Published in final edited form as:

Chembiochem. 2009 May 4; 10(7): 1234–1242. doi:10.1002/cbic.200800741.

Enzymatic glycosylation of triazole-linked GlcNAc/Glc-peptides. Synthesis, stability, and anti-HIV activity of triazole-linked HIV-1 gp41 glycopeptide C34

Wei Huang^[a], Stan Groothuys^[b], Alonso Heredia^[a], Brian H.M. Kuijpers^[b], Floris P.J.T. Rutjes^[b], Floris L. van Delft^[b], and Lai-Xi Wang^[a]

^[a] Institute of Human Virology University of Maryland School of Medicine Baltimore, MD 21201, USA
Fax: (+1) 410 706 5068

^[b] Institute for Molecules and Materials Radboud University Nijmegen Heyendaalseweg 135 6525 AJ, Nijmegen, The Netherlands Fax: (+31) 24 365 3393

Abstract

Endoglycosidase-catalyzed transglycosylation of triazole-linked glucose (Glc) and N-acetylglucosamine (GlcNAc)-containing dipeptides and polypeptides was achieved, using synthetic sugar oxazoline as the donor substrate. It was found that both N- and C-linked Glc/GlcNAc-containing triazole derivatives were effective substrates for endo- β -N-acetylglucosaminidase from *Arthrobacter* (Endo-A) for transglycosylation, demonstrating a broad acceptor substrate specificity for Endo-A. This chemoenzymatic method was successfully used for the synthesis of a novel triazole-linked C34 glycopeptide derived from HIV-1 envelope glycoprotein gp41. It was shown that the synthetic C34 glycopeptide possessed potent anti-HIV activity with an IC₅₀ of 21 nM. The triazole-linked C34 glycopeptide demonstrated much enhanced stability toward protease- and glycoamidase-catalyzed digestion, showing the protective effects of glycosylation and the stability of the triazole linkage. These favorable properties suggest that the triazole-linked C34 glycopeptide may be valuable for further development as an anti-HIV drug candidate.

Keywords

HIV-1 gp41; triazole-linked glyco-C34; fusion inhibition; protease resistance; chemoenzymatic synthesis

Introduction

Synthetic peptides derived from the HIV-1 inner envelope glycoprotein gp41 represent a new class of HIV-1 entry inhibitors that exert their antiviral activity by blocking viral membrane fusion.[1,2] In 2003, a 36-mer peptide derived from the C-terminal ectodomain of HIV-1 gp41, called Enfuvirtide or T20, was first approved as a drug for the treatment of AIDS patients in US and Europe.[3,4] However, there are in general several limitations for polypeptides as drug candidates, including poor oral bioavailability and short serum's half-life because of their susceptibility to protease digestion. It is well-known that glycosylation can significantly influence a peptide's pharmacokinetic properties such as its protease susceptibility.

We became interested to investigate the influence of glycosylation on the activity and stability of peptide C34, a 34-mer peptide derived from the C-ectodomain of gp41 that also demonstrates nanomolar anti-HIV activity.[5,6] In addition to the potential susceptibility to protease digestion *in vivo*, another barrier for developing the C34 peptide as a drug candidate is its poor solubility. Earlier work from one of us indicated that glycosylation at the natural glycosylation site (N637) of peptide C34 led to significant enhancement of solubility with only moderate loss of activity, as reflected by a fusion assay.[7] It was speculated that the carbohydrate portion of the molecule could furthermore confer substantial protection against protease digestion and, as such, we were intrigued by the potential of glycosylated isosteres of C34 for application as fusion inhibitors with enhanced pharmacokinetic profile. An enhanced serum's half-life is particularly important for this class of anti-HIV drugs, as a prolonged for therapeutic efficacy means the use of relatively low dose and less frequency of intravenous injection. In this paper, we describe a chemoenzymatic synthesis, stability, and anti-HIV activity of glycosylated isosteres of C34, in which the glycan and peptide are linked through a triazole moiety.

