Available online at www.sciencedirect.com
SCIENCE DIRECT®
ACADEMIC PRESS
Bioorganic Chemistry 30 (2002) 356–370
BIOORGANIC CHEMISTRY
www.elsevier.com/locate/bioorg

N-Glycosylation pattern of the zymogenic form of human matrix metalloproteinase-9

Lakshmi P. Kotra,[a,1] Li Zhang,[b] Rafael Fridman,[c] Ron Orlando,[b,2] and Shahriar Mobashery[a,*]

[a] *Department of Chemistry, Wayne State University, Detroit, MI 48202, USA*
[b] *Complex Carbohydrate Research Center, The University of Georgia, Athens, GA 30602, USA*
[c] *Department of Pathology, School of Medicine, Wayne State University, Detroit, MI 48201, USA*

Received 7 August 2002

Abstract

Glycosylation of proteins has profound consequences on the activities of macromolecules and their interactions with inhibitors/substrates. Matrix metalloproteinase-9 (MMP-9, also known as gelatinase B) is a member of the MMP family of zinc-dependent endopeptidases, with critical functions in both physiological and pathological processes. MMP-9, a glycosylated MMP, is implicated in inflammation, angiogenesis and tumor metastasis. We have determined by the use of mass spectrometry that of the three possible N-glycosylation sites in human MMP-9 only two are glycosylated. The N-glycosylation sites are at asparagines in positions 38 and 120, the first site within the propeptide domain of the zymogenic form (pro-MMP-9) of the enzyme and the second in the catalytic domain. The chemical nature of the sugar attachments to both these sites was determined by mass spectrometry. Both N-glycosylation sites have NeuAcα(1,2)-Galβ(1,4)-GlcNAcβ(1,2)-Manα(1,3)-[NeuAcα(1,2)-Galβ(1,4)-GlcNAcβ(1,2)-Manα(1,6)-]Manβ(1,4)-GlcNAcβ(1,4)-

* Corresponding author. Fax: 1-313-577-8822.
E-mail address: som@chem.wayne.edu (S. Mobashery).
[1] Present address: Faculty of Pharmacy, University of Toronto, Toronto, Ont., Canada M5S 2S2.
[2] Also Corresponding author.

0045-2068/02/$ - see front matter
PII: S0045-2068(02)00501-1

[Fucα(1,6)-]GlcNAcβ oligosaccharide chains. A computational model of glycosylated pro-MMP-9 was generated and it was studied by dynamics simulations.

Keywords: Matrix metalloproteinases-9; Glycosylation; Mass spectrometry; Molecular simulations; Zymogen activation

1. Introduction

The implications of posttranslational modifications of proteins are just beginning to be understood [1,2]. Glycosylation in proteins, especially in mammalian systems, is ubiquitous and is clearly important for biological functions of many proteins [1,2]. A challenge in study of many of these glycosylated proteins is that their structures are often difficult to determine by X-ray or NMR analyses. The lack of structural information has impeded understanding of the details of the biochemistry of these enzymes.

The matrix metalloproteinase family of endopeptidases plays key roles in multiple physiological and pathological processes [3]. The unregulated activities of these enzymes result in pathological conditions such as cancer growth, tumor metastasis and angiogenesis, arthritis, connective tissue diseases, inflammation, cardiovascular and autoimmune diseases [4]. MMPs are expressed by a wide variety of eukaryotic and prokaryotic cells [5] and to date, 26 members of this family of enzymes have been identified. Traditionally, MMPs are grouped into four major classes: collagenases, stromelysins, gelatinases, and membrane-type MMPs, based on their substrate specificity and domain structures [5,6]. Pro-MMP-2 and pro-MMP-9 belong to gelatinases and are unique among MMPs for preferentially cleaving denatured collagen (gelatin) efficiently. Both enzymes have been implicated in the process of tumor metastasis and angiogenesis [7].

The gelatinases are multi-domain enzymes, which, like other MMPs, are comprised of a propeptide domain, a catalytic domain, a hinge region, and a hemopexin-like C-terminal domain (Fig. 1). In addition, gelatinases contain a unique domain, known as the gelatin-binding domain, comprised of three head to tail repeats homologous to the type II repeat found in fibronectin, which is inserted into the catalytic domain. Pro-MMP-9 also contains a unique region of 54-amino acid long, rich in proline residues, that is similar to the α2(V) chain of collagen V, known as collagen-like hinge region. Another interesting feature of pro-MMP-9 is its high degree of glycosylation with both N- and O-glycosylation sites. The sequence of the protein indicates that there are three potential N-glycosylation sites and several O-glycosylation sites (the latter mainly in the hinge region).

Rudd and coworkers [8,9] recently published the *O*-glycan analysis of the human neutrophil pro-MMP-9 using mass spectrometry. Their analyses did not reveal any

Abbreviations used: MMP, matrix metalloproteinase; DTT, dithiothreitol; TIMP, tissue inhibitor of matrix metalloproteinase.

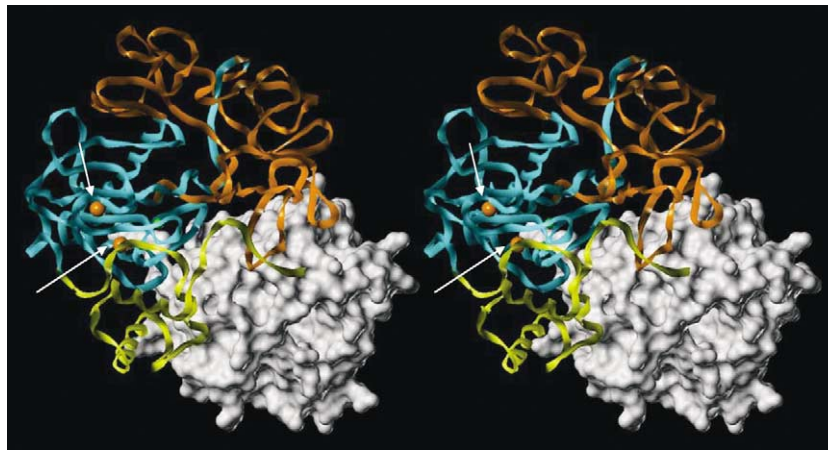


Fig. 1. Domain structures in pro-MMP-2 shown for the X-ray crystal structure (PDB code: 1CK7). The propeptide domain is in yellow, the catalytic domain is in cyan and the gelatin-binding domain is shown in orange. The hemopexin-like domain is rendered as a white surface, the zinc ions as orange spheres and the calcium ion in the catalytic domain as a green sphere. The arrow at 8 o'clock points to the catalytic zinc ion and that at 9 o'clock shows the structural zinc ion.

