Published in final edited form as:

Tetrahedron. 2016 October 6; 72(40): 6091–6098. doi:10.1016/j.tet.2016.07.062.

Synthesis of multivalent glycopeptide conjugates that mimic an HIV epitope

Jennifer K. Bailey[†], Dung N. Nguyen[†], Satoru Horiya, and Isaac J. Krauss^{*}

Department of Chemistry, Brandeis University, 415 South St. MS 015, Waltham, MA 02454-9110, USA

Abstract

Recently, we reported a directed evolution method which enabled us to discover sequences of glycopeptides that bind with picomolar affinity to HIV antibody 2G12 and are of interest as HIV vaccine candidates. In this manuscript, we describe the syntheses of several of these large (~11–12 kDa) glycopeptides by a combination of fast flow peptide synthesis and click chemistry. We also discuss the optimization of their attachment to carrier protein CRM197, affording antigenic and immunogenic conjugates ready for animal vaccination.

Graphical Abstract

Keywords

2G12; HIV; Directed Evolution; Glycopeptides; Solid-phase peptide synthesis

1. Introduction

10–30 % of HIV-infected individuals develop broadly neutralizing antibodies (bnAbs), which neutralize many strains of HIV^{1–3} and can protect against infection in animal models.^{4,5} These antibodies bind to conserved epitopes on HIV envelope proteins gp120 and gp41, and thus provide clues for vaccine design. In “epitope-focused” or “structure-based” vaccine design, a vaccine immunogen (including a protein, protein fragment or peptide) is designed to mimic closely the structure of the epitope targeted by a known protective antibody (Figure 1).^{6–8} If the immunogen successfully recreates this structure and presents it

^{*}Corresponding author. kraussi@brandeis.edu (I. Krauss).

[†]Equal contribution.

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

Supplementary data

Supplementary data associated with this article can be found online at

to the immune system, the vaccinated animal or person may develop antibodies against this epitope; in principle, these vaccine-elicited antibodies could be “bnAb-like”, i.e., exhibit specificity and protective activity similar to the known protective bnAb.

Recently, HIV bnAbs have increasingly been found to target carbohydrate cluster epitopes.^{9–11} In order to design structures that mimic these epitopes, we have prepared libraries in which carbohydrates are clustered in diverse ($\sim 10^{13}$) presentations on random backbones.^{12–15} The bnAb of interest – in our case 2G12¹⁶ – is used to recognize and isolate from the library those carbohydrate presentations which mimic the epitope. With mRNA-displayed glycopeptide libraries, we have been able to discover glycopeptides that are recognized by 2G12 with K_D s in the low nanomolar to picomolar range.¹⁴ Although these directed evolution experiments provide glycopeptide sequences that might have potential for vaccine development, the glycopeptides in these libraries are prepared from ribosomally translated peptides in picomole scale; thus, synthesis is necessary to produce the mg quantities of glycopeptides needed for vaccine studies in animals. Herein, we describe our efforts to synthesize these materials, incorporate them into carrier protein conjugates ready for vaccination, and verify the antigenic presentation of the glycopeptides in the context of the conjugates.

2. Results and Discussion

2.1. Fast-flow synthesis of alkynyl peptides

Our overall strategy for the synthesis of glycopeptides (Figure 2a,b) was to produce the desired alkynyl peptide precursors (**1a–4a**) by solid-phase peptide synthesis (SPPS) and then to use copper-assisted alkyne azide cycloaddition (CuAAAC)¹⁷ to “click” glycosylate with oligomannosyl azide **5**. Because our required alkynyl peptides were quite long (40 amino acids), we were not surprised to obtain very complex mixtures (Figure 2c) from our preliminary synthesis attempts using an automated batch synthesizer operating at room temperature. For the sake of rapid troubleshooting, we opted to utilize Pentelute’s thermally heated rapid flow-based peptide synthesis platform.¹⁸ In this method, deblock, wash or activated amino acid solutions travel through a heated reactor containing SPPS resin after first being pumped through a narrow stainless-steel coil that is also submerged in the heating bath. This flow setup results in rapid heating of reagents seconds prior to contact with the resin, and ensures a continuous and fresh supply of reagents that have been exposed only to brief heating throughout the coupling reaction. Coupling times in our setup are 30 seconds and the entire SPPS cycle is a few minutes, with synthesis yields comparable to microwave synthesis. In our hands, major advantages of this method have been 1) the ability to rapidly test the conditions for long peptide syntheses and thus make corrections on a practical timeframe, and 2) the fact that the only major equipment needed is an HPLC pump, rather than a peptide synthesizer.

2.2. Adjustments to the C-terminal linker

For the purposes of conjugation to carrier proteins, we chose to synthesize our peptides with a flexible C-terminal linker, GSGSGC appended to the active sequence (Figure 2a); the cysteine sulfhydryl (–SH) would eventually couple to maleimided protein. Anticipating

