An Amino-terminal Extension Is Required for the Secretion of Chick Agrin and Its Binding to Extracellular Matrix

Alain J. Denzer, Matthias Gesemann, Beat Schumacher, and Markus A. Ruegg
Department of Pharmacology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland

Abstract. Agrin is an extracellular matrix (ECM) protein with a calculated relative molecular mass of more than 200 kD that induces the aggregation of acetylcholine receptors (AChRs) at the neuromuscular junction. This activity has been mapped to its COOH terminus. In an attempt to identify the functions of the NH2-terminal end, we have now characterized full-length chick agrin. We show that chick agrin encoded by a previously described cDNA is not secreted from transfected cells. Secretion is achieved with a construct that includes an additional 350 bp derived from the 5' end of chick agrin mRNA. Recombinant agrin is a heparan sulfate proteoglycan (HSPG) of more than 400 kD with glycosaminoglycan side chains attached only to the NH2-terminal half. Endogenous agrin in tissue homogenates also has an apparent molecular mass of >400 kD. While the amino acid sequence encoded by the 350-bp extension has no homology to published rat agrin, it includes a stretch of 15 amino acids that is 80% identical to a previously identified bovine HSPG. The extension is required for binding of agrin to ECM. AChR aggregates induced by recombinant agrin that includes the extension are considerably smaller than those induced by agrin fragments, suggesting that binding of agrin to ECM modulates the size of receptor clusters. In addition, we found a site encoding seven amino acids at the NH2-terminal end of agrin that is alternatively spliced. While motor neurons express the splice variant with the seven amino acid long insert, muscle cells mainly synthesize isoforms that lack this insert. In conclusion, the cDNAs described here code for chick agrin that has all the characteristics previously allocated to endogenous agrin.

Proper synaptic transmission requires a high degree of cell specialization at the site of contact between pre- and postsynaptic cells. The best studied chemical synapse is the neuromuscular junction (NMJ), where the nerve terminal of a motor neuron synapses on a muscle fiber. The ingrowing motor axon initiates in muscle fibers a local accumulation of several proteins of the cytosol, the cell membrane and the extracellular matrix (ECM; for a review see Hall and Sanes, 1993). The formation of these specializations in the muscle fiber is triggered by molecules released from the nerve (Frank and Fischbach, 1979; Dahm and Landmesser, 1991). The best-characterized molecule known to be involved in this process is agrin. When added to cultured myotubes, agrin induces the aggregation of AChRs and other molecules concentrated at the NMJ (Nikitin et al., 1987; Wallace, 1989). This process is believed to be mediated by binding of agrin to specific signaling receptors on the muscle cell surface (for reviews see Fallon and Hall, 1994; Bowe and Fallon, 1995). Active agrin is synthesized by motor neurons and is transported to the nerve terminal where it is released and becomes incorporated into the nascent muscle cell basal lamina (reviewed by McMahan, 1990). Nevertheless, agrin is also found in nonsynaptic basal laminae of muscle, in kidney and capillaries (Reist et al., 1987; Fallon and Gelfman, 1989; Godfrey, 1991; Rupp et al., 1991; Ruegg et al., 1992), suggesting that it may also have other functions.

The molecular mass of agrin deduced from cDNAs in rat, chick, and the marine ray is more than 200 kD (Rupp et al., 1991; Tsim et al., 1992; Smith et al., 1992). However a COOH-terminal, 45-kD fragment of agrin is sufficient to induce AChR clustering on cultured myotubes with high efficacy (Gesemann et al., 1995). Within this active fragment, two sites (called A and B in chick and ray, y and z in rat) whose positions and amino acid sequences are conserved among the three species are subject to alternative mRNA splicing. When expressed in eukaryotic cells, only recombinant agrin isoforms with inserts at both sites are active in inducing AChR clusters, while those lacking in-
serts are much less or not at all active (Ruegg et al., 1992; Ferns et al., 1992, 1993; Gesemann et al., 1995). In vivo, transcripts encoding agrin isoforms are differentially distributed. The most active ones are expressed by cholinergic neurons, such as motor neurons of the spinal cord and ciliary ganglia neurons, and the inactive ones are highly expressed by nonneuronal cells like myotubes and Schwann cells (Ruegg et al., 1992; McMahan et al., 1992; Hoch et al., 1993; Thomas et al., 1993; Smith and O‘Dowd, 1994; Ma et al., 1994, 1995). Thus, expression of specific isoforms is most likely the basis of the observed difference between muscle cell-derived and motor neuron-derived agrin in inducing AChR aggregates (Reist et al., 1992; Cohen and Godfrey, 1992).

As in other extracellular matrix proteins, agrin consists of several repeated regions. There are nine Kazal inhibitor-like domains that are also found in follistatin, six EGF-like domains, and three globular domains homologous to COOH-terminal domains of laminin α-chains. Homologous domains in other ECM molecules have been implicated in particular functions, such as cell-cell recognition and growth factor binding, which has led to many proposals for additional functions of agrin (e.g., Patthy and Nikolies, 1993). The NH₂-terminal half of agrin also contains several potential sites for carbohydrate attachment. There are five potential N-glycosylation sites, two serine/threonine-rich regions, and at least six potential glycosaminoglycan (GAG) chain attachment sites.

As a first step in searching for additional functions of agrin isoforms, we aimed to express full-length chick agrin and to characterize its biochemical properties in more detail. When we transfected COS-7 cells with the previously described agrin cDNA (Tsim et al., 1992) the recombinant protein was not targeted to the ER and hence, was not secreted from the cells. Here we present a chick agrin cDNA, whose coding sequence at the 5' end is extended by ~350 bp. Agrin with this extension is secreted from transfected cells. The new sequence includes a stretch of 15 amino acids that is highly homologous to a peptide previously characterized in a heparan sulfate proteoglycan (HSPG) of bovine kidney (Hagen et al., 1993). Only recombinant agrin that includes this extension binds to Matrigel™, a solubilized mixture of ECM components. Recombinant chick agrin, depending on the inserts at site A and B, is active in inducing AChR aggregates on cultured myotubes. However, AChR clusters are considerably smaller than those induced by fragments of agrin. In addition, we have found a new 21-bp-long site at the 5' end that is alternatively spliced. At early stages of synaptogenesis, motor neurons synthesize agrin mRNA with the 21-bp-long insert and the majority of agrin transcripts expressed in muscle lack this insert. In conclusion, these 5' extended constructs code for chick agrin isoforms that appear to fulfill all criteria to be full-length.

