Efficient Semisynthesis of a Tetraphosphorylated Analogue of the Type I TGFβ Receptor

ARTICLE *in* ORGANIC LETTERS · DECEMBER 2001

Impact Factor: 6.36 · DOI: 10.1021/ol016859i

6 AUTHORS, INCLUDING:

Robert R. Flavell
University of California, San Francisco
12 PUBLICATIONS **225** CITATIONS
SEE PROFILE

Michelle Trester-Zedlitz
University of California, San Francisco
14 PUBLICATIONS **899** CITATIONS
SEE PROFILE

ORGANIC LETTERS

2002
Vol. 4, No. 2
165−168

Efficient Semisynthesis of a Tetraphosphorylated Analogue of the Type I TGFβ Receptor

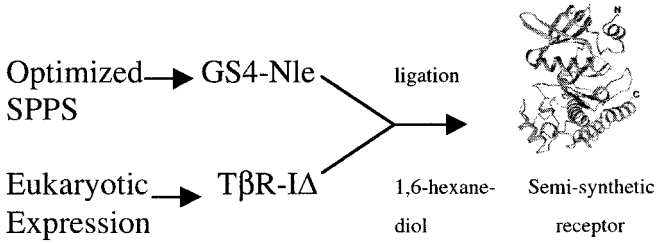
Robert R. Flavell,[†] Morgan Huse,[†] Mike Goger,[†] Michelle Trester-Zedlitz,[†] John Kuriyan,[‡] and Tom W. Muir*,[†]

Laboratory of Synthetic Protein Chemistry, and Laboratory of Molecular Biophysics, The Rockefeller University, New York, New York 10021

muirt@rockvax.rockefeller.edu

Received October 4, 2001

ABSTRACT



Semisynthesis of an active tetraphosphorylated analogue of the Type I TGFβ receptor is reported. An efficient native chemical ligation protocol was developed to link a tetraphosphopeptide and a recombinant receptor fragment. Synthesis of the peptide α-thioester on a 4-sulfamylbutyryl resin was optimized following the characterization of a major side reaction and subsequent substitution of norleucine for methionine in the peptide sequence. These optimized protocols will be applicable to the semisynthesis of related protein kinases.

Protein phosphorylation plays an important role in signal transduction pathways in both prokaryotes and eukaryotes. One example of considerable biomedical interest is the transforming growth factor β (TGFβ) signaling pathway.[1] TGFβ signaling plays a prominent role in the development, homeostasis, and repair of virtually all tissues in organisms from fruit fly to man, and mutations in the pathway are associated with a variety of human cancers.[1] The cytokine TGFβ transduces signals via a receptor complex composed of two transmembrane serine-threonine kinases, the TGFβ receptors I and II (TβR-I and TβR-II). Oligomerization induced by TGFβ binding causes TβR-II to multiply phosphorylate TβR-I within a highly conserved N-terminal regulatory region called the GS region, named for the ^{185}TTSGSGSG^{192} sequence within it. Activated TβR-I then phosphorylates members of the Smad family of transcription factors, which carry the signal to the nucleus. It is thought that at least four of the serines and threonines within the TTSGSGSG motif must be phosphorylated to fully activate TβR-I.[2]

To fully understand the molecular mechanisms underlying signal transduction processes such as the TGFβ signaling pathway, it is necessary to have access to chemically defined phosphoproteins for biochemical and biophysical study. This requirement is complicated by the inability to genetically encode a phospho-amino acid using standard protein expression technologies. One solution to this problem is to use protein ligation strategies in which a synthetic phosphopeptide segment is chemically ligated to a recombinant polypeptide corresponding to the remainder of the target

[†] Laboratory of Synthetic Protein Chemistry.
[‡] Laboratory of Molecular Biophysics.

(1) Massague, J. *Ann. Rev. Biochem.* **1998**, *67*, 753−91.

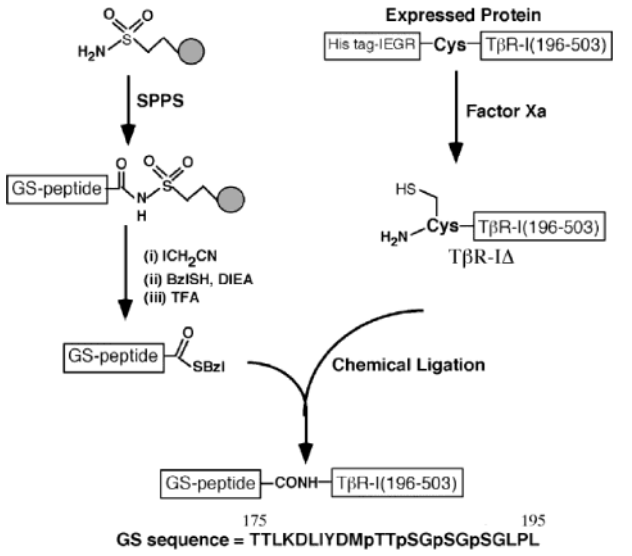
(2) Wieser, R.; Wrana, J. L.; Massague, J. *EMBO J.* **1995**, *14*, 2199−2208.

