Investigations of the role of peroxisomes in sterol biosynthesis in the slime mould *Dictyostelium discoideum*

MURTAKAB YOUNIS AL-HEJJAJ (MSc.)

A thesis submitted to the University of Sheffield for the degree of Doctor of Philosophy

Department of Molecular Biology and Biotechnology

University of Sheffield

November 2017

Declaration

I hereby declare that this thesis is my own work and effort. Whereas other sources of information have been used, they have been acknowledged.

Murtakab Y. Al-hejjaj

Abstract

Dictyostelium discoideum is a good model organism to study a variety of cellular processes. Its complete genome sequence is known (dictybase.org). It has a unique life style, with motile unicellular and multicellular stages, and multiple cell types (Annesley and Fisher 2009). It is widely used to study chemotactic motility and cytoskeletal dynamics. It is also used for studying molecular pathogenesis and treatment of human disease. Furthermore, since *Dictyostelium* cells readily take up drugs, it has led to identification of drug targets. For instance, farnesyl diphosphate synthase was found, by use of *D. discoideum*, to be the primary target of bisphosphonate drugs that are widely used to prevent and treat osteoporosis.

A major challenge with this organism is the difficulty of making targeted mutations by homologous recombination. It is difficult to generate targeting constructs because of the high AT contents (77.4%) of its genome. Introduction of the Clustered Regularly Interspaced Short Palindromic Repeats system (CRISPR) into *D. discoideum* for obtaining gene knockouts would therefore be potentially most helpful and here I report on our progress in setting this up.

All previous investigations of the intracellular location of the first four enzymes involved in sterol biosynthesis from farnesyl diphosphate (FDP) have found that they locate to the ER membrane in eukaryotic cells. However, we provide strong evidence that these enzymes are peroxisomal in *D. discoideum*. The first step in the pathway is catalysed by squalene synthase SQS. Surprisingly, (*Dd*SQS) contains a typical peroxisomal targeting signal type 1 (PTS1). All known PTS1-containing proteins are localized to the peroxisomal matrix whereas membrane proteins use a different pathway. However, we found that *Dd*SQS behaves as a membrane protein. Interestingly, as in other organisms, the *Dd*SQS has a C-terminal amino acid sequence potentially forming a hydrophobic helix but which in *D. discoideum* is located immediately upstream of the PTS1. Deletion of this helix does not affect peroxisomal targeting but does affect DdSQS association with the membrane and may therefore serve as a tail anchor. Furthermore, the helix plays an important role in forming a SQS homodimer. SQS is the first example of a peroxisomal membrane protein that makes use of the PTS1 pathway for its localization.

We have shown that the first four enzymes of sterol biosynthesis from FDP are associated with the peroxisomal membrane by using structured illumination microscopy. This is a relatively new technique that allows imaging beyond the resolution limit for light microscopy. Furthermore, the topology of the enzymes with respect to the peroxisomal membrane was experimentally determined. We provide a model of how sterol biosynthesis from FDP takes place in the peroxisomes of *D. discoideum*.

Acknowledgment

In the name of Allah most gracious most merciful.

Unlimited thanks to Allah, for his protect and care that I would not be as I am without them.

I would like to thank my supervisors Dr Donald J. Watts and Dr Ewald H Hettema for their support, guidance, advice and huge knowledge that I have received throughout my PhD journey.

My thanks to my sponsor The Higher Education Committee of Education development in Iraq -Iraqi Prime minister office (HCED) for offering me the opportunity to do my PhD in the United Kingdom. Many thanks to the Ministry of Higher Education and Scientific Research, University of Basrah, College of Veterinary Medicine, for their support.

Special thanks to Adnan Al-mousawi, Ehssan Al-Bermany, Yasir Al-abdali, Naer Al-Kabbawi for being great friends and for supporting me throughout the PhD. Thanks a lot, to all my Iraqi colleagues in Sheffield for their support, and love that I have achieved during my study. A huge thank to all Hettema's lab member; Alison, Nadal Alsayi, Joanne, Mayyadah, John, Gemma, Paul, Lakhan, and Georgia for being good colleagues. I would like to thank Dr Phil Mitchell for his help for doing Northern blot.

My greatest indebtedness to my Mum and Dad for their emotional, love and continues prayers. I am grateful also to my mother and father in law for their support and encouragement.

Big thanks to my brother (Murtada) and my sisters; Maalim, Hanaan and Batool for their support and endless encouragement.

Finally, I would like to express my warm and great regards and thanks to the woman that I could not achieved my PhD without her unlimited help, support and love, my beloved wife "Yessar". My great thanks extend to my children Fatimah, Mohammed, and Yousef for making my live shiny and hopeful.

Dedication

To the souls of

My sister (Rasha), who left us very early

My uncle (Abdulsaheb), who was my example

All of those who gave their lives defending my beloved home "Iraq, the cradle of civilizations"

Table of Contents

Abbreviations

A15	Actin 15 promoter
aa	Amino acids
APS	Ammonium persulphate
A. subglobosum	*Acetostelium subglobosum*
ATP	Adenosine triphosphate
A.thaliana	*Arabidopsis thaliana*
BF	Bright field
CHO	Chinese hamster ovary
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
C-terminal	Carboxyl-terminal
DNA	Deoxyribonucleic acid
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
dNTPs	Equimolar mixture of dTTP, dATP, dCTP and dGTP
DTT	Dithiothreitol
Dd	*Dictyostelium discoideum*
D. fasciculatum	*Dictyosteleum fasciculatum*
D. lacteum	*Dictyostelium lacteum*
D. fasciculatum	*Dictyostelium fasciculatum*
D. purpureum	*Dictyostelium purpureum*
E.coli	*Escherichia coli*
ECL	Enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
FDPS	Farnesyl diphosphate synthase
GFP	Green fluorescent protein
hr	Hour(s)
HRP	Horseradish peroxidase
Hs	*Homo sapiens*
IgG	Immunoglobulin G
IRD	Infantile Refsum disease

min	minute(s)
mPTS	Peroxisomal membrane protein targeting signal
mRFP	monomeric Red fluorescent protein
NALD	Neonatal adrenoleukodystrophy
NEB	New England Biolabs
N-terminal	Amino-terminal
nt	Nucleotides
OD	Optical density
ORF	Open reading frame
OSC	Oxidosqualene Cyclase
PAM	Protospacer adjacent motif
PEG	Polyethylene glycol
PBD	Peroxisome biogenesis disorder
PCR	Polymerase chain reaction
SPED	Single peroxisomal enzyme deficiency
PMP	Peroxisomal membrane protein
P.pallidum	*Polysphondylium pallidum*
PTS	Peroxisome targeting signal
RCDP	Rhizomelic chondrodysplasia punctata
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SQE	Squalene epoxidase
SMT	Cycloartenol -C-24-methyltransferase
SQS	Squalene synthase
TEMED	N,N,N-Tetramethylethylenediamine
WT	Wild type
ZS	Zellweger syndrome

Amino acids

Ala	A	Alanine
Cys	C	Cysteine
Asp	D	Aspartic acid
Glu	E	Glutamic acid
Phe	F	Phenylalanine
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Lys	K	Lysine
Leu	L	Leucine
Met	M	Methionine
Asn	N	Asparagine
Pro	P	Proline
Gln	Q	Glutamine
Arg	R	Arginine
Ser	S	Serine
Thr	T	Threonine
Val	V	Valine
Trp	W	Tryptophan
Tyr	Y	Tyrosine

List of figures:

List of tables

Chapter one

1.1 Introduction

Osteoporosis is a complex health problem and is also considered as a polygenic disease in elderly people. It is especially common in postmenopausal women (García-Unzueta et al., 2008; Geng et al., 2007). It is characterised by systemic loss of bone mass and reduction in bone mineral density (Woo et al., 2012). This loss results in increased risk of spontaneous fracture at the vertebra, wrist, humerus, distal forearms and hip. These fractures frequently result in a decline in the quality of life owing to pain and loss of mobility. They also lead to serious and great disability in addition, to increasing the mortality rate and causing premature deaths (Cummings et al., 1985).

Nitrogen-containing bisphosphonates (NBPs), such as alendronate, are the most widely-used drugs for prevention and treatment of osteoporosis. Their use has dramatically increased in recent years (Udell et al., 2006; Usher et al., 2006). For instance, since 2005 prescribing these drugs increased rapidly in the UK especially in over 70-years old women (Watson et al., 2007). They decrease further bone loss by inducing the bone-resorbing osteoclasts to die owing to apoptosis. The target of these drugs was identified firstly in the slime mould *Dictyostelium discoideum* and is the essential intracellular enzyme farnesyl diphosphate synthase (FDPS) (Grove et al., 2000). Then the same target (FDPS) for NBPs was described in osteoclasts (Rogers et al., 2011). In order to reach their target, BPs have to enter the cell cytosol. BPs bind tightly to bone mineral and then osteoclasts take up these drugs selectively by an unclear mechanism. According to Nuttall et al. (2012) the target (FDPS) in the amoebae of *D. discoideum* is a peroxisomal enzyme. The previous research shows also that a peroxisomal targeting signal type 2 (PTS2) is responsible for guiding FDPS to the peroxisome.

Dictyostelium discoideum has a different sensitivity to individual bisphosphonate (BP) drugs. If they are ranked according to their potency for inhibiting *D. discoideum* growth, the ranking is the same as for inhibition of bone resorption. It was this that first suggested that the target has the same cellular location in both cell types. Since FDPS is a PTS2 protein and therefore is transported to peroxisomes by Pex7, the attempt we made to change the location of the FDPS into the cytosol by knocking out Pex7 and test whether this would affect sensitivity to NBPs.

Up to now, a drawback in working with *Dictyostelium* has been the difficulty in making specified changes to this organism's genome. Although there are some techniques for doing this, they are difficult to perform. Recently, a new system for manipulating the genome of a wide range of organisms has been described. The Clustered Regulatory Interspaced Short Palindromic Repeats (CRISPR) and CRISPR associated protein system was discovered as a bacterial defence system against invader nucleic acids such as phage and plasmid. The CRISPR system has already been shown to be of huge biotechnological benefit such as gene knock out by making a specific single strand DNA break (nick) or double strand DNA breaks (cut), gene expression knockdown by interfering with gene transcription, gene expression activation which may be done by fusing a catalytically inactive Cas9 to a transcription activator peptide, genomic DNA purification by fusing a catalytically inactive Cas9 to epitope tags, and genomic loci visualization via fusion of a catalytically inactive Cas9 to a fluorescent protein.

At the same time as we attempted to knock out Pex7, we discovered that the next four enzymes following FDPS in the sterol biosynthesis pathway have an unusual location in *D. discoideum*, which is the peroxisome. This was further investigated in detail in chapters 4 and 5.

1.2 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-Cas associated system

1.2.1 CRISPR definition

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) is an immune defence system located in prokaryotes. CRISPR and CRISPR associated protein systems are small non-coding RNAs associated with special Cas protein endonucleases and the respective locus construct adaptive immune system, which protects prokaryotes from foreign genetic material like plasmids and viruses (Carte et al., 2014). CRISPR loci containing short sequences acquired from different invaders are stored in-between specific sequence repeats of 24-37 base pairs (Grissa et al., 2007). CRISPR biogenesis is a transcription process to produce multiple CRISPR RNA (crRNA). Each one contains

two different sequence segments, one derived from the invader and the other comes from the repeat elements. In addition to CRISPR associated protein (Cas), they form an assembly which is guided be crRNA to attack and destroy the complementary foreign nucleic acid (Jore et al., 2012).

1.2.2 CRISPR Defence steps

The CRISPR-Cas immune systems requires three important steps to occur consecutively (Carte et al., 2014). (1) First, an adaptation process whereby short sequences of invader's DNA is inserted into the host's CRISPR array (Datsenko et al., 2012; Yosef et al., 2012). (2) Second, expression of the appropriate Cas protein in addition to expression of the CRISPR array and processing of the product into a guide RNA which has a complementary region to the sequence of the invader DNA (Deltcheva et al., 2011; Nam et al., 2012; Richter et al., 2012). (3) Third, mature guide RNA next guides the Cas protein to attack and cut the invader nucleic acid (Garneau et al., 2010; Gasiunas et al., 2012; Sinkunas et al., 2013; Zhang et al., 2012).

1.2.3 CRISPR–Cas classification

According to the new classification, there are three main CRISPR-Cas system types; type I, type II and type III (Figure 1.1), in addition to many subtypes belonging to each one of them (Carte et al., 2014; Shmakov et al., 2017). For this thesis, only the type II CRISPR-CAS9 system is relevant.

In the type II system, pre-crRNA (a single strand RNA contains the acquired spacers and the original repeats) and a transRNA (small RNA containing 25 base pair sequences complementary to the repeat fragments of pre-crRNA transcripts) bind together forming an RNA duplex. Then maturation of crRNA takes place by cleavage within the repeats by the host endogenous RNase III (Deltcheva et al., 2011). To adapt the system for use in genome editing approaches, these two RNAs (pre-crRNA and *trans*RNA) were fused to generate a synthetic guide RNA (sgRNA), which has the same function (Jinek et al., 2012).

The CRISPR system type II requires only two components; a Cas9 endonuclease and a synthetic guide RNA (sgRNA). The Cas9 is directed by the sgRNA to a specific sequence site in the genome and then introduces either a single (nick) or a double strand (cut) DNA break (Jinek et al., 2012). The sgRNA is responsible for the target site specificity, which is obtained by adding a 20bp of the target complimentary DNA sequence to the 5′ end of the sgRNA. Furthermore, the Cas9 requires a specific DNA sequence called protospacer adjacent motif (PAM), which is located immediately after the target sequence in the genome, generally it is composed the NGG nucleotides sequence (Figure 1.2) (Hsu et al., 2013).

The DNA double strand break (DSB) created by the Cas9 can be repaired within cells by one of the two DNA repair mechanisms; homology-directed repair (HR) or non-homologous end-joining (NHEJ) in the presence or absence of a corresponding homologous DNA template, respectively (Sander and Joung, 2014). The NHEJ introduces deletions and/or insertions of a varying size in the genomic DNA sequence, which leads to a frame shift mutation in the targeted open reading frame (Su et al., 2016). Whereas repair by HR requires a DNA template. This can be used to ones advantage by first introducing a break with CRISPR and repairing it with a homologous sequence that contains point mutations or tags for instance to edit the genome. Thus, adding a donor DNA template containing a mutated PAM sequence and homologous regions flanking both sides of the target site improves the gene editing (DiCarlo et al., 2013; Song and Stieger, 2017). Generally, homologous recombination mechanism is used to repair single DNA break (nicks) introduced by cas9 (Bothmer et al., 2017).

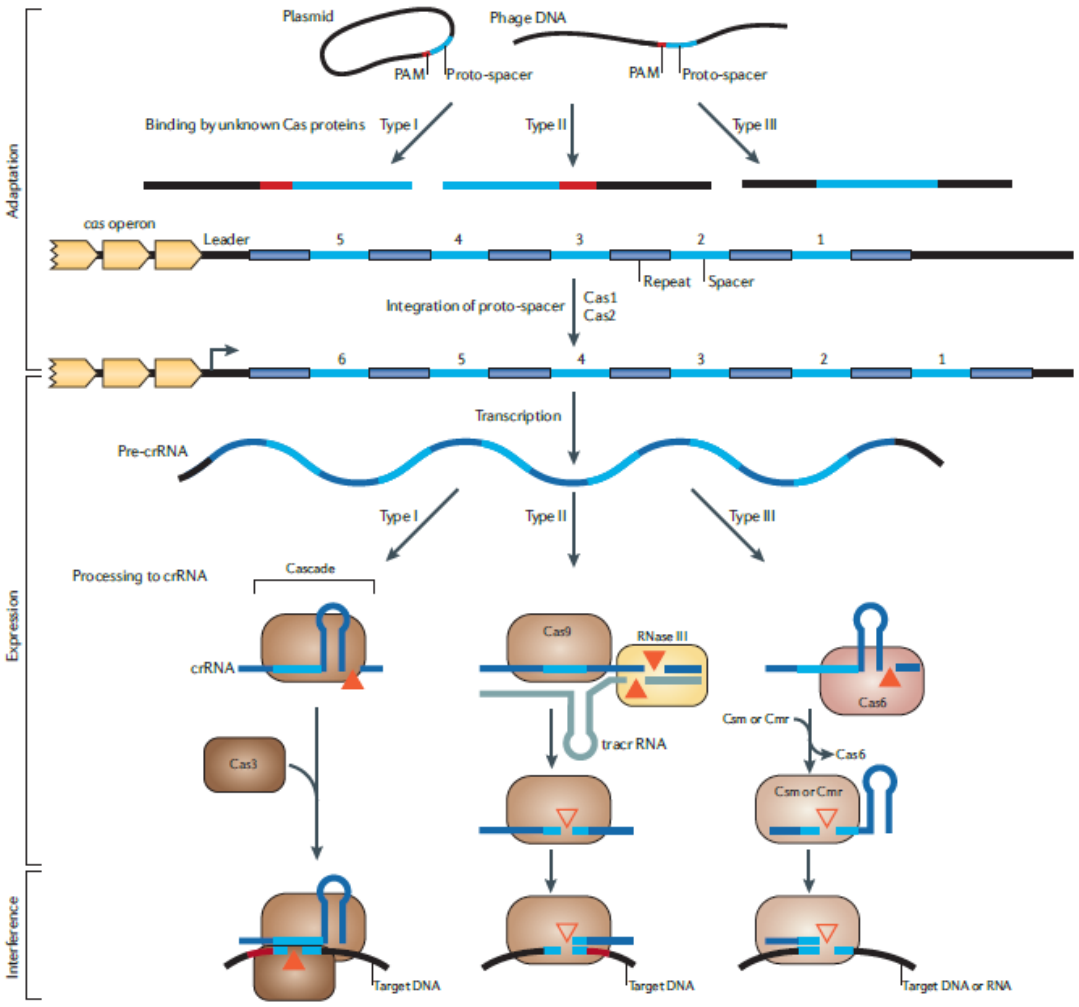


Figure 1.1: Schematic diagram of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) steps. Adaptation, expiration, and interference. And CRISPR types; typeI, typeII, and typeIII. Figure reproduced from (Makarova et al., 2011) with permission.

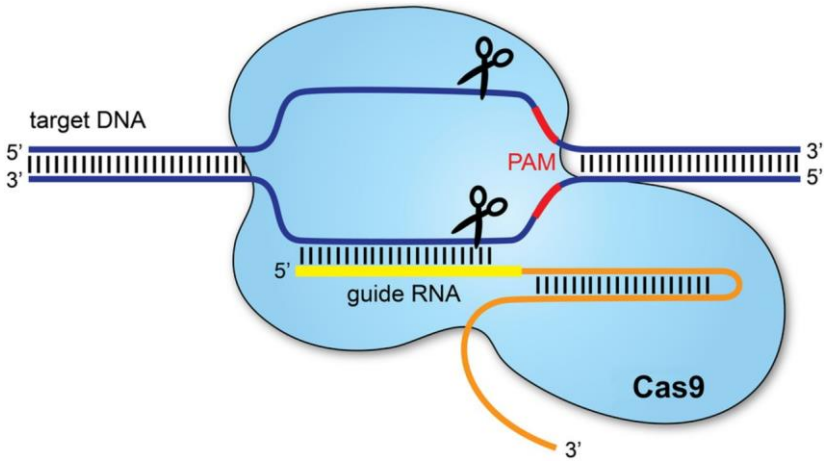


Figure 1.2: Schematic diagram for the CRISPR typeII. The Protospacer Adjacent Motif (PAM) in addition to Cas9 and the guide RNA play an essential role in this immune system. Figure reproduced from (Redman et al., 2016) with permission.

1.3 Bisphosphonate drugs

Bisphosphonate drugs (BPs) were first synthesized in Germany in 1865, (Fleisch, 1998). BPs are divided into two different groups; the nitrogen-containing group (NBPs) such as alendronate, risedronate and ibandronate, and the non-nitrogen containing group (N-NBPs) like etidronate and clodronate (Benford et al., 1999; Green, 2004). Both these groups have been used to treat and prevent some diseases, for example, osteoporosis. Different potencies for the two groups have been noted, with NBPs being more potent than N-NBPs on osteoclasts. NBPs inhibit the farnesyl diphosphate synthase (FDPS) enzyme. It has been demonstrated that FDPS, which catalyses the final two steps of the mevalonate pathway, is the target of nitrogen-containing bisphosphonates (Nuttall, et al, 2014). Inhibition of FDPS synthase induces two important effects in the cell. First, isopentenyl diphosphate accumulates and this is subsequently converted into a toxic nucleotide metabolite (ApppI). On the other hand, isoprenoid lipid synthesis will be prevented, which is essential for post-translational prenylation of small GTPase, thus, disrupting the function of these survival proteins (Rogers et al., 2011). It has been shown that the N-NBP drugs target is different from NBPs, which is not FPP (Grove et al., 2000). The N-NBPs induce osteoclast apoptosis by incorporation into non-hydrolysable analogues of adenosine triphosphate (ATP) (Rogers et al., 2011).

1.4 Mevalonate pathway

The mevalonate (MVA) pathway which leads to the pathway of sterol biosynthesis is a vital cellular metabolic pathway engaging in many different cell functions (Yeganeh et al., 2014). Sterols are present in all eukaryotic cells, such as mammals and higher plants, and in some protozoa like *Leishmania* and *Trypanosoma* in different forms (Goldstein and Brown, 1990). For example, in human, cholesterol is essential for cell viability as it is an essential component of the plasma membrane (Endo, 2010). Together with cholesteryl esters it constitutes the majority of the plasma membrane lipids (Vance, 2012). Furthermore, the majority of plasma lipoprotein is cholesterol. It is also the precursor for steroid hormone biosynthesis (Simons and Vaz, 2004).

The mevalonate pathway comprises irreversible steps, starting from two acetyl CoA molecules, which are obtained from glucose metabolism and mitochondrial fatty acid beta-oxidation. Several enzymes play crucial roles in this pathway, some of them are involved in other metabolite pathways by providing them with essential precursors. For instance, farnesyl diphosphate synthase uses isopentenyl pyrophosphate to produce farnesyl pyrophosphate, which is important for synthesis of heme a, prenylated proteins, ubiquinone, and dolichol pyrophosphate, in addition to its main pathway the cholesterol biosynthesis. It has been suggested that the MVA pathway reactions take place in different organelles such as peroxisomes and the endoplasmic reticulum (ER) (Kovacs et al., 2002). The intracellular locations of mevalonate pathway enzymes are controversial, but it appears that farnesyl diphosphate synthase is in the peroxisome (Nuttall et al., 2014).

1.5 Sterol biosynthesis pathway

Sterols are types of lipid essential for all known eukaryotic plasma membranes, due to their role in the regulation of membrane permeability and fluidity. In addition, they interact with other membrane component such as proteins and lipids (Darnet and Rahier, 2004; Hartmann, 1998). Organisms vary in the sterols they produce. For instance ergosterol, dictysterol and cholesterol are made in fungi, *Dictyostelium* and humans,

respectively (Benveniste, 2004). Although cholesterol is the main sterol in mammals, many fundamental compounds are derived from cholesterol such as bile acids, steroid hormones and vitamin D (Nes, 2011). In plants, several types of sterols are synthesized including campesterol, stigmasterol and sitosterol (Benveniste, 2004). Those sterols supply plants cells with growth hormone, triterpenoids and brassinosteroids.

Sterols are a diverse group of lipids, but they are all synthesised from squalene, which is produced by squalene synthase from two molecules of farnesyl diphosphate. Then squalene is converted by squalene epoxidase to squalene epoxide. Subsequently, the pathway can follow one of two routes depending on organism. Non-photosynthetic and photosynthetic plants make lanosterol and cycloartenol, respectively. Lanosterol synthesis is catalysed by lanosterol synthase, whereas cycloartenol is produced by cycloartenol synthase. Both these sterol molecules have a similar chemical structure except for one more cyclopropane ring in cycloartenol, which is cleaved later on in the pathway (Figure 1.3) (Brumfield et al., 2017). Remarkably, it has been shown that *Arabidopsis thaliana* and *Dictyostelium discoideum* mutated in cycloartenol synthase will shift from cycloartenol to lanosterol synthesis (Segura et al., 2003). In *D. discoideum*, squalene epoxide is converted to cycloartenol (Nes et al., 1990). This suggests that the sterol biosynthesis pathway in *D. discoideum* is similar to that in plants.

Among the long pathway of sterol biosynthesis, there are slight differences between forming phytosterol in plant and sterols formed in mammals or fungi. In plants, conversion of cycloartenol to phytosterol requires removal of two methyl groups at the C-4 position. The first group is catalysed in the early steps of the pathway, while, the second group is removed in a very late step. Even though similar chemical changes in lanosterol take place in mammals and fungi, the two C-4 methyl groups are removed from lanosterol early in the pathway (Hartmann, 1998; Nes et al., 1990).

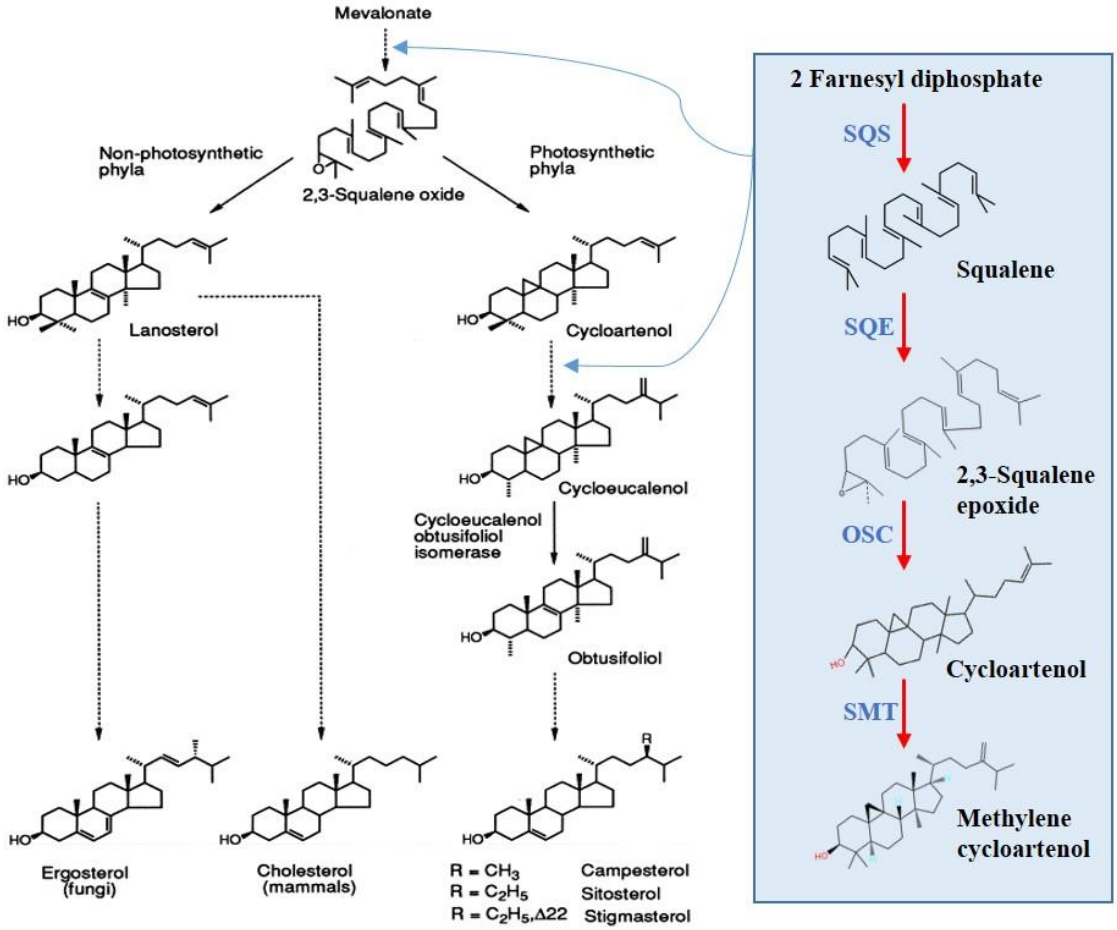


Figure 1.3: Sterol biosynthesis in fungi, mammals and plant. Initially, sterols are biosynthesised from mevalonate. Squalene epoxide converts to cycloartenol or lanosterol based on whether the organisms are photo- or non-photosynthetic, respectively. More than thirty enzymes are involved in the sterol biosynthesis pathway, starting from mevalonate until the sterol production. The highlighted box containing the enzymes of interest for this study (SQS, squalene synthase; SQE, squalene epoxidase; OSC, oxidosqualene cyclase and SMT, cycloartenol-C-24-methyltransferases). Dotted arrows indicate multiple biosynthetic pathway steps. Figure adapted from (Hartmann, 1998) with permission.

1.6 Peroxisomes

Peroxisomes, which include several terms such as microbodies, glycosomes glyoxysomes, and Woronin-bodies, are single membrane-bound subcellular organelles present in almost all eukaryotic cells including single, multicellular microorganisms, plants and animals (van den Bosch et al., 1992). However, among parasitic eukaryotic protozoa, some species such as *Entamoeba*, *Trichomonas* and *Giardia* lack peroxisomes (Parsons et al., 2001). Moreover, Žárský and Tachezy (2015) reported that all the members of the order Trichocephalida (tapeworms, flukes, and parasitic roundworms) which have been studied are devoid of peroxisomes due to absence of many genes related to this organelle (Tsai et al., 2013; Žárský and Tachezy, 2015).

Several different enzymes are located in this ubiquitous organelle, that help catalyze many important metabolic pathways, for instance; very long-chain fatty acids β-oxidation, plasmalogens biosynthesis, and bile acid metabolism (Wanders and Waterham, 2006). Initially, microbodies were described in mouse kidney cells using electron microscopy by a PhD student in Karolinska Institute, Stockholm (Rhodin, 1954). They are spherical in shape with a size range between 0.1 to 1μm in diameter (van den Bosch et al., 1992). Peroxisomes vary in abundance, ranging from a few till hundreds per a single cell in yeast and mammal cells, respectively (Gould and Valle, 2000). Furthermore, numbers of peroxisomes vary among the tissues in the same system, for example, liver, kidneys and central nervous system cells have more peroxisomes than any others cells (Mast et al., 2011). Depending on the organism, cell type and activity, varied biochemical reactions can take place in peroxisomes. Unlike the DNA containing organelles; nucleus, mitochondria, and chloroplasts, peroxisomes depend on the nuclear coding system to make their proteins (Lazarow and Fujiki, 1985).

1.6.1 Peroxisome biogenesis

Generally, proteins associated with peroxisome can be divided into two groups; proteins related to peroxisome biogenesis and replication, and enzymes that catalyzed different biological activities (Distel et al., 1996). Peroxisome biogenesis can be clustered by three fundamental cellular processes. First, peroxisomal membrane-associated proteins import. Second, import of peroxisomal matrix proteins and third, the control mechanism

of peroxisomes proliferation and abundance. First identification of the crucial genes for peroxisome biogenesis was started in the second half of 20th century. Peroxisomes are essential organelles for *S. cerevisiae* growth on media containing oleic acid as the sole carbon source. This was used to identified the first genes required for peroxisome assembly (Erdmann et al., 1989). In further studies, genetic screens for peroxisome assembly mutants were developed in Chinese hamster ovary (CHO) cells and various yeast species (Tsukamoto et al., 1990; Gould et al., 1992; Liu et al., 1992; Purdue and Lazarow, 2001). Using the genes identified via these genetic methods, led to the identification of the mutations in Zellweger spectrum disorder patients.

Proteins required for peroxisome biogenesis are called peroxins. The main function of this group of proteins is to control and maintain the peroxisome biogenesis and the peroxisomal import of membrane and matrix proteins. Peroxin nomenclature was unified in 1996 as PEX genes (Distel et al., 1996) see table (1.1).

Table 1.1: Peroxisomal biogenesis proteins (PXEs) and their function in *H. sapiens, S. cerevisiae and D. discoideum*

Peroxin	*Hs*	*Sc*	*Dd*	General Function	Molecular Biological Activity	References
Pex1	*	*	*	Peroxisomal matrix protein import	AAA+ATPase that is required for recycling of the receptor Pex5	(Erdmann et al., 1991)
Pex2	*	*	*	Peroxisomal matrix protein import	RING finger PMP and required for Pex5 ubiquitination	(Tsukamoto et al., 1991)
Pex3	*	*	*	PMP import and *de novo* formation	Acts as the docking complex for pex19p and controls peroxisomes *de novo* formation	(Höhfeld et al., 1991)
Pex4		*	*	Peroxisomal matrix protein import	Ubiquitin-conjugating enzyme required for ubiquitination and recycling Pex5	(Wiebel and Kunau, 1992)
Pex5	*	*	*	Peroxisomal matrix protein import	Receptor for PTS1 containing proteins	(Van der Leij et al., 1993)
Pex6	*	*	*	Peroxisomal matrix protein import	AAA+ATPase required for recycling of the receptor Pex5	(Spong and Subramani, 1993)
Pex7	*	*	*	Peroxisomal matrix protein import	Receptor for PTS2 containing proteins	(Marzioch et al., 1994)
Pex8		*		Peroxisomal matrix protein import	Binding the Pex5 and connecting the docking proteins with the RING finger	(Liu et al., 1995)
Pex10	*	*	*	Peroxisomal matrix protein import	Ring-finger PMP forms part of the importomer and required for Pex5 ubiquitination	(Kalish et al., 1995)
Pex11	*	*	*	Regulation of Peroxisome size and abundance	Involved in recruitment of the fission machinery and peroxisomal membrane elongation	(Erdmann and Blobel, 1995)
Pex12	*	*	*	Peroxisomal matrix protein import	Ring-finger PMP forms part of the importomer and required for Pex5 recycling	(Kalish1 et al., 1996)
Pex13	*	*	*	Peroxisomal matrix protein import	Forms part of the docking complex of importomer for matrix proteins translocation	(Elgersma et al., 1996)
Pex14	*	*	*	Peroxisomal matrix protein import	Forms part of the docking complex of importomer for matrix proteins translocation	(Huhse et al., 1998)
Pex15		*		Peroxisomal matrix protein import	Membrane receptor and anchor for Pex6 (Pex26 homologue in humans)	(Elgersma et al., 1997)
Pex16	*		*	PMP import and *de novo* formation	Acts as a membrane receptor for Pex3, involves in PMPs recruitment	(Honsho et al., 1998a)
Pex17		*		Peroxisomal matrix protein import	Component of the importomer docking complex	(Huhse et al., 1998)
Pex18		*		Peroxisomal matrix protein import	Co-receptor for PTS2 containing protein	(Purdue et al., 1998)
Pex19	*	*	*	PMPII targeting and *de novo* formation	Chaperone and import receptor for PMP type2 import pathway, required for *de novo* biogenesis	(Götte et al., 1998)

Pex20				Peroxisomal matrix protein import	Co-receptor for PTS2 containing protein import pathway, present most yeast	(Titorenko et al., 1998)
Pex21		*		Peroxisomal matrix protein import	Co-receptor for PTS2 containing protein import pathway	(Purdue et al., 1998)
Pex22		*		Peroxisomal matrix protein import	Membrane anchor protein for Pex4p, involved in Pex5 recycling	(Koller et al., 1999)
Pex23				Peroxisome abundance	homologous to Pex30, Pex31 and Pex32 in function, present in *Yarrowia lipolytica*,	(Brown et al., 2000)
Pex24				Peroxisome abundance	Involved in peroxisome assembly formation, resent in *Y. Lipolytica*,	(Tam and Rachubinski, 2002)
Pex25		*		Regulation of Peroxisome size and abundance	Membrane elongation	(Smith et al., 2002)
Pex26	*			Peroxisomal matrix protein import	Peroxisomal membrane receptor for Pex1-Pex6 protein complex and anchor pex6p in human cellls	(Matsumoto et al., 2003)
Pex27		*		Regulation of Peroxisome size and abundance	Acts as a regulator for peroxisome fission, negatively	(Rottensteiner et al., 2003)
Pex28		*		Regulation of Peroxisome size and abundance	Acts as a regulator for peroxisomal size, abundance and distribution	(Vizeacoumar et al., 2003)
Pex29		*		Regulation of Peroxisome size and abundance	Acts as a regulator for peroxisomal size, abundance and distribution	**(**Vizeacoumar *et al*., 2003)
Pex30		*		Regulation of Peroxisome size and abundance	Regulates peroxisome abundance	(Vizeacoumar et al., 2004)
Pex31		*		Regulation of Peroxisome size and abundance	Controls peroxisome size	(Vizeacoumar *et al*., 2004)
Pex32		*		Regulation of Peroxisome size and abundance	Controls peroxisome size	(Vizeacoumar *et al*., 2004)
Pex33				Peroxisomal matrix protein import	Involved in forming peroxisomal docking complex, present in *Neurospora crassa*.	(Managadze et al., 2010)
Pex34		*		Regulation of Peroxisome size and abundance	Acts as a regulator for peroxisome fission, positively	(Tower et al., 2011)

A * is displayed if the peroxin has identified in any of the above species. *Hs, Homo sapiens; Sc, Saccharomyces cerevisiae and Dd,Dictyostelium discoideum* Peroxins present were checked in Entrez, Saccharomyces genome database and Dictybase for humans, yeast, and amoebae respectively.

1.6.2 Peroxisomal abundance and size

Peroxisomal size and number are controlled and maintained by some *pex* genes. However, in mutants where abundance and size are affected, peroxisomes are still functional (Yuan et al., 2016). The Pex11 family of proteins are required for peroxisome fission and maintenance in all eukaryotes tested. Other peroxisomal membrane proteins also contribute to size and abundance but are not conserved among eukaryotes. (Erdmann and Blobel, 1995; Rottensteiner et al., 2003; Smith et al., 2002; Tam et al., 2003). Furthermore, external signals can regulate peroxisomal abundance. For instance, growth of yeast on carbon sources that require peroxisomes for their utilisation in fungi can lead to massive peroxisome proliferation of number and size.

1.6.3 Peroxisomal functions

Peroxisomes have both common functions and either organism, condition or cell type specific functions. Such significant metabolic processes require a collaboration with other cellular organelles, like the mitochondria, in order to share metabolites and create the final biological product (Schrader and Yoon, 2007).

The universal function of peroxisomes is β-oxidation of fatty acid and fatty acid like molecules, in which the fatty acids (FAs) are oxidised to acetyl-coenzyme A. A product principally used in the citric acid (Krebs) cycle. In yeast and plant cells, β-oxidation takes place in peroxisomes only, whereas in higher eukaryotes such as mammalian cells fatty acids can also be oxidised in mitochondria. Peroxisomal and mitochondrial β-oxidation pathways have the same enzymatic reactions. Nevertheless, distinct enzymes encoded by different genes are involved (Poirier et al., 2006). One of the major differences between the two β-oxidation pathways is that they degrade different length of fatty acids. For instance, the very long chained fatty acids (VLCFAs) are metabolized in peroxisomes, but not by mitochondria and therefore prevents the toxic effects of these compounds accumulation. Whereas short, medium and long chained fatty acids are handled in mitochondria (Wanders & Waterham 2006). Another difference is that mitochondria β-oxidation supplies acetyl-CoA to generate ATP by utilizing the cofactor FAD. In contrast, peroxisomes produce hydrogen peroxide (H_2O_2) and supplies acetyl-CoA for bile acid and sterol biosynthesis (Schrader and Yoon, 2007). Another example

of a universal peroxisomal function is the detoxification of H_2O_2 that is generated during β-oxidation. Peroxisomes are hosting many reactive oxygen species (ROS) treating enzymes including catalase, which is highly abundant in the peroxisomal matrix (Wanders & Waterham 2006).

Peroxisomal specific metabolic functions are dependent on different factors, such as the organism and type of the cells. In higher eukaryotes, such as humans, peroxisomes catalyze additional functions including ether phospholipid (plasmalogens, for example, myelin sheath lipids) biosynthesis, glyoxylate detoxification, bile acids, fatty acids α-oxidation and cholesterol. Rat liver peroxisomes contain urate oxidase (UO), which is an enzyme involved in purine catabolism (van den Bosch et al., 1992; Wanders and Waterham, 2006). They also play a role in β-lactam antibiotics biosynthesis, such as (penicillin) in filamentous fungi *Penicillium chrysogenum* by hosting the last two enzymes in the pathway (Meijer et al., 2010). In trypanosomes, peroxisomes designated as glycosomes, in which they contain specific glycolytic enzymes (Michels et al., 2005). Moreover, firefly *Photinus pyralis* harbour luciferase in their peroxisomes. Which is an enzyme responsible for the light-producing bioluminescent reaction (Keller et al., 1987).

1.6.4 Peroxisomal disorders

As peroxisomes are responsible for several important metabolic activities, peroxisomal disorders lead to health problems in humans. Till the present, two groups of peroxisomal diseases have been classified (Steinberg et al., 2006). Firstly, the peroxisome biogenesis disorders (PBDs), which are characterized by the complete loss of the peroxisomal function. Generally, PBD occurs as a reflection of a defect in one or more *PEX* genes. Secondly, the single peroxisomal enzyme deficiencies (SPED), which are caused by total loss or reduced activity of a single enzyme (Table 1.2) (Wanders and Waterham, 2004).

The PBDs can be subclassified into three groups; neonatal adrenoleukodystrophy (NALD), Zellweger syndrome (ZS) and infantile refsum disease (IRD). Because of their similar phenotype, those disorders were gathered together forming Zellweger syndrome spectrum (ZSS) (Waterham and Ebberink, 2012). In term of severity, the Zellweger syndrome was considered as the most severe disease of the ZSS, whereas IRD has the

less severity of them (Braverman et al., 2013). Children with ZS generally do not survive more than one year. NALD patient may live until they reach their teens, however patients with IRD may reach adulthood (Waterham and Ebberink, 2012). The first time ZS was linked to peroxisomes was when patient's liver samples were analysed and appeared to lack peroxisomes (Goldfischer et al., 1973). The common clinical signs and symptoms of PBD patients are neurological abnormalities, which appear early in their life after birth (Wanders and Waterham, 2004).

Many of the single enzyme deficiencies occur due to a defect in the β-oxidation pathway. In addition, some other enzymatic functional defects related to glyoxylate detoxification, ether lipid biosynthesis, α-oxidation, and H_2O_2 metabolism have been described. For instance, the common peroxisomal inherited disease is X-linked adrenoleukodystrophy (X-ALD) (Wanders and Waterham, 2006). X-ALD happens as a result of a mutation in *ABCD1* gene which encodes ALDP. As an ABC transporter, this protein is responsible for transport of very long-chain fatty acids (VLCFAs) across the peroxisomal membrane prior to their degradation by β-oxidation. Thus, as a result of this disorder, VLCFAs accumulate which becomes toxic over a life time (Kemp and Wanders, 2007). However, Rhizomelic chondrodysplasia Type 2 (RCDP2) displays a total DHAPAT enzyme deficiency leading to failure in ether phospholipids synthesis. Plasmalogens are distinct ether phospholipids representing 70% of the myelin sheaths. Therefore, as a result of plasmalogens deficiency, patients could develop ischemia, Alzheimer's disease and spinal cord trauma (Farooqui and Horrocks, 2001).

Table 1.2: Human peroxisomal diseases: A) Peroxisomal biogenesis disorder. B) Single peroxisomal enzyme deficiency. Table adapted from (Fujiki et al., 2012; Wanders, 2014; Waterham et al., 2016).

Protein	Disease	Responsible function
A) Peroxisomal biogenesis disorder		
Pex26, Pex19, Pex16, Pex14, Pex13, Pex12, Pex11, Pex10, Pex6, Pex5, Pex3, Pex2 and Pex1	Zellweger Syndrome	Import of peroxisome matrix or membrane protein
Pex26, Pex13, Pex12, Pex10, Pex6, Pex5 and Pex1	Neonatal adrenoleukodystrophy	Import of peroxisome matrix or membrane protein
Pex26, Pex12, Pex6, Pex2 and Pex1	Infantile refsum disease	Import of matrix protein
Pex7	Rhizomelic chondrodysplasia punctata type 1	Import of matrix protein
B) Single peroxisomal enzyme deficiency		
ALDP	X linked adrenoleukodystrophy	β-oxidation
ACOX1	Acyl-CoA oxidase deficiency	β-oxidation
DBP	D-bifunctional protein deficiency	β-oxidation
DHAPAT	Rhizomelic chondrodysplasia Type 2	Ether lipid biosynthesis
ADHAPS	Rhizomelic chondrodysplasia Type 3	Ether lipid biosynthesis
AMACR	2-Methylacyl-CoA racemase deficiency	β-oxidation
SCPx	Sterol carrier protein X deficiency	β-oxidation
PHYH/PAHX	Adult Refsum disease	α-oxidation
AGT	Hyperoxaluria Type 1	Glyoxylate detoxification
CAT	Acatalasaemia	H_2O_2 metabolism

1.6.5 Peroxisomal protein import pathways

All peroxisomal proteins (matrix and membrane) are encoded by the nucleus and synthesised in the cytoplasm in the free ribosomes and then imported post-translationally (Baker and Sparkes, 2005). Peroxisomes show an interesting protein import mechanism that is unique among the eukaryotic organelles. They have two main distinct pathways; one for matrix protein and the other for membrane protein import. Among different organisms, peroxisomal protein import shows huge similarity and some differences as well, of which more detail is given in the sections below.

1.6.5.1 Peroxisomal matrix proteins

Peroxisomal matrix proteins are synthesized in the cytosol on the free polyribosomes (Lazarow and Fujiki, 1985). Then special import systems control the protein translocation from the source to the peroxisome. This mechanism includes three steps; (1) specific receptors bind to distinct regions of the protein called the peroxisomal targeting signal (PTS). There are two types of PTSs (PTS1 and PTS2). After that (2) the complex binds to the peroxisomal membrane. Finally (3) the proteins are translocated across the membrane dependent on ATP hydrolysis (Liu et al., 2012).

Previously two types of the peroxisomal matrix targeting signals have been recognized. The first peroxisomal targeting signal (PTS1) comprises a tripeptide located at the carboxy terminus (C-terminal) of the majority of the peroxisomal matrix proteins. This signal contains a consensus amino acid sequence (S/A/C), (K/H/R) and (L/M) (Gould et al., 1988; Hettema et al., 1999), while the second peroxisomal signal type (PTS2) comprises nine amino acids arranged according to the following consensus sequence R, (L/V/I/Q), XX, (L/V/I/H), (L/S/G/A), X, (H/Q) and (L/A) which may be located at various distances from the amino terminus (N-terminal) end but always near the N-terminus (Petriv et al., 2004).

Pex5 is the cycling receptor for PTS1-containing proteins and also acts as a co-receptor for Pex7 in most organisms except fungi where Pex7 is associated with Pex20 (Lanyon-Hogg et al., 2014). While PTS2-containing proteins are transported by Pex7 and Pex5 complex (Rodrigues et al., 2014). Pex7 delivers cargo to the peroxisomal membrane by

docking onto the Pex13/14 complex. Translocation of cargo is poorly characterised. Receptors subsequently recycle into the cytosol for another round of targeting (Figure 1.4) (Rodrigues et al., 2014).

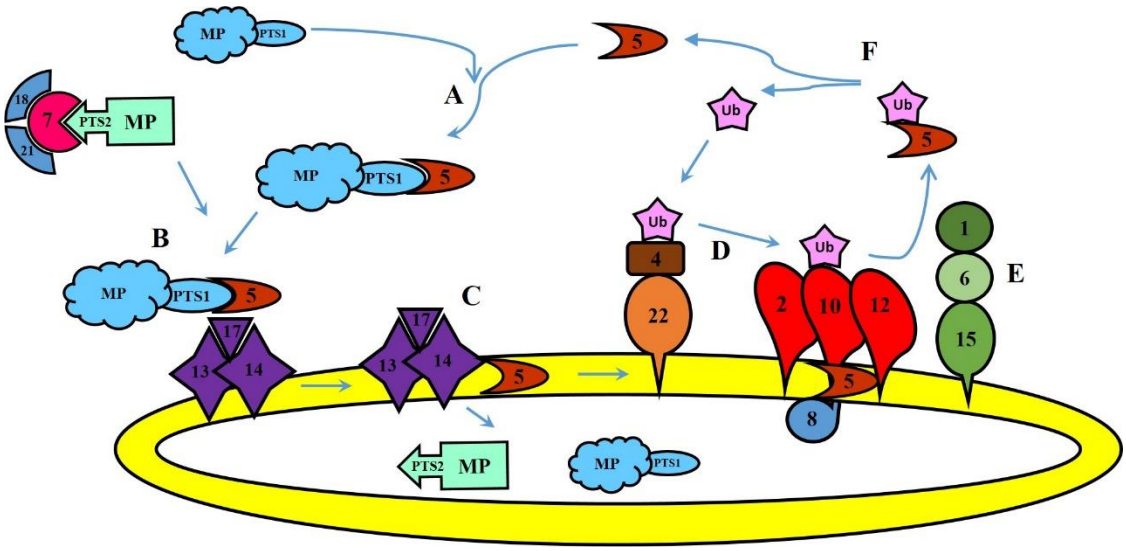


Figure 1.4: Cartoon of peroxisomal matrix proteins import model. Adapted from (Kim and Hettema, 2015) with permission. A) Recognition of newly synthesised PTS containing protein by the receptor Pex5. B) Cargo-loaded receptor docks onto the docking complex (Pex13, Pex14 and Pex17). C) Receptor inserts in the membrane in complex with Pex14 and cargo is released into the lumen. D) Pex2, Pex10 and Pex12 consisted the ubiquitin ligase complex in conjunction with the ubiquitin conjugation complex (Pex4 and Pex22) modifies the receptor (Pex5). E) Ubiquitinated Pex5 is extracted from the membrane by the extraction complex (Pex1, Pex6 and Pex15). F) Then it deubiquitinated for another round of targeting in the cell cytosol.

1.6.5.2 Peroxisomal membrane proteins

Peroxisomal membrane proteins (PMPs) just like the matrix proteins are nuclear encoded, most of them synthesized on free ribosomes in the cytoplasm and post-translationally imported into the peroxisome membrane. However, they require a different machinery for their targeting and integration into the peroxisome membrane.

Even though this mechanism is not well characterised, it is very clear that it is distinct from the peroxisome matrix protein import pathways (Figure 1.5). All the known PMPs lack either a PTS1 or PTS2. However, many PMPs possess a membrane peroxisomal targeting signal (mPTS) which guides them to the peroxisome membrane. In general, mPTSs consist of a cluster of positively charged amino acids or a mixed with hydrophobic residues surrounded by one or two transmembrane region where the pex19 recognize and bind (Dyer et al., 1996; Fransen et al., 2001; Honsho and Fujiki, 2001; Honsho et al., 2002; Murphy et al., 2003; Rottensteiner et al., 2004). The second group of PMPs do not require binding to Pex19 for their transport to peroxisomes (Jones et al., 2004). Depending on their ability to follow the Pex19 pathways, PMPs have been divided into two groups. The targeting signal of Class 1 (mPTS1) is recognized by Pex19 in contrast to those of Class 2 (mPTS2) PMPs (Jones et al., 2001).

1.6.5.3 Import of Class1 PMPs

Initially, newly synthesized class 1 PMPs import requires Pex19 recognition to their mPTS1 in the cytosol. Whereas Pex19 functions both as a chaperone and soluble import receptor for the peroxisome membrane (Hettema et al., 2000). After that, the Pex19-PMP complex is directed to the peroxisome membrane where it binds to Pex3p the docking receptor (Fang et al., 2004; Matsuzono et al., 2006). In mammalian cells, the Pex3-Pex19 complex is required Pex16 as a docking site on the peroxisomal membrane, however, no orthologue of Pex16 has been discovered in *S. cerevisiae* yet (Figure 1.5) (Matsuzaki and Fujiki, 2008).

1.6.5.4 Import of Class 2 PMPs

The mechanism of the class 2 PMP transport is unclear. Instead of targeting directly to the peroxisome, these proteins are proposed to pass through the ER (Hua et al., 2015; Tam et al., 2005). The first description of protein trafficking from ER to peroxisome was in *Yarrowia lipolitica* (Titorenko and Rachubinski, 1998; Titorenko et al., 1997). The only examples have been shown for this type of transport are Pex2p, Pex3p, Pex15p, Pex16p, and Pex22p. Pex16, in mammalian cells, is first directed to the ER were it is co-

translationally inserted into the membrane before it is targeted to peroxisomes (Kim et al., 2006). Other studies show that Pex16p directs to the peroxisome through the ER in *A. thaliana* and *Y. lipolytica* (Karnik and Trelease, 2007; Hua et al., 2015). In *S. cerevisiae*, the class 2 membrane protein Pex3p was shown to traffic from ER to peroxisomes dependent on the receptor Pex19 (Figure 1.5) (Hoepfner et al., 2005).

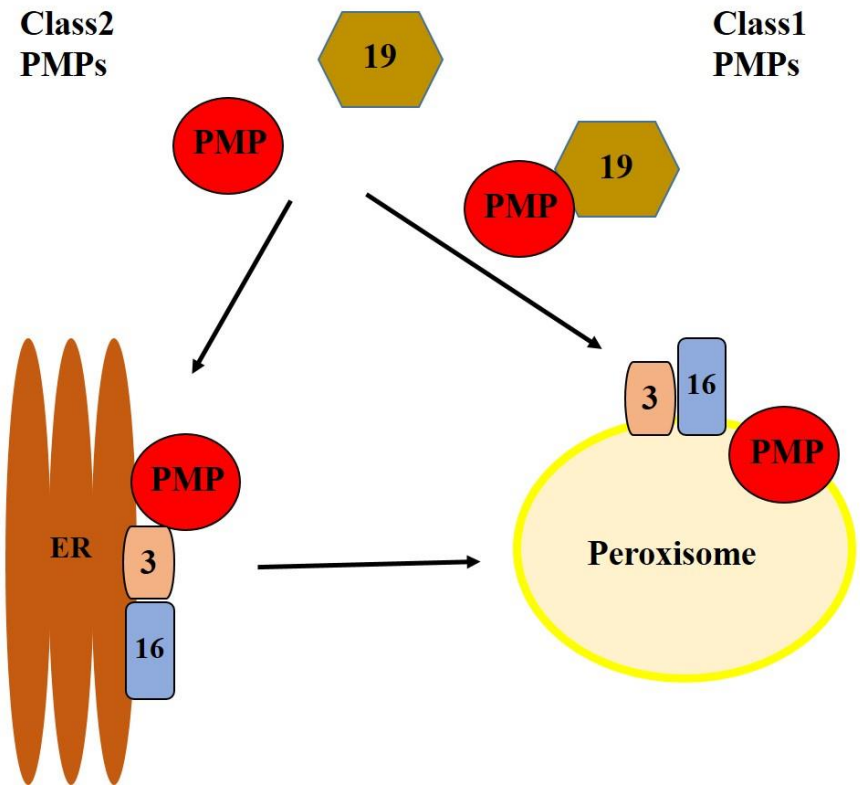


Figure 1.5: Schematic diagram of peroxisomal membrane proteins (PMPs) import. The PMPs proposed to be inserted into the peroxisomal membrane either directly or indirectly through the ER. Pex19 role in this pathway is binding, stabilising and transporting the PMP to the peroxisome. Then the Pex19-PMP complex docks to the membrane via Pex3 or Pex16 in some eukaryotes, and inserted there. Figure adapted from (Hoepfner et al., 2005) with permission.

1.6.6 Peroxisomal membrane proteins topology

As for other organelles, peroxisomes are bounded by a lipid bilayer membrane. The peroxisomal membrane contains both integral and peripheral membrane proteins. Integral peroxisomal membrane proteins are post-translationally inserted into the membrane from the cytosolic side. For example, the Pxa1 protein has 6 transmembrane helices which are inserted into the peroxisomal membrane from cytosolic side by the action of Pex19. Whereas Pex15 possesses a tail anchor domain that also depends on Pex19 for its insertion into the peroxisomal membrane. The only example of the peroxisomal membrane protein following Pex5 pathway is Pex8, which has a common PTS1 (-SKL). However, this protein is peripherally associated with the inside of the peroxisomal membrane. So thus far no proteins have been described that are inserted into the peroxisomal membrane that follow the PTS1 pathway (Kim and Hettema, 2015).

1.7 Aims of the study:

1) To knock out the *PEX7* gene particularly in *D. discoideum* to investigate and compare the effect of the two nitrogen-containing bisphosphonate drugs (alendronate and risedronate) on *D. discoideum* in the absence of *PEX7*. This will put their target, farnesyl diphosphate synthase, into the cytosol and the NBPs will no longer have to cross the peroxisomal membrane to inhibit the enzyme.
2) Introduce the Clustered Regularly Interspaced Short Palindromic Repeats system (CRISPR) into *Dictyostelium* by first constructing the relevant *Dictyostelium CAS9* expression plasmid, using type II CRISPR system. That will be a powerful tool to test candidate genes and understand their roles.
3) Identify the protein import pathway and topology of peroxisomal enzymes required for sterol biosynthesis in *D. discoideum*. In addition, identify sterol pathway partner enzymes such as the electron provider for squalene epoxide formation.

Chapter two “Material and methods”

2.1 Chemicals and enzymes

The majority of chemicals and materials used in this study were supplied by Sigma Aldrich unless stated otherwise. Protein and DNA markers in addition to polymerase chain reaction reagents were obtained from Bioline. Restriction enzymes and buffers were purchased from New England Biolabs (NEB). Plasmid miniprep and gel extraction kits were obtained from Qiagen. Equipment used for biochemical assays were supplied by Bio-Rad.

