Characterization of the Glycosylation Sites in Cyclooxygenase-2 Using Mass Spectrometry †

ARTICLE *in* BIOCHEMISTRY · APRIL 2001

Impact Factor: 3.02 · DOI: 10.1021/bi015133r · Source: PubMed

5 AUTHORS, INCLUDING:

Jennifer F Nemeth

Janssen Research & Development, LLC

28 PUBLICATIONS **484** CITATIONS

Characterization of the Glycosylation Sites in Cyclooxygenase-2 Using Mass Spectrometry†

Jennifer F. Nemeth,‡ G. Phillip Hochensang, Jr.,§ Lawrence J. Marnett,§ and Richard M. Caprioli*,‡,§

Mass Spectrometry Research Center, and Department of Biochemistry, Vanderbilt University Medical Center, Nashville, Tennesee 37232-6400

Received October 3, 2000; Revised Manuscript Received January 8, 2001

ABSTRACT: Cyclooxygenase is involved in the biosynthesis and function of prostaglandins. It is a glycoprotein located in the endoplasmic reticulum and in the nuclear envelope, and it has been found to have two isoforms termed COX-1 and COX-2. This paper reports on the glycosylation site analysis of recombinant COX-2 using matrix-assisted laser desorption/ionization (MALDI) time-of-flight (TOF) mass spectrometry (MS) and nanoelectrospray (nanoESI) quadrupole-TOF (Q-TOF) MS. The nanoESI MS analysis of COX-2 revealed the presence of three glycoforms at average molecular masses of 71.4, 72.7, and 73.9 kDa. Each glycoform contained a number of peaks differing by 162 Da indicating heterogeneity and suggesting the presence of high-mannose sugars. The masses of the glycoforms indicate that oligosaccharides occupy two to four sites and a single *N*-acetylglucosamine (GlcNAc) residue occupied up to two sites. The MALDI MS analysis of a tryptic digest of the protein showed a number of potential glycopeptides. The peptides differed by 162 Da which further suggested high-mannose sugars. Nanoelectrospray MS/MS experiments confirmed glycosylation at the Asn 53 and Asn 130 sites and confirmed the presence of the peptides Asn 396−Arg 414 + GlcNAc and Thr 576−Arg 587 + GlcNAc containing Asn 580. It was not possible to conclusively determine whether the Asn 396 site was glycosylated via an MS/MS experiment, so the tryptic digest was deglycosylated to confirm the presence of the glycopeptides. Finally, a non-glycosylated tryptic peptide was observed containing the Asn 592.

Cyclooxygenases (COX)[1] are key enzymes in the biosynthesis and function of prostaglandins. The enzymes are membrane glycoproteins located in the endoplasmic reticulum and in the nuclear envelope and have been found to have two isoforms. Cyclooxygenase-1 (COX-1) is the product of a constitutively expressed gene and is widely distributed in mammalian tissues. Cyclooxygenase-2 (COX-2) is the product of an immediate-early gene and is only found in tissues stimulated with cytokines or growth factors (e.g., inflammatory lesions or cancers). The most abundant source of COX-1 is the sheep seminal vesicle which has been used to purify protein for biochemical and crystallographic analyses. Sheep seminal vesicle COX-1 contains four potential glycosylation sites, and extensive research into the oligosaccharide structure using NMR showed some occupation by two *N*-acetylglucosamines (GlcNAc) and nine Man ($Man_9GlcNAc_2$) sugars (*1*). Site-directed mutagenesis established that Asn 68, Asn 144, and Asn 410 are glycosylated, but that the fourth site, Asn 104, is not glycosylated (*2*, *3*).

Purified, recombinant COX-2 is most commonly produced by expression from baculovirus vectors in insect cells. The human and mouse proteins contain five consensus glycosylation sites, four of which appear to be occupied. Gel electrophoresis analysis reveals heterogeneity in glycosylation (*2*, *3*), likely due to the use of the polyhedron promotor on the baculovirus vector which is expressed late in the viral life cycle. Site-directed mutagenesis and peptide mapping have identified some of the glycosylation sites that are occupied on recombinant COX-2 (*2*−*4*). The actual occupation of the individual glycosylation sites has not been determined. This is of considerable interest because the extent of glycosylation appears important for catalytic activity in both COX-2 and COX-1 (*3*, *5*). Furthermore, recombinant COX-2 has been crystallized and its structure solved by X-ray analysis (*6*, *7*). The murine protein is particularly amenable to crystallization, and several structures of COX-2 substrate and COX-2- inhibitor complexes have been solved (*7*, *8*). Glycosyl residues appear in the crystal structures but only the first residue or two are detectable because of the flexibility of the oligosaccharide chain.

Mass spectrometry (MS) has been used as an effective method for the analysis of protein glycosylation (*9*−*12*); and tandem mass spectrometry (MS/MS), in combination with electrospray ionization, has been valuable in identifying

† The authors would like to acknowledge the National Institutes of Health/NCI (5R35 CA47479) for research funding, National Institutes of Health/NIGMS (GM 58008-02) for instrument funding, and the National Institutes of Health/NIEHS (T32 ES07028) for personal research funding.

* To whom correspondence should be addressed. Dr. Richard M. Caprioli, Vanderbilt University School of Medicine, Mass Spectrometry Research Center/824A MRBI, 23rd Ave. S. @ Pierce Ave., Nashville, TN 37232-6400. Phone: 615-322-4336, fax: 615-343-8372, e-mail: r.caprioli@vanderbilt.edu.

‡ Mass Spectrometry Research Center.

§ Department of Biochemistry.

[1] Abbreviations: COX, cyclooxygenase; MALDI MS, matrix-assisted laser desorption ionization mass spectrometry; TOF, time-of-flight; QTOF, quadrupole time-of-flight; MS/MS, tandem mass spectrometry; DTT, dithiothreitol; TFA, trifluoroacetic acid; HCCA, α-cyano-4-hydroxy-cinnamic acid; DHB, dihydroxybenzoic acid; HPLC, high performance liquid chromatography; EDTA, ethylenediamine tetraacetic acid; Hac, acetic acid; ESI, electrospray ionization.

