Published in final edited form as:
Biochemistry. 2009 September 22; 48(37): 8891–8898. doi:10.1021/bi9011429.

Mutagenesis and Evolution of Sulfated Antibodies using an Expanded Genetic Code†

Chang C. Liu[‡], **Hyeryun Choe**[§], **Michael Farzan**[‖], **Vaughn V. Smider**[⊥], and **Peter G. Schultz**[‡,*]

[‡] Department of Chemistry and Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

[§] Children's Hospital, Department of Pediatrics, Harvard Medical School, Boston, MA 02115

[‖] Department of Microbiology and Molecular Genetics, Harvard Medical School, New England Primate Research Center, Southborough, MA 01772

[⊥] Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Abstract

To facilitate the biochemical study of posttranslationally modified proteins, we have developed a strategy in which otherwise posttranslationally modified amino acids are genetically encoded in *E. coli* in response to unique nonsense or frameshift codons. Here, we illustrate the utility of this approach through the characterization of the doubly tyrosine-sulfated anti-gp120 antibody, 412d. By expressing selectively sulfated variants of 412d directly in *E. coli* with an orthogonal aminoacyl-tRNA synthetase/tRNA pair specific for sulfotyrosine, we were able to determine the contribution of each of the sulfates in 412d to gp120 binding affinity: tyrosine sulfation of 412d at position H100, position H100c, or dual sulfation at both positions (Kabat numbering where H designates heavy chain) leads to an increase in affinity for gp120 of 4.5-fold, 212-fold, or 500-fold, respectively. We also conducted directed evolution experiments to evolve 412d beyond the known sequence constraints required for posttranslational sulfation, while retaining the two tyrosine-sulfates essential for function, yielding novel doubly sulfated antibodies, one of which binds gp120 with subnanomolar affinity. Taken together, our studies provide a more complete understanding of the role of 412d sulfation in gp120 binding, and highlight the utility of genetically encoded unnatural amino acids in exploring the effects of posttranslational modifications on protein function.

Posttranslational modifications (PTMs)[1] such as phosphorylation, glycosylation, and tyrosine sulfation regulate many complex biological processes, ranging from signal transduction and transcription to protein trafficking and degradation. Unfortunately, the detailed biochemical characterization of the roles of these modifications is often complicated by difficulties involved in isolating posttranslationally modified proteins in defined states. Moreover, the expression

†This work was supported by the US National Institutes of Health (GM62159) and by predoctoral fellowships from the Fannie and John Hertz Foundation (C.C.L.) and the National Science Foundation (C.C.L.).

*To whom correspondence should be addressed: P.G.S., schultz@scripps.edu, 858-784-9273 (ph), 858-784-9440 (fax).

SUPPORTING INFORMATION

Protein sequences, mass spectra, and Sulfinator data is available free of charge on the web (www.pubs.acs.org).

[1]Abbreviations: PTM, posttranslational modification; sY, sulfotyrosine; aaRS, aminoacyl-tRNA synthetase; sTyrRS, sulfotyrosine-specific aminoacyl-tRNA synthetase; Fab, antigen binding fragment; scFv, single-chain variable fragment; SPR, surface plasmon resonance; PAGE, polyacrylamide gel electrophoresis; in this paper, Kabat numbering for antibodies where H designates heavy chain is used.

and mutagenesis of proteins containing PTMs typically necessitates additional enzymes that are restricted to the modification of specific sequences in specific cell types or organisms. One strategy to overcome these limitations is to genetically encode an unnatural amino acid that already contains the modification corresponding to the PTM of interest (1). This requires an orthogonal aaRS/tRNA pair that uniquely recognizes the free unnatural amino acid and incorporates it site-specifically into proteins in bacteria, yeast, or mammalian cell hosts in response to unique nonsense or frameshift codons.

A common PTM that can be studied using this approach is tyrosine sulfation. Sulfotyrosine is found in numerous eukaryotic proteins including those involved in cell adhesion (2–4), ligand/receptor association (5–7), and viral entry (8,9), and acts to enhance interaction strength at protein interfaces, often through the formation of strong salt bridges and hydrogen bonds by the sulfate group (10–12). Most notably, sulfation plays a central role in endogenous chemokine signaling and consequently is involved in the entry of HIV through binding of gp120 to its sulfated coreceptor CCR5 (9) or CXCR4 (5). As receptor sulfation is essential for gp120 association, sulfated anti-gp120 antibodies that exploit this dependence have been identified in human patients (13). A similar paradigm has also been characterized for malaria, which enters cells through the sulfated Duffy antigen/receptor for chemokines (14,15).

One anti-gp120 antibody, 412d, found in human HIV patients, has two sulfotyrosine residues both of which participate in strong interactions with gp120. However, difficulties in expressing site-specifically sulfated proteins (16,17) have limited the detailed characterization of this antibody. Here, we biosynthetically introduce sulfotyrosine into 412d by encoding it in response to the amber nonsense codon TAG in *E. coli* (18,19) and characterize the contribution of each sulfate to the total free energy of gp120 binding. We also evolve, using our recently described phage-based system (20), new sulfotyrosine-containing antibodies that extend beyond the known consensus features for posttranslational sulfation (16,21) and determine their affinity for gp120.

EXPERIMENTAL PROCEDURES

Synthesis of Sulfotyrosine

Sulfotyrosine was synthesized from L-tyrosine (Aldrich) and chlorosulfonic acid (Fisher). L-tyrosine (10 g, 1.1 M) was dissolved in trifluoroacetic acid (50 mL, Fisher) in a dry round bottom flask containing a stir bar, and cooled to −10 °C. While stirring, 5 mL (1.37 M) of chlorosulfonic acid was added over 2 minutes. The reaction was stirred for an additional 5 minutes, quenched by the slow addition of ethanol (3 mL), then stirred for 2 minutes at room temperature. Diethylether (175 mL) was added to precipitate sulfotyrosine, which was then filtered and washed 3X with diethylether (75 mL per wash). The product, a white powder, was dried under high vacuum to remove residual ether, and then dissolved in 2 M aqueous NaOH until the solution reached pH 7. The solution was filtered through a 0.22 μm sterile filter (Millipore) and the concentration of sulfotyrosine was measured by UV absorbance (λ_{max} = 263 and ε = 224 M^{-1} cm^{-1}, as determined from a sulfotyrosine standard purchased from Bachem). This concentrated sulfotyrosine solution was used directly in *E. coli* growth media. 95% yield; λ_{max} = 263 and ε = 224 M^{-1} cm^{-1} at pH 7.5, aqueous; ^{1}H-NMR (500 MHz, D_2O) δ3.13 (dd, *J* = 5.5 Hz, *J* = 14.5 Hz, 1H), 3.34 (dd, *J* = 9 Hz, *J* = 14.5 Hz, 1H), 4.01 (dd, *J* = 5.5 Hz, *J* = 9 Hz, 1H), 7.33 (m, 2H), 7.38 (m, 2H); LCMS (ESI) for calculated $C_9H_{11}NO_6S$ (M + 1) 262.25, observed 262.2.

