Neurexophilins Form a Conserved Family of Neuropeptide-Like Glycoproteins

Markus Missler and Thomas C. Südhof

Howard Hughes Medical Institute and Department of Molecular Genetics, The University of Texas Southwestern Medical School, Dallas, Texas 75235

Neurexophilin was discovered as a neuronal glycoprotein that is copurified with neurexin Iα during affinity chromatography on immobilized α-latrotoxin (Petrenko et al., 1996). We have now investigated how neurexophilin interacts with neurexins, whether it is post-translationally processed by site-specific cleavage similar to neuropeptides, and whether related neuropeptide-like proteins are expressed in brain. Our data show that mammalian brains contain four genes for neurexophilins the products of which share a common structure composed of five domains: an N-terminal signal peptide, a variable N-terminal domain, a highly conserved central domain that is N-glycosylated, a short linker region, and a conserved C-terminal domain that is cysteine-rich. When expressed in pheochromocytoma (PC12) cells with a replication-deficient adenovirus, neurexophilin 1 was rapidly N-glycosylated and then slowly processed to a smaller mature form, probably by endoproteolytic cleavage. Similar expression experiments in other neuron-like cells and in fibroblastic cells revealed that N-glycosylation of neurexophilin 1 occurred in all cell types tested, whereas proteolytic processing was observed only in neuron-like cells. Finally, only recombinant neurexin Iα and IIIα but not neurexin Iβ interacted with neurexophilin 1 and were preferentially bound to the processed mature form of neurexophilin. Together our data demonstrate that neurexophilins form a family of related glycoproteins that are proteolytically processed after synthesis and bind to α-neurexins. The structure and characteristics of neurexophilins indicate that they function as neuropeptides that may signal via α-neurexins.

Key words: neurexins; α-latrotoxin; synapse; gene duplication; proteolytic processing; neuropeptides; adenovirus expression

Neurexins are neuronal membrane proteins with a domain structure similar to that of cell-surface receptors (for review, see Missler and Südhof, 1998). Neurexins were discovered when neurexin Iα was purified on immobilized α-latrotoxin as an affinity matrix (Ushkaryov et al., 1992). The binding of neurexin Iα to α-latrotoxin suggested that neurexin Iα is a receptor for this toxin. This suggestion was confirmed in studies with recombinant neurexin Iα and neurexin Iα knock-out mice, demonstrating that neurexin Iα functions as a high-affinity α-latrotoxin receptor (Davletov et al., 1995; Geppert et al., 1998).

There are at least three genes for neurexins (Geppert et al., 1992; Ushkaryov et al., 1992, 1994; Ushkaryov and Südhof, 1993). Each gene has two independent promoters that direct transcription of longer α-neurexins and shorter β-neurexins. The neurexin transcripts are subject to extensive alternative splicing. The alternative splicing may result in the synthesis of >1000 neurexin isoforms that are differentially expressed in subpopulations of neurons (Ullrich et al., 1995). Although the domain structure of the neurexins suggests a receptor function, their overall biological role is unknown. It seems likely that α- and β-neurexins and some of their various alternatively spliced forms perform distinct functions. The discovery of a family of ligands for β-neurexins called neuroligins supports this hypothesis (Ichtchenko et al., 1995,

1996). Similar to neurexins, neuroligins are neuronal cell-surface proteins that are type 1 membrane proteins. They bind only to β-neurexins and only to one particular splice variant of β-neurexins. The binding of neuroligins to β-neurexins causes cell adhesion, suggesting a mechanism for the formation of a novel intercellular junction between neurons (Nguyen and Südhof, 1997). Intracellularly, β-neurexins and neuroligins are associated with different PDZ-domain proteins; CASK binds to neurexins, and postsynaptic density (PSD)-95 binds to neuroligins (Hata et al., 1996; Irie et al., 1997). Thus a subset of β-neurexins functions as cell adhesion molecules by binding to neuroligins, thereby forming an intercellular junction flanked by PDZ-domain proteins. Other neurexins, however, do not interact with neuroligins, suggesting that they have other ligands and different functions.

After purification on immobilized α-latrotoxin, part of neurexin Iα is isolated in a tight complex with a 29 kDa glycoprotein called neurexophilin (Petrenko et al., 1993). cDNA cloning showed that the primary translation product of neurexophilin is larger than the neurexophilin protein bound to neurexin Iα in the eluate from the α-latrotoxin affinity matrix (Petrenko et al., 1996). This suggested that neurexophilin may be physiologically processed by proteolytic cleavage or that it may have been partially degraded during protein purification. The mouse genome contains two genes for neurexophilin, but only one of these is transcribed (neurexophilin 1). In contrast, mRNAs corresponding to the second gene neurexophilin 2 were only found in bovine brain that, however, contained no detectable mRNAs for neurexophilin 1 (Petrenko et al., 1996). Thus there may be a species-specific expression of the two genes characterized up to now.

The structure of neurexophilin and its purification as a smaller protein in a complex with neurexin Iα suggested the hypothesis that neurexophilin may be post-translationally cleaved from a prepropeptide similar to a neuropeptide (e.g., see Eipper and Mains, 1980; Jacobs et al., 1981; Noda et al., 1982; Maisonpierre et al., 1990). The processed protein could then function as a ligand for α-neurexins. However, at present there is little direct evidence of this hypothesis, and many questions remain. First, virtually all neuropeptides are part of protein families with several related members (e.g., see Wimalawansa, 1997). By contrast, neurexophilin has no sequence homology to other proteins, and only a single neurexophilin gene was found to be expressed in mouse, rat, and bovine brain. Thus the question arises whether neurexophilin is also part of a gene family. Second, neurexophilin was purified in a complex with neurexin Iα, but its binding specificity is unclear. Does neurexophilin bind to all or only to a subset of neurexins? Third, the neurexophilin complexed to neurexin Iα is shorter than the primary translation product of the cDNA. Is neurexophilin physiologically processed by proteolytic cleavage, or is the smaller protein purified in the complex with neurexin Iα on immobilized α-latrotoxin the result of an artifact?

In the present paper, we have attempted to address these questions. We found that neurexophilins form a large gene family, that neurexophilin 1 specifically binds to α-neurexins, and that neurexophilin 1 is physiologically processed in neuronal but not in fibroblastic cells, indicating cell-specific proteolytic cleavage. Our results support the notion that neurexophilins represent a family of signaling molecules that resemble neuropeptides and that act by binding to α-neurexins and possibly other receptors.

MATERIALS AND METHODS

Molecular cloning. GenBank was searched using the BLAST programs of the NCBI (Altschul et al., 1997). Using the amino acid sequence of neurexophilin (Petrenko et al., 1996), we identified four classes of homologs in the human expressed sequence tag (EST) data bank, two of which represent the human homologs of the previously cloned mouse genes. The corresponding clones were obtained from the IMAGE consortium (clones 381246, 309825, 381764, and 704436), mapped with restriction endonucleases, and sequenced. The following restriction fragments were used as probes to screen rat brain and intestinal λZAP cDNA libraries (Stratagene, La Jolla, CA): 0.62 kb *Not*I/*Bgl*II fragment from clone 381764, 0.6 kb *Sfi*I fragment from clone 309825, and 0.42 kb *Eco*RI/*Nhe*I fragment from clone 704436. λZAP clones were plaque-purified and sequenced after *in vivo* excision and subcloning into pBluescript II and M13 vectors using standard procedures (Sambrook et al., 1989). We isolated three distinct full-length neurexophilin 3 cDNA clones extending up to 480 bp into the 5′-untranslated region and four distinct clones representing the C-terminal half of neurexophilin 3. For neurexophilin 4, the library contained two distinct full-length clones that include up to 190 bp of 5′-untranslated region and an additional set of clones coding for all but the signal peptide. All neurexophilin sequences were submitted to GenBank (accession numbers: AF042713, AF042714, AF043467, AF043468, and AF043469).

