



Published in final edited form as:

Tetrahedron. 2009 November 7; 65(45): 9047–9065. doi:10.1016/j.tet.2009.09.032.

Recent Departures in the Synthesis of Peptides and Glycopeptides

Cindy Kan¹ and Samuel J. Danishefsky^{1,2}

¹Laboratory for Bioorganic Chemistry, Sloan-Kettering Institute for Cancer Research, 1275 York Avenue, New York, NY 10065, USA

²Department of Chemistry, Columbia University, Havemeyer Hall, 3000 Broadway, New York, NY 10027, USA

Abstract

In this account, we describe the results of a research program directed to the proposition that chemical synthesis can play a valuable role in identifying biologic level molecules worthy of pharma level development. We recount our journey towards the chemical synthesis of homogeneous erythropoietin, the challenges we encountered, and our efforts to address deficiencies in the current “state of the art” of glycopeptide synthesis. Here we describe new methods for the synthesis of glycopeptides that have emerged from the erythropoietin adventure, including the development of unique C-terminal acyl donors, novel amide bond forming methods, and new ligation and coupling strategies.

Keywords

peptides; glycopeptides; glycoproteins; ligation; desulfurization; isonitriles; chemical synthesis; erythropoietin

INTRODUCTION

Within the traditional mindset of the pharmaceutical sciences there prevails an unstated but sharp division between the worlds of “biologics” and “small molecules.” The dominant and typical role for chemistry has been limited to the realm of “small molecules,” where its capacities for synthesis have traditionally been well recognized. In contrast, within the “biologics” domain, chemistry has played relatively minor roles. In this paper, we describe the results of a research program directed to the proposition that chemical synthesis can actually play a valuable role in identifying biologic level molecules worthy of further pharma level development.¹ A few remarks of a historical nature are perhaps in order before dealing with hardcore specifics.

In 1982 we described what was then a new reaction that ushered in a large body of chemistry. The reaction was the Lewis Acid Catalyzed Diene Aldehyde Cyclocondensation reaction that

© 2009 Elsevier Ltd. All rights reserved.

s-danishefsky@ski.mskcc.org; 212-639-5502 (p); (212)-772-8691 (f).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

affords dihydropyrones in a remarkably general and straightforward fashion.ⁱⁱ Explorations along these lines carried a special appeal to us since, almost a decade earlier, our laboratory had launched the idea of 1,3-dioxygenated (synergistic) dienes in “all-carbon” Diels-Alder reactions.ⁱⁱⁱ Reduction of the dihydropyrones afforded synthetic glycals. This led to a decade-long fascination with glycals, both synthetic glycals as well as those derivable from naturally available carbohydrates, particularly hexoses.^{iv} In time, we learned how to take advantage of nuances in the chemistry of glycals. Control of the chemistry of glycals can be more discriminating than is the case with traditional pentafunctional hexoses. Indeed, we learned how to exploit glycals not only as donors, but as acceptors – a very new idea for its time.^{iv} There then ensued a long term involvement directed to the chemical synthesis of oligosaccharides by what we called “glycal assembly” (Figure 1).^v A particular focus of glycal assembly was that of fully synthetic vaccines, several of which have emerged from successful Phase I clinical trials.^{vi}

In ca. 2002, there began to emerge in our group an interest in structures even more complex than oligosaccharides, i.e., asparagine linked glycopeptides and glycopolypeptides.^{vii} Indeed, we were able to apply our earlier findings and later discoveries to reach such systems including structures mimicking parts of PSA^{viii} and GP120.^{ix} These triumphs, in time, led us to begin *fantasizing* about the complex, multiply glycosylated protein, erythropoietin (vide infra), in particular, about synthesizing erythropoietin (EPO) in a chemical laboratory as an alternative to the Golgi of CHO cells. In time, the fantasizing gave way to planning, and in due course, some visionary pioneers in our laboratory actually began experimentation directed toward the total synthesis of homogeneous EPO. In particular, we were drawn to EPO by its therapeutic value (vide infra) and its complex structure. It was in this context that we set for ourselves the formidable long-term challenge of preparing fully synthetic erythropoietin in homogeneous form. While our focus was on EPO, we anticipated the likelihood that the program would provide valuable new insights into the synthesis of glycopeptides and even insights into the nature of these fascinating bi-domainal “biologics”.

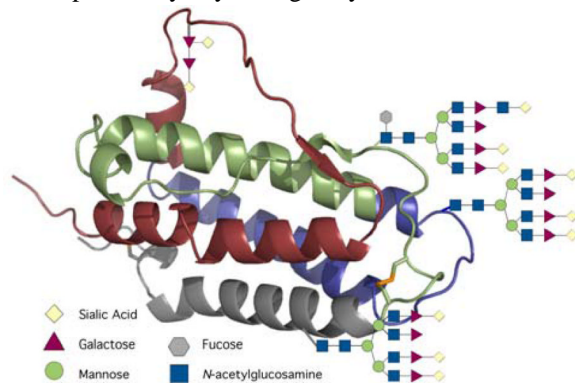
ERYTHROPOIETIN

Glycoproteins arise from protein glycosylation, a complex post-translational process that plays a critical role in a wide range of cellular functions. Protein glycosylation has been implicated in mediating protein folding, in protecting against proteolysis, in cellular differentiation, and in cell-cell communication.^x In mammalian systems, glycoproteins are commonly expressed as heterogeneous mixtures of glycoforms.^{xi} Efforts to determine the exact biological role of a defined glycosylation profile have been seriously complicated by difficulties associated with isolating significant quantities of homogeneous glycoforms.^{xii} While advances in genetic engineering have enabled access to glycoproteins, they do not as yet allow for homogeneous expression of glycoproteins. Furthermore, the emergence of their oligosaccharide regimes are structurally limited both by the host glycosylation machinery and by the genetically encoded amino acid template. Consequently, access to homogeneous and structurally definable glycoproteins could be of significant value for the systematic study of the effects of protein glycosylation on biology-level performance.^{xiii} Of particular interest would be issues of biostability and the nature of the crosstalk between the two (polypeptidic and carbohydrate) domains and function.^{xiv} In particular, de novo chemical synthesis might provide a powerful path to designed glycoproteins as it offers, in principle, precise structural control while providing opportunities for systematic variation of both the structures of the glycodomains and their location within the peptide sequence.

Erythropoietin (EPO) is a 30 kDa glycoprotein hormone with a 166 amino acid backbone and features three *N*-linked glycans (at Asn²⁴, Asn³⁸, and Asn⁸³), one *O*-linked glycan (at Ser¹²⁶), and disulfide bonds at Cys⁷ and Cys¹⁶¹, as well as Cys²⁹ and Cys³³.^{xv,xvi} EPO is

produced primarily in the kidneys. As a hematopoietic growth factor, erythropoietin regulates and maintains red blood cell, or erythrocyte, levels.^{xvii} The production of erythropoietin is regulated by hypoxia following a classical feedback mechanism. While the major role of EPO was initially thought to be limited to erythrocyte production in the kidney, EPO and its receptor have been found in the nervous system and the EPO receptor is also present in the endothelium.^{xviii}^{xix} Natural EPO exists as a heterogeneous mixture of glycoforms and is isolated in only low yield and after difficult purification.^{xx} However, subsequent isolation, cloning of the EPO gene, and expression in mammalian cells have enabled the engineering of recombinant human erythropoietin (rhEPO).^{xxi}

Moreover, heterogeneous EPO has found clinically valuable application as an adjuvant used in the treatment of anemia associated with renal failure and cancer chemotherapy.^{xxii} Extensive efforts have been directed to the study and understanding of the structure and function of EPO, including the role of glycosylation. Specifically, Higuchi et al. demonstrated that erythropoietin's (rhuEPO) three N-linked glycosyl groups were not required for *in vitro* activity in terms of binding to the EPO receptor but were required for *in vivo* activity.^{xxiii} In contrast, the single O-linked glycoporphin apparently was not required for *in vitro* or *in vivo* activity.^{xxiii} Furthermore, the *in vivo* activity of EPO was shown to be more than casually associated with its sialic acid content.^{xxiv} Interestingly, Kent and coworkers have synthesized a polymer-modified EPO that apparently demonstrated incremental *in vivo* activity compared to the glycosylated rhuEPO.^{xxv} Despite impressive biological and chemical studies that have been conducted to understand the structure and function of EPO, efforts to determine the exact biological roles of defined EPO glycoforms have been thwarted by difficulties associated with isolating significant quantities of homogeneous glycoforms. In fact, this goal has not been accomplished by any biologically enabled means.



Chemical synthesis of glycopeptides or glycopolypeptides bearing complex carbohydrates has only recently been realized.^{xxvi} This situation is in sharp contrast to the plethora of non-glycosylated proteins that have been prepared by chemical synthesis, enabled by the discovery and development of native chemical ligation.^{xxvii}^{xxviii} Indeed, using native chemical ligation (NCL), Kajihara and Dawson reported the first synthesis of a complex glycoprotein, i.e., a single glycoform of monocyte chemotactic protein-3 containing human complex sialyloligosaccharide.^{xxix} Erythropoietin glycopeptide fragments containing complex sialyloligosaccharides have been prepared by our group using either NCL or direct condensation methods.^{xxx}^{xxxi}^{xxxii}^{xxxiii}^{xxxiv} Recently, an asialo erythropoietin glycopeptide fragment was reported by Kajihara,^{xxxv} for which serine ligation was effectively employed. The chemical syntheses of ribonuclease B and CTLA-4 glycopeptide fragments using NCL were also reported by the groups of Unverzagt and Kajihara, respectively.^{xxxvi}^{xxxvii} Hojo and co-workers have also employed direct condensation methods in their chemical synthesis of an N-glycosylated Ig domain of emmprin.^{xxxviii}

Notwithstanding these clear advances, the chemical synthesis of complex glycoproteins and glycopeptides remains a significant and daunting challenge.

PROSPECTS

We were fully aware of the long road before us and of the many obstacles that would inevitably be faced in seeking to complete such a program. While our defining “line in the sand” goal is the synthesis of homogeneous erythropoietin, it seemed likely that the struggle itself would prove valuable, as well as stimulating. Such a journey itself would help us to identify and then address deficiencies in the current “state of the art” of glycopeptide and glycopolymer synthesis. As matters transpired, the attempted total synthesis of homogeneous EPO has been, and continues to be, a highly productive undertaking that has resulted in the development of unique C-terminal acyl donors, novel amide bond forming methods, and new ligation and coupling strategies. The scope and applicability of these new methods have been demonstrated in multiple contexts, most recently in our syntheses of erythropoietin glycopeptide fragments.^{xxxii·xxxiii·xxxiv} Here we emphasize new methods for the synthesis of glycopeptides and glycopolymers that have already emerged from the EPO adventure.

As we were pondering possible strategies for the chemical synthesis of glycoproteins, we recognized the need for a highly convergent approach to glycopeptides that would enable access to multiply-glycosylated glycoproteins. Ideally, the synthesis would entail the coupling of separate glycopeptide fragments, which would offer not only maximum convergency for reaching a single glycoform of a glycoprotein, but would also provide a modular synthesis from which new glycoforms could be prepared. Furthermore, during the course of our studies on glycopeptide synthesis, we observed that Lansbury aspartylation reactions^{xxxix} between complex oligosaccharides and peptides were more efficient when the peptidic aspartyl acyl donors were of relatively modest length and presented relatively stable side-chain functional groups. Thus, we required methods that would not only enable the synthesis of glycopeptide domains from their respective glycosyl and peptide fragments, but would also enable the subsequent ligation of these product glycopeptide domains.

RESULTS AND DISCUSSION

In the early phases of the effort, it had been demonstrated that a peptide α thioester and a N-terminal cysteinyl glycopeptide fragment could be joined through native chemical ligation.^{xi} An analogous approach could be employed to join two glycopeptide segments. Clearly, implementation of such a format would require, at some stage, the synthesis of a glycopeptide α thioester. Faced with the practical challenges of preparing such a α thioester, we developed a type of construct which would hopefully be more accessible and functionally even more valuable as an acyl donor. This is first presented in the form of a phenolic ester equipped with an unsymmetrical disulfide (see **1** Figure 2).^{xi} It was presumed that the phenolic ester itself would be of relatively modest acyl donor proclivity. However, one could anticipate that reduction of the disulfide moiety would reveal a free ortho benzene thiol function (**2**) that could enhance the acylating ability of the phenolic ester. Alternatively, a bona fide O→S acyl transfer could serve to generate the corresponding aryl thioester (**3**). At the outset, it was presumed that the thioester valence tautomer would be present, at best, as a minor component in equilibrium with the O-ester. We anticipated that either of these two isomeric ester forms might act as viable acyl donors, under the appropriate reaction conditions.

Happily, the utility of this latent acyl donor in glycopeptide couplings has been well demonstrated, for instance, in the synthesis of **9** (Scheme 1).^{xi} Two glycopeptide segments, one bearing our ortho-disulfide phenolic ester at the C-terminus (**7**) and the other a *t*-butylthio-protected cysteine at the N-terminus (**8**), were subjected to reductive cleavage of their disulfide bonds. The resulting glycopeptide intermediates underwent ligation to afford the doubly

glycosylated polypeptide **9**. The initial success in bringing together two glycopeptide segments through the use of our newly crafted ortho-disulfide phenolic ester via chemical ligation was certainly encouraging.

In terms of EPO, we remained apprehensive as to whether the requirement of an N-terminal cysteine for NCL could be a limiting factor. Cysteine is one of the least abundant amino acids found in proteins, second only to tryptophan.^{xliii} In the context of erythropoietin, which has four glycosylation sites, a convergent synthesis would entail the sequential ligation of the four separate glycopeptide domains. However, the cysteine residues in EPO are unevenly distributed, thereby limiting their utility in ligating different glycopeptide domains. To deal with the paucity of cysteine residues in EPO, it was deemed essential to develop cysteine-independent ligation methods that would also complement our cysteine-based ligation method.

