



# University of HUDDERSFIELD

## University of Huddersfield Repository

Dodds, Kristian

Compartmentalisation of Mevalonate Biosynthesis in Evolutionarily Divergent Eukaryotes

### Original Citation

Dodds, Kristian (2020) Compartmentalisation of Mevalonate Biosynthesis in Evolutionarily Divergent Eukaryotes. Masters thesis, University of Huddersfield.

This version is available at <http://eprints.hud.ac.uk/id/eprint/35440/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: [E.mailbox@hud.ac.uk](mailto:E.mailbox@hud.ac.uk).

<http://eprints.hud.ac.uk/>

Compartmentalisation of Mevalonate  
Biosynthesis in Evolutionarily Divergent  
Eukaryotes

**Kristian Dodd**

Supervised by **Professor Michael Ginger** and **Dr Jane Harmer**

MSc by Research- Biological Sciences

University of Huddersfield

September 2018-January 2020

A thesis submitted to the University of Huddersfield in partial  
fulfilment of the requirements for the degree of Master of Science by Research

## Abstract

Mevalonate is a committed precursor of sterols and other isoprenoids in mammals, yeast, plants and many other eukaryotes. A key enzyme in the mevalonate pathway of isoprenoid biosynthesis is HMG-CoA reductase (HMGR), which represents a key regulatory step in mammals and potentially other organisms, responsible for the reduction of HMG-CoA (3-hydroxy-3-methylglutaryl-CoA). The classical localisation of this enzyme is within the endoplasmic reticulum (ER) as an integral membrane protein with a cytosolic facing enzymatic domain. However, the trypanosomatid HMGR does not fit this description. It is instead localised within the mitochondrion. There is also indication from genome analyses that in a variety of other microbial eukaryotes HMGR may be present as soluble, rather than integral membrane protein, but where within cells these in silico-identified HMGRs localise is not known. The mitochondrial localisation of HMGR in trypanosomatids may be an adaption to allow the amino acid leucine to be utilised as a major carbon source for sterol biosynthesis. Leucine breakdown occurs within mitochondria in eukaryotes and results in HMG-CoA as an intermediate however, due to the organellar separation of these processes in most eukaryotes leucine cannot be utilised as in trypanosomatids.

In order to investigate a possible adaption to the mevalonate pathway in different divergent eukaryotes, two organisms and two genes were selected. Squalene epoxidase (SqE), an enzyme near the end of the pre-squalene section of the mevalonate pathway was investigated in the trypanosomatid *Leishmania tarentolae*. The localisation of this enzyme was not clear from past studies and literature comparisons, being possibly peroxisomal, ER or lipid-body localised. Homologous recombination via a two-step PCR originated tagging amplicon was utilised, resulting in an N-terminal green fluorescent protein (GFP) tag for the SqE gene in *L. tarentolae* cells. These cells were then analysed using confocal microscopy, and although the homologous recombination appeared to be successful, the imaging data was too unclear as to provide accurate localisation information.

*Dictyostelium discoideum* was the major organism of focus for this study. This is an early diverging eukaryote, realistically sharing its last common ancestor with the trypanosomes at the very root of all eukaryotic lineages. The genes of focus in this organism encoded the HMGR isoforms, *DdHMGA* and *DdHMGB*. Myc-tagged synthetic genes for both isoforms were cloned into *Dictyostelium*-specific expression vector *pDM1039* and transfected into *D*. Transformed cells, including those also expressing Red Fluorescent Protein (RFP) targeted to peroxisomes, were analysed by immunoblotting (to size expressed myc-tagged HMGR proteins) and immunofluorescence confocal microscopy. *DdHMGA* was shown to possess a possible peroxisomal localisation while *DdHMGB* showed a punctate possible cytosolic localisation.

Alongside the laboratory studies, computational studies were carried out in order to investigate the presence or lack of various mevalonate pathway enzymes and their localisation within a wide range of eukaryotes. In this way I aimed to provide new insight into the organisation and evolution of the important mevalonate pathway in eukaryotes.

# Table of Contents

<b>Abstract</b>	<b>2</b>
<b>Abbreviations</b>	<b>5</b>
<b>Thesis Introduction Draft</b>	<b>8</b>
General Introduction	8
Isoprenoid Biosynthesis	8
Peroxisomes and Peroxisomal Targeting	15
Aims and Objectives	21
<b>Materials and Methods</b>	<b>22</b>
<i>Dictyostelium discoideum</i> Culturing	22
Growing <i>Dictyostelium discoideum</i>	22
Freezing <i>Dictyostelium discoideum</i>	23
Thawing <i>Dictyostelium discoideum</i> Cells	23
<i>Dictyostelium discoideum</i> DNA Manipulation	23
Electroporation of <i>Dictyostelium discoideum</i>	23
Preparation of <i>pDM 1039</i> Expression Plasmid	24
Sub-cloning <i>DdHMGR</i> isoforms into <i>pDM 1039</i> plasmids	25
Fusion PCR-based GFP tagging of <i>L. tarentolae</i> SqE	26
PCR to Create a Tagging Amplicon for <i>Leishmania tarentolae</i>	27
Amplification of pPlot and Homology Flanks for <i>LtSqE</i>	27
PCR to Create SqE Homology Flanks	27
First Attempt at Fusion PCR protocol	28
Final attempt at creating the Fusion tagging amplicon	29
<i>Leishmania</i> genomic DNA purification	29
New round of PCR amplifications of Fusion amplicon components	30
Fusion cassette PCR using squalene epoxidase ORF and UTR sequences	30
MitoTracker™ Deep Red FM and ER-Tracker™ Red for Visualisation of Mitochondria and the ER in <i>D. discoideum</i>	32
Preparation of slides and immunofluorescence	32
Immunoblotting	33
<b>Results</b>	<b>34</b>
Preparation of <i>D. discoideum</i> Plasmids Encoding HMGR Isoforms	34
Immunoblot analysis of HMGA::myc and HMGB::myc expression in <i>Dictyostelium discoideum</i>	35
Confocal microscopy analysis of DdHMGA::myc and DdHMGB::myc localisation	36
<i>Leishmania tarentolae</i> cells transformed with a GFP-tagged squalene epoxidase enzyme	41
Phylogenomic distribution of the Early Enzymes of the Mevalonate Pathway in Microbial Eukaryotes	42
Determination of Mitochondrial Localisation of HMG-CoA Reductase Enzymes Using Prediction Software	45
Localisation of the Enzymes of the MVA Pathway of Isoprenoid Biosynthesis	47

<b>Discussion</b>	<b>49</b>
Aims and Objectives	49
Preparation of the Dictyostelium discoideum plasmids	49
Localisation of Two Different Isoforms of HMGR in Dictyostelium discoideum Based Upon Fluorescence Microscopy	50
Leishmania tarentolae cells transformed with a GFP-tagged squalene epoxidase enzyme	51
The Importance of Early-Branching Eukaryotes in the Whole Eukaryotic Phylogenetic Tree	53
Additional Work and Comments	56

## Abbreviations

ATP	Adenosine triphosphate
Bp	Base pairs
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
<i>C. merolae</i>	<i>Cyanidioschyzon merolae</i>
DAPI	4',6-diamidino-2-phenylindole
Dci1p	$\Delta$ (3,5)- $\Delta$ (2,4)-dienoyl-CoA isomerase
<i>Dd</i>	<i>Dictyostelium discoideum</i>
<i>D. discoideum</i>	<i>Dictyostelium discoideum</i>
DLP	Dynamain-like protein
DMAPP	Dimethylallyl pyrophosphate
DMSO	Dimethyl sulphoxide
dNTPs	Equimolar mixture of dTTP, dATP, dCTP, dGTP
DOXP	1-deoxy-D-xylulose 5-phosphate
DPMP	Diphosphomevalonate decarboxylase
DRP	Dynamain related protein
Eci1p	$\Delta$ 3, $\Delta$ 2-enoyl-CoA isomerase
<i>E. coli</i>	<i>Escherichia coli</i>
<i>E. histolytica</i>	<i>Entamoeba histolytica</i>
ER	Endoplasmic reticulum
FITC	Fluorescein isothiocyanate
GAP	d-glyceraldehyde 3-phosphate
gDNA	Genomic DNA
GFP	Green fluorescent protein
<i>G. sulphuraria</i>	<i>Galdieria sulphuraria</i>
HMG	3-hydroxy-3-methylglutaryl
HMGA	<i>Dictyostelium discoideum</i> HMG-CoA reductase 2
HMGB	<i>Dictyostelium discoideum</i> HMG-CoA reductase 1
HMG-CoA	3-hydroxy-3-methylglutaryl-CoA
HMGL	HMG-CoA lyase

HMGR	HMG-CoA reductase
HMGS	HMG-CoA synthase
IPP	Isopentyl diphosphate
IVDH	Isovaleryl dehydrogenase
<i>L. major</i>	<i>Leishmania major</i>
<i>Lt</i>	<i>Leishmania tarentolae</i>
<i>L. tarentolae</i>	<i>Leishmania tarentolae</i>
MEP	methylerythritol 4-phosphate
MK	Mevalonate kinase
mTP	Mitochondrial transport peptides
MVA	Mevalonate
ORF	Open reading frame
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PMP	Peroxisome membrane protein
Pox1p	Acyl-CoA oxidase
PTS	Peroxisomal targeting signal
qPCR	Quantitative polymerase chain reaction
RFP	Red fluorescent protein
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SqE	Squalene epoxidase
<i>T. brucei</i>	<i>Trypanosoma brucei</i>
<i>T. cruzi</i>	<i>Trypanosoma cruzi</i>
TEMED	Tetramethylethylenediamine
TPR	Tetratricopeptide repeats
UTR	Untranslated region

## Acknowledgements

Firstly, I would like to express gratitude to my supervisors Professor Michael Ginger and Dr Jane Harmer for their supervision and guidance throughout my Masters by Research.

Further thanks to the research team at the University of Huddersfield, and to Asma Albelazi in particular for her assistance and teaching.

Finally, special thanks to Dr Jason King from the University of Sheffield for the generous donation of *Dictyostelium discoideum* and plasmids used in this study.

## Introduction

### General Introduction

Isoprenoids and isoprenoid-derived compounds are vital end-products of biosynthesis and often intriguing secondary metabolites in prokaryotes and eukaryotes. The acquisition or production of these compounds is beginning to appear as varied as the myriad of functions these molecules occupy. Of significant interest is the evolution of the pathways of isoprenoid synthesis and subsequent modifications in divergent eukaryotes such as *Leishmania tarentolae* and *Dictyostelium discoideum*, the focus organisms of my study. Study of one of the key enzymes; HMGR, and its various forms and localisations in currently known eukaryotes may provide insight into how the compartmentalisation of this pathway evolved in eukaryotes

### Isoprenoid Biosynthesis

#### Isoprenoids: An Introduction

Isoprenoids are also known as terpenoids and encompass an incredibly diverse group of compounds with a multitude of functions within all known families of life.

One of the most common isoprenoid-derived molecules found within eukaryotes are sterols. Sterols are found in a wide range of forms in eukaryotes. Cholesterol is the major sterol in mammals and many other eukaryotes, with numerous functions. For instance, cholesterol acts as a membrane reinforcement against shear stress in animal cell membrane (Ribeiro et al., 2007), and reduces Na<sup>+</sup> and K<sup>+</sup> ion leakage into the cytosol from sub-cellular compartments (Haines, 2001). Lipid rafts, primarily composed of cholesterol and sphingolipids act as anchor-points for proteins within the bilayer structure of cell membranes (Simons and Ikonen, 1997). These lipid rafts and their associated proteins are responsible for initiation of signal transduction pathways involved in cell adhesion, migration, and cell survival and proliferation (Mollinedo and Gajate, 2015). Cholesterol also acts as a precursor for the production of steroids (Holst et al., 2004), which act as signalling hormones in most animals, with insects possessing a unique class of steroid called ecdysteroids (Lafront and Mathieu, 2007). Plants possess a similar class of cholesterol-derived hormones called brassinosteroids (Clouse, 2011). These hormones have similar physiological functions in plants to animal steroids, involved in lipid rafts, mediating development, growth and stress responses (Valitova et al., 2016). Interestingly

unlike vertebrates, arthropods and nematode worms are unable to synthesise cholesterol and must acquire all required sterols from nutrients (Kurzchalia and Ward, 2003).

Some bacteria also produce a class of isoprenoid, sterol-related molecules called hopanoids (Pearson et al., 2007). These molecules act in much the same way as eukaryote sterols and occupy similar roles and functions (Mangiarotti et al., 2019). In eukaryotes, sterols are formed through oxygen-dependent cyclisation of the aliphatic hydrocarbon squalene, whereas hopanoids are formed through oxygen-independent squalene cyclisation (Belin et al., 2018).

Isoprenoid quinones act as electron and proton carriers on both the respiratory and photosynthetic electron transport chains (Kruk and Nowicka, 2010). Dolichol, an isoprenoid lipid is utilised in protein glycosylation in eukaryotes, bacteria and archaea (Aebi, 2013). All isoprenoid molecules are constituted via the same five-carbon intermediate, isopentenyl pyrophosphate (IPP) (Clomburg et al., 2019).

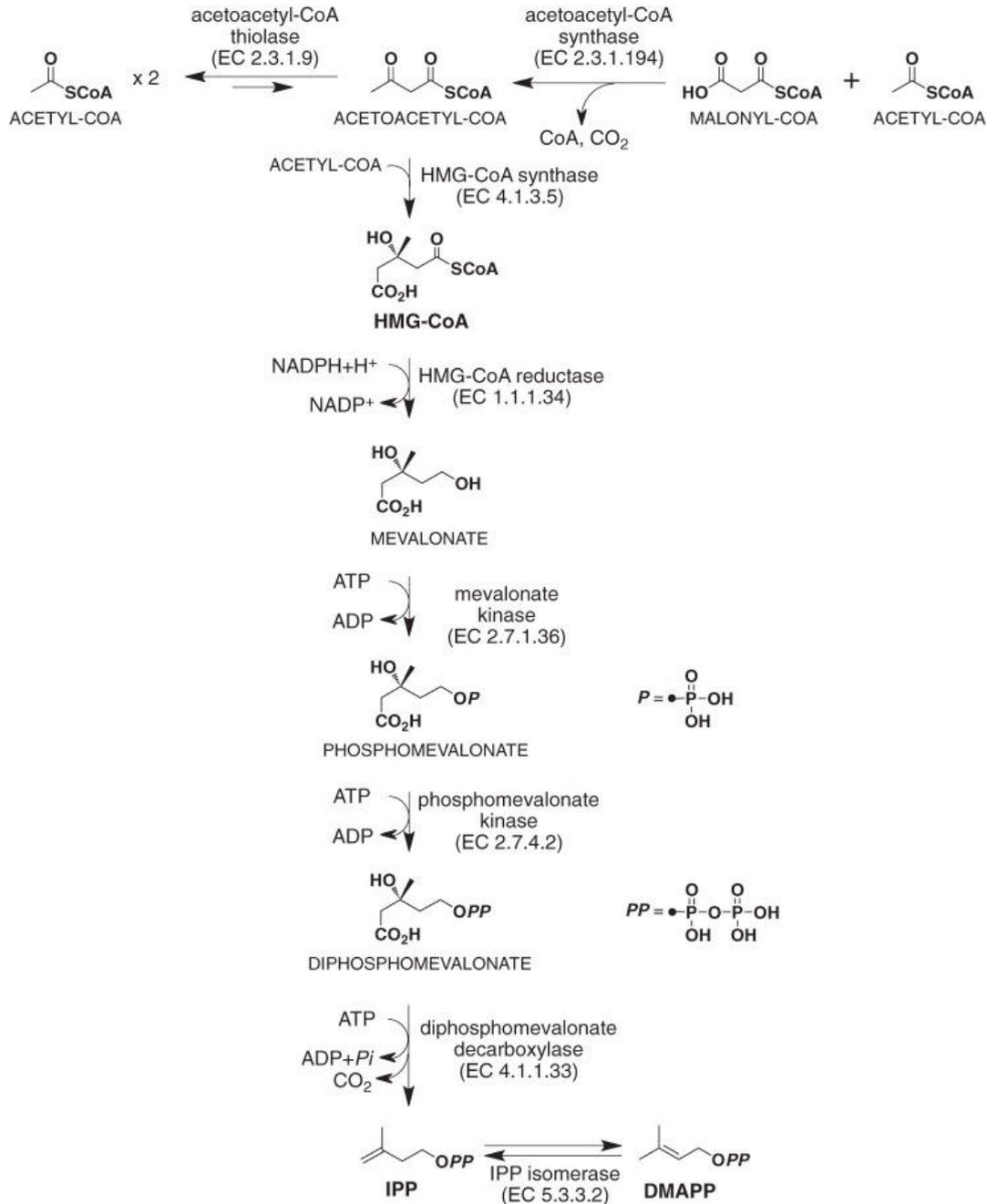
There are two major enzyme pathways which are responsible for the production of IPP, the mevalonate (MVA) pathway and a mevalonate-independent pathway, also called the MEP pathway.

#### The Different Pathways of Isoprenoid Biosynthesis

There are two major enzyme pathways which are responsible for the production of IPP; the mevalonate (MVA) pathway and a mevalonate-independent pathway, also called the MEP pathway. The MVA pathway was the first identified biosynthetic route resulting in the production of the isoprenoid family of molecules. The MVA pathway for biosynthesis of cholesterol was first discovered in the 1950s and 1960s resulting in award of Nobel Prize in Physiology or Medicine to Konrad Bloch and Feodor Lynen for its discovery (Bloch and Katsuki, 1967) (Lynen, 1967).

## The MVA Pathway

The MVA pathway is the principle pathway of isoprenoid biosynthesis within eukaryotes and archaea and is also present in some bacteria (Hoshino and Gaucher, 2018).



**Figure 1. MVA pathway of IPP/DMAPP biosynthesis (Kuzuyama and Seto, 2012).**

Through the action of HMG-CoA synthase, three molecules of acetyl-CoA are using in the synthesis of HMG-CoA. HMG-CoA is converted into the key precursor of the MVA pathway; mevalonate. Mevalonate then goes through two phosphorylation steps and a decarboxylation step before resulting in IPP. IPP is altered by IPP isomerase to become the IPP-isomer DMAPP. DMAPP is the product of the pathway and a precursor to isoprenoid synthesis.

Within the MVA pathway, acetyl-CoA is converted into 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) by the enzymes acetyl-CoA thiolase and HMG-CoA synthase (HMGS). The product of the HMGR-catalysed reaction is mevalonic acid, which is then converted into mevalonate-5-diphosphate and finally into IPP (Miziorko, 2011).

There are three known variants to this route via mevalonate to IPP. The classical pathway is highly conserved in eukaryotes and involves the phosphorylation of mevalonate 5-phosphate via the enzyme mevalonate kinase (MK) before conversion into isopentenyl diphosphate (IPP) by the enzyme diphosphomevalonate decarboxylase (DPMD). This pathway is also utilised by the archaeon *Sulfolobus solfataricus* (Nishimura et al., 2013). Other archaea utilise one of two variations on the classical MVA pathway.

One variation, discovered in the halophile archaea *Haloferax volcanii* is the conversion of mevalonate 5-phosphate into isopentenyl phosphate which is then phosphorylated into IPP (VanNice et al., 2014). This variation, involving isopentenyl phosphate as an intermediate, is also present in some species within the bacterial genus *Chloroflexi* (Dellas et al., 2013).

The other variation of the MVA pathway in archaea was discovered in the thermophilic acidophile *Thermoplasma acidophilum*. This variation involves bypassing the direct conversion of mevalonate into mevalonate 5-phosphate, in favour of a stepwise phosphorylation to mevalonate 3-phosphate by mevalonate 3-kinase followed by a second phosphorylation by mevalonate 3-phosphate 5-kinase. The mevalonate 3,5-bisphosphate is then converted into isopentenyl phosphate and then into IPP in the same process as the MVA pathway variant present in *H. volcanii* (Vinokur et al., 2014).

### HMG-CoA Reductase

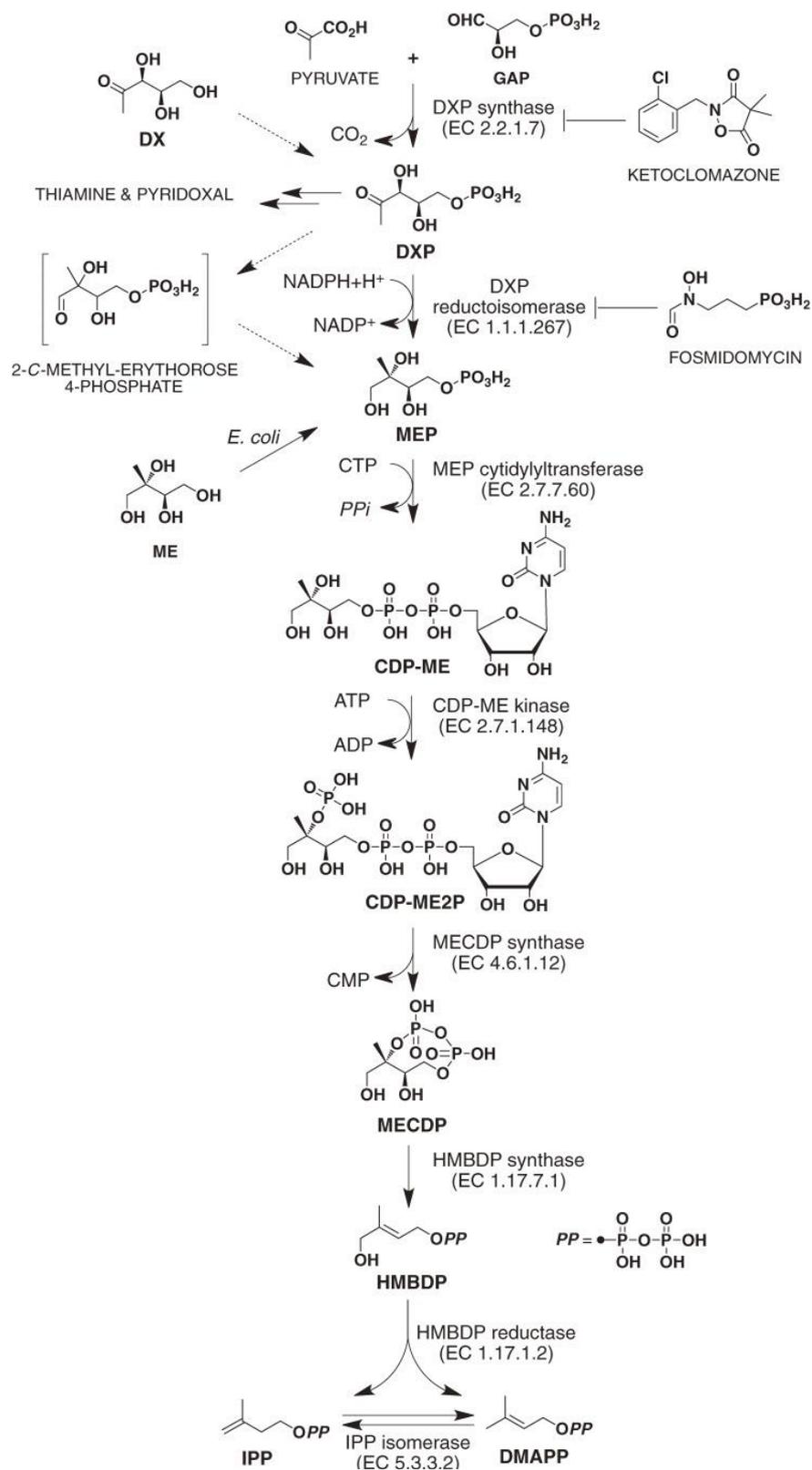
HMGR is a key enzyme in the MVA pathway and catalyses the rate limiting step in mammals: the reduction of HMG-CoA into mevalonate (Friesen and Rodwell, 2004). The textbook eukaryotic form of the enzyme HMGR possesses multiple transmembrane-spanning regions (Xu and Simoni, 2003). These regions target the enzyme for localisation in the ER membrane with the active site facing the cytosol. Statins, now famous and widely used for their application in reduction cholesterol in patients, are a class of drugs that all possess an inhibitory effect on HMGR. This inhibition works by way of a competitive affinity for the HMG-CoA binding region of the enzyme (Istvan, 2003). HMGR broadly exists in two different forms, class 1 enzymes and class 2 enzymes (Gophna et al., 2006). The HMG-CoA binding site is present in both of these classes; however, the sequences show significant variance between the classes despite little difference between orthologues within the same class

(Bochar et al., 1999). Class 1 HMGRs are the enzymes found in most eukaryotes and archaea possessing varying numbers of transmembrane regions in eukaryotes whereas Class 2 enzymes are soluble and present in the relatively few bacterial taxa that possess an MVA pathway (Istvan et al., 2000). However, this is not a rule and exceptions exist such as *Trichomonas vaginalis* which is a eukaryote but, as a consequence of presumably lateral gene transfer possesses a class 2 soluble HMGR enzyme (Ginger et al., 2010).

