Identification of Endogenous Phosphorylation Sites of Bovine Medium and Low Molecular Weight Neurofilament Proteins by Tandem Mass Spectrometry†

Sarah Trimpin,‡ April E. Mixon,‡ Martha D. Stapels,‡ Moo-Young Kim,‡ Peter S. Spencer,§ and Max L. Deinzer*,‡

Department of Chemistry, Oregon State University, Corvallis, Oregon 97331, and CROET, Oregon Health & Science University, Portland, Oregon 97201

Received August 27, 2003

ABSTRACT: Neurofilament proteins (NFP) are intermediate filaments found in the neuronal cytoskeleton. They are highly phosphorylated, a condition that is believed to be responsible for the assembly and stability of the filaments. Matrix-assisted laser desorption/ionization (MALDI) time-of-flight (TOF) mass spectrometry (MS) shows molecular masses for bovine NFP subunits of 63, 105, and 125 kDa for NFL, NFM, and NFH. Mass spectrometric de novo sequencing was used to determine the N-terminal sequence of bovine NFM (115 amino acids), which was previously unknown. Molecular mass information shows that there is one-half equivalent phosphate group on NFL and 24 on NFM. For the first time, it is shown that bovine NFL has three phosphorylation sites (Ser^{55}, Ser^{66}, and Ser^{472}) and NFM has 22 (Ser^{512}, Ser^{546}, Ser^{554}, Ser^{560}, Thr^{627}, Ser^{629}, Ser^{634}, Ser^{639}, Thr^{646}, Ser^{649}, Ser^{654}, Ser^{664}, Ser^{669}, Thr^{676}, Ser^{679}, Ser^{684}, Ser^{694}, Ser^{726}, Ser^{750}, Ser^{756}, Ser^{770}, and Ser^{846}) and two tentative sites (Ser^{659}/Thr^{661} and Thr^{840}). Ser^{66} was previously not known to be phosphorylated in NFL of other species, while two sites (Ser^{55} and Ser^{472}) are consistent with the phosphorylations observed in other mammalian NFLs. The three sites, Ser^{55}, Ser^{66}, Ser^{472}, are heterogeneously phosphorylated. Phosphorylation in bovine NFM occurs mainly in the Lys-Ser-Pro (KSP) region, but the Val-Ser-Pro and Ser-Glu-Lys motifs are also phosphorylated. Most of the phosphorylation sites are in accordance with those previously identified in other mammalian NFMs. In bovine NFM, 16 out of the 22 sites are always phosphorylated (Ser^{512}, Thr^{627}, Ser^{629}, Ser^{634}, Ser^{639}, Thr^{646}, Ser^{649}, Ser^{654}, Ser^{664}, Ser^{669}, Thr^{676}, Ser^{679}, Ser^{684}, Ser^{694}, Ser^{726}, and Ser^{750}), all of which are contained in the KSP region, and six are sometimes phosphorylated (Ser^{546}, Ser^{554}, Ser^{560}, Ser^{756}, Ser^{770}, and Ser^{846}). The NFPs have other modifications, including deamidation, oxidation, and N-terminal acetylation. Pyroglutamic acid formation also occurs.

NFPs[1] are the major axonal cytoskeletal components of neurons and are essential for establishing the correct diameters of large myelinated motor and sensory axons (*1*−*3*). The NFP assembly is composed of three different polypeptide subunits in the adult neuron: NFL (low molecular mass), NFM (intermediate molecular mass), and NFH (high molecular mass) (*4*−*7*) with apparent molecular masses as measured by SDS-polyacrylamide gel electrophoresis (PAGE) of 70, 150, and 200 kDa, respectively. Each polypeptide subunit is organized into three domains: a variable amino-terminal head domain, a central highly conserved segmented α-helical rod domain of approximately 310 amino acids, and a hypervariable C-terminal tail domain (*8*). NFPs primarily located in the axon are reported to be highly phosphorylated in vivo (*9*). Static as well as dynamic properties of NFPs, including their assembly, stability, plasticity, and transport, are regulated by phosphorylation/dephosphorylation through kinase and phosphatase activities (*10*, *11*). Phosphorylation in the N-terminal head domain has been shown to have a regulatory role in subunit assembly and disassembly. Phosphorylation in the carboxy-terminal domains, particularly in NFM and NFH, has been implicated in the maintenance of axonal caliber (*12*, *13*), in determining the rate of transport of neurofilaments in the axon (*14*), in the promotion of cross-linking between the filaments, and in the maintenance of interactions with other cytoskeletal proteins (*15*).

Various approaches have been used to evaluate the extent and sites of NFP phosphorylation in vivo. Proteolytic digestion of proteins followed by mass spectral analysis using HPLC-MS/MS and database searching have become popular for characterizing posttranslational phosphorylation (*16*). Phosphorylated sites of human NFH have been characterized by this method (*9*, *17*). Within the head domain of mouse NFM, Ser^{23} was found to be a major phosphorylation site resulting from protein kinase A activity (*18*). In-gel tryptic digestion, IMAC, nanoelectrospray ionization MS, and

† This work was supported by NIEHS Grant ES10338 to M.L.D. and P.S.S. The Environmental Health Sciences Center Grant from NIH/NIEHS ES00210 is also acknowledged. The Mass Spectrometry Core of the EHSC was used in support of this work.

* To whom the correspondence should be addressed. Tel: 541-737-1773. Fax: 541-737-0497. E-mail: max.deinzer@orst.edu.

‡ Oregon State University.

§ Oregon Health & Science University.

[1] Abbreviations: AA, acetic acid; ACN, acetonitrile; EDTA, ethylenediaminetetraacetic acid; EGTA, ethylene glycol-bis(β-aminoethyl ether)tetraacetic acid; ESP, glutamic acid-serine-proline; IMAC, immobilized metal affinity chromatography; KSP, lysine-serine-proline; MS/MS, tandem mass spectrometry; NFP, neurofilament protein; NFH, NFM, and NFL, high, medium, and low molecular weight neurofilament proteins, respectively; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; VSP, valine-serine-proline.

10.1021/bi030196q CCC: $27.50
Published on Web 01/31/2004

matrix-assisted laser desorption/ionization (MALDI)-MS have all been used to identify phosphorylation sites of rat NFL and NFM (*15*, *19*). Early investigations suggested that Lys-Ser-Pro (KSP) repeats in the tail domains of NFM and NFH containing the phosphorylated epitopes (*20*). In vivo ^{32}P-labeling experiments identified phosphorylated sites in NFL and NFM subunits (*9*). Six Ser residues, located mainly in KSP repeats in the C-terminal tail domain NFM of the rat, were identified as being phosphorylated in vivo (*21*). For porcine NFH, one Ser residue in the AKSPEK repeat was reported to be phosphorylated in vivo (*22*). Of nine phosphorylated peptides, four serines were clearly identified as being phosphorylated (*22*), three of them in the KSP and variant KSP motifs in the tail domain of porcine NFM. An electrospray mass spectrometric approach (*9*) identified 12 Ser residues within KSPXKX, KSPXXX, and KSPXXK motifs in the NFH tail domain of the rat as being endogenously phosphorylated, whereas Ser residues in the non-KSP region were not phosphorylated. Further MS approaches for determination of phosphorylation sites include studies on chicken NFM (*23*), where 20 phosphorylation sites were identified on Ser and Thr residues, and squid NFH (*14*), where 53 phosphorylation Ser sites were characterized.

Despite a number of attempts at direct peptide mapping of proteolytic digests of NFPs by LC-MS and LC-MS/MS, the results have been less than satisfactory as relatively limited sequence coverage has been achieved (*15*, *19*). The relatively small numbers of peptides that have been detected by these mass spectrometric methods may in part be due to the presence of a high percentage of negatively charged residues such as glutamic (Glu) and aspartic acids (Asp). In addition, the high numbers of phosphorylated serines and threonines may further limit the numbers of peptides that can be analyzed. Indeed, modification of phosphoproteins by β-elimination/ethanethiol addition was used to obtain complementary sequence information and more extensive sequence coverage (*17*, *24*).

Trypsin is known to exhibit considerably reduced activity when acidic residues are present on either side of a potentially susceptible bond. Other proteolytic enzymes, Glu C or Lys C, can, of course, be used to complement the results from tryptic digestion (*24*). However, the large numbers of acidic residues in some peptides may still preclude their detection by ESI mass spectrometry. Thus, previous studies using rat NFM and NFL resulted in peptide maps covering just 64 and 81%, respectively (*15*).

Assembly and disassembly of the NFP subunits are believed to be regulated by phosphorylation (*9*, *15*, *25*). The identification of phosphorylated sites in these proteins is important for understanding the mechanisms and biological implications that result from aberrant phosphorylation, which may play a role in the etiology of Parkinson's and Lewy body dementia (*15*). Accumulation of NFPs within the nerve cells has been seen in patients diagnosed with amyotrophic lateral sclerosis (*26*), Alzheimer's, Charcot-Marie-Tooth disease, and some dementias. Previous studies have correlated neurofilament transport with the degree of their phosphorylation. Therefore, phosphorylation on NFM and NFH may be the regulatory mechanism for transport along the axon. Defective axonal transport and accumulation of NFPs within the axon may lead to the development of several neurological diseases and eventual nerve cell death (*15*). In this paper, we describe the results of our mass spectrometric investigations of bovine neurofilament peptides, which led to the identification of definitive and tentative phosphorylation sites in NFL and NFM. Full sequence information is only available in the literature for bovine NFL. The sequence of the central α-helical rod domain and the C-terminal domain was known for bovine NFM, but that of the N terminus has not yet been reported. Currently, there is no published information on the sequence of bovine NFH. In general, there is no information on the in vivo phosphorylation sites for bovine NFPs. The lack of this critical information drew our attention as we were utilizing readily available bovine NFPs as models for determining the sites of adduction by alkane ketones and diones. We wish to report the results of our investigations on the identification of phosphorylation of NFL and NFM.

MATERIALS AND METHODS

Materials. The NFP mixture was provided by Cytoskeleton Inc. (Denver, CO), which was isolated and purified from bovine spinal cord tissue (catalog no. NFS01-C). The proteins were supplied in a nonpolymerized form in 5.0 mM Tris buffer (pH 8.7, 1 mM EDTA and 1 mM EGTA[2]) und used without further purification. Acetic acid (AA), TFA, bromophenol blue, citric acid, monobasic ammonium phosphate, ammonium bicarbonate, glycerol, freeze-dried pepsin, α-cyano-4-hydroxycinnamic acid, Tris-HCl, Trizma (Tris-base), and freeze-dried lysyl endopeptidase (Lys C) were purchased from Sigma Chemical Co. (St. Louis, MO). Acrylamide (>99% pure), ammonium persulfate, 2-mercaptoethanol, N,N′-methylene-bis-acrylamide, and N,N,N′N′-tetramethylethylene-diamine were from Bio-Rad (Hercules, CA). Sequencing-grade trypsin was purchased from Promega (Madison, WI). The GelCode Blue (Coomassie Blue) stain reagent and the phosphopeptide isolation kit were from Pierce Biotechnology (Rockford, IL). HPLC grade acetonitrile (ACN), glycine, and SDS were supplied by Fisher Scientific (Pittsburgh, PA). Water was generated with a Milli-Q Ultrapure water purification system from Millipore Corp. (Bedford, MA).

Slab Gel Electrophoresis. Small scale denaturing PAGE was conducted using the Mini-Protean III mini-gel system (Bio-Rad) with 1 mm built-in spacers. Stacking gels (2 cm) were composed of 3% acrylamide, 0.1% N,N′-methylene-bis-acrylamide, 125 mM Tris-HCl (pH 6.8), and 0.1% SDS. Resolving gels were made with 7% acrylamide in 0.2% N,N′-methylene-bis-acrylamide, 375 mM Tris-HCl (pH 8.8), and 0.1% SDS. Samples were mixed 1:5 with cracking buffer [0.125 M Tris-HCl (pH 6.8), 360 mM 2-mercaptoethanol, 25% (w/v) glycerol, 2.5% SDS, and 0.1% bromophenol blue], heated at 100 °C for 15 min, and then loaded onto the gel. Electrophoresis was conducted with 25 mM Tris-base, 200 mM glycine, and 0.1% SDS as the running buffer at 200 V and at room temperature until the tracking dye had migrated 0.5 cm from the bottom of the gel. Gels were then stained with 125 mL of GelCode Blue and partially destained with water to visualize protein bands.

Proteolytic Digestions. Digestions were performed both in-gel and in-solution (Table 1). For in-solution tryptic

[2] Vendors information: This buffer is essential for the proper dilution of NFP and prevents the protein from aggregating or forming aberrant filaments.

