HETEROLOGOUS EXPRESSION OF EXTRACELLULAR PROTEINS BY *YARROWIA LIPOLYTICA*

by

Nokukhanya Hlengiwe Mfumo

Submitted in fulfillment of the requirements for the degree of

MAGISTER SCIENTIAE

In the Faculty of Natural and Agricultural Sciences
Department of Microbial, Biochemical and Food Biotechnology
University of the Free State
Bloemfontein
South Africa

June 2014

Supervisor: Prof MS Smit
Co-Supervisor: Dr MB Nthangeni

DECLARATIONS

I declare that the dissertation hereby submitted by me for the **Magister Scientiae** degree at the University of the Free State is my own independent work and has not previously been submitted by me at another university/faculty. I further more cede copyright of the dissertation in favour of the University of the Free State.

_____ _____

Nokukhanya Hlengiwe Mfumo Date

DEDICATIONS

This dissertation is dedicated to my mother **Mildred Mfumo** for seeing it that I get the best life and education that she could afford. Thank you mom!

ACKNOWLEDGEMENTS

The study was possible due to the support that I received from different people, and if it were not them, the work would not have come to completion:

- Bheki Ncube, my husband for all the support, patience and understanding.
- My kids Sesiphelele and Sicelo for enduring the lonely times away from the warm and comforting hands of me their mother.
- Dr Nobalanda Mabizela for guidance in certain experiments and for assistance in writing up of the dissertation.
- Dr Faranani Ramagoma for the supervision, guidance and support.
- Dr Bethuel Nthangeni for believing in me despite all the disappointments that came to you from my side; you never gave up on me. This work is a result of your invaluable inputs and merciless critique of the work.
- Prof Smit for continued interest in the work, provision of research advice, assistance and patience with almost endless requests to register, and re-register.
- Dr Stoyan Stoychev for assistance with the design and execution of peptide mapping experiments.
- The CSIR Biosciences through Dr Bethuel Nthangeni in providing me with the opportunity to carry out this study in their laboratories, and provision of research funds.
- My employer, OBP for keeping a blind eye on the conflict between the time for work and for completing the study.
- Wendy Limani for encouraging me to push and finish writing up and not to loose hope.
- Shirley Lukhwareni for assisting me with access to research articles, the endless request for research articles did not bother you to a point of rejection.
- Above all, the almighty God for blessing me with all the people mentioned above who came and graced my life, for providing me with the wisdom, endurance and patience enough to see this project through!

LIST OF ESSENTIAL ABREVIATIONS

3D	-	Three Dimensional
aa	-	amino acid residues
Ac No	-	Accession number
ACN	-	Acetonitrile
AEP	-	Acyl-coenzyme A oxidase 2
AIX	-	Ampicillin IPTG-X-Gal
ARS	-	Autonomous Replication Sequences
BCA	-	Bicinchoninic Acid
BLAST	-	Basic Local Alignment Search Tool
bp	-	base pair
BSA	-	Bovine serum albumin
CAI	-	Codon Adaptation Index
CAPS	-	*N*-cyclohexyl-3-aminopropanesulfonic acid
CPO	-	Chloroperoxidase
CPR	-	Cytochrome P450 reductase
CTAB	-	Cetyl Trimethylammonium Bromide
dH_2O	-	Distilled water
DIG	-	Digoxigenin
DNA	-	Deoxyribonucleic Acid
DTT	-	dithiothreitol
EC	-	Enzyme Commission
EDTA	-	Ethylenediaminetetraacetic acid
ER	-	Endoplasmic reticulum
EXPASY	-	Expert Protein Analysis System
FDA	-	Food and Drug Administration
g	-	Gravitational force

G3P	-	Glycerol-3-phosphate dehydrogenase
GCSF	-	Granulocyte colony stimulating factor
GRAS	-	Generally Regarded As Safe
HRPO	-	Horse Reddish peroxidase
ICL	-	Isocitrate lyase
IPTG	-	Isopropyl β-D-thiogalactoside
kDa	-	kilo Dalton
LB	-	Luria Bertani
LC-MS/MS	-	Liquid chromatography tandem mass spectrometry
LTR	-	Long terminal repeat
MOPS	-	3-(N-morpholino) propanesulfonic acid
M_r	-	Relative Molecular weight
mRNA-	-	Messenger Ribonucleic Acid
NCBI	-	*National Center for Biotechnology Information*
Ni-NTA	-	Nickel-nitriloacetic acid
OD	-	Optical density
ORF	-	Open Reading Frame
PAGE	-	Polyacrylamide gel electrophoresis
PCR	-	Polymerase Chain Reaction
PDI	-	Proteins disulfide isomerase
pI	-	Isoelectric point
POX2	-	*Yarrowia lipolytica* gene encoding Acyl-coenzyme A oxidase 2
rDNA	-	Ribosomal Deoxyribonucleic Acid
RNA	-	Ribonucleic acid
rpm	-	revolution per minute
RSL	-	Rapid separation liquid chromatography
SDS	-	Sodium dodecyl sulphate
SRP -	-	Signal Recognition Particles

SSC	-	Saline-Sodium Citrate
STET	-	Sodium chloride ethylenediaminetetraacetic acid Tris Triton X-100
TAE	-	Tris acetic acid ethylenediaminetetra acetic acid
TE	-	Tris -Ethylenediaminetetraacetic acid
TEF	-	Elongation factor 1-alpha
TMB	-	3,3′,5,5′-Tetramethylbenzidine
Tris	-	Tris (hydroxymethyl) aminomethane
Tween 20	-	Polysorbate 20
UAS	-	Upstream Activating Sequences
UPBRC	-	University of Pretoria Biomedical Research Centre
X-GAL	-	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside
XPR2	-	Extracellular alkaline protease

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER 1

1. INTRODUCTION AND BACKGROUND TO THE STUDY

1.1. Introduction

The advent of recombinant DNA technology has provided routes alternate to natural sources for the production of industrial and therapeutic proteins (Skerker *et al*., 2009). Recombinant protein production has become a multibillion-dollar market (Mattanovich *et al*., 2012). Heterologous gene expression is of considerable interest for the production of industrial and pharmaceutical proteins (Domínguez *et al*., 1998). The expression of foreign genes and production of proteins of interest are very important for both the basic research such as elucidation of physiological activity, its modulation, analysis of structure-function relationship of control elements and regulation of gene expression as well as practical applications related to the production of pharmaceuticals and chemicals (Nasser *et al.,* 2003). The demand for expression systems suitable for high-level synthesis of functional foreign gene products is evidenced by the large number of publications on heterologous protein production. The expression systems consist of combinations of various genetic elements of host and vector. While one single perfect host for every protein does not exist, a number of expression systems, some using bacterial hosts and others using fungi, yeasts, insect cells, mammalian cells and other eukaryotes have been described in literature. In general, the expression of mammalian genes using bacteria as host may sometimes result in an inactive product due to incorrect folding or lack of certain post-translational modifications, though the manipulation of bacteria is easy and the production cost is relatively low. In contrast, most of these problems can easily be solved through expression of genes using animal cells as a host. However, their manipulation is not easy, the production levels are low and the cost is high. Moreover, the mammalian cell expression systems sometimes have the problem of viral infections (Nasser *et al*., 2003).

Yeasts have been developed as host organisms for the production of foreign (heterologous) proteins (Domínguez *et al*., 2010). Yeasts combine the ease of genetic

manipulation and up-scaling of microbial cultures with the ability to secrete and modify proteins with the major eukaryotic post-translational modifications. However, yeasts suffer the drawback of modifying glycoproteins with non-human high mannose-type *N*-glycans, limiting their application as hosts for therapeutic proteins production (Wang *et al*., 2013). *Saccharomyces cerevisiae* has usually been the yeast of choice (Nevoigt, 2008), but an increasing number of alternative non-*Saccharomyces* yeasts have now become accessible for modern molecular genetics techniques (Domínguez *et al*., 1998; Gellissen *et al*., 2005). The best-known alternatives to *S. cerevisiae* are *Kluyveromyces lactis*, *Yarrowia lipolytica*, *Hansenula polymorpha* and *Pichia pastoris* (Domínguez *et al*., 1998). The direct comparison of different yeast platform expression systems such as *S. cerevisiae*, *P. pastoris*, *Hansensula polymorpha*, *K. lactis, Schizosaccharomyces pombe* and *Arxula adeninivorans* established the *Y. lipolytica* yeast as an attractive host for heterologous proteins production (Gellissen *et al*., 2005). Müller *et al*., (1998) investigated alternative hosts to *S. cerevisiae* for heterologous protein expression. They compared the capacity of *S. cerevisiae*, *H. polymorpha*, *K. lactis, S. pombe* and *Y. lipolytica* to express and secrete six fungal enzymes in their active forms. The *Y. lipolytica* yeast was found to be the most efficient especially in terms of performance reproducibility (Muller *et al*., 1998).

Y. lipolytica is characterized by several advantageous traits for heterologous protein production (Gellissen *et al*., 2005). The yeast has been found to have many attributes which make it an attractive host for heterologous protein production. It has been found to have high capabilities to secrete large amount of high molecular weight proteins into the medium (Nicaud *et al*., 2002; Juretzek *et al*., 2001; Pignède *et al*., 2000). The translocation of nascent protein through the endoplasmic reticulum membrane in *Y. lipolytica* has been well studied and is more representative of vesicular secretion of animals and other fungi (Swennen and Beckerich, 2007). The *Y. lipolytica* yeast has been the subject of glycoengineering studies with the ultimate objective of developing it as an efficient expression system for the production of glycoproteins with humanized glycans (De Pourcq *et al*., 2012). The results provided further support for the idea to use the yeast as host for expression of therapeutic proteins of animal and mammalian origins. There is sufficient information regarding its molecular tools and genetic markers for easy of manipulation of *Y. lipolytica* (Madzak *et al*., 2000; Nicaud *et al*.,

2002; Madzak *et al*., 2004). Efficacy and safety studies have demonstrated the safe use of *Yarrowia*-derived products containing significant proportions of *Yarrowia* biomass or with the yeast itself as the final product (Groenewald *et al*., 2014). The yeast is considered non-pathogenic to humans (Spencer *et al.,* 2002), and has been granted the GRAS status by the American Food and Drug Administration (FDA) for citric acid production (Barth and Gaillardin 1996; 1997).