10.1021/bi002313c CCC: $20.00
Published on Web 02/10/2001

glycopeptides. Through the use of precursor (parent) ion scans and marker ions such as *m*/*z* 163 [protonated hexose residue (Hex)], *m*/*z* 204 [protonated *N*-acetylhexosamine residue (HexNAc)], or *m*/*z* 366 (HexHexNAc) (*13*−*16*), it is possible to easily identify glycopeptides. Dissociation of protonated glycopeptides in a collision-induced decomposition (CID) experiment (*17*−*22*) also provides structural information regarding the amino acid sequence of the peptide, the types of sugars attached, and the residue that carries the glycosyl group. This paper reports the tryptic peptide mapping of recombinant mouse COX-2 (mCOX-2) and the glycosylation analysis of the protein. The experiments were performed using a combination of matrix-assisted laser/desorption (MALDI) time-of-flight (TOF) MS and nanoelectrospray (nanoESI) quadrupole-TOF (Q-TOF) MS.

EXPERIMENTAL PROCEDURES

Chemicals. Modified porcine trypsin was obtained from Promega Corporation (Madison, WI); recombinant *N*-glycanase (*N*-glycosidase F) was from Glyko (Novato, CA); ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide, trifluoroacetic acid (TFA), acetic acid (HAc), α-cyano-4-hydroxy-cinnamic acid (HCCA), and 2,5-dihydroxybenzoic acid (DHB) were from Sigma (St. Louis, MO). All solvents used were HPLC grade, and the water was purified with a MilliQ filtration system.

COX-2 Construction/Expression/Purification. The expression of mCOX-2 in a pVL1393 baculovirus expression vector, followed by amplification and purification, have been described elsewhere (*22*, *23*) with the following modifications: (i) after expression, 8 L of cells were grown up instead of 4 L, and (ii) the protein was purified using a 620-mL Sephacryl column as opposed to a 90-mL Superose 12 column. After purification, the protein was stored at −70 °C in 20 mM Tris-HCl, 0.4% CHAPS, 0.15 M NaCl, pH 8.

Before mass spectrometric analyses, aliquots of the protein (approximately 7−10 pmol/μL) were prepared in 100 mM Tris (pH 7.0) containing 5 mM phenol. The sample was diluted to 250 μL with 60:40 H_2O/acetonitrile containing 0.1% TFA and further purified using reverse-phase high-performance liquid chromatography (HPLC) to remove salts. The HPLC separation was performed using a Vydac C4 analytical column (0.46 × 25 cm) with 5 μm particles, 300 Å pore size (Hesperia, CA). Samples were loaded into a 250-μL sample loop, and a Varian model 9010 HPLC system (Palo Alto, CA) was used to produce the solvent gradient at a flow rate of 1.0 mL/min. Solvent A was 0.1% TFA and solvent B was 80:20 acetonitrile/H_2O with 0.1% TFA. The gradient increased from 50% B to 75% B in 30 min, and then 75% B to 100% B in 10 min. The protein was monitored at 230 nm with a Varian model 2050 UV detector (Palo Alto, CA). The protein was collected in an eppendorf tube and lyophilized to approximately one-third its original volume, after which, it was stored at −80 °C until use.

Tryptic Digestion. The protein (10 μL containing approximately 100 pmol of protein) was mixed with 25 μL of 0.4 M NH_4HCO_3 and 5 μL of 45 mM DTT and incubated for 20 min at 60 °C. The solution was allowed to cool to room temperature (RT), after which 5 μL of a 100 mM solution of iodoacetamide in water was added, and the solution was incubated at RT for 30 min in the dark. Five microliters of modified trypsin (0.1 μg/μL) was then added, and the solution was incubated at 37 °C for 24 h. The digest was stored at −80 °C until use. Before analysis, aliquots of digest were purified/desalted using Millipore C18 ZipTips (Bedford, MA) using 50:50:0.5:0.1 2-propanol/acetonitrile/HAc/TFA (v/v/v/v) as the elution solution. The resulting solution was used for both MALDI and nanoESI MS analyses.

Deglycosylation of COX-2 Tryptic Digest. The tryptic digest of COX-2 (20 μL) was incubated with 1.25 mU recombinant *N*-glycanase in 20 mM sodium phosphate (pH 7.5), 50 mM EDTA, 0.02% (w/v) sodium azide for 24 h at 37 °C. The digest was purified using a ZipTip before analysis as described above.

MS Analyses. All MALDI mass spectra were acquired on an Applied Biosystems Voyager STR MALDI MS equipped with delayed extraction. All spectra were acquired at 25 kV in the linear mode. The MALDI matrixes used were either HCCA or DHB prepared at a concentration of 10 mM in 50:50 H_2O/acetonitrile with 0.1% TFA. All electrospray MS and MS/MS mass spectra were acquired on a SCIEX QSTAR (Q-TOF) MS/MS instrument equipped with a nanoESI source supplied by Protana (Odense, Denmark). Purification emitters (Protana) were used to transfer the samples into the mass spectrometer. The emitters were manually opened before sample loading by tapping the tip of the capillary onto the bottom of a beaker filled with methanol while a 20 psi backpressure was applied. After sample loading, a stainless steel wire was inserted into the capillary almost to the tip to allow for voltage to be applied to the solution. An O-ring was fixed over the end of the capillary and the tail end of the wire to keep it in place. A 3-psi backpressure was applied to the emitter to aid in starting, and maintaining, the electrospray. The accelerating voltage was 700 V, the orifice voltage was 75 V, the ring voltage was 275 V, and the Q_0 voltage was maintained at 20 V for MS scans and varied from 40 to 65 V for MS/MS scans. The collision gas pressure was 3×10^{-5} PSI for MS scans and varied from 4 to 6×10^{-5} PSI for MS/MS scans. The curtain gas was maintained at 15 psi.