of the glycosylation sites, but identified the types of sugars that exist at the O-glycosylation sites. It was assumed that all the potential glycosylation sites would be glycosylated. Here, we present for the first time the complete mass spectral and computational analyses of N-glycosylated human pro-MMP-9. Molecular dynamics simulations of the model of glycosylated pro-MMP-9 revealed that the oligosaccharide chains may not have any direct interactions with the protein or with one another and are extended into the milieu.

2. Materials and methods

Human pro-MMP-9 free of TIMP-1 was expressed and purified to homogeneity as described earlier [10,11]. Recombinant *endo*glycosidase [peptide *N*-glycosidase F from *Chryseobacterium meningosepticum* (PNGase F, glycerol-free)] was purchased from Genzyme (Cambridge, MA). TPCK (*N*-tosyl-L-phenylalanine chloromethyl ketone)-treated trypsin and TLCK (*N*-tosyl-L-lysine chloromethyl ketone)-treated α-chymotrypsin from bovine pancreas, iodoacetamide, DL-dithiothreitol, and trifluoroacetic acid (protein sequence grade) were purchased from Sigma (St. Louis, MO), *Exo*glycosidases of *Clostridium perfringens* α2-3,6 sialidase C and *Streptococcus pneumoniae* α2-3 sialidase S were purchased from Prozyme (San Leandro, CA). *Exo*glycosidases of *S. pneumoniae* β1-4 galactosidase, *S. pneumoniae* β1-2 *N*-acetylhexosaminidase, and bovine kidney α1-6 fucosidase were purchased from Oxford GlycoScience (Bedford, MA). Jack bean α-mannosidase was obtained from Roche Molecular Biochemicals (Indianapolis, IN). A 150 × 1.0 mm Haisil 300 C18 5 μm LC column was purchased from Higgins Analytical (Mountain View, CA).

Millipore Ultrafree-MS centrifugal filter units [30,000 nominal molecular weight limit (NMWL), low-binding, regenerated cellulose] were obtained from Millipore Corporation (Bedford, MA).

2.1. Denaturation, reduction, and S-alkylation of pro-MMP-9

Approximately 60 μg of pro-MMP-9 (~600 pmol) were dissolved in 85 μl buffer solution at pH 8.0 containing 0.3 M Tris–HCl, 6 M guanidine hydrochloride, and 1 mM EDTA. A freshly prepared solution (15 μl of 100 mM) of DTT was added to the pro-MMP-9 reduction solution. The mixture was then incubated at 37 °C for 1 h. After the denaturation and reduction processes were completed, 40 μl of freshly prepared iodoacetamide (190 mM) in ammonium bicarbonate solution (pH 8.0, 50 mM) was added to the pro-MMP-9 solution. The mixing solution was then incubated in the dark for 30 min.

2.2. Trypsin digestion of pro-MMP-9

The reduced and *S*-alkylated form of pro-MMP-9 was desalted and concentrated on a 30,000 NMWL Millipore Ultrafree-MC centrifugal filter unit. Deionized water (~3 ml) was used to remove excessive amounts of salts and reagents. Then, the tube was exchanged with 1 ml of ammonium bicarbonate buffer (50 mM, pH 8.0). The final reduced and *S*-alkylated form of pro-MMP-9 was dissolved in 50 μl of ammonium bicarbonate buffer (50 mM, pH 8.0). TPCK-treated trypsin (1.2 μl of 1 μg/1 μl) in ammonium bicarbonate buffer (50 mM at pH 8.0) was added to the pro-MMP-9 solution to give a final trypsin to pro-MMP-9 ratio of 1:50 (w/w). Digestion of pro-MMP-9 by trypsin was allowed to proceed overnight at 37 °C.

2.3. α-Chymotrypsin digestion of the purified glycopeptide T28

α-Chymotrypsin (0.2 μl) solution (0.1 μg/μl) in ammonium bicarbonate buffer (50 mM at pH 8.0) was added to fraction T28 (1.5 μg) in 0.5 μl buffer and incubated for 4 h at 37 °C.

2.4. Microbore RPLC separation of mixtures of trypsin digestion of pro-MMP-9

Trypsin-digested pro-MMP-9 was diluted to 150 μl with 0.1% TFA in water followed by centrifugation. Glycopeptides and peptides were separated on a Haisil 300, 5 μm (C18, 100 × 1 mm) column. Peptide and glycopeptide mixtures were separated by RPLC on a Hewlett-Packard 1100 Series LC system (Palo Alto, CA). The LC system was operated at a flow rate of 75 μl/min with a linear gradient from solvent A (0.1% TFA in water) to 70% of solvent B (0.085% TFA in 80% acetonitrile) over 110 min and then to 90% of solvent B in 10 min. Peptide and glycopeptide peaks were detected at 214 nm and fractions were collected manually. All fractions were then concentrated to dryness at room temperature using a Speed-Vac centrifuge (Savant Instruments, Farmingdale, NY). Deionized water (20 μl) was added to redissolve samples of each fraction and they were then ready for MALDI-MS analysis.

