Published in final edited form as:

Biopolymers. 2010 ; 94(4): 405–413. doi:10.1002/bip.21486.

Alternative Chemistries for the Synthesis of Thrombospondin-1 Type 1 Repeats

Theresa K. Tiefenbrunn, Juan Blanco-Canosa, and Philip E. Dawson*

The Departments of Cell Biology and Chemistry, 10550, N. Torrey Pines Road, The Scripps Research Institute, La Jolla, CA 92037

Introduction

The synthesis of a novel protein target can be technically demanding due to the large number of peptide functional groups and numerous sequence-specific chemical and conformational challenges that may be encountered. The assembly of polypeptides by native chemical ligation (NCL) is typically robust when peptides are solubilized using one of a range of chaotropic agents, detergents, or organic cosolvents. In contrast, the synthesis of the peptide fragments can be subject to challenges inherent in solid-phase peptide synthesis (SPPS) as well as limitations specific to the generation of peptides containing C-terminal thioesters or N-terminal cysteines. Herein we describe a series of synthetic approaches for the total chemical synthesis of the second type 1 repeat of thrombospondin-1 (**TSR2**) that vary in both ligation site and α -amine protecting group strategy (Boc/Fmoc). The results demonstrate the chemistry and context-dependence of challenging sequences, as well as how such challenges can be overcome by altering the chemistry used for SPPS, the choice of ligation site, and the resin used as a solid support.

The first synthetic route to the ~7 kD TSR2 domain involved a two-piece ligation of fragments synthesized using Boc/Bzl solid phase peptide synthesis (Boc-SPPS)¹. However, despite attempts to improve synthetic yields by changing the ligation site and resins used for the synthesis, efforts to scale up to produce protein for biophysical analysis were hampered by low purification yields of peptide fragments. As a result, alternative synthetic routes to these ligation fragments were explored, including the use of Fmoc/tBu SPPS for both N- and C-terminal fragments. Compared to Boc-SPPS, Fmoc-based SPPS is more widely practiced, avoids the need for hydrofluoric acid (HF), and is more compatible with automation. The use of Fmoc-SPPS for the synthesis of peptide precursors for NCL has been encouraged by the development of new approaches for the synthesis of thioesters by Fmoc-SPPS^{2–5}. We report the first application of the *N*-acylurea (Nbz) activation strategy to the synthesis of a protein domain.

The second type 1 repeat from human thrombospondin-1 is a representative TSR domain with known structure⁶ and activity⁷. While the folded domain is structurally defined as residues Q416-I473, the leading KRFLK sequence has been implicated in the activation of TGF- β ⁸ and adds positively charged residues that improve the solubility of unprotected peptide intermediates¹. The synthetic target described in this work, mTSR2, comprises residues K412-I473 of human thrombospondin-1 with Met451Nle and Asn452Gly and amidated termini, as previously described¹.

*To whom correspondence should be addressed: Philip E. Dawson, Department of Cell Biology and of Chemistry, The Skaggs Institute of Chemical Biology, La Jolla, CA 92037, Ph: (858) 784-7015, dawson@scripps.edu.

A two-piece ligation strategy was designed to facilitate modifications to one half of the protein without requiring resynthesis of the entire protein. Inspection of the peptide sequence revealed that, of the six cysteines in the protein, three were centrally located and had good potential as ligation sites¹. In order to determine which ligation site would be the most useful for the synthesis of multiple TSR analogs, three distinct pairs of N and C-terminal peptides were synthesized (Figure 1).

Results

Fmoc-Based SPPS of Ligation Fragments

Synthetic access to five of the six peptides, **2–6**, corresponding to the three potential ligation sites **A–C** for TSR2 (Figure 1) was obtained. Synthesis of peptide **1a**, the longest N-terminal fragment (ligation site **A**), encountered obstacles that were not easily overcome, leading us to focus on the two alternative ligation sites, **B** and **C**, presented in the molecule. The most robust synthetic access was established using N- and C-terminal fragments corresponding to ligation site **B** and a combination of Boc and Fmoc-SPPS.

C-terminal Peptide Fragments

C-terminal fragments **2**, **4**, and **6** were synthesized on a 0.1 mmol scale on either ChemMATRIX resin (0.6 meq/g) with a Rink amide linker or PAL-PEG-PS (Peptide Amide Linker-Polyethylene glycol-polystyrene, 0.2 meq/g). Peptides were assembled manually or on a C.S. Bio CS336X synthesizer using HCTU/HOBt/DIEA activation (10-fold excess amino acid/HCTU/HOBt and 15-fold excess DIEA to resin) and 20 % 4-methyl-piperidine in DMF for deprotection (3 × 10 min for automated syntheses; 2 × 15 min for manual syntheses unless otherwise noted). Peptide resins were cleaved with TFA:TIS:H₂O:EDT 94:1:2.5:2.5 (v/v) for 2 hours at room temperature, and then precipitated in diethyl ether.

