DOCTOR OF PHILOSOPHY

STUDIES ON KINETOPLASTID MITOCHONDRIAL FUCOSYLTRANSFERASES

Paredes Franco, Jose Carlos

Award date:
2023

STUDIES ON KINETOPLASTID MITOCHONDRIAL FUCOSYLTRANSFERASES

JOSE CARLOS PAREDES FRANCO

Division of Biological Chemistry and Drug Discovery

University of Dundee

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

March 2023

Declaration

I declare that I am the author of this thesis; all references cited have been consulted by myself; the work of which this thesis is a record, unless specifically stated, has been done by myself and this work has not been previously accepted for a higher degree.

Jose Carlos Paredes Franco

I confirm that Jose Carlos Paredes Franco has performed the research described in this thesis under my supervision and has fulfilled the conditions of the relevant Ordinance and Regulations of the University of Dundee.

Professor Sir Michael A. J. Ferguson

CONTENTS.

LIST OF FIGURES.

LIST OF TABLES.

LIST OF ABBREVIATIONS.

AFKP80	Bifunctional arabino/fucokinase/pyrophosphorylase
Aq. HF	Aqueous hydrogen fluoride
ATP	Adenosine triphosphate
BHI	Brain heart infusion
BSA	Bovine serum albumin
BSL3	Biosecurity level 3
BZ	Benznidazole
C. fasciculata	*Crithidia fasciculata*
CAZy	Carbohydrate-Active enZYmes database
CD	Chagas disease
CMP	Cytidine monophosphate
CPM	Counts per minute
CSS	Cleavable signal sequence
D-Ara	D-arabinopyranose
D-Gal	D-mannose
D-Glc	D-glucose
D-GlcN	Glucosamine
DHFR-TS	Dihydrofolate reductase-thymidylate synthase
DIG	Digoxigenin
dKO	Double knock-out
D-Man	D-mannose
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid

DPM	Disintegrations per minute
DTU	Discrete Typing Unit
DXD	Aspartate – any amino acid – aspartate
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EFG/TSR	Epidermal growth factor/thrombospondin type repeat sequences
ER	Endoplasmic reticulum
ESI-MS	Electrospray ionization mass spectrometry
EtBr	Ethidium bromide
F6P	Fructose-6-phosphate
FKP40	Bifunctional fucokinase/pyrophosphorylase
FSG	Fish gelatine skin
FT	Flow-through
Fuc1P	Fucose-1-phosphate
FUT	Fucosyltransferase
FUT1	kinetoplastid mitochondrial fucosyltransferase
G418r or NEOr	Neomycin/geneticin resistance gene
GalNAc	N-acetyl-D-galactosamine
GC-MS	Gas chromatography mass spectrometry
gDNA	Genomic DNA
GDP	Guanosine diphosphate
GDP-[^{3}H]Ara	Guanosine 5’-diphospho-β-D-tritiated arabinopyranose
GDP-[^{3}H]Fuc	Guanosine 5’-diphospho-β-L-tritiated fucose
GDP-Ara	GDP-Arabinopyranose
GDP-Fuc	GDP-β-L-Fucose

GDP-Man	GDP-mannose
GFAT	Glucosamine-fructose-6-phosphate aminotransferase
gGAPDH	Glycosomal glyceraldehyde-3-phosphate dehydrogenase
GIPL	Glycosylinositolphospholipids
Glc6P	Glucose-6-phosphate
GlcNAc	N-acetyl-D-glucosamine
GMD	GDP-mannose dehydratase
GMER	GDP-4-hydro-6-deoxy-D-mannose epimerase/reductase
GMP	Guanosine monophosphate
GPI	Glycosylphosphatidylinositol
GPI10	GPI mannosyltransferase 3
GPI12	GlcNAc-phosphatidylinositol (PI) de-N-acetylase
GST	Glutathione-S-transferase
GT	Glycosyltransferase
GT1-3	*Leishmania* spp. D-Glc transporters
GTP	Guanosine triphosphate
HEPES	4-(2-hydrpxyethyl)-1-piperazineethanesulfonic acid
HexNAc	N-acetyl-hexosamine
HIC	Hydrophobic interaction chromatography
(HI-)FBS	(Heat-inactivated-) fetal bovine serum
HK1	Hexokinase
HPLC	High-pressure liquid chromatography
HPTLC	High-performance thin layer chromatography
HYGr	hygromycin resistance gene
IFA	Immunofluorescence analysis

IgG	Immunoglobulin G
IMAC	Immobilized metal affinity chromatography
IP	Immunoprecipitation
IPTG	isopropyl-β-D-thiogalactopyranoside
KPSI	Thousand pounds per square inch
L. major	*Leishmania major*
Lac	Lactose
LacNAc	N-acetyllactosamine
LAG	Lipoarabinogalactan
LC-MS/MS	Liquid chromatography with tandem mass spectrometry
L-Fuc	L-fucopyranose
LNB	Lacto-N-biose
LNB-OMe	Lacto-N-biose-β-O-methyl
LNT	Lacto-N-tetraose
LPG	Lipophosphoglycan
LPG1	UDP-Gal*f* : aMan β-galactofuranosyltransferase
LPG2	Golgi GDP-Man/GDP-Ara/GDP-Fuc nucleotide sugar transporter
Man1P	Mannose-1-phosphate
Man6P	Mannose-6-phosphate
MBP	Maltose-binding protein
MOPS	3-(N-morpholino)propanesulfonic acid
MPGT	Mannose phosphate guanyl transferase
MTS	Mitochondrial targeting sequence
MW	Molecular weight
Ni-NTA	Nickel-nitriloacetic acid

NTDs	Neglected tropical diseases
NX	Nifurtimox
(OE-)MP	Overexpressing drug-resistant mixed population
OGA	O-GlcNAcase
OGT	O-GlcNAc transferase
ORF	Open reading frame
PAS	Periodic acid-Schiff
PCR	Polymerase chain reaction
^{1}H-NMR	Proton nuclear magnetic resonance
PFA	Paraformaldehyde
PG1	Phospho-glucose isomerase
PI-PLC	Phosphatidylinositol-specific phospholipase C
PMI	Phosphomannose isomerase
PMM	Phosphomannomutase
PPase	Inorganic pyrophosphatase
PP_i	inorganic pyrophosphate
PV	Parasitophorous vacuole
R.P.	Radiolysis product(s)
RE	Restriction enzymes
RNA	Ribonucleic acid
RPM	Rotations per minute
RP-SPE	Reverse-phase solid phase extraction
rRNA	Ribosomal RNA
rTcFUT1cat	Recombinant MTS-less T. cruzi FUT1
sKO	Single knock-out

SNP	Single nucleotide polymorphism
SOC	Super optimal broth
SSC	Sodium chloride, sodium citrate buffer
Ta	Annealing temperature
TCEP	Tris(2-carboxyethyl)phosphine
TcNTR-1	*T. cruzi* mitochondrial type 1 nitroreductase
TFA	Trifluoroacetic acid
[^{3}H]Ara	Tritium-labelled D-arabinopyranose
[^{3}H]Fuc	Tritium-labelled L-fucose
[^{3}H]Fuc-1P	Tritium-labelled Fuc-1-PO_4
[^{3}H]Man	Tritium-labelled D-mannose
THT1 & THT2	*T. brucei* D-Glc transporters
TLC	Thin layer chromatography
Tm	Melting temperature
TMS	Trimethylsilyl
TS	TEV protease cleavage site
UDP	Uridine diphosphate
5'/3' UTR	Upstream/downstream untranslated region
WAX-SPE	Weak anion exchanger resin as stationary phase
WHO	World Health Organisation
WT	Wild-type
αFuc	α-L-fucose
αGlc	α-D-glucose
βGal	β-D-galactose
βGal*f*	β-D-galactofuranose

ACKNOWLEDGEMENTS.

My deepest gratitude to my supervisor, Prof. Sir Michael A. J. Ferguson, or Mike, for all those who have had the opportunity of interacting with him. I consider myself immensely lucky of have been the last of a long list of PhD students under his incomparable guidance, continuous encouragement and rigorous yet empathetic supervision. His many lessons in so many aspects of science, and life, will forever be treasured, and, I believe, will benefit my research attitude and critical scientific approach during my research career. Thank you, Mike.

My special thanks to Dr. M. Lucia Sampaio Güther and Dr. Samuel M. Duncan, without whose supervision in the lab I would have not been able to learn about, perform, and optimize many experiments and analyses (many more than the ones that ended up in this thesis) that were, in greater or lesser degree, critical for the realization of this work. I also thank Dr. Rupa Nagar for her lessons in GC-MS and carbohydrate analysis, which I am sure will be a good basis for my future development in different aspects of chemical research. Thanks to Dr Zhe Ji for general help and fruitful discussions, and to Dr. Michele Tinti for his advice in the data analysis. Also, thanks to the former members of the group Dr. Elda Iljazi and Dr. Giulia Bandini, whose previous work has been in many aspects the basis of mine, and for their advice and helpful discussions.

Thanks to the Protein Production Team: Sharon Shepherd, Dr. Thomas Eadsforth, Greg Stewart, and Dr. Alain-Pierre Petit. They have suffered me the most during my (many) attempts of protein expression and purification, and yet, they were always helpful and patient with me, despite a very busy schedule.

I would like to thank DH, MoA, and MCF group members, who have seen my scientific development since I joined the BCDD/WCAIR division. This place has represented to me a wonderful working environment, where everybody was always willing to help, whether it was with a scientific discussion, an advice in one or another technique, or just a chat to relax for a bit, before going back to work.

From my past in the Autonomous University of Madrid, I deeply thank Prof. Pedro Manuel Bonay Miarons. He made me realize that I wanted to be a biochemist and was the first one to show me the wonders of Parasitology and Glycobiology. His supervision was, in many ways, what created the fundaments of my beliefs on what a good praxis in science is, and his personal appreciation and dedication of me made me

see him almost as a paternal figure during hard times in the past. After almost 5 years under his supervision, I came to Dundee much more prepared, scientifically and personally, thanks to him. I also thank Dr. Marcello Salvatore Rossi, who was an incredible support during my time in the Centre of Molecular Biology Severo Ochoa (CBMSO) and keeps being so. I thank Prof. María José Hazen de San Juan and Prof. Vicente Mazimpaka Nibarere, who were the very first academics to see "something" in me and encourage me to follow the scientific career.

I thank Santiago and Celia, who after so many years are still by my side, giving me all their support and receiving me with open arms whenever I go back to Madrid. I also thank all the friends I have made in Dundee and with whom I have shared so many happy moments with: Carmen, Clara, Cristian, Mirtha, and many others. This journey would have been much more difficult without them.

My deepest respect, filial love, and gratitude, to my mother, who faced so many difficulties in her life just so I could have a good future. It is due to her sacrifice, and hers only, that I am here. She will always be in my mind, telling me that despite our humble origins, with effort and dedication everything can be possible. Thank you also to my sister Itziar, whose warmth and joyful spirit has always brought me so much happiness.

Finally, my love and gratitude to the woman who has been by my side during all this challenging time. (Soon to be Dr.) Markéta Novotná and I have built a relationship based in love, respect, trust, and shared interests in science and art. Our little home, with our cat Daphne, and her forever remembered sister Calypso, has been my safe place, my resting point, where I could relax, be and express myself, and only receive love and support, as I hope she also did.

"What I have achieved by industry and practice, anyone else with tolerable natural gift and ability can also achieve." - J. S. Bach.

Magna opera Domini exquisita in omnes voluntates eius.

Grandes son las obras del Señor, dignas de estudio para los que las aman.

Great are the works of the Lord, they are pondered by all who delight in them.

Psalmus 111(110):2

SUMMARY.

Protozoan parasites from the Trypanosomatidae family cause terrible diseases that affect human and animal health worldwide. The causative agents of leishmaniases (*Leishmania* spp.), Chagas disease (*Trypanosoma cruzi)* and African trypanosomiases (*Trypanosoma brucei* subsp.) rely heavily on glycoconjugates for their survival and infectivity. Thus, the study of the glycobiology of these parasites represents an opportunity for the identification of novel and much needed therapeutic targets. Glycosylation processes in eukaryotic cells are carried out by glycosyltransferases located mostly in the secretory pathway, and some in the cytoplasm. Consequently, most of the synthesized glycans, attached to proteins and lipids, are present on their cell surfaces, or in their endosome/lysosome systems, or are secreted.

The discovery of mitochondrial glycosyltransferases in trypanosomatids has challenged this established consensus. Previous work in our group and by collaborators showed that the protein product of an essential fucosyltransferase gene (*FUT1*) is located inside the mitochondrion of the parasites *Leishmania major* and *Trypanosoma brucei*. Although *in vitro* fucosyltransferase activity has been demonstrated for recombinant *LmjFUT1* and *TbFUT1*, to date, the endogenous substrate(s) of these fucosyltransferases have not been identified. Furthermore, prior to this thesis, nothing was known about the orthologous fucosyltransferase (TcFUT1) in *Trypanosoma cruzi*. In this study, I detail approaches to obtain an active form of TcFUT1, both by heterologous (bacteria and human cell) expression and homologous overexpression, in order to characterize its activity and substrate specificity using a panel of synthetic glycan acceptors. This defined TcFUT1 as a GDP-Fuc : βGal α1-2 fucosyltransferase. In parallel, I exploited the metabolism of the non-pathogenic trypanosomatid *Crithidia fasciculata* to synthesize a critical component for glycosylation activity assays (GDP-[^{3}H]Fuc). Furthermore, initial steps were taken to characterize a high-molecular weight endogenous fucosylated substrate by biosynthetic labelling of *C. fasciculata* with [^{3}H]-labelled fucose. To assist this approach, we have also generated a genetically modified *C. fasciculata* cell line that should be deficient in the synthesis of lipoarabinogalactan (LAG). Additionally, based on immunolocalization assays and comparison with similar reports on in *T. cruzi*, we propose the mitochondrial localization of TcFUT1, at least in the epimastigote form.

1. INTRODUCTION.

1.1 Phylogenetics: position of kinetoplastids in the Eukaryota and the position of *Trypanosoma cruzi* and *Crithidia fasciculata* in the kinetoplastids.

From the two domains of cellular life (Bacteria and Archaea/Eukaryota), animals, plants and fungi are the macroscopic representatives of the eukaryotic domain. Nevertheless, the microscopic eukaryotes, or protists, represent the majority of the major lineages of the eukaryotic tree of life (Keeling and Leander, 2009). Analyses of multiple genes by next-generation sequencing allowed for precise phylogenetic categorizations that currently classify the eukaryotes in five "super groups": Amoebozoa, Opisthokonta, Excavata, SAR (Stramenopiles/Alveolata/Rhizaria) and Archaeplastida (Eme *et al.*, 2014). Animal protists, or protozoa, have developed free-living, symbiotic, and parasitic lifestyles, the latter ones belonging to three of these "super groups". Thus, Excavata includes kinetoplastids (*T. brucei*, *T. cruzi*, *Leishmania spp.*, *C. fasciculata,* etc.), trichomonads (*Trichomonas* spp.) and diplomonads (*Giardia* spp.); Alveolata includes apicomplexans (*Plasmodium spp.*, *Toxoplasma gondii*, *Cryptosporidium parvum*, etc.) and Amoebozoa includes the parasitic amoebas (*Entamoeba* spp.). Indeed, parasitic protozoa are an extremely diverse group that collectively cause various pathologies in human and animals worldwide (Bandini and Ferguson, 2022) (Fig. 1.1).

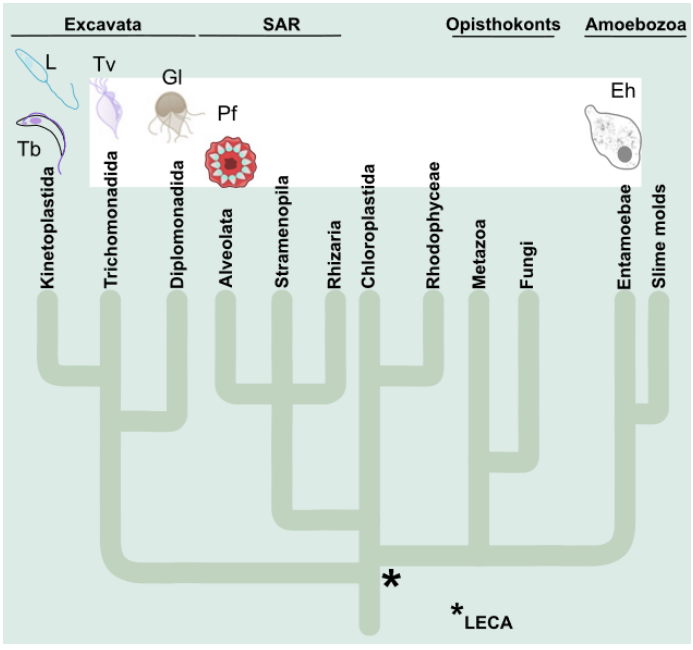


Figure 1.1: Evolutionary diversity of eukaryotic parasites. LECA: last eukaryotic common ancestor; L: *Leishmania* spp.; Tb: *T. brucei*; Gl: *G. lamblia*; Tv: *T. vaginalis*; Pf: *P. falciparum*; Eh: *E. histolytica*. Image taken from (Bandini and Ferguson, 2022).

From these, Kinetoplastea represents one of the most ancestral protozoan classes, since they diverged from a free-living ancestor several hundred million years ago (Yazaki *et al.*, 2017). Kinetoplastids can be separated into two major groups, a monophyletic one composed of only obligate parasites, the Trypanosomatida, and a second one composed of four clades of free-living biflagellates (Eubodonida, Parabodonida, Neobodonida and Prokinetoplastida) (Harmer *et al.*, 2018). The kinetoplastids acquire their name due to a particular subcellular structure known as kinetoplast, a network of circular DNAs that contains many copies of the mitochondrial genome, inside their single, large mitochondrion. Whereas in the free-living members of this group the mitochondrial DNA is composed of numerous free, non-catenated relaxed or supercoiled DNA circles, the Trypanosomatida has thousands of relaxed circles, mutually interlocked into a single giant network composed of DNA maxicircles and minicircles that together with proteins are packaged into a single compact disk (Faktorová *et al.*, 2016).

Multiple advances in different areas of evolutionary biology have provided data suggesting that at some point after diverging from the archaeal lineage, a protoeukaryote acquired different features of typical eukaryote cells such as a tubulin-based flagellum, an endomembrane system, a cytoskeleton, and particularly, a mitochondrion, due to an endosymbiosis event with an ancestral α-proteobacterium (Margulis *et al.*, 2000). There is strong molecular evidence that this Last Eukaryotic Common Ancestor (LECA) (Makarova *et al.*, 2005) diverged into the trypanosomatids very early (Burki *et al.*, 2020; Padilla-Mejia *et al.*, 2021)

The trypanosomatids infect a wide variety of vertebrates, invertebrates, and plants. Although they are predominantly monoxenous parasites (i.e., restricted to a single, mainly invertebrate host), such as the insect parasite *Crithidia*, the trypanosomatids are better known for their dixenous members (i.e., those with an invertebrate and vertebrate host) such as *Leishmania* and *Trypanosoma*, which are causative agents of some of the most important neglected tropical diseases (NTDs) in humans, including leishmaniasis, sleeping sickness and Chagas disease (Kaufer *et al.*, 2017). Many attempts to systematize this family have been made in the last years (Kaufer *et al.*, 2020), the most recent one being based on the coding region of the kinetoplast DNA from the maxicircles of both monoxenous and dixenous trypanosomatids (Kaufer *et al.*, 2019) (Fig. 1.2). A later analysis, only focussed on the *Trypanosoma* genus, was based on both the 18S rRNA sequences and their secondary structure (Borges *et al.*, 2021).

The two parasites which this thesis is focused on, *Crithidia fasciculata* and *Trypanosoma cruzi* (Sylvio strain), are indicated in (Fig. 1.2), where it is also shown that monoxenous trypanosomatids (highlighted in grey) evolved into the dixenous members of the family several times throughout evolution, leading to the emergence of *Trypanosoma*, *Phytomonas*, *Leishmania*, *Porcisia* and *Endotrypanum* (Kaufer *et al.*, 2019). In addition, it is shown that the earliest diverging lineage is the genus *Paratrypanosoma*, represented by only one known species: *P. confusum* (Flegontov *et al.*, 2013).

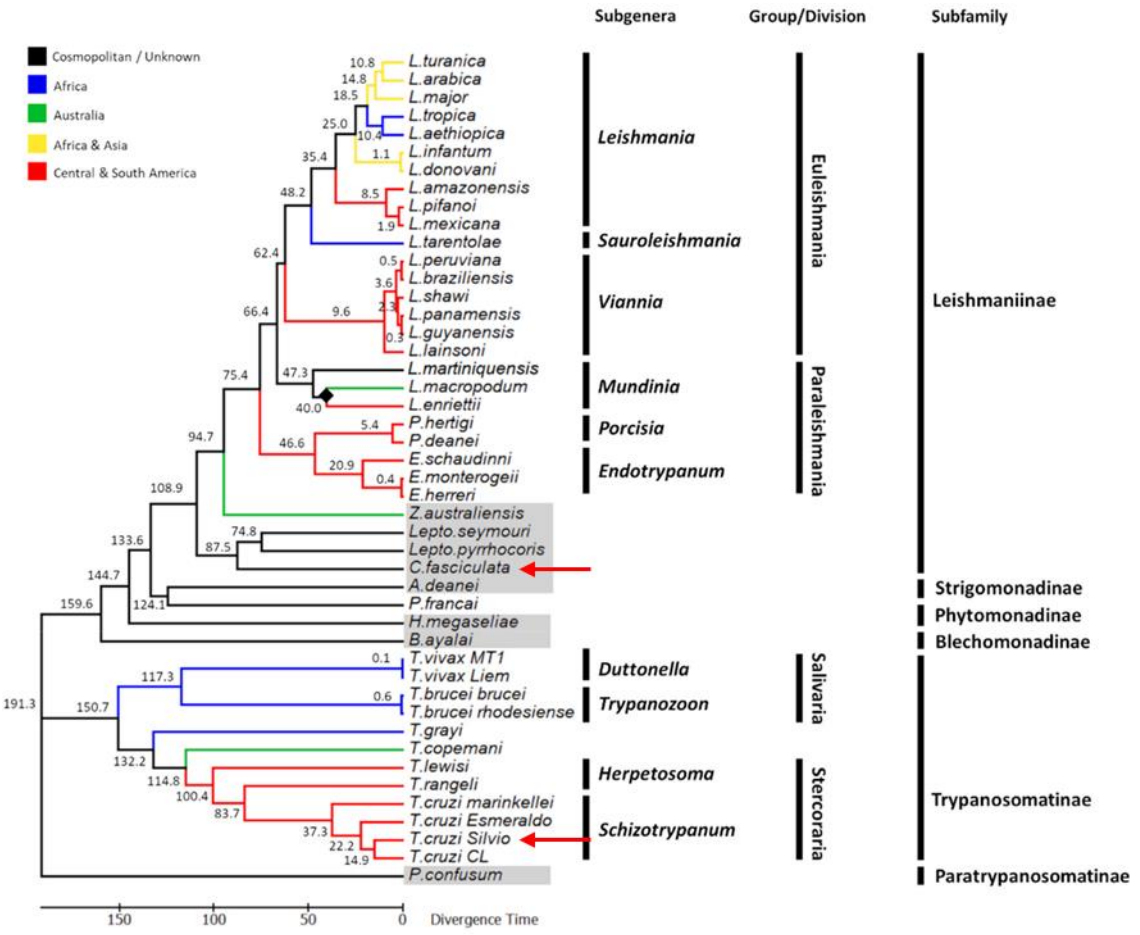


Figure 1.2: Phylogenetic tree of the trypanosomatid family based on the coding region from their mitochondrial maxicircles. The divergence time (in million years) was calculated using as the calibration date, the separation of *Leishmania macropodum* and *Leishmania enriettii* by approximately 40 million years (solid black diamond). Monoxenous trypanosomatid species are highlighted in grey. The parasites used in this thesis are indicated by red arrows. Image adapted from (Kaufer *et al.*, 2019).

The transmission of an insect-living trypanosomatid into a warm-blooded host has likely occurred many times during the evolution of this family (Flegontov *et al.*, 2013; Lukeš *et al.*, 2014; Hamilton and Stevens, 2017). Once this happened, it was the members of the *Trypanosoma* genus that most successfully radiated and adapted to many vertebrate species, reflected in their current evolutionary success in being able to prosper in

essentially all vertebrate animals, from fishes to mammals (Lemos *et al.*, 2015; Hamilton and Stevens, 2017; Fermino *et al.*, 2019).

1.2 *T. cruzi* and Chagas disease.

Chagas disease (CD) is a neglected tropical disease (NTD) that has been endemic of the Americas for millions of years (Hamilton *et al.*, 2012) and that results from the infection with the protozoan parasite *T. cruzi*. Discovered over 100 years ago by the Brazilian physician Carlos Chagas, the causative agent, the clinical aspects of the disease and the insect vector have long been known (Pérez-Molina and Molina, 2018). Human transmission occurs when one of >140 competent species of *Triatominae* insect vector feed on human blood leaving parasite-contaminated vector faeces on the skin. The parasites are then rubbed into the bite wound or a nearby mucous membrane, which allows inoculation (Marcus *et al.*, 2021). Oral infection is also possible from contamination with vector faeces of food or drink, such as fruit juices (Shikanai-Yasuda and Carvalho, 2012). About 6 – 7 million people are currently infected worldwide, mostly in Latin America, which makes CD the single greatest burden of any parasitic infection in this part of the world (Chagas disease, WHO, 2022). The high number of people who remain undiagnosed and/or untreated, combined with the areas with remaining active transmission, lead to an estimation of 75 million people at risk of infection (World Health Organization, 2017). Migration from endemic countries has increased the number of cases detected in Canada, the United States of America (USA), many European countries, Japan and Australia (Lidani *et al.*, 2019) (Fig. 1.3).

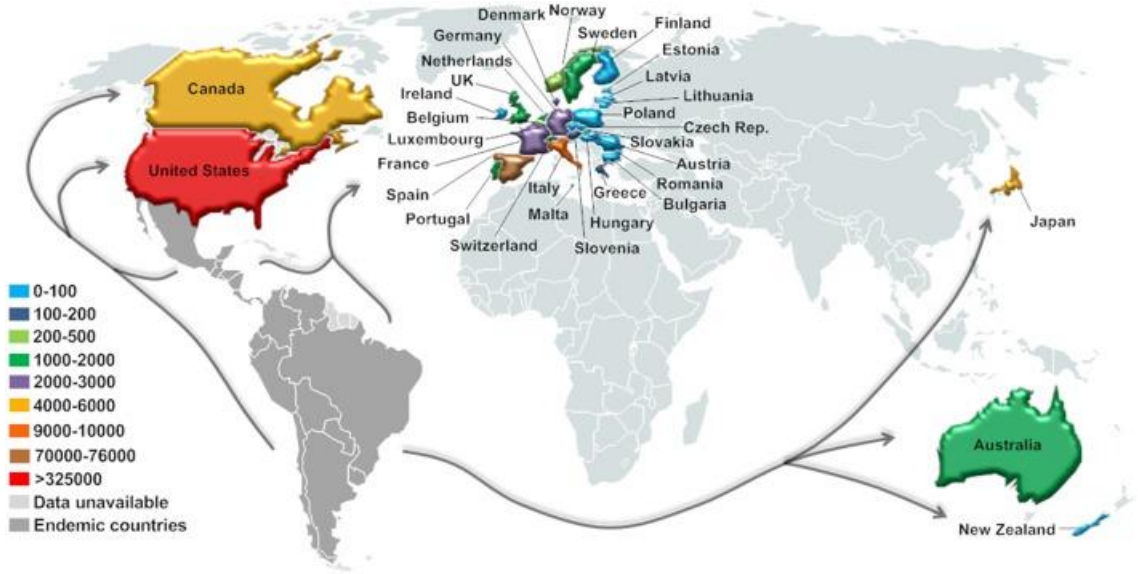


Figure 1.3: Map estimating the number of migrants with *T. cruzi* infection from the endemic countries in Latin America to non-endemic countries, leading to the global spread of infections. The USA has a higher burden of the disease than any other non-endemic country, possibly due to migration in addition to recently identified autochthonous transmission in the southern states. Image taken from (Lidani *et al.*, 2019).

The current estimate of the global economic cost of CD is $7 billion per year (Lee *et al.*, 2013). Overall, with only a few exceptions (Gurevitz *et al.*, 2013), the global incidence of the disease is declining due to large-scale control programmes (Abras *et al.*, 2022). However, with the number of cases outside of South America steadily increasing, or rather with increased screening and realisation of the true burden of the infection (Hernandez *et al.*, 2016; Magri *et al.*, 2021), there is an increasing risk of new epidemic emergence (Schmunis and Yadon, 2010; Lidani *et al.*, 2019).

T. cruzi is a dixenous parasite with a complex life cycle (Fig. 1.4), undergoing several morphological changes and having intracellular and extracellular forms.

Inoculation into a new mammalian host occurs on contact with the faeces of the infected vector, as described above, in which the metacyclic trypomastigotes are present. These infecting life-cycle stage cells are slender, flagellated, highly motile, and capable of infecting a wide range of nucleated cells. Entry into host cells has been extensively investigated and reviewed (Epting *et al.*, 2010; Fernandes and Andrews, 2012; Ferri and Edreira, 2021). Briefly, contact with the host glycocalyx reduces the motility of trypomastigotes, which secrete pore-forming toxins that affect the host plasma membrane integrity and cause an influx of extracellular Ca^{2+}. This triggers a repair mechanism through which cells mobilise lysosomes from the peri-nuclear region to fuse with the plasma membrane and resolve the detected breach. It is hypothesized that *T. cruzi* "hijacks" the endocytic process by entering the inserted lysosomal membrane. Once inside this parasitophorous vacuole (PV), escape is triggered by increased H^+ concentration. Still inside of the PV, parasites start differentiating into the non-motile intracellular life-cycle stage known as amastigote, and later escape into the cytosol. Amastigogenesis occurs via cell division, during which the flagellum is released and two small spherical daughter cells are produced (Kurup and Tarleton, 2014). After several cycles of binary division, the amastigotes transform back to long, slender, flagellated forms known as bloodstream trypomastigotes (Dvorak and Hyde, 1973), which are morphologically similar to metacyclic trypomastigotes but express different surface molecules (Yoshida and Cortez, 2008). Lysis of the host cell membrane releases these non-replicative, motile and complement resistant parasites that disseminate in the lymph and blood invading further host cells. In addition, bloodstream trypomastigotes, and some amastigotes present after the rupture of nearby cells, can be uptaken by the triatomine bug from the skin or in a blood meal (Zuma *et al.*, 2021).

Inside the vector, parasites differentiate to epimastigotes through their journey in the digestive system of the insect. Epimastigotes are the largest forms of the parasite, long, slender, and flagellated. They are replicative and divide by binary fission. In the weeks following a bloodmeal, the number of epimastigotes in the small intestine gradually increases. In the rectum, these life forms will turn into the non-replicative metacyclic form, ready to exit the bug and infect a new host (Kollien and Schaub, 2000; Nogueira *et al.*, 2015).

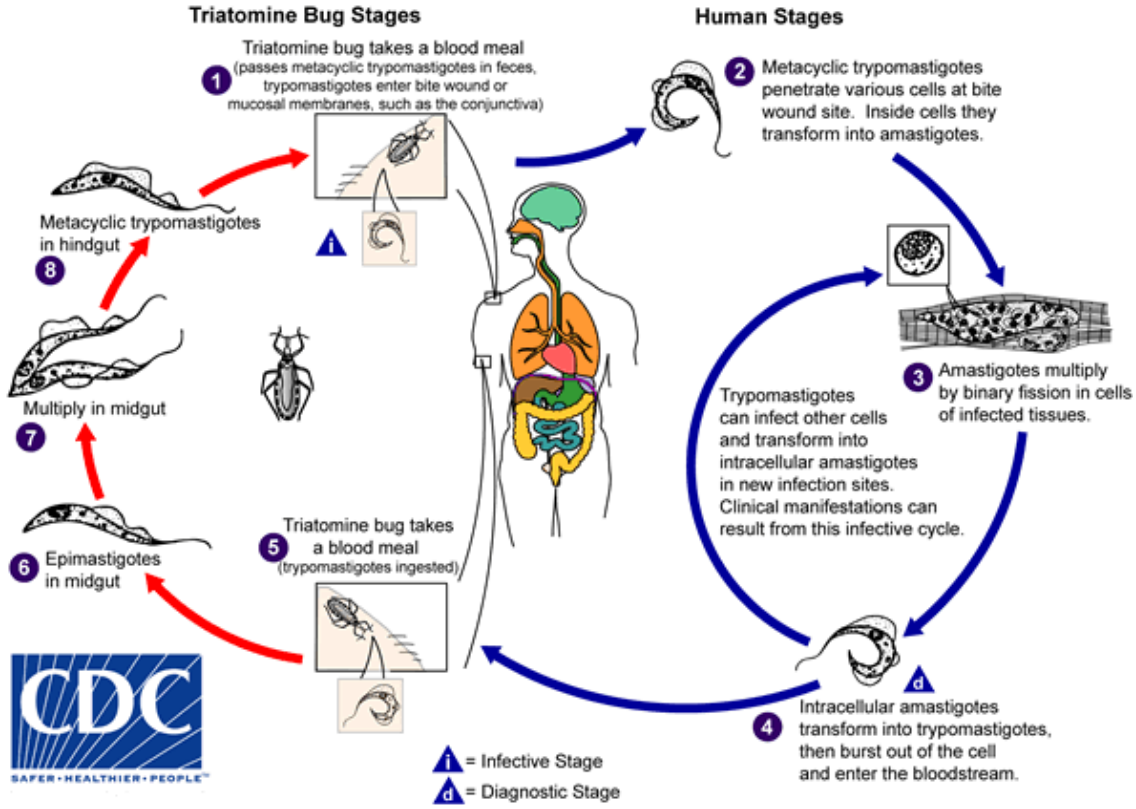


Figure 1.4: Life cycle of *T. cruzi* in the vertebrate and invertebrate host. Image taken from (CDC-Centers for Disease Control and Prevention, 2022)

T. cruzi comprises a single species, however there are great differences between strains, with genetic diversity, distinct geographical distribution, and clinical manifestations, all of it contributing to the complexity of the epidemiology of CD. The variation in the amount of both nuclear and kinetoplast DNA among *T. cruzi* strains is about 40%, which is a significant difference for the same species (Zingales *et al.*, 2012; Zingales, 2018).

After different attempts to classify these parasite strains (Herreros-Cabello *et al.*, 2020), loci genotyping analyses revealed six distinct Discrete Typing Units (DTUs) (Fig. 1.5) identifiable by common molecular markers (Zingales *et al.*, 2009). These are clusters of *T. cruzi* strains that are genetically more similar to one another than they are to parasites

belonging to any other DTU, allowing for the differentiation between strains despite the intraspecific diversity.

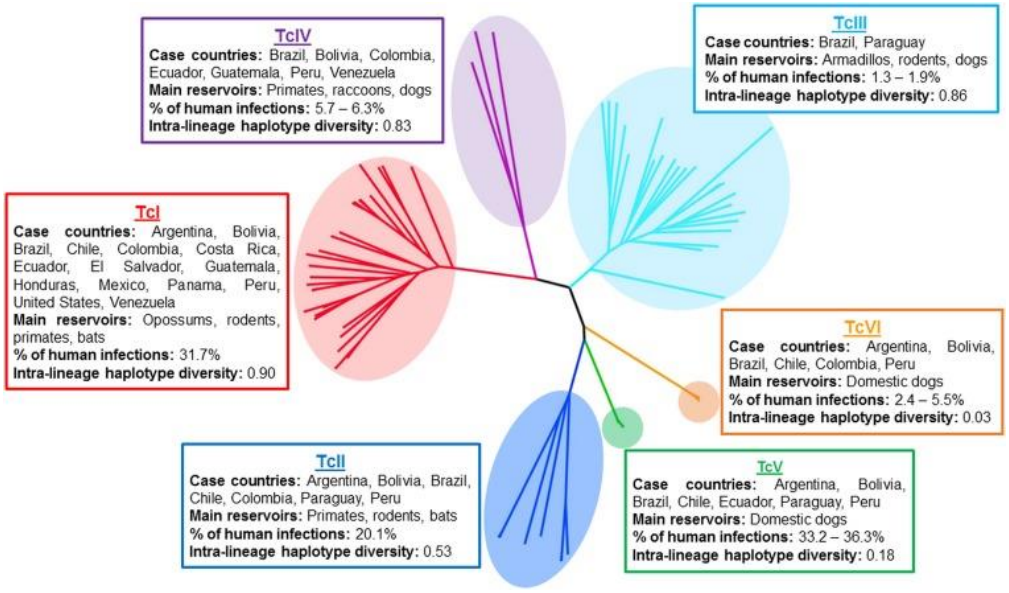


Figure 1.5: Key characteristic of *T. cruzi* DTUs. Tree represents the inter- and intra-lineage diversity and was constructed using multi-locus microsatellite genotype data from parasites collected from diverse geographical locations. The intra-lineage haplotype diversity represents the likelihood that two randomly selected haplotypes from the same lineage will be different Image taken from (Francisco *et al.*, 2017)

Indeed, these differences in parasite genetics are reflected in the geographical distribution and host specificity of strains (Zingales *et al.*, 2012). Some DTUs are commonly sampled from the sylvatic cycle of transmission, with human infection being uncommon (TcIII and TcIV). Other DTUs are more commonly sampled from human infections, with infection of wild animals being less common (TcII, TcV and TcVI). TcI is the most widely dispersed of all the DTUs and is sampled from wild and domestic cycles. The strain of *T. cruzi* used in this thesis, Sylvio X10/7A, belongs to TcI DTU.

The clinical aspects of CD have been extensively reviewed (Perez-Molina and Molina, 2018; Lidani *et al.*, 2019). This disease is categorised in three distinct stages during its progression (Fig. 1.6). The symptomatic acute stage, an asymptomatic stage named indeterminate stage, and a symptomatic chronic stage. Acute infection occurs upon infection and may last for several weeks. Parasite entry into the eye can lead to monocytic infiltration and swelling, referred to as Romaña' sign or "chagoma" if the swelling has taken place at the insect bite site. The acute stage is characterised by uncontrolled proliferation of parasites, during which these are detectable in the blood by microscopic examination. Non-specific symptoms can appear during this stage, such as fever, lethargy, diarrhoea and drowsiness (Carter *et al.*, 2012). The acute stage is not usually life-threatening and is self-resolving due to the generation of a robust (but non-sterilizing) immune response. Acute phase symptoms, when present, resolve over several

weeks/months to be replaced by the intermediate stage. However, a small percentage of cases in the acute stage lead to fatality due to meningoencephalitis or acute heart failure (Medeiros *et al.*, 2008).

The indeterminate stage is characterised by the absence of clinically recognisable pathology, and the number of parasites is generally reduced to levels not detectable in the blood by methods such as PCR. In 30-40% of these cases, clinical signs, mostly a spectrum of cardiac complications referred as chronic chagasic cardiomyopathies, develop at a rate of ~2% per year (Sabino *et al.*, 2013; Bonney and Engman, 2015). These chronic symptoms can include arrythmias and infarctions, dilated cardiomyopathy, and congestive heart failure. The current consensus is that they are the result of a chronic inflammatory damage, although it is still not resolved how this relates to infection and possible autoimmune processes (De Bona *et al.*, 2018; Kölliker *et al.*, 2018; Nunes *et al.*, 2018).

In ~10% of cases, sometimes along with the cardiac symptoms, digestive pathologies, referred as "mega-syndromes" also develop (Matsuda *et al.*, 2009). These involve hyperdilation of organs of the digestive tract, most commonly colon and oesophagus, and are responsible for significant morbidity and mortality. The cause of these complications is not well understood, but it has been suggested to be the result of inflammatory damage combined with age-related loss of neurones in the gut wall (Jabari *et al.*, 2014).

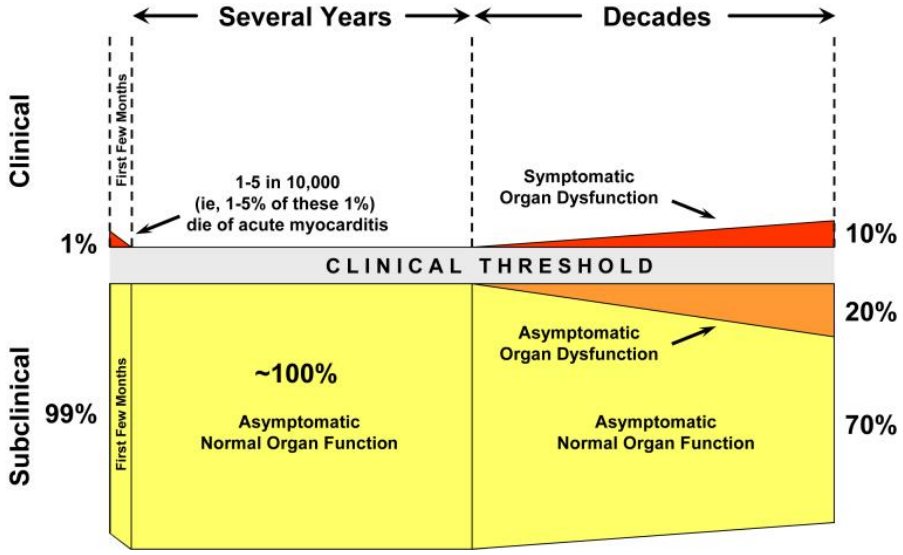


Figure 1.6: Clinical progression of CD. Image taken from (Bonney and Engman, 2015)

It is thought that once the clinical pathologies are established, drug treatment may not have much benefit (Morillo *et al.*, 2015). Treatment is mostly palliative, aiming to improve the life quality of patients (Viotti *et al.*, 2014). Heart transplants have been

conducted with some successes, but there are issues with infection re-appearance (Gray *et al.*, 2018). Drug cure seems to have a positive effect only if done before the onset of clinical pathology and is currently recommended by the World Health Organization (WHO) for all patients with positive serology. Nevertheless, only ~1% of cases in the endemic regions are treated (Alonso-Padilla *et al.*, 2019), due to a lack of diagnostic testing and different socioeconomic circumstances affecting the rural endemic areas. On the other hand, different initiatives in the Americas have helped achieve significant reductions in the number of acute cases of disease and the presence of domiciliary triatomine vectors in endemic areas (PAHO/WHO, 2022).

There are only two available chemotherapeutic drugs for CD patients, benznidazole (BZ) and nifurtimox (NX). The challenges in drug development for CD have been recently reviewed (Francisco *et al.*, 2020; Soeiro, 2022). Both BZ and NX are heterocyclic nitro-prodrugs activated by the parasite-specific mitochondrial enzyme type 1 nitroreductase (TcNTR-1), which catalyses the reduction of the nitro group on both drugs using flavin mononucleotide as the electron donor, to be then used as substrates of downstream reductive metabolism (Fig. 1.7A and B), yielding a series of cytotoxic compounds with multiple targets of which the mode of action, for both drugs, is poorly characterised (Wilkinson *et al.*, 2008, 2011; Trochine *et al.*, 2014). Reduced expression of TcNTR-1 *in vitro* induces cross-resistance to both drugs (Mejia *et al.*, 2012), however natural parasite isolates show variable drug susceptibility that is not linked to polymorphisms at the TcNTR-1 loci (Teston *et al.*, 2013). It is hypothesized that these resistance phenotypes are due to either alternative catabolic pathways for the pro-drugs, which yield less toxic metabolites (Murta *et al.*, 2006), or up-regulation of oxidative defence (Nogueira *et al.*, 2006), or enhanced DNA-repair pathways (Rajão *et al.*, 2014).

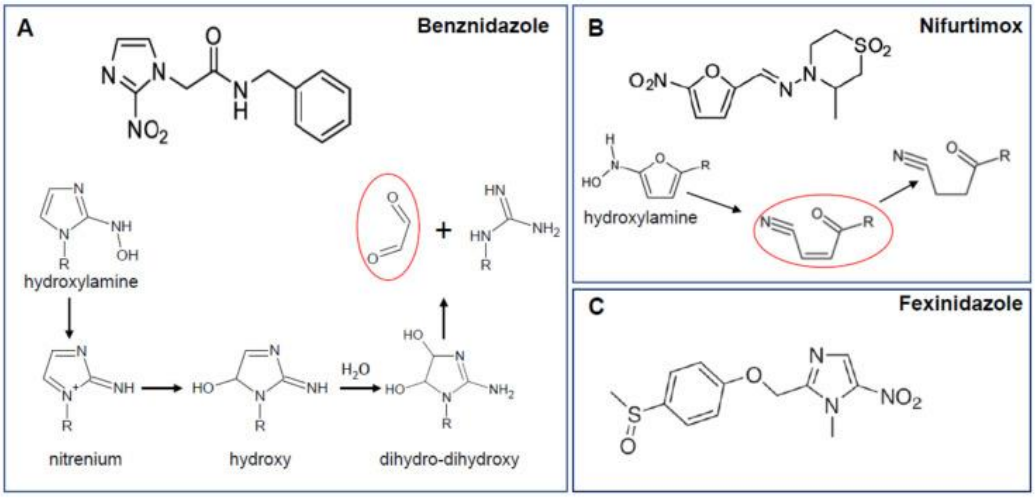


Figure 1.7: Nitroaromatic drugs used to treat CD or undergoing clinical trial. (A) TcNTR-1-dependent reduction of BZ generates hydroxylamine, which ends up yielding highly reactive glyoxal (circled in red). This and all the intermediates affect diverse proteins, DNA and small molecules. (B) TcNTR-1-dependent reduction of NX leads to the generation of unstable hydroxylamine, which decomposes into unsaturated (circled in red) and saturated open-chain nitriles that have trypanocidal activity. (C) Chemical structure of fexinidazole, TcNTR-1-activated prodrug, currently undergoing clinical trials against CD,. Image taken from (Francisco *et al.*, 2020)

Treatment with these compounds needs long regimens of 30 – 60 days and several adverse side effects are common (dermatitis, digestive tract intolerance, anorexia, sleeping disorders, headache and others) (Viotti *et al.*, 2009). Less commonly, the treatment also affects to bone marrow haematopoiesis, leading to neutropenia (Crespillo-Andújar *et al.*, 2018). A leading compound, fexinidazole (Fig. 3C), another nitroaromatic compound, has been licenced for treatment of the related kinetoplastid *T. brucei*, and is undergoing clinical trial against CD (Francisco *et al.*, 2020). More recently, an orally active benzoxaborole prodrug has been proved effective in the treatment of non-human primates with long-term naturally acquired infections. This compound was found to be active *in vitro* and *in vivo* against extra- and intracellular stages of a range of genetically distinct *T. cruzi* lineages (Padilla *et al.*, 2022).

1.3 *Crithidia fasciculata* as a model organism.

The first description of the organism called *C. fasciculata* was by Léger, 120 years ago (Léger, 1902), to denominate a flagellate found in the gut of a mosquito (*Anopheles maculipennis*) collected in the former southeast province of the Dauphiné (France). This insect was responsible for most of the malaria transmission in European countries until the 1970s (European Centre for Disease Prevention and Control, 2018). Currently, it is known that *C. fasciculata* infects *Anopheles* and *Culex* mosquitoes throughout the world, and it is usually found in the mid- and hind-gut of the insects (Olsen, 1974). *C. fasciculata* is 4-10 µm long, shaped like a barley corn, and with a funnel-like depression at the truncate anterior (Léger, 1902; Wallace, 1943). This monoxenous trypanosomatid that

infects mosquitos is non-pathogenic to humans. Clark and colleagues (Clark *et al.*, 1964) suggested a life cycle in which the parasite, present in flowers, fruits, mosquito faeces or on dead adult mosquitos that have fallen into the water, infects mosquito larvae, pupae, and adults, and grow inside their gut, multiplying into large numbers until leaving the host through faecal deposition or death of the host, to start a new cycle. During this process, two life forms have been described, a non-motile one attached in clusters to the lining of the insect gut (haptomonad), and a more elongated form that swims freely in the gut lumen (nectomonad) (Wallace, 1943; Brooker, 1970). These forms resemble the morphotypes amastigote and choanomastigote, typical of trypanosomatids (Kaufer *et al.*, 2017), and so other authors have denominated the life forms of *C. fasciculata* as such, with no current clear consensus.

C. fasciculata is easy to culture in large numbers (see section 3.5.1). The research performed during the first ~60 years after its discovery in order to define the culture conditions and research utility of this and other monoxenous trypanosomatids is out of the scope of this introduction, but has been reviewed in the past (Guttman, 1963). Over the past six decades, *C. fasciculata* has been heavily used for studying biochemical, genetic, and cellular processes unique to the family Trypanosomatida, and which have been later the basis of many studies on the search for drug targets in the dixenous members causing terrible human and animal diseases. Thus, seminal research on understanding the different aspects of DNA metabolism in trypanosomatids (examples in (Podestá *et al.*, 2003; Sinha *et al.*, 2006)), the replication of the kinetoplast (reviewed in (Liu *et al.*, 2005)), the structure, functionality and biogenesis of the flagellum (reviewed in (Gadelha *et al.*, 2004; Sahin *et al.*, 2004)), initial attempts to purify an enzymatically active mitochondrion (Hill *et al.*, 1968; Nichols and Cross, 1977) or the redox mediator trypanothione (reviewed in (Comini *et al.*, 2005)), were in all cases performed using this parasite as the experimental platform. Currently, *C. fasciculata* is proposed as model to study the molecules and biochemical processes involved in the attachment of trypanosomatids to the gut of their vectors, due to its capacity of attaching itself to inert surfaces such as plastics from culture flasks (Filosa *et al.*, 2019; Denecke *et al.*, 2022). However, it is important to note that the specificities of interactions between trypanosomatid cell membranes and the insect gut epithelial cell layer may not be comparable with those involving the attachment to inert surfaces.

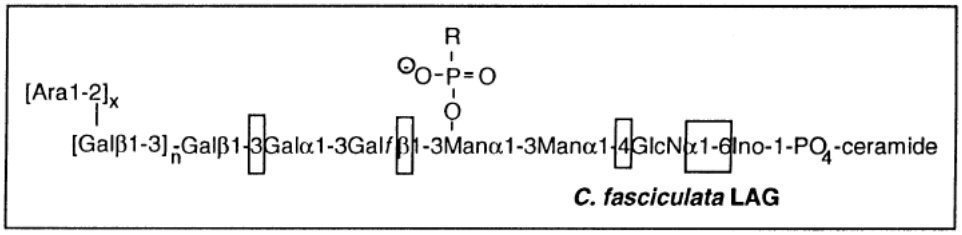
Regarding the cell biology of *C. fasciculata*, the bulk of analyses were done in the decades before the 1980s, mainly through the development of methods of

homogenization, separation, and purification of subcellular organelles and structures with a resulting rapid development in our understanding of some of these components. For example, an array of subpellicular microtubules functionally resembling the cytoskeleton of other eukaryotic cells were described under the plasma membrane of *C. fasciculata* (Brooker, 1972), whose lipid composition analysis showed the presence of phospholipids phosphatidylethanolamine and phosphatidylcholine and only ergosterol as the sterol present in its membrane, unless the parasite was grown in serum-containing media, in which case a small percentage of cholesterol was also found (Korn, Von Brand *et al.*, 1969). In ultrastructural studies (Brooker, 1971b; Brun, 1974), an extensive cytostomal system, involved in the endocytosis of ferritin by the generation of various vesicles, was described in *C. fasciculata*, along with the presence of large numbers or cytoplasmic ribosomes and polysomes attached to the outer nuclear membrane. The rest of the organelles showed a typical eukaryotic morphology. For example, the endoplasmic reticulum (ER) showed both smooth and rough types, and relationships with the outer nuclear membrane. The Golgi apparatus was located anterior to the nucleus and near the base of the flagellar pocket. As for all trypanosomatids, the unique single mitochondrion, consisting of several interconnected tubular elements ramifying throughout the cell, was also reported for *C. fasciculata* (Paulin, 1977). Others cellular aspects such as the replication of the kinetoplast, the structure of the nucleus, cell division, flagellar structure, and others, have also been reviewed (McGhee and Cosgrove, 1980).

For the survival and infection success of trypanosomatids, they must adapt to their variable habitat, especially important for dixenous parasites where both insect and mammal hosts present such a different biochemical environment. This adaptation is very well reflected by the interaction between the interfaces of host and parasite, i.e., between their plasma membranes. Glycoconjugates present in the trypanosomatid plasma membrane are critical components for this interaction (Ferguson, 1997; Valente *et al.*, 2019). The structures of these glycoconjugates diverge among the different parasites and contribute to their virulence and survival (Rodrigues *et al.*, 2015). (Brooker, 1976) showed that in *C. fasciculata* cells from the hindgut of a mosquito host, the plasma membrane was covered by a layer of electron-dense, filamentous material giving a positive reaction for polysaccharide. This layer was absent from the membranes of the flagellum and the flagellar pocket. Later, a major surface glycoconjugate, a glycolipid, was found in the cell membrane of this parasite, named lipoarabinogalactan (LAG), and fully characterized (Schneider *et al.*, 1996). Additionally, a glycoconjugate containing

sialic acid has also been described on *C. fasciculata* cell surface (do Valle Matta *et al.*, 1999), although the full structural characterization has not been reported, and two independent agglutination experiments using sialic acid-binding lectins indicated no positive reaction (Esteves *et al.*, 1987; Motta *et al.*, 1991).

LAG has a glycan and a lipid moiety (Schneider *et al.*, 1996). It is composed of an inositolphosphoceramide in which the C6 position of the inositol is substituted with a glycan core and an arabinogalactan oligosaccharide. The LAG structure is:



where the average size of the β1-3 D-galactan chain (n) is 12 residues and where on average 8 (x) out of 12 β-galactose (βGal) residues are substituted with D-arabinopyranose (D-Ara) residues in α1-2 linkage. Based on overall similarity of the core glycan to that of *Leishmania* spp. LPGs, the R substituent is most like α–glucose (αGlc). The structural details in boxes have not been directly determined but are inferred from similarity with *Leishmania* spp. lipophosphoglycan (LPG), its major surface glycoconjugate.

D-Ara is an unusual sugar that has only been found in two other protein-free glycophosphatidylinositol species from closely related genera of trypanosomatids, the LPG from *Leishmania major* (Descoteaux and Turco, 1999), and the glycosylinositolphospholipids (GIPLs) of *Endotrypanum schaudinni* (Previato *et al.*, 1993; Xavier Da Silveira *et al.*, 1998).

In this thesis, we will describe in sections 5.1 and 6.3 how *C. fasciculata* has provided the most suitable experimental platform for the second of our aims (section 2). Our experimental approach has been based on biosynthetic labelling with tritiated fucose ([^{3}H]Fuc) of a trypanosomatid in order to then detect any radiolabelled potential mitochondrial product. This required an organism in which this procedure could be feasible and efficient. Of the different species considered, *C. fasciculata* was elected as the best candidate. This was not only because of its non-pathogenicity and ease to be cultured, but most importantly, because based in previous experiments using D-Ara, which in the pyranose form is structurally very similar of L-fucose (L-Fuc), and recent

genetic analysis on the metabolism of sugars in the related genus *Leishmania* spp., we predicted that *C. fasciculata* should be able to take up and incorporate the radiolabelled sugar [^{3}H]Fuc into the nucleotide sugar guanosine diphosphate-[^{3}H]Fuc (GDP-[^{3}H]Fuc) and thence into any glycoconjugate containing L-Fuc. Indeed, our results confirmed the radiolabelling of LAG, and at least one more potential intracellular fucosylated species. Furthermore, it will also be described later (section 4.10) how we used the same biosynthetic machinery from the fucose salvage pathway of this parasite to synthesize our own GDP-[^{3}H]Fuc, which was a critical component for the first of our aims (section 2). Indeed, *C. fasciculata* has represented for us a versatile tool with which we have been able to contribute to the study of kinetoplastid mitochondrial fucosyltransferases.

1.4 Notes on glycosylation, with special reference to fucosylation and its related enzymes.

Carbohydrate chains, that can be free or attached to proteins or lipids to form simple or complex glycoconjugates, are known as glycans (Varki *et al.*, 2022). Most glycans are found on plasma membrane-associated or secreted macromolecules and are remarkably diverse. Protein-bound monosaccharides and short glycans are also found in the nucleus and cytoplasm of cells, where they have important regulatory functions. Apart from forming important structural features, the sugar components of glycoconjugates modulate a wide variety of functions in physiological and pathophysiological states (Varki, 2017).

Glycoconjugates are formed through the process of glycosylation, which consists of the sequential addition of sugars to proteins and lipids or the *en bloc* transfer of glycans to proteins (Varki *et al.*, 2022). In eukaryotes, this process involves typically up to nine monosaccharide types which are combined to create a variety of unique glycan structures on lipid and protein molecules that modulate their function (West *et al.*, 2021). This requires numerous glycosyltransferases (GTs), such that final glycan structures differ in their sequences, connections between their sugar components (including branching) and length. In addition, there are several other transferases that can modify glycans by the addition of acetyl, methyl, phosphate, sulphate, and other groups.

Whereas GTs are responsible for the assembly of glycans, their processing and turnover is mediated by glycosidases. The degradation of glycan structures by glycosidases is necessary for the uptake and metabolism of sugars and for the turnover of glycoconjugates in various cellular processes. Glycosidases, or glycosyl-hydrolases, are also involved in the formation of intermediates that are used as substrates for GTs in the

biosynthesis of glycans (Varki *et al.*, 2022). Different aspects of this family of enzymes have been extensively reviewed (Bojarova and Kren, 2009; Cerqueira *et al.*, 2012; Elferink *et al.*, 2020).

In eukaryotes, many GTs are type II transmembrane proteins located in the Golgi apparatus with a short N-terminal domain, a transmembrane domain followed by a stem region, and a large C-terminal catalytic domain at the luminal side of the Golgi membrane (Breton *et al.*, 2012). Currently, GTs are classified into 115 families in the Carbohydrate-Active enZYmes (CAZy) database (Drula *et al.*, 2022) (accessed on 1 November 2022).

Research on GT functionality based on their expression and purification for structural characterization has numerous challenges, discussed extensively in (Gloster, 2014; Videira *et al.*, 2017; Moremen *et al.*, 2018; Moremen and Haltiwanger, 2019). For example: (i) the crystallization process is challenging since GTs are often multi-domain proteins and undergo considerable conformational changes; (ii) it is difficult to produce them in a high yield in recombinant form, especially in their membrane-bound form; (iii) characterization is laborious as the identification of both donor and acceptor substrates is required; and (iv) GTs can undergo a series of post-translational modification events, such as N-glycosylation, disulphide bond formation, and proper folding may require the assistance of chaperones. Consequently, eukaryotic expression systems, such as HEK293-derived cells, are often the most successful for GT expression (Moremen *et al.*, 2018).

The nomenclature of GTs reflects their preferred donor monosaccharide. Thus, families of glucosyltransferases, galactosyltransferases, sialyltransferases, etc. are differentiated. Further, GTs are defined by their donor substrate, the structure of their aglycone acceptor, the anomericity of the transferred sugar (α or β) and the inter-sugar glycosidic linkage stereochemistry (eg. 1-2, 1-3, 1-4 or 1-6). For example, a UDP-Gal : βGlcNAc β1-4 galactosyl transferase makes Galβ1-4GlcNAcβ1-O-R from a UDP-Gal donor and a GlcNAcβ1-O-R acceptor.

In this thesis we will focus on fucosyltransferases (FUTs), i.e., those which transfer the sugar Fuc from the nucleotide sugar donor GDP- β-L-Fucose (GDP-Fuc) to a variety of acceptors.

GTs account for about 1–2% of the gene products of an organism, including archaea, bacteria, or eukaryotes (Albesa-Jove and Guerin, 2016). More than 750,000 glycosyltransferase sequences are currently known across all phyla. Nevertheless, based

on the solved structures of members from 59 families, GTs are grouped into 4 fold-types, known as GT-A, GT-B, GT-C, and lysozyme-type folds. The GT-A and GT-B enzymes use nucleotide sugar donor substrates, whereas the GT-C and lysozyme-type enzymes use lipid-linked sugar donors (Albesa-Jove and Guerin, 2016; Varki *et al.*, 2022). FUTs belong to the GT-B fold family. A general characteristic of the GT-B enzymes is the presence of two β/α/β Rossmann-fold domains separated by a large cleft that contains the reaction centre (Lairson *et al.*, 2008; Breton *et al.*, 2012). In general, the N-terminal domain is involved in acceptor substrate recognition whereas the nucleotide-sugar donors mainly bind to the C-terminal domain. Since acceptors are much more chemically diverse compared with nucleotide-sugar donors, the N-terminal domains of these GTs show different rearrangements of secondary structural elements. Acceptor substrates are typically bound in the cleft between the two domains, and unlike the GT-A enzymes, the GT-B GTs are metal-ion-independent and do not possess a Asp-X-Asp (DXD) motif, which in GT-A GTs is responsible for divalent cation binding (usually Mn^{2+} or Mg^{2+}) and catalysis (de Vries *et al.*, 2001; Lairson *et al.*, 2008; Breton *et al.*, 2012; Albesa-Jove and Guerin, 2016; Varki *et al.*, 2022).

