Top Curr Chem (2015) 362: 27–88
DOI: 10.1007/128_2014_584

Published online: 18 December 2014

Modern Extensions of Native Chemical Ligation for Chemical Protein Synthesis

Lara R. Malins and Richard J. Payne

Abstract Over the past 20 years, native chemical ligation has facilitated the synthesis of numerous complex peptide and protein targets, with and without post-translational modifications, as well as the design and construction of a variety of engineered protein variants. This powerful methodology has also served as a platform for the development of related chemoselective ligation technologies which have greatly expanded the scope and flexibility of ligation chemistry. This chapter details a number of important extensions of the original native chemical ligation manifold, with particular focus on the application of new methods in the total chemical synthesis of proteins. Topics covered include the development of auxiliary-based ligation methods, the post-ligation manipulation of Cys residues, and the synthesis and utility of unnatural amino acid building blocks (bearing reactive thiol or selenol functionalities) in chemoselective ligation chemistry. Contemporary applications of these techniques to the total chemical synthesis of peptides and proteins are described.

Keywords Deselenization · Desulfurization · Native chemical ligation · Post-translational modifications · Protein synthesis

Contents

L.R. Malins (✉) and R.J. Payne (✉)
School of Chemistry, The University of Sydney, Camperdown, NSW 2006, Australia
e-mail: lara.malins@sydney.edu.au; richard.payne@sydney.edu.au

1 Introduction

Proteins are an important class of biomolecules which play a crucial role in a number of diverse physiological processes, including structure, transport, storage and catalysis. The immense structural and functional complexity of proteins is a consequence not only of the unique, genetically-encoded primary sequence of constituent amino acids, but also of the potential introduction of covalent post-translational modifications (PTMs), which occurs enzymatically after protein translation on the ribosome. Each of these factors has a profound impact on the three-dimensional structure, and ultimately the function and/or activity of a protein target.

The study of proteins and their use as novel therapeutics has been aided by the ability to access significant quantities of target molecules using biological expression of programmed genetic sequences. However, such technologies alone are insufficient to enable the flexible design of protein therapeutics, or even the detailed study of all native proteins comprising an organism's proteome, which is considerably more diverse than would be predicted simply from the finite size of the genome. The difference is due, in part, to the enzyme-mediated post-translational modification of proteins. The diverse array of PTMs, including acetylation, phosphorylation, glycosylation, methylation, hydroxylation, and ubiquitylation, among others, occur on the majority of the 20 common proteinogenic amino acid side-chains and, in some cases, on the amide backbone [1]. These modifications serve to greatly enhance proteome diversity beyond the intrinsic size of the genome, and also have significant effects on protein conformation, localization, and function. It is predicted that approximately 5% of the human genome is dedicated to encoding enzymes responsible for PTMs [2].

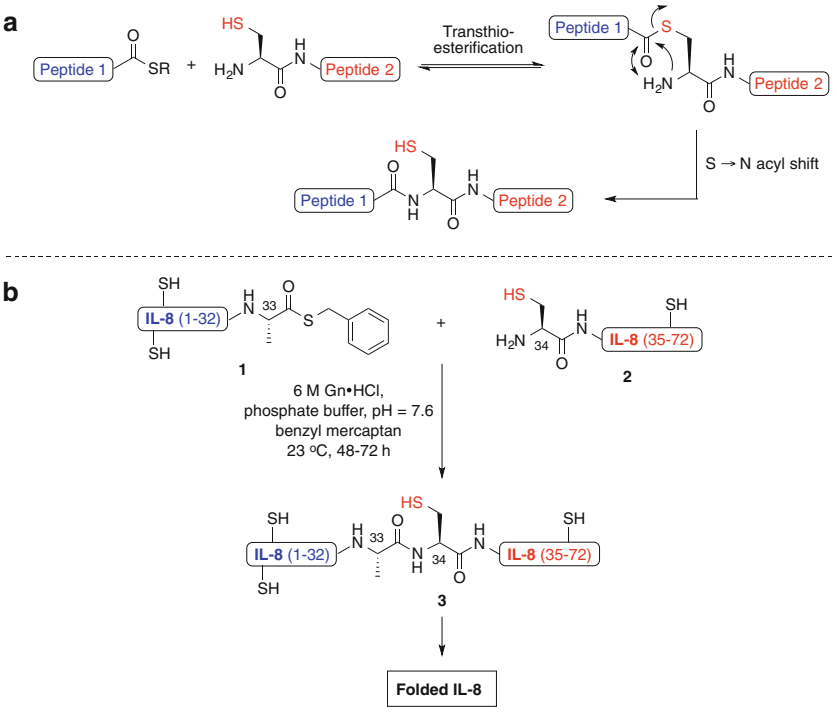
As a consequence of their incorporation through enzyme-mediated processes, PTMs, unlike the primary structure of proteins, are not under direct template control and are instead dictated by the relative intracellular concentration of processing enzymes. This has important implications in cell-signaling pathways, in which reversible modifications, such as phosphorylation and acetylation, dictate both the strength and duration of a signaling process [2]. As a result, the composition of the proteome is inherently dynamic, and the variability of PTMs results in the concomitant heterogeneity of native proteins derived or isolated from living systems. Given these considerations, the role of synthetic chemistry, in the context of the

total synthesis of proteins, becomes paramount to obtaining homogenous, post-translationally modified proteins for the purpose of understanding the biological role of such modifications. The therapeutic potential of modified peptide and protein targets may also benefit from the design and synthesis of engineered variants accessed in homogeneous form via chemical synthesis. Indeed, our understanding of the link between protein structure and function, and our ability to exploit this knowledge for the development of novel therapeutics, is closely connected to our aptitude in the chemical synthesis of complex targets. This chapter therefore focuses on the development of synthetic methods, specifically peptide ligation methods, for the convergent assembly of large polypeptides and proteins from smaller peptide fragments. Particular emphasis is placed on extensions to the native chemical ligation reaction, which has found enormous application in the chemical synthesis of proteins, with and without modifications, since the first report of the method two decades ago [3].

2 Native Chemical Ligation

The development of native chemical ligation by Kent and coworkers in 1994 marked a substantial step in the area of chemical protein synthesis by allowing, for the first time, the chemoselective construction of a native amide linkage between two fully unprotected peptides under mild, aqueous conditions [3]. Building upon a 1953 report by Wieland et al. for the linking of amino acids [4], the method involves the condensation of a peptide bearing an N-terminal cysteine (Cys) residue with a peptide bearing a C-terminal thioester (Scheme 1a). Mechanistically, the reaction occurs through a reversible transthioesterification reaction between the Cys thiol functionality and the C-terminal peptide thioester, generating an intermediate thioester. Spontaneous rearrangement of the bridged thioester moiety in a proximity-induced intramolecular $S \rightarrow N$ acyl shift then occurs through a five-membered ring intermediate to afford a new amide bond [3]. Importantly, the reaction proceeds in the presence of all native amino acid side-chains (including Cys residues distant from the ligation site) and enables a modular approach to the construction of proteins from peptide fragments prepared via iterative solid-phase peptide synthesis (SPPS) [5], a highly robust method for the construction of defined polypeptide sequences.

Kent and coworkers first demonstrated the power of the native chemical ligation methodology toward protein synthesis in the construction of human interleukin-8 (IL-8), a 72-amino acid polypeptide chain bearing 18 of the 20 common proteinogenic amino acids and 4 Cys residues, which form disulfide linkages in the native protein. This was accomplished via efficient ligation of a 33-amino acid peptide bearing a C-terminal benzyl thioester **1** with a 39-residue peptide **2** containing an N-terminal Cys residue. The reaction was performed in aqueous buffer (pH 7.6) containing a denaturing chaotropic salt (6 M guanidine hydrochloride) and afforded a single peptide product **3**, despite the presence of additional



Scheme 1 (**a**) Native chemical ligation. (**b**) Synthesis of human IL-8 by native chemical ligation

unprotected Cys residues in both peptide fragments (Scheme 1b). Selective reaction at the N-terminal Cys residue was attributed to the reversible nature of the initial transthioesterification reaction. Although internal Cys residues can react with the C-terminal benzyl thioester, the resultant thioester intermediates were unproductive in the ligation reaction because of the lack of a proximal amine group to promote the amide bond-forming $S \rightarrow N$ acyl shift, an irreversible process under the conditions employed for the ligation reaction. In the synthesis of IL-8, excess benzyl mercaptan was added to the reaction media to promote thiol exchange with the unproductive thioester intermediates, thereby regenerating thioester **1** and enabling productive transthioesterification with the desired N-terminal Cys residue of peptide **2**. Conveniently, the presence of excess thiol also served as a reducing agent to prevent the formation of Cys disulfides [3].

While the concept of peptide ligation chemistry had previously been explored by others, including Kemp and coworkers [6–8], Liu and Tam [9, 10], and Schnölzer and Kent for the construction of backbone-modified proteins [11], the total chemical synthesis of IL-8 by native chemical ligation marked the first preparation of a protein target, bearing a fully native polypeptide backbone, through the chemoselective condensation of unprotected peptide fragments. It was predicted that this modular synthetic approach would enable nearly unlimited variation in the covalent structure of proteins, thereby facilitating the systematic study of the

structure and function of these important macromolecules [3]. This has indeed proved to be the case with hundreds of proteins prepared using this robust methodology.

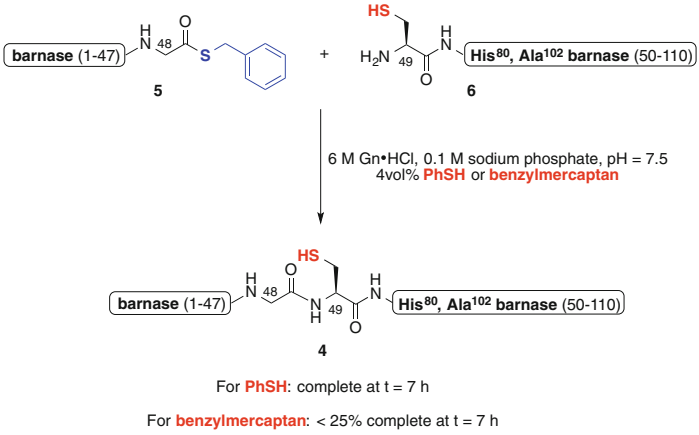
2.1 Scope and Mechanism

Shortly after the initial discovery of native chemical ligation, a number of studies were subsequently conducted with the aim of establishing the scope of the ligation methodology and deciphering the subtleties of the reaction mechanism. In particular, the dependence of the rate of ligation on the nature of the C-terminal peptide thioester moiety has been extensively studied.

Historically, C-terminal peptide thioesters for use in native chemical ligation have been prepared as the alkyl thioester derivatives using optimized in situ neutralization Boc-strategy SPPS [12]. The C-terminal benzyl thioesters initially employed by Kent and coworkers, for example, were prepared via Boc-SPPS using an HF-labile thioacid linker followed by a subsequent alkylation with benzyl bromide [3]. Advances in thioester linkers developed shortly thereafter facilitated the direct synthesis of peptide alkyl thioesters upon cleavage from the resin [13, 14]. The facile handling and preparation and the general stability of alkyl thioesters render these functionalities a convenient choice for use in native chemical ligation reactions. Nonetheless, in terms of acylating power, alkyl thioesters are relatively unreactive acyl donors, a result of the modest leaving group ability of the alkyl thiol component. The use of a more reactive peptide thioester, containing a better thiol leaving group, was first investigated in the seminal report of native chemical ligation, whereby a thioester bearing a 5-thio-2-nitrobenzoic acid (the reduced form of Ellman's reagent) leaving group was shown to facilitate rapid ligation [3].

As a corollary to this observation, one of the earliest general advances to the native chemical ligation method was the development and exploration of thiol ligation catalysts with a view to modulating thioester reactivity [15]. In 1997, Kent and coworkers found that the inclusion of thiophenol as an exogenous ligation additive served to facilitate thiol-thioester exchange with preformed alkyl thioesters, thereby promoting the in situ formation of the considerably more reactive peptide aryl thioester (Scheme 2) to facilitate more rapid native chemical ligation reactions. To enable a direct comparison between benzyl thiol and thiophenol as effective ligation additives, the authors explored the ligation-based assembly of a Cys-containing analogue of barnase **4**, a 110-amino acid microbial ribonuclease. Construction of the barnase analogue from barnase(1–48)-COS-benzyl **5**, a preformed peptide C-terminal benzyl thioester, and a modified N-terminal fragment, [Cys^{49}, His^{80}, Ala^{102}]barnase(49–110) **6**, in the presence of thiophenol proceeded to completion in approximately 7 h (Scheme 3). In contrast, in the presence of benzyl mercaptan, the ligation reaction was less than 25% complete after the same period of time. The observed rate enhancement for the ligation of a preformed benzyl thioester in the presence of exogenous thiophenol established in

Scheme 2 Thiophenol as an exogenous thiol additive for native chemical ligation



Scheme 3 The effect of thiophenol as an exogenous thiol additive for the construction of barnase analogue **4**

situ transthioesterification with thiol additives as a practical and general means of modulating the reactivity of C-terminal peptide alkyl thioesters for use in native chemical ligation [15].

More recently, a detailed study of the ability of various thiol additives to promote native chemical ligation identified the water-soluble aryl thiol mercaptophenylacetic acid (MPAA) as the optimal thiol catalyst [16]. This study also established an important relationship between ligation rate and the pK_a of the thiol additive. The most effective additives, such as thiophenol and MPAA, had pK_a values between approximately 6 and 8, thus maintaining sufficient nucleophilicity to promote the initial thiol-thioester exchange with the less reactive peptide alkyl thioester component, and adequate leaving group ability for effective transthioesterification with the Cys-containing peptide fragment [16]. Thiol additives with higher pK_a values (e.g., the water-soluble alkyl thiol sodium 2-mercaptoethanesulfonate (MESNa) [17], commonly employed in expressed protein ligation (EPL) [18]) rapidly exchanged with preformed alkyl thioester derivatives but were poor leaving groups in the reaction with a peptide bearing an N-terminal Cys residue. Conversely, thiol additives with a pK_a value less than six were inefficient ligation additives as a result of their inability to effectively activate preformed alkyl thioesters through thiol-thioester exchange [16]. It should also be noted that highly electrophilic acyl donors may increase the likelihood of thioester hydrolysis or epimerization at the α-center of the C-terminal amino acid residue,

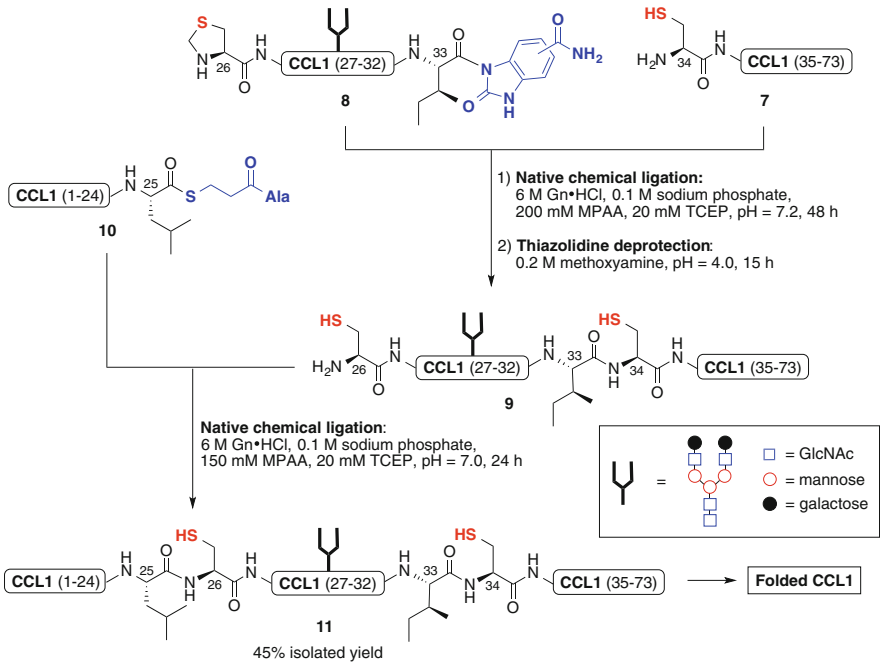
particularly with challenging ligation junctions. In addition to the pK_a of the thiol used in the reaction, these factors must be carefully considered in the design of new ligation additives. The development of other thiol catalysts for application in ligation chemistry is further discussed later in this chapter.

An equally important determinant of thioester reactivity in native chemical ligation reactions is the nature of the C-terminal residue bearing the reactive thioester functionality. In 1999, Dawson and coworkers carried out a comprehensive study of the rate of ligation as a function of the identity of the C-terminal amino acid residue [13]. In a series of model ligation reactions, whereby C-terminal thioesters derived from each of the 20 proteinogenic amino acids were examined, the reaction rate was shown to be closely correlated with the steric and electronic nature of the C-terminal thioester moiety. For example, reactions at sterically unencumbered C-terminal amino acid thioesters proceeded much faster than reactions employing bulky, β-branched amino acid thioesters, such as those derived from Ile, Val, and Thr, which, under the ligation conditions employed, required 48 h or more to reach completion. Electronic effects were likewise apparent, with Cys and His thioesters reacting at remarkably similar rates to the model Gly thioester, despite the increased steric bulk, and much faster than the corresponding Ala thioester [13]. Poor reactivity observed with the model Pro thioester was subsequently explained by the decreased electrophilicity of the Pro thioester carbonyl carbon resulting from an orbital overlap ($n \rightarrow \pi^*$) with the adjacent N-terminal amide oxygen atom [19].

The power of native chemical ligation together with an extended knowledge of the specific reactivity of C-terminal peptide thioesters have been exploited in the total chemical synthesis of numerous proteins to date [20–28], greatly expanding the scope and size of native protein targets within the grasp of synthetic chemists. In combination with SPPS as a robust approach to access peptide fragments bearing native amino acids as well as unnatural amino acid building blocks (derived from modern organic synthesis), the method has advanced the opportunity for protein engineering and structural remodeling. As a testament to the power of native chemical ligation, 20 years after its seminal report the technique is often employed, in near original form, for the total synthesis of complex and post-translationally modified proteins.

2.2 Modern Application

A recent, illustrative example from Okamoto et al. of the extensive applicability of native chemical ligation toward the study of protein structure and function is the total chemical synthesis of the glycosylated and non-glycosylated forms of the 73-amino acid chemokine protein CCL1 and a 74-amino acid, N-terminally extended variant, Ser-CCL1 [29]. The construction of these chemokines was accomplished using a convergent, three-component iterative native chemical ligation approach (see Scheme 4 for a representative example), with the glycosylated



Scheme 4 Total chemical synthesis of glycosylated CCL1 using an iterative native chemical ligation strategy

variants each containing a complex, *N*-linked asialo-nonasaccharide at the native glycosylation site, Asn[29].

Notably, preparation of the requisite peptide fragments for the construction of the target proteins strategically utilized a number of recent and powerful advances to the native chemical ligation motif. Using the synthesis of glycosylated CCL1 as a representative model (Scheme 4), the first step in the iterative ligation strategy involved the reaction of C-terminal Cys-containing peptide CCL1(34–73) **7**, prepared by Boc-SPPS, with bifunctional glycopeptide **8**. To enable chemoselective iterative ligation, the N-terminal Cys residue of glycopeptide **8** was masked as the corresponding thiazolidine (Thz) residue [30], which is orthogonal to the general conditions employed in solid-phase peptide synthesis, but is mildly removed in the presence of methoxyamine at pH 4.0. Furthermore, due to the incompatibility of glycans with the strongly acidic deprotection conditions (e.g., HF) employed in Boc-SPPS, glycopeptide fragment **8** was synthesized via 9-fluorenylmethoxycarbonyl (Fmoc)-SPPS, in which cleavage from the resin is accomplished using milder conditions. Nonetheless, thioester linkers are incompatible with the basic conditions employed for unmasking the N-terminal Fmoc group in standard Fmoc-SPPS. As such, the authors employed a diaminobenzoic acid (Dbz) linker for the Fmoc-based assembly of bifunctional peptide **8**, which bears a C-terminal peptide *N*-acylbenzimidazolinone (Nbz), a novel peptide thioester

precursor accessible from resin-bound *o*-aminoanilides, originally described by Dawson and coworkers in 2008 [31]. In the presence of thiol additives under standard native chemical ligation conditions, the peptide-Nbz group undergoes facile thiolysis to afford the corresponding peptide thioester.

Following the preparation of CCL1 peptide fragments **7** and **8**, the fragments were subsequently joined using native chemical ligation in aqueous denaturing buffer in the presence of 200 mM MPAA as a thiol catalyst and 20 mM of the water soluble phosphine reductant, tris(2-carboxyethyl)phosphine (TCEP), to prevent the formation of Cys disulfides (Scheme 4). The reaction required 48 h to reach completion because of the use of an acyl donor fragment bearing a bulky, β-branched isoleucine (Ile) residue as the C-terminal amino acid. Completion of the ligation, as monitored by HPLC-MS, was directly followed by thiazolidine deprotection in the presence of methoxyamine. Crude fragment **9**, bearing an unmasked N-terminal Cys residue, was then ligated to peptide thioester **10**, corresponding to CCL 1(1–24), affording the full-length glycosylated CCL1 **11** in 45% isolated yield for the two ligation steps following HPLC purification. Similar ligation protocols were employed for the synthesis of the non-glycosylated CCL1 variant and for the glycosylated and non-glycosylated Ser-CCL1 derivatives. Each of the CCL1 variants were folded and evaluated in a chemotaxis assay, allowing for a systematic study of the effect of glycosylation on the chemotactic activity of the proteins. Notably, Kent and coworkers were also able to prepare a non-glycosylated mirror image Ser-CCL1 protein composed entirely of D-amino acids using an analogous synthetic strategy [32]. This D-protein analogue was used to obtain the X-ray crystal structures of both Ser-CCL1 and glycosyl-Ser-CCL1 using racemic and quasi-racemic crystallization, respectively [32].

As the preparation of large quantities of complex, homogeneous post-translationally modified proteins and the construction of unnatural D-proteins represent synthetic feats currently unachievable using biological expression or the majority of other chemical methods, the work described by Okamoto et al. on CCL1 [29, 32] highlights the critical importance of native chemical ligation as an enabling tool for understanding protein structure and function. Importantly, native chemical ligation has also served as a platform for the development of a myriad of related synthetic technologies for application in chemoselective peptide ligations. The remainder of this chapter discusses a number of powerful extensions to the original native chemical ligation manifold, with a particular focus on the development of new technologies for the total chemical synthesis of proteins. Topics covered include the development of auxiliary-based methods for peptide ligation, the post-ligation manipulation of Cys residues, and the synthesis and utility of unnatural amino acid building blocks in native chemical ligation-like reactions. In the process, contemporary applications of these techniques to the total chemical synthesis of peptides and proteins are described.

