Toward Homogeneous Erythropoietin: Fine Tuning of the C-Terminal Acyl Donor in the Chemical Synthesis of the Cys[29] – Gly[77] Glycopeptide Domain

Yu Yuan[†], **Jin Chen**[†], **Qian Wan**[†], **Zhongping Tan**[†], **Gong Chen**[†], **Cindy Kan**[†], and **Samuel J. Danishefsky**[*,†,‡]

[†]Laboratory for Bioorganic Chemistry, Sloan-Kettering Institute for Cancer Research, 1275 York Avenue, New York, New York 10065

[‡]Department of Chemistry, Columbia University, Havemeyer Hall, 3000 Broadway, New York, New York 10027

Abstract

Described herein is the chemical synthesis of the Cys[29]-Gly[77] glycopeptide domain (**22**) of erythropoietin. Our initial ligation strategy targeted a C→N termini condensation between glycopeptide **3** and peptide **4**. However, the reaction was hindered by the "unattainable" reactivity, mismatched polarity, and severe aggregation of the (glyco)peptide substrates. In contrast, by tuning the C-terminal acyl donor and using smaller peptide fragments, the Cys[29]-Gly[77] glycopeptide domain of erythropoietin was prepared through unconventional N→C termini condensation reactions. The use of a *p*-cyanonitrophenyl ester and the development of a masked thiophenyl ester as acyl donors enabled us to promptly access glycopeptides bearing complex carbohydrates and offer potential synthetic applications beyond our current work.

Introduction

There is currently a major initiative underway in our laboratories to gain access to the realm of "biologics" through chemical synthesis.[1,2] While synthesis has been traditionally limited to small molecules, we believe that the fields of chemical synthesis and biologics may mutually benefit from such an effort. Our interest lies primarily in the preparation of fully functional homogeneous glycoproteins. Unlike their carbohydrate-free counterparts, glycoproteins exist as heterogeneous mixtures, a feature which significantly impedes elucidation of the carbohydrate structure-function relationship. We are especially interested in the chemical synthesis of homogeneous glycoproteins in the context of complex carbohydrates, and the focal point of this initiative is our proposed synthesis of the glycoprotein erythropoietin (EPO). [3,4]

To date, a tremendous number of peptides and proteins have been synthesized by both solid phase and solution phase methods. It is generally accepted that peptides with circa 50 amino acid residues represent the upper limit for solid phase peptide synthesis by Fmoc chemistry. For the synthesis of larger peptides or glycopeptides, solution phase methods become more advantageous due to the availability of powerful ligation methodologies and the ease of isolating pure products by reverse phase HPLC. However, a major challenge arises when the inherent physical properties of peptides prevent the use of these traditional methods for their synthesis. For example, during our studies on the synthesis of EPO(29–77), the aggregation-

s-danishefsky@ski.mskcc.org .

prone nature of this sequence created problems during the critical solution phase ligation step. Herein, we present the synthesis of EPO(29–77) and disclose our strategy for coping with this otherwise problematic aggregation issue.

Our overarching assembly plan for the total synthesis of homogeneous erythropoietin is implied in Scheme 1, line 2. For the titled fragment at issue in this report, we envisioned the sequential ligation of two glycopeptide subfragments, from the C→N termini. For maximum convergency, an *N*-linked carbohydrate fragment was to be assembled through Lansbury aspartylation[5] of dodecasaccharide **1**[6] and a short peptide, followed by a ligation with a non-carbohydrate containing peptide segment (Scheme 1, line 3).

Results and Discussion

Our initial attempt to assemble the EPO(29–77) fragment commenced with the union of dodecasaccharide **1** and the partially protected EPO(29–42) fragment, **2** (Scheme 2). Under modified Lansbury conditions,[5] the anomeric amine **1** was covalently joined to Asp^{38} through a newly formed amide bond to afford glycopeptide **3** in 27% yield. However, direct condensation between glycopeptide **3** and peptide **4** proved to be surprisingly difficult to accomplish in our hands. Possible explanations for the failed condensation included "unattainable" reactivity, mismatched polarity, and severe aggregation. It is possible that after TCEP-induced disulfide cleavage and concomitant O→S acyl migration,[2b] the resulting proline thioester could not act as a competent acyl donor in such a sterically demanding coupling reaction. Furthermore, although a barely discernible amount of glycopeptide was produced, the isolation was hindered by the overlapping retention times shared by the ligated product and peptide **4** in various solvent gradients on reverse phase HPLC. To further complicate matters, peptide **4** easily aggregated in DMSO within a couple of hours. This is based on our observations, whereby **4** was initially consumed but little or no product could be detected. After 6–8 hours, LC-MS analysis of the reaction mixture indicated complete "disappearance" of compound **4**; however, its signal could be "regenerated" by addition of hexafluoroisopropyl alcohol to the reaction mixture, a reagent known to disrupt peptide aggregation.[7] Apparently, this aggregation problem not only resulted in the observed diminished reactivity, but also complicated the product isolation process.