#### The MEP Pathway: A Non-Mevalonate Pathway

The other major pathway through which isoprenoids are produced, and identified in the 1990s, is to present in many bacteria and virtually all phototrophs. This pathway is widely known as the MEP pathway (named for the MEP molecule used as a precursor in the production of IPP) but has also previously been known as the DOXP pathway. This pathway represents a mevalonate-independent route of IPP production, utilising different enzymes and intermediate compounds (Rodriguez-Concepcion, 2004). Cyanobacteria utilise the MEP pathway to produce isoprenoids used in the production of a wide variety of molecules, including the bacteria-exclusive membrane components, hopanoids (Pattanaik and Lindberg et al., 2015).

The enzymes of this pathway are localised in the plastids of eukaryotic phototrophic organisms (Ahn and Pai, 2008) and the pathway is also utilised by many bacteria as the sole method of isoprenoid biosynthesis (Hoshino and Gaucher, 2018). The presence of this second pathway in eukaryotes is explained by the endosymbiotic origin of chloroplasts whereby an ancestral heterotrophic eukaryotic host organism acquired and then entered into a long-lasting symbiotic relationship with a cyanobacterium (Lange et al., 2000).



**Figure 2. MEP pathway of IPP/DMAPP biosynthesis (Kuzuyama and Seto, 2012).**

The MEP pathway represents an alternate route to the production of DMAPP. It shares no precursors with the MVA pathway. This pathway involves the initial production of DXP from pyruvate and d-glyceraldehyde 3-phosphate (GAP), and its subsequent conversion into MEP. MEP then goes through multiple enzyme-mediated steps before final conversion into IPP. As with the MVA pathway, IPP is isomerised into DMAPP.

Apicomplexa are a phylum of unicellular eukaryotes in which all are obligate intracellular parasites. All known taxa, with the exception of *Cryptosporidium*, possess a relic non-photosynthetic chloroplast (or plastid) called an apicoplast. This plastid was likely obtained via a secondary endosymbiosis event involving a red alga (Striepen, 2011). The photosynthetic chloroplast that was obtained through this event lost the pathways and processes required for the photosynthesis pathway as the apicomplexan ancestors adopted pathogenic lives (van Dooren and Striepen, 2013). The apicoplast now confers essential advantage towards the pathogenicity of these organisms through the presence of essential apicoplast-localised, plastid specific metabolic pathways. Thus, despite the loss of the photosynthetic pathway in the apicoplasts, the MEP pathway enzymes utilised by the endosymbiont were conserved (He et al, 2018) and is indeed the sole essential apicoplast pathway required by the pathogenic asexual bloodstream stages of the malaria parasite *Plasmodium falciparum*. Apicomplexa therefore produce isoprenoids solely using the MEP pathway, localised within the apicoplast (Kadian et al., 2018). However, some apicomplexans also scavenge isoprenoids from a host that utilises the MVA-pathway (Imlay and Odom, 2014).

#### Exceptions to the 'Typical' Possession of MVA or MEP Pathways

Although in most photosynthetic organisms both the MVA and MEP pathways are present, there are some exceptions. Green algae (Chlorophyta) have been shown to lack the MVA pathway altogether, producing sterols, as well as plastidic isoprenoids, solely using the MEP pathway (Schwender et al., 2001). This adaptation however also required changes in cellular transport in order to move IPP from the chloroplasts into the cytosol for further reactions.

The Apicomplexan *Cryptosporidium* genus, unlike other Apicomplexa, lack the unique apicoplast organelle and thus the MEP pathway machinery (Zhu et al., 2000). As mentioned earlier, all Apicomplexan parasites possess a mechanism for obtaining isoprenoid precursors and products from the host organisms and through this method *Cryptosporidium* species obtained necessary molecules.

The pathogenic amoebozoan (and aetiological agent of amoebic dysentery) *Entamoeba histolytica* also lacks known enzyme machinery from both the MVA and MEP pathways for the production of IPP, but possesses a candidate gene for IPP isomerase (a key enzyme in the final stage of both the MVA and MEP pathway) and farnesyl transferase (a key enzyme in conversion of MVA pathway products). This could suggest either a novel pathway for the production of IPP or a scavenging

pathway for the organisms to obtain host IPP (Loftus et al., 2005). Interestingly despite lacking a gene for HMGR, it has been found that statins (an HMGR inhibitor) provided a synergistic growth inhibition of *E. histolytica* when used in conjunction with a farnesyl transferase inhibitor (Probst et al., 2019). This could again hint towards a currently unknown HMG-CoA binding enzyme present in *E. histolytica* and thus a novel isoprenoid biosynthetic pathway.

## Peroxisomes and Peroxisomal Targeting

### Peroxisome Overview

Eukaryotic cells are distinct from prokaryotes with their wide array of subcellular organelles. These organelles vary between species with differing evolutionary histories and specialised functions. One of these varied organelles is the peroxisome. Despite the diversity of functions, the peroxisome and associated organelles serve in the range of organisms in which they are found, they possess a highly conserved morphology and mode of biogenesis (Gabaldon et al., 2016). The structure of the peroxisome is a single membraned compartment with a widely varying contents depending on the organism.

The biogenesis of peroxisomes can occur via two processes, either *de novo* creation (budding) of peroxisomes from the ER or from the division of existing peroxisomes (Smith and Aitchison, 2013). In bakers' yeast and model microbial eukaryote *S. cerevisiae*, *de novo* peroxisome biogenesis does not occur in most circumstances with peroxisomes being formed via fission. Only when improper cell segregation occurs during cell division and the daughter cell is left without peroxisomes is the *de novo* pathway utilised (Motley and Hettema, 2007). In a study performed on dendritic cells from mice, cells were propagated *in vitro* and observed. It was observed that lamellar structures connected to the ER formed, with these lamellae possessing known peroxisomal proteins (Geuze et al., 2003). In another study, a range of mutant strains of *S. cerevisiae* were created in order to disrupt and thus elucidate the formation of peroxisomes (Erdmann and Kunau, 1992). Upon transformation of these mutants with the wild-type genes, peroxisome function was quickly restored. This again supports the presence of a *de novo* pathway of peroxisome biosynthesis.

The process of *de novo* peroxisome biogenesis occurs via vesicular intermediates. Peroxisomal membrane proteins are first inserted into the ER (van der Zand et al., 2010) before being organised into inactive sub-complexes. Vesicle budding occurs, creating two separate vesicles each containing one half of the peroxisomal membrane subcomplexes. These fuse forming a functional peroxisome

complete with protein import machinery which is then used to import peroxisomally destined proteins (van der Zand et al., 2012).

The more common process of peroxisome proliferation is the fission of existing peroxisomes. This process is mediated by the Pex11 family of proteins (Tower et al. 2011). These insert into the peroxisome membrane, causing an elongation of the structure. A collection of proteins forms the fission factors, responsible for membrane scission. These include three dynamin-related proteins (DRP3A, DRP3B and DRP5B) and a membrane receptor for these dynamin-related proteins (FISSION1) (Pan and Hu, 2011). Dynamin-related proteins or dynamin-like proteins (DLPs) are GTPases involved in membrane-related events in prokaryotes and eukaryotes including membrane fusion and scission (Jilly et al., 2018) and involved in endocytosis and cytokinesis in *Trypanosoma brucei* (Benz et al., 2017). Some proteins involved in peroxisome membrane scission are also involved in the proliferation of mitochondria, such as dynamin-related protein 1 (DRP1) also called DLP1 in mammals (Koch et al., 2005) hinting at conserved fission machinery linked to different subcellular compartments (Pan and Hu, 2011). Fusion of ER-originated vesicles with mature peroxisomes carrying a protein called Pex3 is necessary for the proliferation of peroxisomes. Pex3 deficient *S. cerevisiae* mutants show a complete lack of peroxisome structures (Motley and Hettema, 2007).

Biogenesis of peroxisomes is under the control of proteins known as peroxins encoded by PEX genes. Peroxins can be separated into three classes based on their role: peroxisome membrane synthesis, matrix protein import and DLP co-ordinated peroxisome proliferation (Fujiki et al., 2014). The exact number of encoded PEX genes can vary between organisms, with some proteins conserved and some absent between the different families of life. In both mammalian cells and yeast, Pex3 and Pex19 are involved in the targeting of peroxisomal membrane proteins (PMPs) where Pex3 acts as a docking receptor in the peroxisomal membrane and Pex19 acts as a receptor for the newly synthesised proteins and targets them to the membrane via a Pex3 receptor (Jansen and van der Klei, 2019). Interestingly, it has been shown that in *pex3* and *pex19* mutants of various yeast strains possess peroxisomes with varying levels of correctly imported PMPs, suggesting Pex3-Pex19 independent pathways for targeting of PMPs (Jansen and van der Klei, 2019). Mammalian cells possess an extra peroxin responsible for protein import when compared to *S. cerevisiae*, Pex16. (Cross et al., 2016). Mammalian Pex16 binds to Pex3 via a receptor site and acts as a chaperone for the Pex3-Pex19 complex targeting the complex to the peroxisomal membrane. (Matsuzaki and Fujiki, 2008). Mammalian Pex16 has also been seen to be ER-localised as well as peroxisomal, via co-translational targeting (Kim et al., 2006) where it functions in the targeting of other PMPs to peroxisomes via the ER (Aranovich et al., 2014). A Pex16 homolog is present in *Arabidopsis thaliana*

called SSE1 (Lin et al., 1999). Like mammalian Pex16, SSE1 is localised to both the peroxisome and ER in equal measure (Karnik and Trelease, 2005) and has been shown to be involved in a similar ER-peroxisome protein transport pathway via intermediate compartments (Karnik and Trelease, 2007). Despite the presence of a Pex16 protein in the yeast *Yarrowia lipolytica* (Eitzen et al., 1997), a Pex16 homolog is absent in *Saccharomyces cerevisiae* and all other yeast species (Kiel et al., 2006).

Peroxisomes display a wide range of functions within the organisms in which they appear, which are dictated by the various enzymes localised within them. Due to this varied nature, there are multiple sub-categorisations for peroxisomes present in different organisms based on their function.

In mammalian organisms, peroxisomes are widely distributed in almost all cell types with red blood cells being the only known exception. They are responsible for a range of different functions, but generally all linked by the theme of lipid metabolism, such as plasmalogen (ether-linked glycerolipids) synthesis and the  $\beta$ -oxidation of fatty acids (Lazarow, 1987). Peroxisomally located catalase enzymes are also responsible for the degradation of the oxidative  $H_2O_2$ . Oxidative stress is effectively reduced within the cell by isolating the  $H_2O_2$  and safely degrading it within the peroxisome (Fritz et al., 2007).

For example, within higher plants there are three classes of peroxisome. These are glyoxysomes, leaf peroxisomes and unspecialised peroxisomes whose presence varies dependant on the cell type in which they are present (Hayashi et al., 2000). The glyoxysomes are responsible for the sole localisation of an enzyme pathway, a specific pathway of gluconeogenesis in which fatty acids are converted into succinate.

Protists, in general, provide an interesting insight into the conservation or loss of peroxisomes as an adaptation to different nutrient sources and lifestyles. Apicomplexan parasites including the *Plasmodium* genus containing the malaria-causing parasite, completely lack peroxisomes (Schluter et al., 2006). There is strong evidence that some other apicomplexans (e.g. *Toxoplasma gondii* and *Eimeria tenella*) do however possess peroxisomes. This evidence includes comparative genomics to detect genes encoding Pex proteins in some apicomplexans (Moog et al., 2017). Life-stage specific proteomics and transcriptomics indicated upregulation of peroxisome biogenesis proteins and candidate peroxisomal proteins in tachyzoite, oocyst, and sporozoite stages of a complex life cycle, suggesting peroxisomes may only be important at certain points in the *Toxoplasma* life cycle including those stages which occur in the cat and are not yet possible to reproduce into tractable culture.

Glycosomes are class of peroxisome found in trypanosomatids. Much of the glycolytic pathway vital to the survival of these parasites is localised within glycosomes, with relative levels of glycolytic enzymes within the compartment varying during the different life stages of the organism. In the mammalian blood-stream form of *T. brucei*, 90% of glycosomal proteins are enzymes involved in glycolysis (Michels et al., 2006). The insect-borne parasites *Trypanosoma* and *Leishmania* undergo a wide range of cellular adaptations to their different hosts during their life cycles. Changes in glycosome composition are among these changes as the importance of glycolysis for the survival of the cell varies. When the parasites have transferred into a mammalian host bloodstream, this glucose-rich environment prompts a change to an exclusively glycolytic metabolism of ATP (Michels et al., 2006). The trypanosomatid glycosome also houses other enzymatic pathways essential to cell survival. For instance, trypanosomatids are unable to synthesis purine nucleotides and require salvaged host nucleotides to survive; key enzymes responsible for interconversion of salvaged purines are localised to the glycosome (Zarella-Boitz et al., 2004).

#### Peroxisomal Targeting Signals

Peroxisomes contain no DNA or translational machinery therefore, all proteins destined for this location must be translated in the cytoplasm and transported into peroxisomes. It was found that a specific C-terminal motif could determine the transport of proteins to the peroxisome (Gould et al., 1987). This was termed the peroxisomal targeting signal (PTS1). The original and thus textbook amino acid sequence was a C-terminal –SKL tripeptide motif i.e., a sequence corresponding to serine, lysine and leucine at the C-terminal end of the protein that directed peroxisomal import. This sequence was subsequently found to be slightly variable, where the first serine could be replaced with cysteine or methionine and the lysine could be replaced by either arginine or histidine. The leucine was found to be far more conserved than the other amino acids in the motif (Swinkels et al., 1992).

The PTS1 allows the binding of a polypeptide sequence to the peroxin PEX5. PEX5 binds to PTS1 sequences via a series of repeated 34 amino acid motifs called TPRs at the C-terminal end of the protein (Gatto et al., 2000). 95% of cellular PEX5 in both yeast and human cells is found in the cytosol, with the other 5% localised to the peroxisomes (Dodt et al., 1996). This suggest PEX5 is a cycling receptor, directing proteins to the peroxisome import machinery and then returning to the cytosol afterwards (Harper et al., 2003). PEX5 proteins within different organisms possess varying affinity for different variations of the PTS1, perhaps explaining the reason for variable PTS1

sequences. PEX5 may function based on threshold affinity, where specific PTS1 sequences may require an affinity over a set threshold in an organism to allow binding (Gatto et al., 2003). There is also evidence suggesting that residues adjacent to the PTS1 are also involved with PEX5 binding, constituting a 12 amino acid motif at the C-terminus (Neuberger et al., 2003).

A second set of consensus sequences was found to be necessary for some peroxisomally localised proteins, albeit a far smaller number than PTS1-containing proteins. They were termed PTS2 sequences. Despite being rarer than PTS1 sequences, PTS2-containing protein peroxisomal export is no less vital (Legakis and Terlecky, 2001). Like PTS1, PTS2 sequences are consensus sequences being more varied than PTS1 sequences and possessing different tendencies depending on the organism. (Kunze et al., 2015). The motif;  $-(R/K)(L/V/I/Q)XX(L/V/I/H/Q)(L/S/G/A/K)X(H/Q)(L/A/F)-$  comprises all known variants of the PTS2 (Petriv et al., 2004). Unlike PTS1 sequences, PTS2 sequences are located at the N-terminus of proteins and interact with a different peroxin receptor, PEX7. After interaction with the PTS2 containing protein, PEX7 interacts with another peroxin called PEX5L. This is responsible for peroxisomal import of the PTS2 containing protein.

### *Dictyostelium discoideum*

An interesting organism to consider, regarding the evolution and divergence of the biosynthesis of isoprenoids with eukaryotes is *Dictyostelium discoideum*. The Dictyostelid class of organisms belongs to the large taxonomic group Amoebozoa, a monophyletic group with a diverse range of organisms (Cavalier-Smith et al., 2015). The Amoebozoa are a sister-group to all opisthokonts (that group of organisms that include fungi, animals and a variety of protists, including choanoflagellates). A widely held view is that the common ancestor of extant amoebozoan protists diverged from other organisms at a very early point in eukaryotic evolution (Burki et al., 2020). Due to this early divergence in highly conserved biosynthetic pathways, some light could be shed on early eukaryotic adaptations and help elucidate the ancestral state of the metabolic pathway organisation in the first eukaryotes. This, of course applies to the pathway of isoprenoid biosynthesis and the origins of the various localisations of the enzymes of this pathway.

*Dictyostelium discoideum* serves as a popular model organism for eukaryotic cell differentiation and development due to a complex life cycle that incorporates multicellularity. (Huber, 2016). Thus, when in its vegetative amoebal state, *D. discoideum* feeds on bacteria. If prey sources in the environment (aerobic leaf litter on the forest floor) begin to deplete, *Dictyostelium* adopts a social lifestyle. This includes aggregation of 100,000s of cells guided by cAMP-lead signalling, thereby

forming a large mobile slug. Eventually, cells begin to form a stalk and at the head of the stalk a fruiting body from which spores are released to begin new *D. discoideum* colonies in areas where bacterial prey are present (Escalante and Vicente, 2000). During this stage of the lifestyle, cells begin to terminally differentiate into the various roles required for final fruiting body formation. This lifecycle exhibits similar processes to those carried out in multicellular organisms, such as cell migration, adhesion and cAMP signalling (Loomis, 2015). This hints to a key position in the phylogenetic tree, with the last common ancestor of social-amoeba such as *D. discoideum* being an early adopter of a multicellular lifestyle (Heidel et al., 2011). The *Dictyostelium* genome takes the form of 6 gene-dense chromosomes, encoding ~12,500 proteins in total (Eichinger et al., 2005).

### *Leishmania tarentolae*

A second organism used in this study was *Leishmania tarentolae*. It is a parasitic protist belonging to the large genus *Leishmania*. Although *L. tarentolae* is considered non-pathogenic to humans, various other species of the genus *Leishmania* are responsible for the disease leishmaniasis. Transmission of the parasite between people (or other animals) is typically vector-mediated. The vectors for *Leishmania* are female sandflies from the genera *Phlebotomus* and *Lutzomyia* with the female requiring her bloodmeal specifically for egg maturation (Akhoundi et al., 2016). *L. tarentolae* specifically infects reptiles and is thought to be non-pathogenic to humans. However, an intriguing study performed on a 300-year-old Brazilian mummy showed molecular signatures of *L. tarentolae* which raises the possibility of survival and replication of *L. tarentolae* cells within human macrophages and to allow systemic infection (Novo et al., 2015). *Leishmania* belongs to the Trypanosomatidae family within the order Kinetoplastida and phylum Euglenozoa. These protists are, like the *Amoebozoa*, also considered to be descended from an early-branching eukaryotic ancestor and together with a wide variety of related unusual protists belong to a wider group of protist organisms known as the *Excavata*. Organisms within *Leishmania* and the related genus *Trypanosoma* have been shown to possess interesting adaptations to the mevalonate pathway of isoprenoid biosynthesis.

The HMGR enzymes in both *Trypanosoma cruzi* and *Leishmania major* have been shown to possess interesting characteristics in a range of studies. One study isolated an HMGR gene from *T. cruzi* that was found to be smaller than other eukaryote HMGR enzymes and lacked the N-terminal ER membrane-spanning domains. An enzyme activity assay was also carried out in this study after fractionating *T. cruzi* cells, which showed 95% of HMGR activity was present in the cytosolic fraction

suggesting a soluble cytosolic HMGR form (Peña-Díaz et al., 1997). *Leishmania major* was found to possess a similar HMGR to the *T. cruzi* enzyme. This enzyme again showed similarity to the C-terminal regions (containing the catalytic domains) of HMGR enzymes from many other eukaryotes, but again lacked the N-terminal membrane-spanning region that would normally anchor the enzyme in the ER membrane (Peña-Díaz et al., 1997). A further study on HMGR in both trypanosomatids determined a mitochondrial localisation for these enzymes. These enzymes possessed an N-terminal mitochondrial signal, when removed resulted in accumulation of the enzyme in the *Leishmania* cytosol. This study also introduced *T. cruzi* HMGR into *S. cerevisiae*, where it was also localised in the mitochondria (Peña-Díaz et al., 2004).

In experiments performed on *Leishmania mexicana*, a variety of C<sub>14</sub>-labelled molecules were used to determine the preference of carbon sources for isoprenoid biosynthesis. It was found that although the usual isoprenoid precursor of mevalonic acid, acetyl-CoA was used, the amino acid leucine was utilised at a far higher rate (Ginger et al., 1999). These studies were then expanded by investigation of leucine and acetate usage as precursors in the biosynthesis of isoprenoids among various *Leishmania* species and *Trypanosoma cruzi*. It was shown that the *Leishmania* species utilised leucine more than acetate in varying quantities and the *T. cruzi* utilised some leucine but used mainly acetate as a carbon source for sterol biosynthesis (Ginger et al., 2000). This again highlights adaptations in the MVA pathway involving varying localisations and target molecules of enzymes within the pathway.

The remaining enzymes of the mevalonate pathway of isoprenoid biosynthesis in *Leishmania* are thought to occupy varying compartments in a similar fashion to other eukaryotes (Souza and Rodrigues, 2009). Some of the localisations of these enzymes are thought to be known in *Leishmania* such as farnesyl pyrophosphate synthase in the cytosol (Ortiz-Gómez et al., 2006). However, some localisations are less certain with squalene synthase being shown to possess both a PTS1 and a PTS2 (Opperdoes and Szikora, 2006) and dual peroxisomal/mitochondrial localisation being reported (Urbina et al., 2002). Even less certain is the localisation of squalene epoxidase which follows squalene synthase in the isoprenoid biosynthetic pathway. Considering this, *L. tarentolae* squalene epoxidase was chosen as a target protein for investigation in this study.

## Aims and Objectives

To attempt to determine the intracellular localisation of two different isoforms of HMGR found within *D. discoideum*. Another major aim is to carry out an analysis of the presence of peroxisomal

targeting signals and transmembrane spanning regions found with the two isoforms and other enzymes in the MVA pathway of *D. discoideum* and other eukaryotes. And finally, an analysis of the general organisation of the early reactions of the MVA pathway and the presence of lack of enzymes within a diverse range of protists.

