Structural Characterization of Recombinant Soluble Rat Neuroligin 1: Mapping of Secondary Structure and Glycosylation by Mass Spectrometry†

Ross C. Hoffman,‡,⊥ Lori L. Jennings,*,§,⊥ Igor Tsigelny,§ Davide Comoletti,§ Robyn E. Flynn,§ Thomas C. Sudhof,|| and Palmer Taylor§

Department of Pharmacology and Howard Hughes Medical Institute Mass Spectrometry Facility, University of California, San Diego, La Jolla, California 92093, and Department of Molecular Genetics and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75235

Received July 18, 2003; Revised Manuscript Received December 4, 2003

ABSTRACT: Neuroligins (NLs) are a family of transmembrane proteins that function in synapse formation and/or remodeling by interacting with β-neurexins (β-NXs) to form heterophilic cell adhesions. The large N-terminal extracellular domain of NLs, required for β-NX interactions, has sequence homology to the α/β hydrolase fold superfamily of proteins. By peptide mapping and mass spectrometric analysis of a soluble recombinant form of NL1, several structural features of the extracellular domain have been established. Of the nine cysteine residues in NL1, eight are shown to form intramolecular disulfide bonds. Disulfide pairings of Cys 117 to Cys 153 and Cys 342 to Cys 353 are consistent with disulfide linkages that are conserved among the family of α/β hydrolase proteins. The disulfide bond between Cys 172 and Cys 181 occurs within a region of the protein encoded by an alternatively spliced exon. The disulfide pairing of Cys 512 and Cys 546 in NL1 yields a structural motif unique to the NLs, since these residues are highly conserved. The potential N-glycosylation sequons in NL1 at Asn 109, Asn 303, Asn 343, and Asn 547 are shown occupied by carbohydrate. An additional consensus sequence for N-glycosylation at Asn 662 is likely occupied. Analysis of N-linked oligosaccharide content by mass matching paradigms reveals significant microheterogeneous populations of complex glycosyl moieties. In addition, O-linked glycosylation is observed in the predicted stalk region of NL1, prior to the transmembrane spanning domain. From predictions based on sequence homology of NL1 to acetylcholinesterase and the molecular features of NL1 established from mass spectrometric analysis, a novel topology model for NL three-dimensional structure has been constructed.

NLs[1] (1−4) comprise a family of membrane proteins that interact with β-NXs (*1*−*3*) to function as heterophilic cell adhesion molecules (*4*). NL1 localizes to postsynaptic densities of excitatory synapses (*5*). The production of NL1 and NL2 in nonneuronal cells has been shown to induce the in vitro formation of presynaptic structures in axons of cocultured neurons from the central nervous system (*6*). NXs present in presynaptic elements of neuronal pontine explants have been shown to function as NL partners in synapse formation; stimulation of synaptic differentiation by NL1 recruits NXs to newly forming synapses (*7*).

NLs are type 1 membrane proteins believed to be composed of five distinct domains: a cleaved N-terminal signal peptide, a large extracellular domain, a linker domain resembling an O-glycosylation cassette, a single transmembrane region, and a cytoplasmic C-terminal tail (*1*). An N-terminal hydrophobic sequence of NLs, followed by a typical cleavage site, was suggestive of a signal peptide and confirmed by amino acid sequence analysis (*2*). The large extracellular domain of NLs shares significant sequence identity (32−36%) with certain members (e.g., AChE) of the α/β hydrolase fold protein superfamily (*1*, *8*). However, the NLs lack amino acid residues critical for the characteristic catalytic triad that confers hydrolase activity. Through comparisons with the crystal structure of AChE (*9*, *10*), tertiary protein structures for NLs have been proposed (*1*, *2*) and topology models constructed (*7*). The hydrolase homology domain of NLs is required for their interaction with β-NXs, whereas the intracellular C-terminus region contains a PDZ binding domain and functions to associate with partnering proteins (*6*).

NLs contain several potential N-linked glycosylation sequons that result in heterogeneously glycosylated proteins. An *O*-carbohydrate-rich stalk is predicted in the linker region between the hydrolase homology domain and the transmembrane region, on the basis of the homology to O-linked glycosylation cassettes in other cell surface receptors (*2*). For NL1, purified from rat brain by affinity chromatography

† This work was supported by Grants USPHS P42-ES10337 and R37 GM-18360 to P.T.

* To whom correspondence should be addressed: Phone: (858) 534-1366. Fax: (858) 534-8248. E-mail: ljennings@ucsd.edu.

‡ Howard Hughes Medical Institute Mass Spectrometry Facility, University of California, San Diego.

§ Department of Pharmacology, University of California, San Diego.

|| University of Texas Southwestern Medical Center.

⊥ These authors contributed equally to this work.

[1] Abbreviations: AChE, acetylcholinesterase; ACN, acetonitrile; AMU, atomic mass unit; CNBr, cyanogen bromide; DMEM, Dulbecco's modified Eagle's medium; DTT, dithiothreitol; IAA, iodoacetic acid; LC−MS, liquid chromatography−mass spectrometry; *m/z*, mass (Da) to charge ratio; NLs, neuroligins; NXs, neurexins; TIC, total ion current; TFA, trifluoroacetic acid; TOF, time of flight; 4-VP, 4-vinylpyridine; XIC, extracted ion current.

10.1021/bi035278t CCC: $27.50
Published on Web 01/21/2004

with immobilized β-neurexin, deglycosylation with glycohydrolases for N- and O-linked sugars accounted for the approximate 21 kDa size difference between the native protein and its primary translation product (*1*).

The disulfide bonding pattern of NLs has been proposed from sequence homology to the α/β hydrolase fold protein superfamily (*1*, *2*), since four of the six disulfide-bonded cysteines in AChE are conserved in all esterases in this family. An alternatively spliced sequence in NL1 contains two cysteines that may form an internal disulfide loop. The alternatively spliced regions of NLs are found at loop positions in the AChE crystal structure (*9*, *10*) and may indicate these regions form a tertiary structure of extended surface loops. NLs contain a conserved cysteine residue between predicted disulfide loops that does not appear involved in intermolecular disulfide bonds, since NLs have similar migration when resolved by reducing and nonreducing SDS gel electrophoresis (*2*). An additional cysteine pair is present in the C-terminal region of the ectodomain, conserved among the NLs, but in a region where sequence identity to the hydrolase protein family diverges.

Although the configuration of the NLs has been predicted on the basis of the primary sequence, consensus sequons for glycosylation, and homology to esterases and other α/β hydrolase proteins, direct biochemical evidence establishing structural features of NLs has yet to be reported. We have made a detailed structural characterization of soluble NL (NL1-691), truncated prior to the membrane-spanning region, that contains the entire hydrolase-like domain and the perimembrane linker region. A systematic approach of peptide mapping and mass spectrometric analysis has been performed to (i) deduce disulfide pairing of cysteine residues, (ii) determine the occupancy and major oligosaccharide content of predicted N-linked glycosyl sequons, and (iii) establish the presence of O-linked carbohydrates in the Ser/Thr-rich region near the membrane-spanning segment.

MATERIALS AND METHODS

Chemicals and Reagents. Sequence-grade modified trypsin was purchased from Promega (Madison, WI), and endoproteinase Glu-C was from Roche Diagnostics (Indianapolis, IN). Peptide-*N*-glycosidase F (PNGase F), endo-α-*N*-acetylgalactosaminidase (*O*-glycanase), sialidase A, β-(1−4)-galactosidase, β-*N*-acetylhexosaminidase, and α-(1−2,3,4,5,6)-fucosidase were purchased from Prozyme (San Leandro, CA). pFLAG-CMV-1 expression plasmid, M2 anti-FLAG immunoaffinity gel, and other general chemical supplies were purchased from Sigma (St. Louis, MO). Mutagenesis of recombinant DNA constructs was performed with a Stratagene Quickchange kit (San Diego, CA). Dulbecco's modified Eagle's medium (DMEM) and pcDNA3 expression plasmid were purchased from Invitrogen (Carlsbad, CA). Ultraculture serum-free medium was purchased from Biowhittaker (Walkersville, MD). Centrifugal spin filter devices were purchased from Millipore (Bedford, MA).

Expression and Purification. The sequence encoding rat NL1 was subcloned into pFLAG-CMV1 expression plasmid to generate a construct that encoded a preprotrypsin leader peptide, a FLAG octapeptide epitope tag at the N-terminus, a 10-residue linker peptide, and NL1. The coding sequence for NL1 begins at Gln 46, the first amino acid residue of the mature protein (*1*, *2*). The construct was subcloned into pcDNA3 expression plasmid to generate NL1/FLAG-pcDNA3. For production of soluble NL1, a stop codon was introduced at Tyr 692 to truncate the protein prior to the transmembrane-spanning segment. Soluble recombinant rat NL1 (NL1-691) was produced by transfection of NL1/FLAG-pcDNA3 expression plasmid into human embryonic kidney (HEK293) cells, maintained in DMEM with 10% fetal bovine serum at 37 °C, 10% CO_2. By selection of expressing cells using G418 resistance, cultures were established that secreted milligram quantities of soluble NL1 into the growth media. FLAG-tagged NL1-691 was purified from serum-free growth media by immunoaffinity chromatography on M2 anti-FLAG gel according to the manufacturer's instructions. For subsequent analysis by mass spectrometry, purified NL1-691 was buffer exchanged into 50 mM NH_4HCO_3, pH 7.8, by spin filtration, aliquoted, lyophilized, and stored at −20 °C. Purification to homogeneity was confirmed by SDS−polyacrylamide gel electrophoresis and silver staining (results not shown). The protein quantity was determined by the method of Bradford (*11*).