Table 1: Employed Proteases, Digestion Methods (In-gel, In-solution), Incubation Times, and Sample Treatment (IMAC)

method	trypsin	LysC	pepsin
in-gel	NFL, NFM	not obtained	NFM
in-solution	NFP mixture 10, 30, and 40 min; 1, 2, 6, and 24 h	NFP mixture 10 and 20 min; 1, 2, 4, 8, 17, and 24 h	NFP mixture 2 and 4 h
in-solution, IMAC	NFP mixture 10 min	not obtained	not obtained

digestions, 1.5 μL of 0.5 μg/μL trypsin (in 0.1 M AA) was added to 10 μL of NFP mixture and the digestion mixture was placed on a 37 °C water bath. To perform time-dependent digestions (10 min to 24 h), aliquots of the reaction mixture were pipetted off periodically and the reactions were stopped by placing the reaction mixtures in a −20 °C freezer. For Lys C digestions, 50 μL of MilliQ was added to the freeze-dried Lys C to give a stock solution of 0.1 μg/μL and then 12.5 μL of this solution was combined with 12.5 μL of the NFP mixture. The time-dependent Lys C in-solution digestion was analogous to the tryptic digestion. A stock solution of 3 μg/μL pepsin in citric acid buffer (pH 2) was prepared for pepsin digestions. For in-solution peptic digestions, 5 μL of the NFP mixture was mixed with 5 μL of 0.1 M citric acid buffer to adjust the pH to approximately 2. Pepsin solution (3 μL) was added, and the reaction mixture was placed on a 37 °C water bath for 1.5 h. For in-gel digestions, gel bands were cut and destained with water if necessary. Bands were then dehydrated with ACN before being immersed in 50 μL of a solution containing the enzyme. Incubation was carried out in an ice bath for 30 min. The enzyme solutions were removed, appropriate digestion buffers (0.1 M citric acid buffer for pepsin and 20 mM ammonium bicarbonate for trypsin) were added to the gel pieces, and the digestion solutions were then placed in a 37 °C water bath for 1.5 h (pepsin) or overnight (trypsin).

Phosphopeptide Isolation. To isolate phosphopeptides from the rest of the peptides in the tryptic digest of the NFPs, an IMAC-based phosphopeptide isolation kit was used. The kit contains a resin with chelated gallium(III) ions, which specifically bind to phosphopeptides (*27*). The peptide mixtures were adjusted to pH 2 with 5% AA. The sample was then mixed with the resin for 10 min at room temperature. After nonspecifically bound peptides were washed from the resin with 0.1 M AA, 50:50 ACN:0.1 M AA, and water, phosphopeptides were eluted with 50 mM ammonium phosphate.

Chromatography and Mass Spectrometry. Molecular masses of NFP subunits were obtained by HPLC-MALDI-time-of-flight (TOF)-MS off line. Prior to HPLC fractionation, the NFP mixture was centrifuged to remove particulate matter. A 200 μL loop with a C_4-guard column (Michrom) was inserted in the Rheodyne (Cotati, CA) injector. An analytical grade C_4-column (Vydak 4.6 mm i.d.) was used with a Waters 2487 dual λ detector (AUFS = 2.0). The column was equilibrated with 0.05% TFA/H_2O at a flow rate of 1 mL/min. NFP (90 μL, 450 μg) was loaded onto the guard column and eluted over 40 min (10% TFA:ACN to 60% ACN). The fractions containing the proteins were collected, concentrated to remove ACN, and buffer exchanged into 100 μL of storage buffer (50 mM NH_4HCO_3 and 5 mM $CaCl_2$). For LC-MS/MS, 5 μL of a sample was mixed with 5 μL of solvent A and 3 μL of this solution was injected onto a 5 mm × 0.32 mm C_{18}-LC packing trap. Both solvents A and B contained 0.1% formic acid and 0.005% TFA. Solvent A contained 3% ACN, while solvent B contained 80% ACN. The column was a 15 cm long picofrit column (75 μm i.d.) packed with Jupiter C_{18} from Phenomenex (Phenomenex, Inc., Torrance, CA). The initial LC condition was 3% solvent B for 5 min to wash the sample, followed by a gradient up to 40% solvent B over 40 min, to 70% solvent B at 50 min, to 90% solvent B at 52 min, and then maintained at 90% solvent B to 60 min. A Waters (Milford, MA) CapLC system was used with a flow rate estimated to be 300 nL/min.

Mass Spectrometry. The mass spectrometer used for LC-MS/MS was a Quadrupole time-of-flight (Q-TOF) Global MS from Micromass (Manchester, U.K.) operated in the electrospray mode with a spray voltage of 3.5 kV. Data-dependent MS/MS was used with a 0.5 s survey scan and 2.5 s MS/MS scans on the three most abundant signals in the MS survey scan. To perform neutral loss scanning, the collision energy was varied from 7 to 30 V to specifically monitor for a loss of 98 Da (*28*), and then, MS/MS was performed on the precursors that yielded that loss. Samples for MALDI-TOF-MS were mixed in a 1:3 ratio of α-cyano-4-hydroxycinnamic acid in 50% ACN and 0.1% TFA. For in-gel digests, crystallized material was rinsed with 1 μL of Milli-Q H_2O to remove impurities. MALDI-TOF-MS was performed on an Applied Biosystems ABI4700 TOF/TOF mass spectrometer (Applied Biosystems, Inc., Framingham, MA) with an accelerating voltage of 20 kV. Mass spectra were acquired in the MALDI reflector mode using internal calibration standards supplied with the Applied Biosystems ABI4700 CalMix. Mass spectra in the CID mode were calibrated using fragment ions from Glu-fibrinopeptide B. Tandem mass spectra were acquired using 1 kV CID with an acceleration of precursor ions at 8 kV and selection with a timed gate window of 3 Da, unless two precursors were within 3 Da of each other, in which case the gate was set to 1 Da. Gas pressure (air) in the CID cell was set at 0.2 μTorr. Fragment ions were accelerated by 14 kV into the reflector.

Data Analysis. Mascot (Matrix Science, London, U.K.) was used to interpret tandem mass spectra for protein and peptide identification (*27*). First, a nonredundant database (NCBI) was searched in order to obtain the sequences of the proteins in the neurofilament mixture. To shorten data analysis time, a custom database was created with Mascot containing the sequences of the top 17 proteins that matched those in the NFP mixture. These included, among others, myelin basic protein, glial filaments, and tubulin. The custom database included the sequences of bovine NFL, of the rod and carboxy-terminal tail domain of bovine NFM, various mammalian NFH and NFM sequences (e.g., human and pig) as well as contaminants found in the mixture. In order for a peptide to be considered a good match, it had to have characteristics corresponding to the enzyme used for digestion. Peptide masses also had to be within the mass tolerance characteristic of the mass spectrometer used for data acquisition. The mass errors for the mass spectrometers are usually within 0.1 Da for Q-TOF data and 0.3 Da for the MALDI-TOF data. For both instruments, searches within 0.4 Da were performed, but to accept a peptide as well-matched, it had to be within 0.1 or 0.2 Da depending on the instrument.

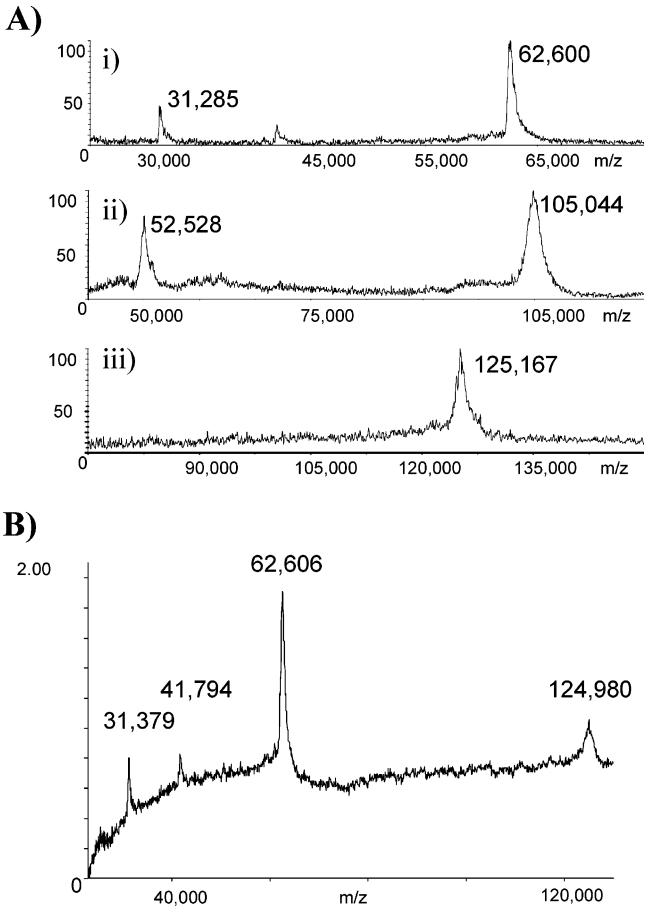

FIGURE 1: Molecular mass determination of bovine NFP subunits: (A) (i) NFL, 62 600 Da; (ii) NFM, 105 044 Da; (iii) NFH, 125 167 Da. (B) NFH, 125 167 Da. The signal at *m*/*z* 62 606 Da is due to admixture of NFL and the doubly charged ion of NFH.

Generally, peptides with individual ion scores below 30 were disregarded, except when the MS/MS spectrum was visually well-matched to the theoretical mass spectrum.

RESULTS

Molecular Weight Determination. MALDI-TOF-MS off line of HPLC-separated fractions was used to determine the molecular masses of each NFP. The molecular mass determined for bovine NFL is 62 600 Da (Figure 1Ai). The reported molecular mass of bovine NFM for a partial sequence is 90 799 Da (*29*). The mass found by mass spectrometry in the present study is 105 044 Da (Figure 1Aii). The sequence of bovine NFH has not been reported. MALDI-TOF-MS analysis in the present study showed a molecular mass of 125 167 Da (Figure 1Aiii). The mass spectrum also showed three signals with molecular masses below 100 kDa (Figure 1B). The signal at 62 kDa was more intense than the one at 125 kDa. SDS−PAGE separation of the fraction containing NFH and subsequent tryptic in-gel digestion of the two strongest bands showed that the band with the lower molecular mass is bovine NFL and the higher one bovine NFH. These results confirmed that the signal in Figure 1B with an apparent mass of 62 kDa is the superposition of the doubly charged NFH and singly charged NFL. Therefore, NFL and NFH coelute under the HPLC conditions used.

Peptide Mapping. A determination of the state of phosphorylation of the primary sequence of bovine NFP subunits was performed by MS peptide mapping under varying experimental conditions, i.e., various proteases, incubation times, in-gel and in-solution digestions, and IMAC isolation. Using the Mascot search engine, the protein sequence was confirmed and then all possible modifications for NFL and NFM other than phosphorylations were identified. These were then searched together with the phosphorylation sites. MS/MS experiments were used to obtain a higher degree of confidence in the sequence information and in the modifications of all of the specific amino acids. The data that were initially obtained by mass spectrometric analyses were used to search the database for the available sequence of NFL and the partial sequence of NFM. The N terminus of bovine NFM, which had not previously been sequenced, also was completed by MS/MS.

Verification of NFP Sequences Present in the Database. The NFP subunits were successfully separated by SDS−PAGE. In-gel tryptic digestions of NFL (Figure 2A) and NFM (Figure 2B) led to the peptide maps. In-solution digestion of the NFP mixture was also performed as it was considered to be a faster and more expeditious approach than in-gel digestions of the separated proteins. Despite the significantly greater number of peptides that resulted from digestion of the NFP mixture, no significant detrimental impact on the quality of the data was noted. Hence, in-solution digestions were performed on all further investigations.

Bovine NFL consists of 73 tryptic cleavage sites, and NFM has 153 sites including those in the N terminus. Longer digestion times resulted in greater sequence coverage in the KSP regions of the proteins, but shorter times improved the coverage in the N-terminal protein regions. The tryptic peptide maps of in-gel and in-solution digestion led to 85% coverage for bovine NFL (Figure 3) and 79% for bovine NFM (Figure 4).

Lys C in-solution digestions produced valuable additional primary structural information from NFL (Figure 3) and NFM (Figure 4); for example, the peptide $L_{314-326}$ EEIAEYRRQLQSK, ion score 36, was recorded. More importantly, in NFL, the peptide, $L_{316-331}$ (TLEIEAC^{322}RGMNEALEK) was detected with an ion score of 83. Similarly in NFM, $L_{529-556}$ (EEEEGQEEEEEEEEAAKS^{546}DQAEEGGS^{554}EK) with various modifications were detected with ion scores ranging between 34 and 39. Neither of these sequences was observed in the tryptic digests. Variation in digestion times using Lys C did not affect the coverage as greatly as it did when trypsin was used. The total Lys C peptide map gave coverage of 59% for NFL (Figure 3) and 61% for NFM (Figure 4), thereby adding about 7−8% coverage for the two proteins. More importantly, the Lys C maps verified to a great extent the tryptic peptide mapping results. Thus, whole regions were covered through the combined tryptic and Lys C maps (e.g., NFL residue 463 to 551 and in NFM residue 452 to 511).