2.2 Culture media

All media used during this work were dissolved in Millipore water and sterilised by autoclaving at121°C for 20 mins (Table 2.1). Amino acids (adenine, lysine, leucine, histidine, tryptophan and methionine) and uracil were prepared as 100x stocks and also sterilised by autoclaving. They were added at 1x final concentration to sterile yeast minimal medium when needed. Antibiotics were added to media at appropriate concentrations (Table 2.2) if required. Amino acids and antibiotics were added after the media had been autoclaved and cooled to 45-50C. Solid media was supplemented with 2% (w/v) Bacto-agar prior to autoclaving. The sterile media were poured into sterile Petri dishes and allowed to solidify at room temperature and then stored at 4˚C.

Table 2.1: Culture media used during this study.

Culture media	Organism	Description
HL5 glucose medium	*Dictyostelium discoideum*	1.43% bacteriological peptone, 0.715% yeast extract, 1.54% D-glucose, 0.128% Na_2HPO_4 $12H_2O$, 0.0486 % KH_2PO_4
SM agar	*D. discoideum*	1% glucose, 1% proteose peptone, 0.1% yeast extract, 0.1% $MgSO_4$ 7H2O, 0.19% KH_2PO_4, 0.06% K_2HPO_4. Adjusted to pH 6.5
Drosophila medium	S2R+ cells	Schneider's growth medium(Gibco), 10%heat-inactivated fetal calf serum, and 1000U/ml Streptomycin Penicillin solution
2TY	Bacteria	1.6%Bacto tryptone, 1%yeast extract, 0.5%NaCl.
Rich yeast medium YPD	Yeast	1% yeast extract, 2% peptone, 2% glucose
Minimal Medium 1 (YM1).	Yeast	0.17% yeast nitrogen base (without amino acids and ammonium sulphate), 0.5% ammonium sulphate, 2% glucose (or galactose, as required). Adjusted to pH 6.5.
Minimal Medium 2 (YM2).	Yeast	0.17% yeast nitrogen base (without amino acids and ammonium sulphate), 0.5% ammonium sulphate, 1% casamino acids, 2% glucose (or galactose, as required). Adjusted to pH 6.5.
Oleate Media	Yeast	Oleate stock solution contained 12ml oleic acid + 20ml Tween40. 320µl of the stock solution was added to 100ml YM2. 1ml of 10% yeast extract stock was added and any required amino acids.
All percentage concentrations are stated as w/v. For plates, 2% agar was added		

Table 2.2: Antibiotics used in this study

Antibiotic	Final concentration	Organism
Ampicillin	75µg/ml	*E. coli*
Kanamycin	50µg/ml	*E. coli*
G418	200 µg/ml and 15µg/ml	Yeast and *D. discoideum*
Hygromycin B	500µg/ml and 25µg/ml	Yeast and *D. discoideum*
Blasticidin	10 µg/ml	*D. discoideum*

2.3 Escherichia coli procedures

2.3.1 Growth and maintenance of *Escherichia coli*

Escherichia coli was routinely cultured on 2TY agar plates or inoculated into the 2TY medium and shaken at 200rpm, and grown at 37°C overnight. Strains that had been transformed with plasmids containing an ampicillin resistance gene were grown on 2TY containing 75µg/ml ampicillin. Strains were kept at 4°C for use for a maximum of 1 week. For long-term storage, 800µl overnight culture was mixed with the same volume of sterile 30% (v/v) glycerol in a 2ml cryo-tube and stored at -80°C.

2.3.2 Preparing chemically-competent *E. coli* HD5α

The method described by (Hanahan, 1983) was used in this study to prepare competent *E. coli* DH5α. A single colony of the DH5α strain was used to inoculate 5ml 2TY medium, incubated overnight at 37ºC with shaking. A secondary culture was started by inoculating 200ml 2TY medium with an appropriate amount of the primary culture (overnight) to give absorbance 0.05 density at 600. Then the new culture was grown at 30ºC with shaking until the cells reached mid-log phase (OD_{600} = 0.5 -0.6). The growing cells were chilled by putting the culture on ice for 15 mins. After that, the culture was divided into four portions, each one was added to a sterile 50ml falcon tube and centrifuged at 180xg for 10 mins at 4ºC. The supernatants were discarded while the pellets cells were re-suspended in 70ml of ice-cold RF1 solution (100mM rubidium chloride, 50mM manganese chloride, 30mM potassium acetate, 10mM calcium chloride, 15% w/v glycerol, pH 5.8). All tubes were centrifuged at 180xg for 10 mins at 4ºC. The supernatants were discarded, the cells were re-suspended with 16ml ice-cold RF2 solution (10mM MOPS, 10mM Rubidium chloride, 75mM Calcium chloride, 15% w/v Glycerol, pH 6.8). The suspension was divided into 200µl or 400µl samples which were added to sterile Eppendorf tubes. All Eppendorf tubes were put in liquid nitrogen to freeze the cells quickly. Finally, the freeze cells were stored at -80ºC.

2.3.3 Transformation of chemically-competent *E. coli* DH5α

The chemically-competent *E. coli* cells were thawed on ice and100µl cells were added to 1µl of plasmid DNA (100ng/µl) or 10µl of a ligation reaction (section 2.7.5). The mixture was placed on ice for 20 mins after being mixed gently. Samples were heat shocked at 42ºC for 90seconds and then incubated on ice for 2-3minutes. An appropriate amount of 2TY medium (750µl) was added to each sample and the samples were incubated at 37ºC for 45 minutes before being centrifuged in a micro-centrifuge (Sigma 1-14K bench top centrifuge, rotor 12094) at full speed for 1min. The majority of the supernatant (700µl) was removed and the pellet was resuspended in the remaining medium. The cell suspension was spread on a 2TY plate containing ampicillin and incubated at 37°C overnight.

2.3.4 Selection of Colonies

A couple of colonies were selected randomly and used to inoculate 5ml 2TY + Amp broth. All tubes were incubated at 37ºC for overnight. Plasmid DNA was extracted and purified by using a QIAgen miniprep kit.

2.4 *Dictyostelium discoideum* procedures

2.4.1 Growth and maintenance of *D. discoideum*

2.4.1.1 Axenic growth and stock of *D. discoideum*

D. discoideum amoebae strain Ax-2 were grown axenically in HL5 glucose medium at 22°C with shaking at 160rpm (Watts and Ashworth, 1970). Stocks of amoebae were kept in a 22°C incubator for a maximum of 14 days without shaking, before being sub-cultured into a fresh HL5 glucose medium. For long-term storage, amoebae were grown in HL5 glucose medium till they reached the mid-log phase. Amoebae ($\approx 10^8$) were harvested by centrifugation at 4°C for 5 mins at 180xg. The pellet was resuspended in ice-cold freezing medium (50:50 HL5 glucose medium: horse serum) containing 10% (v/v) DMSO, after the supernatant had been removed. The mixture was either aliquoted as 200µl in Eppendorf tubes or as 1ml samples in 1.8ml cryotubes (Nunc). The tubes

were kept at -80°C overnight in a pre-cooled polystyrene container surrounded by isopropanol, before being stored at -80°C.

2.4.1.2 Development growth of *D. discoideum*

D. discoideum was grown on SM agar medium in the presence of *Klebsiella aerogenes* bacteria to test the development of the fruiting body formation. Bacteria (200µl) of an overnight culture were spread on a SM plate. A diluted culture of *D. discoideum* at 200 amoebae per ml was also spread on the plate to give around 20 colonies after the plate had been incubated at 22ºC for 7-10 days.

2.4.2 Isolation of genomic DNA from *D. discoideum* amoebae

Genomic DNA of *D. discoideum* was isolated by using a QIAamp DNA Mini Kit. A 70ml HL5 glucose medium was inoculated with *D. discoideum* cells and incubated at 22°C with shaking at 160 rpm for 3 days. Approximately, 2×10^7 amoebae were harvested by centrifugation at 180xg for 5mins. The supernatant was discarded, and the pellet was washed in 10ml 1xPBS buffer and then centrifuged for 5mins at 180xg. The supernatant was removed before the remaining was resuspended in 200µl 1xPBS buffer and 20µl proteinase K (20mg/ ml) was then added. Following steps were carried out following the manufacturer's instructions.

2.4.3 Isolation of *D. discoideum* cDNA

D. discoideum cDNA was isolated from a cDNA library constructed by Stratagene in λzap by using *D. discoideum* strain Ax-2 mRNA prepared from amoebae harvested during the exponential phase of axenic growth and reverse transcribed into DNA. A QIAamp DNA Mini Kit was used and manufacturer's instructions were followed in extraction cDNA from a 200µl sample of the library.

2.4.4 Transformation of *D. discoideum* by electroporation

D. discoideum amoebae were transformed by electroporation. The method was performed according to Pang et al., (1999). Amoebae (5×10^6) were harvested by centrifugation in a sterile pre-cooled 15ml Falcon tube at 4˚C for 5mins at 180xg. The supernatant was removed before the pellet was washed twice in 8 ml sterile ice-cold H50 medium (20mM HEPES, 10mM NaCl, 5mM $NaHCO_3$, 1mM $MgSO_4$, 1mM NaH_2PO_4, 50mM KCl, pH 7.0) and centrifuged at 4°C for 5mins at 180xg. After the last centrifugation, a sterile pipette was used to remove the last residual supernatant, and the pellet was resuspended in 100µl H50 medium at 0ºC. A pre-cooled 10µl of the appropriate plasmid DNA was added to the resuspended cells and mixed gently. Subsequently, the whole mixture was transferred to a sterile chilled electroporation cuvette (1mm gap). Cells were pulsed twice with an electric current using an electroporator (Bio-Rad Gene Pulser) set at 0.65 kV and 25µFD. After the cuvette had been placed on ice for 5 minutes, the electroporated amoebae were transferred into a sterile Petri dish containing 10ml HL5 glucose medium and incubated overnight at 22°C. Next morning, the appropriate antibiotic was added to the culture and incubated at 22°C for approximately 21 days.

2.4.5 *D. discoideum* harvesting method

The previous medium was changed by a new fresh 10ml HL5 medium. Amoebae cells were scraped by using a sterile spreader until the medium becomes turbid. Subsequently, the culture was translocated to a sterile 100ml flask. Finally, ten microliter blasticidin was added, the flask was incubated at 22ºC without shaking. This culture was prepared to transform with the second CRISPR plasmid.

2.4.6 Fluorescence microscopy of *D. discoideum*

Images were captured on a Zeiss Axiovert 200M microscope (Carl Zeiss MicroImaging, Inc.) equipped with an Exfo X-cite 120 excitation light source, band pass filters (Carl Zeiss MicroImaging, Inc. and Chroma Technology Corp.), an α plan-Fluor 100x/1.45 NA, plan-pochromat 63x /1.40 oil immersion lens, or A-plan 40x/0.65 NA Ph2 objective

lens (Carl Zeiss MicroImaging, Inc.), and a digital CCD camera (Orca ER, Hamamatsu). 525nm and 700nm filters were used for Green Channel (GFP) and Red Channel (RFP and mCherry) respectively. Image acquisition was performed using Volocity software (PerkinElmer). Amoebae in late exponential growth were washed, resuspended in 0.7% (w/v) NaCl and imaged at room temperature. Unless otherwise stated, fluorescence images were collected as 0.5µm Z-stacks, merged into one plane after contrast enhancing in Openlab (Improvision), and processed further in Photoshop (Adobe). Bright field images were collected in one plane and added into the blue channel in Photoshop (Adobe). The level of the bright field images was modified, and the image was blurred and sharpened before further level adjustment.

2.4.7 Structured Illumination Microscopy of *D. discoideum*

The same growth conditions as for Fluorescence microscopy were used to prepare the SIM samples. Live cells were harvested, washed twice then imaged in 0.7% NaCl at room temperature. Special imaging plate type µ-Dish$^{35mm,\ high}$ Glass Bottom (provided by Ibidi Company, Germany) were used. Approximately 3 ml of 2% agarose dissolved in HEPES buffer (pH 7) were added to a new plate and left for 2hs at room temperature to solidify and cool. A thin scalpel was used to pick up one side of the sold agarose and then 15-20µl of the growing sample was added directly to the bottom of the plate under the agar. Finally, the gel was pressed gently, and the plate was covered with its lid. The prepared samples were imaged directly by SIM. This microscope (GE, OMX optical microscope, version 4 with 15 bit sCNOS camera and a laser as a source of light) is based on the use of patterned illumination. Using structured (patterned) illumination, rather than standard uniform illumination, generates an image in which information from the sample and the pattern are combined, bringing small variations in sample fluorescence within the diffraction limit of the microscope. By using multiple different patterns on the same area of sample and with precise knowledge of what those patterns are, software can be used to reconstruct a super-resolution image. Possible resolution: two-fold improvement compared to diffraction limited imaging (wavelength dependent) ~120nm (xy) and ~320nm (z). Images were collected as 0.125µm Z-stacks and processed further in FiJi programme.

2.5 *Saccharomyces cerevisiae* procedures

2.5.1 Growth and maintenance of *S. cerevisiae*

The *S. cerevisiae* strains were grown either in rich yeast medium (YPD) or yeast minimal media (YM1 or YM2) for 2 days at 30°C. Liquid media cultures were shaken at 200rpm. 2% glucose or 2% raffinose was used as a carbon source. After transformation, yeast cells were selected on agar minimal medium lacking the appropriate amino acids. Plates were incubated at 30°C for 2 days and then stored at room temperature for approximately 14 days before yeast colonies were re-streaked onto fresh plates. For long-term storage, yeast strains were grown overnight and then 800µl of each strain was added to 800µl sterile 30% glycerol (final concentration 15%) and left at -80°C.

2.5.2 One step yeast transformation

Yeast cells were routinely transformed with plasmid DNA following the one-step method described by Chen et al. (1992). A sample (200µl) of a yeast overnight culture was harvested by centrifugation in a micro-centrifuge for 1min at full speed. After the supernatant had been removed, the pellet was resuspended in 1µl plasmid DNA (200-400µg/ µl) by vortexing. 50µl one-step buffer (0.2M LiAc pH 5.0, 40% (w/v) polyethylene glycol (PEG) 4000, 100mM DTT) and 5µl (50µg) single stranded carrier DNA (Salmon testes) were added to the cells. The mixture was vortexed and incubated at room temperature. For higher transformation rates, the mixture was left for several hours and vortexed periodically. Subsequently, the cells were heat-shocked at 42°C for 30 minutes. Then they were spread on a selective minimal medium agar plate and incubated for 2 days at 30°C.

2.5.3 High efficiency yeast transformation

High efficiency yeast transformations were used for experiments requiring homologous recombination, for example, the creation of knockout strains and building plasmids. The high efficiency transformations were performed by use of the lithium acetate method (Gietz and Woods, 2002).The yeast strain was grown at 30°C overnight in a 3ml

appropriate liquid medium as a primary culture. Next day, a secondary culture was set up by inoculating fresh liquid media (5ml for each sample) with enough of the primary culture to start growth from OD_{600}= 0.1. The secondary culture was incubated at 30ºC with shaking at 200rpm until the cells reached the mid-log phase (OD_{600} = 0.5-0.6). The cells were harvested by centrifuging for 5mins at 180xg and each supernatant was discarded. Each pellet was resuspended in 1ml sterile water and transferred to a 1.5ml Eppendorf tube and then centrifuged in a micro-centrifuge for 1min at full speed. The pellets were resuspended in 1ml TE/LiAc buffer (100mM LiAc pH 7.5, 10mM Tris-Cl pH 7.5, and 0.1 mM EDTA) after the supernatants had been removed. The cells were centrifuged again (1min at full speed). The supernatant was removed before the pellet was resuspended in TE/LiAc (50µl for each sample). The samples were either used for gene knock out or building plasmids. For the latter, 5µl linearized vector (25µg), 5-10µl PCR product (0.5-1µg), 5µl (50µg) single stranded Salmon testes DNA and 300µl 40% PEG 4000 (40 % (w/v) PEG 4000, 10mM Tris-Cl pH 7.5, 0.1 mM EDTA pH 8.0, and 100mM LiAc pH 7.5) were added. The samples were vortexed, kept at room temperature for 30 minutes and then incubated for 30minute at 30°C. Subsequently, the samples were heat shocked in 42°C for 15 mins. The cells were centrifuged for 1min at 100xg and the supernatants removed. This was followed by resuspending each pellet in 50µl TE buffer (10mM Tris-Cl pH 7.5, 0.1mM EDTA). Finally, the cells were plated onto selective agar media and incubated at 30°C for 2 days. For gene knock out, LiAc treated sample was transformed with 5-10µl PCR product containing the appropriate selectable marker cassette then the same steps were followed.

2.5.4 Isolation of yeast genomic DNA

Plasmids that had been made by homologous recombination in yeast cells were isolated by yeast genomic DNA isolation. All the centrifugation steps in this procedure were carried out a micro-centrifuge at 15,700xg at room temperature.

Yeast cells grown to confluence on a plate were scraped off and washed in 1ml sterile water. The supernatant was discarded out and the pellet was resuspended in 1ml sterile water before being transferred to a 2ml screw-cap tube. The yeast cells were centrifuged for 1min, the supernatant was discarded, and the pellet was resuspended in 50µl sterile

water. 200µl sterile TENTS solution (20mM Tris-HCl pH8.0, 1mM EDTA, 100mM NaCl, 2% (v/v) Triton X-100, 1% (w/v) SDS), 200µl acid washed glass beads and 200µl (25:24:1) phenol:chloroform:isoamyl alcohol were added to each sample. The samples were placed in a mini bead beater (Biospec Products) and the cells were disrupted at full speed for 45 seconds. The samples then were centrifuged for 1min to pellet the beads and another 200µl of TENTS was added. The samples were vortexed and subsequently centrifuged for 5mins. The resulting aqueous phase (≈350µl) was transferred to new Eppendorf tube. 200µl (25:24:1) phenol:chloroform:isoamyl alcohol were added. The samples were vortexed thoroughly and centrifuged as previously. 300µl of the aqueous phase was transferred to a new 1.5ml Eppendorf tube and the nucleic acids were precipitated by adding 30µl 3M sodium acetate pH5.2 and 750µl 100% ethanol. The mixtures were incubated for 1hour at -20°C, followed by centrifugation for 15 mins. The pellets were washed in 500µl 70% (v/v) ethanol after the supernatant had been removed. The pellets were subsequently resuspended in 200µl 1xTE (10mM Tris-Cl pH 7.5, 0.1mM EDTA) containing RNAse (5µg/ml) and incubated at room temperature for 10 minutes. DNA was precipitated by adding 20µl 3M sodium acetate, pH 5.2 and 500µl 100% ethanol, which was followed by incubation for 30 mins at -20°C. The samples were centrifuged for 15 mins. The supernatants were discarded after the pellets, which had the precipitated DNA, were washed in 200µl 70% (v/v) ethanol and then dried by incubation for 15 mins at 50°C. Finally, the pellet was resuspended in a final volume of 50µl 1xTE.

2.5.5 Preparation of spheroplasts from yeast cells

A 50ml yeast culture was grown to log phase (OD_{600} = 1.0). The cells were harvested by centrifugation in a Sigma 4-16KS centrifuge for 5mins at 180xg at room temperature and the supernatant was removed. The pellet was resuspended in freshly prepared 10ml 10mM Tris-Cl pH 9.4, 10mM DTT, transferred to a 50ml Falcon tube, and incubated for 15mins at room temperature on a roller, before being centrifuged again as previously described. The supernatant was removed, and the pellet resuspended in 5ml 1.2M sorbitol, 50mM potassium phosphate pH7.4, zymolyase (zymolyase added as 5mg/g harvested cells). Following a 45 mins incubation at room temperature on a roller, cells were checked for spheroplast formation by light microscopy (spheroplasts appear more

spherical than untreated yeast cells and form clusters. They also easily lyse upon addition of drops of water to the microscope sample). All the next steps were carried out at 0ºC (cold conditions). The sample was centrifuged for 4 mins in a Sigma 4-16KS centrifuge as described above. The spheroplasts were then washed twice with 10 ml 1.2 M sorbitol, 50 mM potassium phosphate pH 7.4 and re-harvested. After the final wash, the supernatant was removed, and the pellet was either flash-frozen in liquid nitrogen and then stored at -80°C or used immediately.

2.5.6 Yeast knockout construction by homologous recombination

Yeast genes deletion were achieved by integrating a PCR product cassette containing an auxotrophic marker or antibiotic resistance gene. The PCR product was containing 50bp of homology to the upstream and downstream of the gene of interest in the yeast genome. This to allowed replacement of the gene of interest via homologous recombination (Figure 2.1) (Longtine et al., 1998). Yeast cells that have genome integrated cassette were selected by grown on selective media.

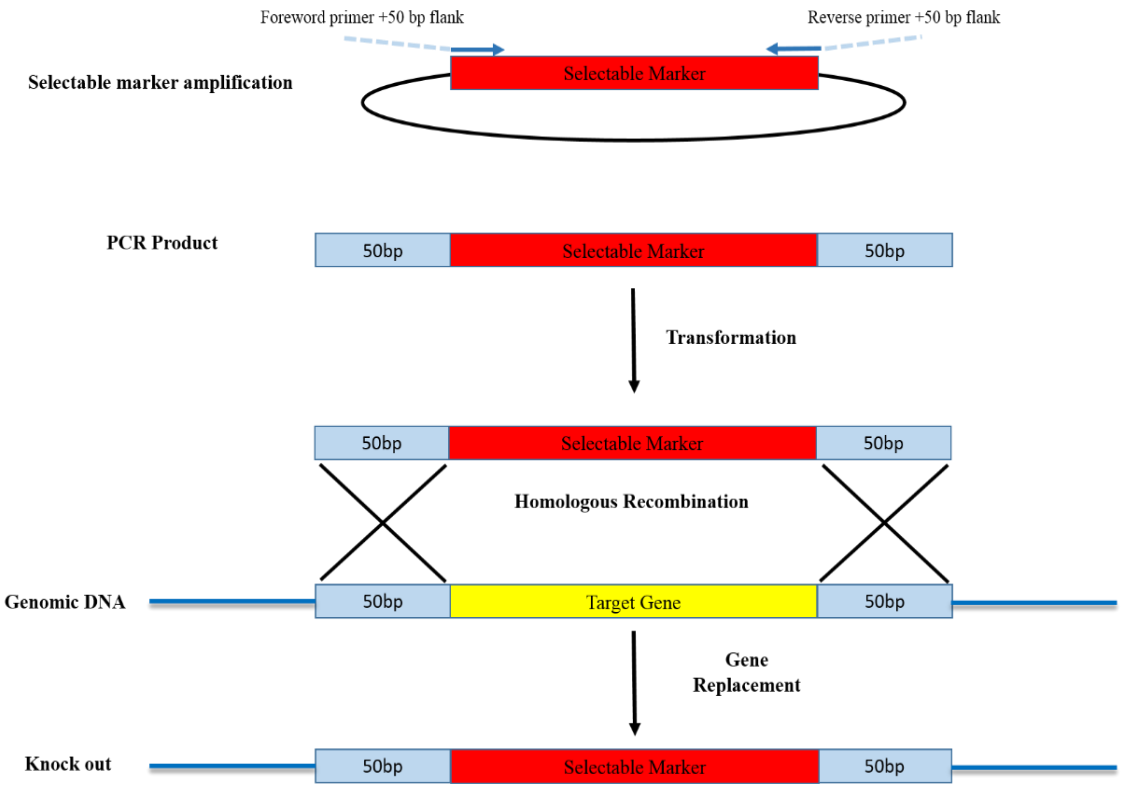


Figure 2.1: Schematic diagram of the construction of yeast gene knockouts by the homologous recombination method. Blue arrows refer to the forward and reverse primers which started with the interrupted light blue line (50bp) as a homolog sequences located upstream and downstream of the gene of interest in the yeast gDNA respectively. The yellow box is the gene of interest, which will be replaced by the marker (red box) as a result of the gene knock out.

2.5.7 *Saccharomyces cerevisiae* strains

Several *S. cerevisiae* strains were used in this study. A list of these strain is given in Table 2.3). Strains that were made during this study were derived from the wild type (WT) BY4742 strain. WT yeast was used for plasmid construction, protein expression for microscopy and as a source of gDNA. Gene deletions were obtained by homologous recombination.

Table 2.3: *S. cerevisiae* yeast strains used in this study.

Strain	Genotype	Source
WT BY4742	*MATα, his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Euroscarf
*Pex1*Δ	BY4742, *Pex1*Δ	Euroscarf
Pex3	BY4742, *Pex3*Δ	Euroscarf
*Pex5*Δ	BY4742, *Pex5::kanMX*	This study
*Pex7*Δ	BY4742, *Pex7::kanMX*	This study
*Pex13*Δ	BY4742, *Pex13*Δ	Euroscarf
*Pex14*Δ	BY4742, *Pex14*Δ	Euroscarf
*Pex19*Δ	BY4742, *Pex19*Δ	Euroscarf
*vps1*Δ*/fis1*Δ	BY4742 *vps1*Δ and *fis1*Δ*::HIS5*	Lab stock
BY25598 (YEH 315)	*MATa ura3-1::ADH1-OsTIR1-9Myc(URA3) ade2-1 his3-11,15 leu2-3,112 trp1-1 can1-100*	Lab stock
YEH 381	BY25598 Δ*Pex1::his*	Lab stock
YEH 404	BY25598 *HIS3p-AID-GFP-PTS1::trp*	Lab stock
YEH 405	YEH 381 *HIS3p-AID-GFP-PTS1::trp*	Lab stock

Euroscarf reference is Johann Wolfgang Goethe University, Frankfurt, Germany

2.5.8 Fluorescent microscopy of fusion proteins expressed in yeast

Routinely, yeast cultures were grown in YM1 or YM2 medium containing the appropriated amino acids and a suitable source of sugar and harvested in mid-log phase. Images were captured on a Zeiss Axiovert 200M microscope (section 2.4.6). Fluorescence images were collected as 0.5µm Z-stacks. Every single Z-stack image was further processed and final figures were prepared using Openlab software and Photoshop programme.

2.5.9 Pulse-chase experiments

Pulse-chase experiments were used in order to identify the molecular localization and to control the expression of the fluorescently tagged protein (expressed under GAL1 promoter control) with time (Motley and Hettema, 2007). Yeast cells were grown overnight in a selective glucose minimal medium before being transferred to 5ml selective galactose minimal medium at an OD_{600} of 0.1 to start induction. Protein expression was induced for 3-5 hours (pulse). Subsequently, expression was shut down by switching the culture to selective glucose minimal medium (chase) for 1 h. Samples were either used for live cell imaging or for biochemical experiments.

2.6 *Drosophila melanogaster* cell line S2R+ procedures

2.6.1 S2R+ cells growth and maintenance

Drosophila S2R+ cells were grown as a flat adherent monolayer and incubated at 25°C. T-25cm2 flasks containing 5ml *Drosophila* growth medium were used to grow the S2R+ cells. Cell growth was observed every day using light microscopy (Nikon Eclipse TS100 microscope). Cells were sub-cultured when growth was approximately 80% confluent, which occurred after approximately 5 days. During each subculturing, 1ml of resuspended cells was transformed to a new T-25cm2 flask containing 4ml of fresh growth medium.

2.6.2 Transfection of S2R+ cells

Drosophila S2R+ cells were added to give ~70% confluence in each well of a 6-well plate containing 2ml *Drosophila* growth medium and a coverslip. The cells were briefly transfected with plasmid DNA using the Effectene kit (provided by Qiagen) and following the manufacturer's instructions. For each transfection, 5µl of the appropriate plasmid DNA at 100ng/µl, (when transfecting with two plasmids, 2.5µl at 100ng/µl of each plasmid), 9µl of EC buffer, and 8µl of enhancer (provided in the kit) were added to 1.5ml Eppendorf tube, vortexed for 2 secs and left for 5 mins at RT. Afterwards, 10µl of Effectene were added to the mixture, vortexed for 10 secs and left for 10 mins at RT. Finally, 600µl *Drosophila* growth media were added. The solutions were mixed by

pipetting and added dropwise onto cell culture. The cells were incubated for 24 hours at 25°C. Subsequently, the growth medium was aspirated off and 2ml growth medium containing copper sulphate at a final concentration of 50µM were added gently. The cells were incubated for 16 hours at 25°C, fixed and observed by fluorescence microscopy.

2.6.3 Fixing S2R+ cells for fluorescence microscopy

Drosophila S2R+ cells were fixed with 3.7% (v/v) formaldehyde in 1xPBS for 15mins before being rinsed three times in 1xPBS. Cells were left for 10mins in 100mM ammonium chloride 0.1% Triton X-100. The coverslips were mounted onto slides in a drop of mounting medium and left for 1-2hs at RT to set. Images were taken and processed as same as *D. discoideum*.

2.7 Molecular cloning

2.7.1 Oligonucleotide primers

All primers used in this study were provided by Sigma (Table 2.4). The coding sequences of *D. discoideum* genes were obtained from the Dictybase website. Primers were designed to be around 18-30 base pairs in length, although in some cases, they were longer. For gene amplification, primers were annealed at both ends of the chosen ORF and contained unique sequences. Primers with a GC content of 45-55% were desired. Furthermore, all primers contained either a G or C nucleotide at the 5' end and two of them at the 3' end to control miss-priming. Each forward and reverse primers was contained a proper restriction site to allowed digestion and cloning. To allow appropriate annealing of primers, each set of primers were designed to have similar melting temperatures (Tm), close to 62°C as much as possible. Calculation of the TM was carried out by using the following formula.

$$TM= (A+T)\ 2°C + (C+G)\ 4°C$$

For amplification and expression of all *D. discoideum* ORF, a multi As (AAAAA) Kozak translation initiation sequence was added just upstream the ATG start codon to enhance the expression.

2.7.1.1 Primer design for homologous recombination

Primers were designed for plasmid cloning or gene deletion by homologous recombination. The 5' ends of the primers matched at least 20 nucleotides flanking cloning or50 nucleotides for gene deletion.

2.7.1.2 Primer design for northern blot

A 20 base oligonucleotide having a unique sequence was designed to be used as a probe to detect the presence of specific RNA molecule (guide RNA) on Northern blots. The probe had been radio-labelled with (2.12.1).

2.7.1.3 Primer design for single stranded guide RNA (sgRNA)

A 20bp oligonucleotide with a specific sequence of the target gene just upstream a PAM region (NGG) in the genome was chosen to be used as guide RNA in CRISPR system, either manually or web site. The complementary overhang was added for ligation (section 2.7.9). The forward primers prefer to have a G nucleotide at the 5' end to reduce the off-target.

Table 2.4: Oligonucleotides used in this study

Name	Use	Sequences 5′ ⟶ 3′
VIP2226	Amplify hCas9 (F)	GACAAGCTTAAAAAATGGACTATAAGG ACCACGACG
VIP2227	Amplify hCas9 (R)	GACGGATCCTTACTTTTTCTTTTTTGCCTGG
VIP2293	Target A of pex7 KO (F) CRISPR	AATTGTGATATGTCAGTTATTGTA
VIP2294	Target A of pex7 KO (R) CRISPR	AAACTACAATAACTGACATATCAC

VIP2295	Target B of pex7 KO (F) CRISPR	AATTGTATAAAAATATGGAATCCA
VIP2297	Target B of pex7 KO (R) CRISPR	AAACTGGATTCCATATTTTTATAC
VIP2299	Oligo to insert RS in pMUR3=> pMUR10 (F)	AGCTTGTCGACTCTAGACTGCAGAT
VIP2300	Oligo to insert RS in pMUR3=> pMUR10 (R)	CGATCTGCAGTCTAGAGTCGACA
VIP2292	Seq. pMUR4 and the CRISPR inset (R)	CTCTAGATTAATTTAAATCTTCTTC
VIP2301	Amplify *wtCas9* +Flag +RS (F)	GACGTCGACAAAATGGATTATAAAGAT GATGATGATAAATCTAGAGATAAGAAA TACTCAATAGGCTTAG
VIP2329	Amplify *wtCas9*+NLS+ stop codon + PstI RS (R)	GACCTGCAGTTAACCAACTTTTCTTTTTTT TTTTGGGTCACCTCCTAGCTGACTC
VIP2303	hCas9 Sequencing R	CTGTTGTCGGGGTTCAGGTCG
VIP2304	hCas9 Sequencing F	GGACGAGGTGGCCTACCACG
VIP2305	hCas9 Sequencing F	GAGATACGACGAGCACCACC
VIP2306	hCas9 Sequencing F	GAGTACTTCACCGTGTATAACG
VIP2307	hCas9 Sequencing F	GAAGGTGGTGGACGAGCTCG
VIP2308	hCas9 Sequencing F	GGTGTCCGATTTCCGGAAGG
VIP2309	hCas9 Sequencing F	CTGAAGAGTGTGAAAGAGCTGC
VIP2310	C.PCR *hCas9* R	GATCTGATATCATCGATGAATTCG
VIP2311	wtCas9 Sequencing R	CAATGACAAAGCAATGAGATTCC
VIP2312	wtCas9 Sequencing F	GGACAAACTATTTATCCAGTTGG
VIP2313	wtCas9 Sequencing F	CGTGAAGATTTGCTGCGCAAGC
VIP2314	wtCas9 Sequencing F	GAAGATAGGGAGATGATTGAGG
VIP2315	wtCas9 Sequencing F	CTATTGGAGACAACTTCTAAACG
VIP2316	wtCas9 Sequencing F	GACAGGCGGATTCTCCAAGG
VIP2317	C.PCR wtCas9	GATGGATTGATGGATAAGAGTGG
VIP2326	Seq. *Dd*CRISPR p1 and the inset (F)	GGTTTATACATATTTATGTTCGTACTGAAG
VIP2328	Seq. *Dd*CRISPR p1 and the inset (R)	GCTCCAGACTCTCAGGCAATGACC
VIP2408	Target A of GFP KO(F) CRISPR	AATTGAATTAGATGGTGATGTTAA
VIP2409	Target A of GFP KO(R) CRISPR	AAACTTAACATCACCATCTAATTC
VIP2410	Target B of GFP KO(F) CRISPR	AATTGATACCCAGATCATATGAAA
VIP2411	Target B of GFP KO (R) CRISPR	AAACTTTCATATGATCTGGGTATC
VIP2412	Target C of GFP KO (F) CRISPR	AATTGTTAATAGAATCGAGTTAAA
VIP2413	Target C of GFP KO (R) CRISPR	AAACTTTAACTCGATTCTATTAAC
VIP2542	Seq of left flank of pex7 KO by HR (R)	CTT CGT ATA ATG TAT GCT ATA CG

VIP2554	Checking the KO cassette insertion in gDNA (F)	GCAATTCTTTTATCAGAATAAAAATCTTC
VIP2555	On Blasticidin (R) PCR	CGA ATT GCC GCT CCC ACA TG
VIP2556	On Blasticidin (F) PCR	CGA TTG AAG AAC TCA TTC CAC TC
VIP2557	Seq of right flank of pex7 KO by HR (R)	CTTCTTCATCATCTTCATCATCTTG
VIP2533	Amplify *Dd*SQSΔH +SKL (R)	GCATCTAGATTATAATTTCGAATCATTGAAAC CAGAAGGTCTATAACC
VIP2534	Amplify *Dd*SQSΔHΔSKL (R)	GCATCTAGATTAATCATTGAAACCAGAAGGTC TATAACC
VIP1892	Amplify *Dd*SQS+L (F)	GACGGATCCAGCACCAGCACCATGCAATATAT GAAATCACTTGCTC
VIP2563	Amplify *Dd*SQS to be inserted in yeast by pAS113 (F)	GGCATGGATGAACTATACAAAGGATCCGCTGG TGCTGGTATGCAATATATGAAATCACTTGCTC
VIP2564	Amplify *Dd*SQS to be inserted in yeast by pAS113 (R)	AAAAAATTGATCTATCGATAAGCTTTTATAAT TTCGAAAAAAAGTTTGGACC
VIP2565	Amplify *Dd*SQE to be inserted in yeast by pAS113 (F)	GGCATGGATGAACTATACAAAGGATCCGCTGG TGCTGGTATGGAAGATATTCAATTTGAAAATG
VIP2566	Amplify *Dd*SQE to be inserted in yeast by pAS113 (R)	AAAAAATTGATCTATCGATAAGCTTTTATTTTG TTAATCTTAAAATAATACA
VIP2567	Amplify *Dd*OSC to be inserted in yeast by pAS113 (F)	GGCATGGATGAACTATACAAAGGATCCGCTGG TGCTGGTATGACAACTACTAATTGGAGT
VIP2568	Amplify *Dd*OSC to be inserted in yeast by pAS113 (R)	AAAAAATTGATCTATCGATAAGCTTTTAAATT TTAGATTTTAAATATAATTGATTATATC
VIP2569	Amplify *Dd*SMT to be inserted in yeast by pAS113 (F)	GGCATGGATGAACTATACAAAGGATCCGCTGG TGCTGGTATGGTATTAATTCAAACTAATCATC
VIP2570	Amplify *Dd*SMT to be inserted in yeast by pAS113 (R)	AAAAAATTGATCTATCGATAAGCTTTTAAAGT TTAGCAGCATTTGGTTTTG
VIP478	Seq on PGK ter (R)	GATAAATAATAGTCTATATATACG
VIP464	Seq on GFP (F)	CATTGAAGATGGAGGCGTTC
VIP2595	Seq *Dd*SQE (F)	GAGCATCTTCATGTGAAAATG
VIP2596	Seq. *Dd*SMT (F)	GGTGCAACAGGTATTCC
VIP2597	Seq. *Dd*SMT (F)	CTTGACATCTCTCAAGTACC
VIP2598	Seq. *Dd*SMT (F)	CTGGAATGCCTTCCTCG
VIP2639	NB Prob gRNA	CTTTTTCAAGTTGATAACGG
VIP2563 pMUR30	Amplify *Dd*SQSΔH to be inserted in yeast by pAS113 (F)	GGCATGGATGAACTATACAAAGGATCCGCTGG TGCTGGTATGCAATATATGAAATCACTTGCTC
VIP2650 pMUR30	Amplify *Dd*SQSΔH to be inserted in yeast by pAS113 (R)	AAAAAATTGATCTATCGATAAGCTTTTATAAT TTCGAATCATTGAAACCAGAAGGTCTATAACC

VIP2651 pMUR31	Amplify *Dd*SQSΔPTS1 to be inserted in yeast by pAS113 (R)	AAAAAATTGATCTATCGATAAGCTTTTAAAAA AAGTTTGGACCATGACG
VIP2652 pMUR32	Amplify *Dd*SQSΔHΔPTS1 to be inserted in yeast by pAS113 (R)	AAAAAATTGATCTATCGATAAGCTTTTAATCA TTGAAACCAGAAGGTCTATAACC
VIP2653 pMUR29	Amplify *Dd*SQE to be inserted in yeast by pAS52 (*Dd*SQE-GFP) (F)	CGTCAAGGAGAAAAAACTATAGAGCTCATGG AAGATATTCAATTTGAAAATG
VIP2654 pMUR29	Amplify *Dd*SQE to be inserted in yeast by pAS52 (*Dd*SQE-GFP) (R)	CTGCAGGTCGACTCTAGAGGATCCTTTTGTTA ATCTTAAAATAATACA
VIP2720	Amplify U6 Pro from *Dd* gDNA. 450bp +BglII (F)	GACAGATCTTAACATGGCAATTTTGGGACACA TATG
VIP2721	Amplify U6 Pro from *Dd* gDNA 450bp+BbsI+ KpnI (R)	CTCGGTACCCGAAGACCCAATTAGTTAATTGT TTATTTTCCTTTTATGTTTTTTATTTTATTATTT TTTTTTTTTTATGG
VIP2746	Seq *Dd*Cas9 (R)	CAGGATTTAAATCACCTTCG
VIP2747	Seq *Dd*Cas9 (F)	GTTGCTTACCATGAAAAATATCC
VIP2748	Seq *Dd*Cas9 (F)	CATCAAGATTTAACTTTGTTAAAAGC
VIP2749	Seq *Dd*Cas9 (F)	CGAAGGAATGAGAAAACCAGC
VIP2750	Seq *Dd*Cas9 (F)	GATGAATTGGTAAAAGTTATGGG
VIP2751	Seq *Dd*Cas9 (F)	GAATGATAAGTTAATTCGTGAAG
VIP2752	Seq *Dd*Cas9 (F)	GTAGTTGCTAAGGTAGAAAAAGG
VIP2871	Amplify *Dd*SQSΔPTS1 EcoRI (R)	GACGAATTCTTAAAAAAAGTTTGGACCATGAC G
VIP2872	Amplify *Dd*SQSΔ1-150aa BamHI (F)	GACGGATCCGCAAATGGTATGTCAGAGTTTCT CC
VIP2873	Amplify *Dd*SQSΔ1-251aa BamHI (F)	GACGGATCCCAAATTCCAGCACTCCACTGTC
VIP2874	Amplify *Dd*SQSΔ1-312aa BamHI (F)	GACGGATCCGTAGTCAAAATTCGTAAGGGTC
VIP2875	Amplify *Dd*SQSΔ1-351aa BamHI (F)	GACGGATCCCCACCAAATGATCCATCTGC
VIP2887	Amplify GFP for NLS (F)	GAGTCTAGAATGAGTAAAGGAGAAGAACTTTT CAC
VIP2888	Amplify GFP for NLS (R)	GACGGATCCTTTGTATAGTTCATCCATGCC
VIP2889	Seq GFP (R)	GCAGATTGTGTGGACAGG
VIP2898	Amplify GFP for NLS4N (pMAA-2) (F)	GACAAGCTTAAAATGGCTCCAAAAAAAAAAA GAAAAGTAGGTGGATCCatgagtaaaggagaagaacttttc
VIP2899	Amplify GFP for NLS4N (pMAA-2) (R)	GACGGTACCttaCTCGAGtttgtatagttcatccatgcc
VIP2939	Amplify *Dd*cas9 to form *Dd*Cas9-GFP (F)	GACAAGCTTAAAATGGATTATAAAGATCATG
VIP2940	Amplify *Dd*cas9 to form *Dd*Cas9-GFP (R)	GACGAGCTCACCAGCACCAGCTTTTTTTTTCTT TGCTTGACCAGCTTTC
VIP2958	Amplify *Dd*RedA, XhoI (F)	CAGCTCGAGAAAAAATGAAATCAGTAATTTTA AAACCAAAAAATTTAG

VIP2959	Amplify *Dd*-L-RedA, SacI (R)	CAGGAGCTCACCAGCACCAGCAAACCAAACA TCTTTTTGATAACG
VIP2960	Amplify *Dd*RedB, XhoI (F)	CAGCTCGAGAAAAAATGGAAATTTTAGAATCA ATTGATTTTATTG
VIP2961	Amplify *Dd*-L-RedB BamHI, (R)	CAGGGATCCACCAGCACCAGCAAACCAAACA TCTTGTAAATATCTACC
VIP2962	Seq1 *Dd*RedA (R)	CACGTTTTTCAACTCCACCC
VIP2963	Seq2 *Dd*RedA (F)	CAAGAAGATTGGACTTGTTTCG
VIP2964	Seq3 *Dd*RedA (F)	CCACGTTACTATAGTATTGC
VIP2965	Seq1 *Dd*RedB (R)	GTTACCATTCTAAATCTTGGTACG
VIP2966	Seq2 *Dd*RedB (F)	CCTGTAGTTTGTAAATTCTTAGG
VIP2967	Seq3 *Dd*RedB (F)	GAATTTACACCACGTCTACC
VIP2976	Amplify *Dd*RedA, XhoI (F)	CAGCTCGAGAAAATGAAATCAGTAATTTTAAA ACCAAAAAATTTAGTATTATTAGGTGCTGGTG TAACAACAAC
VIP2977	Amplify *Dd*-L-RedA SacI (R)	CAGGAGCTCACCAGCACCAGCAAACCAAACA TCTTTTTGATAACGTTTTTCTTTTTCTAATTTGT GAAGTAAAGCTTGAGC
VIP2978	Amplify *Dd*RedC, XhoI (F)	CAGCTCGAGAAAATGAATAAAAAATGTACAA TTATTTATGC
VIP2979	Amplify *Dd*-L-RedC SacI, (R)	CAGGAGCTCACCAGCACCAGCCCAAGTTTCAG TAATAAACCTTTTTTC
VIP2996	Seq *Dd*RedA (R)	CCAAGAGTTTACTAGAGTATGG
VIP2997	Seq *Dd*RedB (R)	GGTTTTAATGGAGTTGAAACG
VIP2603	Seq on Mcherry (F)	CACCATCGTGGAACAGTACG
VIP418	Seq. on Gal Pro (F)	GTATTACTTCTTATTCAAATG
VIP3029	Target A of Atg1 KO (F) CRISPR	AATTGCATTTGCTCAAGTTTATAA
VIP3030	Target A of Atg1 KO (R) CRISPR	AAACTTATAAACTTGAGCAAATGC
VIP3031	Target B of Atg1 KO (F) CRISPR	AATTGACAAAAACAGATTGTTCAT
VIP3032	Target B of Atg1 KO (R) CRISPR	AAACATGAACAATCTGTTTTTGTC
VIP3033	Target C of Atg1 KO (F) CRISPR	AATTGATTCATATAAAAATAATAG
VIP3034	Target C of Atg1 KO (R) CRISPR	AAACCTATTATTTTTATATGAATC
VIP1594	Amplify the LF of Pex7 (gDNA) KO by HR (F)	GACGGTACCGGAAGATAAACTCTTTCTTCTAT TTTG
VIP1595	Amplify the LF of Pex7 (gDNA) KO by HR (R)	GACAAGCTTGCATGTCAAAATTTCATGATCAT GTG
VIP1596	Amplify the RF of Pex7 (gDNA) KO by HR (F)	GACCTGCAGGATTGGAATAAATATAATGATAA GGAG
VIP1597	Amplify the RF of Pex7 (gDNA) KO by HR (R)	GACGGATCCCCTCTATATCATAAATTTTCTTCA TCC
VIP3068	Amplify the *Dd*RedC (long primer) (F)	CAGCTCGAGAAAAAATGAATAAAAAATGTAC AATTATTTATGCAACTGAATCAGGTACATCAC AAGAGGTAGCAGAAAAGTTATCACG

VIP49	M13 (F) Yeast	GTTTTCCCAGTCACGACG
VIP50	M13 (R) Yeast	GGAAACAGCTATGACCATG
VIP3134	Amplify *Dd*SQS 6aaH7aaPTS1 (F)	GACGGATCCCCTTCTGGTTTCAATGATTTCATC
VIP3135	To form *Dd*SQS24aa (F)	GATCCGCTGTTACTTCTTTAGCTGTTTCTTCTG CTTTCTTAATTGCTCGTCATGGTCCAAACTTTT TTTCGAAATTATAAA
VIP3136	To form *Dd*SQS24aa (R)	AGCTTTTATAATTTCGAAAAAAAGTTTGGACC ATGACGAGCAATTAAGAAAGCAGAAGAAACA GCTAAAGAAGTAACAGCG
VIP3137	Amplify the first 187aa of *Dd*SMT (*Dd*)	GCATCTAGATTAGGTGTTGTAGTTTTCGACAC G
VIP3138	Amplify *Dd*SMTΔ1-187aa (*Dd*)	GACGGATCCAGCACCAGCACCATGGTCAGTGA TTACTATGATATCG
VIP1460	Seq *Dd*SMT (F)	GTCACTTTCTTTGAATCCAC
VIP3227	Seq *Dd*SQS (R)	GATTCATCATCTGCTGTTGC
VIP3449	Amplify the first 187aa of *Dd*SMT (Y) F	GACGAGCTCAAAATGGTATTAATTCAAACTAA TCATC
VIP3450	Amplify the first 187aa of *Dd*SMT (Y) R	GACGGATCCGGTGTTGTAGTTTTCGACACG

F, forward; R, reverse; L, peptide linker; LF, left flank; RL, right flank; RE, restriction enzyme; Dd, Dictyostelium discoideum; Y, yeast.

2.7.2 Polymerase chain reaction (PCR)

Routinely, the PCR was used to amplify a specific DNA region. Three different types of DNA polymerase were used. The Taq DNA polymerase, which does not have a 3′-5′ proofreading ability, was used for colony PCR only. The Velocity polymerase provides high fidelity was used for most of the PCR in this study. Each individual PCR reaction consists of 4µl 5µM forward primer, 4µl 5µM reverse primer (10µl 5µM for Velocity), 1.5mM $MgCl_2$, 5µl 2.5mM dNTP mixture containing dATP, dTTP, dCTP, dGTP (3µl for Taq), 1µl 1/50 diluted plasmid miniprep (or 1µl of gDNA) was used as a DNA template. Individual colonies were used for the colony PCR. 10µl 5x velocity/ Taq buffer and 0.2µl 5u/µl or 1µl 2u/µl of Taq or Velocity DNA polymerase respectively. The final volume was made up to 50µl with sterile water. Subsequently, the PCR reaction was placed in a thermocycler (Biometra) and the following program was run (Table 2.5), after the lid had preheated to 100°C.

Table 2.5: Polymerase chain reaction programme

Step	Definition	Temperature	Time	Cycles
1	Initial denaturation	98°C	2-5minutes	1
2	Denaturation	98°C	30 seconds	30-35
3	Annealing	56°C	30 seconds	
4	Extension	72°C	30 seconds- 1min/1Kb DNA	
5	Final extension	72°C	10 minutes	1

For amplifying *D. discoideum* genomic DNA that contains A/T rich regions, MiFi polymerase which has 3′-5′ proofreading ability and provides high fidelity was used. A modified PCR program was run. The initial denaturation and the denaturation temperatures were reduced to 95°C for 3mins. The annealing and extension temperature was raised to 60°C and 62°C for 30sec and 3mins respectively. The thermocycler was set up for 30 cycles. The PCR reaction mixture was prepared as described in (Table 2.6)

Table 2.6: MyFi PCR reaction compounds

Reagent	Concentration	Amount
MyFi buffer	5x	10µl
Forward primer	5µM	4µl
Reverse primer	5µM	4µl
DNA template (gDNA)		5µl
Polymerase enzyme	1u/µl	2µl
H_2O	-	25 µl
Final volume 50 µl per reaction		

2.7.3 Agarose gel electrophoresis

Amplified DNA, as well as digested DNA, were analysed by agarose gel electrophoresis and size-fractionated. An appropriate amount of agarose was dissolved in TBE buffer (1mM Tris-base, 0.1M Boric acid, 1mM EDTA, pH 8.0) to form 0.5%-1% (w/v) gels depending on the size of the DNA fragments. Ethidium bromide was added at 0.5µg/ml final concentration. DNA loading buffer (5x loading buffer, NEB) was added at 1x final concentration before samples were loaded onto the agarose gel. To estimate the size of the DNA fragments, a DNA 1Kb ladder (Hyperladder from Bioline) was loaded alongside the samples. The gel was run in 1xTBE running buffer at constant voltage

(100V) for 45min. the DNA bands were viewed on a UV transilluminator imaging system and an image was captured using Genesnap software (SynGene). DNA size was estimated by comparison to the ladder.

2.7.4 DNA recovery from agarose gels

The DNA fragment(s) of interest were excised from the gel using a scalpel, while the gel was on a UV transilluminator. The DNA was extracted from the gel by using a QIAquick Gel Extraction Kit (Qiagen), following the manufacturer's instructions.

2.7.5 DNA ligation

Routinely, the ligation reaction mixture consisted of 2µl 10x ligation buffer, linearized vector (~50ng), varying amount of double-digested PCR product as the insert (depended on the insert length to give ratio of four times more of insert molecules per one molecule vector), 1µl T4 DNA ligase (3units from Promega) and the volume was made up with water to 20µl. A negative control which contained all the ligation mixture ingredients except the insert was used. For a positive control, the vector was linearized using only one restriction endonuclease and no insert was added. The ligation mixture was incubated for 2hours at room temperature followed by transformation *E. coli* competent cells (section 2.3.3).

2.7.6 Isolation of plasmid DNA

Plasmids DNA were isolated from transformed *E. coli* using the QIAgene miniprep kit following the manufacturer's instructions. 5ml 2TY containing 75µg/ml ampicillin were inoculated with a single transformed *E. coli* colony and incubated at 37°C overnight. Next day, the bacterial cells were lysed, the plasmid DNA was bound to the spin column, washed and eluted into a 1.5ml Eppendorf tube with 50µl elution buffer. Plasmid concentration was routinely measured by using a Nanodrop apparatus.

2.7.7 Restriction digestion

Restriction digestion was used either for diagnostic purposes or in constructing new plasmids. Routinely, the restriction digestion reaction was performed in a 50µl final volume containing; 1000ng plasmid DNA, 5µl appropriate 10x buffer, 2µl of each enzyme. The volume was made up to 50µl with water. The digest was incubated for 2hs at 37°C.

2.7.8 Sequencing of DNA constructs

The sequencing of cloned constructs prepared in this study was performed by SourceBioscience, UK. The plasmids and primer concentration were sent for sequencing following SourceBioscience's instructions.

2.7.9 Annealing of oligonucleotides

Each oligonucleotide was dissolved in a suitable amount of Tris HCl pH 8 according to the manufacturing company's instruction to give a solution at 100µM. Then 5µl of each oligonucleotide was added to 90µl sterile distilled water and samples were incubated in a heating block at 98ºC for 5mins, followed by slow cooling to room temperature. Then 6µl of the annealed oligonucleotides mixture were used as insert in a standard DNA ligation reaction (section 2.7.5)

2.7.10 Constructing plasmids

Plasmids were generated by either homologous recombination (*Saccharomyces cerevisiae*) or restriction digestion method (classical cloning). A list of all plasmids used in this study can be found in Table 2.7.

2.7.10.1 Plasmid construction by homologous recombination

The open reading frame (ORF) of interest was amplified by PCR (section 2.7.2). A pair of primers designed to anneal to each end of the ORF was used. The primers additionally contained 20 nucleotides that were homologous to regions flanking the insertion site in the plasmid. After the vector plasmid had been linearized by using appropriate restriction enzyme(s), both the PCR product of the ORF and the linearized plasmid were transformed into a wild type yeast (section 2.5.3) and cloning occurred by homologous recombination. This method was used to tag the ORF of interest either to N- or C-terminal (Figure 2.2). For C-terminal tagging, the stop codon was deleted from the ORF. The vector plasmids usually contain a 3x Glycine-Alanine linker to allow tag protein and the encoded polypeptide to fold up separately.

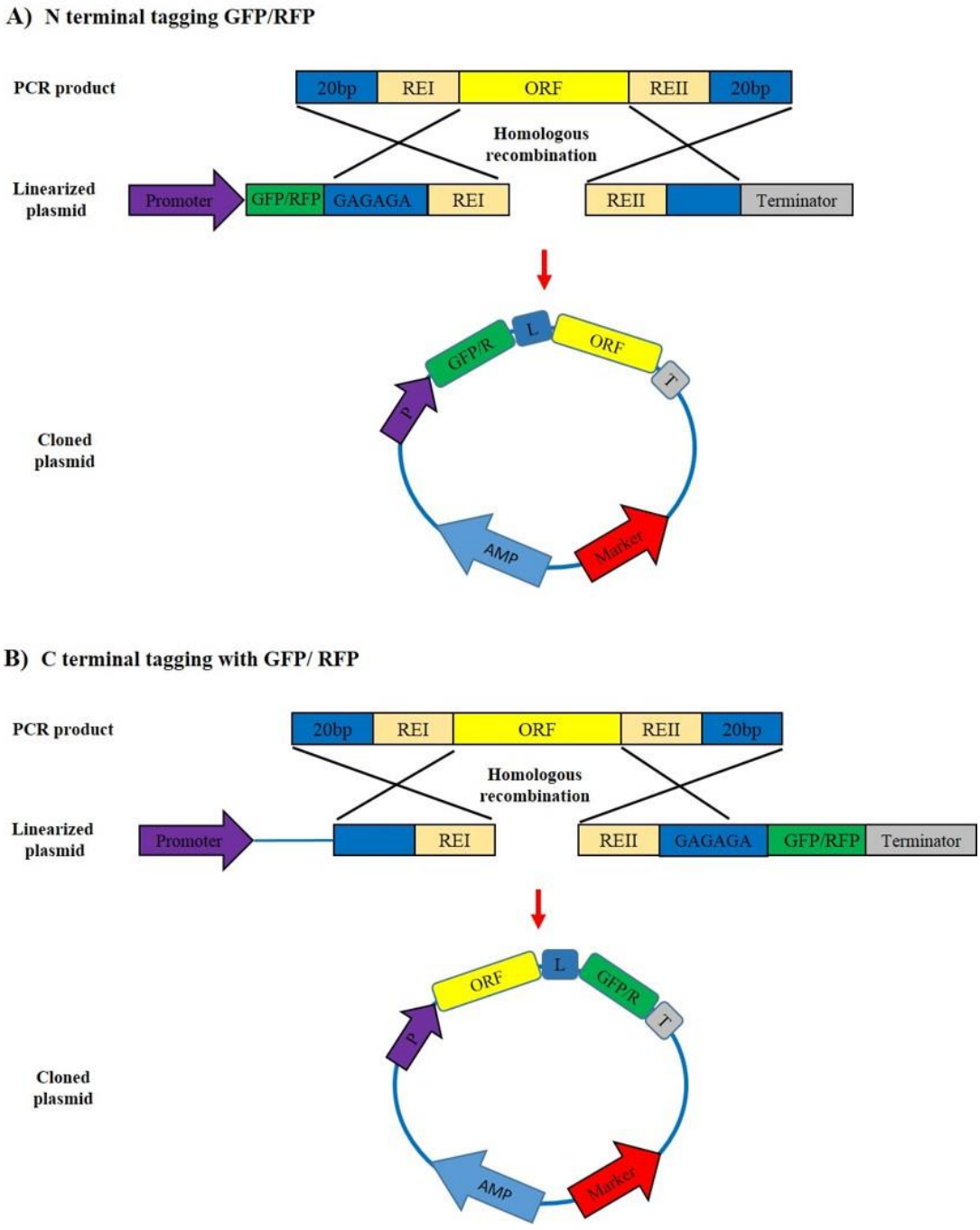


Figure 2.2: Schematic diagram for construction of plasmids in *S. cerevisiae* by homologous recombination. A) Making plasmids with N terminal GFP or RFP tagging. B) Forming plasmids with C terminal GFP or RFP tagging.