Trypsin and Endoproteinase Glu-C Digestion. NL1-691 (50 μg) was digested with sequence-grade modified trypsin using the conditions 1.0 mg/mL protein in 100 mM Tris−HCl, pH 8.0, and 1:50 trypsin/protein (w/w) and incubated at 37 °C for 20 h. Digestion of NL1-691 (50 μg) with endoproteinase Glu-C was performed in a similar manner in 25 mM NH_4HCO_3, pH 8.0, buffer and 1:50 Glu-C/protein (w/w). Simultaneous digestion with protease and removal of N-linked oligosaccharides by PNGase F were carried out on an aliquot of NL1-691 (50 μg) using the same conditions as described, with the addition of 15 mU of PNGase F to the incubation mixture. Digests were stored at −20 °C and thawed as needed.

CNBr Cleavage. NL1-691 (50 μg) was cleaved with CNBr using the conditions 1.0 mg/mL protein in 70% TFA and a 500-fold molar excess of CNBr to Met residues in protein (nine Met residues, 0.322 mg of CNBr) and incubated at room temperature in the dark for 20 h. Immediately following incubation, the digest was spin vacuumed to dryness, resolubilized to 1.0 mg/mL protein in 100 mM NH_4HCO_3, pH 8.0, buffer, and stored at −20 °C. A sample of 20 μg of CNBr-cleaved NL1-691 was de-N-glycosylated by treatment with PNGase F using the conditions 1.0 mg/mL protein in 100 mM NH_4HCO_3, pH 8.0, and 6 mU PNGase F and incubated at 37 °C for 20 h. Digests were stored at −20 °C and thawed as needed.

Reduction of Disulfide Bonds. Proteolytic digests of NL1-691 were reduced by the addition of an aliquot of freshly prepared DTT (1.0 mg/mL solution), such that a 10-fold molar excess of DTT to cysteine residues was present. DTT-treated samples were incubated at 37 °C for 1 h prior to LC−MS.

Differential Alkylation of Cys Residues. NL1-691 (100 μg) was treated with IAA to carboxymethylate any free cysteine residues using the conditions 1.0 mg/mL protein in 50 mM NH_4HCO_3, pH 7.8, buffer and 50 mM IAA and incubated for 2 h. The samples were spin filtered to remove excess IAA, buffer exchanged to 6 M guanidine−HCl, pH 6.5, and incubated for 1 h. Disulfide linkages were reduced by treatment with 10 mM DTT for 1 h, after which the remaining cysteine residues were pyridylethylated by the

addition of 50 mM 4-VP and incubated for 2 h. All incubations were performed at 37 °C in the dark. The protein was spin filtered to remove the alkylating reagents, buffer exchanged with 50 mM NH_4HCO_3, pH 7.8, buffer, and lyophilized. Also, samples of NL1-691 were denatured by the addition of 6 M guanidine−HCl, pH 6.5, to the lyophilized protein and incubated for 1 h, prior to differential alkylation in a manner similar to that described above. Samples of alkylated NL1-691 were stored at −20 °C, until they were thawed for subsequent digestion with trypsin and PNGase F.

O-Deglycosylation. NL1-691 (100 μg) was treated with a cocktail of glycosidases for the removal of O-linked glycans using the conditions 1.0 mg/mL protein in 50 mM Na_2HPO_4, pH 5.5, buffer, 4 mU *O*-glycanase, 20 mU sialidase A, 12 mU β(1−4)-galactosidase, 160 mU β-*N*-acetylhexosaminidase, and 100 mU α-fucosidase and incubated for 5 d at 37 °C. The de-O-glycosylated protein was buffer exchanged to 50 mM NH_4HCO_3, pH 7.8, by spin filtration and lyophilized. Samples were stored at −20 °C, until they were thawed for subsequent digestion with proteases and de-N-glycosylation, in a manner similar to that described previously.

Liquid Chromatography−Mass Spectrometry (LC−MS). Proteolytic digests of NL1-691 were separated by reversed-phase fractionation on a Michrom Magic 2002 microbore HPLC instrument (Michrom BioResources, Inc., Auburn, CA). A 1.0 × 150 mm, 5 μM, 200A Magic C_{18} column, operated at a flow rate of 50 μL/min and column temperature of 22 °C, was used. A standard gradient was employed for all LC−MS/MS runs of 2% A isocratic for 5 min, immediately followed by a 2−62% B gradient over 40 min. A is 2% (v/v) ACN and 0.05% (v/v) TFA, and B is 90% (v/v) ACN and 0.0475% (v/v) TFA. Typical injections were of 5−10 μg of digested protein. The eluant was run unsplit, directly into the electrospray mass spectrometer.

Mass spectra were recorded on an Applied Biosystems QSTAR, a hybrid quadrupole TOF mass spectrometer, equipped with the electrospray source (Applied Biosystems, Foster City, CA). Standard electrospray operating conditions were curtain gas 20, nebulizer gas (GS1) 80, and ion spray voltage (ISV) 5.20 kV. The instrument was set to scan mass (Da) to charge ratios of 400−2000 every 1.00071 s. The data were analyzed with the Analyst & BioAnalyst software packages (Applied Biosystems).

Topology Modeling of Neuroligin. A structural model of NL1 was generated with the programs Homology and InsightII (Accelrys, Inc., 2002) using the crystal structure of mouse AChE (1mah in the PDB database) as a template for the extracellular domain. For the modeling of the C-terminal ectodomain of NL1, the following templates were used: hydrolase (1blp), nitrogenase (2nip), binding protein (1ola), and citrate synthase (1csc). After homology modeling, the structure of NL1 was energy minimized for 10000 iterations using the distance-dependent dielectric constant with the program Discover (Accelrys, Inc., 2002).

RESULTS

Peptide Mapping. Peptide mapping was performed on the 664 amino acid residue recombinant soluble rat NL1-691 using LC−MS/MS techniques. Digestion of the protein with trypsin, GluC, or CNBr was followed by injection of peptides to a microbore HPLC system plumbed directly into a hybrid quadrupole TOF mass spectrometer, QSTAR. The resultant datasets were analyzed and the peptides assigned on the basis of mass. This was facilitated by the excellent mass accuracy of the QSTAR; typical errors for mass-based assignments were below 25 ppm. Trypsin, GluC, and CNBr digests were analyzed, in nonreduced and reduced conditions, to elucidate disulfide pairings and obtain lower mass peptides for the characterization of glycosylation sites. Aliquots of NL1-691 were de-N-glycosylated by treatment with PNGase F to characterize the occupancy of N-glycosylation sites. PNGase F hydrolyzes N-linked oligosaccharides from Asn side chains, converting the Asn to Asp, if *N*-glycan is present. PNGase F treated tryptic digests of NL1-691, in nonreduced, reduced, and differentially alkylated states, were analyzed by LC−MS/MS techniques to confirm peptide assignments based on mass. Additionally, NL1-691 was subjected to a cocktail containing *O*-glycanase and other glycosidases to effect de-O-glycosylation of the protein and to characterize the O-glycosylation sites.

Following this strategy, 97% of the amino acid sequence of NL1-691 was accounted for with assignment of the peptides from trypsin LC−MS/MS datasets (see the Supporting Information) and confirmed by GluC and CNBr LC−MS datasets. Three tryptic peptides were not resolved, spanning the contiguous sequence of K661−K679. The released amino acid K661 and tryptic peptide Q675−K679 are most likely too hydrophilic to be retained on the C18 column and eluted from HPLC separations in the void volume. The tryptic peptide N662−K674 contains a potential N-glycosylation site at Asn 662 and several sites that are predicted to be O-glycosylated by the algorithm NetOGlyc of Hansen and co-workers (*12*−*14*). The combination of N- and O-linked glycosylation most likely accounts for the absence of the peptide in resolved datasets.

Disulfide Mapping. The disulfide pairing of nine cysteine residues present in NL1-691 was determined by comparison of the nonreduced, reduced, and differentially alkylated trypsin + PNGase F digests and the nonreduced and reduced GluC + PNGase F digests. PNGase F was utilized to remove *N*-glycans attached to asparagine residues (Asn 109, Asn 343, Asn 547) close in sequence to three cysteine residues (Cys 117, Cys 342, Cys 546). This strategy simplified the disulfide assignment process by removing microheterogeneity introduced by *N*-glycans and decreased the mass of the peptides involved.

