

Identification of Tyrosine Sulfation in Extracellular Leucine-rich Repeat Proteins Using Mass Spectrometry*

Received for publication, August 6, 2003, and in revised form, October 3, 2003
Published, JBC Papers in Press, October 8, 2003, DOI 10.1074/jbc.M308689200

Patrik Önnerfjord, Terrence F. Heathfield, and Dick Heinegård‡

From the Department of Cell and Molecular Biology, Section for Connective Tissue Biology, Lund University, BMC C12, 221 84 Lund, Sweden

Multiple and variable tyrosine sulfation in extracellular class II leucine-rich repeat proteins/proteoglycans were characterized by mass spectrometry. The sulfogroup on tyrosine is labile and is released from peptides under normal mass spectrometric conditions. Thus, special approaches must be considered in order to identify this modification. By using a combination of mass spectrometry studies operating in negative and positive ion mode, tyrosine sulfation could be identified. In positive mode, the peptides normally appeared non-sulfated, whereas in negative mode a mixture of sulfated and non-sulfated species was observed. A combination of peptides released by different proteinases was used to obtain details on the locations of sulfate groups. Multiple tyrosine sulfates were observed in the N-terminal region of fibromodulin (up to 9 sites), osteoadherin (up to 6 sites), and lumican (2 sites). Osteoadherin contains two additional sulfated tyrosine residues close to its C terminus. We also identified an error in the published sequence of bovine fibromodulin, resulting in the replacement of Thr³⁷ by Tyr³⁷-Gly³⁸, thus increasing its homology with its human counterpart.

Tyrosine sulfation is a post-translational modification found on many secreted and membrane-bound proteins. As much as 1% of all tyrosine residues of the total protein in an organism can be sulfated, making this the most common post-translational modification of this residue (1). A number of proteins have been reported to contain sulfated tyrosines (2). Sulfation is suggested to occur at tyrosines located in close proximity to acidic residues (3). The existence of two different tyrosylprotein sulfotransferases, TPST-1¹ and TPST-2, might explain the diversity of sequences that are sulfated. Each enzyme may have a different substrate specificity and act on a different subset of target proteins. Tyrosine sulfation of chemokine receptor CCR-5 by TPST-1 and TPST-2 follows a discrete pattern and

temporal sequence (4). The N terminus contains the sulfation sites with 0–4 sulfogroups present. TPST-1 null mice (5) appeared healthy but had a ~5% lower average body weight than wild type animals. In addition, although the fertility of (–/–) males and females is normal, they have significantly smaller litters because of fetal death between 8.5 and 15.5 days postcoitum.

Regulation of tyrosine sulfation *in vivo* by enzymatic sulfate removal is still largely unknown. However, human arylsulfatase A and B reside in the lysosome where they are thought to participate in the degradation of tyrosine sulfated proteins (6). In contrast, arylsulfatase E is found in the Golgi compartment (7). Mutations in this gene cause chondrodysplasia punctata, a congenital disorder characterized by abnormalities in cartilage and bone development (7–12).

The biological role of tyrosine sulfation has been unclear, but recent studies have shown its functional importance in leukocyte adhesion (13–16), hormone synthesis (17–19), chemokine receptor signaling (20–26), and hemostasis (27–33).

Mass spectrometry (MS) has previously been used for detection of tyrosine-sulfated peptides in a number of applications. An early report describes a singly sulfated peptide (factor VIII light chain) using matrix-assisted laser/desorption ionization time-of-flight (MALDI-TOF) MS in positive and negative ion modes (34). The sulfated species was more prominent in the negative ion mode. In another publication, a comparison of a sulfated or phosphorylated peptide of cholecystokinin-8 with identical sequences was studied using MALDI-TOF MS, illustrating the difference in lability of the different modifications in negative ion mode using linear or reflector detectors (35). Loss of the phosphate group was difficult to observe in the reflector mode, whereas loss of sulfate was prominent. In linear mode the sulfopeptide was observed intact. In addition, the phosphotyrosine immonium ion of m/z 216.043 can be used to search for tyrosine phosphorylations using a quadrupole-TOF (Q-TOF) instrument, and in the same experiment MS-MS can be used to determine the phosphorylation site (36). This is not the case for tyrosine sulfation sites, because the sulfate group is much more labile than the peptide bonds, which are broken upon a MS-MS experiment. Multiply sulfated peptides (containing 2 and 3 sulfates, respectively) of cionin and cholecystokinin-associated nonapeptides were studied using secondary ion mass spectrometry (37). In these experiments, ladders of up to three sulfates were obtained in negative ion mode. Tyrosine sulfation and phosphorylations have also been reported in the extracellular bone sialoprotein (38). Multiple phosphorylations were observed in the negative ion mode, but in contrast only singly sulfated peptides were detected. The tyrosine sulfate residues in antihemophilic recombinant factor VIII have been reported using triple quadrupole-MS in positive ion mode (39). A neutral loss of 40 Da (m/z = 2) was used to identify sulfated peptides; only ions that lose this specific mass pass through

* Grants for this study were obtained from the Medical Faculty of Lund University, the Swedish Medical Research Council, the IngaBritt & Arne Lundberg Research Foundation (funding for MALDI-TOF MS equipment), the Greta & Johan Kock Foundation, the King Gustaf V 80-year Foundation, Reumatikerförbundet, the Magnus Bergvall Foundation, and The Royal Physiographic Society in Lund. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. E-mail: dick.heinegard@medkem.lu.se.

¹ The abbreviations are: TPST, tyrosylprotein sulfotransferase; MS, mass spectrometry; MS-MS, tandem mass spectrometry; MALDI, matrix-assisted laser/desorption ionization; TOF, time-of-flight; Q, quadrupole; LRR, leucine-rich-repeat; ESI, electrospray ionization; PSD, post-source decay; ISD, in-source decay; sY, sulfated tyrosine; pQ, pyroglutamic acid.

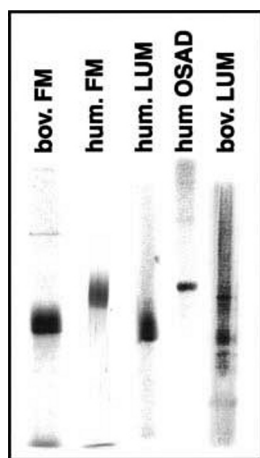


FIG. 1. SDS-PAGE of the purified proteins used in this study: bovine fibromodulin (bov. FM) (8–16% SDS), human fibromodulin (10%), human lumican (hum. LUM) (4–16%), human osteoadherin (hum. OSAD) (4–16%), and bovine lumican (bov. LUM) (4–16%). In-gel digests were performed with human osteoadherin and lumican, and in-solution digests were carried out with the other proteins.

both Q1 and Q3. The doubly charged species was observed to have the most intense signal with up to three sulfates observed within a single peptide.

The sulfation pattern of recombinant human B-domain deleted factor VIII with 6 well documented sulfation sites was also studied using Q-TOF MS (40). Seven peptides with single and multiple +80 modifications were isolated and analyzed for their post-translational modifications. Q-TOF MS-MS in positive mode was used together with in-source dissociation (MS^3). Each sulfated peptide lost all SO_3 before backbone fragmentation occurred. Also, in this case a maximum of 3 tyrosine sulfates were detected within a peptide.

In the present work we have studied a class of extracellular matrix proteins, leucine-rich repeat (LRR) proteins, that contain a large number of potential tyrosine sulfation sites clustered together at either terminal. These proteins are present in cartilage, where they bind other matrix constituents, particularly collagen, and stabilize the fibrillar network that constitutes the framework of the tissue. The study includes the use of different mass spectrometric techniques for identification and characterization of tyrosine sulfates in these proteins.