Results and Discussion

Transglycosylation of triazole-linked GlcNAc-peptide

Methods for the chemical synthesis of glycopeptides have advanced considerably, and peptides carrying relatively simple carbohydrates can now be prepared by a variety of procedures.[8, 9] However, the construction of polypeptides carrying a larger oligosaccharide moiety is still a difficult task.[10] One of the most promising techniques that avoids the typical problems associated with the assembly of a glycopeptide carrying a complex glycan portion involves the endo- β -*N*-acetylglucosaminidase-catalyzed transglycosylation of a peptide.[11] In this approach, only a monosaccharide moiety such as *N*-acetylglucosamine (GlcNAc) needs to be incorporated during solid-phase peptide synthesis, and the sugar chain is subsequently extended by the attachment of an intact oligosaccharide in a single step without the need of any protecting groups.[11] Previous studies have demonstrated that the endo- β -*N*-acetylglucosaminidase from *Arthrobacter* (Endo-A) could recognize both a GlcNAc and a glucose (Glc) moiety in a peptide context as an acceptor substrate for transglycosylation.[11, 12] It was also shown that C-linked GlcNAc peptide could serve as acceptor substrate for Endo-A catalyzed transglycosylation to form a C-glycopeptide.[13,14] The substrate flexibility of Endo-A was further demonstrated by the ability of Endo-A to introduce N-glycans into various natural products, with a corresponding synthetic sugar oxazoline as the donor substrate.[15] However, it has not been demonstrated whether Endo-A could tolerate triazole-linked Glc or triazole-linked GlcNAc moiety as an acceptor. Therefore, we first evaluated the feasibility of two triazole-linked Glc-containing dipeptides for Endo-A catalyzed transglycosylation. A triazole is of particular interest as a bridging moiety for an *N*-glycopeptide or *N*-glycoprotein, since triazoles have early on been proposed and corroborated as promising amide bioisosteres.[16-18] Moreover, such an *N*-glycopeptide analog is expected to be resistant to glycoamidase-catalyzed de-glycosylation. Thus, it will be more stable *in vivo* than the corresponding natural *N*-glycopeptide as drug candidate. To establish an efficient chemoenzymatic method for the synthesis of these types of glycopeptides, we first evaluated the feasibility of two model triazole-linked Glc-containing dipeptides for Endo-A catalyzed transglycosylation. The C-linked glycodipeptide **3** was synthesized by Cu(I)-catalyzed 1,3-dipolar cycloaddition[19] between the azido-containing dipeptide derivative **1** and the glucosylalkyne **2** according to our previous report.[20] Sequential deprotection of the peptide and carbohydrate moieties gave the triazole-linked Glc-dipeptide **4** in 71% yield (Scheme 1). It must be noted that in Glc-dipeptide **4**, the peptide was linked to the glucose moiety through a C-glycosidic linkage at the anomeric center of the sugar. Another model triazole-linked Glc-dipeptide, compound **7**, with glucose being N-linked to the triazole moiety, was synthesized by 1,3-dipolar cycloaddition between the alkyne-containing dipeptide derivative **5** and the glucosyl azide **6**, following our previous

procedure.[21] The copper-catalyzed 1,3-dipolar cycloaddition gave the desired product **7** in 49% yield after deprotection (Scheme 1).

Having two complementary glycopeptides at hand, the transglycosylation under the action of Endo-A was ready to be tested. The tetrasaccharide oxazoline Man₃GlcNAc-oxazoline (**8**) was used as the donor substrate, as it was shown to be an excellent substrate for Endo-A catalyzed transglycosylation.[15,22]. Much to our satisfaction, it was found that both the triazole-linked Glc-dipeptides (**3** and **7**) could serve as acceptor substrate for Endo-A catalyzed transglycosylation to give the corresponding pentasaccharide derivatives **9** and **10**, respectively (Scheme 2). Interestingly, it was observed that transglycosylation with N-linked triazole derivative **7** proceeded much faster than the reaction with the C-linked derivative **4**, as revealed by the HPLC monitoring of the enzymatic reactions (Figure 1) (donor: acceptor = 3:1, pH 7.0, 23 °C). Under the same enzymatic glycosylation conditions, the N-linked compound **7** reached a conversion of around 80%, whereas the C-linked derivative **4** gave only 52% conversion after 45 min. It should be pointed out that the hydrolysis of products **9** and **10** by Endo-A was not observed within 24 h, which is in line with our earlier findings.[22] These results indicate that the N-linked triazole Glc-derivative **7** serves as a better substrate than the C-linked triazole Glc-derivative **4**. A plausible explanation of this observed difference in reactivity is that the N-linked triazole more closely resembles the amide linkage found in natural N-glycopeptide, a finding not earlier described which may have an important impact in the future design and synthesis of triazoles as amide isosteres. Our experiments further confirm the very broad substrate specificity of Endo-A. Clearly, the endoglycosidase not only efficiently accepts a glucose moiety as the acceptor, but it could also tolerate the modification at the anomeric position with a triazole heterocycle, implicating a flexibility of Endo-A in organic synthesis. To attest the broad substrate specificity, we have recently demonstrated that Endo-A could take a range of glucose-containing natural products as substrates for transglycosylation to provide novel glycosylated natural products.[15]

