Structures, Alternative Splicing, and Neurexin Binding of Multiple Neuroligins*

(Received for publication, July 12, 1995, and in revised form, November 8, 1995)

Konstantin Ichtchenko, Thai Nguyen, and Thomas C. Südhof†

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

Neuroligin 1 is a neuronal cell surface protein that binds to a subset of neurexins, polymorphic cell surface proteins that are also localized on neurons (Ichtchenko, K., Hata, Y., Nguyen, T., Ullrich, B., Missler, M., Moomaw, C., and Südhof, T. C. (1995) Cell 81, 435–443). We now describe two novel neuroligins called neuroligins 2 and 3 that are similar in structure and sequence to neuroligin 1. All neuroligins contain an N-terminal hydrophobic sequence with the characteristics of a cleaved signal peptide followed by a large esterase homology domain, a highly conserved single transmembrane region, and a short cytoplasmic domain. The three neuroligins are alternatively spliced at the same position and are expressed at high levels only in brain. Binding studies demonstrate that all three neuroligins bind to β-neurexins both as native brain proteins and as recombinant proteins. Tight binding of the three neuroligins to β-neurexins is observed only for β-neurexins lacking an insert in splice site 4. Thus, neuroligins constitute a multigene family of brain-specific proteins with distinct isoforms that may have overlapping functions in mediating recognition processes between neurons.

Neurexins are neuronal cell surface proteins that exhibit a high degree of diversity (Ushkaryov et al., 1992, 1994; Ushkaryov and Südhof, 1993). Three genes for neurexins have been described, each of which is transcribed from two promoters. This results in six principal neurexin transcripts that are subject to extensive alternative splicing, generating a family of thousands of differentially expressed proteins (Ullrich et al., 1995).

The diversity of the neurexins led to the hypothesis that different neurexin isoforms may have distinct binding activities and mediate recognition events between neurons. This hypothesis predicts that alternative splicing regulates the ligand interactions of neurexins and that neurexins express neurexin ligands that are specific for certain splice variants of neurexins. The interactions between these ligands and the neurexins could mediate recognition events between neurons. The description of neuroligin 1 provided the first support for this hypothesis (Ichtchenko et al., 1995). Neuroligin 1 was purified by affinity chromatography on immobilized neurexin 1Δj; cloning revealed that it constitutes a cell surface protein with a single transmembrane region and an extracellular domain homologous to esterases such as acetylcholinesterase. However, neuroligin 1 lacks an active site serine, suggesting that it is not catalytically active. Interestingly, neuroligin 1 binds tightly only to β-neurexins but not to α-neurexins, and only to those β-neurexins that lack an insert in splice site 4 (Ichtchenko et al., 1995). Together, these data suggested a model whereby neuroligin 1 and β-neurexins lacking an insert in splice site 4 mediate specific interactions between neurons. This model is supported by the recent discovery of gliotactin in Drosophila. Gliotactin also contains an esterase homology domain and a single transmembrane region and is distantly related to neuroligin 1 (Auld et al., 1995). Gliotactin is expressed transiently in glia cells during development and is essential for the formation of the peripheral blood-nerve barrier, suggesting a function in a transient cell-cell recognition event.

We now describe two new forms of neuroligin called neuroligins 2 and 3 and have systematically investigated the alternative splicing of all neuroligins. All three neuroligins are similar in structure and expressed at high levels only in brain. Furthermore, all neuroligins are alternatively spliced at the same position but with different insert patterns. Binding studies revealed that they exhibit comparable binding properties of neurexins. Thus, similar to neurexins, neuroligins form a multigene family of brain proteins that may collaborate with neurexins in mediating cell-cell interactions between neurons.