RESULTS/DISCUSSION

There are five possible glycosylation sites in mCOX-2. Three of these sites are conserved in COX-1 (Table 1). An analysis of the intact COX-2 protein was performed to determine the number of occupied sites. The nanoESI mass spectrum of the protein is shown in Figure 1, panel a, and the deconvolution of this mass spectrum, Figure 1, panel b, shows the presence of three protein glycoforms (indicated by the 1, 2, and 3) at average masses of 71.4, 72.7, and 73.9 kDa. The mass of glycoform 1 corresponds to oligosaccharides with an average sugar residue of $Man_8GlcNAc_2$ at two sites and a GlcNAc residue at two sites. Glycoform 2 corresponds to oligosaccharides with an average sugar residue of $Man_7GlcNAc_2$ at three sites and a GlcNAc residue at one site. Glycoform 3 corresponds to oligosaccharides with an average sugar residue of $Man_7GlcNAc_2$ at four sites. The mass difference between peaks is nominally 162 Da indicating that the sugar residues are hexose, most likely mannose. This is in agreement with the literature report that indicates that the oligosaccharides are high-mannose (*3*). Finally, there

Table 1: A List of Tryptic Peptides that Contain Potential Glycosylation Sites[a]

ID	tryptic peptides containing potential glycosylation sites	amino acids	glycosylation site	peptide mass (Da)	conserved in COX-1	references
C5	TGFYGE***NCT***TPEFLTR	47−62	Asn 53	1891.84	YES	
C12	SWEAFS***NLS***YYTR	124−136	Asn 130	1622.73	YES	*4*
C39	QFLY***NNS***ILLEHGLTQFVESFTR	392−414	Asn 396	2755.39	YES	
C54	TATI***NAS***ASHSR	576−587	Asn 580	1214.60	NO	*2, 3*
C55	LDDI***NPT***VLIK	588−598	Asn 592	1239.70	NO	*2, 3*

[a] The peptide ID indicates the peptide's position within the protein sequence (i.e., C5 is the fifth tryptic peptide from the *N*-terminal end of COX-2). The listing in column 6 indicates conservation of the site in COX-1, and column 7 lists references that have reported on glycosylation information for the site.

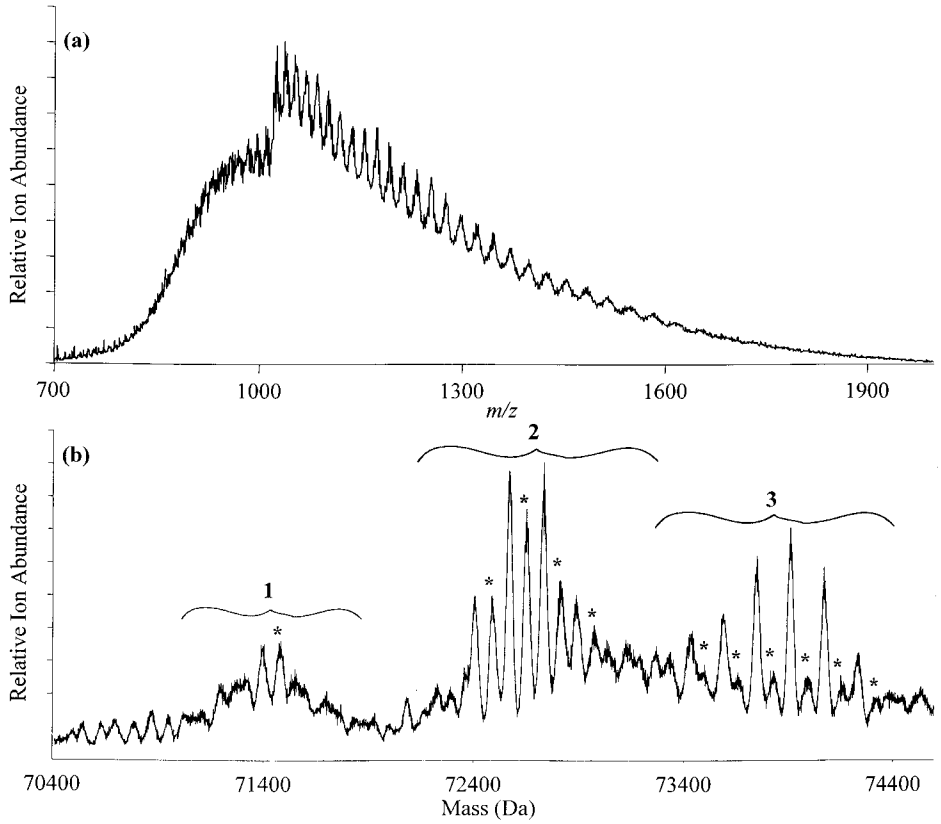

FIGURE 1: Protein analysis of mCOX-2: (a) nanoESI MS mass spectrum, and (b) the zero-charge deconvolution of panel a (100 iterations with a 0.5 step size). Three glycoforms of the protein are observed corresponding to oligosaccharides occupying two to four sites and a single GlcNAc residue at up to two sites. In addition, there appears to be a second form of the protein (labeled * in panel b), indicating the possible presence of a phosphate or sulfate group.

appears to be two forms of the protein for each glycoform (peaks from second form indicated by the *). The difference between the two sets of peaks is approximately 80 Da and could indicate the presence of either a phosphate or sulfate group. Although not observed in the tryptic peptides recorded, this is currently the subject of further investigation.

Further analyses were performed on a tryptic digest of mCOX-2 to identify glycopeptides using MALDI MS and ESI MS. The MALDI mass spectrum of the tryptic digest is shown in Figure 2 , and the data are tabulated in Table 2. Of the few peptides that were not observed, three of them were tryptic peptides C5, C12, and C54, and they all contain potential glycosylation sites (columns 4 and 5, Table 2). In addition, a peptide was recorded at *m*/*z* 2757.0 Da (ave. mass). This ion has a *m*/*z* value that falls between the theoretical *m*/*z* for two peptides, non-glycosylated C39 and tryptic peptide C51 (Figure 2, panel a). An MS/MS analysis revealed C51 to be in high abundance and C39 to be in low abundance (data not shown). Also, a peptide was observed at *m*/*z* 2836.2 (ave. mass) that corresponds to (C39 + GlcNAc) (Figure 2, panel a). No peaks were identified that would correspond to the addition of GlcNAc at the other four sites in the MALDI analysis; however, a doubly charged ion was observed for (C54 + GlcNAc) in the nanoESI MS analysis. An ion was observed for tryptic peptide C55 that contains Asn 592, the reported non-glycosylated site (Figure 2, panel a) (*2*, *3*). It is not surprising that this site is free because of the proline in the second position in the NX(S/T) sequence. The presence of a proline in the second position is generally thought to prevent glycosylation (*24*).