Materials and Methods

Polymerase Chain Reaction

Conditions for PCR were as described elsewhere (Gesemann et al., 1995). Numbering of oligonucleotides is according to the sequences published by Tsim et al. (1992). Oligonucleotides of new 5' sequences are negatively numbered. All PCR products used to generate expression constructs were cloned into PCR II cloning vector (Invitrogen, San Diego, CA) and sequenced to ensure that no mutations were introduced by DNA amplification.

mRNA Isolation, Reverse Transcription and PCR

Poly A⁺ RNA was isolated with the Micro-FastTrack mRNA Isolation Kit (Invitrogen) according to the manufacturer's advice. RACE experiments (Frohmann et al., 1988; Loh et al., 1989) were performed with the 5' RACE System (GIBCO BRL, Gaithersburg, MD) according to the manufacturer's advice. The oligonucleotide used for the reverse transcription reaction was as366 (CCTGTCGCTAGGGTGGGATAA) and the oligonucleotide as223 (ATCCCTGGCAGGATTCG) served as primer in the subsequent PCR reaction. Reverse transcriptase-polymerase chain reaction was conducted as described in Ruegg et al. (1992) except that Superscript™ II reverse transcriptase (GIBCO BRL) was used. To construct cDNAs encoding complete agrin isoforms, first strand synthesis was primed with as366 and subsequent PCR was conducted using EcoRI s-289 (GCATAGAATTCGGCTGGGCGAGTGG) as the sense primer and as232 as the antisense primer (see below).

Chicken Genomic Library Screening

A chicken genomic library from adult chicken liver (Clontech, Palo Alto, CA) was screened according to the manufacturer's advice. Recombinant phage (~9 × 10⁵) from the amplified EMBL-3 library were screened with the 32p-labeled DNA probe (nucleotides 134 to 547; for location see Fig. 3). Hybridization was at 42°C in 5× SSC, 0.1% SDS, 5× Denhardt's reagent, 30% formamide, and 0.1 mg/ml salmon sperm DNA. Filters were washed at a maximal stringency of 0.2× SSC, 0.1% SDS at 65°C. Two rounds of dilution were applied for cloning. DNA of positive recombinant phage was prepared as follows: 450 ml of 50 mM Tris-HCl, pH 7.5, 10 mM NaCl, 8 mM MgSO₄, 0.01% gelatin were inoculated with a single positive plaque. 60 μl of maltose-induced NM-538 (Sambrook et al., 1989) in 10 mM MgSO₄ (OD₆₀₀ = 2) were added and incubated for 15 min at 37°C. After adding 30 ml of growth medium, 10 mM MgSO₄ cells were incubated at 37°C until lysis. 300 μl chloroform were added and incubation was continued for another 45 min. The lysate was centrifuged for 15 min at 15,000 g to pellet cell debris. The supernatant was incubated with approximately 100 μg DNAse (Boehringer Mannheim, Mannheim, Germany) and 100 μg RNase (Boehringer Mannheim) for 1 h at 37°C. 7 ml of 20% PEG (6,000) in 2.5 M NaCl were added and incubation was continued for 1 h on ice. After centrifugation for 15 min at 13,000 g, the pellet was dissolved in 450 μl 10 mM Tris-HCl, pH 8.0, 0.5% SDS, 10 mM EDTA, 10 mM NaCl and 20 μg protease K (Boehringer Mannheim). Digestion was performed for 30 min at 60°C. After phenol/chloroform extraction, DNA was precipitated with ethanol. Inserts were excised by SalI digestion, analyzed on agarose gel, and subcloned into Bluescript KS (Stratagene, La Jolla, CA).

Alternative mRNA Splicing

To examine the relative abundance of agrin mRNA, randomly primed cDNAs were subjected to PCR using either primers s86 (CCCAACC-CAACGAGCTGATG) and as232, which flank the novel splice site, or primers s377 (TTTATGATTTAGGCGAGTACATG) and as511 (CT-TCTGTTTTTGATGCTCAG), which flank the B site. PCR reaction was performed in 50 μl containing 10% of reverse transcription mixture and 3 × 10⁵ cpm of 32p-end labeled oligonucleotide as232 or s537 (Sambrook et al., 1989). After 30 rounds of amplification (94°C, 1 min; 56°C, 2 min; 72°C, 2 min; final extension step of 10 min), 30% of the reaction mixture was fractionated by electrophoresis on a 8% polyacrylamide gel. The relative amount of each PCR product was determined by analysis on a PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

Expression Constructs

To distinguish expression constructs from recombinant protein, cDNA constructs are named with the prefix p followed by the name of the protein they encode. All partial chick agrin cDNA constructs, with the exception of pcFulIP, pcFulPlP, pcNAgrin, and pcAgrin (see below), contain the ER signal sequence of hemagglutinin of the avian influenza virus described by Tsim et al. (1992).

cDNA constructs encoding the NH₂-terminal part of chick agrin were made as follows: for future use a myc epitope (Evan et al., 1985) was inserted at the COOH terminus of chick agrin. This was achieved by PCR