An opportunity to develop a cysteine-free ligation presented itself during our mechanistic studies on ligations with our newly crafted ortho-disulfide phenolic ester.^{xliii} Under our reaction conditions, reduction of the ortho-disulfide group results in a free ortho benzene thiol that rearranges in a unidirectional O→S acyl transfer, to the corresponding thioester (Figure 2, **2**→**3**). We recognized that such an intermediate thioester could potentially undergo a second acyl transfer event, with a proximal amine group, culminating in a cysteine-free ligation strategy (Figure 3). By introducing a second peptide substituent (**12**), two glycopeptide fragments could, in principle, be joined in sequential acyl transfers. As described below, removal of the thiol protecting group from **13** initiated the first acyl transfer, from O→S, to afford thioester **15**. In the presence of the proximal amine group, a second acyl transfer, this time from S→N, gives rise to **16**. Removal of the organizing auxiliary unit affords **17**, in which the fragments are now joined by a new amide bond.

This novel cysteine-free ligation strategy was, in fact, used to prepare glycopeptide **19**, which features two protected Tn antigens, from the differentially glycosylated substrate **18** (Scheme 2).^{xliii} Following acidic removal of the thiol protecting group in **18**, sequential acyl transfers afforded the desired product **19**, wherein the new amidic nitrogen remains connected to the phenolic auxiliary. At this writing, we have not yet dealt with the cleavage of the now-vestigial 1-thiyl meta cresyl residue in **19**.

Following the principles of native chemical ligation, combined with what we had learned from our studies on cysteine-free ligation, we developed a new method for utilizing the electron rich auxiliary **21** (Figure 4).^{xliv} The role of the auxiliary, which would be situated at a glycopeptide N-terminus, was to provide a thiol handle that would essentially mimic cysteine.^{xliv} In a manner analogous to native chemical ligation, the thiol auxiliary should join two glycopeptide fragments through an initial transthioesterification, thus enabling the critical S→N acyl transfer to furnish the new amide bond. In this approach, we first dealt with installation of auxiliary **21** onto a glycopeptide N-terminus (**20**) via reductive amination. The acyl donor would be our ortho-disulfide phenolic ester, situated at the C-terminus of a second glycopeptide fragment (**22**). Under reducing conditions, the ortho-thiol phenolic ester would undergo an O→S acyl transfer and the resulting thioester would react with the thiol auxiliary to provide a new thioester (**24**). Subsequent S→N acyl transfer and removal of the auxiliary would provide the desired bidomainal glycopeptide adduct **26**.

Our new cysteine-free ligation method was successfully applied to the synthesis of several complex glycopeptide fragments.^{xliv} In the synthesis of the bidomainal glycopeptide **31**, glycopeptides **27** and **28** are joined together at a glycine-glutamine junction (Scheme 3). Initial disulfide bond reduction of **27** and **28** resulted in formation of the ligated product **29**. The auxiliary was removed through a two-step sequence whereby the free thiol was first methylated to prevent any acid-mediated N→S acyl transfer. Subsequent treatment with TFA resulted in

cleavage of the auxiliary to give the desired glycopeptide adduct **31**. It is appropriate to point out that during the course of our studies, we observed that as the ligation site becomes more hindered, the S→N acyl transfer becomes increasingly more difficult. Consequently, this method, as currently practiced, is best applied to junctions where at least one of the amino acids is glycine or alanine.

In the context of erythropoietin, the synthesis would require the reiterative coupling of multiple glycopeptide and peptide fragments. To demonstrate the utility of our cysteine-free and cysteine-based ligations in reiterative couplings, we identified a tridomainal glycopeptide (**32**) containing three glycosyl units for total synthesis (Figure 5).^{xliv} The target glycopeptide (**32**) would be assembled from the C→N terminal, using our cysteine-free ligation to first form the glycine-glutamine junction. This merger would be followed by a cysteine-based ligation to form the glycine-cysteine junction. The requisite glycopeptide starting materials are **27**, **33**, and **34**, which can be readily prepared using techniques developed in our lab.

The reiterative coupling began with glycopeptides **33** and **34** (Scheme 4). Disulfide bond reduction followed by cysteine-free ligation provided the adduct **35**. Subsequent cleavage of the the Fmoc and thiazolidine groups served to expose the N-terminal cysteine (**36**). The latter could be joined by cysteine-based ligation with glycopeptide **27** to afford the tridomainal glycopeptide **32**.

While our cysteine-free ligation method has thus far been focused on the synthesis of multidomainal glycopeptides, it was appreciated that this method could, in principle, offer a potential alternative to the preparation of cyclic peptides. Compared to their linear counterparts, cyclic peptides often exhibit enhanced specificity, activity, and stability, making them an important class of biologically relevant compounds. Indeed, in contrast to the traditional methods used for cyclic peptide synthesis, Tam and co-workers have shown that native chemical ligation can also be used for this purpose.^{xlvi} Considering the rarity of cysteine in proteins, the development of cysteine-free ligation methods to gain access to cyclic peptides is of particular importance. Our cysteine-free ligation format was thus applied to the synthesis of cyclic peptide **39** (Scheme 5).^{xliv} In starting peptide **38**, the C-terminus features an ortho-disulfide phenolic ester while the N-terminus was equipped with the auxiliary. Upon exposure to reducing conditions, **38** underwent intramolecular ligation to afford the cyclized product **39**.

With the demonstration of our cysteine-free and cysteine-based ligations on model glycopeptide fragments, we were emboldened to apply these methods to fragments from which one could hypothesize about a synthesis of erythropoietin. Two glycopeptide fragments of EPO, Ala²²-Glu³⁷ (**44**) and Ala¹¹⁴-Arg¹⁶⁶ (**48**) were prepared using the cysteine-based ligation and cysteine-free ligation, respectively.^{xxx,xxxi} In the case of Ala²²-Glu³⁷ (**44**), the necessary glycopeptide fragment was assembled via Lansbury aspartylation between dodecasaccharide **40**^{xxx,xlvii} and peptide **41** (Scheme 6). Ligation of glycopeptide **42** and peptide **43** proceeded under reducing conditions to afford the Ala²²-Glu³⁷ glycopeptide domain (**44**) featuring the fucosylated biantennary N-glycan.

It is, of course, noted that the Ala¹¹⁴-Arg¹⁶⁶ (**48**) glycopeptide fragment of EPO contains an O-linked glycoporphin glycan at Ser¹²⁶ (Scheme 7). By utilizing the glycoporphin-presenting serine as a cassette in combination with Fmoc-based solid phase peptide synthesis, the glycopeptide fragment **45** was assembled.^{xxxii} Applying cysteine-free ligation conditions, glycopeptide **45** and peptide **46** were joined to provide **47**. Selective methylation of the free thiol group followed by acidic cleavage of the auxiliary afforded the Ala¹¹⁴-Arg¹⁶⁶ (**48**) glycopeptide fragment, which should be suitable for the EPO target.

At this juncture in our journey, our cysteine-free and cysteine-based ligation methods had enabled the synthesis of multidomain glycopeptides, including the preparation of two glycopeptide fragments of EPO (**44** and **48**). Using these methods, the reiterative coupling of (glyco)peptide segments had also been demonstrated, leading to the synthesis of a trifunctional glycopeptide (**32**). While encouraging, these methods are not without their limitations. As mentioned earlier, during the cysteine-free ligation, the quality of the culminating S→N acyl transfer step is highly dependent on the state of steric hindrance at the ligation site. Moreover, beyond the critical acyl transfer step lies another potential complication: cleavage of the auxiliary under strongly acidic conditions could compromise the vulnerable glycosidic bonds. Furthermore, our glycopeptide/peptide ligations thus far have been limited to a single direction, from the C to the N terminus. The capacity to couple glycopeptide and peptide segments in a reiterative manner could well be of value in attempting to reach EPO by chemical synthesis. A unidirectional synthesis implies that glycopeptide assembly will be limited to a linear and nonconvergent synthesis whereby short (glyco)peptide fragments will be sequentially coupled. By contrast, a convergent synthesis would require the development of methods that enable glycopeptide coupling from the N to the C terminus, thereby requiring accommodation with existing ligation strategies.^{xlvi}

In exploring alternative glycopeptide coupling methods that could circumvent the need for a cysteine or auxiliary based routes (permitting a new coupling direction), we wondered if our ortho-disulfide phenolic ester (i.e., **22**) could serve as a viable acyl donor with a free amine acceptor (e.g., **20**). The fragment condensation studies of Blake and Aimoto, which are mediated by silver ion, cast new mechanistic light on our ortho-disulfide phenolic ester. We were led to speculate whether the uniquely situated ortho-thiol group could facilitate the generation of an enhanced acyl donor (Figure 6).^{xlix} It was thought that in the presence of an additive, such as HOOBt, the metal-activated ortho-thiol phenolic ester could give rise to a reactive acyl donor (**49**). The new acyl donor so generated could undergo fragment coupling with the free N-terminus of a glycopeptide (**20**), wherein the resulting adduct (**26**) would be free of any auxiliary.

Indeed, it was found that the use of AgCl, in combination with HOOBt and Hunig's base, resulted in the merger of the N-linked glycopeptide **50** and peptide **51** to afford the desired product **52** in 80% yield (Scheme 8).¹ It will be noted that in the example shown, the only amino acid side chains that are protected are those of lysine.^{li} The other amino acid side chains, in particular those of aspartic acid and glutamic acid, remained free. This method enables an unprecedented direct fragment coupling in the absence of an N-terminal cysteine or an auxiliary by exploiting the latent acyl donor potential of the ortho-disulfide phenolic ester.

Happily, the enhancement of the acyl donor capacity of our ortho-disulfide phenolic ester is not, however, limited to the use of silver salts. The use of tris(2-carboxyethyl)phosphine (TCEP), a water-soluble phosphine reducing agent, was also effective in activating the ortho-disulfide phenolic ester to permit fragment coupling. As shown in the example below, the O-linked glycopeptide **53** and peptide **54** were coupled to afford ligated glycopeptide **55** under TCEP-mediated conditions (Scheme 9). It will be noted that the fragment couplings in Schemes 8 and 9 occur at C-terminal glycine and proline residues, respectively. As these residues are not vulnerable to epimerization problems, they are currently the preferred C-terminal amino acids for such fragment couplings.

Having two different actuation protocols for the assembly of EPO directed fragments, we explored their joint potential application to reiterative glycopeptide ligations. In practice it proved possible to exploit differences in the reactivity of thioesters during the two protocols. Specifically, under TCEP-mediated activation conditions, C-terminal alkyl thioesters remain intact whereas they are activated under the AgCl protocol, thereby undergoing fragment

coupling.^{lii} The implications of such functional group orthogonality are not insignificant. By maintaining the C-terminal alkyl thioester during a TCEP-mediated fragment coupling, the resulting product can be used in a second fragment coupling mediated by AgCl to enable glycopeptide assembly from the N to the C terminus.

As shown in Scheme 10, glycopeptide **50** is equipped with our C-terminal ortho-disulfide phenolic ester while peptide **56** has a C-terminal alkyl thioester.^l Fragment coupling of the two was accomplished using the TCEP protocol to provide glycopeptide **57**, which has an intact C-terminal alkyl thioester. The resulting glycopeptide product was then subjected to a AgCl-mediated fragment coupling with peptide **51** to afford glycopeptide **58**. The final product, **58**, was prepared from the N to the C terminus using sequential TCEP/AgCl fragment couplings.

A state of the art example of a glycopeptide assembly utilizing the sequential TCEP/AgCl fragment couplings is seen in the synthesis of erythropoietin glycopeptide domain Gln⁷⁸-Arg¹⁶⁶ (**63**, Scheme 11).^{xxxii} This fragment, prepared in homogeneous form, contains both the N-linked dodecasaccharide at Asn⁸³ and the O-linked glycoporphin at Ser¹²⁶. A TCEP-mediated fragment coupling between glycopeptide **59** and peptide **60** resulted in the formation of erythropoietin glycopeptide domain Gln⁷⁸-Gly¹¹³ (**61**). In the light of its exploitable linkage, C-terminal alkyl thioester **61** was treated with AgCl in the presence of HOObt and *i*Pr₂NEt to react with glycopeptide **62**, thereby furnishing the erythropoietin Gln⁷⁸-Arg¹⁶⁶ glycopeptide domain **63**, in 30% yield.

By exploiting differences in the acyl donor capacities of alkyl thioesters and ortho-disulfide phenolic ester thioesters, the TCEP and AgCl protocols were successfully incorporated in reiterative glycopeptide couplings. We recognized that viable acyl donors need not necessarily be limited to thioesters. Accordingly, we explored the use of activated oxo-esters in peptide ligations. It was found that C-terminal *para*-nitrophenol esters^{liiii} are highly effective acyl donors, undergoing ligation with peptides containing N-terminal cysteine in good yield and in relatively short reaction times with minimal racemization (Scheme 12).^{liv} In one study, the *para*-nitrophenol ester mediated ligation proceeded at a significantly faster rate than the corresponding native chemical ligation, demonstrating the enhanced reactivity of the *para*-nitrophenol ester in these ligations. However, competing hydrolysis was a complication. Fortunately, such undesired hydrolysis could be suppressed by addition of *para*-nitrophenol and application to more hindered C-terminal residues. Most notably, studies with the very hindered unnatural amino acid penicillamine (**68**) and C-terminal *para*-nitrophenol esters also resulted in successful ligations (Scheme 13). By analogy to chemistry used in alanine ligation as described below, the penicillamine residue could be transformed to valine (**69**→**70**) via radical desulfurization (vide infra). For instance, ligation product **69** was exposed to thiol reduction conditions, converting the penicillamine residue to the corresponding valine residue in **70**.