To address the aforementioned problems, we first focused our efforts on tuning the acyl donor reactivity. After an extensive survey of the literature and screening of what we thought to be the most promising prospects, our primary candidates for acyl donors focused on differentially substituted phenyl esters. Model studies led us to the hope that *p*-nitrophenyl esters[8] might serve as viable acyl donors for the peptide ligations required herein. For example, peptide **5**, which contains a disaccharide and a *C*-terminal *p*-nitrophenyl ester, was smoothly condensed with the long peptide **4** within 12 hours to afford **6**, a simplified version of the EPO(29–77) fragment.[9] It is also worth mentioning that during the preparation of **5**, there was no observed competition between the Asp^{38} HOAt ester and the terminal Pro^{42} *p*-nitrophenyl ester; in other words, the disaccharide reacted exclusively at the Asp^{38} side chain.

Encouraged by these results, we next had to determine whether this protocol could be applied to the synthesis of a glycopeptide containing a dodecasaccharide. To our disappointment, when peptide **7** was employed in the Lansbury aspartylation with dodecasaccharide **1**, no desired product could be detected (Scheme 4). Instead, a compound with a molecular weight corresponding to the desired molecular weight minus nitrophenol was produced as the sole product. Although the point of carbohydrate attachment has not yet been rigorously determined, it is still apparent that the *p*-nitrophenyl ester was too reactive as an acyl moiety and is accordingly unsuitable to serve in the Lansbury reaction.

Nonetheless, we persisted in seeking to find a new acyl donor, one that allows for effective condensation but is much less susceptible to nucleophilic attack by the anomeric amine. By comparing the Hammett constants of various phenyl esters,[10] the small reactivity window of *p*-cyanophenyl esters prompted us to prepare **8**, the C-terminal of which was anticipated to serve as an appropriate donor. Upon mixing dodecasaccharide **1** and peptide **8**, followed by addition of DMSO solutions of DIEA and HATU, glycopeptide **11** was produced and subsequently isolated in 28% yield (Scheme 4).

The applicability of the cyanophenyl ester ligation to the synthesis of the EPO(29–77) fragment hinges on the ability to separate peptide **4** from the final product. We expected that modifying the EPO(29–42) fragment could result in changed retention times. In our previous studies, we noticed that the retention time of the EPO(29–42) fragment varied by as much as 5 min when the N-terminal Fmoc group on Cys[29] was removed. Because the carbohydrate and the *C*-terminal *p*-cyanophenyl ester are both sensitive to acidic and basic conditions, as well as to nucleophilic attack, the liberation of the Cys[29] amino group must be executed before the carbohydrate is introduced. Thus we chose peptide **9** as the substrate for Lansbury aspartylation. While we were aware that anomeric amines are less nucleophilic than normal amines, we remained confident that the amino group of **1** would prove more reactive than the amino group of Cys[29], which is substituted by a β heteroatom. We were pleased to find that the reaction proceeded to afford the desired glycopeptide **12**, possessing a free amino group and a relatively reactive *C*-terminal cyanophenyl ester in 33% yield.

With a suitable glycopeptide in hand, we turned our attention to the coupling of glycopeptide **12** and peptide **4**. At the peptide level, the analogous ligation between **9** and **4** proceeded with acceptable conversion in 8–12 hours; but unfortunately, this reactivity did not extend to the corresponding glycopeptide **12**. We attributed this failure to a conformational change within EPO(29–42) (induced by the presence of the large carbohydrate), which serves to prevent large peptides from reacting at the C-terminal cyanophenyl ester. Consequently, either the conformation of **4** must be changed or a subunit of **4** must be used to overcome the accessibility problem at the ligation site.

After evaluating the overall situation, we decided to revise our initial strategy for the synthesis of Epo 29–77 (Scheme 5). Instead of dividing the sequence into two fragments, the new retrosynthetic analysis called for an additional disconnection at Gly[57]-Glu[58]. Thus, we envisioned four short building blocks, including the dodecasaccharide. These would be assembled through sequential *N*-terminal to *C*-terminal bond constructions.

