Mass Spectrometric Characterization of the Glycosylation Pattern of HIV-gp120 Expressed in CHO Cells

Xuegong Zhu,[‡] Christoph Borchers,[‡] Rachelle J. Bienstock,[§] and Kenneth B. Tomer*,[‡]

Laboratory of Structural Biology and Scientific Computing Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, P.O. Box 12233, Research Triangle Park, North Carolina 27709

Received February 25, 2000; Revised Manuscript Received June 6, 2000

ABSTRACT: An analytical approach is reported for the characterization of the specific glycans found on highly glycosylated proteins based on a combination of specific proteolysis and deglycosylation combined with two different mass spectrometric approaches, matrix-assisted laser desorption/ionization mass spectrometry, and nanoelectrospray mass spectrometry/tandem mass spectrometry using a hybrid quadrupole-time-of-flight tandem mass spectrometer. The high resolution and mass accuracy of the mass spectrometric data obtained on the hybrid instrument combined with the high parent mass capabilities are shown to be extremely useful in the site-specific assignment of heterogeneous glycans. Using this methodology, 25 of 26 consensus glycosylation sites on HIV-1_{SF2} gp120, expressed in Chinese hamster ovary cells, could be assigned. Good correlations between the relative abundances of members of heterogeneous series in the matrix-assisted laser desorption/ionization mass spectra and the nanoelectrospray mass spectra were observed, indicating that the mass spectrometric data reflected the actual abundances of the members of the series. These data were incorporated with molecular modeling based on the solved structure of a mutant truncated, highly deglycosylated gp120 to propose a structural model for the completely glycosylated form.

There have been a variety of approaches to the characterization of glycopeptides by mass spectrometry (MS)[1] in the literature. The most recently utilized methods rely upon either matrix-assisted laser desorption/ionization (MALDI) or electrospray ionization (ESI).

MALDI has been employed by several groups (*1*, *2*). Yang and Orlando have reported on the identification of glycopeptides after enzymatically digesting the glycoprotein. The digest was first analyzed by MALDI, then enzymatically deglycosylated by PNGase F and reanalyzed by MALDI. Comparison of the MALDI spectra before and after deglycosylation provided the molecular weight of the carbohydrate chain. Identification of the peptide chain was based on the known sequence and the mass of the peptide after deglycosylation. The peptide masses, however, needed to be corrected for the transformation of glycosylated Asn to Asp during the enzymatic deglycosylation (plus 1 Da mass difference) (*3*). Juhasz et al., used nonspecific proteases, such as Pronase, to digest the glycoprotein, followed by high mass accuracy delayed extraction MALDI measurements (*3*). The glycans were identified by subtracting the masses of the possible combinations of carbohydrate residues from the observed glycopeptide masses and then comparing the results with peptide sequences from the protein. Tsarbopoulos et al. used post source decay (PSD) analysis of the glycopeptide ions obtained by ESI of enzymatic digests to provide information about the glycans (*4*). These workers observed that the PSD spectra were dominated by carbohydrate ions with little amino acid sequence information. HPLC separation of reduced *S*-alkylated, Lys C-digested human urinary erythropoietin followed by MALDI analysis before and after neuraminidase and endoprotease digestion provided site-specific information about glycan identification and heterogeneity (*5*).

Several groups have also investigated the utility of electrospray ionization (ESI) MS for glycoprotein analysis. Carr and co-workers have published several papers in which they explored the use of ions characteristic of specific glycans, for example, the *m*/*z* 204 ion arising from the GN residue, generated either by high cone voltage or MS/MS during an LC/ESI/MS analysis to identify glycopeptides (*6*, *7*). In these studies, treatment with selective deglycosylase enzymes was used to differentiate between N- and O-linked glycans. Burlingame and co-workers have also used LC/ESI/MS with sequential exoglycosidase digestion (*8*, *9*) and carbohydrate-specific ion formation monitoring with sequential exoglycosidase digestion to characterize the site-specific glycosylation of glycoproteins (*10*, *11*).

* To whom correspondence should be addressed. Dr. Kenneth B. Tomer, National Institute of Environmental Health Sciences, National Institutes of Health, P.O. Box 12233 MD F0-03, Research Triangle Park, NC 27709. Phone: (919) 541-1966; fax: (919) 541-0220; e-mail: tomer@niehs.nih.gov.

[‡] Laboratory of Structural Biology.

[§] Scientific Computing Laboratory.

[1] Abbreviations: HIV, human immunodeficiency virus; MS, mass spectrometry; MALDI, matrix-assisted laser desorption/ionization; ESI, electrospray ionization; PNGase F, peptide N-glycosidase F; PSD, post source decay; HPLC, high performance liquid chromatography; GN, *N*-acetylglucosamine; MS/MS, tandem mass spectrometry; LC, liquid chromatography; CHO, Chinese hamster ovary; SIV, simian immunodeficiency virus; FAB, fast atom bombardment; Q-Tof, quadrupole/time-of-flight tandem mass spectrometer; DTT, dithiothreitol; RCM, reduced, carboxymethylated; M, mannose; G, galactose; nano-ESI, nanoelectrospray; MBL, mannose binding lectin.

10.1021/bi000432m
Published on Web 08/22/2000

Our laboratory has been applying mass spectrometric methodologies to the determination of structural features relevant to HIV infection, including fine mapping of epitopes on HIV proteins recognized by antibodies and structural aspects of the interactions of HIV envelope glycoprotein gp120 with cellular receptors. A key step in these studies is the correlation of the observed masses of the proteolytic products with the DNA-derived amino acid sequence of the protein. This approach is relatively straightforward for nonglycosylated proteins of known sequence. In the case of a glycosylated protein, however, there is no simple correspondence between the observed mass and the amino acid sequence of a proteolytic peptide from that protein when the proteolytic peptide contains one or more glycosylation sites.

Our recent efforts have been directed toward probing the interaction of HIV gp120 with cellular receptors and antibodies (*12*−*14*). HIV gp120 is the surface glycoprotein of the HIV virus that binds to CD4 and a second chemokine receptor prior to entrance of the virus into a cell. It is highly glycosylated, with nearly 50% of its mass due to carbohydrate residues. On the basis of the MALDI spectrum of gp120, it is also very heterogeneous with a molecular mass ranging from 90 000 to 110 000 Da (*13*). Furthermore, the system used to express the gp120 significantly impacts the nature of the glycosylation pattern, with the gp120 expressed in baculovirus differing significantly from that in mammalian Chinese hamster ovarian (CHO) cells (*15*). It has been demonstrated that the N-linked glycosylation in the closely related simian immunodeficiency virus (SIV) limits the neutralizing antibody response and shields the virus from immune recognition (*16*).

There have been several reports of studies on the determination of glycosylation sites and of the types of glycans found on gp120. Mizuochi et al. probed the glycan structures found on gp120 cloned from both an HTLV-III B isolate expressed in CHO cells and in lymphoblastoid (H9) cells, and found high mannose, hybrid, and complex structures to be present (*17*, *18*). These workers estimated that more than 100 different glycans are present, with many still unidentified. Leonard et al. have reported the glycosylation sites on gp120 expressed in CHO cells. Tryptic peptides were identified by FAB-MS after enzymatic deglycosylation, and the glycan type was determined by comparison of the tryptic map before and after specific enzymatic deglycosylation (*19*). More detailed glycosylation patterns at each site were not identified. Yeh et al. also determined the specific glycosylation sites and oligosaccharide structures of HIV-1 gp120 BH8 expressed in baculovirus using Con-A Sepharose to extract the glycopeptides from a tryptic digest (*20*). They found only high-mannose glycans, which may have been a reflection of the preferential binding of high-mannose glycans to Con-A relative to the binding of hybrid and complex glycans (*21*), in addition to or instead of differences in glycans due to the differences in expression systems. Yeh also observed different extents of processing of the high mannose glycans depending on the site of glycosylation.

We have approached the problem of determining the site-specific glycosylation pattern of recombinant HIV-I_{SF2} gp120 expressed in CHO cells using a combination of enzymatic digestion, HPLC separation, MALDI and nanoelectrospray/MS and MS/MS on a hybrid quadrupole/time-of-flight (Q-Tof) tandem mass spectrometer. Although the parent ion mass range of peptides from which interpretable MS/MS data can be acquired on most tandem instruments is typically considered to be limited to less than 3000 Da, useful MS/MS data relative to the glycan can be obtained from glycopeptides with masses approaching 6000 Da on the Q-Tof. The high mass resolution and accuracy of data obtained on this instrument, combined with the high parent mass capabilities, makes it extremely attractive for the analysis of complex glycopeptide mixtures from glycoprotein digests.