Cloning, Expression, and Protein Purification of Fabs

The light chain and heavy chain variable regions for all Fabs of interest were amplified by PCR with primers that contained a 5′ extension of 5′-CCTTACGCGTATGCT-3′ and a 3′ extension

of 5′-GGTTCGTACG-3′ for the light chain, and primers that contained a 5′ extension of 5′-CCAACCATGGCT-3′ and a 3′ extension of 5′-GGAAGCTAGC-3′ for the heavy chain. The variable regions were separately inserted into a modified pBC expression vector that encodes human heavy chain gamma and light chain kappa constant regions (synthesized by Blue Heron). To insert the light chain variable region, the restriction sites MluI and BsiWI (underlined in primers) were used; to insert the heavy chain variable region, restriction sites NcoI and NheI (underlined in primers) were used. The resulting plasmids were termed pBC-412dYY, pBC-131-1, and pBC-131-3. To generate singly and doubly sulfated 412d variants, pBC-412dYY was mutated to pBC-412dTAGH100, pBC-412dTAGH100c, or pBC-412d2TAG by site-directed mutagenesis using Quikchange (Stratagene). For mutagenesis, sites corresponding to the desired location(s) of sulfotyrosine were replaced by the codon TAG. All recombinant DNA manipulations involving pBC plasmids were carried out in Top10F′ *E. coli* cells grown in the presence of 1% glucose to repress the lac promoter. All of the resulting plasmids encode Fabs in a bicistronic construct (light and heavy chains) under the control of the lac promoter.

To express Fabs, Top10F′ cells containing pSUPAR6-L3-3SY (19) were transformed with the pBC plasmid of interest and grown at 37 °C at 250 rpm in TB (Terrific Broth, Fisher) supplemented with 30 μg/mL chloramphenicol, 100 μg/mL ampicillin, and 10 mM sulfotyrosine when applicable. When the culture reached an optical density (OD_{600}) of 1.0, aaRS expression was induced with 0.2% L-arabinose (Fisher), followed by induction of Fab expression with 1 mM IPTG (Fisher) at OD_{600} = 1.5. The culture was then shaken at 250 rpm at 18 °C for 36 hours, after which cells were pelleted and lysed three times with 1/20 culture volume of periplasmic lysis buffer (20% sucrose, 30 mM Tris-HCl, 1 mM EDTA, 1 mg/mL lysozyme, pH 7.4). The periplasmic lysates were combined and passed through a 0.22 μm sterile membrane (Millipore).

To purify periplasmic lysates, 1 mL of protein G resin (Pierce) was packed into a 1 mL polypropylene column (Qiagen). After equilibration of the column with 5 mL binding buffer (50 mM MES, 100 mM NaCl, pH 5.5), periplasmic lysate was loaded onto the column and allowed to pass through the resin by gravity flow. The column was then washed with binding buffer (15 mL) and eluted with elution buffer (5 mL, 100 mM glycine, pH 2.7), which was immediately neutralized to pH 7.4. The eluted Fab was dialyzed into PBS and concentrated for further use. Fab yields were determined by UV absorbance at λ = 280; ε = 1.37 mg^{-1} mL.

Binding Kinetics and Free Energy Analysis

Association and dissociation rate constants were obtained by SPR using a BIAcore S51 biosensor. Gp120 (ADA isolate), produced and purified using established methods (13), was covalently immobilized onto a CM5 chip using amine coupling and the manufacturer's protocols. After exchanging Fabs into running buffer (PBS, pH 7.4), establishing conditions that did not show signs of mass transfer limitations, and optimizing regeneration conditions, we conducted Fab binding experiments. Analyte Fabs were injected onto the chip at a rate of 10 μL min^{-1} for 120 seconds and allowed to dissociate for 320 seconds. Concentrations between 0 nM and 400 nM were chosen to provide acceptable association curves (for those Fabs that bound with low k_{on}, concentration ranges above K_d were used to clearly observe association). Data from the 0 nM concentrations were used to subtract noise, and double referenced sensorgrams were fit using a 1:1 binding model with a term for mass transport and local R_{max} values. Local R_{max} values were used for fitting because the tight binding of several Fabs required harsh regeneration conditions (400 mM NaCl, 100 mM glycine, pH 2.7, 1 minute) that slightly decreased R_{max} of duplicate samples from run to run, though no effect on baseline noise was observed. Data were fit using the S51 BIA evaluation 1.2 software.

To calculate free energy contributions of the sulfates in 412d, the following equations were used: contribution of sulfate at position H100 = $\Delta\Delta G_{412d\text{-}YY,\ 412d\text{-}H100sY} = \Delta G_{412d\text{-}H100sY} - \Delta G_{412d\text{-}YY} = RT ln(K_{d412d\text{-}H100sY}/K_{d412d\text{-}YY})$; contribution of sulfate at position H100c = $\Delta\Delta G_{412d\text{-}YY,\ 412d\text{-}H100csY} = \Delta G_{412d\text{-}H100csY} - \Delta G_{412d\text{-}YY} = RT ln(K_{d412d\text{-}H100csY}/K_{d412d\text{-}YY})$; simultaneous contribution of sulfates at both positions = $\Delta\Delta G_{412d\text{-}YY,\ 412d\text{-}sYsY} = \Delta G_{412d\text{-}sYsY} - \Delta G_{412d\text{-}YY} = RT ln(K_{d412d\text{-}sYsY}/K_{d412d\text{-}YY})$. Temperature (T) was 298 K.

412d Library Construction

To create a 412d library with NNK randomization at heavy chain residues Asp99, Asn100a, Asp100b, Pro100e, Glu100f, and Met100i, we first inserted the scFv version of 412d into the pSEX backbone (20), yielding pSEX-412d2TAG. From this plasmid, a fragment of the heavy chain was amplified with the primers: 412dLib2-1: 5′-CCGCTGGCTTGCTGCTGCTG-3′ and 412dLib2-2: 5′-GTAAGGGCTCGCACAGTAAAATACGGCC-3′. The resulting product was then extended with overlap PCR with the primers: 412dLib2-3: 5′-GCCGTATTTTACTGTGCGAGCCCTTACCCGAATNNKTAGNNKNNKTAGGCCNNK NN KGAGGGANNKAGCTGGTACTTCGATCTCTG-3′ and 412dLib2-4: CAGAGATCGAAGTACCAGCT where N = A, T, G, or C and K = G or T. A second fragment was amplified from pSEX-412d2TAG with the primers: 412dLib2-5: AGCTGGTACTTCGATCTCTG and 412dLib2-6: CTCTGATATCTTTGGATCCA, and the two fragments were assembled by overlap PCR to give the 412d gene library. This library was digested with NcoI and BamHI and inserted into a similarly digested pSEX81 phagemid vector (22). The ligation product was heat denatured, ethanol precipitated with tRNA assistance, and transformed into electrocompetent Top10F′ cells. After overnight growth at 37 °C in 2YT supplemented with 100 μg/mL ampicillin, 30 μg/mL tetracycline, and 1% glucose, the supercoiled DNA was isolated. The library DNA mixture and individual library clones were sequenced to ensure library quality. Library DNA was then transformed into SY-X-*E. coli* for phage display and selection of anti-gp120 antibodies.

