

### Proteoglycan form and function: A comprehensive nomenclature of proteoglycans

### Renato V. lozzo<sup>1</sup> and Liliana Schaefer<sup>2</sup>

 Department of Pathology, Anatomy and Cell Biology and the Cancer Cell Biology and Signaling Program, Kimmel Cancer Center, Sidney Kimmel Medical College at Thomas Jefferson University, Philadelphia, PA 19107, USA
Pharmazentrum Frankfurt/ZAFES, Institut für Allgemeine Pharmakologie und Toxikologie, Klinikum der Goethe-Universität Frankfurt am Main, Frankfurt am Main, Germany

Correspondence to Renato V. lozzo and Liliana Schaefer: renato.iozzo@jefferson.edu; schaefer@med.uni-frankfurt.de http://dx.doi.org/10.1016/j.matbio.2015.02.003 Edited by R. Sanderson

#### Abstract

We provide a comprehensive classification of the proteoglycan gene families and respective protein cores. This updated nomenclature is based on three criteria: Cellular and subcellular location, overall gene/protein homology, and the utilization of specific protein modules within their respective protein cores. These three signatures were utilized to design four major classes of proteoglycans with distinct forms and functions: the intracellular, cell-surface, pericellular and extracellular proteoglycans. The proposed nomenclature encompasses forty-three distinct proteoglycan-encoding genes and many alternatively-spliced variants. The biological functions of these four proteoglycan families are critically assessed in development, cancer and angiogenesis, and in various acquired and genetic diseases where their expression is aberrant.

© 2015 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

#### Introduction

It has been nearly 20 years since the original publication of a comprehensive classification of proteoglycan gene families [1]. For the most part, these classes have been widely accepted. However, a broad and current taxonomy of the various proteoglycan gene families and their products is not available. In contrast to the classification of glycosaminoglycans (GAGs), primarily based on the chemical structure of their repeating disaccharide units, classifying proteoglycans is a much more complex task [2]. We propose a comprehensive and simplified nomenclature of proteoglycans based on three criteria including: Cellular and subcellular location, overall gene/protein homology, and the presence of specific protein modules within their respective protein cores. Whereas the first two attributes have been utilized in the past for various nomenclatures, the third attribute is of more recent development and represents a sort of "intrinsic signature" for various protein cores. Indeed,

modular design is based on the simple concept that protein cores are made up of finite units, like pieces of Lego. The units represent a minimum level of organization and a module can be thought of as a functional domain that affects cell-matrix dynamics. Another key feature is that each module/functional unit can be stable and can fold on its own, without being part of the large precursor protein. Thus, a module is a self-contained component. An example of this is the LG3 domain of endorepellin, the C-terminal globular-like domain of perlecan, which has recently been crystallized [3]. Below, we will critically assess the field of proteoglycans which now encompass forty three distinct genes and a much higher number of proteoglycans due to alternative splicing, thereby providing a very rich and biologically-active group of molecules. As hyaluronan and the enzymes involved in the synthesis and degradation of various GAGs are not covered in this review, readers are referred to recent reviews covering these closely-related subjects [4-18].

0022-2836/© 2015 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). *Matrix Biol* (2015) xx, xxx–xxx

Revie

Please cite this article as: Iozzo Renato V., Schaefer Liliana, Proteoglycan form and function: A comprehensive nomenclature of proteoglycans, Matrix Biol (2015), http://dx.doi.org/10.1016/j.matbio.2015.02.003