Digestion was performed with pepsin, which although fairly nonspecific, is a protease that yields peptides quickly and reproducibly. Cleavage sites were found mainly at Leu, Glu, Ala, Phe, Tyr, Asp, and Met residues and generated yet another set of smaller, complementary peptide fragments. Although 72% of NFL (Figure 3) and 57% of NFM (Figure 4) were mapped using pepsin in-gel and in-solution digestions, this enzyme yielded valuable information in regions inaccessible with the other proteases. For example, several

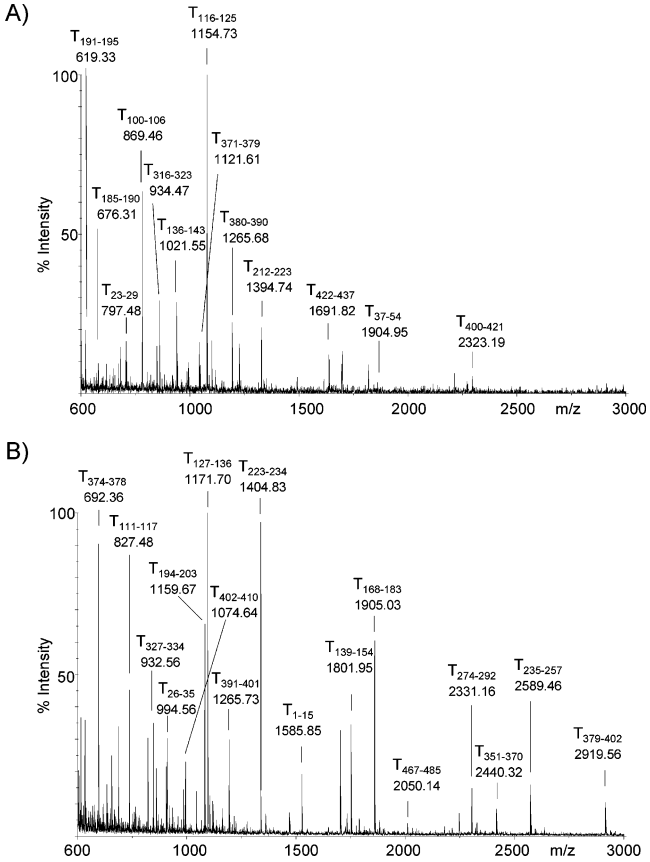

FIGURE 2: MALDI-TOF mass spectrum of tryptic peptides from an in-gel digestion: (A) bovine NFL and (B) bovine NFM. Numbering includes the sequenced N terminus.

serines that are potential phosphorylation sites were identified in NFL sequence aa 128 to 135 and in NFM sequence aa 903 to 916.

The results of tryptic peptide mapping were confirmed and greater sequence coverage was achieved by employing pepsin and Lys C digestions. The combination of enzymatic digestions and the availability of primary sequence information in the database resulted in a 99% coverage for NFL and 93% coverage for the partial sequence of NFM. These results give the data a great deal of credibility.

Sequencing of the N Terminus of NFM. Because of expected sequence homologies, the known sequences for the N termini of human, mouse, rat, and pig NFMs were included in the custom Mascot database and used to map out the N terminus of bovine NFM. To ensure that only the sequence of bovine NFM was investigated, the NFP mixture was separated by SDS−PAGE and the band for bovine NFM was subjected to in-gel digestion. Tryptic and peptic in-gel digestions were employed to map out NFM. By MS/MS with Q-TOF and TOF/TOF instruments, the N terminus of NFM was sequenced (Table 3). The ion scores for P_{51-60} and T_{54-59} are marked with an asterisk and correspond to those investigations utilizing MALDI-TOF/TOF-MS. Some fragments from the tandem MS experiments were manually excluded from the Mascot search, since Mascot does not include certain high energy fragment ions, i.e., d, v, w, x, and b + H_2O. Because of the extensive sequence homologies of these proteins from various mammalian species, the identification of the unknown N-terminal region of NFM was straightforward. Only residues 26 and 104 were in doubt as residue 104 in human, mouse, and rat NFM is Leu and in the pig it is Ile, and residue 26 is Arg in human, rat, and pig and Ser in mouse. Two peptides were identified with Ile residues from the N terminus, aa 24 to 34 (SRI^{26}SGSPSSGF) and 98 to 110 (SNEKEQI^{104}QGLNDR) d_{a3} and the d_{b3} ions for the former and d_{a7} and d_{b7} ions for the latter were observed that appeared to correspond to an Ile residue. On the basis of the observed molecular mass and high energy CID fragmentation by MALDI-TOF/TOF-MS, residue 26 was tentatively identified as Ile.

To confirm the results of in-gel mapping that lead to the sequence of the N-terminal region of NFM, data from in-solution digestions were carefully investigated. The in-solution tryptic and peptic digests strongly support the data obtained by in-gel digestion with these proteases. With Lys C, an additional peptide, L_{36-52} (SQSWSRGSPSTVSSSYK), was detected that served to complete the coverage aa 36 to 41 of the N-terminal NFM head. The overall coverage of the N-terminal region of bovine NFM provided by the proteases was 99%; trypsin yield, 64%; pepsin, 70%; and Lys C, 15%. Bovine NFM including the N-terminal region has 925 amino acids (Figure 6). The peptide coverage including the N terminus was 94%.

Identification of Modification Sites Other Than Phosphorylation. By identifying these modifications on the NFPs first, a large degree of possible confusion was eliminated before attempting to determine the phosphorylation sites. In NFM, peptides $T_{827-851}$, $T_{829-851}$, $T_{829-852}$, $T_{837-851}$, $T_{837-852}$, and $L_{837-851}$ were found in numerous different modification combinations with generally high ion scores ranging between 46 and 117 (Table 4B). All tryptic peptides also showed deamidation of Asn^{841}. Thus, this protein region would have a poorer coverage based only on peptide $L_{837-851}$ (unmodified, ion score 89) because the sequence region 827−831 and residue 852 would not have been observed. However, because of both trypsin and Lys C peptides, the entire sequence 827 to 852 with and without modifications could be analyzed. The ion scores with the additional modifications were usually high. For example, peptide $T_{829-851}$ was detected in deamidated (Asn^{841}, ion score 95) as well as in deamidated and phosphorylated forms (Asn^{841}, Thr^{850}, ion score 66). Thus, the importance of other modifications cannot be underestimated. Most of the modified peptides were observed when trypsin was used to digest the proteins rather than when Lys C or pepsin was used (Table 2). On the basis of 554 amino acids and the coverage provided by the individual enzymes, the modified residues mapped for NFL were 5.6% with trypsin, 2.1% with Lys C, and 0.5% with pepsin. NFM has 925 aa, and modifications observed were 6.2% with trypsin, 3.1% with Lys C, and 0.8% with pepsin. Peptide mapping and identification of phosphorylated residues of the neurofilaments (Table 4) were complicated by these modifications.

Overall, N-terminal serines of both NFL and NFM (Figure 6, Table 4) were found acetylated. Deamidations were observed for numerous asparagines (NFL, Asn^{80} and Asn^{154}; and NFM, Asn^{164}, Asn^{236}, Asn^{283}, Asn^{352}, Asn^{813}, and Asn^{841}), and glutamine (NFL, Gln^{74}, Gln^{157}, Gln^{234}, Gln^{236}, Gln^{238}, Gln^{241}, Gln^{332}, Gln^{340}, Gln^{348}, Gln^{448}, and Gln^{451}; and NFM, Gln^{144}, Gln^{151}, Gln^{171}, Gln^{227}, Gln^{247}, Gln^{249}, Gln^{282}, Gln^{286}, Gln^{359}, and Gln^{363}) residues. Methionine residues (NFL, Met^{64}, Met^{216}, Met^{246}, Met^{273}, Met^{347}, Met^{380}, and Met^{434}; and

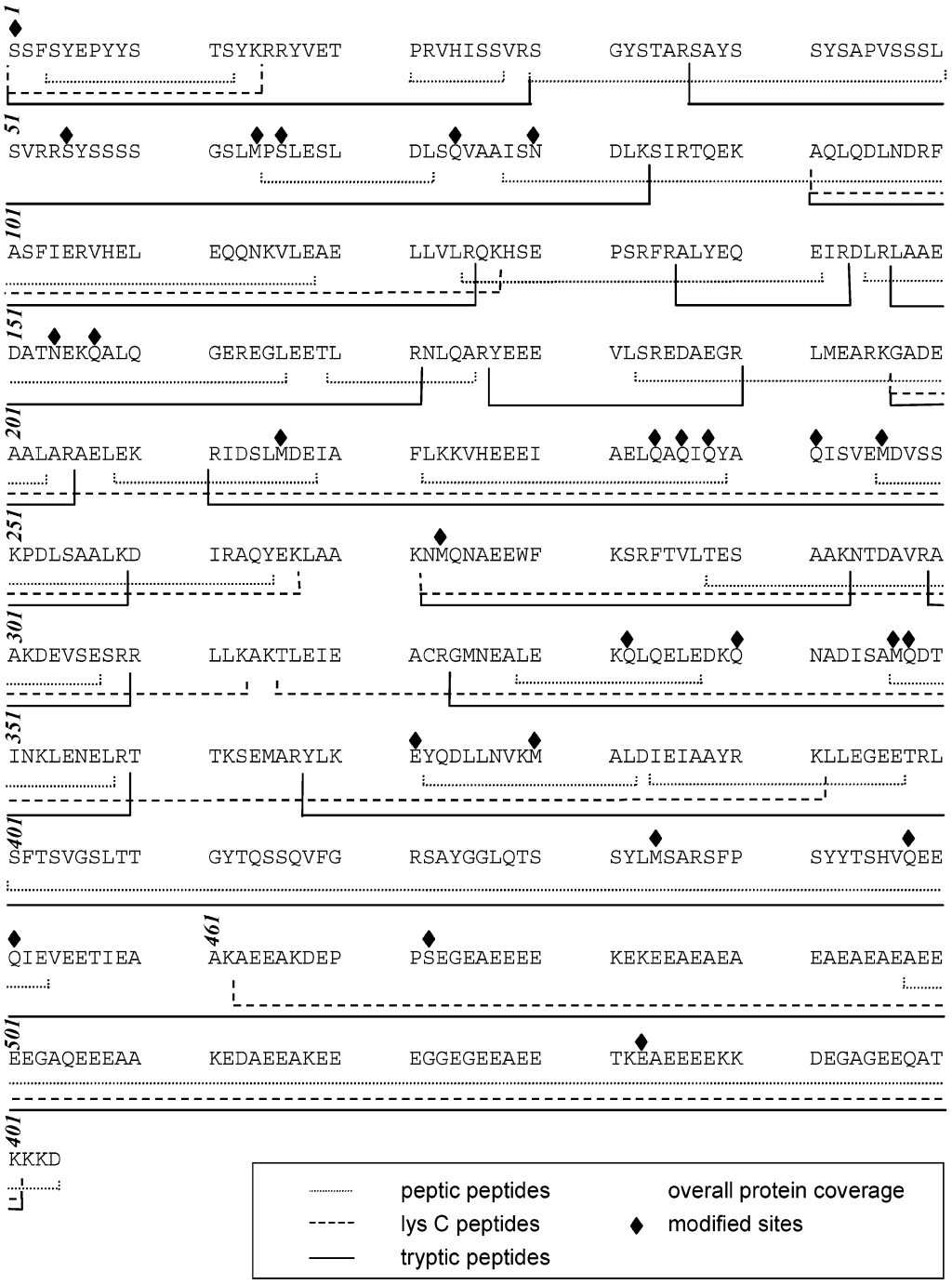

FIGURE 3: Peptide map of bovine NFL. The sequence coverage for the individual enzymes is 68% for pepsin, 59% for Lys C, and 85% for trypsin for an overall coverage of 99%. The modified sites that were identified include phosphorylation, deamidation, oxidation, N-acetylation, and pyroglutamic formation and are indicated with a diamond.

NFM, Met^{66}, Met^{284}, Met^{391}, and Met^{469}) were identified in their oxidized and nonoxidized forms. Finally, cyclization reactions resulting in pyro-glu residues that occurred during or after proteolytic digestion (NFL, Gln^{332}, Gln^{340}, Glu^{371}, and Glu^{333}; and NFM, Glu^{127}, Gln^{139}, Glu^{382}, Glu^{500}, and Glu^{557}) were also detected.