Traditionally, peptide elongation is performed from the *C*-terminus to the *N*-terminus. Kent and co-workers have reported an alternate approach for peptide synthesis, namely kinetically controlled ligation,[11] that enables peptide coupling in the opposite direction, from the *N*-terminus to the *C*-terminus. However, this non-traditional strategy requires a *C*-terminal thioester exchange for each ligation. This protocol introduces an extra step for each fragment extension. In contrast, the masked thioester recently developed in our group offers a unique alternative.[12] When peptide **9** and the EPO(43–57) peptide containing the masked thioester (**13**) were treated with HOOBt and diisopropylethyl amine (DIEA), the ligated peptide **14** was formed; however, the masked thioester was completely hydrolyzed under the basic reaction conditions (Scheme 6).[13]

It was soon realized that the selective reinstallation of the phenyl ester at Gly[57] in the presence of other carboxylic groups would be difficult. Furthermore, efforts to adjust the pH of the reaction to prevent ester hydrolysis were futile. Nevertheless, we reasoned that the unmanageable susceptibility to hydrolysis might be a consequence of the sterically accessible Gly[57] carbonyl group. To suppress vulnerability to nucleophilic attack, our *o*-disulfide phenol

was modified to include an additional *ortho*-substituent. We anticipated that the new phenolic substrate would effectively block both π faces of the Gly[57] carbonyl group. The modified phenol was synthesized from the commercially available propyl phenol **15** (Scheme 7). Treatment of chlorocarbonylsulfenyl chloride with phenol **15** in the presence of Bu_3N and $AlCl_3$ afforded the corresponding 1,3-benzoxathiol-2-one **16** in good yield. Subsequent basic hydrolysis gave thiophenol **17**.[14] Following our previously reported procedure,[11] sequential oxidative dimerization and disulfide metathesis afforded the desired phenol **18** in 43% overall yield.

With phenol **18** in hand, the modified EPO(43–57) peptide **19** was prepared. We proceeded to investigate the feasibility of joining the EPO(29–42) glycopeptide (**12**) with the EPO(43–57) peptide (**19**) while retaining the masked thioester. We were pleased to find that, under the exact same conditions as described in Scheme 6, the desired glycopeptide **20** was isolated in 48% yield without any detectable loss of *C*-terminal phenyl ester.

With this success, we were ready at last to complete the synthesis of the EPO(29–77) fragment. All that remained was the coupling of the EPO(58–77) peptide fragment (**21**) with glycopeptide **20**. Thus, upon treatment of **20** with TCEP and DIEA, the reactive thiophenyl ester, generated *in situ*, reacted smoothly with peptide **21** (EPO(58–77)) in the presence of HOOBt, to furnish the desired EPO(29–77) fragment in 51% isolated yield.[15]

Conclusion

In conclusion, we have completed the synthesis of the EPO(29–77) glycopeptide fragment through consecutive condensations from the *N*-terminal to the *C*-terminal. The highlights of the present work include the employment of a *p*-cyanophenyl ester as an effective acyl donor and the development of a masked thiophenyl ester that is resistant to basic hydrolysis yet sufficiently reactive in subsequent ligations. While the established methods (for instance, para-nitrophenyl ester) could have provided the EPO(29–77) fragment containing *simple* carbohydrate groups, the use of the cyanophenyl ester and the development of a masked thiophenyl ester were critical to the preparation of the EPO(29–77) fragment containing the *complex dodecasaccharide unit*. With this accomplishment, our program toward building homogeneous erythropoietin can advance to the final stage.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. See preceding paper: Tan Z, Shang S, Halkina T, Danishefsky SJ, Yuan Y. J. Am. Chem. Soc 2008;130 pp.

2. (a) Wilson RM, Danishefsky SJ. Pure Appl. Chem 2007;79:2189. (b) Chen G, Wan Q, Tan ZP, Kan C, Hua ZH, Ranganathan K, Danishefsky SJ. Angew. Chem. Int. Ed 2007;46:7383. (c) Chen G, Warren JD, Chen JH, Wu B, Wan Q, Danishefsky SJ. J. Am. Chem. Soc 2006;128:7460. [PubMed: 16756298] (d) Chen JH, Warren JD, Wu B, Chen G, Wan Q, Danishefsky SJ. Tetrahedron Lett 2006;47:1969. (e) Wan Q, Danishefsky SJ. Angew. Chem. Int. Ed 2007;46:9248. (f) Wu B, Chen JH, Warren JD,

Chen G, Hua ZH, Danishefsky SJ. Angew. Chem. Int. Ed 2006;45:4116. (g) Wu B, Warren JD, Chen JH, Chen G, Hua ZH, Danishefsky SJ. Tetrahedron Lett 2006;47:5219.