EXPERIMENTAL PROCEDURES

cDNA Cloning, Construction of Expression Vectors, and COS Cell Expression of Neuroligins and Neurexins—Sequencing of PCR products obtained with oligonucleotide primers corresponding to the neuroligin 1 sequence ELNNEILVPVIQFLGVPYAAPPT (PCR primer sequences (letters in parentheses indicate redundant positions): CGCGAGCTCAA(C,T)AA(C,T)GA(G,A)AT(A,T,C)(T,C) and GGCTGTGAGCNGC(A,G)TANGGNACNCC) resulted in the isolation of novel sequences that were related to, but distinct from, neuroligin 1. cDNA libraries were screened with oligonucleotides and with randomly labeled probes derived from the PCR sequences. Multiple overlapping clones for neuroligin 2 and 3 were isolated and sequenced. Since variability between the sequences of different cDNA clones indicated the possibility of alternative splicing, primers corresponding to the N and C termini of all three neuroligins were used in PCRs on total rat brain cDNA. cDNA libraries of the individual neuroligins were established by directly cloning the PCR products into the Clal-Sall sites of the mammalian expression vector pCMV5 (primer sequences for neuroligin 1: TCTTCCAGAGGACATGGCAGCTTCCAGGATGC and GGCTGCA-CATCGTGCTACCTCTCGTTTTGTTGTA; 2: GCTTCTGAAATCAGCAT-GTGGTTCTGGGCGGTT and TCTGTCGACCCCCTATACCCGAGTG-GTTGAGTG; and 3: CGCTTCGAACGGAACATGTGGCTGCAGTGGCTGATGGCTGAGC-GCTGGGC and CCGGTTCGAGCGCTGATTACATACCGGCTATGGGAGTG). 60 independent PCR-derived full-length neuroligin clones were analyzed for neuroligin 1, 52 for neuroligin 2, and 46 for neuroligin 3. All

* The abbreviations used are: PCR, polymerase chain reaction; PAGE, polyacrylamide electrophoresis; CHAPS, 3-[3-cholamidopropyl]-dimethylammonio]-1-propanesulfonic acid.

† To whom correspondence should be addressed.

‡ The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) U41662 and U41663.
DNA sequencing was performed with Taq DNA polymerase using fluorescently labeled primers and an ABI370 DNA sequencer. Sequence analysis was performed on a PC using PCGene software (Intelligentics). Expression vectors were created by subcloning cDNA clones containing or lacking alternatively spliced sequences into pCMV-based vectors. Construction of α- and β-neurexin IgG fusion protein constructs and production of protein by transfection in COS cells were performed as described (Ushkaryov et al., 1994), with the following expression vectors used for binding reactions: neuroligin 1, pCMVNL14; neuroligin 2, pCMVNL2-1; and neuroligin 3, pCMVNL3-17 and 3-128.

**Binding of Native Neuroligins from Brain and Recombinant Neuroligins from Transfected COS Cells to β-Neurexin IgG-Fusion Proteins—**

Rat brains were homogenized in buffer A (0.1 M NaCl, 4 mM KCl, 5 mM CaCl₂, 2.5 mM EDTA, 20 mM NaHCO₃, 20 mM Tris-HCl, pH 7.5, 0.1 mM phenylmethylsulfonyl fluoride, 1 mM leupeptin, 1 mM pepstatin A, and 1 mM aprotinin). CHAPS was added to a final concentration of 2% (w/v) and dissolved under stirring at 4 °C for 1 h. Insoluble material was removed by low (2,000 × g) and high speed centrifugations (80,000 × g). The supernatant was diluted with an equal volume of buffer A and chromatographed on different IgG proteins attached to protein A-Sepharose pre-equilibrated in buffer A containing 0.8% CHAPS. Columns were eluted with either with a gradient of buffer A to buffer B (1× NaCl, 4 mM KCl, 5 mM CaCl₂, 2.5 mM EDTA, 20 mM NaHCO₃, 20 mM Tris-HCl, pH 7.5, 0.8% CHAPS) or directly with buffer B, followed by buffer B containing 25 mM EDTA and 25 mM EGTA. Samples were analyzed by SDS-PAGE and silver staining. Coomassie staining, and immunoblotting. Neuroligins 1, 2, and 3 expressed in COS cells were analyzed similarly by homologizing COS cells approximately 72 h after transfection.