2.7.10.2 Plasmid construction by classical cloning

The insert (ORF of interest) was obtained either by PCR amplification (section 2.7.2) or by digesting plasmid containing the ORF with restriction enzymes. The vector plasmid and the insert were digested with the same enzymes, to form similar DNA sticky ends. The released ORF and the linearized vector were separated on an agarose gel (section 2.7.3) and extracted from the gel (section 2.7.4). T4 DNA ligase was used to ligate the ORF with the vector as described in sector (section 2.7.5).

2.8 CRISPR plasmids

2.8.1 First CRISPR plasmid

2.8.1.1 The plasmid vectors

Plasmid 339-3mRFP mars which contains the red fluorescent protein gene located between several restriction sites was modified by deletion of the RFP gene. A new multiple cloning site (MCS) was created by inserting two annealed oligonucleotides (VIP2299 and VIP2300). The eventual plasmid is called pMUR10 (Figure 2.3).

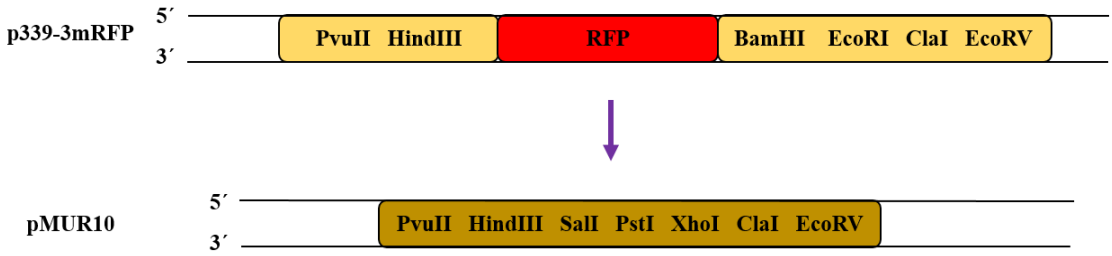


Figure 2.3: Schematic diagram of the restriction sites in p339-3mRFP and p339-3MUR10 plasmids. New restriction sites were added to the original plasmid p339-3mRFP forming pMUR10 (or p339-3MUR10).

Plasmid 339-3mRFP was isolated from an overnight culture of transformed *Escherichia coli* DH5α and two restriction enzymes, HindIII and ClaI, were used to digest the plasmid DNA in order to insert the new restriction sites.

Three versions of the first DdCRISPR plasmid were constructed by inserting *wt*Cas9 into pMUR10, whereas *h*Cas9 and *Dd*Cas9 were inserted into p339-3mRFP separately.

2.8.1.2 The insert

Wild type cas9 and humanised *Cas9* endonuclease open reading frames were amplified by PCR, VIP2301+VIP2302 and VIP2226+VIP2227 primers respectively. A FLAG epitope was added to the N terminal end of wtCas9 during the PCR encoded by the forward primer VIP2301. The flag epitope has the sequence DYKDDDDK. The same nuclear localisation signal (NLS) was added to the 3´ end of the wtCas9 by insertion of the suitable sequence into the reverse primer VIP2329 (Figure 2.4).

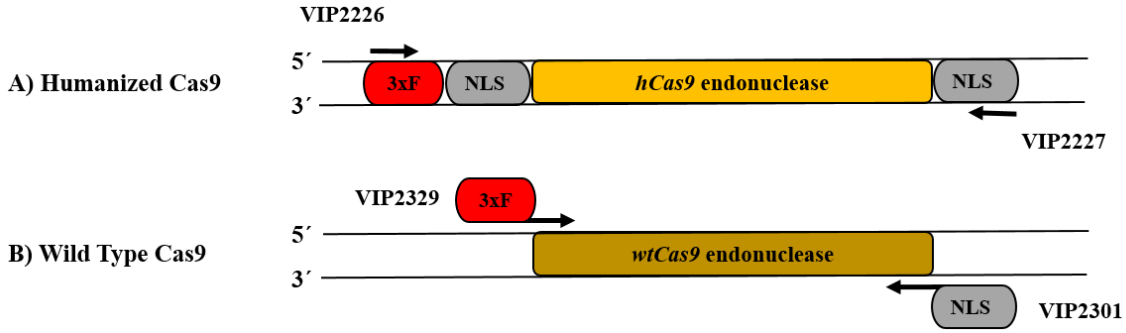


Figure 2.4: Schematic diagram of CRISPR first plasmid. The diagram shows the structure of two versions of the first CRISPR system plasmid. A) Human codon optimized Cas9, which already contained an NLS and 3x flag epitope (commercially provided by Addgene). B) Wild type (bacterial) Cas9 gene (Addgene). The NLS and a FLAG epitope were added to the original gene by PCR and classical cloning.

Another version of cas9 endonuclease was made. It was designed to include the codons commonly used by *D. discoideum*, and it was called *Dictyostelium discoideum Cas9* (*DdCas9*). The same amino acid sequence of *wtCas9* was used as a template to designed *DdCas9*. Amino acid codons were optimized to be expressed in *D. discoideum*. Most of the amino acid codons rarely used by *D. discoideum*, if not all, were changed by the commonly used codons. The SV40 NLS and nucleoplasmin NLS were added to the N and C terminal of the *Ddcas9* respectively. Furthermore, a 3x FLAG epitope was inserted into the extreme N terminal end. This gene was made by GenScript Company (http://www.genscript.com)

2.8.2 Second CRISPR system plasmid

The second DdCRISPR system plasmid was designed to express the guide RNA (gRNA). It consisted of several different parts: the *D. discoideum* U6 promoter (100bp only, because Eurofins Genomics could not make longer than that because of high content of A/T nucleotides), followed by 18bp which could be changed by the BbsI restriction enzyme and replaced by a unique sequence of 20 nucleotides from the target gene, and then the gRNA, finally, the U6 terminator (Figure 2.5). The construct was synthesized by the Eurofins Genomics. The complete sequence was then ligated into a pDXA3C-hygromycin resistant plasmid (section 2.7.5) after it had been digested with BglII and XhoI.

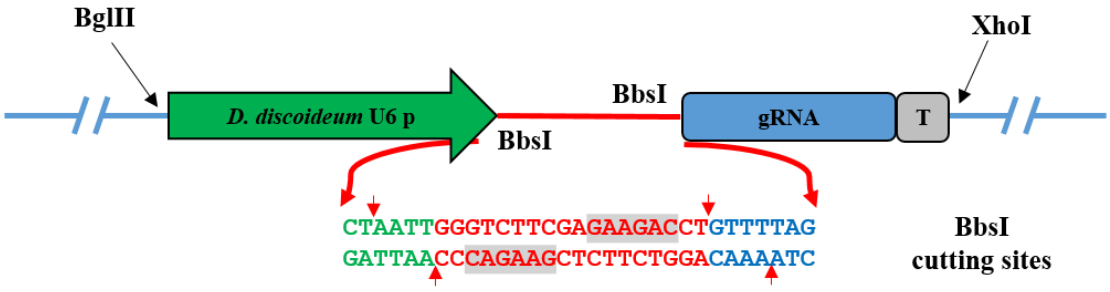


Figure 2.5: Schematic diagram of the second CRISPR system plasmid. The diagram shows the locations of the restriction enzymes and the sequence recognized by the BbsI enzyme, which is used to insert 20bp of the target gene, in between the *Dictyostelium* discoideum U6 promoter and the gRNA which is followed by the terminator. The cluster

was inserted between the BglII and XhoI sites in the pDXA3C-hygromycin resistant plasmid.

Another version of the U6 promoter was made by PCR using the *D. discoideum* gDNA template. A new DNA polymerase (MyFi, Bioline) was used and the PCR protocol was modified. The MyFi polymerase apple to amplify an A/T rich regions of *D. discoideum* gDNA. The following changed were applied in the standard MyFi provided PCR conditions, which they allowed MyFi to amplify 450bp of the *D. discoideum* U6 promoter. First, the annealing temperature was increased to 60ºC and longer forward and reverse primer were designed to ensure the right annealing will be formed between the primers and the gDNA. Secondly, decrease the extension temperature to be 62ºC instead of 72ºC. That would allow the newly synthesised DNA to be in contact with the templet DNA, which would help the polymerase hanging on the DNA. Finally, the extension time was increased from 15sec/ 1Kbp to 10mins/ 1Kbp.

Table 2.7: Plasmids used in this study.

Plasmid name	Promoter	ORF Description	Parental plasmid	Source
Yeast plasmid				
Ycplac33	---	Empty plasmid *URA3*/Centromeric		Lab stock
Ycplac111	---	Empty plasmid *LEU2*/Centromeric		Lab stock
pEH001	*HIS3*	GFP-PTS1	Ycplac33	E Hettema
pEH012	*TPI*	GFP-PTS1	Ycplac33	E Hettema
pAS63	*HIS1*	*Hc*-Red-PTS1	Ycplac111	A Motley
pNA019	*PEX11*	*PEX11-mCherry*	Ycplac33	N Alsaray
pMUR23	*GAL1/10*	*GFP-DdSQS*	pAS113	This study
pMUR24	*GAL1/10*	*GFP-DdSQE*	pAS113	This study
pMUR25	*GAL1/10*	*GFP-DdOSC*	pAS113	This study
pMUR26	*GAL1/10*	*GFP-DdSMT*	pAS113	This study
pMUR29	*GAL1/10*	*DdSQE-GFP*	pAS52	This study
pMUR30	*GAL1/10*	*GFP-DdSQSΔH*	pAS113	This study
pMUR31	*GAL1/10*	*GFP-DdSQSΔPTS1*	pAS113	This study
pMUR32	*GAL1/10*	*GFP-DdSQSΔHΔPTS1*	pAS113	This study
pMUR37	*GAL1/10*	*mCerry-DdSQS*	pNA40	This study
pMUR42	*GAL1/10*	*mCerry-DdSQS*	pAS52	This study

pMUR57	*GAL1/10*	*GFP-104aaDdSQS*	pMUR23	This study
pMUR58	*GAL1/10*	*GFP-65aaDdSQS*	pMUR23	This study
pMUR61	*GAL1/10*	*GFP-37aaDdSQS*	pMUR23	This study
pMUR62	*GAL1/10*	*GFP-24aaDdSQS*	pMUR23	This study
pMUR70	*GAL1/10*	*GFP-DdSQSΔ1-150ΔPTS1*	pMUR21	This study
pMUR71	*GAL1/10*	*GFP-DdSQSΔ1-251ΔPTS1*	pMUR21	This study
pMUR72	*GAL1/10*	*GFP-DdSQSΔ1-312ΔPTS1*	pMUR21	This study
pMUR73	*GAL1/10*	*GFP-DdSQSΔ1-351ΔPTS1*	pMUR21	This study
pMUR80	*GAL1/10*	*3HA-IAA17DdSQS-DdSQS*	pNUT96	This study
pMUR94	*GAL1/10*	*mCerry-DdSQSΔH*	pMUR42	This study
Dictyostelium discoideum plasmids				
pMUR5	*100bpU6*	gRNA		Eurofins
pMUR6	---	*wtCas9*		Addgene
pMUR7	*100bpU6*	20bp of the target +gRNA	pMUR5	This study
pMUR8	*100bpU6*	Target A *PEX7* +gRNA	pMUR7	This study
pMUR9	*100bpU6*	Target B *PEX7* +gRNA	pMUR7	This study
pMUR10	*Actin15*	Multi cloning site	339-3:mRFP in pBsrH	This study
pMUR11	*Actin15*	*wtCas9*	pMUR10	This study
pMUR12	*Actin15*	*hCas9*	339-3:mRFP in pBsrH	This study
pMUR13	*100bpU6*	Target A *GFP* +gRNA	pMUR7	This study
pMUR14	*100bpU6*	Target B *GFP* +gRNA	pMUR7	This study
pMUR15	*100bpU6*	Target C *GFP* +gRNA	pMUR7	This study
pMUR20				
pMUR21	*Actin15*	*DdSQSΔH*	pDXA-3C-MA	This study
pMUR22	*Actin15*	*DdSQSΔHΔPTS1*	pDXA-3C-MA	This study
pMUR33	*Actin15*	*DdCas9*	339-3:mRFP in pBsrH	This study
pMUR35	*450bpU6*	20bp of the target +gRNA	pMUR7	This study
pMUR35I	*450bpU6*	Target A *PEX7* +gRNA	pMUR35	This study
pMUR35II	*450bpU6*	Target B *PEX7* +gRNA	pMUR35	This study
pMUR35III	*450bpU6*	Target A *ATG1* +gRNA	pMUR35	This study
pMUR35IV	*450bpU6*	Target B *ATG1* +gRNA	pMUR35	This study
pMUR35V	*450bpU6*	Target C *ATG1* +gRNA	pMUR35	This study
pMUR49	*Actin15*	*DdCas9-GFP*	pMAA-1	This study
pMUR50	*Actin15*	*NLS-GFP-NLS*	pMAA-2	This study
pMUR52	*Actin15*	*DdREDA-GFP*	pMAA-1	This study
pMUR53	*Actin15*	*DdREDB-GFP*	pMAA_1	This study
pMUR63	*Actin15*	*GFP-1-187aaDdSMT*	pMUR21	This study
pMUR64	*Actin15*	*GFP-DdSMTΔ1-187*	pMUR21	This study
pMUR136	*Actin15*	*DdSQE-mRFP*	339-3:mRFP in pBsrH	This study
Drosophila plasmids				
pJL315	*MT*	*mRFP-PTS1*	pPD006	J. Lacey
pMUR143	*MT*	*eGFP-DdSQE*	pJL005	This study

2.9 Protease protection:

Saccharomyces cerevisiae expressing *Dictyostelium discoideum* proteins (squalene synthase, squalene epoxidase and cycloartenol methyltransferase) behind a conditional promoter (galactose10) were used to test the protein orientation in the peroxisomal membrane.

A protease protection technique was used to identify the protein orientation and side direction in the peroxisomes membrane whether they are exposed to the cell cytosol or facing the peroxisomal matrix. *S. cerevisiae* expressing *D. discoideum* squalene synthase (*Dd*SQS) joined N terminus of GFP was chosen to test the orientation of GFP-DdSQS. Yeast cells were grown overnight in 3ml YM2 medium containing 2% glucose in the absence of leucine, the selectable marker. Then next morning, the culture was used to inoculate a 10ml culture of the same medium containing 0.3% glucose, starting from OD_{600} 0.15. The culture was grown for 9h. Oleate medium (100ml) was inoculated with all of the previous cultures and incubated at 30ºC overnight with shaking at 180tpm. GFP-*Dd*SQS production was induced by adding 10ml galactose (2%) to the culture. GFP-*Dd*SQS expression was checked under a fluorescent microscopy after 2hs. The cells were centrifuged at 180xg for 4min and washed with fresh YM2 medium once. The final pellet was resuspended in YM2 medium containing 2% glucose and incubated for 1h at 30ºC with shaking at 180xg. Spheroplasts were prepared as described in (section 2.5.5). The yeast spheroplasts were resuspended in 2ml ice-cold 650mM sorbitol, 5mM MES pH 5.5, 1mM KCl, 1mM EDTA, and then disrupted in dounce homogeniser at 4°C. The sample was then centrifuged for 5 mins at 400xg at 4°C to remove any intact cells. The supernatant was added to a new pre-cooled 2ml Eppendorf tube and centrifuged again in the same conditions. The final clear supernatant was used as a homogenate.

Proteinase K (from *Trititachium album*) attached to agarose beads was used for testing the membrane-associated protein orientation. Homogenates were treated with proteinase K-agarose beads in the presence and absence of Triton X-100 and were incubated at room temperature for different periods of time (0, 5, 15, 30 and 60mins). Each reaction was stopped by adding TCA to 10% final concentration. The samples were incubated for 10 mins on ice (at 0ºC) and then centrifuged at 4ºC for 5 mins at 15,700xg. The supernatants were discarded and the pellets were resuspended in 10µl 1M Tris (pH 9.5).

A 90µl 1x protein loading dye (PLD) were added to each sample. Samples were boiled at 95ºC for 5 mins by using a heating block, incubated for 2 mins on ice, centrifuged at 4ºC for 1min at full speed. After that, Samples were loaded on 7.5% and 10% SDS-PAGE and western blot analyses.

2.10 Auxin degron system analysis

Plasmid encoded *Dd*SQS tagged to 3x HA-IAA17 at the N terminus end (3x HA-IAA17-*Dd*SQS) was transformed to YEH315 and YEH381 (YEH315Δ*pex1*) yeast strains (expressed TIR1 protein), whereas YEH404 and YEH405 strains (expressed AID-GFP-PTS1 in peroxisome and cytoplasm respectively) used as a control (Table 2.3). Transformed cells and YEH404 and YEH405 strains were grown in 5ml YM1 minimum medium containing 2% glucose (in the absence of leucine -as a selectable marker- in case of the transformed strain) overnight at 30ºC with shaking 180rpm. Subcultures of the same fresh 100ml medium were started next morning from OD_{600} 0.2 for 5-6hs in the same conditions. Cultures were treated either with or without 1ml of stock solution indole-3-acetic acid (auxin, Sigma-Aldrich) at final concentration 500µm. Samples of 10 OD_{600} were taken from each grown culture strain after 0, 30, 60, 90, 120 and150min. Samples then were centrifuged at 1100xg for 2 mins and flash froze in liquid nitrogen before had been stored in -80ºC. Total protein was extracted by following TCA protein precipitate method (section 2.11.1). Samples were loaded onto SDS-PAGE gel (section 2.11.2), analysed by western blot technique and visualised by GenSnap software as described briefly in section (section 2.11.4)

2.11 Protein procedures

2.11.1 Trichloroacetic acid (TCA) protein precipitate method

For total protein precipitation, *D. discoideum* culture was grown on HL5 glucose medium overnight (yeast culture was grown until OD_{600} 0.5). 10 ODs of each cells culture was harvested by centrifugation at 15,700xg for 5min at room temperature. The supernatants were discarded. Each pellet was either flash frozen in liquid nitrogen and stored at -80°C or used directly. All the next steps were carried out on at 4°C or on ice

with precooled buffers and centrifuge. Frozen pellets were resuspended in 0.5ml 0.2M NaOH and 0.2% (v/v) β-mercaptoethanol, prepared freshly, and incubated for 10mins on ice. 71µl 40% (v/v) TCA were added to each sample to a final concentration of 5% and the samples were incubated for 10 mins at 0°C. After centrifugation in a micro-centrifuge at 15,700xg for 5mins, the supernatants were removed totally, and the pellets were resuspended in 10µl 1M Tris (pH9.4). Subsequently, 90µl 1x protein loading buffer (PLB) (which consists of 62.5mM Tris-Cl pH 6.8, 10% (v/v) glycerol, 2.3% (w/v) SDS, 0.05% (w/v) bromophenol blue, and 0.025M DTT) was added to each sample. Samples pH were checked from the final colour. A blue samples had a suitable pH, while green or yellow coloured samples had to be adjusted to a blue colour by adding further 1M Tris. Finally, samples were boiled at 95ºC for 5 mins, incubated 2 mins on ice, centrifuged at 15,700 in a microfuge for 1min at 4ºC and were then ready for loading onto the SDS-PAGE gel.

2.11.2 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed as described by Sambrook and Russell (2006). The 7.5%, 10% or 12% polyacrylamide resolving gels, as well as the stacking gels were made by mixing stock solutions as shown in Table 2.8. Protein samples were denatured in 4x protein loading buffer (200mM Tris-HCl pH 6.8, 400mM DTT, 8% (w/v) SDS, 0.4% (w/v) bromophenol blue and 40% (v/v) glycerol), heated for 5-10 mins at 95°C and then centrifuged at full speed in a bench centrifuge for 1min. The gels were loaded with 2-25µl protein sample. To estimate the size of the detected proteins, a pre-stained protein ladder (Pageruler) was also loaded onto the gel. Electrophoresis was run at 150V in 1x protein running buffer (25mM Tris, 250mM glycine, 0.1% (w/v) SDS) until the bromophenol blue dye had just reached the bottom of the resolving gel. Subsequently, gels were either stained with Coomassie blue staining solution or used for western blotting.

Table 2.8: Showing the composition of individual SDS-PAGE gels.

Stock solution	Resolving gel (10ml)			Stacking gel (10ml)
	7.5%	10%	12%	
4x Resolving buffer (1.5M Tris-HCl, 0.4% (w/v) SDS, pH8.8)	2.6ml	2.6ml	2.6ml	
Stacking buffer (0.5M Tris-HCl, 0.4% (w/v) SDS, pH6.8)				2.5ml
Protogel (30% (w/v) Acrylamide: 0.8% (w/v) Bis-Acrylamide Stock solution)	2.5ml	3.3ml	4ml	1.3ml
Water	4.9ml	4.1ml	3.4ml	6.1ml
10% (w/v) ammonium persulphate APS	100µl	100µl	100µl	100µl
N,N,N',N'-tetramethylethylenediamine TEMED	10µl	10µl	10µl	10µl

2.11.3 Coomassie Blue staining

The polyacrylamide gels were stained with Coomassie Blue stain (50% (v/v) methanol, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie Blue R-250) for 1 hour at room temperature with gentle shaking, followed by washing in destain solution (10% (v/v) methanol, 10% (v/v) acetic acid) until the pale blue background of the gel had been removed and the protein bands became visible.

2.11.4 SDS-PAGE protein analysis by western blotting

Protein samples separated by 7.5% or 10% SDS-PAGE, were blotted onto a nitrocellulose membrane (Schleicher and Schuell) using a Trans-Blot Mini Cell (Bio-Rad), as described in the manufacturer's manual (www.bio-rad.com). Samples were transferred at constant 200mA for 2hs in a pre-cooled transfer buffer (12.5mM Tris, 125mM glycine, 20% (v/v) methanol). The protein transfer efficiency was detected by using Ponceau red staining solution (0.1% Ponceau S (w/v) in 5% (v/v) acetic acid). The nitrocellulose membrane was covered with the Ponceau stain for 1min before washing in 1xTBS-T buffer (1xTris buffered saline (TBS) composing 20mM Tris, 500mM NaCl pH 7.6, 0.05% (v/v) Tween20). The blotted membrane was blocked by incubation in 10ml blocking buffer (2% (w/v) fat-free Marvel milk in TBS-T) for 1h with constant mixing at room temperature. Subsequently, the membrane was incubated with a fresh

10ml blocking buffer containing the desired dilution of the primary antibody (determined experimentally) (Table 2.9) and incubated with mixing for 1h at room temperature. To remove unbound primary antibody, the membrane was washed three times, 5min each, with 10ml 1xTBS-T. Next, the membrane was incubated with 10ml blocking buffer containing the appropriate dilution of the secondary antibody (Table 2.9) for 1h with mixing at room temperature. The washing steps were repeated again as described above. Finally, the protein samples were detected by using a Syngene GBox imaging system after 1min incubation of the membrane with proteins enhanced chemiluminescence (ECL) reagent following the manufacturer's instructions. The exposure times were determined experimentally.

Table 2.9: The following antibodies were used in this study.

Antibody	Purpose	Dilution	Source
Mouse monoclonal anti-GFP	Primary antibody to detect GFP and GFP tagged proteins	1:3000	Roche
Mouse monoclonal anti-Flag	Primary antibody to detect Flag tagged proteins	1:4000	Sigma
Mouse monoclonal anti-Pgk1	primary antibody detect 3-phosphoglycerate kinase	1:7000	Invitrogen
Anti-thiolase	Primary antibody to detect yeast thiolase	1:5000	
Anti-Pxa1	Primary antibody to detect yeast proteins	1:5000	BioServ
Anti-Cas9	Primary antibody to detect Cas9 and *Dd*Cas9	1:2000	New England Biolabs
Polyclonal goat anti-mouse HRP (GαM)	Secondary antibody to detect mouse primary antibodies	1:4000	Sigma
Polyclonal rabbit anti-rat HRP (RARP)	Secondary antibody to detect rat primary antibodies	1:4000	Sigma
Polyclonal goat anti-rabbit HRP (GARP)	Secondary antibody to detect goat primary antibodies	1:5000	Eurogentec

2.11.5 Analysis of peroxisomal membrane proteins topology by sodium chloride treatment

Membranes were incubated in high salt concentrations (1M NaCl) to release the peripheral membrane proteins from the peroxisomal membrane. Approximately 10^6- 10^7 cells were used per each topology experiment for either yeast or *D. discoideum*. Cells were harvested at 180xg for 5mins at 4°C and washed twice with 8ml HEPES buffer (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH7.5. After the supernatants had been discarded, the pellets were resuspended in 1ml 50mM HEPES buffer containing protease inhibitors. Cells were then lysed by either sonication on ice for 10 secs at 10 amplitude microns (Sanyo sonicator) or bead beating (Biospec Products) for amoebae and yeast cells respectively. Samples were centrifuged at 180xg for 5 mins at 4°C, to remove any intact cells. The supernatants (homogenate) were transferred to a new pre-cooled Eppendorf tube. 75µl homogenate was added to 25µl 4x protein loading buffer (PLB) and used as a total lysate, the remaining homogenates were centrifuged in a polycarbonate tube (Beckman) at 80,000rpm for 1h at 4°C using a Beckman Optima MAX-E Ultracentrifuge (TLA 100.2 rotor). Subsequently, the supernatants were discarded after 75µl had been taken for analysis as the first supernatant. Then the pellets had been resuspended in 1ml 1M NaCl/ 50mM HEPES pH7.5 and incubated for 30 minutes on ice with mixing occasionally. The suspensions were centrifuged by ultracentrifugation as describe above. Proteins in the second supernatant were precipitated by adding, 250µl of 40% (v/v) trichloroacetic acid to 750µl of the supernatant and the mixture was kept on ice for 30 mins. Subsequently, samples were centrifuged for 15,700xg at 4°C for 5mins. The pellet was resuspended in 10µl 1M Tris pH9.4, after the supernatant been discarded. 90µl of 1x protein loading buffer were added to the mixture. A further 10x dilution was carried out in 1x PLB this gave the second supernatant. The membrane pellet fraction was resuspended in 1ml 50mM HEPES containing 1% (v/v) Triton X-100 and a 75µl sample was kept as the final pellet. The samples were analyzed by 10% SDS-PAGE and followed be western blot.

2.11.6 Analysis of membrane topology by sodium carbonate

The sodium carbonate peroxisomal membrane protein extraction procedure was adapted from Fujiki *et al.*, 1982 and Elgersma *et al.*, 1997. This technique was used to distinguish between the peripheral and integral membrane proteins by resuspension the membranes in 0.1M Na_2CO_3 pH 11.4. Cells were harvested, washed twice, resuspended in 1ml 50mM HEPES pH7.5 and lysed as described previously. After 75µl had been kept as a total lysate, the sample was centrifuged at 4°C for 1h at 80,000rpm (using Optima MAX-E Ultracentrifuge, TLA 100.2 rotor). 75µl of the first supernatant was collected. The pellet was resuspended in 12.5ml 100mM Na_2CO_3 pH11.4 (in order to dilute the protein concentration to less than 10mg/ml). The mixture was incubated for 30 minutes on ice. To pellet the integral membrane proteins, 1ml of the previous suspension was centrifuged again at 4°C for 1h at 80,000rpm (using Optima MAX-E Ultracentrifuge, TLA 100.2 rotor). 500µl of the second supernatant was concentrated by centrifugation in the ultra-0.5mL centrifugal filter with a 10K cut-out (Millipore) at full speed in a micro-centrifuge for 20 mins at 4°C. The final volume was made up to 80µl and a 75µl sample was taken as the second supernatant. The pellet was resuspended in 80µl 50mM HEPES containing 1% (v/v) Triton X-100, and a 75µl sample was taken for analysis as the final pellet. Equivalent amounts (material from the same number of cells) of each fraction were examined on 10% SDS-PAGE and analyzed by the western blot technique.

2.12 Northern blotting procedure

2.12.1 Designing short oligonucleotide probe and 5′-end labelling

Short synthetic DNA oligonucleotide was designed so that it would anneal with a specific sequence of the guide RNA (gRNA). The oligonucleotide had the following properties; 18bp in length, annealing temperature 60ºC, and 50% C/G ratio. Use of the oligonucleotide followed the procedure described by Maniatis et al, (1982). The probe had been labelled at the 5' hydroxyl group using $\gamma[^{32}P]$-ATP (provided by PerkinElmer, PerkinElmer, Mass, USA). 1µl 5µM DNA oligonucleotide sample was incubated at 37°C for 30min in the labelling reaction, which consisted of 14µl DEPC-H_2O, 1µl polynucleotide kinase (PNK), and 2µl 10x PNK buffer, and 1µl γ $[^{32}P]$-ATP. Finally,

the mixture was mixed with 1 ml hybridisation buffer, heat-inactivated for 5 mins at 65°C and then filtered by using a 0.22µm Millipore filter before use.

2.12.2 RNA extraction and purification from *D. discoideum*

RNA was extracted from growing amoebae according to the hot phenol method described in (Maniatis, at, el. 1982) with the addition of some changes adapted from (Sambrook, 2012). Amoebal cells expressing CRISPR gRNA were grown in 100 ml medium at 22ºC and shaking at 180rpm for 3days. $5x10^7$ cells were harvested by centrifugation at 180xg for 5 mins. After the supernatants had been removed completely 1000µl glass beads, 500µl phenol and 500µl GTC mix (6.1 M guanidine thiocyanate, 15mM EDTA (pH 8.0), 75mM Tris-HCl (pH 8.0), 3% sarkosyl, 1.5% β-mercaptoethanol in DEPC-treated water) were added. The mixture was vortexed for 5min. After that, 3.5ml phenol and 3.5ml of GTC mix were added, mixed briefly and heated at 65°C for 10min. The mixture was cooled on ice for 5min before 4ml chloroform and 2ml sodium acetate (NaAc) mix (10mM Tris-HCl (pH 8.0), 1mM EDTA, and 100mM NaAc (pH 5.0)) were added and the samples mixed by inverting the tube several times. Following this, the aqueous phase was isolated after centrifugation for 5 mins at 2000xg, transferred to a new tube and then re-extracted twice with an equal volume phenol/chloroform. RNA was precipitated by adding two volumes 100% ethanol to the aqueous phase and the mixtures were left at -80ºC for 1-3 hs. Precipitated RNA was pelleted by centrifugation for 10min in a microfuge. The resulting pellet was washed twice in 70% ethanol then air-dried for 10min and dissolved in 100µl DEPC-treated water. RNA concentration was measured by OD_{600} spectrophotometry. RNA samples were either used directly or saved in -80°C.

2.12.3 Acrylamide gel electrophoresis

Total RNA was fractionated by acrylamide gel electrophoresis. An 8% acrylamide gel was prepared in 0.5x TBE buffer containing 50% (w/v) urea (30g urea, 12ml 40% acrylamide (19:1), 6ml 6xTBE, 18ml H_2O, 400µl 10x APS and 40µl TEMED). Each RNA sample was prepared by mixing in 4µl DEPC-H_2O to give contain approximately 5µg total extracted RNA. 4µl 2x RNA loading buffer (95% formamide, 20mM EDTA,

0.05% (w/v) xylene cyanol, and 0.05% (w/v) bromophenol blue) were added. Then samples were denatured by heating for 5min at 70°C, before they were loaded and run on the gel. Equal amount of RNA per sample were loaded on the acrylamide gel which was run at 200 volts in 0.5x TBE buffer. Electrophoresis was stopped when the cyanol blue dye reached the bottom of the gel. The separated RNAs in the polyacrylamide gel were transferred to a Hybond-N+ membrane (GE Healthcare, GE Healthcare Life Sciences, Bucks, UK) at 15 volts overnight in 0.5x TBE buffer. After transfer, RNAs were cross-linked to the membrane using UV light (1200 joules) at about 10cm in distance. After that, the membrane was incubated in a freshly prepared oligo-hybridisation buffer (6xSSPE, 5xDenhardts, 0.2% SDS) at 37°C for 1h before hybridisation overnight with radio-labelled probe (section 2.12.1). Excess probe was discharged, and to remove non-specifically bound material, the membrane was washed three times in 6xSSPE at 37°C for 15 mins. Subsequently, the membrane was wrapped in Saran wrap and then exposed to MS film (Kodak, Kodak Ltd, Herts, UK) for overnight. For safety reason, the radio-labelled probes were subsequently removed. Blot was stripped by incubating in 100ml northern stripping solution (0.1% SSPE, and 0.1% SDS, heated to boiling before use), three times for 15 mins.

2.12.4 Northern blot analysis

After overnight exposure, the films were scanned by using a Typhoon FLA 7000 scanner (GE Healthcare, UK). Images were visualized by the Typhoon FLA 7000 software and the final figures were prepared using the PowerPoint programme.

Chapter three "Attempt to develop new gene editing system in the slime mould *Dictyostelium discoideum*".

3.1 Introduction

Nitrogen containing bisphosphonate drugs have been used widely to treat and prevent bone resorption diseases (e.g. osteoporosis). The target for these drugs was originally discovered by using the slime mould *Dictyostelium discoideum*. The drugs inhibit growth of *D. discoideum* and the target was found to be the enzyme farnesyl diphosphate synthase (FDPS). Among this group of drugs, the most-widely used are alendronate and residronate but, in comparison with residronate, alendronate is at least 50-times less potent at inhibiting FDPS although it is only two to three times less potent at inhibiting *D. discoideum* growth (Grove et al., 2000). Moreover, it has been shown that the target for these drugs in human osteoclasts is also FDPS and there is the same difference as in *D. discoideum* between the potency of the two drugs as inhibitors of bone resorption and of FDPS (Bergstrom et al., 2000).

In this chapter we try to investigate the reasons for the difference in potency of alendronate and risedronate as inhibitors of cell activity and of FDPS.

At least in *D. discoideum*, uptake of the drugs is by micropinocytosis (Rogers et al., 1997), which is a non-selective process (Marsh, 2001), and the apparently anomalous high potency of alendronate in comparison with that of residronate at inhibiting amoebal growth cannot arise from a greater ability of amoebae to take up alendronate. However, FDPS is a peroxisomal enzyme in *D. discoideum* (Nuttall et al., 2012) and human cells (Biardi et al., 1994; Krisans et al., 1994) and it seemed possible that it is the process by which the bisphosphonate drugs enter the peroxisomes that is responsible for modulating the relative potencies of the two drugs. In *D. discoideum* entry of the enzyme into peroxisomes is dependent on a PTS2 sequence (Nuttall et al., 2012) that is recognised by the cytosolic receptor Pex7 (Kim and Hettema, 2015). Hence, we aimed to eliminate any effect of the peroxisomal membrane on bisphosphonate potency by knocking out activity Pex7 so that FDPS would no longer be in the peroxisomes but in the cytosol. If the peroxisomal membrane is affecting the relative potencies of alendronate and

residronate, growth of the *Pex7*Δ mutant strains should now be considerably less affected by alendronate than by residronate.

In this chapter two main approaches were followed to knock out *Pex7*: classical homologous recombination and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) system.

We attempted to develop the new genome editing system (CRISPR) for use in *D. discoideum*. CRISPR is a bacterial and archaeal immune defence system against foreign invasive viral and plasmid DNA. The system has been recently developed successfully to be a genetic editing system in most known organisms such as bacteria, fungi, plants, animals and human cell lines (Böttcher et al., 2014; DiCarlo et al., 2013; Fujihara and Ikawa, 2014; Vyas et al., 2015; Yang et al., 2014). We aimed to use it in this study, but it would also be a good tool for use in the *D. discoideum* field. Thus, knocking out *Pex7* would have two purposes; first, establishing the CRISPR system and secondly, providing the *D. discoideum Pex7*Δ strain (*DdPex7*Δ) for bisphosphonate potency analysis assay.

3.2 Results

3.2.1 Building up the testing system

A fluorescence test system for the Pex7 pathway was set up for use in *D. discoideum*. Farnesyl diphosphate synthase (FDPS) was tagged with green fluorescence protein (GFP) at the carboxyl end FDP-GFP. This enzyme enters peroxisomes by binding to Pex7p, following the peroxisomal targeting 2 pathway (PTS2). Hence a successful knockout of *Pex7* would leave FDP-GFP in the cytosol (Figure 3.1).

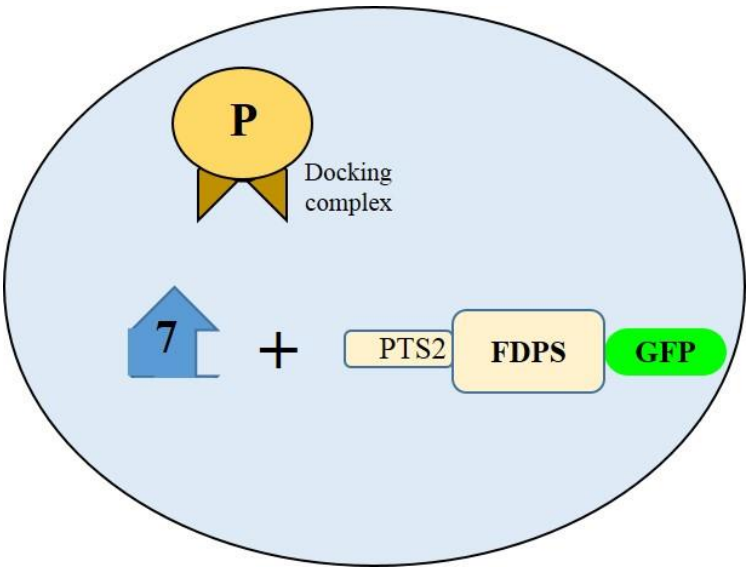


Figure 3.1: Pex7 knock out testing system for peroxisomal PTS2 containing proteins. GFP tagged PDFS following the PTS2 pathway of peroxin 7 protein (Pex7) from the cytosol to the peroxisomes in *D. discoideum*. Afterword, the transporter (Pex7) and the cargo (GFP-PDFS) were landing on the docking complex in the peroxisomal membrane. P, peroxisome and 7, Pex7p

3.2.2 Pex7 knock out using the classical approach (homologous recombination)

A blasticidin resistance cassette in the pLPBLP plasmid was used to try to knock out *Pex7* in *D. discoideum*. Approximately 1753bp of 5′ and 1469bp of the 3′ of the mid genomic sequence of the *pex7* gene were amplified individually using MyFi DNA polymerase in addition to primers VIP1594+VIP1595 and VIP1596+VIP1597, respectively. The two PCR products were ligated individually into pLPBLP following the digestion ligation method. KpnI and HindIII were used for the first digestion, while PstI and BamHI were used for the second digestion. T4 DNA ligase was used to ligate the digested PCR products into pLPBLP. The presence and order of all inserts were checked by enzyme digestion and sequencing of the construct. The resulting (Figure 3.2B) was used for the gene knock out. After *D. discoideum* had been transformed with the knockout cassette, the amoebae were able to grow in a medium containing 10µg/ml blasticidin antibiotic. However, fluorescence microscopy of the transformed amoebae clearly showed green fluorescent puncta in the amoebae. This showed that the *Pex7* gene

was active and was being used for production of the Pex7p that is essential for entry the PDFS-GFP into peroxisomes. Genomic DNA (gDNA) of the transformed amoebae was extracted as described in section (2.). The extracted gDNA was used as a template for gene knock out PCR diagnosis to determine the presence of the insert in the template and its location. Three PCR reactions were carried out 1) to check for the *Pex7* upstream sequence 2) the downstream sequence and 3) both sequences. The PCR results showed the presence of the left and right flanks of the inserted cassette, but the PCR of the whole cassette gave a band similar in size to that obtained when wild type (control) gDNA was used as the template. This suggested that the resistance cassette had been integrated into the genome but that there was also gDNA that did not contain the knock out cassette (Figure 3.3).

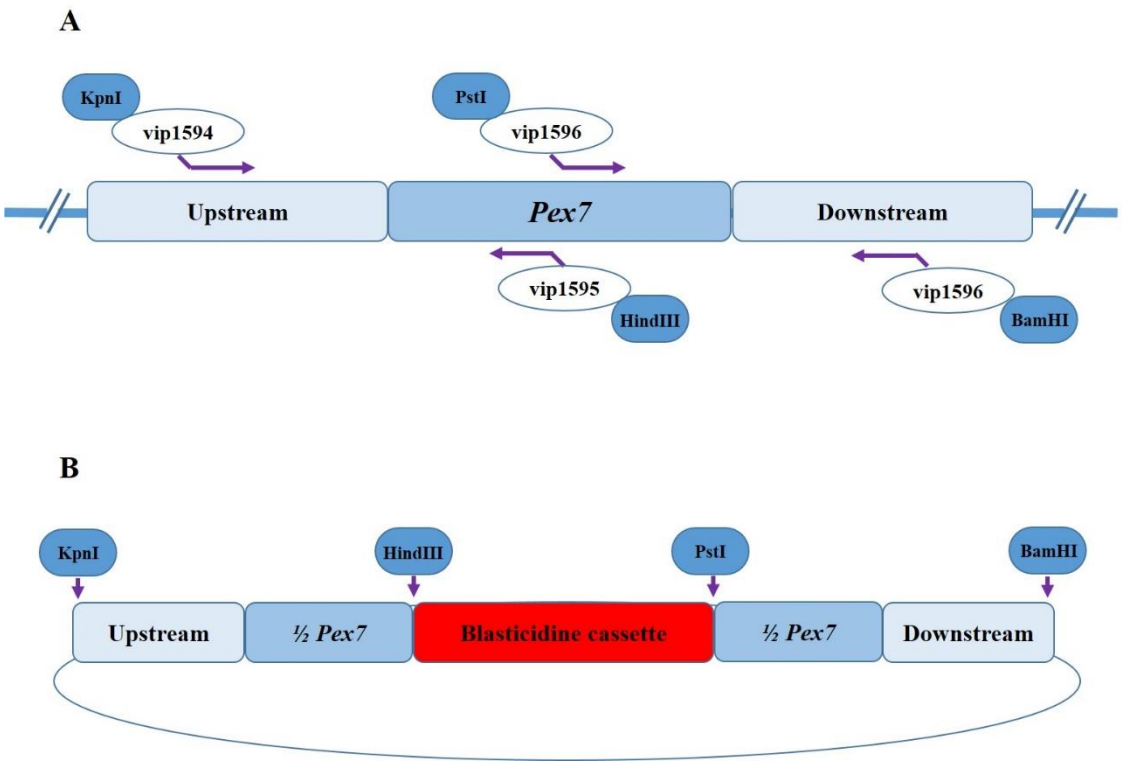


Figure 3.2: Schematic diagram of the blasticidin resistance cassette for *Pex7* knock out in *D. discoideum*. A) Amoebal genomic DNA containing *Pex7* gene. Primers used to form the left and right flanks are shown as small arrows carrying the restriction site of the enzymes used. B) *D. discoideum Pex7* knock out construct structure.

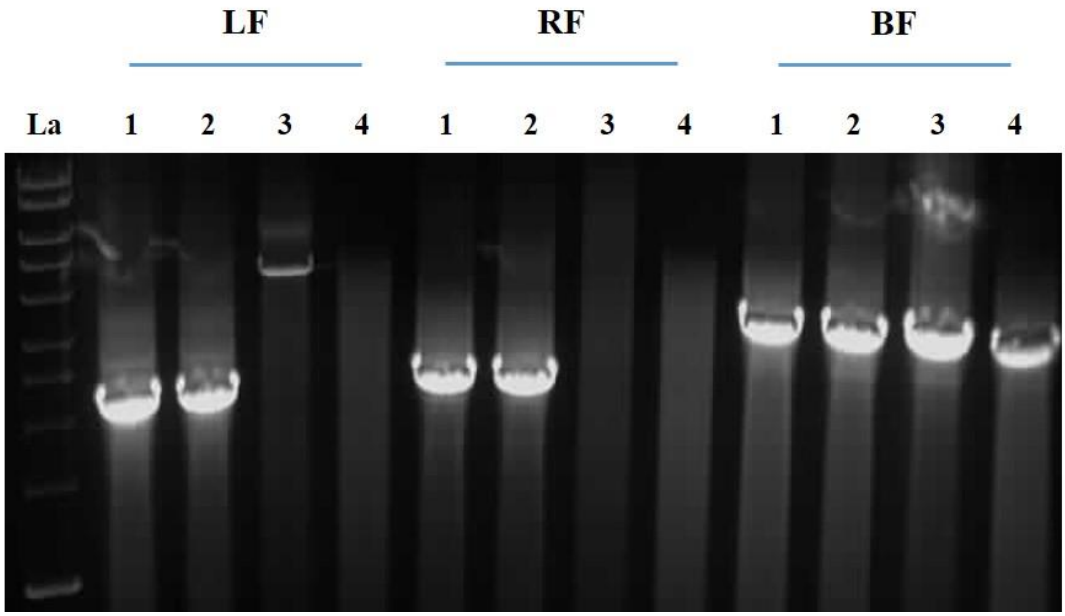


Figure 3.3: PCR diagnostic results for the *Pex7* knock out. Two samples are for *D. discoideum* that had been transformed with the *Pex7* knock out (KO) cassette (lanes 1 and 2). In addition, there are two controls. The first one used gDNA from amoebae transformed with an empty KO cassette as the template (lane 3). The second used gDNA from un-transformed amoebae as the template (lane 4). Lf, RF and BF refer to the left flank, right flank and both flanks of the *Pex7* KO cassette, respectively.

3.2.3 Testing the CRISPR knockout system

It seemed that either we could not knockout the *pex7* gene in *D. discoideum* by using the classical method (homology recombination) or that there is more than one copy of the *Pex7* gene in each amoeba. Thus, we decided to develop a new technique called the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) system and CRISPR Cas9 protein. The CRISPR system has been recently successfully applied to several organisms. We attempted to develop a CRISPR system for *D. discoideum* that might overcome the *Pex7* knock out difficulty by allowing successful genome editing. Our system consisted of two plasmids. The first CRISPR plasmid carried the Cas9 nuclease cDNA and the second encoded a synthetic guide RNA (gRNA) that leads Cas9

to the target regions in the genomic DNA that are to be edited. This hybridizes to the 20 base pairs of the guide and is cleaved if it is followed by the protospacer adjacent motif (PAM) sequence NGG.

Function of the system depends on the presence of Cas9 endonuclease which then reaches the target nucleic acid sequence with help from the single stranded guide RNA (gRNA).

3.2.3.1 First *D. discoideum* CRISPR plasmid (*Dd*CRISPR)

To generate a CRISPR system for *D. discoideum*, we decided to carry the system on two plasmids. The first *Dd*CRISPR system plasmid contains the open reading frame of the *Cas9* gene, and the second encodes the guide RNA (gRNA). Three versions of *Cas9* were used in this study. We started with the wild type *Cas9*, which is a bacterial gene extracted originally from *Streptococcus pyogenes* (Qi et al., 2013). In addition, we used the humanized *Cas9* (*hCas9*), in which the amino acid codons have been optimized for expression in mammals (Ran et al., 2013). Both were commercially available and obtained from Addgene. The bacterial cas9 was modified before use, codons for three SV40 nuclear localization sequences (NLS) were added to the 5′ end, while three FLAG epitope sequence were added to the 3′. However, the *hCas9* open reading frame already had these modifications and used as it is. Both cas9 genes were cloned (either directly or after modification of the restriction sites) into the 339-3: mRFP Blasticidin resistance (***Bs^R***) plasmid between the actin 15 promoter (A15prom) and the actin 8 terminator (A8term) by use of the digestion ligation method. The third *Cas*9 was designed to be expressed particularly in *D. discoideum* and we called it *DdCas9*. The sequence encoding *D. discoideum* codon optimized *S. pyogenes Cas9* was fused with three 5′ SV40 NLS and 3x flag epitope sequences, in addition to one 3′ nucleoplasmin NLS (Figure 3.4). The whole construct was synthesised by GenScript Company and then amplified and cloned into the 339-3: mRFP plasmid.

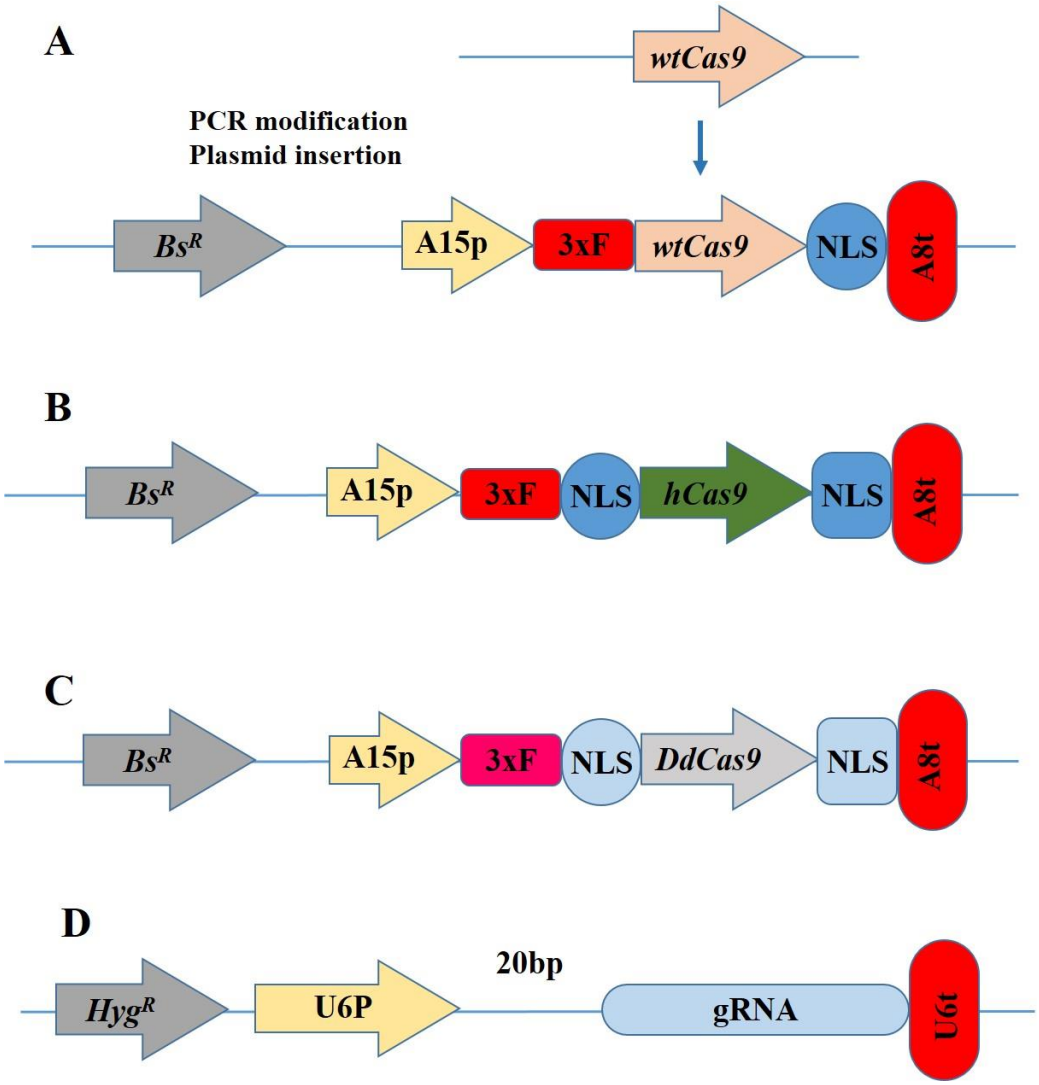


Figure 3.4: Schematic structure of all versions of Cas9 and gRNA used in this study. All *Cas*9 genes were carried on a blasticidin resistance (Bs^R) plasmid, while the gRNA was cloned into a hygromycin resistance (Hyg^R) plasmid. A) Wild type *Cas9* was modified by PCR by adding 3x Flag epitope and SV40 NLS (Blue circle) before been constructed into the first CRISPR. B) Humanized Cas9 was amplified and cloned into the plasmid. C) Dictyosteliumized *Cas9*, which consists 3x Flag epitope and the SV40 NLS (light blue circle) and the nucleoplasmin NLS (light blue squire), was incorporated into the plasmid. D) The U6 promoter was either made by Eurefin or modified PCR technique, followed by 20bp homologue of the target gene and the gRNA scaffold.

3.2.3.2 Second *Dd*CRISPR plasmid

The RNA polymerase III promoter U6 was used to express a small single stranded RNA, the guide RNA (gRNA), which is required for Cas9 targeting. The second *Dd*CRISPR plasmid was designed to express the gRNA. It consists of the *D. discoideum* U6 promoter (*Dd*U6), two restriction sites for the BbsI restriction enzyme (to be used for insertion of the 20bp targeting homologue), and the scaffold RNA and finally the *Dd*U6 terminator (Figure 2.5). Three difference lengths (100bp or 450bp) of the U6 promoter were used. The 100bp was synthesised by Eurofin, while the 450bp fragments created by PCR using developed conditions and new type of proofreading DNA polymerase MyFi (sector 2.8.2).

First, the *Dd*CRISPR plasmid containing either *hCas9* (pMUR8) or *wtCas9* (pMUR9) were used. Only the 100bp U6 promoter was available at this stage of the study. Three different loci from the *Pex7* gene followed by the NGG PAM sequence, were chosen as the 20bp sequence for insertion into the gRNA to target the Cas9 endonuclease. Both CRISPR system plasmids were successfully constructed and checked by enzyme digestion and sequencing. The *D. discoedium* knock out testing system was transformed initially with the first *Dd*CRISPR plasmid. Cultures of the transformed amoebae were then retransformed, each with a different version of the second *Dd*CRISPR plasmid. Successful transformation with each pair of plasmids was indicated by the ability of the transformants to grow in medium containing G418 and blasticidin. Nevertheless, there were no significant differences between the green fluorescent peroxisomal puncta in the transformed and untransformed amoebae (Figure 3.5).

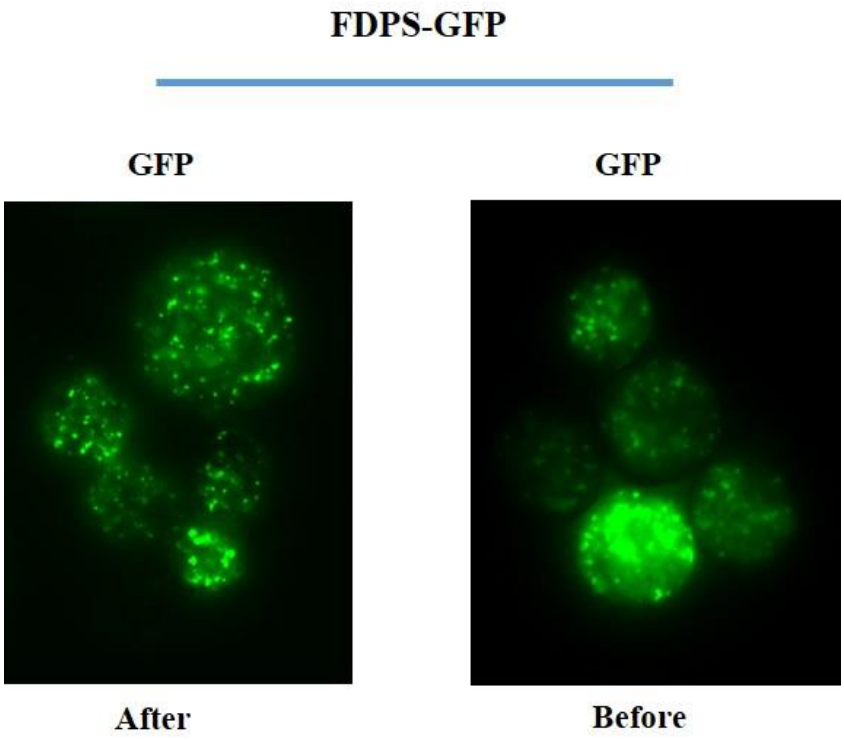


Figure 3.5: Expression of CRISPR plasmids in the *D. discoideum* knock out testing system. Amoebae expressing FDPS-GFP showed green fluorescent peroxisomes both before and after transformation with CRISPR plasmids for *Pex7* knockout.

The *Dd*CRISPR system was tested again against *GFP* itself in amoeba expressed GFP. Unfortunately, no difference was generated between the GFP expression level before and after transformation with the *Dd*CRISPR plasmids.

Why did the *Dd*CRISPR system not work?

1. Is the *wt*Cas9 and *h*Cas9 protein expressed in *D. discoideum*? Expression may not occur because there are codons in the *h*Cas9 and *wt*Cas9 that are rarely used by *D. discoideum*? If so, how many are there in each?
2. Is the guide RNA expressed?
3. Was Cas9 located in the nucleus?

Biochemical experiments and bioinformatics analysis were carried out to try to answer these questions. *Cas9* expression was determined by western blotting analysis. *D. discoideum* transformed with the first *Dd*CRISPR plasmid (either *h*Cas9 or *wt*Cas9) were grown in HL5 medium in the presence of the appropriate antibiotic, and harvested proteins were separated by the SDS-PAGE protein separation system. Immunoblot analysis using an anti-FLAG antibody was then used to detect Cas9 expression.

The gRNA transcription was investigated by using northern blot analysis. Four culture of *D. discoideum*, each transformed with a different version of the Second CRISPR plasmid were used for northern blot analysis. Total RNA was extracted (section 2.12.2) and loaded on to a polyacrylamide gel. RNA molecules were transferred to a nitrocellulose membrane as a specific radio-labelled 20bp probe was then used to detect the gRNA by hybridization. The radioactivity was detected on AGFA photo film when was scanned with BIO-RAD molecular imager. None of the tested samples showed a radioactive signal except for the positive control. Amino acid codon sequences analysis of the *wtCas9* and *hCas9* showed a high number of codons that are infrequently used be *D. discoideum* (Figure 3.6).