Nucleotide sugar dependent GTs catalyse a glycosyl transfer reaction either by retention or inversion (Fig. 1.8A) of stereochemistry at the anomeric reaction centre of the donor substrate (Breton *et al.*, 2012). FUTs transfer L-Fuc from GDP-β-L-Fuc to oligosaccharide or peptide acceptors with inversion of the anomeric configuration of the sugar (Breton *et al.*, 1998; Grewal *et al.*, 2021). For most of the FUTs, the mechanism of inversion is denominated as an S_N2 single-displacement reaction mechanism, in which the side chain of one of the amino acids in the active site catalyses the reaction by deprotonating the acceptor and facilitating the direct inversion of the activated phosphate-leaving group (Fig. 1.8B). In other FUTs, an S_N1 mechanism has been proposed, where an ion pair is formed between acceptor and sugar after cleaving of the glycosidic bond of the nucleotide sugar and before the nucleophilic attack. Amino acid residues in the in the active site of the enzyme aid in this process (Fig. 1.8C).

Fuc can be transferred to different acceptor substrates at diverse positions. Thus, α1-2-, α1-3-, α1-4-, α1-6- and protein O-FUTs have been defined. Two superfamilies have been proposed (Coutinho *et al.*, 2003), one containing all α1-3- and α1-4-FUTs (in CAZy, GT10) and a second superfamily containing all the other FUTs (GT11, GT37, GT64 for α1-2-FUTs; GT23 for α1-6-FUTs; and GT41, GT65 and GT68 for O-FUTs).

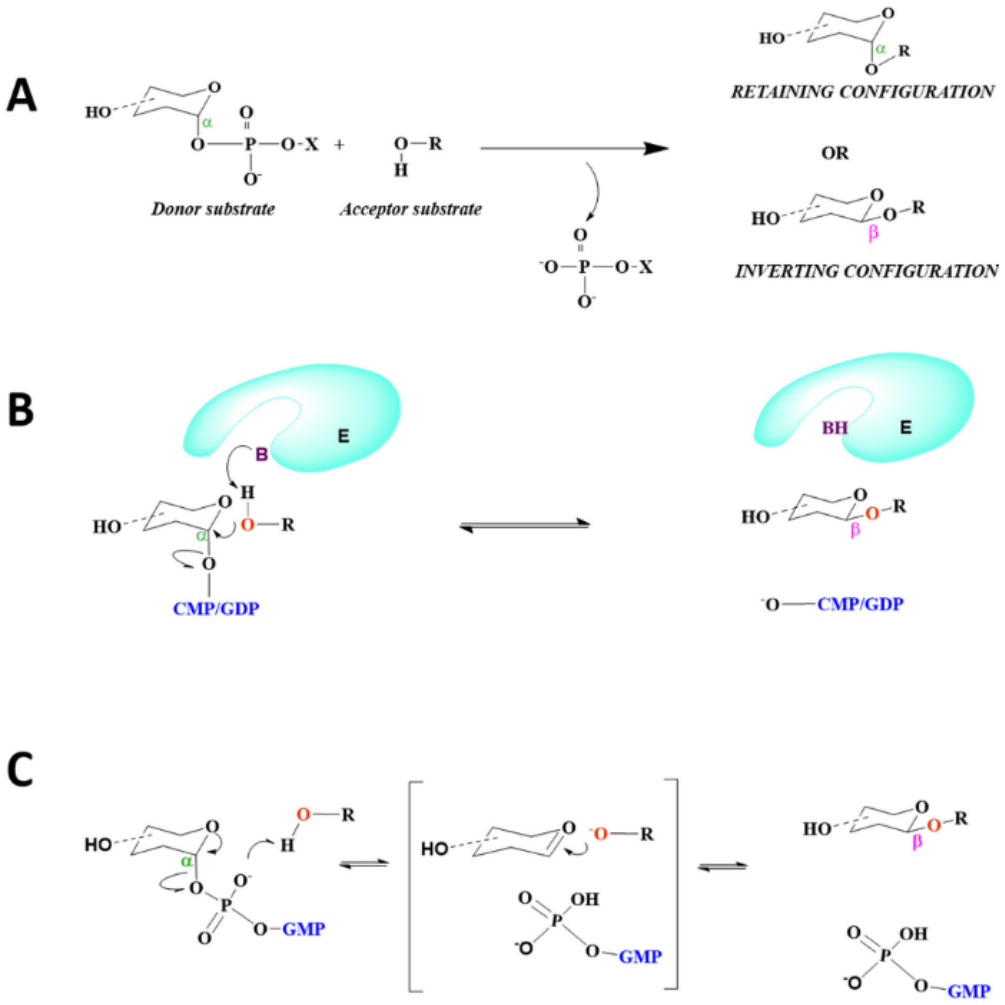


Figure 1.8: Reaction mechanism of GTs. (A) GTs catalyse the transfer of a sugar donor to an acceptor substrate either by retaining or inverting the anomericity of the donor. (B) S_N2 inversion mechanism. (C) S_N1 inversion mechanism. E represents the GT. Purple B represents the catalytic base. Red O represents the oxygen at the acceptor substrate that performs the nucleophilic attack. CMP, GMP and GDP (blue) are cytidine monophosphate, guanosine monophosphate, and guanosine diphosphate, respectively. Note that whereas in (B) and (C) nucleotide-bound α anomeric sugars are shown, the donor substrate of FUTs is GDP-L-β-Fuc, which is then inverted to an α-anomer when transferred to the acceptor. Image taken from (Grewal *et al.*, 2021).

In the majority of cases, α1-2-FUTs transfer Fuc to the galactose moiety of Galβ1-4GlcNAc or Galβ1-3GlcNAc structures which are localized at the peripheral (antennae) position of acceptor molecules (Ma *et al.*, 2006). In humans, the ABO blood group antigens (Cooling, 2015) are generated from the activity of two α1-2-FUTs, the H-transferase (FUT1) and the Secretor (Se) transferase (FUT2), which synthesize the glycan known as the H antigen by adding Fuc to a terminal galactose residue. Then, *ABO* locus-encoded GTs can modify the H antigen to generate A and B antigens in A, B or AB blood type individuals. These antigens are highly immunogenic and are found in large quantities on glycoproteins and glycolipids of red blood cells.

Interestingly, increased α1-2 fucosylation of synapsin I (Murrey *et al.*, 2006) and of neural cell adhesion molecule (Liedtke *et al.*, 2001) regulates their turnover and

stability, which has been reported to be involved in neurite outgrowth and synapse formation in sensory neurons of the olfactory bulb of rats and mice. Furthermore, multiple olfactory bulb proteins involved in cell adhesion, ion and solute transport, ATP binding, synaptic vesicle formation, and cell signalling are all modified with α1-2-Fuc. Their fucosylation contributes to the olfactory bulb development (Murrey *et al.*, 2009).

Furthermore, an unusual non-Golgi located α1-2-FUT in *Dictyostelium discoideum* (van Der Wel *et al.*, 2001) and *Toxoplasma gondii* and other protists (Xu *et al.*, 2012) is localized in the cytoplasm and modifies Skp1 protein, a subunit of the SCF-E3 ubiquitin ligase (van Der Wel *et al.*, 2001). When facing a hypoxic environment, a proline hydroxylation pathway is activated, in which a single proline in the protein Skp1 is targeted by this activity, to then be glycosylated by a pentasaccharide that contains an α1-2-Fuc, transferred by a dual β1-3-galactosyltransferase/α1-2-FUT known as PgtA. This modification influences poly-ubiquitination and targeting to the proteasome of many other proteins (West *et al.*, 2010; Bandini *et al.*, 2019).

The α1-3-/α1-4-FUTs add fucose at α1-3 or α1-4 linkage to the GlcNAc moiety of Galβ1-4GlcNAc or Galβ1-3GlcNAc structures, localized at the antennae of acceptor molecules, or to the innermost GlcNAc moiety at the core of N-glycans in an α1-3 linkage (Varki *et al.*, 2022). Members of these group of FUTs contribute to the synthesis of Lewis blood antigens and Lewis-related carbohydrate antigens, which are present in glycoproteins and glycolipids involved in the processes of leukocyte adhesion, rolling and extravasation during an immune response (Cooling, 2015; Li *et al.*, 2018). In plants, the glycoconjugates containing α1-3-/α1-4-Fuc play an important role in the reproductive development and in cell-to-cell communication and/or recognition (Ma *et al.*, 2006; Schneider *et al.*, 2017). In some intestinal bacteria, these same fucosylated epitopes have a function of molecular mimicry to evade the host immune system and maintain a long-term infection (Schneider *et al.*, 2017).

The α1-6-FUTs add Fuc to the innermost GlcNAc moiety of the core unit in N-glycans. Interestingly, although being GT-B fold GTs they do not need divalent cations for their activity, it has been shown that their activity is increased in the presence of Mn^{2+}, Mg^{2+}, and Ca^{2+} (Ihara *et al.*, 2006). In humans, the only FUT with this activity (FUT8) is widely expressed in all tissues, and when knocked-out in mice, 70% of the new-borns die within 3 days of birth due to major developmental growth and respiratory defects, whereas survivors display severe growth retardation and lung complications (Schneider *et al.*, 2017). Core α1-6 fucosylation plays an important role in the antibody dependent

cellular toxicity, which is inhibited in α1-6-fucosylated IgG antibodies in human cells. Furthermore, different proteins involved in developmental and structural aspects of *Schistosoma mansoni* (a blood fluke human parasite), insects, *Caenorhabditis elegans* and soil bacteria are also fucosylated by these FUTs (Ma *et al.*, 2006).

L-Fuc can also be directly transferred to the hydroxyl group of serine a threonine residues of glycoprotein acceptors containing either the epidermal growth factor (EFG) or thrombospondin type repeat (TSR) sequences. The FUTs involved in this process are O-FUTs, known as POFUT1 and POFUT2 in humans (Holdener and Haltiwanger, 2019). Unlike the other FUTs, these enzymes are predominantly ER-located soluble proteins due to the presence of a KDEL ER-retaining sequence in them (Holdener and Haltiwanger, 2019). The EGF-repeat sequence has been found in over 100 proteins such as Notch, Notch ligands, plasminogen activators, and coagulation factors, whereas the TSR sequence, found in around 50 proteins, is present in extracellular matrix proteins involved in cell–cell and cell–matrix interactions (Schneider *et al.*, 2017; Holdener and Haltiwanger, 2019). O-fucosylation has also been found in members of the Apicomplexa protozoa, in which its functional, evolutionary and biochemical aspects are currently under study (Bandini *et al.*, 2019).

1.5 Mitochondrial FUTs, and other GTs, in kinetoplastids.

In general, based on the extensive glycobiological knowledge built up over many years, eukaryotic GTs are located in the ER, the Golgi apparatus, and also cytosolic (Ma *et al.*, 2022), and as a consequence, the vast majority of glycans are present on their cell surfaces or are secreted (Varki *et al.*, 2022).

Initial analyses on the biosynthesis of the nucleotide sugar GDP-Fuc in the Kinetoplastids (trypanosomatids) *T. brucei* (Turnock *et al.*, 2007) and *L. major* (Guo *et al.*, 2017) indicated that it was required for the parasite proliferation, suggesting the presence of one or more essential FUTs. Later, reverse genetic approaches identified a single FUT from the GT11 family (an α1-2-FUT), named as TbFUT1, in the genome of *T. brucei*, which was essential for the bloodstream and procyclic life stage forms of the parasite (Bandini *et al.*, 2021). Its orthologue in *L. major*, LmFUT1, was also found, functionally characterized and demonstrated to be essential for the promastigote form (Guo *et al.*, 2021). Recombinant TbFUT1 was used in radioactivity (tritium)-based activity assays and later, fine chemical characterization of the glycan product defined its

activity as a GDP-Fuc : βGal α1-2-FUT, typical of GT11 enzymes, with an apparent preference for acceptor substrates containing a terminal Galβ1-3GlcNAc (LNB) motif (Bandini *et al.*, 2021). Strikingly, the subcellular localisation of both orthologues was shown to be in the mitochondrion of these parasites (Bandini *et al.*, 2021; Guo *et al.*, 2021).

The GT11 family has been recently identified as one of 63 initially viral GT families that eukaryotes acquired during their evolution through a horizontal gene transfer event. The genomic incorporation of such enzymes from different viruses correlates with different events of functional and structural diversity in the involved eukaryotic groups. Thus, in the case of kinetoplastids, a mitochondrial structural divergence (sole tubular and ramified structure containing kinetoplastid DNA) may correlate with the horizontal gene transfer of a GT11 activity via a nucleocytoplasmic large DNA virus, which has then been conserved throughout all kinetoplastids, implying functional relevance under long-term selection (Irwin *et al.*, 2022).

The GT11 family is characterized by 4 distinctive motifs (Martinez-Duncker *et al.*, 2003; Li *et al.*, 2008) , which were also identified in TbFUT1 (Bandini *et al.*, 2021), even though it showed relatively low sequence identity to previously characterized GT11 FUTs, such as *Helicobacter pylori* FUT (26%) (Wang *et al.*, 1999) and human FUT2 (21%) (Kelly *et al.*, 1995). Motif I (amino acids 153 – 159) is shared with α1-6-FUTs and in that family it has been implicated in GDP-Fuc binding (Takahashi *et al.*, 2000). No clear functions have been yet assigned to motifs II, III and IV (amino acids 197 – 207, 265 – 273, and 13 – 18, respectively).

Eukaryotic FUTs are type II transmembrane proteins, except in exceptional cases such as the *D. discoideum* α1-2-FUT (van Der Wel *et al.*, 2001; Breton *et al.*, 2012). Nevertheless, a clear consensus on the presence of a transmembrane domain (TM) at the N-terminus of TbFUT1 was not detected by secondary structure prediction analyses, and when a low-confidence prediction was obtained, the residues involved in the potential TM domain contained motif IV (amino acids 6 – 20) (Bandini *et al.*, 2021). Further, since this region is predicted to include a cleavable mitochondrial import sequence, we predict that once transported into the matrix of the mitochondrion, trypanosomatid FUT1s are processed to non-membrane bound proteins.

Since LNB (Galβ1-3GlcNAc) is the preferred substrate for recombinant TbFUT1 (Bandini *et al.*, 2021), this opens the possibility of the existence of other GTs that

assemble the acceptor on a protein, a lipid, or other type of molecule prior to the TbFUT1 action. These putative GTs could be cytosolic, implying an import of the LNB-containing molecule(s) into the mitochondrion for fucosylation, or they could be mitochondrial themselves (Duncan and Ferguson, 2022). Importantly, two GT25 GTs known as TbGTX (Tb927.11.3900) and TbGTZ (Tb927.11.9990), have been predicted to be localised in the mitochondrion of *T. brucei* (Aslett *et al.*, 2010; Peikert *et al.*, 2017), and there is proteomic evidence of TbGTX expression in both bloodstream and procyclic forms of *T. brucei* (Duncan and Ferguson, 2022).

An orthologue of the genes encoding the *T. brucei* and *L. major* mitochondrial FUTs has been found in *T. cruzi*, which we have named TcFUT1. Currently, one of the challenges in developing anti-trypanosomatid therapeutics is in generating drug candidates that demonstrate sterile cure in mouse models of *T. cruzi* infection. This, in turn, requires the identification and characterisation of new molecular drug targets with which to pursue structure-based drug discovery campaigns. Thus, studying the *T. cruzi* orthologue activity and finding the potential endogenous substrates of the mitochondrial FUTs may provide an opportunity to uncover the fundamental function(s) and biology of this novel class of enzymes, and the novel concept of mitochondrial glycosylation, and to provide a much-needed drug target for CD for translational research.

2. AIMS.

The aim of this thesis was to gain more knowledge on the novel mitochondrial fucosylation activities found in kinetoplastids. Different experimental approaches were taken, to address:

Aim 1: Characterisation of *T. cruzi* FUT1 activity.

Aim 2: Use of *C. fasciculata* to identify kinetoplastid mitochondrial substrates of FUT1.

3. MATERIALS AND METHODS.

3.1 Materials.

All general reagents were purchased from Merck (Sigma), Thermo Fisher Scientific and VWR unless otherwise indicated. Enzymes were purchased from NEB and Thermo Fisher Scientific.

3.2 Bioinformatics.

3.2.1 Genome sequence retrieval.

DNA and amino acid sequences for genes of interest were retrieved from TriTrypDB in FASTA format (https://tritrypdb.org/tritrypdb/app). TriTrypDB was also used to identify orthologue genes between *Leishmania spp.*, *T. brucei*, *T. cruzi* and *C. fasciculata* for each gene of interest.

3.2.2 Sequence manipulation and vector and primers design.

In silico manipulation of DNA and amino acid sequences was conducted using CLC Main Workbench version 8.0 software. Features of this software allowed the design of oligonucleotide primers, sequence alignment, *in silico* cloning, restriction digest mapping and reverse translation of proteins. A number of vector fragments were designed using the CLC software and subsequently synthesised by GenScript USA Inc. or Twist Bioscience.

3.3 Prokaryotic expression and related protocols.

3.3.1 *E. coli* strains used.

For transformations of restriction enzyme-based and Gibson assembly ligation reactions, NEB 5-α competent cells were used. For expression and purification of rTcFUT1, ArcticExpress (DE3) (Agilent Technologies), BL21 (DE3) pLysS (Promega) and Shuffle T7 Express *lysY* (NEB) competent cells were used. Their choice as expression systems is detailed in sections 4.3-4.5.

3.3.2 Bacterial culture and storage.

E. coli were grown in LB medium containing 100 µg/ml carbenicillin, and 20 µg/ml gentamycin for ArcticExpress (DE3) cells, at 37ºC and 200 RPM in Infors HT Multitron Standard incubators. After transformation, colonies were selected on LB + agar plates grown at 37ºC for an average of 16 h. For bacterial stabilates, 450 µl of log phase cells were mixed with 50 µl of sterile 80% glycerol and frozen on dry-ice before storage at -80ºC. All media and plates were obtained from the Central Technical Services in the School of Life Sciences, University of Dundee.

3.3.3 Transformations.

Aliquots of 50 µl competent cells were thawed on ice before adding 1 – 5 µl (0.1 – 10 ng) DNA and incubating on ice for 30 min. Cells were subjected to heat shock (25 secs for ArcticExpress (DE3), 30 secs for NEB 5-α and Shuffle T7 Express *lysY*, or 45 secs for BL21 (DE3) pLysS) in a 42ºC water bath and then incubated on ice for 2 min. Either 950 µl (for NEB 5-α) or 500 µl (rest of strains) of pre-warmed to room temperature (RT) SOC medium were added to the tube, which was then incubated for 1 hour at 37ºC with shaking at 210-220 rotations per minute (RPM), inside a 50 ml Falcon tube. Aliquots (100 µl) of grown culture, or from different dilutions, were plated out onto selective LB agar plates and incubated overnight at 37ºC for an average of 16 h. The components of SOC medium are 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose added after autoclaving solution with the remaining components, and letting it cool down; after addition, the solution was filtered by passing through a 0.2 µm filter.

3.3.4 Protein purification of prokaryotic-expressed TcFUT1.

As discussed in sections 4.3-4.5, different bacteria strains and conditions were used in order to express TcFUT1 and obtain sufficient protein for enzyme assays. For the successful expression and large-scale purification of (inactive) recombinant TcFUT1, 50 ng of construct rTcFUT1-2.b were used to transform SHuffle T7 Express lysY *E. coli* (NEB). One colony was picked and cultured for 24 h at 37ºC, 200 RPM in 100 ml LB with ampicillin (100 µg/ml) and chloramphenicol (33 µg/ml). Then, the culture was diluted to an OD600 of 0.1 in up to 1L of same media, induced with 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and incubated at 20ºC, 200 RPM for 24 h.

After the induction period, cells were harvested by centrifugation (3,200 x g 30 min 4°C) (Beckman J6-MC). Pellets were resuspended in 10 ml of 150 mM Tris-HCl, pH 7.2, 150 mM NaCl, 1 mM TCEP (buffer A), to which one mini-tablet of EDTA-free

protease inhibitor cocktail (Merck) were added before being disrupted in a cell disruptor (Constant Systems One Shot Cell Disruptor) at 30 thousand pounds per square inch (KPSI). The soluble fraction was obtained by centrifugation at 45,000 x g 30 min 4°C (Beckman Avanti J-25i), filtered with a 0.45 µm pore size filter and injected in an AKTA system (Cytiva AKTA Pure) to perform immobilized metal affinity chromatography (IMAC). Pre-packed column (HisTrap 5 ml nickel column, Cytiva) was pre-equilibrated in buffer A before incubation with the sample. After injection, washing was done with 10 column volumes of 8% buffer B (buffer A plus 250 mM imidazole) and then with 9 column volumes of 10% buffer B. Elution was performed with a gradient from 10 to 100% of buffer B over 20 column volumes and a final linear elution at 100% buffer B for 4 column volumes. Fractions of interest that could contain rTcFUT1 after analysing chromatogram were used for SDS-PAGE (section 3.7.3) and the gel analysed by Quick Coomassie staining (section 3.7.4), Ponceau S staining and LICOR Western blotting (section 3.7.6), using mouse monoclonal antibody α-Hisx6 tag-HRP 1:3,000 as primary antibody, which was developed with 1:15,000 goat α-mouse green LI-COR secondary antibody. If used for activity assays, recombinant protein was dialysed against Tris-HCl 50 mM pH 7.2, 25 mM KCl at 4ºC O/N and protein concentration estimated by BCA assay (ThermoFisher Scientific) prior to use.

3.4 Eukaryotic expression and related protocols

3.4.1 Expi293F cell stabilate recovery, culture, cell count and storage.

Expi293F cells (1 ml vial) was defrosted on a 37ºC water bath and transferred to 29 ml pre-warmed (to 37ºC) Expi293 Expression Medium in a 125 ml PETG polyester flask with vented closure. Cells were incubated for up to 7 days to recover at 37ºC, 8% CO_2 atmosphere on an Infors HT Multitron Pro shaking at 125 RPM. Although required by manufacturer's user manual, we determined that a ≥80% humidified atmosphere was not necessary for a good and consistent growth of these cells. When original culture reached 1 - 3 × 10^6 viable cells/ml and viability of culture was above 90%, cells were ready for first subculture.

For general maintenance of cells, cells were split when cultures reached 3 - 4 x10^6 viable cells/ml, and only when viability was ≥95%, to 2 – 4 x 10^5 cells/ml. Cells divide once per day. Cultures of different volumes were made following manufacturer's recommendations to select the appropriate flask for the volume needed.

To count cells, 180 µl of culture was mixed with 20 µl Trypan Blue solution (0.4% w/v); 20 µl of this mix was then pipetted into a counting chamber (Nexcelom Bioscience) and put inside a Cellometer Auto 1000 (Nexcelom Bioscience) for cell concentration and viability quantifications, using a programme that defines and quantifies viable cells as any detected 7 – 35 µm diameter, translucent sphere in a 1:100 dilution.

To generate new stabilates, cells were counted and centrifuged at 300 x g, 5 min, RT (Sigma Laboratory Centrifuge 4K15). The supernatant was discarded, and cells resuspended with gentle pipetting to 1 x 10^7 viable cells/ml in Expi293 Expression Medium containing 10% DMSO; 1 ml aliquots were transfer to 1 ml cryogenic vials with external thread (Starlab), kept for 16 h at -80ºC in a Mr Frosty freezing container previously filled with 250 ml 2-propanol, and then transferred to liquid nitrogen tank for long-term storage. From experience, vials stored for longer than 18 months are no longer recoverable.

3.4.2 Transfections.

For 30 ml transfections, cells were subcultured to 3 x 10^5 viable cells/ml and left culturing for 4 days. When cells reached ~5 x 10^6 cells/ml, cells were diluted to 2.94 x 10^6 viable cells/ml in 25.5 ml in a 125 ml flask and left in incubator while preparing the DNA-reagent complex. For this, two Bijou bottles were used. In the first one, 30 µg (~1 µg/µl) of plasmid DNA or positive intracellular and extracellular expression controls, or no DNA for negative control, and 1.5 ml Opti-MEM I Reduced Serum Medium, were mixed. In the second one, 80 µl of ExpiFectamine 293 Reagent and 1.5 ml Opti-MEM I Reduced Serum Medium were mixed. Both mixes were left incubating for 5 min at RT, and then mixed together, and left incubating for 20 min at RT to allow the DNA-reagent complexes to form. After this, the ~3 ml complex solution was added to the cell culture, bringing the total volume to approx. 28.5 ml. Flask was left in shaking incubator for 18 h to then add 150 µl of ExpiFectamine 293 Transfection Enhancer 1 and 1.5 ml of ExpiFectamine 293 Transfection Enhancer 2. The ~30 ml culture was left in a shaking incubator for further 2 – 7 days before harvesting for construct expression analysis or α-Myc immunoprecipitation (IP) (section 3.4.3).

For 500 ml transfections (maximum cultivable volume due to the capacity of largest flask), cells were subcultured to 3 x 10^5 viable cells/ml and left culturing for 4 days. When cells reached ~5 x 10^6 cells/ml, cells were diluted to 2.94 x 10^6 viable cells/ml in 433.5 ml in a 2 L flask and left in incubator while preparing the DNA-reagent complex.

For this, in a first Bijou bottle 510 µg (≥1 µg/µl) of plasmid DNA and 25.5 ml Opti-MEM I Reduced Serum Medium were mixed, whereas a second Bijou bottle contained a mix of 1377 µl ExpiFectamine 293 Reagent and 25.5 ml Opti-MEM I Reduced Serum Medium. After being left for 5 min at RT, both contents were mixed and incubated for further 30 min, RT. Generated DNA-reagent complex solution (~51 ml) was added to cell culture, flask left or 18 h in shaking incubator, and then 2.55 ml of ExpiFectamine 293 Transfection Enhancer 1 and 25.5 ml of ExpiFectamine 293 Transfection Enhancer 2 were added. The ~500 ml culture was left in shaking incubator for further 4 days before harvesting for a large-scale purification attempt (Fig. 4.26A).

3.4.3 Immunoprecipitation of eukaryotic-expressed TcFUT1.

Myc-tagged TcFUT1 was either expressed recombinantly in transfected Expi293F cells with construct rTcFUT1-6.a and secreted into the medium or overexpressed in an electroporated population of *T. cruzi* with construct TcFUT1-7.a. Either rTcFUT1-containing Expi293F medium from one or two 30 ml transfected culture, or overexpressing *T. cruzi* lysate of 1 x 10^9 cells produced by lysing in 2% Triton X-100 in 1X PBS containing 1X protease inhibitor cocktail (Merck), were pre-cleared by spinning at 16,000 x g, 10 min, 4ºC (Eppendorf Minispin Plus). Supernatants were then incubated with 30 - 40 µl 50% α-Myc agarose beads slurry for 1 - 2 h at 4ºC with gentle shaking or rotating to avoid sedimentation of the beads. The beads were previously equilibrated by adding 1 ml 1X PBS and centrifuging at 2,500 x g, 5 min, 4ºC (Eppendorf Minispin Plus), discarding most of the supernatant and resuspending the beads in the residual buffer. After incubation, bead-protein complexes were sedimented by centrifugation at 2,500 x g, 5 min, 4ºC, and supernatant removed with a vacuum aspirator to avoid disturbing the pellet. Then, beads were washed once with 10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Igepal CA-630/Nonidet P40, 0.5 mM EDTA, 0.018% NaN_3. Washed bead-protein complexes were used to set up GDP-glo (section 3.7.14) or [^{3}H]-based fucosyltransferase activity assays (section 3.7.12).

3.5 Parasite cell culture and protocols.

3.5.1 *C. fasciculata* cell cultures.

C. fasciculata HS6 strain parasites (Shim and Fairlamb, 1988) were grown in SDM-79 medium (Brun and Schönenberger, 1979) supplemented with 2g/L $NaHCO_3$, 1X

GlutaMAX as L-glutamine source, and 20 µg/ml hemin in 50 mM NaOH instead of 7.5 µg/ml as in the original formula. When growing electroporated cells, the medium was supplemented with 10% heat-inactivated fetal bovine serum (HI-FBS). When performing subcloning, the $NaHCO_3$ concentration was increased to 4 g/L. Incubation was done without shaking at 28ºC, and when subcloning, 5% CO_2 was established inside the incubator. Culture flasks (CytoOne T-25, T-75 and T-225) were non-tissue culture treated and non-vented. Parasite growth was monitored by haemocytometer and cells were grown for 4-5 days up to late log phase (1 x 10^7 cells/ml) before being diluted to 1 x 10^4 cells/ml, showing a division time of 9 - 10 h for wild-type (WT) cells. Cells reach stationary phase at 1.5 - 1.8 x 10^7 cells/ml and were grown to these concentrations only when needed for LAG purification.

An alternative culture medium was also used. In this case, cells were grown in brain heart infusion (BHI) broth (3.7% solution w/v), autoclaved, 0.45 µm pore size sterile-filtrated and supplemented with 20 µg/ml hemin in 50 mM NaOH, in this order. The pH was consistently in a range of 7.4 ± 0.2 and did not need to be adjusted. Culture flasks were as described above when growing without shaking, or borosilicate glass Erlenmeyer flasks when growing at 100 RPM. Incubation was done at 28ºC, plus 5% CO_2 only when growing subcloning plates. Cells reached stationary phase at 1 - 2 x 10^8 cells/ml in these conditions.

3.5.2 *T. cruzi* cell cultures.

T. cruzi Sylvio X10/7A strain (Roberts *et al.*, 2014) epimastigote parasites were grown in RTH/FBS medium (Gibson and Miles, 1986; Pereira and Hoff, 1986; Hunter *et al.*, 1994) prepared as follows: to 500 ml RPMI 1640 (Merck) was added 10 ml 1 M HEPES pH 7.4, 2.45 g tryptone peptone (Thermo Fisher Scientific), 4 ml hemin solution at 2.5 mg/ml in 0.01 N NaOH and 60 ml heat-inactivated fetal bovine serum (Cytiva), with no need to adjust final pH. Incubation was done without shaking at 28ºC, 5% CO_2. Culture flasks (CytoOne T-25, T-75 and T-225) were non-tissue culture treated and non-vented. Parasite growth was monitored with a CASY cell counter with a 60 µm capillary, following manufacturer's recommendations. Cells were diluted in 100 µl of 1:100 solution with 1% paraformaldehyde (PFA) in 1X PBS, and then added to 10 ml of CASYton isotonic buffer for automated cell counting after purging the machine and normalizing against a blank the average of three counts. Cells were grown for 5-6 days up to late log phase (1 x 10^7 cells/ml) before being diluted to 2 - 5 x 10^5 cells/ml, showing

a division time of 24 h for WT cells. Cells reach stationary phase at 4 x 10^7 cells/ml but were always harvested before they reached stationary phase.

3.5.3 Generation of stabilates.

Parasites were grown to mid-late log phase (5 x 10^6 to 1 x 10^7 cells/ml) to then mix 700 µl cell culture with 100 µl 80% glycerol sterile solution in 1 ml cryogenic vials with external thread. These were then transferred directly to a -80ºC freezer for short-term storage, or liquid nitrogen tank for long-term storage.

3.5.4 Stabilate recovery.

Cryogenic vials were taken from liquid nitrogen storage, left defrosting inside hood (for *C. fasciculata*) or inside a 50 ml Falcon tube placed in an incubator (for *T. cruzi*) and then mixed with 9 ml of their appropriate medium. In the case of *C. fasciculata*, antibiotics were added directly, when required, whereas for *T. cruzi*, cells were left recovering for 24 h before addition of antibiotic.

3.5.5 Electroporation of *C. fasciculata* for gene deletion.

C. fasciculata cells were grown in SDM-79 with 10% HI-FBS to near late log phase (6 – 8 x 10^6 cells/ml). Cells were centrifuged (800 x g, 10 min, RT) (Sigma Laboratory Centrifuge 4K15) and washed three times with Cytomix buffer (Table 3.1) (van den Hoff *et al.*, 1992). Then, cells were resuspended to 3 – 4 x 10^7 cells/ml in the same buffer, and 450 µl were transferred to 4 mm electrode gap electroporation cuvettes (BTX) containing 1 – 10 µg of plasmid DNA, or no DNA as negative control. The plasmid had been previously digested with SacI-HF and BamHI-HF, according to NEB protocol (New England Biolabs, UK) to cut out the drug resistance cassette containing the GOI UTRs of different lengths (300 – 1,000 bp) away from the vector backbone. This linearized DNA is then available for homologous recombination. Digestion was performed for 16 h at 37ºC, after which enzymes were inactivated in a water bath at 80ºC for 30 min, precipitated with ethanol in the presence of 0.3 M sodium acetate O/N at -20C. The DNA was pelleted by spinning for 10 min at 16,000 x g at 4 ºC, followed by 3 subsequent washes with cold 70% ethanol. The DNA in the washed pellet was dissolved in sterile water and an aliquot used to confirm the correct digestion had taken place by agarose gel electrophoresis. The cells were electroporated using a voltage of 1,700 volts for 3 pulses of 0.1 ms each, with an interval of 0.2 secs between each pulse, using a BTX Gemini SC device.

After electroporation, the cells from each cuvette were transferred with a sterile plastic pipette to 10 ml of SDM-79 and left recovering in incubator at 28ºC for 5 – 6 h. Then, 10 ml of medium containing 10% heat-inactivated (HI)-FBS and double concentration of the selection antibiotics (1X concentrations: 50 µg/ml hygromycin and 25 µg/ml neomycin/G418) for the constructs, were added, followed by plating out in 24-well plates (non-tissue culture treated) with 1 ml in each well.

The 24-well plates were incubated for up to three weeks at 28ºC. Negative control cells died around day 5 – 7 post-electroporation. Resistant cells were diluted in subcloning medium (SDM-79 containing 10% HI-FBS and extra $NaHCO_3$ to be final 4 g/L) to 2.5 cells/ml (0.5 cells/well) and plated 200 µl/ well in 96-well plates (non-tissue culture treated plastic) and grown at 28ºC, 5% CO_2 until reasonably dense (around 10 days). Alternatively, 100 µl aliquots of drug-resistant cultures, or dilutions thereof, were spread on solid media following previous reports of clonal growth in this manner (Scolaro *et al.*, 2005; DiMaio *et al.*, 2018) and incubated at 28ºC. Plates contained solidified BHI medium with 0.65% agarose and 10 µg/ml haemin, to which selection antibiotics were added by spreading drug solution at working concentrations (see above) in sterile conditions, less than 24 h prior to being used and kept at 4ºC. Stabilates were generated from the resistant mixed populations (from 24-well plates) or clonal populations and their genotypes were verified by PCR and Southern blot analysis.

For the generation of Δ*CfLPG1*::*HYGr*/Δ*CfLPG1*::*NEOr* double knock-out (dKO) cell line in *C. fasciculata*, a modified electroporation protocol had to be used, after numerous failures of the original protocol. Δ*CfLPG1*::*NEOr* single knock-out (sKO) cell line was grown in SDM-79 containing 10% HI-FBS and 25 µg/ ml G418 and harvested at culture density of about 5 x 10^6 cells/ml. Washed cells were resuspended at 4 x 10^7 cells in 100 µL Cytomix, transferred into 2mm gap cuvettes containing 1 µg of LPG1-HYGr 1Kb UTRs cassette digested DNA (in 5µL of sterile water), or water only, as negative control, and electroporated using programme X-001 in an Amaxa electroporator (Lonza). Cells from each cuvette were transferred into 10 ml SDM-79 containing 10% HI-FBS, 4g/L $NaHCO_3$ and 25 µg/ ml G418 and recovered for 4h at 28 ºC, 5% CO2. Subsequently, equal volume of media containing 25 µg/ml G418 and 50 µg/ml hygromycin was added and immediately plated into 96-well plates as described above.

Table 3.1: Cytomix buffer composition.

Reagent	**Concentration**

EGTA	2 mM
KCl	120 mM
$CaCl_2$	0.15 mM
K_2HPO_4/KH_2PO_4	10 mM
HEPES	25 mM
$MgCl_2 \cdot 6H_2O$	5 mM
D-glucose (dextrose)	0.5 % (w/v)
Fatty acid-free BSA	100 µg/ml
Hypoxanthine	1 mM

Note: pH adjusted to 7.6 with NaOH. After preparation, buffer is sterile-filtered and kept at 4ºC.

3.5.6 Electroporation of *T. cruzi* for gene overexpression.

For this protocol, reagents and cuvettes from Human T Cell Nucleofector Kit (Lonza) were used. *T. cruzi* cells were grown to 5 x 10^6 cells/ml (mid log phase). 10 ml (5 x 10^7 cells) were used per transfection. Cells were centrifuged at 1,550 x g, 10 min, RT (Sigma Laboratory Centrifuge 4K15) and supernatant discarded. After resuspension in residual volume, they were transferred to microcentrifuge tubes for further centrifugation at 1,200 x g, 90 secs, RT (Eppendorf Minispin Plus) and supernatant discarded. The cell pellet was then resuspended in 100 µl of Amaxa buffer (previously prepared following kit's recommendation) and transferred to a cuvette from the kit. 10 µg of NotI-linearized or circular DNA containing the GOI to be overexpressed, C-tagged with 6xMyc, and G418 resistance gene, were transferred to the cuvette and cells were electroporated in an Amaxa Nucleofector II device set up in programme U-033. No DNA was added for the negative control.

Electroporated cells were added to 25 ml of pre-warmed RTH/FBS medium with a plastic pipette, and left incubating for 24 h at 28ºC, 5%CO_2. After it, neomycin was added to a final concentration of 200 µg/ml to each electroporated culture and left incubating in same conditions for up to 2 weeks. Negative control cells died after at least one week post-electroporation. If cells were alive, but growth was not evident after 2 weeks, cells were pelleted by centrifugation at 1,550 x g, 10 min, RT (Sigma Laboratory Centrifuge 4K15), and resuspended in 10 ml of RTH/FBS medium with fresh drug, and left incubating for up to a month.

When resistant cells were obtained, cells were diluted in same medium to 1 cell/600 µl and plated out in one to three 96-well plates containing 200 µl in each well, in order to obtain clonal populations. Plates' perimeter was taped up to avoid cell culture contamination and reduce infection risk by evaporation of culture. Plates were left up to two months incubating at 28ºC, 5%CO_2. Stabilates were generated from the resistant mixed populations and overexpression of tagged gene was analysed by α-Myc Western blot.

3.6 Molecular biology

3.6.1 Primers.

The primers used are listed in (Table 3.2). When possible, primers were designed to contain 20 – 25 bp of complementary sequence, have a guanine (G) or cytosine (C) at their 3' end and a melting temperature T_m around 60ºC. The T_m was calculated according to the following formula: $T_m = (G + C) \times 4ºC + (A + T) \times 2ºC$, where only the nucleotides in the complementary sequence of the primer were considered in the calculations. Primers were all synthesized by Thermo Fisher Scientific.

Table 3.2: List of primers

JC1	TCCACAGGTGTCCAGCGGCCATGGGCGTGCCGGCCGTCTTA
JC2	GCATGATGATGATGATGATGATGGCGGCCATCCCCCTGAAAGTACAGGTTCTCC AGTAAAATCCAGTTTTCACAGTAGAGGTGGT
JC3	TGAACAGTTCCTCGCCCTTGCTCATGCGGCCATCCCCCTGAAAGTACAGGTTCTCCA GTAAAATCCAGTTTTCACAGTAGAGGTGGT
JC4	TCCACAGGTGTCCAGCGGCCATGGGCGTGCCTGCCGTG
JC5	GCATGATGATGATGATGATGATGGCGGCCATCCCCCTGAAAGTACAGGTTCTCC TCGAGCAGCAGGATCCAGTTC
JC6	TGAACAGTTCCTCGCCCTTGCTCATGCGGCCATCCCCCTGAAAGTACAGGTTCTCCT CGAGCAGCAGGATCCAGTTC
JC7	GATCCATATGGGCGTGCCGGCCGTCTTA
JC8	GATCCTCGAGTCACAGTAAAATCCAGTTTTCACAGTAGAGGTGGT
JC9	GATCCATATGGGAGTTCCCGCTGTATTAAAAAGG
JC10	GATCCTCGAGTCACAGCAGAATCCAGTTCTCGCA
JC11	GATCCATATGAACAGCAGGGCATACCTC
JC12	TCCACAGGTGTCCAGCGGCCATGAACAGCAGGGCATACCTC
JC13	GCATCACCCACGGCGAATTCGAGAACCTGTACTTTCAGGGGATGAACAGCAGGGCA TACCTCACCACCAAC
JC14	TGATTGGATCCAAGCTATCATCACAGTAAAATCCAGTTTTC
JC15	GCATCACCCACGGCGAATTCGAGAACCTGTACTTTCAGGGGATGGGCGTGCCGGCC GTCTTA
JC16	CCTCTCTTCGCCAGGTACCCGGGGATCCTCTAG
JC17	CTTTGGGGTCGAGAGCTCGAATTCACTGGCC
JC18	TGAATTCGAGCTCTCGACCCCAAAGGTTTGAAG
JC19	AGGCTTTTTCATTGCAGCGACATCAGCAGTG

JC20	TGATGTCGCTGCAATGAAAAAGCCTGAACTCAC
JC21	CCGTGAAGCACCGCTATTCCTTTGCCCTCGG
JC22	GCAAAGGAATAGCGGTGCTTCACGGAAAATAGAGAG
JC23	TCCCCGGGTACCTGGCGAAGAGAGGCGGGA
JC24	TTGTTCAATCATTGCAGCGACATCAGCAGTG
JC25	TGATGTCGCTGCAATGATTGAACAAGATGGATTG
JC26	CCGTGAAGCACCGTCAGAAGAACTCGTCAAG
JC27	GAGTTCTTCTGACGGTGCTTCACGGAAAATAGAGAG
JC28	AGGCTTTTTCATCGCTCCAGTCGTGAAAGTG
JC29	CACGACTGGAGCGATGAAAAAGCCTGAACTCAC
JC30	TTGTTCAATCATCGCTCCAGTCGTGAAAGTG
JC31	CACGACTGGAGCGATGATTGAACAAGATGGATTG
JC32	GATCAAGCTTAGGCGTGCCGGCCGTCTTAAA
JC33	GATCCTCGAGTCAGTAAAATCCAGTTTTCAC
JC34	GCGCAGTCTGTCATCTGCTG
JC35	GAAGATGTTGGCGACCTCGT
JC36	ATGAACGGCTACTCGTGGGCC
JC37	CTAAAGCTCTTCCAGTTCGTTGACG
JC38	ATGAAAAAGCCTGAACTCAC
JC39	CTATTCCTTTGCCCTCGGAC
JC40	GCCAAGCTCTTCAGCAATATCACGG
JC41	ATGATTGAACAAGATGGATT
JC42	TCAGAAGAACTCGTCAAGAA
JC43	CGATGCAATTCTAGACACAC
JC44	AAAATGGCCGCTTTTCTGGA
JC45	TTTGCTGGATTCTGGCTTTT
JC46	ATGTGCGGCATCTTCGGCTA
JC47	TCAGCTCGGCATCCGGGTCG
JC48	ATCTTTCGTGGGTACCCGGGGATCCTCTAG
JC49	TTGGAACACTGAGCTCGAATTCACTGGCC
JC50	ATTCGAGCTCAGTGTTCCAATCGCTCCC
JC51	GCTTTTTCATCCCGCTAAGTTGACAACAC
JC52	AAAGGAATAGAGAGGTGATAATGCTATTGC
JC53	CCCGGGTACCCACGAAAGATGTCACACC
sJC54	ACTTAGCGGGATGAAAAAGCCTGAACTCAC
JC55	TATCACCTCTCTATTCCTTTGCCCTCGG
JC56	GTTCAATCATCCCGCTAAGTTGACAACAC
JC57	ACTTAGCGGGATGATTGAACAAGATGGATTG
JC58	TATCACCTCTTCAGAAGAACTCGTCAAG
JC59	GTTCTTCTGAAGAGGTGATAATGCTATTGC
JC60	GACGCTTTTGGGTACCCGGGGATCCTCTAG
JC61	AAAAATAGTGGAGCTCGAATTCACTGGCC
JC62	ATTCGAGCTCCACTATTTTTATTTCTCCCCTTCG
JC63	GTTCAATCATGGCTTAGAGTAAGGCGGATATG
JC64	ACTCTAAGCCATGATTGAACAAGATGGATTG
JC65	CTCATCCGCATCAGAAGAACTCGTCAAG

JC66	GTTCTTCTGATGCGGATGAGTCGAAAAAAAAAAAG
JC67	CCCGGGTACCCAAAAGCGTCAAAAGTTCG
JC68	GCTTTTTCATGGCTTAGAGTAAGGCGGATATG
JC69	ACTCTAAGCCATGAAAAAGCCTGAACTCAC
JC70	CTCATCCGCACTATTCCTTTGCCCTCGG
JC71	AAAGGAATAGTGCGGATGAGTCGAAAAAAAAAAAG
JC72	ATGTCGTCACGAACGGGACC
JC73	TCAGCTCGGCGACACATTAAATTTTAAAAG
JC74	ATGGTCGGCCCTAACCTTGC
JC75	TTACTCCTTCTTCGACGTGCTGGTC
JC76	GGCAGATTGGCAGTCTTTGA
JC77	GAGATGCCACGCTGCTGAGA
JC78	TTCACAAGGACGGTACTTGC
JC79	AGTCTCTAGAAGGTCACTCGGTAAG
JC80	ACTCGGTCATCCCGCTAAGTTGACAACAC
JC81	ACTTAGCGGGATGACCGAGTACAAGCCCACGG
JC82	TATCACCTCTTCAGGCACCGGGCTTGCG
JC83	CGGTGCCTGAAGAGGTGATAATGCTATTGC
JC84	ACTCGGTCATGGCTTAGAGTAAGGCGGATATG
JC85	ACTCTAAGCCATGACCGAGTACAAGCCCACGG
JC86	CTCATCCGCATCAGGCACCGGGCTTGCG
JC87	CGGTGCCTGATGCGGATGAGTCGAAAAAAAAAAAG
JC88	ATGACCGAGTACAAGCCCAC
JC89	TCAGGCACCGGGCTTGCGGG
JC90	TCATCTACAATTTTCATTCCCATC
JC91	GAGCTCGAATTCACTGGC
JC92	TTGGGGATCGGGTACCCGGGGATCCTCTAG
JC93	ACGGTGCGATGAGCTCGAATTCACTGGCC
JC94	ATTCGAGCTCATCGCACCGTTCTCGGCTAC
JC95	CCCGGGTACCCGATCCCCAATGTCAGAG
JC96	ACATGTACCACCGCCTCGGTTACATTGTGT
JC97	GAAGACGCAGCACCTCTTTGTTGTCGTTGT
JC98	CTTGATGGTCGACGGAGTAAAAATTCTGCA
JC99	CGACTCCATTGTGTGCGTGGGTATCTCTTG
JC100	TGTCTAGAATTCATGAACAGCAGGG
JC101	GTTGCCTTGGAGTCGTAAATGGCTC
JC102	ACTCCAAGGCAACGCTGCAGAACAA
SMD220	GGCCGCTGGACACCTGTG
SMD221	GATGGCCGCCATCATCATC
SMD222	ATGAGCAAGGGCGAGGAAC
SMD249	TGATAGCTTGGATCCAATCAAC
SMD270	CGCTGGACACCTGTGGAGAGAAAG
SMD309	GAATTCGCCGTGGGTGAT
pGEcSeq F	ATCAACGGGACTTTCCAAAATG
pGEc Seq R1	AGGAGCAACATAGTTAAGAATACC
pGEc Seq R2	ATGAACTTCAGGGTCAGCTTG

pGEnSeq F	GGCATCAAGGCCAACTTCAAG
pGEnSeq R	CCAGGATTTATACAAGGAGGA

3.6.2 Plasmids.

Vectors used in this project were:

- pET15b: *E. coli* overexpression of constructs rTcFUT1-1.a, -2.a, -2.b, -3.a and -3.b. Different versions were used, all containing an N-terminal 6xHis tag, and some containing either MBP or GST additionally. Their schemes are in (Fig. 4.7).
- pGEc1, pGEc2 and pGEn2: for Expi293F eukaryotic overexpression of constructs rTcFUT1-1.b, -1.c, -4.a, -5.a and -5.b. Their schemes are in (Fig. 4.14) and (Fig. 4.16).
- pSecTAG2A: for Expi293F eukaryotic overexpression of secreted recombinant product from construct rTcFUT1-6.a (scheme in (Fig. 4.23)).
- pTREX-G418: for TcFUT1 overexpression in *T. cruzi*. Kind gift from Mode of Action group (Susan Wyllie, School of Life Sciences, University of Dundee).
- pUC19: backbone for drug resistance cassettes with homologous recombination flanks of different lengths used for gene deletion in *C. fasciculata.*

3.6.3 Purification of plasmid DNA and quantification of DNA concentration and purity.

Plasmid DNA transformed in NEB 5-α *E. coli* strain was purified and prepared from 5 ml culture (for miniprep) grown at 37ºC, 200 RPM shaking using Monarch Plasmid Miniprep Kit. For the large scape preparation 200 ml were grown as above, aliquoted in 5 ml/tube in 14 ml Falcon round-bottom tubes and sent for robot (Qiagen Biorobot 9600s) DNA sequencing and miniprep purification by the MRC PPU DNA Sequencing and Services in the School of Life Sciences, University of Dundee. The final purified DNA was eluted with 10 mM Tris-HCl, pH 9.0, 0.1 mM EDTA, or alternatively with autoclaved water when further experiments could be affected by the presence of chelating agent. DNA was quantified on a BioDrop spectrophotometer, through absorption at 260 nm (A_{260}) and purity through the ratio A_{260}/ A_{280} (nucleic acids to proteins) and A_{260}/ A_{230} (nucleic acids to organic compounds and chaotropic agents) obtained by the same device.

3.6.4 Polymerase chain reaction (PCR).

The PCRs were performed with *Taq* DNA polymerase (using Thermopol buffer) and Q5 High-Fidelity DNA polymerase in 25 or 50 µl reactions. The conditions and programs used are shown in (Table 3.3). The annealing temperature (Ta) depends on polymerase used and primers Tm and was calculated using the "Tm and Ta calculator NEB online tool" (https://tmcalculator.neb.com/#!/main). For elongation the time was chosen based on product length. Thermocycler used were a PTC-225 Gradient Thermal Cycler (MJ Research) and ProFlex PCR system (Thermo Fisher Scientific). The reactions were analysed by agarose gel electrophoresis, bands of the correct size were excised aided with an UV lamp, and PCR products purified as described in sections 3.6.5 and 3.6.7.

Table 3.3: 25 µl PCR reactions

A) Taq polymerase reaction

Component	**Amount (µl)**	**Final concentration**
10X Thermopol	2.5	1X
10 mM dNTPs	0.5	200 µM
10 µM Forward primer	0.5	0.2 µM
10 µM Reverse primer	0.5	0.2 µM
Template DNA	variable	<1,000 ng
Taq DNA polymerase	0.125	25 mU/µl PCR
Nuclease-free water	To 25 µl	

PCR conditions: 95ºC (30'', 2' if gDNA) – 30 to 33 x [95ºC (30'') – Ta (50'') - 68ºC (1'/Kb product)] - 68ºC (5') - 4ºC (∞).

B) Q5 polymerase reaction

Component	**Amount (µl)**	**Final concentration**
5X Q5 reaction buffer	5	1X
10 mM dNTPs	0.5	200 µM
10 µM Forward primer	1.25	0.5 µM
10 µM Reverse primer	1.25	0.5 µM
Template DNA	variable	<1,000 ng
Q5 High-Fidelity DNA polymerase	0.25	20 mU/µl PCR
5X Q5 High GC Enhancer (optional)	(5)	(1X)
Nuclease-free water	To 25 µl	

PCR conditions: 98ºC (30'', 2' if gDNA) – 30 to 35 x [98ºC (10'') – Ta (50'') - 72ºC (30''/Kb product)] - 72ºC (5') - 4ºC (∞).

3.6.5 PCR purification.

PCR products were purified with Monarch PCR & DNA Cleanup kits. The final purified DNA was eluted with 10 mM Tris-HCl, pH 9.0, 0.1 mM EDTA, or alternatively with autoclaved water when further experiments could be affected by the presence of chelating agent.

3.6.6 Agarose gel electrophoresis.

Agarose gel electrophoresis was used for DNA separation. For DNA analysis 1% w/v agarose gels in 1X TAE buffer (40 mM Tris-acetate, 1 mM EDTA) (50X solution obtained from Central Technical Services) containing 0.4 µg/ml ethidium bromide were used. For Southern blotting, 0.8% gels were used. DNA samples were run in the presence of 1X Gel Loading Dye Purple and a 1 kb DNA ladder (Promega) to estimate DNA fragment size. Gels were run at 80 V in 1X TAE using BioRad Mini-Sub Cell GT tanks connected to PowerPac 300 power packs, with the exception of agarose gels for Southern blotting, which were run at 40 V. The gels were imaged with a Gbox EF Gel Documentation System (Syngene) and GeneSnap version 7.12 software (Syngene).

3.6.7 DNA gel extraction.

Plasmid vectors or DNA fragments obtained from PCR products or endonuclease digestions were purified, after separation on agarose gel, using the Monarch DNA Gel Extraction kit. The final purified DNA was eluted with 10 mM Tris-HCl, pH 9.0, 0.1 mM EDTA, or alternatively with autoclaved water when further experiments could be affected by the presence of chelating agent.

3.6.8 Restriction endonuclease digestion.

Restriction enzymes (RE) were all purchased from NEB. DNA digestions were performed in 1X CutSmart buffer, following NEB recommendations for RE concentration required used in different digestions. Analytical digestions were performed for 2 h at 37ºC water bath. Cloning digestions, or digestions of large amounts of DNA were incubated for 16 h.

3.6.9 Plasmid dephosphorylation.

After RE digestions, plasmids were dephosphorylated in the 5'-ends to reduce re-circularization events. Plasmids were incubated with 5 units of Antarctic Phosphatase (NEB) in 1X Antarctic Phosphatase Reaction Buffer for 30 min at 37ºC, and then reaction was heat-inactivated by incubation at 80ºC for 2 min.

3.6.10 DNA ligation.

The concentrations of purified, digested vector and insert(s) were measured as defined before. The insert to vector ratios were calculated by the following formula: (ng of vector x kb size of insert) / kb size of vector x insert:vector ratio = ng of insert required. The 3-fold amount of this value (in ng) was the amount of insert used for a given amount of vector (as standard, 100 ng). As a negative control the reaction was performed in the absence of insert to assess the amount of background re-ligation. Plasmid and insert(s), at a 1:3 ratio, were incubated along with 1X T4 DNA Ligase Buffer and 400 U of T4 DNA Ligase for 30 min at RT, followed by heat-inactivation at 65ºC for 10 min. Ligation reaction was then chilled on ice and 1 – 5 µl used to transform NEB 5-α cells, as described in section 3.3.3.

3.6.11 Gibson assembly.

For Gibson assembly of DNA fragments amplified with primers designed with NEBuilder Assembly online tool (https://nebuilder.neb.com/#!/), 0.1 pmols of each fragment were used when performing 2-fragment assemblies (constructs rTcFUT1-1.b, -1.c, -4.a, -5.a and -5.b), and 0.05 pmols per fragment when performing 4-fragment assemblies (pUC19-based drug resistance cassettes for gene deletion in *C. fasciculata*). To calculate the nanograms needed of each fragment the following formula was used: (pmols used x fragment length x 650 Da) / 1,000 = ng needed. The volume corresponding to the ng needed of each fragment were mixed, along with 10 µl of 2X NEBuilder HiFi DNA Assembly Master Mix and sterile water until reach 20 µl. Fragments were diluted, when needed, to not increase final reaction volume of 20 µl, or if volumes were too minute to be pipetted accurately, same 2X master mix volume as the combined volume of all fragments was used. Reactions were incubated at 50ºC for 1 hour, chilled on ice and either kept at -20ºC, or used immediately for NEB 5-α transformation using 2 µl of assembled product, following manufacturer's recommendation.

3.6.12 Site-directed mutagenesis.

25µl reaction was set containing 1 ng template along with 1X Q5 Hot Start High-Fidelity 2X Master Mix, and 0.5 µM forward and reverse primers designed in NEBaseChanger online tool (https://nebasechanger.neb.com/) for site-directed mutagenesis of *HYGr* and *NEOr* cassettes for *CfLPG1* gene knock-out. Final volume was reached with sterile water. PCR formula was the following: 98ºC (30'') – 25 x [98ºC (10'') – 61ºC (30'') - 72ºC (30''/Kb product) - 72ºC (2') - 4ºC (∞). After it, reaction product was used for kinase, ligase and DpnI (KLD) treatment by mixing 1 µl of PCR product with 5 µl 2X KLD Reaction Buffer, 1 µl 10X KLD Enzyme Mix and 3 µl sterile water. This reaction was incubated at RT for 5 min and then 5 µl used to transform NEB 5-α cells, as defined previously.

3.6.13 DNA sequencing.

DNA sequencing was performed by MRC PPU DNA Sequencing and Services in the School of Life Sciences, University of Dundee using 3730XL DNA analyser (Applied Biosystems). The regions of interest in the plasmid DNA were sequenced two to four times in each direction. The consensus sequence was built using CLC Main Workbench version 8.0. Comparison between the consensus and the *in silico* sequences was performed in the same software.

3.6.14 Genomic DNA isolation from parasites.

Parasites were grown to end of log phase and 2 – 5 x 10^7 (*T. cruzi*) or 5 x 10^7 to 1 x 10^8 cells (*C. fasciculata*) were harvested by centrifugation for 10 min, 1,000 x g (*C. fasciculata*) or 800 x g (*T. cruzi*), RT (Sigma Laboratory Centrifuge 4K15). The pellet was resuspended in residual medium, transferred to Eppendorf tube, and centrifuged (RT, 1,000 or 800 x g, 5 min) (Eppendorf Minispin Plus for *T. cruzi* and Eppendorf centrifuge 5418 for *C. fasciculata*). 200 µl lysis buffer (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 25 mM EDTA, 0.5% SDS, 0.1 mg/ml proteinase K in 10 mM Tris-HCl pH 8.0) were added to pelleted cells. At this point, *T. cruzi* lysed cells could be taken out of BSL3 culture room. Then, sample was incubated at 56ºC for 16 h to complete lysis, followed by addition of 0.5 ml of pre-cooled absolute ethanol. Cell lysate was centrifuged at RT for 20 min, 16,000 x g (Eppendorf microcentrifuge 5415R), supernatant discarded, and pellet washed with 70% ethanol followed by a similar centrifugation for 5 min. After discarding the wash, the pelleted genomic DNA was left to air-dry for ~30 min before redissolving in sterile water.

3.6.15 Ethanol precipitation of DNA.

This protocol was used to concentrate or store plasmid and genomic DNA. For each volume of solubilized DNA, 1/10 of the volume of 3 M sodium acetate pH 5.2 was added, followed by 3 times the volume of DNA and sodium acetate of cold absolute ethanol. Mix was incubated at -20ºC for at least 16 h, or at -80ºC for at least 2 h. To resuspended precipitated DNA, the mix was centrifuged at 4ºC for 10 min at 16,000 x g (Eppendorf microcentrifuge 5415R). Supernatant was decanted and pellet washed 3 times with cold 70% ethanol spinning under same conditions as before, but for 5 min. After the last wash, pellet was left to air-dry for approx. 30 min before resuspending in sterile water.

3.6.16 DIG-labelled DNA probe synthesis by PCR.

The DNA probes used in this study were digoxigenin (DIG)-labelled. These were generated using PCR DIG Probe Synthesis Kit (Roche) according to the protocol summarised in (Table 3.4). For each probe, a reaction which uses normal dNTP solution was included as a control to indicate the specificity and efficiency of the DIG-labelled probe synthesis. Aliquots of the reactions (10 µl) were analysed by agarose gel electrophoresis to check the size increase in the labelled probe and estimate the amount. When a bright band was observed, 15 µl of probe/25 ml of hybridization buffer were used. When band was faint, 30 – 40 µl were used for hybridization. Before mixing with hybridization solution, probe was denatured by heating at 100ºC for 5 min.

Table 3.4: Amplification of DIG-labelled probes.

Reagent	**Volume DIG-labelled probe (µl)**	**Volume unlabelled control probe (µl)**	**Final quantity**
10× PCR buffer with $MgCl_2$	5	5	1X
PCR DIG probe synthesis mix	5	-	200 µM dATP, dCTP, dGTP, 130 µM dTTP, 70 µM DIG-dUTP
dNTP stock solution	-	5	200 µM each dNTP
10 µM forward primer	5	5	1 µM
10 µM reverse primer	5	5	1 µM
Enzyme mix	0.75	0.75	
DNA template (100 pg/µl)	1	1	100 pg

PCR conditions: 95ºC (2') – 30 x [95ºC (30'') – 60ºC (30'') - 72ºC (40'')] - 72ºC (7') - 4ºC (∞).

