Heterogeneous Glycosylation of Immunoglobulin E Constructs Characterized by Top-Down High-Resolution 2-D Mass Spectrometry†

Einar K. Fridriksson,‡ Andrew Beavil,§ David Holowka,‡ Hannah J. Gould,§ Barbara Baird,*,‡ and Fred W. McLafferty*,‡

Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301, and Randall Institute, Kings College, London, WC2B 5RL, U.K.

Received August 16, 1999; Revised Manuscript Received January 4, 2000

ABSTRACT: Posttranslational glycosylation is critical for biological function of many proteins, but its structural characterization is complicated by natural heterogeneity, multiple glycosylation sites, and different forms. Here, a top-down mass spectrometry (MS) characterization is applied to three constructs of the Fc segment of IgE: Fcϵ(3-4) (52 kDa) and Fcϵ(2-3-4)$_2$ (76 kDa) disulfide-bonded homodimers. Fourier transform MS of a reduced sample of Fcϵ(2-3-4) gave molecular masses of 37 527, 37 689, 37 851, and 38 014 Da, directly characterizing multiple glycoforms (hexose = 162 Da) without chromatographic separation. Limited proteolysis of the nonreduced Fcϵ(2-3-4)$_2$ protein yielded a peptide mixture with molecular weight values that agreed with those expected from the DNA sequence. The single glycosylation site in these constructs was identified, and quantities were determined of five glycoforms that agreed within ±2% of the molecular ion values. The 2-D mass spectrum of two glycosylated peptides showed these to have high-mannose structures, −GlcNAc−(hex)$_n$, demonstrating that Fcϵ(2-3-4) has a single such structure of $n = 5-9$. For a mutated sample of Fcϵ(3-4), in addition to five glycoforms, MS showed a molecular discrepancy that could be assigned with proteolysis and 2-D mass spectra to the oxidation of two methionines and an additional residue difference.

Mass spectrometry (MS) in multiple dimensions (2-D, *n*-D, or MS/MS) (*1*) is now an extremely useful tool for the characterization of proteins (*2*−*8*) and other biomolecules. Although the DNA sequence can predict directly the amino acid sequence of its expressed protein, this contains no information on posttranslational modifications such as glycosylations that are typically heterogeneous. With MS, each species of a mixture can produce discrete molecular ion peaks (*9*−*12*) utilizing either matrix-assisted laser desorption (MALDI) (*13*) or electrospray ionization (*14*, *15*). The most common MS characterization method subjects the protein to extensive proteolysis to produce a complex mixture of small (0.5−2 kDa) peptides; MS of this mixture directly or after partial separation (*16, 17*) yields the masses of these molecular ions, from which 2-D MS can provide sequence information. We designate this the "bottom up" approach (*18*). High-resolution Fourier transform MS (*19*−*21*) that can yield isotopically resolved spectra even for 112 kDa proteins (*22*) makes possible a "top-down approach" (*18*); accurate protein molecular weights and fragment mass values from subsequent gas-phase and proteolytic dissociation provide efficient location of sequence errors and derivatized active sites (*23*−*25*). Here the top-down approach is applied to characterize posttranslational modifications due to glycosylation and oxidation, as well as errors in the DNA-predicted amino acid sequence.

Glycosylation of immunoglobulin G (IgG) is found to be both cell- and species-specific, and the pattern of glycoforms is altered in pregnancy and some diseases, e.g., rheumatoid arthritis, as reviewed in (*26*). Immunoglobulin E (IgE) (190 kDa) differs from IgG (150 kDa) by a longer ϵ heavy chain containing an additional Ig domain (cϵ2, Figure 1) and more extensive glycosylation at seven potential sites (Asn-X-Ser/Thr) (*27*). The function of IgE is to mediate antigen stimulation of mast cells and basophils such as occurs in allergic responses. IgE binds to its high-affinity receptor (FcϵRI) on the cell surface; when the IgE−FcϵRI complex is cross-linked by a multivalent antigen, a cascade of signaling events is initiated (*28*). A detailed knowledge of the interaction between the antibody and the receptor could potentially lead to antagonists that block the interaction and thereby stop the allergic reaction (*29, 30*). This study uses MS to examine the structure of the Fc segment of IgE which binds directly to FcϵRI on the cell surface. The constructs examined retain a single glycosylation site (Figure 1): Fcϵ-(2-3-4)$_2$, Fcϵ(3-4)$_2$, and a mutated form, Fcϵ(3-4)$_2$ R334S, with much lower affinity for FcϵRI (*31*). Our MS approach allowed us to elucidate directly the glycosylation structure of these polypeptides.

Characterization of protein glycosylation is typically a tedious and time-consuming process, requiring multiple steps of enzymatic reactions and chromatography. The most common approach is to release all carbohydrates and analyze the carbohydrate pool, intact and after enzymatic hydrolysis.

† This work was supported, in part, by Grants GM16609 (F.W.M.) and AI18306 (B.B. and D.H.) from the National Institutes of Health (US) and by project grants from the Medical Research Council (U.K.), the National Asthma Campaign (U.K.), and the Wellcome Trust (A.B. and H.J.G.).

‡ Cornell University.

§ Randall Institute, Kings College.

10.1021/bi9919091 CCC: $19.00
Published on Web 02/26/2000

Sensitive HPLC protocols using fluorescence labeling have been described (*32*). Mass spectrometry has been used to analyze released glycans (*33*−*35*), and 2-D MS of carbohydrates that have been methylated can give branching and linkage information (*36, 37*). Immunoglobulins G are the largest biomolecules (150 kDa) extensively characterized by MS and MS/MS. Although only average glycoform molecular masses have been reported (*12, 38*−*40*), the individual molecular ions of heterogeneous glycoforms have been resolved for the reduced heavy chain of IgG (49 kDa) (*9, 10*), and the reduced Fc segment of the IgG heavy chain (26 kDa) (*12*). However, the most detailed MS sequence characterizations of IgGs have used the "bottom-up" approach, analyzing proteolytic fragments by LC/MS/MS peptide mapping (*40*−*42*). The current level of accomplishment is represented by the impressive recent work of Hirayama and co-workers (*41*), who subjected an IgG of known sequence to reduction, S-carboxymethylation, and Lys-C proteolysis; after LC separation, ESI/ion trap mass spectra gave molecular mass values of 55 peptides, 9 of 3−9 kDa (although not isotopically resolved, accuracy ±2 Da). Of these, 53 gave partial MS/MS sequence information that demonstrated complete sequence coverage by the 55 peptides. The reduced protein was also subjected to V8 proteolysis, and LC/MS/MS identified 43 peptides representing 72% coverage. A separate LC separation identified 12 14-mer peptides with different types of O-glycosylation at the same Thr-residue; MS/MS indicated sugar sequences for 10 of these peptides. Three 8 kDa peptides N-glycosylated at the same Asn residue were found, but these were too large for analysis by ion trap MS/MS.

