Published in final edited form as:

Methods Mol Biol. 2015 ; 1248: 23–38. doi:10.1007/978-1-4939-2020-4_2.

Synthesis of Macrocyclic Organo-Peptide Hybrids from Ribosomal Polypeptide Precursors via CuAAC-/hydrazide-mediated cyclization

Jessica M. Smith and Rudi Fasan

Department of Chemistry University of Rochester Rochester NY 14627

Abstract

Macrocyclic peptides have attracted increasing attention as a potential new source of chemical probes and therapeutics. In particular, their conformationally restricted structure combined with a high degree of functional and stereochemical complexity make them promising scaffolds for targeting biomolecules with high affinity and selectivity. The exploration of this structural class rely on the availability of efficient and versatile methods for the generation of large and diversified libraries of macrocyclic peptide-based molecules. To this end, we have developed a methodology for the synthesis of hybrid organo-peptide macrocycles via the cyclization of ribosomally derived polypeptide sequences with non-peptidic organic linkers. This strategy relies on the chemoselective and bioorthogonal ligation of azide/hydrazide-based “synthetic precursors” with intein-fused polypeptides harboring a side-chain alkyne functionality. This macrocyclization approach was found to proceed with high efficiency across a range of different target peptide sequences spanning 4 to 12 residues as well as across multiple mono- and di-aryl-based synthetic precursors. This versatility combined with the possibility to integrate non-proteinogenic scaffolds into genetically encoded peptide sequences makes this methodology of particularly high value toward the creation and screening of highly diverse libraries of peptide-based macrocycles.

Keywords

Macrocyclic peptide; Organo-peptide hybrids; Azide-alkyne Huisgen cycloaddition; Click chemistry; Intein; Intein-mediated ligation; Bioorthogonal ligation

1. Introduction

Owing to their inherent capability to interact with extended biomolecular interfaces, macrocyclic peptides have recently emerged as promising scaffolds for targeting historically challenging targets such as protein-protein and protein-nucleic acid interactions.^{1,2,3,4} By virtue of their restricted conformational flexibility, these compounds often exhibit higher affinity and selectivity to their targets as compared to their linear counterparts as a result of the reduced entropic costs upon binding and pre-organization into a bioactive conformation.^{5,6,7,8} In addition, backbone and/or side-chain cyclization have proved useful toward improving the proteolytic stability^{9,10} and/or membrane permeability^{11,12,13} of

peptide-based molecules. Not surprisingly, many biologically active peptide natural products harbor a ring topology (e.g. cyclosporine A (immunosuppressant), caspofungin (antifungal), and polymixin (antibiotic)).

In light of the potential advantages of this structural class, methods to access macrocyclic peptides, and combinatorial libraries thereof, have become of particularly high value.^{14,4} In particular, methods that rely on the cyclization of ribosomally derived peptides offer high combinatorial potential toward making vast collections of cyclic peptides (10^7 – 10^{12} members). Additionally, genetic encoding allows for coupling these libraries to various high-throughput screening platforms (e.g. phage, yeast, or mRNA display) as well as the rapid deconvolution of these libraries via sequencing of the peptide-encoding gene. As discussed in a recent review,¹⁵ these methods include the formation of cyclic peptides via split intein-catalyzed trans splicing,^{16,17} cross-linking of genetically encoded peptide sequences with amine- or cysteine-reactive reagents,^{18,19} cyclization of in vitro translated peptides containing unnatural residues,^{20, 21} or manipulation of biosynthetic machineries involved in the construction of naturally occurring cyclopeptides.^{22,23}

In an effort to provide new ways to access structurally diverse cyclic peptides, we have devised a chemobiosynthetic strategy for the synthesis of organo-peptide macrocycles via the chemo- and regioselective cyclization of genetically encoded peptide sequences with bifunctional synthetic agents.²⁴ As schematically described in Figure 1, this approach relies on the use of an engineered precursor protein (called “biosynthetic precursor” or “BP”) in which a variable peptide sequence (called “target sequence”) is framed between an N-terminal unnatural amino acid containing a side-chain alkyne functionality and a C-terminal intein protein (i.e. *Mxe* GyrA(N198A)). Owing to the N198A mutation, the latter is still able to catalyze the formation a reactive thioester bond at the C-terminal end of the target sequence but it is unable to undergo self-splicing from the precursor polypeptide. This arrangement equips the biosynthetic precursor with two reactive functionalities with orthogonal (and bioorthogonal) reactivity, namely the side-chain alkynyl group and the internal thioester linkage (Figure 1). The desired macrocyclic organo-peptide hybrids are then obtained upon reaction of such biosynthetic precursors with a bifunctional, azide/hydrazide-based “synthetic precursor” (“SP”) in the presence of Cu(I) as catalyst.²⁴ As determined by control experiments and mechanistic studies,²⁴ the macrocyclization process involves the formation of a protein-synthetic precursor 1,4-triazole adduct via Cu(I)-catalyzed Huisgen azide/alkyne cycloaddition (CuAAC) followed by attack of the SP-derived hydrazide onto the thioester linkage at the junction between the target peptide sequence and the intein, thus resulting in the formation of the desired macrocyclic product (Figure 1).