Northern blots. Human, rat, and mouse multitissue RNA blots (obtained from Clontech, Cambridge, UK) were sequentially probed with uniformly $[\alpha\text{-}^{32}P]dCTP$-labeled fragments from neurexophilin 1–4 cDNAs. A 400 bp *Pst*I/*Eco*RI fragment from the 5′-coding region of rat neurexophilin 1 (Petrenko et al., 1996), a 620 bp *Not*I/*Bgl*II fragment from the 5′-coding region of human neurexophilin 2 (EST clone 381764), a 530 bp *Eco*RI/*Bgl*II fragment from the 5′-region of rat neurexophilin 3 (cDNA), and a 420 bp *Eco*RI/*Kpn*I fragment from the 5′-region of rat neurexophilin 4 (cDNA) were used as probes. Additionally, a 620 bp *Sfi*I fragment from the 3′-region of human neurexophilin 3 (EST clone 309825) and a 420 bp *Eco*RI/*Nhe*I fragment from the 3′-region of human neurexophilin 4 (EST clone 704436) were used to reprobe the human RNA blots to confirm the specificity of cross-species hybridization signals. Prehybridizations and hybridizations were performed at 42°C overnight in 50% formamide plus 5× Denhardt's solution containing salmon sperm DNA at 0.1 mg/ml. Filters were washed twice for 20 min at 61–63°C in 2× SSC and 0.5% SDS and were exposed for 1–5 d. A second set of RNA blots was used for some hybridizations with virtually identical results.

Cell culture and transfections. COS cells were cultured in DMEM with 10% fetal bovine serum (FBS) and transfected using DEAE-dextran with chloroquine and a 2 min glycerol shock (Gorman, 1985). Human embryonic kidney 293 cells and STO cells (stably transfected fibroblasts routinely used in embryonic stem cell culture) were maintained under similar conditions and transfected by calcium phosphate precipitation. The human neuroblastoma cell line SH-SY5Y was grown in 84% F12 medium supplemented with 15% FBS and 1% nonessential amino acids. The human teratocarcinoma cell line NT2 (Ntera2/D1) cultures were differentiated into mature hNT cells that have some properties of CNS neurons, using retinoic acid according to the protocol of the supplier (Stratagene). Pheochromocytoma (PC12) cells were maintained in RPMI medium with 10% horse serum and 5% FBS and were plated on collagen-coated dishes before adenovirus infection experiments.

Recombinant adenovirus construction and infection studies. A 1.4 kb *Kpn*I fragment containing the entire rat neurexophilin 1 coding sequence was inserted into the multicloning site of the shuttle plasmid pAC-CMVpLqA. Replication-deficient adenovirus expressing neurexophilin (AdNph1) or neurexin Iβ and neuroligin 1 (as controls for antibody specificity) were made by homologous recombination and were propagated essentially as described (Graham and Prevec, 1991). The viral DNA used for cotransfection into replication-permissive 293 cells was from the 32 kb plasmid pJM17 (virus type Ad5 with a deletion of the E1a region). Viral plaques were screened by Western and/or Southern blots. Positive plaques were amplified to high titer stocks and were purified through cesium chloride gradients. Various cell lines were infected with serial dilutions of the different virus stocks to determine the optimum expression conditions for a protein. In a typical experiment, subconfluent cell cultures were inoculated with different titers of AdNph1 in the respective growth medium, the medium was replaced with virus-free medium after 8–24 hr, and cells were analyzed after the indicated time periods.

Western blot analysis and deglycosylation of neurexophilin. Immunoblot analyses of cell extracts were performed with ECL detection. We tested the expression products of AdNph1 for glycosylation by denaturing infected PC12 cells with SDS and β-mercaptoethanol, digesting with recombinant PNGase F (New England Biolabs, Beverly, MA), and analyzing by SDS-PAGE and immunoblotting using anti-neurexophilin antibody F508 (Petrenko et al., 1996).

Binding of neurexophilin 1 to IgG fusion proteins of neurexins. Processed and unprocessed forms of neurexophilin 1 were solubilized from infected PC12 cells using 2% 3-[(3-cholamidopropyl)dimethylammonio]-1-propane-sulfonate (CHAPS). PC12 cells were harvested, subjected to a hypo-osmolar shock (20 mM HEPES containing protease inhibitors for 10–15 min on ice), broken apart with a Dounce homogenizer (10–15 strokes on ice), and centrifuged to pellet insoluble material. Because initial experiments showed that neurexophilin 1 is in the pellet fraction of this centrifugation step (data not shown), the pellet was solubilized in 2% CHAPS in extraction buffer (40 mM Tris-HCl, pH 8.0, 0.15 M NaCl, and protease inhibitors). The supernatant of the solubilization step was used for binding experiments with different neurexin–IgG fusion proteins immobilized on protein A (Ushkaryov et al., 1994; Davletov et al., 1995; Ichtchenko et al., 1996). Binding experiments were performed in extraction buffer containing 1% CHAPS and were analyzed by SDS-PAGE and immunoblotting.

