Published in final edited form as:
Angew Chem Int Ed Engl. 2008 ; 47(44): 8521–8524. doi:10.1002/anie.200803523.

Native Chemical Ligation at Valine: A Contribution to Peptide and Glycopeptide Synthesis**

Jin Chen, **Qian Wan**, **Yu Yuan**, **Jianglong Zhu**, and **Samuel J. Danishefsky***
Laboratory for Bioorganic Chemistry, Sloan-Kettering Institute for Cancer Research, 1275 York Avenue, New York, NY 10065 (USA)

Keywords

native chemical ligation; glycopeptide; peptide; desulfurization; valine

Our laboratory has been pursuing the total synthesis of naturally occurring glycoproteins bearing multiple oligosaccharide domains. Specifically, efforts are well underway to accomplish a *de novo* total synthesis of erythropoietin alpha (EPO), in homogeneous form.[1] Although a variety of peptide ligation strategies have been developed to facilitate the merger of large, complex peptide and glycopeptide fragments,[2–9] the need for highly efficient methodology continues to motivate the chemical community to develop more powerful strategies. Our pursuit of the total synthesis of homogeneous erythropoietin (EPO) as well as other biologically active glycopeptides has inspired new glycopeptide ligations.[10] To achieve our most complex goals, we must learn how to overcome the serious obstacles in joining glycopeptides in an iterative fashion.

Native chemical ligation, developed by Kent and coworkers,[3] constituted a fundamental advance, allowing for the joining of two substantial peptide domains. A further development provided by our laboratory as well as others serves to extend NCL to the assembly of peptides bearing multiple sites of glycosylation.[10a,11] These methods currently require the presence of a cysteine residue at the *N*-terminus of the peptide coupling partner.

Application of NCL is often limited by the paucity of cysteine residues in naturally occurring proteins and glycoproteins. Several strategies have been investigated to circumvent the need for a cysteine in the target at the proposed ligation site. One approach involves the appendage of an auxiliary thiol to the *N*-terminal amino acid. Following ligation the auxiliary is cleaved.[10b,10c,4] This approach suffers from certain limitations, as the reaction may be inefficient at hindered ligation sites and difficulties can arise at the stage of auxiliary removal.

A second conceptual departure involves the use of an amino acid surrogate containing a thiol moiety. Following ligation, the product is converted to the desired amino acid. In this context, ligation at a serine site has been achieved through post-ligational conversion of cysteine to serine.[5] Similarly, formal methionine ligation has been accomplished through

**Support for this research was provided by the National Institutes of Health (CA28824). We thank Dr. George Sukenick, Ms. Hui Fang and Ms. Sylvi Rusli of the Sloan-Kettering Institute's NMR core facility for mass spectral and NMR spectroscopic analysis (SKI core grant no.: CA02848). We would like to express our appreciation to Rebecca Wilson for her assistance in preparing the manuscript.

*Prof. S. J. Danishefsky, Department of Chemistry, Columbia University, 3000 Broadway, New York, NY 10027 (USA), Fax: (+) 212-772-8691, s-danishefsky@mskcc.org.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

homocysteine coupling, followed by post-ligational methylation.[6] Recently, a cysteine-free thioester ligation[2] has been exemplified by Wong's protocol[7] which allows for peptide assembly. Although lysines must be protected, and the reaction usually takes more than 48 hours even at rather unhindered ligation sites, the method is quite promising.

In addition, two-step ligation/metal-based thiol reduction protocols which formally serves to accomplish ligation at *N*-terminal alanine[8] and phenylalanine[9] residues have been developed. However, these methods may not be compatible with a large range of functionalities, particularly sulfur-containing groups which are frequently present in peptide sequences. Alternatively, our group recently disclosed a mild and highly versatile free-radical cysteine reduction protocol, which tolerates all thiol-containing groups mentioned above, as well as oligosaccharide domains.[10e]

Theoretically, native chemical ligation could be achieved at any amino acid site, in the sense proposed by Dawson,[8] with a sulfhydryl group temporarily installed at a non cysteine site. The implementation of this concept requires that the key mechanistic steps (i.e. *trans*-thioesterification followed by S→N acyl transfer) were operative. Post-ligation desulfurization would afford the desired peptide or glycopeptide adduct possessing the natural amino acid residue at the ligation site.