3. Lai PH, Everett R, Wang FF, Arakawa T, Goldwasser E. J. Biol. Chem 1986;261:3116. [PubMed: 3949763]

4. (a) Chen JH, Chen G, Wu B, Wan Q, Tan ZP, Hua ZH, Danishefsky SJ. Tetrahedron Lett 2006;47:8013. [PubMed: 17426819] (b) Wu B, Tan ZP, Chen G, Chen JH, Hua ZH, Wan Q, Ranganathan K, Danishefsky SJ. Tetrahedron Lett 2006;47:8009.

5. Cohen-Anisfeld ST, Lansbury PT. J. Am. Chem. Soc 1993;115:10531.

6. Wu B, Hua ZH, Warren JD, Ranganathan K, Wan Q, Chen G, Tan ZP, Chen JH, Endo A, Danishefsky SJ. Tetrahedron Lett 2006;47:5577.

7. Milton SCF, Milton RCD. Int. J. Pept. Protein Res 1990;36:193. [PubMed: 2272756]

8. Bodanszky M. Nature 1955;175:685. [PubMed: 14370194]

9. Based on the LCMS trace, there was approximately 35–45% conversion after 12 hours.

10. Hansch C, Leo A, Taft RW. Chem. Rev 1991;91:165.

11. Bang D, Pentelute BL, Kent SBH. Angew. Chem. Int. Ed 2006;45:3985.

12. Warren JD, Miller JS, Keding SJ, Danishefsky SJ. J. Am. Chem. Soc 2004;126:6576. [PubMed: 15161285]

13. Based on the LCMS trace, there was approximately 40% conversion after 5 hours and the ratio of hydrolyzed to unhydrolyzed products was 10:1.

14. Yoshida Y, Ogura M, Tanabe Y. Heterocycles 1999;50:681.

15. See the Supporting Information for a discussion of impurities, which are often observed in glycopeptides displaying the N-linked dodecasaccharide.