This paper describes the use of the top-down approach, without any exterior (e.g., LC) separation, and with or without S−S bond reduction. High-resolution MS/MS provides higher mass accuracy and confident charge state assignment from the spacing of the resolved isotopic peaks. We could characterize the glycosylation of IgE constructs as large as 76 kDa (Figure 1), including heterogeneity, substitution sites, and type [complex carbohydrate, hybrid carbohydrate, or high mannose (see, e.g., ref *43*)]. In addition to elucidating the glycosylation properties, our analysis revealed unexpected variations in particular amino acids.

EXPERIMENTAL PROCEDURES

IgE−*Fc Constructs.* Fcϵ(2-3-4)$_2$ was prepared as described previously (*44*). The IgE(ND) Fc region cDNA sequence in the vector pJJ71 (*45*) was modified by substituting glutamine for asparagine at the two surface-exposed glycosylation sites unique to IgE (N265Q and N371Q), leaving the conserved glycosylation site at N394 (homologous to N297 in IgG). The N-terminal portion of the cϵ2 domain was reconstructed with synthetic oligonucleotides to include the substitution of Cys-225 by Ala and to allow the incorporation of an *Eco*RV restriction site at the 5′ end of the cDNA. A mouse variable κ-chain leader sequence was ligated to this site. For expression in NS-0 cells, the hIgE−Fc(N265Q,N371Q) construct was subcloned into the *Eco*RI−*Bcl*I-cut pEE6h-CMVgpt expression vector (*46*) containing glutamine synthetase cDNA as a selectable marker (*47*). After transfection of NS-0 cells according to Bebbington et al. (*47*), stable lines expressing this vector, called pRY22, were selected in glutamine-free media and screened for the highest expression of Fcϵ(2-3-4)$_2$. This protein was purified from the supernatants of roller bottle cultures of this cell line by affinity chromatography on anti-IgE−Fc mAb 7.12−Sepharose 4B. The anti-IgE was prepared from the corresponding hybridoma cell line (American Type Culture Collection, Rockville, MD) and linked to the affinity matrix according to the manufacturers' instructions. Secretion of the Fcϵ(2-3-4)$_2$ is accompanied by processing to yield a final product containing Asp-Ile attached to Ala-225 at the N-terminal end of cϵ2 (*44*).

The construct for the Fcϵ(3-4) (*48*) expressing the sequence from Cys-328 (the penultimate amino acid in cϵ2) to Lys-547 (the C-terminus of the secreted form of the ϵ-chain) was derived from pRY22 by PCR, incorporating the *Eco*RV site at the 5′-end for linking to the κ-leader sequence, as described above. This was cloned, expressed, and purified as described above for pRY22, but was further purified by HPLC on a Superdex 75 HR 10/30 column, as previously reported (*48*).

For Fcϵ(3-4)$_2$ R334S, the Arg334Ser mutation was introduced by PCR into human IgE(ND) heavy chain cDNA (*44*), as presented in the vector pSC213 (*49*). A mutagenic primer was synthesized that incorporated the Arg334Ser mutation, the *Eco*RV restriction site, and the intervening sequence. A reverse primer, JH-1, internal to cϵ4 and downstream of the unique *Hpa*I site in cϵ3, was also synthesized. The primers were: ATTGATATCTGTGCAGATTCCAACCCGAGCGGGGTGAGC and CTCCGGCGTCGCAAACGCATA. The PCR product (430 bp) was blunted with T4 DNA polymerase and subcloned into the *Eco*RV site of pBlueScript. Positive colonies were identified by colony PCR, and the integrity of the *Eco*RV and *Hpa*I restriction sites was ascertained by restriction digestion. Clones that mapped correctly were fully sequenced to demonstrate the absence of any mutations. Because the wild-type cϵ(3-4) expression vector has an additional *Hpa*I restriction site outside of the expression cassette, the entire expression cassette was transferred to the cloning vector pK19 (*50*) as an *Eco*RI−*Bam*HI fragment (new vector designated pK19/19). The mutagenic cassette was removed from the pBluScript vector as an *Eco*RV−*Hpa*I fragment, and inserted into pK19/19. The whole expression cassette was then reinserted as an *Eco*RI−*Bam*HI fragment into a similarly digested NS-0 cell expression vector constructed to express wild-type Cϵ(3-4) (*48*).

Mass Spectral Analysis. Samples were dialyzed in 5% acetic acid; reduced samples were first treated with 10 mM dithiothreitol and 2 M urea. Some samples were deglycosylated by incubating with PNGase F (New England Biolabs, VT) at an enzyme/substrate ratio of 1/200 to 1/100 w/w at 37 °C for 48 h. For proteolysis, 10 or 20 μg of protein was incubated with 0.5 μg of Lys-C or 0.2 μg of Asp-N (Sigma, St. Louis, MO) in 50−100 μL of 2 M urea for 2−4 h at 37 °C, in some cases adding DTT to 10 mM for an additional hour. For pepsin proteolysis, the sample was acidified to 3% acetic acid; then pepsin was added at an enzyme/substrate ratio of 1/1 and reacted at 0 °C for 10 min. For partial-reduction experiments, aliquots of Fcϵ(2-3-4)$_2$ were reduced with 10 or 50 mM DTT at room temperature, alkylated with iodoacetamide, and dialyzed before proteolysis in 2 M urea and a second addition of 10 mM DTT. All the resulting product mixtures were desalted by loading on reverse-phase peptide traps (Michrom Bioresources, Auburn, CA), which were washed with 96/2/2 H_2O/MeOH/AcOH and eluted with