Chemoenzymatic synthesis of triazole-linked GlcNAc-C34

Having established the suitability of Endo-A catalyzed transglycosylation for the preparation of small triazole-linked glycopeptides analogues, the stage was set to prepare a C34 peptide with a large glycan. As the N-linked triazole Glc-dipeptide **7** was found to be a better substrate for Endo-A than the C-linked derivative **4**, we chose to incorporate a GlcNAc moiety into the anti-HIV peptide C34 via an N-triazole linkage. Two principal strategies were attempted for synthesis of glyco-C34, *i.e.* incorporation of a preassembled triazole-linked glycoamino acid glycine in the growing peptide chain or a two-step approach involving initial incorporation of propargylglycine followed by introduction of a monosaccharide moiety by [3+2] cycloaddition. Since the latter strategy was successfully applied in the preparation of compounds **4** and **7**, and attachment of a sugar post-solid phase synthesis is more modular in nature, it was investigated first (Scheme 3). Solid phase peptide synthesis was performed by using the standard Boc chemistry.[23] A propargylglycine moiety was incorporated into the C34 peptide sequence at the conserved glycosylation site (N-637) using a propargylglycine derivative[24] as the building block, followed by ligation with 2-acetamido-2-deoxy-glucosyl azide **13**[25] by copper-catalyzed cycloaddition (Scheme 3). Unfortunately, the copper-catalyzed [3+2] cycloaddition on the resin did not succeed, either on the terminal Boc-protected propargylglycine (**12**) or on the internal propargylglycine (**14**).

In contrast, the desired triazole-linked GlcNAc-C34, GlcNAc-T-C34 (**16**), was smoothly obtained by incorporation of the glycoamino acid **15**, readily derived in two steps from Boc-L-propargylglycine and 2-acetamido-2-deoxy-glucosyl azide **13** into the solid phase peptide synthesis starting with resin-bound peptide **11** (Scheme 4). After solid phase synthesis, the polypeptide was retrieved from the resin by treatment with HF with concomitant removal of

all the protecting groups except the 3- and 4-*O*-acetyl groups on the GlcNAc moiety. Finally, the two *O*-acetyl groups were effectively removed by treatment with hydrazine. The final GlcNAc-T-C34 product (**16**) was purified by preparative reverse-phase HPLC and characterized by ESI-MS.

Chemoenzymatic synthesis of C34 glycopeptides with native *N*-linkage and triazole-linkage

The suitability of GlcNAc-T-C34 (**16**) as the acceptor substrate for Endo-A for the transglycosylation with Man₃GlcNAc-oxazoline (**8**) as the donor substrate was performed and compared to the naturally *N*-linked derivative, GlcNAc-C34, which was prepared according to our previously reported procedure.[7] Gratifyingly, after 2 h, essentially complete conversion was observed by HPLC for both acceptors GlcNAc-C34 and GlcNAc-T-C34, to give the corresponding transglycosylation products Man₃GlcNAc₂-C34 (**17**) and Man₃GlcNAc₂-T-C34 (**18**), respectively (Scheme 5). The products were purified by HPLC and the identity was characterized by ESI-MS (Figure 2).

Anti-HIV activity of the synthetic C34 glycopeptides

The antiviral activities of the synthetic C34 glycopeptides, together with the control C34 peptide, were assayed by inhibiting the infection of TZM-bl cells with HIV-1 IIIB. TZM-bl is a cell line expressing CD4 and viral coreceptors (CCR5 and CXCR4) and containing integrated copies of the luciferase gene under control of the HIV-1 promoter. All the glycoforms of C34 have demonstrated potent inhibitory activities against HIV-1 infection at *nanomolar* concentrations. Results from a typical inhibition assay experiment were demonstrated in Figure 3. The estimated IC₅₀ data for the respective C34 peptides and glycopeptide are: C-34, 7.0 nM; GlcNAc-C34: 14.3 nM; GlcNAc-T-C34, 4.5 nM; Man₃GlcNAc₂-C34, 16.4 nM, and Man₃GlcNAc₂-T-C34, 21.0 nM. It was found that attachment of a sugar moiety resulted in a 2-3 fold decrease in the inhibitory activity based on the current cell-based infectivity assay. The results are consistent with the previous observations on the antiviral activity of C34 glycopeptides based on a fusion assay.[7] Why the anti-HIV activity of GlcNAc-C34 was reduced to some extent, the triazole-linked compound, GlcNAc-T-C34, actually showed a slight enhancement in inhibitory activity in comparison with the non-modified C34 (IC₅₀, 4.5 nM vs. 7.0 nM). These results suggest that attachment of a small monosaccharide residue at the glycosylation site did not have significant impact on the anti-HIV activity. The slight variation in the *in vitro* anti-HIV activity might result in the different nature of the linkers in the two compounds. However, as expected, the attachment of a larger oligosaccharide moiety such as Man₃GlcNAc₂ led to about 3-fold decrease in the anti-HIV activity, presumably caused by steric hindrance of the sugar moiety during binding. Nevertheless, for the perspective of anti-HIV drug development, the glycopeptides might be superior to C34 in two aspects. First, the two glycosylated C34, Man₃GlcNAc₂-C34 and Man₃GlcNAc₂-T-C34, showed much better water solubility under physiological conditions than C34, which overcomes a major drawback for the poorly soluble C34 as a drug candidate. Secondly, a major concern for polypeptide therapeutics is their sensitivity to digestion by proteases and other enzymes. Because of the general protective effect of glycosylation, the glycopeptides may be more resistant to protease digestion *in vivo* than the non-glycosylated C34. This effect was demonstrated by the protease digestion experiments shown below.