EXPERIMENTAL PROCEDURES

Proteins were isolated as described previously (41–44). Human osteoadherin was expressed in human kidney cells, EBNA-293 (Invitrogen) by transfection with a pCEP4 vector (Invitrogen) with the transcription cytomegalovirus promoter included.² Purity of the proteins used in this report was determined by SDS-PAGE.

Purified bovine fibromodulin, 400–600 μ g, was dissolved in 200 μ l of Tris-buffered saline, pH 7.4, and 30 μ l of 500 μ g/ml LysC was added followed by incubation for 16 h at 37 °C. The reaction was stopped by adding 2 μ l of 1 mg/ml 1-chloro-3-tosylamido-7-amino-2-heptanone or N^α -p-tosyl-L-lysine chloromethyl ketone (TLCK) and incubation for 30 min at 37 °C. N-Linked oligosaccharides were removed by the addition of 5 units of N-glycosidase F (Roche Applied Science) and incubation for 16 h at 37 °C. Before loading onto a MiniQ SMART column (Amersham Biosciences), 1500 μ l of 10 mM Tris-HCl, pH 7.6, containing 6 M urea (buffer A for anion exchange chromatography) was added. A gradient of 0–600 mM NaCl in the buffer over 20 column volumes was used to elute bound peptides. The column was regenerated with 2 ml of NaOH following each run. Fractions deemed by Western blot and MALDI analysis to contain the N-terminal peptide were pooled and loaded onto a Sephasil C8 mini reversed-phase chromatography column (Amer-

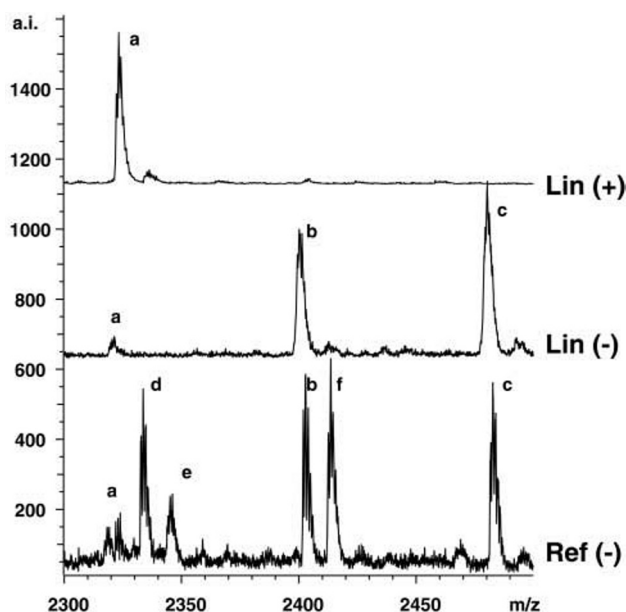


FIG. 2. Behavior of tyrosine-sulfated peptides in MALDI-TOF MS. A sulfopeptide, TYPDYEEYDFPQALYGR, from bovine lumican with two tyrosine sulfation sites was analyzed in positive linear mode (top), negative linear mode (middle), and negative reflector mode (bottom). In linear positive mode, only ISD fragmentation is detected. In negative mode, both singly and doubly sulfated species can be detected. The spectrum in positive reflector mode is not shown because it appeared the same as in linear mode with a single unsulfated peak. In the reflector mode, PSD fragmentation is detected in peaks *e* and *f*, whereas ISD fragmentation is detected in peaks *a* and *b*. The combination of ISD and PSD fragmentation is observed in peak *d*. Annotated peaks: *a*, $[M \pm H - 2SO_3]^{+/-}$, ISD; *b*, $[M - H - SO_3]^-$, ISD; *c*, $[M - H]^-$, *d*, $[M - H - 2SO_3]^-$, ISD + PSD; *e*, $[M - H - 2SO_3]^-$, PSD; *f*, $[M - H - SO_3]^-$, PSD.

sham Biosciences). Bound material was eluted with a gradient of 0–100% acetonitrile in 20 column volumes. The resulting peak was dried and stored at –20 °C. This N-terminal peptide was used for all experiments involving fibromodulin in this work.

Sample Preparation for MALDI-TOF MS—Coomassie-stained bands of human lumican and osteoadherin were excised and washed extensively using 40% acetonitrile in 25 mM NH_4HCO_3 . Gel pieces were then dried in a SpeedVac concentrator before digestion overnight at 37 °C. Digestion was terminated by addition of 2 μ l 2% trifluoroacetic acid, which also extracted the peptides from the gel. After a 1-h extraction at room temperature, peptides were purified from the digestion buffer using miniaturized C18 reversed-phase tips (Ziptips™, Millipore). Purified peptides were eluted directly onto the sample target. The matrix, 2,5-dihydroxybenzoic acid, was used on an Anchorchip™ target (Bruker Daltonics), which confines the sample to a smaller area, thereby increasing the sensitivity (45). Bovine fibromodulin and lumican were digested in solution. The tryptic digest of bovine lumican was further purified by C8 reversed-phase chromatography as described previously for bovine fibromodulin. The fibromodulin samples for electrospray ionization (ESI)-MS were cleaned on Ziptips™ before analysis using off-line nano-electrospray MS.

Enzymatic and Chemical Treatments—Digestions with sequencing grade trypsin (Roche Applied Science), endoproteinase Lys-C (Wako), Asp-N (Roche Applied Science), and chymotrypsin (Roche Applied Science) were generally performed in 50 mM NH_4HCO_3 at pH 7.8. Enzymatic digestions of hirudin fragment (54–65) (Sigma) and lumican peptide (fraction from a reversed-phase separation) using arylsulfatase (Roche Applied Science) and sulfatase (Sigma; type IV (S8504), Type VI (S1629), Type H1 (S9626), and type H2 (S9751)) were performed in phosphate-buffered saline, pH 7.4, using 40–200 milliunits of enzyme with or without a protease inhibitor mixture (Pierce). Digestions were also performed using different buffers including 50 mM NH_4HCO_3 at pH 7.8, 50 mM Tris-HCl at pH 7.4, and 100 mM, 250 mM, 0.5 M and 1 M NH_4CH_3COOH at pH 6.0. For dephosphorylation, peptides were incubated with calf intestinal alkaline phosphatase (1 unit, MBI Fermentas) in 50 mM NH_4HCO_3 for 45 min at 37 °C. The tyrosine-phosphorylated control peptide CRDSNpYISKGS, from platelet-derived growth factor β -receptor, was a generous gift from U. Engström and U. Hellman (Ludwig Institute for Cancer Research, Uppsala, Sweden). The

² V. Tillgren, L. Camper, and D. Heinegård, manuscript in preparation.

FIG. 3. PSD spectrum of a singly sulfated peptide in positive mode obtained after in-gel digestion with chymotrypsin. Switching to the negative mode identifies the tyrosine sulfation site, whereas the PSD spectrum in the positive mode confirms the identity as the N-terminal peptide of bovine fibromodulin, pQsYEEDSHWWF.