that a C-terminal cysteine would be prone to epimerization during SPPS,¹⁹ we initially utilized sterically bulky trityl-protected cysteine (Scheme 1a) anchored to a chloro-trityl linker (on commercially pre-loaded NovaPEG^{®20} resin). Under the conditions of our deblock (60 °C, 20% piperidine), trial syntheses of a short sequence WSGSGC afforded a significant amount (~10%) of 3-(1-piperidinyl)-alanine byproduct **8**, resulting from elimination of sulfur from the C-terminal cysteine, followed by addition of piperidine.²¹ To avoid this problem altogether, we at first included an additional glycine residue at the C terminus before the cysteine residue. However, synthesis starting with commercial glycine-loaded resin turned out to be problematic. For instance, syntheses of model sequence FALFAG starting with Gly-HMPB-NovaPEG[®] resin afforded peptide impurities containing additional unwanted glycine residues (Scheme 1b). Analysis of glycine cleaved from the commercial resin revealed a significant amount (~10%) of di- and triglycine (data not shown). Although glycine is known to be prone to oligomerization during DMAP/DCC-mediated loading,²² it was nevertheless surprising to see this problem in a commercially available resin. Although this could in principle be addressed by in-house loading of C-terminal glycine and a judicious choice of loading chemistry, we opted to circumvent the problem by changing our C-terminal residue to alanine. After making this choice, we also determined with a test sequence, FGSGGCA, that it was necessary to use HBTU rather than HATU for glycine couplings, in order to avoid incorporation of extra glycines (not shown). As a precaution to avoid epimerization of the cysteine (and later histidine and tryptophan), the DIEA base concentration for these couplings was lowered from 0.95M to 0.29M.²³

2.3. Suppression of aspartimide formation

Having established an optimal linker design and synthesis protocol, we then proceeded to synthesize peptide sequence 10F2 (**1a**, Scheme 2), monitoring by periodic removal of a few beads from the reactor for cleavage and LC-MS analysis. Synthesis of mYALALFHRILGSGGCA proceeded smoothly using HBTU for glycine couplings and HATU for all other couplings. For coupling of *M* (homoproparglycine, HPG), we opted to conserve the very expensive Fmoc amino acid building block Fmoc-HPG-OH by removing the resin from the flow reactor and running the coupling in batch in a conical plastic tube heated to 60 °C. Although this batch coupling proceeded smoothly, further extension to QVTDMYALALFHRILGSGGCA in flow afforded product contaminated with significant amounts of aspartimide, which worsened as the synthesis progressed (**11b/12b**, Scheme 2 and Figure 3). This well-known side reaction occurs during the Fmoc deprotection steps with basic piperidine, and can be prevented by the use of specialized aspartate protecting groups, or by acidic additives to buffer the piperidine.^{24,25} Although N-hydroxylamine derivatives such as HOBt are known to be effective as buffers, we found it particularly economical to use formic acid, as reported by Mier.²⁶ Changing the deprotection solution from 80/20 DMF/piperidine to 80/19/1 DMF/piperidine/HCOOH for all deblock steps following aspartate incorporation, aspartimide formation was prevented without any obvious decrease in Fmoc deprotection efficiency.

2.4. Completion of the alkynyl peptides and click glycosylation

Incorporation of the remaining eighteen N-terminal residues to complete the 10F2 sequence (**1a**) proceeded without major problems. N-terminal formylation with 4-nitrophenylformate did not proceed to completion when run in flow in a manner analogous to our standard amino acid couplings, but was complete after three treatments with this reagent in batch at 60 °C. To retain cysteine protection in the final peptide and avoid oxidative dimerization during purification and handling, the synthesis was repeated with the cysteine protecting group changed from Trt (**b** series) to the non-acid-labile tBuS disulfide (**a** series). Cleavage with 87.5/5/5/2.5 TFA/phenol/H₂O/iPr₃SiH afforded the desired cysteine-protected alkynyl 10F2 peptide **1a** in 11 % overall yield, following HPLC purification. Using an identical protocol, three other alkynyl peptides 10F5M (**2a**), 10F6 (**3a**) and 10F8 (**4a**), were synthesized in similar crude purities (Figure 4) with isolated yields ranging from 8–11 %. These peptides were then “click” glycosylated with 4.4–5.5 equivalents of sugar azide **5** in degassed solvents under nitrogen, affording the desired glycopeptides **18–21** in 21–43 % yield (Scheme 3b).

2.5. Preparation of glycopeptide-CRM197 conjugates

Carbohydrates alone do not generally stimulate the CD4 T-cell-dependent mechanisms of antibody maturation and class switching that lead to long-lasting high affinity IgG responses.²⁷ Carbohydrates are therefore commonly conjugated to carrier proteins, which not only provide T epitope peptides to recruit T cell help, but also present the carbohydrate antigen in a multimeric format that can result in B cell receptor clustering and therefore enhanced B cell activation.²⁸ We opted to conjugate our glycopeptides to CRM197, a mutant of the diphtheria toxin, which is detoxified by a single amino acid substitution of Glu for Gly. CRM197 is used commercially as the carrier protein in glycoconjugate vaccines for pneumonia, meningitis, and bacterial influenza²⁹ and was recently shown to be an effective carrier to promote IgG responses against the Globo-H tumor-associated carbohydrate antigen.³⁰ To facilitate attachment of our peptide cysteines to CRM197, we generated CRM197 maleimide **17** (Scheme 3a) by treating the carrier with 1 mg/mL (142 equivalents) of N-[ε-maleimidocaproyloxy] sulfo-succinimide ester (sulfo-EMCS) (**16**). After 2.5 hours of incubation and removal of excess sulfo-EMCS by buffer exchange through a 10 kDa-cutoff centrifugal filter, MALDI-TOF analysis of **17** showed an average molecular weight increase of ~4000 daltons, corresponding to an average of 21 maleimides attached to roughly half of the protein's 40 amino groups.