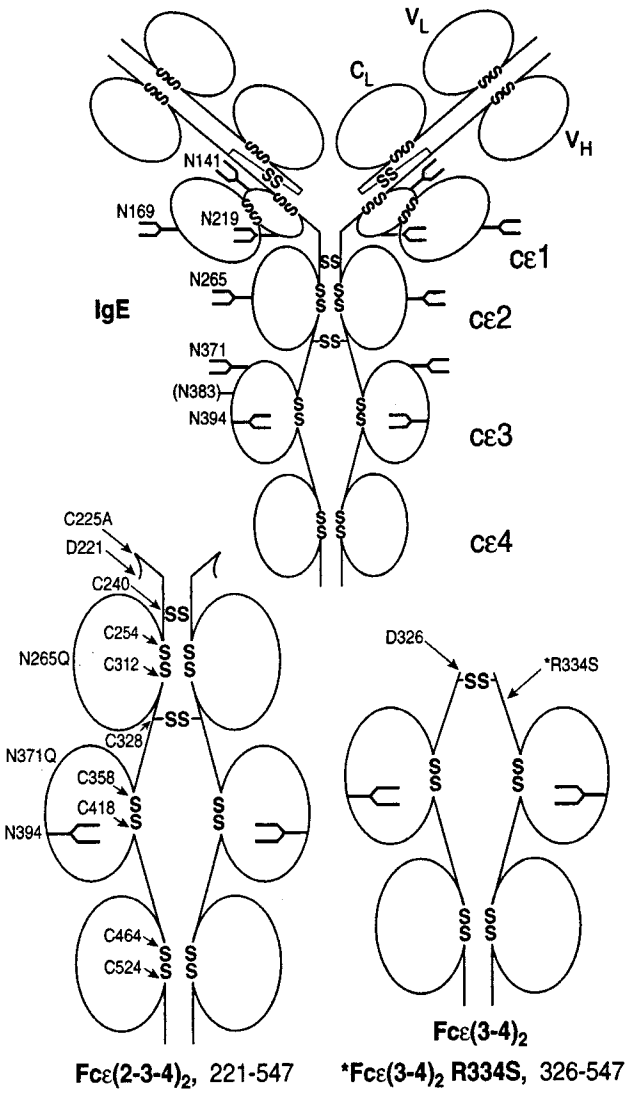


FIGURE 1: Top: Schematic model of IgE with the residue numbering of Dorrington and Bennich (25), and potential Asn (N) glycosylation sites on each ϵ heavy chain. Bottom: IgE constructs studied here; amino-terminal residues, disulfide binding cysteines, and mutated amino acids are indicated.

18/80/2 H_2O/MeOH/AcOH, typically in a volume of 80−120 μL, giving a nominal peptide concentration of about 1.5−5 μM.

Mass spectra were acquired on a modified Finnigan 6 T FTMS described previously (*51*), using nano-ESI for sample introduction (*52*) which requires only a few microliters of the sample loaded for a series of MS experiments. Ion dissociation for 2-D mass spectra was achieved with SWIFT [stored waveform inverse Fourier transform (*53*)] ion isolation and collisionally activated dissociation (CAD) by sustained off-resonance resonance excitation (*20*, *21*) or electron capture dissociation (ECD) (*54*, *55*). Spectral data were reduced with automated data analysis software (*56*). This was also used for quantitative comparisons; briefly, for a given isotopic distribution, all isotopic peak intensities in a fitted distribution are added and divided by the charge state.

RESULTS AND DISCUSSION

The protein constructs of IgE analyzed here were originally engineered to study interactions with the high-affinity receptor, FcϵRI, without added complications from the Fab segments. The constructs contain the Fc segment of the IgE (bottom, Figure 1), either the whole Fc with three Ig domains per chain [Fcϵ(2-3-4)$_2$; 76 kDa] or the truncated Fcϵ(3-4)$_2$ with only two Ig domains per chain (52 kDa). Both of these (but not the single Fc domain cϵ2) bind to the IgE receptor (*44*, *57*). Of the seven potential N-glycosylation sites in the ϵ heavy chain (Asn-X-Ser/Thr, Figure 1), two glycosylated sites, Asn-371 and Asn-394, have been identified in the cϵ3 domain; binding activity is lost if Asn-394 but not if Asn-371 is replaced (*58*). Fc constructs still bind after complete enzymatic deglycosylation (*59*), indicating that glycosylation at Asn-394 is important for proper folding of the protein but not for direct binding to the receptor. [Asn-394 is homologous to the only glycosylation site found in IgG at Asn-297 (*26*).] For all of the constructs examined, Asn-265 in cϵ2 and Asn-371 in cϵ3 are mutated to Gln. One construct, Fcϵ(3-4)$_2$ R334S, has Arg-334 additionally mutated to Ser. Thus, if Asn-383 in cϵ3 is not glycosylated in IgE, as reported previously (*27*, *44*, *60*), only Asn-394 should be available for glycosylation. Both complex and high mannose type glycosylations have been reported for both IgE and IgG (*26*, *60*). Our FTMS analysis allowed direct examination of the expected glycosylation sites of the Fc constructs as well as the unknown degree of glycosylation heterogeneity.

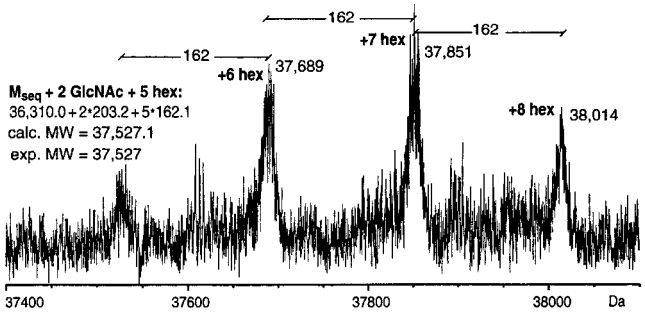


FIGURE 2: ESI/FT mass spectrum of Fcϵ(2-3-4) (reduced), after deconvolution of charges 39+ to 43+.

Table 1: Glycoform Compositions of IgE Constructs[a]

	Fcϵ(2-3-4)$_2$		Fcϵ(3-4)$_2$	Fcϵ(3-4)$_2$ R334S	
	$(M+nH)^{n+}$ red.	peptide nonred.	peptide red.	peptide nonred.	red.
2 hexNAc + 5 hex	13	14	6	13	16
2 hexNAc + 6 hex	29	27	13	31	28
2 hexNAc + 7 hex	36	36	56	27	36
2 hexNAc + 8 hex	23	21	25	16	18
2 hexNAc + 9 hex	−	2	−	12	1

[a] Percentage of total for each sample, based on normalized peak intensities. No peaks with less than 5 hexoses were detected.

Characterization of Heterogeneous Glycosylations. The deconvoluted (*61*) spectrum of DTT-reduced Fcϵ(2-3-4) (Figure 2) shows a heterogeneous series of molecular ions separated by the mass of a hexose unit, $m = 162$. The high mass accuracy ($\pm$1 Da) of FTMS is clearly advantageous in the direct determination of glycosylation heterogeneity. The relative peak heights predict the abundances shown (Table 1) for the four species. However, as considered further below, partial hexose loss during ESI could shift these component proportions. The DNA-predicted sequence for the construct of the whole Fcϵ(2-3-4)$_2$ before any posttranslational modification corresponds to $M_r = 72\,604$. Thus, reduction of the