Identification of Phosphorylation Sites. An example of the collisionally activated tandem mass spectrum of the phosphorylated peptide $T_{463-481}$ (AEEAKDEPPSEGEAEEEEK) of NFL obtained from the Q-TOF instrument shows the selected parent ion with *m*/*z* 1092 Da (Figure 5 arrow) and the most intense product ion in the spectrum, which is the doubly charged $[M - 98]^+$. This is a characteristic neutral loss fragmentation of a phosphorylated residue by which phosphorylation sites are identified (*30*). A series of b-ions and y-ions, particularly the y_8 and y_{10}, showed unambiguously that residue Ser^{472} was phosphorylated. Confirmation of phosphorylation was achieved by two techniques: isolation of the phosphopeptide with a Gallium-chelated resin and neutral loss scanning using high and low energy collisional activation on the Q-TOF instrument. Using these techniques, all of the phosphorylation sites in NFL and NFM were identified.

The sequence coverage (Figure 6) and ion scores (Table 4) for the phosphopeptides were greatest for trypsin, and least for pepsin, but complementary results were obtained for

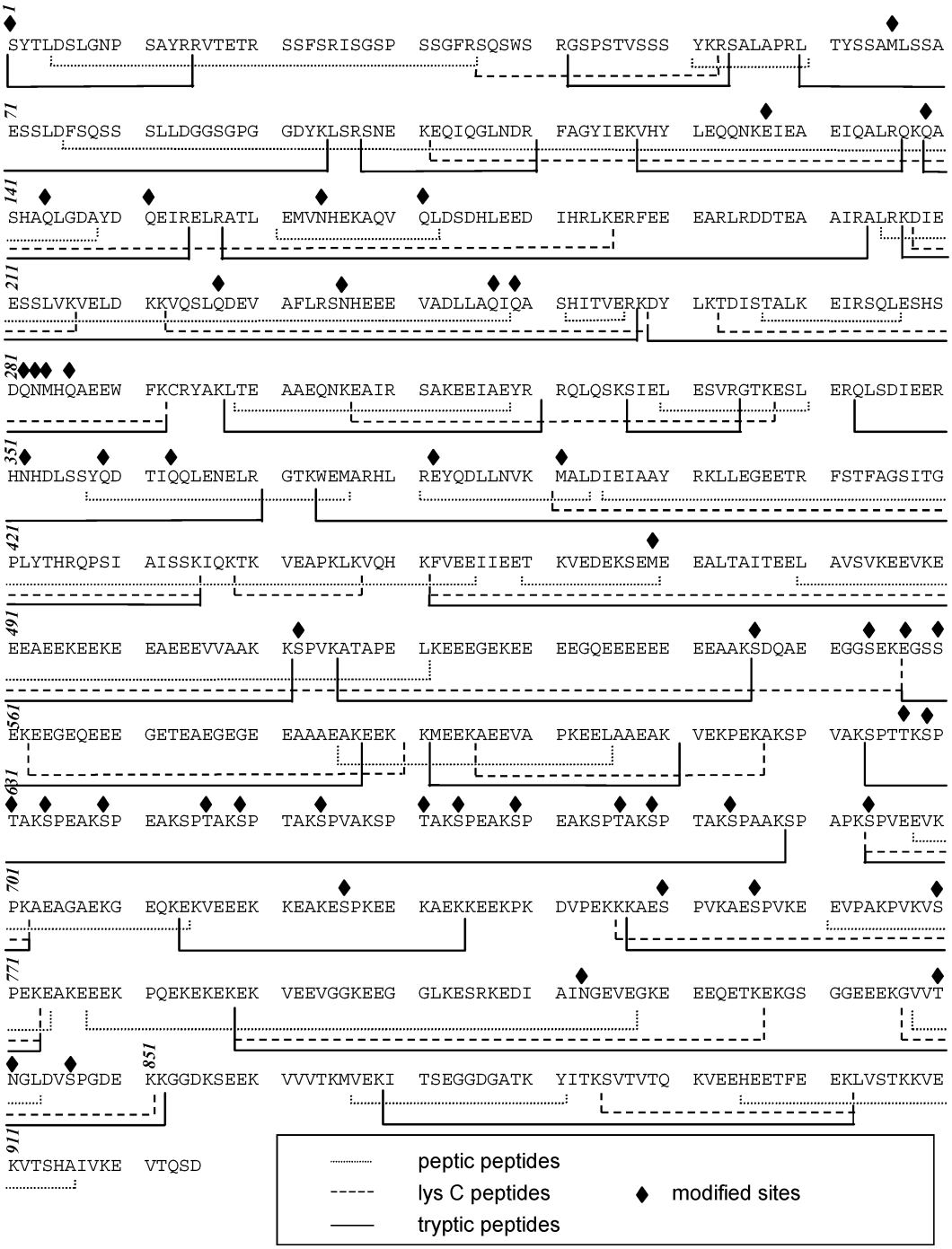

FIGURE 4: Peptide map of bovine NFM. The sequence coverage (including MS-sequenced N terminus) for the individual enzymes is 51% (53%) for pepsin, 61% (57%) for Lys C, and 80% (78%) for trypsin for an overall coverage of 93% (94%). The modified sites that were identified include phosphorylation, deamidation, oxidation, N-acetylation, and pyroglutamic formation and are indicated with a diamond.

many of the phosphorylation sites with the different enzymes. For example, phosphorylation of residue Ser^{472} in NFL was identified through both $L_{463-481}$ and $T_{463-483}$ and phosphorylation of residue Ser^{512} in NFM through $L_{500-515}$ and $P_{509-521}$. In general, the identification of a specific phosphorylation site was considered secure if the ion score was ≥ 30 and tentative if the automatic and subsequent manual data interpretation was still ambiguous because of the lack of other supportive information.

Most of the phosphopeptides were observed (Table 2) when trypsin was used to digest the proteins. Thus, in NFL, there were three sites, and in NFM, there were 21 that were identified in the tryptic digests. In the Lys C digests, one site was found in NFL and seven in NFM of which three were additional ones that were not detected in the tryptic digest. Pepsin digests revealed just one site in NFM. Thus, the total number of phosphorylation sites in NFL was three and in NFM 24.

Phosphorylation sites, Ser^{55}, Ser^{66}, and Ser^{472}, were identified in NFL (Figure 6A and Table 4A) and 22 in NFM (Ser^{512}, Ser^{546}, Ser^{554}, Ser^{560}, Thr^{627}, Ser^{629}, Ser^{634}, Ser^{639}, Thr^{646}, Ser^{649}, Ser^{654}, Ser^{664}, Ser^{669}, Thr^{676}, Ser^{679}, Ser^{684}, Ser^{694}, Ser^{726}, Ser^{750}, Ser^{756}, Ser^{770}, and Ser^{846}) with high confidence (Figure 6B and Table 4B). In modified $T_{827-851}$ and $T_{829-851}$, ion scores of 66 and 80, respectively, were highly suggestive of phosphorylation at Ser^{846}. Phosphory-

Table 2: Summary of the Peptide Mapping Results

(**A**) NFL

	trypsin	Lys C	pepsin	overall
total no. of identified aa (out of 554 possible)	468	329	376	546
total coverage (in %)	**84.5**	**59.4**	**67.9**	**98.6**
unspecific cleavage no./no. of peptides	n.o.	n.o.	35/48	35/48
total no. of modification sites[a]	26	7	2	26
no./554 (in %)	4.7	1.3	0.4	4.7
no./possible sites (in %)	5.6	2.1	0.5	4.8
total no. of phosphorylation sites	**3**[b]	**1**[b]	n.o.	**3**[b]
Ser (number of aa)	**3**	**1**	n.o.	**3**
confident				3
tentative				n.o.
homogeneous				n.o.
heterogeneous				3
Thr (number of aa)	**n.o.**	**n.o.**	**n.o.**	**n.o.**

(**B**) NFM

	trypsin	Lys C	pepsin	overall
total no. of identified aa (out of 810 possible)	646	492	410	753
total coverage for partial NFM sequence (in %)	79.3	60.7	50.6	93
total no. of identified aa (out of 925 possible)	720	523	491	868
total coverage for total NFM sequence (in %)	**77.8**	**56.5**	**53.1**	**93.8**
total no. of modification sites in total NFM sequence[c]	44	16	4	50
no./925 (in %)	4.8	1.7	0.4	5.4
no./possible sites including N term (in %)	6.2	3.1	0.8	5.9
total no. of phosphorylation sites	**21**[d]	**7**[d]	**1**	**24**
secure (+tentative)				**22 (+2)**
Ser (no. of aa)	**16**	**7**	**1**	**20**
secure				19
tentative				"0.5"
homogeneous (+tentative)				13 (+"0.5")
heterogeneous				6
Thr (no. of aa)	**5**	**n.o.**	**n.o.**	**5**
secure				3
tentative				"1.5"
homogeneous (+tentative)				3 (+"0.5")
heterogeneous (+tentative)				0 (+1)

[a] **S**[1], N-acetyl; **S**[55], phosphorylation; **M**[64], M-oxid; **S**[66], phosphorylation; **Q**[74], deamid; **N**[80], deamid; **N**[154], deamid; **Q**[157], deamid; **M**[216], M-oxid; **Q**[234], deamid; **Q**[236], deamid; **Q**[238], deamid; **Q**[241], deamid; **M**[246], M-oxid; **M**[273], M-oxid; **Q**[332], pyro-glu or deamid; **Q**[340], pyro-glu or deamid; **M**[347], M-oxid; **Q**[348], deamid; **E**[371], pyro-glu; **M**[380], M-oxid; **M**[434], M-oxid; **Q**[448], deamid; **Q**[451], deamid; **S**[472], phosphorylation; and **E**[533], pyro-glu. [b] Only specific cleavage sites were considered. [c] **S**[1], N-acetyl; **M**[66], M-oxid; **E**[127], pyro-glu; **Q**[139], pyro-glu; **Q**[144], deamid; ***Q***[151], *deamid*; ***N***[164], *deamid*; **Q**[171], deamid; **Q**[227], deamid; **N**[236], deamid; **Q**[247], deamid; **Q**[249], deamid; **Q**[282], deamid; **N**[283], deamid; **M**[284], M-oxid; **Q**[286], deamid; **N**[352], deamid; **Q**[359], deamid; **Q**[363], deamid; **E**[382], pyro-glu; **M**[391], M-oxid; **M**[469], M-oxid; ***E***[500], *pyro-glu*; ***S***[512], *phosphorylation*; ***S***[546], *phosphorylation*; ***S***[554], *phosphorylation*; **E**[557], pyro-glu; **S**[560], phosphorylation; **T**[627], phosphorylation; **S**[629], phosphorylation; **S**[634], phosphorylation; **S**[639], phosphorylation; **T**[646], phosphorylation; **S**[649], phosphorylation; **S**[654], phosphorylation; (**S**[659], phosphorylation; **T**[661], phosphorylation)[*tent*]; **S**[664], phosphorylation; **S**[669], phosphorylation; **T**[676], phosphorylation; **S**[679], phosphorylation; **S**[684], phosphorylation; **S**[694], phosphorylation; **S**[726], phosphorylation; **S**[750], phosphorylation; **S**[756], phosphorylation; **S**[770], phosphorylation; **N**[813], deamid; (**T**[840], phosphorylation)[*tent*]; **N**[841], deamid; **S**[846], phosphorylation (normal font = trypsin; italic = additional Lys C). [d] Only specific cleavage sites were considered.

lation appeared to be at Thr[840] in peptide $T_{827-851}$ because of an ion score of 46. A manual examination of the mass spectra of peptide $T_{827-851}$ indicated that either Thr[840] or Ser[846] could be phosphorylated. With the present data, it is not possible to distinguish between these two alternatives. However, we conclude that Ser[846] is the phosphorylated residue on the basis of the higher ion scores and multiple peptides showing phosphorylation at that site. Simultaneous phosphorylation at Thr[840] cannot be ruled out, however.

