Neuroglycan C, a Part-time Proteoglycan, in the Central Nervous System

Sachiko AONO, Hiroomi KEINO, Yoshihito TOKITA, Shinobu YAMAUCHI and Atsuhiko OOHIRA
Department of Perinatology, Institute for Developmental Research, Aichi Human Service Center

Received, 17 January 2000; accepted, 8 February 2000.

Abstract: Neuroglycan C (NGC) is a 150-kDa transmembrane chondroitin sulfate proteoglycan with a 120-kDa core glycoprotein and is expressed predominantly in the brain. The core protein is divided into five structurally different domains; an N-terminal domain to which chondroitin sulfate chain(s) may be attached, a cluster of acidic amino acids, a cysteine-rich domain with an EGF-like motif, a transmembrane domain, and a C-terminal cytoplasmic domain with two potential phosphorylation sites for protein kinase C. The mouse and human NGC genes comprise six exons, and are approximately 17 kb in size. The NGC gene is assigned to chromosomal band 9F1 in mouse, 3p21.3 in human, and 8q32.1 in rat. NGC is expressed on the surface of many neuronal cells in the cerebrum, whereas it is expressed only on the Purkinje cells in the cerebellum. The structure of NGC changes from a proteoglycan form to a nonproteoglycan form without chondroitin sulfate side chains as the cerebellar development proceeds. Immunohistological observation suggests that NGC is involved in selective synaptogenesis on the Purkinje cells in the cerebellar development.

Key words: neuroglycan C, chondroitin sulfate, proteoglycan, EGF, central nervous system

INTRODUCTION

Proteoglycans are a group of proteins that bear covalently bound sulfated glycosaminoglycan chains. They are located at strategic sites such as the cell surface and extracellular space in various animal tissues including the central nervous system.

In the vertebrate central nervous system, there are many species of proteoglycans with different structural features. Some of them are obtained from the soluble fraction of the brain, others are from the membrane fraction either in transmembrane form or glycosylphosphatidylinositol-linked form. The occurrence of a variety of proteoglycans in the brain could be due to the existence of a large number of cell types that constitute many neuronal circuits, and to their multiple roles in various cellular events including mitogenesis, migration, differentiation, axonal outgrowth and synaptogenesis.

Recently, we found a new transmembrane chondroitin sulfate proteoglycan, named neuroglycan C (neural proteoglycan with chondroitin sulfate, NGC), that is expressed in the brain, especially on the surfaces of neuronal cells, but not in other tissues in the rat and human. Although the precise physiological functions are not known at present, considering that NGC is expressed exclusively in the brain, especially in the immature brain where the neuronal circuits are actively being formed, this proteoglycan may play some roles in neuronal circuit formation.

In this review, we describe the structure of NGC core protein, organization and chromosomal assignment of the mouse NGC gene, and distribution of NGC in the brain. We will also discuss the possible biological functions of NGC in the neuronal circuit formation of the cerebellum.

Structure of NGC Core Protein

NGC is a 150-kDa transmembrane chondroitin sulfate proteoglycan with a 120-kDa core glycoprotein that was originally isolated from the early postnatal rat brain. The structure of NGC core protein is deduced based on the sequence of rat cDNA, and the protein is divided into five structurally different domains; an N-terminal domain to which chondroitin sulfate chain(s) may be attached, a cluster of acidic amino acids, a cysteine-rich domain with an EGF-like motif, a transmembrane domain, and a C-terminal cytoplasmic domain with two potential phosphorylation sites for protein kinase C (Fig.1). There are also potential phosphorylation sites for casein kinase 2 in the extracellular domain. Both the sequence and the multidomain structure are quite different from those of any known proteoglycan families.

The NGC cDNAs of human and mouse have already been isolated. The cDNA and amino acid sequences and the domain-structure are highly conserved in these species; the homologies of rat NGC in amino acid sequence are 86.4% and 94.3% with the human and mouse counterparts, respectively.