3.6.17 Southern blotting.

For genotyping clonal populations of *C. fasciculata* during the gene deletion attempts, 5 µg of genomic DNA were digested at 16 h with restriction enzymes and run on a 0.8% agarose gel. Then, the gel was prepared for transfer by washing it for 10 min in 0.2 M HCl, then for 15 min in 1.5 M NaCl, 0.5 M NaOH, and then for 20 min in 1 M Tris-HCl, 1.5 M NaCl pH 7.5. All washes were performed at RT, on a glass tray under gentle shaking (Stuart mini gyro-rocker SSM3). The DNA was then transferred on a positively charged nylon membrane (Roche) by reverse capillary action for 24 h with 10X sodium chloride, sodium citrate buffer (SSC) buffer. 1X SSC = 150 mM NaCl, 15 mM sodium citrate, 0.2 µm filtered. Transferred DNA fragments were covalently cross-linked to the membrane by UV cross-linking in a CL-100 (UVP) UV crosslinker at 1200 mJoules. Then, membrane was placed on a roller bottle (Techne) and pre-hybridised in EasyHyb solution (Roche) at 42ºC for 1 hour inside hybridization oven (Techne). Membrane was then hybridized with PCR-synthesized, DIG-labelled DNA denatured probe diluted as described previously in EasyHyb solution by incubating for 16 h at 60ºC. If needed to optimize this step, optimal hybridization temperature (T_{hybopt}) was calculated following the formula: T_{hybopt} = theoretical hybridization temperature (T_{hybthe}) – 20 to 25ºC, where T_{hybthe} = 49.82 + 0.41 x (% G + C of probe) – (600/length of probe in base pairs). After hybridization, membrane was washed twice for 5 min in 1X SSC, 0.1% SDS at 42 to 45ºC, and twice in 0.5X SSC, 0.1% SDS at 65ºC. Following steps were performed at RT using buffers generated from stocks of Roche DIG Wash and Block Buffer Set. The membrane was equilibrated for 5 min at RT in 1X wash buffer before being blocked for 30 min in 1X blocking buffer. Then, it was incubated for another 30 min with a 1:10,000 dilution in blocking buffer of Anti-DIG AP-conjugate antibody (Roche), followed by two 15 min washing steps in 1X wash buffer, and 5 min equilibration step in 1X detection buffer. At this point, membrane was taken from the roller bottle and placed inside a plastic folder, to then apply dropwise over it CSPD detection reagent (Roche). Plastic folder was closed, excess solution and bubbles removed with gentle pressure, and the membrane inside the folder was incubated for 5 min at RT, followed by 10 min in a 37ºC incubator, always covered from light. Closed plastic folder with membrane inside was placed in a exposure cassette and expose in dark to a high performance chemiluminescence film

(Cytiva Amersham Hyperfilm ECL) for different exposure times (normally 10, 20 and 30 min) to then develop films with a Protec ECOMAX X Ray Film Processor in a dark room. For stripping the blot, 0.4 M NaOH, 0.1% SDS was applied to the membrane and washed twice for 5 min at 42 °C in the hybridisation oven. The blot was then washed three times with 5 x SSC buffer for 10 min at 42 °C before re-probing. Membrane was kept in 5X SSC at 4ºC in glass tray.

3.7 Biochemical protocols

3.7.1 Extraction and purification of lipoarabinogalactan from *C. fasciculata*.

C. fasciculata cells (1.5 - 2 x 10^9 cells) were harvested in stationary phase by centrifugation (5,525 x g, 10 min, 4ºC) (Sigma Laboratory Centrifuge 4K15), washed twice with PBS in same conditions, and then extracted by incubation with chloroform:methanol:water (C:M:W, 1:2:0.8) considering the weight of the pellet as part of the water fraction, sonicated for 30 min and incubated for 3 h at RT. The extract was then centrifuged (as previously), and chloroform supernatant removed. The pellet was extracted again, this time adding pre-mixed solution C:M:W (v:v:v), sonicating for 30 min and incubating for 4 h at RT to finally centrifuge again (as previously), and remove chloroform. Then , the delipidated pellet was dried under a N_2 steam, extracted by adding water saturated with 1-butanol (9%), sonicated for 30 min and centrifuged (as previously) to then transfer supernatant to a microcentrifuge tube. This process was repeated, both supernatants combined and then span at 16,000 x g for 10 min, RT in and Eppendorf centrifuge 5418 to remove insoluble material. The clean supernatant was then brought to 0.1 M ammonium acetate (NH_4Ac). The sample was used for hydrophobic interaction chromatography (HIC) by loading into a homemade column of 0.5 ml octyl-Sepharose CL-4B (Sigma), pre-equilibrated with 2.5 ml of 5% 1-propanol in 0.1 M NH_4Ac and 50% 1-propanol, used sequentially twice, and then the first solution once. Sample was loaded onto the equilibrated column and followed with 1 column volume of 5% 1-propanol in NH_4Ac, and then washed with 6 column volumes of 5% propanol in NH_4Ac, 3 column volumes of 5% 1-propanol in water (v:v) and eluted with 10 column volumes of 50% 1-propanol in water. Samples of 5% and 10% of the eluate containing LAG, were used for periodate-Schiff (PAS) staining (section 3.7.5), and 15% for GC-MS composition analyses (section 3.7.2).

3.7.2 Carbohydrate composition analysis by GC-MS.

15% of the eluate containing LAG was dried in vacuum centrifuge at 40ºC. 30 µl of 0.1 mM *scyllo*-inositol and 100 µl of 50% 1-propanol were added to the dried material. 40 µl aliquots of this mix were added to 3 sterilized glass capillaries (Hilgenberg) previously flame sealed in a Bunsen burner. In 2 other glass capillary, 20 µl of a premade mix of 20 µl MiliQ water and 20 µl of 0.1 mM *scyllo*-inositol were added, so each blank capillary contained 1 nmole *scyllo*-inositol. Finally, in 3 glass capillaries, 20 µl of a homemade sugar standard mix were added. This sugar mix contains 0.05 mM of D-ribose, L-rhamnose, D-Ara, L-Fuc, D-xylose, D-mannose (D-Man), D-galactose (D-Gal), D-glucose (D-Glc), *scyllo*-inositol, N-acetyl-D-galactosamine (GalNAc), N-acetyl-D-glucosamine (GlcNAc) and N-acetylneuraminic acid in water. The tubes were dried in a vacuum centrifuge at 40ºC to concentrate the samples, after which 20 µl of methanol were added to each tube, to be dried again, in order to completely dehydrate the samples. To these, 50 µl of 0.5 M HCl in methanol (Merck) were added, the tubes briefly centrifuged to avoid any trace in their walls, and then flame sealed in a Bunsen burner under slight vacuum, using a homemade device. The sealed tubes were incubated at 85ºC for 4 h in order to perform acid methanolysis in the samples. Tubes then were opened with a glass knife, and 10 µl of pyridine added to neutralize the HCl. 10 µl of acetic anhydride were also added to re N-acetylate (by incubation at RT for 30 min) any free primary amine which could have lost an acetate group during acid methanolysis. After this, tubes were dried again, sealed with Teflon tape and left at RT for 10 min. During this time, fresh trimethylsilyl (TMS) reagent was made by mixing 500 µl pyridine, 150 µl hexamethyldisilazane and 50 µl chlorotrimethylsilane (10:3:1) in a glass vial followed by centrifugation at 1,000 x g, 15 min, 4ºC (Sigma Laboratory Centrifuge 4K15) and storge over desiccant at 4ºC. 15 µl of TMS reagent was then added to each sample sequentially, taping glass capillary with Teflon tape, incubated for 30 min at RT and then using 1 µl to inject it into the GC-MS system (Agilent Technologies). The separation was performed using as stationary phase a silica-based capillary column (30m x 250µm x 0.25µm, Agilent), helium steam as the mobile phase at 0.7 ml/min, and a temperature-based gradient from 80ºC to 260ºC. The GC-MS programme used set an injector temperature of 280ºC and the temperature in interface between the GC and MS is held at the same value. Column pressure was 5 psi. Gradient parameters were: initial temperature 80ºC; hold 2 min; 30ºC/min to 140ºC; 10ºC/min to 200ºC; 2.5ºC/min to 260ºC; 20ºC/min to 280ºC, hold 10 min.

3.7.3 SDS-PAGE.

Sodium dodecyl sulphate polyacrylamide gel electrophoresis was performed using pre-cast NuPAGE 10%, 12% and 4-12% Bis-Tris gels with MOPS running buffer (50 mM (3-(N-morpholino)propanesulfonic acid, 50 mM Tris base, 0.1% SDS, 1 mM EDTA, pH 7.7) in a XCell SureLock mini tank (Thermo Fisher Scientific) at 200 V using a BioRad PowerPac 300. Samples were prepared using either 4X NuPAGE LDS sample buffer (Invitrogen) or 4X SDS homemade sample buffer (986 mM Tris base, 2 mM EDTA, SDS 4% (w/v), Orange G 0.2% (w/v), glycerol 40% (w/v) pH 8.5) and a final concentration of 0.1 M DTT as reducing agent. Before loading, samples were boiled at 95ºC for 10 min or heated at 50-55ºC for 20 min.

3.7.4 Coomassie staining.

Proteins present on SDS-PAGE (section 3.7.3) were stained by incubating for 2 h with Quick Coomassie stain (Neo Biotech) in gentle shaking (Stuart mini gyro-rocker SSM3), and afterwards removing the excess stain by incubating for ~30 min in sterile water in gentle shaking. For staining prior to fluorography, the gel was stained with Coomassie Blue R250 (0.15% w/v in water:methanol:acetic acid 4:4:1 v:v:v) for 20 min shaking. Excess stain was removed by incubating in water:ethanol:acetic acid 4:4:1 v:v:v solution for 12 – 16 h shaking. After destain, the gel was rehydrated with 5 quick washes with sterile water.

3.7.5 Periodic acid-Schiff (PAS) staining.

Either 5% or 10% of eluted volume containing LAG after HIC (section 3.7.2) were dried in vacuum centrifuge, resuspended in 30 µl 1X NuPAGE LDS sample buffer and used for SDS-PAGE (section 3.7.3). Subsequently, the gel was rinsed in water, fixed for 25 min in acetic acid:methanol:water (10:35:25), washed 3 times with water, oxidized for 30 min in 1% $NaIO_4$, washed 3 times, incubated for 1 hour with Schiff reagent (Sigma), and then reduced with 1% $Na_2S_2O_5$ for 30 min and finally washed 3 times with water. All steps were done in gentle shaking.

3.7.6 LICOR Western blotting.

For the LICOR system, samples were prepared using 4X SDS homemade sample buffer, since it does not contain blue dye which otherwise would be detected by the scanning system. After SDS-PAGE (section 3.7.3), proteins were transferred on a nitrocellulose membrane (Invitrogen iBlot™ 2 Transfer Stacks) using a iBlo2-Dry Blotting System (Invitrogen) on a 7 min, 25 V programme. Afterward, membrane was

incubated for 5-15 min in gentle shaking with 0.1% Ponceau S (w/v) in 5% acetic acid (Sigma) to verify successful transfer of proteins and correct loading. Next steps were performed with the different components of SNAP i.d. 2.0 Protein Detection System (Merck). The membrane was place inside a blot holder and this inside a plastic frame on top of a base unit connected to a vacuum pump. 30 ml of blocking solution (50 mM Tris-Hcl pH 7.4, 150 mM NaCl, 0.25% bovine serum albumin (BSA), 0.05% (w/v) Tween-20, 0.05% NaN_3, 2 % (w/v) fish skin gelatine (FSG)) were poured on the closed plastic frame and passed through the membrane by vacuum aspiration to block the membrane. Primary antibody was diluted in blocking buffer and incubated with the membrane for either 10 min pouring the solution on the framed membrane and covering from light, or for 60 – 90 min at RT rotating inside a 50 ml Falcon tube. After the incubation, the primary antibody solution was decanted or vacuum aspirated. When needed, membrane was placed back into a holder and this into the frame. 3 PBST (1X PBS, 0.1% Tween-20) washes were performed with continuous vacuum aspiration and then LICOR secondary antibody (IRDye 800CW for green signal or IRDye 680RD for red signal), also diluted in blocking buffer, was added to the framed membrane, and left incubating for 10 min covered from light. When using IRDye 680RD antibodies, 0.01% SDS was added to the solution in order to reduce the background. After incubation, secondary antibody solution was aspirated, and 3 more washes were performed as before. The imagining systems used to detect the fluorescent signal from secondary antibodies were Odyssey Sa and Odyssey CLx scanners (LICOR). To analyse the scanned images the software Image Studio Lite version 5.2 was used.

3.7.7 [^{3}H]Fuc biosynthetic labelling of *C. fasciculata*.

A depleted SDM-79 medium (no-sugar medium) was obtained by adding myo-inositol (1.33 mg/l), L-hydroxyproline (2 mg/l), L-proline (615 mg/l), L-serine (71.4 mg/l), L-methionine (89 mg/l), L-lysine (91 mg/l), L-arginine (248 mg/l), GlutaMAX (1X), 7.5 µg/ml hemin in 50 mM NaOH, and 2 g/l $NaHCO_3$ to a depleted SDM-79 medium (Caisson Labs) without neither of these components nor D-Glc (1.86 g/l) or D-glucosamine (GlcN) (50 mg/l). After preparation, pH was adjusted to 7.3 with NaOH. Prior to be used, the no-sugar medium was supplemented with penicillin-streptomycin to a final concentration of 10 U/ml and 10 µg/ml, respectively. Medium was not sterile-filtered at any time to avoid loss of any nutrient that could hinder the effectivity of the labelling.

C. fasciculata cells were grown to 4 – 6 x 10^6 cells/ml, i.e., in the middle of log phase, when cells are biosynthetically very active. Cells were harvested by centrifugation for 10 min, 800 x g, RT (Sigma Laboratory Centrifuge 4K15). The pellet was washed three times with no-sugar medium (same centrifugation conditions as above) to then resuspend the washed cells to 3 x 10^7 cells/ml in the same medium. 1.5 x 10^7 cells (500 µl) were transferred to a 15 ml Falcon tube in which previously, 5 µCi of [^{3}H]Fuc in 70 or 90% EtOH (American Radiolabeled Chemicals) were dried under gentle gas nitrogen steam placing the tube in an appropriate holder on a heating module (Thermo Fisher Scientific Reacti-Therm) with an evaporator containing PTFE-coated needles (Thermo Fisher Scientific Reacti-Vap) connected to the nitrogen gas line. Then, cells were incubated at 28ºC for 4-5 h with gentle shaking (300-400 RPM) (Grant PMS-1000i Microplate Shaker) in the 15 ml Falcon tube with lid closed. After labelling, cell pellet is recovered by spinning at 16,000 x g, 30-40 secs, 4ºC (Eppendorf microcentrifuge 5415R), and used for lysis and different treatments before being prepared for SDS-PAGE (section 3.7.3).

3.7.8 Chemical/enzymatic treatment of [^{3}H]-labelled *C. fasciculata* lysates.

Pellets of labelled *C. fasciculata* cells were processed differently depending on the treatment to be performed:

- Phosphatidylinositol-specific phospholipase C (PI-PLC): pellet of 1.5 x 10^7 labelled cells was resuspended in PI-PLC digestion buffer (20 mM Tris-HCl pH 7.4, 0.1% Triton X-100). 5 x 10^6 cells of the suspension were then incubated with or without 2 U of *Bacillus thuringiensis* PI-PLC (gift from Dr. Martin Low, Columbia University, New York, USA) for 1 h at RT and then from 2 h to O/N at 37ºC.
- α1-2 exofucosidase: pellet of 1.5 x 10^7 labelled cells was resuspended in commercial 1X Glycobuffer I (5 mM $CaCl_2$, 50 mM sodium acetate pH 5.5) supplemented with 0.1% Triton X-100. 5 x 10^6 cells of the suspension were then incubated with or without 20 U of α1-2 exofucosidase for 1 h at RT and 2 h at 37ºC, and then 40 U of enzyme were added for an O/N incubation at 37ºC.
- 40 mM trifluoroacetic acid (TFA): pellet of 7.5 x 10^6 labelled cells was resuspended in residual medium to then snap-freeze in dry ice before freeze-drying them in a lyophilizer for 1 h (Christ Alpha 1-2 LD plus). After it, dried material was resuspended in 200 µl of freshly prepared 40 mM TFA in sterile water, or just water for the negative control, and incubated at 90ºC for 18 min.

After treatment, tubes were cooled in ice, snap-frozen in dry ice and freeze-dried O/N in lyophilizer.

- 50% aqueous hydrogen fluoride (aq. HF): pellet of 7.5 x 10^6 labelled cells was resuspended in residual medium to then snap-freeze in dry ice before freeze-drying it in a lyophilizer (Christ Alpha 1-2 LD plus). After it, dried material was kept for at least 1 h at -20ºC, to then be resuspended in 50 µl ice-cold 48% aq. HF solution (hydrofluoric acid), or just water for the negative control, and incubated at 0ºC by placing the tubes on ice/water slurry inside a Dewar flask for 26 to 48 h. After treatment, tubes are snap-freeze in dry ice and freeze-dried O/N in lyophilizer.

After treatments, samples from PI-PLC and fucosidase treatments were prepared for SDS-PAGE (section 3.7.3) by mixing with equal volume of 2X homemade sample buffer and 0.2 M DTT, or resuspended in 30 µl of 1X buffer and 0.1 M DTT for samples treated with TFA and aqueous HF.

3.7.9 Fluorography.

Due to the weak β particles emitted by the [^{3}H] isotope, an indirect autoradiography, known as fluorography, was necessary to be performed in order to visualize in a photographic emulsion (X-ray film) the different [^{3}H]-labelled species present in the HPTLC plates and SDS-PAGE gels generated during this project. To achieve this, samples had to be impregnated with a scintillator, being the processing different depending on the platform used in each experiment.

- High-performance thin layer chromatography (HPTLC) plate: after chromatographic run, silica plates were fluorographed by spraying 3 times with En3Hance Spray scintillation solution (PerkinElmer), leaving it dry between each spray for 15 min, and place inside an exposure cassette against an X-ray film (Sigma Carestream Biomax XAR film) and an intensifying screen (Kodak). Exposure cassette was left at -80ºC for few days to up to 2 months. The films were developed on a Protec ECOMAX X Ray Film Processor in a dark room.
- SDS-PAGE gel: gels were stained with Coomassie blue R-250 (section 3.7.4), to then be incubated shaking for 30 min with En^{3}Hance scintillation solution. The gel was rinsed with water 3 times until gel became milky, indicating the precipitation of scintillant inside the gel. Then, it was soaked in 10% glycerol for 30 min shaking and later placed on top of two stacked pieces of Whatman paper

onto the porous gel support of a gel dryer (BioRad Gel Dryer Model 583). Gel was dried under vacuum (BioRad Hydrotech Vacuum Pump) for 2 h in a gradient heating cycle from RT up to 80ºC. After it, dried gel attached to upper filter paper was placed inside an exposure cassette against an X-ray film (Sigma Carestream Biomax XAR film) and an intensifying screen (Kodak). Exposure cassette was left at -80ºC for few days to up to 2 months. The films were developed on a Protec ECOMAX X Ray Film Processor in a dark room.

3.7.10 GDP-[^{3}H]Fuc synthesis.

C. fasciculata cells (1 x 10^9 cells) were harvested in late log phase by centrifugation (1,000 x g, 10 min, 4ºC) (Sigma Laboratory Centrifuge 4K15), washed twice with 1X PBS and then resuspended in 600 µl 10 mM Tris-HCl pH 7.5, 1X protease inhibitor cocktail (Merck cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail). Resuspended cells were lysed in two cycles of freeze-thawing by placing in dry ice for 15 min and then holding tubes under cold water stream until material was thawed again. Lysate was then centrifuged (16,000 x g, 10 min, 4ºC) (Eppendorf microcentrifuge 5415R), the supernatant centrifuged again in same conditions, and then soluble fraction was transferred to a 1 ml Leur slip syringe with a 21G-wide gauge needed. Needle was replaced with a 0.45 µm Millex-HV PVDF syringe filter (Merck) and soluble fraction was filter-expelled into a microcentrifuge tube, to then be separated into two samples of same volume placed inside two 100 kDa, 0.5 ml capacity, diafiltration/concentration devices (Thermo Fisher Scientific), previously prepared as defined by manufacturer. These were centrifuged (12,000 x g, 7.5 min, 4ºC), flow-through (FT) discarded, and 400 µl of ice-cold 10 mM Tris-HCl pH 7.2, 1X protease inhibitor cocktail were added. Then, four consecutive centrifugations (12,000 x g, 10 min, 4ºC) were performed in order to concentrate and diafiltrate the samples, discarding the FT after each one. After first centrifugation, 400 µl of ice-cold 10 mM Tris-HCl pH 7.2, 1X protease inhibitor cocktail were added. After second one, 400 µl of ice-cold 10 mM Tris-HCl pH 7.2 were added. After third one, 400 µl of 10 mM Tris-HCl pH 7.2, 20 mM NaF were added. After fourth one, samples were recovered and pooled along with 2 x 100 µl washes with NaF-containing buffer of both devices. The protein content of this crude enzymatic preparation was measured by the Christian-Warburg method in a BioDrop spectrophotometer, and it always indicated a concentration of 1.5 - 2 µg/µl. This activity source for the synthesis of GDP-[^{3}H]Fuc was either used immediately, kept for no longer than 24 h at 4ºC before use, or supplemented with glycerol for a final concentration of 10% and snap-frozen in

liquid nitrogen, to then be kept at -20ºC until use. Later experiments proved that activity was kept for at least 3 weeks being stored in this way.

For small-scale reactions, 1 µCi [^{3}H]Fuc was dried in a PCR tube with gentle N2 stream. Then, 2.5 µl of 20 mM $MgSO_4$, 8 mM ATP, 8 mM GTP dissolved in 10 mM Tris-HCl pH 7.2 were added to resuspend the material. In parallel, 1 µl of 400 mU/µl of yeast inorganic pyrophosphatase was added to 9 µl of crude enzymatic preparation. Of this mix, 2.5 µl were added to the resuspended [^{3}H]Fuc for a total volume of 5 µl in the PCR tube. Reaction was incubated at 30ºC in a PTC-225 Gradient Thermal Cycler (MJ Research) and 0.8 µl aliquots were taken at different times to analyse the reaction. To these aliquots, 1 µl 100 mM EDTA was added to stop the reaction, 1 µl of a mix containing 1 µg/µl of fucose, fucose-1-phosphate and GDP-fucose was also added to act as carriers to aid the chromatography, and 2 µl of 1-propanol were added to facilitate the absorption of the sample by the silica plate. The time point aliquots were analysed by HPTLC as define below.

For large-scale reactions, 45-50 µCi [^{3}H]Fuc were dried in an Eppendorf tube in similar manner as in section 3.7.7. Once all the radioactive material was dried, 10 µl of 20 mM $MgSO_4$, 8 mM ATP, 8 mM GTP dissolved in 10 mM Tris-HCl pH 7.2 were added to resuspend the material. In parallel, 2 µl of 400 mU/µl of yeast inorganic pyrophosphatase were added to 18 µl of crude enzymatic preparation. Of this mix, 10 µl were added to the resuspended [^{3}H]Fuc for a total volume of 20 µl in the PCR tube. Reaction was incubated in a thermoblock at 30ºC for 3 h (shorter times were tested but had lower yields) to then stop the reaction by adding 380 µl of 2 mM EDTA.

3.7.11 Recovery of GDP-[^{3}H]Fuc and [^{3}H]Fuc by weak anion exchange solid phase extraction (WAX-SPE)

To recover the GDP-[^{3}H]Fuc generated after the synthesis reaction, and the [^{3}H]Fuc that was not converted by the system a solid-phase extraction was performed, using a weak anion exchanger resin as stationary phase. A 1 ml WAX cartridge with 30 mg sorbent (Waters) was prepared by passing through 0.5 ml sterile water, 4 times 0.5 ml of 1 M ammonium acetate and 4 times 0.5 ml sterile water. Stopped GDP-[^{3}H]Fuc synthesis reaction by the addition of EDTA was passed through the cartridge and eluate recovered, along with two washes of 150 µl sterile water. This was the FT fraction. After it, 3 washes (W1 - W3) with 0.5 ml sterile water were done and collected in separate tubes. Finally, 4 elutions (E1 - E4) with 0.5 ml 1 M ammonium acetate were also

collected. The eight obtained fractions were counted (see liquid scintillation counting), their volume estimated, and their radioactive concentration calculated. [^{3}H]Fuc appeared in the FT fraction since being an uncharged molecule it did not interact with the stationary phase. Synthesized GDP-[^{3}H]Fuc appeared in fraction E1 in all cases, although a minimal portion appeared in fraction E2 in one of the experiments.

Once recovered, fractions FT and E1 were processed in order to desalt and resuspend them prior to use. Fraction E1 was snap-frozen in dry ice and then freeze-dried O/N in lyophilizer. Dried material was resuspended in 100 µl sterile water and freeze-drying repeated for 3 h. After it, double-lyophilized material not containing any traces of ammonium acetate was redissolved in 50% ethanol. For fraction FT, sample was passed through a mixed-bed ion exchange resins to be desalted (see Radioactive-based fucosyltransferase activity assay), followed by three washes of 0.4 ml sterile water. The eluate was then separated in microcentrifuge tubes and concentrated in a centrifugal vacuum dryer (Christ RVC 2-25 CD plus) at 50ºC for 3 - 4 h. Dried material was recovered by washing the tubes sequentially with two washes of 70% ethanol and pooling the washes in a single tube. Ready-to-use GDP-[^{3}H]Fuc and [^{3}H]Fuco were kept at -20ºC.

3.7.12 [^{3}H]-based fucosyltransferase activity assay.

As enzymatic source, either 1 - 2 µg of affinity-purified recombinant TcFUT1 (for constructs rTcFUT1-2.b and rTcFUT1-4.a), or α-Myc agarose beads slurry incubated with medium of 1.4 x 10^7 Expi293F cells (for construct rTcFUT1-6.a) or lysate of 4 x 10^8 *T. cruzi* cells (for construct TcFUT1-7.a) were used. These were incubated with 1 - 0.3 µCi of commercial (American Radiolabeled Chemicals) or home-made GDP-[^{3}H]Fuc and 1 mM synthetic saccharide acceptor (Dextra Laboratories and Toronto Research Chemicals) in 50 mM Tris-HCl, 25 mM KCl, pH 7.2 to a final volume of 25 µl for either 2 h at 37ºC when using purified protein, or O/N at RT shaking vigorously (SI Vortex Genie 2) to maintain beads-protein complexes in suspension. A reaction without acceptor and without acceptor and enzyme were run as negative controls. At the end of the incubation, reactions were cooled on ice and sterile water was added to 200 µl final volume. The reactions were desalted on a mixed-bed column made of 100 µl each of 100% slurry Chelex 100, AG 50W X12, AG 4 (BioRad) and QAE-Sephadex A25 (Merck) from top to bottom. The column was pre-washed with 10 column volumes of sterile water, the samples loaded (in case of reactions containing beads, these were first pelleted by centrifugation at 2,500 x g, RT, 10 min and supernatant passed through column) and the glycans eluted with 4 times 400 µl of water. The FT and the elutions were combined. 5%

of the eluates was counted (see liquid scintillation counting), and 30% was concentrated and dried in in a centrifugal vacuum dryer (Christ RVC 2-25 CD plus) at 40ºC for 4 - 5 h. Dried material was resuspended in 150 µl of sterile water and vacuum-dried again 40ºC for 2 - 3 h. After it, samples were redissolved in 20% 1-propanol and analysed by TLC (see TLC analysis), together with non-radioactive (cold) and radioactive standards.

3.7.13 Liquid scintillation counting.

To quantify the radioactivity of [^{3}H]-containing samples, 1 µl was mixed with 2 ml of Opti-Fluor (PerkinElmer) liquid scintillation cocktail in a 5 ml plastic tube. Samples were then counted in a LS 6500 liquid scintillation counter (Beckman) or a Tri-Carb 4910TR 110 V liquid scintillation counter (PerkinElmer) using 1 µl sterile water, or the appropriate solvent used on the samples, to be counted as blank. The counts per minute (CPM) output from the Beckman counter was converted into nCi considering the counting efficiency of the machine to be 50%, i.e., 1,110 CPM being equal to 1 nCi. In the case of the PerkinElmer counter, a disintegrations per minute (DPM) output was obtained from the QuantaSmart version 5.01 software (PerkinElmer), which was converted into nCi considering 1 nCi [^{3}H] isotope being equal to 2,220 DPM.

3.7.14 GDP-glo activity assay.

The commercial GDP-Glo Glycosyltransferase Assay (Promega) was used. Recombinant TcFUT1 (purified, present in Expi293F expression medium or immunoprecipitated with α-Myc agarose beads) or purified TbFUT1 (Bandini *et al.*, 2021) was incubated along with 50 µM Ultra-Pure GDP-Fuc (Promega) with or without 1 mM synthetic sugar acceptor in 50 mM Tris-HCl, 25 mM KCl, pH 7.2 to a final volume of 25 µl, in a solid white, flat bottom 96-well plate (Corning), covered from light inside a humid box O/N at RT. After it, 25 µl of freshly prepared GDP detection reagent (generated following manufacturer's manual) was added to stop the reaction, and incubated covered from light inside a humid box for 1 h at RT. At the same time, a GDP standard curve (from 25 µM to 20 nM) was generated, following manufacturer's manual and using GDP solution (Promega) diluted in same buffer as the reaction, to which GDP detection reagent was also added and incubated as described before. Then, luminescence from the wells was quantified in a Pherastar FS luminometer (BMG Labtech) using an integration time for photon measurement of 0.5 secs/well. Hydrolysed GDP generated in the reaction was inferred from the GDP standard curve. All reactions and standard curve were made in triplicates, and two measurements were performed for each plate.

3.7.15 Thin layer chromatography (TLC) analysis.

TLC was performed on 10 cm height HPTLC silica gel 60 plate (Merck) with 1-propanol:acetone:H_2O 9:6:4 (v:v:v) as mobile phase unless otherwise stated in Results. All solvents used were AnalaR NORMAPUR grade (VWR). To improve resolution, in some cases plates were run twice before imaging. Non-radioactive sugars were visualized by orcinol/H_2SO_4 staining (section 3.7.16). In the case of radiolabelled material, the TLC plates were fluorographed (section 3.7.9).

3.7.16 Orcinol/H_2SO_4 staining.

To stain non-radiolabelled carbohydrates on the HPTLC plates, these were sprayed using an Aldrich flask-type sprayer (Merck) lined to a N2 tank with an orcinol/H_2SO_4 solution, left drying for 10 min, and then heated with a heat gun (Steinel HL1610 S) until the spots appeared. The staining solution was prepared by dissolving 180 mg of orcinol (Merck) in 5 ml of water, and then adding 75 ml absolute ethanol. After placing solution on ice, 10 ml concentrated H_2SO_4 (18.4 M) were added dropwise and carefully mixed in. Final solution was kept covered from light and at 4ºC and brought to RT before using.

3.8 Immunofluorescence.

3.8.1 Cell fixing, staining, and imaging.

Primary antibodies used in the immunofluorescence analyses (IFAs) were rabbit polyclonal α-TbFUT1 (1:1,000) (Bandini *et al.*, 2021) and mouse monoclonal α-Myc tag (1:1,000) (Cell Signaling Technology).

All solutions used in this protocol were filtered on a 0.2 µm Minisart filtering unit (Sartorius). 1 x 10^7 cells of *T. brucei* procyclic form or *T. cruzi* epimastigotes were harvested by centrifugation at 800 x g, 10 min, RT (Sigma Laboratory Centrifuge 4K15) and resuspended in 1 ml ice-cold 1X PBS. Equal volume of 8% methanol-free PFA in 1X PBS was added to fix the cells by incubating at RT for at least 15 min before taking them out of BSL3 culture room. Round coverslips (13 mm diameter, WVR) were washed by dipping them in 70% ethanol and letting them air dry. 150 µl of poly-L-lysine hydrobromide solution at 0.1 mg/ml was added to each coverslip place in a piece of parafilm. Coverslips were covered from light at RT for 15 min, after which solution was pipetted away and coverslips washed twice with sterile water. Then, they were dried in a

pre-heated hybridization oven (same as used for Southern blot) at 60ºC and turned off to avoid over-heating of coverslips, for 1 hour. To these coverslips, 10 µl of fixed cells (5 x 10^4 cells) were placed on the centre of each coverslip and left O/N on a piece of parafilm covered from light at RT. After it, cells were re-hydrated by adding 200 µl of 1X PBS and incubating for 2 min inside a humid chamber. 1X PBS was taken away and cells permeabilized with 100 µl of 0.1% Triton X-100 in 1X PBS for 10 min, RT inside humid chamber (always covered with aluminium foil). After it, coverslips were washed once with 50 µl 1X PBS and then cells blocked by incubating for 1h at RT with 50 µl blocking buffer (5% w/v FSG, 10% goat serum, 0.05% Triton X-100 in 1X PBS, 0.05% NaN_3 in 1X PBS). Blocking solution was then removed and coverslips washed once with 50 µl washing buffer (1% w/v FSG, 0.05% Triton X-100 in 1X PBS, 0.05% NaN_3 in 1X PBS). Primary antibodies mentioned before were prepared by diluting antibody stock in washing buffer (1:1,000) and 50 µl of the solution were added to the coverslips to be incubated inside humid chamber for 1 h, RT. After the dilution, primary antibodies solutions were centrifuged 16,000 x g, 10 min, 4ºC (Eppendorf microcentrifuge 5415R) prior to be used. Coverslips were then washed 3 times with 50 µl washing buffer, and 50 µl 1:500 dilution of secondary antibody (goat anti-rabbit or goat anti-mouse Alexa Fluor 488, Thermo Fisher Scientific) in washing buffer were added. Coverslips were incubated inside humid chamber for 1 h, RT to then be washed 3 times with 50 µl washing buffer. Coverslips were then mounted in frosted glass slides (VWR). ProLong Gold antifade reagent with DAPI (Thermo Fisher Scientific) was spotted on the slide and the coverslip lowered slowly on the drop of mounting agent. The slides were left curing O/N in the dark at RT after being sealed with nail varnish. Microscopy was performed on a Zeiss Axiovert 200M microscope.

4. RESULTS I: Expression and assays on *T. cruzi* mitochondrial fucosyltransferase (TcFUT1), synthesis of GDP-[^{3}H]Fuc sugar donor, and TcFUT1 immunolocalization.

4.1 Identification and analysis of the putative *TcFUT1* gene

The CAZy database (http://www.cazy.org) lists nine distinct FUT families. Previously (Bandini *et al.*, 2021), one or more sequences from the eight families of FUTs that use GDP-Fuc as donor substrate (GT10, 11, 23, 37, 41, 65, 68 and 74), and the one that uses TDP-N-acetyl-D-fucosamine (TDP-Fuc4NAc) (GT56), were selected for searching putative FUTs in the genome of *T. brucei* using the BLASTp search algorithm (Altschul *et al.*, 1997). Only one putative FUT was identified in the *T. brucei* genome (GeneDB ID: Tb927.9.3600), belonging to the GT11 family. This was recombinantly expressed and characterized as a functional FUT (TbFUT1) by Bandini (Bandini *et al.*, 2021).

The *TbFUT1* gene was then used as the query sequence to search the TriTryp database (TriTrypDB) (Aslett *et al.*, 2010) in order to find orthologue(s) in the available *T. cruzi* strain genomes. Every genome of *T. cruzi* contained a *TbFUT1* orthologue. For this study we selected the gene from *T. cruzi* Sylvio X10 strain (GeneDB ID: TCSYLVIO_003556), which is referred to as encoding a hypothetical protein of 307 amino acids in TriTrypDB. We named the corresponding protein sequence TcFUT1.

Analysis of TcFUT1 sequence revealed that it is also characterized by the 4 GT11 family motifs, a low-confidence predicted TM domain including motif IV, and a mitochondrial import sequence, as previously defined for TbFUT1 (Bandini *et al.*, 2021) (see section 1.5) (Fig. 4.0).

Figure 4.0: Multiple sequence alignment of TcFUT1 (red framed) and other GT11 family fucosyltransferases (putative or characterized). Sequencies from protozoa (*T. cruzi*, *T. brucei brucei*, *L. major*, *C. fasciculata*, *B. saltans*), bacteria (*E. coli*), and mammals (*Mm*FUT2 and *Hs*FUT2) were used for the alignment. Predicted TM domain for TcFUT1 and TbFUT1 is marked with purple square. GT11 4 conserved motifs are indicated (black square), being motif I (asterisk) the only one with an assigned activity. The cleavage position of the predicted mitochondrial targeting sequences in TcFUT1 and TbFUT1 are marked with red arrows.

4.2 Cloning of the *TcFUT1* ORF for recombinant expression in *E. coli*

As observed previously (Bandini *et al.*, 2021), the mitochondrial targeting sequence (MTS) of TbFUT1 can affect its heterologous expression. Therefore, we PCR-amplified a truncated genomic open reading frame (ORF) of *TcFUT1* lacking its 5' encoded MTS (*TcFUT1cat*) and cloned it into the bacterial pET15b expression vector. We defined the MTS of TcFUT1 as its first 32 amino acids by using the online MTS predictor tools MitoProt II (Claros and Vincens, 1996) and MitoFates (Fukasawa *et al.*, 2015). At the same time, the partial ORF was codon-optimised for expression in bacteria, synthesized and sub-cloned into the same vector. The construct pET15b-6xHis-TS-TcFUT1cat, named rTcFUT1-1.a (Fig. 4.7), encoded for the TcFUT1 ORF without MTS,

with a 6xHis tag at its N-terminus and a TEV protease cleavage site (TS) between the 6xHis tag and the ORF sequence.

These two constructs containing the *TcFUT1* sequence, either amplified from *T. cruzi* Sylvio X10 genomic DNA by PCR or sub-cloned from a purchased vector including the codon-optimised sequence, were sequenced to confirm successful amplification and/or sub-cloning. In the case of the PCR-amplified construct, there was a single nucleotide polymorphism (SNP) between the PCR-amplified and the genome database sequences at position 324 of the *TcFUT1* DNA sequence (adenine → thymine). This results in a change in a GAA (Glu) codon to a GAT (Asp) codon which, while not silent, is a conservative acidic amino acid change. We presumed that this represented an immaterial SNP between the genome sequence of the reference strain and our laboratory strain.

4.3 Attempts to express rTcFUT1-1.a in ArcticExpress (DE3) *E. coli*

Following the precedent of TbFUT1, we proceeded to transform ArcticExpress (DE3) *E. coli* cells with the genomic and codon-optimised pET15b constructs. Initially, low-volume culture aliquots were analysed by SDS-PAGE and Western blotting. These showed the presence of recombinant TcFUT1 (rTcFUT1) only in cultures induced with IPTG (Fig. 4.1). The presence of recombinant 6xHis-tagged rTcFUT1 was inferred from the presence of a protein of the correct apparent molecular weight (MW) (34 kDa) in eight clones of IPTG-induced rTcFUT1-1.a transformed cells (Fig. 4.1A, lanes 1-8), but not in uninduced cells (Fig. 4.1A, lane 9), and the immunoreactivity of these band with an anti-6xHis tag antibody (a-6xHis) (Fig. 4.1B).

We also detected the expression of another protein at about 60 kDa (Fig. 1A). According to manufacturer, this other protein is likely to be Cpn60, a chaperonin originally found on the psychrophilic bacterium *Oleispira antarctica*, and expressed in these ArcticExpress *E. coli* cells, which shows a high protein refolding activity at temperatures of 4-12°C.

Since no advantage was found in the expression of constructs containing the codon-optimised sequence, we decided to work only with the genomic sequence onwards.

We performed a pulldown from the soluble fraction from a whole cell lysate of rTcFUT1-1.a transfectant cells from a small scale culture using cobalt-charged magnetic

beads. This revealed a protein of the correct apparent molecular weight (Fig. 4.2A, lane 1) which, when excised and digested with trypsin was identified as rTcFUT1-1.a by proteomics (Fig. 4.2B). However, this protein was not successfully eluted from the cobalt charged magnetic beads with imidazole (Fig. 4.2A, lane 2).

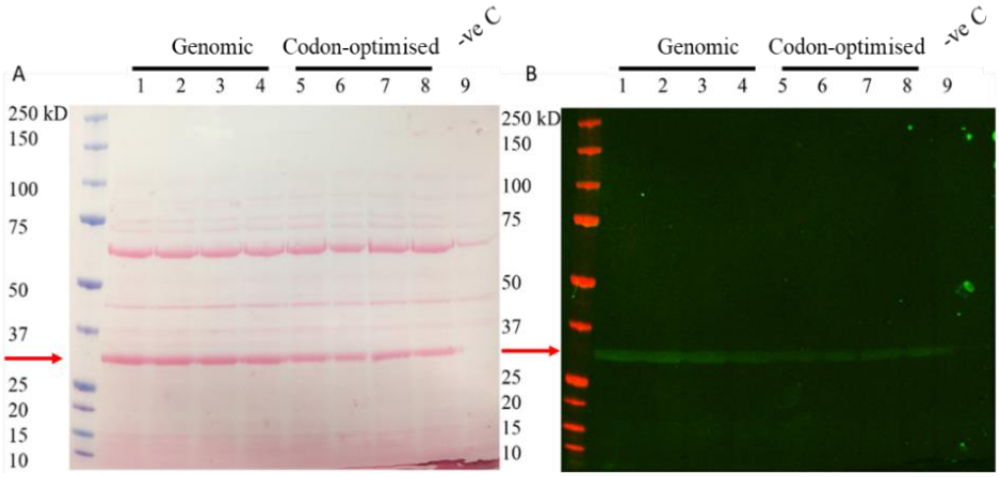


Figure 4.1: Expression of rTcFUT1-1.a in *E. coli* Arctic Express cells. Ponceau S staining (A) and α-6xHis tag Western blot (B) of whole culture aliquots from different transformed bacterial colonies after 24 h IPTG induction (lanes 1 to 8) or without induction (lane 9). Red arrows indicate the recombinant protein. Lanes 1 to 4 correspond to four different colonies transformed with the *T. cruzi* genomic sequence. Lanes 5 to 9 show the protein expression of four different colonies transformed with the bacterial codon-optimised sequence.

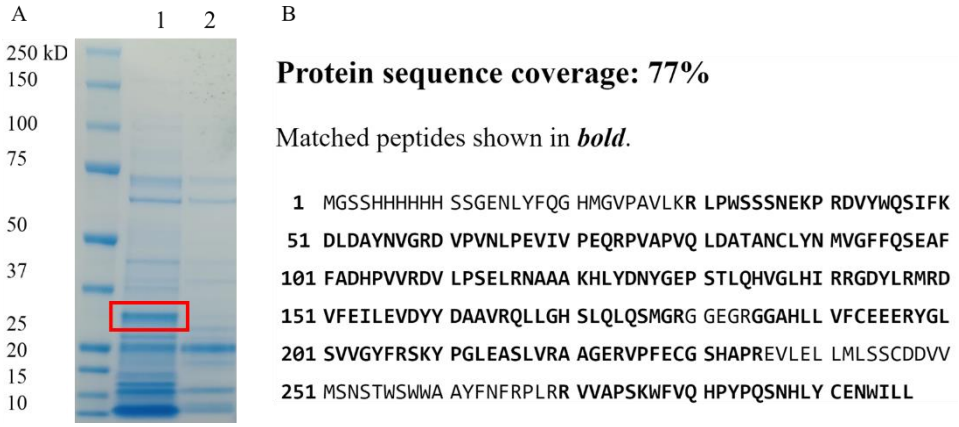


Figure 4.2: Trial small-scale protein purification and mass spectrometry identification of rTcFUT1-1.a. (A) SDS-PAGE and Coomassie blue staining of material eluted from cobalt-charged magnetic bead pulldowns with SDS sample buffer (lane 1) or 300 mM imidazole (lane 2). (B) Mascot software output mapping the detected peptides from trypsin digestion of the band indicated in the red box of (A, lane 1) to input sequence rTcFUT1-1.a, indicating 77% sequence coverage.

Following these results, we decided to try to scale-up protein expression of rTcFUT1-1.a in ArcticExpress *E. coli* and to perform IMAC pull-down purification trials using cobalt-charged magnetic beads and nickel-nitrilotriacetic acid (Ni-NTA)-charged agarose resin.

We observed a significant decrease in the apparent efficiency of protein expression upon scale-up, since less 34 kDa material was found in the soluble and pellet fractions (Fig. 4.3A, lanes 1 and 2) than had been previously observed in whole cell

lysates from small scale cultures (Fig. 4.1A. lanes 1-8). Nevertheless, some 34 kDa α-6xHis-positive protein was found in the soluble fraction (Fig. 4.3A and B, lane 1) and we proceeded to try to trap and elute that material on cobalt charged magnetic beads. However , the majority of the tagged rTcFUT1 appeared in the supernatant of the pulldown (Fig. 4.3B, lane 3) and lower molecular weight α-6xHis-positive bands were also apparent, suggestion that protein degradation might have occurred.

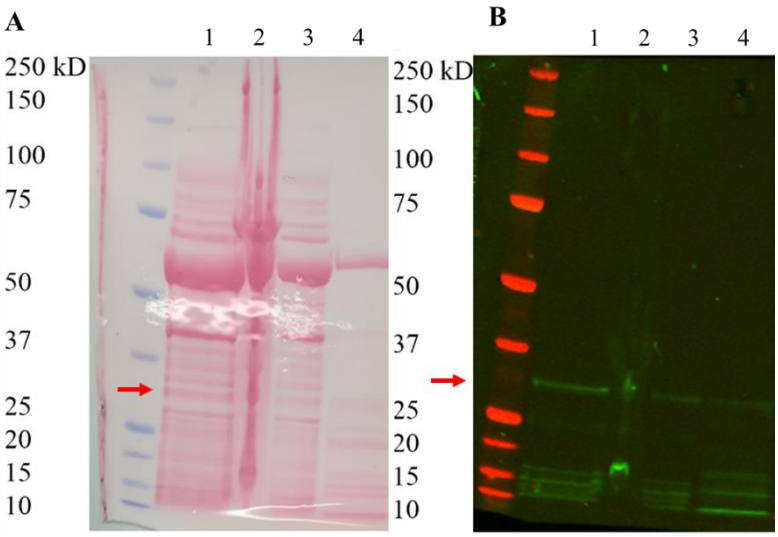


Figure 4.3: Large-scale purification trials of rTcFUT1-1.a by IMAC on cobalt charged magnetic beads. Ponceau S staining (A) and α-6xHis tag Western blot (B) of the soluble (lane 1) and insoluble pellet (lane 2) material. The rTcFUT1 band is indicated with red arrow. The soluble fraction was incubated with cobalt charged magnetic beads and the supernatant of this pulldown (lane 3) and the material eluted with SDS sample buffer (lane 4) were also analysed.

We then moved to try Ni-NTA agarose resin instead of cobalt charged magnetic beads for protein purification. The SDS-PAGE analysis of fractions from two purifications are shown in (Fig. 4.4). Most of the rTcFUT1 appeared not to bind to the Ni-NTA beads (Fig. 4.4A, lane 2). However, some of the recombinant product was recovered, mainly in the last four of fourteen consecutive 300 mM imidazole elutions (Fig. 4.4A, lanes 11-14), along with a considerable amount of Cpn60 and several other contaminating proteins. We tried to resolve the Cpn60 contamination in another purification by including a high KCl treatment (20 mM HEPES pH 7.0, 5 mM ATP, 10 mM $MgCl_2$, 150 mM KCl) (Joseph and Andreotti, 2008) before elution (Fig. 4.4B). In this purification, most of the rTcFUT1 remained in the supernatant of the pulldown (Fig. 4.4B, lane 2), and even though beads were washed with the high KCl buffer before elution, a significant amount of the chaperonin remained, being eluted in consecutive elutions of increasing imidazole concentration (Fig. 4.4B, lanes 3-7).

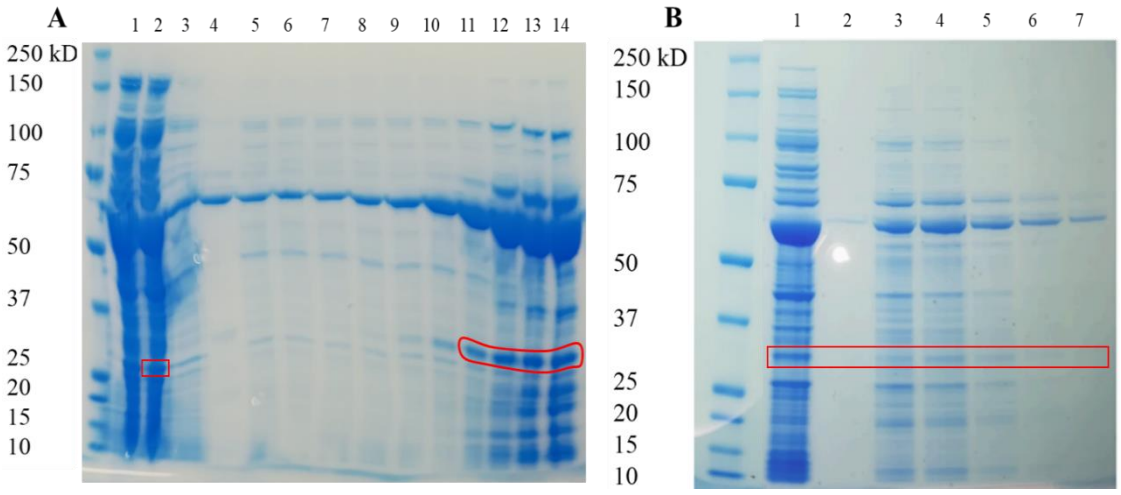


Figure 4.4: Ni-NTA resin-based purification trials. SDS-PAGE and Coomassie blue staining of fractions from large-scale purification attempts. (A) Cell lysate soluble fraction (lane 1), FT after incubation with Ni-NTA beads (lane 2), bead washings (lane 3) and ten consecutive 300 mM imidazole elutions (lanes 4 to 14) were analysed. The rTcFUT-1.a bands are indicted by the red boxes. (B) Soluble fraction from another culture was incubated with Ni-NTA resin, treated with high KCl as an attempt to dissociate Cpn60, and the supernatant of the pulldown (lane 1), the resin after being eluted (lane 2), and 5 consecutive elutions with 50, 100, 150, 200 and 250 mM imidazole (lanes 3-7) were analysed. The rTcFUT-1.a bands are indicated by the red boxes.

4.4 Attempts to express rTcFUT1-1.a in BL21 (DE3) pLysS *E. coli*

To avoid the Cnp60 chaperonin contamination problem described above, we decided to move to the BL21 (DE3) pLysS *E. coli* expression system, which lowers the background expression of the transformed construct without interfering with expression after induction.

These cells were transformed with the construct rTcFUT1-1.a, and tests were performed to define the best expression conditions regarding IPTG induction concentration, time of culturing before induction, and salt concentration of the lysis buffer (Fig. 4.5). These data suggested that 16 h of cell growth was better than 24 h of cell growth before IPTG induction (compare Fig. 4.5A and B), that 0.25 mM IPTG was the best induction condition and that 150 mM NaCl was the best salt concentration in the lysis buffer to assist rTcFUT-1.a solubilisation (Fig. 4.5C).

After this, several large-scale purification trials were performed, trying to optimize the conditions of IMAC purification performed in an HPLC system. Unfortunately, in all cases the recombinant protein did not bind the nickel IMAC column and always appeared predominantly in the FT, see (Fig. 4.6) as an example.

These results suggested that the r*Tc*FUT1-1.a construct expressed in BL21 (DE3) pLysS E. coli was unable to bind to immobilised nickel matrices (see section 6.1).

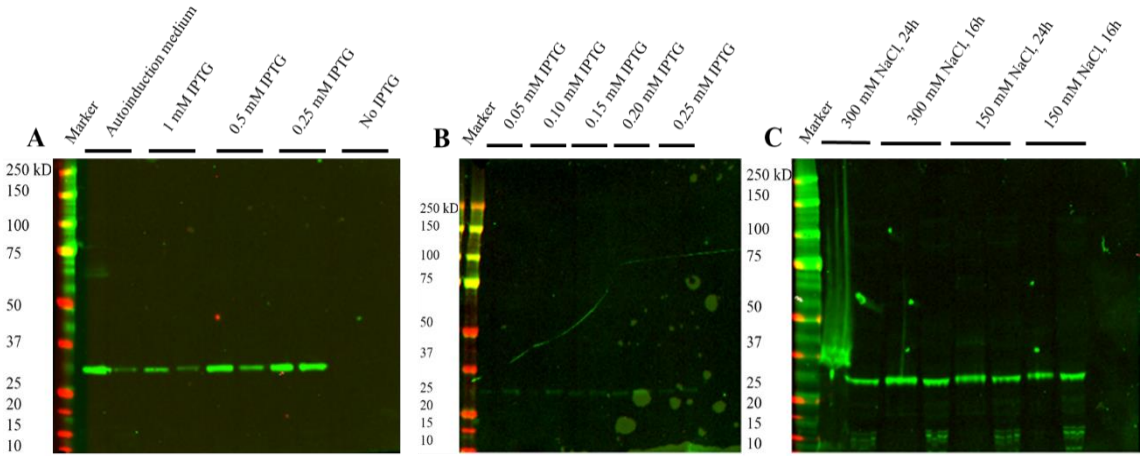


Figure 4.5: α-6xHis tag Western blot analyses of different expression trials conditions for rTcFUT1-1.a (34 kDa) in BL21 (DE3) pLysS *E. coli*. Small (10 mL) cultures were used, and aliquots of the insoluble pellet (left lane) and soluble supernatant (right lane) fractions were analysed for each condition tested. (A) IPTG induction at 16ºC after 16 h of cell growth. (B) IPTG induction at 16ºC after 24 h of cell growth. (C) Different salt concentrations in the lysis buffer following 0.25 mM IPTG induction following 16 h of cell growth.

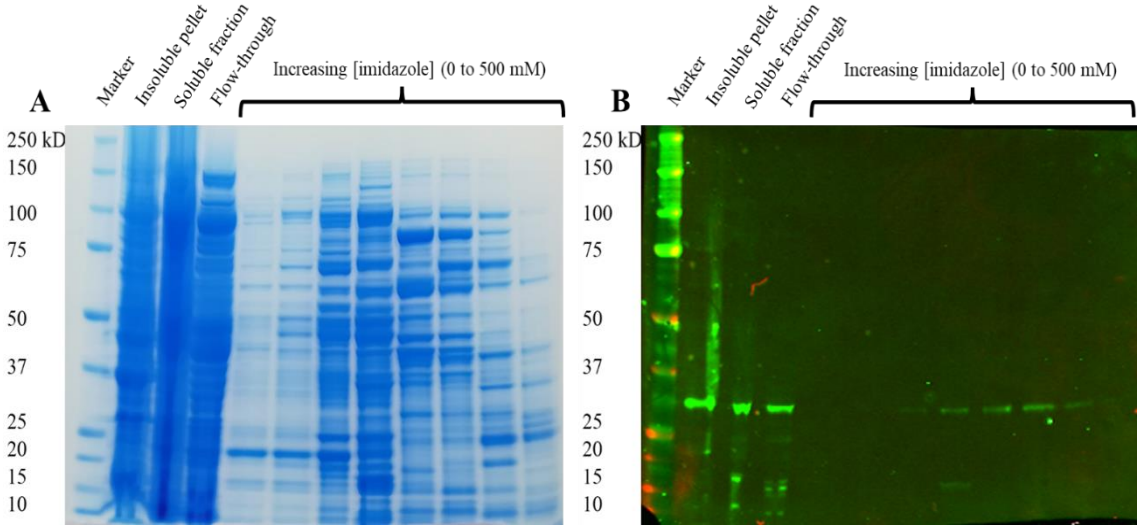


Figure 4.6: HPLC IMAC purification of rTcFUT1-1.a (34 kDa) from BL21 (DE3) pLysS *E. coli* large (2 L) cultures. (A) SDS-PAGE and Coomassie blue staining and (B) α-6xHis tag Western blot of fractions from a representative purification attempt.

4.5 Attempts to express a new batch of constructs in SHuffle T7 Express lysY

Based on these previous results with both *E. coli* strains, we generated a new set of constructs with the same pET15b vector and containing the same N-terminal 6xHis tag for nickel affinity purification, but with either TcFUT1cat or the full-length version of the protein and one of two different solubility enhancers, maltose-binding protein (MBP) and glutathione-S-transferase from *Schistosoma japonicum* (GST). As previously, these constructs were confirmed by DNA sequencing. Construct rTcFUT1-2.a encodes a recombinant product of 76 kDa; rTcFUT1-3.a, 60 kDa; rTcFUT1-2.b, 79.2 kDa; and rTcFUT1-3.b, 63.3 kDa. They are schematized below (Fig. 4.7).

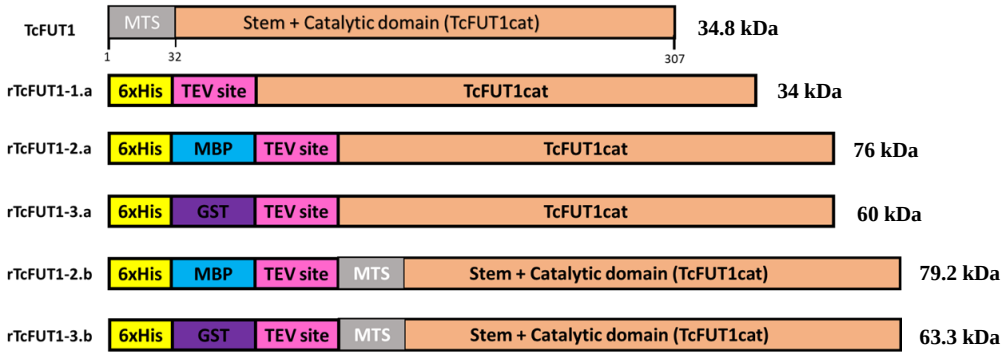


Figure 4.7: Scheme of TcFUT1 bacteria-expression constructs.

In addition, we decided to use a new *E. coli* strain: SHuffle T7 Express lysY. This strain has the same tight regulation of construct expression as the pLysS strain, but also has an enhanced capacity to correctly fold proteins with multiple disulphide bonds, such as GTs (Moremen and Haltiwanger, 2019). Initial trials were performed to define the best conditions for expression (Fig. 4.8). Small (10 mL) cultures of bacteria transformed with the new batch of constructs were used to test conditions of induction temperature and IPTG concentration in 24 h induction experiments. Soluble fractions and insoluble pellets of the lysates were analysed by α-6xHis tag Western blot. This showed that the majority of the recombinant 6xHis tagged proteins were found in the soluble fraction for the full-length TcFUT1 constructs (rTcFUT1-2.b and 3.b) (Fig. 4.8A, B and C), whereas the TFUT1cat constructs (rTcFUT1-2a and rTcFUT1-3a) showed poor protein expression.

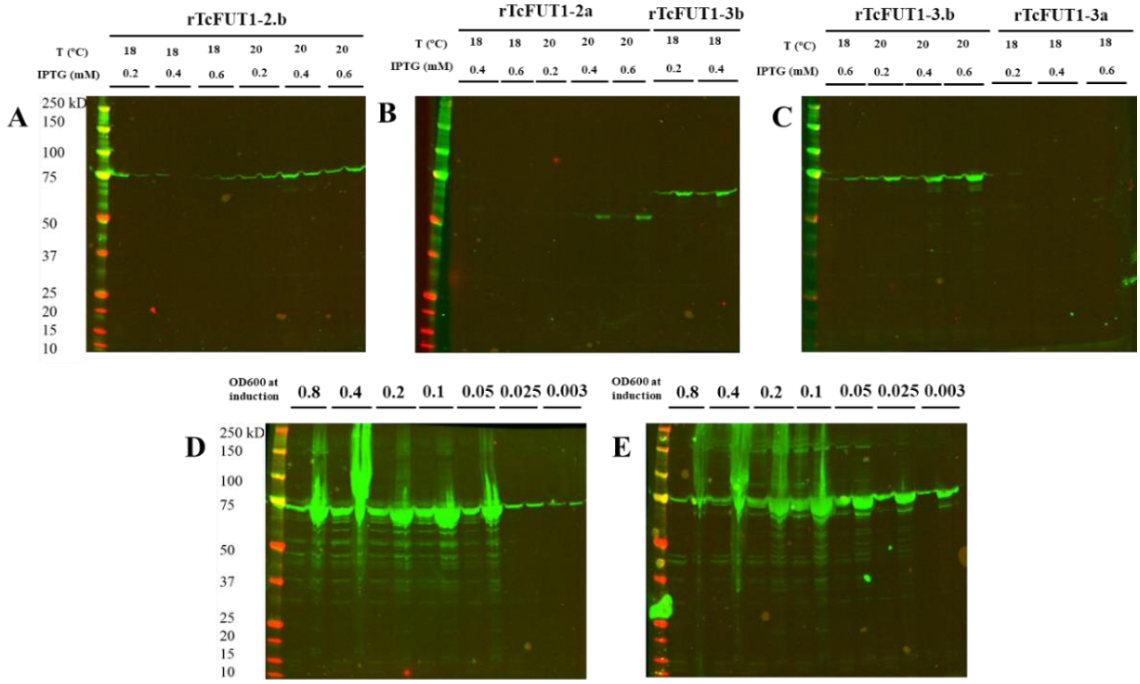


Figure 4.8: α-6xHis tag Western blot analysis of different expression trials with new batch of constructs in SHuffle T7 Express lysY *E. coli*. Small (10 mL) cultures were used, and the soluble supernatant (left lane) and insoluble pellet (right lane) fractions were analysed for each construct and condition tested, as described above the gel lanes. (A-C) Tests of induction temperature and IPTG concentration for 24 h induction performed for the four constructs. Additionally, initial OD600 values for IPTG induction were tested for constructs rTcFUT1-2.b (D) and rTcFUT1-3.b (E).

Based on these results, we defined 24 h induction with 0.4 mM IPTG at 20ºC as the conditions that yielded the most soluble recombinant protein, especially for construct rTcFUT1-2.b (Fig. 4.8A). We also tested different initial OD600 values for IPTG induction of transformants with full-length TcFUT1 (Fig. 4.8D and E). Based on the results, we defined 0.1 as the OD600 to use for further purification assays.