Phage Production

Phage experiments are based on previously described protocols for protein evolution with sulfotyrosine (20). SY-X-*E. coli* transformed with either phagemid library or a single phagemid clone was grown overnight at 37 °C in 2YT (2X yeast extract tryptone media, Difco) supplemented with 100 μg/mL ampicillin, 30 μg/mL tetracycline, 30 μg/mL chloramphenicol, and 1% glucose. After saturation was reached, a total volume equal to 20% of the desired final phage expression volume was removed from the culture, pelleted, and resuspended in a volume of 2YT equal to 40% of the desired final phage expression volume. The 2YT was supplemented with 100 μg/mL ampicillin, 30 μg/mL tetracycline, and 30 μg/mL chloramphenicol. To this culture was added hyperphage (Progen) at 20 MOI (MOI = multiplicity of infection). The resulting culture was then incubated at 100 rpm for 1 hour at 37 °C after which the cells were pelleted and the media was removed. The cells were then resuspended in the desired final phage expression volume of 2YT supplemented with 100 μg/mL ampicillin, 50 μg/mL kanamycin, and 30 μg/mL chloramphenicol. Sulfotyrosine (10 mM) was added directly to the media if necessary and the phage culture was incubated at 280 rpm for 18 hours at 30 °C. After phage production, the media was collected and cells were discarded. The media was then concentrated to a convenient volume (0.5–3 mL) using 10 kD cutoff concentrators (Amicon). The concentrated phage was then precipitated by the addition of 5X phage precipitation buffer (20% PEG 8000, 2.5 mM NaCl) and dissolved in PBS, pH 7.4. Precipitation was done twice.

Two methods were used to titer phage according to published protocols (20,23). For selection experiments where infective phage are the relevant population, phage was quantified by infection of Top10F′ cells. Specifically, a small volume of phage corresponding to ~10^4 particles was digested with 1.75 μg/mL trypsin (Fisher) and used to infect Top10F′ cells. The

amount of infected cells was then determined by plating dilutions on selective ampicillin plates. For phage ELISAs where the total particles of phage is the relevant population, phage in PBS (100 μL) were coated into wells of a MaxiSorp microtiter plate (Nunc) for 2 hours at 37 °C, blocked with 2% milk (Biorad) in PBS (200 μL) for 2 hours at 37 °C, washed several times with 200 μL PBST (PBS + 0.025% Tween 20), bound by an HRP conjugated anti-M13 polyclonal antibody (NEB) in 2% milk PBST (105 μL) for 2 hours at 37 °C, and detected with a QuantaBlu fluorogenic substrate (110 μL, Pierce) whose signal (excitation/emission, 325/420 nm) was read on a fluorescence plate reader (SpectraMax Gemini). This sample was compared to a standard curve where known amounts of hyperphage were adsorbed onto the ELISA plate and similarly treated.

Selection of gp120 binding phage

Soluble gp120 (0.5 μg, ADA isolate) was coated onto the surface of a MaxiSorp microtiter plate well in PBS (100 μL) for 12 hours at 4 °C. After blocking for 2 hours with 2% milk in PBS (200 μL) and washing 3X with 200 μL PBST, concentrated phage library was added in 2% milk PBST (100 μL) and incubated at 37 °C for 2 hours. Washing was done with 10X PBST and 2X PBS (200 μL per wash) over 30 minutes after which phage were eluted with 1.75 μg/mL trypsin in PBS (105 μL) for 12 minutes. Eluted phage were used to infect 20 mL of an OD_{600} = 0.4 culture of SY-X-*E. coli* in 2YT supplemented with 30 μg/mL chloramphenicol and 30 μg/mL tetracycline. Infection was carried out for 1 hour at 37 °C (100 rpm shaking) after which a small aliquot of cells was plated to determine the number of eluted phage. The remaining cells were pelleted and resuspended in 25 mL 2YT supplemented with 30 μg/mL chloramphenicol, 30 μg/mL tetracycline, 100 μg/mL ampicillin, and 1% glucose. After overnight growth, the enriched library was used to produce phage for the subsequent round.

ELISA for gp120 binding

For each phage clone tested, soluble gp120 (0.33 μg, ADA isolate) was coated onto the surface of a MaxiSorp microtiter plate well in PBS (100 μL) for 12 hours at 4 °C. After blocking for 2 hours with 2% milk in PBS (200 μL), equal amounts of concentrated phage previously expressed, precipitated, and quantified by ELISA according to the procedures described above, were loaded in 2% milk PBST (100 μL) and incubated at 37 °C for 2 hours. After washing 8X with 200 μL PBST over 10 minutes, an HRP conjugated anti-M13 polyclonal antibody was added in 2% milk PBST (105 μL) and incubated at 37 °C for 2 hours. After washing 8X with 200 μL PBST over 10 minutes, QuantaBlu fluorogenic substrate (110 μL) was added and the ELISA signal was determined using a fluorescence plate reader.

RESULTS AND DISCUSSION

The effect of PTMs on a given protein's structure and function depends on the site(s) at which the protein is modified. Unfortunately, the generation of specific posttranslationally modified protein variants is limited both by the lack of heterologous expression systems that accommodate PTMs and by the target protein sequence requirements of modifying enzymes (transferases). These limitations can be circumvented by the cotranslational incorporation into proteins of free amino acids that contain a modification corresponding to the desired PTM (18,24,25). In this strategy, a PTM is genetically encoded at any site by unique nonsense or frameshift codons. Thus, PTMs can be added at virtually any site in a protein, and PTM-containing proteins can be evolved beyond their traditional sequence constraints, resulting in new modified sequences that nature has not explored. We applied this strategy to the study and evolution of the doubly sulfated 412d using an expanded genetic code that allows for the direct incorporation of sulfotyrosine in *E. coli* (18,20).

Expression of 412d Variants

HIV's coreceptor, CCR5, contains several tyrosine-sulfates at its N-terminus, and those at residues 10 and 14 facilitate HIV entry (26). The IgG antibody 412d binds HIV in part by mimicking CCR5 and therefore contains two tyrosine-sulfates (positions H100 and H100c) required for its gp120 binding activity (13). Structural data shows that these sulfate groups increase affinity for gp120 through strong salt bridge and hydrogen bonding interactions (12), but the contribution of these interactions to binding energy is unknown. We were interested in determining both the independent and combined effects of these two PTMs on gp120 binding by 412d. Because traditional methods for posttranslational sulfation cannot produce 412d variants where only one of two tyrosines is sulfated, we turned to genetic methods in which an orthogonal sTyrRS/tRNA pair, which specifies the incorporation of sulfotyrosine in response to the TAG codon in *E. coli*, could be used to generate all possible sulfated forms of 412d – two sulfates (412d-sYsY), one sulfate at either location (412d-H100sY and 412d-H100csY), and no sulfates (412d-YY).