RESULTS

Cloning of neurexophilins: definition of a gene family

To identify neurexophilin homologs, we searched EST data banks with the neurexophilin sequence using the BLAST program. Human EST clones corresponding to neurexophilins 1 and 2 were identified and sequenced, revealing that neurexophilins 1 and 2 are highly conserved between rats, mice, and humans (>90% sequence identity; Fig. 1). The presence of human EST clones for both neurexophilin 1 and 2 indicates that unlike in bovine and rodent tissues, both neurexophilins are expressed in human tissues. In addition to EST clones for neurexophilins 1 and 2, we found human EST clones coding for two novel neurexophilins, neurexophilins 3 and 4. We used probes from these EST clones to
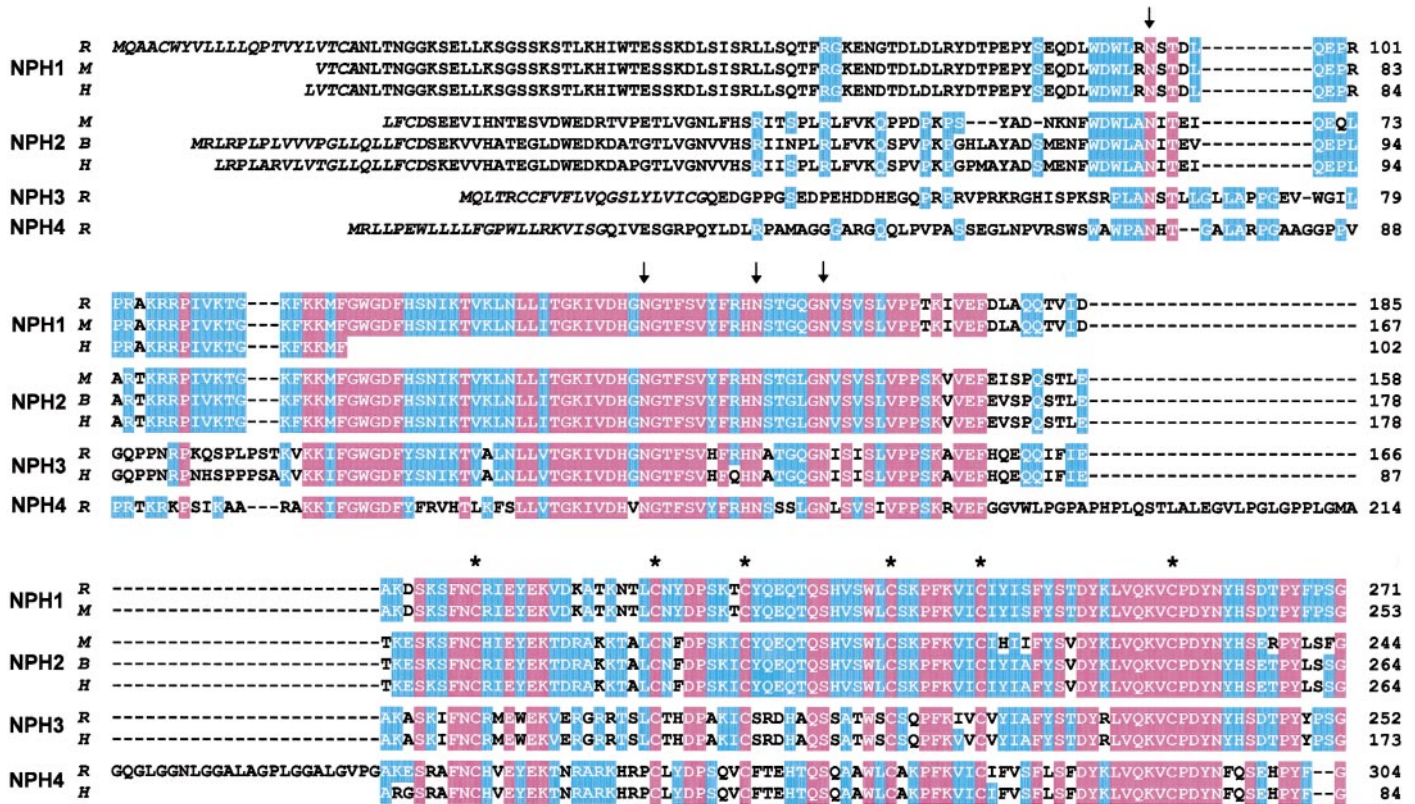
Figure 1. Primary structure of neurexophilins from rat (*R*), mouse (*M*), bovine (*B*), and human (*H*). The amino acid sequences of the four neurexophilins (NPH1–NPH4) are aligned for maximal homology, with *hyphens* indicating gaps. Sequences are identified on the *left* and numbered on the *right*. Residues that are identical in all sequences are shown on a *red* background, and residues identical in at least two isoforms are shown on a *blue* background. The four conserved N-glycosylation sites in neurexophilins are identified by *arrows*, and the six cysteines present in all neurexophilins with identical spacing are marked by *asterisks above* the sequences. The putative signal sequence is shown in *italics*. The linker sequence in neurexophilin 4 between the N-glycosylation domain and the cysteine-rich domain contains seven imperfect GGxL repeats (residues 201–234). The human neurexophilin 1, 3, and 4 sequences were obtained from incomplete EST clones and are only partial. No rat neurexophilin 2 cDNA could be identified, leading to the absence of a rat neurexophilin 2 sequence; the mouse neurexophilin 2 sequence was deduced from the genomic sequence (Petrenko et al., 1996).

screen a rat brain cDNA library and isolated clones containing the full coding region for both new neurexophilins. In addition, we rescreened the rat brain cDNA library for neurexophilin 2 to confirm that this neurexophilin is absent in rats but again were unable to isolate positive clones.

The translated amino acid sequences for all currently described neurexophilins, each with at least partial sequences from two or more species, are aligned with each other in Figure 1. This alignment demonstrates that each neurexophilin is highly conserved evolutionarily. Comparisons between different neurexophilins reveal that they are closely related to each other in their C-terminal regions but diverge considerably in their N-terminal sequences. Thus neurexophilins form a family of at least four related evolutionarily conserved proteins with divergent N terminals.

Domain structure of neurexophilins

The alignment of the neurexophilin sequences reveals a pattern of similarity and diversity that suggests a pronounced domain structure. Five domains can be distinguished (Fig. 2): I, an N-terminal hydrophobic sequence with the properties of a signal peptide (Fig. 1, *italicized region*); II, an N-terminal region that exhibits little sequence similarity between neurexophilins, although it is highly conserved evolutionarily for each individual neurexophilin and includes one conserved N-glycosylation consensus sequence; III, a central domain that represents the most conserved part of the
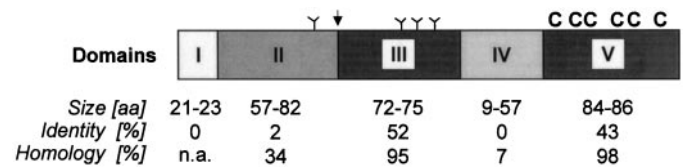

Domains	I	II	III	IV	V
Size [aa]	21-23	57-82	72-75	9-57	84-86
Identity [%]	0	2	52	0	43
Homology [%]	n.a.	34	95	7	98

Figure 2. Domain model of neurexophilins based on their primary structures. Five domains are proposed (identified by roman numerals): I, an N-terminal signal peptide; II, a variable N-terminal region that may be cleaved off during proteolytic processing; III, a highly conserved domain containing three N-glycosylation sites; IV, a linker sequence that varies in size and composition between different neurexophilins; and V, a conserved C-terminal domain with six cysteine residues that are identically spaced in all neurexophilins. The putative site of proteolytic cleavage is indicated by an *arrow above* the diagram, positions of N-glycosylation sequences are marked by *branched lines*, and the conserved cysteines are identified by *letters C*. The size range of the different domains in the neurexophilins and the sequence identities and homologies between neurexophilins in the domains are shown *below* the diagram. *n.a.*, Not applicable.

neurexophilins and contains three N-glycosylation sites (Fig. 1, *arrows*); IV, a linker sequence that is only nine amino acids in neurexophilins 1, 2, and 3 but 57 amino acids in neurexophilin 4, with the 57 amino acid linker sequence of neurexophilin 4 primarily composed of an imperfect Gly-Gly-Xxx-Leu repeat; and V, a C-terminal conserved domain that contains six identically spaced cysteine residues (Fig. 1, *asterisks*).