In this paper, strategies for implementing the logic of native chemical ligation at valine residues are described. Valine is a rather abundant amino acid, with ca 6.6% frequency in nature, (compared to 1.7% for cysteine). There are two potential valine surrogates, β-thiol containing valine (penicillamine) and a γ-thiol valine. We envisioned that the γ-thiol valine containing a more reactive primary thiol group would serve as a more suitable precursor. As illustrated in Scheme 1, an *N*-terminal thiol-modified valine derivative, installed on peptide 2, would react with the peptide 1 thioester through *trans*-thioesterification. The resultant thioester-linked intermediate would undergo a rapid intramolecular acyl transfer thus creating an amide bond. Radical based desulfurization would serve to remove the thiol moiety and provide the desired peptide with valine at the ligation site.

Penicillamine (see **2**) containing a β-thiol moiety is a commercially available valine surrogate. As shown in Scheme 2, penicillamine was introduced to the N-terminus of the peptide as the acyl acceptor, and NCL was performed under standard conditions. Due to the tertiary nature of the thiol group, ligation proceeded rather slowly. Though the ligation yield was reasonable in a case in which glutamine was present at the C-terminus (Scheme 2A), when the more bulky threonine residue was incorporated at the C-terminus, the reaction became prohibitively slow (Scheme 2B). During our preparation of this communication, Seitz and co-workers reported a ligation culminating in valine, using penicillamine as the valine precursor.[12] They successfully achieved peptide bond formation when glycine, histidine, methionine, and leucine were presented at the C-terminus. Based on their results and our own (vide infra) in the case of peptides bearing less hindered C-terminal amino acids, penicillamine can serve in fostering valine ligation. However, our primary goal was that of developing a logic which would encompass more generally applicable ligations including certain β-branched C-terminal acyl donors. Our successes in this regard are described below.

Given the greater reactivity and diminished steric hindrance of primary thiol relative to tertiary thiol, we felt that it was worthwhile to develop a protocol for the efficient synthesis of a modified valine surrogate with a thiol group installed at its γ-position. Drawing from the earlier work of Rapoport,[13] in which a γ-hydroxyvaline was obtained from L-aspartic acid, we modified the route thereby enabling the preparation of γ-thiol valine derivative (**11**), albeit as two diastereomers. Either epimer of **11** could, in principle, be used for our

purposes, since removal of thiol group following ligation would eliminate the chirality at the β-position. Thus, as outlined in Scheme 3, the protected L-aspartic acid diester **6** was prepared from Fmoc-Asp-O*t*Bu (See supporting information for details). The PhFl (9-(9-phenylfluorenyl)) group was used to block the α-center, providing exclusively β-alkylated compound **7**. Regioselective reduction of **7** furnished alcohol **8** and *epi*-**8** as a mixture of diastereomers. They were readily separated by chromatography. Following mesylation and treatment with the DBU salt of thioacetic acid, acetylated thiol **9** was in hand. The latter was advanced to the target compound, **11**, in a straightforward manner, as shown.

