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 Analogues

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]

Endoglycosidase-catalyzed transglycosylation of triazole-linked glucose (Glc) and *N*-acetylglucosamine (GlcNAc)-containing dipeptides and polypeptides was achieved by 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; this demonstrates 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 the HIV-1 envelope glycoprotein, gp41. We found that the synthetic C34 glycopeptide possesses potent anti-HIV activity with an IC_{50} of 21 nM. The triazole-linked C34 glycopeptide demonstrated a much enhanced stability against protease- and glycoamidase-catalyzed digestion; this shows the protective effects of glycosylation and the stability of the triazole linkage. These favorable properties suggest that the triazole-linked C34 glycopeptide might be valuable for further development as an anti-HIV drug candidate.

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 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 half-life because of their susceptibility to protease digestion. It is well known that glycosylation can significantly influence peptide pharmacokinetic properties, such as its protease susceptibility.

We are interested in investigating the influence of glycosylation on the activity and stability of peptide C34—a 34-mer peptide derived from the C ectodomain of gp41—which 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 half-life is particularly important for this class of anti-HIV drugs, as prolonged therapeutic efficacy means the need for a relatively low dose with a lower frequency of intravenous injection. In this paper, we describe the 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

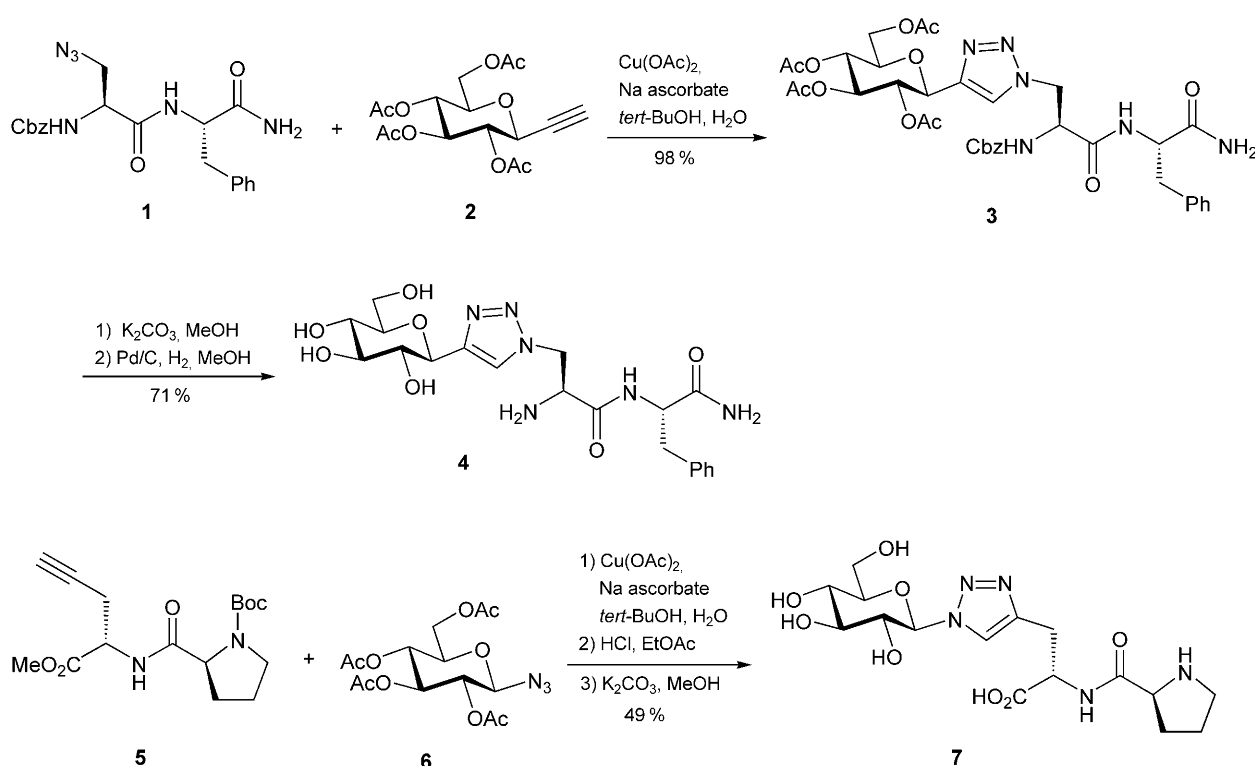
[a] W. Huang, A. Heredia, Dr. L.-X. Wang
Institute of Human Virology, University of Maryland
School of Medicine, Baltimore, MD 21201 (USA)
Fax: (+1) 410-706-5068
E-mail: lwang3@ihv.umaryland.edu

[b] S. Groothuys, B. H. M. Kuijpers, Dr. F. P. J. T. Rutjes, Dr. F. L. van Delft
Institute for Molecules and Materials, Radboud University Nijmegen
Heyendaalseweg 135, 6525 AJ Nijmegen (The Netherlands)
Fax: (+31) 24-365-3393
E-mail: f.vandelft@science.ru.nl

 WILEY InterScience DISCOVER SOMETHING GREAT
© 2009 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