We constructed a plasmid (pBC-412d2TAG) containing TAG codons at the two locations of sY (positions H100 and H100c) for the expression of doubly sulfated 412d (412d-sYsY) as a soluble Fab from a bicistronic construct under the control of the lac promoter. Three variants of this plasmid were also generated: pBC-412dYY encodes tyrosines at both positions H100 and H100c, pBC-412dTAGH100 encodes sulfotyrosine at position H100 and tyrosine at position H100c, and pBC-412dTAGH100c encodes a tyrosine at position H100 and a sulfotyrosine at position H100c. These plasmids were then transformed into a Top10F′ *E. coli* strain containing pSUPAR6-L3-3SY, a plasmid that encodes the orthogonal sTyrRS/tRNA pair (19). Cultures grown from individual colonies were used for protein expression in TB media containing 10 mM sulfotyrosine (except in the case of pBC-412dYY), 30 μg/mL chloramphenicol, and 100 μg/mL ampicillin. After induction and 36 hours of shake flask expression at 18 °C, cells were harvested, their periplasms were lysed, and Fabs were purified using Protein G affinity chromatography, yielding 0.2, 0.25, 0.5, and 0.7 mg/L for 412d-sYsY, 412d-H100sY, 412d-H100csY, and 412d-YY, respectively. These yields are all within the range expected for Fabs, but differ significantly across the variants. This is due in part to the fact that TAG suppression by the evolved sTyrRS/tRNA pair is not as efficient as sense codon suppression by natural aaRS/tRNA pairs, and to the fact that permeability of sY into cells is low (18) (though despite these effects, peptides containing up to 5 sulfotyrosines have been efficiently expressed (17)). However, the yields of 412d-sYsY and 412d-H100sY were similar to each other but lower than the yields of 412d-H100csY and 412d-YY. Since 412d-H100sY and 412d-H100csY both require suppression of one TAG codon, the discrepancy between their yields suggests that suppression at position H100 is likely the limiting step for the translation of 412d-sYsY. Therefore, silent mutagenesis of the codons surrounding position H100 may lead to higher expression levels as TAG suppression is known to be context dependent (27, 28).

PAGE analysis was used to confirm the purity of the various modified proteins and ESI-LCMS analysis showed single [M+H] masses of 48,817 Da ($[M+H]_{calculated}$ = 48,823 Da), 48,736 Da ($[M+H]_{calculated}$ = 48,743 Da), 48,738 Da ($[M+H]_{calculated}$ = 48,743 Da), and 48,665 Da ($[M+H]_{calculated}$ = 48,663 Da) for 412d-sYsY, 412d-H100sY, 412d-H100csY, and 412d-YY, respectively. In each case, the observed masses are within the ESI error range of the calculated masses and confirm the presence of two sulfates for 412d-sYsY, and one sulfate for 412d-H100sY and 412d-H100csY.

Binding Kinetics and Free Energy Analysis of 412d Variants

We next examined the kinetics of gp120 binding by the four 412d variants using SPR. Soluble gp120 (ADA isolate) was immobilized on a BIAcore biosensor chip and the association (k_{on})

and dissociation (k_{off}) rate constants for each of the 412d Fabs were measured (Figure 1), from which dissociation equilibrium (K_d) constants were calculated (Table 1). Consistent with the recent X-ray crystal structure of the gp120/CD4/412d complex (12), which shows that the two sulfotyrosines in 412d bind gp120 in dissimilar ways, we observed a large difference between the effects of the two sulfates on the kinetics of gp120 binding. For example, sulfation at position H100c increases the affinity of 412d for gp120 approximately 212-fold whereas sulfation of 412d at position H100 only improves affinity by approximately 4.5-fold. The strength of the single interaction at H100c can be explained by the fact that sY H100c is mostly buried, with the sulfate group making a salt bridge interaction with $Arg298^{gp120}$ and the oxygens of the sulfate moiety making hydrogen bonding interactions with $Asn302^{gp120}$ and $Thr303^{gp120}$ side chains as well as the backbone amides of residues 302^{gp120}, 303^{gp120}, and 441^{gp120}. In contrast, sY H100 is largely exposed, with the sulfate group making only peripheral electrostatic interactions with gp120. The binding interactions between H100c's sulfate and gp120 are reminiscent of the sulfate-mediated interactions in the extensively characterized sulfo-hirudin/thrombin complex, whose structure shows a salt bridge and an extensive network of hydrogen bonds between the tyrosine-sulfate of hirudin and thrombin (10). However, in this case, sulfation is responsible for only a 10-fold increase in affinity (18,29). It is likely that because H100c's sulfate is buried, the value of the sulfate-mediated interactions, especially the salt bridge, is greater than those in the sulfo-hirudin/thrombin complex. The doubly sulfated variant, 412d-sYsY, has a 500-fold decrease in K_d from that of 412d-YY, yielding a 0.64 nM antibody with an exceptionally slow off rate (6.76×10^{-5} s^{-1}) that gives the 412d/gp120 complex a half-life of 171 minutes.

The kinetic data can be used to determine the ΔΔG attributable to each sulfate in 412d – *i.e*, $[\Delta G_{412d\text{-}H100sY} - \Delta G_{412d\text{-}YY}]$ or $[RTln(K_{d412d\text{-}H100sY}/K_{d412d\text{-}YY})]$ gives the free energy contribution of position H100's sulfate. The contribution to gp120 binding by the sulfate at position H100 at 25 °C is −0.889 kcal/mol; the contribution to binding by the sulfate at position H100c is −3.142 kcal/mol; and the contribution to binding by both sulfates when simultaneously present is −3.683 kcal/mol. This analysis reveals that the two sulfates bind in a largely independent manner, though a slight deviation (0.348 kcal/mol) from additivity is observed. Since 412d-sYsY is a CD4-induced antibody whose affinity for gp120 increases by ~2-fold (−0.41 kcal/mol at 25 °C) upon addition of CD4 (the primary receptor for gp120 binding) (13), it is possible that CD4's association with gp120 would restore complete additivity of the two sulfates.

412d Library Construction and Phage Selection

The sequence determinants of sulfotransferases limit the range of sulfated antibodies that the immune system can generate. We reasoned that because the tyrosine-sulfates of 412d have such a large effect on its activity, 412d sulfation would occur even if the surrounding sequence determinants necessary for sulfation were not ideal for gp120 binding. Therefore, we asked whether there exist sulfated 412d variants that would not result from natural posttranslational sulfation but that might have equivalent or enhanced binding to gp120. Toward this end, we used SY-X-*E. coli*, which allows multivalent phage display of sulfotyrosine-containing proteins, to subject a phage displayed library of sulfated 412d mutants to selection for gp120 binding. Our library contained two fixed sulfotyrosines at positions H100 and H100c and several randomized surrounding residues. We hypothesized that by randomizing positions that likely determine the posttranslational sulfation of 412d in mammalian cells (these include acidic residues in proximity to the locations of sulfation), we would explore sulfated sequences that nature had not.