3 Development of New Cysteine Ligation Surrogates

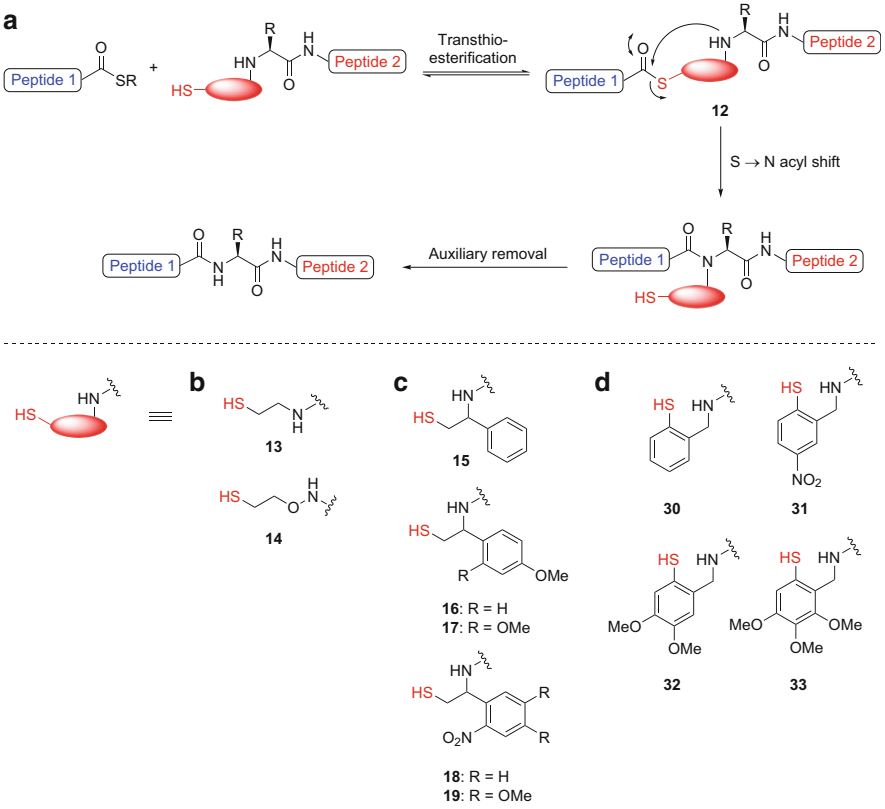
The successful disconnection of proteins using native chemical ligation is predicated on the presence of appropriately placed Cys residues within a target sequence. While there are countless targets that fulfill this unique requirement, the relative scarcity of Cys in naturally occurring proteins (1.1%) [33] precludes the ligation-based assembly of a number of desirable targets and represents a limitation to native chemical ligation in its original form. A number of innovative strategies have therefore been developed to address this intrinsic limitation. The following discussion provides an overview of some of the most significant discoveries, based on the overall logic of native chemical ligation, toward circumventing the reliance on N-terminal Cys residues.

3.1 Auxiliary-Based Methods

3.1.1 N-Terminal Ligation Auxiliaries

One of the earliest and most heavily explored means of circumventing the need for an N-terminal Cys residue has been the use of removable N-terminal thiol ligation auxiliaries to mimic the role of the Cys thiol group in the ligation reaction [27]. A summary of effort in this area is shown in Scheme 5. The general mechanism of N-terminal auxiliary-based ligation is analogous to that of native chemical ligation and is conceptually similar to the original "prior thiol capture" technique employed by Kemp and coworkers [6, 7], whereby a proximity-induced acyl shift event is templated by a traceless reaction auxiliary. In the case of N-terminal auxiliaries, the N-linked thiol reaction auxiliary participates in an initial transthioesterification step, generating the intermediate thioester **12**. A subsequent acylation of the auxiliary-bound secondary amine occurs via an $S \rightarrow N$ shift to generate an amide bond at the ligation site (Scheme 5a). Cleavage of the tethered reaction auxiliary following the ligation event then generates native peptide products. A variety of N-terminal reaction auxiliaries have been explored for this purpose (Scheme 5b–d), allowing access to select non-Cys ligation junctions.

The first N-terminal thiol auxiliary approach was developed by Kent and coworkers in 1996, with the installation of ethanethiol **13** and oxyethanethiol **14** auxiliaries (Scheme 5b) on the terminal amine functionality of peptides bearing N-terminal Gly and Ala residues [34]. These auxiliaries were shown to promote ligation with various C-terminal peptide thioesters in a number of model ligation reactions. Interestingly, when the oxyethanethiol auxiliary **14** was used to ligate peptides bearing steric bulk on either the N-terminal or C-terminal residue of the ligation junction, the authors detected substantial amounts of the unrearranged thioester intermediate (see **12**, Scheme 5a) in the ligation mixture. This observation suggested a decrease in the rate of acyl transfer of the thioester intermediate to the auxiliary-bound, secondary amine. It was postulated that the added steric bulk at the



Scheme 5 (**a**) Native chemical ligation facilitated by N-terminal ligation auxiliaries. (**b**) Oxyethanethiol and ethanethiol auxiliaries. (**c**) 1-Phenylethanethiol auxiliaries. (**d**) 2-Mercaptobenzyl auxiliaries

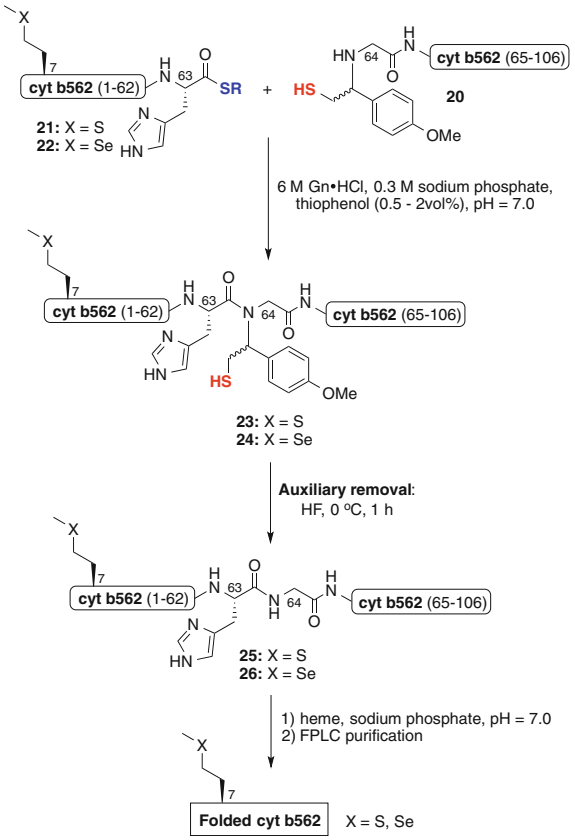
ligation junction in combination with the requirement for a six-membered ring intermediate in the intramolecular rearrangement of the oxyethanethiol-derived thioester (rather than the five-membered ring intermediate generated at unsubstituted Cys residues) served to slow substantially, or in some cases halt, the rate of the $S \rightarrow N$ acyl shift. Nonetheless, in ligation reactions in which amide bond formation was successfully mediated by the oxyethanethiol linker, removal of the ligation auxiliary was readily achieved through cleavage of the N–O bond upon treatment with Zn dust in acidic media. Despite some limitations, this technique served as a proof of concept for the generation of native peptide products following an auxiliary-based ligation strategy [34].

With the aim of further generalizing auxiliary ligations and increasing reaction rates, a number of 1-phenylethane thiol-based auxiliaries (**15**–**19**) (Scheme 5c) [35–39] were also explored. Importantly, these scaffolds were designed to facilitate intramolecular acyl shift through a five-membered ring intermediate and were

proposed to increase the rate of $S \rightarrow N$ acyl transfer relative to oxyethanethiol linker **14** [35]. An additional consideration in the strategic development of the methoxy benzyl auxiliaries **16** and **17** was the ease of removal following the ligation reaction [36, 40]. When bound to the N-terminal α-amino moiety of a peptide fragment, auxiliaries **16** and **17** were not labile to treatment with HF, and were thus compatible with standard Boc-SPPS conditions. The observed stability was attributed to the positive charge ensuing from protonation of the terminal amine under acidic cleavage conditions, which possibly disfavored the formation of a proximal benzylic cation during cleavage of the methoxybenzyl moiety [37]. However, amide bond formation as a consequence of the ligation reaction effectively increased the acid lability of the thiol auxiliary, which was readily cleaved in a post-ligation HF treatment. It is also important to note that, despite the potential for chirality in the 1-phenylethane thiol scaffold, auxiliaries **16** and **17** were synthesized as a mixture of diastereomers [36]. In a contemporary study by Dawson and coworkers on 1-phenylethane thiol derivative **15**, configuration at the benzylic position was determined to have no effect on the rate of ligation [35].

The first application of an N-terminal ligation auxiliary to the total synthesis of proteins was demonstrated through the use of *p*-methoxy auxiliary **16** in the ligation-based assembly of the 106-amino acid metalloprotein cytochrome b562 and an engineered analogue, [SeMet^{7}]cyt b562, containing a strategically placed selenomethionine residue to serve as an axial ligand for the cytochrome heme iron atom [37]. The synthesis of these two proteins was accomplished via ligation of N-terminal auxiliary bound peptide **20** to peptide thioesters **21** and **22**, bearing a native methionine residue or a selenomethionine residue, respectively, at position 7 (Scheme 6). The reaction was performed under aqueous, denaturing conditions in the presence of thiophenol as an exogenous thiol additive. Following the successful ligation, treatment of **23** and **24** with anhydrous HF facilitated clean removal of the N-terminal auxiliary to afford the full-length polypeptide chains **25** and **26**. The target protein molecules were then obtained after folding in the presence of heme, thus enabling a detailed analysis of the spectroscopic and electrochemical properties of wild-type cyt b562 relative to the engineered, selenomethionine-containing protein [37].

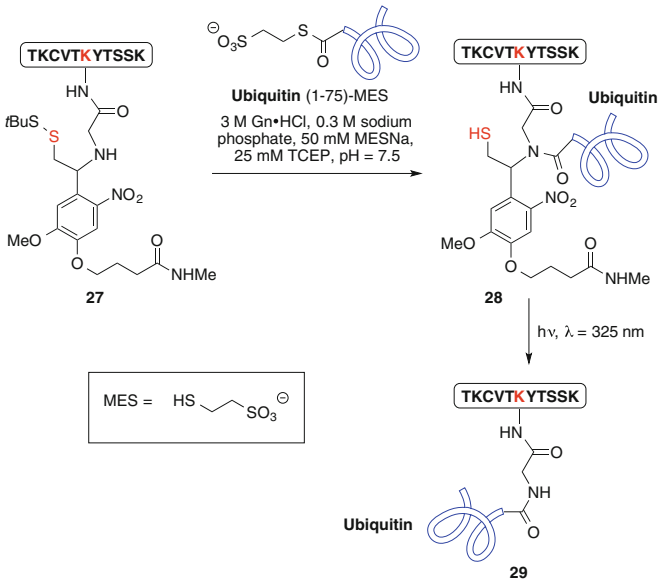
The successes of the 1-phenylethane thiol auxiliary scaffold in protein synthesis prompted the subsequent exploration of photolabile derivatives [35] (such as **18** and **19**) [39, 38] bearing the *ortho*-nitrobenzyl motif (see Scheme 5c), to allow for auxiliary removal without the use of harsh acidic conditions. These reaction auxiliaries were shown to facilitate model ligation reactions at Gly–Gly [38] and Ala–Gly [39] junctions, and could be removed by facile photolysis upon irradiation with UV light. Muir and coworkers have also reported the innovative application of a photolabile linker in auxiliary-mediated expressed protein ligation (EPL) [41] for the synthesis of a ubiquitylated peptide target [42]. In this study, the authors ligate recombinantly produced ubiquitin and SUMO (small ubiquitin-related modifier) thioesters to a small peptide fragment **27**, from the mammalian histone nucleoprotein H2B, bearing an auxiliary-modified Lys residue (see Scheme 7 for ubiquitin example). Following ligation, irradiation of the ubiquitylated product **28** at 325 nm



Scheme 6 Synthesis of cyt b562 and [SeMet^{7}]cyt b562 via native chemical ligation facilitated by an N-terminal ligation auxiliary

effected rapid removal of the photolabile auxiliary to afford chemically modified peptide **29**. Importantly, the final ubiquitin-modified protein was also shown to be a suitable substrate for the ubiquitin-dependent hydrolase UCH-L3, thus confirming the structural and functional integrity of the semi-synthetic protein [42]. A subsequent report from the Muir laboratory utilized the same photolytic ligation auxiliary, in combination with a second EPL reaction, to achieve the semi-synthesis of the full-length, chemically ubiquitylated H2B histone [43].

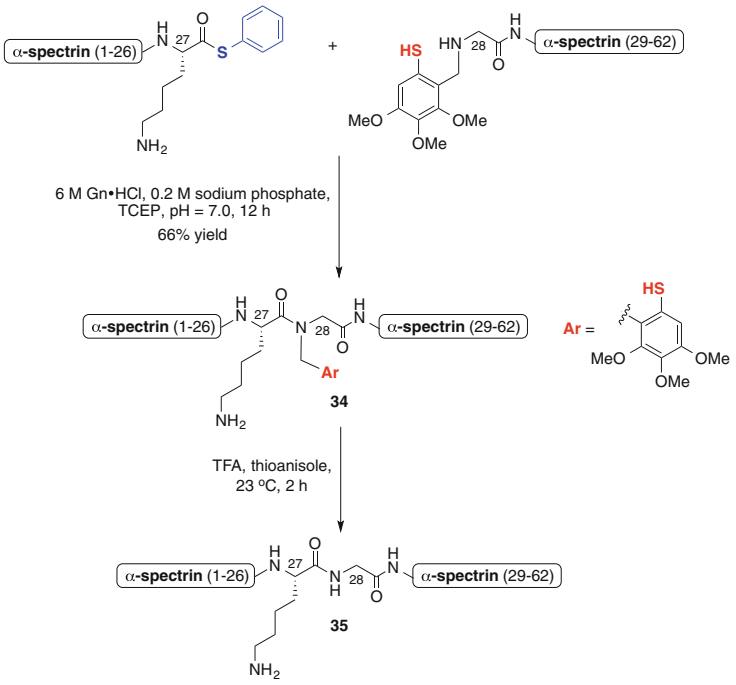
A third class of auxiliaries probed for use in ligation reactions employed the 2-mercaptobenzyl motif (e.g., **30**–**33**, Scheme 5d) [44–47]. The rationale for this tethered thiophenol-based scaffold was to exploit the rate enhancement observed in the presence of aryl thiol additives [15]. It was envisaged that transthioesterification of C-terminal peptide thioesters with the 2-mercaptobenzyl thiol auxiliaries would generate a highly acylating aryl thioester intermediate capable of facilitating rapid $S \rightarrow N$ acyl shift. Furthermore, substitution of the aromatic ring could be used to modulate the nucleophilicity of the thiol or enhance the lability of the auxiliary [44]. For example, increasing substitution of the aromatic ring with electron-donating functionalities (e.g., methoxy groups) was shown to enhance greatly the



Scheme 7 Peptide ubiquitylation using a photolabile N-terminal ligation auxiliary

acid lability of the auxiliaries, with the 4,5-dimethoxy-2-mercaptobenzyl auxiliary **32** [45, 46] cleaved upon treatment with strongly acidic trifluoromethane sulfonic acid (TFMSA) or bromotrimethylsilane (TMSBr) and the more electron-rich 4,5,6-trimethoxy-2-mercaptobenzyl derivative **33** effectively removed in the presence of TFA [47]. Notably, the comparatively mild conditions for removal of trimethoxybenzyl (Tmb) auxiliary **33** enabled its application in the synthesis of glycopeptides [48], including fragments derived from human erythropoietin (EPO) bearing complex *O*- and *N*-linked glycans [49, 50]. Tmb derivative **33** has likewise been applied to the synthesis of a 62-amino acid SH3 domain of the actin cross-linking protein α-spectrin using a Lys-Gly disconnection site (Scheme 8) [47]. This ligation reaction was complete in 12 h, affording auxiliary-bound peptide **34** in 66% isolated yield. Cleavage of the Tmb auxiliary in the presence of TFA subsequently provided the target SH3 domain **35** [47].

Although *N*-linked thiol ligation auxiliaries have expanded the scope of ligation chemistry to include non-Cys ligation sites, such methods are often plagued by additional limitations. Harsh conditions for the removal of some auxiliaries (particularly strongly acidic conditions such as HF and TFMSA) limit the application of these methodologies to peptides and proteins bearing acid labile functionalities, including post-translational modifications such as glycosidic linkages. Furthermore, the techniques generally exhibit poor sequence tolerance at both the C-terminus of the thioester component and the N-terminus of the auxiliary-bound peptide, with the majority of successful auxiliary-mediated ligations exclusively utilizing junctions where one or both termini are Gly residues [23, 27]. Additional steric bulk at the ligation junction resulting from the tethered auxiliary, together



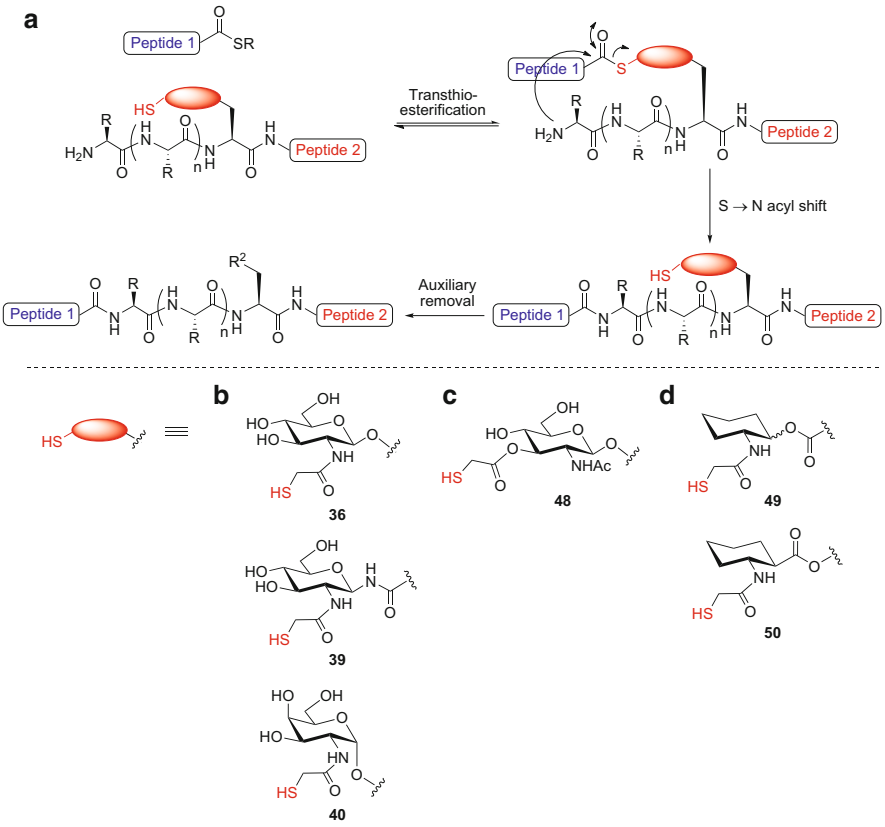
Scheme 8 Synthesis of the SH3 domain of α-spectrin using 4,5,6-trimethoxy-2-mercaptobenzyl ligation auxiliary **33**

with the requirement, in some cases, for larger ring-sized intermediates in the $S \rightarrow N$ acyl migration, greatly decrease the overall rate of ligation and limit general application of these auxiliaries in protein synthesis.

3.1.2 Side-Chain Ligation Auxiliaries

The development of side-chain ligation auxiliaries (Scheme 9) bearing reactive thiol functionalities tethered to the side-chain of an N-terminally located amino acid circumvented many of the issues associated with *N*-linked auxiliaries, including additional steric bulk at the terminal amine moiety [27, 51]. Mechanistically, the side-chain mediated process is very similar to native chemical ligation and *N*-linked auxiliary ligation, consisting of a thiol exchange reaction between a C-terminal peptide thioester and the side-chain thiol auxiliary, followed by an $S \rightarrow N$ acyl shift of the thioester intermediate to generate the new amide bond. A final step is then required to effect auxiliary removal and generate a native peptide product (Scheme 9a).

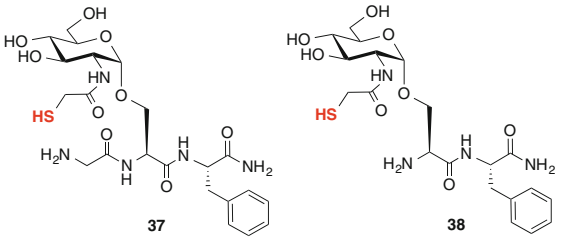
The first example of ligation via side-chain auxiliary, termed sugar-assisted ligation (SAL), was developed by Wong and coworkers in 2006 and involved the use of a glycopeptide whereby the reactive thiol auxiliary was appended to the C-2



Scheme 9 (**a**) Generalized ligation using side-chain auxiliaries. (**b**) First generation sugar-assisted ligation. (**c**) Second generation sugar-assisted ligation. (**d**) Side-chain assisted ligation

position of a β-*O*-linked carbohydrate moiety **36** (Scheme 9b) [52]. Ligation reactions were performed in aqueous buffer and, in contrast to ligation mediated by N-terminal auxiliaries, demonstrated a relatively broad sequence tolerance at the ligation junction [53]. Interestingly, reaction rates increased when the auxiliary-appended glycosylamino acid unit was incorporated as the penultimate residue as in model peptide **37**, rather than the terminal residue of the peptide fragment (compound **38**), despite the reliance on a larger ring-sized intermediate in the $S \rightarrow N$ acyl shift (Scheme 10) [52]. Following amide bond construction, the side-chain thiol functionality could be cleaved using a reductive desulfurization protocol (see below) [54] to yield glycopeptide products bearing a native acetamide at the C-2 position of the carbohydrate. The method was subsequently expanded to include *N*-linked [55] and α-*O*-linked [56] glycans (e.g., **39** and **40**, respectively, Scheme 9b). Enzymatic manipulation of the glycosylated peptide products (with or without the thiol handle present), including removal of the glycan [55] or elaboration of the appended monosaccharide unit through the action of glycosyltransferases [55, 57],

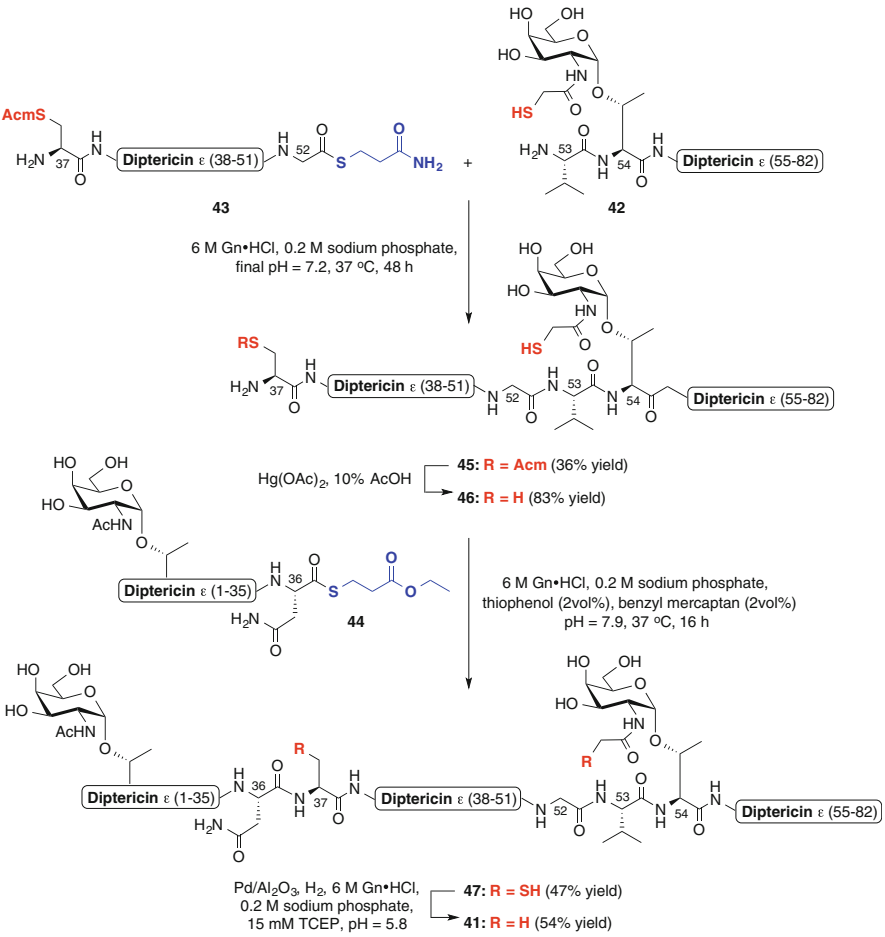
Scheme 10 Glycopeptides originally employed in sugar-assisted ligation (SAL)



further increased the scope and complexity of peptide and glycopeptide targets accessible using SAL.