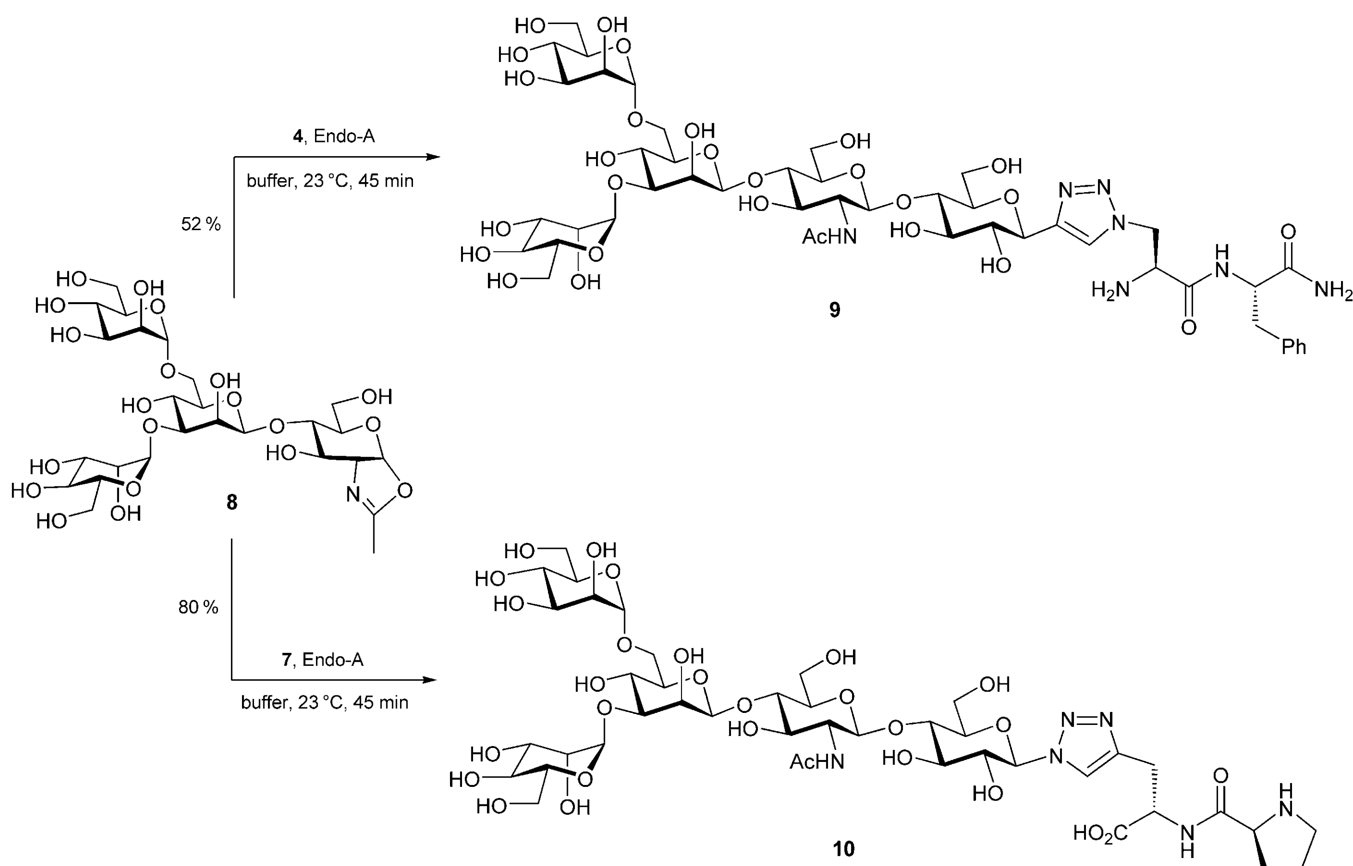
single step without the need of any protecting groups.[11] Previous studies have demonstrated that in a peptide context the *endo*-β-*N*-acetylglucosaminidase from *Arthrobacter* (Endo-A) can recognize both a GlcNAc and a glucose (Glc) moiety as an acceptor substrate for transglycosylation.[11,12] It 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 can tolerate either a triazole-linked Glc or a 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. Triazole is of particular interest as a bridging moiety for either a *N*-glycopeptide or *N*-glycoprotein, since triazoles have been proposed and corroborated to be promising amide bioisosteres.[16–18] Moreover, such an *N*-glycopeptide analogue is expected to be resistant to glycoamidase-catalyzed deglycosylation. 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 reports.[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**, in which glucose is 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**, by following our previous procedure.[21] The copper-catalyzed 1,3-dipolar cycloaddition gave the desired product **7** in 49% yield after deprotection (Scheme 1).

With two complementary glycodipeptides at hand, the transglycosylation under the action of Endo-A was ready to be tested. The tetrasaccharide oxazoline, Man_3GlcNAc–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, we found that both 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 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



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

 © 2009 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 2. Transglycosylation with triazole-linked Glc–dipeptides.

