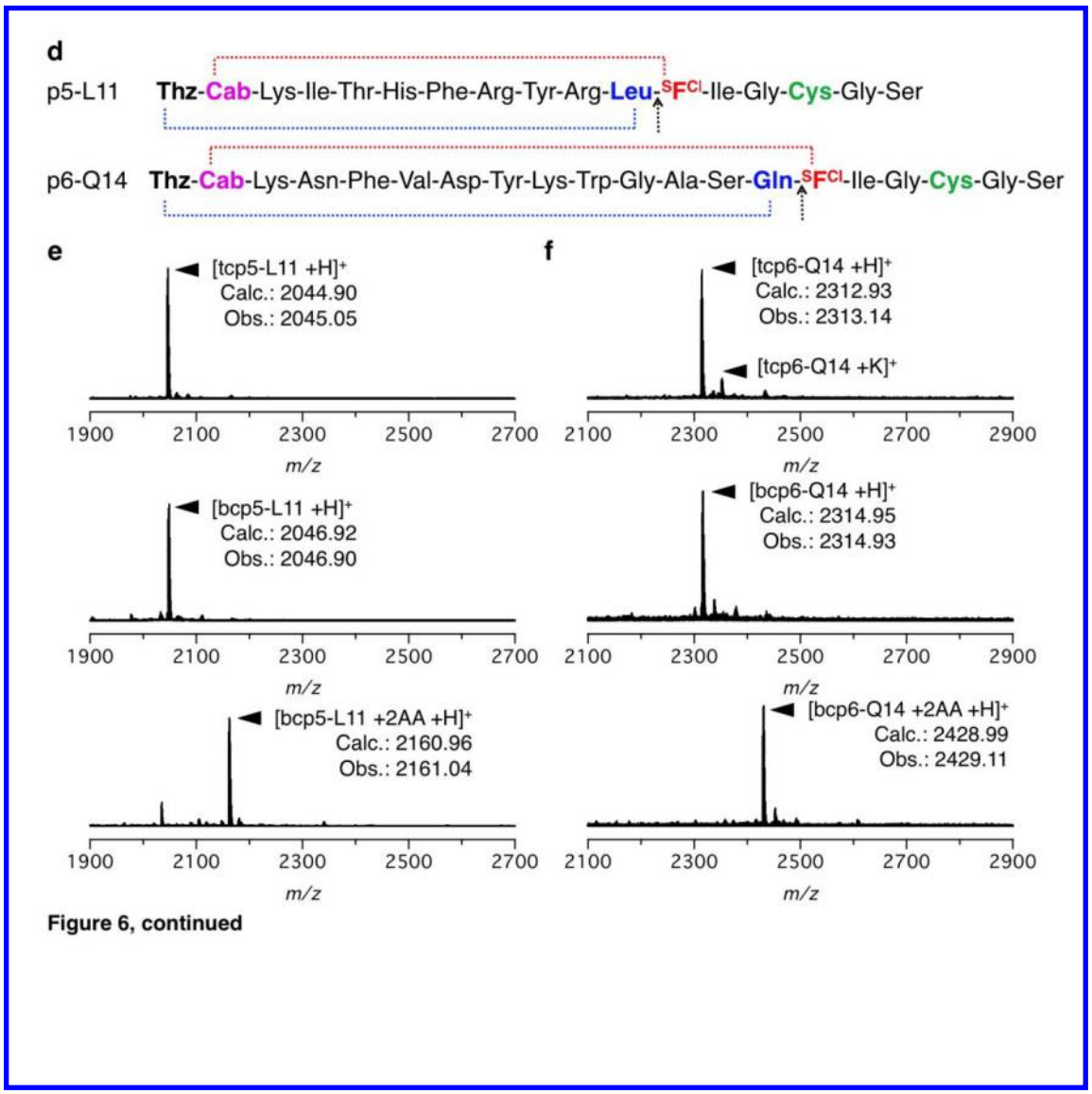

Figure 6, continued

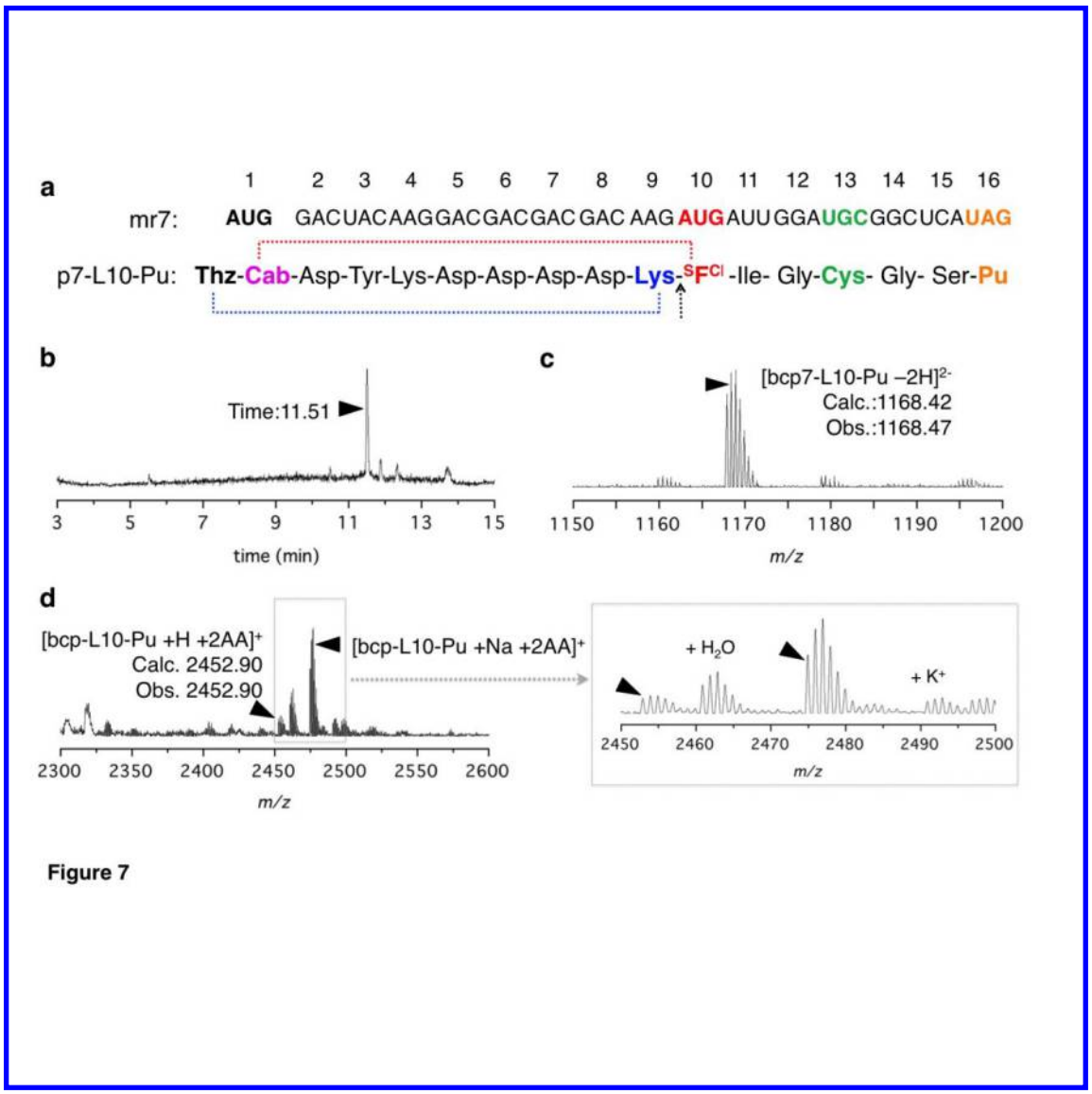
a
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
mr7: AUG GACUACAAGGACGACGACGAC AAG AUG AUU GGA UGC GGC UCA UAG
p7-L10-Pu: Thz-Cab-Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys-SFCl-Ile- Gly-Cys- Gly- Ser-Pu
b
Time:11.51
3 5 7 9 11 13 15
time (min)
c
[bcp7-L10-Pu –2H]2-
Calc.:1168.42
Obs.:1168.47
1150 1160 1170 1180 1190 1200
m/z
d
[bcp-L10-Pu +H +2AA]+
Calc. 2452.90
Obs. 2452.90
[bcp-L10-Pu +Na +2AA]+
2300 2350 2400 2450 2500 2550 2600
m/z
+ H2O
+ K+
2450 2460 2470 2480 2490 2500
m/z

Figure 7

Table of Contents (TOC) graphic