Notably, the utility of SAL for the construction of glycoprotein targets was confirmed through the total chemical synthesis of diptericin ε **41**, an 82-amino acid, Cys-free antibacterial glycoprotein containing two galactosamine moieties α-*O*-linked to Thr^{10} and Thr^{54} (Scheme 11) [56]. It was envisaged that the glycoprotein could be synthesized from three segments, **42**, **43**, and **44**, in the *C*-to-*N* direction by employing a sequential SAL-native chemical ligation sequence, whereby the former is facilitated by a side-chain glycan auxiliary, and the latter a Cys residue installed as a temporary mutation. To this end, glycopeptide **42**, corresponding to the C-terminal region of diptericin ε and bearing a side-chain α-*O*-linked carbohydrate auxiliary at Thr^{54}, was first ligated to thioester **43** bearing an N-terminal Cys protected as the corresponding Cys acetamidomethyl (Acm) residue [58]. This sugar-assisted ligation was conducted in aqueous denaturing buffer at 37 °C and was complete in 48 h to afford peptide product **45**. Removal of the N-terminal Cys (Acm) group using mercury salts then provided glycopeptide **46**, bearing a free N-terminal thiol at position 37, poised for further functionalization. Accordingly, reaction between **46** and glycopeptide thioester **44** proceeded in 47% isolated yield to provide the 82-residue polypeptide chain **47**. A final global desulfurization reaction using conditions first described by Yan and Dawson [54] (explored in detail in Sect. 4.1) facilitated removal of the glycan-tethered thiol auxiliary as well as cleavage of the Cys side-chain thiol to generate diptericin ε **41**, with the native Ala residue at position 37 [56].

Although SAL in its original form greatly increased the number and flexibility of accessible ligation junctions, a major limitation of the technique was the reliance on a post-ligation reductive desulfurization protocol to facilitate cleavage of the glycan-appended auxiliary. These methods are incompatible with unprotected Cys thiols, and, as such, preclude the use of SAL for synthesis of peptides and proteins bearing native Cys residues. To address this shortfall, a second-generation SAL protocol was developed by Wong and coworkers in 2007, employing a modified auxiliary which could be removed in a mild and selective manner (Scheme 9c) [59]. In particular, this technique employed a thiol auxiliary with an ester linkage to the C-3 hydroxyl group of an *O*-linked glucosamine moiety **48**, which was easily removed via hydrazinolysis following ligation and without affecting unprotected Cys residues in the peptide sequence. Another important modification to the original SAL technique was the use of an organic cosolvent



Scheme 11 Synthesis of diptericin ε using sugar-assisted ligation (SAL)

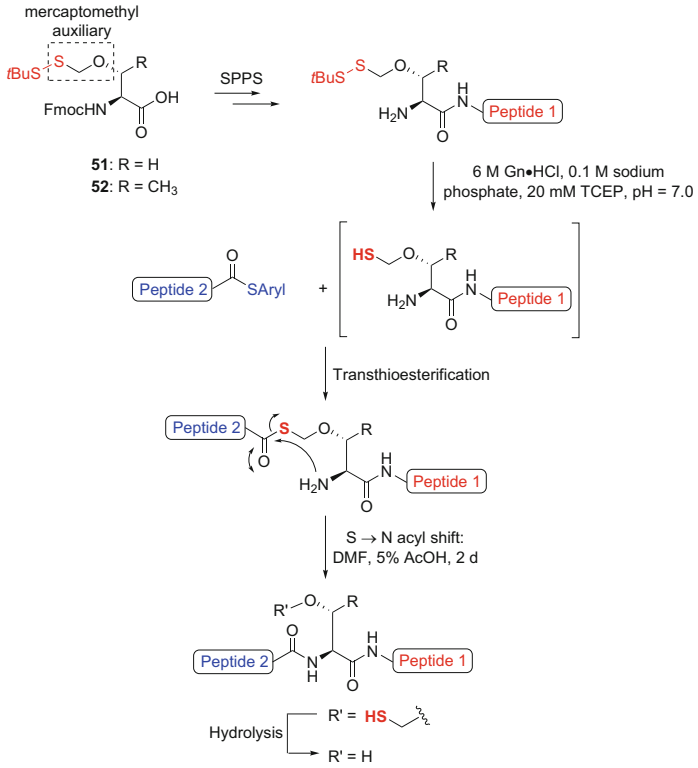
(e.g., *N*-methyl-2-pyrrolidone, NMP) in ligation reactions, serving to reduce hydrolysis of both the thioester component and the ester-linked auxiliary and therefore enhance ligation yields [59].

An additional SAL-inspired ligation strategy, developed in 2008 by Brik and coworkers, utilized a cyclohexyl or cyclopentyl ring auxiliary bound by an ester linkage to the side-chain of aspartic acid (Asp), glutamic acid (Glu), or serine (Ser) (e.g., **49** and **50**, Scheme 9d) [60]. In this method, the simplified side-chain carbocycle served an analogous role to the carbohydrate moiety in SAL. Following ligation, the auxiliary was rapidly cleaved in situ by the addition of NaOH [60]. Application of a Ser-linked cyclohexyl auxiliary using a sequential side-chain assisted ligation-native chemical ligation approach enabled the successful construction of the 86-residue polypeptide backbone of the regulatory protein

HIV-1 Tat [61]. Unfortunately, complications in the removal of the side-chain auxiliary hampered the total synthesis of the native protein in this study.

A remarkable feature of side-chain carbohydrate and cyclohexyl reaction auxiliaries is the ability to promote ligation effectively, despite the reliance of such systems on considerably larger ring size intermediates (14–15-membered) in the $S \rightarrow N$ acyl shift than those required for native chemical ligation or for N-terminal auxiliary-mediated ligation. As previously noted, SAL reactions employing the original carbohydrate bearing a C-2-linked thiol auxiliary proceeded faster with a single amino acid extension N-terminal to the glycosylamino acid moiety than when the glycosylamino acid was N-terminally located (see Scheme 10). These results suggest that the side-chain auxiliary may play an important role in appropriately positioning the intermediate bridged thioester for attack by the N-terminal amine [52]. Detailed studies probing the effect of multiple amino acid extensions N-terminal to the glycosylamino acid auxiliary in SAL [59, 62] demonstrated that ligation is feasible (though substantially slower) with as many as six amino acids appended to the N-terminus of the auxiliary-bound residue. However, it should be noted that ligation reactions bearing such large N-terminal extensions are likely to proceed, at least in part, via a direct aminolysis pathway [63, 64].

Another novel side-chain thiol auxiliary approach was developed in 2010 by Hojo et al. for ligation reactions at N-terminal Ser and Thr residues [65]. This method optimized the use of a mercaptomethyl group attached as a thiohemiacetal to the side-chain of Ser or threonine (Thr) (Scheme 12). As with native chemical ligation, the auxiliary was found to promote ligation through initial transthioesterification with a C-terminal peptide thioester followed by an $S \rightarrow N$ acyl shift, in this instance through a seven-membered ring intermediate, to generate a new amide bond. The inherent instability of the thiohemiacetal functional group was overcome, in part, through incorporation of the auxiliary-bound residue into peptides as the corresponding asymmetric disulfide **51** or **52**, allowing an in situ generation of the free auxiliary upon treatment with TCEP in the ligation reaction (Scheme 12). Furthermore, ligations were performed with preformed aryl thioesters to enhance the rate of the initial transthioesterification reaction. Under these conditions, rearrangement of the thioester intermediate through $S \rightarrow N$ acyl migration proved to be the rate-limiting step. To prevent hydrolysis of the unrearranged intermediate, the aqueous reaction buffer was diluted with DMF containing 5% acetic acid after the initial transthioesterification and left for 2 days to promote rearrangement. Following ligation, the susceptibility of the auxiliary to hydrolysis became a strategic advantage, whereby spontaneous cleavage of the thiohemiacetal afforded native products in a one-pot fashion. The method was utilized in the synthesis of the glycopeptide toxin contulakin-G, derived from the venom of *Conus geographus*, and for the preparation of human calcitonin [65].

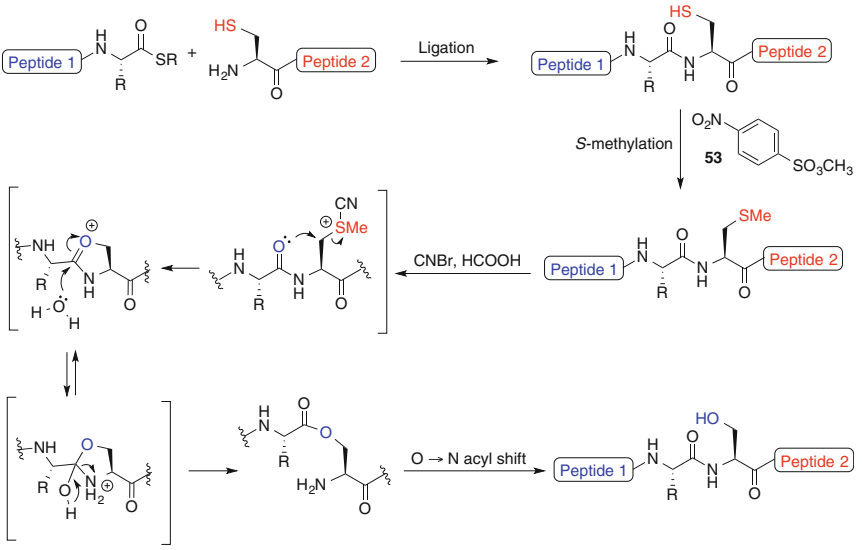


Scheme 12 Ligation at Ser and Thr using a mercaptomethyl side-chain auxiliary

4 Post-Ligation Manipulations

The contributions of N-terminal and side-chain auxiliary-mediated ligations have served to enhance greatly the scope of ligation chemistry beyond the original reliance on N-terminal Cys residues. Despite these successes, auxiliary methods generally require the multi-step preparation of specialized thiol auxiliaries and auxiliary-bound peptides, which reduces the overall efficiency of the techniques. As previously discussed, auxiliary-mediated ligations are also slower than native chemical ligation at Cys, requiring lengthy reaction times whereby hydrolysis and epimerization become significant competing pathways. As such, a conceptually appealing approach to increasing the scope of native chemical ligation without sacrificing the simplicity and efficiency of the technique is to explore the post-ligation modification of Cys residues [54, 66] for the generation of target peptides and proteins.

An innovative demonstration of the manipulation of Cys residues to expand the scope of accessible ligation junctions was reported in 2008 by Okamoto and Kajihara [67]. With the aim of preparing complex glycopeptides and proteins, the

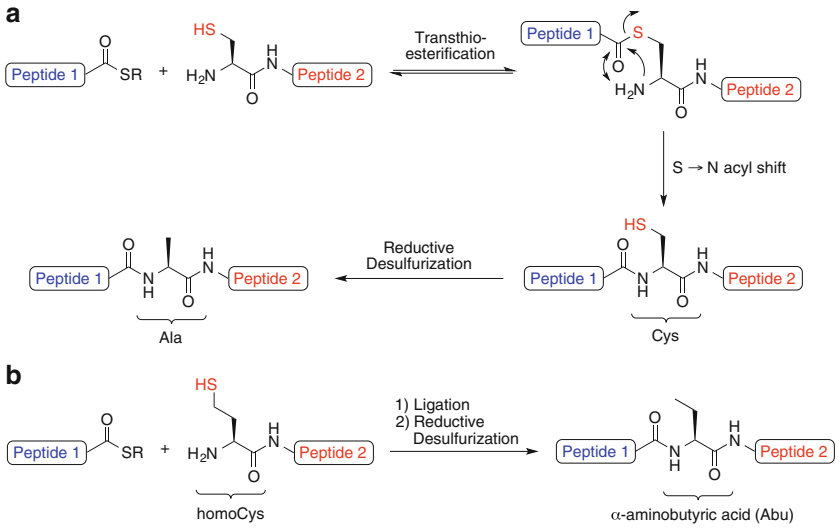


Scheme 13 Conversion of Cys to Ser following native chemical ligation

authors demonstrated a post-ligation conversion of Cys to the corresponding Ser residue using a three-step protocol (Scheme 13). The method first involved *S*-methylation of the Cys residue using methyl 4-nitrobenzenesulfonate **53** [68]. Subsequent reaction with CNBr in the presence of formic acid leads to the activation of the *S*-methyl group [69], facilitating attack of the *S*-methyl Cys β-carbon by the carbonyl oxygen of the neighboring amide bond. The resultant five-membered ring is converted to the *O*-ester peptide intermediate, which undergoes a subsequent O → N acyl shift at slightly basic pH (7–8) to generate the new amide bond. Initial application of this protocol on model systems demonstrated the overall feasibility of the approach. The potential for reaction of CNBr with Met residues, however, required the incorporation of Met as the corresponding sulfoxide so as to allow selectivity for the methyl Cys residues. Reduction of the sulfoxide to yield a native Met residue was accomplished following the conversion of Cys to Ser. The utility of the Cys to Ser transformation in the synthesis of complex glycopeptide fragments was also demonstrated through the construction of a glycopeptide fragment of erythropoietin, bearing a complex, *N*-linked glycan, along with the synthesis of a MUC1 40-mer peptide containing two copies of the T_N antigen [67].

4.1 Ligation-Desulfurization

In 2001, Yan and Dawson reported an elegant approach for the disconnection of peptide and protein targets at alanine (Ala) residues [54]. The concept combined a typical native chemical ligation reaction between a peptide bearing an N-terminal



Scheme 14 (**a**) Ligation at Cys and post-ligation reductive desulfurization to Ala. (**b**) Ligation at homoCys followed by reductive desulfurization to α-aminobutyric acid (Abu)

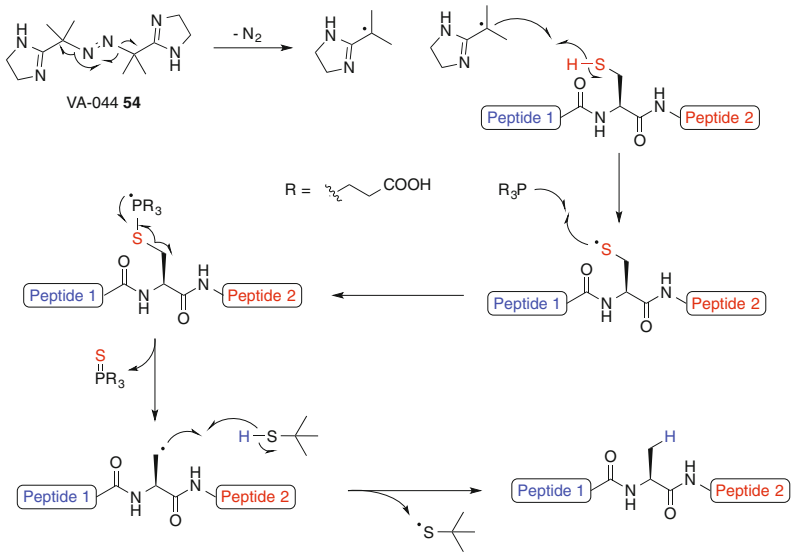
Cys residue and a C-terminal peptide thioester with a post-ligation, reductive desulfurization protocol which selectively removed the sulfhydryl moiety in the Cys side-chain to generate the corresponding Ala residue at the ligation junction (Scheme 14a). The technique combined the efficiency of Cys-promoted ligations while enabling access to ligation junctions containing the significantly more abundant Ala residue (8.9%) [33], thereby enabling the synthesis of a variety of previously inaccessible peptide and protein targets. To optimize the desulfurization protocol, the authors treated Cys-containing ligation products with a variety of metal reagents, including Pd on Al_2O_3, Pd on carbon, PdO and Raney Ni, in the presence of hydrogen gas. Excellent results were obtained with Pd on Al_2O_3, which effected high-yielding, global desulfurization of Cys residues while minimizing the over-reduction of the aromatic proteinogenic amino acids tyrosine (Tyr), phenylalanine (Phe), and tryptophan (Trp). Rapid and efficient desulfurization with Raney Ni was also demonstrated, although demethylthiolation of Met residues was evident with prolonged reaction times [54].

In the original publication, Yan and Dawson also demonstrated the post-ligation reduction of a homocysteine (homoCys) residue to the corresponding α-aminobutyric acid (Abu) (Scheme 14b) [54]. This application built upon earlier work by Tam and Yu [70], which demonstrated that ligation at homoCys followed by *S*-methylation provided a feasible approach to ligation at Met residues. The extension of the desulfurization protocol to include non-Cys thiols established the generality of the technique and led to the prescient notion of utilizing unnatural, thiol-derived amino acid derivatives to effect ligation (see Sect. 4.3 for more details) [54]. Importantly, the authors also demonstrated the utility of the Cys

ligation-desulfurization strategy in diverse peptide and protein systems through the syntheses of a cyclic microcin J25-like peptide, the 56-amino acid streptococcal protein G B1 (PGB1) domain, and [Ala^{49}]barnase, prepared by desulfurization of the previously reported 110-amino acid [Cys^{49}]barnase analogue [15].

Following this seminal report, research efforts have shifted to improving the selectivity of the desulfurization protocol to preclude the desulfurization of Cys residues (and other sulfur-containing functionalities) crucial to the native peptide sequence. Pentelute and Kent reported the desulfurization of unprotected Cys residues with Raney Ni in the presence of acetamidomethyl (Acm)-protected Cys residues, which could be unmasked following desulfurization to generate the free sulfhydryl side-chain of Cys [58]. In their work on sugar-assisted ligation (SAL), Wong and coworkers independently reported that the use of hydrogen gas and Pd/Al_2O_3 could effect selective removal of a glycan-bound thiol auxiliary in the presence of a Cys(Acm) residue [56].

However, broader issues with the use of metal-based reagents, including the adsorption of specific peptide sequences to the metal surface [71] and the undesired reduction of Met and thiazolidine-protected Cys residues [72], prompted the search for a milder, metal-free desulfurization protocol. To this end, Danishefsky and coworkers turned to a radical-promoted desulfurization method [72] based on an initial report by Hoffman et al. 50 years earlier on the desulfurization of thiols, under thermal or photochemical conditions, in the presence of trialkylphosphites [73]. The subsequent use of trialkylphosphines to promote desulfurization was also reported [74]. With the goal of mild and selective desulfurization of peptide systems in mind, Danishefsky and coworkers specifically employed the water-soluble phosphine TCEP, owing to its stability, ease of handling, and proven compatibility with proteinogenic amino acid side-chains and glycopeptide functionalities. Indeed, in the presence of TCEP, the water-soluble radical initiator 2,2′-azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride (VA-044) **54**, and *t*BuSH in aqueous media, the selective desulfurization of Cys residues was readily achieved. The mechanism for this transformation presumably mirrors the one proposed by Walling et al. [75, 74] and involves the initial formation of a Cys thiyl radical in the presence of radical initiator **54** (Scheme 15). Addition of the sulfur-centered radical to TCEP generates a TCEP-adduct, which undergoes β-scission to produce an alanyl radical and a phosphine sulfide byproduct. Hydrogen abstraction from an exogenous thiol by the alanyl radical then generates the native Ala residue. Importantly, these conditions were shown to tolerate Cys(Acm) groups, Met residues, thiazolidine groups, and C-terminal thioester moieties. In the initial report, Danishefsky and coworkers employ the metal-free desulfurization protocol in a ligation-desulfurization approach to an *N*-linked glycopeptide fragment and the cyclic nonapeptide crotogossamide [72].

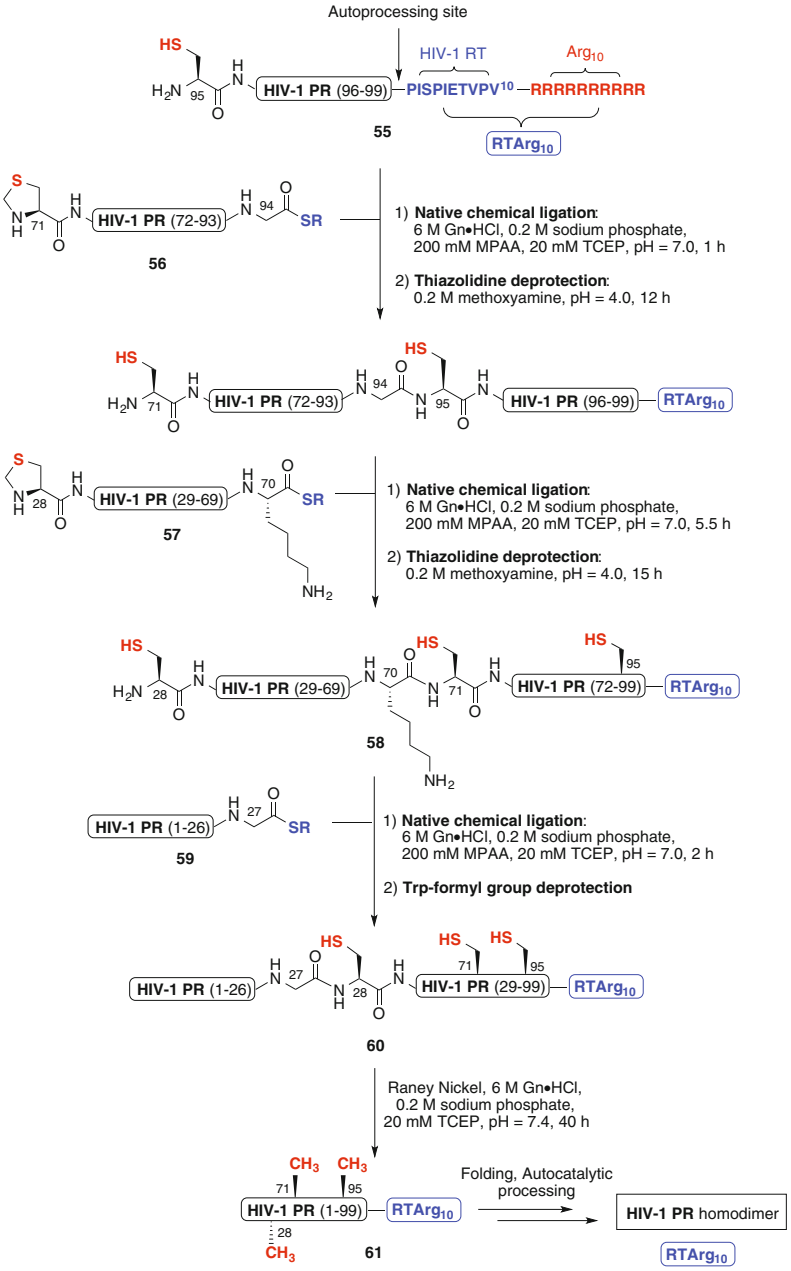


Scheme 15 Radical desulfurization of Cys residues initiated by VA-044 **54**

4.2 Ligation-Desulfurization in Protein Synthesis

Since its inception, the concept of ligation-desulfurization chemistry [76, 77] at Cys residues has been widely adopted for the synthesis of an enormous variety of peptides and proteins, including targets bearing complex PTMs [20, 21, 28, 78]. The following discussion outlines a number of recent, illustrative examples of the technique for the ligation-based assembly of protein targets.