2. Materials

2.1 Reagents

- 1 Chemically competent BL21(DE3) cells
- 2 Ampicillin
- 3 Chloramphenicol

- 4 Sodium hydrogen phosphate (Na_2HPO_4)
- 5 Potassium dihydrogen phosphate (KH_2PO_4)
- 6 Sodium chloride (NaCl)
- 7 Ammonium chloride (NH_4Cl)
- 8 Luria Bertani (LB) media (per liter: 10 g tryptone, 5 g yeast extract, 10 g NaCl)
- 9 5x M9 salt solution (per liter: 64 g Na_2HPO_4 , 15 g KH_2PO_4 , 2.5 g NaCl , 5 g NH_4Cl)
- 10 Magnesium sulfate (MgSO_4)
- 12 Calcium chloride (CaCl_2)
- 13 Glycerol
- 14 *O*-propargyl tyrosine ($^{\text{OPG}}\text{Y}$)
- 15 Arabinose
- 16 Isopropyl β -D-1-thiogalactopyranoside (IPTG)
- 17 Tris-hydrochloride (Tris-HCl)
- 18 Imidazole
- 19 Nickel nitrilotriacetic acid (NTA) resin
- 20 Centrifugal filter units, 10 kDa cutoff
- 21 Copper sulfate (CuSO_4)
- 22 4-(bromomethyl)-benzoate
- 23 Sodium azide (NaN_3)
- 24 Hydrazine monohydrate
- 25 *tris*(2-carboxyethyl)phosphine (TCEP)
- 26 Sodium ascorbate
- 27 Ethylenediaminetetraacetic acid (EDTA)
- 28 Sodium borohydride (NaBH_4)
- 29 2,5-pyridinediacarboxylic acid dimethyl ester
- 30 *para*-Toluenesulfonyl chloride
- 31 Triethylamine (TEA)
- 32 (L)-*N*-*tert*-butoxycarbonyl-tyrosine
- 33 Potassium carbonate (K_2CO_3)
- 34 Propargyl bromide
- 35 Acetyl chloride

2.2 Solvents

1. N-N-Dimethylformamide (DMF)
2. Trifluoroacetic acid (TFA)
3. Acetonitrile
4. Ethanol (EtOH)
5. Methanol (MeOH)
6. Ethyl Acetate (EtOAc)
7. Tetrahydrofuran (THF)
8. Dichloromethane (DCM)
9. Diethyl ether

3. Methods

3.1 Expression of the precursor protein (“Biosynthetic Precursor”)

Precursor proteins for use within this method can be expressed in *E. coli* and purified via Ni-affinity chromatography using a polyhistidine (His₆) fused to the C-terminus of the intein. Suitable biosynthetic precursors consists of a N-terminal tail, followed by O-propargyl-Tyrosine (OpgY), followed by a variable peptide target sequence ((AA)_n, with n = 4 to 12 amino acids), followed by *Mxe* GyrA(N198A) fused to a C-terminal polyhistidine tag. The N-terminal tail can be as small as a Met-Gly dipeptide (Met is post-translationally cleaved during expression in *E. coli*), an oligopeptide (e.g. 4–10 amino acids), or a whole protein (e.g. chitin binding domain). Ribosomal incorporation of OpgY into the biosynthetic precursor is achieved via amber stop codon (TAG) suppression²⁵ using an engineered *Methanococcus jannaschii* tRNA / aminoacyl-tRNA synthetase (aaRS) pair evolved for recognition of OpgY²⁶.

Recombinant expression of the biosynthetic precursor is carried out using *E. coli* BL21(DE3) cells co-transformed with (1) a pET22-based plasmid vector encoding for the biosynthetic precursor protein under ITPG-inducible T7 promoter (“pBP” plasmid), and (2) a pEVOL vector²⁷ encoding for two copies of the OpgY-specific *Mj* aminoacyl-tRNA synthetase under a constitutive GlnS promoter and an arabinose-inducible AraC promoter, and a copy of the cognate *Mj* tRNA^{Tyr}_{CUA} (“pEVOL_OpgY plasmid”). *E. coli* cultures are normally grown in minimal (M9) media. However, we established that OpgY incorporation is tolerant to the presence of 1% yeast extract in the minimal medium, which helps increasing the expression yield of the biosynthetic precursor (30–40 mg/L culture cp. to 5–8 mg/L culture).

The following section describes a typical protocol for the expression of the precursor protein. In our original report,²⁴ the biosynthetic precursors were designed to contain a Thr residue at the last position of the target sequence preceding the intein ('I-I' site) to minimize the amount of premature intein splicing during recombinant expression of the protein in *E. coli*. In this case, less than 10% of the intein fusion protein undergoes *in vivo* splicing after

16–18 hour culture time at 27° C. However, we later determined that the majority (14/20) of the twenty natural amino acids can be installed at this site without incurring into significant premature intein splicing under these expression conditions.²⁸ Suboptimal residues at the 'I-1' site are Leu, Lys, Asp, His, Ile or Val. If required, decrease in the extent of premature splicing can be achieved by lowering the post-induction temperature (e.g. 20–25° C) and/or by shortening the culture time after induction of the protein expression (e.g. to 4–6 hours).