Glycopeptides **22–25** were prepared for conjugation to maleimided CRM197 (**17**) by reductive deprotection of the cysteine tBuS disulfide (Scheme 3b) using TCEP (tris(carboxyethyl)phosphine), and freshly deprotected glycopeptides were incubated with freshly-maleimided CRM197. Initially, conjugation to CRM197 resulted in very unsatisfactory loading of 0–2 molecules of glycopeptide **20** per CRM197 molecule, as assessed by MALDI-TOF (data not shown), even with 10 equivalents of **20** per CRM197 molecule (0.5 equiv. **20** per maleimide). After some investigation, we determined that excess TCEP was at fault: conjugation was greatly improved (average loading of 6, data not shown) by reducing the equivalents of TCEP used for cysteine deprotection from 300 to ten. Although excess TCEP is > 95 % removed by several rounds of buffer exchange in Amicon

centrifugal filters (3 kDa cutoff), even a few equivalents of residual TCEP are apparently able to inhibit the conjugation.³¹ A further improvement in conjugation efficiency (to a loading of 7–9) was realized by decreasing the pH of the PBS buffer from 7.5 to 6.5 (Scheme 3b,c). Although cysteine nucleophilicity is expected to decrease at more acidic pH, this effect may be counteracted by fuller utilization of the glycopeptide, owing to a decrease in competing oxidative dimerization at cysteine. This optimized procedure was applied to the conjugation of glycopeptides **18**, **19**, and **21**, resulting in loadings of 7–9 in all cases (**26–29**). Finally, all the unreacted maleimide groups were capped with β -mercaptoethanol (BME) to yield conjugates **30–33**.

2.6. Antigenicity of glycopeptide-CRM197 conjugates

Prior to use in immunogenicity studies in animals, it is important to verify that the presentation of the glycopeptide on the carrier does not interfere with the accessibility of the epitope. Therefore, we performed ELISA assays (Figure 5) to assess the binding of 2G12 to glycopeptide conjugates. Plates were coated with 12 ng of glycopeptide-CRM197 conjugates, JRFL HIV gp120, or maleimided CRM197 and incubated with 2G12; bound 2G12 was then detected with horseradish peroxidase (HRP)-conjugated goat anti-human antibody. As expected, this assay detected 2G12 binding to JRFL HIV gp120 envelope protein, and not to CRM197 alone (Figure 5). 10F2, 10F6 and 10F8 conjugates were all recognized by 2G12 more tightly than gp120, with EC_{50} s of 0.54–2.4 nM, vs. 3.1 nM for gp120. 10F5M conjugate was recognized somewhat more weakly, with an EC_{50} of 34 nM; this result does not quite correlate with the trend in K_D values for the four glycopeptides, since 10F5M glycopeptide alone binds to 2G12 as tightly as does 10F8 (K_D s of 2.4 nM and 2.6 nM, respectively). It is therefore likely that 10F5M is coupled to CRM197 in an orientation which interferes somewhat with its accessibility or conformation for binding to 2G12.

3. Conclusion

In conclusion, we have described herein how the use of fast flow peptide synthesis enabled rapid optimization of the preparation of peptide 40-mers, which were readily utilized in the synthesis of antigenic glycopeptides. We further detailed the “click” glycosylation of these peptides, the optimization of glycopeptide conjugation to carrier protein CRM197, and ELISA studies to verify the antigenicity of the resulting conjugates. Although these glycopeptide conjugates are tightly recognized by antibody 2G12, the more relevant question for vaccine development purposes is whether they can elicit an antibody response with 2G12-like specificity. Rabbit immunogenicity studies of these conjugates are currently in progress and will be reported in due course.

4. Experimental Section

4.1. General procedure

For peptide synthesis, DMF (certified ACS grade) and NMP (extra dry grade) were purchased from Fisher Scientific. N,N-diisopropylethylamine (DIEA) was distilled from CaH_2 . Trityl-OH ChemMatrix[®] resin was purchased from Biotage. All coupling reagents

(HATU, HBTU, and Fmoc-protected amino acids) were acquired from Chem-Impex International, Inc. Fmoc-(Homopropargyl)Gly-OH was obtained from ChemPep Inc. Man₉-cyclohexyl-azide used in click reaction was prepared in house.¹² MilliQ-grade water was used for all the reactions in aqueous solution. PBS buffer was prepared in-house with the following composition: 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄ (final concentrations). pH was adjusted with either NaOH or HCl.

Crude peptides and glycopeptides were analyzed by UPLC-ESI-MS on a Waters Acquity UPLC coupled to a ZQ4000 single quad mass spectrometer. All synthetic peptides were purified by RP HPLC (Waters Symmetry 300 C4, 5 μ m, 10 $\times$ 250 mm, 4 mL/ min, 2–42% MeCN in H₂O with 0.1% formic acid, over 60 min). All synthetic glycopeptides were purified by RP HPLC (Waters Symmetry 300 C4, 5 μ m, 10 $\times$ 250 mm, 4 mL/ min, 10–45% MeCN in H₂O with 0.1% formic acid, over 60 min). After HPLC purification, purity and identity of peptides and glycopeptides were assessed on the same UPLC-ESI-MS instrument. Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF-MS) spectra of glycopeptide-CRM197 conjugates were obtained on a Voyager DE-Pro instrument using sinapinic acid as matrix.