After these expression trials, we selected the construct rTcFUT1-2.b (containing an MBP tag) for further purification not only because it gave higher yields than the rTcFUT1-3.b (containing a GST tag) (Fig. 4.8D and E, lanes 0.1) but also because MBP has a higher capacity to enhance the solubility of co-expressed recombinant proteins than GST (Hammarström *et al.*, 2006; Ohana *et al.*, 2009; Costa *et al.*, 2014). Furthermore, rTcFUT1-3.b showed similar problems to rTcFUT1-1.a in a large-scale purification attempt, i.e., i t was mostly unable to bind to the nickel column: As we can see in (Fig. 4.9) in an HPLC IMAC purification of rTcFUT1-3.b, the 63.3 kDa recombinant product was expressed, mainly as insoluble form (Fig. 4.9A), whereas the soluble portion remained in the flow through after passing it through the nickel column (Fig. 4.9B, lane FT). Minimal signal was obtained from the elutions with highest imidazole concentration (Fig. 4.9B), indicating an aggregated, and therefore most probably inactive, form.

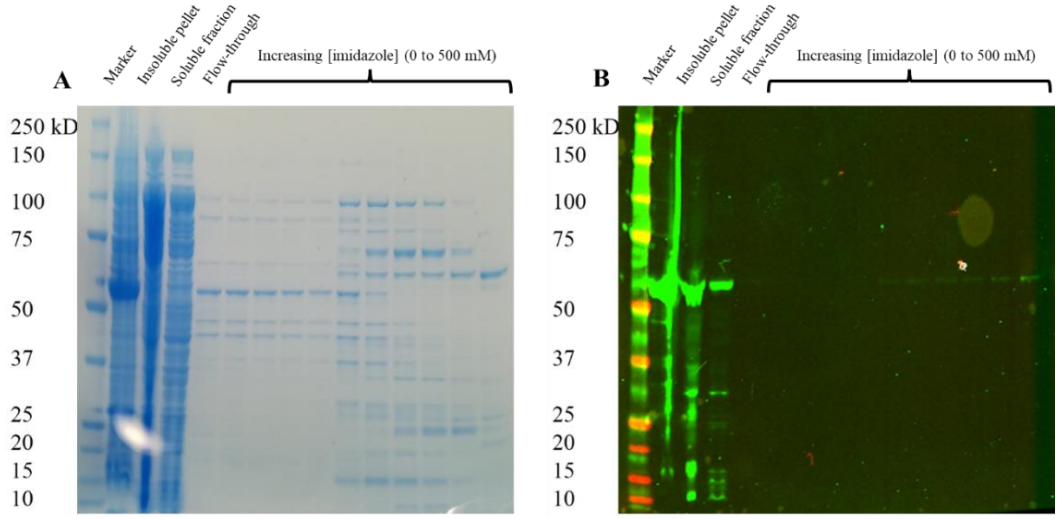


Figure 4.9: HPLC IMAC purification of rTcFUT1-3.b (63.3 kDa) expressed in SHuffle T7 Express lysY *E. coli* large (2 L) cultures. (A) SDS-PAGE and Coomassie blue staining and (B) α-6xHis tag Western blot of fractions from the purification attempt.

During the subsequent rTcFUT1-2.b large-scale purification trials by HPLC IMAC, different imidazole concentrations, concentration steps and longer washes were tested, among other modifications. Although in most of them rTcFUT1 was obtained in the eluted fractions, the lowest number of contaminants were obtained along with the recombinant product following the procedure described in detail in section 3.3.4, despite the low yield (~0.1 mg/L) obtained in all cases. As shown in (Fig. 4.10), following this

procedure, rTcFU1 from rTcFUT1-2.b construct (79.2 kDa) was present as both soluble and insoluble form, and a significant portion of it was not able to bind the nickel column (Fig. 4.10B, lane flow through). Despite this, some material was bound, and it eluted either early in the imidazole gradient together with substantial contamination with 70 kDa and 60 kDa proteins (most likely *E. coli* chaperonins), or late in the gradient along with a 15 kDa *E. coli* protein. Some degraded protein was also seen between these early and late eluting forms (lower bands in Fig. 4.10B). The late-eluting material with the 15 kDa contaminant was otherwise reasonably pure.

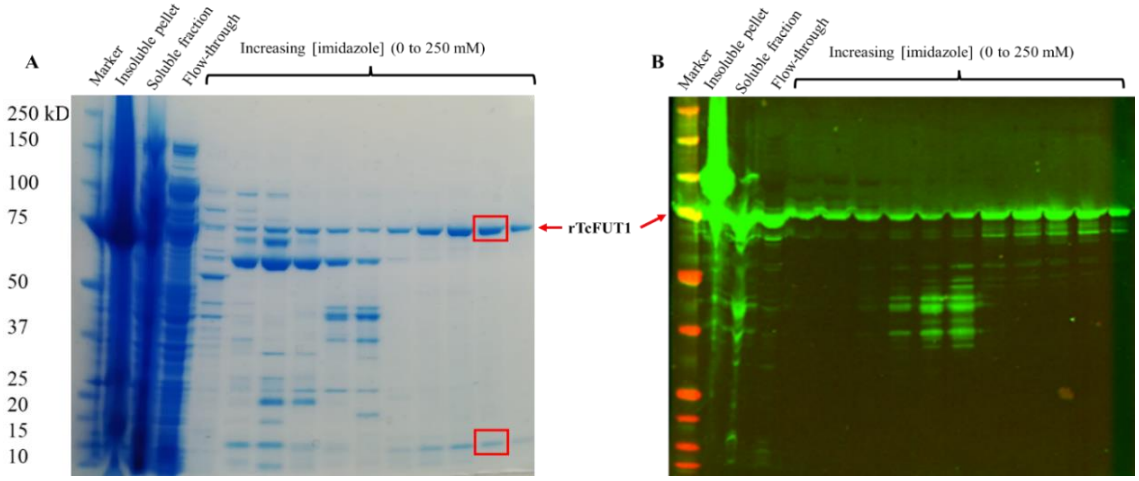


Figure 4.10: HPLC IMAC purification of rTcFUT1-2.b (79.2 kDa) in SHuffle T7 Express lysY *E. coli* large (2 L) cultures. (A) SDS-PAGE and Coomassie blue staining analysis of different fractions along the purification process. (B) α-6xHis tag Western blot analysis of the same fractions. Bands framed in red were submitted for MS peptide fingerprinting and identified as TcFUT1 and *E. coli* IbpA (Fig. 4.11).

Further, it eluted at imidazole concentrations (150-250 mM; Fig. 4.10A, last 5 lanes from right) typical of IMAC, suggesting the protein was not aggregated. Samples of both the rTcFUT1-2b and the contaminating 15 kDa protein were excised from the gel for trypsin digestion and MS fingerprinting by proteomics. The peptide data were used to search the *T. cruzi* Sylvio X10 strain predicted protein database (supplemented with the sequence of rTcFUT1-2.b) and the *E. coli* BL21 (DE3) strain (parental strain of the one used for these purification trials) predicted protein database. This identified the 79 kDa band as rTcFUT1-2b and the 15 kDa band as IbpA, a 16 kDa bacterial chaperone protein (Figure 4.11A and B, respectively)

After some repetitions to confirm the reproducibility of this purification procedure, the late-eluting fractions containing the cleanest eluted rTcFUT1-2.b were dialysed to remove imidazole, a well-known hindrance for many type of activity assays, and to equilibrate the rTcFUT1-2.b protein in the activity assay buffer.

rTcFUT1-2.b

Protein sequence coverage: 74%

Matched peptides shown in ***bold***.

```
1   MGSSHHHHHH GSSMKIEEGK LVIWINGDKG YNGLAEVGKK FEKDTGIKVT
51  VEHPDKLEEK FPQVAATGDG PDIIFWAHDR FGGYAQSGLL AEITPDKAFQ
101 DKLYPFTWDA VRYNGKLIAY PIAVEALSLI YNKDLLPNPP KTWEEIPALD
151 KELKAKGKSA LMFNLQEPYF TWPLIAADGG YAFKYENGKY DIKDVGVDNA
201 GAKAGLTFLV DLIKNKHMNA DTDYSIAEAA FNKGETAMTI NGPWAWSNID
251 TSKVNYGVTV LPTFKGQPSK PFVGVLSAGI NAASPNKELA KEFLENYLLT
301 DEGLEAVNKD KPLGAVALKS YEEELAKDPR IAATMENAQK GEIMPNIPQM
351 SAFWYAVRTA VINAASGRQT VDEALKDAQT NSSSNNNNNN NNNNLGENLY
401 FQGHMNSRAY LTTNLVGGLG NQLFLVANLL ATSRRNGVPA VLKRLPWSSS
451 NEKPRDVYWQ SIFKDLDAYN VGRDVPVNLP EVIVPEQRPV APVQLDATAN
501 CLYNMVGFFQ SEAFFADHPV VRDVLPSELR NAAAKHLYDN YGEPSTLQHV
551 GLHIRRGDYL RMRDVFEILE VDYYDAAVRQ LLGHSLQLQS MGRGGEGRGG
601 AHLLVFCEEE RYGLSVVGYF RSKYPGLEAS LVRAAGERVP FECGSHAPRE
651 VLELLMLSSC DDVVMSNSTW SWWAAYFNFR PLRRVVAPSK WFVQHPYPQS
701 NHLYCENWIL L
```

***E. coli* IbpA**

Protein sequence coverage: 79%

Matched peptides shown in ***bold***.

```
1   MRNFDLSPLY RSAIGFDRLF NHLENNQSQS NGGYPPYNVE LVDENHYRIA
51  IAVAGFAESE LEITAQDNLL VVKGAHADEQ KERTYLYQGI AERNFERKFQ
101 LAENIHVRGA NLVNGLLYID LERVIPEAKK PRRIEIN
```

Figure 4.11: Mascot software output mapping the detected peptides from trypsin digestion of the band indicated in the red boxes of (Fig. 4.10A) to input sequence rTcFUT1-2.b (upper band) and *E. coli* B strain genome (lower band) and indicating 74% and 79% sequence coverage, respectively.

4.6 [^{3}H]-based radioactive FUT assays with rTcFUT1-2.b product.

Despite the low purification yield (~0.1 mg/L), sufficient protein was obtained to perform fucosyltransferase activity assays. Previously (Bandini *et al.*, 2021), the TbFUT1 activity was defined as a GDP-Fuc : β-D-galactose α1-2-fucosyltransferase, and this was independent of the presence of co-purified contaminants (mainly bacterial chaperones) and tags in the recombinant product. Following this previous work, we hoped to see a similar activity for TcFUT1. Aliquots (2 µg) of recombinant product (6xHis-MBP-rTcFUT1) from pooled and dialysed eluted fractions were tested by incubation with commercial GDP-[^{3}H]Fuc (1 µCi) as a donor substrate and a panel of commercially available mono- to tetrasaccharides (1 mM each) as α1-2 potential FUT acceptor substrates, or sub-structures thereof, selected from the literature (Wang *et al*., 1999; Li *et al.*, 2008; Zhang *et al.*, 2010). Lacto-N-biose (LNB, Galβ1-3GlcNAc), Galβ1-3GalNAcβ1-O-methyl glycoside (GalGalNAc-OMe), lacto-N-tetraose (LNT, Galβ1-3GlcNAcβ1-3Galβ1-4Glc), lactose (Lac, Galβ1-4Glc), N-acetyllactosamine (LacNAc, Galβ1-4GlcNAc) and Galβ1-O-methyl glycoside (βGal-OMe) met these criteria. In later assays including recombinant protein obtained from a eukaryotic expression system, Lacto-N-biose-β-O-methyl glycoside (LNB-OMe, Galβ1-3GlcNAc-OMe), and Galβ1-6GlcNAc (GNG) were also added to the tested acceptors. These saccharides are schematized in (Fig. 4.12).

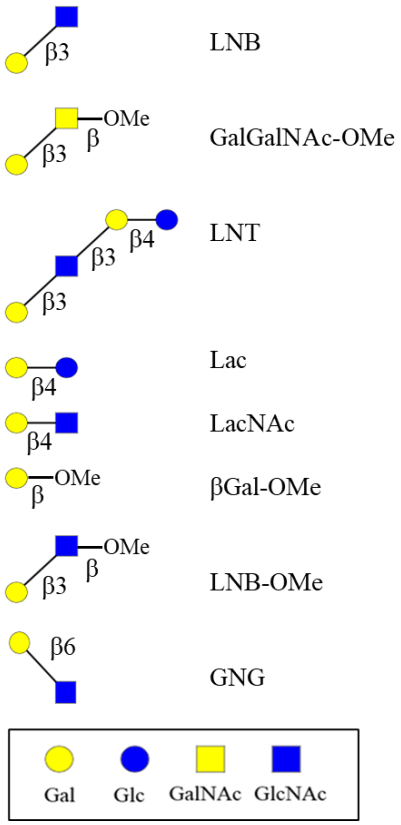


Figure 4.12: Representation of synthetic acceptors used in the radioactivity-based FUT activity assays. Legend (framed) describe the symbols used for the 4 different sugars forming these saccharides.

The products of these rTcFUT1-2.b + GDP-[^{3}H]Fuc + glycan acceptor assays were desalted and stripped of excess GDP-[^{3}H]Fuc by passage through mixed-bed ion exchange resins (section 3.7.12), dried and analysed by normal-phase HPTLC. The radioactive products were detected by fluorography (section 3.7.9). Unfortunately, no evidence of activity was detected in any trial. As shown in a representative example (Fig. 4.13), mainly two bands appeared. One that had the same retention factor as L-Fuc, and an upper one initially undetermined. After analysing the stock GDP-[^{3}H]Fuc by the same means, we concluded that the radioactive donor substrate was partially degraded due to its long term storage, into free [^{3}H]Fuc, and another, less polar, radiolysis product(s) (R.P.) Minor spots closer to the origin (spotting place in the silica plate, corresponding to the bottom of the image) were also detected, most probably corresponding to traces of GDP-[^{3}H]Fuc that was not retained during the desalting process. Most importantly, no radioactive bands running below [^{3}H]Fuc were seen. This is where fucosylated glycan products are expected to chromatograph (Bandini *et al*., 2021).

From these data, we concluded that rTcFUT1-2.b was either inactive or that we had failed to provide it with a useable acceptor substrate. The possible reasons for our failure to obtain active recombinant TcFUT1 in *E. coli* are considered in section 6.1.

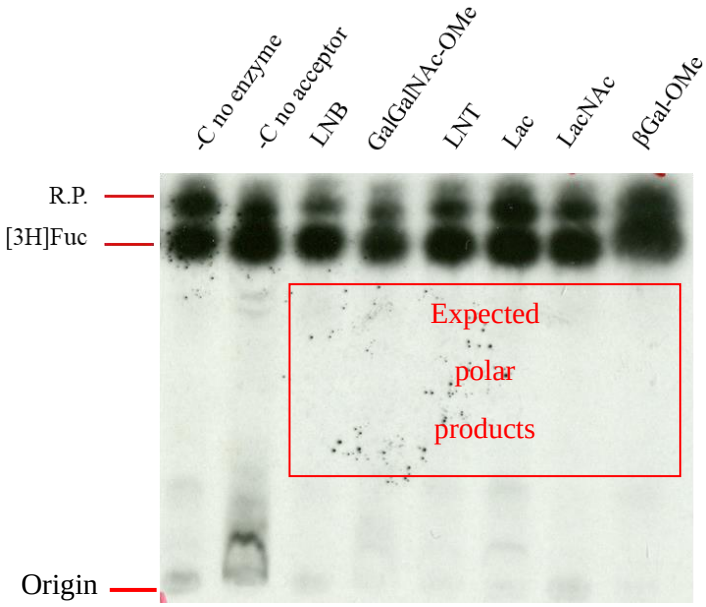


Figure 4.13: rTcFUT1-2.b activity assay. Fluorograph of the HPTLC plate showing the desalted products from the radioactive FUT assay performed with the acceptor substrates from (Fig. 4.12). No enzyme and no acceptor controls were included, as indicated. Tritium-labelled fucosylated glycan products would run well below the position of free [^{3}H]Fuc. R.P.: radiolysis product(s).

4.7 Cloning of the *TcFUT1* ORF, eukaryotic expression attempts and activity assays.

Eukaryotic GTs are known to be problematic when trying to express and purify them. To obtain well-folded, active recombinant protein molecular machinery for disulphide bond formation, N-glycosylation, and appropriate chaperones are often required (Moremen and Haltiwanger, 2019). Although there have been successful attempts to purify GTs from bacteria (Li *et al.*, 2008; Zhang *et al.*, 2010; Bandini *et al.*, 2021), most reports show that eukaryotic expression systems are more suitable (Ihara *et al.*, 2006; Brzezinski *et al.*, 2007; Miyazaki *et al.*, 2010; Urbanowicz *et al.*, 2017).

Based on this literature, we decided to generate, in parallel to the rTcFUT1-1.a construct, two other constructs, named rTcFUT1-1.b and rTcFUT1-1.c (Fig. 4.14). Therefore, the truncated ORF (amplified from genomic DNA or from a purchased plasmid containing a codon-optimised sequence for mammalian cells expression) of *TcFUT1* (*TcFUT1cat*) was cloned into two eukaryotic expression vectors containing different tags (pGEc1 and pGEc2). These were used for transfection into Expi293F cells, a well-characterized eukaryotic system suitable for the expression of many eukaryotic glycosyltransferases and other glycosylation-related proteins (Subedi *et al.*, 2015; K W Moremen *et al.*, 2018). Construct rTcFUT1-1.b consisted of *TcFUT1cat* in the pGEc1 vector C-terminally tagged with 8xHis and StrepTag peptide, separated from the 3'end of the ORF by a TEV protease cleavage site (TS). Its recombinant product had a predicted

molecular weight of 35 kDa. Construct rTcFUT1-1.c consisted of *TcFUT1cat* followed by a TS, a super folder green fluorescent protein (sGFP), a AviTag peptide and an 8xHis tag at its C-terminus, all included in plasmid pGEc2. Its recombinant product had a predicted molecular weight of 61.8 kDa.

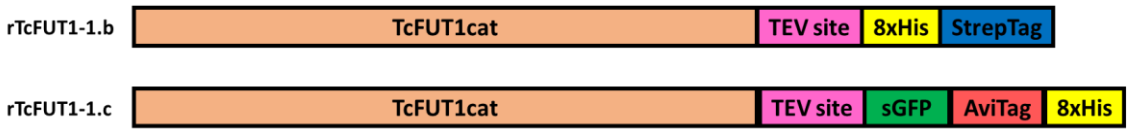


Figure 4.14: Scheme of first constructs generated for eukaryotic expression of TcFUT1. rTcFUT1-1.b (35 kDa) consisted on the cloned TcFUT1cat in pGEc1, whereas rTcFUT1-1.c (61.8 kDa) uses pGEc2 as the backbone. Tags are all included in the plasmids.

In an initial experiment, Expi293F cells were transfected with the genomic and codon-optimised versions of the rTcFUT1-1.b construct. Soluble fractions of cell lysates from the transfected cells were analysed by α-6xHis tag Western blot three, four, and six days post-transfection (Fig. 4.15). Unfortunately, no sign of recombinant product (of around 35 kDa) was detected at any point, independently of the TcFUT1cat version used.

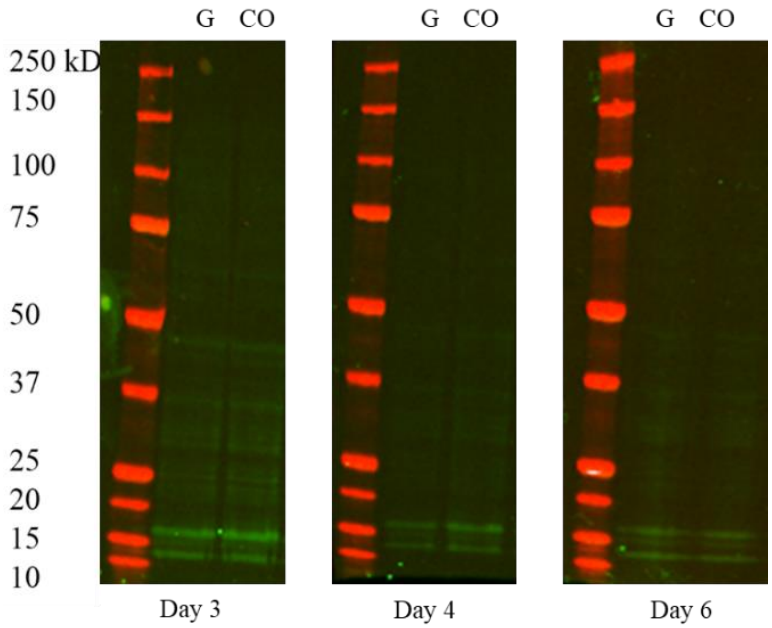


Figure 4.15: α-6xHis tag Western blot analysis of first eukaryotic expression trial of rTcFUT1-1.b (35 kDa) in Expi293F cells. G = genomic TcFUT1cat, CO = codon-optimised TcFUT1cat.

After this first failed attempt of eukaryotic expression, we focused on bacterial expression of rTcFUT1 (sections 4.3-4.6). When we eventually found that despite our efforts using different bacterial strains, only inactive rTcFUT1 was obtained, we decided to go back to our attempts at eukaryotic expression.

Basing our next strategy on Moremen's work (Moremen *et al.*, 2018), we generated a new batch of constructs (Fig. 4.16). rTcFUT1-4.a is similar to construct rTcFUT1-1.c but includes the full-length protein (MTS + TcFUT1cat). Constructs rTcFUT1-5.a and 5.b were generated by cloning either full-length TcFUT1 or

TcFUT1cat, respectively, into pGEn2 which contains the same tags as pGEc2 (Fig. 4.14, rTcFUT1-1.c) but at the N-terminus of the protein of interest, plus a cleavable signal sequence (CSS), which should cause the protein to be secreted to the medium.

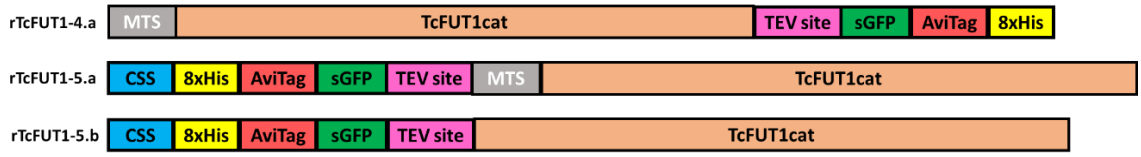


Figure 4.16: 2nd set of eukaryotic expression constructs. rTcFUT1-4.a expresses a recombinant product of 65.1 kDa; rTcFUT1-5.a, 68 kDa; and rTcFUT1-5.b, 64.6 kDa.

Even though the recommended expression procedure (section 3.4.2) is standardized, the day of highest expression must be determined empirically. With these three new constructs, plus the previously generated rTcFUT1-1.c (with genomic sequence) (Fig. 4.14), we performed an expression trial in parallel to compare the expression of each construct (and of a secreted IgG heavy chain control construct) over the days following transfection by α-6xHis tag (green) and α-rabbit IgG (red) Western blot (Fig. 4.17). From day 3 to day 6 after transfection, aliquots of the cell lysates were taken to analyse the soluble fraction and insoluble pellet of each. Additionally, for constructs rTcFUT1-5.a and rTcFUT1-5.b, an extracellular medium aliquot was also included in the analyses to detect any possible secreted recombinant product. Along with these, positive controls for intracellular expression (kind gift from Protein Production Team, unknown 6xHis tagged product of approximately 75 kDa) and extracellular expression (heavy chain rabbit IgG, 50 kDa) of our eukaryotic system, plus a negative control (empty plasmid) were used. On day 3 (Fig. 4.17A), the best result was achieved: The positive controls worked, proving our system was functional, and a signal from the soluble fraction of three of the constructs was detected (rTcFUT1-4a, -1.c, and -5.b). However, most of the signal for rTcFUT1-5.b was found in the insoluble fraction, and for construct rTcFUT1-1.c, the detected band in the soluble fraction had an approximate size of 50 kDa, whereas the recombinant product from this construct should be around 61.8 kDa. In contrast, a band of the correct size was detected in the soluble fraction from rTcFUT1-4.a-transfected cells.

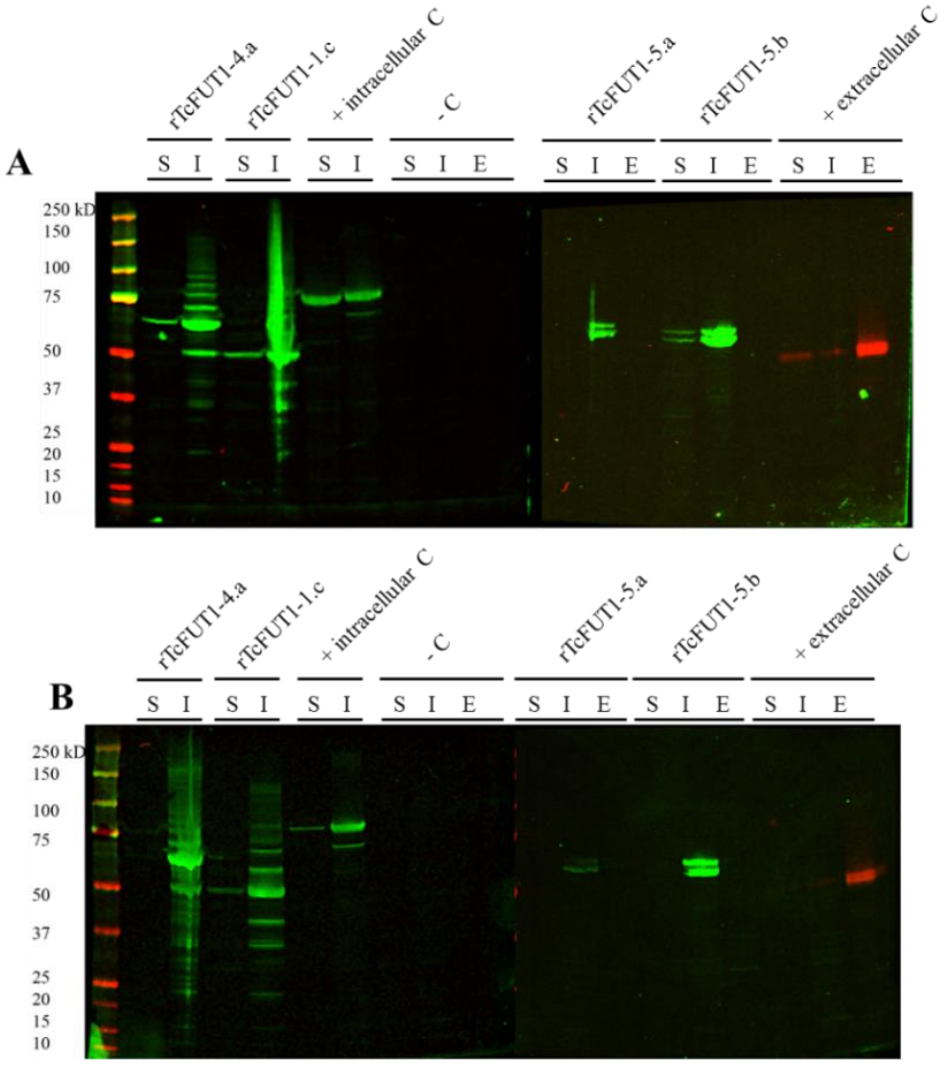


Figure 4.17: α-6xHis tag (green) and α-rabbit IgG (red) Western blot analysis of expression tests of eukaryotic constructs rTcFUT1-4.a, -1.c, -5.a and -5.b. Shown are results for day 3 (A) and day 5 (B). For each cell lysate soluble fraction and insoluble pellet were analysed. In the case of the N-terminally tagged constructs containing a CSS (-5.a and -5.b), extracellular medium aliquots were also tested, to determine if a secreted product was expressed.

The Western blot analyses were repeated over the subsequent three days, but the best results were seen on day 3. For example, on day 5 (Fig. 4.17B): there was no signal in the soluble fraction of rTcFUT1-4.a, the same aberrant molecular weight was seen for rTcFUT1-1.c, and only insoluble products were expressed from cells transfected with pGEn2-based constructs. From these data, we concluded that rTcFUT1-4.a was the most promising construct, and that 3 days after transfection was the best time to harvest rTcFUT1 for purification from a large-scale Expi293F transfected culture.

Multiple HPLC IMAC attempts were subsequently performed to try to optimize the process of purifying rTcFUT1 from rTcFUT1-4.a-transfected Expi293F cells from 1 L cultures. A representative result (Fig. 4.18) showed the elution of rTcFUT1, although a significant number of other co-eluted proteins were always obtained in the fractions.

The presence of rTcFUT1 was confirmed by α-6xHis tag Western blot (Fig. 4.18B) and peptide mass fingerprinting (Fig. 4.19).

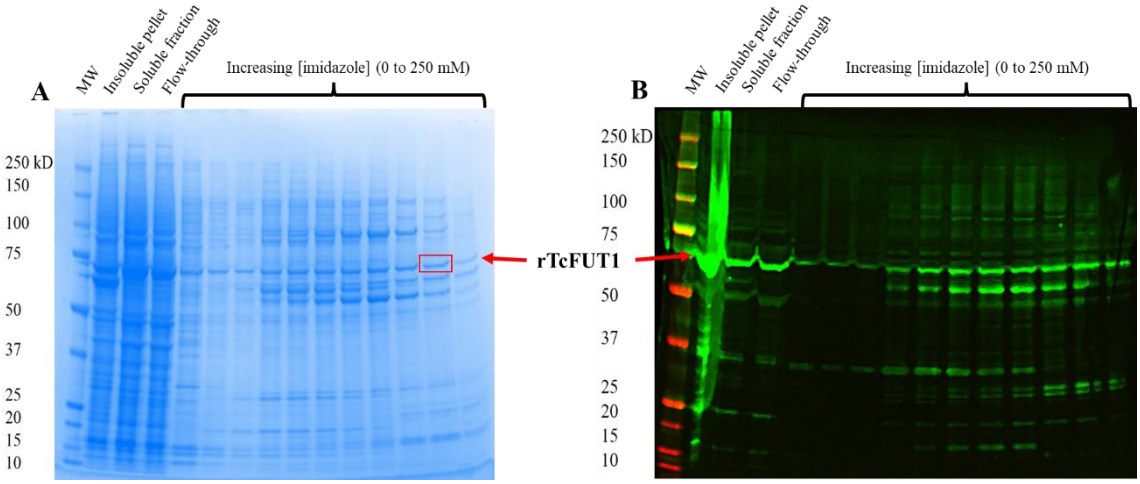


Figure 4.18: HPLC IMAC purification of rTcFUT1-4.a (65.1 kDa) from Expi293F cells large (1 L) culture. (A) SDS-PAGE and Coomassie blue staining and (B) α-6xHis tag Western blot of fractions from the purification attempt. Red framed band was submitted for MS peptide fingerprinting and identified as TcFUT1 (Fig. 4.19)

rTcFUT1-4.a

Protein sequence coverage: 81%

Matched peptides shown in ***bold***.

1 MNSR**AYLTTN LVGGLGNQLF LVANL**LATSR **RNGVPAVLKR LPWSSSNEKP**
51 **RDVYWQSIFK DLDAYNVGRD VPVNLPEV**IV PEQR**PVAPVQ LDATANCLYN**
101 **MVGFF**QSEAF FADHPVVR**DV LPSELRNAAA KHLYDNYGEP STLQHVGLHI**
151 **RRGDYLRMRD VFEILEVDYY DAAVRQLLGH SLQLQSMGRG GEGRGGAHLL**
201 **VFCEEERYGL SVVGYFRSKY PGLEA**SLVRA **AGERVPFECG SHAPREVLEL**
251 **LMLSSCD**DVV MSNSTWSWWA AYFNFR**PLRR VVAPSKWFVQ HPYPQSNHLY**
301 **CENW**ILLENL YFQGDGR**MSK GEELFTGVVP** ILVELDGDVN GHK**FSVRGEG**
351 **EGDATNGKLT LKFICTTGKL PVPWPTLVTT LTYGVQ**CFSR **YPDHMKRHDF**
401 **FKSAMPEGYV QERTISFKDD GTYKTRAEVK FEGDTLVNRI ELKGIDFKED**
451 **GNILGHKLEY NFNSHNVYIT ADKQKNGIKA NFKIRHNVED GSVQLADHYQ**
501 **QNTP**IGDGPV LLPDNHYLST QSVLSK**DPNE KRDHMVLLEF VTAAGITHGM**
551 **SGLNDI**FEAQ **KIEWHEHHHH HHHH**

Figure 4.19: Mascot software output mapping the detected peptides from trypsin digestion of the band indicated in the red box of (Fig. 4.18A) to input sequence rTcFUT1-4, indicating 81% sequence coverage (in bold type).

Despite the impurity of the preparation, we decided to proceed to perform a FUT activity assay, based on previous work with TbFUT1 showing that activity was not necessarily compromised by the presence of co-eluted proteins (Bandini *et al.*, 2021).

We decided to perform in parallel activity assays in which either the donor or the acceptor substrates were radiolabelled with tritium. The first of these ("hot donor assay"), similar to that previously performed with bacteria-expressed TcFUT1, used a radioactive donor substrate, GDP-[^{3}H]Fuc, and an unlabelled acceptor substrate LNB (Fig. 4.12). The second assay ("hot acceptor assay") used unlabelled GDP-Fuc as donor substrate and tritiated lacto-N-tetritol as the acceptor substrate ([^{3}H]LNTol; Galβ1-3GlcNAcβ1-3Galβ1-4[1-^{3}H]glucitol, prepared by Michael Ferguson by the reduction of LNT (Fig.

4.12) with NaB[^{3}H]$_4$). Both LNB and LNT are known to be good acceptor substrates for *T. brucei* FUT1 (Bandini *et al.*, 2021).

The products of these reactions were separated on an HPTLC plate; radioactive products were visualized by fluorography, and unlabelled sugar and glycan standards by spraying the plate with orcinol/H_2SO_4 reagent and heating at 100ºC. Standards are necessary to correlate a band on the X-ray film with a particular sugar or glycan based on their retention factor. The result of one of these experiments showed that no reaction product was obtained when using the [^{3}H]LNTol acceptor substrate (Fig. 4.20A). However, there was a difference in intensities of the fast migrating signal between the sample and its negative control in the hot donor experiment (Fig. 4.20B). In these lanes we can see that the radioactive signal that co-migrates with the fucose standard (Fig. 4.20A) is much stronger when the hot donor (GDP-[^{3}H]Fuc) is incubated with partially purified rTcFUT1 than with buffer alone.

We hypothesized that this could be due to the effective hydrolysis of GDP-Fuc by rTcFUT1, i.e., the transfer of Fuc to water. Such a GDP-Fuc hydrolase activity has been already reported for TbFUT1 (Bandini *et al.*, 2021). As described before for the assays with bacteria-expressed protein, the GDP-[^{3}H]Fuc stock showed degradation due to long-term storage, explaining the presence of radiolysis products (R.P.) in these assays.

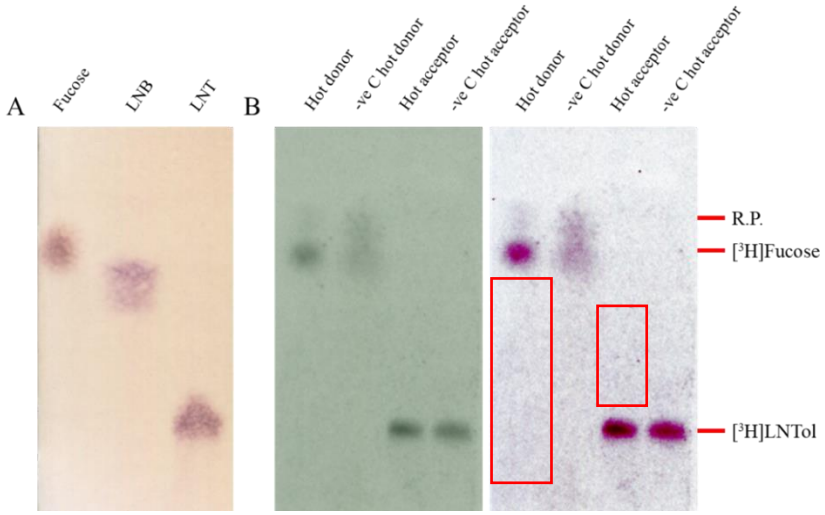


Figure 4.20: HPTLC analysis of rTcFUT1 (from rTcFUT1-4.a) hot donor and hot acceptor activity assays. (A) Orcinol/H_2SO_4 staining of plate with 2 µg of each sugar standard. (B) Developed X-ray film exposed against fluorographic TLC plate. Original image (left) and enhanced contrast image (right) are shown. R.P.: radiolysis product. Position where polar products should appear is framed in red.

At this point, our GDP-[^{3}H]Fuc stock was depleted. As an alternative to our radiochemical assays, we decided to confirm the effective GDP-Fuc hydrolytic activity (i.e., transfer of Fuc to water) of rTcFUT1-4.a recombinant product by using a commercial luminescence-based assay (GDP-glo) that can detect and quantify released

GDP during a GDP-sugar donor dependent glycosyltransferase reaction (Engel *et al.*, 2021). To test whether this system works under our conditions, we decided to first generate recombinant TbFUT1 as in Bandini's work (Bandini, 2011; Bandini *et al.*, 2021). Following her well-established procedure, rTbFUT1 from 1 L culture of *E. coli* BL21 (DE3) strain transformed with plasmid pGEX6P1-GST-PP-TbFT was purified by immobilised glutathione affinity purification, and the SDS-PAGE and Coomassie blue staining analysis (Fig. 4.21) showed a similar profile to that in Bandini's work, and from which band identifications were based.

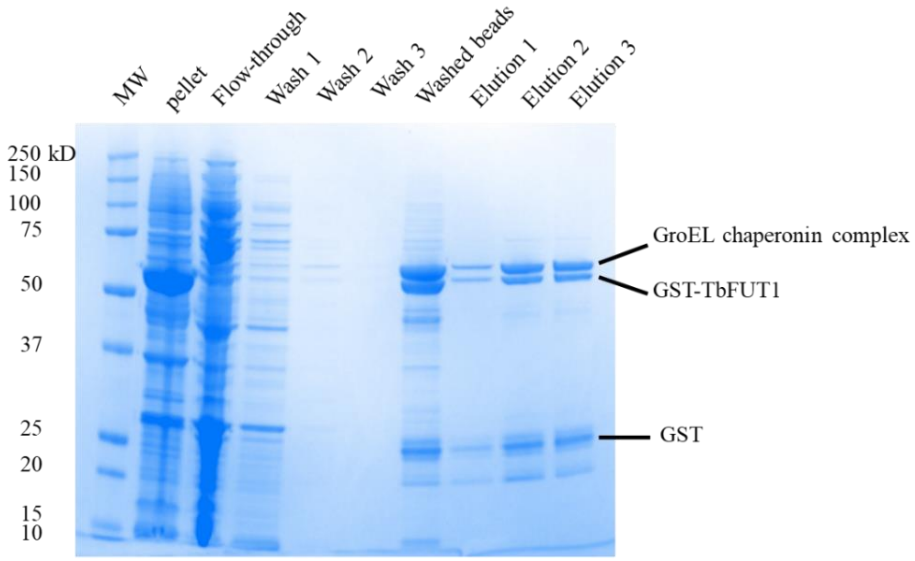


Figure 4.21: SDS-PAGE and Coomassie blue staining analysis of rTbFUT1 purification. Bands were identified based on Bandini's work (Bandini, 2011; Bandini *et al.*, 2021).

Purified rTbFUT1 and partially pure rTcFUT1-4.a product were then used to set up a GDP-glo assay. The effective hydrolysis of GDP-Fuc by recombinant TbFUT1 and TcFUT1 preparations was tested by performing the reactions without a glycan acceptor substrate. In this assay, a GDP standard curve is generated to calculate the amount of GDP generated in the reactions using the average luminescence value of three technical replicates, after normalizing against a blank. The result (Fig. 4.22) allowed us to validate this assay through the detection of released GDP by the reported GDP-Fuc hydrolytic activity of rTbFUT1. However, we did not detect any hydrolysis of GDP-Fuc in the presence of rTcFUT1 from rTcFUT1-4.a-transfected Expi293F cells. We concluded from this that our recombinant product rTcFUT1-4.a was not active.

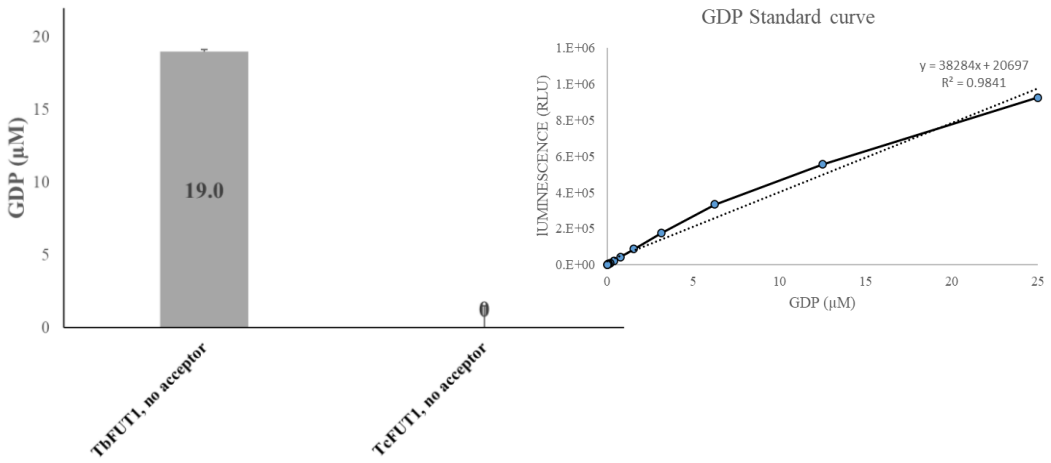


Figure 4.22: GDP-glo assay to confirm the hydrolytic activity of rTbFUT1 (from *E. coli*) and to test for hydrolytic activity in rTcFUT1 from rTcFUT1-4.a-transfected Expi293F cells. The bars show the detected free GDP concentration (µM) for each reaction (mean of three technical replicates, standard error of the mean is indicated). A GDP standard curve was generated to determine the GDP concentration based on the luminescence values obtained in the assay.

In a new attempt to obtain active rTcFUT1, we decided to clone the genome-amplified, truncated form of TcFUT1 (TcFUT1cat) into a commercially available vector (pSecTAG2A) designed for the secretion of the recombinant product out of Expi293F cells. Previous research has demonstrated its utility for the expression and purification of other parasite proteins (Chauché *et al.*, 2020; Meiners *et al.*, 2020). This new construct was named rTcFUT1-6.a (Fig. 4.23). At the N-terminus of TcFUT1cat, there is a CSS (as in constructs rTcFUT1-5.a and 5.b), and the protein is C-terminally tagged with Myc and 6xHis tags. The recombinant product has a predicted molecular weight of 38 kDa, and if processed by the removal of the CSS in the secretory pathway, of 35. 7 kDa.

Figure 4.23: Scheme of construct rTcFUT1-6.a (35.7 kDa). Used for expression of a secreted recombinant TcFUT1cat in transfected Expi293F cells.

As done before for the previous set of eukaryotic constructs, a first expression test was performed, analysing aliquots of insoluble pellet, soluble fraction, and extracellular medium by α-6xHis tag Western blot over the course of five days (from day 3 to day 7 post-transfection). Results from day 3, 5 and 7, depicted in (Fig. 4.24), suggested that the majority of the protein was being expressed in an insoluble form (lane I). Nevertheless, on day 5, a small signal from the extracellular medium aliquot (lane E) was detected, which made us decide to incubate the medium from transfected cells with α-Myc agarose beads to see whether we could capture any soluble, secreted rTcFUT1 from the culture supernatant (Fig. 4.25). From day 2 to 7 post-transfection, media was collected from independent transfections and incubated with α-Myc agarose beads. Both beads and media post-incubation were analysed by α-6xHis tag Western blot. Based on this result, we defined day 5 (Fig. 4.25, bold) as the best day post-transfection for obtaining soluble,

secreted rTcFUT1. Interestingly, two bands were detected, suggesting a possible difference in post-translational modifications present in the recombinant product. However, this idea was later dismissed when we detected the lower band alone in non-transfected cultures (Fig. 4.27), indicating that a protein of slightly lower apparent MW than rTcFUT1 was present in the medium and detected by the α-6xHis antibody.

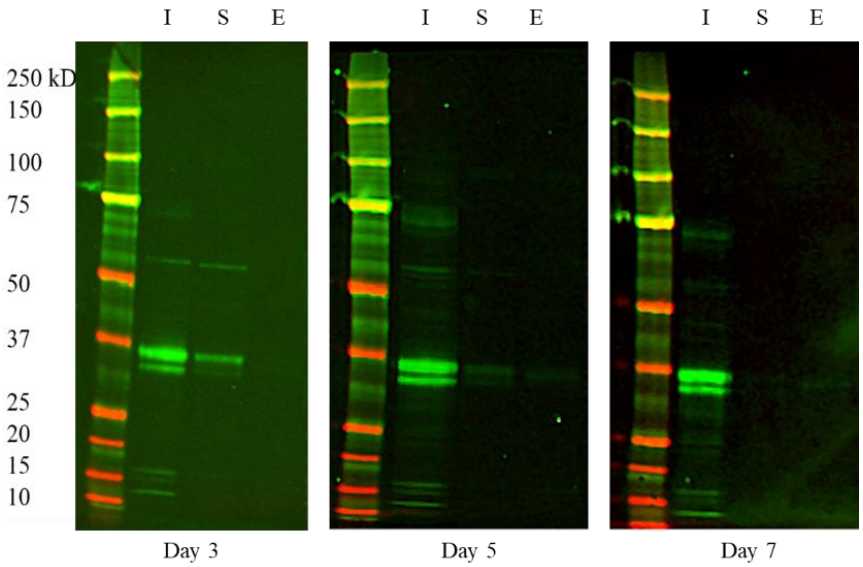


Figure 4.24: α-6xHis tag Western blot analysis of rTcFUT1-6.a (35.7 kDa) expression trial over the course of 5 days (from day 3 to day 7) post-transfection. I = insoluble pellet. S = soluble fraction. E = culture supernatant corresponding to extracellular medium.

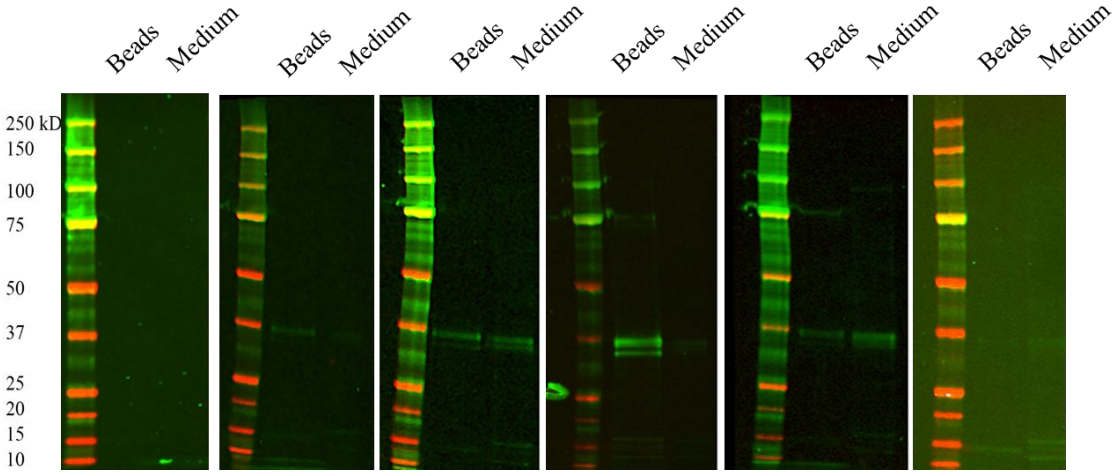


Figure 4.25: α-6xHis tag Western blot analysis of rTcFUT1-6.a (35.7 kDa) expression trial in the medium of Expi293F transfected cells over the course of 6 days (from day 2 to day 7) post-transfection. Beads = material eluted from α-Myc agarose beads with SDS sample buffer. Medium = aliquot of medium post-incubation with beads.

Based on these results, we tried to purify the secreted rTcFUT1 as previously reported (Chauché *et al.*, 2020) by performing an HPLC IMAC purification with the medium of a large-scale (2 L) transfection. However, analysis by SDS-PAGE Coomassie blue staining and α-6xHis tag Western blot indicated that although our protein was present in the medium (pre-injection) it did not all bind to the nickel column and that which did was not eluted with imidazole (Fig. 4.26A). Nevertheless, we decided to use an aliquot

of original transfected medium and column flow through to use in the GDP-glo assay to look for the potential GDP-Fuc hydrolytic activity of rTcFUT1. We detected some GDP in the assay performed with the whole medium but not with the column FT (Fig. 4.26B).

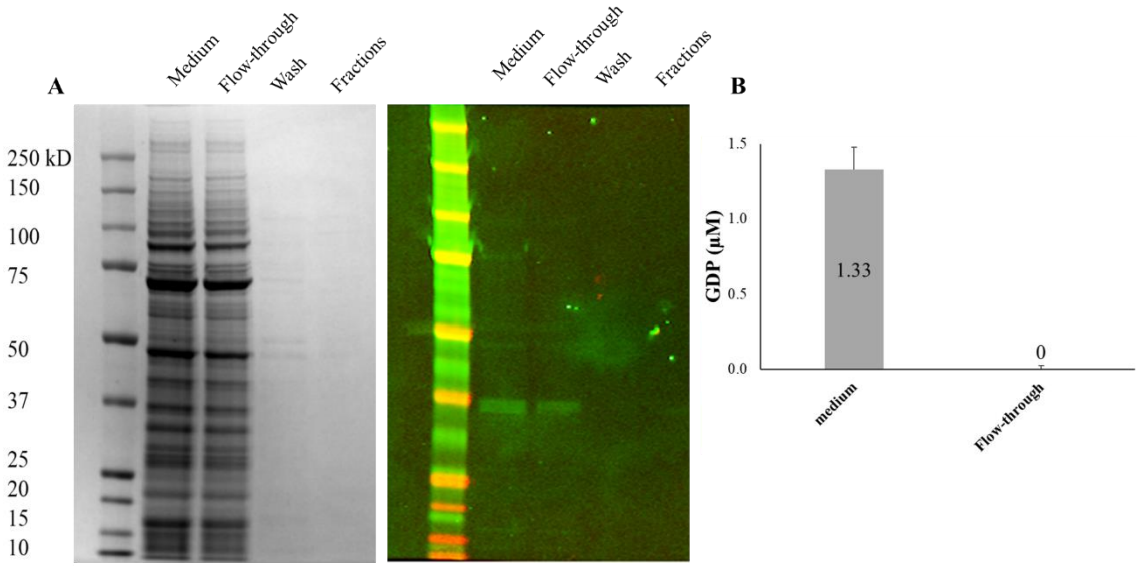


Figure 4.26: HPLC IMAC purification attempt of secreted rTcFUT1 (35.7 kDa) from 2 L of medium form transfected Expi293F cells, and GDP-glo activity assay. (A) SDS-PAGE Coomassie blue staining (left) and α-6xHis tag Western blot (right) analyses of medium pre-injection and fractions from imidazole elution. (B) GDP-glo assay with medium and FT from the purification attempt (mean of triplicate technical measurements with the standard error of the mean)

In order to determine if the GDP detected in (Fig. 4.26B) was due to the hydrolytic activity of secreted rTcFUT1, and not to free GDP in the medium, we decided to obtain the recombinant product using α-Myc beads in an IP assay as an alternative purification strategy. In this experiment, rTcFUT1-containing medium aliquots were incubated with α-Myc beads. Also, a non-transfected medium was used as a negative control. Elution of the α-Myc bead IPs with Myc peptide from two commercial sources (named “elution A” and “elution B”) showed that the majority of rTCFUT1 remained bound to the beads with some being eluted in elution A and little eluted in elution B (Fig. 4.27A). Both eluates (A and B) and non-eluted beads with captured rTcFUT1 were used in a GDP-glo assay with GDP-Fuc and without glycan acceptor (Fig. 4.27B).

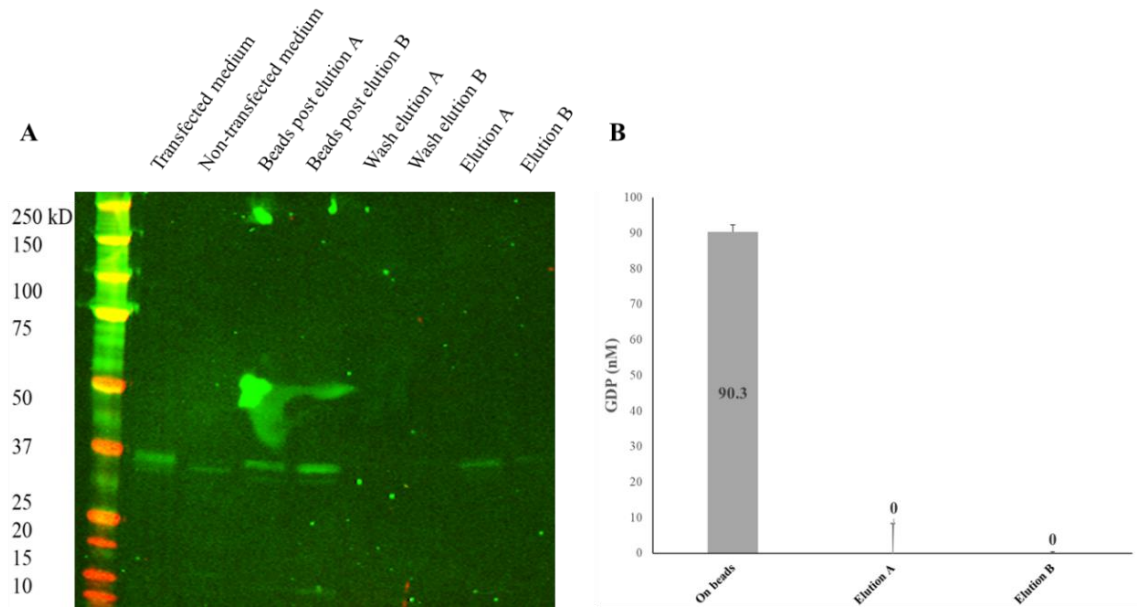


Figure 4.27: α-Myc beads IP and GDP-glo assay with bead-bound rTcFUT1 (35.7 kDa) and Myc-peptide eluates. (A) α-6xHis tag Western blot analysis of medium pre-incubation with beads, non-transfected medium and different steps in elution attempts. Note: the "non-transfected medium" lane contains a lower molecular weight band that reacts with α-6xHis tag antibodies that may also be present in previous experiments (Fig. 10 and Fig. 11). (B) GDP-glo assay with rTcFUT1-captured beads ("On beads") and Myc-peptide eluates ("elution A" and "elution B"). The data are means of three technical replicates and standard error of the mean is indicated.

Although the Myc-peptide eluates appeared to be inactive, the bead-bound rTcFUT1 did appear to be able to turnover GDP-Fuc to GDP. Together, these results lead us to suggest that although we are not able to purify secreted rTcFUT1 efficiently, we could at least capture it in its active form on α-Myc beads and measure its activity by the GDP-glo assay.

Knowing this, we decided to analyse whether the GDP-Fuc hydrolytic activity was affected by the presence of putative glycan acceptors. Thus, media from more cultures transfected with construct rTcFUT1-6.a were incubated with α-Myc agarose beads, and the washed beads were resuspended in GDP-glo assay buffer supplemented with GDP-Fuc with or without various synthetic glycan acceptors, as defined previously (Fig. 4.12). All these reactions, along with a negative control, were then used for free GDP quantification (Fig. 4.28).

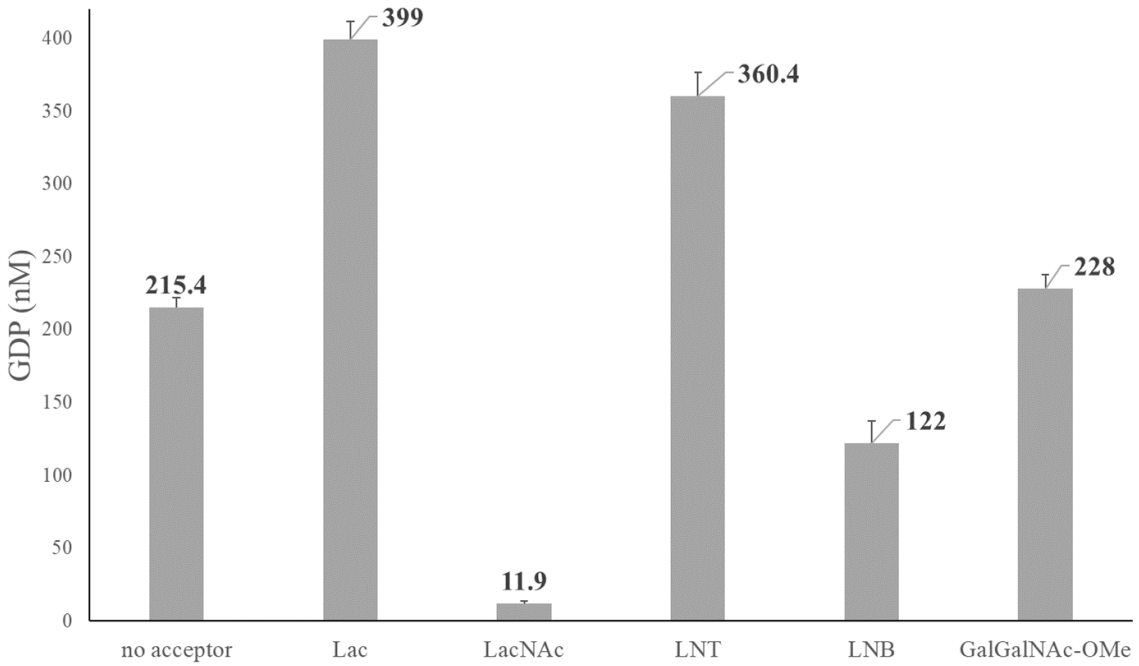


Figure 4.28: GDP-glo assay with α-Myc bead-bound rTcFUT1 and different sugar acceptors. Standard error was calculated and added to the bars showing the detected free GDP (nM) for each reaction. Result was calculated from the average of three technical replicates.

The result from this, and another experiment with similar results, confirmed that we can generate secreted, soluble, and active rTcFTU1, which can be then bound to α-Myc agarose beads and used in GDP-glo activity assays which show an effective hydrolytic activity from the enzyme on GDP-Fuc, and a variation on the free GDP generated when different sugar acceptors are present. These may be acting either as substrates or inhibitors by increasing or reducing the rate of turnover of GDP-Fuc to GDP, the latter most likely by competing with water molecules for access to the active site. To probe this further, we wanted to return to the radiochemical FUT1 assay where it is possible to observe and characterise any [^{3}H]fucosylated-glycan acceptor reaction products. Unfortunately, at this time our commercial GDP-[^{3}H]Fuc stocks were depleted and more GDP-[^{3}H]Fuc (which had been on-order for over a year) was unavailable. We therefore elected to synthesise our own GDP-[^{3}H]Fuc (section 4.10).

Once we had successfully synthesized and purified GDP-[^{3}H]Fuc, we decided to check it by performing a FUT activity assay using a fresh preparation of *E. coli*-expressed rTbFUT1 (similar to that shown in Fig. 4.21). This was incubated with 400 nCi of synthesized GDP-[^{3}H]Fuc in the presence of 1 mM LNB as the acceptor saccharide, which are lower concentrations than in (Bandini *et al.*, 2021). The result (Fig. 4.29) reproduced the reported activity of rTbFUT1: In the absence of LNB, rTbFUT1 turned over GDP-[^{3}H]Fuc to free [^{3}H]Fuc. In the presence of LNB, turnover to free [^{3}H]Fuc is reduced and [^{3}H]fucosylated products of LNB are clearly visible migrating just below

LNB itself (compare fluorograph with orcinol stained HPTC plate). The presence of two product bands is because the α- and β-anomers of LNB and [^{3}H]Fuc-LNB separate on this HPTLC system. Using our home-made fresh GDP-[^{3}H]Fuc, no signs of radiolysis products were detected (which run ahead of free Fuc) but direct analysis of an aliquot of our GDP-[^{3}H]Fuc preparation showed a trace of a low mobility component that we think may be a [^{3}H]Fuc-1-pyrophosphate degradation product.

