ChemBioChem

Review
doi.org/10.1002/cbic.202500095

Chemistry Europe
European Chemical Societies Publishing

www.chembiochem.org

Recent Advances in Enzymes and Chemoenzymatic Synthesis of Tetrasaccharide Linkage Region of Proteoglycans

Po-han Lin and Xuefei Huang*

Proteoglycans (PGs) play roles in many important biological events, including growth factor signaling, wound repair, blood coagulation, brain development, and neural stem cell migration. As the glycosaminoglycan chain of PGs is attached to the core protein through a tetrasaccharide linkage region, it is important to decipher the intricate control of the multitude of enzymes engaged in biosynthesis of the linkage region. During the past two decades, significant advances have been achieved in our understanding of the identity and interactions of these enzymes as well as the control of their activities. Furthermore, the knowledge gained on biosynthesis of the linkage region has enabled the synthesis of structurally well-defined proteoglycan-associated glycopeptides. The expression, activity, and substrate profile of enzymes and the latest development of chemoenzymatic synthesis of linkage region-bearing glycopeptides are summarized in this review.

1. Introduction

Proteoglycans (PGs), a class of highly glycosylated proteins, are ubiquitous in vertebrates. Located mostly in the extracellular matrix and on the cell surface,[1–3] PGs have diverse biological activities, including growth factor signaling, wound repair, blood coagulation, brain development, and neural stem cell migration.[3–7] The structure of PGs commonly consists of a core protein, which is covalently linked to a glycosaminoglycan (GAG) chain through a tetrasaccharide linkage region composed of glucuronic acid (GlcA)β1-3galactose (Gal)β1-3Galβ1-4xylose (Xyl)-β1-O-. The glycan chain is attached to a serine residue of a serine-glycine (SG) or serine-alanine (SA, much more rare than SG) dipeptide motif within the core protein backbone.[8,9] A PG can have one or more GAG chains, which can be heparan sulfate (HS), chondroitin sulfate (CS), or dermatan sulfate (DS). The GAG chains bear multiple sulfate moieties at various hydroxyl and amino groups, rendering them highly negatively charged in nature.

P. Lin, X. Huang
Department of Chemistry
Michigan State University
578 S. Shaw Lane, East Lansing, MI 48824, USA
E-mail: huangxu2@msu.edu

P. Lin, X. Huang
Institute for Quantitative Health Science and Engineering
Michigan State University
East Lansing, MI 48824, USA

X. Huang
Department of Biomedical Engineering
Michigan State University
East Lansing, MI 48824, USA

The attachment of glycan chains to form PGs mainly occurs through post-synthetic modifications of the nascent protein backbone in the endoplasmic reticulum and Golgi apparatus.[10] The biosynthesis is accomplished through a series of well-orchestrated enzymatic reactions (**Figure 1**). The synthesis is initiated with an xylosyltransferase-1 or −2 (XT-1 or XT-2) recognizing a peptide or protein substrate with an SG or SA dipeptide sequence and transferring a xylose in β configuration to the serine residue.[9] Subsequently, β-1,4-galactosyltransferase 7 (B4GALT7) installs a galactose residue to the 4-O position of xylose[11,12] followed by the phosphorylation of the 2-O of xylose by kinase family with sequence similarity 20 member B (FAM20B)[13] and the addition of another galactose unit by β-1,3-galactosyltransferase 6 (B3GALT6).[14] Following the trisaccharide assembly, 2-phosphoxylose phosphatase 1 (XYLP) cleaves the phosphate from xylose,[15] and β-1,3-glucuronyltransferase 3 (B3GAT3)[16] adds a GlcA residue to the trisaccharide glycopeptide, completing the tetrasaccharide linkage region.
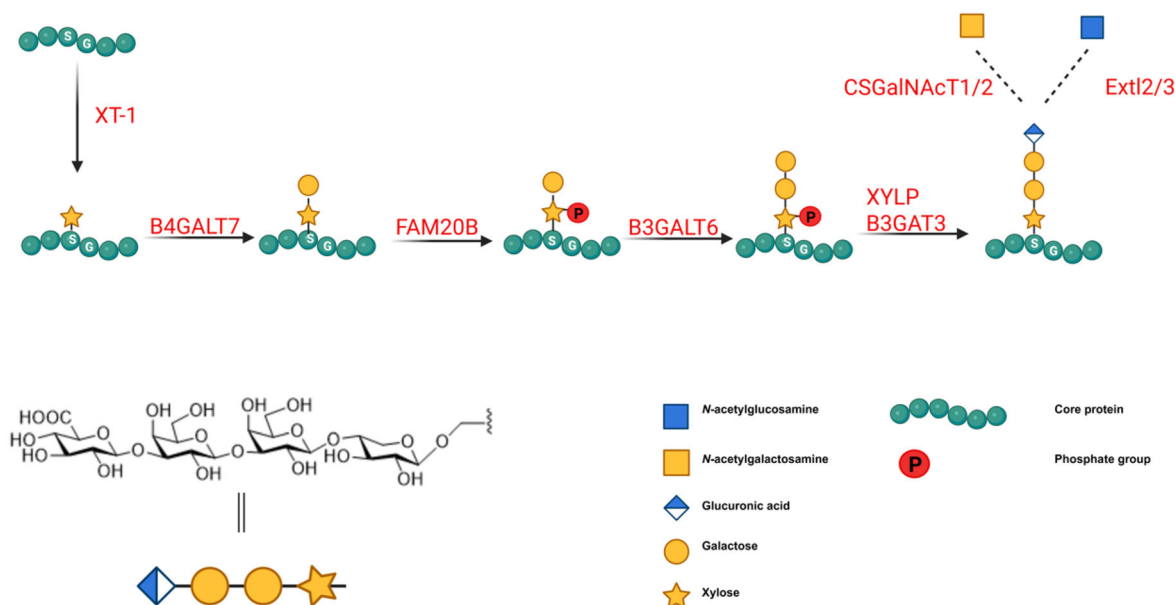
Upon synthesis of the complete tetrasaccharide linkage region, the glycan chain is extended by chondroitin sulfate N-acetylgalactosaminyltransferase 1 (CSGalNAcT1) or chondroitin sulfate N-acetylgalactosaminyltransferase 2 (CSGalNAcT2) to add a β-1,4-N-acetylgalactosamine (GalNAc) (Figure 1). The GalNAc-containing pentasaccharides are further elongated by bifunctional enzymes like chondroitin sulfate synthase 1 (Chsy1), chondroitin polymerizing factor (Chpf), and chondroitin polymerizing factor 2 (Chpf2), resulting in extended CS chains as part of a chondroitin sulfate proteoglycan (CSPG).[17] Alternatively, the tetrasaccharide linkage region can be modified by exostosin-like glycosyltransferase 2 (Extl2) or exostosin-like glycosyltransferase 3 (Extl3) to install an α-1,4-N-acetylglucosamine (GlcNAc) residue.[17–20] The GlcNAc-bearing pentasaccharides are elongated by exostosin glycosyltransferase 1 (Ext1) or exostosin glycosyltransferase 2 (Ext2), leading to HS chains for HSPG.[20] The length, sulfation, and epimerization pattern of HS and CS chains

ChemBioChem

Review
doi.org/10.1002/cbic.202500095

Chemistry Europe
European Chemical
Societies Publishing

Figure 1. Schematic demonstration of the biosynthesis of the tetrasaccharide linkage region of PGs.

can vary significantly depending on the extent of enzymatic modifications, resulting in a highly heterogeneous glycan pattern.[21]

While the biological functions of PGs are traditionally viewed to be dominated by the glycan chains, it is increasingly appreciated that both the core protein and GAGs can significantly influence the activities of PGs.[22–26] Hence, it is imperative to consider both the protein and the glycan in biological studies, necessitating the use of structurally well-defined PGs or glycopeptides to enhance the understanding of their structure-activity relationships. Due to the inherent heterogeneity of naturally existing PGs, the synthesis of PGs is needed to produce homogeneous PGs or glycopeptides.

As the tetrasaccharide linkage region is the bridge connecting the core protein with GAG chains, it is important to understand the enzymes involved in its biosynthesis. Recently, a review focused on the efforts made on XT and B4GALT7, the first two enzymes engaged in the biosynthesis pathway.[27] In this review, we will summarize the understanding and the progress made in the expression and substrate specificity of other major enzymes responsible for tetrasaccharide linkage region synthesis. In addition, efforts in applying the enzymes to successfully synthesize the full linkage region are presented.