1. Add 1.0 µL of the pBP plasmid and 1.0 µL of the pEVOL_OpgY plasmid to a 100 µL aliquot of chemically competent BL21(DE3) cells. Flick the tube of cells gently several times to mix the DNA. Incubate in ice for 30 minutes.
2. Place the cells in a water bath at 42° C for 35 seconds. Remove and place back on ice for 5 minutes.
3. Add 300 µL of Luria Bertani (LB) media to the cells. Place at 37° C for 45 minutes.
4. Dilute 1 µL of the cell mixture into 100 µL LB and plate 10 µL of the diluted cells onto an LB agar plate with 50 µg/mL ampicillin and 30 µg/mL chloramphenicol. Grow for 14–17 h at 37° C until colonies are visible.
5. Take one colony from step 4 and place into 5 mL of LB media containing with 50 µg/mL ampicillin and 30 µg/mL chloramphenicol in an overnight culture tube with cap loose. Place in an incubated shaker for 16 hours at 37° C. After 16 hours of growth, the media should be opaque.
6. Prepare 500 mL M9 minimal media as follows (note, all solutions must be pre-sterilized through autoclaving or filter sterilizing, see Note 1): measure out 370 mL distilled water and add to 1 L sterile flask, add 100 mL M9 salt solution, 500 µL 2 M MgSO₄, 50 µL 1 M CaCl₂, 5 mL glycerol, 50 mL 10% yeast extract. Add ampicillin to 50 µg/mL concentration and chloramphenicol to 30 µg/mL concentration. Mix well.
7. Inoculate the 500 mL M9 culture with the 5 mL LB culture containing the cells from step 4. Place an incubated shaker at 37° C, shaking at 220 rpm, until OD₆₀₀ reaches approximately 0.6 (~ 3 hours).
8. Add 4.0 mL of 10% arabinose and 3.0 mL of 0.4 M *O*-propargyl tyrosine (suspension in water) to the culture. Place culture back in incubator at 27° C for 1 hour.
9. Add 0.5 mL of 0.5 M IPTG to the culture. Place culture back in incubator at 27° C for 16 hours.
10. Harvest the cells by centrifugation at 5000 x *g*. Resuspend cell pellet in a 50 mL tube with 25 mL A20 buffer (50 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, pH 7.4) and store cell suspension at –80° C.

¹All of the ingredients for the M9 minimal media must be made and autoclaved or filter sterilized separately first before combining. If reagents are mixed together and then autoclaved, the salts will be precipitate from solution.

3.2 Purification of the Biosynthetic Precursor

1. Thaw frozen cell suspension. Sonicate cell/buffer mixture to lyse cells. A typical method is 30 pulses, 1 second per pulse, and 10 seconds between pulses. This method is run three times with the cells on ice and temperature maintained below 18° C (see Note 2). From this point on, keep lysate at 4° C.
2. Centrifuge cell lysate at 18,000 x g for 30 minutes at 4° C.
3. Collect supernatant which should not be cloudy at this point. Pass the lysate over a column of 3 mL Ni-NTA beads (pre-equilibrated in A20 buffer, kept on ice) at a rate of ~ 3.0 mL/minute. Wash the column by flowing A20 buffer through at rate of 2.0 mL/minute for 25 minutes. To elute the His-tagged protein, flow A300 buffer (50 mM Tris-HCl, 150 mM NaCl, 300 mM imidazole, pH 7.4, kept on ice) through the column at a rate of 2.0 mL/minute. Collect the fractions containing protein and pool together (see Note 3).
4. Place the combined fractions in a centrifugal filter (MW cutoff: 10,000 Da). Perform a buffer exchange into a potassium phosphate (KPi) buffer (50 mM potassium phosphate, 150 mM NaCl, pH 7.4) using a centrifuge (4° C). First, spin down the sample which is in A300. Then, add ~ 12 mL of KPi buffer and centrifuge 25 minutes. Repeat this two more time.
5. Determine the protein concentration using a UV spectrometer. First, run a blank scan of KPi buffer in a quartz cuvette. Then, add the protein and scan from 350 to 250 nm. Using Beer's Law (Absorbance = molar absorptivity x pathlength (cm) x concentration (M)) to determine concentration of samples. Typical yield for a biosynthetic precursor is 30 mg/L culture.

3.3 Synthesis of the synthetic precursor SP_H1 (4-(azidomethyl)-benzohydrazide)

Several azide/hydrazide-containing synthetic precursors were found to be compatible with the present methodology, including mono-aryl SPs, in which the azido and hydrazido groups are in *ortho*, *meta*, and *para* to each other, as well as biaryl and di-aryl scaffolds.²⁴ Outlined below is a general method for making 4-(azidomethyl)-benzohydrazide from the commercially available *para*-(bromomethyl)benzoate (Figure 2).