The identification of phosphorylated sites in NFM was most difficult in the region between residue 624 and 688. First of all, it is the most heavily phosphorylated region in NFM and all of the peptides in this region were found triply phosphorylated with ion scores between 36 and 53. Second, none of the other enzymes except trypsin provided sequence coverage in that region. Thus, no confirmatory or complementary information was available and results are based entirely on tryptic peptide maps. Third, there are segments within the KSP region with identical sequences (aa 628−655 and 663−685; Figure 6 with dotted boxes). The tryptic peptide, SPTAKSPEAKSPEAK, is triply phosphorylated. Only positions six and 11 within the sequence are unequivocally phosphorylated; the third site is located either at the first or the third residue. In the case of $T_{629-643}$ (S[629]-PTAKSPEAKSPEAK[643]), phosphorylation appears on Ser[629] and not Thr[631]. This is based on the acceptable ion score of 36 for the peptide and the associated MS/MS data that shows phosphorylation on Ser[629]. An analogous determination cannot be made for $T_{659-673}$ (S[659]PTAKSPEAKSPEAK[673]) because the sequence $T_{644-663}$, which was suggested to contain four phosphate groups, has an ion score of only 14, and thus, a distinction between phosphorylation at Ser[659] or Thr[661] cannot be made. As a matter of fact, only $T_{629-643}$ has sequence overlap with another tryptic peptide, $T_{624-638}$, within the core of the KSP region. Therefore, the assignment of phosphorylation sites is fairly secure for this peptide but not for $T_{659-673}$. In fact, for the latter, we cannot be sure that this peptide was even identified nor is there any certainty

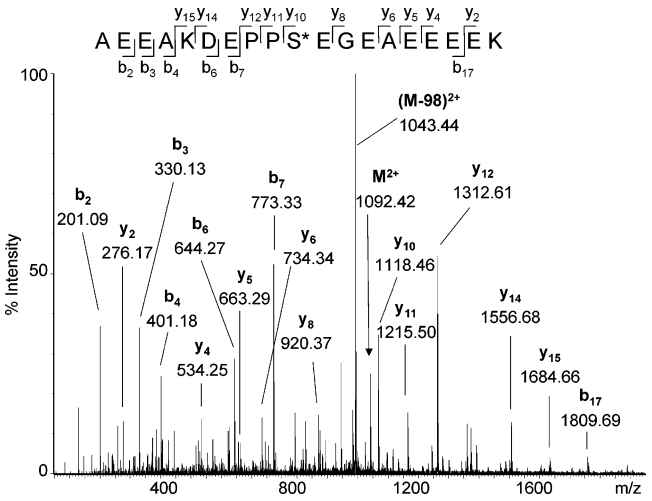

FIGURE 5: MS/MS mass spectrum of phosphopeptide $T_{463-481}$ (AEEAKDEPPSEGEAEEEEK) in bovine NFL obtained with the Q-TOF mass spectrometer. The selected parent ion has a doubly charge ion peak with *m*/*z* 1092 Da and is marked with an arrow.

that it is phosphorylated, and in particular, any possible modifications at Ser^{659} or Thr^{661} cannot be assigned. The other two phosphopeptides $T_{634-643}$ and $T_{654-673}$ that are identified in that region are only shorter peptide sequences of $T_{629-643}$ and $T_{659-673}$ and do not give any additional information to solve the uncertainty. Thus, although one might realistically conclude that since the two peptides, $T_{629-643}$ and $T_{659-673}$, have identical sequences, they might also be phosphorylated similarly, but the ion scores possible for the latter are simply too low to extract that kind of conclusion.

DISCUSSION

Mass Spectrometric Viewpoint. The molecular masses (M_r) of rat NFP subunits have been reported to be 68, 150, and 200 kDa based on SDS−PAGE analysis (*15*). On the basis of the primary sequences, the molecular masses should be 68 (*31*), 95 (*32*), and 115 kDa (*33*), respectively. These large differences in molecular masses provide indirect evidence for extensive posttranslational modifications. The off-line HPLC-MALDI-TOF-MS analysis performed in this study shows that NFL has a mass of 62 600 kDa. The theoretical nonphosphorylated average molecular mass of bovine NFL is 62 514.6 Da. The molecular mass difference of about 80 Da suggests the presence of one phosphate group on NFL. However, the mass of all of the total modifications observed amounts to about 168 Da, which would be the equivalent of two phosphate groups. Except for the N-acetylation, these modifications were all heterogeneous, and since phosphorylation was clearly indicated, it is concluded that all these modifications outside of N-acetylation and phosphorylation arose during gel separation or digestion. N-Acetylation is a common modification of proteins, and the present studies found it to be a homogeneous modification. Therefore, when the mass of the acetyl group (43 Da) is taken into consideration, it is concluded in the absence of any other modifications that there is an equivalent of one-half phosphate group on NFL, which is heterogeneously at Ser^{55}, Ser^{66}, and Ser^{472}.

The theoretical nonphosphorylated average molecular mass of the partial NFM is 90 799.3 Da, and the N terminus determined in this study is 12 297.3 Da for a total average mass of 103 078.6 Da. The measured molecular mass is 105 044 Da. If the same arguments about the modifications as used for NFL are applied to NFM, the mass difference amounts to one acetyl group and 24 phosphate groups. Indirectly, the peptide mapping results suggest that there are two additional phosphorylation sites within the KSP motifs. This conjecture is based on the observation that at least 82% of serines are phosphorylated within the KSP motifs. Unfortunately, sequence coverage inclusive of two Ser-containing sites was not available. Thus, it is concluded that overall there are 22 secure phosphorylation sites (Figure 6), two that are tentatively phosphorylated (Ser^{659} or Thr^{661} and Thr^{840}) and two others that we speculate may be phosphorylated (Ser^{619} and Ser^{689}).

The primary sequence for bovine NFH is not available, but if one extrapolates from the molecular mass of the primary sequence of mouse NFH (116 542.7110 Da) and compares this value to the experimental mass (125 167 Da) determined in this study for bovine NFH, then one would conclude that there is considerable phosphorylation. The very large difference between the M_r value and the mass determined by MALDI-TOF-MS is certainly consistent with a high degree of phosphorylation. The difference between the theoretical mass from the primary sequence and the spectrometrically determined mass suggests that there are approximately 108 phosphate groups on NFH. Other studies have also shown a high degree of phosphorylation (*9*, *14*, *25*).

The peptide mapping approach used in this investigation allowed for the unambiguous determination of three phosphorylation sites in NFL and 22 in NFM on the basis of one to three phosphate groups on any given peptide. These were all identified with high ion scores. Peptides (Table 4) were found with up to six phosphate groups per peptide but with maximum ion scores of only 17. Although the ion scores were low, some of the indicated phosphorylation sites correlated with ones that were determined by the unambiguous approaches described above. No protease resistant region was observed in these studies contrary to previous observations (*25*). Even the highly phosphoryaletd KSP region was accessible through the application of a conventional trypsin digestion protocol involving 24 h incubation periods.

Overall, the one-half equivalent phosphate group in NFL and 24 in NFM are in perfect agreement with the peptide mapping identified and characterized phosphorylation sites: three heterogeneous sites in NFL; 16 homogeneous, six heterogeneous, two tentative, and two speculative sites in NFM. Thus, this presents a comprehensive presentation of protein analysis that integrates both intact protein measurement ("top-down") and proteolytic fragment characterization ("bottom-up") mass spectrometric approaches, emphasizing the distinct capabilities of each method leading to unique biological information that will be discussed as follows.

Biological Viewpoint. *Heterogeneity of Modified Sites*. There is evidence that phosphorylation of NFPs is heterogeneous (*34*) and that the state of phosphorylation changes as the protein is transported down the axon (*35*). The results presented here (Figure 6) support these findings as some sites were found heterogeneously phosphorylated, although loss of phosphate groups during experimental workup or mass spectrometric investigations cannot be ruled out (*23*). In NFL,

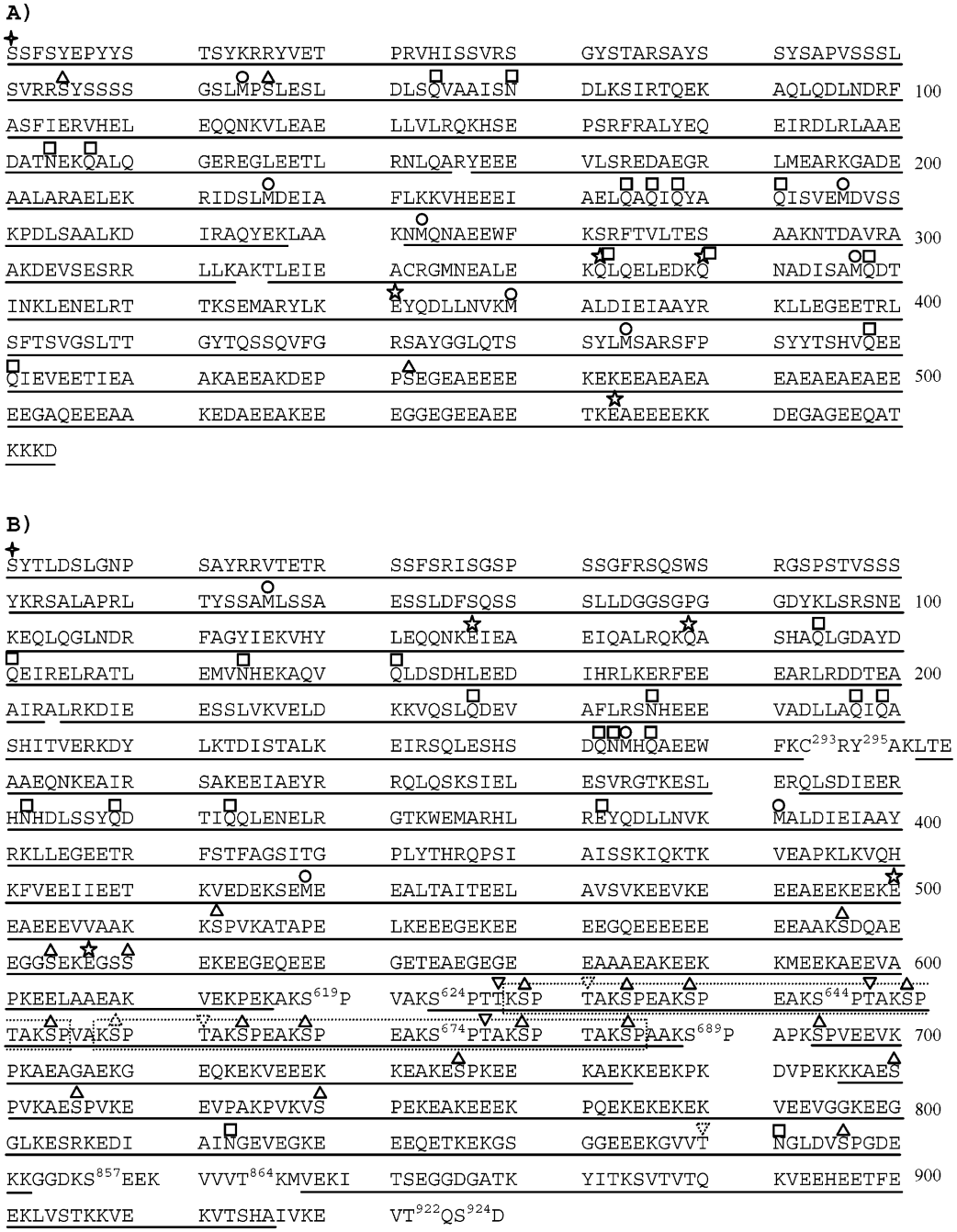

FIGURE 6: Peptide map showing modifications of (A) bovine NFL and (B) bovine NFM. Modifications are shown with the following symbols: △, phosphorylated Ser; ▽, phosphorylated Thr; ☆, pyro glutamic acid of Glu and Gln; dotted rectangle, acetylated Ser; ○, oxidized Met; □, deamidated Gln and Asp; —, overall coverage; ■, Ser motifs; dotted rectangle, identical sequences; shape with straight line, secure site; shape with dotted line, tentative site; empty shape, heterogeneous site; full shape, homogeneous site.

three phosphopeptides (Table 4A) were found heterogeneously phosphorylated. For example, peptides T_{54-83} and T_{55-83} only differ by Arg^{54}. In case of peptide T_{55-83}, two phosphorylated peptide forms were identified, where Ser^{55} (ion score 78) and Ser^{66} (ion score 54), respectively, carry the phosphate group. The same peptide T_{55-83} was not phosphorylated in the unmodified (ion score 235), deamidated (Asn^{80}, ion score 35; Gln^{74}, Asn^{80}, ion score 78), and the oxidized forms (Met^{64}, ion score 232). Similarly, peptide T_{54-83} was detected unmodified (ion score 90) and in oxidized forms (Met^{64}, ion score 102). The highest ion scores for the phosphorylated Ser^{55} (ion score 78) or Ser^{66} (ion score 54) and nonphosphorylated polypeptide (ion score 235) are convincingly high. These results suggest that the corresponding sites at Ser^{55} and Ser^{66} may not always be phosphorylated. Additional evidence for heterogeneous phosphorylation was obtained for residue Ser^{472} as both the phosphorylated (ion score 88, 83, 55, and 32) and nonphosphorylated forms (ion score 43) were found within various peptide sequences. The Lys C map confirmed these results as peptide $L_{463-483}$ was found in its unmodified (ion score 108) and phosphorylated (Ser^{472}, ion score 69) forms and peptide $L_{468-481}$ also was found unmodified (ion score 52) and phosphorylated (Ser^{472}, ion score 69).