There were a number of unidentified ions recorded in the MALDI MS mass spectrum that did not correspond to theoretical (unmodified) tryptic peptides and therefore may be glycopeptides. Figure 2, panels b and c, are enlarged portions of Figure 2, panel a, where it is seen that several peaks differ by 162 Da, the mass difference of a hexose residue. The first two sets overlap in Figure 2, panel b, and the third set is observed in Figure 2, panel c. Assuming that

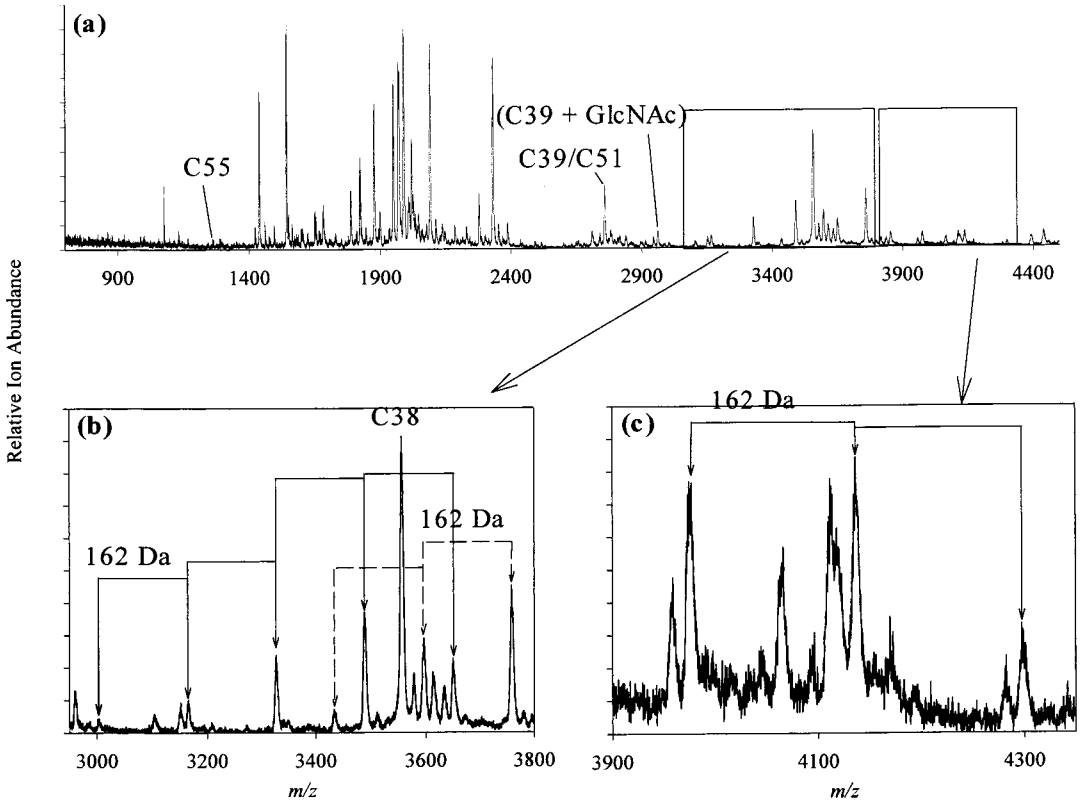

FIGURE 2: MALDI MS analysis of the tryptic digest of COX-2: (a) the mass spectrum of the digest analysis with relevant peaks and mass ranges marked; (b) inset region containing the glycopeptide series C12 (−) and C5 (- - -); (c) inset region containing the glycopeptide series C39. The matrix used in this analysis was HCCA.

Table 2: A List of Tryptic Peptides for COX-2[a]

ID	amino acids	mass (Da)	observed by MS	contains NX(S/T)	ID	amino acids	mass (Da)	observed by MS	contains NX(S/T)
C2	N−29[b,c]	1477.6	YES		C33	304−319	1985.9	YES	
C3	30−41[c]	1421.6	YES		C34	320−328	998.6	YES	
C4	42−46[c]	710.2			C35	329−344	1947.0	YES	
C5	47−62[c]	1891.8		NCT	C37	347−362	2086.0	YES	
C7	65−82	2134.2	YES		C38	363−391	3553.7	YES	
C8	83−95	1540.9	YES		C39	392−414	2755.5	?[d]	NNS
C9	96−100	590.3			C39 + GlcNAc	392−414	2834.4	YES	NNS
C10	101−106	737.4	YES		C40	415−419	543.3		
C11	107−123	2016.0	YES		C41	420−424	458.3		
C12	124−136	1622.8		NLS	C42	425−435	1108.7	YES	
C13	137−152[c]	1666.8	YES		C43	436−442	775.4		
C15	156−161	687.3			C45	446−453	1071.5	YES	
C16	162−166	616.3			C48	456−471	1874.9	YES	
C17	167−170	499.3			C49	472−478	790.4	YES	
C19	172−197	3148.4	YES		C50	479−518	4437.3	YES	
C20	198−201	499.2			C51	519−543	2756.3[c]	?[d]	
C22	203−208	633.3	YES		C52	544−559	1787.0[c]	YES	
C23	209−226	1965.0	YES		C53	560−575	1821.8[c]	YES	
C29	240−253	1548.8	YES		C54	576−587	1214.6		NAS
C30	254−293	4602.3	YES		C54 + GlcNAc	576−587	1293.6	YES[e]	NAS
C31	294−297	554.3			C55	588−598	1239.7	YES	NPT
C32	298−303[c]	746.4							

[a] Only peptides within the analyzed mass range (450−5000) are listed. Column 3 lists the theoretical masses, column 4 lists whether the peptides were observed using MALDI MS, and column 5 shows any potential glycosylation sites. [b] *N*-terminal processing. [c] Carboxyamidated. [d] Peptides were within one mass unit and could not be conclusively identified. [e] Only observed by ES MS.