Y. lipolytica has been developed to be an attractive yeast for heterologous production of proteins ranging from simple recombinant proteins to more complex antibodies (Madzak *et al*., 2004; Madzak and Beckerich, 2013). Notable is the reported ability of the *Y. lipolytica* yeast to express eukaryotic cytochrome P450 protein complexes (Mauersberger *et al*., 2013). The cultivation conditions for biomass optimization and for scaled-up batch to fed-batch processes to improve yeast growth and enhance production of heterologous proteins are widely reported in literature (Kim *et al*., 2000; Galvagno *et al*., 2011). However most of the studies describing recombinant protein production by this yeast rely on the use of complex media which is not convenient for large scale production particularly for products intended for pharmaceutical applications. An efficient defined medium for large scale production of heterologous proteins by *Y. lipolytica* suitable for expression of therapeutics has been reported (Gasmi *et al*., 2011). Genetic mutants of *Y. lipolytica* with enhanced capacity to secrete extracellular proteins have been identified in literature (Fickers *at al*., 2003, 2005; Darvishi, 2011; Ghezelbash *et al*., 2014). The amount of literature spanning basic, applied and commercial applications is a demonstration of the reliability and versatility of *Y. lipolytica* as attractive host for heterologous production.

1.2. Background of the study

The fungus *Caldariomyces fumago* chloroperoxidase (EC 1.11.1.10) or CPO is a versatile heme-containing heavily gycosylated enzyme with a molecular mass of about 42 kDa, that exhibits peroxidase, catalase and cytochrome P450-like activities in addition to catalyzing halogenation reactions (Morris and Hager, 1966; Pickard, 1991, Hofrichter and Ullrich, 2006). The chloroperoxidase catalyzes a wide range of

biotechnologically important oxidations and molecular conversions. The chloroperoxidase oxidation reactions are highly desirable in a multitude of topical applications for the development and large-scale manufacture of pharmaceuticals to industrial wastewater treatment (Manoj *et al*., 2000; Osborne *et al*., 2006; Hofrichter and Ullrich, 2006). The filamentous fungus *C. fumago* is currently the only source for this highly versatile biocatalyst and the enzyme is commercially available. The main drawback that CPO has not yet been used in large-scale industrial processes is that its production costs are far too high (Buchhaupt, 2011). The expression of CPO in *Escherichia coli* had inadequate success for synthesis of the active form of CPO; the results showed that the non-glycosylated enzyme was obtained in its apo-form without the heme being incorporated into the enzyme (Zong *et al*., 1995). Only under high pressure, was the protein refolded with heme to generate an active enzyme in very low yield. The expression of the CPO in the yeasts *S. cerevisiae* and *P. pastoris* also did not lead to the production of active protein (Zong *et al*., 1995; Conesa, *et al*., 2001a). Expression of a mutant CPO in the parental host *C. fumago* has also been reported (Yi *et al.,* 1999). However, the presence of residual wild type CPO complicated the selection of recombinant CPO from native enzyme for further investigations. The only successful heterologous expression of the CPO was with the use of the *Aspergillus niger* fungal expression system (Conesa *et al*., 2001a). The attempts to increase CPO production have now focused on the generation of *C. fumago* mutant strains with superior production capacities for CPO activity (Buchhaupt *et al*., 2011).

Human granulocyte colony-stimulating factor (hG-CSF) is a proinflammatory cytokine and hematopoietic growth factor. Recombinant human granulocyte-macrophage colony-stimulating factor (hG-CSF) serves as a biotherapeutic agent in bone marrow stimulations, vaccine development, gene therapy approaches, and stem cell mobilization (Li, 2011; Srinivasa Babu *et al*., 2014). Since its isolation, the human granulocyte- colony stimulating factor has been proposed as a new class of therapeutic biological products in the treatment of various diseases. The therapeutic protein has been a subject of a number of heterologous production studies in *P. pastoris* (Jacobs *et al*., 2010; Srinivasa Babu *et al*., 2014), *S. cerevisiae* (Bae *et al*., 1998; Bae *et al*., 1999; Pozzuolo *et al*., 2008), and *E. coli* (Das *et al*., 2011; Khasa *et al*., 2011; Kim *et al*., 2014). The major obstacles to bioprocess development for large scale production is the

toxicity towards the *E. coli* expression host (Wang *et al*., 2008; Das *et al*., 2011; Khasa *et al*., 2011), protein insolubility due to aggregation in *P. pastoris* (Lasnik *et al*., 2001, Bahrami *et al.,* 2009; Srinivasa Babu *et al*., 2008), and formation of undesirable multimers in the culture broth of *S. cerevisiae* (Bae *et al*., 1999).

1.3. Research problem

The hypothesis that was formulated based on the large number of literature available was of possible successful exploitation of *Y. lipolytica* expression systems for heterologous production and secretion of recombinant CPO and human G-GSF proteins. The proteins have in common that they require post-translation glycosylation modifications, are characterized by presence of disulphide bonds, and have presented challenges when heterologously expressed in other yeast and bacterial expression systems. The successful expression would indicate the possibilities of further exploring the economic ways of producing at commercial scales the CPO and hG-CSF proteins.

1.4. Objectives of the study

The objective of the study is to explore the versatility of the *Y. lipolytica* as host for extracellular production of *C. fumago* chloroperoxidase and the human granulocyte colony-stimulating factor. The genes encoding the two commercially important proteins will be codon optimized for expression in the *Y. lipolytica* yeast, cloned in multi-copy expression and secretion vectors. Protein production and secretion will be evaluated through the use of Western Blot detection, purification and enzyme activity assays.

CHAPTER 2

Literature review: *Yarrowia lipolytica* as host system for heterologous protein expression

2.1. The *Y. lipolytica* yeast

Yarrowia lipolytica belongs to a family of hemiascomycetous yeasts, originally classified as *Candida lipolytica* and later reclassified as *Saccharomycopsis lipolytica* (Yarrow, 1972). *Yarrowia* was proposed in acknowledgement of a new genus identified by David Yarrow from the Delft Microbiology Laboratory (van der Walt and von Arx, 1980). The species name lipolytica originated from the ability of this yeast to hydrolyze lipids (van der Walt and von Arx, 1980). The yeast *Y. lipolytica* is a dimorphic fungus which forms yeast-like cells or true mycelium and pseudo hyphae, with transition highly dependent on the growth medium (Pérez-Campo and Domínguez *et al*., 2001). *Y. lipolytica* wild type strains are broadly isolated from dairy products (Jacques and Casaregola, 2008), and are strictly aerobe with a unique capability to utilize aliphatic hydrocarbons such as alkanes as well as fatty acids as carbon sources (Kerscher *et al.*, 2001). Its recommended temperature for growth and sporulation is 20-30°C (Barth and Gaillardin, 1997).

The natural occurrence of the *Y. lipolytica* species in food, particularly dairy products and meat, augmented its Generally Regarded As Safe (GRAS) classification. Efficacy and safety studies have demonstrated the safe use of *Yarrowia*-derived products containing significant proportions of *Yarrowia* biomass or with the yeast itself as the final product (Groenewald *et al*, 2014). *Y. lipolytica* has consequently been developed as a production host for a large variety of biotechnological applications. This yeast's potential application in the production of industrial bioproducts (Bankar *et al*., 2009; Chi *et al*., 2010; Max *et al*., 2010; Gonçalves *et al*., 2014), production of single cell oil proteins (Beopoulos *et al*., 2009; Ageitos *et al*., 2011), bioremediation processes (Bankar *et al*., 2009; Zinjarde *et al*., 2014), as biocontrol agent (Chi *et al*. 2010) and in food processing technologies (Ogrydziak DM, 1993; Waché *et al*., 2003; Fickers *et al*., 2011; Sabirova *et al*., 2011) has been extensively reviewed.

The *Y. lipolytica* yeast is widely used in production of recombinant proteins of medical or industrial interest. The characteristics of interest in *Y. lipolytica* as host for recombinant protein production include its ability to rapidly reach high cell densities (Kim *et al*., 2000) and the capacity to utilize unusual hydrocarbons as carbon sources (Thevenieau *et al*., 2007; Fukuda *et al*., 2013; Palande *et al*., 2014). Several strains of *Y. lipolytica* have been engineered to have further advantages such as humanized glycosylation pathways (Moon *et al*., 2013; De Pourcq *et al*., 2012) or lack of proteases (Nicaud *et al*., 2002). The availability of a large variety of vectors, promoters, selection markers and protein localization signals to choose from (Nicaud *et al*., 2002; Madzak *et al*., 2000; Madzak *et al*., 2004; Juretzek *et al*., 2001), combined with the accumulated knowledge on industrial-scale fermentation techniques (Rywińska *et al*., 2012; Gonçalves *et al*., 2014) and the current advances in the post-genomic technology (Casaregola *et al*., 2000; Kerscher *et al*, 2001), has made *Y. lipolytica* yeast an attractive host for heterologous protein production. This chapter reviews the current understanding of the *Y. lipolytica* expression systems focusing on its applicability as a host system of choice for heterologous protein production.