In the mouse brain, there existed three splicing...
variants (designated NGC-I, -II, and -III) of NGC\(^8\). NGC-I is the major isoform and corresponds to the counterpart of rat NGC. Only the N-terminal portion of pro-NGC, including the signal peptide, is structurally different between NGC-I and NGC-II. NGC-III has a peptide insertion, which is composed of 27 amino acid residues and gives a new putative phosphorylation site for protein kinase C, in the cytoplasmic domain of NGC-I. mRNA for NGC-II is also expressed in the rat and human brains (unpublished observations).

A transmembrane protein whose EGF-like domain and cytoplasmic segment are highly related to rat NGC (85% amino acid identity) is identified in the chicken central nervous system and designated chicken acidic leucine-rich EGF-like domain containing brain protein (CALEB)\(^9\). CALEB is composed of a 140/130-kDa doublet, an 80-kDa band, and a chondroitin sulfate-containing 200-kDa component. CALEB has a larger extracellular domain and a smaller cytoplasmic domain than NGC\(^10\). The difference in size of the extracellular domain between these two molecules is mainly due to the insertion of a leucine-proline-rich segment.

The putative signal sequence is shown as a black box, the cluster of acidic amino acids is shaded with dots, and the transmembrane domain is diagonally striped. The potential site of glycosaminoglycan attachment is indicated by a solid circle, and the putative phosphorylation site by protein kinase C in the cytoplasmic domain are indicated by open triangles. A closed triangle indicates the position of the insertion composed of 27 amino acid residues.

**Fig. 1. Schematic representation of the structure of rat NGC core protein.**

Structure of NGC Gene

The NGC gene (CSPG5) of mouse is comprised of six exons and has a size of approximately 17 kb\(^8\) (Fig. 2). Human NGC gene shows almost the same structure (unpublished observation). There are two alternative splicing sites. First, exon 1 covers the 5'-untranslated region and a part of the coding region common to NGC-I and -III mRNAs, and exon 1' covers them unique to NGC-II mRNA. Second, exon 5 codes an insertion unique to NGC-III in the cytoplasmic domain. No introns exist between exons 1' and 2, and between exons 4 and 5.

Exon 2 encodes the chondroitin sulfate-attachment domain, the acidic amino acid cluster, and a part of the EGF-like module. Exon 3 codes the remaining part of the EGF-like module, and the transmembrane domain. The cytoplasmic domain is encoded by exons 4 and 6 in the cases of NGC-I and II, and exons 4, 5, and 6 in the case of NGC-III.

**Chromosomal Mapping of NGC Gene**

By fluorescence in situ hybridization, CSPG5 is assigned to chromosomal band 9FL in mouse\(^9\), 3p21.3 in human\(^3\), 8q32.1 in rat, and 4q26 in Chinese hamster (unpublished observations). A computer search reveals that no mutant animals nor hereditary diseases whose responsible genes are assigned to these chromosomal regions have been reported.

**NGC in the Cerebrum**

On Northern blots, the NGC cDNA hybridizes with a single transcript of 3.1 kb in rat\(^6\) and 2.4 kb in human\(^7\) that is detectable in the brain but not in other tissues such as the kidney, liver, lung, and muscle. The amount of NGC in the rat brain begins to increase from perinatal stages, reaches a maximum level around 20 days after birth, and thereafter, decreases. In the mature brain, it is approximately half of the peak level\(^6\).

Immunohistological studies reveal that this proteoglycan exists on the surface of developing neuronal cells throughout the cerebral cortex of young rat. As the brain matures, the intensity of immunostaining of the cortex decreases, suggesting that NGC plays some roles in

---

**Fig. 2. Structure of genomic DNA of mouse NGC.**

Schematic diagram shows the scaled relationship of the introns and exons of mouse NGC gene.
neuronal circuit formation in the immature brain where the neuronal circuits are actively being formed\textsuperscript{10}.

It has been shown by in vitro antibody perturbation experiments that CALEB, probably a chicken homolog of NGC, participates in neurite formation in a permissive environment as a novel member of the EGF family of differentiation factors\textsuperscript{9}. Since the homology between CALEB and rat NGC is very high as mentioned above\textsuperscript{7}, NGC could have the same function as CALEB.