## Materials and Methods

### *Bioinformatic Queries*

#### BlastP

NCBI BlastP (Altschul et al., 1990, States and Gish, 1994) was used as part of a bioinformatic study aimed at identifying the presence of orthologues of the HMGR, HMGS and HMG-CoA lyase (HMGL) and their associated isoforms from *D. discoideum*. cDNA sequences (obtained with thanks to Dr. Jason King of The University of Sheffield) were used to conduct the queries for all three genes and their isoforms. The “Expect Threshold” (E-value) of 1 was utilised for these searches, allowing for a wide range of sequences to be obtained in consideration of the wide range of organisms included

#### UniProt

The UniProt database (UniProt Consortium, 2019) was utilised for searching for more divergent examples of the MVA pathway enzymes. These may possess limited identity to, or exceed the E-value cut-off for, the *D. discoideum* reference sequences.

#### HMGR Mitochondrial Localisation Prediction

Three different mitochondrial localisation prediction software were used for this study: MitoProt II (Claros and Vincens, 1996), Predotar (Small et al., 2004) and TargetP 1.1 (Emanuelsson et al., 2000). HMGR sequences for the organisms of interest were entered into the prediction software. The prediction outcomes were recorded and divided into low, medium and high probabilities of mitochondrial localisation.

### *Dictyostelium discoideum* Culturing

#### *Growing Dictyostelium discoideum*

*Dictyostelium discoideum* cells (generously donated by Dr Jason King from the University of Sheffield) were grown in 10 ml liquid HL-5 media in petri dishes at 22°C, with cultures being sub-passaged when an approximate density of  $3 \times 10^6$  cells/ml was reached. Cells were counted using a Bright-Line hemocytometer (Sigma-Aldrich, US) and the density was calculated per millimetre. The sub-passage was carried out every two days at a 1:10 dilution. ‘HL-5 media with glucose’ was used

(Formedium, UK), with 35.5g of the powdered medium dissolved in 1L of water. When sub-culturing, care was taken in order to displace amoebae from their adhered position on the bottom of the petri-dish; this was done by using 10ml pipettes to draw up and then release the overlying medium.

#### Freezing *Dictyostelium discoideum*

*Dictyostelium discoideum* can be stored for up to 2 years at  $-80^{\circ}\text{C}$ . One confluent ( $\sim 3 \times 10^6$  cells/ml) dish can be used to make four dense 250 $\mu\text{l}$  aliquots of cells. The cells first needed to be resuspended, as done when passaging by washing the bottom of the plate with the liquid media already in the plate. This dislodged adhered cells from the surface of their plate. The 10ml of liquid HL-5 media was then transferred to a 50ml Falcon tube, which was centrifuged at  $500 \times g$  for 2 minutes in a bench-top swing out centrifuge. A freezing medium comprised of 45% HL-5 media, 45% BSA and 10% DMSO (Sigma-Aldrich) was then prepared and kept on ice. The supernatant from harvested cultures was discarded, and the cell pellet resulting from centrifugation re-suspended in 1ml of the freezing medium. This was then divided into four separate 250 $\mu\text{l}$  aliquots in Nalgene™ System 100™ Cryogenic Vials (Thermo Fisher Scientific, US) and kept at  $-20^{\circ}\text{C}$  for 2 hours before longer term storage at  $-80^{\circ}\text{C}$ .

#### Thawing *Dictyostelium discoideum* Cells

Frozen cells were taken and thawed quickly in a  $37^{\circ}\text{C}$  water bath until only a small amount of frozen material was left in the tube to ensure the temperature within the cryovial was below  $4^{\circ}\text{C}$ . This reduced stress on the cells from DMSO. These cells were then transferred to a petri-dish containing 10ml of normal HL-5 media. These cells were allowed to stand for 60 minutes at room temperature. The cells typically adhered to the bottom of a petri-dish within the first 15 min, and the additional time allowed the cells to acclimatise before a media change. A change into fresh HL-5 media was done to remove DMSO. The cells were left for 24 hours before antibiotics were added for selection; 10 $\mu\text{l}$  of 50mg/ml Hygromycin B (Sigma-Aldrich, US) and 10 $\mu\text{l}$  of 50mg/ml of Gibco™ Geneticin™ Selective Antibiotic (G418 Sulphate) (Thermo-Fisher Scientific, US).

#### *Dictyostelium discoideum* DNA Manipulation

##### Electroporation of *Dictyostelium discoideum*

3ml of resuspended cells taken from a confluent dish ( $3 \times 10^6$  cells/ml) were centrifuged at  $2,000 \times g$  for 2 mins. The supernatant was discarded and the pelleted cells were resuspended in 400 $\mu\text{l}$  of ice

cold E-buffer (10mM Potassium phosphate, 50mM sucrose). This was prepared for 5 samples, ready for transfection with the following plasmids, as displayed in Table 1.

**Table. 1 The different samples of *D. discoideum* cells after electroporation**

*D. discoideum* cells were split into five samples and transfected with different combinations of plasmids. + indicates the sample was transfected with this plasmid, -- indicates that they were not.

<i>D. discoideum</i> Transfectants	Plasmids		
	1039+DdHMGA	1039+DdHMGB	PTS1-RFP Plasmid
<b>1</b>	+	--	+
<b>2</b>	--	+	+
<b>3</b>	+	--	--
<b>4</b>	--	+	--
<b>5</b>	--	--	+

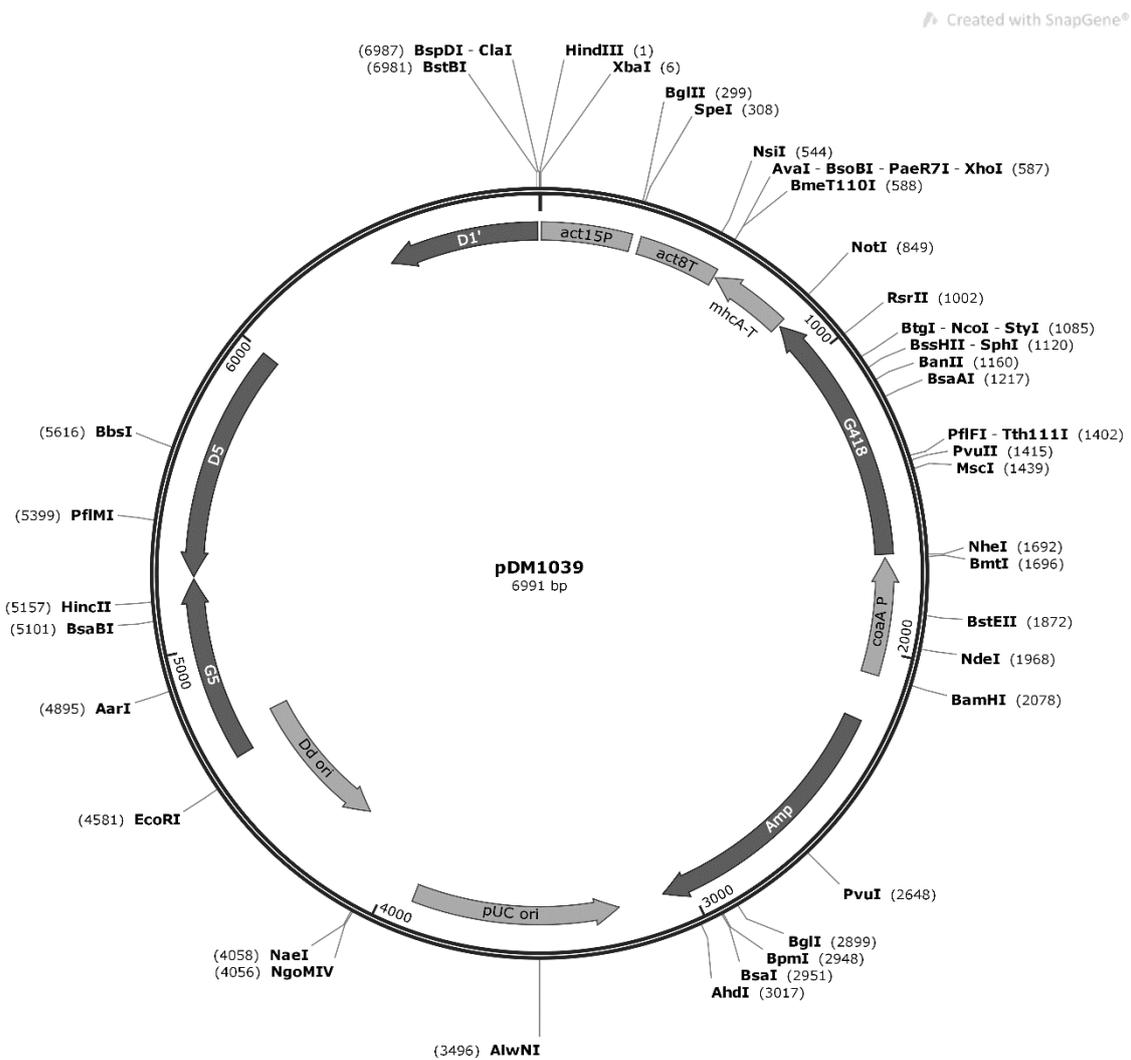
Plasmid DNA was added into the respective electroporation cuvettes (5µl of a 100ng/µl solution) (Bio-Rad). The electroporation was carried out using a Gene Pulser Xcell electroporation system (Bio-Rad) with an additional 5 ohm resistor. Electroporations were done with a capacitance of 3µF and a voltage of 1.2kV.

The cells were then transferred to 10ml of standard HL5 media in a petri dish. 24 hours later, 10µl of 50mg/ml Hygromycin B (Sigma-Aldrich, US) was added to the cells transfected with the PTS1-RFP plasmid. 10µl of 50mg/ml of Gibco™ Geneticin™ Selective Antibiotic (G418 Sulphate) (Thermo-Fisher Scientific, US) was added to the cells transfected with either of the *pDM* 1039 plasmids. Both drugs were added to the cells that were transfected with both plasmids.

#### Preparation of *pDM* 1039 Expression Plasmid

*pDM* 1039 is an untagged extrachromosomal plasmid designed as an expression vector for *Dictyostelium discoideum* (Veltman et al., 2009). It includes two selection genes, ampicillin for selection in *E. coli* and G418 for selection in *Dictyostelium*.

50µl of *E. coli* XL1-Blue competent cells (Agilent) were used in a standard heat-shock method of transformation with 0.5µl of purified *pDM* 1039. The plasmid DNA was added to the cells after thawing from -80°C. This mixture was then incubated on ice for 30 minutes. The heat shock was performed at 42°C for 45 seconds. Another incubation step on ice was performed for 3 minutes. These cells were then grown on an ampicillin-containing agar petri dish for selection. Plasmid DNA was then extracted from these cells using a Thermo Scientific™ GeneJET Plasmid Miniprep Kit.



**Figure 3. Map of the *pDM* 1039 plasmid used for transformation of *Dictyostelium discoideum*.** Includes the origin of replication sequences for both *E. coli* and *D. discoideum*, G418 and ampicillin selection markers and the expression cassette.

### Sub-cloning *DdHMGR* isoforms into *pDM* 1039 plasmids

Synthetic gene constructs encoding each isoform of *D. discoideum* HMGR (*DdHMGA* and *DdHMGB*) were sourced. Each synthetic gene used an endogenous (AT-rich) coding sequence except for recoding across two *XbaI* recognition sites in *DdHMGA* and one *XbaI* recognition site in *DdHMGB* and the additional presence of a C-terminal epitope double-myc tag and then a stop codon. These

synthetic genes were prefixed with a *Bgl*II recognition site and finished (downstream of the stop-codon) with a *Spe*I recognition site. The *Bgl*II and *Spe*I sites were to facilitate directional cloning into *Bgl*II-*Spe*I-digested *pDM* 1039. Following gel extraction of *Bgl*II-*Spe*I digested inserts and vector, ligation reactions were then carried out using T4 ligase (Roche) and transformed into competent *E. coli*. Nanodrop spectrophotometry was used to measure the concentration of insert and vector DNA prior to ligation.

From the transformation of competent *E. coli*, colonies were picked from overnight plates and grown overnight at 37°C in liquid LB media broth with ampicillin (100µg/ml). Plasmid DNA was isolated from these bacterial growths using a Thermo Scientific™ GeneJET Gel Extraction Kit. In this way, two constructs were prepared for transfection of *Dictyostelium*: p1039A (*pDM* 1039 + DdHMGA) and 1039B (*pDM* 1039 + DdHMGB). Plasmids prepared by mini-prep were analysed for the presence of cloned insert via diagnostic digests with *Bgl*II and *Spe*I and also sequenced using ABI technology.

### Fusion PCR-based GFP tagging of *L. tarentolae* SqE

In order to identify the localisation of the enzyme squalene epoxidase within the organism *Leishmania tarentolae* a Fusion PCR-based approach (Dean et al., 2015) was used in an attempt to express *Lt*SqE with a GFP sequence was inserted at the C-terminal GFP tag from its endogenous chromosomal C-terminal GFP tag from its endogenous chromosomal location. In this approach, partially open reading frame (ORF) for SqE and intergenic sequence are amplified alongside GFP

**Table 2. Primers used for the replication of homology flanks for *L. tarentolae* squalene epoxidase and the pPlotv1 NeonGreen Blast plasmid.**

DNA Sequence	Forward Primer	Reverse Primer	Nested Primer
N-terminal squalene epoxidase homology flank (ORF):	ttccatcaagggtctatgttg	gaggccgacaggcttgcgttgttg	ggttgcaaacagtttgaagg
C-terminal squalene epoxidase homology flank (UTR):	tgagcggcatctatttcttcagg	cacagacgtgtatatgtgtcc	actcctaacagacgtaacagg
pPlotv1 NeonGreen Blast:	ttccggttccggttctaagc	aagccaactaaatgggcactcg	-

coding sequence and a selectable marker gene for integration in the *L. tarentolae* genome via homologous recombination. In order to achieve the construct for transfection and homologous recombination into the *Leishmania* genome a specially developed Fusion PCR protocol was used (Dean et al., 2015). This construct features two 500bp homology flanks for the squalene epoxidase

gene, alongside a pPlot plasmid sequence consisting of a GFP gene and a Blastocidin resistance cassette. The fusion PCR protocol required a significant period of optimisation and required multiple attempts to successfully create the correct gene construct.

### PCR to Create a Tagging Amplicon for *Leishmania tarentolae*

#### Amplification of pPlot and Homology Flanks for *LtSqE*

A PCR reaction was carried out using Phusion® High-Fidelity DNA Polymerase to amplify both homology flanks for the squalene epoxidase gene and the C-terminal tagging pPlot plasmid. This reaction involved an annealing temperature of 65.1°C and an extension time of 35 seconds. This reaction resulted in an amplified sample of the plasmid, but no successful amplification of either of the homology flanks.

#### PCR to Create SqE Homology Flanks

Another protocol was required to achieve the amplification of the two 500bp homology flanks for the gene squalene epoxidase gene to allow homologous recombination. The two homology flanks include an ORF sequence that will occupy the N-terminal region of the tagging amplicon and an UTR sequence that occupy the C-terminal region.

This reaction used a lower annealing temperature at 48°C and an extension time of 7 seconds. This reaction successfully amplified the SqE ORF sequence however, untranslated region (UTR) sequence was unfortunately unsuccessful.

**Table 3. Reaction mixture for the second PCR attempt, amplifying the ORF and UTR sequences for the squalene epoxidase gene.**

Component	Volume
<b>H<sub>2</sub>O</b>	16µl
<b>dNTPS</b>	0.4µl
<b>5X Phusion Green HF Buffer</b>	4µl
<b>Phusion DNA Polymerase</b>	0.2µl
<b>Forward primer (1:10)</b>	0.5µl
<b>Reverse primer (1:10)</b>	0.5µl
<b><i>Leishmania tarentolae</i> genomic DNA</b>	2µl

**Table 4. Thermocycling conditions for the second PCR attempt, amplifying the ORF and UTR sequences for the squalene epoxidase gene.**

Cycle Step	Temperature	Time	Cycles
<b>Denaturation</b>	94°C	15s	30
<b>Annealing</b>	55°C	30s	
<b>Extension</b>	68°C	7 s	

A second attempt was carried out to amplify the SqE UTR sequence, using the same reaction mix and a lower annealing temperature at 46°C. Two reaction mixtures were made up (Table. 3), one with normal forward and reverse primers and one with the forward and nested primers for the UTR. The mixture with the nested primer used resulted in amplification of the sequence while the other did not.

#### First Attempt at Fusion PCR protocol

The purpose of the second round of PCR is to create a single tagging amplicon, composed of both of the homology flanks and the pPlot plasmid (complete with a GFP region and Blastocidin resistance cassette) between them (Dean et al., 2015). *Leishmania tarentolae* cells could then be transfected with this cassette and the pPlot sequence would be introduced at the C-terminus of the squalene epoxidase gene via homologous recombination.

**Table 5. Reaction mixture for the first Fusion PCR protocol, involving the GFP containing pPlotv1 and the SqE gene.**

Component	Volume	Time added
<b>H<sub>2</sub>O</b>	37.7µl	Start
<b>pPlotv1 sequence</b>	1.7µl (20ng)	Start
<b>SqE ORF sequence</b>	1.3µl (20ng)	Start
<b>SqE UTR sequence</b>	3.3µl (20ng)	Start
<b>dNTPS</b>	1µl	Start
<b>5X Phusion Green HF Buffer</b>	5µl	Start
<b>Phusion DNA Polymerase</b>	0.5µl	After first annealing step
<b>SqE ORF nested primer (B_ORF_Ltsqe)</b>	1µl	At the start of the 6 <sup>th</sup> denaturation step
<b>SqE UTR nested primer (UTR3n_Ltsqe)</b>	1µl	At the start of the 6 <sup>th</sup> denaturation step

**Table 6. Thermocycler conditions for the first Fusion PCR protocol, involving the GFP containing pPlotv1 and the SqE gene.**

Cycle Step	Temperature	Time	Cycles
<b>Initial Denaturation</b>	94°C	2 mins	1
<b>Annealing</b>	55°C	3 mins	1
<b>Extension</b>	68°C	3 mins	30 cycles
<b>Denaturation</b>	94°C	30 secs	
<b>Annealing</b>	55°C	3 mins	
<b>Final Extension</b>	68°C	7 mins	1

### Final attempt at creating the Fusion tagging amplicon

Previous attempts failed to produce significant amounts of the correct Fusion DNA sequence, with varying annealing temperatures in the previously mentioned PCR protocol and  $MgCl_2$ . The first step determined to resolve this was to use a specialised genomic DNA (gDNA) purification system, to improve the quality of the *L. tarentolae* template gDNA. A different DNA polymerase was used in the following protocols.

### *Leishmania* genomic DNA purification

*Leishmania tarentolae* gDNA was purified using the Promega Wizard SV Genomic DNA Purification System (Promega, USA). 10ml of *L. tarentolae* cells were taken at a culture density of  $\sim 10^7$  cells/ml and pelleted in a centrifuge at  $312 \times g$  for 10 minutes. After discarding the supernatant, the pellet was suspended in 1ml of phosphate-buffered saline (PBS). This was then centrifuged in a bench-top microfuge at  $3463 \times g$  for 3 minutes. The supernatant was again discarded and  $150\mu l$  of lysis buffer from the Wizard SV Genomic DNA Purification System was added to the pellet and after suspension the sample was added to a Wizard SV Minicolumn assembly. This sample was then spun in a microfuge at  $16,259 \times g$  for 3 minutes and the collection tube was emptied.  $650\mu l$  of Column Wash Solution (with ethanol) was added to the column before another spin in the centrifuge at  $16,259 \times g$  for 1 minute and the collection tube was emptied before being replaced. This step was repeated 4 times before a final spin at  $16,259 \times g$  for 2 minutes to dry the binding matrix. The Wizard SV Minicolumn was then taken from the collection tube and placed into a 1.5ml Eppendorf tube.  $250\mu l$  of nuclease-free water was added to the column and incubated for 2 minutes at room temperature. The Wizard SV Minicolumn assembly was then spun in a centrifuge at  $16,259 \times g$  for 1 minute. This causes the DNA sample to elute into the Eppendorf tube, which was then stored at  $-20$  until being used in later PCR protocols.

### New round of PCR amplifications of Fusion amplicon components

A new round of PCR protocols were devised, attempting to amplify all necessary homology flank sequences in a single PCR protocol. The newly purified *Leishmania* template gDNA was also used. Included in this PCR run were primers for amplifying ORF and UTR sequences not just for the squalene epoxidase gene but also ORF and UTR sequences for both the mevalonate kinase gene (MK) and the isovaleryl dehydrogenase gene (IVDH). MK is thought to be localised within the peroxisome (Carrero-Lerida et al., 2009) and IVDH is a known mitochondrial matrix enzyme (Shen et al., 2009). These can serve as reference signals when imaging *Leishmania* cells with tagged squalene epoxidase enzymes when attempting to determine the enzymes localisation.

The pPlotV1 sequence was also amplified using the same protocol and mixture as above, however with a slightly different annealing temperature of 63°C and a longer extension time of 1 minute

**Table 7. Reaction mixture for the second round of PCR, involving the GFP containing pPlotv1 and the SqE, MK and IVDH genes.**

Component	Volume
<b>H<sub>2</sub>O</b>	11µl
<b>dNTPS</b>	0.5µl
<b>5X Phusion Green HF Buffer</b>	4µl
<b>Phusion DNA Polymerase</b>	0.2µl
<b>Forward primer (1:10)</b>	1µl
<b>Reverse primer (1:10)</b>	1µl
<b><i>Leishmania tarentolae</i> genomic DNA</b>	2µl

**Table 8. Thermocycler conditions for the second round of PCR, involving the GFP containing pPlotv1 and the SqE, MK and IVDH genes.**

Cycle Step	Temperature	Time	Cycles
<b>Initial Denaturation</b>	98°C	30s	1
<b>Denaturation</b>	98°C	10s	10
<b>Annealing</b>	62°C	30s	30
<b>Extension</b>	72°C	15s	30
<b>Final Extension</b>	72°C	10 min	1

### Fusion cassette PCR using squalene epoxidase ORF and UTR sequences

The samples of SqE ORF and UTR and the pPlotV1 sequence were measured using a NanoDrop spectrophotometer to determine the concentration of DNA in each sample.

The Fusion PCR protocol requires 20ng of each previously amplified sequence to be carried out. This was measured (Table 9.) before being added in the correct concentration for a final Fusion PCR reaction (Table. 10).

**Table 9. Sequences obtained via the second round of PCR reactions were measured to determine the DNA concentration.**

Sample	DNA concentration (ng/ $\mu$ l)
<b>SqE ORF</b>	14.9
<b>SqE UTR</b>	6.1
<b>pPlotv1</b>	11.5

**Table 10. Reaction mixture for the second Fusion PCR protocol, involving the GFP containing pPlotv1 and the SqE gene.**

Component	Volume	Time added
<b>H<sub>2</sub>O</b>	37.7 $\mu$ l	Start
<b>pPlotv1 sequence</b>	1.7 $\mu$ l (20ng)	Start
<b>SqE ORF sequence</b>	1.3 $\mu$ l (20ng)	Start
<b>SqE UTR sequence</b>	3.3 $\mu$ l (20ng)	Start
<b>dNTPS</b>	1 $\mu$ l	Start
<b>Expand High Fidelity PCR system 10X buffer</b>	5 $\mu$ l	Start
<b>Expand High Fidelity PCR system DNA polymerase</b>	0.5 $\mu$ l	After first annealing step
<b>SqE ORF nested primer (B_ORF_Ltsqe)</b>	1 $\mu$ l	At the start of the 6 <sup>th</sup> denaturation step
<b>SqE UTR nested primer (UTR3n_Ltsqe)</b>	1 $\mu$ l	At the start of the 6 <sup>th</sup> denaturation step

**Table 11. Thermocycler conditions for the second Fusion PCR protocol, involving the GFP containing pPlotv1 and the SqE gene.**

Cycle Step	Temperature	Time	Cycles
<b>Initial Denaturation</b>	94	2 mins	1
<b>Annealing</b>	55	3 mins	1
<b>Extension</b>	68	3 mins	30 cycles
<b>Denaturation</b>	94	30 secs	
<b>Annealing</b>	55	3 mins	
<b>Final Extension</b>	68	7 mins	1

A Fusion tagging amplicon was only produced for the squalene epoxidase gene as time constraints prevented the production of constructs for the MK and IVDH genes. However, as ORF and UTR

sequences were partially produced for these genes, with slight variation this method could also be applied to those genes.