During our studies on the synthesis of the erythropoietin Cys²⁹-Gly⁷⁷ glycopeptide domain (**76**), it was found that the *para*-nitrophenyl ester was actually too reactive as an acyl donor. This high reactivity had to be attenuated to avoid unwanted side reactions in the critical aspartylation reaction.^{xxxiii} In the context of efforts to tune the reactivity of the acyl donor, it was found that *para*-cyanophenyl esters manifested decreased but still sufficient reactivity for our requirements. As shown in Scheme 14, the aspartylation reaction between dodecasaccharide **40** and peptide **71** had afforded the desired glycopeptide fragment **72** in 33% yield. Direct aminolysis of the C-terminal *para*-cyanophenyl ester in glycopeptide **72** with peptide **73** provided the erythropoietin Cys²⁹-Gly⁵⁷ glycopeptide domain (**74**). It was noted that the C-terminal masked thioester within peptide **73** possesses two ortho-functional groups, a disulfide group *and* a propyl group. The propyl substitution was introduced to suppress undesired hydrolysis during the aminolysis reaction, thus permitting the synthesis of the

Cys²⁹-Gly⁵⁷ glycopeptide (**74**) while retaining the masked thioester. This example serves to demonstrate how subtle adjustments can modulate the capacity of an acyl donor. Indeed, a TCEP-mediated fragment coupling between glycopeptide **74** and peptide **75** afforded the potential erythropoietin Cys²⁹-Gly⁷⁷ glycopeptide domain **76**. The reader will note that this fragment possesses a C-terminal alkyl thioester that could well serve as a means for future glycopeptide coupling during the assembly of erythropoietin.

The strategies, methods, and tools discussed in the previous sections focused on circumventing cysteine dependence in various assemblies. As described, we investigated the use of auxiliaries as a cysteine mimic and developed a modified Blake-Aimoto direct fragment coupling; both of these methods have been applied to the reiterative synthesis of complex glycopeptide fragments. An alternative to these methods that also addresses the problem of cysteine dependence is the direct reduction of the cysteine sidechain to an alanine sidechain, following native chemical ligation. This idea was first proposed by Yan and Dawson and utilized primarily Raney nickel or Pd/Al₂O₃ to effect the desired cysteine reduction.^{lv} Essentially, cysteine serves as a surrogate for alanine, which occurs with considerably greater frequency than cysteine in peptides and proteins.^{xlvi} The potential opportunities presented by this approach are significant and the groups of Kent^{lvii} and Wong^{lviii} have also made contributions. However, given our interest in the synthesis of complex glycoproteins, which often contain extensive and sensitive functional groups, we were interested in developing a mild and metal-free reduction for the conversion of cysteine to alanine.

After taking note of a 1956 disclosure by Hoffmann describing a desulfurization reaction between mercaptan and trialkylphosphite,^{lxviii} in a totally different context, we explored the use of trialkylphosphines, in particular TCEP, to effect a cysteine reduction.^{lix} TCEP was an attractive choice for several reasons. Thus, it can tolerate a range of glycopeptide functionality, it is water soluble, and it is easily manipulated in air. In conjunction with the water soluble 2,2'-azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride (VA-044) and *t*-butylthiol, TCEP proved to be a very effective reducing agent. As shown in Scheme 15, glycopeptide **77**, which has an N-terminal thiazolidine, an Ac_m-protected cysteine, and our C-terminal ortho-thiophenolic ester, was coupled with peptide **78**, which has a methionine, a C-terminal thioester, and an N-terminal cysteine, in a kinetically controlled ligation.^{lxviii} The resulting peptide (**79**) was subjected to the desulfurization conditions described above, thereby accomplishing clean reduction of the free cysteine to alanine and giving glycopeptide **80**. It will be noted that the sulfur atoms within the thiazolidine, Ac_m-protected cysteine, methionine, and C-terminal thioester are all intact. The reduction protocol is both mild and versatile, tolerating all thiol-containing groups as well as oligosaccharide domains.

Given the versatility of this method, we anticipated that it could accommodate the different functional groups present within the erythropoietin Ala¹-Gly²⁸ glycopeptide domain (**87**).^{xxxiv} This glycopeptide fragment lacks any exploitable cysteine and glycine/proline residues, thus precluding the use of either NCL or direct condensation methods for its assembly. However, the presence of several alanine residues in close proximity to the glycosylation site suggested that native chemical ligation/radical desulfurization could address the synthetic challenges of this fragment. As shown in Scheme 16, peptide **81** and glycopeptide **82** were joined via native chemical ligation to give **83**. In situ thiolactone formation during the reaction resulted in formation of **84**. Removal of the allyl protecting group followed by opening of the thiolactone with thiopropionic acid provided glycopeptide **86**, in which the cysteine side chain is revealed for desulfurization. Treatment of **86** with VA-044, TCEP, and thiopropionic acid in buffered conditions at 37 °C cleanly afforded the reduced product **87** in 67% yield. The use of thiopropionic acid as the radical propagator also served to open any thiolactone produced during the reaction. The erythropoietin Ala¹-Gly²⁸ glycopeptide domain, **87**, features both the

biantennary glycan and a C-terminal α -thioester, two critical features necessary for the convergent preparation of synthetic homogeneous EPO.

Following our results in “alanine ligation,” native chemical ligation/metal-based thiol reduction strategies have also been developed to enable ligations at phenylalanine residues.^{lx} In principle, the native chemical ligation strategy could be applied to any amino acid site by introducing a sulfhydryl group within the sidechain of the desired N-terminal residue. Following ligation and desulfurization, the natural amino acid is then revealed at the ligation site. This logic led us to develop a ligation at valine, which is a relatively abundant amino acid.^{lxiii} We selected γ -thiol valine to serve as the valine precursor,^{lxii} instead of the hindered penicillamine.^{lxii} As shown in Scheme 17, peptide **89** is equipped with a protected γ -thiol valine residue at the N-terminus and peptide **88** has a C-terminal ortho-thiophenolic ester. Under native chemical ligation conditions, the two are joined through transthioesterification followed by S \rightarrow N acyl transfer to afford **90**. Radical desulfurization using our TCEP/VA-044/*t*BuSH protocol at 37 °C provided peptide **91** with the natural valine residue at the ligation site.

Another outcome of our interest in amide bond formation driven by our EPO journey was a revisitation of the chemistry of isonitriles. Our isonitrile program has already matured into a field of inquiry of its own. Here we merely recapitulate for the reader some of our introductory findings. In particular we developed what we call the two-component coupling between carboxylic acids and isonitriles to give an N-formyl amide.^{lxiii} It is proposed that under microwave heating, carboxylic acid **92** and isonitrile **93** react to give rise to formimidate carboxylate mixed anhydride **94**. The latter then undergoes a 1,3-O \rightarrow N acyl transfer giving rise to the observed N-formyl amide **95** (Figure 7). The utility of this two-component coupling in building asparagine-linked glycopeptides was explored (Scheme 18). The reactions between glycosyl isonitriles and aspartic acid were anomericly specific, that is, the respective α - and β -isonitriles (**99** and **96**) produce, correspondingly, the α - and β -N-linked glycosyl amino acids (**100** and **98**). It is important to note that the C2 functional group must be non-participatory, or, as shown in equation 3 (Scheme 18), a highly reactive β -GlcNAc donor can form that results in ester bond formation (**101** \rightarrow **102**).

The N-formyl amide unit is a highly versatile functional group (Figure 8). It can be cleaved under basic conditions to provide the corresponding secondary amide (**98** \rightarrow **108**). Even more interesting is its reduction giving rise to an N-methyl amide via the hydroxymethyl intermediate (**98** \rightarrow **107**). Alternatively, the N-formyl group undergoes a Lewis-acid induced nucleophilic aminomethylation reaction with allyl trimethylsilane (**104** \rightarrow **103**). These transformations provide novel routes to tertiary amides that are otherwise only difficultly accessible. Moreover, the resulting N-formyl mixed imide can be readily converted to dihydropyridones (**105** \rightarrow **106**). The isonitrile based methods briefly described above have sparked major new opportunities to generate a range of amidic subtypes. This work will be described in a separate treatment.

CONCLUSION

In this account, we have provided some of the key findings that have brought us to the threshold of realizing our impractical goal of 2003, i.e., a total synthesis of homogeneous erythropoietin. Obviously, there are still significant remaining issues. It may well be necessary to develop new chemistry to deal with additional difficulties as they arise. Nonetheless, it is fair to say that the long term goal of a total synthesis of erythropoietin has encouraged exploration of new ideas in amide-directed acylation. Many of these discoveries are likely to prove applicable to other biologic-level targets. In that sense, it can well be argued that the EPO journey has already been a success.

Biography



Cindy Kan began her chemistry training as an undergraduate at Barnard College. While at Barnard, she worked with Professor Christian M. Rojas on the use of acyl azides in amidoglycosylation reactions. She then went on to Stanford University to pursue her doctoral studies under the guidance of Professor Paul A. Wender. Her graduate work focused on the development and synthesis of bryostatins analogs as well as synthetic studies on daphnane diterpene orthoester natural products. After receiving her doctoral degree, she returned to New York City for a postdoctoral position with Professor Samuel J. Danishefsky at Memorial Sloan-Kettering Cancer Center. Her work in the Danishefsky lab has centered on the chemical synthesis of homogeneous glycoproteins.



Samuel J. Danishefsky received his B.S. degree at Yeshiva University. He did his graduate research under the late Professor Peter Yates. He then joined the laboratory of Professor Gilbert Stork at Columbia University as an NIH Postdoctoral Associate. His first academic position was at the University of Pittsburgh, where he joined as Assistant Professor in 1963. He was promoted to Associate Professor, Professor, and University Professor. In January 1980, he moved to Yale University and was named Eugene Higgins Professor in 1981. Appointed by President A. Bartlett Giamatti as Chairman of the Department of Chemistry, he served until 1987. He became Sterling Professor at Yale in 1990. In 1993, Professor Danishefsky moved back to New York as Professor of Chemistry at Columbia University and the Eugene Kettering Chair and Head of the Laboratory of Bioorganic Chemistry at Memorial Sloan-Kettering Cancer Center. In 1996, he shared the Wolf Prize in Chemistry with Professor Gilbert Stork. In 2006, he was the recipient of the Franklin Medal in Chemistry, the Bristol Myers Squibb Lifetime Achievement Award in Chemistry, and the National Academy of Sciences Award in the Chemical Sciences.

References

- i. a Zhu J, Warren JD, Danishefsky S. J. *Expert Rev. Vaccines*. 2009 manuscript accepted. b Warren JD, Geng XD, Danishefsky SJ. *Top. Curr. Chem* 2007;267:109–141. c Wilson RM, Danishefsky SJ. *Pure Appl. Chem* 2007;79:2189–2216.
- ii. a Danishefsky SJ, Kerwin JF, Kobayashi S. J. *Am. Chem. Soc* 1982;104:358–360. b Danishefsky SJ, Kato N, Askin D, Kerwin JF. *J. Am. Chem. Soc* 1982;104:360–362.
- iii. Danishefsky SJ, Kitahara T. *J. Am. Chem. Soc* 1974;96:7807–7808.
- iv. For a review, see: Danishefsky SJ, Bilodeau MT. *Angew. Chem., Int. Ed* 1996;35:1380–1419.
- v. For a review, see: Seeberger PH, Bilodeau MT, Danishefsky SJ. *Aldrichim. Acta* 1997;30:75–92.
- vi. a Slovin SF, Ragupathi G, Adluri S, Ungers G, Terry K, Kim S, Spassova M, Bornmann WG, Fazzari M, Dantis L, Olkiewicz K, Lloyd KO, Livingston PO, Danishefsky SJ, Scher HI. *Proc. Natl. Acad. Sci. U.S.A* 1999;96:5710–5715. [PubMed: 10318949] b Gilewski T, Ragupathi G, Bhuta S, Williams LJ, Musselli C, Zhang X-F, Bencsath KP, Panageas KS, Chin J, Hudis CA, Norton L, Houghton AN, Livingston PO, Danishefsky SJ. *Proc. Natl. Acad. Sci. U.S.A* 2001;98:3270–3275. [PubMed: 11248068] c Krug LM, Ragupathi G, Hood C, Kris MG, Miller VA, Allen JR, Keding SJ, Danishefsky SJ, Gomez J, Tyson L, Pizzo B, Baez V, Livingston PO. *Clin. Cancer Res* 2004;10:6094–6100. [PubMed: 15447995] d Sabbatini PJ, Kudryashov V, Ragupathi G,