4.2. Synthesis of peptides and glycopeptides

4.2.1. Standard procedure for peptide synthesis—Peptides were prepared by Fmoc solid phase peptide synthesis, employing Pentelute's rapid flow-based method.¹⁸ HO-Trityl ChemMatrix[®] resin was loaded with alanine by 1) chlorination with acetyl chloride, treatment with Fmoc-Ala-OH/diisopropylethylamine followed by Fmoc removal. Loading was measured at 0.32 mequiv/g, as assessed by UV absorption of the Fmoc cleavage product. 20–40 μ mol of loaded resin was subjected to 39 cycles of flow-based peptide coupling (0.33M Fmoc-AA-OH, 0.33 M HATU, 0.95M DIEA for 6 mL/min for 30 sec at 60 °C) and Fmoc deprotection (80 seconds of 10 mL/min 20% piperidine in DMF, changing to 19% piperidine / 1% HCO₂H after incorporation of aspartate). Each coupling or deprotection was followed with 2 minutes of wash with 10mL/min DMF. There were exceptions to this procedure for a few amino acids: 1) Homopropargylglycine couplings were run as batch reactions in order to conserve valuable amino acid (0.15 mmol Fmoc-HPG-OH, 0.15 mmol HATU, 0.43 mmol (75 μ L) DIEA, stirring in 425 μ L DMF in 15-mL conical Falcon tube at 60 °C for 10 minutes). In flow reactions, a lower base concentration was used in the couplings of His and Cys(StBu) to prevent racemization (0.29M instead of 0.95M DIEA). Glycine couplings were conducted with HBTU instead of HATU to avoid double coupling. Finally, the N terminus was formylated in batch (0.25 mmol 4-nitrophenylformate, 125 μ L DIEA in 632 μ L DMF), with stirring at 60 °C in a 15-mL conical Falcon tube. Formylation was usually repeated 1–2 additional times before complete conversion was evident by LC/MS analysis of peptide cleaved from a few beads. Peptides were cleaved and deprotected using cleavage cocktail (87.5/5/5/2.5 TFA/H₂O/Phenol/iPr₃SiH) and then subjected to ether precipitation to afford crude peptide that was purified by RP HPLC (Waters Symmetry 300 C4, 5 μ m, 10 $\times$ 250 mm, 4 mL/ min, 2–42% MeCN in H₂O with 0.1% formic acid, over 60 min).

4.2.2. Standard procedure for “click” glycosylation of peptides—The synthesis of 10F8 glycopeptide (**21**) is representative: Peptide 10F8 (**4a**) (2.5 mg, 0.58 μmol , 1 equiv) and Man₉-cyclohexyl-azide **5** (4.1 mg, 2.6 μmol , 4.4 equiv) were combined and lyophilized into a ½-dram glass vial. A 0.5 mL Eppendorf tube, containing 14.5 μL (1.5 μmol , 2.5 equiv) of a 100 mM solution of CuSO_4 and 40 μL (1.6 μmol , 0.6 equiv) of 40 mM solution of THPTA ligand, was evaporated to dryness in a SpeedVac. Solid sodium ascorbate (5.8 mg, 29 μmol , 50 equiv) was placed in a second Eppendorf tube. The two Eppendorf tubes and the glass vial containing peptide and sugar azide (**5**) were placed in a two-neck flask with a conical bottom and the apparatus was evacuated and backfilled with nitrogen several times. Under nitrogen efflux, 200 μL of degassed DMSO was added to the glass vial and 45 μL of degassed water was added to each of the other tubes. Under nitrogen efflux, a syringe needle was inserted and used to transfer the CuSO_4 /THPTA mixture, and then the sodium ascorbate, to the peptide/sugar solution. The mixture stirred for 3h, at which time UPLC/MS showed nearly complete reaction. The reaction was quenched by addition of tetramethylethylenediamine (0.9 μL , 5.8 μmol , 10 equiv). The mixture was purified by RP-HPLC (Water Symmetry 300 C4, 5 μm 10 $\times$ 250 mm, 4 mL/min, 10–45% MeCN in H_2O with 0.1% formic acid, over 60 min, retention time 27 min) to afford 2.5 mg of pure glycopeptide **21** (quantified by BCA assay, see below), corresponding to 40% yield.

Each purified glycopeptide was quantified once by AAA (amino acid analysis), and each AAA result was compared to a BCA (Bicinchonic acid) assay measurement (Pierce) on the same glycopeptide. Using this data, a correction factor for the BCA assay was calculated for each glycopeptide, and subsequent batches of glycopeptide were then quantified by BCA assay.

4.2.3. Peptide 10F2 (1a)—Following the standard procedure, 60 mg of crude peptide 10F2 (**1a**) was obtained from 140 mg of trityl ChemMatrix[®] resin loaded with 0.30 mequiv/g alanine (42 μmol scale). 9.8 mg of this crude peptide 10F2 was purified by HPLC (general procedure, retention time 51 min) to afford 3.4 mg of pure peptide 10F2. This corresponds to 11 % of overall yield if all crude peptide had been purified. LR ESI-MS: observed average m/z of multiply charged ions 1086.08 $[\text{M} + 4\text{H}]^{4+}$, 1448.1 $[\text{M} + 3\text{H}]^{3+}$, corresponding to 4340.81 observed average mass, calculated average mass for $\text{C}_{196}\text{H}_{297}\text{N}_{53}\text{O}_{55}\text{S}_2$: 4339.91.