2. Family with Sequence Similarity 20 Member B

FAM20B is a protein belonging to the FAM20 family, which plays a significant role in various biological processes.[28–30] The FAM20 family consists of three members, namely FAM20A, FAM20B, and FAM2C. FAM20A acts as a pseudo kinase that interacts with FAM2C, enhancing its activity.[31] FAM2C functions as a Golgi casein kinase, preferentially phosphorylating serine in the serine-x-aspartic acid tripeptide motifs (x can be any amino acid residue).[32] In comparison, FAM20B functions as an xylosylkinase, phosphorylating the 2-O position of xylose of the Galβ1-4Xyl- disaccharide.[13]

Po-han Lin earned his Ph.D. in organic chemistry in Dr. Xuefei Huang's laboratory at Michigan State University, where he focused on chemical and enzymatic syntheses of glycopeptides and complex oligosaccharides. In addition, he developed new methods for conformational analysis and profiling of chondroitin sulfate-bearing glycopeptides and proteoglycans. He is currently a postdoctoral researcher at Scripps Research.

Xuefei Huang is a MSU Foundation Professor at Michigan State University. He is a full professor jointly appointed in the Departments of Chemistry and Biomedical Engineering, as well as the Chief of the Chemical Biology Division of the Institute for Quantitative Health Science and Engineering. His main research interests are aimed at studying the chemistry and biology of carbohydrates. Major focuses of his research include the synthesis of complex carbohydrates and glycoconjugates, the development of novel approaches to boost immune responses against carbohydrate antigens as potential anticancer and anti-microbial vaccines, as well as the integration of glycoscience with nanotechnology.

The function of FAM20B has been shown to be essential in the body.[13] When FAM20B is depleted or dysfunctional, it can lead to abnormalities in cartilage matrix organization, early-stage chondrocyte development, development of supernumerary teeth, chondrosarcoma with major postnatal ossification defects, and severe craniofacial defects.[28,30,33–36] These observations have provided valuable insights into the mechanisms underlying these developmental disorders and underscore the significance of FAM20B and its relationship with these disorders.

2.1. Expression of FAM20B

The discovery of the FAM20 series dates back to 2005[37] with the expression of human FAM20B achieved later. In 2009, the Kitagawa group expressed FAM20B in HeLa cells using a stable transfection system.[13] At that time, FAM20B was only known as an xylosylkinase and its relationship with other linkage region biosynthesis enzymes such as B3GALT6 was not clear.

In 2011, the Kimmel group reported that FAM20B could be expressed in vivo when Tol 2 expression plasmids were injected into the embryo of zebrafish.[28] In 2013, Kitagawa group again expressed FAM20B but using the African green monkey kidney fibroblast-like cell line (COS-1 cells) with the plasmid inserted with a cleavable immunoglobulin G (IgG) domain. Upon harvest, the enzyme was purified from the culture medium with IgG-Sepharose.[38]

In 2014, the Dixon group expressed FAM20B using insect cells Hi5.[39] The desired gene containing truncated FAM20B(aa 42-409) was fused to a maltose binding protein (MBP)-6X His tag and generated as a baculovirus plasmid. Hi5 cells were infected and then incubated for 2 days. The medium was collected and purified with Ni-NTA resin. When necessary, the MBP fusion protein can be cleaved by Tobacco Etch Virus protease to obtain pure FAM20B.[39] In 2018, Dixon and Xiao's group together reported another expression of FAM20B.[40] The expression system was very similar to their published study in 2014, with the exception that the protein sequence was changed to 55–402. This change in the sequence did not alter the protein function.

2.2. Substrate Selectivity of FAM20B

The substrate specificity of FAM20B was initially documented by the Kitagawa group.[13] Their study employed α-thrombomodulin (α-TM) as the substrate for FAM20B, which is a glycoprotein found on the surface of endothelial cells.[41] Unlike β-TM, the CS variant of thrombomodulin, α-TM only possesses the tetrasaccharide linkage region, which can potentially serve as a substrate for FAM20B. FAM20B recognizes and accepts the tetrasaccharide-bearing PG α-TM as a substrate with a k_{cat} value of 102 pmol h mL^{-1}. When the acceptor was switched to a trisaccharide with serine as the aglycon (Galβ1-3 Galβ1-4Xyl-Ser), similar activity of 128 pmol h mL^{-1} was observed. This indicates that FAM20B does not require a protein/peptide aglycon for acceptor binding and can readily phosphorylate tri- or tetra-saccharides.

In 2014, the Dixon group conducted a kinetic assay of FAM20B using [γ-^{32}P]ATP to measure its rates of reaction.[39] Three substrates were employed: tetra-Bn (GlcAβ1-3 Galβ1-3Galβ1-4Xylβ1-Bn), Gal-Xyl-Bn (Galβ1-4Xylβ1-Bn), and Xyl-Bn (Xylβ1-Bn), all containing a β-linked benzyl (Bn) group as the aglycon of the substrate. Tetra-Bn and disaccharide Gal-Xyl-Bn served as the substrate with K_m values of 40 and 42 μM, respectively. In contrast, the enzyme did not have much activity toward the monosaccharide Xyl-Bn. By combining the results from these two studies, it is clear that FAM20B can carry out phosphorylation without a stringent requirement on the aglycon. Instead, the number of saccharides primarily influences its activity. The substrate must possess a minimum disaccharide motif (Gal-Xyl), implying that galactose potentially plays a crucial role in acceptor binding.

2.3. Structural Analysis of FAM20B

To better understand the activity of Fam20B, Xiao and Dixon groups performed structural studies of this enzyme.[40] Although human Fam20B could be purified to homogeneity, it failed to crystallize. Rather, they focused on Hydra magnipapillata FAM20B (hmFam20), a homolog of human Fam20B, which readily crystallized. Treatment of hmFam20 with endoglycosidase F3 removed N-linked glycans and significantly improved crystal quality, enabling structure determination at 2.2 Å resolution. The overall fold of hmFam20 closely resembled Fam2C, indicating strong structural conservation within this enzyme family.

To better understand substrate recognition, the structure of hmFam20 complexed with the Gal-Xyl disaccharide was obtained in the presence of adenylyl imidodiphosphate (AMP-PNP) as a nonhydrolyzable ATP analog (**Figure 2**a). Only the adenosine (ADN) moiety of AMP-PNP showed clear electron density, with its phosphate groups poorly resolved. Structural comparison with ADP-bound caenorhabditis elegans FAM2C (ceFam2C) revealed that the ATP-binding site is highly conserved. The Asp152 residue in hmFam20 aligned with Glu213 in ceFam2C and formed a salt bridge with Lys133, a key feature for catalytic activity in canonical kinases.