Several challenges were encountered during the synthesis of the C-terminal fragments by Fmoc-SPPS. Synthesis of the shorter C-terminal fragment **2** on either PAL-PEG-PS or ChemMATRIX resin yielded the desired product and an impurity of ~10% corresponding to a mass of -129 Da, or deletion of a Glu or Lys residue. A sample reserved for analysis at Fmoc-TSP-1(462-473)-NH₂, or Fmoc-ETKACKKDACPI, revealed a deletion of 351 Da corresponding to the terminal Fmoc-Glu, suggesting that coupling of Fmoc-Glu rather than deprotection of the preceding residue was the synthetic problem. In order to limit the aggregation leading to this deletion, the Fmoc-Glu(OtBu)-Thr($\psi^{\text{Me,Me}}_{\text{pro}}$)-OH pseudoproline dimer was incorporated at this position. With this pseudoproline dimer and synthesis on ChemMATRIX resin, **2** was obtained in 90% analytical purity (by integration of the main product peak) and 40% recovered yield after HPLC purification.

For the synthesis of the longer peptide **4**, a second pseudoproline dimer was incorporated at Ile438-Thr439. Additionally, 4-methyl-piperidine treatment times were reduced to 2 × 5 minutes after incorporation of Asp435 to reduce potential aspartimide formation at Asp435-Gly436. Synthesis on Rink-ChemMATRIX resin yielded the desired product in 80% analytical purity (by integration of the main product peak) (Figure 2a), while an identical synthesis on PAL-PEG-PS resin led to an impure mixture containing only ~20% desired product (Figure 2b). Synthesis of the slightly longer **6** was achieved in 70% analytical purity when synthesized on ChemMATRIX resin with incorporation of the two pseudoproline residues previously described. Only minor amounts of -18 Da product resulting from aspartimide formation were observed.

N-terminal Peptide Fragments

N-terminal fragments were synthesized as *N*-acyl-benzimidazole (Nbz) thioester precursors². In this approach, a peptide is synthesized on a diaminobenzoic acid (Dbz) linker **7**, which is then treated with *p*-nitrochloroformate **11** and base to yield the Nbz linker **13** after peptide elongation (Scheme 1). Under ligation conditions, the Nbz moiety exchanges with a thioester to generate a reactive C-terminal thioester peptide. One concern with this approach was the acetylation of the N-terminus in the presence of a free aromatic amine in the Dbz linker. Although this amine is resistant to acylation by activated Fmoc amino acids, it is reactive towards nitrophenylchloroformate for activation and would likely be modified by standard acetic anhydride/DMF acetylation cocktails. Atkinson et. al. describe mild acetylation conditions that are selective for primary amines in the presence of more hindered secondary amines using *N*-Acetyl-*N*-acyl-3-aminoquinazolinones⁹, for example commercially-available DAAQ (**9**), which we reasoned would have selectivity for the N-terminus over the Dbz linker (Scheme 1).

An initial synthesis of the N-terminal fragment **3b** corresponding to ligation site **B** led to a mixture of the desired product with a significant –18 Da product and several products produced from incomplete loading of the C-terminal Thr onto the Dbz moiety, as well as a minor amount of unacetylated product, resulting from unoptimized conditions for acetylation. The –18 Da product was attributed to aspartimide formation at the Asp-Gly sequence (417-418). In order to improve synthetic yields, the backbone protected Fmoc-Asp(OtBu)-(Dmb)Gly-OH dipeptide was incorporated for positions 417-418 to eliminate aspartimide formation. Loading of the hindered Fmoc-Thr-OH onto the Dbz linker was performed using a double coupling with HATU as an activating agent in place of the previously reported conditions utilizing HBTU². Finally, the acetylation conditions were modified from an overnight treatment with a 1.5-fold excess of DAAQ to a 30-minute treatment with a 5-fold excess of DAAQ to resin. With these modifications, the desired product **3b** was obtained in 45% purity (by peak integration)(Figure 3). Similar results were obtained for synthesis of the shorter peptide **5b**.