4.2.4. Peptide 10F5M (2a)—Following the standard procedure and starting with 100 mg of trityl ChemMatrix[®] resin loaded with 0.30 mequiv/g alanine (30 μmol scale), 47 mg of crude peptide 10F5M (**2a**) was obtained. HPLC purification of 4.2 mg crude peptide (general procedure, retention time 47 min) yielded 1.1 mg of pure peptide 10F5M (**2a**), corresponding to 9.4 % overall yield. LR ESI-MS: observed average m/z of multiply charged ions 1082.66 $[\text{M} + 4\text{H}]^{4+}$, 1443.22 $[\text{M} + 3\text{H}]^{3+}$, corresponding to 4326.65 observed average mass, calculated average mass for $\text{C}_{189}\text{H}_{311}\text{N}_{57}\text{O}_{55}\text{S}_2$: 4325.97.

4.2.5. Peptide 10F6 (3a)—Following the standard procedure and starting with 125 mg of trityl ChemMatrix[®] resin, loaded with 0.2 mequiv/g alanine (25 μmol scale), 85 mg of crude peptide 10F6 (**3a**) was obtained. HPLC purification of 59 mg of crude peptide (general

procedure, retention time 52 min), afforded 8.5 mg of pure peptide 10F6 (**3a**), corresponding to 11% overall yield. LR ESI-MS: observed m/z of multiply charged ions 1117.83 [M + 4H]⁴⁺, 1490.13 [M + 3H]³⁺, corresponding to 4467.36 observed average mass, calculated average mass for C₂₁₀H₃₁₈N₅₂O₅₂S₂: 4467.22.

4.2.6. Peptide 10F8 (4a)—Following the standard procedure and starting with 72 mg of trityl ChemMatrix[®] resin loaded with 0.3 mequiv/g alanine (23 μmol scale), 36 mg of crude peptide 10F8 (**4a**) was obtained. HPLC purification of 59 mg of crude peptide (general procedure, retention time 50 min) afforded 8.5 mg of pure peptide 10F8 (**4a**), corresponding to 8 % overall yield. LR ESI-MS: observed average m/z of multiply charged ions 1077.51 [M + 4H]⁴⁺, 1436.21 [M + 3H]³⁺, corresponding to 4305.84 observed average mass, calculated average mass for C₁₉₈H₃₁₂N₅₀O₅₃S₂: 4305.03.

4.2.7. Gly copeptide 10F2 (18)—The standard click procedure was followed with 2.3 mg (0.53 μmol, 1 equiv) of peptide 10F2 (**1a**) and 4.7 mg (2.9 μmol, 4.4 equiv) of sugar azide **5**. HPLC purification (general procedure, retention time 30 min) afforded 1.4 mg of pure glycopeptide 10F2 (**18**), corresponding to 21 % yield. LR ESI-MS: observed average m/z of multiply charged ions 1543.9 [M + 8H]⁸⁺, 1764.36 [M + 7H]⁷⁺, corresponding to 12342.22 observed average mass, calculated average mass for C₄₉₆H₈₀₂N₆₈O₂₈₅S₂: 12343.36.

4.2.8. Gly copeptide 10F5M (19)—The standard click procedure was followed with 3.5 mg (0.81 μmol, 1 equiv) of peptide 10F5M (**2a**) and 5.7 mg (3.6 μmol, 5.5 equiv) of sugar azide **5**. HPLC purification (general procedure, retention time 20 min) afforded 3.7 mg of pure glycopeptide 10F5M (**19**), corresponding to 43 % yield. LR ESI-MS: observed average m/z of multiply charged ions 1533.3 [M + 7H]⁷⁺, 1788.92 [M + 6H]⁶⁺, corresponding to 10726.81 observed average mass, calculated average mass for C₄₂₉H₇₁₅N₆₉O₂₃₉S₂: 10727.72.

4.2.9. Gly copeptide 10F6 (20)—The standard click procedure was followed with 2.5 mg (0.56 μmol, 1 equiv) of peptide 10F6 (**3a**) and 4.9 mg (3.1 μmol, 5.5 equiv) of sugar azide **5**. HPLC purification (general procedure, retention time 25 min) afforded 2.1 mg of pure glycopeptide 10F6 (**20**), corresponding to 36 % yield. LR ESI-MS: observed average m/z of multiply charged ions 1559.85 [M + 8H]⁸⁺, 1782.43 [M + 7H]⁷⁺, corresponding to 124.82 observed average mass, calculated average mass for C₅₁₀H₈₂₃N₆₇O₂₈₂S₂: 12470.41.

4.2.10. Gly copeptide 10F8 (21)—The standard click procedure was followed with 2.5 mg (0.58 μmol, 1 equiv) of peptide 10F8 (**4a**) and 4.1 mg (2.6 μmol, 4.4 equiv) of sugar azide **5**. HPLC purification (general procedure, retention time 27 min) afforded 2.5 mg of pure glycopeptide 10F8 (**21**), corresponding to 40 % yield. LR ESI-MS: observed average m/z of multiply charged ions 1530.56 [M + 7H]⁷⁺, 1785.58 [M + 6H]⁶⁺, corresponding to 10707.20 observed average mass, calculated average mass for C₄₃₈H₇₁₆N₆₂O₂₃₇S₂: 10706.78.

4.3. K_D determination of 10F5M-2G12 interaction

The plasmid encoding 10F5M was prepared by site-directed mutagenesis of the wild type 10F5 plasmid. The radiolabeled glycopeptide was prepared and the binding constant for 2G12 was determined as previously described.¹⁴ Detailed information is described in the Supporting Information.