2.2. Protein expression and molecular tools in *Yarrowia lipolytica*

The use of *Y. lipolytica* as host for heterologous protein production has been made possible by the availability of molecular and protein expression and secretion technologies to manipulate the host. The introduction of recombinant proteins for heterologous expression is generally done by transforming the *Y. lipolytica* host with the expression vector carrying the DNA fragment of interest. In addition to the gene or ORF of interest, the expression system usually contains a selection marker, a promoter, secretion signal, transcription terminator and sequences to localize and maintain the expression cassette within the *Y. lipolytica* host. The structural components that characterize common *Y. lipolytica* expression systems are summarized in Table 2.1.

Table 2.1: Common structural components of *Y. lipolytica* expression constructs.

Integrative sites	Selection markers	Promoters	Secretion signals	Terminators
ARS	*Leu*2	XPR2	Native	*LIP2*
rDNA	*Ura*3	HP4D	*LIP2*	*XPR*2
pBR322	*Ura*3d1	TEF	*XPR2*	
Zeta	*Ura*3d4	POX2		
	Hygromycin	ICL		

2.2.1. Expression vectors

The *Y. lipolytica* expression vectors are usually propagated in *E. coli*, and as such are made up of a bacterial moiety carrying the plasmid origin of replication and selection marker encoding antibiotic resistance in addition to the yeast expression cassette (Nicaud *et al.,* 2002). The *Y. lipolytica* expression vectors can be distinguished into episomal and integrative vectors. The episomal vectors replicate autonomously while the integrative vectors lack autonomous replication.

2.2.1.1. Episomal replicative vectors

Autonomous Replication Sequences or ARS (Table 2.1) genetic elements displaying extrachomosomal and autonomous replication activity have been described in literature (Fournier *et al*., 1993; Vernis *et al*., 1997). The cloning of *Y. lipolytica ARS1* and *ARS2* into the *LEU2* selective integrative plasmid conferred on the hybrid plasmids high transformation efficiency and enabled extrachromosomal transmission of the plasmids in 1 or 2 copies per yeast cell under selective conditions (Matsuoka *et al*., 1993). The ARS-carrying plasmids exhibit relative mitotic stability due to the presence of the centromere (CEN) sequences, and are usually in low copy numbers (Madzak *et al*., 2004). The use of plasmids based on these ARS/CEN elements is impractical for higher amplification of gene expression (Nicaud *et al.,* 1991). The copy numbers of genes expressed using these vectors are limited to 1-3 copies per cell since they are only stable as ARS/CEN plasmids. The other main drawback for industrial use of these vectors is the need for maintenance by selective pressure (Madzak *et al.,* 2004).

2.2.1.2. Integrative vectors

The integration of vectors into the *Y. lipolytica* genome can be achieved by either homologous recombination, which is strongly stimulated by the linearization of the plasmid within the targeting region or non-homologous integration.

Expression vectors which integrate into host genome preferably by homologous recombination have been described (Barth and Gaillardin, 1996). These integrative vectors are usually constructed based on the availability of a target integration site with in the *Y. lipolytica* genome sequences. Homologous recombination by integrative vectors typically occur as a single copy event. However, methods to integrate multiple copies have been developed using DNA sequences available in the *Y. lipolytica* genome in multiple repeats such as the ribosomal DNA (*rDNA*) cluster. In the first attempts to increase copy numbers in *Y. lipolytica* genome, homologous rDNA (Table 2.1) unit clusters located on the various chromosomes were used as target sites in conjunction with the *URA3* defective selection marker (Le Dall *et al.,* 1994). The rDNA cluster consists of about 140 tandem repeats of a 9.1 kb unit on several chromosomes (Casaregola *et al.,* 1997). The total number of rDNA units per *Y. lipolytica* genome were evaluated and confirmed to be more than 200 (Casaregola *et al.,* 1997). Tandem multi-copy inserts are mostly attained when high DNA concentrations of integration vectors are used; resulting in multi-copy transformed clones that are stable due to repeated recombination events. Integrative vectors based on *Y. lipolytica* strains fitted with the bacterial DNA from the pBR322 plasmid have been constructed (Madzak *et al.,* 2000, 2004). The expression vectors are designed to contain the pBR322 DNA sequences, and the target site for integration is the pBR322 docking platform recombinantly introduced within the genome of the *Y. lipolytica* recipient yeast strain (Madzak *et al*, 2004).

The *Ylt*1 *zeta* element which has been used for both homologous and non-homologous integration of expression cassettes is an interesting feature of *Y. lipolytica*. The vectors carrying *zeta* elements allow integration to be homologous in *Y. lipolytica* strains carrying *Ylt*1 sequences and to be non-homologous in strains devoid of *Ylt*1 genetic elements (Pignede *et al.,* 2000; Juretzek *et al.,* 2001). The *Ylt*1 is a 9.6 kb long retrotransposon bound by a long terminal repeat (LTR) referred to as the *zeta* element (Schmid-Berger *et al.,* 1994). The *zeta* element is a 714 bp long, highly conserved

genetic element capable of solo existence. The *Ylt*1 and solo *zeta* regions are flanked by a 4 bp directly repeated DNA sequence. These repetitive elements provide potential targeting sites to direct the integration of expression cassettes into the yeast genome (Pignede *et al.,* 2000; Juretzek *et al.,* 2001; Schmid-Berger *et al.,* 1994). There are at least 35 copies of *Ylt*1 present per haploid genome in a dispersed manner and about 50-60 copies of solo *zeta* elements, although the number of *Ylt*1 and solo *zeta* elements differ per strain (Juretzek *et al.,* 2001). Autocloning vectors derived from the 714 bp *zeta* elements flanking the expression cassettes have been developed (e.g. Figure 2.1) and used as genome integrating elements in *Y. lipolytica* strains devoid of homologous *Ylt*1 or solo *zeta* units (Pignède *et al*., 2000; Emond *et al*., 2010).

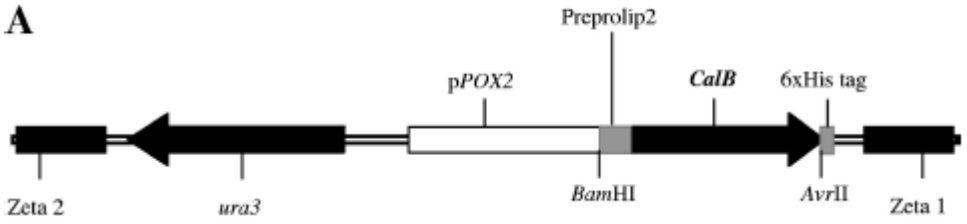

Figure 2.1: Schematic diagram of the expression cassette showing the *zeta* elements flanks of the expression cassette. The flanking *zeta* elements facilitate non-homologous recombination within the *Y. lipolytica* genome (adapted from Emond *et al.,* 2010).

2.2.2. Selection Markers

The yeast selection markers can be classified into two types: dominant and complementation selection markers. Dominant selection markers are typified by antibiotic resistant markers that can be used in yeast selection such as the hygromycin-B resistance gene (Cordero Otero and Gaillardin, 1996). The direct selection of hygromycin-B resistant *Y. lipolytica* transformants on complete medium resulted in transformation frequencies comparable to those observed with conventional auxotrophic markers (Cordero Otero and Gaillardin, 1996). The selection marker based on hygromycin resistance found application in integrating single copies of plasmid DNA, gene disruption and selection marker rescue in the *Y. lipolytica* yeast (Cordero Otero and Gaillardin, 1996; Fickers *et al*., 2003). The complementation markers for use in *Y. lipolytica* are complementary to host strain auxotrophy. These auxotrophic markers are used in selection of recombinants with all the types of integrations, as well as episomal and centromeric plasmids expression systems. The *Y. lipolytica LEU2* and *URA3* are non-leaky and non-reverting auxotrophic markers and recipient strains are available (Madzak *et al.,* 2004). The wild-type *URA3* allele (*ura3d*1), is used for single-copy integration, and a mutant *URA3* allele, *ura3d*4, is used to select for multi-copy integrations (Le Dall *et al.,* 1994; Nicaud *et al*., 2002). The copy numbers for *ura3d4* defective marker selections avarages 10-13 copies, reflecting the optimal auxotrophy complementation (Juretzek *et al.,* 2001).

2.2.3. Transcriptional promoters and terminators

The highly active transcriptional genetic elements to initiate and drive heterologous gene expression in *Y. lipolytica* are available to function as promoters (Table 2.1). Regulated and poorly regulated promoters have been developed into inducible and constitutive expression systems, respectively. The most commonly used regulated promoter element for heterologous gene transcription in *Y. lipolytica* has been the XPR2 derived from the promoter for the gene encoding alkaline extracellular protease (*AEP*) (Ogrydziak *et al.,* 1977). The high transcription activity and full induction of the XPR2 promoter require high levels of peptone in the culture medium lacking preferred nitrogen and carbon sources at pH above 6 (Ogrydziak *et al.,* 1977; Hamsa and Chatto 1994). These functional regulatory requirements and complexity limit its practical usage

since availability of peptone in the medium complicates the purification and recovery of the recombinant product (Gellissen *et al.,* 2005). To circumvent XPR2 promoter difficulties, a synthetic hybrid promoter based on the XPR2 was developed. The upstream activating sequences UAS1 and UAS2 of the XPR2 promoter are essential for promoter activity irrespective of environmental conditions (Blanchin-Roland *et al.,* 1994). A hybrid promoter was constructed by combining four tandem copies of the UAS1 elements upstream of a minimal LEU2 promoter, reduced to its TATA box (Madzak *et al*., 2000). The resulting recombinant hybrid promoter was termed HP4D, and its highly active transcriptional function was reported to be independent of environmental conditions (Madzak *et al*., 2000). The heterologous gene expression driven by HP4D contains unidentified elements that drive growth phase dependent gene expression given that its gene expression was found to occur at the beginning of the stationary phase (Madzak *et al.,* 2000). In addition to the HP4D promoter, there exists two strong constitutive promoters, the TEF and the RPS7, derived from the *Y. lipolytica* genes encoding translation elongation factor-1 alpha and the ribosomal protein, respectively (Müller *et al*., 1998).