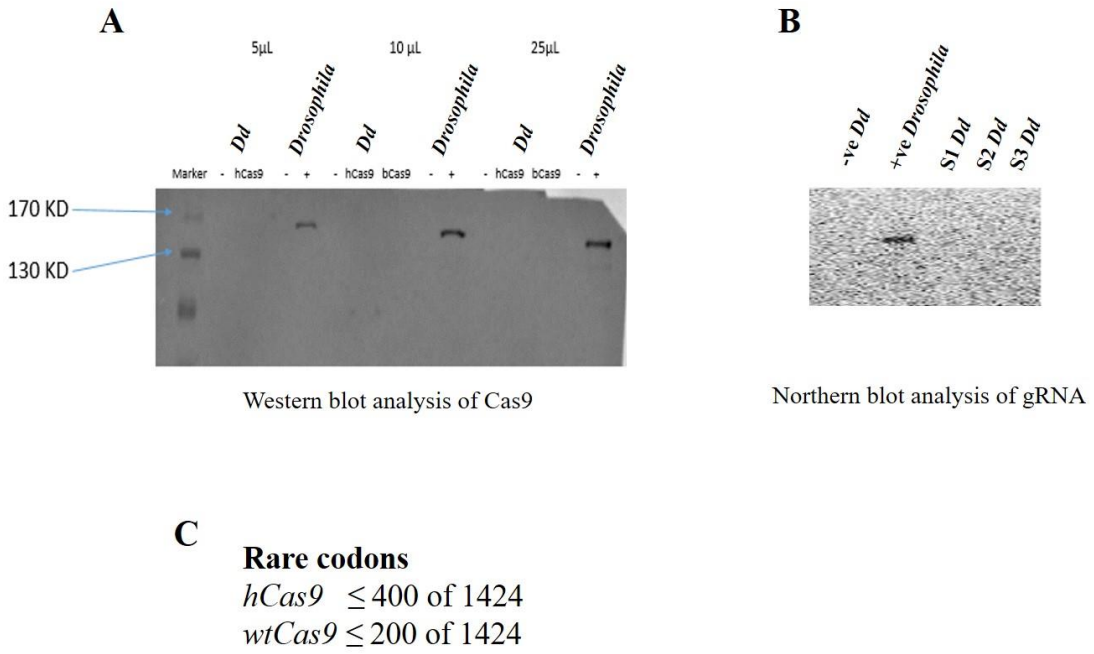


Figure 3.6*:* Western and northern blot and bioinformatics analysis of the *Dd*CRISPR system. A) This part of the figure shows the western blot analysis of humanised *hCas9* and bacterial cas9 *wtCas9* protein expression in transformed *D. discoideum*. Three volumes (5, 10, 25µl) of each sample were loaded on the SDS-PAGE gel. It is clear that only the positive (+ve) control, Drosophila S2R+ cells expressing CRISPR system, showed a protein band at the expected size for the Cas9 protein (150kD). B) Northern blot analysis showed only one band which was present in the RNA from the S2R+ cell line (+ve control). C) Bioinformatics analysis of all Cas9 sequences showed that there are at least 400 and 200 codons out of a total of 1424 codons in *h*Cas9 and *wt*Cas9 respectively that are rarely used by *D. discoideum* Messenger RNA.

To try to solve these problems, we decided to make our own *Cas9* gene which we called *DdCas9*. It was designed to include only codons commonly used by *D. discoideum*. In addition, codons in the SV40 NLS, neocleoplasmin NLS and the FLAG epitope were also optimised for expression in *D. discoideum* (Figure 3.4). Furthermore, new conditions for the PCR reaction were developed and a special DNA polymerase (MyFi) were used to obtain a longer sequence of *D. discoideum* U6 promoter, which is 450bp

in length. Four long nucleotide sequences were obtained, each similar but only slightly different in sequence from the *D. discoideum* U6 promoter. The differences arose because the PCR reaction catalysed by MyFi DNA polymerase did not give entirely faithful copies of the DNA template. All of them consist the *Dictyostelium* Upstream Sequence Element (DUSE) motif sequences (Figure 3.7).

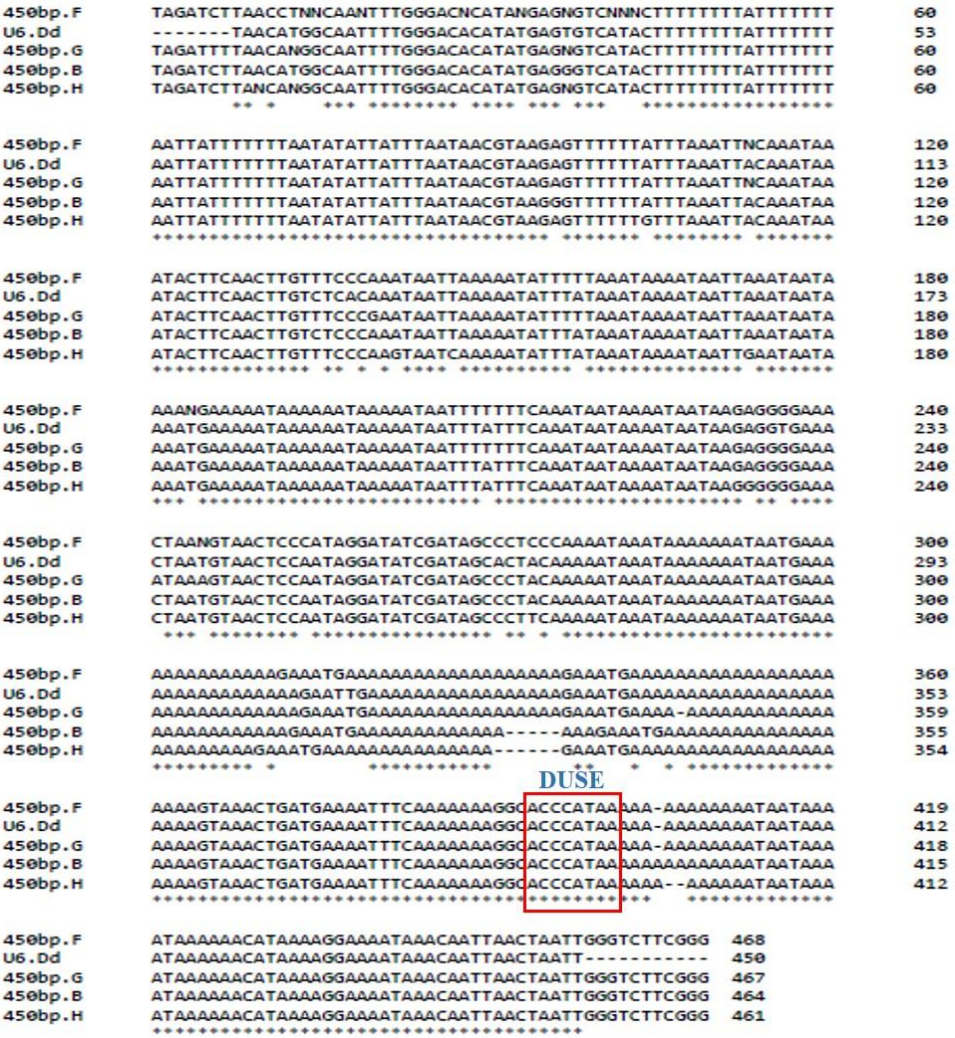


Figure 3.7*:* Alignment of *D. discoideum* U6 promoter sequences. Four long PCR products (450bp) of the *D. discoideum* U6 promoter alignment with the U6 promoter sequence from data base (Dictybase). The *Dictyostelium* Upstream Sequence Element (DUSE) domain was detected in all of them and highlighted in red box region.

DdCas9 and each of the four long U6 promoter PCR product (450bp) were cloned into *Dictyostelium* plasmids following the digestion ligation methods, to give pMUR33 and pMUR35A, B, C and D respectively. The resulting plasmids were introduced into *D. discoideum*. Cultures of the transformed amoebae were grown and used for western analysis. The results, as shown in figure 3.8 indicate that we were able to express *Dd*Cas9 and to generate an active 450bp U6 promoter in *D. discoideum*.

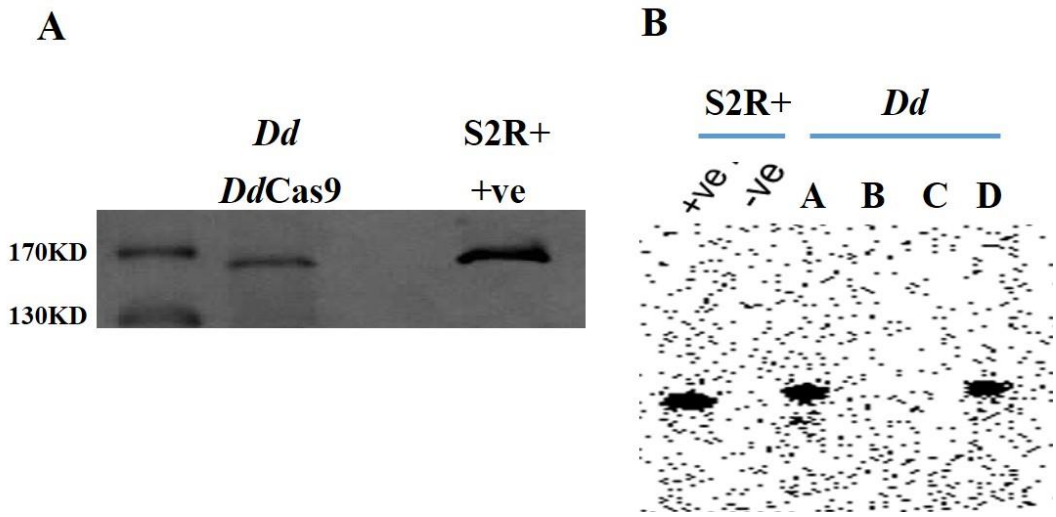


Figure 3.8*:* Biochemical analysis of the *Dd*CRISPR system. A) Western blot analysis of *D. discoideum Cas9* (*DdCas9*) expression. Amoebae expressed *DdCas9* as it shown by the clear band detected with the anti-FLAG antibody. It was of the expected size for the Cas9 protein. A Drosophila S2R+ cell line expressing Ca9 was used as a positive control. Approximately, the same number of cells were harvested and loaded on the SDS-PAGE gel and detected by immune blotting analysis. B) Northern blot analysis using the radio-active $\gamma[^{32}P]$ of the four gRNA (450bp U6 promoter in pMUR35). Only versions A and D were found to be expressed the gRNA in *D. discoideum*, in addition to the positive control (S2R+), while the untransformed S2R+ cell as a negative control (-ve)

The *Dd*Cas9 and the long U6 promoter (A and D) plasmids were used to attempt to knock out the *D. discoideum Pex7* and *atg1* gene. The *Atg1* gene, which is one of several members of the autophagy family of genes. Macroautophagy in eukaryotic cells is a mechanism for degradation of non-essential cellular components in starvation and development (Otto et al., 2003). The *Atg1* mutant strain of *D. discoideum* grows slower than the wild type and multicellular development is severely reduced. Disfunction in the *Atg1* gene would prevent *D. discoideum* aggregation and keep the social amoebae as single cell in starvation conditions (Mesquita et al., 2015). After the 20bp sequence of the target genes had been cloned to the gRNA plasmid, the two *Dd*CRISPR plasmids were introduced into *D. discoideum* individually. The transformed cells were grown with appropriate antibiotics. Unfortunately, there were no significant differences between the green fluorescent peroxisomal puncta in the transformed and untransformed amoebae in the attempt the knock out *Pex*7. Furthermore, no significant differences were found in fruiting body formation in attempt the knock out *Atg1*. Those results would suggest that the system failed to knock out the *Pex7* or the *atg1* genes. Although, the *DdCas9* and the gRNA were expressed in *D. discoideum* (Figure 3.8), it seems we could not make the CRISPR system operate in this organism.

3.3 Discussion

Homologous recombination was first used in this study in an attempt to knock out the *Pex7* gene in *Dictyostelium discoideum*. The blasticidin resistance gene was used as a selectable marker to create the knock out cassette. The cassette consisted of the blasticidin resistance gene flanked by more than 1400 nucleotides homologues to *Pex7* sequence and immediately upstream and downstream sequences (Figure 3.2). After the *D. discoideum* amoebae had been transformed with the *Pex7* knockout cassette, the transformed amoebae were able to grow in a medium containing 10μg/ml blasticidin. Furthermore, insertion of the knock out cassette into the gDNA was confirmed by using diagnostic PCR for the left and right flanks of the blasticidin gene. The cassette was found to be inserted in the open reading frame sequence of the *Pex7* gene. However, using primers annealing upstream and downstream (outside the targeting region) of the

Pex7 gene showed only a wild type pattern in the diagnostic PCR (Figure 3.3). Fluorescence microscopy showed similar green fluorescence puncta in amoebae before and after the knock out cassette transformation. Together those results demonstrated that the *Pex7* gene was still active.

Why did the classical knockout method (homologous recombination) not work?

The most plausible reason why *Pex7* remained active after the knockout cassette had been integrated successfully into the appropriate region of the gDNA, is that there are more than one copy of the *Pex7* gene in each amoeba. This would explain the results of the diagnostic PCR for the transformed and untransformed amoebae. Even though, the left and right flanks gave products when the gDNA of the transformed strain was used as a template the whole tested sequence (*Pex7* plus upstream and downstream flanks) showed only a wild type pattern in all samples (Figure 3.3). This suggests the presence of more than one copy of the *Pex7* gene in each amoeba, and that the blasticidin cassette was able to be inserted in at least one of them. As insertion of a single copy of a blasticidin resistance gene give resistance to the concentration of the blasticidin used in our selection procedure (Barth et al., 1998). In this case, two bands were expected to have been detected by the diagnostic PCR. The reason why the diagnostic PCR failed to show a product with the size of the mutant *Pex7*, might have been because of the difficulty in amplifying a very long A/T rich DNA sequence.

Because of these difficulties, we decided to develop a new technique -CRISPR- to attempt to knock out *Pex*7. We expected that using this system would solve the difficulty of knocking out multiple copies of a gene which cannot be done using homologous recombination.

CRISPR type II was set up in *D. discoideum*. This type of the system requires two main components; the Cas9 endonuclease and the gRNA, in which the latter directs the Cas9 to the target sequence. Each part of the system was carried on a separate plasmid. Three different versions of *Cas9* genes (wild type, humanized and Dictyosteliumized) endonucleases were used. The *wtCas9* and the *hCas9* could not be expressed in the amoebae (because they containing many codons rarely used by *D. discoideum*), whereas the *Dd*Cas9 expressed successfully (Figure 3.8A). It was not possible to check directly

that Cas9 was entering nuclei. However, it was possible to show that the NLS sequences being used on Cas9 were able to direct GFP into the nuclei. Thus, it appeared that the SV40 NLS and nucleoplasmin NLS sequences were functional in *D. discoideum* and should have been sending Cas9 into nuclei. In addition, two lengths of U6 promoter (100bp and 450bp) were used to induce expression of the gRNA. Northern blot analysis showed that only when amoebae were transformed with the long version was the gRNA expressed (Figure 3.8B).

Nevertheless, although the two components of the CRISPR system (*Dd*Cas9 and gRNA) were successfully introduced into *D. discoideum*, we could not knockout the *Pex7* gene. This might have been because *Pex7* is an essential gene in *D. discoideum*. The Pex7 protein is a soluble receptor and transporter of PTS2-containing proteins to the peroxisomes (Petriv et al., 2004). Knocking out this gene would lead to proteins depending on this pathway in the cytoplasm instead of in the peroxisomes. However, it is unclear for which *D. discoideum* protein using a PTS2 pathway to enter peroxisomes, a peroxisomal location is essential.

If the *Pex*7 gene is essential, it would not be possible to knock it out either of homologous recombination or CRISPR. This led us to test the CRISPR system by attempting to knock out a different gene. Autophagy1 (*Atg1*) was chosen to test the CRISPR system activity in *D. discoideum* for the following reasons; it is a non-essential gene and has been deleted before from amoebae. Furthermore, the *Atg1* mutant cells showed the clear phenotype that they could not form fruiting bodies. For attempts to inactivate the *Atg*1 gene, three different sequences in *Atg1* were chosen as targets. The 20bp of each target were included in the gRNA plasmid and amoebae expressing *Dd*Cas9 were transformed separately with the gRNA plasmids. Even though, all the three transformed cultures of amoebae were able to grow in the presence of the selective antibiotics, they formed fruiting bodies on starvation i.e. *Atg1* had not been inactivated.

All the experiments taken together appeared to indicate that it is not possible to introduce a successful CRISPR system into *D. discoideum*. Because it was shown that the amoebae transformed with CRISPR system plasmids expressed cas9 and the gRNA and it appeared that the NLS would send the Cas9 into the nucleus, it is unclear why CRISPR does not operate in the way as expressed in *D. discoideum*.

Chapter four “A novel topogenic route inserts a tail anchored membrane protein to the inside of the peroxisomal membrane”.

4.1 Introduction

Sterol biosynthesis is an essential pathway in all eukaryotic kingdoms, which provides cells with the fundamental material for controlling plasma membrane permeability. Furthermore, based on the type of the cells, it supplies them with assorted products such as cholesterol in humans, which are used as a precursor for the production of other important products such as bile acids and steroid hormones (Nes, 2011).

Squalene synthase (SQS) is the first enzyme in sterol biosynthesis. It is a bifunctional enzyme that first forms pre-squalene diphosphate (PSPP) from two hydrophilic farnesyl diphosphate (FDP) molecules. Then it catalyses the reduction of PSPP into one hydrophobic squalene (SQ) molecule using nicotinamide adenine dinucleotide phosphate (NADPH) as cofactor (Figure 4.1) (Jiang et al., 2017; Ohtake et al., 2014; Rong et al., 2016; Takatsuji et al., 1982; Zhao et al., 2010). In most eukaryotes, including yeast, plants, animals and humans, SQS is localised to the ER (Busquets et al., 2008; Fegueur et al., 1991; Robinson et al., 1993; Stamellos et al., 1993). It contains a C-terminal hydrophobic amino acids sequence that serves as a tail anchor, which integrates the protein into the ER membrane (Jiang et al., 2015; LoGrasso et al., 1993). Structural analysis of SQS revealed the presence of a channel that passes through the middle of the protein. The tail anchor allows one opening of the channel to be very close to the membrane, whereas the opening on the opposite side of the channel faces the cytosol. This architecture facilitate directed movement of the hydrophilic precursor (FDP) from the cytosol, through the active site where it is converted to a hydrophobic product that is subsequently released in the membrane (Pandit et al., 2000).

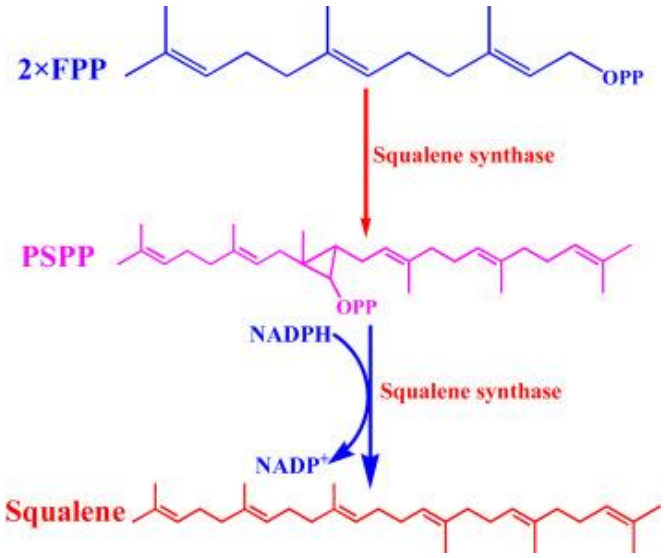


Figure 4.1: Reaction overview of squalene synthase. SQS is a bifunctional enzyme that catalyses the reaction whereby two FDP molecules form PSPP. Subsequently, PSPP is reduced to squalene using NADPH as an electron donor. Figure reproduced from (Liu et al., 2017; Nakashima et al., 1995) with permission.

Studies in a variety of organisms have shown that sterol biosynthesis occurs at the ER membrane. However, we have found that *D. discoideum* farnesyl diphosphate synthase is a peroxisomal enzyme (Nuttall et al., 2012). This prompted us to further investigate the organisation of the sterol biosynthesis pathway in this amoeba and initial studies established that the first four enzymes of the pathway are indeed located in peroxisomes (Mayyadah Al-Kuwayti, Donald Watts, unpublished data).

Most peroxisomal matrix proteins and membrane proteins are posttranslationally targetted from the cytosol. Genetic dissection of the targeting and import pathways have revealed that matrix protein import and membrane protein import depend on separate machineries. This is further reflected by the presence of targeting signals that are specific for either matrix proteins (PTS1 and PTS2) or membrane proteins (mPTS) (Kim and Hettema, 2015). Furthermore, some matrix proteins lacking a PTS are delivered to peroxisomes by piggy-backing onto other PTS-containing proteins such as Pnc1(Saryi et al., 2017). However, no similar evidences of PMPs has been found (for review see (Kim and Hettema, 2015). PMPs are therefore thought to be synthesised in the cytosol, picked up by a specific receptor (Pex19) that delivers them to the peroxisomal membrane

by docking onto Pex3. Subsequently, the PMPs integrates directly into membrane from the cytosolic side. No PMPs have been described that insert into the peroxisomal membrane from the matrix side. We therefore studied the topology of *Dd*SQS, taking into account that its hydrophilic substrate is synthesised inside peroxisomes.

The aim of this chapter is to understand *Dd*SQS topogenesis. To address this aim, I had the following objectives:

1) Investigating the transport pathway of the protein to peroxisomes.

3) Identifying the peroxisomal targeting signal in *Dd*SQS.

4) Determining the topology of the enzyme in peroxisomes.

4.2 Result

4.2.1 Squalene synthase (the first enzyme in the Sterol biosynthesis) isa peroxisomal enzyme in *D. discoideum* amoebae

Farnesyl diphosphate synthase (FDPS) is the last enzyme in the mevalonate pathway is localised to the peroxisomal matrix (Nuttall et al., 2012). Its product is used in the first step in the sterol biosynthesis, by squalene synthase (SQS). Interestingly, *D. discoideum* SQS (*Dd*SQS) contains a putative peroxisomal targeting signal type 1 (PTS1) at its extreme C-terminus consisting of amino acids sequence Ser, Lys, and Leu (S, K, and L) which is conserved to other *Dictyostelium* species (Figure 4.2).

Figure 4.2: Schematic structure of alignment of the last 10 amino acid of the SQS in *Dictyostelium* species. *Dictyostelium* species and two other species (*D. discoideum*, *D. lacteum*, *D. fasciculatum*, *D. purpureum*, *Polysphondylium pallidum* and *Acetostelium subglobosum*) have a conserved putative PTS1 at the very C-terminal end

Indeed, previously, it was found that a fusion of GFP with *Dd*SQS localises to peroxisomes in *D. discoideum* amoebae (Alkuwyti, 2014). This experiment was repeated and confirmed this observation (Figure 4.3).

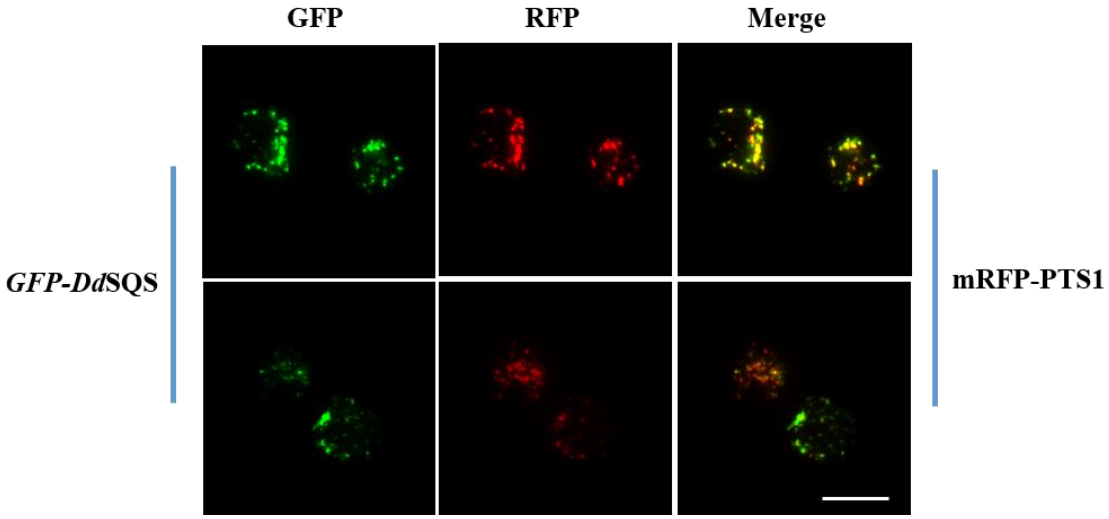


Figure 4.3: The intracellular location of *Dd*SQS in *D. discoideum*. Amoeba expressing GFP-*Dd*SQS and the peroxisomal marker mRFP-PTS1 were imaged with epifluorescence microscopu. The green and red fluorescence pattern overlap extensively. Bar 10µm.

4.2.2 Amino acids sequence analysis of *Dd*SQS

In previously studies, SQS has been reported to be an ER membrane protein (Busquets et al., 2008; Fegueur et al., 1991; Robinson et al., 1993; Stamellos et al., 1993). Therefore, our observation that *Dd*SQS is a peroxisomal protein containing a putative peroxisomal targeting signal for the peroxisomal matrix was surprising.

An alignment of human SQS and *Dd*SQS clearly shows that the proteins are highly conserved, including the presence of functional and structural motifs that were previously shown to be important see Figure 4.4, (Pandit et al., 2000).

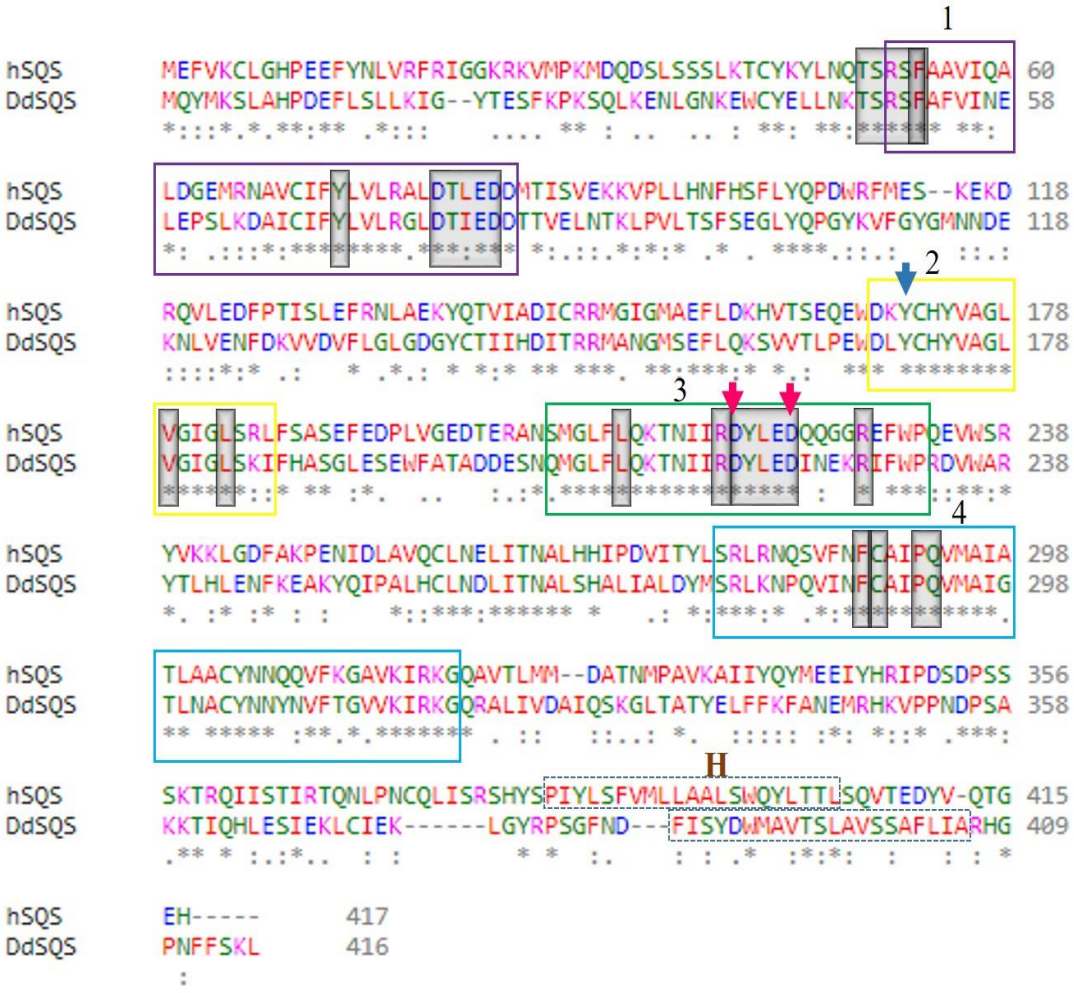


Figure 4.4: alignment of the amino acid sequences of *Dd*SQS and *h*SQS. The Conserved domains are surrounded by four coloured rectangles and labelled as 1, 2, 3 and 4. The transmembrane domain helix sequences were marked with capital English letter H. blue arrow, catalyses first reaction to form pre-SQ-PP: red arrows, FDP binding, grey boxes, important for structural function.

The *Dd*SQS amino acids sequence was analysed using different sequence analyses programmes. Using the Phyr2 online services for protein prediction, modelling and analysis, the three-dimensional structure for *Dd*SQS was predicted based on the structure of other SQSs (Kelley et al., 2015) (Figure 4.5). The predicted 3D structure of the protein showed that like other SQSs, *Dd*SQS comprises a bundle of helices. Interestingly, a hydrophobic helix is predicted seven amino acids N-terminal of the PTS1. The equivalent helix in mammalian SQS anchors the enzyme to the ER membrane. Also in *Dd*SQS, this hydrophobic helix is predicted to be a transmembrane domain (TMD) according to the TMD prediction programme TMHMM (Figure 4.6) and UniProtKB web site.

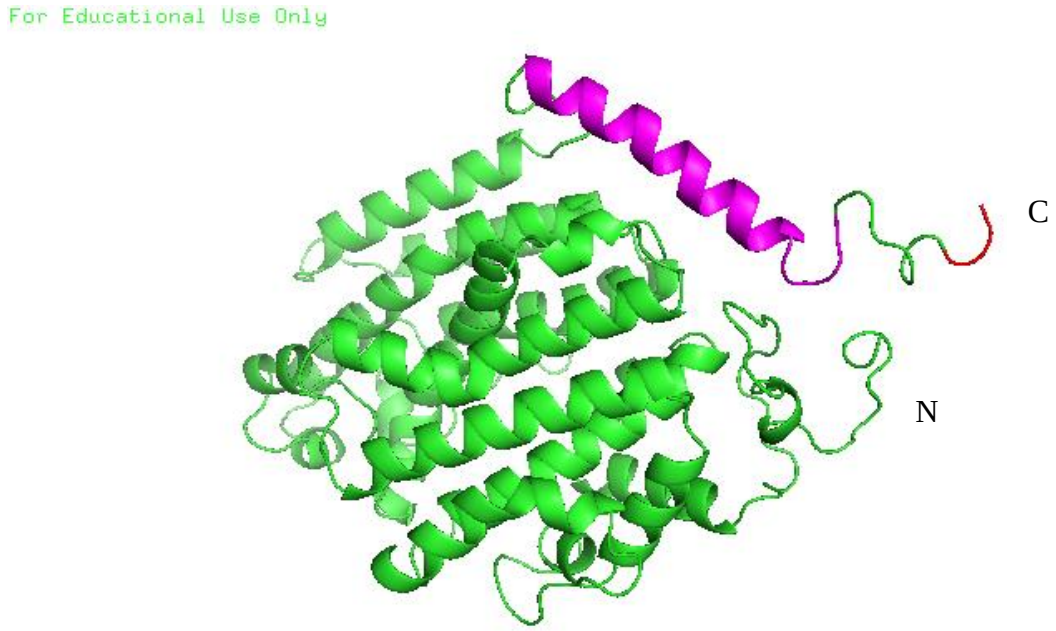


Figure 4.5: A cartoon structure of the three-dimensional (3D) structure of *Dd*SQS. The 3D predicted structure of DdSQS was displayed using Pymol to analyse the Phyr2 predicted profile. The letter N and C refer to the amino and carboxyl terminus of the protein, respectively. The PTS1 is drawn as a red stick at the very C-terminal end. The predicted TMD, which is displayed as a pink helix, ends seven amino acids N-terminal to the PTS1.

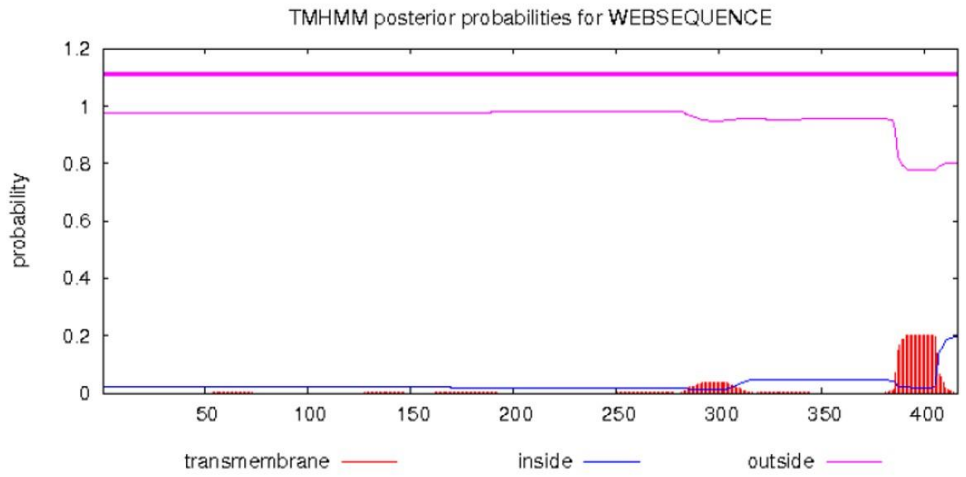


Figure 4.6: Transmembrane domain (TMD) prediction of *Dd*SQS. The red area represents the hydrophobic region of the protein and the probability of forming a TMD. The x-axis represents the amino acid sequence.

4.2.3 Identification of the peroxisomal targeting signal in *Dd*SQS and its peroxisomal targeting pathway

Although the PTS1 is conserved, deletion of it does not prevent import in amoebae. (Alkuwayti 2014 and Figure 4.7).

To identify other regions in the protein required for targeting, two more deletion mutants were constructed and analysed. First, a C-terminal truncation was constructed where the putative TMD and the PTS1 were deleted (*Dd*SQSΔHΔPTS1). The second construct comprises an internal deletion of the putative TMD (*Dd*SQSΔH).

The above constructed expression plasmids were transformed to amoebae already expressing RFP-PTS1. The full-length protein GFP-*Dd*SQS, GFP-*Dd*SQSΔH, and GFP-*Dd*SQSΔPTS1 showed a green fluorescence punctate pattern that mostly co-localised with the red fluorescence pattern of the peroxisomal marker. However, the GFP-*Dd*SQSΔHΔPTS1 showed no structure-linked fluorescence as it was displaying a diffuse fluorescence pattern typical of a cytosolic protein (Figure 4.7). The ability of *Dd*SQSΔPTS1 to reach peroxisomes, while the GFP-*Dd*SQSΔHΔPTS1 mislocalised to the cytosol in *D. discoideum* implicates a role for the helix in protein transport.

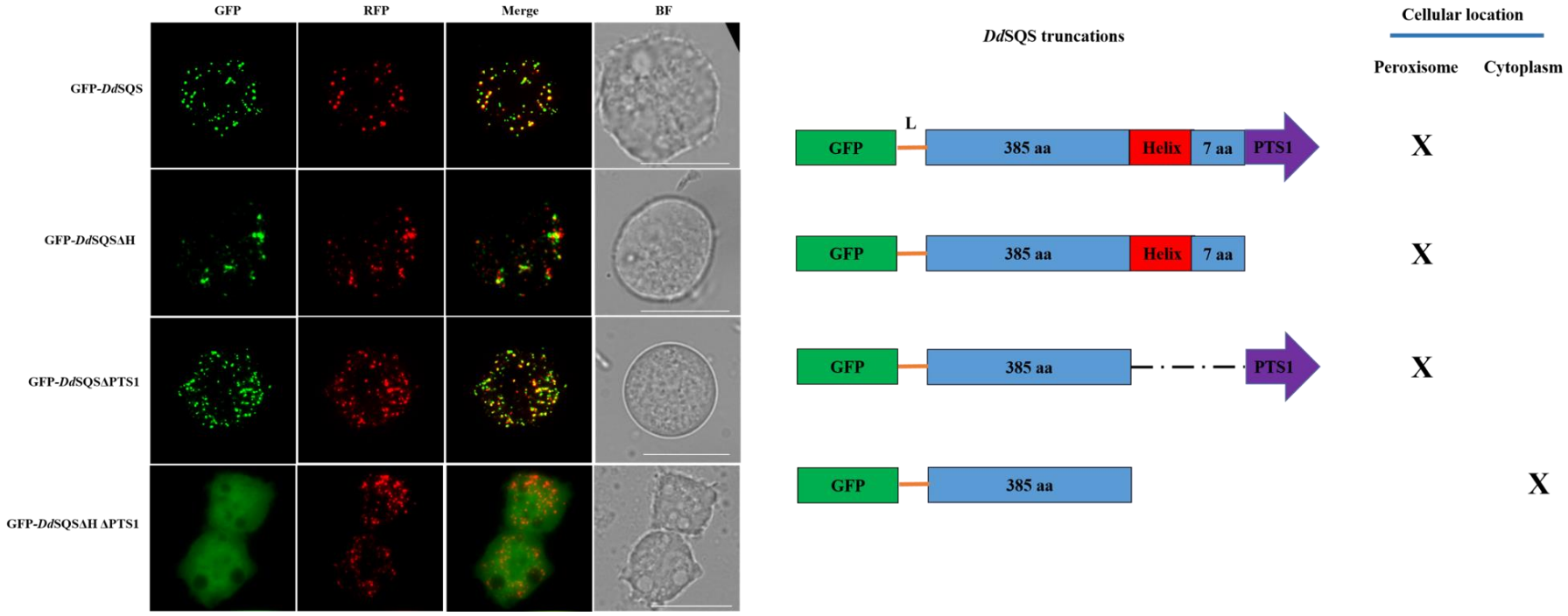


Figure 4.7: Truncation analysis of DdSQS and the effect on the subcellular location. The figure shows the role of the PTS1 import pathway in *D. discoideum*. A) Amoebae expressing mRFP-PTS1 as a peroxisomal marker and either full length GFP-*Dd*SQS, GFP-DdSQSΔPTS1, GFP-*Dd*SQSΔHelix and GFP-*Dd*SQSΔHΔPTS1. The first three colocalise with mRFP-PTS1. Only the last one localises to cytosol. B) cartoon of truncation constructs and the cellular location of each one.

4.2.4 GxxxG and (Small)xxx(Small) motif

The amino acid sequence comprising the helix was analysed extensively. It contains many small amino acids such as Ala and Ser, (8 of 21 in total). The sequence AxxxA and SxxxS are related to GxxxG motifs. These motifs consist of any three amino acid residues surrounded by two small amino acid residues. These motifs are supposed to allow dimer formation in TMDs (for a review see (Teese and Langosch, 2015)). The AxxxA motif extended to form a AxxxAxxxA motif, which we called a Gly zipper-like motif. Interestingly, the motif is highly conserved in all *Dictyostelium* species and in some very closely related species as well (Figure 4.8). These findings suggest that the Gly-zipper-like motif plays a particularly significant role in the structure of this putative TMD. According to this and the results described in the previous section, we postulate that the helix allows *Dd*SQS to dimerise prior to import. *Dd*SQS lacking its PTS1 still localises to peroxiosmes as it can dimerise with endogenous *Dd*SQS, which allows piggy-back import. The first step to test the homo-dimer formation in living *D. discoideum* cells would be by knocking out the endogenous gene of interest. That would stop formation of a dimer between the endogenous and the extra-chromosomal gene. As SQS is a key enzyme in the sterol biosynthesis (Jiang et al., 2017), it is highly expected to be an essential gene. This co-import hypothesis can also be tested by expression of the various version of DdSQS in a heterologous system and this approach is described below.

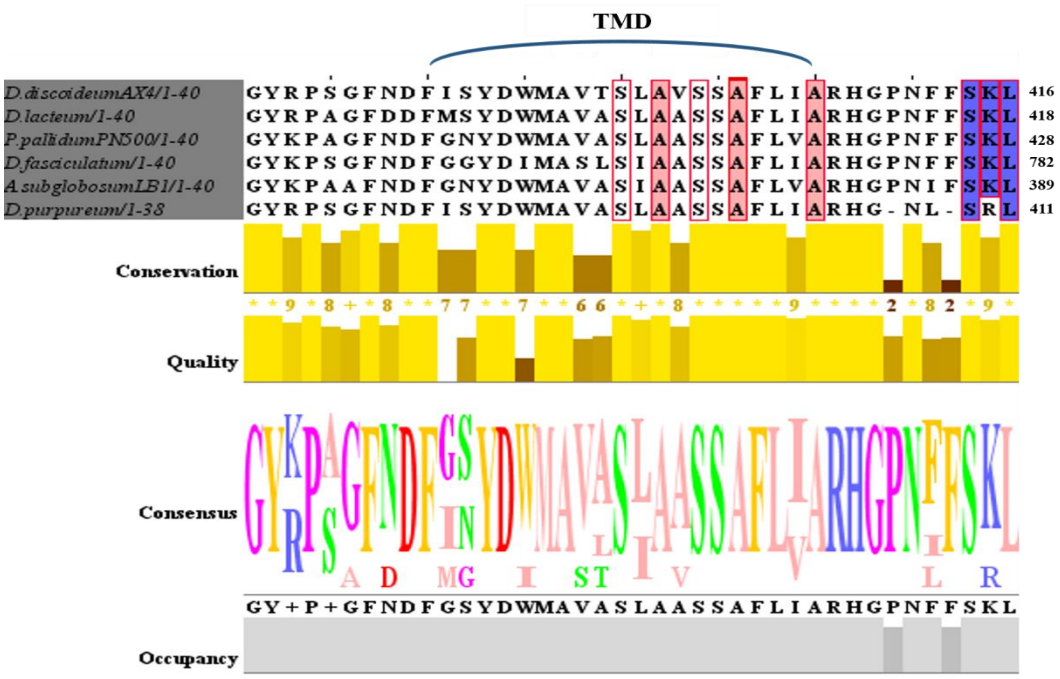


Figure 4.8: The putative transmembrane domain of *Dictyostelium* SQSs has a conserved Gly-zipper-like motif. Amino acid sequence alignment of the C-terminal 40 amino acids containing the putative TM domains (helix) and PTS1 of SQS in *Dictyostelium* species and two other organisms. All the analysed organisms (*D. discoideum*, *D. lacteum*, *D. fasciculatum*, *D. purpureum*, *Polysphondylium pallidum* and *Acetostelium subglobosum*) show a highly conserved G-zipper-like motif, which is identified as AxxxAxxxA (boxed in light red), PTS1 indicated in blue. B) Sequence logo of the alignment of all the species above.

4.2.5 Characterisation of the role of *Dd*SQS C-terminus in peroxisomal targetting using a heterologous system.

We used *Saccharomyces cerevisiae* as heterologous expression host to test whether *Dd*SQS lacking its PTS1 is co-imported with full length *Dd*SQS. First however, we confirmed that DdSQS is imported into yeast peroxisomes via the PTS1 pathway (Figure 4.9).

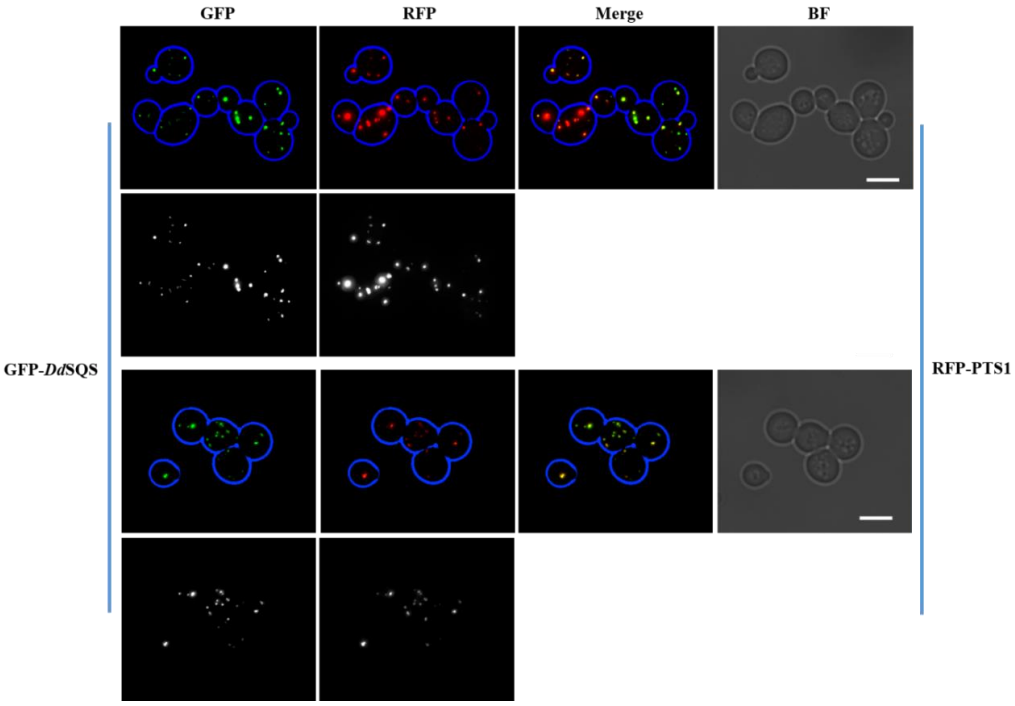


Figure 4.9: *Dd*SQS expressed in *S. cerevisiae* cells localises to peroxisomes. Yeast cells were cotransformed with a plasmid encoding mRFP-PTS1 and GFP-*Dd*SQS. The latter was under control of the inducible *GAL1/10* promoter. Cells were grown for 2h on galactose medium to induce expression. An overlapping red and green punctate pattern is observed with epifluorescence microscopy. BF, bright field. Bar 5µm.

Subsequently, GFP-*Dd*SQS was expressed in cells with defects in peroxisomal protein import such as *Pex1Δ*, *Pex5Δ*, *Pex7Δ* and *Pex13Δ* or peroxisome membrane biogenesis such as *Pex19Δ*. In *Pex19* cells, when peroxisomal membranes are absent, GFP-*Dd*SQS was, as predicted, mislocalised. In mutants where all matrix protein import is blocked but membrane biogenesis is functional, such as in *Pex1* and *Pex13* cells, again *Dd*SQS mislocalised. This implies that *Dd*SQS targets to peroxisomes via a matrix protein import pathway. To investigate whether *Dd*SQS follows the PTS1 or PTS2 pathway, we analysed its localisation in *Pex5* and *Pex7* cells. In *Pex7* cells, import is unaffected but in *Pex5* cells, *Dd*SQS is mis-localised. Collectively, these observations indicate that *Dd*SQS is targeted to peroxisomes via the PTS1 pathway. This is in line with the presence of a PTS1 (Figure 4.10).

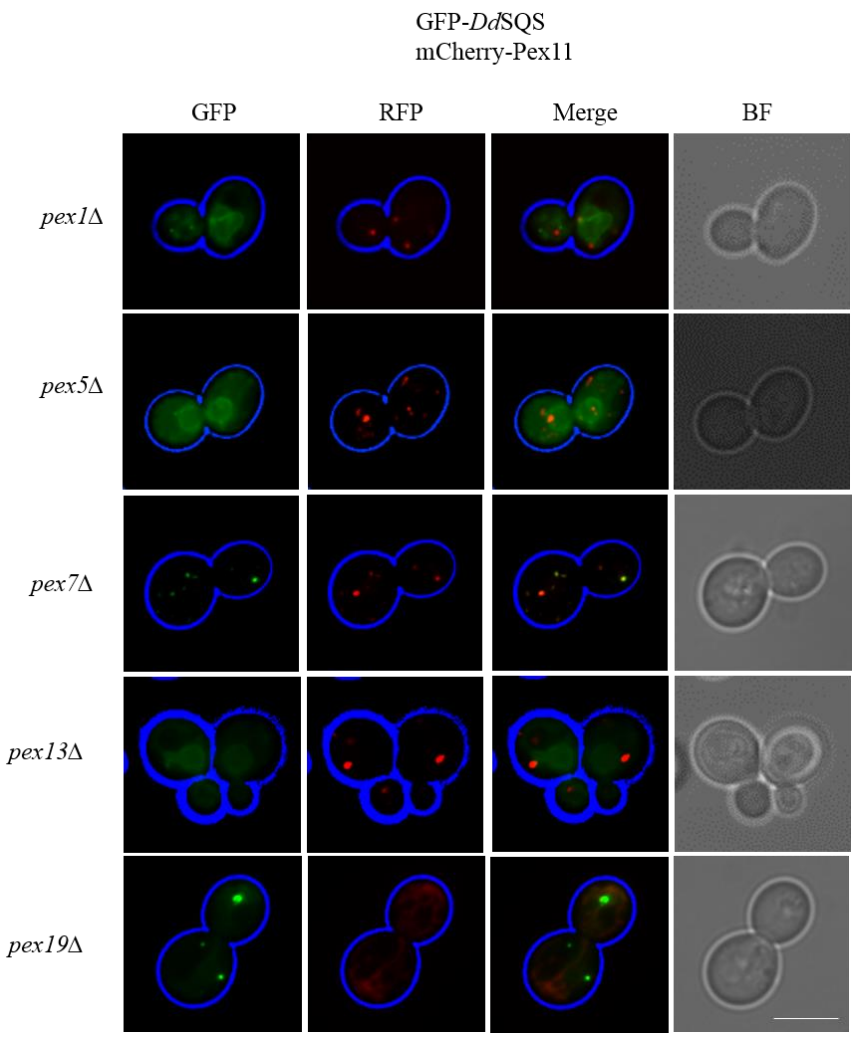


Figure 4.10: Investigation of the pathway used by *Dd*SQS to reach the peroxisomes. The various mutant yeast strains expressed GFP tagged *Dd*SQS (GFP-*Dd*SQS) together with mCherry-Pex11 as a peroxisomal marker.

Next we tested whether *Dd*SQS import is dependent on its PTS1 in yeast. All the protein truncations mentioned before were transferred to conditional yeast expression plasmids, transformed individually into wild type *S. cerevisiae* that are also expressing mRFP-PTS1. Full-length GFP-*Dd*SQS and GFP-*Dd*SQSΔH localise to yeast peroxisomes as revealed by their co-localisation to mRFP-PTS1 (Figure 4.11). However, the GFP-*Dd*SQSΔPTS1 was located in the ER, plasma membrane and cytosolic pool. Remarkably, the *Dd*SQSΔHΔPTS1 showed a clear cytosolic green fluorescence only (Figure 4.11, Table 4.1). We conclude that in yeast cells, the PTS1 is essential for import of *Dd*SQS.

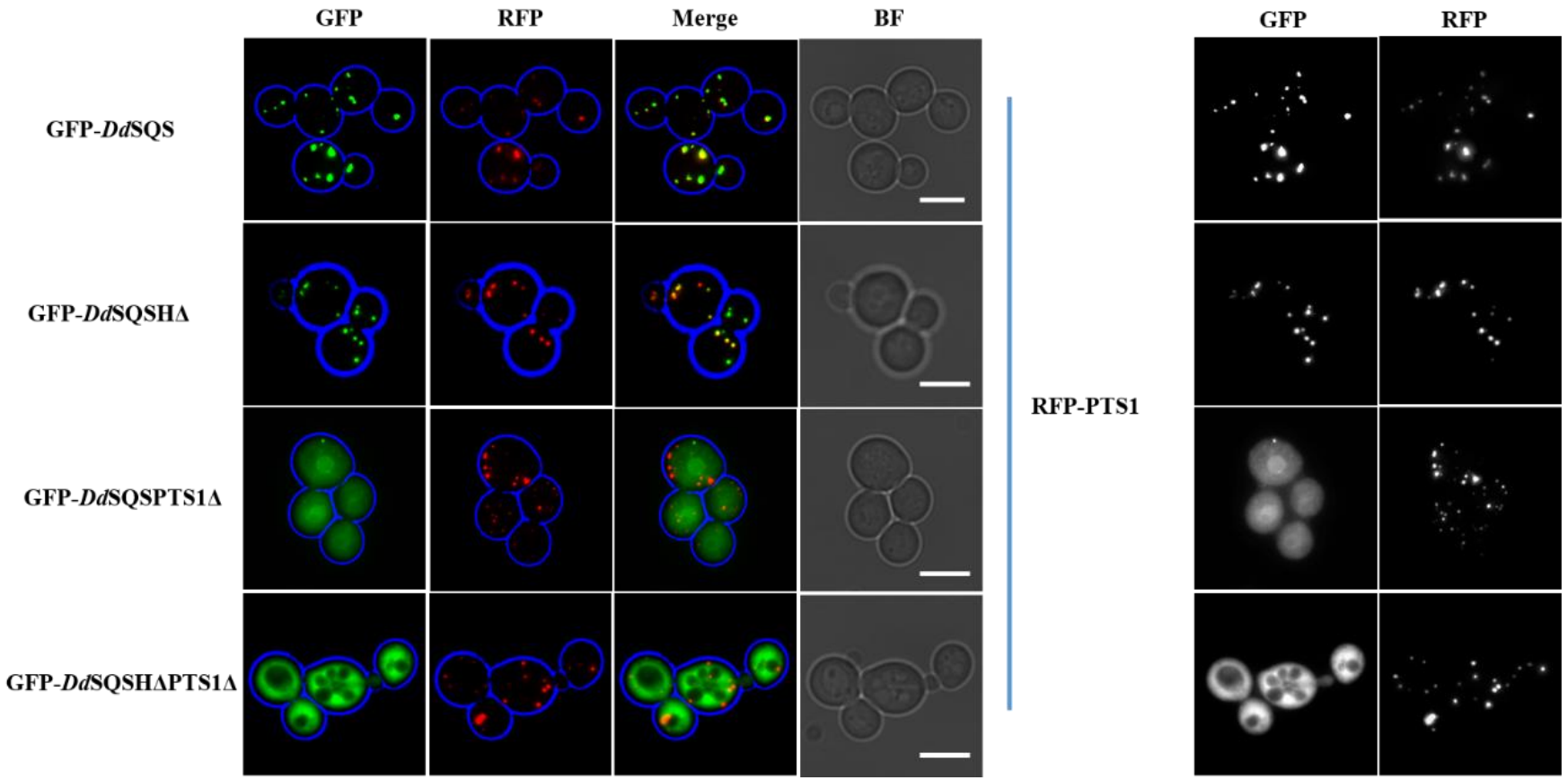


Figure 4.11: investigation of the cellular location of full length *Dd*SQS and some truncations. Fluorescence microscopy of a wild type yeast expressed full length GFP-DdSQS, GFP-DdSQSΔH, GFP-DdSQSΔPTS1, and GFP-DdSQSΔHΔPTS1 individually, beside peroxisomal marker (RFP-PTS1). full length GFP-*Dd*SQS and GFP-*Dd*SQSΔH were forming a green fluorecence punctate colocalised with the red fluorecence of the marker. However, GFP-*Dd*SQSΔPTS1 was located in the ER, plasma membrane and the rest of it either forming an agregate or remainded in the cytoplasm. While the cells expressed GFP-*Dd*SQSΔHΔPTS1 showed a pure cytoplasm green fluorecence. Bar 5µm.

Table 4.1: Displaying the cellular location and the level of expression of the full length, ΔH, ΔPTS1 and ΔHΔPTS1 of *Dd*SQS truncations.

*Dd*SQS truncations tagged to GFP	Molecular location	Level of expression
Whole *Dd*SQS	Peoxisomes	Good +++
*Dd*SQSΔH	Peoxisomes	Good +++
*Dd*SQSΔPTS1	Another membrane (plasma membrane), ER, and the rest in the cytoplasm	Good +++
*Dd*SQSΔHΔPTS1	Cytosol	Good +++

If *Dd*SQS can be imported as a homo-dimer then one prediction would be that import of GFP-*Dd*SQSΔPTS1 is restored upon expression of full-length *Dd*SQS. The construct was built by amplifying the *Dd*SQS open reading frame and inserting it into a yeast conditional expression plasmid containing an mRFP tag at the N-terminal end (RFP-*Dd*SQS). The resulting plasmid is called pMUR42 (table 2.). Co-expression of mRFP-*Dd*SQS and GFP-*Dd*SQSΔPTS1 restored the green fluorescent punctate pattern, that co-localises with the mRFP pattern. This indicates that indeed *Dd*SQS can dimerise in the cytosol and be co-imported as such into the peroxisom (Figure 4.12).

In light with our observations that the helix plays a role in targeting to peroxisomes in amoebae, that *Dd*SQS can be imported as a dimer, and the helix contains a potential smallxxxsmall dimerisation motif, we tested whether the helix is required for dimer import using further fluorescence microscopical investigations. A *Dd*SQS open reading frame without helix was created and inserted into a yeast N-terminal mRFP tagging plasmid. The resulting plasmid (pMUR94) was used to express *Dd*SQSΔH fused to RFP (RFP-*Dd*SQSΔH). The co-transformation method was repeated again. All GFP tagged *Dd*SQS truncations were transformed into a wild type *S. cerevisiae* in addition to pMUR94. As expected, the GFP-*Dd*SQS and GFP-*Dd*SQSΔH were located in yeast peroxisomes, which is driven by the PTS1. In contrast, GFP-*Dd*SQSΔHΔPTS1 remained in the cytoplasm. The most surprising aspect of the data is the molecular location of the GFP-*Dd*SQSΔPTS1. This recombinant protein does not reach the peroxisome. It is located in the ER, cytoplasm and some times forming an aggregate (Figure 4.13). This shows that neither *Dd*SQSΔPTS1 is able to reach the peroxisome, nor that the *Dd*SQSΔH could take it there by piggy-backing. From this, it can be concluded that *Dd*SQS can be imported as a dimer and that the helix is important for this dimer dependent co-import.