Scheme 1.
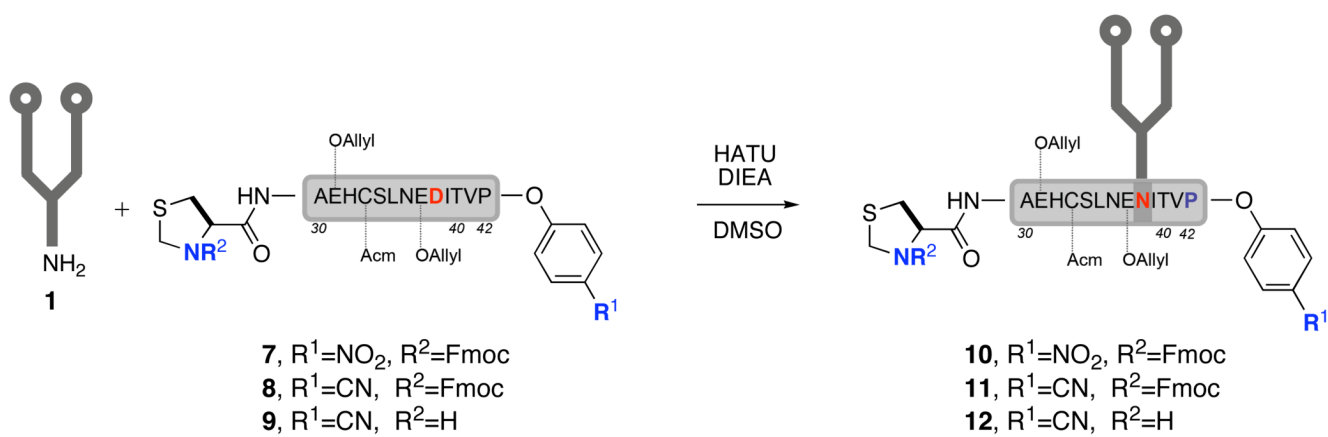Retrosynthetic analysis of Epo and fragment Cys[29]-Gly[77].
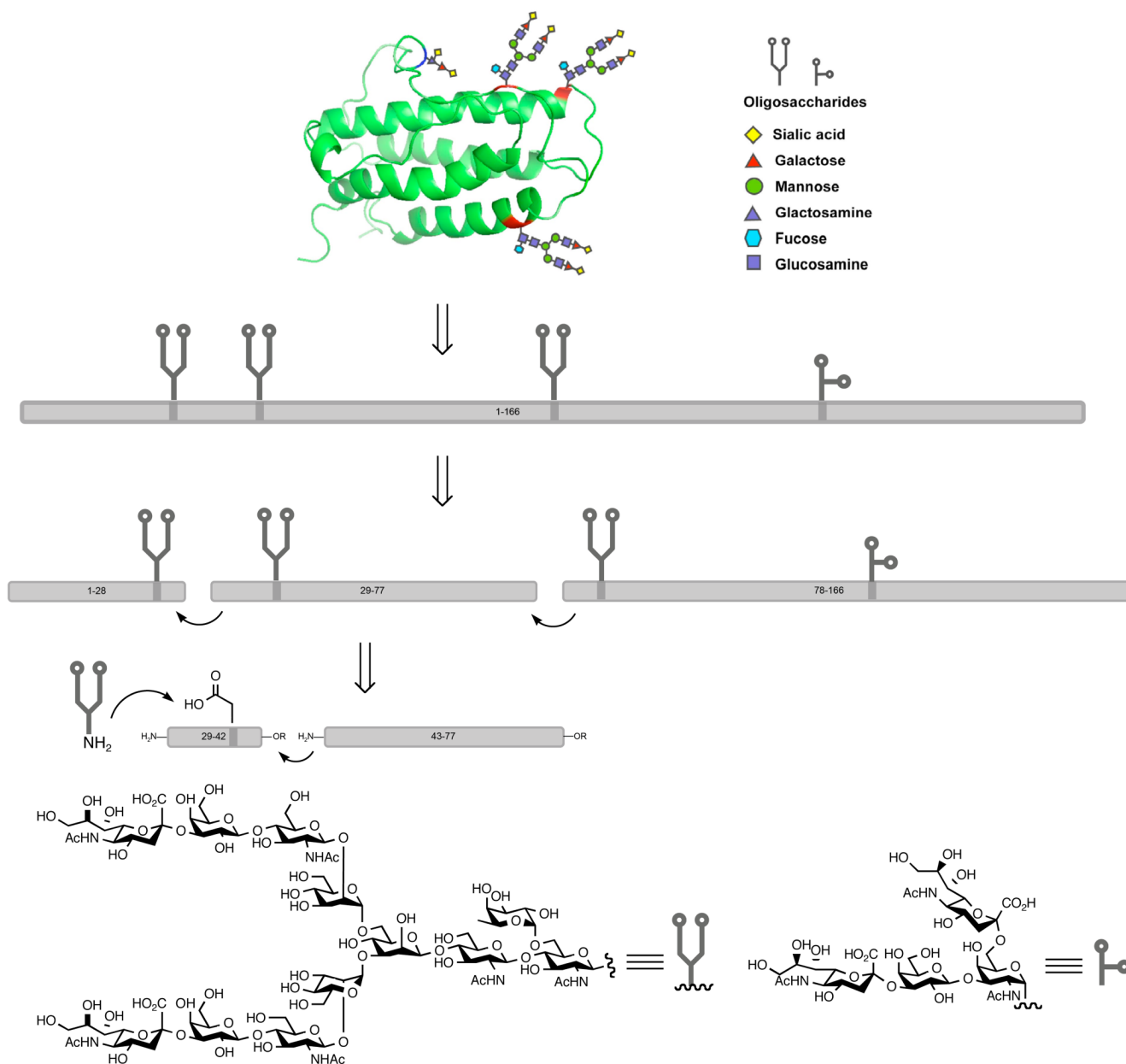