The genes encoding for enzymes required for the assimilation of hydrophobic substrates (Gellissen *et al.,* 2005) have been used as sources of inducible promoters. Notable are the promoters transcribing the isocitrate lyase (ICL), 3-oxo-acyl-CoA thiolase (POT1), and acyl-CoA oxidases (*POX1, POX2 and POX5*) genes (Juretzek *et al.,* 2000, Gellissen *et al.,* 2005). The POX promoters are induced by the addition of fatty acids and repressed by availability of glucose or glycerol. Besides fatty acids, the ICL promoter is also induced by ethanol and acetate, and repression by glucose or glycerol does not result in complete repression (Juretzek *et al.,* 2000). The XPR2, POX2 and ICL promoters have found wide applications in recombinant protein expression in the *Y. lipolytica* yeast (Table 2.3).

An efficient termination of transcription is required for maximal gene expression (Zaret and Sherman, 1982). The yeast expression vectors are developed with transcriptional terminators for efficient mRNA 3' end formation. The terminators from the *XPR2* and lipase (*LIP2*) genes respectively encoding for endogenous extracellular alkaline protease and lipase enzymes have been used for application in the *Y. lipolytica* yeast expression vectors (Nicaud *et al*, 2002; Madzak *et al*, 2004).

2.2.4. Protein secretion and localisation signals

Heterologous protein secretion in *Y. lipolytica* can be attained using secretion signals; and this could either be a foreign signal derived from the protein being secreted (native), or derived from a protein encoded by a gene endogenous to *Y. lipolytica* such as the *LIP2* or *XPR2* (Table 2.1). In *Y. lipolytica* the signal sequence of AEP encoded by the *XPR2* gene and the lipase (*LIP2)* have been used as model for heterologous protein secretion (Nicaud *et al.,* 2002; Ogrydziak and Nicaud 2012). *Y. lipolytica* can secrete very large quantities of AEP (1-2 g per liter under appropriate inducing conditions) (Tobe *et al.,* 1976; Ogryziak, 1993). The AEP is synthesized as a prepro protein with a short pre-sequence followed by a stretch of 10 dipeptides (X-Ala or X-Pro) and a larger pro-region ending with the Lys-Arg recognition site of the Kex2-like endoprotease encoded by *XPR6* (Enderlin and Ogrydziak, 1994). The secretion signal of the extracellular lipase encoded by the *LIP2* gene present features similar to those of the XPR2 signal: short pre-sequence (13 amino acids) followed by four dipeptides X-Ala/X-Pro, a short pro-region (10 amino acids) and Lys-Arg (KR) cleavage site (Pignéde *et al.,* 2000). The XPR2 and LIP2 secretion signals have been extensively used to drive heterologous protein secretion in the *Y. lipolytica* host (Table 2.3).

The initial step of protein secretion from the cytoplasm to the ER can follow two pathways, either co-translational or post-translational translocation (Delic *et al.,* 2013). In the co-translation pathway, the signal recognition particles (SRP) bind to the ribosome, slowing down translation and then translation starts with concomitant translocation into the ER lumen with the implication of the chaperone. In *Y. lipolytica* co-translation is reportedly the predominant pathway and the glycosylation pattern in

this regard resembles that of mammalian high-mannose type glycosylation (Boisramé *et al*., 2002). This is in contrast to *S. cerevisiae* where the post-translational pathway is predominant (Boisramé *et al*., 2002). The similarity to co-translational translocation of proteins in mammalian cells highlight the potential of the *Y. lipolytica* as a suitable host for heterologous production of therapeutic proteins from mammalian origins and for therapeutic applications. In addition, display systems for heterologous protein localization to the yeast cell surface have been developed in *Y. lipolytica* using the cell wall proteins YlPir1 (Yuzbashev *et al*., 2012) and the carboxyl terminus anchor domain of YlCWP1 proteins (Yue *et al*., 2008). The constructed surface display systems have been described to have applications in fields as different as the immobilization of biocatalysts, bioconversion, bioremediation, live vaccine development and ultra-high-throughput screening for the identification of novel biocatalysts (Yue *et al*., 2008).

2.3. *Yarrowia lipolytica* host strains

The *Y. lipolytica* strains are very diverse and available for a variety of uses, and can allow either homologous or non-homologous integration of auto-cloning vectors, or integration of pBR322-based vectors. The *Y. lipolytica* Po1d, Po1f, Po1g, and Po1h are the most commonly used strains for the production of several heterologous proteins (Madzak *et al.,* 2000; 2001; Nicaud *et al.,* 2002; Swennen *et al.,* 2002; Laloi *et al.,* 2002; Gasmi *et al.,* 2010; Boonvitthya *et al.,* 2013). The characteristics of these strains are presented in Table 2.2.

Table 2.2: *Yarrowia lipolytica* host strains commonly used for heterologous protein production.

Y. lipolytica strains	Genotype	Phenotype	References
Po1d	*MatA, leu2-270, ura3-302, xpr2-322*	Leu$^-$, Ura$^-$, ΔAEP, Suc$^+$	Le Dall *et al.,* 1994
Po1f	*MatA, leu2-270, ura3-302,xpr2-322, axp1*	Leu$^-$, Ura$^-$, ΔAEP, ΔAXP, Suc$^+$	Madzak *et al.,* 2000
Po1g	*MatA, leu2-270, ura3-302::URA3, xpr2-322, axp-2*	Leu-, DAEP, DAXP, Suc+, pBR docking platform	Madzak *et al.,* 2000
Po1h	*MatA, ura3-302, xpr2-322, axp1-2*	Ura$^-$, AEP, AXP, Suc$^+$	Madzak *et al*., 2004

Notable is the availability of *Y. lipolytica* host strains deleted of genes encoding extracellular proteolytic activity to prevent proteolytic degradation during heterologous protein expressions (Table 2.2). The *Y. lipolytica* Po1g strain permits homologous recombination of the pBR322-based vectors to this plasmid DNA's docking platform recombinantly introduced into the yeast genome (Madzak *et al*., 2000). The *Y. lipolytica* Po1f and Po1h strains were designed for both homologous and random integration using either *rDNA* or *zeta* based vectors and are transformable into Leu or Ura prototrophy (Madzak *et al.,* 2000; Nicaud *et al.,* 2002).

2.4. Heterologous protein production in *Yarrowia lipolytica*

More than 100 heterologous proteins have been successfully produced in *Y. lipolytica* yeast expression systems (Madzak and Beckerich, 2013). A selection of some of the proteins heterologously produced in *Y. lipolytica* yeasts is presented in Table 2.3. The expressed proteins originated from the various phylogenetic origins such as fungi, bacteria, plants, animals and humans. The structural complexities of the proteins expressed in the *Y. lipolytica* yeast varied from simple industrial polypeptides such as the amylase, to complex therapeutic antibody proteins (Table 2.3). Many recombinant proteins of interest contain cysteine disulphide bond that are difficult to fold accurately

to yield an active protein. Although Muller *et al*., (1998) found *Y. lipolytica* to be the most efficient host in heterologous protein expression especially in performance reproducibility, there is no study in literature that has systematically investigated the general utility of the *Y. lipolytica* expression systems to argue for its prolificacy in expression of complex proteins in terms of disulphide bond content, size, structural and functional diversities. Therefore, succesful heterologous protein production in *Y. lipolytica* still remains to be determined empirically. However, the *Y. lipolytica* expression systems have been used successfully to catalyze oxidation processes for disulphide bonds formation for proper protein folding (Swennen *et al*., 2002, Gasmi *et al*., 2012). The *Y. lipolytica* systems co-expressing complex oxidation-reduction protein partners have also been reported in literature (Nthangeni *et al.*, 2004; Braun *et al*., 2012). The scales for heterologous protein production ranged from test tube cultures at levels of micrograms per liter to bioreactors yielding the biomass of 83 g/L dry cell weight and over 100 g/L of recombinant protein concentration (Kim *et al*., 2000).

2.5. Strategies to improve heterologous expression of functional protein in *Y. lipolytica*

Recombinant proteins could still be secreted at low levels even though the transcription or translation levels of the target protein is optimized for overexpression in the host system (Macauley-Patrick *et al.,* 2005; Porro *et al.,* 2005). As a result, many studies on *Y. lipolytica* expression systems have focused on systematic engineering of the yeast strains for effective protein secretion and post-translational modifications (Idiris *et al.,* 2010). The factors that affect the expression level of foreign genes that have been considered in literature include codon adaptation, gene co-expression, and genetic engineering of host strains.