### MitoTracker™ Deep Red FM and ER-Tracker™ Red for Visualisation of Mitochondria and the ER in *D. discoideum*

In order to allow visualisation of the mitochondria and ER in *D. discoideum*, MitoTracker™ Deep Red FM and ER-Tracker™ Red (Invitrogen) were used. These stains were both used on different cultures of *D. discoideum* due to possible signal overlap during fluorescence microscopy but the method for utilisation of both is the same.

*D. discoideum* cultures at an approximate density of  $3 \times 10^6$  cells/ml were used for this method. Cultures for both myc-tagged *DdHMGR* isoforms were used, with separate cultures used for MitoTracker™ Deep Red FM and ER-Tracker™ Red, so four in total.  $5 \mu\text{l}$  of the 1mM stock solution of either MitoTracker™ Deep Red FM or ER-Tracker™ Red was added to the 10ml of HL-5 media in the *D. discoideum* culture petri-dish and left for one hour at  $22^\circ\text{C}$ . The *D. discoideum* cells were then resuspended in fresh HL-5 media and left for a further 1 hour at  $22^\circ\text{C}$ . These cells can then be used for slide preparation and immunofluorescence as seen below.

### Preparation of slides and immunofluorescence

1ml of *D. discoideum* cells were taken from the cultures at confluent densities ( $3 \times 10^6$  cells/ml) and pelleted in a centrifuge ( $385 \times g$  for 3 minutes). After discarding the supernatant, the pelleted cells were resuspended in PBS (Melford) After two PBS washes,  $50 \mu\text{l}$  of the cells were added to Thermo Scientific™ SuperFrost™ glass slides and were fixed in 3.7% paraformaldehyde (PFA) (Alfa Aesar).

For the *D. discoideum* control cultures that did not require immunofluorescence to image (wild-type and PTS1-RFP transfected cultures) and the *L. tarentolae* cultures, VECTASHIELD (Vector Laboratories) with 4',6-diamidino-2-phenylindole (DAPI) was added, followed by a glass cover slide.

For the cultures that required immunofluorescence, an additional step was required. Firstly, the slides are incubated with  $50 \mu\text{l}$  of blocking buffer for 1 hour (200ml PBS, 0.05% Tween 20 (Melford), 100mg BSA). Then anti-myc mouse primary antibody (ab18185) (Abcam) was diluted to 1:50 in blocking buffer before being added to the slide and allowed to incubate for 1 hour. After  $3 \times 5$  min washes in PBS, the goat anti-mouse IgG FITC-conjugated antibodies (Merck Millipore) diluted to 1:200 in blocking buffer were added and again an incubation step of 40 minutes was allowed.

Twelve types of *D. discoideum* slides were prepared. This included three slides featuring wild-type *D. discoideum* cells, with slides treated with primary antibodies only, secondary antibodies only and no antibody treatment. Slides were made up of *D. discoideum* cells transfected with; 1039+HMGA and the PTS1-RFP plasmid, 1039+HMGA and the PTS1-RFP plasmid, 1039+HMGA only and 1039+HMGB only. These slides were treated with both primary and secondary antibodies. The remaining four slides were cells transfected with 1039+HMGA or 1039+HMGB and incubated with MitoTracker™ Deep Red FM or ER-Tracker™ Red (Invitrogen). These slides were also treated with both primary and secondary antibodies.

Two *Leishmania tarentolae* slide types were prepared; a wild-type control and cells transfected with the GFP-tagged squalene epoxidase.

### Immunoblotting

SDS-PAGE (Sodium dodecyl sulphate-polyacrylamide gel electrophoresis) was used to separate cell lysate samples from wild type *D. discoideum* and *D. discoideum* transfected with either 1039+HMGA or 1039+HMGB. An 8% resolving gel was used made using 9.3ml of H<sub>2</sub>O, 5.3ml of 30% Acrylamide/Bis Solution (Bio Rad), 5ml of 1.5M Tris, 200µl of 10% SDS, 200µl of 20% ammonium persulphate and 12µl of TEMED (Tetramethylethylenediamine). The 5% stacking gel constituted 2.7 H<sub>2</sub>O, 670µl of 30% Acrylamide/Bis Solution (Bio Rad), 500µl of 1.0M Tris, 40µl of 10% SDS, 40µl of 20% ammonium persulphate and 4µl of TEMED.

For immunoblots proteins were transferred from gels onto Immuno-Blot® PVDF Membrane (Bio-Rad) following a standard Immunoblotting protocol. Membranes were probed using an anti-myc mouse antibody in a 1:50 dilution in PBS-Tween 20 containing 5% powdered, skimmed milk for one hour, followed by goat anti-mouse IgG conjugated with horse radish peroxidase antibodies diluted (Abcam) in a 1:1000 dilution in PBS-Tween 20 containing 5% powdered milk for a further one hour.

## Results

### Preparation of *D. discoideum* Plasmids Encoding HMGR Isoforms

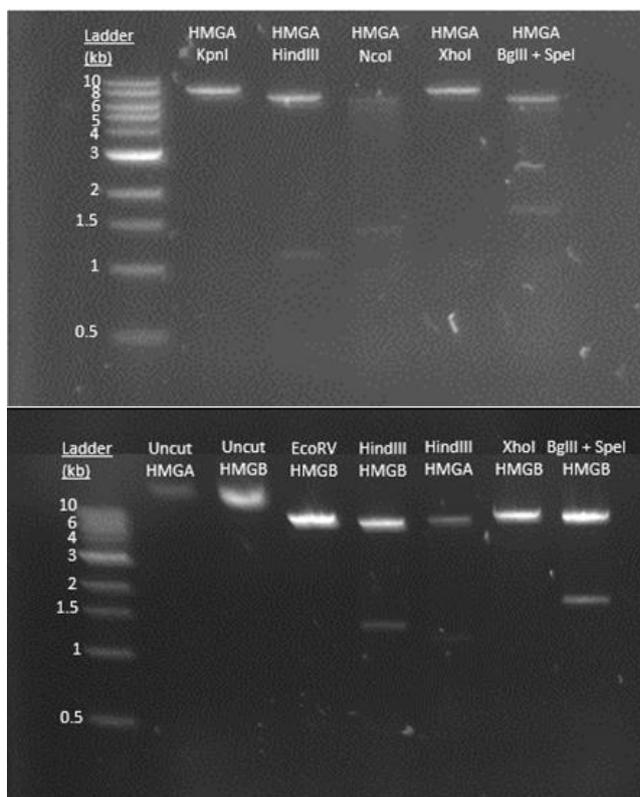
As a first step to expressing epitope-tagged HMGR isoforms in *D. discoideum* PCR was used in an attempt to amplify coding sequences for DdHMGA and DdHMGB. Unfortunately, these initial attempts at amplification of the HMGR isoforms failed, due to the AT rich nature of *Dictyostelium* protein-coding genes (Rosengarten et al., 2015) either making it difficult to design primers that allowed full length coding sequences to be amplified or allowing Taq polymerases to progress through AT-rich HMGR coding sequences. To address this problem, synthetic genes were designed that also coded, in frame, a tandem myc-epitope tag at the end of each coding sequence (Thermo Scientific).

The two genes encoding myc-tagged HMGR isoforms were each subcloned into *BglII-SpeI* digested *pDM* 1039 (Veltman et al., 2009) for transfection into *D. discoideum* and expression of recombinant gene. To confirm the sub-cloning of each HMGR insert into *pDM* 1039 diagnostic digests were carried out on purified mini-prep plasmids.

The diagnostic digests carried out are shown in Table 12 together with indication of the expected sizes of DNA fragments that would be evident following DNA agarose gel electrophoresis. These digests included the release of HMGA or HMGB insert, but also included digests that took advantage of restriction sites unique to DdHMGA (*KpnI*, *HindIII*, *NcoI*) or DdHMGB (*EcoRV*). In this way I could be more confident of cloning each similarly sized ORF into *pDM* 1039 prior to sending promising looking plasmids for DNA sequencing to confirm the presence of HMGR inserts. The results of DNA agarose gel electrophoresis following restriction digestion are shown in Figure 4.

**Table 12. Predicted fragment sizes for both of the insert+plasmid combinations with a range of restriction enzymes.**

Restriction Enzyme	Plasmid + Insert and size of fragment(s) (BP)	
	1039 + DdHMGA	1039 + DdHMGB
<i>KpnI</i>	8740	-
<i>HindIII</i>	1188, 7552	-
<i>NcoI</i>	1531, 7209	-
<i>XhoI</i>	314, 8426	314, 8348
<i>EcoRV</i>	-	8662
<i>BglII+SpeI</i>	1749, 6991	1671, 6991

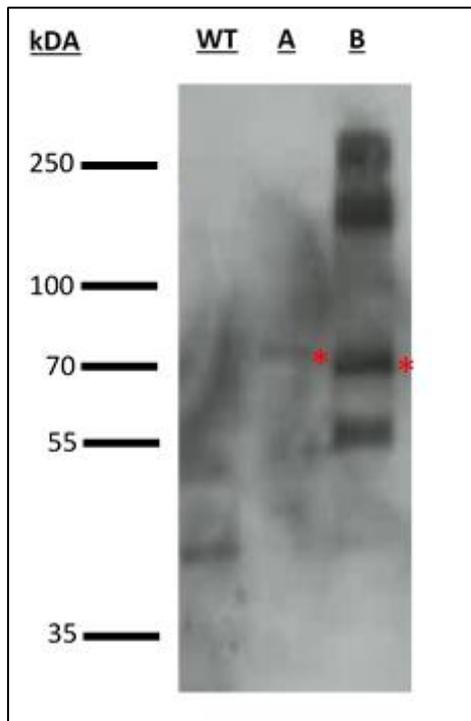


**Figure 4. Agarose gel electrophoresis of DNA fragments of HMGR isoform insert + 1039 plasmid after incubation with a range of restriction enzymes.**

The sequences of insert and plasmid are known and thus the restriction sites. The predicted sizes for these fragments are given in Table 12.

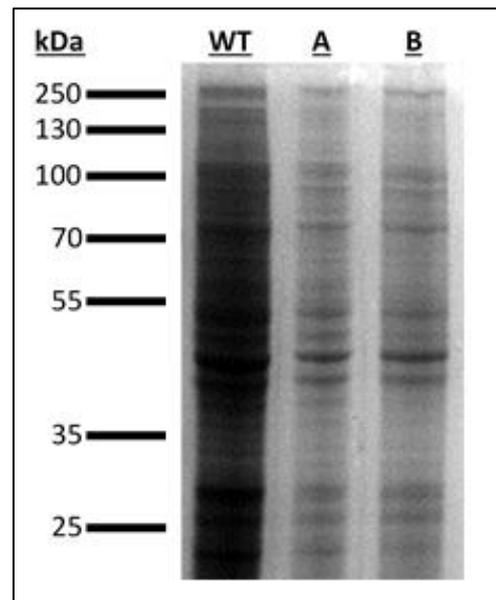
### Immunoblot analysis of HMGA::*myc* and HMGB::*myc* expression in *Dictyostelium discoideum*

Following the electroporation of *pDM 1039* plasmids encoding *DdHMGA::*myc** or *DdHMGB::*myc** and selection of stable transformants with G418, cell lysates were prepared to confirm expression of HMGR isoforms at the expected molecular mass. Amoebae from each transformed culture (*DdHMGA*+1039 and *DdHMGB*+1039) and wild type amoebae were collected for lysate preparation. Lysates were prepared by re-suspending amoebae pellets in boiling SDS-PAGE loading buffer for 3 min. From the Immunoblots shown in Figure 5, the primary anti-*myc* antibody detects protein of the predicted mass for both at the correct sizes predicted for each *myc*-tagged HMGR isoform. However, the band for *DdHMGA* was significantly less pronounced than that of *DdHMGB* (Fig. 6); since an equal amount of protein was loaded for SDS-PAGE from each transformed culture the Western blot suggested *DdHMGA::*myc** was expressed at a lower level than *DdHMGB::*myc**. Additionally, despite a greater amount of total protein in the loaded wild-type cell lysates (Fig. 6), there are few visible bands on the immunoblot suggesting successful targeting of the *DdHMGR* isoforms (Fig. 5). Additional bands may be due to non-specific binding of either of the antibodies or post-translationally modified proteins in the case of the higher molecular weight bands and proteolytically degraded fragments in the case of the lower molecular weight bands.



**Figure 5. An Immunoblot analysis of *D. discoideum* cells.**

WT refers to wild-type cells. A is *D. discoideum* cells transformed with the myc-tagged *DdHMGA*. B is *D. discoideum* cells transformed with the myc-tagged *DdHMGB*. The asterisks mark correctly sized *DdHMGA* and *DdHMGB* bands.

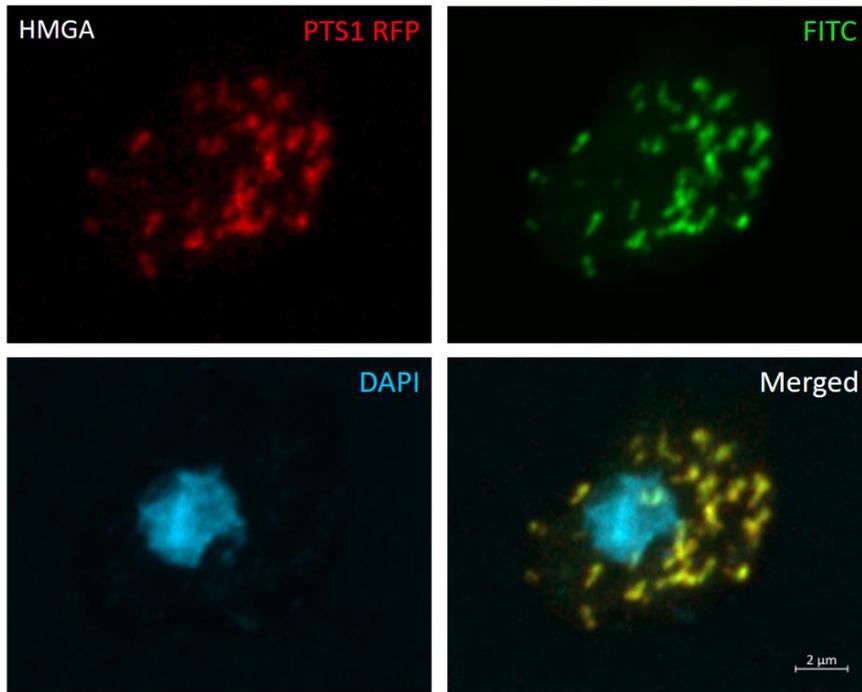


**Figure 6. An SDS-PAGE of lysate of three different cultures of *D. discoideum* stained with Coomassie Blue.**

WT refers to wild-type cells. A is *D. discoideum* cells transformed with the myc-tagged *DdHMGA*. B is *D. discoideum* cells transformed with the myc-tagged *DdHMGB*.

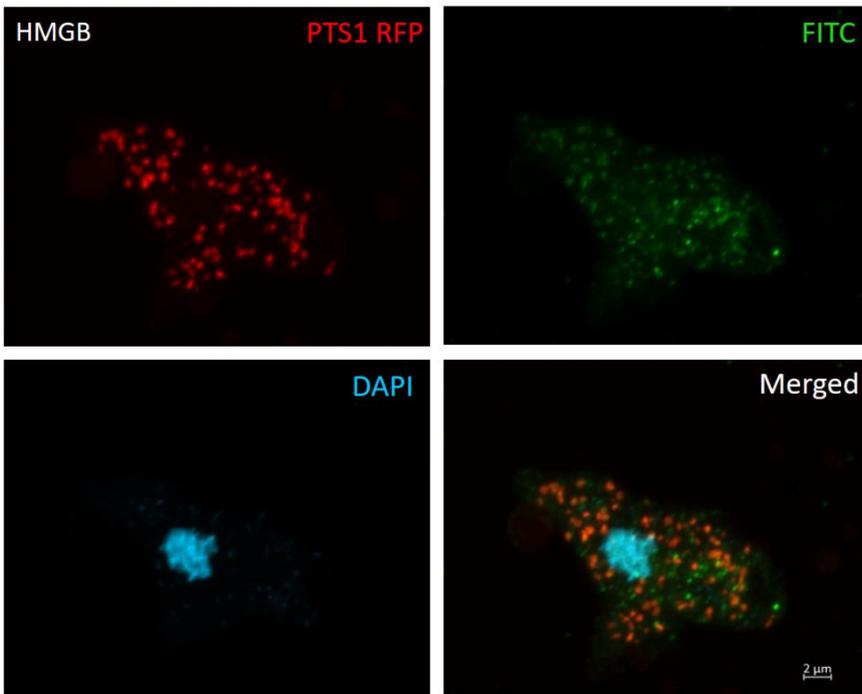
### Confocal microscopy analysis of *DdHMGA::myc* and *DdHMGB::myc* localisation

The possible reason underlying the presence of two HMGR isoforms in *D. discoideum* could perhaps be elucidated by identification of their site of intracellular localisation in amoebae. In an attempt to resolve this, *D. discoideum* cells transformed with either of the two myc-tagged HMGR isoforms were analysed via immunofluorescence. In addition to *D. discoideum* cells subject to expression of a single ectopic gene (either *DdHMGA::myc* or *DdHMGB::myc*), additional transformations were carried out to obtain amoebae that expressed either *DdHMGA* or *DdHMGB* plus a peroxisome-targeted variant of red fluorescent protein (PTS1-RFP). For indirect immunofluorescence detection of *DdHMGA::myc* or *DdHMGB::myc*, a FITC-conjugated secondary monoclonal antibody was used to detect primary anti-myc monoclonal antibody.



**Figure 7a:** *D. discoideum* cells transformed with a myc-tagged *DdHMGA* and a PTS1-linked RFP.

Fluorescence microscopy images of the same cell showing red fluorescence for the PTS1-linked RFP, green for the FITC conjugated secondary antibodies bound to anti-myc primary antibodies and blue for the DAPI stain applied to the live *Dictyostelium* cells.

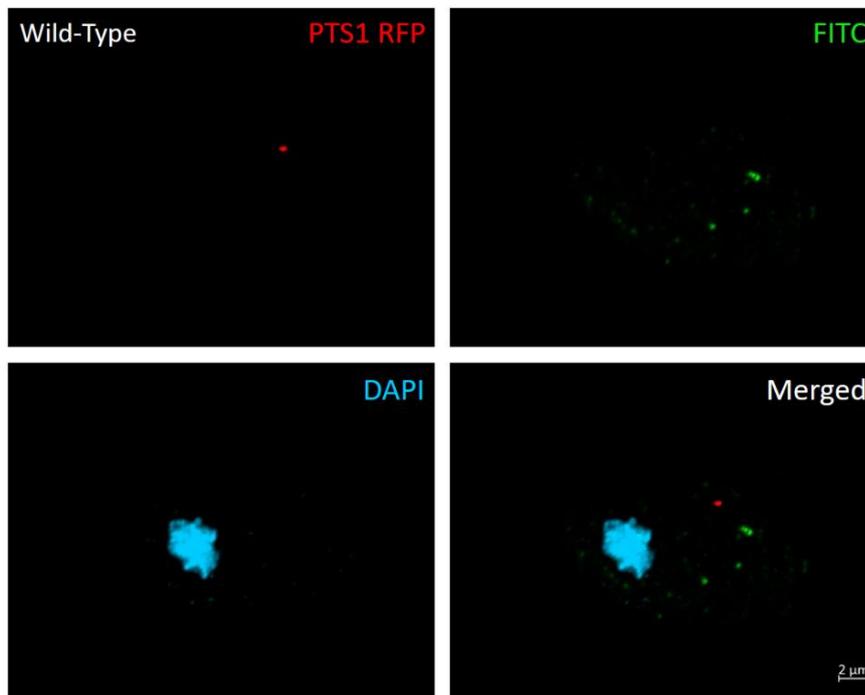


**Figure 7b:** *D. discoideum* cells transformed with a myc-tagged *DdHMGB* protein, and a PTS1-linked RFP.

Fluorescence microscopy images of the same cell showing red fluorescence for the PTS1-linked RFP, green for the FITC conjugated secondary antibodies bound to anti-myc primary antibodies and blue for the DAPI stain applied to the live *Dictyostelium* cells.

The myc-tagged *DdHMGA* expressing amoebae (Fig. 7a) showed a punctate signal for both the PTS1-RFP and *DdHMGA::myc*. This punctate signal for *DdHMGA::myc* did not resemble the ER to which eukaryotic class I HMGR is classically localised (Xu and Simoni, 2003). Instead, there was good co-localisation between the myc-tagged *DdHMGA* protein and PTS1-RFP strongly suggestive of peroxisomal localisation for the former. The myc-tagged *DdHMGB* amoeba also showed a punctate signal for both the indirect immunofluorescent detection of *DdHMGB::myc* and the direct

fluorescence of PTS1-RFP. However, in these amoebae there was little indication of extensive overlap between the two signals suggesting no evidence of co-localisation between the *Dd*HMGB and the PTS1-linked RFP (Fig. 7b). The patterns of fluorescence appeared similar to those acquired in a study concerning immunofluorescence microscopy of a *D. discoideum* PTS2-linked enzyme acetoacetyl-CoA thiolase (Isezaki et al., 2015). Wild-type cells show little native red or green fluorescence, indicating a successful transformation of the two HMGR isoforms in the other cultures (Fig. 8).