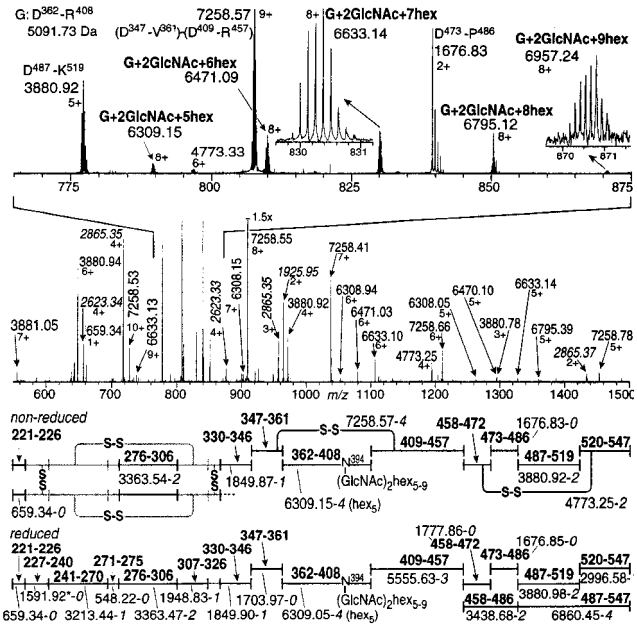


FIGURE 3: ESI mass spectrum of an unseparated Asp-N digest of nonreduced Fcϵ(2-3-4)$_2$. Mass labels, except those in italics, are assigned (within 20 ppm) to masses predicted from the DNA sequence, as shown (bottom) in the peptide map; assigned peptides in boldface type. The italic numeral -*n* following the mass value indicates that it refers to the peak Δn Da above the monoisotopic peak.

8 S−S bonds of this symmetrical dimer should yield Fcϵ-(2-3-4) monomer with M_r = 36 310, but this mass was not detected. The mass 37 527 isotopic cluster of Figure 2 could correspond to 36 310 + (2 × 203) + (5 × 162): a high-mannose modification at a single Asn site would have two acetylglucosamine (GlcNAc) units, each 203 Da, followed by multiple mannose units. Thus, the spectrum is consistent with monomeric Fcϵ(2-3-4) as a high-mannose glycoconjugate with heterogeneity of at least 5, 6, 7, and 8 mannoses. If the sequence was not known, Figure 2 data would also be consistent with glycosylation at multiple protein sites with an unknown total number of hexoses.

The spectrum of a proteolytic Asp-N digest of Fcϵ(2-3-4)$_2$ (nonreduced sample, no chromatographic product separation) gives the spectrum in Figure 3. M_r values could be assigned to give a partial coverage (all but 227−275 and 307−329 out of 221−547) peptide map of the 37 kDa monomer (Figure 3, bottom), including peptides containing the intrachain disulfide bonds Cys-358 to Cys-418 and Cys-464 to Cys-524. Only one glycosylated peptide (five mass values of Δm = 162 Da) is observed, and this corresponds to the predicted D^{362}−R^{408} containing Asn-394; a nonglycosylated form of this peptide was not observed. This analysis also confirms the N371Q mutation, but the peptide containing the N265Q mutation was not observed. In a DTT-reduced digest, however, the peptide D^{241}−E^{270} was observed, confirming also the latter mutation (data not shown). The peptide data clearly show a low-abundance (2%) glycoform with nine hexoses, and the relative amounts of the different glycoforms (Table 1) agree well for Fcϵ(2-3-4) in Figures 2 [reduced molecular ion, $(M + nH)^{n+}$] and 3 (nonreduced proteolysis peptides). This indicates that ESI did not induce appreciable hexose loss; it is unlikely that the ESI losses would be the same for the protein and peptide.

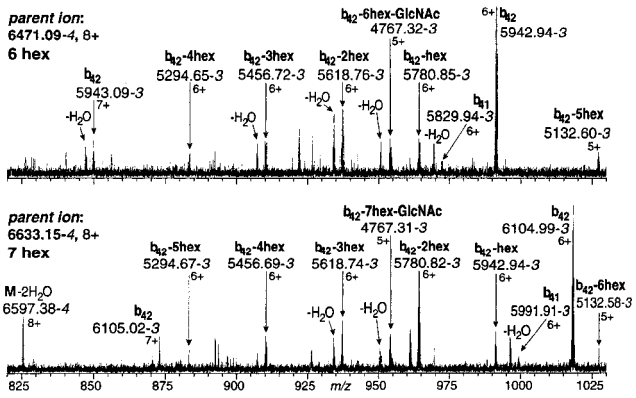


FIGURE 4: 2-D CAD mass spectra of the ions represented by the isotopic clusters centered at 6471-*4* and 6633-*4* Da ions in the Figure 3 spectrum.

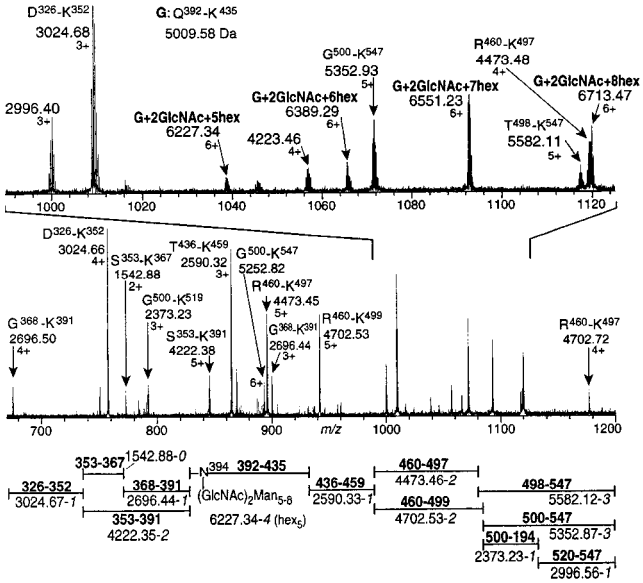
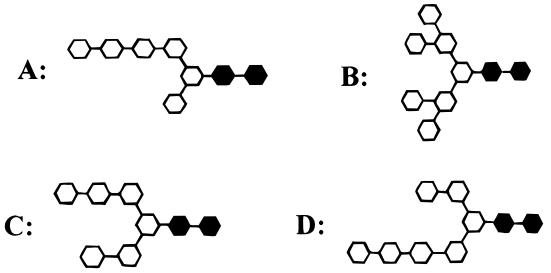


FIGURE 5: ESI/FT mass spectrum of a Lys-C digest of a Fcϵ(3-4)$_2$ (without R334S mutation) construct, reduced with DTT after proteolysis and before desalting/acidification; labels and peptide map as in Figure 3.

Such analysis also assumes equal ionization efficiencies for the glycoforms; this and possible ESI losses should be checked with standards if rigorous quantitation is desired.