Using the above approach, we have characterized 25 of the 26 glycosylation sites on HIV-I_{SF2} gp120. The results obtained from the mass spectrometic experiments were then used in conjunction with molecular modeling based on the known crystal structure of a truncated, highly deglycosylated mutant of HIV HxBc2 gp120 (*22*) to propose a structural model for the completely glycosylated intact protein.

EXPERIMENTAL PROCEDURES

Materials. HIV-1_{SF2} gp120 expressed in mammalian Chinese hamster ovary cells was purchased from Austral Biologicals (San Ramon, CA). Trypsin immobilized on POROS media was purchased from PerSeptive Biosystems (Framingham, MA). Dithiothreitol (DTT) was purchased from Schwarz/Mann Biotech (Cleveland, OH). Iodoacetic acid was purchased from Sigma Chemical Co. (St. Louis, MO). Neuraminidase from *Vibrio cholerae* was purchased from Boehringer Mannheim (Indianapolis, IN). Recombinant peptide-N-glycosidase F (PNGase F) was purchased from Oxford GlycoSciences (Abingdon, UK).

Purification and Digestion of gp120. The gp120 was purified by HPLC on a model 1100 HPLC (Hewlett-Packard, Wilmington, DE) using a Vydac 214 TP54 C_4 column (5μm, 4.6 × 250 mm). Peptides were eluted from the column using A: water/0.1% TFA and B: acetonitrile/0.1%TFA with a gradient that changed from 20% B to 60% B over 50 min, followed by a ramp from 60 to 90% B over 10 min. The purified gp120 (100 μg) was reconstituted in 152 μL of 0.36 M Tris-HCl buffer, pH 8.6, containing 8 M urea, 3 mM EDTA, and 10 mM DTT. The mixture was incubated under nitrogen for 4 h at room temperature. Iodoacetic acid was added to a final concentration of 25 mM. The mixture was incubated in the dark for 45 min at room temperature. The reaction was quenched by adding DTT to give a final concentration of 40 mM, and the resulting reduced, carboxymethylated gp120 (RCM-gp120) was promptly purified by HPLC under the same conditions used previously. Tryptic digestion of the RCM-gp120 was carried out using trypsin immobilized on POROS media. The POROS media was washed three times with 0.1 M ammonium bicarbonate containing 50% acetonitrile, followed by three times with 0.1 M ammonium bicarbonate containing 5% acetonitrile. The immobilized trypsin media was then transferred to a solution of RCM-gp120 in 100 μL of 0.1 M ammonium bicarbonate.

Fractionation of the RCM-gp120 Trypic Digest. The tryptic peptides were fractionated using a Vydac 218 TP54 C_{18} column (5μm, 4.6 × 250 mm). Peptides were eluted from the column using A: water/0.1% TFA and B: acetonitrile/0.1%TFA with a gradient that changed from 8% B to 40% B over 50 min, followed by a ramp from 40 to 90% B over

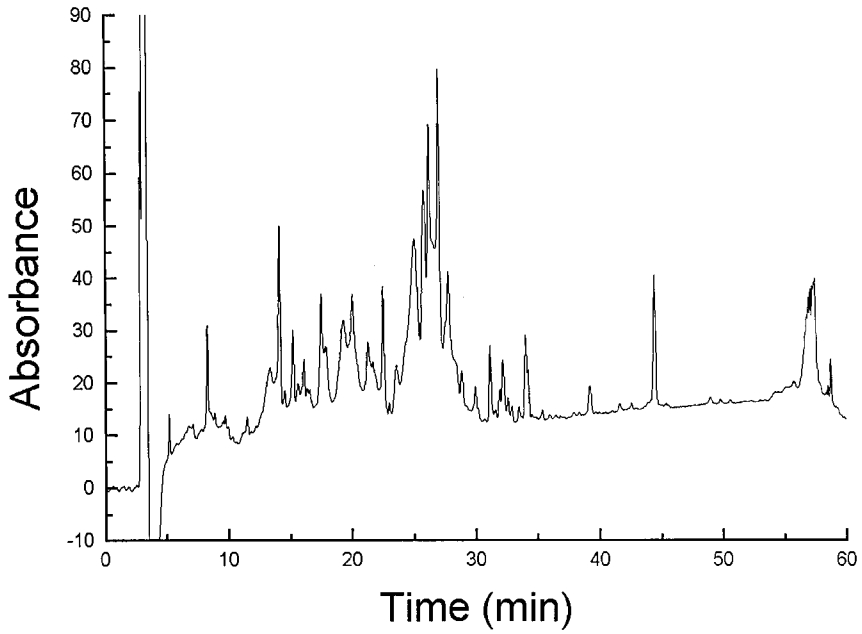


FIGURE 1: HPLC chromatogram of the tryptic digest of rgp120. Conditions are as in Experimental Procedures. One fraction was collected per minute. Fractions in which specific glycopeptides were observed are listed in Table 1. Other peaks in the chromatogram contained nonglycosylated peptides and/or trypsin autolysis products.

10 min. The HPLC chromatogram obtained by variable wavelength detector set at 215 nm is shown in Figure 1. Under these conditions, highly acidic peptides/glycopeptides eluted in the void volume.

Neuraminidase Digestion. Individual HPLC fractions were reconstituted in 10 mL of 50 mM ammonium acetate containing 5 mM $CaCl_2$, pH 6. One milliliter of neuraminidase (0.85 mU) was added, and the mixture was incubated for 24 h at 37 °C.

PNGaseF Digestion. Individual HPLC fractions were reconstituted in 10 mL of 50 mM sodium phosphate buffer, pH 7.5, containing 50 mM EDTA and 0.1% (w/v) sodium azide. One milliliter of PNGase F solution (0.5 U) was added, and the mixture was incubated for 24 h at 37 C.

Sample Cleanup Prior to ESI/MS Analysis. All samples were subjected to cleanup prior to ESI/MS analysis via the method reported by Anderegg and co-workers (*23*). Briefly, a microcolumn was prepared using a gel-loading tip filled with POROS R2 slurry (sufficiently swollen in methanol). The column was washed four times with 0.2% formic acid. The sample was applied to the column which was then washed once with 0.2% formic acid. The peptides were eluted with 75% methanol in 0.2% formic acid.

MALDI. MALDI spectra were obtained on a PerSeptive Voyager RP MALDI/time-of-flight mass spectrometer (PerSeptive Biosystems, Framingham, MA) equipped with a nitrogen laser and at an accelerating voltage of 30 kV. The matrix used was α-cyano-4-hydroxycinnamic acid that had been recrystallized from hot methanol and stored in the dark. All samples were analyzed using an external calibrant covering the analyte mass range. Mass accuracy was ±0.5%. A total of 128 individual scans were averaged.

ESI/MS and ESI/MS/MS. All ESI spectra were obtained on a Q-Tof hybrid mass spectrometer (Micromass, Altrincham, UK). This instrument is a tandem mass spectrometer consisting of a quadrupole mass analyzer for MS1 and an orthogonal time-of-flight analyzer as MS2. This instrument was equipped with a nanoelectrospray Z-spray source. Type A borosilicate glass capillaries (Micromass, Beverly, MA) were used for introduction of nanoelectrospray samples. Experimental conditions for acquisition of nanoelectrospray spectra were as follows: source block temperature, 60 C; desolvation temperature, 80 C; spray voltage, 1400−1500 V; cone voltage, 22−28 V; accelerating voltage, 200 V; TOF voltage, 7200 V. From 50 to 200 individual 2-s scans were averaged.