Neurexophilins 1 and 2 are more homologous to each other than they are to neurexophilins 3 or 4 (Fig. 1). Neurexophilin 3 is closer to neurexophilins 1 and 2 than to neurexophilin 4, and neurexophilin 4 is the most divergent neurexophilin. Many extracellular domains, for example, Ig- and EGF-like domains, are composed of sequences containing an even number of cysteine residues that are spaced in a characteristic pattern and are disulfide bonded. Such domains can be defined by consensus sequences that are present in many similar domains in a large number of proteins. With four neurexophilins, a comparable consensus sequence can now be defined. However, we were unsuccessful in detecting any sequences in the data banks that are related to the cysteine-rich or the central conserved domain of neurexophilins except for a short sequence in *Caenorhabditis elegans* (accession number, U41995), suggesting that the cysteine-rich domain of neurexophilins does not constitute a widely present extracellular motif.

Tissue-specific expression of neurexophilins

Cloning and analysis of EST data banks suggested that all four neurexophilins are expressed in humans, and neurexophilins 1, 3, and 4 are expressed in rats and mice. To obtain a more accurate assessment of the expression patterns of different neurexophilins in mouse, rat, and human tissues, we analyzed the tissue distribution of their mRNAs (Fig. 3). Although all neurexophilins were preferentially expressed in brain, we observed a remarkable variability between species. This is surprising because the sequences of the individual neurexophilins are highly conserved evolutionarily.

In mice and rats, neurexophilin 1 was expressed at high levels only in brain, whereas in humans, the strongest hybridization signals were detected in spleen (Fig. 3*A,E,I*) without a specific signal in brain. Neurexophilin 2 was detected in humans in brain and kidney but in rodents only in mouse liver; no rat tissue tested was positive (Fig. 3*B,F,J*). Neurexophilins 3 and 4 were most highly expressed in brain in all three species analyzed. However, there were again major differences between species with regard to other tissues. In humans, neurexophilin 3 mRNA was almost brain-specific; in rats, low levels could be detected in several other tissues; and in mice, high levels were present in lung, kidney, and testis (Fig. 3*C,G,K*). Similarly, mRNA for neurexophilin 4 was detected in mice only in brain, in rats also in kidney, and in humans in spleen and testis (Fig. 3*D,H,L*). Together these data show that neurexophilins are preferentially expressed in brain but that mRNAs for some neurexophilins are also present in nonneural tissues in a species-specific pattern.

The tissue distribution of an mRNA is usually not simultaneously analyzed in multiple species. It is thus unknown whether other genes, similar to neurexophilins, exhibit differences in expression pattern between species. The variations in the expression patterns of different neurexophilin mRNAs indicate either that neurexophilins have distinct, species-specific functions or, more likely, that expression of a neurexophilin can occur in a tissue without necessarily providing a function.

Proteolytic processing of neurexophilin in PC12 cells

Previous studies showed that neurexophilin is purified on immobilized α-latrotoxin in a tight complex with neurexin Iα. In this complex, neurexophilin is present as an N-glycosylated 29 kDa protein that has a protein core of ~19 kDa (Petrenko et al., 1996). In contrast, the size of unglycosylated neurexophilin predicted from the cDNA sequence would be ~35 kDa (Fig. 1). This size discrepancy of 16 kDa could be explained by two hypotheses. (1)

Neurexophilin may be physiologically processed by site-specific proteolytic cleavage in the secretory pathway. In this case, the size of the protein complexed to neurexin Iα corresponds to that of mature neurexophilin. (2) Neurexophilin may have been partially degraded during purification of neurexin Iα on immobilized α-latrotoxin.

The domain structure of neurexophilins with a highly variable N-terminal sequence resembles that of a prepropeptide (e.g., see Eipper and Mains, 1980; Jacobs et al., 1981; Noda et al., 1982; Maisonpierre et al., 1990). This supports the notion that neurexophilin is physiologically processed by proteolytic cleavage. To address this issue directly, we used an adenovirus expression system to direct synthesis of neurexophilin 1 in PC12 cells, a neuron-like cell line. The size and glycosylation of neurexophilin in the infected cells was then analyzed as a function of time. We used an adenovirus expression system in these experiments because it is very difficult to transiently transfect neuron-like cell lines with a high enough efficiency to allow analysis of expressed proteins. Permanent transfection of neuron-like cells, however, often leads to selection of dedifferentiated cells that have lost some of the more interesting neuron-like properties. In this situation, protein expression with an adenovirus or similar system offers a unique advantage in that all cells in the dish will synthesize the protein encoded by the recombinant virus.

After infection, we followed the apparent size of neurexophilin as a function of time by SDS-PAGE and immunoblotting to test whether it is processed after synthesis (Fig. 4). As a control for protein loads and for artifactual proteolysis, the same blots were probed for synaptotagmin I, a synaptic vesicle protein that is sensitive to proteolysis (Perin, 1991). In addition, cells were also infected with virus encoding neurexin Iβ and neuroligin to control for adenovirus-induced changes unrelated to neurexophilin. No neurexophilin immunoreactivity was detected in noninfected PC12 cells or in PC12 cells in the first hours after infection (Fig. 4, *lanes 1–3*). However, ~22 hr after infection, a protein of 50 kDa that was immunoreactive with neurexophilin antibodies became detectable; the size of this protein corresponded to that of glycosylated full-length neurexophilin (Fig. 4, *lanes 4–6*). With a delay of ~4 additional hours, a second protein band reactive with neurexophilin antibodies developed. This band had a size of ~29 kDa and comigrated with neurexophilin isolated in a complex with neurexin Iα on immobilized α-latrotoxin (Fig. 4, *lanes 6–9*; data not shown) (Petrenko et al., 1996). The lower band became more intense than the upper band with time, suggesting that most of the full-length neurexophilin was converted to the smaller form (Fig. 4, *lanes 10–12*).