Antibodies—Four antibodies raised against synthetic peptides from the cytoplasmic domains of neuroligins were used for the current study. Peptides were coupled to keyhole limpet hemocyanin, and antibodies were produced as described (Johnston et al., 1989). The antibody that reacts with all neuroligins (L067) was raised against the C terminus of neuroligin 1 (sequence: CHPHPHPSHSTTRV) that is highly homologous to the C terminus of neuroligins 2 and 3; the neuroligin 1-specific antibody (L066) was raised against the sequence CRRILSPGGSGGVPGGG; and the neuroligin 3 specific antibody (L069) against the sequence CEAGPPHDTLRLTALPDYT.

**Miscellaneous Procedures—** SDS-PAGE, immunoblotting, and silver staining were performed using standard procedures and antibodies described previously (Johnston et al., 1989; Ushkaryov et al., 1992). RNA blots were purchased from Clontech and hybridized with uniformly labeled cDNA probes (Ushkaryov et al., 1992).

**RESULTS**

PCRs were performed with degenerate primers corresponding to a sequence from neuroligin 1, and the products were sequenced (Ichtchenko et al., 1995). The sequences obtained suggested the presence of new neuroligins in addition to neuroligin 1 and were used to isolate multiple overlapping cDNA clones. The cDNA clones encoded two new forms of neuroligin called neuroligins 2 and 3. Full-length sequences for neuroligins 2 and 3 were assembled from multiple overlapping cDNA clones (Fig. 1).

Alignment of the translated amino acid sequences of the three neuroligins revealed that they are highly homologous. Overall, the three neuroligins share 52% identical residues. In pairwise comparisons, the neuroligins are almost equally sim-
Similar to each other (neuroligin 1 and 2, 59% identity; 1 and 3, 66% identity; and 2 and 3, 59% identity; Fig. 1). The sequence similarity between the neuroligins is distributed over the whole protein, with the extracellular domains (55% identity) and the transmembrane regions (91% identity) being more conserved than the cytoplasmic sequences (31% identity). As is often observed in gene families, sequence stretches of almost 100% identity alternate with clusters of divergence. The most conserved sequences between neuroligins were observed in the esterase homology domain (see below), and the most divergent sequences in the segment linking the esterase homology domain to the transmembrane region and in parts of the cytoplasmic domains (Fig. 1).

All neuroligins contain an N-terminal hydrophobic sequence suggestive of a signal peptide. Analysis of these sequences for signal sequence cleavage sites indicated preferred cleavage.
sites that are shown in Fig. 1. In order to confirm these sites directly, we purified neuroligins 1, 2, and 3 from rat brain as a mixture by affinity chromatography on immobilized neurexin 1μ (see below) and attempted to obtain N-terminal sequences. Only a single unequivocal sequence was obtained (XXGGG-PGGGAP in single-letter amino acid code) that corresponds to the N terminus of neuroligin 2. This sequence shows that the N terminus of neuroligin 2 contains a cleaved sequence. The N-terminal sequence starts one residue after the N terminus suggested by the computer analysis of signal peptide cleavage sites (Fig. 1). However, it seems likely that the predicted cleavage site at residue 14 is correct and that the N-terminal glutamine resulting from the signal sequence cleavage is removed after cleavage because a cleavage site corresponding to residue 15 would be atypical (von Hejne, 1986). The N termini of the other two neuroligins appear to be blocked and could not be sequenced.

Sequencing of multiple independent cDNA clones for neuroligins 2 and 3 demonstrated that at a single identical position, the cDNAs either contained or lacked a sequence with an intact reading frame, suggesting that this sequence is alternatively spliced (Fig. 1). This suggestion was supported by the previously observed alternative splicing of neuroligin 1 at the same position (Ichtchenko et al., 1995). Interestingly, neuroligins 1 and 2 each contains a single type of alternatively spliced insert that is not homologous between neuroligins 1 and 2. By contrast, neuroligin 3 had two types of alternatively spliced inserts, one of which is homologous to that of neuroligin 2, whereas the second contains a sequence homologous to the alternatively spliced neuroligin 1 insert in addition to the sequence homologous to the neuroligin 2-type insert (Fig. 1).