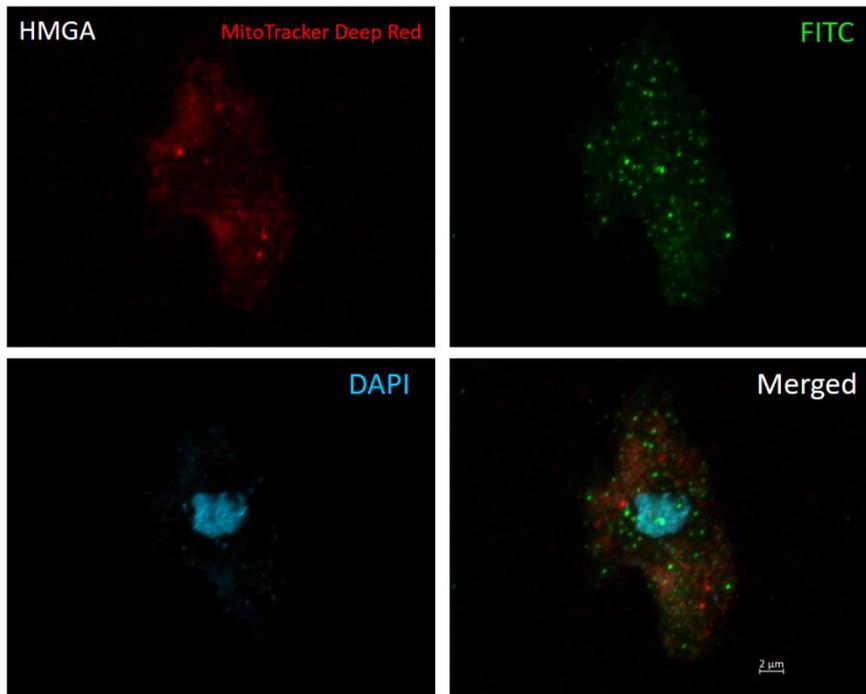


**Figure 8: Wild-type *D. discoideum* cells incubated with the same primary and secondary antibodies as the experimental cells.**

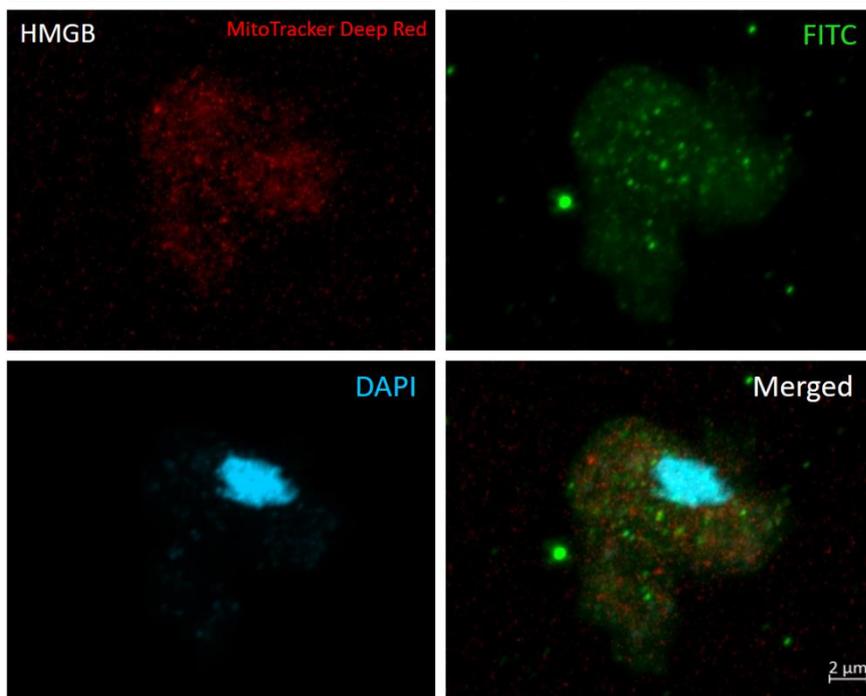
Fluorescence microscopy images of the same cell showing red fluorescence for the PTS1-linked RFP, green for the FITC conjugated secondary antibodies bound to anti-myc primary antibodies and blue for the DAPI stain applied to the live *Dictyostelium* cells.

To further elucidate the nature of the punctate signal of both tagged HMGR isoforms, myc-tagged *Dd*HMGA and *Dd*HMGB transformed cells were incubated with both MitoTracker™ Deep Red FM (Invitrogen™) and ER-Tracker™ Red (Invitrogen™)

A punctate signal can be seen for the MitoTracker™ Deep Red FM, suggesting a correct mitochondrial fluorescence signal for both sets of cells incubated with the stain. It can be seen that for both sets of cells, there is no obvious co-localisation between the stained mitochondria and either the myc-tagged *Dd*HMGA (Fig 9a) or *Dd*HMGB (Fig. 9b).

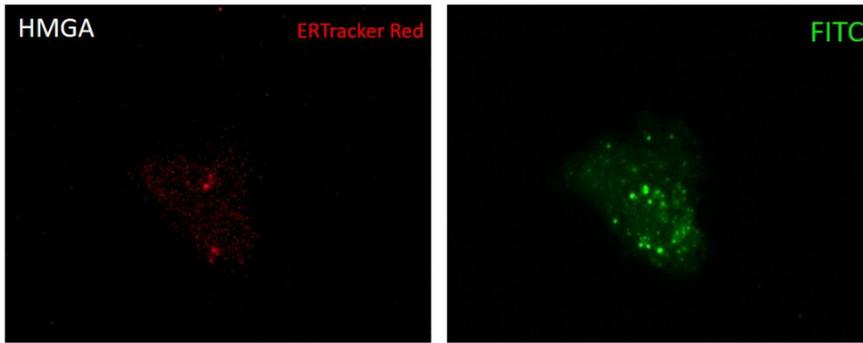


**Figure 9a: *D. discoideum* cells transformed with a myc-tagged *DdHMGA* and incubated with MitoTracker™ Deep Red FM.** Fluorescence microscopy images of the same cell showing red fluorescence for the MitoTracker™ Deep Red FM, green for the FITC conjugated secondary antibodies bound to anti-myc primary antibodies and blue for the DAPI stain applied to the live *Dictyostelium* cells.



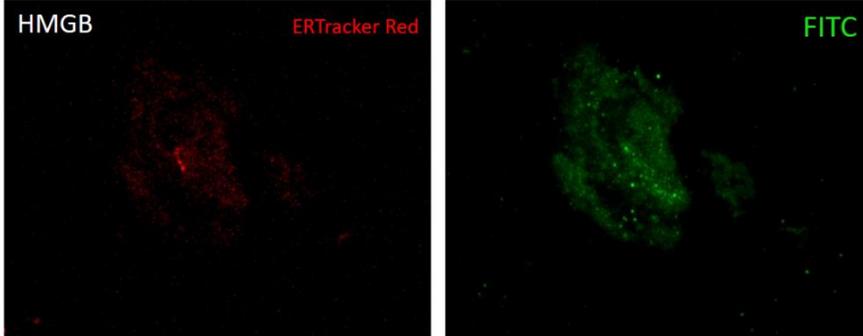
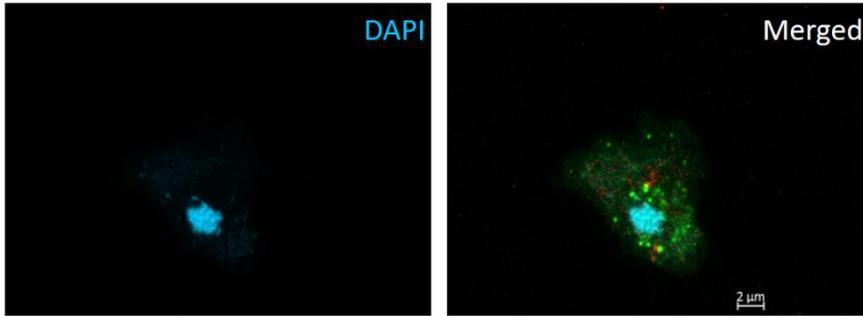
**Figure 9b: *D. discoideum* cells transformed with a myc-tagged *DdHMGB* protein, and incubated with MitoTracker™ Deep Red FM.** Fluorescence microscopy images of the same cell showing red fluorescence for the MitoTracker™ Deep Red FM, green for the FITC conjugated secondary antibodies bound to anti-myc primary antibodies and blue for the DAPI stain applied to the live *Dictyostelium* cells.

The cells incubated with ER-Tracker™ Red show a poor fluorescence for the stain, with no visible ER structure (Fig. 10a and Fig. 10b). This suggested that the staining process was not successful within the timeline of my project. However a similar punctate distribution of FITC-detected *DdHMGA*::myc or *DdHMGB*::myc was seen in both sets of cells as in previous samples.



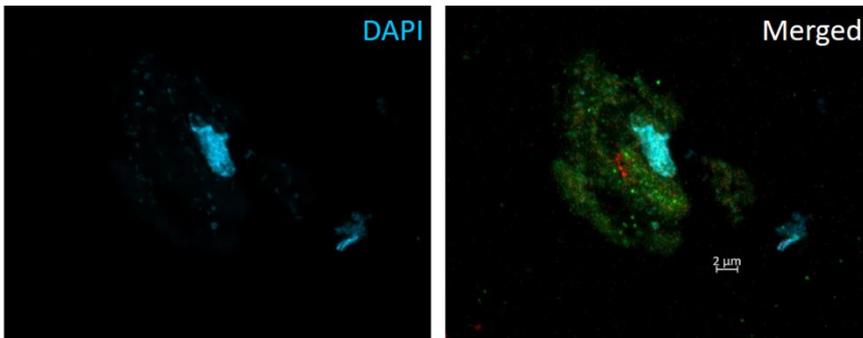
**Figure 10a: *D. discoideum* cells transformed with either a myc-tagged *DdHMGA*, and incubated with ER-Tracker™ Red.**

Fluorescence microscopy images of the same cell showing red fluorescence for the ER-Tracker™ Red stain, green for the FITC conjugated secondary antibodies bound to anti-myc primary antibodies and blue for the DAPI stain applied to the live *Dictyostelium* cells.



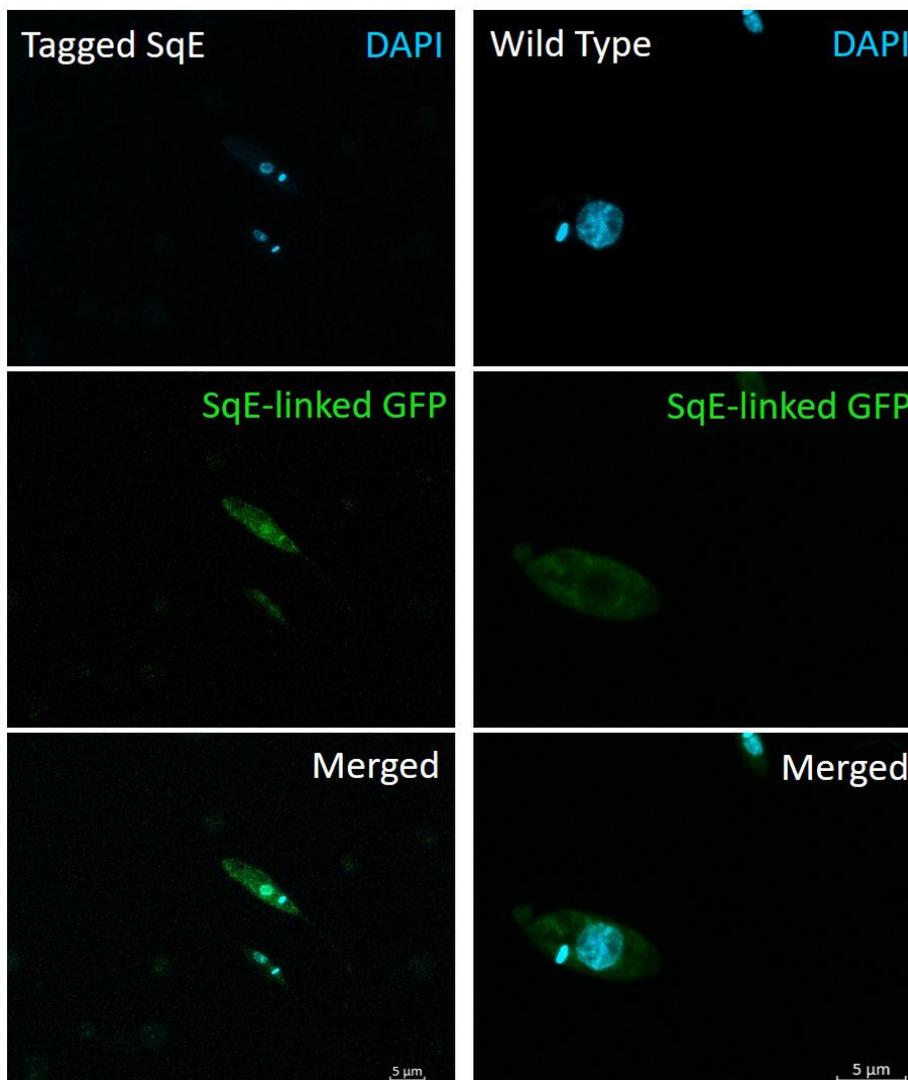
**Figure 10b: *D. discoideum* cells transformed with either a myc-tagged *DdHMGB* protein, and incubated with ER-Tracker™ Red.**

Fluorescence microscopy images of the same cell showing red fluorescence for the ER-Tracker™ Red stain, green for the FITC conjugated secondary antibodies bound to anti-myc primary antibodies and blue for the DAPI stain applied to the live *Dictyostelium* cells.



*Leishmania tarentolae* cells transformed with a GFP-tagged squalene epoxidase enzyme

A PCR-derived construct was created, allowing for the insertion via homologous recombination of a GFP ORF upstream of *LtSqE*. This would allow for the detection via fluorescence microscopy of the (Neon-green) GFP-tagged SqE. A ZEISS confocal microscope was used to visualise the fluorescence and capture images for analysis. DAPI staining highlighted the DNA in both the nucleus and the kinetoplast, an identifying characteristic of Trypanosomatids (Sunter and Gull, 2017) (Fig. 11). There is a degree of native green fluorescence for the wild-type *L. tarentolae* cells, however this is less than the fluorescence presented by the cells transformed with the GFP-tagged SqE (Fig. 11).



**Figure 11. *L. tarentolae* cells transformed with the artificial Fusion tagging construct and wild type *L. tarentolae* cells.**

Fluorescence microscopy images of both *L. tarentolae* cells transformed with the squalene epoxidase (SqE) tagging amplicon and wild-type cells. The green fluorescence is the SqE linked GFP, and the blue fluorescence is the DAPI stain applied to live *L. tarentolae* cells taken using the same exposure and settings.

Overall, the fluorescence microscopy of the *L. tarentolae* shows no conclusive evidence of successful translation of the GFP-tagged SqE despite the amoebae imaged surviving the necessary ampicillin selection.

## Phylogenomic distribution of the Early Enzymes of the Mevalonate Pathway in Microbial Eukaryotes

One of the first steps of the bioinformatic study was to identify principally within a range of microbial eukaryotes key enzymes within the MVA pathway of isoprenoid synthesis: HMGR, HMGS and HMGL. The presence or lack of these enzymes is in itself informative and can help in understanding relationships between and within clades. The organism of focus for much of the experimental work, *Dictyostelium discoideum*, was found to possess two isoforms for each of HMGR and HMGS. In order to determine both the presence and possible isoforms of these key enzymes in a range of other eukaryotic organisms, sequence-based searches were carried out. The online software NCBI BLAST (Altschul et al., 1990, States and Gish, 1994) and EUPathDB (Aurrecochea et al., 2017) were used to gather this data. This data can be seen in Table 13, including details of the protists included in the survey. Most taxa in the list were found to possess a single HMGR form. In the eukaryotic organisms, this is likely the typical ER bound, cytosol facing form (Friesen and Rodwell, 2004). However, in plants there is typically more than one gene, with *Arabidopsis thaliana* possessing two (Antolín-Llovera et al., 2011). These genes can be alternatively spliced to produce a larger number of isoforms, but Table 13 shows only the two fully transcribed proteins.

The HMGR enzymes in plants have been shown to possess dual localisation to both the ER and cytosol, shown using GFP tagging (Leivar et al., 2005). The other two Archaeplastida chosen for this study are red algae species, *Cyanidioschyzon merolae* and *Galdieria sulphuraria*. *Cyanidioschyzon merolae* does not appear to possess an HMGR enzyme but has preserved the other two enzymes that have involvement in this pathway. The loss of HMGR in this species suggests a loss of the MVA pathway for isoprenoid biosynthesis, instead favouring the conserved MEP pathway (Qiu et al., 2016). The species possesses an HMGS (HMG CoA synthase) enzyme, but this protein has multiple functions within eukaryotes including serving as a rate limiting step in the production of ketone bodies (Shimazu et al., 2010). HMGL is also present, suggesting a preservation of leucine catabolism. *Galdieria sulphuraria* however, appears to have preserved the HMGR and HMGS enzymes, suggesting utilisation of the MVA pathway. This is interesting considering both *Cyanidioschyzon merolae* and *Galdieria sulphuraria* both form a monophyly within the Cyanidiales (Yoon et al., 2002), suggesting either a separate loss of the MVA pathway by *C. merolae* than that of the other red algae

categorised by (Qiu et al., 2016), or a reacquisition of the MVA pathway in *G. sulphuraria*. The former alternative is more likely. For the two animal species analysed, the range and number of enzymes present is as expected. Both *Homo sapiens* and *Mus musculus* is known to utilise the MVA pathway. Both these organisms possess two isoforms of HMGS; orthologues of both isoforms have also been isolated and characterised in rats (Ayte et al., 1990). One isoform, HMGS-1, functions with the MVA pathway and the other is involved within the production of ketone bodies (Cotter et al., 2014).

*Blastocystis hominis*, a member of the stramenopiles, interestingly, according to the searches done, was found to possess two isoforms of HMGR. The presence of HMGR is also supported by a study showing a prevention of isoprenoid synthesis via the MVA pathway when statins were introduced being the inhibitor of HMGR (Basyoni et al., 2018). Interestingly, there is a lack of HMGS in *Blastocystis hominis*, raising the question of the carbon source for the MVA pathway with a possibility of leucine catabolism as seen in some trypanosomes (Ginger et al., 2000).

*Pythium* and *Phytophthora* species did not appear to possess significantly similar enzymes to *D. discoideum* as BlastP queries utilising the *D. discoideum* reference sequences produced no similar proteins below the E-value cut-off. However additional searches were made using the UniProt database in order to determine the presence of possible MVA pathway enzymes. These were included as percentage similarity to *D. discoideum* is not the only criteria of interest in this study, and the presence of MVA pathway enzymes in and of themselves are of sufficient interest. Most of the *Phytophthora* species possess one of each of the enzymes. *Phytophthora cactorum*, *Phytophthora kernoviae* and *Phytophthora parasitica* show no HMGL enzyme, however this may simply reflect absence from draft, rather than complete genomes. Interestingly, the three *Phytophthora* species lacking HMGL are each from different subclades (in the case of *P. cactorum* and *P. parasitica*) or from a different clade all together (in the case of *P. kernoviae*) (Yang et al., 2017). *P. cactorum*, *P. parasitica* and *P. infestans* are all grouped within different subclades within clade alongside a number of other species. As HMGL is present in *P. infestans* but not in the other two species mentioned, suggests multiple gene loss events for each of these species.

The species *Phytophthora nicotianae* and *Phytophthora infestans* possess two copies of HMGL and HMGS respectively. *Saccharomyces cerevisiae* possess two isoforms of both HMGR and HMGS, with each of these isoforms being involved in the MVA pathway of isoprenoid synthesis (Bröker et al., 2018). *D. discoideum* itself appears to possess two isoforms of both HMGR and HMGS and is known to utilise the MVA pathway (Kasahara et al., 2004), in which at least one of the isoforms of each enzyme must function. Information regarding the localisation of the HMGR isoforms has been

gathered as part of this project. The problem with considering this information alone is that although some of the missing enzymes may be genuinely informative, there is always the possibility that in some less well studied organisms, this apparent absence is due instead to the enzyme not yet being identified.

**Table 13. The presence of each of the enzymes of interest (HMGR, HMGS and HMGL) is shown for a selection of different species.**

1 indicates presence of one isoform of the enzyme and 2 indicates the presence of two. – indicates that this enzyme has not been identified within the organism. Highlighted black organisms were found to possess these enzymes via UniProt searches, but sequences exceeded the E-value cut-off when using the *D. discoideum* reference sequences.

<b>List of different organisms and the presence of enzymes related to the mevalonate pathway of isoprenoid synthesis</b>			
<b>Organism</b>	<b>HMG-CoA Reductase</b>	<b>HMG-CoA Synthase</b>	<b>HMG-CoA Lyase</b>
<b>Amoebozoa</b>			
<i>Dictyostelium discoideum</i>	2	2	1
<i>Acanthamoeba castellanii</i>	1	1	1
<b>Stramenopiles</b>			
<i>Pythium insidiosum</i>	1	1	-
<i>Phytophthora cactorum</i>	1	1	-
<i>Phytophthora kernoviae</i>	1	1	-
<i>Phytophthora sojae</i>	1	1	1
<i>Phytophthora megakarya</i>	1	1	1
<i>Phytophthora parasitica</i>	1	1	-
<i>Phytophthora infestans</i>	1	2	1
<i>Thalassiosira pseudonana</i>	1	1	-
<i>Phaeodactylum tricornutum</i>	1	1	-
<i>Blastocystis hominis</i>	2	-	-
<b>Alveolates</b>			
<i>Tetrahymena thermophila</i>	-	1	-
<i>Paramecium tetraurelia</i>	1	1	-
<b>Excavata</b>			
<i>Trypanosoma brucei</i>	1	1	1
<i>Naegleria gruberi</i>	2	1	-
<b>Animalia</b>			
<i>Homo sapiens</i>	3	2	1
<i>Mus musculus</i>	2	2	1
<b>Fungi</b>			
<i>Saccharomyces cerevisiae</i>	2	2	-
<i>Histoplasma capsulatum</i> (Strain G186AR)	1	-	-
<i>Histoplasma capsulatum</i> (Strain H88)	1	1	1
<b>Archaeplastida</b>			
<i>Arabidopsis thaliana</i>	2	1	1

<i>Cyanidioschyzon merolae</i>	-	1	1
<i>Galdieria sulphuraria</i>	1	1	-
<b>Bacteria</b>			
<i>Pseudomonas mevalonii</i>	1	-	1

## Determination of Mitochondrial Localisation of HMG-CoA Reductase Enzymes Using Prediction Software

One of the major focuses of this study is to attempt to determine localisation of various isoforms of enzymes within the MVA pathway of isoprenoid synthesis. To this end, a study was performed on the HMGR enzymes from a range of organisms, including the two isoforms within *Dictyostelium discoideum*. The study itself used three different prediction programs, MitoProt II (Claros and Vincens, 1996), Predotar (Small et al., 2004) and TargetP 1.1 (Emanuelsson et al., 2000). This was done in order to increase the reliability of the results and allow further confidence in drawing conclusions from results. These programs search sequences for Transit Peptides, in this case Mitochondrial Transport Peptides (mTP) (Emanuelsson et al., 2007). These sequences are very variable, but never include acidic amino acids and contain high proportions of alanine, arginine and serine (von Heijne et al., 1989). *Dictyostelium discoideum* showed moderate potential for mitochondrial localisation for one of the isoforms and very unlikely mitochondrial localisation for the other. As seen from these predictions, *T. brucei* shows mitochondrial localisation for HMGR. This information is supported by studies performed on the enzyme in this organism, which through elimination of the mitochondrial localisation sequence caused cytosolic accumulation of the enzyme (Peña-Diaz et al., 2004). The results from several of the organisms highlight the possible unreliability of using such software. For example, the enzymes for the two strains of *Histoplasma capsulatum* showed a wide range of predictions, from very low to very high so it is difficult to draw any conclusions for those particular organisms. *Blastocystis hominis* shows a very low prediction for mitochondrial localisation for HMGR1 and a very high probability of mitochondrial localisation for HMGR2, which appears similar to the probabilities seen for the HMGR isoforms of *Dictyostelium discoideum*. For the *Phytophthora* genus, the HMGRs for all species selected were not predicted to be mitochondrial by any of the software used.

**Table 14. The software Mitoprot II, Predotar and TargetP 1.1 were used to predict mitochondrial localisation of HMGR enzymes from various species.**

White and \* dictates low (<0.3), grey and \*\* moderate (>0.3 and <0.7) and black and \*\*\* high (>0.7) probability of mitochondrial localisation.