4.4. Synthesis of CRM197-maleimide (**17**)

CRM197 (Fina Biosolutions EcoCRM) (1 mg, 17 nmol) was first desalted by buffer exchange through an Amicon centrifugal filter (10 kDa cutoff, Ultra-0.5). The commercial solution was diluted with PBS buffer (pH 7.5) to a volume of 500 μ L, and then concentrated in the filter to a volume of 42 μ L. This was repeated two more times. The concentrate was then dissolved in 1 mL of PBS buffer (pH 7.5) and transferred into a 1.5 mL Eppendorf tube. Sulfo-EMCS (1 mg, 2.4 μ mol, 142 eq) was added to the solution and the reaction was incubated at room temperature for 2.5 h. Excess Sulfo-EMCS was removed by buffer exchange through an Amicon centrifugal filter as detailed above (10 kDa cutoff, Ultra-0.5, at least 3 rounds of dilution with PBS pH6.5). The molecular weight of the activated CRM197-maleimide (**17**) was determined by MALDI-TOF analysis, showing a typical increase of ~3900, corresponding to 21 maleimide linkers. BCA assay quantification indicated a yield of 1.1 mg.

4.5. Synthesis of glycopeptide-CRM197 conjugates (**30–33**)

Representative procedure with 10F2 sequence: Glycopeptide 10F2 (**18**) (1.3 mg, 105 nmol) in 500 μ L water was treated with 2 μ L of 0.5 M TCEP-HCl / 1M Tris-HCl buffer (pH 7.8, 10 eq). The reaction proceeded overnight at room temperature under nitrogen atmosphere, after which time UPLC-ESI-MS indicated complete deprotection of the cysteine. Excess TCEP was removed by buffer-exchange through an Amicon centrifugal filter (3-kDa cutoff, Ultra-0.5) with PBS buffer (pH 6.5) (3-kDa cutoff, 20 mins in the first round of filtration and 30 mins for the second round). After the first round of filtration, the reaction mixture was concentrated to 62 μ L. It was then diluted with PBS (pH 6.5) to a volume of 500 μ L, following by second round of filtration to concentrate to a final volume of 48 μ L. The freshly-produced CRM197-maleimide **17** (0.60 mg as quantified by BCA assay, 10 nmol, corresponding to ~220 nmol of maleimide groups) was diluted to 1 mg/mL (550 μ L) in PBS (pH 6.5). The CRM197-maleimide solution was transferred into a 1.5 mL screwcap tube inside a 2-neck flask flushed with nitrogen and the freshly deprotected glycopeptide solution (48 μ L in PBS pH 6.5) was then added. After overnight incubation under nitrogen, the glycopeptide-CRM197 conjugate was purified by using Amicon Ultra-0.5 (50 kDa cutoff) to remove salts and unreacted glycopeptide. MALDI-TOF MS analysis indicated the distribution of conjugates of various loadings, 7–9 glycopeptides per CRM197 molecule being the most prevalent. The glycopeptide-CRM197 conjugate was then dissolved in PBS buffer (pH 6.5, concentration = 1mg/mL) and capped by treatment with 20 equivalents of BME. The mixture was incubated at room temperature for 2h. Excess BME and salts were then removed from the glycopeptide-CRM197 by buffer exchange into water using Amicon Ultra-0.5 (50 kDa cutoff, 10 min each round, 5 rounds). Corrected BCA assay quantitation indicated 0.94 mg of 10F2 conjugate (**30**). Analogous procedures starting with 2.1, 1.8, and

1.6 mg of glycopeptides 10F5M (**19**), 10F6 (**20**), and 10F8 (**21**) respectively, afforded 2.6, 2.3, and 2.3 mg of the respective conjugates **31–33**.

4.6. ELISA procedure

High-protein-binding flat-bottomed ELISA plates (Nunc-Immuno) were coated with 12 ng of antigen per well in 60 μ L coating buffer at 4°C overnight. The wells were washed twice with PBS-0.05% Tween 20 (PBS-T) and then blocked for 2h at room temperature with PBS-T 5% fat-free milk (200 μ L/ well). After the wells were again washed twice with PBS-T, they were then incubated with 2.5-fold serial dilutions of 2G12 (starting at different concentrations: 50 nM for 10F2 (**30**), 750 nM for 10F5M (**31**), 20 nM for 10F6 (**32**), and 120 nM for 10F8 (**33**), JRFL gp120, CRM197-maleimide) in PBS-T 1% fat-free milk for 2h at room temperature. The wells were washed 3 times before incubating with 60 μ L of a horseradish peroxidase (HRP) conjugated goat anti-human antibody (Invitrogen, part number 81–7120) at 1:20,000 dilution for 1h at room temperature. After 3 washes, the wells were developed by adding 60 μ L of 3,3',5,5'-tetramethylbenzidine (TMB solution, Abcam Ab171522) for 2 min. The reaction was stopped by adding 60 μ L of 1M sulfuric acid and absorbance was measured at 450 nm wavelength by a plate reader. All measurements were performed in triplicate.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Support of the NIH (R01-AI090745 and R01-AI113737) is gratefully acknowledged. We also wish to thank J. Sebastian Temme for synthesis of Mang-cyclohexyl-azide, the Kritzer group at Tufts University for allowing us to use their batch peptide synthesizers, and the Pentelute group at MIT for their advice and help in setting up fast flow peptide synthesis.