Together these results suggest that in PC12 cells, neurexophilin expressed with a recombinant adenovirus is first produced as an N-glycosylated full-length protein and then proteolytically processed to a mature peptide with a delay of ~4 hr. However, an alternative explanation for these data would be that after the 4 hr delay, full-length neurexophilin that is not glycosylated is produced. Full-length nonglycosylated neurexophilin would have a size of ~35 kDa, and its size may not be distinguishable from that of neurexophilin purified in a complex with neurexin Iα. To eliminate this possibility, we investigated the N-glycosylation status of the different forms of neurexophilin using endoglycosidase F (Fig. 5). Endoglycosidase F treatment caused a similar shift in the apparent size of full-length neurexophilin and processed neurexophilin. Thus both the unprocessed and processed proteins were N-glycosylated, and the smaller protein must be a product of proteolytic processing.
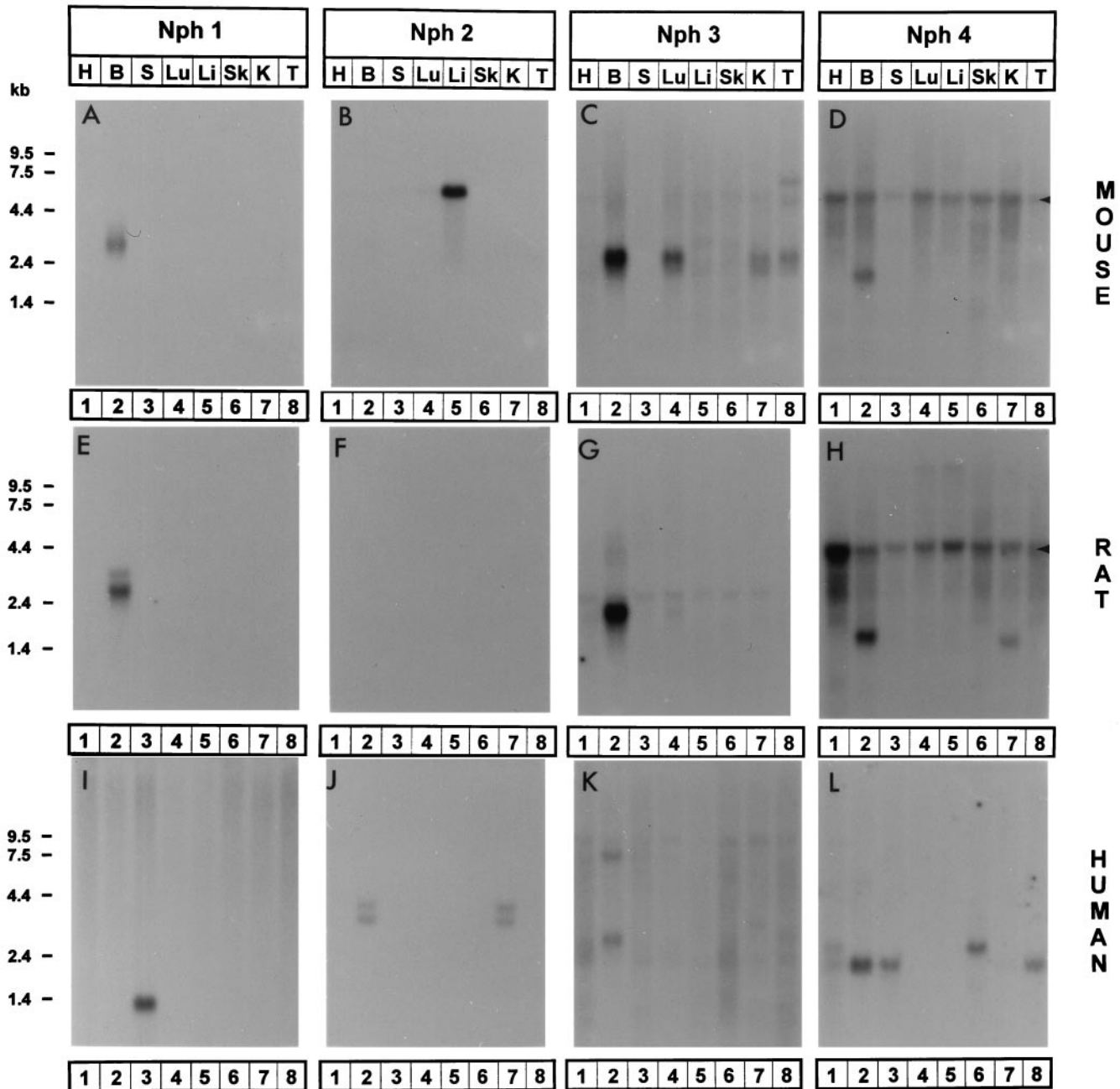
Figure 3. Tissue-specific expression of neurexophilins in rats, mice, and humans. RNA blots are hybridized with neurexophilin 1 (*Nph 1*; *A, E, I*), neurexophilin 2 (*Nph 2*; *B, F, J*), neurexophilin 3 (*Nph 3*; *C, G, K*), and neurexophilin 4 (*Nph 4*; *D, H, L*) probes. RNA blots from mouse (*A–D*), rat (*E–H*), and human (*I–L*) tissues contained total RNA from heart (*H*; *lane 1*), brain (*B*; *lane 2*), spleen (*S*; *lane 3*), lung (*Lu*; *lane 4*), liver (*Li*; *lane 5*), skeletal muscle (*Sk*; *lane 6*), kidney (*K*; *lane 7*), and testis (*T*; *lane 8*). A single multitissue RNA blot (obtained from Clontech) from each species was consecutively hybridized with probes for each neurexophilin. Rat probes were used for all hybridizations except for the blots for human neurexophilins 3 and 4 (*K* and *L*) that were hybridized with human probes. *Arrowheads* for *D* and *H* mark the position of the 28S RNA that cross-hybridizes with the glycine- and cysteine-rich rat neurexophilin 4 probe. Positions of molecular size markers are indicated on the *left*.

Cell-type specificity of proteolytic processing of neurexophilin

Inspection of the neurexophilin sequences revealed a conserved polybasic motif (KxKK) at the border between the N-terminal variable and the central conserved domains (Fig. 1). Cleavage of neurexophilin at this motif would result in a processed form with a size of ~19 kDa, suggesting that the boundary between variable and conserved domains in neurexophilin represents the site of cleavage. Similar polybasic motifs in many proteins, for example,

the insulin receptor or the surface receptor LRP (Bravo et al., 1994; Willnow et al., 1996), are recognized by ubiquitous processing proteases such as furin. In addition, many peptide hormones and neuropeptides are also processed at similar polybasic cleavage sites by specific processing enzymes that are restricted to a few specialized cell types (Thomas et al., 1986; Smeekens and Steiner, 1990; Rehemtulla and Kaufman, 1992). To investigate whether the processing of neurexophilin occurs by a ubiquitously present cellular processing enzyme or by a specialized enzyme
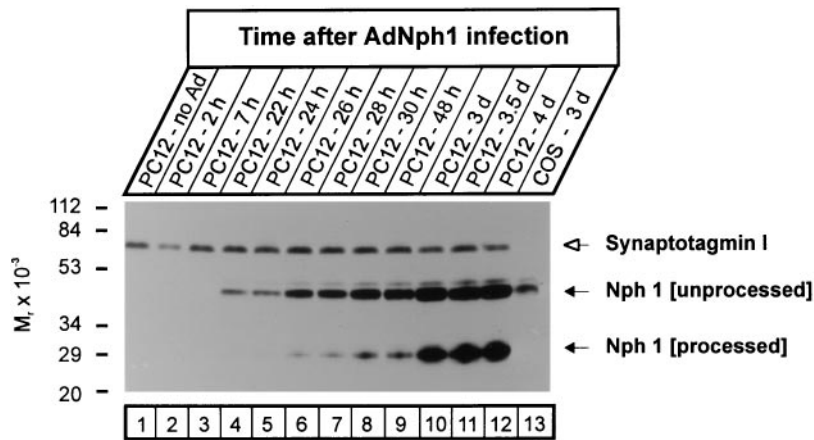
Figure 4.   Time-dependent proteolytic processing of neurexophilin 1. PC12 cells were infected with recombinant adenovirus encoding full-length neurexophilin 1 (*AdNph1*), and neurexophilin expression was analyzed as a function of time by SDS-PAGE and immunoblotting. To control for protein loads, we then reprobed the same blot for synaptotagmin I. Without adenovirus infection (*lane 1*), only synaptotagmin I but no neurexophilin can be detected. Full-length neurexophilin corresponding to the N-glycosylated form (*Nph 1* [*unprocessed*]) is detected within 24 hr of infection (*lanes 4–6*), whereas the shorter, processed form of neurexophilin (*Nph 1* [*processed*]) appears later (*lanes 7–9*) and becomes the dominant form only after 3 d (*lanes 10–12*). No processing is observed with neurexophilin expressed in COS cells (*lane 13*).
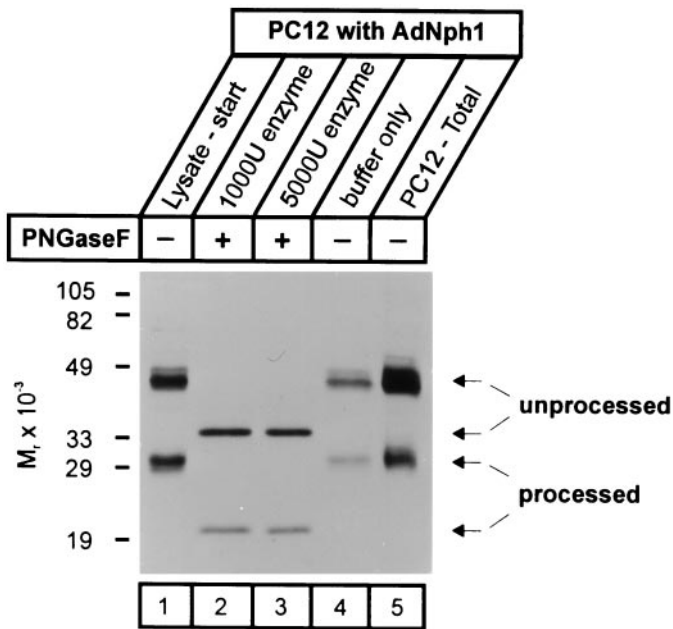