In trypsin + PNGase F digests, two interpeptide disulfides and two intrapeptide disulfides were clearly indicated by comparing the masses, fragment ions observed in MS/MS spectra, and shifts in retention time between the nonreduced and reduced datasets. Tryptic peptides D109−R125 and L126−K166 were linked in nonreduced conditions with an average mass of 6620.24 Da and separated upon reduction with monoisotopic masses of 1843.89 and 4774.35 Da, respectively. Note that the native Asn 109 was converted to Asp by PNGase F.

Similarly, tryptic peptides E171−R174 and K179−R182 were linked in the nonreduced datasets with a monoisotopic mass of 993.49 Da and separated upon reduction with a monoisotopic mass of 518.31 Da observed for K179−R182. Tryptic peptide E171−R174 was not observed in the reduced datasets, most likely due to its hydrophilicity and elution in

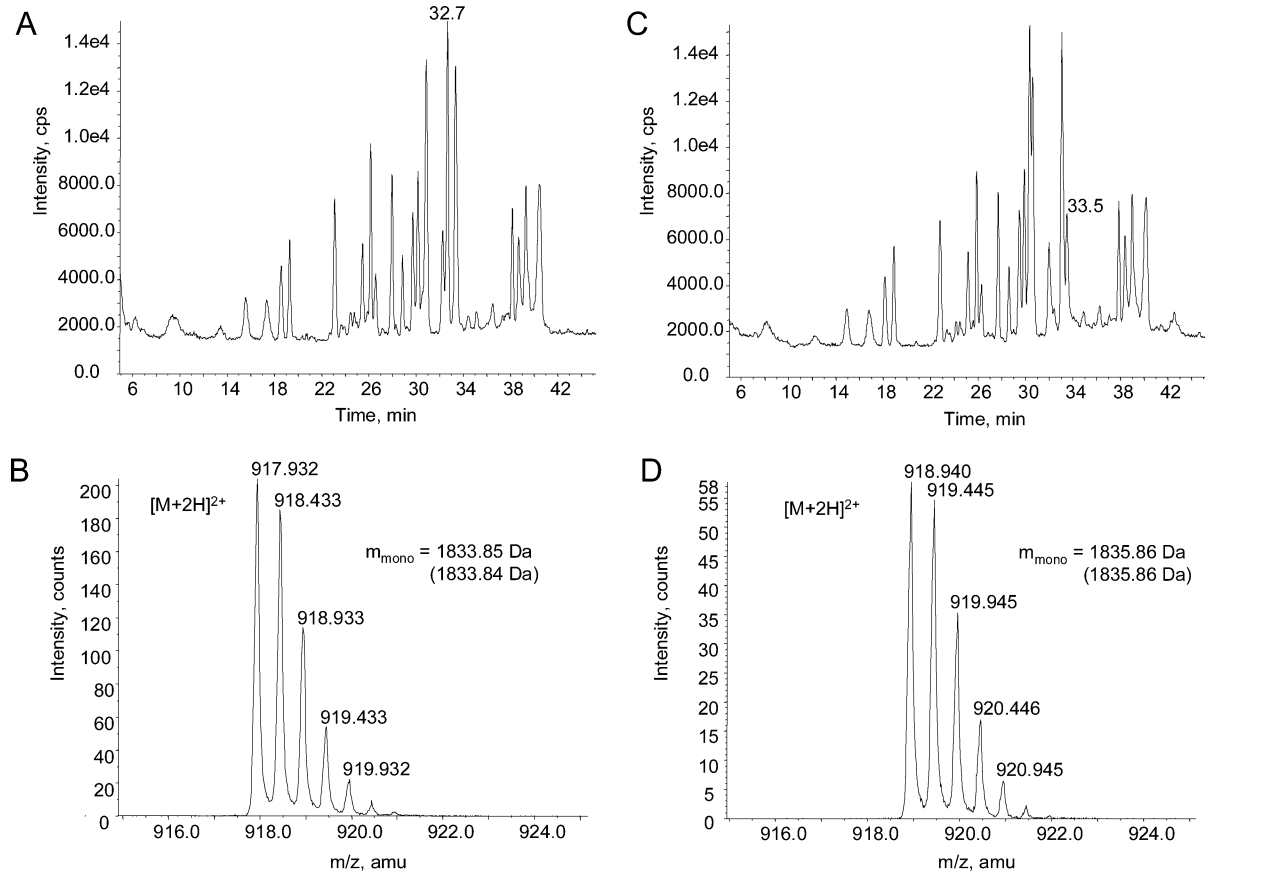

FIGURE 1: Disulfide linkage in the de-N-glycosylated peptide V340−K356. (A) LC−MS TIC chromatogram (*m*/*z* 400−2000) for the nonreduced PNGase F treated trypsin digest of NL1-691. The elution position of disulfide-linked tryptic peptide V340−K356 is indicated at 32.7 min. (B) Mass spectrum (*m*/*z* 915−925) of the disulfide-linked V340−K356 eluting at 32.7 min. The doubly charged ion gives a monoisotopic mass of 1833.85 Da. The predicted mass for the disulfide-linked peptide is shown in parentheses. (C) LC−MS TIC chromatogram (*m*/*z* 400−2000) for the reduced PNGase F treated trypsin digest of NL1-691. The elution position of the reduced tryptic peptide V340−K356 is indicated at 33.5 min. (D) Mass spectrum (*m*/*z* 915−925) of the reduced V340−K356 peptide eluting at 32.7 min. The doubly charged ion gives a monoisotopic mass of 1835.86 Da. The predicted mass for this reduced peptide is shown in parentheses. Note that the V340−K356 peptide is 0.98 Da heavier than the native sequence due to the conversion of N343 to D343 by PNGase F.

the HPLC void volume. However, the mass of 993.49 Da was only observed in nonreduced conditions, 518.31 Da was only observed following reduction, and acquired MS/MS spectra confirmed the peptide assignments.

Two intrapeptide disulfides were observed by an increase in mass of 2.01 Da (monoisotopic) between the nonreduced and reduced datasets and shifts in the retention times of affected peptides. Tryptic peptide V340−K356 changed from a monoisotopic mass of 1833.86 Da in nonreduced conditions to 1835.87 Da following reduction and shifted retention time by 0.8 min. Figure 1 presents data used to determine disulfide linkage between Cys 342 and Cys 353. MS/MS analyses for peptide V340−K356, before and after reduction, were consistent with an intrapeptide disulfide linkage (Figure 2). Tryptic peptide T475−K550 changed from an average mass of 8410.09 Da in nonreduced conditions to 8411.87 Da following reduction and shifted retention time by 0.4 min. A weak ionization signal precluded the observation of fragment ions in LC−MS/MS datasets for tryptic peptide T475−K550.

Thus, disulfide bonds present in NL1−691 are Cys 117−Cys 153, Cys 172−Cys 181, Cys 342−Cys 353, and Cys 512−Cys 546. The trypsin + PNGase F data are summarized in Table 1 for all cysteine-containing peptides. Additional support for the assignment of disulfides Cys 117−Cys 153 and Cys 512−Cys 546 was observed in trypsin datasets in the form of incomplete digestion products. In the case of the former, D109−R125 and L126−R167 behaved similarly to D109−R125 and L126−K166, and in the case of the latter, T475−K568 behaved similarly to T475−K550. Moreover, independent confirmation of disulfides Cys 117−Cys 153, Cys 172−Cys 181, and Cys 512−Cys 546 came from analysis of the GluC + PNGase F nonreduced and reduced digests. The disulfide bonding Cys 512−Cys 546 was also confirmed by the analysis of CNBr + PNGaseF digests before and after reduction.

Tryptic peptide I275−R300 contains the single remaining cysteine residue (Cys 286) in the extracellular domain of NL1. This tryptic peptide is observed as a disulfide-bonded homodimer in the nonreduced LC−MS/MS datasets with a monoisotopic mass of 5274.56 Da and as a monomer with a monoisotopic mass of 2638.28 Da in reducing conditions. It is not surprising that a peptide containing an unpaired cysteine residue, once released from the structural context of the native protein, forms a covalent dimer in the nonreducing environment of the trypsin digest. To investigate further the native state of Cys 286, NL1-691 was subjected to differential alkylation. IAA was used to carboxymethylate free cysteines, and following denaturation and reduction of disulfide bonds, 4-VP was used to pyridylethylate the remaining cysteine residues. Carboxymethylation and pyridylethylation of cysteine residues results in an increase in