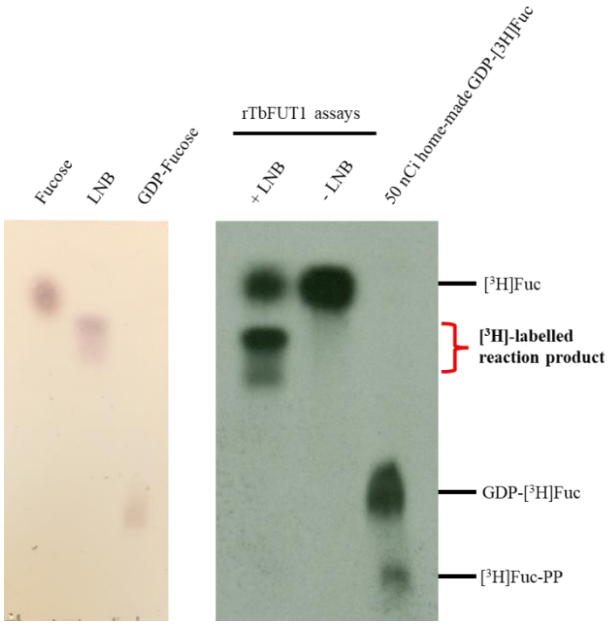


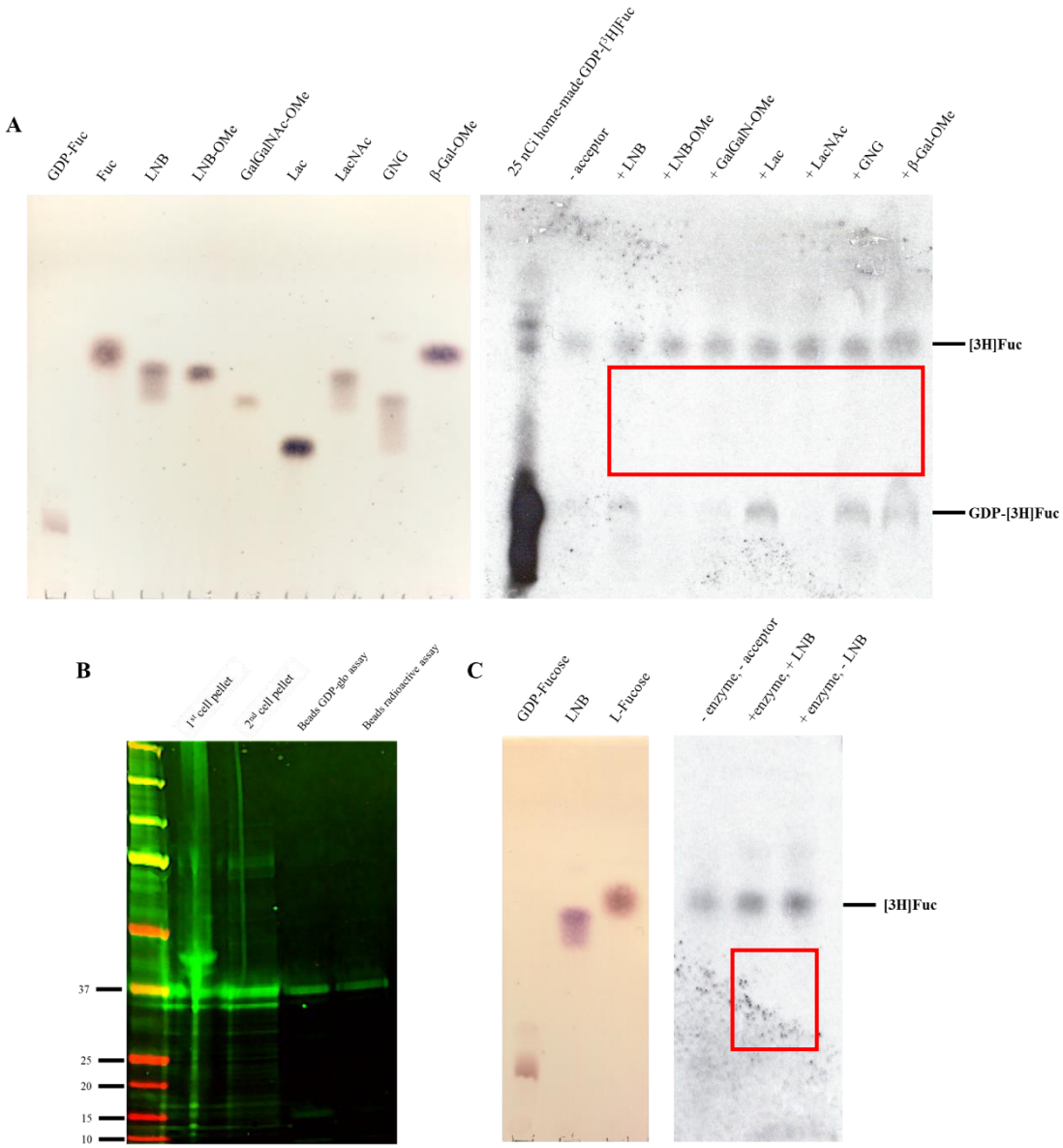
Figure 4.29: *E.coli*-expressed rTbFUT1 activity assay with home-made GDP-[^{3}H]Fuc and LNB. Radioactive products were detected by fluorography (right panel) and sugar standards by orcinol/H_2SO_4 staining (left panel). A 50 nCi aliquot of the synthesized sugar donor was also included in the HPTLC analysis.

Based on these results, we considered that we had validated the use of our synthesized GDP-[^{3}H]Fuc for its use in assays. Thus, secreted rTcFUT1 was expressed in rTcFUT1-6.a-transfected Expi293F cells, the extracellular medium incubated with α-Myc agarose beads, and these were then used to set up on-bead reactions including either 200 nCi GDP-[^{3}H]Fuc alone or in the presence of different glycan acceptors at 1 mM (Fig. 4.12). This first attempt was analysed by HPTLC (Fig. 4.30A) and showed that no clear product was generated in any reaction, based on the lack of fluorographic signal with similar (slightly lower) retention factor as any of the acceptors stained with orcinol.

Further, although [^{3}H]Fuc was present in all reactions, it could not be determined if it was a product of the hydrolytic activity of the recombinant protein, or due to the partial degradation of the synthesized sugar donor, since it showed signs of degradation to radiolabelled L-Fuc, as can be seen in the first lane of the X-ray film. Also, some of the reactions had traces of GDP-[^{3}H]Fuc, most probably due to an incomplete desalting process while preparing the samples for TLC spotting.

A second attempt was performed, in which prior to setting up the reactions, the hydrolytic activity of the bead-bound rTcFUT1 was checked by a GDP-glo assay (using a 2 h rather than overnight incubation). The assay indicated the generation of free GDP at a concentration of 3.53 nM. Knowing that the bead-bound recombinant protein was showing GDP-Fuc hydrolytic activity, on-bead radiochemical reactions were immediately performed using our home-made GDP-[^{3}H]Fuc donor substrate, with and without LNB potential acceptor substrate, and a negative control in which neither enzyme nor acceptor were added to control for any GDP-[^{3}H]Fuc degradation into [^{3}H]Fuc under the conditions used. As shown in (Fig. 4.30B), the enzyme preparation was checked by α-6xHis tag Western blot. The cell pellet analysis when obtaining the extracellular medium showed that part of rTcFUT1 was expressed inside the cells. Also, the lower band detected in previous experiments (Fig. 4.24 and Fig. 4.25) and suggested to be a contaminant (Fig. 4.27) was detected. Further, analysis of pellet after a second high-speed centrifugation showed that more contaminant and most (probably aggregated) rTcFUT1 were cleared by this (Fig. 4.30B, "2nd pellet" lane). The beads used for both the GDP-glo and radioactivity-based assays were also analysed, showing that only the higher band, assumed to be rTcFUT1was bound to the beads and therefore present in the assays. Analysis by fluorography and orcinol staining of the HPTLC plate (Fig. 4.30C) showed an increased presence of free [^{3}H]Fuc in the "+ enzyme" reactions and less free [^{3}H]Fuc in the "+ LNB" compared to "- LNB" reactions, suggesting there may be some reduction in GDP-[^{3}H]Fuc hydrolysis in the presence of LNB. Nevertheless, longer (2 months) exposure of the X-ray film did not show the appearance of any [3^H]fucosylated-LNB products.

Figure 4.30 (*next page*): rTcFUT1-6.a product on-bead FUT activity assays with home-made GDP-[^{3}H]Fuc. (A) First attempt, using the set of glycan acceptors and sugar donor alone. Position where products should appear is framed in red. (B) α-6xHis tag Western blot analysis of bead-bound rTcFUT1 preparation prior to GDP-glo and glycosyltransferase assays. (C) Second attempt, using a negative control to detect the degradation of the home-made sugar donor, and two reactions with and without LNB acceptor. Position where products should appear is framed in red. For (A) and (C), radioactive products (right panel) were detected by fluorography and sugar standards (left panel) by orcinol/H_2SO_4 staining.



Altogether, these results suggested that despite our efforts, rTcFUT1-6.a-transfected Expi293F cells were expressing a soluble, potentially hydrolytic, yet not fully active recombinant protein. As a final attempt to achieve this aim of the project, we decided to set up a homologous overexpression of TcFUT1 in *T. cruzi* itself, hoping to fulfil any requirements this protein may have in order to be fully functional, such as the interaction with potential partners or a necessary molecular machinery which may only be present in its original environment.

4.8 Homologous overexpression of *TcFUT1*.

To achieve homologous overexpression of TcFUT1, a synthetic plasmid was purchased containing a C-terminal 6xMyc-tagged sequence of full-length *TcFUT1* with an NdeI restriction site separating the tag and ORF (TcFUT1-NdeI-6xMyc). This was

cloned into the *T. cruzi* expression vector pTREX-G418r (kind gift from Susan Wyllie, School of Life Sciences, University of Dundee). This vector contains a polymerase I ribosomal RNA (rRNA) promoter and a short sequence (HX1) from the upstream region of locus H1.8 of the ribosomal protein gene *TcP2β* (Vazquez and Levin, 1999), in order to induce a strong expression of the gene of interest, and an integration of the plasmid in the 18S rRNA gene locus (Lorenzi, Vazquez and Levin, 2003). This process is independent of using circular or linearized molecules during electroporation (Kelly *et al.*, 1992; Martínez-Calvillo, López and Hernández, 1997; Vazquez and Levin, 1999). Using restriction enzyme-based cloning, TcFUT1-NdeI-6xMyc was cloned between the HX1 sequence and a 5'-untranslated region (5'UTR) of the *T. cruzi* glycosomal glyceraldehyde-3-phosphate dehydrogenase tandem genes (gGAPDH) present in the vector. The selection gene for neomycin resistance (G418r) was included in the vector and located between the gGAPDH 5'UTR and 5'-untranslated region (3'UTR) sequences. In this way, construct pTREX-G418r-TcFUT1-6xMyc, or TcFUT1-7.a (Fig. 4.31) was created.

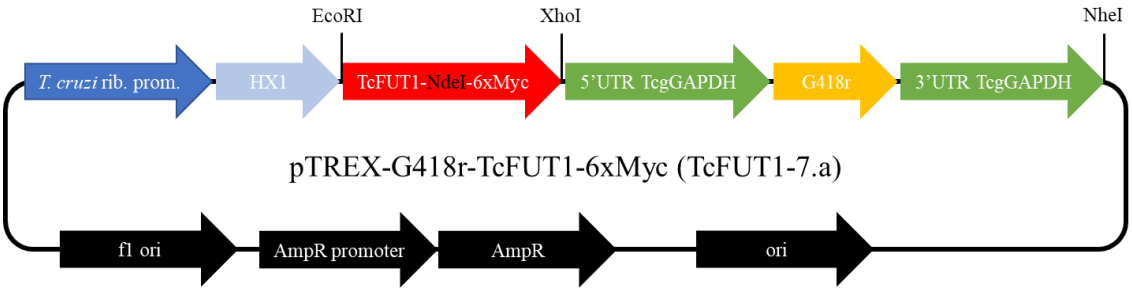


Figure 4.31: Scheme of construct TcFUT1-7.a for overexpression of TcFUT1 in *T. cruzi*. The tagged protein has a predicted molecular weight of 42.2 kDa.

Either circular DNA or pre-linearized by NheI digestion preparations were used for electroporations. After incubation for 3 weeks in drug-containing medium, all four transfected populations recovered. In order to confirm the overexpression of tagged TcFUT1, we performed α-Myc Western blot analysis using lysates of the drug-resistant cultures. As shown in (Fig. 4.32, left panel), all four drug-resistant mixed populations (MPs) showed the detection of a band corresponding to the predicted molecular weight of TcFUT1-6xMyc (42.2 kDa). Additionally, in the lane of MP1, a second band of lower MW was also detected. Although there were multiple detected bands in the WT lane, we considered them to be non-specific due to the difference in loading, which is appreciated in the Ponceau S staining of the membrane prior to blotting (Fig. 4.32, right panel).

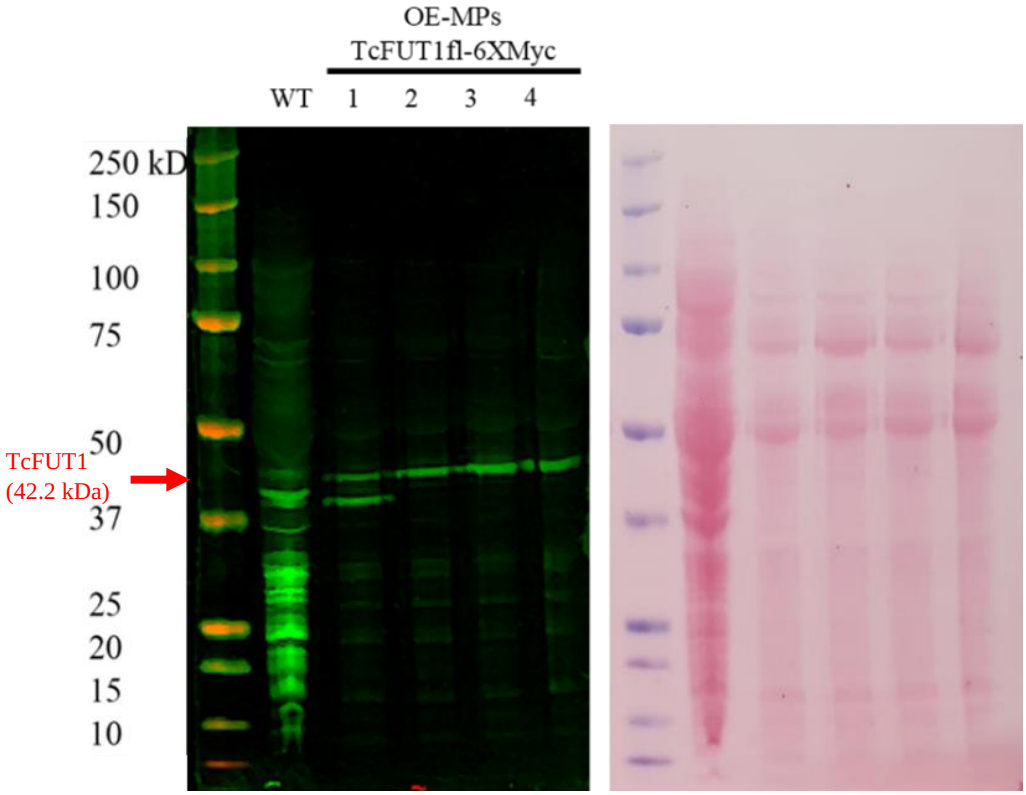


Figure 4.32: α-Myc tag Western blot analysis of WT and potential TcFUT1 overexpressing mixed populations (OE-MPs) (left panel) and Ponceau S pre-staining of the same membrane (right panel).

TcFUT1 contains a predicted MTS, which suggested to us that the overexpressed TcFUT1 product in the MP1 sample could be proteolytically processed by internalisation into the mitochondria and thus detected as an additional band of slightly lower molecular weight. Therefore, we proceeded to subclone by limiting dilution the population MP1, along with one of the other MPs showing only one band in the Western blot analysis (MP4). Despite multiple attempts and two months of incubation, we were not successful on obtaining a clonal population.

Nevertheless, as the MPs could be expressing an active TcFUT1, we decided to perform IPs by incubating lysates of the MP1 and MP4 populations with α-Myc agarose beads. The protein-bead complexes from these IPs were used for on-bead FUT reactions with either 300 nCi of home-made GDP-[^{3}H]Fuc alone or in the presence of 1 mM LNB as acceptor substrate. Different steps during the IP were analysed by SDS-PAGE Coomassie blue staining (Fig. 4.33, left panel) and α-Myc Western blot (Fig. 4.33, right panel), showing that part of the overexpressed product was insoluble in both MPs, and that only the lower MW band was found in the insoluble fraction for MP1. Although some material did not bind, most of the fluorescence signal came from the beads, indicating the presence of TcFUT1-6xMyc bound to them prior to set up of the reactions, even when this was not the main protein species present, as indicated by Coomassie blue staining of the gel lanes containing the beads (Fig. 4.33, left panel). Importantly, the 2 band/1 band

phenotypes were maintained for MP1 and MP4, respectively, as seen previously (Fig. 4.32, left panel).

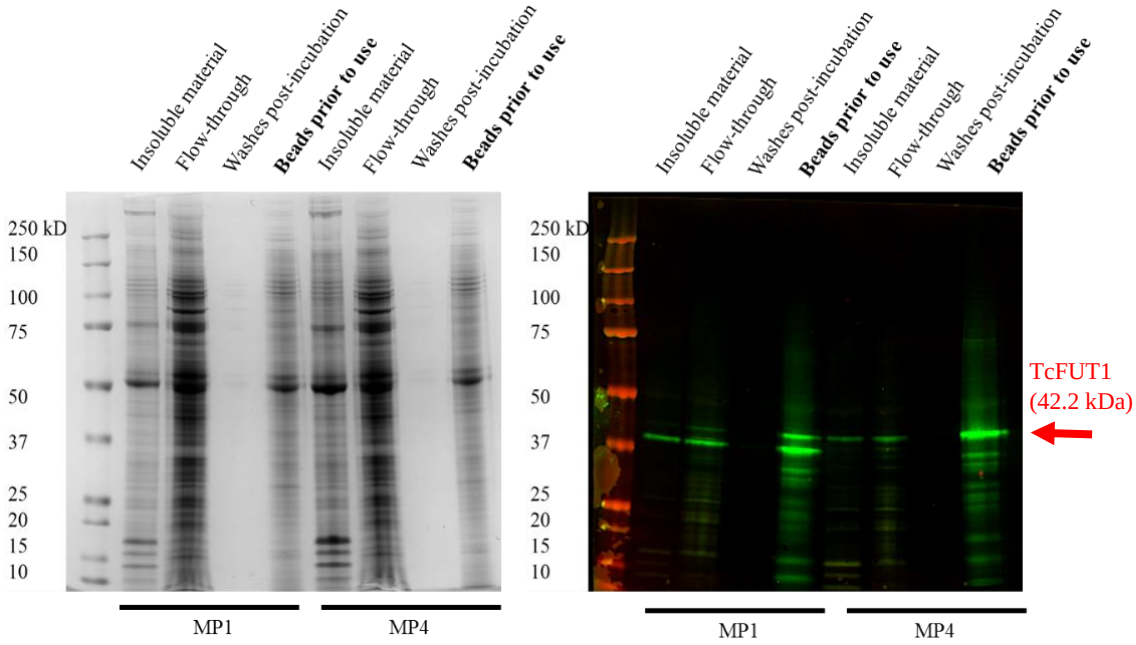


Figure 4.33: SDS-PAGE Coomassie blue staining (left panel) and α-Myc tag Western blot analysis (right panel) of TcFUT1-6xMyc IP of *T. cruzi* MP1 and MP4 lysates.

Subsequently, the fucosyltransferase reaction products were separated by HPTLC and analysed by fluorography and orcinol staining (Fig. 4.34) showing, for the first time, the fluorographic detection of a reaction product when the sugar acceptor LNB was included in the reaction.

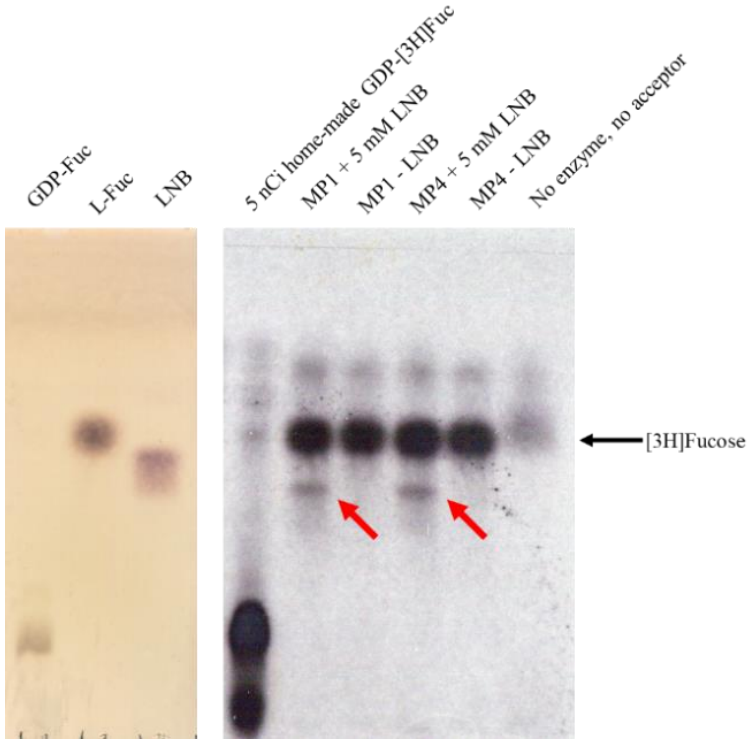


Figure 4.34: TcFUT1-6xMyc from MP1 and MP4 activity assays with home-made GDP-[^{3}H]Fuc and LNB. Radioactive products were detected by fluorography (right panel) and sugar standards by orcinol/H_2SO_4 staining (left panel). A 5 nCi aliquot of the synthesized sugar donor was also included in the HPTLC analysis. Reaction products are marked with red arrows.

When no enzyme was present, some [^{3}H]Fuc was still detected, indicating some degradation of the home-made sugar donor. The [^{3}H]Fuc fluorographic signal was much stronger when enzyme was present, indicating a hydrolysis/water transfer activity of the overexpressed protein. Only when the LNB acceptor was present, a spot running slightly below the LNB appeared, which indicated the generation and detection of a trisaccharide product after transferring the [^{3}H]Fuc to the disaccharide LNB (Fig. 4.34, red arrows). Interestingly, both enzyme preparations, despite their different phenotypes, showed similar activity based on the intensity of the TLC band corresponding to the reaction product. This result is similar to the one reported using TbFUT1 (Bandini *et al.*, 2021).

Now that we were able to obtain active rTcFUT1 and perform FUT assays, we decided to try to define the substrate specificity of the overexpressed rTcFUT1 by incubating the protein-bound IP beads with the panel of acceptor substrates (Fig. 4.12). For this experiment, lysates from MP1 overexpressing parasites were used along with 400 nCi of home-made GDP-[^{3}H]Fuc and 1 mM acceptors (Fig. 4.35).

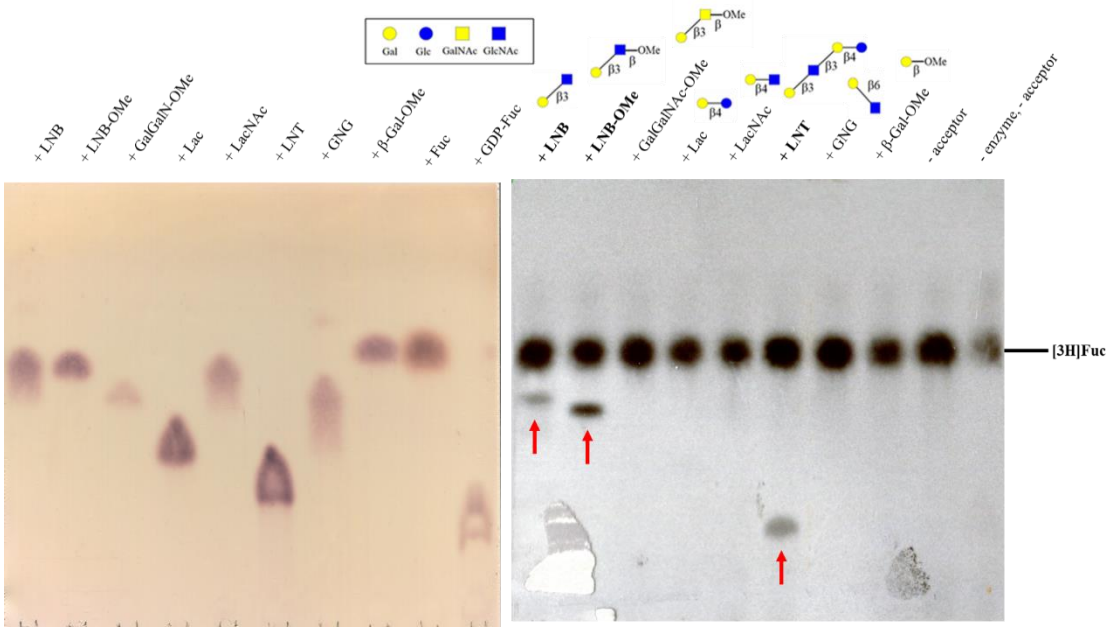


Figure 4.35: MP1 TcFUT1-6xMyc substrate specificity assay with home-made GDP-[^{3}H]Fuc and panel of glycan acceptors. Positive reactions are marked in bold. Radioactive products were detected by fluorography (right panel) and sugar standards by orcinol/H_2SO_4 staining (left panel). Reaction products are marked with red arrows. Schemes of the assayed acceptors and key are shown above the TLC lanes.

The result (Fig. 4.35) indicated the presence of radiolabelled products (marked with red arrows) when the acceptors LNB (Galβ1-3GlcNAc), its β-O-methyl glycoside LNB-OMe (Galβ1-3GlcNAcβ1-O-Me), and LNT (Galβ1-3GlcNAcβ1-3Galβ1-4Glc) were present in the reaction. Based on the fluorographic intensity of the product bands, TcFUT1 showed its best activity with LNB-OMe. Its non-methylated version LNB and the larger oligosaccharide LNT were also good acceptors. Based on the chemical

structures of the assayed acceptors, TcFUT1 was able to transfer [^{3}H]Fuc to carbohydrate acceptors only when Galβ1-3GlcNAc structures, either alone or as a non-reducing terminal moiety, were present. This did not happen when the acceptors were galactose methyl glycoside alone (Galβ1-OMe), or if they contained a βGal residue bound to other positions (4 or 6) of GlcNAc or to the 4-position of Glc (see results with LacNAc, GNG and Lac in (Fig. 4.35). Interestingly, when a βGal residue was bound to the 3-potion of GalNAc as in the GalGalNAc-OMe acceptor, no product was detected either, suggesting a specificity for the presence of a GlcNAc residue in acceptor substrate structure. As in previous assays, some degradation of the home-made sugar donor was detected since when no enzyme or acceptor were present, [^{3}H]Fuc was still detected. An increase in the fluorographic signal of [^{3}H]Fuc when only the enzyme was present further confirmed the ability of the protein to transfer Fuc to water.

Although we did not have time for further experiments, the continuation of this research would be an initial characterization of the reaction products by chemical and enzymatic treatments, such as TFA acid hydrolysis and α1-2 exofucosidase treatments (see section 6.1), and, ideally, then a scale-up of the reactions, product purification by normal phase HPLC and fine chemical characterization by proton nuclear magnetic resonance (^{1}H-NMR), electrospray ionization mass spectrometry (ESI-MS) and GC-MS methylation linkage analysis, as done previously for the analysis of TbFUT1 reaction products (Bandini *et al.*, 2021).

4.9 TcFUT1 immunolocalization

Previously, the mitochondrial localization of TbFUT1 was determined by indirect fluorescent antibody microscopy by co-localizing both C-terminally Myc-tagged TbFUT1 and native TbFUT1 with mitochondrial dye MitoTracker™ and with mitochondrial ATPase (Bandini *et al.*, 2021).

Initial immunolocalization microscopy attempts using α-TbFUT1 rabbit polyclonal antibody were not conclusive, either because TcFUT1 is not well recognised by the anti-TbFUT1 antibody and/or because native TcFUT1 is below the limit of detection with our anti-TbFUT1 preparation, which had been in storage for several years.

We therefore used two of our mixed populations overexpressing TcFUT1-6xMyc (MP1 and MP4) and α-Myc antibodies to perform IFA microscopy along with MitoTracker™ staining of the mitochondrion. Because MitoTracker™ gives a very bright signal in *T. cruzi* epimastigotes, we established conditions where there was no detectable signal leakage in the green channel from the very bright red (MitoTracker™) channel with a no α-Myc 1st antibody control experiment (Fig. 4.36A). Using these conditions, we then performed co-localisation analysis with α-Myc and Mitotracker™ for MP1 and MP4 cells (Fig. 4.36B).

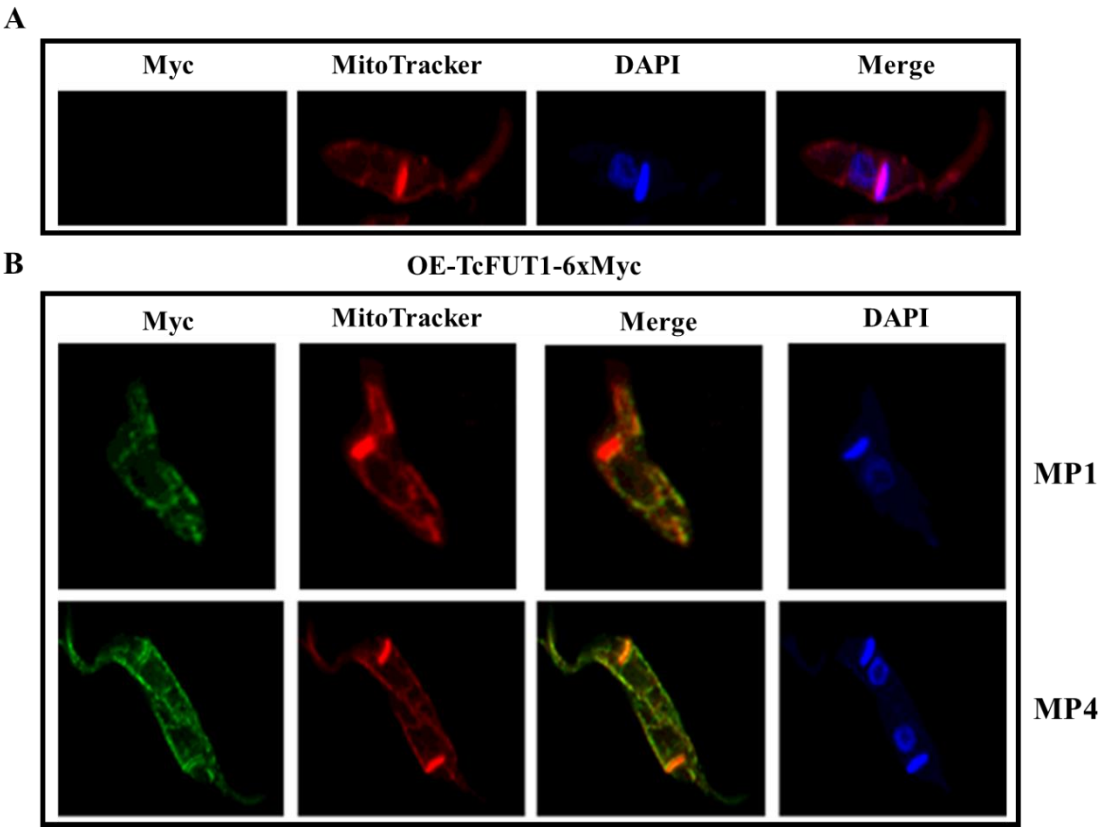


Figure 4.36: IFAs of MP1 and MP4 *T. cruzi* TcFUT1 overexpressing populations to define TcFUT1-6xMyc subcellular localization. (A) Preparation with no α-Myc staining used to define the exposure time at which no signal was detected in the green channel from crossover from the red channel. These conditions were used for the analyses in panel B. (B) MP1 (upper panels) and MP4 (lower panels) *T. cruzi* cells stained with α-Myc after being treated with MitoTracker, and counterstained with DAPI.

Our results using MitoTracker™ were similar to those reported by other authors, for example (Peloso *et al.*, 2016; Bustos *et al.*, 2017; Chiurillo *et al.*, 2019), whereby a very bright signal is observed corresponding to the location of the kinetoplast, along with less bright staining of a reticulated perinuclear network. The α-Myc staining for TcFUT1-6xMyc in these cells gave a similar reticulated labelling but a less pronounced labelling of the kinetoplast itself.

The intense labelling of the kinetoplast by MitoTracker™ in *T. cruzi* epimastigotes, to our knowledge, does not occur in *T. brucei* (bloodstream forms or procyclic forms) or in *Leishmania* spp. promastigotes. Further, it is difficult to reconcile kinetoplast staining with the primary mechanism by which Mitotracker™ selectively stains viable cell mitochondria, i.e., by detecting mitochondrial membrane potential. Nevertheless, this is clearly a feature of MitoTracker™ staining in *T. cruzi* epimastigotes. Interestingly, while some labelling of the kinetoplast with the α-Myc antibody was apparent, this staining (in contrast to that of the small molecule MitoTracker™) was excluded from the interior of the kinetoplast (see MP4 images in Fig. 4.36B).

Despite the unexplained anomalous staining of the kinetoplast with MitoTracker™, we suggest that the results in (Fig. 4.36) are at least consistent with a mitochondrial location for TcFUT1-6xMyc. However, this is not a definitive result and further experiments with well characterised antibodies to *T. cruzi* mitochondrial, cytosolic and endoplasmic reticulum proteins are required to confirm this hypothesis.

4.10 Synthesis of GDP-[^{3}H]Fuc

Due to the necessity of GDP-[^{3}H]Fuc to perform on-bead radiochemical fucosyltransferase assays with rTcFUT1-6.a recombinant protein (to detect the generation of products and validate the preliminary GDP-glo assay results) and the unavailability of commercial GDP-[^{3}H]Fuc, we decided to synthesize our own home-made radioactive sugar donor.

Previous research described the synthesis of the sugar donor GDP-[5-^{3}H]D-arabinopyranose using an enzymatic kinase reaction involving a crude enzymatic preparation from *C. fasciculata* lysates ([5-^{3}H]D-Ara -> [5-^{3}H]D-Arapα1-P) followed by chemical condensation with GMP-morpholidate ([5-^{3}H]D-Arapα1-P -> GDP-[5-^{3}H]D-Arap) (Schneider *et al.*, 1995), discussed in section 6.2. Since D-Ara is structurally very similar to L-Fuc, differing only in an extra sixth carbon in the methyl group bound to the C5 of the L-fucopyranose ring, and since the high-molecular weight (>100 kDa) cytosolic AFKP80/FKP40 protein dimers have been identified to harbour both kinase and pyrophosphorylase activities (Novozhilova and Bovin, 2009) (see sections 5.1 and 6.2), we decided to obtain a crude cytosolic preparation of >100 kDa proteins from *C. fasciculata* lysates, which would then be used to synthesize GDP-[^{3}H]Fuc, using commercial [^{3}H]Fuc.

An initial, small-scale experiment (using 1 μCi [^{3}H]Fuc) was performed (Fig. 4.37), suggesting the successful preparation of cytosolic activity that converted [^{3}H]Fuc into [^{3}H]Fuc-1-PO_4 ([^{3}H]Fuc-1P) and/or GDP-[^{3}H]Fuc (Fig. 4.38A and B, left panels), according to the retention factors of Fuc-1P and GDP-Fuc in the orcinol-stained HTLC plate with standard sugars (Fig. 4.37A and B, right panels). However, due to the similar polarity of Fuc-1P and GDP-Fuc, the separation of these two species proved to be challenging.

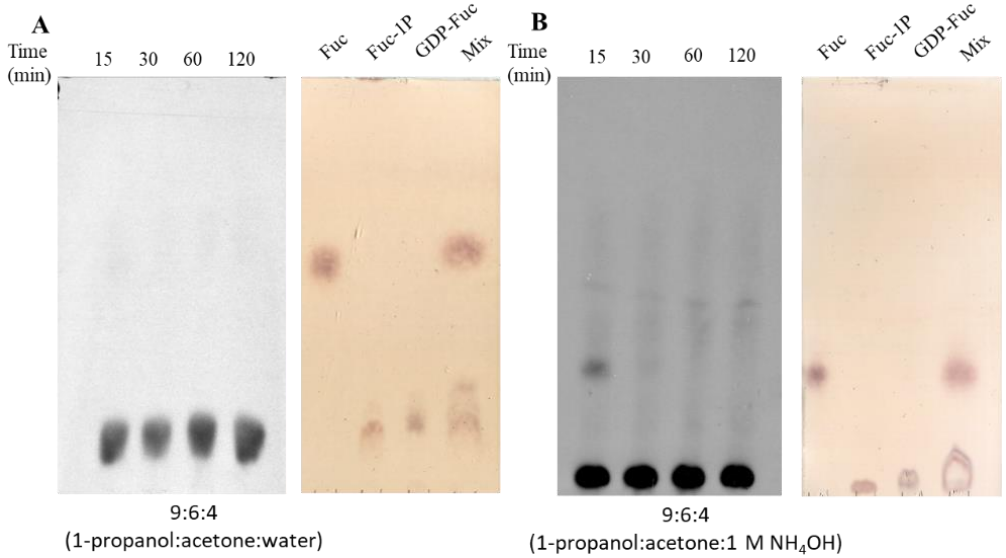


Figure 4.37: First GDP-[^{3}H]Fuc synthesis trial. Cytosolic activity was prepared as detailed in section 3.7.10. 1 μCi [^{3}H]Fuc was used, and aliquots at 15, 30, 60 and 120 min were detected by fluorography after separation on an HPTLC plate (left panels). Sugar standards and a mix of the three of them were run in parallel and stained with orcinol/H_2SO_4 (right panels). Two solvents, either (A) 1-propanol:acetone:water (9:6:4) or (B) 1-propanol:acetone:1 M NH_4OH (9:6:4), were tested to separate Fuc-1P and GDP-Fuc.

Nevertheless, the result suggested a rapid (30 min) conversion of [^{3}H]Fuc into either its phosphorylated form, nucleotide form, or both. In order to improve resolution between Fuc1-P and GDP-Fuc, we investigated using consecutive runs of the same plate to achieve a better separation (Fig. 4.38). As seen in (Fig. 4.38B) two runs with solvent 1-propanol:acetone:water (9:6:4, v:v:v) gave the best overall result, and it was established as the method to analyse further GDP-[^{3}H]Fuc synthesis trials.

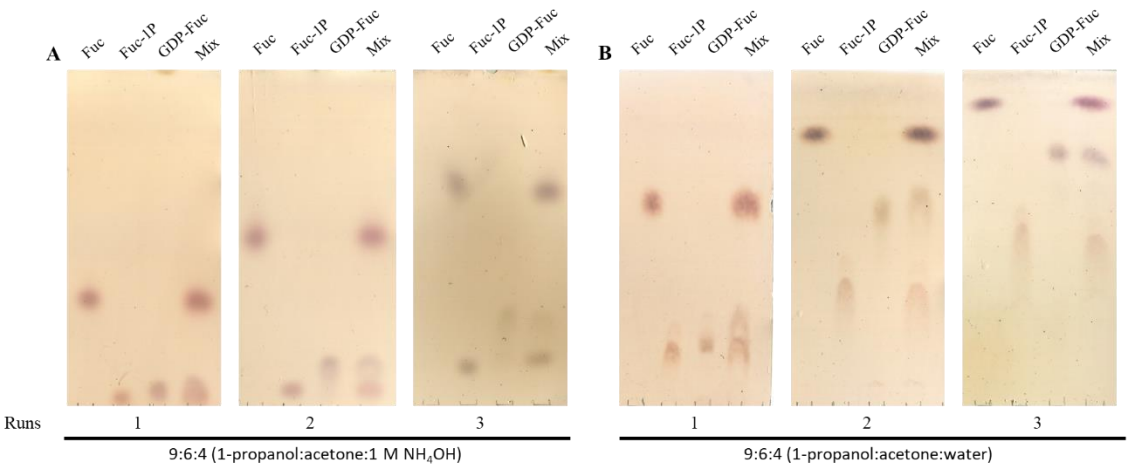


Figure 4.38: Two consecutive TLC runs to achieve better separation of sugars of interest (Fuc, Fuc-1P and GDP-Fuc) by testing mobile phases (A) 1-propanol:acetone:1 M NH_4OH (9:6:4) or (B) 1-propanol:acetone:water (9:6:4). Plates were stained with orcinol/H_2SO_4.

While reviewing the relevant bibliography, we found a similar attempt to synthesize tritiated sugar donors (Mengeling and Turco, 1999). Incorporating aspects of this work to our trials (particularly the inclusion of a pyrophosphatase to degrade pyrophosphate product in the Fuc-1P +GTP -> GDP-Fuc + PPi step to inorganic phosphate), we optimized the synthesis of GDP-[^{3}H]Fuc (Fig. 4.39), achieving an almost 100% synthesis after 30 min incubation. The detailed protocol is defined in section 3.7.10, analysed in section 6.2, and depicted in (Fig. 4.40).

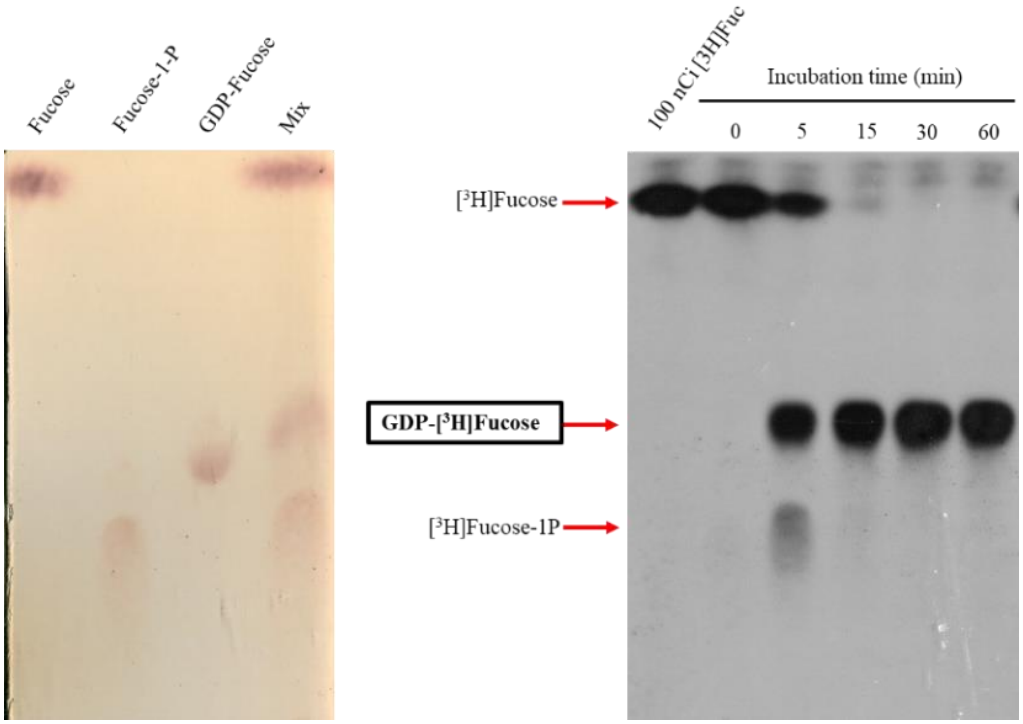


Figure 4.39: Optimized small-scale GDP-[^{3}H]Fuc synthesis. [^{3}H]Fuc, [^{3}H]Fuc-1P and GDP-[^{3}H]Fuc sugars seen by fluorography (right panel) can be determined by the retention factor of the non-radioactive species run in parallel and stained with orcinol/H_2SO_4 (left panel). Two consecutive developments in 1-propanol:acetone:water (9:6:4, v:v:v) (Fig. 4.38) were used.

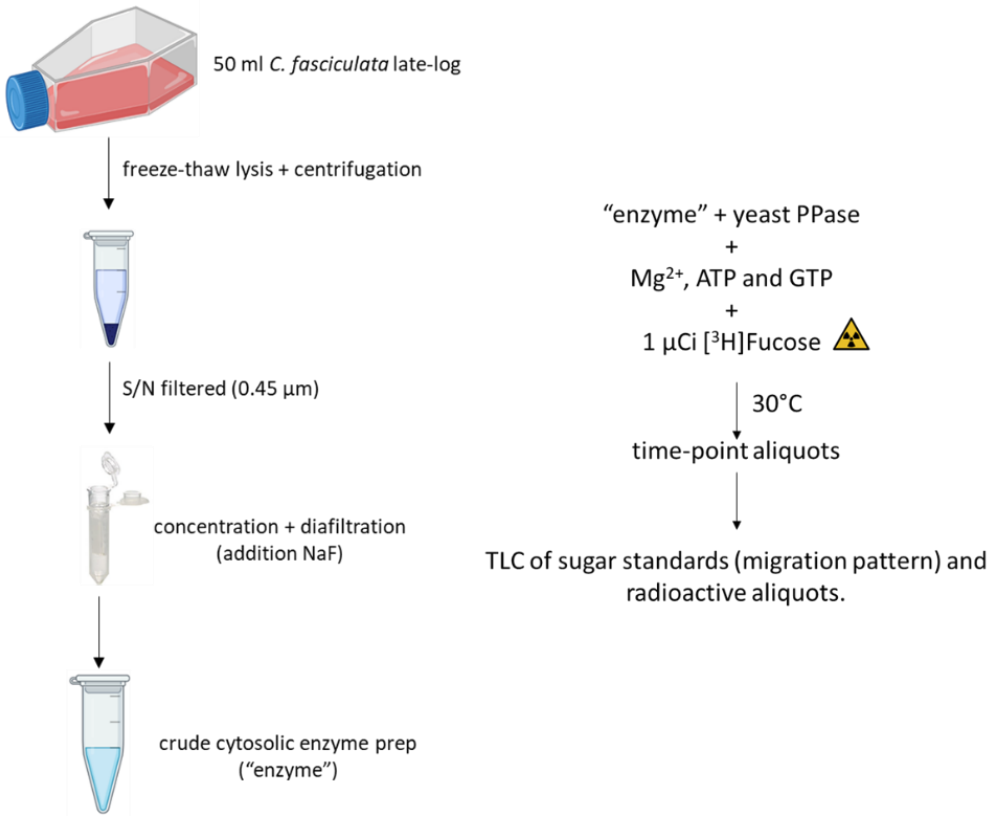


Figure 4.40: Scheme of optimized method for synthesis of GDP-[^{3}H]Fuc using crude cytosolic enzyme preparation from *C. fasciculata*.

The stability of the *C. fasciculata* cytosolic fucokinase/GDP-Fuc pyrophosphorylase activity was tested after a week kept either at 4ºC or after being snap-frozen with 10% glycerol at -20ºC. As seen in (Fig. 4.41), the enzymatic preparation is stable for at least a week kept cold, but freeze-thawing affects its activity.

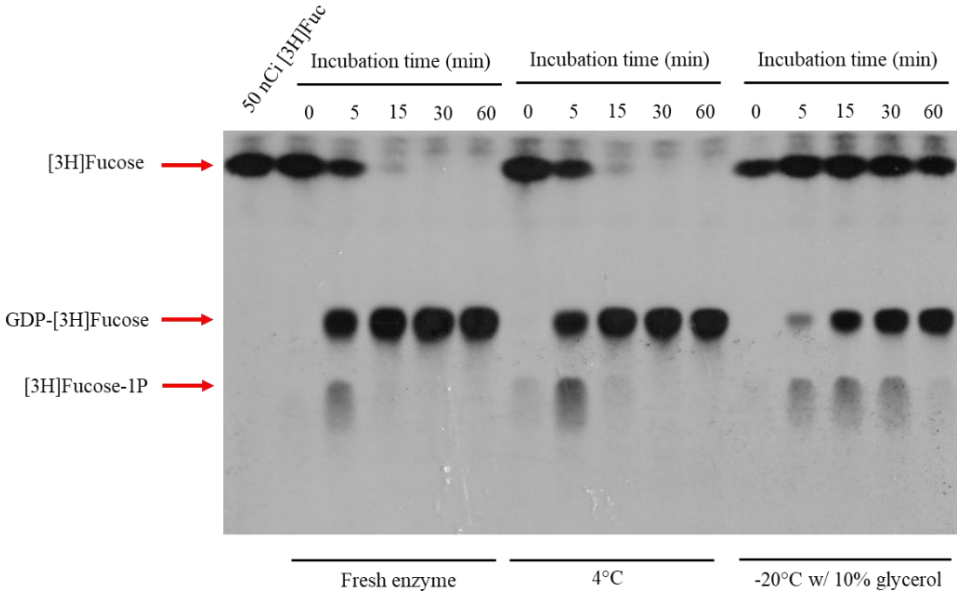


Figure 4.41: Analysis of the *C. fasciculata* fucokinase/GDP-Fuc pyrophosphorylase preparation stability for a week kept in different conditions (either at 4ºC or frozen at -20ºC with 10% glycerol). Each radioactive species detected by fluorography was defined with the orcinol/H_2SO_4-stained plate from (Fig. 4.39).

At this point, we decided to perform large-scale GDP-[^{3}H]Fuc synthesis using 50 µCi [^{3}H]Fuc and either an enzymatic preparation kept frozen for 3 weeks (Fig. 4.42A), or double the amount of freshly prepared enzyme (Fig. 4.42B) for a longer reaction time of 90 min, compar ed to the small-scale experiments (see section 3.7.10). The results were similar in both cases, indicating that the use of more enzyme did not increase the yield of the reaction. However, in both cases in these scale-up experiments, only incomplete conversion into GDP-[^{3}H]Fuc was achieved and, when frozen enzyme was used, a partial degradation of the GDP-[^{3}H]Fuc was detected (indicated by the presence of another radioactive species that we think is [^{3}H]Fuc-PP (Fig. 4.42A). This result also indicated that the rate limiting step of the reaction is the first one, the Fuc-kinase activity, since otherwise we would expect [^{3}H]Fuc-1P to have accumulated. Although not tested, an increase in the ATP concentration might overcome this problem, increasing the yield of the reaction.

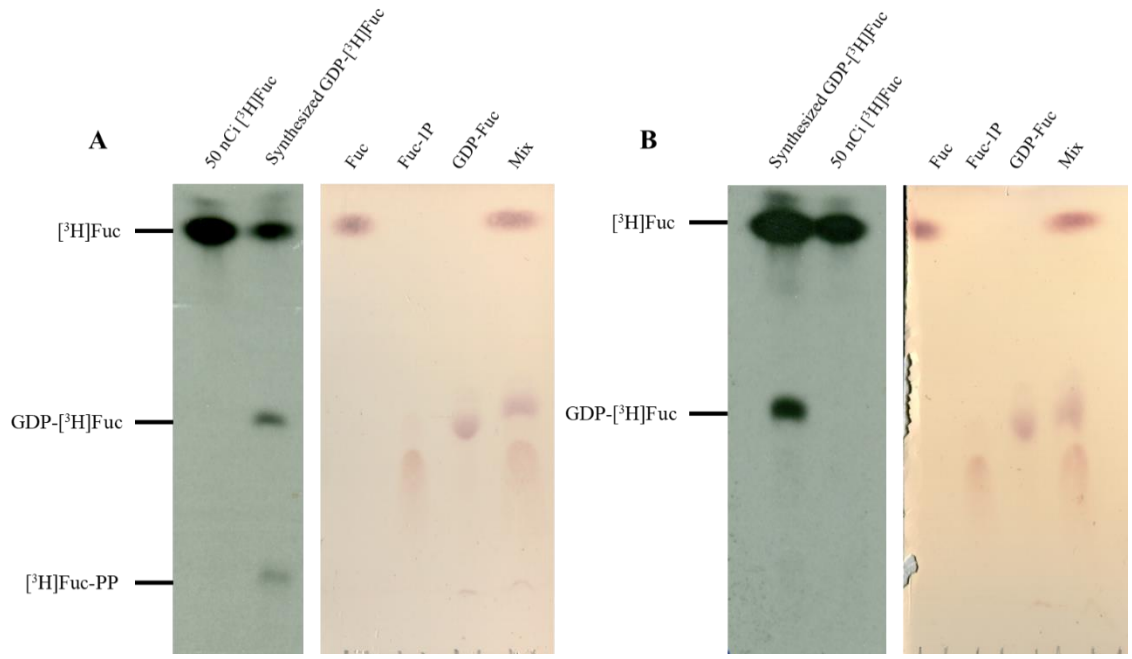


Figure 4.42: HPTLC analysis of large scale GDP-[^{3}H]Fuc synthesis attempts. Either (A) 3-week-old cytosolic enzymatic preparation, or (B) freshly prepared one, were used. Radiolabelled products were visualized by fluorography (left panels), and standard sugars with orcinol/H_2SO_4 (right panels). Two runs for each plate were performed, using 1-propanol:acetone:water (9:6:4) solvent.

Next, we started to investigate how to purify the GDP-[^{3}H]Fuc from the reaction mixture. Initially, we decided to test if a reverse-phase solid phase extraction (RP-SPE) procedure was able to separate a mix of Fuc, Fuc-1P and GDP-Fuc in the GDP-[^{3}H]Fuc synthesis reaction buffer. We used graphitised carbon as the stationary phase, the column was equilibrated first with a solution of 80% acetonitrile (MeCN) and 0.1 % TFA, and then with water. After passing through the equilibrated column the sugar mix, four elutions were performed with (1) water, (2) 25% MeCN, (3) 50 mM triethylamine acetate (TEAA) and (4) 25% MeCN containing 50 mM TEAA. A 10% fraction of each eluate was analysed by HPTLC, along with sugar standards, which was then stained with orcinol (Fig. 4.43). Two plates were run, testing the same solvents analysed in (Fig. 4.39).

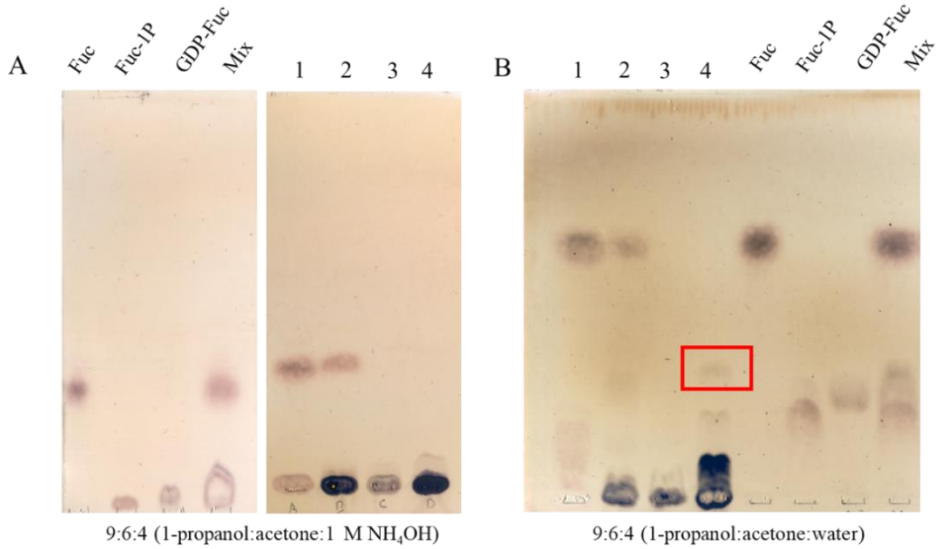


Figure 4.43: HPTLC analysis of RP-SPE test for purification of GDP-Fuc from a sugar mix. Mobile phases (A) 1-propanol:acetone:1 M NH_4OH (9:6:4) and (B) 1-propanol:acetone:water (9:6:4) were tested. Plates were stained with orcinol/H_2SO_4. One run was performed in both cases.

As shown in (Fig. 4.43), solvent containing NH_4OH (Fig. 4.43A) proved again to be inefficient when trying to separate species of similar polarity. When using the neutral solvent (Fig. 4.43B), the 4 fractions showed that although the 3 sugars had been partially separated, GDP-Fuc (Fig. 4.43B, red box) and Fuc-1P had co-eluted mainly in the fourth step. Along with them, and present in the two previous fractions, other unidentified species were also present in the eluates. This result made us discard RP-SPE as the method to purify the home-made radioactive sugar donor.

We next decided to try a weak anion exchange (WAX)-SPE method (section 3.7.11). The procedure was successful, and the unconverted [^{3}H]Fuc was recovered in the flow through and the first of 3 water washes (Fig. 4.44, FT and W1-3). Furthermore, in the first of four consecutive 1 M ammonium acetate elutions (Fig. 4.44 E1-4) 6 µCi of GDP-[^{3}H]Fuc were recovered (quantified by a liquid scintillation counting), representing a 6% yield from the initial 100 µCi [^{3}H]Fuc used in total in both reactions (Fig. 4.45A). A later synthesis experiment using fresh enzyme, 5 µCi [^{3}H]Fuc and a longer incubation time (2 h) yielded 22% final product, i.e., 1.1 µCi GDP-[^{3}H]Fuc (Fig. 4.44B). This was an indication that our system can be improved by increasing the incubation time of the reaction, and possibly, addition of more ATP as suggested before.

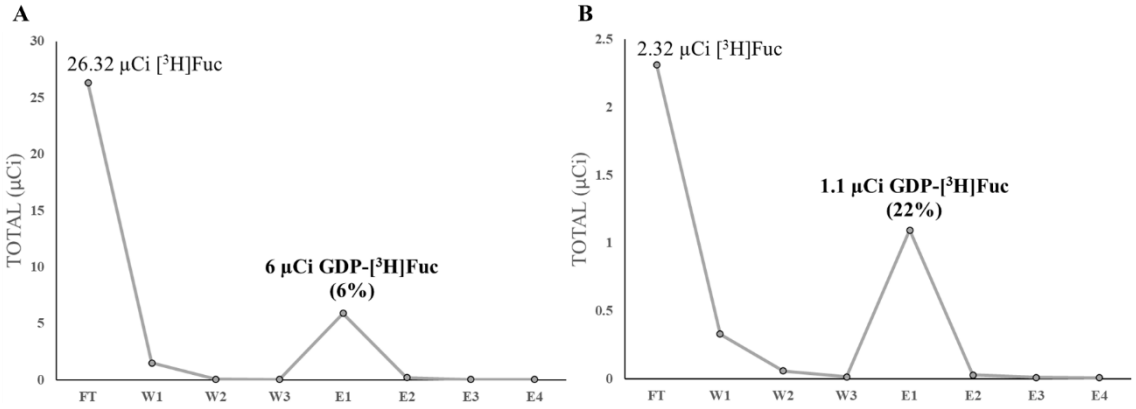


Figure 4.44: WAX-SPE profiles of home-made GDP-[^{3}H]Fuc purification. (A) Purification profile from two combined synthesis reactions, each using 50 µCi [^{3}H]Fuc as initial substrate (Fig. 4.42). (B) Purification profile from a synthesis reaction using 5 µCi [^{3}H]Fuc as initial substrate, and a longer (2 h) incubation time.

These purified 7.1 µCi GDP-[^{3}H]Fuc were lyophilized twice to remove the ammonium acetate, and resuspended in 50% ethanol (see Methods), ready to be used for glycosyltransferase activity assays. Additionally, the [^{3}H]Fuc from the FT fractions of both WAX-SPE purifications were recovered by desalting, vacuum-drying and resuspended in 70% ethanol (see section 3.7.11). Then, it was used for a new GDP-[^{3}H]Fuc synthesis with fresh enzymatic preparation and 3 h incubation. The WAX-SPE profile (Fig. 4.45A) and quantification by liquid scintillation counting of the fractions indicated that more than 10 µCi GDP-[^{3}H]Fuc were obtained, mainly in fraction E1, representing a 23.3% yield from the initial [^{3}H]Fuc used. Almost 1 µCi more was detected in fraction E2, increasing the total yield to 25.5%, indicating that even longer incubation times might be beneficial to the reaction yield. The GDP-[^{3}H]Fuc from fraction E1 was recovered to be used in further activity assays. The unconverted [^{3}H]Fuc from the FT fraction was also recovered for future use. A 50 nCi aliquot of both labelled products were used for HPTLC analysis (Fig. 4.45B), which showed a single fluorographic signals

correlating with the retention factors of the orcinol-stained Fuc and GDP-Fuc, respectively, indicating the uncontaminated, undegraded state of both radioactive samples.

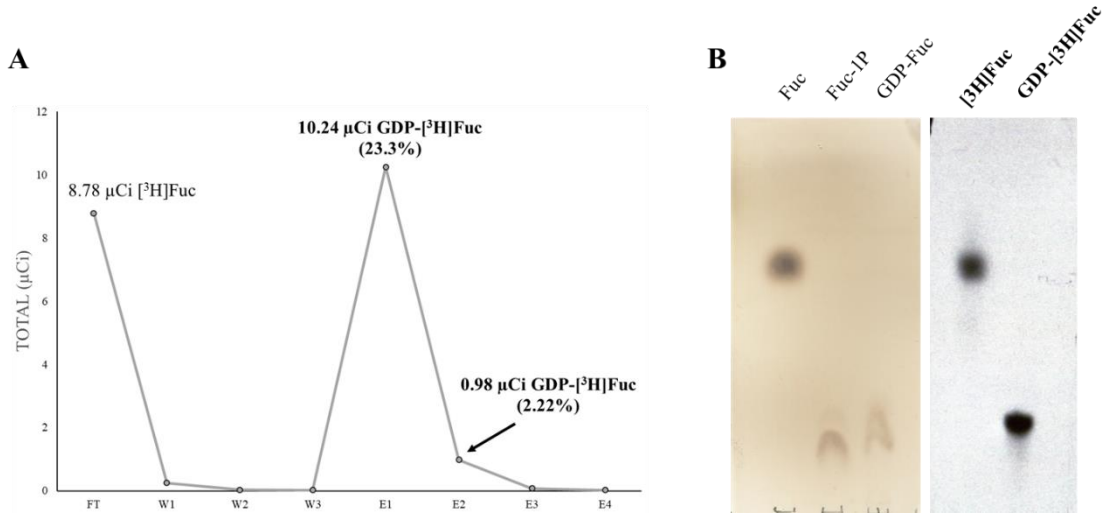


Figure 4.45: Analysis of GDP-[^{3}H]Fuc synthesis using recovered [^{3}H]Fuc from previous purifications (Fig. 4.44). (A) WAX-SE purification profile. (B) HPTLC analysis of fractions F1 and E1 after recovering the radioactive species. Radiolabelled products were visualized by fluorography (right panel), and standard sugars with orcinol/H_2SO_4 (left panel)

5. RESULTS II: Search of kinetoplastid mitochondrial acceptor substrates of FUT1.

5.1 [^{3}H]Fuc labelling of *C. fasciculata* cells.