This structural characterization of the glycoforms was based on the DNA-predicted sequence. To obtain independent information, two of the Δm = 162 Da peaks of the Figure 3 spectrum were SWIFT-isolated and subjected to CAD to yield the 2-D spectra in Figure 4. Fragmentation is only seen at two amide bonds in the peptide backbone, yielding b_{41} and, mainly, b_{42} species from cleavage between Thr−Leu and Leu−Pro, respectively; CAD cleavage of peptide bonds on the N-terminal side of Pro is frequently observed (*21*). The more abundant b_{42} ions show sequential losses of six or seven hexose units (mass 162 Da), respectively, followed by one mass 203 Da unit, thereby positively verifying the high-mannose assignment with mass 162 Da as mannose and mass 203 Da as GlcNAc. The other basic carbohydrate structures, complex and hybrid, do not show such sequential $C_6H_{10}O_5$ losses (*36*). These b_{41}, b_{42} ions also show that the glycosylation is within amino acids 362−403 of the protein. This region contains both the Asn-383 and Asn-394 sites, so an unambiguous assignment of glycosy-

lation to Asn-394 cannot be made from these spectra. A Lys-C digest of the sample was also analyzed and yielded the nonglycosylated peptides $G^{368}-K^{380}$ and $G^{368}-K^{388}$ (spectrum not shown, but see Figure 3, bottom), confirming that Asn-383 is not modified and that Asn-371 has been mutated to Q. In the 2-D spectra for both hex-6 and hex-7 (Figure 4), the most abundant hexose loss peaks correspond to the loss of 2 units; the relative losses are 0.48, 1.0, 0.45, and 0.26 for the loss of 1−4 units from hex-6 and 0.35, 1.0, 0.47, and 0.31 for the corresponding loss from hex-7. This is evidence against high-mannose structures such as A for hex-6 and B for hex-7; structures such as C and D are more likely, assuming that branching favors cleavage (*34*−*37*). More exact structure assignments require spectral correlation with known analogous structures, permethylation (*36*, *37*), or specific deglycosylation enzymes (*56*, *63*).



Another construct, that of Fcϵ(3-4)$_2$, also with the N371D mutation (Figure 1), allowed further restriction of the glycosylation site with the "top down" approach. After DTT reduction, an ESI spectrum (not shown) of poor signal/noise indicated molecular weights of 26 197, 26 358, and 26 520. After digesting with Lys-C followed by DTT reduction, the unseparated product mixture gave the spectrum in Figure 5; only one peptide has mass values (6227, 6389, 6551, and 6713 Da) consistent with heterogeneous glycosylation. In this case, fragment mass values are found whose respective sums (26 033, 26 195, 26 357, and 26 519 Da) are those of the molecular ions, so that peptides could be assigned to cover all of the protein sequence (Figure 5, bottom). The predicted sequence of the glycosylated peptide $Q^{392}-K^{435}$ has only one asparagine residue, Asn-394, so that it is positively identified as the site of glycosylation. Subtracting masses corresponding to the indicated $(GlcNAc)_2Man_n$ results in a deglycosylated mass sum of 24 813.6 Da versus a DNA-predicted value of 24 812.7 Da. The masses of the glycosylated species match the calculated M_r values for high-mannose structures with 5−8 mannoses, although their relative abundances differ significantly from the values determined for Fcϵ(2-3-4) (Table 1).

Characterization of Multiple Amino Acid Variants. For the Fcϵ(3-4) $_2$ R334S construct (Figure 1), a good spectrum for the intact or reduced glycoprotein was not obtained, but after deglycosylation with PNGase, high-resolution spectra were obtained both before and after reduction (Figure 6). However, for neither sample does the molecular mass match the mass calculated from the DNA-predicted sequence; the discrepancy is +40 Da for the intact Fcϵ(3-4)$_2$ and +18.3 Da for the reduced single chain, both consistent with a 19 ± 1 Da error in the latter. Note that the latter spectrum (Figure 6, lower inset) has minor amounts of lower mass isotopic peaks near the calculated M_r value (loss of H_2O could contribute

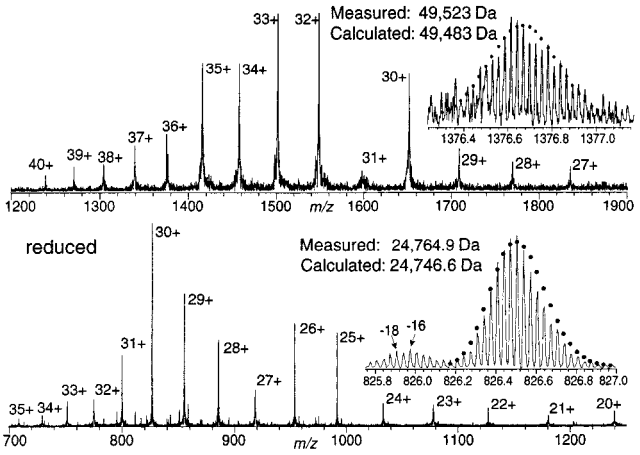


FIGURE 6: ESI/FT mass spectra of the Fcϵ(3-4)$_2$ R334S construct, after deglycosylation with PNGase, before (top) and after (bottom) DTT reduction.

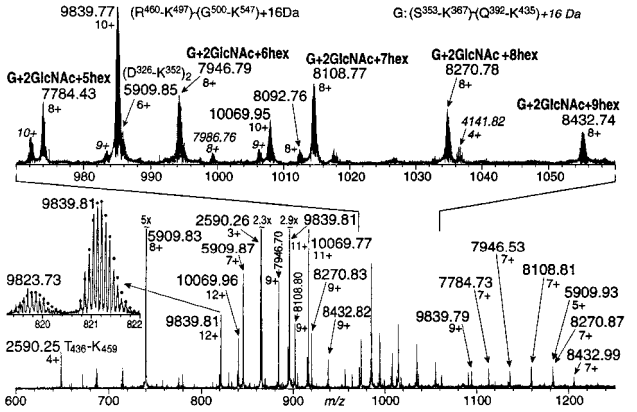


FIGURE 7: ESI/FT mass spectrum of a Lys-C digest of the Fcϵ(3-4)$_2$ R334S construct, without deglycosylation or DTT reduction of the peptides. The inset in the lower spectrum shows an expansion of the 9839.8 Da 12+ distribution and a lower intensity peak at 9823.7 Da, assigned as the (460−497)−(500−547) peptide, with Met-470 oxidized and nonoxidized, respectively.