1. Dissolve 1.0 g 4-(bromomethyl)benzoate into 15 ml DMF in a dry 50 ml round bottom flask with stir bar.
2. Add 0.851 g sodium azide to the solution. Place reflux condenser on flask and heat reaction to 70°C for 12 hours.
3. Let reaction cool back down to room temperature. Quench with 25 ml water and extract with ethyl acetate (3 x 20 ml).
4. Wash the organic fraction with 30 ml brine and dry over Na₂SO₄ for 10 minutes.

²If temperature is allowed to go above 18° C for an extended period of time, premature intein splicing of the biosynthetic precursor may occur.

³It is recommended to not freeze the protein before buffer exchange as the high concentration of imidazole can cause protein precipitation after freezing/thawing.

5. Concentrate *in vacuo* to yield a yellow oil. Proceed to the following step without purification.
6. Dissolve 0.83 g 4-(azidomethyl)benzoate (product from step 5) in 14 ml EtOH.
7. Add 3.2 ml of 20.6 M hydrazine monohydrate. Place reflux condenser on flask and heat reaction to 80° C for 5 hours.
8. Concentrate *in vacuo* to remove excess ethanol. Add 30 ml water and extract the product with ethyl acetate (5 x 30 ml) (see Note 4).
9. Dry organic fraction over Na₂SO₄ and concentrate *in vacuo* to yield a white solid.
10. Recrystallize product using a 2:1 ratio of ethanol:ethylacetate to yield 4-(azidomethyl)-benzohydrazide.

Analytical Data for 4-(azidomethyl)-benzohydrazide: ¹H NMR (400 MHz, (CD₃)₂SO) δ 4.48 (s, 2H), 7.41 (d, *J* = 8.0 Hz, 2H), 7.81 (d, *J* = 8.0 Hz, 2H), 9.77 (br s, 1H); ¹³C NMR (100 MHz, (CD₃)₂SO) δ 53.6, 127.8, 128.7, 133.5, 139.1, 166.0; MS (ESI) calcd for C₈H₉N₅O [M+H]⁺: *m/z* 192.2; found: 192.1. Yield over two steps: 80%.

3.4 Synthesis of the synthetic precursor SP_H2 (6-(azidomethyl)nicotinohydrazide)

The pyridine-based synthetic precursor SP_H2 (Figure 2) was designed to enable the synthesis of macrocyclic organo-peptide hybrids using lower concentration of copper catalyst. The pyridine ring is believed to coordinate Cu(I) during catalysis, enabling CuAAC ligations to proceed with high efficiency at lower copper catalyst concentrations.²⁹ As described in Figure 2, SP_H2 can be prepared from commercially available 2,5-pyridine-dicarboxylic acid dimethyl ester in a few steps. The first step, conversion of the diester to the monoalcohol is low yielding (~30%) due largely to conversion to the diol. The other steps are highly efficient, with isolated yields >90–95%.

1. Place 1.5 g 2,5-pyridine-dicarboxylic acid dimethyl ester and 3.6 g CaCl₂ (anhydrous) into a 100 mL round bottom flask. Add 15 mL anhydrous THF and 30 mL anhydrous methanol and let stir at room temperature for 10 minutes. Reaction will be a white slurry.
2. Cool the mixture in an ice bath for 10 minutes, then add 0.79 g sodium borohydride slowly (over ~10 minutes) to the mixture. Let react 3 hours at 0° C.
3. Let warm to room temperature. Quench with 30 mL ice-cold water. Extract product with 3 x 60 mL chloroform, dry with magnesium sulfate and remove solvent under reduced pressure. Product is a white solid but may appear as a viscous oil if enough of the di-alcohol product is extracted with chloroform.
4. Dissolve product from step 3 (~400 mg) in 50 mL DCM. Add 0.68 g para-toluenesulfonyl chloride and 1.6 mL triethylamine and let reaction stir for 2–3 hours.

⁴It is recommended to test the EtOAc layers during extraction by TLC to ensure all of the product has been recovered from the water layer as the product is very polar.

5. Remove volatiles under reduced pressure. Add 25 mL THF to the crude mixture (does not fully dissolve) and add 1.5 g NaN₃. Let reaction stir at room temperature for 20 hours.
6. Dilute reaction with 30 mL water. Extract product with 3 x 20 mL ethyl acetate, dry over sodium sulfate, and concentrate under reduced pressure.
7. Purify compound from step 6 by silica chromatography with 8:1 hexanes:ethyl acetate. Product is a yellow oil.
8. Take 50 mg of product from step 7 and place in 2 mL ethanol. Add 450 µL of 20.6 M hydrazine monohydrate and heat to 70 °C for 5 hours.
9. Let the reaction cool back down to room temperature. Add 6 mL of water and extract product with 5 x 4 mL ethyl acetate. Wash with brine and then dry with sodium sulfate. Remove volatiles under reduced pressure to afford the product as a yellow-white solid. If salts are isolated with the product, add 3 mL methanol to product mixture. The pyridine product will dissolve, leaving the salts behind.