Figure 5.   N-glycosylation of neurexophilin 1 expressed in PC12 cells. PC12 cells infected with recombinant adenovirus encoding neurexophilin 1 (*AdNph1*) were lysed 2 d after infection, and lysates were treated with 1000 or 5000 units of endoglycosidase F (*PNGaseF*) or with control buffer. Samples were analyzed by SDS-PAGE and immunoblotting, demonstrating a similar shift of the unprocessed and processed forms of neurexophilin by endoglycosidase F treatment but not by control treatment. Numbers on the *left* indicate positions of molecular weight standards.

system restricted to a limited number of cells, we studied the processing of neurexophilin in two additional neuron-like cell lines (hNT and SH-SY5Y cells) and two fibroblastic cell lines (STO and COS cells), with PC12 cells as a positive control (Fig. 6). The processing of neurexophilin 1 from a 50 kDa full-length glycosylated form to a 29 kDa form of identical size was observed in hNT and SH-SY5Y cells similar to that in PC12 cells (Fig. 6*A*,*B*). Fibroblastic STO and COS cells, however, were unable to process neurexophilin expressed either by the recombinant ade-

novirus or by transfection (Fig. 6*C*,*D*), although these cells do cleave polybasic motifs in LRP, for example (Willnow et al., 1996). Thus not all cells can process neurexophilin, suggesting that a special processing protease or accessory chaperone is required.

α-Neurexins preferentially bind processed neurexophilin

Neurexophilin was discovered as a protein complexed to neurexin Iα (Petrenko et al., 1996). To investigate the binding specificity of neurexophilin to different neurexins and to evaluate the relative ability of processed and unprocessed neurexophilin to bind, we purified IgG fusion proteins of α- and β-neurexins from COS cells transfected with the appropriate expression vectors (Ichtchenko et al., 1995). IgG fusion proteins immobilized on protein A were then incubated with lysates from PC12 cells infected with neurexophilin adenovirus. PC12 lysates were recovered early after infection of the cells, at a time when most of the neurexophilin was not yet processed (Fig. 7, *lane 1*). In this manner, preferential binding of processed over nonprocessed neurexophilin was easier to evaluate. Proteins bound to the immobilized fusion proteins were analyzed by SDS-PAGE and immunoblotting using neurexophilin antibodies.

Neurexin Iβ and control IgG fusion proteins were unable to bind either unprocessed or processed neurexophilin 1 (Fig. 7, *lanes 2, 3*). Neurexin Iα and IIIα fusion proteins, however, avidly bound recombinant neurexophilin (Fig. 7, *lanes 4, 5*; data not shown). Although neurexin Iα bound both the unprocessed and processed forms of neurexophilin, the processed forms were highly enriched in the bound fraction compared with the starting material (Fig. 7, *lanes 1* vs *4* and *5*). These data suggest that only α-neurexins bind neurexophilin and that they preferentially bind processed over nonprocessed neurexophilin.

DISCUSSION

Neurexins are neuronal cell-surface proteins that are composed of three α- and three β-neurexins (for review, see Missler and Südhof, 1998). Neurexin Iα was initially discovered because it
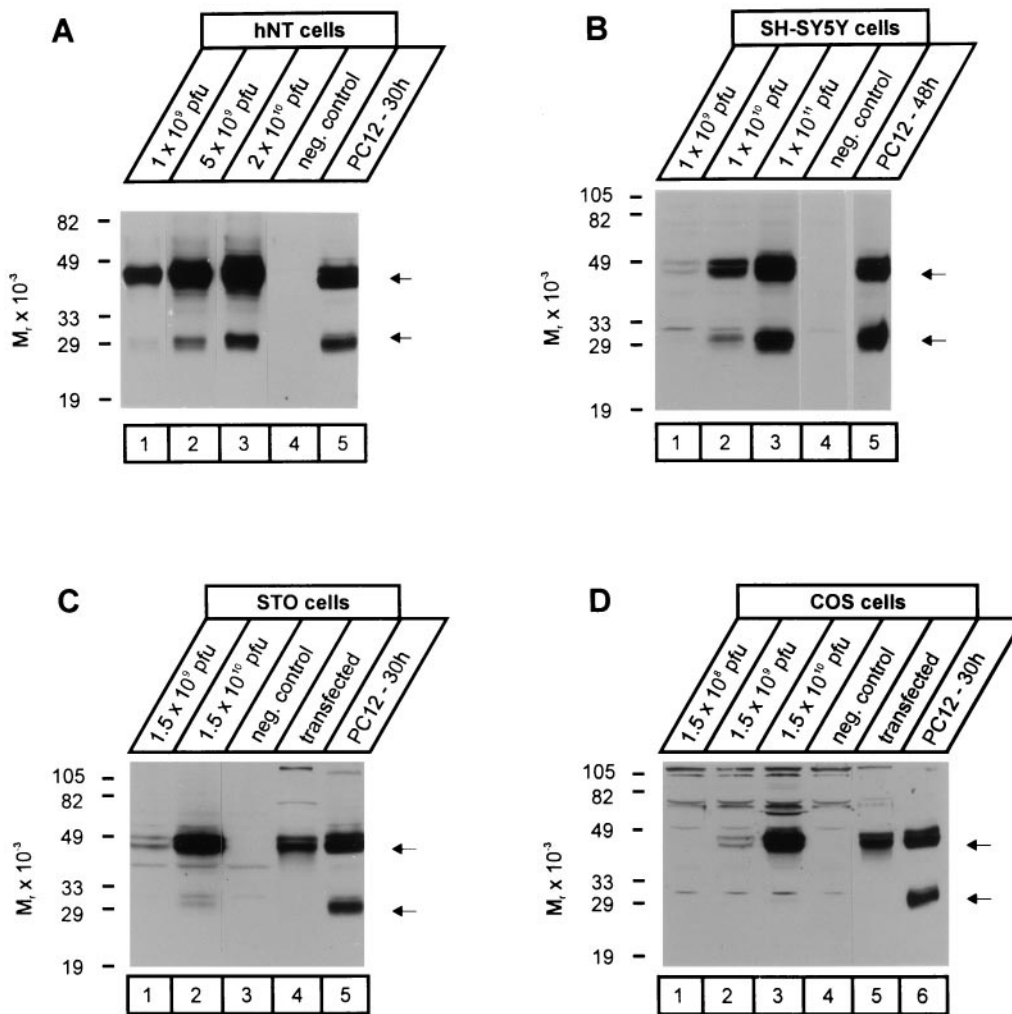
Figure 6. Proteolytic processing of neurexophilin 1 in different cell types. Cell lines of neuronal [hNT (*A*) and SH-SY5Y (*B*) cells] and non-neuronal [STO (*C*) and COS (*D*) cells] origin were infected with different concentrations (*pfu*, plaque-forming units) of recombinant adenovirus expressing neurexophilin 1. In addition, neurexophilin 1 was also expressed in the non-neuronal cells by transfection. The size of recombinant neurexophilin 1 produced in the various cell types was analyzed by SDS-PAGE and immunoblotting and compared with that produced in PC12 cells infected with the neurexophilin 1 adenovirus. Numbers on the *left* of each *panel* indicate positions of molecular weight markers; *arrows* on the *right* mark the migration of unprocessed (*top*) and processed (*bottom*) neurexophilin.