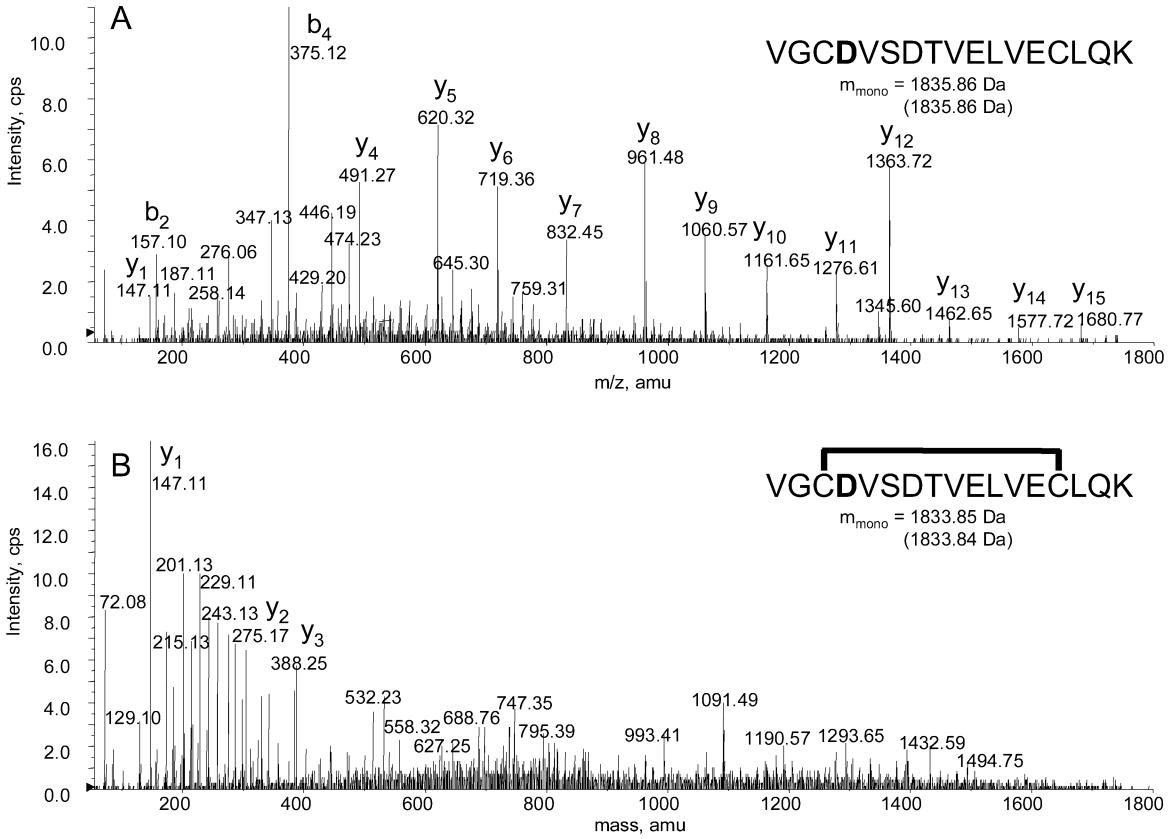

FIGURE 2: 2. MS/MS spectra (*m*/*z* 55−1800) of the (A) reduced and (B) oxidized de-N-glycosylated peptide V340−K356. The identities of the b and y fragment ions are indicated above the fragment ion masses. The sequence of the peptide is shown along with the observed and predicted masses; the predicted masses are shown in parentheses. N343 is converted to D343 by PNGase F treatment. D343 is shown in bold text.

Table 1: Masses[a] and Elution Times Observed for Cys-Containing Peptides in Nonreduced and Reduced LC−MS Datasets of PNGase F Treated Tryptic Digests of NL1-691

disulfides	nonreduced elution time (min)	mass (Da)	reduced elution time (min)	mass (Da)	sequence
C117−C153	40.4	6620.24 (6620.39)[b]	30.3	1843.89 (1843.89)	*D*$_{109}$ATQFAPV**C**PQNIIDGR$_{125}$ [c]
			42.6	4774.35 (4774.31)	L$_{126}$PEVMLPVWFTNNLDVVSS YVQDQSED**C**LYLNIYVPTED VK$_{166}$
C172−C181	9.8	993.49 (993.48)			E$_{171}$**C**AR$_{174}$
			7.2	518.31 (518.30)	K$_{179}$I**C**R$_{182}$
C286[d]	33.7	5274.56 (5274.55)	33.0	2638.28 (2638.28)	I$_{275}$TVFGSGAGGS**C**VNLLTLS HYSEGNR$_{300}$
C342−C353	32.7	1833.86 (1833.85)	33.5	1835.87 (1835.86)	V$_{340}$G**C***D*VSDTVELVE**C**LQK$_{356}$
C512−C546	39.8	8410.09 (8410.35)	39.4	8411.87 (8412.37)	T$_{475}$LLALFTDHQWVAPAVAT ADLHSNFGSPTYFYAFYHH**C** QTDQVPAWADAAHGDEVPY VLGIPMIGPTELFP**C***D*FSK$_{550}$

[a] Monoisotopic masses recorded below 6000 Da, average masses above 6000 Da. [b] Calculated masses based on sequence are shown in parentheses following observed masses. [c] Cys residues are shown in bold type. Asn residues converted to Asp residues by PNGase F are shown in italics. [d] C286 is believed to be unpaired in monomeric NL1-691. However, a dimer of I275−R300 is observed in nonreduced LC−MS data.

monoisotopic mass of 58.01 and 105.06 Da, respectively. Tryptic peptide I275−R300 is resolved in the alkylated trypsin + PNGase F datasets with monoisotopic masses of 2696.29 and 2743.34 Da, corresponding to masses expected for carboxymethylated and pyridylethylated peptides, respectively. The identification of the alkylated I275−R300 peptides was confirmed by MS/MS analysis. The abundances of the alkylated forms, as determined by ion peak areas, indicated Cys 286 was 82.6% carboxymethylated and 17.4% pyridylethylated. The extent of Cys 286 carboxymethylation indicates that the majority of this residue is free in native NL1-691 and suggests that it does not form stable disulfide-linked homodimers of the NLs. The incomplete carboxymethylation observed for Cys 286 may be due to poor accessibility, whereby IAA encounters steric hindrance that limits its ability to interact with this residue in native NL1-691. The denaturation of NL1-691 prior to the sequential treatment with IAA, DTT, and 4-VP results in exclusive pyridylethylation of Cys 286. In addition, denaturation of NL1-691 in the presence of IAA, followed by reduction and

pyridylethylation, results in 11.5% carboxymethyl-Cys 286 and 88.5% pyridylethyl-Cys 286. Taken together, these data suggest that Cys 286 exhibits great propensity to disulfide bond when released from the structural context of the protein. Note that all other Cys-containing tryptic peptides of NL1-691 were observed only as the pyridylethylated form, providing further confirmation of the disulfide linkages as assigned.

N-Glycosylation. NL1-691 contains six potential N-glycosylation sites in the form of N-glycosylation sequons: N(−6)SS, N109AT, N303ST, N343VS, N547FS, and N662ST. Comparison of LC−MS datasets for the reduced trypsin and the reduced trypsin + PNGase F digests indicates the first five potential N-glycosylation sites listed above are all N-glycosylated. Note that the estimates of occupancy are based on the assumption that nonglycosylated and de-N-glycosylated peptides ionize with an equal efficiency. Estimates of *N*-glycan structure predominance are based on the assumption that all *N*-glycopeptide forms at a given site ionize with equal efficiency. Since these assumptions may not be strictly true, the numbers should be viewed as best estimates. Oligosaccharide content, as determined by mass-matching paradigms, is used to propose structures that would be consistent with typical complex glycans commonly observed for mammalian expression systems (see the Supporting Information). These types of analyses have been shown to be fairly accurate when compared with more quantitative carbohydrate analyses (*15*). Each N-glycosylation site will be considered separately below.

Asn−6. This N-glycosylation site is not native to NL1, as it was introduced as part of an 18-residue FLAG peptide and linker region at the N-terminus of the recombinant construct. Occupancy of N-glycosylation site Asn −6 was estimated to be 98% on the basis of a comparison of the $(M + 2H)^{2+}$ ion peak areas of nonglycosylated tryptic peptide D(−18)K47 and de-N-glycosylated peptide in the reduced trypsin + PNGase F datasets. These two peptides differed in mass by 0.98 Da, and nonglycosylated peptide eluted 0.6 min before the de-N-glycosylated form. The *N*-glycopeptide D(−18)K47 elutes as a broad heterogeneous peak centered about 24.5 min, and the de-N-glycosylated peptide elutes as a single sharp peak at 25.9 min. Analysis of the masses associated with tryptic *N*-glycopeptide D(−18)K47 suggest glycans that are mainly consistent with sialylated core fucosylated complex type *N*-glycans of biantennary to tetraantennary type, with biantennary predominating at this site. Glycan masses are observed that would be consistent with several further modifications of antennae, including GalNAc for Gal substitution, typically found at the end of nonsialylated antennae, fucosylation, and sulfation.

Asn 109. Occupancy of N-glycosylation site Asn 109 was estimated to be 100%, as no nonglycosylated tryptic peptide N109−R125 was observed in the reduced trypsin datasets. *N*-Glycopeptide N109−R125 elutes as a broad heterogeneous peak centered about 29.3 min, and the de-N-glycosylated peptide elutes as a single sharp peak at 30.3 min. Analysis of glycan masses associated with tryptic *N*-glycopeptide N109−R125 suggests they are mainly consistent with sialylated core fucosylated complex type *N*-glycans of biantennary to tetraantennary type, with biantennary predominating at this site. We find masses that indicate a number of modifications of antennae, including GalNAc for Gal substitution, fucosylation, and sulfation. Further, masses that can be attributed to an additional GlcNAc are observed for some *N*-glycans, consistent with a β1,4-GlcNAc attached to terminal mannose of the core.