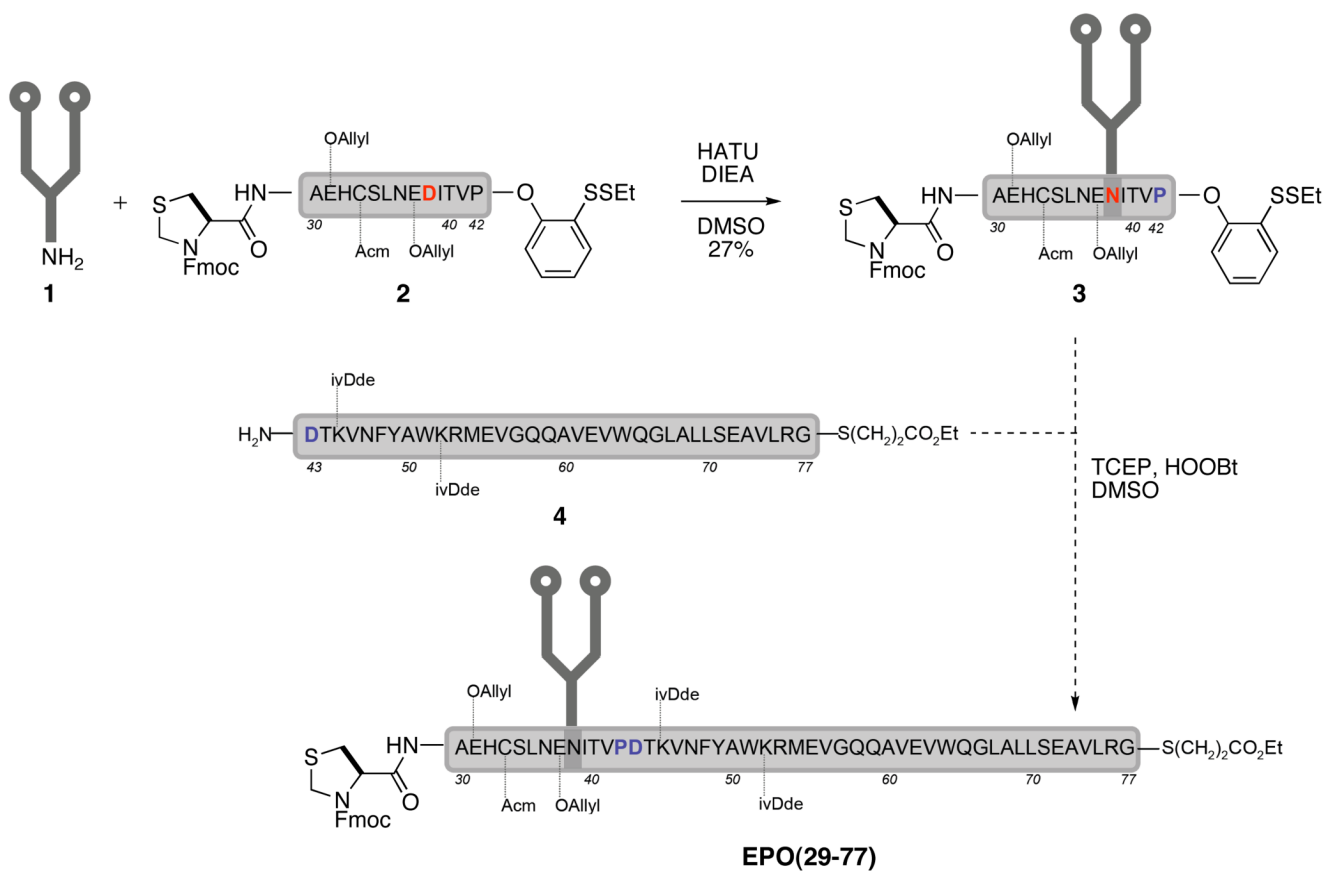
Scheme 2.
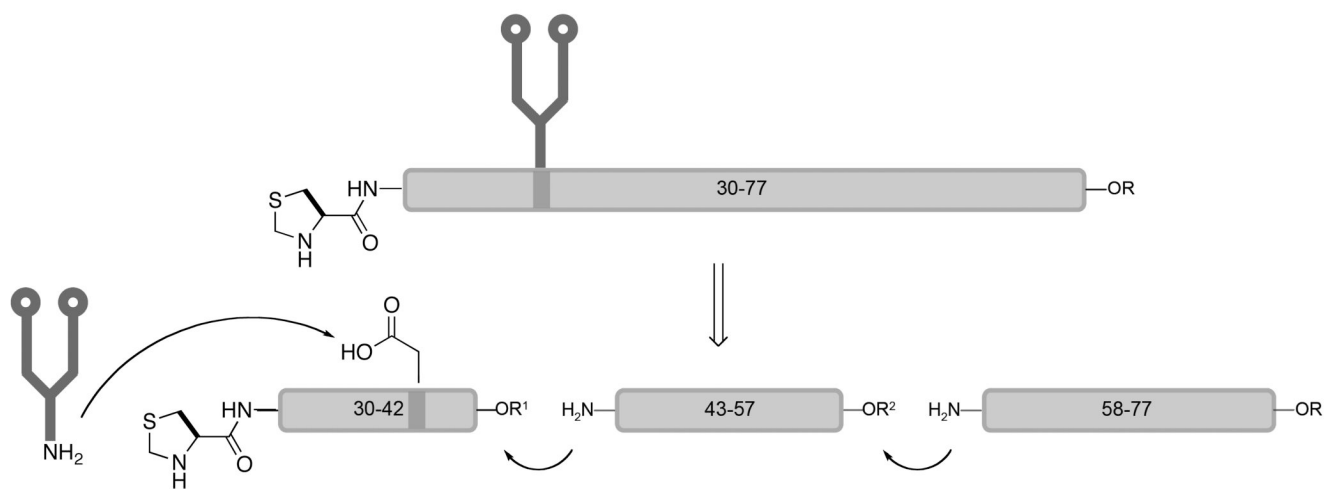Initial construction of fragment Epo 29–77.
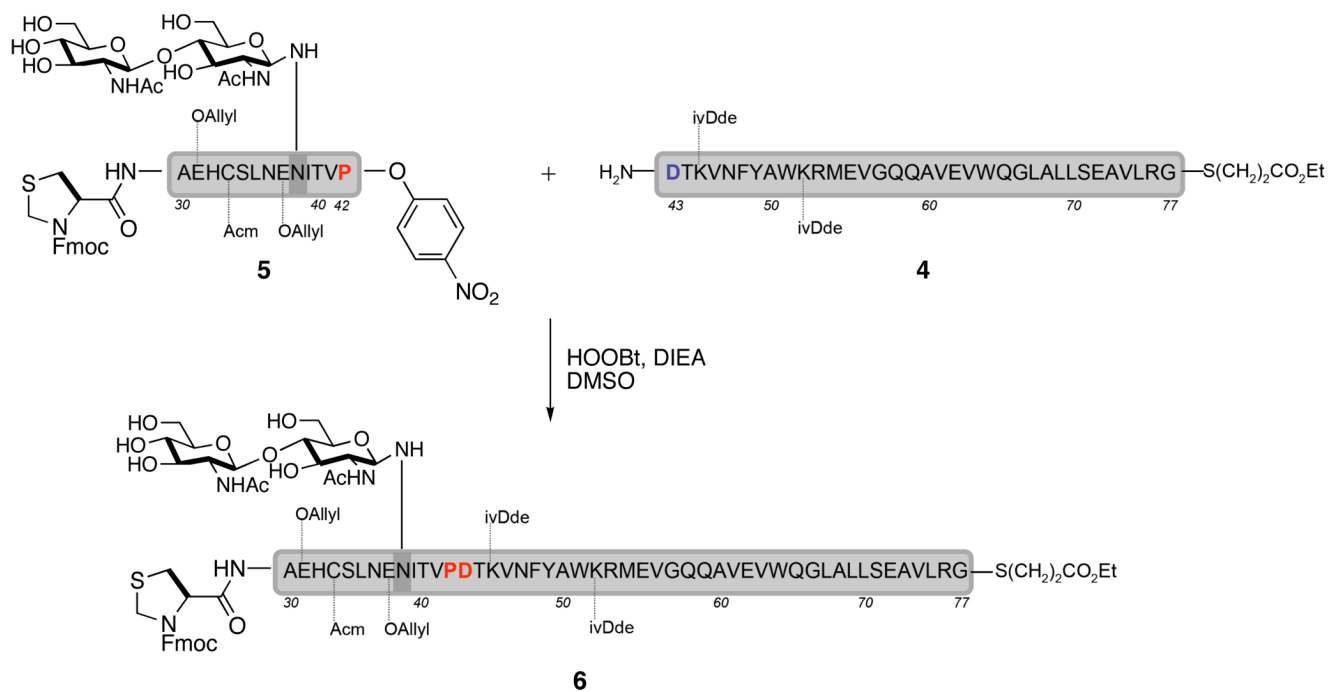