- Danishefsky SJ, Livingston PO, Bornmann W, Spassova M, Zatorski A, Spriggs D, Aghajanian C, Soignet S, Peyton M, O'Flaherty C, Curtin J, Lloyd KO. *Int. J. Cancer* 2000;87:79–85. [PubMed: 10861456] e Slovin SF, Ragupathi G, Musselli C, Fernandez C, Diani M, Verbel D, Danishefsky SJ, Livingston P, Scher HI. *Cancer Immunol. Immunother* 2005;54:694–702. [PubMed: 15726361] f Sabbatini PJ, Ragupathi G, Hood C, Aghajanian CA, Juertzka M, Lasonos A, Hensley ML, Spassova MK, Ouerfelli O, Spriggs DR, Tew WP, Konner J, Clausen H, Abu Rustum N, Danishefsky SJ, Livingston PO. *Clin. Cancer Res* 2007;13:4170–4177. [PubMed: 17634545]
- vii. For examples, see: a Wang Z-G, Zhang X, Visser M, Live D, Zatorski A, Iserloh U, Lloyd KO, Danishefsky SJ. *Angew. Chem., Int. Ed* 2001;40:1728–1732. b Wang Z-G, Zhang X, Live D, Danishefsky SJ. *Angew. Chem., Int. Ed* 2000;39:3652–3656.
- viii. Dudkin VY, Miller JS, Danishefsky SJ. *J. Am. Chem. Soc* 2004;126:736–738. [PubMed: 14733546]
- ix. a Mandal M, Dudkin VY, Geng XD, Danishefsky SJ. *Angew. Chem., Int. Ed* 2004;43:2557–2561. b Geng XD, Dudkin VY, Mandal M, Danishefsky SJ. *Angew. Chem., Int. Ed* 2004;43:2562–2565.
- x. For reviews, see: a Dwek RA. *Chem. Rev* 1996;96:683–720. [PubMed: 11848770] b Wold F. *Annu. Rev. Biochem* 1981;50:783–814. [PubMed: 6791580] c Roth J. *Chem. Rev* 2002;102:285–304. [PubMed: 11841244] d Ohtsubo K, Marth JD. *Cell* 2006;126:855–867. [PubMed: 16959566] e Solá RJ, Rodríguez-Martínez JA, Griebenow K. *Cell. Mol. Life Sci* 2007;64:2133–2152. [PubMed: 17558468] f Mitra N, Sinha S, Ramya TNC, Surolia A. *Trends Biochem. Sci* 2006;31:156–163. [PubMed: 16473013]
- xi. a Spiro RG. *Glycobiology* 2002;12:43R–56R. b Rudd PM, Dwek RA. *Crit. Rev. Biochem. Mol. Biol* 1997;32:1–100. [PubMed: 9063619]
- xii. a Kornfeld R, Kornfeld S. *Annu. Rev. Biochem* 1985;54:631–664. [PubMed: 3896128] b Roth J. *Chem. Rev* 2002;102:285–303. [PubMed: 11841244] c Rudd PM, Dwek RA. *Crit. Rev. Biochem. Mol. Biol* 1997;32:1–100. [PubMed: 9063619]
- xiii. For recent reviews, see: a Rich JR, Withers SG. *Nat. Chem. Biol* 2009;5:206–215. [PubMed: 19295526] b Gamblin DP, Scanlan EM, Davis BG. *Chem. Rev* 2009;109:131–163. [PubMed: 19093879]
- xiv. Live DH, Wang Z-G, Iserloh U, Danishefsky SJ. *Org. Lett* 2001;3:851–854. [PubMed: 11263898]
- xv. For general reviews on EPO, see: a Sytkowski, AJ. *Erythropoietin*. Wiley-VCH Verlag GmbH & Co. KGaA; Weinheim: 2004. b Jelkmann W. *Intern. Med* 2004;43:649–659. [PubMed: 15468961]
- xvi. a Lai P-H, Everett R, Wang F-F, Arakawa T, Goldwasser E. *J. Biol. Chem* 1986;261:3116–3121. [PubMed: 3949763] b Tsuda E, Goto M, Murakami A, Akai K, Ueda M, Kawanishi G. *Biochemistry* 1988;27:5646–5654. [PubMed: 3179269] c Takeuchi M, Takasaki S, Miyazaki H, Kato T, Hoshi S, Kochibe N, Kobata A. *J. Biol. Chem* 1988;263:3657–3663. [PubMed: 3346214]
- xvii. Jelkmann W. *Physiol. Rev* 1992;72:449–489. [PubMed: 1557429]
- xviii. a Marti HH, Wenger RH, Rivas LA, Straumann U, Digicaylioglu M, Henn V, Yonekawa Y, Bauer C, Gassmann M. *Eur. J. Neurosci* 1996;8:666–676. [PubMed: 9081618] b Juul SE, Yachnis AT, Rojiani AM, Christensen RD. *Pediatr. Dev. Pathol* 1999;2:148–158. [PubMed: 9949221]
- xix. a Anagnostou A, Lee ES, Kessimian N, Levinson R, Steiner M. *Proc. Natl. Acad. Sci. U.S.A* 1990;87:5978–5982. [PubMed: 2165612] b Anagnostou A, Liu Z, Steiner M, Chin K, Lee ES, Kessimian N, Noguchi CT. *Proc. Natl. Acad. Sci. U.S.A* 1994;91:3974–3978. [PubMed: 8171022]
- xx. For the first purification of EPO, see: a Miyake T, Kung CKH, Goldwasser E. *J. Biol. Chem* 1977;252:5558–5564. [PubMed: 18467] For an improved purification procedure of EPO, see: b Krystal G, Pankratz HRC, Farber NM, Smart JE. *Blood* 1986;67:71–79. [PubMed: 3753581]
- xxi. Spivak JL. *Annu. Rev. Med* 1993;44:243–253. [PubMed: 8476246]
- xxii. a Spival JL. *Seminars Oncology* 1998;25(S7):7–11. b Cazzola M, Mercuriali F, Brugnara C. *Blood* 1997;89:4248–4267. [PubMed: 9192747]
- xxiii. Higuchi M, Masayoshi O, Kuboniwa H, Tomonoh K, Shimonaka Y, Ochi N. *J. Biol. Chem* 1992;267:7703–7709. [PubMed: 1560005]
- xxiv. a Egrie JC, Grant JR, Gillies DK, Aoki KH, Strickland TW. *Glycoconjugate J* 1993;10:263. b Egrie JC, Browne JK. *Br. J. Cancer* 2001;84(S1):3–10. and references therein. [PubMed: 11308268]
- xxv. Kochendoerfer GG, Chen S-Y, Mao F, Cressman S, Traviglia S, Shao H, Hunter CL, Low DW, Cagle EN, Carnevali M, Gueriguian V, Keogh PJ, Porter H, Stratton SM, Wiedeke MC, Wilken J, Tang J, Levy JJ, Miranda LP, Crnogorac MM, Kalbag S, Botti P, Schindler-Horvat J, Savatski L,

- Adamson JW, Kung A, Kent SBH, Bradburne JA. *Science* 2003;299:884–887. [PubMed: 12574628]
- xxvi. Glycoprotein fragments and glycoproteins bearing complex carbohydrates have been prepared via chemoenzymatic approaches. For examples, see: a O'Connor SE, Imperiali B. *Chem. Biol* 1996;3:803–812. [PubMed: 8939697] b Haneda K, Inazu T, Mizuno M, Iguchi R, Yamamoto K, Kumagai H, Aimoto S, Suzuki H, Noda T. *Bioorg. Med. Chem. Lett* 1998;8:1303–1306. [PubMed: 9871755] c Unverzagt C. *Tetrahedron Lett* 1997;38:5627–5630. d Leppänen A, Mehta P, Ouyang Y-B, Ju T, Helin J, Moore KL, van Die I, Canfield WM, McEver RP, Cummings RD. *J. Biol. Chem* 1999;274:24838–24848. [PubMed: 10455156] e Leppänen A, White SP, Helin J, McEver RP, Cummings RD. *J. Biol. Chem* 2000;275:39569–39578. [PubMed: 10978329] f Koeller KM, Smith MEB, Wong C-H. *Bioorg. Med. Chem* 2000;8:1017–1025. [PubMed: 10882013] g Koeller KM, Smith MEB, Huang R-F, Wong C-H. *J. Am. Chem. Soc* 2000;122:4241–4242. h Li B, Song H, Hauser S, Wang L-X. *Org. Lett* 2006;8:3081–3084. [PubMed: 16805557]
- xxvii. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. *Science* 1994;266:776–779. [PubMed: 7973629]
- xxviii. For the first example of glycoprotein chemical synthesis using NCL, see: a Shin Y, Winans KA, Backes BJ, Kent SBH, Ellman JA, Bertozzi CR. *J. Am. Chem. Soc* 1999;121:11684–11689. For recent examples of protein synthesis using NCL, see: b Torbeev VY, Kent SBH. *Angew. Chem. Int. Ed* 2007;46:1667–1670. c Durek T, Torbeev VY, Kent SBH. *Proc. Natl. Acad. Sci. U.S.A* 2007;104:4846–4851. [PubMed: 17360367] For a review on NCL, see d Dawson PE, Kent SBH. *Annu. Rev. Biochem* 2000;69:923–960. [PubMed: 10966479]
- xxix. Note that the sialyloligosaccharide was isolated from hen's egg yolk and does not contain any fucose unit. Yamamoto N, Tanabe Y, Okamoto R, Dawson PE, Kajihara Y. *J. Am. Chem. Soc* 2008;130:501–510. [PubMed: 18085777]
- xxx. Wu B, Hua Z, Warren JD, Ranganathan K, Wan Q, Chen G, Tan Z, Chen J, Endo A, Danishefsky SJ. *Tetrahedron Lett* 2006;47:5577–5579.
- xxxi. Chen J, Chen G, Wu B, Wan Q, Tan Z, Hua Z, Danishefsky SJ. *Tetrahedron Lett* 2006;47:8013–8016. [PubMed: 17426819]
- xxxii. Tan Z, Shang S, Halkina T, Yuan Y, Danishefsky SJ. *J. Am. Chem. Soc* 2009;131:5424–5431. [PubMed: 19334683]
- xxxiii. Yuan Y, Chen J, Wan Q, Tan Z, Chen G, Kan C, Danishefsky SJ. *J. Am. Chem. Soc* 2009;131:5432–5437.
- xxxiv. Kan C, Trzupke JD, Wu B, Wan Q, Chen G, Tan Z, Yuan Y, Danishefsky SJ. *J. Am. Chem. Soc* 2009;131:5438–5443. [PubMed: 19334679]
- xxxv. Okamoto R, Kajihara Y. *Angew. Chem. Int. Ed* 2008;47:5402–5406.
- xxxvi. Mezzato S, Schaffrath M, Unverzagt C. *Angew. Chem. Int. Ed* 2005;44:1650–1654.
- xxxvii. Yamamoto N, Takayanagi A, Yoshino A, Sakakibara T, Kajihara Y. *Chem. Eur. J* 2007;13:613–625.
- xxxviii. Hojo H, Haginoya E, Matsumoto Y, Nakahara Y, Nabeshima K, Toole BP, Watanabe Y. *Tetrahedron Lett* 2003;44:2961–2964.
- xxxix. Cohen-Anisfeld ST, Lansbury PT Jr. *J. Am. Chem. Soc* 1993;115:10531–10537.
- xl. Miller JS, Dudkin VY, Lyon GJ, Muir TW, Danishefsky SJ. *Angew. Chem., Int. Ed* 2003;42:431–434.
- xli. Warren JD, Miller JS, Keding SJ, Danishefsky SJ. *J. Am. Chem. Soc* 2004;126:6576–6578. [PubMed: 15161285]
- xlii. McCaldon P, Argos P. *Proteins* 1988;4:99–122. [PubMed: 3227018]
- xliii. Chen G, Warren JD, Chen J, Wu B, Wan Q, Danishefsky SJ. *J. Am. Chem. Soc* 2006;128:7460–7462. [PubMed: 16756298]
- xliv. Wu B, Chen J, Warren JD, Chen G, Hua Z, Danishefsky SJ. *Angew. Chem., Int. Ed* 2006;45:4116–4125.
- xlv. a Offer J, Boddy CNC, Dawson PE. *J. Am. Chem. Soc* 2002;124:4624–4626. b Offer J, Dawson PE. *Org. Lett* 2000;2:23–26. [PubMed: 10814236]
- xlvi. Zhang L, Tam JP. *J. Am. Chem. Soc* 1997;119:2363–2370.

- xlvi. a Nagorny P, Fasching P, Li X, Chen G, Aussedat B, Danishefsky SJ. *J. Am. Chem. Soc* 2009;131:5792–5799. [PubMed: 19341309] b Wu B, Tan Z, Chen G, Chen J, Hua Z, Wan Q. *Tetrahedron Lett* 2006;47:8009–8011.
- lviii. Kent has developed a cysteine-based kinetic chemical ligation method that enables peptide coupling from the N to the C terminus; see: a Bang D, Pentelute BL, Kent SBH. *Angew. Chem., Int. Ed* 2006;45:3985–3988. b Torbeev VY, Kent SBH. *Angew. Chem., Int. Ed* 2007;46:1667–1670. c Durek T, Torbeev VY, Kent SBH. *Proc. Natl. Acad. Sci. USA* 2007;104:4846–4851. [PubMed: 17360367]
- lix. a Blake J. *Int. J. Pept. Protein Res* 1981;17:273–274. [PubMed: 7228502] b Aimoto S, Mizoguchi N, Hojo H, Yoshimura S. *Bull. Chem. Soc. Jpn* 1989;62:524–531. c Aimoto S. *Biopolymers* 1999;51:247–265. [PubMed: 10618594]
- l. Chen G, Wan Q, Tan Z, Kan C, Hua Z, Ranganathan K, Danishefsky SJ. *Angew. Chem., Int. Ed* 2007;46:7383–7387.
- li. The side chain of cysteine also requires protection. For example, see preceding reference.
- lii. The difference in thioester reactivity was also reported by Kent. For references, see: a Johnson EC, Kent SBH. *J. Am. Chem. Soc* 2006;128:6640–6646. [PubMed: 16704265] b reference ⁴⁸
- liii. Bodanszky M. *Nature* 1955;175:685. [PubMed: 14370194]
- liv. Wan Q, Chen J, Yuan Y, Danishefsky SJ. *J. Am. Chem. Soc* 2008;130:15814–15816. [PubMed: 18855357]
- lv. Yan LZ, Dawson PE. *J. Am. Chem. Soc* 2001;123:526–533. [PubMed: 11456564]
- lvi. a Pentelute B, Kent SBH. *Org. Lett* 2007;9:687–690. [PubMed: 17286375] b Johnson ECB, Malito E, Shen Y, Rich D, Tang W, Kent SBH. *J. Am. Chem. Soc* 2007;129:11480–11490. [PubMed: 17705484]
- lvii. a Brik A, Yang Y-Y, Ficht S, Wong C-H. *J. Am. Chem. Soc* 2006;128:5626–5627. [PubMed: 16637616] b Brik A, Ficht S, Yang Y-Y, Bennett CS, Wong C-H. *J. Am. Chem. Soc* 2006;128:15026–15033. [PubMed: 17105315] c Yang Y-Y, Ficht S, Brik A, Wong C-H. *J. Am. Chem. Soc* 2007;129:7690–7701. [PubMed: 17523644]
- lviii. Hoffmann FW, Ess RJ, Simmons TC, Hanzel RS. *J. Am. Chem. Soc* 1956;78:6414.
- lix. Wan Q, Danishefsky SJ. *Angew. Chem., Int. Ed* 2007;46:9248–9252.
- lx. a Botti, P.; Tchertchian, S. p. 2006WO 133962 b Crich D, Banerjee A. *J. Am. Chem. Soc* 2007;129:10064–10065. [PubMed: 17658806]
- lxi. Chen J, Wan Q, Yuan Y, Zhu J, Danishefsky SJ. *Angew. Chem., Int. Ed* 2008;47:8521–8524.
- lxii. A valine ligation has also been accomplished using penicillamine as the valine precursor. See: Haase C, Rohde H, Seitz O. *Angew. Chem. Int. Ed* 2008;47:6807–6810.
- lxiii. a Li X, Danishefsky SJ. *J. Am. Chem. Soc* 2008;130:5446–5448. [PubMed: 18370392] b Li X, Yuan Yu, Berkowitz WF, Todaro LJ, Danishefsky SJ. *J. Am. Chem. Soc* 2008;130:13222–13224. [PubMed: 18783224] c Li X, Yuan Y, Kan C, Danishefsky SJ. *J. Am. Chem. Soc* 2008;130:13225–13227. [PubMed: 18783225]