With the required thiol containing amino acids in hand, we next sought to probe the versatility of our newly developed protocol. A variety of substrates incorporating a range of relevant functionalities was evaluated (Table 1). In the cases of less hindered *C*-terminal amino acids, such as Gln (entry 1) and Phe (entry 2), ligation proceeded very rapidly (1h). In the case of penicillamine as an acyl acceptor (peptide **2**), the ligation with the same peptide (peptide **1**) proceeded at a much slower rate (Scheme 2A). As the *C*-terminus became sterically more demanding, (entries 3 and 4), the reaction rate dropped correspondingly, but ligations were still completed within a reasonable time range and in good yield. Table 1 records examples of non cysteine driven NCL with two β-branched amino acids (Thr and Val) presented at the ligation site. While reacting with the same substrate (peptide **4**), penicillamine (peptide **2**) displayed much lower reactivity (Scheme 2B), only 4% product was formed in 10 hours. A range of *C*-terminal esters was investigated, including the standard thiophenyl ester (entry 3 and entry 4) as well as an ortho-thiophenolic ester, developed in our laboratory to allow for glycopeptide ligation (entry 1). We recently described a novel direct oxo-ester ligation protocol, which circumvents the need for a thioester intermediate, and allows ligation at sterically hindered C-terminal sites.[14] In this context, we were pleased to find that both C-terminal *p*-CN-phenyl ester (entry 2), and the *p*-NO_2-phenyl ester (entry 5) readily participate in our two-step valine ligation protocol. In addition, combination of C-terminal *p*-NO_2-phenyl ester and N-terminal γ-thiol valine furnished accomplished ligation at a Pro-Val site (entry 5), which is extremely steric demanding. Furthermore, substrates incorporating unprotected lysine residues (entry 2), as well as thiazolidine and cysteine (Acm) moieties (entry 1) were examined in an effort to probe the limits of our system. In each of these cases, ligation and subsequent reduction proceeded smoothly, providing the desired peptide adducts in moderate to good overall yield.

Interestingly, it was found that, although both γ-thiol valine epimers (i.e. **11** and *epi*-**11**) participate in the ligation protocol, epimer **11** is not markedly more reactive, as evidenced by a comparison of entries 3 and 4 (Table 1). Thus, when **11** serves as the *N*-terminal amino acid, ligation proceeds in 4h to yield product in 77% yield. Analogous ligation with *epi*-**11** is somewhat slower presumably due to a 1,2-*cis* steric interaction between the β-methyl and the amide moiety in the S-N acyl transfer step. That the difference in rates is relatively modest, suggests that the rate determining step in the coupling is the initial intermolecular transthioesterification rather than the intramolecular S→N acyl transfer.

We next investigated the efficiency of glycopeptide coupling employing the γ-thiol valine **11** as a key partner. Thus, glycopeptide **20**, with an N-linked glycan, and a C-terminal *ortho*-thiophenolic ester was prepared and coupled with peptide **14** (Scheme 4). The ligation proceeded smoothly and provided adduct **21'** in 1 h. Subsequent thiol reduction generated desired product **21** in nearly quantitative conversion (Figure 1).

In summary, an efficient and broadly useful two-step valine ligation protocol has been developed. The γ-thiol valine (**11**), containing a primary thiol group serves as a valine surrogate. The combination of NCL and metal-free reduction provides an efficient, two-step

protocol by which to accomplish a formal valine ligation. This method represents an important extension of existing NCL techniques, and should be valuable in our ongoing programs to bring synthesis to bear in domains which had previously been considered to be inaccessible to chemistry based interventions.[15]