to these peaks). The protein was also digested with Lys-C without DTT reduction. In the spectrum of the resulting peptide mixture (Figure 7; Figure 8, top), the glycans observed are similar (with the number of hexose units ranging from 5 to 9) but relative abundances are substantially different from those of the Fcϵ(3-4) $_2$ construct without the R334S mutation (Table 1). The major glycopeptide peaks appear 16 Da (not 19 Da) higher than the masses calculated for the DNA sequence of $(S^{353}-K^{367})-(Q^{392}-K^{435})$, with mass additions corresponding to $GlcNAc_2(hex)_n$ as in the other Fc constructs (other common glycan types would not account for the extra +16 Da). However, for some of these, lower abundance peaks of the predicted 16 Da lower mass appear. The mass of the glycosylated peptide from a reduced sample (Figure 8, middle and spectrum not shown) again is 16 Da higher than that predicted for $Q^{392}-K^{435}$, and it is consistent with an Asn-394 glycosylation site (although the indicated abundance of hex-9 is inexplicably smaller than in the spectrum of nonreduced sample, Table 1). A 16 Da mass shift is also present (inset, Figure 7) for the disulfide-bonded species $(R^{460}-K^{497})$–$(G^{500}-K^{547})$ (9839 Da), with a lower intensity peak at the correct mass also apparent for some charge states. CAD 2-D MS of the 9839 Da species restricted the location of the 16 Da addition to the disulfide-

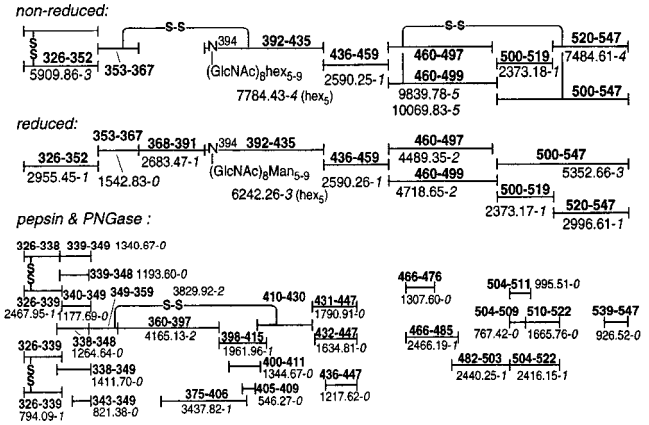


FIGURE 8: Peptide maps of Fcϵ(3-4)$_2$ R334S. Top: Assigned peptides from the nonreduced Lys-C digest of the Figure 7 spectrum. The possible peptides (460−499)−(500−547) and (460−497)−(498−547) are isomeric, so both correspond in mass to the observed 10069.83-*5* Da value. Middle: Peptides from a Lys-C digest that was reduced with DDT after proteolysis. Bottom: Peptides from pepsin proteolysis (nonreduced) of a deglycosylated portion of the sample.

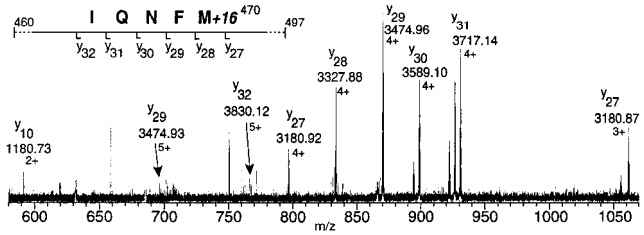


FIGURE 9: 2-D mass spectrum of the 4489 Da peak in the reduced Lys-C digest of Fcϵ(3-4)$_2$ R334S corresponding to R^{460}−K^{497} (see middle peptide map, Figure 8), DNA-predicted mass 4473 Da. The inset shows the sequence information provided by the sequential y ion fragments. Small peaks just below prominent peaks represent losses of (in parentheses): y_{27}^{3+} (18.00 Da), y_{30}^{4+} (17.09 Da), and y_{31}^{4+} (17.00 Da, 34.05 Da), with H_2O = 18.01 Da and NH_3 = 17.03 Da. Note that the N-terminal residues for y_{31} are Gln and Asn.

bonded species (R^{460}−M^{470})–(Q^{518}−S^{532}), and the 2-D mass spectrum from electron capture dissociation (ECD) (*54*) that preferentially cleaves disulfide bonds restricted this further to the R^{460}−M^{470} segment (*55*). After reducing the Lys-C peptide mixture, 2-D MS of the resulting R^{460}−K^{497} peptide gave fragment ions corresponding to F^{469}−K^{497} and M^{470}−K^{497} to locate the extra 16 Da at Met-470 (Figure 9).

The amino acid substitution Met → Phe (131 → 147 Da) could cause a +16 Da error, but incomplete substitution would also have to be postulated to account for the observed minor peak at the correct mass, and this is unlikely. The masses of glycosylated Q^{392}−K^{435} and of (S^{353}−K^{367})−(Q^{392}−K^{435}) in the nonreduced digest are also 16 Da high (Figure 7), with a minor peak at the predicted mass value for some glycan species (e.g., 8092.76 vs 8108.77 Da). The 2-D MS spectrum of its mass 7946.70-*4* (9+) Da peak (as well as showing losses of 1, 2, and 3 hexose units) has a small 7+ peak whose mass of 6769.62-*4* Da corresponds to that predicted (without a +16 Da error) for (S^{353}−K^{367})−(Q^{392}−X^{425}), restricting the 16 Da addition to the C-terminal P^{426}RALMRSTTK^{435}. The most reasonable explanation is that Met-430 and Met-470 are largely oxidized. Although under certain conditions extensive methionine oxidation can occur in the electrospray process (*64*), that is unlikely here,

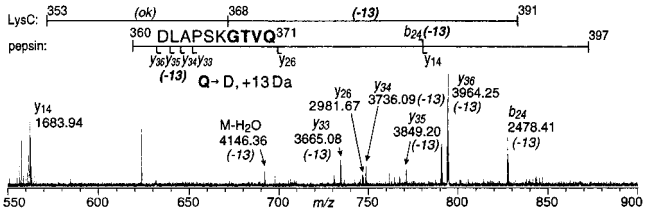


FIGURE 10: 2-D mass spectrum of a 4165 Da peptide from a pepsin digest of Fcϵ(3-4)$_2$ R334S corresponding to V^{360}−L^{397} (see bottom peptide map, Figure 8), DNA-predicted mass 4178 Da. Top: Peptide map of the S^{353}−L^{397} region; the complementary digests and the 2-D MS fragments locate the −13 Da discrepancy to the residues G^{367}TVQ^{371}.

as significant oxidation is only seen for one out of three constructs, in multiple runs each. For the same reasons, it is unlikely that the oxidation would occur during proteolysis, as the molecular ion spectra demonstrate similar oxidation. Oxidation during preparation of the IgE construct can be checked quickly with molecular ion spectra (Figure 6).

The two methionine oxidations add up to +32 Da, not the +19 Da discrepancy seen in the Figure 6 spectra of the deglycosylated protein. This additional −13 Da difference must come from region 368−391 (calculated mass 2696.44-*1* Da), as all other regions are covered by the peptide mapping; in fact, a small peak of mass 2683.47-*1* Da in the Lys-C reduced spectrum (Figure 8, middle) could be a −13 Da error for that predicted peptide, but its intensity was too low for 2-D MS. In a different experiment using pepsin protease (as part of the conformational studies to be reported later), a peptide identified as V^{360}−L^{397} (Figure 8, bottom) showed a −13 Da low mass, and by 2-D MS that error is confined to 365−371 (Figure 10). This and the 368−391 restriction described above pinpoint the error to the four residues G^{367}-TVQ^{371}. The only single amino acid substitution to give a −13 Da shift is Gln → Asp. A subsequent thorough check on the original Fcϵ(3-4)$_2$ R334S preparation supported this assignment, with the error occurring in a data transmission step.