Scheme 3.
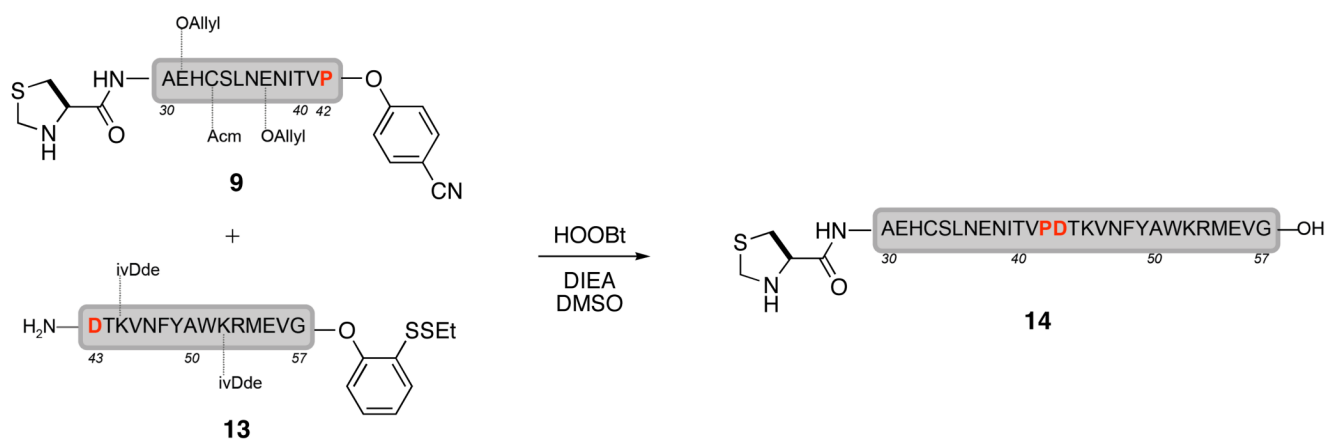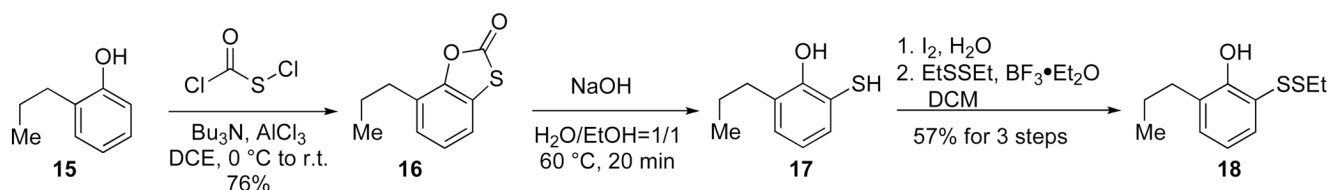Successful fragment condensation to deliver EPO(29–77) with disaccharide.