With the aim of finding endogenous substrates of mitochondrial fucosylation in trypanosomatids, we initially designed a strategy to do so in *T. cruzi* (see section 6.3). However, analysis of the literature and the genomes of *T. cruzi*, as well for *T. brucei* and the *Leishmania* spp., revealed significant impediments to perform biosynthetic labelling using any of these organisms based on the uptake of [^{3}H]Fuc from the medium, and its conversion into GDP-[^{3}H]Fuc to be used as sugar donor to radiolabel acceptor substrate(s).

T. brucei not only lacks a cytoplasmic salvage pathway (to convert L-Fuc to Fuc-1-P and then to GDP-Fuc), it is also unable to take up L-Fuc from growth medium because it cannot transport pyranose sugars with axial C-3 and/or C-4 hydroxyl groups, such as D-Gal and L-Fuc (Eisenthal *et al.*, 1989). Consequently, *T. brucei* relies completely the *de novo* pathway from D-Glc to GDP-Fuc (Turnock *et al.*, 2007) (Fig. 5.1). This means that the only way to radiolabel GDP-Fuc and downstream acceptor substrates in *T. brucei* would be via radiolabelled D-Glc or D-Man, making the labelling procedure non-selective for Fuc-containing glycoconjugates.

While *Leishmania* spp. are able to take up L-Fuc and incorporate it into the major surface glycoconjugate, lipophosphoglycan (LPG) (Wilson *et al.*, 1999) via the AFKP80/FKP40 salvage pathway (Guo *et al.*, 2017), they also contain genes encoding an intact *de novo* pathway. This could significantly dilute the specific activity of GDP-[^{3}H]Fuc made by the salvage pathway with non-labelled GDP-Fuc from the *de novo* pathway, affecting the efficiency of labelling.

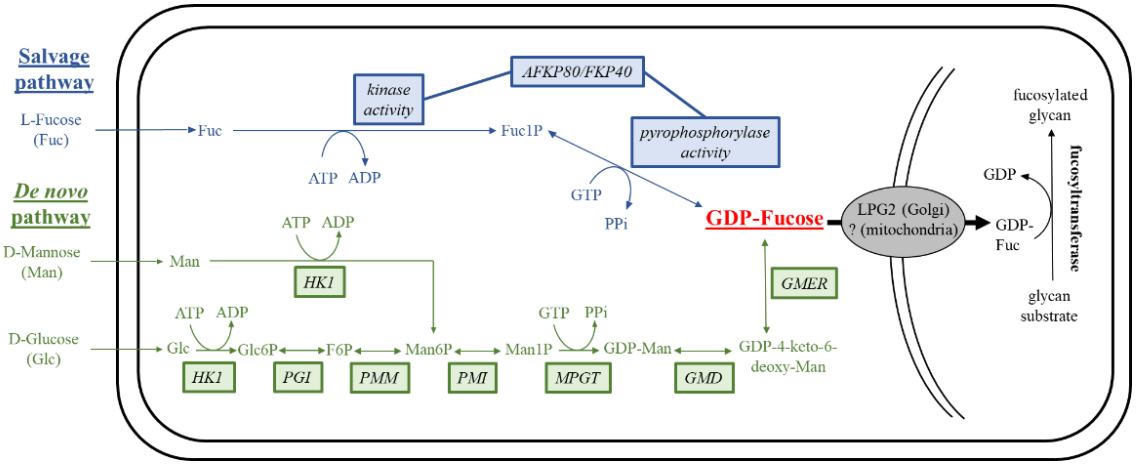


Figure 5.1 (*previous page*): Scheme of the nucleotide sugar GDP-Fuc biosynthesis in trypanosomatids. GDP-Fuc can be synthesized either by fucose salvage pathway in which the dual kinase/pyrophosphorylase activity of the cytoplasmic proteins AFPK80/FKP40 (characterized in *L. major*) is the only protein involved (Guo *et al.*, 2017), or from D-Glc or D-Man by a multi-step *de novo* pathway that, at least for *T. brucei*, has all the involved proteins located in the glycosomes (Sampaio Guther *et al.*, 2021). The presence of each pathway varies in the different trypanosomatids (sections 5.1 and 6.3). The mechanisms for internalization of the different sugars present outside the cells are not fully characterized for all trypanosomatids (see section 6.3). Once GDP-Fucose is synthesized, it is transported inside the Golgi apparatus by the LPG2 transporter (with orthologues in all trypanosomatids) or inside mitochondria by an unknown mechanism, in order to be used as the sugar donor of a FUT activity to synthesize fucosylated products. The enzyme abbreviations (in *italics*) are the following: hexokinase (*HK1*), phospho-glucose isomerase (*PGI*), phospho-mannose mutase (*PMM*), phospho-mannose isomerase (*PMI*), mannose phosphate guanyl transferase (*MPGT*), GDP-mannose dehydratase (*GMD*), GDP-4-hydro-6-deoxy-D-mannose epimerase/reductase (*GMER*), bifunctional arabino/fucokinase/pyrophosphorylase (*AFKP80*), bifunctional fucokinase/pyrophosphorylase (*FKP40*). The abbreviated products are the following: glucose-6-phosphate (Glc6P), fructose-6-phosphate (F6P), mannose-6-phosphate (Man6P), mannose-1-phosphate (Man1P), guanosine diphosphate mannose (GDP-Man), fucose-1-phosphate (Fuc1P). Scheme based on (Marquardt *et al.*, 1999; Turnock and Ferguson, 2007; Guo *et al.*, 2017; Sampaio Guther *et al.*, 2021).

By contrast, the monoxenous parasite *Crithidia fasciculata* appears to lack a competing *de novo* pathway from D-Glc to GDP-Fuc, since genes for last two enzymes of this pathway (*GMD* and *GMER*, Fig. 5.1), which convert GDP-Man into GDP-Fuc, are not present in its genome. *C. fasciculata* is easy to grow, is not infectious to mammals and has been used as a model organism in research into trypanosomatid biology that then may be applied to understand the human infective species (Alcolea *et al.*, 2014; Rojas *et al.*, 2014) (section 1.3). The main surface glycoconjugate in this parasite, LAG, it is anchored to the membrane by a inositol phosphoceramide moiety, and is composed of D-Ara, D-Gal and D-Man (Schneider *et al.*, 1996). To generate the donor substrate GDP-Arabinopyranose (GDP-Ara), the parasite converts D-Glc into D-Ara which is converted into GDP-Ara (via D-Ara1-P), which is used then used in the synthesis of LAG (Schneider *et al.*, 1995). The D-Ara to GDP-Ara activity is attributed to the *C. fasciculata* orthologue of *Leishmania* spp. AFKP80 (Fig. 5.1) which can process both L-Fuc and D-Ara due to their structural similarity (Fig. 6.6A) (Guo *et al.*, 2017). Furthermore, *C. fasciculata* can biosynthetically incorporate tritium labelled D-Ara ([^{3}H]Ara) into GDP-[^{3}H]Ara and LAG (Schneider *et al.*, 1995), and cytoplasmic extracts of *C. fasciculata* have been shown to convert [^{3}H]Fuc into GDP-[^{3}H]Fuc (Mengeling and Turco, 1999, and this thesis). Importantly, *C. fasciculata* also contains a FUT1 gene sequence in its genome. Based on this knowledge, and by determining whether LAG can be efficiently labelled and radiolabelled by uptake of L-Fuc and [^{3}H]L-Fuc, respectively, present in the medium, we decided to assess *C. fasciculata* as a suitable system to find other radiolabelled intracellular substrates.

Initially, we extracted and purified LAG and analysed it by GC-MS and confirmed the presence of the expected sugar components. LAG was purified from WT *C. fasciculata* cultures (section 3.7.1) by a first extraction with chloroform:methanol:water (1:2:0.8) followed by a second extraction with water saturated with 1-butanol (9% 1-butanol in water). This second extract was used to perform hydrophobic interaction chromatography (HIC) in which LAG was successfully bound and then eluted. Fractions (5% and 10%) of the eluate were used for periodic acid Schiff (PAS) staining of an SDS-PAGE gel in which they were loaded, along with 10 µg asialofetuin (corresponding to 1 µg carbohydrate mass), as positive control. Asialofetuin glycans and the sugar moiety of LAG were detected, the latter appearing as a polydisperse molecule (Fig. 5.2A). A 15% fraction of the purified LAG was used for GC-MS carbohydrate composition analysis, which indicated the presence of expected sugar elements of LAG (D-Ara, D-Man and D-Gal) (Fig. 5.2B) as previously reported (Schneider *et al.*, 1996). Multiple peaks can be detected for each sugar as mutarotation of the released sugars leads to the formation of α and β methyl glycosides of both pyranose and furanose forms of the sugars, the relative proportions of which are sugar-specific.

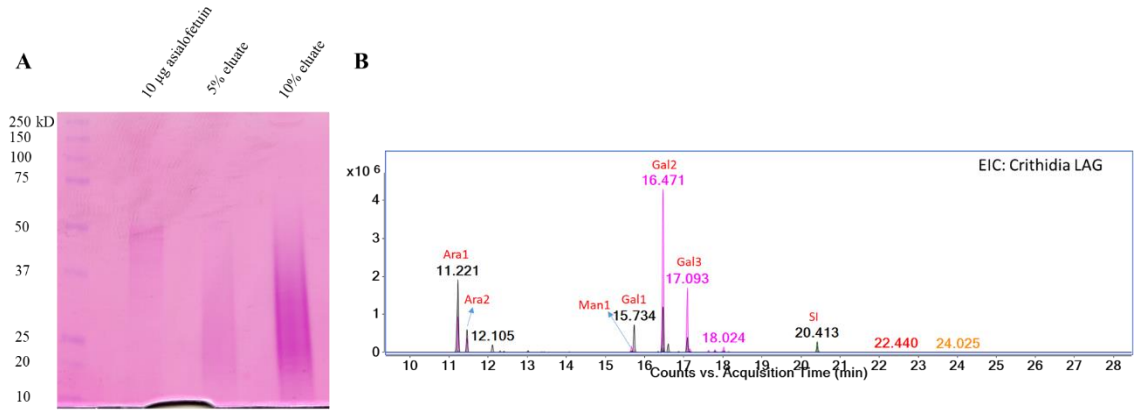


Figure 5.2: SDS-PAGE and PAS staining and composition analysis of LAG purified from *C. fasciculata*. (A) SDS-PAGE PAS staining of 5% and 10% fractions of purified LAG. 10 µg commercial asialofetuin, corresponding to 1 µg carbohydrate, were used as positive control. (B) Extracted ion chromatogram (m/z 204 and 217) of purified LAG on GC-MS. 1 nmole of *scyllo*-inositol (s-I) was used as an internal standard.

Next, we grew *C. fasciculata* in medium supplemented with different concentrations of L-Fuc (0, 2, 10 and 50 mM) prior to LAG purification. Aliquots (10%) of the purified LAG preparations were analysed by SDS-PAGE and PAS staining (Fig. 5.3A), and 15% aliquots were used for GC-MS analysis. In (Fig. 5.3B), the extracted ion chromatogram of LAG purified from a *C. fasciculata* culture supplemented with 10 mM L-Fuc is shown, confirming the efficient incorporation of L-Fuc into LAG, displacing all

of the D-Ara from the glycoconjugate composition. This confirmed our hypothesis of L-Fuc uptake and salvage into GDP-Fuc for glycosylation in *C. fasciculata*.

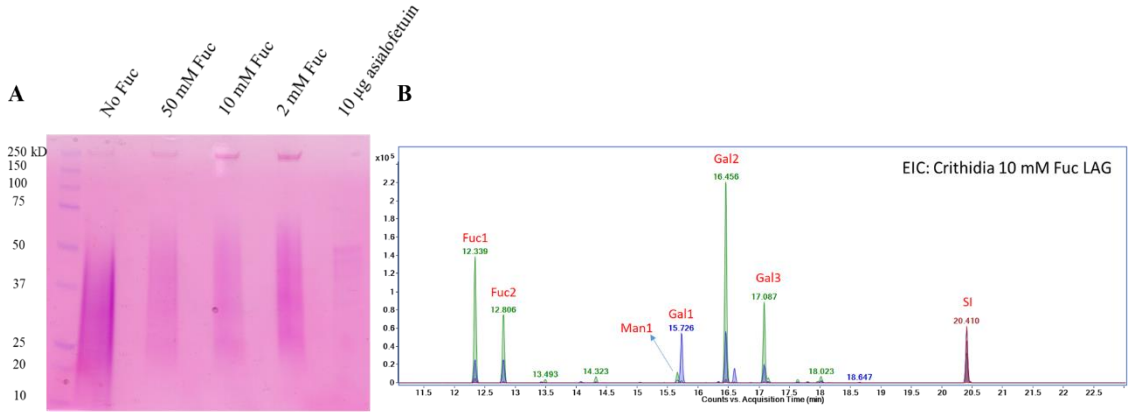


Figure 5.3: SDS-PAGE and composition analyses of LAG from *C. fasciculata* grown without (No Fuc) or supplemented with different L-Fuc concentrations. (A) SDS-PAGE PAS staining of 10% fractions of purified LAG from all cultures. 10 µg commercial asialofetuin, corresponding to 1 µg carbohydrate, was used as positive control. (B) GC-MS extracted ion chromatogram of purified LAG from *C. fasciculata* grown in 10 mM Fuc-supplemented medium. 1 nmole of *scyllo*-inositol (s-I) was used as an internal standard.

Even though this parasite does not harbour a *de novo* pathway for synthesis of GDP-Fuc that would affect the efficiency of the [^{3}H]Fuc labelling, the uptake of [^{3}H]Fuc might have to compete with D-Glc and/or D-GlcN in the standard growth medium for uptake. We therefore decided to compare biosynthetic labelling of LAG with [^{3}H]Fuc in normal growth medium and D-Glc/D-GlcN-free medium. First, we tested by light microscopy if parasite survival was affected when cultured in medium depleted of D-Glc and D-GlcN and determined that the cells survived and appeared healthy for at least 24 h. Subsequently, a first labelling experiment was performed in which [^{3}H]Fuc was added to either standard or D-Glc and D-GlcN-depleted medium (- sugar medium). Cell lysates from time points of 1, 2, 4 and 6 h incubation in standard or - sugar radiolabelling (containing [^{3}H]Fuc) medium were subjected to SDS-PAGE and Coomassie blue staining, followed by incubation in En^{3}Hance fluorophore and exposure to X-ray film. In the developed X-ray films (Fig. 5.4B and 5.4C), we could detect how the vast majority of the label went to a polydisperse molecule, suggesting that [^{3}H]Fuc was incorporated into LAG. Also, this process was much more favourable in cells incubated in sugar-depleted medium, suggesting that [^{3}H]Fuc (at 0.16 µM in labelling medium) competes for uptake with glucose and/or glucosamine through the same transporter(s). D-Glc and D-GlcN are present at 5.6 mM (35,000 times more concentrated) and 0.28 mM (1,750 times more concentrated), respectively, than [^{3}H]Fuc in these experiments in standard medium. This difference in labelling efficiency is more evident in another X-ray film exposed to the gel for shorter time (Fig. 5.4C). Importantly, a faint, high-molecular weight signal could be

detected in cells labelled for 6 h in – sugar medium (Fig. 5.4B, red box). In order to confirm the presence of this other labelled molecule, further experiments were performed.

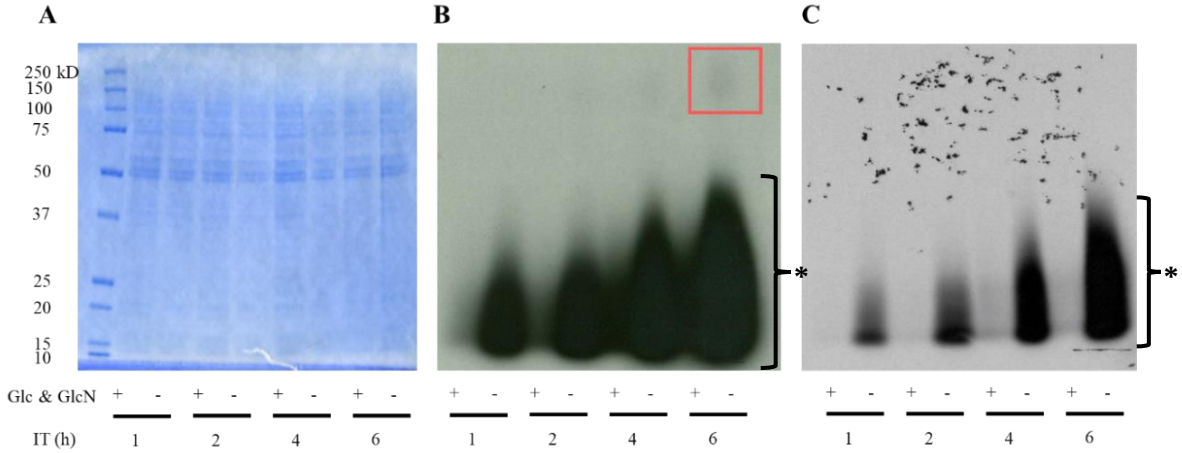


Figure 5.4: [^{3}H]Fuc biosynthetic labelling in *C. fasciculata*. Aliquots of cell cultures at time points (1, 2, 4 and 6 h) of [^{3}H]Fuc labelling in medium with or without glucose and glucosamine (Glc & GlcN) were taken for analysis. (A) Vacuum-dried SDS-PAGE gel processed for fluorography containing all aliquots. (B) 1 week exposure developed X-ray film. Red box indicates a high-molecular weight species that was detected along with the [^{3}H]-labelled polydisperse LAG. (C) 2-day exposure of X-ray film evidencing a higher efficiency of labelling in - Glc & GlcN medium. The **"*"** indicates [^{3}H]Fuc-containing LAG position.

LAG is a glycosylphosphatidylinositol (GPI)-anchored glycoconjugate. Therefore, it can be separated into its sugar and lipid moieties by treatment with phosphatidylinositol-specific phospholipase C (PI-PLC), which in this case cleaves the *sn*-3-phosphodiester bond between the *myo*-inositol and ceramide (Fig. 5.5A). Thus, by treating cells that were incubated with [^{3}H]Fuc for 5 h in - sugar medium with this enzyme, we aimed to remove the radioactive signal from [^{3}H]-LAG from the fluorograph in order to reveal any other signals that could represent potential intracellular fucosylated products. The result showed that when there is no treatment (Fig. 5.5B, lane -), and a bigger number of labelled cells were lysed and loaded in the gel, the high-molecular weight species was clearly present. When PI-PLC treatment was performed (Fig. 5.5B, lane +), the [^{3}H]-LAG signal disappears, but the higher molecular weight species remains and actually appears to increase in intensity. This increase in intensity of higher molecular weight material might be due to some of the PI-PLC cleaved LAG (arabinogalactan) entering the gel but migrating much more slowly due to lack of SDS-binding lipid. This, in turn, might suggest that high molecular weight labelled material observed before PI-PLC digestion might itself be due to non-lipid anchored arabinogalactan. Such a situation is known in *L. donovani*, where a lipid-free form of LPG, called extracellular phosphoglycan (exPG), has been characterised (Greis *et al.*, 1992).

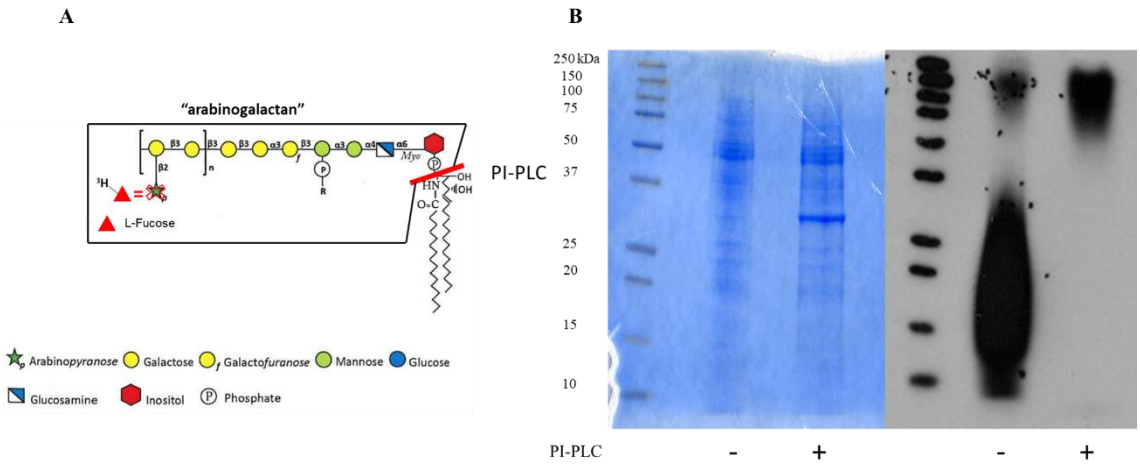


Figure 5.5: [^{3}H]-LAG scheme and PI-PLC treatment of radiolabelled lysates. (A) Scheme of LAG structure, based on (Schneider *et al.*, 1996; Iljazi, 2018). PI-PLC digestion separates the glycan moiety ("arabinogalactan") from the ceramide. (B) SDS-PAGE gel Coomassie blue staining (left) and fluorographic detection (right) of [^{3}H]-labelled species from lysates before (-) and after (+) treatment with PI-PLC.

In order to test this hypothesis, we performed a series of different treatments along with PI-PLC on the radioactive lysates after labelling more cells with [^{3}H]Fuc, and analysed in a similar manner as before to see whether both species, LAG and the high-molecular weight material, behaved similarly or not. The additional treatments were dephosphorylation with 48% aq. HF at 0ºC, mild acid hydrolysis with 40 mM TFA at 90ºC for 18 min to cleave sugar-1-phosphate bonds, and α1-2 exofucosidase digestion to remove non-reducing terminal α-L-fucose (αFuc) residues (Fig. 5.6). The PI-PLC digest result was similar to that described in (Fig.5.5), as expected.

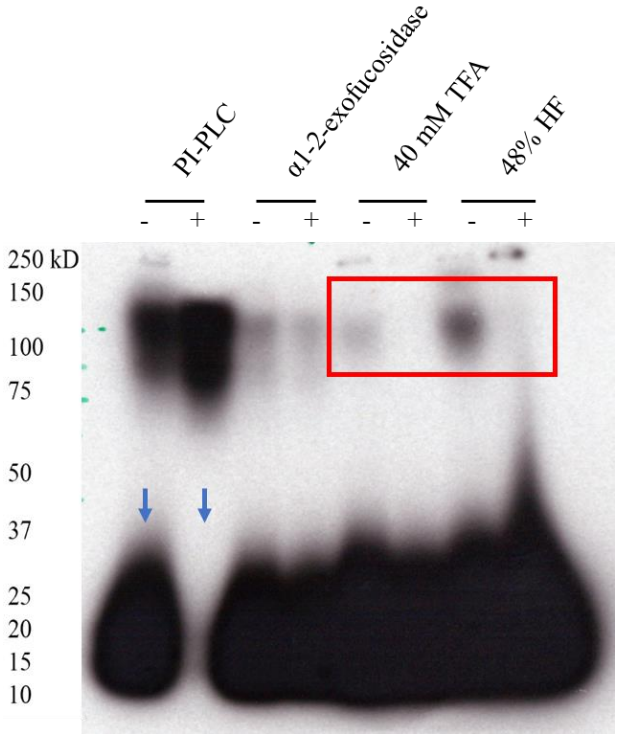


Figure 5.6: Different enzymatic and chemical treatments of radiolabelled *C. fasciculata* lysates. After treatments, radioactive signal from samples was detected by fluorography of SDS-PAGE gel. PI-PLC reproducible result (see Fig. 5.5) is marked with blue arrows. High molecular weight labelled species sensitivity to TFA and HF treatments is emphasized with a red box.

The α1-2 exofucosidase digestion was inconclusive since both LAG and the high-molecular weight material remained after digestion (albeit there may have been some reduction in signal for the high-molecular weight material). Native LAG contains Araα1-2 residues decorating a poly (-3Galβ1-3Galβ1-) β-galactan chain (Schneider et al., 1996) (Fig. 5.5A and Fig. 6.7) and we presume that the [^{3}H]Fuc residues incorporated into it are similarly linked. Thus, LAG might be a particularly difficult substrate for α1-2 exofucosidase based on the steric constraints to access the [^{3}H]Fucα1-2Gal glycosidic bond due to the adjacent βGal residue in the [^{3}H]Fucα1-2(Galβ1-3)Gal environment (Fig. 5.5A and Fig. 6.7). Since it is not known what structures are present in the high-molecular weight material it is hard to predict the accessibility of incorporated [^{3}H]Fuc in this material. In any case, a positive control digestion with a known α-fucoside was not performed in this experiment so we must be careful not to overinterpret this result.

The mild acid hydrolysis experiment suggests that LAG is relatively unaffected, whereas the high-molecular weight material lost all detectable [^{3}H]Fuc radiolabel. Although αFuc and αArabinopyranose glycosidase linkages, and also the β-D-galactofuranose (βGal*f*) linkage in the LAG core, are relatively acid labile, they require ≥1 h of hydrolysis in 40 mM TFA at 100^{o}C to approach quantitative cleavage (McConville *et al.*, 1990; Schneider *et al.*, 1996). The conditions used here (18 min 40 mM TFA at 90^{o}C) are selective for much more labile linkages, such as the Manα1-PO$_3$H-glycosidic phosphodiester bond of molecules like the *Leishmania* LPGs. The loss of label from the high-molecular weight material observed here may indicate that the incorporated [^{3}H]Fuc is attached to macromolecules, either directly or indirectly, via a sugar-1-phosphodiester linkage. In any case, this result provides the first evidence that the high-molecular weight material behaves differently from LAG and is, therefore, likely to be a different class of molecule.

The aq. HF dephosphorylation experiment also showed the same differential sensitivity between LAG and the high-molecular weight material as the mild acid treatment (Fig. 5.6), again suggesting that they belong to different classes of macromolecule. In this case, we were surprised that aq. HF-digested LAG did not behave in a similar way to PI-PLC digested LAG; aq. HF should cleave the *myo*-inositol-1-PO$_3$H-ceramide phosphodiester linkage on both sides of the phosphate group and the rest of the molecule (with the exception of the single phosphoryl substituent on the LAG core (Fig. 5.5A and 6.7)) should be resistant to aq. HF. The reason for this discrepancy is unclear at this time.

In summary, from these experiments we may conclude:

(i) *C. fasciculata* can be efficiently biosynthetically radiolabelled with [^{3}H]Fuc in medium depleted of glucose and glucosamine.

(ii) [^{3}H]Fuc radiolabel is incorporated principally into *C. fasciculata* LAG (as expected) and, to a much lesser extent, into high (apparent) molecular weight material, as resolved by SDS-PAGE.

(iii) The [^{3}H]Fuc radiolabel in LAG and the high (apparent) molecular weight materials behave differently when digested by PI-PLC, mild acid and aq. HF.

(iv) Some PI-PLC digested LAG may enter SDS-PAGE gels, obscuring the high (apparent) molecular weight material.

From this, we decided to explore ways to genetically ablate LAG synthesis in *C. fasciculata* in order to simplify the task of selectively labelling the high (apparent) molecular weight material, which would greatly aid its purification by following incorporated [^{3}H]Fuc radiolabel throughout extraction, chromatographic and electrophoretic separations.

5.2 Gene knock-out attempts in *C. fasciculata.*

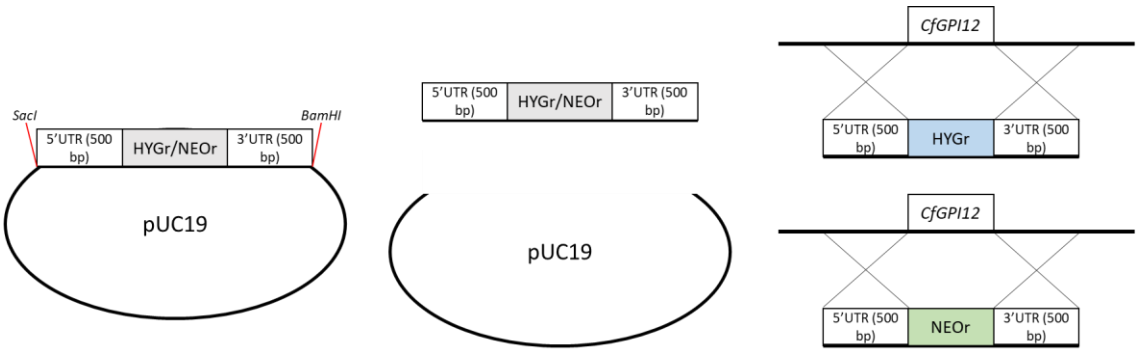
GPI-anchored proteins and GPI-related glycoconjugates (like *C. fasciculata* LAG) are synthesized in the secretory pathway since they are destined for the cell surface of eukaryotes. The preceding data indicated that the [^{3}H]Fuc-labelled high-MW material behaved differently to LAG upon digestion with PI-PLC, mild acid and aq. HF, suggesting that it might represent non-GPI glycoconjugate material, and possibly the product of kinetoplastid mitochondrial FUT1 action. Based on this, we decided to generate a double knock-out (dKO) cell line in *C. fasciculata* unable to synthesize GPI anchors and GPI-related glycoconjugates, and therefore unable to synthesize LAG. This would make the [^{3}H]Fuc-labelled high molecular weight material much easier to track through fractionation attempts by following [^{3}H]Fuc-radiolabel.

As the targeted gene for deletion, we chose *CfGPI12*. The GPI12 gene product is the GlcNAc-PI de-N-acetylase that catalyses the second step of the GPI biosynthetic pathway (Doering *et al.*, 1989; Watanabe *et al.*, 1999; Chang *et al.*, 2002). The *TbGPI12*

gene is not essential for the insect-dwelling, procyclic form of *T. brucei* (Güther *et al.*, 2006) or for *L. major* promastigotes (Almani *et al.*, 2016). Since *C. fasciculata* is a monoxenous culicid parasite (section 1.3) and evolutionarily very close to the *Leishmania* genus (both form the subfamily Leishmaniinae) (Jirků *et al.*, 2012), we hypothesized that, most probably, *CfGPI12* would also not be essential for *C. fasciculata*.

To delete *CfGPI12*, we decided to follow a gene replacement strategy by homologous recombination with a DNA template containing a drug resistance gene (Fig. 5.7). For this, we generated our gene replacement constructs by Gibson assembly (Gibson *et al.*, 2009). This cloning method allows for the successful assembly of multiple DNA fragments, regardless of fragment length or end compatibility. Each fragment is PCR-amplified with primers that grant them overlapping ends, each primer being designed to allow for the desired concatenation of the different fragments, i.e., for the desired order of fragments in the final, assembled product. This method has been proven to be suitable for the generation of large DNA constructs, including the assembly of the whole mouse mitochondrial genome from 600 overlapping 60-mers (Gibson *et al.*, 2010). In a single isothermal reaction, the different DNA fragments with overlapping ends are joined thanks to three different enzymatic activities (Gibson, 2011). Thus, the 500-bp 5' and 500-bp 3' UTRs directly adjacent to the 5' and 3' ends of *CfGPI12* ORF, the hygromycin resistance gene ORF (*HYGr*) from the pTRIX2-Luc::Neon-HYG plasmid (Costa *et al.*, 2018), the *NEOr* ORF from pGEM-T Easy VEX2 NEOr (kind gift from Joana Faria, former postdoc in David Horn group), and the whole sequence of commercial plasmid pUC19, were PCR-amplified using primers with hanging overlapping ends to then be ligated by Gibson assembly (see section 3.6.11). In this way, two plasmids containing each drug resistance gene were generated and sequenced to confirm successful cloning. The overlapping ends for the assembly of the pUC19 backbone with the repair template, i.e., with the 5' end of the 5'UTR, and the 3' end of the 3' UTR sequences, contained restriction sites (SacI and BamHI) that, after digestion, allowed for the separation of the DNA region needed for gene replacement from the rest of the plasmid. This digested DNA was used for electroporations of *C. fasciculata* (see section 3.5.5) without further purification of the template.

Figure 5.7 (*next page*): *CfGPI12* gene replacement strategy. (Left scheme) Homologous recombination template was generated by Gibson assembly containing restriction enzyme sites (in *italics*) between the pUC19 backbone and the DNA region of interest. (Middle scheme) Digestion products were used for electroporations, since the pUC19 backbone should not affect the recombination process. (Right scheme) For the replacement of *CfGPI12*, a sequential replacement of the two original alleles with drug resistance genes was attempted.



We initiated the process by electroporating the cells with the hygromycin resistance template. When eventually drug-resistant populations were recovered, we decided to obtain clonal populations initially by limiting dilution in standard medium. Unfortunately, this method did not yield any clonal population, which made us try as an alternative a cloning protocol using solid medium (section 3.5.5). In this way, some colonies were obtained, which were expanded in liquid culture and their genomic DNA (gDNA) extracted for PCR analysis using primers for the *CfGPI12* 5'UTR region and part of the *HYGr* ORF (JC34 + JC35, Fig. 5.9). The gDNA of 4 PCR-positive clones, and of WT cells, were used for genotype analysis by Southern blot using DIG-labelled probes for the *CfGPI12* and *HYGr* full ORFs (Fig. 5.8).

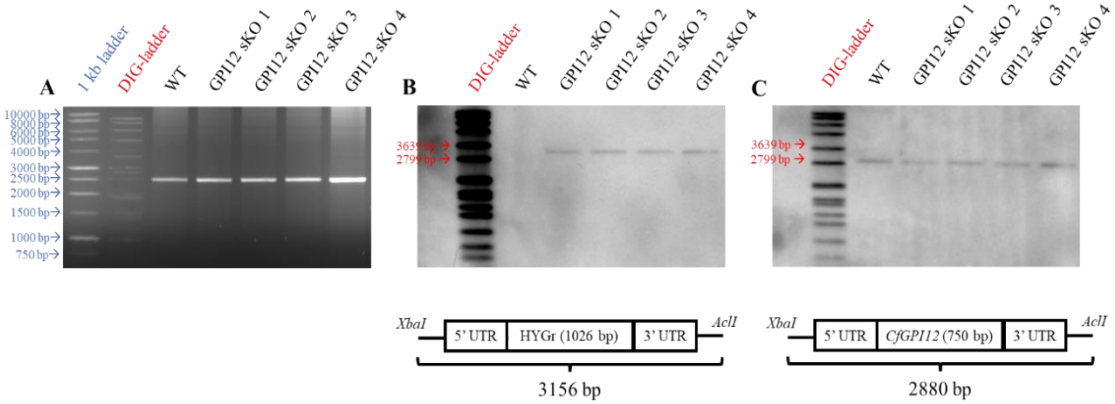


Figure 5.8: Genotype analysis by Southern blot of the Δ*CfGPI12*::*HYGr* sKO mutants. (A) Ethidium bromide (EtBr)-stained 0.8% agarose gel containing XbaI and AclI-digested gDNAs of WT and sKO cells. After transferring DNA to positively charged nylon membrane, it was probed against *HYGr* ORF (B), stripped and re-probed against *CfGPI12* ORF (C). DIG-labelled ladder was used to determine the size of the detected band in (B) and (C). Schemes of the detected loci are shown under the developed chemiluminescence films in (B) and (C).

Using a *HYGr* ORF probe, this experiment showed the detection of a unique band (absent in the WT sample) of the correct size to suggest that the *HYGr* gene had replaced the *CfGPI12* gene in one locus in all four of the putative sKO clones (Fig. 5.8B). Further, when the same blot was stripped and probed with a *CfGPI12* probe, a single band of the predicted size for the *CfGPI12* locus was seen in the wild type and all four putative sKO clones (Fig. 8C). The latter result suggested that no genetic rearrangements involving the *CfGPI12* gene occurred upon single allele replacement. Interestingly, all of the XbaI and

AclI digested gDNA samples showed an EtBr-stained band of around 2,500 bp (Fig. 5.8A) of unknown origin.

Following this initial promising result, we carried on with the replacement of the second *CfGPI12* allele with the *NEOr* gene. After electroporations, the recovered double drug-resistant mixed populations were spread in solid media plates and the colonies obtained were expanded for gDNA extraction. Initially, we performed a PCR analysis in a similar manner as for the Δ*CfGPI12*::*HYGr* sKO clones, i.e., using primers targeting the *CfGPI12* 5'UTR region and part of the *NEOr* ORF (JC34 + JC40, Fig. 5.9). Although this experiment was positive for the 12 potential dKO colonies analysed, amplification of the original *CfGPI12* gene (JC36 + JC37, Fig. 5.9) was also positive for all of them. Due to this incongruent result, further PCR analyses were performed with these and other gDNAs extracted from double drug-resistant colonies. A map of the primers used in the multiple PCR analyses of the *CfGP12* locus is shown in (Fig. 5.9)

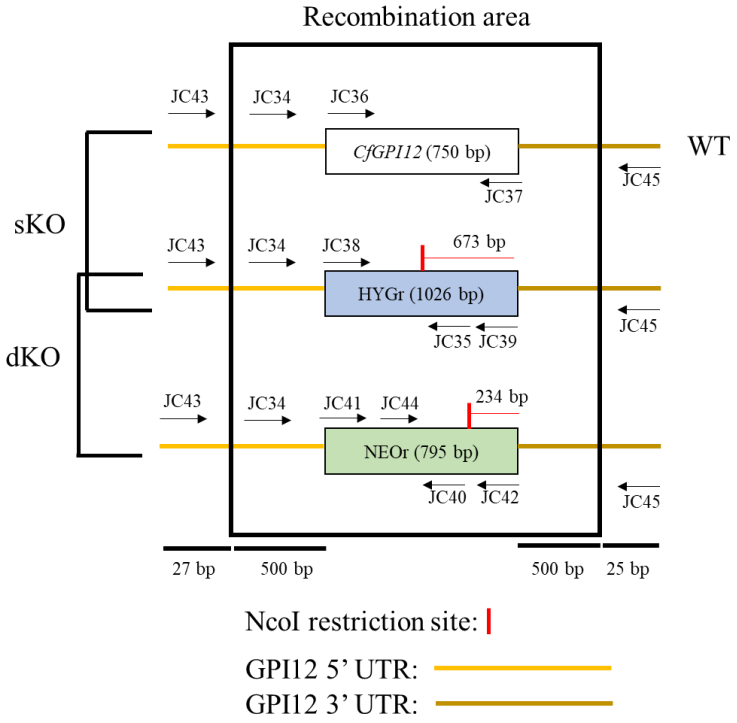


Figure 5.9: Map of primers used in PCR analyses for detection of Δ*CfGPI12*::*HYGr*/Δ*CfGPI12*::*NEOr* dKO colonies.

Numerous attempts were performed trying different amplifications and, in some cases, including the further digestion of the PCR products with NcoI restriction enzyme to confirm these contained the *CfGPI12* ORF sequence, since the drug resistance genes contained this restriction site (Fig. 5.9). Nevertheless, all obtained clones supported the PCR amplification of the original *CfGPI12* ORF, along with the presence of both drug resistance genes, leading us to hypothesize either the wrong insertion of the *NEOr* gene

in the genome of the parasites after electroporations or the duplication and transposition of the *CfGPI12* gene when the second allele was correctly replaced by *NEOr*. Overall, the results were inconclusive. We carried on with further electroporations and PCR analyses of colony gDNAs, but no Δ*CfGPI12*::*HYGr*/Δ*CfGPI12*::*NEOr* dKO clones were obtained. In parallel, we also tried to obtain Δ*CfGPI12*::*NEOr* sKO cells by electroporating WT cells with the respective template, and to induce gene conversion of the second allele of the Δ*CfGPI12*::*HYGr* sKO by growing some of the clonal populations in increasing concentrations of hygromycin. In all cases, no dKO population were obtained. The possible reasons for not being able to generate a Δ*CfGPI12*::*HYGr*/Δ*CfGPI12*::*NEOr* dKO cell line, including the possible essentiality of the gene in *C. fasciculata*, are commented on in section 6.4.

Due to this situation, we decided to redirect our efforts to ablate LAG synthesis by knocking out expression of *CfLPG1* or *CfLPG2*. Originally identified in *L. major* and *L. donovani* (Ryan *et al.*, 1993; Descoteaux *et al.*, 1995; Sacks *et al.*, 2000), both genes are essential for the synthesis of *Leishmania* spp. surface membrane glycoconjugate lipophosphoglycan (LPG). LPG1 is a β-galactofuranosyl transferase, involved in the glycan core synthesis of LPG. LPG2 is a Golgi GDP-Man/GDP-Ara/GDP-Fuc transporter necessary for the availability of these sugar nucleotides during the synthesis of LPG in the secretory pathway. Both proteins have orthologues in *C. fasciculata*, and therefore, we hypothesized that by generating a *CfLPG1* KO cell line, the parasite would synthesize a partial version of LAG, without a galactofuranose (Fig. 5.5A and Fig. 6.7) present in its glycan core, and therefore unable to further synthesize the repetitive chain of β1-3-linked D-galactan (Fig. 5.5A and Fig. 6.7) to which [^{3}H]Fuc would be transferred. In the case of *CfLPG2*, a null mutant of this gene would still be able to synthesize LAG, but the salvaged GDP-[^{3}H]Fuc would not be transported inside the Golgi apparatus, preventing the transfer of [^{3}H]Fuc to the C2 positions of the β1,3-linked D-galactan Gal residues (Fig. 5.5A and Fig. 6.7). In either case, no [^{3}H]-LAG would be synthesized, fulfilling our aim.

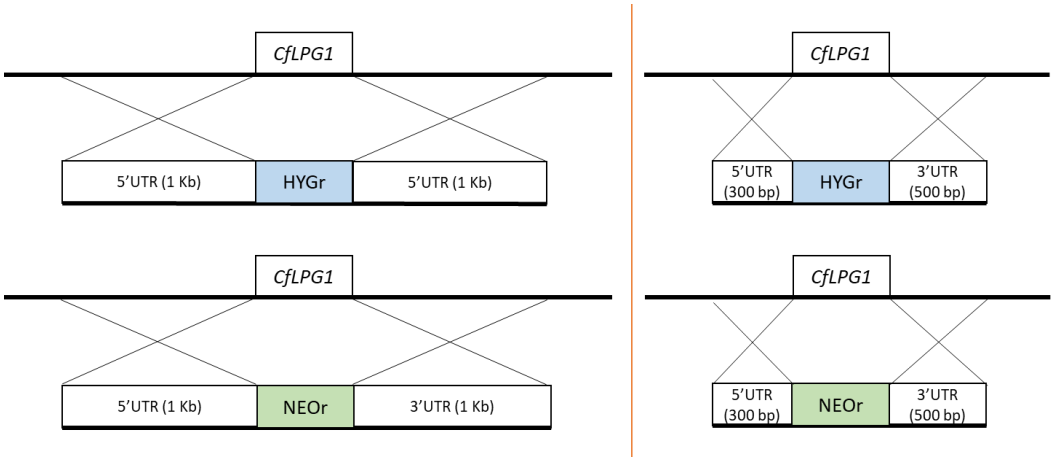
For the deletion of *CfLPG1* and *CfLPG2*, we initially followed the same strategy as for *CfGPI12* KO. Constructs on a pUC19 backbone in which a *HYGr* or *NEOr* genes were flanked by 500-bp 5' and 3' UTRs of the original genes were generated by Gibson assembly. After digestion of both templates from the rest of the plasmid, multiple electroporations were performed targeting both genes. Although some drug-resistant mixed populations were obtained and these were cloned by spreading in solid media and regrown for gDNA extraction, the PCR analysis of the *CfLPG1* and *CfLPG2* loci in all

cases indicated that the original gene was the only amplicon detected. Furthermore, PCR amplification of *CfLPG1/CfLPG2* 5'UTR or 3'UTR region and *HYGr/NEOr* ORF was in all cases negative, but amplification of the drug-resistance ORFs was always positive. Altogether, these results indicated that the templates conferring drug resistance to the cells had been wrongly inserted in the genome.

Protocol variations such as decreasing the amount of DNA template electroporated or the addition of 10% FBS in the recovering media post-electroporation to avoid a potential "sticky" phenotype of sKO cells with less LAG being synthesized (see section 6.4), were also tested and further mixed populations and colonies were obtained and analysed by PCR, but the results were the same as described above.

Due to this situation, we hypothesized that a possible explanation for the wrong insertion of our drug resistance genes might be due to the selectivity for the targeted locus by the homologous recombination regions. We decided to focus on the LPG1 gene, since its deletion would be more specific than for LPG2 in terms of ablation of only LAG synthesis, and therefore less likely to be essential. By performing a nucleotide BLAST analysis of the 500-bp 5'UTR and 3'UTR regions of *CfLPG1* against the genome of the parasite, we detected less cross-homology with other regions of the annotated *C. fasciculata* genome when the 5'UTR region was reduced to 300-bp instead of 500-bp, so we tried that as a targeting strategy. At the same time, we also tried to increase targeting efficiency by increasing the homologous recombination sequences to 1 kb, a strategy that is more in line with numerous successful attempts of homologous gene replacement in the related parasite *Leishmania* spp., including for *AFKP80* and *FKP40* (Guo *et al.*, 2017) and *LPG1* (Zhang *et al.*, 2004). To generate the shorter 5'UTR, we performed site-directed mutagenesis (section 3.6.12). Primers (JC90 + JC91) were used for the amplification of the whole HYGr/NEOr cassette without the inclusion of the first 200 bp of the 500-bp long *CfLPG1* 5'UTR, and the shorter constructs were sequenced prior to their use. For the generation of longer cassettes, a new Gibson assembly was designed, amplifying 1 kb long *CfLPG1* adjacent UTRs from the parasite's genome with primers including hanging overlapping ends for the ligation with the pUC19 backbone, which were later sequenced to confirm the successful assembly. The four generated constructs are schematized in (Fig. 5.10).

Figure 5.10 (*next page*): New batch of DNA templates for homologous replacement of *CfLPG1* gene. Constructs with either longer UTRs (left scheme), or shorter 5'UTR (right scheme) were generated by Gibson assembly or site-directed mutagenesis, respectively.



After many electroporations with the NEOr replacement constructs with shorter or longer *CfLPG1* UTRs, various neomycin-resistant mixed populations were recovered. In parallel with the spreading of these cultures on solid media plates to obtain colonies, this time we decided to apply a well-established procedure in this group for generating *T. brucei* clones, i.e., the *C. fasciculata* populations were used for limit-dilution in medium containing 10% FBS and 4 g/L $NaHCO_3$ and incubated in an atmosphere with 5% CO_2 (see section 3.5.5). In this way, multiple clones and colonies were obtained, and their gDNA was extracted for PCR analysis of the *CfLPG1* locus (Fig. 5.11A).

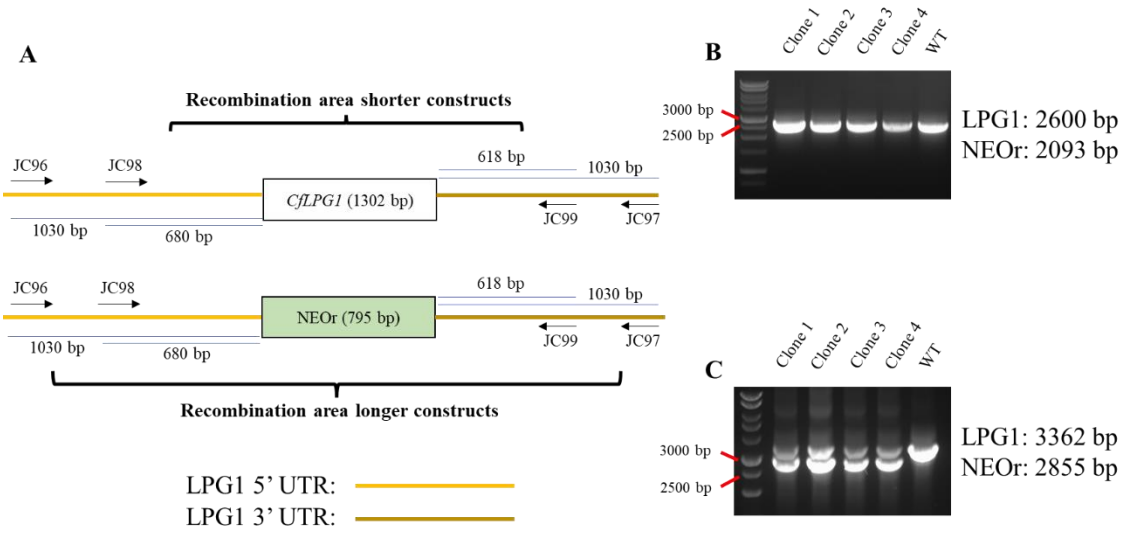


Figure 5.11: PCR analysis to identify potential Δ*CfLPG1*::*NEOr* sKOs after electroporations with gene replacement constructs of different length. (A) Primers ("JC-number") used for this analysis are shown in their complementary position in the *CfLPG1* locus. (B) EtBr-stained 1% agarose gel loaded with JC98 + JC99 PCR product using as template the gDNA from 4 clones transfected with shorter constructs (Fig. 10, right scheme) and WT cells. The length of the expected amplicons for a sKO gDNA are specified. (C) EtBr-stained 1% agarose gel loaded with JC96 + JC97 PCR products using as template the gDNA from 4 clones transfected with longer constructs (Fig. 10, left scheme) and WT cells. The length of the expected amplicons for a sKO gDNA are specified.

In the case of clones electroporated with the smaller constructs, only the amplicon of the original gene was detected in all cases (Fig. 5.11B), similar to the positive controls using WT gDNA (Fig. 5.11B, WT lane). In contrast, all clones from the experiment using

the NEOr construct containing 1 kb *CfLPG1* UTRs were positive (Fig. 5.11C), amplifying both expected amplicons for Δ*CfLPG1*::*NEOr* sKO clones.

Following the identification by PCR of putative Δ*CfLPG1*::*NEOr* sKO clones, the gDNA of 4 such clones, and of WT cells, were used for genotype analysis by Southern blot using DIG-labelled full-length probes for the *CfLPG1* and *NEOr* ORFs (Fig. 5.12).

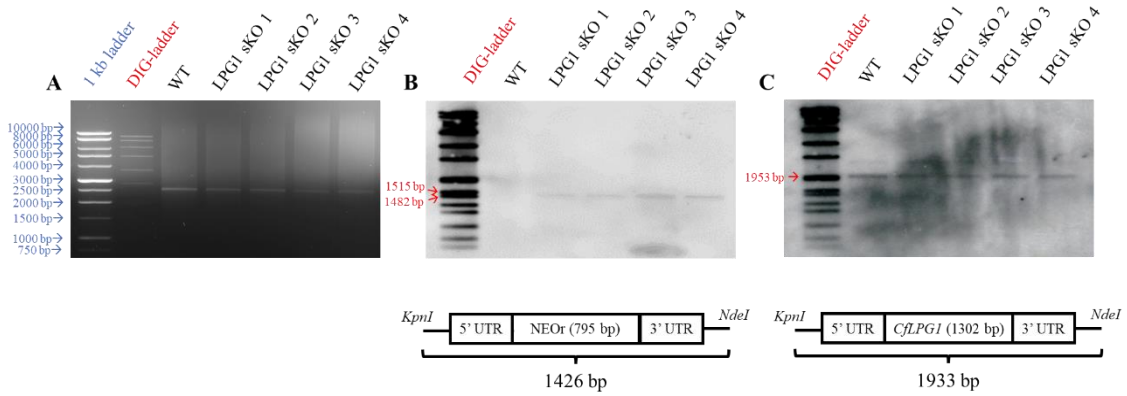


Figure 5.12: Genotype analysis by Southern blot of the Δ*CfLPG1*::*NEOr* sKO mutants. (A) EtBr-stained 0.8% agarose gel containing KpnI and NdeI-digested gDNAs of WT and sKO cells. After transferring DNA to positively charged nylon membrane, it was probed against *NEOr* ORF (B), stripped and re-probed against *CfLPG1* ORF (C). DIG-labelled ladder was used to determine the size of the detected band in (B) and (C). Schemes of the detected loci are shown under the developed chemiluminescence films in (B) and (C).

The Southern blot analysis showed a unique band of correct size corresponding for the *NEOr*-containing gene locus in the sKO clones (Fig. 5.12B) and, following blot stripping and re-probing, for the remaining *CfLPG1* gene locus (Fig. 5.12C). Interestingly, similar to (Fig. 5.8A), a band of around 2,500 bp was detected in the EtBr-stained agarose gel prior to DNA transfer (Fig. 5.12A).

Having obtained several Δ*CfLPG1*::*NEOr* sKO clones, we carried on with the replacement of the second allele with the *HYGr* gene. In the following electroporations of these sKOs cells, the *HYGr*-harbouring constructs with 1 kb UTRs were the only ones used. Despite several electroporations performed using different DNA quantities of the digested recombination template, none of the cultures survived in the presence of both drugs.

Due to the lack of time to perform further experiments, we were generously assisted by Dr. Maria Lucia Sampaio Güther (Dr MLSG), who carried on with the attempts of *CfLPG1* dKO generation with the following protocol modifications (see section 3.5.5): (i) the quantity of cells per cuvette were increased (4 x 10^7 cells instead of standard quantity of 1.35 – 1.8 x 10^7 cells), (ii) cells were suspended in 100 µl home-made Cytomix buffer, instead of 450 µl, (iii) 1 µg of digested 1 kb UTRs *HYGr* gene replacement template was used for each electroporation to avoid non-specific insertion

the resistance gene, (iv) electroporation was performed in a Lonza Nucleofector device (instead of BTX device), (v) cloning was by limiting dilution in 96-well plates (0.5 cells/well or 2.5 cells/ml) with media containing both drugs after only 4 h recovery without antibiotics to separate potentially slow growing dKO mutants away from sKO cells that might otherwise take over the cultures. Using this approach, Dr. MLSG was able to obtain some hygromycin + neomycin-resistant clones.

Of these, one clone was expanded to generate stabilates and to extract gDNA. after which an initial PCR analysis of the *CfLPG1* locus was performed. Our PCR strategy (Fig. 5.13A) consisted of locus amplification and further NcoI digestion of the purified amplicons from WT, Δ*CfLPG1*::*NEOr* sKO and potential Δ*CfLPG1*::*NEOr*/Δ*CfLPG1*::*HYGr* dKO gDNAs. This experiment was performed by Dr. MLSG, and the result (Fig. 5.13B) showed that all the expected DNA fragments were present in the gel: The original *CfLPG1* band (3362 bp) was amplified from WT and sKO gDNAs, but not from the dKO gDNA. An amplicon corresponding to the *NEOr*-containing locus (2855 bp) appeared in both sKO and dKO lanes, and a third (3086 bp) amplicon was only detected in the dKO lane. All three ORFs (*CfLPG1*, *HYGr* and *NEOr*) have a single NcoI restriction site. Based on this, the PCR products were purified and subjected to NcoI digestion, to then be included in the agarose gel. The resulting fragments of calculated size were detected in all cases, further pointing towards the obtention of a true Δ*CfLPG1*::*NEOr*/Δ*CfLPG1*::*HYGr* dKO cell line.

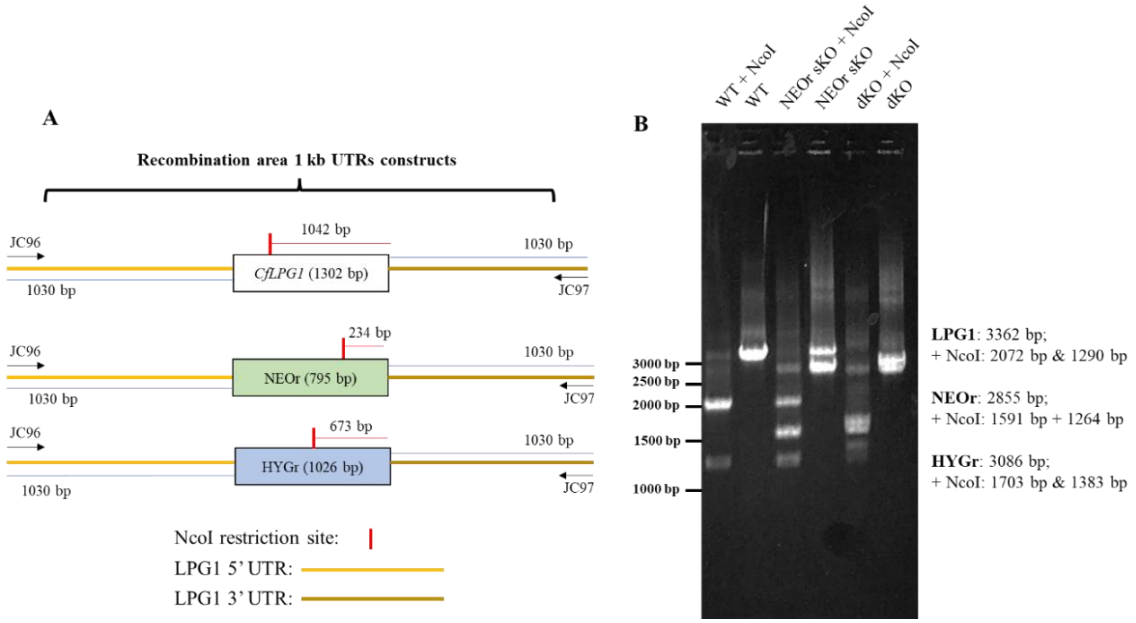


Figure 5.13: PCR analysis to identify potential Δ*CfLPG1*::*NEOr*/Δ*CfLPG1*::*HYGr* dKOs. (A) JC96 and JC97 primers used for this analysis are shown in their complementary position in the *CfLPG1* locus. The location of the single NcoI restriction site in the amplified sequences is showed in red. (B) EtBr-stained 1% agarose gel loaded with JC96 + JC97 PCR product using as template the gDNA from WT, NEOr sKO clones. The length of the expected amplicons before or after being NcoI-digested are specified. Experimental design belongs to this thesis' author. Experiment was generously performed by Dr. MLSG.

To fully confirm this, the dKO gDNA was genotyped along with WT and Δ*CfLPG1*::*NEOr* sKO gDNAs by Southern blot using DIG-labelled probes for the *CfLPG1*, *NEOr* and *HYGr* ORFs (Fig. 5.14). In this way, a unique band of correct size corresponding to the digested *CfLPG1* gene locus in the WT and sKO gDNAs was detected, whereas no such band was visible in the lane containing dKO gDNA (Fig. 5.14A). This same nylon membrane was used for a second probing against *HYGr* ORF. As shown in (Fig. 5.14B), the selected restriction enzymes NdeI and KpnI cut this resistance marker ORF in two smaller fragments of 693 bp and 964 bp (Fig. 5.14B, green and red arrow), which were successfully detected. Two chemiluminescence films of different exposure time are shown to confirm the presence of both fragments detected by the *HYGr* DIG-labelled probe. In parallel, another nylon membrane containing the same amount and loading order of the three analysed gDNAs was used to probe against the *NEOr* ORF (Fig. 5.14C). In this case, the *NEOr* undigested ORF was detected only in the sKO and dKO lanes. The EtBr-stained gel pre-transfer to the nylon membrane confirms the correct digestion of similar quantity of DNA in all lanes (Fig. 5.14A).

This experiment, also performed by Dr. MLSG, confirmed us that the *HYGr* ORF replaced the second *CfLPG1* allele correctly, that no extra copies of either the original gene or of the resistance markers are present, and that, therefore, a *CfLPG1* null mutant has been successfully generated. This mutant should be unable to synthesize [^{3}H]-labelled LAG when cultured in [^{3}H]Fuc-containing glucose & glucosamine-depleted medium (to be confirmed).