Based on the recent crystal structure of gp120 bound to 412d (12), we chose several residues in the vicinity of the sulfotyrosines of 412d for randomization (Figure 2). Three criteria were

used to guide this process: first, the residues should be within 5 amino acids of either tyrosine-sulfate; second, acidic residues are preferential; and third, residue side chains should point toward gp120. The first two criteria were set because we wanted to evolve beyond known consensus features for sulfation. Since the sequence determinants for posttranslational sulfation are in close proximity to the site of sulfation and always include acidic residues, choosing these positions for randomization would most likely access sequences beyond the natural motifs for sulfation (16,21). The third criterion was expected to increase the likelihood of mutations that would directly contribute to gp120 interactions. We therefore chose the following randomization pattern: …P N D sY N D sYA P E E G M S W… (residues H97-H100k are shown), in which the two sulfotyrosines are encoded by TAGs, and codons corresponding to the underlined residues are subject to site-saturation mutagenesis using NNK (N = A, T, G, C; K = G, T). This randomized section was generated and assembled into the scFv form of 412d by overlap PCR. It was then inserted into the pSEX phagemid (which contains a phage origin that allows the production and packaging of single stranded DNA) as an N-terminal fusion to the phage coat protein, pIII, yielding pSEX-412dTAG2Lib. The phagemid library was transformed into SY-X-*E. coli* for phage display with 21 amino acids, affording a maximal diversity of 2.0×10^9.

To generate the phage library, members of which display a 412d mutant with two fixed sulfotyrosines, we infected SY-X-*E. coli* cells containing pSEX-412dTAG2Lib with hyperphage and produced phage in media supplemented with 10 mM sulfotyrosine. Since the only source of pIII in multivalent phage display is the library fusion protein, all phage produced should contain a doubly sulfated 412d variant displayed on its surface. This was confirmed by sequencing of 30 clones from the initial phage pool, which revealed the presence of two TAG codons in each clone. A purified phage preparation (1.6×10^9) was then incubated with immobilized gp120 and washed. Phage that remained bound were eluted with trypsin (a trypsin site joins the displayed antibody with pIII) and used to infect SY-X-*E. coli* cells for the next round. In total, four rounds of selection were conducted, during which a significant enrichment in the percentage of eluted phage was observed (Table 2). Sequencing of the selected phage population showed a collection of doubly sulfated clones (a representative set is shown in Table 3), with 131-1 occurring most frequently. Phage from clone 131-1 and seven other selected clones (131-2 to 131-8) from after rounds 3 and 4 (both rounds 3 and 4 showed similar loaded/eluted ratios) were then individually produced in the presence and absence of sulfotyrosine in the expression media. In all 8 cases, phage production was dependent on the presence of sulfotyrosine (Table 4), confirming that the doubly sulfated mutants were properly displayed. The 8 selected phage clones were then tested individually by ELISA. As expected, all exhibit gp120 binding over the control antigen BSA (Figure 3). However, 131-1, which corresponds to the most frequently selected clone, did not afford the highest ELISA signal, which was instead observed for clone 131-3. This can be explained by the fact that clone 131-1 expresses much better than clone 131-3 with yields of 5.0×10^8 phage/mL versus 8.0×10^7 phage/mL, respectively. Therefore, the lower gp120 binding function of 131-1 is more than compensated by its higher phage yield. Clones 131-1 and 131-3 were both selected for further characterization.

The V_H and V_L regions of 131-1 and 131-3 were inserted into the pBC vector, yielding pBC-131-1 and pBC-131-3 for the expression of soluble Fabs 131-1-sYsY and 131-3-sYsY. As with the 412d variants previously described, pBC-131-1 and pBC-131-3 were cotransformed with pSUPAR6-L3-3SY into Top10F′ *E. coli* cells for shake flask expression (36 hours at 18 °C). Periplasmic lysis and purification using Protein G affinity chromatography gave protein yields of 0.65 and 0.3 mg/L for 131-1-sYsY and 131-3-sYsY, respectively. PAGE analysis confirmed the purity of the Fabs, and ESI-LCMS analysis revealed [M+H] masses of 48,792 Da for 131-1-sYsY ($[M+H]_{calculated}$ = 48,798 Da) and 48,839 Da for 131-3-sYsY ([M

$+H]_{calculated}$ = 48,847 Da). In both cases, the presence of two sulfates is observed in the mass spectra.

Binding Kinetics and Analysis of Selected Mutants

We determined the gp120 binding kinetics of 131-1-sYsY and 131-3-sYsY by SPR (Figure 4); the k_{on}, k_{off}, and K_d values for 131-1-sYsY and 131-3-sYsY are presented in Table 1. Comparison of these kinetic constants to those for 412d-sYsY shows that the best Fab, 131-3-sYsY, has a K_d that is slightly lower than that of 412d-sYsY, suggesting that selection achieved a doubly sulfated gp120 binding antibody as good as 412d-sYsY but not significantly better. The amino acid sequence for the Fab 131-3-sYsY is …P N F sY T G sYA E N E G Y S W… (residues H97-H100k are shown), where the underlined residues correspond to mutations from the original 412d sequence. In this selected sequence, three of the acidic residues subject to randomization were replaced by uncharged ones, which suggests that this sequence would not be the result of sulfation in mammalian cells since it no longer adheres to the known consensus features of sulfation (16,21). This was supported by bioinformatic analysis of 131-3-sYsY's full heavy chain sequence using the Sulfinator algorithm (www.expasy.org/tools/sulfinator/), a query of which predicted that a variant of 131-3-sYsY with tyrosines in the place of sulfotyrosines would not be sulfated in mammalian cells (see Supporting Information).

We note that the evolved 131-3-sYsY binds gp120 with a K_d of 0.51 nM, which is quite similar to the K_d of 412d-sYsY (0.64 nM). We initially expected a much larger improvement to 412d-sYsY's affinity for gp120, but a comparison of the k_{off} rate constants of 131-3-sYsY and 412d-sYsY (1.24×10^{-4} s^{-1} and 6.76×10^{-5} s^{-1}, respectively) reveals that their complex half-lives (93 minutes and 171 minutes, respectively) both exceed 30 minutes, which was the duration of the washing phase of selection. This means that selection was essentially neutral in the k_{off} range of 412d-sYsY, thus resulting in no improvement to 412d-sYsY's k_{off} by the selected hits, though an improvement in k_{on} was obtained. Selections with a much longer washing phase (several hours) are likely required to yield sulfated antibodies that significantly outperform 412d-sYsY in gp120 binding. Nonetheless, the selection of subnanomolar sulfated antibodies against gp120 that would not likely be the result of PTM highlights the utility of using an expanded genetic code to evolve tyrosine-sulfated proteins beyond the known sequence constraints for posttranslational sulfation.