2.5. MALDI-MS analysis of trypsin-digested fractions of pro-MMP-9

The mass spectrometer used in this work was the MALDI-MS HP G2025A (Hewlett-Packard, Palo Alto, CA). The instrument was operated with an accelerating voltage of 28 kV, an extractor voltage of 7 kV, and under vacuum of $\sim 1 \times 10^{-6}$ Torr. Samples were desorbed and ionized from the probe tip with a nitrogen laser ($\lambda = 337$ nm) having a pulse of 3 ns and delivering ~ 10 μJ energy/pulse. α-Cyano-4-hydroxycinnamic acid was used as the matrix and a saturated solution of the matrix was prepared in 0.1% TFA of 80% acetonitrile solution. The instrument was calibrated externally with a mixture of known peptides before each set of experiments.

2.6. Endo- and exoglycosidase digestion conditions for the purified glycopeptides

All digestions were performed in 0.65 ml Eppendorf tubes. The endoglycosidase PNGase F was diluted 500-fold in ammonium bicarbonate buffer (pH 8.0, 30 mM) to make a solution of 10 mU/μl. Glycopeptides (20–50 pmol) in water were mixed with an equal volume of PNGase F solution and incubated for 4 h. *Exo*glycosidases were diluted 50–100-fold at different pH levels in ammonium acetate buffer solutions as required by the specific activities of each enzyme [12].

A series of stepwise digestions was performed with different *exo*glycosidases to sequence the N-linked glycan on the identified glycopeptides. In the first step of the experiment, the N-linked glycopeptides were digested with α2-3,6 sialidase C from *C. perfringens* and α2-3 sialidase S from *S. pneumoniae*. About 2 μl of aqueous solution of the glycopeptides (~50 pmol) was mixed with 2 μl (0.1 mU) sialidase C and S, respectively, in buffer solutions of 30 mM NH_4OAc at pH 5.0. Digestions were performed at 37 °C for 4 h. After this, a 0.5 μl portion of the digested products was analyzed by MALDI-MS. In the second step of the experiment, 1 μl (0.4 mU) of β1-4 galactosidase from *S. pneumoniae* in buffer solution of 30 mM NH_4OAc at pH 5.0 was added to the rest of the solution from the first step of the digestion, and the mixture was incubated for 4 h at 37 °C. The digested product (0.5 μl) was analyzed by MALDI-MS. In the third step of the experiment, 1 μl (0.5 mU) β1-2 *N*-acetylhexosaminidase from *S. pneumoniae* was added to the rest of the solution remaining after the second digestion and incubated for 4 h at 37 °C. The digested product (0.5 μl) was analyzed by MALDI-MS. In the fourth step of the experiment, the solution was vacuum-dried. The materials were solubilized in 2 μl of buffer solution (30 mM NH_4OAc in 2 mM $ZnCl_2$ at pH 4.5) and 0.5 μl of α-mannosidase from Jack bean was added to the sample and incubated for 4 h at 37 °C. After MALDI-MS analysis, the sample was vacuum-dried again and 1 μl of 30 mM NH_4OAc at pH 5.0 was added. In the final step of the *exo*glycosidase digestion experiment, 1 μl (1 mU) α1-6 fucosidase from bovine kidney in 30 mM NH_4OAc at pH 5.0 was added to the solution and incubated overnight at 37 °C. Ammonium acetate buffer solutions were used for all the *endo*- and *exo*glycosidase digestions of glycopeptides in order to minimize any interference from salts in the MALDI-MS analysis.

2.7. Model building and molecular dynamics simulations

The three-dimensional model of pro-MMP-9 that was used in this paper was constructed as described below. The three-dimensional model of the MMP-9 catalytic domain lacking the gelatin-binding domain was published earlier [5]. The gelatin-binding domain and propeptide domain of pro-MMP-9 were modeled based on the three-dimensional X-ray structure of pro-MMP-2 [13] [PDB code: 1CK7], using the COMPOSER module in Sybyl [14]. The three-dimensional model of the gelatin-binding domain of pro-MMP-2 was inserted between the residues F^{222} and V^{398} in the catalytic domain of pro-MMP-9 to obtain the complete model of the pro-MMP-9 from residues F^{107} to G^{444}, excluding the hemopexin-like domain. The propeptide region from Q^{25} to R^{106} was added to the above model in such a way that C^{99} of the propeptide coordinates with the catalytic zinc, analogous to the zinc coordination present in pro-MMP-2. In the homology modeling, the sequence A^{19}–E^{450} of pro-MMP-9 when compared to the sequence of pro-MMP-2 as found in the X-ray crystal structure of MMP-2 (PDB code: 1CK7), showed an identity score of 56.6% at a significance level of 58.5% during the composer homology modeling run. This assembly provided the model of the three-dimensional structure of pro-MMP-9 from the residues Q^{25} to G^{444}. The terminal residues in the pro-MMP-9 sequence were eliminated because there was no homologous sequence in the X-ray structure of pro-MMP-2. This model was energy-minimized using the Amber 5.0 software suite [15] and was utilized in construction of the N-glycosylation sites.

The N-glycosylation sites on the catalytic domain of pro-MMP-9 were determined to be at positions N^{38} and N^{120} by mass spectrometry (*vide infra*). The oligosaccharides that were present at these two positions were identical, as determined by mass spectrometry, and were introduced into the computational model. This *N*-glycosylated pro-MMP-9 model was subjected to energy-minimization to relax any potentially undue geometrical constraints. The model was further refined by molecular dynamics simulations. In the starting model, the orientations of the oligosaccharide chains were set in an extended confirmation and the complete model was immersed in a cubic box of water that spanned at least 10 Å from the surface of the glycoprotein. This entire system, including the solvent molecules, contained approximately 70,000 atoms. Molecular dynamics simulations were carried out on this system using particle mesh Ewald option as implemented in Amber 5.0, with cut-off for non-bonded interactions at 12 Å at constant temperature 300 K in the time steps of 2 fs. The system was first equilibrated at constant volume for 10 ps and then at constant pressure of 1 atm for an additional 40 ps. Then, snapshots were collected for every picosecond until the end of the simulations at 250 ps.