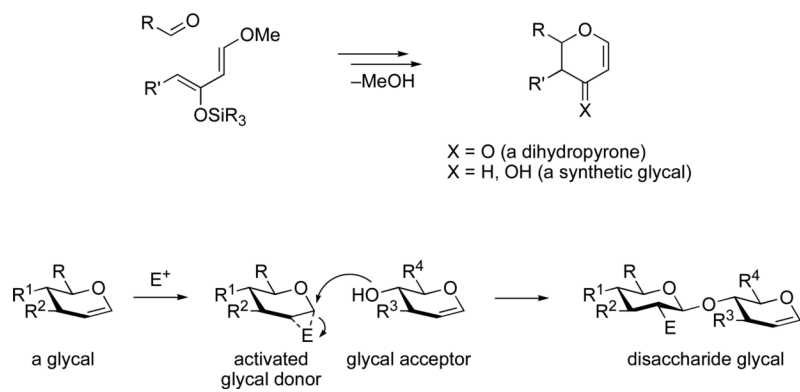


Figure 1.
Glycal assembly

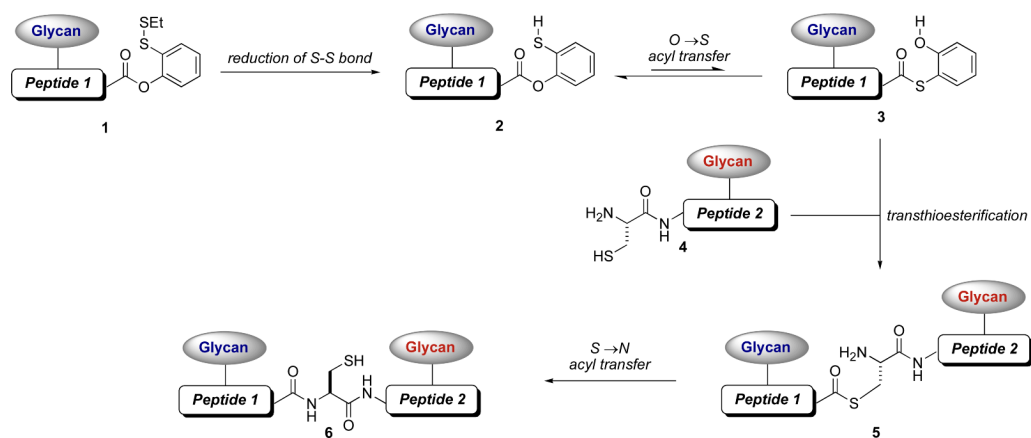
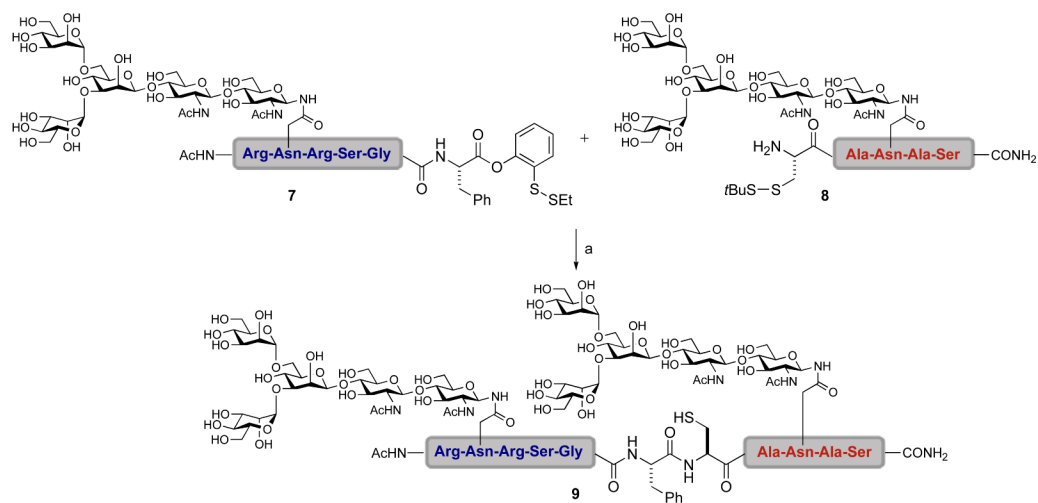


Figure 2. Strategy for glycopeptide-glycopeptide coupling using a masked thioester



^a Reagents and conditions: (a) MesNa (excess), 0.2 M Phosphate, 0.2 M NaCl, pH = 7.4, 75%.

Scheme 1.

^a Glycopeptide-glycopeptide coupling using a ortho-disulfide phenolic ester as a latent acyl donor

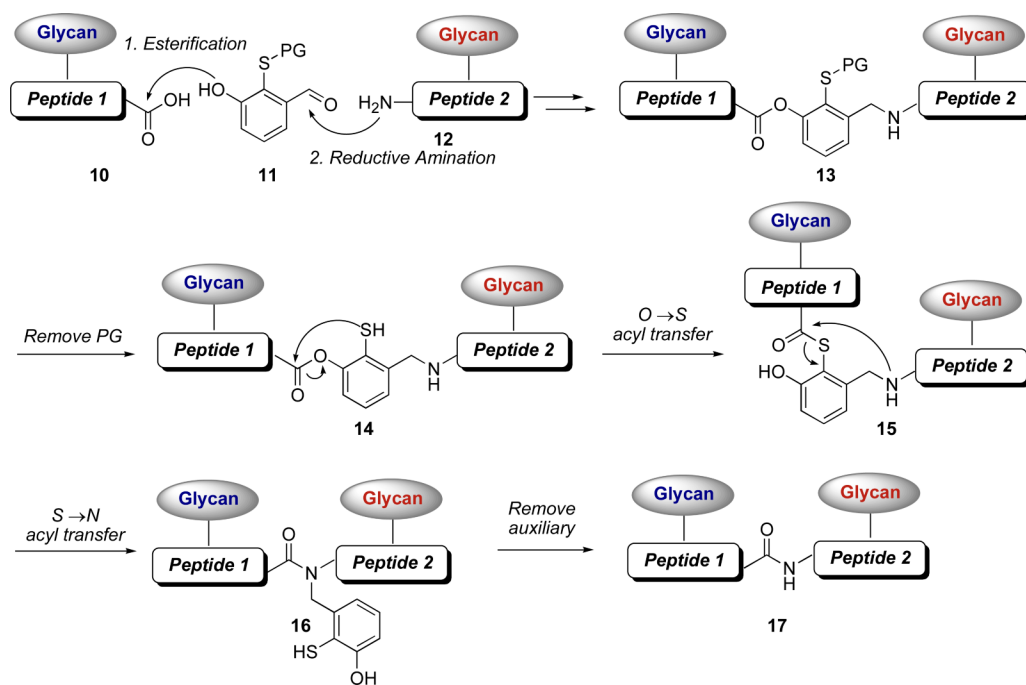
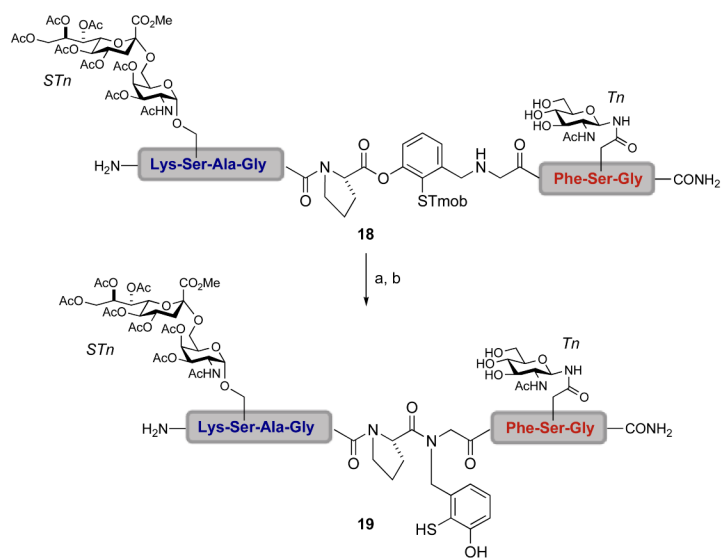


Figure 3.
Strategy for a cysteine-free glycopeptide-glycopeptide ligation



^a Reagents and conditions: (a) 50% TFA/0.5% TIPS, CH₂Cl₂; (b) MeOH, Na₂HPO₄, 72% over two steps.

Scheme 2.

^a Synthesis of bifunctional glycopeptide **19** using a cysteine-free ligation strategy

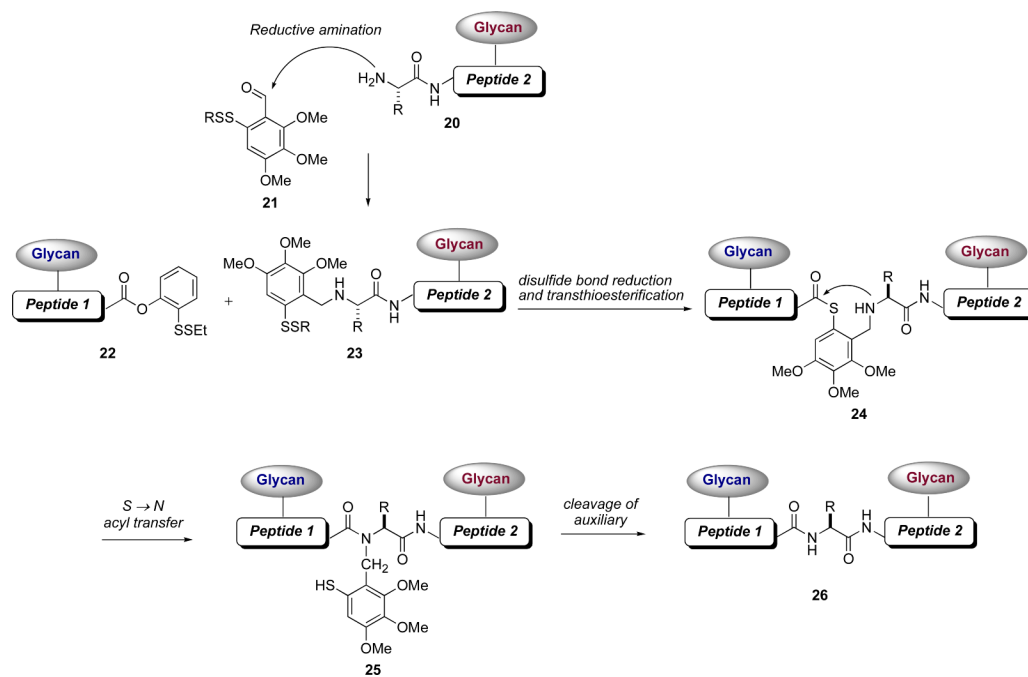
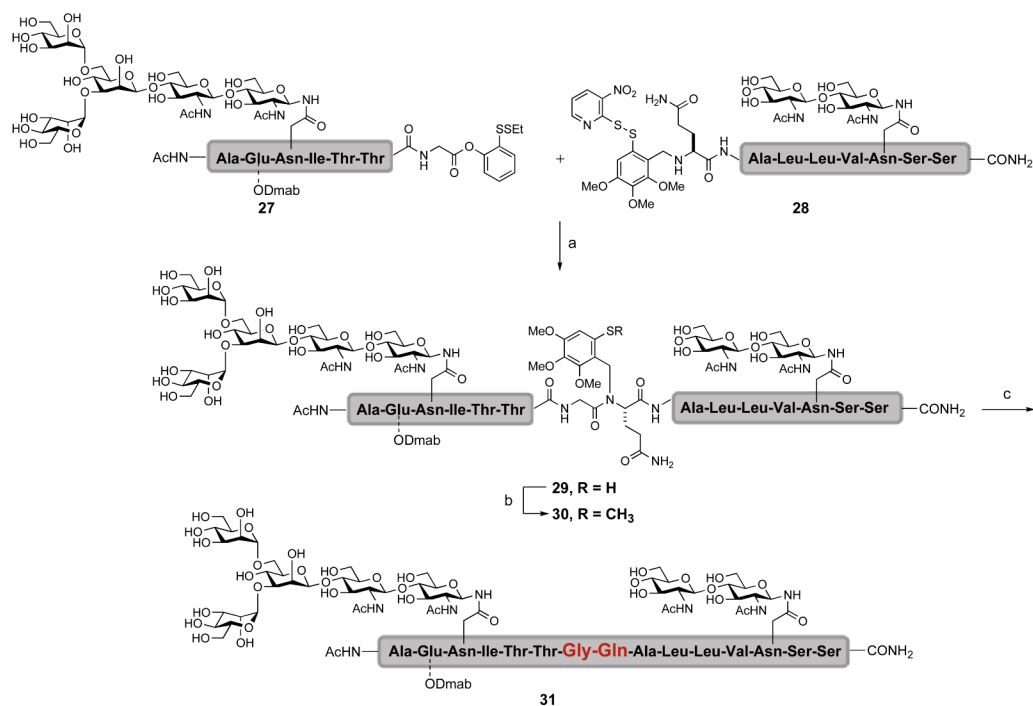


Figure 4.
Second-generation cysteine-free glycopeptide-glycopeptide ligation strategy



^a Reagents and conditions: (a) TCEP, DMF, Na₂HPO₄, 32 °C, 54%; (b) methyl *p*-nitrobenzene sulfonate;

(c) 95% TFA.

Scheme 3.