In 2007, Kent and coworkers reported an elegant synthesis of a fully functional, homodimeric HIV-1 protease (PR) by combining ligation-desulfurization chemistry with a biomimetic autoprocessing strategy [79], taking advantage of the ability of the HIV-1 protease to catalyze its own removal from the Gag-Pol polyprotein precursor during HIV-1 maturation in vivo. In their initial studies toward the total chemical synthesis of the 99-amino acid HIV-1 PR monomer using ligation chemistry, the authors encountered considerable difficulties in solubilizing peptide intermediates, particularly the C-terminal fragment. A revised strategy therefore incorporated a C-terminal poly-Arg tag to aid in the solubility of the terminal fragment and subsequent ligation intermediates (Scheme 16). In order to facilitate removal of the solubility tag following construction of the monomeric polypeptide backbone, the poly-Arg sequence was attached to the C-terminus of the protease using a ten-residue linker sequence derived from the HIV-1 reverse transcriptase (RT) protein (which is proximally located in the Gag-Pol polyprotein precursor). The C-terminal construct **55** would therefore contain an autoprocessing site, and it was envisaged that the folded protease would readily cleave the modification.

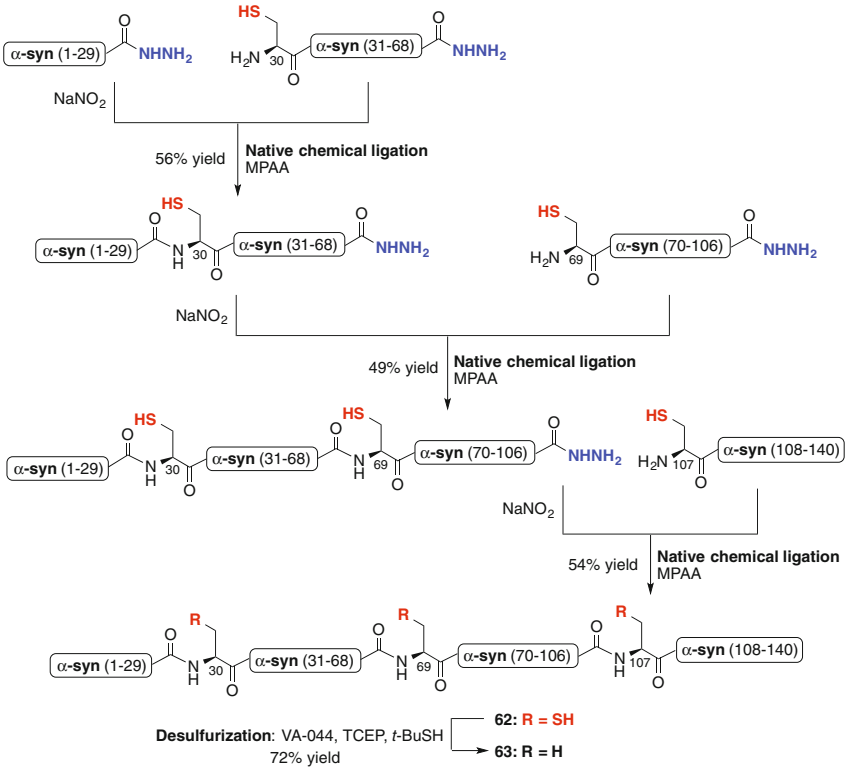


Scheme 16 Total synthesis of HIV-1 protease using an iterative ligation-autoprocessing strategy

Accordingly, construction of the full-length peptide was accomplished using a four-component iterative ligation-deprotection strategy in the *C*-to-*N* direction, whereby bifunctional thioesters **56** and **57** were prepared as N-terminal thiazolidines which could be easily unmasked following a ligation reaction to facilitate extension in the N-terminal direction (Scheme 16). As such, iterative ligation-thiazolidine deprotections employing peptide fragments **55**, **56**, and **57** rapidly afforded 28-99RTArg$_{10}$ **58**. After ligation of N-terminal thioester **59** to 28-99RTArg$_{10}$ **58** and removal of Trp formyl protecting groups, the 119-residue polypeptide **60** was globally desulfurized in the presence of Raney Ni to afford 1-99RTArg$_{10}$ **61** in 26% overall yield based on starting peptide **55**. As anticipated, autoprocessing of the C-terminal modification occurred concomitantly with folding of the protease to generate the target homodimeric protein (Scheme 16). The biological activity of the synthetic enzyme was further confirmed in a kinetic assay and the structure validated using X-ray crystallography [79].

Liu and coworkers recently reported the total chemical synthesis of α-synuclein, a Cys-free protein implicated in the development of Parkinson's disease, in another impressive application of ligation-desulfurization chemistry to the total chemical synthesis of proteins [80]. Herein, the authors utilize a four-component *N*-to-*C* ligation approach which takes advantage of C-terminal peptide acyl hydrazides, a technology pioneered in the Liu laboratory, as thioester surrogates for ligation chemistry [81, 82]. Briefly, the method first involves the activation of fully deprotected C-terminal peptide hydrazides with the oxidant $NaNO_2$, which chemoselectively affords an intermediate acyl azide. Thiolysis of the acyl azide using an aryl thiol (e.g., MPAA) then promotes the in situ formation of the peptide thioester, which is poised for use in ligation chemistry. Importantly, peptide hydrazides are easily prepared using Fmoc-SPPS and are able to serve as masked thioesters in iterative ligation strategies because of the requirement for an initial activation step (oxidation to the acyl azide). As such, the reactivity of bifunctional peptides containing a C-terminal hydrazide and an N-terminal Cys residue are carefully controlled to promote protein synthesis using iterative ligation chemistry. The ability of hydrazides to serve as latent thioesters was strategically exploited in the *N*-to-*C* construction of α-synuclein using an iterative oxidation-ligation approach (Scheme 17). To facilitate ligation, Cys mutants were temporarily introduced at residues 30, 69, and 107 in place of the native Ala residues. Following the iterative ligation sequence and construction of the full-length α-synuclein (1–140, A^{30,69,107}C) **62**, high-yielding conversion to the native protein **63** was accomplished using a global radical desulfurization protocol employing the conditions developed by Wan and Danishefsky [72].

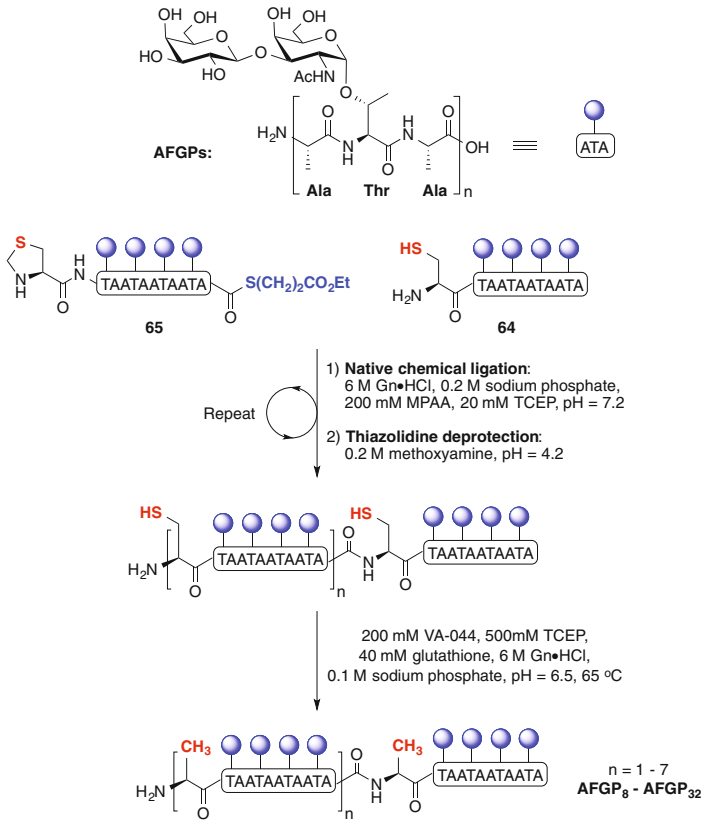
Post-ligation desulfurization has also been extensively applied to the synthesis of proteins bearing post-translational modifications. In 2012, Wilkinson et al. reported the construction of a library of homogeneous antifreeze glycoproteins (AFGPs) using ligation-desulfurization chemistry [83]. AFGPs are mucin-type glycoprotein natural products isolated from select Arctic and Antarctic fish, where they play a critical role in preventing the growth of ice crystals. Structurally, AFGPs are composed of multiple copies of the repeating tripeptide Ala-Thr-Ala/Pro, in



Scheme 17 Total synthesis of α-synuclein through an iterative ligation-desulfurization strategy employing peptide acyl hydrazides

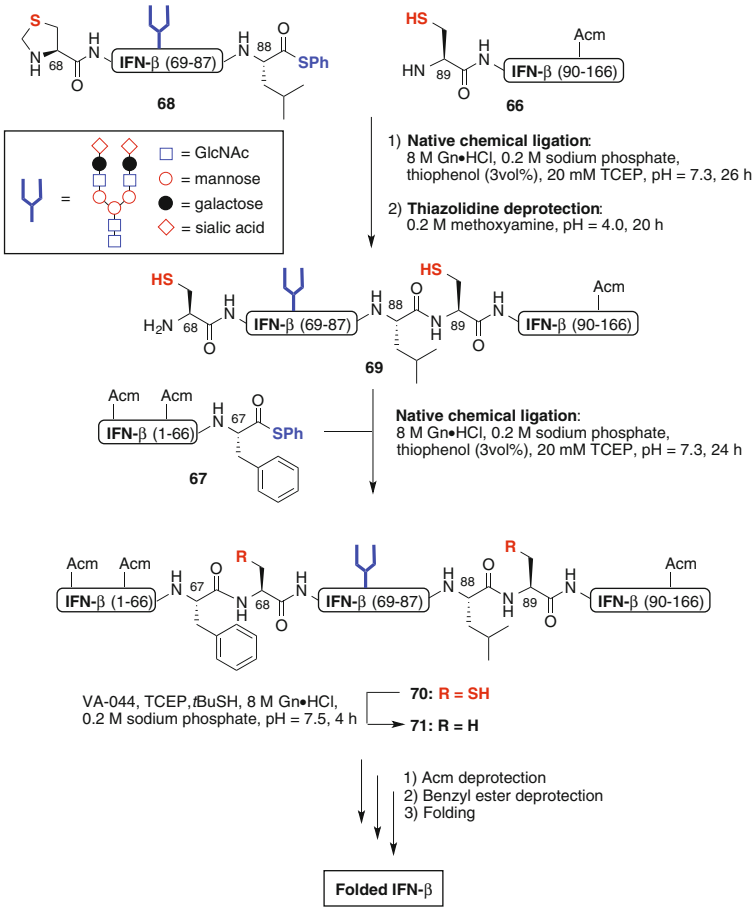
which each Thr residue is α-*O*-linked to the disaccharide β-D-galactosyl-(1→3)-α-*N*-acetyl-D-galactosamine (Scheme 18) and range in size from 4 to 50 repeat units [84, 85]. In an effort to access large quantities of homogeneous AFGPs for biological studies and applications in materials science, the authors designed a convergent ligation approach to homogeneous AFGPs bearing between 4 and 32 repeat units. Specifically, peptide fragment **64**, bearing an N-terminal Cys residue, and bifunctional peptide **65**, bearing an N-terminal thiazolidine and a C-terminal thioester, were used in iterative ligation chemistry to assemble increasingly large AFGP repeat units (Scheme 18). Upon reaching the desired chain length, the Cys ligation handles were readily converted to the native Ala residues via radical desulfurization [72] in the presence of VA-044, TCEP and glutathione [71] as a source of hydrogen atoms. The resulting library of homogeneous AFGPs (ranging in size from 1.2 to 19.5 kDa) enabled a comprehensive study of the effect of chain length on thermal hysteresis and ice recrystallization inhibition activities [83].

Kajihara and coworkers recently reported the total chemical synthesis of two glycoforms of the 166-amino acid human glycosyl-interferon-β, bearing a complex



Scheme 18 Construction of homogeneous antifreeze glycoproteins (AFGPs) using ligation-desulfurization chemistry

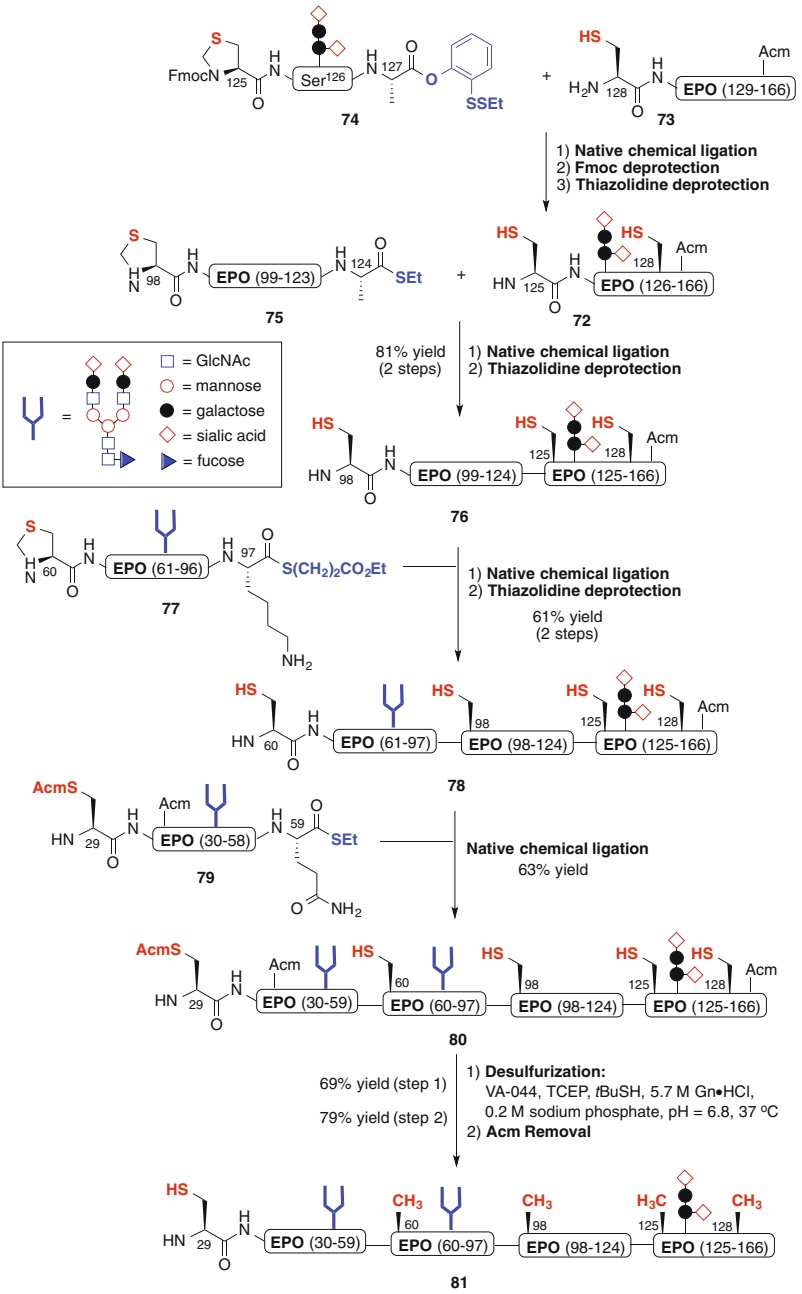
N-linked sialyl or asialo biantennary oligosaccharide [86]. Using a three-component synthetic strategy, the target was disconnected at two Ala ligation junctions, accessible via ligation-desulfurization chemistry (Scheme 19). Notably, interferon-β also contains three native Cys residues in the full-length sequence. However, the location of these residues was deemed unsuitable for the facile construction of the target protein. As such, the native Cys residues were incorporated as the corresponding Cys(Acm) residues in the target peptide fragments **66** and **67** to facilitate construction of the protein using ligation-desulfurization chemistry (Scheme 19). Initial ligation of peptide **66** and glycopeptide thioester **68**, bearing an N-terminal thiazolidine, was accomplished under standard aqueous conditions in the presence of thiophenol as a ligation catalyst. Removal of the thiazolidine afforded **69**, which was subsequently ligated with fragment **67** to afford the 166-amino acid polypeptide **70**. At this stage, radical desulfurization enabled conversion of the two ligation site Cys residues to the native Ala residues, affording glycopeptide **71**. Removal of the Cys(Acm) groups upon treatment with AgOAc in 90% AcOH and saponification of the benzyl ester protecting group (in the synthesis



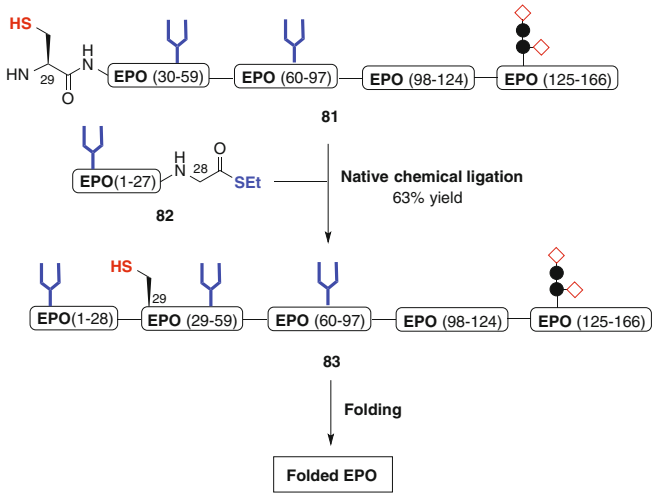
Scheme 19 Construction of homogeneous human glycosyl interferon-β (IFN-β) using ligation-desulfurization chemistry

of the sialylated glycoprotein variant) afforded the native protein glycoforms. Upon folding, both glycosylated variants of interferon-β were shown to suppress tumor growth in vivo [86].

Perhaps one of the most demonstrative applications to date of the power of ligation-desulfurization chemistry has been in the total chemical syntheses of the human glycoprotein hormone, erythropoietin (EPO) [87–90]. The most recently reported example described the first synthesis of the 166-amino acid glycoprotein as single glycoforms bearing natively linked glycans [87, 88]. In this groundbreaking work [91], the Danishefsky laboratory accessed multiple Ala ligation sites by employing a post-ligation metal-free radical desulfurization protocol [72] together with judicious protection of the native Cys residues, including those inappropriately positioned for ligation chemistry. A summary of the ligation strategy is shown in Schemes 20 and 21 [88].

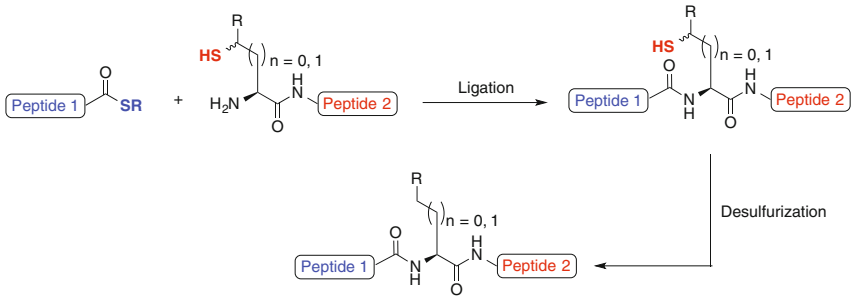


Scheme 20 Synthesis of EPO(29–166) **81** using ligation-desulfurization chemistry



Scheme 21 Total synthesis of homogenous EPO(1–166) as a single glycoform

To begin, C-terminal fragment EPO(125–166) **72**, bearing a Ser-linked glycophorin moiety, was first prepared via ligation of fragment **73** with the short glycopeptide **74** bearing a latent thioester moiety [92, 93] followed by an Fmoc deprotection and unmasking of the N-terminal thiazolidine (Scheme 20). Ligation of **72** with thioester EPO(98–124) **75** and subsequent thiazolidine deprotection then provided peptide **76** in 81% yield over the two steps. Glycopeptide thioester **77**, corresponding to EPO(60–97) and bearing a complex *N*-linked sialyl biantennary glycan, was next ligated to peptide fragment **76**. Another thiazolidine deprotection, followed by ligation with glycopeptide thioester EPO(29–59) **79**, afforded peptide **80** corresponding to residues 29–166 of the protein sequence. At this point, radical desulfurization of the four ligation-site Cys residues was accomplished in 69% yield to afford the corresponding Ala residues (Scheme 20). Removal of the Acm protecting groups then unmasked the native Cys residues, including N-terminal Cys^{29} to afford **81**, which was appropriately positioned for a subsequent ligation with glycopeptide thioester EPO(1–28) **82** to afford the full-length, native polypeptide sequence **83** (Scheme 21). Importantly, upon folding, chemically-derived EPO displayed potent erythropoietic activity in both in vitro and in vivo assays [88]. The size and complexity of the target EPO glycoform pushed the limits of modern protein synthesis and thus served as a potent validation of the utility of ligation-desulfurization chemistry for the construction of post-translationally modified proteins.



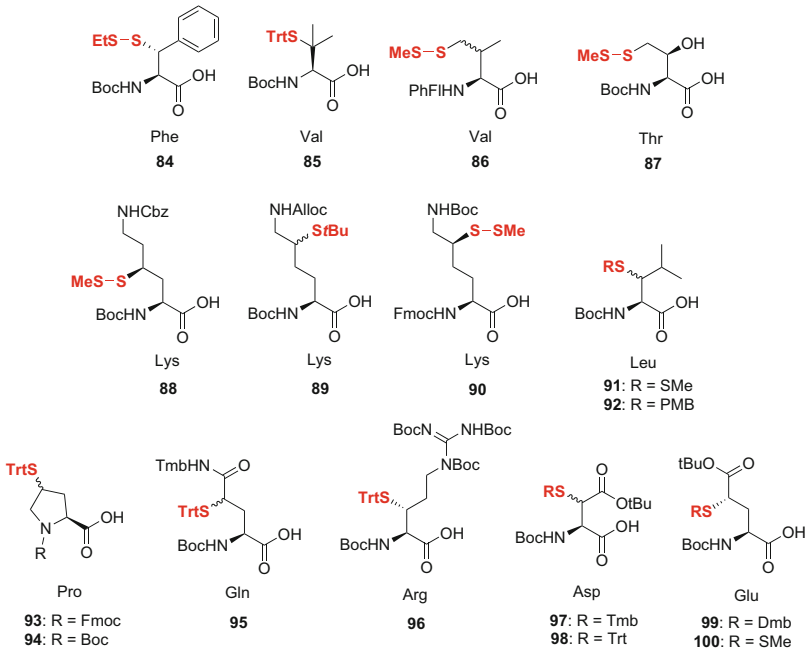
Scheme 22 Ligation-desulfurization chemistry at β- ($n = 0$) or γ-thiol ($n = 1$) amino acids

4.3 Ligation-Desulfurization at Thiol-Derived Amino Acids

The impact of Yan and Dawson's seminal work [54] on post-ligation desulfurization has extended far beyond access to Ala ligation junctions. As previously discussed in their seminal report, the authors also established the intellectual framework for ligation at a variety of additional non-Cys sites by demonstrating ligation-desulfurization at homoCys residues. These results prompted the proposal that unnatural, β- or γ-thiol amino acid derivatives could be utilized in a similar manner to enable disconnections at other proteinogenic amino acids (Scheme 22) [54]. This idea has fuelled an intense focus on the development of unnatural, thiol-derived amino acids for use in ligation reactions (Schemes 22 and 23) [76, 77]. A concise overview of the synthetic strategies employed in the construction of these valuable building blocks has recently been reported [94]. As such, the following discussion serves to outline the application of these novel building blocks (e.g., **84**–**100**, Scheme 23) in ligation-desulfurization chemistry for the synthesis of peptides and proteins.