3. Results and discussion

3.1. Identification of glycopeptides from trypsin-digested pro-MMP-9

There are three potential N-linked sites in pro-MMP-9, at N^{38}, N^{120}, and N^{127}, as predicted from the protein sequence. Position N^{38} is in the pro-domain of

pro-MMP-9, while N^{120} and N^{127} are both in the catalytic domain. Trypsin digestion of pro-MMP-9 generated two potential glycopeptides, TN^{38} LTDR and WHHHN^{120}ITYWIQN^{127}YSEDLPR. To determine if any of these sites were occupied, pro-MMP-9 was digested with trypsin, and the resulting mixture of peptides and glycopeptides was separated by microbore RPLC.

MALDI-MS analysis of fraction T9 (Fig. 2A) showed three major peaks at 2488.1, 2779.3, and 3070.6 Da. The mass difference between the adjacent peaks was 291 Da. The fact that 291 Da is the incremental mass of a sialic acid residue indicated that this fraction contained a glycopeptide having heterogeneity with its sialic acid residues. To confirm the presence of an N-linked glycopeptide, fraction T9 was digested with PNGase F to release the N-linked carbohydrate chains in the sample, and the resulting products were analyzed by MALDI-MS. After this digestion, the three major peaks in fraction T9 disappeared and a new peak appeared at 719.9 Da. The mass of the new peak matched the tryptic peptide sequence of TD^{38} LTDR that contained the first N-linked peptide consensus sequence in which N^{38} was converted to D^{38} in the oxidative process. The mass differences between the three major peaks in fraction T9 and the peak after PNGase F digestion (1768.2, 2059.4, and 2350.7 Da) must be due to the loss of the N-linked oligosaccharides that were

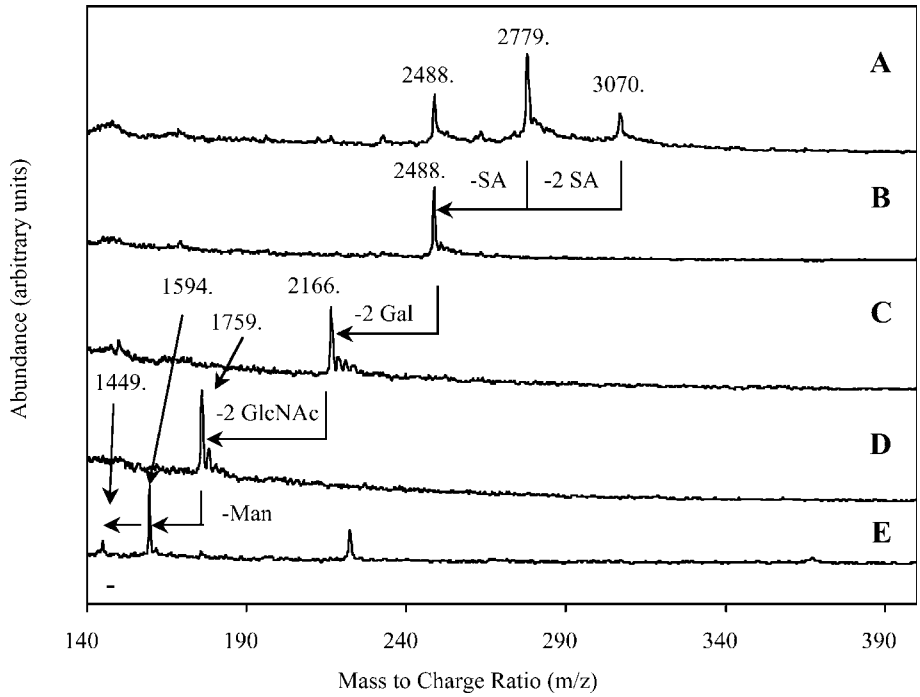


Fig. 2. MALDI-MS analysis of fraction T9 (A), after the digestion of fraction T9 with sialidase S (B), followed by β1-4 galactosidase (C) and then by β1-2-*N*-acetylglucosaminidase digestion (D). MALDI-MS analysis of fraction T9 after treatment with α1-6-fucosidase and α-mannosidase (E).

attached to N^{38} in the peptide TN^{38} LTDR. The smallest of these mass differences, 1768.2 Da, corresponds to the composition of a fucosylated, biantennary complex-type oligosaccharide structure (Gal2GlcNAc2Man3GlcNAc2Fuc). The other two oligosaccharides (2059.4 and 2350.7 Da) appear to correspond to the same carbohydrate structure with one and two sialic acid residues attached.

The next step was to determine the structures of the N-linked oligosaccharides that are attached to the N^{38} residue. This was accomplished with a series of sequential exoglycosidase digestions/MALDI-MS on the purified glycopeptide T9. Each of these exoglycosidases released only one type of monosaccharide residue, with specific linkage and configuration, from the nonreducing termini of the oligosaccharides on the glycopeptide. Stepwise experiments were performed so that in each round of digestion only one type of monosaccharide was removed. MALDI-MS analysis was performed on the products of each digestion so that the structural information from the digestion was obtained before the next round.

Fraction T9 was digested with α2-3,6 sialidase C from *C. perfringens* with cleavage rates of sialic acid residues of α2-3 > α2-6 in the first round experiment. MALDI-MS analysis of the resulting mixture showed that sialidase C caused the two peaks at the higher masses (2779 and 3070 Da) to disappear, while the peak at 2488 Da was not affected. This result demonstrates that these two higher mass peaks in fraction T9 contain one and two sialic acid residues, respectively, at their nonreducing termini, while the peak at 2488 Da does not have a nonreducing terminal sialic acid residue. To identify the specific sialic acid linkage, fraction T9 was digested with sialidase S from *S. pneumoniae*, which only releases α2-3-linked sialic acid residues from the nonreducing termini. MALDI-MS analysis of the digestion (Fig. 2B) showed that both sialic acid residues had been released. Therefore, all the outermost sialic acid residues in fraction T9 are attached to the oligosaccharide chains only through α2-3 linkages.