the oligosaccharides are high-mannose, and the composition of the chains includes two GlcNAc residues and 6−10 Man residues, it is possible to postulate compositions consistent with these masses. The first group of peaks in Figure 2, panel b (solid line) were calculated to correspond to the C12 peptide (containing Asn 130) with sugar compositions ranging from $Man_6GlcNAc_2$ to $Man_{10}GlcNAc_2$, the second group of peaks in Figure 2, panel b (dashed line) were calculated to correspond to the C5 peptide (containing Asn 53) with sugar compositions ranging from $Man_7GlcNAc_2$ to $Man_9GlcNAc_2$, and the third group of peaks in Figure 2, panel c, were calculated to correspond to the peptide C39

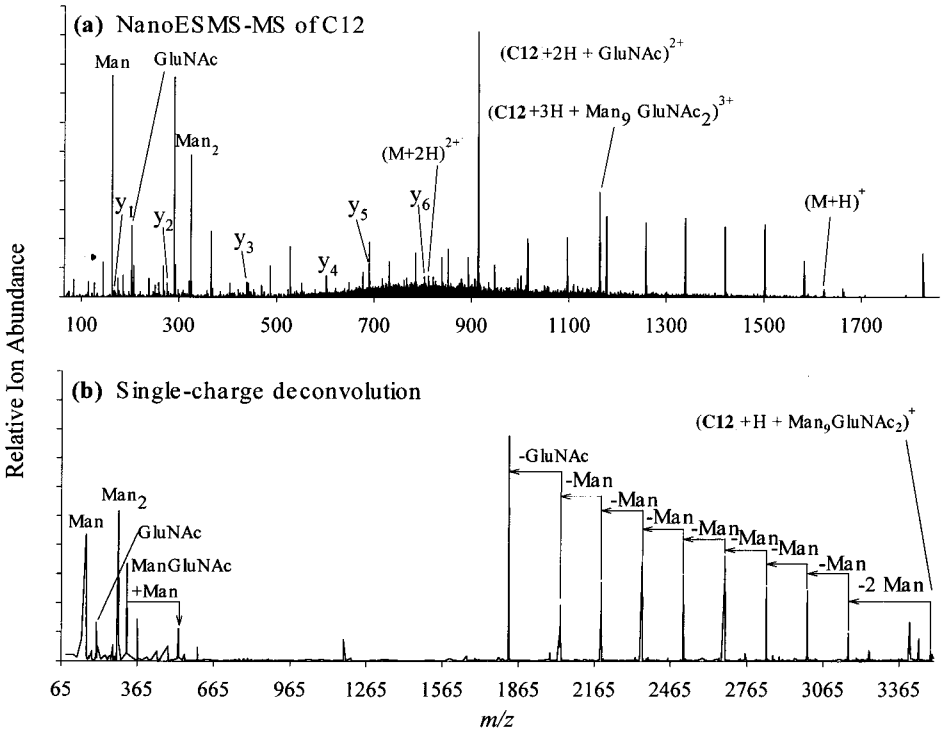

FIGURE 3: NanoESI MS/MS analysis of $(C12 + 3H + Man_9GlcNAc_2)^{3+}$ ions: (a) the mass spectrum where y ions are indicated, and (b) the singly charged deconvoluted mass spectrum of panel a where the sugar-loss ion series is indicated. Refer to Table 3 for a complete listing of the observed product ions.

Table 3: A List of the Observed Product Ions from the nanoES MS/MS Analysis of $(C12 + 3H + Man_9GlcNAc_2)^{3+}$ Ions

observed free peptide ions	observed *m/z*	observed free sugar ions[a,b]	observed *m/z*	observed (peptide + sugar)$^{2+}$ ions[a,b]	observed *m/z*	observed (peptide + sugar)$^{3+}$ ions[a,b]	observed *m/z*
y_1^+	175.0	$(M + H)^+$	163.1	$(C12 + 2H + G)^{2+}$	913.9	$(C12 + 3H + G_2)^{3+}$	677.1
y_2^+	276.2	$(G + H)^+$	204.1	$(C12 + 2H + G_2)^{2+}$	1015.4	$(C12 + 3H + MG_2)^{3+}$	731.1
y_3^+	439.2	$(M_2 + H)^+$	325.1	$(C12 + 2H + MG_2)^{2+}$	1096.4	$(C12 + 3H + M_2G_2)^{3+}$	785.1
y_4^+	602.3	$(MG + H)^+$	366.1	$(C12 + 2H + M_2G_2)^{2+}$	1177.4	$(C12 + 3H + M_3G_2)^{3+}$	839.2
y_5^+	689.3	$(M_2G + H)^+$	528.2	$(C12 + 2H + M_3G_2)^{2+}$	1258.5	$(C12 + 3H + M_4G_2)^{3+}$	893.2
y_6^+	802.8	$(M_3G + H)^+$	690.2	$(C12 + 2H + M_4G_2)^{2+}$	1339.5	$(C12 + 3H + M_5G_2)^{3+}$	947.2
$(C12 + H)^+$	1624.7	$(M_4G + H)^+$	852.2	$(C12 + 2H + M_5G_2)^{2+}$	1421.0	$(C12 + 3H + M_6G_2)^{3+}$	1001.2
$(C12 + 2H)^{2+}$	812.3	$(M_5G + H)^+$	1014.3	$(C12 + 2H + M_6G_2)^{2+}$	1502.0	$(C12 + 3H + M_7G_2)^{3+}$	1055.2
		$(M_6G + H)^+$	1176.4	$(C12 + 2H + M_7G_2)^{2+}$	1583.1	$(C12 + 3H + M_8G_2)^{3+}$	1109.6
		$(M_7G + H)^+$	1338.4			$(C12 + 3H + M_9G_2)^{3+}$	1163.6
		$(M_8G + H)^+$	1500.5				
		$(M_9G + H)^+$	1662.5				

[a] Mannose = M. [b] *N*-acetylglucosamine = G.

(containing Asn 396) with sugar compositions ranging from $Man_5GlcNAc_2$ to $Man_7GlcNAc_2$.

Nanoelectrospray MS/MS analyses were used to confirm that these were indeed glycopeptides. The most abundant multiply charged ions in each peptide series were dissociated in CID experiments. It was possible to quickly determine if the peptide under dissociation was a glycopeptide by the presence of fragment residue ions at *m/z* 163 and *m/z* 204 corresponding to protonated free Hex and GlcNAc, respectively.