4.3.1 Phenylalanine

The first application of post-ligation desulfurization beyond Cys to Ala conversions was in the demonstration of ligation disconnections at Phe residues [95, 96]. In 2007, Crich and Banerjee reported the synthesis of β-thiol derivative **84** (Scheme 23), beginning with L-Phe methyl ester. The synthetic pathway utilized chemistry originally developed for the bromination of the benzylic position of aromatic amino acid residues and subsequent conversion to the β-hydroxy analogues [97–99]. Synthetic Phe building block **84** was first shown to facilitate ligation reactions with amino acid thioesters. Following incorporation into a model peptide **101** using Fmoc-SPPS, the building block was successfully used to mediate ligation with peptide thioesters **102** and **103**, bearing C-terminal Gly and Met residues, respectively, in good yields (Scheme 24). Removal of the β-thiol



Scheme 23 β- and γ-thiol amino acid derivatives for ligation-desulfurization chemistry

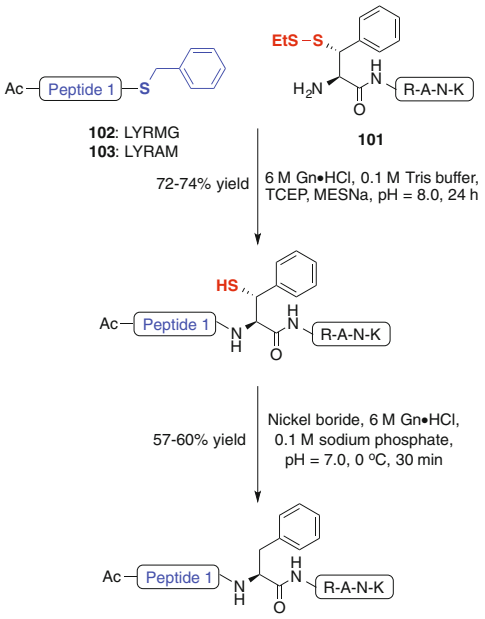
moiety following ligation was achieved via hydrogenolytic desulfurization with nickel boride, thereby generating the native Phe residue at the ligation junction [96]. The results of this initial study served as a critical proof-of-concept for the development of subsequent thiol-derived proteinogenic amino acids.

4.3.2 Valine

Shortly after the report of ligation at Phe building block **84**, Seitz and coworkers reported the first strategy for ligation at Val by employing a commercially available penicillamine building block (Boc-Pen(Trt)-OH, **85**, Scheme 23) [71]. Following incorporation into peptides, penicillamine-mediated ligations were successfully demonstrated for peptide thioesters bearing C-terminal Gly, His, Met, and Leu residues, proceeding in 12–48 h. The relatively lengthy reaction times were attributed to the additional steric bulk associated with the use of a tertiary thiol. Following ligation, application of a slight variation on Wan and Danishefsky's metal-free radical desulfurization protocol [72], using glutathione (rather than *t*BuSH) as the hydrogen atom donor, afforded native peptide products in excellent yields [71].

Danishefsky and coworkers independently reported access to Val ligation junctions via a synthetic γ-thiol Val building block **86** [100]. Using a ten-step

Scheme 24 Ligation-desulfurization at β-thiol phenylalanine



synthesis beginning with Fmoc-Asp-O*t*Bu, the authors were able to access both diastereomers of **86**. Following incorporation into model peptides, comparative rate studies revealed that ligations mediated by both diastereomers of γ-thiol Val proceeded significantly faster than the corresponding reactions mediated by penicillamine (β-thiol Val), despite the requirement for a six-membered ring intermediate in the $S \rightarrow N$ acyl shift for the γ-thiol variants. The enhanced rate of ligation relative to penicillamine is likely owing to the greater reactivity and decreased steric bulk of the primary thiol of building block **86** relative to the tertiary thiol of **85** (Scheme 23). Ligation at γ-thiol Val was also utilized in the high-yielding synthesis of a glycopeptide bearing an *N*-linked disaccharide through ligation of a peptide bearing an N-terminal γ-thiol Val residue with a glycopeptide *ortho*-thiophenolic ester [100].

4.3.3 Threonine

The Danishefsky laboratory also reported a ten-step synthetic approach to γ-thiol Thr **87** from H-Met-OMe (Scheme 23), expanding the repertoire of ligation-desulfurization chemistry to include Thr ligation sites [101]. Following incorporation of the building block into peptides, its utility in ligation-desulfurization chemistry was demonstrated in a number of diverse model systems. Notably, the γ-thiol Thr residue was capable of facilitating ligation with a variety of C-terminal acyl donors, including at sterically encumbered C-terminal Val, Ile, and Pro residues, through activation as the corresponding *p*-nitrophenyl ester acyl donors.

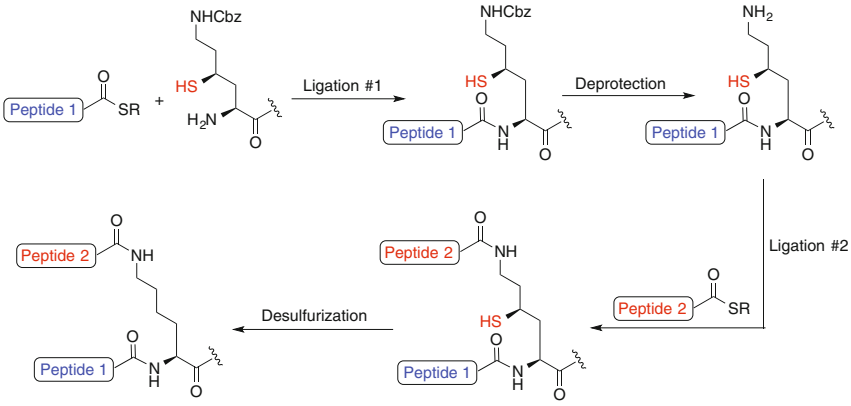
Application of the Thr ligation strategy followed by radical desulfurization also enabled the synthesis of a glycopeptide bearing a complex *N*-linked hexasaccharide moiety [101].

4.3.4 Lysine

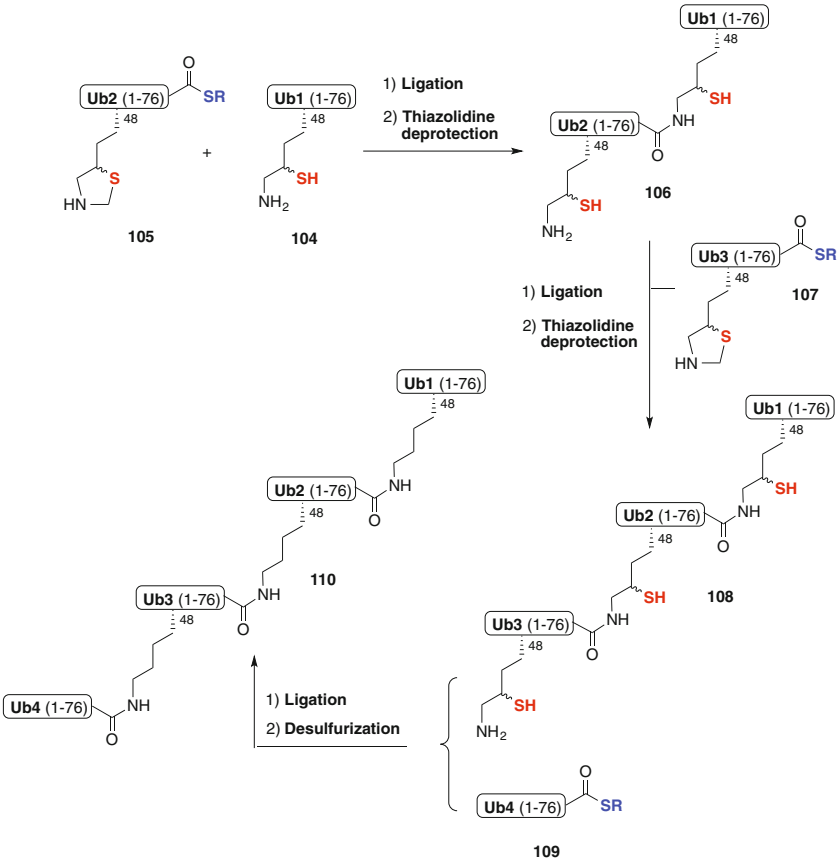
In 2009, Liu and coworkers published the synthesis of a γ-thiol Lys derivative **88** which was capable of mediating ligation at both the α- and ε-amino groups of Lys (Scheme 25), each via a six-membered ring intermediate in the $S \rightarrow N$ acyl shift [102]. Specifically, a side-chain Cbz-protected γ-thiol Lys derivative was first ligated at the α-amino group. Deprotection of the Cbz group unmasked the ε-amino group for a subsequent ligation, and a final desulfurization protocol rendered a native Lys residue, derivatized at both the α- and ε-positions (Scheme 25). This dual ligation protocol is particularly attractive for the synthesis of post-translationally modified peptides and proteins given the variety of functionalization occurring naturally at the ε-amino moiety of Lys, including acetylation, ubiquitylation, and methylation. Access to the key γ-thiol Lys building block **88** (Scheme 23) was accomplished in 16 steps from Fmoc-Asp-O*t*Bu by first employing the method of Guichard and coworkers [103] to access a 4-hydroxy-Lys derivative bearing a side-chain Cbz protecting group. Lys building block **88** was shown to effectively mediate a number of ligations at both the α- and ε-amino groups, and was used in the preparation of side-chain ubiquitylated and biotinylated peptide products [102].

Two δ-thiol Lys derivatives, **89** [104] and **90** [105], have been independently reported for use in side-chain ligation approaches to ubiquitylation. In particular, δ-thiol Lys building block **89** has been extensively employed by Brik and coworkers to study the ubiquitylation of α-synuclein [104, 106], a 140-amino acid presynaptic protein implicated in a number of neurodegenerative diseases, through protein semi-synthesis. Ovaa and coworkers have also reported the incorporation of Lys building block **90** in place of native Lys residues in the 76-amino acid ubiquitin protein sequence to facilitate the synthesis of a library of diubiquitin conjugates using ligation-desulfurization chemistry [105]. The same group recently reported a concise synthetic route to an additional γ-thiol Lys derivative and determined that both γ- and δ-thiol Lys are equally efficient in facilitating the synthesis of diubiquitin conjugates [107].

Another impressive display of the applicability of thiol-derived amino acid building blocks in the synthesis of large proteins through ligation-desulfurization chemistry was the total chemical synthesis of a 304-amino acid tetraubiquitin protein by Brik and coworkers in 2011 [108]. The polyubiquitin chain was constructed using an iterative ligation approach facilitated by a δ-thiol Lys residue positioned at Lys^{48} in the ubiquitin chain (Scheme 26). Ligation of Ub1 fragment **104**, bearing a δ-thiol Lys residue, with Ub2 thioester **105**, containing a thiazolidine protected δ-thiol Lys residue, was accomplished using standard ligation conditions in the presence of exogenous benzylmercaptan and thiophenol. Removal of the



Scheme 25 Dual native chemical ligation at Lys mediated by a γ-thiol Lys derivative



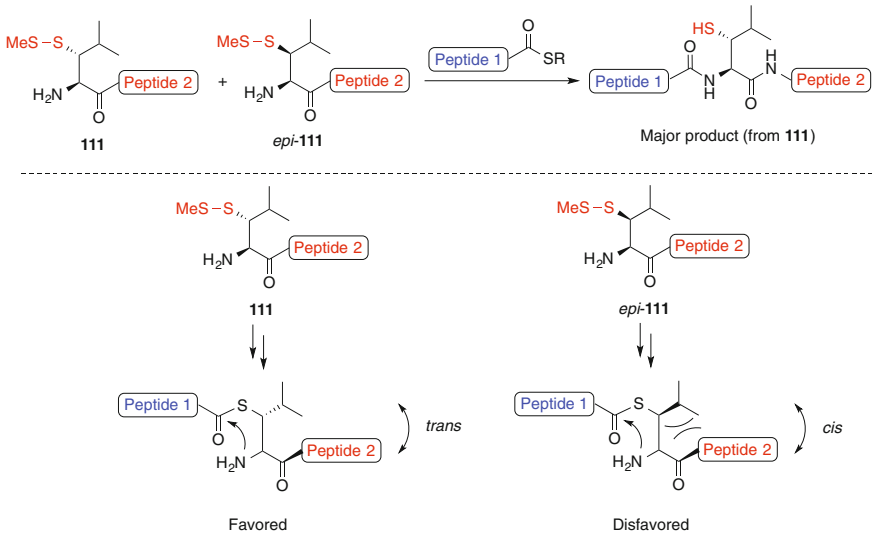
Scheme 26 Synthesis of a 304-amino acid tetraubiquitin protein using ligation-desulfurization at δ-thiol Lys

thiazolidine positioned diubiquitin **106** for subsequent reaction with another Ub thioester, Ub3 **107**. Thiazolidine deprotection and a final ligation of triubiquitin **108** with ubiquitin thioester Ub4 **109**, bearing a native Lys residue at position 48, then afforded a tetraubiquitin adduct containing three unnatural δ-thiol Lys residues. A final conversion of the unnatural amino acid derivatives to the native Lys residues using a global radical desulfurization protocol furnished the 304-amino acid tetraubiquitin **110** [108], currently the second largest protein prepared to date by total chemical synthesis.

4.3.5 Leucine

Ligation at Leu has been independently demonstrated by Danishefsky and coworkers [109] and Brik and coworkers [110], each developing a seven-step synthetic approach to suitably protected Leu building blocks, **91** and **92** (Scheme 23), respectively, beginning with commercially available β-hydroxy-L-Leu. Danishefsky and coworkers prepared both β-epimers of the target Leu building block **91** by beginning the synthesis with both *erythro*- and *threo*-β-hydroxy-L-Leu precursors [109]. In contrast, Brik and coworkers reported the synthesis of a single diastereomer of β-thiol Leu **92**, beginning exclusively with *threo*-β-hydroxy-L-Leu [110]. Following the preparation of the target building blocks, both groups demonstrated the utility of the Leu derivatives in ligation-desulfurization chemistry. Interestingly, in a competition experiment, Danishefsky and coworkers were able to show that diastereomeric β-thiol Leu derivatives reacted at substantially different rates, with peptide **111** reacting approximately 20 times faster than the β-epimer *epi*-**111** (Scheme 27). This selectivity was thought to be a result of a *trans* relationship between the β-isopropyl group and the peptide chain imposed by the putative five-membered ring transition state in the $S \rightarrow N$ acyl shift in peptide **111**, while a corresponding *cis* relationship would predominate for *epi*-**111** (Scheme 27). The authors also postulate that the rate of initial transthioesterification for the two epimeric thiols with the C-terminal peptide thioester is altered by the ability of the proximal α-amino group to participate in a base catalysis step to generate the reactive thiolate at the β-position. This proposed intramolecular proton transfer step would be similarly affected by the *cis* and *trans* orientation of the β-isopropyl group and the peptide chain [109].

Brik and coworkers further demonstrate the synthetic capability of ligation-desulfurization at β-thiol Leu through the total synthesis of the HIV-1 Tat protein, which the authors had previously attempted using a side-chain auxiliary approach [61] (see Sect. 3.1.2). In this instance, HIV-1 Tat was disconnected at a β-thiol Leu residue and a native Cys residue for an iterative ligation strategy employing fragments **112**, **113**, and **114** prepared via Fmoc-SPPS (Scheme 28) [110]. Ligation between peptide **112**, containing an N-terminal β-thiol Leu residue, and C-terminal peptide thioester **113**, bearing an N-terminal Cys thiazolidine, was first carried out to generate an intermediate peptide ligation product, which was subsequently desulfurized to generate the native Leu residue in peptide **115**. It should be noted



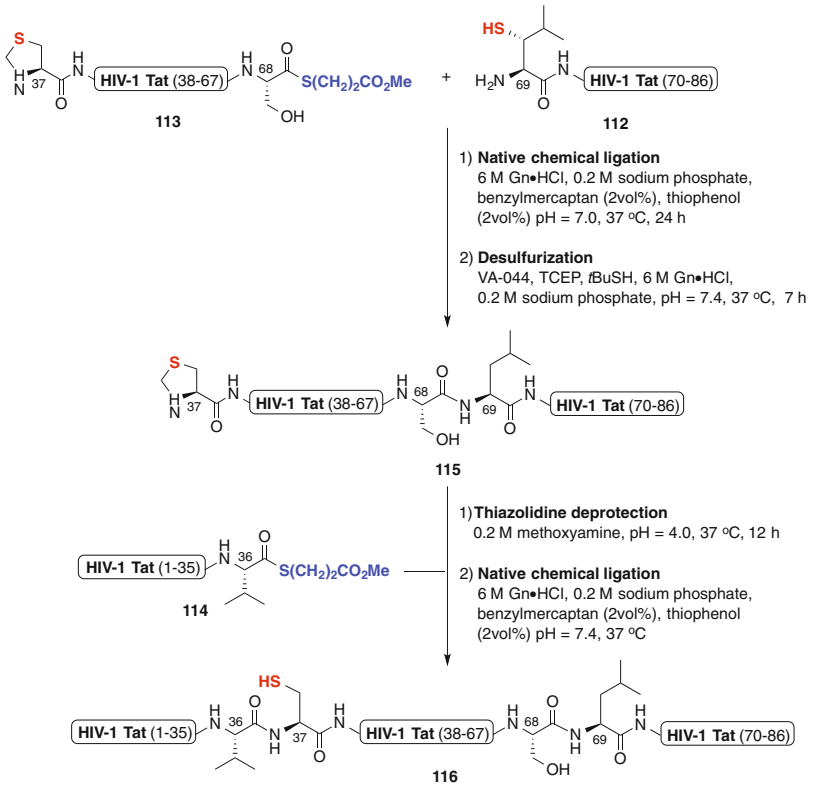
Scheme 27 Competition experiment between β-thiol Leu epimers and proposed origin of the observed selectivity

that a final, global desulfurization protocol was not employed to avoid the use of protecting groups on the N-terminal HIV-1 Tat fragment **114**, which is rich in Cys residues. Deprotection of intermediate thiazolidine **115** followed by ligation with thioester **114** afforded the full-length HIV-1 Tat protein **116** [110].

4.3.6 Proline

An approach to Pro ligation junctions employing the commercially available, protected γ-thiol Pro derivatives **93** and *epi*-**93** (Scheme 23) was offered by Danishefsky and coworkers in 2011 [111]. As with the diastereomeric Leu derivatives, a substantial rate difference was observed for the two γ-epimers, with only the *trans* derivative of **93** capable of facilitating ligation with C-terminal peptide thioesters. Once again, steric hindrance in the cyclic transition state of the $S \rightarrow N$ acyl shift was implicated as an explanation for the rate differential. Notably, in subsequent reports, Danishefsky and coworkers also demonstrated the utility of *trans* Pro derivative **93** in the ligation-based assembly of a large glycopeptide fragment of EPO [112, 113].

A synthetic approach to both diastereomers of the Boc-protected γ-thiol Pro derivative **94** was also reported by Otaka and coworkers [114]. The authors of this study confirmed the finding that use of the *trans* Pro derivative **94** was essential to facilitate ligation. The relative rate differential between rapid Cys-mediated ligations and ligation reactions mediated by *trans* γ-thiol Pro residues was also exploited to effect a one-pot, dual kinetically controlled ligation reaction.



Scheme 28 Total synthesis of the HIV-1 Tat protein using ligation-desulfurization at β-thiol Leu

This methodology strategically utilized an *N*-sulfanylethylanilide (SEAlide) peptide as a latent thioester moiety (see [115] for more details on SEA peptide technologies) [114].

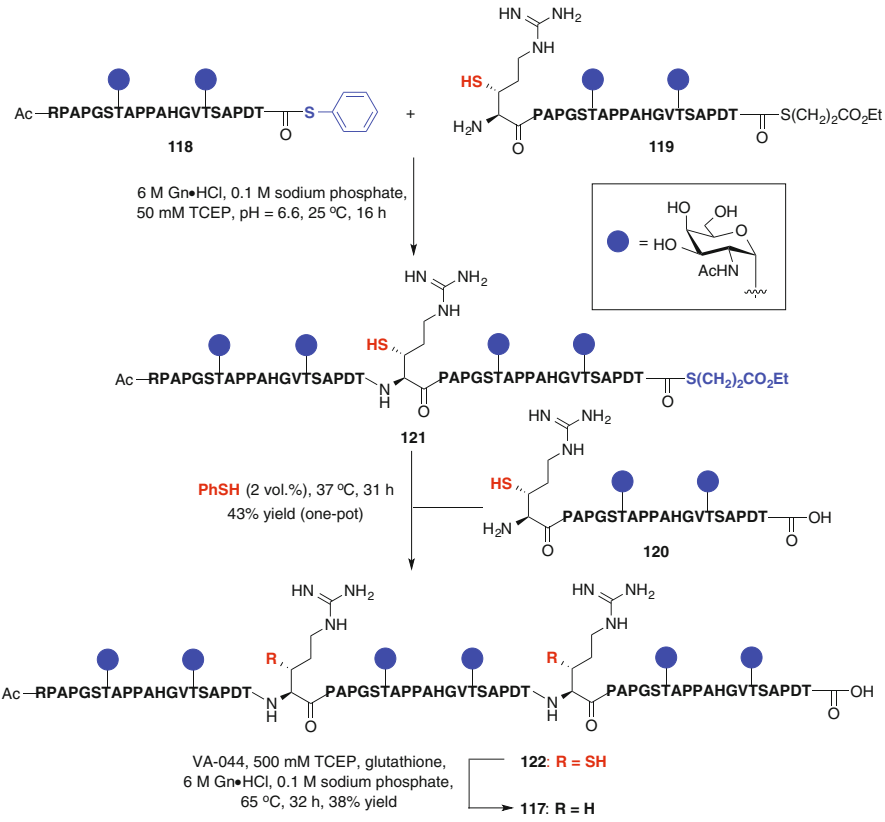
4.3.7 Glutamine

A ligation-desulfurization approach to Gln ligation junctions using a γ-thiol Gln derivative was reported by Brik and coworkers in 2012 [116]. The preparation of a diastereomeric mixture of the requisite γ-thiol Gln building block **95** was accomplished in ten steps from L-Asp. Ligation chemistry with a variety of model C-terminal peptide thioesters was unhindered by the diastereomeric γ-thiols. Interestingly, attempted radical desulfurization of γ-thiol Gln-containing peptides produced a complex mixture of products. As such, removal of the ligation auxiliary was accomplished via reductive desulfurization with nickel boride, affording enantiomerically pure peptide products [116].