Analytical data for 6-(azidomethyl)nicotinohydrazide: ¹H NMR (500 MHz, CD₃ OD) δ 4.55 (s, 2H), 7.55 (d, *J* = 8.2 Hz, 1H), 8.20 (dd, *J* = 8.2, 2.0 Hz, 1H), 8.92 (s, 1H); ¹³C NMR (125 MHz, CD₃OD) δ 54.4, 121.8, 128.2, 136.1, 147.6, 158.6, 165.5. Isolated yield over four steps: 20%.

3.5 Synthesis of O-propargyl-tyrosine

As outlined in Figure 3, the unnatural amino acid OpgY can be prepared in three steps from commercially available N-Boc-tyrosine according to the following procedure, which is an improved version of a previously reported protocol.³⁰

1. Place 6.0 g (L)-*N*-*tert*-butoxycarbonyl-tyrosine and 9.0 g potassium carbonate into a 100 mL round bottom flask equipped with stir bar.
2. Add 30 mL anhydrous DMF and 6.3 mL propargyl bromide and stir at room temperature for 20 hours.
3. Quench the reaction with 150 mL of water. Extract the product using diethyl ether (2 x 100 mL) and dry the combined organic layers with magnesium sulfate. Remove solvent under reduced pressure to yield the dialkylated product as a yellow oil.
4. Add 21 mL acetyl chloride to 180 mL methanol and cool in ice bath. Add the crude product from step 3 and stir for 4 hours while warming to room temperature. Remove volatiles under reduced pressure to yield the Boc-protected amino acid precursor.
5. Mix 42 mL of 2 N NaOH and 30 mL methanol. Add the crude product from step 4 and stir for 2 hours. After TLC confirms complete hydrolysis, adjust the pH to 7.0 using concentrated HCl. Stir the mixture overnight at 4° C.

6. The final product will precipitate from solution after overnight incubation. Filter and wash the precipitate with cold water, then dry overnight under reduced pressure to yield the final *O*-propargyl tyrosine as an off-white powder.

Analytical data for *O*-propargyl tyrosine: ^1H NMR (400 MHz, D_2O) δ 2.78 (s, 2H), 2.94 (dd, J = 6.8, 22.4 Hz, 1 H), 3.08 (dd, J = 9.6, 20 Hz, 1 H), 3.81 (dd, J = 2.0, 12.8 Hz, 1 H), 6.92 (d, J = 8.8 Hz, 2 H), 7.13 (d, J = 8.4 Hz, 2 H); ^{13}C NMR (100 MHz, D_2O) δ 35.4, 56.0, 76.6, 78.7, 115.6, 128.5, 130.6, 156.1, 173.9. MS (ESI) calcd for $\text{C}_{12}\text{H}_{13}\text{NO}_3$ $[\text{M}+\text{H}]^+$: m/z 220.1; found: 220.3. Isolated yield over three steps: 72%.

3.6 Standard Reaction Conditions for Macrocyclization

To generate the hybrid macrocycles, the biosynthetic precursor protein is made to react with the bifunctional azide/hydrazide synthetic precursor in phosphate buffer at physiological pH in the presence of Cu(I) generated *in situ* from copper sulfate (CuSO_4) and sodium ascorbate. The reaction can be monitored by MALDI-TOF mass spectrometry and LC-MS as exemplified by the representative reaction between synthetic precursor $\text{SP}_{\text{H}1}$ and biosynthetic precursor MG6 (MG-(OpgY)-TGSgt-GyrA) in Figure 4. Depending on the target peptide sequence, some acyclic product 'a' may form as shown in Figure 1. This by-product derives from the hydrolysis of the biosynthetic precursor-synthetic precursor (BP-SP) adduct as shown in Figure 1. Hydrolysis of the precursor protein, leading to the by-product 'h' (Figure 1), may also occur but this is rarely observed. In the following, a typical protocol for the synthesis of organo-peptide macrocycles at an analytical scale is provided.

A modification of this protocol is also included in which the pyridine-based synthetic precursor $\text{SP}_{\text{H}2}$ is used. $\text{SP}_{\text{H}2}$ allows for the use of a significantly (4-fold) lower concentration of the Cu(I) catalyst (i.e. 250 μM vs. 1 mM) for the CuAAC coupling, which can be advantageous in various settings (e.g. macrocyclization in cell lysate or in the context of a cellular/viral display system).