The first putative glycopeptide series to be analyzed was tryptic peptide C12 which contained Asn 130 (Figure 2, panel b, solid line). Multiply charged ions corresponding to the peptide plus $Man_9GlcNAc_2$ and $Man_{10}GlcNAc_2$ were dissociated using nanoESI MS/MS. The nanoESI MS/MS spectrum generated from the dissociation of *m/z* 1163.6, corresponding to $(C12 + 3H + Man_9GlcNAc_2)^{3+}$ ions, is shown in Figure 3, panel a. Ions from Man and GlcNAc were clearly visible at the lower end of the mass spectrum and a number of y ions were observed (Figure 3, panel a). The y ion series goes up to, but does not include, the glycosylated asparagine (refer to Table 1 for the amino acid sequence). Deconvolution of the triply charged mass spectrum to a singly charged mass spectrum made possible by the high-resolution QTOF technology (Figure 3, panel b) showed the stepwise loss of sugars down to the $(C12 + H + GlcNAc)^+$. Table 3 provides a list of ions identified from Figure 3, panel a.

The second glycopeptide series to be analyzed was the tryptic peptide C5 containing Asn 53. The peptides suspected of containing the $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ oligosaccharides were collisionally dissociated using nanoESI MS/MS. The nanoESI MS/MS spectrum generated from the dissociation of *m/z* 1199.3, corresponding to $(C5 + 3H + Man_8GlcNAc_2)^{3+}$ ions, is shown in Figure 4, panel a. As before, Man and GlcNAc fragment ions are present in the mass spectrum and there are y ions that preceded up to, but did not include, the glycosylated asparagine (refer to Table

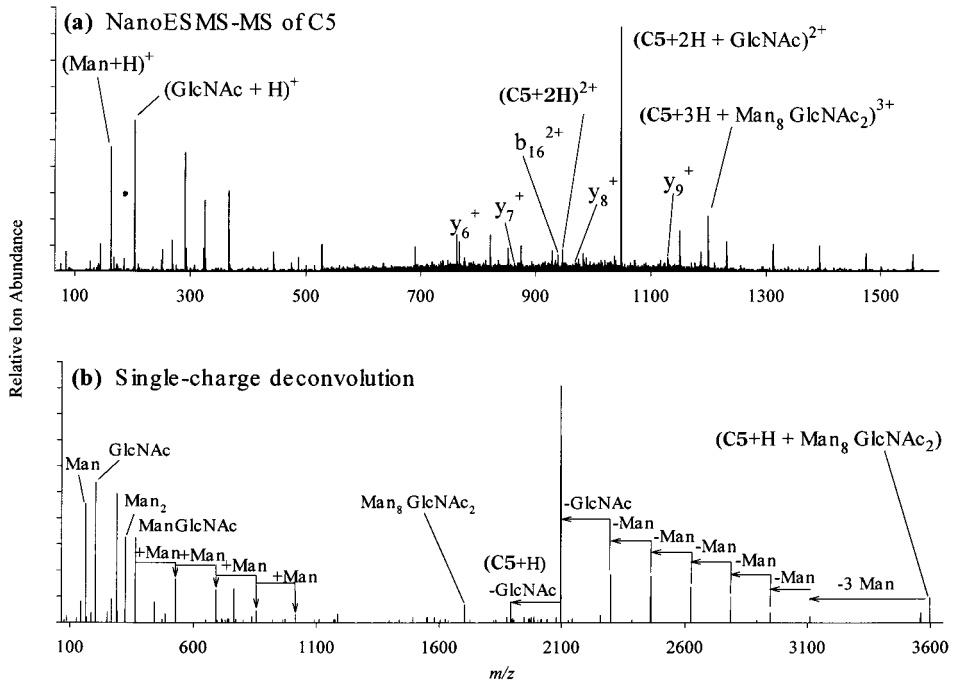

FIGURE 4: NanoESI MS/MS analysis of $(C5 + 3H + Man_8GlcNAc_2)^{3+}$ ions: (a) the mass spectrum where y ions are indicated, and (b) the singly charged deconvoluted mass spectrum of panel a where the sugar-loss ion series is indicated. Refer to Table 4 for a complete listing of the observed product ions.

Table 4: A List of the Observed Product Ions from the nanoES MS/MS Analysis of $(C5 + 3H + Man_8GlcNAc_2)^{3+}$ Ions

observed free peptide ions	observed *m/z*	observed free sugar ions[a,b]	observed *m/z*	observed (peptide + sugar)$^{2+}$ ions[a,b]	observed *m/z*	observed (peptide + sugar)$^{3+}$ ions[a,b]	observed *m/z*
y_6^+	762.5	$(M + H)^+$	163.1	$(C5 + 2H + G)^{2+}$	1048.5	$(C5 + 3H + G_2)^{3+}$	767.0
y_7^+	863.5	$(G + H)^+$	204.1	$(C5 + 2H + G_2)^{2+}$	1150.1	$(C5 + 3H + MG_2)^{3+}$	821.1
y_8^+	964.5	$(M_2 + H)^+$	325.1	$(C5 + 2H + MG_2)^{2+}$	1231.1	$(C5 + 3H + M_2G_2)^{3+}$	875.1
y_9^+	1124.5	$(MG + H)^+$	366.1	$(C5 + 2H + M_2G_2)^{2+}$	1312.1	$(C5 + 3H + M_3G_2)^{3+}$	929.1
b_{16}^{2+}	938.4	$(M_2G + H)^+$	528.2	$(C5 + 2H + M_3G_2)^{2+}$	1393.1	$(C5 + 3H + M_4G_2)^{3+}$	983.1
$(C5 + 2H)^{2+}$	947.5	$(M_3G + H)^+$	690.2	$(C5 + 2H + M_4G_2)^{2+}$	1474.6	$(C5 + 3H + M_6G_2)^{3+}$	1091.1
		$(M_4G + H)^+$	852.3	$(C5 + 2H + M_5G_2)^{2+}$	1558.8	$(C5 + 3H + M_8G_2)^{3+}$	1199.3
		$(M_5G + H)^+$	1014.3	$(C5 + 2H + M_6G_2)^{2+}$	1634.2		
		$(M_6G + H)^+$	1176.4				
		$(M_8G + H)^+$	1500.5				

[a] Mannose = M. [b] *N*-acetylglucosamine = G.