The Journal of Cell Biology, Volume 131, 1995 1548
using s377 as the sense primer (TTTGATGGTAGGACGTACTAGT), as5945 myc Xbal as the antisense primer (GGCCCTCTAGAACTAAAT- TCAATGTCCTCCTCGCTGATTAACTITFGTTCTITGGCTA/GG/A-
using s5377 as the sense primer (TTTGATGGTAGGACGTACATG),
NH2-terminal part of cFull hs, a BamHI restriction site was introduced at 5'
uct was cleaved with endonucleases Sail and Xbal and inserted into
as5945 myc XbaI as the antisense primer (GGCCCTCTAGAACTAAT-
gous NH2-terminal construct (pcNFull~) containing the putative ER sig-
replacing the SalI/XhoI fragment of pcNFull hs with that of pcFull pS.
The analogous NH2-terminal construct (pcNFull~) containing the putative
of the region encoding the NH2-termi-
constructed by replacing the Sail/Xhol fragment of pcNFull~ with that of pcFull~
Construction of complete chick agrin cDNA was accomplished by
to isolate overlapping cDNAs (Tsim et al. 1992) were fused. This
Construction of complete chick agrin cDNA was accomplished by
to isolate overlapping cDNAs (Tsim et al. 1992) were fused. This
Tissues from E11-E15 White Leghorn chick embryos were harvested and
on ice with a 15 ml Dounce-homogenizer in the appropriate
volume of PBS containing 1 mM PMSF, 1 μg/ml peptatin, 1 μg/ml anti-
pain, 10 μg/ml aprotinin, 1 μg/ml leupeptin, and 1.75 μg/ml benzamidin.
Homogenized tissues were subjected to SDS-PAGE on a 3-12% gradient
100 μl) were added and incubated for 3 h at 37°C. Samples, released by adding 100 μl SDS-PAGE sample buffer and boiling, were then analyzed by SDS-PAGE on 3-12% gradient gel followed by fluorography.

Antibodies
To raise polyclonal antisera against cFull~AB159, the procedures described by Gesemann et al. (1995) were used.

Tissue Culture and Transfections
COS-7 cells (Gluzman, 1981) were cultured in DME (GIBCO BRL) sup-
pplemented with 10% newborn calf serum, 10 mM sodium pyruvate, 100 U/ ml penicillin, and 100 μg/ml streptomycin (all from GIBCO BRL). COS-7 cells were transiently transfected with the DEAE-dextran–based method as described by Cullen (1987). Chick myotube cultures were prepared as described (Gesemann et al., 1995).

Metabolic Labeling, Immunoprecipitation, and Deglycosylation
Metabolic labeling of intracellular proteins and immunoprecipitation fol-
Immunoprecipitation of agrin in the conditioned medium was done as
follows: 12 h after transfection, growth medium was replaced by 1 ml of
labeling medium (25% DME, 65% MEM without methionine, 10% IgG-depleted FCS, 10 mM sodium pyruvate, 100 U/ml penicillin, 100 μg/ml streptomycin; all from GIBCO BRL), and 50 μCi/ml [35S]methionine. 48-72 h after transfection, 0.5 ml of the conditioned medium was subjected to
Immunoprecipitation of agrin in the conditioned medium was done as
follows: 12 h after transfection, growth medium was replaced by 1 ml of
labeling medium (25% DME, 65% MEM without methionine, 10% IgG-depleted FCS, 10 mM sodium pyruvate, 100 U/ml penicillin, 100 μg/ml streptomycin; all from GIBCO BRL), and 50 μCi/ml [35S]methionine. 48-72 h after transfection, 0.5 ml of the conditioned medium was subjected to
Immunoprecipitation of agrin in the conditioned medium was done as
follows: 12 h after transfection, growth medium was replaced by 1 ml of
labeling medium (25% DME, 65% MEM without methionine, 10% IgG-depleted FCS, 10 mM sodium pyruvate, 100 U/ml penicillin, 100 μg/ml streptomycin; all from GIBCO BRL), and 50 μCi/ml [35S]methionine. 48-72 h after transfection, 0.5 ml of the conditioned medium was subjected to
Immunoprecipitation of agrin in the conditioned medium was done as
follows: 12 h after transfection, growth medium was replaced by 1 ml of
labeling medium (25% DME, 65% MEM without methionine, 10% IgG-depleted FCS, 10 mM sodium pyruvate, 100 U/ml penicillin, 100 μg/ml streptomycin; all from GIBCO BRL), and 50 μCi/ml [35S]methionine. 48-72 h after transfection, 0.5 ml of the conditioned medium was subjected to

To study properties of the entire chick agrin, the overlapping cDNAs isolated by Tsim et al. (1992) were fused. This cDNA construct is expected to encode full-length agrin.
which we called cFullp (for chick full-length with putative signal sequence; Fig. 1). COS-7 cells were transiently transfected with this cDNA construct to obtain recombinant protein. The presence of cFullp in the conditioned medium was tested by immunoprecipitation after labeling with [35S]methionine. No immunoreactive protein was detected in the supernatant with either mAb 5B1 (Reist et al., 1987; data not shown) or a polyclonal antiserum raised against the 95 kD, COOH-terminal fragment of chick agrin (Gesemann et al., 1995; Fig. 2), indicating that cFullp is not secreted from COS-7 cells. No agrin-like immunoreactivity, as assayed by immunohistochemistry, was found associated with the cell surface of transfected COS cells (data not shown).

Soluble secretory molecules usually contain a signal sequence located at their NH2 terminus that targets protein synthesis to the endoplasmatic reticulum (for review see Verner and Schatz, 1988). Such sequences are characterized by a stretch of 16-26 residues with a polar, basic NH2 terminus and a central hydrophobic core of 7-15 residues (von Heijne, 1983; Perlman and Halvarson, 1983). Because the NH2-terminal sequence of the cDNA of Tsim et al. (1992) does not meet these criteria, we replaced it with the signal sequence of hemagglutinin. The latter sequence is capable of targeting protein synthesis of agrin fragments to the secretory pathway (Tsim et al., 1992; Ruegg et al., 1992; Gesemann et al., 1995). As shown in Fig. 2, cFullh (for chick full-length with hemagglutinin signal sequence; Fig. 1) accumulates in the medium of transfected cells. As reported previously (Gesemann et al., 1995), it has an apparent molecular mass is more than 400 kD, indicating the high content of carbohydrates. Molecular weights in kD of standard proteins are indicated.