The three variants of alternative splicing observed for neuroligin 3 with their segmental homology to the alternatively spliced sequences in neuroligins 1 and 2 suggested the possibility that neuroligins 1 and 2 have additional splice variants. Furthermore, one might expect that neuroligin 3 mRNAs should also occur with alternatively spliced inserts consisting only of the neuroligin 1-similar insert but not the neuroligin 2-similar insert. To systematically investigate these possibilities and potential additional events of alternative splicing, we performed PCRs on rat brain cDNA using oligonucleotides that amplify the complete coding region of the three neuroligins and produced cDNA libraries for each neuroligin. More than 45 independent cDNA clones of each of the three neuroligins were isolated and mapped, and some clones were sequenced and used for expression and neurexin binding studies (see below). These experiments confirmed the alternative splicing pattern deduced from the cDNA clones isolated by library screening but failed to discover additional splice variants, suggesting that the most common splice variants had in fact been identified.

Data bank searches revealed that similar to neuroligin 1, neuroligins 2 and 3 are also homologous to esterase domain proteins (Fig. 2). The esterase domain starts immediately after the signal sequence. The active site serine of esterases is changed to a glycine in all neuroligins, suggesting that the neuroligins are not catalytically active. The esterase domains of the neuroligins are much more homologous to each other than they are to other member of the esterase gene family, indicating that they form a separate gene family (Table I). Six cysteines form three intramolecular disulfide bonds in acetylcholinesterase (Sussman et al., 1991). The four cysteines forming the two N-terminal disulfide bonds are conserved precisely in the neuroligins, suggesting that the corresponding disulfide bonds are also conserved. The two cysteines of acetylcholinesterase that form the C-terminal disulfide bond are shifted to different positions in the neuroligins, indicating differences between neuroligins and acetylcholinesterase in the C-terminal disulfide bonds. In addition, neuroligins contain a conserved cysteine residue between the first and second disulfide loop. This cysteine does not appear to be used for intermolecular disulfide bonds since neuroligins have very similar migrations on reducing and nonreducing SDS-polyacrylamide gels (data not shown).

To investigate the tissue distribution of expression of different neuroligins, RNA blots were performed. Hybridization of RNAs from adult rats revealed a strongly positive signal for all three neuroligins that could only be detected in brain (Fig. 3; Ichtchenko et al., 1995). This indicates that, in adults, neuroligins are expressed at high levels only in the nervous system. Neuroligin 1 was purified and cloned based on its binding to β-neurexins. To determine if neuroligins 2 or 3 also bind to neurexins, we raised anti-peptide antibodies against each neuroligin. The three neuroligins were expressed by transfection in COS cells, and the specificity of the antibodies was examined using proteins from the transfected COS cells and from total rat brain homogenates. The results demonstrated that an antibody against the C terminus of neuroligin 1 reacted with all neuroligins, presumably because of their high degree of homology to each other in this region (top panel, Fig. 4). All other peptide antibodies, however, were specific for their respective neuroligin (lower three panels, Fig. 4). In rat brain only, the antibody specific for neuroligin 1 and the antibody against all neuroligins but not the neuroligin 2 and 3 antibodies detected a band at the sensitivity of the immunoblot shown, suggesting that neuroligin 1 is more abundant than neuroligins 2 or 3.

The antibodies were used to study the binding properties of neuroligins. For this purpose, we expressed β-neurexins as IgG-fusion proteins in COS cells and purified them from the medium as described (Ushkaryov et al., 1994). The purified β-neurexin IgG-fusion proteins immobilized on protein A-Sepharose were then used as an affinity matrix to isolate binding proteins. As shown in Fig. 5, elution of a neurexin 1μ
affinity matrix loaded with rat brain proteins with increasing salt concentrations in the presence of Ca$^{2+}$ resulted in the release of a fraction of the bound neuroligins. However, the majority of neuroligins remained attached to the matrix until Ca$^{2+}$ was removed by application of EGTA and EDTA. This result suggests that all three neuroligins bind to neurexin 1β in a Ca$^{2+}$-dependent manner. The Ca$^{2+}$-dependent binding of neuroligins 2 and 3 to neurexin 1β was confirmed using recombinant neuroligins expressed as full-length proteins, showing that it is a direct interaction (data not shown).