# **ARTICLE IN PRESS**

Proteoglycan nomenclature

LOCATION		CLASSIFICATION			GENE SYMBOL	UniProt ENTRY NAME	EPONYM	PREDOMINANT GAG	PROTEIN MODULE
INTRACELLULAR -	_	SECRE	TORY GRANULES		SRGN	SRGN	- Serglycin	— Нер —	
			1		sdc -	SDC	Syndecan, 1-4	— нѕ	
CELL SURFACE		TRA	ANSMEMBRANE 7 PI-ANCHORED		- CSPG4 -	CSPG4	- NG2	cs	
					- TGFBR3 -	TGFBR3	Betaglycan	CS/HS	
					PTPRZ1 -	PTPRZ	- Phosphacan	cs	
		01			GPC	GPC	- Glypican, 1-6	— нз	
		Gr	6		GFC	Gro	Chypican, 1-0	110	
PERICELLULAR -		E	BASEMENT MBRANE ZONE 4		F HSPG2 -	PGBM	Perlecan	— нs —	$(\bigcirc \Diamond \Diamond \Omega \bigcirc \bigcirc)$
					AGRN -	AGRIN -	- Agrin	— нз —	
		MEN			- COL18A1 -	COL18A1	Collagen XVIII	<u>    нs    </u>	
					COL15A1	COFA1	Collagen XV	CS/HS	$\bigcirc$
					ACAN -	PGCA	Aggrecan	— ся/кя —	
EXTRACELLULAR -		F	HYALECTAN LECTICAN 4		VCAN _	CSPG2	- Versican	CS	
					- NCAN -	NCAN	- Neurocan	CS	
			CANONICAL 13		L BCAN -	PGCB	Brevican		Ω <b>●]</b>
						PGS1	- Biglycan	cs	
		SMALL LEUCINE- RICH PROTEOGLYCANS (SLRPS)		CLASS I	DCN -	PGS2	- Decorin	DS	
					ASPN -	ASPN	- Asporin		
					ECM2	ECM2	ECM2		$\bigcirc$
				-	ECMX —	ECMX	- ECMX		$\bigcirc$
					FMOD -	FMOD	- Fibromodulin	— кs —	
					- LUM -	LUM	- Lumican	— кs —	
				CLASS II	PRELP -	PRELP	- PRELP		
					KERA -	KERA	- Keratocan	— кs —	
					L OMD _	OMD	Osteoadherin	<u> </u>	
					EPYC -	EPYC	- Epiphycan	— DS/CS —	
				CLASS III	ОРТС —	ОРТ	- Opticin		$\overline{\bigtriangleup}$
					L OGN -	MIME	- Osteoglycin		$\bigcirc$
			-NON-CANONICAL 5		CHAD -	CHAD	- Chondroadherin		
				CLASS IV	NYX -	NYX	- Nyctalopin		
							- Tsukushi		
		L			F PODN -	PODN	- Podocan		
				CLASS V	PODNL1 -	B7Z3M0	- Podocan-Like 1		
			SPOCK		- SPOCK -	TICN	- Testican, 1-3	— нѕ —	
			3						
KEY			Ser-Gly rep	eat 🚺	PDZ Bind	ling Domain	- CSPG Repe	eats 🦳 Lam	inin Like Domain
GAG - Glycosam	inog	lycan						)	
			OO Zona Pellu	cida 🥑	Carbonic An	hydrase 🔵 🔴 Pr	otein Tyrosine Pho	sphatase D1/2	Fibronectin TIII
<b>Hep</b> – Heparin				Sperm Pr	otein. Enter	okinase and Agrin		Receptor Class A	Repeat
HS - Heparan S	Sulfa	ite			_		V		
CS – Chondroiti	n Sı	ulfate	<mark>Ι Ω</mark> IgG []E	GF-Like	Follis	statin S/T	Rich 🛧 Frizz	led <b>O</b> Thron	nbospondin-Like 1
<b>KS</b> – Keratan S			Proteoglyc	an Tander	n Repeat	C-type Lee	ctin 🚫 Co	mplement Binding	Protein
NO - Neratali S	unal		Leucine-Ric	h Denset	Ear re	peat Acidic	Thyroglobu		Calcium-Binding
DS - Dermatan									

**Fig. 1.** A comprehensive classification of proteoglycans. The four families are based on their cellular and subcellular location, homology at the protein and genomic levels and the presence of unique protein modules which are often shared by members of a given class. The key for the various modules is provided in the bottom panel. For additional details about structure and function, please consult the text.

#### **General features**

Four major proteoglycan classes encompass nearly all the known proteoglycans of the mammalian genome (Fig. 1). Observing the types of proteoglycans based on cellular and subcellular localization, we can see that there is only one intracellular proteoglycan, serglycin. This unique proteoglycan forms a class on its own as it is the only proteoglycan that carries heparin side chains.

Please cite this article as: lozzo Renato V., Schaefer Liliana, Proteoglycan form and function: A comprehensive nomenclature of proteoglycans, Matrix Biol (2015), http://dx.doi.org/10.1016/j.matbio.2015.02.003

2

Download English Version:

## https://daneshyari.com/en/article/8455258

Download Persian Version:

https://daneshyari.com/article/8455258

Daneshyari.com