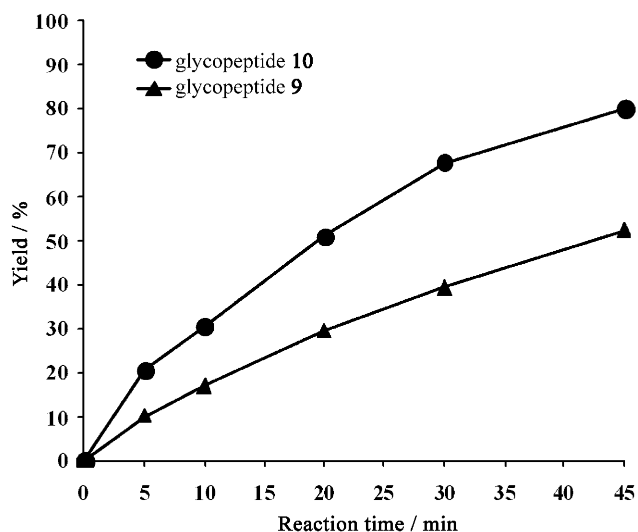


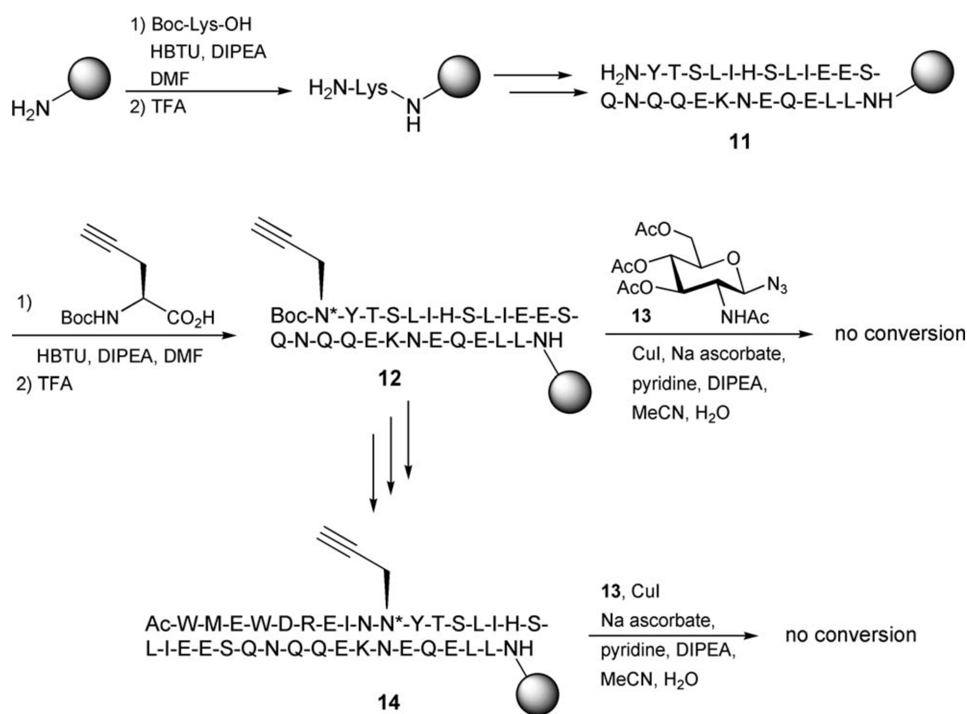
Figure 1. Formation of glycopeptides **9** and **10**. The molar ratio of donor to acceptor was 3:1 for the reactions.

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 for this observed difference in reactivity is that the N-linked triazole resembles the amide linkage found in natural *N*-glycopeptides more closely. Furthermore, this finding has not been described previously and could 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 accepted a glucose moiety as the acceptor, but it also tolerated the modification at the anomeric position with a triazole heterocycle; this implies that Endo-A can be a flexible enzyme in organic synthesis. To attest the broad substrate specificity, we have recently demonstrated that Endo-A can 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 glycopeptide 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 a *N*-triazole linkage. Two principal strategies were attempted for synthesis of glycol–C34, that is, either 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 success-

 © 2009 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 3. Attempted incorporation of triazole-linked GlcNAc by on-resin copper-catalyzed cycloaddition.

fully applied in the preparation of compounds **4** and **7**, and attachment of a sugar postsolid-phase synthesis is more modular in nature, it was investigated first (Scheme 3). Solid phase peptide synthesis was performed by using standard Boc chemistry.[23] A propargylglycine moiety was incorporated into the C34 peptide sequence at the conserved glycosylation site (N637) by 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**).

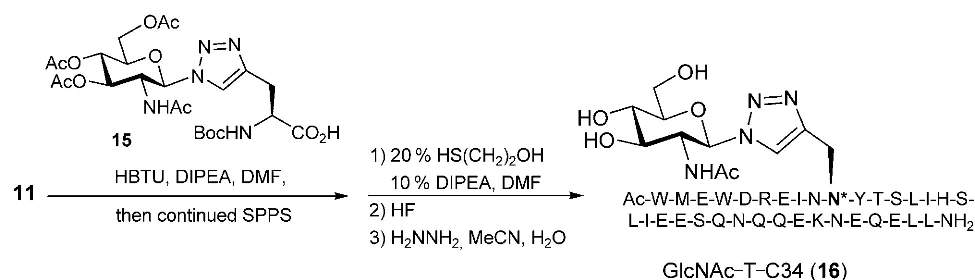
In contrast, the desired triazole-linked GlcNAc–C34, GlcNAc–T–C34 (**16**), was smoothly obtained by incorporation of the glycoamino acid **15**, which was readily derived in two steps from Boc-L-propargylglycine and 2-acetamido-2-deoxy-glucosyl azide (**13**), into the solid phase peptide synthesis 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 reversed-phase HPLC and characterized by ESI-MS.

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

The suitability of GlcNAc–T–C34 (**16**) as the acceptor substrate for Endo-A for the transglycosylation with $Man_3GlcNAc$–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, and the corresponding transglycosylation products $Man_3GlcNAc_2$–C34 (**17**) and $Man_3GlcNAc_2$–T–C34 (**18**), respectively, were obtained (Scheme 5). The products were purified by HPLC and their identities were characterized by ESI-MS (Figure 2).

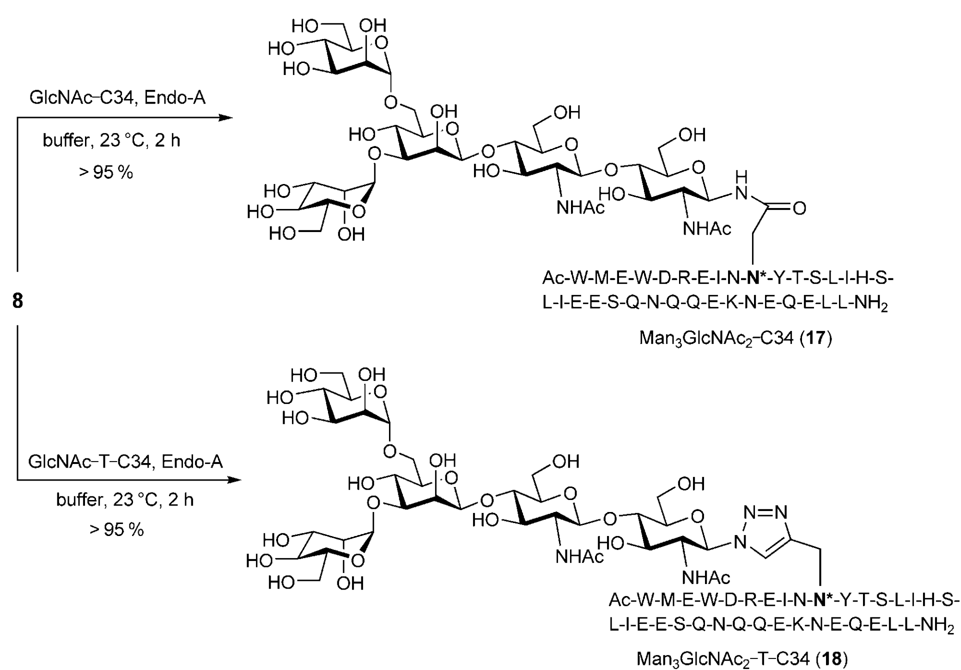
Anti-HIV activity of the synthetic C34 glycopeptides