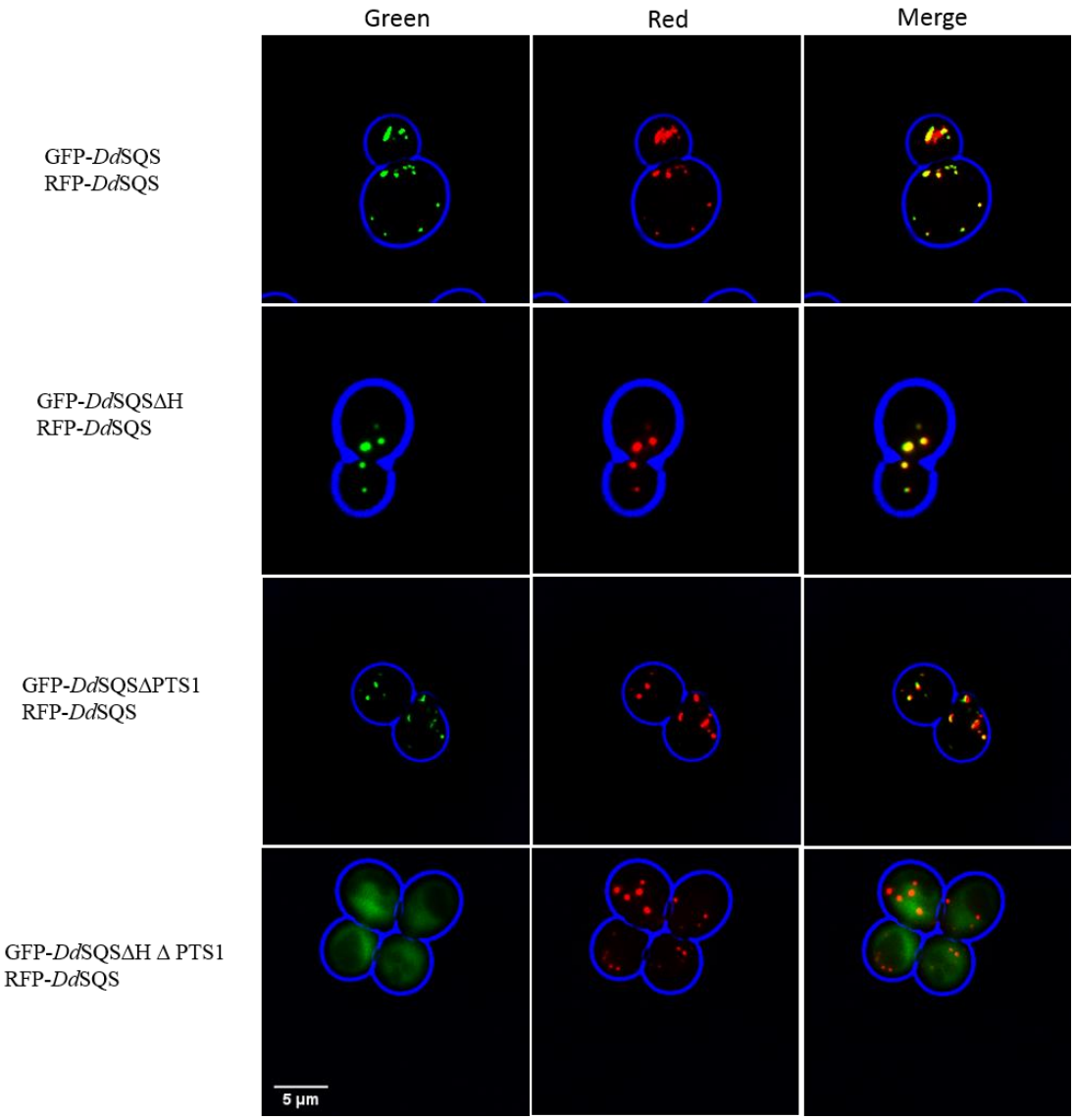


Figure 4.12: Role for helix in co-import. *Dd*SQS truncations tagged with GFP at the N-terminus were expressed in *S. cerevisiae* in the presence of full-length mRFP-*Dd*SQS. All truncations that contain either their own PTS1 or the helix localised to peroxisomes. Bar 5µm

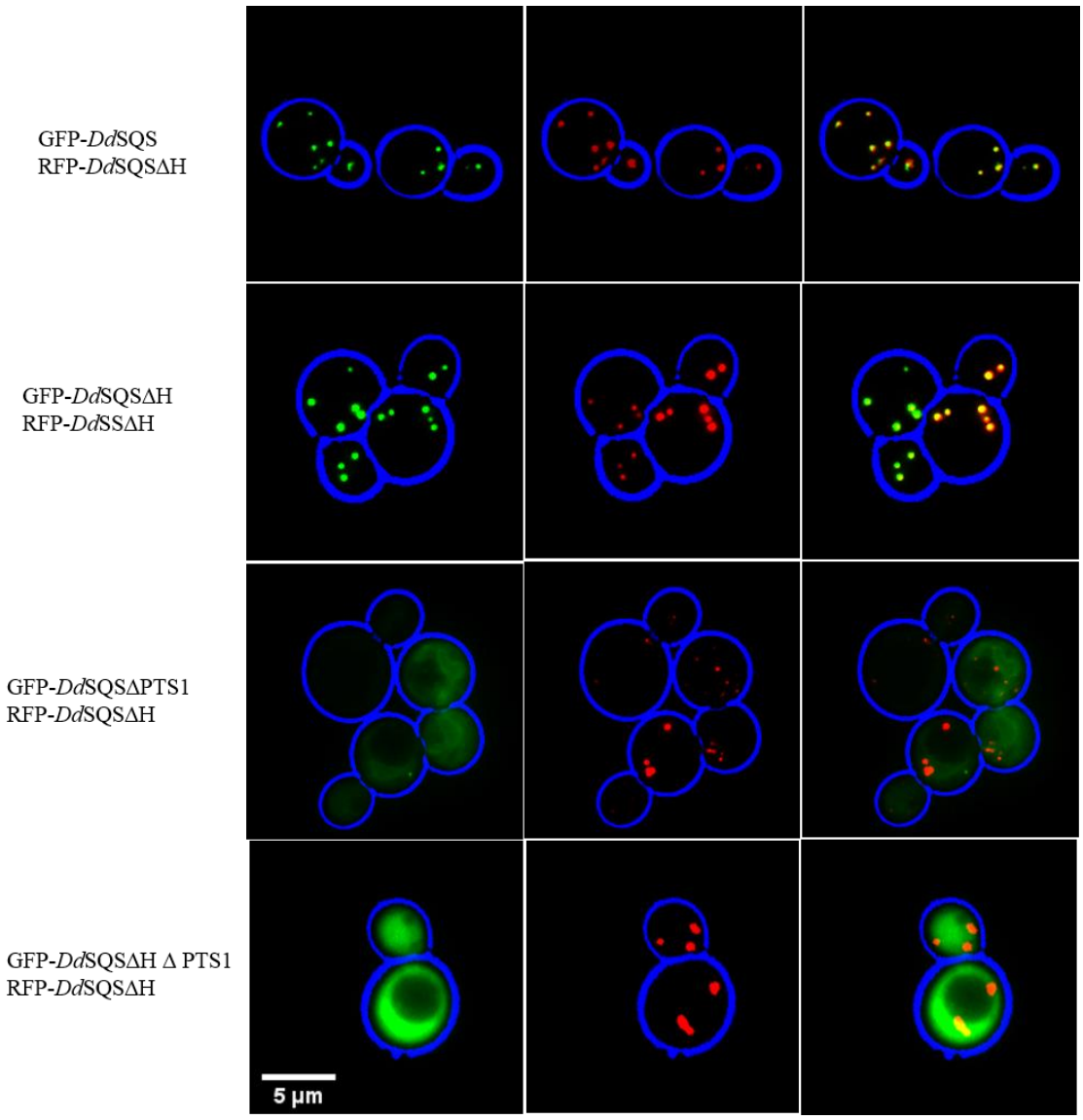


Figure 4.13: Piggy-back import of *Dd*SQS lacking a PTS1 depends on the c-terminal helix in partner protein containing the PTS1. *Dd*SQS truncations tagged with GFP were expressed in S. cerevisiae in presence of the *Dd*SQSΔH tagged with red fluorescence protein RFP-*Dd*SQSΔH. It is clear that RFP-*Dd*SQSΔH could neither be co-imported with GFP-*Dd*SQSΔPTS1 nor GFP-DdSQSΔHΔPTS1. Bar 5µm.

All the previous data support the model that the carboxy terminal end of *Dd*SQS contains a complex peroxisomal targeting signal that directs homo-dimerization and PTS1-dependent import.

4.2.6 Further analysis of the requirements for *Dd*SQS co-import.

Is the full length of *Dd*SQSΔPTS1 necessary for forming the dimer?

*Dd*SQS co-import (bipartail signalling) was investigated further in yeast to identify the minimal length of *Dd*SQSΔPTS1 that is still co-imported with full-length *Dd*SQS using fluorescence microscopy. For this purpose, four truncations of *Dd*SQS were made. The following amino acids were deleted from the N-terminal end of the protein (Δ1-150, Δ1-251, Δ1-312, and Δ1-351). Furthermore, the PTS1 were removed from all the truncations by PCR. The amplified DNA fragments were ligated into a yeast expression plasmid, which encoded a GFP at N terminal end of each insert. Yeast cells were transformed with each truncation in the presence of full-length RFP-*Dd*SQS. Remarkably, only the full-length GFP tagged protein lacking its PTS1 (*Dd*SQSΔPTS1) was located in the peroxisome. However, no green fluorescence was detected from all other truncations, suggesting that the fusions are unstable (Figure 4.14).

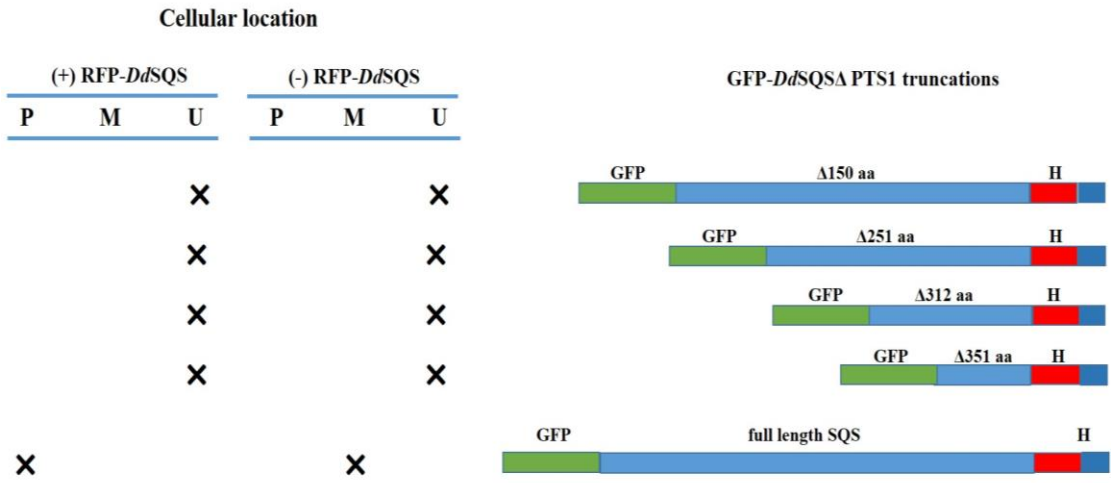


Figure 4.14: Cellular location of four GFP-*Dd*SQSΔPTS1 N-terminal truncations. Four different truncations of GFP-*Dd*SQSΔPTS1 in addition to the full length were tagged with GFP, expressed in S. cerevisiae in the presence and absence of full-length RFP-*Dd*SQS. Only *Dd*SQSΔPTS1 was located in the peroxisome when it has been expressed together with RFP-*Dd*SQS. However, it mislocalised when it was expressed alone. None of the other truncations could be detected by the fluorescence microscope in either the presence or absence of RFP-*Dd*SQS. Cellular location P, M and U mean peroxisome, mislocalised and undetectable fluorescence signal, respectively.

4.2.7 *D. discoideum* Squalene synthase topology.

Generally, all the known PTS1-containing proteins are localised to the peroxisomal matrix. However, SQS of mammals and plants and fungi are tail-anchored proteins of the ER membrane (Pandit et al., 2000). In line with, bioinformatic analysis using TMHMM server v 2, phyer2 and Uniprot predicted that the helix preceding the PTS1 that is required for co-import, to be a transmembrane helix. This suggests that DdSQS is a peroxisomal tail-anchored membrane protein. The transmembrane helix is predicted to be 21 amino acid in length starting from phenylalanine in position 386 (F^{386}) till alanine at position 406 (A^{406}) which is 7 amino acid before the PTS1 (see figure 4.8). First, we tested whether DdSQS is a membrane protein. For this purpose, two different truncations of *Dd*SQS (*Dd*SQSΔH and *Dd*SQSΔPTS1) in addition to the full length tagged to GFP were used. Cells from exponential growing cultures of *D. discoideum* amoebae expressing each one of the above mentioned fusion proteins were harvested and disrupted by sonication. A crude total homogenate, which contains all organellar fractions, was centrifuged at ultra-speed to separate the soluble from the insoluble fraction. The pellet, which contains membranes and associated proteins, was extracted with high salt (1M NaCl). This treatment releases many unspecific membrane-binding proteins and some of the specific membrane-associated proteins (Fujiki et al., 1982). Western blot analysis was used to determine the distribution of the fusion proteins among the different fractions. Full-length GFP-*Dd*SQS was present in the final pellet, which strongly suggests that DdSQS is tightly associated with membranes. Interestingly, GFP-*Dd*SQSΔH was detected in the first supernatant, just like the matrix protein control (GFP-PTS1). While GFP-*Dd*SQSΔPTS1 showed the same behaviour as the full-length protein (Figure 4.15). This means the predicted transmembrane helix is required for fractionation with membranes and can potentially serve as a tail anchor. This hypothesis was further tested by studying heterologously expressed *Dd*SQS in yeast.

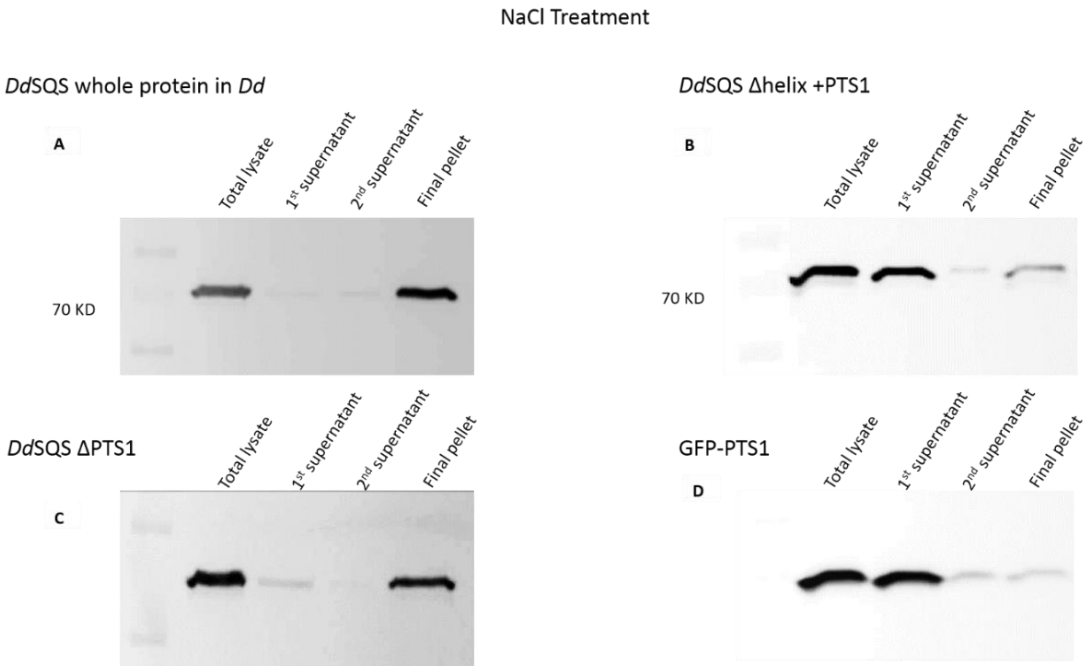


Figure 4.15: 1M NaCl treatment of amoebae transformed with full length and different truncations of GFP tagged squalene synthase. Amoebae were broken open by sonication for 10 secs, at 10 amplitude microns. Total lysate (TL) was taken directly from the sonicated cells. The first supernatant refers to the supernatant left after ultra-speed centrifugation at 4ºC for 1 h at 80,000 rpm (using Optima MAX-E Ultracentrifuge, TLA 100.2 rotor) of the TL. The resulting pellet was treated with 1M NaCl for 30mins at 0ºC and centrifuged at 80,000 rpm (as above) to give the second supernatant and the final pellet. Fractions were analysed by western blotting using anti-GFP. Equivalent fractions were loaded. A) GFP-*Dd*SQS, B) GFP-*Dd*SQSΔHelix, C) GFP-*Dd*SQSΔPTS1 and D) GFP-PTS1 as a matrix protein control. Anti GFP antibody was used as a primary antibody in this western blot experiment.

4.2.8 Using *Saccharomyces cerevisiae* to study *Dd*SQS membrane association

Our results on the targeting of *Dd*SQS so far are remarkable as no peroxisomal membrane proteins have been described that require the matrix protein import pathways for their targeting. In mutants where matrix protein import is blocked, membrane proteins still assemble into peroxisomal membranes. However, *Dd*SQS appears to

follow the PTS1 pathway (see Figure 4.10 and Figure 4.2) but also behaves as a membrane protein (Figure 4.15).

S. cerevisiae cells were used in this part of the study for the following reasons:

1) To test whether the topogenic information can function in different organisms.

2) Using an organism free of the endogenous *Dd*SQS to avoid homodimerisation

3) Techniques for fractionation are better developed and a good range of controls is available.

4) Yeast cells grow 5-6 times faster than *D. discoideum*, which means saving time.

As described above (Figure 4.10), GFP-DdSQS localises to yeast peroxisomes dependent on the PTS1 pathway. Cells expressing *Dd*SQS were harvested and broken open using the glass bead method (see Material and methods). Samples were centrifuged at 4ºC for 1 h at 80,000rpm (using Optima MAX-E Ultracentrifuge, TLA 100.2 rotor), the pellets were treated with 1M NaCl for 30min at 0ºC as described above. Western blot analysis of the various fractions showed that GFP-*Dd*SQS fractionates mainly in the membrane pellet (Figure 4.16). This is in line with the results obtained in amoebae and reveals that the topogenic information od *Dd*SQS is functional in yeast.

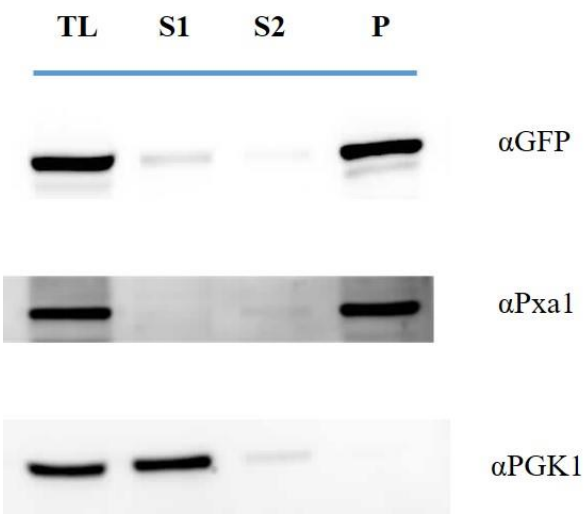


Figure 4.16: Investigation of membrane association of *Dd*SQS in yeast. Total lysate (TL) was taken directly from homogenised yeast cells. First supernatant refers to the supernatant left after ultra-speed centrifugation at 80,000 rpm (using Optima MAX-E Ultracentrifuge, TLA 100.2 rotor) of the TL. The resulting pellet was treated with 1M NaCl for 30 mins at 0ºC and centrifuged at 80,000 rpm (at same conditions) to give the second supernatant and the final pellet. Fractions were analysed by western blotting using anti-GFP, anti-Pgk1 and anti-Pxa1. Equivalent fractions were loaded. Pgk1 is marker for cytosolic proteins. Pxa1 was used as a marker for integral peroxisomal membrane proteins. TL, total cell lysate; S1, first supernatant; S2, second supernatant and P, final pellet (insoluble fraction).

In order to characterise the minimal sequence that direct GFP-*Dd*SQS to the peroxisomal membrane, a series of N-terminal deletion constructs were generated and analysed by fluorescence microscopy and fractionation experiments.

Depending on the length of each truncation, the following fused protein GFP-104aa*Dd*SQS, GFP-65aa*Dd*SQS, GFP-37aa*Dd*SQS and GFP-24aa*Dd*SQS were made. First of all, we analysed the location of each truncation. Plasmids carrying all the above open reading frame of GFP fused proteins were transformed to *S. cerevisiae*, already expressing the red peroxisomal marker (RFP-PTS1). Remarkably, all of them showed a clear co-localization of the green and red fluorescence (Figure 4.17), indicating that these fusions are targeted to peroxisomes and are stably expressed. This is in contrast to the version without PTS1, as these were undetectable (see Figure 4.14)

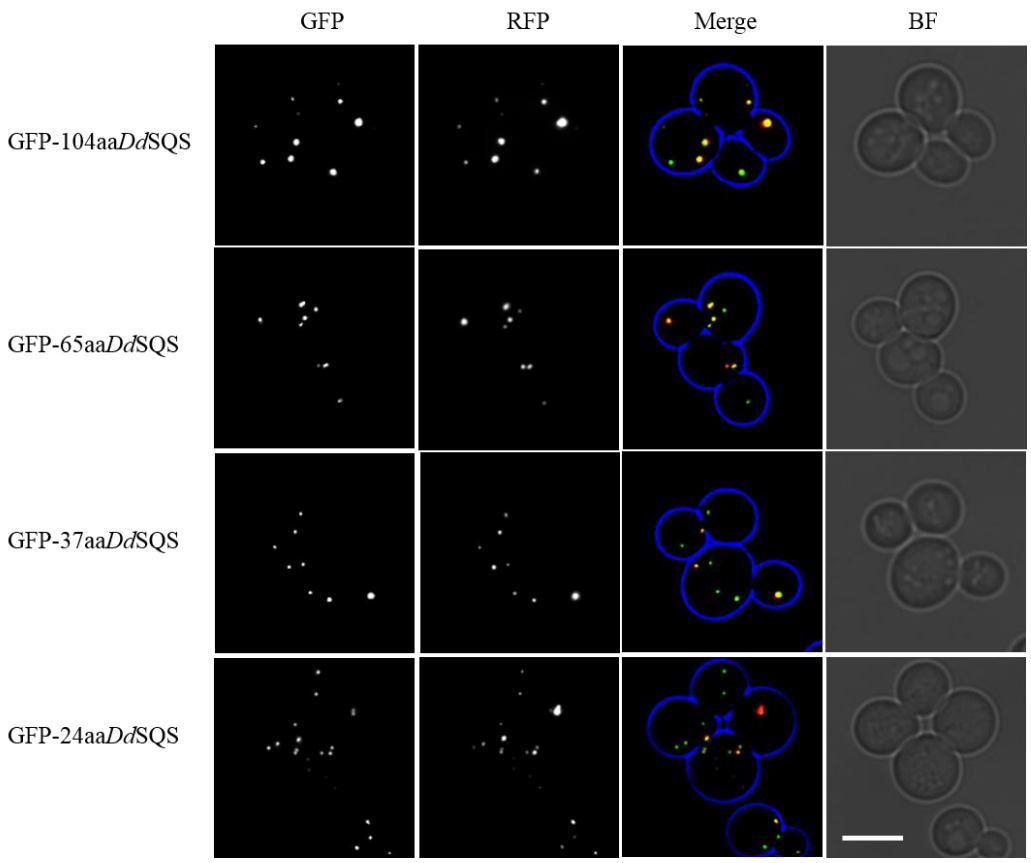


Figure 4.17: Cellular location of *Dd*SQS truncations expressed in the S. cerevisiae. Wild type yeast cells expressing four different truncations of GFP-*Dd*SQS. All the truncations colocalised with the red fluorescence of the peroxisomal marker (RFP-PTS1). Bar, 5µm.

Subsequently, the membrane association of the various versions of GFP-*Dd*SQS was tested using high salt (1M NaCl) and alkaline (100mM Na_2CO_3, pH 11.5) extractions of membrane fractions. Western blot analysis of NaCl-treated membranes using anti-GFP antibody showed that all *Dd*SQS truncations remained in the final pellet fractions suggesting that those truncations are tightly associated with the peroxisomal membrane (Figure 4.18A). Furthermore, carbonate treatment could not release the protein from the peroxisomal membrane either (Figure 4.18B). Together, all these findings suggest that all *Dd*SQS truncations are strongly bound to the membrane behave like peroxisomal integral membrane proteins. Note that GFP-appended with the last 37 or 24 amino acid residues of SQS migrate as two bands, with the top band being membrane associated and the lower band soluble. At this moment, we don't know whether the lower band is a breakdown product, but we consider this likely.

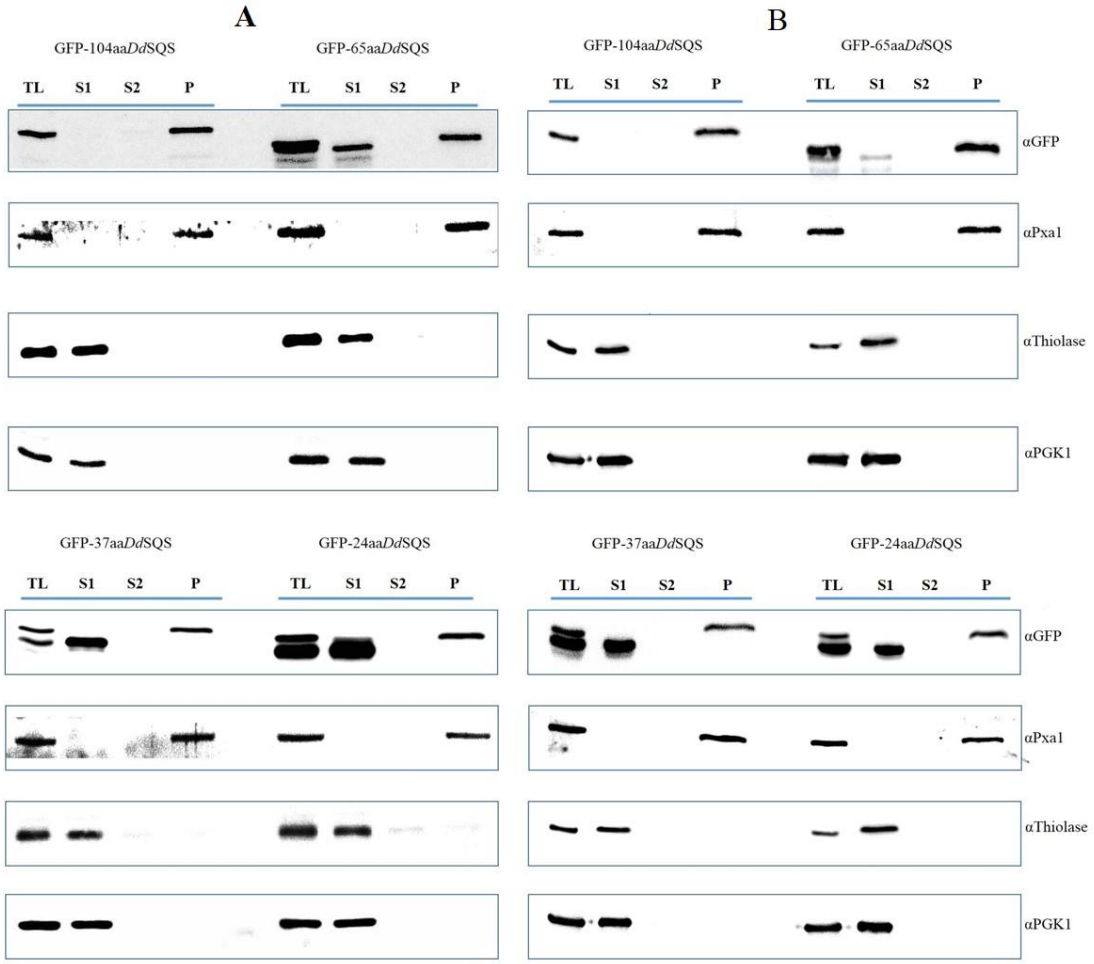


Figure 4.18: *Dd*SQS truncations are membrane associated proteins. Yeast cells were transformed with the following plasmids pMUR57 (GFP-*Dd*SQS104aa), pMUR58 (GFP-*Dd*SQS65aa), pMUR61 (GFP-DdSQS37aa) and pMUR62 (GFP-*Dd*SQS24aa). All open reading frames were expressed behind the *GAL1/10* promoter, induced for 2hour by adding 2% galactose. Total cell lysate was obtained by beads beating. The insoluble membrane fraction was separated from the cytosol by ultracentrifugation. Membrane fractions were extracted either with: A) high salt solution, 1M NaCl, where all the GFP-*Dd*SQS truncations showed a membrane association, or B) 100mM Na_2CO_3 where they all behaved like an integral membrane protein, indicates breakdown product. Each sample was derived from approximately the same cells number. All samples were applied to SDS-PAGE, followed by western blot analysis using anti-GFP antibody. Peroxisomal membranes were detected with anti-Pxa1, soluble cytosolic and peroxisomal matrix proteins were detected by anti-Pgk1 and anti-thiolase, respectively. TL (total cell lysate), S1 (first supernatant), S2 (second supernatant) and P (final pellet, insoluble fraction).

4.2.9 Topology of *Dd*SQS

Since the *Dd*SQS follows the PTS1 import pathway but behaves like a membrane protein, we investigated the topology of the protein in the peroxisomal membrane. We aimed to elucidate the orientation of *Dd*SQS with respect to the peroxisomal membrane. In other words, is the enzyme inserted into the membrane from the peroxisomal matrix side or the cytosolic side (Figure 4.19)? For this purpose, *in vitro* protease protection analysis and an *in vivo* auxin degron system technique were used.

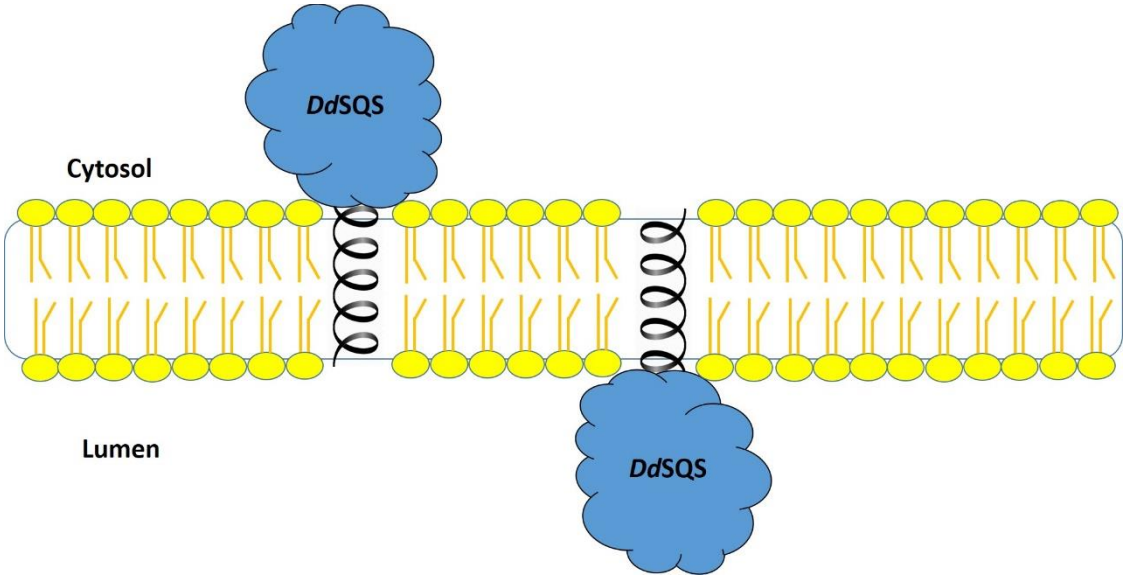


Figure 4.19: schematic diagram of the potential orientation *Dd*SQS in the peroxisomal membrane.

4.2.9.1 Protease protection assay

Protease protection analysis was used to determine the protein orientation in the peroxisomal membrane. *Dd*SQS samples for protease protection assay were prepared by generating spheroplast of *S.* cerevisiae cells overexpressing GFP-*Dd*SQS (section 2.5.9). The cells were homogenised gently by douncing (section 2.5.5). Homogenates were incubated with 0.05U proteinase K agarose beads (proteinase K-agarose from Tritirachium album, P9290-10UN, Lot041M5162V Sigma) at room temperature for the indicated times (for detail see section 2.9). GFP-*Dd*SQS was largely protected from proteinase K digestion in the absence of the detergent Triton X-100 after 60min incubation at room temperature. However, the addition of detergent resulted in rapid breakdown of the protein. This result demonstrates that *Dd*SQS was protected by the peroxisomal membrane which means it is located inside the peroxisome facing the lumen.

Pgk1 is a cytosolic protein in yeast. It is used as a positive control for the proteinase K enzymatic activity in the presence and absence of the Triton X-100. It showed a high sensitivity to the proteinase K in both cases. While the enzyme thiolase, which is a well known peroxisomal matrix protein, was chosen to determine the proteinase K activity against membrane-protected proteins in the absence of detergent. Indeed, thiolase was protected from the proteinase K digestion activity in the absence of detergent, however, it was very sensitive to proteinase K in the presence of detergent. Pxa1 is a yeast peroxisomal integral membrane protein. The C-terminal half of this protein is exposed to the cytoplasm. Pxa1 was used as a membrane-binding protein control to determine the difference in the digestion sensitivity between the membrane and soluble proteins. Pxa1 was digested after 5min incubation with proteinase K in both presence and absence of detergent, and was undetectable after 30 mins. Importantly, *Dd*SQS and all the control proteins used in this assay remained intact in untreated samples throughout the experiment (Figure 4.20).

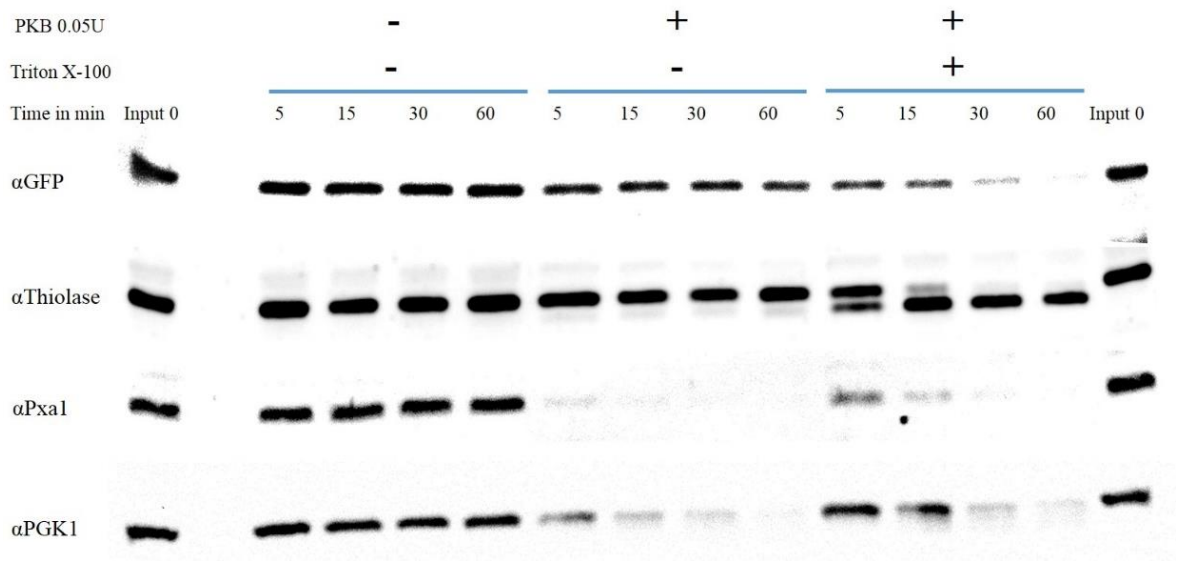


Figure 4.20: Protease protection assay testing the orientation of *Dd*SQS in yeast peroxisomes. Yeast cells expressing GFP-*Dd*SQS were grown on oleate medium, converted to spheroplast and broken open using a dounce homogenizer. Equal volumes of homogenate were incubated with 0.05U proteinase K in the presence (+) or absence (-) of 1% Triton X-100 for 0, 5, 15, 30 and 60 mins at RT. In addition, untreated samples were incubated at RT for the same incubation time points. The reaction was stopped by addition of trichloroacetic acid (TCA) to a final concentration of 10%. Equal fractions of reaction samples were analysed by SDS-PAGE and immunoblotting using anti-GFP, anti-thiolase, anti-Pxa1 and anti-Pgk1antibodies.

4.2.9.2 Auxin degron system as a tool to study protein topology *in vivo*

Plants have a unique protein depletion system in which the plant hormone auxin induces the breakdown of specific proteins containing the IAA domain. Auxin stabilises the interaction between Tir1 and an IAA domain. Tir1 recruit the SCF E3-ubiquitin ligase complex. This ubiquitin ligase then ubiquitinates the IAA-domain-containing proteins which lead to their auxin-dependent degradation. This auxin response is absent in other eukaryotic kingdoms; however, they share the SCF E3-ligase-dependent ubiquitination pathway of protein degradation. This allows for the use of the auxin-inducible

degradation (AID) system in non-plant cells. To get this to work in other organisms, two requirements have to be met. First, the cells need to express Tir1 heterologously and second, a protein of interest will need to be fused to the IAA domain. Upon addition of auxin, Tir1 forms a bridge between an IAA-tagged substrate protein and the SCF E3-ligase, thereby resulting in the auxin-dependent degradation of the target protein by the proteasome. A target protein that is localised inside an organelle, will be inaccessible for Tir1, SCF and proteasome and will, therefore, be protected from auxin-dependent degradation *in vivo*.

WT and *pex1* strains expressing Tir1 and expressing either IAA-GFP-PTS1 or 3HA-IAA-*Dd*SQS were used in an *in vivo* protease protection experiment. Pgk1 was used as a control cytosolic protein that is expected to be insensitive to auxin-induced degradation as it is not tagged with the IAA domain. Growing cells were treated with 500µM auxin and samples were examined at 30min intervals. IAA-GFP-PTS1 was stable upon auxin incubation in WT cells but was rapidly degraded in *pex1* cells. This shows that we can use the auxin degron system to test whether a protein is present inside peroxisomes or in the cytosol *in vivo*. Next, we applied this to 3HA-IAA-*Dd*SQS. In wt cells, it remained easily detectable throughout the 2h auxin treatment, indicating it was at least partially protected. In *pex1* cells, where 3HA-IAA-*Dd*SQS is no longer targeted to peroxisomes, it was almost completely degraded (Figure 4.21).

These experiments support the in vitro protease protection experiments and strongly suggest that *Dd*SQS is localised to the matrix side of the peroxisomal membrane.

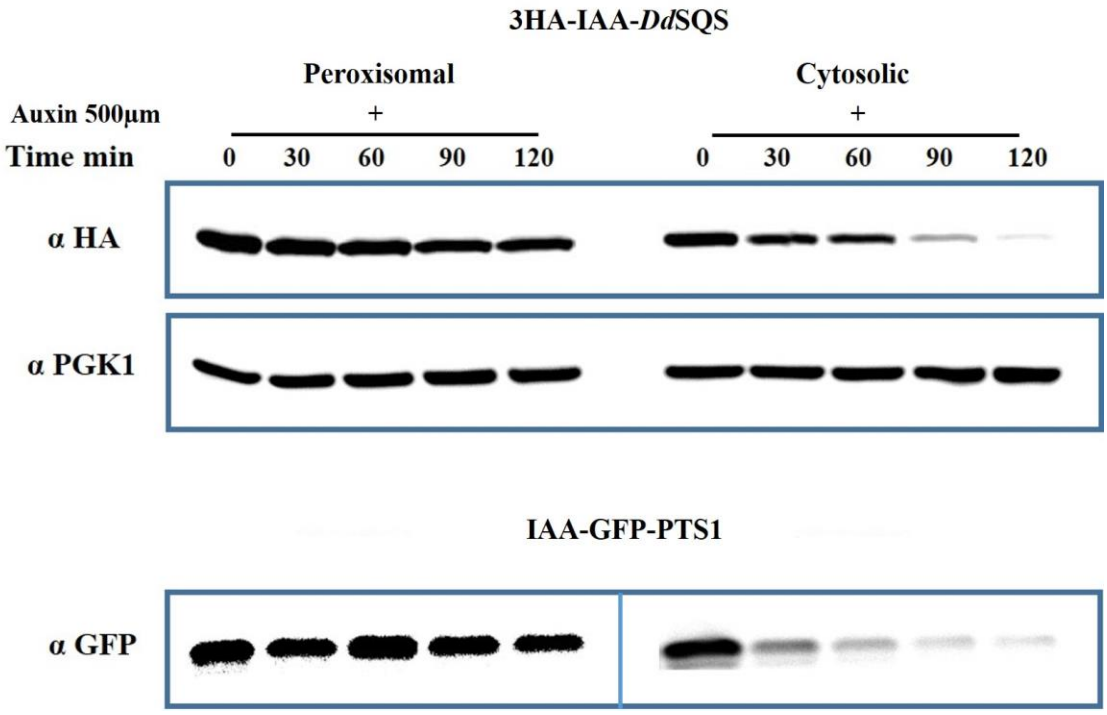


Figure 4.21: Auxin degron analysis of *Dd*SQS investigating the protein orientation in the peroxisomal membrane. The indicated yeast strain expressing IAA-GFP-PTS1 or 3HA-IAA-*Dd*SQS were treated with auxin at final concentration 500µM. Samples were collected after 0, 30, 60, 90 and 120 mins. Total protein was extracted followed TCA method. Approximately equal amount of total protein was analysed by western blotting using anti-HA, anti-Pgk1 and anti-GFP antibodies.

4.2.10 Intracellular location and protein pattern

Since we have confirmed the cellular location and membrane association of *Dd*SQS in *D. discoideum* and showed that the protein has a conserved and active PTS1 in yeast. Furthermore, it is membrane-associated protein in yeast as well. We aim to use an independent method to test membrane localisation. For this we can't use conventional fluorescence microscopy as the size of peroxisomes is below the resolution of our microscopy set up. Therefore, we used structured illumination microscopy that can achieve a much higher resolution and should be able to resolve the peroxisomal membrane from the matrix. This increased resolution also allows for detection of segregation of proteins in membrane domains of a single organelle. Using SIM helped

to identify more details of the peroxisomes shape and size in *D. discoideum* and the GFP-*Dd*SQS pattern in the peroxisome. *D. discoideum* was co-transformed with two plasmids encoding GFP-*Dd*SQS and peroxisomal marker mRFP-PTS1. The transformed amoebae were imaged using SIM. Solid red fluorescence spheres were detected reflecting the peroxisomal matrix. Interestingly, all these spheres were surrounded by green ring-like structures of GFP-*Dd*SQS. Peroxisomes showed a variety in diameter (0.15-0.78µm) and approximately the average number per cell (140-210) per cell (Figure 4.22).

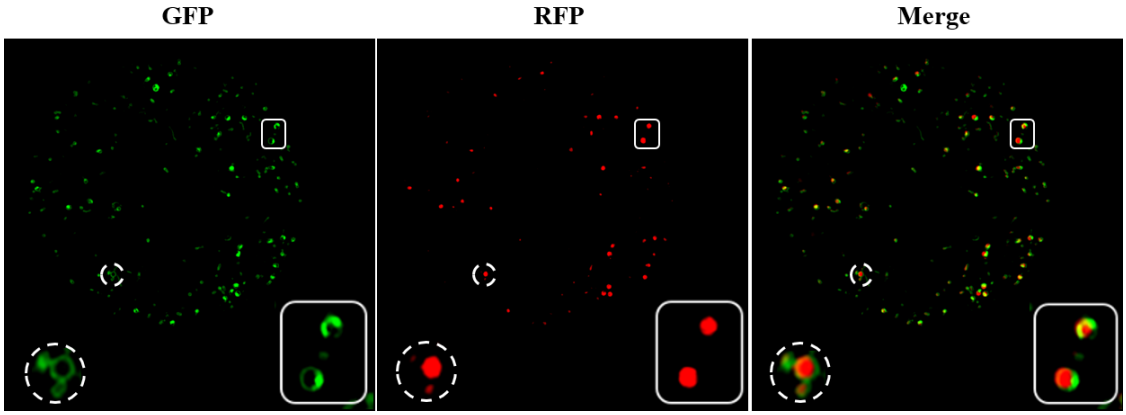


Figure 4.22: Structural Illumination Microscopy of *D. discoideum* expressing GFP-*Dd*SQS. Amoeba expressing GFP-*Dd*SQS and mRFP-PTS1 were imaged. The peroxisomes with low level green fluorescence showed an equal distribution of GFP-*Dd*SQS around the mRFP-PTS1 labelled matrix. Peroxisomes with higher levels of GFP signal, the signal concentrated into one or two green clumps. These clumps are associated with the membrane and giving a wedding ring-like structure.

4.3 Discussion

4.3.1 *Dd*SQS peroxisomal targeting signal

In this chapter of the present study, we confirmed that the first sterol biosynthesis enzyme SQS is a peroxisomal enzyme in the slime mould *D. discoideum*. Furthermore, we showed that it has a bipartite signal to target the protein to the inner leaflet of the peroxisomal membrane. The first one is a canonical peroxisomal targeting signal 1, which consists of the amino acid triplet serine, lysine and leucine (-SKL) located at the extreme C-terminal end of the protein. This amino acid sequence is a highly efficient signal (Gould et al., 1987, 1988, 1989, 1990; Lametschwandtner et al., 1998). The second one is a helix located just N-terminal of the PTS1. The helix appears to have a dual function, one during the targeting process in the cytosol and one in the insertion into the peroxisomal membrane. In targeting, it contributes to co-import of SQS versions lacking a PTS1, most likely by stimulating dimerization with SQS containing a PTS1 (Figure 4.23). Both partners of co-import require the presence of the helix. We found that the helix had an unusual number of small amino acids resembling the GxxxG dimerization motif. Our findings agree with (Russ and Engelman, 2000) who found that a GxxxG-like motif can mediate hydrophobic helix-helix associations. Remarkably, *h*SQS was found as a monomeric protein (Pandit et al., 2000), interestingly, it does not contain any GxxxG like motif at the TMD in the C terminal end.

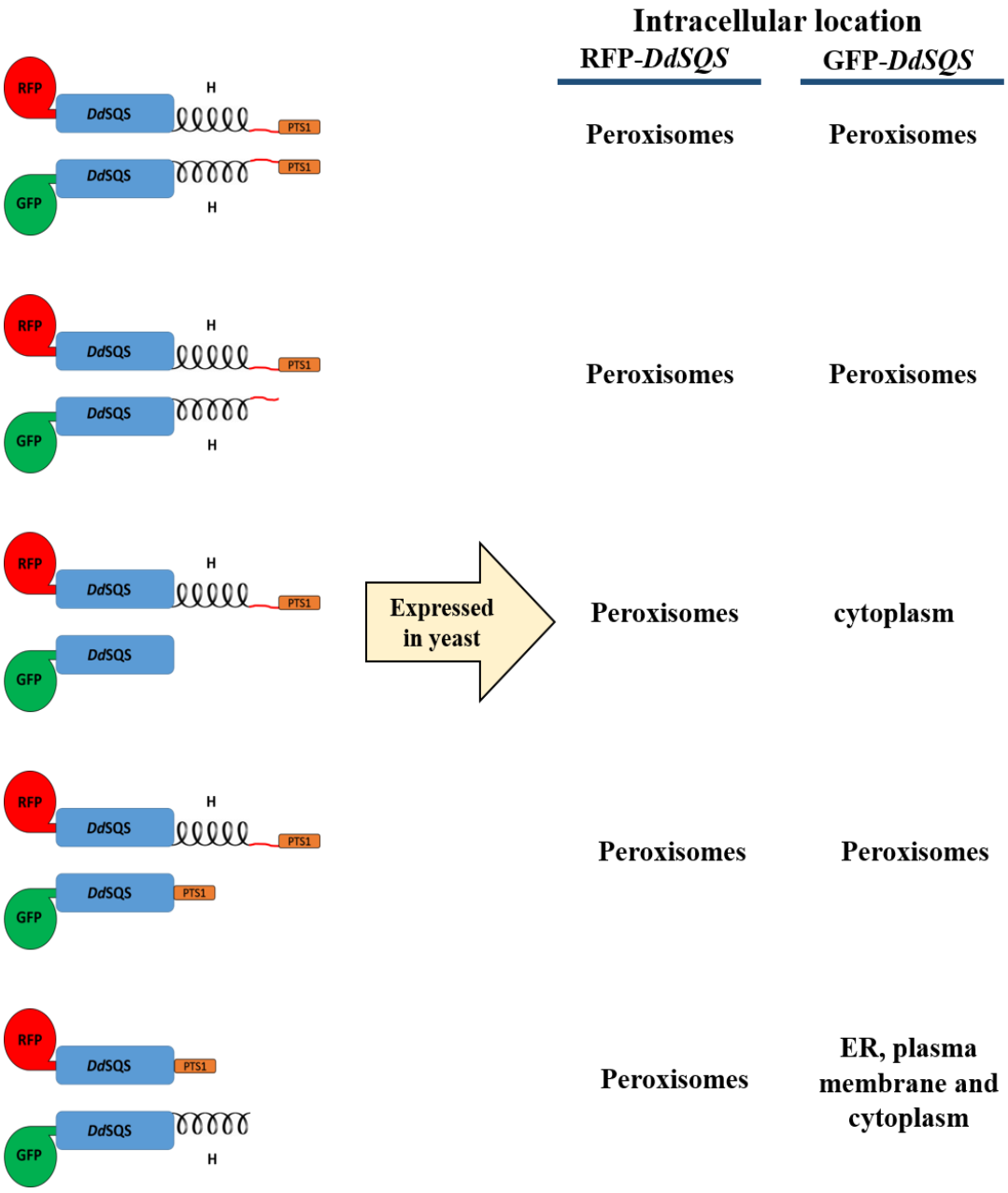


Figure 4.23: Schematic diagram showing the ability of full-length *Dd*SQS interaction with *Dd*SQS truncations. *S. cerevisiae* co-expressing the full-length RFP-*Dd*SQS in the presence of different *Dd*SQS truncations tagged with GFP. In addition to the subcellular location of each truncation.

4.3.2 Membrane association

Location of *Dd*SQS to the peroxisomal membrane was established by high salt extraction (1M NaCl), carbonate extraction and structured illumination microscopy (SIM). SIM imaging showed *Dd*SQS as a ring-like structure around the peroxisomal matrix. Results of protein extraction by 1M NaCl treatment of the membrane fraction of *D. discoideum* and *S. cerevisiae* revealed that the protein is not a peripheral membrane protein but instead in tightly bound to the membrane. Carbonate extraction further confirmed and extended the SIM and NaCl treatment findings and demonstrated that *Dd*SQS behaves like an integral membrane protein. Both in vitro protease protection experiments and the in vivo auxin-induced degradation experiments show that *Dd*SQS is protected from degradation by the peroxisomal membrane. This is the first example of an integral membrane protein that faces the peroxisomal matrix. Furthermore, as we showed in Figure 4.17Figure 4.18, the C-terminus of *Dd*SQS contains all information necessary and sufficient to target a protein to peroxisomes and to insert it into the peroxisomal membrane of yeast cells. The shortest construct consists of the C-terminal 24 amino acids which comprise 14 amino acids of the predicted hydrophobic helix, which includes the Gly-zipper-like motif (AxxxAxxxA), followed by ten more residues that include the PTS1.

*Dd*SQS has a unique localisation as a peroxisomal membrane protein, as it is anchored from the inside of the organelle. In contrast to other PMPs, *Dd*SQS needs to be imported first before it can be inserted into the membrane. Import, we find, is mediated by the PTS1 pathway. Since no other peroxisomal membrane proteins have been identified that follow such route, it is unlikely that yeast has got a specific machinery that mediates this insertion. Therefore we propose that insertion occurs spontaneously as has been described in *in vitro* experiments for artificial tail-anchors (Honsho et al., 1998b; Whitley et al., 1996; Yang et al., 1997), for review see (Borgese et al., 2003).

Why is *Dd*SQS localised inside peroxisomes? A previous study by Nuttal et al (2014) reported that the farnesyl diphosphate synthase is a peroxisomal enzyme in *D. discoideum* and this enzyme catalyse isopentenyl diphosphate to FDP. Therefore, the FDP should be located in the peroxisome. In sterol biosynthesis, SQS catalysed the initial reaction, which converts two molecules of FDP into squalene. That means converting hydrophilic molecules to a hydrophobic molecule. Thus, the enzyme needs

to be located very close to a membrane. That would facilitate the movement of the precursor and the product (Figure 4.24).

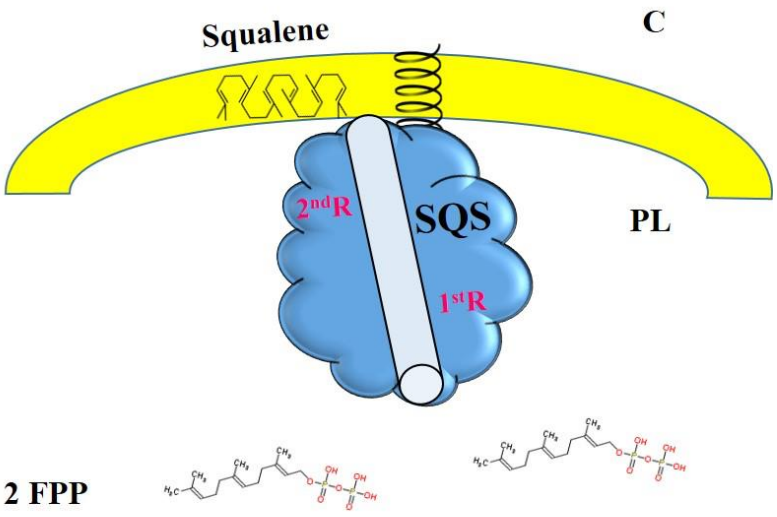


Figure 4.24: Model for the structure of *Dd*SQS in the peroxisomal membrane and the flow of substrate and product. 1stR and 2ndR are proposed SQS first and second reaction location in the central channel. PL: peroxisomal lumen, C: cell cytoplasm side.

All studied SQSs are tightly associated with the membrane as an integral membrane protein. They all have a tail-anchor. Furthermore, Pandit et al in 2000 solved a structure of human SQS (*h*SQS). Five helices create a channel that passes through the centre of the protein. Moreover, *h*SQS is anchored to the membrane in a way that allows one end of the channel to be closer to the membrane while the second end faces the aqueous environment. Likewise, we found that the *Dd*SQS anchored to the peroxisomal membrane by its C terminal TMD. Furthermore, it has a conserved amino acid sequence and showed 48% identity to the *h*SQS, including high similarity to functional and structurally important motifs (Figure 4.4) (Pandit et al., 2000). Thus, it might have a similar functional structure. Since FDPS is localised inside the peroxisomal matrix in *D. discoideum*, the substrate for SQS is produced inside peroxisomes. This is in line with the model of how SQS functions, where the hydrophilic substrate molecules FDPP enter the lower end of the channel. The hydrophobic (lipophilic) product SQ is released from

the end of the channel which is closest to the peroxisomal membrane allowing it to diffuse into the membrane (Jiang et al., 2017; Linscott et al., 2016; Liu et al., 2017; Rong et al., 2016).

The amino acid sequence alignment of *Dd*SQS with *h*SQS shows four highly conserved functional motifs, which they are associated with binding, regulatory and catalytic activities of SQS (Kalra et al., 2013; McKenzie et al., 1992; Merkulov et al., 2000). Sectors 1 and 3 were proposed to be the two FDP binding sites (Tansey and Shechter, 2000). The two active sites Aspartic acid (Asp) residues at positions 213 (Asp^{213}) and 217 (Asp^{217}) in sector 3, indicated by red arrows, play essential roles in the squalene synthase activity. Their putative function is to bind the substrate by forming magnesium salt bridges. A single mutation of all these residues leads to a strong reduction in the enzymatic activity of SQS. Sector 2 was proposed to play a role in the first reaction of the enzyme whereby the intermediate pre-squalene diphosphate (PSPP) is formed (Gu et al., 1998). The first Tyrosine residue in sector 2 at position 171 (Y^{171}), which marked in blue arrow, is crucial for catalysis and is supposed to be involved in the first reaction. Sector 4 was speculated to be involved in the second reaction step of SQS formation, which converts PSPP to SQ (Kalra et al., 2013; Tansey and Shechter, 2000). Amino acid residues of structural importance are boxed in with grey. As mentioned above, the catalytic important residues and those important for structure are conserved in *Dd*SQS (Pandit et al., 2000).