One of the most striking aspects of the interaction of neurexins with neuroligin 1 is that only β-neurexins and only those β-neurexins lacking an insert in splice site 4 bind neuroligin 1 (Ichtchenko et al., 1995). To determine if neuroligins 2 and 3 exhibit similar properties, we tested the binding of brain neuroligins to neurexins 1β, 2β, and 3β lacking inserts in splice site 4, and to neurexins 1β and 2β containing inserts in splice site 4 (Fig. 6). All neuroligins bound preferentially to β-neurexins lacking an insert, and all bound to all three neurexins. However, slight differences in binding were observed, especially for neurexin 2β which bound much better to neuroligin 2 than to the other two neuroligins (lane 5, Fig. 6). Together, these results extend the previous description of a splice site specific interaction of neuroligin 1 with β-neurexins to the other two neuroligins and suggest the possibility that different neurexins and neuroligins may have distinct affinities.

As described above, neuroligins are also alternatively spliced similar to neurexins, although much less extensively. In order to determine if the alternative splicing of neuroligins had an effect on their interactions with neurexins, we transfected neuroligin 3 with and without an alternatively spliced insert into COS cells and examined its interaction with neurexin 1β. No effect of alternative splicing was observed, suggesting that the alternative splicing of neuroligins does not regulate its interaction with neurexins (data not shown).

**DISCUSSION**

The current study describes the structures, alternative splicing, tissue distribution, and expression of two new neuroligins, neuroligins 2 and 3. Neuroligins are expressed at high levels only in brain where in situ hybridization localized them pri-
Marilytoneurons. Structurally, neuroligins resemble cell surface receptors and are exposed on the surface of transfected cells. They are composed of five domains: an N-terminal cleaved signal sequence, a large extracellular domain homologous to esterases, a linker domain between the transmembrane region and the esterase homology domain that may be O-glycosylated, a single transmembrane region, and a cytoplasmic tail. Sequence comparisons place the neuroligins into the large family of esterase homology domain proteins that includes thyroglobulin, acetylcholinesterase, and gliotactin. However, the neuroligins are only distantly related to these proteins and form a unique subset of this protein family (Fig. 2 and Table I). Functionally, neuroligins bind tightly to the extracellular domains of β-neurexins in a Ca\(^2+\)-dependent manner. This interaction depends on the alternative splicing of the β-neurexins since only those β-neurexins that lack an insert in splice site 4 bind tightly to neuroligins. Together, these data suggest that neuroligins are neuronal cell surface proteins that interact with a specific subset of β-neurexins.

Similar to neurexins, neuroligins are alternatively spliced, although to a much lesser extent. All neuroligins are alternatively spliced at the same position in the esterase homology domain; the position of alternative splicing maps to a loop in the crystal structure of acetylcholinesterase (Sussman et al., 1991) and to a position where Drosophila acetylcholinesterase contains an insert compared with other esterases (Fournier et al., 1988). In neuroligins 1 and 2, alternative splicing exists in the presence or absence of a single sequence that shares no homology between neuroligins 1 and 2. By contrast, neuroligin 3 has two types of alternatively spliced inserts. The first is homologous to that observed in neuroligin 1; the second contains in addition to the first insert sequence a sequence homologous to the alternatively spliced neuroligin 2 sequence. Extensive screening failed to reveal similar double inserts for the other neuroligins. The neuroligin 1-type alternatively spliced insert contains two cysteines, suggesting that it may form a disulfide-bonded ring, whereas the neuroligin 2-type insert has no cysteines (Fig. 1). Binding studies with recombinant neuroligins showed that the alternative splicing of neuroligins has no effect on their interaction with neurexins; however, this finding does not exclude the possibility that the alternative splicing may have other functional consequences.