Organism and Enzyme	Probability of export to the mitochondria		
	MitoProt II	Predotar	TargetP 1.1
<i>Dictyostelium discoideum</i> HMGR B (XP_643058.1)	0.8516***	0.49**	0.416**
<i>Dictyostelium discoideum</i> HMGR A (XP_646489.1)	0.0224*	0.00*	0.009*
<i>Phytophthora cactorum</i> HMGR (RAW37381.1)	0.0587*	0.00*	0.044*
<i>Phytophthora kernoviae</i> HMGR (RLN20925.1)	0.0495*	0.00*	0.032*
<i>Phytophthora sojae</i> HMGR (EGZ20536.1)	0.0368*	0.00*	0.070*
<i>Phytophthora megakarya</i> HMGR (OWZ01501.1)	0.0399*	0.00*	0.054*
<i>Phytophthora infestans</i> HMGR (XP_002897124.1)	0.1386*	0.07*	0.080*
<i>Thalassiosira pseudonana</i> HMGR (XP_002289576.1)	-	-	0.155*
<i>Phaeodactylum tricornutum</i> HMGR (XP_002185302.1)	-	-	0.078*
<i>Blastocystis hominis</i> HMGR 1 (XP_012897778.1)	-	-	0.041*
<i>Blastocystis hominis</i> HMGR 2 (XP_014528476.1)	0.9579***	0.41**	0.903***
<i>Paramecium tetraurelia</i> HMGR (XP_001438919.1)	0.8543***	0.85***	0.732***
<i>Trypanosoma brucei brucei</i> HMGR (XP_845571)	0.4331**	0.01*	0.105*
<i>Naegleria gruberi</i> HMGR (XP_002670914.1)	-	0.15*	0.657**
<i>Naegleria gruberi</i> HMGR (XP_002668160.1)	0.1438*	0.01*	0.233*
<i>Acanthamoeba castellanii</i> HMGR 1 (ELR17548.1)	0.9890***	0.71***	0.969***
<i>Acanthamoeba castellanii</i> HMGR 2 (ELR15390.1)	0.0168*	0.00*	0.068*
<i>Homo sapiens</i> HMGR 1 (NP_001351116.1)	0.2128*	0.29*	0.281*
<i>Homo sapiens</i> HMGR 2 (NP_001124468.1)	0.1162*	0.29*	0.281*
<i>Homo sapiens</i> HMGR 3 (XP_011541661.1)	0.0409*	0.00*	0.620**
<i>Mus musculus</i> HMGR 1 (NP_001347094.1)	0.1449*	0.30*	0.281*
<i>Mus musculus</i> HMGR 2 (NP_001347095.1)	0.1916*	0.01*	0.240*
<i>Saccharomyces cerevisiae</i> S288C HMGR (NP_013636.1)	0.4131**	0.74***	0.624**

<i>Saccharomyces cerevisiae</i> S288C HMGR 2 (NP_013555.1)	0.7843***	0.67**	0.323**
<i>Histoplasma capsulatum</i> Strain G186AR HMGR (EEH10959.1)	0.1833*	0.19*	0.799***
<i>Histoplasma capsulatum</i> Strain H88 HMGR (EGC45925.1)	0.3278**	0.20*	0.793***
<i>Arabidopsis thaliana</i> HMGR 1 (AAA76821.1)	0.1636*	0.03*	0.432**
<i>Arabidopsis thaliana</i> HMGR 2 (NP_179329.1)	0.2252*	0.01*	0.295*
<i>Galdieria sulphuraria</i> HMGR (EME29007.1)	0.0813*	0.00*	0.010*
<i>Pseudomonas mevalonii</i> HMGR (AAA25837.1)	0.1001*	0.17*	0.453**

### Localisation of the Enzymes of the MVA Pathway of Isoprenoid Biosynthesis

In eukaryotes lacking the MEP pathway, the synthesis of isoprenoids is reliant solely on the mevalonate pathway, which produces dimethylallyl pyrophosphate (DMAPP). The DMAPP can then be used in order to produce squalene. There are nine enzymes that carry out this process in eukaryotes. The localisation of these various enzymes varies between organisms. Localisation of enzymes within a particular organelle can have significant effects on its function and efficiency, as certain molecules are unavailable or available in lesser concentrations within different organelles. The original theory claimed that the enzymatic reactions with this pathway took place in the cytosol and the ER (Kovacs et al., 2007). However, a range of more recent evidence has been produced that claims peroxisomal localisation for many of the enzymes within the pathway.

A study was carried out, analysing the sequences of these enzymes from a selection of different organisms. This study aimed to gather information regarding the localisation of enzymes in the MVA pathway within different organelles. TrypTag.org (Dean et al., 2017), was used to collect localisation information for *T. brucei* (Table 15). This study used fluorescent protein tags placed at either the N- or C-terminus of a target protein before being analysed and imaged.

Important sequences to consider in regard to this study are the peroxisomal transport sequences (PTS1 and PTS2). PTS1 is a C-terminal consensus sequence, consisting of three amino acids (S/A/C)(K/H/R)(L/M) (Kovacs et al., 2002). The PTS1 is found in most known peroxisomally localised proteins. PTS2 is a nine amino acid consensus sequence (R)-(A/L/Q/I)-X5-(H)-(I/L/F) (Wang et al., 2008). This sequence is less common than the PTS1, and present in less proteins known to be peroxisomal localised (Brocard and Hartig et al., 2006). Both of these sequences can be used as an indicator for peroxisomal localisation, but total confidence cannot be placed in this as presence or

lack of PTS sequences does not necessarily always determine protein localisation. A survey was performed on 37,080 genes from *L. major*, *T. brucei* and *T. cruzi* with consensus PTS1 and PTS2 sequences being identified (Oppendoes and Szikora, 2006) (Table 16).

The HMGR enzyme was predicted at one point to be glycosomal, in both *T. brucei* and *L. major*. These conclusions were based on sub-cellular fractionation experiments, but which can be prone cross-contamination of target organelles with other sub-cellular compartments. However, a similar pattern of peroxisomal localisation has been predicted in both these organism for mevalonate kinase (Carrero-Lerida et al., 2009). After this step in the pathway, for both *L. major* and *T. brucei* all other reactions take place outside of peroxisomes, with the further enzymes being predicted as either cytosolic or ER localised. In contrast, the enzymes mevalonate diphosphate decarboxylase and isopentenyl-diphosphate delta isomerase were predicted to be cytosolic for the all of the other organisms were a prediction was made.

**Table 15. Localisation of the enzymes that function as part of the MVA pathway in *T. brucei***

Localisations obtained from TrypTag.org (Dean et al., 2017) based upon fluorescent microscopy of *T. brucei* amoeba using N-terminal and C-terminal fluorescent tagging proteins. NT\* indicates the amoeba imaged were tagged at the N-terminus of the protein in question, CT\* refers to tagging occurring at the C-terminus.

<b>Enzymes in the Mevalonate Pathway</b>	<b>Localisation within <i>T. brucei</i></b>
<b>HMG-CoA synthase</b>	NT*- Cytoplasm CT*- Mitochondrion
<b>HMG-CoA reductase</b>	CT*-Kinetosome
<b>Mevalonate kinase</b>	NT*-Glycosome CT*- Cytoplasm
<b>Phosphomevalonate kinase</b>	-
<b>Mevalonate diphosphate decarboxylase</b>	NT*- Cytoplasm CT*- Cytoplasm
<b>Squalene synthase</b>	NT*- Cytoplasm
<b>Squalene epoxidase</b>	NT*- Cytoplasm CT*- ER

**Table 16. Presence of PTS1 or PTS2 in enzymes in the mevalonate pathway of isoprenoid biosynthesis.**

Presence of PTS1 or PTS2 taken from Opperdoes and Szikora (2006) survey of 37,080 proteins from *L. major*, *T. brucei* and *T. cruzi* for presence of consensus PTS sequences.

Enzymes in the Mevalonate Pathway	Peroxisomal Targeting Sequence Present		
	<i>L. major</i>	<i>T. brucei</i>	<i>T. cruzi</i>
HMG-CoA reductase	-	-	-
Mevalonate kinase	PTS1, PTS2	PTS1, PTS2	PTS1, PTS2
Phosphomevalonate kinase	-	-	-
Mevalonate diphosphate decarboxylase	PTS1	PTS1	PTS1
Isopentenyl-diphosphate delta-isomerase	PTS1	-	PTS1
Squalene synthase	PTS1	-	-

## Discussion

### Aims and Objectives

To investigate the intracellular localisation of two different isoforms of HMGR found within *D. discoideum*. Another major aim is to carry out an analysis of the presence of peroxisomal targeting signals and transmembrane spanning regions found with the two isoforms and other enzymes in the MVA pathway of *D. discoideum* and other eukaryotes. Finally, an analysis of the general organisation of the early reactions of the MVA pathway and the presence of lack of enzymes within a diverse range of protists.

### Preparation of the *Dictyostelium discoideum* plasmids

The plasmid used for transformation of the *D. discoideum* cells was *pDM1039*, a plasmid created by Veltman et al. comprising; an *Escherichia coli* replication region, replication regions, selection sequences and a protein expression region in *Dictyostelium*. The restriction sites chosen for both the inserted *DdHMGR* isoforms and the *D. discoideum* plasmid were for the restriction enzymes *Bgl*III and *Spe*I. Initially, this process was met with multiple setbacks with a range of differing PCR protocols and optimisation attempts providing no solution to the problem of HMGR gene amplifications. This highlights the difficulty with AT-rich genomes in regard to PCR (Dhatterwal et al., 2017). These attempts were abandoned in favour of obtaining synthetic genes, which also allowed insertion of restriction sites for controlled directional insertion of each of the HMGR isoforms into the expression

plasmid. After further introduction of these enzymes to the plasmid-insert vector (and then subsequently analysing these samples via electrophoresis), two bands can be seen on the agarose gel (Fig. 4). These two bands correspond to the separate *DdHMGR* inserts and the plasmid. This shows that the correct insert has been ligated into the plasmid. After subsequent incubation with a range of restriction enzymes with known restriction sites in both *DdHMGR* isoforms and the plasmid, the fragments all show expected sizes based on prediction for the two *DdHMGR* isoforms and the plasmid (Fig 4. Table 12.).

### Localisation of Two Different Isoforms of HMGR in *Dictyostelium discoideum* Based Upon Fluorescence Microscopy

Based upon the fluorescence microscopy evidence obtained in this study *DdHMGA* appears to possess a degree of peroxisomal localisation, where FITC tagged enzyme co-localised with an RFP attached to a PTS1 sequence (Fig. 7a). In contrast to this, *DdHMGB* appears to not possess co-localisation with the PTS1-RFP (Fig. 7b).

Analysis of the amino acid sequences of *DdHMGA* shows no obvious N-terminal PTS2 sequence, but a possible weak PTS1 sequence. Nine amino acids from the C-terminus of *DdHMGA* the peptide KSHL is present. The tripeptide SHL is similar to the classic SKL PTS1, with one positively charged amino acid traded for another. SHL has been shown to be a putative peroxisomal targeting domain in plants, enhanced by the presence of a basic amino acids such as two Arginine residues. (Ma and Reumann, 2008). It was also shown that as only one basic amino acid is present adjacent to this tripeptide, it may be too weak to serve as an effective peroxisomal targeting signal, however this may not eliminate the possibility of a peroxisomal localisation for *DdHMGA*. Interestingly, the same study also found that a PTS1 can be modulated by the presence of several acidic residues resulting in a cytosolic destination for the protein.

There has been some evidence of protein localisation within the peroxisome in absence of a PTS1 or PTS2 signal. This has been shown to occur in a number of ways. It has been shown that the enzyme  $\Delta 3, \Delta 2$ -enoyl-CoA isomerase (Eci1p) can form a complex with a PTS1 containing enzyme with 50% identity (Dci1p) in order to induce transport into the peroxisome via Pex5p (Yang et al., 2001).

The *S. cerevisiae* peroxisome protein acyl-CoA oxidase (Pox1p) lacks a PTS1 yet still interacts with the peroxisome import protein Pex5p. It does this via an internal binding site, in contrast to the usual C-

terminal PTS1 site (Klein et al., 2002). Other proteins have shown the ability to localise within peroxisomes without the presence of PTS1 or PTS2 sequences.

Alcohol oxidase in methylotrophic yeast normally contains a PTS1 but when artificially removed, the enzyme continues to associate with Pex5p and localise within peroxisomes as shown via immunocytochemical analysis which shows similar punctate appearance to the immunofluorescent analysis for DdHMGA (Fig. 7a). Interestingly, when the PTS1 binding region of Pex5p is also removed, alcohol transport is unaffected suggesting an alternate binding mechanism for this protein (van der Klei and Veenhuis, 2006).

Following the input of *DdHMGR* sequences into three different mitochondrial targeting prediction programs (Mitoprot II, Predotar and TargetP 1.1), *DdHMGA* in *D. discoideum* showed a low probability of mitochondrial export based upon all the software used (Table 14). Based upon this, it is unlikely that *DdHMGA* is a mitochondrial enzyme. The punctate appearance of the *DdHMGA* seen using immunofluorescent analysis (Fig. 7a) can then not be explained by a mitochondrial localisation and further suggests a peroxisomal localisation. Comparable patterns of punctate distribution were seen in a study imaging enzymes with known PTS1 and PTS2 sequences (Nuttall et al., 2012), further supporting *DdHMGA* peroxisomal localisation.

The resulting probability given for the mitochondrial targeting prediction software for *DdHMGB* in *D. discoideum* showed moderate to high for all software use (Table 14). The sequence for *DdHMGB* in *D. discoideum* also contains no apparent PTS2 sequence. The sequence does however possess the peptide RSHL, 27 amino acids from the C-terminus. In the study in which the tripeptide nature of the PTS1, the PTS1 signal was found to be the last 3 residues in the 550 amino acid long firefly luciferase protein (Gould et al., 1989). This is similar to the possible PTS1 in *DdHMGA*, however it is further from the C-terminus of the peptide which may interfere with its function as a PTS1. *DdHMGB* also shows a punctate signal when analysed using immunofluorescence (Fig.7b), which in combination with the software prediction, weak possible PTS1 and lack of obvious co-localisation with the PTS1-RFP suggests a mitochondrial localisation for the HMGR B isoform in *D. discoideum*.

### *Leishmania tarentolae* cells transformed with a GFP-tagged squalene epoxidase enzyme

Construction of the GFP-tagging amplicon was fraught with many problems due to both the two-step nature of the PCR protocol required to create the tagging amplicon and the required size of

homologous regions for effect recombination. In yeast, two 80bp length homologous regions are required for high efficiency gene recombination (Hua et al., 1997). However, this is not the case for *Leishmania* which was shown to require far larger regions of homology, greater than 200bp in length (Dean et al., 2015). During this study, the two flanking homologous regions were 393bp in length. When combined with the two-step PCR protocol utilised to produce the tagging amplicon (Dean et al., 2015) this created significant optimisation problems during the study.

Considering this, the immunofluorescence images obtained (Fig. 11) do not give much information regarding the localisation of the GFP-tagged squalene epoxidase. However, there is a range of data including PTS1 and PTS2 sequence prediction and immunofluorescence assays on the enzymes within the mevalonate pathway including squalene epoxidase for *Leishmania major* a cousin of *L. tarentolae*. The study performed by Opperdoes and Szikora in 2006 analysed PTS1 and PTS2 sequences in 37,080 genes and the results of this prediction for MVA pathway enzymes were used and displayed in Table 15.

Two studies used a combination of immunofluorescence and digitonin permeabilization to determine mitochondrial localisations for both HMGR (Peña-Diaz et al., 2004) and HMGS (Carrero-Lerida et al., 2009) in *L. major* and *T. brucei*. Mevalonate kinase (MK), another MVA pathway enzyme was found to be a glycosomal enzyme in both *L. major* and *T. brucei* (Carrero-Lerida et al., 2009). Interestingly, these localisations differ from those based upon sequence prediction displayed in Table 15 and those obtained by Opperdoes and Szikora in 2006. If the mitochondrial localisations are correct, this creates the possibility for multiple differing localisation across groups of Eukaryotes. Mammalian cells possess a peroxisomal isoform and an ER-localised isoform, *Leishmania* possesses a mitochondrial HMGR and *Dictyostelium* possess a peroxisomal isoform. Alternately, MK was found to be glycosomal and is similarly peroxisomally localised in mammalian cells (Kovacs et al., 2007).

Studies performed on *Leishmania* using labelled substrates showed interesting use of precursors for isoprenoid biosynthesis (Ginger et al., 2001 and Ginger et al., 2000). These studies showed the usage of both leucine and acetate as carbon sources for the production of sterols. The enzyme 3-methylglutaconyl-coenzyme-A hydratase functions in the catabolism of the amino acid leucine into HMG-CoA and was found to be located at the mitochondrial matrix (de Lima Stein et al., 2017). This HMG-CoA could then be utilised by the HMGR enzyme within the same mitochondrial compartment. The utilisation of acetate agrees with the mitochondrial presence of HMGS, as acetate can be used in the formation of acetyl-CoA which is utilised by HMGS to create HMG-CoA. Acetyl-CoA is regularly transported into the mitochondria (Shi and Tu, 2015). This combination of enzymes and substrates

localised within the mitochondria could allow more energy efficient isoprenoid biosynthesis, requiring less membrane transport.

The importance on the positioning of the GFP in relation to the target protein can have a significant impact on the protein, with possible disruption to the normal localisation of the protein. This is highlighted in the localisation data from Table 15 in *T. brucei* (Dean et al., 2017), where different placements of the GFP at either N- or C-terminus resulted in different localisations for the same enzyme. This is most likely due to a localisation signal that requires positioning at the peptide terminus, being blocked by the GFP. Examples include the N-terminal signal sequence required for Signal Recognition Particle-related ER protein transport (Akopian et al., 2013) and the C-terminal PTS1. This would also be the case for the C-terminal GFP tag for *LtSqE* used in this study, requiring further tagging with an N-terminal GFP to validate possible localisation results and fully elucidate any leader sequences required for correct targeting.

## The Importance of Early-Branching Eukaryotes in the Whole Eukaryotic Phylogenetic Tree

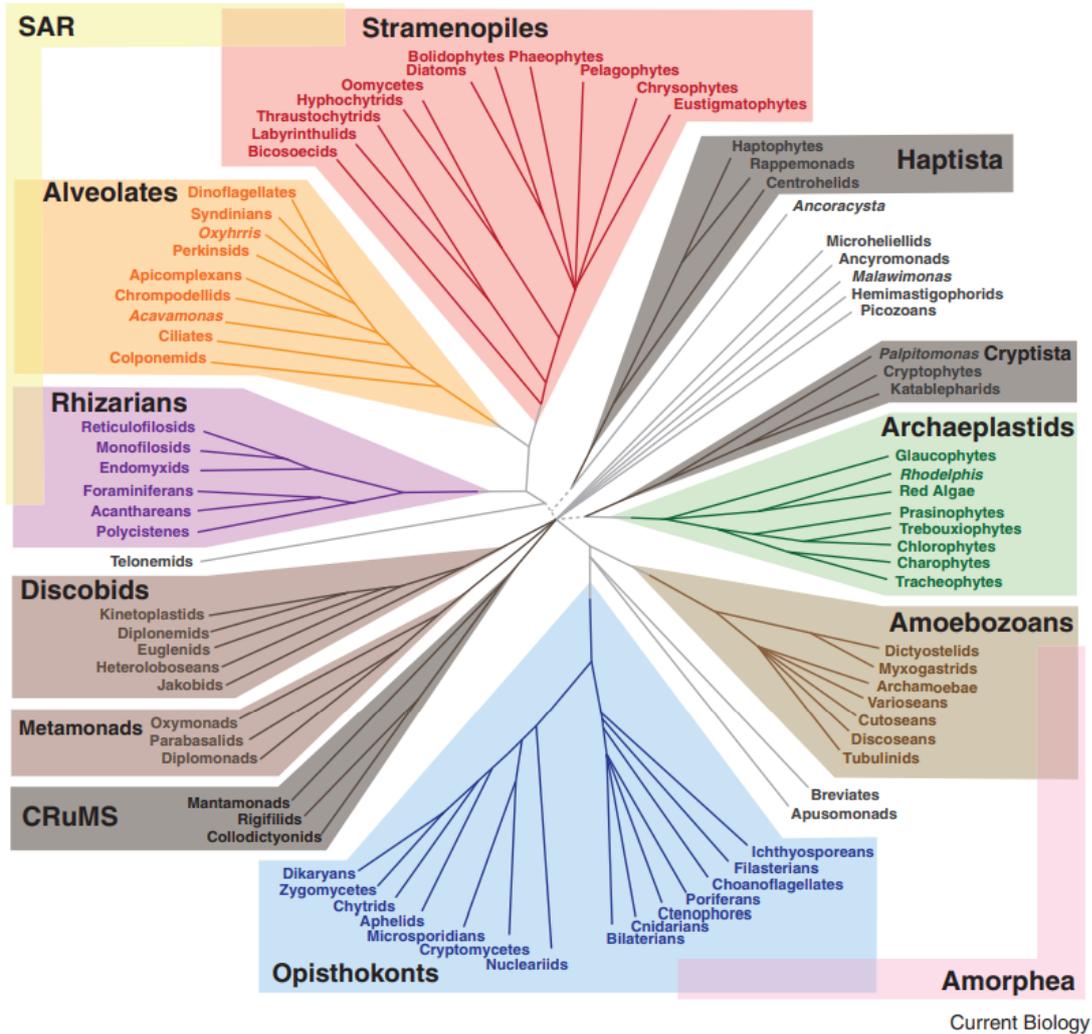
The focus of this study was the determination of possible alternate localisations of enzyme isoforms involved in a heavily conserved biosynthetic pathway in eukaryotes. Any interesting localisations within either *D. discoideum* or *L. tarentolae* would be made more significant due to the positioning of both of these organisms in the eukaryote phylogenetic tree. Any difference to the localisation of mevalonate pathway enzymes could elucidate the nature of the original acquisition of the MVA pathway, adaptations to confer enhanced pathogenicity or additional functions of the enzyme.

*D. discoideum* belongs to the Amoebozoan group, which represents one of the earliest branching lineages from the last common ancestor of all eukaryotes (Eichinger et al., 2005). This position particularly highlights any adaptations to enzyme pathways, especially those as conserved as the MVA pathway of isoprenoid biosynthesis. This early separation from organisms that would become ancestors to all opisthokonts allows for comparison and identification of genetic divergence between these two lineages. *D. discoideum* represents a particularly interesting Amoebozoan example based upon its distinct lifestyle and extensive information regarding its genome. As mentioned, the MVA pathway of isoprenoid is well conserved being found with slight variance in archaea, eukaryotes and some bacteria which raises the question of where the pathway originated and how it is present in all these lineages. Interestingly, one of the major variations between these lineages regarding the MVA pathway is the two classes of HMGR with class 1 present largely in

eukaryotes and archaea, and class 2 largely present in bacteria. This variation leads to three separate hypotheses for the origin of the MVA pathway. The first being evolution within the last common ancestor of eukaryotes and archaea followed by horizontal gene transfer to the various bacterial lineages that possess the MVA pathway. The second involving the same origin of the pathway, but a horizontal gene transfer event to the ancestor of all bacterial lineages followed by gene loss events for the ancestors of bacterial lineages that now do not possess the MVA pathway. The final hypothesis is the evolution of the MVA pathway in the cenancestor of archaea, eukaryotes and bacteria with gene loss event and adaptations occurring later accounting for the variations across the families of life (Lombard and Moreira, 2011).

*Leishmania tarentolae* belongs to the Trypanosomatids, themselves belonging to the monophyletic Euglenozoa. The Euglenozoa group despite being a monophyletic group possess species with a widely varying niche and lifestyle, for example the photosynthetic euglenids and the parasitic kinetoplastids (to which *L. tarentolae* belongs) (Zakrys et al., 2017). The Excavata were a super group to which the Euglenozoa were thought to belong (Hapl et al., 2009) and has since been reconsidered and split into three different groups; Discoba (including the kinetoplastids), metamonads (including malawimonas) (Heiss et al., 2018) and the remaining CRuMS found to be related via molecular taxonomies but possessing varying morphologies (Keeling and Burki, 2019).