References and Notes

1. Doria-Rose NA, Klein RM, Daniels MG, O'Dell S, Nason M, Lapedes A, Bhattacharya T, Migueles SA, Wyatt RT, Korber BT, Mascola JR, Connors M. J. Virol. 2010; 84:1631–1636. [PubMed: 19923174]
2. Sather DN, Armann J, Ching LK, Mavrantoni A, Sellhorn G, Caldwell Z, Yu X, Wood B, Self S, Kalams S, Stamatos L. J. Virol. 2009; 83:757–769. [PubMed: 18987148]
3. Simek MD, Rida W, Priddy FH, Pung P, Carrow E, Laufer DS, Lehrman JK, Boaz M, Tarragona-Fiol T, Miiro G, Birungi J, Pozniak A, McPhee DA, Manigart O, Karita E, Inwoley A, Jaoko W, DeHovitz J, Bekker L-G, Pitisuttithum P, Paris R, Walker LM, Poignard P, Wrin T, Fast PE, Burton DR, Koff WC. J. Virol. 2009; 83:7337–7348. [PubMed: 19439467]
4. Moldt B, Rakasz EG, Schultz N, Chan-Hui P-Y, Swiderek K, Weisgrau KL, Piaskowski SM, Bergman Z, Watkins DI, Poignard P, Burton DR. Proc. Natl. Acad. Sci. U. S. A. 2012; 109:18921–18925. [PubMed: 23100539]
5. Hessel AJ, Rakasz EG, Poignard P, Hangartner L, Landucci G, Forthal DN, Koff WC, Watkins DI, Burton DR. PLoS Pathog. 2009; 5:e1000433. [PubMed: 19436712]
6. Correia BE, Bates JT, Loomis RJ, Baneyx G, Carrico C, Jardine JG, Rupert P, Correnti C, Kalyuzhniy O, Vittal V, Connell MJ, Stevens E, Schroeter A, Chen M, MacPherson S, Serra AM, Adachi Y, Holmes MA, Li Y, Kleit RE, Graham BS, Wyatt RT, Baker D, Strong RK, Crowe JE, Johnson PR, Schief WR. Nature. 2014; 507:201–206. [PubMed: 24499818]

7. Kulp DW, Schief WR. *Curr. Opin. Virol.* 2013; 3:322–331. [PubMed: 23806515]
8. Horiya S, MacPherson IS, Krauss IJ. *Nat. Chem. Biol.* 2014; 10:990–999. [PubMed: 25393493]
9. Lavine CL, Lao S, Montefiori DC, Haynes BF, Sodroski JG, Yang XZ. *Immunol N. C. H. A. V. J. Virol.* 2012; 86:2153–2164. [PubMed: 22156525]
10. Landais E, Huang X, Havenar-Daughton C, Murrell B, Price MA, Wickramasinghe L, Ramos A, Bian CB, Simek M, Allen S, Karita E, Kilembe W, Lakhi S, Inambao M, Kamali A, Sanders EJ, Anzala O, Edward V, Bekker L-G, Tang J, Gilmour J, Kosakovsky-Pond SL, Phung P, Wrin T, Crotty S, Godzik A, Poignard P. *PLoS Pathog.* 2016; 12:e1005369. [PubMed: 26766578]
11. Krauss IJ. *Glycobiology*. 2016 *accepted article*.
12. Temme JS, Drzyzga MG, MacPherson IS, Krauss IJ. *Chem. Eur. J.* 2013; 19:17291–17295. [PubMed: 24227340]
13. Temme JS, MacPherson IS, DeCoursey JF, Krauss IJ. *J. Am. Chem. Soc.* 2014; 136:1726–1729. [PubMed: 24446826]
14. Horiya S, Bailey JK, Temme JS, Guillen Schlippe YV, Krauss IJ. *J. Am. Chem. Soc.* 2014; 136:5407–5415. [PubMed: 24645849]
15. MacPherson IS, Temme JS, Habeshian S, Felczak K, Pankiewicz K, Hedstrom L, Krauss IJ. *Angew. Chem. Int. Ed.* 2011; 50:11238–11242.
16. Calarese DA, Scanlan CN, Zwick MB, Deechongkit S, Mimura Y, Kunert R, Zhu P, Wormald MR, Stanfield RL, Roux KH, Kelly JW, Rudd PM, Dwek RA, Katinger H, Burton DR, Wilson IA. *Science.* 2003; 300:2065–2071. [PubMed: 12829775]
17. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. *Angew. Chem. Int. Ed.* 2002; 41:2596–2599.
18. Simon MD, Heider PL, Adamo A, Vinogradov AA, Mong SK, Li X, Berger T, Policarpo RL, Zhang C, Zou Y, Liao X, Spokoiny AM, Jensen KF, Pentelute BL. *ChemBioChem.* 2014; 15:713–720. [PubMed: 24616230]
19. Han Y, Albericio F, Barany G. *J. Org. Chem.* 1997; 62:4307–4312. [PubMed: 11671751]
20. García-Martín F, Quintanar-Audelo M, García-Ramos Y, Cruz LJ, Gravel C, Furic R, Côté S, Tulla-Puche J, Albericio F. *J. Comb. Chem.* 2006; 8:213–220. [PubMed: 16529516]
21. Lukszo J, Patterson D, Albericio F, Kates SA. *Lett. Pept. Sci.* 1996; 3:157–166.
22. Pedroso E, Grandas A, Saralegui MA, Giralt E, Granier C, van Rietschoten J. *Tetrahedron.* 1982; 38:1183–1192.
23. Mong SK, Vinogradov AA, Simon MD, Pentelute BL. *ChemBioChem.* 2014; 15:721–733. [PubMed: 24616257] . We have observed epimerization at the higher base concentration during coupling of cysteine and histidine, but not other amino acids, consistent with Pentelute's original report.
24. Subirós-Funosas R, El-Faham A, Albericio F. *Tetrahedron.* 2011; 67:8595–8606.
25. Mergler M, Dick F, Sax B, Weiler P, Vorherr T. *J. Pept. Sci.* 2003; 9:36–46. [PubMed: 12587881]
26. Michels T, Dölling R, Haberkorn U, Mier W. *Org. Lett.* 2012; 14:5218–5221. [PubMed: 23025410]
27. Coutinho A, Möller G. *Eur. J. Immunol.* 1973; 3:608–613. [PubMed: 4148779]
28. Pollard AJ, Perrett KP, Beverley PC. *Nat. Rev. Immunol.* 2009; 9:213–220. [PubMed: 19214194]
29. Shinefield HR. *Vaccine.* 2010; 28:4335–4339. [PubMed: 20452430]
30. Huang Y-L, Hung J-T, Cheung SKC, Lee H-Y, Chu K-C, Li S-T, Lin Y-C, Ren C-T, Cheng T-JR, Hsu T-L, Yu AL, Wu C-Y, Wong C-H. *Proc. Natl. Acad. Sci. U. S. A.* 2013; 110:2517–2522. [PubMed: 23355685]
31. Shafer DE, Inman JK, Lees A. *Anal. Biochem.* 2000; 282:161–164. [PubMed: 10860517]