Glycoamidase stability of triazole-linked glycosylated C34 versus amide-linked glyco-C34

Peptide-*N*⁴-(*N*-acetyl-β-D-glucosaminy)asparagine amidases (PNGases) such as PNGase F is a class of glycoamidases that cleave the β-aspartylglucosamine bond of asparagine-linked glycopeptides or glycoproteins, thereby converting the asparagine residue to an aspartic acid. The enzyme PNGase F has a broad substrate specificity, with the only restriction that both the amino and carboxyl groups of the asparagine residue must be engaged in a peptide linkage,

while the oligosaccharide must consist of at least the *N,N'*-diacetylchitobiose core, GlcNAc-(β 1 $\rightarrow$ 4)-GlcNAc[26]. Therefore, subjection of either **17** or **18** to the action of PNGase F provides a good insight into the enzymatic stability of the triazole-linked glycopeptide **18**. Much to our satisfaction, the triazole-linked glycopeptide **18** was completely resistant to the glycoamidase catalyzed hydrolysis, whereas the native glycopeptide **17** was completely hydrolyzed within 1 hour by the glycoamidase under the same conditions (Figure 4). These results suggest that the triazole-linked glyco-C34 would not be metabolized *in vivo* by glycoamidases. It should be pointed out that although it is known that PNGases exist in cytosol, it is not clear how significant the PNGase activity is in serum.

Protease stability of Man₃GlcNAc₂-C34 and Man₃GlcNAc₂-T-C34

We also investigated the stability of C34 as well as the synthetic glycosylated C34 derivatives during protease-catalyzed digestion by two prototypical proteases, *i.e.* trypsin and chymotrypsin. While trypsin prefers to hydrolyze the amide bond of the basic amino acid residues, chymotrypsin is specific for cleavage of the peptide linkages next to an aromatic amino acid residue (Figure 5). Since C34 has several basic and aromatic residues, it is interesting to test whether the glycan attached at the N-637 glycosylation site (N25 in the 34-mer peptide) confers protection against protease digestion.

First of all, it was observed that the native, non-glycosylated C34 was nearly completely digested by trypsin after 4 h (Figure 6A). Attachment of a monosaccharide at N637 had little effect on proteolysis, for either amide- or triazole-linked glycopeptide, but the large glycan clearly exerted some protective effects on the polypeptide against trypsin digestion. For example, treatment with trypsin for 4 h resulted in 85% digestion of the non-glycosylated peptide C34, whereas under the same conditions, treatment of the Man₃GlcNAc₂-C34 and Man₃GlcNAc₂-T-C34 led to hydrolysis of about 60% and 50% of the corresponding glycosylated C34, respectively, showing the protective effect from the attached glycan. In the case of chymotrypsin, the difference between the peptide bearing the pentasaccharide and the plain C34 peptide were found to be even more dramatic (Figure 6B). As demonstrated, the nonglycosylated C34 was completely digested by chymotrypsin within 4 h and again, little protective effect is observed by attachment of a single sugar. However, under the same digestion conditions, only about 10% and 20% of the Man₃GlcNAc₂-C34 and Man₃GlcNAc₂-T-C34 were hydrolyzed after 4 h. These results clearly implicate that the glycosylated C34 carrying a larger N-glycan could be much more stable *in vivo* than the non-glycosylated C34. The enhanced protease stability, together with their potent anti-HIV activity and their enhanced water solubility, suggests that glycosylated C34 may be a valuable candidate for further development as an anti-HIV drug.

Conclusion

A novel triazole-linked glycopeptide C34 was efficiently synthesized by a chemoenzymatic approach. It was found that the synthetic glycopeptide possessed potent anti-HIV activity, demonstrated dramatically enhanced water solubility, and enhanced protease stability in comparison with the non-glycosylated peptide C34. These findings reveal favorable properties for the glycosylated C34, which may be valuable for further development as an anti-HIV drug candidate.

Experimental Section

Materials

The recombinant *Arthrobacter protophormiae* endo- β -N-acetylglucosaminidase (Endo-A) was overproduced in *E. coli* and purified by affinity chromatography according to the literature.

(K. Fujita, N. Tanaka, M. Sano, I. Kato, Y. Asada, K. Takekawa, Biochem. Biophys. Res. Commun., 2000, 267, 134-138). The pGEX-2T/Endo-A plasmid used for the overexpression was kindly provided by Prof. K. Takekawa. The peptide N-glycosidase F (PNGase F) were purchased from New England Biolabs Inc. Peptide C34 and GlcNAc-C34 were synthesized on a solid-phase peptide synthesizer using Fmoc approach, as previously reported.[7] All other reagents were purchased from Sigma/Aldrich and used as received.