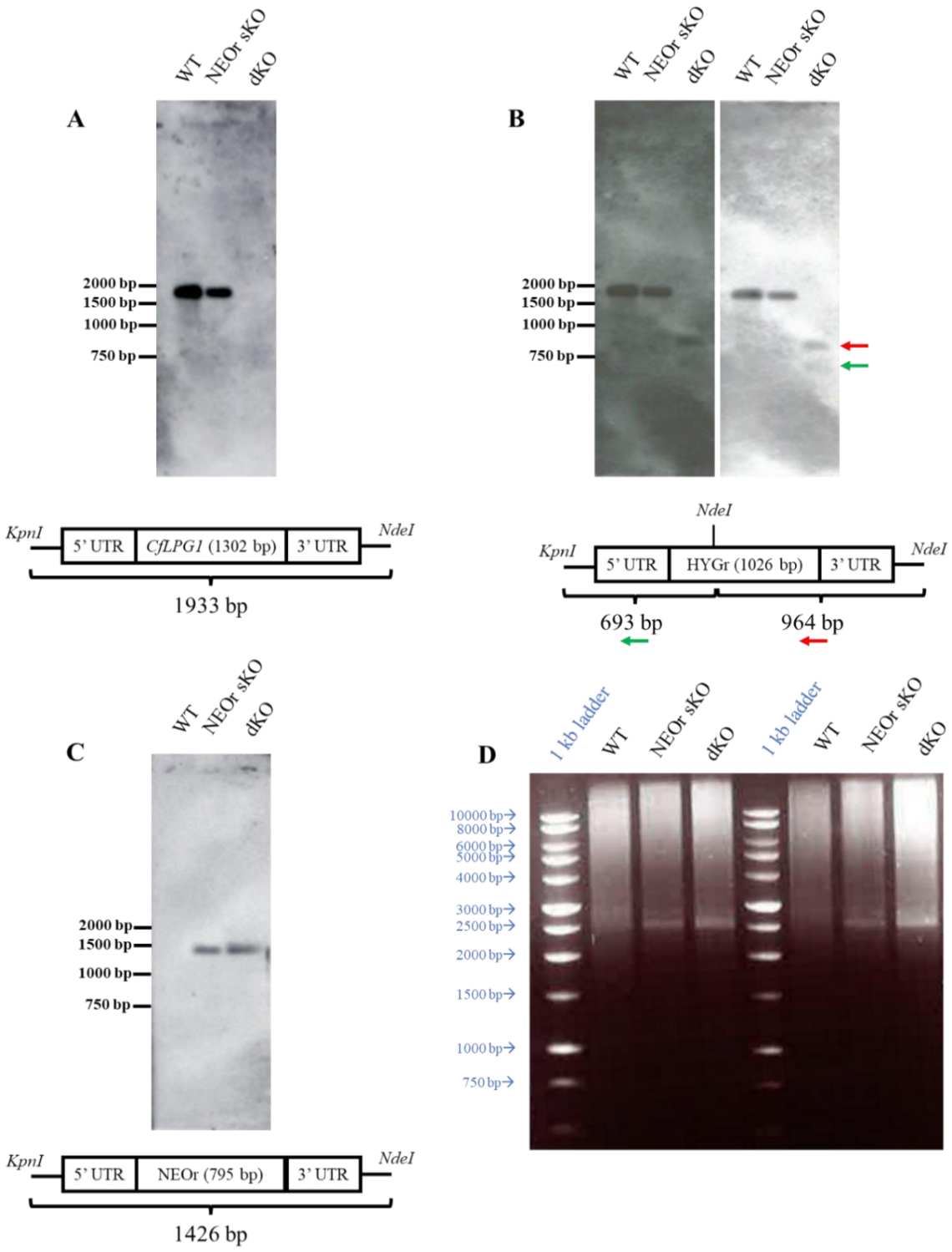


Figure 5.14: Genotype analysis by Southern blot of the Δ*CfLPG1*::*NEOr*/Δ*CfLPG1*::*HYGr* dKO mutant. After transferring DNA to a positively charged nylon membrane, this was cut in two pieces containing the gDNAs of WT, NEOr sKO and potential dKO. One of the pieces was probed first against *CfLPG1* ORF (A), and then re-probed, without stripping, against HYGr ORF (B). Due to the presence of a NdeI site in the sequence of this resistant marker gene, two fragments were detected. The other nylon piece was probed against NEOr ORF (C). 1 kb ladder in the original agarose gel (D) was used to determine the size of the detected band in A, B and C. Schemes of the detected loci are shown under the developed chemiluminescence films in A, B and C. In all cases, DIG-labelled probes synthesized by amplification of the whole gene ORFs were used. (D) Ethidium bromide (EtBr)-stained 0.8% agarose gel containing KpnI and NdeI-digested gDNAs of WT, NEOr sKO and dKO cells. Experimental design belongs to this thesis' author. Experiment was generously performed by Dr. MLSG.

6. DISCUSSION

6.1 Attempts to express active TcFUT1 in *E. coli* and eukaryotic expression systems.

We were unable to achieve the expression of active rTcFUT1 in *E. coli* using the methodology reported for rTbFUT1 in (Bandini *et al*., 2021). To confirm that this was not for any technical reason, we reproduced the expression of soluble, active TbFUT1 in *E. coli* (Fig. 4.21) and used this as a positive control in several experiments.

Although TcFUT1 and TbFUT1 are orthologous proteins with 68.7% sequence identity, rTcFUT1 made in *E. coli* generally failed to bind to IMAC matrices. The molecular modelling of rTcFU11-1.a based on its Robetta-predicted structure (Baek *et al.*, 2021), analysed with ChimeraX v1.4 software (Pettersen *et al.*, 2021), may explain this since the 6xHis tag is predicted to form an a-helix buried within the protein structure (Fig. 6.1). Another issue may be misfolding and aggregation of the rTcFUT1 proteins. Hints that this might have been a problem came from our inability to efficiently elute the small amounts of rTcFUT1 that did bind to IMAC matrices and from the co-elution of rTcFUT1 constructs with *E. coli* chaperonins (see Fig. 4.3, Fig. 4.4, Fig. 4.9 and Fig. 4.10).

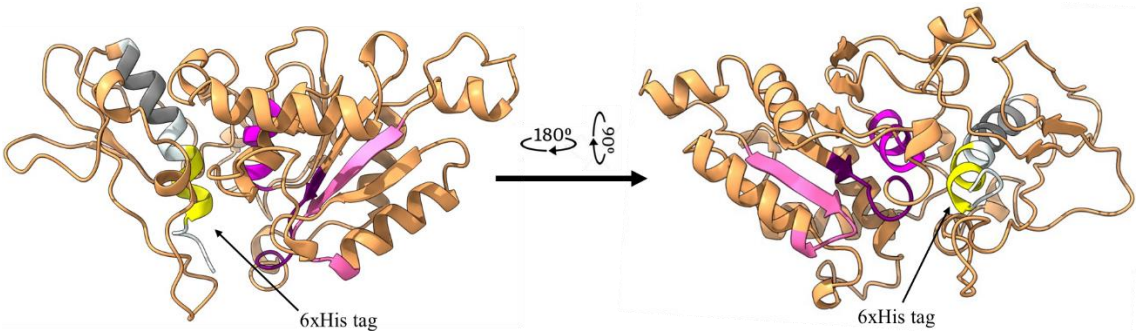


Figure 6.1: 3D model of rTcFUT1-1.a recombinant product (6xHis-TEV site-TcFUT1cat). Protein sequence is coloured in orange, containing a N-terminal 6xHis tag (yellow) followed by a TEV site (grey). GT11 conserved motifs I, II and III (see Introduction) are depicted in different tones of pink. Motif IV is absent since it is located in the predicted MTS sequence. RoseTTAFold was used as modelling method. Global Distance Test (GDT) prediction score: 0.78.

For example, in the case of construct rTcFUT1-3.b (6xHis-GST-TEV-TcFUT1) the overexpressed product was not recovered from the matrix unless very high imidazole concentrations were used (Fig. 4.9). The 3D model of this recombinant product (Fig. 6.2A) predicted the 6xHis tag to be exposed and therefore presumably available for binding to nickel. However, in this case, the presence of the strong solubilising tag GST may have induced the formation of micelles (Costa *et al.*, 2014) perhaps causing the protein to detach from the column, even in the absence of imidazole, and thus mostly

appear in the flow through. On the other hand, when MBP was used as a solubilising tag (construct rTcFUT1-2.b) (Fig. 6.2B), purification of an almost pure product was achieved (Fig. 4.10). In this case the modelling predicted that the 6xHis tag is exposed in a helical structure (Fig. 6.2B), but the TcFUT1 structure is predicted to be less well folded in the context of this chimeric protein, which may explain the inactivity of this product. Indeed, TcFUT1 seems to form a trapezoidal prism shape which contains a potential catalytic pocket in its apex (Fig. 3), where the four motifs (section 1.5) colocalize. In contrast, in the case of rTcFUT1-2.b (Fig. 6.2B), the motifs don't seem to form a pocket, particularly motif IV (cyan).

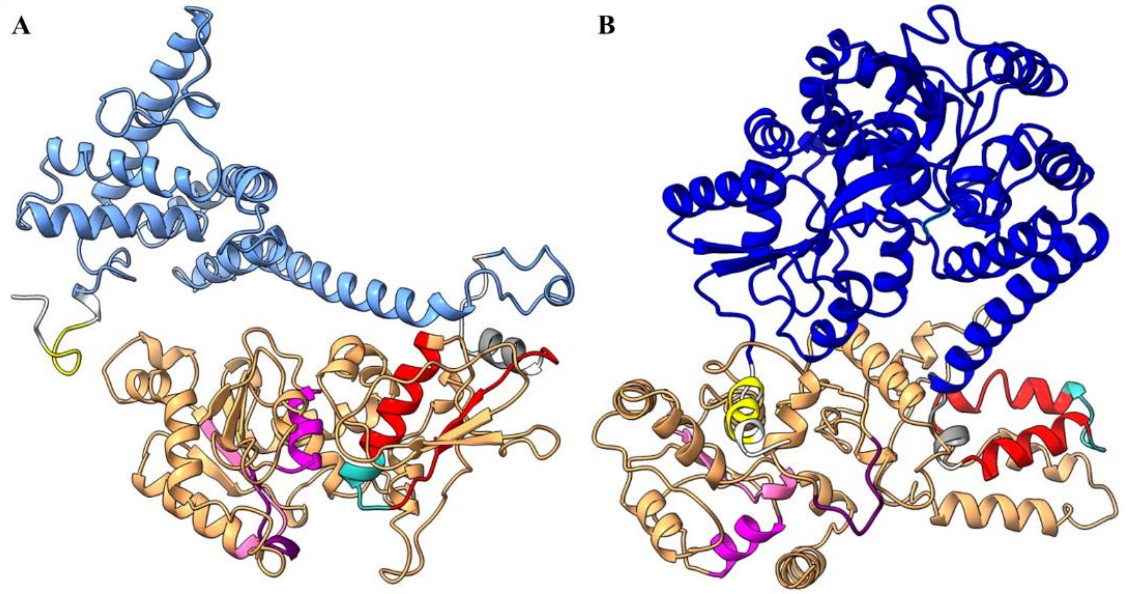


Figure 6.2: 3D model of (A) rTcFUT1-3.b recombinant product (6xHis-GST-TEV site-TcFUT1) and (B) rTcFUT1-2.b recombinant product (6XHis-MBP-TEV site-TcFUT1). 6xHis: yellow, GST: light blue, MBP: dark blue, TEV site: grey, MTS: red, containing motif IV (cyan). Motifs I, II and III are coloured in different tones of pink. RoseTTAFold was used as modelling method. Global Distance Test (GDT) prediction score: 0.79 (A) and 0.80 (B).

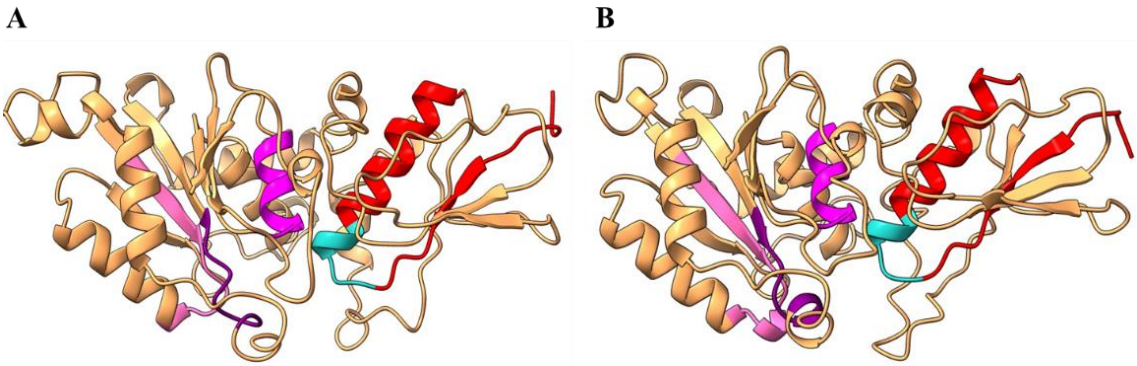


Figure 6.3: 3D model of (A) TcFUT1 and (B) TbFUT1. Same scheme colours as in previous models. RoseTTAFold was used as modelling method. Global Distance Test (GDT) prediction score: 0.74 (A) and 0.75 (B).

While other experiments could have been done to probe these suggestions (for example, gel filtration to assess for protein aggregation) we moved on to eukaryotic protein expression systems at this point.

Using the eukaryotic expression platform previously reported to be successful for the characterization of many human GTs (Moremen *et al.*, 2018), we attempted to obtain rTcFUT1 from transfected HEK293-derived, suspension-cultured, Expi293F cells. After an initial screening of the different constructs generated, rTcFUT1-4.a seemed to be the most promising one. Its predicted structure (Fig. 6.4A) showed an accessible 8xHis tag and a "pocket"-like cavity including all the different motifs. Even though we were unable to purify it completely by IMAC, we decided to use it for activity assays, the results of which tentatively suggested a possible GDP-Fuc → GDP + L-Fuc hydrolytic activity for the recombinant product. However, when used alongside recombinant TbFUT1 in GDP-glo assays for the detection and quantification of hydrolysed GDP, no activity was detected.

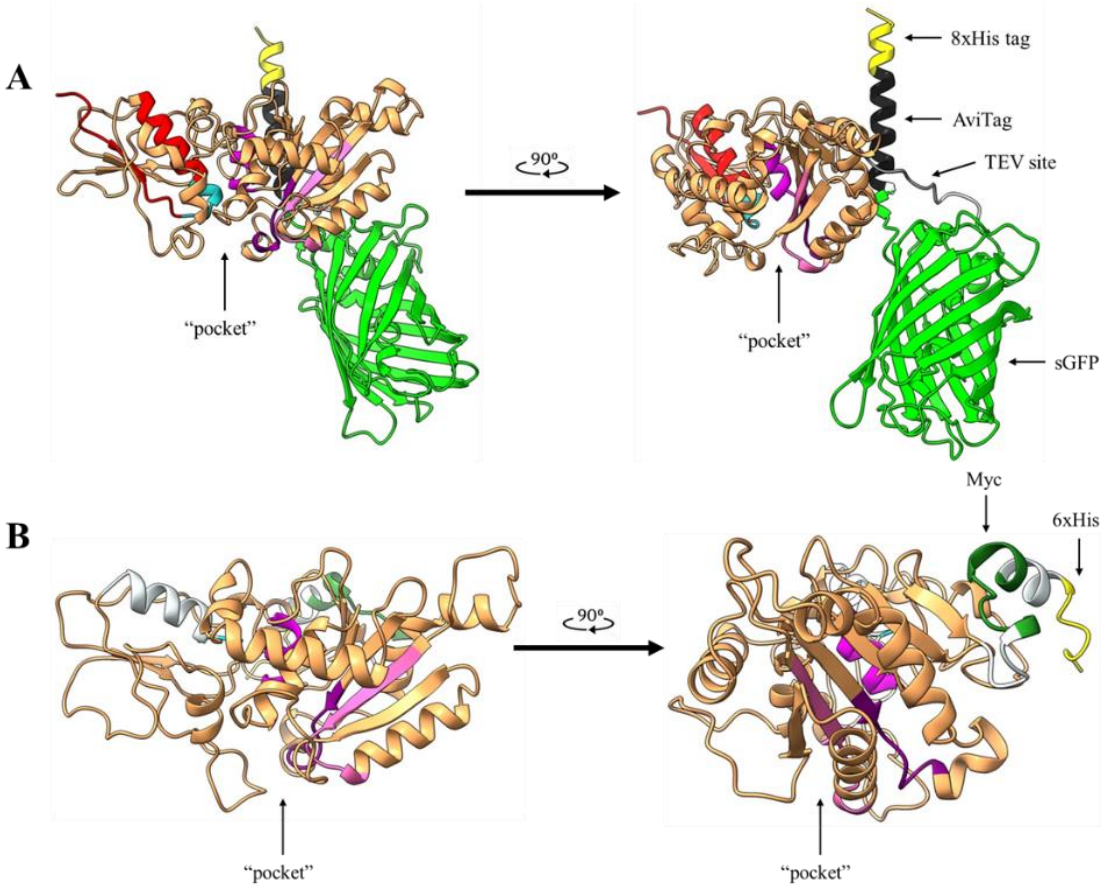


Figure 6.4: 3D model of (A) rTcFUT1-4.a (TcFUT1-TEV site-AviTag-8xHis) and (B) rTcFUT1-6.a (TcFUT1cat-Myc-6xHis). In (B), a small α-helix peptide adjacent to the N-terminus of TcFUT1cat is coloured in white and links with the 2 amino acids left after the cleavage of the CSS. RoseTTAFold was used as modelling method. Global Distance Test (GDT) prediction score: 0.83 (A) and 0.77 (B).

Unfortunately, the recombinant product from construct rTcFUT1-6.a, which was secreted to the culture medium, was mostly unable to bind to nickel matrix when attempting a large-scale purification, appearing mostly in the flow through. Analysis of the predicted model (Fig. 6.4B) indicates that the 6xHis tag should have been exposed and available for binding to nickel. It is possible that the rTcFUT1-6.a protein product is

aggregated in some way that shields the His tag. In contrast, the Myc tag was available for interaction with anti-Myc antibodies for successful IP, but elution from the immobilised antibodies with Myc peptide was not successful. This latter observation may also be consistent with rTcFUT1-6.a aggregation, such that the effective high-valency of the aggregates for the anti-Myc antibodies prevents the elution with Myc peptide ligand. In retrospect, this aggregation hypothesis could have been tested by gel-filtration studies. Despite the caveat of possible aggregation, on-bead anti-Myc immunoprecipitated rTcFUT1-6.a gave us our first preparation with GDP-Fuc hydrolytic activity, although no [^{3}H]-labelled product was detected in later activity assays.

In summary, the different recombinant products that were successfully expressed in Expi293F cells were either unable to bind to nickel matrices for purification (1.a, 3.b, and 6.a), or largely or completely inactive following enrichment (2.b, 4.a, and 6.a when immunoprecipitated).

Taken together, these multiple attempts with different constructs made us hypothesise that TcFUT1 may obligatorily require its original environment in order to be fully functional. Thus, we generated TcFUT1-overexpressing *T. cruzi* cells and tested this hypothesis by using the immunoprecipitated overexpressed product as the enzyme source for activity assays. Eventually, we were able to detect the generation of a [^{3}H]-labelled product obtained in these assays, supporting our hypothesis.

Specifically, when overexpressed and immunoprecipitated TcFUT1 was incubated along with LNB and home-made radioactive sugar donor, we detected for the first time the generation of a radiolabelled product. Subsequently, an assay was designed to shed light on the substrate specificity of TcFUT1, involving the use of different glycan acceptors (Fig. 4.12). The features of acceptor substrate recognition by TcFUT1 that we can extract form the substrate specificity study are summarised in (Fig. 6.5). The components of these saccharides and a scheme of the fucosyltransferase reactions performed are shown in (Fig.6.5A and B).

The results (Fig. 4.35) demonstrated a preference for acceptors with non-reducing terminal Galβ1-3GlcNAc motifs (such as LNB, LNB-OMe and LNT) (Fig.6.5C). Previously, the substrate specificity of TbFUT1 was tested in a similar manner (Bandini, 2011; Bandini *et al.*, 2021), which showed a similar preference for non-reducing terminal Galβ1-3GlcNAc motifs but which could also recognise Galβ1-4GlcNAc terminating acceptors. Nevertheless, for both TcFUT1 and TbFUT1 the preferred substrate of those

tested was clearly LNB-OMe, (Fig. 6.5C), based on the TLC fluorographic band intensity. Indeed, for TbFUT1 the presence of LNB-OMe greatly reduced the background hydrolysis of GDP-Man, suggesting it can outcompete water in the acceptor binding site (Bandini *et al.*, 2021), whereas this was not the case for in the case of TcFUT1 (Fig. 4.35. lane +LNB-OMe).

The superiority of β-methyl glycoside LNB-OMe (Galβ1-3GlcNAcβ1-O-Me) over LNB as an acceptor substrate suggests that both enzymes prefer that the GlcNAc residue is present as the β-anomer (LNB with a non-methylated, free reducing terminus, will undergo mutarotation between open chain and α- and β-anomeric forms) (Fig. 6.5E). Nevertheless, we did not perform the same reactions using the α-glycoside forms of the methylated synthetic acceptors, therefore we cannot discard the possibility of TcFUT1 substrate recognition not being dependent of the GlcNAc residue anomericity. The reason for not including these substrates in our assays was due to their unavailability from any commercial source. Furthermore, for both TbFUT1 and TcFUT1, Galβ1-O-Me glycoside (Fig. 6.5D) was not a suitable acceptor, indicating that the GlcNAc of the Galβ1-3GlcNAc disaccharide motif is necessary for substrate recognition.

For TbFUT1, characterisation of the trisaccharide product defined the enzyme as a GDP-Fuc : βGal α1-2 fucosyltransferase catalysing the reaction: GDP-Fuc + Galβ1-3GlcNAcβ1-OR → GDP + Fucα1-2Galβ1-3GlcNAcβ1-OR (Bandini, 2011; Bandini *et al.*, 2021). Most likely the same is true for TcFUT1, although product characterisation has not yet been performed. In the future, a simple digestion experiment on the putative [^{3}H]Fucα1-2Galβ1-3GlcNAcβ1-OMe radioactive reaction product with a specific α1-2 fucosidase (like the one from *Xanthomonas manihotis* (Wong-Madden and Landry, 1995) could resolve this. The Galβ1-3GlcNAcβ1-O-Me acceptor has only one free 2-hydroxyl group, that of the βGal residue, since the βGlcNAc C2-position is occupied by the N-acetamido group. Therefore, a positive digestion with *X. manihotis* α1-2 fucosidase, releasing free [^{3}H]Fuc (detected by HPTLC and fluorography) would define the product as [^{3}H]Fucα1-2Galβ1-3GlcNAcβ1-O-Me. This experiment is currently being designed.

Significantly, when TcFUT1 was incubated with Galβ1-3GalNAcβ1-O-Me (GalGalNAc-OMe) (Fig. 6.5D) no product was detected. This acceptor differs from LNB-OMe only on the N-acetyl-hexosamine (HexNAc) residue: N-acetyl-galactosamine (GalNAc) instead of N-acetyl-glucosamine (GlcNAc). These two HexNAcs are 4-epimers of each other, suggesting that the acceptor binding site of TcFUT1 recognises the 4-hydroxyl group of GlcNAc in the equatorial configuration, and not that of GalNAc,

which is axial. In a similar manner, this result suggests that a D-Gal residue, also defined by an axial 4-hydroxyl group, would not be recognised.

Furthermore, no product was detected when TcFUT1 was incubated with lactose, LacNAc, or GNG (Fig. 6.5D), suggesting that the enzyme may be only able to recognize the disaccharide motif when it involves a β1-3 glycosidic bond, but not when they are linked by a β1-4, or a β1-6 bond.

Finally, although our results suggest that the Galβ1-3-linked D-glucopyranose residue may need to be in its N-acetylated amido form to be recognised (Fig. 6.5E), an additional Galβ1-3Glc acceptor could have been included in our tests. If no product was detected using this disaccharide, this would be a confirmation of the necessity of an N-acetamido group in the C2-position of the D-glucopyranose residue. This reaction was not included in our assays due to the extreme rarity of this sugar, which has not been identified in Nature, although it can be synthetically manufactured (CAS: 28447-38-3) by some specialized companies (Biosynth AG, CymitQuimica S.L and BOC Sciences).

In addition, once we obtained mixed populations of *T. cruzi* overexpressing TcFUT1-6xMyc, these were used for IFA microscopy together with MitoTrackerTM. While these results are preliminary and a wider range of organelle markers need to be included (including antibodies to mitochondrial proteins as an alternative to MitoTrackerTM), they are at least consistent with a mitochondrial location for TcFUT1-6xMyc. As noted for TbFUT1, the predicted presence of a cleavable mitochondrial targeting sequence (MTS) in the first 32 amino acids of TcFUT1 suggests a mitochondrial matrix specific localization of the imported and processed protein (Kunze and Berger, 2015; Dimogkioka *et al*., 2021), although it is important to note that one of its four motifs (motif IV – a motif of unknown function) is predicted in this same MTS sequence, as shown in (Fig. 6.3). Whether the MTS and motif IV are cleaved in vivo during mitochondrial import in any of the kinetoplastids is currently unknown. The lower MW version of TcFUT1-6xMyc in the MP1 cell population (Fig. 4.32) may be the first indication of this. However, the latest proteome-wide analysis of T. cruzi (performed in the strain CL Brener) (Avila *et al*., 2018) failed to identify TcFUT1 in exponential and stationary phase epimastigotes, suggesting, as predicted for TbFUT1 (Bandini *et al*., 2021) and found by proteomics (Tinti *et al*., 2019), that native TcFUT1 is a very low-abundance protein. Therefore, we must be careful not to overinterpret our TcFUT1-6xMyc overexpression data.

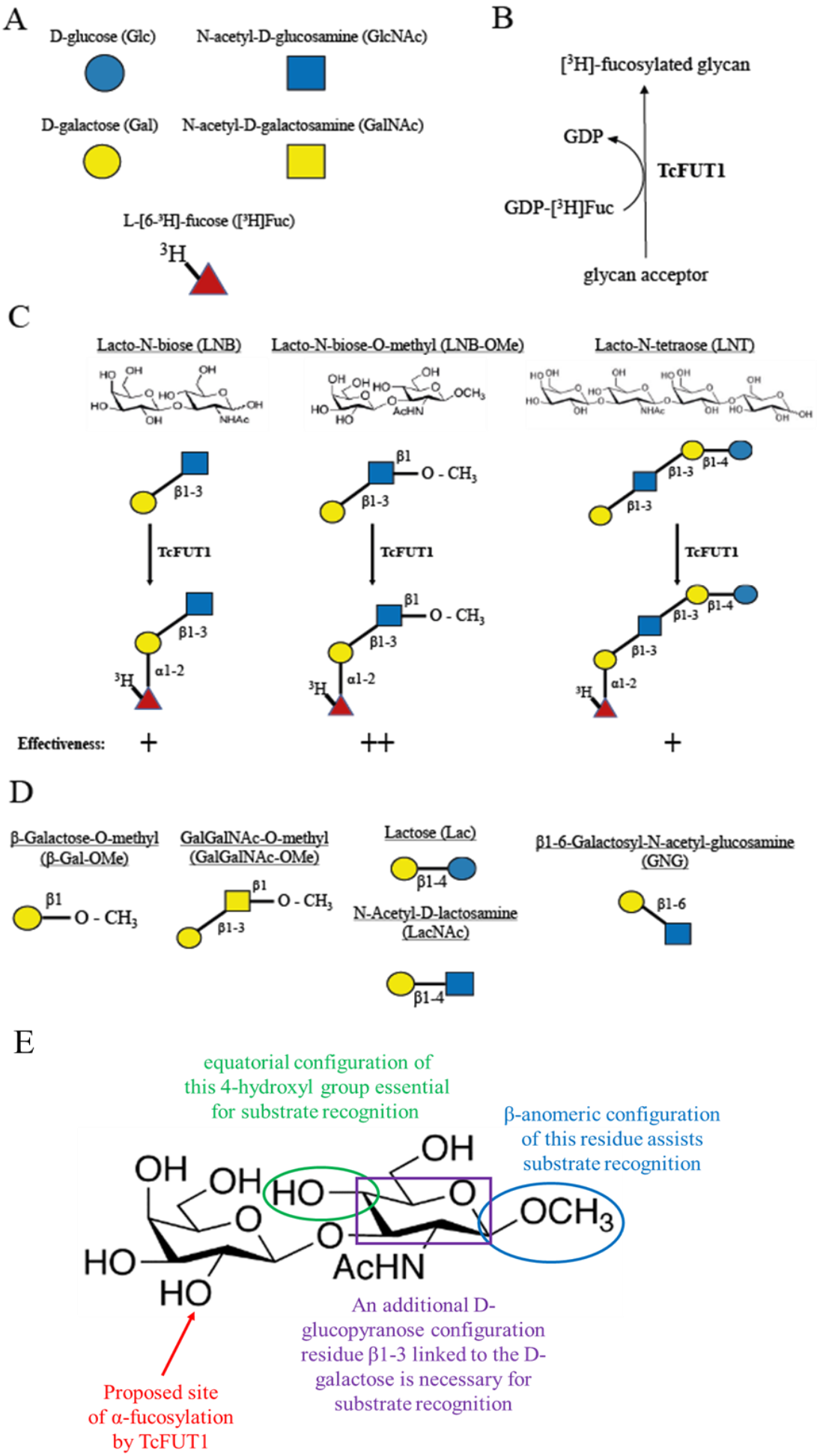

A
D-glucose (Glc)
N-acetyl-D-glucosamine (GlcNAc)
D-galactose (Gal)
N-acetyl-D-galactosamine (GalNAc)
L-[6-³H]-fucose ([³H]Fuc)
B
[³H]-fucosylated glycan
GDP
TcFUT1
GDP-[³H]Fuc
glycan acceptor
C
Lacto-N-biose (LNB)
Lacto-N-biose-O-methyl (LNB-OMe)
Lacto-N-tetraose (LNT)
TcFUT1
Effectiveness: + ++ +
D
β-Galactose-O-methyl (β-Gal-OMe)
GalGalNAc-O-methyl (GalGalNAc-OMe)
Lactose (Lac)
N-Acetyl-D-lactosamine (LacNAc)
β1-6-Galactosyl-N-acetyl-glucosamine (GNG)
E
equatorial configuration of this 4-hydroxyl group essential for substrate recognition
β-anomeric configuration of this residue assists substrate recognition
An additional D-glucopyranose configuration residue β1-3 linked to the D-galactose is necessary for substrate recognition
Proposed site of α-fucosylation by TcFUT1

Figure 6.5 (*previous page*): Summary of TcFUT1 substrate specificity assay. (A) Monosaccharide symbol nomenclature. (B) Schematic representation of the fucosyltransferase reactions performed in our radioactivity-based assays. The detection of the [^{3}H]-fucosylated glycan, or any other radioactive molecule, was done by fluorography. (C) The three glycan acceptors recognised by TcFUT1 in our assay. The substrate effectiveness was defined based on the fluorographic intensity of the TLC bands corresponding to the radiolabelled reaction products. (D) The five glycan acceptors that were not recognised by TcFUT1 in our assay. (E) LNB-OMe chair conformation with indications of the chemical features needed for TcFUT1 recognition, inferred from our substrate specificity assay.

6.2 Enzymatic synthesis of GDP-[^{3}H]Fuc.

Due to the unavailability of commercially sourced tritiated nucleotide sugar GDP-[^{3}H]Fuc, necessary for our activity assays with TcFUT1, we aimed to establish a robust protocol for its synthesis. The synthesis of nucleotide sugar donors in trypanosomatids, and other eukaryotes, occurs through *de novo* and/or salvage pathways, depending on which enzymes are encoded in the genome and expressed in the cell (see section 5.1 and Fig. 5.1). Nucleotide sugars are then used by GTs, which transfer the sugar component to a protein, lipid, or growing glycan chain.

For the generation of GDP-Fuc and GDP-Ara in *L. major*, the *FKP40* and *AFKP80* genes encode bifunctional (sugar 1-kinase and pyrophosphorylase) enzymes which salvage L-Fuc alone (FKP40) or both L-Fuc and D-Ara (AFKP80) and convert them into GDP-Fuc and GDP-Ara (Guo *et al.*, 2017). The configuration of the hydroxyl groups of D-Ara and L-Fuc are identical at carbon positions C2 to C5, and they differ only in that D-Ara has an extra methyl group bound to the C5 of L-Fuc (Fig. 6.6A). AFKP80 is a homodimer, composed of two 126.5 kDa monomers, and was identified in the cytoplasm of *L. major* (Novozhilova and Bovin, 2009). *L. major* AFKP80 catalyses the reactions of phosphorylation and pyrophosphorylation required for nucleotide sugar synthesis (Fig. 6.6B). First, an ATP-dependent kinase activity phosphorylates L-Fuc/D-Ara at C1 position. Second, the pyrophosphorylase activity converts the sugar phosphate and one GTP to GDP-Ara/GDP-Fuc and inorganic pyrophosphate (PP_i). The PP_i product is known to generally cause product-inhibition of the forward pyrophosphorylase reaction (Kornberg, 1957, 1962), drastically reducing the generation of nucleotide sugars. There are orthologues of *L. major* AFKP80 and FKP40 in other *Leishmania* spp., *T. cruzi* and *C. fasciculata.*

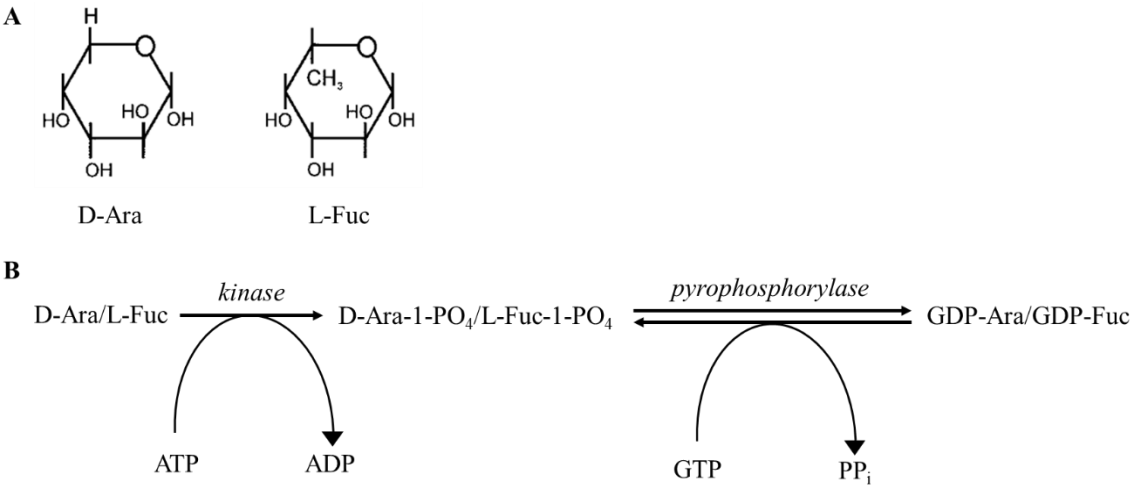


Figure 6.6: Structural similarity and salvage pathway of D-Ara and L-Fuc. (A) Structures of D-Ara and L-Fuc, used by AFKP80 and FKP40 in *L.major* for the synthesis of GDP-Ara and GDP-Fuc or GDP-Fuc alone, respectively. (B). Enzymatic scheme of GDP-Ara and GDP-Fuc synthesis. The two activities of the bifunctional enzymes are written in *italics*.

(Schneider *et al.*, 1996) used *C. fasciculata* cytosolic fraction for the chemo-enzymatic synthesis of tritium-labelled GDP-Ara (GDP-[^{3}H]Ara). In this case, the cytosolic fraction was used as a crude source of Ara1-kinase activity to generate [^{3}H]Ara1-P and a chemical step using GMP-morpholidate (Moffatt, 1966; Feingold and Barber, 1990) was used to convert it to GDP-[^{3}H]Ara. It was also in (Schneider *et al.*, 1996) where the enzymatic pathway depicted in (Fig. 6.6B) was first postulated. Later, (Mengeling and Turco, 1999) were able to perform the Ara1-kinase and pyrophosphorylase reactions simultaneously by adding GTP to the kinase reaction mixture. These authors also used an ultracentrifugation step to remove all membrane-bound proteins from the *C. fasciculata* cytosolic fraction. Two other important components were added to the reaction: the phosphatase inhibitor sodium fluoride (NaF), to avoid the degradation of the sugar phosphate intermediate product, and yeast inorganic pyrophosphatase (yeast PPase), which hydrolyses PP_i to inorganic phosphate and therefore shifts the equilibrium of the pyrophosphorylase reaction (Fig. 6.6B) towards the right. Importantly, it was in this work where, based on the structural similarity between D-Ara and L-Fuc, that GDP-[^{3}H]Fuc was also synthesized following the same procedure.

Based on this previous research, we generated an alternative protocol for the synthesis of GDP-[^{3}H]Fuc (Fig. 4.41), which most probably is equally useful for the synthesis of GDP-[^{3}H]Ara. In our procedure, after lysing similar number of cells (1×10^9) by two freeze-thaw cycles, the centrifuged and 0.45 µm-filtered supernatant is subjected to diafiltration in a 100 kDa cut-off centrifugal diafiltration device in order to remove salts, small metabolites and proteins <100 kDa. It is during these diafiltration steps that

NaF is added. The resulting crude cytosolic enzyme preparation is mixed with [^{3}H]Fuc, yeast PPase (at double the concentration originally used), $MgSO_4$, ATP and GTP. Reactions are stopped by the addition of EDTA which inhibits the Mg^{2+}-dependent kinase step.

Initial tests, analysed by HPTLC after optimization of the solvent used, confirmed the previously reported (Mengeling and Turco, 1999) total conversion of [^{3}H]Fuc to GDP-[^{3}H]Fuc in 30 min (Fig. 4.40). After scaling-up the reaction using up to 50 µCi [^{3}H]Fuc, HPTLC analysis showed a similar result, although only a partial conversion into GDP-[^{3}H]Fuc was achieved. Since these HPTLC analyses showed the [^{3}H]Fuc starting material and the GDP-[^{3}H]Fuc, but not the [^{3}H]Fuc1-P intermediate, we concluded that the rate limiting step in the reaction is the Fuc1-kinase activity, which might be overcome by the addition of higher concentration of ATP. This could be tested in future optimisation work. In a later synthesis reaction (using unreacted [^{3}H]Fuc from previous experiments), a longer incubation time with freshly-prepared enzyme gave a higher yield of GDP-[^{3}H]Fuc. Altogether, there are number of parameters that could be optimised in the synthesis of GDP-[^{3}H]Fuc using *C. fasciculata* AFKP80 activity, and this could include replacing the crude enzyme preparation (which was somewhat labile on storage) with a recombinant version. While beyond the scope of this thesis, which only required sufficient GDP-[^{3}H]Fuc to characterise out TcFUT1 preparations, a fully optimised process to synthesise GDP-[^{3}H]Fuc from [^{3}H]Fuc with a stable, recombinant, AFKP80 preparation, could re-establish commercial source of GDP-[^{3}H]Fuc for an appropriate manufacturer.

6.3 Biosynthetic [^{3}H]Fuc-labelling of *C. fasciculata* to find endogenous fucosylated substrates.

There are a few reports of glycosylation in mitochondria, for example, glycoproteomic approaches suggest the presence of glycosylated isoforms of several mitochondrial proteins in yeast and bovine hearts (Kung *et al.*, 2009; Burnham-Marusich and Berninsone, 2012). However, it has not been determined whether they were imported into the organelle already glycosylated or whether they were modified within the mitochondria by as yet unknown GTs. The only characterised mitochondrial GT is the mitochondrial isoform of the mammalian O-GlcNAc transferase (OGT), whose cell biology and potential as a therapeutical target for different types of cancer and metabolic

disease therapy has been extensively reviewed (Bond and Hanover, 2015; Gonzalez-Rellan *et al.*, 2022; Liu and Chiu, 2022). This protein catalyses the glycosyl linkage of a single β-GlcNAc, an uncharged acetylated hexosamine, to hydroxylic amino acids (Ser/Thr). As with other GTs, OGT uses the nucleotide sugar UDP-GlcNAc as the donor substrate. This sugar moiety can be removed by the glycosidase activity of O-GlcNAcase (OGA). Thus, O-GlcNAcylation is a cycling post-translational modification that has been implicated in gene expression, metabolism, cellular stress responses, insulin signalling and proteostasis in response to nutrient availability (Levine and Walker, 2016). It has been postulated that a role of GlcNAcylation may be as a sensor of metabolic homeostasis in the cell because the biosynthesis of UDP-GlcNAc requires D-Glc, glutamine, acetyl-CoA and uridine (Ryczko *et al.*, 2016). The active forms of mitochondrial OGT and OGA, as well as a UDP-GlcNAc transporter, have been reported to be present in mammalian mitochondria, and putative targets have been identified (Banerjee *et al.*, 2015; Sacoman *et al.*, 2017). However, OGT and OGA orthologues or not findable in trypanosomatid genomes.

Although the mitochondrial localization of the novel fucosyltransferase activity has been confirmed for the trypanosomatids *T. brucei* (Bandini *et al.*, 2021) and *L. major* (Guo *et al.*, 2021), and suggested for *T. cruzi* (this thesis), there is currently no knowledge on the endogenous substrates for this activity, or their biological role(s). To shed light on this matter, we initially decided to use the parasite *T. cruzi* as our experimental platform to try to radiolabel fucosylated mitochondrial molecules. As previously mentioned (section 5.1), *T. cruzi* has *AFKP80* and *FKP40* genes, and so is probably able to uptake Fuc from the medium and salvage it to GDP-Fuc (Fig. 5.1). In order to confirm this hypothesis, a [^{3}H]Fuc labelling of *T. cruzi* epimastigotes followed by SDS-PAGE of the cell lysates and fluorography of the gel could have been done. Importantly, there is a single characterized fucose-containing glycoconjugate in trypanosomatids, *T. cruzi* gp72 (Nozaki and Cross, 1994; Haynes *et al.*, 1996; Allen *et al.*, 2013). Using it as a positive control, we could determine the efficiency of [^{3}H]Fuc-labelling of gp72 by SDS-PAGE and fluorography. Due to the low abundance of this glycoprotein, we could additionally perform an IP using the mouse monoclonal antibody WIC29.26, which specifically targets the glycan moiety (Ferguson *et al.*, 1983), to analyse if the [^{3}H]-labelling was successful.

T. cruzi also has a functional *de novo* pathway for the synthesis of GDP-Fuc (Fig. 5.1). Therefore, *T. cruzi* GDP-Fuc can be synthesized from GDP-Man by the action of

GMD followed by GMER (Fig. 5.1) (Turnock and Ferguson, 2007). This GDP-Man is generated from either D-Glc or D-Man uptaken from the media. Thus, an alternative way to label the fucosylated glycoconjugates of this parasite would be by addition of tritiated mannose ([^{3}H]Man) to an epimastigote culture. D-Man has been found only in N-glycans and GPI-anchor structures in trypanosomatids, therefore if the labelling is done in the presence of tunicamycin, which inhibits the first step of the N-glycans synthesis in the ER (Heifetz *et al.*, 1979), only cell surface-located GPI-anchored, and any other fucose-containing glycoconjugates should become radiolabelled. Since the former do not occur in mitochondria, any [^{3}H]-labelled glycoproteins or glycolipids found in enriched mitochondrial subcellular fractions would be good candidates for endogenous *T. cruzi* mitochondrial fucosylation substrates, which could be further enriched chromatographically following the [^{3}H]-label, for proteomic/mass spectrometric identification.

These strategies were considered initially but, while viable, they have important drawbacks. First, due to the presence of both the *de novo* and salvage pathways. Thus, the specific activity of GDP-[^{3}H]Fuc made by either would be diluted by the other, affecting the efficiency of the labelling. Second, *T. cruzi* is a BSL3 pathogen and less convenient to grow in bulk than other parasites.

Due to this situation, we considered the use of other human pathogenic parasites for our labelling experiments, such as *T. brucei* and *Leishmania* spp. However, the genome of the former predicts that it has no salvage pathway, therefore we could only use [^{3}H]-labelled D-Glc or D-Man in order to induce the intracellular generation of GDP-[^{3}H]Fuc in *T. brucei*, which would make the labelling non-selective for fucose-containing glycoconjugates. In the case of *Leishmania* spp., they are able to uptake Fuc from the medium, but they contain both nucleotide sugar synthesis pathways, which again would dilute the specific activity of intracellular GDP-[^{3}H]Fuc.

Very evolutionarily close to the *Leishmania* spp. is the non-pathogenic trypanosomatid *C. fasciculata*, both belonging to the same subfamily Leishmaniinae, a monophyletic group of monoxenous insect parasites (genera *Crithidia*, *Leptomonas*, and *Wallaceina*) and the dixenous parasites of insects and vertebrates included in the genera *Leishmania* and *Endotrypanum*, characterized by their highly conserved sequences of the small subunit ribosomal ribonucleic acid (SSU rRNA) and gGAPDH genes (Jirků *et al.*, 2012). *C. fasciculata* parasite is easy to grow, safer to handle since it only infects insects, and conveniently, uses a salvage pathway that converts D-Ara into GDP-Ara, one of the

sugar donors needed for the synthesis of its major surface glycoconjugate, LAG (Fig. 5.5A, and Fig. 6.7) (Schneider *et al.*, 1996). As discussed previously, the chemical structure of D-Ara and L-Fuc are very similar (Fig. 6.6A), differing only by a methyl group. Indeed, it has been reported that *L. major* can uptake L-Fuc from the medium and incorporate it into its major surface glycoconjugate, LPG, in place of the D-Ara residues (Wilson *et al.*, 1999) via AFKP80 and/or FKP40. There are orthologues of this protein machinery in *C. fasciculata* and so, we hypothesized that this parasite could also uptake L-Fuc in the medium, convert it to GDP-Fuc and transfer it to its glycoconjugates, i.e., LAG and other potential intracellular (mitochondrial) substrates. It has been reported that *C. fasciculata* can uptake and utilise [^{3}H]Ara in this way (Schneider *et al.*, 1995) and, as reported (Mengeling and Turco, 1999) and demonstrated in this work, cytoplasmic extracts of this parasite can convert [^{3}H]Fuc into GDP-[^{3}H]Fuc.

Based on these precedents, we initially tested the utilization of L-Fuc when present in *C. fasciculata* cultures, by detecting its presence in the sugar content of purified LAG from the parasite. After detecting the D-Ara, D-Man and D-Gal composition of octyl-Sepharose-purified LAG by the methanolysis/TMS derivatization method and analysis by GC-MS (Chambers and Clamp, 1971; Fukuda and Kobata, 1993), we purified LAG from *C. fasciculata* cultures supplemented with different concentrations of L-Fuc. Their GC-MS analysis showed the interchange of D-Ara to L-Fuc as one of the sugar components of the glycoconjugate, which confirmed that L-Fuc is uptaken when present in the medium and salvaged into GDP-Fuc for use as a sugar donor.

Although there is no *de novo* pathway to GDP-Fuc in *C. fasciculata* to dilute GDP-[^{3}H]Fuc specific activity made by the salvage pathway, it was not known whether the transport of [^{3}H]Fuc might be competing with the much more abundant Glc and GlcN in the labelling medium, predominant energy sources for trypanosomatids (Barrett *et al.*, 1998). After confirming the survival of *C. fasciculata* in D-Glc- and D-GlcN-depleted medium for at least 24 h, we performed parallel labellings in standard and depleted media, observing a much higher labelling efficiency in the absence of these abundant sugars, suggesting that the uptake of L-Fuc might be competing for uptake with D-Glc and/or D-GlcN through the same transporter(s).

In considering this inhibition of [^{3}H]Fuc biosynthetic labelling in the presence of D-Glc and D-GlcN, we wondered whether there was sufficient literature precedent to support our postulate that L-Fuc might compete with the same sugar transporter(s). No reports are currently published about the particularities of sugar transport in *C.*

fasciculata. However, in the evolutionarily close *Leishmania* spp., D-Glc transport is important in both insect-stage promastigotes and intramacrophage mammalian-host stage amastigotes (Landfear, 2010). Initially identified in *L. mexicana* (Burchmore and Landfear, 1998) and later in other *Leishmania* spp., their genomes encode a cluster of three D-Glc transporters designated *LmGT1*, *LmGT2*, and *LmGT3*. After a null mutant was created by targeted disruption of the whole cluster (Burchmore *et al.*, 2003), the promastigote null mutants were significantly reduced in size compared to wild type promastigotes, and they were significantly more sensitive to oxidative stress, which could in part explain their non-viability as amastigotes, since these have to face the oxidative environment of the phagolysosomes. A later phenotypic characterization of this promastigote null mutant (Rodriguez-Contreras *et al.*, 2007) demonstrated its inability to import the hexoses D-Glc, D-Man, D-fructose and D-Gal, and that complementation with any of the isoforms restores the import of the four hexoses. However, it is unclear from these data whether LmGT1-3 can transport L-Fuc and/or D-Ara.

In the case of *T. brucei*, its genome contains a cluster of hexose transporter genes (Bringaud and Baltz, 1993) which encompasses six copies of the *THT1* gene, most abundantly expressed in the bloodstream form (Barrett *et al.*, 1998), followed by five copies of the *THT2* gene, mostly expressed in procyclics (Barrett *et al.*, 1998), but which are unable to transport D-Gal (Eisenthal *et al.*, 1989; Tetaud *et al.*, 1997). Again, it is unclear (but unlikely) whether THT1 or THT2 could transport L-Fuc and/or D-Ara.

In *T. cruzi*, a single gene present in a cluster of multiple copies has been described, which is similar in sequence to *TbTHT1* (Tetaud *et al.*, 1994), and that, in epimastigotes and trypomastigotes, transports D-Glc with high affinity and other hexoses, such as D-fructose, with lower affinities (Barrett *et al.*, 1999). In the case of amastigotes, much less studied, D-Glc transport is down regulated compared with trypomastigotes (Silber *et al.*, 2009).

In the TriTrypDB annotated genome of *C. fasciculata*, there are three copies of a putative D-Glc transporter in a cluster located in the chromosome 28. BLAST analyses showed that each *L. mexicana GT* sequence is 80% similar to each of these three copies of the *C. fasciculata* putative transporters. Interestingly, also the sequences of *TbTHT1* and *TbTHT2* are 66 and 65% similar, respectively, to one of the three copies in *C. fasciculata*. The trypanosome D-Glc transporters belong to the D-Glc transporter superfamily (Gould and Holman, 1993), and therefore present a certain structural similarity between them and with the reference gene for this family, the mammalian

erythrocyte transporter, GLUT1 (Tetaud *et al.*, 1997). Therefore, the sequence similarity of *T. brucei* THT1 and THT2 and *L. mexicana* GT1-3 to the putative *C. fasciculata* transporters also suggests structural similarity to GLUT1. Indeed, using the transmembrane topology predictor DeepTMHMM (Hallgren *et al.*, 2022), the three *C. fasciculata* proteins are predicted to have 12 transmembrane domains, as in the case of other trypanosomatid D-Glc transporters, or the mammalian GLUT1 (Barrett *et al.*, 1998; Landfear, 2010). The higher similarity of the *C. fasciculata* putative transporters to the *Leishmania* spp. transporters may be indicating that there is a functional similarity between both genera.

In summary, we lack sufficient information from the literature to conclude whether or not the LmGT1-3 like transporters can transport L-Fuc and/or D-Ara. Further, it is worth noting that both bacteria (Bradley *et al.*, 1987; Gunn *et al.*, 1994) and mammalian cells (Wiese *et al.*, 1994; Leck and Wiese, 2004) have entirely separate transport systems for L-Fuc versus D-Glc and related sugars. A BLASTp search using as the query the amino acid sequence of the *E. coli* L-Fuc transporter (the only available) found no significant similarity in the *C. fasciculata* and *Leishmania* spp. genomes. Therefore, while a common D-Glc/D-GlcN and L-Fuc transport mechanism is an attractive hypothesis to explain our biosynthetic labelling data, we should also consider alternative hypotheses. One of these would be that in the presence of a rich D-Glc source from the medium, *C. fasciculata* may be able to produce much more D-Ara from D-Glc, through the oxidative and non-oxidative branches of the pentose phosphate pathway (Ferguson *et al.*, unpublished), and that this then competes with [^{3}H]Fuc at the level of the AFKP80 kinase/pyrophosphorylase step. Further experiments would be necessary to resolve these possibilities, such as direct analysis of whether L-Fuc can compete with D-Glc for uptake, using for example classical 2-deoxy-Glc uptake experiments, and the measurement of GDP-Ara and GDP-Fuc levels by liquid chromatography with tandem mass spectrometry (LC-MS/MS) in *C. fasciculata* extracts under different media conditions (e.g. permutations of high- and low-D-Glc and high- and low-L-Fuc). Regardless of mechanism, the [^{3}H]Fuc biosynthetic labelling experiments clearly established that high biosynthetic labelling efficiency requires a low D-Glc/D-GlcN labelling medium.

The detection of a high-MW, [^{3}H]Fuc-labelled species in our initial experiments was the first clue that another, potentially intracellular (mitochondrial), substrate was being labelled. Such a high-MW, fucose-containing substrate has been previously

identified, but not characterized, in *T. brucei* procyclic form, and it was not a GPI-anchored species (Güther *et al.*, 2009). Therefore, we decided to pre-treat the *C. fasciculata* [^{3}H]-labelled lysates with PI-PLC, which would selectively cleave the GPI anchor and remove the radioactive signal from LAG from SDS-PAGE analysis, and to see whether our high-MW species was also not affected by PI-PLC treatment. Indeed, as shown in (Fig. 5.5B), the radiolabelled LAG signal disappeared after the treatment, whereas the signal at the top of the gel increased. This suggested that the radiolabelled material at the top of the gel (after PI-PLC) was most likely some fraction of PI-PLC-digested LAG plus the original high-MW material. Since this could be implying that the high-MW species was non-lipidated, [^{3}H]Fuc-containing "arabinogalactan" that did not resolve well and remained at the top of the gel, and not a novel intracellular fucosylated glycoconjugate, further enzymatic and chemical assays were performed. The results after performing very mild acid and chemical dephosphorylation treatments previously used in the compositional analysis of *Leishmania* spp. LPG and *C. fasciculata* LAG surface glycoconjugates (McConville *et al.*, 1990; Schneider *et al.*, 1996) demonstrated the different response of [^{3}H]-LAG and the high-MW species to these treatments (Fig. 5.6), leading us to suggest that the high-MW molecule is a putative intracellular fucosylated substrate.

6.4 Gene KO attempts in *C. fasciculata* to halt LAG synthesis.

Due to the strong, polydisperse [^{3}H]-LAG signal, the further characterization of the high-MW, non-GPI anchored, mild acid hydrolysis- and aq. HF-sensitive, potentially intracellular fucosylated species was hindered. Therefore, we initially decided to halt the LAG synthesis by blocking GPI anchor synthesis in *C. fasciculata*.

To achieve this, we decided to sequentially replace the two alleles encoding *CfGPI12* with gene cassettes conferring resistance to hygromycin and neomycin. *GPI12* encodes the enzyme that catalyses the second step in the biosynthesis of GPI anchors. Originally identified in *T. brucei*, it has a N-acetylglucosamine-phosphatidylinositol (GlcNAc-PI) de-*N*-acetylase activity, which converts the GlcNAc residue to GlcN and generates the $GlcNH_2$-PI precursor, which is flipped to the luminal side of the ER before continuing the synthesis (Chang *et al.*, 2002; Borges *et al.*, 2021). Therefore, a *CfGPI12* null mutant would not be able to synthesize LAG (Fig. 6.7), along with any other GPI-anchored membrane surface molecule.

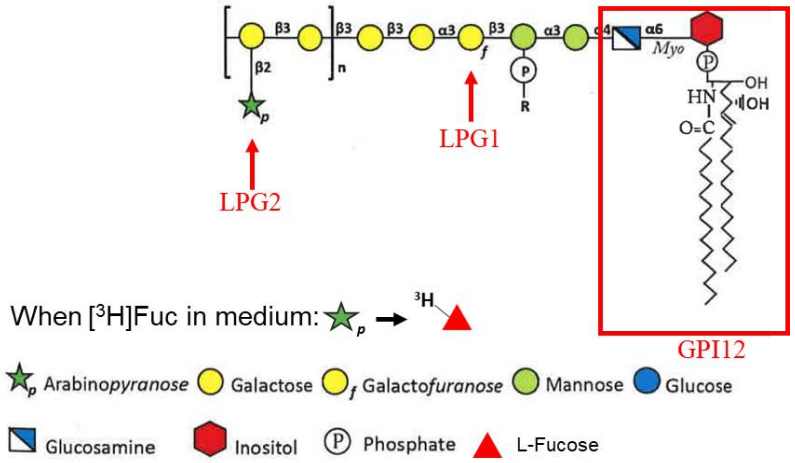


Figure 6.7: Scheme of LAG structure based on (Schneider *et al.*, 1996; Iljazi, 2018) indicating the targeted components in this work during the gene KO attempts. When *C. fasciculata* is grown in D-Glc & D-GlcN-depleted medium supplemented with [^{3}H]Fuc, the tritiated sugar is incorporated into the structure, displacing the arabinopyranose residues from the glycolipid.

To our knowledge, only four successful attempts of gene disruption in *C. fasciculata* have been published (Ray and Hines, 1995; Lukeš *et al.*, 2001; Mittra *et al.*, 2003; Avliyakulov *et al.*, 2004). In all cases, a similar strategy was followed based on the generation of gene replacement constructs where drug resistance cassettes (for hygromycin, neomycin, phleomycin or blasticidin) were flanked by the 5' and 3' UTRs of *L. major* dihydrofolate reductase-thymidylate synthase (DHFR-TS), to assist the expression of the selection markers, and the constructs were targeted to disrupt the gene of interest using sequences either side of a specific restriction site within the gene of interest. The 5' and 3' UTR sequences of *L. major* DHFR-TS sequences were obtained from the pX vectors (LeBowitz *et al.*, 1990; Cruz *et al.*, 1991) in which they were successfully used for mediating the expression of neomycin phosphotransferase (NEOr) or hygromycin phosphotransferase (HYGr). A fifth, unpublished, successful gene disruption, was achieved in this group (Iljazi, 2018). In this case, the drug resistance gene sequences (*HYGr* and *NEOr*) were cloned between the 500-bp long sequences of the 5' and 3' UTRs directly adjacent to the targeted gene (glucosamine-fructose-6-phosphate aminotransferase, GFAT), allowing for sequential allele replacement by homologous recombination and successful deletion of the *CfGFAT* gene during studies focused on the D-Ara sugar metabolism in this parasite.

Based on this more recent approach, we decided to follow the same strategy and generate constructs where the *HYGr* or *NEOr* genes were cloned between the 500-bp long 5' and 3'UTRs directly adjacent to *CfGPI12*. We initiated the process using the *HYGr*-harbouring replacement cassette in our electroporations. When drug-resistant, mixed populations were recovered, we were able to obtain clonal populations grown on solid

media (colonies). The approach of using growth on agar for *C. fasciculata* stems from the report on the first isolates of this parasite when a blood agar slant supplemented with different sugars was used (Noguchi and Tilden, 1926). The use of agar plates with BHI medium and haemin, to which the selection antibiotic(s) are added during the setting phase or later spread on the solidified surface, is now well established for the growth and selection of *C. fasciculata* (Brooker, 1971a; Glassberg *et al.*, 1985; Pasion *et al.*, 1994; Scolaro *et al.*, 2005; DiMaio *et al.*, 2018). Thus, clonal populations/colonies can be picked and expanded in liquid culture like yeast and bacteria. We used this approach to obtain Δ*CfGPI12::HYGr* sKO clones that were confirmed by PCR and Southern blot.

Multiple attempts were henceforth performed to generate a null mutant cell line. Despite obtaining many double drug-resistant colonies and clonal cultures, their PCR analyses always indicated the presence of a remaining *CfGPI12* gene in the targeted locus, implying insertion of the second resistance cassette (NEOr) in the wrong locus. These data suggest that *CfGPI12* may be an essential gene for *C. fasciculata* grown in culture. This would be in contrast to *T. brucei* procyclics (Güther *et al.*, 2006) and *L. major* promastigotes (Almani *et al.*, 2016) and more like *T. brucei* bloodstream forms (Chang *et al.*, 2002) and *T. cruzi* epimastigotes (Cardoso *et al.*, 2013). In the latter study, the authors were also able to obtain a single knockout of *TcGPI12*, but not a null mutant leaving them to conclude that *TcGPI12* may be essential for *T. cruzi* epimastigotes. In the absence of additional data for *T. cruzi* and *C. fasciculata*, the essentiality of their *GPI12* genes must remain a supposition. On the other hand, whereas the evolutionary proximity between *C. fasciculata* and *Leishmania* spp. fed our expectation that *CfGPI12* would not be essential, it is quite possible that it is. For example, a *C. fasciculata* cell surface devoid of LAG and all GPI-anchored proteins and GPI-related free glycolipids beyond GlcNAc-PI could well be highly adsorbent or "sticky". Previously, such a phenomenon has been reported when deleting *TbGPI10*, a mannosyltransferase needed in the 6th step of the GPI-anchor biosynthesis in *T. brucei*. The deletion of this non-essential gene in *T. brucei* procyclics yielded cells with an abnormally adhesive surface. Thus, the null mutants were only recovered and expanded when using flasks treated for nonadherent cultures (Nagamune *et al.*, 2000). The same flasks were used in the successful deletion of *TbGPI12* in *T. brucei* (Güther *et al.*, 2006) and were used the studies reported here. However, it is possible that the flasks successfully used for *T. brucei* procyclics might not be sufficient for supporting GPI-less *C. fasciculata*. In this way, it is possible that, even if dKO colonies were obtained on our solid media cultures, when these were grown back

in liquid medium to generate enough material for gDNA extraction, only the suspension cells with a wrongly inserted NEOr cassette grew out for genotyping by PCR.

Another possibility is that if deletion of *CfGPI12* causes a growth deficit, then those in which the second allele remained intact could outgrow the few null mutants during post-electroporation recovery and dual drug-selection step prior cloning by limiting dilution or agar colony growth. It is important to note that no growth deficit has been observed in previous genetic studies where GPI12 was determined as non-essential (Güther *et al.*, 2006, Almani *et al.*, 2016). Despite this, further repetitions were made, and flasks were carefully scrutinized every day after the electroporation and used for spreading on agar plates as soon as a single alive cell was detected. This did not yield a single PCR-positive dKO clonal population. In parallel, the NEOr repair cassette was analysed by restriction enzymes digestion and sent again for sequencing, but everything indicated that the DNA used for the generation of a dKO had been correctly generated.