The product after sialidase digestion was digested with a second enzyme, β1-4 galactosidase. MALDI-MS analysis (Fig. 2C) after this step demonstrated that the oligosaccharide had lost two β1-4-linked galactose residues from the pervious product, as indicated by the mass loss of 324 Da (from 2488.6 to 2166.8 Da). This product was then digested with a third enzyme, β1-2-*N*-acetylglucosaminidase, which released two β1-2 GlcNAc residues from the product of the second digestion, as indicated by the mass loss of 406 Da (2166.8–1759.6 Da) shown in the MALDI-MS spectrum (Fig. 2D). In the last procedure, the glycopeptide was digested with a mixture of α1-6-fucosidase from bovine kidney and α-mannosidase from Jack bean. Within the core structure of the *N*-linked oligosaccharide, Manα(1,3)[Manα(1,6)]Manβ(1,4)-GlcNAcβ(1,4)-GlcNAcβ, this α-mannosidase removed the α1-3-linked mannose residue from the core β-linked mannose residue, while leaving the α1-6-linked mannose residue intact. MALDI-MS analysis of the final digestion products (Fig. 2E) showed two new peaks at 1594.6 and 1449.4 Da that corresponded to the sequential mass losses of an α1-3 mannose residue (162 Da) and a β1-6-fucose residue (146 Da), respectively, from their precursor at 1759.6 Da. Through these experiments, the N-linked oligosaccharides at the N^{38} residue were unambiguously determined to have α1-6 fucosylated bianteannary complex structures with zero, one, and two α2-3 sialic acid residues on the nonreducing end of the structure,

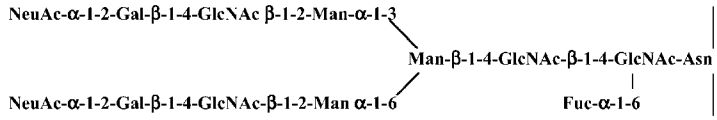
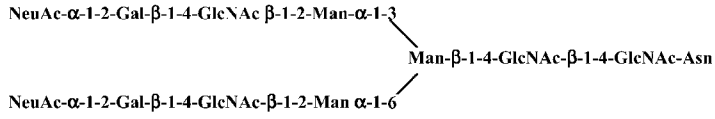
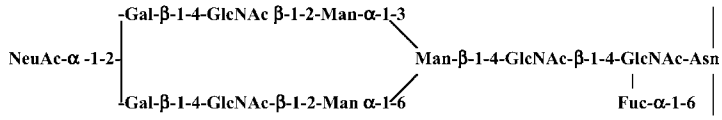
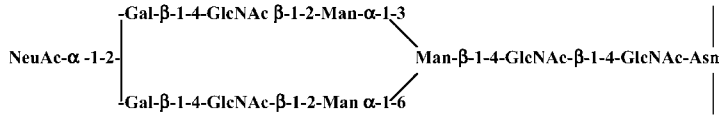
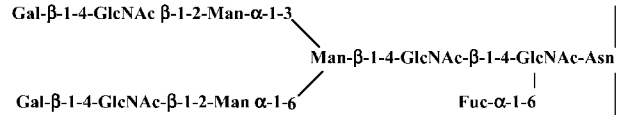
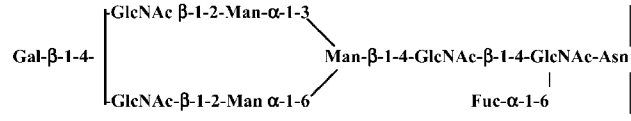
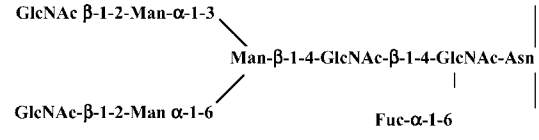
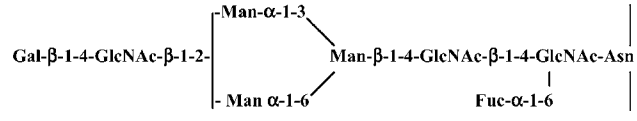
as shown in Table 1. There were also some minor glycosylated forms at the N-linked glycosylation site of N^{38}, as listed in Table 1.

MALDI-MS analysis of fraction T28 from the RPLC showed three peaks at 4280.2, 4571.8, and 4863.1 Da with a mass difference of 291 Da between the adjacent peaks (Fig. 3A). As in the discussion of the characterization of fraction T9, fraction T28 was probably a set of glycopeptides. MALDI-MS analysis of the digested products of fraction T28 by PNGase F treatment showed that in addition to the disappearance of these glycopeptide peaks, the new peak at 2878.8 Da matched the sequence of the tryptic peptide mass that contained the two potential N-linked sites in pro-MMP-9 (WHHH**N**^{120}ITYWIQ**N**^{127}**Y**SEDLPR). The mass losses during PNGase F digestion of T28 resulted from the N-linked oligosaccharides with masses of 1776.9, 2059.5, and 2350.8 Da, respectively. The smallest mass oligosaccharide corresponded to the mass of the fucosylated and biantennary complex structure, while the other two peaks with masses of 291 and 582 Da higher, corresponded to the presence of one and two sialic acid residues attached to the fucosylated and biantennary complex structure. A series of *exo*glycosidase digestions, identical to that discussed above for fraction T9, were performed on fraction T28. α2-3 Sialidase S from *S. pneumonia* removed one and two α2-3 sialic acid residues (291 and 582 Da) from 4571.8 and 4863.1 Da, respectively, to give a peak at 4279.7 Da (Fig. 3B). β1-4 Galactosidase from *S. pneumoniae* removed two adjacent β1-4 galactose residues (324 Da) to give the peak at 3958.1 Da (Fig. 3C). β1-2-*N*-acetylglucosaminidase from *S. pneumoniae* removed two more β1-2 GlcNAc residues (406 Da) to give the peak at 3552.8 Da (Fig. 3D). It was clear from these experiments that there was a set of N-linked oligosaccharides that had fucosylated bianteannary complex structures with heterogeneity of zero, one, and two sialic acid residues attached to this peptide.