Preparation of **L-T1M[4-(β-D-Glc)]-L-Phe-NH₂ (4)**

Cbz-L-T1M{4-[β-D-Glc(Ac)₄]}-L-Phe-NH₂ (**3**; 73 mg, 0.095 mmol)²⁰ was dissolved in MeOH (1 mL), and K₂CO₃ (2 mg, 9 μmol) was added. The reaction was stirred for 1.5 h at room temperature. The crude mixture was neutralized with Amberlite IR 120 (which was pre-washed with MeOH), filtered and evaporated the solvent to obtain the deacetylated product. This was dissolved again in MeOH (2 mL) and Pd-C (12 mg, 10 μmol) was added. The reaction was stirred overnight under an atmospheric pressure of H₂. The suspension was filtered and the solvent was evaporated. The product was lyophilized from AcOH (1 M in H₂O) to yield the desired product **4** (35 mg, 71%). ¹H NMR (400 MHz, D₂O) δ = 8.04 (s, 1 H), 7.51-7.31 (m, 5 H), 4.72-4.50 (m, 4 H), 4.06-3.90 (m, 2 H), 3.84-3.74 (m, 2 H), 3.72-3.55 (m, 3 H), 3.20 (dd, *J* = 13.9, 6.5 Hz, 1 H), 3.06 (dd, *J* = 13.9, 8.4 Hz, 1 H). ¹³C NMR (75 MHz, D₂O) δ = 174.7, 172.6, 172.6, 144.2, 135.8, 128.7, 128.2, 126.7, 125.1, 79.5, 76.5, 72.9, 72.4, 69.0, 60.3, 54.0, 52.5, 36.5. HRMS (ESI) calculated for C₂₀H₂₉N₆O₇ [M+H]⁺ 465.20977; found 465.20851.

L-Pro-L-T4M[1-(β-D-Glc)]-OH (**7**)

Boc-L-Pro-L-T4M{4-[β-D-Glc(Ac)₄]}-OH was prepared from **5** and **6** as earlier described.²¹ Next, the Boc-L-Pro-L-T4M{1-[β-D-Glc(Ac)₄]}-OH (40 mg, 0.06 mmol) was dissolved in 2.6 M HCl in EtOAc (2 mL) and stirred for 30 min, the solvent was evaporated *in vacuo*. The crude product was dissolved in MeOH (3 mL), a catalytic amount of K₂CO₃ was added and the mixture was stirred overnight. Purification of the product with an acidic ion exchange column (IRA-120) afforded **20** (16 mg, 0.03 mmol, 53%) as a white solid. ¹H NMR (400 MHz, D₂O): δ = 7.97 (s, 1H), 5.58 (d, *J* = 9.1 Hz, 1H), 4.28-4.16 (m, 1H), 3.84 (t, *J* = 9.2 Hz, 1H), 3.77 (d, *J* = 11.6 Hz, 1H), 3.67-3.51 (m, 4H), 3.47 (t, *J* = 9.2 Hz, 1H), 3.36-3.09 (m, 4H), 2.34-2.20 (m, 1H), 1.96-1.81 (m, 3H). ¹³C NMR (75 MHz, D₂O): δ = 172.9, 171.7, 168.9, 142.5, 142.4, 123.2, 86.8, 78.3, 75.4, 71.7, 71.6, 68.4, 59.9, 59.1, 52.7, 52.3, 52.2, 46.0, 45.9, 29.1, 26.5, 25.9, 23.1. HRMS (ESI) *m/z* calculated for C₁₆H₂₆N₅O₈ (M+H)⁺: 416.1781, found: 416.1791.

Boc-L-T4M(1-[β-D-GlcNAc(Ac)₃])-OH (**15**)

To a solution of GlcNAc(Ac)₃N₃ (**13**²⁵, 0.93 g, 2.5 mmol) and Boc-L-propargylglycine²⁴ (0.53 g, 2.5 mmol) in *tert*-butanol (5 mL) was added a mixture of Cu(OAc)₂ (0.10 g, 0.50 mmol) and sodium ascorbate (0.20 g, 1.0 mmol) in H₂O (5 mL). The reaction was stirred overnight, water was added and the product was extracted with CH₂Cl₂. The combined organic layers were washed aqueous NaCl, dried over Na₂SO₄ and evaporated *in vacuo*. The crude product was purified by flash chromatography using EtOAc/heptane to give **15** (1.25 g, 2.14 mmol, 86%) as a white solid. FTIR (ATR): ν = 1744, 1364, 1213 cm⁻¹. ¹H NMR (400 MHz, CDCl₃): δ = 7.75 (s, 1H), 6.48 (br d, *J* = 9.6 Hz, 1H), 5.95 (d, *J* = 10.0 Hz, 1H), 5.62 (d, *J* = 8.0 Hz, 1H), 5.39 (dd, *J* = 10.8, 10.0 Hz, 1H), 5.24 (t, *J* = 9.6 Hz, 1H), 4.71-4.55 (m, 2H), 4.30 (dd, *J* = 12.8, 4.8 Hz, 1H), 4.18-4.10 (m, 1H), 4.07-4.01 (m, 1H), 3.34 (dd, *J* = 15.2, 5.2 Hz, 1H), 3.23 (dd, *J* = 15.2, 4.8 Hz, 1H), 2.08 (s, 3H), 2.07 (s, 3H), 2.06 (s, 3H), 1.80 (s, 3H), 1.45 (s, 9H). ¹³C NMR (75 MHz, CDCl₃): δ = 171.6, 170.7, 170.4, 170.3, 169.1, 155.4, 142.9, 121.7, 86.2, 80.0, 75.1, 72.3, 68.2, 61.9, 53.7, 53.3, 52.6, 28.8, 28.6, 23.0, 21.0, 21.0, 20.9. HRMS (ESI): *m/z* calculated for C₂₄H₃₅O₁₂N₅Na (M+Na)⁺ 608.21799, found 608.21924.