Remarkably, the transmembrane domain helix amino acid sequence, which is marked by the capital letter H, does not show a high similarity between the *Dd*SQS and the *h*SQS. That may be related to the differences in the sub-cellular location of each one of them. Furthermore, *h*SQS does not contain a GxxxG-like, therefore there is no chance to form helix-helix dimer (Figure 4.4).

All the previous finding points to the ability of *Dd*SQS reaching the peroxisome via the Pex5 pathway. Then, it will be anchored into the peroxisomal membrane from the inner side (Figure 4.25).

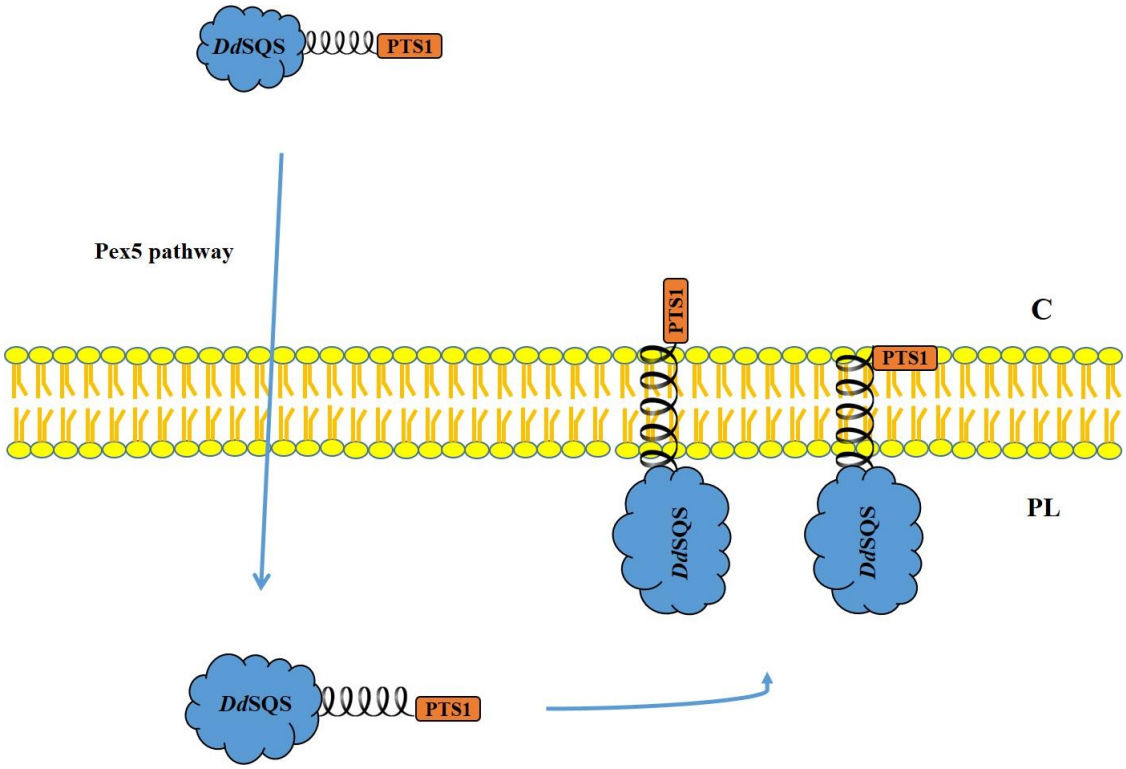


Figure 4.25: Diagram depicting our model of how *Dd*SQS is targeted to the inside of the peroxisomal membrane via the sequential steps of PTS1-mediated targeting and import and spontaneous insertion of the tail-anchor. Whether the PTS1 is exposed to the cytosol or partially embedded in the membrane is unknown and needs further investigation.

Taken all observations together, we conclude that SQS is a peroxisomal protein that is anchored to the inside of the peroxisomal membrane and we propose that its product is released in the peroxisomal membrane. This raises the question as to where subsequent steps in sterol biosynthesis occur? This is the topic described in chapter 5.

Chapter five "Topology of squalene epoxidase, cycloartenol synthase and sterol methyl transferase"

5.1 Introduction

The next three peroxisomal enzymes in the sterol biosynthesis pathway are squalene epoxidase (SQE), cycloartenol synthase and cycloartenol-C-24-methyltransferase (sterol methyl transferase SMT).

Having established that *Dd*SQS is peroxisomal, it seemed important to determine whether the other three peroxisomal enzymes involved in sterol biosynthesis are also membrane proteins. However, the structure of cycloartenol synthase indicates that it is a monotopic membrane protein (Alkuwayti, 2014). Most attention was therefore focused on the *Dd*SQE and the *Dd*SMT.

Squalene epoxidase (squalene monooxygenase, SQE) is the second enzyme involved in the sterol biosynthesis pathway. Initially, it was identified in rat liver microsomes (Yamamoto and Bloch, 1970). After that, it was demonstrated in fungi and plants as an enzyme involved in the ergosterol and phytosterol synthesis pathways, respectively (Godio et al.; Han et al., 2009; He et al., 2008; Rasbery et al., 2007; Uchida et al., 2007). Even though sterol biosynthesis does not occur in prokaryotic organisms, SQE was isolated from some bacteria. Instead of sterol, these bacteria make pentacyclic triterpenes (Nakano et al., 2007).

Squalene epoxidase is classified as a member of the flavoprotein monooxygenase enzyme family. These enzymes catalyse a huge variety of different kinds of oxidative reactions (Van Berkel et al., 2006; Joosten et al., 2007). SQE catalyses the conversion of squalene to a squalene epoxide (squalene 2,3-oxide). This reaction requires the flavin adenine dinucleotide (FAD) bind to the enzyme, nicotinamide adenine dinucleotide phosphate (NADPH), molecular oxygen, and NADPH-cytochrome P450 reductase (figure 5.1) (Shibata et al., 2001; Tai and Bloch, 1972). The final products are one molecule each of squalene epoxide, $NADP^+$ and a water molecule (Figure 5.1) (Entsch and van Berkel, 1995; Ghisla and Massey, 1989; Massey, 1994). The requirement for cytochrome P450 reductase arises because the enzyme-bound FAD has to be reduced by the NADPH but there appears not to be an NADPH binding domain on the SQE enzyme.

Hence, it is believed that the NADPH is used to reduce the FAD bound to the cytochrome P450 reductase and that the reduced flavin on the reductase then reduces the FAD bound to the SQE. (Tchen & Bloch 1957; Ono et al. 1977; Belter et al. 2011).

Figure 5.1: Squalene epoxidase catalyses the epoxidation reaction. The formula shows the oxidation reaction partners and the result of the SQE activity. The figure is adapted from Belter et al., (2011).

5.2 Results,

5.2.1 Squalene epoxidase

5.2.1.1 Cellular location

Squalene epoxidase was found to be located in the endoplasmic reticulum (ER) in all organisms, such as human, plants (e.g. *Arabidopsis*), and microalgae (e.g. *Chlamydomonas reinhardtii*) (Kajikawa et al., 2015). In some cases, it could be assigned to two different places, such as the ER and lipid bodies as shown in *Saccharomyces cerevisiae* (Leber et al., 1998; Müllner et al., 2004). However, a previous study in our laboratory showed that *Dd*SQE is a peroxisomal protein in the amoebae even though it lacks both a PTS1 and a PTS2. We have carried out further investigations to confirm the intracellular location of the *Dd*SQE. First of all, we confirmed that the enzyme is peroxisomal in *D. discoideum* (Figure 5.2) by tagging the *Dd*SQE with GFP and co-expressing the fused protein *Dd*SQE-GFP with a peroxisomal marker (mRFP-PTS1) in the amoebae. Moreover, we have identified the cellular location of the Dictyostelium enzyme in other organisms by expressing the enzyme tagged with a fluorescent marker (GFP) in baker yeast *S cerevisiae* and the *Drosophila melangasta* cell line S2R+.

The plasmids that could be used to transform yeast or the S2R+ cells were constructed as follows. The open reading frame encoding the *Dd*SQE was amplified by PCR using VIP2653 and VIP2654 the forward and reverse primers respectively, velocity DNA polymerase (a proofreading enzyme) and cDNA prepared from a *D. discoideum* cDNA library as a template. The PCR product was digested with SacI and HindIII, before being ligated into either the SacI and HindIII digested yeast plasmid pAS52 or the plasmid (pJL005) that was used to transfect the S2R+ cells. The pAS52 plasmid encoded GFP and an amino acid linker sequence Ala-Gly-Ala-Gly-Ala-Gly, which would be located between the *Dd*SQE and the GFP in the same open reading frame. The resulting plasmid (pMUR29) encoded *Dd*SQE-GFP. Similarly, the plasmid (pMUR143) for transforming the S2R+ cells encoded *Dd*SQE-GFP with a Gly-Ala-Gly-Ala-Gly-Ala linker between the *Dd*SQE and the GFP.

When expressed in wild type yeast, the green fluorescence from the *Dd*SQE-GFP was punctate (Figure 5.3A). Western blot analyses of the yeast-expressed *Dd*SQE-GFP showed that the GFP tagged enzyme had the expected molecular size (82kD (Figure 5.3B). The *Dd*SQE-GFP was then co-expressed with *Hc*RFP-PTS1 in yeast and it was apparent that the green and the red fluorescence co-located (figure 5.4), indicating that the *Dd*SQE-GFP was peroxisomal. Further confirmation that the *Dd*SQE-GFP was peroxisomal was sought by expression in the *Vps1*Δ*Fis1*Δ mutant yeast strain (Figure 5.3A), in which the peroxisomes have difficulty in division and are thus sausage-shaped (Motley et al., 2008). The green fluorescence was seen in a sausage-like structures, again indicating that the epoxidase is peroxisomal in yeast.

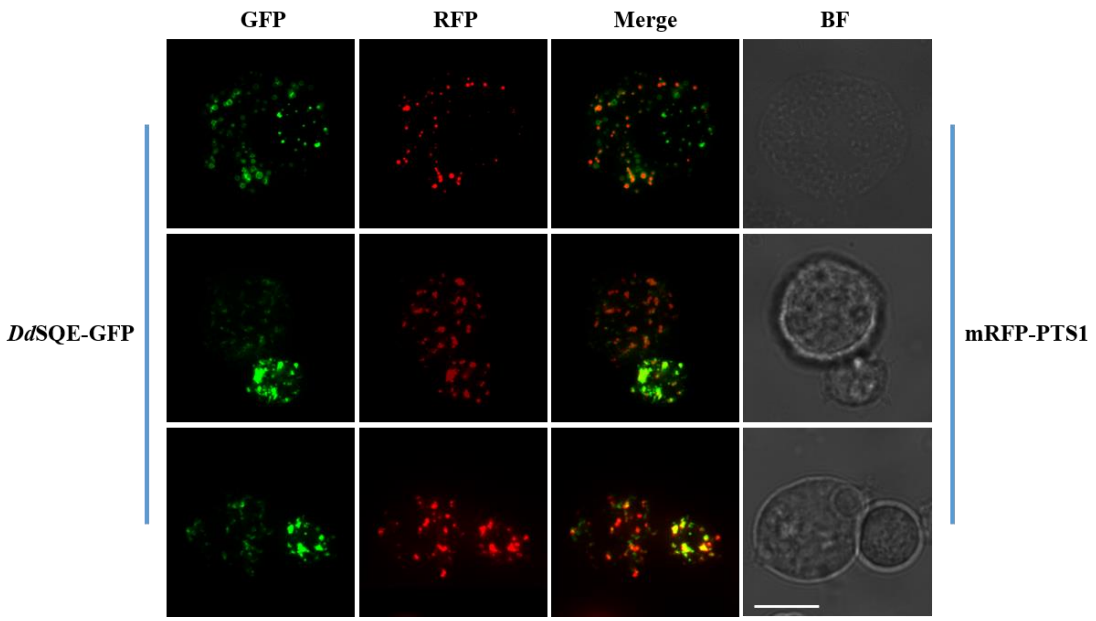


Figure 5.2: Fluorescence microscopy images of *D. discoideum*. Amoebae co-expressing *Dd*SQE-GFP and mRFP-PTS1 as a peroxisomal marker. An obvious co-localization was obtained between the green and the red fluorescence colours of the *Dd*SQE fusion protein and the peroxisomal marker respectively.

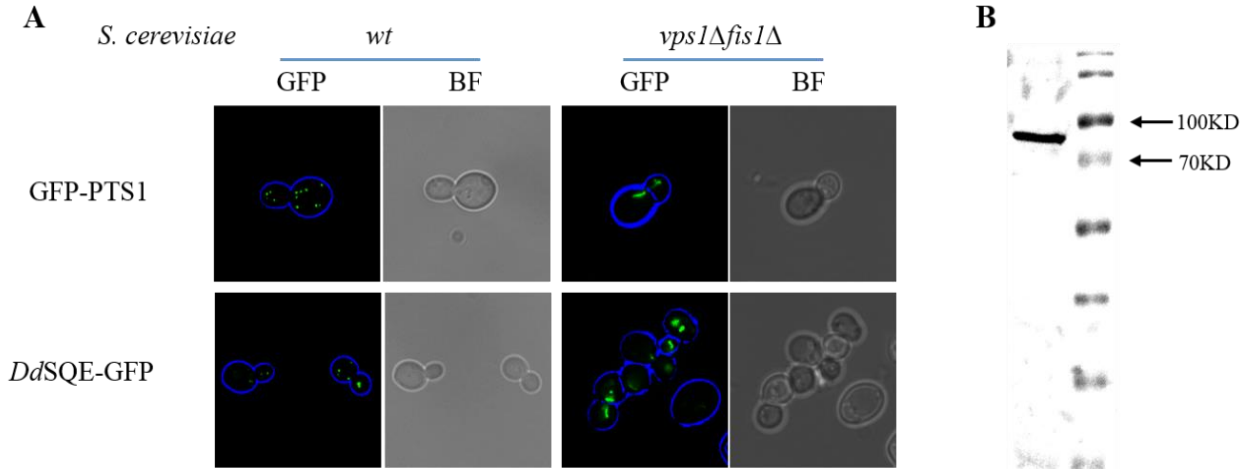


Figure 5.3: Analysis of the location and size of the *D. discoideum* squalene epoxidase expressed in *Saccharomyces cerevisiae*. A) Fluorescence microscopy images of wild type and *vps1*Δ*fis*Δ mutant strain *S. cerevisiae* expressing full-length *Dd*SQE-GFP. The GFP tagged enzyme appeared punctate in the wild type cells, and sausage-shaped in the mutant strain. B) Western blot analyses of yeast expressed *Dd*SQE-GFP. The TCA cell extract contained a product of the expected size for *Dd*SQE-GFP. The *Dd*SQE-GFP was detected by using an anti-GFP anti body.

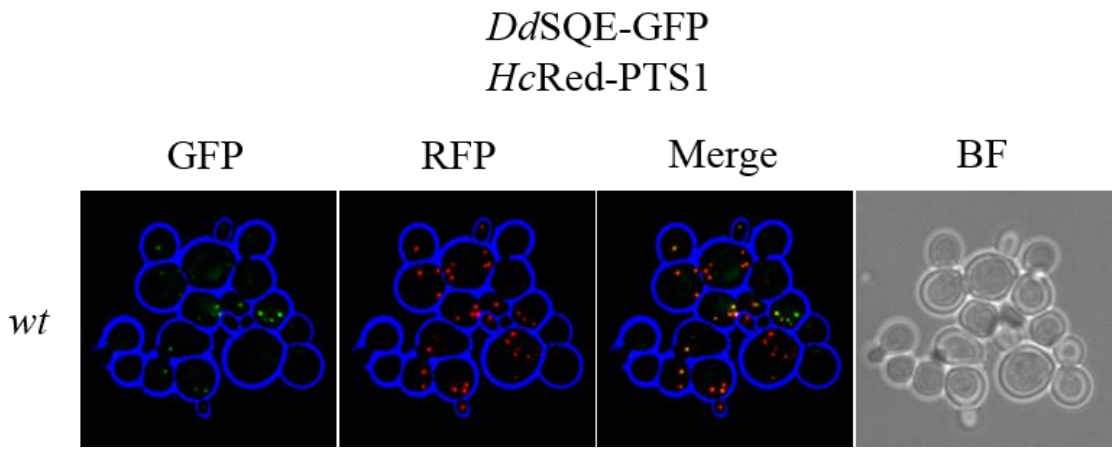


Figure 5.4: Demonstration of the peroxisomal location of *D. discoideum* squalene epoxidase when expressed in *S. cerevisiae*. Yeast was expressing both *Dd*SQE-GFP and a red peroxisomal marker *Hc*Red-PTS1. The green fluorescence of the *Dd*SQE-GFP co-localises with red fluorescence of the peroxisomal marker.

The *Dd*SQE-GFP transfected *Drosophila* S2R+ cells also expressed the enzyme well. They contained green fluorescent puncta when viewed under the fluorescence microscope similar to those expected of peroxisomes containing GFP (Figure 5.5). In addition, there was a little cytosolic fluorescence background.

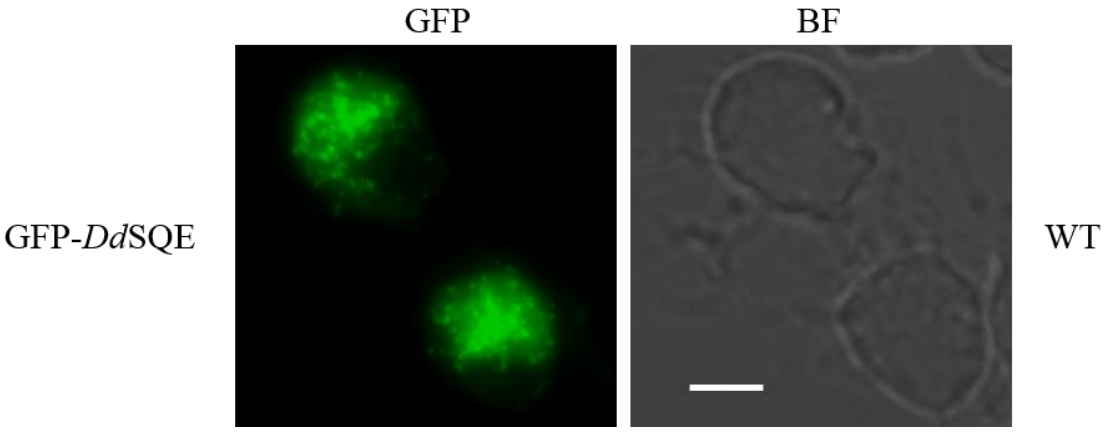


Figure 5.5: Analysis of the intracellular location of GFP-*Dd*SQE expressed in *Drosophila* S2R+ cells, visualized by fluorescence microscopy. A pattern of a green fluorescent puncta was observed in the S2R+ cells. Bar 10µm

Further investigation of the cellular location of *Dd*SQE in S2R+ cells included use of red fluorescent protein tagged with a peroxisomal targeting signal type 1 (mRFP-PTS1) as a peroxisomal marker for *Drosophila* cells' peroxisomes. The green puncta generated by GFP-*Dd*SQE co-localised with the red fluorescence of the peroxisomal marker (Figure 5.6).

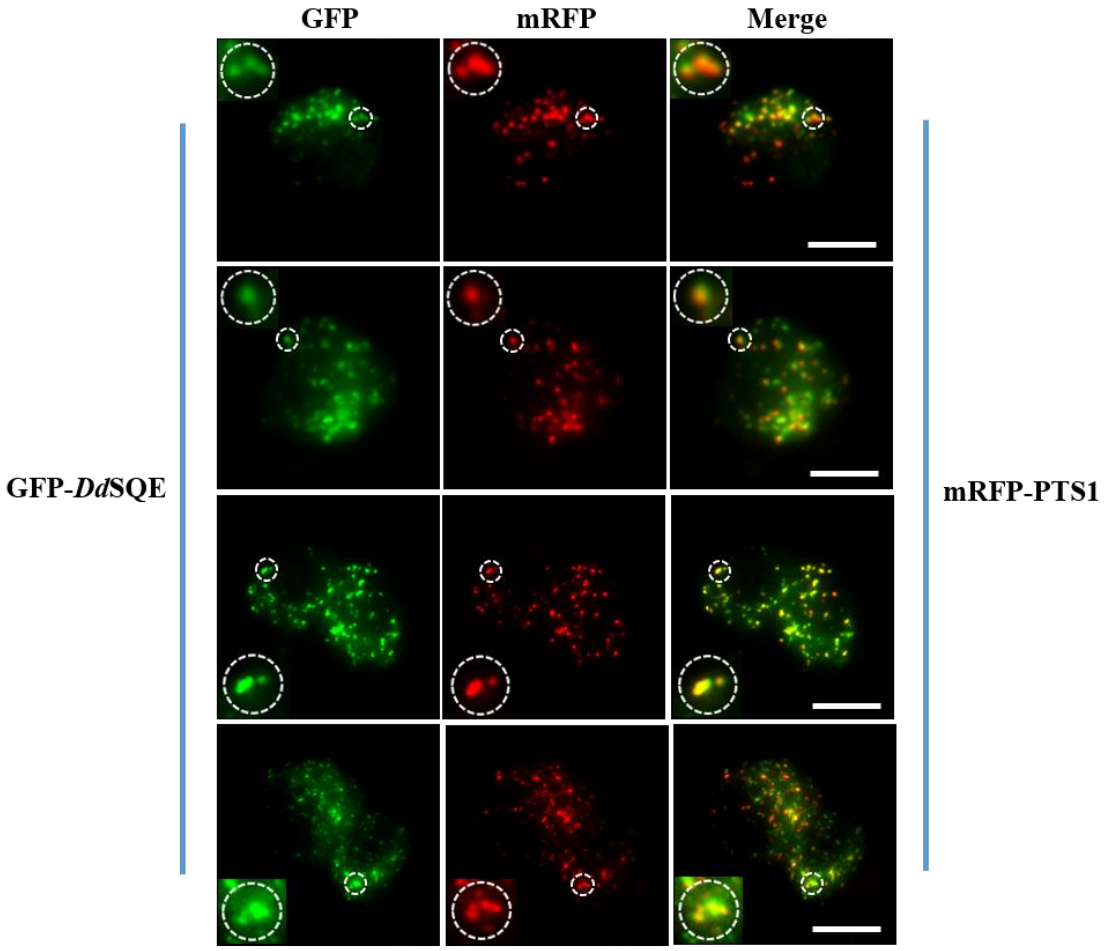


Figure 5.6: Fluorescence microscopy to determine the peroxisomal location of GFP-*Dd*SQE in *Drosophila* S2R+ cells expressing GFP-*Dd*SQE together with mRFP-PTS1. The fluorescence of the GFP-*Dd*SQE and the mRFP-PTS1 co-localised in puncta. Highlighted circles show magnified peroxisomes. Bar 10µm.

It was apparent that the *Dd*SQE is peroxisomal despite lacking an obvious PTS. *D. discoideum* squalene epoxidase was therefore expressed together with a peroxisomal membrane marker in two yeast mutant strains in an attempt to establish the targeting pathway. Strains *Pex5*Δ or *Pex7*Δ were used to block the PTS dependent pathways. These yeast strains were used in these experiments because we could not make these knockouts in *D. discoideum*. The enzyme was associated with the peroxisomes in both the *Pex5Δ* and *Pex7*Δ yeast strains as in the wild type strain, (Figure 5.4 and Figure 5.7). This confirmed that neither the Pex5 nor the Pex7 pathway is involved in transporting *Dd*SQE to the peroxisome.

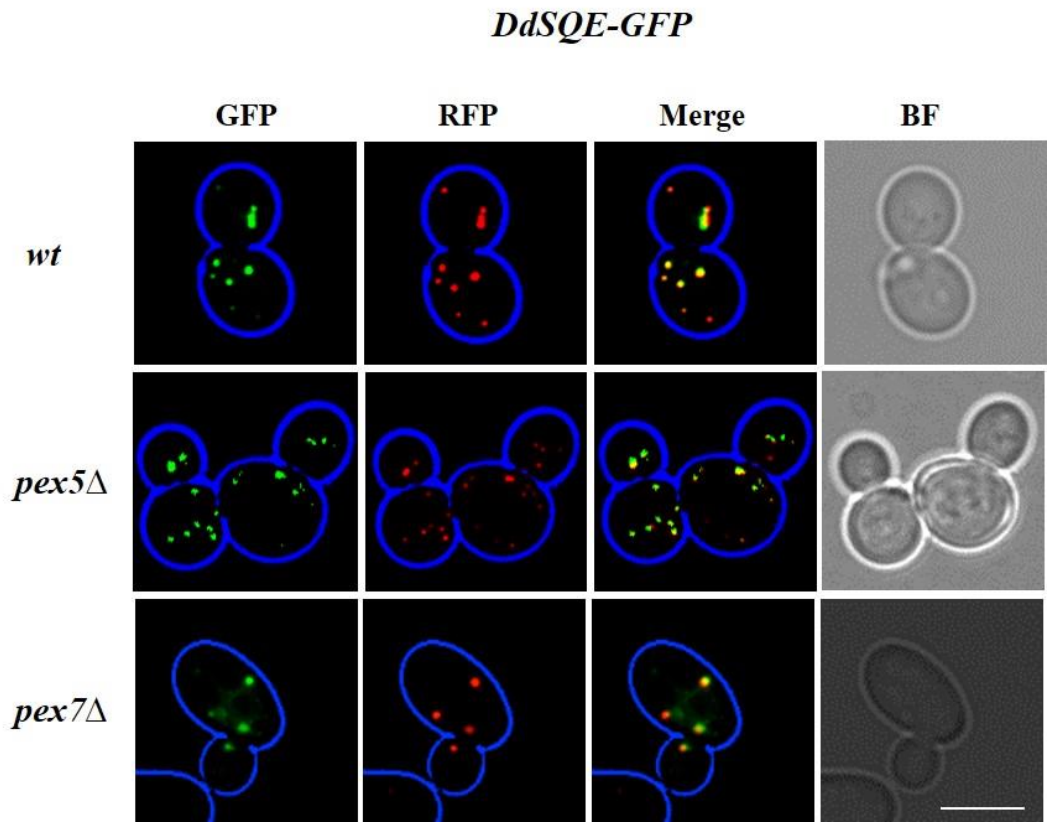


Figure 5.7: Determination of the intracellular location of *Dd*SQE expressed in different *S. cerevisiae* mutant backgrounds. *Dd*SQE-GFP was expressed in wild type and two mutant (Pex5, Pex7) yeast strains along with a peroxisomal marker (Mcherry-pex11). In the wild-type and the two mutant strains, there was the same pattern of green fluorescent puncta of *Dd*SQE-GFP co-localising with the red puncta of the peroxisomal membrane (Mcherry-Pex11) marker. Bar 5µm.

5.2.1.2 Squalene epoxidase function

Squalene epoxidase catalyses conversion of squalene into squalene epoxide by adding an oxygen. As it is a FAD-dependent enzyme, it needs an external electron donor (NADPH). Bioinformatics analysis of the amino acid sequence of the *D. discoideum* squalene epoxidase (Figure 5.8) predicted that *Dd*SQE lacks an NAD(P)H binding domain. However, it has a FAD binding site as in the epoxidases from other organisms (Shibata et al., 2001). This would suggest that the enzyme (SQE) needs to interact with another enzyme that can bind NADPH and donate the electrons needed for the epoxidase reaction. This donor enzyme was identified in human cells, mice and yeast as a cytochrome P450 reductase (Emmerstorfer et al., 2015; Gutierrez et al., 2003; Masaki et al., 1987; Porter, 2015). A previous study (Gonzalez-Kristeller et al., 2008) showed that *D. discoideum* is unusual in having two cytochrome P450 reductases; *redA* and *redB* and also, as in many other organisms an enzyme *redC* with some similarity to a cytochrome P450 reductase but with a different function as a component of a cytosolic iron-sulphur protein assembly machinery. Bioinformatics analysis showed that all three of the *D. discoideum* reductases are predicted to have the same domains which are the flavodoxin-like domain, the FAD binding domain and the NADPH binding domain. Furthermore, these domains are predicted to be present in the same order in all three enzymes. (Figure 5.9). *Dd*RedA and *Dd*RedB are also predicted to possess a TMD located in the N terminal end while *Dd*RedC is not predicted to have a TMD (Figure 5.10).

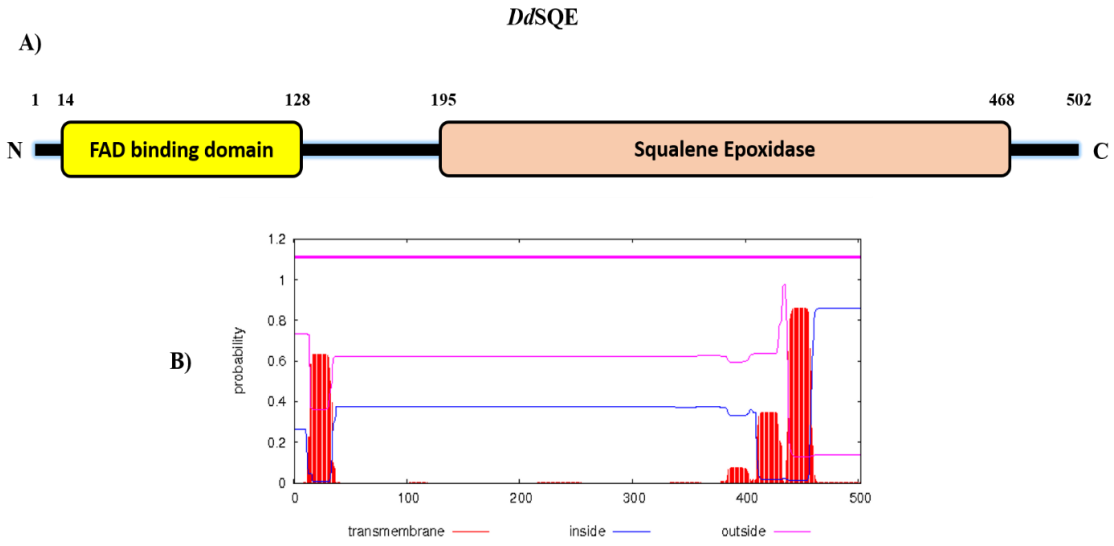


Figure 5.8: Bioinformatics analysis of the full length (502 amino acid) *D. discoideum* squalene epoxidase. A) The domains in *Dd*SQE were predicted by using the Pfam web site. Only two domains were identified; the FAD-binding domain and the squalene epoxidase active site domain. B) Transmembrane domain prediction for *Dd*SQE. The red columns display the probability of transmembrane domains (TMD) in the protein structure. The protein active domains were analysed by using the pfam web site (http://pfam.xfam.org/). The transmembrane domains were predicted by using the TMHMM Server v. 2.0 web site (http://www.cbs.dtu.dk/services/TMHMM/).

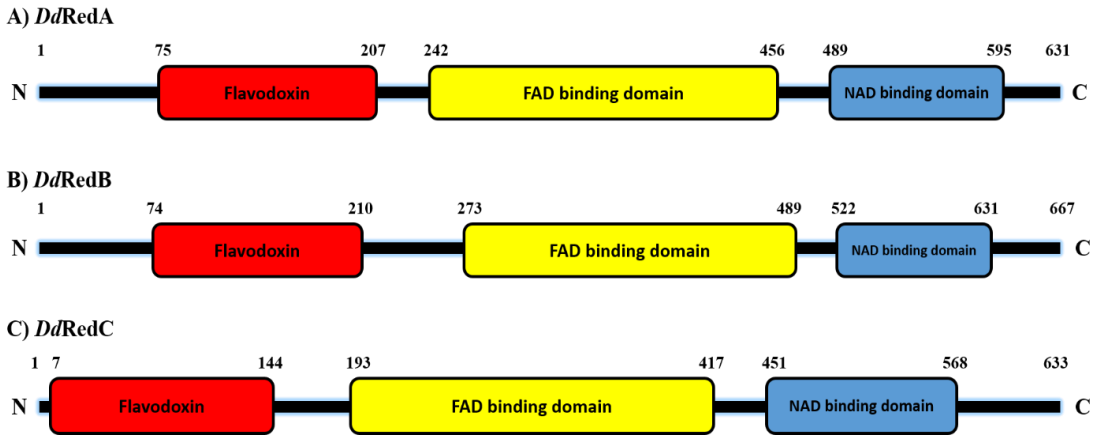


Figure 5.9: Domain prediction for *D. discoideum* RedA, RedB and RedC. All three enzymes were predicted to have the same domains in the same order.

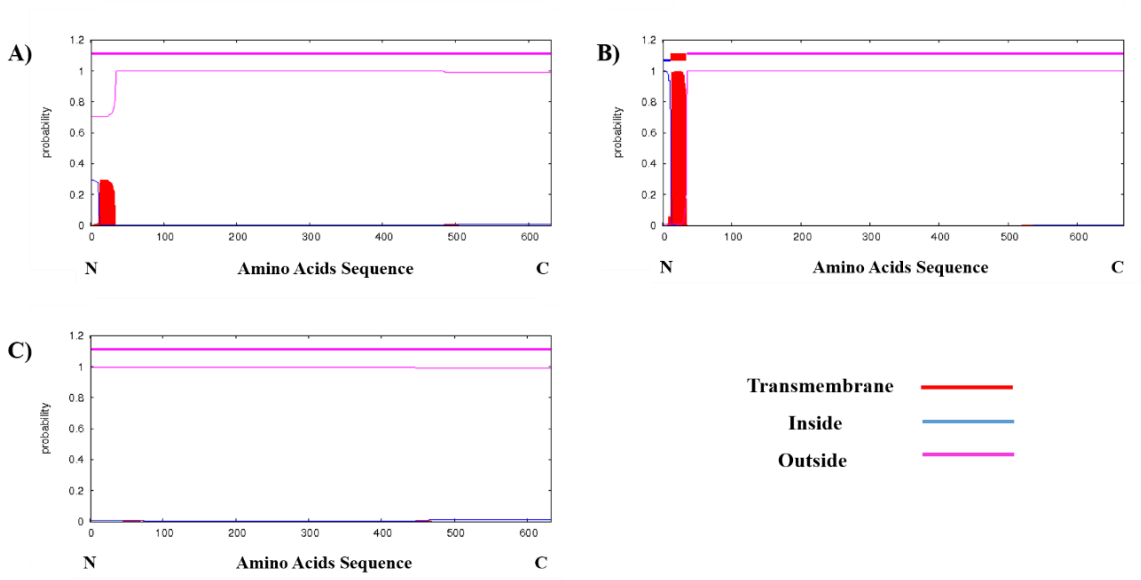


Figure 5.10: Transmembrane topology prediction for *D. discoideum* RedA, RedB and RedC. The red regions predict transmembrane domains in *Dd*RedA and *Dd*RedB. However, *Dd*RedC is not predicted to have a transmembrane domain.

To investigate the intracellular location of these enzymes, *RedA* and *RedB* were amplified from the *D. discoideum* cDNA library by PCR using VIP2976 +VIP2977 and VIP2960 +VIP2961 primers, respectively. Unfortunately, *RedC* could not be amplified. The *RedA* and *RedB* proteins were expressed tagged with GFP at the C terminal ends (*Red*A-GFP and *Red*B-GFP). In amoebae, the GFP fluorescence associated with the reductases appeared as rings and as a network (Figure 5.11). This appeared similar to the pattern of fluorescence given by GFP tagged ER proteins.

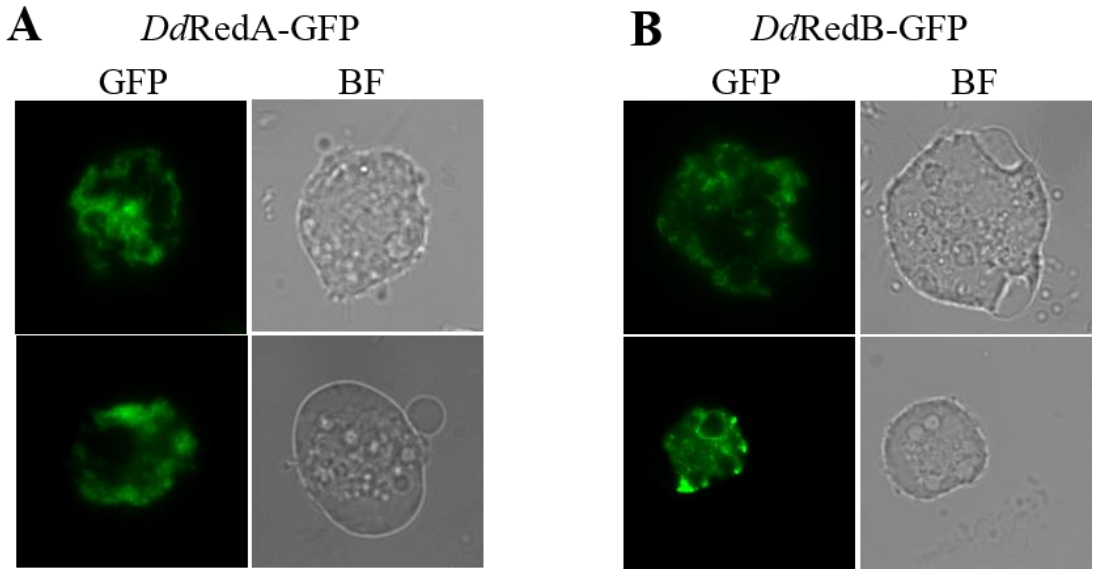


Figure 5.11: The intracellular localization of *D. discoideum* cytochrome P450 reductases. A) Amoebae over-expressing the *Dd*RedA-GFP showed a non-peroxisomal pattern of green fluorescence. Bright field (BF) microscopy displays cell morphology. B) Amoebae over-expressing the *Dd*RedB-GFP had a similar appearance to amoebae expressing *Dd*RedA-GFP.

In order to confirm the intracellular location of *Dd*RedA and *Dd*RedB, calnexin, which is known to be an ER integral membrane calcium binding phosphoprotein (Bergeron et al., 1994), was also expressed in amoebae. Calnexin has previously been used as a specific ER marker in *D. discoideum* amoebae (Hervet et al., 2011). It was tagged with mRFP at the C terminal end (calnexin-mRFP). *D. discoideum* cells expressing *Dd*RedA-GFP were transformed with the plasmid encoding the calnexin-mRFP as an ER marker. A clear co-localization between the *Dd*RedA-GFP and the calnexin-mRFP was observed. The same result was obtained from expressing *Dd*RedB-GFP and the ER marker (Figure 5.12A and B).

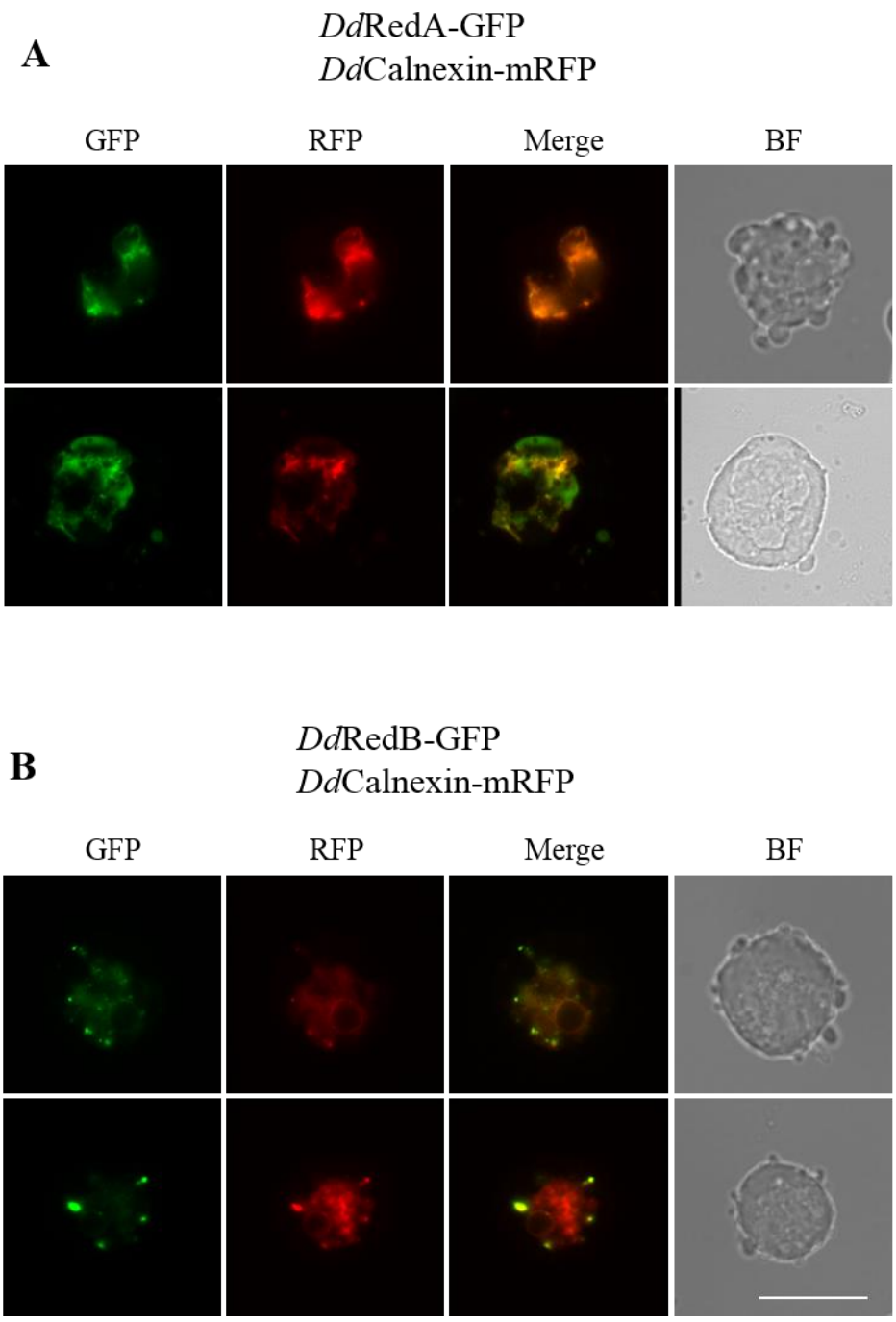


Figure 5.12: Investigation of the intracellular location of *Dd*RedA and *Dd*RedB in *D. discoideum*. A) Amoebae co-expressing *Dd*RedA-GFP and *Dd*calnexin-mRFP as an ER marker. An obvious co-localization of the GFP and mRFP fluorescence was obtained. B) The green and red fluorescence of *Dd*RedB-GFP and *Dd*calnexin-mRFP co-localized. Bar 10µl.

It had seemed possible that *D. discoideum* is unusual in having two cytochrome P450 reductases because one is in the peroxisomes to interact with *Dd*SQE and the other is in the ER to catalyse reduction of cytochrome P450. The intracellular location of the two reductases indicates that both enzymes are in the ER and that there is not a peroxisomal reductase. This indicates that, if *Dd*SQE is dependent on a cytochrome P450 reductase, there must be close contact between the peroxisomes and the ER to allow interaction between the *Dd*SQE and a cytochrome P450 reductase.

5.2.1.3 Investigations of *Dd*SQE as a peroxisomal membrane protein

SIM microscopy was used to investigate further whether *Dd*SQE is a peroxisomal membrane protein (Figure 5.13). Amoebae expressing RFP-PTS1 to define the peroxisomal matrix and *Dd*SQE-GFP were examined. Red fluorescence due to the expression of the RFP-PTS1 appeared as round blobs. The most striking result to emerge from these data was that these red blobs appeared to be surrounded by a green ring like structure, which must have contained the *Dd*SQE-GFP. The green rings, which varied in diameter from approximately 0.14µm to 0.8µm, nevertheless always closely surrounded the red blobs. Moreover, the fluorescence from the *Dd*SQE-GFP appeared to be evenly distributed throughout the rings and therefore most probably over the whole peroxisomal surface (Figure 5.13). The SIM images were thus consistent with a conclusion that the *Dd*SQE-GFP is a peroxisomal membrane protein. It also appears that the peroxisomes in *D. discoideum* vary markedly in size.

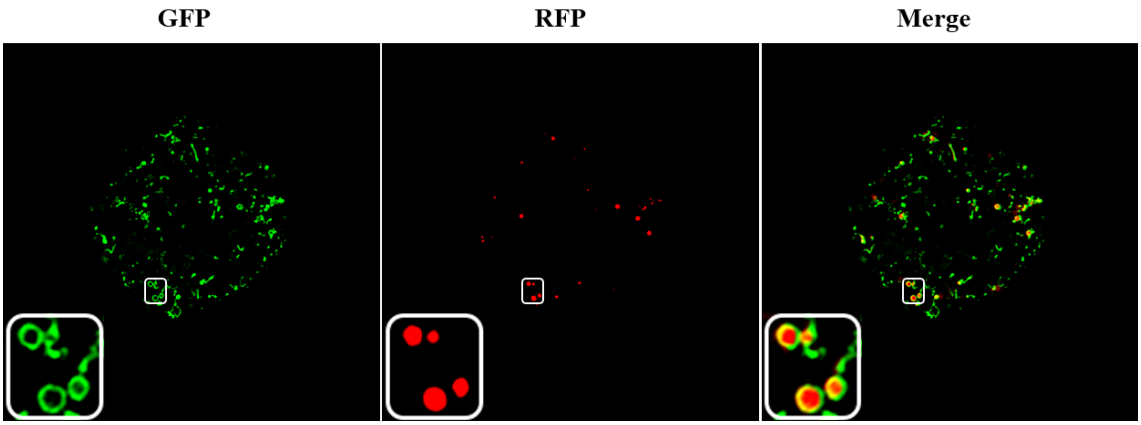


Figure 5.13: Structured Illumination Microscopy (SIM) of *D. discoideum* amoebae expressing mRFP-PTS1 as a peroxisomal matrix marker and *Dd*SQE-GFP. Red fluorescent blobs identified the peroxisomal matrix. Each blob was surrounded by a green ring like structure of the *Dd*SQE-GFP. The highlighted box shows a magnification of the selected point of each image.

5.2.1.4 *Dd*SQE orientation in the peroxisomal membrane

The orientation of *Dd*SQE in the peroxisomal membrane was investigated by using the protease protection technique. The investigations were carried out in the yeast *S. cerevisiae* because it is possible to isolate well-preserved peroxisomes from these cells whereas no method has been developed to isolate intact peroxisomes from *D. discoideum*. Spheroplasts were prepared from yeast cells that were expressing *Dd*SQE-GFP after growth on oleate medium. The spherolasts were gently broken open by using a Dounce homogeniser and the homogenate was digested with proteinase K attached to agarose beads.

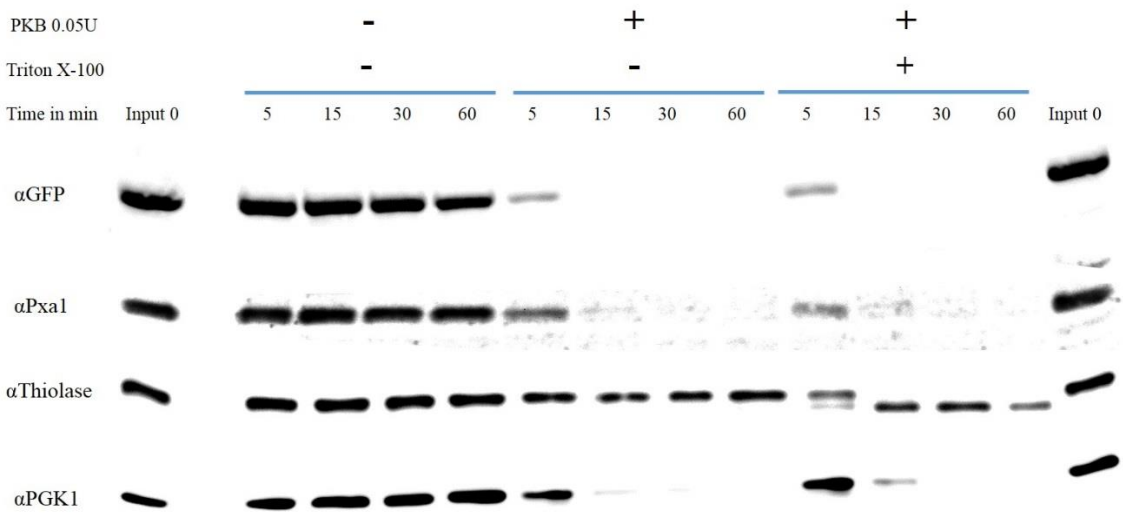


Figure 5.14: Testing the orientation of *D. discoideum* squalene epoxidase by using the protease protection technique. Wild type yeast cells expressing the *Dd*SQE-GFP were used to investigate the enzyme orientation in the peroxisomal membrane, after the yeast had been grown on oleate medium. Several controls were used: pxa1 protein as a peroxisomal membrane protein part of which is exposed to the cytosol; thiolase as a peroxisomal matrix protein (which is expressed in yeast only during growth on the oleate); pgk1 as a cytosolic protein. Anti-GFP antibody was used to detect the *Dd*SQE-GFP, while specific monoclonal antibodies were used to identify each control protein.

As shown in Figure 5.14, the *Dd*SQE-GFP was digested by the proteinase K both in the presence and the absence of the detergent, Triton X-100. However, a very faint band at the expected size of *Dd*SQE-GFP could be detected after 5min treatment with proteinase K in the presence and absence of the detergent. No intact *Dd*SQE-GFP could be detected after 15min or longer digestion. By contrast the anti-GFP antibody identified a clear undigested band of *Dd*SQE-GFP in the proteinase K untreated samples (control) between 0 and 60 mins incubation. The yeast pxa1 protein which is a peroxisomal membrane protein was also digested by the proteinase K in the presence and absence of the detergent. Part of pxa1 is exposed to the cell cytoplasm (Takahashi et al., 2007) and it is this part of pxa1 that is recognised by the anti-pxa1 antibody used in developing the western blot. Thiolase was used as a positive control. This is a peroxisomal matrix

protein in yeast (Glover et al., 1994) and was therefore protected from proteinase K digestion but only in the absence of the detergent (Figure 5.14). This showed that in the absence of Triton the peroxisomes were intact throughout the experiment. Phosphoglycerate kinase (PGK1), a cytosolic protein, was used as a control to show that the proteinase K was active and able to digest protein in the spheroplast homogenate both in the presence and absence of detergent. Taken together, the results of SIM imaging and the protease protection test showed that the *Dd*SQE-GFP is a peroxisomal membrane protein exposed to the cell cytosol. This would suggest that it might be possible for this enzyme to interact with one or other of the extra-peroxisomal cytochrome P450 reductases in *D. discoedium*. This depends on the reasonable assumption that *Dd*SQE is in the same location in *D. discoideum* peroxisomes as in yeast.

A live cell imaging method (SIM) was used again but, in this case, to determine whether *Dd*SQE and the cytochrome P450 reductases are in close proximity in Dictyostelium amoebae. The amoebae were transformed with two plasmids, each one having a different selectable marker. One of them conferred G418 resistance and was used to express either *Dd*RedA or *Dd*RedB labelled with a GFP as *Dd*RedA-GFP or *Dd*RedB-GFP respectively. The second plasmid had a blasticidine resistance gene and was used to express *Dd*SQE tagged with mRFP (*Dd*SQE-RFP). Red fluorescence in ring-like structures was seen in both cases and reflected expression of *Dd*SQE-RFP (Figure 5.15A and B). The green fluorescence was from the *Dd*RedA or *Dd*RedB. Only when green and red fluorescence are emitted in extremely close proximity does the fluorescence appear yellow in merged images. Thus, in Figure 5.15, the yellow fluorescence in the merged images indicates a very close interaction between the red fluorescent rings (i.e. *Dd*SQE) and the green fluorescent *Dd*RedA and *Dd*RedB.

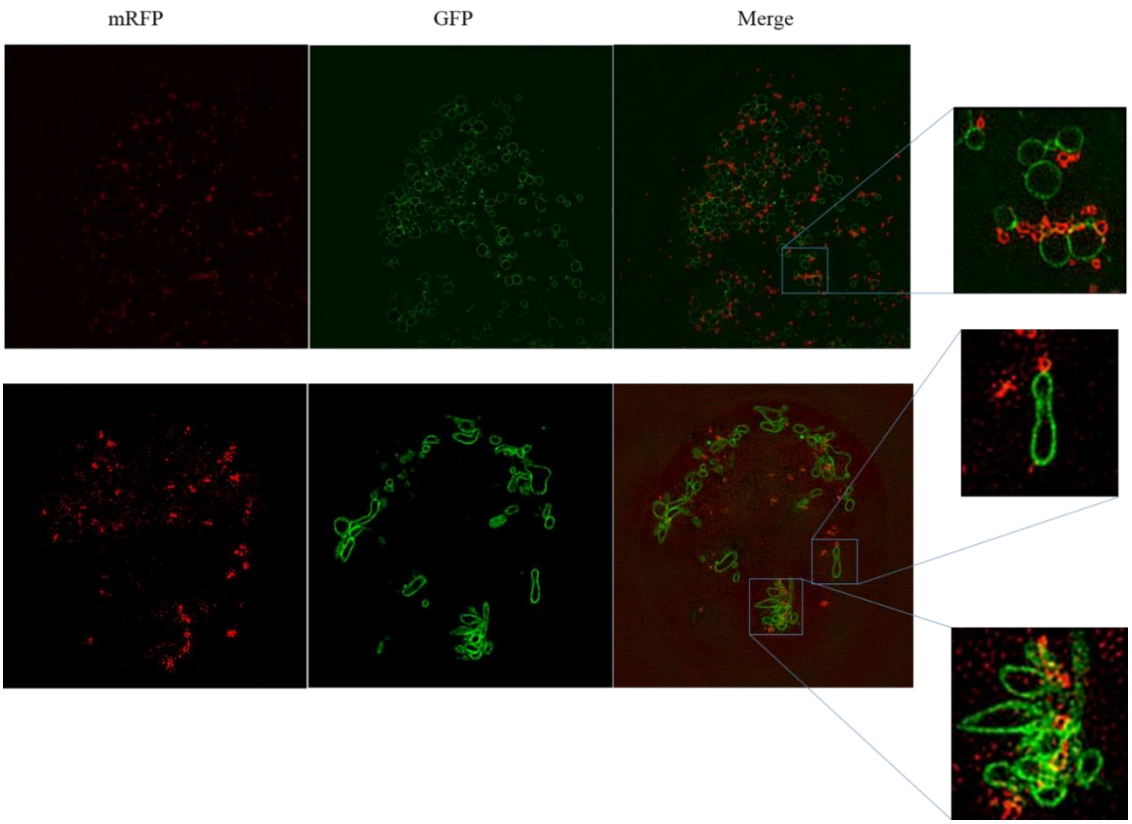


Figure 5.15: **Structured Illumination Microscopy of *D. discoideum*.** Amoebae were expressing either *Dd*RedA or *Dd*RedB tagged with GFP (*Dd*RedA-GFP/ *Dd*RedB-GFP) and also *Dd*SQE tagged with mRFP (*Dd*SQE-RFP). A) *Dd*RedA-GFP plus *Dd*SQE-RFP. B) *Dd*RedB-GFP plus *Dd*SQE-RFP. The SIM images indicate that *Dd*RedA or *Dd*RedB is extremely close to *Dd*SQE in live amoebae. This is shown by the yellow fluorescence in the merged images.

5.2.1.5 How does *Dd*SQE become associated with the peroxisomal membrane?

To try to determine how the *Dd*SQE associates with the peroxisome and particularly in the membrane, we attempted to identify whether it needs a special machinery found in *D. disocideum* only or whether it reaches the peroxisomes by following a universal pathway. Since *Dd*SQE-GFP expressed successfully in yeast and S2R+ cells and was associated with the peroxisomes, it was apparent that the enzyme follows a pathway common to several species to reach the peroxisomes (Figure 5.4 and Figure 5.6). It had been found that blocking the peroxisomal targeting signal type 1 and 2 pathways did not stop the *Dd*SQE-GFP reaching the peroxisome when it was expressed in *Pex5Δ* or *Pex7Δ*

yeast strain respectively. Therefore, it seems more likely that *Dd*SQE follows the pex19 pathway to be inserted into the peroxisomal membrane from the cytosolic side (as we showed it is exposed to the cell cytoplasm). Pex19 does exist in *D. discoideum* but, it has a somewhat unusual amino acid sequence (according to dictybase website) because it lacks a cysteine at the C-terminal end that is farnesylated in the pex19 in other organisms (Kammerer et al., 1997; Matsuzono et al., 1999; Rucktäschel et al., 2009). This is a very important modification for pex19 to be functional (Emmanouilidis et al., 2017).

Immunoprecipitation of *Dd*SQE-GFP (GFP trap) followed by mass spectrophotometry were used to determine whether pex19 can be found bound to *Dd*SQE or any other proteins. Neither pex19 nor any other potential transporter protein could be identified (Figure 5.16). The mechanism by which *Dd*SQE associates with the cytosolic face of the peroxisomal membrane has still to be established.