10.1021/ol016859i CCC: $22.00
Published on Web 12/21/2001

protein.[3] Indeed, we recently reported the semisynthesis of a hyperphosphorylated version of TβR-I containing four phospho-amino acids within the GS region.[4] Although this synthesis provided homogeneous material for biochemical analysis,[5] the yields obtained were rather low and insufficient material was isolated for detailed structural studies. Two major problems were encountered in the original synthesis, namely, the generation of a major side-product during the solid-phase peptide synthesis (SPPS) of the phosphopeptide and nonspecific protein aggregation during the ligation reaction. In the present study, both of the issues are addressed, allowing the final yield of active semisynthetic receptor to be greatly improved.

In the previous study, native chemical ligation[6] was employed to link together a synthetic peptide α-thioester, corresponding to the tetraphosphorylated N-terminal GS fragment, and a recombinant C-terminal fragment possessing the requisite N-terminal cysteine for ligation (ref 4 and Scheme 1). The phosphopeptide (hereafter referred to as GS4) was synthesized on the commercially available 4-sulfamylbutyryl AM "safety catch" resin. This resin allows the synthesis of α-thioester peptides using the Fmoc-SPPS strategy; the sulfonamide linkage is stable to the basic conditions used for N^{α} deprotection but can be activated to nucleophilic cleavage by thiols upon N-alkylation with agents such as iodoacetonitrile or trimethylsilyldiazomethane (TMS-CHN_2).[7] Importantly, this strategy offers a synthetic route to α-thioester phosphopeptides for use in protein ligation. In the case of GS4, activation of the linker was achieved using iodoacetonitrile since treatment of the protected peptide-resin with TMS-CHN_2 resulted in apparent O-methylation of the benzyl protected phosphothreonine and phosphoserine residues (data not shown). Analysis of the crude GS4 peptide after thiolysis with benzyl mercaptan and deprotection with TFA revealed the presence of a major side-product of mass 25 Da greater than expected (hereafter referred to as GS4*). The amount of this adduct varied from cleavage to cleavage but even in the best case the ratio of GS4 to GS4* was approximately 1:1 (Figure 2A).[8]

Scheme 1. Semisynthesis of TβR-I(4P); Tetraphosphorylated GS Sequence Shown Below



To improve the yield and reproducibility of the GS4 synthesis, we first sought to characterize the GS4* side-product. Treatment of GS4* with calf intestinal phosphatase resulted in a mass reduction of 320 Da as would be expected for the removal of four phosphate groups. This result rules out the possibility that the modification is on one of the phosphate groups. Digestion of GS4* with AspN endoproteinase followed by LC-MS analysis of the resulting peptides localized the +25 modification to the C-terminal DMpTTpSGpSGpSGLPL-SBzl region of the peptide. Consistent with the mapping data, N-terminal sequencing of GS4* indicated that an unknown amino acid was present at the 10th residue where methionine was expected. From these experiments we concluded that the +25 Da modification in GS4* was most likely located on the single methionine residue in the sequence.

To further characterize this +25 Da modification, a model dipeptide, Fmoc-Met-Gly-SBzl, was synthesized on the 4-sulfamylbutyryl AM resin. Activation of this resin with iodoacetonitrile and cleavage with benzyl mercaptan yielded both the desired benzyl thioester product **1** and a +25-modified adduct **2**, in a 1:1 ratio. Following purification, the two products were analyzed by ^{1}H and ^{13}C NMR spectroscopy. The proton chemical shifts were very similar for the two compounds with the exception of the ϵ-methyl protons in **2** that were shifted significantly downfield compared to those of **1** (3.7 ppm versus 2.0 ppm). Similarly, the ^{13}C shifts for **1** and **2** were nearly identical, again with the exception of the ϵ-methyl group of the methionine (16.7 ppm in **2** versus 15.2 ppm in **1**). The ^{13}C spectrum of **2** also contained a resonance at 119.1 ppm that was not present in the spectrum of **1**. Importantly, $^{1}H\{^{13}C\}$ heteronuclear multiple bond correlation (HMBC) NMR indicated that this new carbon signal cross-coupled with the ϵ-methyl protons of the methionine residue in **2** (Figure 1), indicating that the modification is on this group.