4.3.8 Arginine

Payne and coworkers recently reported the synthesis of a β-thiol arginine (Arg) building block from commercially available Garner's aldehyde [117], a configurationally stable α-amino aldehyde, for use in ligation-desulfurization chemistry [118]. The Arg derivative **96** (Scheme 23) was shown to facilitate ligation with peptide thioesters bearing a range of functionality at the C-terminal position, and detailed kinetic studies indicated that ligation rates were inversely proportional to steric bulk at the C-terminal thioester residue, largely mirroring studies performed by Dawson and coworkers on native chemical ligation at Cys residues [13]. Interestingly, removal of the β-thiol auxiliary using radical desulfurization was substantially slower than the corresponding desulfurization of Cys to Ala. It was proposed that the Arg guanidine side-chain moiety was interfering with the standard radical desulfurization mechanism. Nonetheless, the Arg building block was successfully utilized in the ligation-based assembly of a 60-amino acid homogeneous MUC1 glycopeptide **117**, corresponding to three copies of the 20-residue MUC1 variable number tandem repeat (VNTR) sequence and bearing six *O*-linked glycans (Scheme 29).

Construction of the 60-amino acid MUC1 glycopeptide was achieved in a one-pot fashion from peptide fragments **118**, **119**, and **120** using a kinetically-controlled ligation strategy, first reported by Kent and coworkers for the construction of crambin [119]. The technique capitalizes on the innate reactivity difference between alkyl and aryl thioesters to perform iterative ligations in a rapid and efficient manner, while minimizing protecting group manipulations and intermediary purifications. This strategy has been successfully applied to the synthesis of a number of complex protein targets, including human lysozyme [120], HIV-1 protease [121], and full-length glycosylated EPO [87]. For the synthesis of MUC1 60-mer **117**, it was envisaged that kinetically-controlled, β-thiol Arg-mediated ligations between three functionalized MUC1 VNTR fragments would facilitate rapid construction of the target glycopeptides (Scheme 29). To this end, peptide thiophenyl thioester **118** was first ligated with bifunctional peptide **119** bearing an N-terminal β-thiol Arg and a C-terminal peptide alkyl thioester. The large increase in reactivity associated with aryl thioester **118** relative to the corresponding alkyl thioester **119** effectively promoted the intermolecular condensation of the two fragments to yield the 40-residue intermediate **121** rather than the competing cyclization or oligomerization of peptide **119** by reaction at the alkyl thioester. Importantly, this reaction took place in the absence of a thiol additive to avoid in situ activation of the alkyl thioester through transthioesterification. Without isolation, ligation of intermediate product **121** with peptide **120** containing an N-terminal β-thiol Arg residue was accomplished via addition of the final peptide fragment, along with 2 vol.% thiophenol, to generate the 60-amino acid polypeptide **122** in 43% isolated yield. Removal of the Arg β-thiol auxiliaries using a double desulfurization reaction then afforded the native glycopeptide **117** [118].



Scheme 29 Synthesis of a MUC1 glycopeptide oligomer using a kinetically-controlled ligation-desulfurization strategy at β-thiol Arg

4.3.9 Aspartic Acid

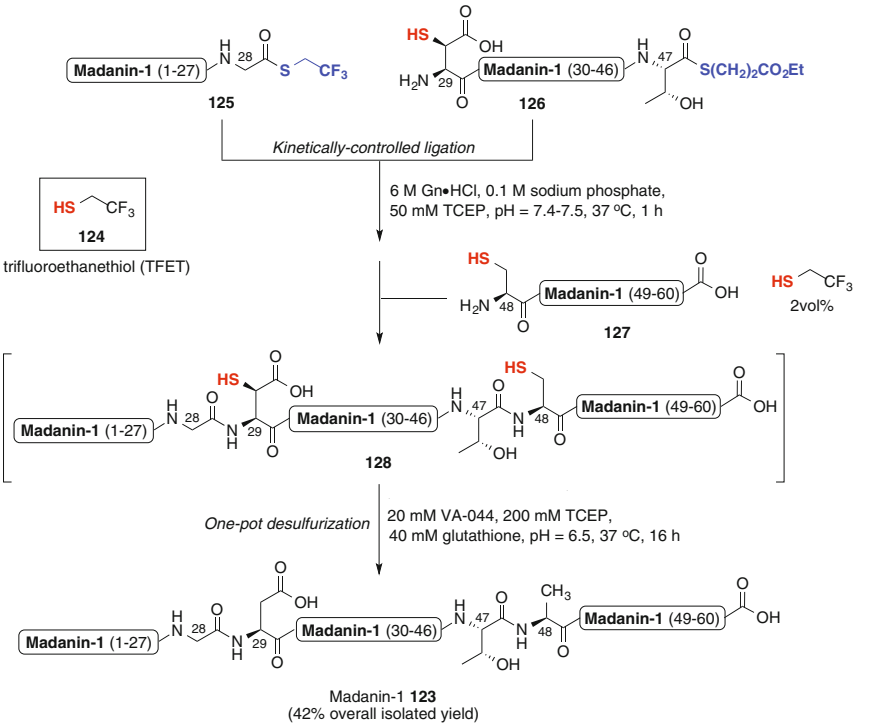
A concise, three-step synthesis of β-thiol Asp derivative **97** (Scheme 23) was developed in 2013 by Thompson et al. from commercially available Boc-Asp (O*t*Bu)-OH by employing a key sulfenylation reaction to install the requisite protected-thiol moiety [122]. Building block **97** was shown to facilitate the high-yielding synthesis of native peptides through ligation-desulfurization chemistry, and ligation reactions were found to proceed with equal efficiency, regardless of the configuration at the β-position. Interestingly, the authors of this report also demonstrated that β-thiol Asp residues could be selectively desulfurized in the presence of unprotected Cys residues upon treatment with TCEP and dithiothreitol (DTT) at 65 °C and pH 3, in the absence of a radical initiator. As standard reductive and radical-based desulfurization methods are unselective, application of these techniques requires the protection of all native Cys residues in the target sequence. In contrast, the ability to chemoselectively remove the β-thiol Asp ligation auxiliary

abrogates the need for protecting group manipulation in protein targets with functionally important Cys residues. The utility of the ligation-chemoselective desulfurization protocol was demonstrated through an efficient, one-pot synthesis of the N-terminal, extracellular domain of the chemokine receptor CXCR4. The target CXCR4(1–38) fragment contained a native Cys residue (although at an intractable Pro-Cys junction) and two post-translational modifications – an *N*-linked glycan and Tyr *O*-sulfation [122].

Tan and coworkers subsequently reported the 7-step synthesis of a modified β-thiol Asp derivative **98** and the application of this building block in the total synthesis of the 60-amino acid neuropeptide human galanin-like peptide (hGALP) using ligation-desulfurization chemistry [123]. Notably, the authors of this report utilize the β-epimer of the thiol Asp derivative previously employed for the ligation-based assembly of CXCR4(1–38) [122]. In contrast to the rate differential observed with epimeric β-thiol Leu derivatives **91** [109], these results confirm that configuration at the β-position does not have a large impact on the efficiency of ligation reactions at β-thiol Asp residues [122, 123].

Very recently, the application of the β-thiol Asp derivative **97** [122] has been reported in the one-pot, three-component total synthesis of madanin-1 **123** (Scheme 30), a small, Cys-free thrombin inhibitory protein derived from the hard tick *H. longicornis* [124]. In this study, the target 60-amino acid protein **123** was disconnected at Asp and Ala ligation junctions using ligation-desulfurization chemistry mediated by β-thiol Asp and Cys, respectively. Interestingly, for the construction of the target protein, the authors also reported the use of a novel thiol additive, trifluoroethanethiol (TFET) **124**, which facilitates the facile application of a one-pot ligation-desulfurization protocol [124]. Previous attempts to streamline the ligation-desulfurization strategy into a one-pot process have been hampered by the use of aryl thiol ligation additives which, because of their radical quenching ability [125], complicate the radical desulfurization process [126–128]. Given the importance of aryl thiol additives in modulating thioester reactivity and promoting rapid ligation reactions [15, 16], a number of approaches have aimed to facilitate the post-ligation removal of aryl thiols. The liquid–liquid extraction of thiophenol [128] and the development of bifunctional aryl thiol catalysts which can be captured using a resin-bound aldehyde following the ligation reaction [126] have recently been employed to facilitate one-pot ligation desulfurization reactions.

The alkyl thiol TFET was designed as an alternative additive in ligation reactions to circumvent the issues posed by aryl thiol additives in radical desulfurization reactions while maintaining efficiency as a ligation catalyst [124]. Because the pK_a of TFET (=7.30) is comparable to the pK_a of common aryl thiol ligation additives, it was postulated that this alkyl thiol would be sufficiently nucleophilic to promote initial transthioesterification with the unactivated C-terminal peptide thioester and would also maintain sufficient leaving group ability to promote the acylation of the Cys thiol moiety. The efficacy of the additive was showcased in a kinetically-controlled one-pot total synthesis of madanin-1 **123** (Scheme 30). To this end, preformed TFET thioester **125**, corresponding to residues 1–28 of madanin-1 was ligated with bifunctional peptide madanin-1 (29–47) **126**, bearing an N-terminal



Scheme 30 One-pot total synthesis of madanin-1 **123** using kinetically-controlled ligation-desulfurization chemistry with trifluoroethanethiol (TFET) **124**

β-thiol Asp and an unactivated C-terminal alkyl thioester. Without isolation, madanin-1 (48–60) **127**, containing an N-terminal Cys residue, was added to the ligation mixture, along with 2 vol.% TFET to activate the latent C-terminal alkyl thioester and promote a second ligation reaction. The 60-amino acid ligation product **128** was then subjected directly to radical desulfurization, cleanly affording the target protein **123** in 42% isolated yield over the three steps. The one-pot total synthesis of madanin-1 **123** therefore served as a proof-of-concept for the efficiency of β-thiol Asp residues in ligation-desulfurization chemistry, as well as application of the new thiol additive, TFET, in the one-pot construction of Cys-free proteins [124].

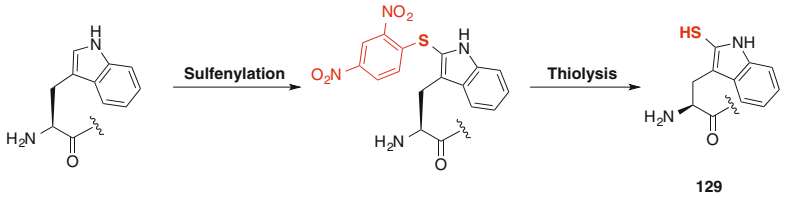
4.3.10 Glutamic Acid

Using a similar synthetic approach to the one employed for the synthesis of β-thiol Asp building block **97**, Cergol et al. reported the preparation of γ-thiol Glu derivatives **99** and **100** (Scheme 23) and their application in ligation-desulfurization chemistry [128]. Initial attempts to incorporate building block **99** into model

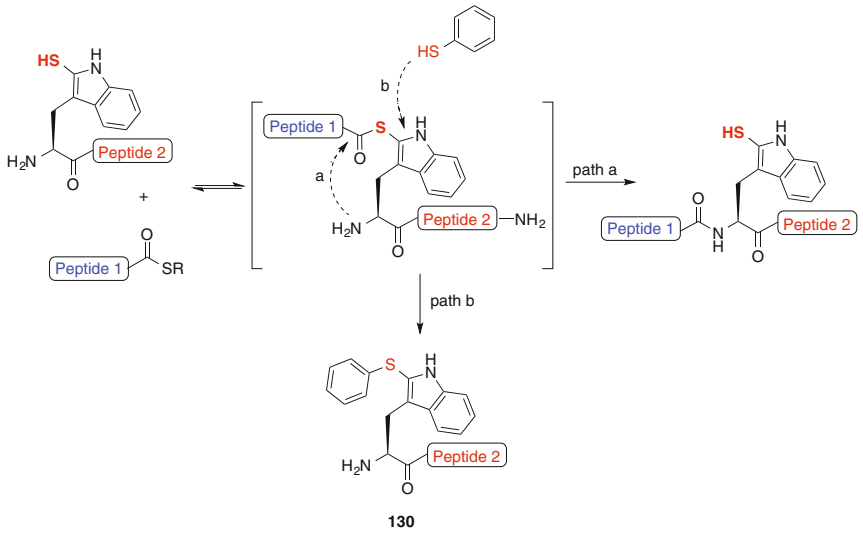
peptides using Fmoc-SPPS were complicated by the instability of the γ-thiol Glu derivative in the acidic cleavage cocktail. It was determined that, at acidic pH, the deprotected γ-thiol could facilitate nucleophilic cleavage of the amide backbone, resulting in loss of the terminal thiol-derived amino acid. Subsequent incorporation of the asymmetric disulfide building block **100** circumvented this issue, allowing for the efficient construction of model peptides. Ligation reactions proceeded in high isolated yields and could be followed by desulfurization using either a one- or two-pot protocol. To facilitate one-pot desulfurization, ligation reactions were initially performed using thiophenol as an additive, which could be removed almost entirely from the reaction mixture by extraction with diethyl ether [128]. Subsequently, the application of γ-thiol Glu derivative **100** in a TFET-mediated iterative ligation-desulfurization strategy enabled the efficient one-pot total chemical synthesis of chimadanin-1, another small thrombin inhibitory protein isolated from the hard tick *H. longicornis* [124].

4.3.11 Tryptophan

An innovative approach to ligation-desulfurization at N-terminal Trp residues has recently been reported [129]. Based on results from Scoffone et al. in the late 1960s [130, 131], the method utilizes a chemoselective sulfenylation protocol, followed by a mild and selective thiolysis reaction [132], to install a thiol ligation auxiliary at the 2-position of the Trp indole ring (e.g., **129**, Scheme 31) in unprotected peptides or resin-bound peptide intermediates. This methodology therefore eliminated the need to synthesize a protected thiol-derived Trp amino acid building block. Ligation reactions at model peptides bearing the 2-thiol Trp moiety **129** were found to proceed in good yields with preformed thiophenyl thioesters in the absence of an exogenous thiol additive. Interestingly, in the presence of exogenous aryl thiol, significant quantities of the 2-thioether byproduct **130** were observed (Scheme 32). Mechanistically, following the initial transthioesterification step, the positioning of the thiol auxiliary at the 2-position of the Trp indole ring required the $S \rightarrow N$ acyl shift to proceed through a seven-membered ring intermediate (path a, Scheme 32). It was postulated that, in the presence of excess thiophenol, a slower $S \rightarrow N$ acyl migration step allowed the intermediate bridged thioester to be intercepted at the C-2 position of the indole ring with exogenous aryl thiol (path b) before the intramolecular acyl shift could occur (Scheme 32). Following optimization of the ligation conditions, reductive desulfurization using Pd on Al_2O_3 in the presence of H_2 gas cleanly afforded native peptide products. The technique was also applied to the synthesis of a glycopeptide fragment of the N-terminal extracellular domain of the chemokine receptor CXCR1. In this example, installation of the key thiol auxiliary was performed using an efficient solid-phase sulfenylation protocol [129].



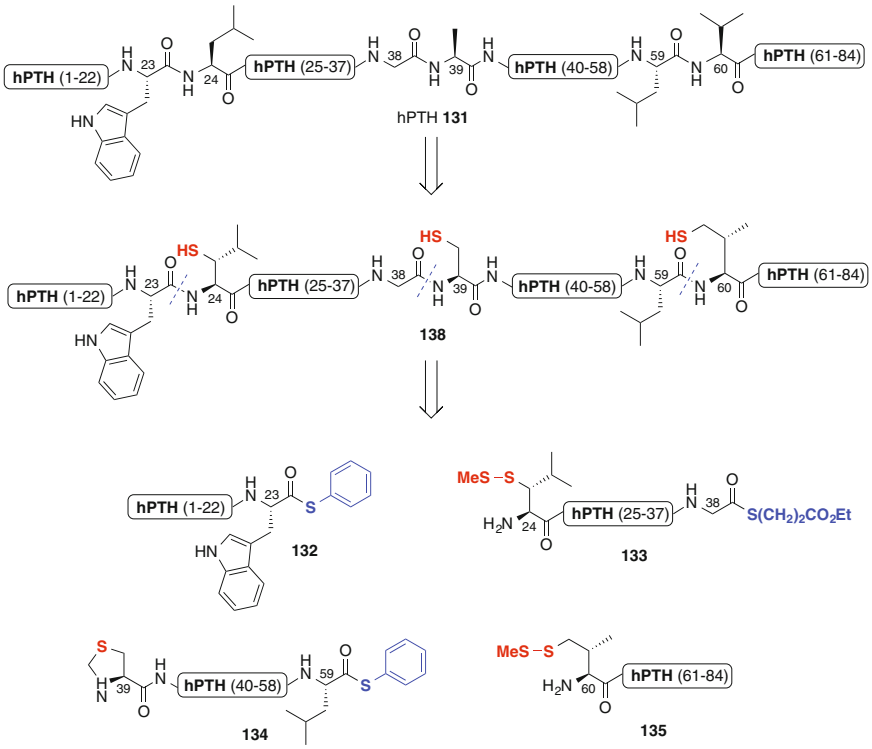
Scheme 31 Synthesis of 2-thiol Trp derivatives (e.g., **129**) through the sulfenylation of Trp



Scheme 32 Transthioesterification of 2-thiol Trp and subsequent $S \rightarrow N$ acyl shift (path a) or interception with exogenous thiophenol (path b)

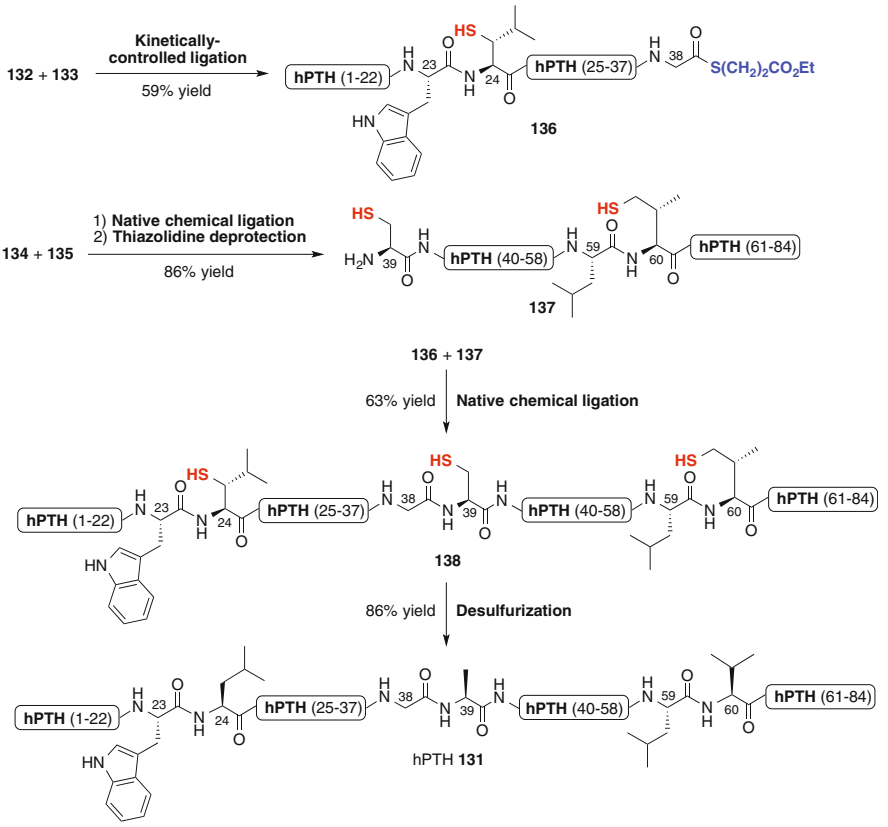
4.3.12 Protein Synthesis via Ligation at Non-Cys Sites

The flourish of activity in the development of building blocks for ligation-desulfurization chemistry has provided enormous flexibility for the disconnection of target proteins. As discussed, these building blocks are used in combination with ligation at Cys or ligation-desulfurization at Ala for the construction of complex proteins. A powerful example of the scope of ligation using a combination of these building blocks was provided by Danishefsky and coworkers in 2011 with the construction of human parathyroid hormone (hPTH) **131** using a convergent ligation-global desulfurization strategy employing Cys, β-thiol Leu, and γ-thiol Val ligation disconnections (Scheme 33) [133]. It was envisaged that the target hPTH protein could be synthesized from peptide fragments **132–135** by combining iterative ligation technologies for the construction of proteins in both the *N*-to-*C* and *C*-to-*N* directions (Scheme 34).



Scheme 33 Retrosynthetic analysis for the total chemical synthesis of hPTH **131**

First, thiophenyl thioester **132** and bifunctional peptide **133**, bearing an N-terminal β-thiol Leu residue and a C-terminal alkyl thioester were ligated in a kinetically-controlled ligation reaction, affording hPTH(1–38) **136** as the alkyl thioester derivative in 59% yield (Scheme 34). Peptide fragments **134** and **135** were then joined using a γ-thiol Val-mediated ligation reaction. Subsequent thiazolidine deprotection afforded hPTH(39–84) **137** bearing a free N-terminal Cys residue for further functionalization. Accordingly, standard native chemical ligation of peptide **137** with thioester **136** facilitated construction of the full-length polypeptide sequence **138** in 63% yield. Radical desulfurization of the three thiol ligation auxiliaries finally afforded the native protein hPTH **131** in 86% yield following HPLC purification. Notably, the hPTH prepared in this study using total chemical synthesis was shown to be significantly more pure than the comparative recombinant hPTH reference sample used to confirm the identity of the protein. The superior quality of the product obtained using chemical methods attests to the importance of ligation chemistry, including non-Cys ligation strategies, for the production of homogeneous proteins for use in biological studies. It is therefore envisaged that the ligation-desulfurization motif will continue to be widely utilized in the total chemical synthesis of target proteins, particularly as thiol-derived amino

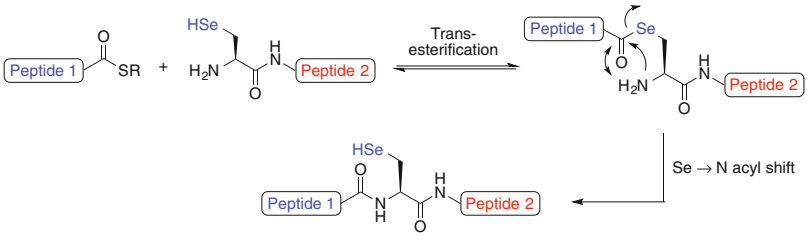


Scheme 34 The synthesis of hPTH **131** using a convergent, iterative ligation strategy followed by global desulfurization

acid building blocks become increasingly accessible, e.g., through commercial sources.

4.4 Ligation at Selenocysteine

Selenocysteine (Sec) is often considered to be the 21st amino acid [134], and there are a number of naturally occurring proteins that contain structurally and functionally crucial Sec residues [135]. These considerations, together with the structural similarities between Cys and Sec, render native chemical ligation at N-terminal Sec residues a logical extension of chemoselective ligation technologies. Indeed, in 2001, the laboratories of van der Donk, Hilvert, and Raines independently reported native chemical ligation at Sec [136–138]. Mechanistically, Sec-mediated ligations were proposed to proceed in an analogous fashion to ligation at Cys (Scheme 35),



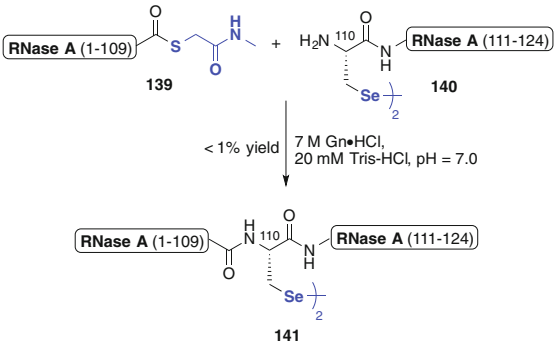
Scheme 35 Native chemical ligation at selenocysteine (Sec)

whereby an initial, reversible transesterification reaction between a C-terminal peptide thioester and an N-terminal Sec residue first proceeds to form a bridged selenoester intermediate. Intramolecular rearrangement of the selenoester in a *Se* → *N* acyl shift through a five-membered ring intermediate subsequently generates a new amide bond.