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 contains integrated copies of the luciferase gene under control of the HIV-1 promoter. All the glycoforms of C34 demonstrated potent inhibitory activities against HIV-1 infection at *nanomolar* concentrations. Results from a typical inhibition assay experiment are demonstrated in Figure 3. The estimated IC_{50} data for the C34 peptide and glycopeptides are: C34: 7.0 nM; GlcNAc–C34: 14.3 nM; GlcNAc–T–C34: 4.5 nM; $Man_3GlcNAc_2$–C34: 16.4 nM, and $Man_3GlcNAc_2$–T–C34: 21.0 nM. It was found that attachment of a sugar moiety resulted in a two- to threefold decrease in the inhibitory activity based on the current cell-based infectivity assay. The results are consistent with previous observations on the antiviral activity of C34 glycopeptides based on a fusion assay.[7] Whereas the anti-HIV ac-



Scheme 4. Incorporation of a triazole-linked GlcNAc moiety into C34 by using a preassembled building block, **15**.

 © 2009 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 5. Transglycosylation to form C34 glycopeptides with natural amide and triazole linkages.

during binding. Nevertheless, from the perspective of anti-HIV drug development, the glycopeptides might be superior to C34 in two aspects. First, the two glycosylated C34 peptides, $Man_3GlcNAc_2$–C34 and $Man_3GlcNAc_2$–T–C34, showed much better water solubility under physiological conditions than C34; this 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 might be more resistant to protease digestion in vivo than the nonglycosylated C34. This effect was demonstrated by protease digestion experiments.

tivity 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 unmodified C34 (IC_{50} 4.5 vs. 7.0 nM). These results suggest that attachment of a small monosaccharide residue at the glycosylation site does not have a significant impact on the anti-HIV activity. The slight variation in the in vitro anti-HIV activity might be the result of the different types of linkers used in the two compounds. However, as expected, the attachment of a larger oligosaccharide moiety, such as $Man_3GlcNAc_2$, led to about threefold decrease in the anti-HIV activity, which is presumably caused by steric hindrance of the sugar moiety

Glycoamidase stability of triazole-linked glycosylated C34 versus amide-linked glycol–C34

Peptide-N^4-(*N*-acetyl-β-D-glucosaminyl)asparagine amidases (PNGases), such as PNGase F, are a class of glycoamidases that cleave the β-aspartylglucosamine bond of asparagine-linked glycopeptides or glycoproteins, and thereby convert 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

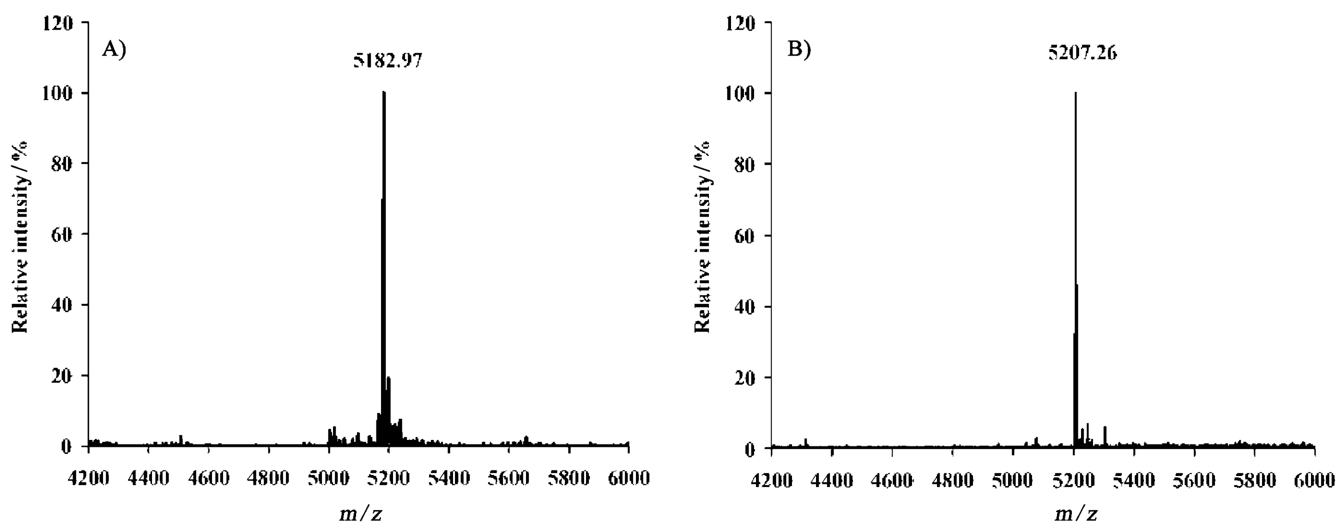


Figure 2. ESI-MS of synthetic A) $Man_3GlcNAc_2$–C34 (**17**) and B) $Man_3GlcNAc_2$–T–C34 (**18**).