1 for the amino acid sequence). The deconvolution of this mass spectrum to the singly charged spectrum, Figure 4, panel b, gave a stepwise loss of sugar residues from the precursor down to the free peptide. In addition, the free sugar chain was observed. Table 4 provides a complete list of identified ions from Figure 4, panel a.

The final conserved site, Asn 396 (peptide C39), could not be completely sequenced using nanoESI MS/MS. The ion intensities for this glycopeptide series were far lower than observed for the C5 and C12 glycopeptides (Figure 2, panel c). Furthermore, the C39 peptide mass is almost twice that of either the C5 or C12 peptide. It is suspected that either ion suppression effects, the increased peptide mass, or a combination of both prevented efficient dissociation and/or detection of the product ions. What was detectable in the MS/MS analysis was fragment ions of Man and GlcNAc, and these marker ions are a good indication of a glycopeptide (data not shown). To confirm that the ions observed in the MALDI MS analyses (Figure 2, panel c) are glycopeptides, the sample was deglycosylated using a specific glycosidase. The MALDI MS mass analyses are shown in Figure 5: Figure 5, panel a, shows the mass spectrum with suspected glycopeptides for C5, C12, and C39, and Figure 5, panel b, shows the mass spectrum after deglycosylation. In all cases, the suspected glycopeptides were no longer detected and new peaks at *m/z* 1891.5, 1622.7, and 2756.0 appeared corresponding to non-glycosylated C5, C12, and C39, respectively. This, in addition to the MS/MS data, strongly indicated that C39 is glycosylated.

Finally, MS/MS experiments were performed on the peptides suspected of being $(C39 + GlcNAc)^{2+}$ and $(C54 + GlcNAc)^{2+}$ (data not shown). In both cases, the MS/MS mass spectra contained fragment ions of GlcNAc at *m/z* 204. In addition, two sets of y ions were observed in the upper *m/z* region of the mass spectra for C39 and C54 corresponding to the peptide with and without the GlcNAc residue, and the sequence ions confirm the location of the sugar to be on the asparagine. On the basis of the precursor ion intensities, it would appear that the $(C54 + GlcNAc)^{2+}$ ions are in greater abundance than the $(C39 + GlcNAc)^{2+}$ ions. The presence of these truncated glycopeptides may have resulted from processing of the oligosaccharides after expression or

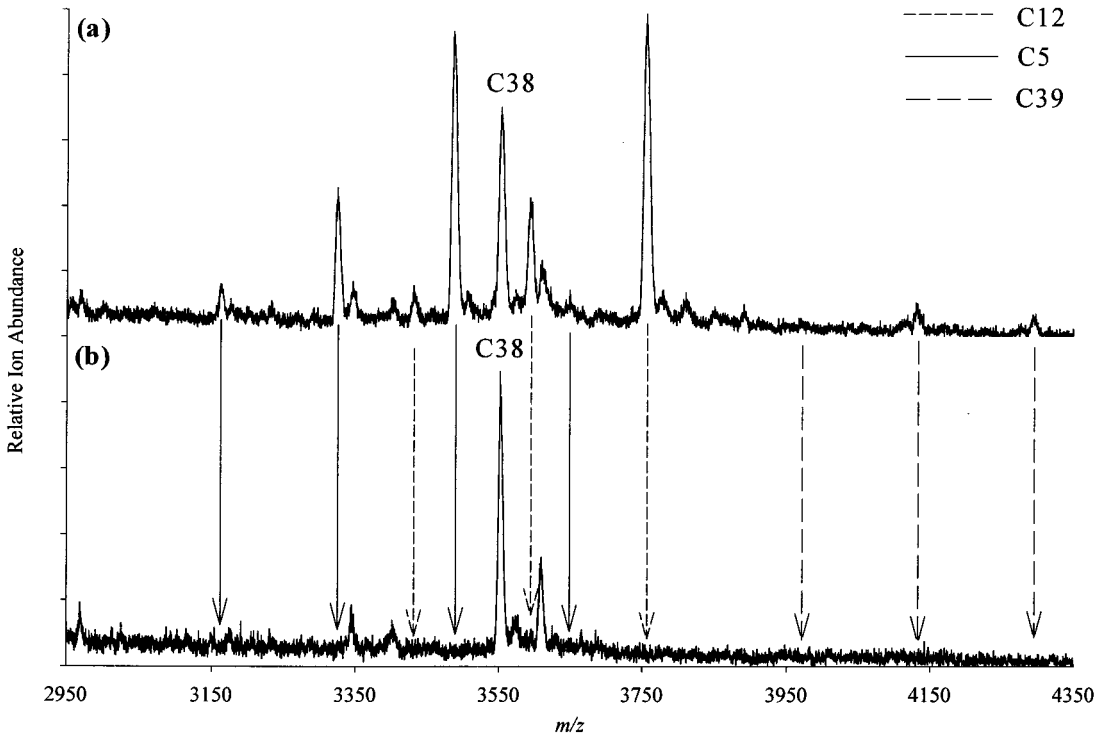

FIGURE 5: MALDI MS analyses of (a) COX-2 tryptic digest before deglycosylation, and (b) after deglycosylation. All of the suspected glycopeptides were not detected after deglycosylation.

Glycoform	MW* (kDa)	Asn 53	Asn 130	Asn 396	Asn 580	Asn 592
1	71.4	$Man_{7-9}GlcNAc_2$	$Man_{6-10}GlcNAc_2$	GlcNAc	GlcNAc	None
2	72.7	$Man_{7-9}GlcNAc_2$	$Man_{6-10}GlcNAc_2$	GlcNAc	$Man_7GlcNAc_2$	None
2	72.7	$Man_{7-9}GlcNAc_2$	$Man_{6-10}GlcNAc_2$	$Man_{5-7}GlcNAc_2$	GlcNAc	None
3	73.8	$Man_{7-9}GlcNAc_2$	$Man_{6-10}GlcNAc_2$	$Man_{5-7}GlcNAc_2$	$Man_7GlcNAc_2$	None

* Average molecular weight for the glycoform

FIGURE 6: Summary of the glycosylation sites and extent of glycosylation observed for each site in COX-2.

the result of deglycosylation by a glycosidase that cuts between the two GlcNAc residues.