Our experiments establish that neuroligins form a gene family encoding a series of alternative spliced proteins. The data provide further support for the notion that the interaction between neurexins and neuroligins may be important in neuronal cell-cell interactions. With three genes, the number of

---

2 B. Ullrich and T. C. Südhof, unpublished observation.
neuroligin genes mirrors that of neurexins and adds to the growing list of neuronal proteins that are expressed in a large number of isoforms. The question arises if this serves to impart distinct functional properties onto subsets of neurons that express certain combinations of neurexins and/or neuroligins, or if this is an evolutionary accident of random gene duplications. It is striking that in spite of their similarity, the different neurexins and neuroligins are on the average only 60% identical and exhibit major structural differences, suggesting an appreciable evolutionary distance and functional diversification. Although the different neuroligins have similar binding activities for neurexins, it is possible that they have different functions either due to distinct affinities for different neurexins (which could not be detected in our assays) or to interactions with other proteins. Future experiments including studies using knockout mice lacking one or several neuroligins or neurexins should help to clarify these issues.

Acknowledgments—We thank A. Roth, E. Borowicz, C. Moomaw, S. Afendis, and I. Leznicki for excellent technical assistance, Dr. C. A. Slaughter for help in the N-terminal sequencing, Dr. P. Taylor for helpful comments regarding acetylcholinesterase structure, and Drs. M. S. Brown and J. L. Goldstein for advice.

REFERENCES

Auld, V. J., Fetter, R. D., Brodie, K., and Goodman, C. S. (1995) Cell 81, 757–767
Di Lauro, R., Obici, S., Condliffe, D., Ursini, V. M., Musti, A., Moscatelli, C., and Avvedimento, V. E. (1985) Eur. J. Biochem. 148, 7–11
Fournier, D., Bride, J.-M., Karch, F., and Berge, J.-B. (1988) FEBS Lett. 238, 333–337
Han, J. H., Stratowa, C., and Rutter, W. J. (1987) Biochemistry 26, 1617–1625
Hortsch, M., Patel, N. H., Bieber, A. J., Traquina, Z. R., and Goodman, C. S. (1990) Development 110, 1327–1340
Ichtchenko, K., Hata, Y., Nguyen, T., Ullrich, B., Missler, M., Moomaw, C., and Südhof, T. C. (1995) Cell 81, 435–443
Johnston, P. A., Jahn, R., and Südhof, T. C. (1989) J. Biol. Chem. 264, 1268–1273
Nagao, T., Shimada, Y., Sugihara, A., and Tominaga, Y. (1993) J. Biochem. (Tokyo) 113, 776–780
Robbi, M., Beaufay, H., and Octave, J.-N. (1990) Biochem. J. 269, 451–458
Schumacher, M., Maujet, Y., Camp, S., and Taylor, P. (1988) J. Biol. Chem. 263, 18979–18987
Schumacher, M., Camp, S., Maujet, Y., Newton, M., and McPhee-Quigley, K. (1996) Nature 319, 407–409
Sussman, J. L., Harel, M., Frolov, F., Oefner, C., Goldman, A., Toker, L., and Silman, I. (1991) Science 253, 872–879
Ullrich, B., Ushkaryov, Y. A., and Südhof, T. C. (1995) Neuron 14, 497–507
Ushkaryov, Y. A., and Südhof, T. C. (1993) Proc. Natl. Acad. U. S. A. 90, 6410–6414
Ushkaryov, Y. A., Petrenko, A. G., Geppert, M., and Südhof, T. C. (1992) Science 257, 50–56
Ushkaryov, Y. A., Hata, Y., Ichtchenko, K., Moomaw, C., Afendis, S., Slaughter, C. A., and Südhof, T. C. (1994) J. Biol. Chem. 269, 11987–11992
von Heijne, G. (1986) Nucleic Acids Res. 11, 4683–4690