 © 2009 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

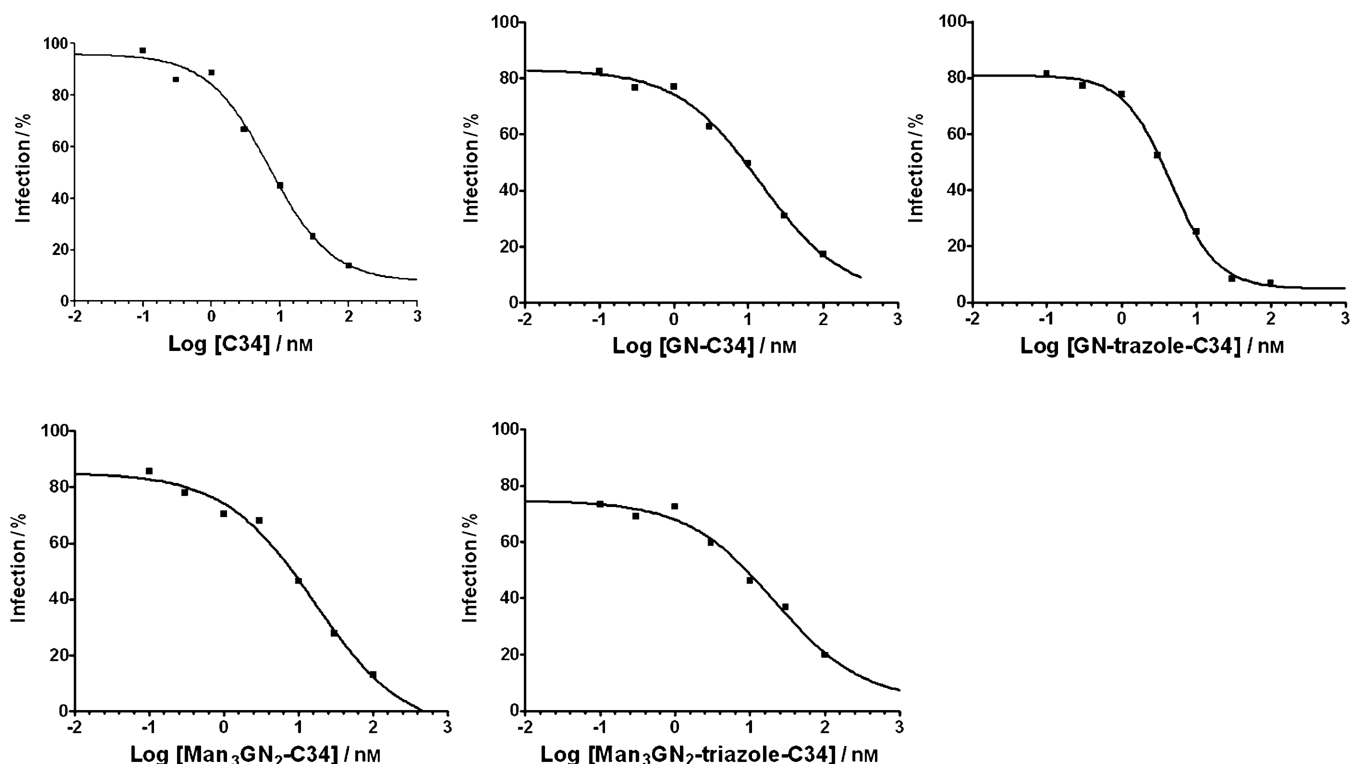


Figure 3. Anti-HIV activity of C34 and glycosylated C34 variants; GN denotes GlcNAc.

must consist of at least the *N,N'*-diacetylchitobiose core, GlcNAc-(β1→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 h by the glycoamidase under the same conditions (Figure 4). These results suggest that the triazole-linked glycol–C34 cannot be metabolized in vivo by glycoamidases. It should be pointed out that although it is known that PNGases exist in the cytosol, it is not clear how significant their activity is in serum.

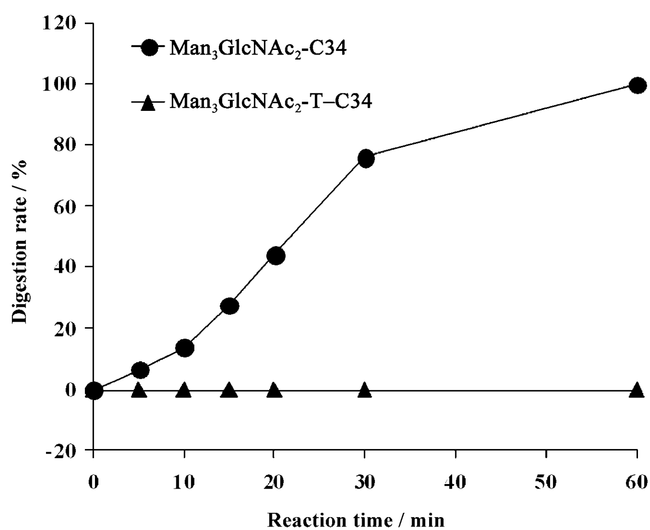


Figure 4. PNGase stability of $Man_3GlcNAc_2$–C34 (**17**) and $Man_3GlcNAc_2$–T–C34 (**18**).

Protease stability of $Man_3GlcNAc_2$–C34 and $Man_3GlcNAc_2$–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, that is, 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 was interesting to test whether the glycan attached at the N637 glycosylation site (N25 in the 34-mer peptide) confers protection against protease digestion.

First, it was observed that the native, nonglycosylated 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 the 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 nonglycosylated peptide C34, whereas under the same conditions, treatment of $Man_3GlcNAc_2$–C34 and $Man_3GlcNAc_2$–T–C34 led to hydrolysis of about 60 and 50% of the corresponding glycosylated C34, respectively; this shows the protective effect of the attached glycan. In the case of chy-

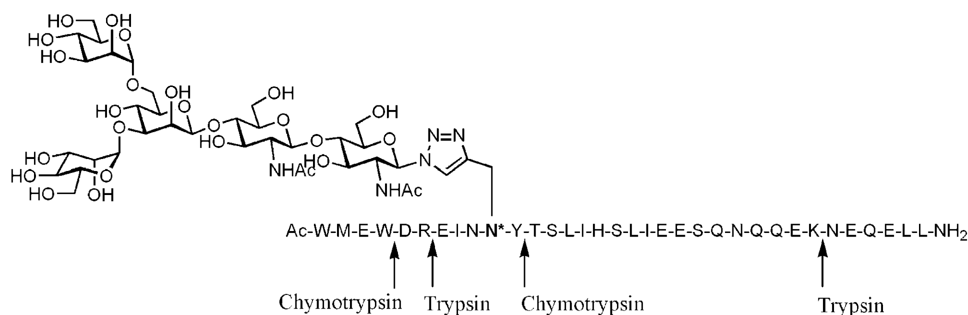


Figure 5. Potential protease cleavage sites in peptide C34.