Fig. 1.

"Structure-based" or "epitope-focused" vaccine design. bnAbs bind to conserved epitopes that are accessible on trimeric gp120 present on diverse live virus. However, vaccination with recombinant gp120 protein monomer largely elicits ineffective antibodies directed to other epitopes that may be inaccessible or mutated in live virus. In epitope-focused vaccine design, directed evolution or other design tools are used to create structural mimics of the bnAb epitope; vaccination with bnAb epitope mimics should in principle be able to elicit antibodies that bind the bnAb epitope, and thus bind effectively to diverse viral strains and prevent infection.

Fig. 2.

a) Sequences of glycopeptides to be synthesized. 10F5M is a mutant of the originally evolved sequence 10F5, in which the cysteines at positions 10 and 25 have been changed to serines, (indicated by a bold S); b) Strategy for glycopeptide synthesis. After Fmoc SPPS and N-terminal formylation, peptide is cleaved from the resin and click glycosylated with Man₉-cyclohexyl-azide (**5**) to yield glycopeptides; * X's denote the amino acids other than M present in peptides **1a–4a**. m = 3–4 sugar sites. c) UV chromatogram of a complex crude mixture resulting from attempted synthesis of a peptide using a standard batch automated

peptide synthesizer at room temperature. Little of the desired product was detectable by MS analysis.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 3. ESI-MS chromatogram of crude peptide **11b** showing significant aspartimide formation just three SPPS cycles after incorporation of aspartate (D).

Fig. 4. UV chromatograms of crude 40-mer alkynyl peptides **1a–4a**. ^a280 nm wavelength. ^b220 nm wavelength.

Analyte	EC ₅₀ (nM)	K _D (nM) ^a
10F2 (30)	0.64 ± 0.05	0.60 ± 0.0045
10F5M (31)	34 ± 0.05	2.4 ± 0.49 ^b
10F6 (32)	0.54 ± 0.02	0.54 ± 0.043
10F8 (33)	2.4 ± 0.13	2.5 ± 0.23
gp120 (JRFL)	3.12 ± 0.12	ND ^c
CRM197-maleimide	NB ^d	ND ^c

Fig. 5.

Glycopeptide and CRM197 conjugate binding to 2G12. EC₅₀ values refer to the 2G12 concentrations that gave half-maximal binding to adsorbed glycopeptide-CRM197 conjugates in ELISA format. ^aK_D values are for glycopeptides binding to 2G12, as reported in Ref. 14, except where noted. ^bSee SI for K_D determination. ^c Not determined. ^d No binding.

Scheme 1.

a) Undesired formation of 3-(1-piperidinyl)alanine with C-terminal Cys(Trt) anchored on chlorotrityl resin; b) Syntheses beginning with commercial H-Gly-HPMB-NovaPEG resin yielded mono-, di-, and tri-glycine products.

Scheme 2.

Synthesis of 10F2 alkynyl peptide (**1a**). For peptides still on resin, amino acid letter abbreviations have implicit protecting groups, except Cys, for which the protecting group is drawn. ^aGlycine residues were coupled using HBTU instead of HATU.

Scheme 3.

a) Maleimidation of CRM197 and MALDI-TOF MS of maleimidated CRM197; b) Glycosylation of peptides and conjugation to CRM197. *All four peptide sequences are illustrated generically with X's, and it is implicit that the *M* sites are located at the positions specified in Figure 2a (*m* = 3–4); c) MALDI-TOF MS indicating the extent of 10F2 glycopeptide loading on CRM197. “*n*” represents the qualitatively estimated average loading based on this data. Because MALDI-TOF is insensitive to detection of very large molecular weights, this “*n*” is a slight underestimate of the true average loading.