^a Synthesis of glycopeptide **31** using a cysteine-free ligation strategy

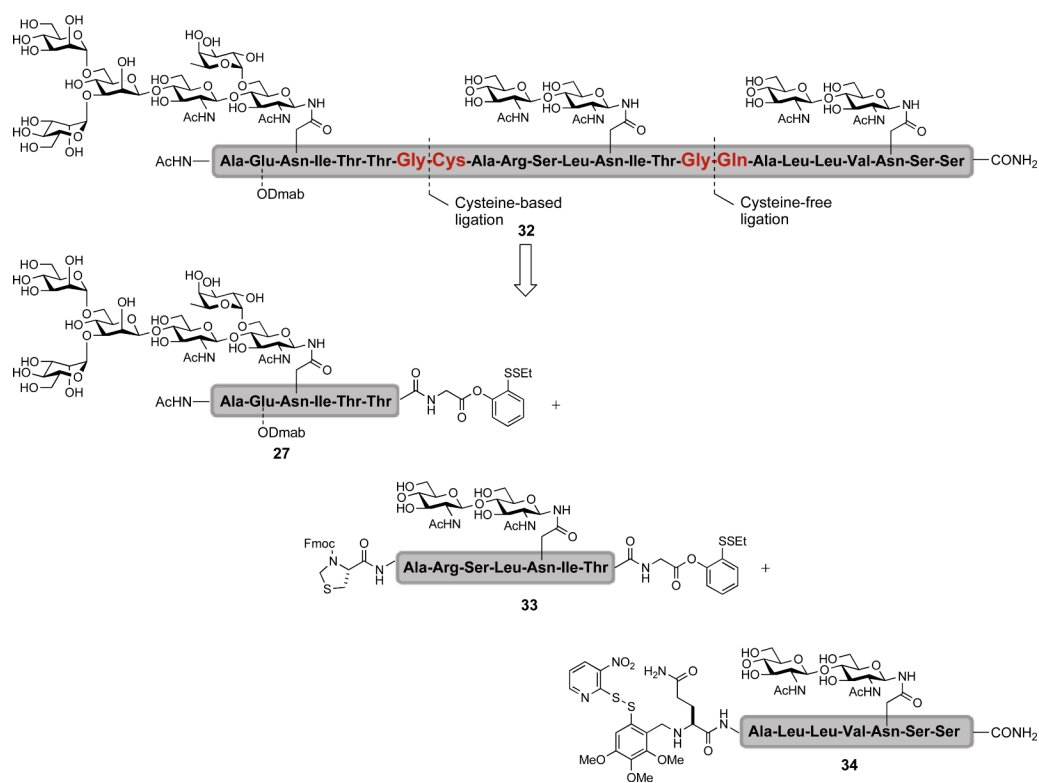
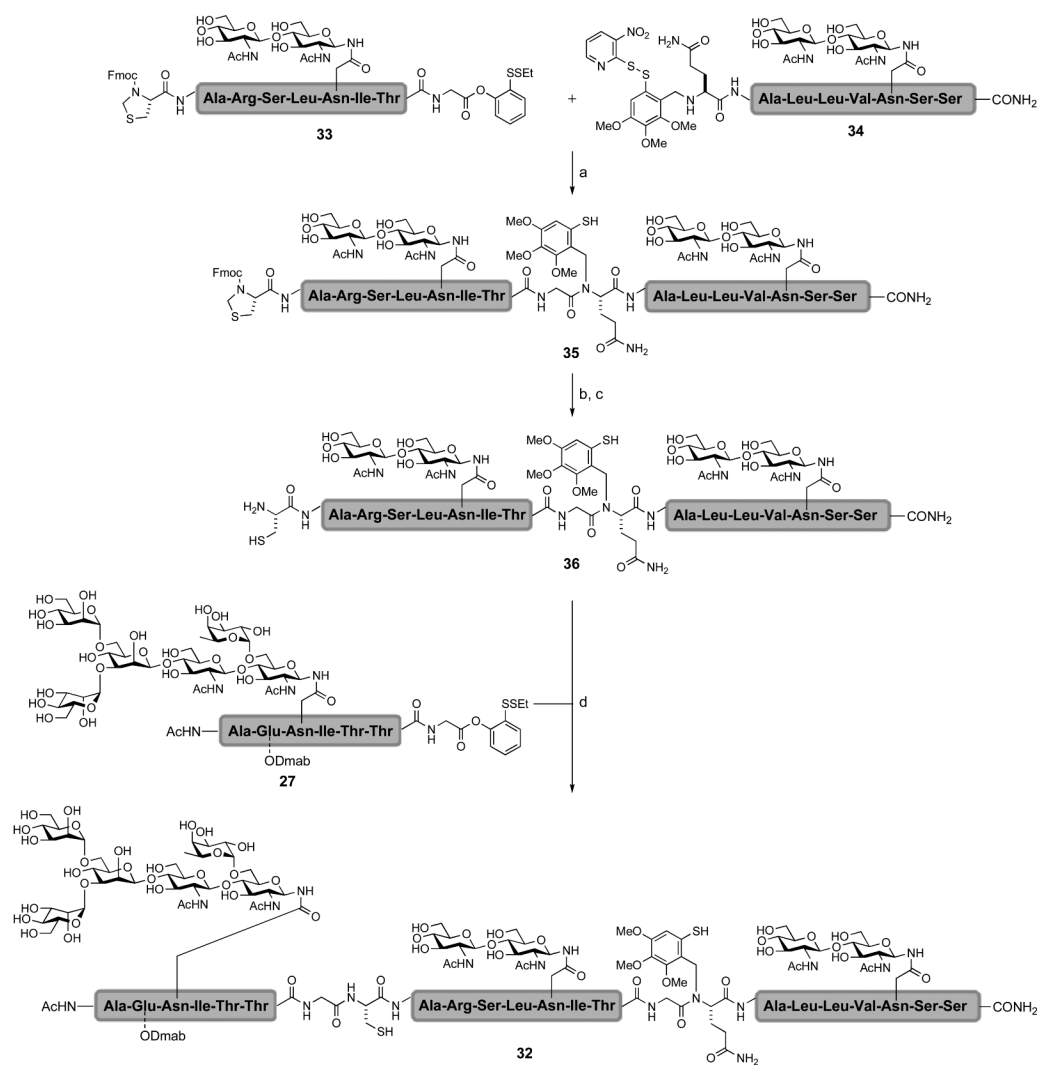


Figure 5.
A retrosynthetic route to glycopeptide **32** via reiterative cysteine-based and cysteine-free ligations

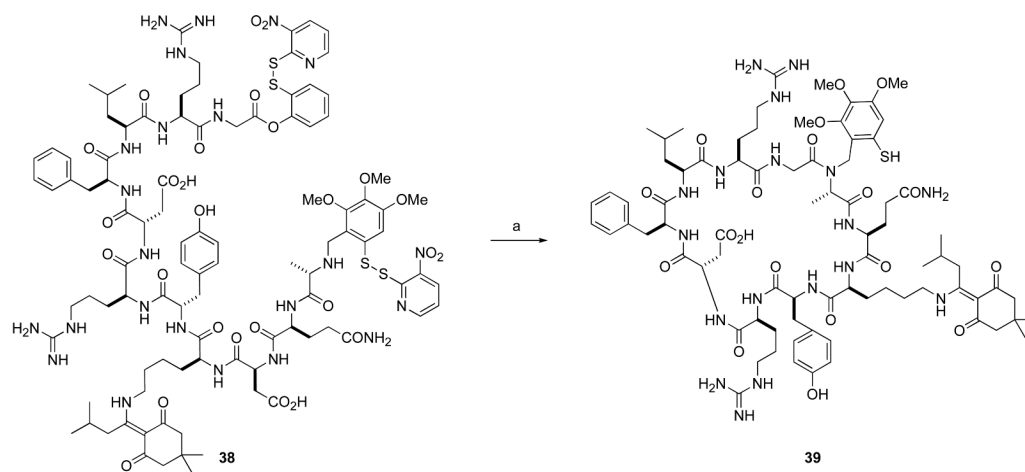


^a Reagents and conditions: (a) TCEP, DMF, Na₂HPO₄, 58%; (b) 10% morpholine in DMF; (c) 0.4 M

MeONH₂·HCl, 60%; (d) MesNa, TCEP, PBS (pH 8.0), 57%.

Scheme 4.

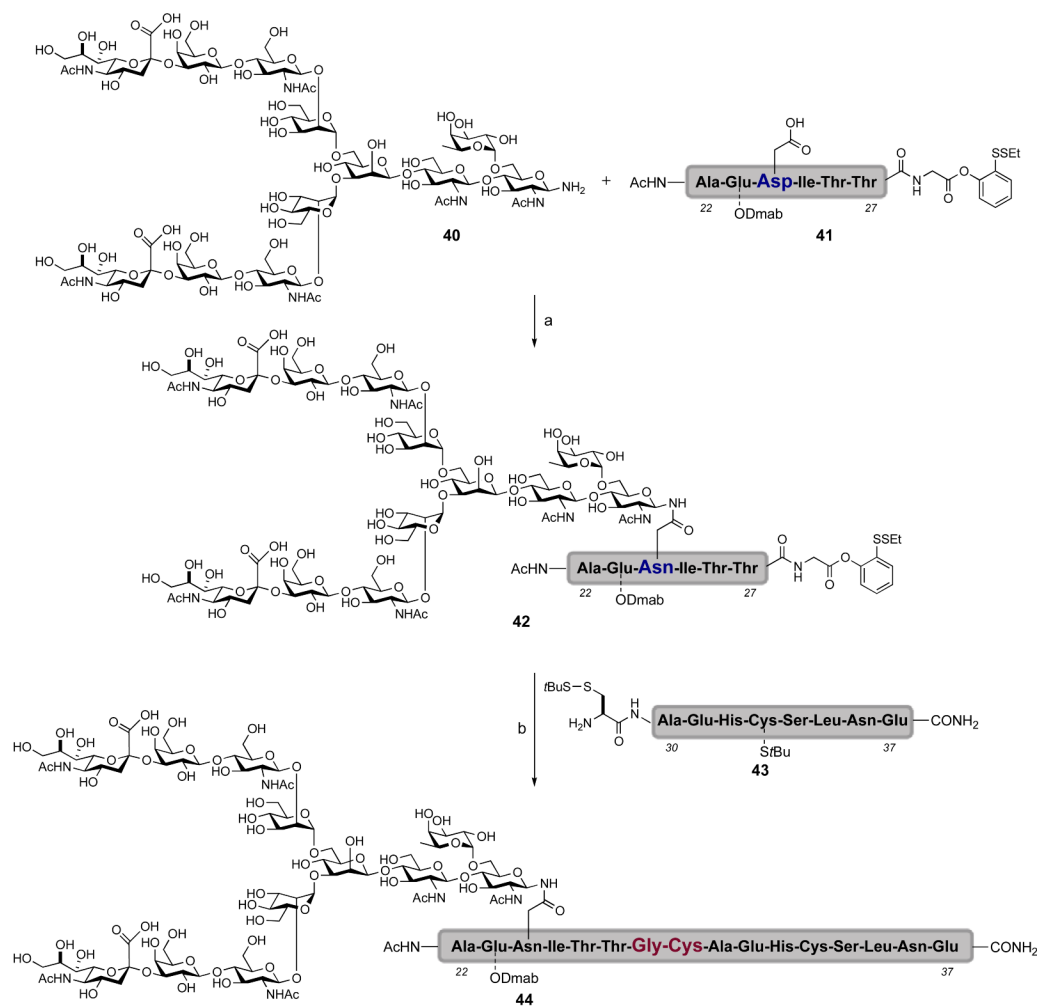
^a Synthesis of trifunctional glycopeptide **32** via reiterative cysteine-based and cysteine-free ligations



^a Reagents and conditions: (a) TCEP, DMF, Na₂HPO₄, 78%.

Scheme 5.

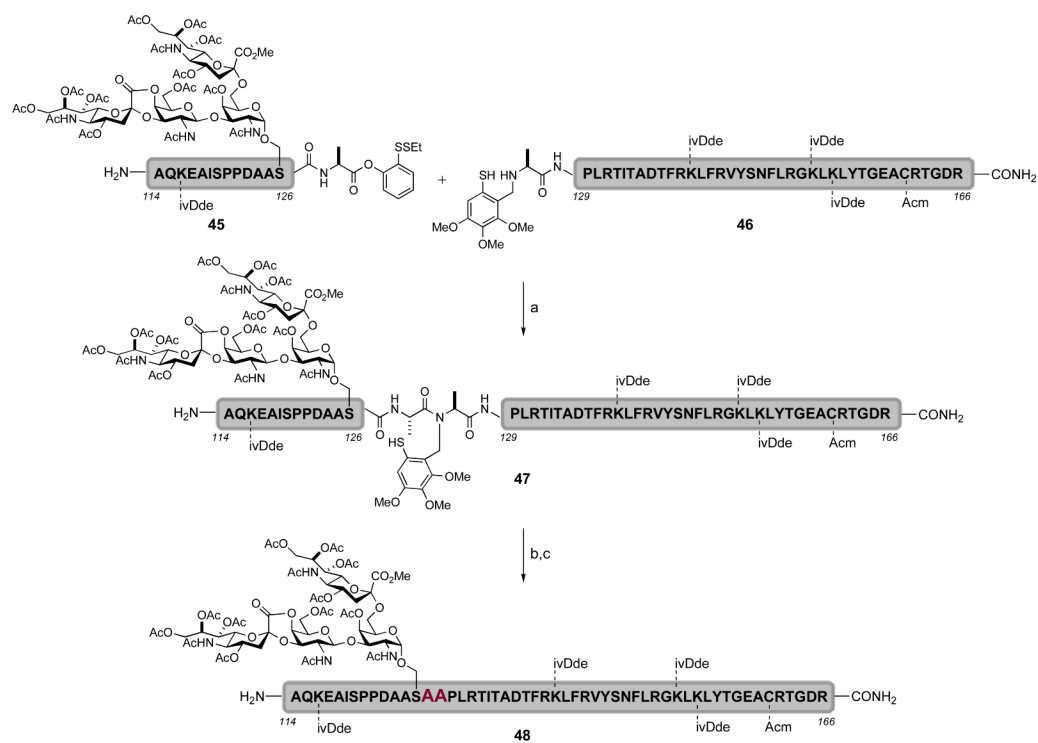
^a Synthesis of cyclic peptide **39** via an intramolecular cysteine-free ligation



^a Reagents and conditions: (a) HATU, *i*Pr₂NEt, DMSO, 40%; (b) PBS, 1% PhSH, TCEP, 56%.