Table 2.3: Examples of industrial and therapeutic recombinant proteins expressed in *Y. lipolytica.*

Proteins	Promoters and secretion signal	References
Industrial proteins		
Rhizopus oryzae lipase (ROL	XP2 promoter: hybrid of AEP and native pre-sequence	Yuzbasheva *et al.,* 2012
Aspergillus oryzae tyrosinase	HP4D: none	Rao *et al.,* 2011
Thermbifida fusca alpha amylase	LEU2 promoter: XPR2 prepo	Yang *et al.,* 2010
Aspergillus aculeatus Endo-1,4-β-mannanase	HP4D promoter: LIP2 and native secretion signal	Roth *et al.,* 2009
Rhodotorula araucaria epoxide hydrolase (EH)	HP4D: none	Maharajh *et al.,* 2008
T. versicolor laccase IIIb	HP4D:native	Jolivalt *et al.,* 2005
T. versicolor laccase IIIb	TEF: native	Theerachat *et al.,* 2012
Therapeutic Proteins		
Co-expression of Human cytochromes P450 2D6 and 3A4 genes together with human cytochrome P450 reductase (hCPR) or *Y. lipolytica* P450 reductase (YlCPR)	ICL1 promoter:none	Braun *et al.,* 2012
Interferon alpha	POX2 promoter: LIP2 prepro	Gasmi *et al.,* 2011
Cytochrome P450 1A1	POX2 promoter:none	Nthangeni *et al.,* 2004
Anti-Ras single chain antibody	HP4D promoter: XPR2 pre sequence	Swennen *et al.,* 2002.
scFv (30 kDa)	HP4D promoter:XPR2 prepro	Swennen *et al.,* 2002
Epidermal growth factor	XPR2 promoter: XPR2 pre sequence	Hamsa *et al.,*1998
Blood coagulation factor XIIIa	XPR2 promoter:XPR2 prepro	Tharaud *et al.,* (1992)
Alternaria alternata recombinant Alta1p allergen	POX2 promoter: LIP2 prepro	Gasmi *et al.,* 2012
Human granulocyte colony stimulating factor	POX2 promoter: LIP2 prepro	Gasmi *et al.,* 2012
Tissue plasmogen activator	XPR2 promoter:XPR2 prepro	Pfizer (USA)

2.5.1. Codon optimization

Rare codons have been considered as a significant hindrance for foreign protein expression (Outchkourov *et al.,* 2002; Lanza *et al.,* 2014). The rare codons in foreign genes are usually translated by the host at very slow rate, causing ribosomal pausing and disturbing polypeptide translation and elongation, in-frame deletion of some amino acids, and reduction of the amount of the produced protein (Hartfield and Roth, 2007; Gustafsson *et al.,* 2004). To overcome this hindrance, strategies such as synonymous codon usage bias analysis prior to heterologous protein expression (Zhao *et al.,* 2000; Outchkourov *et al.,* 2002; Sinclair and Choy, 2002; Lü *et al*, 2005), and supplementation of rare tRNAs (Hatfield and Roth, 2007; Gustafsson *et al.,* 2004) have been considered. Recently, the effect of synonymous codon usage bias and the consensus ATG for translation initiation for enhancement of protein expression in *Y. lipolytica* has been investigated using hIFN-α2b protein (Gasmi *et al.,* 2011). Codon optimization resulted in a 11-fold increase in hIFN-α2b protein production, and the insertion of CACA sequence upstream of the initiation codon of IFN-optimized construct resulted in 16.5-fold increase of the expression level (Gasmi *et al.,* 2011).

2.5.2. Gene co-expressions in *Yarrowia lipolytica*

The availability of *Y. lipolytica* strains with more than one auxotrophic genotypes has made the yeast amenable for protein co-expressions. Several investigations on gene co-expression in the yeast *Y. lipolytica* have been reported to enhance production and activity of the target heterologous protein. For example, the optimal activity of human cytochrome P450 was attained by co-expression with cytochrome P450 reductase enzyme (Nthangeni *et al.,* 2004; Braun *et al.,* 2012). The similar approach was used to improve production of trans-10, cis-12 conjugated linoleic acid (CLA), by co-expression of the delta 12-desaturase gene from *Mortierella alpina* together with the linoleic acid isomerase gene from *Propionibacterium acne* (Zhang *et al*., 2012; Zhang *et al.,* 2013). The co-expression of heterologous genes with genes encoding molecular chaperones and foldases play an important role in *in vivo* folding, assembling and secretion of proteins in endoplasmic reticulum (Gasser *et al.,* 2008). A eukaryotic protein disulfide isomerase catalyzes the protein cysteine oxidation and disulfide bond isomerization

and also exhibits chaperone activity (Wilkinson and Gilbert, 2004). The strategy to increase activity of the cysteine rich proteins was demonstrated in *P. pastoris* (Li *et al.,* 2010). The strain of *P. pastoris* that overproduces protein disulphide isomerase transformed with the expression vector containing multiple copies of human secretory leukocyte protease inhibitor (SLPI) gene resulting in enhanced SLPI specific activity (Li *et al.,* 2010). However, there have been no reports in literature of *Y. lipolytica* expression platforms with exogenous disulphide isomerases, foldases or chaperonic gene co-expressions.

2.5.3. Genetic engineering of *Yarrowia lipolytica* for enhanced heterologous protein production

Apart from the availability of *Y. lipolytica* host strains deficient in the biosynthesis of major proteolytic enzymes to prevent heterologous protein degradation, the yeast has been engineered to have strains with enhanced protein production and secretion capacity and also as host for production of proteins with modified post-translational modifications (Madzak *et al*., 2004; De Pourcq *et a*l., 2012; Zang *et al*., 2013). The yeast has been engineered for enhanced expression of membrane proteins by deleting the phosphatidic acid phosphatase, which led to improvement in membrane protein quantity and quality in terms of proper protein folding and biological activity (Guerfal *et al*., 2013). The *N*-linked glycosylation and the glycan profile of proteins is one of the most common post-translational modifications determining biological activity, pharmacokinetics, protein clearance, and immunogenicity (Li and d'Anjou, 2009). The nature of the desired *N*-linked glycosylation profile of a heterologous protein is a crucial parameter in selecting the host for production of therapeutic proteins for human applications. The *Y. lipolytica* yeast is currently the subject of genetic engineering studies to establish strains capable of producing glycoproteins with humanized glycans (Park *et al*., 2011; De Pourcq *et al*., 2012). The *Y. lipolytica* glycosylated proteins contain a single core *N*-linked oligosaccharide chain and unlike in other yeasts the degree of hypermannosylation is low (Barnay-Verdier *et al*., 2004; Park, *et al*., 2011). The production of heterologous glycoproteins that are homogenously glycosylated with either $Man_8GlcNAc_2$ or $Man_5GlcNAc_2$ *N*-glycans has been reported (De Pourcq *et al*., 2012). This platform expanded the utility of *Y. lipolytica* as a heterologous expression

host and makes it possible to produce glycoproteins with homogeneously glycosylated *N*-glycans of the human high-mannose-type. Although the complete glycoengineering of *Y. lipolytica* for production of humanized glycoproteins is several steps away, its successful completion will greatly broaden the application scope of the yeast as host for human therapeutic protein production. The development by chemical mutagenesis of *Y. lipolytica* gene with enhanced abilities to produce endogenous extracellular proteins has been described, suggesting the potential use of such strains as host for heterologous protein production (Fickers *at al*., 2003; Ghezelbash *et al.,* 2014). A *Y. lipolytica* strain deleted for the GPI7 strain has been constructed and demonstrated to have enhanced endogenous and heterologous protein production capacity (Ramagoma, 2011).

2.6. Concluding remarks

Yarrowia lipolytica, a GRAS and nonconventional yeast has been presented as an attractive host for the production of industrial and therapeutic recombinant proteins. The availability of a wide range of tools such as, single and multicopy expression vectors, highly transcribed inducible and constitutive promoters, dominant and auxotrophic selection markers, protein localization signals, in-depth knowledge concerning genetics, physiology, and biochemistry as well as genetic engineering and fermentation technologies has propelled this yeast as an attractive host for heterologous protein production. The inherent ability of this yeast to secrete a variety of proteins via a cotranslational translocation similar to that of mammalian systems, low overglycosylation, high secretion efficiency, good product yield, and performance reproducibility are additional features that make *Y. lipolytica* attractive as a host for heterologous protein production. A wide variety of simple and complex proteins from phylogenetically diverse origins have been successfully expressed in the *Y. lipolytica* yeast. The attributes associated with *Y. lipolytica* make it worthy to explore this yeast for heterologous expression and secretion of the *C. fumago* chloroperoxidase and human Granulocyte-Colony Stimulating Factor proteins. The proteins require post-translation glycosylation modifications, contain intra molecular disulfide bonds required for proper folding and biological activity and have presented challenges when heterologously expressed in other yeast and bacterial expression systems.

CHAPTER 3

Cloning and expression of *Caldariomyces fumago* chloroperoxidase gene in *Yarrowia lipolytica*

3.1. INTRODUCTION

The chloride hydrogen-peroxide oxidoreductase (EC 1.11.1.10) or chloroperoxidase enzyme from *C. fumago* was discovered as a peroxidative chlorination catalyst involved in the biosynthesis of caldariomycin (Morris and Hager, 1966). The *C. fumago* chloroperoxidase (CPO) enzyme is secreted to the extracellular medium by the *C. fumago* fungus as a heavily glycosylated glycoprotein of 40-42 kDa containing 25-30% carbohydrate content (Morris and Hager, 1966). It is regarded as one of the most diverse of the known heme enzyme catalysts due to the versatility of the reactions that it catalyses (Sundaramoorthy *et al.,* 1998). In addition to its biological function as a peroxide-dependent chlorinating enzyme, CPO also acts as a cytochrome P450 enzyme and a potent catalase (Yi, *et al*., 1999). CPO behaves as a catalase in terms of catalyzing the dismutation of hydrogen peroxide and the oxidation of alcohol (Yi *et al*., 1999). It mimics cytochrome P450s in catalysing heteroatom dealkylation (Kedderis *et al*., 1980; Kedderis and Hollenberg, 1984, Yi *et al*., 1999), benzylic hydroxylations (Miller *et al*., 1995; Yi *et al*., 1999) and oxygen transfer to alkenes (Allian *et al*., 1993; Hu and Hager, 1999; Yi *et al*., 1999), alkynes (Hu and Hager, 1998; Hu and Hager, 1999, Yi *et al*., 1999), sulfides (Colonna *et al*., 1990; Casella *et al*., 1992; Yi *et al*., 1999), and arylamines (Kedderis *et al*., 1986; Doerge and Corbett, 1991; Yi *et al*., 1999). It is noteworthy that the oxidation reaction with CPO is not dependent on co-factors such as NAD(P)H or other electron donor as in cytochrome P450 catalyzed reactions, but it involves hydrogen peroxide (H_2O_2) or hydroxyl radicals (ROOH) (Sundaramoorthy *et al*., 1995). The CPO enzyme is especially adept in the stereoselective epoxidation of alkenes (Lakner and Hager, 1996; Lakner *et al*., 1998; Hu and Hager 1999; Yi *et al*., 1999) hydroxylation of alkynes (Hu and Hager, 1998; Hu and Hager, 1999; Yi *et al*., 1999) and in the production of chiral sulfoxides (Colonna *et al*., 1990; Yi *et al*., 1999).