Figure 1. Structure of chick agrin and of the constructs used in this study. Conserved sites for glycosaminoglycan side chain attachment and potential N-glycosylation are shown. Positions, designation and protein sequence of alternatively spliced inserts are given. On the right of each construct, the name and the isoforms used in this study are indicated. For future use, a myc-tag (Evan et al., 1985) was engineered to the COOH terminus of some constructs.

Figure 2. Recombinant agrin is secreted from COS cells after replacement of the putative signal sequence by the signal sequence of hemagglutinin. COS-7 cells were transfected with cDNA described by Tsim et al. (1992), encoding full-length agrin (cFullp) or with cDNA construct where the first 38 amino acids at the NH2-terminus had been replaced by the signal sequence of hemagglutinin (cFullh). A mock transfection (Mock) provided a control. Conditioned medium from [35S]methionine labeled COS cells was immunoprecipitated with anti-agrin antiserum, analyzed on 3-12% SDS-PAGE, and proteins were visualized by fluorography. No agrin-like protein is detected with cFullp while cFullh is efficiently secreted into the medium. As reported previously (Gesemann et al., 1995), its
This result shows that the NH₂-terminal sequence of amplification of cDNA ends; Frohmann et al., 1988; Loh

full ps does not serve as a signal sequence in COS cells.

position 119 homology between chick and rat agrin begins only at Lys 158 (L). These sequence data are available from GenBank/EMBL/DDBJ under accession number U35613.

Parent molecular mass of 400-600 kD on SDS-PAGE. This result shows that the NH₂-terminal sequence of cFullp does not serve as a signal sequence in COS cells.

In an attempt to isolate the endogenous signal sequence of chick agrin, we performed RACE experiments (rapid amplification of cDNA ends; Frohmann et al., 1988; Loh et al., 1989) on poly A⁺ RNA from E5 to E6 chick spinal cord. Agrin-derived primers as366 were used for reverse transcription, and as232 for the subsequent anchored PCR were both designed to anneal at the 5' end of the agrin mRNA where chick and rat agrin are still highly homologous (see Fig. 3). We analyzed more than 40 individual PCR products. The longest PCR product comprised sequences from position 107 through 620 (Fig. 3). In all the clones analyzed, an additional thymidine residue was inserted at position 419 (V, Fig. 3). This insertion neutralizes the in-frame stop codon upstream of the initiator methionine. Several attempts with altered conditions did not yield longer RACE products (data not shown). One reason may be that the secondary structure of agrin mRNA induced a premature stop of reverse transcription (Sambrook et al., 1989).

We therefore screened a chick genomic library with a DNA probe spanning nucleotides 134 to 547 (for the location see Fig. 3). Out of 9 × 10⁶ phages screened, two positive clones were analyzed. In both clones, an in-frame ATG was found 16 nucleotides upstream of the longest PCR product. This ATG is part of a consensus for initiation of translation (Kozak, 1986) and is followed by 24 mainly hydrophobic amino acids indicative of a signal sequence (von Heijne, 1983; Perlman and Halvarson, 1983), but no new initiator methionine was found. Several attempts with altered conditions did not yield the expected size and could be re-amplified using the primer s-368 and a nested primer (data not shown).

**Recombinant Agrin with the Extension Is Secreted**

To generate a cDNA construct that covers the entire coding sequence of chick agrin, RT-PCR was performed with the sense oligonucleotide EcoRI_s-289, spanning the putative initiator methionine, and the antisense oligonucleotide as232 (see Fig. 3). The PCR product was then fused to the cDNA encoding cFullp. The recombinant protein encoded by this construct was called cAgrin (Fig. 1). To show that cAgrin was targeted to the secretary pathway of COS cells, we investigated the endo-1,3-N-acetylglucosaminidase (endo H) sensitivity of freshly synthesized protein. Endo H removes N-linked carbohydrates that become attached to the protein backbone in the ER. A decrease of the apparent molecular mass upon endo H treatment indicates that the protein is targeted to the ER. To see this shift more clearly, we transfected COS cells with cDNA constructs encoding only the NH₂-terminal half (termed cNAgrin; see Fig. 1), which contains all the potential sites for N-linked glycosylation. After a pulse of [³⁵S]methionine for 30 min, the cells were lysed and agrin-like proteins were immunoprecipitated. After incubation with endo H, agrin was analyzed on SDS-PAGE. As shown in Fig. 4, the construct with the signal sequence of hemagglutinin, cNFullp, is sensitive to endo H, indicating that the proteins had reached the lumen of the ER where they became N-glycosylated. In contrast, cNFullp is only weakly expressed and it is not sensitive to endo H (Fig. 4). The relative molecular mass of cNAgrin is also shifted by endo H digestion (Fig. 4). These results show that the new NH₂-
Chick agrin containing the novel NH2-terminal sequence is targeted to the endoplasmatic reticulum. COS-7 cells that were transiently transfected with cDNA constructs encoding NH2-terminal agrin fragments (see Fig. 1) were pulsed with [35S]methionine for 30 min. Recombinant agrin was then isolated from the COS cell lysates by immunoprecipitation with anti-agrin antiserum. Immunoprecipitates were incubated in buffer alone (−) or in the presence of endo H (+) to determine the presence of N-linked carbohydrates. Proteins were analyzed on a 4–10% SDS-PAGE. The apparent molecular mass of both cNFull hs and cN-Agrin is decreased upon Endo H treatment, indicating that the proteins had reached the lumen of the ER. In contrast, cN-Full p~ is only weakly expressed and is not sensitive to endo H. The higher molecular mass of cN-Agrin compared to both, cNFull hs and cN-Full p~ is consistent with its 123-amino acid-long extension.