For the Gal-Xyl substrate, it was bound in a defined pocket adjacent to the ATP-binding site, formed by residues including Thr114, Gln115, Tyr148, Gly150, Tyr253, Asp299, His301, and Lys321 (Figure 2b). Multiple hydrogen bonds stabilized the disaccharide: Thr114 and Gln115 interacted with xylose, Gly150 with galactose, and Asp299—analogous to the conserved catalytic Asp in protein kinases, coordinated with the substrate's hydroxyl group, acting as a catalytic base. These substrate-interacting residues were evolutionarily conserved, and mutations at the corresponding sites in human Fam20B significantly impaired kinase activities. Collectively, these results indicated that Fam20B employed a conserved structural and mechanistic framework for specific recognition of Gal-Xyl, aligning with its functional role in glycan phosphorylation.
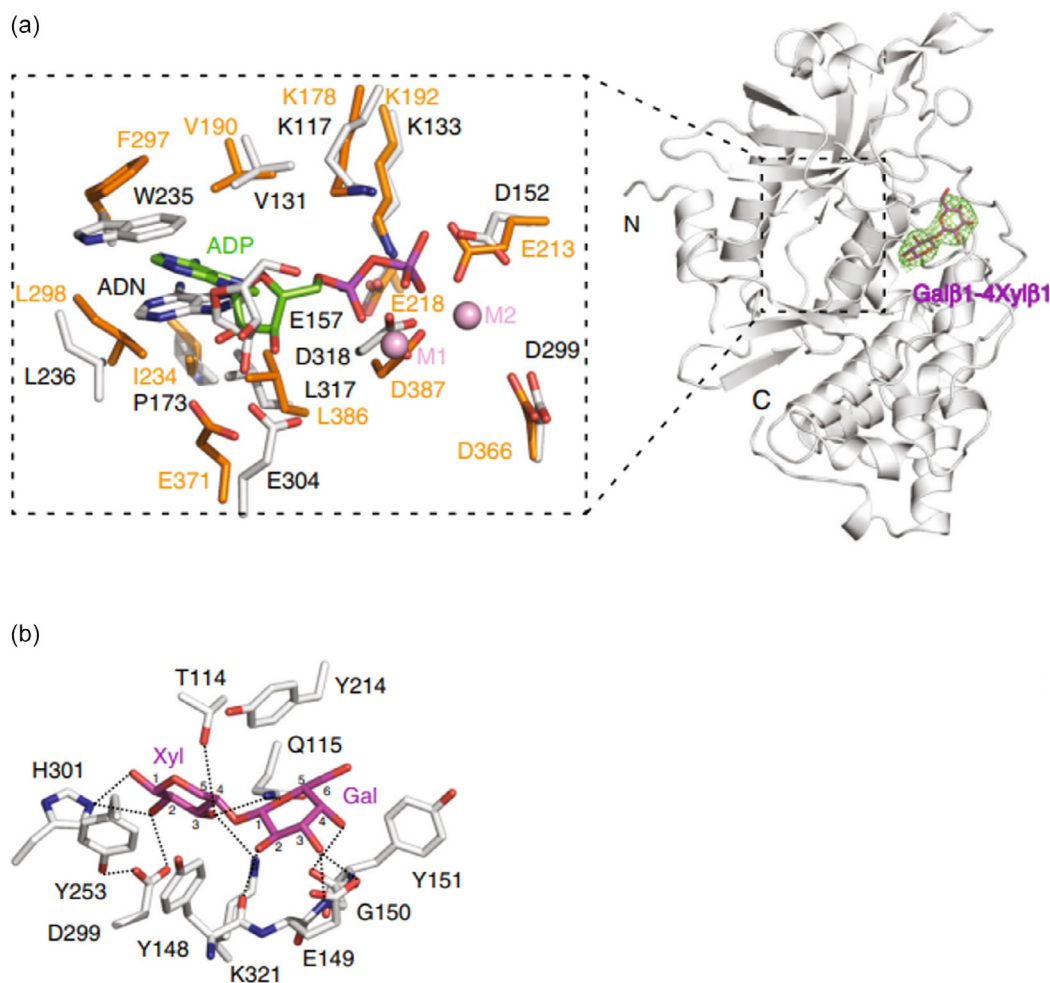
Figure 2. a) The structure of hmFam20 in complex with Gal-Xyl. Gal-Xyl was shown as a green mesh. A close view of the ATP-binding site is displayed on the left, with the Mn^{2+}/ADP-bound ceFam2C structure (shown in orange) and hmFam20B (shown in white) superimposed for comparison. ADN, adenosine. The two Mn^{2+} ions in ceFam20 were labeled as M1 and M2. b) Disaccharide-binding pocket showing the detailed molecular interactions between hmFam20 and Gal-Xyl disaccharide substrate. Adapted with permission.[40] Copyright 2018, Nature Publishing Group.

3. β-1,3-Galactosyltransferase 6

β-1,3-Galactosyltransferase 6, also referred to as B3GALT6, is an enzyme of significant importance in the biosynthesis of the tetrasaccharide linkage region. The discovery and comprehension of the activities of B3GALT6 have significantly advanced our understanding of PG biosynthesis and its profound impact on various biological processes.

In 2013, the Ikegawa group reported the connection between B3GALT6 mutations and several types of connective tissue disorders, such as lax skin, muscle hypotonia, joint dislocation, skeletal dysplasia, and deformities.[42–44] This investigation highlighted the crucial role of B3GALT6 in the development and maintenance of various tissues, including the skin, bones, cartilage, tendons, and ligaments. These findings shed light on the vital contributions of B3GALT6 to the intricate processes involved in tissue growth, organization, and overall physiological homeostasis. Further research continues to explore the functions and precise regulatory mechanisms by which B3GALT6 operates and its implications for understanding and potentially treating connective tissue disorders.[44]

3.1. Expression of B3GALT6

Although the initial discovery of B3GALT6 was reported in 2001,[14] the expression of B3GALT6 was achieved much later. In 2014, the Dixon group published the expression of B3GALT6 utilizing a baculovirus expression system.[39] A truncated version of B3GALT6, consisting of amino acids 31 to 329, was inserted into a bacmid along with a His-MBP affinity tag. The constructed bacmid was then transfected into Hi5 cells, and the resulting medium was collected after a 2-day period. The fusion protein was subsequently purified with a Ni-NTA column.

In 2018, the Moremen lab outlined their approach to construct a library of glycosyltransferases and glycoside hydrolases.[45] In their study, a truncated form of B3GALT6 comprising amino acids 35 to 329 was cloned into the mammalian expression vector pGen2-DEST. To enhance solubility, a His tag and an avidin tag were introduced to the N-terminus of B3GALT6 with a superfolded green fluorescent protein inserted between the tags and the B3GALT6 protein sequence. This plasmid was successfully utilized to express B3GALT6, which was found to be enzymatically active.[46]

ChemBioChem

Review
doi.org/10.1002/cbic.202500095

Chemistry
Europe
European Chemical
Societies Publishing

3.2. Substrate Selectivity of B3GALT6

The acceptor profile of B3GALT6 exhibited a strong dependence on the phosphorylation state of the acceptor. Dixon and coworkers reported that B3GALT6 displayed minimal activities when tested with Gal-Xyl-Bn as the acceptor.[13,39,47] Interestingly, when the disaccharide was phosphorylated by FAM20B prior to B3GALT6-mediated reaction, the K_m value for this reaction decreased by ≈230-fold, suggesting the phosphorylated xylose is critical for substrate binding by B3GALT6.

The role of FAM20B in modulating PG synthesis was underappreciated, with the original recognition limited to its kinase activity in phosphorylating xylose. In 2014, the Dixon group provided insights into the relationship between FAM20B and B3GALT6.[39] They generated FAM20B knockout (KO) human bone osteosarcoma epithelial cells (U2OS). Two distinct methods were employed to assess the overall GAG content in the cell lysate. The first involved the widely used 3G10 antibody, which specifically detects HS. The second method utilized the isotope [^{35}S] to quantify the sulfate content within GAGs. Comparative analysis with wild-type (WT) U2OS cells against the FAM20B KO U2OS cells revealed a 95% reduction in GAG levels as detected by the 3G10 antibody, and a similar outcome was observed using the ^{35}S isotope measurement approach. Furthermore, the researchers determined the linkage region structure of glypican 1 (GPC1) in the FAM20B KO cells.[39] Through β-elimination and mass spectrometry analysis, they discovered that instead of the expected tetrasaccharide linkage region, a trisaccharide linkage region comprising N-acetyl neuraminic acid (Neu5Ac)α2–3Galβ1–4Xylβ1 was observed. In 2018, the Yang group conducted a similar experiment by knocking out FAM20B in Chinese hamster ovary cells (CHO).[47] High-performance liquid chromatography analysis demonstrated a roughly threefold decrease in CS and a sixfold decrease in HS levels compared to those from WT CHO cells. This finding suggests that the impact on GAG levels may vary in different cell lines when specific GAG-associated enzymes are altered. Subsequent experiments involving the KO of B3GALT6 in CHO cells revealed that CS and HS were still detectable, although the overall content of GAG was significantly reduced compared to WT CHO cells. Similarly, in 2018, the Malfait group reported decreases in GAG levels in patients with mutated B3GALT6 genes.[48] Additionally, in 2019, the Larson group reported the identification of a trisaccharide linkage region (GlcA-Gal-Xyl) in GAGs, indicating that cells are capable of elongating GAGs with an incomplete linkage region.[49] These findings suggest that with its strong preference for phosphorylated disaccharide (Gal-Xyl) acceptor generated by FAM20B, B3GALT6 plays important roles in PG biosynthesis. In the absence of B3GALT6, GAG synthesis is not completely halted, but instead occurs in a lower yield, leading to the formation of a shortened trisaccharide linkage region GlcA-Gal-Xyl.

4. Phosphoxylose Phosphatase

2-Phosphoxylose phosphatase, also known as XYLP, is a phosphatase responsible for the dephosphorylation of xylose. It is one of the least studied enzymes among those involved in the linkage region synthesis, with the most systematic study of this enzyme reported in 2014 by the Kitagawa group.[15]

4.1. Expression of XYLP

In the Kitagawa study,[15] a truncated version of XYLP(aa 38-480) was cloned into the pEF-BOS vector. Transfection reagent FuGENE six was used to transfect XYLP into COS-1 cells. Two days after the transfection, the medium was collected, and the enzyme was purified with IgG-Sepharose beads.