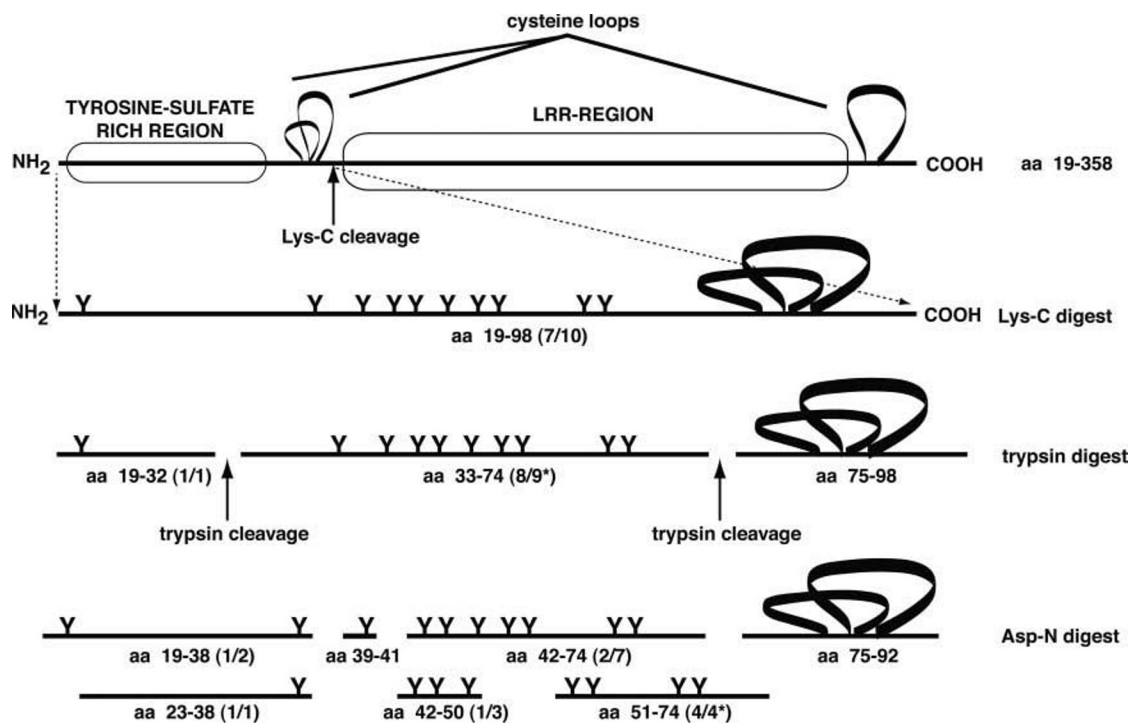
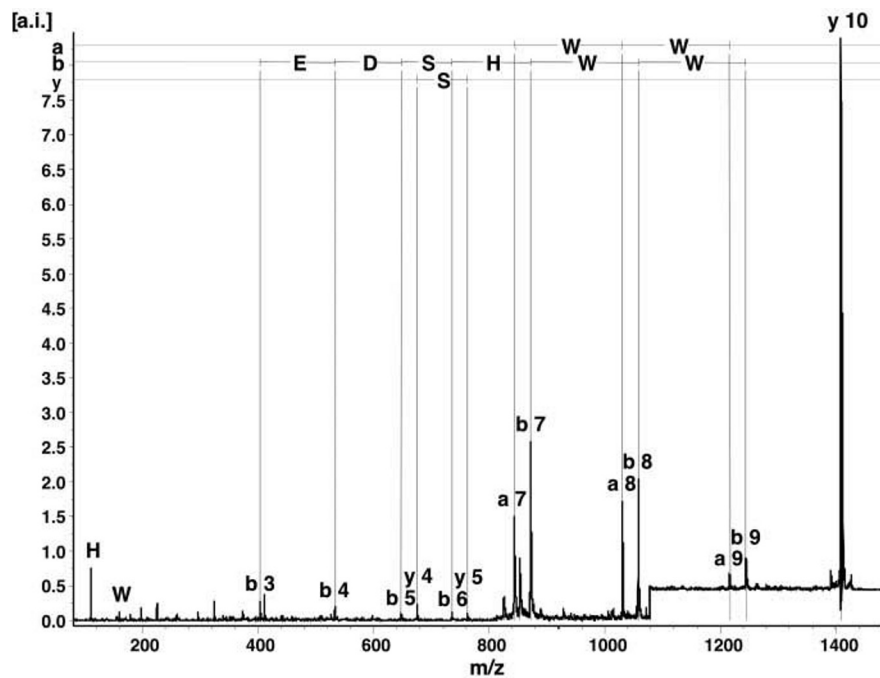


FIG. 4. Schematic representation of bovine fibromodulin containing a tyrosine sulfate-rich region, three disulfide loops, and an LRR region. The intact protein was first digested with endoproteinase Lys-C, and the N-terminal peptide was purified. Trypsin or endoproteinase Asp-N was used to further cleave this peptide. Peptides are labeled with residue numbers, and tyrosines are marked Y. The maximum number of tyrosine sulfates found by MALDI-TOF MS or ESI ion-trap MS (*)/total number of tyrosines present in each peptide is shown in parentheses. Sequences on all detected peptides are listed in Table I. In total 9 of 10 potential tyrosines in the N-terminal peptide were detected with tyrosine sulfate. The residue undetected for this modification is located in the sequence of amino acids (aa) 39–50.

tryptic serine-phosphorylated peptide from α -S1-casein (Sigma), VPQLEIVNPpSAEER, was used as a reference.

Chemical treatments for desulfation of tyrosine sulfates were performed by acid hydrolysis using 1 M HCl at 99 °C for 0.1, or 5 min (46). Samples were then cooled to room temperature and immediately cleaned up on Ziptips™ before application onto the MALDI target.

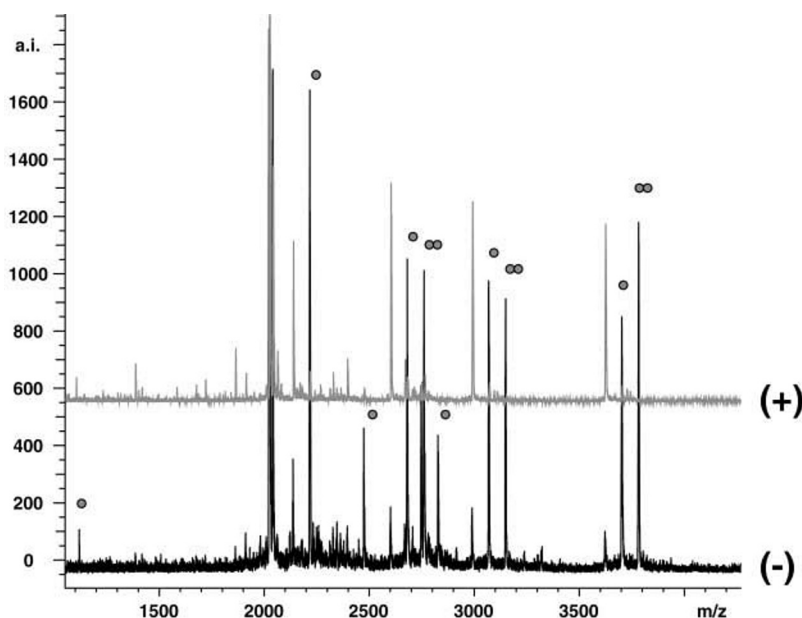
Mass Spectrometry—Mass spectrometric studies were performed using a Bruker Scout 384 Reflex III MALDI-TOF mass spectrometer. The instrument was used in the positive or negative ion mode with delayed extraction and an acceleration voltage of 20–26 kV. Peptide samples were analyzed by collecting 50–150 single-shot spectra for improved

signal-to-noise ratio. Some data were obtained using ESI ion-trap mass spectrometers (Esquire 3000+ (Bruker Daltonics, Bremen, Germany) and LCQ Deca XP+ (ThermoFinnigan, London, UK)).

RESULTS

The purity of the protein preparations used in this work, as tested by SDS-PAGE, is depicted in Fig. 1. Post-translational modifications made the protein bands appear heterogeneous on the gel. These modifications include various oligosaccharide chains as well as tyrosine sulfation.