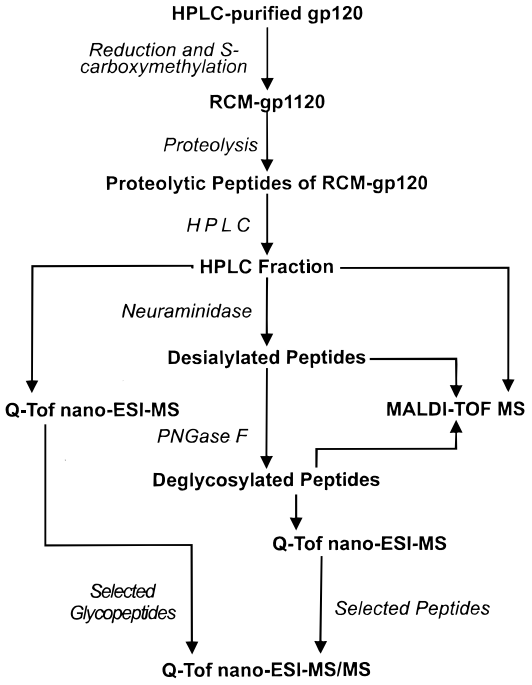
Homology Modeling of the HIV gp120 Protein. The model for the fully glycosylated gp120 protein structure was based on the coordinates for the HIV gp120-CD4-FAB17B antibody complex deposited in the Protein Data Bank as entry 1GC1. The gp120 isolate used in our study is the SF2 isolate, while 1GC1 represents the HXB2 variant. These HIV isolate sequences share 84.5% sequence identity and 87.3% sequence similarity. The homology model constructed is based upon the sequence alignment shown in Figure 1. Coordinates for the missing V1/V2 loop were inserted based upon a Fasta and Blast search of entries in the Protein Data Bank. The structure of the loop representing residues 211−280 of RNAse H from bacteriophage T4 (Databank entry 1TFR) was selected as the structure for the inserted V1/V2 loop. This structure shares 28% sequence homology with the HIV-1_{SF2} V1/V2 loop and was selected by both Fasta and Blast protein searches as the top matching loop represented in the Protein Data Bank. Coordinates for the HIV V3 loop were inserted based on the reported NMR structure [Protein Data Bank Entry: 1CE4 (A chain)] (*24*−*29*). The V3 loop in the 1CE4 structure shared 86% sequence identity to the V3 loop of the HIV isolate used in this study. Recent NMR studies have indicated that the V3 loop conformation is antibody-dependent (*29*) and may be heterogeneous in unbound gp120. As the purpose of our model is to highlight the conformational distribution of glycans on gp120, we have chosen to include in our model a single NMR determined conformation of the V3 loop and to explore flexibility in this loop through

molecular dynamics simulations. The significant flexibility in the V3 loop is subtle and involves the GPGR motif. This subtle change is more significant when discussing conformation in the light of antibody interaction or drug design rather than the distribution of gp120 glycans. An eight-residue loop LNHTEGTK was inserted in the gp120 model based upon the loop RGHTNGTK from the Protein Data Bank entry 1FSU. The model was constructed using the Molecular Simulations Homology program. The system was minimized using the Molecular Simulations Discover 3 molecular dynamics program with the CVFF force field (*30*) to minimize bad steric conflicts, bumps, and overlaps prior to glycosylation of the system.

Construction of Complex and High Mannose Glycosylation Model. The structure of each individual simple sugar and each complex carbohydrate and high mannose glycosylation was constructed using the Molecular Simulations Biopolymer Builder program. Every complex carbohydrate and high mannose structure was then minimized individually and appended to the appropriate asparagine residue (in the minimized gp120 protein model described above) as determined by the mass spectrometric results. As branching within the glycan was not determined in this study, the branched structures previously determined by Mizouchi et al. were used (*18*, *19*). The entire complete glycosylated molecule was minimized initially in vacuo to eliminate bad steric contacts, overlaps, and bumps present from construction. This entire system (12 592 atoms) was solvated in a 5-Å sphere of water and reminimized. Minimizations were conducted using an extended AMBER force field that included the Homans parametrization for oligosaccharides and glycoproteins (*31*, *32*). Calculations were conducted using a distant dependent dielectric and nonperiodic boundary conditions with the cell multipole method to evaluate electrostatic interactions. The RMS deviation between the complete model C_α backbone with *N*-glycans after dynamic minimization and original PDB structure (1GC1) backbone (2024 atoms) excluding the inserted loops is 4.55 Å.

Analytical Approach. The general procedure we have used to identify the glycopeptides from rgp120 is shown in Scheme 1. The protein was first reduced, carbomethoxylated, then proteolytically digested. The protein digest was fractionated by HPLC, and all eluent fractions were collected. Each HPLC fraction was then analyzed by MALDI and by nanoelectrospray/MS. Selected ions of interest observed in the nanoelectrospray spectra are subjected to MS/MS analysis. The HPLC fraction was also treated with neuraminidase to remove any sialic acid residues, and then it was reanalyzed by both MALDI and nanoelectrospray. MS/MS spectra were obtained from any new ions that appeared only upon desialylation. The fractions were also completely deglycosylated by treatment with PNGase F and further analyzed by MALDI and/or nanoelectrospray/MS. Structural assignments for the ions observed in the spectra of the deglycosylated fractions are made based on the mass of the ion and/or on MS/MS spectra obtained for the ion.

Scheme 1: Analytical Methodology



RESULTS

In this section, we will discuss a limited set of specific examples in detail to illustrate our analytical approach. Examples include high mannose, multiple glycosylation sites, hybrid and complex glycans. Consensus glycosylation sites (NxS/T) are underlined in Figure 2.

High Mannose Glycopeptides. The MALDI spectrum and the nanoelectrospray/MS spectrum of HPLC fraction 20 from the tryptic digest are shown in Figure 3, panels a and b, respectively. The MALDI spectrum (Figure 3a) shows two series of peaks separated by 162 Da [designated A (*m*/*z* 3318, 3153, 2991, 2826, and 2664) and B (*m*/*z* 5451, 5289, 5127, and 4965)] and adjacent peaks 32 Da higher in mass. The 32 Da higher peaks are assigned as due to 2-fold oxidation of the peptides. We have previously observed that peptides/proteins containing methionines and/or tryptophanes can undergo oxidation at one or more steps in the isolation procedure, e.g., during lyophilization. The nanoelectrospray spectrum (Figure 3b) shows multiply charged ions corresponding to the MALDI ions of mass 5451 (*m*/*z* 1362.7, +4 charge state), 3318 (*m*/*z* 1103.42, +3 charge state), 3153 (*m*/*z* 1049.42, +3 charge state), and 2991 (*m*/*z* 995.41, +3 charge state) as well as sodiated and oxidized ions corresponding to the peaks 32 Da higher observed in the MALDI spectrum. The components of mass 3318, 3153, and 2991 observed in the MALDI spectrum are separated by 162 ± 2 Da (the mass of a mannose residue), indicating the presence of a peptide glycosylated with differing numbers of mannoses. In the ESI spectrum, the triply charged ions corresponding to these species (*m*/*z* 1103.42, 1049.42, and 995.41) show a separation of 54 Da (162/3) indicative of mass differences due to mannose residues.

An MS/MS spectrum of the highest abundance ion in the nanoelectrospray mass spectrum, *m*/*z* 1103.42 (triply charged from the isotope pattern), was acquired. This ion was also the highest mass ion identified as part of that specific series. The multiply charged MS/MS spectrum was transformed into a singly charged spectrum using MaxEnt software, and the resulting spectrum is shown in Figure 4. This transformed spectrum shows a series of losses of one through nine mannoses ($\Delta m = 162$) from the $(M + H)^+$ ion. From

```
  1 EKLWVTVYYGVPVWKEATTTLFCASDARAYDTEVHNVWATHACVPTDPNP 50  HIVI SF2
    ||||||||||||||||||||||||||:||||||||||||||||||||||
 32 EKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNP 81  HXB2 (CRYSTAL)

 51 QEVVLGNVTENFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNC 100
    ||||| |||||||||||.||||| ||||||||||||||||||||.| |
 82 QEVVLVNVTENFNMWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKC 131

101 TDLGKATNTNSSNWKEEI.KGEIKNCSFNITTSIRDKIQKENALFRNLDV 149
    |||   ||||||. :  . ||||||||||.|||| |:||| | |  ||:
132 TDLKNDTNTNSSSGRMIMEKGEIKNCSFNISTSIRGKVQKEYAFFYKLDI 181

150 VPIDNASTTTNYTNYRLIHCNRSVITQACPKVSFEPIPIHYCTPAGFAIL 199
    :||||     . |.|:|  || |||||||||||||||||| |||||||
182 IPIDN.....DTTSYKLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAIL 226

200 KCNNKTFNGKGPCTNVSTVQCTHGIRPIVSTQLLLNGSLAEEEVVIRSDN 249
    |||||||| ||||||||||||||||:|||||||||||||||||||| |
227 KCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSVN 276

250 FTNNAKTIIVQLNESVAINCTRPNNNTRKSIYI..GPGRAFHTTGRIIGD 297
    ||.|||||||||| || |||||||||||| | |  |||||| | |: ||.
277 FTDNAKTIIVQLNTSVEINCTRPNNNTRKRIRIQRGPGRAFVTIGK.IGN 325

298 IRKAHCNISRAQWNNTLEQIVKKLREQFGNNKTIVFNQSSGGDPEIVMHS 347
    .|.|||||||.|||||.||  ||||||||||||:| |||||||||| ||
326 MRQAHCNISRAKWNNTLKQIASKLREQFGNNKTIIFKQSSGGDPEIVTHS 375

348 FNCRGEFFYCNTTQLFNNTWRLN..HTEG...TKGNDTIILPCRIKQIIN 392
    ||| |||||||.|||||.||  .   |||   |.|.||| ||||||||||
376 FNCGGEFFYCNSTQLFNSTWFNSTWSTEGSNNTEGSDTITLPCRIKQIIN 425

393 MWQEVGKAMYAPPIGGQISCSSNITGLLLTRDGGTNVTNDTEVFRPGGGD 442
    |||.|||||||||| ||| |||||||||||||| |  |:.|:|||||||
426 MWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGG.NSNNESEIFRPGGGD 474

443 MRDNWRSELYKYKVIKIEPLGIAPTKAKRRVVQREKR 479
    |||||||||||||:||||||:||||||||||||||||
475 MRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKR 511
```