Terminal extension of cAgrin targets biosynthesis to the ER. In addition, they demonstrate that agrin encoded by the previously proposed full-length chick agrin cDNA is not targeted to the ER confirming that its NHE-terminal sequence does not serve as signal sequence.

**Glycosaminoglycan Side Chain Attachment**

Several lines of evidence indicate that agrin is a HSPG: (a) Recombinant cFull with a calculated molecular mass of 206 kD, when synthesized by COS-7 and 293 cells, has an apparent molecular mass between 400 and 600 kD (Gesse-mann et al., 1995; Fig. 2 and 7). (b) Cloning of a previously characterized major HSPG from chick brain (Halfter, 1993) has revealed that this HSPG is chick agrin (Tseng et al., 1995a). (c) Tryptic peptides derived from a HSPG isolated from bovine renal tubular basement membranes are highly homologous to chick agrin (Hagen et al., 1993; see below).

When we tested the sensitivity of cFull with chondroitinase ABC and keratanase, its molecular weight was not decreased (data not shown), while heparitinase treatment decreased the apparent molecular mass (Fig. 5). Agrin contains at least six potential GAG chain attachment sites with the consensus sequence E/DGSGE/D or SGXG (Bourdon et al., 1987; Zimmermann and Ruoslahti, 1989). Three sites are located in the 95-kD, COOH-terminal part of agrin and three sites, whose positions are conserved across species, are within the NH2-terminal half (see Fig. 1). To see which of these attachment sites in fact carry GAG side chains, COS cells were transiently transfected with cDNA constructs, encoding cAgrin, cFull, and the NH2-terminal half of cAgrin (c95; see Fig. 1). As expected, incubation with heparitinase resulted in the reduction of the apparent molecular mass of both cFull and cAgrin (Fig. 5). While cNFull, the NH2-terminal portion of cFull, was deglycosylated with heparitinase, the COOH-terminal fragment was not sensitive to the enzyme (Fig. 5). From this we conclude that HS-GAG chains are only attached to the NH2-terminal portion of agrin.

**Agrin in Tissue Homogenates**

Independent evidence confirms that agrin is a HSPG in both mammals and birds. Hagen et al. (1993) isolated a HSPG from basement membranes of bovine kidney and subsequently determined amino acid sequences of three tryptic peptides. As shown in Fig. 6, two peptides are similar to both chick and rat agrin. Peptides 2 and 3 are homologous to sequences between the first and the second follistatin (PSTI)-like domain and to sequences preceding the first follistatin (PSTI)-like domain of agrin, respectively. Peptide 1 maps however to the extension presented here. These data provide additional evidence that the novel sequences described in this report are indeed contained in mature agrin and that they are conserved be-
tides (single letter code) of a HSPG from bovine renal tubular basement membranes (BTBM; Hagen et al., 1993) with rat and chick agrin sequences. Conserved amino acids are shaded. Note that peptides 2 and 3 are homologous to both chick and rat agrin. Peptide 1 maps to a region encoded by the novel 5' extension (box) and hence is homologous only to chick. The position of the initiator methionine of Tsim et al. (1992) is indicated (A).

Figure 6. Amino acids of the novel 5' sequence are conserved between chick and bovine agrin. Alignment of three tryptic peptides (single letter code) of a HSPG from bovine renal tubular basement membranes (BTBM; Hagen et al., 1993) with rat and chick agrin sequences. Conserved amino acids are shaded. Note that peptides 2 and 3 are homologous to both chick and rat agrin. Peptide 1 maps to a region encoded by the novel 5' extension (box) and hence is homologous only to chick. The position of the initiator methionine of Tsim et al. (1992) is indicated (A).

The calculated molecular mass of the protein backbone of chick agrin that includes the extension is 225 kD and agrin synthesized by transfected cells has an apparent molecular mass of more than 400 kD. However, agrin-like proteins purified from basal lamina extracts of different tissues and several species appear in two discrete doublets with molecular mass of 150/135 and 95/70 kD (Nitkin et al., 1987; Godfrey et al., 1988a,b; Godfrey, 1991). Similarly, Hagen et al. (1993) reported a calculated relative molecular mass of the core protein after heparitinase treatment of 145 and 125 kD. Hence, agrin seems to be degraded during purification. To examine whether the high molecular weight form is found in vivo and to determine whether proteolytic fragments of the reported size are also detected in freshly homogenized tissue, we performed Western blot analyses. To prevent proteolytic degradation, tissues were homogenized on ice in the presence of several protease inhibitors. Tissues examined were embryonic day 14 (E14) brain (Fig. 7, lanes B), spinal cord (Fig. 7, lanes S), hindlimb muscle (Fig. 7, lanes M), E10 vitreous humor (Fig. 7, lanes V), and E15 liver (Fig. 7, lanes L). Conditioned medium of 293 cells that had been stably transfected with cDNA encoding cFullhs provided a positive control (Fig. 7, lanes C). Agrin-like proteins were detected with antiserum raised either against the 95 kD, COOH-terminal part of agrin (Fig. 7 A), or against cFullhs (Fig. 7 B). No immunoreactivity was observed with the corresponding preimmune sera (data not shown) except in the vitreous humor where the preimmune serum of the anti-c95 antiserum recognized two distinct bands of ~100 and 125 kD (* in Fig. 7 A, lane V). Agrin-like immunoreactivity was associated with a high molecular weight band in all tissues examined except liver (Fig. 7). The high molecular weight agrin was most abundant in the vitreous humor, followed by brain, muscle and spinal cord. It was seen more clearly with antibodies raised against cFullhs that carries HS-GAG chains (Fig. 7 B). Since this band was also detected with antibodies against c95 (Fig. 7 A), we conclude that the antigen recognized by anti-cFullhs antibo-