Asn 303. Occupancy of N-glycosylation site Asn 303 was estimated to be 100%, as no nonglycosylated tryptic peptide W301−K306 was observed in the reduced trypsin datasets. *N*-Glycopeptide W301−K306 elutes as a broad heterogeneous peak centered about 7.9 min, and the de-N-glycosylated peptide was not observed in the reduced trypsin + PNGase F datasets. Analysis of the masses associated with the weakly ionizing tryptic *N*-glycopeptide W301−K306 suggests glycans that are mainly consistent with sialylated core fucosylated complex type *N*-glycans of tetraantennary type. We find no mass-based evidence for further modifications of antennae at this site.

Asn 343. Occupancy of N-glycosylation site Asn 343 was estimated to be 100% as no nonglycosylated tryptic peptide V340−K356 was observed in the reduced trypsin datasets. *N*-Glycopeptide V340−K356 elutes as a broad heterogeneous peak centered about 32.2 min, and the de-N-glycosylated peptide elutes as a single sharp peak at 33.5 min. Figure 3 shows the mass spectrum of the heterogeneous *N*-glycopeptide V340−K356, indicating the doubly and triply charged ions and the reconstructed mass spectrum showing the *N*-glycopeptide masses derived from the raw mass spectrum. Analysis of glycan masses associated with tryptic *N*-glycopeptide V340−K356 suggests they are mainly consistent with sialylated core fucosylated complex type *N*-glycans of biantennary to tetraantennary type, with biantennary predominating at this site. Additional modifications of antennae are suggested, by observed masses that would be consistent with GalNAc for Gal substitution, fucosylation, and sulfation. The mass for an additional GlcNAc was observed for some glycan structures.

Asn 547. Occupancy of N-glycosylation site Asn 547 was estimated to be 100%, as no nonglycosylated tryptic peptide T475−K550 was observed in the reduced trypsin datasets. Although the N-glycosylated form of the peptide containing Asn 547 was not observed in the reduced trypsin LC−MS data, the de-N-glycosylated peptide T475-K550 elutes as a single peak at 39.4 min in the reduced trypsin + PNGase F datasets. The de-N-glycosylated peptide has a mass of over 8000 Da and weakly ionizes; it is therefore not surprising that the more massive and heterogeneous *N*-glycopeptide does not resolve. However, this precludes characterization of the type of *N*-glycans associated with Asn 547 by mass analysis.

Asn 662. The tryptic peptide containing Asn 662 (N662−K674) was not observed in any of the trypsin or GluC datasets. Treatment with both PNGase F and a glycosidase cocktail for de-O-glycosylation failed to produce a mass that could be attributed to deglycosylated peptide N662−K674 or incomplete cleavage products containing this peptide. The absence of the parent mass for this peptide in both the trypsin + PNGase F digested NL1-691 and the trypsin + PNGase F digested de-O-glycosylated NL1-691 suggests that both N- and O-glycosylations are occurring on this peptide. This may render the peptide partially resistant to enzymatic deglycosylation and/or confer sufficient heterogeneity that it elutes over an extended time with a poor ionization signal.

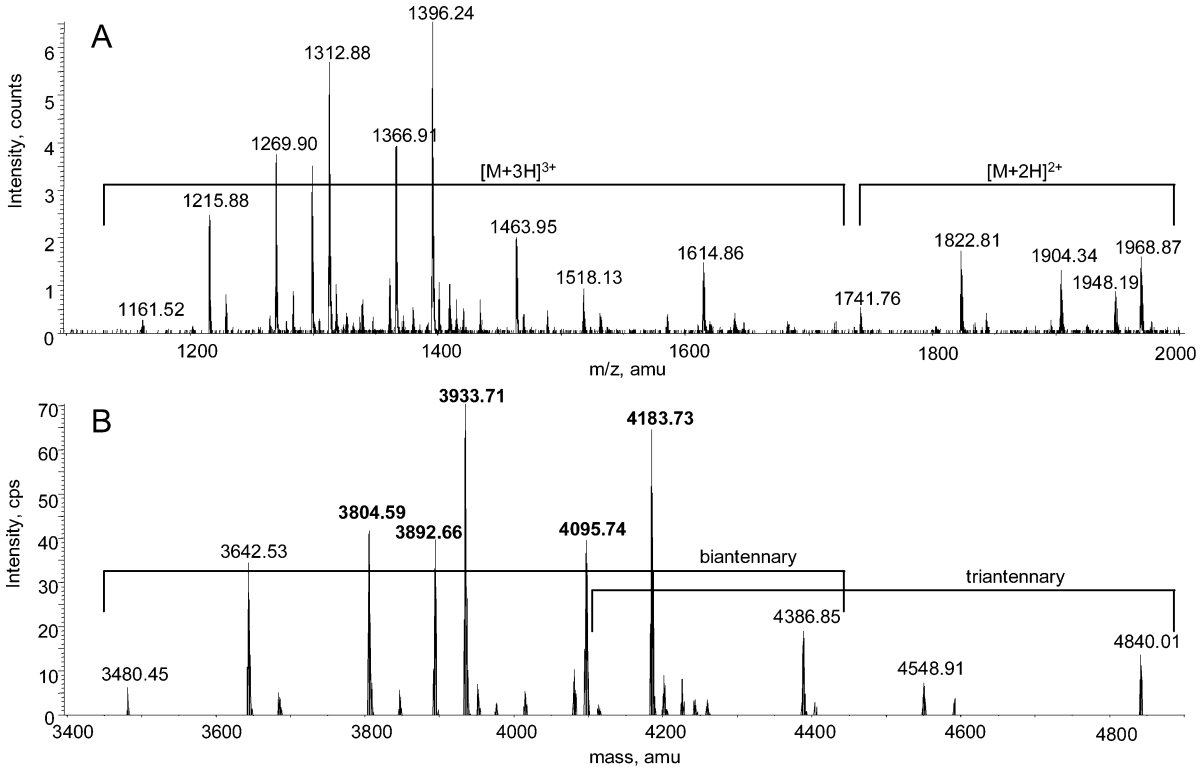

FIGURE 3: *N*-Glycopeptide V340−K356. (A) Summed mass spectrum (*m*/*z* 1100−2000, 31.4−31.9 min) from the LC−MS dataset of the nonreduced trypsin digest of NL1-691. The ranges of doubly and triply charged ions of the heterogeneous tryptic *N*-glycopeptide V340−K356 are indicated by bars. (B) Reconstructed mass spectrum from 3400 to 4900 AMU of the mass spectrum shown in panel A. Monoisotopic masses are shown above a representative set of peaks, and mass ranges of biantennary and triantennary complex *N*-glycans are indicated by bars. The five most prevalent masses of *N*-glycopeptide V340−K356 are indicated in bold text.

O-Glycosylation. We find masses for tryptic peptide Q680−R690 that would be consistent with O-glycosylation by mucin-type *O*-glycans, likely typical mono- and disialylated Galβ1−3GalNAcα1-S/T glycosylation. Figure 4 shows the total ion current (TIC) chromatogram for the nonreduced trypsin dataset as well as the extracted ion currents (XICs) for doubly charged ions of Q680−R690 in various O-glycosylation states. On the basis of doubly charged ion intensities, it is estimated that the *O*-glycan occupancy of this peptide is about 21%. Of the O-glycosylated peptide, approximately 85% resolves as masses that can be attributed to a single mucin-type *O*-glycan and 15% to two mucin-type *O*-glycans. There are only two possible O-glycosylation sites present in this peptide, indicating Ser 683 and Ser 686 are O-glycosylated in NL1-691.

Indirect evidence suggests O-glycosylation to occur in tryptic peptides V623−K648 and V649−R660. These two tryptic peptides were observed only in the trypsin digests of enzymatically de-O-glycosylated NL1-691. Upon de-O-glycosylation, two new masses appear in the trypsin datasets, 3035.56 Da at 30.6 min and 1354.75 Da at 15.6 min, corresponding to masses expected for V623−K648 and V649−R660, respectively. The mucin-type O-glycosylation prediction algorithm (*12*−*14*) predicts four O-glycosylation sites in these two tryptic peptides, Ser 644, Thr 647, Thr 652, and Thr 659. This does not preclude the possibility that *O*-glycans other than the mucin type are attached to the two *O*-glycopeptides. Four mucin-type O-glycosylation sites are also predicted in tryptic peptide N662−K674. The predicted sites are Thr 664, Thr 667, Ser 668, and Thr 672. Note that no other O-glycosylation sites are predicted N-terminal to residue Val 623.