FIG. 5. MALDI-TOF MS spectrum of an Asp-N digest of bovine fibromodulin using positive (+) and negative (−) ion mode. The tyrosine-sulfated species are clearly seen; each sulfate is annotated with a small circle as +80 adducts in the negative mode. In this spectrum 7 different peptides are detected with 1 or 2 tyrosine sulfates each.



The sulfate group that modifies tyrosine is easily lost when exposed to low pH, high temperature, or energetic collisions such as those used in mass spectrometers. Today, MALDI-TOF MS instruments commonly have both linear and reflector detector capabilities and can be operated in both the positive and negative ion mode. These different options facilitate the studies of tyrosine-sulfated peptides. In the linear mode, fragments that form in the flight tube, post-source decay (PSD), are not separated and hence are detected as if they were still intact. This simplifies the spectral interpretation because only the fragments formed inside the ion source, in-source decay (ISD), would be observed as different from the intact fragments.

This observation is typical for sulfated peptides as illustrated in Fig. 2 using peptide Thr¹⁹-Arg³⁶ from bovine lumican as a model. The mass spectra were acquired in the linear positive (*top*) and negative ion modes (*middle*), respectively. In the positive mode only the non-sulfated peptide was observed (*peak a*) because of in-source fragmentation loss of SO₃, whereas in the negative ion mode both one and two sulfate adducts were seen (*peaks b* and *c*). This is the general behavior of sulfated peptides with MALDI-TOF MS. In reflector mode the picture becomes somewhat more complex because PSD fragments and ISD fragments were observed in the same spectrum. As before, the peptide was detected in both the positive and negative ion mode. The spectrum in the positive ion mode again resulted in detection of the non-sulfated species alone (not shown). However, the negative ion mode spectrum (Fig. 2, *bottom*) appears more complex with three additional peaks (*d*, *e*, and *f*). These peaks correspond to the metastable PSD fragment with losses of SO₃ (~69 Da). The mass difference of the PSD fragment was not equal to 80 Da because the calibration step was omitted in order to separate the ISD and PSD fragments within the same spectrum. The generated spectrum displayed a mixture of ISD (*peaks a* and *b*) and PSD fragments (*peaks d–f*) due to loss of the labile SO₃ group. Because sulfated peptides behaved as non-sulfated species in the positive ion mode, sequence confirmation could be performed using conventional PSD experiments, as illustrated with the N-terminal fibromodulin peptide in Fig. 3. The fragments were matched to the theoretical sequence.

Characterization and Identification of Tyrosine Sulfation in LRR Class II Proteins—To locate the position of individual tyrosine sulfates, a number of enzyme digestions were performed. The strategy for bovine fibromodulin is shown sche-

matically in Fig. 4, where the intact protein is illustrated together with the N-terminal digestion products of endoproteinase Lys-C, trypsin, and Asp-N. The trypsin and Asp-N digestions were performed using the purified large N-terminal fragment from Lys-C (amino acids 19–98). The positive and negative ion spectra from the Asp-N digest are shown in Fig. 5. In a single spectrum we were able to detect 7 different peptides with varying degrees of tyrosine sulfation. The number of tyrosine sulfates detected in each peptide varied between 0 and 2. Detailed information on all detected sulfotyrosine peptides is listed in Table I. Some sites are more constantly sulfated than others, *e.g.* Tyr³⁸ is invariably modified because several peptides containing only this single sulfotyrosine were detected. However, three sulfopeptides in this digest (*m/z* = 2608, 2997, and 3633 Da) did not match the theoretical sequence. Additionally, the sulfated species detected could be a heterogeneous mixture of different sulfation states or a fully sulfated species that readily fragments.

The labile nature of tyrosine sulfate puts high demands on the ionization process, and thus soft ionization methods are of utmost importance in minimizing fragmentation. The use of atmospheric pressure MALDI has been reported to decrease the in-source fragmentation of positive ions as compared with conventional MALDI (47). Other soft ionization techniques such as electrospray ionization (39, 40, 47) and liquid secondary ion mass spectrometry (37) were used to study tyrosine sulfation in the past. However, the complete absence of tyrosine sulfate fragmentation has not generally been observed.

To our knowledge there are no previous reports on sulfation studies using ESI in combination with ion-trap mass spectrometers. In this study, we analyzed tyrosine sulfation of fibromodulin using two different ion-trap instruments that are ideal for tandem MS experiments. The behavior of sulfated peptides in the ion-trap instrument was tested using a sulfated peptide standard (hirudin fragment) in both the positive and negative mode. MS in the positive mode resulted in partial loss of an SO₃ group, and in MS-MS mode the major event was loss of SO₃. In the negative MS mode there was no loss of SO₃, whereas in MS-MS loss of SO₃ and/or water was the main fragmentation event.

The Asp-N digest of fibromodulin was analyzed by static nanospray MS. The data obtained for the unknown peptides indicated that the theoretically available primary sequence

TABLE I
Peptides from LRR class II proteins that have been detected containing at least one post-translational modification in the form of tyrosine sulfation

Peptide sequences are listed, and all tyrosines are labeled in boldface with unambiguously modified tyrosines denoted as "sY." The expected unmodified mass, $[M+H]^+$, is listed together with the number of tyrosine sulfates detected and the MS method, enzyme, and protein used. The residue numbering in human recombinant osteoadherin includes the His tag, and the mature protein begins at residue 20. FM, fibromodulin; OSAD, osteoadherin; LUM, lumican; bov, bovine; hum, human.