CONCLUSION

Mutagenesis and phage display using an expanded genetic code represents a unique method for studying the biochemistry associated with sulfated proteins as it allows for the efficient production and evolution of homogenously modified mutants, including those that cannot be accessed posttranslationally. To demonstrate the utility of this approach, we selectively mutated 412d and found that the two tyrosine-sulfates of 412d contribute approximately additively to increase its affinity for gp120 by 500-fold. We also evolved 412d beyond the sequence determinants required for its posttranslational sulfation, yielding novel doubly sulfated antibodies from selection, one of which binds gp120 in the subnanomolar range. This strategy should be general to other biologically important PTMs such as glycosylation, phosphorylation, and methylation where mutagenesis and adaptation too can be limited by the efficiency and the sequence constraints of endogenous modification. Engineering novel orthogonal aaRS/tRNA pairs that allow the efficient incorporation of these PTMs is underway in our laboratory.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Antha Mack, Raymond Melikian, Jeremy Mills, Susan Cellitti, and Bernhard Geierstanger for helpful discussions and experimental assistance.

References

1. Xie J, Schultz PG. Adding amino acids to the genetic repertoire. Current opinion in chemical biology 2005;9:548–554. [PubMed: 16260173]
2. Bernimoulin MP, Zeng XL, Abbal C, Giraud S, Martinez M, Michielin O, Schapira M, Spertini O. Molecular basis of leukocyte rolling on PSGL-1. Predominant role of core-2 O-glycans and of tyrosine sulfate residue 51. J Biol Chem 2003;278:37–47. [PubMed: 12403782]
3. Pouyani T, Seed B. PSGL-1 recognition of P-selectin is controlled by a tyrosine sulfation consensus at the PSGL-1 amino terminus. Cell 1995;83:333–343. [PubMed: 7585950]
4. Fong AM, Alam SM, Imai T, Haribabu B, Patel DD. CX3CR1 tyrosine sulfation enhances fractalkine-induced cell adhesion. J Biol Chem 2002;277:19418–19423. [PubMed: 11909868]
5. Farzan M, Babcock GJ, Vasilieva N, Wright PL, Kiprilov E, Mirzabekov T, Choe H. The role of post-translational modifications of the CXCR4 amino terminus in stromal-derived factor 1 alpha association and HIV-1 entry. J Biol Chem 2002;277:29484–29489. [PubMed: 12034737]
6. Kehoe JW, Bertozzi CR. Tyrosine sulfation: a modulator of extracellular protein-protein interactions. Chem Biol 2000;7:R57–61. [PubMed: 10712936]
7. Preobrazhensky AA, Dragan S, Kawano T, Gavrilin MA, Gulina IV, Chakravarty L, Kolattukudy PE. Monocyte chemotactic protein-1 receptor CCR2B is a glycoprotein that has tyrosine sulfation in a conserved extracellular N-terminal region. J Immunol 2000;165:5295–5303. [PubMed: 11046064]
8. Dorfman T, Moore MJ, Guth AC, Choe H, Farzan M. A tyrosine-sulfated peptide derived from the heavy-chain CDR3 region of an HIV-1-neutralizing antibody binds gp120 and inhibits HIV-1 infection. J Biol Chem 2006;281:28529–28535. [PubMed: 16849323]
9. Farzan M, Mirzabekov T, Kolchinsky P, Wyatt R, Cayabyab M, Gerard NP, Gerard C, Sodroski J, Choe H. Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. Cell 1999;96:667–676. [PubMed: 10089882]
10. Liu CC, Brustad E, Liu W, Schultz PG. Crystal structure of a biosynthetic sulfo-hirudin complexed to thrombin. Journal of the American Chemical Society 2007;129:10648–10649. [PubMed: 17685615]
11. Veldkamp CT, Seibert C, Peterson FC, De la Cruz NB, Haugner JC 3rd, Basnet H, Sakmar TP, Volkman BF. Structural basis of CXCR4 sulfotyrosine recognition by the chemokine SDF-1/CXCL12. Science signaling 2008;1:ra4. [PubMed: 18799424]
12. Huang CC, Lam SN, Acharya P, Tang M, Xiang SH, Hussan SS, Stanfield RL, Robinson J, Sodroski J, Wilson IA, Wyatt R, Bewley CA, Kwong PD. Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. Science (New York, NY) 2007;317:1930–1934.
13. Choe H, Li W, Wright PL, Vasilieva N, Venturi M, Huang CC, Grundner C, Dorfman T, Zwick MB, Wang L, Rosenberg ES, Kwong PD, Burton DR, Robinson JE, Sodroski JG, Farzan M. Tyrosine sulfation of human antibodies contributes to recognition of the CCR5 binding region of HIV-1 gp120. Cell 2003;114:161–170. [PubMed: 12887918]
14. Choe H, Moore MJ, Owens CM, Wright PL, Vasilieva N, Li W, Singh AP, Shakri R, Chitnis CE, Farzan M. Sulphated tyrosines mediate association of chemokines and Plasmodium vivax Duffy binding protein with the Duffy antigen/receptor for chemokines (DARC). Molecular microbiology 2005;55:1413–1422. [PubMed: 15720550]
15. Hans D, Pattnaik P, Bhattacharyya A, Shakri AR, Yazdani SS, Sharma M, Choe H, Farzan M, Chitnis CE. Mapping binding residues in the Plasmodium vivax domain that binds Duffy antigen during red cell invasion. Molecular microbiology 2005;55:1423–1434. [PubMed: 15720551]
16. Seibert C, Sakmar TP. Toward a framework for sulfoproteomics: Synthesis and characterization of sulfotyrosine-containing peptides. Biopolymers 2008;90:459–477. [PubMed: 17680702]
17. Choe H, Farzan M. Chapter 7. Tyrosine sulfation of HIV-1 coreceptors and other chemokine receptors. Methods in enzymology 2009;461:147–170. [PubMed: 19480918]

18. Liu CC, Schultz PG. Recombinant expression of selectively sulfated proteins in Escherichia coli. Nature biotechnology 2006;24:1436–1440.

19. Liu CC, Cellitti SE, Geierstanger BH, Schultz PG. Efficient expression of tyrosine-sulfated proteins in E. coli using an expanded genetic code. 2009 Submitted.

20. Liu CC, Mack AV, Tsao ML, Mills JH, Lee HS, Choe H, Farzan M, Schultz PG, Smider VV. Protein evolution with an expanded genetic code. Proceedings of the National Academy of Sciences of the United States of America 2008;105:17688–17693. [PubMed: 19004806]

21. Moore KL. The biology and enzymology of protein tyrosine O-sulfation. J Biol Chem 2003;278:24243–24246. [PubMed: 12730193]

22. Rondot S, Koch J, Breitling F, Dubel S. A helper phage to improve single-chain antibody presentation in phage display. Nature biotechnology 2001;19:75–78.

23. Broders O, Breitling F, Dubel S. Hyperphage. Improving antibody presentation in phage display. Methods in molecular biology (Clifton, NJ) 2003;205:295–302.