Scheme 4.
Lansbury aspartylation between dodecasaccharide and reactive acyl donors
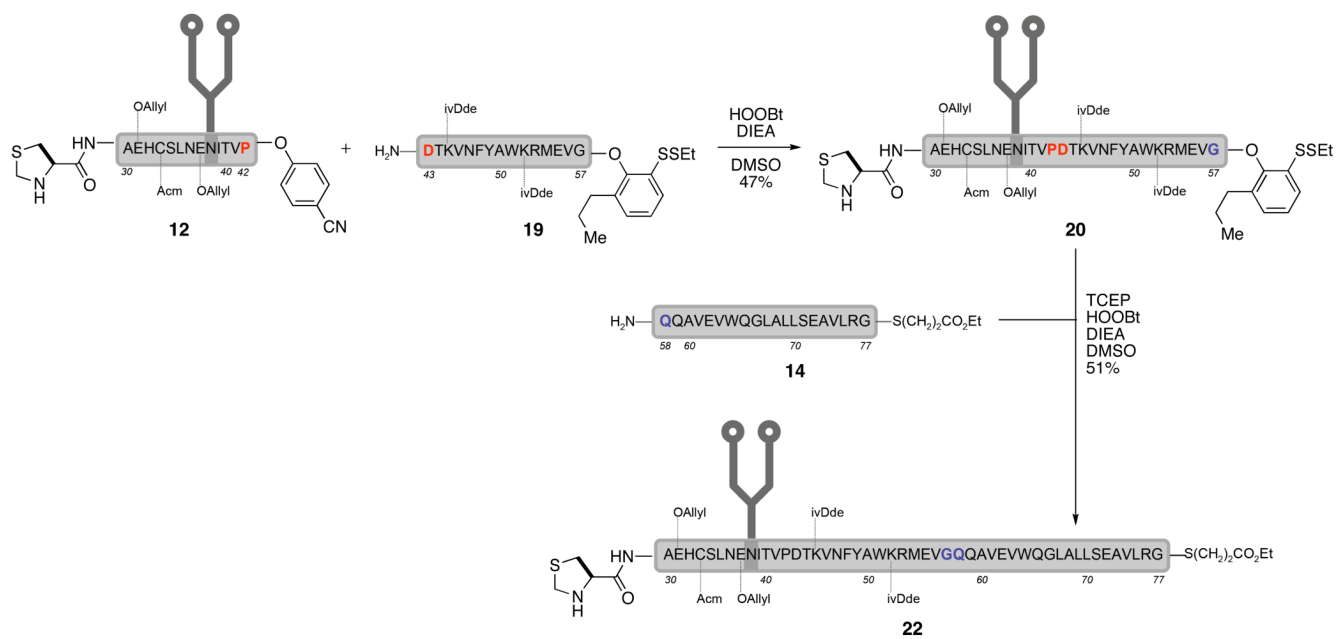
Scheme 5.
Revised retrosynthetic analysis for Epo 29–77

Scheme 6.
Potential problem associated with hydrolysis of the masked thioester.

Scheme 7.
Preparation of masked thiophenol **18**

Scheme 8.
Completion of synthesis EPO(29–77)