Parallel or Crossed Interchain Disulfide Bonds? A high-resolution structural model of IgE, based on the crystal structure of IgG, included crossed interchain bonds between Cys-240 and Cys-328 in the cϵ2 domain (*65*). Subsequently, the model was refined to be consistent with mutation experiments which indicated that these bonds are parallel (*66*), as depicted in Figure 1. Variations in the S−S bond connectivities would not change the measured M_r values for the molecular ions, but this can be determined from the peptide masses from the nonreduced protein. No peptides containing these disulfide bonds were observed in the nonreduced digest of Fcϵ(2-3-4)$_2$ (Figure 3). In an attempt to test the parallel connectivity, aliquots were partially reduced with 10 or 50 mM DTT, subsequently alkylated, subjected to Asp-N proteolysis, and finally completely reduced under denaturing conditions. No peptide containing Cys-328 was isolated, and D^{227}−C^{240} that contains Cys-240 was found only as the alkylated species. At the higher concentration of DTT, the intensity of peptide D^{227}−C^{240} increases. All peptides resulting from reduction of intrachain disulfide bonds are nonalkylated, confirming that intrachain S−S bonds were little affected by the first reduction step. Thus, our results are consistent with preferential reduction of a terminal C^{240}−C^{240} S−S bond in Fcϵ(2-3-4)$_2$.

CONCLUSIONS

As illustrated here with IgE, top-down MS is a powerful method for characterization of posttranslationally modified proteins and verification of their assigned sequences. The first step, measuring an accurate molecular weight for the intact protein, gives an instant check on both the cDNA-derived amino acid sequence and the extent of modifications. Second, the location of indicated modifications can be specified more closely with limited proteolysis producing fairly large peptides (>20 kDa is still easily resolved isotopically by FTMS) that represent all, or nearly all, the protein sequence. Digests of nonreduced proteins provide data on the connectivity of disulfides. Molecular masses differing by hexose (162 Da) units for both the protein and its peptides specify glycosylation heterogeneity. Amino acid modifications can be located more exactly by 2-D MS spectra, which can also give definitive information on the type, as well as extent, of glycosylation. Our MS analysis of IgE constructs confirmed the glycosylation site and quantitatively characterized the glycoform heterogeneity. For the Fcϵ(3-4) R334S mutant, these detailed analyses also revealed the site of an amino acid misassignment as well as unexpected oxidation of two specific methionines.

ACKNOWLEDGMENT

We are grateful to D. M. Horn for assisting with the FTMS analysis and to R. Calvert for retrieving from previous laboratory notebooks detailed information about preparations of IgE−Fc constructs.