24. Xie J, Supekova L, Schultz PG. A genetically encoded metabolically stable analogue of phosphotyrosine in Escherichia coli. ACS chemical biology 2007;2:474–478. [PubMed: 17622177]

25. Neumann H, Hazen JL, Weinstein J, Mehl RA, Chin JW. Genetically encoding protein oxidative damage. Journal of the American Chemical Society 2008;130:4028–4033. [PubMed: 18321101]

26. Cormier EG, Persuh M, Thompson DA, Lin SW, Sakmar TP, Olson WC, Dragic T. Specific interaction of CCR5 amino-terminal domain peptides containing sulfotyrosines with HIV-1 envelope glycoprotein gp120. Proceedings of the National Academy of Sciences of the United States of America 2000;97:5762–5767. [PubMed: 10823934]

27. Albertini S, Chetelat AA, Miller B, Muster W, Pujadas E, Strobel R, Gocke E. Genotoxicity of 17 gyrase- and four mammalian topoisomerase II-poisons in prokaryotic and eukaryotic test systems. Mutagenesis 1995;10:343–351. [PubMed: 7476271]

28. Bossi L. Context effects: translation of UAG codon by suppressor tRNA is affected by the sequence following UAG in the message. Journal of molecular biology 1983;164:73–87. [PubMed: 6188841]

29. Stone SR, Hofsteenge J. Kinetics of the inhibition of thrombin by hirudin. Biochemistry 1986;25:4622–4628. [PubMed: 3768302]

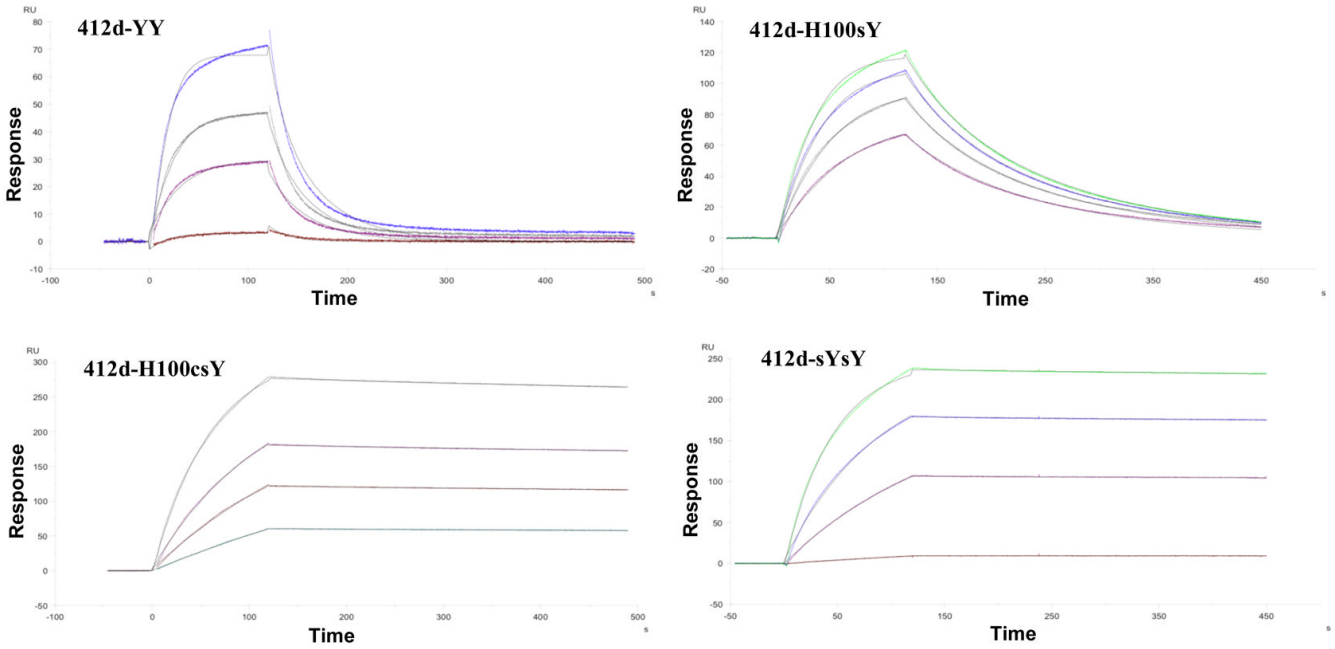

Figure 1.

SPR sensorgrams (determined at 25 °C) for 412d-YY, 412d-H100sY, 412d-H100csY, and 412d-sYsY showing binding to gp120. Colored lines represent sensorgram data and overlaid black lines represent nonlinear fits.

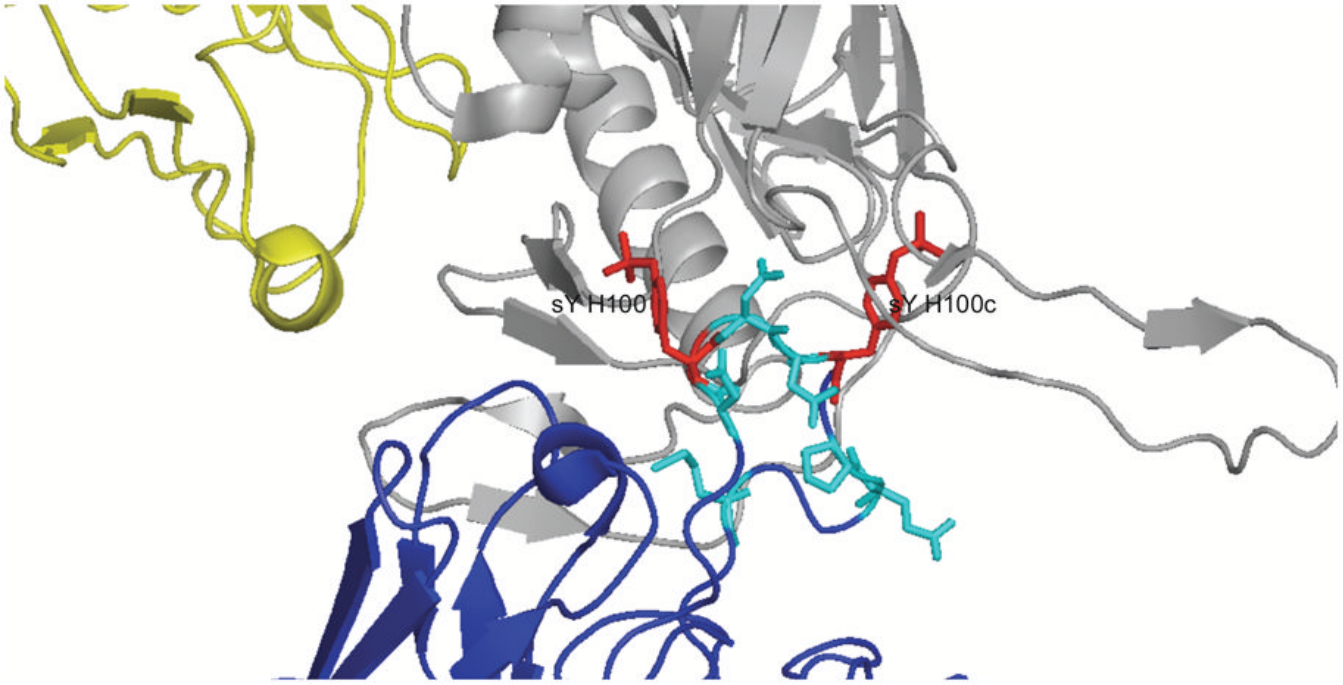

Figure 2.
Sulfotyrosines (red) at the binding interface between 412d (blue) and gp120 (grey). Residues chosen for library randomization by site-saturation mutagenesis are colored cyan. CD4 is colored yellow. (Adapted from PDB: 2QAD.)