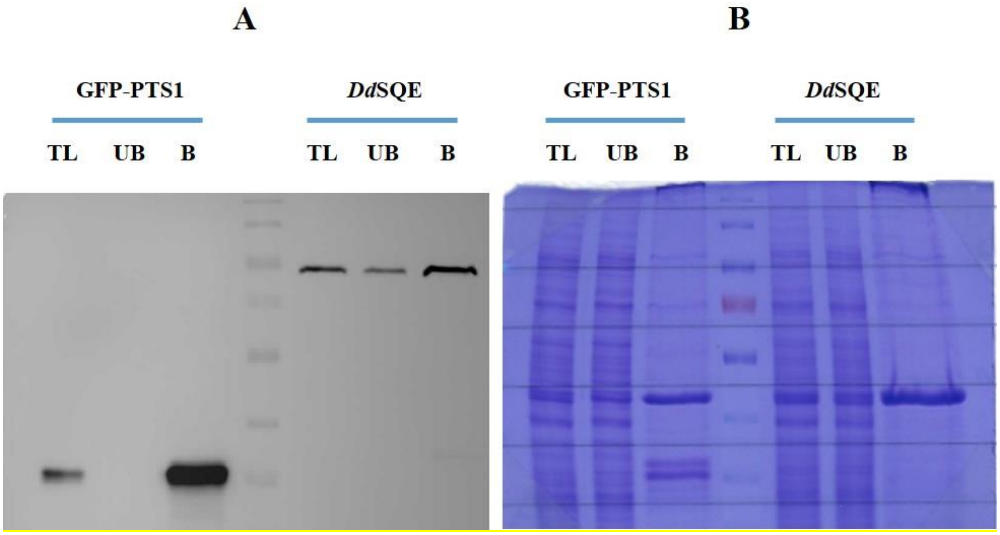


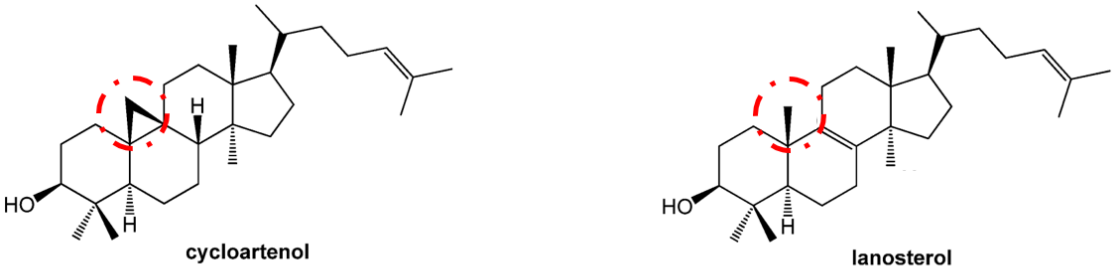
Figure 5.16: The GFP trap (co-IP) of *Dd*SQE-GFP. A) Western blot of amoebae expressing GFP-PTS1 (positive control) and *Dd*SQE-GFP. GFP was detected with anti-GFP antibody. B) SDS-PAGE gel of the same samples as in (A) stained with Coomassie blue stain. TL: total lysate, UB: unbound to the GFP trap beads, B: bound to the GFP trap beads.

5.2.1.6 Overview

In the previous chapter, it was demonstrated that *Dd*SQS is located on the matrix side of the peroxisomal membrane. Presumably this is to allow the highly hydrophobic product (squalene) of the synthase reaction to be temporarily stored in the peroxisomal membrane (Pandit et al., 2000). Squalene is then the substrate for *Dd*SQE which is again a peroxisomal membrane protein but, provided that it is located in *Dictyostelium* in the same way as when expressed in yeast, it is on the cytosolic side of the peroxisomal membrane. It is proposed that this orientation of the *Dd*SQE allows interaction with an NADPH-cytochrome P450 reductase so that reducing equivalents from NADPH may be used in the reaction catalysed by the *Dd*SQE. It is not known whether the interaction is with RedA or RedB (or both) but, provided that there is such an interaction, it must be able to involve RedB because a mutant strain of *D. discoideum* in which RedA has been knocked out (Gonzalez-Kristeller et al., 2008) is fully viable.

5.2.2 *D. discoideum* Oxidosqualene Cyclase

*Dd*oxidosqualene cyclase (*Dd*OSC) is the third enzyme on the sterol biosynthesis pathway. It converts oxidosqualene (squalene epoxide) into cycloartenol in plants. However, in animals and fungi the enzyme produces lanosterol, which is a slightly different sterol (Figure 5.17).



Plant Fungi and animal

Figure 5.17: Chemical structure of cycloartenol and lanosterol. The different in the structure highlighted by the red circle. Figure adapted from (Brycki et al., 2014).

5.2.2.1 Intracellular location of *Dd*oxidosqualene cyclase

*Dd*OSC has the following amino acid sequence Ser, Lys and Ile (SKI) at its C terminal end. This is believed to be a peroxisomal targeting signal type 1 (PTS1) triplet in some organisms especially plants (Reumann et al., 2016). We expressed the *Dd*OSC tagged with GFP at the N-terminal end in its normal host *D. discoideum* and in a foreign host *S. cerevisiae*. RFP-SKL was used as the peroxisomal marker. In *D. discoideum* there were very clear green fluorescent puncta (of the GFP-*Dd*OSC) in addition to the red fluorescent puncta given by the peroxisomal marker. An obvious co-localization between the red and green fluorescence was obtained (

Figure 5.18) and indicates that *Dd*OSC is peroxisomal. Overexpressing the GFP-*Dd*OSC in yeast showed that some of this enzyme was in the peroxisomes, because there were green fluorescent puncta co-localising with the red punctate fluorescence of the peroxisomal marker. Nevertheless, the majority of the GFP-*Dd*OSC was not peroxisomal. It was either in the cytosol or in the ER or it formed large ring-like structure around the cell presumably in the plasma membrane. In addition, it accumulated in puncta which did not co-localise with the peroxisomal marker. These may have been aggregates (

Figure 5.19). These results would suggest that SKI function well as a PTS1 in *D. discoideum* but not in *S. cerevisiae*. Bioinformatics investigation of the whole genomic data base of the *S. cerevisiae* showed that only three proteins end in SKI triplets (Table 5.1). None appears to be a peroxisomal protein. This would suggest that SKI is not used as a functional peroxisome targeting signal in the yeast *S. cerevisiae*. Use of SKI as a functional PTS1 in *D. discoideum* would further emphasise that Dictyostelium has plant-like characteristics.

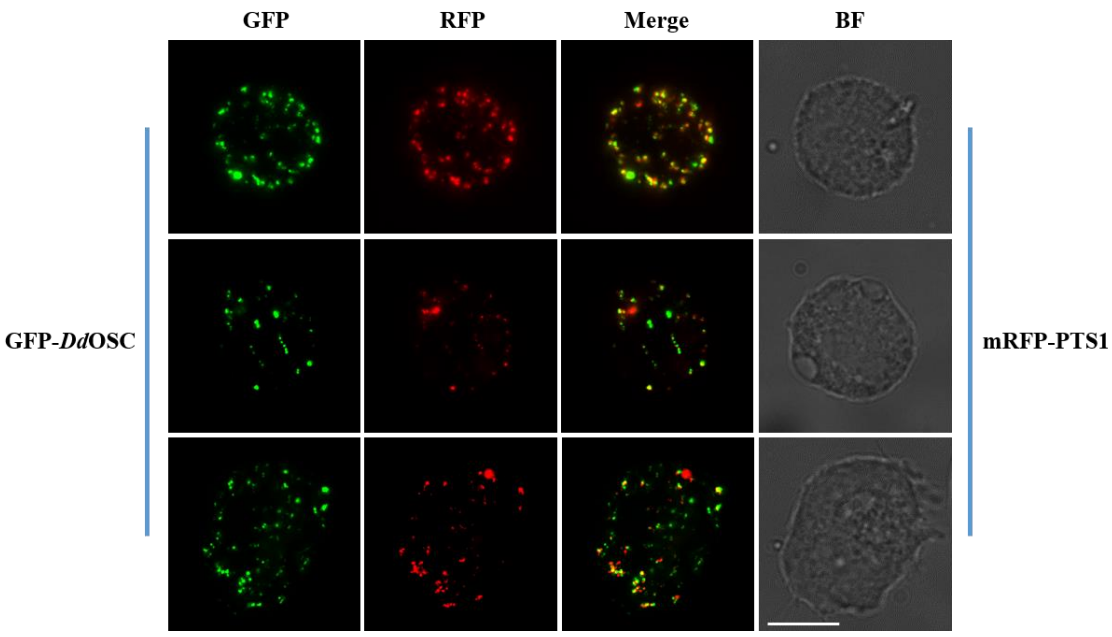


Figure 5.18: Determination of the intracellular location of *Dd*OSC in *D. discoideum*. Amoebae were co-expressing GFP fused *Dd*OSC (GFP-*Dd*OSC) and mRFP-PTS1 as a peroxisomal marker. The green fluorescence, reflecting expression of GFP-*Dd*OSC, appeared as puncta and showed a clear co-localization with the red fluorescent puncta of the peroxisomal marker (mRFP-PTS1). Bar 5µm.

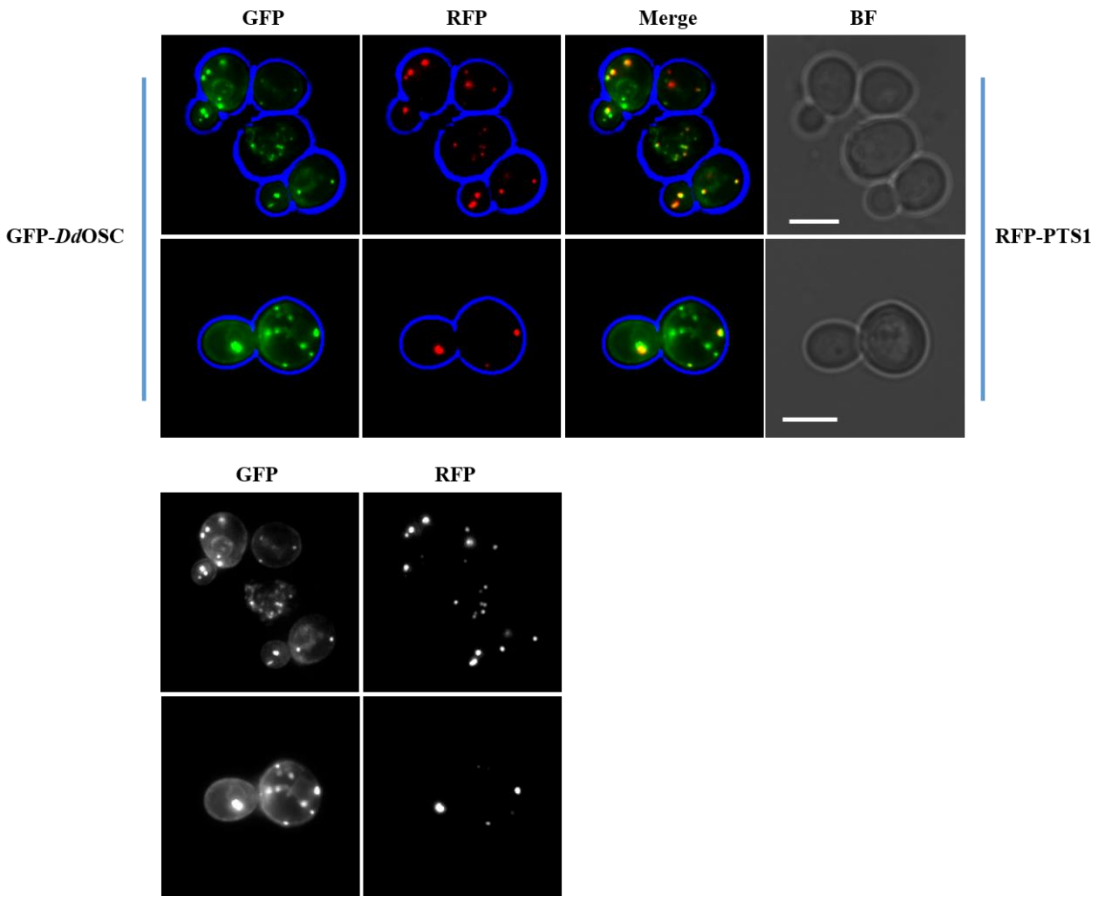


Figure 5.19: Fluorescence microscopy analysis to determine the intracellular location of *Dd*OSC in the foreign host *S. cerevisiae*. The yeast cells were transformed to express GFP-*Dd*OSC and a peroxisomal marker RFP-PTS1. The marker appeared as red puncta randomly distributed in the cell. The GFP-*Dd*OSC was presents in several intracellular locations. The lower part of the figure shows the fluorescence signals without their being converted to coloured signals. Bar 5µm.

Table 5.1: Search among all yeast proteins for those ending with a SKI amino acid sequence (C-terminal end).

Feature Name	Gene Name	Match Pattern	Match Start	Match Stop	Locus Info
YDR242W	AMD2	SKI	547	549	Putative amidase
YGL111W	NSA1	SKI	461	463	Constituent of 66S pre-ribosomal particles; involved in 60S ribosomal subunit biogenesis
YPL089C	RLM1	SKI	674	676	MADS-box transcription factor; component of the protein kinase C-mediated MAP kinase pathway involved in the maintenance of cell integrity; phosphorylated and activated by the MAP-kinase Slt2p; RLM1 has a paralog, SMP1, that arose from the whole genome duplication

5.2.2.2 Observation of *Dd*oxidosqualene cyclase by use of SIM

Structured Illumination microscopy (SIM) was used to identify the intracellular location and the morphology pattern of *Dd*OSC. *D. discoideum* was transformed to express GFP-*Dd*OSC along with mRFP-PTS1. The transformed amoebae were used for SIM analysis. The mRFP-PTS1 appeared as roughly circular blobs (peroxisomal matrix) whereas the fluorescence from the GFP- *Dd*OSC appeared as a single spot at the edge of the peroxisomal matrix. It was already known that *Dd*OSC is a somewhat unusual peroxisomal membrane protein because it would appear that in common with the similar human protein, lanosterol synthase (Ruf et al., 2004; Thoma et al., 2004), it is a monotopic membrane protein (Alkuwayti, 2014). The SIM microscopy suggests that the *Dd*OSC is also unusual because it appears that it occupies a single location in the membrane of each peroxisome (Figure 5.20).

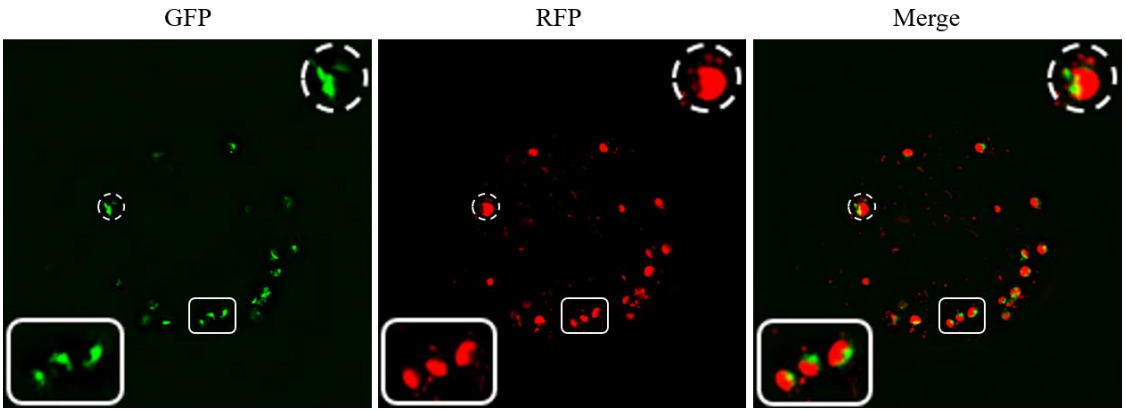


Figure 5.20: Use of Structured Illumination Microscopy to identify the intracellular location and morphology pattern of GFP-*Dd*OSC.

5.2.3 Cycloartenol-C-24-methyltransferase

Cycloartenol-C-24-methyltransferases is the fourth enzyme involved in sterol biosynthesis from FDP. Among different species, there are some differences in this step of the pathway. In plants there are two families of transferases (Bouvier-Navé et al., 1998). One family is believed to add the first methyl group to the sterol side chain while the second family adds, at the end of the sterol biosynthesis pathway a further methyl group to give an ethyl group. In fungi (e.g. yeast), only a single methyl group is added to the side chain and this occurs as the final step of sterol (ergosterol) biosynthesis. In animal, there are no sterol methyl transferases because cholesterol lack a side chain methyl group.

In *D. discoideum*, the major sterol has an ethyl group at C-24. Therefore, two transferases families were expected to be involved as in plants. Analysis of the *D. discoideum* genome showed that there are only two probable cycloartenol-C-24-methyltransferases (SMT). One of them possesses a putative PTS1 at the C- terminal end, comprising the amino acids sequence Ala, Lys and Leu (AKL) and is likely to be peroxisomal. The cycloartenol substrate for the methyltransferase reaction is located in the peroxisome as well. Thus, it seems that this enzyme probably adds the first methyl group to the cycloartenol. The other non-peroxisomal enzyme presumably adds the second methyl group to form an ethyl group. We confirmed that the *D. discoideum* SMT1 (*Dd*SMT1) is a peroxisomal enzyme in *D. discoideum* and also in the foreign host

S. cerevisiae. Thus, fluorescence microscopy of *D. discoideum* expressing GFP-*Dd*SMT1 together with mRFP-PTS1 (peroxisomal marker) showed a clear co-localization between the green fluorescence of GFP-*Dd*SMT and the red fluorescence of the marker. Similar results were obtained from the GFP-*Dd*SMT1 transformed *S. cerevisiae* (Figure 5.21A and B). Furthermore, the enzyme behaved as an integral membrane protein, because most of it remained in the final pellet when amoebal membranes were treated with 0.1M Na_2CO_3 (Figure 5.22). However, some of the protein was present in the second supernatant.

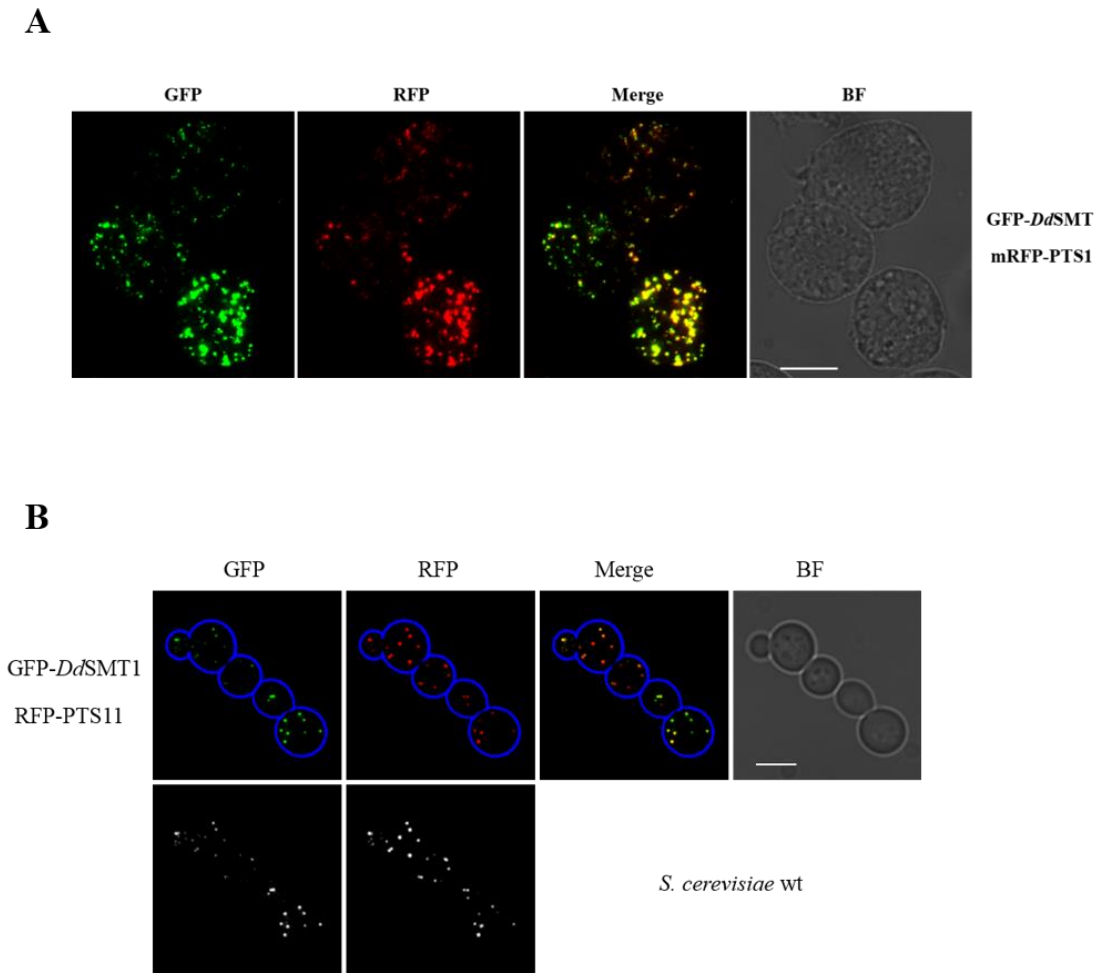


Figure 5.21: Fluorescence microscopy images of *D. discoideum* sterol methyl transferase. A) Amoebae expressing the GFP tagged *Dd*SMT1. The enzyme appeared as green puncta. These puncta co-localised with those of the peroxisomal marker (mRFP-PTS1). Bar 10µm. B) The co-localization of GFP-*Dd*SMT1 and the peroxisomal marker in *S. cerevisiae* indicated that the enzyme is also peroxisomal in yeast. Bar 5µm.

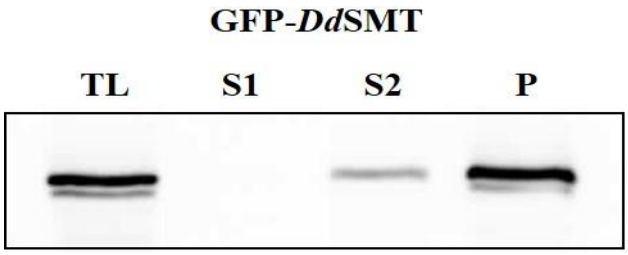


Figure 5.22: Western blot analyses of GFP-*Dd*SMT. Amoebae expressing GFP-*Dd*SMT were treated with 0.1M Na_2CO_3. The majority of the protein was detected in the final pellet by using an anti-GFP antibody.

5.2.3.1 Analysis of the ability of -AKL to act as a PTS1

The *Dd*SMT1 peroxisomal targeting signal (AKL) was tested in *S. cerevisiae* by expressing GFP fused *Dd*SMT1 (GFP-*Dd*SMT1) in a pex5Δ yeast. The GFP tagged enzyme was cytosolic, while the peroxisomal membrane marker (Mcherry-pex11) appeared as red fluorescent puncta (Figure 5.23). This would suggest that the AKL triplet acts as a PTS1 in DdSMT and that there is no other peroxisomal signal in the *Dd*SMT even though DdSMT is a peroxisomal membrane protein and not a matrix protein.

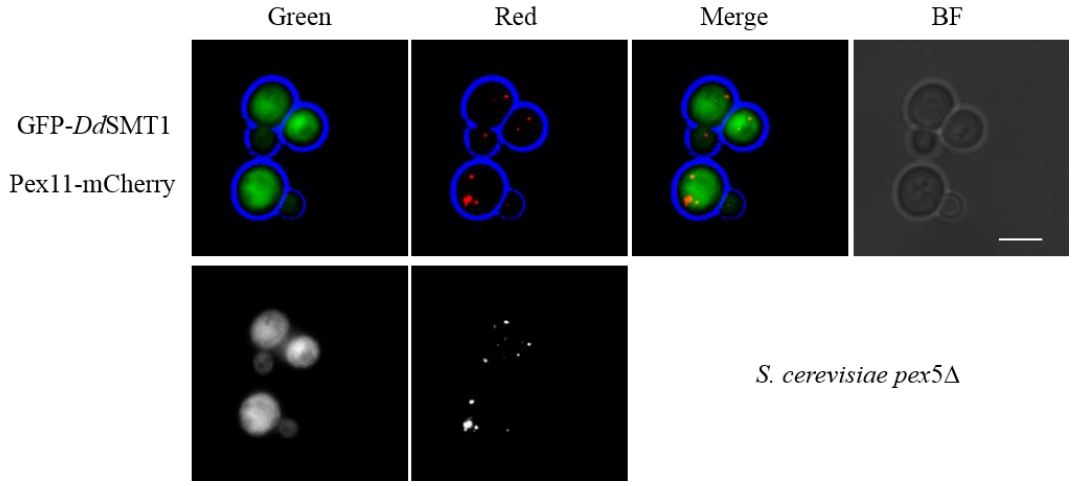


Figure 5.23: Testing the *Dd*SMT peroxisomal targeting signal. Fluorescence microscopy showed that the GFP-*Dd*SMT remained in the cytosol when it was expressed in a *Pex5*Δ mutant strain of *S. cerevisiae*. The peroxisomal membrane marker (MCherry-Pex11) was located in the peroxisome because it was not using the Pex5 pathway to reach the peroxisomes. Bar 5µm.

5.2.3.2 Domain structure of *Dd*SMT

Although there is considerable conservation of amino acid sequence in the SMT of *Dictyostelium discoideum* and other organisms, *Dd*SMT has an additional motif located at the N- terminal end, which is not found in the SMT sequences of other organisms. This motif contains 187 amino acids starting from the first methionine and then the methyltransferase domain starts with another methionine at position aa188. Thus, the active domains analysis using the Pfam online analysis tool showed two different domains in *Dd*SMT1 (

Figure 5.24). Furthermore, in the predicted three-dimensional (3D) structure, *Dd*SMT1 appeared to comprise two different fused polypeptides. The 3D structure was obtained by using the Pher2 3D analysis programme (Figure 5.25).

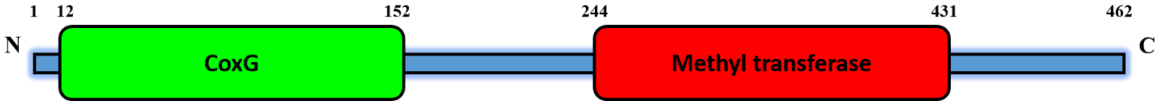


Figure 5.24: Predicted domain analysis of peroxisomal *Dd*SMT. The methyl transferase (MT) domain was predicted to be a second domain in *Dd*SMT between amino acid 244 to 431. A carbon monoxide dehydrogenase sub unit G-like domain was predicted to be located upstream from the MT domain between amino acids 12 and 152.

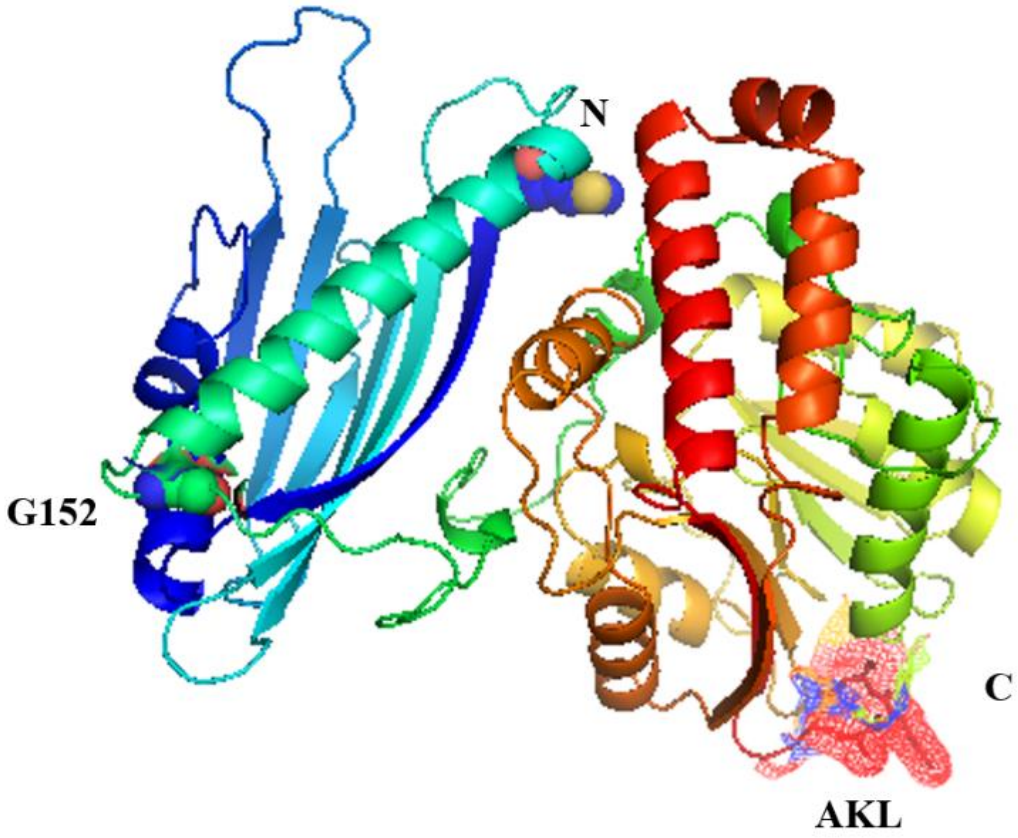


Figure 5.25: Prediction of the three-dimensional structure of *Dd*SMT. The Phyer2 program was used to predict the three-dimensional structure of *Dd*SMT. The data obtained were analysed by using the PyMol software program. The two polypeptide domains were predicted to remain essentially separate. The first polypeptide runs from M^{1} to G^{152}. These are shown as spheres. The remaining amino acids were predicted to form another polypeptide, which is the methyl transferase. The peroxisomal targeting signal (AKL) is highlighted as a mesh.

The first domain has *structural* similarity with a large number of proteins having different functions. One of these is the G sub-unit of carbon monoxide dehydrogenase (CoxG). CoxG is a bacterial protein but it has the closest similarity in amino acid sequence with the first domain of peroxisomal DdSMT. The Cox enzymes utilize carbon monoxide (CO) as a carbon source in an aerobic reaction giving CO_2. The sub-units are encoded as a cluster on a mega-plasmid in *Oligotropha carboxidovorans* and it has been shown that most of the sub-units are essential for enzyme activity. However the CoxG

sub-unit was found not to be essential for enzyme activity but, when the CoxG gene was deleted from the plasmid, the dehydrogenase enzyme was no longer anchored to the bacterial plasma membrane but was cytosolic (Fuhrmann et al., 2003). Hence, it was concluded that the CoxG sub-unit is necessary to anchor the enzyme to the plasma membrane where the respiratory chain is located. Because it had already been demonstrated that the GFP-*Dd*SMT (GFP-*Dd*SMT) is strongly attached to the peroxisomal membrane (Figure 5.22), it appeared possible that the first 187aa of the *Dd*SMT might have a similar role in anchoring the transferase in *D. discoideum* to the peroxisomal membrane.

To investigate the possible role of the individual domains, each was expressed alone either in *D. discoideum* or in *S cerevisiae*. The first 187aa domain was tagged with GFP at the C- terminal end, whereas the SMT domain was tagged with the same marker but at the N-terminus to leave the PTS1 at the C-terminus. The first 187aa domain was cytosolic when expressed in *D. discoideum* or in *S. cerevisiae* whereas the SMT domain was peroxisomal. The most significate aspect of these data was that the GFP-SMT domain appeared in the first supernatant of the sodium carbonate treatment experiment (Figure 5.26). Thus, by itself, this domain behaves as a peroxisomal matrix protein instead of being a peroxisomal membrane protein. This would be consistent with the theory that the function of the first 187aa domain is to anchor *Dd*SMT to the peroxisomal membrane.

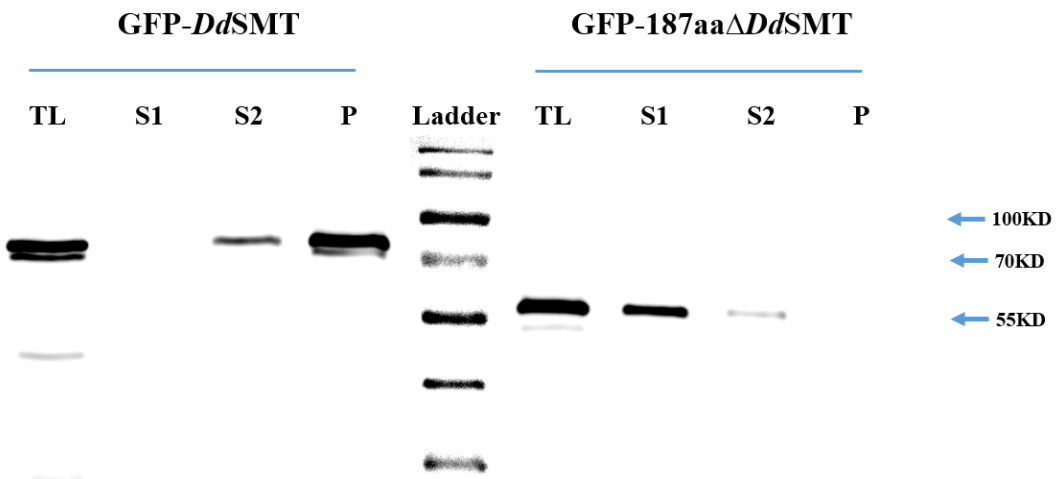


Figure 5.26: Sodium carbonate analysis of the full-length and truncated *Dd*SMT protein. Western blot analyses of lysates of *D. discoideum* expressing the full length GFP-*Dd*SMT or the SMT domain (GFP-187aaΔ*Dd*SMT) treated with 0.1M Na_2CO_3. TL: total lysate, S1: first supernatant, S2: second supernatant, P: pellet. An anti-GFP antibody was used to detect the GFP-DdSMT and the GFP-187aaΔ*Dd*SMT. Bands at the expected size, 78.3 and 57.3, were detected for the full length GFP-DdSMT and the GFP-tagged truncated *Dd*SMT respectively. Most of the full-length protein was found in the final pallet (left side of the figure). The majority of the truncated protein was detected in the first supernatant (right side of the figure).

Investigations were continued to determine whether the CoxG-like domain can bind to the peroxisomal membrane in the absence of the SMT domain. The GFP tagged CoxG-like domain, which had been extended with the PTS1 triplet -SKL so that it would be peroxisomal, was expressed in amoebae or S. cerevisiae. When lysates of the amoebae or yeast had been treated with 0.1M Na_2CO_3, it was found that only about 25% of the GFP-CoxG domain was membrane bound in amoebae and even less was membrane bound in yeast (Figure 5.27). It appeared that the CoxG-like domain has only a limited ability to bind to the matrix side of the peroxisomal membrane. However, it remains possible that both the CoxG-like and the SMT domains are synergistic in allowing the full-length DdSMT to bind to the peroxisomal membrane.

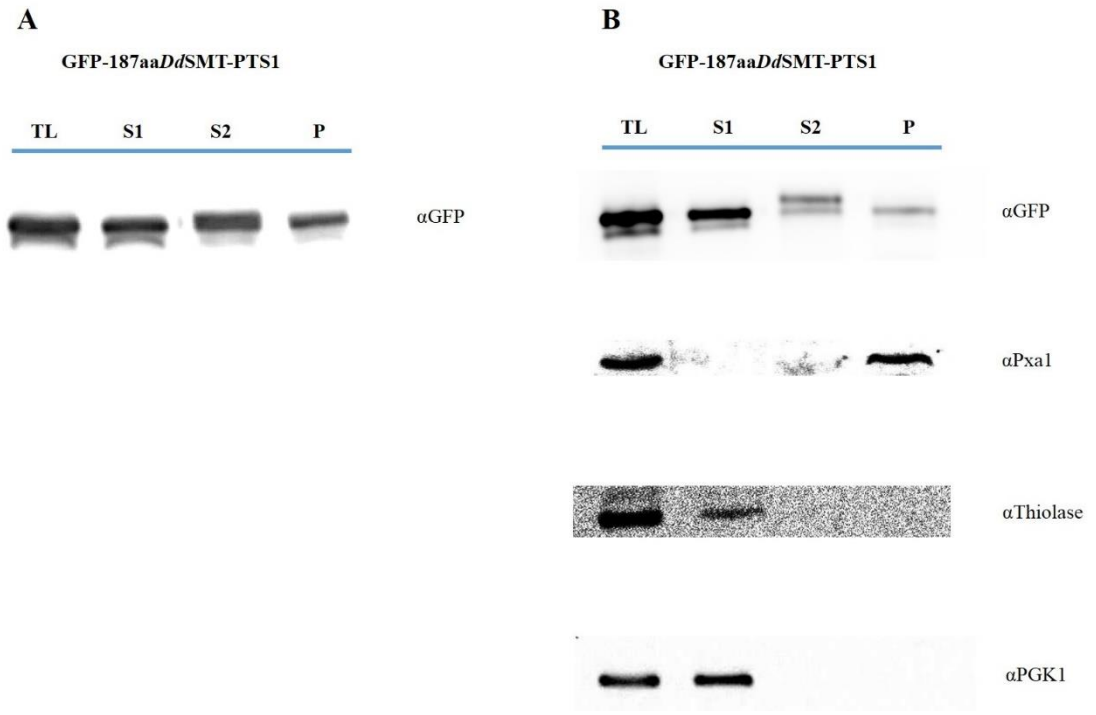


Figure 5.27: Association of the first 187aa*Dd*SMT-PTS1 with the peroxisomal membrane. The sodium carbonate technique was applied to A) *D. discoideum*, B) *S. cerevisiae.*

5.2.3.3 Use of SIM to investigate the localization of *Dd*SMT in peroxisomes

Structured Illumination Microscopy was used to identify GFP-*Dd*SMT1 in amoebae that were also expressing the red peroxisomal marker (mRFP-PTS1). Green fluorescence (GFP-*Dd*SMT) appeared as spherical blobs whereas red fluorescence was seen as a crescent around the green fluorescence. The green fluorescence was always close to only one place on the peroxisomal membrane while the red fluorescent matrix marker filled the rest of the space in the peroxisome (Figure 5.28). One possible explanation for this result is that, when overexpressed in amoebae, the *Dd*SMT enters the peroxisome and forms a membrane-bound aggregate. As a result, the matrix becomes highly compressed.

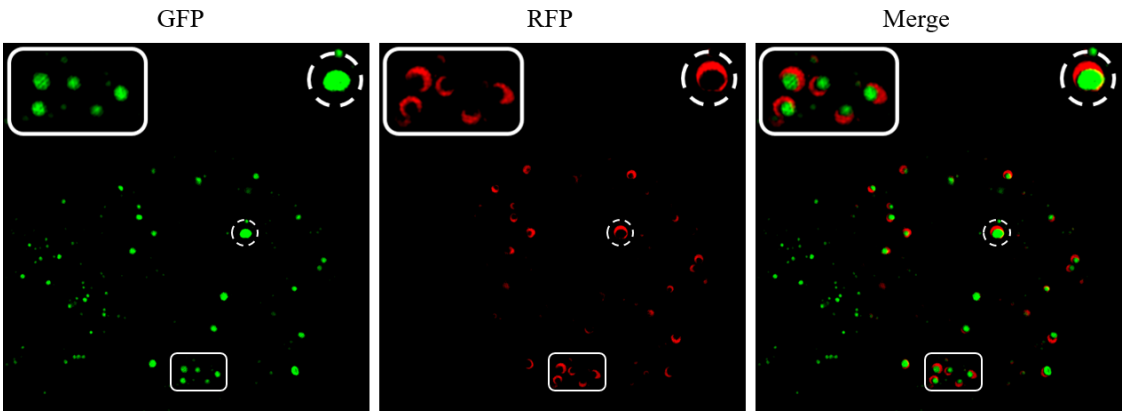


Figure 5.28: Structured Illumination Microscopy used to determine the intracellular location of *Dd*SMT. Images are of GFP-*Dd*SMT co-expressed with mRFP-PTS1. Areas outlined by the small circles and boxes are shown at higher magnification in the larger boxes and circles

5.2.3.4 Protease protection

The orientation of the *Dd*SMT enzyme in the peroxisomal membrane was investigated by using protease protection. The GFP-*Dd*SMT was digested by the proteinase K in the presence of detergent (Triton X-100) but was resistant in the absence of the detergent. The enzyme behaved during the digestions like thiolase, which is peroxisomal matrix enzyme. Since it was already known that *Dd*SMT is a membrane protein, the protease protection assay indicated that it must be bound to the matrix side of the peroxisomal membrane. The behavior of thiolase also confirmed that the peroxisomes were intact in the absence of detergent (Figure 5.29). In the absence of detergent, the pxa1 protein and PGK1 were digested by the proteinase K which was expected because part of pxa1 and all of PGK1 are extra-peroxisomal.

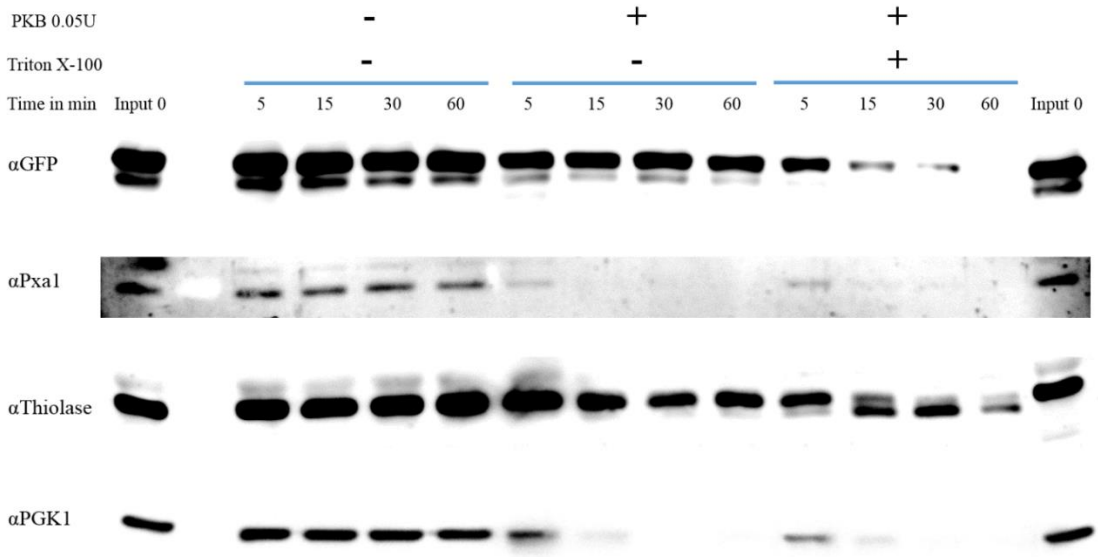


Figure 5.29: Protein orientation analysis using the protease protection technique. *D. discoideum* amoebae over-expressing GFP-*Dd*SMT were used to determine the orientation of *Dd*SMT in the peroxisomal membrane.

5.3 Discussion

5.3.1 *D. discoideum* squalene epoxidase

Two aspects of *D. discoideum* squalene epoxidase (*Dd*SQE) have been addressed in this investigation. First, the peroxisomal targeting signal on this enzyme and the pathway by which it is carried to the peroxisomes were sought and, secondly, the precise peroxisomal location of the enzyme was investigated.

Our finding confirmed that the *Dd*SQE is a peroxisomal enzyme in *D. discoideum* even though it lacks either a peroxisomal targeting signal type I or II. Furthermore, it is located in the peroxisomes when it is expressed in two different foreign hosts, S. ceivisaea and Drosophila melanogaster. These results would suggest that the enzyme makes use of a universal peroxisomal targeting signal that is conserved in the three different organisms. Moreover, tagging with a fluorescent protein either at the N- or the C-terminal end had no effect on the targeting signal. In addition, the GFP tagged *Dd*SQE showed co-localization with the red fluorescence peroxisomal marker in *Pex5*Δ and

Pex7Δ yeast strains, which lack the peroxisomal targeting signal type I and II pathway. All these results confirm that neither the PTS1 nor the PTS2 pathway is involved in taking DdSQE to the peroxisome.

Unlike the preceding enzyme, *Dd*SQS, *Dd*SQE is located on the cytosolic side of the peroxisomal membrane. This would suggest the possibility that the enzyme follows the Pex19 pathway to the peroxisome. It is known that the Pex19 protein is a chaperon and receptor responsible for aiding peroxisomal membrane protein (PMP) import. With the exception of Pex3 all newly synthesized PMPs appear to be inserted directly into the peroxisomal membrane following the Pex19 pathway (Kim and Hettema, 2015; Mayerhofer, 2016). However, immune precipitation of *Dd*SQE followed by mass amperometry did not detect co-precipitation of Pex19. Therefore, the pathway that *Dd*SQE takes to the peroxisome has still to be established.

The second aspect of *Dd*SQE that has been investigated is the role of its intracellular location in its enzymatic function. The predicted transmembrane domains of *Dd*RedA and *Dd*RedB showed a possible and strong TMD at the N terminal ends. It is predicted to start from amino acid 20 till 41 in *Dd*RedA. In a similar pattern, it runs from 12 to 23 in *Dd*RedB. However, there is not any predicted TMD in the *Dd*RedC. These results agree with the active domain prediction finding, which showed that the first predicted domain (binding domain) start from the amino acid in the position number 75 and 74 for *Dd*RedA and *Dd*RedB respectively. However, the flavodoxin domain in *Dd*RedC starts at amino acid in position number 7.

It has been showed that squalene epoxidase is an FAD dependent enzyme and also requires NADPH for its activity (Ono et al., 1977, 1980). Because SQE lacks an NADPH binding site, it needs to operate together with an enzyme that can bind NADPH and donate reducing equivalents from the NADPH to the FAD on the SQE. The donor protein has been shown to be cytochrome P450 reductase (Ono et al., 1977, 1980, 1982).

It has been shown (Figure 5.12) that, of the two cytochrome p450 reductases (Gonzalez-Kristeller et al., 2008) in *D. discoideum*, neither is peroxisomal but that both are located in the ER. This may explain why *Dd*SQE lacks a PTS and is on the cytosolic side of the peroxisomal membrane whereas the preceding enzyme (*Dd*SQS) and the two following

enzymes (*Dd*OSC and *Dd*SMT) on the sterol biosynthesis pathway are on the matrix side of the peroxisomal membrane. *Dd*SQE has to be on the cytosolic side of the peroxisomal membrane if it is to interact with cytochrome P450 reductase and has to be associated with the peroxisomal membrane because this is where its substrate, squalene, will be located.

The SIM imaging of amoebae co-expressing both *Dd*SQE and *Red*A/ *Red*B showed that the enzymes are probably close enough for transport of electrons from either *Dd*RedA or *Dd*RedB to the *Dd*SQE. However, these results do not prove that the reductases are definitely the electron donors to *Dd*SQE.

5.3.2 *Dd*oxidosqualene cyclase (*Dd*OSC)

*Dd*oxidosqualene cyclase has the peroxisomal targeting signal 1 (SKI), which appears to be limited to use in plants (Reumann et al., 2016). Even among the slime moulds, -SKI is found in only two cycloartenol synthases - in those of *Dictyostelium lacteum* and *Polysphondylium pallidum*. However, alternative PTS1 triplets are present in the cyclase sequences of some of the other slime moulds.

It has previously been shown that *Dd*OSC is a peroxisomal membrane protein and is most probably a monotopic protein (Alkuwayti 2014). Remarkably, SIM imaging has shown that whereas *Dd*SQS and *Dd*SQE are present throughout the peroxisomal membrane, the *Dd*OSC is apparently located only at a specific point on the peroxisomal membrane. The reasons for this remain unknown.

5.3.3 *D. discoideum* SMT

The study of the peroxisomal location of *Dd*SMT gave somewhat puzzling results. It was very clear that the enzyme is bound to the peroxisomal membrane (Figure 5.22) but it was not possible to establish unequivocally which domain of the enzyme is involved in the membrane binding. It appeared, when the SMT domain was expressed alone, that this domain is unable to bind to the peroxisomal membrane (Figure 5.26). On the other hand, although it was possible to demonstrate that the CoxG-like domain has some ability to bind to the peroxisomal membrane, this ability appeared rather limited (Figure

5.27). However, these results could imply that both domains are involved in the membrane binding.

The other puzzling aspect of *Dd*SMT was revealed by SIM imaging where the GFP tagged enzyme appeared as a large blob against a limited region of the peroxisomal membrane. The peroxisomal matrix was also largely filled by the GFP-*Dd*SMT (Figure 5.28). It would seem improbable that all of the GFP-*Dd*SMT can be involved in methylation of cycloartenol because much of the enzyme appears not to be in contact with the peroxisomal membrane where the cycloartenol substrate is located. It is possible, however, that the large blobs of the GFP-*Dd*SMT are formed only because of overexpression of the enzyme. It could be helpful to investigate the normal localization of the endogenous *Dd*SMT by use of a specific antibody. It would also be interesting to determine how close the peroxisomal membrane binding site of *Dd*SMT is the binding site for *Dd*OSC.

Chapter six "General discussion and future directions"

6.1 General discussion

A remarkable feature of sterol biosynthesis from mevalonate in *D. discoideum* is the location of the first four enzymes in the peroxisomes (Figure 6.1). In all other organisms in which these enzymes have been previously investigated, it is been found that they are integral membrane proteins on the ER (Nes, 2011). It is believed that these enzymes are usually associated with the ER because the reaction products are highly hydrophobic and need to be temporarily stored in a lipid environment (Pandit et al., 2000). It would seem reasonable to conclude that, as shown in earlier chapters, the four enzymes in *D. discoideum* are associated with the peroxisomal membrane for the same reason. It is, however, difficult to explain why these enzymes are peroxisomal in *D. discoideum*, instead of being in the ER, although it may be advantageous to have SQS in the organelle in which its substrate, farnesyl diphosphate, is synthesized (Nuttall et al., 2012). It is more difficult to account for the presence of *Dd*SQE, *Dd*OSC and DdSMT in the peroxisomes. Thus, as described in section (5.2.1.4), the association of *Dd*SQE with the peroxisome appears to be potentially disadvantageous because of the requirement to interact with a cytochrome P450 reductase in the ER. The localization of the *Dd*SQE to the cytosolic side of the peroxisomal membrane, which contrasts with the localization of the other three enzymes to the matrix side of the peroxisomal membrane, appears necessary to allow the *Dd*SQE reaction to proceed.

The other remarkable feature of sterol biosynthesis in *D. discoideum* is the use of PTS1 pathway to allow *Dd*SQS, *Dd*OSC and *Dd*SMT to associate with the peroxisomal membrane. There are no other examples of peroxisomal membrane proteins that make use of a PTS pathway to reach the peroxisomal membrane. Previously, it has been thought that the pathways by which proteins are assembled into the peroxisomal membrane are independent of the PTS pathways. They also always insert the membrane proteins from the cytosolic side of the peroxisomal membrane (Kim and Hettema, 2015). It is probable, therefore, that *Dd*SQS, *Dd*OSC and *Dd*SMT make use of the PTS1 pathway because they need to enter the peroxisome and associate with the peroxisomal membrane from the matrix side. It is apparent that *Dd*SQS has to be inserted into the matrix side of the peroxisomal membrane so that its active site is able to bind to farnesyl diphosphate produced in the peroxisomal matrix. However, it is more difficult to

understand why *Dd*OSC and *Dd*SMT have to be associated with the matrix side of the peroxisomal membrane.

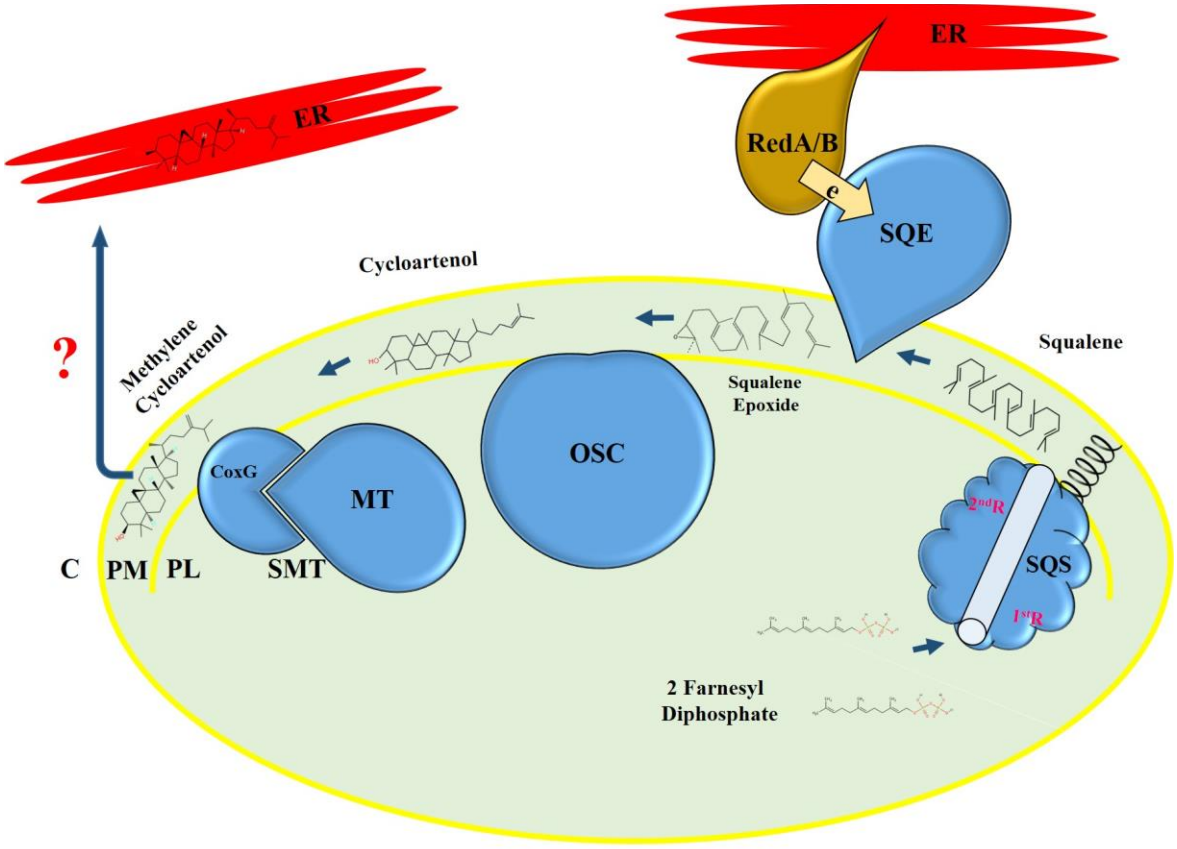


Figure 6.1: Diagrammatic summary of the first four steps of sterol biosynthesis from FDP in *D. discoideum*. C, cytosol; PM, peroxisomal membrane; PL, peroxisomal lumen; 1^{st}R and 2^{nd}R, first and second reactions catalysed by squalene synthase.

The fifth reaction on the pathway from FDP to end-product sterols is catalyzed by methylsterol monooxygenase (*Dd*MSM). There are two *Dd*MSM isozymes and it has been shown previously (Alkuwayti, 2014) that both are located in the ER. The methylcycloartenol synthesized in the peroxisomes has therefore to pass to the ER. There are two possible ways in which this might occur. First, lipid vesicles containing the methylcycloartenol could bud off from the peroxisomal membrane and carry the sterol to the ER. However, there is, at present no evidence that vesicles can form from the peroxisomal membrane. The second mechanism by which sterols are known to move around the cell and pass through membranes involves transport by steroidogenic acute regulatory protein (StAR) and StAR-like proteins (Clark, 2012; Lev, 2012; Miller, 2017; Soccio and Breslow, 2004). These proteins have a three-dimensional structure similar to that of CoxG (Radauer et al., 2008; Thorsell et al., 2011). Hence the first domain of *Dd*SMT has also strong similarity with StAR proteins (Figure 6.2). It is, therefore, possible that the CoxG domain of *Dd*SMT is involved in export of the methylcycloartenol from the peroxisome to the ER as well as in anchoring the *Dd*SMT to the peroxisomal membrane.

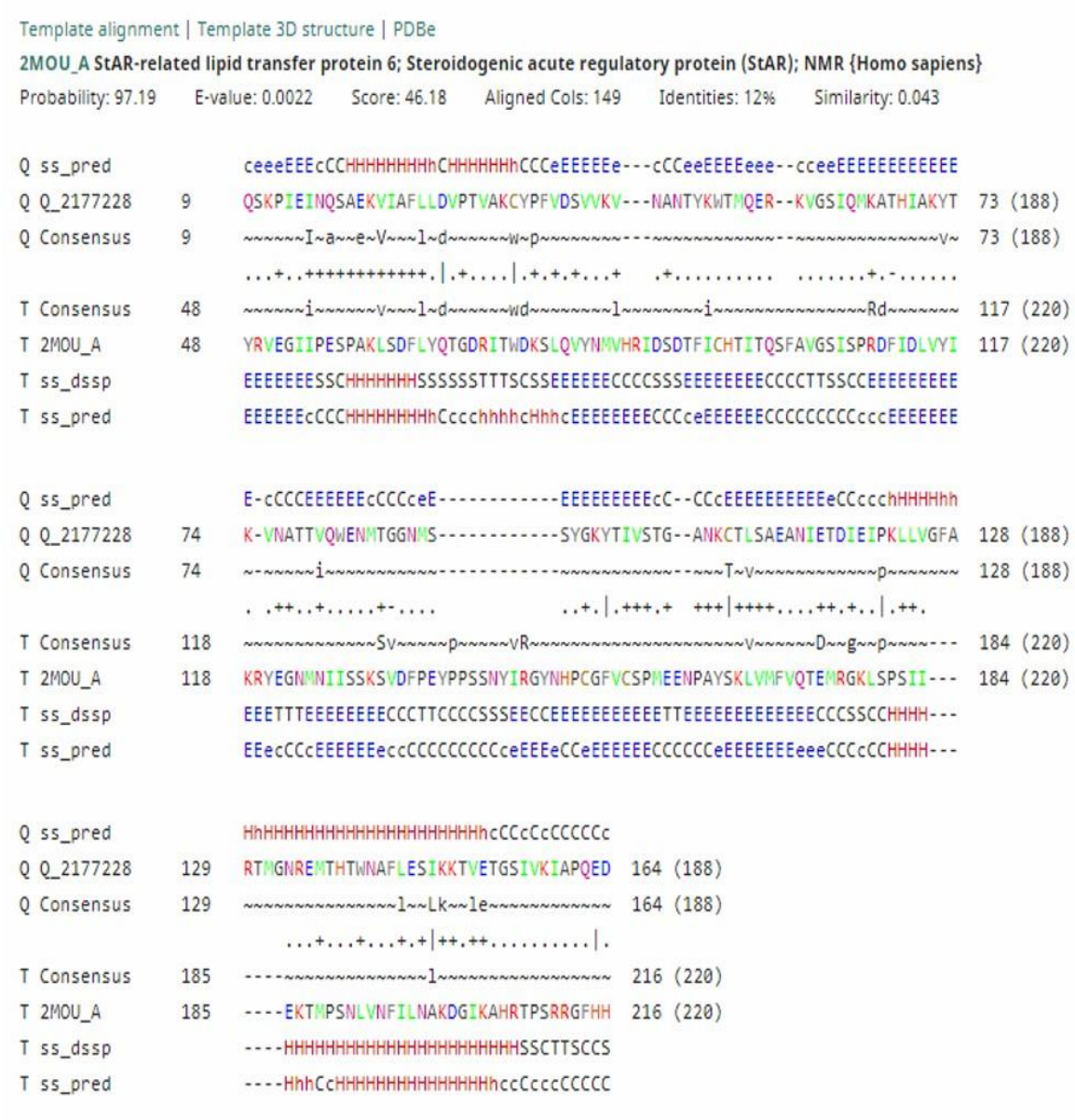


Figure 6.2: comparison of the structures of the CoxG domain of *Dd*SMT and a StAR domain. The secondary structure of the 187aa of the *Dd*SMT CoxG domain was obtained by using HHpred from sequence analysis in the toolkit web site (Alva et al., 2016). The figure shows that there is a 97% probability that the secondary structure of the *Dd*SMT CoxG domain is similar to that of the *Homo sapiens* StAR domain.