FIGURE 2: Sequence alignment comparing the HIV SF2 GP120 with the HXB2 isolate (crystal structure PDB Entry: 1GC1). The sequence alignment was performed using the Wisconsin GCG package program Bestfit with the Blosum62 comparison matrix. The sequences are 87% similar and 84% identical. (|) indicates identical residues and (:) and (.) indicate closely related residues, respectively. Consensus glycosylation sites are underlined.

previous reports, nine mannoses are the maximum number observed on a high mannose glycan in gp120 (*18*, *19*). High mannose glycans have two *N*-acetylglucosamine (GN, mass of residue = 203.08 Da) residues as part of the core of the glycan bound to an Asn. After loss of the 9 mannoses, one GN residue has been reported to be lost (*33*). Thus, these data indicated that the ion of *m*/*z* 1646.78 corresponds to the peptide plus one GN. Subtracting the second GN leaves the peptide chain with a mass of 1443.70. The ions observed in the spectrum with a mass less than that of the peptide chain are primarily due to carbohydrate ions and do not provide amino acid sequence information. From a listing of the predicted tryptic peptide fragments from nonglycosylated gp120, there is only one predicted peptide within ± 50 Da of *m*/*z* 1443.70. That peptide is T29, residues 308−319 (AQWNNTLEQIVK), which has a calculated mass of 1443.76 (41 ppm error). Thus, the high resolution and mass accuracy of the MS/MS data obtained on the QTof hybrid mass spectrometer accurately identifies the peptide.

The ratios of the molecular ion species corresponding to a high mannose glycopeptide with 9, 8, 7, and 6 M are 100: 52:28:17 in the nanoelectrospray mass spectrum (Figure 3, panel b) and 100:51:32:20 in the MALDI mass spectrum (Figure 3, panel a). This demonstrates that there is an excellent correlation between the relative abundances of high mannose species observed by nanoelectrospray and by MALDI and directly indicates the relative abundances of the heterogeneic high mannose glycan. Although suppression effects and differences in ionization efficiencies can affect the relative abundances of ions in both MALDI- and ESI/MS, the relative abundances of ions from structurally similar molecules will show less of an effect. The glycan types found at specific consensus glycosylation sites and their pattern of heterogeneity are listed in Table 1.

Multiple Glycosylation Sites. Several gp120 tryptic peptides contain more than one glycosylation site. An example of this was found in tryptic digest HPLC fraction 25 (Figure 5, panels a and b). A series of triply charged ions (*m*/*z* 1977/

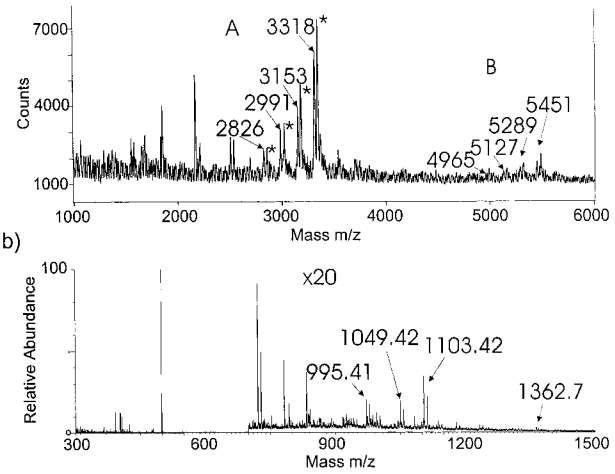


FIGURE 3: (a) MALDI mass spectrum of HPLC fraction 20 of the tryptic digest of rgp120. (b) Nanoelectrospray mass spectrum of HPLC fraction 20 of the tryptic digest of rgp120. Peaks marked with an asterisk (*) are due to 2-fold oxidation of the adjacent peaks

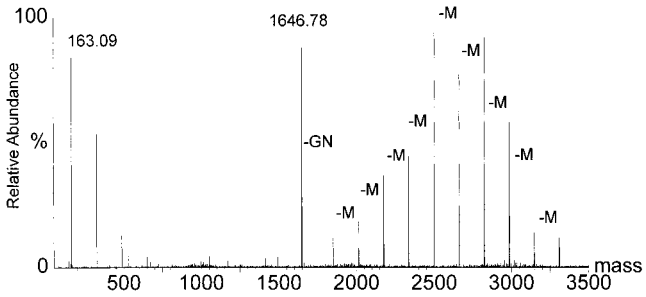


FIGURE 4: MS/MS spectrum of *m*/*z* 1103.42 (+3) observed in the nanoelectrospray mass spectrum of HPLC fraction 20 after transformation into a singly charged spectrum.

1923/1869/1815/1761/1707/1653/ 1599/1545) and the corresponding quadruply charged ions (*m*/*z* 1483/1442.5/1402/ 1361.5/ 1321/1280.7) separated by 162/*z* (*z* = charge of ion) Da are observed (as well as ions due to the corresponding oxidized and sodiated glycopeptides). The quadruply charged ion of *m*/*z* 1482.71 was selected for MS/MS analysis. The corresponding transformed singly charged MS/MS spectrum is shown in Figure 6. The first fragment ion observed corresponds to loss of two mannose units from the calculated molecular ion species, *m*/*z* 5927.8. A further loss of seven mannoses is observed leading to ion *a* at *m*/*z* 4469. Losses of both M and of GN from ion *a* can also be observed. Loss of GN leads to another series of 9 M losses (*b* series) followed by another GN loss, while loss of M leads to a series due to loss of a further four M residues. In total, 18 M and 2 GN residues are lost from the protonated molecule. Thus, the ion of *m*/*z* 2604.94 corresponds to the amino acid sequence plus two GN residues (mass = 406.16), indicating that the mass of the peptide alone is 2198.78. From the table of tryptic peptides this corresponds to T34, residues 352−367, (GEFFYBNTTQLFNNTWR) with a calculated mass of 2198.95 (calculated error 0.008%). This peptide has two glycosylation sites. The relative abundances of the ions in the full-scan mass spectrum indicates that the major glycan at both sites contain 9 M with ions due to successively one less mannose being observed at 83, 68, 41, 36, 24, 17, 9, and 7% relative abundance. The MALDI spectrum of this fraction shows similar ratios for the glycan series (data not shown).

In some specific cases, information can be obtained about the amino acid sequence of a glycosylated peptide by using a high cone voltage in the source to form ions due to the amino acid sequence plus a GN residue. This, typically, did not succeed in the present set of experiments due to the large number of components in any one fraction. In the case of one component in fraction 16, however, conditions were favorable for this approach. Figure 7, panel a shows the full scan nanoelectrospray spectrum of this fraction. The high mass region of the spectrum (> *m*/*z* 950) is dominated by a series of triply charged ions separated by 162 Da/3 (*m*/*z* 54) along with the corresponding sodiated ions (labeled with *). The MS/MS spectrum of the ion of *m*/*z* 1179.6 (Figure 7, panel b) contains a doubly charged ion of *m*/*z* 1017.9 that was assigned as the "peptide + GN" ion. This ion does not correspond to any of the predicted tryptic ions. An MS spectrum (Figure 7, panel c) obtained at an in-source collision (cone) energy of 80 eV contains a relatively abundant doubly charged ion of *m*/*z* 1017.9 identical to that observed in the MS/MS spectrum of *m*/*z* 1179.6. The MS/MS spectrum of the source-formed 1017.9 ion (Figure 8) shows an abundant ion of *m*/*z* 1829.93 arising from loss of GN and a series of y ions interpretable as coming from the sequence NQSSG-GD. On the basis of this sequence and on the mass of the peptide, this peptide can be assigned the sequence TIVFN-QSSGGDPQIEVMH, aa 330−346. Thus, the glycan on N334 is a high mannose type.

Complex Glycans. Although the nano-ESI analysis of the tryptic digest HPLC fractions provided information about the high mannose glycans, ions due to complex glycans were not observed with any certainty. The MALDI spectra of several of the fractions, however, especially fraction 4, suggested the presence of complex glycans as evidenced by ions separated by 291 Da (sialic acid) (Figure 9). These fractions were then subjected to neuraminidase treatment to remove the sialic acids. The fractions were again analyzed by MALDI and nano-ESI, and those fractions that contained new ions not observed prior to neuraminidase treatment were investigated further.