GlcNAc-T-C34 (16)

The peptide was obtained via manual solid phase peptide synthesis (SPPS) using the *in situ* neutralization/ HBTU activation procedure for Boc chemistry on a *p*-methylbenzhydrylamine (MBHA) resin. *N*-Terminal acetylation was performed by 2×2 min treatment with 1/1 *v/v* Ac₂O/pyridine (5 mL 0.5 M in DMF). DNP was removed from the His with 2×30 min treatment with a mixture of DIPEA (1% *v/v*) and mercaptoethanol (4% *v/v*) in DMF (7 mL). Formyl groups were removed by treatment with a continuous flow of piperidine (250 mL; 20% *v/v* in DMF) in 8 min. The Boc groups were removed with a 2×1 min treatment with TFA. After a DMF, DCM, and a 1:1 *v/v* MeOH/DCM flow-wash, the resin was dried under vacuum. HF cleavage (4% *v/v* of *p*-cresol added as a scavenger) and subsequent lyophilization gave the acetylated crude product. The acetyl-protected glycopeptide was dissolved in 10% MeCN in H₂O (1 mg/1 mL) and treated with hydrazine hydrate (10% *v/v* in H₂O) for 2 h. The fully deprotected GlcNAc-T-C34 **16** was analyzed by RP HPLC on a C18 column (0.5 $\times$ 15 cm) by elution with a linear gradient of 0–60% CH₃CN containing 0.1% TFA (Method A) in 30 min, flow rate 1 mL/min (*T_R* = 14.7 min). Preparative RP HPLC was performed over a C18 column (2.5 $\times$ 20 cm, 10 mL/min), by elution with a linear gradient of 0–60% CH₃CN containing 0.1% TFA in 90 minutes. The product was analyzed by HRMS (ESI): calculated for C₁₉₅H₂₉₉N₅₄O₆₈S [*M*+3*H*]³⁺ 4517.13193; found 4517.11479 (based on deconvolution of data).

Synthesis of glycopeptide 9 and 10 by enzymatic transglycosylation

A mixture of the Man₃GlcNAc-oxazoline **8** (290 nmol) and the C-linked Glc-dipeptide **4** (100 nmol) or the N-linked Glc-dipeptide **7** (100 nmol) in a phosphate buffer (40 μ L, pH 7.0, 50 mM) was incubated at 23 °C with the enzyme Endo A (5 mU). The reaction was monitored by analytical HPLC on a Waters Nova-Pak C18 column (3.9 $\times$ 150mm) at 40 °C eluted by method B (a linear gradient of 0–30% MeCN containing 0.1% TFA in 18 min, flow rate 1 mL/min) for the reaction with Glc-dipeptide **4** or method C (isocratic elution with 100% water containing 0.1% TFA in 10 min then a linear gradient of 0–70% MeCN containing 0.1% TFA in 20 min, flow rate 1 mL/min) for the reaction with Glc-dipeptide **7**. The Glc-dipeptide was converted to a new species that was eluted slightly earlier than the starting material. The enzymatic reaction was stopped by heating in a boiling water bath for 3 min. The product was purified by preparative HPLC on a Waters preparative column (Symmetry 300, 19 $\times$ 300 mm) to afford the transglycosylation product.

(*L*-T1M[4-(Man₃GlcNAcGlc)]-*L*-Phe-NH₂) (**9**), 52% yield; *t_R* = 11.41 min (method B); ESI-MS calculated for C₄₆H₇₂N₇O₂₇, *M* = 1154.45; found, 1155.67 [*M*+*H*]⁺.

(*L*-Pro-*L*-T4M[4-(Man₃GlcNAcGlc)]-OH) (**10**): 80% yield; *t_R* = 15.41 min (method C); ESI-MS calculated for C₄₂H₆₈N₆O₂₈, *M* = 1104; found, 1105 [*M*+*H*]⁺.