These NMR observations can be reconciled with the observed mass increase of 25 Da through the addition of a nitrile group on the ϵ-methyl group of the methionine.

(3) For a recent review, see: Muir, T. W. *Synlett* **2001**, *6*, 733−740.
(4) Huse, M.; Holford, M. N.; Kuriyan, J.; Muir, T. W. *J. Am. Chem. Soc.* **2000**, *122*, 8337−8338.
(5) Huse, M.; Muir, T. W.; Xu, L.; Chen, Y.-G.; Kuriyan, J.; Massague, J. *Mol. Cell* **2001**, *6*, 671−682.
(6) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776−779.
(7) (a) Backes, B. J.; Virgilio, A. V.; Ellman, J. A. *J. Am. Chem. Soc.* **1996**, *118*, 3055−3056. (b) Backes, B. J.; Ellman, J. A. *J. Org. Chem.* **1999**, *64*, 2322−2330. (c) Shin, Y.; Winans, K. A.; Backes, B. J.; Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1999**, *121*, 11684−11689. (d) Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A. *J. Am. Chem. Soc.* **1999**, *121*, 11369−11374.
(8) The final yield of homogeneous GS4 was further reduced because GS4 and GS4* could only be partially resolved on preparative scale HPLC, even using optimized gradients and elevated temperatures.

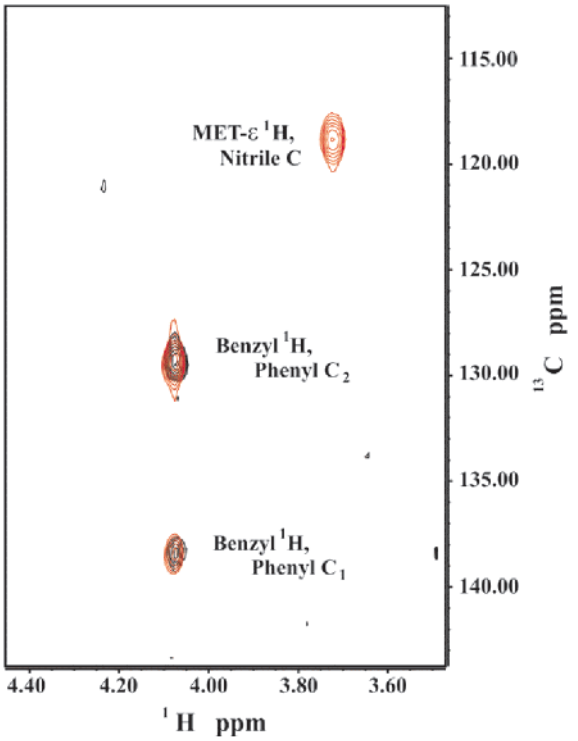
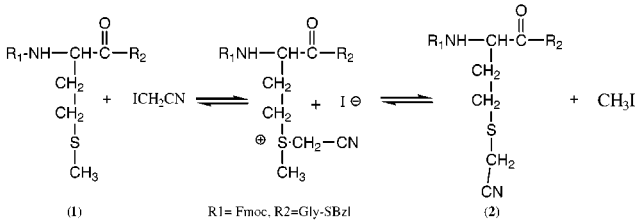


Figure 1. Region of the^{1}H{^{13}C} HMBC spectra of **1** (black contours) and **2** (red contours).

Presumably the cyanilation of the ϵ-methyl group occurs during the activation step via the *S*-alkylation mechanism shown in Scheme 2.[9] Consistent with this mechanism, the

Scheme 2. Proposed Mechanism of Methionine ϵ-Cyanilation



+25 Da adduct was not observed when TMS−CHN_2 was used in the linker activation step (data not shown).

As mentioned above, TMS−CHN_2 cannot be used in the synthesis of the GS4 peptide. The cyanilation problem can, however, be overcome by conservative substitution of norleucine for methionine, i.e., replacement of the sulfur atom with a methylene group. Accordingly, we undertook the solid-phase synthesis of the norleucine substituted GS4 phosphopeptide (GS4-Nle) on the safety catch resin. HPLC and mass analysis of the crude material obtained after thiolysis and deprotection indicated that, as expected, the cyanilation adduct had been eliminated (Figure 2B). Elimination of this major side-reaction allowed the final yield of the purified peptide to be greatly improved (60 mg from a 0.5 mmol synthesis, 4.6%).