DISCUSSION

Greater than 97% of the NL1-691 amino acid sequence has been accounted for and characterized by LC−MS and LC−MS/MS analysis. N-glycosylation has been confirmed at five of the six possible sites, and the *N*-glycan content has been characterized at four of the observed sites. The masses of *N*-glycans are typical of those observed in other recombinant proteins expressed in mammalian cell lines, sialylated core fucosylated complex *N*-glycans with variable antennae numbers and structures. O-glycosylation has been confirmed at two C-terminal serine residues, and other likely O-glycosylation sites have been predicted (*12*−*14*). A graphical summary of peptide, disulfide, and glycosylation data is presented in Figure 5.

Given the homology of the extracellular domain of NLs to α/β hydrolase fold proteins, the pattern of cysteine connectivity and sites of oligosaccharide attachment of NL1 provide a foundation to deduce the tertiary structure of NLs. Moreover, the experimental evidence acquired here supports refining of a structural model for NL (Figure 6) constructed by predictions based mainly on homology to AChE.

The truncation of rat NL1 prior to the transmembrane segment produces a soluble protein comprising the entire ectodomain. Recombinant soluble rat NL1-691 can be purified to homogeneity, by use of an epitope tag, facilitating detailed biophysical analysis by peptide mapping and mass spectrometric strategies. It is not known whether the glycosyl

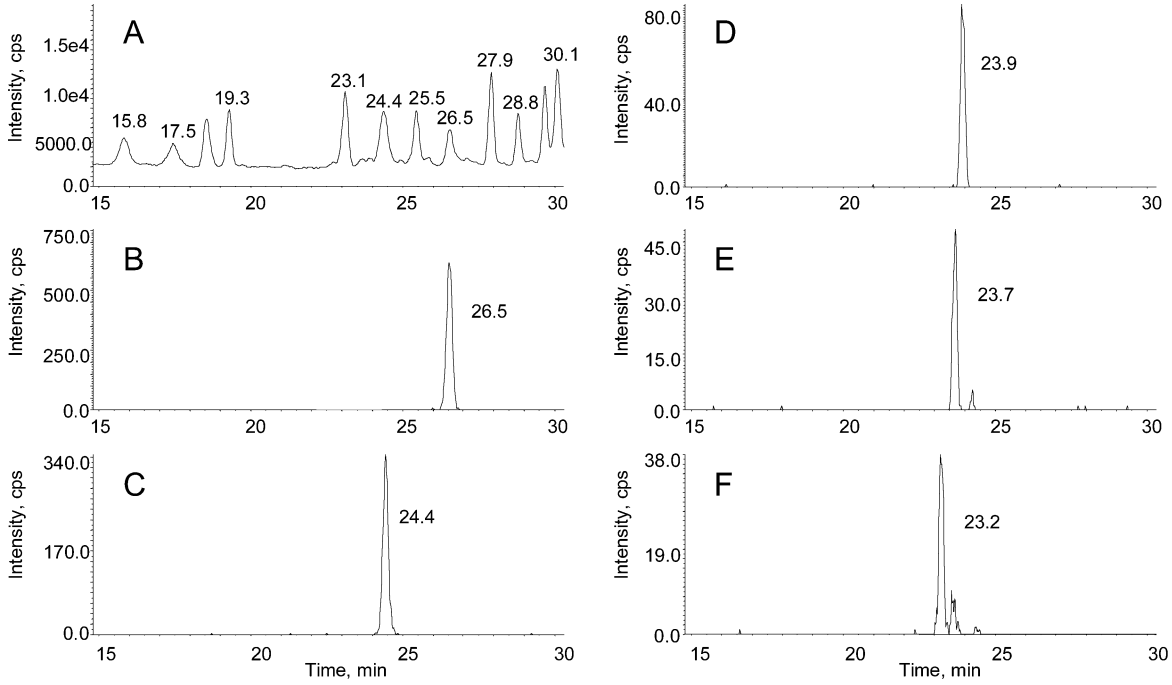

FIGURE 4: O-glycopeptide Q680−R690. (A) LC−MS TIC chromatogram (*m*/*z* 400−2000) for the nonreduced trypsin digest of NL1-691 from 15 to 30 min. (B) XIC of *m*/*z* 636.1−636.5 from the chromatogram in panel A showing the elution position of the doubly charged ion of nonglycosylated peptide Q680−R690. Note that the mass of this tryptic peptide is 17.03 Da lower than that expected from the native sequence due to conversion of the N-terminal residue Q680 to a pyroglutamic acid, as confirmed by MS/MS analysis. (C) XIC of *m*/*z* 964.4−964.8 from the chromatogram in panel A showing the elution position of the doubly charged ion of Q680−R690 + monosialylated HexNAcHex. (D) XIC of *m*/*z* 1110.0−1110.4 from the chromatogram in panel A showing the elution position of the doubly charged ion of Q680−R690 + disialylated HexNAcHex. (E) XIC of *m*/*z* 1147.1−1147.5 from the chromatogram in panel A showing the elution position of the doubly charged ion of Q680−R690 + monosialylated $(HexNAcHex)_2$. (F) XIC of *m*/*z* 1292.7−1293.1 from the chromatogram in panel A showing the elution position of the doubly charged ion of Q680−R690 + disialylated $(HexNAcHex)_2$. Note that the scale of the *y*-axis (intensity, cps) is variable.

content resulting from production of recombinant NL1-691 in HEK293 cells is similar to those of NLs of neurons in the central nervous system. However, NLs produced by heterologous expression in HEK293 cells have been shown to interact with β-NXs (*16*) and to induce synapse formation in cocultured neurons (*6*, *7*, *17*). Such functionality indicates that the recombinant protein is properly folded and post-translationally processed in a manner similar to that of native NL.

Proteolytic digests of NL1 generated by treatment with trypsin, Glu-C, and CNBr were resolved to obtain an overlapping, comprehensive coverage of NL1 peptides. The amino acid sequence of recombinant NL1 was confirmed to be the expected primary translation product encoded by the cDNA sequence; cleavage of the signal peptide was confirmed, and no modifications of amino acids were evident. Peptides of NL1 generated by specific proteolysis were resolved in native and reduced conditions. Analysis of peptides by mass-matching paradigms and fragment ion data from MS/MS spectra elucidated the disulfide bridging pattern of the nine cysteines in the NL1 ectodomain. Eight cysteines are disulfide bonded, and one, Cys 286, is likely present as a free cysteine. Cys 286 does not appear to participate in intermolecular disulfide bonding between NL dimers, since a majority of this residue is converted to carboxymethyl-cysteine on treatment with alkylating reagent. Two disulfide bonds, Cys 117 to Cys 153 and Cys 342 to Cys 353, are highly conserved among the family of hydrolase proteins (α/β hydrolase protein database: http://bioweb.ensam.inra.fr/ESTHER/general?what=index). The disulfide linkage of Cys 117 to Cys 153 in NL1 is homologous to an ω loop in AChE, shown to undergo distinctive conformational fluctuations that are coupled to enzyme activity (*18*). While NLs lack hydrolase activity, it is tempting to speculate the homologous loop domain may undergo conformational changes that affect functional properties. Disulfide bonding of Cys 172 to Cys 181, presumably forming an exposed surface loop, is encoded within an alternatively spliced exon. This region has been identified in NL1 and NL2 isoforms of rat brain (*2*) and may be involved in adhesion or recognition events. The cysteine pair in the C-terminal region of the NL ectodomain, distinct from the majority of hydrolases in the family, is also shown to disulfide bond. It is this region which likely possesses a tertiary structure unique to the NL members. The overall cysteine pattern in the nonspliced region is highly conserved among the NLs, suggesting these three disulfide-bonded loops in the invariant regions are essential for defining the structural motif.

NL1 peptides containing N-glycosylation sequons have been mapped, in native conditions and following de-N-glycosylation, to demonstrate carbohydrate occupancy at Asn 109, Asn 303, Asn 343, and Asn 547. An additional N-glycosyl attachment site occurs in the linker peptide between the epitope tag and NL1 in the recombinant construct (Asn −6) and is occupied. N-glycosylation at Asn 662 could not be confirmed, since the peptide containing the sequon was not observed in proteolytic digests of NL1, native or deglycosylated. It is likely that both N- and