Synthesis of C34 glycopeptides 17 and 18 by enzymatic transglycosylation

A mixture of the Man₃GlcNAc-oxazoline **8** (290 nmol) and the GlcNAc-C34 (45 nmol) or GlcNAc-T-C34 (45 nmol) in a phosphate buffer (40 μ L, pH 7.0, 50 mM) was incubated at 23 °C with the enzyme Endo A (10 mU). The reaction was monitored by analytical HPLC on a Waters Nova-Pak C18 column (3.9 $\times$ 150mm) at 40 °C eluted by method D (a linear gradient of 0–90% MeCN containing 0.1% TFA in 18 min, flow rate 1 mL/min). After 1h, the residue was subject to preparative HPLC on a Waters preparative column (Symmetry 300, 19 $\times$ 300 mm) to afford the transglycosylation product.

Man₃GlcNAc₂-C34 (**17**): 95% yield; *t_R* = 19.60 min (method D); ESI-MS, calculated for C₂₂₀H₃₄₂N₅₃O₈₉S, *M* = 5182.36; found, 5182.97 (deconvolution by MaxEnt).

Man₃GlcNAc₂-T-C34 (**18**): 95% yield; t_R = 19.94 min (method D); ESI-MS, calculated for C₂₂₁H₃₄₂N₅₅O₈₈S, M = 5206.37; found, 5207.26 (deconvolution by MaxEnt).

Infectivity Assays

The antiviral activities of the synthetic C34 peptides and glycopeptides were assayed by inhibiting the infection of TZM-bl cells with HIV-1 IIIB. TZM-bl is a cell line expressing CD4 and viral coreceptors (CCR5 and CXCR4) and containing integrated copies of the luciferase gene under control of the HIV-1 promoter. This cell line was obtained from the NIH AIDS Repository (Germantown, Maryland). TZM-bl cells were cultured in DMEM medium supplemented with 10 % heat inactivated fetal bovine serum and antibiotics. Tissue culture reagents were purchased from Invitrogen (Carlsbad, California). For infection, 2×10^4 cells were plated per well in triplicate wells of 96-well tissue culture plates one day before infection. Cells were then infected with virus using a multiplicity of infection (m.o.i) of 0.001 in the absence (control) and presence of 3-fold dilutions of each peptide (range from 0.1–100 nM). Infected cells were incubated at 37°C with 5 % CO₂ in a humidified incubator for 3 days. Infection was measured by determining luciferase activity in cell lysates using a luciferase detection kit (Promega, Madison, Wisconsin) and following the manufacturer's directions. Mock-infected wells were used to determine background luminescence, which was subtracted from the sample wells. Viral infectivity was determined by dividing luciferase units at each peptide concentration by the luciferase units obtained in the control wells containing no peptide.

PNGase digestion

Glycosylated C34 (**17** or **18**) (30 µg) in a Tris-Cl buffer (20 µL, pH 7.5, 20 mM) was incubated at 37 °C with PNGase F (6 U). The digestion was monitored by analytic HPLC under the conditions as described in the enzymatic transglycosylation reactions (method D). The digestion rates were calculated based on the integration of C34 peptides and their digestion fragments (characterized by mass spectrum) in HPLC profiles.

Trypsin and chymotrypsin digestion

C34 (6 nmol) or glycosylated C34 peptides (**17** or **18**, 6 nmol) in a Tris-Cl buffer (200 µL, pH 8.0, 25 mM) was incubated with trypsin (0.5 µg) or chymotrypsin (0.5 µg) at 23 °C. The digestion was monitored by analytic HPLC under the conditions as described in the enzymatic transglycosylation reactions (method D). The digestion rates were calculated based on the integration of C34 peptides and their digestion fragments (characterized by mass spectrum) in HPLC profiles.

Acknowledgments

Dr. Tilman Hackeng (Cardiovascular Research Institute Maastricht Department of Biochemistry, University Maastricht) is kindly acknowledged for his contribution in the solid phase peptide synthesis. SenterNovem (Ministry of Economic Affairs, The Netherlands) is gratefully acknowledged for providing financial support. The work is supported in part by the National Institutes of Health (NIH grant R01 AI067111 to LXW).

References

1. Chan DC, Kim PS. Cell 1998;93:681–684. [PubMed: 9630213]
2. Liu S, Wu S, Jiang S. Curr. Pharm. Des 2007;13:143–162. [PubMed: 17269924]
3. Robertson D. Nat. Biotechnol 2003;21:470–471. [PubMed: 12721558]
4. Fletcher CV. Lancet 2003;361:1577–1578. [PubMed: 12747873]
5. Chan DC, Fass D, Berger JM, Kim PS. Cell 1997;89:263–273. [PubMed: 9108481]
6. Ernst JT, Kutzki O, Debnath AK, Jiang S, Lu H, Hamilton AD. Angew. Chem. Int. Ed 2002;41:278–281.