For the peptides that were sequenced using tandem mass spectrometry, the majority of the product ions were either free sugar ions or intact peptide ions containing some portion of the oligosaccharide moiety (Tables 3 and 4). It is evident that the energy required to remove sugar residue(s) from the peptide is far lower than the energy required for peptide backbone cleavages (*17*), and this accounts for the limited number of backbone product ions that are present. In both cases, the y ion series were limited to the C-terminal end of the peptide most likely because of the charge residing on the basic C-terminal amino acid.

The stepwise loss of sugars from the oligosaccharide chain offers a convenient way to confirm the number, and in some cases type(s), of sugars in the structure. Of course, it is not possible to differentiate between glucose, mannose, and galactose as they all give a mass shift of 162 Da. In addition, because low energy CID conditions generate multiple bond cleavages, it is not possible to derive a great deal of connectivity information on the sugar branching as is the case when using high-energy CID and single ion collisions (*25*).

CONCLUSIONS

From the protein analysis, it was possible to determine that COX-2 is heterogeneously glycosylated, contains variations in sugar residues that differ by 162 Da indicating hexose sugars, and has three glycoforms. Form 1 contains oligosaccharides at two sites and a GlcNAc residue at two sites, form 2 contains oligosaccharides at three sites and a GlcNAc residue at one site, and form 3 contains oligosaccharides at four sites.

The MS and MS/MS analyses of the tryptic digest of COX-2 revealed that Asn 53 and Asn 130 are close to being 100% glycosylated, and that Asn 396 and Asn 580 are partially glycosylated. The sugar compositions for the sites ranged from $Man_5GlcNAc_2$ to $Man_{10}GlcNAc_2$. In addition, a percentage of the Asn 396 and Asn 580 sites were found to contain a single GlcNAc residue. Finally, Asn 592 site is not glycosylated. These findings are also summarized in Figure 6.

REFERENCES

1. Mutsaers, J. H., Van Halbeek, H., Kamerling, J. P., and Vliegenthart, J. F. (1985) *Eur. J. Biochem 147*, 569.

2. Otto, J. C., DeWitt, D. L., and Smith, W. L. (1993) *J. Biol. Chem. 268*, 18234.
3. Percival, M. D., Bastien, L., Griffin, P. R., Kargman, S., Ouellet, M., and O'Neill, G. P. (1997) *Protein Expr. Purif. 9*, 388.
4. Wennogle, L. P., Liang, H., Quintavalla, J. C., Bowen, B. R., Wasvary, J., Miller, D. B., Allentoff, A., Boyer, W., Kelly, M., and Marshall, P. (1995) *FEBS Lett. 371*, 315.
5. Shimokawa, T., and Smith, W. L. (1992) *Biochem. Biophys. Res. Commun. 183*, 975−982.
6. Luong, C., Miller, A., Barnett, J., Chow, J., Ramesha, C., and Browner, M. F. (1996) *Nat. Struct. Biol. 3*, 927−933.
7. Kurumbail, R. G., Stevens, A. M., Gierse, J. K., McDonald, J. J., Stegeman, R. A., Pak, J. Y., Gildehaus, D., Miyashiro, J. M., Penning, T. D., Seibert, K., Isakson, P. C., and Stallings, W. C. (1996) *Nature 384*, 644−648.
8. Kiefer, J. R., Pawlitz, J. L., Moreland, K. T., Stegeman, R. A., Hood, W. F., Gierse, J. K., Stevens, A. M., Goodwin, D. C., Rowlinson, S. W., Marnett, L. J., Stallings, W. C., and Kuruimbail, R. G. (2000) *Nature 405*, 97−101.
9. Carr, S. A., and Roberts, G. D. (1986) *Anal. Biochem. 157*, 396.
10. Medzihradszky, K. F., Maltby, D. A., Hall, S. C., Settineri, C. A., and Burlingame, A. L. (1994) *J. Am. Soc. Mass Spectrom. 5*, 350.
11. Denzinger, T., Diekmann, H., Bruns, K., Laessing, U., Stuermer, C. A., and Przybylski, M. (1999) *J. Mass Spectrom. 34*, 435.
12. Zeng, C., and Biemann, K. (1999) *J. Mass Spectrom. 34*, 311.
13. Hunter, A. P., and Games, D. E. (1995) *Rapid Commun. Mass Spectrom. 9*, 42.
14. Jedrezejewski, P. T., and Lehmann, W. D. (1997) *Anal. Chem. 69*, 294.
15. Colangelo, J., Licon, V., Benen, J., Visser, J., Bergmann, C., and Orlando, R. (1999) *Rapid Commun. Mass Spectrom. 13*, 1448.
16. Colangelo, J., Licon, V., Benen, J., Visser, J., Bergmann, C., and Orlando, R. (1999) *Rapid Commun. Mass Spectrom. 13*, 2382.
17. Conboy, J., and Henion, J. D. (1992) *J. Am. Soc. Mass Spectrom. 3*, 804.
18. Swiderek, K. M., Pearson, C. S., and Shively, J. E. (1993) *Technol. Protein Chem. IV. 4*, 127.
19. Bauer, S. H. J., Zhang, X.-Y., Van Dongen, W., and Claeys, M. (1999) *Anal. Biochem. 274*, 69.
20. Kimura, Y., Hess, D., and Sturm, A. (1999) *Eur. J. Biochem. 264*, 168.
21. Fridriksson, E. K., Beavil, A., Holowka, D., Gould, H. J., Baird, B., and McLafferty, F. W. (2000) *Biochem*istry. *39*, 3369.
22. Rowlinson, S. W., Crews, B.C., Goodwin, D.C., Schneider, C., Gierse, J. K., and Marnett, L. J. (2000) *J. Biol. Chem. 275*, 6586.
23. Rowlinson, S. W., Crews, B. C., Lanzo, C. A., and Marnett, L. J. (1999) *J. Bio. Chem. 274*, 23305.
24. Lennarz, W. J. (1980) in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, W. J., Ed.) p 48, Plenum Press: New York.
25. Gillece-Castro, B., and Burlingame, A. L. (1990) *Methods Enzymol. 193*, 689.

BI002313C