These experiments demonstrated that only one of the two possible glycosylation sites on this peptide (N^{120} or N^{127}) was glycosylated, but it did not indicate which. Hence, the next goal was to distinguish the site that was glycosylated in the peptide in fraction T28. Since there were two hydrophobic amino acid residues, Thr and Tyr, in the peptide backbone between the two potential N-linked sites in the peptide of fraction T28, α-chymotrypsin was chosen to further digest fraction T28. In order to simplify the spectrum of the MALDI-MS analysis of the glycopeptides, fraction T28 was first digested with α2-3 sialidase S to remove the heterogeneity of the oligosaccharides in fraction T28, as shown in Fig. 4B. Subsequently, the desialylated fraction T28 was digested with α-chymotrypsin and the resulting products were analyzed by MALDI-MS (Fig. 4C). A new peak at 2878.8 Da was generated by α-chymotrypsin digestion, which corresponded to the sequence WHHH**N**^{120}**ITY** plus the oligosaccharide with the fucosylated bianteannary complex structure that must be attached to the N^{120} residue. In addition, the mass loss from the glycopeptide (1402.5 Da) matches the mass of the amino acid sequence of WIQ**N**^{127}**Y**SEDLPR that contains the Asn^{127} residue. This clearly demonstrates that N^{120} is glycosylated and that N^{127} is unmodified by any carbohydrate.

In order to understand the structural implications of the N-glycosylation, we built a computational model of pro-MMP-9. The residues N^{38} and N^{120} were glycosylated in silico and a molecular dynamics simulation was carried out for 250 ps. The average structure from the dynamics simulations was calculated by first averaging the

Table 1
Oligosaccharide profiles on the first N-linked glycopeptide of pro-MMP-9

Mass	The structures of the carbohydrate chains on Asn-38
3070.6 Major Peak	
2924.2 Minor peak	
2779.3 Major Peak	
2633.7 Minor peak	
2488.1 Major Peak	
2326.7 Minor peak	
2164.5 Minor peak	
2123.2 Minor peak	

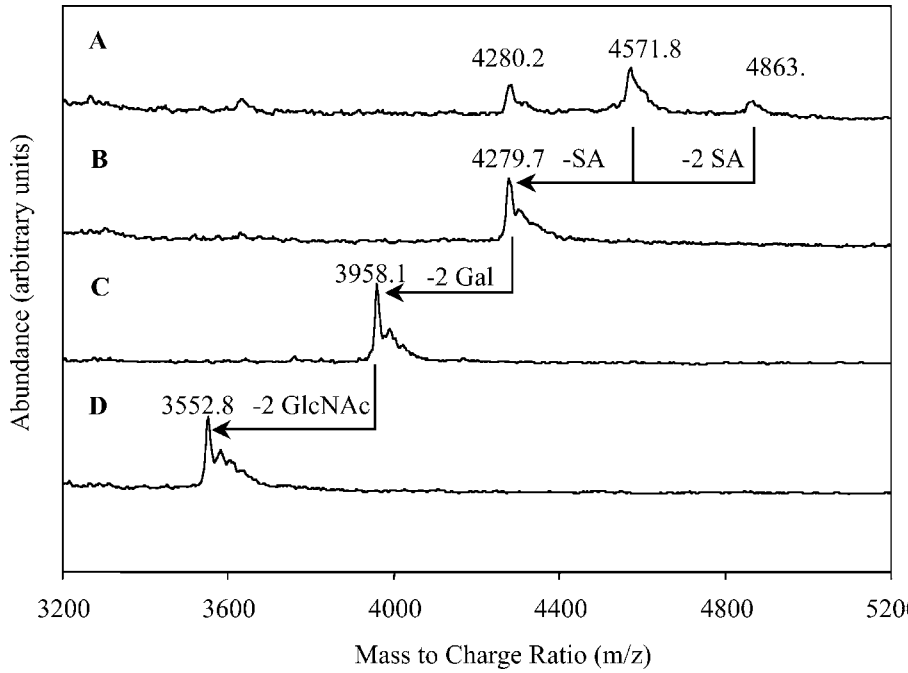


Fig. 3. MALDI-MS analysis of fraction T28 (A), after subsequent treatments with α2-3 sialidase S (B), β1-4 galactosidase (C) and β1-2-*N*-acetylglucosaminidase (D).

coordinates of all the atoms in the snapshots from 50 to 250 ps, and then subjecting the model to energy-minimization. This computational model shows that the oligossaccharides at positions N^{38} and N^{120} extend into the solution from the globular structure of pro-MMP-9. During these dynamics simulations, we were interested to see if the oligosaccharide "antennas" interact with the amino acid residues in the propeptide or the catalytic domains of pro-MMP-9 (Fig. 5). However, during the 250 ps of simulations, we did not see any such interactions between the oligosaccharides and the peptide portions of pro-MMP-9. Whereas it is conceivable that there might be such interactions in solution, it would appear that for the low-energy conformations such structural possibilities are relatively less populated, if at all present. Upon activation of pro-MMP-9 [11], the propeptide region is removed, hence the oligosaccharides at N^{38} will be absent in the active form of pro-MMP-9. The active pro-MMP-9 will have the oligosaccharides at position N^{120} as the only N-glycosylation site.

Determination of the N-linked glycosylation sites and structural characterization of the N-linked oligosaccharides of the recombinant human matrix metalloproteinase-9 (pro-MMP-9) was carried out by a combination of specific *endo*proteinase digestions, microbore RPLC, specific *endo*- and *exo*glycosidase digestions, and MALDI-MS analysis. The first N-linked site (N^{38}), which was in the pro-domain of pro-MMP-9, was determined to contain fucosylated and biantennary complex-type oligosaccharides with heterogeneity of zero, one, and two α2-3-linked sialic acid