Peptide sequence	<i>m/z</i>	No. of sY	MS/enzyme	Protein	Amino acids
pQsYEEDSHWWFQFLR	1953.86	0–1	MALDI/trypsin	FM_bov	19–32
NQQSTYDDPYDPYEPYEPYYPYGGE GPAYAYGSPPQPEPR	4844.06	6–8 ^a	ESI-IT/trypsin	FM_bov	33–74
pQsYEEDSHWW	1262.47	0–1	MALDI/CT	FM_bov	19–27
pQsYEEDSHWWF	1409.54	0–1	MALDI/CT	FM_bov	19–28
DPYYPYEPY	1043.44	0–1	MALDI/Asp-N	FM_bov	42–50
DSHWWFQFLRNQQSTsY	2142.98	0–1	MALDI/Asp-N	FM_bov	23–38
EDSHWWFQFLRNQQSTsY	2272.02	0–1	MALDI/Asp-N	FM_bov	22–38
EEDSHWWFQFLRNQQSTsY	2401.06	0–1	MALDI/Asp-N	FM_bov	21–38
pQYEEDSHWWFQFLRNQQSTY	2675.16	0–1	MALDI/Asp-N	FM_bov	19–38
EPsYpSYGGE GPAsYAsYGSPPQ PEPR	2608.16	0–4 ^b	ESI-IT/Asp-N	FM_bov	51–74
EPYYPYGGE GPAYAYGSPPQPEPR	2608.16	0–2	MALDI/Asp-N	FM_bov	51–74
EPYEPYYPYGGE GPAYAYGSPPQPEPR	2997.32	0–2	MALDI/Asp-N	FM_bov	48–74
DPYDPYEPYEPYYPYGGE GPAYAYGSPPQPEPR	3632.58	0–2	MALDI/Asp-N	FM_bov	42–74
pQYEEDSHWWFQFLRNQQSTYDDPYDPYEPYEPYYPYGGE EGPAYAYGSPPQPEPRDCPQECDCPPNFPTAMYCDNRNLK	9529.94	1–7 ^c	MALDI/Lys-C	FM_bov	19–98
PYDPYEPYEPYYPYGGE GPAYAYGSPPQPEPRDCPQECDC PPNFPTAMYCDNRNLK	6271.64	0–4 ^c	MALDI/LysC + HCOOH	FM_bov	43–98
PYPYEPYEPYYPYGGE GPAYAYGSPPQPEPRDCPQECDCPPN FPTAMYCDNRNLK	6646.79	0–4 ^c	MALDI/LysC + HCOOH	FM_bov	46–98
pQYEDDPHWWFHYLR	1974.86	0–1	MALDI/trypsin	FM_hum	19–33
pQsYEDDPHWW	1258.48	0–1	MALDI/CT	FM_hum	19–27
pQsYEDDPHWWFH	1542.61	0–1	MALDI/CT	FM_hum	19–29
EPsYpSYGV	824.38	0–2	MALDI/Asp-N	FM_hum	51–57
DPYYPYET	1047.43	0–1	MALDI/Asp-N	FM_hum	43–50
DPYYPYETYPYYPYGV	1852.79	0–1	MALDI/Asp-N	FM_hum	43–57
DDPHWWFHYLRSSQSTYY	2429.07	0–2	MALDI/Asp-N	FM_hum	22–39
APQLHHHHHDLDDDDKLQYETYQWDEDYDQEPDDDYQT GFPRQNVHYGVFPHQYTLGCVSECFCTNFPSSMYCDNRK	10033.22	0–6 ^d	MALDI/LysC	OSAD_hum	1–81
TQVFRFPDDDDSESDHDDPDNAHESPEQEGAGHFDLH sYsYENQE	5361.72	0–2 ^c	MALDI/Lys-C	OSAD_hum	376–420
DDDDKLQYETYQW	1718.72	0–1	MALDI/CT	OSAD_hum	14–26
DKLQYETYQW	1373.64	0–1	MALDI/CT	OSAD_hum	17–26
DLHsYsYENQE	1210.50	0–1	MALDI/Asp-N	OSAD_hum	412–420
DDsYQTGFPRQNV	1586.72	0–1 ^c	MALDI/Asp-N	OSAD_hum	17–29
DDDsYQTGFPRQNV	1701.75	0–1 ^c	MALDI/Asp-N	OSAD_hum	16–29
DYGVPHQYTLGCVS	1742.78	0–1 ^c	MALDI/Asp-N	OSAD_hum	30–44
DQEPDDDsYQTGFPRQNV	2170.93	0–1 ^c	MALDI/Asp-N	OSAD_hum	12–29
DYGVPHQYTLGCVSTNFPSSMYC	3523.42	0–1 ^c	MALDI/Asp-N	OSAD_hum	30–58
TYPDYYEYDFPQALYGR	2324.05	0–2 ^c	MALDI/trypsin	LUM_bov	19–36
TYPDYYEYDFPQALY	2110.90	0–1	MALDI/CT	LUM_bov	19–34
TYPDYYEYDFPQAL	1947.83	0–1	MALDI/CT	LUM_bov	19–33
pQYYDYDFPLSIY	1569.68	0–1	MALDI/CT	LUM_hum	19–30
pQYYDYDFPLSIYQSSPNCAPCNCPEsYPSAMYCDE	4232.6	0–2	MALDI/Glu-C	LUM_hum	19–55

^a MS analysis made at the demonstration lab of Bruker Daltonics, Bremen, Germany.

^b MS analysis made at the demonstration lab of Thermofinnigan, London, UK.

^c Acid treatment was performed, confirming the presence of tyrosine sulfation.

^d Recombinant human protein from EBNA cells containing a His tag.

EPYPTGEGPAYAYGSPPQPEPR was incorrect or post-translationally modified. The previously detected peptide at *m/z* 2608.16 with MALDI was suspected to contain a sequence different from the theoretical sequence when we observed the peptide present with up to four tyrosine sulfates in the negative mode. The theoretical sequence contained only three tyrosines, and therefore at least one amino acid appeared incorrect. *De novo* sequencing of the mismatching peptide at 2608 Da (theoretical mass 2489 indicated a mass shift of 119 Da) revealed the sequence PPQPEPR with MS³. However, no significant sequence series were obtained using only MS². By combining data from other MS³ experiments using different *m/z* values, more information about the sequence was obtained. In addition, comparisons with the human sequence suggested a correction of Thr³⁷ to Tyr³⁷ as well as the addition of one amino acid, which in this case corresponded to the mass of glycine (Gly³⁸). These changes resulted in the following sequence: EPsYpSYGGE GPAsYAsYGSPPQPEPR (tyrosines are boldface). When fitting the obtained data to this sequence, most y

ions and several b ions matched the collected MS² spectrum (Fig. 6).

Considering the tryptic digest of fibromodulin, it became apparent that the sample should be analyzed in the negative ion mode. A high signal was observed at *m/z* 1080 (Fig. 7) corresponding to a charge state of -5 (isotopic peaks were separated by a mass difference of 0.2 Da). When this peak was selected for MS-MS fragmentation, it lost 16 Da corresponding to the loss of SO₃. The advantage of using this technology is that tandem MS (MS)^{*n*}, where *n* < 12, can be performed. Consequently the *m/z* 1064 peak was isolated for another MS fragmentation (MS)³, where again we detected loss of SO₃. This was repeated until no more losses of SO₃ could be detected and loss of water became apparent. In this case the peptide of interest lost five SO₃ groups before the signal became too low for further analysis. The previously corrected sequence suggests that seven sulfogroups are present in this peptide in order for it to match the experimental data. This peptide then lost SO₃ in a stepwise manner until

FIG. 6. MS-MS spectrum, in positive mode on an ESI ion-trap MS, of a bovine fibromodulin peptide obtained by endoproteinase Asp-N digestion. The peptide sequence EPsYPsYGGEEG-PA_sYAsYGSP_sQPEPR was confirmed to fit the obtained data, and matching b/y ions are annotated accordingly.