motrypsin, 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 was observed by attachment of a single sugar. However, under the same digestion conditions, only about 10 and 20% of $Man_3GlcNAc_2$–C34 and $Man_3GlcNAc_2$–T–C34, respectively, were hydrolyzed after 4 h. These results clearly imply that the glycosylated C34 carrying a larger *N*-glycan could be much more stable in vivo than the nonglycosylated C34. The enhanced protease stability, together with the potent anti-HIV activity and enhanced water solubility, suggests that glycosylated C34 might be a valuable candidate for further development as an anti-HIV drug.

Conclusions

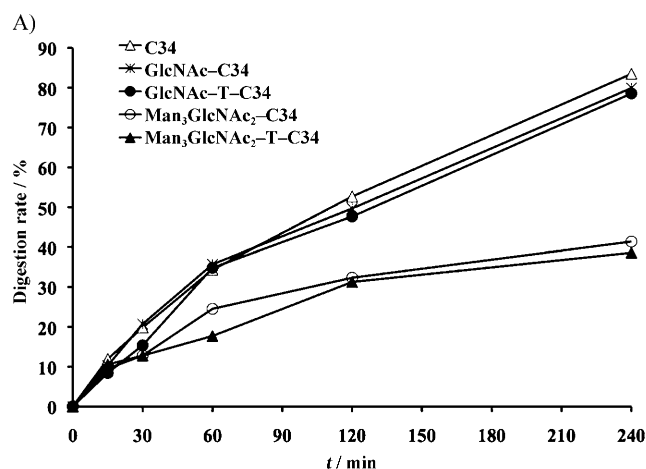
A novel triazole-linked glycopeptide C34 was efficiently synthesized by a chemoenzymatic approach. It was found that the synthetic glycopeptide possesses potent anti-HIV activity, and demonstrates dramatically enhanced water solubility and protease stability in comparison with the nonglycosylated peptide C34. These findings reveal favorable properties for the glycosylated C34, which could 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.[27] The pGEX-2T–Endo-A plasmid used for over-expression was kindly provided by Prof. Kaoru Takegawa (Kagawa University, Japan). The peptide *N*-glycosidase F (PNGase F) was purchased from New England Biolabs, Inc. Peptide C34 and GlcNAc–C34 were synthesized on a solid-phase peptide synthesizer by using the Fmoc approach, as previously reported.[7] All other reagents were purchased from Sigma–Aldrich and were used as received.

Preparation of L-T1M[4-(β-D-Glc)]-L-Phe-NH$_2$ (4): Cbz-L-T1M{4-[β-D-Glc(Ac)$_4$]}-L-Phe-NH$_2$ (**3**; 73 mg, 0.095 mmol)[20] was dissolved in MeOH (1 mL), and K_2CO_3 (2 mg, 9 μmol) was added. The reaction was stirred for 1.5 h at room temperature. The crude mixture was neutralized with Amberlite IR120 (prewashed with MeOH), filtered and the solvent evaporated 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 H_2. The suspension was filtered and the solvent was evaporated. The product was lyophilized from AcOH (1 M in H_2O) to yield the desired product **4** (35 mg, 71%). ^{1}H NMR (400 MHz, D_2O) $\delta = 8.04$ (s, 1H, triazole-H), 7.51–7.31 (m, 5H, H-arom), 4.72–4.50 (m, 4H), 4.06–3.90 (m, 2H), 3.84–3.74 (m, 2H), 3.72–3.55 (m, 3H), 3.20 (dd, $J = 13.9$, 6.5 Hz, 1H, CH_{2a}, Phe), 3.06 (dd, $J = 13.9$, 8.4 Hz, 1H, CH_{2b}, Phe). ^{13}C NMR (75 MHz, D_2O): $\delta =$ 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) calcd for $C_{20}H_{29}N_6O_7$ $[M+H]^+$: 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 described earlier.[21] Next, the Boc-L-Pro-L-T4M{1-[β-D-Glc(Ac)$_4$]}-OH (40 mg, 0.06 mmol) was dissolved in HCl (2.6 M) in EtOAc (2 mL) and stirred for 30 min; the solvent was then evaporated in vacuo. The crude product was dissolved in MeOH (3 mL), a catalytic amount of K_2CO_3 was added



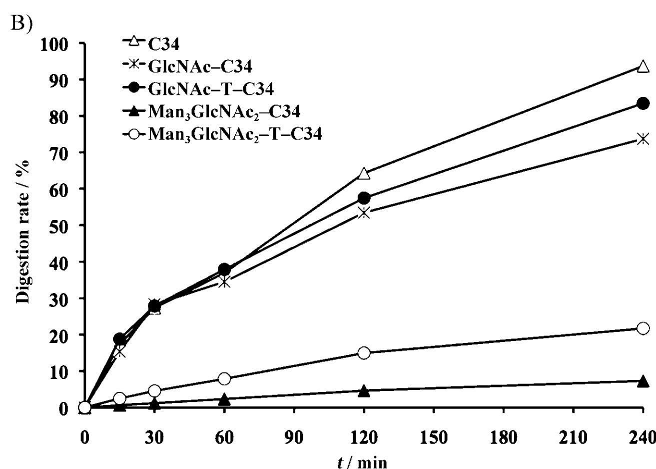


Figure 6. The digestion of C34 peptide and glycopeptides by A) trypsin and B) chymotrypsin.