The elucidation of the structure of the CPO has provided some insights into how the CPO carries out the diverse reactions (Sundaramoorthy *et al.,* 1995). The structural characteristics of the unique active site of CPO have been postulated to account for the versatility of CPO as an oxidation catalyst (Yi *et al*., 1999). Even though CPO folds into a unique structure that does not resemble either a peroxidase or a cytochrome P450, CPO does share certain structural features with both enzymes. Like the cytochrome P450, CPO has a cysteine-thiolate group as the proximal heme ligand (Sundaramoorthy *et al.,* 1998). Moreover, the local hydrogen-bonding environment involving the cysteine sulphur is the same in both the cytochrome P450s and CPO (Sundaramoorthy *et al.,* 1998). As in heme peroxidases, but unlike cytochrome P450s, the distal heme surface that forms the peroxide binding site in CPO is polar (Sundaramoorthy *et al.,* 1998). The CPO also is unique in the hemoprotein family by virtue of having a glutamic acid residue distal to the heme-iron (Yi *et al*., 1999). It has been postulated that this glutamic acid residue functions as an inflexible acid-base catalyst and plays a mechanistic role similar to that played by the distal histidine in traditional peroxidase chemistry (Sundaramoorthy *et al*., 1998; Yi *et al*., 2003).

The molecular genetics of the *C. fumago* CPO has been extensively characterised. The gene encoding *C. fumago* CPO has been cloned and sequenced (Nuell *et al.,* 1988). The gene codes for a protein of 352 amino acids with a calculated molecular mass of 42 kDa. A 21-aa signal sequence is removed from the CPO precursor and the resulting *N*-terminal glutamic acid is converted into a pyroglutamic acid. The wild-type CPO contains a relatively high degree of posttranslational modifications (Kenigsberg *et al.,* 1987). Carbohydrate moieties are incorporated at *N*- and *O*-glycosylation sites. Different isozymes are found in the extracellular medium of *C. fumago* cultures (Sae *et al.,* 1979; Pickard and Hashimoto, 1982) which are indicated to be glycosylation variants of the same polypeptide (Kenigsberg *et al.,* 1987). In the major CPO isozyme, the most heavily glycosylated, three *N*- and eleven *O*-glycosylation sites are occupied with a total of 21 sugar groups (Kenigsberg *et al.,* 1987; Sundaramoorthy *et al.,* 1995). Other modifications include deamidation of three amidic residues to the corresponding acids (Kenigsberg *et al.,* 1987), one important disulphide bond (Blanke and Hager, 1988) and the incorporation of a protoporphyrin XI (haem b) molecule (Sundaramoorthy *et al.,* 1995). The crystal structure of the CPO protein also revealed

that the last 52 aa of the primary translation product are lacking in the mature protein, suggesting that CPO undergoes proteolytic processing at the carboxyl terminus with the processing apparently occurring downstream of a dibasic Lys-Arg (KR) processing site (Sundaramoorthy *et al.,* 1995). The function of the carboxyl terminal propeptide of CPO is considered to be in the maturation of the CPO protein during biosynthesis (Conesa *et al*., 2001).

The mechanism of CPO catalysis has been described (Figure 3.1). In resting state, native ferric enzyme has a ferric protoporphyrin IX. Binding of hydrogen peroxide at the heme iron (III) initiates the reaction. Two electrons are then transferred from the heme center to hydrogen peroxide, resulting in the heterolytic cleavage of the peroxide bond and the formation of a water molecule and activated heme iron (IV) oxo-ferryl porphyrin cation radical [heme (Fe^{4+}=O)·+] intermediate which is compound I [(Step 3, Figure 3.1)]. High valent iron (IV) oxo is the key reactive intermediate in the catalytic cycle of oxygen activating thiolate heme-peroxidase (Groves *et al.,* 2006). In the chlorination reactions (Liby *et al.,* 1982), compound I interacts with chloride ion to form a ferric complex compound X (heme Fe^{3+}–O–X), which is unstable and can easily decompose to the resting state of the enzyme, halonium ion (X+) and hydroxyl ion (OH^-) can react to form hypohalous acid (HOX). In the catalase reactions, compound I interact with the second molecule of hydrogen peroxide, to form dioxygen and reduce heme iron to the ferric resting state (Sun *et al.,* 1994). In peroxidation reactions, compound I convert organic molecules (HA) to radical products (A·) and be reduced to compound II by transferring one electron to the organic substrate (Nakajima *et al.,* 1985).

The *C. fumago* CPO enzyme has applications in industrial chemical synthesis, chiral synthons for drugs, and the detection and inactivation of chemical warfare agents (Pickard, *et al*., 1991; Hager, 1998; Prokop *et al*., 2006). CPO is capable of regio- and enantioselective oxygenations and halogenations of organic substrates. The halogenation is the first type of reaction catalysed by CPO, hydrogen peroxide is used as the initiator, and halides (Cl, Br^-) and susceptible compounds serve as the substrates. A variety of organic compounds and structures are susceptible to halogenation by CPO, and these substrates include phenols (Wannstedt *et al.,* 1990),

polycyclic hydrocarbons (Niu and Yu, 2004), and flavonoids (Yaipakdee and Robertsons, 2001). The halogenations reactions catalyzed by CPO are shown in Figure 3.2.

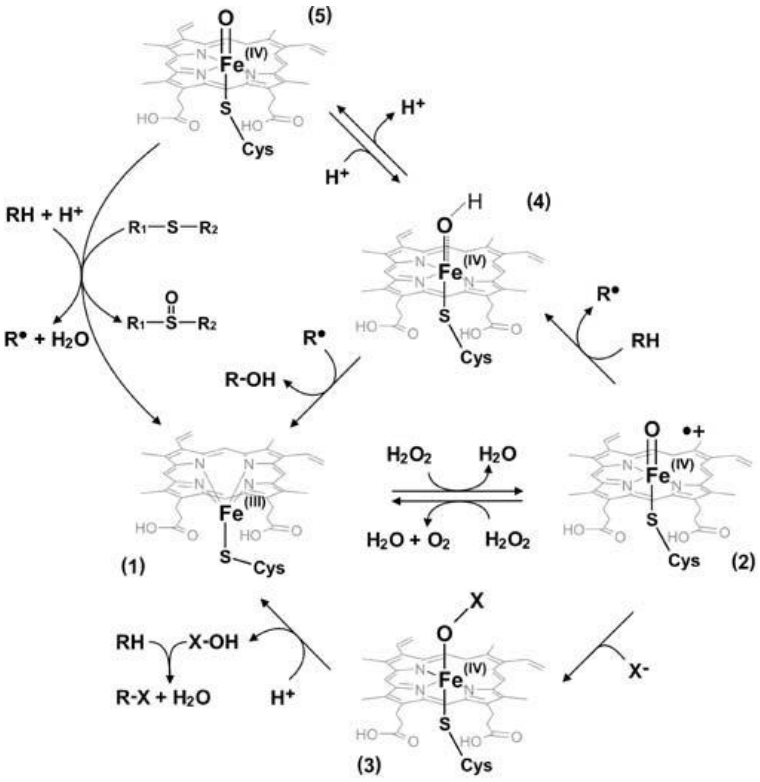

Figure 3.1: Mechanism of CPO-catalyzed reactions. RH represents the substrate, Compounds I and II represent the ferryl intermediates; X represents halides involved in the halogenation pathway (Andersson and Dawson, 1991; Green *et al*., 2004). (1) Resting state ferric enzyme, (2) compound I, (3) compound X releases hypohalous acid (HOX), (4) protonated compound II, (5) compound II.

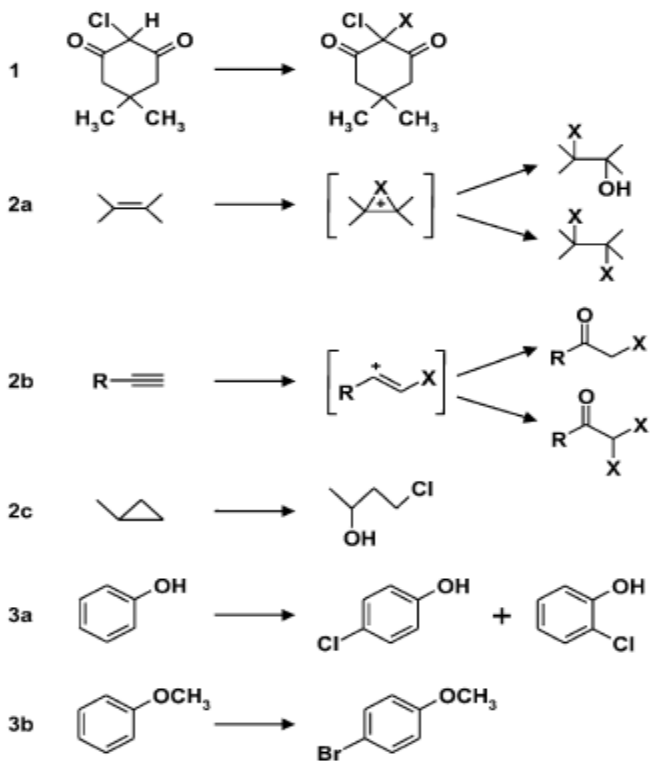

Figure 3.2: Halogenation of organic compound substrates catalyzed by *C. fumago* CPO. Chlorination of alicyclic ketones such as β-diketones reaction 1 mostly as monochlorodimedone used to assay haloperoxidases (Morris and Hager 1966), halogenation of alkanes 2a (Geigert *et al.,* 1983c), alkynes 2b (Geigert *et al.,* 1983a, c), halogenation of cycloalkanes eg methylcyclopropane (Geigert *et al.,* 1983a). Halogenation of phenols 3a (Wannstedt *et al.,* 1990), anisol 3b (Pickard *et al.,* 1991).