4.2. Substrate Selectivity of XYLP

The substrate profile of XYLP was reported by the Kitagawa group through comparisons using different substrates: a phosphorylated trisaccharide (Gal-Gal-Xyl(2P)-TM) and a tetrasaccharide (GlcA-Gal-Gal-Xyl(2P)-TM), both linked to α-TM.[15] The results showed that only the phosphorylated trisaccharide with α-TM was accepted as a substrate by XYLP, with little activity observed for the tetrasaccharide. Furthermore, when comparing GalNAc-type (GalNAc-GlcA-Gal-Gal-Xyl(2P)-TM, a precursor to CSPG) and GlcNAc-type (GlcNAc-GlcA-Gal-Gal-Xyl(2P)-TM, a precursor to HSPG) compounds, both bearing a phosphorylated pentasaccharide with α-TM, XYLP only dephosphorylated the GalNAc-type compound, while no reaction was observed when GlcNAc-type was used as a substrate.

It is important to note that XYLP is highly selective toward the linkage region and does not accept other phosphorylated glycans on glycoproteins such as osteopontin or matrix extracellular phosphoglycoprotein.[15] This observation suggests that XYLP's substrate preference is specific to 2-phosphoxylose. In contrast, when alkaline phosphatase, a commercially available phosphatase, was used, all the substrates mentioned above could be dephosphorylated.

5. β-1,3-Glucuronyltransferase 3

Human B3GAT3, also known as glucuronyltransferase I (GlcAT-I), is an enzyme that plays a crucial role in assembling the tetrasaccharide linkage region.[16] B3GAT3 specifically catalyzes the transfer of a GlcA to the trisaccharide Gal–Gal–Xyl linkage region. GlcA is an important component of GAGs such as HS and CS. The addition of GlcA to the PG chain by B3GAT3 is a key step in the modification and maturation of GAG molecules. The synthesis of HS and CS is completely abolished when B3GAT3 is knocked out in cells,[47] which subsequently affects the integrity and function of tissues and contributes to the development of certain diseases such as recessive joint dislocations and congenital heart defects, including bicuspid aortic valve and aortic root dilatation.[50]

Overall, B3GAT3 is an important enzyme for PG biosynthesis, contributing to the structural diversity and functional versatility of these complex carbohydrates in the body. Its study has implications in understanding developmental processes, disease

mechanisms, and potentially identifying therapeutic targets for GAG-related disorders.

5.1. Expression of B3GAT3

In 1998, the Sugahara group reported the first expression of B3GAT3, a truncated version lacking the 43 amino acids from the N-terminus of the transferase.[51] This enzyme was cloned into vector pSVL containing an insulin signal sequence and a protein A sequence, and then transfected into COS-1 cells using Lipofectamine. Two days after the transfection, the cell culture medium was separated, and the desired enzyme was purified with IgG-bearing Sepharose.

In 1999, the Esko group reported a similar method using stable transfection.[52] In their study, cDNA encoding amino acids 30–335 was cloned into pcDNA1, then further fused into vector pRK5-F10-PROTA with the C-terminal domain of protein A. This plasmid was then transfected into COS-7 cells using Lipofectamine, and stable transfectants were selected using 0.2 mg mL^{-1} of antibiotic G418. Supernatants after transfection were subjected to rabbit IgG beads to yield the purified enzyme.

During the same year, the Esko group reported another expression of B3GAT3 using a stable transfection method.[53] Gene containing the sequence of B3GAT3 was cloned from CHO cells and inserted into the vector pCDNA3. Plasmid was then transfected into CHO cells with Lipofectin with appropriate colony was selected.

In 2000, to study the role of cysteine in B3GAT3 function, the Fournel–Gigleux group reported the expression of multiple truncated or mutated human B3GAT3 using yeast as the expression host.[54] Constructs included one lacking the predicted N-terminal cytoplasmic tail (B3GAT3ΔNT) or one further fused with the yeast prepro-α-factor secretion leader peptide (B3GAT3ΔNT/TMD). Similar sequences but lacking the first 25 N-terminal amino acids were fused with yeast prepro-α-factor secretion leader peptide and an antisense primer corresponding to the coding sequence for the last six amino acids (αF-B3GAT3ΔNT/TMD). Another two mutants were constructed with the full sequence of B3GAT3 and C33 and C301 mutated to alanine, respectively. These enzymes were cloned into yeast vector pPICZB plasmids and transformed into Pichia pastoris SMD 1168 by lithium chloride. Downstream analysis of the expressed B3GAT3 showed that this enzyme, fused with the yeast prepro-α-factor secretion leader peptide, was successfully cleaved during expression, and the glycosylation state of B3GAT3ΔNT/TMD was similar to the WT membrane-bounded B3GAT3. The result also showed that without the N-terminal cytoplasmic tail, B3GAT3ΔNT did not alter the enzyme's ability to target and associate with the yeast membrane or its enzymatic activity. In contrast, deletion of the cytoplasmic tail allowed B3GAT3 to enter the secretory pathway, suggesting the lack of a retention signal in the stem. Furthermore, the mutation of C33A ($K_m = 67.23$ μM to the Gal-Gal disaccharide) suppressed dimer formation through disulfide bond, eventually leading to impaired activity compared to WT B3GAT3 ($K_m = 37.03$ μM), indicating that dimerization was not absolutely required for enzymatic activity but was necessary for the optimal function of the enzyme. Mutation of the strictly conserved C301 to alanine completely abolished its activity, suggesting that Cys301 is an essential residue for catalytic activity.

In the same year, Negishi group reported an expression of B3GAT3 using Escherichia coli as the host.[16] Truncated human B3GAT3 with amino acids from 76 to 335 was cloned into pET-28a with an N-terminal 6X His tag. Plasmid was then transformed into BL21(DE3), and the transformed colony was selected on Luria–Bertani (LB) agar plates. Upon induction and incubation, cells were lysed, and the desired protein was purified by a Ni-NTA column. This B3GAT3 was further concentrated to 30.8 mg mL^{-1} and used for crystallization. It was found that the aforementioned C301 was not involved in binding with the acceptor and donor, and E281 acted as the base to deprotonate the Gal acceptor.

5.2. Substrate Selectivity of B3GAT3

The Sugahara group reported the acceptor specificity of B3GAT3 first.[51] They compared several acceptor molecules: Galβ1–3Galβ1–4Xylβ1-O-Ser, GalNAcβ1–4GlcAβ1–3Galβ1–3Galβ1–4Xylβ1-O-Ser, GalNAcβ1–4GlcAβ1–3GalNAcβ1–4GlcAβ1–3Galβ1–3Galβ1–4Xylβ1-O-Ser, chondroitin, and asialoorosomucoid (Galβ1–4GlcNAc-R). Reactions were only observed when the trisaccharide linkage region Galβ1–3Galβ1–4Xylβ1-O-Ser was used as an acceptor. Penta- and hepta-saccharides with GalNAc at the non-reducing end showed no reactions, demonstrating that B3GAT3 can only transfer GlcA to the linkage region, rather than elongating the chondroitin-bearing backbone. Similarly, chondroitin and asialoorosomucoid were not active as acceptors of B3GAT3.

In 2002, the Fournel–Gigleux group published the donor substrate specificity of B3GAT3.[55] They utilized UDP-GlcA, UDP-galacturonic acid (GalA), UDP-glucose (Glc), UDP-Gal, UDP-N-acetyl glucosamine (GlcNAc), UDP-mannose (Man), and GDP-Man as donors, with Gal-Gal as the acceptor. Only UDP-GlcA and UDP-GalA exhibited reactions, while other donors gave minimal to no reactivity. To further elucidate the reaction mechanism, two mutants, H308A and H308R, of B3GAT3 were generated based on the crystal structure of the protein. Mutation of histidine 308 to alanine completely abolished reactions for all donors, suggesting that this histidine residue is crucial for catalytic activity. Interestingly, when histidine 308 was mutated to arginine, all donors tested displayed activities except for UDP-Gal. When the native donor UDP-GlcA was used, the mutant exhibited a decreased V_{max} of 23.78 mol min^{-1} mgP^{-1}, significantly lower than that for the native B3GAT3 (68.03 mol min^{-1} mgP^{-1}).

5.3. Structural Analysis of B3GAT3

The crystal structure of B3GAT3 was reported by the Negishi group,[16] which has a seven-stranded mixed β-sheet structure divided into two subdomains. The active site was found to be located in a cleft spanning both subdomains. The N-terminal subdomain featured an α/β motif with alternating β-strands and α-helices, which comprised most of the residues for binding the donor substrate. The UDP molecule bound in a long cleft of B3GAT3, with the uridine and ribose rings positioned in the

ChemBioChem

Review

doi.org/10.1002/cbic.202500095

Chemistry Europe
European Chemical
Societies Publishing

N-terminal subdomain and forming multiple interactions, notably Asp113, Tyr84, and Asp195 (**Figure 3**). Two arginine residues (Arg156, Arg310) stabilized the phosphate groups, while a Mn^{2+} ion coordinated with UDP phosphates and conserved DXD motif residues (Asp194–Asp196), aided by water molecules. The C-terminal subdomain is responsible for binding the acceptor substrate, which continues the β-sheet from the N-terminal region. It also includes a mixed β-sheet and extends at the C-terminus along another molecule in the crystal. The C-terminal subdomain bound the acceptor substrate (Gal-Gal-Xyl), primarily through interactions with the non-reducing terminal Gal (Gal-2) via Glu227, Asp252, and Glu281—key residues conserved among glucuronyltransferases.

The proposed mechanism for B3GAT3-catalyzed glucuronidation involved Glu281 acting as a catalytic base to deprotonate the 3-OH of Gal-2, facilitating nucleophilic attack on the glucuronic acid of UDP-GlcA and forming the GlcAβ1–3Gal linkage. His308 could assist in stabilizing the transition state or orienting the donor. Structural comparison revealed that B3GAT3 shares a similar fold and catalytic features with other inverting glycosyltransferases despite little to no sequence identity, suggesting a conserved reaction mechanism across this enzyme family.

To determine the acceptor substrate preference of B3GAT3, Sugahara group co-crystalized the catalytic domain of B3GAT3 with UDP and Galβ1–3Gal(6S)β1–4Xyl(2P)β1-O-Ser or UDP and Galβ1–3Galβ1–4Xyl.[56] The substrate Galβ1–3Gal(6S)β1-O-Ser bound to B3GAT3 in the same orientation as the previously studied nonsulfated trisaccharide,[16] though only the Galβ1–3Gal(6S) portion showed clear electron density. The 2-O-phosphorylated xylose extended into the solvent and did not interact specifically with the enzyme. The non-reducing end Gal-2 binds deeply in the active site, maintaining conserved hydrogen bonds with Glu227, Arg247, Asp252, and Glu281 (the proposed catalytic base). Sulfation at the O-6 of the internal Gal

(Gal-1) induced a conformational shift in Gln318, allowing interaction with the sulfate group. However, mutagenesis studies suggested Lys317, not Gln318, may enhance binding, though in the crystal it appeared disordered. In contrast, the isomeric Gal(6S)β1–3Galβ1–4Xyl(2P)β1-O-Ser was not a substrate for B3GAT3. Structural analysis indicated that sulfation of the Gal-2 created steric clashes with Glu227 and Arg247, preventing proper binding and catalysis, thus explaining the lack of enzymatic activity with this trisaccharide.

5.4. Relationship Between FAM20B, B3GAT3 and XYLP

Following the synthesis of the phosphorylated linkage region trisaccharide, B3GAT3 transfers GlcA to the phosphorylated trisaccharide (Gal-Gal-Xyl(2P)-Ser) linked to a serine residue (**Figure 4**). At the same time, xylose dephosphorylation is initiated by XYLP. The co-expression of XYLP and B3GAT3 led to rapid dephosphorylation of the linkage region. This is because B3GAT3 can form oligomers with XYLP to enhance the dephosphorylation activity more than 16-fold.[15] Excessive acceleration of linkage region phosphorylation by FAM20B and/or attenuated Xyl dephosphorylation by XYLP may lead to the accumulation of biosynthetic intermediates, specifically phosphorylated linkage tetrasaccharides. Upon completion of the tetrasaccharide linkage region synthesis, CS or HS polymerases facilitate the polymerization of disaccharide chains onto the tetrasaccharide (Figure 4).
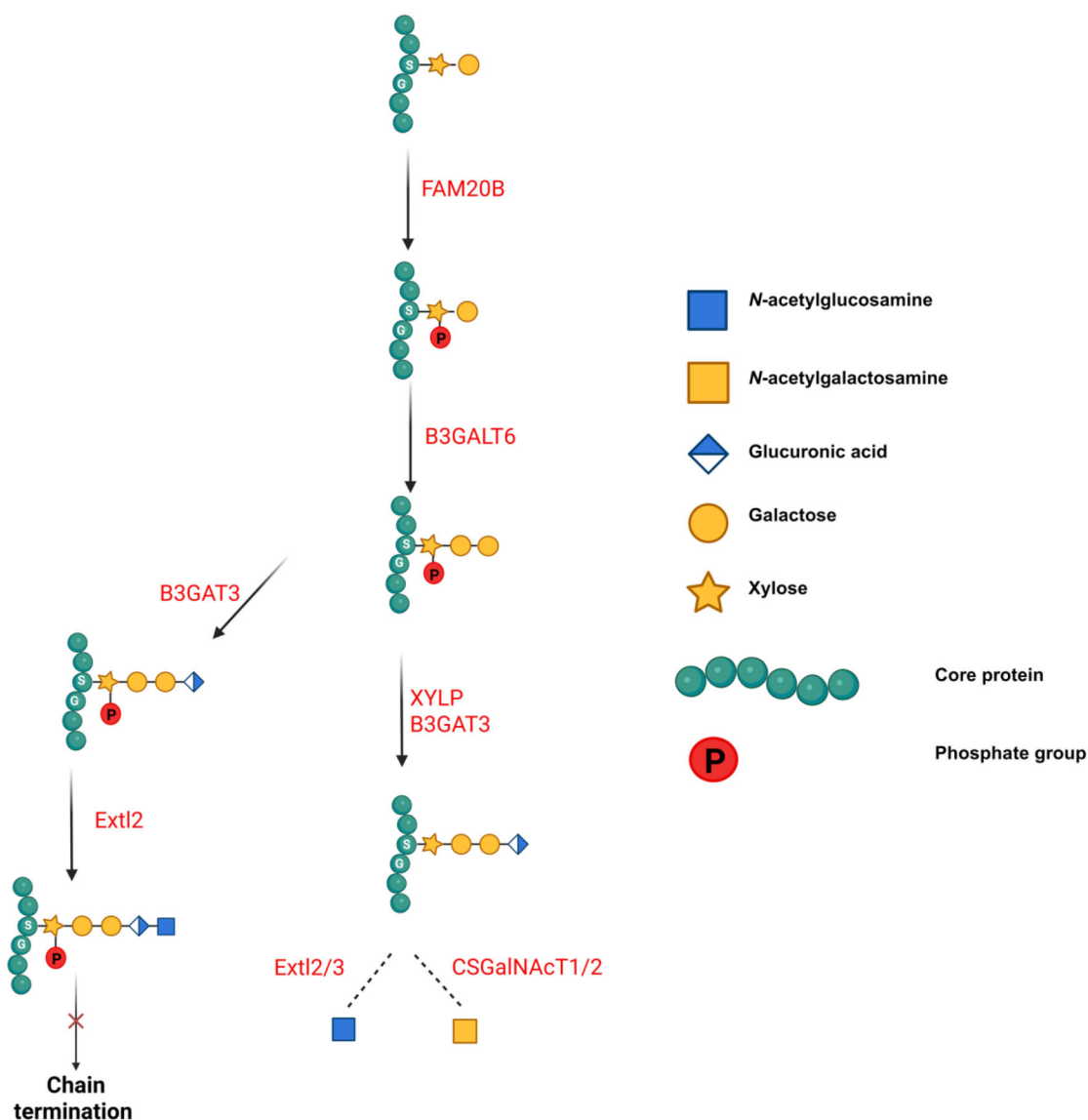
6. Enzymatic and Chemo-Enzymatic Synthesis of the Tetrasaccharide Linkage Region

With the appreciation of the importance of PGs, there has been significant interest in producing homogeneous PGs to facilitate the structure and activity relationship studies.[26,46] Due to the complexity of PG structures, chemical synthesis is highly tedious and time-consuming because of the extensive protective group manipulations needed and the challenging nature of chemical glycosylation reactions.[57–59] The knowledge gained on the enzymes involved in the tetrasaccharide linkage formation can significantly expedite the synthesis of homogeneous glycopeptides.

In order to synthesize the glycopeptides, the Huang group explored a chemoenzymatic strategy.[46] The requisite enzymes, including XT-1, B4GALT7, B3GALT6, and B3GAT3, were recombinantly prepared. They found that B4GALT7 and B3GAT3 could be expressed in *E. coli* with the yields of 16.7 and 5 mg L^{-1}, respectively, while XT-1, FAM20B, and B3GALT6 were expressed in HEK293F cells with the desired enzymes isolated from the supernatant in 10, 2.3 and 16.7 mg L^{-1}, respectively. The peptide acceptors bearing one or more SG glycosylation sites from PGs, including bikunin, syndecan-3, and syndecan-4 (peptides **1–5**) were produced by solid phase peptide synthesis (**Figure 5**). As the glycopeptide product can be difficult to purify from the aqueous reaction media, a new solid phase supported strategy was developed where the peptide substrates were anchored on Sepharose resin via a diethyl square linker that
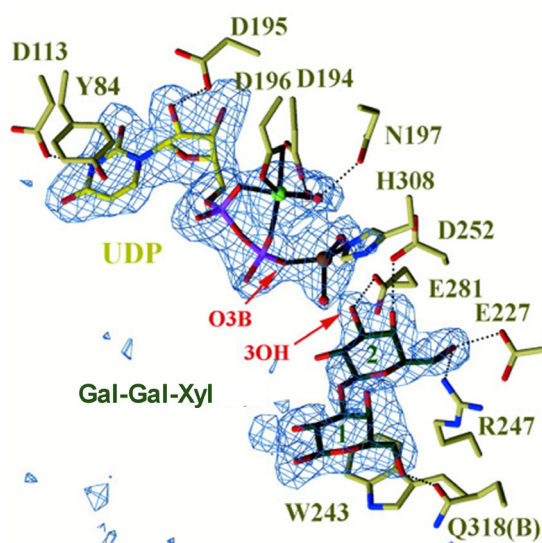
Figure 3. Active site of B3GAT3 with interactions between the enzyme and UDP and the Gal-Gal-Xyl trisaccharide acceptor highlighted. Reproduced with permission.[16] Copyright 2000, ASBMB Publishing Group.

ChemBioChem

Review

doi.org/10.1002/cbic.202500095

**Chemistry Europe**
European Chemical
Societies Publishing

Figure 4. Phosphorylation and dephosphorylation of Xyl residues regulate the formation of the linkage region and GAG biosynthesis.[15]

can be cleaved under mild conditions to release the glycopeptide product after enzymatic reactions (Figure 5).

The chemoenzymatic synthesis was first explored by incubating the peptide-immobilized resins sequentially with XT-1, B4GALT7, and B3GALT6.[46] However, only glycopeptide bearing Gal-Xyl disaccharide was isolated from the reaction mixture upon cleavage from the solid phase support, with little Gal-Gal-Xyl trisaccharide glycopeptide detected. To overcome this, upon completion of B4GALT7-promoted galactosylation reaction, the resin was incubated with FAM20B. Subsequent galactosylation catalyzed by B3GALT6 successfully led to the formation of Gal-Gal-Xyl(2-P) bearing glycopeptide. This observation is consistent with the report that 2-O phosphorylation of Gal-Xyl disaccharide significantly enhanced the rate of B3GALT6 reaction.[39] The Gal-Gal-Xyl(2-P) bearing glycopeptide was then dephosphorylated with XYLP or alkaline phosphatase, which was followed by B3GAT3 and hydrazine-promoted linker cleavage, leading to glycopeptide **6** bearing the full tetrasaccharide linkage region

on a mg scale (Figure 5). The synthesis protocol was robust, which was successfully applied to the syntheses of glycopeptides **6**–**10** in 6–20% overall yields from the peptide backbone, including glycopeptides **8** and **9** bearing multiple sites of glycosylation. Only one purification was needed in this multi-step synthesis, enhancing the overall efficiency. The glycopeptides were then further treated with CS biosynthetic enzymes, forming a series of CS glycopeptides.[46] This was the first time such homogeneous CSPG-based glycopeptides have been synthesized, highlighting the power of chemoenzymatic synthesis. The interactions between the synthetic CS glycopeptides and cathepsin G were analyzed. The sulfation of the glycan chain was found to be important for binding to cathepsin G.[46]

In another recent study, the Wild group explored a chemoenzymatic strategy to synthesize fluorescently labeled tetrasaccharide linkage regions bearing glycopeptides.[60] XT-1, B4GALT7, B3GALT6, and B3GAT3 were expressed recombinantly in *E. coli*. Peptide fragments from PGs such as syndecan-3, TGFβ
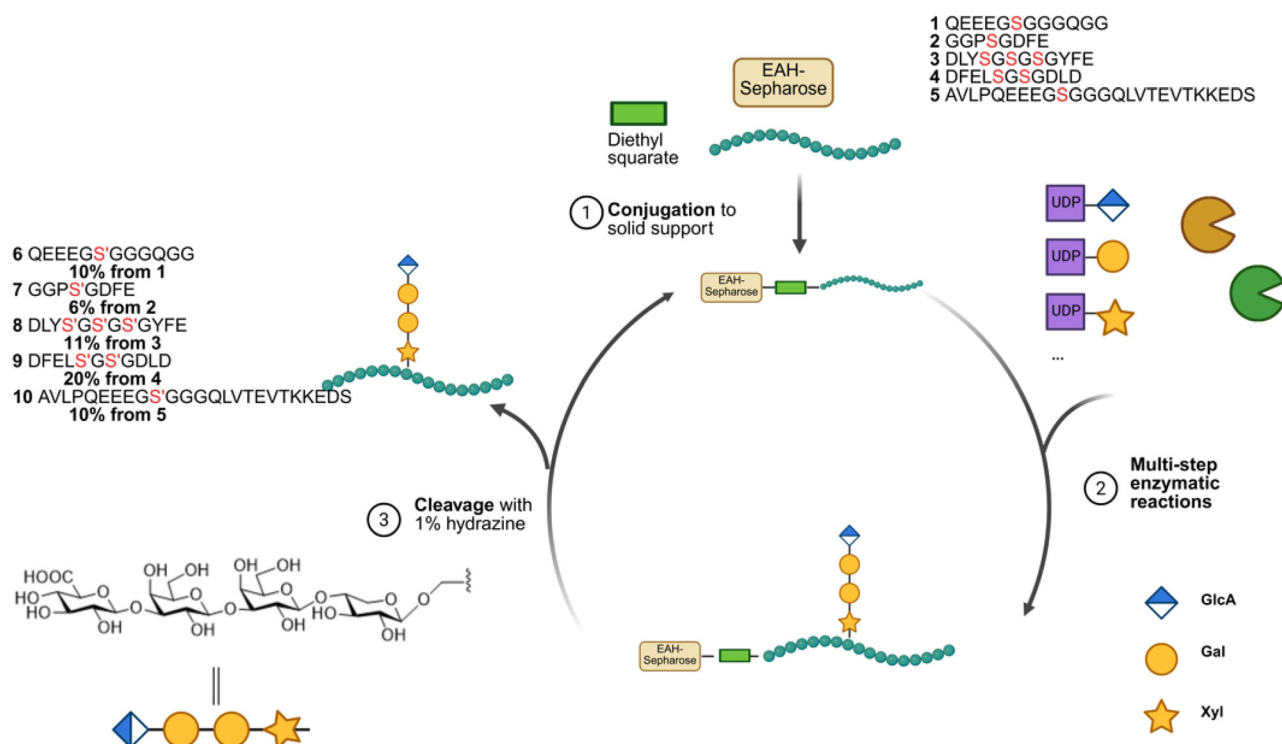
Figure 5. Schematic demonstration of solid phase (Sepharose) supported chemoenzymatic synthesis of glycopeptides **6–10**. bearing the full tetrasaccharide linkage region.[46]

type III receptor, and macrophage colony-stimulating factor 1 were treated with these enzymes in a one-pot two-step manner. First, XT-1 and B4GALT7, along with glycosyl donors UDP-Xyl and UDP-Gal, were incubated with the fluorophore-bearing peptides. Following a 16-hour incubation, B3GALT6 and B3GAT3 were added with UDP-GlcA and additional UDP-Gal. Interestingly, in this reaction, without the need for FAM20B and 2-*O* phosphorylation of xylose, the full tetrasaccharide GlcA-Gal-Gal-Xyl was successfully installed onto the peptide backbone in μg scale reactions. This could be due to the relatively high amounts of enzymes used and the long reaction time applied for their μg scale reactions.

Another interesting chemoenzymatic synthesis was reported by Hohenester and coworkers.[61] They expressed the requisite PG biosynthetic enzymes (XT-1, B4GALT7, B3GALT6, B3GAT3, FAM20B) as well as enzymes responsible for constructing CS and HS chains in Expi293F cells.[61] Remarkably, they discovered that the tetrasaccharide glycopeptide synthesis could be performed in a one-pot multiple enzyme (OPME) reaction. In their synthesis, a PG peptide acceptor was co-incubated with all the GAG linker enzymes (XT-1, B4GALT7, B3GALT6, B3GAT3, and FAM20B), their cognate UDP-sugars, and ATP, and near-complete (≈90%) conversion of the peptide to the phosphorylated tetrasaccharide peptide, GlcA-Gal-Gal-Xyl2P-peptide, was achieved. The 2-*O* phosphorylation of xylose was found to dramatically increase the catalytic activities of B3GALT6 (632-fold increase in k_{cat}/K_m) and modestly increase the rate of B3GAT3 reaction (6.4-fold increase in k_{cat}/K_m). This synthetic protocol was applied to the

syntheses of several PG glycopeptides, including those from bikunin, CS proteoglycan 4, TGFβ type III receptor, syndecan-2, syndecan-4, and glypican-1, demonstrating the broad applicability of the method. The availability of these PG glycopeptides enabled the probing of the molecular mechanism for downstream extension of the tetrasaccharide linkage region for biosynthesis of HSPG versus CSPG.[61]

7. Conclusions and Outlook

PGs play important roles in many biological events. Significant advances have been achieved during the last two decades regarding the mechanisms of tetrasaccharide linkage biosynthesis. The requisite enzymes have been successfully cloned and expressed. This has led to syntheses of PG glycopeptides bearing the full tetrasaccharide. The OPME protocol can be particularly attractive with high overall synthetic efficiency. The availability of homogeneous PG glycopeptides will enable more in-depth studies on the intricate regulatory machinery governing the biosynthesis of PGs bearing full GAG chains as well as in vitro assembly of biologically active GAG. Furthermore, it would be intriguing to develop engineered enzymes[62] that can introduce novel chemically modified sugars with reactive handles into PGs to facilitate further applications. This offers exciting potential to not only shine light on fundamental understanding of PG functions, but also develop novel therapeutics to address PG-associated biomedical challenges.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Po-han Lin: conceptualization (equal); writing—original draft (lead); writing—review and editing (supporting). **Xuefei Huang**: conceptualization (equal); supervision (lead); writing—original draft (supporting); writing—review and editing (lead).

Keywords: biosynthesis · chemoenzymatic synthesis · enzymes · linkage region · proteoglycan

[1] V. H. Pomin, B. Mulloy, *Pharmaceuticals* **2018**, *11*, 27.

[2] S. Tumova, A. Woods, J. R. Couchman, *Int. J. Biochem. Cell Biol.* **2000**, *32*, 269.

[3] J. R. Bishop, M. Schuksz, J. D. Esko, *Nature* **2007**, *446*, 1030.

[4] L. T. Galindo, A. M. Mundim, A. S. Pinto, G. M. D. Chiarantin, M. E. S. Almeida, M. L. Lamers, A. R. Horwitz, M. F. Santos, M. Porcionatto, *Mol. Neurobiol.* **2018**, *55*, 3185.

[5] L. Li, M. Ly, R. J. Linhardt, *Mol. BioSyst.* **2012**, *8*, 1613.

[6] A. Oohira, F. Matsui, Y. Tokita, S. Yamauchi, S. Aono, *Arch. Biochem. Biophys.* **2000**, *374*, 24.

[7] D. Xu, J. D. Esko, *Annu. Rev. Biochem.* **2014**, *83*, 129.

[8] L. Fransson, I. Silverberg, I. Carlstedt, *J. Biol. Chem.* **1985**, *260*, 14722.

[9] D. C. Briggs, E. Hohenester, *Structure* **2018**, *26*, 801.

[10] K. Prydz, K. T. Dalen, *J. Cell Sci.* **2000**, *113*, 193.

[11] A. Siegbahn, S. Manner, A. Persson, E. Tykesson, K. Holmqvist, A. Ochocinska, J. Rönnols, A. Sundin, K. Mani, G. Westergren-Thorsson, G. Widmalm, U. Ellervik, *Chem. Sci.* **2014**, *5*, 3501.

[12] J. Gao, P.-H. Lin, S. T. Nick, J. Huang, E. Tykesson, U. Ellervik, L. Li, X. Huang, *Org. Lett.* **2021**, *23*, 1738.

[13] T. Koike, T. Izumikawa, J. Tamura, H. Kitagawa, *Biochem. J.* **2009**, *421*, 157.

[14] S. E. Cole, M. S. Mao, S. H. Johnston, T. F. Vogt, *Mamm. Genome* **2001**, *12*, 177.

[15] T. Koike, T. Izumikawa, B. Sato, H. Kitagawa, *J. Biol.Chem.* **2014**, *289*, 6695.

[16] L. C. Pedersen, K. Tsuchida, H. Kitagawa, K. Sugahara, T. A. Darden, M. Negishi, *J. Biol. Chem.* **2000**, *275*, 34580.

[17] T. Izumikawa, T. Uyama, Y. Okuura, K. Sugahara, H. Kitagawa, *Carbohydr. Polym.* **2007**, *403*, 545.

[18] T. Uyama, H. Kitagawa, J.-I. Tamura, K. Sugahara, *J. Biol.Chem.* **2002**, *277*, 8841.

[19] H. Kitagawa, H. Shimakawa, K. Sugahara, *J. Biol.Chem.* **1999**, *274*, 13933.

[20] T. Lind, F. Tufaro, C. McCormick, U. Lindahl, K. Lidholt, *J. Biol. Chem.* **1998**, *273*, 26265.

[21] N. S. Gandhi, R. L. Mancera, *Chem. Biol. Drug Des.* **2008**, *72*, 455.

[22] W. Yang, Y. Eken, J. Zhang, L. E. Cole, S. Ramadan, Y. Xu, Z. Zhang, J. Liu, A. K. Wilson, X. Huang, *Chem. Sci.* **2020**, *11*, 6393.

[23] S. K. Kim, M. A. Henen, A. P. Hinck, *Exp. Biol. Med.* **2019**, *244*, 1547.

[24] R. V. Iozzo, *J. Clin. Invest.* **2001**, *108*, 165.

[25] M. E. Herndon, C. S. Stipp, A. D. Lander, *Glycobiology* **1999**, *9*, 143.

[26] T. R. O'Leary, M. Critcher, T. N. Stephenson, X. Yang, A. A. Hassan, N. M. Bartfield, R. Hawkins, M. L. Huang, *Nat. Chem. Biol.* **2022**, *18*, 634.

[27] J. Gao, X. Huang, *Advances in Carbohydrate Chemistry and Biochemistry*, (Ed: D. C. Baker), *80*, Academic Press, Cambridge, MA **2021**, pp. 95–119.

[28] B. F. Eames, Y.-L. Yan, M. E. Swartz, D. S. Levic, E. W. Knapik, J. H. Postlethwait, C. B. Kimmel, *PLoS Genet.* **2011**, *7*, e1002246.

[29] P. Vogel, G. Hansen, R. Read, R. Vance, M. Thiel, J. Liu, T. Wronski, D. Smith, S. Jeter-Jones, R. Brommage, *Vet. Pathol.* **2012**, *49*, 998.

[30] P. Ma, W. Yan, Y. Tian, J. Wang, J. Q. Feng, C. Qin, Y.-S. L. Cheng, X. Wang, *Sci. Rep.* **2016**, *6*, 29814.

[31] C. A. Worby, J. E. Mayfield, A. J. Pollak, J. E. Dixon, S. Banerjee, *J. Biol. Chem.* **2021**, *296*, 100267.

[32] V. S. Tagliabracci, J. L. Engel, J. Wen, S. E. Wiley, C. A. Worby, L. N. Kinch, J. Xiao, N. V. Grishin, J. E. Dixon, *Science* **2012**, *336*, 1150.

[33] P. Vogel, G. M. Hansen, R. W. Read, R. B. Vance, M. Thiel, J. Liu, T. J. Wronski, D. D. Smith, S. Jeter-Jones, R. Brommage, *Vet. Pathol.* **2012**, *49*, 998.

[34] Y. Tian, P. Ma, C. Liu, X. Yang, D. M. Crawford, W. Yan, D. Bai, C. Qin, X. Wang, *Eur. J. Oral Sci.* **2015**, *123*, 396.

[35] J. Wu, A. T. Bollinger, X. He, G. D. Gu, H. Miao, M. P. M. Dean, I. K. Robinson, I. Božović, *J. Supercond. Nov. Magn.* **2020**, *33*, 87.

[36] X. Liu, N. Li, H. Zhang, J. Liu, N. Zhou, C. Ran, X. Chen, Y. Lu, X. Wang, C. Qin, J. Xiao, C. Liu, *Eur. J. Oral Sci.* **2018**, *126*, 433.

[37] D. Nalbant, H. Youn, S. I. Nalbant, S. Sharma, E. Cobos, E. G. Beale, Y. Du, S. C. Williams, *BMC Genom.* **2005**, *6*, 11.

[38] S. Nadanaka, S. Zhou, S. Kagiyama, N. Shoji, K. Sugahara, K. Sugihara, M. Asano, H. Kitagawa, *J. Biol.Chem.* **2013**, *288*, 9321.

[39] J. Wen, J. Xiao, M. Rahdar, B. P. Choudhury, J. Cui, G. S. Taylor, J. D. Esko, J. E. Dixon, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15723.

[40] H. Zhang, Q. Y. Zhu, J. X. Cui, Y. X. Wang, M. J. Chen, X. Guo, V. S. Tagliabracci, J. E. Dixon, J. Y. Xiao, *Nat. Comm.* **2018**, *9*, https://doi.org/10.1038/s41467-41018-03615-z.

[41] S. Nadanaka, H. Kitagawa, K. Sugahara, *J. Biol. Chem.* **1998**, *273*, 33728.

[42] P. Trejo, F. Rauch, F. H. Glorieux, J. Ouellet, T. Benaroch, P. M. Campeau, *Mol. Syndromol.* **2017**, *8*, 303.

[43] S. Han, X. Xu, J. Wen, J. Wang, S. Xiao, L. Pan, J. Wang, *Front. Pediatr.* **2022**, *10*, https://doi.org/10.3389/fped.2022.1073748.

[44] M. Nakajima, S. Mizumoto, N. Miyake, R. Kogawa, A. Iida, H. Ito, H. Kitoh, A. Hirayama, H. Mitsubuchi, O. Miyazaki, R. Kosaki, R. Horikawa, A. Lai, R. Mendoza-Londono, L. Dupuis, D. Chitayat, A. Howard, Gabriela F. Leal, D. Cavalcanti, Y. Tsurusaki, H. Saitsu, S. Watanabe, E. Lausch, S. Unger, L. Bonafé, H. Ohashi, A. Superti-Furga, N. Matsumoto, K. Sugahara, G. Nishimura, et al., *Am. J. Hum. Genet.* **2013**, *92*, 927.

[45] K. W. Moremen, A. Ramiah, M. Stuart, J. Steel, L. Meng, F. Forouhar, H. A. Moniz, G. Gahlay, Z. Gao, D. Chapla, S. Wang, J.-Y. Yang, P. K. Prabhakar, R. Johnson, M. D. Rosa, C. Geisler, A. V. Nairn, J. Seetharaman, S.-C. Wu, L. Tong, H. J. Gilbert, J. LaBaer, D. L. Jarvis, *Nat. Chem. Biol.* **2018**, *14*, 156.

[46] P.-H. Lin, Y. Xu, S. K. Bali, J. Kim, A. Gimeno, E. T. Roberts, D. James, N. M. S. Almeida, N. Loganathan, A. K. Wilson, J. Amster, K. W. Moremen, J. Liu, J. Jiménez-Barbero, X. Huang, *Angew. Chem. Int. Ed.* **2024**, *63*, e202405671.

[47] Y.-H. Chen, Y. Narimatsu, T. M. Clausen, C. Gomes, R. Karlsson, C. Steentoft, C. B. Spliid, T. Gustavsson, A. Salanti, A. Persson, A. Malmström, D. Willén, U. Ellervik, E. P. Bennett, Y. Mao, H. Clausen, Z. Yang, *Nat. Methods* **2018**, *15*, 881.

[48] T. Van Damme, X. Pang, B. Guillemyn, D. Gulberti, D. Syx, R. De Rycke, O. Kaye, C. E. M. de Die-Smulders, R. Pfundt, A. Kariminejad, S. Nampoothiri, G. Pierquin, S. Bulk, A. A. Larson, K. C. Chatfield, M. Simon, A. Legrand, M. Gerard, S. Symoens, S. Fournel-Gigleux, F. Malfait, *Hum. Mol. Genet.* **2018**, *27*, 3475.

[49] A. Persson, J. Nilsson, E. Vorontsov, F. Noborn, G. Larson, *Glycobiology* **2019**, *29*, 366.

[50] S. Baasanjav, L. Al-Gazali, T. Hashiguchi, S. Mizumoto, B. Fischer, D. Horn, D. Seelow, Bassam R. Ali, Samir A. A. Aziz, R. Langer, Ahmed A. H. Saleh, C. Becker, G. Nürnberg, V. Cantagrel, Joseph G. Gleeson, D. Gomez, J.-B. Michel, S. Stricker, Tom H. Lindner, P. Nürnberg, K. Sugahara, S. Mundlos, K. Hoffmann, *Am. J. Med. Genet.* **2011**, *89*, 15.

[51] H. Kitagawa, Y. Tone, J.-I. Tamura, K. W. Neumann, T. Ogawa, S. Oka, T. Kawasaki, K. Sugahara, *J. Biol. Chem.* **1998**, *273*, 6615.

[52] G. Wei, X. Bai, A. K. Sarkar, J. D. Esko, *J. Biol. Chem.* **1999**, *274*, 7857.

[53] X. Bai, G. Wei, A. Sinha, J. D. Esko, *J. Biol. Chem.* **1999**, *274*, 13017.

[54] M. Ouzzine, S. Gulberti, P. Netter, J. Magdalou, S. Fournel-Gigleux, *J. Biol.Chem.* **2000**, *275*, 28254.

[55] M. Ouzzine, S. Gulberti, N. Levoin, P. Netter, J. Magdalou, S. Fournel-Gigleux, *J. Biol. Chem.* **2002**, *277*, 25439.

[56] T. Yuko, C. P. Lars, Y. Tomoko, I. Tomomi, K. Hiroshi, N. Junko, T. Jun-ichi, N. Masahiko, S. Kazuyuki, *J. Biol. Chem.* **2008**, *283*, 16801.

[57] M. Mende, C. Bednarek, M. Wawryszyn, P. Sauter, M. B. Biskup, U. Schepers, *Chem. Rev.* **2016**, *116*, 8193, and references cited therein.

[58] B. Yang, K. Yoshida, Z. Yin, H. Dai, H. Kavunja, M. H. El-Dakdouki, S. Sungsuwan, S. B. Dulaney, X. Huang, *Angew. Chem. Int. Ed.* **2012**, *51*, 10185.

[59] W. Yang, K. Yoshida, B. Yang, X. Huang, *Carbohydr. Res.* **2016**, *435*, 180.

[60] M. Bourgeais, F. Fouladkar, M. Weber, E. Boeri-Erba, R. Wild, *Glycobiology* **2024**, *34*, cwae016.

[61] D. Sammon, A. Krueger, M. Busse-Wicher, R. M. Morgan, S. M. Haslam, B. Schumann, D. C. Briggs, E. Hohenester, *Nat. Commun.* **2023**, *14*, 6425.

[62] Z. Li, L. Di Vagno, H. Chawla, A. Ni Cheallaigh, M. Critcher, D. Sammon, D. C. Briggs, N. Chung, V. Chang, K. E. Mahoney, A. Cioce, L. D. Murphy, Y.-H. Chen, Y. Narimatsu, R. L. Miller, L. I. Willems, S. A. Malaker, M. L. Huang, G. J. Miller, E. Hohenester, B. Schumann, *bioRxiv* **2024**, 2023.12.20.572522.