Both the MALDI and ESI spectra of fraction 4 were found to show the greatest difference between sialylated and desialylated samples. The MALDI and nano-ESI spectra of desialylated HPLC fraction 4 are shown in Figure 10, panels a and b, respectively. Several series of ions related by differences of 365/z, attributable to the *N*-acetylglucosamine-galactose (GN-G) sequence found in complex glycans, can easily be seen. For example, the triply charged ions of *m*/*z* 1104.4, 982.8, and 861.1 (ion series D) are separated by 121.67 (365/3) Da and comprise one series of ions due to a complex glycopeptide differing by GN-G, while the quadruply charged ions of *m*/*z* 991.3, 1082.6, and 1173.8 (ion series A) comprise another complex glycopeptide series. Nano-ESI/MS/MS spectra of the major ions from the observed series were obtained.

Attention is drawn to the ESI ion series 805.1, 926.8, 1048.5, and 1170.5 labeled F in Figure 10, panel b, which is a triply charged series of ions separated by 365/*z*. Note that this ion series is not observed above the background level in the MALDI spectrum (Figure 10, panel a). The ratios of these ions is 60:100:22:10. The MS/MS spectrum of *m*/*z* 927 is shown in Figure 11. In contrast to the MS/MS spectra obtained from ions due to high mannose glycans, the MS/

Table 1: Type and Relative Abundance of Glycans Observed at Specific Glycosylation Sites

amino acid	HPLC fraction	glycan type[a]									
N57	4	C	C6[b]−F[c]:20[d]	C5−F:58	C4−F:100	C3−F:73	C2−F:75				
N99	4	C	C4:24	C3:33	C2:100	C1:15					
N99 (minor)		HM	9M[a,d]:	8M:	7M:	6M:					
N110	4	C	C4−F: <10	C3−F:53	C2−F:100						
N124/128	4/16	C	C6:57	C5:100	C4:43						
N154/N160	4/16	C	C9:40	C8:100							
N170	4	C	C5:3	C4:29	C3:40	C2+M:9	C2:100	C2-G:10			
N170	24	HM	9M:90	8M:100	7M:80						
N203	24	HM	9M:100	8M:77	7M:28	6M:41	5M:15	4M:7			
N214	25, 4	HM	9M:75	8M:80	7M:45	6M:100	5M:55	4M:45			
N235	4	C	C5−F:14	C4−F:29	C3−F:40	C2−F:100					
N235	25	HM	9M:77	8M:80	7M:100	6M:55	4M:53				
N249	4	C	C4:6	C3:24	C3−F:4	C2:100	C2−F:45	C2-G:44	C2−F-G:21	C1:29	C1−F:20
N262/268	16	HM	*e*								
N274	4	C	*e*								
N304	24	HM	9M:100	8M:33	7M:10	6M:9	5M:22	4M:9			
N311	20	HM	9M:100	8M:51	7M:28	6M:19					
N328	19	HM	*e*								
N334	16, 17	HM	9M:100	8M:95	7M:95	6M:100	5M:70	4M:100			
N358/N364	25, 26	HM	18M:83	17M:68	16M:41	15M:36	14M:24	13M:17	12M:9	11M:7	
N370	24	HM	*e*								
N378	18/4	HM/C	9M:28	8M:100	7M:68	6M:62	5M:29	4M:8			
N415	24		*e*								
N428/N431	4/15	C/HM/	C7:32	C6:41	C5:64	C4:100					
		(428P)	C4:26	C3:65	C2+M:21	C2:100					
N431	42(AspN)[f]	C	C1:30	C2:30	C3:100						

[a] C = complex; HM = high mannose; P = partial. [b] Number of GlcNAcGal units; C, for complex glycans or number of mannoses, M, for high mannose glycans. [c] F = fucose. [d] Relative abundance; where no numbers are present, the relative abundances could not be determined accurately. [e] Heterogeneity could not be determined accurately. [f] Isolated from an AspN digest.

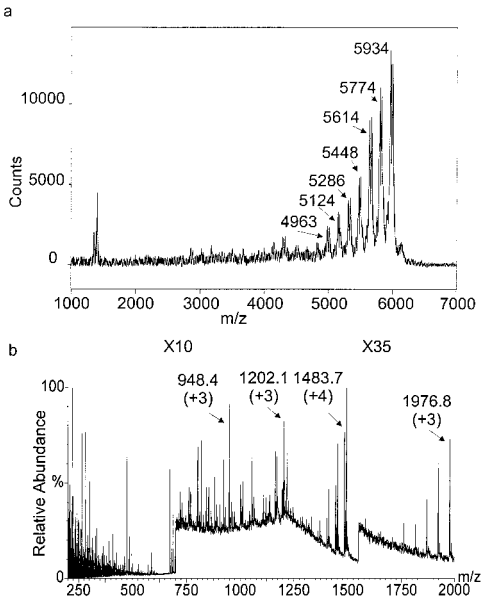


FIGURE 5: (a) MALDI mass spectrum of HPLC fraction 25 of the tryptic digest of rgp120. (b) Nanoelectrospray mass spectrum of HPLC fraction 25 of the tryptic digest of rgp120.

MS spectra of ions due to complex glycans were relatively uninformative about the mass of the peptide chain. Loss of one and two GNs (mass 2203.08 and 406.16, respectively) from the molecular ion are observed, as well as ions due to loss of G (162 Da). Ions due to loss of fucose (146 Da) can also be observed, suggesting that the complex ions are typically fucosylated. No ion stands out as being due to the amino acid sequence plus GN as seen in the high mannose series. The ion of *m*/*z* 1261.6, as well as the ions of lower mass observed, arises from the glycan moiety and is observed

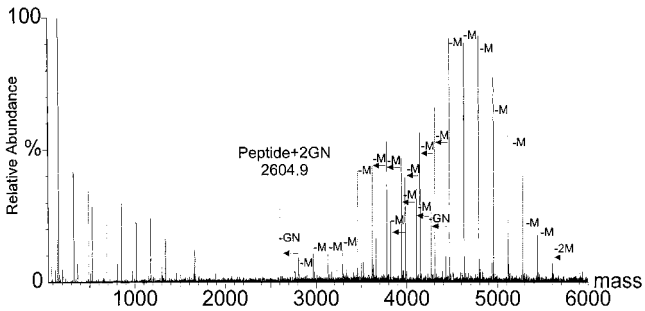


FIGURE 6: MS/MS spectrum of *m*/*z* 1482.71 (+4) ion observed in the nanoelectrospray mass spectrum of HPLC fraction 25 after transformation into a singly charged spectrum.

in most of the MS/MS spectra of complex glycopeptides. The observation of losses of two GN-G residues from the molecular ion does indicate the presence of at least two GN-G sequences. The maximum mass for the peptide inferred based on these data is 1009.4 Da (MW = 926.8 × 3 = 2780.4 − 3H − F − 2(GN-G) − 3M − 2GN).

To provide further information about the structure of the glycopeptide, the fractions (after desialylation) were treated with PNGaseF to remove all the carbohydrates. This process also converts all glycosylated asparagine residues into aspartic acid residues. The MALDI and nano-ESI spectra of desialylated/deglycosylated fraction 4 are shown in Figure 12, panels a and b, respectively. Because no ions clearly assignable to the peptide portion of the glycopeptide were observed in the MS/MS spectra of complex glycopeptides, the calculated masses of possible combinations of peptide ions observed in the spectra of the deglycosylated products plus the masses of various possible complex glycan structures were then matched against the ions observed in the spectra obtained for the desialylated fractions to identify the glyco-

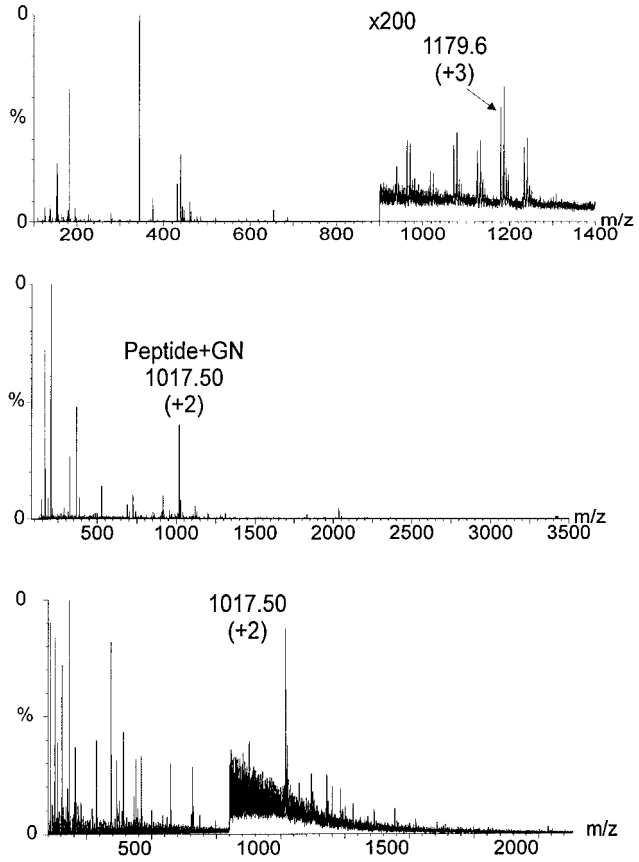


FIGURE 7: (a) Nanoelectrospray mass spectrum of HPLC fraction 16 of the tryptic digest of rpg120. (b) MS/MS spectrum of the *m*/*z* 1179.6 (+3) ion observed in panel a. (c) Nanoelectrospray mass spectrum of HPLC fraction 16 of the tryptic digest of rpg120 obtained using a cone voltage of 80 eV.