Evidence of heterogeneous phosphorylation was also obtained for bovine NFM (Table 4B). Peptide $T_{747-768}$ was found singly phosphorylated (Ser^{750}, ion score 72) and doubly phosphorylated (Ser^{750}, Ser^{756}, ion score 140). Other tryptic

Table 3: Mapping of the N Terminus of Bovine NFM: Trypsin (In-gel, In-solution), Pepsin (In-gel, In-solution), and Lys C (In-solution); Numbers in Bold Are Sufficient for Sequence Coverage

no.	sequence	score	NFM species
T_{1-15}	**SYTLDSLGNPSAYRR**	90	human, mouse, rat
P_{5-12}	DSLGNPSA	56	human, mouse, rat
P_{13-23}	**YRRVTETRSSF**	49[a]	human
P_{24-34}	**SRISGSPSSGF**	42	not observed[b]
P_{26-35}	**ISGSPSSGFR**	31*	not observed
L_{36-52}	**SQSWSRGSPSTVSSSYK**	74	human, rat, pig
T_{42-52}	GSPSTVSSSYK	82	human, rat, pig
T_{42-53}	GSPSTVSSSYKR	74	human, rat, pig
P_{51-60}	**YKRSALAPRL**	31[a]	rat, pig
T_{54-59}	SALAPR	29[a]	rat, pig
T_{60-94}	**LTYSSAMLSSAESSLDFSQ-SSSLLDGGSGPGGDYK**	130	pig
P_{76-92}	FSQSSSLLDGGSGPGGD	55	pig
P_{83-92}	LDGGSGPGGD	60	pig
P_{93-114}	**YKLSRSNEKEQIQ-GLNDRFAGY**	38	pig
T_{98-110}	SNEKEQIQGLNDR	82	pig
{**T_{98-110}**	SNEKEQLQGLNDR	82	human, rat}[c]
$P_{101-109}$	KEQIQGLND	71	pig
$T_{102-110}$	EQIQGLNDR	53	pig
{**$T_{102-110}$**	EQLQGLNDR	53	human, rat}[c]
$L_{102-117}$	EQIQGLNDRFAGYIEK	86	pig
$T_{111-117}$	FAGYIEK	60	human, mouse, rat pig
$P_{115-129}$	**IEKVHYLEQQNKEIE**	55	human, mouse, rat, pig

[a] The ion scores of peptides found with MALDI-TOF-TOF are marked with an asterisk. [b] Error tolerant search. [c] Peptide sequence excluded by MALDI-TOF-TOF data because of side chain fragmentation of Ile.

peptides confirm the singly phosphorylated Ser^{750} form ($T_{747-759}$, ion score 55; $T_{748-759}$, ion score 27) as well as the doubly phosphorylated form ($T_{748-768}$, ion score 94). The phosphopeptides generated by Lys C were doubly phosphorylated ($L_{746-768}$, ion score 57; $L_{746-773}$, ion score 60; $L_{747-768}$, ion score 56; $L_{747-773}$, ion score 43). Thus, Ser^{750} is detected in all of the various phosphopeptides and was always found phosphorylated, whereas Ser^{756} was only sometimes phosphorylated. It is concluded that Ser^{750} is a homogeneous and Ser^{756} is a heterogeneous phosphorylation site in bovine NFM. Overall, including the two tentative sites (Table 2), 17 of the 24 phosphopeptides were homogeneous phosphorylation sites [Ser^{512}, Thr^{627}, Ser^{629}, Ser^{634}, Ser^{639}, Thr^{646}, Ser^{649}, Ser^{654}, (Ser^{659}, Thr^{661}), Ser^{664}, Ser^{669}, Thr^{676}, Ser^{679}, Ser^{684}, Ser^{694}, Ser^{726}, and Ser^{750}] and seven of them were heterogeneous [Ser^{546}, Ser^{554}, Ser^{560}, Ser^{756}, Ser^{770}, (Thr^{840}), and Ser^{846}]. Thus, all of the phosphorylated sites in NFM within KSP motifs are homogeneous modifications with the exception of Ser^{756} and Ser^{770}.

The state of phosphorylation of NFP subunits is different in different developmental stages as well as in different cellular components of animals. Hence, phosphorylation sites may differ in soluble and assembled NFP subunits. It must be noted that these investigations were carried out on purchased spinal cord bovine NFP mixture. Therefore, one must assume that the more abundant, assembled, axonal forms in adult cows were predominant, and the results should be extrapolated with caution when considering soluble NFP subunits.

Motifs of Phosphorylation Sites. The dynamic behavior of these cytoskeletal proteins depends on their state of phosphorylation (*25*, *36*−*40*). The function of phosphorylation on the Glu-rich subdomains in the N-terminal portion of the tail domains (E segments) of all NFP subunits is not understood yet (*41*−*44*). In vivo, it is likely that several different kinases may phosphorylate KSP motifs simultaneously. These motifs are only found in bovine NFM but not in bovine NFL (Figure 6). Phosphorylation of the head domain of NFL (*45*−*48*) is believed to be important for the regulation of the NFP assembly in vivo (*49*). Various different sites in this region have been shown to be phosphorylated in different mammals. For example, of the possible sites Ser^{41}, Ser^{55}, and Ser^{62} in the head region of rat NFL, in vitro experiments with cyclic AMP-dependent protein kinase (PKA) suggest that Ser^{55} is the preferred phosphorylation site (*50*). It was also shown (*51*) that PKA can stimulate in vitro phosphorylation of Ser^{2} in rat NFL. Nevertheless, in various mammals, the most common phosphorylation site in the head domain of NFL is the site corresponding to the bovine residue, Ser^{55}, which in the present study was unambiguously identified as an in vivo phosphorylation site of bovine NFL. The results do not support in vivo phosphorylation of Ser^{2} in NFL. In vivo or in vitro phosphorylation of Ser^{66} found in the present study for bovine NFL has not been reported previously for any other mammalian species. In the C-terminal tail domain of rat NFL, it has been shown that residue Ser^{473} is the major phosphorylation site in vivo (*9*, *52*). Residue Ser^{473} in rat NFL corresponds to Ser^{472} in bovine NFL. The present results on the in vivo phosphorylation of bovine NFL of residue Ser^{472} strongly support this finding. Overall, two of the three identified phosphorylation sites in bovine NFL are known phosphorylation sites in other mammalian species.

Phosphorylation of rat (*9*) and porcine (*22*) NFM occurs mainly on Ser residues in the KSP motifs within the C-terminal domains. The types, number, and distribution of these KSP repeats differ among species (*53*). The KSP multiple repeats in tail domains constitute various candidates for phosphorylation (*9*). The phosphorylated KSP motifs in bovine NFM identified in this study (Figure 6) are [experimentally identified/possible phosphorylation site]: KSP [11-(+1) out of 17; 65(+71)%]; KXSP [3 out of 3; 100%] with X = K, E, V; KXXSP [2 out of 3; 67%] with XX = AK, AE; KSD [1 out of 1; 100%]; VSP [2 out of 2; 100%]; and SEK [2 out of 2; 100%]. $KS^{659}P$ is a tentative phosphorylation site. Three Ser residues (Ser^{624}, Ser^{644}, and Ser^{674}) within the 17 KSP motifs were found not to be phosphorylated. Two Ser residues (Ser^{619} and Ser^{689}) within and two (Ser^{857} and Ser^{924}) outside of the KSP motifs were not within the coverage of the combination of enzymes (trypsin, Lys C, and pepsin) in the peptide maps. Thus, because of the fact that 14 of 17 serines (82%) within the KSP motifs are shown to be phosphorylated, it is reasonable to assume that there might be two additional sites that are phosphorylated within the sequence not covered by peptide mapping. There were no Ser residues outside the KSP motifs of NFM that were found phosphorylated. Thus, two serines in the protein for which there was not sequence coverage are assumed not to be phosphorylated. The results for the overall number of phosphorylation sites in KSP motifs (21 out of 24; 88%) are in good agreement with previous investigations in which the overwhelming evidence indicates that all Ser residues within the KSP motifs are potentially phosphorylated. Nearly

Table 4: Modified Peptides Identified in Bovine NFL and NFM with Corresponding Ion Scores from Mascot Searches

(**A**) NFL

no.	peptide sequence in bovine NFL	modifications (aa, ions score[a])
T$_{1-14}$	**S**^{Ho}SFSYEPYYSTSYK	N-acetyl (**S**^{1}, 77)
L$_{1-14}$	**S**^{Ho}SFSYEPYYSTSYK	N-acetyl (**S**^{1}, 68)
T$_{1-15}$	**S**^{H}SFSYEPYYSTSYKR	N-acetyl (**S**^{1,conf}, **79**)
T$_{54-83}$	RS$^{(He)}$YSSSSGSL**M**^{He}PS$^{(He)}$LESLDLSQVAAISNDLK	unmodif (90); M-oxid (**M**^{64}, 102)
T$_{55-83}$	**S**^{He}YSSSSGSL**M**^{He}P**S**^{He}LESLDLS**Q**^{He}VAAISN^{He}DLK	unmodif (235); phospho (**S**^{55,conf}, 78); M-oxid (**M**^{64,conf}, **232**); phospho (**S**^{66,conf}, **54**); 2 deamid (**Q**^{74,conf}, **N**^{80,conf}, **78**); deamid (**N**^{80}, 35)
T$_{147-156}$	LAAEDATN$^{(He)}$EK	unmodif (78)
T$_{147-163}$	LAAEDATN^{He}EK**Q**^{He}ALQGER	deamid (**N**^{154,conf}, **93**); deamid (**Q**^{157,conf}, **125**)
P$_{145-166}$	LRLAAEDATNEKQ$^{(He)}$ALQGEREGL	unmodif (**32**)
T$_{211-223}$	RIDSL**M**^{He}DEIAFLK	unmodif (83); M-oxid (**M**^{216}, 38)
L$_{211-223}$	RIDSL**M**^{He}DEIAFLK	M-oxid (**M**^{216}, 37)
T$_{211-224}$	RIDSL**M**^{He}DEIAFLKK	unmodif (75); M-oxid (**M**^{216}, 52)
T$_{212-223}$	IDSL**M**^{He}DEIAFLK	unmodif (100); M-oxid (**M**^{216,conf}, **79**)
T$_{212-224}$	IDSL**M**^{He}DEIAFLKK	unmodif (101); M-oxid (**M**^{216}, 77)
T$_{225-259}$	VHEEEIAEL**Q**^{He}A**Q**^{He}I**Q**^{He}YA**Q**^{He}ISVE**M**^{He}DVSSKPDLSAALK	unmodif (247); deamid (**Q**^{234,conf}, **89**); deamid (**Q**^{236,conf}, **65**); deamid (**Q**^{238,conf}, **55**); deamid (**Q**^{241,conf}, **84**); M-oxid (**M**^{246,conf}, **56**)
P$_{246-265}$	**M**^{He}DVSSKPDLSAALKDIRAQY	unmodif (56); M-oxid (M^{246}, 45)
T$_{272-281}$	N**M**^{He}QNAEEWFK	unmodif (53); M-oxid (**M**^{273,conf}, **58**)
L$_{272-281}$	N**M**^{He}QNAEEWFK	unmodif (65); M-oxid (**M**^{273}, 48)
P$_{329-337}$	LEK**Q**^{He}LQELE	deamid (Q^{332,conf}, 58)
L$_{332-339}$	**Q**^{He}LQELEDK	unmodif (59); pyro-glu (**Q**^{332}, 40)
T$_{332-353}$	Q$^{(He)}$LQELEDK**Q**^{He}NADISA**M**^{He}QDTINK	unmod (119); deamid (**Q**^{340}, 83); M-oxid (**M**^{347}, 40)
T$_{332-359}$	**Q**^{He}LQELEDK**Q**^{He}NADISA**M**^{He}**Q**^{He}DTINKLENELR	unmodif (91); pyro-glu (**Q**^{332,conf}, **96**); pyro-glu, M-oxid (**Q**^{332}, **M**^{347}, 57); 2 deamid (**Q**^{340,conf}, **Q**^{348,conf}, 106); M-oxid (**M**^{347}, 55)
T$_{340-353}$	QNADISA**M**^{He}QDTINK	unmodif (101); M-oxid [**M**^{347,conf}, **75**)
T$_{340-359}$	QNADISA**M**^{He}QDTINKLENELR	unmodif (178); M-oxid (**M**^{347}, 45)
L$_{340-362}$	**Q**^{He}NADISAM**Q**^{He}DTINKLENELRTTK	unmodif (76); pyro-glu (**Q**^{340,conf}, 61); deamid (**Q**^{348}, 66)
T$_{368-379}$	YLKE$^{(He)}$YQDLLNVK	unmodif (93)
T$_{371-379}$	**E**^{He}YQDLLNVK	unmodif (56); pyro-glu (**E**^{371,conf}, **78**)
T$_{380-390}$	**M**^{He}ALDIEIAAYR	unmodif (93); M-oxid (**M**^{380,conf}, **85**)
T$_{422-437}$	SAYGGLQTSSYL**M**^{He}SAR	unmodif (120); M-oxid (**M**^{434,conf}, **56**)
T$_{438-462}$	SFPSYYTSHV**Q**^{He}EE**Q**^{He}IEVEETIEAAK	unmodif (115); deamid (**Q**^{448,conf}, **72**); 2 deamid (**Q**^{448}, **Q**^{451}, 51); deamid (**Q**^{451,conf}, 61)
T$_{463-481}$	AEEAKDEPP**S**^{He}EGEAEEEEK	unmodif (43); phospho (**S**^{472}, **83**)
L$_{463-481}$	AEEAKDEPP**S**^{He}EGEAEEEEK	unmodif (108); phospho (**S**^{472}, **69**)
T$_{463-483}$	AEEAKDEPP**S**^{He}EGEAEEEEKEK	phospho (**S**^{472,conf}, **88**)
L$_{468-481}$	DEPP**S**^{He}EGEAEEEEK	unmodif (52); phospho (**S**^{472}, 64)
T$_{468-481}$	DEPP**S**^{He}EGEAEEEEK	phospho (**S**^{472}, 55)
T$_{468-483}$	DEPP**S**^{He}EGEAEEEEKEK	phospho (**S**^{472}, 32)
T$_{533-551}$	**E**^{He}AEEEEKKDEGAGEEQATK	unmodif (151); pyro-glu (**E**^{533,conf}, **91**)