7. Wang LX, Song H, Liu S, Lu H, Jiang S, Ni J, Li H. *ChemBioChem* 2005;6:1068–107. [PubMed: 15883971]
8. Seitz O. *ChemBioChem* 2000;1:214–246. [PubMed: 11828414]
9. Herzner H, Reipen T, Schultz M, Kunz H. *Chem. Rev* 2000;100:4495–4538. [PubMed: 11749356]
10. Unverzagt, C.; Eller, S.; Mezzato, S.; Schubert, R. 2008. p. 1304–1311. references cited herein
11. a Yamamoto K, Kadowaki S, Watanabe J, Kumagai H. *J. Biochem. Biophys. Res. Commun* 1994;203:244–252. b Haneda K, Inazu T, Yamamoto K, Kumagai H, Nakahara Y, Kobata A. *Carbohydr. Res* 1996;292:61–70. [PubMed: 8870237] c Mizuno M, Haneda K, Iguchi R, Muramoto I, Kawakami T, Aimoto S, Yamamoto K, Inazu T. *J. Am. Chem. Soc* 1999;121:284–290. d Wang LX. *Carbohydr. Res* 2008;343:1509–1522. [PubMed: 18405887]
12. Deras IL, Takegawa K, Kondo A, Kato I, Lee YC. *Bioorg. Med. Chem. Lett* 1998;8:1763–1766. [PubMed: 9873430]
13. Wang LX, Tang M, Suzuki T, Kitajima K, Inoue Y, Inoue S, Fan JQ, Lee YC. *J. Am. Chem. Soc* 1997;119:11137–11146.
14. Wang LX, Fan JQ, Lee YC. *Tetrahedron Lett* 1996;37:1975–1978.
15. Huang W, Ochiai H, Zhang X, Wang LX. *Carbohydr. Res* 2008;343:2903–2913. [PubMed: 18805520]
16. Horne WS, Yadav MK, Stout CD, Ghadiri MR. *J. Am. Chem. Soc* 2004;126:15366–15367. [PubMed: 15563148]
17. van Maarseveen JH, Horne WS, Ghadiri MR. *Org Lett* 2005;7:4503–4506. [PubMed: 16178569]
18. Bock VD, Speijer D, Hiemstra H, van Maarseveen JH. *Org. Biomol. Chem* 2007;5:971–975. [PubMed: 17340013]
19. a Tornøe CW, Christensen C, Meldal M. *J. Org. Chem* 2002;67:3057–3064. [PubMed: 11975567] b Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. *Angew. Chem. Int. Ed* 2002;41:2596–2599.
20. a Groothuys S, Kuijpers BHM, Quaedflieg PJLM, Roelen HCPF, Wiertz RW, Blaaup RH, van Delft FL, Rutjes FPJT. *Synthesis* 2006;3146–3152. b Kuijpers BHM, Groothuys S, Hawner C, ten Dam J, Quaedflieg PJLM, Schoemaker HE, van Delft FL, Rutjes FPJT. *Org. Proc. Res. Dev* 2008;12:503–511.
21. Kuijpers BH, Groothuys S, Keereweer AB, Quaedflieg PJ, Blaaup RH, van Delft FL, Rutjes FPJT. *Org Lett* 2004;6:3123–3126. [PubMed: 15330603]
- 22a. Li B, Zeng Y, Hauser S, Song H, Wang LX. *J. Am. Chem. Soc* 2005;127:9692–9693. [PubMed: 15998066]
- 22b. Li B, Song H, Hauser S, Wang LX. *Org. Lett* 2006;8:3081–3084. [PubMed: 16805557]
23. Fan JQ, Lee YC. *J. Biol. Chem* 1997;272:27058–27064. [PubMed: 9341145]
24. van Esseveldt, van Delft FL, Smits JMM, de Gelder R, Schoemaker HE, Rutjes FPJT. *Adv. Synth. Cat* 2004;346:823–834.
25. a Paulsen H, Gyorgyde Z, Friedman M. *Chem. Ber* 1974;107:1590–1613. b Bianchi A, Bernardi A. *J. Org. Chem* 2006;71:4565–4577. [PubMed: 16749790]
26. Schnölzer M, Alewood P, Jones A, Alewood D, Kent SBH. *Int. J. Pept. Prot. Res* 1992;40:180–193.

Scheme 1.
Synthesis of triazole-linked Glc-dipeptides.

Figure 1. Formation of glycopeptides **9** and **10**. The molar ratio of donor to acceptor was 3:1 for the reactions.

Scheme 3.

Attempted incorporation of triazole-linked GlcNAc by on-resin copper-catalyzed cycloaddition.

Scheme 4.

Incorporation of a triazole-linked GlcNAc moiety into C34 by using a preassembled building block **15**.

Scheme 5.

Transglycosylation to form C34 glycopeptides with natural amide and triazole linkages.

Figure 2.
ESI-MS of synthetic Man₃GlcNAc₂-C34 (**17**) (A) and Man₃GlcNAc₂-TC34 (**18**) (B)

Figure 3.
Anti-HIV activity of C34 and glycosylated C34 variants

Figure 4.
PNGase stability of Man₃GlcNAc₂-C34 (**17**) and Man₃GlcNAc₂-T-C34 (**18**)

Figure 5.
Potential protease cleavage sites in peptide C34.

Figure 6.
The digestion of C34 peptide and glycopeptides by trypsin (A) and chymotrypsin (B).