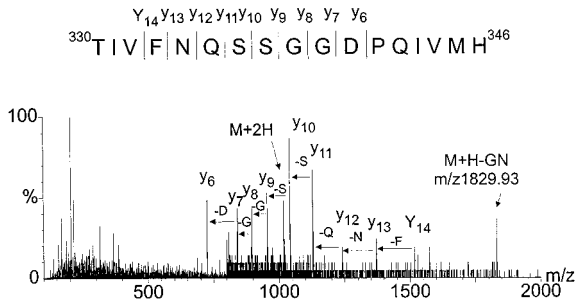


FIGURE 8: MS/MS spectrum of the *m*/*z* 1017.50 ion observed in Figure 7, panel c.

sylated peptide. Thus, for the example given above (*m*/*z* 2778.87, Figure 11), a complex glycan with one F and two GN-G residues fits with the peptide of mass 506.2 observed in the ESI spectrum of deglycosylated fraction 4 (F in Figure 12, panel b). This ion is consistent with the predicted mass of T21, SDNFTNNAK, after deglycosylation. (In the deglycosylation process, N is transformed into D increasing the mass of the peptide by 1 Da). The ESI/MS/MS spectrum of the *m*/*z* 506 doubly charged ion found in the nano-ESI spectrum of desialylated/deglycosylated fraction 4 was clearly consistent with this assignment (data not shown).

Other tryptic peptides identified in desialylated/deglycosylated fraction 4 (and the corresponding ion series in the spectrum of the desialylated fraction) include T6 (residues 95−105; Figure 10, G series), T7 (Figure 10, H series), T9−10 (Figure 10, C series), T13 (no glycosylation site), T14

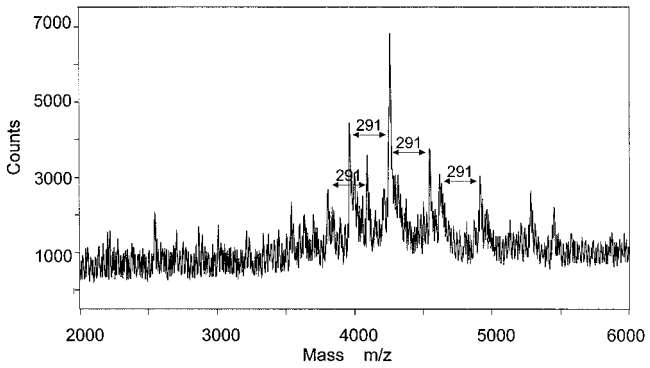


FIGURE 9: MALDI mass spectrum of fraction 4 of the tryptic digest of rgp120.

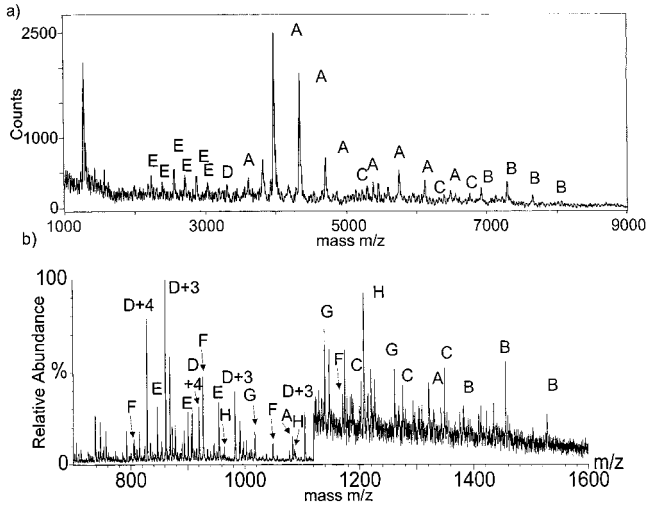


FIGURE 10: (a) MALDI mass spectrum of desialylated fraction 4 of the tryptic digest of rgp120. (b) Nanoelectrospray mass spectrum of desialylated fraction 4 of the tryptic digest of rgp120. A = T40 (residues 424−444); B = T14 (residues 146−165); C = T9−10 (residues 120−134); D = T15 (residues 166−171); E = T36 (residues 377−386); F = T21 (residues 247−255); G = T6 (partial, residues 95−105); H = T7 (residues 106−115).

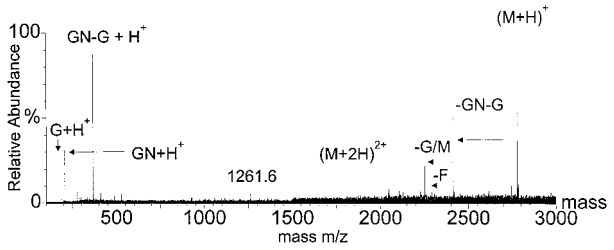


FIGURE 11: MS/MS spectrum of the ion of *m*/*z* 926.8 observed in Figure 10, panel b.

(Figure 10, B series), T15 (Figure 10, D series), T16 (no glycosylation site), T21 (Figure 10, F series), T24 (no glycosylation site), T26 (no glycosylation site), T28, T35, T36 (Figure 10, E series), and T40 (Figure 10, A series). In addition, ions corresponding to glycosylated T25 and T36 were observed in the MALDI spectrum of nondesialylated fraction 4 (Figure 9), while T22 was observed in the MALDI spectrum of deglycosylated fraction 4 (Figure 12, panel a).

Heterogeneity Due to Partial Glycosylation. Glycopeptide heterogeneity can be due to the presence of some sites that are only partially glycosylated as well as due to heterogeneity in the glycan moiety itself. Useful information about this type of heterogeneity can be obtained from the MS/MS

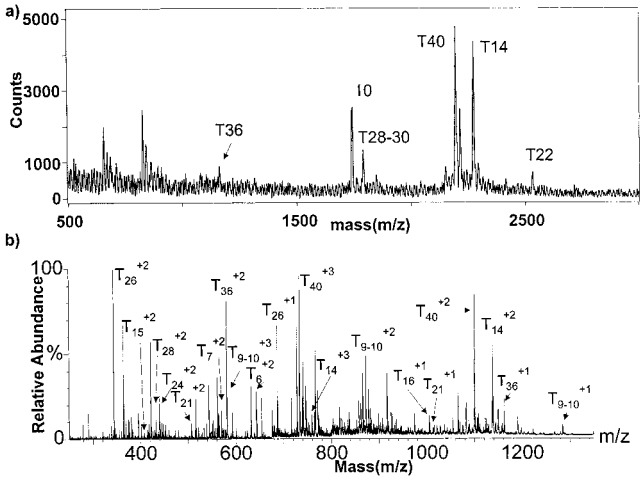


FIGURE 12: (a) MALDI mass spectrum of deglycosylated fraction 4 of the tryptic digest of rgp120. (b) Nanoelectrospray mass spectrum of deglycosylated fraction 4 of the tryptic digest of rgp120.