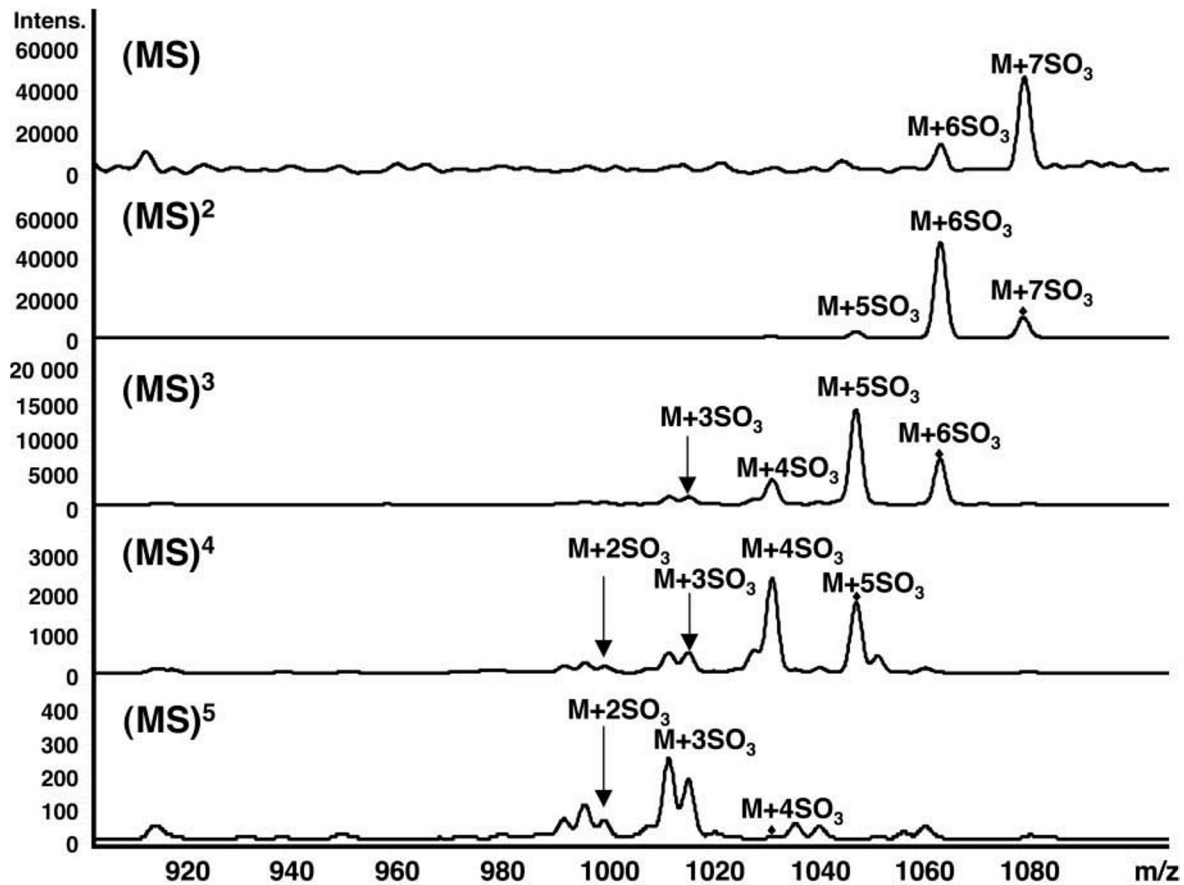
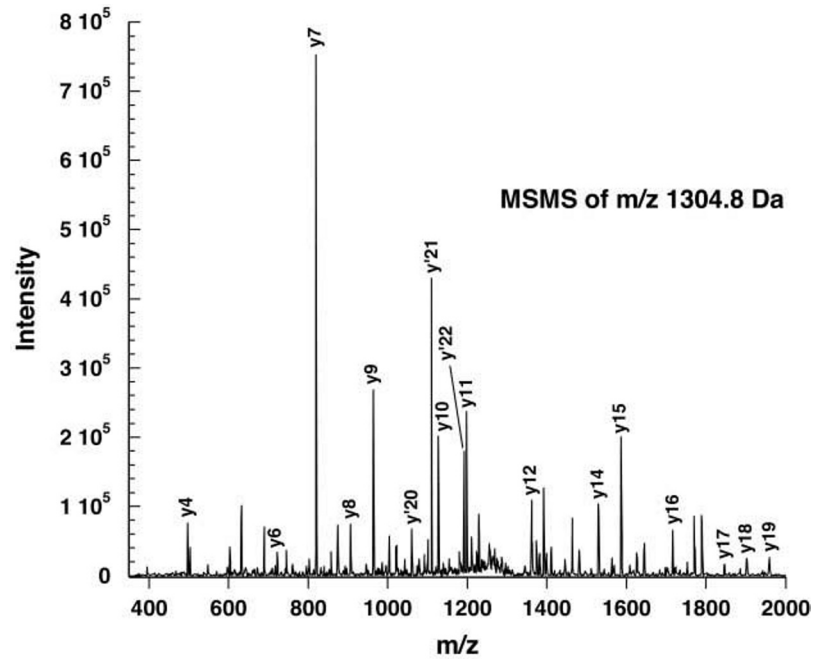


FIG. 7. Multiple tyrosine sulfations studied with ESI ion-trap in negative mode. $m/z = 1080$ corresponds to the tyrosine-sulfated peptide NQQSTYDDPYDPYPYEPYPYGGEEGPAYAYGSP_sQPEPR of a high pressure liquid chromatography fraction of trypsin-digested bovine fibromodulin. The tandem MS experiment revealed a stepwise loss of a sulfate group.

only two sulfogroups remained. The reason that the peptide did not fragment further is unclear, but it might be that the -5 charge was no longer stable when removing too many negatively charged SO_3 groups. A lower, but significant, signal was also detected for the same peptide containing 8 tyrosine sulfates at the charge state -6 . Apparently, in this

tryptic peptide we detected 8 of the 9 possible sites that carry the modification.

In the case of human fibromodulin the same experiments were performed using MALDI-TOF MS. The data obtained suggest a similar sulfation pattern with at least 6 tyrosine sulfates. The trypsin digest indicated that Tyr²⁰ was sulfated

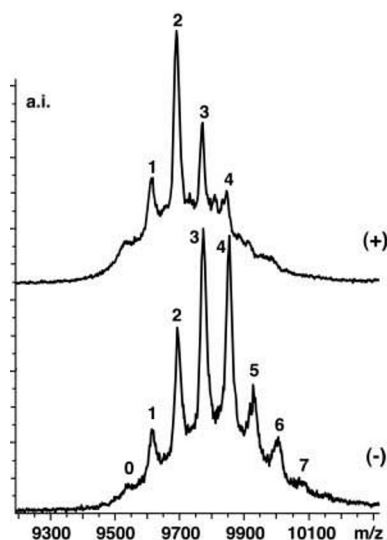


FIG. 8. Multiple tyrosine sulfation sites in bovine fibromodulin, pQYEDSHWVFQFLRNQQSTY DDPYDPYPYEPYEPYPYG-GEETPAYAYGSPQPEPRDCPQECDCPPNFPTAMYCDNR-NLK, illustrates a different behavior as compared with smaller, less sulfated peptides. The sulfated species are more stable in the positive mode where multiple sulfations (peaks 0–5, most abundant peak at 2) are present. In negative mode the peptides are even more stable (peaks 0–7, most abundant peak at 3–4) but the difference is not as significant as previously shown on smaller, less sulfated peptides (e.g. as shown in Fig. 2).

similarly to bovine fibromodulin. The Asp-N digest of human fibromodulin was most informative resulting in the detection of 5 additional tyrosine sulfates.

The other member of the class II LRR family, osteoadherin, was analyzed in a similar manner. The results are listed in Table I and show that both the N-terminal and the C-terminal parts of the protein are sulfated. The N-terminal Lys-C peptide (amino acids 1–81), including the His tag, was identified as carrying 0–6 tyrosine sulfate residues, whereas the C-terminal peptide (amino acids 376–420) contained 0–2 such modifications. These sulfation sites were confirmed in the peptides of the Asp-N digest. The sulfated tyrosines in osteoadherin are similar in number compared with fibromodulin but are more evenly distributed. For bovine and human lumican, two tyrosine sulfates were detected close to the N terminus.

Large sulfopeptides Are More Stable—The behavior of sulfated peptides described in this work, where sulfated species are detected as if non-sulfated in the positive ion mode, seems to hold true for peptides below 3000 Da. Larger peptides from 3–10 kDa appeared more stable and underwent less ISD in the positive ion mode. In this case, switching polarities did not alter the observed peaks significantly. This is illustrated in Fig. 8, which shows that multiple tyrosine sulfates were observed in both the positive and negative ion modes for the N-terminal peptide (amino acids 19–98) of fibromodulin. A number of sulfate adducts (0–7) in both modes and a small shift toward higher sulfation states in the negative mode were detected.