The dehalogenation of halophenols by CPO leads to its use as a bioremediation catalyst for aromatic dehalogenation reactions (Osborne *et al.,* 2006). Effective oxidation of chlorinated phenol and aromatic hydrocarbons has been performed using the CPO enzyme and hydrogen peroxide as a co-substrate (Vázquez-Duhalt *et al.,* 2001; La Rotta and Bon, 2002). This application resulted in the application of CPO as detoxification agent of the otherwise toxic, mutagenic and carcinogenic chlorophenol derivatives commonly found in pesticide preparations (Olaniran and Igbinosa, 2011).

The oxidation reactions catalyzed by CPO are depicted in Figure 3.3. The reactions are important for chiral compounds synthesis. The greatest industrial interest is in the

synthesis of enantioselective products such as cyclopropylmethanols (Hu and Dordick 1992, Hager, 1998), chiral epoxidation of olefins (Allain *et al.,* 1993), epoxidation oxidation of sulfides (Colonna *et al.,* 1990) and hydroxylation of alkynes (Hu and Hager, 1998).

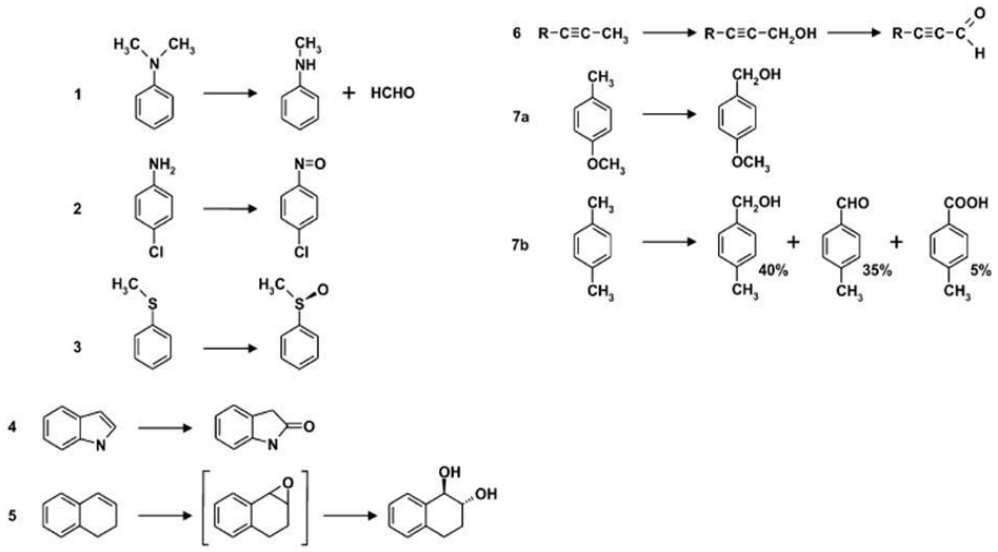

Figure 3.3: Oxidation reactions catalyzed by CPO. Oxidation reaction 1 demethylation of N,N-dimethylaniline (Kedderis *et al.,* 1980); 2 oxidation of the amino group of chloroaniline into nitroso compound (Corbett *et al.,* 1978, Doerge and Corbett, 1991); 3 enantioselective sulfoxidation of thioanisole into (R)-sulfoxide (van de Velde *et al.,* 2001); 4 indole oxidation into oxindole (Seelbach and Kragl, 1997); 5 epoxidation/hydroxylation of 1,2-dihydronaphthalene (Sanfilippo *et al.,* 2004); 6 propargylic hydroxylation (Hu and Hager, 1998); 7a benzylic hydroxylation of p-methylanisole (Miller *et al.,* 1995); 7b selective hydroxylation and subsequent oxidation of one methyl group in p-xylene (Morgan *et al.,* 2002)

The widespread adoption of enzyme-catalyzed synthetic strategies using *C. fumago* CPO is hindered by the high cost of purified proteins. Numerous attempts to produce the *C. fumago* CPO recombinantly have been described in literature. The expression of CPO in *E. coli* had inadequate success for synthesis of the active form of chloroperoxidase (Zong *et al*., 1995). The enzyme was obtained in a non-glycosylated apo-form without the heme being incorporated into the enzyme (Zong *et al*., 1995). Only under high pressure, was the protein refolded with heme to generate an active

enzyme in very low yields (Zong *et al*., 1995). Eukaryotic expression systems have been considered as heterologous production hosts for the CPO protein. The expression of CPO using the baculovirus system resulted in the production of extracellular inactive CPO, which could not be reconstituted to active protein (Conesa *et al*., 2001a). The attempts to express CPO in the yeasts *S. cerevisiae* and *P. pastoris* also did not lead to the production of active protein (Zong *et al*., 1995; Conesa *et al*., 2001a). The expression of a recombinant CPO in the parental host *C. fumago* has been reported (Yi *et al.,* 1999). However, the presence of residual wild type CPO complicated the selection of recombinant CPO from native enzyme for further investigations (Yi *et al.,* 1999, Conesa *et al*., 2001a). The attempts to increase CPO production in *C. fumago* also focused on the generation and isolation of mutant strains with superior production capacities for CPO activity (Buchhaupt, 2011). The only report on successful heterologous expression of the CPO was with the use of the filamentous fungus *A. niger* as host (Conesa *et al.,* 2001a). Although the recombinant enzyme was overglycosylated, the excess of glycosyl groups did not have a major effect on enzyme properties (Conesa *et al.,* 2001a).

The aim of the study was to investigate heterologous expression and secretion of *C. fumago* chloroperoxidase metalloprotein in *Y. lipolytica*. The *Y. lipolytica* expression system was chosen because of its advanced genetic tools (Madzak *et al*., 2000), high capacities to secrete proteins to the extracellular space (Madzak *et al.,* 2004), and effectiveness in heterologous expression of complex cytochrome P450s (Nthangeni *et al.*, 2004; Shiningavamwe *et al*., 2006; Braun *et al.,* 2012) that share catalytic and structural characteristics with *C. fumago* CPO (Sundaramoorthy *et al.,* 1995). The successful expression of active CPO enzyme will enable development of cost-effective production processes leading to adoption as industrial catalyst.

3.2. MATERIAL AND METHODS

3.2.1. General chemicals, reagents, kits and enzymes.

Yeast extract was purchased from Oxoid (Basingstoke, Hampshire, England), peptone (Belton, Dickison, Le Pont de Claix, France), ammonium sulphate, casamino acids and yeast nitrogen base (YNB) were purchased from Difco Laboratories (Detroit, MI, USA), restriction endonucleases, lambda (λ) DNA, Taq DNA Polymerase, deoxyribonucleotide triphosphates (dNTPs), and 100 bp and 1 kb DNA molecular weight markers were supplied by New England Biolabs (NEB) (Beverley, MA, USA). Expand HiFi DNA polymerase and were purchased from Roche (Mannheim GmbH, Germany). Oligonucleotides were purchased from Whitehead Scientific (Cape Town, South Africa). The DNA Ladder, T4 DNA ligase and pGemT-Easy vector were obtained from Promega (Madison, USA). Protein molecular weight markers were supplied by Thermo scientific (Burlington, Canada). The PCR purification, DNA isolation and gel band extraction kits were supplied by BioFlux (Bioer Technology Co. Ltd, Japan). The agarose gel matrix, commercial *C. fumago* chloroperoxidase, hemin and thionin were obtained from Sigma Aldrich (St. Louis, USA). Unless stated otherwise all other chemicals were from Merck (Pty) Ltd (Modderfontein, South Africa).

3.2.2. Plasmids, microbial strains and cultivation conditions

The original plasmids used in this study are described in Table 3.1. The plasmids were propagated in *E. coli* XL-10 Gold cells (Table 3.2). The *E. coli* cells transformed with plasmid DNA were propagated by growing at 37°C in Luria Bertani medium (g/L; yeast extract 5g, tryptone 10g and sodium chloride 10g) supplemented with appropriate antibiotic; 50 µg/ml ampicillin or 100 µg/ml kanamycin. The *E. coli* cells transformed with plasmid DNA were selected on LB agar medium containing appropriate concentration of the antibiotic. The AIX LB agar media for blue/white screening of colonies containing recombinant pGemT-Easy plasmids contained ampicillin, isopropyl-β-D-thiogalactonidase (IPTG) and 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) to final concentrations of 0.2 mM and 40 µg/ml respectively. The *Y. lipolytica* Po1f strain (Table 3.2.) was used as host for heterologous CPO expression. The yeast strain was maintained on YPD medium (10 g yeast extract, 20 g peptone and 20 g

glucose; per litre of distilled water) at 28°C. The *Y. lipolytica* yeast transformed with the desired expression cassette was selected on YNB casa media (1.7g YNB without ammonium sulphate and amino acids, 4g Ammonium Chloride, 10 g glucose, 2 g casamino acids and 15 g agar per litre distilled water) and the pH of the media was adjusted to 6.5 using sodium hydroxide. A single colony was used to prepare a 5 ml overnight preinoculum. The preinoculum was used to inoculate 50 ml YPD and the culture was grown at 28°C on a shaker at 250 rpm. The solid media for culturing yeast or bacterial cells contained 15 g/L agar.