A Difficult Sequence in Several Contexts

The attempted synthesis of the six potential ligation fragments corresponding to three ligation sites for mTSR2 led to the identification of a difficult sequence near the middle of the domain, ⁴²⁹Cys-Ser-Val-Thr-Cys-Gly-Asp-Gly⁴³⁶. Under conditions of Fmoc-SPPS, difficulties in chain assembly were observed for this sequence regardless of its position relative to the resin. When this sequence is distant from the resin, as for peptide **6**, synthesis on PAL-PEG-PS resin with standard amino acids leads to a complex mixture of deletion products with a near complete truncation of the final Cys-Val-Ser-Thr sequence (Supporting Information, Figure S1). As described above for the synthesis of the slightly shorter peptide **4**, the use of aggregation-disrupting pseudoproline residues at positions Glu⁴⁶²-Thr⁴⁶³ and Ile⁴³⁸-Thr⁴³⁹ in combination with the fully PEGylated ChemMATRIX resin can overcome these chain assembly problems.

When this sequence is located closer to the resin, as for N-terminal fragment **1b**, its synthesis is even more challenging. In an initial test synthesis of Fmoc-CSVTGDGVITRIRL-NH₂, on PAL-PEG-PS resin, only ~20% was the desired product, and the remainder of peptide was a variety of deletion and truncation products (Supporting Information, Figure S2). To improve this synthesis, PAL-PEG-PS resin was substituted with ChemMATRIX resin, Ile⁴³⁸-Thr⁴³⁹ was incorporated as the pseudoproline dimer, Fmoc-IleThr(ψ^{Me,Me}pro)-OH, and Asp⁴³⁵-Gly⁴³⁶ was incorporated as Fmoc-Asp(OtBu)-(Dmb)-Gly-OH. Unfortunately, these aggregation-disrupting techniques were not sufficient to solve the chain assembly problems, and a sample from a synthesis of Fmoc-CSVTGDGVITRIRL-

NH₂ revealed a mixture of four products resulting from deletions within the CSVT region (Supporting Information, Figure S3). As we had previously synthesized fragment **1a** by Boc-SPPS with good results¹, and an alternative synthetic route to the desired protein domain had already been established, no further attempts were made to optimize the synthesis of this fragment by Fmoc chemistry.

Native Chemical Ligation

After purification of individual fragments, native chemical ligation¹⁰ was used to form the full length TSR2 domain. Conditions for native chemical ligation depended on whether the N-terminal fragment was synthesized as a mercaptopropionic acid (Mpa) thioester by Boc chemistry or as an acyl urea thioester precursor (Nbz) by Fmoc chemistry. Using standard native chemical ligation conditions^{11,12} of 1% thiophenol/1% benzyl mercaptan in 6 M GnHCl, 200 mM sodium phosphate buffer adjusted to pH ~7 and Boc-synthesized Mpa thioester fragments (~2 mM), ligations were complete in 18 hours at 37° C (e.g. Figure 4a) consistent with the poorly soluble thiophenol catalyst producing trace quantities of aryl thioester. For the Fmoc Nbz N-terminal peptides, the more soluble thiol, 4-mercaptophenylacetic acid (MPAA) was used to ensure rapid conversion of the Nbz group quantitatively to an aryl thioester¹³ that is more resistant to hydrolysis. The ligation reaction of **3b** with **4** (2 mM) was complete in 1–2 hours at room temperature in 6M GnHCl, 200 mM sodium phosphate, 200 mM MPAA and 20 mM tris-carboxyethylphosphine (TCEP), pH 7.5 (Figure 4b). Ligation reactions of Boc-synthesized Mpa thioester fragments under these conditions (200 mM MPAA/20mM TCEP in 6M GnHCl, pH 7.5 buffer) are complete in a similar time frame; however, standard NCL conditions utilizing 1% thiophenol/benzylmercaptan were preferred for these ligations due to challenges associated with removing the large excess of MPAA from the ligated protein product. Thiophenol/benzylmercaptan is a convenient thiol additive since the concentration of thiol in solution is only ~0.02 mM and the disulfide products precipitate, allowing relatively simple HPLC separation from peptide products.