References

1. a) Lai P, Everett R, Wang F, Arakawa T, Goldwasser E. J. Biol. Chem. 1986; 261:3116–3121. [PubMed: 3949763] b) Ridley DM, Dawkins F, Perlin E. J. Natl. Med. Assoc. 1994; 86:129–135. [PubMed: 8169988]
2. a) Blake J. Int. J. Pept. Protein Res. 1981; 17:273–274. [PubMed: 7228502] b) Aimoto S, Mizoguchi N, Hojo H, Yoshimura S. Bull. Chem. Soc. Jpn. 1989; 62:524–531. c) Aimoto S. Biopolymers. 1999; 51:247–265. [PubMed: 10618594]
3. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. Science. 1994; 266:776–779. [PubMed: 7973629]
4. Low DW, Hill MG, Carrasco MR, Kent SBH. Proc. Natl. Acad. Sci. USA. 2001; 98:6554–6559. [PubMed: 11390992]
5. Okamoto R, Kajihara Y. Angew. Chem. 2008; 120:5482–5486.Angew. Chem. Int. Ed. 2008; 47:5402–5406.
6. a) Tam JP, Yu Q. Biopolymer. 1998; 46:319–327. b) Pachamuthu K, Schmidt RR. Synlett. 2003; 5:659–662. c) Saporito A, Marasco D, Chambery A, Botti P, Pedone C, Ruvo M. Biopolymers. 2006; 83:508–518. [PubMed: 16888773]
7. Payne RJ, Fichet S, Greenberg WA, Wong C-H. Angew. Chem. 2008; 120:4483–4487.Angew. Chem. Int. Ed. 2008; 47:4411–4415.
8. Yan LZ, Dawson PE. J. Am. Chem. Soc. 2001; 123:526–533. [PubMed: 11456564]
9. Crich D, Banerjee A. J. Am. Chem. Soc. 2007; 129:10064–10065. [PubMed: 17658806]
10. a) Warren JD, Miller JS, Keding SJ, Danishefsky SJ. J. Am. Chem. Soc. 2004; 126:6576–6578. [PubMed: 15161285] b) Chen G, Warren JD, Chen J, Wu B, Wan Q, Danishefsky SJ. J. Am. Chem. Soc. 2006; 128:7560–7462. c) Wu B, Chen J, Warren JD, Chen G, Hua Z, Danishefsky SJ. Angew. Chem. 2006; 118:4222–4231.Angew. Chem. Int. Ed. 2006; 45:4116–4125. d) Chen G, Wan Q, Tan Z, Kan C, Hua Z, Ranganathan K, Danishefsky SJ. Angew. Chem. 2007; 119:7527–7531.Angew. Chem. Int. Ed. 2007; 46:7383–7387. e) Wan Q, Danishefsky SJ. Angew. Chem. 2007; 119:9408–9412.Angew. Chem. Int. Ed. 2007; 46:9248–9252.
11. For reviews on glycopeptide and glycoprotein synthesis, see a) Davis BG. Chem. Rev. 2002; 102:579–601. [PubMed: 11841255] b) Grogan MJ, Pratt MR, Marcaurelle LA, Bertozzi CR. Annu. Rev. Biochem. 2002; 71:593–634. [PubMed: 12045107] c) Liu L, Bennett CS, Wong C-H. Chem. Commun. 2006:21–33.
12. Haase C, Rohde H, Seitz O. Angew. Chem. Int. Ed. 2008; 47 Early View.
13. a) Rapoport H, Wolf JP. J. Org. Chem. 1989; 54:3164–3173. b) Gmeiner P, Feldman PL, Chu-Moyer MY, Rapoport HJ. J. Org. Chem. 1990; 55:3068–3074.
14. Wan Q, Chen J, Yuan Y, Danishefsky SJ. 2008 submitted.
15. Wilson RM, Danishefsky SJ. Pure and Applied Chem. 2007; 79:2189–2216.