(**B**) NFM

no.	peptide sequence in bovine NFM	modifications (aa: ions score[a])
T$_{1-14}$	**S**^{Ho}YTLDSLGNPSAYR	N-acetyl (**S**^{1,conf}, **90**)
{**T**$_{1-14}$	(**S**)YTLDSLGNPSAYRR	N-acetyl (S^{1}, 28)}[e]
{**P**$_{1-12}$	(**S**)YTLDSLGNPSA	N-acetyl (S^{1}, 16)}[c]
T$_{60-94}$	LTYSSA**M**^{He}LSSAESSLDFSQSSSLLDGGSGPGGDYK	unmodif (130); M-oxid (**M**^{66,conf}, **94**)
T$_{127-136}$	**E**^{He}IEAEIQALR	unmodif (76); pyro-glu (**E**^{127,conf}, **55**)
T$_{139-154}$	**Q**^{He}ASHA**Q**^{He}LGDAYDQEIR	unmodif (129); pyro-glu (**Q**^{139,conf}, **99**); deamid (**Q**^{144,conf}, **83**)
L$_{139-167}$	**Q**^{He}ASHAQLGDAYD**Q**^{He}EIRELRATLE(**M**)V**N**^{He}HEK	unmodif (59); pyro-glu (**Q**^{139}, 34); 2 deamid (**Q**^{151,conf}, **N**^{164,conf}, 53); {M-oxid (M^{162}, 26)}[e]
T$_{168-183}$	AQV**Q**^{He}LDSDHLEEDIHR	unmodif (143); deamid (**Q**^{171,conf}, **95**)
T$_{223-234}$	VQSL**Q**^{He}DEVAFLR	unmodif (81); deamid (**Q**^{227,conf}, **53**)
T$_{235-257}$	SN^{He}HEEEVADLLA**Q**^{He}I**Q**^{He}ASHITVER	unmodif (199); 3 deamid (**N**^{236,conf}, **Q**^{247}, **Q**^{249,conf}, **72**); deamid (**Q**^{247,conf}, 134)
P$_{245-248}$	LA**Q**^{He}I	deamid (Q^{247}, **32**)
L$_{271-292}$	EIRSQLESHSD**Q**^{He}**N**^{He}MHQAEEWFK	unmodif (50); 2 deamid (**Q**^{282}, **N**^{283}, **62**)

Table 4: Continued

no.	peptide sequence in bovine NFM	modifications (aa: ions score[a])
T$_{274-292}$	SQLESHSD**Q**^{He}**N**^{He}**M**^{He}H**Q**^{He}AEEWFK	unmodif (133); 2 deamid (**Q**^{282,conf}, **N**^{283,conf}, **113**); 2 deamid (**Q**^{282}, **Q**^{286,conf}, 70); deamid (N^{283}, 58); deamid, M-oxid (**Q**^{283}, **M**^{284}, 39); M-oxid (**M**^{284,conf}, **51**)
T$_{351-370}$	H**N**^{He}HDLSSY**Q**^{He}DTI**Q**^{He}QLENELR	unmodif (137); deamid (**N**^{352,conf}, **54**); deamid (**Q**^{359,conf}, **120**); deamid (**Q**^{363,conf}, **106**)
T$_{382-390}$	**E**^{He}YQDLLNVK	unmodif (56); pyro-glu (**E**^{382,conf}, **78**)
T$_{391-401}$	**M**^{He}ALDIEIAAYR	unmodif (93); M-oxid (**M**^{391,conf}, **85**)
T$_{462-485}$	VEDEKSE**M**^{He}EEALTAITEELAVSVK	unmodif (191); M-oxid (**M**^{469}, 59)
T$_{467-485}$	SE**M**^{He}EEALTAITEELAVSVK	unmodif (135); M-oxid (**M**^{469,conf}, **132**)
L$_{467-496}$	SE**M**^{He}EEALTAITEELAVSVKEEVKEEEAEEK	unmodif (105); M-oxid (**M**^{469}, 83)
L$_{500-510}$	**E**^{He}EAEEEVVAAK	unmodif (76); pyro-glu (**E**^{500,conf}, **34**)
L$_{500-515}$	EEAEEEVVAAKK**S**^{Ho}PVK	phospho (**S**^{512,conf}, **64**)
P$_{495-519}$	EKEEKEEAEEEVVAAKK**S**^{Ho}PVKA(T)AP	phospho (S^{512}, (T^{517})^{w}, **38**)
P$_{509-521}$	AKK**S**^{Ho}PVKATAPEL	phospho (S^{512}, **60**)^{i,k}
L$_{529-556}$	EEEEGQEEEEEEEEAAKS**S**^{He}DQAEEGGS**S**^{He}EK	phospho (**S**^{546,conf}, **39**); phospho (**S**^{554,conf}, **34**)
{**L**$_{546-556}$	(**S**)D(**Q**)AEEGG(**S**)EK	deamid, 2 phospho (S^{546}, Q^{548}, S^{554}, 14)}^{c}
{**L**$_{546-562}$	(**SD**)QAEEGG(**S**)EKEGS**S**^{He}EK	methylester, 3 phospho (S^{546}, D^{547}, S^{554}, **S**^{560,conf}, 4)}^{a,c,f}
T$_{557-587}$	**E**^{He}G(**SS**)^{He}EKEEGEQEEEGETEAEGEGEEAAAEAK	unmodif (111); pyro-glu, phospho (**E**^{557,conf}, (**S**^{559} or **S**^{560})^{d}, **118**); phospho (**S**^{560}, 102)
T$_{624-638}$	SPT**T**^{Ho}K**S**$^{Ho,(He)}$P(T^{He})AK**S**^{Ho}PEAK	3 phospho (**T**^{627,conf}, **S**^{629,conf}, **S**^{634,conf}, **36**)
T$_{629-643}$	**S**$^{Ho,(He)}$P(T^{He})AK**S**^{Ho}PEAK**S**^{Ho}PEAK	3 phospho (**S**^{629}, (T^{631,tent})^{r}, **S**^{634}, **S**^{639}, **45**)
T$_{634-643}$	**S**^{Ho}PEAK**S**^{Ho}PEAK	2 phospho (**S**^{634,conf}, **S**^{639,conf}, **53**)
T$_{644-658}$	SP**T**^{Ho}AK**S**^{Ho}PTAK**S**^{Ho}PVAK	3 phospho (**T**^{646,conf}, **S**^{649,conf}, **S**^{654,conf}, **36**)
{**T**$_{644-663}$	SP(**T**)AKSP(**T**)AK(**S**)PVAK(**S**)PTAK	4 phospho (T^{646}, T^{651}, S^{654}, S^{659}, 14)}^{e}
T$_{659-673}$	(**ST**)^{Ho}AK**S**^{Ho}PEAK**S**^{Ho}PEAK	3 phospho ((S^{659,tent}, T^{661,tent}), **S**^{664}, **S**^{669}, **45**)
T$_{664-673}$	**S**^{Ho}PEAK**S**^{Ho}PEAK	2 phospho (**S**^{664,conf}, **S**^{669,conf}, **53**)
T$_{674-688}$	SP**T**^{Ho}AK**S**^{Ho}PTAK**S**^{Ho}PAAK	3 phospho (**T**^{676,conf}, **S**^{679,conf}, **S**^{684,conf}, **39**)
{**T**$_{674-702}$	SPTAK(**S**)P(**T**)AKSPAAK(**S**)PAPK(**S**)PVEEVKPK	4 phospho (S^{679}, T^{681}, S^{689}, S^{694}, 5)}^{e}
T$_{694-702}$	**S**^{Ho}PVEEVKPK	phospho (**S**^{694,conf}, **69**)
L$_{694-702}$	**S**^{Ho}PVEEVKPK	phospho (**S**^{694}, 51)
T$_{714-734}$	EKVEEEKKEAKE**S**^{Ho}PKEEKAEK	phospho (**S**^{726,conf}, **49**)
L$_{746-768}$	KKAE**S**^{Ho}PVKAE**S**^{He}PVKEEVPAKPVK	2 pospho (**S**^{750}, **S**^{756}, 57)
L$_{746-773}$	KKAE**S**^{Ho}PVKAE**S**^{He}PVKEEVPAKPVKVS$^{(He)}$PEK	2 pospho (**S**^{750}, **S**^{756}, 60)
T$_{747-759}$	KAE**S**^{Ho}PVKAES$^{(He)}$PVK	phospho (**S**^{750}, 55)
T$_{747-768}$	KAE**S**^{Ho}PVKAE**S**^{He}PVKEEVPAKPVK	phospho (**S**^{750}, 72); 2 phospho (**S**^{750,conf}, **S**^{756,conf}, **140**)
L$_{747-768}$	KAE**S**^{Ho}PVKAE**S**^{He}PVKEEVPAKPVK	2 pospho (**S**^{750}, **S**^{756}, 56)
L$_{747-773}$	KAE**S**^{Ho}PVKAE**S**^{He}PVKEEVPAKPVKVS$^{(He)}$PEK	2 pospho (**S**^{750}, **S**^{756}, 43)
{**T**$_{748-759}$	AE(**S**)PVKAES$^{(He)}$PVK	phospho (S^{750}, 27)}^{e}
T$_{748-768}$	AE**S**^{Ho}PVKAE**S**^{He}PVKEEVPAKPVK	2 phospho (**S**^{750}, **S**^{756}, 94)
T$_{754-768}$	AE**S**^{He}PVKEEVPAKPVK	phospho (**S**^{756}, 74)
T$_{754-773}$	AE**S**^{He}PVKEEVPAKPVKV**S**^{He}PEK	2 phospho (**S**^{756}, **S**^{770,conf}, **57**)
L$_{754-768}$	AE**S**^{He}PVKEEVPAKPVK	phospho (**S**^{756}, 54)
L$_{754-773}$	AE**S**^{He}PVKEEVPAKPVKVS$^{(He)}$PEK	phospho (**S**^{756}, 33)
L$_{804-819}$	ESRKEDIAI**N**^{He}GEVEGK	unmodif (49); deamid (**N**^{813}, 90)
L$_{804-826}$	ESRKEDIAI**N**^{He}GEVEGKEEEQETK	deamid (**N**^{813}, 70)
T$_{807-819}$	KEDIAI**N**^{He}GEVEGK	deamid (**N**^{813}, 87)
T$_{807-826}$	KEDIAI**N**^{He}GEVEGKEEEQETK	deamid (**N**^{813,conf}, **159**)
T$_{807-828}$	KEDIAI**N**^{He}GEVEGKEEEQETKEK	deamid (**N**^{813}, 34)
T$_{827-851}$	EKGSGGEEEKGVV**T**^{He}**N**^{He}GLDV**S**^{He}PGDEK	deamid (**N**^{841}, 74); phospho, deamid (**T**^{840,conf}, **N**^{841}, **46**); deamid, phospho (**N**^{841}, **S**^{846}, 80)
T$_{829-851}$	GSGGEEEKGVVT$^{(He)}$**N**^{He}GLDV**S**^{He}PGDEK	deamid (**N**^{841}, 95); deamid, phospho (**N**^{841}, **S**^{846}, 66)
T$_{829-852}$	GSGGEEEKGVVT$^{(He)}$**N**^{He}GLDVS$^{(He)}$PGDEKK	deamid (**N**^{841}, 77)
T$_{837-851}$	GVVT$^{(He)}$**N**^{He}GLDV**S**^{He}PGDEK	deamid (**N**^{841}, 77); deamid, phospho (**N**^{841}, **S**^{846}, 69)
T$_{837-852}$	GVVT$^{(He)}$**N**^{He}GLDV**S**^{He}PGDEKK	deamid (**N**^{841}, 97); deamid, phospho (**N**^{841}, **S**^{846}, 106)
L$_{837-851}$	GVVT$^{(He)}$**N**^{He}GLDV**S**^{He}PGDEK	unmodified (89); deamid (**N**^{841}, 109); deamid, phospho (**N**^{841,conf}, **S**^{846,conf}, **117**)

[a] Bold numbers indicate the highest score for the individual modification irrespective of peptide length and its origin. [b] Secure phosphorylation site based on the highest ion score (>30) found for the individual site irrespective of the enzymatic cleavage. [c] Based on mass spectra. [d] Ser^{560} is secure due to L$_{546-562}$. [e] Because of the low ions, the score was not taken into account for coverage but discussed for different reasons. ^{Ho} Homogeneous modified site. ^{He} Heterogeneous modified site. [w] Less presumable because of other observed modified peptides. [r] Ruled out because of T$_{624-638}$.

all Ser residues in the KSP and KXXSP motifs in in vivo chicken NFM appear to be phosphorylated (*23*). Eight of nine phosphopeptides of porcine NFM carried various KSP as well as variant KAESP and VSP motifs (*22*). Only four phosphorylated sites were determined unambiguously, three within those motifs (*22*). The KSD and SEK motifs of bovine

NFM (aa 546, 554, and 560) show high homology with chicken, human, rat, and mouse NFM, all of which were also found phosphorylated (*44*). The KSD motif is present in the acidic N-terminal side of the chicken NFM tail and can be phosphorylated by a casein kinase I-like NFP-associated kinase (*23*). Residue Ser^{23} is located in a non-KSP motif in the N-terminal head domain of mouse NFM and is reported to be a major protein kinase A phosphorylation site in vitro (*54*). In the present investigation on in vivo bovine NFM, Ser^{23} was found not to be phosphorylated.