Chemical Desulfation of Tyrosine—It is well documented that tyrosine sulfates are labile to acid hydrolysis (46, 48). A short acid treatment can be used in combination with the results from positive/negative ion mode analysis in order to prove existence of tyrosine sulfates. Results from bovine lumican using 0, 1, or 5 min hydrolysis times were obtained by MALDI-TOF MS (Fig. 9). The zero-time control indicated the presence of two tyrosine sulfates. The desulfation at 1 min was not complete, but 5 min incubations were sufficient for complete sulfate removal. It is of major importance to avoid excess exposure to acid, because acid hydrolysis of the peptide back-

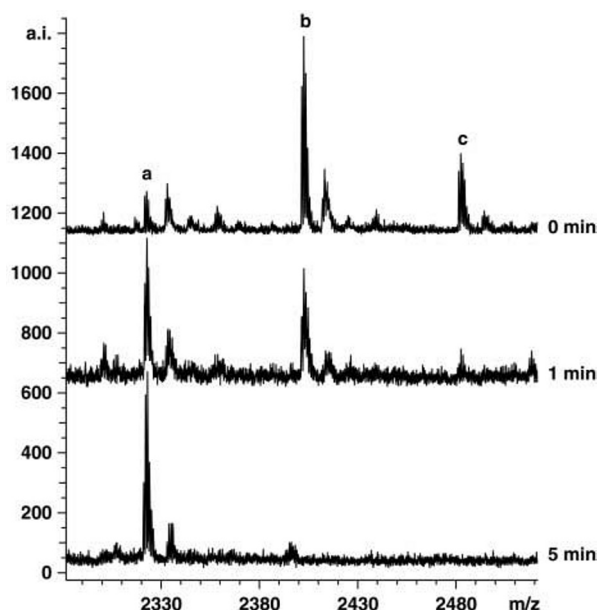


FIG. 9. Acid treatment of tyrosine sulfates on the same lumican N-terminal tryptic peptide as shown in Fig. 1. Treatment with 1 M HCl at 99 °C for 0, 1, and 5 min is shown, illustrating the acid lability of a sulfogroup being removed completely after 5 min of treatment.

bone can occur, predominantly between Asp–Pro bonds (49, 50). This was also observed in our samples to some extent. Desulfation with acid was used to confirm tyrosine sulfation in bovine lumican, fibromodulin, and human osteoadherin. The treatment successfully desulfated the peptides as shown in Table I.

Enzymatic Removal of Tyrosine Sulfate—Another way to distinguish sulfation from phosphorylation is by enzymatic treatments, which specifically remove these respective modifications. Alkaline phosphatase is a readily available enzyme that dephosphorylates Ser, Thr, and Tyr residues. This treatment had no effect upon sulfated control peptides but worked well on phosphorylations (data not shown). Alkaline phosphatase treatment of osteoadherin digests caused no shifts in the MS spectra, indicating that the +80 adducts were not caused by tyrosine phosphorylations. The enzymatic removal of tyrosine sulfate is unfortunately not as straightforward. A number of enzymes of different origin are available commercially and have been tested previously with some sulfated standards (51). Several sulfatases were evaluated for desulfation of hirudin. In our hands, the efficiency of these was not satisfactory and resulted in only partial desulfation, despite the use of large amounts of enzyme. Without protease inhibitors, samples were in some cases completely degraded, probably due to activity by peptidases. In experiments performed on biological samples such as lumican and fibromodulin, the desulfation failed. Seemingly, no general solution for enzymatic or chemical removal exists today, which suggests that each individual sulfated site must be tested by specifically adapted procedures.

DISCUSSION

Tyrosine sulfation in three LRR class II proteins has been characterized using various mass spectrometric techniques including MALDI-TOF MS and ESI ion-trap instruments. The labile sulfogroup was studied in both the positive and negative ion modes. Use of the negative ion mode appeared necessary to minimize the fragmentation event in ISD and PSD events. Bovine fibromodulin is sulfated with up to 9 tyrosine sulfates in the N-terminal part of the protein. Only one of 10 tyrosines was not found to carry the modification. This tyrosine residue is

encompassed within amino acids 39–50, but the obtained data could not verify whether a particular residue was never sulfated. For most sulfopeptides it appeared that ESI was the optimal ionization mode with less fragmentation than MALDI-TOF MS. For larger peptides, the stability of tyrosine sulfates was enhanced in the positive mode using MALDI. Within the same peptide, up to 7 tyrosine sulfates were detected. However, by ESI a smaller peptide was detected with up to 8 sulfates. Acid treatment is a possible way to remove tyrosine sulfate, but the presence of acid sensitive Asp–Pro bonds in the protein complicates this approach. In human fibromodulin at least 6 tyrosine sulfations were detected within the same region as in the bovine sequence. Osteoadherin was sulfated in both its N-terminal (6 sulfates) and C-terminal (2 sulfates) parts. Treatment of these peptides with alkaline phosphatase caused no mass shift, whereas the α -casein control was successfully dephosphorylated. This showed that these tyrosines were not phosphorylated. However, a smaller peptide (amino acids 376–420) was observed to shift in mass and appeared desulfated after 5 min of acid treatment. Both bovine and human lumican contained at least two tyrosine sulfates in the N-terminal part of the protein. The clustering of sulfated tyrosines and the acidic nature of these regions precludes digestions with trypsin. However, endoproteinase Asp-N is a suitable alternative in this case, with its ability to cleave acidic residues (Asp and Glu) also next to sulfated tyrosine residues.

There is a software program, The Sulfinator, that can be used to predict tyrosine sulfation (52). When the sequences of the proteins used in this study were analyzed by this program, the following number of tyrosine sulfates were predicted: bovine fibromodulin, 6; human fibromodulin, 5; osteoadherin, 2; bovine lumican, 6; and human lumican, 0; however, our data support 9, 6, 8, 2, and 2 sulfates, respectively. Apparently, this program does not accurately predict tyrosine sulfation in this class of proteins.

The biological role of tyrosine sulfation in these proteins is as yet unknown, but the presence of negatively charged clusters around the termini of these proteins suggests a role in protein interactions. It could also be important for increased stability or correct protein folding. Future developments in derivatization of aryl sulfates may allow tandem MS to be performed on sulfated peptides without fragmentation. To our knowledge this has not been achieved. The methodology described in this work uses present technology to obtain a great deal of information. However, the exact location of a single sulfation in a peptide with multiple tyrosine residues by mass spectrometry remains a challenging problem.

Acknowledgments—The skillful technical assistance at the demonstration lab of Bruker Daltonics (Dirk Wunderlich and Andrea Kiehne) and Thermofinnigan (Gary Woffendin) is acknowledged and much appreciated.