Folding and Purification

Ligation products were purified by RP-HPLC and lyophilized. Initial attempts at folding were performed in 10 mM Tris, 150 mM NaCl, pH 7.4 buffer with a unusually oxidizing 3:1 ratio of cystine to cysteine (3 mM/1 mM), as previously reported¹. However, under these conditions, it was observed that aggregation was a significant competing process, even at protein concentrations of 0.1 mg/mL. A number of folding conditions were screened, including varying the redox ratio from 1:5 oxidized: reduced up to 3:1 oxidized: reduced, using the cystine/cysteine or oxidized/reduced glutathione (GSH/GSSG) pair, and several different buffers and pH conditions. A 3:1 ratio of oxidized: reduced redox pair led to the highest yields of folded product, with lower ratios leading to a mix of products containing only one or two disulfide bonds. In order to improve folding methods over those previously reported, the cysteine/cystine pair was substituted with the more soluble reduced/oxidized glutathione (GSH/GSSG) pair and several buffer additives were tested. The use of GSH/GSSG improved recovery yields, but some aggregation was still observed. While the use of the osmolyte TMAO¹⁴ did not lead to significant improvements, it was observed that the use of a buffer containing 100mM Tris•HCl, adjusted to pH 7.4, significantly decreased aggregation and improved folding recovery yields after preparative HPLC from below 10% to greater than 50%. Folding was observed as a shift in HPLC retention time to a single peak (Figure 5) and confirmed by far UV circular dichroism (Figure 6a). Misfolded TSR2 species, i.e. those that eluted at a later time than folded TSR2, do not have the characteristic CD spectra of the folded product (Figure 6b). HPLC purification performed on a 4.6 mm × 150 mm column provided superior yields to that performed at higher flow rates on a 10 mm × 250 mm column, even for protein loads of up to 3 mg. Lyophilized protein could be

dissolved in a variety of buffers for subsequent biophysical analysis. Concentrated stock solutions were used since the small size of the protein made concentration using ultrafiltration membranes impractical.

Discussion

Boc and Fmoc Approaches to Thioester Peptides

Synthetic yields for the N-terminal fragment are comparable for Boc-SPPS using an MpaK thioester linker on ChemMATRIX resin and Fmoc-SPPS using a Dbz thioester precursor on ChemMATRIX resin. Synthesis by Boc-SPPS is faster, with an average coupling cycle (including deprotection) lasting <25 minutes, while our average automated Fmoc cycle takes about an hour. For thioester fragments, Boc-SPPS also has the advantage that the mercaptopropionic acid linker is commercially available (Peptides International), while the diaminobenzoic acid linker for Fmoc-SPPS must be synthesized. Because of these advantages, Boc-SPPS was chosen as the preferred method to obtain N-terminal peptide fragments for future analogs of mTSR2. However, these fragments are also accessible by Fmoc-SPPS, which will be particularly useful for the future generation of acid sensitive glycoforms of the TSR2 domain.

Boc and Fmoc Approaches to N-terminal Cysteine Peptides

Synthetic yields for the C-terminal fragments are significantly higher for an optimized synthesis using Fmoc-SPPS than were achieved using standard Boc-SPPS conditions for chain assembly on polystyrene-based resin¹. The optimized synthesis of this fragment by Fmoc-SPPS includes the use of pseudoproline dimers at positions Glu⁴⁶²-Thr⁴⁶³ and Ile⁴³⁸-Thr⁴³⁹, as well as the use of ChemMATRIX resin. The use of pseudoproline residues in combination with ChemMATRIX resin has previously been successfully applied to the synthesis of chemokines¹⁵. Synthesis of C-terminal fragments by Boc-SPPS on ChemMATRIX resin was not attempted, although it is postulated that this could lead a reduction in side products during HF cleavage. However, the use of ChemMATRIX resin for Boc-SPPS is not without drawbacks, including the need for larger volumes of TFA due to the intense swelling of ChemMATRIX in acid and the need for either double couplings or longer coupling times¹.

Overcoming a Difficult Sequence in Multiple Contexts

The availability of several potential ligation sites in a protein provides the option to choose a ligation site corresponding to the most synthetically accessible peptide fragments and avoid the need to synthesize particularly challenging peptides. For the TSR2 domain, the choice of ligation site **B** limits challenges associated with synthesis of the problematic ⁴²⁹Cys-Ser-Val-Thr-Cys-Gly-Asp-Gly⁴³⁶ sequence. Interestingly, this sequence presents unique challenges depending on the α -amine protecting group strategy used for synthesis. In the context of the N-terminal ⁴¹²Lys-TSP-1(413-443)-COR peptide **1** (R = MpaK for Boc synthesis (**1a**), R = Nbz for Fmoc synthesis (**1b**), R = NH₂ for Fmoc test synthesis), this sequence falls within the aggregation prone region of 8–15 amino acids from the resin.