1. Dissolve the biosynthetic precursor to a concentration of 100 μM in KPi buffer (typical reaction volume is 40 μL). Add in the synthetic precursor $\text{SP}_{\text{H}1}$ or $\text{SP}_{\text{H}2}$ (pre-dissolved as a 0.3 M solution in DMSO) to a final concentration of 50 mM.
2. Add sodium ascorbate to a concentration of 5 mM, and lastly, add CuSO_4 to a final concentration of 1 mM (if using $\text{SP}_{\text{H}2}$, add CuSO_4 to 0.25 mM). Mix gently by pipetting and leave at room temperature. Note: care must be taken to not add more than the given amount of copper as protein precipitation may occur. In addition, copper sulfate should be added as the final reagent and mixed immediately to prevent any localized precipitation.
3. After 15 minutes, add TCEP to a final concentration of 10 mM and EDTA to a final concentration of 30 mM. These additives are useful to facilitate copper removal and prevent cysteine oxidation.
4. After 5 minutes, add 80 μL KPi buffer to reaction. Remove copper ions, excess synthetic precursor, and sodium ascorbate via fast desalting chromatography (removal of Cu ions is particularly important for obtaining high quality MALDI-TOF MS spectra). Place ~100 μL of gel filtration resin in a fritted microcentrifuge

column, and insert the column into a 1.5 mL microcentrifuge tube. Equilibrate the resin in KPi buffer by washing with 3 x 500 μ L KPi and centrifuging at 1500 x *g* for one minute each. Apply sample to the resin and let sit for one minute. Then centrifuge at 1500 x *g* for two minutes to elute the sample. At this point, quantitative conversion to the triazole cycloaddition product occurs with either SP as evidenced by MS analysis (step 5).

5. To analyze products via LC-MS, dilute protein reaction to 10 μ M in 90% acetonitrile in water. Use a C₄ column to separate protein from the peptide products (e.g. 100 x 4.6 mm, 3 μ m C₄ column).
6. Use a binary mobile phase system consisting of A: water + 0.1% formic acid and B: acetonitrile + 0.1% formic acid. Perform a 12 minute run at a flow rate of 0.5 ml/min with a linear gradient from 5 to 95% of B. If the sample is not desalted prior to LCMS analysis, the first 3 minutes of the run should be diverted to waste to protect MS from salts in the sample as this can significantly diminish signal intensity.
7. Determine the extent of splicing of the BP-SP adducts after overnight incubation based on the corresponding LC-MS chromatograms using deconvolution software for mass determination of the high molecular weight analytes (BP-SP adduct, spliced GyrA) and peak assignment. Additionally, protein splicing can be determined by SDS-PAGE when macrocycle is > 1,500 Da.
8. Analyze the peptide products by performing MS-MS and generating the corresponding ion extract chromatograms (Figure 4) based on the m/z ($M+H^+$) of the macrocyclic product 'm' and the acyclic product 'a' (Figure 1). The chromatograms can be integrated and the area-under-the-curve (AUC) used to estimate the relative abundance of the two species (see Note 5).

3.7 Large-Scale Macrocyclization of MG6 and Purification

The macrocyclization reaction can be scaled up to yield milligram amounts of the desired macrocycle, as needed for further characterization studies (e.g. binding affinity, proteolytic stability, and cell permeability). It is important to keep the concentration of protein and copper ions as low as stated in the protocol below, or precipitation of the protein may occur. Alternatively, the pyridine-based SP_H2 can be used.

After macrocyclization, the spliced intein by-product (and uncyclized protein precursor, if present) can be removed by passing the reaction mixture through a Ni-NTA resin. The macrocycle can be further purified via solid-phase extraction followed by HPLC purification. This procedure allows for the efficient separation of the desired macrocyclic product from potential by-products such as (a) the branched acyclic peptide 'a' (Figure 1) resulting from hydrolysis of the SP-BP adduct after CuAAC coupling, and (b) the linear peptide 'h' (Figure 1), resulting from hydrolysis of the intein from the precursor protein. In

⁵For many of the macrocycles we analyzed by MS-MS, there is usually a second broad peak in the ion extract chromatogram whose retention time corresponds to the GyrA intein. This is caused by unrelated multicharged ions from the spliced GyrA protein and can be ignored.

practice, the latter byproduct is rarely observed as the CuAAC coupling proceeds with high efficiency (>90–100%) regardless of the composition and length of the target sequence (and N-terminal tail).²⁴ The amount of the branched acyclic peptide by-product was found to vary from 0 to about 50% of the total product amount depending on the length of the target peptide sequence (highest yields were achieved with 5 amino acid-long sequences) and type of synthetic precursor (highest yields were achieved with azide/hydrazide spacing distance being 7 Å).²⁴

Below, a typical protocol for the large-scale synthesis and isolation of an organo-peptide macrocycle generated from SP_H1 and the precursor protein MG6 (MG-(OpgY)-TGSGT-GyrA) is described:

1. Add 1.0 ml of a 1.3 mM solution of biosynthetic precursor MG6 (in KPi) to 8.0 ml KPi buffer in a test tube with cap. To this solution, add 38 mg (500 µL) of SP_H1 (dissolved in DMSO), and 125 µL of a 400 mM sodium ascorbate solution. Lastly, add 120 µL of 100 mM CuSO₄ and immediately invert the tube several times (gently) to mix.
2. Let incubate at room temperature for 20 minutes with gentle shaking. Solution will be slightly cloudy.
3. Add 10 ml KPi buffer with 30 mM EDTA. Let mix for 5 minutes.
4. Split volume by 4 (5 ml each) and place into 4 centrifugal filters to remove copper ions, TCEP, and EDTA. Centrifuge at 4000 x g for 25 minutes.
5. Repeat wash/filter 3 times with KPi buffer alone (no EDTA). Final combined volume should be raised to 10 ml by adding KPi after filtrations and TCEP to a final concentration of 3 mM. Let reactions incubate at room temperature for two days.
6. Add Ni-NTA resin and incubate 5 hr at room temperature with gentle mixing to bind the His-tagged intein protein. Centrifuge and pour off supernatant (separate from beads). Lyophilize solution.
7. Resuspend lyophilized solution in 500 µL water and load onto a C18 solid phase extraction column (after pre-equilibrating the column with methanol first and then water). MOrPH product elutes from the column at 50% acetonitrile in water. Check that the correct product has eluted using MALDI-TOF MS. Lyophilize elution.
8. Dissolve lyophilized peptide mixture in 75% water, 25% acetonitrile, 0.1% TFA solution. Purify MOrPH using HPLC. The MOrPH is purified on a C18 column with a binary mobile phase system consisting of A: water + 0.1% TFA and B: acetonitrile + 0.1% TFA. Perform a 10 minute run at a flow rate of 1.0 ml/min at 5% A for 1 minute, then with a linear gradient from 5 to 60% of B. For the MG6-SP_H1 MOrPH, the product elutes at 5.2 minute retention time.

Acknowledgments

This work was supported by the U.S. National Science Foundation grant CHE-1112342 awarded to R.F.. J.M.S. is grateful to the NSF Graduate Research Fellowship Program for financial support. MS instrumentation was supported by the U.S. National Science Foundation grants CHE-0840410 and CHE-0946653.

References

1. Robinson JA, Demarco S, Gombert F, Moehle K, Obrecht D. The design, structures and therapeutic potential of protein epitope mimetics. *Drug Discov Today*. 2008; 13:944–51. [PubMed: 18725320]
2. Driggers EM, Hale SP, Lee J, Terrett NK. The exploration of macrocycles for drug discovery--an underexploited structural class. *Nat Rev Drug Discov*. 2008; 7:608–24. [PubMed: 18591981]
3. Marsault E, Peterson ML. Macrocycles Are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery. *J Med Chem*. 2011; 54:1961–2004. [PubMed: 21381769]
4. Smith JM, Frost JR, Fasan R. Emerging strategies to access peptide macrocycles from genetically encoded polypeptides. *J Org Chem*. 2013; 78:3525–31. [PubMed: 23517465]
5. Hayashi Y, Morimoto J, Suga H. In vitro selection of anti-Akt2 thioether-macrocyclic peptides leading to isoform-selective inhibitors. *ACS Chem Biol*. 2012; 7:607–13. [PubMed: 22273180]
6. 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*. 2012; 51:3423–7.
7. Dias RL, Fasan R, Moehle K, Renard A, Obrecht D, Robinson JA. Protein ligand design: from phage display to synthetic protein epitope mimetics in human antibody Fc-binding peptidomimetics. *J Am Chem Soc*. 2006; 128:2726–2732. [PubMed: 16492060]
8. Cardoso RM, Brunel FM, Ferguson S, Zwick M, Burton DR, Dawson PE, Wilson IA. Structural basis of enhanced binding of extended and helically constrained peptide epitopes of the broadly neutralizing HIV-1 antibody 4E10. *J Mol Bio*. 2007; 365:1533–44. [PubMed: 17125793]
9. Fairlie DP, Tyndall JD, Reid RC, Wong AK, Abbenante G, Scanlon MJ, March DR, Bergman DA, Chai CL, Burkett BA. Conformational selection of inhibitors and substrates by proteolytic enzymes: implications for drug design and polypeptide processing. *J Med Chem*. 2000; 43:1271–81. [PubMed: 10753465]
10. Satoh T, Li S, Friedman TM, Wiaderkiewicz R, Korngold R, Huang Z. Synthetic peptides derived from the fourth domain of CD4 antagonize off function and inhibit T cell activation. *Biochem Biophys Res Commun*. 1996; 224:438–43. [PubMed: 8702407]
11. Walensky LD, Kung AL, Escher I, Malia TJ, Barbuto S, Wright RD, Wagner G, Verdine GL, Korsmeyer SJ. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science*. 2004; 305:1466–70. [PubMed: 15353804]
12. Rezai T, Yu B, Millhauser GL, Jacobson MP, Lokey RS. Testing the conformational hypothesis of passive membrane permeability using synthetic cyclic peptide diastereomers. *J Am Chem Soc*. 2006; 128:2510–1. [PubMed: 16492015]
13. Gudmundsson OS, Vander Velde DG, Jois SD, Bak A, Siahaan TJ, Borchardt RT. The effect of conformation of the acyloxyalkoxy-based cyclic prodrugs of opioid peptides on their membrane permeability. *J Pept Res*. 1999; 53:403–13. [PubMed: 10406218]
14. White CJ, Yudin AK. Contemporary strategies for peptide macrocyclization. *Nat Chem*. 2011; 3:509–24. [PubMed: 21697871]
15. Frost JR, Smith JM, Fasan R. Design, synthesis, and diversification of ribosomally derived peptide macrocycles. *Curr Opin Struct Biol*. 2013; 23:571–80. [PubMed: 23856642]
16. Scott CP, Abel-Santos E, Wall M, Wahn DC, Benkovic SJ. Production of cyclic peptides and proteins in vivo. *Proc Natl Acad Sci U S A*. 1999; 96:13638–13643. [PubMed: 10570125]
17. Tavassoli A, Lu Q, Gam J, Pan H, Benkovic SJ, Cohen SN. Inhibition of HIV budding by a genetically selected cyclic peptide targeting the Gag-TSG101 interaction. *ACS Chem Biol*. 2008; 3:757–64. [PubMed: 19053244]