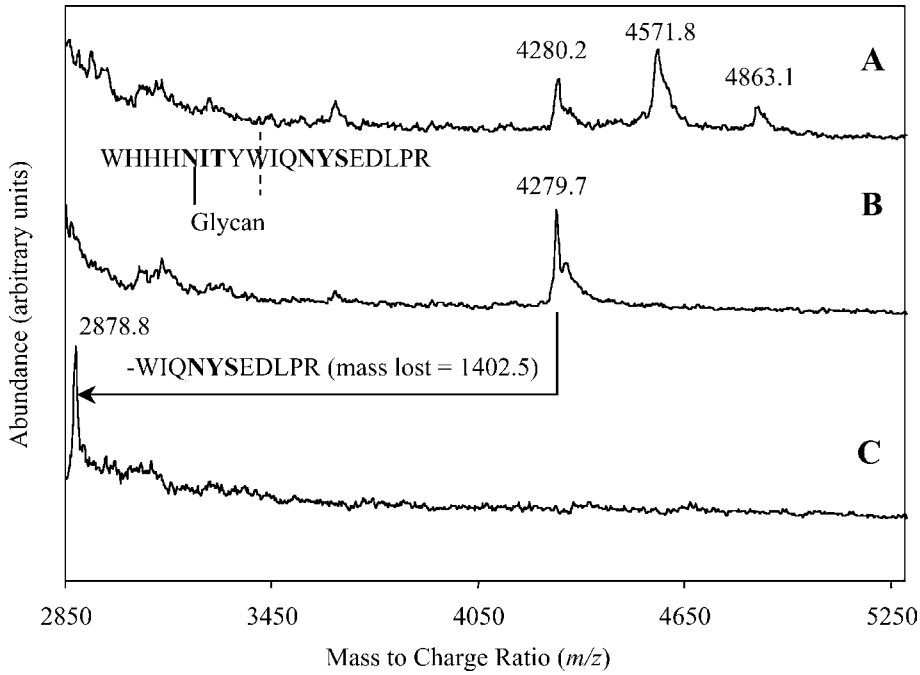


Fig. 4. MALDI-MS analyses of α2-3 sialidase S treatment of fraction T28 (B) and of subsequent α-chymotrypsin-digested sample (C).

residues. Through sequential digestions by trypsin and α-chymotrypsin, the second N-linked site (N^{120}), which was in the catalytic domain in pro-MMP-9, was determined to contain the identical carbohydrate structure as found in the first N-linked site (N^{38}). The third potential N-linked site (N^{127}), also in the catalytic domain in pro-MMP-9, was found not occupied by carbohydrates. This observation is further supported by our computational model of pro-MMP-9, where N^{127} is in close proximity to the gelatin-binding domain (Fig. 5B). The N^{127} is sheltered by the protein and hence not available for posttranslational modification. This is in sharp contrast to the situation of N^{38} and N^{120}, which are both exposed on the surface of the protein and do not pose any steric problems for glycosylation.

Sang and coworkers [16] recently reported that MMP-26 (human endometase) might contain three potential N-glycosylation sites, one in the propeptide at N^{64} and the other two in the catalytic domain at N^{133} and N^{22}. There is no experimental data yet to confirm this prediction. We have generated a preliminary model of the MMP-26 to investigate the structural features pertinent to glycosylation. Our model shows that N^{64} and N^{133} are predisposed spatially for potential glycosylation. However, the third site at N^{221} appears to be sterically encumbered by the propeptide domain, thus it is likely to be precluded as a potential glycosylation site. This observation awaits experimental demonstration.

It is intriguing as to why pro-MMP-9 is glycosylated, but not its close cousin pro-MMP-2. In general, proteins are glycosylated in a defined fashion for various

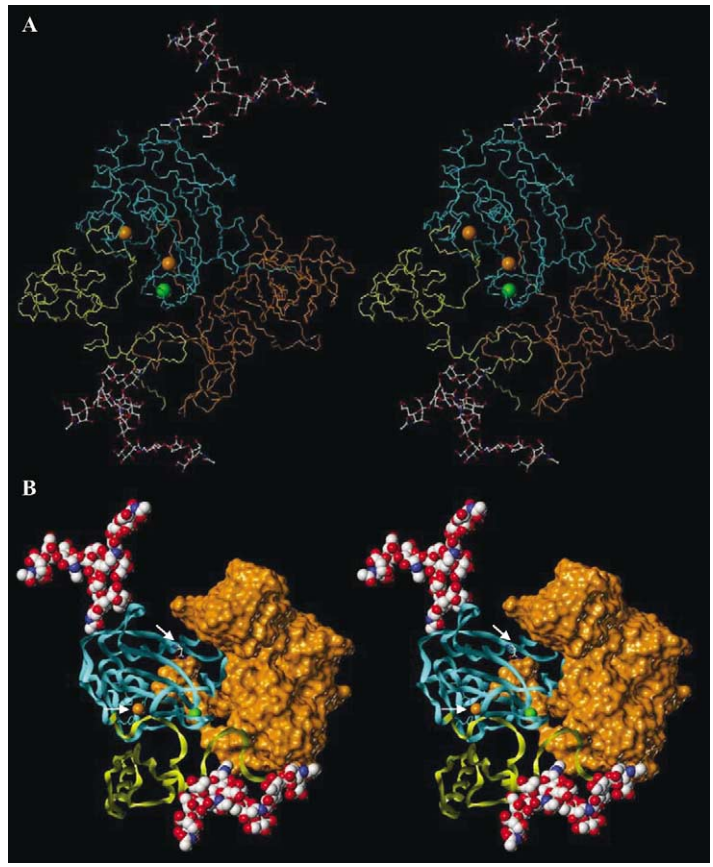


Fig. 5. (A) The time-averaged computational model of pro-MMP-9 derived from the snapshots collected from 51 ps to 250 ps of the molecular dynamics simulations. The propeptide domain (yellow), the catalytic domain (cyan) and the gelatin-binding domain (orange) are shown as backbone trace. The orange spheres represent the zinc ions and the green sphere represents the calcium ion. The oligosaccharide chains at positions N^{38} and N^{120} (at 6 and 12 o'clock positions, respectively) are shown in ball-and-stick representation and are color-coded by atom types (carbon:white, oxygen:red, nitrogen: blue). (B) A different perspective of the time-averaged structure of pro-MMP-9 shown in panel A. The oligosaccharide chains at N^{38} and N^{120} (shown in space-fill representation and color-coded by atom types) are at 6 and 11 o'clock positions, respectively, and the gelatin-binding domain is shown as an orange surface. The position N^{127} in the catalytic domain is shown by a white arrow at 12 o'clock position and the arrow at 9 o'clock position points to the catalytic zinc ion in the active site of pro-MMP-9.