 © 2009 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

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. ^{1}H NMR (400 MHz, D_2O): $\delta = 7.97$ (s, 1H, triazole-H), 5.58 (d, $J = 9.1$ Hz, 1H, H-1), 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). ^{13}C NMR (75 MHz, D_2O): $\delta = 172.9$, 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* calcd for $C_{16}H_{26}N_5O_8$ $[M+H]^+$: 416.1781; found: 416.1791.

Boc-L-T4M(1-[β-D-GlcNAc(Ac)$_3$])-OH (15): $Cu(OAc)_2$ (0.10 g, 0.50 mmol) and sodium ascorbate (0.20 g, 1.0 mmol) in H_2O (5 mL) were added to a solution of GlcNAc(Ac)$_3$N$_3$ (**13**,[25] 0.93 g, 2.5 mmol) and Boc-L-propargylglycine[24] (0.53 g, 2.5 mmol) in *tert*-butanol (5 mL). The reaction was stirred, overnight, water was added and the product was extracted with CH_2Cl_2. The combined organic layers were washed with aqueous NaCl, dried over Na_2SO_4 and evaporated in vacuo. The crude product was purified by flash chromatography by using EtOAc/heptane (3:1) to give **15** (1.25 g, 2.14 mmol, 86%) as a white solid. FTIR (ATR): $\nu = 1744$, 1364, 1213 cm^{-1}. ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 7.75$ (s, 1H, triazole-H), 6.48 (brd, $J = 9.6$ Hz, 1H, NH), 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). ^{13}C NMR (75 MHz, $CDCl_3$): $\delta = 171.6$, 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* calcd for $C_{24}H_{35}O_{12}N_5Na$ $[M+Na]^+$: 608.21799; found: 608.21924.

GlcNAc–T–C34 (16): The peptide was obtained by manual solid phase peptide synthesis (SPPS) by 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*) Ac_2O/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) for 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*, *p*-cresol added as a scavenger) and subsequent lyophilization gave the acetylated crude product. The acetyl-protected glycopeptide was dissolved in MeCN (10%) in H_2O (1 mg per 1 mL) and treated with hydrazine hydrate (10%, *v/v*, in H_2O) for 2 h. The fully deprotected GlcNAc–T–C34 (**16**) was analyzed by RP-HPLC on a C18 column (0.5×15 cm) by elution with a linear gradient of 0–60% CH_3CN containing TFA (0.1%; method A) for 30 min, flow rate 1 mL min^{-1} ($t_R = 14.7$ min). Preparative RP-HPLC was performed over a C18 column (2.5×20 cm, 10 mL min^{-1}) by elution with a linear gradient of 0–60% CH_3CN containing TFA (0.1%) in 90 min. The product was analyzed by HRMS (ESI): calcd for $C_{195}H_{299}N_{54}O_{68}S$ $[M+3H]^{3+}$: 4517.13193; found: 4517.11479 (based on deconvolution of data).

Synthesis of glycopeptides 9 and 10 by enzymatic transglycosylation: A mixture of Man$_3$GlcNAc–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 µL, 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 (3.9×150 mm) at 40 °C and eluted by using method B (a linear gradient of 0–30% MeCN containing 0.1% TFA in 18 min, flow rate 1 mL min^{-1}) 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^{-1}) for the reaction with Glc–dipeptide **7**. The Glc–dipeptide was converted to a new species that 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, 19×300 mm) to afford the transglycosylation product.

L-T1M[4-(Man$_3$GlcNAcGlc)]-L-Phe-NH$_2$ (**9**): 52% yield; $t_R = 11.41$ min (method B); ESI-MS calcd for $C_{46}H_{72}N_7O_{27}$, $M = 1154.45$; found: 1155.67 $[M+H]^+$.

L-Pro-L-T4M[4-(Man$_3$GlcNAcGlc)]-OH (**10**): 80% yield; $t_R = 15.41$ min (method C); ESI-MS calcd for $C_{42}H_{68}N_6O_{28}$, $M = 1104$; found: 1105 $[M+H]^+$.

Synthesis of C34 glycopeptides 17 and 18 by enzymatic transglycosylation: A mixture of the Man$_3$GlcNAc–oxazoline (**8**; 290 nmol) and GlcNAc–C34 (45 nmol) or GlcNAc–T–C34 (45 nmol) in a phosphate buffer (40 µL, 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 (3.9×150 mm) at 40 °C and eluted by using method D (a linear gradient of 0–90% MeCN containing 0.1% TFA in 18 min, flow rate 1 mL min^{-1}). After 1 h, the residue was subject to preparative HPLC on a Waters preparative column (Symmetry 300, 19×300 mm) to afford the transglycosylation product.

Man$_3$GlcNAc$_2$–C34 (**17**): 95% yield; $t_R = 19.60$ min (method D); ESI-MS calcd for $C_{220}H_{342}N_{53}O_{89}S$, $M = 5182.36$; found: 5182.97 (deconvolution by MaxEnt).

Man$_3$GlcNAc$_2$–T–C34 (**18**): 95% yield; $t_R = 19.94$ min (method D); ESI-MS calcd for $C_{221}H_{342}N_{55}O_{88}S$, $M = 5206.37$; 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 contains 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, USA). TZM-bl cells were cultured in DMEM medium supplemented with heat inactivated fetal bovine serum (10%) and antibiotics (100 units penicillin, 0.1 mg mL^{-1} streptomycin). Tissue culture reagents were purchased from Invitrogen (Carlsbad, California, USA). 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 by using a multiplicity of infection (m.o.i) of 0.001 in the absence (control) and presence of threefold dilutions of each peptide (range from 0.1–100 nM). Infected cells were incubated at 37 °C with 5% CO_2 in a humidified incubator for three days. Infection was measured by determining luciferase activity in cell lysates by using a luciferase detection kit (Promega, Madison–Wisconsin, USA) 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

 © 2009 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

same conditions as described for the enzymatic transglycosylation reactions (method D). The digestion rates were calculated based on the integration of C34 peptides and their digestion fragments (characterized by MS) 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) were incubated with trypsin (0.5 μg) or chymotrypsin (0.5 μg) at 23 °C. The digestion was monitored by analytic HPLC under the same conditions as described for the enzymatic transglycosylation reactions (method D). The digestion rates were calculated based on the integration of C34 peptides and their digestion fragments (characterized by MS) in HPLC profiles.