REFERENCES

- Huttner, W. B. (1987) *Trends Biochem. Sci.* **12**, 361–363
- Niehrs, C., Beisswanger, R., and Huttner, W. B. (1994) *Chem. Biol. Interact.* **92**, 257–271
- Bundgaard, J. R., Vuust, J., and Rehfeld, J. F. (1997) *J. Biol. Chem.* **272**, 21700–21705
- Seibert, C., Cadene, M., Sanfiz, A., Chait, B. T., and Sakmar, T. P. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11031–11036
- Ouyang, Y. B., Crawley, J. T. B., Aston, C. E., and Moore, K. L. (2002) *J. Biol. Chem.* **277**, 23781–23787
- Parenti, G., Meroni, G., and Ballabio, A. (1997) *Curr. Opin. Genet. Dev.* **7**, 386–391
- Daniele, A., Parenti, G., d'Addio, M., Andria, G., Ballabio, A., and Meroni, G. (1998) *Am. J. Hum. Genet.* **62**, 562–572
- Brunetti-Pierri, N., Andreucci, M. V., Tuzzi, R., Vega, G. R., Gray, G., McKeown, C., Ballabio, A., Andria, G., Meroni, G., and Parenti, G. (2003) *Am. J. Med. Genet.* **117**, Suppl. A, 164–168
- Dahl, H. H. M., Osborn, A. H., Hutchison, W. M., Thorburn, D. R., and Sheffield, L. J. (1999) *Mol. Genet. Metab.* **68**, 503–506
- Sheffield, L. J., Osborn, A., Hutchison, W., Sillence, D. O., and Dahl, H. H. M. (1997) *Am. J. Hum. Genet.* **61**, A27
- Parenti, G., Buttitta, P., Meroni, G., Franco, B., Bernard, L., Rizzolo, M. G., Brunetti-Pierri, N., Ballabio, A., and Andria, G. (1997) *Am. J. Med. Genet.* **73**, 139–143
- Horton, W. A. (1995) *Eur. J. Hum. Genet.* **3**, 357–373
- Bernimoulin, M. P., Zeng, X. L., Abbal, C., Giraud, S., Martinez, M., Michielin, O., Schapira, M., and Spertini, O. (2003) *J. Biol. Chem.* **278**, 37–47
- Leppänen, A., Mehta, P., Ouyang, Y. B., Ju, T. Z., Helin, J., Moore, K. L., van Die, I., Canfield, W. M., Mcever, R. P., and Cummings, R. D. (1999) *J. Biol. Chem.* **274**, 24838–24848
- Leppänen, A., White, S. P., Helin, J., Mcever, R. P., and Cummings, R. D. (2000) *J. Biol. Chem.* **275**, 39569–39578
- Wilkins, P. P., Moore, K. L., Mcever, R. P., and Cummings, R. D. (1995) *J. Biol. Chem.* **270**, 22677–22680
- Costagliola, S., Panneels, V., Bonomi, M., Koch, J., Many, M. C., Smits, G., and Vassart, G. (2002) *EMBO J.* **21**, 504–513
- Nlend, M. C., Cauvi, D., Venot, N., and Chabaud, O. (1999) *Biochem. Biophys. Res. Commun.* **262**, 193–197
- Venot, N., Nlend, M. C., Cauvi, D., and Chabaud, O. (2002) *Biochem. Biophys. Res. Commun.* **298**, 193–197
- Bannert, N., Craig, S., Farzan, M., Sogah, D., Santo, N. V., Choe, H., and Sodroski, J. (2001) *J. Exp. Med.* **194**, 1661–1673
- Farzan, M., Mirzabekov, T., Kolchinsky, P., Wyatt, R., Cayabyab, M., Gerard, N. P., Gerard, C., Sodroski, J., and Choe, H. (1999) *Cell* **96**, 667–676
- Farzan, M., Vasilieva, N., Schnitzler, C. E., Chung, S., Robinson, J., Gerard, N. P., Gerard, C., Choe, H., and Sodroski, J. (2000) *J. Biol. Chem.* **275**, 33516–33521
- Farzan, M., Schnitzler, C. E., Vasilieva, N., Leung, D., Kuhn, J., Gerard, C., Gerard, N. P., and Choe, H. (2001) *J. Exp. Med.* **193**, 1059–1065
- Farzan, M., Babcock, G. J., Vasilieva, N., Wright, P. L., Kiprilov, E., Mirzabekov, T., and Choe, H. (2002) *J. Biol. Chem.* **277**, 29484–29489
- Fong, A. M., Alam, S. M., Imai, T., Haribabu, B., and Patel, D. D. (2002) *J. Biol. Chem.* **277**, 19418–19423
- Preobrazhensky, A. A., Dragan, S., Kawano, T., Gavrilin, M. A., Gulina, I. V., Chakravarty, L., and Kolattukudy, P. E. (2000) *J. Immunol.* **165**, 5295–5303
- Dong, J. F., Li, C. Q., and Lopez, J. A. (1994) *Biochemistry* **33**, 13946–13953
- Dong, J. F., Ye, P., Schade, A. J., Romo, G. M., Gao, S., McIntire, L. V., and Lopez, J. A. (1999) *Blood* **94**, 441A
- Dong, J. F., Ye, P., Schade, A. J., Gao, S., Romo, G. M., Turner, N. T., McIntire, L. V., and Lopez, J. A. (2001) *J. Biol. Chem.* **276**, 16690–16694
- Leyte, A., Vanschijndel, H. B., Niehrs, C., Huttner, W. B., Verbeet, M. P., Mertens, K., and Vanmourik, J. A. (1991) *J. Biol. Chem.* **266**, 740–746
- Marchese, P., Murata, M., Mazzucato, M., Pradella, P., Demarco, L., Ware, J., and Ruggeri, Z. M. (1995) *J. Biol. Chem.* **270**, 9571–9578
- Tait, A. S., Dong, J. F., Lopez, J. A., Dawes, I. W., and Chong, B. H. (2002) *Blood* **99**, 4422–4427
- Meh, D. A., Siebenlist, K. R., Brennan, S. O., Holyst, T., and Mosesson, M. W. (2001) *Thromb. Haemostasis* **85**, 470–474
- Talbo, G., and Roepstorff, P. (1993) *Rapid Commun. Mass Spectrom.* **7**, 201–204
- Talbo, G., and Mann, M. (1994) *Techniques in Protein Chemistry V* (Crabb, J. W., ed) pp. 105–113, Academic Press, San Diego
- Steen, H., Kuster, B., Fernandez, M., Pandey, A., and Mann, M. (2001) *Anal. Chem.* **73**, 1440–1448
- Yagami, T., Kitagawa, K., and Futaki, S. (1995) *Rapid Commun. Mass Spectrom.* **9**, 1335–1341
- Zaia, J., Boynton, R., Heinegård, D., and Barry, F. (2001) *Biochemistry* **40**, 12983–12991
- Severs, J. C., Carnine, M., Eguizabal, H., and Mock, K. K. (1999) *Rapid Commun. Mass Spectrom.* **13**, 1016–1023
- Nemeth-Cawley, J. F., and Rouse, J. C. (2002) *J. Mass Spectrom.* **37**, 270–282
- Wendel, M., Sommarin, Y., and Heinegård, D. (1998) *J. Cell Biol.* **141**, 839–847
- Axelsson, I., and Heinegård, D. (1975) *Biochem. J.* **145**, 491–500
- Funderburgh, J. L., and Conrad, G. W. (1990) *J. Biol. Chem.* **265**, 8297–8303
- Heinegård, D., Larsson, T., Sommarin, Y., Franzen, A., Paulsson, M., and Hedbm, E. (1986) *J. Biol. Chem.* **261**, 3866–3872
- Schuerenberg, M., Luebbert, C., Eickhoff, H., Kalkum, M., Lehrach, H., and Nordhoff, E. (2000) *Anal. Chem.* **72**, 3436–3442
- Ecarot-Charrier, B., Bouchard, F., and Delloye, C. (1989) *J. Biol. Chem.* **264**, 20049–20053
- Wolfender, J. L., Chu, F. X., Ball, H., Wolfender, F., Fainzilber, M., Baldwin, M. A., and Burlingame, A. L. (1999) *J. Mass Spectrom.* **34**, 447–454
- Huttner, W. B. (1984) *Methods Enzymol.* **107**, 200–223
- Landon, J. (1977) *Methods Enzymol.* **47**, 145–149
- Marcus, F. (1985) *Int. J. Pept. Protein Res.* **25**, 542–546
- Suiko, M., Tojo, T., Fernando, P. H. P., Sakakibara, Y., Kawano, J., and Liu, M. C. (1996) *Biosci. Biotechnol. Biochem.* **60**, 137–138
- Monigatti, F., Gasteiger, E., Bairoch, A., and Jung, E. (2002) *Bioinformatics* **18**, 769–770