18. Millward SW, Takahashi TT, Roberts RW. A general route for post-translational cyclization of mRNA display libraries. *J Am Chem Soc.* 2005; 127:14142–3. [PubMed: 16218582]
19. Schlippe YV, Hartman MC, Josephson K, Szostak JW. In vitro selection of highly modified cyclic peptides that act as tight binding inhibitors. *J Am Chem Soc.* 2012; 134:10469–77. [PubMed: 22428867]
20. Hipolito CJ, Suga H. Ribosomal production and in vitro selection of natural product-like peptidomimetics: the FIT and RaPID systems. *Curr Opin Chem Biol.* 2012; 16:196–203. [PubMed: 22401851]
21. Hofmann FT, Szostak JW, Seebeck FP. In vitro selection of functional lantipeptides. *J Am Chem Soc.* 2012; 134:8038–41. [PubMed: 22545861]
22. Levensgood MR, Knerr PJ, Oman TJ, van der Donk WA. In vitro mutasynthesis of lantibiotic analogues containing nonproteinogenic amino acids. *J Am Chem Soc.* 2009; 131:12024–5. [PubMed: 19655738]
23. Maksimov MO, Pan SJ, Link AJ. Lasso peptides: structure, function, biosynthesis, and engineering. *Nat Prod Rep.* 2012; 29:996–1006. [PubMed: 22833149]
24. Smith JM, Vitali F, Archer SA, Fasan R. Modular Assembly of Macrocyclic Organo-Peptide Hybrids Using Synthetic and Genetically Encoded Precursors. *Angew Chem Int Ed.* 2011; 50:5075–5080.
25. Liu CC, Schultz PG. Adding new chemistries to the genetic code. *Annu Rev Biochem.* 2010; 79:413–444. [PubMed: 20307192]
26. Deiters A, Schultz PG. In vivo incorporation of an alkyne into proteins in *Escherichia coli*. *Bioorg Med Chem Lett.* 2005; 15:1521–4. [PubMed: 15713420]
27. Young TS, Ahmad I, Yin JA, Schultz PG. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J Mol Bio.* 2010; 395:361–74. [PubMed: 19852970]
28. Frost JR, Vitali F, Jacob NT, Brown MD, Fasan R. Macrocyclization of Organo-Peptide Hybrids through a Dual Bio-orthogonal Ligation: Insights from Structure-Reactivity Studies. *Chembiochem.* 2013; 14:147–60. [PubMed: 23203912]
29. Uttamapinant C, Tangpeerachaikul A, Grecian S, Clarke S, Singh U, Slade P, Gee KR, Ting AY. Fast, cell-compatible click chemistry with copper-chelating azides for biomolecular labeling. *Angew Chem Int Ed.* 2012; 51:5852–6.
30. Deiters A, Cropp TA, Mukherji M, Chin JW, Anderson JC, Schultz PG. Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*. *J Am Chem Soc.* 2003; 125:11782–3. [PubMed: 14505376]

Figure 1.

Strategy for synthesis of macrocyclic organo-peptide hybrids (**'m'**) via chemoselective cyclization of intein-fusion precursor polypeptides with azide/hydrazide synthetic precursors (blue). The unnatural amino acid O-propargyl tyrosine is highlighted in red. R' is a variable N-terminal tail. Possible side-products of the reaction are also indicated (dotted lines), which can derive from hydrolysis of the biosynthetic precursor (**'h'**) or SP-BP adduct (**'a'**).

Figure 2.
Synthetic routes for the preparation of synthetic precursors SP_H1 (A) and SP_H2 (B).

Figure 3.
Synthesis of O-propargyl tyrosine from N-Boc-tyrosine.

Figure 4. Monitoring and analysis of macrocyclization reaction between synthetic precursor SP_H1 and precursor protein MG6 (MG-(OpgY)-TGSYGT-GyrA). (A) MALDI-TOF MS analysis of biosynthetic precursor MG6 before the reaction (left panel), after formation of the BP-SP_H1 adduct via CuAAC (central panel), and after macrocyclization to release the GyrA intein (right panel). c: calculated mass, o: observed mass. (B) LC-MS ion extract chromatogram and MS-MS spectrum of the macrocyclic product (**1**) formed upon reaction of MG6 with SP_H1. (C) LC-MS ion extract chromatogram and MS-MS spectrum of the minor 'acyclic' by-product (**2**, = 'a') formed upon hydrolysis of the intein from the BP-SP_H1 adduct.