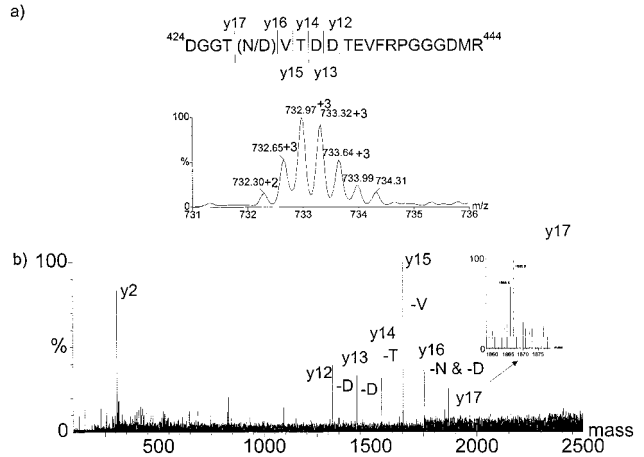


FIGURE 13: (a) Expansion of the spectrum in Figure 10b over the range *m*/*z* 731−736. (b) MS/MS spectrum of the *m*/*z* 732 cluster after transformation into a singly charged spectrum. The inset shows the doublet observed for y_{17}.

spectra of the deglycosylated peptides. As an example, a triply charged ion of mass 732.65−734.31 was observed in the nano-ESI spectrum, which correlates with the predicted mass of T40, residues 424−444. The lowest mass ion of this cluster, however, was 0.37 Da less than expected for the deglycosylated peptide, while the ratios of the ions in the cluster were not consistent with that expected for either one or two N to D transformations (Figure 13, panel a). The MS/MS spectrum of this ion (Figure 13, panel b) was obtained, and the sequence ions observed were consistent with that expected for T40. Close inspection of the fragment ions showed that y_{13}−y_{16} ion series containing residue 431 were single peaks with no indication of the presence of N rather than D. A doublet would be expected to be observed if residue 431 were initially only partially glycosylated, resulting in the presence of N from nonglycosylated N-431 and D from glycosylated N-431. The y_{17} ion, which contained residue 428, however, was observed to be a doublet due to the presence of both N and D in relatively equal amounts separated by one dalton. Because the glycosylation process converts glycosylated N into D, the MS/MS data indicated that N428 is partially glycosylated, while N431 is completely glycosylated. The ion cluster in the full scan mass spectrum is also consistent with that predicted on the basis of two ions being present (Figure 13, panel a).

The data obtained from this study gives us a unique opportunity to develop a model of the fully glycosylated, full sequence protein based on the crystal structure (*23*). This model was based on the gp120 protein structure coordinates for the HIV gp120-CD4-FAB 17B antibody complex. The resulting structure is shown in Figure 14.

DISCUSSION

Using a combination of specific glycosidases, proteolytic digestion, and MALDI and nanoelectrospray MS/MS, we have identified the type of glycans found on 25 of 26 consensus glycosylation sites in HIV-1_{SF2} rgp120 and in most cases the pattern of glycan heterogeneity found at these sites (Table 1). Those sites where the glycan heterogeneity pattern could not be determined were characterized by ions of relatively low abundance and/or by ions that were overlapped by other ions.

Leonard et al. have reported the assignment of glycan type at potential glycosylation sites on recombinant HIV 1_{IIIB} gp120. The SF2 and IIIB strains are similar in that the 24 consensus glycosylation sites in IIIB show a similar distribution to 24 of the 26 sites in the SF2 strain (± 7 residues). Our results for glycan identification for the SF2 strain are consistent with those of Leonard for similarly located glycosylation sites. In addition, the presence of additional lysines and arginines in SF2 in the region encompassing Asn 358/364/370/378 that were not present in the corresponding region of IIIB, Asn 356/362/367/376, permitted assignment of the glycan type for each residue. Biller et al. analyzed three N-linked glycans in HIV-1 BRU containing mutant HIV-1 BH10 *env* sequences by releasing the glycans and then analyzing by normal phase HPLC (*34*). They observed that the glycan on N448 had a high mannose structure, while the glycans on N406 and N463 were heterogeneous. As with the IIIB strain, there is not an exact correspondence between the sequence of the SF2 and BH10 strains. When the sequences are aligned, BH10 N448 corresponds most closely to SF2 N415, which is the one site that we could not determine, while BH10 N406 and N463 correspond to SF2 N378 and N431. We also observed heterogeneity for both of these residues.

From our results, we found that the high mass glycopeptides above ca. *m*/*z* 4000, (e.g., T39 containing N415, which would have a mass of 4446 for a high mannose glycan) or those with multiple glycosylation sites were not readily observed in the complex mixtures analyzed by nanoelectrospray, presumably due to suppression effects. In general, we have observed ions from high mass species to be suppressed in the nanoelectrospray analysis of mixtures relative to their abundance when analyzed individually.

Structural Implications. HIV gp120 plays a key role in the viral infection process through a tertiary complex involving CD4 and a chemokine secondary receptor (*35*−*37*). The V-3 loop and its associated glycans, as well as other regions of the gp120, have been implicated in the binding of gp120 in this complex (*38*, *39*). Recently, a crystal structure of a complex between a mutant gp120, a mutant CD4, and an antigen-binding fragment of an antibody has been reported (*23*). To obtain crystals of the protein for

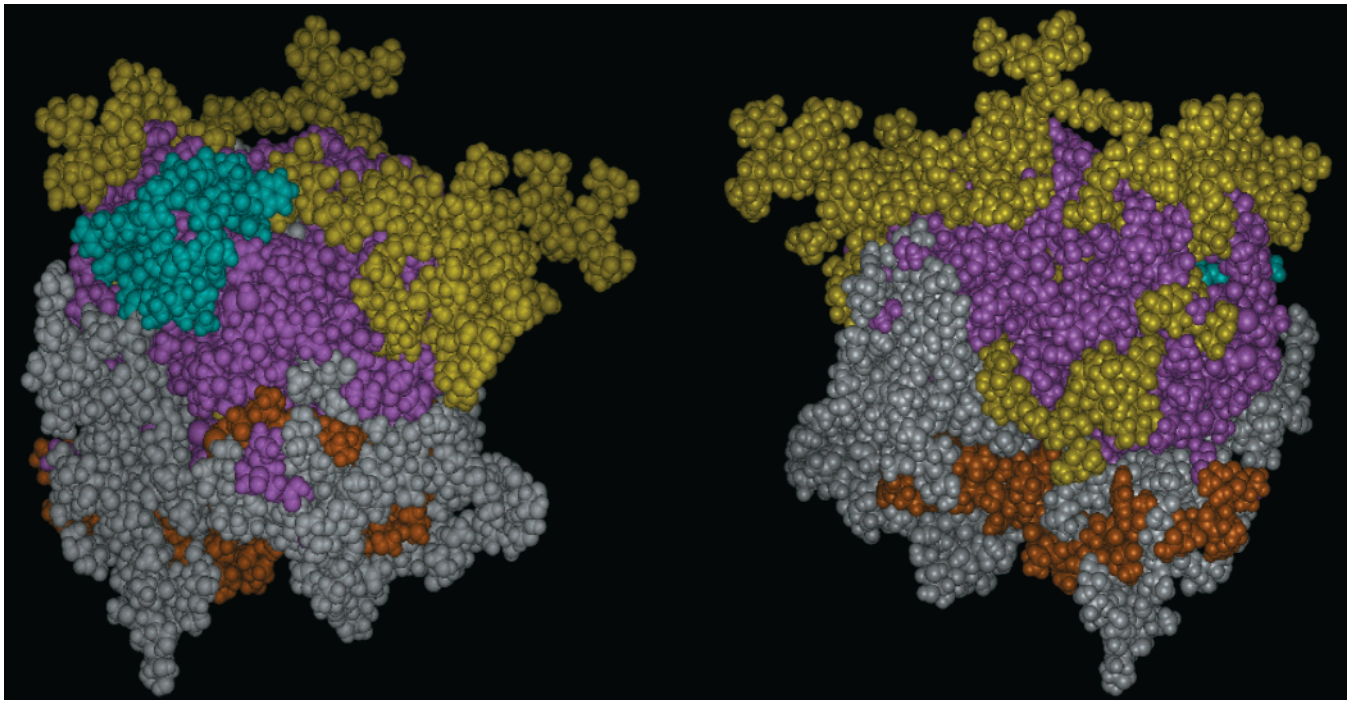


FIGURE 14: Modeled structure of the complete HIV GP120 glycoprotein including the V1/V2 and V3 protein loops and the complete carbohydrate glycosylation on all sites identified experimentally by mass spectrometry. In white are shown the modeled complex carbohydrate glycosylation sites and in yellow the high mannose glycosylation sites. The modeled V1/V2 protein loops are shown in orange, and the modeled V3 loop is shown in blue. The remainder of the protein is shown in magenta. Two views of the protein are shown so that the entire surface of the protein can be seen.

structural studies, it was necessary to extensively deglycosylate the gp120. More than 90% of the carbohydrate structure was removed for crystallization, and all sugar groups beyond the linkages between the core *N*-acetylglucosamine residues were removed. Additionally, the mutant protein used for the crystal structure determination had tripeptides, Gly−Ala−Gly, substituted for the 67 residues of the combined V1 and V2 loops and for the 32 residues of the V3 loop. The HIV gp120 V3 loop is of particular significance, as it plays a major role in eliciting neutralizing anti-HIV antibodies, and has been an important target for AIDS vaccine development (*26*−*30*). The HIV SF2 variant gp120 structure reported here is the first model including the V1, V2, and V3 loops and complete structural characterization for 25 of the protein glycosylation sites. Consistent with available information, the V-3 loop is highly exposed. This model (Figure 14) illustrates the significant role that mass spectrometric techniques combined with molecular modeling can play in facilitating the characterization of glycoprotein structures.

