CHAPTER

8 Heterologous expression of plant glycosyltransferases for biochemistry and structural biology

Pradeep K. Prabhakar[a,b,c], Hsin-Tzu Wang[a,b,c], Peter J. Smith[a,b,c], Jeong-Yeh Yang[a,b], William J. Barnes[a,b], Maria J. Peña[b,c], Kelley W. Moremen[a,b], and Breeanna R. Urbanowicz[a,b,c,*]

[a]*Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, United States*

[b]*Complex Carbohydrate Research Center, University of Georgia, Athens, GA, United States*

[c]*Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oakridge, TN, United States*

**Corresponding author: e-mail address: breeanna@uga.edu*

Chapter outline

Methods in Cell Biology, Volume 160, ISSN 0091-679X, https://doi.org/10.1016/bs.mcb.2020.05.002

Abstract

Much of the carbon captured by photosynthesis is converted into the polysaccharides that constitute plant cell walls. These complex macrostructures are composed of cellulose, hemicellulose, and pectins, together with small amounts of structural proteins, minerals, and in many cases lignin. Wall components assemble and interact with one another to produce dynamic structures with many capabilities, including providing mechanical support to plant structures and determining plant cell shape and size. Despite their abundance, major gaps in our knowledge of the synthesis of the building blocks of these polymers remain, largely due to ineffective methods for expression and purification of active synthetic enzymes for *in vitro* biochemical analyses. The hemicellulosic polysaccharide, xyloglucan, comprises up to 25% of the dry weight of primary cell walls in plants. Most of the knowledge about the glycosyltransferases (GTs) involved in the xyloglucan biosynthetic pathway has been derived from the identification and carbohydrate analysis of knockout mutants, lending little information on how the catalytic biosynthesis of xyloglucan occurs *in planta*. In this chapter we describe methods for the heterologous expression of plant GTs using the HEK293 expression platform. As a demonstration of the utility of this platform, nine xyloglucan-relevant GTs from three different CAZy families were evaluated, and methods for expression, purification, and construct optimization are described for biochemical and structural characterization.

1 Introduction

The plant cell wall is a complex matrix that is largely composed of polysaccharides, including cellulose, hemicelluloses, and pectin. The synthesis of these polysaccharides involves the coordinated action of Leloir glycosyltransferases (GTs), which catalyze the formation of glycosidic bonds by transferring sugars from activated nucleotide-sugar donors (termed glycosyl donors) to a saccharide, protein, lipid, DNA or other acceptor molecule (Rini & Esko, 2015). GTs have been classified into sequence-based families in the carbohydrate-active enzyme (CAZy) database (Lombard, Golaconda Ramulu, Drula, Coutinho, & Henrissat, 2014). Due to the complexity and structural diversity of the hemicelluloses and pectins present in plant cell walls, their biosynthesis requires the participation of many GTs that function together and with other enzymes to form complex synthetic pathways, many of which remain poorly understood.

Hemicelluloses are structurally diverse, and generally consist of a β-1,4-linked glycosyl backbone that is often, but not always, decorated with glycosyl and/or acetyl substituents. The common types of hemicelluloses found in plants include xyloglucan, xylans, mixed-linkage β-(1→3,1→4)-glucans, and mannans (Scheller & Ulvskov, 2010). Xyloglucan is the major hemicellulose present in the primary cell walls of dicots. The basic structure of xyloglucan consists of a linear backbone of β-1,4-linked D-glucan regularly substituted with α-1,6-linked D-xylose. The xylosyl residues can be further decorated with mono and disaccharides, and the structure of these sidechains is often variable across species and even within different tissues

of the same plant (Tuomivaara, Yaoi, O'Neill, & York, 2015). Xyloglucan is synthesized in the Golgi through the concerted actions of GTs with different acceptor and donor specificities. The importance of hemicelluloses is realized outside of the plant cell wall as well, as they are some of the most abundant biopolymers on Earth and act as dietary fiber for humans, confer useful rheological properties in the food industry, and are used as natural gelling agents, sugar substitutes, coating agents, and sources of high-value molecules (Saha, 2003). In order to fully valorize plant biomass for use as a renewable industrial feedstock, it is critical to better understand how we can synthesize and modify cell wall polysaccharides. However, in order to study the molecular pathways involved in the synthesis of plant cell walls, robust workflows to probe these biocatalysts *in vitro* are required.

Extensive studies of xyloglucan structural variants isolated from different plant species have demonstrated that the different sidechains on the xyloglucan backbone are reliably repeated every four to five backbone residues. Thus, to simplify the representation of xyloglucan structures, a single-letter nomenclature was developed, yielding a code that conveys the structure of a xyloglucan repeating unit (Fry et al., 1993). For example, most vascular plants contain XXXG-type xyloglucan composed of XXXG, XXFG, XXLG, or XLFG units. Here, the letter G represents an unsubstituted backbone β-D-glucose, X represents a glucose backbone residue substituted with an α-1,6-linked D-xylose, L represents a β-1,2-linked D-galactose attached to an X sidechain, and F represents an α-1,2-linked L-fucose attached to an L sidechain (Fig. 1B). To date, 19 different XyG sidechain structures have been identified (Tuomivaara et al., 2015).

The GTs involved in synthesis of xyloglucan have been identified by a combination of biochemistry and plant reverse genetics. In the model species *Arabidopsis thaliana*, previous studies have shown that XYLOGLUCAN XYLOSYL TRANSFERASE1 (XXT1), XXT2, and XXT5 are α-1,6-xylosyltransferases that

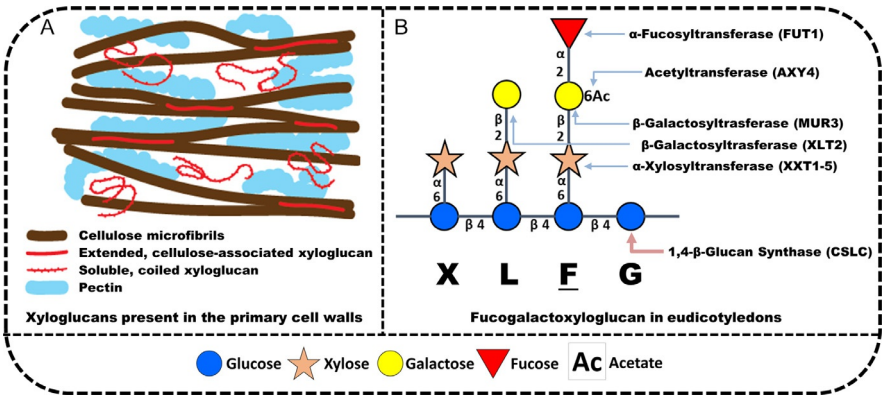


FIG. 1

Schematic of a primary cell wall where xyloglucan is the main hemicellulose (A) and a fucosylated XLFG subunit characteristic of XXXG-type xyloglucan from dicots (B).

add xylose to the β-1,4-glucan backbone produced by members of the CELLULOSE SYNTHASE-LIKE C (CSLC family; Cocuron et al., 2007). XXTs are retaining GTs from GT34 that adopt a GT-A fold (Culbertson, Ehrlich, Choe, Honzatko, & Zabotina, 2018; Faik, Price, Raikhel, & Keegstra, 2002). From there, galactosylation of the two central X sidechains can be performed by two β-(1,2)-galactosyltransferases: XYLOGLUCAN L-SIDE CHAIN GALACTOSYLTRANSFERASE POSITION 2 (XLT2) and MURUS3 (MUR3). Based on genetic experiments, XLT2 (Jensen, Schultink, Keegstra, Wilkerson, & Pauly, 2012) and MUR3 (Madson et al., 2003) proteins regiospecifically add galactose to the second and third xylosyl residues within the XXXG core structure, generating XLXG and XXLG, respectively. In specialized cell types such as the root hairs of *Arabidopsis thaliana*, a unique acidic xyloglucan is synthesized that contains galacturonic acid (GalA) in place of Gal, and these sidechains are designated by the letter Y in the single-letter nomenclature. Genetic experiments have demonstrated the involvement of XYLOGLUCAN-SPECIFIC GALACTURONOSYLTRANSFERASE (XUT1) in the formation of the β-D-galactosyluronic acid-(1→2)-α-D-xylosyl linkage at the O_2-position to the first or third xylosyl residue within the XXXG core structure (Pena, Kong, York, & O'Neill, 2012). XLT2, MUR3, and XUT1 are inverting GTs from GT47, and there are no structural determinations of representatives from this family. Finally, FUCOSYLTRANSFERASE1 (FUT1) can fucosylate the galactose in the third sidechain from the non-reducing end (XX**L**G) to produce an F sidechain (XX**F**G) (Faik et al., 2000). Structural analysis of FUT1 determined that it adopts the glycosyltransferase B (GT-B) fold, is metallo-independent, and uses an atypical water-mediated fucosylation mechanism (Rocha et al., 2016; Urbanowicz et al., 2017).

Despite genetic approaches facilitating an understanding of these and other biosynthetic enzymes, it is necessary to utilize heterologously expressed enzymes to validate that syntheses of specific glycosidic linkages are catalyzed by specific GTs (Amos & Mohnen, 2019). Many plant GTs predominantly function in membrane-bound complexes within the Golgi, and as such have been challenging to express, purify, and biochemically characterize using traditional approaches. Heterologous expression of GTs has mostly been accomplished through expression of the soluble portion of the protein containing the catalytic domain(s) and lacking the membrane anchored region (Fig. 2). One of the most popular systems for protein expression is *E. coli*. However, *E. coli* lacks the necessary biosynthetic machinery required for proper assembly, folding, and glycosylation to generate many functional eukaryotic proteins. This obstacle is broadly avoided by heterologously expressing GTs in eukaryotic expression systems such as yeast, insect cells, transgenic plants, or mammalian cells (Gupta, Dangi, Smita, Dwivedi, & Shukla, 2019). The utilization of mammalian-cell expression platforms is advantageous compared to the other systems due to better expression of glycosylated proteins with high yields. Of these, Chinese hamster ovary (CHO) and human embryonic kidney 293 cells (HEK293) have been widely used for protein production for structural biology (Bussow, 2015; Cuozzo & Soutter, 2014; Dyson, 2016; Geisse & Fux, 2009; Geisse & Voedisch, 2012;

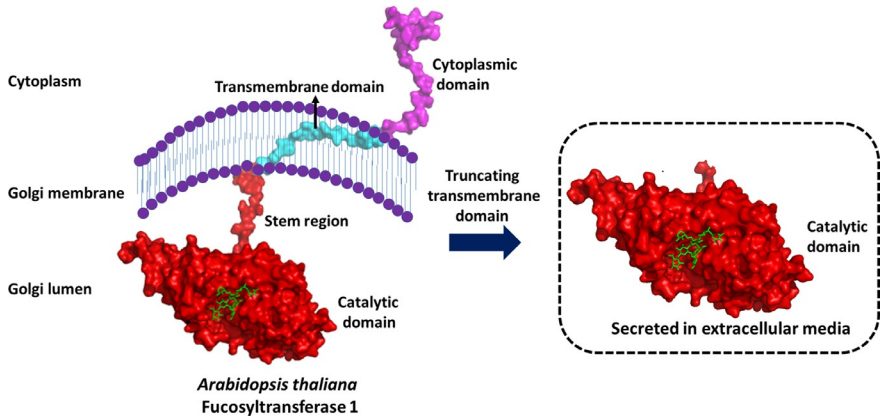


FIG. 2

Schematic representation of a type-II membrane anchored glycosyltransferase. (A) Schematic of *Arabidopsis thaliana* FUCOSYLTRANSFERASE 1 (FUT1) depicts the amino-terminal cytoplasmic region (pink), the predicted transmembrane domain (amino acids 41–63, shown in aqua blue; TMHMM Server v. 2.0), and the lumenal catalytic domain (residues 81–552). The structure of the catalytic domain (92–558 residues) of FUT1 structure is shown (5KOE) and the amino-terminal region where no structural information is available is depicted as a disordered structure for representation purpose only. An XXLG (in green) oligosaccharide is shown in the active site of FUT1.

He, Wang, & Yan, 2014; Jain et al., 2017; Owczarek, Gerszberg, & Hnatuszko-Konka, 2019; Wang & Lomino, 2012). HEK293F cells in particular have recently gained further momentum for heterologous gene expression as they are easy to grow and maintain as a suspension culture in serum-free medium, are simple to transfect, and express reproducible levels of protein from small scale flasks to large fermenters (Nettleship et al., 2015; Nigi, Fairall, & Schwabe, 2017). Glyco-engineered insect cells also provide alternative methods to express glycosylated proteins, particularly single large-domain proteins, integral membrane proteins, or proteins that require chaperones (Geisler, Mabashi-Asazuma, & Jarvis, 2015; Jarvis, 2009; Kost, Condreay, & Jarvis, 2005; Trometer & Falson, 2010).

Given the inherent complexity of xyloglucan and the number of GT families involved in its synthesis, xyloglucan biosynthetic enzymes are ideal candidates to showcase methods for expression and purification of plant GTs in HEK293 cells. The robust HEK293 protein expression system advanced by Moremen and coworkers at the University of Georgia (Moremen et al., 2018) has been used to heterologously express several plant enzymes, including GTs (Jensen et al., 2018; Ruprecht, Dallabernardina, Smith, Urbanowicz, & Pfrengle, 2018; Urbanowicz et al., 2017; Urbanowicz, Peña, Moniz, Moremen, & York, 2014), GT complexes (Amos et al., 2018), polysaccharide modifying enzymes (Urbanowicz et al., 2014), and plant lectins (Muchero et al., 2018). Previously, we have applied this methodology to identify and

characterize challenging and novel pectin biosynthetic enzymes (Amos et al., 2018; Voiniciuc et al., 2018), study the enzymatic characteristics of protein complexes (Amos et al., 2018), and determine enzyme structure and mechanisms (Amos & Mohnen, 2019; Amos et al., 2018; Urbanowicz et al., 2017, 2014). Here, we describe methods for heterologous expression and purification of plant GTs for biochemical and structural analyses using nine xyloglucan biosynthetic enzymes from three different CAZy families as a proof of concept.

2 Methods

2.1 Materials, equipment, and reagents

Freestyle™ 293 expression medium (ThermoFisher Cat. No. 12338026), Ex-cell 293 serum-free medium (Sigma Cat. No. 14571C), polyethylenimine (PEI; Polysciences Inc. Cat. No. 23966), valproic acid (VPA; Sigma Cat. No. P4543), 500 mL Erlenmeyer flasks (Corning), a biological safety cabinet II, a platform shaking (150 rpm) incubator (37 °C) with humidity (85%) and CO_2 (5%), Countess™ automated cell counter (ThermoFisher), Trypan Blue Stain (0.4% Catalog number: T10282), NaCl, HEPES, imidazole, glycerol (Sigma), 47 mm membrane filters (0.45 μm, Millipore), ultrafiltration membranes (preferably polyethersulfone (PES)) with 10-kDa molecular weight cut off (Millipore), 10% Tris-glycine SDS polyacrylamide gels and electrophoresis chamber (NuPAGE, ThermoFisher Scientific; Mini-PROTEAN®, Bio-Rad), electrophoresis power supply, Ni-NTA Resin (GE Healthcare), HisTrap HP columns (GE Healthcare), Sephadex G-75, diafiltration devices (Millipore), peristaltic pump, liquid chromatography system in a cold cabinet with a fraction collector (NGC Chromatography System, Bio-Rad; ÄKTA pure, Cytiva), Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific), SDS-PAGE Gel Running units with power pack, 4× Laemmli Sample Buffer (Bio-Rad) with reducing agent (dithiothreitol is preferred for storage), EZNA® Plasmid Mini Kit (Omega Bio-tek), and Amicon ultrafiltration cells (Millipore) were obtained from the specified vendors. For chromatography buffer preparation a glass filtration assembly is recommended comprised of the following components: a coarse porosity fritted glass filter support base with a ground glass connection (47 mm, Wheaton) with a fitted silicone stopper, regenerated cellulose membrane filters (47 mm, 0.45 μm pore size, Whatman), an aluminum clamp (Wheaton), a 250–500 mL glass funnel (Wheaton), a 1 L vacuum flask, a rubber stopper, and a vacuum source.

2.2 Cloning and construct design

Cloning was performed as previously described (Moremen et al., 2018) for expression of human GTs in HEK293 cells. GTs from three different CAZy GT families involved in the synthesis of xyloglucan sidechain structures in *Arabidopsis thaliana*, including FUT1 (At2g03220, GT37), MUR3 (At2g20370, GT47), XLT2 (At5g62220, GT47), XUT1 (At1g63450, GT47), XXT1 (At3g62720, GT34), XXT2 (At4g02500, GT34), XXT3 (At5g07720, GT34), XXT4 (At1g18690, GT34), and XXT5

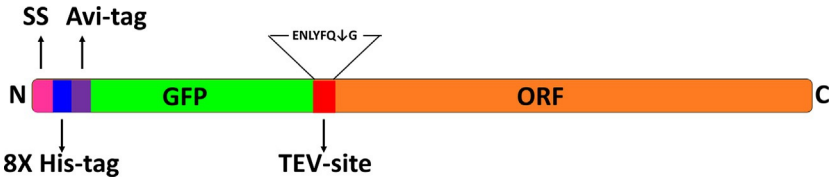


FIG. 3

Schematic representation of a typical construct used for expression of plant GTs. SS, signal sequence; GFP, green fluorescent protein; TEV-site, recognition sequence of the tobacco etch virus (TEV) protease; ORF, open reading frame.

(At1g74380, GT34) were used in this study. Final pGEn2 expression constructs are designed to encode for a fusion protein comprised of an NH_2-terminal signal sequence (SS), an 8× His-tag, an Avi-tag recognition site, "superfolder" GFP, and the recognition sequence of the tobacco etch virus (TEV) protease in frame with respect to the ORF of the catalytic domain sequence (Fig. 3).

Clone the truncated coding sequences by amplifying the region encoding the catalytic domain (includes all or part of the predicted stem region) by PCR (referred to as PCR1) using Phusion® High-Fidelity DNA Polymerase (NEB) and *Arabidopsis thaliana* cDNA as template with the primers indicated in Table 1. *Att*B sites are added to PCR1 products by a second PCR reaction (called PCR2) using universal primers and the PCR1 product as template. Gel purify the completed *attB*-PCR products using the NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) and clone into pDONR221 (ThermoFisher Scientific) using Gateway BP Clonase II Enzyme Mix (ThermoFisher Scientific) according to the manufacturer's instructions. Screen resultant colonies to identify positive transformants by colony PCR and confirm by sequence analysis using M13 universal primers: M13 (-20) Forward and M13 Reverse (ThermoFisher Scientific). Generate expression clones by recombining the entry clones into pGEn2-DEST (Moremen et al., 2018) using Gateway LR Clonase II Enzyme Mix (ThermoFisher Scientific) and confirm by colony PCR using restriction fragment analysis.

2.3 DNA isolation and preparation

For transfection, purify plasmids using the PureLink HiPure Plasmid Filter Maxiprep Kit (ThermoFisher) according to the manufacturer's instructions with the following modifications. For Maxiprep-scale DNA isolation (~1–1.5 mg) of pGEn2-based constructs, inoculate 300–500 mL of LB media supplemented with ampicillin (100 mg/L) in a 2 L baffled flask with an overnight culture and grow for 16 h. Use the bacterial pellet from 250 mL for DNA isolation. After treating the final DNA pellet with 80% ethanol, carry out all steps under sterile conditions in a laminar flow hood. Dissolve the ethanol washed DNA pellet in 500 µL cell culture grade sterile water (Corning), not TE buffer. To verify sterility of DNA prior to transfections, add 10 µL of the DNA in solution, or sterile water as a control, to 1 mL of LB broth

Table 1 List of primers used to amplify genes in this study.

Gene	Details	Primer sequence 5′-3′
FUT1	AtFUT1_81F[a]	AACTTGTACTTTCAAGGCGGAGTTTTCCCAAATGTTA
	AtFUT1_558R[a]	ACAAGAAAGCTGGGTC**CTA**TACTAGCTTAAGTCCCCA
MUR3	AtMUR3 57F	AACTTGTACTTTCAAGGCAGTAACATTGATAAACAGC
	AtMUR3 619R	ACAAGAAAGCTGGGTC**CTA**CTGTGTCTTATCTCTCTG
XLT2	AtXLT2 91F	AACTTGTACTTTCAAGGCGTCACCACCACCGTCACA
	AtXLT2 517F	ACAAGAAAGCTGGGTC**CTA**TCTCCATTTGTACCATTTCTC
XUT1	AtXUT 105F	AACTTGTACTTTCAAGGCAAGCTCACTGACCTTTAT
	AtXUT 661R	ACAAGAAAGCTGGGTC**CTA**TGCAATCTTCTTGAATAAAC
XXT1	AtXXT1AA51F[b]	AACTTGTACTTTCAAGGCATCGAGGAGATCCGTGAG
	AtXXT1 AA457R[b]	ACAAGAAAGCTGGGTC**CTA**CGTACTAAGCTTGGCCG
XXT2	AtXXT2 AA48F	AACTTGTACTTTCAAGGCGAGCAAGATCTTGACGAG
	AtXXT2 AA462R	ACAAGAAAGCTGGGTC**CTA**GTACAAACCAATCAAGTT
XXT3	AtXXT3 AA71F	AACTTGTACTTTCAAGGCAACTTCGGAACTTCCGAC
	AtXXT3 AA457R	ACAAGAAAGCTGGGTC**CTA**GCTTCGTGCTTCAATCTT
XXT4	AtXXT4 AA62F	AACTTGTACTTTCAAGGCATCCGTGTTGGGAACCTT
	AtXXT4 AA509R	ACAAGAAAGCTGGGTC**CTA**CATGTGCATTCTTCGAAT
XXT5	AtXXT5 AA71F	AACTTGTACTTTCAAGGCGGTAACCTAGGAAGCTCA
	AtXXT5 AA457R	ACAAGAAAGCTGGGTC**CTA**GTTCTGTGGTTTGGTTTC
OptXLT2	OptXLT2 92F	AACTTGTACTTTCAAGGCACCACAACTGTGACCAC
	OptXLT2 517R	ACAAGAAAGCTGGGTC**CTA**CCTCCACTTATACCA
Optv2XLT2	OptXLT2 103F	AACTTGTACTTTCAAGGCGCTGCATCTAGTAACC
	OptXLT2 517R	ACAAGAAAGCTGGGTC**CTA**CCTCCACTTATACCA
Universal Primer 1	–	GGGGACAAGTTTGTACAAAAAAGCAGGCTC TGAAAACTTGTACTTTCAAGGC
Universal Primer 2	–	GGGGACCACTTTGTACAAGAAAGCTGGGTC

The primers were designed to amplify the amino acid (AA) residues according to the full-length peptide sequence indicated by the numbering in the primer name. Underlined sequences denote the partial attB *adapter sequences appended to the primers used in the first round of PCR amplification, and the bold sequence indicates the inserted STOP codon*

[a]*Primers are published in Dallabernardina et al. (2017).*

[b]*Primers equivalent to FUT1_pDONR_GS-F and FUT1_pDONR_GS-R from Urbanowicz et al. (2017).*

in a culture tube, shake at 250 rpm overnight at 37 °C, and check for bacterial growth. Clear cultures indicate that the DNA is sterile. Dilute DNA 1:10 in sterile water and quantify its concentration using a NanoDrop Microvolume Spectrophotometer (ThermoFisher).

Note: When larger quantities of DNA (~10–15 mg) are needed, the PureLink™ HiPure Expi Plasmid Gigaprep Kit (ThermoFisher) can be used. All procedures are the same as those above using 2.5 L of an overnight bacterial culture, and the final pellet is resuspended in 2 mL of cell culture grade water.

2.4 Transfection

1. Grow and maintain HEK293F cells at 0.5–3.0 × 10^6 cells/mL in a humidified CO_2 platform shaker incubator at 37 °C in Freestyle 293 expression medium (ThermoFisher). Check cell density and viability daily using a Countess automated cell counter (ThermoFisher) by mixing 10 µL of sample with 10 µL of trypan blue according to the manufacturer's instructions. Perform transfections only when culture viability is >95%.
2. Two days before transfection, seed HEK293F cells at a cell density of 1.0 × 10^6 cells/mL in Freestyle 293 expression medium.
3. Day of transfection—Count cells using an automated cell counter and calculate the total number of cells required for a 50 mL (for 100 mL transfection, you require 50 mL of cells) transfection using the following formula:

$$Total\ cells\ required = \frac{2.5 \text{ x } 10^6 \times \text{Final volume}}{\text{Cell count}}$$

4. Spin cells at 330 × *g* for 10 min at room temperature and resuspend cells in 50 mL fresh Freestyle 293 expression medium.

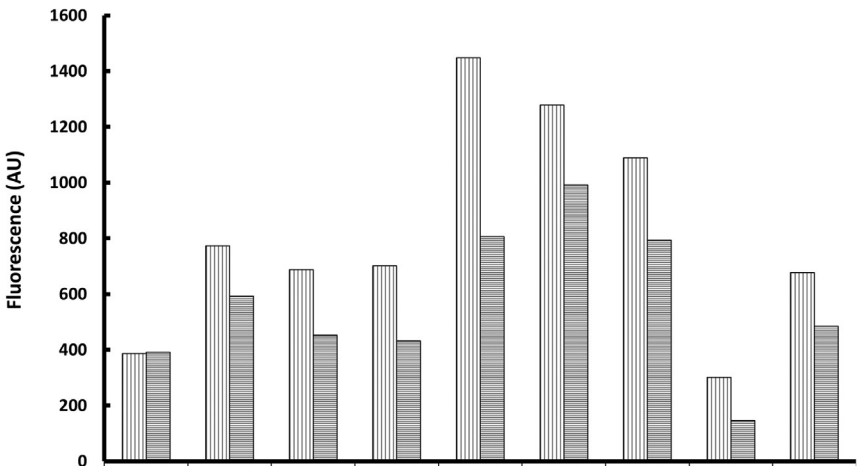
   Note: Resuspension in fresh media prior to transfection is critical. Conditioned medium contains metabolites that inhibit transfection.
5. Remove medium from the cell pellet and resuspend the pellet in 50 mL of fresh Freestyle 293 expression medium.
6. Add plasmid DNA to the cells at a final concentration of 4 µg/mL of transfection volume and swirl to mix the DNA. Return the flask to the shaker platform in the incubator for 5 min.
7. Add polyethylenimine (PEI) to a final concentration of 9 µg/mL of transfection volume and swirl the culture to mix the PEI, return the flask to the shaker platform in the incubator.
8. After 24 h, dilute the cells 1:1 with (50 mL for a 100 mL transfection) pre-warmed (37 °C) 9:1 Freestyle 293 expression and Ex-cell 293 serum-free medium supplemented with valproic acid (VPA) to a final concentration of 2.2 mM.

   Note: No further media supplementation is required for the duration of the transfection.
9. Return the culture flask(s) to the CO_2 incubator on the orbital shaker. Harvest the transfected cells 6 days after transfection.

2.5 Estimation of protein production

The incorporation of the "superfolder" GFP (sfGFP) domain has been shown to improve protein production and secretion for several mammalian GTs tested in both human and insect cell culture expression systems (Moremen et al., 2018). The pGEn2 vector was selected for expression of Golgi-localized plant GTs that are predicted to be single pass Type-II transmembrane proteins. Genes cloned into pGEn2 are NH_2-terminal GFP fusion proteins constructed by replacing the predicted NH_2-terminal transmembrane domains of candidate GTs. In addition to its positive effect on fusion protein solubility, sfGFP as a fusion partner facilitates protein quantitation during expression and purification, aiding in estimation of the total amounts of protein produced and secreted into the extracellular medium.

1. Six days after transfection, transfer 100 µL of the culture into a 96-well black flat bottom polystyrene microplate. This is referred to as the Media + Cells fraction (Fig. 4), and represents total protein expression.
2. Transfer 200 µL of the culture to a 1.5 mL microcentrifuge tube and centrifuge at $150 \times g$ for 5 min, then carefully transfer the supernatant to a new tube and centrifuge again at $1500 \times g$. Next, transfer 100 µL of the supernatant (extracellular medium) into a well of a 96-well black flat bottom polystyrene microplate, being careful not to introduce any bubbles. This is referred to as the Media fraction (Fig. 4) and represents total secreted sfGFP-fusion protein.



	FUT1	MUR3	XLT2	XUT1	XXT1	XXT2	XXT3	XXT4	XXT5
Media + Cells	386	773	687	701	1448	1278	1089	301	677
Media	391	593	452	432	806	992	793	146	484

FIG. 4

Expression and secretion of GTs involved in xyloglucan biosynthesis as measured by sfGFP fluorescence.

3. Measure fluorescence of the Media + Cells and the Media fractions at 515 nm emission (excitation at 450 nm) using a microplate reader equipped to measure fluorescence, such as a BioTek Synergy H1 Hybrid Multi-Mode Reader (Fig. 4).
4. Using a standard curve of purified sfGFP, which can be generated by expression of the empty pGEn2-DEST vector and purification of the 8× His-tagged "superfolder" GFP protein, it was estimated that 13.1 Fluorescent Units (FU) represents 1 mg/L of fusion protein. Results will vary on different plate readers; thus, this value should be empirically determined by comparing the florescence value to the concentration of a protein standard curve calculated by another method, such as a BCA assay.
5. Estimate the amount of the protein with the following formula.

$$\textit{Estimated amount of protein in the media in mg/L} = \frac{\text{Measured Media Fluorescence}}{13.1}$$

Note: This method can also be used to check and optimize protein production by aseptically removing 100 µL aliquots from transfected cultures at 24 h intervals, clarifying the media by centrifugation, and analyzing by SDS-PAGE and/or fluorescence measurement to determine the optimal time needed for expression.

2.6 Purification of heterologously expressed protein from extracellular media

Prepare three buffers for purification of recombinant proteins. All buffers used for chromatography must be filtered and degassed. Filter solutions using a setup comprised of a coarse porosity fritted glass filter support base (Wheaton) topped with a regenerated cellulose membrane filter (47 mm, 0.45 µm pore size, Whatman) clasped with an aluminum clamp (Wheaton) to a 500 mL glass funnel (Wheaton) that is fitted onto a 1 L vacuum flask attached to vacuum source via a non-collapsible rubber hose. Transfer buffer to upper glass funnel and filter into the vacuum flask by applying low vacuum. Remove filtration device, add a stir-bar, seal top of flask with rubber stopper, attach vacuum hose to sidearm of vacuum flask, turn on vacuum to low, and stir solution on a stir plate (100 rpm) to promote release of gases for 60 min to overnight. Degassing is complete when it no longer appears to "boil." Chill buffers to 4 °C prior to use.

1. 10× Buffer A: 500 mM HEPES free acid, 4 M NaCl, and 200 mM Imidazole, pH 7.2

   Buffer A: 50 mM HEPES free acid, 400 mM NaCl, and 20 mM Imidazole, pH 7.2

   Buffer B: 50 mM HEPES free acid, 400 mM NaCl, and 500 mM Imidazole, pH 7.2

   Note: Proper preparation of buffers is critical for chromatography experiments. Dissolved gases in buffers can come out of solution (outgas) and

form bubbles within the resin bed, interfering with column function and resulting in decreased binding capacity and elution efficiency.

2. Harvest extracellular medium 6 days post-transfection by sequential centrifugation at $377 \times g$ for 15 min, $2683 \times g$ for 15 min, and $22{,}100 \times g$ for 30 min at 4 °C to remove any remaining cellular debris and larger aggregates. Corning® 250 mL polypropylene (PP) centrifuge tubes or equivalent are ideal for applications requiring large volume centrifugation. All samples must be kept on ice.
3. Measure the total volume of supernatant, and slowly add 10 × Buffer A to reach a final concentration of 1 ×. For example, for 90 mL of cell culture supernatant, 10 mL of 10 × Buffer A would be added to reach a final concentration of 1 ×, which is referred to as buffer-adjusted extracellular media. Pass buffer-adjusted extracellular media through a 5-μm cellulose filter (Pall Corporation, https://www.pall.com) followed by a 0.45-μm filter (Pall Corporation, https://www.pall.com).

   Note: This filtration step is critical to avoid damaging or locking up chromatography columns in the next steps.
4. *IMAC1 purification:* Carry out small scale purification of 8 ×-His-tagged fusion proteins from buffer-adjusted extracellular media with a HisTrap HP (1 mL column volume (CV); GE Healthcare) prepacked column on an ÄKTA Pure 25 L protein purification system (GE Healthcare) using a step gradient. Maintain the flow rate at $1\,mL\,min^{-1}$ throughout the purification. If large scale purification is required, a HisPrep FF 16/10 column (20 mL CV; GE Healthcare) can be used by adjusting the protocol scheme according to the recommended CV values with flow rates of up to $4\,mL\,min^{-1}$ (recommended). To eliminate the possibility of protein contamination, purification of different enzymes or enzyme variants should be carried out on individual columns.

   Note: For new columns, it is important to perform a blank run to remove any weakly bound Ni^{2+} ions. Wash columns in line at $1\,mL\,min^{-1}$ beginning with 5 CV of water, followed by 5 CV of 100% Buffer A, followed by a linear gradient of Buffer A to 100% Buffer B for 5 CV. Then, wash the columns with 100% Buffer B for 5 CV, followed by a linear gradient of 100% Buffer B to 100% Buffer A in 5 CV, prior to re-equilibrating with Buffer A.

   Note: Columns can be washed with water containing 0.04% (*w*/*v*) sodium azide and stored at 4 °C in 20% ethanol between uses.
5. For purification, all steps are carried out at $1\,mL\,min^{-1}$. Load adjusted media onto a HisTrap HP column pre-equilibrated with at least 5 CV Buffer A. Wash and elute the column with a step gradient of 5 CV per condition, consisting of three sequential wash steps of 0%, 10%, and 20% Buffer B, followed by two elution steps of 60% and 100% Buffer B (Fig. 5). Fractions containing GFP fluorescence (typically 60% Buffer B elution fractions), which will also be visibly green in natural light, represent those containing the target fusion protein.
6. After purification, take 30 μL aliquots of adjusted media, flow-through, wash, and eluted fractions and dilute each with 10 μL of 4 × Laemmli sample buffer (Bio-Rad) containing reducing agent. Heat the mixed samples at 95 °C for 6 min

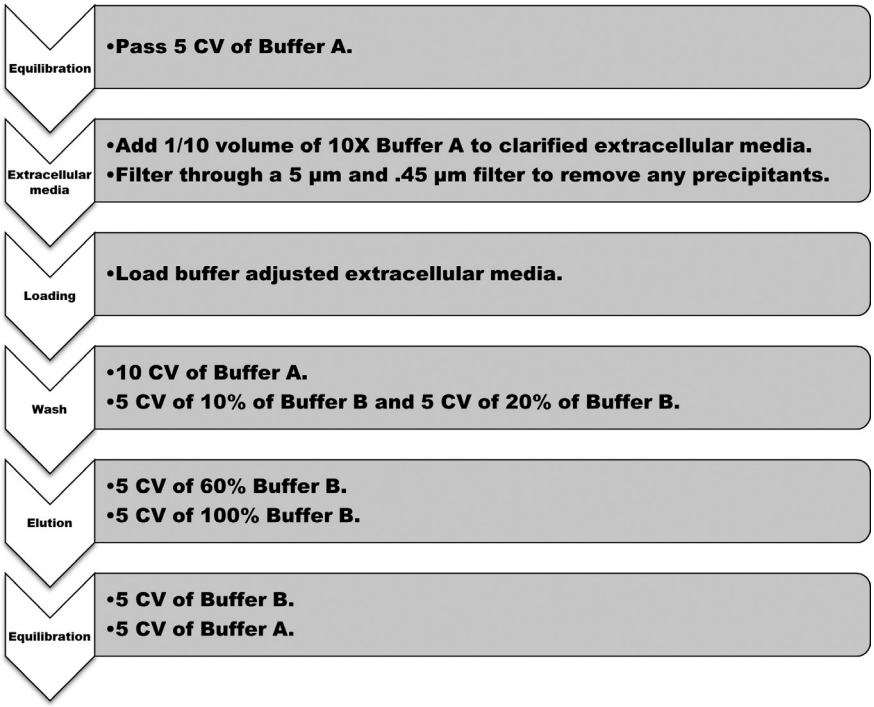


FIG. 5

Flow chart for protein purification on a HisTrap HP (1 mL).

and centrifuge at 12,500 × *g* for 2 min before performing SDS-PAGE to assess protein purity (Fig. 6).

7. *Dialysis and buffer exchange*: Further concentrate eluted protein from the IMAC1 step to approximately 2 mg/mL using a 5 mL Amicon concentrator with a 10-kDa molecular weight cutoff. At this point, if necessary, the protein can be further purified using the size exclusion chromatography protocol described in step 5 in Section 2.7 or exchanged into a new buffer for biochemical analysis or storage in step 8.
8. Exchange the protein buffer in 75 mM HEPES sodium salt pH 6.7 via dialysis (3.5 kDa MWCO) overnight in 500 mL of cold buffer at 4 °C stirring at 100 rpm. Decant buffer and replace with 500 mL of fresh cold (4 °C) buffer and continue dialysis for an additional 2–4 h an additional two times. Dialyzed samples are ready to use for assays or flash frozen.

Note: Buffer conditions for dialysis and storage must be empirically determined for each protein.

Note: Do not use regenerated cellulose dialysis membranes when working with glycoside hydrolases that cleave β-1,4-glycosidic linkages.

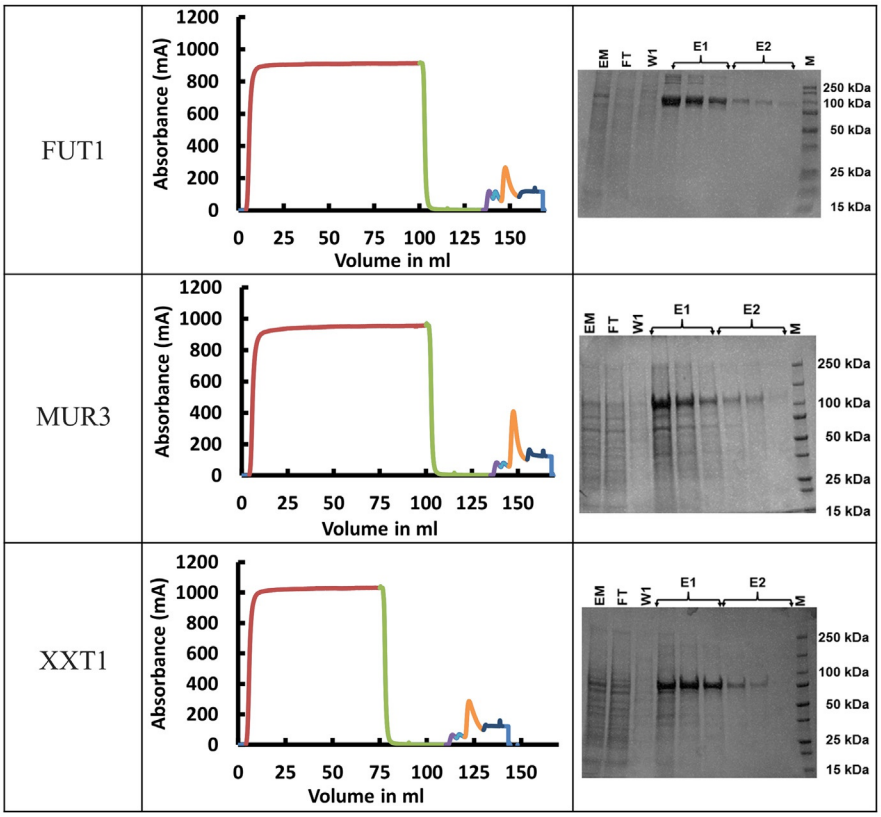


FIG. 6

IMAC1 purification of XyG glycosyltransferases. AKTA UV chromatogram (left) and SDS-PAGE of the eluted fractions to assess protein purification (right). Color code: unbound protein (red), wash (green), wash with 10% Buffer B (purple), wash with 20% Buffer B (light blue), elution with 60% Buffer B (orange), elution with 100% Buffer B (dark blue). UV is shown in milli-absorbance (mA). EM, FT, W1, E1, E2, and M represent extracellular media, flow through, wash (20% of Buffer B), elution (60% of Buffer B), elution (100% of Buffer B) and marker, respectively.

2.7 De-glycosylation and NH_2-terminal tag removal

X-ray crystallography is currently the most favored technique for studying detailed molecular interactions of acceptors and donors with GTs, such as those involved in xyloglucan biosynthesis. In this scheme, the NH_2-terminal tags are removed via cleavage by the tobacco etch virus (TEV) protease (Wu et al., 2009), which recognizes the amino acid sequence Glu-Asn-Leu-Tyr-Phe-Gln-(Gly/Ser) and cleaves between the Gln and Gly/Ser residues (Fig. 3). For structural biology, GTs are expressed in HEK293S GnTI- cells (ATCC® CRL-3022™), which limit

N-glycosylation to $Man_5GlcNAc_2$ structures. To further reduce protein complexity for structural analyses, N-glycans are cleaved with endoglycosidase F1 (EndoF1; Tarentino, Gomez, & Plummer, 1985), resulting in a single GlcNAc residue at the glycosylation site.

1. Both TEV and EndoF are heterologously expressed in *E. coli* and contain 8 ×-His tags, and thus can be generated in-house and purified using Ni-NTA (Moremen et al., 2018; Tarentino et al., 1985; Wu et al., 2009). Purified enzymes are buffer exchanged into 50 mM HEPES, pH 8.0, 100 mM NaCl and 10% glycerol, flash frozen, and stored at −80 °C for further use.
2. Pool enzymes eluted in the 60% B fractions from IMAC1, concentrate to 3 mg/mL, and exchange buffer to 50 mM HEPES buffer, pH 7.5 and 300 mM NaCl via dialysis.
3. Mix 10 μg of IMAC1 eluted protein (60% Buffer B elution, step 5 in Section 2.6) with TEV protease and EndoF in 10:1, 20:1 and 40:1 ratios (by weight) and incubate at 4 °C for 24–36 h followed by 2 h at room temperature to determine the optimum ratio for large scale digestion. Assess the extent of digestion by SDS-PAGE. The lowest ratio with complete digestion is used for the final digestion. Incubate the final TEV and EndoF digestion at 4 °C for 24–36 h, followed by 2 h at room temperature, and dilute 15-fold with 25 mM HEPES, 300 mM NaCl, pH 7.0.
4. *IMAC2:* To remove the cleaved NH_2-terminal fusion tag, His-tagged GFP-TEV protease, and His-tagged EndoF1, pass the treated sample through a second, separate IMAC column using the same chromatography scheme described for IMAC1. In this case, the unbound protein in the flow-through fraction contains the tag-free target protein, as the His-tag has been removed and the target protein should no longer associate with the IMAC resin.
5. *Size exclusion chromatography:* Concentrate fractions containing the tag-free target protein to ∼2 mg/mL using a 20 mL Amicon centrifugal filter device before loading. Load the sample using a 5 mL syringe loop on a HiLoad 16/600 Superdex 75 pg (Cytiva) column pre-equilibrated with 50 mM HEPES, 400 mM NaCl, and 20 mM Imidazole, pH 7.2. After injecting the sample, run 1.5 CV of buffer at a flow rate of 1 mL min^{-1}. Collect peak fractions based on absorbance at 280 nm and analyze by SDS-PAGE as previously described.

2.8 Considerations for optimization

In this chapter, a suite of GTs involved in xyloglucan biosynthesis were used as an example to showcase optimized methods for heterologous expression of plant proteins in HEK293 cells and associated downstream purification protocols. All of the vectors described throughout this chapter, thus far, were constructed using unmodified plant coding sequences amplified from cDNA templates. However, not all constructs display medium to high levels of expression and/or secretion using our initial, more conservative schema. For high-value targets, construct optimization can often result in improved protein secretion and increased efficiency of NH_2-terminal tag

removal for biochemical and structural studies. In our lab, we follow a general optimization strategy that is illustrated below using the xyloglucan galactosyltransferase XLT2 as an example. As a reference, the first round of expression using the original XLT2 construct (GFP-OriXLT2) resulted in expression of 34.5 mg/L (452 FU) of secreted protein based on GFP fluorescence (Fig. 5).

As a first pass, we generally clone the ORF directly from cDNA generated from the plant species of interest to generate expression constructs. In some cases, these constructs do not perform well, and we have found that generating new constructs via commercial gene synthesis with codon optimization for human cells often, but not

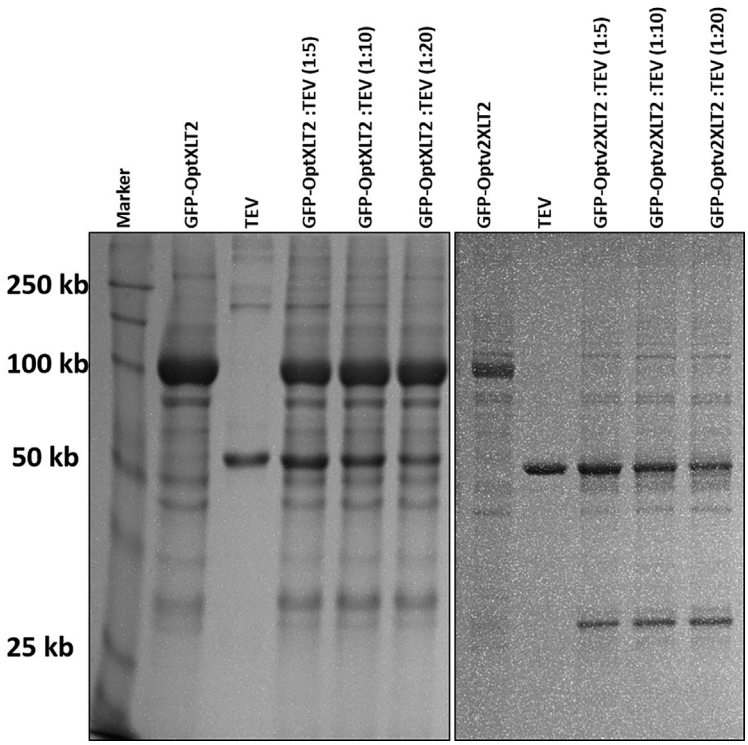


FIG. 7

SDS-PAGE of TEV protease cleavage of GFP-OptXLT2 and GFP-Optv2XLT2. Cleavage of the NH_2-terminal 8×-His-sfGFP tag (MW 29 kDa) is carried out by treatment of GFP-OptXLT2 and GFP-Optv2XLT2 with different ratios (fusion protein:TEV) of TEV protease (right). Data indicate that GFP-Optv2XLT2 is more susceptible to TEV cleavage, apparent as the appearance of a protein band (27 kDa) consistent with the size of NH_2-terminal 8×-His-sfGFP tag. In contrast, little to no cleavage of the NH_2-terminal 8×-His-sfGFP tag is observed after TEV treatment of GFP-OptXLT2, indicating that the GFP-OptXLT2 fusion protein is resistant to cleavage by TEV.

always, results in improved levels of expression and/or secretion. In this example, a codon optimized version of the full-length ORF of XLT2 (henceforth referred to as OptXLT2) was synthesized by a commercial supplier (GenScript) using their in-house codon optimization parameters and used as a PCR template to generate different truncation variants. The presence and location of transmembrane domain(s) can be predicted using servers like TMHMM 2.0 (Krogh, Larsson, von Heijne, & Sonnhammer, 2001) or the ARAMEMNON plant membrane protein database (Schwacke et al., 2003). Using the AramTmCon prediction in ARAMEMNON, XLT2 has three potential transmembrane domains spanning the following amino acid regions with the probability indicated in parentheses: 50–70 (0.94), 82–102 (0.16), and 205–225 (0.34). Two truncation variants were generated using the optimized gene sequence: $XLT2^{\Delta 92}$ (GFP-OptXLT2) and $XLT2^{\Delta 103}$ (GFP-Optv2-XLT2). $XLT2^{\Delta 92}$ was designed in between the first two hydrophobic patches, while $XLT2^{\Delta 103}$ was designed to bypass both NH_2-terminal hydrophobic regions. Fluorescence intensity of the secreted GFP-OptXLT2 and GFP-Optv2XLT2 was measured at 718 and 518 FU (HEK293S cells), which is nearly a 59% and 14% increase relative to GFP-OriXLT2, respectively.

As stated earlier, the NH_2-terminal tag is cleaved via TEV protease for further studies, such as crystallography. There are several routes we have used to increase TEV cleavage efficiency of resistant constructs, including truncation of the stem region, adding linkers, or moving a NH_2-terminal GFP to the C-terminus of the ORF. The first route we pursued was to design and evaluate different truncation variants in the stem region. As visualized by SDS-PAGE, GFP-OptXLT2 ($XLT2^{\Delta 92}$) is resistant to TEV cleavage at all tested ratios, while the slightly more truncated variant, GFP-Optv2XLT2 ($XLT2^{\Delta 103}$), is highly susceptible even at the lowest XLT2:TEV ratio of 1:20 (Fig. 7).

3 Conclusions

The Arabidopsis genome alone encodes more than 1200 carbohydrate-active enzymes, including 561 GTs (nearly 2% of total genes) distributed across 42 GT families in the CAZy database at the time of publication. Among these GTs, only a select few have been biochemically characterized due to the historical difficulties associated with expressing and purifying functional enzymes *in vitro*, recently reviewed in detail by Amos and Mohnen (2019). This chapter highlights methods based in the use of HEK293 cells to express plant GTs relevant to the biosynthesis of xyloglucan as an example. The methods laid out herein can facilitate strategies to express and screen large numbers of uncharacterized GTs and glycoside hydrolases for functional characterization, and be used to test their importance in the generation and modification of plant polysaccharide structures.

It is also worth mentioning that not all GTs express in quantities required for crystallography, or even biochemistry experiments. In these cases, several parameters can be tinkered with to optimize protein expression and secretion, including

media components, transformation methods, time of harvest, and purification conditions. Unfortunately, despite extensive efforts some protein targets still present a challenge for recombinant expression and characterization, and require additional advances in our understanding of the production of functional proteins for downstream experiments. Future optimization of expression systems and protocols may involve, for example, the generation of stable lines, the use of different cell lines or expression hosts, engineering of new transgenic cell lines, and/or altering fusion tags or secretion signals to facilitate proper folding and export. However, these challenges are outside of the scope of this chapter and should be followed up with further reading and research.

Acknowledgments

Work on plant glycosyltransferases in the authors' laboratories has been supported by the Center for Bioenergy Innovation (Oak Ridge National Laboratory), a US Department of Energy (DOE) Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science, the National Science Foundation under Grant No. 1817697, and NIH grants R01GM130915, P41GM103390, P01GM107012.

References

Amos, R. A., & Mohnen, D. (2019). Critical review of plant cell wall matrix polysaccharide glycosyltransferase activities verified by heterologous protein expression. *Frontiers in Plant Science*, *10*, 915. https://doi.org/10.3389/fpls.2019.00915.

Amos, R. A., Pattathil, S., Yang, J. Y., Atmodjo, M. A., Urbanowicz, B. R., Moremen, K. W., et al. (2018). A two-phase model for the non-processive biosynthesis of homogalacturonan polysaccharides by the GAUT1:GAUT7 complex. *The Journal of Biological Chemistry*, *293*(49), 19047–19063. https://doi.org/10.1074/jbc.RA118.004463.

Bussow, K. (2015). Stable mammalian producer cell lines for structural biology. *Current Opinion in Structural Biology*, *32*, 81–90. https://doi.org/10.1016/j.sbi.2015.03.002.

Cocuron, J.-C., Lerouxel, O., Drakakaki, G., Alonso, A. P., Liepman, A. H., Keegstra, K., et al. (2007). A gene from the cellulose synthase-like C family encodes a β-1,4 glucan synthase. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(20), 8550–8555.

Culbertson, A. T., Ehrlich, J. J., Choe, J. Y., Honzatko, R. B., & Zabotina, O. A. (2018). Structure of xyloglucan xylosyltransferase 1 reveals simple steric rules that define biological patterns of xyloglucan polymers. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(23), 6064–6069. https://doi.org/10.1073/pnas.1801105115.

Cuozzo, J. W., & Soutter, H. H. (2014). Overview of recent progress in protein-expression technologies for small-molecule screening. *Journal of Biomolecular Screening*, *19*(7), 1000–1013. https://doi.org/10.1177/1087057114520975.

Dallabernardina, P., Ruprecht, C., Smith, P. J., Hahn, M. G., Urbanowicz, B. R., & Pfrengle, F. (2017). Automated glycan assembly of galactosylated xyloglucan oligosaccharides and their recognition by plant cell wall glycan-directed antibodies. *Organic & Biomolecular Chemistry*, *15*(47), 9996–10000. https://doi.org/10.1039/c7ob02605f.

Dyson, M. R. (2016). Fundamentals of expression in mammalian cells. *Advances in Experimental Medicine and Biology*, *896*, 217–224. https://doi.org/10.1007/978-3-319-27216-0_14.

Faik, A., Bar-Peled, M., DeRocher, A. E., Zeng, W., Perrin, R. M., Wilkerson, C., et al. (2000). Biochemical characterization and molecular cloning of an alpha-1,2-fucosyltransferase that catalyzes the last step of cell wall xyloglucan biosynthesis in pea. *Journal of Biological Chemistry*, *275*(20), 15082–15089.

Faik, A., Price, N. J., Raikhel, N. V., & Keegstra, K. (2002). An *Arabidopsis* gene encoding an alpha-xylosyltransferase involved in xyloglucan biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(11), 7797–7802.

Fry, S. C., York, W. S., Albersheim, P., Darvill, A., Hayashi, T., Joseleau, J.-P., et al. (1993). An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiologia Plantarum*, *89*(1), 1–3. https://doi.org/10.1111/j.1399-3054.1993.tb01778.x.

Geisler, C., Mabashi-Asazuma, H., & Jarvis, D. (2015). An overview and history of glyco-engineering in insect expression systems. *Methods in Molecular Biology*, *1321*, 131–152.

Geisse, S., & Fux, C. (2009). Recombinant protein production by transient gene transfer into mammalian cells. *Methods in Enzymology*, *463*, 223–238. https://doi.org/10.1016/s0076-6879(09)63015-9.

Geisse, S., & Voedisch, B. (2012). Transient expression technologies: Past, present, and future. *Methods in Molecular Biology*, *899*, 203–219. https://doi.org/10.1007/978-1-61779-921-1_13.

Gupta, S. K., Dangi, A. K., Smita, M., Dwivedi, S., & Shukla, P. (2019). Effectual bioprocess development for protein production. In P. Shukla (Ed.), *Applied microbiology and bioengineering* (pp. 203–227). Academic Press. (chapter 11).

He, Y., Wang, K., & Yan, N. (2014). The recombinant expression systems for structure determination of eukaryotic membrane proteins. *Protein & Cell*, *5*(9), 658–672. https://doi.org/10.1007/s13238-014-0086-4.

Jain, N. K., Barkowski-Clark, S., Altman, R., Johnson, K., Sun, F., Zmuda, J., et al. (2017). A high density CHO-S transient transfection system: Comparison of ExpiCHO and Expi293. *Protein Expression and Purification*, *134*, 38–46. https://doi.org/10.1016/j.pep.2017.03.018.

Jarvis, D. L. (2009). Baculovirus-insect cell expression systems. In R. R. Burgess & M. P. Deutscher (Eds.), *Vol. 463. Methods in enzymology* (pp. 191–222). Academic Press. (chapter 14).

Jensen, J. K., Busse-Wicher, M., Poulsen, C. P., Fangel, J. U., Smith, P. J., Yang, J. Y., et al. (2018). Identification of an algal xylan synthase indicates that there is functional orthology between algal and plant cell wall biosynthesis. *The New Phytologist*, *218*(3), 1049–1060. https://doi.org/10.1111/nph.15050.

Jensen, J. K., Schultink, A., Keegstra, K., Wilkerson, C. G., & Pauly, M. (2012). RNA-Seq analysis of developing nasturtium seeds (Tropaeolum majus): Identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis. *Molecular Plant*, *5*(5), 984–992. https://doi.org/10.1093/mp/sss032.

Kost, T. A., Condreay, J. P., & Jarvis, D. L. (2005). Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nature Biotechnology*, *23*(5), 567–575. https://doi.org/10.1038/nbt1095.

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology*, *305*(3), 567–580. https://doi.org/10.1006/jmbi.2000.4315.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, *42*(Database issue), D490–D495. https://doi.org/10.1093/nar/gkt1178.

Madson, M., Dunand, C., Li, X., Verma, R., Vanzin, G. F., Caplan, J., et al. (2003). The MUR3 gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. *Plant Cell*, *15*(7), 1662–1670. https://doi.org/10.1105/tpc.009837.

Moremen, K. W., Ramiah, A., Stuart, M., Steel, J., Meng, L., Forouhar, F., et al. (2018). Expression system for structural and functional studies of human glycosylation enzymes. *Nature Chemical Biology*, *14*(2), 156–162. https://doi.org/10.1038/nchembio.2539.

Muchero, W., Sondreli, K. L., Chen, J. G., Urbanowicz, B. R., Zhang, J., Singan, V., et al. (2018). Association mapping, transcriptomics, and transient expression identify candidate genes mediating plant-pathogen interactions in a tree. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11573–11578. https://doi.org/10.1073/pnas.1804428115.

Nettleship, J. E., Watson, P. J., Rahman-Huq, N., Fairall, L., Posner, M. G., Upadhyay, A., et al. (2015). Transient expression in HEK 293 cells: An alternative to E. coli for the production of secreted and intracellular mammalian proteins. *Methods in Molecular Biology*, *1258*, 209–222. https://doi.org/10.1007/978-1-4939-2205-5_11.

Nigi, I., Fairall, L., & Schwabe, J. W. (2017). Expression and purification of protein complexes suitable for structural studies using mammalian HEK 293F cells. *Current Protocols in Protein Science*, *90*(1), 5–28. https://doi.org/10.1002/cpps.44.

Owczarek, B., Gerszberg, A., & Hnatuszko-Konka, K. (2019). A brief reminder of systems of production and chromatography-based recovery of recombinant protein biopharmaceuticals. *BioMed Research International*, *2019*, 4216060. https://doi.org/10.1155/2019/4216060.

Pena, M. J., Kong, Y., York, W. S., & O'Neill, M. A. (2012). A galacturonic acid-containing xyloglucan is involved in Arabidopsis root hair tip growth. *Plant Cell*, *24*(11), 4511–4524. https://doi.org/10.1105/tpc.112.103390.

Rini, J. M., & Esko, J. D. (2015). Glycosyltransferases and glycan-processing enzymes. In A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, … P. H. Seeberger (Eds.), *Essentials of glycobiology* (3rd ed., pp. 65–75). Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright 2015–2017.

Rocha, J., Ciceron, F., de Sanctis, D., Lelimousin, M., Chazalet, V., Lerouxel, O., et al. (2016). Structure of Arabidopsis thaliana FUT1 reveals a variant of the GT-B class fold and provides insight into xyloglucan fucosylation. *Plant Cell*, *28*(10), 2352–2364. https://doi.org/10.1105/tpc.16.00519.

Ruprecht, C., Dallabernardina, P., Smith, P. J., Urbanowicz, B. R., & Pfrengle, F. (2018). Analyzing xyloglucan endotransglycosylases by incorporating synthetic oligosaccharides into plant cell walls. *Chembiochem*, *19*(8), 793–798. https://doi.org/10.1002/cbic.201700638.

Saha, B. C. (2003). Hemicellulose bioconversion. *Journal of Industrial Microbiology and Biotechnology*, *30*(5), 279–291.

Scheller, H. V., & Ulvskov, P. (2010). Hemicelluloses. *Annual Review of Plant Biology*, *61*, 263–289. https://doi.org/10.1146/annurev-arplant-042809-112315.

Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., et al. (2003). ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. *Plant Physiology*, *131*(1), 16–26. https://doi.org/10.1104/pp.011577.

Tarentino, A. L., Gomez, C. M., & Plummer, T. H., Jr. (1985). Deglycosylation of asparagine-linked glycans by peptide:N-glycosidase F. *Biochemistry*, *24*(17), 4665–4671. https://doi.org/10.1021/bi00338a028.

Trometer, C., & Falson, P. (2010). Mammalian membrane protein expression in Baculovirus-infected insect cells. In I. Mus-Veteau (Ed.), *Heterologous expression of membrane proteins: Methods and protocols* (pp. 105–117). Totowa, NJ: Humana Press.

Tuomivaara, S. T., Yaoi, K., O'Neill, M. A., & York, W. S. (2015). Generation and structural validation of a library of diverse xyloglucan-derived oligosaccharides, including an update on xyloglucan nomenclature. *Carbohydrate Research*, *402*, 56–66. https://doi.org/10.1016/j.carres.2014.06.031.

Urbanowicz, B. R., Bharadwaj, V. S., Alahuhta, M., Pena, M. J., Lunin, V. V., Bomble, Y. J., et al. (2017). Structural, mutagenic and in silico studies of xyloglucan fucosylation in Arabidopsis thaliana suggest a water-mediated mechanism. *The Plant Journal*, *91*(6), 931–949. https://doi.org/10.1111/tpj.13628.

Urbanowicz, B., Peña, M., Moniz, H., Moremen, K., & York, W. (2014). Two Arabidopsis proteins synthesize acetylated xylan *in vitro*. *The Plant Journal*, *80*, 197–206. https://doi.org/10.1111/tpj.12643.

Voiniciuc, C., Engle, K. A., Günl, M., Dieluweit, S., Schmidt, M. H.-W., Yang, J.-Y., et al. (2018). Identification of key enzymes for pectin synthesis in seed mucilage. *Journal of Plant Physiology*, *178*(3), 1045–1064. https://doi.org/10.1104/pp.18.00584.

Wang, L. X., & Lomino, J. V. (2012). Emerging technologies for making glycan-defined glycoproteins. *ACS Chemical Biology*, *7*(1), 110–122. https://doi.org/10.1021/cb200429n.

Wu, X., Wu, D., Lu, Z., Chen, W., Hu, X., & Ding, Y. (2009). A novel method for high-level production of TEV protease by superfolder GFP tag. *Journal of Biomedicine & Biotechnology, 2009*. https://doi.org/10.1155/2009/591923.