Despite similar mechanistic considerations, Sec displays a number of distinct physicochemical properties from its sulfur counterpart which have important implications in ligation chemistry. In particular, the pK_a of the Sec selenol functionality is approximately 5.24–5.63 [139, 140], implying that at physiological pH, Sec exists primarily in the anionic selenolate form. The corresponding Cys thiol, however, has a pK_a of 8.25 [139], and remains largely protonated at neutral pH. Selenolates also display enhanced nucleophilicity [139] and leaving group abilities relative to their thiolate analogues [141]. Collectively, these unique chemical properties suggest that ligation at Sec residues might proceed faster than ligation at the corresponding Cys residues, particularly at slightly acidic pH. Raines and coworkers indeed demonstrated that ligation reactions with Sec in the presence of ligation buffer and TCEP proceeded 10^3-fold faster than the analogous Cys ligation at pH 5.0, leading to the proposition that Sec ligations may be performed chemoselectively in the presence of Cys residues [138].

However, in the absence of TCEP, van der Donk and coworkers observed sluggish reactivity in ligation reactions facilitated by N-terminal Sec residues. The authors speculated that the decreased rate of reaction was a consequence of a low steady-state concentration of reactive selenol in the ligation mixture, with the starting Sec peptide existing primarily as the symmetrical diselenide or the selenylsulfide variant under the ligation conditions employed [136]. The unique redox properties of diselenides, particularly their relative stability and large negative reduction potential [142, 143], support this hypothesis. Furthermore, the observation by Hilvert and coworkers that Sec-mediated ligation reactions did not proceed in the absence of TCEP and exogenous thiophenol also suggests that the overall rate of reaction at Sec is heavily dependent on the ability to reduce diselenides in the ligation mixture [137]. The observation that the inclusion of TCEP in Sec-mediated ligations resulted in the formation of deselenization byproducts [136], however, has fuelled the adoption of thiol additives, such as MPAA, as mild reductants in Sec ligation chemistry [144].

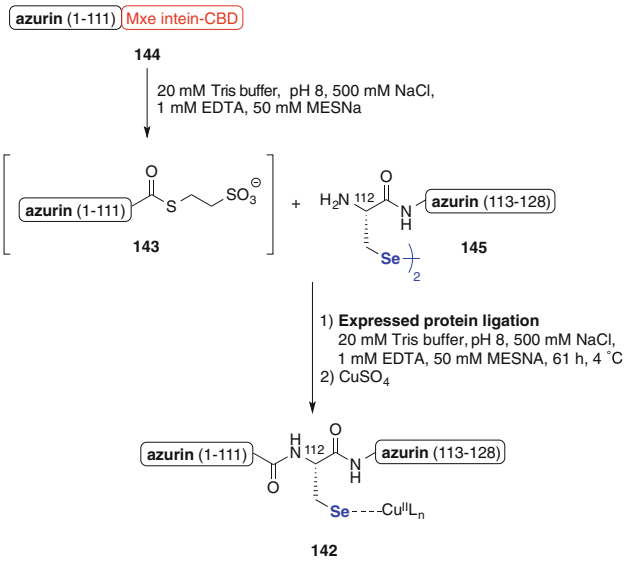
Scheme 36 Synthesis of RNase A using a Sec-mediated expressed protein ligation reaction



There have been a number of applications of Sec-mediated ligation chemistry for the synthesis of diverse peptide and protein targets, including a Sec analog of the C-terminal fragment of ribonucleotide reductase [136] and a fully folded Cys38Sec mutant of the 58-amino acid bovine pancreatic trypsin inhibitor (BPTI) [137]. The ligation technology has also been used in conjunction with expressed protein ligation to construct a Sec analog of ribonuclease A (RNase A) [138]. To this end, a 109-amino acid recombinant thioester **139** was produced in *E. coli* and subsequently ligated to the short, synthetic selenopeptide **140**, prepared using SPPS, to provide the target semi-synthetic protein **141** (Scheme 36). The Cys110Sec RNase A analogue **141** was isolated in very low yields because of the poor recovery of recombinant thioester, but nonetheless displayed ribonucleolytic activity consistent with the wild-type enzyme, suggesting that the Sec analog was properly folded [138].

A novel application of Sec ligation chemistry in protein engineering has been the preparation of a semi-synthetic analogue of the type 1 blue copper protein, azurin, containing a Cys112Sec mutation [145]. Cys^{112} is involved in the active site of the wild-type redox metalloprotein through coordination to the copper ligand. It was therefore envisaged that synthesis of a Sec^{112} variant might provide important insight into the structure and function of the protein. Synthesis of the target protein **142** was accomplished using expressed protein ligation of an N-terminal recombinant thioester **143** corresponding to residues 1–111 of azurin (Scheme 37). The recombinant thioester was generated in situ from the corresponding fusion protein **144** upon treatment with MESNa. Ligation with selenopeptide **145**, bearing an N-terminal Sec residue afforded the full-length Cys112Sec azurin. Addition of copper sulfate then produced the ligand-bound protein **142** in a yield of ~0.4 mg/L of culture. Synthesis of the engineered protein enabled a detailed comparison of the electronic absorption spectra and reduction potential of the engineered variant with wild-type azurin [145].

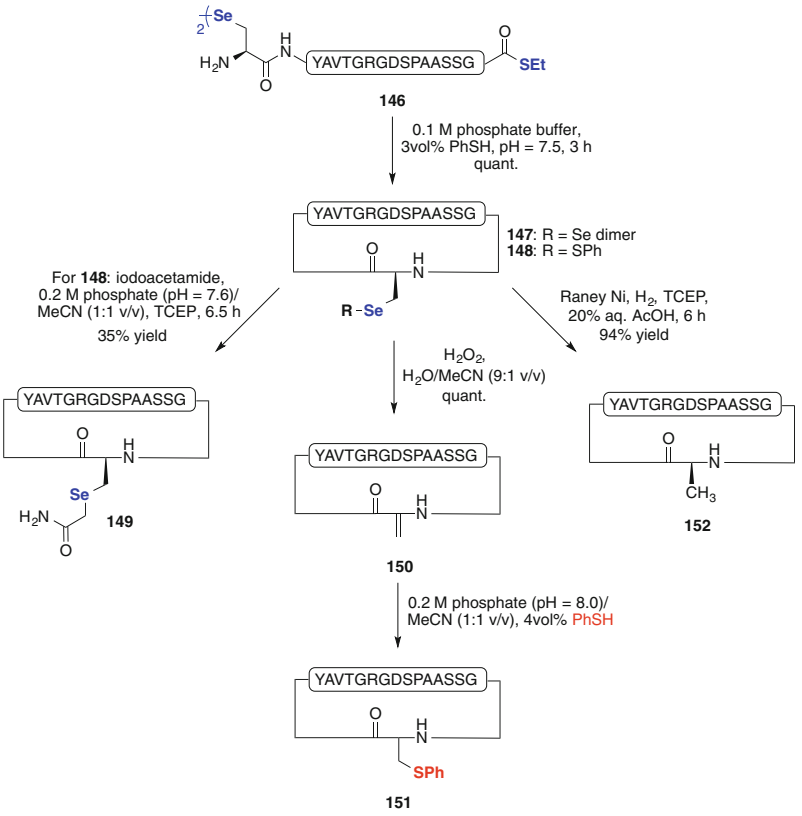
As with ligation at Cys, there has been considerable interest in the post-ligation manipulation of Sec residues to facilitate access to diverse ligation junctions. In 2002, Hilvert and coworkers reported the synthesis of a cyclic peptide from



Scheme 37 Synthesis of Cys112Sec azurin using expressed protein ligation

bifunctional precursor **146** using a Sec ligation strategy (Scheme 38) [146]. Following construction of the cyclic selenopeptide (isolated as the symmetrical diselenide **147** and the seleno-thiophenyl sulfide **148**), the Sec residue was used as a handle for chemical manipulation. Alkylation of selenylsulfide **148** with iodoacetamide in the presence of TCEP generated the selenoether derivative **149** in 35% yield. Oxidative elimination with hydrogen peroxide afforded the dehydroalanine (Dha) derivative **150**, which could be further functionalized in a thiol Michael reaction to generate thioethers such as **151**. Finally, reductive deselenization of **146** in the presence of Raney Ni, akin to the post-ligation desulfurization of Cys residues [54], afforded peptide **152**, bearing the corresponding Ala residue at the ligation junction [146]. Interestingly, in their original report of metal-free radical desulfurization, Danishefsky and coworkers also reported the extension of the radical protocol to the deselenization of Sec to Ala [72].

Despite the potential for diverse post-ligation modifications at Sec residues, Sec-mediated ligation chemistry has not been widely adopted for the routine construction of target peptides and proteins which do not contain Sec residues in the final product [147]. The lack of commercially available Sec building blocks for direct incorporation into peptides using standard SPPS and the ability to modify readily the side-chain of Cys residues [54, 66, 148] have generally favored the use of standard native chemical ligation at Cys.

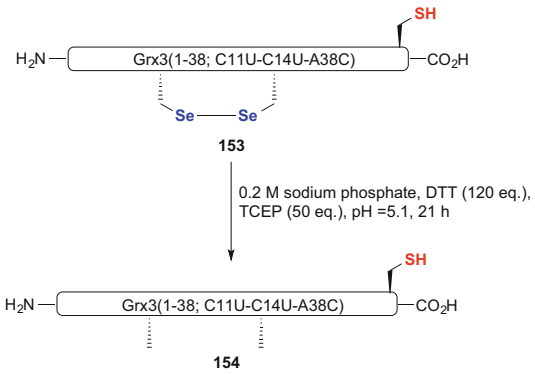


Scheme 38 Sec-mediated backbone cyclization followed by side-chain functionalizations

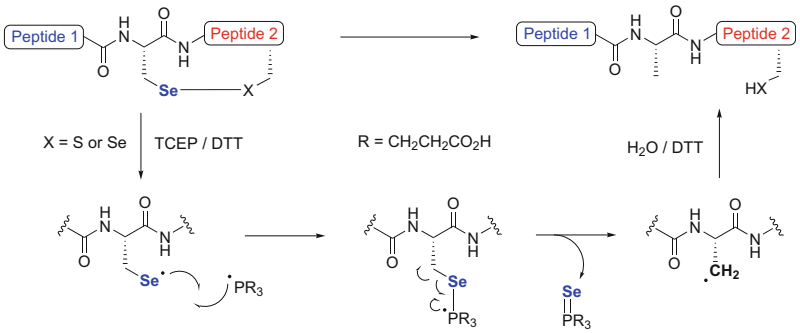
4.5 Ligation-Deselenization Chemistry

A 2010 report by Dawson and coworkers outlining the mild and selective deselenization of Sec in the presence of unprotected Cys residues [144] has contributed to a resurgence of interest in Sec-mediated ligation chemistry as a general strategy for the synthesis of peptides and proteins [147]. At the time of this seminal report, all known protocols for the reductive or radical-based desulfurization of Cys (see above) effected global cleavage of unprotected thiols within the target sequence. As such, the synthesis of peptides and proteins bearing native, non-ligation site Cys residues using ligation-desulfurization chemistry demanded the use of side-chain Cysprotecting groups. By providing the first chemoselective approach to Ala ligation junctions, Dawson and coworkers provided an important tool for the construction of proteins from fully deprotected peptide precursors [144].

The deselenization protocol involved the treatment of Sec-containing peptides, at room temperature and in aqueous media, with excess TCEP in the presence of dithiothreitol (DTT). The utility of the reaction was demonstrated on a model

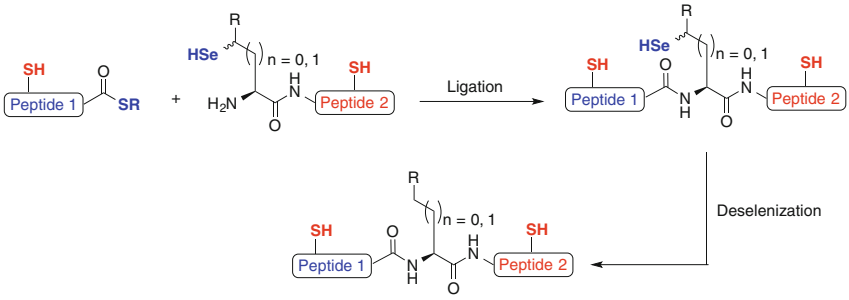


Scheme 39 Chemoselective deselenization of Sec to Ala in the presence of an unprotected Cys residue

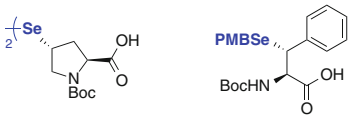


Scheme 40 Proposed mechanism for the phosphine-mediated deselenization of Sec

system and in a larger polypeptide corresponding to Grx(1–38) **153** containing two Cys to Sec mutations at positions 11 and 14, and a single Ala to Cys mutation at residue 38 (Scheme 39). Under the optimized deselenization conditions, the two Sec to Ala conversions were accomplished to afford compound **154** with only minor amounts of the globally reduced Grx (1–38) product observed (Scheme 39). Notably, a Grx(1–38) Cys mutant, containing no Sec residues, was completely stable to the deselenization conditions [144]. Mechanistically, the phosphine-mediated deselenization reaction was thought to proceed via a radical pathway (Scheme 40), similar to the one proposed for the radical desulfurization of thiols by trialkylphosphites and trialkylphosphines [74, 75] and implicated by Danishefsky and coworkers in their development of the metal-free radical desulfurization protocol [72]. Although proceeding through similar pathways, the observed selectivity of the deselenization reaction for Sec over Cys might be attributed to the preferential formation of selenium-centered radicals over the corresponding sulfur-centered radicals, particularly under the mild conditions employed for deselenization.



Scheme 41 Chemoselective ligation-deselenization at β- ($n = 0$) and γ- ($n = 1$) selenol amino acids



155 156

Scheme 42 γ-Selenoproline **155** and β-selenophenylalanine **156** building blocks

By allowing for the chemoselective conversion of Sec to Ala in the presence of unprotected Cys residues, the ligation-deselenization strategy effectively provided a means of accomplishing, without protecting group manipulations, ligation at Ala and Cys junctions in the same protein target. In a similar manner to the extension of ligation-desulfurization chemistry to include thiol-derived proteinogenic amino acid building blocks (see Scheme 23) [54], Dawson and coworkers also proposed that the logic of chemoselective ligation-deselenization chemistry could be extended to include synthetic selenol-derived amino acids [144]. Specifically, non-proteinogenic building blocks bearing a suitably positioned β- or γ-selenol auxiliary could facilitate a Sec-mediated ligation reaction to afford an unnatural selenopeptide product. The mild and chemoselective removal of the selenol auxiliary in the presence of TCEP could then be effected in the presence of unprotected Cys residues elsewhere in the target sequence (Scheme 41), thereby expanding the scope and flexibility of chemoselective ligation chemistry.

Chemoselective ligation-deselenization chemistry has recently been extended through the preparation of a γ-selenoproline building block **155** [113] and a β-selenophenylalanine derivative **156** [149] (Scheme 42). The *trans*-γ-selenoproline derivative **155** was prepared by Danishefsky and coworkers in three steps from a protected hydroxyproline precursor. The *trans*-derivative was chosen over the corresponding γ-epimer on the basis of favorable results obtained from earlier studies with the analogous *trans*-γ-thiol derivatives **93** and **94** [111, 113, 114]. Following incorporation into model peptides, building block **155** was shown to mediate ligation effectively with a variety of C-terminal peptide thioesters, including Gly, Ala, Phe, and Val. Notably, ligation reactions were performed in

the absence of TCEP to avoid premature deselenization of the selenol ligation auxiliary. The authors instead used MPAA [16] as both an exogenous thiol exchange catalyst and a mild reductant for the generation of free selenol from the starting diselenide [144]. Deselenization reactions were performed in a one-pot fashion through the sequential addition of DTT and TCEP. Importantly, the chemoselectivity of the deselenization protocol was also confirmed in the presence of an unprotected γ-thiol Pro residue [113].

The preparation of β-selenophenylalanine derivative **156** and its application in ligation-deselenization chemistry was subsequently reported [149]. The synthesis of the key protected amino acid building block was accomplished in seven steps from Garner's aldehyde [117]. Notably, Garner's aldehyde has also served as a synthetic precursor for the preparation of β-thiol Arg derivative **96** by Payne and coworkers [118] and has been proposed as a common starting point for the divergent synthesis of both β-thiol and β-selenol amino acid derivatives [149]. Following the synthesis of **156** and its incorporation into model peptides using standard Fmoc-SPPS, a number of ligation reactions were performed to evaluate the utility of the building block in ligation-deselenization chemistry. Reactions proceeded in moderate to good yields for the majority of C-terminal peptide thioesters examined, requiring 24–48 h to reach completion. The slow rates of reaction were attributed to the relative stability of the starting diselenide and the reliance on MPAA, rather than the more powerful reductant TCEP, to liberate free selenol to promote the ligation reaction. Interestingly, deselenization of the purified ligation materials with TCEP and DTT led to substantial formation of peptide derivatives bearing diastereomeric β-hydroxy Phe at the ligation junction. The prevalence of this byproduct was dramatically reduced when ligation-deselenization reactions were performed as a one-pot protocol in the presence of exogenous MPAA. Under these conditions, the rate of deselenization also dramatically decreased, perhaps because of the ability of aryl thiols such as MPAA to act as competitive radical scavengers [126, 127, 144]. Nonetheless, native peptide products, including those bearing unprotected, non-ligation site Cys residues, were isolated in good yields following the one-pot protocol for ligation-deselenization at β-selenophenylalanine [149].

Although the application of ligation-deselenization chemistry has thus far been limited to model systems and small polypeptides, it is envisaged that this technology will also be amenable to the construction of proteins. By minimizing the need for late-stage protecting group manipulations, chemoselective ligation-deselenization chemistry will enhance the availability of ligation junctions and serve to expedite the construction of complex targets [147].

5 Conclusion

Over the last 20 years, native chemical ligation has ushered in a new era in the total chemical synthesis of proteins by enabling the efficient, programmed construction of native structures, including those bearing post-translational modifications, as

well as the modular synthesis of strategically engineered protein variants. This chapter has summarized the importance of native chemical ligation and highlighted a number of modern extensions to the original technology which have facilitated these synthetic feats. In particular, novel methods for Cys-free ligation, including auxiliary-mediated ligation and ligation-desulfurization chemistry, employing both Cys residues and synthetic thiol-derived amino acid variants, have been extensively explored. These techniques have served to increase dramatically the availability of synthetically viable ligation junctions and have expanded the flexibility of modern ligation chemistry for the construction of diverse targets. In addition, the advent of chemoselective ligation-deselenization chemistry has provided a promising new strategy for the manipulation of proteins in the absence of protecting groups. It is predicted that these powerful new tools, and additional developments in chemoselective ligation technologies, will continue to fuel the construction of increasingly more complex protein targets in the years to come.

References

1. Walsh CT, Garneau-Tsodikova S, Gatto GJ Jr (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. Angew Chem Int Ed 44:7342–7372
2. Walsh C (2006) Posttranslational modification of proteins: expanding nature's inventory. Roberts and Co Publishers, Englewood, Colo
3. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH (1994) Synthesis of proteins by native chemical ligation. Science 266:776–779
4. Wieland T, Bokelmann E, Bauer L, Lang HU, Lau H (1953) Über peptidsynthesen. 8. Mitteilung bildung von S-haltigen peptiden durch intramolekulare wanderung von aminoacylresten. Justus Liebigs Ann Chem 583:129–149
5. Merrifield RB (1963) Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. J Am Chem Soc 85:2149–2154
6. Kemp DS, Leung S-L, Kerkman DJ (1981) Models that demonstrate peptide bond formation by prior thiol capture. I. Capture by disulfide formation. Tetrahedron Lett 22:181–184
7. Kemp DS, Kerkman DJ (1981) Models that demonstrate peptide bond formation by prior thiol capture. II. Capture by organomercury derivatives. Tetrahedron Lett 22:185–186
8. Fotouhi N, Galakatos NG, Kemp DS (1989) Peptide synthesis by prior thiol capture. 6. Rates of the disulfide-bond-forming capture reaction and demonstration of the overall strategy by synthesis of the C-terminal 29-peptide sequence of BPTI. J Org Chem 54:2803–2817
9. Liu C-F, Tam JP (1994) Chemical ligation approach to form a peptide bond between unprotected peptide segments. Concept and model study. J Am Chem Soc 116:4149–4153
10. Liu C-F, Tam JP (1994) Peptide segment ligation strategy without use of protecting groups. Proc Natl Acad Sci U S A 91:6584–6588
11. Schnoelzer M, Kent SBH (1992) Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. Science 256:221–225
12. Schnoelzer M, Alewood P, Jones A, Alewood D, Kent SBH (1992) In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. Int J Pept Protein Res 40:180–193
13. Hackeng TM, Griffin JH, Dawson PE (1999) Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. Proc Natl Acad Sci U S A 96:10068–10073

14. Camarero J, Adeva A, Muir T (2000) 3-Thiopropionic acid as a highly versatile multidetachable thioester resin linker. Lett Pept Sci 7:17–21
15. Dawson PE, Churchill MJ, Ghadiri MR, Kent SBH (1997) Modulation of reactivity in native chemical ligation through the use of thiol additives. J Am Chem Soc 119:4325–4329
16. Johnson ECB, Kent SBH (2006) Insights into the mechanism and catalysis of the native chemical ligation reaction. J Am Chem Soc 128:6640–6646
17. Evans TC Jr, Benner J, Xu MQ (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. Protein Sci 7:2256–2264
18. Muir TW (2003) Semisynthesis of proteins by expressed protein ligation. Annu Rev Biochem 72:249–289
19. Pollock SB, Kent SB (2011) An investigation into the origin of the dramatically reduced reactivity of peptide-prolyl-thioesters in native chemical ligation. Chem Commun 47:2342–2344
20. Kent SB (2009) Total chemical synthesis of proteins. Chem Soc Rev 38:338–351
21. Payne RJ, Wong CH (2010) Advances in chemical ligation strategies for the synthesis of glycopeptides and glycoproteins. Chem Commun 46:21–43
22. Raibaut L, Ollivier N, Melnyk O (2012) Sequential native peptide ligation strategies for total chemical protein synthesis. Chem Soc Rev 41:7001–7015
23. Macmillan D (2006) Evolving strategies for protein synthesis converge on native chemical ligation. Angew Chem Int Ed 45:7668–7672
24. Hackenberger CPR, Schwarzer D (2008) Chemoselective ligation and modification strategies for peptides and proteins. Angew Chem Int Ed 47:10030–10074
25. Unverzagt C, Kajihara Y (2013) Chemical assembly of N-glycoproteins: a refined toolbox to address a ubiquitous posttranslational modification. Chem Soc Rev 42:4408–4420
26. Haase C, Seitz O (2008) Extending the scope of native chemical peptide coupling. Angew Chem Int Ed 47:1553–1556
27. Dirksen A, Dawson PE (2008) Expanding the scope of chemoselective peptide ligations in chemical biology. Curr Opin Chem Biol 12:760–766
28. Gamblin DP, Scanlan EM, Davis BG (2009) Glycoprotein synthesis: an update. Chem Rev 109:131–163
29. Okamoto R, Mandal K, Ling M, Luster AD, Kajihara Y, Kent SBH (2014) Total chemical synthesis and biological activities of glycosylated and non-glycosylated forms of the chemokines CCL1 and Ser-CCL1. Angew Chem Int Ed 53:5188–5193
30. Bang D, Kent SB (2004) A one-pot total synthesis of crambin. Angew Chem Int Ed 43:2534–2538
31. Blanco-Canosa JB, Dawson PE (2008) An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. Angew Chem Int Ed 47:6851–6855
32. Okamoto R, Mandal K, Sawaya MR, Kajihara Y, Yeates TO, Kent SB (2014) (Quasi-) racemic X-ray structures of glycosylated and non-glycosylated forms of the chemokine Ser-CCL1 prepared by total chemical synthesis. Angew Chem Int Ed 53:5194–5198
33. UniprotKB/TrEMBL Protein Database Release 2014_07 Statistics (2014) http://www.ebi.ac.uk/uniprot/TrEMBLstats. Accessed 1 Sept 2014
34. Canne LE, Bark SJ, Kent SBH (1996) Extending the applicability of native chemical ligation. J Am Chem Soc 118:5891–5896
35. Marinzi C, Bark SJ, Offer J, Dawson PE (2001) A new scaffold for amide ligation. Bioorg Med Chem 9:2323–2328
36. Botti P, Carrasco MR, Kent SBH (2001) Native chemical ligation using removable Nα-(1-phenyl-2-mercaptoethyl) auxiliaries. Tetrahedron Lett 42:1831–1833
37. Low DW, Hill MG, Carrasco MR, Kent SB, Botti P (2001) Total synthesis of cytochrome b562 by native chemical ligation using a removable auxiliary. Proc Natl Acad Sci U S A 98:6554–6559