REFERENCES

1. McLafferty, F. W., Kelleher, N. L., Begley, T. P., Fridriksson, E. K., Zubarev, R. A., and Horn, D. M. (1998) *Curr. Opin. Chem. Biol. 2*, 571−578.
2. Biemann, K., and Papayannopoulos, I. A. (1994) *Acc. Chem. Res. 27*, 370−378.
3. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) *Nature 379*, 466−469.
4. Hunt, D. F., Henderson, R. A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A. L., Appella, E., and Engelhard, V. H. (1992) *Science 255*, 1261−1263.
5. Qin, J., and Chait, B. T. (1997) *Anal. Chem. 69*, 4002−4009.
6. Andersen, J. S., Svensson, B., and Roepstorff, P. (1996) *Nat. Biotechnol. 14*, 449−457.
7. Annan, R. S., and Carr, S. A. (1997) *J. Protein Chem. 16*, 391−402.
8. McLafferty, F. W., Fridriksson, E. K., Horn, D. M., Lewis, M. A., and Zubarev, R. A. (1999) *Science 284*, 1289−1290.
9. Lewis, D. A., Guzzetta, A. W., Hancock, W. S., and Costello, M. (1994) *Anal. Chem. 66*, 585−595.
10. Ashton, D. S., Beddell, C. R., Cooper, D. J., Craig, S. J., Lines, A. C., Oliver, R. W. A., and Smith, M. A. (1995) *Anal. Chem. 67*, 835−842.
11. Duffin, K. L., Welply, J. K., Huang, E., and Henion, J. D. (1992) *Anal. Chem. 64*, 1440−1448.
12. Bennett, K. L., Smith, S. V., Lambrecht, R. M., Truscott, R. J. W., and Sheil, M. M. (1996) *Bioconj. Chem. 7*, 16−22.
13. Beavis, R. C., and Chait, B. T. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6873−6877.
14. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) *Science 246*, 64−71.
15. Chowdhury, S. K., Katta, V., and Chait, B. T. (1990) *Biochem. Biophys. Res. Commun. 167*, 686−692.
16. Kassel, D. B., Luther, M. A., Willard, D. H., Fulton, S. P., and Salzmann, J. P. (1993) in *Techniques in Protein Chemistry* (Angeletti, R., Ed.) pp 55−64, Academic Press, San Diego.
17. Tomer, K. B., Moseley, M. A., Deterding, L. J., and Parker, C. E. (1994) *Mass Spectrom. Rev. 13*, 431−457.
18. Kelleher, N. L., Lin, H. Y., Valaskovic, G. A., Aaserud, D. J., Fridriksson, E. K., and McLafferty, F. W. (1999) *J. Am. Chem. Soc. 121*, 806−812.
19. Comisarow, M. B., and Marshall, A. G. (1974) *Chem. Phys. Lett. 25*, 282−283.
20. Williams, E. R. (1998) *Anal. Chem. 70*, 179A−185A.
21. Loo, J. A., Edmonds, C. G., and Smith, R. D. (1993) *Anal. Chem. 65*, 425−438.
22. Kelleher, N. L., Senko, M. W., Siegel, M. M., and McLafferty, F. W. (1997) *J. Am. Soc. Mass Spectrom. 8*, 380−383.
23. Kelleher, N. L., Costello, C. A., Begley, T. P., and McLafferty, F. W. (1995) *J. Am. Soc. Mass Spectrom. 6*, 981−984.
24. Wood, T. D., Guan, Z. Q., Borders, C. L., Chen, L. H., Kenyon, G. L., and McLafferty, F. W. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 3362−3365.
25. Taylor, S. V., Kelleher, N. L., Kinsland, C., Chiu, H. J., Costello, C. K., Backstrom, A. D., McLafferty, F. W., and Begley, T. P. (1998) *J. Biol. Chem. 273*, 16555−16560.
26. Rudd, P. M., and Dwek, R. A. (1997) *Crit. Rev. Biochem. Mol. Biol. 32*, 1−100.
27. Dorrington, K. J., and Bennich, H. H. (1978) *Immunol. Rev. 41*, 3−25.
28. Kinet, J.-P. (1999) *Annu. Rev. Immunol. 17*, 931−972.
29. Sutton, B. J., Beavil, R. L., Beavil, A. J., Young, R. J., and Gould, H. J. (1997) in *Allergic Mechanisms and Immunotherapeutic Strategies* (Roberts, A. M., and Walker, M. R., Eds.) pp 17−31, John Wiley & Sons, New York.
30. Baird, B., Zheng, Y., and Holowka, D. (1993) *Acc. Chem. Res. 26*, 428−434.
31. Henry, A. J., Cook, J. P. D., McDonnell, J. M., Mackay, G. A., Shi, J. G., Sutton, B. J., and Gould, H. J. (1997) *Biochemistry 36*, 15568−15578.
32. Guile, R. G., Rudd, P. M., Wing, D. R., and Dwek, R. A. (1997) in *A laboratory Guide to Glycoconjugate Analysis* (Jackson, P., and Gallagher, J. T., Eds.) pp 199−234, Birkhauser Verlag, Basel.
33. Sutton, C. W., Oneill, J. A., and Cottrell, J. S. (1994) *Anal. Biochem. 218*, 34−46.
34. Yang, Y., and Orlando, R. (1996) *Anal. Chem. 68*, 570−572.
35. Olsthoorn, M. M. A., Haverkamp, J., and Thomas-Oates, J. E. (1999) *J. Mass Spectrom. 34*, 622−636.
36. Reinhold, V. N., Reinhold, B. B., and Costello, C. E. (1995) *Anal. Chem. 67*, 1772−1784.
37. Solouki, T., Reinhold, B. B., Costello, C. E., O'Malley, M., Guan, S., and Marshall, A. G. (1998) *Anal. Chem. 70*, 857−864.
38. Feng, R., and Konishi, Y. (1992) *Anal. Chem. 64*, 2090−2095.
39. Verentchikov, A. N., Ens, W., and Standing, K. G. (1994) *Anal. Chem. 66*, 126−133.
40. Akashi, S., Noguchi, K., Yuji, R., Tagami, U., and Hirayama, K. (1996) *J. Am. Soc. Mass Spectrom. 7*, 707−721.
41. Hirayama, K., Yuji, R., Yamada, N., Kato, K., Arata, Y., and Shimada, I. (1998) *Anal. Chem. 70*, 2718−2725.
42. Roberts, G. D., Johnson, W. P., Burman, S., Anumula, K. R., and Carr, S. A. (1995) *Anal. Chem. 67*, 3613−3625.
43. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994) *Molecular Biology of the Cell*, 3rd ed., Garland, New York.
44. Young, R. J., Owens, R. J., Mackay, G. A., Chan, C. M., Shi, J., Hide, M., Francis, D. M., Henry, A. J., Sutton, B. J., and Gould, H. J. (1995) *Protein Eng. 8*, 193−199.
45. Kenten, J. H., Molgaard, H. V., Houghton, M., Derbyshire, R. B., Viney, J., Bell, L. O., and Gould, H. J. (1982) *Proc. Nat. Acad. U.S.A. 79*, 6661−6665.
46. Stephens, P. E., and Cockett, M. I. (1989) *Nucleic Acids Res. 17*, 7110.
47. Bebbington, C., Renner, G., Thomson, S., King, D., Abrams, D., and Yarranton, G. (1992) *BioTechnology 10*, 169−175.
48. Shi, J., Ghirlando, R., Beavil, R. L., Beavil, A. J., Keown, M. B., Young, R. J., Owens, R. J., Sutton, B. J., and Gould, H. J. (1997) *Biochemistry 36*, 2112−2122.

49. Kenten, J. H., Helm, G., Ishizaka, T., Cattini, P., and Gould, H. J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2955−2959.
50. Pridmore, R. D. (1987) *Gene 56*, 309−312.
51. Beu, S. C., Senko, M. W., Quinn, J. P., Wampler, F. M., and McLafferty, F. W. (1993) *J. Am. Soc. Mass Spectrom. 4*, 557−565.
52. Wilm, M., and Mann, M. (1996) *Anal. Chem. 68*, 1−8.
53. Marshall, A. G., Wang, T. C. L., and Ricca, T. L. (1985) *J. Am. Chem. Soc. 107*, 7893−7897.
54. Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) *J. Am. Chem. Soc. 120*, 3265−3266.
55. Zubarev, R. A., Kruger, N. A., Fridriksson, E. K., Lewis, M. A., Horn, D. M., Carpenter, B. K., and McLafferty, F. W. (1999) *J. Am. Chem. Soc. 121*, 2857−2862.
56. Horn, D. M., Zubarev, R. A., and McLafferty, F. W. (2000) *J. Am. Soc. Mass Spectrom.* (in press).
57. Keown, M. B., Ghirlando, R., Mackay, G. A., Sutton, B. J., and Gould, H. J. (1997) *Eur. Biophys. J. 25*, 471−476.
58. Nettleton, M. Y., and Kochan, J. P. (1995) *Int. Arch. Allergy Immunol. 107*, 328−329.
59. Basu, M., Hakimi, J., Dharm, E., Kondas, J. A., Tsien, W. H., Pilson, R. S., Lin, P., Gilfillan, A., Haring, P., Braswell, E. H., Nettleton, M. Y., and Kochan, J. P. (1993) *J. Biol. Chem. 268*, 13118−13127.
60. Baenziger, J., Kornfeld, S., and Kochwa, S. (1974) *J. Biol. Chem. 249*, 1889−1896.
61. Reinhold, B. B., and Reinhold, V. N. (1992) *J. Am. Soc. Mass Spectrom. 3*, 207−215.
62. Harvey, D. J., Rudd, P. M., Bateman, R. H., Bordoli, R. S., Howes, K., Hoyes, J. B., and Vickers, R. G. (1994) *Org. Mass Spectrom. 29*, 753−766.
63. Kuster, B., Wheeler, S. F., Hunter, A. P., Dwek, R. A., and Harvey, D. J. (1997) *Anal. Biochem. 250*, 82−101.
64. Morand, K., Galbo, G., and Mann, M. (1993) *Rapid Commun. Mass Spectrom. 7*, 738−743.
65. Padlan, E. A., and Davies, D. R. (1986) *Mol. Immunol. 23*, 1063−1075.
66. Helm, B. A., Ling, Y., Teale, C., Padlan, E. A., and Bruggemann, M. (1991) *Eur. J. Immunol. 21*, 1543−1548.

BI9919091