mechanistic reasons [17]. These include promoting proper protein folding, recognition of improperly folded proteins and their degradation, sorting events and perhaps protection from proteases after secretion. In the case of pro-MMP-9, protein folding does not seem to be the likely reason because its close cohort MMP-2 (pro-MMP-2) is not glycosylated, yet serves as a functional protein. It is worth noting that the site of N-glycosylation in the catalytic domain (N^{120}) in the active form of pro-MMP-9 is

away from the enzyme active site and N-glycosylation at that site would not directly participate in substrate recognition and the catalytic process. The function for glycosylation of pro-MMP-9 is yet to be elucidated, but it is clear that as the only gelatinase that is glycosylated, glycosylation is intimately linked to the function of this enzyme. Although a direct comparison between fully glycosylated and unglycosylated pro-MMP-9 has yet to be made, studies using bacterially expressed pro-MMP-9, which is unglycosylated, demonstrated maintenance of some catalytic activity after organomercurial- and trypsin-mediated activation [18]. Thus, glycosylation appears not to be critical for activation or catalysis, consistent with our findings. Interestingly, pro-MMP-9, in contrast to pro-MMP-2, is not readily activated in cellular systems and purified pro-MMP-9 is resistant to autocatalysis activation in solution, likely due to a stabilizing effect of glycosylation. The role of the pro-MMP-9 glycosylation on the interaction with TIMP-1, which is also glycosylated, and in binding of the enzyme to the collagen matrix and the cell surface remain to be explored.

Acknowledgments

This research was supported by grants from the US National Institutes of Health (P41RR05351 and CA-82298) and the National Science Foundation (9626835).

References

[1] P.M. Rudd, T. Elliott, P. Cresswell, I.A. Wilson, R.A. Dwek, Science 291 (2001) 2370–2376;
L. Wells, K. Vosseller, G.W. Hart, Science 291 (2001) 2376–2378.
[2] C.R. Bertozzi, L.L. Kiessling, Science 291 (2001) 2357–2364.
[3] M.-A. Forget, R.R. Desrosier, R. Béliveau, Can. J. Physiol. Pharmacol. 77 (1999) 465–480.
[4] A.R. Nelson, B. Fingleton, M.L. Rothenberg, L.M. Matrisian, J. Clin. Oncol 18 (1999) 1135–1149;
L. Ravanti, V.M. Kahari, Int. J. Mol. Med. 6 (2000) 391–407.
[5] I. Massova, L.P. Kotra, R. Fridman, S. Mobashery, FASEB J. 12 (1998) 1075–1095.
[6] H. Nagase, J.F. Woessner Jr., J. Biol. Chem. 274 (1999) 21491–21494.
[7] B. Bodey, B. Bodey, S.E. Siegel, H.E. Kaiser, In Vivo 14 (2000) 659–666;
S. Ye, Matrix Biol. 19 (2000) 623–629;
B.P. Himelstein, R. Canete-Soler, E.J. Bernhard, D.W. Dilks, R.J. Muschel, Invasion Metastasis 14 (1994) 246–258;
G. Sehgal, J. Hua, E.J. Bernhard, I. Sehgal, T.C. Thompson, R.J. Mischel, Am. J. Pathol. 152 (1998) 591–596.
[8] T.S. Mattu, L. Royle, J. Langridge, M.R. Wormald, P.E. Van den Steen, J. Van Damme, G. Opdenakker, D.J. Harvey, R.A. Dwek, P.M. Rudd, Biochemistry 39 (2000) 15695–15704.
[9] P.M. Rudd, T.S. Mattu, S. Masure, T. Bratt, P.E. Van den Steen, M.R. Wormald, B. Küster, D.J. Harvey, N. Borregaard, J. Van Damme, R.A. Dwek, G. Opdenakker, Biochemistry 38 (1999) 13937–13950.

[10] R. Fridman, T.R. Fuerst, R.E. Bird, M. Hoyhtya, M. Oelkuct, S. Kraus, D. Komarek, L.A. Liotta, M.L. Berman, W.G. Stetler-Stevenson, J. Biol. Chem. 267 (1992) 15398–15405.
[11] L.P. Kotra, J.B. Cross, Y. Shimura, R. Fridman, H.B. Schlegel, S. Mobashery, J. Am. Chem. Soc. 123 (2001) 3108–3113;
M.W. Olson, M.M. Bernardo, M. Pietila, D.C. Gervasi, M. Toth, L.P. Kotra, I. Massova, S. Mobashery, R. Fridman, J. Biol. Chem. 275 (2000) 2661–2668.
[12] Detailed digestion conditions of each enzyme and their specificities are given in the supplement.
[13] E. Morgunova, A. Tuuttile, U. Bergmann, M. Isupov, Y. Lindqvist, G. Schneider, K. Tryggvason, Science 284 (1999) 1667–1669.
[14] Tripos Inc., St. Louis, MO, USA.
[15] D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, W.S. Ross, C.L. Simmerling, T.A. Darden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley, D.M. Ferguson, R.J. Radmer, G.L. Seibel, U.C. Singh, P.K. Weiner, P.A. Kollman, AMBER 5. University of California, San Francisco, CA, 1997;
D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham III, S. DeBolt, D. Ferguson, G. Seibel, P.A. Kollman, Comp. Phys. Commun. 91 (1995) 1–41.
[16] H.I. Park, J. Ni, F.E. Gerkema, D. Liu, V.E. Belozerov, Q.X.A. Sang, J. Biol. Chem. 275 (2000) 20540–20544.
[17] A. Helenius, M. Aebi, Science 291 (2001) 2364–2369.
[18] C.H. Bu, T. Pourmotabbed, J. Biol. Chem. 270 (1995) 18536–18569.