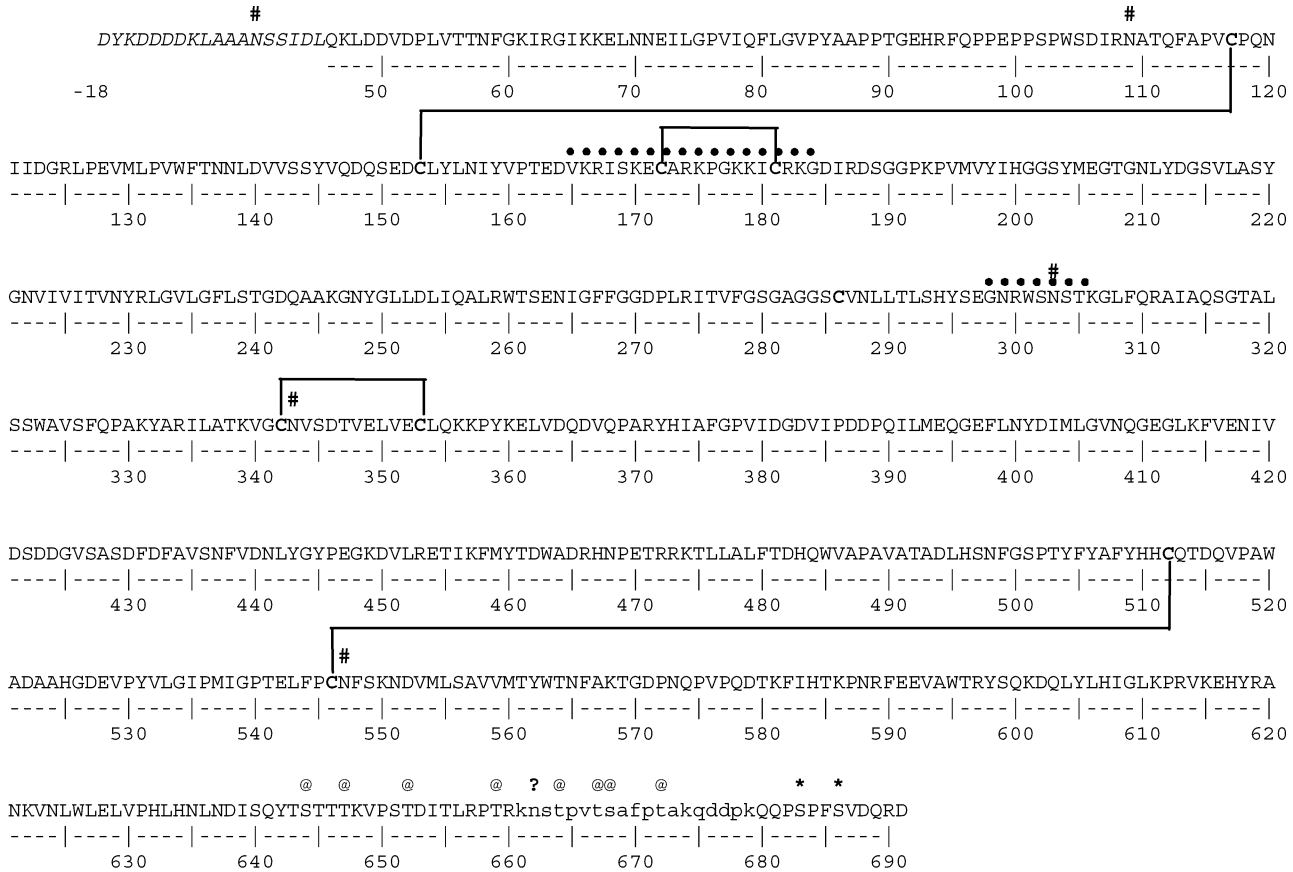

FIGURE 5: Peptide, disulfide, and glycosylation map of NL1-691. The sequence of rat NL1-691 expressed from Q46−D691 is shown, including the N-terminal FLAG and linker peptide designated as residues −18 to −1 and contiguous with Q46; the FLAG peptide grafted onto NL1-691 sequence is shown in italics. The alternatively spliced regions of the NLs (V165−G184 and G298−T305) are denoted by a dotted line. The region from K661 to K679, shown in small letters, was not observed. Cysteine residues are shown in bold text, and disulfide bridges are shown by connecting lines. The experimentally characterized glycosylation sites are marked by asterisks (O-glycosylation) and by pound signs (N-glycosylation). The single potential N-glycosylation site in the unresolved region is indicated by a question mark. Possible O-glycosylation sites, as predicted by the mucin-type O-glycosylation algorithm of Hansen and co-workers (*12*−*14*), are marked by ampersands.

O-linked glycosylations on the peptides comprising this region lead to inferior resolving capabilities by LC−MS methods (discussed further below).

An analysis of the carbohydrate content at the first four N-glycosylation sites has been performed by mass-matching paradigms. A significant microheterogeneous population of glycosyl residues is identified. The major species suggested by the observed masses are consistent with trimannosyl core complex carbohydrates, core fucosylated, bi-, tri- and tetraantennary structures, that can be sialylated, sulfated, and/or capped by neutral *N*-acetylhexosamine. The most abundant species, as determined by ion intensity in resolved mass spectra, were consistent with highly branched, core fucosylated and sialylated carbohydrates.

The occupancy of N-linked glycosylation sequons leads us to conclude that these domains are surface exposed and facilitates the prediction of tertiary protein structure from the amino acid sequence. Moreover, N-linked glycosylation may play an important functional role. Studies on binding interactions of NLs with β-NXs, by surface plasmon resonance techniques, have shown glycosylation processing of NL to inhibit this interaction (*16*). Interestingly, N-glycosylation at Asn 303 occurs in a region alternatively spliced among NL isoforms and may represent a mechanism for modulating NL and NX interactions. In previous studies, large N-linked oligosaccharides have been shown to influence the recognition of protein−protein interactions and to down modulate the activity of enzymes and signaling molecules (*19*, *20*).

O-glycosylation in the segment of the NL1 ectodomain that links the hydrolase homology and transmembrane-spanning regions was confirmed by the direct observation of masses that are consistent with mucin-like sugars on the peptides containing Ser 683 and Ser 686. The prediction of O-linked glycosylation by the NetOGlyc algorithm (*12*−*14*) indicates Ser 644, Thr 647, Thr 652, Thr 659, Thr 664, Thr 667, Ser 668, and Thr 672 have a high probability for modification. Interestingly, Ser 683 and Ser 686, directly observed to carry O-linked mucin-type sugars, were predicted to have low probability of occupancy. O-linked carbohydrates in the region spanning Ile 639 to Arg 660 were inferred by the ability to observe the peptides only after deglycosylation by a cocktail of glycohydrolases specific for O-linked oligosaccharides. It was not possible to resolve O-glycosylated peptides within this region. Additionally, the peptide from Asn 662 to Lys 674, also containing an *N*-glycosyl sequon, was not detectable. A high content of oligosaccharides would contribute to inadequate separation by reversed-phase chromatography and/or peptides with poor ionization potential for detection by mass spectrometry. Further, we presume that a high carbohydrate content may interfere, by steric hindrance, with the exposure of appropriate bonds to

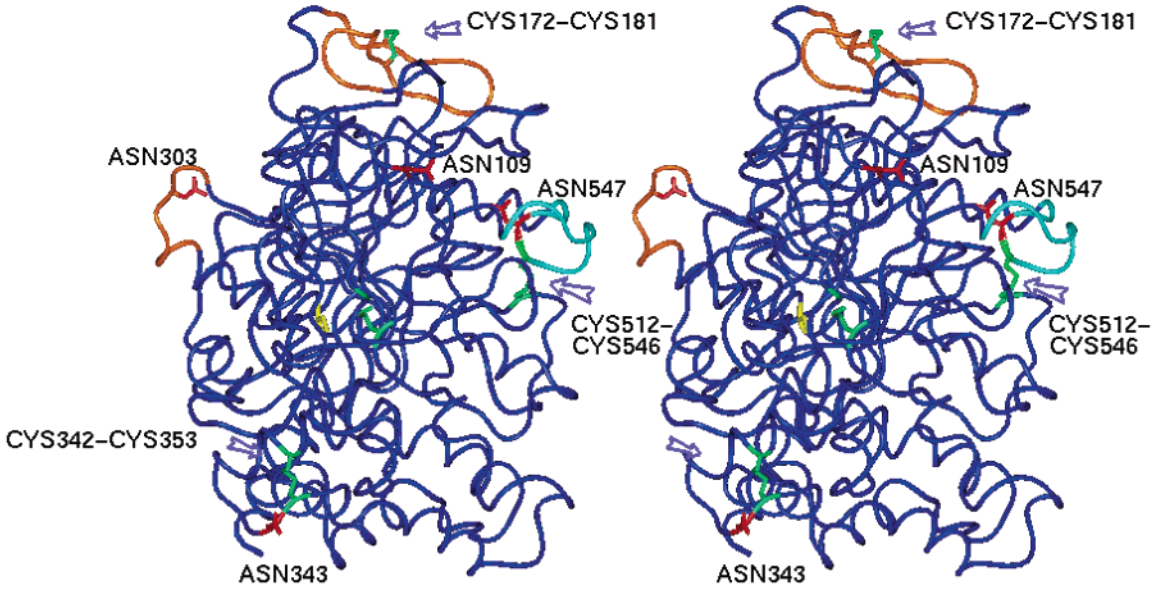

FIGURE 6: Topology model of NL1 from amino acids Q46−S640, generated with the programs Homology, InsightII, and Discover (Accelrys, Inc., 2002) shown as a stereo image. Disulfide linkages are indicated in green. In the center of the molecule, the unlabeled yellow residue on the left is Cys 286 and the disulfide bridge between Cys 117 and Cys 153 is to its right. N-linked glycosylation sites are indicated in red. Alternatively spliced regions of NL1 are colored orange. The region where the novel structure is proposed to account for disulfide linkage of Cys 512 to Cys 546 is indicated in cyan.

the glycosidase active site, thereby preventing deglycosylation.