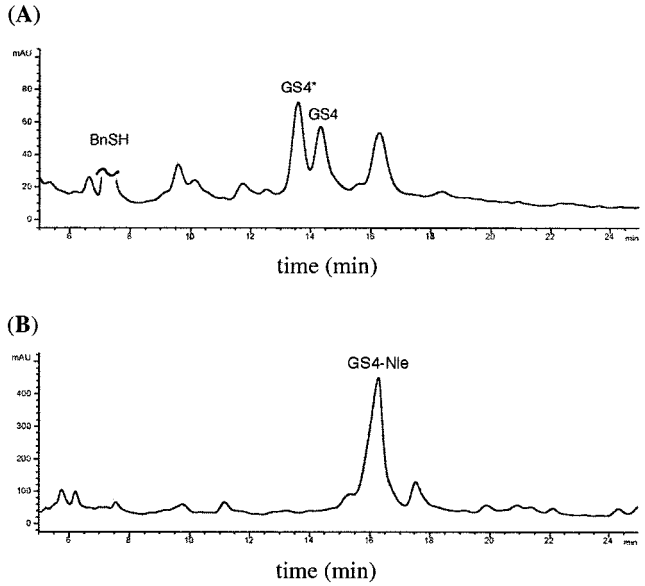


Figure 2. Region of the analytical HPLC traces for crude GS4 (A) and crude GS4-Nle (B). Indicated peaks were characterized by electrospray mass spectrometry.

The next step in the semisynthesis, the ligation of GS4-Nle to the recombinantly expressed C-terminal region of TβR-I, was also optimized. In the initial study,[4] a significant fraction of the ligation product, TβR-I(4P), and the recombinant protein fragment, TβR-IΔ,[10] were found to aggregate nonspecifically and irreversibly during the course of the ligation reaction.

In an effort to minimize this aggregation problem, we studied the effect of a number of small molecule additives in the ligation mixture. These included cosolvents known to stabilize protein structures, as well as a series of detergents.[11] Gel filtration chromatography was used to gauge the effectiveness of the additives at reducing aggregation; protein aggregates could easily be resolved from monomeric protein using a Superdex 75 column. Of the large number of conditions examined, the addition of 1,6-hexanediol to the ligation mixture proved to be the most effective at reducing aggregation.[12] Following ligation, the desired product TβR-I(4P,Nle) was purified using a combination of gel filtration and anion exchange chromatography (Figure 3).

(9) In contrast with the present model study, Backes and Ellman did not observe significant S-alkylation in a Met-containing dipeptide (ref 7b).

(10) TβR-IΔ, which contains an N-terminal Cys, was prepared using a baculovirus expression/proteolysis strategy as previously described.[4]

(11) The following additives were tested at various concentrations: 1,6-hexanediol; 1,4-butanediol; 2-methyl-2,4-pentanediol; 1,2,3-heptanetriol; dioxane; 2-propanol; DMSO; PEG-400; sucrose; spermidine; CHAPS; Triton X-100; NP-40; glycine; arginine; bovine serum albumin; $MgCl_2$ and $MnCl_2$.

(12) Addition of 350 mM 1,6-hexanediol was found to reduce nonspecific aggregation to 25% of total protein, as compared to 60% in its absence.

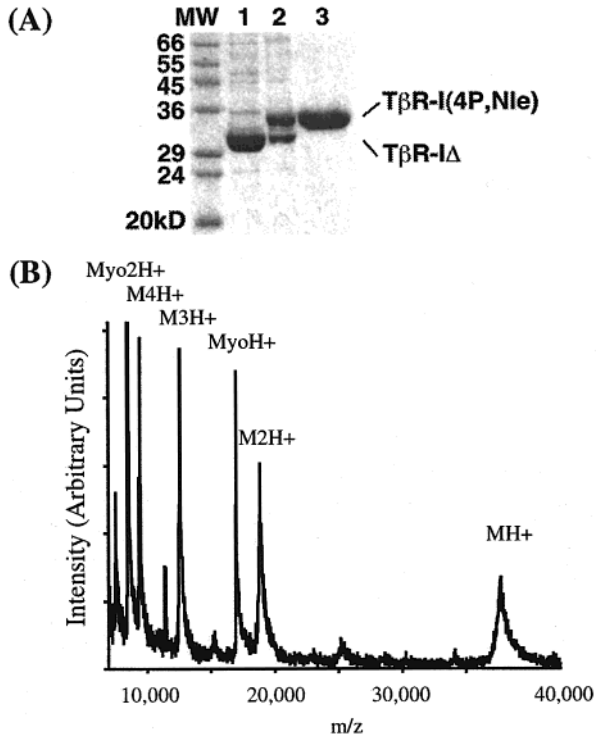


Figure 3. (A) SDS−PAGE gel showing molecular weight markers (MW), purified TβR-1Δ (lane 1), crude TβR-I(4P,Nle) ligation mixture (lane 2), and purified TβR-I(4P,Nle) (lane 3). (B) The identity of the ligation product was confirmed by MALDI-MS with myoglobin as an internal calibrant; measured mass 37,710.5 Da, calculated = 37,713.9 Da.