Scheme 6.

a Synthesis of the erythropoietin Ala²²-Glu³⁷ glycopeptide domain (**44**) via cysteine-based ligation



^a Reagents and conditions: (a) TCEP, DMF, Na₂HPO₄, 57%; (b) methyl *p*-nitrobenzenesulfonate, PBS, 7.7, 81%; (c) TFA, TIPSH.

Scheme 7.

^a Synthesis of the erythropoietin Ala¹¹⁴-Arg¹⁶⁶ glycopeptide domain (**48**) via cysteine-free ligation

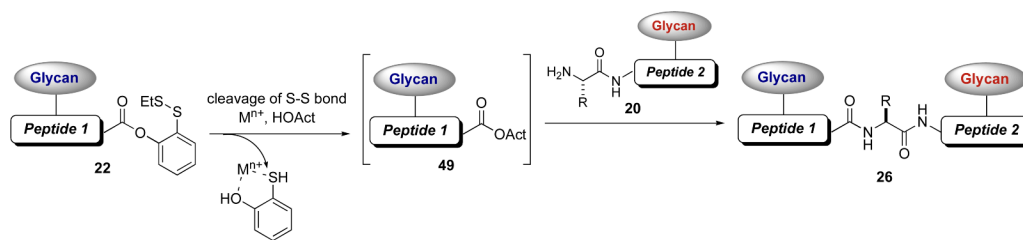
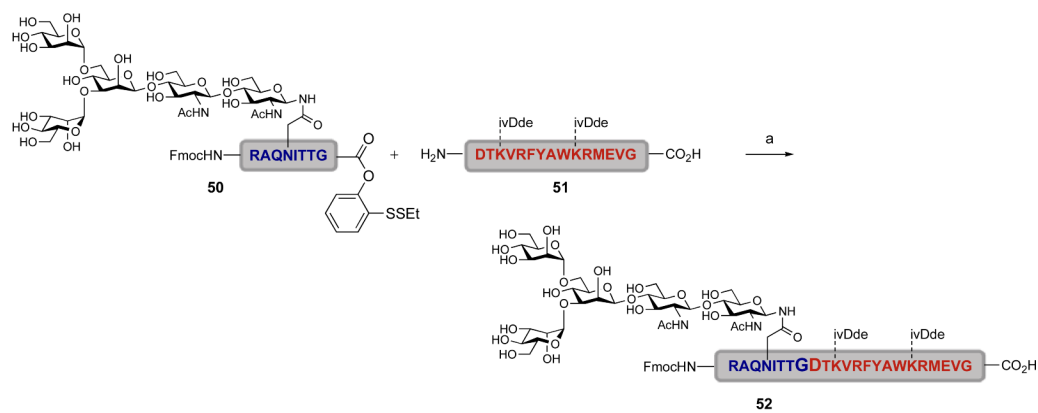


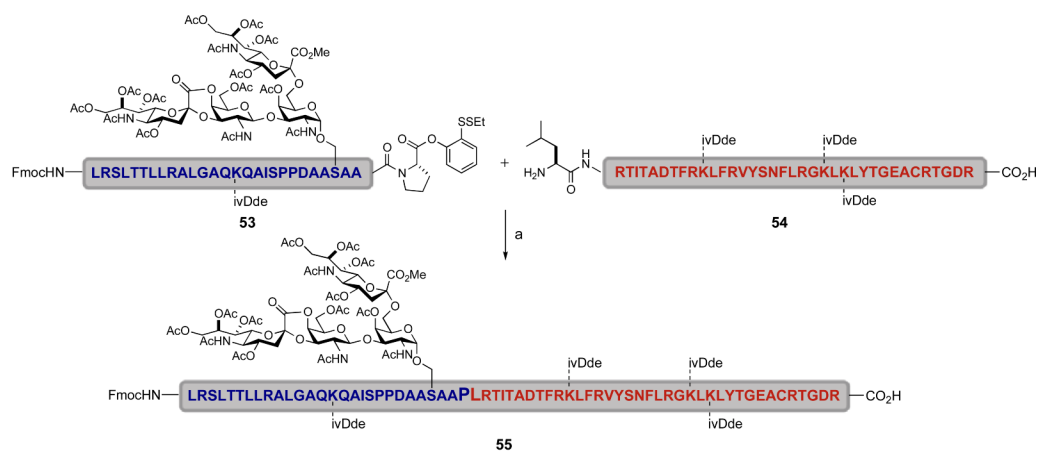
Figure 6.
Direct fragment condensation via metal-mediated activation of an ortho-disulfide phenolic ester



^a Reagents and conditions: (a) AgCl, HOObt, *i*Pr₂NEt, DMSO, 10 h, rt, 80%.

Scheme 8.

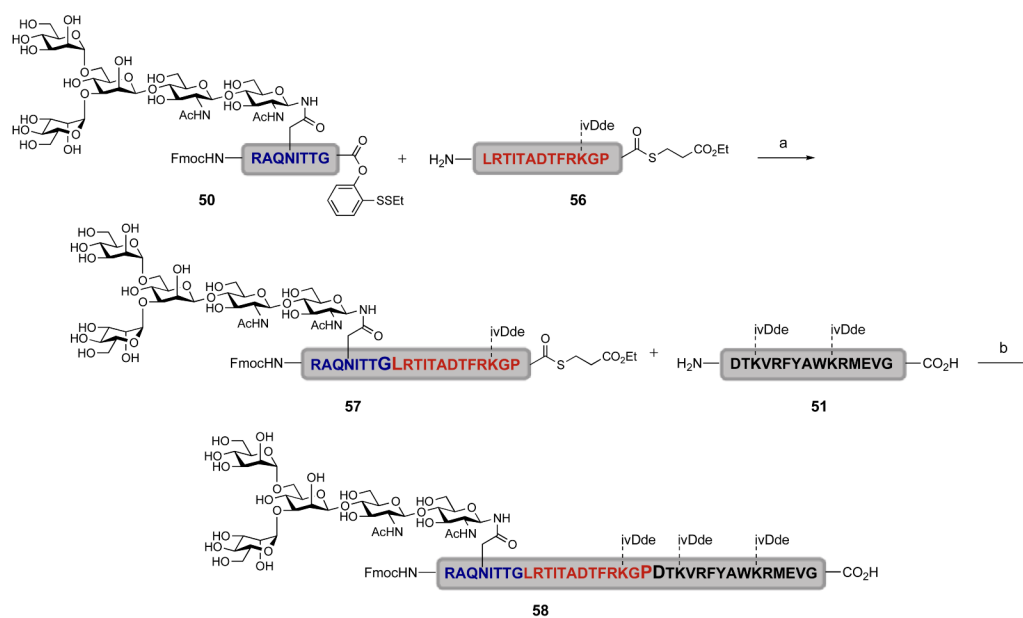
^a Synthesis of glycopeptide **52** via silver-mediated fragment condensation



^a Reagents and conditions: (a) TCEP·HCl, HOObt, *i*Pr₂NEt, DMSO, 48 h, rt, 80%.

Scheme 9.

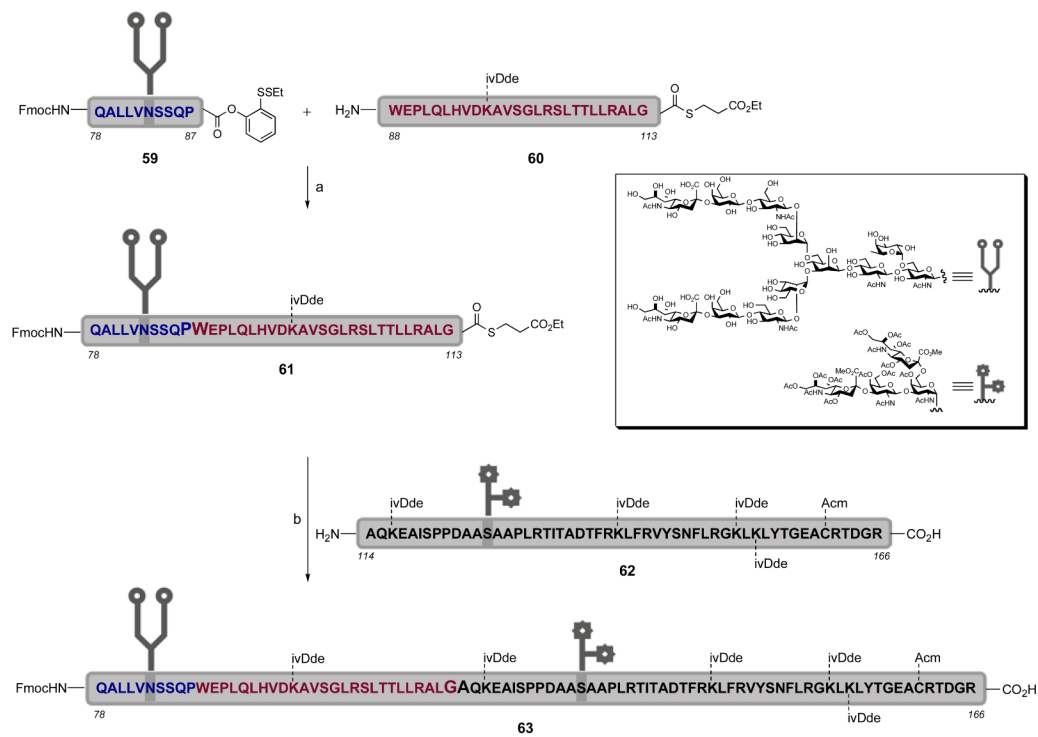
^a Synthesis of glycopeptide **55** via TCEP-mediated direct condensation



^a Reagents and conditions: (a) TCEP·HCl, HOObt, *i*Pr₂NEt, DMSO, 10 h, rt, 80%; (b) AgCl, HOObt, *i*Pr₂NEt, DMSO, 24 h, rt, 65%.

Scheme 10.

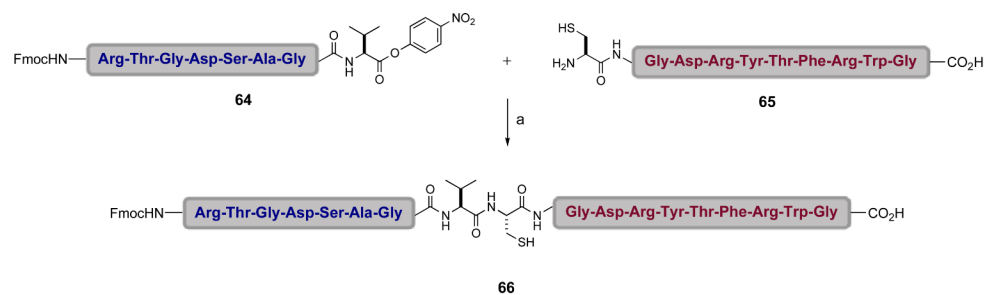
^a Reiterative TCEP/AgCl (glyco)peptide coupling



^a Reagents and conditions: (a) TCEP·HCl, HOObt, *i*Pr₂NEt, DMSO, 33%; (b) AgCl, HOObt, *i*Pr₂NEt, DMSO, 30%.

Scheme 11.

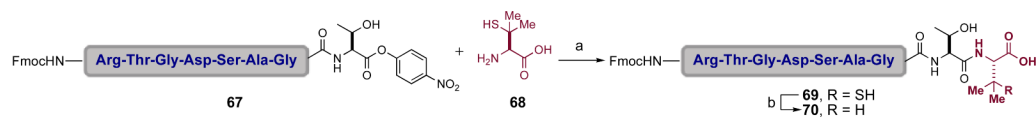
^a Synthesis of the erythropoietin glycopeptide domain Gln⁷⁸-Arg¹⁶⁶ (**63**) via reiterative TCEP/AgCl fragment coupling



^a Reagents and conditions: (a) 6.0 M Guanidine-HCl, 188.8 mM Na₂HPO₄, 7.2 mM *p*-NO₂-PhOH, pH = 6.3 to 6.5, TCEP, rt, 6 h, 69%.

Scheme 12.

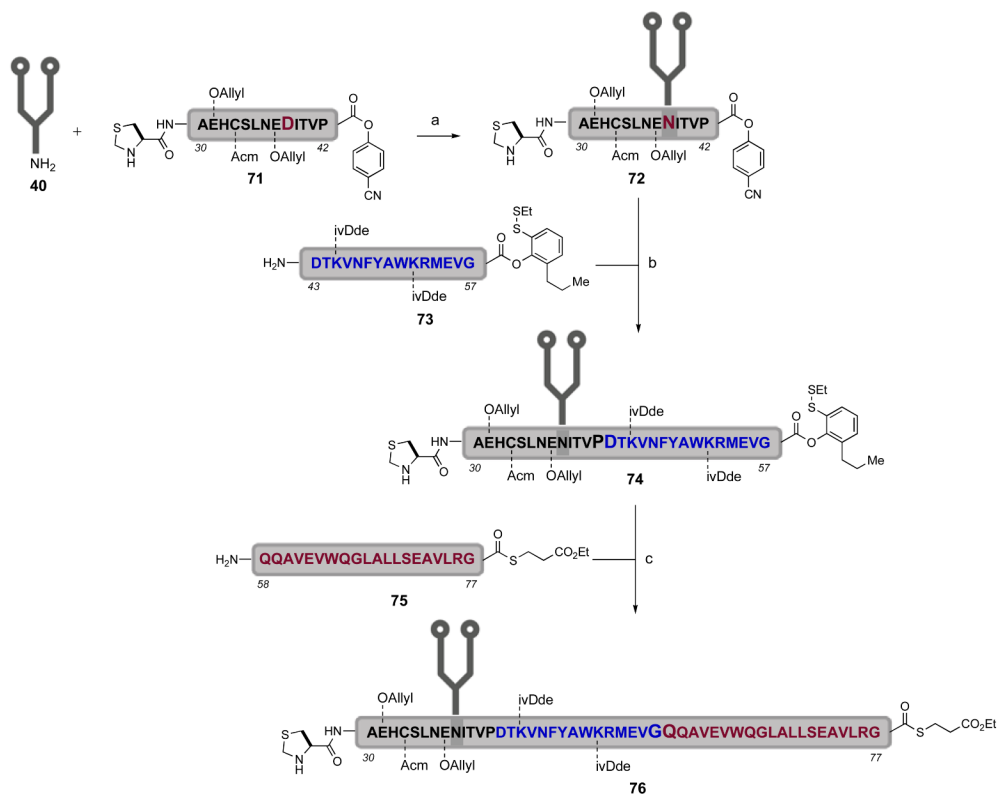
^a A *para*-nitrophenyl ester as acyl donor in native chemical ligation



^a Reagents and conditions: (a) 6.0 M Guanidine-HCl, 188.8 mM Na₂HPO₄, 7.2 mM *p*-NO₂-PhOH, pH = 6.3 to 6.5, TCEP, 30 °C, 2.5 h, 54%; (b) 6.0 M Guanidine-HCl, 188.8 mM Na₂HPO₄, 18.8 mM TCEP-HCl, pH = 6.3 to 6.5, TCEP, VA-044, 37 °C, 2.5 h, 83%.