O-glycosylation in protein domains with high Ser/Thr/Pro content and limited N-linked carbohydrates, such as that found in the NL1 linker region, causes significant perturbations in secondary protein structure and has a striking influence on tertiary conformations (*21*). The high number of oligosaccharides on Ser and Thr, combined with an abundance of Pro residues, creates a bottlebrush-like structure. Multiple interactions of GalNAc at Ser/Thr residues and the peptide backbone lead to an elongation of the peptide chain, as demonstrated in the P-selectin molecule (*22*). For the neurotrophin receptor, an O-glycosylation cassette that is heavily sialylated in a stalk-like region, similar to NLs, helps to orient the ligand binding domain away from the cell surface and is possibly involved in intercellular signaling (*23*). Further, O-glycosylation cassettes can function in conferring protein stability and protease resistance. Of particular interest is that *O*-glycans have been shown to influence the aggregation of proteins and their quaternary structure, as for lactase phlorizin hydrolase (*24*). A recent study demonstrated that oligomerization of NL is required for functional activity (*7*). Perhaps O-linked sugars serve to modulate NL and NX associations.

A topology model predicting the tertiary structure of the NL1 ectodomain (Figure 6) has been refined to incorporate the disulfide bonding pattern and position of extrafacial domains, inferred from carbohydrate attachment sites. The unequivocal disulfide connectivity between Cys 512 and Cys 546 indicates NL1 topology must deviate from predictions based solely on AChE homology. This is not surprising, since significant sequence divergence from AChE occurs in the C-terminal portion of the esterase-like domain of NLs. Therefore, a substantial difference in tertiary structure in this region of the NL1 ectodomain is proposed to place the cysteine pair in close proximity to support disulfide bonding.

Mutations of X-linked genes encoding NLs have been found to be associated with autistic disorders and Asperger syndrome (*25*). The ability of NL1 to induce synapse formation requires the specific interaction of NL1 with β-NXs and a homooligomerization of NL1, involving extracellular protein domains that are discrete from the NL and NX binding interface (*7*). Defining the molecular topology of the extracellular region of NL1 and posttranslational modifications, such as N- and O-glycosylation, is an important advance toward efforts to elucidate structure−activity relationships essential for NL-induced synaptic organization in the central nervous system.

ACKNOWLEDGMENT

We thank Dr. Elizabeth Komives for expert advice on mass spectrometric analyses.

SUPPORTING INFORMATION AVAILABLE

Tables S1 and S2 describing LC−MS/MS datasets for trypsin + PNGaseF treated NL1-691, in nonreduced and reduced conditions, Tables S3−S5 describing the heterogeneous N-linked glycosyl species observed at Asn 109, Asn 303, and Asn 343 of NL1-691, as predicted by the average mass (Da) of oligosaccharide chains, and Table S6 describing the O-linked glycosyl structures of the mucin type found at Ser 683 and Ser 686 in NL1-691. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Ichtchenko, K., Hata, Y., Nguyen, T., Ullrich, B., Missler, M., Moomaw, C., and Sudhof, T. C. (1995) Neuroligin 1: a splice site-specific ligand for beta-neurexins, *Cell 81*, 435−443.
2. Ichtchenko, K., Nguyen, T., and Sudhof, T. C. (1996) Structures, alternative splicing, and neurexin binding of multiple neuroligins, *J. Biol. Chem. 271*, 2676−2682.
3. Bolliger, M. F., Frei, K., Winterhalter, K. H., and Gloor, S. M. (2001) Identification of a novel neuroligin in humans which binds to PSD-95 and has a widespread expression, *Biochem. J. 356*, 581−588.

4. Nguyen, T., and Sudhof, T. C. (1997) Binding properties of neuroligin 1 and neurexin 1beta reveal function as heterophilic cell adhesion molecules, *J. Biol. Chem. 272*, 26032−26039.
5. Song, J. Y., Ichtchenko, K., Sudhof, T. C., and Brose, N. (1999) Neuroligin 1 is a postsynaptic cell-adhesion molecule of excitatory synapses, *Proc. Natl. Acad. Sci. U.S.A. 96*, 1100−1105.
6. Scheiffele, P., Fan, J., Choih, J., Fetter, R., and Serafini, T. (2000) Neuroligin expressed in nonneuronal cells triggers presynaptic development in contacting axons, *Cell 101*, 657−669.
7. Dean, C., Scholl, F. G., Choih, J., DeMaria, S., Berger, J., Isacoff, E., and Scheiffele, P. (2003) Neurexin mediates the assembly of presynaptic terminals, *Nat. Neurosci. 6*, 708−716.
8. Holmquist, M. (2000) Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms, *Curr. Protein Pept. Sci. 1*, 209−235.
9. Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., and Silman, I. (1991) Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding protein, *Science 253*, 872−879.
10. Marchot, P., Ravelli, R. B., Raves, M. L., Bourne, Y., Vellom, D. C., Kanter, J., Camp, S., Sussman, J. L., and Taylor, P. (1996) Soluble monomeric acetylcholinesterase from mouse: expression, purification, and crystallization in complex with fasciculin, *Protein Sci. 5*, 672−679.
11. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248−254.
12. Hansen, J. E., Lund, O., Engelbrecht, J., Bohr, H., and Nielsen, J. O. (1995) Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase, *Biochem. J. 308(Part 3)*, 801−813.
13. Hansen, J. E., Lund, O., Tolstrup, N., Gooley, A. A., Williams, K. L., and Brunak, S. (1998) NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility, *Glycoconjugate J. 15*, 115−130.
14. Hansen, J. E., Lund, O., Nilsson, J., Rapacki, K., and Brunak, S. (1998) O-GLYCBASE Version 3.0: a revised database of O-glycosylated proteins, *Nucleic Acids Res. 26*, 387−389.
15. Hoffman, R. C., Andersen, H., Walker, K., Krakover, J. D., Patel, S., Stamm, M. R., and Osborn, S. G. (1996) Peptide, disulfide, and glycosylation mapping of recombinant human thrombopoietin from Ser1 to Arg246, *Biochemistry 35*, 14849−14861.
16. Comoletti, D., Flynn, R., Jennings, L. L., Chubykin, A., Matsumura, T., Hasegawa, H., Sudhof, T. C., and Taylor, P. (2003) Characterization of the interaction of a recombinant soluble neuroligin-1 with neurexin-1-beta, *J. Biol. Chem. 90*, 3950−3957.
17. Fu, Z., Washbourne, P., Ortinski, P. I., and Vicini, S. (2003) Functional excitatory synapses in HEK293 cells expressing neuroligin and glutamate receptors, *J. Neurophysiol.* (in press).
18. Shi, J., Tai, K., McCammon, J. A., Taylor, P., and Johnson, D. A. (2003) Nanosecond dynamics of the mouse acetylcholinesterase Cys69−Cys96 omega loop, *J. Biol. Chem. 278*, 30905−30911.
19. Opdenakker, G., Rudd, P. M., Ponting, C. P., and Dwek, R. A. (1993) Concepts and principles of glycobiology, *FASEB J. 7*, 1330−1337.
20. Shen, L., and Kane, K. P. (1995) Differential ability of isolated H-2 Kb subsets to serve as TCR ligands for allo-specific CTL clones: potential role for N-linked glycosylation, *J. Exp. Med. 181*, 1773−1783.
21. Van den Steen, P., Rudd, P. M., Dwek, R. A., and Opdenakker, G. (1998) Concepts and principles of O-linked glycosylation, *Crit. Rev. Biochem. Mol. Biol. 33*, 151−208.
22. Li, F., Erickson, H. P., James, J. A., Moore, K. L., Cummings, R. D., and McEver, R. P. (1996) Visualization of P-selectin glycoprotein ligand-1 as a highly extended molecule and mapping of protein epitopes for monoclonal antibodies, *J. Biol. Chem. 271*, 6342−6348.
23. Chapman, B. S., Eckart, M. R., Kaufman, S. E., and Lapointe, G. R. (1996) O-linked oligosaccharide on the 75-kDa neurotrophin receptor, *J. Neurochem. 66*, 1707−1716.
24. Naim, H. Y., and Lentze, M. J. (1992) Impact of O-glycosylation on the function of human intestinal lactase-phlorizin hydrolase. Characterization of glycoforms varying in enzyme activity and localization of O-glycoside addition, *J. Biol. Chem. 267*, 25494−25504.
25. Jamain, S., Quach, H., Betancur, C., Rastam, M., Colineaux, C., Gillberg, I. C., Soderstrom, H., Giros, B., Leboyer, M., Gillberg, C., Bourgeron, T., Nyden, A., Philippe, A., Cohen, D., Chabane, N., Mouren-Simeoni, M. C., Brice, A., Sponheim, E., Spurkland, I., Skjeldahl, O. H., Coleman, M., Pearl, P. L., Cohen, I. L., Tsiouris, J., Zappella, M., Menchetti, G., Pompella, A., Aschauer, H., and Van Maldergem, L. (2003) Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism, *Nat. Genet. 34*, 27−29.

BI035278T