Both the Discoba and Amoebozoans can then be seen to represent early divergences in the eukaryotic tree of life. Importantly, the last common ancestor of *D. discoideum* and *L. tarentolae* was also the last common ancestor of all extant eukaryote lineages and the root of all eukaryotic life on Earth. This particularly highlights any similarities these organisms possess in regard to the localisation of mevalonate pathway enzymes, in contrast to other eukaryotes. The question can be raised as to the origin of the peroxisomal isoforms of HMGR and squalene epoxidase; obtained via the last common ancestor of all eukaryotes and adapted to a dual ER and peroxisomal localisation by the Opisthokont ancestor or adapted to solely peroxisomal in unrelated events in both the Kinetoplastids and *Dictyostelium*.



**Figure 12. Phylogenetic tree of eukaryotes.**

A schematic tree of life for eukaryotes based on a phylogenomic data obtained from a range of studies, supported by morphological information (Keeling and Burki, 2019).

A possible peroxisomal localisation for the HMGR isoform *DdHMGA* is very interesting when put into the context of previously known peroxisomal MVA pathway enzymes in other eukaryotes. Two studies that investigated human (Kovacs et al., 2007) and both human and Chinese hamster ovarian cells (Olivier et al., 2000) found the MVA pathway enzymes to be spread across various organelles (mitochondria, ER, peroxisomes and the cytosol) with some disagreements regarding several enzyme localisations. However, both studies have a feature in common; that throughout the MVA pathway, precursor molecules were moved across organelle membranes multiple times in the pre-squalene steps of the process. Importantly the first study (Kovacs et al., 2007) found the HMGR enzyme to be dual localised at the ER and the peroxisomes.

Farnesyl diphosphate synthase (FDPS) is a mevalonate pathway enzyme downstream of the DMAPP precursor to squalene. This enzyme was found to contain a PTS2 and was localised at the peroxisomes in *D. discoideum* via immunofluorescence studies (Nuttall et al., 2012). Interestingly, FDPS-GFP constructs lacking the complete PTS2 were shown to also localise at the peroxisomes with a reduced efficiency suggesting alternate mechanisms similar to those mentioned earlier.

### Additional Work and Comments

Following on from the research carried out in this study, there is much additional research to be done on the topic in order to further elucidate the findings. Critically, an attempt at greater optimisation of the various methods already utilised in the study would give more confidence to current data. A clearer signal could be obtained for the ER-Tracker™ Red study, with a repeat of the cell incubation stage with the dye. This could give a clearer indication of non-ER localised *Dd*HMGR.

Although it can be determined through immunoblotting and immunofluorescent whether the isoforms of HMGR are being translated, there is no insight into the relative levels of transcription. Further work could include qPCR investigation of *D. discoideum* cells at different time points and in different growth conditions in order to monitor the levels of transcription of the HMGR isoforms.

The green fluorescence in GFP-tagged squalene epoxidase *Leishmania tarentolae* appeared greater than in the wild-type cells (Fig. 11) which is an indication of correctly transfected cells. However, using the immunofluorescence imagery, the signal is not clear enough to accurately determine the localisation of the GFP-tagged SqE. Despite the successful replication of an ORF for MK and a UTR for IVDH, due to time constraints additional PCR reactions were not carried out to obtain the remaining ORF and UTR sequences. MK is a known peroxisomal enzyme and IVDH is a known mitochondrial matrix enzyme, thus allowing *L. tarentolae* cell transfected with tagging amplicons for these genes to serve as positive indicators for a peroxisomal or mitochondrial localisation pattern.

Cellular fractionation studies would provide further data concerning subcellular localisation of the enzymes of interest in this study. There is a range of potential commercial organellar markers available that allow determination of successful isolation of the desired organelle, such as Anti-Calreticulin antibody [EPR3924] - ER Marker, Anti-ATP5A antibody [15H4C4] - Mitochondrial Marker (ab14748) and Recombinant Anti-Catalase antibody [EP1929Y] - Peroxisome Marker (ab76024) (abcam). Mechanical fractionation of trypanosomes can be difficult as the amoebae possess a layer of microtubules beneath the plasma membrane, preventing easy disruption of the plasma

membrane without enough force to disrupt other organelles and isolation of the complete mitochondrion is difficult due to the size and branching structure (de Souza and Cunha-e-Silva, 2003). A solution to these problems is the use of digitonin for selective permeabilization of plasma membranes based upon varying concentrations, which can allow isolation of the mitochondrion (Lai et al., 2012) and permeabilization of peroxisomes (Rondón-Mercado et al., 2017).

## References

- Aebi, M. (2013). N-linked protein glycosylation in the ER. *Biochim Biophys Acta*, 1833(11), 2430-2437. doi:10.1016/j.bbamcr.2013.04.001
- Ahn, C. S., & Pai, H. S. (2008). Physiological function of *IspE*, a plastid MEP pathway gene for isoprenoid biosynthesis, in organelle biogenesis and cell morphogenesis in *Nicotiana benthamiana*. *Plant Mol Biol*, 66(5), 503-517. doi:10.1007/s11103-007-9286-0
- Akhoundi, M., Kuhls, K., Cannet, A., Votýpka, J., Marty, P., Delaunay, P., & Sereno, D. (2016). A Historical Overview of the Classification, Evolution, and Dispersion of Leishmania Parasites and Sandflies. *PLoS Negl Trop Dis*, 10(3). doi:10.1371/journal.pntd.0004349
- Akopian, D., Shen, K., Zhang, X., & Shan, S. (2013). Signal Recognition Particle: An essential protein targeting machine. *Annu Rev Biochem*, 82, 693-721. doi:10.1146/annurev-biochem-072711-164732
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol*, 215(3), 403-410. doi:10.1016/s0022-2836(05)80360-2
- Antolín-Llovera, M., Leivar, P., Arró, M., Ferrer, A., Boronat, A., & Campos, N. (2011). Modulation of plant HMG-CoA reductase by protein phosphatase 2A: Positive and negative control at a key node of metabolism. *Plant Signal Behav*, 6(8), 1127-1131. doi:10.4161/psb.6.8.16363
- Aranovich, A., Hua, R., Rutenberg, A. D., & Kim, P. K. (2014). PEX16 contributes to peroxisome maintenance by constantly trafficking PEX3 via the ER. *J Cell Sci*, 127(17), 3675-3686. doi:10.1242/jcs.146282
- Aurrecochea, C., Barreto, A., Basenko, E. Y., Brestelli, J., Brunk, B. P., Cade, S., . . . Zheng, J. (2017). EuPathDB: the eukaryotic pathogen genomics database resource. *Nucleic Acids Res*, 45(Database issue), D581-591. doi:10.1093/nar/gkw1105
- Ayte, J., Gil-Gomez, G., Haro, D., Marrero, P. F., & Hegardt, F. G. (1990). Rat mitochondrial and cytosolic 3-hydroxy-3-methylglutaryl-CoA synthases are encoded by two different genes. *Proc Natl Acad Sci U S A*, 87(10), 3874-3878.
- Basyoni, M. M. A., Fouad, S. A., Amer, M. F., Amer, A. F., & Ismail, D. I. (2018). Atorvastatin: In-Vivo Synergy with Metronidazole as Anti-Blastocystis Therapy. *Korean J Parasitol*, 56(2), 105-112. doi:10.3347/kjp.2018.56.2.105
- Bauer, S., & Morris, M. T. (2017). Glycosome biogenesis in trypanosomes and the de novo dilemma. *PLoS Negl Trop Dis*, 11(4). doi:10.1371/journal.pntd.0005333

- Belin, B. J., Busset, N., Giraud, E., Molinaro, A., Silipo, A., & Newman, D. K. (2018). Hopanoid lipids: from membranes to plant–bacteria interactions. *Nat Rev Microbiol*, 16(5), 304-315. doi:10.1038/nrmicro.2017.173
- Benz, C., Stribrna, E., Hashimi, H., & Lukes, J. (2017). Dynamin-like proteins in *Trypanosoma brucei*: A division of labour between two paralogs? *PLoS One*, 12(5), e0177200. doi:10.1371/journal.pone.0177200
- Bochar, D. A., Stauffacher, C. V., & Rodwell, V. W. (1999). Sequence comparisons reveal two classes of 3-hydroxy-3-methylglutaryl coenzyme A reductase. *Mol Genet Metab*, 66(2), 122-127. doi:10.1006/mgme.1998.2786
- Brocard, C., & Hartig, A. (2006). Peroxisome targeting signal 1: is it really a simple tripeptide? *Biochim Biophys Acta*, 1763(12), 1565-1573. doi:10.1016/j.bbamcr.2006.08.022
- Bröker, J. N., Müller, B., van Deenen, N., Prüfer, D., & Schulze Gronover, C. (2018). Upregulating the mevalonate pathway and repressing sterol synthesis in *Saccharomyces cerevisiae* enhances the production of triterpenes. *Appl Microbiol Biotechnol*, 102(16), 6923-6934. doi:10.1007/s00253-018-9154-7
- Burki, F. (2014). The Eukaryotic Tree of Life from a Global Phylogenomic Perspective. *Cold Spring Harb Perspect Biol*, 6(5). doi:10.1101/cshperspect.a016147
- Carrero-Lerida, J., Perez-Moreno, G., Castillo-Acosta, V. M., Ruiz-Perez, L. M., & Gonzalez-Pacanowska, D. (2009). Intracellular location of the early steps of the isoprenoid biosynthetic pathway in the trypanosomatids *Leishmania major* and *Trypanosoma brucei*. *Int J Parasitol*, 39(3), 307-314. doi:10.1016/j.ijpara.2008.08.012
- Cavalier-Smith, T., Fiore-Donno, A. M., Chao, E., Kudryavtsev, A., Berney, C., Snell, E. A., & Lewis, R. (2015). Multigene phylogeny resolves deep branching of Amoebozoa. *Mol Phylogenet Evol*, 83, 293-304. doi:10.1016/j.ympev.2014.08.011
- Claros, M. G., & Vincens, P. (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem*, 241(3), 779-786.
- Clomburg, J. M., Qian, S., Tan, Z., Cheong, S., & Gonzalez, R. (2019). The isoprenoid alcohol pathway, a synthetic route for isoprenoid biosynthesis. *Proc Natl Acad Sci U S A*, 116(26), 12810-12815. doi:10.1073/pnas.1821004116
- Clouse, S. D. (2011). Brassinosteroids. *Arabidopsis Book*, 9. doi:10.1199/tab.0151

- Cotter, D. G., Ercal, B., Huang, X., Leid, J. M., d'Avignon, D. A., Graham, M. J., . . . Crawford, P. A. (2014). Ketogenesis prevents diet-induced fatty liver injury and hyperglycemia. *J Clin Invest*, 124(12), 5175-5190. doi:10.1172/jci76388
- Cross, L. L., Ebeed, H. T., & Baker, A. (2016). Peroxisome biogenesis, protein targeting mechanisms and PEX gene functions in plants. *Biochim Biophys Acta*, 1863(5), 850-862. doi:10.1016/j.bbamcr.2015.09.027
- de Duve, C. (1982). Peroxisomes and related particles in historical perspective. *Ann N Y Acad Sci*, 386, 1-4. doi:10.1111/j.1749-6632.1982.tb21402.x
- de Lima Stein, M. L., Icimoto, M. Y., de Castro Levatti, E. V., Oliveira, V., Straus, A. H., & Schenkman, S. (2017). Characterization and role of the 3-methylglutaconyl coenzyme A hydratase in *Trypanosoma brucei*. *Mol Biochem Parasitol*, 214, 36-46. doi:10.1016/j.molbiopara.2017.03.007
- de Souza, W., & da Cunha-e-Silva, N. L. (2003). Cell fractionation of parasitic protozoa: a review. *Mem Inst Oswaldo Cruz*, 98(2), 151-170. doi:10.1590/s0074-02762003000200001
- de Souza, W., & Rodrigues, J. C. (2009). Sterol Biosynthesis Pathway as Target for Anti-trypanosomatid Drugs. *Interdiscip Perspect Infect Dis*, 2009, 642502. doi:10.1155/2009/642502
- Dean, S., Sunter, J., Wheeler, R. J., Hodkinson, I., Gluenz, E., & Gull, K. (2015). A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol*, 5(1). doi:10.1098/rsob.140197. (Accession No. 25567099)
- Dean, S., Sunter, J. D., & Wheeler, R. J. (2017). TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. *Trends Parasitol*, 33(2), 80-82. doi:10.1016/j.pt.2016.10.009
- Dellas, N., Thomas, S. T., Manning, G., & Noel, J. P. (2013). Discovery of a metabolic alternative to the classical mevalonate pathway. *eLife*, 2. doi:10.7554/eLife.00672
- Dhatterwal, P., Mehrotra, S., & Mehrotra, R. (2017). Optimization of PCR conditions for amplifying an AT-rich amino acid transporter promoter sequence with high number of tandem repeats from *Arabidopsis thaliana*. *BMC Res Notes*, 10. doi:10.1186/s13104-017-2982-1
- Dodt, G., & Gould, S. J. (1996). Multiple PEX genes are required for proper subcellular distribution and stability of Pex5p, the PTS1 receptor: evidence that PTS1 protein import is mediated by a cycling receptor. *J Cell Biol*, 135(6 Pt 2), 1763-1774. doi:10.1083/jcb.135.6.1763

Edwards, P. A., & Ericsson, J. (1999). Sterols and isoprenoids: signaling molecules derived from the cholesterol biosynthetic pathway. *Annu Rev Biochem*, 68, 157-185.

doi:10.1146/annurev.biochem.68.1.157

Eichinger, L., Pachebat, J., Glöckner, G., Rajandream, M. A., Sucgang, R., Berriman, M., . . . Kuspa, A. (2005). The genome of the social amoeba *Dictyostelium discoideum*. *Nature*, 435(7038), 43-57.

doi:10.1038/nature03481

Eitzen, G. A., Szilard, R. K., & Rachubinski, R. A. (1997). Enlarged peroxisomes are present in oleic acid-grown *Yarrowia lipolytica* overexpressing the PEX16 gene encoding an intraperoxisomal peripheral membrane peroxin. *J Cell Biol*, 137(6), 1265-1278. doi:10.1083/jcb.137.6.1265

Emanuelsson, O., Brunak, S., von Heijne, G., & Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc*, 2(4), 953-971. doi:10.1038/nprot.2007.131

Emanuelsson, O., Nielsen, H., Brunak, S., & von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol*, 300(4), 1005-1016.

doi:10.1006/jmbi.2000.3903

Erdmann, R., & Kunau, W. H. (1992). A genetic approach to the biogenesis of peroxisomes in the yeast *Saccharomyces cerevisiae*. *Cell Biochem Funct*, 10(3), 167-174. doi:10.1002/cbf.290100306

Escalante, R., & Vicente, J. J. (2000). *Dictyostelium discoideum*: a model system for differentiation and patterning. *Int J Dev Biol*, 44(8), 819-835.

Friesen, J. A., & Rodwell, V. W. (2004). The 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductases. *Genome Biol*, 5(11), 248. doi:10.1186/gb-2004-5-11-248

Fritz, R., Bol, J., Hebling, U., Angermüller, S., Volkl, A., Fahimi, H. D., & Müller, S. (2007).

Compartment-dependent management of H<sub>2</sub>O<sub>2</sub> by peroxisomes. *Free Radic Biol Med*, 42(7), 1119-1129. doi:10.1016/j.freeradbiomed.2007.01.014

Fujiki, Y., Okumoto, K., Mukai, S., Honsho, M., & Tamura, S. (2014). Peroxisome biogenesis in mammalian cells. *Front Physiol*, 5. doi:10.3389/fphys.2014.00307

Gabalton, T., Ginger, M. L., & Michels, P. A. (2016). Peroxisomes in parasitic protists. *Mol Biochem Parasitol*, 209(1-2), 35-45. doi:10.1016/j.molbiopara.2016.02.005

Gatto, G. J., Jr., Geisbrecht, B. V., Gould, S. J., & Berg, J. M. (2000). Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nat Struct Biol*, 7(12), 1091-1095.

doi:10.1038/81930

- Gatto, G. J., Jr., Maynard, E. L., Guerrero, A. L., Geisbrecht, B. V., Gould, S. J., & Berg, J. M. (2003). Correlating structure and affinity for PEX5:PTS1 complexes. *Biochemistry*, 42(6), 1660-1666. doi:10.1021/bi027034z
- Geuze, H. J., Murk, J. L., Stroobants, A. K., Griffith, J. M., Kleijmeer, M. J., Koster, A. J., . . . Tabak, H. F. (2003). Involvement of the endoplasmic reticulum in peroxisome formation. *Mol Biol Cell*, 14(7), 2900-2907. doi:10.1091/mbc.e02-11-0734
- Ginger, M. L., Chance, M. L., & Goad, L. J. (1999). Elucidation of carbon sources used for the biosynthesis of fatty acids and sterols in the trypanosomatid *Leishmania mexicana*. *Biochem J*, 342(Pt 2), 397-405.
- Ginger, M. L., Chance, M. L., Sadler, I. H., & Goad, L. J. (2001). The biosynthetic incorporation of the intact leucine skeleton into sterol by the trypanosomatid *Leishmania mexicana*. *J Biol Chem*, 276(15), 11674-11682. doi:10.1074/jbc.M006850200
- Ginger, M. L., McFadden, G. I., & Michels, P. A. M. (2010). Rewiring and regulation of cross-compartmentalized metabolism in protists. *Philos Trans R Soc Lond B Biol Sci*, 365(1541), 831-845. doi:10.1098/rstb.2009.0259
- Ginger, M. L., Prescott, M. C., Reynolds, D. G., Chance, M. L., & Goad, L. J. (2000). Utilization of leucine and acetate as carbon sources for sterol and fatty acid biosynthesis by Old and New World *Leishmania* species, *Endotrypanum monterogeei* and *Trypanosoma cruzi*. *Eur J Biochem*, 267(9), 2555-2566. doi:10.1046/j.1432-1327.2000.01261.x
- Gophna, U., Thompson, J. R., Boucher, Y., & Doolittle, W. F. (2006). Complex histories of genes encoding 3-hydroxy-3-methylglutaryl-CoenzymeA reductase. *Mol Biol Evol*, 23(1), 168-178. doi:10.1093/molbev/msj019
- Gould, S. G., Keller, G. A., & Subramani, S. (1987). Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. *J Cell Biol*, 105(6 Pt 2), 2923-2931. doi:10.1083/jcb.105.6.2923
- Gould, S. J., Keller, G. A., Hosken, N., Wilkinson, J., & Subramani, S. (1989). A conserved tripeptide sorts proteins to peroxisomes. *J Cell Biol*, 108(5), 1657-1664. doi:10.1083/jcb.108.5.1657
- Gualdron-Lopez, M., Brennand, A., Hannaert, V., Quinones, W., Caceres, A. J., Bringaud, F., . . . Michels, P. A. (2012). When, how and why glycolysis became compartmentalised in the Kinetoplastea. A new look at an ancient organelle. *Int J Parasitol*, 42(1), 1-20. doi:10.1016/j.ijpara.2011.10.007

- Güther, M. S., Urbaniak, M. D., Tavendale, A., Prescott, A., & Ferguson, M. A. J. (2014). High-Confidence Glycosome Proteome for Procyclic Form *Trypanosoma brucei* by Epitope-Tag Organelle Enrichment and SILAC Proteomics. *J Proteome Res*, 13(6), 2796-2806. doi:10.1021/pr401209w
- Haines, T. H. (2001). Do sterols reduce proton and sodium leaks through lipid bilayers? *Prog Lipid Res*, 40(4), 299-324. doi:10.1016/s0163-7827(01)00009-1
- Hampl, V., Hug, L., Leigh, J. W., Dacks, J. B., Lang, B. F., Simpson, A. G., & Roger, A. J. (2009). Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". *Proc Natl Acad Sci U S A*, 106(10), 3859-3864. doi:10.1073/pnas.0807880106
- Harper, C. C., Berg, J. M., & Gould, S. J. (2003). PEX5 binds the PTS1 independently of Hsp70 and the peroxin PEX12. *J Biol Chem*, 278(10), 7897-7901. doi:10.1074/jbc.M206651200
- Hayashi, M., Toriyama, K., Kondo, M., Kato, A., Mano, S., De Bellis, L., . . . Nishimura, M. (2000). Functional transformation of plant peroxisomes. *Cell Biochem Biophys*, 32 Spring, 295-304.
- He, L., He, P., Luo, X., Li, M., Yu, L., Guo, J., . . . Zhao, J. (2018). The MEP pathway in *Babesia orientalis* apicoplast, a potential target for anti-babesiosis drug development. *Parasit Vectors*, 11. doi:10.1186/s13071-018-3038-7
- Heidel, A. J., Lawal, H. M., Felder, M., Schilde, C., Helps, N. R., Tunggal, B., . . . Glöckner, G. (2011). Phylogeny-wide analysis of social amoeba genomes highlights ancient origins for complex intercellular communication. *Genome Res*, 21(11), 1882-1891. doi:10.1101/gr.121137.111
- Heiss, A. A., Kolisko, M., Ekelund, F., Brown, M. W., Roger, A. J., & Simpson, A. G. B. (2018). Combined morphological and phylogenomic re-examination of malawimonads, a critical taxon for inferring the evolutionary history of eukaryotes. *R Soc Open Sci*, 5(4), 171707. doi:10.1098/rsos.171707
- Holst, J. P., Soldin, O. P., Guo, T., & Soldin, S. J. (2004). Steroid hormones: relevance and measurement in the clinical laboratory. *Clin Lab Med*, 24(1), 105-118. doi:10.1016/j.cll.2004.01.004
- Hoshino, Y., & Gaucher, E. A. (2018). On the Origin of Isoprenoid Biosynthesis. *Mol Biol Evol*, 35(9), 2185-2197. doi:10.1093/molbev/msy120
- Hua, S. B., Qiu, M., Chan, E., Zhu, L., & Luo, Y. (1997). Minimum length of sequence homology required for in vivo cloning by homologous recombination in yeast. *Plasmid*, 38(2), 91-96. doi:10.1006/plas.1997.1305

- Huber, R. J. (2016). Using the social amoeba *Dictyostelium* to study the functions of proteins linked to neuronal ceroid lipofuscinosis. *J Biomed Sci*, 23. doi:10.1186/s12929-016-0301-0
- Imlay, L., & Odom, A. R. (2014). Isoprenoid metabolism in apicomplexan parasites. *Curr Clin Microbiol Rep*, 1(3-4), 37-50. doi:10.1007/s40588-014-0006-7
- Isezaki, N., Sekiba, A., Itagaki, S., Nagayama, K., Ochiai, H., & Ohmachi, T. (2015). *Dictyostelium* acetoacetyl-CoA thiolase is a dual-localizing enzyme that localizes to peroxisomes, mitochondria and the cytosol. *Microbiology*, 161(7), 1471-1484. doi:10.1099/mic.0.000102
- Istvan, E. (2003). Statin inhibition of HMG-CoA reductase: a 3-dimensional view. *Atheroscler Suppl*, 4(1), 3-8.
- Istvan, E. S., Palnitkar, M., Buchanan, S. K., & Deisenhofer, J. (2000). Crystal structure of the catalytic portion of human HMG-CoA reductase: insights into regulation of activity and catalysis. *EMBO J*, 19(5), 819-830. doi:10.1093/emboj/19.5.819
- Jansen, R. L. M., & van der Klei, I. J. (2019). The peroxisome biogenesis factors Pex3 and Pex19: multitasking proteins with disputed functions. *FEBS Lett*, 593(5), 457-474. doi:10.1002/1873-3468.13340
- Jilly, R., Khan, N. Z., Aronsson, H., & Schneider, D. (2018). Dynamin-Like Proteins Are Potentially Involved in Membrane Dynamics within Chloroplasts and Cyanobacteria. *Front Plant Sci*, 9, 206. doi:10.3389/fpls.2018.00206
- Kadian, K., Gupta, Y., Singh, H. V., Kempaiah, P., & Rawat, M. (2018). Apicoplast Metabolism: Parasite's Achilles' Heel. *Curr Top Med Chem*, 18(22), 1987-1997. doi:10.2174/1568026619666181130134742
- Karnik, S. K., & Trelease, R. N. (2005). Arabidopsis peroxin 16 coexists at steady state in peroxisomes and endoplasmic reticulum. *Plant Physiol*, 138(4), 1967-1981. doi:10.1104/pp.105.061291
- Karnik, S. K., & 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(7), 1677-1693. doi:10.1093/jxb/erm018
- Kasahara, H., Takei, K., Ueda, N., Hishiyama, S., Yamaya, T., Kamiya, Y., . . . Sakakibara, H. (2004). Distinct isoprenoid origins of cis- and trans-zeatin biosyntheses in Arabidopsis. *J Biol Chem*, 279(14), 14049-14054. doi:10.1074/jbc.M314195200