Scheme 13.

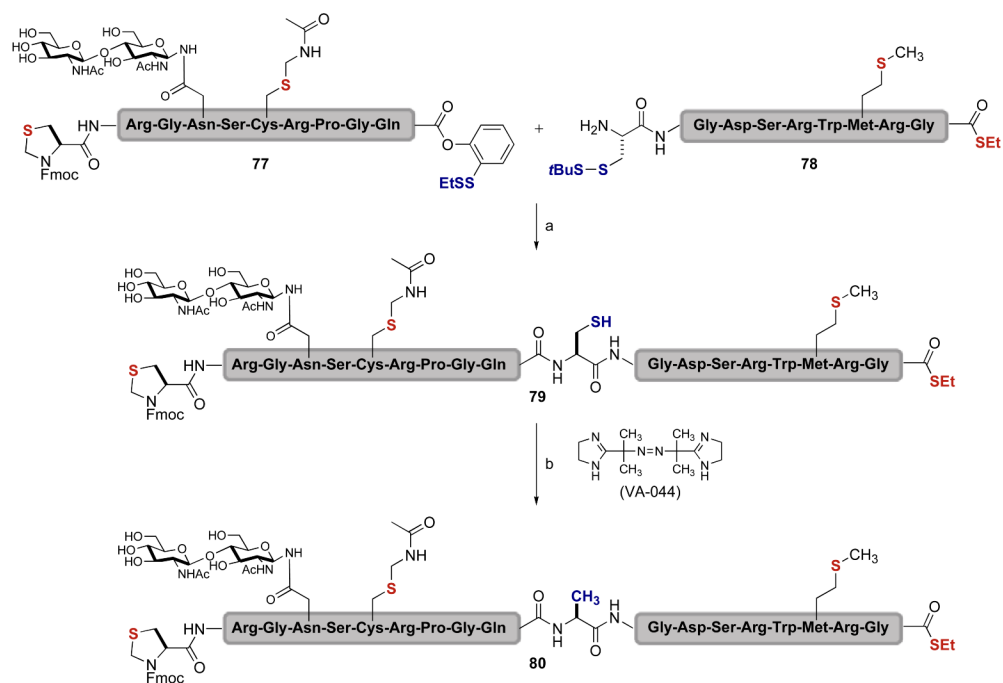
^a Peptide ligation with hindered penicillamine



^a Reagents and conditions: (a) HATU, *i*Pr₂NEt, DMSO, rt, 33%; (b) HOObt, *i*Pr₂NEt, DMSO, rt, 47%; (c) TCEP, HOObt, *i*Pr₂NEt, DMSO, rt, 51%.

Scheme 14.

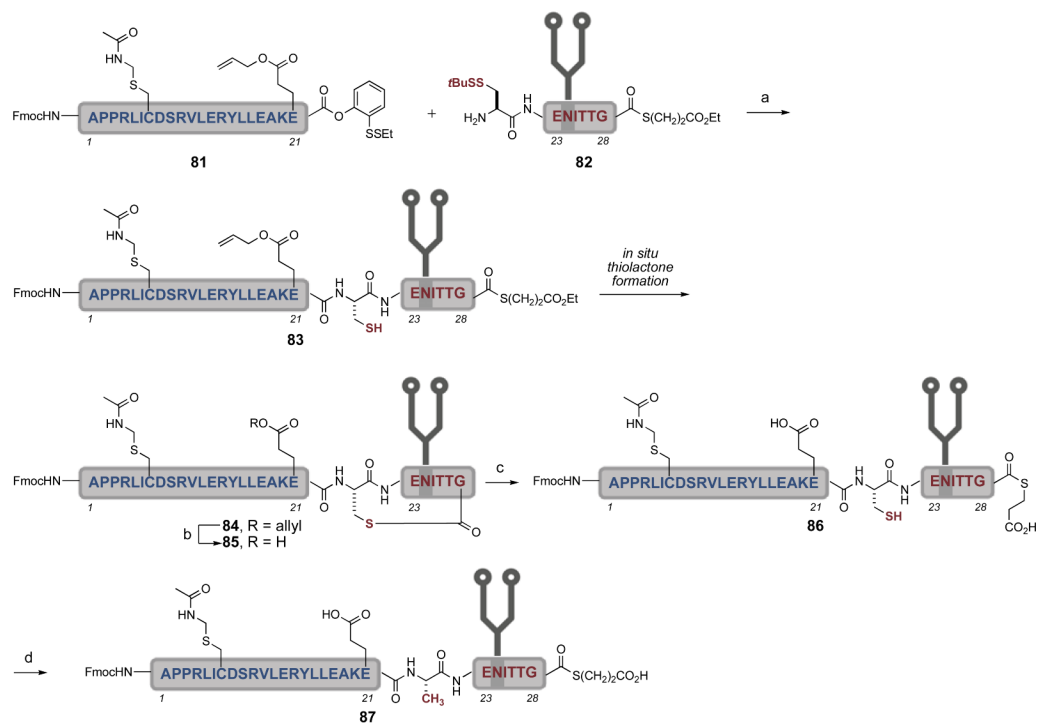
^a Fine tuning acyl donors in the synthesis of the erythropoietin Cys²⁹-Gly⁷⁷ glycopeptide domain (76)



^a Reagents and conditions: (a) 6M Guanidine·HCl, Na₂HPO₄, TCEP·HCl, pH 6.3, 2 h, 67%; (b) TCEP, VA-044, EtSH/*t*BuSH, H₂O, 37 °C, 2 h, 87%.

Scheme 15.

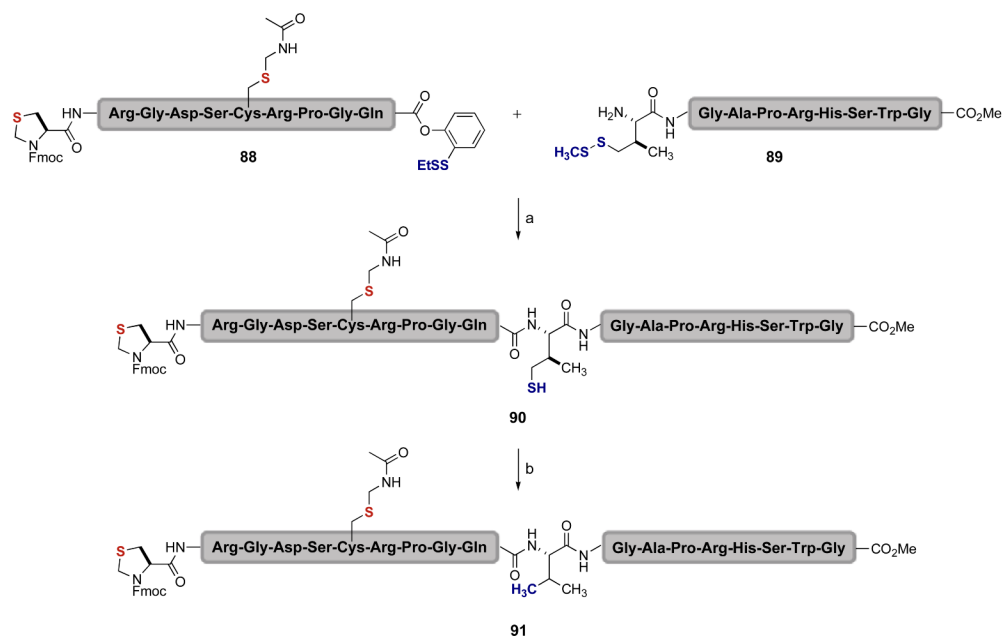
^a Glycopeptide synthesis via native chemical ligation and radical desulfurization



^a Reagents and conditions: (a) 6M Guanidine-HCl, Na₂HPO₄, TCEP-HCl, TCEP, 20%; (b) Pd(PPh₃)₄, PhSiH₃, DMSO, 90%; (c) 6M Guanidine-HCl, Na₂HPO₄, TCEP-HCl, TCEP, thiopropionic acid, 65%; (d) 6M Guanidine-HCl, Na₂HPO₄, TCEP-HCl, TCEP, VA-044, thiopropionic acid, 37 °C, 67%.

Scheme 16.

^a Synthesis of the erythropoietin Ala¹-Gly²⁸ glycopeptide domain (**87**) via native chemical ligation and radical desulfurization



^a Reagents and conditions: (a) 6M Guanidine·HCl, Na₂HPO₄, TCEP·HCl, pH 6.5, 1 h, 78%; (b) TCEP, VA-

044, *t*BuSH, H₂O/MeCN, 37 °C, 3 h, 84%.

Scheme 17.

^a Native chemical ligation at valine

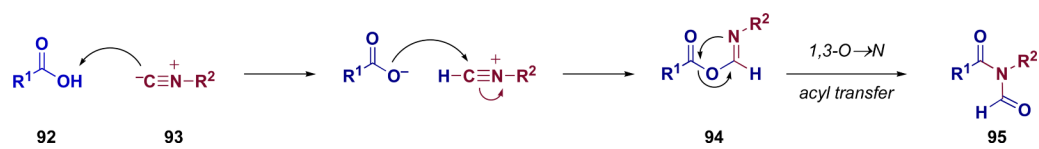
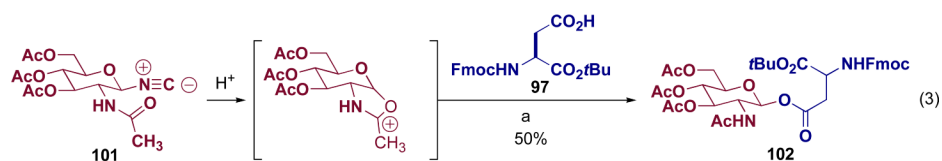
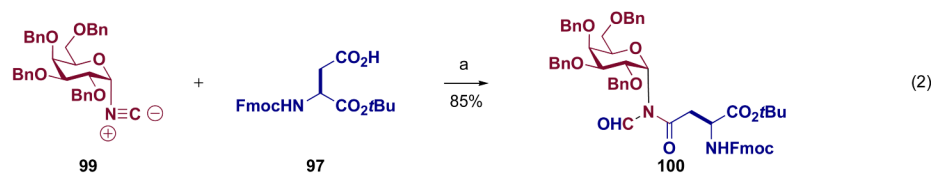
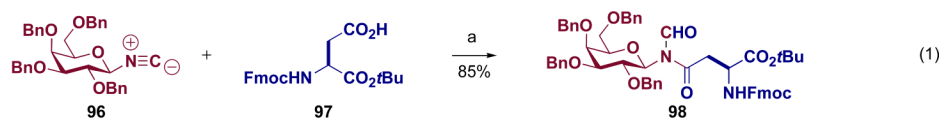


Figure 7.
Two-component coupling reaction of carboxylic acids and isonitriles



^a Reagents and conditions: (a) CHCl₃, 150 °C, microwave, 30 to 45 min.

Scheme 18.

^a Studies on the two-component coupling in building asparagine-linked glycopeptides

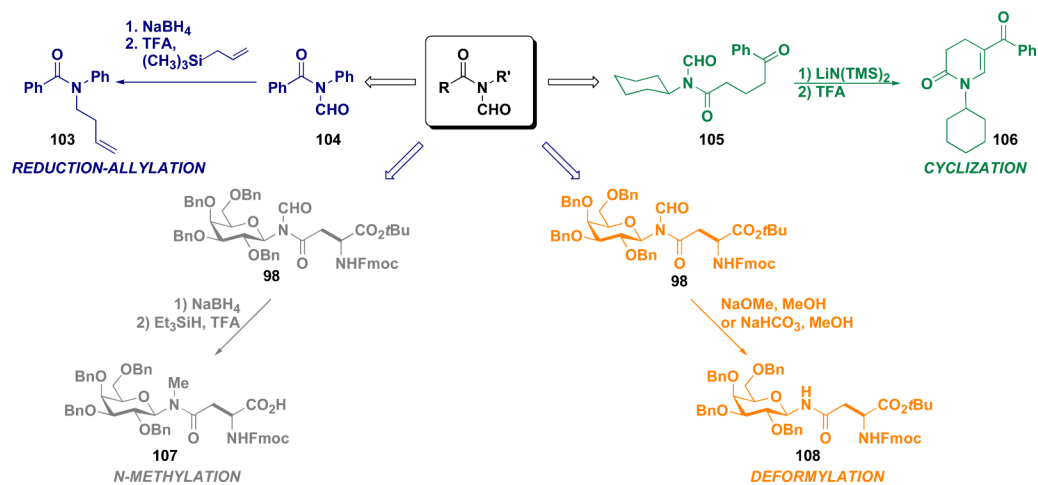


Figure 8.
 The versatility of the N-formyl mixed imide