Figure 7. Agrin in vivo is a high molecular weight protein of 400-600 kD. Western blot analysis of different tissue homogenates using either antisera raised against the 95 kD, COOH-terminal half of agrin (A) or cFullhs (B). Samples correspond to: conditioned medium of 293 cells stably transfected with cDNA encoding cFullhs (lane C), E14 brain (lane B), E14 and E15 spinal cord (lane S), E14 hindlimb muscle (lane M), vitreous humor of E10 eyes (lane V), and E15 liver (lane L). No immunoreactive proteins were detected with the preimmune sera, except in vitreous humor, where the anti-c95 preimmune serum recognized two protein bands (* in A, lane V). Except in liver, agrin-like immunoreactivity in the high molecular weight range is detected in all the tissues examined, suggesting that agrin in vivo is also a proteoglycan. Note that immunoreactive proteins of smaller size are present in most tissues examined. Arrows in lane V indicate agrin-like proteins that have similar molecular mass as agrin fragments purified from Torpedo californica (see Nitkin et al., 1987).
Figure 8. Tissue distribution of the new splice variant. (A) Nucleotide and deduced amino acid sequence of cDNAs around the novel splice site. Numbering is according to Fig. 3. (B) Autoradiogram of PCR products derived from randomly reverse transcribed mRNA of E5-E6 chick spinal cord (spinal cord) and of cultured chick myotubes (muscle). PCR was performed using oligonucleotides flanking the novel splice site. Control PCR using no template (1), cDNA with 21 bp (2), cDNA without 21 bp (3), or a mixture of both cDNAs (4). Motor neurons of total spinal cord (T) were enriched by panning spinal cord cells with mAb SC1 (S), by fractionating them on a metrizamide gradient (M), or a combination of both methods (S+M). Finally, motor neurons were also determined the relative amount of agrin mRNA containing inserts at site B (data not shown). The relative abundance of individual splice variants for site B was similar to that reported by Smith and O'Dowd (1994) for single ciliary ganglia neurons and obtained earlier with the same preparation (McMahan et al., 1992; Honig, L. S., M. J. Werle, S. E. Horton, and M. A. Ruegg, manuscript in preparation). In summary, we conclude that splicing of agrin mRNA at the novel site is regulated. While motor neurons express agrin mRNA lacking this insert, muscle cells synthesize mainly agrin mRNA including this insert.

**Binding of Recombinant Agrin to Extracellular Matrix Requires the NH2-terminal Extension**

Agrin was originally purified from basal lamina extracts of the electric organ of _Torpedo californica_ (Nitkin et al., 1987). Moreover, agrin-like immunoreactivity remains associated with synaptic basal lamina for at least three weeks after degeneration of the cells at the NMJ (Reist et al., 1987).
Figure 9. Binding of recombinant agrin to extracellular matrix requires the NH₂-terminal extension. Double immunofluorescence staining of transiently transfected COS-7 cells grown on Matrigel™. The constructs and agrin isoforms used are indicated on the left of each row (see Fig. 1 for structure of the constructs). Extracellular agrin (nonpermeabilized) was visualized by incubating COS cells with antibodies directed against the 95 kD, COOH-terminal fragment at 4°C for 30 min and an appropriate, fluorescein-conjugated secondary antibody. To identify transfected cells, cells were permeabilized and agrin was subsequently visualized with mAb 5B1 and a rhodamine-conjugated secondary antibody (permeabilized). No deposition of agrin in Matrigel™ is observed with c95ₐₐB₈ and cFullhsₐ₈ (A and B), while agrin is deposited in a gradient with both constructs that include the extension, cAgrin₀ₐ₈₈ and cAgrin₇ₐ₈₈ (C and D). Bar, 40 μm.
Hence, in addition to inducing AChR aggregates, agrin also must bind to basal lamina (McMahan, 1990). To localize the region in agrin that confers its binding to ECM and to test whether alternative mRNA splicing affects the binding, transiently transfected COS cells were grown overnight on MatrigelTM, a solubilized basement membrane extracted from the Engelbreth-Holm-Swarm (EHS) mouse sarcoma (Kleiman et al., 1982). Deposition of agrin isoforms was assayed by extracellular staining with antibodies directed against the 95-kD COOH-terminal fragment of agrin. COS cells were then permeabilized and stained with mAb 5B1 (Reist et al., 1987) to identify transfected cells. No deposits in MatrigelTM are detected with cells that express the 95-kD COOH-terminal fragment of agrinAABB (Fig. 9 A) although intracellular staining shows that c95AABB is synthesized (Fig. 9 A'). The cDNA construct encoding recombinant agrin, cFullbSA4B8, that does not include the 5' extension shows the same staining pattern as c95AABB (Fig. 9, B and B'). In contrast, cAgrin7A4B8, a splice variant that includes the 7-amino acid insert at the new site, is efficiently deposited in the MatrigelTM predominantly in the vicinity of transfected COS cells (Fig. 9 C). Exclusion of the 7-amino acid insert at the new site did not change the staining pattern on MatrigelTM (Fig. 9 D, but see below). Similarly, no effect of splicing at sites A and B was observed (data not shown). No staining was observed on the cell surface of COS cells. This contrasts results by Campanelli et al. (1991) where full-length rat agrin remained cell surface of COS cells. This contrasts results by Cam-

To quantify the binding of recombinant agrin to MatrigelTM and to see whether alternative splicing at the novel site may influence this binding, we next compared the concentration of recombinant protein in the conditioned medium of transiently transfected COS cells that were either grown on normal tissue culture dishes or on MatrigelTM-coated dishes. As shown in Table I, no significant difference in the concentration of recombinant agrin on either substrate was seen with c95AABB and cFullbAABB. In contrast, the concentration of recombinant agrin that remained soluble on MatrigelTM-coated dishes was less than 50% with both constructs that include the 5' extension. Moreover, in two independent experiments, considerably more cAgrin0AABB was bound to MatrigelTM than cAgrin7AABB. Hence, alternative splicing at the novel site may influence binding of agrin to extracellular matrix.