38. Kawakami T, Aimoto S (2003) A photoremovable ligation auxiliary for use in polypeptide synthesis. Tetrahedron Lett 44:6059–6061
39. Marinzi C, Offer J, Longhi R, Dawson PE (2004) An o-nitrobenzyl scaffold for peptide ligation: synthesis and applications. Bioorg Med Chem 12:2749–2757
40. Clive DL, Hisaindee S, Coltart DM (2003) Derivatized amino acids relevant to native peptide synthesis by chemical ligation and acyl transfer. J Org Chem 68:9247–9254
41. Muir TW, Sondhi D, Cole PA (1998) Expressed protein ligation: a general method for protein engineering. Proc Natl Acad Sci U S A 95:6705–6710
42. Chatterjee C, McGinty RK, Pellois JP, Muir TW (2007) Auxiliary-mediated site-specific peptide ubiquitylation. Angew Chem Int Ed 46:2814–2818
43. McGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW (2008) Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. Nature 453:812–816
44. Offer J, Dawson PE (2000) N-alpha-2-Mercaptobenzylamine-assisted chemical ligation. Org Lett 2:23–26
45. Kawakami T, Akaji K, Aimoto S (2001) Peptide bond formation mediated by 4,5-dimethoxy-2-mercaptobenzylamine after periodate oxidation of the N-terminal serine residue. Org Lett 3:1403–1405
46. Vizzavona J, Dick F, Vorherr T (2002) Synthesis and application of an auxiliary group for chemical ligation at the X-gly site. Bioorg Med Chem Lett 12:1963–1965
47. Offer J, Boddy CNC, Dawson PE (2002) Extending synthetic access to proteins with a removable acyl transfer auxiliary. J Am Chem Soc 124:4642–4646
48. Macmillan D, Anderson DW (2004) Rapid synthesis of acyl transfer auxiliaries for cysteine-free native glycopeptide ligation. Org Lett 6:4659–4662
49. Wu B, Chen J, Warren JD, Chen G, Hua Z, Danishefsky SJ (2006) Building complex glycopeptides: development of a cysteine-free native chemical ligation protocol. Angew Chem Int Ed 45:4116–4125
50. Chen J, Chen G, Wu B, Wan Q, Tan Z, Hua Z, Danishefsky SJ (2006) Mature homogeneous erythropoietin-level building blocks by chemical synthesis: the EPO 114–166 glycopeptide domain, presenting the O-linked glycophorin. Tetrahedron Lett 47:8013–8016
51. Kumar KS, Brik A (2010) Accessing posttranslationally modified proteins through thiol positioning. J Pept Sci 16:524–529
52. Brik A, Yang YY, Ficht S, Wong CH (2006) Sugar-assisted glycopeptide ligation. J Am Chem Soc 128:5626–5627
53. Brik A, Wong CH (2007) Sugar-assisted ligation for the synthesis of glycopeptides. Chemistry 13:5670–5675
54. Yan LZ, Dawson PE (2001) Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. J Am Chem Soc 123:526–533
55. Brik A, Ficht S, Yang YY, Bennett CS, Wong CH (2006) Sugar-assisted ligation of N-linked glycopeptides with broad sequence tolerance at the ligation junction. J Am Chem Soc 128:15026–15033
56. Yang YY, Ficht S, Brik A, Wong CH (2007) Sugar-assisted glycoprotein synthesis. J Am Chem Soc 129:7690–7701
57. Bennett CS, Dean SM, Payne RJ, Ficht S, Brik A, Wong CH (2008) Sugar-assisted glycopeptide ligation with complex oligosaccharides: scope and limitations. J Am Chem Soc 130:11945–11952
58. Pentelute BL, Kent SBH (2007) Selective desulfurization of cysteine in the presence of Cys (Acm) in polypeptides obtained by native chemical ligation. Org Lett 9:687–690
59. Ficht S, Payne RJ, Brik A, Wong CH (2007) Second-generation sugar-assisted ligation: a method for the synthesis of cysteine-containing glycopeptides. Angew Chem Int Ed 46:5975–5979
60. Lutsky MY, Nepomniaschiy N, Brik A (2008) Peptide ligation via side-chain auxiliary. Chem Commun 10:1229–1231

61. Ajish Kumar KS, Harpaz Z, Haj-Yahya M, Brik A (2009) Side-chain assisted ligation in protein synthesis. Bioorg Med Chem Lett 19:3870–3874
62. Payne RJ, Ficht S, Tang S, Brik A, Yang YY, Case DA, Wong CH (2007) Extended sugar-assisted glycopeptide ligations: development, scope and applications. J Am Chem Soc 129 (44):13527–13536
63. Payne RJ, Ficht S, Greenberg WA, Wong CH (2008) Cysteine-free peptide and glycopeptide ligation by direct aminolysis. Angew Chem Int Ed 47:4411–4415
64. Thomas GL, Hsieh YSY, Chun CKY, Cai ZL, Reimers JR, Payne RJ (2011) Peptide ligations accelerated by N-terminal aspartate and glutamate residues. Org Lett 13:4770–4773
65. Hojo H, Ozawa C, Katayama H, Ueki A, Nakahara Y, Nakahara Y (2010) The mercaptomethyl group facilitates an efficient one-pot ligation at Xaa-Ser/Thr for (glyco) peptide synthesis. Angew Chem Int Ed 49:5318–5321
66. Chalker JM, Bernardes GJL, Lin YA, Davis BG (2009) Chemical modification of proteins at cysteine: opportunities in chemistry and biology. Chem Asian J 4:630–640
67. Okamoto R, Kajihara Y (2008) Uncovering a latent ligation site for glycopeptide synthesis. Angew Chem Int Ed 47:5402–5406
68. Heinrikson RL (1970) Selective S-methylation of cysteine in proteins and peptides. Biochem Biophys Res Commun 41:967–972
69. Gross E, Morell JL (1974) The reaction of cyanogen bromide with S-methylcysteine: fragmentation of the peptide 14–29 of bovine pancreatic ribonuclease A. Biochem Biophys Res Commun 59:1145–1150
70. Tam JP, Yu Q (1998) Methionine ligation strategy in the biomimetic synthesis of parathyroid hormones. Biopolymers 46:319–327
71. Haase C, Rohde H, Seitz O (2008) Native chemical ligation at valine. Angew Chem Int Ed 47:6807–6810
72. Wan Q, Danishefsky SJ (2007) Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. Angew Chem Int Ed 46:9248–9252
73. Hoffmann FW, Ess RJ, Simmons TC, Hanzel RS (1956) The desulfurization of mercaptans with trialkyl phosphites. J Am Chem Soc 78:6414
74. Walling C, Basedow OH, Savas ES (1960) Some extensions of the reaction of trivalent phosphorus derivatives with alkoxy and thiyl radicals; a new synthesis of thioesters. J Am Chem Soc 82:2181–2184
75. Walling C, Rabinowitz R (1957) The reaction of thiyl radicals with trialkyl phosphites. J Am Chem Soc 79:5326
76. Rohde H, Seitz O (2010) Ligation-desulfurization: a powerful combination in the synthesis of peptides and glycopeptides. Biopolymers 94:551–559
77. Dawson PE (2011) Native chemical ligation combined with desulfurization and deselenization: a general strategy for chemical protein synthesis. Isr J Chem 51:862–867
78. Chalker JM (2013) Prospects in the total synthesis of protein therapeutics. Chem Biol Drug Des 81:122–135
79. Johnson EC, Malito E, Shen Y, Rich D, Tang WJ, Kent SB (2007) Modular total chemical synthesis of a human immunodeficiency virus type 1 protease. J Am Chem Soc 129:11480–11490
80. Zheng J-S, Tang S, Qi Y-K, Wang Z-P, Liu L (2013) Chemical synthesis of proteins using peptide hydrazides as thioester surrogates. Nat Protocols 8:2483–2495
81. Fang GM, Li YM, Shen F, Huang YC, Li JB, Lin Y, Cui HK, Liu L (2011) Protein chemical synthesis by ligation of peptide hydrazides. Angew Chem Int Ed 50:7645–7649
82. Fang GM, Wang JX, Liu L (2012) Convergent chemical synthesis of proteins by ligation of peptide hydrazides. Angew Chem Int Ed 51:10347–10350
83. Wilkinson BL, Stone RS, Capicciotti CJ, Thaysen-Andersen M, Matthews JM, Packer NH, Ben RN, Payne RJ (2012) Total synthesis of homogeneous antifreeze glycopeptides and glycoproteins. Angew Chem Int Ed 51:3606–3610

84. Garner J, Harding MM (2010) Design and synthesis of antifreeze glycoproteins and mimics. ChemBioChem 11:2489–2498
85. Peltier R, Brimble MA, Wojnar JM, Williams DE, Evans CW, DeVries AL (2010) Synthesis and antifreeze activity of fish antifreeze glycoproteins and their analogs. Chem Sci 1:538–551
86. Sakamoto I, Tezuka K, Fukae K, Ishii K, Taduru K, Maeda M, Ouchi M, Yoshida K, Nambu Y, Igarashi J, Hayashi N, Tsuji T, Kajihara Y (2012) Chemical synthesis of homogeneous human glycosyl-interferon-beta that exhibits potent antitumor activity in vivo. J Am Chem Soc 134:5428–5431
87. Wang P, Dong S, Brailsford JA, Iyer K, Townsend SD, Zhang Q, Hendrickson RC, Shieh J, Moore MAS, Danishefsky SJ (2012) At last: erythropoietin as a single glycoform. Angew Chem Int Ed 51:11576–11584
88. Wang P, Dong S, Shieh J-H, Peguero E, Hendrickson R, Moore MAS, Danishefsky SJ (2013) Erythropoietin derived by chemical synthesis. Science 342:1357–1360
89. Murakami M, Okamoto R, Izumi M, Kajihara Y (2012) Chemical synthesis of an erythropoietin glycoform containing a complex-type disialyloligosaccharide. Angew Chem Int Ed 51:3567–3572
90. Liu S, Pentelute BL, Kent SBH (2012) Convergent chemical synthesis of [lysine24, 38, 83] human erythropoietin. Angew Chem Int Ed 51:993–999
91. Payne RJ (2013) Total synthesis of erythropoietin through the development and exploitation of enabling synthetic technologies. Angew Chem Int Ed 52:505–507
92. Warren JD, Miller JS, Keding SJ, Danishefsky SJ (2004) Toward fully synthetic glycoproteins by ultimately convergent routes: a solution to a long-standing problem. J Am Chem Soc 126:6576–6578
93. Chen G, Warren JD, Chen J, Wu B, Wan Q, Danishefsky SJ (2006) Studies related to the relative thermodynamic stability of C-terminal peptidyl esters of O-hydroxy thiophenol: emergence of a doable strategy for non-cysteine ligation applicable to the chemical synthesis of glycopeptides. J Am Chem Soc 128:7460–7462
94. Wong CTT, Tung CL, Li X (2013) Synthetic cysteine surrogates used in native chemical ligation. Mol Biosyst 9:826–833
95. Botti P, Tchertchian S (2006) Side chain extended ligation. WO/2006/133962
96. Crich D, Banerjee A (2007) Native chemical ligation at phenylalanine. J Am Chem Soc 129:10064–10065
97. Easton CJ, Hutton CA, Tan EW, Tiekink ERT (1990) Synthesis of homochiral hydroxy-α-amino acid derivatives. Tetrahedron Lett 31:7059–7062
98. Easton CJ, Hutton CA, Roselt PD, Tiekink ERT (1994) Stereocontrolled synthesis of β-hydroxyphenylalanine and β-hydroxytyrosine derivatives. Tetrahedron 50:7327–7340
99. Crich D, Banerjee A (2006) Expedient synthesis of threo-beta-hydroxy-alpha-amino acid derivatives: phenylalanine, tyrosine, histidine, and tryptophan. J Org Chem 71:7106–7109
100. Chen J, Wan Q, Yuan Y, Zhu JL, Danishefsky SJ (2008) Native chemical ligation at valine: a contribution to peptide and glycopeptide synthesis. Angew Chem Int Ed 47:8521–8524
101. Chen J, Wang P, Zhu JL, Wan Q, Danishefsky SJ (2010) A program for ligation at threonine sites: application to the controlled total synthesis of glycopeptides. Tetrahedron 66:2277–2283
102. Yang RL, Pasunooti KK, Li FP, Liu XW, Liu CF (2009) Dual native chemical ligation at lysine. J Am Chem Soc 131:13592–13593
103. Marin J, Didierjean C, Aubry A, Casimir JR, Briand JP, Guichard G (2004) Synthesis of enantiopure 4-hydroxypipecolate and 4-hydroxylysine derivatives from a common 4,6-dioxopiperidinecarboxylate precursor. J Org Chem 69:130–141
104. Ajish Kumar KS, Haj-Yahya M, Olschewski D, Lashuel HA, Brik A (2009) Highly efficient and chemoselective peptide ubiquitylation. Angew Chem Int Ed 48:8090–8094
105. El Oualid F, Merkx R, Ekkebus R, Hameed DS, Smit JJ, de Jong A, Hilkmann H, Sixma TK, Ovaa H (2010) Chemical synthesis of ubiquitin, ubiquitin-based probes, and diubiquitin. Angew Chem Int Ed 49:10149–10153

106. Hejjaoui M, Haj-Yahya M, Kumar KS, Brik A, Lashuel HA (2011) Towards elucidation of the role of ubiquitination in the pathogenesis of Parkinson's disease with semisynthetic ubiquitinated alpha-synuclein. Angew Chem Int Ed 50:405–409
107. Merkx R, de Bruin G, Kruithof A, van den Bergh T, Snip E, Lutz M, El Oualid F, Ovaa H (2013) Scalable synthesis of γ-thiolysine starting from lysine and a side by side comparison with δ-thiolysine in non-enzymatic ubiquitination. Chem Sci 4:4494–4498
108. Kumar KS, Bavikar SN, Spasser L, Moyal T, Ohayon S, Brik A (2011) Total chemical synthesis of a 304 amino acid K48-linked tetraubiquitin protein. Angew Chem Int Ed 50:6137–6141
109. Tan ZP, Shang SY, Danishefsky SJ (2010) Insights into the finer issues of native chemical ligation: an approach to cascade ligations. Angew Chem Int Ed 49:9500–9503
110. Harpaz Z, Siman P, Kumar KSA, Brik A (2010) Protein synthesis assisted by native chemical ligation at leucine. ChemBioChem 11:1232–1235
111. Shang SY, Tan ZP, Dong SW, Danishefsky SJ (2011) An advance in proline ligation. J Am Chem Soc 133:10784–10786
112. Dong S, Shang S, Tan Z, Danishefsky SJ (2011) Toward homogeneous erythropoietin: application of metal free dethiylation in the chemical synthesis of the Ala79-Arg166 glycopeptide domain. Isr J Chem 51:968–976
113. Townsend SD, Tan Z, Dong S, Shang S, Brailsford JA, Danishefsky SJ (2012) Advances in proline ligation. J Am Chem Soc 134:3912–3916
114. Ding H, Shigenaga A, Sato K, Morishita K, Otaka A (2011) Dual kinetically controlled native chemical ligation using a combination of sulfanylproline and sulfanylethylanilide peptide. Org Lett 13:5588–5591
115. Raibaut L, El Mahdi O, Melnyk O (2014) Solid phase protein chemical synthesis. Topics Curr Chem. doi:10.1007/128_2014_609
116. Siman P, Karthikeyan SV, Brik A (2012) Native chemical ligation at glutamine. Org Lett 14:1520–1523
117. Garner P (1984) Stereocontrolled addition to a penaldic acid equivalent: an asymmetric synthesis of threo-β-hydroxy-L-glutamic acid. Tetrahedron Lett 25:5855–5858
118. Malins LR, Cergol KM, Payne RJ (2013) Peptide ligation-desulfurization chemistry at arginine. ChemBioChem 14:559–563
119. Bang D, Pentelute BL, Kent SB (2006) Kinetically controlled ligation for the convergent chemical synthesis of proteins. Angew Chem Int Ed 45:3985–3988
120. Durek T, Torbeev VY, Kent SB (2007) Convergent chemical synthesis and high-resolution X-ray structure of human lysozyme. Proc Natl Acad Sci U S A 104:4846–4851
121. Torbeev VY, Kent SB (2007) Convergent chemical synthesis and crystal structure of a 203 amino acid "covalent dimer" HIV-1 protease enzyme molecule. Angew Chem Int Ed 46:1667–1670
122. Thompson RE, Chan B, Radom L, Jolliffe KA, Payne RJ (2013) Chemoselective peptide ligation-desulfurization at aspartate. Angew Chem Int Ed 52:9723–9727
123. Guan X, Drake MR, Tan Z (2013) Total synthesis of human galanin-like peptide through an aspartic acid ligation. Org Lett 15:6128–6131
124. Thompson RE, Liu X, Alonso-García N, Pereira PJB, Jolliffe KA, Payne RJ (2014) Trifluoroethanethiol: an additive for efficient one-pot peptide ligation – desulfurization chemistry. J Am Chem Soc 136:8161–8164
125. Rohde H, Schmalisch J, Harpaz Z, Diezmann F, Seitz O (2011) Ascorbate as an alternative to thiol additives in native chemical ligation. ChemBioChem 12:1396–1400
126. Moyal T, Hemantha HP, Siman P, Refua M, Brik A (2013) Highly efficient one-pot ligation and desulfurization. Chem Sci 4:2496–2501
127. Siman P, Blatt O, Moyal T, Danieli T, Lebendiker M, Lashuel HA, Friedler A, Brik A (2011) Chemical synthesis and expression of the HIV-1 Rev protein. ChemBioChem 12:1097–1104
128. Cergol KM, Thompson RE, Malins LR, Turner P, Payne RJ (2014) One-pot peptide ligation-desulfurization at glutamate. Org Lett 16:290–293

129. Malins LR, Cergol KM, Payne RJ (2014) Chemoselective sulfenylation and peptide ligation at tryptophan. Chem Sci 5:260–266
130. Scoffone E, Fontana A, Rocchi R (1966) Selective modification of tryptophan residue in peptides and proteins using sulfenyl halides. Biochem Biophys Res Commun 25:170–174
131. Scoffone E, Fontana A, Rocchi R (1968) Sulfenyl halides as modifying reagents for polypeptides and proteins. I. Modification of tryptophan residues. Biochemistry 7:971–979
132. Wilchek M, Miron T (1972) Conversion of tryptophan to 2-thioltryptophan in peptides and proteins. Biochem Biophys Res Commun 47:1015–1020
133. Shang S, Tan Z, Danishefsky SJ (2011) Application of the logic of cysteine-free native chemical ligation to the synthesis of human parathyroid hormone (hPTH). Proc Natl Acad Sci U S A 108:5986–5989
134. Bock A, Forchhammer K, Heider J, Leinfelder W, Sawers G, Veprek B, Zinoni F (1991) Selenocysteine: the 21st amino acid. Mol Microbiol 5:515–520
135. Lobanov AV, Hatfield DL, Gladyshev VN (2009) Eukaryotic selenoproteins and selenoproteomes. Biochim Biophys Acta 1790:1424–1428
136. Gieselman MD, Xie L, van Der Donk WA (2001) Synthesis of a selenocysteine-containing peptide by native chemical ligation. Org Lett 3:1331–1334
137. Quaderer R, Sewing A, Hilvert D (2001) Selenocysteine-mediated native chemical ligation. Helv Chim Acta 84:1197–1206
138. Hondal RJ, Nilsson BL, Raines RT (2001) Selenocysteine in native chemical ligation and expressed protein ligation. J Am Chem Soc 123:5140–5141
139. Huber R, Criddle RS (1967) Comparison of the chemical properties of selenocysteine and selenocystine with their sulfur analogs. Arch Biochem Biophys 122:164–173
140. Arnold AP, Tan KS, Rabenstein DL (1986) Nuclear magnetic resonance studies of the solution chemistry of metal complexes. 23. Complexation of methylmercury by selenohydryl-containing amino acids and related molecules. Inorg Chem 25:2433–2437
141. Pleasants JC, Guo W, Rabenstein DL (1989) A comparative study of the kinetics of selenol/diselenide and thiol/disulfide exchange reactions. J Am Chem Soc 111:6553–6558
142. Besse D, Siedler F, Diercks T, Kessler H, Moroder L (1997) The redox potential of selenocystine in unconstrained cyclic peptides. Angew Chem Int Ed 36:883–885
143. Nauser T, Dockheer S, Kissner R, Koppenol WH (2006) Catalysis of electron transfer by selenocysteine. Biochemistry 45:6038–6043
144. Metanis N, Keinan E, Dawson PE (2010) Traceless ligation of cysteine peptides using selective deselenization. Angew Chem Int Ed 49:7049–7053
145. Berry SM, Gieselman MD, Nilges MJ, Van der Donk WA, Lu Y (2002) An engineered azurin variant containing a selenocysteine copper ligand. J Am Chem Soc 124:2084–2085
146. Quaderer R, Hilvert D (2002) Selenocysteine-mediated backbone cyclization of unprotected peptides followed by alkylation, oxidative elimination or reduction of the selenol. Chem Commun 2620–2621
147. Malins LR, Mitchell NJ, Payne RJ (2014) Peptide ligation chemistry at selenol amino acids. J Pept Sci 20:64–77
148. Chalker JM, Bernardes GJL, Davis BG (2011) A "tag-and-modify" approach to site-selective protein modification. Acc Chem Res 44:730–741
149. Malins LR, Payne RJ (2012) Synthesis and utility of beta-selenol-phenylalanine for native chemical ligation-deselenization chemistry. Org Lett 14:3142–3145