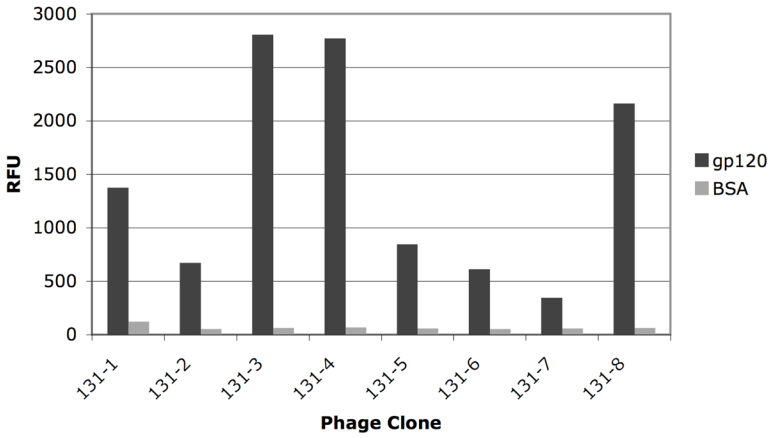

Figure 3.
Binding of selected phage clones to gp120 and BSA based on ELISA analysis (only single measurements were taken). For each sample, 3.2×10^7 phage particles were loaded.

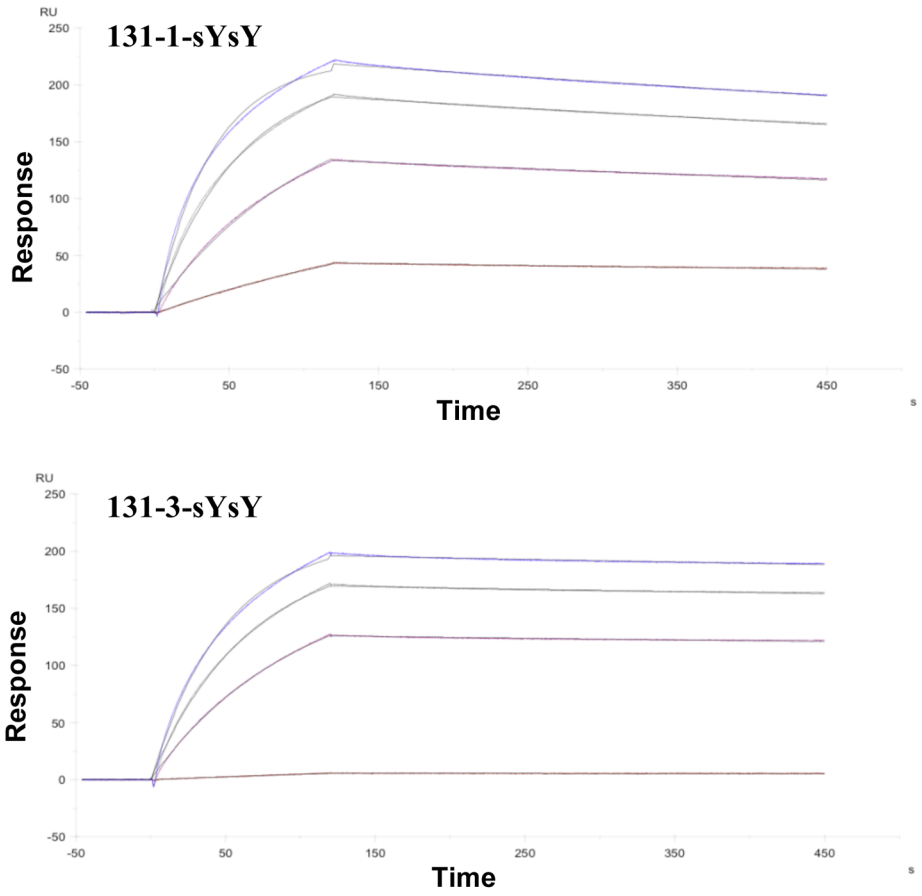

Figure 4.
SPR sensorgrams (determined at 25 °C) for 131-1-sYsY and 131-3-sYsY showing binding to gp120. Colored lines represent sensorgram data and overlaid black lines represent nonlinear fits.

Table 1

Kinetics of Binding between Fabs and gp120

Fab	k_{on} (s^{-1} M^{-1})	k_{off} (s^{-1})	K_d (nM)
412d-YY	4.47×10^5	1.43×10^{-1}	320
412d-H100sY	2.08×10^5	1.49×10^{-2}	71.3
412d-H100csY	8.57×10^4	1.29×10^{-4}	1.51
412d-sYsY	1.06×10^5	6.76×10^{-5}	0.63
131-1-sYsY	1.04×10^5	4.16×10^{-4}	4.00
131-3-sYsY	2.42×10^5	1.24×10^{-4}	0.51

Table 2

Phage Enrichment Statistics

Round	Loaded	Eluted	Loaded/Eluted
1	1.6×10^9	5.7×10^4	2.8×10^4
2	3.2×10^8	4.6×10^4	7.0×10^3
3	4.5×10^8	2.8×10^6	1.6×10^2
4	1.4×10^8	5.2×10^5	2.7×10^2

Table 3

Amino Acid Sequences of Chosen Hits from Selection

Clone	CDR3 sequence (residues H99-H100i shown)
131-1	…A sY N N sY A S E E G F…
131-2	…E sY Q N sY A P R E G L…
131-3	…F sY T G sY A E N E G Y…
131-4	…L sY T G sY A S D E G I…
131-5	…S sY N E sY A K Q E G Q…
131-6	…Q sY N S sY A E P E G K…
131-7	…L sY E M sY A A S E G M…
131-8	…F sY N S sY A E P E G F…
412d-sYsY	…D sY N D sY A P E E G M…
	(not a product of selection, but shown for comparison)

Table 4

Phage Yield Dependence on Sulfotyrosine

	Phage Yield (particles mL^{-1})	
Clone	**+ *Sulfotyrosine***	**− *Sulfotyrosine***
131-1	5.0×10^8	2.0×10^7
131-2	5.0×10^7	2.5×10^5
131-3	8.0×10^7	$< 2.0 \times 10^5$
131-4	3.2×10^8	8.0×10^5
131-5	3.6×10^8	4.0×10^5
131-6	2.8×10^8	$<2.0 \times 10^5$
131-7	3.2×10^8	$<2.0 \times 10^5$
131-8	2.8×10^8	4.0×10^5