**The NH$_2$-terminal Extension Influences Size But Not Extent of Agrin-induced AChR Clusters**

AChR-aggregating activity of agrin is strongly regulated by the alternatively spliced inserts A and B (Ruegg et al., 1992; Ferns et al., 1992, 1993; Gesemann et al., 1995). The dependence of activity on these inserts was observed with 95-kD COOH-terminal fragments and those that included the NH$_2$-terminal part of agrin (cFullbB; Gesemann et al., 1995). To test whether recombinant agrin isoforms with the extension were also active in AChR aggregation in an isoform-specific manner we incubated cultured chick myotubes for 16 h with an excess of recombinant protein. AChRs were visualized using rhodamine-labeled α-bungarotoxin. No AChR clustering was detected after incubation with conditioned medium of mock-transfected cells (Fig. 10 A). Agrin isoforms that lacked the A and the B inserts also did not induce receptor clustering (cFullbA0B0; Fig. 10 C; cAgrin7A0B8; Fig. 10 E; cAgrin0A0B2; Fig. 10 G). As shown previously, clustering of AChRs was clearly induced by 500 pM of c95AABB (Fig. 10 B) and cFullbAABB (Fig. 10 D). Agrin isoforms that included the extension were also active in inducing AChR clusters (cAgrin7AABB; Fig. 10 F; cAgrin0AABB; Fig. 10 H). AChR clusters induced by these constructs were however smaller and more frequent (Fig. 10, F and H) than those induced by agrin fragments (Fig. 10, B and D). The average size of AChR clusters induced by cFullbAABB in 20 myotube segments was 26.2 ± 2.3 μm$^2$ (mean ± SEM), while induction by cAgrin7AABB yielded clusters with an average size of 10.8 ± 0.6 μm$^2$ (mean ± SEM). To determine whether the change in size was based on decreased AChR-aggregating activity, we measured the area of AChR clusters with a computerized image analysis system. As shown in Fig. 10 I, there was no significant difference between active agrin isoforms that have the extension and their corresponding fragments. The distribution of AChR clusters (aligned versus dispersed; compare Fig. 10, F and H) sometimes varied between cAgrin7AABB and cAgrin0AABB. However, we could not reproduce this phenomenon consistently. In summary, these results show that AChR-aggregating activity of complete chick agrin depends on splice sites at sites A and B and that the most NH$_2$-terminal sequence of agrin alters the size of the induced clusters but not the overall extent of the induction.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Plastic</th>
<th>Matrigel</th>
</tr>
</thead>
<tbody>
<tr>
<td>c95AABB</td>
<td>100.0 ± 3.4</td>
<td>99.2 ± 3.2</td>
</tr>
<tr>
<td>cFullbAABB</td>
<td>100.0 ± 4.8</td>
<td>100.7 ± 2.4</td>
</tr>
<tr>
<td>cAgrin7AABB</td>
<td>100.0 ± 1.7</td>
<td>45.2 ± 1.2</td>
</tr>
<tr>
<td>cAgrin0AABB</td>
<td>100.0 ± 3.4</td>
<td>22.3 ± 1.8</td>
</tr>
</tbody>
</table>

COS cells were transfected with cDNAs encoding the indicated agrin constructs. One day after transfection, cells were replated in the same density on regular tissue culture dishes (plastic) or on MatrigelTM-coated dishes (Matrigel). After 18 h, conditioned medium was collected and the concentration of recombinant agrin was determined by ELISA as described in Materials and Methods. Concentration of agrin with cells grown on regular tissue culture dishes was set as 100%. Values are means ± SEM of two independent experiments with triplicate cultures for each condition. The concentration of c95AABB and cFullbAABB does not change, indicating that these constructs do not bind to MatrigelTM. In contrast, both cAgrin7AABB and cAgrin0AABB bind to MatrigelTM. Note also the difference in binding to MatrigelTM between agrin constructs with different inserts at the novel splice site.
Agrin with the NH$_2$-terminal extension induces smaller AChR clusters. Fluorescence micrographs of cultured chick myotubes labeled with rhodamine-α-bungarotoxin. Myotubes were incubated for 16 h with conditioned medium of COS cells, that had been mock-transfected (A) or with cDNAs encoding C95A4BS (B), cFullhSAoB0 (C), cFullhSA4B8 (D), cAgrin7n0B0 (E), cAgrin7A4B8 (F), cAgrin0g0B0 (G) and cAgrin0a4Bs (H). The concentration of recombinant agrin was always 500 pM. Only agrinn4B8 isoforms induced AChR aggregates (B, D, F, and H), whereas agring0ao isoforms were inactive (C, E, and G). Active agrin including the extension (cAgrin7ggBs, F; cAgrin0A4BS, H) induced smaller clusters. Bar, 50 μm. (/) Quantification of AChR aggregation. Due to the smaller size of AChR clusters induced by active agrin isoforms that include the extension AChR aggregates with the longer axis >2 μm are included. Activity is expressed as percentage area of AChR clusters. Each data point represents mean ± SEM of duplicate cultures, where 10 myotube segments were counted in each. The control corresponds to myotubes incubated with conditioned medium of mock-transfected COS cells. Results of one representative experiment are shown.