Acknowledgements

Dr. Tilman Hackeng (Cardiovascular Research Institute Maastricht Department of Biochemistry, University Maastricht) is kindly acknowledged for his contribution to 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 Institute of Health (NIH grant R01 AI067 111 to L.X.W.).

Keywords: antiviral agents · chemoenzymatic synthesis · glycoproteins · glycosylation · protease resistance

[1] D. C. Chan, P. S. Kim, *Cell* **1998**, *93*, 681–684.
[2] S. Liu, S. Wu, S. Jiang, *Curr. Pharm. Des.* **2007**, *13*, 143–162.
[3] D. Robertson, *Nat. Biotechnol.* **2003**, *21*, 470–471.
[4] C. V. Fletcher, *Lancet* **2003**, *361*, 1577–1578.
[5] D. C. Chan, D. Fass, J. M. Berger, P. S. Kim, *Cell* **1997**, *89*, 263–273.
[6] J. T. Ernst, O. Kutzki, A. K. Debnath, S. Jiang, H. Lu, A. D. Hamilton, *Angew. Chem.* **2002**, *114*, 288–291; *Angew. Chem. Int. Ed.* **2002**, *41*, 278–281.
[7] L. X. Wang, H. Song, S. Liu, H. Lu, S. Jiang, J. Ni, H. Li, *ChemBioChem* **2005**, *6*, 1068–1107.
[8] O. Seitz, *ChemBioChem* **2000**, *1*, 214–246.
[9] H. Herzner, T. Reipen, M. Schultz, H. Kunz, *Chem. Rev.* **2000**, *100*, 4495–4538.
[10] C. Unverzagt, S. Eller, S. Mezzato, R. Schuberth, *Chem. Eur. J.* **2008**, *14*, 1304–1311; and references cited therein.
[11] a) K. Yamamoto, S. Kadowaki, J. Watanabe, H. Kumagai, *J. Biochem. Biophys. Res. Commun.* **1994**, *203*, 244–252; b) K. Haneda, T. Inazu, K. Yamamoto, H. Kumagai, Y. Nakahara, A. Kobata, *Carbohydr. Res.* **1996**, *292*, 61–70; c) M. Mizuno, K. Haneda, R. Iguchi, I. Muramoto, T. Kawakami, S. Aimoto, K. Yamamoto, T. Inazu, *J. Am. Chem. Soc.* **1999**, *121*, 284–290; d) L. X. Wang, *Carbohydr. Res.* **2008**, *343*, 1509–1522.
[12] I. L. Deras, K. Takegawa, A. Kondo, I. Kato, Y. C. Lee, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1763–1766.
[13] L. X. Wang, M. Tang, T. Suzuki, K. Kitajima, Y. Inoue, S. Inoue, J. Q. Fan, Y. C. Lee, *J. Am. Chem. Soc.* **1997**, *119*, 11137–11146.
[14] L. X. Wang, J. Q. Fan, Y. C. Lee, *Tetrahedron Lett.* **1996**, *37*, 1975–1978.
[15] W. Huang, H. Ochiai, X. Zhang, L. X. Wang, *Carbohydr. Res.* **2008**, *343*, 2903–2913.
[16] W. S. Horne, M. K. Yadav, C. D. Stout, M. R. Ghadiri, *J. Am. Chem. Soc.* **2004**, *126*, 15366–15367.
[17] J. H. van Maarseveen, W. S. Horne, M. R. Ghadiri, *Org Lett* **2005**, *7*, 4503–4506.
[18] V. D. Bock, D. Speijer, H. Hiemstra, J. H. van Maarseveen, *Org. Biomol. Chem.* **2007**, *5*, 971–975.
[19] a) C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064; b) V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708–2711; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.
[20] a) S. Groothuys, B. H. M. Kuijpers, P. J. L. M. Quaedflieg, H. C. P. F. Roelen, R. W. Wiertz, R. H. Blaauw, F. L. van Delft, F. P. J. T. Rutjes, *Synthesis* **2006**, 3146–3152; b) B. H. M. Kuijpers, S. Groothuys, C. Hawner, J. ten Dam, P. J. L. M. Quaedflieg, H. E. Schoemaker, F. L. van Delft, F. P. J. T. Rutjes, *Org. Process Res. Dev.* **2008**, *12*, 503–511.
[21] B. H. Kuijpers, S. Groothuys, A. B. Keereweer, P. J. Quaedflieg, R. H. Blaauw, F. L. van Delft, F. P. J. T. Rutjes, *Org Lett.* **2004**, *6*, 3123–3126.
[22] a) B. Li, Y. Zeng, S. Hauser, H. Song, L. X. Wang, *J. Am. Chem. Soc.* **2005**, *127*, 9692–9693; b) B. Li, H. Song, S. Hauser, L. X. Wang, *Org. Lett.* **2006**, *8*, 3081–3084.
[23] J. Q. Fan, Y. C. Lee, *J. Biol. Chem.* **1997**, *272*, 27058–27064.
[24] B. C. J. van Esseveldt, F. L. van Delft, J. M. M. Smits, R. de Gelder, H. E. Schoemaker, F. P. J. T. Rutjes, *Adv. Synth. Catal.* **2004**, *346*, 823–834.
[25] a) H. Paulsen, Z. Györgydeák, M. Friedman, *Chem. Ber.* **1974**, *107*, 1590–1613; b) A. Bianchi, A. Bernardi, *J. Org. Chem.* **2006**, *71*, 4565–4577.
[26] M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pept. Protein Res.* **1992**, *40*, 180–193.
[27] K. Fujita, N. Tanaka, M. Sano, I. Kato, Y. Asada, K. Takekawa, *Biochem. Biophys. Res. Commun.* **2000**, *267*, 134–138.

Received: November 11, 2008
Published online on April 7, 2009

 © 2009 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim