HETEROLOGOUS EXPRESSION OF HUMAN ST6GAL1 IN *TRICHODERMA REESEI*

Aneesh Chandran Bhuvanachandran Pillai

Department of Chemistry and Biomolecular Sciences,

Macquarie University, Sydney, Australia

Supervisors

Prof. Helena Nevalainen

Prof. Nicki Packer

MACQUARIE University
SYDNEY·AUSTRALIA

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

July 2016

"Set your heart upon your work, but never on its reward"

- *The Bhagavad Gita*

"And when you want something, all the universe conspires in helping you to achieve it"

- *Paulo Coelho*

TABLE OF CONTENTS

ABSTRACT

Sialyltransferases can be used to add sialic acid to glycoproteins to study and modify the effects of sialylation on, for example, the metastatic potential of cancer cells, drug resistance, differentiation of stem cells and the half-life and *in vivo* bioactivity of therapeutic proteins. A functional human sialyltransferase, ST6Gal1, was expressed in the high protein secreting filamentous fungus *Trichoderma reesei* with a view of improving availability of the enzyme. Two expression cassettes were designed (i) a cassette with the *mCherry* gene which encodes a fluorescent protein to enable initial screening of transformants as well as tracking of secretion (ST6Gal1 produced as a fusion protein), and (ii) a cassette without mCherry where the ST6Gal1 production was monitored using ST6Gal1 specific antibodies. Both constructs also contained a *Strep*-tag^{TM} II and an enterokinase cleavage site to facilitate the purification of recombinant ST6Gal1 (rST6Gal1) and to cleave the rST6Gal1 from the fusion protein respectively.

A truncated and codon optimised ST6Gal1 cDNA was expressed under the *T. reesei cbh1* promoter. Recombinant ST6Gal1 was found to be secreted into the culture medium at 48 h, established by Western blotting with ST6Gal1 specific antibodies. Two main forms, 40 kDa and 60 kDa of the secreted rST6Ga11 were observed indicating processing of the fusion protein.

The secreted rST6Gal1 was eventually degraded during cultivation. A total protease activity assay revealed an increase in the extracellular protease activity over a period of 120 h of cultivation. The transformants exhibited a higher protease activity compared to the non-transformant RUT-C30 host. Subsequent analysis of the intracellular fraction by Western blotting revealed that processing of the fusion protein occurred inside the cells. Transcriptional analysis of the *ST6Gal1* gene by qRT-PCR revealed that the gene was expressed at 24 h, 48 h and 96 h.

Confocal microscopy suggested co-localisation of mCherry in the endoplasmic reticulum (ER) possibly indicating the activity of ER associated quality control mechanism. Further analysis of the RNA levels by qRT-PCR revealed an increase in the transcription levels of genes encoding proteins associated with the unfolded protein response (UPR) and ER associated degradation (ERAD) pathways, such as BiP1, PDI, calnexin, Sec61, ubiquitin and proteasome subunits RPN8 and beta3. This indicated the cellular stress that might have led to the disintegration of rST6Gal1 resulting in lower secretion.

Southern blot analysis revealed integration of multiple copies of the *ST6Gal1* gene into the fungal genomic DNA, with at least one of the transformants having the expression cassette integrated in the *cbh1* locus. A lectin based enzyme activity assay was developed and established to be functioning with commercially available recombinant human ST6Gal1. The results were further validated using mass spectrometry. As a summary, this project serves as a proof of concept for the successful expression of a recombinant human sialyltransferase ST6Gal1 in *Trichoderma reesei*.

STATEMENT OF DECLARATION

I certify that the work presented in this thesis has not been submitted as a part of the requirement for a degree or course to any institution or university other than Macquarie University.

I also certify that this thesis is an original piece of research conducted by me between July 2013 and July 2016 and it contains no material previously published or written by any other person except where due reference is made in the text.

I hereby give my consent to allow a copy of my thesis to be deposited in the University Library for consultation, loan and photocopying forthwith.

Aneesh Chandran Bhuvanachandran Pillai

July 2016

ACKNOWLEDGEMENTS

"No one who achieves success does so without acknowledging the help of others. The wise and confident acknowledge this help with gratitude" **– Alfred North Whitehead**

This thesis marks the end of my PhD work and the beginning of a new career. But, this moment would not have come true without the help and support that I got during my PhD life.

Firstly, I want to thank God for giving me the inner strength, patience and the resilience that I needed during those stressful times. I would like to express my earnest gratitude to Prof. Helena Nevalainen for giving me the opportunity to do a PhD under her supervision. Without her help and support my ambition of doing a PhD would have remained a dream. I also express my sincere thanks to Prof. Nicki Packer and Dr. Junior Te'o for their timely advices and valuable guidance.

A big thanks to Dr. Liisa Kautto for her encouragement and altruistic support throughout the course of my project. I also want to express my deepest thanks to Dr. Chi-Hung Lin for his guidance, and teaching me the basics of glycan analysis by mass spectrometry. I want to thank Dr. Angela Sun for her timely advices and help during my project. I would like to extend my gratitude to Dr. Tom Williams for helping me with qRT-PCR analysis. Thank you, Dr. Robyn Peterson for your time proof reading my thesis and fishing out those critical errors, which otherwise would have missed my eyes.

I acknowledge the use of the Microscopy unit, Faculty of Science and Engineering at Macquarie University. My deepest thanks to Dr. Nadia Suarez Bosche and Ms. Nicole Vella for their guidance and technical assistance with fluorescence microscopy. A special thanks to Ms. Elsa Mardones, Mr. Joe Gatt and Ms. Catherine Wong for their help and support whenever I needed. I also want to thank my colleagues Mafruha, Zhiping, Edward, Karthik, Ketan, Shalini, Imam, Ishan, Anna and Vinoth and my friends Vinod Krishna, Rukshana

Tadwalla, Divya Vinod, Cherag Tadwalla and Chris Wang for their wonderful companionship and the good memories that they have left in me.

Finally, I desire to express my whole-hearted gratitude to everyone in my family for their selfless love and support that kept me going during the hardest of times. Saying a mere thanks to my wife Shreeja would be unjust, who went through the pain for taking good care of my baby girl Smriti, all alone, in my absence.

1

INTRODUCTION

1.1. Glycans

Glycans are any sugar or assembly of sugars attached to a molecule, such as a glycoprotein, glycolipid or a proteoglycan. Being one of the four basic components of cells glycans are the most profuse and diverse of nature's biopolymers. They play an important role in protein folding, quality control, oligomerisation, sorting, and transport in the endoplasmic reticulum (ER) and in the early secretory pathway. Glycans also play a vital role in the mass and structural difference in biological systems (Ohtsubo and Marth, 2006). By virtue of their highly ordered structures, these sugar chains exert multiple bio-specific functions such as cell–cell communication and signal transduction.

Glycosylation is the covalent attachment of a carbohydrate to a polypeptide, lipid, polynucleotide, carbohydrate, or other organic compound, utilising specific sugar donor substrates (Dwek, 1995; Malissard *et al*., 1999). Catalysed by enzymes such as glycosyltransferases and glycosidases, glycosylation produces a diverse and highly regulated collection of cellular glycans. There are five types of known glycosylation (Table 1-1), namely, *N*-, *O*- and *C*-linked glycosylation, glypiation and phosphoglycosylation, of which *N*-glycosylation constitutes about 90% of glycosylation found in nature (Spiro, 2002). Specific chemical, enzymatic and analytic approaches including mass spectrometric (MS) and nuclear magnetic resonance (NMR) spectroscopic methods can be used in determining glycan structures (Mulloy *et al*., 2009).

Table 1-1: Various types of glycosylation found in the nature.

Types of glycosylation	
N- linked	Glycan binds to the amino group of an asparagine (Asn) residue (Stanley *et al*., 2009)
O- linked	N-acetylgalactosamine (GalNAc) monosaccharides bind to the hydroxyl group of serine or threonine residue (Lodish *et al*., 2000)
C-Mannosylation	α-mannose attaches to the C2 carbon atom in the indole ring of a tryptophan (Trp) residue to produce C-mannosyl-tryptophan (Ihara *et al*., 2015)
Glypiation	Involves the attachment of glycosylphosphatidylinositol (GPI anchor) to a protein on plasma membrane (Kobayashi *et al*., 1997)
Phosphoglycosylation	Glycan binds to serine or threonine residues via phosphodiester bond (Haynes, 1998)

1.2. Structure of glycans

1.2.1. *N*- linked glycans

N-glycans are assembly of sugars attached to the asparagine (Asn) residues of proteins or polypeptides by an *N*-glycosidic linkage (Stanley *et al*., 2009). They are formed co-translationally by adding blocks of 14 sugars, comprised of *N*-acetylglucosamine (GlcNAc), mannose (Man) and glucose (Glc). The first seven sugars are donated by sugar nucleotides UDP and GDP in the cytoplasm, and are attached to dolichol (long chain of unsaturated organic compound made up of varying numbers of isoprene units containing an alcohol functional group), by means of a pyrophosphate linkage to make the $Man_5GlcNAc_2$-PP-dolichol intermediate. After the synthesis of dolichol intermediate the final addition of sugars takes place in the ER lumen. These sugars come from Man- and Glc-P-dolichol molecules to make the $Glc_3Man_9GlcNAc_2$-PP-dolichol precursor glycan molecule. Following the transfer of the newly synthesised polypeptide into the ER, the precursor glycan molecule is transferred from $Glc_3Man_9GlcNAc_2$-PP-dolichol to the carboxamido nitrogen on asparagine of the

growing polypeptide. It is bound through an N-glycosidic bond with the nitrogen on the asparagine residue that is part of the Asn-X-Ser/Thr consensus sequence (where X is any amino acid other than proline) and is mediated by the enzyme oligosaccharide transferase (OSTase). This is followed by the removal of terminal glucose and mannose residues, in the ER, by various glucosidases and mannosidases (Stanley *et al*., 2009; Roth, 2002).

All precursor *N*-glycans share a common core sugar sequence, Manα1–6(Manα1–3)Manβ1–4GlcNAcβ1–4GlcNAcβ1-Asn-X-Ser/Thr (where X is any amino acid other than proline). The structural diversity of the mammalian glycome is produced predominantly within the Golgi apparatus in the secretory pathway of the cell during which glycans become increasingly oligomeric and branched, and are bound mostly for the cell surface and extracellular compartments. Maturation of precursor *N*-glycans occur in the Golgi apparatus (Fig.1-1). In order to differentiate the glycans on individual glycoproteins, glycans are processed in the Golgi apparatus by both trimming and adding sugars, by the action of mannosidase I and II and GlcNAc transferase respectively (Ohtsubo and Marth, 2006; Stanley *et al*., 2009).

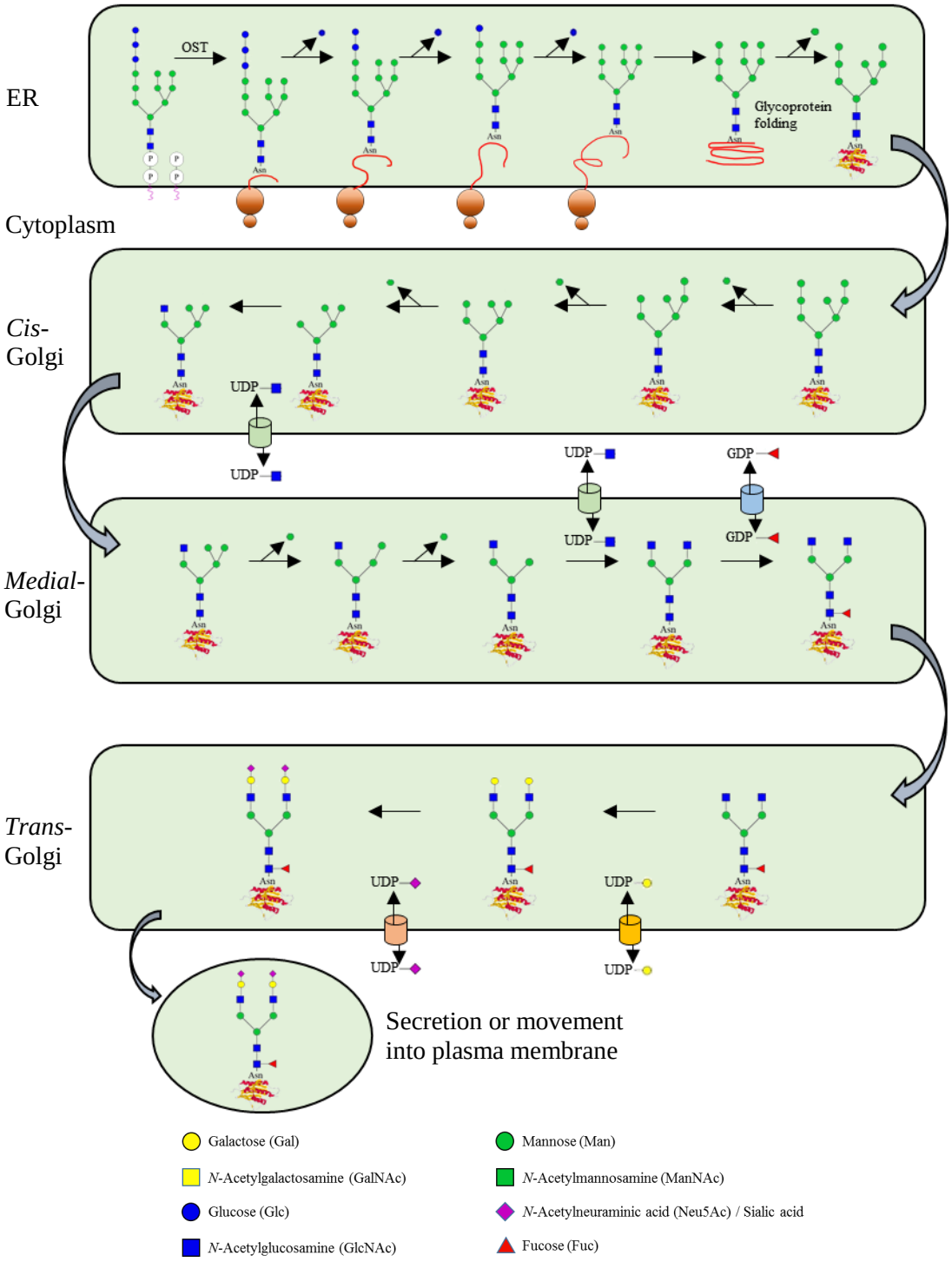


Fig 1-1: Processing and maturation of *N*-glycans. The $Glc_3Man_9GlcNAc_2$-PP-dolichol synthesised in the ER lumen is transferred to Asn-X-Ser/Thr sequons of a protein. Subsequently, glucosidases and mannosidases trim the Dol-P-P-glycan sequentially, a process often associated with proper folding of proteins assisted by the lectins calnexin and calreticulin. The properly folded proteins are then translocated to Golgi, where further trimming of mannose residues by mannosidases occur in the cis compartment to yield $Man_5GlcNAc_2Asn$. The action of GlcNAcT-1 on the $Man_5GlcNAc_2Asn$ to add a GlcNAc residue initiates the first branch of an *N*-glycan. The whole complex is afterwards translocated to the medial compartment where further trimming of mannose by mannosidase II and subsequent addition of GlcNAc by GlcNAcT-II occur to produce a biantennary structure. The resulting biantennary *N*-glycan is extended by the addition of fucose, galactose and sialic acid in the trans-Golgi compartment to generate a complex *N*-glycan with two branches (Modified from Stanley *et al*., 2009).

N-linked glycans can classified into three broad categories (Fig 1-2):

i. oligomannose, in which only mannose residues are attached to the core;

ii. complex, in which "antennae" initiated by *N*-acetylglucosaminyltransferases (GlcNAcTs) are attached to the core; and

iii. hybrid, in which only mannose residues are attached to the Manα1–6 arm of the core and one or two antennae are on the Manα1–3 arm.

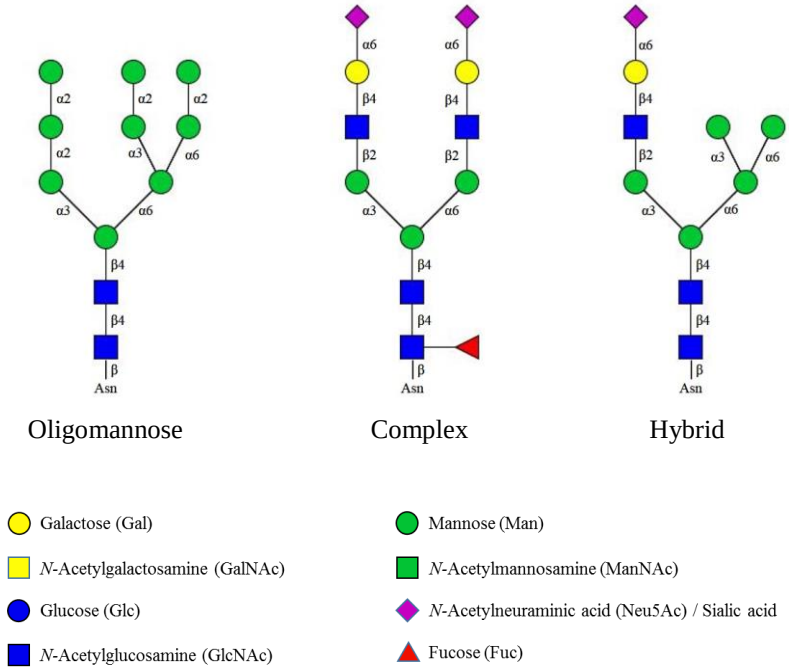


Fig 1-2: Antennae structure of *N*-glycans. **A:** oligomannose **B:** complex and **C:** hybrid structures (adapted from Stanley *et al*. (2009)). Glycan structures are generated using GlycoWorkbench.

Golgi mannosidase I and II trim the glycans that are destined to form complex oligosaccharides and are subsequently glycosylated by GlcNAc transferase, resulting in a common core region. The core then becomes the substrate for multiple glycosyltransferases that consecutively transfer sugar moieties from sugar nucleotides to build variable-length and -branched oligosaccharide chains of *N*-acetylneuraminic acid (NANA or sialic acid), galactose (Gal), GlcNAc, and fucose (Stanley *et al*., 2009; Roth, 2002; Trombetta, 2003).

1.2.2. *O*- linked glycans

While *N*-glycosylation is the most common glycosidic linkage on proteins, *O*-glycoproteins also play a key role in cell biology and are essential in the biosynthesis of proteins such as mucins. Mucins are a family of high-molecular weight proteins that are heavily *O*-glycosylated and form mucus secretions on cell surfaces and in body fluids. *O*-glycosylation is the attachment of a sugar molecule to an oxygen atom of serine and threonine residues of glycoproteins. Even though, by definition, the *O*-glycans are referred to as mucin type glycans, they are found on many secreted and membrane bound glycoproteins in higher eukaryotes. Differently to the *N*-glycosylation, *O*-glycosylation does not begin with a dolichol precursor (Van den Steen *et al*., 1998). *O*-glycosylation occurs post-translationally in the Golgi apparatus and involves much simpler oligosaccharide structures than *N*-glycosylation and is vastly different in fungi than in mammalian cells (Gemmill and Trimble, 1999).

In *O*-glycosylation, a single GalNAc residue is transferred from UDP-GalNAc to the β-OH group of serine (Ser) or threonine (Thr) of a polypeptide chain and is carried out by *N*-acetylgalactosamine (GalNAc) transferase. Following the first sugar, a number of different sugars are consecutively added to the growing glycan chain. Just like *N*-glycosylation, sugar nucleotides are the monosaccharide donors for *O*-glycosylation. (Spiro, 2002; Van den Steen *et al*., 1998; Lamarre-Vincent and Hsieh-Wilson, 2003). The GalNAc may extend with sugars including GalNAc, galactose, sialic acid, or fucose, but not mannose, glucose, or xylose residues, generating eight distinct core structures with four common core structures designated 1 through 4 (Fig 1-3) and additional core structures designated 5 through 8. It is believed that the addition of the monosaccharide residues to GalNAc take place in the Golgi complex. However, in contrast to mammalian *O*-glycosylation, mostly oligomannoses are found in yeast and the first step in the biosynthesis occurs in the endoplasmic reticulum (Van den Steen *et al*., 1998; Brockhausen *et al*., 2009).

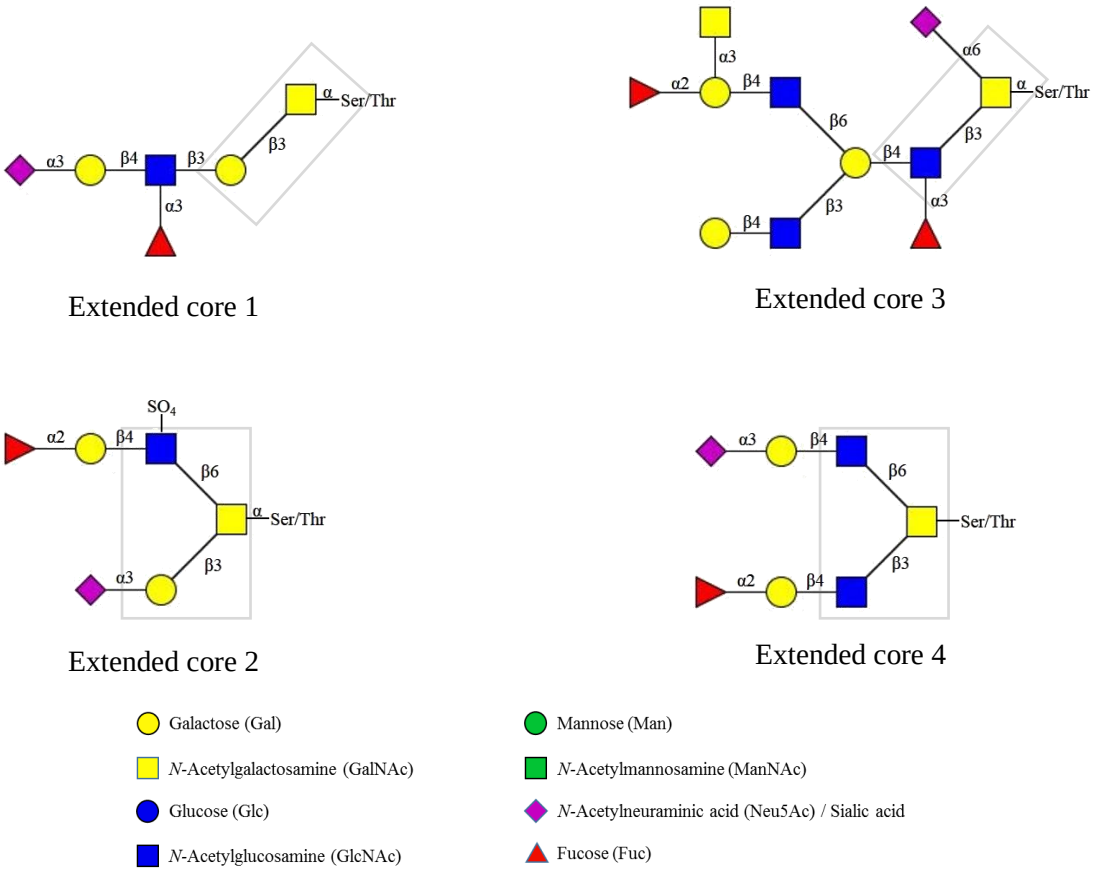


Fig 1-3: Complex *O*-glycan structures with core structures 1-4. The core structures (in boxes) can be extended, branched, and terminated by fucose, α2-3 sialic acid linkage, or blood group antigenic determinants. Core structures 1 and 3 may also carry α2-6 sialic acid linked to the core GalNAc (adapted from Brockhausen *et al*. (2009)). Glycan structures are generated using GlycoWorkbench.

1.3. Functions of glycans

The biological roles of glycans can be classified into two broad categories (Varki and Lowe, 2009):

i. structural and modulatory properties;
ii. specific recognition of glycans by other molecules such as glycan-binding proteins (GBPs).

Structural variations in the glycans at the cell surface produce numerous biomarkers, some of which are associated with differentiation, cell activation, and disease (Hakomori, 2002). Cell

surface carbohydrates are involved in a variety of interactions between a cell and its extracellular environment as they are located on the outermost layer of the cell, where *N*-glycans play a vital role in physiology and pathology (Montefiori *et al*., 1988; Wang *et al*., 2001). Naturally occurring oligosaccharides bound to asparagine (*N*- linked) or serine or threonine (*O*- linked) residues of glycoproteins play significant roles in protein folding, protein stability and enzyme activity, and also in physiological processes such as cell growth and development, cell–cell communication, cancer-cell metastasis, anticoagulation, immune response and host–pathogen interactions. The structural diversity of these glycans is regulated by 200 glycosyltransferases (Moremen *et al*., 2012).

1.3.1. Functions of *N*-glycans

N-glycans of signalling molecules such as receptors and adhesion molecules are recognised by other cells, invading bacteria, viruses, antibodies, hormones and toxins that bind to carbohydrate receptors on the cell surface. Moreover, most receptors on the cell surface including the epithelial growth factor receptor (EGFR), transforming growth factor β receptor (TGFβR) and integrins are *N*-glycosylated. In humans, the ABO blood group antigens are formed by the tissue-specific glycosylation of poly-*N*-acetyllactosamine chains and their β1–6-branched variants. Protein folding and stability are affected by oligosaccharide modifications, and have the ability to interfere with carbohydrate–protein, carbohydrate–carbohydrate, and glycoprotein–glycoprotein interactions. Subsequently, many physiological and pathological events such as cell growth, migration, differentiation, host–pathogen interactions, tumour invasion, cell trafficking and transmembrane signalling are regulated as a result. Glycans are also essential in the formation of proteoglycan core proteins that are used to make extracellular matrix components (Esko *et al*., 2009).

Antibodies are often glycosylated. Studies show that glycosylation on antibodies improves the thermal as well as overall stability (He *et al*., 2010; Zheng *et al*., 2011). Further, glycosylation largely reduces aggregation of antibodies since the carbohydrates block hydrophobic sites on the surface (Kayser *et al*., 2011). Studies on glycoconjugates present in human milk give strong evidence on the significance of oligosaccharide structures in imparting disease resistance to infants (Peterson *et al*., 2013). Glycosylation also plays a remarkable role in complement-dependent cytotoxicity (CDC) and antibody-dependent cell-mediated cytotoxicity (ADCC) by modulating the binding to the Fcγ receptor (Ravetch and Bolland, 2001; Jefferis, 2009).

Studies indicate that the development and activation of the mammalian immune system can be altered by glycans of non-vertebrate organisms (Cobb and Kasper, 2005). Mammalian glycans are found to be species-specific and are well conserved across different species. This uniqueness in glycan structures has aided in the emergence of species-specific infectious pathogens (Gagneux and Varki, 1999). Most microbial parasites such as protozoans and helminthes, synthesise abnormal glycans and glycan binding proteins, which aid in parasitism and are antigenic. For example, Tyvelose found in complex *N*-glycans of larvae from *Trichinella spiralis* is a critical antigenic determinant recognised by antibodies in infected animals. Strong immunity to these glycans provides protective resistance resulting in removal of the invading larvae from the intestine (Cummings and Turco, 2009).

Further, studies by Cobb and Kasper (2005) describe that not only antigenic proteins, but also sugars play a critical role in immune recognition. Their findings on conditions like rheumatoid arthritis, a systemic autoimmune disease, establish an exemplary paradigm on the dependency of sugar and differences in protein glycosylation on immune recognition. Studies illustrate that the normal tissues primarily contain glycosylated Type II collagen (CII), whereas arthritic tissues contain a mixture of glycosylated and non-glycosylated CII proving

that a lack of glycosylation is the key component of arthritis progression. Accordingly, immunodominant glycopeptide remnant epitopes as well as glycosylation changes on self-proteins can elicit autoimmunity (Cobb and Kasper, 2005).

1.3.2. Functions of *O*-glycans

The functions of *O*-glycans are as diverse as their structures. Like the *N*-glycans, they also have a significant role contributing to the physiology and functioning of an organism. *O*-glycans have been identified to play a key role in protecting epithelial cell surfaces and underlying proteins, maintaining protein structure, controlling antigenicity, mediating sperm–egg binding, binding to microbes and functioning of the immune system (Brockhausen, 1999). *O*-linked glycosylation may also play a major role in determining the secondary, tertiary, and quaternary structure of the fully folded protein. Studies have revealed that the *O*-glycans at specific sites are required to prevent proteolytic cleavage of the extracellular domain of low density lipoproteins (Gabius, 2011). Additionally, it is understood that the *O*-linked glycan at Thr 27 prevents the proteolysis of the mammalian copper transporter protein hCTR1 (Maryon *et al*., 2007).

Abnormal glycosylation patterns are often found associated with disease manifestation, cancer invasion and metastasis (Zhao *et al*., 2008; Dennis *et al*., 1999; Stanley P, 2009). The proximity of an *O*-linked sugar is closer to the peptide backbone compared to an *N*-linked glycan. This closeness of sugar can lead to alterations to α-helicity of a peptide, leading to perturbation of the protein secondary structure. Otvos *et al*. (1995) observed this remarkable property of glycans in their studies with a synthetic peptide homologous to a part of a rabies virus glycoprotein.

1.4. Sialic acids

Sialic acids are naturally occuring acidic monosaccharides that occupy the end of sugar chains present on cell surfaces and soluble proteins. Sialic acids are commonly present as components of oligosaccharide chains of glycoproteins, mucins, and glycolipids. The sialic acids rarely occur free in nature and are functionally important. They occupy the terminal, non-reducing positions of oligosaccharide chains of complex carbohydrates and are linked mainly to galactose, *N*-acetylgalactosamine, and other sialic acid moieties, where they are highly exposed (Wang and Brand-Miller, 2003). They are ubiquitously expressed in all cell surfaces of vertebrates and higher invertebrates and certain bacteria that interact with vertebrates (Varki and Gagneux, 2012).

Free sialic acids can be used for glycan synthesis only after converting them into the nucleotide donor cytidine monophosphate-sialic acid (CMP-sialic acid), a reaction catalysed by CMP-sialic acid synthatases using cytidine triphosphate (CTP) as a donor. This reaction occurs in the nucleus of all eukaryotes known to date. Sialic acids comprise a family of more than 50 naturally occurring carbohydrates. The two common primary sialic acids are 5-acetamido-2-keto-3,5-dideoxy-D-glycero-D-galactonononic acid (*N*-acetylneuraminic acid, Neu5Ac) and 2-keto-3-deoxy-D-glycero-D-galactonononic acid (2-keto-3-deoxynononic acid, Kdn (Varki and Schauer, 2009; Bull *et al*., 2014). The carbon C5 in neuraminic acid is modified by an *N*-acetyl group to form *N*-acetylneuraminic acid (Neu5Ac) which can be further hydroxylated to form *N*-glycolylneuraminic acid (Neu5Gc) (Varki, 2009). Other kinds of modifications include acetylation, methylation and sulphation of its hydroxyl groups.

Sialic acids contribute to a rich diversity of glycan structures that mediate cell surface biology. Occupying the terminal positions on *N*- and *O*- glycan chains of glycoproteins and glycolipids, sialic acids are recognised as ligands by glycan binding proteins. In mammals,

sialic acids are usually found at the non-reducing terminal position of glycoconjugate sugar chains that are α2-3- or α2-6-linked to a β-D-galactopyranosyl (Gal) residue, α2-6-linked to a β-D-*N*-acetylgalactosaminyl (GalNAc) residue or a β-D-Nacetylglucosaminyl (GlcNAc) residue. Sialic acids are also found linked to other sialic acid units, such as in gangliosides and polysialic acids (PSA), with an α2-8 linkage. PSAs are linear α2-8-homopolymers observed on several glycoproteins (Harduin-Lepers *et al*., 1995; Harduin-Lepers *et al*., 2001). This diversity in α- linkages facilitate for their potential in biologically significant diversity.

In humans, sialic acids are found in large abundance in the brain where they are an integral part of ganglioside structure in synaptogenesis and neural transmission. Recent studies by Salcedo *et al*. (2013) have showcased the inhibitory effect of sialic acids to the adhesion of diarrhoeal microbes such as *Campylobacter jejuni*, *Helicobacter pylori*, *Listeria monocytogenes*, *Salmonella entericasero-vartyphi*, and *Shigella sonnei* to Caco-2 cells (human epithelial colorectal adenocarcinoma cells**)**. Human milk contains a high concentration of sialic acid attached to the terminal end of free oligosaccharides (Wang and Brand-Miller, 2003). In 1994, studies by Vacca-Smith *et al*. (1994) and later Ward *et al*. (1997) elucidated the role of sialic acid residues in human milk glycoproteins, especially in κ casein, and their interactions with gut microbes in preventing infections, thereby validating the findings of Fiat *et al*. (1980).

Two families of mammalian glycan binding proteins utilise sialic acid containing glycans as cell surface ligands: the selectin sub-group of the C-type lectins that mediate leukocyte trafficking, and the **s**ialic acid binding Ig-like lectins, Siglecs, that are differentially expressed on white blood cells of the immune system, and modulate cell signaling receptors involved in innate and adaptive immune responses (Karlsson, 1998; Taylor and Drickamer, 2007; Crocker *et al*., 2007).

Polysialic acid (PSA), found on the neural cell adhesion molecule (NCAM) plays an integral part in maintaining developmental plasticity by interfering with both homotypic and heterotypic interactions involving neuronal cells (Varki and Schauer, 2009). NCAM is a glycoprotein containing, five Immunoglobulin (Ig) like domains and six *N*-glycosylation sites. PSA is a polyanionic molecule with high water binding capacity and increases the intercellular space creating conditions that are helpful for cellular plasticity. PSA on microbial polysaccharides may also contribute to conserving water (Vimr *et al*., 2004; Muhlenhoff *et al*., 2009).

1.4.1. Altered sialylation in disease manifestation and inflammation

Pathogens including viruses (influenza virus, polyoma virus), bacteria, and a variety of bacterial toxins (e.g cholera toxin) utilise sialylated glycans as cell surface receptors for attachment to the host cell. This can be illustrated by the mode of infection by Herpes Simplex Virus (HSV) that requires sialic acids for efficient infection of the host cell via surface proteoglycans. Studies suggest that α2-6-linked sialic acids are required for HSV infection on Hep-2 cells (a human epithelial cancer cell line) (Teuton and Brandt, 2007). Wu *et al*. (2006) also confirmed in their studies, that the *N*-linked α2-3 and α2-6 sialic acids facilitated efficient binding and transduction by adeno-associated virus types 1 and 6.

Influenza virus, a segmented single-stranded RNA virus, infects mammals and birds and is common in humans. The trimeric viral hemagglutinin (HA) protein adhere to host cells by binding to sialic acid, usually Neu5Ac. Influenza viruses engage with α2-3-linked and α2-6-linked sialic acid attached to a sub-terminal galactose of the glycan receptor. While the human influenza viruses primarily bind to α2-6-linked sialic acid, avian influenza viruses bind to α2-3-linked sialic acid (Rogers *et al*., 1983a; Rogers *et al*., 1983b; Stencel-Baerenwald *et al*., 2014). One of the key factors governing the adaptation of a virus to human host is the glycan

receptor-binding specificity of its HA. The human-adapted subtypes such as H1N1, H2N2, and H3N2 prefer α2-6 sialylated glycan receptors (or human receptors), and has a high relative binding affinity to human receptors over avian receptors (Tharakaraman *et al*., 2013). The human H7N9 virus has a significantly higher affinity towards α2-3 and α2-6-linked sialic acids, compared to the avian counterpart, aiding HA to bind with the human receptor protein (Xiong *et al*., 2013). H10N8 virus, latest in the line causing human influenza has a strong affinity towards sialic acid and galactose components of the receptor protein (Vachieri *et al*., 2014).

Pathogenic bacteria have evolved to use sialic acids beneficially in either providing resistance against the host's immune response by coating themselves in sialic acids, or utilising them as a nutrient for survival (Severi *et al*., 2007). Although, glycans constitute the majority of cell surface components on animal and plant cells and viral and bacterial surface, the polysaccharide capsule that covers the surface is the main factor that aids in disease manifestation by bacteria. For instance, the bacterial immune evasion by capsule-based host mimicry is well demonstrated by the homopolymeric sialic acid capsules of *Neisseria meningitidis* (meningococcus). *N. meningitidis* is an important cause of sepsis and meningitis. The group B meningococcal capsule resembles a motif present on neural cell adhesion molecules (NCAMs) found in human neural tissues and is composed of an α2-8-linked sialic acid polymer. In contrast, the group C meningococcal capsule is unique for the bacterium and is composed of an α2-9-linked sialic acid polymer (Nizet and Esko, 2009).

The microbial molecules responsible for disease manifestations are known as virulence factors (Nizet and Esko, 2009). Studies have shown that the sialylated lipopolysaccharide (LPS) on *C. jejuni*, *N. meningitidis*, and the capsule on *Streptococcus agalactiae* can interact with sialic acid-specific lectins of the Siglec family, which are expressed on the surface of cell types of the immune system (Avril *et al*., 2006; Carlin *et al*., 2007; Jones *et al*., 2003). The

high density of sialic acids on polysaccharides of pathogenic bacteria can alter their biophysical properties. For example, by altering the glycosylation of an immuno-dominant antigen, such as flagellin, certain pathogenic bacteria like *Campylobacter* spp., are able to alter their surface antigenicity, thereby avoiding or delaying the host clearance mechanism (Vimr *et al*., 2004). Most of the microbial pathogens have evolved a certain degree of molecular mimicry of host sialic acids to hijack host Siglecs and inhibit antibody formation (Varki and Gagneux, 2012).

In addition to the role of surface expressed sialic acids, microbial sialidases also have a significant role in pathogenesis and survival. Bacterial sialidases remove the negatively charged sugars thereby aiding them to break down epithelial mucins and surface glycoconjugates to facilitate their entry into the cells. Bacteria such as *Haemophilus influenza* utilise these free SA as a source of energy (Varki and Gagneux, 2012).

Many studies have reported that increased sialic acid content was found associated with a number of autoimmune diseases. Janega *et al*. (2002) concluded that there is a significant increase in α2-3 linked sialic acid in autoimmune diseases of the thyroid gland. In a recent study, Chrostek *et al*. (2014) referred to the significance of utilising the increased levels of sialic acid as a marker of inflammation activity in rheumatoid arthritis, systemic sclerosis and systemic lupus erythematosus.

1.4.2. Abnormal sialylation in cancer and metastasis

Aberrant sialylation is associated with tumour growth and progression. Tumour cells are often seen to feature an increased expression of sialic acids on membrane glycoproteins and glycolipids. Sialic acids are also found secreted into the tumour micro environment (Bull *et al*., 2014). There are three mechanisms by which aberrant sialylation occur in cancer cells. Firstly, the overexpression of sialyltransferases lead to increased sialylation of glycans and

expression of tumour associated carbohydrate antigens such as sLex/a, STn, GD2, GD3, or PSA, facilitating tumor cell motility (Bull *et al*., 2014; Hauselmann and Borsig, 2014).

Secondly, an increased substrate availability or over-expression of sialic acid biosynthesis genes can enhance their synthesis in cancer cells. Almaraz *et al*. (2012), in their studies proved that the bulk metabolic flux through the sialic acid pathway leads to the increase of certain sialoglycoproteins and not just the transcription of genes encoding the biosynthetic enzymes or the availability of *N*-glycan sequons alone. Thirdly, improper removal of sialic acids due to a decrease in the production of innate sialidases can lead to an increase in terminal sialic acids on proteins in the tumour cells (Miyagi *et al*., 2012). Recent studies indicate that the overexpression of sialyltransferases, enzymes involved in sialylation, especially ST6Gal1, and the resulting aberrant sialylation contribute to resistance in cancer therapy (Schultz *et al*., 2013; Swindall *et al*., 2013).

A significant feature of cancer cells is their ability to resist apoptosis or programmed cell death. Swindall and Bellis (2011) observed that the hypersialylation of Fas receptors (also known as apoptosis antigen 1) hindered the Fas mediated apoptosis. Moreover, α2-6-sialylation impairs the internalisation of the Fas receptor preventing the formation of death-inducing signaling complex (DISC) (Lee *et al*., 2006). Studies have revealed that galectins regulate cell-cell and cell-matrix adhesion (Liu and Rabinovich, 2005). Hypersialylation, particularly α2-6 sialylation, is known to negatively affect galectin-3 recognition and binding to its β-galactoside ligands, thereby leading to detachment of cells from the primary tumour promoting metastasis (de Oliveira *et al*., 2011). It has also been reported that hypersialylation mediates resistance to anoikis, cell-detachment-induced apoptosis (Guadamillas *et al*., 2011). Although not conclusively proven, some studies have claimed that long term ingestion of mammalian red meat rich in Neu5Gc (a non-human sialic acid), promotes inflammation and cancer progression in humans (Samraj *et al*., 2014; Samraj *et al*., 2015).

1.4.3. Role of sialic acids in host immunity and pathogen recognition

Owing to their significant diversity in structures, glycosidic linkage, and underlying glycan chains, as well their exposed location, sialic acids have numerous roles in the many facets of immunity. Sialic acids are ubiquitously expressed at the surface of all cell types including those of the immune system and have major biophysical effects. (Varki and Gagneux, 2012). Pilatte *et al*. (1993) observed that removing the surface sialic acids using sialidases changed the behaviour of immune cells.

Sialic acids modify structures that form the key fundamental determinants for a number of receptors at the cell surface with known involvement in cellular adhesiveness and cell trafficking, such as the Selectins and the Siglec families of carbohydrate recognising receptors. They play a key role in immunity by modifying structures involved in all dendritic cell (DC) functioning, such as, antigen uptake, DC migration, pathogen and tumour cell recognition, cell recruitment and capacity to prime T cell responses. It has been studied that the sialic acid content changes during DC differentiation and activation and has important implications in DC functions (Crespo *et al*., 2013). Apart from NCAM, polysialic acids are also discovered on dendritic cells and in some stages of T cell development attached to specific proteins such as neurophilins. Mammalian sialic acid mimicry by pathogens may raise a concern of stealth invasion into the host cells. However, the immune system is evolved to identify and evade such incursions. The Siglec-1 is highly specific for bacterial α2-3 or α2-8 linked Neu5Ac. None of the microorganisms is known to synthesise Neu5Gc and α2-6 linked sialic acids that are common to mammals. Furthermore, host sialic acids on proteins such as mucins can act as decoys preventing the pathogens from reaching a target cells (Varki and Gagneux, 2012).

A classic study by Rosen *et al.* (1985, 1989) has demonstrated that the pre-treatment of lymph nodes with sialidase eliminated the interaction of lymphocytes with the high-endothelial venules (HEV), via which recirculating lymphocytes attach as they migrate from the blood into the parenchyma of the lymphoid organs. This suggested that the sialic acid on endothelial cells serves as a recognition factor for lymphocyte attachment.

1.5. Human sialyltransferases

Sialyltransferases (STs) are a family of enzymes that catalyses the addition of terminal sialic acids to oligosaccharide sequences. STs are type II transmembrane glycoproteins belonging to the CAZy 29 glycosyltransferase family, members of which catalyse the transfer of sialic acid from Cytidine 5-Prime-Monophosphate *N*-Acetylneuraminic Acid (CMP-Neu5Ac or CMP-sialic acid), a nucleotide monophosphosugar, to the glycan moiety with a galactose, *N*--acetylgalactosamine (GalNAc) or sialic acid as the acceptor. The human genome encodes at least 20 different sialyltransferases (Harduin-Lepers *et al*., 2005; Kuhn *et al*., 2013). STs are classified based on the position of attachment of donor sialic acid to the acceptor, being either α2-3 (ST3), α2-6 (ST6) or α2-8 (ST8) (Rao *et al*., 2009).

Differential expression of sialyltransferases is biologically important during the development of mammalian physiology. This has been established by knockout studies in mice models in which inactivation of the α2-6 ST (ST6Gal1) gene led to a severe decrease in immune function, while the inactivation of α2-3 ST (ST3Gal1) led to deficiency of mature cytotoxic T lymphocytes (Hennet *et al*., 1998; Martin *et al*., 2002). As discussed before, sialic acids, and hence sialyltransferases, play a significant role in cell interaction, immunity and other physiological conditions such as inflammation, disease and cancer. Sialyltransferase genes are differentially expressed in a tissue-, cell type-, and stage-specific manner to regulate the sialylation pattern of cells (Harduin-Lepers *et al*., 2001). Normal regulation of

sialyltransferase expression is crucial for the regular functioning of cells. Low oxygen levels and high hormone levels have been found to upregulate expression of sialyltransferases (Bull *et al*., 2014).

1.5.1. Human α-3 sialyltransferases

The α-3 sialyltransferases catalyse the transfer of sialic acid from CMP-sialic acid to galactose-containing substrates. Six different α2-3 sialyltransferase cDNAs have been identified and cloned from human cells (GlycoGene DataBase, http://jcggdb.jp/rcmg/ggdb/). Two Galβ1-3GalNAc α2-3 sialyltransferases (ST3Gal1 and ST3Gal2) mediate the transfer of sialic acid residues to a Gal residue of the terminal Galβ1-3GalNAc disaccharide found on glycolipids or glycoproteins. Human α2-3 sialyltransferase ST3Gal3 was shown to be highly expressed in skeletal muscle, whereas human ST3Gal4 cDNA (Galβ1-4GlcNAc α2-3 sialyltransferase) was abundantly expressed in placenta (Kitagawa and Paulson, 1994a). Despite the ubiquitous distribution of hST3Gal4, otherwise known as GM3 synthase, in human tissues, the ST3Gal5 gene is found to be expressed in a tissue-specific manner with predominant expression in brain, skeletal muscle and testes while expression at very low levels is found in liver (Ishii *et al*., 1998). In 1999, Okajima *et al*. identified the human ST3Gal6 gene and the cDNA was isolated from a human melanoma line SK-MEL-37. It is primarily expressed in tissues found in placenta, liver, heart and skeletal muscles (Okajima *et al*., 1999; Harduin-Lepers *et al*., 2001). There are six human α-3 sialyltransferases identified to date. Table 1-2 shows the human α-3 sialyltransferases and the source of cDNA for the cloning of the gene.

Table 1-2: The human α-3 sialyltransferases that have been identified to date (adapted from Harduin-Lepers *et al*. (2001)).

STs	Structures formed	Initial source of cDNA	Reference
ST3GalI	Neu5Acα2-3Galβ1-3GalNAc-	Placenta	(Kitagawa and Paulson, 1994b)
ST3GalII	Neu5Acα2-3Galβ1-3GalNAc	Liver	(Kim *et al*., 1996)
ST3GalIII	Neu5Acα2-3Galβ1-3/4GlcNAcβ-	Placenta	(Kitagawa and Paulson, 1993)
ST3GalIV	Neu5Acα2-3Galβ1-4GlcNAc	Placenta	(Kitagawa and Paulson, 1994a)
	Neu5Acα2-3Galβ1-3GalNAc	B lymphoma	(Sasaki *et al*., 1993)
ST3GalV	Neu5Acα2-3Galß1-4Glc-Cer	HL-60	(Ishii *et al*., 1998)
ST3GalVI	Neu5Acα2-3Galβ1-4GlcNAcβ	SK-MEL-37	(Okajima *et al*., 1999)

1.5.2. Human α-6 sialyltransferases

The α-6 sialyltransferases mediate two types of sialic acid linkages. The first one is the α2-6-linkage to a terminal Gal residue of type 2 (Galβ1-4GlcNAc) disaccharide of an *N*- or *O*-linked oligosaccharide. The second is the α2-6- linkage to GalNAc. The cDNA encoding α-6 sialyltransferase was first isolated and cloned from humans in 1990 (Grundmann *et al*., 1990). Eight members of α-6 sialyltransferases have been identified to date. Two of them, namely the ST6Gal1 and ST6Gal2, promote α-6 sialic acid linkage to a galactose residue. The six remaining α-6 sialyltransferases (ST6GalNAc1 to 6), link sialic acid to the GalNAc residue (GlycoGene DataBase http://jcggdb.jp/rcmg/ggdb/).

While ST6Gal1 is ubiquitously expressed in human tissues, ST6Gal2 shows a restricted tissue specific expression mainly confined to embryonic and adult brain (Lehoux *et al*., 2010). In addition, a recent study by Lee *et al*. (2014) has reported that the human breast cancer cell line SKBR3 over-expressed the ST6Gal1. Human ST6Gal2 has been identified to have a 48.9% amino acid sequence identity to hST6Gal1, where the former has a strong substrate specificity for the free disaccharide Galβ1-4GlcNAc structure and the latter to the Galβ1-4GlcNAc-R disaccharide sequence linked to a protein (Takashima *et al*., 2002).

Human ST6GalNAc1 is involved in the synthesis of cancer-associated sialyl-Tn (sTn) antigens, expression of which is related to the onset of cancer (Ikehara *et al*., 1999). In adult tissues, ST6GalNAc2 is found abundant in skeletal muscles and to a lesser extent in the heart, and has been found to be expressed in various human cancer cell lines. While heart, skeletal muscle, kidney, liver and small intestine show a 2.0 kb gene transcript of ST6GalNAc2, placenta, lung and leukocytes are observed to show a smaller transcript size (Samyn-Petit *et al*., 2000). Tsuchida *et al*. (2005) observed that hST6GalNAc3 has a restricted expression pattern confined to kidney and brain, but has a high specificity towards GM1b (an α2-3 sialylated acceptor substrate). The hST6GalNAc4 on the other hand is constitutively expressed in various adult human cells and also at lower levels in brain, heart and skeletal muscles. Investigation of various cancer cells by reverse transcription PCR revealed its expression in most of the cancer cell lines except T47D and pro-monocyte THP cells (Harduin-Lepers *et al*., 2000). ST6GalNAc5 and ST6GalNAc6 have been found to be involved in the synthesis of α-series gangliosides such as GD1α, GT1aα and GQ1bα. In addition, ST6GalNAc6 has been found to be responsible for the synthesis of disialyl Lewis a (Le(a)) structure in colon cancer cell lines. ST6GalNAc5 is specifically expressed in the brain, (Bos *et al*., 2009) whereas ST6GalNAc6 is expressed in a wide range of tissues such as kidney and proximal tubule epithelial cells (Tsuchida *et al*., 2003; Senda *et al*., 2007). Table

1-3 lists the various human α-6 sialyltransferases and the source of cDNA for the cloning of the gene.

Table 1-3: The human α-6 sialyltransferases that have been identified to date (adapted from Harduin-Lepers *et al*. (2001)).

STs	Structures formed	Initial source of cDNA	Reference
ST6Gal1	Neu5Acα2-6Galβ1-4GlcNAcβ-R	Human placenta	(Grundmann *et al*., 1990)
ST6Gal2	Neu5Acα2-6Galβ1-4GlcNAcβ-	Human colon	(Takashima *et al*., 2002)
ST6GalNAc1	Neu5Acα2-6 (Neu5Acα2-3)(Galβ1-3)GalNAc-Ser	Human non-tumorous pyloric mucosa	(Ikehara *et al*., 1999)
ST6GalNAc2	Neu5Acα2-6 (Neu5Acα2-3)Galβ1-3GalNAc-Ser	MDA-MB-231	(Samyn-Petit *et al*., 2000)
ST6GalNAc3	Neu5Acα2-6Neu5Acα2-3Galβ1-3GalNAc-R	Human brain cDNA library	(Tsuchida *et al*., 2005)
ST6GalNAc4	Neu5Acα2-6 Neu5Acα2-3Galβ1-3GalNAc-R	HepG2 cells	(Harduin-Lepers *et al*., 2000)
ST6GalNAc5	Neu5Acα2-6GD1α	Human gene not yet cloned	(Bos *et al*., 2009)
ST6GalNAc6	GD1α , (GT1α)	Human gene not yet cloned	(Senda *et al*., 2007)

1.5.3. Human α-8 sialyltransferases

The α-*N*-acetylneuraminate α2-8-sialyltransferases otherwise known as α-8 sialyltransferases include a set of six enzymes identified to date (GlycoGene DataBase). As with the other sialyltransferases, human α-8 sialyltransferases have a substantial role in modulating cellular functions. Human sialyl α2-3Galβ1-4Glcβ1-*O*-Cer α-2-8-sialyltransferase (hST8Sia1) catalyses the synthesis of disialoganglioside GD3, hence known as GD3 synthase, that

functions in an intracellular signalling cascade leading to leukocyte apoptosis (Harduin-Lepers *et al*., 2001; Simon *et al*., 2002). ST8SiaI is the only enzyme known to catalyse the transfer of sialic acid onto GM3 to synthesise GD3. ST8Sia I and GD3 are expressed in foetal tissues at an early developmental stage and in adult human brain tissues (Steenackers *et al*., 2012). Cazet *et al*. (2009) reported that GD3 synthase overexpression enhanced proliferation and migration of MDA-MB-231 breast cancer cells.

Human ST8Sia2, ST8Sia3 and ST8Sia4 are called polysialyltransferases for their ability to transfer multiple number of sialic acid residues (as many as 55) on to a glycan chain. Neural cell adhesion molecule (NCAM) in mammals is found to be polysialylated, a reaction catalysed by ST8Sia2 and ST8Sia4. Knockout studies in mice have revealed the significance of polysialylation on NCAM and its critical role in neural cell migration, axonal growth, neural plasticity and development. The hST8Sia3 on the other hand, forms polysialic acid on the enzyme itself but not on NCAM (Angata *et al*., 2000; Angata *et al*., 2007). The hST8Sia2 is expressed in foetal brain and kidneys but also in adult heart, thymus and brain. The hST8Sia3 is expressed in foetal and adult brain and liver cells and the hST8Sia4 in adult spleen, thymus, heart, small intestine and leukocytes (Harduin-Lepers *et al*., 2001).

The gene encoding human sialyltransferase hST8Sia5 was first cloned by Kim *et al*. in 1997 and has been found to have activity towards gangliosides GM1b, GD1a, GT1b, and GD3. Two sizes of hST8Sia5 transcripts, 11 kb and 2.5 kb in length are expressed in human foetal and adult brain, whereas a transcript of 2.5 kb was detected in adult heart and skeletal muscle (Kim *et al*., 1997). Human ST8Sia6 gene was identified on human chromosome 10 and is responsible for the synthesis of a disialic acid (diSia) motif on *O*-glycoproteins. Teintenier-Lelievre *et al*. (2005) cloned the hST8SiaVI first and reported high level of expression in cancer cell lines MSF-7 and Dami megakaryocyte cell line, whereas human tissues were

expressing the enzyme in lower levels. Table 1-4 lists different human α-8 sialyltransferases identified to date and the source of cDNA for the cloning of the gene.

Table 1-4: The human α-8 sialyltransferases that have been identified to date (adapted from Harduin-Lepers *et al*. (2001)).

STs	Structures formed	Initial source of cDNA	Reference
ST8Sia1	Neu5Acα2-8Neu5Acα2-3Galβ1-4Glc-Cer	Human melanoma	(Nara *et al*., 1994; Sasaki *et al*., 1994; Haraguchi *et al*., 1994)
ST8Sia2	Neu5Acα2-8Neu5Acα2-3Galβ1-4GlcNAc	NCI-H69/F3 cell line	(Scheidegger *et al*., 1995)
ST8Sia3	Neu5Acα2-8Neu5Acα2-3Galβ1-4GlcNAc	Human brain	(Lee *et al*., 1998)
ST8Sia4	Neu5Acα2-8(Neu5Acα2-8)$_n$Neu5Acα2-3Galβ1-R	Embryonic brain	(Nakayama *et al*., 1995)
ST8Sia5	GD1c, GT1a, GQ1b, GT3	Human brain	(Kim *et al*., 1997)
ST8Sia6	Neu5Acα2-8Neu5Acα2-3Galβ1-4GlcNAc	MCF-7 cell line	(Teintenier-Lelievre *et al*., 2005)

1.6. Structure of sialyltransferases

Sialyltransferases are Golgi type II transmembrane glycosyltransferases that predominantly reside in the trans-Golgi compartment and have a similar structure. They have a short N-terminal cytoplasmic tail, transmembrane domain signal anchor, and a stem region composed of 20 to 200 amino acids followed by a large C-terminal catalytic domain that resides in the lumen (Fig 1-4) (Harduin-Lepers *et al*., 2005). Human sialyltransferases share a limited 15 to 57% sequence identity, but share four peptide conserved motifs called the sialylmotifs which are L (large), S (small), motif III, and motif VS (very small) (Fig 1-5) (Harduin-Lepers *et al*.,

2001; Harduin-Lepers *et al*., 2005). A disulphide bond stabilising the L- and the S-motifs is also conserved across sialyltransferases (Datta *et al*., 2001).

A conserved histidine residue located in the VS-motif has been identified as the catalytic base from the X-ray crystal structures of porcine ST3Gal1 which has the GT-A fold with a single Rossmann domain (a protein structural motif found in proteins that bind to nucleotides) (Rao *et al*., 2009). In addition, the primary amino acid sequences of the sialyltransferases consist of several consensus sites for phosphorylation and *N*-glycosylation that may influence the catalytic activity and intracellular traffic of sialyltransferases (Harduin-Lepers *et al*., 2001).

1.6.1. Structure of the human ST6Gal1

The human β-galactoside α2-6-sialyltransferase1 (ST6Gal1) establishes the final glycosylation pattern of glycoproteins by transferring a sialic acid residue to a terminal galactose of *N*-glycans. The crystal structure of human ST6Gal1 was resolved by Kuhn *et al*. (2013) and was determined from a truncated version of the human ST expressed in the human embryonic kidney (HEK) cells. The truncated protein lacked 89 amino acids in the N terminus. The secondary structure revealed that the human ST6Gal1 adopts a GT-A variant two fold, seven-stranded central β-sheet flanked by α-helices (Fig 1-4). Comparison of hST6Gal1 with porcine ST3Gal1 revealed that CMP binds to the proteins in the same conformation but with a different chemical environment (Kuhn *et al*., 2013).

The human ST6Gal1 is composed of 406 amino acids, with four characteristic regions. The first nine amino acids form the N-terminal cytoplasmic tail protrudes into the cytoplasm. Following this is the transmembrane domain (TMD) that extends into a stem region. The TMD and stem region consists of signals that localise the protein in the Golgi membrane. The stem region leads to the C-terminal catalytic domain that faces the Golgi lumen (Kuhn *et al*., 2013) (Fig 1-5).

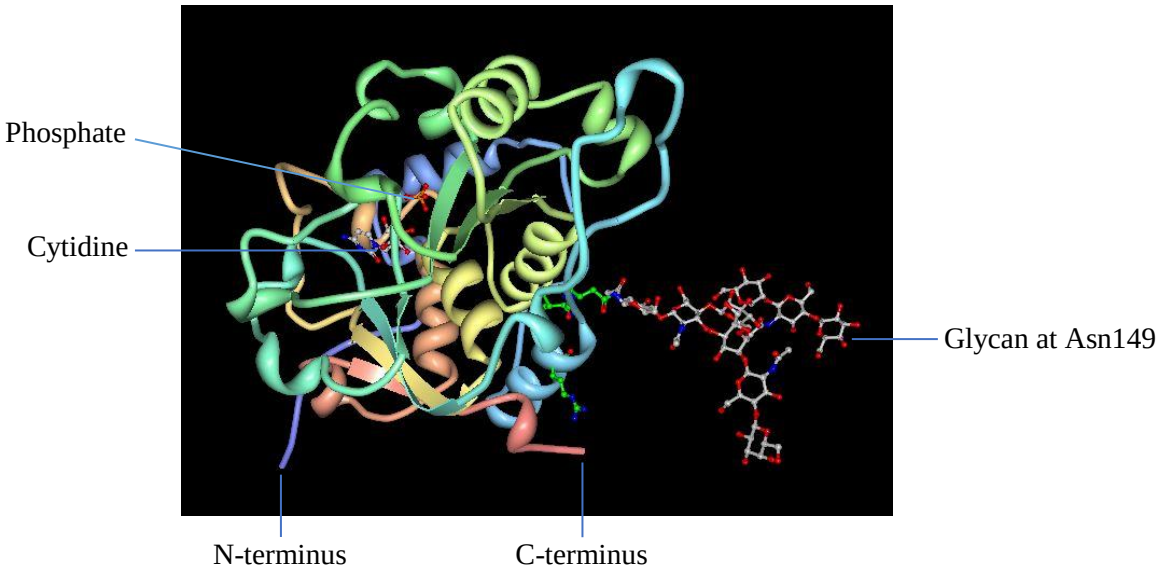


Fig 1-4: Structure of human ST6Gal1 complex with cytidine and phosphate revealing an *N*-glycosylation site at Asn149 (Moreland *et al*., 2005; Kuhn *et al*., 2013). The protein model is coloured based on a rainbow colouring scheme. The colours from violet to red (VIBGYOR) represent the orientation of the polypeptide chain from the N-terminus to the C-terminus.

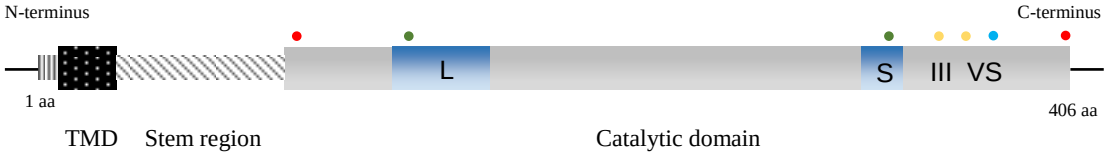


Fig 1-5: Schematic representation of the structure of human ST6Gal1. – Short N-terminal cytoplasmic residue comprised of 9 amino acids, – the transmembrane domain (TMD) contains Golgi membrane localisation signals, – the stem region contains Golgi localisation signals and extends into the Golgi lumen, and – the catalytic domain consisting of L, S, III and VS motifs. The blue dot represents the catalytic His370. Three disulphide bonds formed between Cys142-406, Cys184-335 and Cys353-364 are represented by red, green and yellow dots respectively.

The catalytic domain of human ST6Gal1 also consists of four conserved motifs. The L (large) motif extends from Trp181 to Gly224, the S (small) motif is composed of amino acid residues between Pro321 to Phe343, motif III comprises amino acids Tyr354 to Gln357 and the VS (very small) motif consisting of five amino acids from His370 to Glu375. Motifs L and III are attributed to participate in binding of the sugar donor CMP-Neu5Ac, motif S is involved in donor and acceptor substrate binding and motif VS take part in the catalytic reaction (Datta,

2009; Paulson and Rademacher, 2009; Rao *et al*., 2009; Audry *et al*., 2011; Kuhn *et al*., 2013). There are two *N*-glycosylation sites at Asn149 and Asn161 (Kuhn *et al*., 2013) and a phosphorylation site at Tyr369 within the human ST6Gal1 sequence (NCBI, 2015). One of salient features that distinguishes the ST6Gal1 structure from other ST structures determined to date is the presence of a large glycan at the Asn149 that mimics a substrate complex with the central mannose bifurcating into a 1,3-branch and a1,6-branch. The terminal galactose residue on the 1,3-branch binds to the active site of ST6Gal1 (Kuhn *et al*., 2013). Even though the basic structures remain the same, the human ST6Gal2 sub-member of the family shares only 48.9% sequence identity to ST6Gal1 and differ in their detailed acceptor specificities (Takashima *et al*., 2002). Interestingly, Kuhn *et al*. (2013) found that, the ST6Gal1 exhibits sialidase activity and could be its own substrate for sialidase activity at high concentrations.

1.7. Glyco-engineering

As the most common and complex posttranslational modification, glycosylation contributes to micro-heterogeneity, in a qualitative and quantitative sense. Recent reports on the significance of glycosylation in many diseases (Maguire and Breen, 1995; Wu *et al*., 2006) have paved the way for extensive studies into the glycosylation mechanisms as potential targets for therapeutic purposes. Therapeutic proteins have reformed the treatment of many diseases. New approaches have been taken to design drugs with enhanced *in vivo* activity and half-life to reduce injection frequency, increase convenience, and improve patient compliance. The demand for these recombinant human-like proteins is mainly governed by their medical need. The *in vivo* counterparts of most of the therapeutic proteins are glycoproteins with single carbohydrates or large carbohydrate moieties attached to the protein. It has been shown that glycosylation affects the bioactivity of a protein, physical features such as tertiary structure and stability (Elliott *et al*., 2003). Further, the momentous role played by the carbohydrate

components of proteins in solubility, primary functional activity, half-life and immunogenicity, expound the significance of glycans (Macdougall, 2002; Egrie *et al*., 2003; Son *et al*., 2011).

With more than one-third of the approved bio-therapeutics being glycoproteins, the manufacture of therapeutic proteins with a reproducible and consistent glycoform profile remains a considerable challenge to the biopharmaceutical industry (Walsh and Jefferis, 2006). Frequent administration, rapid clearance and low activity of therapeutic proteins are often a challenge in treatment. In a study by Elliott *et al*. (2003), re-engineering the glycan moieties on recombinant human erythropoietin (rHuEPO), leptin and Mpl ligand, by introducing a new sialic acid containing carbohydrate, was shown to have significantly increased and prolonged the *in vivo* activity of these proteins. Other studies also indicate the role of sialic acid-containing carbohydrate content of darbepoetin alfa (a synthetic form of HuEPO) in increased half-life and *in vivo* bioactivity (Egrie *et al*., 2003; Macdougall, 2002; Son *et al*., 2011). Darbepoetin alfa is commercialised under the name Aranesp® and is approved by the FDA (Walsh and Jefferis, 2006).

Cerezyme, a recombinant glucocerebrosidase used to treat Gaucher disease, was engineered to have exposed mannose residues, thus facilitating its uptake by macrophages, which otherwise would have been quickly removed by hepatocytes after administration (Walsh and Jefferis, 2006). Heterogeneity in the *N*-glycans on therapeutic proteins causes difficulties for protein purification and process reproducibility and can lead to variable therapeutic efficacy. "GlycoDelete" modification strategy reported by Meuris *et al*. (2014) shortens the Golgi *N*-glycosylation pathway in mammalian cells, thus reducing the complexity of mammalian cell glycoproteins. Therapeutic human IgG expressed in GlycoDelete cells had properties such as reduced initial clearance.

One of the most important considerations to have while producing antibodies is that the glycosylation can be influenced by the cell in which it is expressed. The cell, culture conditions and the conformation of the antibodies also greatly contribute to the glycosylation of antibodies (Wright and Morrison, 1997). The success of antibody therapeutics in oncology largely depends on antibody-dependent cellular cytotoxicity (ADCC) mediated by natural killer (NK) cells (Seidel *et al*., 2013). Various studies have shown that the ability to elicit ADCC can vary depending on the antibody glycoforms. For instance, the glycosylation pattern of chCH7, an antineuroblastoma chimeric IgG1, was modified by expressing β(1,4)-*N*-acetylglucosaminyltransferase III enzyme. This enzyme catalysed the formation of bisected oligosaccharides on the antibody, thus optimising ADCC (Umana *et al*., 1999).

Metabolic oligosaccharide engineering (MOE) is a method used to modify cellular glycans with non-natural sugars (Tan *et al*., 2010). A range of non-natural sugar analogs of NeuAc, GlcNAc and GalNAc have been successfully put into organisms such as zebrafish and *Caenorhabditis elegans* (Griffin and Hsieh-Wilson, 2016). Laughlin *et al*. (2008) applied MOE to image cell surface glycans in developing zebrafish. The embryos of zebrafish were incubated with peracetylated *N*-azidoacetylgalactosamine (Ac4GalNAz), followed by two different cyclooctyne-conjugated fluorophores, so that the newly synthesised glycans could be distinguished from the older glycans and be tracked *in vivo*. Apart from imaging other applications of MOE include identifying binding partners of lectins such as Siglec CD22 and remodelling of surface glycans to modulate cellular response (Griffin and Hsieh-Wilson, 2016). Further to this, Kaewsapsak *et al*. (2013) used MOE to trigger an artificial immune response against a gut bacteria *Helicobacter pylori* that is responsible for stomach ulcers. This method exploited the differential metabolic utilisation of glycans to selectively target and elicit an immune response against bacteria.

Alternative to MOE, which involves modification of natural glycans by manipulating biosynthetic machinery, *de novo* glycan display concentrates on inserting defined glycans directly into the plasma membrane. In an approach using the *de novo* modification, *O*-type erythrocytes (that lacked A and B type antigens) were incubated with synthetic glycolipid mimics containing A and B antigen types to allow for lipid insertion in the plasma membrane. The newly remodelled cells acquired a strong immune response to anti-A and anti-B antibodies (Frame *et al*., 2007). However, a shortcoming of this approach is that this method of glycan presentation is often short lived (Griffin and Hsieh-Wilson, 2016).

The enzymatic method of the synthesis of complex oligosaccharides on proteins and lipids or in living cells are beneficial over chemical methods because, degradation or denaturation of active proteins, lipids, and living cells can be eluded. Since it is difficult to replicate the synthesis of mammalian oligosaccharides with branched and sialylated glycans with chemical methods or bacterial enzymes, enzymes of human origin are mostly ideal for this purpose (Shimma *et al*., 2006). With the emerging field of glycomics and glycan research in place, great level of advancements has been achieved in the field of medicine and research in the recent past. Research on glycan remodeling and glycan engineering has been applied *in vitro* with very few exceptions of *in vivo* applications. Recent developments in mass spectrometry, high performance liquid chromatography and lectin microarray systems for the structural analysis of glycans demand a library composed of a variety of structure-defined glycans. Since the human glycosyltransferases have strict specificity towards donor and acceptor substrates, it is reasonable to employ *in vitro* synthesis of glycans by recombinant human glycosyltransferases for producing glycan standards (Chiba *et al*., 2009). Developments in molecular biotechnology and recombinant protein expression offer great potential for *in vitro* glycan engineering and also for research purposes.

1.8. Heterologous protein expression

The expression of a gene of interest in a host organism that naturally lacks this gene is referred as heterologous expression. The recent past has seen many advances in heterologous expression of proteins especially for therapeutic and diagnostic purposes and the technique has indirectly become indispensable for the treatment of many diseases. Recombinant proteins are expressed in a range of systems including bacteria, fungi, insect and mammalian cells (Lauber *et al*., 2015). About one-third of the currently approved recombinant therapeutic proteins are produced in *Escherichia coli* strains (Huang *et al*., 2012). The ability to express and purify the desired protein in a large quantity allows for its industrial and commercial viability. In theory, recombinant protein expression is straightforward and is based on the assumption that basic principles of protein expression and function are similar in all organisms. However, many things can go wrong in practice. For example, while eukaryotic organisms share many principles of cellular functions such as cell compartmentation, intracellular transport, and regulation there are also significant differences across fungal, plant and animal cells (Frommer and Ninnemann, 1995).

The first hurdle in heterologous protein expression is the choice of host organism. Microorganisms such as bacteria, yeast, filamentous fungi and unicellular algae offer a good range of hosts for heterologous protein expression. The choice largely depends on the complexity of protein of interest. For instance, prokaryotic systems may not be compatible with the expression of eukaryotic proteins, which may have elaborate and more complicated posttranslational modifications such as glycosylation. One the other hand *E. coli* has been a popular choice of interest for heterologous expression (where possible) because of its fast growth rate, inexpensive culture conditions and fast and easy transformation with an exogenous DNA (Rosano and Ceccarelli, 2014).

One has to note that variables within the open reading frame also affect heterologous gene expression. Some of these variables include the relative frequencies of different amino acid codons, the ability of the 5′-end of the mRNA to fold into stable secondary structures, the relative abundance of G and C in the base composition of mRNA and the presence of cryptic transcriptional terminators (Gustafsson *et al*., 2012). The amino acid codons are read in the ribosomes by complementary tRNAs during translation. Hence, the preferred codon usage by an organism is reflected by the abundance of cognate tRNAs. Expression of heterologous proteins in the expression host can be limited by non-optimal codon content due to the limiting of available cognate tRNAs. Codon usage has been identified as the single most important factor in heterologous gene expression (Gustafsson *et al*., 2004). Discrepancies in codon usage can introduce significant stress on metabolic and translational processes of the host-cell and can lead to reduced translation rates and overall expression levels (Angov, 2011).

Various strategies have been used to minimise the bias in codon usage for heterologous expression. In bacteria and simple eukaryotes, by introducing the most abundant codons throughout the length of the sequence, the resulting protein would be expressed at high levels, primarily because cognate isoacceptor tRNA molecules are not rate limiting (Angov, 2011). For instance, Deng (1997) and Feng *et al*. (2000) have documented the significance of codon optimisation for enhanced heterologous expression of proteins in *E. coli*. Codon harmonisation is another approach by which codon usage frequencies from the native host is best approximated and adjusted for use in the heterologous system thereby improving the solubility and functionality of proteins during expression in the heterologous host (Angov *et al*., 2011). For example, codon harmonisation was proven the best approach in expressing a recombinant liver-stage antigen 1 (LSA 1) malarial vaccine against *Plasmodium falciparum* in *E. coli.* The problem of loss of the plasmid containing a codon-optimised LSA gene during

the exponential growth phase of *E. coli* was also solved by codon harmonisation. Moreover, the LSA protein expressed from the codon-harmonised gene was not degraded, aggregated, or detrimental to cell growth, unlike the product from the codon-optimised gene (Hillier *et al*., 2005).

Although, choice of the host and codon bias play a crucial role in heterologous protein expression, transcriptional promoters also have a critical role in recombinant protein production (Gustafsson *et al.*, 2004). Promoters used in biotechnology can be classified into four groups: (i) constitutive promoters – active in most of the tissues and developmental stages; (ii) spatiotemporal promoters – provide tissue-specific or stage-specific expression; (iii) inducible promoters – external chemical or physical signal induce them and; (iv) synthetic promoters – contain specifically defined elements within the promoter or adjacent to it. (Hernandez-Garcia and Finer, 2014).

Heterologous expression of proteins has facilitated low cost production improving the commercial viability ranging from food industry to biopharmaceuticals. Recent growth in the field of biopharmaceuticals has increased the demand of recombinant proteins for structural and biophysical studies (Assenberg *et al*., 2013). However, it should be noted that the heterologous polypeptide chain has to form a proper tertiary conformation comprehensively by appropriate disulphide bonds along with appropriate posttranslational modifications such as glycosylation that contributes to the functionality of recombinant proteins (Gustafsson *et al*., 2004).

1.8.1. Heterologous expression of human glycosyltransferases

Owing to the significance of glycosylation in maintaining the intricate structural and functional integrity of proteins, the potential of glycan modification or glyco-engineering is very promising. Glycosyltransferases are increasingly used for *in vitro* synthesis of

oligosaccharides and human glycosyltransferases have strict regioselectivity and stereospecificity. However, the difficulty of purifying these enzymes from their natural sources has led to the development of recombinant protein expression systems (Huang *et al*., 2012).

Recombinant glycosyltransferases can be used as the potential biocatalysts for the combined chemical and enzymatic synthesis of a compound library of oligosaccharides, glycosphingolipids, glycopeptides, and various artificial glycoconjugates. The structurally defined glycan-related compound library is a key resource for the discovery of new diagnostic biomarkers and therapeutic reagents and also in the basic studies of their functional roles in various biological processes (Ito *et al*., 2010). Many human glycosyltransferases are known to be glycosylated and this must be considered when choosing an appropriate host for recombinant expression (Breen, 2002). *E. coli* serves as the most widely used heterologous expression system and has been reported for the expression of recombinant glycosyltransferases (Bajazawa *et al*., 1993; Hidari *et al*., 2005; Skretas *et al*., 2009). A major drawback of *E. coli* system is its incapability of protein glycosylation and proper protein folding. To overcome these issues, *E. coli* strains have been genetically engineered to express glycosylation machinery capable of glycosylating heterologous proteins. Further, genetic mutations have been induced to provide an increased oxidative cytoplasmic environment that favours disulphide bond formation. Expressing molecular chaperones in the cells are also found effective for the proper folding of recombinant proteins (Schwarz *et al*., 2010; Valderrama-Rincon *et al*., 2012; Merritt *et al*., 2013). Apart from *E. coli*, several other systems such as mammalian cells, yeast, insect cells and plant tissues have been used to express recombinant glycosyltransferases successfully. *Saccharomyces cerevisiae* and *Pichia pastoris* are the two yeast species mainly used for heterologous glycosyltransferase

production. Table 1-5 lists some of the heterologously expressed human glycosyltransferases that were deemed functional.

Table 1-5: Examples of heterologously expressed human glycosyltransferases in different expression hosts.

Recombinant human glycosyltransferase	Expression host	Reference
Beta 1,4-galactosyltransferase (β4GalT4,T5)	Tobacco cells, *S. cerevisiae*	(Bakker *et al*., 2001; Palacpac *et al*., 1999; Shimma *et al*., 2006)
Beta 1,3-*N*-acetylglucosaminyltransferase 1 (β3GnT1)	*P. pastoris*, *E. coli*	(Hamilton *et al*., 2003; Fujiyama *et al*., 2001)
Beta 1,3-*N*-acetylglucosaminyltransferase 2 (β3GnT2)	*P. pastoris*, *S. cerevisiae*	(Hamilton *et al*., 2003; Shimma *et al*., 2006)
Beta 1,3-*N*-acetylglucosaminyltransferase 3 (β3GnT3)	*S. cerevisiae*	(Shimma *et al*., 2006)
GalNAc alpha-2, 6-sialyltransferase 1 (ST6GalNAc1)	*E. coli*	(Skretas *et al*., 2009)
Beta-gal alpha-2,6-sialyltransferase 1 (ST6Gal1)	*E. coli*, *S. cerevisiae*, *P. pastoris*	(Hidari *et al*., 2005; Shimma *et al*., 2006; Ribitsch *et al*., 2014; Luley-Goedl *et al*., 2016)
Beta-gal alpha-2,6-sialyltransferase 2 (ST6Gal2)	*S. cerevisiae*	(Shimma *et al*., 2006)

Table 1-5 *contd…*

GalNAc alpha-2, 6-sialyltransferase 2 (ST6GalNAc2)	COS-7 cells, *S. cerevisiae*	(Samyn-Petit *et al*., 2000; Shimma *et al*., 2006)
GalNAc alpha-2, 6-sialyltransferase (ST6GalNAc1,4)	*S. cerevisiae*	(Shimma *et al*., 2006)
Beta-gal alpha 2,3 sialyltransferase (ST3Gal1,3,4,6)	*S. cerevisiae*	(Shimma *et al*., 2006)
CMP-NeuAc:beta-galactoside α2-3-sialyltransferase (ST3Gal6)	Mouse L cells	(Okajima *et al*., 1999)
Alpha1,2-fucosyltransferase (FUT 1,3,6,7,8,9)	*S. cerevisiae*	(Shimma *et al*., 2006)
N-acetylgalactosaminyltransferase (pp-GalNAcT1,T3,T4,T6,T9,T10,T12,T13,T4,T15, β3GalNAcT1)	*S. cerevisiae*	(Shimma *et al*., 2006)
Polypeptide *N*-acetylgalactosaminyltransferase 2 (GalNAcT2)	*S. cerevisiae*, *E. coli*	(Shimma *et al*., 2006; Lauber *et al*., 2015)

1.8.2. Drawbacks of the commonly used heterologous expression systems

Despite the advantages of the conventional expression hosts mentioned in the previous section, there are some drawbacks that limit their application as universal hosts for recombinant protein expression. Those proteins with glycosylation or complex tertiary structure require more advanced expression systems such as yeast or mammalian cells. The production of recombinant proteins in yeast such as *S. cerevisiae* and *P. pastoris* combines simple, inexpensive and large-scale culture conditions of bacteria with the protein processing abilities of eukaryotic cells, thereby increasing the likelihood of proper folding and posttranslational modification of the recombinant protein of interest. Yet, the over-mannosylation of recombinant proteins generated in *S. cerevisiae* (Vervecken *et al*., 2004; Hou *et al*., 2012) and the β- mannose residues on proteins generated in *P. pastoris* (Hopkins *et al*., 2011) are recognised by mannose receptors on macrophages, dendritic cells and hepatic and lymphatic endothelia and are removed when injected into the mammalian circulatory system. Further, the polymannans consisting of 50 to 100 mannose residues may increase the bulkiness of recombinant proteins impairing their biological activity (Malissard *et al*., 1999; Jayaraj and Smooker, 2009). To overcome this limitation yeasts have been genetically engineered to carry mammalian like *N*-glycosylation machinery. Choi *et al*. (2003) reported the genetic engineering of *P. pastoris* to mimic the early processing of *N*-glycans in humans. This was achieved by deleting the α1-6-mannosyltransferase gene from *P. pastoris* and expressing the catalytic domain of α1-2-mannosidase from *Homo sapiens*, *Mus musculus*, *Aspergillus nidulans*, *Chlamydomonas elegans*, *Drosophila melanogaster* and *Penicillium citrinium* and the β1-2-*N*-acetylglucosaminyltransferase I (GnTI) from *H. sapiens*, *C. elegans*, *Xenopus laevis* and *D. melanogaster* in the secretory pathway.

Two important considerations to have while manufacturing recombinant proteins are the culture conditions and the downstream processing. Manipulation of cell culture conditions can

greatly influence recombinant protein productivity and posttranslational modifications (Walsh and Jefferis, 2006). For instance, a monoclonal IgG-1 produced in murine hybridoma displayed varied glycosylation pattern when cultured in three different media (Patel *et al*., 1992). The glycoform profile of recombinant therapeutic antibodies produced in expressions systems such as Chinese Hamster Ovary cells (CHO) and mouse myeloma cells can vary widely from clone to clone depending on the production and culture conditions (Lund *et al*., 1993; Cabrera *et al*., 2005; Gramer *et al*., 2011). These diversities can greatly affect the potency of recombinant protein production. As explained in previous sections, prokaryotic expression systems such as *E. coli* does not feature human-like posttranslational modifications such as glycosylation and appropriate protein folding. Moreover, the recombinant proteins can be confined in inclusion bodies (Lilie *et al*., 1998). Further downstream processing to solubilise and refold the protein *in vitro* may be required to rescue the trapped proteins (Walsh and Jefferis, 2006).

The fastest way to produce recombinant proteins in mammalian cells is by transient transfection of DNA or RNA constructs. Albeit being the most efficient transformation method, transient transfection is very expensive for large scale applications, diminishing its feasibility (Kollewe and Vilcinskas, 2013). Insect cells, on the other hand, offer an adequate compromise compared to the other systems. Even though the insect cells are less demanding compared to mammalian cells in cultivation, the use of this system comes with certain drawbacks. For example, since the insect system requires a virus mediated gene transfer, every round of heterologous protein production requires infection of new insect cells due to the eventual death of virus infected cells rendering it inferior to the bacterial and yeast systems. Also the difference in the glycosylation pattern can vary the bioactivity of the recombinant proteins (Yin *et al*., 2007). Limitations of the conventional expression systems have necessitated the search for an alternative solution that will enable large scale production

of glycoproteins with post translational modifications that are more human like. Prospective candidates for heterologous expression addressing these issues are filamentous fungi.

1.8.3. Heterologous expression of proteins in filamentous fungi

Filamentous fungi have long been used for the production of industrial enzymes, antibiotics and specialty chemicals by employing the essential elements of fungal gene expression systems similar to those that have been developed for yeasts and bacteria. An expression unit is comprised of transcriptional, translational, and secretory signals from the host organism, joined to the DNA sequence of interest. The expression unit is introduced into the host organism using a suitable transformation system (Berka and Barnett, 1989). Recent advances in fungal genomics and related experimental technologies such as gene arrays and proteomics are further enabling heterologous expression of various proteins in filamentous fungi (Nevalainen *et al*., 2005). Filamentous fungal systems have several advantages over the more conventional expression systems because of their high-level secretion of enzymes owing to their decomposer life style. Filamentous fungi have also become the host organism of choice for large scale production of recombinant proteins of eukaryotic origin. However, the complexity and relative difficulty of understanding the complex cellular functions, regulation of gene expression and protein secretory pathways in detail has impeded the rapid development of filamentous fungi as highly efficient cell factories for the production of heterologous proteins compared to the bacterial systems (Su *et al*., 2012).

Filamentous fungi in general are excellent secretors of homologous proteins, like cellulases with a reported yield of up to 100 g/L of native enzymes using advanced fermentation techniques (Cherry and Fidantsef, 2003) making them attractive candidates for heterologous gene expression (Sharma *et al*., 2009). Mutant strains of *T. reesei* have been found to secrete up to tens of grams per litre of cellulases into the culture medium while *A. niger* is shown to

secrete up to 20 grams of glucoamylase per litre (Durand *et al*., 1988; Finkelstein, 1987). For efficient production of a desired protein, it is essential to have a strong and efficient secretion mechanism. Hence, the gene of interest is placed under a strong promoter and a secretion signal sequence, which directs the translated protein to the secretory pathway. The secretion signal sequence is required for targeting the protein to the endoplasmic reticulum (ER) and its subsequent translocation across the ER membrane (Gierasch, 1989). There are two groups of fungal promoters used for heterologous expression; promoters from housekeeping genes with a high level of expression and inducible promoters (Sharma *et al*., 2009). Some of the well-known inducible promoters used in many applications of filamentous fungi include cellobiohydrolase I (*cbh1*) from *T. reesei* (Harkki *et al*., 1991) glucoamylase A (*glaA*) from *A. niger* (Smith *et al*., 1990), Taka-amylase (*amyA*) from *A. oryzae* (Tsuchiya *et al*., 1992) and 1,4-beta-xylanase A (*exlA*) from *A. awamori* (Gouka *et al*., 1996a).

Aspergillus niger, *A. oryzae*, *Trichoderma reesei*, *Acremonium chrysogenum* and *Penicillium chrysogenum* are some of the industrially important filamentous fungi. These fungi are classified GRAS (generally regarded as safe), and can thrive in relatively inexpensive substrates secreting a great amount of proteins making them suitable vehicles for heterologous protein expression and feasible workhorses in industry. The secretory nature of fungi is beneficial for easier purification of recombinant proteins from the culture supernatant eliminating the need to break cells and removing the intracellular proteins, a process which is often tedious and expensive. Moreover, fermentation techniques are well developed for filamentous fungi making it easy to grow them in large quantities (Sharma *et al*., 2009). Table 1-6 illustrates some of the heterologously expressed human proteins in filamentous fungi. Despite the development of techniques for heterologous expression of human proteins, human glycosyltransferases have not been expressed in filamentous fungi to date.

Table 1-6: Examples of filamentous fungal systems used to produce human proteins.

Protein	Expression system	Yield (mg/mL)	Reference
Interferon-a 2	*Aspergillus nidulans*	2	(MacRae *et al*., 1993)
Lactoferrin	*Aspergillus oryzae*	2	(Ward *et al*., 1995)
Human interleukin 6	*Aspergillus niger*	150	(Punt *et al*., 2002)
IgG1 (κ), trastuzumab	*Aspergillus niger*	900	(Ward *et al*., 2004)
IgG1 (κ), Hu1D10		200	
Fab′ antibody fragment		1200	
Fab′ antibody fragment	*Trichoderma reesei*	8200	(Landowski *et al*., 2016)

1.9. *Trichoderma reesei* as an expression host for recombinant protein production

Trichoderma reesei is an asexually reproducing mesophilic filamentous fungus. Previously classified as *T. viride*, it was renamed *T. reesei* due to its distinct difference in morphology (Simmons, 1977). The first strain to be isolated was named QM6a and was found to be a good secretor of cellulase enzymes (Mandels and Reese, 1957). *T. reesei* is currently being developed as a heterologous protein expression host for many proteins of interest, including pharmaceuticals that are conventionally produced by mammalian cell culture. The efficiency of production of antibody fragments in *T. reesei* has already been documented (Nyyssönen *et al*., 1993; Landowski *et al*., 2015)

The major *N*-glycan form $GlcNac_2Man_5$ synthesised by *T. reesei* offers to be a suitable precursor for mammalian glycosylation thus opening an avenue for the production of

therapeutic proteins (Ferrer-Miralles *et al*., 2015). Host strains used for the production of recombinant proteins are typically high-protein secreting mutant strains (Nevalainen and Peterson, 2014b).

Consequently, in an attempt to produce a strain with desirable protein secretion traits, a hypercellulolytic mutant was developed by a three step random mutagenesis of *T. reesei* QM6a using UV radiation and *N*-nitrosoguanidine, and was termed RUT-C30. The RUT-C30 strain exhibited a high amount of cellulase activity, relatively low protease production and catabolite de-repression by resistance to the antimetabolite 2-deoxyglucose (2DG) compared to its wild strain counterpart (Montenecourt and Eveleigh, 1977a, b, 1979). Table 1-7 illustrates production and activity of the native proteins in both the strains (Ryu and Mandels, 1980; Peterson and Nevalainen, 2012).

Table 1-7: The extracellular protein and enzyme activities of wild-type *T. reesei* QM6a and RUT-C30 following growth on 6% (w/v) roll-milled cotton in a 10 L fermenter for 14 days. FPU: Filter paper units; CMC: carboxymethylcellulase (endoglucanase); βGl: β-glucosidase; U: Enzyme units, µmol glucose produced per min in standard assay (Ryu and Mandels, 1980) (adapted from Peterson and Nevalainen, 2012).

Strain	Soluble protein (mg/mL)	FPU (U/mL)	Productivity (FPU/L/h)	CMC (U/mL)	βGl (U/mL)
QM6a	7	5	15	88	0.3
RUT-C30	19	14	42	150	0.3

1.9.1. The *Trichoderma reesei cbh1* promoter

Just as in any other system, the choice of promoter is critical in heterologous protein expression in *T. reesei*. The *cbh1* promoter is one of the most prolific inducible promoters known in fungi and is commonly used for obtaining high expression levels (Uusitalo *et al*.,

1991; Singh *et al*., 2015) even though other promoters such as the *cbh2*, *xyn2*, and *egl2* have been employed for heterologous expression (Meng *et al*., 2013; Miyauchi *et al*., 2013). Consequently, it has also been shown that introduction of multiple copies of binding sites for transcriptional activators into the *cbh1* promoter can enhance the expression rate of heterologous proteins (Liu *et al*., 2008). However, Zhang *et al*. (2010) reported that a four-copy *cbh1* promoter may have a negative effect on heterologous gene expression.

Generally, the *cbh1* promoter is induced when the natural inducer cellulose is the only carbon source available. However, studies indicate that it is possible to induce the *cbh1* promoter by other soluble carbon sources such as cellobiose, sophorose and lactose (Mandels and Reese, 1960; Mandels *et al*., 1962; Bisaria and Mishra, 1989). Multiple studies indicate that the *cbh1* promoter is regulated by multiple mechanisms under different inducers in the culture medium (Aro *et al*., 2001; 2003).

Moreover, in the presence of different carbon sources, regulation of *cbh1* is mediated by various transcription factors including CRE1. Ries *et al*. (2014) reported that under repressing conditions, that is, in the presence of glucose, the *cbh1* promoter and the protein coding regions are occupied by several positioned nucleosomes whereas in the presence of an inducer called sophorose the nucleosomes were lost from the coding region preventing CRE1 from binding to its recognition sites within the promoter region. Strains with mutated/deleted CRE1 exhibited a loss of positioned nucleosomes within the *cbh1* coding region under repressing conditions, indicating that CRE1 is important for correct nucleosome positioning. The high cellulase secreting, catabolite de-repressed variant of *T. reesei*, RUT-C30, carries a mutant variant of *cre1* gene (Ries *et al*., 2014).

The major cellulase cellobiohydrolase I (CBHI) represents more than 50 % of all protein secreted by *T. reesei* under the influence of the strong inducible *cbh1* promoter (Nyyssönen *et*

al., 1995). The CBHI molecule is composed of a catalytic domain and a carboxy-terminal cellulose-binding domain (CBD) separated by a flexible *O*-glycosylated linker (Keränen and Penttilä, 1995; Nakari-Setälä *et al*., 2009). The CBD domain can subsequently be replaced with a foreign protein of interest. Production of heterologous proteins, such as calf chymosin and human interleukin-6, has been proven to have considerably improved by fusing the foreign protein to CBHI. Table 1-8 summarises a list of heterologous proteins expressed in *T. reesei* RUT-C30 strain, under the *cbh1* promoter.

Table 1-8: Examples of heterologous proteins expressed in non-modified *T. reesei* RUT-C30, under the *cbh1* promoter (adapted from Nevalainen and Peterson (2014a)

Heterologous protein	Yield (g/L)	Culture condition	Reference
Calf chymosin	.04	Fermenter (10 L)	(Harkki *et al*., 1989)
Glucoamylase from *Hormoconis resinae*	0.7	Shake flask (50 mL)	(Joutsjoki *et al*., 1993)
Fab antibody	0.001 (Fab fragment only)	Fermenter (10 L)	(Nyyssönen *et al*., 1993)
	0.15 (CBHI-Fab fusion)		
Trichoderma harzianum endochitinase	0.13	Shake flask (50 mL)	(Margolles-Clark *et al*., 1996)
Barley endopeptidase B	0.5	Fermenter (1 L)	(Saarelainen *et al*., 1997)
Xylanase XynB from *Dictyoglomus thermophilum*	0.5	Shake flask (50 mL)	(Bergquist *et al*., 2002)
3 beta-glucosidase from *Talalaromyces emersonii*	0.0027	Shake flask (50 mL)	(Murray *et al*., 2004)
Laccase from *Melanocarpus albomyces*	0.230	Shake flask (50 mL)	(Kiiskinen *et al*., 2004)
	0.920	Fermenter (20 L)	
Xylanase (XynVI) from *Acrophialophora nainiana*	0.172	Shake flask (50 mL)	(Salles *et al*., 2007)
Cinnamoyl esterase EstA from *Piromyces equi*	0.033	Shake flask (50 mL)	(Poidevin *et al*., 2009)
Human α-galactosidase A	0.020 (secreted)	Fermenter (10L)	(Smith *et al*., 2014)
	0.636 (intracellular)		
Mammalian obestatin	0.0055	Shake flask (50 mL)	(Sun *et al*., 2016)

1.9.2. Protein secretory pathway in *Trichoderma reesei*

The successful expression of heterologous proteins in fungal hosts depends not only on the techniques employed for designing and construction of heterologous genes but also on the successful transport of the heterologously expressed protein through the intracellular secretory pathway. Secretory proteins pass through a series of intracellular compartments where they undergo posttranslational modifications (PTM) that assist with the proper folding of a protein before they are secreted outside the cells, the schematic of which is shown in Fig 1-6.

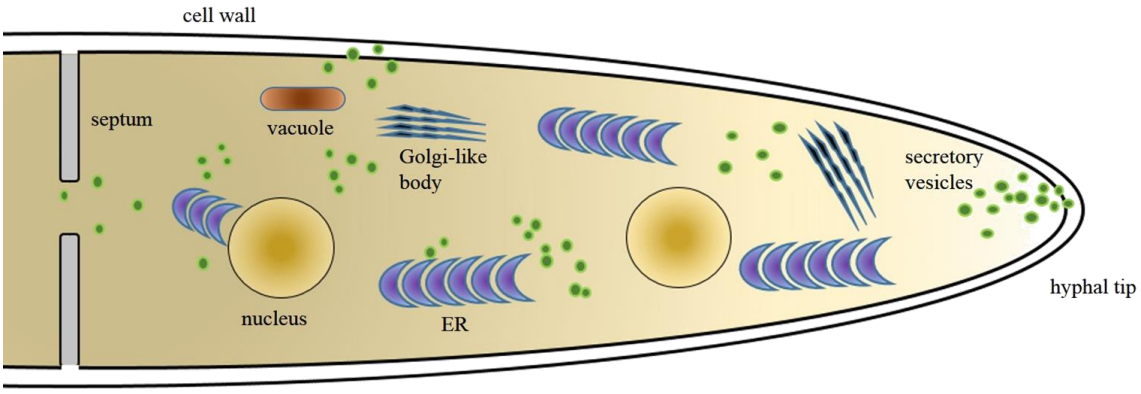


Fig 1-6: The secretory pathway of *T. reesei* showing cell organelles involved during the secretion of a protein. Modified from Saloheimo and Pakula (2012). See text for more details.

Most of the information applied in the studies on fungal secretory pathway comes from models developed in yeast and higher eukaryotes. The secretory proteins begin their journey to the extracellular medium by entering the endoplasmic reticulum (ER) where the proteins undergo distinct modifications such as glycosylation, disulphide bridge formation, phosphorylation, and subunit assembly (Conesa *et al*., 2001).

1.9.2.1. Cellular trafficking of proteins

Those proteins destined for secretion into the culture medium essentially carry a short peptide signal sequence on the *N*-terminus. In *T. reesei*, the signal sequence (ss) of the main cellulase CBHI is represented by the sequence MYRKLAVISAFLATARA. As in most other cases, the

ss of *T. reesei* is not glycosylated (Madhavan and Sukumaran, 2014). In fact, the *cbh1* ss has been successfully used for the secretion of heterologous proteins in *T. reesei* (Te'o *et al*., 2000; Sun *et al*., 2016), as well as in other hosts such as *Candida glabrata, Pichia pastoris, Saccharomyces cerevisiae, Schizosaccharomyces pombe and Kluyveromyces lactis* (Eiden-Plach *et al*., 2004; Madhavan and Sukumaran, 2014). The ss plays a significant role in determining the translocation of a protein through the ER in eukaryotes often mediated by a ribonucleoprotein called the signal recognition particle (SRP) (Koch *et al*., 2003; Halic *et al*., 2004). Two translocation routes have been postulated in *S. cerevisiae*. The first one is the SRP-dependent pathway, in which the protein is transported through the ER membrane co-translationally, which begins upon the binding of the ribosome containing a nascent polypeptide to the SRP. The ribosome-peptide-SRP complex is then directed towards the SRP receptor on the ER membrane. At the ER membrane, the nascent polypeptide enters into the ER lumen via the Sec61p channel. The bound SRP is released at this stage (Conesa *et al*., 2001).

The second route is the SRP-independent pathway in which proteins are targeted to the ER posttranslationally in a process that involves the ER chaperone BiP. This pathway initiates with the binding of the Hsp70 chaperone to a polypeptide. The complex then binds with the Sec62p-Sec72p-Sec73p translocon complex on the ER membrane. Subsequently, binding of the chaperone BiP to Sec62p facilitates the entry of the polypeptide through the Sec61p channel (Conesa *et al*., 2001). Moreover, it is the hydrophobicity of the ss peptide that determines either of these routes (Ng *et al*., 1996). Upon entry into the ER, ss on "preproteins" form a hairpin structure and is usually cleaved by signal peptidase, a protease complex residing at the lumenal face of the ER membrane (Fewell and Brodsky, 2009). The cleaved ss is further broken down by the ER resident signal peptide peptidase (Weihofen *et al*., 2002).

Once inside the ER, the proteins undergo maturation and folding to their native tertiary or quaternary conformation by ubiquitously expressed helper proteins such as chaperones and foldases. Foldases are essential in disulphide bond formation and proline isomerisation that are in turn essential for obtaining functional conformation, whereas molecular chaperones bind transiently and non-covalently to non-native or misfolded proteins thus preventing nonproductive protein–protein interactions thereby promoting correct folding (Chevet *et al*., 2001; Gething and Sambrook, 1992). Binding Protein 1 (BiP) and protein disulphide isomerase (PDI) are the two most important molecular chaperones that are characterised in *T. reesei* (Saloheimo and Pakula, 2012).

Activities localised in the ER ensure that the proteins are folded correctly and delivered into the subsequent cellular compartments. There are two mechanisms that essentially participate in the protein quality control inside the ER. The unfolded protein response (UPR) detects the presence of unfolded proteins in the ER and induces the synthesis of folding enzymes, and the ER-associated protein degradation (ERAD) degrades those proteins that fail to reach the correct conformation (Conesa *et al*., 2001). The ERAD system eliminates misfolded proteins via degradation by translocating them to the cytosol where they become ubiquitylated and degraded by a giant protease, the 26S proteasome (Thrower *et al*., 2000).

Studies indicate that UPR is induced by various factors such as inability to form disulphide bonds, Ca^{2+} depletion in ER, expression of aberrant proteins, and overexpression of normal proteins (Chapman *et al*., 1998). Travers *et al*. (2000) found that UPR not only regulates the ER-resident chaperones, but also the genes involved in secretion and ERAD, indicating an intimate coordination between these responses for maintaining a proper folding mechanism. On an additional note, Ghosh *et al*. (1982) discovered that the ER was more abundant and more developed in *T. reesei* RUT-C30 than in QM6a suggesting this could be one of the reasons for better protein productivity in RUT-C30.

Once the proteins are correctly folded in the ER, they are transported to the Golgi like structures in a filamentous fungi since structurally identifiable Golgi apparatus with the classic dictyosome organisation of stacked disc-shaped cisternae are rare among filamentous fungi. The oligomannose *N*- and *O*-glycosylation of proteins occur in these Golgi like organelles in filamentous fungi (Markham, 1995). Correct folding and glycosylation of proteins in the ER and Golgi are essential for the secretion of a protein. Following the passage through the Golgi compartment, proteins are targeted either to the plasma membrane or to the vacuoles for secretion. The trafficking of proteins to Golgi or plasma membrane occurs via a vesicle transport system. The protein is packaged inside membrane vesicles coated by a specific protein and are subsequently transported to their destinations in a process assisted by the fungal cytoskeleton. The vesicles then dock and fuse with the target membrane and subsequently release their contents to the Golgi matrix or outside the plasma membrane (Conesa *et al*., 2001).

1.9.2.2. Secretion of proteins from the fungal hyphae

Previous studies have indicated that the secretion of proteins occurs at the apical or subapical regions of the hyphae (Wösten *et al*., 1991; Gordon *et al*., 2000). However, studies on expression of calf chymosin as a stand-alone protein and in conjunction with CBHI revealed that there were spatial restrictions in secretion of foreign proteins. Further studies indicated that the expression and secretion of recombinant barley cysteine endopeptidase (EPB) was mostly confined to apical and subapical compartments of young hyphae while the presence and secretion of native CBHI was established even in older hyphal compartments in *T. reesei* (Nykänen *et al*., 1997; Nykänen *et al*., 2002). Therefore, fusion of a foreign protein with the native CBHI is expected to greatly enhance its secretion out of the cells (Nykänen, 2002; Saloheimo and Pakula, 2012).

1.9.3. Degradation of recombinant proteins by proteolysis

Fungal proteases have long been identified as bottlenecks in heterologous protein production. Even antibodies, which are thought to be relatively stable molecules, are susceptible to protease degradation. Expression of recombinant proteins in filamentous fungi requires the protease activity to be drastically reduced (van den Hombergh *et al*., 1997; Braaksma and Punt, 2008). In fact, antibodies are the most successfully produced proteins of human origin.

The total number of proteases predicted in *T. reesei* amounts to 383 (Druzhinina *et al*., 2012). The extensive ensemble of proteases secreted by *T. reesei* is found to be active across a wide range of pH (Table 1-9). For instance, at pH 5 and lower, the major protease of *T. reesei* has been shown to be acid aspartic protease (Haab *et al*., 1990), whereas subtilisin like serine protease has been seen active across a pH of 6 to 11 (Dunaevsky *et al*., 2000; Landowski *et al*., 2015). Haab *et al*. (1990) further demonstrated that the rate of proteolysis is pH dependent. The *T. reesei* QM9414 proteases showed highest activity towards CBHI at a pH of 3.0-4.0, while at pH 6.0 it dropped to 30% of the maximal protease activity.

Table 1- 9: Proteases secreted by *T. reesei* are active across a wide range of pH.

Protease group	Active pH	References
Aspartic acid protease	2.5-5.5	(Haab *et al*., 1990; Eneyskaya *et al*., 1990; Landowski *et al*., 2015)
Sedolisins	1.5-7.0	(Reichard *et al*., 2006; Druzhinina *et al*., 2012)
Trypsin like Serine protease	7.0-8.0	(Sharon *et al*., 2011; Landowski *et al*., 2015)
Glutamic protease	2.0-6.0	(Sims *et al*., 2004; Sriranganadane *et al*., 2011; Landowski *et al*., 2015)
Subtilisin like serine protease	6.0-11.0	(Dunaevsky *et al*., 2000; Landowski *et al*., 2015)
Metalloprotease	7.5-8.0	(Markaryan *et al*., 1994; Landowski *et al*., 2015)

Landowski *et al*. (2015) reported that they have identified 13 major secreted proteases including aspartic, glutamic, subtilisin-like, trypsin-like, sedolisin and metalloproteases, that were associated with the degradation of therapeutic antibodies, interferon alpha 2b and insulin like growth factor expressed in *T. reesei.* The sequential deletion of these major proteases not only increased the overall yield of the recombinant mammalian immunoglobulin and STT3 subunit of OSTase (from *Leishmania major, Trichomonas vaginalis, Leishmania infantum* and *Entamoeba histolytica*) up to 5 g/L but also promoted an increased growth rate of the deletion strains (Natunen *et al*., 2015). Table 1-10 summarises the improved yield of other human proteins expressed in the *T. reesei* protease deficient strain that lacked the 13 major secreted proteases (Landowski *et al*., 2016).

Table 1- 10: Yield of human proteins expressed in the protease deficient *T. reesei* strain lacking 13 major secreted proteases (adapted from Landowski *et al*. (2016)).

Protein	Yield (g/L)
Full length antibodies	7.6
Antibody Fab fragments	8.2
Single-chain antibodies	5.9
Single domain antibody multimers	1.9
Interferon α-2b	7.9
IGF1	8.0
Elastin like polypeptide (ELP5)	0.02-0.05

Even though extracellular or secreted proteases are considered a major threat to recombinant protein production, intracellular degradation is also possible within the lysosomes, ER, or by membrane associated proteases. A study of expression of human interleukin 6 (hIL-6) in *A. awamori* specified various proteolytic events occurring during heterologous expression. The hIL-6 was expressed under endogenous 1,4-beta-xylanase A (*exlA*) promoter with a single

gene copy integrated into the *A. awamori pyrG* locus. It was observed that the level of protein secreted was not in accordance with the high level of transcribed mRNA. Moreover, the purified hIL was found to be stable in the culture medium indicating that the extracellular proteases are not responsible for the low yield. In addition to this, the BiP mRNA expression levels in the recombinant strain were also similar to the control strain indicating that protein degradation through UPR did not occur (Gouka *et al*., 1996b). However, later when hIL-6 was expressed in conjunction with glucoamylase joined by a Kex2 cleavage site, it was identified that 60 to 90% of the hIL-6 protein was degraded intracellularly or by mycelium associated proteases (Gouka *et al*., 1997). This finding confirms that there are multiple factors associated with disintegration of a recombinant protein when expressed in a filamentous fungus and extracellular proteases is just one among them.

1.10. Fluorescent reporter proteins

Fluorescent reporter proteins are widely used for labelling cells, recombinant proteins and as reporters of gene expression. Since the reporter proteins can be easily quantitated, expression of a particular gene can be evaluated by fusing it to a reporter gene. Hence, in principle, the expression of the reporter gene should accurately reflect the expression of the gene of interest since the genes are fused to each other (Albano *et al*., 1998). The discovery of *Aequorea victoria* green fluorescent protein (avGFP) capable of forming an intrinsic chromophore has enabled the use of fluorescent proteins in cell biology applications. The non-toxic characteristic of GFP further makes it an ideal candidate for recombinant fusion protein expression (Koike *et al*., 2013). The successful use of GFP as fluorescent marker has assisted in understanding the interaction of fungal phytopathogens such as *Ustilago maydis* and *Cochliobolus heterostrophus* with host plants (Maor *et al*., 1998; Spellig *et al*., 1996). However, some of the limitations of GFP include a low turnover rate, 2 h lag time for the auto activation of its chromophore, improper folding at higher temperatures (37°C) leading to non-

fluorescence and the requirement of oxygen for chromophore maturation and activation (Lorang *et al*., 2001). It has been reported that GFP tagged hydrophobin has been successfully expressed in *T. reesei* as a fusion protein (Mustalahti *et al*., 2013).

Subsequently, another protein that differed from GFP in its spectral property by emitting red fluorescence was discovered and was named the DsRed. Initially isolated from the *Discosoma* species, it became the first commercially available red fluorescent protein (RFP). Further discoveries and the introduction of a number of desirable changes to the physical and biochemical properties of RFP were achieved that enriched its brightness, maturation efficiency, photostability, pH stability and minimised cytotoxicity thereby significantly improving the utilisation of RFPs for live-cell microscopy (Piatkevich and Verkhusha, 2011).

Photoactivatable fluorescent proteins (PAFPs), even though a small category, are beginning to find extensive uses far from the normal fluorescent proteins. PAFPs have little or no initial fluorescence but can be switched on by irradiation with violet light. One such reporter protein is the red fluorescent protein, mCherry (Subach *et al*., 2009). First developed by Shaner *et al*. (2004), mCherry has an ideal excitation/ emission wavelength of 587/610 nm and an extinction coefficient of 72, 000 $M^{-1}cm^{-1}$. The mCherry is advantageous in generating fusion proteins due to the increased photostability and their relative tolerance to N- and C-terminal fusions (Shaner *et al*., 2004). In addition to this, Li *et al*. (2008) describes the relative tolerance of mCherry to retain the fluorescence property despite the insertions of five-residue peptides at various regions within the protein. In filamentous fungi, successful expression of mCherry has been reported for *T. reesei* strain RUT-C30 and *Botrytis cinerea* (Sun, 2008; Leroch *et al*., 2011).

The protein complexes carrying mCherry reporter proteins require a more reliable method for detection than conventional optical microscopy techniques. This is because the cellular

processes are often localised with a relatively shorter half-life. Moreover, a high spatial resolution is required along with data acquisition within the timescale of the biological process under investigation (Digman *et al*., 2009). Confocal laser scanning microscopy increases the capability of mainstream fluorescent microscopy and is much more advantageous. Excitation by a He/Ar laser is focused onto a particular x/y coordinate of the specimen and the image is scanned on a single plane. This further prevents exposure of the entire sample that may result in photobleaching. Images can also be acquired by scanning in a vertical (z) direction to create a three dimensional image. The images are then recorded with an imaging software, which allows various degrees of analysis and manipulation (Paddock, 1999).

Fluorescent tagging of secreted proteins not only aid in quantitation but also provide an efficient strategy to identify the intracellular location of an expressed protein, its secretory pathway and also to determine any bottlenecks during secretion. The exact subcellular localisation of a protein of interest can be established by the simultaneous staining of cell organelles in the secretory pathway (ER, Golgi, vesicles) with fluorescent dyes (Nevalainen and Peterson, 2014a). Further, Kaur *et al*. (2015) elaborates the utilisation of two different fluorescent reporter proteins, the yellow fluorescent protein (YFP) and mCherry, expressed in the bacterium *Pseudomonas aeruginosa* and the fungus *Scedosporium aurantiacum* respectively, in an effort to study the interaction between these organisms which can coexist in the lungs of cystic fibrosis patients.

1.11. Aims of the project

Trichoderma reesei offers an excellent expression system for a wide range of recombinant proteins. However, not many proteins of human origin have been produced in this fungal system. To add to the knowledge, the broad aim of the current research was to explore expression of human sialyltransferases in *T. reesei*.

More specific aims of this study included:

1. Generation of *T. reesei* RUT-C30 transformants containing genes encoding the biologically important human sialyltransferases (i) ST3 beta-galactoside alpha2-3 sialyltransferase 3 (ST3Gal3), (ii) ST6 beta-galactoside alpha2-6-sialyltransferase 1 (ST6Gal1) and (iii) ST8 alpha-*N*-acetyl-neuraminide alpha2-8-sialyltransferase 3 (ST8Sia3). From these, the most promising transformants were selected for further studies.

2. Monitoring production, secretion and processing of the recombinant ST6Gal1-mCherry fusion protein in comparison with the recombinant ST6Gal1 produced without the reporter protein.

3. Investigating the role of fungal proteases and the Endoplasmic Reticulum associated quality control mechanisms in degradation of the recombinant ST6Gal1.

4. Devising a lectin based activity assay to examine the functionality of the recombinant ST6Gal1 expressed in *T. reesei.*

In this project, successful expression of the human sialyltransferase ST6Gal1 in *T. reesei* is reported for the first time.

2

MATERIALS AND METHODS

2.1. Culture media

Both liquid and solid culture media with necessary nutrients were used in this work to support the growth of the bacterial and fungal strains. All ingredients were purchased from Sigma Aldrich, USA. The following methods are shared with all parts of the research and specific methods are introduced in each chapter.

2.1.1. Bacterial culture media

Luria broth (LB) was prepared to propagate bacterial cells. To make 600 mL of the medium, tryptone (6 g), yeast extract (3 g) and NaCl (6 g) were dissolved in 600 mL of sterile Milli-Q water (Bertani, 1951). Super Optimal Broth (SOB) medium was prepared by dissolving tryptone (20 g), yeast extract (5 g), NaCl (0.58 g), KCl (0.2 g), $MgCl_2 \cdot 6H_2O$ (2g), $MgSO_4 \cdot 7H_2O$ (2.5 g) in 1 L Milli-Q water (Hanahan, 1983). LB-agar was used to grow bacterial cultures and was prepared by adding 12 g of agar to 1 L of LB medium. All media were sterilised by autoclaving at 121ºC for 20 min.

2.1.2. Fungal culture media

Potato dextrose agar (PDA, BD Difco, USA) plates were prepared to for plate cultures. The cellobiose/lactose/soy hydrolysate medium (CLS) was prepared for liquid culturing of fungal transformants. Cellobiose is a soluble carbon source and it promotes gene expression under the *cbh1* promoter (Lim *et al*., 2001). CLS medium is composed of a mineral salt solution that contains KH_2PO_4 (15 g/L), $(NH_4)_2PO_4$ (5 g/L) and 100x mineral stock (10 mL/L). Mineral stock (100x) was made of $FeSO_4 \cdot 7H_2O$ (500 mg/L), $MnSO_4 \cdot H_2O$ (100 mg/L), $ZnSO_4$ (100 mg/L) and $CoSO_4 \cdot 7H_2O$ (200 mg/L). The pH of the CLS medium was adjusted to 6.5 using 1 M KOH. The pH was measured at room temperature on a Cyberscan pH510 (Eutech Instruments). Further, cellobiose (10 g/L) and 15% (w/v) soybean extract stock (200mL/L, prepared by autoclaving 15% (w/v) soybean flour in 1 L Milli-Q water) were added to the

mineral salt solution and topped up to the desired volume by adding Milli-Q water; the medium was autoclaved at 121ºC for 20 min. The medium was then cooled down to room temperature and 20% (w/v) lactose (50 mL/L, filter sterilised), 1M $MgSO_4$ (2.4 mL/L, filter sterilised) and 1M $CaCl_2$ (5.4 mL/L, filter sterilised) were added to make the final medium.

2.2. Antibiotics used for the selection of transformants

Recombinants were selected based on their antibiotic resistance. Ampicillin was used at a final concentration of 100 µg/mL in the LB agar plates to select the recombinant bacterial colonies. Culture medium with hygromycin B (AG Scientific) at a final concentration of 80 µg/mL was used to select for fungal transformants.

2.3. Microbial strains and culture conditions

Escherichia coli strain DH5α (New England Biolabs, USA) was used as a host for propagation and amplification of plasmids. A high-protein producing strain of the mesophilic filamentous fungus *Trichoderma reesei,* RUT-C30 (Montenecourt and Eveleigh, 1979a), was used as host strain for the expression of recombinant human ST6Gal1. *T. reesei* RUT-C30 is a strain with low protease activity (Sheir-Neiss and Montenecourt, 1984; reviewed in Peterson *et al*., 2011).

2.3.1. Harvesting of fungal conidia

T. reesei RUT-C30 was cultured on PDA plates for the formation of conidia. PDA plates inoculated with fungal conidia were incubated at 28ºC in the dark for 72-96 h, until the plates were covered with hyphae, before they were exposed to light at room temperature for enhancing conidiation. The conidia were harvested by adding 5 mL of harvesting solution (0.9% (w/v) NaCl, 0.01% (v/v) Tween 80) and gently scraping off with a sterile plastic spreader (Copan, Italy, cat# 174CS05). The conidial suspension was filtered through sterile cotton wool to remove any hyphae. Conidia were counted on a haemocytometer (BRAND®,

Germany). The filtered conidia were stored at -80ºC in an equal amount of sterile glycerol and Revco solution (42 mM K_2HPO_4, 22 mM KH_2PO_4, 1.7 mM sodium citrate, 0.4 mM $MgSO_4$, pH 8.5).

2.3.2. Fungal cultures

Liquid fungal cultures generally contained 50 mL of CLS medium with a standard inoculum of 1×10^8 conidia/50 mL in 250 mL flat bottom conical flasks. Fungal cultures for preliminary screening were grown in 3 mL CLS medium in 24 well polypropylene plates (Enzyscreen, Netherlands). The cultures were incubated at 28ºC with constant shaking at 220 rpm. Cultures were grown in the dark to limit photobleaching of the fluorescent reporter protein present in the transformant strains carrying the sialyltransferase genes (ST3Gal3 and ST6Gal1). No alterations to the initial culture conditions were introduced during the entire course of the work. All experiments with the exception of preliminary screening and qRT-PCR were performed using three biological replicates.

2.3.3. Bacterial cultures

E. coli DH5α was grown in SOC medium (SOB medium with 20 mM glucose) at 37ºC post transformation for one hour with constant shaking at 250 rpm. Heat shock transformation method described by Inoue *et al*. (1990) was used for transforming *E. coli* DH5α cells with suitable plasmids. The transformed cells were selected by growing them on LB-agar plates containing ampicillin (100 µg/mL) at 37ºC for 10-12 h. Single colonies from the selection plate were inoculated into 5 mL LB medium and grown for 12-14 h at 37ºC with constant shaking at 250 rpm.

2.4. Molecular techniques

2.4.1 Restriction digestion and ligation of DNA

All restriction endonucleases were from New England Biolabs (USA). All (restriction) digestions were carried out at 37ºC using suitable buffers from the manufacturer as per their instructions. DNA ligation reactions were performed at room temperature using T4 DNA Ligase from Fermentas (USA) according to manufacturer's instructions.

2.4.2. Agarose gel electrophoresis

DNA analyses and PCR product confirmations were carried out on 1% (w/v) agarose gels (Amresco, USA) prepared in 1× TBE buffer (unless stated otherwise). The GelRed™ stain (Biotium, USA) was used for visualisation of nucleic acids as instructed by the manufacturer. DNA was mixed with 5× gel loading dye before electrophoresis was performed at 80V for 50 min in 1× TBE buffer. The 1 Kb plus DNA ladder (Fermentas USA) was used in all DNA gel electrophoreses. Qiagen QIAquick gel extraction kit was used to extract and purify DNA from the gels when required.

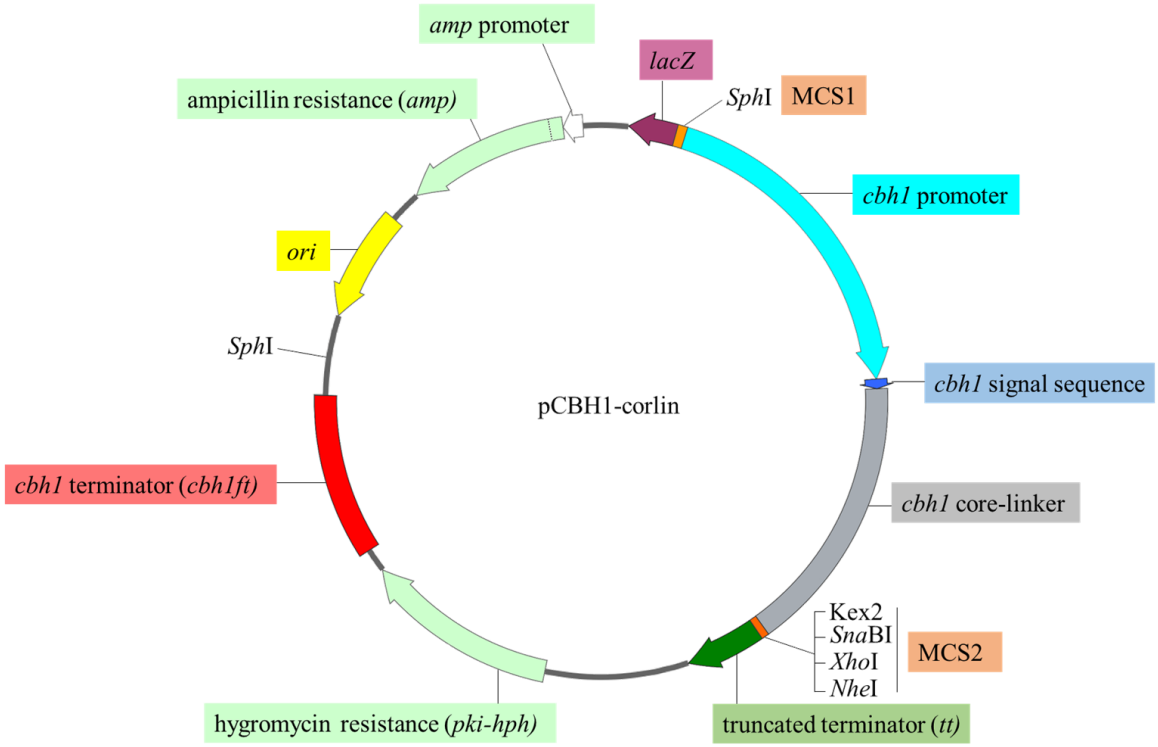
2.4.3. Polymerase chain reaction (PCR)

Polymerase chain reaction was used to amplify small segments of DNA. PCR amplifications were performed on bacterial colonies, plasmid DNAs, genomic DNAs or reverse transcribed RNAs in the Eppendorf MasterCycler EP S system. PCR reactions were performed using One*Taq*® DNA polymerase from New England Biolabs. A standard reaction contained 1 µL each of 10 pM forward and reverse primers, 5 µL of 5× One*Taq*® standard reaction buffer, 1 µL of 10 mM dNTPs and approximately 10 ng of template DNA. The reaction mixture was then made up to 50 µL using sterile Milli-Q H_2O. Standard reaction conditions with the following temperature gradient were used. Denaturation temperature of 95ºC for 10 min was used on an initial cycle and 30 s for the rest of the reaction cycles. Annealing temperatures of

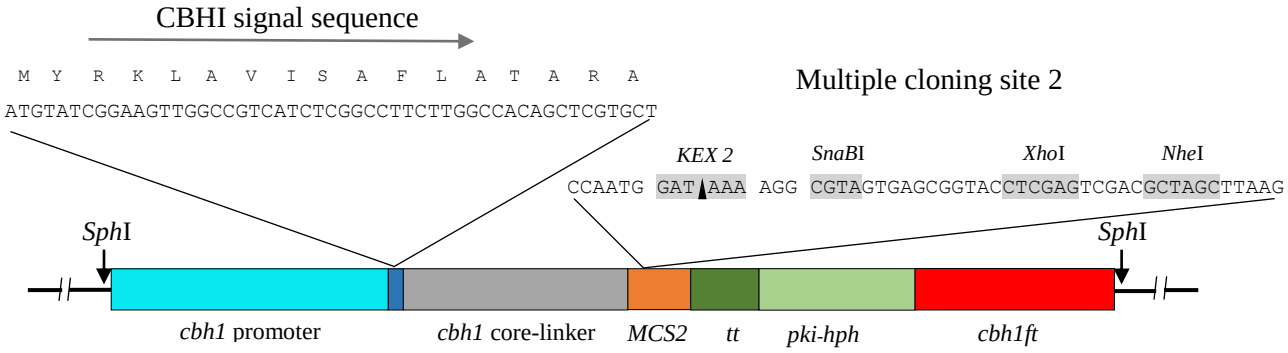
56-60ºC, depending on the primers (Table 2-1), were used for a period of 30 s. The extension was performed at 72ºC and for 30 s to 1 min depending on the length of the amplified segment and up to 2 min for the final cycle. The amplification steps were repeated for up to 40 cycles. PCR products were excised and extracted from agarose gel using Qiagen QIAquick gel extraction kit when required.

2.4.4. Expression vector

The plasmid pCBH1corlin (Te'o and Nevalainen, 2011) was used to build the expression vector for heterologous expression of the sialyltransferase genes (*ST3Gal3, ST6Gal1* and *ST8Sia3*). Parts of the plasmid include a *cbh1* promoter, *cbh1* signal sequence (or secretion sequence), *cbh1* core-linker sequence that facilitates protein secretion and stability, a truncated transcription terminator sequence followed by the *E. coli* hygromycin resistance gene (*hph)* under the *T. reesei pki* promoter, that conveys antibiotic resistance to transformants, and *cbh1* transcription terminator sequence (*cbh1ft*). The plasmid is also designed to have the ampicillin resistance gene (a*mp*) that enables propagation and amplification of plasmids in *E.coli* DH5α. There are two multiple cloning sites MCS 1 and MCS 2 out of which MCS 2 was utilised for cloning of cDNA encoding ST3Gal3, ST6Gal1 and ST8Sia3. The structure of pCBH1corlin vector is shown in Fig 2-1.



A



B

Fig 2-1: Structure of the pCBH1corlin plasmid (Te'o and Nevalainen, 2011) **A:** Circular representation of the plasmid containing the *cbh1* promoter, *cbh1* signal sequence, *cbh1* core-linker sequence, a truncated terminator (*tt)* sequence followed by hygromycin resistance gene (*hph*) expressed under the *T.reesei pki* promoter. The ampicillin (antibiotic) resistance gene aids in the cloning of the plasmid in *E. coli*. The multiple cloning site 2 (MCS 2) was employed in this project for inserting the sialyltransferase cDNAs into the plasmid (schematic generated using SnapGene®). **B:** The linear representation of pCBH1corlin vector showing the CBHI signal sequence and the multiple cloning site 2 (MCS2). The MCS2 carries gene sequence encoding the cleavage site for KEX2 like protease and the restriction sites for the restriction endonucleases *Sna*BI, *Xho*I and *Nhe*I.

*Diagrams not according to scale

2.4.5. Designing the human sialyltransferase expression constructs

The cDNA sequences for human sialyltransferases were generated by reverse translating the amino acid sequence of the full sized protein (protein id: P15907) obtained from UniprotKB. A truncated version of the cDNA was designed based on the available nucleotide sequence, excluding the nucleotides encoding the transmembrane domain and the cytoplasmic tail at the 5′ end. The amino-terminus and the transmembrane domain have been previously shown not to be required for the catalytic activity of hST6Gal1 (Donadio *et al.*, 2003). The cDNA sequences are shown in Fig 2-2 to 2-5.

Strep-tag^{TM} II is a short peptide tag, composed of eight amino acids, with a molecular weight of 1000 Da. The specific affinity of *Strep*-tag^{TM} II to *Strep*-Tactin® resin aids in the purification of recombinant proteins (Schmidt and Skerra, 2007). Nucleotide sequence encoding the *Strep*-tag^{TM} II was added to the N-terminus of the sialyltransferase cDNA (lacking the transmembrane domain) using a short linker sequence consisting of four amino acids in order to render flexibility. Restriction sites for *Xho*I and *Cla*I were engineered to the 5′- and 3′-termini of the *Strep*-tag^{TM} II sequence respectively. A stop codon was inserted at the 3′-terminus of the sialyltransferase cDNA followed by a *Nhe*I restriction site. It was required to insert a random nucleotide (red highlight in Fig 2-4 and 2-5) in two of the sialyltransferase cDNA constructs, between *Xho*I and *Strep*-tag^{TM} II to make sure that there were no stop codons introduced within the open reading frame after it was cloned into the pCBH1corlin plasmid vector. All gene manipulations were performed *in silico* using the software Snapgene® ver. 2.2.2 (from GSL Biotech; available at www.snapgene.com) and Geneious ver. R7 (http://www.geneious.com; Kearse *et al.*, 2012). The synthesis of the *in silico* designed cDNA was outsourced to Genscript (USA).

*Xho*I *Strep*-tag™ II *Cla*I Short linker

```
   1 CTCGAGTGGAGCCATCCCCAGTTCGAGAAGATCGATGGCGGCAGCGGCAAGGAGAAGAAG  60
  61 AAGGGCAGCTACTACGACAGCTTCAAGCTGCAGACCAAGGAGTTCCAGGTCCTGAAGAGC 120
 121 CTGGGCAAGCTGGCCATGGGCAGCGACAGCCAGAGCGTCAGCAGCAGCACCCAGGAC 180
 181 CCCCATCGCGGCCGCCAGACCCTGGGCAGCCTGCGCGGCCTGGCCAAGGCCAAGCCCGAG 240
 241 GCCAGCTTCCAGGTCTGGAACAAGGACAGCAGCAGCAAGAACCTGATCCCCCGCCTGCAG 300
 301 AAGATCTGGAAGAACTACCTGAGCATGAACAAGTACAAGGTCAGCTACAAGGGCCCCGGC 360
 361 CCCGGCATCAAGTTCAGCGCCGAGGCCCTGCGCTGCCATCTGCGCGACCATGTCAACGTC 420
 421 AGCATGGTCGAGGTCACCGACTTCCCCTTCAACACCAGCGAGTGGGAGGGCTACCTGCCC 480
 481 AAGGAGAGCATCCGCACCAAGGCCGGCCCCTGGGGCCGCTGCGCCGTCGTCAGCAGCGCC 540
 541 GGCAGCCTGAAGAGCAGCCAGCTGGGCCGCGAGATCGACGACCATGACGCCGTCCTGCGC 600
 601 TTCAACGGCGCCCCCACCGCCAACTTCCAGCAGGACGTCGGCACCAAGACCACCATCCGC 660
 661 CTGATGAACAGCCAGCTGGTCACCACCGAGAAGCGCTTCCTGAAGGACAGCCTGTACAAC 720
 721 GAGGGCATCCTGATCGTCTGGGACCCCAGCGTCTACCATAGCGACATCCCCAAGTGGTAC 780
 781 CAGAACCCCGACTACAACTTCTTCAACAACTACAAGACCTACCGCAAGCTGCATCCCAAC 840
 841 CAGCCCTTCTACATCCTGAAGCCCCAGATGCCCTGGGAGCTGTGGGACATCCTGCAGGAG 900
 901 ATCAGCCCCGAGGAGATCCAGCCCAACCCCCCCAGCAGCGGTATGCTGGGCATCATCATC 960
 961 ATGATGACCCTGTGCGACCAGGTCGACATCTACGAGTTCCTGCCCAGCAAGCGCAAGACC 1020
1021 GACGTCTGCTACTACTACCAGAAGTTCTTCGACAGCGCCTGCACCATGGGCGCCTACCAT 1080
1081 CCCCTGCTGTACGAGAAGAACCTGGTCAAGCATCTGAACCAGGGCACCGACGAGGACATC 1140
1141 TACCTGCTGGGCAAGGCCACCCTGCCCGGCTTCCGCACCATCCATTGCTAGGCTAGC     1197
                                                    NheI
```

Fig 2-2: cDNA sequence of ST6Gal1 for cloning into the pCBH1corlin vector with mCherry (section 2.4.6). Double underlining indicates the nucleotide sequence of *Strep*-tag™ II; nucleotide sequence with bold letters indicate the sequence for a short linker; nucleotide sequences with a yellow highlight indicates sites for restriction endonucleases. The nucleotide sequence in red letters indicates a stop codon to stop translation. The truncated human ST6Gal1 is encoded by the nucleotide sequence indicated by green letters and is comprised of 1140 base pairs.

*Xho*I *Strep*-tag™ II *Cla*I Short linker

```
   1 CTCGAGTGGAGCCATCCCCAGTTCGAGAAGATCGATGGCGGCAGCGGCAAGCTGCATCTG  60
  61 CTGCAGTGGGAGGAGGACAGCAACAGCGTCGTCCTGAGCTTCGACAGCGCCGGCCAGACC 120
 121 CTGGGCAGCGAGTACGACCGCCTGGGCTTCCTGCTGAACCTGGACAGCAAGCTGCCCGCC 180
 181 GAGCTGGCCACCAAGTACGCCAACTTCAGCGAGGGCGCCTGCAAGCCCGGCTACGCCAGC 240
 241 GCCCTGATGACCGCCATCTTCCCCCGCTTCAGCAAGCCCGCCCCCATGTTCCTGGACGAC 300
 301 AGCTTCCGCAAGTGGGCCCGCATCCGCGAGTTCGTCCCCCCCTTCGGCATCAAGGGCCAG 360
 361 GACAACCTGATCAAGGCCATCCTGAGCGTCACCAAGGAGTACCGCCTGACCCCCGCCCTG 420
 421 GACAGCCTGCGCTGCCGCCGCTGCATCATCGTCGGCAACGGCGGCGTCCTGGCCAACAAG 480
 481 AGCCTGGGCAGCCGCATCGACGACTACGACATCGTCGTCCGCCTGAACAGCGCCCCCGTC 540
 541 AAGGGCTTCGAGAAGGACGTCGGCAGCAAGACCACCCTGCGCATCACCTACCCCGAGGGC 600
 601 GCCATGCAGCGCCCCGAGCAGTACGAGCGCGACAGCCTGTTCGTCCTGGCCGGCTTCAAG 660
 661 TGGCAGGACTTCAAGTGGCTGAAGTACATCGTCTACAAGGAGCGCGTCAGCGCCAGCGAC 720
 721 GGCTTCTGGAAGAGCGTCGCCACCCGCGTCCCCAAGGAGCCCCCCGAGATCCGCATCCTG 780
 781 AACCCCTACTTCATCCAGGAGGCCGCCTTCACCCTGATCGGCCTGCCCTTCAACAACGGC 840
 841 CTGATGGGCCGCGGCAACATCCCCACCCTGGGCAGCGTCGCCGTCACCATGGCCCTGCAT 900
 901 GGCTGCGACGAGGTCGCCGTCGCCGGCTTCGGCTACGACATGAGCACCCCCAACGCCCCC 960
 961 CTGCATTACTACGAGACCGTCCGCATGGCCGCCATCAAGGAGAGCTGGACCCATAACATC 1020
1021 CAGCGCGAGAAGGAGTTCCTGCGCAAGCTGGTCAAGGCCCGCGTCATCACCGACCTGAGC 1080
1081 AGCGGCATCTAGGCTAGC                                           1098
```

*Nhe*I

Fig 2-3: cDNA sequence of ST3Gal3 for cloning into the pCBH1corlin vector with mCherry (section 2.4.6). Double underlining indicates the nucleotide sequence of *Strep*-tag™ II; nucleotide sequence in bold letters indicates the sequence for a short linker; nucleotide sequences with a yellow highlight indicate sites for restriction endonucleases. The nucleotide sequence in red font indicates a stop codon to stop translation. The truncated human ST3Gal3 is encoded by the nucleotide sequence indicated by green font and is comprised of 1041 base pairs.

*Xho*I *Strep*-tag™ II *Cla*I Short linker

```
   1 CTCGAGATGGAGCCATCCCCAGTTCGAGAAGATCGATGGCGGCAGCGGCAAGGAGAAGAA   60
  61 GAAGGGCAGCTACTACGACAGCTTCAAGCTGCAGACCAAGGAGTTCCAGGTCCTGAAGAG  120
 121 CCTGGGCAAGCTGGCCATGGGCAGCGACAGCCAGAGCGTCAGCAGCAGCAGCACCCAGGA  180
 181 CCCCCATCGCGGCCGCCAGACCCTGGGCAGCCTGCGCGGCCTGGCCAAGGCCAAGCCCGA  240
 241 GGCCAGCTTCCAGGTCTGGAACAAGGACAGCAGCAGCAAGAACCTGATCCCCCGCCTGCA  300
 301 GAAGATCTGGAAGAACTACCTGAGCATGAACAAGTACAAGGTCAGCTACAAGGGCCCCGG  360
 361 CCCCGGCATCAAGTTCAGCGCCGAGGCCCTGCGCTGCCATCTGCGCGACCATGTCAACGT  420
 421 CAGCATGGTCGAGGTCACCGACTTCCCCTTCAACACCAGCGAGTGGGAGGGCTACCTGCC  480
 481 CAAGGAGAGCATCCGCACCAAGGCCGGCCCCTGGGGCCGCTGCGCCGTCGTCAGCAGCGC  540
 541 CGGCAGCCTGAAGAGCAGCCAGCTGGGCCGCGAGATCGACGACCATGACGCCGTCCTGCG  600
 601 CTTCAACGGCGCCCCCACCGCCAACTTCCAGCAGGACGTCGGCACCAAGACCACCATCCG  660
 661 CCTGATGAACAGCCAGCTGGTCACCACCGAGAAGCGCTTCCTGAAGGACAGCCTGTACAA  720
 721 CGAGGGCATCCTGATCGTCTGGGACCCCAGCGTCTACCATAGCGACATCCCCAAGTGGTA  780
 781 CCAGAACCCCGACTACAACTTCTTCAACAACTACAAGACCTACCGCAAGCTGCATCCCAA  840
 841 CCAGCCCTTCTACATCCTGAAGCCCCAGATGCCCTGGGAGCTGTGGGACATCCTGCAGGA  900
 901 GATCAGCCCCGAGGAGATCCAGCCCAACCCCCCCAGCAGCGGTATGCTGGGCATCATCAT  960
 961 CATGATGACCCTGTGCGACCAGGTCGACATCTACGAGTTCCTGCCCAGCAAGCGCAAGAC 1020
1021 CGACGTCTGCTACTACTACCAGAAGTTCTTCGACAGCGCCTGCACCATGGGCGCCTACCA 1080
1081 TCCCCTGCTGTACGAGAAGAACCTGGTCAAGCATCTGAACCAGGGCACCGACGAGGACAT 1140
1141 CTACCTGCTGGGCAAGGCCACCCTGCCCGGCTTCCGCACCATCCATTGCTAGGCTAGC    1198
```

*Nhe*I

Fig 2-4: cDNA sequence of ST6Gal1 for cloning into the pCBH1corlin vector without mCherry (section 2.4.6). Double underlining indicates the nucleotide sequence of *Strep*-tag™ II; nucleotide sequence in bold letters indicates the sequence for a short linker; nucleotide sequences with a yellow highlight indicate sites for restriction endonucleases. The nucleotide sequence in red font indicates a stop codon to stop translation. The red highlight indicates an extra nucleotide that had to be manually inserted to avoid any stop codons being introduced in the open reading frame (ORF). The truncated human ST6Gal1 is encoded by the nucleotide sequence indicated by green letters and is comprised of 1140 base pairs.

*Xho*I *Strep*-tag™ II *Cla*I Short linker

```
   1 CTCGAGATGGAGCCATCCCCAGTTCGAGAAGATCGATGGCGGCAGCGGCAAGAAGGAGAA   60
  61 CATCTTCACCACCCCCAAGTACGCCAGCCCCGGCGCCCCCCGCATGTACATGTTCCATGC  120
 121 CGGCTTCCGCAGCCAGTTCGCCCTGAAGTTCCTGGACCCCAGCTTCGTCCCCATCACCAA  180
 181 CAGCCTGACCCAGGAGCTGCAGGAGAAGCCCAGCAAGTGGAAGTTCAACCGCACCGCCTT  240
 241 CCTGCATCAGCGCCAGGAGATCCTGCAGCATGTCGACGTCATCAAGAACTTCAGCCTGAC  300
 301 CAAGAACAGCGTCCGCATCGGCCAGCTGATGCATTACGACTACAGCAGCCATAAGTACGT  360
 361 CTTCAGCATCAGCAACAACTTCCGCAGCCTGCTGCCCGACGTCAGCCCCATCATGAACAA  420
 421 GCATTACAACATCTGCGCCGTCGTCGGCAACAGCGGCATCCTGACCGGCAGCCAGTGCGG  480
 481 CCAGGAGATCGACAAGAGCGACTTCGTCTTCCGCTGCAACTTCGCCCCCACCGAGGCCTT  540
 541 CCAGCGCGACGTCGGCCGCAAGACCAACCTGACCACCTTCAACCCCAGCATCCTGGAGAA  600
 601 GTACTACAACAACCTGCTGACCATCCAGGACCGCAACAACTTCTTCCTGAGCCTGAAGAA  660
 661 GCTGGACGGCGCCATCCTGTGGATTCCCGCCTTCTTCTTCCATACCAGCGCCACCGTCAC  720
 721 CCGCACCCTGGTCGACTTCTTCGTCGAGCATCGCGGCCAGCTGAAGGTCCAGCTGGCCTG  780
 781 GCCCGGCAACATCATGCAGCATGTCAACCGCTACTGGAAGAACAAGCATCTGAGCCCCAA  840
 841 GCGCCTGAGCACCGGCATCCTGATGTACACCCTGGCCAGCGCCATCTGCGAGGAGATCCA  900
 901 TCTGTACGGCTTCTGGCCCTTCGGCTTCGACCCCAACACCCGCGAGGACCTGCCCTACCA  960
 961 TTACTACGACAAGAAGGGCACCAAGTTCACCACCAAGTGGCAGGAGAGCCATCAGCTGCC 1020
1021 CGCCGAGTTCCAGCTGCTGTACCGTATGCATGGCGAGGGCCTGACCAAGCTGACCCTGAG 1080
1081 CCATTGCGCCTAGGCTAGC                                          1140
```

*Nhe*I

Fig 2-5: cDNA sequence of ST8Sia3 for cloning into the pCBH1corlin vector without mCherry (section 2.4.6). Double underlining indicates the nucleotide sequence of *Strep*-tag™ II, nucleotide sequence in bold letters indicates the sequence for short linker; nucleotide sequences with a yellow highlight indicate sites for restriction endonucleases. The nucleotide sequence in red font indicates a stop codon to stop translation. The red highlight indicates an extra nucleotide that had to be manually inserted to avoid any stop codons being introduced in the open reading frame (ORF). The truncated human ST8Sia3 is encoded by the nucleotide sequence indicated by green font and is comprised of 1041 base pairs.

2.4.6. Expression cassette for sialyltransferase gene expression

The cDNA constructs of sialyltransferases were obtained from Genscript (USA) in a pUC57 plasmid vector. The plasmid pCBH1corlin was engineered to have the human sialyltransferase cDNA and mCherry cDNA (encodes for a fluorescent reporter protein) inserts in the MCS 2 (Fig 2-6 A) in such a way that the 5′ end of the sialyltranferase cDNA was attached to the 3' end of the mCherry cDNA, so that subsequent translation will yield a fusion protein that can be tracked in the fungal hyphae. The mCherry protein further aids in the identification of any bottlenecks in the secretory pathway. A nucleotide sequence that provides a protease cleavage site by the enterokinase enzyme was also engineered between the mCherry and sialyltransferase cDNAs. This aids in cleaving the mCherry protein from the recombinant sialyltransferase. The mCherry cDNA was amplified from the mCherry plasmid (Clontech, USA, cat# 632522) using respective primers (Table 2-1). The presence of genes encoding sialyltransferases (ST3Gal3, ST6Gal1 and ST8Sia3) and mCherry in the plasmid vector pCBH1corlin were confirmed by PCR using respective primers (Table 2-1). Two plasmid constructs were essentially engineered for this project; one with mCherry cDNA and the other without. The plasmid vectors were digested using *Sph*I to generate linear DNA (expression cassette) for transformation into *T. reesei* (Fig 2-6). The linear expression cassette lacked the parts of the plasmid that are required for its propagation in *E. coli*.

2.4.7. Sequencing of plasmid DNA

Plasmid DNA sequencing was carried out to confirm that the cDNAs were in frame with the *cbh1* promoter. The samples were sent to Australian Genome Research Facility Ltd (AGRF), NSW, Australia and Macrogen, Korea for sequencing. The DNA sequences were analysed using the software Snapgene® and Geneious at Macquarie University, Sydney, Australia.

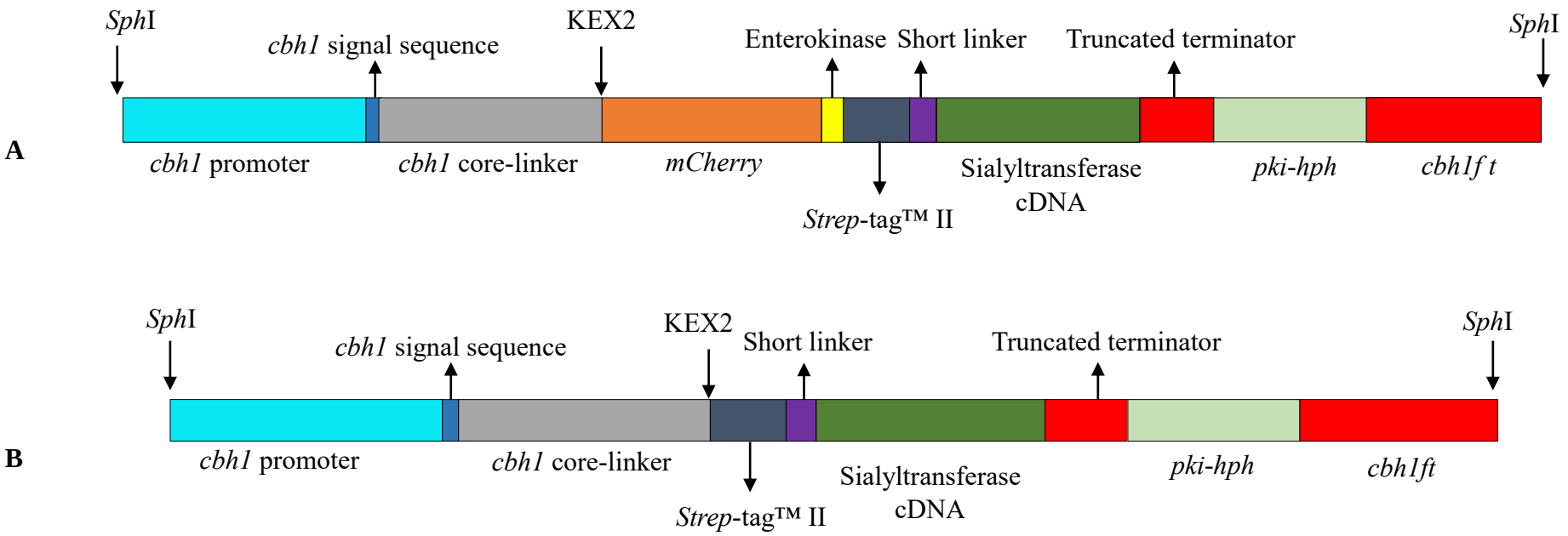


Fig 2-6: Schematic representation of the expression cassettes. **A:** Represents the expression cassette with mCherry. **B:** Represents the expression cassette without mCherry. Both expression cassettes consist of the basic components necessary for recombinant gene expression in *T. reesei*. Sialyltransferases are expressed under the *T. reesei cbh1* promoter. The *cbh1* signal sequence and the *cbh1* core-linker facilitates the secretion of heterologous proteins. *Strep*-tag™ II helps in the purification of recombinant silayltransferases. The *E. coli hph* gene codes for hygromycin phosphotransferase and conveys hygromycin (antibiotic) resistance to the transformants and is expressed under the *T. reesei pki* promoter. The *mCherry* gene codes for a fluorescent reporter protein. Enterokinase is a protease that aids in cleaving the recombinant sialyltransferase from mCherry. The *cbh1ft* represents the full sized *cbh1* terminator. The KEX2 like protease is expected to aid in the cleavage of the heterologous protein (mCherry and sialyltransferase) from the native CBHI protein carrier.

*Diagrams not according to scale

Table 2-1: Primers used in this study.

Primer	Sequence 5′ to 3′	Use	Expected PCR product size (bp)	Annealing temperature (°C)
hst6fwdpr	CTGTACAACGAGGGCATCCT	To confirm the presence of hST6Gal1 cDNA by PCR	156	56
hst6revpr	CTGGGGCTTCAGGATGTAGA			
hst3fwdpr	CACCAAGTACGCCAACTTCA	To confirm the presence of hST3Gal3 cDNA by PCR	193	56
hst3revpr	GATGGCCTTGATCAGGTTGT			
hst8fwdpr	CCGCTACTGGAAGAACAAGC	To confirm the presence of hST8Sia3 cDNA by PCR	246	56
hst8revpr	CCATGCATACGGTACAGCAG			
mCherry*fwdpr*	GGCGGCAAGCTTATACGTAATGGTGAGCAAG GGCGAGG	To synthesise mCherry cDNA and to confirm the presence of mCherry cDNA by PCR	760	60
mCherry*revpr*	GCGGCGGGATCCCTCGAGCTTGTCGTCGTCG TCCTTGTACAGCTCGTCCATGC			
hst6fwdpr_2	AGCAGCAGCAAGAACCTGAT	To synthesise probe for Southern blotting by PCR	600	56
hst6revpr	CTGGGGCTTCAGGATGTAGA			
corlin_fwd	TCAAGCAGCTGACTGAGATG	To sequence the plasmid constructs	-	55

2.5. Transformation of *Trichoderma reesei*

Transformation of *T. reesei* conidia was achieved by the biolistic bombardment procedure as previously described (Te'o *et al*., 2002). The single barrel system PDS-1000/He from Bio-Rad was used in this work. All solutions and chemical reagents used were freshly prepared before the experiment.

2.5.1. Preparation of conidia for transformation

Fresh PDA plates, with 20 mL PDA, were kept at 70ºC for 10 min to dry them and then allowed to cool to room temperature. A freshly harvested suspension of 1×10^7 fungal conidia (section 2.3.1) was spotted on at the centre of the dry PDA plate. The conidial solution on the PDA plates was then allowed to dry at room temperature for up to 30 min before biolistic bombardment was performed.

2.5.2. Preparation of gold microparticles for transformation

About 50 mg of gold particles, with a diameter of 0.6 µm (Inbio Gold), was weighed into a 1.5 mL Eppendorf tube. Particles were then washed with 1 mL absolute ethanol three times followed by 1 mL of sterile Milli-Q water. The tube was vortexed and centrifuged at each wash step. The particles were then resuspended in 1 mL of fresh sterile Milli-Q water before use.

At least 2.5 µg of linear expression cassette (section 2.4.6) was added to 50 µL of resuspended gold particles in a new Eppendorf tube and mixed by vortexing for 30 s. While vortexing, 50 µL of 2 M $CaCl_2$ solution and 20 µL of 0.1 M spermidine (Sigma Aldrich) were added into the tube followed by incubation on ice for at least 30 min. Further, the sample was centrifuged at 2500 × g for 10 s and the supernatant discarded. DNA coated microparticles were resuspended and washed using 500 µL of absolute ethanol by gentle mixing. The sample

was spun down and the supernatant was removed gently. The DNA coated gold particles were then resuspended in 60 µL absolute ethanol.

2.5.3. Biolistic bombardment

All components of the biolistic gene gun were sterilised with 70% (v/v) ethanol prior to use. Ten µL of the DNA coated microparticles containing at least 500 ng of DNA was pipetted onto a microcarrier and left to dry. Once dried, the microcarrier was loaded into a holder and assembled in the gene gun. A release pressure of 650 psi was used to propel the DNA coated microparticles. This was achieved by temporarily blocking the helium gas, using a rupture disk (Inbio Gold) that ruptures exactly at a force of 650 psi. PDA plates spotted with conidia were placed at a target distance of 3 cm facing up without the lid in a vacuum chamber at about 28" of Hg. DNA coated gold particles were bombarded towards the fungal conidia on the PDA plates.

2.5.4. Post bombardment

The plates were incubated at 28ºC for 2 h after which the fungal conidia were spread to cover the entire plate by adding 250 µL of 0.9% (w/v) NaCl, 0.01% (v/v) Tween 80, using a sterile spreader. The plates were further incubated at 28°C for 4 h. After incubation, the plates were overlain with 10 mL molten PDA (cooled to ~70ºC before adding hygromycin) with a sufficient amount of hygromycin B (A G Scientific) to give a final concentration of 80 µg/mL. Once the overlay PDA was set, the plates were left incubated at 28ºC until single colonies appeared.

2.5.5. Selection of transformant colonies

Distinct single colonies that appeared were carefully picked and re-plated on fresh PDA plates containing 80 µg/mL hygromycin B. Re-plated colonies were left to grow at 28ºC for 72-96 h. The colonies from the primary selection plates were then selected and re-plated a further two

times onto fresh PDA plates with 80 µg/mL hygromycin B. The colonies that thrived after the third selection were re-plated on to fresh PDA plates and left to incubate at 28ºC for conidiation. Conidia were harvested and stored at -80°C (section 2.2) for future experiments.

2.6. Extraction of intracellular proteins from fungal transformants

Fungal mycelia, harvested from liquid cultures after 24 h, 48 h and 96 h were washed once with PBS. The washed mycelia were resuspended in 2-3 mL protein extraction buffer (PEB) (50 mM Tris, 50 mM EDTA, 3% (w/v) SDS, 1 mM phenylmethylsulfonyl fluoride (PMSF), 10 mM DTT, pH 8). One Roche cOmplete mini EDTA free protease inhibitor cocktail tablet (Sigma Aldrich, USA) was dissolved in 900 µL of Milli-Q water and 100 µL of this solution was added to each sample prior to protein extraction. Intracellular protein was extracted using a Thermo Spectronic French press at 16 kpsi cell lysis pressure. The protein extract was centrifuged at 4ºC at 13, 000 × g to remove cell debris. Collected supernatant was stored at -30ºC.

2.7. Analysis of transformants for heterologous protein expression

The primary reason for having a fluorescent reporter protein like the mCherry was to monitor heterologous expression of a protein of interest and its secretion. Since mCherry cDNA codes for a fluorescent protein and the sialyltransferase (ST3Gal3 and ST6Gal1) is expected to be expressed in fusion to mCherry, the transformants were initially screened by monitoring the intracellular and extracellular fluorescence. The presence of sialyltransferase was then confirmed by Western blotting and mass spectrometry.

2.7.1. Live cell imaging to monitor the intracellular mCherry fluorescence

Live hyphae from the transformant cultures were observed under a confocal microscope for fluorescence. Hyphae from the PDA selection plates (containing hygromycin) of each transformant with the mCherry fusion protein and non-transformant RUT-C30 was inoculated

into 3 mL CLS medium and cultured at 28°C with constant shaking of 220 rpm for 72 h in the dark. Hyphae collected from the 3 mL CLS cultures were mounted on glass slides to be visualised. The Olympus Fluoview FV 1000 system (Olympus, Japan) was used for imaging. An excitation/emission wavelength of 559/612 nm was used to view the intracellular fluorescence. Images were viewed and manipulated using the Olympus Fluoview viewer (ver. 3.1a).

2.7.2. Measurement of secreted mCherry fluorescence

About 500 µL of the culture supernatant was collected from the 3 mL cultures and centrifuged at 13, 000 × g for 10 min to remove any hyphal debris. About 100 µL of the culture supernatant was carefully pipetted to 96 well microtiter plate (Greiner, USA). Fluorescence readings were taken using FluoStar Galaxy plate reader (BMG Lab technologies, Australia) at an excitation/emission wavelength of 560/630 nm (this was the closest matching wavelength filter available to detect mCherry fluorescence at the time of the experiment).

After the preliminary analysis of ST3Gal3 and ST6Gal1 transformants containing mCherry, it was decided to continue the project with the ST6Gal1 transformants with and without the mCherry reporter protein. This was to monitor the difference in behaviour of the two sets of transformants (with and without mCherry) during heterologous expression, and also to monitor for any variation in functionality of the recombinant ST6Gal1 when expressed in fusion with mCherry. All the following experiments were performed only with ST6Gal1 transformants.

2.7.3. One dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Total protein concentration was determined using Bradford method (Bradford, 1976). Bovine serum albumin was used as the standard for analysis. Measurement was performed using

biological triplicate in a 96 well microtiter plate (Greiner, USA). Absorbance was read in a FluoStar Galaxy plate reader (BMG Lab technologies, Australia) at 595 nm.

The culture supernatants were centrifuged at 13, 000 × g for 5 min to remove any cellular debris. Total protein concentration was normalised for SDS-PAGE based on the Bradford assay. The proteins were denatured by mixing 25 µL of protein sample with 5 µL of 4 × NuPAGE® LDS loading buffer and 1 µL 1M DTT and heating at 90°C for 10 min in a heating block in an Eppendorf tube. The samples were spun down before loading into NuPAGE® 4-12% (w/v) gradient gels for electrophoresis. The gels were run in 3-(N-Morpholino)propanesulfonic acid (MOPS) buffer for 50 min at a constant voltage of 200 V. All SDS-PAGE gels were run against 7 µL of Novex® protein ladder standard (Invitrogen, Australia) unless stated otherwise. The gel was then released from the holding cassette for Coomassie Colloidal Blue 250G staining or further used for Western blotting.

2.7.4. Confirmation of heterologous protein expression by Western blotting

The presence of mCherry and ST6Gal1 were confirmed by Western blotting using anti-mCherry and anti-ST6Gal1 antibodies respectively. Screening of non-mCherry transformants had to be performed by Western blotting using an anti-ST6Gal1 antibody. The detection techniques are explained below.

2.7.4.1. Detection based on fluorescently labelled secondary antibody (LI-COR® system)

Protein bands from the SDS-PAGE gels (section 2.7.3) were transferred to a low-fluorescence PVDF membrane (Bio-Rad, USA) for 15 min at 1.3 A and 25 V in Bio-Rad Trans-Blot® Turbo™ buffer using a Bio-Rad Trans-Blot® Turbo™ system (Bio-Rad, USA). Following the transfer, the membrane was blocked for 1 h at room temperature using PBS (Sigma Aldrich, Australia) buffer with 3% (w/v) skim milk and 0.05% (v/v) Tween 20 (BDH Chemicals Ltd., England) with gentle shaking. The membrane was incubated with a suitable

primary antibody (Table 2-2) diluted in 3% (w/v) skim milk in PBS buffer overnight at 4ºC with gentle shaking. The membrane was then washed three times for 10 min each with PBS-Tween 20 buffer at room temperature and incubated for an hour with the secondary antibody (Table 2-2), in PBS buffer with gentle shaking. The secondary antibody is tagged to an infra-red (IR) dye that enables detection using an IR scanner. After incubation with the secondary antibody, the membrane was washed three times for 5 min each in PBS-Tween 20 buffer at room temperature to remove the excess antibody. The washed membrane was then scanned using the LI-COR® Odyssey® infra-red imaging system. Recombinant human ST6Gal1 protein (Abnova, USA, cat# H00006480-P01) was used as positive control for antibody binding (unless stated otherwise).

Table 2-2: Antibodies used for Western blotting

Protein to be detected	Primary antibody	Dilution	Manufacturer	Origin	Secondary antibody	Dilution	Manufacturer
hST6Gal1	Anti-hST6Gal1 (#ARP64078_P050)	1:2,500	Aviva Systems Biology	Rabbit polyclonal	Anti-rabbit IgG	1:10,000	LI-COR®
mCherry	Anti-mCherry (#ab183628)	1:2,500	Abcam	Rabbit polyclonal	Anti-rabbit IgG	1:10,000	LI-COR®

i. The anti-human ST6Gal1 antibody is raised against PPSSGMLGIIIMMTLCDQVDIYEFLPSKRKTDVCYYYQKFFDSACTMGAY sequence located at the C-terminal catalytic domain.

ii. The anti-mCherry antibody is raised against an epitope within the mCherry amino acid sequence 1-236. The exact sequence remains proprietary.

2.7.4.2. *Strep*-Tactin® –AP based detection (for S*trep*-tag™ II detection)

Strep-tag™ II is an eight amino acid peptide sequence (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) having intrinsic affinity to streptavidin. *Strep*-Tactin® is an engineered form of streptavidin (Schmidt and Skerra, 2007). Proteins were transferred to a low-fluorescence PVDF membrane as discussed previously (section 2.7.4.1). The membrane was then blocked using 20 mL tris-buffered saline (TBS) with 3% (w/v) BSA and 0.1% (v/v) Tween 20 for overnight at 4ºC. After blocking, the membrane was washed three times using TBS buffer with 0.1% (v/v) Tween 20 for 5 min each at room temperature. After the last wash step, 2.5 µL of *Strep*-Tactin® alkaline phosphatase (*Strep*-Tactin®-AP) conjugate (IBA-Life Sciences, USA, Cat.no: 2-1503-001) dissolved in10 mL of TBS buffer with 0.1% (v/v) Tween 20 (1:4000 dilution) was added to the membrane and incubated for 60 min at room temperature with gentle shaking. The membrane was washed twice for one min each with TBS-Tween 20 buffer at room temperature and then twice with TBS buffer. The membrane was then incubated with BCIP/NBT substrate kit from Invitrogen (Cat.no: WP20001), following the manufacturer's protocol, for colour development. The chromogenic reaction was continued until the desired intensity of colour was achieved. A *Strep*-tag® protein ladder (IBA-Life Sciences, USA, Cat.no: 2-1011-100) was used as a positive control for detection.

2.7.5. Identification of proteins by mass spectrometry

Molecules are identified by mass spectrometer by comparing the mass to charge ratio of ions. Mass spectrometry was used to advantage in this project to confirm production of the recombinant human ST6Gal1 by the *T. reesei* transformants. Sample preparation procedure for mass spectrometry was based on the protocol described previously (Mirzaei *et al*., 2012).

2.7.5.1. De-staining of SDS-PAGE gel

Gel bands for mass spectrometry analysis were excised from 1D SDS PAGE gels stained with Coomassie Colloidal Blue 250G (section 2.7.3) on a clean glass plate and cut into small pieces of 1 mm^3 size. The pieces were transferred into a clean Eppendorf tube and excess water was removed using a micropipette. The gel pieces were then washed briefly with 100 mM NH_4HCO_3 to normalise the pH of the gels. Washed gel pieces were de-stained by adding 200 µL of 50% (v/v) acetonitrile (ACN) with 50% (v/v) 50mM NH_4HCO_3 and vortexing. The mixture was then incubated for 10 min. The process was repeated until most of the stain was removed. Excess wash buffer was removed after the final wash and 200 µL 100% (v/v) ACN was added to the gel pieces and incubated for 5 min with vortexing to dehydrate them. ACN was removed and the gel pieces were dried briefly without heat in a vacuum centrifuge.

2.7.5.2. Reduction and alkylation of in-gel proteins

The dried gel pieces were incubated for 60 min with 50 µL of 10 mM dithiothreitol (DTT) in 100 mM NH_4HCO_3 at 37ºC. The gel pieces were cooled to room temperature and the DTT-NH_4HCO_3 solution was removed. About 50 µL of 55 mM iodoacetamide (IAA) in 100 mM NH_4HCO_3 was added to cover the gel pieces. The tube was vortexed, centrifuged briefly and incubated in the dark for 45 min at room temperature. IAA was then removed and the gel pieces were washed with 100 mM NH_4HCO_3 followed with two times wash using 50% (v/v) ACN / 50 mM NH_4HCO_3 for 5 min each. Gel pieces were then dehydrated with 100% ACN. Excess ACN was removed and the gel pieces were dried in a vacuum centrifuge.

2.7.5.3. In-gel trypsin digestion of proteins

Trypsin was prepared fresh in 50 mM NH_4HCO_3 shortly before use with a final concentration of 12.5 ng/L. The solution was added to the dried gel pieces to cover them completely. The gel pieces were allowed to rehydrate with the trypsin buffer for at least 30 min at 4ºC. More

trypsin was added if the gel pieces absorbed the trypsin and the samples were centrifuged briefly. The tubes were then incubated overnight at 37ºC for digestion.

2.7.5.4. Extraction of trypsin digested peptides from the SDS-PAGE gel

The digestion mixture was transferred to a fresh Eppendorf tube after overnight incubation. A volume of 30 µL of 50% (v/v) ACN/ 2% (v/v) formic acid solution was added to the gel pieces and incubated for 20 min with vortexing. The tube was centrifuged and the supernatant removed and combined with the initial digest solution. This was repeated for at least two times to have a final extract volume of 60-100 µL. The extracted solution was vortexed, dried in a vacuum centrifuge and reconstituted in 1% (v/v) formic acid (FA) before loading into the mass spectrometer for analysis.

2.7.5.5. Electrospray ionisation mass spectrometry (ESI-MS/MS) based peptide detection

The method of detection has been previously described by Loke *et al*. (2015). The peptide mixtures were analysed by ESI-MS/MS using an HTC 3D ion trap (Bruker Daltonics) coupled to an Ultimate 3000 LC (Dionex, Australia). The sample was directly loaded on to a C18 column (Proteocol HQ303, 300 µm inner diameter x 10 cm length, 3 µm particle size, 300 Å pore size, SGE, Australia). The column was equilibrated using 100% (v/v) solvent A consisting of 0.1% (v/v) FA and a gradient up to 30% (v/v) (0.5% (v/v)/min slope) and a second gradient up to 60% (v/v) (4.2% (v/v)/min slope) of solvent B consisting of 0.1% (v/v) FA in ACN for 60 min and 7 min respectively, before washing the column in 80% (v/v) solvent B for 10 min and re-equilibration. A constant flow rate of 5 µL/injection was used. LC-MS/MS analysis (m/z 300-2200, scan speed: 8100 m/z/s) was followed by collision-induced dissociation (CID). The data were analysed using the Global Proteome Machine (GPM) software (http://www.thegpm.org/).

2.8. Cultivation and maintenance of the SKBR3 cell line

The SKBR3 human breast cancer cell line with enhanced expression of ST6Gal1 (Lee *et al.*, 2014) was obtained from American Type Culture Collection (Manassas, VA). The cells were maintained in RPMI 1640 medium (Life Technologies, Australia) supplemented with 10% (v/v) FBS (Fetal bovine serum, Life Technologies, Australia), 1 mM glutamine, (Life technologies, Australia), 100 U/mL penicillin and 100 µg/mL streptomycin in small cell culture flasks (Sigma-Aldrich, Australia) at 37ºC in a humidified atmosphere with 5% CO_2. Cells were grown to confluence and washed twice with PBS to remove excess medium. Subsequently, cells were separated from the culture flask by adding 0.25% (v/v) trypsin (Life technologies, Australia) and 0.02% (w/v) EDTA in PBS and incubating at 37ºC for 1-5 min. Cell count and viability were calculated with TC20™ Automated Cell Counter (Biorad, Australia) using 10 µL of cells mixed with 0.4% (v/v) Trypan blue in a 1:1 ratio. The cells were preserved in 500 µL aliquotes of freeze mix (FBS in 10% DMSO) at -130ºC in cryovials. Alternatively, cells were lysed by incubating in lysis buffer (50 mM MES buffer pH 6.5, 150 mM NaCl, 1% (v/v) Triton X100) at 4ºC for 1 h. The lysed cells were centrifuged at 4ºC at 6000 × g for 10 min to remove cell debris and the supernatant was stored at -30ºC to be used as positive control for Western blot analysis.

2.9. Expression of the mCherry protein in *E. coli*

A low endonuclease producing *E. coli* strain Origami™ B (DE3) (Novagen, USA) was used for the expression of mCherry protein. The Origami™ B cells were transformed with mCherry plasmid (Clontech, USA, cat# 632522) by heat shock transformation and were propagated as described in section 2.3.3. The propagated cells were pelleted by centrifugation at 6000 × g for 10 min. The pelleted cells were lysed by incubating in lysis buffer (50 mM MES buffer pH 6.5, 150 mM NaCl, 1% (v/v) Triton X100) at 4ºC for 1 h to release intracellular proteins. Cell debris was removed by centrifugation at 6000 × g at 4ºC for 10

min. The supernatant was collected and stored at -30ºC for future use as positive control in Western blotting for mCherry detection.

2.10. Total protease activity assay

Total extracellular protease activity in the *T. reesei* culture medium was measured, as described by Rauscher *et al.* (1995). *T. reesei* cultures (transformants and the non-transformant strain) were grown in triplicate for 120 h and pH was monitored every 24 h. The pH was measured at room temperature using Cyberscan pH510 (Eutech Instruments) in conjunction with Inlab®423 combination pH micro-electrode (Mettler Toledo, Switzerland).

Azocasein, at a final concentration of 5 mg/ml in 50 mM KH_2PO4, at a suitable pH (equal to the pH of the culture medium at a given time) was used as the substrate. To 100 µL of the culture supernatant, 150 µL of substrate solution was added and the sample was incubated at 28ºC overnight with continuous shaking in an Eppendorf tube. Subsequently, the reaction was stopped by adding 30 µL of 50% (w/v) trichloroacetic acid (TCA) in Milli-Q water. The tubes were centrifuged and the supernatant was removed to a new Eppendorf tube. An equal volume of 0.5 N NaOH was added into the tube. Consequently, 100 µL of this solution was pipetted into a 96 well microtiter plate (Greiner, USA) and the absorbance was measured at 440 nm. The assay was carried out in triplicate. The protease activity was defined as the amount of substrate (azocasein) digested (in mg) by the proteases present in 1 mg of the total secreted protein.

2.11. Isolation of fungal genomic DNA

Genomic DNA was extracted from the fungal mycelia based on the technique described previously by Lee *et al*. (1988). The entire procedure was performed in a fume hood. Fungal transformants were grown for seven days (until conidiation) on sterile cellophane discs placed on PDA plates. Hyphae were harvested by scraping off from the surface using a sterile

spatula. Freshly harvested hyphae were ground into a fine powder in liquid nitrogen using a sterile mortar and pestle. To 50-100 mg of ground hyphae, 750 µL of lysis buffer (50 mM Tris-HCl, 50 mM EDTA, 3% (w/v) SDS, 1% (v/v) 2-mercaptoethanol (added just before use)) were added and mixed thoroughly by vortexing. After an hour of incubating the mixture at 65ºC, 700 µL of chloroform: phenol (1:1) was added and mixed by inverting. The mixture was then centrifuged at 13, 000 × g for 15 min.

A volume of 600-650 µL of the aqueous (top) phase was removed to a new 2 mL phase lock gel tube (Eppendorf, Germany), carefully by avoiding any cellular debris at the interface. Further, 700 µL of SEVAG (chloroform: isoamyl alcohol, 24:1) was added to the aqueous phase and was mixed by inverting and centrifuged at 13,000 × g for 5 min. The aqueous phase was then removed to a new Eppendorf tube. Subsequently, 20 µL of 3 M sodium acetate and 1 mL isopropanol was added into the tube and mixed gently by inverting until ropes of DNA appeared. The tube was centrifuged at 13,000 × g for 30 s to pellet the DNA. The pelleted DNA was washed twice with 70% (v/v) ethanol. The tube was centrifuged and ethanol removed. DNA was briefly dried in a vacuum centrifuge and resuspended in fresh sterile Milli-Q water. DNA concentration was determined using NanoDrop 2000 (Thermo Scientific, USA). The quality of extracted genomic DNA was further analysed by electrophoresis on a 1% (w/v) agarose gel at 80V for 50 min.

2.12. Southern blotting to study the insertion of ST6Gal1 cDNA into the fungal genome

Southern blotting was performed to identify the number of copies of the hST6Gal1 cDNA inserted in the fungal genome, as well as to check for potential insertion of the hST6Gal1 cDNA into the *cbh1* locus. This can reveal any correlation between heterologous protein expression, the number of cDNA copies inserted into the fungal genomic DNA and the location of the cDNA insert. The procedure for Southern blotting was based on the technique described previously by Southern (2006).

2.12.1. Blotting of fungal DNA onto positively charged nylon membrane and hybridisation with a DNA probe

Detection of DNA by Southern blotting requires a nucleotide probe, usually radioactively labelled or DIG labelled as in this case. The probe for gene identification was synthesised by PCR amplification of the target fragment in ST6Gal1 cDNA using Roche DIG labelling mix following the manufacturer's protocol employing the respective primers mentioned in Table 2-1. The synthesised probe was 600 bp long. The DIG amplicons were confirmed by gel electrophoresis and were extracted and purified using Qiagen QIAquick gel extraction kit.

At least 10 µg of the fungal genomic DNA (section 2.11) was digested using *Eco*RV for overnight at 37ºC. The digested DNA was loaded on 0.8% (w/v) agarose gel and electrophoresed for 2 h at 90 V in 1 × TBE buffer. The gel was then soaked for 20 min in 0.5 µg/mL acridine orange in Milli-Q water to stain the DNA and subsequently de-stained in Milli-Q water for 30 min with gentle rocking. The de-stained gel was photographed with the ladder in view. The gel was then submerged twice in denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 30 min each. The gel was then transferred to a vacuum blotter, wherein the DNA was transferred on to a positively charged nylon membrane. The blotting was performed in 10 x SSC buffer (1.5 M NaCl, 0.15 M sodium citrate ($Na_3C_6H_5O_7$, pH-7)) for 2 h at a vacuum of 5 mm Hg. After blotting, the membrane was sandwiched between filter papers and baked at 120°C for 20 min to fix the DNA.

The fixed membrane was prepared for hybridisation with the aforementioned probe by incubating it with 20 mL hybridisation solution (prepared using DIG Easy Hybridisation granules from Roche) for one hour at 55ºC inside a glass cylinder with rotation. After one hour of incubation, the solution in the cylinder was discarded and 20 mL of fresh hybridisation solution mixed with 400 ng of the DIG labelled probe was added into the cylinder and incubated at 60ºC with rotation overnight. The solution was then removed and

the membrane was washed twice with 2 x SSC/ 0.1% (w/v) SDS for 5 min each and twice with 0.5 x SSC/ 0.1% (w/v) SDS for 15 min each at 60ºC.

Following hybridisation, the membrane was equilibrated with 20 mL washing buffer (100 mM maleic acid, 150 mM NaCl, 0.3% (v/v) Tween 20, pH-7.5, autoclaved) for one min. The washing buffer was discarded and the membrane was soaked with 20 mL of 1% (w/v) blocking solution (Roche, USA) with rotation at room temperature for 30 min. The membrane was then incubated with 20 mL of fresh 1% (w/v) blocking solution with 2 µL of anti-digoxigenin-AP (1: 10,000 dilution) at room temperature for 30 min with rotation. The blocking solution with anti-digoxigenin-AP was then discarded and washed twice with washing buffer for 15 min each.

2.12.2. Detection of hybridised probe using X-ray film

The hybridised membrane was carefully transferred to a clear plastic bag and equilibrated with 5 mL of detection buffer (100 mM tris-HCl, 100 mM NaCl, pH- 9.5, autoclaved). The buffer was tipped off from the bag completely before adding 5 mL of fresh detection buffer. To this three drops of CDP star (Roche, USA) was added and the membrane was washed thoroughly and the solution was removed completely. Subsequently, the membrane was exposed against CL-XPosure™ film (ThermoFisher, USA) for 15 min. The film was developed for one min in Kodak GBX developer/replenisher (Sigma-Aldrich, USA) following the manufacturer's protocol. Leaving the film in 1% (v/v) acetic acid for 30 s stopped the developing process. Once stopped, the film was fixed for another 30 s in Kodak GBX fixer /replenisher (Sigma-Aldrich, USA), following the manufacturer's protocol. The CL-XPosure™ film was handled in the dark room.

2.13. α2-6 sialyltransferase activity assay

To study the functionality of the recombinant ST6Gal1 enzyme, a lectin based assay was devised. The assay procedure was adapted and modified based on a technique previously described (Mattox *et al*., 1992). The principle of this assay technique is based on the selective affinity of *Sambuccus nigra* Agglutinin (SNA), a lectin derived from elderberry bark, to α-2,6 sialic acids transferred by ST6Gal1 (Shibuya *et al*., 1987). The assay technique is depicted in Fig 2-7.

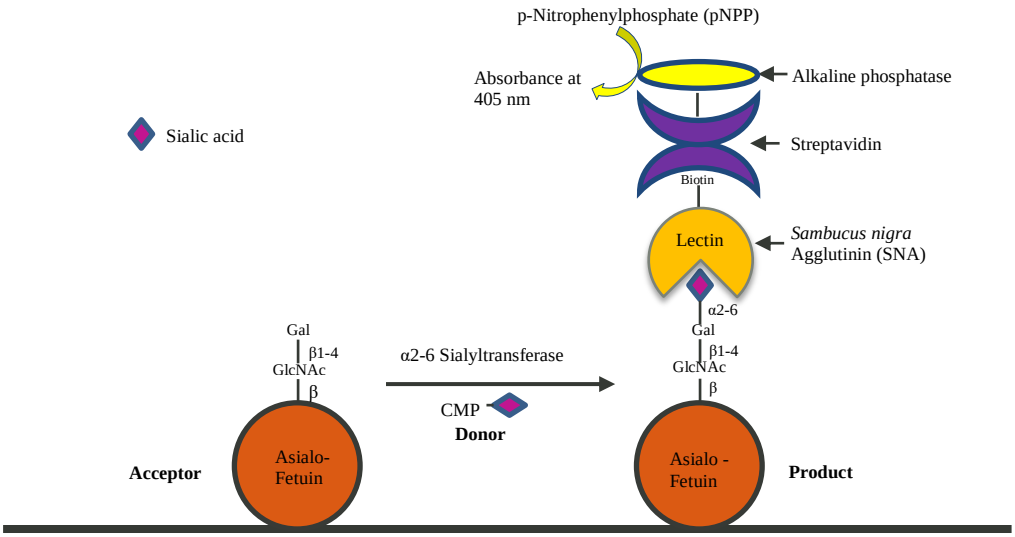


Fig 2-7: Schematic representation of the liquid-phase α2-6 sialyltransferase activity assay. Fetuin desialylated with a broad specific sialidase ABS™ (Glyko®) is used as the acceptor substrate. α2-6 sialyltransferase (ST6Gal1 in this case) transfers sialic acid (SA) residue on to the galactose (Gal) residue on the desialylated fetuin via an α2-6 linkage to which biotinylated plant lectin *Sambuccus nigra* agglutinin (SNA) specifically binds. The biotinylated SNA can then be detected with biotinylated streptavidin conjugated to alkaline phosphatase (AP). AP in the presence of p-Nitrophenylphosphate (pNPP) reacts to produce yellow colour, which can be measured at 405 nm.

2.13.1. Sample preparation for ST6Gal1 activity assay

Culture supernatant from *T. reesei* transformants was collected at 48 h of culturing in CLS medium. The supernatants were centrifuged at 3500 × g for 15 min at 4ºC. The clear supernatant was collected and desalted using a PD-10 desalting column (GE Healthcare, USA) and eluting the proteins using 10 mM HEPES and 50 mM NaCl (pH 7.0), following the

manufacturer's protocol. Desalting and protein elution were performed using chilled buffer on ice. The cleaned up samples were then concentrated up to 5× in a vacuum centrifuge and were stored in the freezer at -80ºC until used. The protein concentration of the processed culture supernatant was determined using Bradford assay.

2.13.2. Recombinant ST6Gal1 activity assay

Desialylated fetuin (asialofetuin) was used as an acceptor substrate for sialic acid. Bovine fetuin (Sigma) was desialylated by a broad specific sialidase, Glyko® Sialidase ABS™ (Prozyme, USA), following the manufacturer's protocol. A final concentration of 1 µg/mL acceptor substrates in PBS/TBS were immobilised on the surface of a 96 well high binding ELISA plate (Greiner, USA) for overnight at 4°C. Wells with immobilised fetuin (with sialylated glycans) served as a control for lectin (SNA) detection. The plates were then blocked using 1% (w/v) BSA in TBS and 0.05% (v/v) Tween 20 for one hour at room temperature with shaking. After an hour, the wells were washed with 200 µL of TBS and 0.05% (v/v) Tween 20 thrice with shaking for a minute.

A volume of 50 µL of 2× reaction buffer (100 mM MES pH 6.5, 40 mM $MnCl_2$) with CMP-sialic acid (donor substrate, 0.1 mM CMP-sialic acid) was added into each well that served as test reactions. Fifty µL of 2× reaction buffer without CMP-sialic acid was added to negative control reactions. Active recombinant human ST6Gal1 (R&D Systems, USA, cat no. 5924-GT) was used as a positive control for the reactions. The wells were topped up with 50 µL transformant culture supernatant (test reactions), from all the samples that were confirmed for the presence of recombinant ST6Gal1 by using Western blotting, active human ST6Gal1 (0.5 µg/100 µL of reaction mixture as positive control) or sterile Milli-Q water (for negative control) and were incubated for 5 h at 37ºC with shaking. After incubation, the reaction mixtures were discarded and the wells were washed thrice with 200 µL TBS and 0.05% (v/v) Tween 20.

Following the wash step, 100 µL of biotinylated SNA lectin (Vector Labs, USA) solution (10 µg/mL) in TBS was added into the wells and incubated at room temperature for one hour. SNA solution was discarded and the wells were washed with 200 µL of TBS and 0.05% (v/v) Tween 20 thrice with shaking. A volume of 100 µL of *Strep*-Tactin® alkaline phosphatase conjugate (IBA-Life Sciences, USA, Cat.no: 2-1503-001) at a dilution of 1:1000 in TBS was added into the well and the sample incubated at room temperature for 30 min with shaking. The wells were then washed thrice with 200 µL of TBS with 0.05% (v/v) Tween 20 with shaking. The wash solution was removed and 100 µL of alkaline phosphatase yellow (pNPP) liquid substrate (Sigma-Aldrich) was added into the wells and the samples were incubated at 37ºC for 15 min. The alkaline phosphatase reaction was stopped by adding 25 µL of 3 N NaOH and absorbance was read at 405nm with Fluostar Galaxy (BMG Labtech, Germany) plate reader.

2.13.3. Glycopeptide analysis by mass spectrometry

The results from the plate assay were validated using mass spectrometry (MS) analysis of glycans present on the immobilised acceptor substrates (fetuin and asialofetuin) on the wells. The alkaline phosphatase reaction mixture was removed from the wells and the wells were rinsed twice with 100 µL of PBS. To the wells, 70 µL of 1% (w/v) Sodium Deoxycholate (SDC) and 10 µL of 10 mM DTT were added and the plate was incubated at 37ºC for one hour. After an hour's incubation, 10 µL of 55 mM iodoacetamide (IAA) was added and the plate was incubated at room temperature for one hour in the dark. Subsequently, trypsin was added to a final concentration of 0.002 µg/µL into the wells and the plate was incubated at 37ºC to digest the acceptor substrates.

On the next day, the digested samples from the wells were removed to 1.5 mL Eppendorf tubes. SDC was removed by pelleting with 1% (v/v) formic acid and the pellet formed was removed by centrifuging at 14, 000 × g in a desktop centrifuge for 5 min. Supernatants were

collected, dried in a speedvac, and dissolved in 2% (v/v) ACN/0.1% (v/v) formic acid before the MS analysis. Samples were injected onto a C18 reverse phase (RP) trap chip using a NanoLC400 in combination with cHiPLC system (Eksigent, part of ABSCIEX) and separated on a 15 cm×200 µm analytical RP chip column with one hour of acetonitrile gradient run (5% (v/v)-40% (v/v)). Eluting peptides were analysed using a TripleTOF 5600 (ABSCIEX) mass spectrometer with a 0.1 s MS1 scanning followed by eight 0.1 s MS2 scanning targeting a fetuin glycopeptide (LCPDCPLLAPLNDSR) with different glycoforms. The results were imported into Skyline for quantitative analysis (MacLean *et al*., 2010).

2.14. Confocal microscopy to investigate co-localisation of mCherry-rST6Gal1 fusion protein in the ER

In order to investigate co-localisation of recombinant fusion proteins (mCherry-ST6Gal1) in the endoplasmic reticulum (ER), the fungal cell ER were stained and observed under a confocal microscope. ER-Tracker™ Green BODIPY® FL Glibenclamide (Life Technologies, Australia**)** was used to visualise the ER and Hoechst 33342 solution (ThermoFisher Scientific, USA) for staining the nuclei. For live cell staining, 100 uL of fresh fungal culture in CLS medium was removed into a 1.5 mL Eppendorf tube and centrifuged at 3000 rpm for 5 min followed by three times washing with PBS. The mycelia were resuspended in 100 µL of one µM ER-Tracker™ Green and incubated at room temperature for 30 min in dark. The staining solution was discarded and washed twice with PBS for 5 min. The washed mycelia were resuspended in 100 µL of Hoechst 33342 solution (1:2000 dilution in PBS) and incubated for 10 min in dark at room temperature. The stain solution was removed, washed with PBS and resuspended in 70 µL PBS. About 10 µL of the resuspended mycelial suspension was loaded on a clean glass slide. The mycelial suspension was covered using a clean cover slip before viewing under the Olympus Fluoview FV 1000 system (Olympus, Japan). An

excitation/emission wavelength of 559/619 nm was used to view mCherry fluorescence, 405/461 nm for Hoechst stain and 473/520 nm for ER Tracker ™ Green.

2.15. Extraction of fungal RNA

Fungal mycelia collected from a liquid culture in 50 mL CLS medium were washed three times with PBS in 0.1% (v/v) diethylpyrocarbonate (DEPC) water to remove the traces of the culture medium. The washed mycelia were ground into a fine powder using a mortar and pestle by adding liquid nitrogen. Care was taken not to thaw the powder. About 100 mg of the mycelial powder was mixed thoroughly with 1 ml of Trizol reagent (Sigma-Aldrich, US) followed by 5 min incubation at room temperature. The sample was subsequently centrifuged and the supernatant was removed to a new tube. Organic extraction was carried out by adding 200 µL of chloroform to the supernatant and incubating at room temperature for 3 min. Subsequently the sample was centrifuged at 12,000 × g at 4ºC for 15 min and the resulting aqueous phase was collected into a 1.5 mL Eppendorf tube. RNA was precipitated by adding 500 µL of isopropanol. The resulting RNA pellet was washed briefly with 75% (v/v) ethanol in 0.1% (v/v) DEPC water. Ethanol was removed by centrifugation and the pellet was dried and resuspended in 20 µL of 0.1% (v/v) DEPC water and was stored at -80ºC until used.

2.16. Quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) analysis was performed on a selection of genes to evaluate their transcription. Total RNA was extracted as described above from the *T. reesei* transformants with and without mCherry gene. All RNA samples were subjected to DNase digestion using RNase free DNase I (New England Biolabs, USA) and reverse transcribed to cDNA using AffinityScript qPCR cDNA Synthesis Kit (Agilent Technologies, USA) following the manufacturer's instructions. A standard 10 µL real time-PCR reaction mixture comprised of 5 µL of 2 x LightCycler® 480 SYBR Green I

Master (Roche, USA), 1 µL each of the forward and reverse primers at a final concentration of 1 µM, 2 µL of DNase free water and 1 µL of cDNA (20 ng/µL). All reactions were carried out in a Roche LightCycler® 480 instrument. The reaction conditions had an initial denaturation step at 95ºC for 10 min followed by 40 cycles of 95ºC for 10 s, annealing at 58ºC for 10 s and extension at 72ºC for 10 s. The primers used are listed in Table 2-3. The JGI Trire id represents the gene id derived from the *Trichoderma reesei* genome sequencing project (http://genome.jgi.doe.gov/Trire2/%20Trire2.home.html) (Martinez *et al*., 2008). The amplification efficiencies of primers were checked prior to commencing the experiment. Reactions with RNA only samples were also performed to check for the presence of gDNA contamination in the DNase treated samples and to confirm the quality of cDNA. The housekeeping gene actin was used an internal reference. The relative abundance of RNA was calculated by importing the cycle threshold (C_T) values into Microsoft Excel and using $2^{-(\Delta\Delta C_T)}$ method (Livak and Schmittgen, 2001).

Table 2-3: Primers used for qRT-PCR. The JGI Trire id indicates the corresponding *T. reesei* gene id in the JGI database (Martinez *et al*., 2008).

Primer name	Primer sequence 5' to 3'	Product size	Target gene function	JGI Trire id
Actin_fwd	GGTCGTACCACCGGTATCGT	107	Encodes actin	44504
Actin_rev	GCCATGTCAACACGAGCAAT			
BIP_fwd	TCTTCCCAGATGAGCACTCG	101	Encodes binding protein (BiP)	122920
BIP_rev	TCCATGTTGAGCTCCTCGAA			
PDI_fwd	GGTCTTGGCCAGGTTCTCAC	100	Encodes protein disulphide isomerase (PDI)	122415
PDI_rev	AGTTGGCCCTGAGACCTACG			
Cal_fwd	AAGGACGCAGAGGAGGAATG	107	Encodes Calnexin	73678
Cal_rev	GCAGGGTTCTTGACAACGAG			
Sec61_fwd	CGCTCCTCATCTTCCTGGTC	103	Encodes the Sec61 translocon on ER membrane	121397
Sec61_rev	TTACTCGCCATGACCATTCG			
UBA_fwd	TCACCATCATCGACCAGACC	108	Encodes ubiquitin-activating enzyme	72606
UBA_rev	CTTCCAAGCCGTGTCGAGTA			
RPN8_fwd	AGCAGCAGGACGAGAAGGAG	105	Encodes subunit 8 of 19S regulatory particle of the 26S proteasome	80843
RPN8_rev	CGCTGCTCTCCTTGGACTCT			
Beta3_fwd	AGCGACCTCTTCCGCTACAA	100	Encodes β-3subunit of 20S core particle of the 26S proteasome	58125
Beta3_rev	CGTAGAGGGAGGAGGAGACG			
ST6Gal1_fwd	CTGTACAACGAGGGCATCCT	156	Encodes ST6Gal1	-
ST6Gal1_rev	CTGGGGCTTCAGGATGTAGA			

3

PREPARATION OF EXPRESSION VECTORS, GENERATION AND CHARACTERISATION OF *TRICHODERMA REESEI* TRANSFORMANTS

3.1. Introduction

The availability of genetic manipulation techniques and the detailed knowledge of gene expression mechanisms have provided a good basis for heterologous expression of proteins of interest. The essence of recombinant protein production is the choice of a suitable vector for the delivery of a heterologous gene of interest into a host organism of choice. Especially important is the promoter that drives the expression.

The pCBH1corlin plasmid (Fig 2-1) was used as a base to construct the transformation vectors for this work. The pCBH1corlin vector was originally constructed based on the pHEN42 plasmid, which has been previously used for the successful expression of enzymatically active bacterial xylanase *xynB* in *T. reesei* (Te'o *et al*., 2000). The vector pCBH1corlin contains a *cbh1* promoter region followed by the *cbh1* signal sequence, the *cbh1* core-linker sequence, a truncated terminator, hygromycin B resistance gene (*pki-hph*) and a *cbh1* full terminator (*cbh1ft*) sequence. The *cbh1* promoter is contained in a 2.2 kb long nucleotide sequence and is strongly induced by cellulose and repressed in the presence of glucose (Ilmén *et al*., 1996).

The *cbh1* core-linker sequence encodes the *T. reesei* cellobiohydrolase 1 (CBHI) catalytic core and a linker. The *cbh1* core-linker is 460 amino acids long constitutive of a 436 amino acids long catalytic domain (the core) and a 24 amino acid linker. The *cbh1* signal sequence codes for a 17 amino acids long hydrophobic peptide sequence involved in the translocation of the protein into the ER (Fewell and Brodsky, 2009; Madhavan and Sukumaran, 2014). The truncated terminator is expected to stop the transcription of the recombinant DNA.

The *pki-hph* fragment allows for the selection of transformants on the antibiotic hygromycin B. The gene coding for bacterial hygromycin phosphotransferase, *hph*, is expressed under the

T. reesei pyruvate kinase promoter (*pki*). It is expected that the regions of the plasmid vector that correspond to the regions of the *T. reesei* genomic DNA such as the *cbh1* promoter and the terminator sequence will facilitate the homologous recombination of the sialyltransferase gene into the *cbh1* locus, thus replacing the endogenous *cbh1* gene. Further, it has been proposed that fusion of a heterologous gene to a native gene or a part of it can stabilise the resulting mRNA. Fusion can also facilitate translocation of the translated recombinant protein into the secretory pathway and impart protection from proteolysis (Gouka *et al*., 1997).

There is a multiple cloning site (MCS2) at the 3' end of the *cbh1* core-linker sequence, which was used to clone the genes coding for mCherry and sialyltransferases ST3Gal3, ST6Gal1 and ST8Sia3. The MCS2 contains cleavage sites for numerous restriction endonucleases; *Sna*BI, *Xho*I and *Nhe*I were utilised for cloning purposes in this project. Furthermore, the MCS2 also contains a sequence targeted by a KEX2 like protease that is expected to aid in the cleavage of heterologous protein from the native CBHI protein carrier (Te'o *et al*., 2000). The schematic of the expression vector and its components are depicted in Fig 2-1 (Chapter 2).

In this chapter, the method of construction of expression vectors containing cDNAs encoding the sialyltransferases ST3Gal3, ST6Gal1 and ST8Sia3 is described in detail. Two types of expression cassettes, with and without mCherry were constructed for transforming *T. reesei* (section 2.4.6, Fig 2-6). The mCherry is a fluorescent protein that is encoded by 230 amino acids. Moreover, the properties of mCherry such as high photostability, fast maturation and excellent pH resistance make it a suitable choice as a reporter protein (Shaner *et al*., 2004). The expression cassettes were used to transform *T. reesei* RUT-C30 conidia using a gene gun. The stable transformants were selected in PDA plates containing hygromycin B. The results of cloning, transformation, selection and preliminary screening of the transformants are discussed in detail.

3.2. Materials and methods

3.2.1. Insertion of mCherry gene into pCBHIcorlin vector

The *mCherry* cDNA was cloned into the MCS2 of the pCBH1corlin vector prior to cloning the sialyltransferase genes. The *mCherry* cDNA was synthesised from the mCherry plasmid using the primers given in Table 2-1 by polymerase chain reaction (PCR). The mCherry primers were designed in such a way that, two restriction sites namely the *SnaB*I and *Xho*I were introduced at the 5′ and 3′ ends respectively and an enterokinase cleavage site was introduced at the 3′ end of the *mCherry* cDNA (Fig 3-1). The *mCherry* cDNA carrying the aforementioned features was designed *in silico* using the software SnapGene® ver. 2.2.2 (from GSL Biotech; available at www.snapgene.com) in such a way that the cDNA is in frame with the *cbh1* core-linker sequence after cloning it into the MCS2.

The *mCherry* cDNA that was synthesised by PCR was subsequently purified using Qiagen QIAquick gel extraction kit following the manufacturer's protocol. Consequently, the *mCherry* cDNA and the pCBH1corlin vector were subjected to restriction digestion using *Sna*BI and *Xho*I restriction endonucleases to generate "sticky ends" to enable cloning. The restriction digested *mCherry* cDNA was then ligated into the MCS2 using T4 DNA Ligase (section 2.4.1).

*SnaB*I

```
   1  GGCGGCAAGCTTATACGTAATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAA   60
  61  GGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGAT  120
 121  CGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGAC  180
 181  CAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTC  240
 241  CAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGA  300
 301  GGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCA  360
 361  GGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTT  420
 421  CCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCG  480
 481  GATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGA  540
 541  CGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCT  600
 601  GCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACAC  660
 661  CATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCT  720
 721  GTACAAGGACGACGACGACAAGCTCGAGGGATCCCGCCGC                      760
```

*Xho*I

Fig 3-1: The cDNA sequence of mCherry for cloning into pCBH1corlin vector. Nucleotide sequences with a yellow highlight indicate sites for restriction endonucleases. The mCherry is encoded by the underlined nucleotide sequence and is comprised of 708 base pairs. The double underline indicates the enterokinase cleavage site.

Once the *mCherry* cDNA was ligated, the whole plasmid was transformed, selected and propagated in the *E. coli* DH5α cells (section 2.3.3). Colonies from the selection plates were subjected to colony PCR (section 2.4.3) to confirm the presence of *mCherry* cDNA. Those colonies that turned out positive for the presence of *mCherry* cDNA were further propagated overnight in LB medium to amplify the plasmid copy numbers. Plasmid DNA was extracted from the cells using Qiagen QIAprep spin miniprep kit according to the manufacturer's instructions. Fig 3-2A illustrates the pCBH1corlin vector containing the *mCherry* cDNA in the multiple cloning site 2 (MCS2).

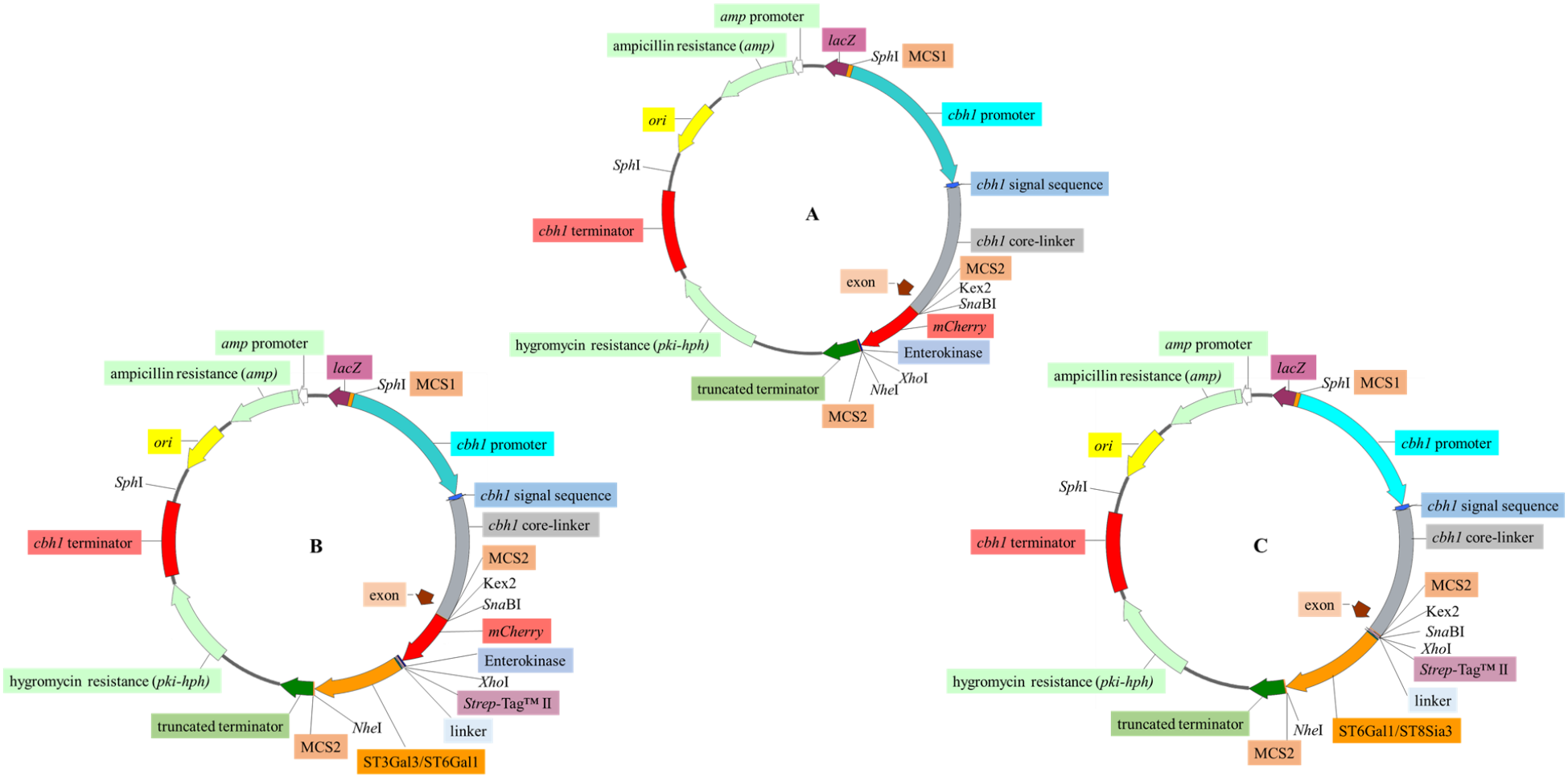


Fig 3-2: Representation of various plasmid constructs used for cloning. **A: S**chematic representation of the pCBH1corlin vector containing the *mCherry* gene cloned into the multiple cloning site 2 (MCS2) employing *SnaB*I and *Xho*I restriction sites. Enterokinase cleavage site was genetically engineered at the 3′ end of the *mCherry* gene followed by the *Xho*I restriction site. **B: S**chematic representation of the *ST3Gal3/ST6Gal1 and mCherry* genes cloned into the pCBH1corlin vector into the multiple cloning site 2 (MCS2). **C: S**chematic representation of the pCBH1corlin vector containing *ST6Gal1/ST8Sia3* cDNA without the *mCherry* gene cloned into the MCS2 (schematics of the plasmids were created using SnapGene®).

3.2.2. Construction of vectors containing genes encoding sialyltransferases ST3Gal3, ST6Gal1 and ST8Sia3

Further to cloning the *mCherry* cDNA into the MCS2, the sialyltransferase cDNAs coding for ST3Gal3 and ST6Gal1 were cloned into the pCBH1corlin plasmid. This was achieved by employing the restriction sites *Xho*I and *Nhe*I that flank the 5′ and 3′ ends respectively of the ST3Gal3 and ST6Gal1 cDNAs (section 2.4.5). The schematic of the final plasmid construct carrying the *mCherry* and sialyltransferase genes is given in Fig 3-2B.

Similarly, two vector constructs carrying the cDNAs encoding ST6Gal1 and ST8Sia3 without the *mCherry* gene were also engineered (Fig 3-2C). Detailed explanation for the construction of sialyltransferase cDNAs is given in section 2.4.5. The plasmid constructs were further amplified in *E. coli* DH5α. The nucleotide sequences were optimised for *Trichoderma reesei* codon usage as previously described (Te'o *et al*., 2000). It may be noted that the potential restriction endonuclease sites for *Sna*BI, *Xho*I, *Nhe*I and *Sph*I that were present within the open reading frames (ORF) of all the three sialyltransferase cDNAs were removed by adjusting the nucleotide bases on the basis of the preferred codon usage by *T. reesei*. Table 3-1 illustrates the comparison of codon usage in *T. reesei* and humans. Codon-optimised cDNA sequences encoding human sialyltransferases are given in Fig 3-3.

3.2.3. Confirmation of heterologous gene inserts in pCBH1corlin and sequencing of plasmid constructs

The sialyltransferase and *mCherry* gene inserts were further confirmed by PCR using the primers given in Table 2-1. Single transformant colonies were picked and were subjected to colony PCR prior to propagating them for plasmid extraction. The plasmids were extracted from the cells using Qiagen QIAprep spin miniprep kit according to the manufacturer's

instructions. The plasmids were sequenced as described in section 2.4.7. Sequencing of the entire plasmid DNA was not undertaken since the cDNAs were sourced from Genscript (USA) according to our specifications and occurrence of an error was considered highly unlikely.

Table 3-1: Codon usage in *Trichoderma reesei* and humans. Higher percentage indicates a higher priority in codon usage (adapted from Te'o *et al*., 2000) and Castro-Chavez (2011).

Amino acid	Codon	Amino acid codon usage in *T. reesei* (%)	Most used codons in humans per amino acid (%)	Amino acid	Codon	Amino acid codon usage in *T. reesei* (%)	Most used codons in humans per amino acid (%)	Amino acid	Codon	Amino acid codon usage in *T. reesei* (%)	Most used codons in humans per amino acid (%)
Gly	GGG	15		Ser	AGT	5		Leu	TTG	10	
	GGA	14			AGC	25	24		TTA	1	
	GGT	18			TCG	22			CTG	36	40
	GGC	60	34		TCA	8			CTA	2	
Glu	GAG	82	58		TCT	16			CTT	14	
	GAA	18			TCC	25			CTC	38	
Asp	GAT	31		Asn	AAT	19		Gln	CAG	81	73
	GAC	69	54		AAC	81	53		CAA	19	
Val	GTG	23	46	Met	ATG	100	100	His	CAT	17	
	GTA	4		Ile	ATA	4			CAC	83	58
	GTT	18			ATT	33		Pro	CCG	22	
	GTC	56			ATC	64	47		CCA	12	
Ala	GCG	17		Thr	ACG	29			CCT	25	
	GCA	12			ACA	10			CCC	40	32
	GCT	22			ACT	20		Phe	TTT	38	
	GCC	49	40		ACC	41	36		TTC	62	54
Arg	AGG	13		STOP	TGA	19	47	Tyr	TAT	25	
	AGA	7	21		TAG	44			TAC	75	56
	CGG	12			TAA	37		Lys	AAG	92	57
	CGA	16		Cys	TGT	24			AAA	8	
	CGT	17			TGC	76	54	Trp	TGG	100	100
	CGC	35									

```
      K  L  H  L  L  Q  W  E  E  D  S  N  S  V  V  L  S  F  D  S
    1 AAGCTGCATCTGCTGCAGTGGGAGGAGGACAGCAACAGCGTCGTCCTGAGCTTCGACAGC   60
      AAGCTGCACCTGCTGCAGTGGGAGGAGGACAGCAACAGCGTGGTGCTGAGCTTCGACAGC
      A  G  Q  T  L  G  S  E  Y  D  R  L  G  F  L  L  N  L  D  S
   61 GCCGGCCAGACCCTGGGCAGCGAGTACGACCGCCTGGGCTTCCTGCTGAACCTGGACAGC  120
      GCCGGCCAGACCCTGGGCAGCGAGTACGACAGACTGGGCTTCCTGCTGAACCTGGACAGC
      K  L  P  A  E  L  A  T  K  Y  A  N  F  S  E  G  A  C  K  P
  121 AAGCTGCCCGCCGAGCTGGCCACCAAGTACGCCAACTTCAGCGAGGGCGCCTGCAAGCCC  180
      AAGCTGCCCGCCGAGCTGGCCACCAAGTACGCCAACTTCAGCGAGGGCGCCTGCAAGCCC
      G  Y  A  S  A  L  M  T  A  I  F  P  R  F  S  K  P  A  P  M
  181 GGCTACGCCAGCGCCCTGATGACCGCCATCTTCCCCCGCTTCAGCAAGCCCGCCCCCATG  240
      GGCTACGCCAGCGCCCTGATGACCGCCATCTTCCCCAGATTCAGCAAGCCCGCCCCCATG
      F  L  D  D  S  F  R  K  W  A  R  I  R  E  F  V  P  P  F  G
  241 TTCCTGGACGACAGCTTCCGCAAGTGGGCCCGCATCCGCGAGTTCGTCCCCCCCTTCGGC  300
      TTCCTGGACGACAGCTTCAGAAAGTGGGCCAGAATCAGAGAGTTCGTGCCCCCCTTCGGC
      I  K  G  Q  D  N  L  I  K  A  I  L  S  V  T  K  E  Y  R  L
  301 ATCAAGGGCCAGGACAACCTGATCAAGGCCATCCTGAGCGTCACCAAGGAGTACCGCCTG  360
      ATCAAGGGCCAGGACAACCTGATCAAGGCCATCCTGAGCGTGACCAAGGAGTACAGACTG
      T  P  A  L  D  S  L  R  C  R  R  C  I  I  V  G  N  G  G  V
  361 ACCCCCGCCCTGGACAGCCTGCGCTGCCGCCGCTGCATCATCGTCGGCAACGGCGGCGTC  420
      ACCCCCGCCCTGGACAGCCTGAGATGCAGAAGATGCATCATCGTGGGCAACGGCGGCGTG
      L  A  N  K  S  L  G  S  R  I  D  D  Y  D  I  V  V  R  L  N
  421 CTGGCCAACAAGAGCCTGGGCAGCCGCATCGACGACTACGACATCGTCGTCCGCCTGAAC  480
      CTGGCCAACAAGAGCCTGGGCAGCAGAATCGACGACTACGACATCGTGGTGAGACTGAAC
      S  A  P  V  K  G  F  E  K  D  V  G  S  K  T  T  L  R  I  T
  481 AGCGCCCCCGTCAAGGGCTTCGAGAAGGACGTCGGCAGCAAGACCACCCTGCGCATCACC  540
      AGCGCCCCCGTGAAGGGCTTCGAGAAGGACGTGGGCAGCAAGACCACCCTGAGAATCACC
      Y  P  E  G  A  M  Q  R  P  E  Q  Y  E  R  D  S  L  F  V  L
  541 TACCCCGAGGGCGCCATGCAGCGCCCCGAGCAGTACGAGCGCGACAGCCTGTTCGTCCTG  600
      TACCCCGAGGGCGCCATGCAGAGACCCGAGCAGTACGAGAGAGACAGCCTGTTCGTGCTG
      A  G  F  K  W  Q  D  F  K  W  L  K  Y  I  V  Y  K  E  R  V
  601 GCCGGCTTCAAGTGGCAGGACTTCAAGTGGCTGAAGTACATCGTCTACAAGGAGCGCGTC  660
      GCCGGCTTCAAGTGGCAGGACTTCAAGTGGCTGAAGTACATCGTGTACAAGGAGAGAGTG
      S  A  S  D  G  F  W  K  S  V  A  T  R  V  P  K  E  P  P  E
  661 AGCGCCAGCGACGGCTTCTGGAAGAGCGTCGCCACCCGCGTCCCCAAGGAGCCCCCCGAG  720
      AGCGCCAGCGACGGCTTCTGGAAGAGCGTGGCCACCAGAGTGCCCAAGGAGCCCCCCGAG
      I  R  I  L  N  P  Y  F  I  Q  E  A  A  F  T  L  I  G  L  P
  721 ATCCGCATCCTGAACCCCTACTTCATCCAGGAGGCCGCCTTCACCCTGATCGGCCTGCCC  780
      ATCAGAATCCTGAACCCCTACTTCATCCAGGAGGCCGCCTTCACCCTGATCGGCCTGCCC
      F  N  N  G  L  M  G  R  G  N  I  P  T  L  G  S  V  A  V  T
  781 TTCAACAACGGCCTGATGGGCCGCGGCAACATCCCCACCCTGGGCAGCGTCGCCGTCACC  840
      TTCAACAACGGCCTGATGGGCAGAGGCAACATCCCCACCCTGGGCAGCGTGGCCGTGACC
      M  A  L  H  G  C  D  E  V  A  V  A  G  F  G  Y  D  M  S  T
  841 ATGGCCCTGCATGGCTGCGACGAGGTCGCCGTCGCCGGCTTCGGCTACGACATGAGCACC  900
      ATGGCCCTGCACGGCTGCGACGAGGTGGCCGTGGCCGGCTTCGGCTACGACATGAGCACC
      P  N  A  P  L  H  Y  Y  E  T  V  R  M  A  A  I  K  E  S  W
  901 CCCAACGCCCCCCTGCATTACTACGAGACCGTCCGCATGGCCGCCATCAAGGAGAGCTGG  960
      CCCAACGCCCCCCTGCACTACTACGAGACCGTGAGAATGGCCGCCATCAAGGAGAGCTGG
      T  H  N  I  Q  R  E  K  E  F  L  R  K  L  V  K  A  R  V  I
  961 ACCCATAACATCCAGCGCGAGAAGGAGTTCCTGCGCAAGCTGGTCAAGGCCCGCGTCATC 1020
      ACCCACAACATCCAGAGAGAGAAGGAGTTCCTGAGAAAGCTGGTGAAGGCCAGAGTGATC
      T  D  L  S  S  G  I
1021 ACCGACCTGAGCAGCGGCATC                                         1041
      ACCGACCTGAGCAGCGGCATC
```

Fig 3-3 A: The human ST3Gal3 cDNA sequence before codon optimisation (indicated by red letters) and after codon optimisation (indicated by green letters) for expression in *Trichoderma reesei*. The corresponding amino acid sequence is given in black letters.

```
      K  E  K  K  K  G  S  Y  Y  D  S  F  K  L  Q  T  K  E  F  Q
    1 AAGGAGAAGAAGAAGGGCAGCTACTACGACAGCTTCAAGCTGCAGACCAAGGAGTTCCAG   60
      AAGGAGAAGAAGAAGGGCAGCTACTACGACAGCTTCAAGCTGCAGACCAAGGAGTTCCAG
      V  L  K  S  L  G  K  L  A  M  G  S  D  S  Q  S  V  S  S  S
   61 GTCCTGAAGAGCCTGGGCAAGCTGGCCATGGGCAGCGACAGCCAGAGCGTCAGCAGCAGC  120
      GTGCTGAAGAGCCTGGGCAAGCTGGCCATGGGCAGCGACAGCCAGAGCGTGAGCAGCAGC
      S  T  Q  D  P  H  R  G  R  Q  T  L  G  S  L  R  G  L  A  K
  121 AGCACCCAGGACCCCCATCGCGGCCGCCAGACCCTGGGCAGCCTGCGCGGCCTGGCCAAG  180
      AGCACCCAGGACCCCCACAGAGGCAGACAGACCCTGGGCAGCCTGAGAGGCCTGGCCAAG
      A  K  P  E  A  S  F  Q  V  W  N  K  D  S  S  S  K  N  L  I
  181 GCCAAGCCCGAGGCCAGCTTCCAGGTCTGGAACAAGGACAGCAGCAGCAAGAACCTGATC  240
      GCCAAGCCCGAGGCCAGCTTCCAGGTGTGGAACAAGGACAGCAGCAGCAAGAACCTGATC
      P  R  L  Q  K  I  W  K  N  Y  L  S  M  N  K  Y  K  V  S  Y
  241 CCCCGCCTGCAGAAGATCTGGAAGAACTACCTGAGCATGAACAAGTACAAGGTCAGCTAC  300
      CCCAGACTGCAGAAGATCTGGAAGAACTACCTGAGCATGAACAAGTACAAGGTGAGCTAC
      K  G  P  G  P  G  I  K  F  S  A  E  A  L  R  C  H  L  R  D
  301 AAGGGCCCCGGCCCCGGCATCAAGTTCAGCGCCGAGGCCCTGCGCTGCCATCTGCGCGAC  360
      AAGGGCCCCGGCCCCGGCATCAAGTTCAGCGCCGAGGCCCTGAGATGCCACCTGAGAGAC
      H  V  N  V  S  M  V  E  V  T  D  F  P  F  N  T  S  E  W  E
  361 CATGTCAACGTCAGCATGGTCGAGGTCACCGACTTCCCCTTCAACACCAGCGAGTGGGAG  420
      CACGTGAACGTGAGCATGGTGGAGGTGACCGACTTCCCCTTCAACACCAGCGAGTGGGAG
      G  Y  L  P  K  E  S  I  R  T  K  A  G  P  W  G  R  C  A  V
  421 GGCTACCTGCCCAAGGAGAGCATCCGCACCAAGGCCGGCCCCTGGGGCCGCTGCGCCGTC  480
      GGCTACCTGCCCAAGGAGAGCATCAGAACCAAGGCCGGCCCCTGGGGCAGATGCGCCGTG
      V  S  S  A  G  S  L  K  S  S  Q  L  G  R  E  I  D  D  H  D
  481 GTCAGCAGCGCCGGCAGCCTGAAGAGCAGCCAGCTGGGCCGCGAGATCGACGACCATGAC  540
      GTGAGCAGCGCCGGCAGCCTGAAGAGCAGCCAGCTGGGCAGAGAGATCGACGACCACGAC
      A  V  L  R  F  N  G  A  P  T  A  N  F  Q  Q  D  V  G  T  K
  541 GCCGTCCTGCGCTTCAACGGCGCCCCCACCGCCAACTTCCAGCAGGACGTCGGCACCAAG  600
      GCCGTGCTGAGATTCAACGGCGCCCCCACCGCCAACTTCCAGCAGGACGTGGGCACCAAG
      T  T  I  R  L  M  N  S  Q  L  V  T  T  E  K  R  F  L  K  D
  601 ACCACCATCCGCCTGATGAACAGCCAGCTGGTCACCACCGAGAAGCGCTTCCTGAAGGAC  660
      ACCACCATCAGACTGATGAACAGCCAGCTGGTGACCACCGAGAAGAGATTCCTGAAGGAC
      S  L  Y  N  E  G  I  L  I  V  W  D  P  S  V  Y  H  S  D  I
  661 AGCCTGTACAACGAGGGCATCCTGATCGTCTGGGACCCCAGCGTCTACCATAGCGACATC  720
      AGCCTGTACAACGAGGGCATCCTGATCGTGTGGGACCCCAGCGTGTACCACAGCGACATC
      P  K  W  Y  Q  N  P  D  Y  N  F  F  N  N  Y  K  T  Y  R  K
  721 CCCAAGTGGTACCAGAACCCCGACTACAACTTCTTCAACAACTACAAGACCTACAGAAAG  780
      CCCAAGTGGTACCAGAACCCCGACTACAACTTCTTCAACAACTACAAGACCTACAGAAAG
      L  H  P  N  Q  P  F  Y  I  L  K  P  Q  M  P  W  E  L  W  D
  781 CTGCATCCCAACCAGCCCTTCTACATCCTGAAGCCCCAGATGCCCTGGGAGCTGTGGGAC  840
      CTGCACCCCAACCAGCCCTTCTACATCCTGAAGCCCCAGATGCCCTGGGAGCTGTGGGAC
      I  L  Q  E  I  S  P  E  E  I  Q  P  N  P  P  S  S  G  M  L
  841 ATCCTGCAGGAGATCAGCCCCGAGGAGATCCAGCCCAACCCCCCCAGCAGCGGTATGCTG  900
      ATCCTGCAGGAGATCAGCCCCGAGGAGATCCAGCCCAACCCCCCCCAGCAGCGGCATGCTG
      G  I  I  I  M  M  T  L  C  D  Q  V  D  I  Y  E  F  L  P  S
  901 GGCATCATCATCATGATGACCCTGTGCGACCAGGTCGACATCTACGAGTTCCTGCCCAGC  960
      GGCATCATCATCATGATGACCCTGTGCGACCAGGTGGACATCTACGAGTTCCTGCCCAGC
      K  R  K  T  D  V  C  Y  Y  Y  Q  K  F  F  D  S  A  C  T  M
  961 AAGCGCAAGACCGACGTCTGCTACTACTACCAGAAGTTCTTCGACAGCGCCTGCACCATG 1020
      AAGAGAAAGACCGACGTGTGCTACTACTACCAGAAGTTCTTCGACAGCGCCTGCACCATG
      G  A  Y  H  P  L  L  Y  E  K  N  L  V  K  H  L  N  Q  G  T
 1021 GGCGCCTACCATCCCCTGCTGTACGAGAAGAACCTGGTCAAGCATCTGAACCAGGGCACC 1080
      GGCGCCTACCACCCCCTGCTGTACGAGAAGAACCTGGTGAAGCACCTGAACCAGGGCACC
      D  E  D  I  Y  L  L  G  K  A  T  L  P  G  F  R  T  I  H  C
 1081 GACGAGGACATCTACCTGCTGGGCAAGGCCACCCTGCCCGGCTTCCGCACCATCCATTGC 1140
      GACGAGGACATCTACCTGCTGGGCAAGGCCACCCTGCCCGGCTTCAGAACCATCCACTGC
```

Fig 3-3 B: The human ST6Gal1 cDNA sequence before codon optimisation (indicated by red letters) and after codon optimisation (indicated by green letters) for expression in *Trichoderma reesei*. The corresponding amino acid sequence is given in black letters.

```
      K  K  E  N  I  F  T  T  P  K  Y  A  S  P  G  A  P  R  M  Y
    1 AAGAAGGAGAACATCTTCACCACCCCCAAGTACGCCAGCCCCGGCGCCCCCCGCATGTAC  60
      AAGAAGGAGAACATCTTCACCACCCCCAAGTACGCCAGCCCCGGCGCCCCCAGAATGTAC
      M  F  H  A  G  F  R  S  Q  F  A  L  K  F  L  D  P  S  F  V
   61 ATGTTCCATGCCGGCTTCCGCAGCCAGTTCGCCCTGAAGTTCCTGGACCCCAGCTTCGTC 120
      ATGTTCCACGCCGGCTTCAGAAGCCAGTTCGCCCTGAAGTTCCTGGACCCCAGCTTCGTG
      P  I  T  N  S  L  T  Q  E  L  Q  E  K  P  S  K  W  K  F  N
  121 CCCATCACCAACAGCCTGACCCAGGAGCTGCAGGAGAAGCCCAGCAAGTGGAAGTTCAAC 180
      CCCATCACCAACAGCCTGACCCAGGAGCTGCAGGAGAAGCCCAGCAAGTGGAAGTTCAAC
      R  T  A  F  L  H  Q  R  Q  E  I  L  Q  H  V  D  V  I  K  N
  181 CGCACCGCCTTCCTGCATCAGCGCCAGGAGATCCTGCAGCATGTCGACGTCATCAAGAAC 240
      AGAACCGCCTTCCTGCACCAGAGACAGGAGATCCTGCAGCACGTGGACGTGATCAAGAAC
      F  S  L  T  K  N  S  V  R  I  G  Q  L  M  H  Y  D  Y  S  S
  241 TTCAGCCTGACCAAGAACAGCGTCCGCATCGGCCAGCTGATGCATTACGACTACAGCAGC 300
      TTCAGCCTGACCAAGAACAGCGTGAGAATCGGCCAGCTGATGCACTACGACTACAGCAGC
      H  K  Y  V  F  S  I  S  N  N  F  R  S  L  L  P  D  V  S  P
  301 CATAAGTACGTCTTCAGCATCAGCAACAACTTCCGCAGCCTGCTGCCCGACGTCAGCCCC 360
      CACAAGTACGTGTTCAGCATCAGCAACAACTTCAGAAGCCTGCTGCCCGACGTGAGCCCC
      I  M  N  K  H  Y  N  I  C  A  V  V  G  N  S  G  I  L  T  G
  361 ATCATGAACAAGCATTACAACATCTGCGCCGTCGTCGGCAACAGCGGCATCCTGACCGGC 420
      ATCATGAACAAGCACTACAACATCTGCGCCGTGGTGGGCAACAGCGGCATCCTGACCGGC
      S  Q  C  G  Q  E  I  D  K  S  D  F  V  F  R  C  N  F  A  P
  421 AGCCAGTGCGGCCAGGAGATCGACAAGAGCGACTTCGTCTTCCGCTGCAACTTCGCCCCC 480
      AGCCAGTGCGGCCAGGAGATCGACAAGAGCGACTTCGTGTTCAGATGCAACTTCGCCCCC
      T  E  A  F  Q  R  D  V  G  R  K  T  N  L  T  T  F  N  P  S
  481 ACCGAGGCCTTCCAGCGCGACGTCGGCCGCAAGACCAACCTGACCACCTTCAACCCCAGC 540
      ACCGAGGCCTTCCAGAGAGACGTGGGCAGAAAGACCAACCTGACCACCTTCAACCCCAGC
      I  L  E  K  Y  Y  N  N  L  L  T  I  Q  D  R  N  N  F  F  L
  541 ATCCTGGAGAAGTACTACAACAACCTGCTGACCATCCAGGACCGCAACAACTTCTTCCTG 600
      ATCCTGGAGAAGTACTACAACAACCTGCTGACCATCCAGGACAGAAACAACTTCTTCCTG
      S  L  K  K  L  D  G  A  I  L  W  I  P  A  F  F  F  H  T  S
  601 AGCCTGAAGAAGCTGGACGGCGCCATCCTGTGGATTCCCGCCTTCTTCTTCCATACCAGC 660
      AGCCTGAAGAAGCTGGACGGCGCCATCCTGTGGATCCCCGCCTTCTTCTTCCACACCAGC
      A  T  V  T  R  T  L  V  D  F  F  V  E  H  R  G  Q  L  K  V
  661 GCCACCGTCACCCGCACCCTGGTCGACTTCTTCGTCGAGCATCGCGGCCAGCTGAAGGTC 720
      GCCACCGTGACCAGAACCCTGGTGGACTTCTTCGTGGAGCACAGAGGCCAGCTGAAGGTG
      Q  L  A  W  P  G  N  I  M  Q  H  V  N  R  Y  W  K  N  K  H
  721 CAGCTGGCCTGGCCCGGCAACATCATGCAGCATGTCAACCGCTACTGGAAGAACAAGCAT 780
      CAGCTGGCCTGGCCCGGCAACATCATGCAGCACGTGAACAGATACTGGAAGAACAAGCAC
      L  S  P  K  R  L  S  T  G  I  L  M  Y  T  L  A  S  A  I  C
  781 CTGAGCCCCAAGCGCCTGAGCACCGGCATCCTGATGTACACCCTGGCCAGCGCCATCTGC 840
      CTGAGCCCCAAGAGACTGAGCACCGGCATCCTGATGTACACCCTGGCCAGCGCCATCTGC
      E  E  I  H  L  Y  G  F  W  P  F  G  F  D  P  N  T  R  E  D
  841 GAGGAGATCCATCTGTACGGCTTCTGGCCCTTCGGCTTCGACCCCAACACCCGCGAGGAC 900
      GAGGAGATCCACCTGTACGGCTTCTGGCCCTTCGGCTTCGACCCCAACACCAGAGAGGAC
      L  P  Y  H  Y  Y  D  K  K  G  T  K  F  T  T  K  W  Q  E  S
  901 CTGCCCTACCATTACTACGACAAGAAGGGCACCAAGTTCACCACCAAGTGGCAGGAGAGC 960
      CTGCCCTACCACTACTACGACAAGAAGGGCACCAAGTTCACCACCAAGTGGCAGGAGAGC
      H  Q  L  P  A  E  F  Q  L  L  Y  R  M  H  G  E  G  L  T  K
  961 CATCAGCTGCCCGCCGAGTTCCAGCTGCTGTACCGTATGCATGGCGAGGGCCTGACCAAG 1020
      CACCAGCTGCCCGCCGAGTTCCAGCTGCTGTACAGAATGCACGGCGAGGGCCTGACCAAG
      L  T  L  S  H  C  A
 1021 CTGACCCTGAGCCATTGCGCC                                        1041
      CTGACCCTGAGCCACTGCGCC
```

Fig 3-3 C: The human ST8Sia3 cDNA sequence before codon optimisation (indicated by red letters) and after codon optimisation (indicated by green letters) for expression in *Trichoderma reesei*. The corresponding amino acid sequence is given in black letters.

3.2.4. Preparation of recombinant plasmids for transformation of *Trichoderma reesei*

The sequenced plasmids were linearised by restriction digestion with *Sph*I restriction endonuclease prior to transformation. The digested DNA fragments containing the expression cassettes were separated by gel electrophoresis and subsequently extracted and purified using Qiagen QIAquick gel extraction kit, prior to transformation. The *T. reesei* conidia were transformed by biolistic bombardment (section 2.5).

3.3. Results and discussion

3.3.1. Preparation of plasmid vectors containing mCherry and sialyltransferase cDNAs

The pUC57 plasmids containing the sialyltransferase coding DNAs (sourced from Genscript, USA) were propagated and amplified in *E. coli* DH5α cells and subsequently purified using Qiagen plasmid purification kit. The extracted pUC57 plasmids containing the sialyltransferase coding DNAs were run in 1% agarose gels for quality check and then subjected to PCR amplification (section 2.4.3) using the primers given in Table 2-1 and the PCR conditions were optimised to get the desired results. The optimum annealing temperatures for the primers were determined by running a gradient PCR. The expected sizes of the PCR amplicons were 193 bp for ST3Gal3, 156 bp for ST6Gal1 and 246 bp for ST8Sia3. The results are shown in Fig 3-4.

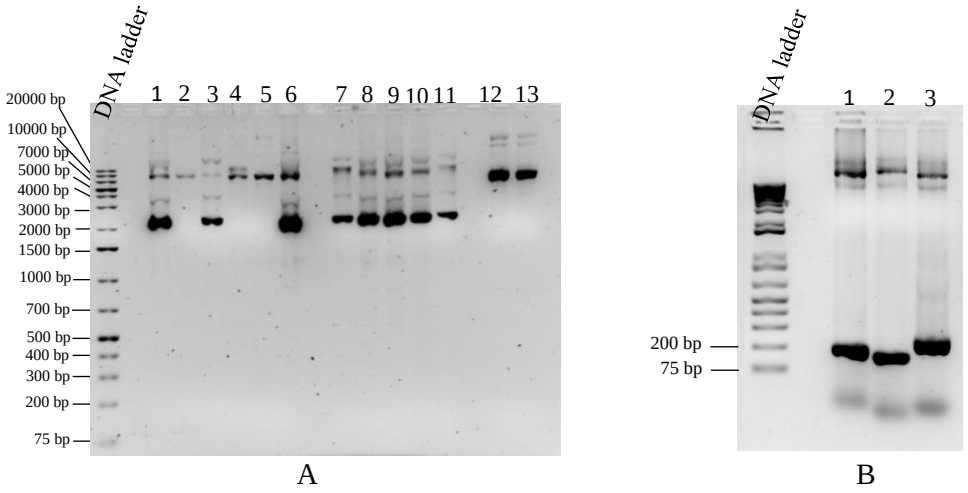


Fig 3-4A: pUC57 plasmids containing sialyltransferase cDNAs. Lane 1-6 indicate the plasmid containing ST3Gal3 cDNA, lane 7-11 indicate the plasmid containing ST6Gal1 cDNA and lane 12-13 indicates the plasmid containing ST8Sia3 cDNA. The different bands indicate the different forms of the plasmid such as linear, circular and supercoiled. **B:** PCR amplification of DNA sequence within the sialyltransferase cDNAs. Lane 1-3 indicates the amplicons from ST3Gal3, ST6Gal1 and ST8Sia3 cDNAs with product sizes of 193 bp, 156 bp and 246 bp respectively. GeneRuler™1 kb plus DNA ladder was used as a reference.

Construction of the plasmid vectors commenced with the synthesis of the mCherry cDNA. The *mCherry* cDNA was synthesised from the mCherry plasmid by PCR as discussed in section 3.2.1. The expected size of *mCherry* cDNA was 760 bp. The cDNA synthesised by PCR that corresponded to 760 bp was excised and extracted from 1% agarose gel using the Qiagen QIAquick gel extraction kit. The PCR result for *mCherry* cDNA synthesis is shown below in Fig 3-5.

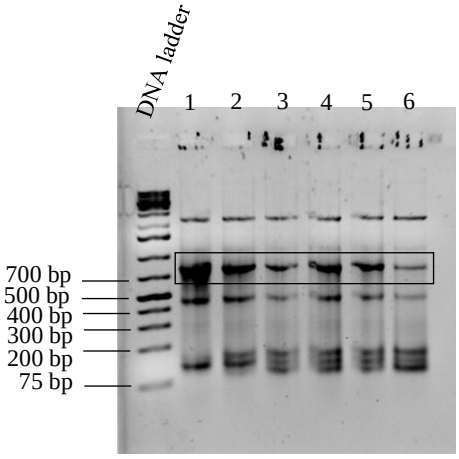


Fig 3-5: Synthesis of *mCherry* cDNA by PCR. The expected product size was 760 bp. The product corresponding 700 bp (indicated by the box) was excised extracted from the agarose gel. The amplicons apart from the ones seen at 700 bp region are probably the products of unspecific binding of the primers. PCR amplification was performed in multiple tubes (lane 1 to 6). The amplification mixture from all the tubes were pooled together and the cDNA fragment corresponding to 700 bp was subsequently extracted from the agarose gel. GeneRuler™1 kb plus DNA ladder was used as a reference.

The gel extracted DNA amplicons that corresponded to 700 bp were subsequently subjected to PCR for further confirmation, using *mCherry* primers (Table 2-1).

After this, the mCherry cDNA synthesised by PCR and the pCBHI corlin vector were digested using *Sna*BI and *Xho*I, to enable ligation and the subsequent generation of a pCBHI corlin plasmid vector containing *mCherry*. *E. coli* DH5α cells were transformed with the ligated plasmids by heat shock transformation (section 2.3.3). The *E. coli* transformants were then selected on LB agar plates containing ampicillin (section 2.2). The transformants that survived the antibiotic selection were subjected to colony PCR to confirm presence of the *mCherry* cDNA in the plasmid used for transformation. Results of colony PCR are shown in Fig 3-6.

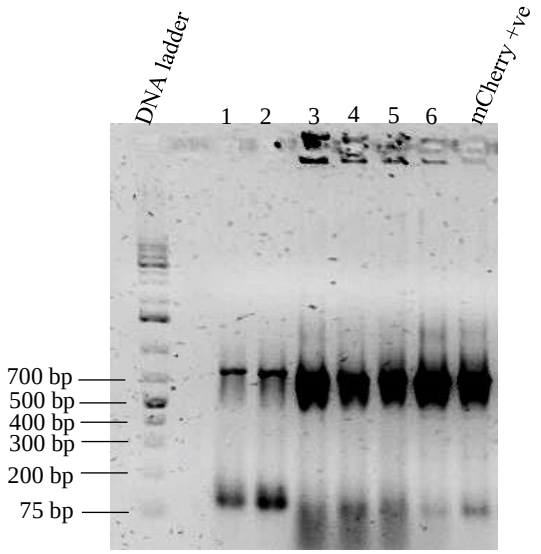


Fig 3-6: Colony PCR of *E. coli* DH5α transformant colonies to confirm the presence of mCherry in transformed pCBHI corlin plasmid. Colony PCR was performed using multiple transformant colonies (lane 1 to 6). The expected PCR product size was 760 bp. GeneRuler™ 1 kb plus DNA ladder was used as a reference. The mCherry plasmid was used as positive control.

Transformant colonies harbouring the pCBHI corlin plasmid containing the *mCherry* cDNA were further propagated in LB medium (section 2.3.3) for the extraction of plasmids. Subsequently, the plasmids containing *mCherry* cDNA were subjected to restriction digestion using *Xho*I and *Nhe*I in order to clone in the sialyltransferase cDNAs (ST6Gal1 and ST3Gal3, section 2.4.5). The recombinant plasmids were transformed into *E. coli* DH5α cells by heat

shock transformation for propagation. Plasmid DNA was then extracted from different transformant colonies (after antibiotic selection) and the presence of ST3Gal3 and ST6Gal1 cDNAs confirmed by PCR. The result is shown in Fig. 3-7.

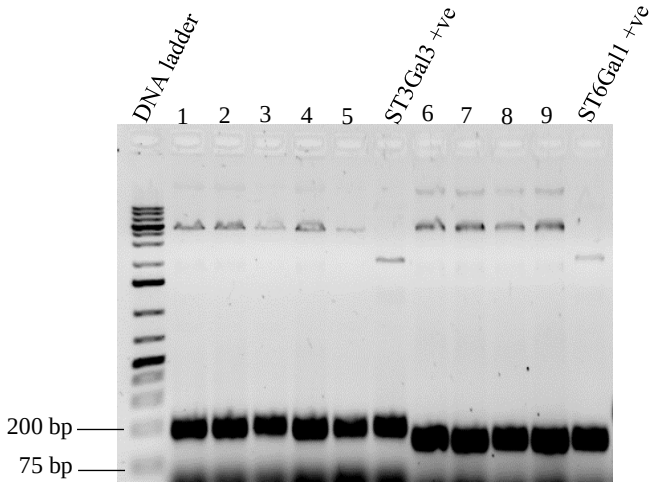


Fig 3-7: PCR amplification of plasmids to confirm the ligation of ST3Gal3 (lane 1-5) and ST6Gal1 (lane 6-9) in pCBHI corlin vector containing *mCherry*. The expected PCR product size was 193 bp for ST3Gal3 and 156 bp for ST6Gal1. The positive controls used were pUC 57 vector containing ST6Gal1 cDNA and ST3Gal3 cDNA respectively. GeneRuler™ 1 kb plus DNA ladder was used as a reference.

Consequently, those plasmid samples that showed the presence of sialyltransferase cDNAs (ST3Sia3 and ST6Gal1) were, once again, subjected to PCR to confirm the presence *mCherry* in them. This result is given in Fig 3-8.

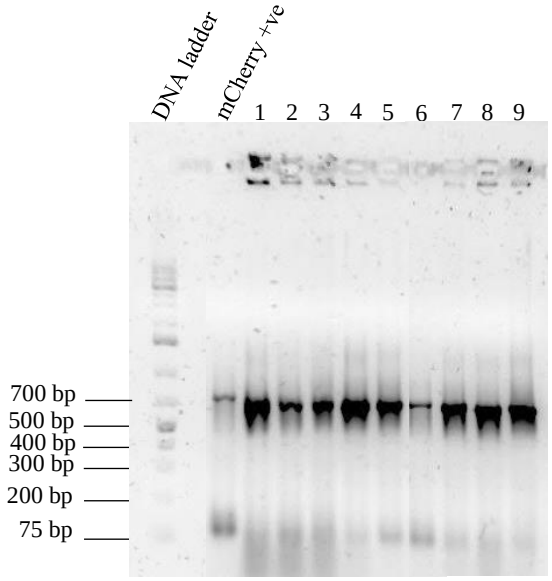


Fig 3-8: PCR amplification of plasmids to confirm the presence of *mCherry* in the pCBHI corlin vector containing sialyltransferase cDNAs (ST3Gal3 and ST6Gal1). Lane 1-5 indicate samples containing ST3Gal3 and lane 6-9 indicate samples containing ST6Gal1. The expected PCR product size was 760 bp. The mCherry plasmid was used as a positive control. GeneRuler™ 1 kb plus DNA ladder was used as a reference.

Two more plasmid constructs carrying ST6Gal1 and ST8Sia3 cDNAs without mCherry were also constructed for transformation. The pCBH1 corlin vector, the ST6Gal1 cDNA and ST8Sia3 cDNA (section 2.4.5) were digested using *Xho*I and *Nhe*1 to enable cloning, as discussed previously. The recombinant plasmids were propagated and amplified in *E. coli* DH5α cells. The plasmid DNAs were subsequently extracted and were subjected to PCR to confirm the ST6Gal1 and ST8Sia3 cDNA inserts. The results are shown in Fig 3-9.

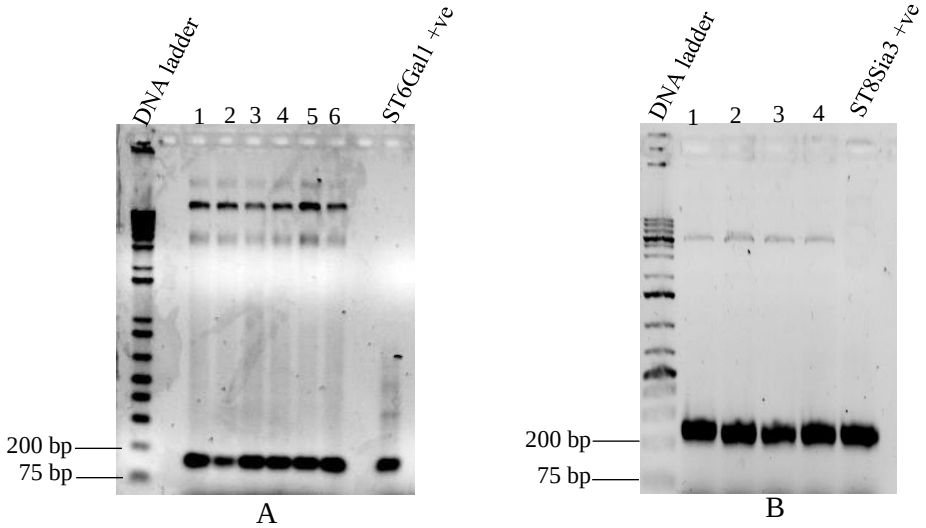


Fig 3-9: PCR amplification of plasmids to confirm the presence of sialyltransferase cDNAs (ST6Gal1 and ST8Sia3). **A**. Lane 1-6 indicate plasmid samples containing ST6Gal. The expected size of the PCR product was 156 bp. The pUC 57 containing ST6Gal1 cDNA was used as a positive control. **B.** Lane 1-4 indicate plasmid samples containing ST8Sia3. The expected size of the PCR product was 246 bp. The pUC 57 plasmid containing ST8Sia3 cDNA was used as a positive control. GeneRuler™ 1 kb plus DNA ladder is used as a reference.

3.3.2. Examination of the alignment of cDNA sequences in the pCBHI corlin plasmid

Plasmids destined to be transformed into cells were sequenced to ensure that all components were in place in the desired configuration. The sequencing results are given as supplementary material (Appendix (i) to (iv)). The *cbh1* core-linker sequence was genetically engineered to be in frame with the *cbh1* promoter at the time of designing the pCBH1corlin vector. Hence, if the ORF's of the heterologous cDNA inserts were in frame to the *cbh1* core-linker sequence, this in turn would mean that the whole configuration was in frame with the *cbh1* promoter. The sequenced plasmids were subsequently prepared for transformation. It can be

seen from the results that the cDNA's coding the mCherry and the sialyltransferases were in frame with the 3′ end of the *cbh1* core-linker sequence. Establishing this was important as reading of a cDNA sequence not in frame with the gene promoter sequence may result in a completely different or truncated protein due to a change in the reading frame.

3.3.3. Generation and selection of *Trichoderma reesei* transformants

Biolistic bombardment was used to generate the *T. reesei* transformants. Care was taken to minimise contamination during the entire procedure and the whole set-up was placed inside a laminar airflow cabinet where the procedure of transformation was carried out. Biolistic bombardment was the chosen mode of transformation because it offers an efficient way to transform *T. reesei* (Te'o *et al*., 2002). Biolistic delivery of transforming DNA into fungal genomes, especially when performed on uninucleate haploid conidia, has proven successful in bypassing the time-consuming repetitive purification of protoplasts used for the widely applied polyethylene glycol-mediated method (Te'o and Nevalainen, 2015). Also, a reduced ability of the protoplasts to regenerate their cell wall makes the protoplast transformation particularly difficult with *T. reesei* RUT-C30 (Nevalainen *et al*., 1995; Bergquist *et al*., 2002). While the transformation frequency for biolistic bombardment does not compare to the high numbers reported in the protoplast approach, transformants generated with this method have been found to be mitotically more stable (Lorito *et al*., 1993; Herzog *et al*., 1996).

The transformants generated were selected based on their hygromycin resistance (attributed by the *E. coli* hygromycin resistance gene (*hph*) in the expression cassette) by cultivating for up to three generations in PDA plates containing hygromycin B (section 2.5.5). This was to ensure mitotic stability of the recombinant expression cassette integrated into the fungal genome, before transferring to fresh PDA plates for conidiation. The number of transformants obtained varied with different expression cassettes and are documented in Table 3-2.

Table 3-2: Number of stable transformants generated by biolistic bombardment using different expression cassettes.

Expression cassette	No. of stable transformants	Transformation frequency (per µg DNA)
ST3Gal3 with mCherry	66	26.4
ST6Gal1 with mCherry	40	16
ST6Gal1	18	7.2
ST8Sia3	46	18.4

The results above indicate a transformation frequency ranging from a minimum of 7.2 to 26.4 stable transformant colonies per µg of DNA used, depending on the expression cassette. Biolistic bombardment has been successfully used to transform various organisms such as bacteria, algae, plant cells and fungi. The first successful demonstration of using biolistic transformation for *T. reesei* was by (Hazell *et al*., 2000) by microprojectile bombardment of the fungal conidia with the fungal gene coding for acetamidase *amdS* and the bacterial hygromycin phosphotransferase (*hph*) gene claiming a transformation frequency of up to 11 stable transformant colonies per µg of DNA used. The mitotic stability of the transformants was reported to be 75–100%. Later, Te'o *et al*. (2002) achieved successful transformation of *T. reesei* conidia by using a Bio-Rad seven barrels Hepta Adapter system claiming transformation frequencies of up to 37 colonies per µg of DNA. Moreover, the mitotic stability of transformants was reported to be 98-100%, a result similar to what was observed in this project. All the transformants obtained by biolistic transformation in this work survived even after three rounds of selection on PDA plates containing hygromycin B, indicating the stable integration of the expression cassettes containing sialyltransferase genes.

It has been reported previously that various factors such as the use of suitable projectile (gold or tungsten), particle velocity (controlled by He gas pressure), size of the micro projectile and the amount of vacuum maintained during the transformation procedure play critical roles in

contributing to a highly efficient transformation (Klein *et al*., 1988). Other factors that are expected to affect the transformation efficiency are temperature, the number of conidia and their ability to regenerate, the number of DNA-coated particles and the amount of DNA that covers each particle (Rivera *et al*., 2014).

It was observed, during the course of this project, that using gold particles resulted in a better transformation efficiency than using tungsten particles (results are not shown). Furthermore, it was also observed that the vacuum applied to the transformation chamber played a critical role in the efficient transformation of fungal conidia. The gold particles used were of the size 0.6 micron, which is almost 15 times smaller than the average size of talcum powder particles. Hence, a slight decrease in vacuum resulting in an increase in residual air inside the shooting chamber from 28" will introduce air drag thereby slowing down the accelerated particles thus affecting the transformation efficiency. This potential technical error could have been a reason for the lower transformation frequency observed with the ST6Gal1 expression cassette without mCherry (Table 3-2). Since, the vacuum maintained in the shooting chamber is also a factor that can result in the death of the conidia this possibility was also considered to explain the low number of the ST6Gal1 transformants without mCherry.

A distinct difference in the phenotype across various transformants was observed while cultivating the transformants in PDA plates containing hygromycin B. Also there were visible differences in the growth, conidiation and the general morphology of the transformant colonies when cultivated for a period of seven days (Fig 3-10). The difference in the morphology of fungal transformant colonies served as an early indication for the diversity in the location of insertion of the expression cassette within the fungal genomic DNA. Non-homologous integration or random integration does not rely on homologous DNA sequences for recombination. Rather, DNA sequences are joined by direct ligation of the ends of DNA

strands without the requirement for sequence homology (Krappmann *et al*., 2006; Ishibashi *et al*., 2006). The different patterns of growth of the fungal colonies may have occurred due to the stress introduced by deletion of genes important for the normal fungal growth and physiology, due to random recombination. It could also be true that the hygromycin B in the PDA plates may have been lethal to the younger mycelia towards the periphery of the colony resulting in its death and subsequent irregular radial growth pattern.

One of the transformants (Fig 3-10 J) exhibited white coloured conidia as opposed to the green pigmentation in normal *T. reesei* RUT-C30 conidia (Fig 3-10 A). This may have happened because of the probable loss of polyketide synthase (*pks4*) gene responsible for the green pigmentation of *T. reesei* conidia (Atanasova *et al*., 2013) due to random insertion of the heterologous expression cassette. However, this speculation can be confirmed only after establishing the loss of *pks4* gene from the genomic DNA by PCR analysis. The variance in phenotypes indicated that the expression cassette was integrating into the genome. In fact, integration of a foreign gene is mainly non-homologous in most fungi except yeast such as *S. cerevisiae* (Ishibashi *et al*., 2006). Also, the frequency of targeted integration by homologous recombination is only up to 5% when using biolistic transformation so most of the integration events in this work were expected to be ectopic (Davidson *et al*., 2000; Te'o *et al*., 2002). While Fig 3-10 displays transformant colonies different to RUT-C30, the colony morphology of almost all other transformants looked similar to the non-transformant RUT-C30. Also, colony morphology did not indicate the expression of mCherry.

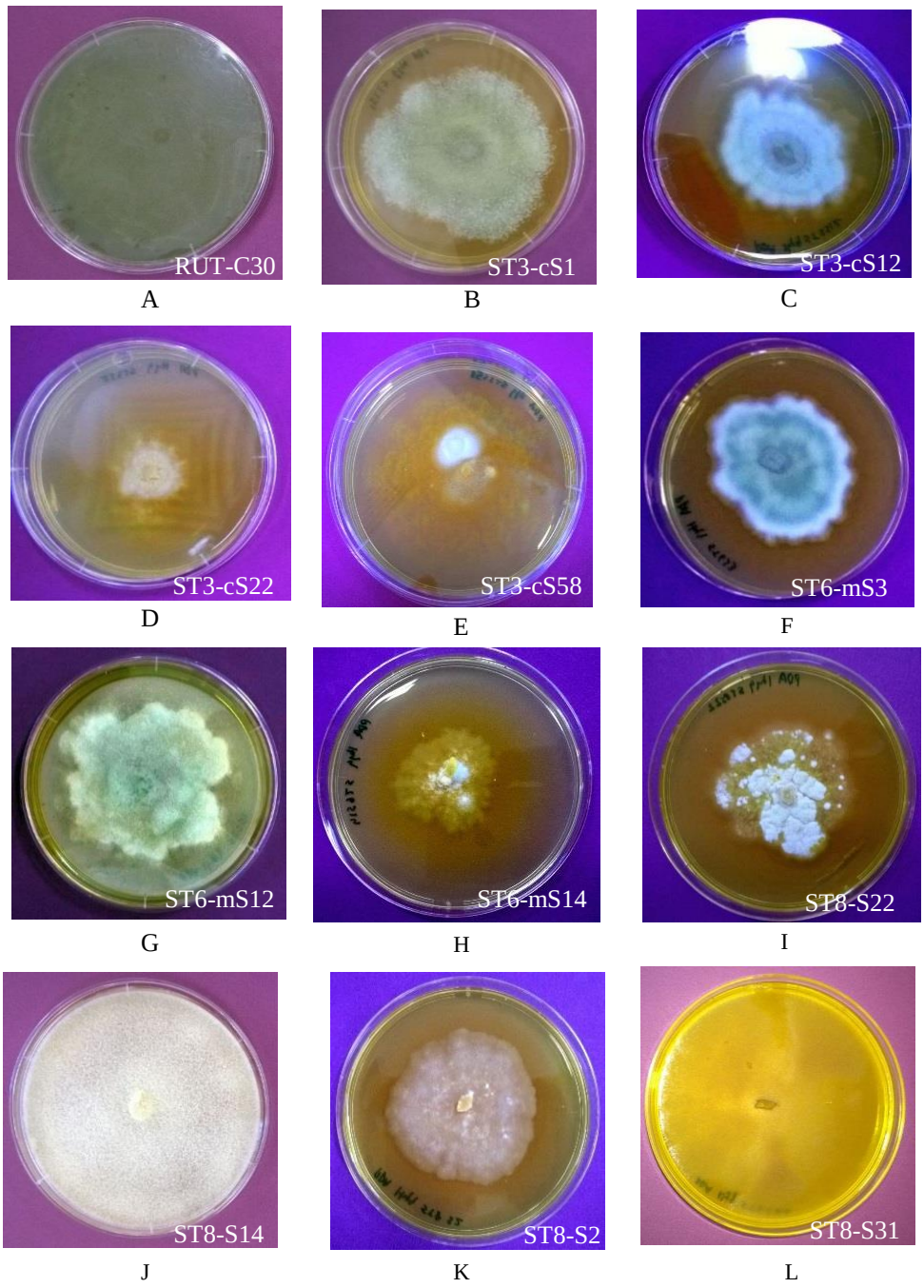


Fig 3-10: *T. reesei* transformant colonies exhibiting phenotypical differences after seven days of cultivation on PDA plates containing hygromycin B. **A**: Normal phenotype of RUT-C30. **B-L**: Variance in phenotype exhibited by ST3Gal3, ST6Gal1 and ST8Sia3 transformants indicating a difference in location of expression cassette insertion within the fungal genomic DNA. The transformants containing mCherry are indicated by the letters "c" and "m".

All mitotically stable transformants raised on PDA plates containing hygromycin B were subsequently cultivated on PDA plates for up to seven days to generate conidia. The plates were exposed to sunlight after five days of cultivation at 28ºC to enhance conidiation (Castellanos *et al*., 2010).

3.3.4. Initial screening of transformants based on mCherry expression

The fluorescent reporter protein, mCherry, enabled the preliminary screening of the *T. reesei* transformants on the basis of mCherry expression provided that the sialyltransferases were expressed in fusion to the mCherry protein as the 5′ end of the sialyltransferase cDNAs were fused to the 3′ end of mCherry cDNA. It has been reported previously that a difference in the codon preferences can impact negatively on the heterologous expression of proteins in *T. reesei* (Te'o *et al*., 2000). Therefore, the sialyltransferase codons were optimised according to *T. reesei* expression. Conversely, mCherry codons have been found compatible with *T. reesei* expression and were not optimised (Sun, 2008).

The 3′ end of sialyltransferase cDNAs codes for the catalytic domain of the enzyme that is responsible for the activity of the enzyme, *i.e*. transfer of sialic acids. Therefore, to avoid any hindrance to the catalytic activity of the recombinant sialyltransferases, the mCherry cDNA was fused to the 5′ end of the sialyltransferase genes. Further, cultivating, maintaining and exploring the entire collection of transformants for heterologous expression of sialyltransferases by Western blotting did not seem feasible. Therefore, preliminary screening based on the expression of mCherry was essential to reduce the number of transformants for further analyses.

The transformants were cultivated in 3 mL CLS medium for 72 h before collecting the samples for analysing the intracellular and extracellular mCherry fluorescence (section 2.7). The intracellular fluorescence in the transformant hyphae indicated successful expression of mCherry-containing fusion protein under the *cbh1* promoter. However, in order to confirm the secretion of the fusion protein, and to shortlist the best secretor, mCherry fluorescence in the culture supernatant was analysed.

3.3.4.1. Monitoring the intracellular expression of mCherry

The intracellular fluorescence analysis of mCherry containing transformants was performed using confocal fluorescence microscopy as explained in section 2.7.1. Confocal fluorescence microscopy is beneficial for the use of fluorescent reporter proteins in filamentous fungi due to the relatively large cell size and thickness of hyphae. Previous studies have indicated that samples grown in liquid cultures had a more uniform distribution of hyphae and were found to provide the more suitable material for monitoring expression of fluorescent protein compared to the cultures obtained from PDA plates (Curach, 2005). Therefore, *T. reesei* transformants were cultivated in liquid CLS medium for fluorescent live cell imaging.

The orange-red fluorescence emitted by live hyphae of transformants containing mCherry, when observed under fluorescent microscopy, indicated the intracellular expression of mCherry. Fig 3-11 A and B show the fluorescent microscopy images of transformants with *ST3Gal3* and *ST6Gal1* genes respectively (note that the microscopy images of only selected transformants are shown). The level of fluorescence emitted varied across different transformants with some of the transformants showing very little to no intracellular mCherry fluorescence. This observation was against the expected outcome because, ideally, all the transformants should be expressing the same amount of mCherry since the same expression cassette was used for transformation.

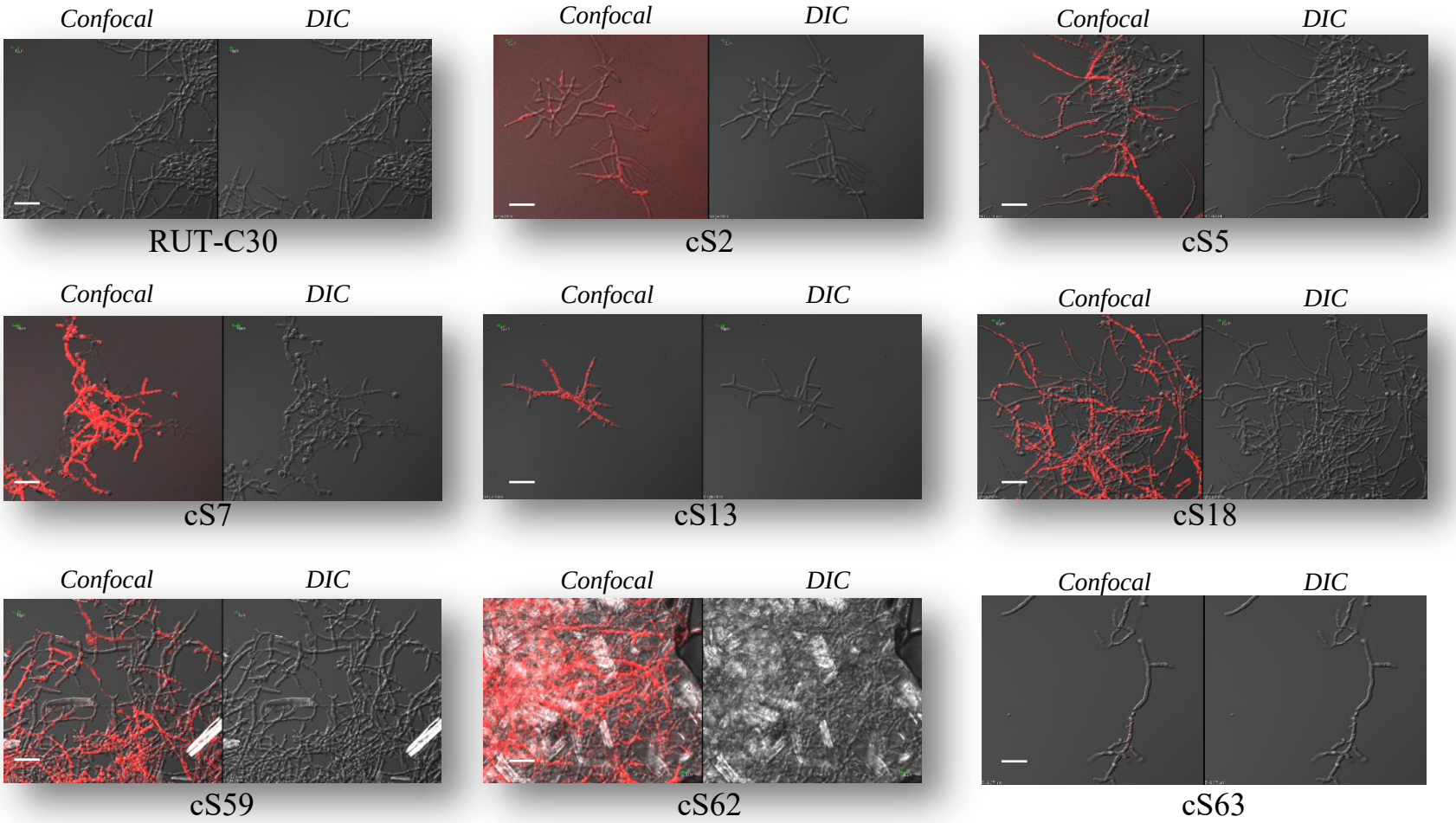


Fig 3-11 A: Live cells of the non-transformant RUT-C30, and selected transformants (cS) containing ST3Gal3 expression cassettes showing intracellular mCherry fluorescence. The transformants were cultivated in 3 mL CLS medium for 72 h at 28°C. Confocal fluorescence microscopy was performed at an excitation/emission wavelength of 559/612 nm to reveal the intracellular presence of mCherry. The hyphal morphology was observed by differential interference contrast (DIC) microscopy. Scale bars represent 30 µm.

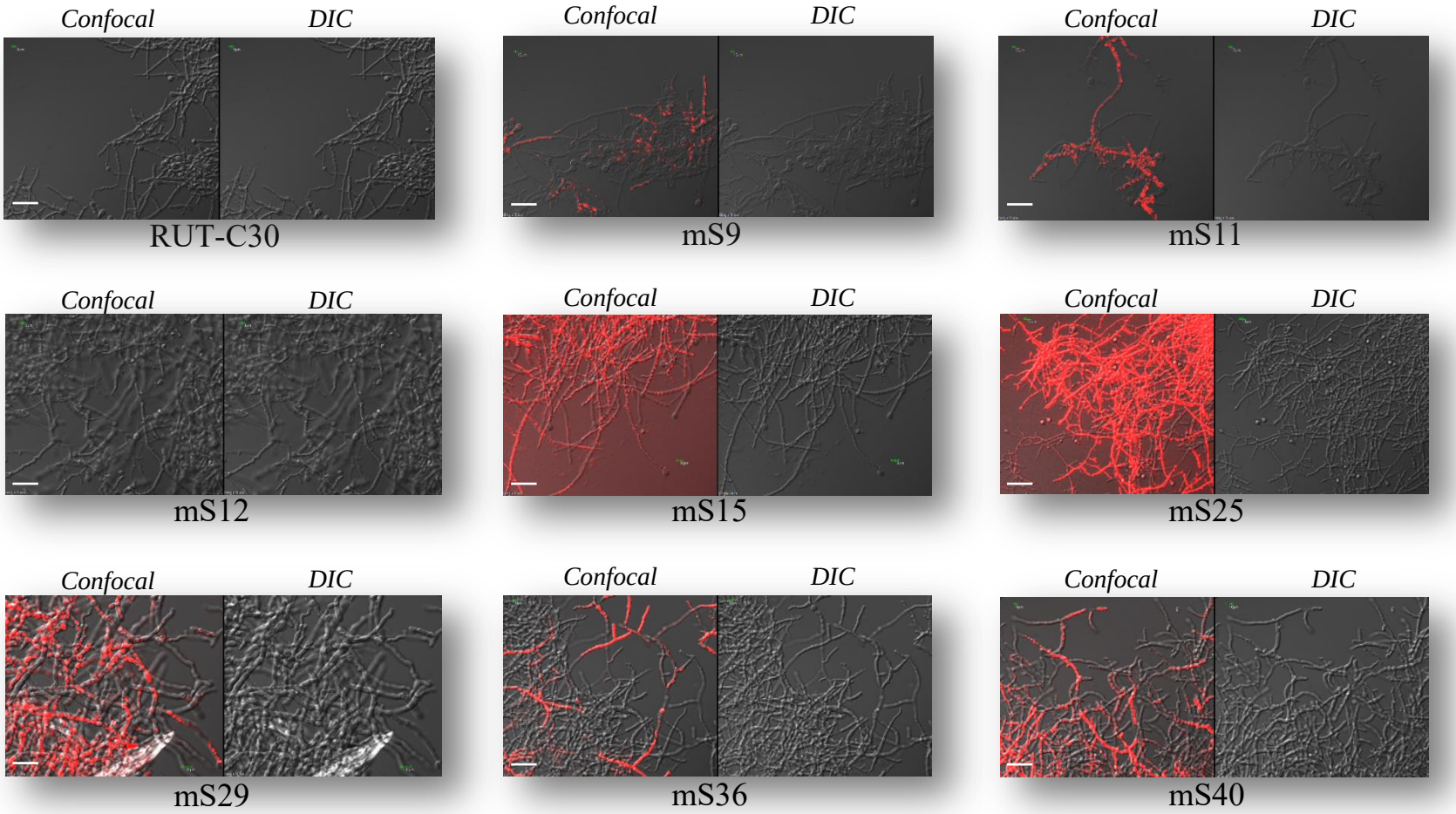
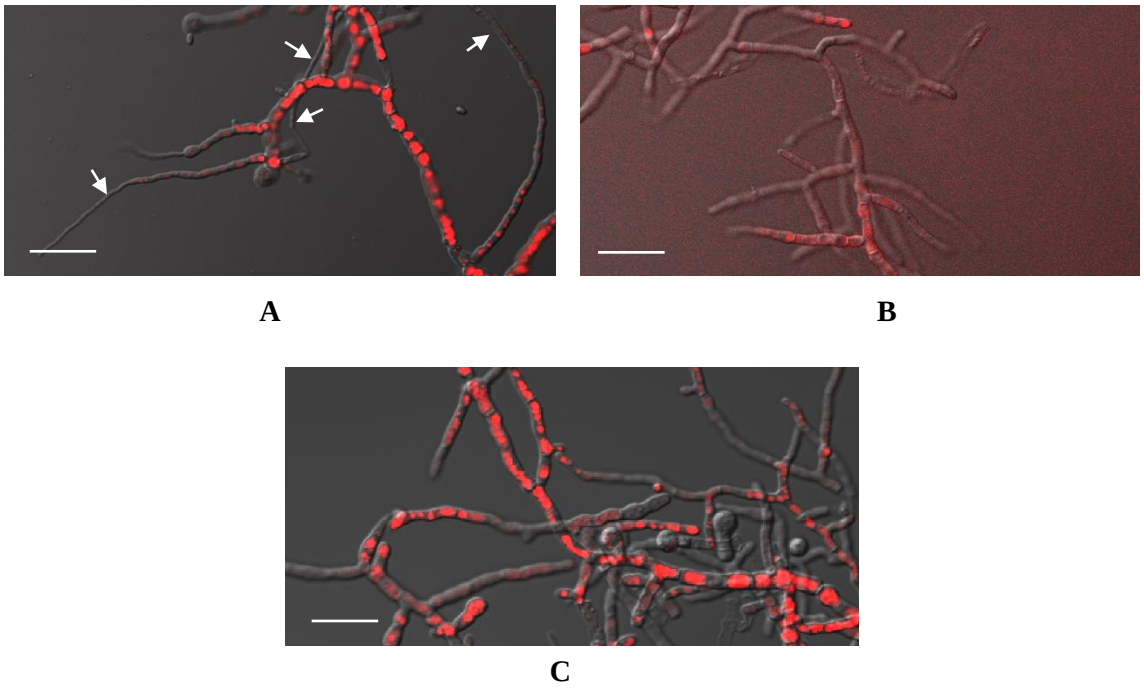


Fig 3-11 B: Live cells of the non-transformant RUT-C30, and selected transformants (mS) containing ST6Gal1 expression cassettes showing intracellular mCherry fluorescence. The transformants were cultivated in 3 mL CLS medium for 72 h at 28°C. Confocal fluorescence microscopy performed at an excitation/emission wavelength of 559/612 nm revealed the intracellular presence of mCherry. The hyphal morphology was observed by differential interference contrast (DIC) microscopy. Scale bars represent 30 µm.

The difference in intracellular fluorescence may have been due to the difference in the number of copies of expression cassettes, and hence the mCherry genes, that were integrated into the fungal genome. The number of *ST3Gal3* and *ST6Gal1* containing transformants that displayed no visible mCherry fluorescence within the hyphae were 10 and 12 respectively. The complete lack of mCherry fluorescence maybe attributed to the lack of expression of the *mCherry* gene probably due to damage that the expression cassette might have incurred during biolistic transformation.

In addition to the difference in intracellular fluorescence, a difference in the morphology of hyphae was observed across different transformants under the microscope. For instance, the mCherry containing ST3Gal3 transformant cS5 exhibited thinning of hyphae towards the tip as they grew (Fig 3-12 A), whereas the hyphae of transformants cS2 and cS18 (Fig 3-12 B and C respectively) were smooth and appeared normal.

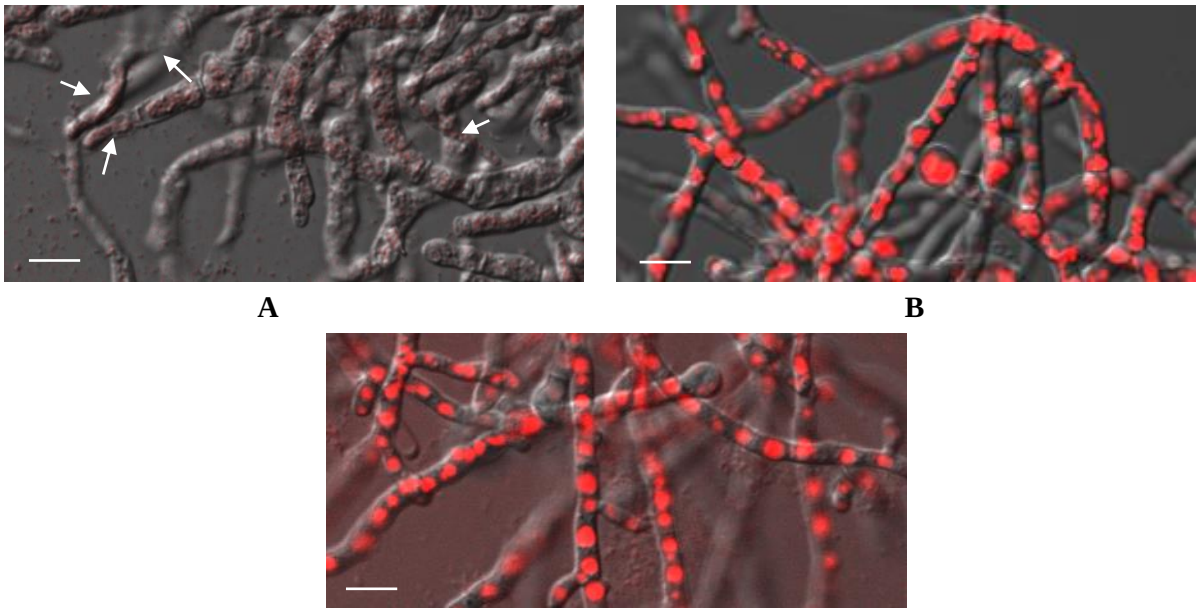


A **B**

C

Fig 3-12 A: Thinning of hyphae (shown by white arrows) exhibited by ST3Gal3 transformant cS5. **B**: Normal looking hyphae exhibited by ST3Gal3 transformant cS2. The background fluorescence might be from the secreted mCherry. **C:** Normal looking hyphae exhibited by ST3Gal3 transformant cS18. ST3Gal3 transformants shown in the figure displayed visible expression of mCherry inside the hyphae. Scale bars represent 30 µm. The images are generated by merging DIC with confocal fluorescence microscopy images.

Similarly, the mCherry containing ST6Gal1 transformants mS3, mS5 and mS15 also exhibited a significant difference in hyphal morphology (Fig 3-13). The hyphae of transformant mS3 exhibited an atypical morphology as opposed to the smooth hyphae exhibited by the transformant mS5.



A **B**

C

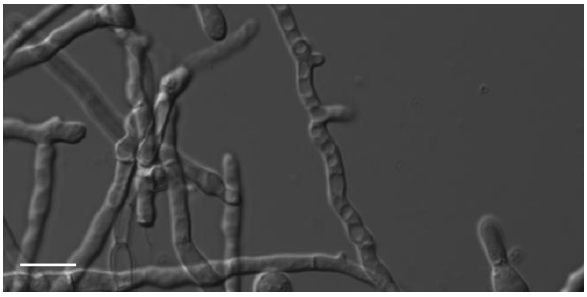
Fig 3-13 A: The coarse appearance of the hyphae of ST6Gal transformant mS3 indicating cell damage (shown by white arrows).**B:** Normal looking hyphae exhibited by ST6Gal1 transformant mS5 **C:** Smooth appearance of mS15 hyphae indicating good expression of mCherry. The background fluorescence might be from secreted mCherry. Scale bars represent 10 µm. The images are generated by merging DIC with confocal fluorescence microscopy images.

More importantly, the ST3Gal3 transformants cS2, cS5, cS18 (Fig 3-12) and the ST6Gal1 transformants mS5 and mS15 (Fig 3-13) displayed a significant amount of mCherry fluorescence within the hyphae which appeared as small globules and seemed to have a high abundance of vesicular structures. This may further indicate that these transformants are expressing the fusion proteins quite proficiently inside the hyphae, provided that the sialyltransferase is still attached to the mCherry protein.

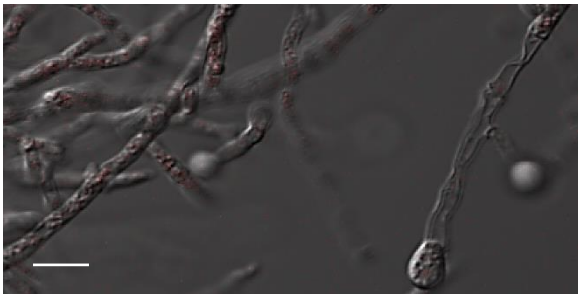
The intracellular mCherry fluorescence of transformant mS3 was feeble indicating a very low expression of mCherry containing fusion protein. Also there were signs of abnormalities on

its hyphae that appeared like lesions (3-13 A). This might have occurred because of the loss of fungal structural genes that was necessary for normal hyphal growth and development, due to biolistic transformation.

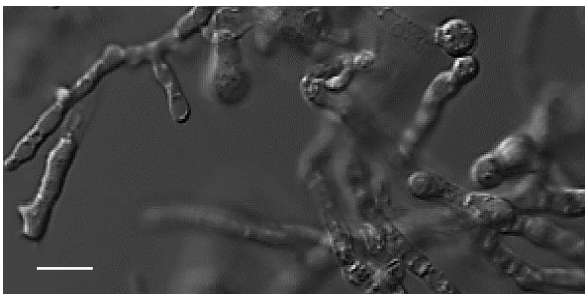
The fungal colonies of the transformants cS1 and mS12 cultivated in the PDA plates containing hygromycin B (Fig 3-10 B and G respectively) had a close resemblance to the normal RUT-C30 colony morphology (Fig 3-10 A). However, there was no visible intracellular mCherry fluorescence in the hyphae of the transformant mS12 and the transformant cS1 showed very weak intracellular mCherry fluorescence. Also, they indicated abnormalities in the hyphae compared to the non-transformant RUT-C30 hyphae when observed under a DIC microscope (Fig 3-14).



A



B



C

Fig 3-14 A: The appearance of the hyphae of non- transformant RUT-C30. **B:** Abnormal looking hyphae of cS1 and **C:** mS12. Scale bars represent 10 µm. The images are generated by merging DIC with confocal fluorescence microscopy images.

Secretion of proteins in filamentous fungi typically occurs through the growing hyphal tips (Wösten *et al*., 1991; Nevalainen and Peterson, 2014b). However, the mCherry was visualised throughout the length of transformant hyphae and not just the hyphal tips in the samples collected at 72 h of cultivation (Fig 3-11, 3-12 and Fig 3-13). Moreover, the visualisation of

mCherry as small globules within the hyphae may indicate that the recombinant mCherry containing fusion protein was trapped in the endoplasmic reticulum (ER), vacuoles or in the secretory vesicles of the fungus. It may be speculated that since the heterologous expression was achieved under the most prolific *cbh1* promoter of *T. reesei*, the fungus might not have been able to cope with the processing and secretion of the high amount of proteins.

It was recently reported by Nykänen *et al*. (2016) that the highly expressed BiP1-Venus YFP fusion protein in *T. reesei* RUT-C30 mutant BV47 was found localised in sporadic punctate bodies within the hyphae. In fact, the presence of punctate bodies was found to be over ten times more in the BV47 mutant compared to the transformation host RUT-C30. These punctate bodies are believed to be functioning as repositories for misfolded proteins waiting for further processing. Therefore, it may be hypothesised that the mCherry containing circular bodies visualised as globules in the hyphae (Fig 3-11, 3-12 and 3-13) might be the aforementioned sporadic punctate bodies restraining misfolded heterologous fusion proteins for further processing or degradation.

3.3.4.2. Monitoring the secretion of mCherry into the cultivation medium

The fluorescence imaging of live cells provided enough evidence of intracellular presence of the heterologous fusion protein containing mCherry (provided that the fusion protein is intact). As the aim was to have the protein secreted outside the cells, the next step was to investigate the secretion of the fusion protein into the culture supernatant and screen the transformants based on the extracellular abundance of mCherry fluorescence. An increased fluorescence in the culture supernatant may reflect the ability of the transformant to secrete high amounts of fusion protein. Fig 3-15 A and B shows the mCherry fluorescence in the culture supernatants of transformants containing ST3Gal3 and ST6Gal1 genes respectively.

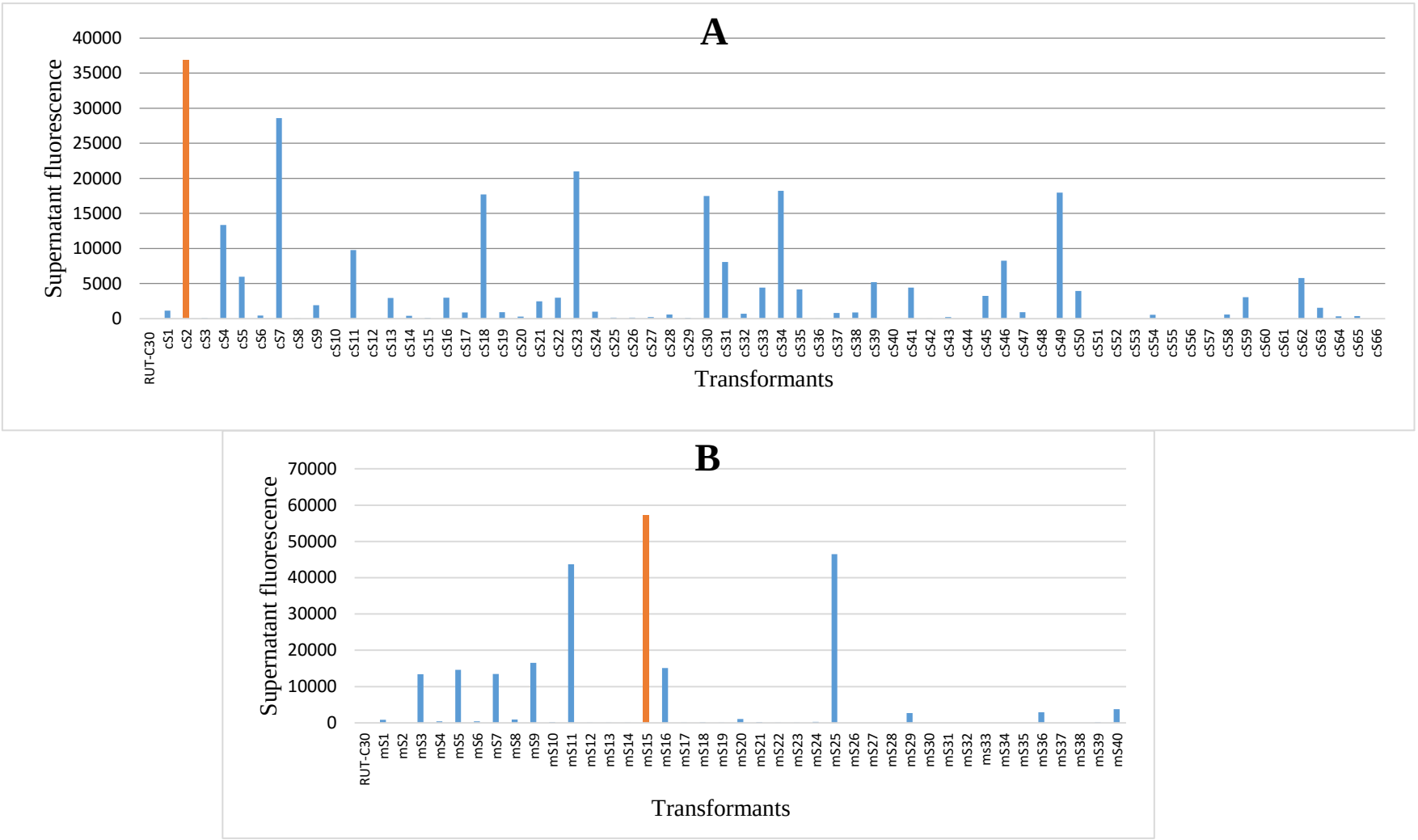


Fig 3-15: A: The mCherry fluorescence observed in the culture supernatant of *ST3Gal3* containing transformants. **B**: The mCherry fluorescence observed in the culture supernatant of *ST6Gal1* containing transformants. The orange coloured bars indicate the transformant that exhibited highest mCherry fluorescence in the culture supernatant (cS2 and mS15).

Fluorescence readings from the transformant culture supernatants were considerably distinct from each other indicating a difference in the amount of mCherry containing fusion protein being secreted into the culture supernatant. It can be observed by comparing Fig 3-11 and 3-15, that even though mCherry fluorescence was observed in the hyphae, many transformants did not secrete the fusion protein outside the cells. For instance ST3Gal3 transformants cS59 and cS62 and ST6Gal1 transformants mS29 and mS40 seemed to show quite proficient expression of mCherry within the cells; however, analysis of the extracellular fluorescence revealed quite a low level of secreted mCherry fluorescence indicating that these transformants had difficulty in secreting the fusion protein. This suggests that the proteins were either localised in the intracellular compartments for further processing/degradation, thus restraining them from being secreted, or these transformants are simply not good secretors of proteins at all. Furthermore, transformants cS2, cS7, mS15 and mS25 seemed to secrete mCherry quite proficiently indicating that these transformants are efficient secretors of proteins. This might have been a reason for the background fluorescence shown by these transformants during fluorescence microscopy of live cells as shown in Fig 3-11, Fig 3-12 B and Fig 3-13 C.

Furthermore, it was also observed that the ST3Gal3 transformant cS5, that exhibited thinning of hyphae (Fig 3-12 A), revealed a low amount of secreted mCherry fluorescence despite the abundance of the intracellular mCherry fluorescence. This thinning of the hyphae, especially towards the tip, may be the reason for weak secretion, since the secretion of proteins mainly occurs through the hyphal tip of a filamentous fungus. In spite of the similarity of the colony morphology of cS1 and mS12 (Fig 3-10 B and G respectively) to RUT C30, the secretion of mCherry was significantly low compared to the efficiently secreting transformants (mS15 and cS2). Furthermore, the transformants that exhibited atypical colony morphologies (Fig 3-10) had considerably weaker mCherry fluorescence inside the cells and in the culture supernatant

(Fig 3-15). This may also indicate the inefficiency of these transformants to express and secrete proteins (fusion protein containing mCherry in this case) due to the stress experienced because of the random integration of the foreign DNA into the fungal gDNA that might have affected the normal and healthy functioning of the cells.

The low level of extracellular fluorescence exhibited by certain transformants with strong intracellular fluorescence may be due to the inefficiency of the transformants to secrete the protein. Degradation of the fusion protein containing the heterologous fluorophore may also have been a reason for the diminished fluorescence in the culture medium by many transformants. A similar observation was previously made by Gordon *et al*. (2000) in *Aspergillus niger* where the recombinant glucoamylase-GFP fusion protein was degraded by the endogenous proteases. Another potential reason could be ectopic or non-homologous integration of the expression cassette into the fungal genome, which may have led to the disruption of genes, for example, involved in the formation of secretory vesicles or the translocons on the ER membrane, important for protein secretion. Gene disturbance has been often reported with transformation using a gene gun (Weld *et al*., 2006). At the same time, ectopic integration may also become handy at times if the resulting transformants gain certain desired qualities such as knocking out of protease encoding genes or enhanced secretion of proteins due to the repression or deletion of a gene.

The transformants containing mCherry fused to the sialyltransferase were selected for further analyses based on the higher culture supernatant fluorescence, which, however, was not necessarily an indication for the expression or secretion of sialyltransferases. In order to establish the secretion of sialyltransferases as well as to screen transformants without the *mCherry* gene (section 3.3.3), a more targeted approach such as Western blotting was necessary.

To maintain simplicity in analysis and consistency in comparison of the results, it was decided to continue the project with only the *ST6Gal1* containing transformants with and without *mCherry*. Investigating the transformants for the production of rST6Gal1 will be discussed in detail in the following chapter. The ST6Gal1 transformants with mCherry were given a prefix “mS” and non-mCherry ones “S”.

3.4. Summary

Genes encoding the human sialyltransferases (ST3Gal3, ST6Gal1 and ST8Sia3) and mCherry fluorescent reporter protein were successfully cloned into the pCBH1corlin vector and transformed into *T. reesei* using biolistic bombardment. The transformation frequency ranged from a lowest of 7.2 to a highest of 26.4 colonies per µg of DNA used. All the transformants generated were found to be mitotically stable even after three rounds of selection on PDA plates containing 80 µg/mL of hygromycin B. A difference in the growth pattern and morphology of colonies of the fungal transformants was observed when cultivated in PDA plates containing hygromycin B indicating the variance in location of integration of the expression cassette in the fungal gDNA.

Examination of the hyphae of the transformants cultivated in 3 mL CLS medium by confocal microscopy revealed the intracellular presence of mCherry fluorescence in the majority of the transformants indicating the expression of the fusion protein (mCherry fused to sialyltransferase). The intracellular mCherry fluorescence was visualised as globules along the length of the hyphae that appeared to have been localised in the fungal cell organelles such as endoplasmic reticulum (ER), vacuoles or punctate bodies. One possible reason for the detection of high level of intracellular mCherry fluorescence could be because of the localisation of the fusion proteins in the ER for correct folding and processing. It may also be

true that the fusion proteins may be trapped in other cellular organelles such as secretory vesicles and are being transported towards the hyphal tip for secretion.

Examination of the culture supernatant of the transformant cultivation medium revealed the presence of mCherry fluorescence indicating the secretion of the fusion protein. However, not all the samples examined exhibited mCherry fluorescence in the culture medium. The ST3Gal3 transformants such as cS59 and cS62 and the ST6Gal1 transformants mS29 and mS40 seemed to have low culture supernatant fluorescence despite the significant presence of mCherry inside the hyphae, indicating bottlenecks in the secretion of the fusion protein. On the other hand, the ST3Gal3 and ST6Gal1 transformants cS2 and mS15 secreted the highest level of mCherry into the cultivation medium respectively indicating that these transformants are efficient secretors of the fusion proteins or at least of mCherry.

In order to maintain simplicity in analysis and consistency in comparison of the results, only the *ST6Gal1* containing transformants with and without mCherry were selected for further studies. The transformants mS3, mS5, mS7, mS9, mS11, mS15 and mS25 containing mCherry (selected on the basis of mCherry fluorescence in the culture supernatant) and all the non-mCherry transformants (S1 to S18) were chosen for subsequent analysis for the expression of rST6Gal1.

4

EXPRESSION OF HUMAN SIALYLTRANSFERASE ST6GAL1 IN *TRICHODERMA REESEI* RUT-C30

4.1. Introduction

Trichoderma reesei is an efficient cell factory for the production of industrially important enzymes such as cellulases and hemicellulases. Despite the high yields of homologous secreted proteins amounting up to 100 g/L, success with the heterologous expression of commercially viable proteins, especially bio therapeutic proteins, still remains a challenge (Nevalainen and Peterson, 2014b). Some of the most notable accomplishments reported recently in *T. reesei* were the successful expression of a human Fab antibody fragment and the mammalian Obestatin at levels of 8.2 g/L (Landowski *et al*., 2016) and 0.0055 g/L (Sun *et al*., 2016) respectively. Examples of successful expression of human sialyltransferases in *T. reesei* were none, at the time of writing this thesis. This was the first known attempt to express a human sialyltransferase in *T. reesei* RUT-C30.

The expression and secretion of the recombinant human ST6Gal1 in *T. reesei* RUT-C30 are studied in detail in this chapter. The *T. reesei* transformants generated using two different expression cassette constructs, one with mCherry (mS) and the other without mCherry (S) (Chapter 3) were analysed for the expression of recombinant ST6Gal1 (rST6Gal1). The expression cassettes (Chapter 2, Fig 2-6) were designed in such a way that ST6Gal1 in fusion to mCherry and without mCherry will be expressed in fusion with the active domain of the *T. reesei* endogenous protein CBHI. It has been previously shown that the expression of heterologous genes can be improved by fusion to a fragment or a full coding sequence of a native gene. For example, the production of recombinant Fab fragments increased over 50 fold when expressed as a fusion protein with the CBHI core-linker in *T. reesei* RUT-C30 (Nyyssönen *et al*., 1993). Further, Paloheimo *et al*. (2003) also reported that expression of a bacterial xylanase in *T. reesei* clearly increased when the recombinant protein was linked to a carrier polypeptide with an intact domain structure.

In this work, analysis and validation of expression and secretion of ST6Gal1 was accomplished by employing Western blotting, SDS-PAGE and mass spectrometry. Southern blotting was employed to determine the correlation between heterologous protein expression and the copy number of the *ST6Gal1* gene integrated into the fungal genome.

4.2. Materials and methods

4.2.1. Collection of conidia

Fresh conidia were harvested from PDA plates for the cultivation of transformants for recombinant ST6Gal1 expression. Extraction and storage of conidia were performed as described in section 2.3.1. The extracted conidia were stored at 4ºC in the harvesting solution. Stored conidia older than two months were not used at any stage of cultivation. A stock suspension was prepared in a Revco medium and stored at -80ºC for long-term storage.

4.2.2. Cultivation of ST6Gal1 containing transformants

Based on the preliminary screening, the mCherry-containing ST6Gal1 transformants mS3, mS5, mS7, mS9, mS11, mS15 and mS25 that exhibited a high level of fluorescence in the culture supernatant (Chapter 3), were cultivated along with non-mCherry transformants S1 to S18 in 3 mL CLS medium for the induction of recombinant ST6Gal1 (rST6Gal1) expression. Expression of rST6Gal1 was confirmed by Western blotting. The selected transformants were further cultivated in 50 mL CLS medium in shake flasks (section 2.3.2) for subsequent analyses. An inoculum of 6×10^6 conidia and 1×10^8 conidia were used for 3 mL and 50 mL cultures respectively. The CLS medium was prepared as described in section 2.1.2 and the transformants were cultivated as described in section 2.3.2.

4.2.3. Confirmation of recombinant ST6Gal1 expression

4.2.3.1. Western blot analysis

Western blot analysis was carried out as described in section 2.7.4. Total protein concentration was determined by Bradford assay. The total protein concentration in culture supernatants from each transformant was kept equal while loading into the gels for polyacrylamide gel electrophoresis (PAGE) prior to Western blotting. Commercially available recombinant human ST6Gal1 (rhST6Gal1, Abnova, USA, cat# H00006480-P01) was used as a positive control for detection. A lysate from human cancer cell line SKBR3 that over-expressed ST6Gal1 (Lee *et al*., 2014) was used as a positive control where rhST6Gal1 was not available. The presence of mCherry and *Strep*-tag™ II were also established by Western blotting. Primary antibody binding to ST6Gal1 and mCherry was confirmed by using LI-COR® based detection (section 2.7.4.1). *Strep*-tag™ II was confirmed by using *Strep*-Tactin® alkaline phosphatase (*Strep*-Tactin®-AP) based detection (section 2.7.4.2).

4.2.3.2. Mass spectrometry

Mass spectrometric (MS) analysis was performed on protein bands present in the SDS-PAGE gels that corresponded to the bands positively identified by Western blotting for the presence of recombinant ST6Gal1. The MS analysis served as an additional confirmation for the production and secretion of recombinant ST6Gal1. The protein bands were excised and processed as described in section 2.7.5. As an example, the transformant samples mS15 and S16 were chosen for MS analysis.

4.2.4. Extraction of genomic DNA from transformants and Southern blot analysis

The genomic DNA (gDNA) from the fungal transformants was extracted as described in section 2.11. The extracted gDNA was subjected to PCR amplification using primers *hst6fwdpr* and *hst6revpr* (Table 2-1) to confirm integration of the *ST6Gal1* gene within the gDNA. Genomic DNA extracted from non-transformant RUT-C30 was used as a negative control. The gDNA concentration was measured using NanoDrop and the quality was examined by agarose gel electrophoresis.

Approximately 10 µg of good quality genomic DNA extracted from the transformants was used for Southern blotting. The gDNA samples were digested with *Eco*RV overnight at 37ºC. The digested gDNA was used for Southern blotting as described in section 2.12. RUT-C30 gDNA digested with *Eco*RV was used as a negative control and pCBH1corlin vector containing *ST6Gal1* gene as a positive control for hybridisation and detection.

4.3 Results and discussion

4.3.1. Analysis of mCherry and non-mCherry transformants for ST6Gal1 expression by Western blotting

Since the aim of the project was to express ST6Gal1 heterologously, a specific detection method was required to prove its presence in the culture supernatant of the transformants. The previous analyses discussed in Chapter 3 served for the screening of transformants based on the expression of mCherry (Fig 3-4 B and 3-8 B). Also, the non-mCherry transformants can be screened only on the basis of rST6Gal1 expression. Therefore, the cultivation supernatant from 3 mL cultures of the mCherry containing transformants mS3, mS5, mS7, mS9, mS11, mS15, mS25 and the non-mCherry transformants S1 to S18 were subjected to Western blotting. Culture supernatants were collected after 72 h of cultivation. The results are given in

Fig 4-1 A and B. A cell lysate prepared from the human cancer cell line SKBR3 that over-expresses ST6Gal1 was used as the positive control for detection.

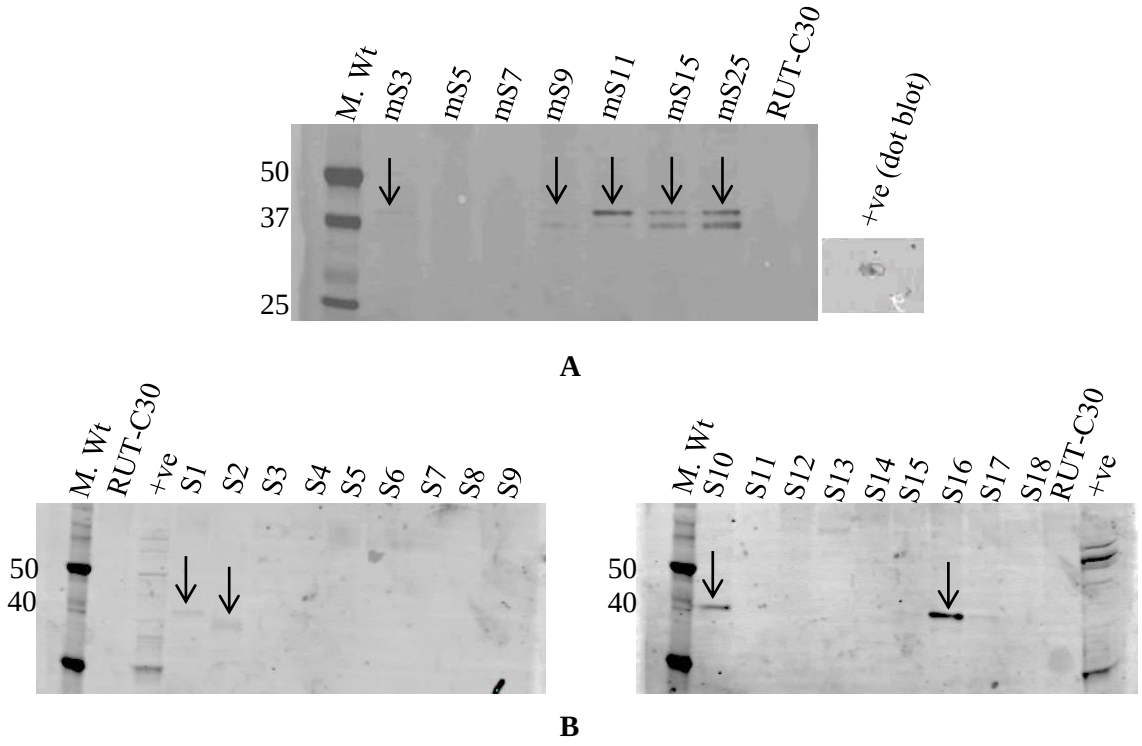


Fig 4-1: Screening of transformants by Western blotting using anti-ST6Gal1 antibodies. **A**: Screening of mCherry containing transformants for ST6Gal1 expression selected based on the mCherry fluorescence in the culture supernatant. The samples mS3, mS9, mS11, mS15 and mS25 indicate the presence of rST6Gal1 in the culture supernatant. The mS15 showed the maximum fluorescence and mS3 and mS7 had the lowest fluorescence in the culture supernatant of the selected transformants. A dot blot using SKBR3 cell lysate was used as positive control for detection. Transformants were cultured in 3 mL CLS medium for 72 h for preliminary screening. **B**: Screening of non-mCherry transformants for ST6Gal1 expression using Western blotting. The sample S1, S2, S10 and S16 indicate the presence of rST6Gal1 in the culture supernatant. SKBR3 cell lysate was used as a positive control for detection. The expected band size in the positive control was approx. 46.6 kDa. Transformants were cultured in 3 mL CLS medium for 72 h for preliminary screening. The rST6Gal1 detected by Western blotting is indicated by arrows.

Among the mCherry containing strains selected for Western blotting, the culture supernatant of mS15 showed the highest mCherry fluorescence followed by mS25 and mS11 whereas the culture supernatants of transformants mS3 and mS7 indicated very low level of fluorescence compared to mS11, mS15 and mS25 (Fig 3-8 B). Analysis of the culture supernatants by Western blotting revealed the presence of rST6Gal1 in mCherry containing strains mS3, mS9, mS11, mS15 and mS25 and in the non-mCherry samples S1, S2, S10 and S16 at a molecular weight of approximately 40 kDa. Moreover, the intensity of Western blot bands was higher

for mS11, mS15 and mS25 (Fig 4-1 A) and was in accordance with the intensity of the mCherry fluorescence in the culture supernatants (Fig 3-8 B).

From Fig 2-6 (Chapter 2), the expression cassette was designed in such a way that it is expected to facilitate the expression of a fusion protein consisting of CBHI core-linker, mCherry and rST6Gal1 (Fig 4-2 A) for the mCherry containing transformants. The CBHI core-linker is expected to be fused with the rST6Gal1 (Fig 4-2 B) for the non-mCherry transformants. The calculated theoretical molecular weight of the fusion proteins using Expasy Compute pI/Mw (Bjellqvist *et al*., 1994) was approximately 120.9 kDa and 93 kDa for the mCherry and non-mCherry containing fusion proteins respectively (Fig 4-2). Also, the calculated theoretical mass of ST6Gal1 lacking the N-terminus transmembrane domain is 43.5 kDa. Therefore, the observation made from the Western blotting (Fig 4-1) indicated the processing of fusion proteins leading to the separation of rST6Gal1.

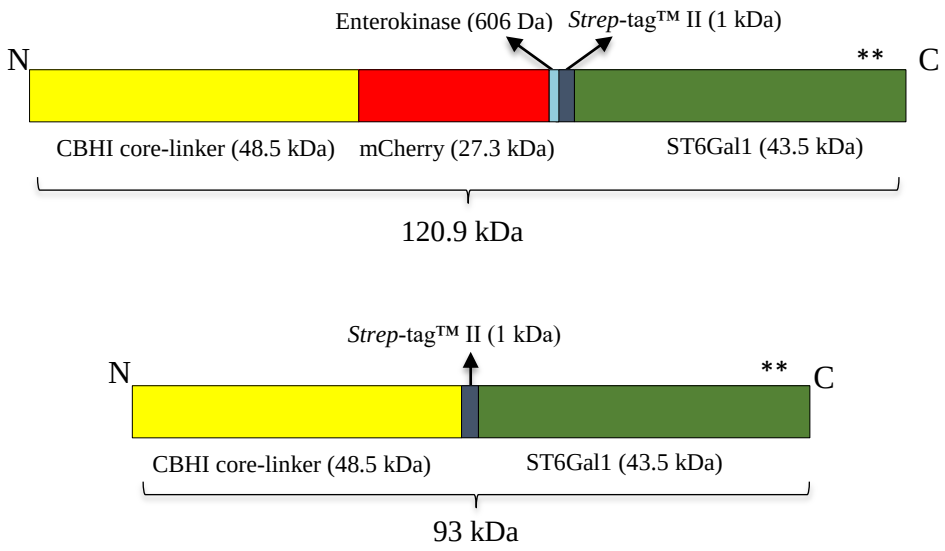


Fig 4-2: Schematic representation of the expected recombinant fusion proteins **A:** fusion protein with mCherry with a theoretical molecular weight of 120.9 kDa (approximately) **B:** fusion protein without mCherry with a theoretical molecular weight of 93 kDa (approximately). ** indicates the anti-ST6Gal1 antibody binding region.

Further analysis of the Western blot image from the mCherry containing transformants mS3, mS9, mS11, mS15 and mS25 revealed two parallel bands close to each other (Fig 4-1 A). A possible cause for this difference in molecular weights can be due to the variance in processing of the glycans present on the rST6Gal1 by *T. reesei* RUT-C30, which could be further studied by the analysis of glycan structures present on the rST6Gal1. There are two *N*-glycosylation sites identified in the human ST6Gal1 protein at the Asn149 and Asn161 residues (Kuhn *et al*., 2013). It has been reported that the removal of carbohydrate chains from ST6Gal1 with *N*-glycanase resulted in a large reduction in its catalytic activity. Moreover, the presence of trimannose core with GlcNAc attached is found to be very important for the catalytic activity of human ST6Gal1 (Fast *et al*., 1993).

However, it may also be true that the cleavage of rST6Gal1 from the fusion protein has occured at two different adjacent locations resulting in cleaved proteins with a slight difference in their molecular weights. N-terminal sequencing (Speicher *et al*., 2009) may be an option to study the exact location of cleavage that has led to the separation of the rST6Gal1 from the fusion proteins.

One potential location of cleavage of the fusion protein is at the KEX2 like cleavage site at the end of the CBHI core-linker in the MCS2 (Fig 2-1). This could be true for the non-mCherry fusion protein as it would produce the rST6Gal1 (attached to *Strep*-tag™ II) with a molecular size of 44.5 kDa. However, for the mCherry-containing fusion protein, the cleavage at the KEX2 site would have resulted in a protein with a larger molecular weight of 70.8 kDa.

The enterokinase cleavage site located between the mCherry rST6Gal1 also seems like a probable cleavage location that has cleaved the rST6Gal1 from the rest of the fusion protein. However, previous studies by Salles *et al*. (2007) indicated that the expression of XynVI in

conjunction with a FLAG-tag containing the enterokinase cleavage site was not cleaved when expressed in *T. reesei*.

It can also be speculated that the short linker sequence consisting of four amino acids incorporated between *Strep*-tag™ II and ST6Gal1 (section 2.4.5) may have introduced cryptic cleavage sites that resulted in the cleavage of rST6Gal1 from the fusion protein. Subsequently, based on the results from Western blotting on the 3 mL cultures, the mCherry containing transformants mS11, mS15, mS25 and non-mCherry transformants S1, S10, S16 were chosen for further analyses. The six selected transformants along with non-transformant RUT-C30 were cultivated in 50 mL CLS medium in shake flasks.

4.3.2. Shake flask culturing of the transformants selected based on Western blotting

The shake flask culturing of transformants mS11, mS15 mS25, S1, S10 and S16 was carried out for 120 h. Culture supernatants were collected at regular intervals of 24 h for subsequent analysis by Western blotting. During the course of cultivation of the transformants, it was noticed that the cultivation medium containing mCherry transformants gradually shifted the colour to orange-red serving as a visual indication for the secretion of mCherry as shown in Fig 4-3.

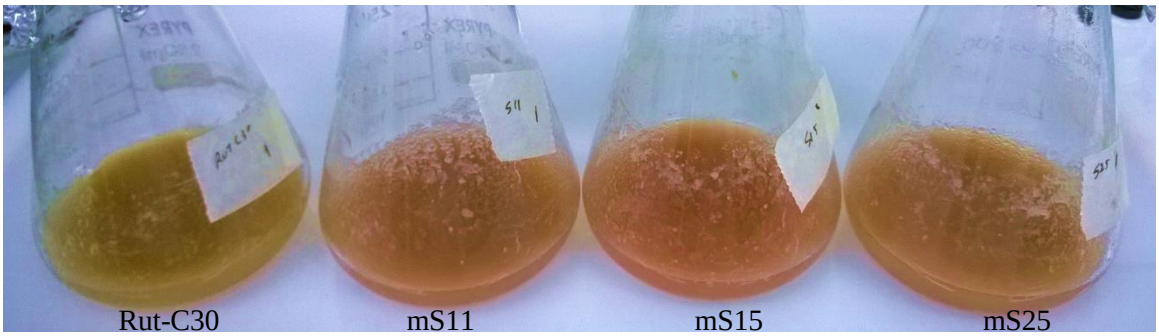


Fig 4-3: The shift in colour of the transformant culture medium after 96 h compared to RUT-C30, indicating the secretion of mCherry in shake flask cultivation.

The mCherry fluorescence of the cultivation medium was measured as described in section 2.7.2. The results indicated a gradual increase in the mCherry fluorescence present in the culture supernatant of the transformants while the non-transformant RUT-C30 indicated no fluorescence in the culture medium in samples collected at various time points (Fig 4.4). As it can be seen, the transformant mS15 still showed the highest amount of culture supernatant fluorescence which was in accordance to what was observed during preliminary screening as described in Chapter 3 (Fig: 3.8 B), indicating a proficient secretion of mCherry. This finding was consistent to the culture supernatant fluorescence observed in Chapter 3 (Fig 3-8 B). The secreted mCherry fluorescence of the selected transformants was measured from three biogical replicates and are indicated by error bars. The possible correlation between the transformant phenotype, intracellular and extracellular mCherry fluorescence was not explored further.

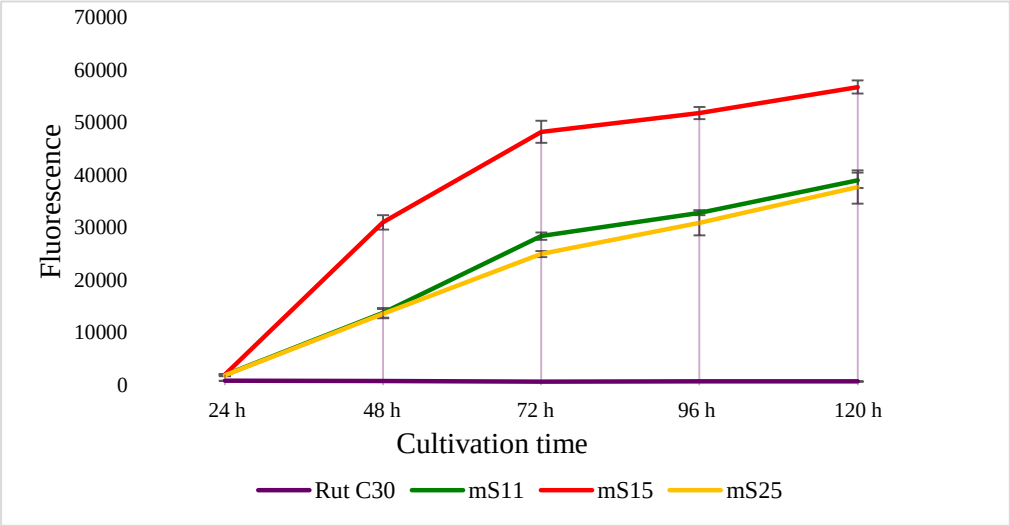


Fig 4-4: Culture supernatants of transformants mS11, mS15 and mS25 showing a gradual increase in mCherry fluorescence over 120 h of cultivation. Excitation/emission wavelength= 560/630 nm. Error bars indicate the standard error of the mean culture supernatant fluorescence and was measured from three biological replicates for each time point.

4.3.3. Profiling of the secreted proteins of *Trichoderma reesei* transformants cultured for 48 h

The total secreted proteins from the culture supernatant of transformants mS11, mS15, mS25, S1, S10 and S16 collected at 48 h from 50 mL shake flask cultures, were analysed by SDS-PAGE (Fig 4-5). Coomassie blue staining of the SDS-PAGE gel revealed the presence of distinct protein bands in the region between 80 and 110 kDa and at the 40 kDa region (shown in boxes a and b, Fig 4-5 A) in the mS11, mS15 and mS25 transformant samples which were absent in the RUT-C30 non-transformant (Fig 4-5 B). The protein bands observed at the 40 kDa molecular weight region could be the cleaved rST6Gal1 and the bands between 80 and 110 kDa may possibly be the remaining fragment of the fusion protein consisting of CBHI core-linker and mCherry. The weak protein band that was visible between 80 and 110 kDa molecular weight region in the gel lane containing RUT-C30 supernatant in Fig 4-5 A was not as prominent as the bands observed in the transformants and might have come from the accidental spillage from the adjacent well containing the sample from mS11.

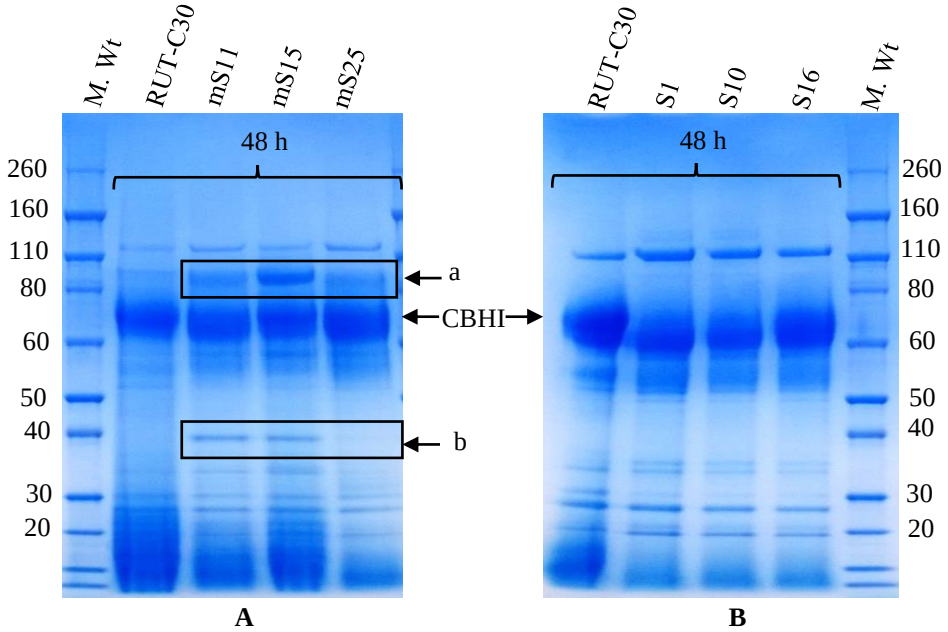


Fig 4-5: Coomassie blue stained SDS-PAGE gels of culture supernatants of *T. reesei* transformants **A:** mCherry containing transformants **B**: non-mCherry transformants. The boxed areas indicate protein bands (a) between 80 and 110 kDa and (b) at 40 kDa found in mCherry containing transformants. The strong band appearing at around 60 kDa molecular weight region is the *T. reesei* cellobiohydrolase 1 (CBHI).

Initial observation of the Coomassie stained SDS-PAGE gel (Fig 4-5 B) did not return any evidence for the secretion of the CBHIcore-linker-ST6Gal1 fusion protein by the non-mCherry transformants that was expected at a molecular weight of 93 kDa. It is possible that the CBHIcore-linker-ST6Gal1 fusion protein produced by the non mCherry transformants S1, S10 and S16 was lower than the detection limit of Coomassie blue staining. It can also be speculated that the presence of rST6Gal1 was masked by the native *T. reesei* RUT-C30 proteins that appeared at the same molecular weight region. Cellobiohydrolase 1 (CBHI), the most abundant protein secreted by *T. reesei* was also seen at a molecular weight close to 60 kDa (Paloheimo *et al*., 2007) indicating that the expression cassette may not have integrated into the *T. reesei cbh1* locus.

4.3.4. Confirmation of the production of ST6Gal1 in shake flask culturing by Western blotting

Western blotting was performed with culture supernatants collected after 48 h from the samples cultivated in 50 mL CLS medium shake flasks. The results. obtained from the Western blot are shown in Fig: 4-6 and were partly in accordance with the observation made in the preliminary analysis (Fig 4-1).

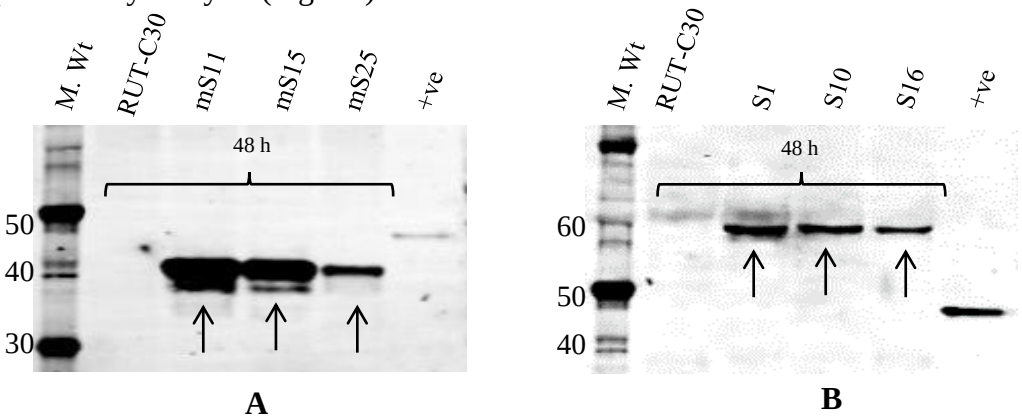


Fig 4-6: Confirmation of expression of recombinant ST6Gal1 in 50 mL shake flask culture after 48 h by Western blotting. **A:** Confirmation of recombinant ST6Gal1 expression in the mS11, mS15 and mS25 transformant (containing mCherry) culture supernatants. The molecular weight of rST6Gal1 secreted was 40 kDa (approximately). **B:** Confirmation of recombinant ST6Gal1 expression in the S1, S10 and S16 transformant (non-mCherry) culture supernatants. The molecular weight of rST6Gal1 secreted was 60 kDa (approximately). Commercially available recombinant human ST6Gal1 was used as positive control for detection. The arrows indicate the rST6Gal1 detected by Western blotting.

The presence of conspicuous bands in the Western blot (Fig 4-6) using anti-ST6Gal1 antibody revealed that the rST6Gal1 was expressed and secreted into the culture supernatant. However, from Fig 4-6 A, the protein band observed in the culture supernatant of the mCherry containing transformants was of approximately 40 kDa (as also seen in the preliminary analysis, Fig 4-1 A) similar to the calculated theoretical mass of 43.5 kDa for the r ST6Gal1. On the other hand, the size of rST6Gal1 secreted by the non-mCherry transformants S1, S10 and S16 was approximately 60 kDa. This could be the reason why a distinct band was not visible in the Coomassie stained SDS-PAGE gel of the non-mCherry transformant, wherein 60 kDa recombinant protein might have been hidden by the *T. reesei* CBHI protein present at the same molecular weight region (Fig 4-5 B).

It may be noted that the anti-ST6Gal1 antibody binds to the C-terminal catalytic domain suggesting that the catalytic domain may still be intact and the cleavage might have occurred in a region close to the N-terminus of rST6Gal1 resulting in the protein with a molecular weight of approximately 40 kDa. The presence of two distinct bands parallel to each other was observed in the Western blot images of mCherry containing transformants (mS11, mS15 and mS25) indicating a variation in processing of the rST6Gal1 (also seen in the preliminary analysis, that was discussed previously, Fig 4-1 A). However, the appearance of the 60 kDa protein was unexpected since none of the expected cleavage sites present in the fusion protein will yield a processed protein with a molecular weight of 60 kDa. Therefore, it can be speculated that cleavage might have occurred in a cryptic location within the CBHI core-linker sequence producing a larger sized protein.

Moreover, the intensity of rST6Gal1 bands in the Western blot of mCherry containing transformants mS11, mS15 and mS25 was stronger compared to S1, S10 and S16. It may be noted that the transformants mS11, mS15 and mS25 were picked for further study based on the mCherry fluorescence. This inference suggests that expressing ST6Gal1 in fusion with a

fluorescent protein such as mCherry was in fact helpful in screening the fungal transformants for rST6Gal1 production based on mCherry fluorescence.

Interestingly, a comparison between the Western blot images given in Fig 4-1 B and Fig 4-6 B, indicates that there was difference in the molecular sizes of rST6Gal1 detected in the 3 mL culture and 50 mL cultures respectively. While the 3 mL culture indicated the presence of rST6Gal1 at 40 kDa (approx.) region, the 50 mL culture indicated it at the 60 kDa (approx.) region. Meanwhile, this variation in the size of rST6Gal1 was unexpected and was confirmed by repeating the Western blotting several times. One potential reason for the discrepancy in the size of rST6Gal1 may be attributed to oxidative stress that the cells underwent due to poorer aeration in the 3 mL culture volume. It may also be speculated that the activity of extracellular proteases may have been different in the 3 mL culture compared to the 50 mL culture, possibly due to a difference in pH, thereby, rendering a shorter fragment of rST6Gal1 in the 3 mL cultures. However, a more fitting explanation for this phenomenon cannot be provided at this point of time since this was not studied further.

4.3.5. Monitoring the production of ST6Gal1 in shake flasks at various time points

Having established the secretion of rST6Gal1 after 48 h of cultivation, the next step was to investigate secretion of rST6Gal1 as a function of time. Samples mS15 and S16 were selected, as an example, based on the observations made from the previous analyses and were cultivated for a period of 120 h in 50 mL CLS medium in shake flasks. The culture supernatant was collected at regular intervals of 24 h. The results are shown in Fig 4-7.

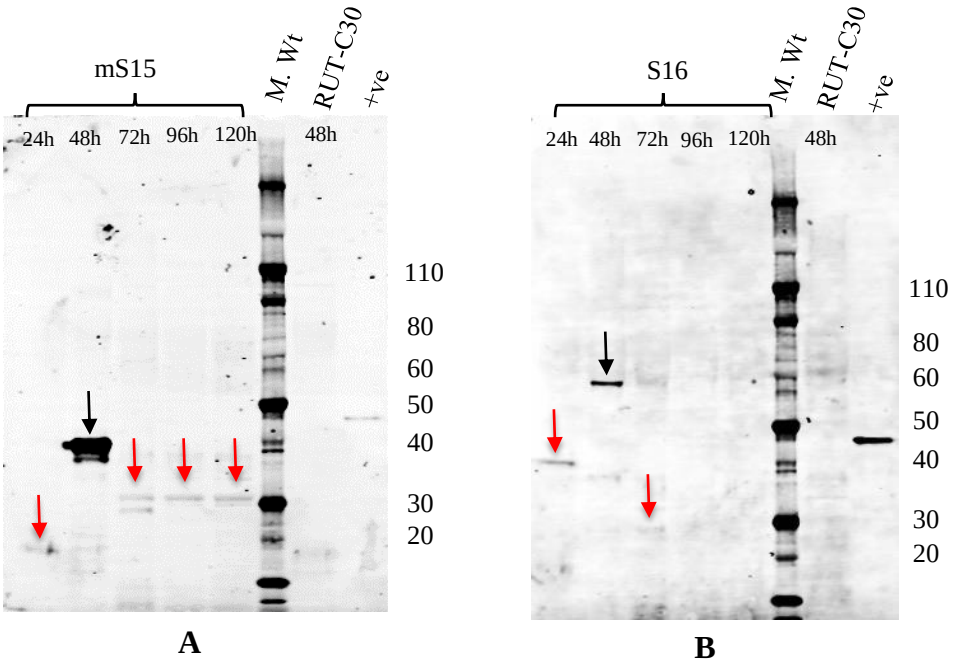


Fig 4-7: Western blots indicating the presence of recombinant ST6Gal1 at 48 h (indicated by black arrows) in **A**: mS15, **B**: S16**.** Weak presence of lower molecular fragments of recombinant ST6Gal1 can be observed at other time points indicating degradation of the protein (shown by red arrows). Commercially available recombinant human ST6Gal1 was used as positive control for detection.

Western blot revealed a strong presence of the rST6Gal1 in culture supernatant collected from mS15 and S16 transformants at 48 h of cultivation. However, lower molecular weight fragments were identified in the mS15 samples collected at 24 h, 72 h, 96 h and 120 h and in the S16 culture medium samples collected at 24 h and 72 h (indicated by red arrows in Fig 4-7). These observations indicated that the rST6Gal1 secreted into the culture medium was disintegrated, possibly due to the activity of extracellular fungal proteases. A similar observation was previously reported by Ribitsch *et al*. (2014) where the seecreted recombinant human ST6Gal1 expressed in *Pichia pastoris* had a molecular size less than the anticipated theoretical size. They also claimed to have detected additional degradation products, presumably by the action of fungal proteases.

4.3.6. Confirmation of ST6Gal1 by mass spectrometry (MS)

In order to validate the results from Western blotting for the presence of rST6Gal1, MS analysis was performed. Also, MS analysis serves as an examination for the presence of rST6Gal in other parts of the SDS-PAGE gel that may have been sterically hindered preventing the antibody binding. Samples collected at 48 h from mS15 and S16 transformants were chosen for further MS examination. Protein bands from the SDS PAGE gel were excised from mS15 and S16 samples as shown in Fig 4-8 A.

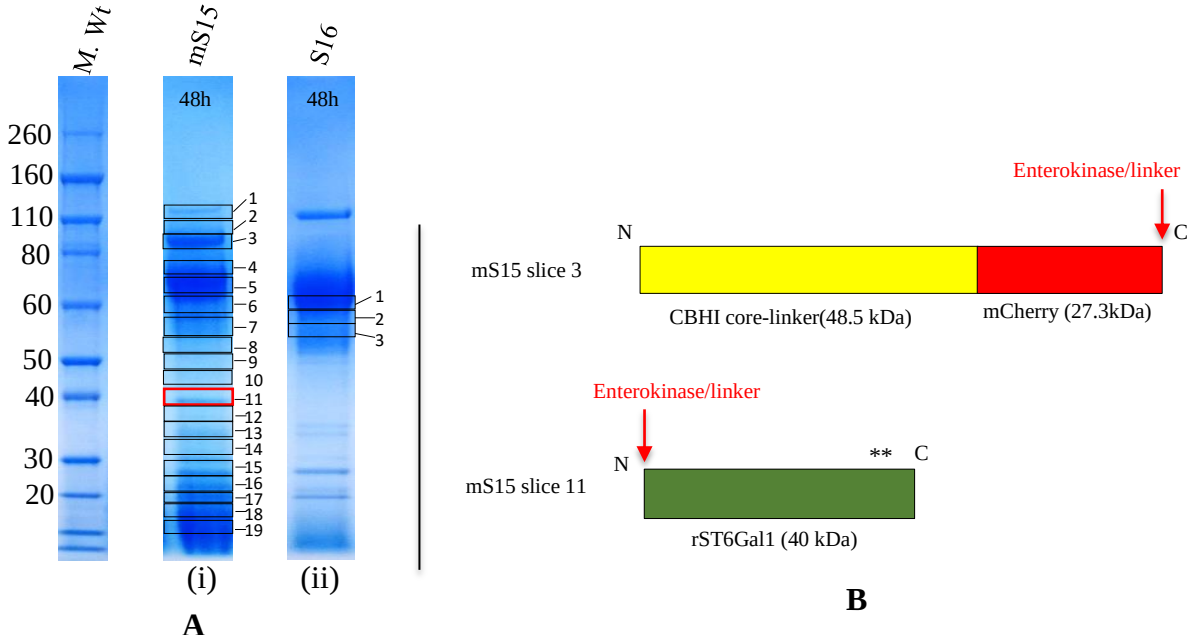


Fig 4-8 A: Protein bands excised from SDS-PAGE gels for mass spectrometry analysis of samples from (i) mS15 and (ii) S16. The presence of rST6Gal1 was confirmed in the protein band from gel slice 11 of transformant mS15 indicated by the red box. **B:** Schematic showing the possible way fusion protein from mS15 has been cleaved. The red arrows indicate the probable site of cleavage (enterokinase or linker). ** indicates the anti-ST6Gal1 antibody binding region. No conclusive evidence was obtained for the validation of rST6Gal1 in the S16 sample by MS (see text for more details).

All visible protein bands from the mS15 sample were excised and analysed by MS. The MS results confirmed the presence of rST6Gal1 in gel slice 11 (red box in Fig 4-8 A (i)) corresponding to a molecular weight of 40 kDa. This observation coincided with the results obtained from the Western blotting, where the rST6Gal1 was detected at a molecular weight

region of 40 kDa (Fig 4-6 A). The proteins identified from MS analysis of slice 11 containing rST6Gal1 from the mS15 transformant is given in Fig 4-9. The smaller the log(e) value, the higher is the confidence of identification of a protein and the confidence of identification of the human ST6Gal1 was very high in the sample analysed.

Rank	Identifier	Protein description	log(I)	rI	log(e)	pI	Mr (kDa)
1	sp|P15907|SIAT1_HUMAN	Human ST6Gal1	6.64	10	-57.8	8.93	45.5
2	sp|P07981|GUN1_HYPJE	Unannotated	4.83	2	-7.5	4.73	48.2
3	sp|Q99034|AXE1_HYPJE	Unannotated	4.62	1	-5.2	5.58	30.7
4	sp|P62694|GUX1_HYPJE	*T. reesei* CBHI	4.66	1	-1.6	4.65	54

A

```
  1   WSHPQFEKIDGGSGKEKKKGSYYDSFKLQTKEFQVLKSLGKLAMGSDSQSVSSSSTQDPH   60
 61   RGRQTLGSLRGLAKAKPEASFQVWNKDSSSKNLIPRLQKIWKNYLSMNKYKVSYKGPGPG  120
121   IKFSAEALRCHLRDHVNVSMVEVTDFPFNTSEWEGYLPKESIRTKAGPWGRCAVVSSAGS  180
181   LKSSQLGREIDDHDAVLRFNGAPTANFQQDVGTKTTIRLMNSQLVTTEKRFLKDSLYNEG  240
241   ILIVWDPSVYHSDIPKWYQNPDYNFFNNYKTYRKLHPNQPFYILKPQMPWELWDILQEIS  300
301   PEEIQPNPPSSGMLGIIIMMTLCDQVDIYEFLPSKRKTDVCYYYQKFFDSACTMGAYHPL  360
361   LYEKNLVKHLNQGTDEDIYLLGKATLPGFRTIHCASEF                        398
```

B

Fig 4-9: Mass spectrometry results confirmed the presence of recombinant ST6Gal1 in the gel slice 11 of mS15 transformant. **A:** Protein accession based on the probability of identification. The rI value of 10 indicates the total number of ST6Gal1 peptides identified. A log(e) value of -57.8 indicates that the confidence level of ST6Gal1 identification was high **B:** Identified peptides of human ST6Gal1 during the mass spectrometry analysis are shown in red letters.

Moreover, subsequent analysis of results from gel slice 3 (corresponding to the box (a) in Fig 4-5) indicated the presence of mCherry and CBHI, suggesting that the rST6Gal1 component was cleaved off from the fusion protein. The confidence of peptide identification (log(e)) was -28.6 and -23.6 for mCherry and CBHI proteins respectively, with four unique peptides identified for each of the proteins. The schematic daiagram represented in Fig 4-8 B depicts the possible manner by which the fusion protein may have been cleaved to generate the free rST6Gal1 that was detected at a molecular weight of 40 kDa. However, the exact location of

the cleavage was not completely understood at this stage. It may be speculated that the possible location of cleavage might be at the enterokinase site between the mCherry and the *Strep*-tag™ II or at the short linker between the *Strep*-tag™ II and the rST6Gal1.

Analysis of the gel slices 1, 2, 4, 5, 6, 9, 12, 13, 14, 15, 16 and 17 (between the 110 kDa and 20 kDa molecular weight region) of mS15 sample revealed the presence of mCherry protein indicating degradation of the fusion protein. The confidence of identification (log(e)) of mCherry in different gel slices were -4 (1), -10.4 (2), -3 (1), -9.9 (2), -6.5 (1), -3 (1), -27 (4), -19.2 (3), -6.5 (1), -20 (3), -20.1 (3) and -33.1 (5) repectively (the number of unique peptides identified are shown in brackets). This result was subsequently confirmed by Western blotting (section 4.3.7). Further, the MS analysis did not return any evidence for the presence of rST6Gal1 in any protein bands at other molecular weight regions.

The rST6Gal1 produced by the non-mCherry transformant S16 was detected by Western blotting at a region close to 60 kDa (Fig 4-6 B). Therefore, it was decided to excise gel slices from the S16 sample corresponding to the 60 kDa region. The gel slices excised from the S16 sample for MS analysis are shown in Fig 4-8 A (ii). However, the data acquired revealed that the MS did not identify rST6Gal1 in any of the three gel slices examined. One possible explanation for this could be the accidental loss of peptides during sample preparation for MS analysis. Another possible reason could be that the most proficiently expressed CBHI protein of *T. reesei*, which was seen at a molecular weight of 60 kDa, might have interfered with the MS detection of rST6Gal1 peptides that might have been relatively lower in abundance, and hence went undetected. The confidence of identification (log(e)) of CBHI protein in the gel slices 1-3 of S16 sample analysed by MS was -178.2 (18 unique peptides), -111.47 (13 unique peptides) and -99.3 (11 unique peptides) respectively.

4.3.7. Western blotting to identify the presence of mCherry

Having identified mCherry from a protein gel band at a molecular size of 80 kDa by MS, the secreted mCherry was probed for in the cultivation medium of the mS15 transformant over 120 h of cultivation time using Western blotting. Culture supernatants collected at 24 h, 48 h, 72 h, 96 h and 120 h were subjected to Western blotting with anti-mCherry antibodies. Cell lysate of Origami™ B (DE3) cells containing recombinant mCherry was used as a positive control for detection (section 2.9).

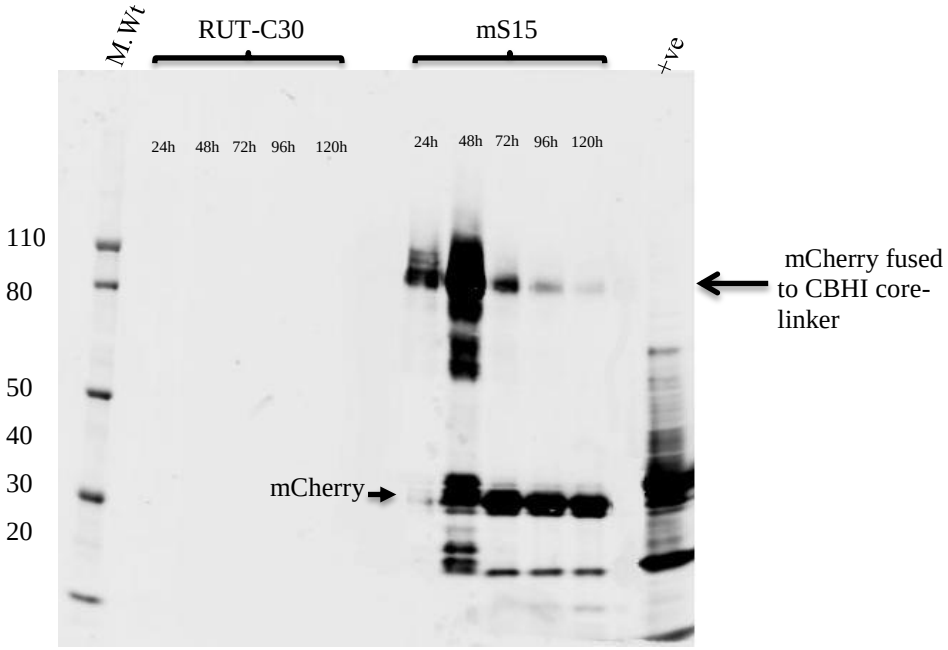


Fig 4-10: Confirmation of mCherry secretion by mS15 transformant using Western blotting indicating the degradation of fusion protein at various time points over the 120 h of cultivation. mCherry (27.3 kDa) is detected around the 30 kDa molecular weight region and mCherry fused to the CBHI core-linker is detected at around the 80 kDa region. Cell lysate from Origami™ B (DE3) cells expressing recombinant mCherry was used as a positive control for detection.

Western bloting (Fig 4-10) revealed the presence of mCherry protein in the culture supernatant of the mS15 transformant at all the time points analysed across 120 h of cultivation. Also, the culture sample collected at 48 h indicated a significantly high amount of mCherry over a wide range of molecular weights and was in accordance with the MS results (section 4.3.6). Moreover, during the course of cultivation for 120 h, mCherry protein was found in the culture supernatant at all time points of sample collection indicating that the

protein was not degraded entirely unlike the rST6Gal1. This suggested that the mCherry might be resistant to degradation in the fungal culture medium. Also, the occurrence of mCherry around the 80 kDa molecular size marker was in turn in accordance with the result obtained from the MS analysis further confirming the possibility of mCherry being fused to the CBHI core-linker particularly in the first 48 hours.

The relative stability of the mCherry protein in the *T. reesei* culture medium compared to the rST6Gal1 might have led to its accumulation over time and hence resulting in a gradual increase in the culture supernatant fluorescence as described in the section 4.3.2. The results obtained may also imply that relying on mCherry fluorescence was not the best approach to monitoring the secretion of the protein of interest, ST6Gal1 in this case, as this would rely on the fusion protein staying together until cleaved by the enterokinase treatment after secretion. Analysis of culture supernatant samples collected after 48 h of cultivation revealed a similar pattern of mCherry detection across mS11, mS15 and mS25 as shown in Fig 4-11.

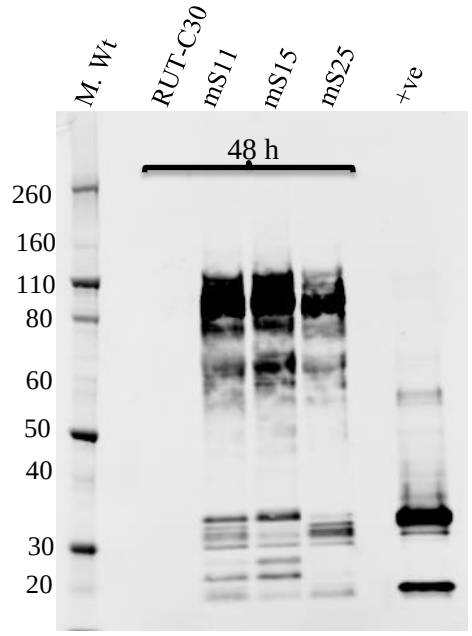


Fig 4-11: mCherry identified at different molecular weight regions in the culture supernatant of mS11, mS15 and mS25 collected at 48 h. Cell lysate from Origami™ B (DE3) cells expressing recombinant mCherry was used as a positive control for detection.

4.3.8. Identification of N-terminal *Strep*-tag™ II attached to rST6Gal1

Having gathered evidence for the cleavage of rST6Gal1, it was necessary to determine if the N-terminal *Strep*-tag™ II was still attached to the sialyltransferase since it facilitates the purification of rST6Gal1. The mCherry and non-mCherry transformant culture supernatants collected at 48 h were subjected to Western blotting with the *Strep*-Tactin® alkaline phosphatase (*Strep*-Tactin®-AP) conjugate. As it can be seen from Fig 4-12 A, the *Strep*-tag™ II was detected at a molecular weight between 110 kDa and 80 kDa and around 30 kDa and below in samples mS11, mS15 and mS25. This finding corresponded to the pattern obtained during mCherry Western blotting (Fig 4-11) indicating that the *Strep*-tag™ II may be attached to mCherry. The Western blot did not reveal any indication of *Strep*-tag™ II at a molecular weight corresponding to 40 kDa at which rST6Gal1 was detected (Fig 4-6 A). This further suggested that the *Strep*-tag™ II was not attached to the rST6Gal1 and the fusion protein might have been cleaved at a location in the N-terminus of rST6Gal1 downstream of the *Strep*-tag™ II.

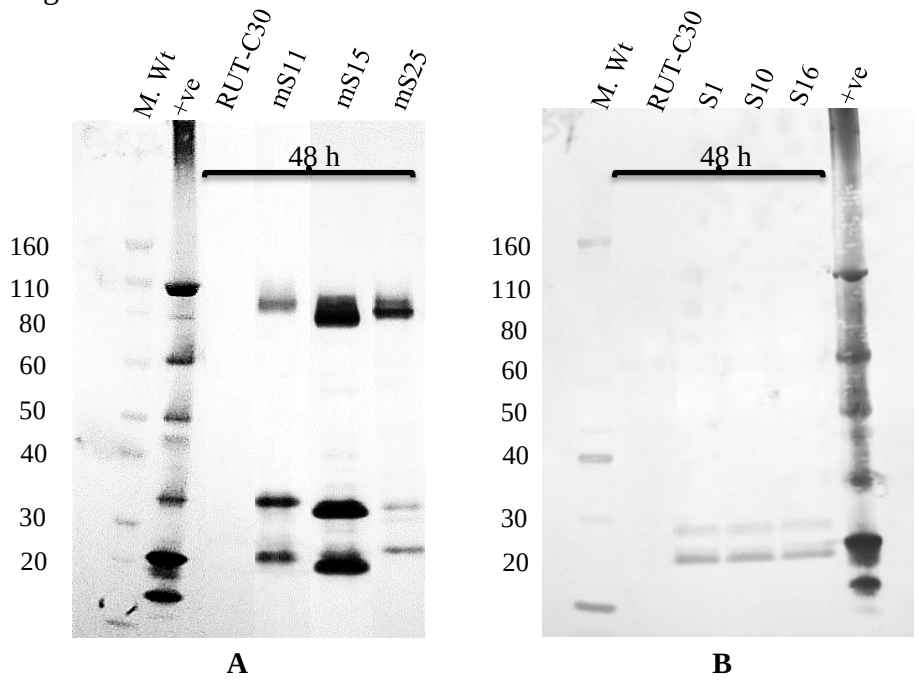


Fig 4-12: *Strep*-tag™ II detection using Western blotting of **A**: Transformants with mCherry showing the presence of *Strep*-tag™ II between molecular mass 80-110 kDa and 20-30 kDa. **B**: Transformants without mCherry showing evidence of *Strep*-tag™ II at a molecular weight region between 20-30 kDa. *Strep*-tag® protein ladder is used as positive control for detection.

From Fig 4-12 B, the *Strep*-tag™ II was not detected at the 60 kDa molecular weight region, where the rST6Gal1 was detected previously (Fig 4-6 B). Moreover, some faint bands were observed at a molecular weight below the 30 kDa marker indicating degraded proteins carrying the *Strep*-tag™ II.

The difference between the molecular weights of the secreted recombinant ST6Gal1 seen in the culture supernatants of mCherry and non-mCherry transformants is not completely understood. However, based on the results presented in this chapter it may be summed up as follows. As previously mentioned, the theoretical molecular weight of ST6Gal1 lacking the N-terminus transmembrane domain has been determined to be 43.5 kDa. Western blotting revealed that the molecular weight of the rST6Gal1 produced by mCherry transformants was approximately 40 kDa. Further, the anti-ST6Gal1 antibody binds to the C-terminal catalytic domain of the protein. This suggests that the cleavage might have occurred at a site close to the N-terminus of the rST6Gal1. Subsequent Western blotting revealed that the *Strep*-tag™ II was not attached to rST6Gal1 further strengthening the possibility of the cleavage downstream to the location of *Strep*-tag™ II. Based on these findings it can be assumed that the cleavage might have occurred at the short linker sequence between the *Strep*-tag™ II and the rST6Gal1 as represented in Fig 4-13 A.

The cleavage of a fusion protein expressed in *T. reesei* in conjunction with the CBHI core-linker has been previously observed and reported by Nyyssönen *et al*. (1993). It is believed that the cleavage of Fab fragment from the CBHI partner of the fusion protein was carried out by a yet uncharacterised fungal protease that cleaves after tyrosine. Furthermore, Smith *et al*. (2014) also reported that the human α-galactosidase A (GLA) expressed in fusion with the CBHI carrier was also cleaved from the CBHI fusion partner. Therefore, it can be assumed that the cleavage of CBHI core-linker might have been resulted by the aforementioned uncharacterised fungal protease. The possibility of cleavage at the KEX2 site in the

pCBH1corlin vector (Fig 2-1) also cannot be ruled out. Based on these assumptions, the possible cleavage locations resulting in the lower molecular weight fragments as seen in mCherry and *Strep*-tag™ II Western blotting of the mS11, mS15 and mS25 transformants can be represented as shown in Fig 4-13 B.

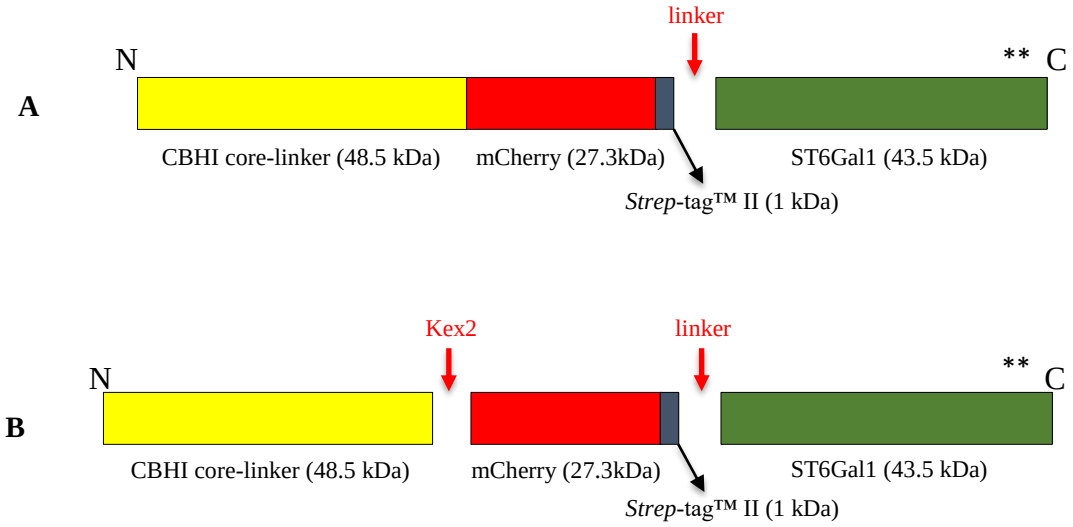


Fig 4-13 Schematic representation of the possible ways of cleavage of the fusion protein produced by mCherry containing transformants mS11, mS15 and mS25. **A:** Cleavage at the linker region cleaving rST6Gal1 from mCherry resulting in the 76.8 kDa fragment consisting of mCherry fused to CBHI-core linker and the 43.5 kDa rST6Gal1. **B:** Cleavage at the linker region and the Kex2 site resulting in smaller fragments detected by Western blotting. ** indicates the anti-ST6Gal1 antibody binding region.

On the other hand, the rST6Gal1 produced by non-mCherry transformants (S1, S10 and S11) was detected at a molecular weight of 60 kDa. However, *Strep*-tag™ II was unidentified in further analysis by Western blotting at a molecular weight of 60 kDa. Two possibilities can be proposed for the lack of *Strep*-tag™ II detection.

Firstly, the *Strep*-tag™ II epitope might have been sterically hindered by the abundance of *T. reesei* CBHI protein (present in a molecular weight of 60 kDa) thereby disabling the anti-*Strep*-tag™ II antibody binding. Proteins tend to form layers on the membrane surface of the Western blot when applied at high concentrations, and they ultimately saturate the membrane resulting in unavailability of the epitope for antibody binding due to steric hindrance (Taylor *et al.*, 2000). Therefore, based on this idea, it can be assumed that the cleavage of the non-

mCherry fusion protein may have occurred at a cryptic site within the CBHI core-linker resulting in 60 kDa protein may have had a part of the CBHI core-linker protein attached to the N-terminus of rST6Gal1 as represented in Fig 4-14 A. A possible way to eliminate the suspected steric hindrance is to selectively remove CBHI protein that runs at 60 kDa region. This may be achieved either by immunopurification using CBHI antibodies or by passing the culture supernatant through a filter paper and expecting CBHI to adhere to the cellulose in filter paper paper. A third and presumambly the best way would be to express ST6Gal1 in a *cbh1* negative transformantion host.

The second possibility is that rST6Gal1 produced by the non-mCherry transformants may have been cleaved at the linker between the *Strep*-tag™ II and the rST6Gal1. However, it can be speculated that during the translation of the sialyltransferase mRNA in the non-mCherry transformants, the stop codon may have been bypassed resulting in a much longer protein. This is as represented in Fig 4-14 B.

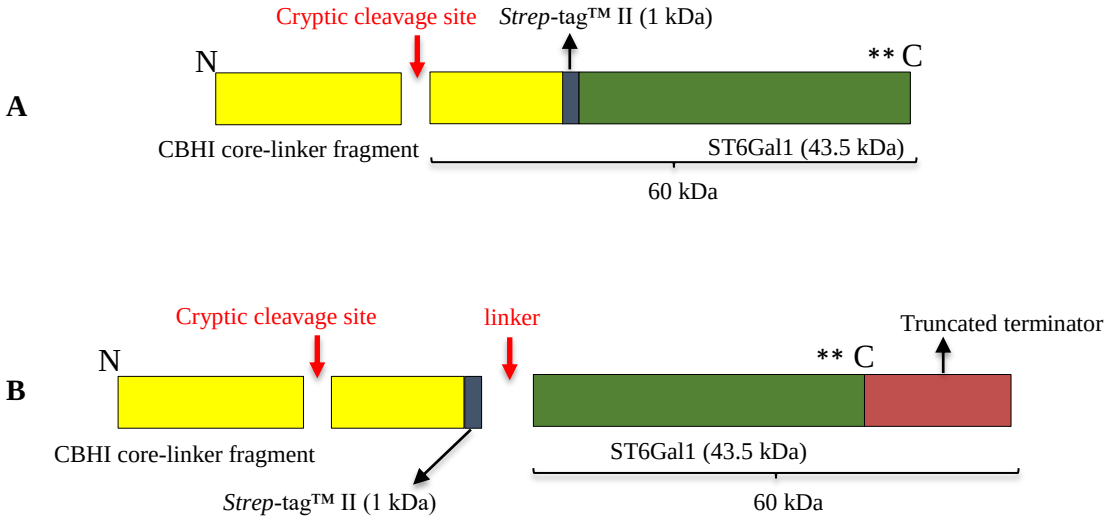


Fig 4-14 Schematic representation of the possible ways of cleavage of the fusion protein produced by non-mCherry transformants S1, S10 and S16. **A:** Cleavage at cryptic cleavage site in the CBHI core-linker region resulting in a 60 kDa protein. **B:** Cleavage at the linker region and the cryptic cleavage site resulting in smaller fragment containing *Strep*-tag™ II detected by Western blotting. The 60 kDa molecular weight protein detected by the anti-ST6Gal1 antibody might have contained the sequence from the truncated terminator (see text for more details). ** indicates the anti-ST6Gal1 antibody binding region.

It has been known since the 1980s that the same gene can yield several mRNAs by variations in splicing (Amara *et al*., 1982; King and Piatigorsky, 1983; Benne *et al*., 1986). Recoding events occurring during translation can challenge the normal decoding rules resulting in the synthesis of multiple related proteins from the same mRNA. The recoding of mRNA is thought to occur in three ways: (a) frameshifting, resulting in two protein products from one mRNA; (b) redirecting the specific stop codons to encode selenocysteine, tryptophan, or glutamine; (c) translating the mRNA by skipping the coding gaps by ribosomes (Gesteland and Atkins, 1996). A study by Namy *et al*. (2003) further described about the bypassing of stop codon during translation in *Saccharomyces cerevisiae*. However, the reason for stop codon bypassing has not been completely understood yet. There are no published articles about this hypothesis in *T. reesei* or other filamentous fungi. Also, the lower molecular weight fragments observed in *Strep*-tag™ II Western blot may have been produced by the cleaved peptide fragment with *Strep*-tag™ II as shown in Fig 4-14 B.

4.3.9. Determination of *ST6Gal1* gene copy numbers using Southern blotting

There was a difference in the amount of rST6Gal1 secreted into the culture medium with the mCherry containing transformants (mS11, mS15 and mS25) having a higher protein abundance compared to the non-mCherry transformants (S1, S10 and S16). Therefore in order to find out whether there is a correlation between the *ST6Gal1* gene copy numbers and the and the production of rST6Gal1, Southern blotting was performed. The transformants chosen for the experiment were the mCherry containing transformants mS11, mS15 and mS25 and the non-mCherry transformants S1, S10 and S16.

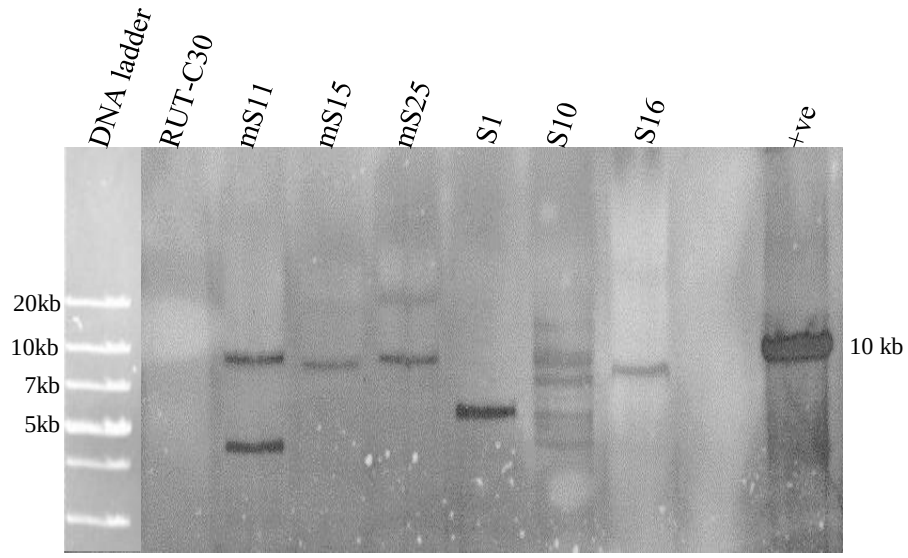


Fig 4-15: Southern blotting of gDNA from the *T. reesei* transformants mS11, mS15, mS25, S1, S10 and S16 (left to right). Each band represents one copy of *ST6Gal1* gene inserted into the gDNA. Plasmid pCBH1corlin containing ST6Gal1 cDNA linearised with *Eco*RV is shown as positive control of detection. Non-transformant RUT-C30 is used as a negative control.

Table 4-1: No. of copies of *ST6Gal1* gene inserted into the genomic DNA of *T. reesei* transformants.

Transformants	mS11	mS15	mS25	S1	S10	S16
Gene copy	2	1	2	1	6	1

The Southern blot (Fig 4-15) indicated a varying number of the *ST6Gal1* gene integrated into the genome of the transformants as shown in Table 4-1. There were at least two copies of *ST6Gal1* gene integrated into the gDNA of transformants mS11 and mS25. The transformant S10 seems to have integrated up to six copies of ST6Gal1 into the fungal genome. There was only one copy of *ST6Gal1* gene inserted into the gDNA of transformant mS15, S1 and S16. Meanwhile, tandem integration of expression cassettes is very common in *T. reesei* (Jovanovic *et al.*, 2014). On this note, the presence of an *Eco*RV restriction site within the nucleotide sequence coding for the CBHI core-linker would result in different sized DNA fragments containing the ST6Gal1 expression cassette; therefore, even though they were tandemly integrated into the fungal gDNA, they would appear as separate bands in the Southern blot.

The transformant gDNA digested with *Eco*RV would yield a DNA fragment of size 6286 bp and 5578 bp containing ST6Gal1 expression cassettes with and without mCherry respectively, if the integration has occurred into the *cbh1* locus. Based on this assumption, subsequent analysis indicated that the integration of *ST6Gal1* gene into gDNA of transformant S1 (without mCherry) was seemingly in the *cbh1* locus. However, from Fig 4-5 (section 4.3.3) the presence of a strong band at 60 kDa region in the Coomassie stained SDS PAGE gel indicates that this assumption might be wrong. Therefore, in order to confirm this finding, additional experimentation *e.g*. using PCR will be required. On an additional note, a study by Te'o *et al*. (2002) had previously shown that the homologous recombination frequency was only up to 5% using biolistic bombardment of *Trichoderma reesei* using a linear DNA. A more targeted DNA delivery method such as CRISPR-Cas9 maybe experimented for ensuring gene integration occurs in the *cbh1* locus (Liu *et al.*, 2015). However, a method such as protoplast transformation followed by heavy metal salt treatment may not be used owing to some disadvantages outlined in section 3.3.3.

The size of the expression vectors linearised by restriction digestion using *Sph*I, before transformation, are approximately 8.2 kb with *mCherry* and approximately 7.5 kb without *mCherry* as shown in Fig 4-16. After subsequent restriction digestion of the gDNA (isolated from the fungal transformants) with *Eco*RV, it is expected that the minimum size of the DNA fragment containing the ST6Gal1 expression cassette detected by Southern blotting is approximately 5.8 kb with *mCherry* (Fig 4-16 A), and approximately 5.1 kb without *mCherry* (Fig 4-16 B). However, it can be seen from the Southern blot image, shown in Fig 4-15**,** that one of the DNA bands in the sample mS11 (containing *mCherry*) and two bands in sample S10 (without *mCherry*) are noticeably smaller than the minimum expected band size of 5.8 kb and 5.1 kb respectively, indicating a possible disruption of the expression cassette during its integration into the fungal gDNA.

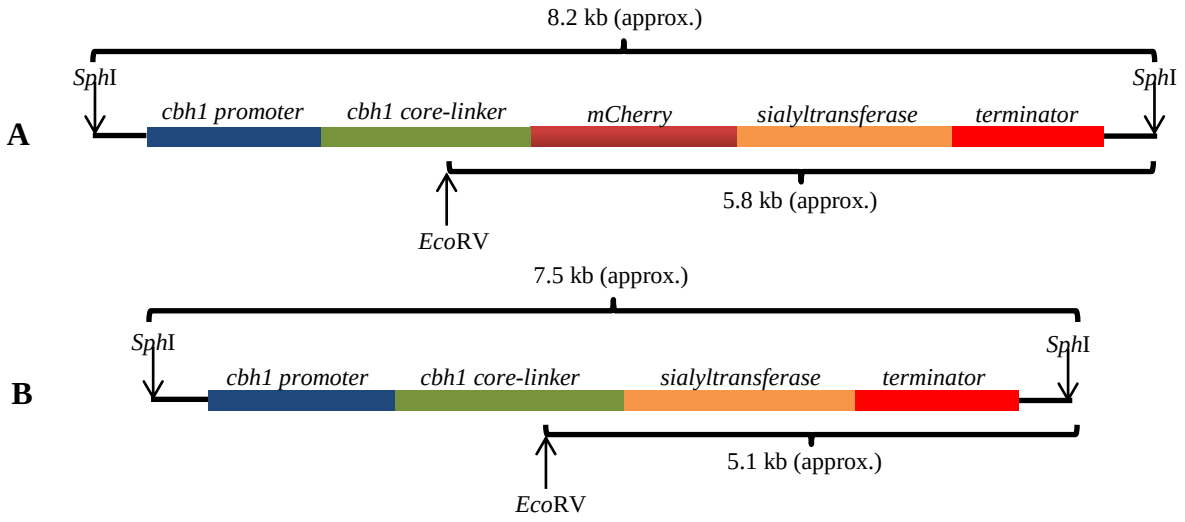


Fig 4-16: Schematic diagram representing the size of the linearised expression vectors using *Sph*I **(A)** vector containing *mCherry* and **(B)** vector without *mCherry*. The minimum expected size of the expression cassettes between the *Eco*RV restriction site within the *cbh1* core-linker region and the *Sph*I restriction site downstream of the expression cassette containing mCherry is about 5.8 kb and about 5.1 kb without *mCherry*.

*Diagram not according to scale

Furthermore, the Western blotting result also revealed that despite the integration of upto six copies of *ST6Gal1* gene into the genome of the S10 transformant, there was no difference in the ST6Gal1 production compared to the transformants S1 and S16. Also, the rST6Gal1 produced by mS15, having only one gene copy, seemed to be much higher than in the S10 transformant. The low level of rST6Gal1 protein produced by the transformant S10 (indicated by Western blot) was not surprising since many studies have indicated that the protein titer does not linearly correlate with the number of gene copies. For instance, it was noted that the titer of heterologous trypsin increased from one copy clone to two-copy clone and decreased significantly with a three-copy clone when expressed in *Pichia pastoris* under the inducible promoter *AOX1* (Hohenblum *et al*., 2004). However, this observation was in contrast to the findings of Sun (2013) who reported a much higher rate of recombinant Obestatin (Obe) expression in *T. reesei* transformants with higher *Obe* gene copy number. Therefore, the results obtained from Southern blotting indicated that the production of rST6Gal1 protein may be dependent on various other factors such as the location of the *ST6Gal1* gene integration in

the fungal genome and effective folding of the heterologous protein, discussed in detail in Chapter 5.

4.3.10. Transcriptional analysis of *ST6Gal1* gene in *Trichoderma reesei* transformants

Transcription of the *ST6Gal1* gene in the fungal transformants mS15 and S16 was explored by employing quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) described in section 2.16. The transformants mS15 and S16 were selected for qRT-PCR analysis because despite each containing only one copy of the *ST6Gal1* gene (section 4.3.9) there was a significant difference in the expression of the rST6Gal1 protein as observed by Western blotting (section 4.3.4). Hence the transcriptional analysis of the *ST6Gal1* gene may reveal if there is any difference in the levels of RNA production in these transformants (mS15 and S16) that had led to the difference in the amount of rST6Gal1 protein produced.

Samples collected at 24 h, 48 h and 96 h were examined for transcription. An initial time point of 24 h was chosen because the induction and mRNA transcription of the *cbh1* gene was reported to have occurred at 24 h in *T. reesei* Rut-C30 when grown in the presence of cellulose or other cellulase inducers such as sophorose or cellobiose (El-Gogary *et al*., 1989).

The qRT-PCR analysis revealed that the mS15 sample collected at 24 h, 48 h and 96 h showed 1.5, 12 and 10 fold higher transcription rate of the *ST6Gal1* gene respectively compared to the S16 samples. This increased transcription level may have been a reason for the increased level of rST6Gal1 protein detected by Western blotting (Fig 4.6). The crossing threshold (C_T) values obtained from the qRT-PCR experiment are provided as supplementary material (Appendix (v)). Furthermore, the expression of *ST6Gal1* gene in mS15 was higher by 1.8 fold at 48 h compared to its expression level at 24 h. Subsequently, the *ST6Gal1* expression level reduced by 4 fold by 96 h of cultivation compared to its expression level at 24 h. At the same time, the expression of *ST6Gal1* gene in the S16 sample reduced by more

than 4 fold and 29 fold in samples collected at 48 h and 96 h of cultivation respectively compared to the gene expression at 24 h (Table 4-2).

Table 4-2: The fold change in *ST6Gal1* expression in transformants mS15 and S16 at 48 h and 96 h compared to expression at 24 h. The green arrow indicates an increase in transcription and the red arrows indicate a decrease in transcription.

Sample / Time point	mS15	S16
48 h	1.857 ↑	4.310 ↓
96 h	4.424 ↓	29.411 ↓

The results derived from the qRT-PCR shows that there is a difference in the rate of expression of the *ST6Gal1* gene in transformants mS15 and S16 even though both had only one copy of gene inserted (section 4.3.9). This inference further indicated that there might be multiple factors associated with the transcription of a gene and subsequent translation to a protein that leads to the variation in its level of expression including the location of the insertion of the gene. However, a more extensive study is required to derive a comprehensive conclusion, some of which will be addressed in the next chapter. The possible bottlenecks in the protein secretory pathway and the protein quality control mechanisms of *T.reesei* will be discussed in detail in Chapter 5.

4.4. Summary

Recombinant human sialyltransferase has been successfully expressed in *Trichoderma reesei* RUT-C30. Transformants mS11, mS15, mS25, S1, S10 and S16 were selected for detailed analyses based on the expression and secretion of rST6Gal1. Expression of rST6Gal1 and its subsequent secretion into the culture medium was confirmed by Western blotting and mass spectrometry. The rST6Gal1 was identified in the culture supernatant at 48 h of cultivation.

The rST6Gal1 seemed to be cleaved from the fusion protein and was detected at a molecular weight of 40 kDa and 60 kDa in the culture supernatant of mCherry and non-mCherry transformants respectively. Also, the rST6Gal1 was detected only at 48 h and was found degraded at other time points.

The mCherry was detected by Western blotting at all the time points in the culture medium indicating that it was more resistant to degradation compared to the rST6Gal1. More importantly, it was noticed that the mCherry was attached to the CBHI core-linker protein and not the rST6Gal1. Notwithstanding this, the amount of rST6Gal1 secreted by the mCherry transformants was higher than the non mCherry transformants. These observations suggested that while the mCherry fluorescence seems a good indicator for selecting a high protein secreting transformant, it cannot be taken as a direct indication of the amount of the rST6Gal1 produced.

Further, the rST6Gal1 seemed to have lost the *Strep*-tag™ II from the N-terminus of rST6Gal1 possibly due to proteolytic cleavage. In order to identify the position of cleavage of the rST6Gal1 from the fusion protein, N-terminal sequencing of a purified protein would have to be carried out. However, since straigthforward purification of the fusion protein using *Strep*-tag™ II was not possible due to loss of the tag, this analysis was not carried out.

The results from Western blotting indicated that the *Strep*-tag™ II was attached to the mCherry protein. This conclusion could be verified by purifying the protein using *Strep*-Tactin® affinity column and subsequently subjecting the purified protein to Western blotting using mCherry antibody.

A Southern blot analysis revealed integration of multiple copies of *ST6Gal1* gene into the fungal gDNA. Transformant S1 seemed to have the integrated the expression cassette containing the *ST6Gal1* gene into the *cbh1* locus suggesting the probability of homologous

recombination. The transformant S10 appeared to have integrated up to six copies of *ST6Gal1* gene into the fungal genome. However, no correlation could be made between the production of rST6Gal1 and the number of copies of *ST6Gal1* gene integrated into the fungal genome.

Subsequent qRT-PCR revealed that the rate of expression of *ST6Gal1* gene was different in mS15 and S16 transformants at 24 h, 48 h, and 96 h despite having only one copy of the gene. The mS15 transformant had a substantially higher expression rate compared to S16 at all the time points analysed indicating that the gene expression may be dependent on other factors such as the location of integration into the gDNA.

5

DEGRADATION OF THE RECOMBINANT ST6GAL1 AND DETERMINATION OF FUNCTIONALITY OF rST6GAL1 PRODUCED IN *TRICHODERMA REESEI*

5.1. Introduction

Despite the well-documented success of heterologous expression of a number proteins in *T. reesei*, the levels of proteins produced vary considerably (Peterson and Nevalainen, 2012). Degradation of heterologous proteins, especially those of mammalian origin, is an ongoing concern for recombinant protein expression in filamentous fungi including *T. reesei*. Proteolysis by intracellular and extracellular proteases, metabolic stress, intracellular protein quality control mechanisms such as the unfolded protein response (UPR) and endoplasmic reticulum associated protein degradation (ERAD) are amongst the proposed bottlenecks suggesting compatibility issues between the host and non-native proteins. The host translational machinery has to cope with the heterologous gene's transcription product in order to achieve optimal expression. Protein incompatibility may be particularly challenging when a simpler eukaryote such as *T. reesei* is used for expression of mammalian proteins, especially of human origin since the host cell should be compatible for the expression of a heterologous protein. The compatibility factors range from codon usage between the host and the donor organism, the ability to fold the protein and to modify the protein after translation (Greene, 2004). However, the mechanism of how the fungus senses an unfamiliar protein is not known.

Protein folding in the ER is monitored by the ER quality control (ERQC) mechanisms and only those proteins that pass the ERQC are subsequently translocated for secretion. The main players in protein folding are the ER resident chaperone BiP substrates (Simons *et al*., 1995) and the foldase PDI, which catalyses the formation of disulphide bonds (Chakravarthi *et al*., 2006). During folding, the glycoproteins interact with another ER resident glycoprotein-specific chaperone system called the calnexin-calreticulin cycle. The calnexin-calreticulin system binds to the carbohydrate moieties of proteins and along with foldases and chaperones promote proper folding of glycoproteins (Simons *et al*., 1995; Chakravarthi *et al*., 2006).

In an event of constant misfolding of a protein, the BiP and the PDI binds and restricts a substrate from further folding or aggregating (Simons *et al*., 1995; Chakravarthi *et al*., 2006). The calnexin-calreticulin binding along with the extensive interaction of foldases and chaperones retains the misfolded glycoprotein in the ER making transport to the ER exit sites difficult (Molinari *et al*., 2004). The accumulation of misfolded or unfolded proteins in the ER elicits the unfolded protein response (UPR). The UPR further triggers the upregulation of the genes important for protein folding (such as *bip1* and *pdi1*), ER expansion, and transport from and to the ER (Heimel, 2015). Subsequently, the unfolded or misfolded proteins that are retained in the ER are degraded by the ERAD pathway.

The UPR pathway has been found to be induced during the expression of heterologous proteins in *T. reesei* and represents a mechanism to preserve the homeostasis of the ER (Arvas *et al*., 2006; Saloheimo *et al*., 2003). Prolonged retention of unfolded proteins in the ER can result in repetitive rounds of attempts for oxidative protein folding resulting in the generation of reactive oxygen species (ROS), and consequent up-regulation of the UPR. The UPR and resulting ERAD can contribute to an elevation in cellular stress and subsequent death (Gasser *et al*., 2008).

During ERAD, misfolded proteins are translocated to the cytoplasm where they undergo disintegration by the ubiquitin-proteasome system (Vembar and Brodsky, 2008). Proteins destined for proteasomal degradation are ubiquitinated (Meusser *et al*., 2005). The ubiquitinated proteins are subsequently translocated to the 26S proteasome via the Sec61 translocon present on the ER membrane (Stolz and Wolf, 2010). A simplified UPR assisted ERAD pathway is depicted in Fig 5-1.

The 26S proteasome complex is composed of two 19S regulatory particles (RP) called the cap and a 20S proteolytic complex called the core particle (CP). The cap or RP is composed of the ATPase (RPT) and the non-ATPase (RPN) subunits. Seven α-subunits make up the two outer

layers and seven β-subunits constitute the two inner rings in the four-layer 20S cylinder structure (Fig 5-1) (Ichihara, 2010; Pereira-Junior *et al*., 2013).

Misfolded proteins that accumulate in the ER can be monitored by confocal microscopy, for example, by genetically tagging the protein of interest with a fluorescent protein such as mCherry and staining the ER with a suitable fluorescent dye. Subsequently, monitoring the transcriptional regulation of the UPR and ERAD associated genes can be used to confirm activity of the ER quality control mechanisms that leads to the protein degradation. Transcriptomic analysis of the genes coding for the proteins BiP1, PDI, calnexin, Sec61, ubiquitin and proteasome subunits RPN8 and beta3 may enable a better understanding of the effect of these mechanisms on the expression of recombinant ST6Gal1 in *T. reesei*.

With an increasing interest in glycobiology and attempts to express recombinant glycosyltransferases underway, practical and rapid methods for the analysis of enzyme activity are essential. The traditional assays for glycosyltransferases involve the use of radioactive donor sugar nucleotide substrates such as CMP-[^{14}C] NeuAc (Rearick *et al*., 1979; Weinstein *et al*., 1982). However, a disadvantage of using radiolabelled donor substrates is the time-consuming step of separating the reaction products from the unused sugar nucleotides by chromatography or gel filtration. In order to address this disadvantage Laroy *et al.* (1997) converted the assays to a microplate format where the labelled product is precipitated, and unused labelled substrates are removed by filtration on glass fibres. However, this method still required radiolabelled donor substrates. Mattox *et al.* (1992) and Yeh and Cummings (1996) developed a solid-phase assay method, which is free of radioactivity and can be used to test sialyltransferase activity in pure and crude samples. The reaction product of the α2-6 sialyltransferase assay described by Mattox *et al.* (1992), is based on the selective affinity of the lectin SNA to α2-6 linked sialic acids. In this project, the solid phase assay method was modified to function in liquid phase and was used to assay the

activity of recombinant ST6Gal1 present in the transformant culture supernatant (Chapter 2, section 2.13 for more details).

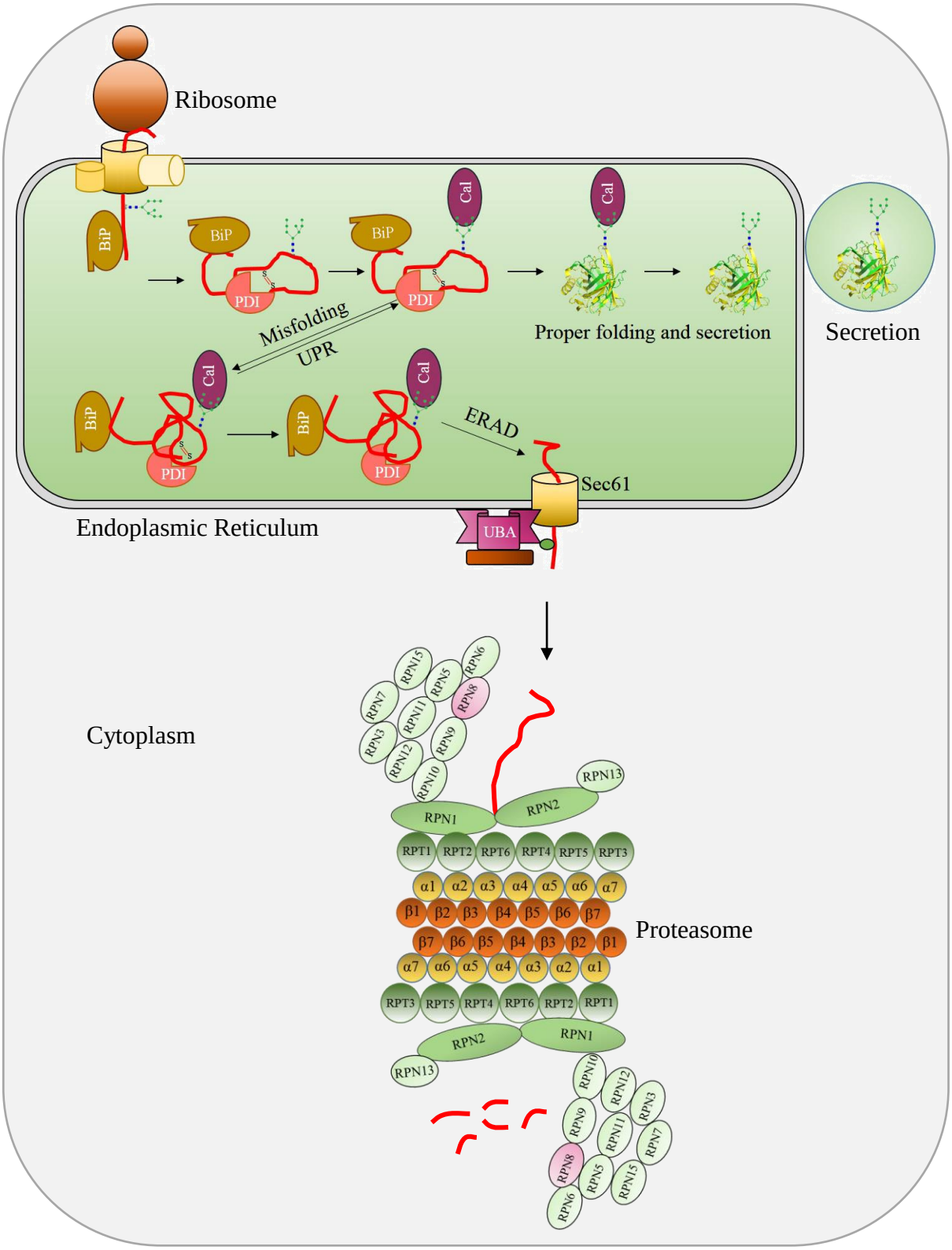


Fig 5-1: Simplified representation of the UPR and ERAD pathway of protein quality control. The newly synthesised protein is folded by the ER resident foldase binding protein 1 (BiP) and the chaperone protein-disulphide isomerase (PDI). The correctly folded protein is translocated for secretion. The misfolded proteins are retained in the ER by the coordinated action of BiP, PDI and calnexin-calreticulin (CAL) until correctly folded or are sent for degradation by the proteasome. A protein destined for the proteasome is ubiquitinated (UBA) and translocated via the Sec61 translocon. The schematic of the proteasome is adapted from Ichihara (2010).

Western blotting (Chapter 4) had revealed that there was more mCherry than rST6Gal1 in the culture supernatant of the mCherry containing transformant. Also, the abundance of rST6Gal1 in the culture medium of mCherry transformants was higher than the non-mCherry transformants. The potential reasons that may have led to the degradation of rST6Gal1 were explored and are discussed in this chapter. Also, the activity of the rST6Gal1 was probed using a suitable activity assay.

5.2. Materials and methods

5.2.1. Extracellular protease activity assay

The total extracellular protease activity assay was performed as explained in section 2.10 in order to investigate the possible reason for the degradation of rST6Gal1 in the culture supernatant. The strains used were mS11, mS15 and mS25 (with mCherry) and S1, S10 and S16 (without mCherry). Protease activity in the cultivation medium was measured every 24 h at a suitable pH using azocasein as substrate. Hydrolysis of azocasein by proteases releases the azo dye, which can be detected by measuring the absorbance at 440 nm.

5.2.2. Confocal microscopy of live cells to establish co-localisation of proteins in the ER

Staining the *T. reesei* cell organelles for confocal microscopy was performed as described in section 2.14 using the mCherry containing transformant mS15. One of the major advantages of live cell imaging over immunofluorescence microscopy is that the cells are not fixed by formaldehyde or ethanol/acetone treatment before microscopy, hence avoiding possible artefacts from the fixation process. Labelling of the cell organelles was carried out to distinguish between membranous organelles in the fungal cell (ER and nuclei). The fluorochrome ER Tracker™ Green is a highly selective stain for the ER of live cells (Wood, 1994). Hoechst 33342 stain was used to stain the nuclei.

5.2.3. Quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Transcriptional levels of genes encoding the selected proteins BiP1, PDI, calnexin, Sec61, ubiquitin and the proteasome subunits RPN8 and beta3 associated with unfolded protein response (UPR) and ER associated degradation (ERAD) pathway, were analysed by qRT-PCR as described in section 2.16. The amplification efficiency of primers was checked by amplifying the corresponding genes using the genomic DNA from non-transformant RUT-C30 as a template. The reactions for determining primer efficiency were performed in duplicate. The test reactions to determine the fold change in the transcription of the corresponding genes were performed in triplicate using reverse transcribed cDNA.

5.2.4. Extraction of intracellular proteins from fungal transformants

The intracellular protein fraction was extracted as described in section 2.6. Cell lysis was performed using a French press. The French press is a strong yet gentle device for breaking hyphae for the extraction of intracellular proteins from filamentous fungi and also introduces less shear to the extracted proteins compared to sonication (Okungbowa *et al.*, 2007). The extraction was performed at room temperature and the cellular fraction was passed through the equipment twice, to maximise the lysis of the cells. The lysed mycelia were centrifuged to remove the cell debris and supernatant that contains the intracellular proteins was stored at -20ºC and subsequently used for Western blotting.

5.2.5. Recombinant ST6Gal1 activity assay

In order to examine the functionality of the rST6Gal1, a lectin based activity assay was executed as described in section 2.13. The activity assay is based on the rationale of the selective affinity of lectins to terminal sialic acids present on a glycan chain (Shibuya *et al.*, 1987; Mattox *et al.*, 1992). In this case, the lectin from *Sambucus nigra* (SNA) is specific for α-2, 6 linked sialic acids present on *N*-glycans. The proposed activity assay can be modified,

depending on the type of linkage of terminal sialic acids, and can be used to assay the functionality of various sialyltransferases in addition to ST6Gal1.

5.3. Results and Discussion

5.3.1. Determination of protein secretion by fungal transformants

The total amount of secreted proteins in the culture supernatant can serve as an assessment of the secretion of proteins by the transformants. The results shown in Fig 5-2 indicate that the protein concentration was the highest at the time of inoculation (0 h). This was expected since the protein from soy hydrolysate in the CLS medium was unused at this stage. After 24 h of cultivation of the transformants, the protein concentration dropped to a lower level, indicating the utilisation of the proteins for the growth and development of fungal mycelia.

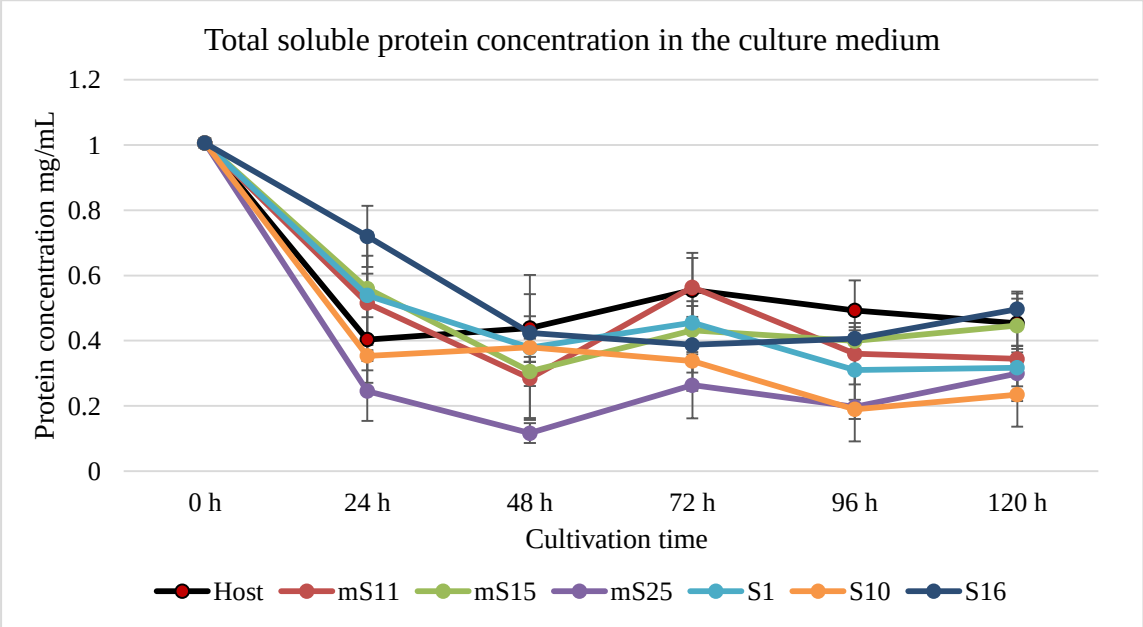


Fig 5-2: Total protein concentration in the cultivation medium of transformants with and without mCherry grown for a time period of 120 h. Error bars indicate the standard error of the mean protein concentration of three biological replicates for each time point.

The variation in the concentration of the proteins may suggest cellular stress with protein folding or transport or variation in cell metabolism that may have been incurred due to heterologous recombination. In a study by Pakula *et al*. (2003), it was identified that when *T. reesei* was treated with chemical drugs such as dithiothreitol (DTT) and brefeldin A (BFA),

they activated the unfolded protein response pathway resulting in the inhibition of secretion of proteins. DTT is a reducing agent and prevents protein folding by inhibiting the formation of disulphide bonds and BFA disrupts the Golgi compartment thereby blocking the transport of proteins in the secretory pathway. Therefore, a decrease in the amount of secreted proteins may be an early indication for the bottlenecks in the secretory pathway.

From 24 h to 120 h of cultivation, the concentration of proteins in the cultivation medium revealed an interesting pattern of fluctuation indicating continuous secretion and degradation of the proteins. The fluctuation or pulsating behaviour in the protein concentration seemed to be continuing at a 24 h interval in all the strains over a 120 h of culturing. A similar pattern in the fluctuation in the amount of secreted proteins has been previously observed (Kautto, 2009; Godlewski *et al*., 2009; Nevalainen and Peterson, 2014a). In general, the overall pattern of protein quantity in the cultivation medium was similar across different transformants and the sample mS25 showed the lowest secreted protein concentration during the course of cultivation. This indicated the probable difficulty of the mS25 transformant in secreting proteins that might have occurred possibly due to the cellular stress incurred by heterologous gene insertion.

At 48 h of cultivation, the transformant mS11 and mS15 had similar concentrations of proteins in the culture supernatant, while mS25 had a lower concentration of protein. A direct correlation can be drawn by comparing this observation to the results obtained from Western blotting (Fig 4-6A). The intensity of rST6Gal1 bands detected by Western blotting in the transformant culture samples from mS11 and mS15 were similar while the intensity of rST6Gal1 band of mS25 was weaker compared to mS11 and mS15. The total protein concentrations of S1, S10 and S16 at 48 h were alike and accorded to the similarity of the rST6Gal1 band intensities in the Western blot (Fig 4-6 B).

5.3.2. Activity of extracellular proteases in the cultivation medium

The secreted rST6Gal1 in the culture supernatant was confirmed at 48 h of cultivation by Western blotting as discussed in Chapter 4. However, the rST6Gal1 was found degraded in the culture supernatant collected at subsequent time points indicating the activity of extracellular proteases. Prior to assaying the activity of extracellular proteases, the pH (Fig 5-3) of the culture supernatant was monitored as the pH of the cultivation medium is a major factor that affects the metabolic pathways, enzyme synthesis efficiency (Adav *et al*., 2011) and the activity of extracellular proteases (Kredics *et al*., 2004).

T. reesei has a strong tendency to decrease the pH of the culture medium when grown on a carbon substrate and then subsequently increase the pH after the exhaustion of the carbon source (Bailey and Tähtiharju, 2003). A similar observation was made in this project, where the pH dropped to a lower point of around 5 at 48 h, from an initial pH value of 6.5, and subsequently increased by 72 h of cultivation (Fig 5-3). Generally speaking, all the transformant cultures exhibited a similar pH curve as the RUT-C30 non-transformant. The lowest pH of the culture medium observed in this project was around 5 as opposed to a pH less than or equal to 4 reported previously (Sun *et al*., 2016).

Total protease activity in the cultivation medium was measured at a pH of 5 for samples collected at 48 h and 6.5 for samples collected at 24 h, 72 h, 96 h and 120 h using three biological replicates. From Fig 5-3 it can be seen that the pH of the cultivation medium at 48 h was close to 5 when rST6Gal1 was detected using Western blotting. It can be speculated that the extracellular proteases that degraded the rST6Gal1 may have had a lower protease activity at pH 5 and higher activity at pH close to 6.5. From Table 1-9 (Chapter 1) it can be seen that the subtilisin like serine proteases are active at a pH of 6 and above. This effect of proteases on the rST6Gal1 may be further investigated by maintaining the pH of the culture medium of the transformants at constant pH (at 5), preferably in a fermenter.

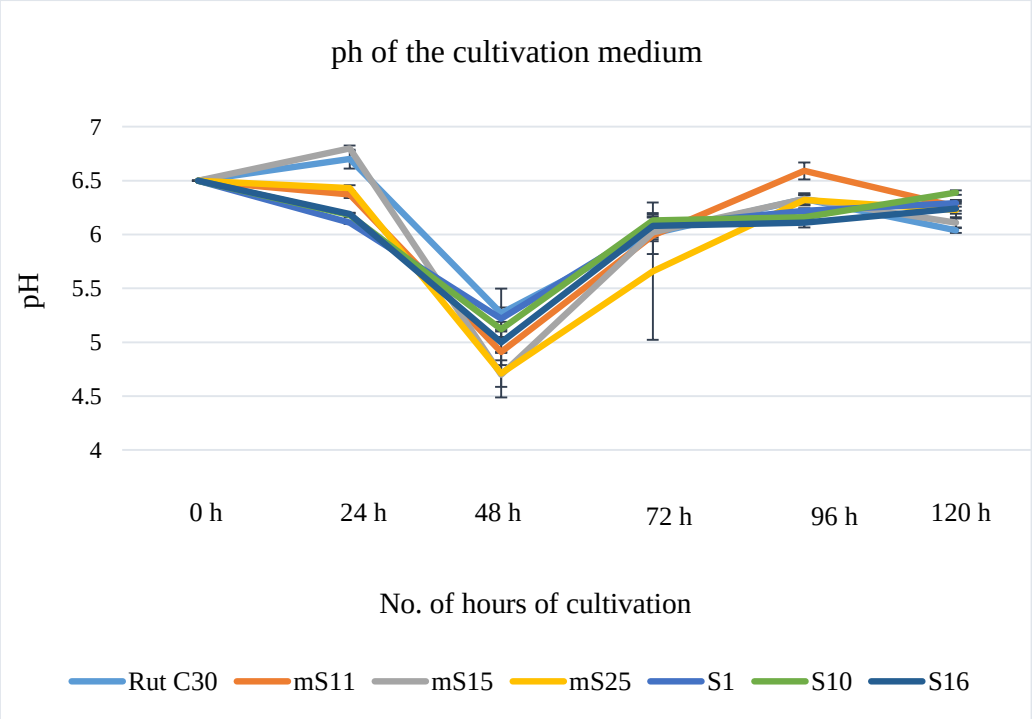


Fig 5-3: Variation in the pH of the cultivation medium (CLS) at different time points. The pH dropped from 6.5 to 5 (approx.) in all the samples by 48 h of cultivation. pH returned to 6 - 6.5 by the time point of 120 h. Error bars indicate the standard error of the mean pH of three biological replicate cultures for each time point.

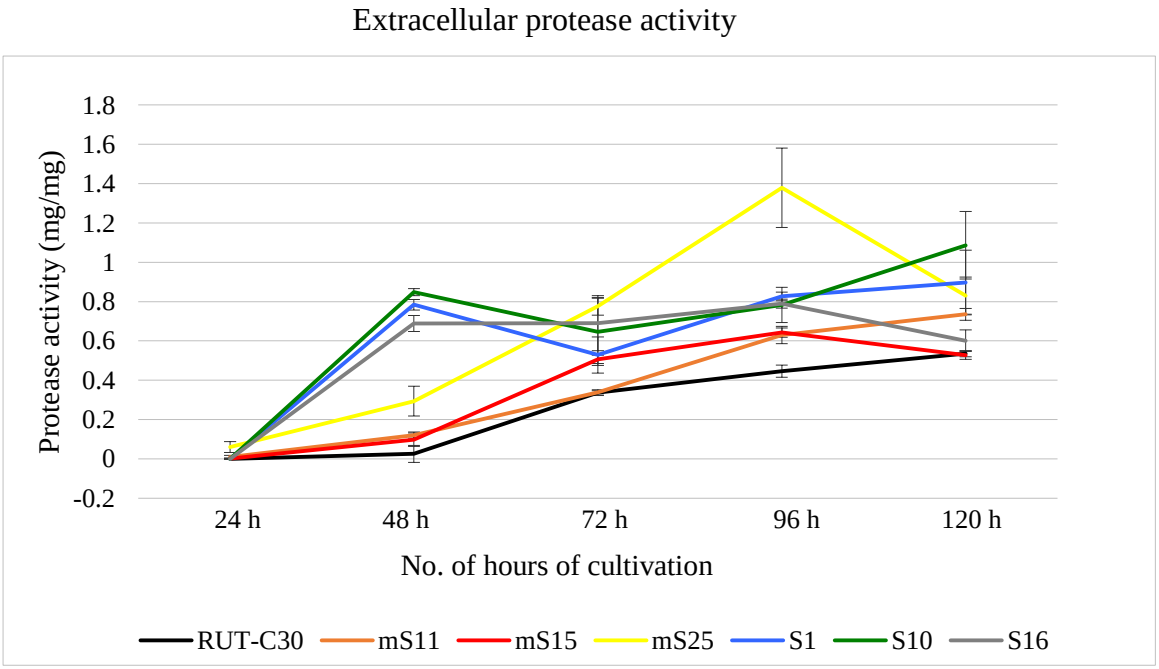


Fig 5-4: Extracellular protease activity in the cultivation medium of transformants showing an increasing trend over 120 h of cultivation. Error bars indicate the standard error of the mean protease activity of three biological replicates for each time point. The protease activity is defined by the amount (mg) of substrate (azocasein) digested by the proteases present in 1 mg of total protein secreted by the fungal transformants.

The mS11, mS15 and mS25 transformants had very low protease activity in the culture medium compared to S1, S10 and S16 after 48 h. This may have been a reason for the higher amount of rST6Gal1 detected in mS11, mS15 and mS25 samples compared to S1, S10 and S16 samples in the Western blots shown in Fig 4-6 A and B (Chapter 4). It is also worthwhile noting that the transformants mS11 and mS15 exhibited the least amount of protease activity, clearly reflecting the amount of protein identified by Western blotting (Fig 4-6 A). The non-transformant RUT-C30 maintained an overall lowest protease activity throughout the cultivation period. A potential reason for the increase in extracellular protease activity by the transformants could be because of an increased protease production as a line of defence to hydrolyse the rST6Gal1 and mCherry protein. There are indications in the published literature that the type and amount of proteases produced by the expression host may change when a foreign protein is produced (discussed in Nevalainen and Peterson, 2014a). For example, Nevalainen *et al*. (2004) reported a nine-fold increase in the amount of a putative aspartic proteinase in a *T. reesei* transformant expressing a heterologous thermophilic xylanase (XynB). Specific protease activity assays can be performed in order to further investigate the protease(s) responsible for the degradation of rST6Gal1.

The total extracellular protease activity assay further revealed an increasing trend of protease activity in the culture supernatant across all the transformant cultures cultivated for 120 h. This increase in the protease activity could be responsible for the subsequent degradation of the rST6Gal1 in the culture supernatant after 48 h, as observed in the Western blots (Chapter 4).

Despite expressing *mCherry* and *ST6Gal1* genes under the same *cbh1* promoter, there was a significant decrease in the amount of rST6Gal1 detected in the culture supernatant by Western blotting compared to mCherry. Therefore, in order to investigate for any bottleneck in the secretory pathway of the fusion protein, especially in the ER, confocal microscopy was employed.

5.3.3. Investigating the intracellular localisation of proteins for degradation

Proteins undergo numerous posttranslational modifications (PTM) in the ER such as (i) the proteolytic cleavage to remove the signal sequence; (ii) proper folding of the proteins by molecular chaperones and the formation of disulphide bonds; and (iii) core glycosylation, before being secreted. Those proteins that fail to undergo proper folding need to be removed from the secretory pathway, as they tend to form aggregates and possibly block the secretory pathway (Peberdy, 1994). The misfolded proteins undergo the ER associated quality control and are eventually removed by the proteasomal degradation via the ER-associated degradation (ERAD) pathway. Potential bottlenecks in the secretory pathway can be examined by co-localisation of fluorescently labelled recombinant proteins within the cell organelles in the secretory pathway by utilising confocal microscopy.

Confocal microscopy images of the transformant mS15 producing rST6Gal1 in fusion to the mCherry protein revealed successful staining of the ER and nuclei in samples collected at 24 h and 48 h as shown in Fig 5-5 and 5-6 respectively. It was decided to stain the fungal nuclei because the identification of nuclei in the cells serves as an indication of a healthy mycelium. Samples collected at 24 h and 48 h showed co-localisation of mCherry in the ER that was revealed by a yellow colour due to the overlapping of green colour from the ER staining and the red fluorescence from mCherry indicating the presence of the fusion protein or mCherry alone in the ER.

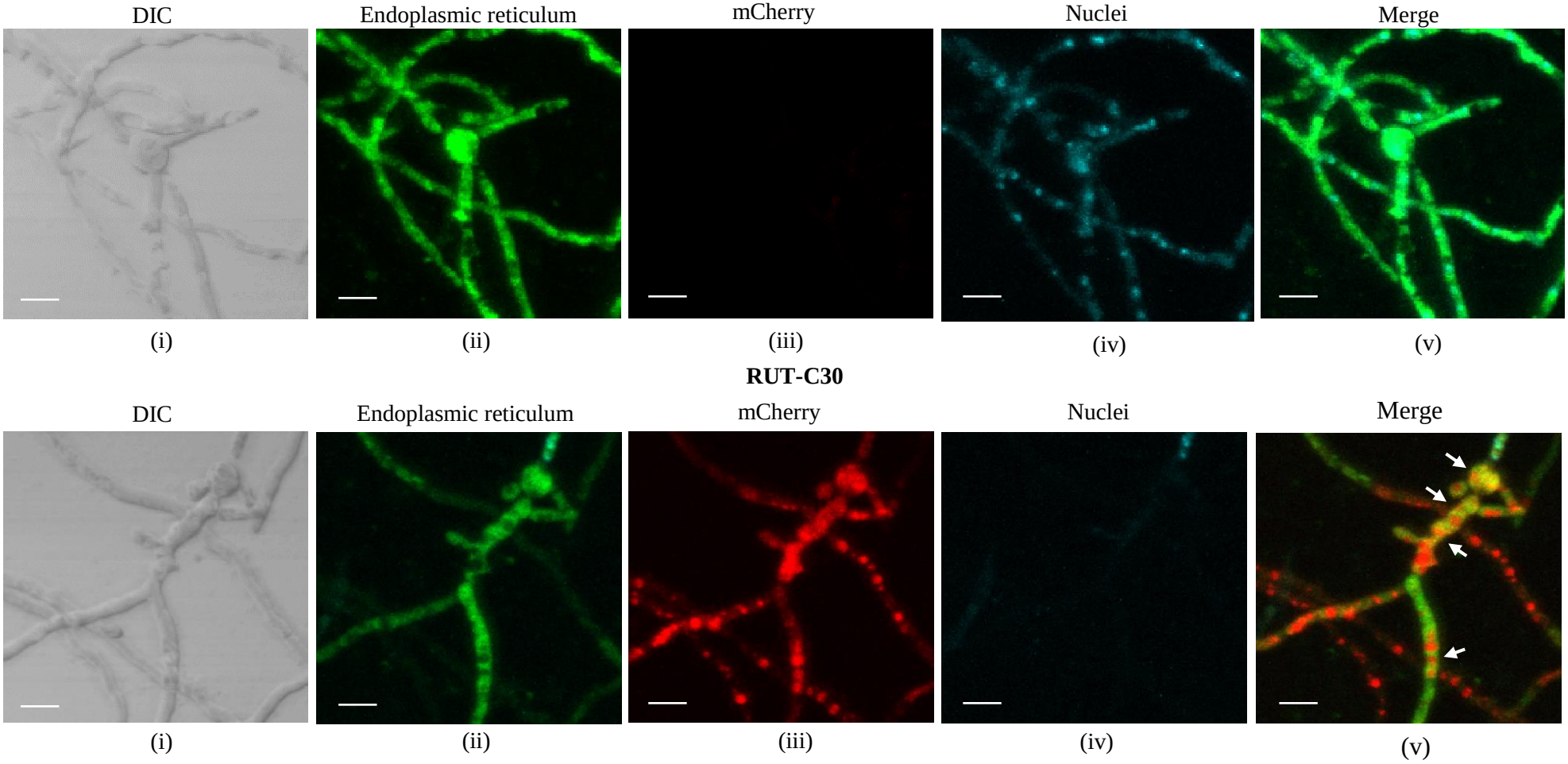


Fig 5-5: Confocal microscopy images of hyphae of RUT-C30 and transformant mS15 cultivated for 24 h. (i) The DIC image for fungal hyphae (ii) The green channel indicates the endoplasmic reticulum (ER) stained with ER-Tracker™ Green BODIPY® FL Glibenclamide, (iii) The red channel indicates mCherry fluorescence visualised in the mycelia, (iv) Nuclei stained with Hoechst 33342 visualised in blue colour and (v) hyphae visualised with all the channels active. It can be seen that there is a significant amount of yellow coloured fluorescence indicating co-localisation of mCherry in the ER. Co-localisation is indicated by white coloured arrows. Scale bars represent 10 µm.

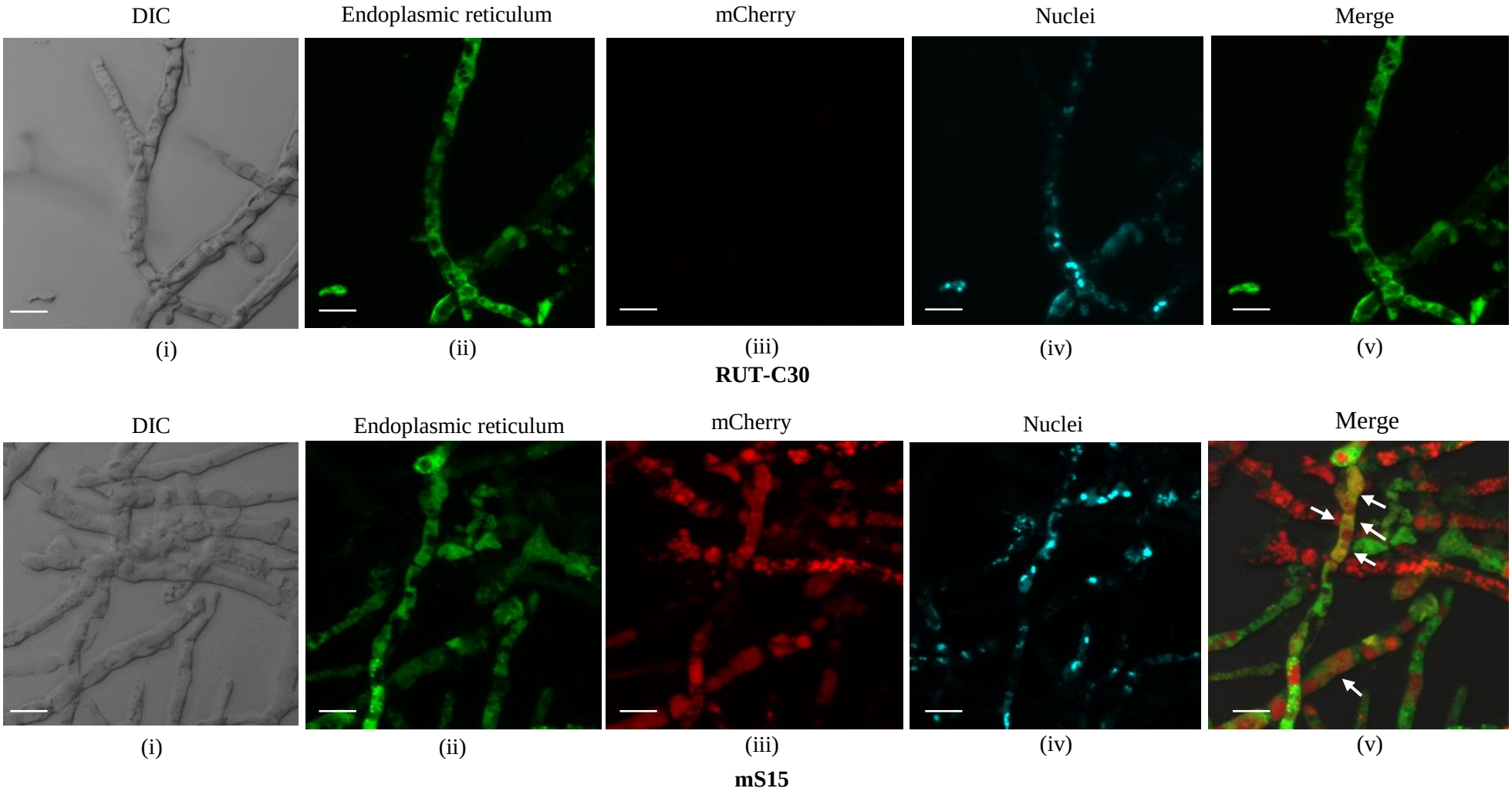


Fig 5-6: Confocal microscopy images of hyphae of RUT-C30 and transformant mS15 cultivated for 48 h. (i) The DIC image for fungal hyphae (ii) The green channel indicates the endoplasmic reticulum (ER) stained with ER-Tracker™ Green BODIPY® FL Glibenclamide, (iii) The red channel indicates mCherry fluorescence visualised in the mycelia, (iv) Nuclei stained with Hoechst 33342 visualised in blue colour and (v) hyphae visualised with all the channels active. It can be seen that there is a significant amount of yellow coloured fluorescence indicating co-localisation of mCherry in the ER. Co-localisation is indicated by white coloured arrows. Scale bars represent 10 µm.

The results indicate that there was co-localisation of mCherry in the ER in hyphae collected at 24 h and 48 h potentially indicating stress in folding and processing of the fusion protein in the ER. A high amount of mCherry occurring in the ER due to expression under the strong *cbh1* promoter may result in 'overcrowding' of the ER by protein. Accumulation of improperly processed proteins, indicated by strong co-localisation, leads to ER stress and subsequent triggering of the ER quality control mechanisms such as UPR and ERAD (Oslowski and Urano, 2011).

However, regardless of seeing co-localised mCherry protein in the ER, there was also an abundance of mCherry visualised as globules outside the ER, along the length of the hyphae at 24 h and 48 h. The abundance of the mCherry protein along the length of the hyphae may indicate that the mCherry was confined in other intracellular organelles of the hyphae such as vacuoles and Golgi bodies. Another possibility is that the already processed fusion protein (in the ER) was localised in the secretory vesicles for subsequent secretion. This can be further studied by selectively staining the vacuoles, Golgi bodies and secretory vesicles using chemical dyes.

Recent studies by Nykänen *et al*. (2016) indicated that the autophagy related structures such as autophagy vacuoles were not obvious in *T. reesei* RUT-C30. Also, it was observed that in a *T. reesei* transformant that overexpressed the recombinant BiP1-Venus YFP fusion protein, the protein was localised in the intracellular punctate bodies. In another study, reported by Kimura *et al*. (2011), the α-amylase expressed in conjunction with mDsRed in an autophagy vacuole deficient strain of *Aspergillus oryzae*, revealed that the misfolded fusion protein was immobilised in large spherical structures surrounded by an ER membrane before they were degraded at a later growth phase. The aforementioned observations had similarities with what was observed in this project.

These revelations may, in turn, indicate that a potentially misfolded fusion protein (rST6Gal1 with mCherry) may have been degraded via the ERAD pathway or were immobilised for future disintegration. In order to explore these options, transcriptional levels of genes encoding proteins associated with the unfolded protein response (UPR) and ER associated degradation (ERAD) pathway were analysed as follows.

5.3.4. Exploring the transcriptional regulation of UPR and ERAD associated genes

The change in the fold of transcription of the genes associated with the UPR and the ERAD pathways relative to actin (housekeeping gene) is given in Table 5-1. The crossing threshold (CT) values obtained from the qRT-PCR experiment are provided as supplementary materials (appendix (vi) to (xiii)).

Table 5-1: The fold change in transcription level of the UPR and ERAD associated genes compared to actin in the transformants mS15 and S16 samples collected at 24 h, 48 h and 96 h. The transformants were cultivated in CLS medium at 28ºC with a constant shaking of 220 rpm. The green arrows and the red arrows indicate upregulation and downregulation of genes respectively. The green dot indicates no fold change in transcription.

Gene / Sample	*BiP*	*PDI*	*CAL*	*Sec61*	*UBA1*	*RPN8*	*Beta3*
mS15 24 h	1.81924 ↑	1.44727 ↑	5.1337 ↑	2.04202 ↑	2.8813 ↑	1.3348 ↑	15.5266 ↑
mS15 48 h	7.04533 ↑	3.79298 ↑	4.5106 ↑	5.53404 ↑	2.07891 ↓	2.46244 ↓	1.90306 ↓
mS15 96 h	1.59844 ↑	1.04488 ●	1.23107 ↓	1.69546 ↓	1.2439 ↓	1.8446 ↑	2.82843 ↑
S16 24 h	1.31039 ↑	1.00696 ●	3.6133 ↑	2.55499 ↓	1.91417 ↑	9.5798 ↑	13.2385 ↑
S16 48 h	3.16015 ↓	5.56607 ↓	4.49034 ↓	2.80891 ↓	7.32332 ↓	2.36183 ↓	3.92218 ↓
S16 96 h	10.90274 ↓	20.58036 ↓	16.36661 ↓	14.91202 ↓	19.49317 ↓	5.1573 ↓	3.63003 ↓

Table 4-2: Relative abundance of ST6Gal1 transcripts mS15 and S16 at 48 h and 96 h compared to its expression at 24 h (Section 4.3.10).

Time point / Sample	mS15	S16
48 h	1.87 ↑	4.310 ↓
96 h	4.424 ↓	29.411 ↓

Analysis of the transcriptional fold change indicated that the UPR associated genes encoding the chaperone BiP and the foldase PDI were highly upregulated at 48 h compared to 24 h and 96 h in the transformant mS15. Such a scenario is often associated with the unfolded protein response (UPR) and indicates the difficulty in processing and folding of the proteins in the ER. The gene encoding calnexin (CAL) was highly expressed at 24 h and 48 h pointing towards the retention of glycoproteins in the ER for assisting with the processing and folding of proteins at 24 h and 48 h.

The ubiquitin activating enzyme (UBA1) and the proteasome subunits RPN8 and beta3 of mS15 were also upregulated by almost one and a half folds and 15 fold respectively at 24 h. The beta3 subunit belongs to the 20S proteolytic core of the proteasome where the disintegration of the misfolded proteins occurs (Fig 5-1). This indicated that the cells were potentially responding to the UPR and preparing for the degradation of the misfolded proteins. However despite an increased *ST6Gal1* expression in the mS15 transformant at 48 h (section 4.3.10), the UBA1, RPN8 and the beta3 were substantially under regulated. This could either mean that the cells had enough ERAD associated proteins to deal with misfolded proteins or the bulk of the proteins produced were properly folded by an increased expression of the chaperones and foldases (BiP and PDI) at 48 h.

The Sec61 channel, through which the misfolded proteins are translocated was found to be up-regulated two fold higher at 24 h and over five-fold higher at 48 h in the mS15 transformant. This might indicate a higher outflow of the misfolded proteins. However, on the contrary, this could also mean that there is more inward flow of newly synthesised proteins into the ER since the Sec61 translocon is also involved in the transport of proteins into the ER from ribosomes (Johnson and van Waes, 1999). Subsequently, by 96 h of cultivation, the transcription levels of genes coding for BiP, PDI, CAL, Sec61 and UBA1 in the mS15 transformant were lower than the other two time points. However, the RPN8 and beta3 were

slightly upregulated possibly to eliminate misfolded proteins in general in order to maintain an overall homeostasis of the cells.

In the case of S16 transformant (without mCherry), the UPR associated genes encoding BiP and CAL were upregulated while PDI had no change in the transcription levels at 24 h. However, the genes encoding UBA1, RPN8 and beta3 were upregulated indicating a strong activity of the ERAD pathway. The upregulation of the gene encoding CAL is a clear indication that the misfolded glycoprotein is retained in the ER for a longer period of time. The expression of the gene coding for Sec61 translocon at 24 h was downregulated, however, despite a very high transcription rate of the proteasomal components RPN8 and beta3. The existing Sec61 channels might have been sufficient to translocate the misfolded proteins to the proteasome for degradation.

Meanwhile, transcription of all the genes regulating the UPR and the ERAD were under expressed at 48 h and 96 h in the S16 transformants. It was previously observed in Chapter 4 (section 4.3.10), that the transcription of the *ST6Gal1* gene was at its maximum at 24 h and compared to 48 h and 96 h in the S16 transformant. One probable reason for the down regulation of the UPR and ERAD associated genes at 48 h and 96 h compared to 24 h can be related to the lower transcription of *ST6Gal1* gene at these time points.

From Chapter 4, it was also observed that the transcription levels of the *ST6Gal1* gene in the mS15 transformant was 1.5, 12 and 10 folds higher at 24 h, 48 h and 96 h compared to the transcription levels in the S16 transformant. One potential reason for the overall up-regulation of the UPR and ERAD pathway could be due to the increased transcription of the *ST6Gal1* gene and the subsequent translation, misfolding and retention of rST6Gal1 in the ER.

The findings described in the previous sections explain the relatively lower amount of rST6Gal1 established in the culture supernatant. However, it is still unclear whether the mCherry and non-mCherry fusion proteins were cleaved inside the fungal hyphae or in the

culture supernatant. The succeeding sections aim in investigating the integrity of fusion proteins inside the cells.

5.3.5. Investigating the expression of intracellular rST6Gal1 in fusion to mCherry

Extracellular fungal proteases are a threat to heterologous protein expression in filamentous fungi. However, proteolysis leading to truncation or degradation of proteins, especially of heterologous origin, can also occur intracellularly within the lysosomes, proteasomes, within the ER or by vacuolar proteases (Pollack *et al*., 2009; Carvalho *et al*., 2011). Having confirmed the co-localisation of recombinant fusion protein in the ER, it was considered worthwhile to explore the intracellular rST6Gal1. This will also serve as a scrutiny for the identification of the anticipated fusion protein inside the hyphae.

It was previously determined (Chapter 4) that the size of the rST6Gal1 detected in the culture supernatant by Western blotting was 40 kDa and 60 kDa for the transforamts mS15 and S16 respectively. The anticipated molecular sizes of the fusion proteins with and without mCherry were 120.9 kDa and 93.6 kDa respectively (Fig 4-2). Western blotting was performed using the intracellular protein fraction from the transformants mS15 and S16, collected after 24 h, 48 h and 96 h the anti-ST6Gal1 antibody (Fig 5-7). Therefore, this will further explain whether the processing of the fusion protein has occurred inside the hyphae or in the culture supernatant.

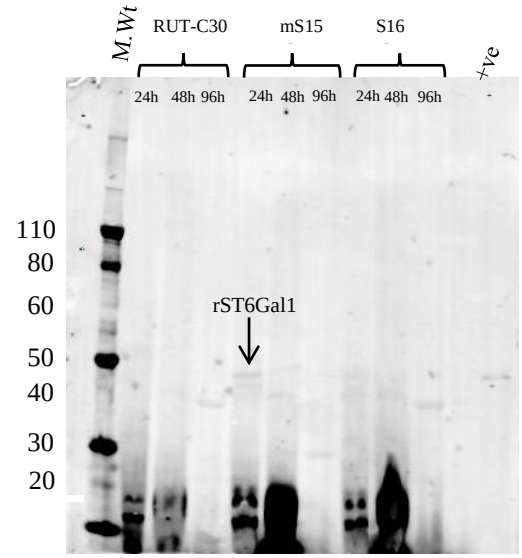


Fig 5-7: Western blot of intracellular proteins of *T. reesei* transformants with anti-ST6Gal1 antibody. Western blotting of the mS15 sample indicated a weak presence of recombinant ST6Gal1 with a molecular weight of 40 kDa (approx.) in the intracellular proteins extracted from hyphae after 24 h of cultivation. Antibody binding was not significant in the S16 sample. Commercially available recombinant human ST6Gal1 is used as positive control for detection.

Western blotting of the intracellular protein fraction revealed a weak indication of the truncated rST6Gal1 in the mS15 sample collected at 24 h. The molecular weight observed was similar to the size of rST6Gal1 (40 kDa), found in the mS15 culture supernatant (Chapter 4) detected in the sample collected at 48 h. However, there was no evidence for the full sized fusion protein in the intracellular protein fraction indicating that the truncation/processing of the fusion protein had already occurred inside the cells. In order to explore this finding, Western blotting using anti-mCherry antibody was performed with the intracellular protein fraction of mS15 collected after 24 h, 48 h and 96 h (Fig 5-8). The mCherry was found in the culture supernatant across all time points analysed (24 h to 120 h) previously (Chapter 4).

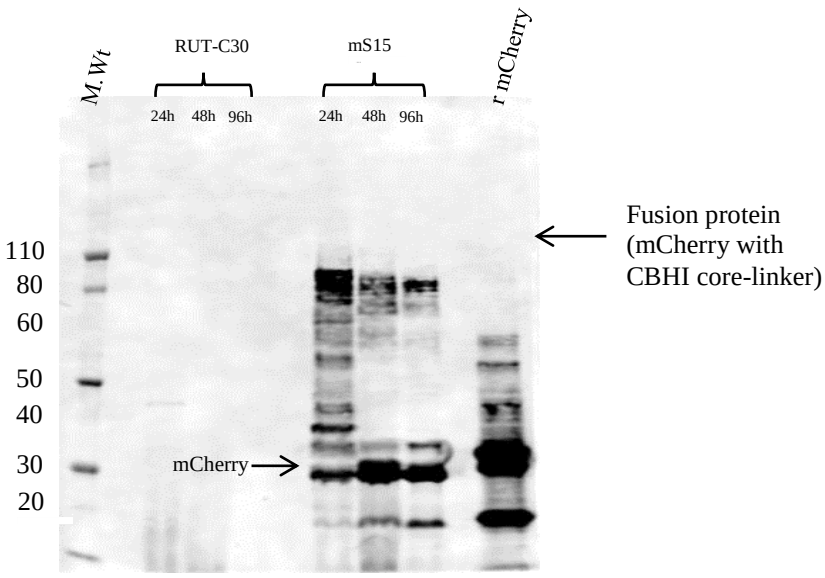


Fig 5-8: Western blot of intracellular fractions of *T. reesei* transformant mS15 with anti-mCherry antibody indicating a high abundance of mCherry. The mCherry was detected at different molecular weight regions indicating intracellular cleavage of fusion protein. Cell lysate from Origami™ B (DE3) cells expressing recombinant mCherry was used as a positive control for detection.

A high abundance of mCherry at different molecular weights was detected in the intracellular fraction. This observation did not accord to the anticipated outcome because ideally the amount of mCherry protein to rST6Gal1 was expected to be in a 1:1 ratio. This observation further indicated that the rST6Gal1 was degraded at a faster rate compared to the mCherry protein. However, there can be more than one reason that can be speculated for the higher relative amount of the mCherry protein inside the cells (and also in the culture supernatant). The recombinant mCherry might have been resistant to the fungal proteases rendering it robust against degradation leading to its detection at all the time points investigated. Previous studies have also indicated that the mCherry is quite stable in proteolytic conditions (Gupta *et al*., 2010; Eiyama *et al*., 2013).

Secondly, the mCherry is a protein with a β-barrel structure that has no cysteine residues and hence is deficient in disulphide bridges (Suzuki *et al.*, 2012). Moreover, there are no predicted *N*-glycosylation sites on the protein. These properties might have aided in the easy processing and folding of mCherry protein (and hence less degradation due to incorrect folding) by *T.*

reesei leading to a stronger presence. On the other hand, the human ST6Gal1 contains two *N*-glycosylation sites and three disulphide bonds (Kuhn *et al*., 2013) and is from a more advanced eukaryote compared to the fungus. Therefore, processing of the rST6Gal1 in the ER may have been cumbersome and potentially inadequate for the human glycoprotein leading to stress build-up and subsequent degradation and hence weaker presence. This can be related to the results from qRT-PCR analysis (Table 5-1), where there was a significant upregulation of the UPR and ERAD associated genes at 24 h in the transformants mS15 and S16, indicating a difficulty in protein folding and subsequent proteasomal degradation.

Finally, the degradation of mRNA during the ribosomal translation by RNA-degrading proteins is also a possibility. The transcribed mRNA that encodes the fusion protein containing mCherry and ST6Gal1 may have been degraded during the ribosomal translation. The degradation of mRNA in eukaryotic cells during translation is thought to proceed in two steps; (i) the translation initiation region on the mRNA is decapped to prevent the assembly of new ribosomal units, (ii) the ribosomes present on the mRNA at the moment of decapping are allowed to conclude the translation leading to incomplete translation of the mRNA sequence. The increasing protein length can reduce the translation rate leading to a decrease in the steady state amount of proteins per mRNA (Nagar *et al*., 2011). This incomplete translation of the mRNA (encoding the fusion protein) by the fungal ribosomes may have led to the higher presence of mCherry compared to rST6Gal1.

On further comparison of the Western blots, it was noticed that the amount of mCherry in the intracellular protein fraction collected was higher at 24 h (Fig 5-8), whereas in the extracellular medium it was stronger at 48 h (Fig 4-10). The intracellular localisation of the mCherry protein may be further attributed to the delay in protein secretion. The qRT-PCR results discussed in Chapter 4 (section 4.3.10) revealed that the transcription of *ST6Gal1* gene was 1.8 folds higher at 48 h compared to 24 h. This could indirectly mean a higher

transcription of the *mCherry* gene also at 48 h, which may perhaps be another reason for the increased presence of mCherry protein in the culture supernatant at 48 h.

Also, the full sized fusion protein produced by translation of the mRNAs may have been proteolitically cleaved by the intracellular proteases leading to the generation of the truncated rST6Gal1 with a molecular weight of 40 kDa (Fig 5-7). In a recent study by Ribitsch *et al*. (2014), it was shown that the rST6Gal1 secreted by *P. pastoris* lacked the polyHis tag due to the proteolytic cleavage of the N- terminus of rST6Gal1. They further reported to have identified five potential locations in the N-terminus of ST6Gal1 protein that are susceptible to cleavage by fungal proteases. This is shown in Fig 5-9.

```
 27  KEKKKGSYYDSFKLQTKEFQVLKSLGKLAMGSDSQSVSSSSTQDPHRGRQTLGSLRGLAK   86
 88  AKPEASFQVWNKDSSSKNLIPRLQKIWKNYLSMNKYKVSYKGPGPGIKFSAEALRCHLRD  146
148  HVNVSMVEVTDFPFNTSEWEGYLPKESIRTKAGPWGRCAVVSSAGSLKSSQLGREIDDHD  206
208  AVLRFNGAPTANFQQDVGTKTTIRLMNSQLVTTEKRFLKDSLYNEGILIVWDPSVYHSDI  266
268  PKWYQNPDYNFFNNYKTYRKLHPNQPFYILKPQMPWELWDILQEISPEEIQPNPPSSGML  326
328  GIIIMMTLCDQVDIYEFLPSKRKTDVCYYYQKFFDSACTMGAYHPLLYEKNLVKHLNQGT  386
388  DEDIYLLGKATLPGFRTIHC                                          406
```

Fig 5-9: Amino acid sequence of ST6Gal1 lacking the 26 aa long cytosolic tail and transmembrane domain. The grey arrows indicate potential cleavage sites on the N-terminus of ST6Gal1 protein identified by Ribitsch *et al*. (2014).

In addition to this, Ribitsch *et al*. (2014) and Luley-Goedl *et al*. (2016) have reported that the intrinsic resistance of the recombinant ST6Gal1 against fungal proteases can be improved by adding an N-terminal FLAG-tag along with the deletion of a putative proteolytic site Lys^{114}-Asn close to the N-terminus.

Apart from the short amino acid linker that was suspected to be the site of cleavage of rST6Gal1 from the fusion protein, as discussed in Chapter 4 (section 4.3.8), the potential N-terminus cleavage sites reported by Ribitsch *et al*. (2014) and Luley-Goedl *et al*. (2016) also seem to be probable sites of rST6Gal1 cleavage that yielded the 40 kDa protein. However, the

cleavage at the aforementioned locations in the N-terminus of the ST6Gal1 protein by the *T. reesei* proteases was not studied further since it was out of the scope of this project.

5.3.6. Lectin based activity assay to determine the functionality of the rST6Gal1

5.3.6.1. Microtiter plate assay

The lectin based ST6Gal1 activity assay that was used during the course of this project is a colorimetric assay in which the intensity of the colour generated is directly proportional to the activity of the ST6Gal1. Subsequent analysis of acceptor substrates for the presence of sialic acid using mass spectrometry (MS) makes the technique robust and reliable since the samples can be analysed for glycan transfer at a higher resolution. An advantage of the assay over the commercially available assays is its compatibility to be used with a crude cell lysate or culture supernatant. Most of the assays available commercially today mostly rely on stable radioisotope labelled sialic acid donor substrates, require a purified enzyme or are very expensive. For example, the sialyltransferase activity assay kit available from R&D Systems, USA (https://www.rndsystems.com/products/sialyltransferase-activity-kit) is very expensive and requires a purified enzyme.

Since the rST6Gal1 produced in *T. reesei* transformants lacked the N-terminal *Strep*-tag (Chapter 4), it was not possible to purify the protein using *Strep*-Tactin® Sepharose® column. Therefore, the rST6Gal1 activity assay was performed using culture supernatant that was processed and enriched by buffer exchange and vacuum centrifugation (section 2.13.1). The presence of rST6Gal1 was later confirmed in the buffer exchanged culture medium by Western blotting.

Reactions containing bovine fetuin (F) and asialofetuin (AF), immobilised onto the wells, served as positive and negative controls for the presence of α-2,6 sialic acids and specific binding of SNA lectin. A reaction mixture containing asialofetuin (AF), recombinant active

human ST6Gal1 (hST6) and CMP-sialic acid (CMP-SA) served as the positive control for the activity assay, *i.e.* an indicator of the expected transfer of sialic acid by ST6Gal1 to the glycans on the protein backbone. The results from the control experiment (Fig 5-10) indicated that the activity assay worked as intended.

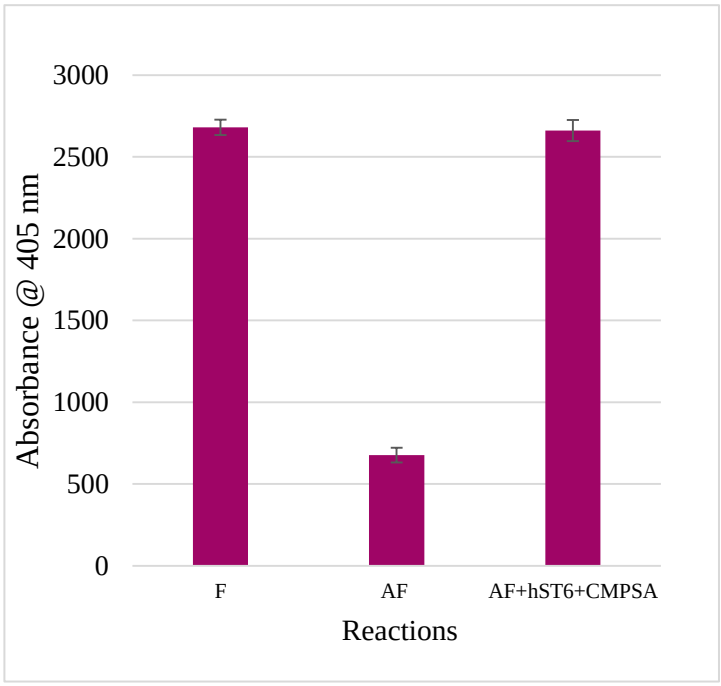


Fig 5-10: Lectin based activity assay using commercially available active recombinant human ST6Gal1 (hST6). F= Reaction with bovine fetuin. AF= Reaction with asialofetuin (desialyltated fetuin). CMP-Sialic acid (CMPSA) served as the sialic acid donor. Reactions with fetuin (F) and assialofetuin (AF) served as control reactions for SNA (lectin) binding. Error bars indicate the standard error of the mean absorbance of three technical replicates. See text for more details.

The terminal sialic acids on the *N*-glycans present on bovine fetuin (F) molecules are linked to the sub-terminal sugar galactose via α-2, 3 or α-2, 6 bonds. About 92.8% of N-glycans in bovine fetuin are sialylated and the ratio of α-2, 6 to α-2, 3 sialylation is determined to be 31:19 (Cointe *et al.*, 1998).

As previously explained, the lectin SNA selectively binds to the α-2, 6 linked sialic acids. From Fig 5-10, the results obtained from reactions with de-sialylated fetuin (asialofetuin, AF) and fetuin (F), provide evidence for the selective affinity of SNA towards terminal sialic acids, wherein, the absorbance reading from a reaction with fetuin was multiple folds higher compared to the reaction with asialofetuin (AF). The absorbance from the asialofetuin may

have occurred due to background signal or from the sialic acids that may have evaded the sialidase treatment (section 2.13.2). Subsequently, analysis of the reaction mixture containing active human ST6Gal1 enzyme (at a concentration of 0.5 µg/100 µL) returned an absorption signal almost equal to the reaction with bovine fetuin (F) as the immobilised substrate. This further indicated that the reaction from commercial enzyme had actively transferred sialic acids onto asialofetuin (AF), almost completely occupying the free acceptor sites. Drawing positive conclusions from the control reactions, the results from the test reactions were analysed (Fig 5-11).

In order to eliminate discrepancies and false signals due to interference from the constituents of the culture supernatant of the samples analysed (Fig 5-11 A), negative control reactions (one for each transformant culture supernatant) without the donor substrate (CMP-SA) was set up (Fig 5-11 B), and passed through the same assay conditions in parallel. Ideally, the absorbance readings from all the sample negative controls would be similar to form a common baseline of absorbance above which rST6Gal1 activity could be measured. Comparing the results from test reactions (Fig 5-11 A), samples from the transformants mS11, mS15, S1 and S10 indicated a higher absorbance reading compared to the transformation host RUT-C30. Subsequent analysis of the results from negative control reactions indicated a variation in absorbance readings across the transformants when compared to the RUT-C30. However, these variations were deemed statistically insignificant by the student's t-test where the probability (p) value determined was over 0.05. In order to further validate the results from the microtiter plate assay, an MS analysis was performed.

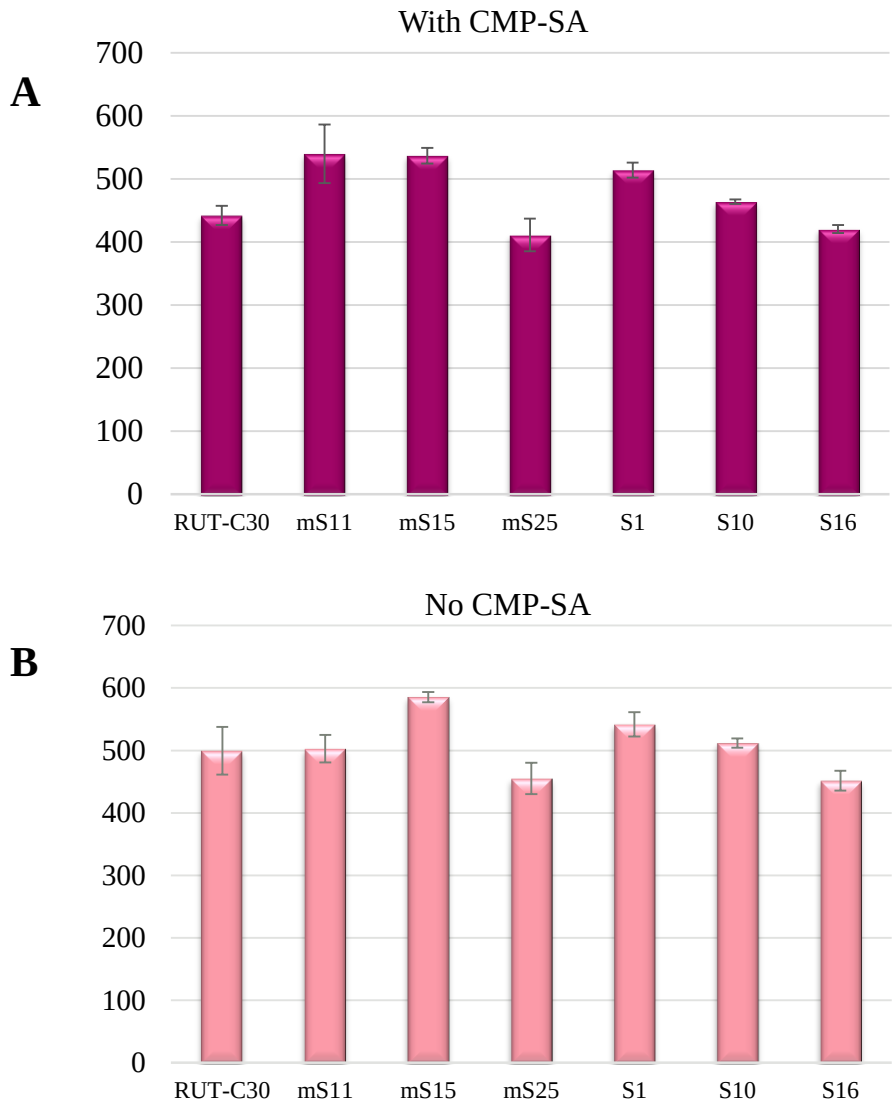


Fig 5-11: Lectin based activity assay using the processed culture supernatant containing recombinant ST6Gal1. **A**: Test reaction mixtures containing sugar donor substrate Cytidine-5′-monophospho-*N*-acetylneuraminic acid (CMP-SA) **B**: Negative control reaction without CMP-SA.

5.3.6.2. Targeted glycan analysis by mass spectrometry

The immobilised acceptor substrates (fetuin, F and asialofetuin, AF) from the reaction mixtures were digested using trypsin, extracted and subsequently subjected to targeted glycan identification by MS analysis as described in section 2.13.3 for the determination of sialic acid transfer. The glycopeptides were analysed using multiple reaction monitoring (MRM) mass spectrometry (MRM-HR) in a TripleTOF 5600 System (APAF, Australia), where

MS/MS spectra are collected at high resolution in a TOF analyser. The fragment ions were then extracted after acquisition to generate high-resolution MRM-like data. The data were subsequently analysed using Skyline to determine sialic acid transfer. The different *N*-glycoforms present on the fetuin peptide sequence LCPDCPLLAPLNDSR, are shown in Fig: 5-12.

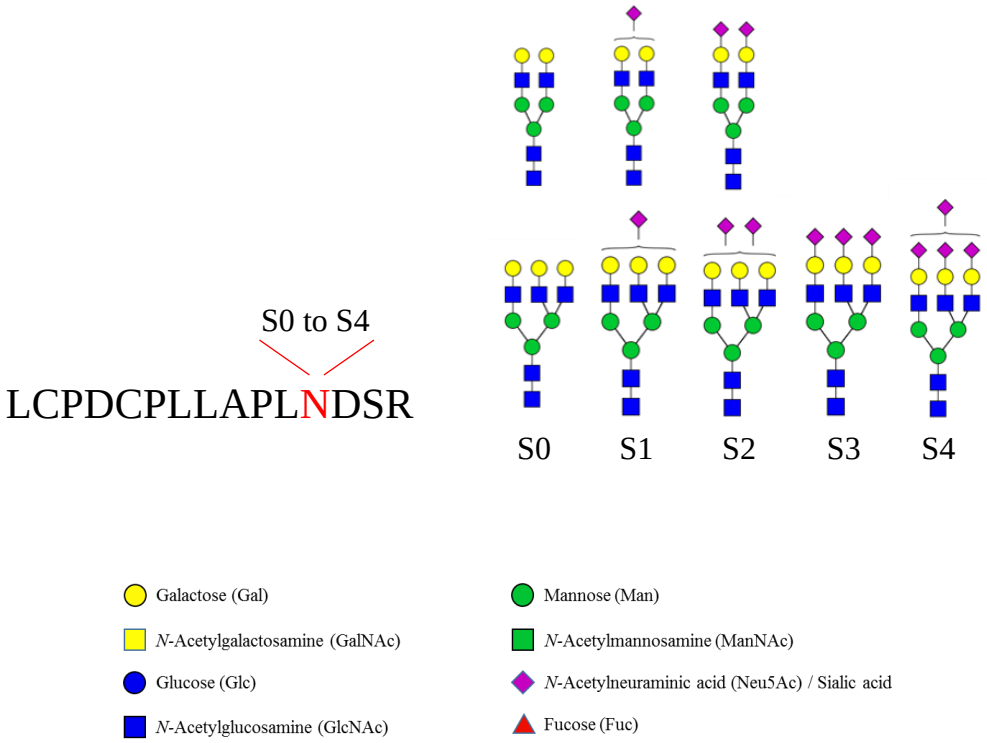


Fig 5-12: The different *N*- glycoforms present on the fetuin peptide LCPDCPLLAPLNDSR that were analysed in the mass spectrometer. The red letter indicates the amino acid asparagine (N) carrying the various glycoforms. S0 to S4 indicates the number of terminal sialic acids present on different glycoforms. Glycan structures were generated using GlycoWorkbench.

Analysis of MS data revealed an active transfer of sialic acid in the reactions containing commercial hST6Gal1. There was an increase in the relative abundance of sialylated glycoforms on the acceptor substrate (asialofetuin, AF) in the positive control reaction containing asialofetuin, CMP-SA and the commercial hST6Gal1 (Fig 5-13). However, the subsequent analysis of the data from the test reactions (with transformant culture supernatants) did not return any conclusive evidence for active sialic acid transfer (results not shown).

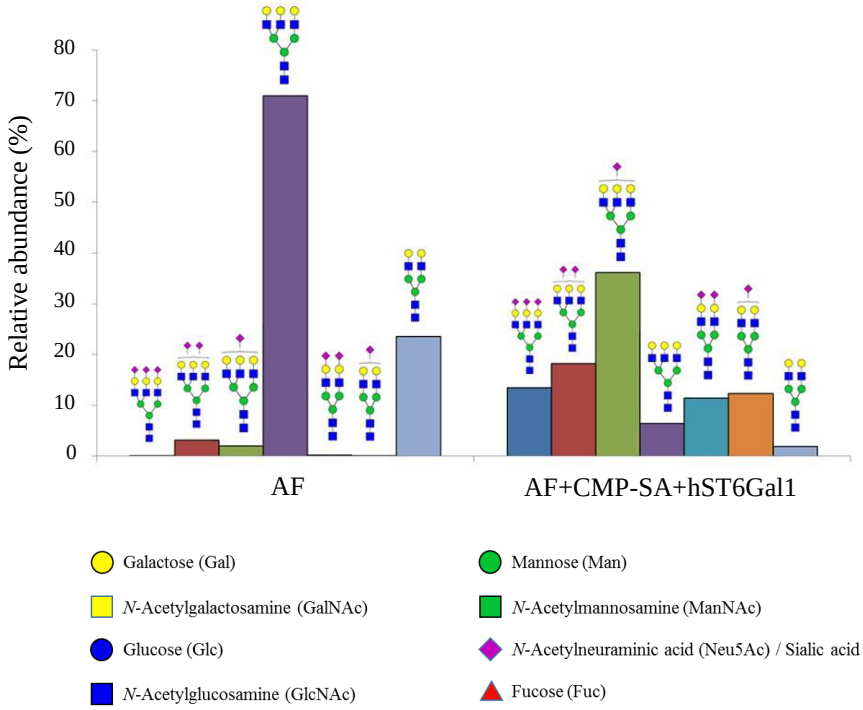


Fig 5-13: MS results indicating the relative abundance of different glycoforms identified on the acceptor substrates from the control reactions. AF indicates the reaction with immobilised asialofetuin (AF) only and AF+CMP-SA+hST6Gal1 indicates the reaction with immobilised asialofetuin (AF), CMP-sialic acid and the commercially available active ST6Gal1 enzyme. Glycan structures are generated using GlycoWorkbench.

One possible reason for the lack of active transfer of sialic acids from the donor (CMP-SA) to the substrate (asialofetuin) by the rST6Gal1 may be associated with the lack of sufficient incubation time, where the samples were incubated only for 5 h. In a study by Thaysen-Andersen *et al.* (2011), a recombinant rat sialyltransferase had to be incubated for up to 24 h to register sialic acid transfer. However, an increased incubation of *T. reesei* culture supernatant containing rST6Gal1 could introduce undesirable results owing to the abundance of proteases in the samples. This could be further examined by performing the experiment with the processed culture supernatant containing protease inhibitors.

As discussed previously in Chapter 1, glycans have been found to have a significant role in protein activity. The *T. reesei* glycosylation may not have been compatible with a protein of human origin, which might have led to its inactivity. It has been previously reported that the presence of a trimannose core with GlcNAc attached was very important for the catalytic

activity of human ST6Gal1 (Fast *et al*., 1993). However, this argument may not be valid since an active human ST6Gal1 was expressed in the yeast *P. pastoris* which doesn't have an advanced glycosylation machinery either (Ribitsch *et al*., 2014; Luley-Goedl *et al*., 2016).

Furthermore, the C-terminus of glycosyltransferases is very sensitive to modification since it generally includes the catalytic domain (Shimma *et al*., 2006). It has also been shown that the loss of one amino acid at the C-terminus of human alpha-1, 3-fucosyltransferase has led to the loss of enzymatic activity (Xu *et al*., 1996). Also, Ribitsch *et al.*, (2014) reported that truncations beyond the 108th amino acid residue of rST6Gal1 (expressed in *P. Pastoris*) in the N-terminus resulted in a loss of enzyme activity. The possibility of loss of amino acids from the N- and C- terminus of the rST6Gal1 cannot be ruled out owing to the random protease activity in the *T. reesei* culture medium, eventually leading to the loss of catalytic activity.

The terminal sialic acids present on the *N*-glycans of bovine fetuin are linked to the sub-terminal galactose residues (Fig 5-12). The galactose residues are exposed by the removal of terminal sialic acids in desialylated fetuin (asialofetuin, AF). The exo-galactosidases of *T. reesei* RUT-C30 present in the culture supernatant (Golubev and Neustroev, 1993; Seiboth *et al.*, 2005; Marx *et al.*, 2013) might remove the galactose residues from the *N*-glycans on the asialofetuin thus eliminating the residues required by the sialic acids to attach during the transfer reaction. This also could have been a reason why the rST6Gal1 activity could not be established by the lectin based activity assay and subsequent MS analysis. The role of exo-galactosidase interfering with the activity of rST6Gal1 can be determined by incubating the active human ST6Gal1 with the fungal culture supernatant. The treated ST6Gal1 can subsequently be assayed for its functionality.

Moreover, the possibility of inactivation of rST6Gal1 produced in *T. reesei* during the sample preparation (section 2.13.1) cannot be ruled out. However, the functionality of the rST6Gal1 can be investigated further with purified rST6Gal1. In this work *Strep*-tag™ II was

incorporated to the 5′ end of the *ST6Gal1* gene with an aim to purify the recombinant protein. However, the loss of *Strep*-tag™ II from the N-terminus of the rST6Gal1 made the purification of rST6Gal1 from the *T. reesei* culture medium difficult. Therefore, purification of the rST6Gal1 was not carried out in this project.

5.4. Summary

The successfully expressed human ST6Gal1 in *Trichoderma reesei* was found to be proteolytically cleaved and disintegrated in the culture supernatant after 48 h of cultivation. This observation was further supported by the detection of an increase in the total extracellular protease activity in the culture medium over a cultivation period of 120 h. More notably, the extracellular protease activity was higher in the culture medium of the transformant strains compared to the non-transformant RUT-C30. Analysis of the intracellular protein fraction did not reveal the presence of the anticipated fusion protein; rather the fusion protein was cleaved to separate rST6Gal from mCherry. The presence of rST6Gal was established by Western blotting in the intracellular protein fraction extracted from hyphae collected at 24 h only. However, the amount of the rST6Gal was found lower compared to the mCherry protein even though both the proteins were expressed under the same *cbh1* promoter. Also, the mCherry protein was identified at other time points analysed indicating its relative stability and resistance against degradation by fungal proteases.

Confocal microscopy analysis revealed the intracellular co-localisation of mCherry in the ER indicating the processing of mCherry-rST6Gal1 fusion protein. Subsequent transcriptional analysis of the genes associated with the unfolded protein response (UPR) and the ER associated degradation (ERAD) pathways revealed an upregulation indicating the intracellular stress in processing and folding of proteins. The relatively lower presence of the rST6Gal1 compared to mCherry might have resulted due to the degradation of the protein by the cellular

quality control mechanisms (UPR assisted ERAD pathway) probably due to the difficulty in processing and folding of the protein from a mammalian origin.

A lectin based activity assay was developed to examine the functionality of the rST6Gal1. The assay was established to be working with the commercially available active human ST6Gal1, which was subsequently confirmed by mass spectrometry analysis. The lack of a *Strep*-tag™ at the N-terminus of the rST6Gal1 made it rather impractical to purify the protein. Therefore, the activity assay was performed using the culture supernatant of transformants which was buffer exchanged and concentrated using vacuum centrifuge in an attempt to enrich the rST6Gal1. However, the assay and the subsequent MS analysis did not return any evidence for the activity of rST6Gal1 expressed in *T. reesei.*

6

SUMMARY, FUTURE DIRECTIONS AND CONCLUSIONS

6.1. Summary of findings made in this project

Recombinant sialyltransferases can be used for *in vitro* modification of glycans on therapeutic proteins to improve their properties such as activity, solubility and half-life and also to reduce immunogenicity. However, the yield of sialyltransferases in the existing production systems such as *E. coli*, *P. pastoris*, *S. cerevisiae*, mouse L cells and COS 7 cells is low with high production cost (Chapter 1). Therefore, a more efficient system for the production of recombinant sialyltransferases is required. The filamentous fungus *Trichoderma reesei* is a suitable candidate for heterologous expression of sialyltransferases owing to its ability to carry out post-translational modification and higher secretion of proteins.

The human sialyltransferase ST6Gal1 was successfully expressed in *Trichoderma reesei* RUT-C30 in this work. The *ST6Gal1* cDNA was expressed under the strong *cbh1* promoter of *T. reesei* in conjunction with an endogenous DNA that encodes the core and linker portion of the cellobiohydrolase I (CBHI) protein. The major outcomes of this project are summarised in this chapter, with suggestions for future work.

6.1.1. Differences exhibited by the *Trichoderma reesei* RUT-C30 transformants (Chapter 3)

Transformants were generated using biolistic bombardment with DNA containing *ST6Gal1* expression cassettes with and without a gene encoding a fluorescent protein mCherry. The expression of ST6Gal1 in fusion with mCherry aided in the screening of high-protein secreting transformants based on the mCherry fluorescence in the culture medium. All the transformants survived three rounds of selection on PDA plates containing hygromycin indicating mitotic stability of the transformants containing the heterologous expression cassette. The transformants generated showed notable differences in the colony morphology,

which served as an early indication for random integration of the expression cassette into the fungal gDNA.

Subsequent analysis of the mCherry containing transformants under a confocal laser scanning microscope further revealed differences in the hyphal morphology compared to the transformation host. While some transformants exhibited lesions on the hyphae, some indicated thinning of the hyphae towards the tips. Most of the transformants displayed normal and healthy looking hyphae. The intracellular mCherry was visualised along the length of the fungal hyphae and appeared as tiny globules indicating that the proteins may be localised in fungal organelles such as ER, Golgi bodies, punctate bodies or secretory vesicles. Certain transformants did not reveal any intracellular fluorescence indicating the inability of these transformants to express the recombinant protein. Further analysis of the culture medium revealed the level of fluorescence varied across the culture supernatants of different transformants. It was also noticed that some transformants with a relatively higher level of intracellular mCherry fluorescence had difficulty in secreting the protein, probably because of intracellular localisation that was seen as globules. The transformants were selected for subsequent analyses based on the level of fluorescence from secreted mCherry.

6.1.2. Human ST6Gal1 was successfully expressed in *Trichoderma reesei* RUT-C30 (Chapter 4)

The transformants that were screened on the basis of the mCherry fluorescence in the culture supernatant were investigated for the production of the recombinant ST6Gal1 (rST6Gal1) by Western blotting using an anti-ST6Gal1 antibody. The rST6Gal1 was detected in the culture supernatant of the transformants with and without mCherry cultivated in 3 mL CLS medium. However, not all the transformants analysed returned indication for the rST6Gal1.

A shake flask cultivation of the transformants selected after the preliminary screening revealed rST6Gal1 at a molecular weight of 40 kDa and 60 kDa in mCherry and non-mCherry transformants respectively in the culture medium at 48 h. This indicated processing of the fusion proteins that were expected at 120.9 kDa (with mCherry) and 93 kDa (without mCherry). However, rST6Gal1 was not detected at any other time points analysed for a period of 120 h of cultivation suggesting its degradation during secretion or in the culture medium. On the contrary, the mCherry protein was detected in the culture supernatant at all the time points at different molecular sizes indicating disintegration of the fusion protein. Also, it was later confirmed that the *Strep*-tag™ II was not attached to the rST6Gal1 indicating a possible cleavage at the N-terminus of the rST6Gal1 or in the short amino acid linker between the rST6Gal1 and the *Strep*-tag™ II.

The amount of rST6Gal1 in the culture medium, as determined by Western blotting, was higher in the mCherry containing transformants compared to the non-mCherry transformants. Therefore, the inclusion of a reporter protein (mCherry) had certainly aided in the screening of transformants that were good secretors of proteins, on the basis of mCherry fluorescence in the culture supernatant. Southern blotting revealed a difference in the number of copies (1-6) of the *ST6Gal1* gene integrated into the gDNA of transformants. One of the transformants seemed to have the *ST6Gal1* gene integrated into in the *cbh1* locus of *T. reesei* (determined by Southern blotting); however, this requires further investigation for confirmation. However, no direct correlation could be established between the number of copies of the *ST6Gal1* gene and the amount of protein detected by Western blotting. A subsequent transcriptomic analysis by qRT-PCR indicated a difference in the *ST6Gal1* mRNA levels across the two transformants (mS15 and S16) that had only one copy of the gene integrated into the gDNA, suggesting that the transcription might be dependent on other factors such as the location of integration.

6.1.3. Recombinant ST6Gal1 was degraded during cultivation (Chapter 5)

The total protein concentration in the culture supernatant of the transformants revealed a pulsating pattern at an interval of 24 h over a period of 120 h of cultivation. This pointed to continuous secretion and degradation of proteins in the culture medium. Subsequent analysis showed a gradual increase in the total extracellular protease activity in the culture medium. More notably, the extracellular protease activity was higher in the transformants compared to the non-transformant RUT-C30, probably to hydrolyse the recombinant ST6Gal1 and mCherry proteins as a line of defence.

Confocal microscopy with selective staining of the endoplasmic reticulum (ER) revealed the co-localisation of mCherry in the ER suggesting processing of the fusion protein. Subsequent analysis of the transcriptomic levels of selected genes associated with the unfolded protein response (UPR) and the ERAD pathway revealed up-regulation at 24 h and 48 h for the mCherry containing transformant (mS15) and at 24 h for the non-mCherry transformant (S16). This indicated cellular stress in folding and processing of proteins, most likely due to the expression of human ST6Gal1.

Western blotting of the intracellular protein fraction using an anti-ST6Gal1 antibody detected the rST6Gal1 in the mCherry containing transformant (mS15) at a molecular size of 40 kDa. This suggested that cleavage of the fusion protein occurred inside the hyphae. However, there was no indication of the fusion protein inside the cells based on the results from Western blotting. Further, the rST6Gal1 was not detected in the intracellular protein fraction of the non-mCherry transformant (S16). Subsequent detection of the mCherry protein at several molecular weights further reinforced the possibility of intracellular cleavage of the fusion protein. Also, the abundance of mCherry inside the hyphae was much higher compared to the rST6Gal1. This was against expectation since the theoretical ratio of expression should ideally be 1:1 owing to the fact that both the genes encoding mCherry and ST6Gal1 were

expressed under the same *cbh1* promoter. This finding further suggested the likelihood of intracellular degradation of rST6Gal1.

Purification and hence quantitation of the rST6Gal1 could not be accomplished due to the lack of the N-terminus *Strep*-tag ™ II that would have aided protein purification. A lectin based ST6Gal1 activity assay was devised and established to be working with commercially available active recombinant human ST6Gal1. The results were subsequently validated using mass spectrometry by targeted analysis of the glycans on the acceptor substrate confirming the sialic acid transfer. The activity of the rST6Gal1 was assessed using a concentrated culture supernatant of the transformants; however, no conclusive evidence could be obtained for the functionality of rST6Gal1. This may be because of lack of sufficient incubation time for the transfer reaction or due to the cleavage of the rST6Gal1 at the N-terminus and/or C-terminal catalytic domain by fungal proteases that might have rendered the protein inactive. It could also be true that the galactose residues present on the asialofetuin might have been eliminated by the fungal exo-galactosidases, thus removing the residue required by the sialic acids to attach during the transfer reaction.

6.2. Future directions

6.2.1. Optimising the culture conditions to improve the rST6Gal1 yield and reduce protease activity

The rST6Gal1 in the culture supernatant was detected only at 48 h of cultivation and was found degraded at other time points. The pH of the culture medium at this time point was approximately 5, suggesting that proteases that had led to the degradation of the rST6Gal1 may be less active around this pH. This can be investigated further by using specific protease activity assay. Also, it will be a worthwhile attempt to keep the pH of the cultivation medium of the transformants constant at pH 5 to check if there is a decrease in the protease activity

and degradation of the rST6Gal1. Also, performing large-scale cultivation of the transformants in a fermenter will maintain a consistent level of pH over a large time period of cultivation (for example 120 h of cultivation), which also may increase the yield of rST6Gal1.

6.2.2. Purification of the rST6Gal1 protein

In the absence of *Strep*-tag™ II, an alternative purification strategy for the rST6Gal1 needs to be implemented. Purification of rST6Gal1 attempted during the course of this project established by previous studies (Miyagi and Tsuiki, 1982; Sticher *et al*., 1991) based on the affinity of ST6Gal1 protein to Cibacron Blue F3GA dye, was not successful. Therefore, a more sophisticated attempt needs to be devised, for example, by using high performance liquid chromatography (HPLC) to separate proteins based on their molecular weights and subsequently performing immunopurification using an anti-ST6Gal1 antibody, provided that the cost of the antibody can be covered. The functionality of the purified rST6Gal1 can be subsequently assayed using the lectin based activity assay (section 2.13). Moreover, N-terminal sequencing can be employed subsequently to identify the exact location of cleavage of the rST6Gal from the fusion protein.

6.2.3. Sequencing of the genomic DNA

The genomic DNA of the fungal transformants producing the rST6Gal1 can be sequenced to draw a clear conclusion on the location of the expression cassettes in the fungal genome. Sequencing the expression cassette integration sites within the fungal genome may bring about more clarity on why there was a difference in the expression of the heterologous *ST6Gal1* gene across different transformants since the location of integration may affect the heterologous gene expression. Efficient transcription of a heterologous gene requires integration of the DNA into transcriptionally active locations within the gDNA of filamentous fungi (Jeenes *et al.*, 1991). Moreover, integration of the transforming DNA into the gDNA

may cause interruption in a gene of which the function is required for, *e.g*., growth and efficient secretion of proteins. Furthermore, sequencing will also show if the integration site of the heterologous DNA is different across the transformants.

6.2.4. Analysis of the transformants containing *ST3Gal3* and *ST8Sia3* genes

The transformants containing *ST3Gal3* and *ST8Sia3* genes that were not analysed in this work could be examined for sialyltransferase expression using the same cultivation conditions and approaches employed for the ST6Gal1 transformants. This will further show if the conclusions made for the ST6Gal1 transformants are true for the ST3Gal3 and ST8Sia3 transformants.

6.2.5. Redesign the expression cassette for transformation

An additional strategy that could be employed is to redesign the expression cassette. The regions in the expression cassettes that were thought to be susceptible to cleavages such as the linker sequence that was introduced between the *ST6Gal1* gene and the *Strep*-tag™ II could be replaced or removed to see if the tag is intact with the rST6Gal1. The N-terminal sequencing will also provide an insight into the potential cleavage site (if any) within the rST6Gal1 that led to its separation from the fusion protein. The gene coding for the mCherry protein may be included as it served the intended purpose of screening of transformants based on the production of mCherry. However, mCherry fluorescence may not be utilised for tracking the secretion since the fusion protein was found cleaved intracellularly.

6.3. Conclusions

While the human ST6Gal1 was successfully expressed in *Trichoderma reesei*, challenges remain for optimising the cultivation conditions of the transformants to improve the yield and minimise degradation, and to purify and examine the functionality of the recombinant ST6Gal1. The project provides a proof of concept for the recombinant expression of a human

sialyltransferase in *T. reesei* and contains valuable information for future work on the production of recombinant human glycosyltransferases and other proteins of human origin in this industrial workhorse.

REFERENCES

Adav, S. S., Ravindran, A., Chao, L. T., Tan, L., Singh, S. & Sze, S. K. 2011. Proteomic analysis of pH and strains dependent protein secretion of *Trichoderma reesei*. *Journal of Proteome Research,* 10**,** 4579-96.

Albano, C. R., Randers-Eichhorn, L., Bentley, W. E. & Rao, G. 1998. Green fluorescent protein as a real time quantitative reporter of heterologous protein production. *Biotechnology Progress,* 14**,** 351-54.

Almaraz, R. T., Tian, Y., Bhattarcharya, R., Tan, E., Chen, S. H., Dallas, M. R., Chen, L., Zhang, Z., Zhang, H., Konstantopoulos, K. & Yarema, K. J. 2012. Metabolic flux increases glycoprotein sialylation: implications for cell adhesion and cancer metastasis. *Molecular and Cellular Proteomics,* 11**,** M112 017558.

Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S. & Evans, R. M. 1982. Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. *Nature,* 298**,** 240-4.

Angata, K., Huckaby, V., Ranscht, B., Terskikh, A., Marth, J. D. & Fukuda, M. 2007. Polysialic acid-directed migration and differentiation of neural precursors are essential for mouse brain development. *Molecular and Cellular Biology,* 27**,** 6659-68.

Angata, K., Suzuki, M., Mcauliffe, J., Ding, Y., Hindsgaul, O. & Fukuda, M. 2000. Differential biosynthesis of polysialic acid on neural cell adhesion molecule (NCAM) and oligosaccharide acceptors by three distinct alpha 2,8-sialyltransferases, ST8Sia IV (PST), ST8Sia II (STX), and ST8Sia III. *Journal of Biological Chemistry,* 275**,** 18594-601.

Angov, E. 2011. Codon usage: nature's roadmap to expression and folding of proteins. *Biotechnology Journal,* 6**,** 650-59.

Angov, E., Legler, P. M. & Mease, R. M. 2011. Adjustment of codon usage frequencies by codon harmonization improves protein expression and folding. *Methods in Molecular Biology,* 705**,** 01-13.

Aro, N., Ilmen, M., Saloheimo, A. & Penttilä, M. 2003. ACEI of *Trichoderma reesei* is a repressor of cellulase and xylanase expression. *Applied and Environmental Microbiology,* 69**,** 56-65.

Aro, N., Saloheimo, A., Ilmen, M. & Penttilä, M. 2001. ACEII, a novel transcriptional activator involved in regulation of cellulase and xylanase genes of *Trichoderma reesei*. *Journal of Biological Chemistry,* 276**,** 24309-14.

Arvas, M., Pakula, T., Lanthaler, K., Saloheimo, M., Valkonen, M., Suortti, T., Robson, G. & Penttilä, M. 2006. Common features and interesting differences in transcriptional responses to secretion stress in the fungi *Trichoderma reesei* and *Saccharomyces cerevisiae*. *BMC Genomics,* 7**,** 32.

Assenberg, R., Wan, P. T., Geisse, S. & Mayr, L. M. 2013. Advances in recombinant protein expression for use in pharmaceutical research. *Current Opinion in Structural Biology,* 23**,** 393-402.

Atanasova, L., Knox, B. P., Kubicek, C. P., Druzhinina, I. S. & Baker, S. E. 2013. The polyketide synthase gene pks4 of *Trichoderma reesei* provides pigmentation and stress resistance. *Eukaryotic Cell,* 12**,** 1499-508.

Audry, M., Jeanneau, C., Imberty, A., Harduin-Lepers, A., Delannoy, P. & Breton, C. 2011. Current trends in the structure-activity relationships of sialyltransferases. *Glycobiology,* 21**,** 716-26.

Avril, T., Wagner, E. R., Willison, H. J. & Crocker, P. R. 2006. Sialic acid-binding immunoglobulin-like lectin 7 mediates selective recognition of sialylated glycans expressed on Campylobacter jejuni lipooligosaccharides. *Infection and Immunity,* 74**,** 4133-41.

Bailey, M. J. & Tähtiharju, J. 2003. Efficient cellulase production by *Trichoderma reesei* in continuous cultivation on lactose medium with a computer-controlled feeding strategy. *Applied Microbiology and Biotechnology,* 62**,** 156-62.

Bajazawa, K., Furukawa, K., Narimatsu, H. & Kobata, A. 1993. Kinetic study of human β- 1,4-galactosyltrasgerase expressed in *E.coli*. *Journal of Biochemistry,* 113**,** 747-753.

Bakker, H., Bardor, M., Molthoff, J. W., Gomord, V., Elbers, I., Stevens, L. H., Jordi, W., Lommen, A., Faye, L., Lerouge, P. & Bosch, D. 2001. Galactose-extended glycans of antibodies produced by transgenic plants. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 2899-904.

Benne, R., Van Den Burg, J., Brakenhoff, J. P., Sloof, P., Van Boom, J. H. & Tromp, M. C. 1986. Major transcript of the frameshifted *coxII* gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell,* 46**,** 819-26.

Bergquist, P., Te'o, V., Gibbs, M., Cziferszky, A., De Faria, F. P., Azevedo, M. & Nevalainen, H. 2002. Expression of xylanase enzymes from thermophilic microorganisms in fungal hosts. *Extremophiles,* 6**,** 177-84.

Berka, R. M. & Barnett, C. C. 1989. The development of gene expression systems for filamentous fungi. *Biotechnology Advances,* 7**,** 127-54.

Bertani, G. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *Journal of Bacteriology,* 62**,** 293-300.

Bisaria, V. S. & Mishra, S. 1989. Regulatory aspects of cellulase biosynthesis and secretion. *Critical Reviews in Biotechnology,* 9**,** 61-103.

Bjellqvist, B., Basse, B., Olsen, E. & Celis, J. E. 1994. Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis,* 15**,** 529-39.

Bos, P. D., Zhang, X. H., Nadal, C., Shu, W., Gomis, R. R., Nguyen, D. X., Minn, A. J., Van De Vijver, M. J., Gerald, W. L., Foekens, J. A. & Massague, J. 2009. Genes that mediate breast cancer metastasis to the brain. *Nature,* 459**,** 1005-09.

Braaksma, M. & Punt, P. J. 2008. *Aspergillus* as a cell factory for protein production: controlling protease activity in fungal production. *The Aspergilli: Genomics, Medical Aspects, Biotechnology, and Research Methods***,** 441-55.

Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry,* 72**,** 248-54.

Breen, K. C. 2002. The role of protein glycosylation in the control of cellular *N*-sialyltransferase activity. *FEBS Letters,* 517**,** 215-18.

Brockhausen, I. 1999. Pathways of *O*-glycan biosynthesis in cancer cells. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1473**,** 67-95.

Brockhausen, I., Schachter, H. & Stanley, P. 2009. *O*-GalNAc Glycans. *In:* Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. & Etzler, M. E. (eds.) *Essentials of Glycobiology.* 2nd ed. Cold Spring Harbor (NY).

Bull, C., Stoel, M. A., Den Brok, M. H. & Adema, G. J. 2014. Sialic acids sweeten a tumor's life. *Cancer Research,* 74**,** 3199-204.

Cabrera, G., Cremata, J. A., Valdes, R., Garcia, R., Gonzalez, Y., Montesino, R., Gomez, H. & Gonzalez, M. 2005. Influence of culture conditions on the *N*-glycosylation of a monoclonal antibody specific for recombinant hepatitis B surface antigen. *Biotechnology and Applied Biochemistry,* 41**,** 67-76.

Carlin, A. F., Lewis, A. L., Varki, A. & Nizet, V. 2007. Group B streptococcal capsular sialic acids interact with siglecs (immunoglobulin-like lectins) on human leukocytes. *Journal of Bacteriology,* 189**,** 1231-37.

Carvalho, N. D., Arentshorst, M., Kooistra, R., Stam, H., Sagt, C. M., Van Den Hondel, C. A. & Ram, A. F. 2011. Effects of a defective ERAD pathway on growth and heterologous

protein production in *Aspergillus niger*. *Applied Microbiology and Biotechnology,* 89**,** 357-73.

Castellanos, F., Schmoll, M., Martinez, P., Tisch, D., Kubicek, C. P., Herrera-Estrella, A. & Esquivel-Naranjo, E. U. 2010. Crucial factors of the light perception machinery and their impact on growth and cellulase gene transcription in *Trichoderma reesei*. *Fungal Genetics and Biology,* 47**,** 468-76.

Castro-Chavez, F. 2011. Most used codons per amino acid and per genome in the code of man compared to other organisms according to the rotating circular genetic code. *Neuroquantology,* 9.

Cazet, A., Groux-Degroote, S., Teylaert, B., Kwon, K. M., Lehoux, S., Slomianny, C., Kim, C. H., Le Bourhis, X. & Delannoy, P. 2009. GD3 synthase overexpression enhances proliferation and migration of MDA-MB-231 breast cancer cells. *Biological Chemistry,* 390**,** 601-09.

Chakravarthi, S., Jessop, C. E. & Bulleid, N. J. 2006. The role of glutathione in disulphide bond formation and endoplasmic-reticulum-generated oxidative stress. *EMBO Rep,* 7**,** 271-5.

Chapman, R., Sidrauski, C. & Walter, P. 1998. Intracellular signaling from the endoplasmic reticulum to the nucleus. *Annual Review of Cell and Developmental Biology,* 14**,** 459-85.

Cherry, J. R. & Fidantsef, A. L. 2003. Directed evolution of industrial enzymes: an update. *Current Opinion in Biotechnology,* 14**,** 438-43.

Chevet, E., Cameron, P. H., Pelletier, M. F., Thomas, D. Y. & Bergeron, J. J. 2001. The endoplasmic reticulum: integration of protein folding, quality control, signaling and degradation. *Current Opinion in Structural Biology,* 11**,** 120-24.

Chiba, Y., Ito, H., Sato, T., Takahashi, Y., Jigami, Y. & Narimatsu, H. 2009. Expression system for human glycosyltransferases and its application. *Journal of Applied Glycoscience,* 57**,** 131-36.

Choi, B. K., Bobrowicz, P., Davidson, R. C., Hamilton, S. R., Kung, D. H., Li, H., Miele, R. G., Nett, J. H., Wildt, S. & Gerngross, T. U. 2003. Use of combinatorial genetic libraries to humanize *N*-linked glycosylation in the yeast *Pichia pastoris*. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 5022-27.

Chrostek, L., Cylwik, B., Gindzienska-Sieskiewicz, E., Gruszewska, E., Szmitkowski, M. & Sierakowski, S. 2014. Sialic acid level reflects the disturbances of glycosylation and acute-phase reaction in rheumatic diseases. *Rheumatology International,* 34**,** 393-99.

Cobb, B. A. & Kasper, D. L. 2005. Coming of age: carbohydrates and immunity. *European Journal of Immunology,* 35**,** 352-56.

Cointe, D., Leroy, Y. & Chirat, F. 1998. Determination of the sialylation level and of the ratio alpha-(2-->3)/alpha-(2-->6) sialyl linkages of *N*-glycans by methylation and GC/MS analysis. *Carbohydrate Research,* 311**,** 51-9.

Conesa, A., Punt, P. J., Van Luijk, N. & Van Den Hondel, C. A. 2001. The secretion pathway in filamentous fungi: a biotechnological view. *Fungal Genetics and Biology,* 33**,** 155-71.

Crespo, H. J., Lau, J. T. & Videira, P. A. 2013. Dendritic cells: a spot on sialic Acid. *Frontiers in Immunology,* 4**,** 491.

Crocker, P. R., Paulson, J. C. & Varki, A. 2007. Siglecs and their roles in the immune system. *Nature Reviews: Immunology,* 7**,** 255-66.

Cummings, R. & Turco, S. 2009. Parasitic Infections. *In:* Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. & Etzler, M. E. (eds.) *Essentials of Glycobiology.* 2nd ed. Cold Spring Harbor (NY).

Curach, N. C. 2005. An investigation into the *hex1* gene and gene promoter for the enhancement of protein production in *Trichoderma reesei*. *Ph.D Thesis*.

Datta, A. K. 2009. Comparative sequence analysis in the sialyltransferase protein family: analysis of motifs. *Current Drug Targets,* 10**,** 483-98.

Datta, A. K., Chammas, R. & Paulson, J. C. 2001. Conserved cysteines in the sialyltransferase sialylmotifs form an essential disulfide bond. *Journal of Biological Chemistry,* 276**,** 15200-07.

Davidson, R. C., Cruz, M. C., Sia, R. A., Allen, B., Alspaugh, J. A. & Heitman, J. 2000. Gene disruption by biolistic transformation in serotype D strains of *Cryptococcus neoformans*. *Fungal Genetics and Biology,* 29**,** 38-48.

De Oliveira, J. T., De Matos, A. J., Santos, A. L., Pinto, R., Gomes, J., Hespanhol, V., Chammas, R., Manninen, A., Bernardes, E. S., Albuquerque Reis, C., Rutteman, G. & Gartner, F. 2011. Sialylation regulates galectin-3/ligand interplay during mammary tumour progression--a case of targeted uncloaking. *International Journal of Developmental Biology,* 55**,** 823-34.

Deng, T. 1997. Bacterial expression and purification of biologically active mouse c-Fos proteins by selective codon optimization. *FEBS Letters,* 409**,** 269-72.

Dennis, J. W., Granovsky, M. & Warren, C. E. 1999. Protein glycosylation in development and disease. *Bioessays,* 21**,** 412-21.

Digman, M. A., Wiseman, P. W., Horwitz, A. R. & Gratton, E. 2009. Detecting protein complexes in living cells from laser scanning confocal image sequences by the cross correlation raster image spectroscopy method. *Biophysical Journal,* 96**,** 707-16.

Donadio, S., Dubois, C., Fichant, G., Roybon, L., Guillemot, J.-C., Breton, C. & Ronin, C. 2003. Recognition of cell surface acceptors by two human α-2, 6-sialyltransferases produced in CHO cells. *Biochimie,* 85**,** 311-321.

Druzhinina, I. S., Shelest, E. & Kubicek, C. P. 2012. Novel traits of *Trichoderma* predicted through the analysis of its secretome. *FEMS Microbiology Letters,* 337**,** 1-9.

Dunaevsky, Y. E., Gruban, T. N., Beliakova, G. A. & Belozersky, M. A. 2000. Enzymes secreted by filamentous fungi: regulation of secretion and purification of an extracellular protease of *Trichoderma harzianum*. *Biochemistry (Moscow)*, 65**,** 723-27.

Durand, H., Clanet, M. & Tiraby, G. 1988. Genetic improvement of *Trichoderma reesei* for large-scale cellulase production. *Enzyme and Microbial Technology*, 10**,** 341-46.

Dwek, R. A. 1995. Glycobiology: "towards understanding the function of sugars". *Biochemical Society Transactions*, 23**,** 1-25.

Egrie, J. C., Dwyer, E., Browne, J. K., Hitz, A. & Lykos, M. A. 2003. Darbepoetin alfa has a longer circulating half-life and greater in vivo potency than recombinant human erythropoietin. *Experimental Hematology*, 31**,** 290-99.

Eiden-Plach, A., Zagorc, T., Heintel, T., Carius, Y., Breinig, F. & Schmitt, M. J. 2004. Viral preprotoxin signal sequence allows efficient secretion of green fluorescent protein by Candida glabrata, *Pichia pastoris*, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe*. *Applied and Environmental Microbiology*, 70**,** 961-66.

Eiyama, A., Kondo-Okamoto, N. & Okamoto, K. 2013. Mitochondrial degradation during starvation is selective and temporally distinct from bulk autophagy in yeast. *FEBS Letters*, 587**,** 1787-92.

El-Gogary, S., Leite, A., Crivellaro, O., Eveleigh, D. E. & El-Dorry, H. 1989. Mechanism by which cellulose triggers cellobiohydrolase I gene expression in *Trichoderma reesei*. *Proceedings of the National Academy of Sciences of the United States of America*, 86**,** 6138-41.

Elliott, S., Lorenzini, T., Asher, S., Aoki, K., Brankow, D., Buck, L., Busse, L., Chang, D., Fuller, J., Grant, J., Hernday, N., Hokum, M., Hu, S., Knudten, A., Levin, N., Komorowski, R., Martin, F., Navarro, R., Osslund, T., Rogers, G., Rogers, N., Trail, G. & Egrie, J. 2003. Enhancement of therapeutic protein in vivo activities through glycoengineering. *Nature Biotechnology*, 21**,** 414-21.

Eneyskaya, E. V., Kulminskaya, A. A., Savel'ev, A. N., Savel'eva, N. V., Shabalin, K. A. & Neustroev, K. N. 1990. Acid protease from *Trichoderma reesei*: limited proteolysis of fungal carbohydrases. *Applied Microbiology and Biotechnology,* 52**,** 226-31.

Esko, J. D., Kimata, K. & Lindahl, U. 2009. Proteoglycans and sulfated glycosaminoglycans. *In:* Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. & Etzler, M. E. (eds.) *Essentials of Glycobiology.* 2nd ed. Cold Spring Harbor (NY).

Fast, D. G., Jamieson, J. C. & Mccaffrey, G. 1993. The role of the carbohydrate chains of Gal beta-1,4-GlcNAc alpha 2,6-sialyltransferase for enzyme activity. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1202**,** 325-30.

Feng, L., Chan, W. W., Roderick, S. L. & Cohen, D. E. 2000. High-level expression and mutagenesis of recombinant human phosphatidylcholine transfer protein using a synthetic gene: evidence for a C-terminal membrane binding domain. *Biochemistry,* 39**,** 15399-409.

Ferrer-Miralles, N., Saccardo, P., Corchero, J. L., Xu, Z. & Garcia-Fruitos, E. 2015. General introduction: recombinant protein production and purification of insoluble proteins. *Methods in Molecular Biology,* 1258**,** 1-24.

Fewell, S. W. & Brodsky, J. L. 2009. Entry into the endoplasmic reticulum: protein translocation, folding and quality control. *Trafficking Inside Cells* [Online].

Fiat, A. M., Jolles, J., Aubert, J. P., Loucheux-Lefebvre, M. H. & Jolles, P. 1980. Localisation and importance of the sugar part of human casein. *European Journal of Biochemistry,* 111**,** 333-39.

Finkelstein, D. B. 1987. Improvement of enzyme production in *Aspergillus*. *Antonie Van Leeuwenhoek,* 53**,** 349-52.

Frame, T., Carroll, T., Korchagina, E., Bovin, N. & Henry, S. 2007. Synthetic glycolipid modification of red blood cell membranes. *Transfusion,* 47**,** 876-82.

Frommer, W. B. & Ninnemann, O. 1995. Heterologous expression of genes in bacterial, fungal, animal, and plant cells. *Annual Review of Plant Physiology and Plant Molecular Biology*, 46**,** 419-44.

Fujiyama, K., Ido, Y., Misaki, R., Moran, D. G., Yanagihara, I., Honda, T., Nishimura, S., Yoshida, T. & Seki, T. 2001. Human *N*-acetylglucosaminyltransferase I. Expression in *Escherichia coli* as a soluble enzyme, and application as an immobilized enzyme for the chemoenzymatic synthesis of *N*-linked oligosaccharides. *Journal of Bioscience and Bioengineering*, 92**,** 569-74.

Gabius, H. J. 2011. *The Sugar Code: Fundamentals of Glycosciences*, Wiley.

Gagneux, P. & Varki, A. 1999. Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology*, 9**,** 747-55.

Gasser, B., Saloheimo, M., Rinas, U., Dragosits, M., Rodriguez-Carmona, E., Baumann, K., Giuliani, M., Parrilli, E., Branduardi, P., Lang, C., Porro, D., Ferrer, P., Tutino, M. L., Mattanovich, D. & Villaverde, A. 2008. Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microbial Cell Factories*, 7**,** 11.

Gemmill, T. R. & Trimble, R. B. 1999. Overview of *N*- and *O*-linked oligosaccharide structures found in various yeast species. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1426**,** 227-37.

Gesteland, R. F. & Atkins, J. F. 1996. Recoding: dynamic reprogramming of translation. *Annual Review of Biochemistry*, 65**,** 741-68.

Gething, M. J. & Sambrook, J. 1992. Protein folding in the cell. *Nature*, 355**,** 33-45.

Ghosh, A., Al-Rabiai, S., Ghosh, B., Trimiño-Vazquez, H., Eveleigh, D. & Montenecourt, B. 1982. Increased endoplasmic reticulum content of a mutant of *Trichoderma reesei* (RUT-C30) in relation to cellulase synthesis. *Enzyme and Microbial Technology*, 4**,** 110-13.

Gierasch, L. M. 1989. Signal sequences. *Biochemistry,* 28**,** 923-30.

Godlewski, M., Kautto, L. & Nevalainen, H. 2009. The elusive pulse in protein production in *Trichoderma reesei*, *Light in Life Sciences (LILS) Conference Poster*. Melbourne, Australia.

Golubev, A. M. & Neustroev, K. N. 1993. Crystallization of α-Galactosidase from *Trichoderma reesei*. *Journal of Molecular Biology,* 231**,** 933-934.

Gordon, C. L., Archer, D. B., Jeenes, D. J., Doonan, J. H., Wells, B., Trinci, A. P. & Robson, G. D. 2000a. A glucoamylase::GFP gene fusion to study protein secretion by individual hyphae of *Aspergillus niger*. *Journal of Microbiological Methods,* 42**,** 39-48.

Gordon, C. L., Khalaj, V., Ram, A. F., Archer, D. B., Brookman, J. L., Trinci, A. P., Jeenes, D. J., Doonan, J. H., Wells, B., Punt, P. J., Van Den Hondel, C. A. & Robson, G. D. 2000b. Glucoamylase::green fluorescent protein fusions to monitor protein secretion in *Aspergillus niger*. *Microbiology,* 146 (Pt 2)**,** 415-26.

Gouka, R. J., Hessing, J. G., Punt, P. J., Stam, H., Musters, W. & Van Den Hondel, C. A. 1996a. An expression system based on the promoter region of the *Aspergillus awamori* 1,4-beta-endoxylanase A gene. *Applied Microbiology and Biotechnology,* 46**,** 28-35.

Gouka, R. J., Punt, P. J., Hessing, J. G. & Van Den Hondel, C. A. 1996b. Analysis of heterologous protein production in defined recombinant *Aspergillus awamori* strains. *Applied and Environmental Microbiology,* 62**,** 1951-57.

Gouka, R. J., Punt, P. J. & Van Den Hondel, C. A. 1997a. Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. *Applied Microbiology and Biotechnology,* 47**,** 1-11.

Gouka, R. J., Punt, P. J. & Van Den Hondel, C. A. 1997b. Glucoamylase gene fusions alleviate limitations for protein production in *Aspergillus awamori* at the transcriptional and (post) translational levels. *Applied and Environmental Microbiology,* 63**,** 488-97.

Gramer, M. J., Eckblad, J. J., Donahue, R., Brown, J., Shultz, C., Vickerman, K., Priem, P., Van Den Bremer, E. T., Gerritsen, J. & Van Berkel, P. H. 2011. Modulation of antibody galactosylation through feeding of uridine, manganese chloride, and galactose. *Biotechnology and Bioengineering,* 108**,** 1591-602.

Greene, J. J. 2004. Host cell compatibility in protein expression. *Methods in Molecular Biology,* 267**,** 3-14.

Griffin, M. E. & Hsieh-Wilson, L. C. 2016. Glycan engineering for cell and developmental biology. *Cell Chemical Biology*.

Grundmann, U., Nerlich, C., Rein, T. & Zettlmeissl, G. 1990. Complete cDNA sequence encoding human beta-galactoside alpha-2,6-sialyltransferase. *Nucleic Acids Research,* 18**,** 667.

Guadamillas, M. C., Cerezo, A. & Del Pozo, M. A. 2011. Overcoming anoikis--pathways to anchorage-independent growth in cancer. *Journal of Cell Science,* 124**,** 3189-97.

Gupta, A., Roy, S., Lazar, A. J., Wang, W. L., Mcauliffe, J. C., Reynoso, D., Mcmahon, J., Taguchi, T., Floris, G., Debiec-Rychter, M., Schoffski, P., Trent, J. A., Debnath, J. & Rubin, B. P. 2010. Autophagy inhibition and antimalarials promote cell death in gastrointestinal stromal tumor (GIST). *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 14333-8.

Gustafsson, C., Govindarajan, S. & Minshull, J. 2004. Codon bias and heterologous protein expression. *Trends in Biotechnology,* 22**,** 346-53.

Gustafsson, C., Minshull, J., Govindarajan, S., Ness, J., Villalobos, A. & Welch, M. 2012. Engineering genes for predictable protein expression. *Protein Expression and Purification,* 83**,** 37-46.

Haab, D., Hagspiel, K., Szakmary, K. & Kubicek, C. P. 1990. Formation of the extracellular proteases from *Trichoderma reesei* QM 9414 involved in cellulase degradation. *Journal of Biotechnology,* 16**,** 187-98.

Hakomori, S. 2002. Glycosylation defining cancer malignancy: new wine in an old bottle. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 10231-33.

Halic, M., Becker, T., Pool, M. R., Spahn, C. M., Grassucci, R. A., Frank, J. & Beckmann, R. 2004. Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature,* 427**,** 808-14.

Hamilton, S. R., Bobrowicz, P., Bobrowicz, B., Davidson, R. C., Li, H., Mitchell, T., Nett, J. H., Rausch, S., Stadheim, T. A., Wischnewski, H., Wildt, S. & Gerngross, T. U. 2003. Production of complex human glycoproteins in yeast. *Science,* 301**,** 1244-46.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology,* 166**,** 557-80.

Haraguchi, M., Yamashiro, S., Yamamoto, A., Furukawa, K., Takamiya, K., Lloyd, K. O., Shiku, H. & Furukawa, K. 1994. Isolation of GD3 synthase gene by expression cloning of GM3 alpha-2,8-sialyltransferase cDNA using anti-GD2 monoclonal antibody. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 10455-59.

Harduin-Lepers, A., Mollicone, R., Delannoy, P. & Oriol, R. 2005. The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology,* 15**,** 805-17.

Harduin-Lepers, A., Recchi, M. A. & Delannoy, P. 1995. 1994, the year of sialyltransferases. *Glycobiology,* 5**,** 741-58.

Harduin-Lepers, A., Stokes, D. C., Steelant, W. F., Samyn-Petit, B., Krzewinski-Recchi, M. A., Vallejo-Ruiz, V., Zanetta, J. P., Auge, C. & Delannoy, P. 2000. Cloning, expression and

gene organization of a human Neu5Ac alpha 2-3Gal beta 1-3GalNAc alpha 2,6-sialyltransferase: hST6GalNAcIV. *Biochemical Journal,* 352 Pt 1**,** 37-48.

Harduin-Lepers, A., Vallejo-Ruiz, V., Krzewinski-Recchi, M. A., Samyn-Petit, B., Julien, S. & Delannoy, P. 2001. The human sialyltransferase family. *Biochimie,* 83**,** 727-37.

Harkki, A., Mäntylä, A., Penttilä, M., Muttilainen, S., Bühler, R., Suominen, P., Knowles, J. & Nevalainen, H. 1991. Genetic engineering of *Trichoderma* to produce strains with novel cellulase profiles. *Enzyme and Microbial Technology,* 13**,** 227-33.

Harkki, A., Uusitalo, J., Bailey, M., Penttilä, M. & Knowles, J. K. C. 1989. A novel fungal expression system: secretion of active calf chymosin from the filamentous fungus *Trichoderma reesei*. *Nature Biotechnology,* 7**,** 596-603.

Hauselmann, I. & Borsig, L. 2014. Altered tumor-cell glycosylation promotes metastasis. *Frontiers in Oncology,* 4**,** 28.

Haynes, P. A. 1998. Phosphoglycosylation: a new structural class of glycosylation? *Glycobiology,* 8**,** 1-5.

Hazell, B. W., Te'o, V. S., Bradner, J. R., Bergquist, P. L. & Nevalainen, K. M. 2000. Rapid transformation of high cellulase-producing mutant strains of *Trichoderma reesei* by microprojectile bombardment. *Letters in Applied Microbiology,* 30**,** 282-86.

He, F., Hogan, S., Latypov, R. F., Narhi, L. O. & Razinkov, V. I. 2010. High throughput thermostability screening of monoclonal antibody formulations. *Journal of Pharmaceutical Sciences*, 99**,** 1707-20.

Heimel, K. 2015. Unfolded protein response in filamentous fungi-implications *in Biotechnology*. *Applied Microbiology and Biotechnology,* 99**,** 121-32.

Hennet, T., Chui, D., Paulson, J. C. & Marth, J. D. 1998. Immune regulation by the ST6Gal sialyltransferase. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 4504-09.

Hernandez-Garcia, C. M. & Finer, J. J. 2014. Identification and validation of promoters and cis-acting regulatory elements. *Plant Science,* 217-218**,** 109-19.

Herzog, R., Daniell, H., Singh, N. & Lemke, P. 1996. A comparative study on the transformation of *Aspergillus nidulans* by microprojectile bombardment of conidia and a more conventional procedure using protoplasts treated with polyethyleneglycol. *Applied Microbiology and Biotechnology,* 45**,** 333-337.

Hidari, K. I., Horie, N., Murata, T., Miyamoto, D., Suzuki, T., Usui, T. & Suzuki, Y. 2005. Purification and characterization of a soluble recombinant human ST6Gal I functionally expressed in *Escherichia coli*. *Glycoconjugate Journal,* 22**,** 1-11.

Hillier, C. J., Ware, L. A., Barbosa, A., Angov, E., Lyon, J. A., Heppner, D. G. & Lanar, D. E. 2005. Process development and analysis of liver-stage antigen 1, a preerythrocyte-stage protein-based vaccine for *Plasmodium falciparum*. *Infection and Immunity,* 73**,** 2109-15.

Hohenblum, H., Gasser, B., Maurer, M., Borth, N. & Mattanovich, D. 2004. Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnology and Bioengineering,* 85**,** 367-75.

Hopkins, D., Gomathinayagam, S., Rittenhour, A. M., Du, M., Hoyt, E., Karaveg, K., Mitchell, T., Nett, J. H., Sharkey, N. J., Stadheim, T. A., Li, H. & Hamilton, S. R. 2011. Elimination of beta-mannose glycan structures in *Pichia pastoris*. *Glycobiology,* 21**,** 1616-26.

Hou, J., Tyo, K. E., Liu, Z., Petranovic, D. & Nielsen, J. 2012. Metabolic engineering of recombinant protein secretion by *Saccharomyces cerevisiae*. *FEMS Yeast Research,* 12**,** 491-510.

Huang, C. J., Lin, H. & Yang, X. 2012. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *Journal of Industrial Microbiology & Biotechnology,* 39**,** 383-99.

Ichihara, A. 2010. Reminiscence of 40-year research on nitrogen metabolism. *Proceedings of the Japan Academy. Series B: Physical and Biological Sciences,* 86**,** 707-16.

Ihara, Y., Inai, Y., Ikezaki, M., Matsui, I.-S. L., Manabe, S. & Ito, Y. 2015. *C*-mannosylation: modification on tryptophan in cellular proteins. *Glycoscience: Biology and Medicine*. Springer.

Ikehara, Y., Kojima, N., Kurosawa, N., Kudo, T., Kono, M., Nishihara, S., Issiki, S., Morozumi, K., Itzkowitz, S., Tsuda, T., Nishimura, S. I., Tsuji, S. & Narimatsu, H. 1999. Cloning and expression of a human gene encoding an *N*-acetylgalactosamine-alpha2,6-sialyltransferase (ST6GalNAc I): a candidate for synthesis of cancer-associated sialyl-Tn antigens. *Glycobiology,* 9**,** 1213-24.

Ilmén, M., Onnela, M. L., Klemsdal, S., Keränen, S. & Penttilä, M. 1996. Functional analysis of the cellobiohydrolase I promoter of the filamentous fungus *Trichoderma reesei*. *Molecular and General Genetics,* 253**,** 303-14.

Inoue, H., Nojima, H. & Okayama, H. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene,* 96**,** 23-8.

Ishibashi, K., Suzuki, K., Ando, Y., Takakura, C. & Inoue, H. 2006. Nonhomologous chromosomal integration of foreign DNA is completely dependent on MUS-53 (human Lig4 homolog) in *Neurospora*. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 14871-6.

Ishii, A., Ohta, M., Watanabe, Y., Matsuda, K., Ishiyama, K., Sakoe, K., Nakamura, M., Inokuchi, J., Sanai, Y. & Saito, M. 1998. Expression cloning and functional characterization of human cDNA for ganglioside GM3 synthase. *Journal of Biological Chemistry,* 273**,** 31652-55.

Ito, T., Sadamoto, R., Naruchi, K., Togame, H., Takemoto, H., Kondo, H. & Nishimura, S. 2010. Highly oriented recombinant glycosyltransferases: site-specific immobilization of unstable membrane proteins by using *Staphylococcus aureus* sortase A. *Biochemistry,* 49**,** 2604-14.

Janega, P., Cerna, A., Kholova, I., Brabencova, E. & Babal, P. 2002. Sialic acid expression in autoimmune thyroiditis. *Acta Histochemica,* 104**,** 343-47.

Jayaraj, R. & Smooker, P. M. 2009. So you need a protein-A guide to the production of recombinant proteins. *Open Veterinary Science Journal,* 3**,** 28-34.

Jeenes, D. J., Mackenzie, D. A., Roberts, I. N. & Archer, D. B. 1991. Heterologous protein production by filamentous fungi. *Biotechnology & Genetic Engineering Reviews,* 9**,** 327-67.

Jefferis, R. 2009. Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action. *Trends in Pharmacological Sciences,* 30**,** 356-62.

Johnson, A. E. & Van Waes, M. A. 1999. The translocon: a dynamic gateway at the ER membrane. *Annual Review of Cell and Developmental Biology,* 15**,** 799-842.

Jones, C., Virji, M. & Crocker, P. R. 2003. Recognition of sialylated meningococcal lipopolysaccharide by siglecs expressed on myeloid cells leads to enhanced bacterial uptake. *Molecular Microbiology,* 49**,** 1213-25.

Joutsjoki, V. V., Torkkeli, T. K. & Nevalainen, K. M. 1993. Transformation of *Trichoderma reesei* with the *Hormoconis resinae* glucoamylase P (*gamP*) gene: production of a heterologous glucoamylase by *Trichoderma reesei*. *Current Genetics,* 24**,** 223-28.

Jovanovic, B., Mach, R. L. & Mach-Aigner, A. R. 2014. Erythritol production on wheat straw using *Trichoderma reesei*. *AMB Express,* 4**,** 34.

Kaewsapsak, P., Esonu, O. & Dube, D. H. 2013. Recruiting the host's immune system to target *Helicobacter pylori*'s surface glycans. *ChemBioChem,* 14**,** 721-26.

Karlsson, K. A. 1998. Meaning and therapeutic potential of microbial recognition of host glycoconjugates. *Molecular Microbiology,* 29**,** 1-11.

Kaur, J., Pethani, B. P., Kumar, S., Kim, M., Sunna, A., Kautto, L., Penesyan, A., Paulsen, I. T. & Nevalainen, H. 2015. Pseudomonas aeruginosa inhibits the growth of *Scedosporium aurantiacum,* an opportunistic fungal pathogen isolated from the lungs of cystic fibrosis patients. *Frontiers in Microbiology,* 6**,** 866.

Kautto, L. 2009. *Trichoderma reesei* proteasome and genome-wide effects of the expression of mutant cellobiohydrolase I. Ph.D, Macquarie University.

Kayser, V., Chennamsetty, N., Voynov, V., Forrer, K., Helk, B. & Trout, B. L. 2011. Glycosylation influences on the aggregation propensity of therapeutic monoclonal antibodies. *Biotechnology Journal,* 6**,** 38-44.

Keränen, S. & Penttilä, M. 1995. Production of recombinant proteins in the filamentous fungus *Trichoderma reesei*. *Current Opinion in Biotechnology,* 6**,** 534-37.

Kiiskinen, L. L., Kruus, K., Bailey, M., Ylosmaki, E., Siika-Aho, M. & Saloheimo, M. 2004. Expression of Melanocarpus albomyces laccase in *Trichoderma reesei* and characterization of the purified enzyme. *Microbiology,* 150**,** 3065-74.

Kim, Y. J., Kim, K. S., Do, S., Kim, C. H., Kim, S. K. & Lee, Y. C. 1997. Molecular cloning and expression of human alpha2,8-sialyltransferase (hST8Sia V). *Biochemical and Biophysical Research Communications,* 235**,** 327-30.

Kim, Y. J., Kim, K. S., Kim, S. H., Kim, C. H., Ko, J. H., Choe, I. S., Tsuji, S. & Lee, Y. C. 1996. Molecular cloning and expression of human Gal beta 1,3GalNAc alpha 2,3-sialytransferase (hST3Gal II). *Biochemical and Biophysical Research Communications,* 228**,** 324-27.

Kimura, S., Maruyama, J., Kikuma, T., Arioka, M. & Kitamoto, K. 2011. Autophagy delivers misfolded secretory proteins accumulated in endoplasmic reticulum to vacuoles in the filamentous fungus *Aspergillus oryzae*. *Biochemical and Biophysical Research Communications,* 406**,** 464-70.

King, C. R. & Piatigorsky, J. 1983. Alternative RNA splicing of the murine alpha A-crystallin gene: protein-coding information within an intron. *Cell,* 32**,** 707-12.

Kitagawa, H. & Paulson, J. C. 1993. Cloning and expression of human Gal beta 1,3(4)GlcNAc alpha 2,3-sialyltransferase. *Biochemical and Biophysical Research Communications,* 194**,** 375-82.

Kitagawa, H. & Paulson, J. C. 1994a. Cloning of a novel alpha 2,3-sialyltransferase that sialylates glycoprotein and glycolipid carbohydrate groups. *Journal of Biological Chemistry,* 269**,** 1394-401.

Kitagawa, H. & Paulson, J. C. 1994b. Differential expression of five sialyltransferase genes in human tissues. *Journal of Biological Chemistry,* 269**,** 17872-78.

Klein, T., Gradziel, T., Fromm, M. & Sanford, J. 1988. Factors influencing gene delivery into zea mays cells by high–velocity microprojectiles. *Nature Biotechnology,* 6**,** 559-563.

Kobayashi, T., Nishizaki, R. & Ikezawa, H. 1997. The presence of GPI-linked protein(s) in an archaeobacterium, *Sulfolobus acidocaldarius,* closely related to eukaryotes. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1334**,** 1-4.

Koch, H. G., Moser, M. & Muller, M. 2003. Signal recognition particle-dependent protein targeting, universal to all kingdoms of life. *Reviews of Physiology Biochemistry and Pharmacology,* 146**,** 55-94.

Koike, M., Yutoku, Y. & Koike, A. 2013. Ku80 attentuates cytotoxicity induced by green fluorescent protein transduction independently of non-homologous end joining. *FEBS Open Bio,* 3**,** 46-50.

Kollewe, C. & Vilcinskas, A. 2013. Production of recombinant proteins in insect cells. *American Journal of Biochemistry and Biotechnology,* 9**,** 255-71.

Krappmann, S., Sasse, C. & Braus, G. H. 2006. Gene targeting in *Aspergillus fumigatus* by homologous recombination is facilitated in a nonhomologous end- joining-deficient genetic background. *Eukaryotic Cell,* 5**,** 212-5.

Kredics, L., Manczinger, L., Antal, Z., Penzes, Z., Szekeres, A., Kevei, F. & Nagy, E. 2004. In vitro water activity and pH dependence of mycelial growth and extracellular enzyme activities of *Trichoderma* strains with biocontrol potential. *Journal of Applied Microbiology,* 96**,** 491-8.

Kuhn, B., Benz, J., Greif, M., Engel, A. M., Sobek, H. & Rudolph, M. G. 2013. The structure of human alpha-2,6-sialyltransferase reveals the binding mode of complex glycans. *Acta Crystallographica. Section D: Biological Crystallography,* 69**,** 1826-38.

Lamarre-Vincent, N. & Hsieh-Wilson, L. C. 2003. Dynamic glycosylation of the transcription factor CREB: a potential role in gene regulation. *Journal of the American Chemical Society,* 125**,** 6612-13.

Landowski, C. P., Huuskonen, A., Wahl, R., Westerholm-Parvinen, A., Kanerva, A., Hänninen, A. L., Salovuori, N., Penttilä, M., Natunen, J., Ostermeier, C., Helk, B., Saarinen, J. & Saloheimo, M. 2015. Enabling Low Cost Biopharmaceuticals: A systematic approach to delete proteases from a well-known protein production host *Trichoderma reesei*. *PloS One,* 10**,** e0134723.

Landowski, C. P., Westerholm-Parvinen, A., Huuskonen, A., Wahl, R., Sommer, B. P., Penttilä, M., Ostermeier, C., Helk, B., Saarinen, J. & Saloheimo, M. 2016. Next generation biotherapeutic production system: the filamentous fungus *Trichoderma reesei*. *13th European Conference on Fungal Genetics,* Poster.

Laroy, W., Maras, M., Fiers, W. & Contreras, R. 1997. A radioactive assay for sialyltransferase activity using 96-well multiscreen filtration plates. *Analytical Biochemistry*, 249**,** 108-111.

Lauber, J., Handrick, R., Leptihn, S., Durre, P. & Gaisser, S. 2015. Expression of the functional recombinant human glycosyltransferase GalNAcT2 in *Escherichia coli*. *Microbial Cell Factories*, 14**,** 3.

Laughlin, S. T., Baskin, J. M., Amacher, S. L. & Bertozzi, C. R. 2008. In vivo Imaging of membrane-associated glycans in developing zebrafish. *Science*, 320**,** 664-67.

Lee, K. H., Feig, C., Tchikov, V., Schickel, R., Hallas, C., Schutze, S., Peter, M. E. & Chan, A. C. 2006. The role of receptor internalization in CD95 signaling. *EMBO Journal*, 25**,** 1009-23.

Lee, L. Y., Thaysen-Andersen, M., Baker, M. S., Packer, N. H., Hancock, W. S. & Fanayan, S. 2014. Comprehensive *N*-glycome profiling of cultured human epithelial breast cells identifies unique secretome *N*-glycosylation signatures enabling tumorigenic subtype classification. *Journal of Proteome Research*, 13**,** 4783-95.

Lee, S. B., Milgroom, M. G. & Taylor, J. W. 1988. A rapid, high yield mini-prep method for isolation of total genomic DNA from fungi. *Fungal Genetics Newsletter*, 35**,** 23-24.

Lee, Y. C., Kim, Y. J., Lee, K. Y., Kim, K. S., Kim, B. U., Kim, H. N., Kim, C. H. & Do, S. I. 1998. Cloning and expression of cDNA for a human Sia alpha 2,3Gal beta 1, 4GlcNA:alpha 2,8-sialyltransferase (hST8Sia III). *Archives of Biochemistry and Biophysics*, 360**,** 41-46.

Lehoux, S., Groux-Degroote, S., Cazet, A., Dhaenens, C. M., Maurage, C. A., Caillet-Boudin, M. L., Delannoy, P. & Krzewinski-Recchi, M. A. 2010. Transcriptional regulation of the human ST6GAL2 gene in cerebral cortex and neuronal cells. *Glycoconjugate Journal*, 27**,** 99-114.

Leroch, M., Mernke, D., Koppenhoefer, D., Schneider, P., Mosbach, A., Doehlemann, G. & Hahn, M. 2011. Living colors in the gray mold pathogen *Botrytis cinerea*: codon-optimized genes encoding green fluorescent protein and mCherry, which exhibit bright fluorescence. *Applied and Environmental Microbiology,* 77**,** 2887-97.

Li, Y., Sierra, A. M., Ai, H. W. & Campbell, R. E. 2008. Identification of sites within a monomeric red fluorescent protein that tolerate peptide insertion and testing of corresponding circular permutations. *Photochemistry and Photobiology,* 84**,** 111-19.

Lilie, H., Schwarz, E. & Rudolph, R. 1998. Advances in refolding of proteins produced in *E. coli*. *Current Opinion in Biotechnology,* 9**,** 497-501.

Lim, D., Hains, P., Walsh, B., Bergquist, P. & Nevalainen, H. 2001. Proteins associated with the cell envelope of *Trichoderma reesei*: a proteomic approach. *Proteomics,* 1**,** 899-909.

Liu, F. T. & Rabinovich, G. A. 2005. Galectins as modulators of tumour progression. *Nature Reviews: Cancer,* 5**,** 29-41.

Liu, R., Chen, L., Jiang, Y., Zhou, Z. & Zou, G. 2015. Efficient genome editing in filamentous fungus *Trichoderma reesei* using the CRISPR/Cas9 system. *Cell Discovery,* 1**,** 15007.

Liu, T., Wang, T., Li, X. & Liu, X. 2008. Improved heterologous gene expression in *Trichoderma reesei* by cellobiohydrolase I gene (*cbh1*) promoter optimization. *Acta Biochimica et Biophysica Sinica (Shanghai),* 40**,** 158-65.

Livak, K. J. & Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods,* 25**,** 402-8.

Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D. & Darnell, J. 2000. Protein glycosylation in the ER and Golgi complex.

Loke, I., Packer, N. H. & Thaysen-Andersen, M. 2015. Complementary LC-MS/MS-based *N*-glycan, *N*-glycopeptide, and intact *N*-glycoprotein profiling reveals unconventional Asn71-glycosylation of human neutrophil cathepsin G. *Biomolecules,* 5**,** 1832-54.

Lorang, J. M., Tuori, R. P., Martinez, J. P., Sawyer, T. L., Redman, R. S., Rollins, J. A., Wolpert, T. J., Johnson, K. B., Rodriguez, R. J., Dickman, M. B. & Ciuffetti, L. M. 2001. Green fluorescent protein is lighting up fungal biology. *Applied and Environmental Microbiology,* 67**,** 1987-94.

Lorito, M., Hayes, C. K., Di Pietro, A. & Harman, G. E. 1993. Biolistic transformation of *Trichoderma harzianum* and *Gliocladium virens* using plasmid and genomic DNA. *Current Genetics,* 24**,** 349-56.

Luley-Goedl, C., Czabany, T., Longus, K., Schmölzer, K., Zitzenbacher, S., Ribitsch, D., Schwab, H. & Nidetzky, B. 2016. Combining expression and process engineering for high-quality production of human sialyltransferase in *Pichia pastoris*. *Journal of Biotechnology*.

Lund, J., Takahashi, N., Nakagawa, H., Goodall, M., Bentley, T., Hindley, S. A., Tyler, R. & Jefferis, R. 1993. Control of IgG/Fc glycosylation: A comparison of oligosaccharides from chimeric human/mouse and mouse subclass immunoglobulin Gs. *Molecular Immunology,* 30**,** 741-748.

Macdougall, I. C. 2002. Darbepoetin alfa: a new therapeutic agent for renal anemia. *Kidney International. Supplement***,** 55-61.

Maclean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C. & Maccoss, M. J. 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics,* 26**,** 966-68.

Macrae, W. D., Buxton, F. P., Gwynne, D. I. & Davies, R. W. 1993. Heterologous protein secretion directed by a repressible acid phosphatase system of *Aspergillus niger*. *Gene,* 132**,** 193-98.

Madhavan, A. & Sukumaran, R. K. 2014. Promoter and signal sequence from filamentous fungus can drive recombinant protein production in the yeast *Kluyveromyces lactis*. *Bioresource Technology,* 165**,** 302-08.

Maguire, T. M. & Breen, K. C. 1995. A decrease in neural sialyltransferase activity in Alzheimer's disease. *Dementia,* 6**,** 185-90.

Malissard, M., Zeng, S. & Berger, E. G. 1999. The yeast expression system for recombinant glycosyltransferases. *Glycoconjugate Journal,* 16**,** 125-39.

Mandels, M., Parrish, F. W. & Reese, E. T. 1962. Sophorose as an inducer of cellulase in *Trichoderma viride*. *Journal of Bacteriology,* 83**,** 400-08.

Mandels, M. & Reese, E. T. 1957. Induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals. *Journal of Bacteriology,* 73**,** 269-78.

Mandels, M. & Reese, E. T. 1960. Induction of cellulase in fungi by cellobiose. *Journal of Bacteriology,* 79**,** 816-26.

Maor, R., Puyesky, M., Horwitz, B. & Sharon, A. 1998. Use of green fluorescent protein (GFP) for studying development and fungal-plant interaction in *Cochliobolus heterostrophus*. *Mycological Research,* 102**,** 491-496.

Margolles-Clark, E., Hayes, C. K., Harman, G. E. & Penttilä, M. 1996. Improved production of *Trichoderma harzianum* endochitinase by expression in *Trichoderma reesei*. *Applied and Environmental Microbiology,* 62**,** 2145-51.

Markaryan, A., Morozova, I., Yu, H. & Kolattukudy, P. E. 1994. Purification and characterization of an elastinolytic metalloprotease from *Aspergillus fumigatus* and

immunoelectron microscopic evidence of secretion of this enzyme by the fungus invading the murine lung. *Infection and Immunity,* 62**,** 2149-57.

Markham, P. 1995. Organelles of filamentous fungi. *The Growing Fungus.* Springer.

Martin, L. T., Marth, J. D., Varki, A. & Varki, N. M. 2002. Genetically altered mice with different sialyltransferase deficiencies show tissue-specific alterations in sialylation and sialic acid 9-*O*-acetylation. *Journal of Biological Chemistry,* 277**,** 32930-38.

Martinez, D., Berka, R. M., Henrissat, B., Saloheimo, M., Arvas, M., Baker, S. E., Chapman, J., Chertkov, O., Coutinho, P. M., Cullen, D., Danchin, E. G., Grigoriev, I. V., Harris, P., Jackson, M., Kubicek, C. P., Han, C. S., Ho, I., Larrondo, L. F., De Leon, A. L., Magnuson, J. K., Merino, S., Misra, M., Nelson, B., Putnam, N., Robbertse, B., Salamov, A. A., Schmoll, M., Terry, A., Thayer, N., Westerholm-Parvinen, A., Schoch, C. L., Yao, J., Barabote, R., Nelson, M. A., Detter, C., Bruce, D., Kuske, C. R., Xie, G., Richardson, P., Rokhsar, D. S., Lucas, S. M., Rubin, E. M., Dunn-Coleman, N., Ward, M. & Brettin, T. S. 2008. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nature Biotechnology,* 26**,** 553-60.

Marx, I. J., Van Wyk, N., Smit, S., Jacobson, D., Viljoen-Bloom, M. & Volschenk, H. 2013. Comparative secretome analysis of *Trichoderma asperellum* S4F8 and *Trichoderma reesei* Rut C30 during solid-state fermentation on sugarcane bagasse. *Biotechnol Biofuels,* 6**,** 172.

Maryon, E. B., Molloy, S. A. & Kaplan, J. H. 2007. *O*-linked glycosylation at threonine 27 protects the copper transporter hCTR1 from proteolytic cleavage in mammalian cells. *Journal of Biological Chemistry,* 282**,** 20376-87.

Mattox, S., Walrath, K., Ceiler, D., Smith, D. F. & Cummings, R. D. 1992. A solid-phase assay for the activity of CMPNeuAc:Gal beta 1-4GlcNAc-R alpha-2,6-sialyltransferase. *Analytical Biochemistry,* 206**,** 430-36.

Meng, F., Wei, D. & Wang, W. 2013. Heterologous protein expression in *Trichoderma reesei* using the *cbhII* promoter. *Plasmid,* 70**,** 272-76.

Merritt, J. H., Ollis, A. A., Fisher, A. C. & Delisa, M. P. 2013. Glycans-by-design: engineering bacteria for the biosynthesis of complex glycans and glycoconjugates. *Biotechnology and Bioengineering,* 110**,** 1550-64.

Meuris, L., Santens, F., Elson, G., Festjens, N., Boone, M., Dos Santos, A., Devos, S., Rousseau, F., Plets, E., Houthuys, E., Malinge, P., Magistrelli, G., Cons, L., Chatel, L., Devreese, B. & Callewaert, N. 2014. GlycoDelete engineering of mammalian cells simplifies *N*-glycosylation of recombinant proteins. *Nature Biotechnology,* 32**,** 485-89.

Meusser, B., Hirsch, C., Jarosch, E. & Sommer, T. 2005. ERAD: the long road to destruction. *Nature Cell Biology,* 7**,** 766-72.

Mirzaei, M., Pascovici, D., Atwell, B. J. & Haynes, P. A. 2012. Differential regulation of aquaporins, small GTPases and V-ATPases proteins in rice leaves subjected to drought stress and recovery. *Proteomics,* 12**,** 864-77.

Miyagi, T., Takahashi, K., Hata, K., Shiozaki, K. & Yamaguchi, K. 2012. Sialidase significance for cancer progression. *Glycoconjugate Journal,* 29**,** 567-77.

Miyagi, T. & Tsuiki, S. 1982. TI - Purification and characterization of β-galactoside (α2→6)sialyltransferase from rat liver and hepatomas. *European Journal of Biochemistry,* 126**,** 253-61.

Miyauchi, S., Te'o, V. S., J., Bergquist, P. L. & Nevalainen, K. M. 2013. Expression of a bacterial xylanase in *Trichoderma reesei* under the *egl2* and *cbh2* glycosyl hydrolase gene promoters. *New Biotechnology,* 30**,** 523-30.

Molinari, M., Eriksson, K. K., Calanca, V., Galli, C., Cresswell, P., Michalak, M. & Helenius, A. 2004. Contrasting functions of calreticulin and calnexin in glycoprotein folding and ER quality control. *Molecular Cell,* 13**,** 125-35.

Montefiori, D. C., Robinson, W. E., Jr. & Mitchell, W. M. 1988. Role of protein *N*-glycosylation in pathogenesis of human immunodeficiency virus type 1. *Proceedings of the National Academy of Sciences of the United States of America,* 85**,** 9248-52.

Montenecourt, B. S. & Eveleigh, D. E. 1977a. Preparation of mutants of *Trichoderma reesei* with enhanced cellulase production. *Applied and Environmental Microbiology,* 34**,** 777-82.

Montenecourt, B. S. & Eveleigh, D. E. 1977b. Semiquantitative plate assay for determination of cellulase production by *Trichoderma viride*. *Applied and Environmental Microbiology,* 33**,** 178-83.

Montenecourt, B. S. & Eveleigh, D. E. 1979a. Production and characterization of high yielding cellulase mutants of *Trichoderma reesei*. *TAPPI Journal,* 28**,** 101-108.

Montenecourt, B. S. & Eveleigh, D. E. 1979b. Selective screening methods for the isolation of high yielding cellulase mutants of *Trichoderma reesei*. *In:* Brown, R. & Jurasek, L. (eds.) *Hydrolysis of Cellulose: Mechanisms of Enzymatic and Acid Catalysis.* Washington, DC: American Chemical Society.

Moreland, J. L., Gramada, A., Buzko, O. V., Zhang, Q. & Bourne, P. E. 2005. The Molecular Biology Toolkit (MBT): a modular platform for developing molecular visualization applications. *BMC Bioinformatics,* 6**,** 21.

Moremen, K. W., Tiemeyer, M. & Nairn, A. V. 2012. Vertebrate protein glycosylation: diversity, synthesis and function. *Nature Reviews: Molecular Cell Biology,* 13**,** 448-62.

Muhlenhoff, M., Oltmann-Norden, I., Weinhold, B., Hildebrandt, H. & Gerardy-Schahn, R. 2009. Brain development needs sugar: the role of polysialic acid in controlling NCAM functions. *Biological Chemistry,* 390**,** 567-74.

Mulloy, B., Hart, G. W. & Stanley, P. 2009. Structural Analysis of Glycans. *In:* Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. & Etzler, M. E. (eds.) *Essentials of Glycobiology.* 2nd ed. Cold Spring Harbor (NY).

Murray, P., Aro, N., Collins, C., Grassick, A., Penttilä, M., Saloheimo, M. & Tuohy, M. 2004. Expression in *Trichoderma reesei* and characterisation of a thermostable family 3 beta-glucosidase from the moderately thermophilic fungus *Talaromyces emersonii*. *Protein Expression and Purification,* 38**,** 248-57.

Mustalahti, E., Saloheimo, M. & Joensuu, J. J. 2013. Intracellular protein production in *Trichoderma reesei* (*Hypocrea jecorina*) with hydrophobin fusion technology. *New Biotechnology,* 30**,** 262-68.

Nagar, A., Valleriani, A. & Lipowsky, R. 2011. Translation by ribosomes with mRNA degradation: Exclusion processes on aging tracks. *Journal of Statistical Physics,* 145**,** 1385-1404.

Nakari-Setälä, T., Paloheimo, M., Kallio, J., Vehmaanperä, J., Penttilä, M. & Saloheimo, M. 2009. Genetic modification of carbon catabolite repression in *Trichoderma reesei* for improved protein production. *Applied and Environmental Microbiology,* 75**,** 4853-60.

Nakayama, J., Fukuda, M. N., Fredette, B., Ranscht, B. & Fukuda, M. 1995. Expression cloning of a human polysialyltransferase that forms the polysialylated neural cell adhesion molecule present in embryonic brain. *Proceedings of the National Academy of Sciences of the United States of America,* 92**,** 7031-35.

Nara, K., Watanabe, Y., Maruyama, K., Kasahara, K., Nagai, Y. & Sanai, Y. 1994. Expression cloning of a CMP-NeuAc:NeuAc alpha 2-3Gal beta 1-4Glc beta 1-1'Cer alpha 2,8-sialyltransferase (GD3 synthase) from human melanoma cells. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 7952-56.

Natunen, J., Landowski, C., Saloheimo, M., Ostermeier, C., Sommer, B. P. & Wahl, R. 2015. Production of glycoproteins having increased *N*-glycosylation site occupancy. Google Patents.

Ncbi 2015. beta-galactoside alpha-2,6-sialyltransferase 1 isoform a [*Homo sapiens*]. NCBI.

Nevalainen, H., Hekelaar, J., Uusitalo, J., Te'o, J., Jonkers, I., P.L.B., B. & Penttilä, M. 2004. Transcriptional and proteomic changes in *Trichoderma reesei* expressing a heterologous bacterial *xynB* gene from the thermophile *Dictyoglomus thermophilum*. *7th European Conference on Fungal Genetics, Copenhagen 17-20 April 2004*, Poster.

Nevalainen, H., Lavygina, I., Neethling, D. & Packer, N. 1995. The biochemical nature of the cell envelope of a high cellulase-secreting mutant differs from that of the *Trichoderma reesei* wild type. *Journal of Biotechnology,* 42**,** 53-59.

Nevalainen, H. & Peterson, R. 2014a. Chapter 7 - Heterologous expression of proteins in *Trichoderma*. *Biotechnology and Biology of Trichoderma,* 1, 89-102.

Nevalainen, H. & Peterson, R. 2014b. Making recombinant proteins in filamentous fungi- are we expecting too much? *Frontiers in Microbiology,* 5**,** 75.

Nevalainen, K. M., Te'o, V. S. & Bergquist, P. L. 2005. Heterologous protein expression in filamentous fungi. *Trends in Biotechnology,* 23**,** 468-74.

Ng, D. T., Brown, J. D. & Walter, P. 1996. Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *Journal of Cell Biology,* 134**,** 269-78.

Nizet, V. & Esko, J. D. 2009. Bacterial and Viral Infections. *In:* Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. & Etzler, M. E. (eds.) *Essentials of Glycobiology.* 2nd ed. Cold Spring Harbor (NY).

Nykänen, M., Birch, D., Peterson, R., Yu, H., Kautto, L., Gryshyna, A., Te'o, J. & Nevalainen, H. 2016. Ultrastructural features of the early secretory pathway in *Trichoderma reesei*. *Current Genetics,* 62**,** 455-65.

Nykänen, M., Saarelainen, R., Raudaskoski, M., Nevalainen, K. & Mikkonen, A. 1997. Expression and secretion of barley cysteine endopeptidase B and cellobiohydrolase I in *Trichoderma reesei*. *Applied and Environmental Microbiology,* 63**,** 4929-37.

Nykänen, M. J. 2002. Protein secretion in Trichoderma reesei. Expression, secretion and maturation of cellobiohydrolase I, barley cysteine proteinase and calf chymosin in Rut-C30. Ph.D thesis, University of Jyväskylä.

Nykänen, M. J., Raudaskoski, M., Nevalainen, H. & Mikkonen, A. 2002. Maturation of barley cysteine endopeptidase expressed in *Trichoderma reesei* is distorted by incomplete processing. *Canadian Journal of Microbiology,* 48**,** 138-50.

Nyyssönen, E., Keränen, S., Penttilä, M., Demolder, J. & Contreras, R. 1995. Protein production by the filamentous fungus *Trichoderma reesei*: secretion of active antibody molecules. *Canadian journal of botany,* 73**,** 885-90.

Nyyssönen, E., Penttilä, M., Harkki, A., Saloheimo, A., Knowles, J. K. & Keränen, S. 1993. Efficient production of antibody fragments by the filamentous fungus *Trichoderma reesei*. *Bio/Technology,* 11**,** 591-95.

Ohtsubo, K. & Marth, J. D. 2006. Glycosylation in cellular mechanisms of health and disease. *Cell,* 126**,** 855-67.

Okajima, T., Fukumoto, S., Miyazaki, H., Ishida, H., Kiso, M., Furukawa, K., Urano, T. & Furukawa, K. 1999. Molecular cloning of a novel alpha2,3-sialyltransferase (ST3Gal VI) that sialylates type II lactosamine structures on glycoproteins and gly*coli*pids. *Journal of Biological Chemistry,* 274**,** 11479-86.

Okungbowa, F., Ghosh, A., Chowdhury, R., Chaudhuri, P., Basu, A. & Pal, K. 2007. Mechanical lysis of *Candida* cells for crude protein and enzymatic activity estimation: Comparison of three methods. *World Journal of Medical Sciences,* 2**,** 101-104.

Oslowski, C. M. & Urano, F. 2011. Measuring ER stress and the unfolded protein response using mammalian tissue culture system. *Methods in Enzymology,* 490**,** 71-92.

Otvos, L., Jr., Krivulka, G. R., Urge, L., Szendrei, G. I., Nagy, L., Xiang, Z. Q. & Ertl, H. C. 1995. Comparison of the effects of amino acid substitutions and beta-*N*- vs. alpha-*O*-

glycosylation on the T-cell stimulatory activity and conformation of an epitope on the rabies virus glycoprotein. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1267**,** 55-64.

Paddock, S. W. 1999. Confocal laser scanning microscopy. *BioTechniques,* 27**,** 992-1007.

Pakula, T. M., Laxell, M., Huuskonen, A., Uusitalo, J., Saloheimo, M. & Penttilä, M. 2003. The effects of drugs inhibiting protein secretion in the filamentous fungus *Trichoderma reesei*. Evidence for down-regulation of genes that encode secreted proteins in the stressed cells. *Journal of Biological Chemistry,* 278**,** 45011-20.

Palacpac, N. Q., Yoshida, S., Sakai, H., Kimura, Y., Fujiyama, K., Yoshida, T. & Seki, T. 1999. Stable expression of human beta1,4-galactosyltransferase in plant cells modifies *N*-linked glycosylation patterns. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 4692-97.

Paloheimo, M., Mäntylä, A., Kallio, J., Puranen, T. & Suominen, P. 2007. Increased production of xylanase by expression of a truncated version of the *xyn11A* gene from *Nonomuraea flexuosa* in *Trichoderma reesei*. *Applied and Environmental Microbiology,* 73**,** 3215-24.

Paloheimo, M., Mäntylä, A., Kallio, J. & Suominen, P. 2003. High-yield production of a bacterial xylanase in the filamentous fungus *Trichoderma reesei* requires a carrier polypeptide with an intact domain structure. *Applied and Environmental Microbiology,* 69**,** 7073-82.

Patel, T. P., Parekh, R. B., Moellering, B. J. & Prior, C. P. 1992. Different culture methods lead to differences in glycosylation of a murine IgG monoclonal antibody. *Biochemical Journal,* 285 (Pt 3)**,** 839-45.

Paulson, J. C. & Rademacher, C. 2009. Glycan terminator. *Nature Structural & Molecular Biology,* 16**,** 1121-22.

Peberdy, J. F. 1994. Protein secretion in filamentous fungi--trying to understand a highly productive black box. *Trends in Biotechnology,* 12**,** 50-7.

Pereira-Junior, O. S., Pereira, R. V., Silva, C. S., Castro-Borges, W., Sa, R. G., Cabral, F. J., Silva, S. H., Soares, C. S., Morais, E. R., Moreira, E. B., Magalhaes, L. G., De Paula, F. M. & Rodrigues, V. 2013. Investigation on the 19S ATPase proteasome subunits (Rpt1-6) conservation and their differential gene expression in *Schistosoma mansoni*. *Parasitology Research,* 112**,** 235-42.

Peterson, R., Cheah, W. Y., Grinyer, J. & Packer, N. 2013. Glycoconjugates in human milk: protecting infants from disease. *Glycobiology,* 23**,** 1425-38.

Peterson, R., Grinyer, J. & Nevalainen, H. 2011. Extracellular hydrolase profiles of fungi isolated from koala faeces invite biotechnological interest. *Mycological Progress,* 10**,** 207-218.

Peterson, R. & Nevalainen, H. 2012. *Trichoderma reesei* RUT-C30--thirty years of strain improvement. *Microbiology,* 158**,** 58-68.

Piatkevich, K. D. & Verkhusha, V. V. 2011. Guide to red fluorescent proteins and biosensors for flow cytometry. *Methods in Cell Biology,* 102**,** 431-61.

Pilatte, Y., Bignon, J. & Lambre, C. R. 1993. Sialic acids as important molecules in the regulation of the immune system: pathophysiological implications of sialidases in immunity. *Glycobiology,* 3**,** 201-18.

Poidevin, L., Levasseur, A., Paes, G., Navarro, D., Heiss-Blanquet, S., Asther, M. & Record, E. 2009. Heterologous production of the *Piromyces equi* cinnamoyl esterase in *Trichoderma reesei* for biotechnological applications. *Letters in Applied Microbiology,* 49**,** 673-78.

Pollack, J. K., Harris, S. D. & Marten, M. R. 2009. Autophagy in filamentous fungi. *Fungal Genetics and Biology,* 46**,** 1-8.

Punt, P. J., Van Biezen, N., Conesa, A., Albers, A., Mangnus, J. & Van Den Hondel, C. 2002. Filamentous fungi as cell factories for heterologous protein production. *Trends in Biotechnology,* 20**,** 200-06.

Rao, F. V., Rich, J. R., Rakic, B., Buddai, S., Schwartz, M. F., Johnson, K., Bowe, C., Wakarchuk, W. W., Defrees, S., Withers, S. G. & Strynadka, N. C. 2009. Structural insight into mammalian sialyltransferases. *Nature Structural & Molecular Biology,* 16**,** 1186-88.

Rauscher, M., Mendgen, K. & Deising, H. 1995. Extracellular proteases of the rust fungus *Uromyces viciae-fabae*. *Experimental Mycology,* 19**,** 26-34.

Ravetch, J. V. & Bolland, S. 2001. IgG Fc receptors. *Annual Review of Immunology,* 19**,** 275-90.

Rearick, J. I., Sadler, J. E., Paulson, J. C. & Hill, R. L. 1979. Enzymatic characterization of beta D-galactoside alpha2 leads to 3 sialyltransferase from porcine submaxillary gland. *Journal of Biological Chemistry,* 254**,** 4444-4451.

Reichard, U., Lechenne, B., Asif, A. R., Streit, F., Grouzmann, E., Jousson, O. & Monod, M. 2006. Sedolisins, a new class of secreted proteases from *Aspergillus fumigatus* with endoprotease or tripeptidyl-peptidase activity at acidic pHs. *Applied and Environmental Microbiology,* 72**,** 1739-48.

Ribitsch, D., Zitzenbacher, S., Augustin, P., Schmolzer, K., Czabany, T., Luley-Goedl, C., Thomann, M., Jung, C., Sobek, H., Muller, R., Nidetzky, B. & Schwab, H. 2014. High-quality production of human alpha-2,6-sialyltransferase in *Pichia pastoris* requires control over *N*-terminal truncations by host-inherent protease activities. *Microbial Cell Factories,* 13**,** 138.

Ries, L., Belshaw, N. J., Ilmén, M., Penttilä, M. E., Alapuranen, M. & Archer, D. B. 2014. The role of CRE1 in nucleosome positioning within the *cbh1* promoter and coding regions of *Trichoderma reesei*. *Applied Microbiology and Biotechnology,* 98**,** 749-62.

Rivera, A. L., Magana-Ortiz, D., Gomez-Lim, M., Fernandez, F. & Loske, A. M. 2014. Physical methods for genetic transformation of fungi and yeast. *Physics of Life Reviews,* 11**,** 184-203.

Rogers, G. N., Paulson, J. C., Daniels, R. S., Skehel, J. J., Wilson, I. A. & Wiley, D. C. 1983a. Single amino acid substitutions in influenza haemagglutinin change receptor binding specificity. *Nature,* 304**,** 76-78.

Rogers, G. N., Pritchett, T. J., Lane, J. L. & Paulson, J. C. 1983b. Differential sensitivity of human, avian, and equine influenza A viruses to a glycoprotein inhibitor of infection: selection of receptor specific variants. *Virology,* 131**,** 394-408.

Rosano, G. L. & Ceccarelli, E. A. 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology,* 5**,** 172.

Rosen, S. D., Chi, S. I., True, D. D., Singer, M. S. & Yednock, T. A. 1989. Intravenously injected sialidase inactivates attachment sites for lymphocytes on high endothelial venules. *Journal of Immunology,* 142**,** 1895-902.

Rosen, S. D., Singer, M. S., Yednock, T. A. & Stoolman, L. M. 1985. Involvement of sialic acid on endothelial cells in organ-specific lymphocyte recirculation. *Science,* 228**,** 1005-07.

Roth, J. 2002. Protein *N*-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. *Chemical Reviews,* 102**,** 285-303.

Ryu, D. D. & Mandels, M. 1980. Cellulases: biosynthesis and applications. *Enzyme and Microbial Technology,* 2**,** 91-102.

Saarelainen, R., Mäntylä, A., Nevalainen, H. & Suominen, P. 1997. Expression of barley endopeptidase B in *Trichoderma reesei*. *Applied and Environmental Microbiology,* 63**,** 4938-40.

Salcedo, J., Barbera, R., Matencio, E., Alegria, A. & Lagarda, M. J. 2013. Gangliosides and sialic acid effects upon newborn pathogenic bacteria adhesion: an in vitro study. *Food Chemistry,* 136**,** 726-34.

Salles, B. C., Te'o, V. S., Gibbs, M. D., Bergquist, P. L., Filho, E. X., Ximenes, E. A. & Nevalainen, K. M. 2007. Identification of two novel xylanase-encoding genes (*xyn5* and *xyn6*) from *Acrophialophora nainiana* and heterologous expression of *xyn6* in *Trichoderma reesei*. *Biotechnology Letters,* 29**,** 1195-201.

Saloheimo, M. & Pakula, T. M. 2012. The cargo and the transport system: secreted proteins and protein secretion in *Trichoderma reesei* (*Hypocrea jecorina*). *Microbiology,* 158**,** 46-57.

Saloheimo, M., Valkonen, M. & Penttilä, M. 2003. Activation mechanisms of the HAC1-mediated unfolded protein response in filamentous fungi. *Molecular Microbiology,* 47**,** 1149-61.

Samraj, A. N., Laubli, H., Varki, N. & Varki, A. 2014. Involvement of a non-human sialic Acid in human cancer. *Frontiers in Oncology,* 4**,** 33.

Samraj, A. N., Pearce, O. M., Laubli, H., Crittenden, A. N., Bergfeld, A. K., Banda, K., Gregg, C. J., Bingman, A. E., Secrest, P., Diaz, S. L., Varki, N. M. & Varki, A. 2015. A red meat-derived glycan promotes inflammation and cancer progression. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** 542-47.

Samyn-Petit, B., Krzewinski-Recchi, M. A., Steelant, W. F., Delannoy, P. & Harduin-Lepers, A. 2000. Molecular cloning and functional expression of human ST6GalNAc II. Molecular expression in various human cultured cells. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1474**,** 201-11.

Sasaki, K., Kurata, K., Kojima, N., Kurosawa, N., Ohta, S., Hanai, N., Tsuji, S. & Nishi, T. 1994. Expression cloning of a GM3-specific alpha-2,8-sialyltransferase (GD3 synthase). *Journal of Biological Chemistry,* 269**,** 15950-56.

Sasaki, K., Watanabe, E., Kawashima, K., Sekine, S., Dohi, T., Oshima, M., Hanai, N., Nishi, T. & Hasegawa, M. 1993. Expression cloning of a novel Gal beta (1-3/1-4) GlcNAc alpha 2,3-sialyltransferase using lectin resistance selection. *Journal of Biological Chemistry,* 268**,** 22782-87.

Scheidegger, E. P., Sternberg, L. R., Roth, J. & Lowe, J. B. 1995. A human STX cDNA confers polysialic acid expression in mammalian cells. *Journal of Biological Chemistry,* 270**,** 22685-88.

Schmidt, T. G. & Skerra, A. 2007. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nature Protocols,* 2**,** 1528-35.

Schultz, M. J., Swindall, A. F., Wright, J. W., Sztul, E. S., Landen, C. N. & Bellis, S. L. 2013. ST6Gal-I sialyltransferase confers cisplatin resistance in ovarian tumor cells. *Journal of Ovarian Research,* 6**,** 25.

Schwarz, F., Huang, W., Li, C., Schulz, B. L., Lizak, C., Palumbo, A., Numao, S., Neri, D., Aebi, M. & Wang, L. X. 2010. A combined method for producing homogeneous glycoproteins with eukaryotic *N*-glycosylation. *Nature Chemical Biology,* 6**,** 264-66.

Seiboth, B., Hartl, L., Salovuori, N., Lanthaler, K., Robson, G. D., Vehmaanperä, J., Penttilä, M. E. & Kubicek, C. P. 2005. Role of the bga1-encoded extracellular {beta}-galactosidase of *Hypocrea jecorina* in cellulase induction by lactose. *Applied and Environmental Microbiology,* 71**,** 851-7.

Seidel, U. J., Schlegel, P. & Lang, P. 2013. Natural killer cell mediated antibody-dependent cellular cytotoxicity in tumor immunotherapy with therapeutic antibodies. *Frontiers in Immunology,* 4**,** 76.

Senda, M., Ito, A., Tsuchida, A., Hagiwara, T., Kaneda, T., Nakamura, Y., Kasama, K., Kiso, M., Yoshikawa, K., Katagiri, Y., Ono, Y., Ogiso, M., Urano, T., Furukawa, K., Oshima, S. & Furukawa, K. 2007. Identification and expression of a sialyltransferase responsible for the synthesis of disialylgalactosylgloboside in normal and malignant kidney cells: downregulation of ST6GalNAc VI in renal cancers. *Biochemical Journal,* 402**,** 459-70.

Severi, E., Hood, D. W. & Thomas, G. H. 2007. Sialic acid utilization by bacterial pathogens. *Microbiology,* 153**,** 2817-22.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E. & Tsien, R. Y. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnology,* 22**,** 1567-72.

Sharma, R., Katoch, M., Srivastava, P. S. & Qazi, G. N. 2009. Approaches for refining heterologous protein production in filamentous fungi. *World Journal of Microbiology and Biotechnology,* 25**,** 2083-94.

Sharon, E., Chet, I. & Spiegel, Y. 2011. *Trichoderma as a Biological Control Agent*.

Sheir-Neiss, G. & Montenecourt, B. 1984. Characterization of the secreted cellulases of *Trichoderma reesei* wild type and mutants during controlled fermentations. *Applied Microbiology and Biotechnology,* 20**,** 46-53.

Shibuya, N., Goldstein, I. J., Broekaert, W. F., Nsimba-Lubaki, M., Peeters, B. & Peumans, W. J. 1987. The elderberry (*Sambucus nigra* L.) bark lectin recognizes the Neu5Ac(alpha 2-6)Gal/GalNAc sequence. *Journal of Biological Chemistry,* 262**,** 1596-601.

Shimma, Y., Saito, F., Oosawa, F. & Jigami, Y. 2006. Construction of a library of human glycosyltransferases immobilized in the cell wall of *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology,* 72**,** 7003-12.

Simmons, E. G. 1977. Classification of some cellulase-producing *Trichoderma* species. *In:* Simmons, E. G. (ed.) *Second International Mycological Congress, Abstracts,*. Tampa, FL: University of South Florida.

Simon, B. M., Malisan, F., Testi, R., Nicotera, P. & Leist, M. 2002. Disialoganglioside GD3 is released by microglia and induces oligodendrocyte apoptosis. *Cell Death and Differentiation,* 9**,** 758-67.

Simons, J. F., Ferro-Novick, S., Rose, M. D. & Helenius, A. 1995. BiP/Kar2p serves as a molecular chaperone during carboxypeptidase Y folding in yeast. *Journal of Cell Biology,* 130**,** 41-9.

Sims, A. H., Dunn-Coleman, N. S., Robson, G. D. & Oliver, S. G. 2004. Glutamic protease distribution is limited to filamentous fungi. *FEMS Microbiology Letters,* 239**,** 95-101.

Singh, A., Taylor, L. E., 2nd, Vander Wall, T. A., Linger, J., Himmel, M. E., Podkaminer, K., Adney, W. S. & Decker, S. R. 2015. Heterologous protein expression in *Hypocrea jecorina*: a historical perspective and new developments. *Biotechnology Advances,* 33**,** 142-54.

Skretas, G., Carroll, S., Defrees, S., Schwartz, M. F., Johnson, K. F. & Georgiou, G. 2009. Expression of active human sialyltransferase ST6GalNAcI in *Escherichia coli*. *Microbial Cell Factories,* 8**,** 50.

Smith, T. L., Gaskell, J., Berka, R. M., Yang, M., Henner, D. J. & Cullen, D. 1990. The promoter of the glucoamylase-encoding gene of *Aspergillus niger* functions in *Ustilago maydis*. *Gene,* 88**,** 259-62.

Smith, W., Jantti, J., Oja, M. & Saloheimo, M. 2014. Comparison of intracellular and secretion-based strategies for production of human alpha-galactosidase A in the filamentous fungus *Trichoderma reesei*. *BMC Biotechnology,* 14**,** 91.

Son, Y. D., Jeong, Y. T., Park, S. Y. & Kim, J. H. 2011. Enhanced sialylation of recombinant human erythropoietin in Chinese hamster ovary cells by combinatorial engineering of selected genes. *Glycobiology,* 21**,** 1019-28.

Southern, E. 2006. Southern blotting. *Nature Protocols,* 1**,** 518-25.

Speicher, K. D., Gorman, N. & Speicher, D. W. 2009. *N*-terminal sequence analysis of proteins and peptides. *Current Protocols in Protein Science,* Chapter 11**,** Unit11 10.

Spellig, T., Bottin, A. & Kahmann, R. 1996. Green fluorescent protein (GFP) as a new vital marker in the phytopathogenic fungus *Ustilago maydis*. *Molecular and General Genetics,* 252**,** 503-09.

Spiro, R. G. 2002. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology,* 12**,** 43R-56R.

Sriranganadane, D., Reichard, U., Salamin, K., Fratti, M., Jousson, O., Waridel, P., Quadroni, M., Neuhaus, J. M. & Monod, M. 2011. Secreted glutamic protease rescues aspartic protease Pep deficiency in *Aspergillus fumigatus* during growth in acidic protein medium. *Microbiology,* 157**,** 1541-50.

Stanley P, C. R. D. 2009. *Structures Common to Different Glycans*, Cold Spring Harbor (NY).

Stanley, P., Schachter, H. & Taniguchi, N. 2009. *N*-Glycans. *In:* Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. & Etzler, M. E. (eds.) *Essentials of Glycobiology.* 2nd ed. Cold Spring Harbor (NY).

Steenackers, A., Vanbeselaere, J., Cazet, A., Bobowski, M., Rombouts, Y., Colomb, F., Le Bourhis, X., Guerardel, Y. & Delannoy, P. 2012. Accumulation of unusual gangliosides G(Q3) and G(P3) in breast cancer cells expressing the G(D3) synthase. *Molecules,* 17**,** 9559-72.

Stencel-Baerenwald, J. E., Reiss, K., Reiter, D. M., Stehle, T. & Dermody, T. S. 2014. The sweet spot: defining virus-sialic acid interactions. *Nature Reviews: Microbiology,* 12**,** 739-49.

Sticher, U., Gross, H. J. & Brossmer, R. 1991. Purification and characterization of alpha (2-6)-sialyltransferase from human liver. *Glycoconjugate Journal,* 8**,** 45-54.

Stolz, A. & Wolf, D. H. 2010. Endoplasmic reticulum associated protein degradation: a chaperone assisted journey to hell. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1803**,** 694-705.

Su, X., Schmitz, G., Zhang, M., Mackie, R. I. & Cann, I. K. 2012. Heterologous gene expression in filamentous fungi. *Advances in Applied Microbiology,* 81**,** 1-61.

Subach, F. V., Patterson, G. H., Manley, S., Gillette, J. M., Lippincott-Schwartz, J. & Verkhusha, V. V. 2009. Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nature Methods,* 6**,** 153-59.

Sun, A. 2008. Fluorescent tagging of the endoplasmic reticulum in *Trichoderma reesei.* Honours thesis, Macquarie University.

Sun, A. 2013. Heterologous expression of the mammalian peptide hormone obestatin in *Trichoderma reesei.* Ph.D thesis, Macquarie University.

Sun, A., Peterson, R., Te'o, J. & Nevalainen, H. 2016. Expression of the mammalian peptide hormone obestatin in *Trichoderm reesei*. *New Biotechnology*, 33**,** 99-106.

Suzuki, T., Arai, S., Takeuchi, M., Sakurai, C., Ebana, H., Higashi, T., Hashimoto, H., Hatsuzawa, K. & Wada, I. 2012. Development of cysteine-free fluorescent proteins for the oxidative environment. *PloS One,* 7**,** e37551.

Swindall, A. F. & Bellis, S. L. 2011. Sialylation of the Fas death receptor by ST6Gal-I provides protection against Fas-mediated apoptosis in colon carcinoma cells. *Journal of Biological Chemistry,* 286**,** 22982-90.

Swindall, A. F., Londono-Joshi, A. I., Schultz, M. J., Fineberg, N., Buchsbaum, D. J. & Bellis, S. L. 2013. ST6Gal-I protein expression is upregulated in human epithelial tumors and correlates with stem cell markers in normal tissues and colon cancer cell lines. *Cancer Research,* 73**,** 2368-78.

Takashima, S., Tsuji, S. & Tsujimoto, M. 2002. Characterization of the second type of human beta-galactoside alpha 2,6-sialyltransferase (ST6Gal II), which sialylates Galbeta 1,4GlcNAc structures on oligosaccharides preferentially. Genomic analysis of human sialyltransferase genes. *Journal of Biological Chemistry,* 277**,** 45719-28.

Tan, E., Almaraz, R. T., Khanna, H. S., Du, J. & Yarema, K. J. 2010. Experimental design considerations for in vitro non-natural glycan display via metabolic oligosaccharide engineering. *Current Protocols in Chemical Biology,* 2**,** 171-94.

Taylor, M. E. & Drickamer, K. 2007. Paradigms for glycan-binding receptors in cell adhesion. *Current Opinion in Cell Biology,* 19**,** 572-77.

Taylor, S., Yee, D., Academia, K., Paulus, A. & Alburo, A. 2000. The dynamic range effect on protein quantitation in polyacrylamide gels and on Western blots.

Te'o, V. S., Bergquist, P. L. & Nevalainen, K. M. 2002. Biolistic transformation of *Trichoderma reesei* using the Bio-Rad seven barrels Hepta Adaptor system. *Journal of Microbiological Methods,* 51**,** 393-9.

Te'o, V. S., Cziferszky, A. E., Bergquist, P. L. & Nevalainen, K. M. 2000. Codon optimization of xylanase gene *xynB* from the thermophilic bacterium *Dictyoglomus thermophilum* for expression in the filamentous fungus *Trichoderma reesei*. *FEMS Microbiology Letters,* 190**,** 13-19.

Te'o, V. S. J. & Nevalainen, K. M. H. 2011. Multiple promoter platform for protein production. Google Patents.

Te'o, V. S. J. & Nevalainen, K. M. H. 2015. Use of the biolistic particle delivery system to transform fungal genomes. *In:* Van Den Berg, A. M. & Maruthachalam, K. (eds.) *Genetic Transformation Systems in Fungi, Volume 1*. Cham: Springer International Publishing.

Teintenier-Lelievre, M., Julien, S., Juliant, S., Guerardel, Y., Duonor-Cerutti, M., Delannoy, P. & Harduin-Lepers, A. 2005. Molecular cloning and expression of a human hST8Sia VI (alpha2,8-sialyltransferase) responsible for the synthesis of the diSia motif on *O*-glycosylproteins. *Biochemical Journal,* 392**,** 665-74.

Teuton, J. R. & Brandt, C. R. 2007. Sialic acid on herpes simplex virus type 1 envelope glycoproteins is required for efficient infection of cells. *Journal of Virology,* 81**,** 3731-39.

Tharakaraman, K., Raman, R., Viswanathan, K., Stebbins, N. W., Jayaraman, A., Krishnan, A., Sasisekharan, V. & Sasisekharan, R. 2013. Structural determinants for naturally evolving H5N1 hemagglutinin to switch its receptor specificity. *Cell*, 153**,** 1475-85.

Thaysen-Andersen, M., Wilkinson, B. L., Payne, R. J. & Packer, N. H. 2011. Site-specific characterisation of densely *O*-glycosylated mucin-type peptides using electron transfer dissociation ESI-MS/MS. *Electrophoresis,* 32**,** 3536-45.

Thrower, J. S., Hoffman, L., Rechsteiner, M. & Pickart, C. M. 2000. Recognition of the polyubiquitin proteolytic signal. *EMBO Journal,* 19**,** 94-102.

Travers, K. J., Patil, C. K., Wodicka, L., Lockhart, D. J., Weissman, J. S. & Walter, P. 2000. Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell*, 101**,** 249-58.

Trombetta, E. S. 2003. The contribution of *N*-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis. *Glycobiology,* 13**,** 77R-91R.

Tsuchida, A., Ogiso, M., Nakamura, Y., Kiso, M., Furukawa, K. & Furukawa, K. 2005. Molecular cloning and expression of human ST6GalNAc III: restricted tissue distribution and substrate specificity. *Journal of Biochemistry,* 138**,** 237-43.

Tsuchida, A., Okajima, T., Furukawa, K., Ando, T., Ishida, H., Yoshida, A., Nakamura, Y., Kannagi, R., Kiso, M. & Furukawa, K. 2003. Synthesis of disialyl Lewis a (Le(a)) structure in colon cancer cell lines by a sialyltransferase, ST6GalNAc VI, responsible for the synthesis of alpha-series gangliosides. *Journal of Biological Chemistry,* 278**,** 22787-94.

Tsuchiya, K., Tada, S., Gomi, K., Kitamoto, K., Kumagai, C., Jigami, Y. & Tamura, G. 1992. High level expression of the synthetic human lysozyme gene in *Aspergillus oryzae*. *Applied Microbiology and Biotechnology,* 38**,** 109-14.

Umana, P., Jean-Mairet, J., Moudry, R., Amstutz, H. & Bailey, J. E. 1999. Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *Nature Biotechnology,* 17**,** 176-80.

Uusitalo, J. M., Nevalainen, K. M., Harkki, A. M., Knowles, J. K. & Penttilä, M. E. 1991. Enzyme production by recombinant *Trichoderma reesei* strains. *Journal of Biotechnology,* 17**,** 35-49.

Vacca-Smith, A. M., Van Wuyckhuyse, B. C., Tabak, L. A. & Bowen, W. H. 1994. The effect of milk and casein proteins on the adherence of *Streptococcus* mutans to saliva-coated hydroxyapatite. *Archives of Oral Biology,* 39**,** 1063-69.

Vachieri, S. G., Xiong, X., Collins, P. J., Walker, P. A., Martin, S. R., Haire, L. F., Zhang, Y., Mccauley, J. W., Gamblin, S. J. & Skehel, J. J. 2014. Receptor binding by H10 influenza viruses. *Nature,* 511**,** 475-77.

Valderrama-Rincon, J. D., Fisher, A. C., Merritt, J. H., Fan, Y. Y., Reading, C. A., Chhiba, K., Heiss, C., Azadi, P., Aebi, M. & Delisa, M. P. 2012. An engineered eukaryotic protein glycosylation pathway in *Escherichia coli*. *Nature Chemical Biology,* 8**,** 434-36.

Van Den Hombergh, J. P., Van De Vondervoort, P. J., Fraissinet-Tachet, L. & Visser, J. 1997. *Aspergillus* as a host for heterologous protein production: the problem of proteases. *Trends in Biotechnology,* 15**,** 256-63.

Van Den Steen, P., Rudd, P. M., Dwek, R. A. & Opdenakker, G. 1998. Concepts and principles of *O*-linked glycosylation. *Critical Reviews in Biochemistry and Molecular Biology,* 33**,** 151-208.

Varki, A. 2009. Multiple changes in sialic acid biology during human evolution. *Glycoconjugate Journal,* 26**,** 231-45.

Varki, A. & Gagneux, P. 2012. Multifarious roles of sialic acids in immunity. *Annals of the New York Academy of Sciences,* 1253**,** 16-36.

Varki, A. & Lowe, J. B. 2009. Biological Roles of Glycans. *In:* Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. & Etzler, M. E. (eds.) *Essentials of Glycobiology.* 2nd ed. Cold Spring Harbor (NY).

Varki, A. & Schauer, R. 2009. *Sialic Acids*, Cold Spring Harbor Laboratory Press.

Vembar, S. S. & Brodsky, J. L. 2008. One step at a time: endoplasmic reticulum-associated degradation. *Nature Reviews: Molecular Cell Biology,* 9**,** 944-57.

Vervecken, W., Kaigorodov, V., Callewaert, N., Geysens, S., De Vusser, K. & Contreras, R. 2004. In vivo synthesis of mammalian-like, hybrid-type *N*-glycans in *Pichia pastoris*. *Applied and Environmental Microbiology,* 70**,** 2639-46.

Vimr, E. R., Kalivoda, K. A., Deszo, E. L. & Steenbergen, S. M. 2004. Diversity of microbial sialic acid metabolism. *Microbiology and Molecular Biology Reviews,* 68**,** 132-53.

Walsh, G. & Jefferis, R. 2006. Post-translational modifications in the context of therapeutic proteins. *Nature Biotechnology,* 24**,** 1241-52.

Wang, B. & Brand-Miller, J. 2003. The role and potential of sialic acid in human nutrition. *European Journal of Clinical Nutrition,* 57**,** 1351-69.

Wang, Y., Tan, J., Sutton-Smith, M., Ditto, D., Panico, M., Campbell, R. M., Varki, N. M., Long, J. M., Jaeken, J., Levinson, S. R., Wynshaw-Boris, A., Morris, H. R., Le, D., Dell, A., Schachter, H. & Marth, J. D. 2001. Modeling human congenital disorder of glycosylation type IIa in the mouse: conservation of asparagine-linked glycan-dependent functions in mammalian physiology and insights into disease pathogenesis. *Glycobiology,* 11**,** 1051-70.

Ward, M., Lin, C., Victoria, D. C., Fox, B. P., Fox, J. A., Wong, D. L., Meerman, H. J., Pucci, J. P., Fong, R. B., Heng, M. H., Tsurushita, N., Gieswein, C., Park, M. & Wang, H. 2004. Characterization of humanized antibodies secreted by *Aspergillus niger*. *Applied and Environmental Microbiology,* 70**,** 2567-76.

Ward, P. P., Piddington, C. S., Cunningham, G. A., Zhou, X., Wyatt, R. D. & Conneely, O. M. 1995. A system for production of commercial quantities of human lactoferrin: a broad spectrum natural antibiotic. *Bio/Technology,* 13**,** 498-503.

Ward, T. J., Honeycutt, R. L. & Derr, J. N. 1997. Nucleotide sequence evolution at the kappa-casein locus: evidence for positive selection within the family *Bovidae*. *Genetics,* 147**,** 1863-72.

Weihofen, A., Binns, K., Lemberg, M. K., Ashman, K. & Martoglio, B. 2002. Identification of signal peptide peptidase, a presenilin-type aspartic protease. *Science,* 296**,** 2215-18.

Weinstein, J., De Souza-E-Silva, U. & Paulson, J. 1982. Sialylation of glycoprotein oligosaccharides *N*-linked to asparagine. Enzymatic characterization of a Gal beta 1 to 3 (4) GlcNAc alpha 2 to 3 sialyltransferase and a Gal beta 1 to 4GlcNAc alpha 2 to 6 sialyltransferase from rat liver. *The Journal of Biological Chemistry,* 257**,** 13845.

Weld, R. J., Plummer, K. M., Carpenter, M. A. & Ridgway, H. J. 2006. Approaches to functional genomics in filamentous fungi. *Cell Research,* 16**,** 31-44.

Wood, E. 1994. Molecular probes: Handbook of fluorescent probes and research chemicals: By RP Haugland. pp 390. Interchim (Molecular Probes Inc, PO Box 22010 Eugene, OR 97402-0414, USA, or 15 rue des Champs, 92600 Asnieres, Paris). 1992–1994. $15. *Biochemical Education,* 22**,** 83-83.

Wösten, H. A., Moukha, S. M., Sietsma, J. H. & Wessels, J. G. 1991. Localization of growth and secretion of proteins in *Aspergillus niger*. *Journal of General Microbiology,* 137**,** 2017-23.

Wright, A. & Morrison, S. L. 1997. Effect of glycosylation on antibody function: implications for genetic engineering. *Trends in Biotechnology,* 15**,** 26-32.

Wu, Z., Miller, E., Agbandje-Mckenna, M. & Samulski, R. J. 2006. Alpha2,3 and alpha2,6 *N*-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. *Journal of Virology,* 80**,** 9093-103.

Xiong, X., Martin, S. R., Haire, L. F., Wharton, S. A., Daniels, R. S., Bennett, M. S., Mccauley, J. W., Collins, P. J., Walker, P. A., Skehel, J. J. & Gamblin, S. J. 2013. Receptor binding by an H7N9 influenza virus from humans. *Nature,* 499**,** 496-99.

Xu, Z., Vo, L. & Macher, B. A. 1996. Structure-function analysis of human alpha1,3-fucosyltransferase. Amino acids involved in acceptor substrate specificity. *Journal of Biological Chemistry,* 271**,** 8818-23.

Yeh, J.-C. & Cummings, R. D. 1996. Absorbance-and light-based solid-phase assays for CMPNeuAc: Galβ1-4GlcNAc-R α-2, 3-sialyltransferase. *Analytical Biochemistry,* 236**,** 126-133.

Yin, J., Li, G., Ren, X. & Herrler, G. 2007. Select what you need: a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. *Journal of Biotechnology,* 127**,** 335-47.

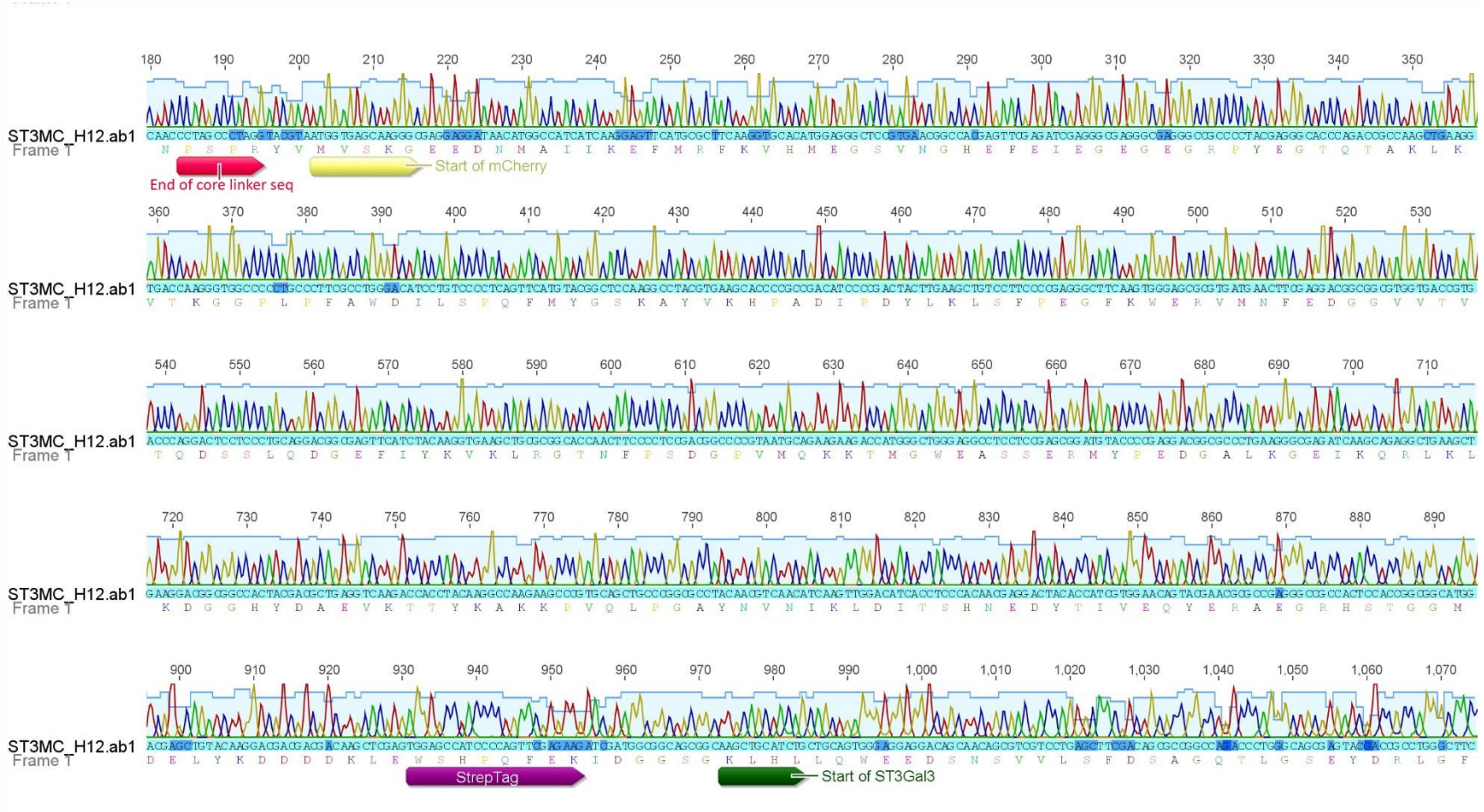
Zhang, J., Zhong, Y., Zhao, X. & Wang, T. 2010. Development of the cellulolytic fungus *Trichoderma reesei* strain with enhanced beta-glucosidase and filter paper activity using strong artificial cellobiohydrolase 1 promoter. *Bioresource Technology,* 101**,** 9815-18.

Zhao, Y.-Y., Takahashi, M., Gu, J.-G., Miyoshi, E., Matsumoto, A., Kitazume, S. & Taniguchi1, N. 2008. Functional roles of *N*-glycans in cell signaling and cell adhesion in cancer. *Cancer Science,* 99**,** 1304-10.

Zheng, K., Bantog, C. & Bayer, R. 2011. The impact of glycosylation on monoclonal antibody conformation and stability. *MAbs,* 3**,** 568-76.

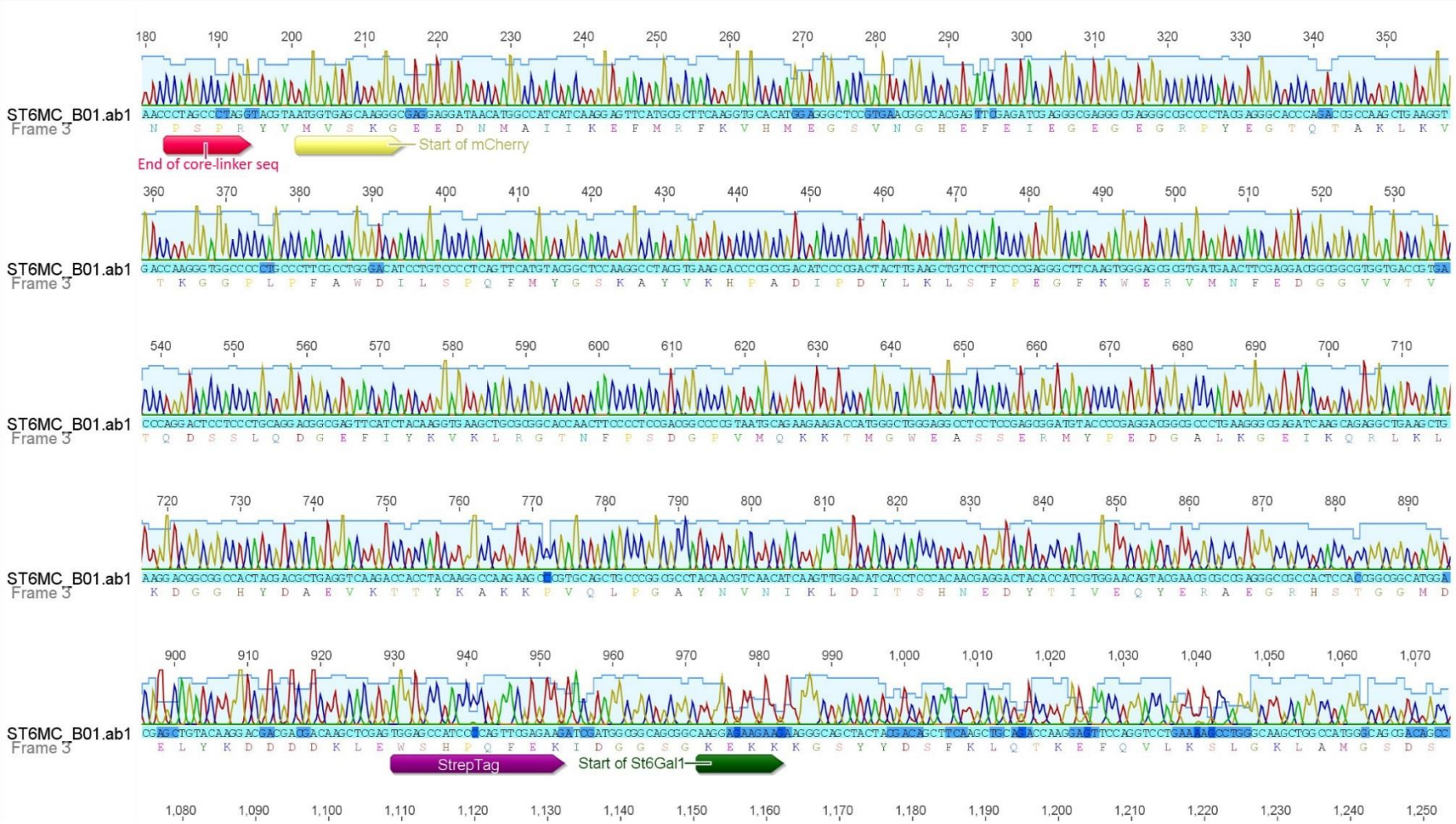
APPENDICES

Appendix (i)



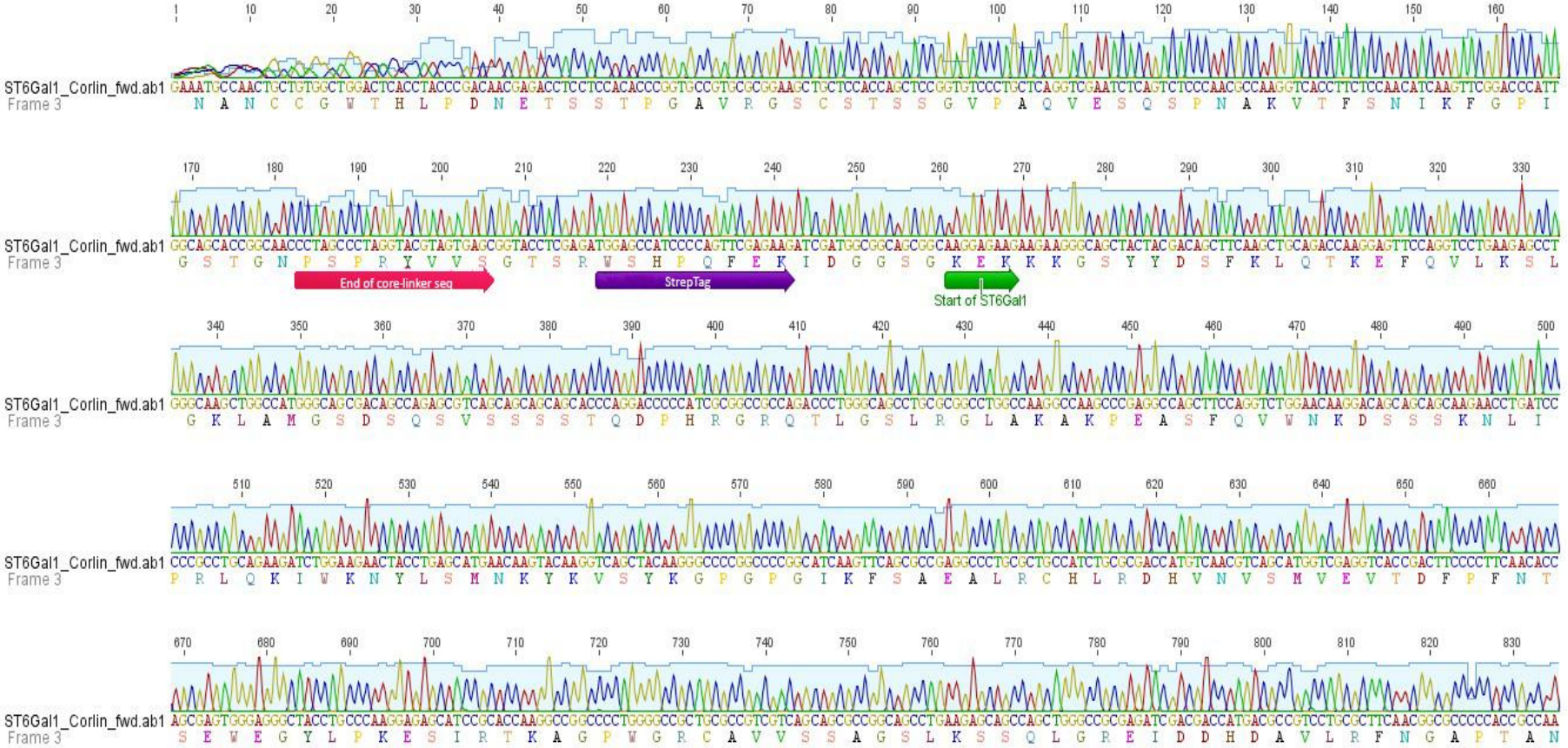
Sequencing result for plasmid containing mCherry and ST3Gal3 cDNAs. ST3Gal3 and mCherry cDNAs are in frame with the core-linker sequence indicating that these cDNAs are in turn in frame with the *cbh1* promoter located upstream of the core-linker sequence.

Appendix (ii)



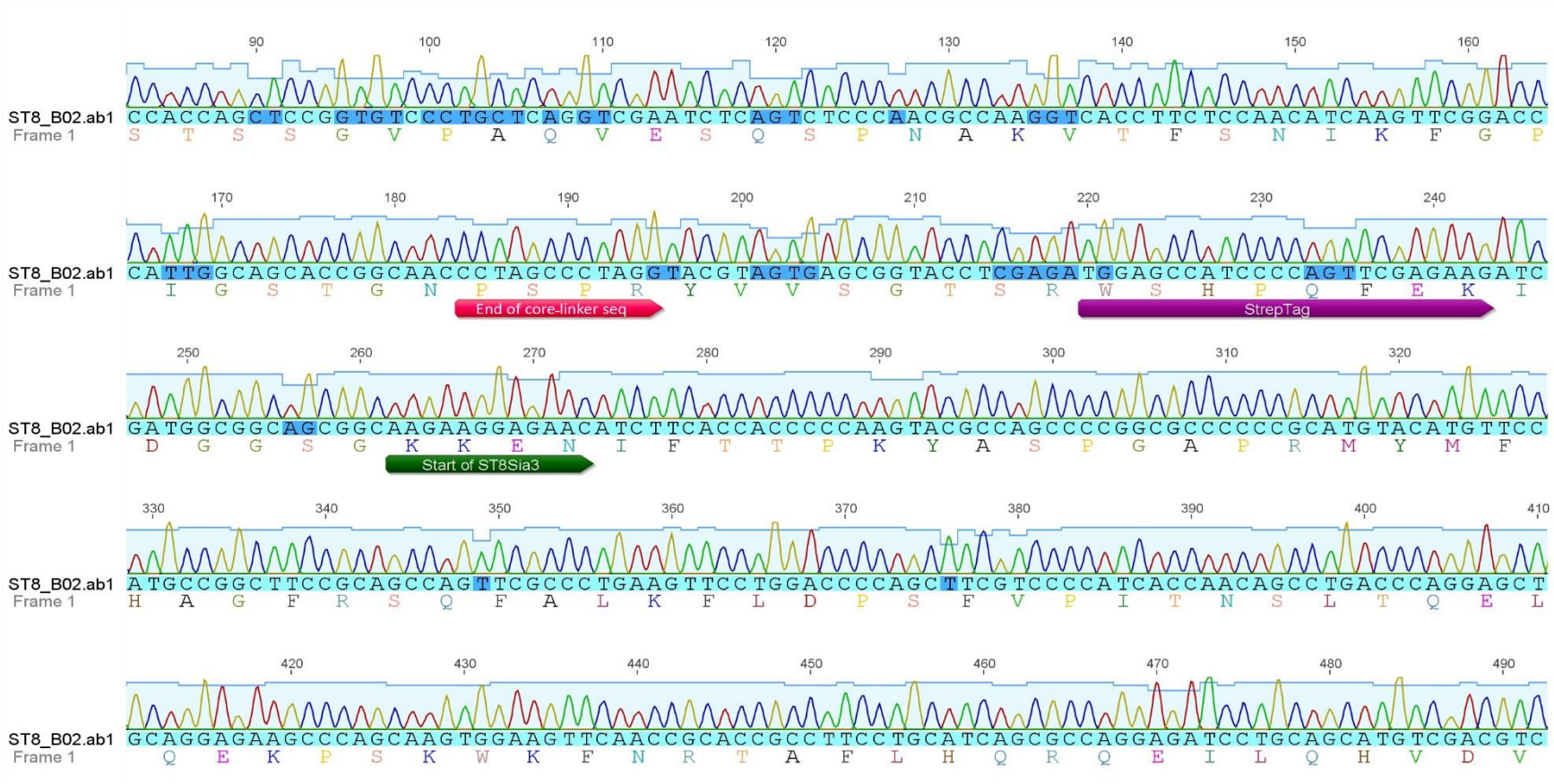
Sequencing result for plasmid containing mCherry and ST6Gal1 cDNAs. ST6Gal1 and mCherry cDNAs are in frame with the core-linker sequence indicating that these cDNAs are in turn in frame with the *cbh1* promoter located upstream of the core-linker sequence.

Appendix (iii)



Sequencing result for plasmid containing ST6Gal1 cDNA. ST6Gal1 cDNA is in frame with the core-linker sequence indicating that the cDNA is in turn in frame with the *cbh1* promoter located upstream of the core-linker sequence.

Appendix (iv)

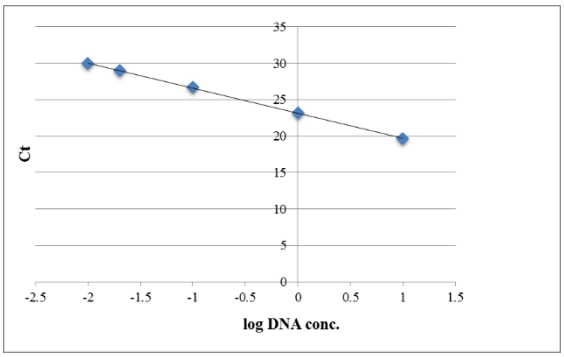


Sequencing result for plasmid containing ST8Sia3 cDNA. ST8Sia3 cDNA is in frame with the core-linker sequence indicating that the cDNA is in turn in frame with the *cbh1* promoter located upstream of the core-linker sequence.

Appendix (v)

DNA conc. (ng /µL)	log DNA conc.	Mean Ct
0.01	-2	29.995
0.02	-1.698970004	28.995
0.1	-1	26.67
1	0	23.195
10	1	19.62
Slope= -3.456344818		
Amplification factor = 1.95		
Primer efficiency = 94.69%		

Standard curve



Determination of efficiency of ST6Gal1 primers using template DNA

Ct values of ST6Gal1		
Samples	Ct	Mean Ct
RC30 (24 h)	- - -	-
RC30 (48 h)	- - -	-
RC30 (96 h)	- - -	-
mS15 (24 h)	18.27 18.63 18.28	18.39333
mS15 (48 h)	22.06 22.06 21.77	21.96333
mS15 (96 h)	31.09 23.61 23.56	26.08667
S16 (24 h)	- - 21.81	21.81
S16 (48 h)	22.95 23 22.98	22.97667
S16 (96 h)	25.72 25.75 25.77	25.74667

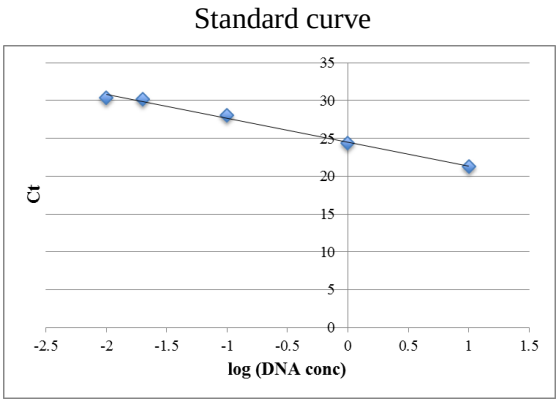
Ct values of Actin		
Samples	Ct	Mean Ct
RC30 (24 h)	22.61 22.25 22.3	22.38666667
RC30 (48 h)	25.29 25.1 25.13	25.17333333
RC30 (96 h)	25.96 25.85 25.95	25.92
mS15 (24 h)	20.58 20.52 20.51	20.53666667
mS15 (48 h)	24.74 25.13 25.13	25
mS15 (96 h)	26.26 26 26	26.08666667
S16 (24 h)	23.19 23.45 23.34	23.32666667
S16 (48 h)	24.97 24.94 24.87	22.38666667
S16 (96 h)	26.19 26.16 25.21	22.38666667

Livak ratio	
mS15	
24 h	-
48 h	1.85746282
96 h	0.22635619
S16	
24 h	-
48 h	0.232182853
96 h	0.034039021

Relative abundance of ST6Gal1 compared to its expression at 24 h, calculated using Livak method (Livak and Schmittgen, 2001).

Appendix (vi)

DNA conc. (ng /µL)	log DNA conc.	Mean Ct
0.01	-2	30.43
0.02	-1.698970004	30.165
0.1	-1	28.015
1	0	24.385
10	1	21.285
Slope= -3.170702114		
Amplification factor = 2.07		
Primer efficiency =106.76%		



Determination of efficiency of actin primers using template DNA

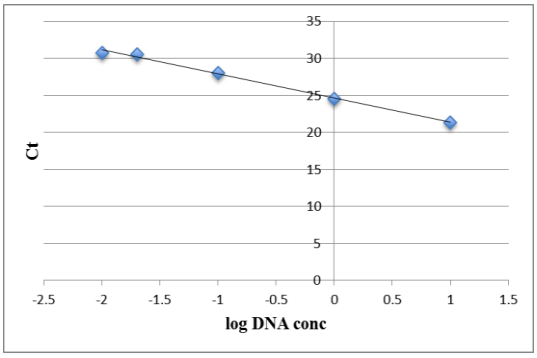
Ct values of Actin		
Samples	Ct	Mean Ct
RC30 (24 h)	22.61 22.25 22.3	22.38666667
RC30 (48 h)	25.29 25.1 25.13	25.17333333
RC30 (96 h)	25.96 25.85 25.95	25.92
mS15 (24 h)	20.58 20.52 20.51	20.53666667
mS15 (48 h)	24.74 25.13 25.13	25
mS15 (96 h)	26.26 26 26	26.08666667
S16 (24 h)	23.19 23.45 23.34	23.32666667
S16 (48 h)	24.97 24.94 24.87	22.38666667
S16 (96 h)	26.19 26.16 25.21	22.38666667

Crossing threshold (Ct) values of actin (housekeeping gene) in the reverse transcribed cDNA samples.

Appendix (vii)

DNA conc. (ng /µL)	log DNA conc.	Mean Ct
0.01	-2	30.795
0.02	-1.698970004	30.555
0.1	-1	28
1	0	24.575
10	1	21.37
Slope= -3.258784284		
Amplification factor = 2.03		
Primer efficiency =102.74%		

Standard curve



Determination of efficiency of BiP primers using template DNA

Ct values of BiP		
Samples	Ct	Mean Ct
RC30 (24 h)	21.41	21.25333
	21.52	
	20.83	
RC30 (48 h)	28.52	28.45
	28.54	
	28.29	
RC30 (96 h)	27.59	28.06667
	27.76	
	28.85	
mS15 (24 h)	18.34	18.54
	18.43	
	18.85	
mS15 (48 h)	25.47	25.46
	25.49	
	25.42	
mS15 (96 h)	27.61	27.55667
	27.33	
	27.73	
S16 (24 h)	21.79	21.80333
	21.84	
	21.78	
S16 (48 h)	27.47	27.32333
	27.26	
	27.24	
S16 (96 h)	28.19	27.98
	27.92	
	27.83	

Ct values of Actin		
Samples	Ct	Mean Ct
RC30 (24 h)	22.61	22.38666667
	22.25	
	22.3	
RC30 (48 h)	25.29	25.17333333
	25.1	
	25.13	
RC30 (96 h)	25.96	25.92
	25.85	
	25.95	
mS15 (24 h)	20.58	20.53666667
	20.52	
	20.51	
mS15 (48 h)	24.74	25
	25.13	
	25.13	
mS15 (96 h)	26.26	26.08666667
	26	
	26	
S16 (24 h)	23.19	23.32666667
	23.45	
	23.34	
S16 (48 h)	24.97	22.38666667
	24.94	
	24.87	
S16 (96 h)	26.19	22.38666667
	26.16	
	25.21	

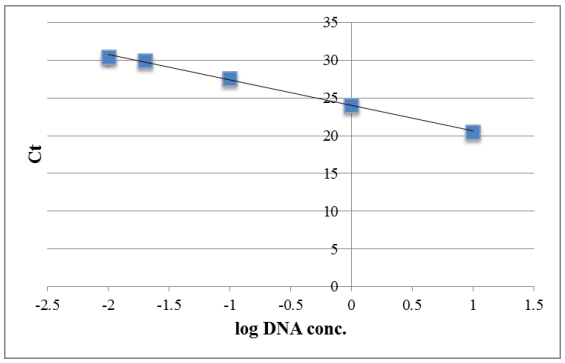
Livak ratio	
mS15	
24 h	1.819237
48 h	7.045327
96 h	1.598442
S16	
24 h	1.310393
48 h	0.316439
96 h	0.091717

Relative abundance of BiP compared to actin, calculated using Livak method (Livak and Schmittgen, 2001).

Appendix (viii)

DNA conc. (ng /µL)	log DNA conc.	Mean Ct
0.01	-2	30.445
0.02	-1.698970004	29.97
0.1	-1	27.58
1	0	24
10	1	20.535
Slope= -3.383439578		
Amplification factor = 1.98		
Primer efficiency = 97.51%		

Standard curve



Determination of efficiency of PDI primers using template DNA

Ct values of PDI		
Samples	Ct	Mean Ct
RC30 (24 h)	21.51 21.63 21.64	21.59333333
RC30 (48 h)	27.32 27.26 32.31	27.29
RC30 (96 h)	25.66 25.64 25.63	25.64333333
mS15 (24 h)	19.23 19.19 32.98	19.21
mS15 (48 h)	25.29 25.14 25.15	25.19333333
mS15 (96 h)	25.76 25.82 25.66	25.74666667
S16 (24 h)	22.55 22.53 22.49	22.52333333
S16 (48 h)	26.97 26.93 27.04	26.98
S16 (96 h)	26.44 26.5 26.48	26.47333333

Ct values of Actin		
Samples	Ct	Mean Ct
RC30 (24 h)	22.61 22.25 22.3	22.38666667
RC30 (48 h)	25.29 25.1 25.13	25.17333333
RC30 (96 h)	25.96 25.85 25.95	25.92
mS15 (24 h)	20.58 20.52 20.51	20.53666667
mS15 (48 h)	24.74 25.13 25.13	25
mS15 (96 h)	26.26 26 26	26.08666667
S16 (24 h)	23.19 23.45 23.34	23.32666667
S16 (48 h)	24.97 24.94 24.87	22.38666667
S16 (96 h)	26.19 26.16 25.21	22.38666667

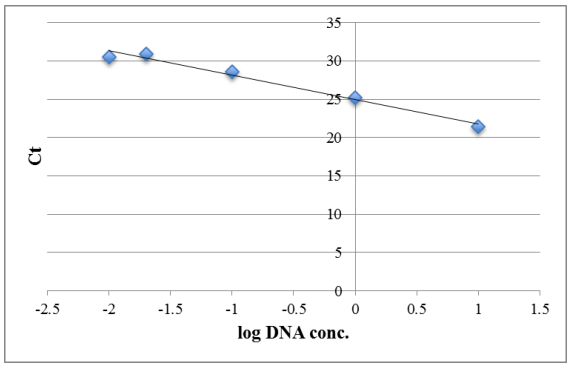
Livak ratio	
mS15	
24 h	1.447269237
48 h	3.792984125
96 h	1.044877153
S16	
24 h	1.00695555
48 h	0.179659028
96 h	0.0485854

Relative abundance of PDI compared to actin, calculated using Livak method (Livak and Schmittgen, 2001).

Appendix (ix)

DNA conc. (ng /µL)	log DNA conc.	Mean Ct
0.01	-2	30.53
0.02	-1.698970004	30.9
0.1	-1	28.55
1	0	25.165
10	1	21.41
Slope= -3.199272998		
Amplification factor = 2.05		
Primer efficiency = 105.39%		

Standard curve



Determination of efficiency of CAL primers using template DNA

Ct values of CAL		
Samples	Ct	Mean Ct
RC30 (24 h)	23.64 22.78 26.92	24.44667
RC30 (48 h)	28.75 28.8 29.29	28.94667
RC30 (96 h)	28.32 28.3 28.28	28.3
mS15 (24 h)	20.29 20.23 20.19	20.23667
mS15 (48 h)	26.53 26.67	26.6
mS15 (96 h)	28.73 28.55 29.02	28.76667
S16 (24 h)	23.53 23.46 23.61	23.53333
S16 (48 h)	28.45 28.19 28.34	28.32667
S16 (96 h)	28.79 28.8 28.81	28.8

Ct values of Actin		
Samples	Ct	Mean Ct
RC30 (24 h)	22.61 22.25 22.3	22.38666667
RC30 (48 h)	25.29 25.1 25.13	25.17333333
RC30 (96 h)	25.96 25.85 25.95	25.92
mS15 (24 h)	20.58 20.52 20.51	20.53666667
mS15 (48 h)	24.74 25.13 25.13	25
mS15 (96 h)	26.26 26 26	26.08666667
S16 (24 h)	23.19 23.45 23.34	23.32666667
S16 (48 h)	24.97 24.94 24.87	22.38666667
S16 (96 h)	26.19 26.16 25.21	22.38666667

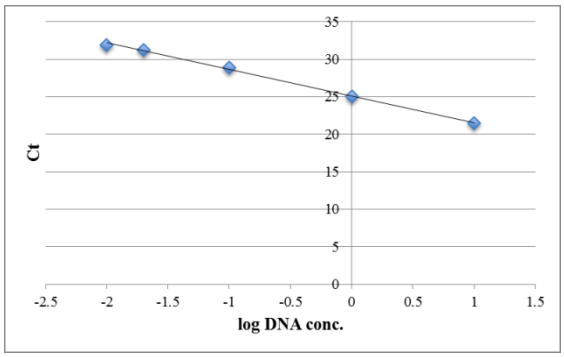
Livak ratio	
mS15	
24 h	5.13370359
48 h	4.510643708
96 h	0.812252396
S16	
24 h	3.613340803
48 h	0.22272468
96 h	0.061072498

Relative abundance of CAL compared to actin, calculated using Livak method (Livak and Schmittgen, 2001).

Appendix (x)

DNA conc. (ng /µL)	log DNA conc.	Mean Ct
0.01	-2	31.925
0.02	-1.698970004	31.255
0.1	-1	28.94
1	0	25.035
10	1	21.5
Slope= -3.547168748		
Amplification factor = 1.91		
Primer efficiency = 91.39%		

Standard curve



Determination of efficiency of UBA primers using template DNA

Ct values of UBA		
Samples	Ct	Mean Ct
RC30 (24 h)	27.94 31.91 27.48	29.11
RC30 (48 h)	35 30.58 35	30.58
RC30 (96 h)	30.28 30.6 30.5	30.46
mS15 (24 h)	25.27 25.15 25.44	25.28667
mS15 (48 h)	30.02 33.2 30.09	31.10333
mS15 (96 h)	30.66 30.66 30.6	30.64
S16 (24 h)	28.53 28.72 28.75	28.66667
S16 (48 h)	30.14 30.29 30.49	30.30667
S16 (96 h)	31.09 30.76 30.88	30.91

Ct values of Actin		
Samples	Ct	Mean Ct
RC30 (24 h)	22.61 22.25 22.3	22.38666667
RC30 (48 h)	25.29 25.1 25.13	25.17333333
RC30 (96 h)	25.96 25.85 25.95	25.92
mS15 (24 h)	20.58 20.52 20.51	20.53666667
mS15 (48 h)	24.74 25.13 25.13	25
mS15 (96 h)	26.26 26 26	26.08666667
S16 (24 h)	23.19 23.45 23.34	23.32666667
S16 (48 h)	24.97 24.94 24.87	22.38666667
S16 (96 h)	26.19 26.16 25.21	22.38666667

Pffaffl ratio	
mS15	
24 h	2.88130
48 h	0.481024
96 h	0.803899
S16	
24 h	1.9141714
48 h	0.1365514
96 h	0.0512994

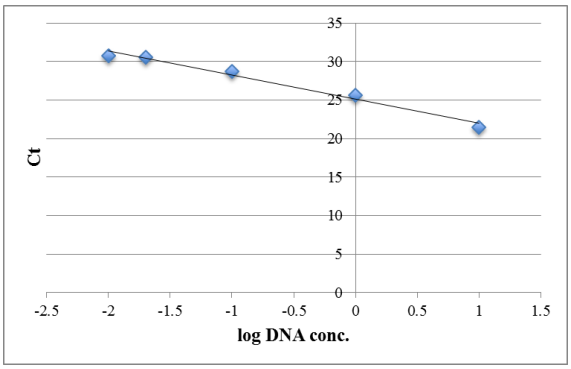
Livak ratio	
mS15	
24 h	3.926743421
48 h	0.616996125
96 h	0.990800613
S16	
24 h	2.60870414
48 h	0.175150459
96 h	0.063226215

Relative abundance of UBA compared to actin, calculated using Livak method (Livak and Schmittgen, 2001).

Appendix (xi)

DNA conc. (ng /µL)	log DNA conc.	Mean Ct
0.01	-2	30.71
0.02	-1.698970004	30.595
0.1	-1	28.68
1	0	25.635
10	1	21.49
Slope= -3.114920224		
Amplification factor = 2.09		
Primer efficiency = 109.43%		

Standard curve



Determination of efficiency of Sec61 primers using template DNA

Ct values of Sec61		
Samples	Ct	Mean Ct
RC30 (24 h)	29.31	25.62
	23.71	
	23.84	
RC30 (48 h)	32.04	29.73667
	28.55	
	28.62	
RC30 (96 h)	-	28.205
	27.98	
	28.43	
mS15 (24 h)	24.75	22.74
	21.73	
	21.74	
mS15 (48 h)	-	27.095
	27.08	
	27.11	
mS15 (96 h)	29.87	29.13333
	28.82	
	28.71	
S16 (24 h)	26.92	27.91333
	31.21	
	25.61	
S16 (48 h)	28.42	28.44
	28.48	
	28.42	
S16 (96 h)	28.54	28.57
	28.56	
	28.61	

Ct values of Actin		
Samples	Ct	Mean Ct
RC30 (24 h)	22.61	22.38666667
	22.25	
	22.3	
RC30 (48 h)	25.29	25.17333333
	25.1	
	25.13	
RC30 (96 h)	25.96	25.92
	25.85	
	25.95	
mS15 (24 h)	20.58	20.53666667
	20.52	
	20.51	
mS15 (48 h)	24.74	25
	25.13	
	25.13	
mS15 (96 h)	26.26	26.08666667
	26	
	26	
S16 (24 h)	23.19	23.32666667
	23.45	
	23.34	
S16 (48 h)	24.97	22.38666667
	24.94	
	24.87	
S16 (96 h)	26.19	22.38666667
	26.16	
	25.21	

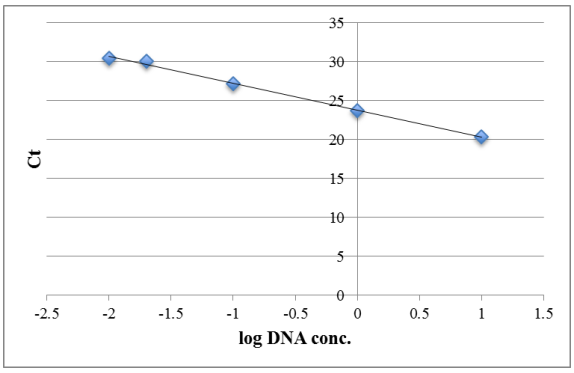
Livak ratio	
mS15	
24 h	2.042024251
48 h	5.534041003
96 h	0.589814556
S16	
24 h	0.391386708
48 h	0.35601255
96 h	0.067063271

Relative abundance of Sec61 compared to actin, calculated using Livak method (Livak and Schmittgen, 2001).

Appendix (xii)

DNA conc. (ng /µL)	log DNA conc.	Mean Ct
0.01	-2	30.425
0.02	-1.698970004	30.045
0.1	-1	27.24
1	0	23.78
10	1	20.36
Slope= -3.452623491		
Amplification factor = 1.95		
Primer efficiency = 94.84%		

Standard curve



Determination of efficiency of RPN8 primers using template DNA

Ct values of RPN8		
Samples	Ct	Mean Ct
RC30 (24 h)	26.92	29.04
	28.33	
	31.87	
RC30 (48 h)	28.99	29.87667
	27.71	
	32.93	
RC30 (96 h)	29.57	30.85
	33.4	
	29.58	
mS15 (24 h)	25.91	26.77333
	28.16	
	26.25	
mS15 (48 h)	31.27	31.00333
	28.81	
	32.93	
mS15 (96 h)	29.1	30.13333
	29.02	
	32.28	
S16 (24 h)	26.71	26.72
	26.76	
	26.69	
S16 (48 h)	28.26	28.33
	28.44	
	28.29	
S16 (96 h)	29.7	29.68333
	29.45	
	29.9	

Ct values of Actin		
Samples	Ct	Mean Ct
RC30 (24 h)	22.61	22.38666667
	22.25	
	22.3	
RC30 (48 h)	25.29	25.17333333
	25.1	
	25.13	
RC30 (96 h)	25.96	25.92
	25.85	
	25.95	
mS15 (24 h)	20.58	20.53666667
	20.52	
	20.51	
mS15 (48 h)	24.74	25
	25.13	
	25.13	
mS15 (96 h)	26.26	26.08666667
	26	
	26	
S16 (24 h)	23.19	23.32666667
	23.45	
	23.34	
S16 (48 h)	24.97	22.38666667
	24.94	
	24.87	
S16 (96 h)	26.19	22.38666667
	26.16	
	25.21	

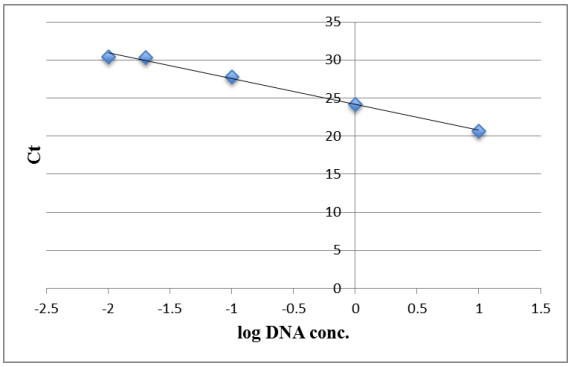
Livak ratio	
mS15	
24 h	1.334839857
48 h	0.4061262
96 h	1.844632391
S16	
24 h	9.579829637
48 h	0.423372658
96 h	0.193893096

Relative abundance of RPN8 compared to actin, calculated using Livak method (Livak and Schmittgen, 2001).

Appendix (xiii)

DNA conc. (ng /µL)	log DNA conc.	Mean Ct
0.01	-2	30.46
0.02	-1.698970004	30.32
0.1	-1	27.755
1	0	24.155
10	1	20.665
Slope= -3.393083355		
Amplification factor = 1.97		
Primer efficiency = 97.12%		

Standard curve



Determination of efficiency of beta 3 primers using template DNA

Ct values of beta 3		
Samples	Ct	Mean Ct
RC30 (24 h)	26 32.14 30.82	29.65333
RC30 (48 h)	28.94 29.17 -	29.055
RC30 (96 h)	29.79 29.75 -	31.51333
mS15 (24 h)	23.71 24.1 23.73	23.84667
mS15 (48 h)	31.68 28.86 28.89	29.81
mS15 (96 h)	30.1 30.14 30.3	30.18
S16 (24 h)	26.83 26.93 26.84	26.86667
S16 (48 h)	28.27 28.21 28.24	28.24
S16 (96 h)	29.79 30.01 29.72	29.84

Ct values of Actin		
Samples	Ct	Mean Ct
RC30 (24 h)	22.61 22.25 22.3	22.38666667
RC30 (48 h)	25.29 25.1 25.13	25.17333333
RC30 (96 h)	25.96 25.85 25.95	25.92
mS15 (24 h)	20.58 20.52 20.51	20.53666667
mS15 (48 h)	24.74 25.13 25.13	25
mS15 (96 h)	26.26 26 26	26.08666667
S16 (24 h)	23.19 23.45 23.34	23.32666667
S16 (48 h)	24.97 24.94 24.87	22.38666667
S16 (96 h)	26.19 26.16 25.21	22.38666667

Livak ratio	
mS15	
24 h	15.5265637
48 h	0.525465032
96 h	2.828427125
S16	
24 h	13.23848996
48 h	0.254958322
96 h	0.27547628

Relative abundance of beta3 compared to actin, calculated using Livak method (Livak and Schmittgen, 2001).