serves as a high-affinity receptor for α-latrotoxin (Ushkaryov et al., 1992; Davletov et al., 1995). Although neurexin Iα is not the only α-latrotoxin receptor and is not essential for the excitotoxic action of α-latrotoxin, neurexin Iα potentiates toxin action (Geppert et al., 1998). The α-latrotoxin receptor activity of neurexin Iα suggests that this neurexin and maybe α-neurexins in general function as receptors for signaling molecules. By contrast, at least a subset of β-neurexins performs a role as a cell adhesion molecule by binding to neuroligins (Ichtchenko et al., 1995, 1996). Binding of β-neurexins to neuroligins creates an intercellular junction flanked by the PDZ domain proteins CASK and PSD-95 (Irie et al., 1997; Nguyen and Südhof, 1997).

If neurexin Iα, as a receptor for α-latrotoxin, functions as a signaling receptor, there must be endogenous ligands for neurexin Iα and other α-neurexins. Previous studies established that a 29 kDa protein called neurexophilin is purified from brain in a tight complex with neurexin Iα and may represent an endogenous ligand for neurexin Iα (Petrenko et al., 1996). Only part of the

neurexin Iα purified on immobilized α-latrotoxin was complexed to neurexophilin, and neurexophilin was not required for α-latrotoxin binding (Davletov et al., 1995). Neurexophilin is a secreted glycoprotein that is expressed in a small subset of neurons. Its structure and tight complex with neurexin Iα suggested the possibility that neurexophilin may be a proteolytically processed novel neuropeptide that represents an endogenous ligand for neurexins. We have now investigated this possibility. The experiments reported in the current paper establish four conclusions. (1) Neurexophilins form a gene family of at least four members that constitute secreted glycoproteins and exhibit a domain structure similar to that of neuropeptides (Figs. 1, 2). (2) In all species studied, either neurexophilin 1 or 2 is expressed in brain together with neurexophilins 3 and 4. In addition, neurexophilin mRNAs are transcribed in some non-neuronal tissues in a highly species-specific pattern (Fig. 3). (3) Neurexophilin 1 and probably also other neurexophilins are N-glycosylated immediately after synthesis and proteolytically processed with a delay of
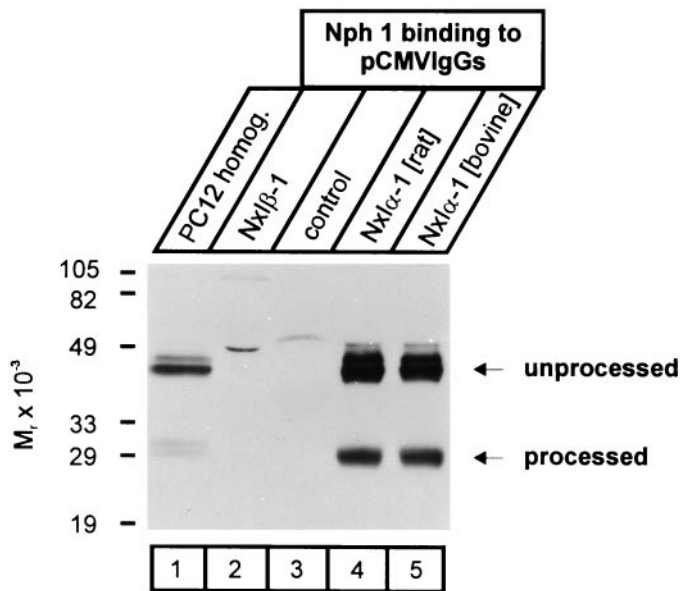
Figure 7. Binding of recombinant neurexophilin 1 to neurexin Iα. PC12 cells infected with recombinant adenovirus expressing neurexophilin 1 were incubated for 24 hr, the earliest time at which mature processed neurexophilin is detected (*lane 1*). Immobilized IgG fusion proteins of either neurexin Iβ (*NxIβ-1*; *lane 2*), a short irrelevant sequence (*control*; *lane 3*), or rat and bovine neurexin Iα (*NxIα-1*; *lanes 4, 5*) were incubated with lysates from infected PC12 cells, and bound proteins were analyzed by SDS-PAGE and immunoblotting for neurexophilin. Numbers on the *left* indicate positions of molecular weight markers. Note that both processed and unprocessed neurexophilin specifically binds only to neurexin Iα but that there is a relative enrichment of processed neurexophilin in the bound fraction.

~4 hr. Proteolytic processing is not ubiquitously performed but takes place only in neuron-like cells (Figs. 4–6). (4) Neurexophilin 1 specifically binds to neurexin Iα and IIIα but not to β-neurexins, with the processed form being preferentially bound, suggesting that neurexophilins are ligands for α-neurexins (Fig. 7).