The present results reveal phosphorylation of Ser^{846}, conforming to phosphorylated Ser^{787} in chicken NFM (*23*), in the far C-terminal region of the NFM tail domain. Early investigations of porcine NFM also identified phosphopeptide GVVTDGLDSPGDRSEEK, but the exact site of phosphorylation was not determined (*22*). The tail domains are thought to protrude as sidearms in assembled NFPs that have been proposed to interact with other NFPs and axonal cytoskeletal proteins. These interactions have been shown to be regulated by phosphorylation (*55*). Phosphorylation of the tail domains is also thought to increase spacing between filaments, hence regulating axonal caliber and, additionally, may play a role in the axonal transport (*56*). This Ser-Pro site at Ser^{846} does not conform to known NF phosphorylation motifs. Nevertheless, it has been reported for chicken Ser^{787} (*15*, *23*), mouse, and human NFM, respectively, in a region of the tail that is highly conserved (*57*, *58*). Because various species show identical phosphorylation sites, it suggests that this an important transduction site because it is a well-preserved region and obviously evolutionary important.

Phosphorylation of Thr residues is generally less likely than phosphorylation of serines. In the present study, four phosphorylated Thr residues (Thr^{627}, Thr^{631}, Thr^{646}, and Thr^{661}) were found, all of them in the highly Ser-phosphorylated KSP region of NFM. Thus, there is solid evidence for Thr phosphorylation in TAK motifs. There are potentially other phosphorylation sites in mammalian NFPs. For example, a SAK motif is described for squid NFH as being phosphorylated at Ser in the C-terminal repeat segment (*14*). Recently, eight Thr residues were described as being phosphorylated in chicken NFM, where seven of them occur within a PXS/TP motif (*23*). In bovine NFM studied here, 20 out of 92 Ser residues (about 22%), four out of 44 Thr residues (about 9%), and zero out of 16 Tyr residues have been found phosphorylated. Coverage for four Ser, two Thr, and one Tyr residues has not yet been achieved through the combined peptide mapping results with trypsin, Lys C, and pepsin digests. Therefore, it is assumed that the degree of phosphorylation (2.6% of all aa) of bovine NFM could be slightly greater than the present results indicate.

In summary, 20 phosphorylation sites were identified within the KSP and variant KSP motifs in bovine NFM. No serines outside the KSP motifs were found to be phosphorylated. Thus, these KSP motifs play a major role in the phosphorylation of bovine NFP. Because there are six different KSP motifs phosphorylated in bovine NFM, it is concluded that various protein kinases are involved in phosphorylation of bovine NFP. For example, SEK and KSD motifs are targeted by casein kinase I and the KSP motif by Pro-dependent protein kinase (for review, see ref *25*). Analogous conclusions have to be made for phosphorylation of Thr where three of the four phosphorylated sites are in the TAK motif. In these investigations, only three phosphorylation sites are found in bovine NFL, all in non-KSP motifs. Thus, different kinases have to be considered for bovine NFL, such as, for example, cyclic AMP-dependent protein kinase, as it was found to be important for rat NFL (*50*, *51*).

ACKNOWLEDGMENT

We thank Brian Arbogast and Elisabeth Barofsky for technical assistance in mass spectrometry.

SUPPORTING INFORMATION AVAILABLE

(1) Extensive figures: peptide maps of trypsin, Lys C, and pepsin and (2) large tables: peptide maps of trypsin, Lys C, and pepsin. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Friede, R. L., and Samorajaski, T. (1970) *Anat. Rec. 167*, 379−388.
2. Hoffman, P. N., Cleveland, D. W., Griffin, J. W., Landes, P. W., Cowan, N. J., and Price, D. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 3472−3476.
3. Lee, M. K., and Cleveland, D. W. (1994) *Curr. Opin. Cell Biol. 6*, 34−40.
4. Liem, R. K. H. (1993) *Curr. Opin. Cell Biol. 5*, 12−16.
5. Julien, J. P., Meijer, D., Hurst, J., and Grosveld, F. (1986) *Mol. Brain Res. 1*, 243−250.
6. Napolitano, E. W., Chin, S. S., Colman, D. R., and Liem, R. K. (1987) *J. Neurosci. 7*, 2590−2599.
7. Chin, S. M., and Liem, R. K. H. (1990) *J. Neurosci. 10*, 3714−3726.
8. Steinert, P. M., and Roop, D. R. (1988) *Annu. Rev. Biochem. 57*, 593−625.
9. Elhanany, E., Jaffe, H., Link, W. T., Sheeley, D. M., Gainer, H., and Pant, H. C. (1994) *J. Neurochem. 63*, 2324−2335.
10. Jaffe, H., Veeranna, and Pant, H. C. (1998) *Biochemistry 37*, 16211−16224.
11. De Mattos-Dutra, A., Meirelles, R., da Rocha, B. B., Kommers, T., Wofxhuk, S. T., Wajner, M., and Pessoa-Pureur, R. (2000) *Brain Res. 856*, 111−118.
12. Leterrier, J. F., Kaes, J., Hartwig, J., Vegners, R., and Janmey, P. A. (1996) *J. Biol. Chem. 271*, 15687−15694.
13. Jacomy, H., Zhu, Q., Couillard-Depres, S., Beaulieu, J.-M., and Julien, J.-P. (1999) *J. Neurochem. 73*, 972−984.
14. Jaffe, H., Sharma, P., Grant, P., and Pant, H. C. (2001) *J. Neurochem. 76*, 1022−1031.
15. Betts, J. C., Blackstocks, W. P., Ward, M. A., and Anderton, B. H. (1997) *J. Biol. Chem. 272*, 12922−12927.
16. Carr, S. A., Huddleston, M. J., and Annan, R. S. (1996) *Anal. Biochem. 239*, 180−192.
17. Balin, B. J., and Miller, M. E. (1995) *J. Neurosci. Res. 40*, 79−88.
18. Chan, S.-O., Peng, D., and Chiu, F.-C. (1997) *Brain Res. 775*, 107−118.
19. Veeranna, Amin, N. D., Ahn, N. G., Jaffe, H., Winters, C. A., Grant, P., and Pant, H. C. (1998) *J. Neurosci. 18*, 4008−4021.
20. Lee, V. M. Y., Carden, M., Schlaepfer, W., and Trojanski, J. (1987) *J. Neurosci. 7*, 3474−3488.
21. Xu, Z.-S., Liu, W.-S., and Willard, M. B. (1992) *J. Biol. Chem. 267*, 4467−4471.
22. Geisler, N., Vandekerckhove, J., and Weber, K. (1987) *FEBS Lett. 221*, 403−407.
23. Bennett, G. S., and Quintana, R. (1997) *J. Neurochem. 68*, 534−543.
24. Cleverley, K. E., Betts, J. C., Blackstock, W. P., Gallo, J. M., and Anderton, B. H. (1998) *Biochemistry 37*, 3917−3930.
25. Pant, H. C., and Veeranna (1995) *Biochem. Cell Biol. 75*, 575−592.
26. Julien, J.-P. (1995), *Biochem. Cell Biol. 73*, 593−597.
27. Poseqitz, M. C., and Tempst, P. (1999) *Anal. Chem. 71*, 2883−2892.

28. Bateman, R. H., Carruthers, R., Hoyes, J. B., Jones, C., Langridge, J. I., Millar, A., and Vissers, J. P. C. (2002) *J. Am. Soc. Mass Spectrom. 13*, 792−803.
29. PROWL: ProteinInfo results of (NF-M) gi|3641350|gb|AAC36357.1|.
30. Schlosser, A., Pipkorn, R., Bossemeyer, D., and Lehmann, W. D. (2001) *Anal. Chem. 73*, 170−176.
31. Chin, S. S. M., and Liem, R. K. H. (1989) *Eur. J. Cell Biol. 50*, 475−490.
32. Napolitano, W. W., Chin, S. S. M., Colman, D. R., and Liem, R. K. H. (1987) *J. Neurosci. 7*, 2590−2599.
33. Chin, S. S. M., and Liem, R. K. H. (1990) *J. Neurosci. 10*, 3714−3726.
34. Lewis, S. E., and Nixon, R. A. (1988) *J. Cell Biol. 107*, 2689−2701.
35. Nixon, R. A., and Lewis, S. E. (1986) *J. Biol. Chem. 261*, 16298−16301.
36. Julien, J. P., and Mushynski, W. E. (1982) *J. Biol. Chem. 257*, 10467−10470.
37. Lazarides, E. (1982) *Annu. Rev. Biochem. 51*, 219−258.
38. Hirokawa, N. (1991) in *The Neuronal Cytoskeleton* (Burgoyne, R. D., Ed.) pp 5−74, New York, Wiley.
39. Nixon, R. A., and Sihag, R. K. (1991) *Trends Neurosci. 14*, 501−506.
40. Nixon, R. A., and Shea, T. B. (1992) *Cell Motil. Cytoskeleton 22*, 81−91.
41. Julien, J. P., and Mushynski, W. E. (1983) *J. Biol. Chem. 258*, 4019−4025.
42. Dosmeci, A., Floyd, C., and Pant, H. C. (1990) *Cell Mol. Neurobiol. 10*, 369−382.
43. Link, W. T., Dosemeci, A., Floyd, C. C., and Pant, H. C. (1993) *Neurosci. Lett. 151*, 89−93.
44. Hollander, B. A., Bennett, G. S., and Shaw, G. (1996) *J. Neurochem. 66*, 412−420.
45. Sihag, R. K., and Nixon, R. A. (1989) *J. Biol. Chem. 264*, 457−464.
46. Sihag, R. K., and Nixon, R. A. (1990) *J. Biol. Chem. 265*, 4166−4171.
47. Gonda, Y., Nishizawa, K., Kitamura, S., Minoura, Y., Nishi, Y., and Inagaki, M. (1990) *Biochem. Biophys. Res. Commun. 167*, 1316−1325.
48. Dosemeci, A., and Pant, H. C. (1992) *Biochem. J. 282*, 477−481.
49. Ando, S., Tanabe, K., Gonda, Y., Sato, C., and Inagaki, M. (1989) *Biochemistry 270*, 27489−27494.
50. Nakamura, Y., Hashimoto, R., Kashiwagi, Y., Aimoto, S., Fukusho, E., Matsumoto, N., Kudo, T., and Takeda, M. J. (2000) *Neurochemistry 74*, 949−959.
51. Giasson, B. I., Cromlish, J. A., Athlan, E. S., and Mushynski, W. E. (1996) *J. Neurochem. 66*, 1207−1213.
52. Xu, Z.-S., Wu-Schyong, L., and Willard, M. B. (1990) *J. Neurosci. 10*, 1838−1846.
53. Shetty, K. T., Link, W. T., and Pant, H. C. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6844−6848.
54. Sihag, R. K., Jaffe, H., Nixon, R. A., and Rong, X. (1999) *J. Neurochem. 72*, 491−499.
55. Gotow, T., Tanaka, T., Nakamura, Y., and Takeda, M. (1994) *J. Cell Sci. 107*, 1949−1957.
56. Nixon, R. A., Paskevich, P. A., Sihag, R. K., and Thayer, C. (1994) *J. Cell Biol. 126*, 1031−1046.
57. Shaw, G. (1989) *Biochem. Biophys. Res. Commun. 162*, 294−299.
58. Shaw, G. (1991) in *The Neuronal Cytoskeleton* (Burgoyne, R. D., Ed.) pp 185−214, Alan R. Liss, New York.

BI030196Q