Using these optimized protocols, milligram quantities of highly purified TβR-I(4P,Nle) could be isolated from a single preparation (16% overall yield).[13] The biological activity of TβR-I(4P,Nle) was assayed using an established in vitro kinase assay.[4] The C-terminal domain of the physiological substrate, Smad2, was used as the substrate in the assay.[14] Under the conditions of this experiment, TβR-I(4P,Nle) displays approximately 1 order of magnitude increase in kinase activity relative to the unphosphorylated receptor. This is consistent with the increase in activity previously found for TβR-I(4P)[4,5] and indicates that the Met → Nle substitution has little or no effect on the activity of the phosphorylated receptor.

The routine synthesis of peptide α-thioesters is essential to broadening the scope of native chemical ligation.[7,15] The sulfonamide-based "safety-catch" resin approach, originally introduced by Kenner[16] and more recently refined by Ellman and co-workers,[7a,b] offers an attractive route to peptide α-thioesters using Fmoc-SPPS. In particular, the milder conditions associated Fmoc-SPPS compared to those with Boc-SPPS allows the synthesis of phosphopeptide[4] and glycopeptide[7c] α-thioesters. In the present study, we have identified a major side-reaction that occurs on methionine during activation of the linker with iodoacetonitrile. This adduct can, however, be avoided by using TMS−CHN_2 as the N-alkylation reagent or by conservative replacement of methionine with norleucine. The later strategy was successfully employed to prepare the phosphopeptide α-thioester, GS4-Nle, which was then used in the preparation of TβR-I(4P,Nle) via native chemical ligation. Importantly, addition of 1,6-hexanediol to the ligation mixture was found to significantly reduce the level of protein aggregation, allowing the final yield of tetraphosphorylated receptor to be greatly improved. This underscores one of the great strengths of native chemical ligation, namely, the tolerance of the chemistry to the presence of a wide variety of additives in the ligation soup. It is likely that these optimized ligation conditions will be applicable to the semisynthesis of related protein kinases.

Acknowledgment. Support for this work was provided by the Irma T. Hirschl/Monique Weill-Caulier Trust and by the PEW Charitable Trusts. We thank Dr. Brian Chait for allowing us to use his MALDI-TOF mass spectrometer (support by NIH grant RR00862).

Supporting Information Available: Full experimental procedures and representative spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

OL016859I

(13) In a typical ligation reaction, 3 μmol of GS4-Nle (7.5 mg) was mixed with 218 nmol of TβR-1Δ (7.5 mg) in 4.5 mL of 100 mM Hepes pH 8.0, 100 mM okadaic acid, 0.5 mM NaCl, 350 mM 1,6 hexanediol, and 50 mM MESNA at 4 °C for 16 h.

(14) Kinase assays were conducted by mixing 10 pmol of TβR-I(4P) with 80 pmol of Smad substrate in 15 μL of kinase buffer containing trace amounts of [γ-^{32P}]ATP. The level of Smad phosphorylation was analyzed by SDS−PAGE and autoradiography.

(15) (a) Canne, L. E.; Walker, S. M.; Kent, S. B. H. *Tetrahedron Lett.* **1995**, *36*, 1217−1220. (b) Futaki, S.; Sogawa, K.; Maruyama, J.; Asahara, T.; Niwa, M. *Tetrahedron Lett.* **1997**, *38*, 6237−6240. (c) Li, X.; Kawakami, T.; Aimoto, S. *Tetrahedron Lett.* **1998**, *39*, 8669−8672. (d) Alsina, J.; Yokum, T. S.; Albericio, F.; Barany, G. *J. Org. Chem.* **1999**, *64*, 8671−8769. (e) Clippingdale, A. B.; Barrow, C. J.; Wade, J. D. *J. Peptide Sci.* **2000**, *6*, 225−234. (f) Swinnen, D.; Hilvert, D. *Org. Lett.* **2000**, *2*, 2439−2442.

(16) Kenner, G. W.; McDermott, J. R.; Sheppard, R. C. *Chem. Commun.* **1971**, 636−637.