There is a belief that there are contact sites between peroxisomes and other cellular organelles (Mattiazzi Ušaj et al., 2015; Schrader et al., 2015) although the nature of these contact sites remains unknown. Such contact sites between the ER and the peroxisomes could therefore be involved in movement of the methylcycloartenol from the peroxisomal membrane to the ER. In addition, contact sites between the peroxisome and the ER could also be of importance in bringing *Dd*SQE and a cytochrome P450 reductase into contact. Further investigation of possible interaction between *Dd*SQE and the cytochrome P450 reductase could therefore help to provide evidence for the existence of contact sites and provide better understanding of their structure.

6.2 Future directions

Further study of a possible interaction between *Dd*SQE and Pex19 could clarify the pathway by which *Dd*SQE is reaching the peroxisome. It would be possible to demonstrate that Pex19 does bind *Dd*SQE if a modified version of Pex19 having NLS would take *Dd*SQE-GFP to the nucleus.

Further experiments could be done in order to confirm the interaction between *Dd*SQE and cytochrome P450 reductase, for example by using the split GFP technique. This may also provide fundamental information for further understanding of contact sites between peroxisomes and other cellular organelles.

Production of specific antibodies for *Dd*OSC and *Dd*SMT would help to determine the peroxisomal location of these two enzymes in the absence of overexpression. By use of the antibodies in indirect immunofluorescence microscopy, it should be possible to observe the natural location of these proteins in the peroxisomal membrane.

References

Alkuwayti, M.A.M. (2014). The role of peroxisomes in sterol biosynthesis by the cellular slime mould *Dictyostelium discoideum*. PhD Thesis, Sheffield University, UK.

Alva, V., Nam, S., Söding, J., and Lupas, A.N. (2016). The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. Nucleic Acids Res. *44*, 410–415.

Baker, A., and Sparkes, I.A. (2005). Peroxisome protein import: some answers, more questions. Curr. Opin. Plant Biol. *8*, 640–647.

Barth, C., Fraser, D.J., and Fisher, P.R. (1998). Co-insertional replication is responsible for tandem multimer formation during plasmid integration into the Dictyostelium genome. Plasmid *39*, 141–153.

Belter, A., Skupinska, M., Giel-Pietraszuk, M., Grabarkiewicz, T., Rychlewski, L., and Barciszewski, J. (2011). Squalene monooxygenase – a target for hypercholesterolemic therapy. Biol. Chem *392*, 1053–1075.

Benford, H.L., Frith, J.C., Auriola, S., Monkkonen, J., and Rogers, M.J. (1999). Farnesol and geranylgeraniol prevent activation of caspases by aminobisphosphonates: biochemical evidence for two distinct pharmacological classes of bisphosphonate drugs. Mol Pharmacol *56*, 131–140.

Benveniste, P. (2004). Biosynthesis and accumulation of sterols. Annu. Rev. Plant Biol. *55*, 429–457.

Bergeron, J.J., Brenner, M.B., Thomas, D.Y., and Williams, D.B. (1994). Calnexin: a membrane-bound chaperone of the endoplasmic reticulum. Trends Biochem. Sci. *19*, 124–128.

Bergstrom, J.D., Bostedor, R.G., Masarachia, P.J., Reszka, A.A., and Rodan, G. (2000). Alendronate is a specific, nanomolar inhibitor of farnesyl diphosphate synthase. Arch. Biochem. Biophys. *373*, 231–241.

Van Berkel, W.J.H., Kamerbeek, N.M., and Fraaije, M.W. (2006). Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. J. Biotechnol. *124*, 670–689.

Biardi, L., Sreedhar, A., Zokaei, A., Vartak, N.B., Bozeat, R.L., Shackelford, J.E., Keller, G.A., and Krisans, S.K. (1994). Mevalonate kinase is predominantly localized in peroxisomes and is defective in patients with peroxisome deficiency disorders. J. Biol. Chem. *269*, 1197–1205.

Borgese, N., Colombo, S., and Pedrazzini, E. (2003). The tale of tail-anchored proteins: Coming from the cytosol and looking for a membrane. J. Cell Biol.

van den Bosch, H., Schutgens, R.B.H., Wanders, R.J.A., and Tager, J.M. (1992). Biochemistry of Peroxisomes. Annu. Rev. Biochem. *61*, 157–197.

Bothmer, A., Phadke, T., Barrera, L.A., Margulies, C.M., Lee, C.S., Buquicchio, F., Moss, S., Abdulkerim, H.S., Selleck, W., Jayaram, H., et al. (2017). Characterization of the interplay between DNA repair and CRISPR/Cas9-induced DNA lesions at an endogenous locus. Nat. Commun. *8*, 13905.

Böttcher, R., Hollmann, M., Merk, K., Nitschko, V., Obermaier, C., Philippou-Massier, J., Wieland, I., Gaul, U., and Förstemann, K. (2014). Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured *Drosophila* cells. Nucleic Acids Res. *42* (11), e89.

Bouvier-Navé, P., Husselstein, T., and Benveniste, P. (1998). Two families of sterol methyltransferases are involved in the first and the second methylation steps of plant sterol biosynthesis. Eur. J. Biochem. *256*, 88–96.

Braverman, N.E., D'Agostino, M.D., and MacLean, G.E. (2013). Peroxisome biogenesis disorders: Biological, clinical and pathophysiological perspectives. Dev. Disabil. Res. Rev. *17*, 187–196.

Brown, T.W., Titorenko, V.I., and Rachubinski, R.A. (2000). Mutants of the Yarrowia lipolytica PEX23 gene encoding an integral peroxisomal membrane peroxin mislocalize matrix proteins and accumulate vesicles containing peroxisomal matrix and membrane proteins. Mol. Biol. Cell *11*, 141–152.

Brumfield, K.M., Laborde, S.M., and Moroney, J. V. (2017). A model for the ergosterol biosynthetic pathway in *Chlamydomonas reinhardtii*. Eur. J. Phycol. *52*, 64–74.

Brycki, B., Koenig, H., Kowalczyk, I., and Pospieszny, T. (2014). Synthesis,

Spectroscopic and Theoretical Studies of New Dimeric Quaternary Alkylammonium Conjugates of Sterols. Molecules *19*, 9419–9434.

Busquets, A., Keim, V., Closa, M., del Arco, A., Boronat, A., Arró, M., and Ferrer, A. (2008). *Arabidopsis thaliana* contains a single gene encoding squalene synthase. Plant Mol. Biol. *67*, 25–36.

Carte, J., Christopher, R.T., Smith, J.T., Olson, S., Barrangou, R., Moineau, S., Glover, C.V.C., Graveley, B.R., Terns, R.M., and Terns, M.P. (2014). The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *S treptococcus thermophilus*. Mol. Microbiol. *93*, 98–112.

Chen, D.C., Yang, B.C., and Kuo, T.T. (1992). One-step transformation of yeast in stationary phase. Curr. Genet. *21*, 83–84.

Clark, B.J. (2012). The mammalian START domain protein family in lipid transport in health and disease. J. Endocrinol. *212*, 257–275.

Cummings, S.R., Kelsey, J.L., Nevitt, M.C., and O'Dowd, K.J. (1985). Epidemiology of osteoporosis and osteoporotic fractures. Epidemiol. Rev. *7*, 178–208.

Darnet, S., and Rahier, A. (2004). Plant sterol biosynthesis: identification of two distinct families of sterol 4alpha-methyl oxidases. Biochem. J. *378*, 889–898.

Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K., and Semenova, E. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nat. Commun. *3*, 945-951.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602–607.

DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. *41*, 4336–4343.

Distel, B., Erdmann, R., Gould, S.J., Blobel, G., Crane, D.I., Cregg, J.M., Dodt, G., Fujiki, Y., Goodman, J.M., Just, W.W., et al. (1996). A unified nomenclature for peroxisome biogenesis factors. J. Cell Biol. *135*, 1–3.

Dyer, J.M., McNew, J.A., and Goodman, J.M. (1996). The sorting sequence of the peroxisomal integral membrane protein PMP47 is contained within a short hydrophilic loop. J. Cell Biol. *133*, 269–280.

Elgersma, Y., Kwast, L., Klein, A., Voorn-Brouwer, T., van den Berg, M., Metzig, B., America, T., Tabak, H.F., and Distel, B. (1996). The SH3 domain of the *Saccharomyces cerevisiae* peroxisomal membrane protein Pex13p functions as a docking site for Pex5p, a mobile receptor for the import PTS1-containing proteins. J. Cell Biol. *135*, 97–109.

Elgersma, Y., Kwast, L., van den Berg, M., Snyder, W.B., Distel, B., Subramani, S., and Tabak, H.F. (1997). Overexpression of Pex15p, a phosphorylated peroxisomal integral membrane protein required for peroxisome assembly in *Saccharomyces cerevisiae*, causes proliferation of the endoplasmic reticulum membrane. EMBO J. *16*, 7326–7341.

Emmanouilidis, L., Schütz, U., Tripsianes, K., Madl, T., Radke, J., Rucktäschel, R., Wilmanns, M., Schliebs, W., Erdmann, R., and Sattler, M. (2017). Allosteric modulation of peroxisomal membrane protein recognition by farnesylation of the peroxisomal import receptor PEX19. Nat. Commun. *8*, 14635-14647.

Emmerstorfer, A., Wimmer-Teubenbacher, M., Wriessnegger, T., Leitner, E., Müller, M., Kaluzna, I., Schürmann, M., Mink, D., Zellnig, G., Schwab, H., et al. (2015). Over-expression of *ICE2* stabilizes cytochrome P450 reductase in *Saccharomyces cerevisiae* and *Pichia pastoris*. Biotechnol. J. *10*, 623–635.

Endo, A. (2010). A historical perspective on the discovery of statins. Proc. Japan Acad. Ser. B *86*, 484–493.

Entsch, B., and van Berkel, W.J. (1995). Structure and mechanism of para-hydroxybenzoate hydroxylase. FASEB J. *9*, 476–483.

Erdmann, R., and Blobel, G. (1995). Giant peroxisomes in oleic acid-induced *Saccharomyces cerevisiae* lacking the peroxisomal membrane protein Pmp27p. J. Cell Biol. *128*, 509–523.

Erdmann, R., Veenhuis, M., Mertens, D., and Kunau, W.H. (1989). Isolation of peroxisome-deficient mutants of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U. S. A. *86*, 5419–5423.

Erdmann, R., Wiebel, F.F., Flessau, A., Rytka, J., Beyer, A., Fröhlich, K.U., and Kunau, W.H. (1991). PAS1, a yeast gene required for peroxisome biogenesis, encodes a member of a novel family of putative ATPases. Cell *64*, 499–510.

Fang, Y., Morrell, J.C., Jones, J.M., and Gould, S.J. (2004). PEX3 functions as a PEX19 docking factor in the import of class I peroxisomal membrane proteins. J. Cell Biol. *164*, 863–875.

Farooqui, A.A., and Horrocks, L.A. (2001). Book Review: Plasmalogens: Workhorse Lipids of Membranes in Normal and Injured Neurons and Glia. Neurosci. *7*, 232–245.

Fegueur, M., Richard, L., Charles, A.D., and Karst, F. (1991). Isolation and primary structure of the ERG9 gene of *Saccharomyces cerevisiae* encoding squalene synthetase. Curr. Genet. *20*, 365–372.

Fleisch, H. (1998). Bisphosphonates : Mechanisms of Action. Endocr. Rev. *19*, 80–100.

Fransen, M., Wylin, T., Brees, C., Mannaerts, G.P., and Van Veldhoven, P.P. (2001). Human pex19p binds peroxisomal integral membrane proteins at regions distinct from their sorting sequences. Mol. Cell. Biol. *21*, 4413–4424.

Fuhrmann, S., Ferner, M., Jeffke, T., Henne, A., Gottschalk, G., and Meyer, O. (2003). Complete nucleotide sequence of the circular megaplasmid pHCG3 of Oligotropha carboxidovorans: function in the chemolithoautotrophic utilization of CO, H2 and CO2. Gene *322*, 67–75.

Fujihara, Y., and Ikawa, M. (2014). CRISPR/Cas9-Based Genome Editing in Mice by Single Plasmid Injection. In Methods in Enzymology, pp. 319–336.

Fujiki, Y., Hubbard, A.L., Fowler, S., and Lazarow, P.B. (1982). Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. J. Cell Biol. *93*, 97–102.

Fujiki, Y., Yagita, Y., and Matsuzaki, T. (2012). Peroxisome biogenesis disorders: molecular basis for impaired peroxisomal membrane assembly: in metabolic functions and biogenesis of peroxisomes in health and disease. Biochim. Biophys. Acta *1822*, 1337–1342.

García-Unzueta, M., Riancho, J., Zarrabeitia, M., Sañudo, C., Berja, A., Valero, C.,

Pesquera, C., Paule, B., González-Macías, J., and Amado, J. (2008). Association of the 163A/G and 1181G/C Osteoprotegerin Polymorphism with Bone Mineral Density. Horm. Metab. Res. *40*, 219–224.

Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67–71.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. *109*, 2579-2586.

Geng, L., Yao, Z., Yang, H., Luo, J., Han, L., and Lu, Q. (2007). Association of CA Repeat Polymorphism in Estrogen Receptor ?? Gene with Postmenopausal Osteoporosis in Chinese. J. Genet. Genomics *34*, 868–876.

Ghisla, S., and Massey, V. (1989). Mechanisms of flavoprotein-catalyzed reactions. Eur. J. Biochem. *181*, 1–17.

Gietz, R.D., and Woods, R.A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. *350*, 87–96.

Glover, J.R., Andrews, D.W., Subramani, S., and Rachubinski, R.A. (1994). Mutagenesis of the amino targeting signal of *Saccharomyces cerevisiae* 3-ketoacyl-CoA thiolase reveals conserved amino acids required for import into peroxisomes in vivo. J. Biol. Chem. *269*, 7558–7563.

Godio, R.P., Fouces, R., and Martín, J.F. (2007). A Squalene Epoxidase Is Involved in Biosynthesis of Both the Antitumor Compound Clavaric Acid and Sterols in the Basidiomycete *H. sublateritium*. J. Chem. and Biol. *14*, (12) 1334-1346

Goldfischer, S., Moore, C.L., Johnson, A.B., Spiro, A.J., Valsamis, M.P., Wisniewski, H.K., Ritch, R.H., Norton, W.T., Rapin, I., and Gartner, L.M. (1973). Peroxisomal and mitochondrial defects in the cerebro-hepato-renal syndrome. Science *182*, 62–64.

Goldstein, J.L., and Brown, M.S. (1990). Regulation of the mevalonate pathway. Nature *343*, 425–430.

Gonzalez-Kristeller, D.C., Farage, L., Fiorini, L.C., Loomis, W.F., and da Silva, A.M.

(2008). The P450 oxidoreductase, RedA, controls development beyond the mound stage in *Dictyostelium discoideum*. BMC Dev. Biol. *8*, 8-18.

Götte, K., Girzalsky, W., Linkert, M., Baumgart, E., Kammerer, S., Kunau, W.H., and Erdmann, R. (1998). Pex19p, a farnesylated protein essential for peroxisome biogenesis. Mol. Cell. Biol. *18*, 616–628.

Gould, S.J., and Valle, D. (2000). Peroxisome biogenesis disorders: genetics and cell biology. Trends Genet. *16*, 340–345.

Gould, S.G., Keller, G.A., and Subramani, S. (1987). Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. J. Cell Biol. *105*, 2923–2931.

Gould, S.J., Keller, G.A., and Subramani, S. (1988). Identification of peroxisomal targeting signals located at the carboxy terminus of four peroxisomal proteins. J. Cell Biol. *107*, 897–905.

Gould, S.J., Keller, G.A., Hosken, N., Wilkinson, J., and Subramani, S. (1989). A conserved tripeptide sorts proteins to peroxisomes. J. Cell Biol. *108*, 1657–1664.

Gould, S.J., Keller, G.A., Schneider, M., Howell, S.H., Garrard, L.J., Goodman, J.M., Distel, B., Tabak, H., and Subramani, S. (1990). Peroxisomal protein import is conserved between yeast, plants, insects and mammals. EMBO J. *9*, 85–90.

Gould, S.J., McCollum, D., Spong, A.P., Heyman, J.A., and Subramani, S. (1992). Development of the yeast Pichia pastoris as a model organism for a genetic and molecular analysis of peroxisome assembly. Yeast *8*, 613–628.

Green, J.R. (2004). Bisphosphonates: Preclinical Review. Oncologist *9*, 3–13.

Grissa, I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics *8*, 172-182.

Grove, J.E., Brown, R.J., and Watts, D.J. (2000). The intracellular target for the antiresorptive aminobisphosphonate drugs in *Dictyostelium discoideum* is the enzyme farnesyl diphosphate synthase. J. Bone Miner. Res. *15*, 971–981.

Gu, P., Ishii, Y., Spencer, T.A., and Shechter, I. (1998). Function-structure studies and identification of three enzyme domains involved in the catalytic activity in rat hepatic squalene synthase. J. Biol. Chem. *273*, 12515–12525.

Gutierrez, A., Grunau, A., Paine, M., Munro, a W., Wolf, C.R., Roberts, G.C.K., and Scrutton, N.S. (2003). Electron transfer in human cytochrome P450 reductase. Biochem. Soc. Trans. *31*, 497–501.

Han, J.-Y., In, J.-G., Kwon, Y.-S., and Choi, Y.-E. (2009). Regulation of ginsenoside and phytosterol biosynthesis by RNA interferences of squalene epoxidase gene in Panax ginseng. Phytochemistry *71*, 36–46.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. *166*, 557–580.

Hartmann, M. (1998). Plant sterols and the membrane environment. Trends Plant Sci. *3*, 170–175.

He, F., Zhu, Y., He, M., and Zhang, Y. (2008). Molecular cloning and characterization of the gene encoding squalene epoxidase in Panax notoginseng. DNA Seq. *19*, 270–273.

Hervet, E., Charpentier, X., Vianney, A., Lazzaroni, J.-C., Gilbert, C., Atlan, D., and Doublet, P. (2011). Protein kinase LegK2 is a type IV secretion system effector involved in endoplasmic reticulum recruitment and intracellular replication of *Legionella pneumophila*. Infect. Immun. *79*, 1936–1950.

Hettema, E.H., Distel, B., and Tabak, H.F. (1999). Import of proteins into peroxisomes. Biochim. Biophys. Acta - Mol. Cell Res. *1451*, 17–34.

Hettema, E.H., Girzalsky, W., van Den Berg, M., Erdmann, R., and Distel, B. (2000). *Saccharomyces cerevisiae* pex3p and pex19p are required for proper localization and stability of peroxisomal membrane proteins. EMBO J. *19*, 223–233.

Hoepfner, D., Schildknegt, D., Braakman, I., Philippsen, P., and Tabak, H.F. (2005). Contribution of the Endoplasmic Reticulum to Peroxisome Formation. Cell *122*, 85–95.

Höhfeld, J., Veenhuis, M., and Kunau, W.H. (1991). PAS3, a *Saccharomyces cerevisiae* gene encoding a peroxisomal integral membrane protein essential for peroxisome biogenesis. J. Cell Biol. *114*, 1167–1178.

Honsho, M., and Fujiki, Y. (2001). Topogenesis of peroxisomal membrane protein requires a short, positively charged intervening-loop sequence and flanking hydrophobic segments. study using human membrane protein PMP34. J. Biol. Chem. *276*, 9375–9382.

Honsho, M., Tamura, S., Shimozawa, N., Suzuki, Y., Kondo, N., and Fujiki, Y. (1998a). Mutation in PEX16 is causal in the peroxisome-deficient Zellweger syndrome of complementation group D. Am. J. Hum. Genet. *63*, 1622–1630.

Honsho, M., Mitoma, J.Y., and Ito, A. (1998b). Retention of cytochrome b5 in the endoplasmic reticulum is transmembrane and luminal domain-dependent. J. Biol. Chem. *273*, 20860–20866.

Honsho, M., Hiroshige, T., and Fujiki, Y. (2002). The membrane biogenesis peroxin Pex16p. Topogenesis and functional roles in peroxisomal membrane assembly. J. Biol. Chem. *277*, 44513–44524.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832.

Hua, R., Gidda, S.K., Aranovich, A., Mullen, R.T., and Kim, P.K. (2015). Multiple Domains in PEX16 Mediate Its Trafficking and Recruitment of Peroxisomal Proteins to the ER. Traffic *16*, 832–852.

Huhse, B., Rehling, P., Albertini, M., Blank, L., Meller, K., and Kunau, W.H. (1998). Pex17p of *Saccharomyces cerevisiae* is a novel peroxin and component of the peroxisomal protein translocation machinery. J. Cell Biol. *140*, 49–60.

Jiang, D., Rong, Q., Chen, Y., Yuan, Q., Shen, Y., Guo, J., Yang, Y., Zha, L., Wu, H., Huang, L., et al. (2017). Molecular cloning and functional analysis of squalene synthase (SS) in Panax notoginseng. Int. J. Biol. Macromol. *95*, 658–666.

Jiang, Y., Chen, H., Chen, X., Köllner, T.G., Jia, Q., Wymore, T.W., Wang, F., and Chen, F. (2015). Volatile squalene from a nonseed plant *Selaginella moellendorffii* : Emission and biosynthesis. Plant Physiol. Biochem. *96*, 1–8.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E.

(2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science (80-.). *337*, 816–821.

Jones, J.M., Morrell, J.C., and Gould, S.J. (2001). Multiple distinct targeting signals in integral peroxisomal membrane proteins. J. Cell Biol. *153*, 1141–1150.

Jones, J.M., Morrell, J.C., and Gould, S.J. (2004). PEX19 is a predominantly cytosolic chaperone and import receptor for class 1 peroxisomal membrane proteins. J. Cell Biol. *164*, 57–67.

Joosten, V., van Berkel, W.J., author, C., and Berkel, van (2007). Flavoenzymes. Curr. Opin. Chem. Biol. *11*, 195–202.

Jore, M.M., Brouns, S.J.J., and van der Oost, J. (2012). RNA in Defense: CRISPRs Protect Prokaryotes against Mobile Genetic Elements. Cold Spring Harb. Perspect. Biol. *4*, (6) 3657–3668.

Kajikawa, M., Kinohira, S., Ando, A., Shimoyama, M., Kato, M., and Fukuzawa, H. (2015). Accumulation of squalene in a microalga Chlamydomonas reinhardtii by genetic modification of squalene synthase and squalene epoxidase genes. PLoS One *10*, (3) e0120446.

Kalish, J.E., Theda, C., Morrell, J.C., Berg, J.M., and Gould, S.J. (1995). Formation of the peroxisome lumen is abolished by loss of *Pichia pastoris* Pas7p, a zinc-binding integral membrane protein of the peroxisome. Mol. Cell. Biol. *15*, 6406–6419.

Kalish1, J.E., Keller2, G.-A., Morrell1, J.C., Mihalik3, S.J., Smith4, B., Cregg4, J.M., and Gould1, S.J. (1996). Characterization of a novel component of the peroxisomal protein import apparatus using fluorescent peroxisomal proteins. EMBO J. *1513*, 3275–3285.

Kalra, S., Kumar, S., Lakhanpal, N., Kaur, J., and Singh, K. (2013). Characterization of Squalene synthase gene from *Chlorophytum borivilianum* (Sant. and Fernand.). Mol. Biotechnol. *54*, 944–953.

Kammerer, S., Arnold, N., Gutensohn, W., Mewes, H.W., Kunau, W.H., Höfler, G., Roscher, A.A., and Braun, A. (1997). Genomic organization and molecular characterization of a gene encoding HsPXF, a human peroxisomal farnesylated protein.

Genomics *45*, 200–210.

Karnik, S.K., and Trelease, R.N. (2007). Arabidopsis peroxin 16 trafficks through the ER and an intermediate compartment to pre-existing peroxisomes via overlapping molecular targeting signals. J. Exp. Bot. *58*, 1677–1693.

Keller, G.A., Gould, S., Deluca, M., and Subramani, S. (1987). Firefly luciferase is targeted to peroxisomes in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. *84*, 3264–3268.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. *10*, 845–858.

Kemp, S., and Wanders, R.J.A. (2007). X-linked adrenoleukodystrophy: very long-chain fatty acid metabolism, ABC half-transporters and the complicated route to treatment. Mol. Genet. Metab. *90*, 268–276.

Kim, P.K., and Hettema, E.H. (2015). Multiple Pathways for Protein Transport to Peroxisomes. J. Mol. Biol. *427*, 1176–1190.

Kim, P.K., Mullen, R.T., Schumann, U., and Lippincott-Schwartz, J. (2006). The origin and maintenance of mammalian peroxisomes involves a de novo PEX16-dependent pathway from the ER. J. Cell Biol. *173*, 521–532.

Koller, A., Snyder, W.B., Faber, K.N., Wenzel, T.J., Rangell, L., Keller, G.A., and Subramani, S. (1999). Pex22p of *Pichia pastoris*, essential for peroxisomal matrix protein import, anchors the ubiquitin-conjugating enzyme, Pex4p, on the peroxisomal membrane. J. Cell Biol. *146*, 99–112.

Kovacs, W.J., Olivier, L.M., and Krisans, S.K. (2002). Central role of peroxisomes in isoprenoid biosynthesis. Prog. Lipid Res. *41*, 369–391.

Krisans, S.K., Ericsson, J., Edwards, P.A., and Keller, G.A. (1994). Farnesyl-diphosphate synthase is localized in peroxisomes. J. Biol. Chem. *269*, 14165–14169.

Lametschwandtner, G., Brocard, C., Fransen, M., Van Veldhoven, P., Berger, J., and Hartig, A. (1998). The difference in recognition of terminal tripeptides as peroxisomal targeting signal 1 between yeast and human is due to different affinities of their receptor

Pex5p to the cognate signal and to residues adjacent to it. J. Biol. Chem. *273*, 33635–33643.

Lanyon-Hogg, T., Hooper, J., Gunn, S., Warriner, S.L., and Baker, A. (2014). PEX14 binding to *Arabidopsis* PEX5 has differential effects on PTS1 and PTS2 cargo occupancy of the receptor. FEBS Lett. *588*, 2223–2229.

Lazarow, P.B., and Fujiki, Y. (1985). Biogenesis of peroxisomes. Annu. Rev. Cell Biol. *1*, 489–530.

Leber, R., Landl, K., Zinser, E., Ahorn, H., Spö, A., Kohlwein, S.D., Turnowsky, F., and Nther Daum, G. (1998). Dual Localization of Squalene Epoxidase, Erg1p, in Yeast Reflects a Relationship between the Endoplasmic Reticulum and Lipid Particles. Mol. Biol. Cell *9*, 375–386.

Van der Leij, I., Franse, M.M., Elgersma, Y., Distel, B., and Tabak, H.F. (1993). PAS10 is a tetratricopeptide-repeat protein that is essential for the import of most matrix proteins into peroxisomes of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U. S. A. *90*, 11782–11786.

Lev, S. (2012). Nonvesicular lipid transfer from the endoplasmic reticulum. Cold Spring Harb. Perspect. Biol. *4*, (10) a013300.

Linscott, K.B., Niehaus, T.D., Zhuang, X., Bell, S.A., and Chappell, J. (2016). Mapping a kingdom-specific functional domain of squalene synthase. Biochim. Biophys. Acta - Mol. Cell Biol. Lipids *1861*, 1049–1057.

Liu, H., Tan, X., Veenhuis, M., McCollum, D., and Cregg, J.M. (1992). An efficient screen for peroxisome-deficient mutants of Pichia pastoris. J. Bacteriol. *174*, 4943–4951.

Liu, H., Tan, X., Russell, K.A., Veenhuis, M., and Cregg, J.M. (1995). PER3, a gene required for peroxisome biogenesis in Pichia pastoris, encodes a peroxisomal membrane protein involved in protein import. J. Biol. Chem. *270*, 10940–10951.

Liu, M., Li, L.-N., Pan, Y.-T., and Kong, J.-Q. (2017). cDNA isolation and functional characterization of squalene synthase gene from Ornithogalum caudatum. Protein Expr. Purif. *130*, 63–72.

LoGrasso, P. V, Soltis, D.A., and Boettcher, B.R. (1993). Overexpression, purification, and kinetic characterization of a carboxyl-terminal-truncated yeast squalene synthetase. Arch. Biochem. Biophys. *307*, 193–199.

Longtine, M.S., McKenzie, A., Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. Yeast *14*, 953–961.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J.M., Wolf, Y.I., Yakunin, A.F., et al. (2011). Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. *9*, 467–477.

Managadze, D., Würtz, C., Wiese, S., Schneider, M., Girzalsky, W., Meyer, H.E., Erdmann, R., Warscheid, B., and Rottensteiner, H. (2010). Identification of PEX33, a novel component of the peroxisomal docking complex in the filamentous fungus *Neurospora crassa*. Eur. J. Cell Biol. *89*, 955–964.

Marsh, M. (2001). Endocytosis (Oxford University Press).

Marzioch, M., Erdmann, R., Veenhuis, M., and Kunau, W.H. (1994). PAS7 encodes a novel yeast member of the WD-40 protein family essential for import of 3-oxoacyl-CoA thiolase, a PTS2-containing protein, into peroxisomes. EMBO J. *13*, 4908–4918.

Masaki, R., Yamamoto, A., and Tashiro, Y. (1987). Cytochrome P-450 and NADPH-cytochrome P-450 reductase are degraded in the autolysosomes in rat liver. J. Cell Biol. *104*, 1207–1215.

Massey, V. (1994). Activation of molecular oxygen by flavins and flavoproteins. J. Biol. Chem. *269*, 22459–22462.

Mast, F.D., Li, J., Virk, M.K., Hughes, S.C., Simmonds, A.J., and Rachubinski, R.A. (2011). A *Drosophila* model for the Zellweger spectrum of peroxisome biogenesis disorders. Dis. Model. Mech. *4*, 659–672.

Matsumoto, N., Tamura, S., and Fujiki, Y. (2003). The pathogenic peroxin Pex26p recruits the Pex1p-Pex6p AAA ATPase complexes to peroxisomes. Nat. Cell Biol. *5*, 454–460.

Matsuzaki, T., and Fujiki, Y. (2008). The peroxisomal membrane protein import receptor Pex3p is directly transported to peroxisomes by a novel Pex19p- and Pex16p-dependent pathway. J. Cell Biol. *183*, 1275–1286.

Matsuzono, Y., Kinoshita, N., Tamura, S., Shimozawa, N., Hamasaki, M., Ghaedi, K., Wanders, R.J., Suzuki, Y., Kondo, N., and Fujiki, Y. (1999). Human PEX19: cDNA cloning by functional complementation, mutation analysis in a patient with Zellweger syndrome, and potential role in peroxisomal membrane assembly. Proc. Natl. Acad. Sci. U. S. A. *96*, 2116–2121.

Matsuzono, Y., Matsuzaki, T., and Fujiki, Y. (2006). Functional domain mapping of peroxin Pex19p: interaction with Pex3p is essential for function and translocation. J. Cell Sci. *119*, 3539–3550.

Mattiazzi Ušaj, M., Brložnik, M., Kaferle, P., Žitnik, M., Wolinski, H., Leitner, F., Kohlwein, S.D., Zupan, B., and Petrovič, U. (2015). Genome-Wide Localization Study of Yeast Pex11 Identifies Peroxisome–Mitochondria Interactions through the ERMES Complex. J. Mol. Biol. *427*, 2072–2087.

Mayerhofer, P.U. (2016). Targeting and insertion of peroxisomal membrane proteins: ER trafficking versus direct delivery to peroxisomes. J. Bioch. et Biophy. Acta. *1863*, (5) 870-880.

McKenzie, T.L., Jiang, G., Straubhaar, J.R., Conrad, D.G., and Shechter, I. (1992). Molecular cloning, expression, and characterization of the cDNA for the rat hepatic squalene synthase. J. Biol. Chem. *267*, 21368–21374.

Meijer, W.H., Gidijala, L., Fekken, S., Kiel, J.A.K.W., van den Berg, M.A., Lascaris, R., Bovenberg, R.A.L., and van der Klei, I.J. (2010). Peroxisomes are required for efficient penicillin biosynthesis in *Penicillium chrysogenum*. Appl. Environ. Microbiol. *76*, 5702–5709.

Merkulov, S., van Assema, F., Springer, J., Fernandez Del Carmen, A., and Mooibroek, H. (2000). Cloning and characterization of the *Yarrowia lipolytica* squalene synthase (SQS1) gene and functional complementation of the *Saccharomyces cerevisiae* erg9 mutation. Yeast *16*, 197–206.

Mesquita, A., Tábara, L.C., Martinez-Costa, O., Santos-Rodrigo, N., Vincent, O., and

Escalante, R. (2015). Dissecting the function of Atg1 complex in *Dictyostelium* autophagy reveals a connection with the pentose phosphate pathway enzyme transketolase. Open Biol. *5*, 150088.

Michels, P.A.M., Moyersoen, J., Krazy, H., Galland, N., Herman, M., and Hannaert, V. (2005). Peroxisomes, glyoxysomes and glycosomes (Review). Mol. Membr. Biol. *22*, 133–145.

Miller, W.L. (2017). Disorders in the initial steps of steroid hormone synthesis. J. Steroid Biochem. Mol. Biol. *165*, 18–37.

Motley, A.M., Ward, G.P., and Hettema, E.H. (2008). Dnm1p-dependent peroxisome fission requires Caf4p, Mdv1p and Fis1p. J. Cell Sci. *121*, 1633–1640.

Müllner, H., Zweytick, D., Leber, R., Turnowsky, F., and Daum, G. (2004). Targeting of proteins involved in sterol biosynthesis to lipid particles of the yeast *Saccharomyces cerevisiae*.

Murphy, M.A., Phillipson, B.A., Baker, A., and Mullen, R.T. (2003). Characterization of the targeting signal of the Arabidopsis 22-kD integral peroxisomal membrane protein. Plant Physiol. *133*, 813–828.

Nakano, C., Motegi, A., Sato, T., Onodera, M., and Hoshino, T. (2007). Sterol Biosynthesis by a Prokaryote: First in Vitro Identification of the Genes Encoding Squalene Epoxidase and Lanosterol Synthase from *Methylococcus capsulatus*.

Nakashima, T., Inoue, T., Oka, A., Nishino, T., Osumi, T., and Hata, S. (1995). Cloning, expression, and characterization of cDNAs encoding *Arabidopsis thaliana* squalene synthase. Proc. Natl. Acad. Sci. U. S. A. *92*, 2328–2332.

Nam, K.H., Haitjema, C., Liu, X., Ding, F., Wang, H., Delisa, M.P., and Ke, A. (2012). Cas5d protein processes Pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg crispr-cas system. Structure *20*, 1574–1584.

Nes, W.D. (2011). Biosynthesis of cholesterol and other sterols. Chem. Rev. *111*, 6423–6451.

Nes, W.D., Norton, R.A., Crumley, F.G., Madigan, S.J., and Katz, E.R. (1990). Sterol phylogenesis and algal evolution. Proc. Natl. Acad. Sci. U. S. A. *87*, 7565–7569.

Nuttall, J.M., Hettema, E.H., and Watts, D.J. (2012). Farnesyl diphosphate synthase, the target for nitrogen-containing bisphosphonate drugs, is a peroxisomal enzyme in the model system *Dictyostelium discoideum*. Biochem. J. *447*, 353–361.

Ohtake, K., Saito, N., Shibuya, S., Kobayashi, W., Amano, R., Hirai, T., Sasaki, S., Nakano, C., and Hoshino, T. (2014). Biochemical characterization of the water-soluble squalene synthase from *Methylococcus capsulatus* and the functional analyses of its two DXXD(E)D motifs and the highly conserved aromatic amino acid residues. FEBS J. *281*, 5479–5497.

Ono, T., Ozasa, S., Hasegawa, F., and Imai, Y. (1977). Involvement of NADPH-cytochrome c reductase in the rat liver squalene epoxidase system. Biochim. Biophys. Acta *486*, 401–407.

Ono, T., Takahashi ', K., Odani ', S., Konno2, H., and Imai ', Y. (1980). Purification of squalene epoxidase from rat liver microsomes. Biochim. Biophys research commuincation 522-528.

Ono, T., Nakazono, K., and Kosaka, H. (1982). Purification and partial characterization of squalene epoxidase from rat liver microsomes. Biochim. Biophys. Acta *709*, 84–90.

Otto, G.P., Wu, M.Y., Kazgan, N., Anderson, O.R., and Kessin, R.H. (2003). Macroautophagy is required for multicellular development of the social amoeba *Dictyostelium discoideum*. J. Biol. Chem. *278*, 17636–17645.

Pandit, J., Danley, D.E., Schulte, G.K., Mazzalupo, S., Pauly, T.A., Hayward, C.M., Hamanaka, E.S., Thompson, J.F., and Harwood, H.J. (2000). Crystal Structure of Human Squalene Synthase. J. Biol. Chem. *275*, 30610–30617.

Pang, K.M., Lynes, M.A., and Knecht, D.A. (1999). Variables Controlling the Expression Level of Exogenous Genes in *Dictyostelium*. Plasmid *41*, 187–197.

Parsons, M., Furuya, T., Pal, S., and Kessler, P. (2001). Biogenesis and function of peroxisomes and glycosomes. Mol. Biochem. Parasitol. *115*, 19–28.

Petriv, O.I., Tang, L., Titorenko, V.I., and Rachubinski, R.A. (2004). A New Definition for the Consensus Sequence of the Peroxisome Targeting Signal Type 2. J. Mol. Biol. *341*, 119–134.

Poirier, Y., Antonenkov, V.D., Glumoff, T., and Hiltunen, J.K. (2006). Peroxisomal beta-oxidation--a metabolic pathway with multiple functions. Biochim. Biophys. Acta *1763*, 1413–1426.

Porter, T.D. (2015). Electron Transfer Pathways in Cholesterol Synthesis. Lipids *50*, 927–936.

Purdue, P.E., and Lazarow, P.B. (2001). Peroxisome Biogenesis. Annu. Rev. Cell Dev. Biol. *17*, 701–752.

Purdue, P.E., Yang, X., and Lazarow, P.B. (1998). Pex18p and Pex21p, a novel pair of related peroxins essential for peroxisomal targeting by the PTS2 pathway. J. Cell Biol. *143*, 1859–1869.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173–1183.

Radauer, C., Lackner, P., and Breiteneder, H. (2008). The Bet v 1 fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands. BMC Evol. Biol. *8*, 286-305.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.

Rasbery, J.M., Shan, H., Leclair, R.J., Norman, M., Matsuda, S.P.T., and Bartel, B. (2007). *Arabidopsis thaliana* Squalene Epoxidase 1 Is Essential for Root and Seed Development. J. Biol. Chem. *282*, (23) 17002-17013.

Redman, M., King, A., Watson, C., and King, D. (2016). What is CRISPR/Cas9? Arch. Dis. Child. Educ. Pract. Ed. *101*, 213–215.

Reumann, S., Chowdhary, G., and Lingner, T. (2016). Characterization, prediction and evolution of plant peroxisomal targeting signals type 1 (PTS1s). Biochim. Biophys. Acta - Mol. Cell Res. *1863*, 790–803.

Rhodin, J. (1954). Correlation of ultrastructural organization and function in normal and experimentally changed proximal convoluted tubule cells of the mouse kidney, an electron microscopic study. PhD Thesis. Karolinska Institute, Stockholm.

Richter, H., Zoephel, J., Schermuly, J., Maticzka, D., Backofen, R., and Randau, L. (2012). Characterization of CRISPR RNA processing in *Clostridium thermocellum* and *Methanococcus maripaludis*. Nucleic Acids Res. *40*, 9887–9896.

Robinson, G.W., Tsay, Y.H., Kienzle, B.K., Smith-Monroy, C.A., and Bishop, R.W. (1993). Conservation between human and fungal squalene synthetases: similarities in structure, function, and regulation. Mol. Cell. Biol. *13*, 2706–2717.

Rodrigues, T.A., Alencastre, I.S., Francisco, T., Brites, P., Fransen, M., Grou, C.P., and Azevedo, J.E. (2014). A PEX7-Centered Perspective on the Peroxisomal Targeting Signal Type 2-Mediated Protein Import Pathway. Mol. Cell. Biol. *34*, 2917–2928.

Rogers, M.J., Xiong, X., Ji, X., Mönkkönen, J., Russell, R.G., Williamson, M.P., Ebetino, F.H., and Watts, D.J. (1997). Inhibition of growth of *Dictyostelium discoideum* amoebae by bisphosphonate drugs is dependent on cellular uptake. Pharm. Res. *14*, 625–630.

Rogers, M.J., Crockett, J.C., Coxon, F.P., and Mönkkönen, J. (2011). Biochemical and molecular mechanisms of action of bisphosphonates. Bone *49*, 34–41.

Rong, Q., Jiang, D., Chen, Y., Shen, Y., Yuan, Q., Lin, H., Zha, L., Zhang, Y., and Huang, L. (2016). Molecular Cloning and Functional Analysis of Squalene Synthase 2(SQS2) in *Salvia miltiorrhiza* Bunge. Front. Plant Sci. *7*, 1274-1284.

Rottensteiner, H., Stein, K., Sonnenhol, E., and Erdmann, R. (2003). Conserved function of pex11p and the novel pex25p and pex27p in peroxisome biogenesis. Mol. Biol. Cell *14*, 4316–4328.

Rottensteiner, H., Kramer, A., Lorenzen, S., Stein, K., Landgraf, C., Volkmer-Engert, R., and Erdmann, R. (2004). Peroxisomal membrane proteins contain common Pex19p-binding sites that are an integral part of their targeting signals. Mol. Biol. Cell *15*, 3406–3417.

Rucktäschel, R., Thoms, S., Sidorovitch, V., Halbach, A., Pechlivanis, M., Volkmer, R., Alexandrov, K., Kuhlmann, J., Rottensteiner, H., and Erdmann, R. (2009). Farnesylation of Pex19p Is Required for Its Structural Integrity and Function in Peroxisome Biogenesis. J. Biol. Chem. *284*, 20885–20896.

Ruf, A., Müller, F., D'Arcy, B., Stihle, M., Kusznir, E., Handschin, C., Morand, O.H., and Thoma, R. (2004). The monotopic membrane protein human oxidosqualene cyclase is active as monomer. Biochem. Biophys. Res. Commun. *315*, 247–254.

Russ, W.P., and Engelman, D.M. (2000). The GxxxG motif: A framework for transmembrane helix-helix association. J. Mol. Biol. *296*, 911–919.

Sambrook, M.R.G. and J.P. (2012). Molecular Cloning. A Laboratory Manual, 4th Edition. Cold Spring Harb. Protoc.

Sambrook, J., and Russell, D.W. (2006). SDS-Polyacrylamide Gel Electrophoresis of Proteins. Cold Spring Harb. Protoc. pdb.prot4540.

Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. *32*, 347–355.

Saryi, N.A. Al, Hutchinson, J.D., Al-hejjaj, M.Y., Sedelnikova, S., Baker, P., and Hettema, E.H. (2017). Pnc1 piggy-back import into peroxisomes relies on Gpd1 homodimerisation. Sci. Rep. *7*, 42579-42590.

Schrader, M., and Yoon, Y. (2007). Mitochondria and peroxisomes: are the "big brother" and the "little sister" closer than assumed? Bioessays *29*, 1105–1114.

Schrader, M., Godinho, L.F., Costello, J.L., and Islinger, M. (2015). The different facets of organelle interplay—an overview of organelle interactions. Front. Cell Dev. Biol. *3*, 56-77.

Segura, M.J.R., Jackson, B.E., and Matsuda, S.P.T. (2003). Mutagenesis approaches to deduce structure-function relationships in terpene synthases. Nat. Prod. Rep. *20*, 304–317.

Shibata, N., Arita, M., Misaki, Y., Dohmae, N., Takio, K., Ono, T., Inoue, K., and Arai, H. (2001). Supernatant protein factor, which stimulates the conversion of squalene to lanosterol, is a cytosolic squalene transfer protein and enhances cholesterol biosynthesis. J. Proc. Natl. Acad. Sci.*98*, (5) 2243-2249.

Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., Abudayyeh, O.O., Gootenberg, J.S., Makarova, K.S., Wolf, Y.I., et al. (2017). Diversity and evolution of class 2 CRISPR–Cas systems. Nat. Publ. Gr. *15*, 169-182

Simons, K., and Vaz, W.L.C. (2004). Model Systems, Lipid Rafts, and Cell Membranes. Annu. Rev. Biophys. Biomol. Struct. *33*, 269–295.

Sinkunas, T., Gasiunas, G., Waghmare, S.P., Dickman, M.J., Barrangou, R., Horvath, P., and Siksnys, V. (2013). In vitro reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. EMBO J. *32*, 385–394.

Smith, J.J., Marelli, M., Christmas, R.H., Vizeacoumar, F.J., Dilworth, D.J., Ideker, T., Galitski, T., Dimitrov, K., Rachubinski, R.A., and Aitchison, J.D. (2002). Transcriptome profiling to identify genes involved in peroxisome assembly and function. J. Cell Biol. *158*, 259–271.

Soccio, R.E., and Breslow, J.L. (2004). Intracellular Cholesterol Transport. Arterioscler. Thromb. Vasc. Biol. *24*, 1150-1160.

Song, F., and Stieger, K. (2017). Optimizing the DNA Donor Template for Homology-Directed Repair of Double-Strand Breaks. Mol. Ther. Nucleic Acids *7*, 53–60.

Spong, A.P., and Subramani, S. (1993). Cloning and characterization of PAS5: a gene required for peroxisome biogenesis in the methylotrophic yeast Pichia pastoris. J. Cell Biol. *123*, 535–548.

Stamellos, K.D., Shackelford, J.E., Shechter, I., Jiang, G., Conrad, D., Keller, G.A., and Krisans, S.K. (1993). Subcellular localization of squalene synthase in rat hepatic cells. Biochemical and immunochemical evidence. J. Biol. Chem. *268*, 12825–12836.

Steinberg, S.J., Dodt, G., Raymond, G. V., Braverman, N.E., Moser, A.B., and Moser, H.W. (2006). Peroxisome biogenesis disorders. Biochim. Biophys. Acta - Mol. Cell Res. *1763*, 1733–1748.

Su, T., Liu, F., Gu, P., Jin, H., Chang, Y., Wang, Q., Liang, Q., and Qi, Q. (2016). A CRISPR-Cas9 Assisted Non-Homologous End-Joining Strategy for One-step Engineering of Bacterial Genome. Sci. Rep. *6*, 37895-37906.

Tai, H.H., and Bloch, K. (1972). Squalene epoxidase of rat liver. J. Biol. Chem. *247*, 3767–3773.

Takahashi, N., Morita, M., Maeda, T., Harayama, Y., Shimozawa, N., Suzuki, Y., Furuya, H., Sato, R., Kashiwayama, Y., and Imanaka, T. (2007). Adrenoleukodystrophy:

subcellular localization and degradation of adrenoleukodystrophy protein (ALDP/ABCD1) with naturally occurring missense mutations. J. Neurochem. *101*, 1632–1643.

Takatsuji, H., Nishino, T., Izui, K., and Katsuki2, H. (1982). Formation of Dehydrosqualene Catalyzed by Squalene Synthetase in *Saccharomyces cerevisiae*1. J. Biochem *91*, 911–921.

Tam, Y.Y.C., and Rachubinski, R.A. (2002). Yarrowia lipolytica cells mutant for the PEX24 gene encoding a peroxisomal membrane peroxin mislocalize peroxisomal proteins and accumulate membrane structures containing both peroxisomal matrix and membrane proteins. Mol. Biol. Cell *13*, 2681–2691.

Tam, Y.Y.C., Torres-Guzman, J.C., Vizeacoumar, F.J., Smith, J.J., Marelli, M., Aitchison, J.D., and Rachubinski, R.A. (2003). Pex11-related proteins in peroxisome dynamics: a role for the novel peroxin Pex27p in controlling peroxisome size and number in *Saccharomyces cerevisiae*. Mol. Biol. Cell *14*, 4089–4102.

Tam, Y.Y.C., Fagarasanu, A., Fagarasanu, M., and Rachubinski, R.A. (2005). Pex3p initiates the formation of a preperoxisomal compartment from a subdomain of the endoplasmic reticulum in *Saccharomyces cerevisiae*. J. Biol. Chem. *280*, 34933–34939.

Tansey, T.R., and Shechter, I. (2000). Structure and regulation of mammalian squalene synthase. Biochim. Biophys. Acta *1529*, 49–62.

Tchen, T.T., and Bloch, K. (1957). On the mechanism of enzymatic cyclization of squalene. J. Biol. Chem. *226*, 931–939.

Teese, M.G., and Langosch, D. (2015). Role of GxxxG Motifs in Transmembrane Domain Interactions. Biochemistry *54*, 5125–5135.

Thoma, R., Schulz-Gasch, T., D'Arcy, B., Benz, J., Aebi, J., Dehmlow, H., Hennig, M., Stihle, M., and Ruf, A. (2004). Insight into steroid scaffold formation from the structure of human oxidosqualene cyclase. Nature *432*, 118–122.

Thorsell, A.-G., Lee, W.H., Persson, C., Siponen, M.I., Nilsson, M., Busam, R.D., Kotenyova, T., Schüler, H., and Lehtiö, L. (2011). Comparative Structural Analysis of Lipid Binding START Domains. PLoS One *6*, (6) e19521-19532.

Titorenko, V.I., and Rachubinski, R.A. (1998). Mutants of the yeast *Yarrowia lipolytica* defective in protein exit from the endoplasmic reticulum are also defective in peroxisome biogenesis. Mol. Cell. Biol. *18*, 2789–2803.

Titorenko, V.I., Ogrydziak, D.M., and Rachubinski, R.A. (1997). Four distinct secretory pathways serve protein secretion, cell surface growth, and peroxisome biogenesis in the yeast Yarrowia lipolytica. Mol. Cell. Biol. *17*, 5210–5226.

Titorenko, V.I., Smith, J.J., Szilard, R.K., and Rachubinski, R.A. (1998). Pex20p of the yeast Yarrowia lipolytica is required for the oligomerization of thiolase in the cytosol and for its targeting to the peroxisome. J. Cell Biol. *142*, 403–420.

Tower, R.J., Fagarasanu, A., Aitchison, J.D., and Rachubinski, R.A. (2011). The peroxin Pex34p functions with the Pex11 family of peroxisomal divisional proteins to regulate the peroxisome population in yeast. Mol. Biol. Cell *22*, 1727–1738.

Tsai, I.J., Zarowiecki, M., Holroyd, N., Garciarrubio, A., Sánchez-Flores, A., Brooks, K.L., Tracey, A., Bobes, R.J., Fragoso, G., Sciutto, E., et al. (2013). The genomes of four tapeworm species reveal adaptations to parasitism. Nature *496*, 57–63.

Tsukamoto, T., Yokota, S., and Fujiki, Y. (1990). Isolation and characterization of Chinese hamster ovary cell mutants defective in assembly of peroxisomes. J. Cell Biol. *110*, 651–660.

Tsukamoto, T., Miura, S., and Fujiki, Y. (1991). Restoration by a 35K membrane protein of peroxisome assembly in a peroxisome-deficient mammalian cell mutant. Nature *350*, 77–81.

Uchida, H., Sugiyama, R., Nakayachi, O., Takemura, M., and Ohyama, K. (2007). Expression of the gene for sterol-biosynthesis enzyme squalene epoxidase in parenchyma cells of the oil plant, Euphorbia tirucalli. Planta *226*, 1109–1115.

Udell, J. a, Fischer, M. a, Brookhart, M.A., Solomon, D.H., and Choudhry, N.K. (2006). Effect of the Women's Health Initiative on osteoporosis therapy and expenditure in Medicaid. J. Bone Miner. Res. *21*, 765–771.

Usher, C., Teeling, M., Bennett, K., and Feely, J. (2006). Effect of clinical trial publicity on HRT prescribing in Ireland. Eur. J. Clin. Pharmacol. *62*, 307–310.

Vance, J.E. (2012). Dysregulation of cholesterol balance in the brain: contribution to neurodegenerative diseases. Dis. Model. Mech. *5*, 746–755.

Vizeacoumar, F.J., Torres-Guzman, J.C., Tam, Y.Y.C., Aitchison, J.D., and Rachubinski, R.A. (2003). YHR150w and YDR479c encode peroxisomal integral membrane proteins involved in the regulation of peroxisome number, size, and distribution in *Saccharomyces cerevisiae*. J. Cell Biol. *161*, 321–332.

Vizeacoumar, F.J., Torres-Guzman, J.C., Bouard, D., Aitchison, J.D., and Rachubinski, R.A. (2004). Pex30p, Pex31p, and Pex32p form a family of peroxisomal integral membrane proteins regulating peroxisome size and number in *Saccharomyces cerevisiae*. Mol. Biol. Cell *15*, 665–677.

Vyas, V.K., Barrasa, M.I., and Fink, G.R. (2015). A Candida albicans CRISPR system permits genetic engineering of essential genes and gene families. Sci. Adv. *1*, e1500248-1500254.

Wanders, R.J.A. (2014). Metabolic functions of peroxisomes in health and disease. Biochimie *98*, 36–44.

Wanders, R., and Waterham, H. (2004). Peroxisomal disorders I: biochemistry and genetics of peroxisome biogenesis disorders. Clin. Genet. *67*, 107–133.

Wanders, R.J.A., and Waterham, H.R. (2006). Biochemistry of Mammalian Peroxisomes Revisited. Annu. Rev. Biochem. *75*, 295–332.

Waterham, H.R., and Ebberink, M.S. (2012). Genetics and molecular basis of human peroxisome biogenesis disorders. Biochim. Biophys. Acta - Mol. Basis Dis. *1822*, 1430–1441.

Waterham, H.R., Ferdinandusse, S., and Wanders, R.J.A. (2016). Human disorders of peroxisome metabolism and biogenesis. Biochim. Biophys. Acta - Mol. Cell Res. *1863*, 922–933.

Watson, J., Wise, L., and Green, J. (2007). Prescribing of hormone therapy for menopause, tibolone, and bisphosphonates in women in the UK between 1991 and 2005. Eur. J. Clin. Pharmacol. *63*, 843–849.

Whitley, P., Grahn, E., Kutay, U., Rapoport, T.A., and von Heijne, G. (1996). A 12-

residue-long polyleucine tail is sufficient to anchor synaptobrevin to the endoplasmic reticulum membrane. J. Biol. Chem. *271*, 7583–7586.

Wiebel, F.F., and Kunau, W.-H. (1992). The Pas2 protein essential for peroxisome biogenesis is related to ubiquitin-conjugating enzymes. Nature *359*, 73–76.

Woo, J., Kim, Y., and Lee, C. (2012). Heterogeneous genetic associations of nucleotide sequence variants with bone mineral density by gender. Mol. Biol. Rep. *39*, 2259–2265.

Yamamoto, S., and Bloch, K. (1970). Studies on squalene epoxidase of rat liver. J. Biol. Chem. *245*, 1670–1674.

Yang, L., Mali, P., Kim-Kiselak, C., and Church, G. (2014). CRISPR-Cas-Mediated Targeted Genome Editing in Human Cells. In Methods in Molecular Biology (Clifton, N.J.), pp. 245–267.

Yang, M., Ellenberg, J., Bonifacino, J.S., and Weissman, A.M. (1997). The transmembrane domain of a carboxyl-terminal anchored protein determines localization to the endoplasmic reticulum. J. Biol. Chem. *272*, 1970–1975.

Yosef, I., Goren, M.G., and Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. Nucleic Acids Res. *40*, 5569–5576.

Yuan, W., Veenhuis, M., and van der Klei, I.J. (2016). The birth of yeast peroxisomes. Biochim. Biophys. Acta - Mol. Cell Res. *1863*, 902–910.

Žárský, V., and Tachezy, J. (2015). Evolutionary loss of peroxisomes--not limited to parasites. Biol. Direct *10*, (1) 74-84.

Zhang, J., Rouillon, C., Kerou, M., Reeks, J., Brugger, K., Reimann, J., Cannone, G., Liu, H., Albers, S., Naismith, H., et al. (2012). Europe PMC Funders Group Structure and mechanism of the CMR complex for CRISPR- mediated antiviral immunity. Mol Cell *45*, 303–313.

Zhao, R.-Y., Xiao, W., Cheng, H.-L., Zhu, P., and Cheng, K.-D. (2010). Cloning and characterization of squalene synthase gene from *Fusarium fujikuroi* (Saw.) Wr. J. Ind. Microbiol. Biotechnol. *37*, 1171–1182.