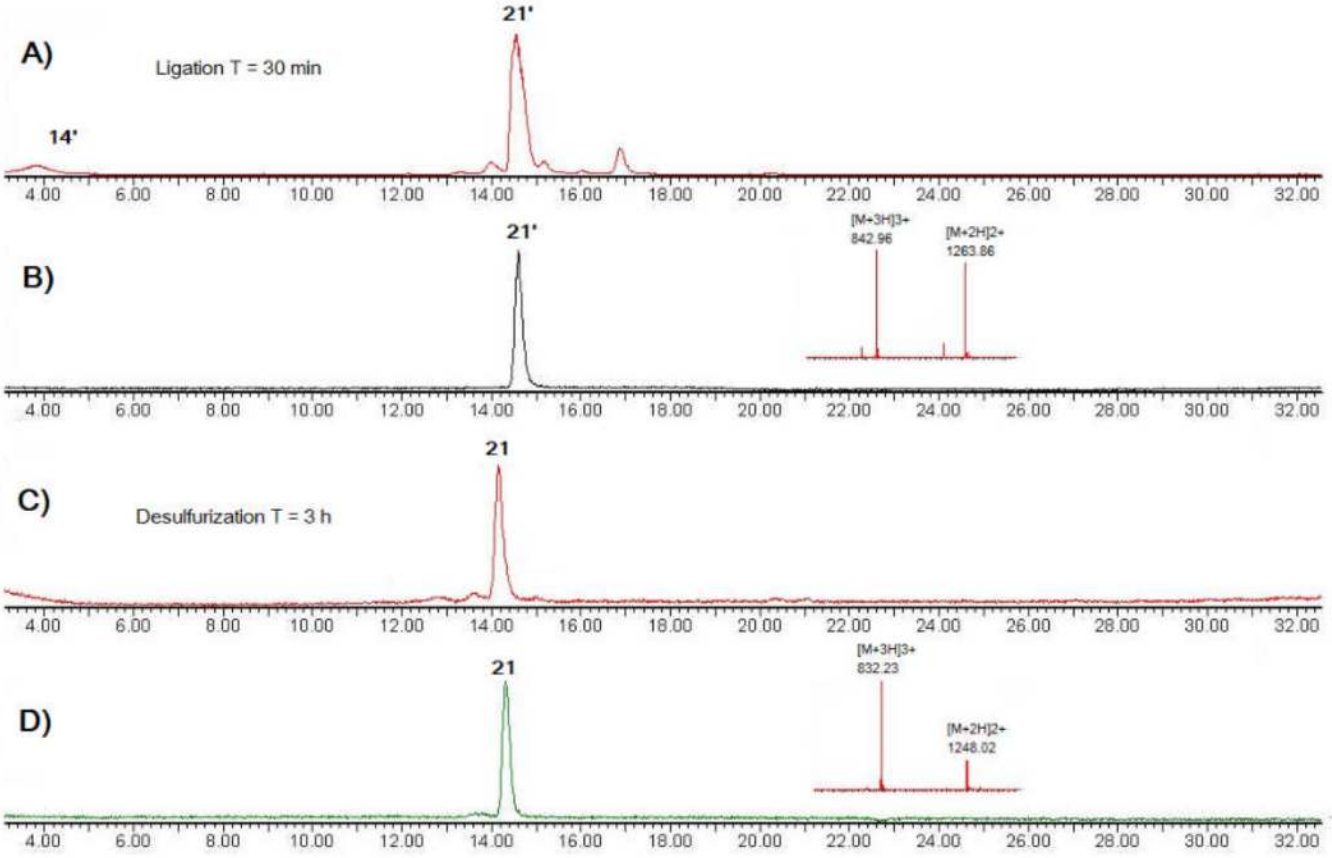
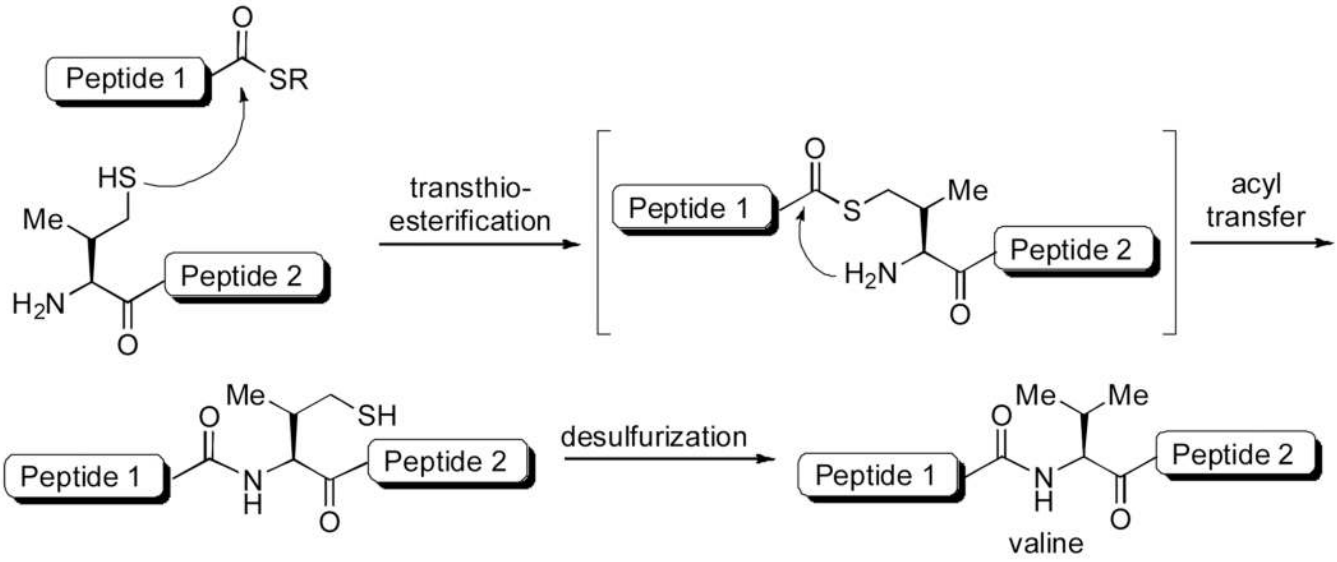
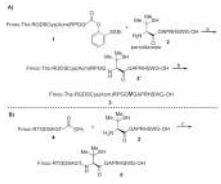