**NGC in the Cerebellum**

Western blot analysis demonstrates that, although NGC occurs in a proteoglycan form in the developing cerebellum of the mouse, most NGC in the mature cerebellum does not bear chondroitin sulfate, indicating that NGC is a typical part-time proteoglycan in the cerebellum\textsuperscript{8}.

Immunohistochemical studies show that only the Purkinje cells are immunostained with NGC antibodies in the mouse cerebellum at various developmental stages. In the immature Purkinje cells, NGC is immuno-localized to the soma and thick dendrites on which the climbing fibers form synapses, but not to the thin branches on which the parallel fibers form synapses. In the matured cerebellum, the reaction products are observed in spots on the large dendrites of the Purkinje cells, whereas the soma is hardly stained. Electron microscopic observation shows that the spots are the hypolemmal cisternae, which supply membrane to the plasma membrane, in the large dendrites of the Purkinje cells\textsuperscript{8}. The distribution pattern of CALEB in the chicken cerebellum is very similar to that of NGC in the mouse cerebellum\textsuperscript{9}.

As described above, the Purkinje cells form synapses with the climbing fibers at the thick stems and with the parallel fibers at the thin branches of the dendrites\textsuperscript{12}. Considering that the developmental change in the localization of NGC on the Purkinje cells correlates well with synaptogenesis of the climbing fiber system with the Purkinje cells, NGC may be involved in this selective synaptogenesis in the developing cerebellum. The proteoglycan form of NGC may mediate the adhesion and synaptogenesis of the climbing fibers or inhibit the adhesion of the parallel fibers to the Purkinje cell dendrites more effectively than the non-proteoglycan form (Fig. 3).

Neuregulins are known as a typical member of the neural EGF family. Neuregulin-$\beta$ increases the expression of a subunit (NR2C) of NMDA receptors by innervation of the glutamatergic mossy fibers to the internal granule cells in the cerebellum\textsuperscript{12}. Likewise, NGC may regulate the expression of some receptors in the cerebellum by innervation of the climbing fibers to the Purkinje cells.

**Fig. 3. Expression of NGC on the immature cerebellar Purkinje cell.**

NGC is expressed on the soma and thick dendrites, but not on the thin branches of the Purkinje cell. The Purkinje cell forms synapses (open circles) with the climbing fibers and the parallel fibers. GC: granule cell; PC: Purkinje cell; CF: climbing fiber; PF: parallel fiber.

NGC also changes its structure from a proteoglycan form to a nonproteoglycan form as the rat retina matures (Inatani et al., unpublished observation). In the case of CALEB, a similar developmental change in the structure is reported in the retina of chick by Schumacher et al.\textsuperscript{10}, supporting the idea that CALEB is a chicken homolog of NGC.

It has been reported that CALEB binds to the extracellular matrix glycoproteins, tenascin-C and tenascin-R. As a part of the ectodomain of CALEB, which probably offers the binding sites to tenascin-C and -R, is quite similar in structure to that of NGC, NGC may also bind to these glycoproteins. A search for the binding molecules of NGC is presently in progress in our laboratory.

**CONCLUSION AND PERSPECTIVES**

Considering that NGC is expressed exclusively in the brain, especially in the immature brain where the neuronal circuits are actively formed, this proteoglycan
may play some roles in neuronal circuit formation. It is of interest that NGC changes its structure from a proteoglycan form to a nonproteoglycan form which does not bear chondroitin sulfate side chains with the development of both the cerebellum and the retina. The introduction of chondroitin sulfate chains to the NGC core protein could be closely related to the maturation of the nervous tissues. The regulatory mechanism of the polysaccharide substitution remains to be examined. In addition, identification of the molecules binding to NGC with or without chondroitin sulfates will lead to a better understanding of NGC functions. Production of NGC-gene targeting animals should also help to elucidate the role of NGC in the development and maintenance of the brain.

Acknowledgement: This work was presented at a symposium at the 46th meeting for the Japanese Matrix Club in Nagoya, June, 1999. Research conducted at the authors' laboratory, which is cited in this review article, was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Culture and Sports of Japan.

References