- Katsuki, H., & Bloch, K. (1967). Studies on the biosynthesis of ergosterol in yeast. Formation of methylated intermediates. *J Biol Chem*, 242(2), 222-227.
- Keeling, P. J., & Burki, F. (2019). Progress towards the Tree of Eukaryotes. *Curr Biol*, 29(16), R808-r817. doi:10.1016/j.cub.2019.07.031
- Kiel, J. A., Veenhuis, M., & van der Klei, I. J. (2006). PEX genes in fungal genomes: common, rare or redundant. *Traffic*, 7(10), 1291-1303. doi:10.1111/j.1600-0854.2006.00479.x
- Kim, P. K., Mullen, R. T., Schumann, U., & 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(4), 521-532. doi:10.1083/jcb.200601036
- Klein, A. T., van den Berg, M., Bottger, G., Tabak, H. F., & Distel, B. (2002). *Saccharomyces cerevisiae* acyl-CoA oxidase follows a novel, non-PTS1, import pathway into peroxisomes that is dependent on Pex5p. *J Biol Chem*, 277(28), 25011-25019. doi:10.1074/jbc.M203254200
- Koch, A., Yoon, Y., Bonekamp, N. A., McNiven, M. A., & Schrader, M. (2005). A Role for Fis1 in Both Mitochondrial and Peroxisomal Fission in Mammalian Cells. *Mol Biol Cell*, 16(11), 5077-5086. doi:10.1091/mbc.E05-02-0159
- Kovacs, W. J., Olivier, L. M., & Krisans, S. K. (2002). Central role of peroxisomes in isoprenoid biosynthesis. *Prog Lipid Res*, 41(5), 369-391.
- Kovacs, W. J., Tape, K. N., Shackelford, J. E., Duan, X., Kasumov, T., Kelleher, J. K., . . . Krisans, S. K. (2007). Localization of the pre-squalene segment of the isoprenoid biosynthetic pathway in mammalian peroxisomes. *Histochem Cell Biol*, 127(3), 273-290. doi:10.1007/s00418-006-0254-6
- Kunze, M., Malkani, N., Maurer-Stroh, S., Wiesinger, C., Schmid, J. A., & Berger, J. (2015). Mechanistic Insights into PTS2-mediated Peroxisomal Protein Import: THE CO-RECEPTOR PEX5L DRASTICALLY INCREASES THE INTERACTION STRENGTH BETWEEN THE CARGO PROTEIN AND THE RECEPTOR PEX7\*. *J Biol Chem*, 290(8), 4928-4940. doi:10.1074/jbc.M114.601575
- Kurzchalia, T. V., & Ward, S. (2003). Why do worms need cholesterol? *Nat Cell Biol*, 5(8), 684-688. doi:10.1038/ncb0803-684
- Kuzuyama, T., & Seto, H. (2012). Two distinct pathways for essential metabolic precursors for isoprenoid biosynthesis. *Proc Jpn Acad Ser B Phys Biol Sci*, 88(3), 41-52. doi:10.2183/pjab.88.41
- Lafont, R., & Mathieu, M. (2007). Steroids in aquatic invertebrates. *Ecotoxicology*, 16(1), 109-130. doi:10.1007/s10646-006-0113-1

- Lai, D. H., Bontempi, E. J., & Lukes, J. (2012). Trypanosoma brucei solanesyl-diphosphate synthase localizes to the mitochondrion. *Mol Biochem Parasitol*, 183(2), 189-192. doi:10.1016/j.molbiopara.2012.02.011
- Lange, B. M., Rujan, T., Martin, W., & Croteau, R. (2000). Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc Natl Acad Sci U S A*, 97(24), 13172-13177. doi:10.1073/pnas.240454797
- Lazarow, P. B. (1987). The role of peroxisomes in mammalian cellular metabolism. *J Inherit Metab Dis*, 10 Suppl 1, 11-22. doi:10.1007/bf01812843
- Legakis, J. E., & Terlecky, S. R. (2001). PTS2 protein import into mammalian peroxisomes. *Traffic*, 2(4), 252-260.
- Leivar, P., González, V. M., Castel, S., Trelease, R. N., López-Iglesias, C., Arró, M., . . . Fernández-Busquets, X. (2005). Subcellular Localization of Arabidopsis 3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase1. *Plant Physiol*, 137(1), 57-69. doi:10.1104/pp.104.050245
- Lin, Y., Sun, L., Nguyen, L. V., Rachubinski, R. A., & Goodman, H. M. (1999). The Pex16p homolog SSE1 and storage organelle formation in Arabidopsis seeds. *Science*, 284(5412), 328-330. doi:10.1126/science.284.5412.328
- Loftus, B., Anderson, I., Davies, R., Alsmark, U. C., Samuelson, J., Amedeo, P., . . . Hall, N. (2005). The genome of the protist parasite Entamoeba histolytica. *Nature*, 433(7028), 865-868. doi:10.1038/nature03291
- Lombard, J., & Moreira, D. (2011). Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life. *Mol Biol Evol*, 28(1), 87-99. doi:10.1093/molbev/msq177
- Loomis, W. F. (2015). Genetic control of morphogenesis in Dictyostelium. *Dev Biol*, 402(2), 146-161. doi:10.1016/j.ydbio.2015.03.016
- Lynen, F. (1967). Biosynthetic pathways from acetate to natural products. *Pure Appl Chem*, 14(1), 137-167. doi:10.1351/pac196714010137
- Ma, C., & Reumann, S. (2008). Improved prediction of peroxisomal PTS1 proteins from genome sequences based on experimental subcellular targeting analyses as exemplified for protein kinases from Arabidopsis. *J Exp Bot*, 59(13), 3767-3779. doi:10.1093/jxb/ern221

- Mangiarotti, A., Galassi, V. V., Puentes, E. N., Oliveira, R. G., Del Popolo, M. G., & Wilke, N. (2019). Hopanoids Like Sterols Form Compact but Fluid Films. *Langmuir*, 35(30), 9848-9857. doi:10.1021/acs.langmuir.9b01641
- Matsuzaki, T., & 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(7), 1275-1286. doi:10.1083/jcb.200806062
- Michels, P. A., Bringaud, F., Herman, M., & Hannaert, V. (2006). Metabolic functions of glycosomes in trypanosomatids. *Biochim Biophys Acta*, 1763(12), 1463-1477. doi:10.1016/j.bbamcr.2006.08.019
- Miziorko, H. M. (2011). ENZYMES OF THE MEVALONATE PATHWAY OF ISOPRENOID BIOSYNTHESIS. *Arch Biochem Biophys*, 505(2), 131-143. doi:10.1016/j.abb.2010.09.028
- Mollinedo, F., & Gajate, C. (2015). Lipid rafts as major platforms for signaling regulation in cancer. *Adv Biol Regul*, 57, 130-146. doi:10.1016/j.jbior.2014.10.003
- Montalvetti, A., Pena-Diaz, J., Hurtado, R., Ruiz-Perez, L. M., & Gonzalez-Pacanowska, D. (2000). Characterization and regulation of *Leishmania major* 3-hydroxy-3-methylglutaryl-CoA reductase. *Biochem J*, 349(Pt 1), 27-34. doi:10.1042/0264-6021:3490027
- Moog, D., Przyborski, J. M., & Maier, U. G. (2017). Genomic and Proteomic Evidence for the Presence of a Peroxisome in the Apicomplexan Parasite *Toxoplasma gondii* and Other Coccidia. *Genome Biol Evol*, 9(11), 3108-3121. doi:10.1093/gbe/evx231
- Motley, A. M., & Hettema, E. H. (2007). Yeast peroxisomes multiply by growth and division. *J Cell Biol*, 178(3), 399-410. doi:10.1083/jcb.200702167
- Neuberger, G., Maurer-Stroh, S., Eisenhaber, B., Hartig, A., & Eisenhaber, F. (2003). Motif refinement of the peroxisomal targeting signal 1 and evaluation of taxon-specific differences. *J Mol Biol*, 328(3), 567-579. doi:10.1016/s0022-2836(03)00318-8
- Nishimura, H., Azami, Y., Miyagawa, M., Hashimoto, C., Yoshimura, T., & Hemmi, H. (2013). Biochemical evidence supporting the presence of the classical mevalonate pathway in the thermoacidophilic archaeon *Sulfolobus solfataricus*. *J Biochem*, 153(5), 415-420. doi:10.1093/jb/mvt006
- Novo, S. P., Leles, D., Bianucci, R., & Araujo, A. (2015). *Leishmania tarentolae* molecular signatures in a 300 hundred-years-old human Brazilian mummy. *Parasit Vectors*, 8. doi:10.1186/s13071-015-0666-z

- Nowicka, B., & Kruk, J. (2010). Occurrence, biosynthesis and function of isoprenoid quinones. *Biochim Biophys Acta*, 1797(9), 1587-1605. doi:10.1016/j.bbabi.2010.06.007
- Nuttall, J., Hettema, E., & Watts, D. (2012). Farnesyl diphosphate synthase, the target for nitrogen-containing bisphosphonate drugs, is a peroxisomal enzyme in the model system *Dictyostelium discoideum*. *Biochem J*, 447(Pt 3), 353-361. doi:10.1042/bj20120750
- Olivier, L. M., Kovacs, W., Masuda, K., Keller, G. A., & Krisans, S. K. (2000). Identification of peroxisomal targeting signals in cholesterol biosynthetic enzymes. AA-CoA thiolase, hmg-coa synthase, MPPD, and FPP synthase. *J Lipid Res*, 41(12), 1921-1935.
- Opperdoes, F. R., & Szikora, J. P. (2006). In silico prediction of the glycosomal enzymes of *Leishmania major* and trypanosomes. *Mol Biochem Parasitol*, 147(2), 193-206. doi:10.1016/j.molbiopara.2006.02.010
- Ortiz-Gómez, A., Jiménez, C., Estévez, A. M., Carrero-Lérida, J., Ruiz-Pérez, L. M., & González-Pacanowska, D. (2006). Farnesyl Diphosphate Synthase Is a Cytosolic Enzyme in *Leishmania major* Promastigotes and Its Overexpression Confers Resistance to Risedronate. *Eukaryot Cell*, 5(7), 1057-1064. doi:10.1128/ec.00034-06
- Pan, R., & Hu, J. (2011). The conserved fission complex on peroxisomes and mitochondria. *Plant Signal Behav*, 6(6), 870-872. doi:10.4161/psb.6.6.15241
- Pattanaik, B., & Lindberg, P. (2015). Terpenoids and Their Biosynthesis in Cyanobacteria. *Life (Basel)*, 5(1), 269-293. doi:10.3390/life5010269
- Pawlowski, J., & Burki, F. (2009). Untangling the phylogeny of amoeboid protists. *J Eukaryot Microbiol*, 56(1), 16-25. doi:10.1111/j.1550-7408.2008.00379.x
- Pearson, A., Flood Page, S. R., Jorgenson, T. L., Fischer, W. W., & Higgins, M. B. (2007). Novel hopanoid cyclases from the environment. *Environ Microbiol*, 9(9), 2175-2188. doi:10.1111/j.1462-2920.2007.01331.x
- Pena-Diaz, J., Montalvetti, A., Camacho, A., Gallego, C., Ruiz-Perez, L. M., & Gonzalez-Pacanowska, D. (1997). A soluble 3-hydroxy-3-methylglutaryl-CoA reductase in the protozoan *Trypanosoma cruzi*. *Biochem J*, 324 ( Pt 2), 619-626. doi:10.1042/bj3240619
- Petřiv, O. I., Tang, L., Titorenko, V. I., & Rachubinski, R. A. (2004). A new definition for the consensus sequence of the peroxisome targeting signal type 2. *J Mol Biol*, 341(1), 119-134. doi:10.1016/j.jmb.2004.05.064

- Peña-Díaz, J., Montalvetti, A., Flores, C. L., Constán, A., Hurtado-Guerrero, R., De Souza, W., . . . Gonzalez-Pacanowska, D. (2004). Mitochondrial Localization of the Mevalonate Pathway Enzyme 3-Hydroxy-3-methyl-glutaryl-CoA Reductase in the Trypanosomatidae. *Mol Biol Cell*, 15(3), 1356-1363. doi:10.1091/mbc.E03-10-0720
- Probst, A., Nguyen, T. N., El-Sakkary, N., Skinner, D., Suzuki, B. M., Buckner, F. S., . . . Debnath, A. (2019). Bioactivity of Farnesyltransferase Inhibitors Against *Entamoeba histolytica* and *Schistosoma mansoni*. *Front Cell Infect Microbiol*, 9. doi:10.3389/fcimb.2019.00180
- Qiu, H., Yoon, H. S., & Bhattacharya, D. (2016). Red Algal Phylogenomics Provides a Robust Framework for Inferring Evolution of Key Metabolic Pathways. *PLoS Curr*, 8. doi:10.1371/currents.tol.7b037376e6d84a1be34af756a4d90846
- Ribeiro, N., Streiff, S., Heissler, D., Elhabiri, M., Albrecht-Gary, A. M., Atsumi, M., . . . Ourisson, G. (2007). Reinforcing effect of bi- and tri-cyclopolyprenols on 'primitive' membranes made of polyprenyl phosphates. *Tetrahedron*, 63(16), 3395-3407. doi:https://doi.org/10.1016/j.tet.2007.01.076
- Rodriguez-Concepcion, M. (2004). The MEP pathway: a new target for the development of herbicides, antibiotics and antimalarial drugs. *Curr Pharm Des*, 10(19), 2391-2400. doi:10.2174/1381612043384006
- Rondon-Mercado, R., Acosta, H., Caceres, A. J., Quinones, W., & Concepcion, J. L. (2017). Subcellular localization of glycolytic enzymes and characterization of intermediary metabolism of *Trypanosoma rangeli*. *Mol Biochem Parasitol*, 216, 21-29. doi:10.1016/j.molbiopara.2017.06.007
- Rosengarten, R. D., Beltran, P. R., & Shaulsky, G. (2015). A deep coverage *Dictyostelium discoideum* genomic DNA library replicates stably in *Escherichia coli*. *Genomics*, 106(4), 249-255. doi:10.1016/j.ygeno.2015.05.008
- Schilde, C., Lawal, H. M., Kin, K., Shibano-Hayakawa, I., Inouye, K., & Schaap, P. (2019). A well supported multi gene phylogeny of 52 dictyostelia. *Mol Phylogenet Evol*, 134, 66-73. doi:10.1016/j.ympev.2019.01.017
- Schluter, A., Fourcade, S., Ripp, R., Mandel, J. L., Poch, O., & Pujol, A. (2006). The evolutionary origin of peroxisomes: an ER-peroxisome connection. *Mol Biol Evol*, 23(4), 838-845. doi:10.1093/molbev/msj103

- Schwender, J., Gemunden, C., & Lichtenthaler, H. K. (2001). Chlorophyta exclusively use the 1-deoxyxylulose 5-phosphate/2-C-methylerythritol 4-phosphate pathway for the biosynthesis of isoprenoids. *Planta*, 212(3), 416-423. doi:10.1007/s004250000409
- Shen, Y. Q., Lang, B. F., & Burger, G. (2009). Diversity and dispersal of a ubiquitous protein family: acyl-CoA dehydrogenases. *Nucleic Acids Res*, 37(17), 5619-5631. doi:10.1093/nar/gkp566
- Shi, L., & Tu, B. P. (2015). Acetyl-CoA and the Regulation of Metabolism: Mechanisms and Consequences. *Curr Opin Cell Biol*, 33, 125-131. doi:10.1016/j.ceb.2015.02.003
- Shimazu, T., Hirsche, M. D., Hua, L., Dittenhafer-Reed, K. E., Schwer, B., Lombard, D. B., . . . Verdin, E. (2010). SIRT3 deacetylates mitochondrial 3-hydroxy-3-methylglutaryl CoA synthase 2 and regulates ketone body production. *Cell Metab*, 12(6), 654-661. doi:10.1016/j.cmet.2010.11.003
- Simons, K., & Ikonen, E. (1997). Functional rafts in cell membranes. *Nature*, 387(6633), 569-572. doi:10.1038/42408
- Small, I., Peeters, N., Legeai, F., & Lurin, C. (2004). Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics*, 4(6), 1581-1590. doi:10.1002/pmic.200300776
- Smith, J. J., & Aitchison, J. D. (2013). Peroxisomes take shape. *Nat Rev Mol Cell Biol*, 14(12), 803-817. doi:10.1038/nrm3700
- States, D. J., & Gish, W. (1994). Combined use of sequence similarity and codon bias for coding region identification. *J Comput Biol*, 1(1), 39-50. doi:10.1089/cmb.1994.1.39
- Striepen, B. (2011). The apicoplast: a red alga in human parasites. *Essays Biochem*, 51, 111-125. doi:10.1042/bse0510111
- Sunter, J., & Gull, K. (2017). Shape, form, function and *Leishmania* pathogenicity: from textbook descriptions to biological understanding. *Open Biol*, 7(9). doi:10.1098/rsob.170165
- Swinkels, B. W., Gould, S. J., & Subramani, S. (1992). Targeting efficiencies of various permutations of the consensus C-terminal tripeptide peroxisomal targeting signal. *FEBS Lett*, 305(2), 133-136. doi:10.1016/0014-5793(92)80880-p
- Tower, R. J., Fagarasanu, A., Aitchison, J. D., & 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(10), 1727-1738. doi:10.1091/mbc.E11-01-0084
- UniProt Consortium (2019). UniProt: A Worldwide Hub of Protein Knowledge. *Nucleic acids research*, 47(D1). doi:10.1093/nar/gky1049

- Urbina, J. A., Concepcion, J. L., Rangel, S., Visbal, G., & Lira, R. (2002). Squalene synthase as a chemotherapeutic target in *Trypanosoma cruzi* and *Leishmania mexicana*. *Mol Biochem Parasitol*, 125(1-2), 35-45. doi:10.1016/s0166-6851(02)00206-2
- Valitova, J. N., Sulkarnayeva, A. G., & Minibayeva, F. V. (2016). Plant Sterols: Diversity, Biosynthesis, and Physiological Functions. *Biochemistry (Mosc)*, 81(8), 819-834. doi:10.1134/s0006297916080046
- van der Klei, I. J., & Veenhuis, M. (2006). PTS1-independent sorting of peroxisomal matrix proteins by Pex5p. *Biochim Biophys Acta*, 1763(12), 1794-1800. doi:10.1016/j.bbamcr.2006.08.013
- van der Zand, A., Braakman, I., & Tabak, H. F. (2010). Peroxisomal membrane proteins insert into the endoplasmic reticulum. *Mol Biol Cell*, 21(12), 2057-2065. doi:10.1091/mbc.E10-02-008210.1091/mbc.e10-02-0082
- van der Zand, A., Gent, J., Braakman, I., & Tabak, H. F. (2012). Biochemically distinct vesicles from the endoplasmic reticulum fuse to form peroxisomes. *Cell*, 149(2), 397-409. doi:10.1016/j.cell.2012.01.054
- van Dooren, G. G., & Striepen, B. (2013). The algal past and parasite present of the apicoplast. *Annu Rev Microbiol*, 67, 271-289. doi:10.1146/annurev-micro-092412-155741
- VanNice, J. C., Skaff, D. A., Keightley, A., Addo, J. K., Wyckoff, G. J., & Miziorko, H. M. (2014). Identification in *Haloferax volcanii* of Phosphomevalonate Decarboxylase and Isopentenyl Phosphate Kinase as Catalysts of the Terminal Enzyme Reactions in an Archaeal Alternate Mevalonate Pathway. *J Bacteriol*, 196(5), 1055-1063. doi:10.1128/jb.01230-13
- Veltman, D. M., Akar, G., Bosgraaf, L., & Van Haastert, P. J. (2009). A new set of small, extrachromosomal expression vectors for *Dictyostelium discoideum*. *Plasmid*, 61(2), 110-118. doi:10.1016/j.plasmid.2008.11.003
- Vinokur, J., Korman, T. P., Cao, Z., & Bowie, J. U. (2014). Evidence of a Novel Mevalonate Pathway in Archaea. *Biochemistry*, 53(25), 4161-4168. doi:10.1021/bi500566q
- von Heijne, G., Steppuhn, J., & Herrmann, R. G. (1989). Domain structure of mitochondrial and chloroplast targeting peptides. *Eur J Biochem*, 180(3), 535-545.
- Wang, J., Wu, X., Zhang, Z., Du, X., Chai, R., Liu, X., . . . Sun, G. (2008). Fluorescent co-localization of PTS1 and PTS2 and its application in analysis of the gene function and the peroxisomal dynamic in *Magnaporthe oryzae*\*§. *J Zhejiang Univ Sci B*, 9(10), 802-810. doi:10.1631/jzus.B0860001

- Xu, L., & Simoni, R. D. (2003). The inhibition of degradation of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase by sterol regulatory element binding protein cleavage-activating protein requires four phenylalanine residues in span 6 of HMG-CoA reductase transmembrane domain. *Arch Biochem Biophys*, 414(2), 232-243. doi:10.1016/s0003-9861(03)00168-1
- Yang, X., Purdue, P. E., & Lazarow, P. B. (2001). Eci1p uses a PTS1 to enter peroxisomes: either its own or that of a partner, Dci1p. *Eur J Cell Biol*, 80(2), 126-138. doi:10.1078/0171-9335-00144
- Yang, X., Tyler, B. M., & Hong, C. (2017). An expanded phylogeny for the genus *Phytophthora*. *IMA Fungus*, 8(2), 355-384. doi:10.5598/imafungus.2017.08.02.09
- Yoon, H. S., Hackett, J. D., Pinto, G., & Bhattacharya, D. (2002). The single, ancient origin of chromist plastids. *Proc Natl Acad Sci U S A*, 99(24), 15507-15512. doi:10.1073/pnas.242379899
- Zakrys, B., Milanowski, R., & Karnkowska, A. (2017). Evolutionary Origin of *Euglena*. *Adv Exp Med Biol*, 979, 3-17. doi:10.1007/978-3-319-54910-1\_1
- Zarella-Boitz, J. M., Rager, N., Jardim, A., & Ullman, B. (2004). Subcellular localization of adenine and xanthine phosphoribosyltransferases in *Leishmania donovani*. *Mol Biochem Parasitol*, 134(1), 43-51. doi:10.1016/j.molbiopara.2003.08.016
- Zhu, G., Marchewka, M. J., & Keithly, J. S. (2000). *Cryptosporidium parvum* appears to lack a plastid genome. *Microbiology*, 146 ( Pt 2), 315-321. doi:10.1099/00221287-146-2-315