For Fmoc-SPPS, the challenges associated with this sequence arise during chain elongation. In the context of the N-terminal fragment **1b**, where the ⁴²⁹Cys-Ser-Val-Thr-Cys-Gly-Asp-Gly⁴³⁶ sequence comprises residues 8–16 from the resin, it causes aggregation and collapse of the synthesis on PAL-PEG-PS resin in the absence of any thioester linker. A similar collapse is seen for the synthesis of the C-terminal peptide **6** on PAL-PEG-PS resin when this sequence is farther from the resin. In the latter case, the synthesis can be rescued by replacing PAL-PEG-PS resin with ChemMATRIX resin containing a Rink amide linker and incorporating two aggregation-disrupting pseudoproline residues, at Glu⁴⁶²-Thr⁴⁶³ and

Ile438-Thr439. However, for synthesis of the N-terminal ⁴¹²Lys-TSP-1(413-443)-COR peptide, significant deletions are observed even with the use of a pseudoproline residue at Ile438-Thr439 and synthesis on ChemMATRIX resin.

It is interesting to note that the chain assembly of this sequence did not present a challenge during chain assembly using *in situ* neutralization Boc-SPPS on polystyrene resin as developed by Kent¹⁶. However, this procedure did encounter a significant challenge at the ⁴²⁹Cys-Ser-Val-Thr-Cys sequence during HF cleavage. An unusual protecting group migration between the benzyl group on threonine and the 4-methylbenzyl group on the adjacent cysteine was observed¹. Replacement of Thr(Bzl) with unprotected threonine mitigated formation of this side product. However, synthetic yields for the longer C-terminal fragments by Boc chemistry were limited by the accumulation of minor side products throughout the synthesis. These results suggest that the CSVT sequence has unusual conformational and chemical properties that complicate synthesis in a variety of contexts.

Alternative Resins

For all of the syntheses described, the use of ChemMATRIX resin proved to be important for obtaining good product yields. For Fmoc-SPPS, the improvements seen with ChemMATRIX resin (0.6 meq/g) compared to a PEGylated polystyrene resin, PAL-PEG-PS (0.2 meq/g), are remarkable (Figure 2). These improvements are complementary to the improvements seen from including aggregation-disrupting pseudoproline residues, and together allow the synthesis of a difficult sequence that is unattainable under other conditions. For Boc-SPPS, moderate improvements seen with ChemMATRIX resin (0.6 meq/g) compared to standard MBHA resin (0.59 meq/g) do not appear to be due to improvements in chain assembly, but rather due to increased access of scavengers to the peptide-resin during HF cleavage¹. In fact, the use of ChemMATRIX resin was as effective in limiting the formation of tryptophan adducts arising during HF cleavage as protecting the three tryptophan residues with formyl groups¹.

Folding and Purification

A final challenge in developing a robust synthesis of mTSR2 was identifying folding and purification conditions that do not lead to significant product loss due to either misfolding or aggregation. A screen of a number of conditions, including ratios of oxidized: reduced glutathione and a variety of buffer conditions, led to the identification of 3:1 GSSG:GSH in 100 mM Tris•HCl, pH 7.4 as improved folding conditions. The use of such a highly oxidizing environment for protein folding has been previously described for conotoxins¹⁷, and the use of 100 mM Tris•HCl has been previously shown to improve protein folding yields¹⁸.

Conclusions

Synthesis of the full-length mTSR2 domain is possible by a variety of synthetic routes utilizing either Boc or Fmoc-SPPS. The preferred synthetic route involves a native chemical ligation between an N-terminal fragment **3a**, ⁴¹²Lys-TSP-1(413-432)-COSR, synthesized by Boc-SPPS and a C-terminal fragment **4**, ⁴³³Cys-TSP-1(434-473)-NH₂ synthesized by Fmoc-SPPS, followed by oxidative folding using a 3:1 ratio of oxidized: reduced glutathione in 100 mM Tris buffer at pH 7.5. While the N-terminal peptide fragments have comparable synthetic yields using either Boc or Fmoc-SPPS, Boc-SPPS with *in situ* neutralization provides for rapid and straightforward chain assembly without the need for backbone protection. In contrast, synthesis of peptide fragments by Fmoc-SPPS was only successful under optimized conditions including the use of ChemMATRIX resin and pseudoproline residues. For the synthesis of mTSR2, Boc-SPPS and Fmoc-SPPS provide complementary

synthetic routes to different ligation fragments. Robust synthetic access to the mTSR2 domain has been established, and investigations are underway using synthetic mTSR2 domains containing nonnatural amino acids to probe the structure of this unique domain.