Together with previous observations, these data provide evidence of the concept that neurexophilins constitute a novel class of neuropeptides with α-neurexins as their receptors. This concept is supported by the following findings. (1) The domain structure of neurexophilins is composed of variable N-terminal sequences and conserved C-terminal sequences and thereby resembles the prepropeptide structure of neuropeptides (e.g., see Eipper and Mains, 1980; Jacobs et al., 1981; Noda et al., 1982; Maisonpierre et al., 1990). (2) Neurexophilin 1 and possibly other neurexophilins are endoproteolytically processed after synthesis by cleavage at a defined position. (3) Neurexophilin purified from brain in a complex with neurexin Iα is exclusively present in the mature processed form (Petrenko et al., 1996; data not shown). (4) The proteolytic cleavage of newly synthesized neurexophilin occurs only in neuron-like cells and not in the non-neuronal cells tested. (5) Similar to other neuropeptides, neurexophilins are N-glycosylated and have a conserved C-terminal glycine residue that may be amidated (Murthy et al., 1986). (6) Neurexophilins are primarily synthesized in brain but exhibit species-specific expression patterns in non-neural tissues, similar, for example, to the aberrant high-level expression of NGF in the mouse salivary gland. (7) Neurexophilin 1 and possibly the other neurexophilins are synthesized in only a subset of neurons (Petrenko et al., 1996). (8) In brain, neurexophilin binds tightly to the extracellular do-

mains of α- but not β-neurexins. (9) The processed form of neurexophilin preferentially binds to α-neurexins.

If neurexophilins are neuropeptides, they could represent the endogenous ligands for α-neurexins that may activate α-neurexins in a manner physiologically similar to the mechanism by which α-latrotoxin activates neurexin Iα pathologically. Signal transduction might be mediated by CASK, a PDZ domain protein that binds to the cytoplasmic domain of neurexins (Hata et al., 1996). Because at least a subset of β-neurexins functions as cell adhesion proteins (Nguyen and Südhof, 1997), it seems likely that α- and β-neurexins have distinct functions in agreement with their different domain structures. This suggests that neurexins generally perform dual functions as signaling receptors and cell adhesion molecules.

REFERENCES

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped blast and PSI-blast: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.

Bravo R, Gleason JB, Sanchez RI, Roth RA, Fuller RS (1994) Accurate and efficient cleavage of the human insulin proreceptor by the human proprotein-processing protease furin. J Biol Chem 269:25830–25837.

Davletov B, Krasnoperov V, Hata Y, Petrenko AG, Südhof TC (1995) High affinity binding of α-latrotoxin to recombinant neurexin Iα. J Biol Chem 270:23903–23905.

Eipper BA, Mains RE (1980) Structure and biosynthesis of pro- adrenocorticotrophin/endorphin and related peptides. Endocr Rev 1:1–27.

Geppert M, Ushkaryov YA, Hata Y, Davletov B, Petrenko AG, Südhof TC (1992) Neurexins. Cold Spring Harb Symp Quant Biol 57:483–490.

Geppert M, Missler M, Petrenko AG, Südhof TC (1998) Neurexin Iα is a major α-latrotoxin receptor that cooperates in α-latrotoxin action. J Biol Chem 273:1705–1710.

Gorman C (1985) In: DNA cloning, Vol II (Glover DM, ed), pp 143–190. Oxford: IRL.

Graham FL, Prevec L (1991) Manipulation of adenovirus vectors. Methods Mol Biol 7:109–128.

Hata Y, Butz S, Südhof TC (1996) CASK: a novel *dlg*/PSD95 homologue with an N-terminal CaM kinase domain identified by interaction with neurexins. J Neurosci 16:2488–2494.

Ichtchenko K, Hata Y, Nguyen T, Ullrich B, Missler M, Moomaw C, Südhof TC (1995) Neuroligin 1: a splice-site specific ligand for β-neurexins. Cell 81:435–443.

Ichtchenko K, Nguyen T, Südhof TC (1996) Structure, alternative splicing, and neurexin binding of multiple neuroligins. J Biol Chem 271:2676–2682.

Irie M, Hata Y, Takeuchi M, Ichtchenko K, Toyoda A, Hirao K, Takai Y, Rosahl TW, Südhof TC (1997) Binding of neuroligins to PSD95. Science 277:1511–1515.

Jacobs JW, Goodman RH, Cein WW, Dee PC, Habener JF (1981) Calcitonin messenger RNA encodes multiple polypeptides in a single precursor. Nature 213:457–459.

Maisonpierre PC, Belluscio L, Squinto S, Ip NY, Furth ME, Lindsay RM, Yancopoulos GD (1990) Neurotrophin-3: a neurotrophic factor related to NGF and BDNF. Science 247:1446–1451.

Missler M, Südhof TC (1998) Neurexins: three genes and 1001 products. Trends Genet 14:20–26.

Murthy ASN, Mains RE, Eipper BA (1986) Purification and characterization of peptidylglycine α-amidating monooxygenase from bovine neurointermediate pituitary. J Biol Chem 261:1815–1822.

Nguyen T, Südhof TC (1997) Binding properties of neuroligin 1 and neurexin 1β reveal function as heterophilic cell adhesion molecules. J Biol Chem 272:26032–26039.

Noda M, Furutani Y, Takahashi H, Toyosato M, Hirose T, Inayama S, Nakanishi S, Numa S (1982) Cloning and sequence analysis of cDNA for bovine adrenal pre-pro-enkephalin. Nature 295:202–206.

Perin M (1991) Domain structure of synaptotagmin (p65). J Biol Chem 266:623–629.

Petrenko AG, Lazaryeva VD, Geppert M, Tarasyuk TA, Moomaw C, Khokhlatchev AV, Ushkaryov YA, Slaughter C, Nasimov IV, Südhof

TC (1993) Polypeptide composition of the α-latrotoxin receptor. J Biol Chem 268:1860–1867.

Petrenko AG, Ullrich B, Missler M, Krasnoperov V, Rosahl TW, Südhof TC (1996) Structure and evolution of neurexophilin. J Neurosci 16:4360–4369.

Rehemtulla A, Kaufman RJ (1992) Protein processing within the secretory pathway. Curr Opin Biotechn 3:560–565.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual, 2nd Edition. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Smeekens SP, Steiner DF (1990) Identification of human insulinoma cDNA encoding a novel mammalian protein structurally related to the yeast dibasic processing protease Kex2. J Biol Chem 265:2997–3000.

Thomas G, Herbert E, Hruby DE (1986) Expression and cell-type specific processing of human proenkephalin with a vaccinia recombinant virus. Science 232:1641–1643.

Ullrich B, Ushkaryov YA, Südhof TC (1995) Cartography of neurexins: more than 1000 isoforms generated by alternative splicing and expressed in distinct subsets of neurons. Neuron 14:497–507.

Ushkaryov YA, Südhof TC (1993) Neurexin IIIα: extensive alternative splicing generates membrane-bound and soluble forms in a novel neurexin. Proc Natl Acad Sci USA 90:6410–6414.

Ushkaryov YA, Petrenko AG, Geppert M, Südhof TC (1992) Neurexins: synaptic cell surface proteins related to the α-latrotoxin receptor and laminin. Science 257:50–56.

Ushkaryov YA, Hata Y, Ichtchenko K, Moomaw C, Afendis S, Slaughter CA, Südhof TC (1994) Conserved domain structure of β-neurexins. J Biol Chem 269:11987–11992.

Willnow TE, Moehring JM, Inocencio NM, Moehring TJ, Herz J (1996) The low-density-lipoprotein receptor-related protein (LRP) is processed by furin in vivo and in vitro. Biochem J 313:71–76.

Wimalawansa SJ (1997) Amylin, calcitonin gene-related peptide, calcitonin, and adrenomedullin: a peptide superfamily. Crit Rev Neurobiol 11:167–239.