Discussion

Differences between Chick and Rat Agrin

We report here the heterologous expression of chick agrin comprising all the functional properties that have been postulated for endogenous agrin (McMahan, 1990). This was achieved by ligating the novel 5’ end sequences to the cDNAs described by Tsim et al. (1992). When grown on noncoated tissue culture dishes, two different cell lines transfected with this construct secrete the recombinant protein into the medium (Figs. 5 and 7). This contrasts the results of Campanelli et al. (1991) where recombinant full-length rat agrin remained associated with the cell surface of transfected COS and CHO cells. Comparison of the protein sequences of chick and rat agrin suggests that the difference resides in the NH$_2$ terminus. Although chick and rat agrin are highly conserved (60% identity), no homology is found in the previously postulated putative signal peptides (Rupp et al., 1991; Tsim et al., 1992). The 80% identity of the 15-amino acid peptide of bovine kidney HSPG with a corresponding stretch in the 5’ extended region in chick (Fig. 6) shows that the formerly postulated signal sequence of chick agrin is in fact part of the mature protein. This, together with our finding that the extension is important for the binding of agrin to ECM, makes it very likely that rat agrin also contains a homologous stretch. The published NH$_2$-terminal sequence of rat agrin (Rupp et al., 1991) is however completely different to our extension and to the bovine peptide and may therefore be derived from a rat agrin mRNA that encodes an alternative 5’ sequence.

The calculated relative molecular mass of chick agrin is 225 kD. When expressed in COS and 293 cells, recombinant chick agrin has a relative molecular mass of 400 to 600 kD (Fig. 5) due to the attachment of carbohydrates. Similarly, agrin-like immunoreactivity in tissue homogenates has the same apparent molecular mass (Tsen et al., 1995a; Fig. 7). Full-length rat agrin expressed by COS cells and agrin-like protein in rat tissue homogenates has however an apparent molecular mass of ~200 kD (Rupp et al., 1991; Campanelli et al., 1991). In addition, ray and chick agrin-like proteins in earlier studies appear as distinct protein bands between 150 and 70 kD (Nitkin et al., 1987; Godfrey et al., 1988a,b; Godfrey, 1991). The difference in the molecular mass observed earlier and our results suggests that the previously described immunoreactive proteins result from degradation. Although we observed several smaller protein bands in our Western blots, the presence of agrin-like protein at a high molecular mass indicates that agrin in vivo can be released from cells without proteolytic processing. One possible explanation for the absence of the high molecular mass bands in earlier studies may be that the antibodies used were directed either against the nonglycosylated, COOH-terminal part of agrin or against fusion proteins expressed in Escherichia coli. As shown in Fig. 7, such antibodies have a lower avidity for the highly glycosylated agrin.

Possible Role of Agrin as a Heparan Sulfate Proteoglycan

We show here that sites within the NH$_2$-terminal half of agrin carry N-linked carbohydrates and HS-GAG side chains (Figs. 4 and 5), and that the COOH-terminal end lacks these glycosylations. The attachment of GAG side chains is not necessary for agrin’s activity to induce AChR clusters (Gesemann et al., 1995). Several lines of evidence suggest however that HSPGs expressed by the muscle cells...
do play a role (Wallace, 1991; Ferns et al., 1992, 1993; Gordon et al., 1993). We have therefore recently postulated that agrin may bind to HSPGs expressed in muscle cells and that these HSPGs may help to present agrin to its postulated signal-transducing receptor (Gesemann et al., 1995). Because agrin expressed by muscle cells is a HSPG (Fig. 7) and because it accumulates at AChR clusters (Lieth et al., 1992), muscle agrin itself may be such a helper protein. Indeed, the 95-kD fragment of the active isoform agrin$_{AB}$ binds weakly to the longer, inactive recombinant cFull$_{AB}$, or the HS-GAG chains (Gesemann, M., A. J. Denzer, and M. A. Ruegg, unpublished observation).

Agrin's role as a HSPG may be similar to those described for other HSPGs. They are believed to play a role in the scaffold for the attachment of various ECM components and to modulate interactions with cells during morphogenesis (for a review see Timpl, 1993). The presence of HS-GAG chains together with a similar arrangement of structural motifs makes agrin a homologue to perlecan. Deposition of agrin-like immunoreactivity on a similar substrate has also been observed with spinal cord neurons (Dahm and Landmesser, 1991). This is also the time when motor neurons express high levels of SCI (also known as DM-GRASP or BEN; Pourquie et al., 1990; Burns et al., 1991). In the purest preparation of motor neurons, 97% of agrin mRNA contains the 21-bp insert. In this cell fraction more than 95% are motor neurons (Bloch-Gallego et al., 1991). Hence, the 3% of agrin mRNA lacking the 21-bp insert most likely represents contamination by nonmotor neurons and we conclude that, early in synaptogenesis, chick motor neurons synthesize agrin mRNA that contains the 21-bp insert on the other hand, in postsynaptic cells, such as E12 (HH stage 38; Hamburger and Hamilton, 1951) hindlimb muscle and myotubes grown in culture, the majority of agrin mRNA lacks the 21-bp insert. E10 dorsal root ganglia also synthesize predominantly agrin mRNA containing the 21-bp insert and Schwann cells of the sciatic nerve have a pattern similar to muscle cells (Denzer, A. J., and M. A. Ruegg, unpublished data). This suggests that neurons other than motor neurons synthesize agrin mRNA containing the 21 bp while nonneuronal cells show a preference for the splice variant lacking this insert. Indeed, Tsen et al. (1995b) have obtained similar results in the central nervous system, suggesting that nonneuronal cells express mainly agrin isoforms lacking the 21-bp insert and neurons synthesize agrin mRNA that contains the 21-bp insert. We have not closely looked at functional differences between the splice variants. Our finding that the isoforms seem to bind to Matrigel with different strength (Table I) suggests that this insert may regulate the interaction of agrin with ECM. We are currently aiming at identifying the binding partners of agrin in the ECM and studying in detail whether splicing at the novel site affects binding.

We thank Drs. W. B. Adams and E. Reinhard for critically reading the manuscript; Dr. H. Tanaka for providing SC1 hybridoma and Dr. P. Sonderegger for the chick genomic library.

This work was supported by a grant No. 31-33697.92 of the Swiss National Science Foundation.

Received for publication 12 June 1995 and in revised form 13 September 1995.

References