Materials and Methods

Materials

Standard Boc amino acids were obtained from C. S. Bio. Standard Fmoc amino acids, MBHA resin, HCTU and HOBt were obtained from Peptides International. Fmoc-Glu(OtBu)-Thr($\psi^{\text{Me,Me}}$ pro)-OH, Fmoc-Ile-Thr($\psi^{\text{Me,Me}}$ pro)-OH, Fmoc-Asp(OtBu)-(Dmb)-Gly-OH, and Fmoc-Rink linker were purchased from Novabiochem. Fmoc-Dbz was obtained from Juan Blanco-Canosa. Diisopropylcarbodiimide (DIC), diisopropylethylamine (DIEA), and N,N-diacetylaminoquinazolinone were obtained from Sigma. 4-Methylpiperidine was from Alfa Aesar. Dimethylformamide (DMF) and dichloromethane (CH_2Cl_2) were from EMD Chemicals. Acetonitrile (CH_3CN) was from J. T. Baker or Fisher. PAL-PEG-PS resin and ninhydrin reagents are from Applied Biosciences. ChemMATRIX resin is from Matrix International.

Boc-Based SPPS

Detailed methods have been previously reported¹. In brief, all peptides were synthesized by manual SPPS using standard Boc protection and HCTU activation with in situ neutralization¹⁶ on a 0.1–0.2 mmole scale. Protecting groups were as follows: Arg(Tos), Asn(Xan), Asp(OcHex), Cys(4-MeBzl), Glu(OcHex), Gln(Xan), His(Dnp), Lys(ClZ), Ser(Bzl), and Thr(Bzl). Tryptophan was used without a side chain protecting group. For thioester containing peptides, S-trityl-mercaptopropionic acid was coupled following a lysine linker. The S-trityl group was removed with 2.5% TIS/2.5% H_2O /95% TFA and the C-terminal amino acid was coupled using standard conditions. For N-terminal fragments, 10% anisole was incubated with resin for 10 minutes before HF cleavage. For C-terminal fragments, 0.2% ethanedithiol (EDT) was included along with 10% anisole during HF cleavage.

Fmoc-Based SPPS

Standard Fmoc amino acids, HCTU and HOBt were obtained from Peptides International. Protecting groups were as follows: Asp(OtBu), Cys(Trt), Glu(OtBu), His(Trt), Lys(Boc), Asn(Trt), Gln(Trt), Arg(Pbf), Ser(tBu), Thr(tBu), Trp(Boc). Fmoc-Glu(OtBu)-Thr($\psi^{\text{Me,Me}}$ pro)-OH, Fmoc-Ile-Thr($\psi^{\text{Me,Me}}$ pro)-OH, Fmoc-Asp(OtBu)-(Dmb)-Gly-OH, and Fmoc-Rink linker were purchased from Novabiochem. Fmoc-Dbz was obtained from Juan Blanco-Canosa. DIC, DIEA, and N,N-diacetylaminoquinazolinone were obtained from Sigma. 4-Methylpiperidine was from Alfa Aesar. DMF and CH_2Cl_2 were from EMD Chemicals. Acetonitrile was from J. T. Baker or Fisher. PAL-PEG-PS resin and ninhydrin reagents are from Applied Biosciences. ChemMATRIX resin is from Matrix International.

Synthesis was performed at 0.1 or 0.05 mmol scale. When ChemMATRIX resin was used, the first step was manual coupling of Fmoc-Rink linker, using 3-fold excess of linker and DIC/HOBt activation. For Fmoc-PAL-PEG-PS resin, this step was omitted. Automated synthesis was carried out on a C. S. Bio CS336X peptide synthesizer with 1.0 mmol amino acid (10 or 20-fold excess) activated with 1 eq 0.4 M HCTU/HOBt in DMF and 1.5 eq 1 M DIEA in DMF. Coupling times were 30 minutes. Fmoc deprotection was carried out with 20% 4-Methylpiperidine in DMF for ~20 minutes.

All synthesis was carried out on the synthesizer except for residues 429–439 of peptide **6** and residues 433–439 of peptide **4** and the exceptions noted below. For residues 429–435 and

residues 433-435 for syntheses where Fmoc-Asp(OtBu)-(Dmb)-Gly-OH was not incorporated at 435-436, 4-Methylpiperidine treatments were limited to 2×5 minutes. For all other manual steps, 4-Methylpiperidine treatment was 2×15 minutes.

Fmoc-Glu(OtBu)-Thr($\psi^{\text{Me,Me}}$ pro)-OH, Fmoc-Ile-Thr($\psi^{\text{Me,Me}}$ pro)-OH, and Fmoc-Asp(OtBu)-(Dmb)-Gly-OH were coupled manually using 3-fold excess of amino acid and DIC/HOBt amino acid activation. If a ninhydrin test showed incomplete coupling after 1 hr, the couplings were repeated. For residues incorporated after a pseudoproline dimer, 1 mmol of amino acid was activated with HATU, and again the coupling was repeated after an hour if necessary. Residues incorporated after Fmoc-Asp(OtBu)-(Dmb)-Gly-OH could be incorporated via standard HCTU coupling conditions.