Based on this situation, we decided to try to inhibit LAG synthesis by deleting other related genes. After fruitful discussions with this group's long-time collaborator Prof. Dr. Stephen Beverley (Washington University School of Medicine, St. Louis, MO, USA) we chose *CfLPG1* and *CfLPG2* as the targeted genes. As explained previously (see section 5.2), these genes encode a UDP-Gal*f* : αMan β-galactofuranosyltransferase and a Golgi GDP-Man/GDP-Ara/GDP-Fuc nucleotide sugar transporter. The disruption of each one of them would disrupt LAG biosynthesis in *C. fasciculata* (Fig. 6.7). In the case of the *CfLPG1* null mutant, we would predict that only a GPI glycolipid of R-P-6Manα1-3Manα1-4GlcNα1-6myo-inositol1-P-ceramide would remain (R most probably being αGlc, see section 1.3). On the other hand, *CfLPG2* dKO cells would not be able to transport GDP-[^{3}H]Fuc to the Golgi and we would predict that these cells might make a LAG molecule lacking the D-Ara/L-Fuc sidechains. In both cases, a [^{3}H]Fuc-labelled LAG should no longer be synthesized, facilitating the detection and analysis of the high-MW species and other potential fucosylated intracellular substrates.

We therefore generated similar constructs as previously, with 500-bp long homologous regions to the 5' and 3' UTRs of the targeted genes flanking a *HYGr* or *NEOr* gene. The latter was used first to generate a Δ*CfLPG1*::*NEOr* sKO cell line, but different PCR analyses of the obtained clones indicated a similar problem as with *CfGPI12* dKO detection, i.e., a wrong insertion of the *NEOr* gene. After several failed attempts, we decided to vary the length of the UTR regions, based on BLAST analysis of these

sequences, and carry on the experiments focused on *CfLPG1*, since its deletion should result in a definitive halt of LAG synthesis but hopefully leave a GPI glycolipid on the cell surface of sufficient size to keep the cells alive. Using new shorter and longer UTR sequences in the NEOr cassettes, we were able to detect by initial PCR analyses the potential generation of Δ*CfLPG1*::*NEOr* sKO clones. These were genotyped by Southern blot, confirming the successful replacement of the first *CfLPG1* allele with *NEOr* sequence when longer (1 Kb) UTR targeting sequences were used. After this, many electroporations using the HYGr 1 kb *CfLPG1* UTRs cassette were performed, but no cultures were recovered in the presence of neomycin and hygromycin. Thanks to the efforts of Dr. MLSG, who carried on these attempts and changed different aspects of the electroporation process, such as the quantity of cells used per electroporation, the device used and an almost immediate subcloning by limit dilution, double drug-resistant clones were recovered. Follow-up genotyping by PCR and Southern blot confirmed that at least one clone with fully deleted *CfLPG1* and no extra copies of the gene or wrong insertion of the drug resistance markers had been obtained.

Having generated a *C. fasciculata* null mutant in which most probably the synthesis of LAG is stopped at the addition of the single galactofuranose residue of the glycan core (Fig. 6), immediate future experiments will involve the [^{3}H]Fuc labelling of these Δ*CfLPG1*::*NEOr*/Δ*CfLPG1*::*HYGr* dKO cells to confirm that LAG is no longer labelled and that the high-MW species is the only, or not, potential intracellular fucosylation substrate.

6.5 Conclusion.

The main aims of this work were to characterise *T. cruzi* mitochondrial fucosyltransferase and identify kinetoplastid fucosylated intracellular glycoconjugates.

The biochemical characterization of TcFUT1 was achieved, and we now define it as a GDP-Fuc : βGal α1-2 fucosyltransferase. For this, endogenously overexpressed and immunoprecipitated protein was incubated with a panel of synthetic sugar acceptors and the radioactive sugar donor GDP-[^{3}H]Fuc, which we synthesized by exploiting the enzymatic machinery of the insect parasite *C. fasciculata*. Due to the commercial unavailability of this molecule, our well-established procedure may result useful for the glycobiological community in assays involving glycosyltransferases or transporters. Furthermore, our preliminary immunolocalization studies in *T. cruzi*, contrasted with

other similar reports (Peloso *et al.*, 2016; Bustos *et al.*, 2017; Chiurillo *et al.*, 2019), seem to indicate the mitochondrial localization of TcFUT1, at least in the epimastigote form.

The [^{3}H]Fuc biosynthetic labelling experiments in *C. fasciculata* detailed in this work indicated a vast incorporation of the radioactive sugar in the structure of the parasite's main surface glycoconjugate, LAG. Importantly, these experiments also lead us to the identification of an intracellular, fucose-containing, non-GPI anchored, HF and TFA sensitive, high-molecular weight species which has been proposed as the substrate of the mitochondrial fucosylation activity present in all kinetoplastids. In order to facilitate the isolation and characterization of this novel molecule, we have generated a *C. fasciculata* cell line in which the synthesis of LAG is inhibited, which will redirect the label only towards our molecule of interest.

Future experiments will include the complete characterization of TcFUT1 + LNB-OMe [^{3}H]Fuc-containing product, by means of an α1-2 exofucosidase treatment, and further labelling experiments with the generated Δ*CfLPG1*::*NEOr*/Δ*CfLPG1*::*HYGr C. fasciculata* cell line along with different treatments, such as pronase, PNGase F, or RNAse A, to determine the molecular nature of the intracellular fucosylated substrate.

Thus, this work has contributed to understanding the biology of a novel class of kinetoplastid enzymes, and the novel concept of mitochondrial glycosylation.

7. REFERENCES.

Abras, A. *et al.* (2022) 'Worldwide Control and Management of Chagas Disease in a New Era of Globalization: a Close Look at Congenital Trypanosoma cruzi Infection', *Clinical Microbiology Reviews*, 35(2), p. e0015221.

Albesa-Jove, D. and Guerin, M.E. (2016) 'The conformational plasticity of glycosyltransferases', *Curr Opin Struct Biol*, 40, pp. 23–32.

Alcolea, P.J. *et al.* (2014) 'An insight into the proteome of Crithidia fasciculata choanomastigotes as a comparative approach to axenic growth, peanut lectin agglutination and differentiation of Leishmania spp. promastigotes', *PLoS ONE*, 9(12), p. e113837.

Allen, S. *et al.* (2013) 'Structure of a complex phosphoglycan epitope from gp72 of Trypanosoma cruzi', *Journal of Biological Chemistry*, 288(16), pp. 11093–11105.

Almani, P.G.N. *et al.* (2016) 'The role of GlcNAc-PI-de-N-acetylase gene by gene knockout through homologous recombination and its consequences on survival, growth and infectivity of Leishmania major in in vitro and in vivo conditions', *Acta Tropica*, 154, pp. 63–72.

Alonso-Padilla, J. *et al.* (2019) 'Strategies to enhance access to diagnosis and treatment for Chagas disease patients in Latin America', *Expert Review of Anti-Infective Therapy*, 17(3), pp. 145–157.

Altschul, S.F. *et al.* (1997) 'Gapped BLAST and PSI-BLAST: a new generation of protein database search programs', *Nucleic Acids Res*, 25(17), pp. 3389–3402.

Aslett, M. *et al.* (2010) 'TriTrypDB: a functional genomic resource for the Trypanosomatidae', *Nucleic Acids Res*, 38(Database issue), pp. D457-62.

Avliyakulov, N.K., Lukeš, J. and Ray, D.S. (2004) 'Mitochondrial histone-like DNA-binding proteins are essential for normal cell growth and mitochondrial function in Crithidia fasciculata', *Eukaryotic Cell*, 3(2), pp. 518–526.

Baek, M. *et al.* (2021) 'Accurate prediction of protein structures and interactions using a three-track neural network', *Science (New York, N.Y.)*, 373(6557), pp. 871–876.

Bandini, G. (2011) *Studies on fucosylation of Trypanosoma brucei*. University of Dundee.

Bandini, G. *et al.* (2019) 'Protein O - And C -Glycosylation pathways in Toxoplasma gondii and Plasmodium falciparum', *Parasitology*, pp. 1–12.

Bandini, G. *et al.* (2021) 'An essential, kinetoplastid-specific GDP-Fuc: β-D-Gal α-1,2-fucosyltransferase is located in the mitochondrion of Trypanosoma brucei', *eLife*. Edited by M.J. McConville, D. Soldati-Favre, and Á. Acosta Serrano, 10, p. e70272.

Bandini, G. and Ferguson, M.A.J. (2022) 'Roles of Glycans in Protozoal Host-Parasite Interactions', in *Reference Module in Life Sciences*. Elsevier.

Banerjee, P.S., Ma, J. and Hart, G.W. (2015) 'Diabetes-associated dysregulation of O-GlcNAcylation in rat cardiac mitochondria', *Proc Natl Acad Sci U S A*, 112(19), pp. 6050–6055.

Barrett, M.P. *et al.* (1998) 'Trypanosome glucose transporters', *Molecular and Biochemical Parasitology*, 91(1), pp. 195–205.

Barrett, M.P., Walmsleyt, A.R. and Gould, G.W. (1999) 'Structure and function of facultative sugar transporters', *Current Opinion in Cell Biology*, 11(4), pp. 496–502.

Bojarova, P. and Kren, V. (2009) 'Glycosidases: a key to tailored carbohydrates', *Trends Biotechnol*, 27(4), pp. 199–209.

Bond, M.R. and Hanover, J.A. (2015) 'A little sugar goes a long way: the cell biology of O-GlcNAc', *J Cell Biol*, 208(7), pp. 869–880.

Bonney, K.M. and Engman, D.M. (2015) 'Autoimmune pathogenesis of Chagas heart disease: looking back, looking ahead', *The American Journal of Pathology*, 185(6), pp. 1537–1547.

Borges, A.R. *et al.* (2021) 'The Glycosylphosphatidylinositol Anchor: A Linchpin for Cell Surface Versatility of Trypanosomatids', *Frontiers in Cell and Developmental Biology*, 9, p. 720536.

Borges, A.R., Engstler, M. and Wolf, M. (2021) '18S rRNA gene sequence-structure phylogeny of the Trypanosomatida (Kinetoplastea, Euglenozoa) with special reference to Trypanosoma', *European Journal of Protistology*, 81, p. 125824.

Bradley, S.A. *et al.* (1987) ‘Proton-linked L-fucose transport in Escherichia coli’, *The Biochemical Journal*, 248(2), pp. 495–500.

Breton, C., Fournel-Gigleux, S. and Palcic, M.M. (2012) ‘Recent structures, evolution and mechanisms of glycosyltransferases’, *Current Opinion in Structural Biology*, 22(5), pp. 540–549.

Breton, C., Oriol, R. and Imberty, A. (1998) ‘Conserved structural features in eukaryotic and prokaryotic fucosyltransferases’, *Glycobiology*, 8(1), pp. 87–94.

Bringaud, F. and Baltz, T. (1993) ‘Differential regulation of two distinct families of glucose transporter genes in Trypanosoma brucei’, *Molecular and Cellular Biology*, 13(2), pp. 1146–1154.

Brooker, B.E. (1970) ‘Desmosomes and hemidesmosomes in the flagellate Crithidia fasciculata’, *Zeitschrift für Zellforschung und Mikroskopische Anatomie*, 105(2), pp. 155–166.

Brooker, B.E. (1971a) ‘Flagellar adhesion of Crithidia fasciculata to Millipore filters’, *Protoplasma*, 72(1), pp. 19–25.

Brooker, B.E. (1971b) ‘The fine structure of Crithidia fasciculata with special reference to the organelles involved in the ingestion and digestion of protein’, *Zeitschrift Fur Zellforschung Und Mikroskopische Anatomie (Vienna, Austria: 1948)*, 116(4), pp. 532–563.

Brooker, B.E. (1972) ‘Modifications in the arrangement of the pellicular microtubules of Crithidia fasciculata in the gut of Anopheles gambiae’, *Zeitschrift für Parasitenkunde*, 40(4), pp. 271–280.

Brooker, B.E. (1976) ‘The cell coat of Crithidia fasciculata’, *Parasitology*, 72(3), pp. 259–267.

Brun, R. (1974) ‘[Ultrastructure and life cycle of Herpetomonas muscarum, “Herpetomonas mirabilis” and Crithidia luciliae in Chrysomyia chloropyga] (article in German)’, *Acta Tropica*, 31(3), pp. 219–290.

Brun, R. and Schönenberger (1979) ‘Cultivation and in vitro cloning or procyclic culture forms of Trypanosoma brucei in a semi-defined medium. Short communication’, *Acta tropica*, 36(3), pp. 289–292.

Brzezinski, K. *et al.* (2007) ‘High-resolution structure of NodZ fucosyltransferase involved in the biosynthesis of the nodulation factor’, *Acta Biochim Pol*, 54(3), pp. 537–549.

Burchmore, R.J.S. *et al.* (2003) ‘Genetic characterization of glucose transporter function in Leishmania mexicana’, *Proceedings of the National Academy of Sciences*, 100(7), pp. 3901–3906.

Burchmore, R.J.S. and Landfear, S.M. (1998) ‘Differential Regulation of Multiple Glucose Transporter Genes in Leishmania mexicana’, *Journal of Biological Chemistry*, 273(44), pp. 29118–29126.

Burki, F. *et al.* (2020) ‘The New Tree of Eukaryotes’, *Trends in Ecology and Evolution*, 35(1), pp. 43–55.

Burnham-Marusich, A.R. and Berninsone, P.M. (2012) ‘Multiple proteins with essential mitochondrial functions have glycosylated isoforms’, *Mitochondrion*, 12(4), pp. 423–427.

Bustos, P.L. *et al.* (2017) ‘A homolog of cyclophilin D is expressed in Trypanosoma cruzi and is involved in the oxidative stress–damage response’, *Cell Death Discovery*, 3(1), pp. 1–7.

Cardoso, M.S. *et al.* (2013) ‘Identification and Functional Analysis of Trypanosoma cruzi Genes That Encode Proteins of the Glycosylphosphatidylinositol Biosynthetic Pathway’, *PLoS Neglected Tropical Diseases*, 7(8), p. e2369.

Carter, Y.L. *et al.* (2012) ‘Acute Chagas disease in a returning traveler’, *The American Journal of Tropical Medicine and Hygiene*, 87(6), pp. 1038–1040.

Cerqueira, N. *et al.* (2012) ‘Glycosidases – A Mechanistic Overview’, in *Carbohydrates - Comprehensive Studies on Glycobiology and Glycotechnology*. IntechOpen.

Chagas disease, WHO (2022). Available at: https://www.who.int/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis) (Accessed: 25 October 2022).

Chambers, R.E. and Clamp, J.R. (1971) ‘An assessment of methanolysis and other factors used in the analysis of carbohydrate-containing materials’, *The Biochemical Journal*, 125(4), pp. 1009–1018.

Chang, T. *et al.* (2002) ‘Cloning of Trypanosoma brucei and Leishmania major genes encoding the GlcNAc-phosphatidylinositol de-N-acetylase of glycosylphosphatidylinositol biosynthesis that is essential to the African sleeping sickness parasite’, *The Journal of Biological Chemistry*, 277(51), pp. 50176–50182.

Chauché, C. *et al.* (2020) ‘A Truncated Form of HpARI Stabilizes IL-33, Amplifying Responses to the Cytokine’, *Frontiers in Immunology*, 11, 1363.

Chiurillo, M.A. *et al.* (2019) ‘Functional analysis and importance for host cell infection of the Ca2+-conducting subunits of the mitochondrial calcium uniporter of Trypanosoma cruzi’, *Molecular Biology of the Cell*, 30(14), pp. 1676–1690.

Clark, T.B. *et al.* (1964) ‘The Transmission of Crithidia fasciculata Leger 1902 in Culiseta incidens (Thomson)’, *The Journal of Protozoology*, 11(3), pp. 400–402.

Claros, M.G. and Vincens, P. (1996) ‘Computational Method to Predict Mitochondrially Imported Proteins and their Targeting Sequences’, *European Journal of Biochemistry*, 241(3), pp. 779–786.

Comini, M. *et al.* (2005) ‘Trypanothione Synthesis in Crithidia Revisited’, *Journal of Biological Chemistry*, 280(8), pp. 6850–6860.

Cooling, L. (2015) ‘Blood Groups in Infection and Host Susceptibility’, *Clinical Microbiology Reviews*, 28(3), pp. 801–870.

Costa, F.C. *et al.* (2018) ‘Expanding the toolbox for Trypanosoma cruzi: A parasite line incorporating a bioluminescence-fluorescence dual reporter and streamlined CRISPR/Cas9 functionality for rapid in vivo localisation and phenotyping’, *PLoS Neglected Tropical Diseases*, 12(4), p. e0006388.

Costa, S. *et al.* (2014) ‘Fusion tags for protein solubility, purification, and immunogenicity in Escherichia coli: The novel Fh8 system’, *Frontiers in Microbiology*, 5(FEB), p. 63.

Coutinho, P.M. *et al.* (2003) 'An evolving hierarchical family classification for glycosyltransferases', *J Mol Biol*, 328(2), pp. 307–317.

Crespillo-Andújar, C. *et al.* (2018) 'Severe immune thrombocytopaenia in a patient taking benznidazole for chronic Chagas disease', *BMJ case reports*, 2018, p. bcr-2017-223788.

Cruz, A., Coburn, C.M. and Beverley, S.M. (1991) 'Double targeted gene replacement for creating null mutants', *Proceedings of the National Academy of Sciences of the United States of America*, 88(16), pp. 7170–7174.

De Bona, E. *et al.* (2018) 'Autoimmunity in Chronic Chagas Disease: A Road of Multiple Pathways to Cardiomyopathy?', *Frontiers in Immunology*, 9, p. 1842.

Denecke, S. *et al.* (2022) 'Adhesion of Crithidia fasciculata promotes a rapid change in developmental fate driven by cAMP signaling'. bioRxiv, p. 2022.10.06.511084.

van Der Wel, H. *et al.* (2001) 'A non-Golgi alpha 1,2-fucosyltransferase that modifies Skp1 in the cytoplasm of Dictyostelium', *The Journal of Biological Chemistry*, 276(36), pp. 33952–33963.

Descoteaux, A. *et al.* (1995) 'A Specialized Pathway Affecting Virulence Glycoconjugates of Leishmania', *Science*, 269(5232), pp. 1869–1872.

Descoteaux, A. and Turco, S.J. (1999) 'Glycoconjugates in Leishmania infectivity', *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1455(2), pp. 341–352.

DiMaio, J. *et al.* (2018) 'The single mitochondrion of the kinetoplastid parasite Crithidia fasciculata is a dynamic network', *PLOS ONE*, 13(12), p. e0202711.

Doering, T.L. *et al.* (1989) 'Biosynthesis of the glycosyl phosphatidylinositol membrane anchor of the trypanosome variant surface glycoprotein. Origin of the non-acetylated glucosamine', *The Journal of Biological Chemistry*, 264(19), pp. 11168–11173.

Drula, E. *et al.* (2022) 'The carbohydrate-active enzyme database: functions and literature', *Nucleic Acids Research*, 50(D1), pp. D571–D577.

Duncan, S.M. and Ferguson, M.A.J. (2022) ‘Common and unique features of glycosylation and glycosyltransferases in African trypanosomes’, *Biochemical Journal*, 479(17), pp. 1743–1758.

Dvorak, J.A. and Hyde, T.P. (1973) ‘Trypanosoma cruzi: interaction with vertebrate cells in vitro. 1. Individual interactions at the cellular and subcellular levels’, *Experimental Parasitology*, 34(2), pp. 268–283.

Eisenthal, R., Game, S. and Holman, G.D. (1989) ‘Specificity and kinetics of hexose transport in Trypanosoma brucei’, *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 985(1), pp. 81–89.

Elferink, H. *et al.* (2020) ‘A comprehensive overview of substrate specificity of glycoside hydrolases and transporters in the small intestine’, *Cellular and Molecular Life Sciences*, 77(23), pp. 4799–4826.

Eme, L. *et al.* (2014) ‘On the Age of Eukaryotes: Evaluating Evidence from Fossils and Molecular Clocks’, *Cold Spring Harbor Perspectives in Biology*, 6(8), p. a016139.

Engel, L. *et al.* (2021) ‘Utility of Bioluminescent Homogeneous Nucleotide Detection Assays in Measuring Activities of Nucleotide-Sugar Dependent Glycosyltransferases and Studying Their Inhibitors’, *Molecules*, 26(20), p. 6230.

Epting, C.L., Coates, B.M. and Engman, D.M. (2010) ‘Molecular mechanisms of host cell invasion by Trypanosoma cruzi’, *Experimental Parasitology*, 126(3), pp. 283–291.

Esteves, M.J.G. *et al.* (1987) ‘Cell Surface Carbohydrate Differences in Wild and Mutant Strains of Crithidia fasciculata’, *The Journal of Protozoology*, 34(2), pp. 226–230.

European Centre for Disease Prevention and Control (2018) *New map shows the presence of Anopheles maculipennis s.l. mosquitoes in Europe*, *European Centre for Disease Prevention and Control*. Available at: https://www.ecdc.europa.eu/en/news-events/new-map-shows-presence-anopheles-maculipennis-sl-mosquitoes-europe (Accessed: 29 October 2022).

Faktorová, D. *et al.* (2016) ‘From simple to supercomplex: mitochondrial genomes of euglenozoan protists’, *F1000Research*, 5, p. F1000 Faculty Rev-392.

Feingold, D.S. and Barber, G.A. (1990) ‘Nucleotide Sugars’, in P.M. Dey (ed.) *Methods in Plant Biochemistry*. Academic Press (Carbohydrates), pp. 39–78.

Ferguson, M.A., Allen, A.K. and Snary, D. (1983) ‘Studies on the structure of a phosphoglycoprotein from the parasitic protozoan Trypanosoma cruzi.’, *Biochemical Journal*, 213(2), pp. 313–319.

Ferguson, M.A.J. (1997) ‘The surface glycoconjugates of trypanosomatid parasites’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 352(1359), pp. 1295–1302.

Fermino, B.R. *et al.* (2019) ‘Shared species of crocodilian trypanosomes carried by tabanid flies in Africa and South America, including the description of a new species from caimans, Trypanosoma kaiowa n. sp.’, *Parasites & Vectors*, 12(1), p. 225.

Fernandes, M.C. and Andrews, N.W. (2012) ‘Host cell invasion by Trypanosoma cruzi: a unique strategy that promotes persistence’, *FEMS microbiology reviews*, 36(3).

Ferri, G. and Edreira, M.M. (2021) ‘All Roads Lead to Cytosol: Trypanosoma cruzi Multi-Strategic Approach to Invasion’, *Frontiers in Cellular and Infection Microbiology*, 11.

Filosa, J.N. *et al.* (2019) ‘Dramatic changes in gene expression in different forms of Crithidia fasciculata reveal potential mechanisms for insect-specific adhesion in kinetoplastid parasites’, *PLOS Neglected Tropical Diseases*, 13(7), p. e0007570.

Flegontov, P. *et al.* (2013) ‘Paratrypanosoma is a novel early-branching trypanosomatid’, *Current biology: CB*, 23(18), pp. 1787–1793.

Francisco, A.F. *et al.* (2020) ‘Challenges in Chagas Disease Drug Development’, *Molecules (Basel, Switzerland)*, 25(12), p. E2799.

Fukasawa, Y. *et al.* (2015) ‘MitoFates: Improved Prediction of Mitochondrial Targeting Sequences and Their Cleavage Sites’, *Molecular & Cellular Proteomics : MCP*, 14(4), pp. 1113–1126.

Fukuda, M. and Kobata, A. (1993) *Glycobiology: A Practical Approach*. IRL Press at Oxford University Press (Practical approach series).

Gadelha, C. *et al.* (2004) 'Relationships between the major kinetoplastid paraflagellar rod proteins: a consolidating nomenclature', *Molecular and Biochemical Parasitology*, 136(1), pp. 113–115.

Gibson, D.G. *et al.* (2009) 'Enzymatic assembly of DNA molecules up to several hundred kilobases', *Nature Methods*, 6(5), pp. 343–345.

Gibson, D.G. *et al.* (2010) 'Chemical synthesis of the mouse mitochondrial genome', *Nature Methods*, 7(11), pp. 901–903.

Gibson, D.G. (2011) 'Enzymatic assembly of overlapping DNA fragments', *Methods Enzymol*, 498, pp. 349–361.

Gibson, W.C. and Miles, M.A. (1986) 'The karyotype and ploidy of Trypanosoma cruzi', *The EMBO journal*, 5(6), pp. 1299–1305.

Glassberg, J., Miyazaki, L. and Rifkin, M.R. (1985) 'Isolation and partial characterization of mutants of the trypanosomatid Crithidia fasciculata and their use in detecting genetic recombination', *The Journal of Protozoology*, 32(1), pp. 118–125.

Gloster, T.M. (2014) 'Advances in understanding glycosyltransferases from a structural perspective', *Current Opinion in Structural Biology*, 28, pp. 131–141.

Gonzalez-Rellan, M.J. *et al.* (2022) 'O-GlcNAcylation: A Sweet Hub in the Regulation of Glucose Metabolism in Health and Disease', *Frontiers in Endocrinology*, 13, p. 873513.

Gould, G.W. and Holman, G.D. (1993) 'The glucose transporter family: structure, function and tissue-specific expression.', *Biochemical Journal*, 295(Pt 2), pp. 329–341.

Gray, E.B. *et al.* (2018) 'Reactivation of Chagas disease among heart transplant recipients in the United States, 2012-2016', *Transplant Infectious Disease: An Official Journal of the Transplantation Society*, 20(6), p. e12996.

Greis, K.D. *et al.* (1992) 'Purification and characterization of an extracellular phosphoglycan from Leishmania donovani', *The Journal of Biological Chemistry*, 267(9), pp. 5876–5881.

Grewal, R.K. *et al.* (2021) 'Structural Insights in Mammalian Sialyltransferases and Fucosyltransferases: We Have Come a Long Way, but It Is Still a Long Way Down', *Molecules (Basel, Switzerland)*, 26(17), p. 5203.

Gunn, F.J., Tate, C.G. and Henderson, P.J.F. (1994) 'Identification of a novel sugar-H+ symport protein, FucP, for transport of L-fucose into Escherichia coli', *Molecular Microbiology*, 12(5), pp. 799–809.

Guo, Hongjie *et al.* (2017) 'Genetic metabolic complementation establishes a requirement for GDP-fucose in Leishmania', *Journal of Biological Chemistry*, 292(25), pp. 10696–10708.

Guo, H *et al.* (2017) 'Genetic metabolic complementation establishes a requirement for GDP-fucose in Leishmania', *J Biol Chem*, 292(25), pp. 10696–10708.

Guo, H., Damerow, S., Penha, L., Menzies, S., Polanco, G., Zegzouti, H., Ferguson, Michael A. J., *et al.* (2021) 'A broadly active fucosyltransferase LmjFUT1 whose mitochondrial localization and activity are essential in parasitic Leishmania', *Proceedings of the National Academy of Sciences of the United States of America*, 118(33), p. e2108963118.

Guo, H., Damerow, S., Penha, L., Menzies, S., Polanco, G., Zegzouti, H., Ferguson, Michael A J, *et al.* (2021) 'A broadly active fucosyltransferase LmjFUT1 whose mitochondrial localization and catalytic activity is essential in parasitic Leishmania major', *bioRxiv*, p. 2021.05.10.443387.

Gurevitz, J.M. *et al.* (2013) 'Intensified Surveillance and Insecticide-based Control of the Chagas Disease Vector Triatoma infestans in the Argentinean Chaco', *PLOS Neglected Tropical Diseases*, 7(4), p. e2158.

Güther, M.L.S. *et al.* (2006) 'GPI-anchored proteins and free GPI glycolipids of procyclic form Trypanosoma brucei are nonessential for growth, are required for colonization of the tsetse fly, and are not the only components of the surface coat', *Molecular Biology of the Cell*, 17(12), p. 5265-75.

Güther, M.L.S. *et al.* (2009) 'Fate of glycosylphosphatidylinositol (GPI)-less procyclin and characterization of sialylated non-GPI-anchored surface coat molecules of procyclic-form trypanosoma brucei', *Eukaryotic Cell*, 8(9), p 1407-17.

Guttman, H.N. (1963) ‘Experimental glimpses at the lower trypanosomatidae’, *Experimental Parasitology*, 14(1), pp. 129–142.

Hallgren, J. *et al.* (2022) ‘DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks’. bioRxiv, p. 2022.04.08.487609.

Hamilton, P.B. and Stevens, J.R. (2017) ‘Classification and phylogeny of Trypanosoma cruzi’, in J. Telleria and M. Tibayrenc (eds) *American Trypanosomiasis Chagas Disease (Second Edition)*. London: Elsevier, pp. 321–344.

Hamilton, P.B., Teixeira, M.M.G. and Stevens, J.R. (2012) ‘The evolution of Trypanosoma cruzi: the “bat seeding” hypothesis’, *Trends in Parasitology*, 28(4), pp. 136–141.

Hammarström, M. *et al.* (2006) ‘Effect of N-terminal solubility enhancing fusion proteins on yield of purified target protein’, *Journal of Structural and Functional Genomics*, 7, pp. 1–14.

Harmer, J. *et al.* (2018) ‘Farming, slaving and enslavement: histories of endosymbioses during kinetoplastid evolution’, *Parasitology*, 145(10), pp. 1311–1323.

Haynes, P.A., Ferguson, A.J. and Cross, G.A.M. (1996) ‘Structural characterization of novel oligosaccharides of cell-surface glycoproteins of Trypanosoma cruzi’, *Glycobiology*, 6(8), pp. 869–878.

Heifetz, A., Keenan, R.W. and Elbein, A.D. (1979) ‘Mechanism of action of tunicamycin on the UDP-GlcNAc:dolichyl-phosphate Glc-NAc-1-phosphate transferase’, *Biochemistry*, 18(11), pp. 2186–2192.

Hernandez, S. *et al.* (2016) ‘Autochthonous Transmission of Trypanosoma Cruzi in Southern California’, *Open Forum Infectious Diseases*, 3(4), p. ofw227.

Herreros-Cabello, A. *et al.* (2020) ‘Trypanosoma cruzi genome: Organization, multi-gene families, transcription, and biological implications’, *Genes* [Preprint].

Hill, G.C., Brown, C.A. and Clark, M.V. (1968) ‘Structure and Function of Mitochondria in Crithidia fasciculata’, *The Journal of Protozoology*, 15(1), pp. 102–109.

van den Hoff, M.J., Moorman, A.F. and Lamers, W.H. (1992) 'Electroporation in "intracellular" buffer increases cell survival', *Nucleic Acids Research*, 20(11), p. 2902.

Holdener, B.C. and Haltiwanger, R.S. (2019) 'Protein O-fucosylation: structure and function', *Current Opinion in Structural Biology*, 56, pp. 78–86.

Hunter, K.J., Le Quesne, S.A. and Fairlamb, A.H. (1994) 'Identification and biosynthesis of N1,N9-bis(glutathionyl)aminopropylcadaverine (homotrypanothione) in Trypanosoma cruzi', *European Journal of Biochemistry*, 226(3), pp. 1019–1027.

Ihara, H., Ikeda, Y. and Taniguchi, N. (2006) 'Reaction mechanism and substrate specificity for nucleotide sugar of mammalian α1,6-fucosyltransferase - A large-scale preparation and characterization of recombinant human FUT8', *Glycobiology*, 16(4), pp. 333–342.

Iljazi, E. (2018) *Defining D-Arabinose Metabolism in Leishmania major and Crithidia fasciculata*. University of Dundee.

Irwin, N.A.T. *et al.* (2022) 'Systematic evaluation of horizontal gene transfer between eukaryotes and viruses', *Nature Microbiology*, 7(2), pp. 327–336.

Jabari, S. *et al.* (2014) 'Chagasic megacolon: enteric neurons and related structures', *Histochemistry and Cell Biology*, 142(3), pp. 235–244.

Jirků, M. *et al.* (2012) 'New species of insect trypanosomatids from Costa Rica and the proposal for a new subfamily within the Trypanosomatidae', *The Journal of Eukaryotic Microbiology*, 59(6), pp. 537–547.

Joseph, R.E. and Andreotti, A.H. (2008) 'Bacterial expression and purification of Interleukin-2 Tyrosine kinase: Single step separation of the chaperonin impurity', *Protein Expression and Purification*, 60(2), pp. 194–197.

Kaufer, A. *et al.* (2017) 'The evolution of trypanosomatid taxonomy', *Parasites and Vectors*, 10(1), p. 287.

Kaufer, A., Stark, D. and Ellis, J. (2019) 'Evolutionary Insight into the Trypanosomatidae Using Alignment-Free Phylogenomics of the Kinetoplast', *Pathogens*, 8(3), p. 157.

Kaufer, A., Stark, D. and Ellis, J. (2020) ‘A review of the systematics, species identification and diagnostics of the Trypanosomatidae using the maxicircle kinetoplast DNA: from past to present’, *International Journal for Parasitology*, 50(6), pp. 449–460.

Keeling, P. and Leander, B.S. (2009) *Eukaryotes, The Tree of Life Project.* Available at: http://tolweb.org/Eukaryotes/3 (Accessed: 21 October 2022).

Kelly, J.M. *et al.* (1992) ‘A shuttle vector which facilitates the expression of transfected genes in Trypanosoma cruzi and Leishmania.’, *Nucleic Acids Research*, 20(15), pp. 3963–3969.

Kelly, R.J. *et al.* (1995) ‘Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype’, *The Journal of Biological Chemistry*, 270(9), pp. 4640–4649.

Kollien, A.H. and Schaub, G.A. (2000) ‘The development of Trypanosoma cruzi in triatominae’, *Parasitology Today (Personal Ed.)*, 16(9), pp. 381–387.

Kölliker, R. *et al.* (2018) ‘Chagas Cardiomyopathy: Role of Sustained Host-Parasite Interaction in Systemic Inflammatory Burden’, Chagas Disease - Basic Investigations and Challenges.

Korn, E.D., Von Brand, T. and Tobie, E.J. (1969) ‘The sterols of Trypanosoma cruzi and Crithidia fasciculata’, *Comparative Biochemistry and Physiology*, 30(4), pp. 601–610.

Kornberg, A. (1957) ‘Pyrophosphorylases and phosphorylases in biosynthetic reactions’, *Advances in Enzymology and Related Subjects of Biochemistry*, 18, pp. 191–240.

Kornberg, A. (1962) ‘On the metabolic significance of phosphorolytic and pyrophosphorolytic reactions’, *M Kasha, and B Pullman, Editors Horizons in Biochemistry Albert Szent-Gyorgyi Dedicatory Volume*. Available at: https://eurekamag.com/research/025/164/025164117.php (Accessed: 11 October 2022).

Kung, L.A. *et al.* (2009) ‘Global analysis of the glycoproteome in Saccharomyces cerevisiae reveals new roles for protein glycosylation in eukaryotes’, *Mol Syst Biol*, 5, p. 308.

Kurup, S.P. and Tarleton, R.L. (2014) ‘The Trypanosoma cruzi flagellum is discarded via asymmetric cell division following invasion and provides early targets for protective $CD8^+$ T cells’, *Cell Host & Microbe*, 16(4), pp. 439–449.

Lairson, L.L. *et al.* (2008) ‘Glycosyltransferases: structures, functions, and mechanisms’, *Annu Rev Biochem*, 77, pp. 521–555.

Landfear, S.M. (2010) ‘Glucose Transporters in Parasitic Protozoa’, *Methods in molecular biology (Clifton, N.J.)*, 637, pp. 245–262.

LeBowitz, J.H. *et al.* (1990) ‘Development of a stable Leishmania expression vector and application to the study of parasite surface antigen genes.’, *Proceedings of the National Academy of Sciences*, 87(24), pp. 9736–9740.

Leck, J.R. and Wiese, T.J. (2004) ‘Purification and characterization of the L-fucose transporter’, *Protein Expression and Purification*, 37(2), pp. 288–293.

Lee, B.Y. *et al.* (2013) ‘Global economic burden of Chagas disease: a computational simulation model’, *The Lancet. Infectious Diseases*, 13(4), pp. 342–348.

Léger, M.L. (1902) ‘Sur un flagellé parasite de l’Anopheles maculipennis’, *Comptes rendus des séances de la Société de biologie et de ses filiales.*, 54, pp. 354–356.

Lemos, M. *et al.* (2015) ‘Phylogenetic and morphological characterization of trypanosomes from Brazilian armoured catfishes and leeches reveal high species diversity, mixed infections and a new fish trypanosome species’, *Parasites & Vectors*, 8(1), p. 573.

Levine, Z.G. and Walker, S. (2016) ‘The Biochemistry of O-GlcNAc Transferase: Which Functions Make It Essential in Mammalian Cells?’, *Annu Rev Biochem*, 85, pp. 631–657.

Li, J. *et al.* (2018) ‘Unmasking Fucosylation: from Cell Adhesion to Immune System Regulation and Diseases’, *Cell Chem Biol*, 25(5), pp. 499–512.

Li, M. *et al.* (2008) ‘Characterization of a novel alpha1,2-fucosyltransferase of Escherichia coli O128:b12 and functional investigation of its common motif’, *Biochemistry*, 47(1), pp. 378–387.

Lidani, K.C.F. *et al.* (2019) 'Chagas Disease: From Discovery to a Worldwide Health Problem', *Frontiers in Public Health*, 7, p. 166.

Liedtke, S. *et al.* (2001) 'Characterization of N-glycans from mouse brain neural cell adhesion molecule', *Glycobiology*, 11(5), pp. 373–384.

Liu, B. *et al.* (2005) 'Fellowship of the rings: the replication of kinetoplast DNA', *Trends in Parasitology*, 21(8), pp. 363–369.

Liu, X. and Chiu, J.C. (2022) 'Nutrient-sensitive protein O-GlcNAcylation shapes daily biological rhythms', *Open Biology*, 12(9), p. 220215.

Lorenzi, H.A., Vazquez, M.P. and Levin, M.J. (2003) 'Integration of expression vectors into the ribosomal locus of Trypanosoma cruzi', *Gene*, 310, pp. 91–99.

Lukeš, J. *et al.* (2001) 'Disruption of the Crithidia fasciculata KAP1 gene results in structural rearrangement of the kinetoplast disc', *Molecular and Biochemical Parasitology*, 117(2), pp. 179–186.

Lukeš, J. *et al.* (2014) 'Evolution of parasitism in kinetoplastid flagellates', *Molecular and Biochemical Parasitology*, 195(2), pp. 115–122.

Ma, B., Simala-Grant, J.L. and Taylor, D.E. (2006) 'Fucosylation in prokaryotes and eukaryotes', *Glycobiology*, 16(12), pp. 158R-184R.

Ma, J., Hou, C. and Wu, C. (2022) 'Demystifying the O-GlcNAc Code: A Systems View', *Chemical Reviews*, 122(20), pp. 15822–15864.

Magri, A., Galuppi, R. and Fioravanti, M. (2021) 'Autochthonous Trypanosoma spp. in European Mammals: A Brief Journey amongst the Neglected Trypanosomes', *Pathogens*, 10(3), p. 334.

Makarova, K.S. *et al.* (2005) 'Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell', *Nucleic Acids Research*, 33(14), pp. 4626–4638.

Marcus, R. *et al.* (2021) 'Recognition and screening for Chagas disease in the USA', *Therapeutic Advances in Infectious Disease*, 8, p. 20499361211046090.

Margulis, L., Dolan, M.F. and Guerrero, R. (2000) 'The chimeric eukaryote: Origin of the nucleus from the karyomastigont in amitochondriate protists', *Proceedings of the National Academy of Sciences*, 97(13), pp. 6954–6959.

Martínez-Calvillo, S., López, I. and Hernández, R. (1997) 'pRIBOTEX expression vector: A pTEX derivative for a rapid selection of Trypanosoma cruzi transfectants', *Gene*, (1-2), p. 71-6.

Martinez-Duncker, I. *et al.* (2003) 'A new superfamily of protein-O-fucosyltransferases, alpha2-fucosyltransferases, and alpha6-fucosyltransferases: phylogeny and identification of conserved peptide motifs', *Glycobiology*, 13(12), pp. 1C-5C.

Matsuda, N.M., Miller, S.M. and Evora, P.R.B. (2009) 'The chronic gastrointestinal manifestations of Chagas disease', *Clinics (Sao Paulo, Brazil)*, 64(12), pp. 1219–1224.

McConville, M.J. *et al.* (1990) 'Structure of the lipophosphoglycan from Leishmania major', *Journal of Biological Chemistry*, 265(32), p. 19611-23.

McGhee, R.B. and Cosgrove, W.B. (1980) 'Biology and physiology of the lower trypanosomatidae', *Microbiological Reviews*, 44(1), pp. 140–173.

Medeiros, M.B. de, Guerra, J.A. de O. and Lacerda, M.V.G. de (2008) 'Meningoencephalitis in a patient with acute Chagas disease in the Brazilian Amazon', *Revista Da Sociedade Brasileira De Medicina Tropical*, 41(5), pp. 520–521.

Meiners, J. *et al.* (2020) 'IL-33 facilitates rapid expulsion of the parasitic nematode Strongyloides ratti from the intestine via ILC2- and IL-9-driven mast cell activation', *PLOS Pathogens*, 16(12), p. e1009121.

Mejia, A.M. *et al.* (2012) 'Benznidazole-resistance in Trypanosoma cruzi is a readily acquired trait that can arise independently in a single population', *The Journal of Infectious Diseases*, 206(2), pp. 220–228.

Mengeling, B.J. and Turco, S.J. (1999) 'A high-yield, enzymatic synthesis of GDP-D-[3H]arabinose and GDP-L-[3H]fucose', *Analytical biochemistry*, 267(1), pp. 227–233.

Mittra, B. *et al.* (2003) 'Presence of Multiple mRNA Cycling Sequence Element-binding Proteins in Crithidia fasciculata', *Journal of Biological Chemistry*, 278(29), pp. 26564–26571.

Miyazaki, T. *et al.* (2010) 'Enzymatic synthesis of Lacto-N-difucohexaose i which binds to helicobacter pylori', in *Methods in Enzymology*, pp. 511–524.

Moffatt, J.G. (1966) 'Sugar nucleotide synthesis by the phosphoromorpholidate procedure', in *Methods in Enzymology*. Academic Press (Complex Carbohydrates), pp. 136–142.

Moremen, K W *et al.* (2018) 'Expression system for structural and functional studies of human glycosylation enzymes', *Nat Chem Biol*, 14(2), pp. 156–162.

Moremen, Kelley W *et al.* (2018) 'Expression system for structural and functional studies of human glycosylation enzymes', *Nature Chemical Biology*, 14(2), pp. 156–162.

Moremen, Kelley W and Haltiwanger, R.S. (2019) 'Emerging structural insights into glycosyltransferase-mediated synthesis of glycans', *Nature Chemical Biology*, 15(9), pp. 853–864.

Moremen, Kelley W. and Haltiwanger, R.S. (2019) 'Emerging structural insights into glycosyltransferase-mediated synthesis of glycans', *Nature chemical biology*, 15(9), pp. 853–864.

Morillo, C.A. *et al.* (2015) 'Randomized Trial of Benznidazole for Chronic Chagas' Cardiomyopathy', *New England Journal of Medicine*, 373(14), pp. 1295–1306.

Motta, M.C. *et al.* (1991) 'Cell surface charge and sugar residues of Crithidia fasciculata and Crithidia luciliae', *Microbios*, 68(275), pp. 87–96.

Murrey, H.E. *et al.* (2006) 'Protein fucosylation regulates synapsin Ia/Ib expression and neuronal morphology in primary hippocampal neurons', *Proceedings of the National Academy of Sciences of the United States of America*, 103(1), pp. 21–26.

Murrey, H.E. *et al.* (2009) 'Identification of the Plasticity-Relevant Fucose-α(1−2)-Galactose Proteome from the Mouse Olfactory Bulb', *Biochemistry*, 48(30), pp. 7261–7270.

Murta, S.M.F. *et al.* (2006) 'Deletion of copies of the gene encoding old yellow enzyme (TcOYE), a NAD(P)H flavin oxidoreductase, associates with in vitro-induced benznidazole resistance in Trypanosoma cruzi', *Molecular and Biochemical Parasitology*, 146(2), pp. 151–162.

Nagamune, K. *et al.* (2000) 'Critical roles of glycosylphosphatidylinositol for Trypanosoma brucei', *Proceedings of the National Academy of Sciences*, 97(19), p. 10336 LP – 10341.

Nichols, J.M. and Cross, G.A.M. (1977) 'Isolation of mitochondria and mitochondrial RNA from Crithidia fasciculata', *Journal of General Microbiology*, 99(2), pp. 291–300.

Noguchi, H. and Tilden, E.B. (1926) 'Comparative studies of Herpetomonads and Leishmanias I: Cultivation of Herpetomonadas from insects and plants', *Journal of Experimental Medicine*, 44(3), pp. 307–325.

Nogueira, F.B. *et al.* (2006) 'Increased expression of iron-containing superoxide dismutase-A (TcFeSOD-A) enzyme in Trypanosoma cruzi population with in vitro-induced resistance to benznidazole', *Acta Tropica*, 100(1–2), pp. 119–132.

Nogueira, N.P. *et al.* (2015) 'Proliferation and differentiation of Trypanosoma cruzi inside its vector have a new trigger: redox status', *PloS One*, 10(2), p. e0116712.

Novozhilova, N.M. and Bovin, N.V. (2009) 'D-Arabinose Methabolism: Characterization of Bifunctional Arabinokinase/Pyrophosphorylase of Leishmania major', *Acta Naturae*, 1(3), pp. 81–83.

Nozaki, T. and Cross, G.A. (1994) 'Functional complementation of glycoprotein 72 in a Trypanosoma cruzi glycoprotein 72 null mutant', *Molecular and Biochemical Parasitology*, 67(1), pp. 91–102.

Nunes, M.C.P. *et al.* (2018) 'Chagas Cardiomyopathy: An Update of Current Clinical Knowledge and Management: A Scientific Statement From the American Heart Association', *Circulation*, 138(12), pp. e169–e209.

Ohana, R.F. *et al.* (2009) 'HaloTag7: A genetically engineered tag that enhances bacterial expression of soluble proteins and improves protein purification', *Protein Expression and Purification*, 68, pp. 110–120.

Olsen, O.W. (1974) *Animal Parasites: Their Life Cycles and Ecology*. University Park Press.

Padilla A.M., *et al.* (2022) 'Discovery of an orally active benzoxaborole prodrug effective in the treatment of Chagas disease in non-human primates' *Nature Microbiology*, 7(10), pp. 1536-46.

Padilla-Mejia, N.E. *et al.* (2021) 'Evolution and diversification of the nuclear envelope', *Nucleus*, 12(1), pp. 21–41.

Pasion, S.G. *et al.* (1994) 'Periodic expression of nuclear and mitochondrial DNA replication genes during the trypanosomatid cell cycle', *Journal of Cell Science*, 107 (Pt 12), pp. 3515–3520.

Paulin, J.J. (1977) 'Crithidia fasciculata: reconstruction of the mitochondrion based on serial thick sections and high-voltage electron microscopy', *Experimental Parasitology*, 41(2), pp. 283–289.

Peikert, C.D. *et al.* (2017) 'Charting organellar importomes by quantitative mass spectrometry', *Nature Communications*, 8(1), p. 15272.

Peloso, E.F. *et al.* (2016) 'Trypanosoma cruzi mitochondrial tryparedoxin peroxidase is located throughout the cell and its pull down provides one step towards the understanding of its mechanism of action', *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1864(1), pp. 1–10.

Pereira, M.E. and Hoff, R. (1986) 'Heterogeneous distribution of neuraminidase activity in strains and clones of Trypanosoma cruzi and its possible association with parasite myotropism', *Molecular and Biochemical Parasitology*, 20(2), pp. 183–189.

Pérez-Molina, J.A. and Molina, I. (2018) 'Chagas disease', *The Lancet*, 391(10115), pp. 82–94.

Perez-Molina, J.A. and Molina, I. (2018) 'Chagas disease', *Lancet*, 391(10115), pp. 82–94.

Pettersen, E.F. *et al.* (2021) 'UCSF ChimeraX: Structure visualization for researchers, educators, and developers', *Protein Science: A Publication of the Protein Society*, 30(1), pp. 70–82.

Podestá, D., Stoppani, A. and Villamil, S.F. (2003) 'Inactivation of Trypanosoma cruzi and Crithidia fasciculata topoisomerase I by Fenton systems', *Redox Report: Communications in Free Radical Research*, 8(6), pp. 357–363.

Previato, J.O. *et al.* (1993) 'Structure of the carbohydrate moiety of the glycophosphosphingolipid of Endotrypanum schaudinni', *Glycoconjugate Journal*, 10(4), pp. 340–340.

Rajão, M.A. *et al.* (2014) 'Unveiling benznidazole's mechanism of action through overexpression of DNA repair proteins in Trypanosoma cruzi', *Environmental and Molecular Mutagenesis*, 55(4), pp. 309–321.

Ray, D.S. and Hines, J.C. (1995) 'Disruption of the Crithidia fasciculata RNH1 gene results in the loss of two active forms of ribonuclease H.', *Nucleic Acids Research*, 23(13), pp. 2526–2530.

Roberts, A.J. *et al.* (2014) 'Biochemical and genetic characterization of Trypanosoma cruzi N-myristoyltransferase', *Biochemical Journal*, 459(2), p. 323-32.

Rodrigues, J.A. *et al.* (2015) 'Parasite Glycobiology: A Bittersweet Symphony', *PLoS Pathog*, 11(11), p. e1005169.

Rodriguez-Contreras, D. *et al.* (2007) 'Phenotypic characterization of a glucose transporter null mutant in Leishmania mexicana', *Molecular and biochemical parasitology*, 153(1), pp. 9–18.

Rojas, R. *et al.* (2014) 'The effect of tunicamycin on the glucose uptake, growth, and cellular adhesion in the protozoan parasite crithidia fasciculata', *Current Microbiology*, 69, pp. 541–548.

Ryan, K.A. *et al.* (1993) 'Isolation of virulence genes directing surface glycosyl-phosphatidylinositol synthesis by functional complementation of Leishmania.', *Proceedings of the National Academy of Sciences*, 90(18), pp. 8609–8613.

Ryczko, M.C. *et al.* (2016) 'Metabolic Reprogramming by Hexosamine Biosynthetic and Golgi N-Glycan Branching Pathways', *Sci Rep*, 6, p. 23043.

Sabino, E.C. *et al.* (2013) 'Ten-year incidence of Chagas cardiomyopathy among asymptomatic Trypanosoma cruzi-seropositive former blood donors', *Circulation*, 127(10), pp. 1105–1115.

Sacks, D.L. *et al.* (2000) 'The role of phosphoglycans in Leishmania–sand fly interactions', *Proceedings of the National Academy of Sciences*, 97(1), pp. 406–411.

Sacoman, J.L. *et al.* (2017) 'Mitochondrial O-GlcNAc Transferase (mOGT) Regulates Mitochondrial Structure, Function, and Survival in HeLa Cells', *J Biol Chem*, 292(11), pp. 4499–4518.

Sahin, A. *et al.* (2004) 'Trypanosomatid flagellum biogenesis: ARL-3A is involved in several species', *Experimental Parasitology*, 108(3–4), pp. 126–133.

Schmunis, G.A. and Yadon, Z.E. (2010) 'Chagas disease: a Latin American health problem becoming a world health problem', *Acta Tropica*, 115(1–2), pp. 14–21.

Schneider, M., Al-Shareffi, E. and Haltiwanger, R.S. (2017) 'Biological functions of fucose in mammals', *Glycobiology*, 27(7), pp. 601-18.

Schneider, P. *et al.* (1996) 'Structural studies on a lipoarabinogalactan of Crithidia fasciculata', *Biochemical Journal*, 313, pp. 963–971.

Schneider, P., Nikolaev, A. and Ferguson, M.A.J. (1995) 'The biosynthesis of GDP-D-arabinopyranose in Crithidia fasciculata: Characterization of a D-arabino-1-kinase activity and its use in the synthesis of GDP-[5-3H]D-arabinopyranose', *Biochemical Journal*, 311, pp. 307–315.

Scolaro, E.J., Ames, R.P. and Brittingham, A. (2005) 'Growth-Phase Dependent Substrate Adhesion in Crithidia fasciculata', *Journal of Eukaryotic Microbiology*, 52(1), pp. 17–22.

Shikanai-Yasuda, M.A. and Carvalho, N.B. (2012) 'Oral transmission of Chagas disease', *Clin Infect Dis*, 54(6), pp. 845–852.

Shim, H. and Fairlamb, A.H. (1988) 'Levels of polyamines, glutathione and glutathione-spermidine conjugates during growth of the insect trypanosomatid Crithidia fasciculata', *Journal of General Microbiology*, 134(3), pp. 807–817.

Silber, A.M. *et al.* (2009) 'Glucose uptake in the mammalian stages of Trypanosoma cruzi', *Molecular and Biochemical Parasitology*, 168(1), pp. 102–108.

Sinha, K.M., Hines, J.C. and Ray, D.S. (2006) 'Cell cycle-dependent localization and properties of a second mitochondrial DNA ligase in Crithidia fasciculata', *Eukaryotic Cell*, 5(1), pp. 54–61.

Soeiro, M. de N.C. (2022) ‘Perspectives for a new drug candidate for Chagas disease therapy’, *Memorias Do Instituto Oswaldo Cruz*, 117, p. e220004.

Subedi, G.P. *et al.* (2015) ‘High yield expression of recombinant human proteins with the transient transfection of HEK293 cells in suspension’, *Journal of Visualized Experiments*, 106(e53568).

Takahashi, T. *et al.* (2000) ‘A sequence motif involved in the donor substrate binding by α1,6-fucosyltransferase: the role of the conserved arginine residues’, *Glycobiology*, 10(5), pp. 503–510.

Teston, A.P.M. *et al.* (2013) ‘In vivo susceptibility to benznidazole of Trypanosoma cruzi strains from the western Brazilian Amazon’, *Tropical medicine & international health: TM & IH*, 18(1), pp. 85–95.

Tetaud, E. *et al.* (1994) ‘Characterization of glucose transport and cloning of a hexose transporter gene in Trypanosoma cruzi’, *Proceedings of the National Academy of Sciences of the United States of America*, 91(17), pp. 8278–8282.

Tetaud, E. *et al.* (1997) ‘Kinetoplastid glucose transporters’, *The Biochemical Journal*, 325 (Pt 3), pp. 569–580.

Trochine, A. *et al.* (2014) ‘Benznidazole Biotransformation and Multiple Targets in Trypanosoma cruzi Revealed by Metabolomics’, *PLOS Neglected Tropical Diseases*, 8(5), p. e2844.

Turnock, D.C. and Ferguson, M.A. (2007) ‘Sugar nucleotide pools of Trypanosoma brucei, Trypanosoma cruzi, and Leishmania major’, *Eukaryot Cell*, 6(8), pp. 1450–1463.

Turnock, D.C., Izquierdo, L. and Ferguson, M.A. (2007) ‘The de novo synthesis of GDP-fucose is essential for flagellar adhesion and cell growth in Trypanosoma brucei’, *J Biol Chem*, 282(39), pp. 28853–28863.

Urbanowicz, B.R. *et al.* (2017) ‘Structural, mutagenic and in silico studies of xyloglucan fucosylation in Arabidopsis thaliana suggest a water-mediated mechanism’, *Plant Journal*, 91(6), pp. 931–949.

Valente, M. *et al.* (2019) 'Overview of the role of kinetoplastid surface carbohydrates in infection and host cell invasion: Prospects for therapeutic intervention', *Parasitology*, 146(14), pp. 1743–1754.

do Valle Matta, M.A. *et al.* (1999) 'Cell-surface sialoglycoconjugate structures in wild-type and mutant Crithidia fasciculata', *Parasitology Research*, 85(4), pp. 293–299.

Varki, A. (2017) 'Biological roles of glycans', *Glycobiology*, 27(1), pp. 3–49.

Varki, A. *et al.* (eds) (2022) *Essentials of Glycobiology*. 4th edn. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press. Available at: http://www.ncbi.nlm.nih.gov/books/NBK579918/ (Accessed: 1 November 2022).

Vazquez, M.P. and Levin, M.J. (1999) 'Functional analysis of the intergenic regions of TcP2beta gene loci allowed the construction of an improved Trypanosoma cruzi expression vector', *Gene*, 239(2), pp. 217–225.

Videira, P.A., Marcelo, F. and Grewal, R.K. (2017) 'Glycosyltransferase inhibitors: a promising strategy to pave a path from laboratory to therapy', in *Carbohydrate Chemistry*, pp. 135–158.

Viotti, R. *et al.* (2009) 'Side effects of benznidazole as treatment in chronic Chagas disease: fears and realities', *Expert Review of Anti-Infective Therapy*, 7(2), pp. 157–163.

Viotti, R. *et al.* (2014) 'Towards a paradigm shift in the treatment of chronic Chagas disease', *Antimicrobial Agents and Chemotherapy*, 58(2), pp. 635–639.

de Vries, T. *et al.* (2001) 'Fucosyltransferases: structure/function studies', *Glycobiology*, 11(10), pp. 119R-128R.

Wallace, F.G. (1943) 'Flagellate Parasites of Mosquitoes with Special Reference to Crithidia fasciculata Léger, 1902', *The Journal of Parasitology*, 29(3), pp. 196–205.

Wang, G. *et al.* (1999) 'Novel Helicobacter pylori α1,2-fucosyltransferase, a key enzyme in the synthesis of Lewis antigens', *Microbiology*, 145(11), pp. 3245–3253.

Watanabe, R. *et al.* (1999) 'Mammalian PIG-L and its yeast homologue Gpi12p are N-acetylglucosaminylphosphatidylinositol de-N-acetylases essential in

glycosylphosphatidylinositol biosynthesis', *The Biochemical Journal*, 339 (Pt 1), pp. 185–192.

West, C.M. *et al.* (2021) 'Glycomics, Glycoproteomics, and Glycogenomics: An Inter-Taxa Evolutionary Perspective', *Molecular & cellular proteomics: MCP*, 20, p. 100024.

West, C.M., Wang, Z.A. and van der Wel, H. (2010) 'A cytoplasmic prolyl hydroxylation and glycosylation pathway modifies Skp1 and regulates O2-dependent development in Dictyostelium', *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1800(2), pp. 160–171.

Wiese, T.J., Dunlap, J.A. and Yorek, M.A. (1994) 'L-fucose is accumulated via a specific transport system in eukaryotic cells', *Journal of Biological Chemistry*, 269(36), pp. 22705–22711.

Wilkinson, S.R. *et al.* (2008) 'A mechanism for cross-resistance to nifurtimox and benznidazole in trypanosomes', *Proceedings of the National Academy of Sciences of the United States of America*, 105(13), pp. 5022–5027.

Wilkinson, S.R. *et al.* (2011) 'Trypanocidal activity of nitroaromatic prodrugs: current treatments and future perspectives', *Current Topics in Medicinal Chemistry*, 11(16), pp. 2072–2084.

Wilson, I.B.H. *et al.* (1999) 'Typing of Leishmania lipophosphoglycans by electrospray mass spectrometry', *Molecular and Biochemical Parasitology*, 100, pp. 207–215.

Wong-Madden, S.T. and Landry, D. (1995) 'Purification and characterization of novel glycosidases from the bacterial genus Xanthomonas', *Glycobiology*, 5(1), pp. 19–28.

World Health Organization (2017) *Integrating neglected tropical diseases into global health and development : 4th WHO report on neglected tropical diseases*. Geneva : World Health Organization.

Xavier Da Silveira, E. *et al.* (1998) 'Glycoinositol phospholipids from Endotrypanum species express epitopes in common with saccharide side chains of the lipophosphoglycan from Leishmania major', *The Biochemical Journal*, 329 (Pt 3), pp. 665–673.

Xu, Y. *et al.* (2012) 'The Skp1 protein from Toxoplasma is modified by a cytoplasmic prolyl 4-hydroxylase associated with oxygen sensing in the social amoeba Dictyostelium', *The Journal of Biological Chemistry*, 287(30), pp. 25098–25110.

Yazaki, E. *et al.* (2017) 'Global Kinetoplastea phylogeny inferred from a large-scale multigene alignment including parasitic species for better understanding transitions from a free-living to a parasitic lifestyle', *Genes & Genetic Systems*, 92(1), pp. 35–42.

Yoshida, N. and Cortez, M. (2008) 'Trypanosoma cruzi: Parasite and Host Cell Signaling during the Invasion Process', in B.A. Burleigh and D. Soldati-Favre (eds) *Molecular Mechanisms of Parasite Invasion: Subcellular Biochemistry*. New York, NY: Springer (Subcellular Biochemistry), pp. 82–91.

Zhang, K. *et al.* (2004) 'The LPG1 gene family of Leishmania major', *Molecular and biochemical parasitology*, 136(1), pp. 11–23.

Zhang, L. *et al.* (2010) 'Helicobacter hepaticus Hh0072 gene encodes a novel alpha1-3-fucosyltransferase belonging to CAZy GT11 family', *Glycobiology*, 20(9), pp. 1077–1088.

Zingales, B. *et al.* (2009) 'A new consensus for Trypanosoma cruzi intraspecific nomenclature: second revision meeting recommends TcI to TcVI', *Memorias Do Instituto Oswaldo Cruz*, 104(7), pp. 1051–1054.

Zingales, B. *et al.* (2012) 'The revised Trypanosoma cruzi subspecific nomenclature: rationale, epidemiological relevance and research applications', *Infect Genet Evol*, 12(2), pp. 240–253.

Zingales, B. (2018) 'Trypanosoma cruzi genetic diversity: Something new for something known about Chagas disease manifestations, serodiagnosis and drug sensitivity', *Acta Tropica*, 184, pp. 38–52.

Zuma, A.A., Dos Santos Barrias, E. and de Souza, W. (2021) 'Basic Biology of Trypanosoma cruzi', *Current Pharmaceutical Design*, 27(14), pp. 1671–1732.