Interestingly, the high mannose glycans appear to be clustered together on the surface, while the complex glycans form another cluster with little structural overlap. This clustering may have a biological significance that had not been recognized because the structural clustering had not previously been suggested. There have been suggestions about potential biological activity of the glycans, e.g., sialylation of the complex glycans has been proposed to reduce the infectivity of human and nonhuman primate lentiviruses (*40*). Several groups have also noted that gp120 interacts with mannose-binding lectin (MBL) in human serum (*41*, *42*). Saiffuddin et al. also showed that high-mannose glycosylation sites of gp120 largely contributed to the MBL-HIV interaction rather than host cell-derived glycoproteins (*43*). The presence of clustered high mannose or complex glycans may also be related to improved interactions between cellular glycans and the viral glycans.

In future developments, we plan to investigate the use of capillary HPLC nanoflow ESI/MS for glycopeptide analysis. The goal will be to obtain profiles of the complex glycans without prior desialylation and of the higher molecular weight glycopeptides.

REFERENCES

1. Huberty, M. C., Vath, J. E., Yu, W., and Martin, S. A. (1993) *Anal. Chem. 65*, 877−884.
2. Sutton, C. W., O'Neill, J. A., and Cottrell, J. S. (1994) *Anal. Biochem. 218*, 34−46.
3. Juhasz, P., and Martin, S. A. (1997) *Int. J. Mass Spectrom. Ion Proc. 169/170*, 217−230.
4. Tsarbopoulos, A., Bahr, U., Pramanik, B. N., and Karas, M. (1997) *Int. J. Mass Spectrom. Ion Proc. 169/170*, 251−261.
5. Rahbek-Nielsen, H., Roepstorff, P., Reischl, H., Wozny, M., Koll, H., and Haselbeck, A. (1997) *J. Mass Spectrom. 32*, 948−958.
6. Carr, S. A., Huddleston, M. J., and Bean, M. F. (1993) *Prot. Sci. 2*, 183−196.
7. Huddleston, M. J., Bean, M. F., and Carr, S. A. (1993) *Anal. Chem. 65*, 877−884.
8. Medzihradszky, K. F., Maltby, D. A., Hall, S. C., Settineri, C. A., and Burlingame, A. L. (1994) *J. Am. Soc. Mass Spectrom. 5*, 350−358.
9. Settineri, C. A., and Burlingame, A. L. (1994) in *Techniques in Protein Chemistry V*. (Crabb, J. W., Ed.) pp 97−104, Academic Press, San Diego, CA.
10. Schindler, P. A., Settineri, C. A., Collet, X., Fielding, C. J., and Burlingame, A. L. (1995) *Prot. Sci. 4*, 791−803.
11. Stimson, E., Hope, J., Chong, A., and Burlingame, A. L. (1999) *Biochem. 38*, 4885−4895.
12. Jeyarajah, S., Parker, C. E., Sumner, M. T., and Tomer, K. B. (1998) *J. Am. Soc. Mass Spectrom. 9*, 157−165.

13. Borchers, C., and Tomer, K. B. (1999) *Biochem. 38*, 11734−11740.
14. Hochleitner, E. O., Gorny, M. K., Zolla-Pazner, S., and Tomer, K. B. (2000) *J. Immunol. 164*, 4156−4161.
15. Luckow, V. A., and Summers, M. D. (1988) *Bio/Technology 6*, 47−55.
16. Reitter, J. N., Means, R. E., and Desrosiers, R. C. (1998) *Nat. Med. 4*, 679−684.
17. Mizuochi, T., Spellman, M. W., Larkin, M., Solomon, J., Basa, L. J., and Feizi, T. (1988) *Biochem. J. 254*, 599−603.
18. Mizuochi, T., Matthews, T. J., Kato, M., Hamako, J., Titani, K., Soloman, J., and Feizi, T. (1990) *J. Biol. Chem. 265*, 8519−8524.
19. Leonard, C. K., Spellman, M. W., Riddle, L., Harris, R. J., Thomas, J. N., and Gregory, T. J. (1990) *J. Biol. Chem. 265*, 10373−10382.
20. Yeh, J.-c., Seals, J. R., Murphy, C. I., van Halbeek, H., and Cummings, R. D. (1993) *Biochem. 32*, 11087−11099.
21. Hunt, L. A. (1982) *J. Virol. Methods 4*, 283−95.
22. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J.. Hendrickson, W. A. (1998) *Nature 393*, 648−659.
23. Blackburn, K., Anderegg, R. (1998) in *Proceedings of the 46th ASMS Conference on Mass Spectrometry and Allied Topics*, p 219, American Society for Mass Spectrometry, Orlando, FL.
24. Vranken, W. F., Budesinsky, M., Fant, F., Boulez, K., and Borremans, F. A. M. (1995) *FEBS Lett. 374*, 117−121.
25. Vu, H. M., de Lorimier, R., Moody, M. A., Haynes, B. F., and Spicer, L. D. (1996) *Biochemistry 35*, 5158−5165.
26. Ghiara, J. B., Ferguson, D. C., Satterthwait, A. C., Dyson, H. J., and Wilson, I. A. (1997) *J. Mol. Biol. 266*, 31−39.
27. Stanfield, R. L., Cabezas, E., Satterthwait, A. C., Stura, E. A., Profy, A.T., and Wilson, I. A. (1999) *Structure 1*, 131−142.
28. Ghiara, J. B., Stura, E. A., Stanfield, R. L., Profy, A. T., Wilson, I. A. (1994) *Science 264*, 82−85.
29. Weliky, D. P., Bennett, A.E., Zvi, A., Anglister, J., Steinback, P. J., and Tycko, R. (1999) *Nat. Struct. Biol. 6*, 141−145.
30. Hagler, A. T., Huler, E., and Lifson, S. (1974) *J. Am. Chem. Soc. 96*, 5319.
31. Homans, S. W. (1990) *Biochemistry 29*, 9110−9118.
32. Ha, S. N., Giammona, A., Field, M., and Brady, J. W. (1988) *Carbohydr. Res. 180*, 207−221.
33. Conboy, J. J., and Henion J. D. (1992) *J. Am. Soc. Mass Spectrom. 3*, 804−814.
34. Biller, M., Bolmstedt, A., Hemming, A., and Olofsson, S. (1998) *J. Virol. Methods 76*, 87−100.
35. Feng, Y., Broder, C. C., Kennedy, P. E., and Berger, E. A. (1996) *Science 272*, 872−877.
36. Trkola, A., Dragic, T., Arthos, J., Binley, J. M., Olson, W. C., Allaway, G. P., Cheng-Mayer, C., Robinson, J., Maddon, P. J., and Moore, J. P. (1996) *Nature 384*, 184−187.
37. Lapham, C. K., Ouyang, J., Chandrasekhar, B., Nguyen, N. Y., Dimitrov, D. S., and Golding, H. (1996) *Science 274*, 602−605.
38. Nakayama, E. E., Shioda, T., Tatsumi, M., Xin, X., Yu, D., Ohgimoto, S., Kato, A., Sakai, Y., Ohnishi, Y., and Nagai, Y. (1998) *FEBS Lett. 426*, 367−372.
39. Valenzuela, A., Blanco, J., Krust, B., Franco, R., and Hovanessian, A. G. (1997) *J. Virol. 71*, 8289−8298.
40. Hu, H., Shioda, T., Moriya, C., Xin, X., Hasan, M. K., Miyake, K., Shimada, T., and Nagai, Y. (1996) *J. Virol. 70*, 7462−7470.
41. Ezkowitz, R. A. B., Kuhlman, M., Groopman, J. E., and Byrn, R. A. (1989) *J. Exp. Med. 169*, 185−196.
42. Larkin, M., Childs, R. A., Matthews, T. J., Thiel, S., Mizouchi, T., Lawson, A. M, Savill, J. S., Haslett, C., Diaz, R., and Feizi, T. (1989) *AIDS 3*, 793−798.
43. Saifuddin, M., Hart, M. L., Gewurz, H., Zhang, Y., and Spear, G. T. (2000) *J. Gen. Virol. 81*, 949−955.

BI000432M