After synthesis, C-terminal peptide-resins were treated with a cocktail of 2.5% TIS, 2.5% H₂O, 4% EDT in 91% TFA for 2 hours at room temperature. Resin was removed by filtration and excess TFA solution was evaporated on rotovap. Peptides were precipitated and washed twice with ice cold diethyl ether, dissolved in 27% CH₃CN, H₂O, 0.1% TFA and lyophilized.

For the synthesis of N-terminal thioester precursor peptides, the first five couplings were performed manually, as described¹⁹. Briefly, 0.05 mmol of ChemMATRIX resin was swollen in DMF. Fmoc-Rink linker was coupled as above. After Fmoc removal with 20% 4-methylpiperidine in DMF, (2 $\times$ 15 min), 0.20 mmol of 3-Fmoc-4-diaminobenzoic acid was activated with 1 eq 0.2 M HBTU/HOBt (1:1) in DMF and 1.5 eq 0.2 M DIEA in DMF and added to the resin. Subsequent amino acid coupling steps were performed manually with 0.5 mmol amino acid (10-fold excess) activated with 1 eq HBTU/HOBt (0.2 M in DMF) and 1.5 eq DIEA (0.2 M in DMF). After the first five couplings, resin was placed on the synthesizer and the synthesis was completed following standard coupling cycles using 0.4 M HBTU/HOBt in place of HCTU/HOBt. After the final Fmoc group was removed, the resin was acetylated by treatment with 5 eq of DAAQ in CH₂Cl₂/DMF for 1 hour.

Native Chemical Ligation

The conditions used for native chemical ligation depend on the C-terminus of the N-terminal peptide fragment. For alkyl thioester peptides synthesized by Boc methods, ligation reactions were carried out using 1 % thiophenol and 1 % benzylmercaptan. Equimolar amounts of each fragment were dissolved to ~10 mg/mL in 6 M GnHCl, 200 mM NaH₂PO₄, pH 8.0 buffer (Final pH after thiol addition is ~7). Thiols were added and the ligation was allowed to occur overnight (16–18 hrs) at 37 °C. For Nbz-peptides, ligation reactions were carried out using 200 mM mercaptophenylacetic acid and 20 mM TCEP in 6 M GnHCl, 200 mM sodium phosphate buffer, adjusted to pH ~7, and were complete within 1 hour at room temperature.

Folding and Purification

Ligation products were purified by RP-HPLC and lyophilized. For folding trials, 0.1 mg of peptide was dissolved in 1 mL of pH 7.4 buffer and folding was monitored by RP-HPLC. For larger-scale protein folding, peptide was dissolved to 0.1–0.2 mg/mL in 100 mM Tris•HCl, pH 7.4 buffer with a 3:1 ratio of oxidized: reduced glutathione. HPLC purification was performed by loading dilute samples over 5–15 minutes at 1 mL/min on a 4.6 mm column equilibrated at 20% B (90% CH₃CN/0.1% TFA/H₂O), and then eluting with a gradient of 28%–32% B over 16 minutes. Major peaks eluted before minor impurities, and corresponded to folded protein as determined by circular dichroism, with disulfide bond formation confirmed by mass spectrometry.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Tiefenbrunn TK, Dawson PE. *Protein Sci.* 2009; 18:970–979. [PubMed: 19384999]
2. Blanco-Canosa JB, Dawson PE. *Angew Chem Int Ed.* 2008; 47:6851–6855.
3. Tofteng AP, Sorensen KK, Conde-Frieboes KW, Hoeg-Jensen T, Jensen K. *J Angew Chem Int Ed.* 2009; 48:7411–7414.
4. Tsuda S, Shigenaga A, Bando K, Otaka A. *Organic Letters.* 2009; 11:823–826. [PubMed: 19161341]
5. Nakamura, Ki; Kanao, T.; Uesugi, T.; Hara, T.; Sato, T.; Kawakami, T.; Aimoto, S. *Journal of Peptide Science.* 2009; 15:731–737. [PubMed: 19735084]
6. Tan K, Duquette M, Liu JH, Dong Y, Zhang R, Joachimiak A, Lawler J, Wang JH. *J Cell Biol.* 2002; 159:373–382. [PubMed: 12391027]
7. Lawler J. *J Cell Mol Med.* 2002; 6:1–12. [PubMed: 12003665]
8. Miao WM, Seng WL, Duquette M, Lawler P, Laus C, Lawler J. *Cancer Research.* 2001; 61:7830–7839. [PubMed: 11691800]
9. Atkinson RS, Barker E, Sutcliffe M. *J Chem Commun.* 1996:1051–1052.
10. Dawson PE, Muir TW, Clark-Lewis I, Kent SB. *Science.* 1994; 266:776–779. [PubMed: 7973629]
11. Dawson PE. *Methods Enzymol.* 1997; 287:34–45. [PubMed: 9330313]
12. Dawson PE, Churchill MJ, Ghadiri MR, Kent SBH. *Journal of the American Chemical Society.* 1997; 119:4325–4329.
13. Johnson ECB, Kent SBH. *Journal of the American Chemical Society.* 2006; 128:6640–6646. [PubMed: 16704265]
14. Doan-Nguyen V, Loria JP. *Protein Sci.* 2007; 16:20–29. [PubMed: 17123958]
15. Garca-Martn F, White P, Steinauer R, Ct S, Tulla-Puche J, Albericio F. *Peptide Science.* 2006; 84:566–575. [PubMed: 16810664]
16. Schnolzer M, Alewood P, Jones A, Alewood D, Kent SB. *Int J Pept Res Ther.* 2007; 13:31–44.
17. Price-Carter M, Gray WR, Goldenberg DP. *Biochemistry.* 1996; 35:15537–15546. [PubMed: 8952508]
18. Rudolph R, Lilie H. *Faseb J.* 1996; 10:49–56. [PubMed: 8566547]
19. Blanco-Canosa JB, Dawson PE. *Angew Chem Int Ed.* 2008; 47:6851–6855.