Figure 1.

Synthesis of glycopeptide **21**. LC-MS trace of: A) Ligation between glycopeptide **20** and peptide **14** after 30 min, **21'** is the ligation product. B) **21'** after HPLC purification with the observed mass $[M+2H]^{2+}$ = 1263.86, $[M+3H]^{3+}$ = 842.96. C) Desulfurization of **21'** after 3h, **21** is the desulfurization product. D) **21** after HPLC purification with the observed mass $[M+2H]^{2+}$ = 1248.02, $[M+3H]^{3+}$ = 832.23.

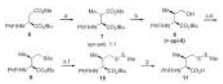


Scheme 1.
Native chemical ligation at valine.



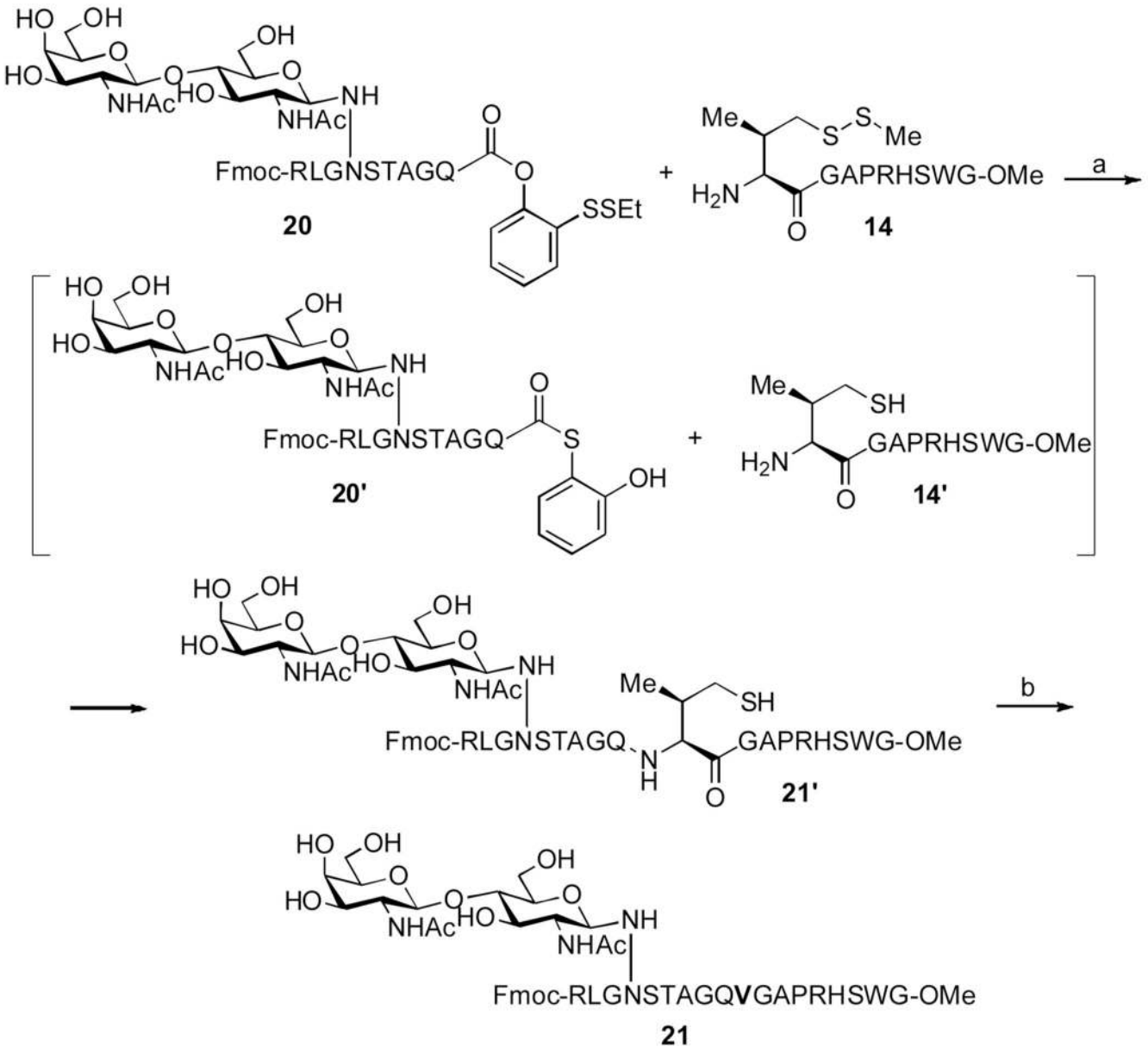
Scheme 2.

Reagents and conditions: A. a) buffer at pH 6.5 (6.0 M Gn•HCl, 188.8 mM Na_2HPO_4), TCEP, RT, 9h, 66% yield; b) TCEP, VA-044, *t*BuSH, 37 °C, 3h, 79% yield. B. c) buffer at pH 6.5 (6.0 M Gn•HCl, 188.8 mM Na_2HPO_4), TCEP, RT, 10h, 4% yield based on LC-MS trace. VA-044 = 2, 2 ’-azobis-[2-(2-imidazolin-2-yl)propane] dihydrochloride. TCEP = tris(2-carboxyethyl)phosphine.



Scheme 3.

Reagents and conditions: a) KHMDS, MeI, THF, −78 °C, 3h, 97%; b) DIBAL-H, THF, −35 °C, 1h, 83% (combined yield for 8 and epi-8); c) MsCl, Et_3N, CH_2Cl_2, 0 °C, 1h; d) AcSH, DBU, DMF, RT



Scheme 4.

Reagents and conditions: a) buffer at pH 6.5 (6.0 M Gn•HCl, 188.8 mM Na_2HPO_4), TCEP, RT, 30min, 90% yield; b) TCEP, VA-044, *t*BuSH, 37 °C, 3h, 89% yield.

Table 1

NCL at valine through ligation followed by free-radical desulfurization.

[a] For detailed reaction conditions, please see supporting information

[b] Isolated yield.