Figure 1.

Sequence of mTSR2 and fragments thereof, corresponding to the three potential ligation sites for TSR2 (A, B, and C). Z represents norleucine. X represents either a thioester (1a, 3a, 5a) or *N*-acyl-benzimidazole (Nbz) thioester precursor (1b, 3b, 5b). Nbz peptides contain a mixture of the two isomers shown.

Figure 2.

Figure 2a. Analytical RP-HPLC analysis of peptide **4** on ChemMATRIX resin with Ile-Thr and Glu-Thr pseudoproline. Product yield was 80% by integration of main product peak. Gradient is 10–50% B over 30 minutes.

Figure 2b. Analytical RP-HPLC of the synthesis of peptide **4** on PAL-PEG-PS resin with Ile-Thr and Glu-Thr pseudoproline. Very little of the resulting peptide was the desired product. Gradient is 10–50% B over 30 minutes.

Figure 3. HPLC showing synthesis of peptide **3b**. Both major peaks correspond to Nbz isomers of the desired product. Gradient is 0–70% B over 30 minutes.

Figure 4.

Figure 4a. Ligation of **3a** with **4** in 6M GnHCl, 200mM phosphate, pH ~7, 1% v/v thiophenol and benzylmercaptan, at 18 hours. Gradient is 0–67% B over 30 minutes. Figure 4b. Ligation of **3b** with **4** in 6M GnHCl, 200 mM phosphate, pH 7.1 with 20 mM TCEP and 200 mM MPAA, at 2 hours. Gradient is 0% for 10 minutes, then 0–50% B over 50 minutes; the only peak eluting in the first 30 minutes corresponds to unoxidized MPAA (not shown).

Figure 5. Analytical RP-HPLC analysis of the folding of native TSR2. Unfolded product (top) elutes at a later time than folded product (bottom). Gradient is 10–50% B over 30 minutes.

Figure 6.

Figure 6a. Far UV CD spectra of folded native TSR2. Three independently folded samples are shown.

Figure 6b. Far UV CD spectra for a sample corresponding to “misfolded” TSR2, a product that elutes later on HPLC than correctly folded product.

Scheme 1. Synthetic strategy for Fmoc-SPPS of thioester peptides. Following SPPS, aminoacid 1 undergoes specific acylation and cyclization to yield the resin-bound

Scheme 1.

Scheme for the synthesis of thioester precursor peptides by Fmoc chemistry. The Dbz linker **7** is added to a standard Fmoc linker and the peptide is synthesized by Fmoc SPPS. DAAQ (**9**) is used as mild acetylation agent for the N-terminus. Prior to cleavage and deprotection, Dbz **10** is activated to Nbz **13**, which exchanges with thiols under NCL conditions. Schematic is only shown for one of the two possible isomers of Nbz.