Table 3.1. Original plasmid DNA used in the study.

Plasmid	Description	Reference
pGem-T Easy	Blue/white screening using AIX agar plates, AmpR	Promega (Madison, USA)
pGA4-CPO	Plasmid carrying customized *C. fumago* CPO gene	This study
JMP62	Source for POX2 promoter	Nicaud *et al.,* 2002
pINA1291	Source for HP4D promoter	Nicaud *et al.,* 2002
pINA1293	Multicopy integrative shuttle vector containing the synthetic promoter HP4D, *Y. lipolytica* signal peptide, Kan^R and *Ura3d4* selective marker, random integration into Po1f genome through *ZETA* transposable elements.	Nicaud *et al.,* 2002
pKOV96	Source of TEF promoter and was used for subcloning of promoters (POX2, HP4D) and secretion signals (LIP2 and LACC)	Labuschagne and Albertyn 2007

Table 3.2: Microbial Strains used in the study.

Strains	Genotype	Phenotype	Reference
E. coli XL10 Gold	endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 tetR F'[proAB lacIqZΔM15 Tn10(TetR Amy CmR)]		Stratagene (Califonia, USA)
Y. lipolytica Po1f	MatA, leu2-270, ura3-302, xpr2-322, axp-2	Leu-, Ura-, ΔAEP, ΔAXP, Suc+	Madzak *et al.,* 2000

3.2.3. Recombinant DNA techniques

3.2.3.1. Polymerase Chain Reaction and conditions

The PCR was performed according to the specifications of the manufacturer using either Expand High Fidelity ^{Plus} PCR system or Taq DNA polymerase. The PCR was performed in a total reaction mixture of 50 µL consisting of the template DNA, 200 µM dNTP mix (dTUP, dATP, dCTP, dGTP), 5 µL of the 10X concentrated buffer with $MgCl_2$, 0.4 µM of each upstream and downstream primer, and 1.0U of the DNA polymerase. The PCR was performed in an MJ Mini Personal Thermal Cycler (Bio-Rad Laboratories GmbH, California, USA). The following general conditions were used for PCR amplifications: initial denaturation cycle at 94°C for 2 min followed by 30 cycles of denaturation (94ºC for 30 seconds, 55-60 °C for 30 seconds and extension 72°C for 1 min per 1000 bp long PCR product target. The final extension was done for 7 min at 72 °C. The annealing temperature was determined by the melting temperature of the specific primers used (Table 3.3). The PCR product was purified using the BioFlux PCR purification kit before further manipulation.

Table 3.3: List of oligonucleotide primers used in the study.

Oligonucleotide	sequence(5'-3')	Restriction sites	Amplicon
*Hind*IIICPO-F	**AAGCTT**GAGCCCGGCTCTGGCATCGGCTA	*Hind*III	mature CPO
CPO*Avr*II-R	**CCTAGG**CTATTAATGATGATGATGATGATGATGTTTCGGGCCTTG	*Avr*II	mature CPO with sequence encoding 6XHis tag
LIP2SP-F	**CGGATCC**ATGTCCACCATCCTTTTCACA	*BamH*I	LACC secretion signal
LIP2SP-R	**GGGTACCAAGCTT**TCGCTTCTGCAGAACTG	*Acc*651/*Hind*III	LIP2 secretion signal
LaccSP-F	**CGGATCC**ATGTCGAGGTTTCACTCTCTTCT	*BamH*I	LIP2 secretion signal
LaccSP-R	**GGGTACCAAGCTT**AGCGTGGGCCACAG	*Acc*651/*Hind*III	LACC secretion signal

The restriction sites introduced are indicated in bold, and the sequence encoding the 6X Histidine tag sequence is underlined.

3.2.3.2. DNA restriction digestion, ligations and cloning

The digestion of DNA with restriction enzymes was performed in a total reaction mixture of 10 µL consisting of target DNA (50-100 ng), 0.2 U/µg of restriction endonucleases in appropriate buffer volume. DNA ligations were done using T4 DNA ligase generally using a ligation ratio of 4:1 insert:plasmid vector backbone followed by incubation overnight at 4 °C. The ligation mixture was used to transform 50 µL aliquots of *E. coli* XL-10 Gold competent cells.

3.2.3.3. Preparation of competent *Escherichia coli* cells and transformation

The *E. coli* X10 Gold competent cells were prepared using rubidium/calcium chloride method (Hanahan, 1985). Briefly, single colony of *E. coli* Gold XL10 Gold cells was inoculated into 5ml LB media as pre-inoculum and incubated for 16 hours at 37 °C with shaking at 250 rpm. Hundred millilitres (100 ml) of Psi broth (20 g tryptone, 5 g yeast extract, 5 g magnesium sulphate per litre, pH 7, 6) was inoculated with 1ml of 16 hour old pre-inoculum and grown at 37 °C to an OD_{550nm} of 0.48. The culture was placed on ice for 15 min and the cells were harvested by centrifugation at 5000 rpm for 5 min. The harvested pellet was resuspended in TFBI buffer (2 manganese chloride; 2.42 g Rubidium chloride; 0.294 g Calcium chloride; 0.58 g Potassium acetate; 30 ml glycerol per 200 ml; pH 5.8) and placed on ice for 15 min. The cells were centrifuged at 5000 rpm for 5 min. The collected pellet was resuspended in minimal TFBII buffer (0.12 g MOPS; 1.1 g calcium chloride; 0.121 g rubidium chloride; 15 ml glycerol per 100 ml; pH 6.5). Fifty microlitre aliquots of competent cells were prepared and stored at -70ºC.

The transformation of was performed by mixing an appropriate aliquot of plasmid DNA or the ligation mixture with the thawed competent *E. coli* XL10 Gold cells followed by incubation on ice for 20 min, then heat shocked at 42 °C for 45 seconds and placed on ice for 10 min. The LB media (750 µL) was added into the mixture and incubated at 37 °C for an hour. The transformed cells were streaked on LB or AIX agar plates supplemented with an appropriate antibiotic followed by overnight incubation at 37°C.

3.2.4. Plasmid isolation

The plasmid DNA isolations were done using 2-5ml of overnight bacterial culture. Plasmid DNA isolations were done using the lysozyme boiling method (Sambrook *et al*., 1989). Briefly, a single colony was inoculated into 5ml LB medium containing appropriate antibiotics. The resultant cell culture was centrifuged at 12000 rpm to harvest the cells. The cell pellet was re-suspended in 400 µL of STET buffer (50 mM EDTA, 50 mM Tris-HCl pH 8.0, 0.1% v/v Triton X-100, 8% w/v sucrose). To lyse the cells, 10 µL of lysozyme (10mg/ml) was added to the cell suspension and the mixture was incubated at 37 °C for 20 min and thereafter immersed in boiling water for 60 seconds. The lysate was cooled on ice for 10 min; the cell debris was separated from the supernatant by centrifugation at 12000 rpm for 15 min. The supernatant was transferred into a new microcentrifuge tube and the plasmid DNA was precipitated using 400 µL cold isopropanol and the contents were mixed by inversion. The DNA was harvested by centrifuging for 15 min at 12000 rpm. The collected DNA was washed with 1 ml of 70% ethanol and spun for 5 min at 12000 rpm. The supernatant was discarded and the ethanol residual from the pellet was removed by air drying. Precipitated DNA was re-suspended in 100 µL TE buffer (10mM Tris-HCl, 1 mM EDTA, pH 8.0) containing 10 µL of RNAse (10 mg/ml). When highly purified plasmid DNA isolations were required, the BioFlux plasmid preparation kit was used in accordance with the manufacturer's instructions.

3.2.5. Agarose Gel Electrophoresis

The DNA electrophoresis was done at 90 V/cm in agarose gel (0.8-1.5 %) dissolved in 1X Tris-acetate EDTA [242 g Tris, 50 mM EDTA and 57.1 ml glacial acetic, per litre for a 50X buffer, pH 8.0]. Ethidium bromide (1 mg/ml) stain was added and the samples were mixed with loading buffer (10 v/v% bromophenol blue in 50 v/v% glycerol solution). The electrophoresed gels were visualised under transillumination radiation using a Gel documentation system (Syngene G-Box gel System, Vacutec, South Africa). The Dark Reader transilluminator DR-45M was used to visualise agarose gel and excision of DNA band for further cloning purposes. The purification of DNA from agarose gel was done using the Bioflux PCR and Gel purification kit according to manufactures instructions.

3.2.6. DNA sequencing

The sequences of the cloned DNA fragments were confirmed by DNA sequencing following PCR amplification and sublconing procedures. DNA sequencing was done at Inqaba Biotechnical Industries (Pretoria, South Africa). The DNA fragments cloned in pGemT-Easy vector were sequenced using T7 promoter and T7 terminator primers. DNA sequence analysis and protein alignments were performed using the ClustalW alignment tool (http://www.ebi.ac.uk./Tools/clustalw2/index.html), the NCBI BLAST search (http://www.ncbi.nlm.nih.gov/blast/BLAST.cgi) and Expasy tools (http://www.expasy.org/tools/).

3.2.7. Transformation of *Yarrowia lipolytica* cells

The *Y. lipolytica* yeast was transformed by the lithium acetate method as previously described (Labuschagne and Albertyn, 2007). Briefly, two day old cells of *Y. lipolytica* Po1f from a YPD agar plate were used to inoculate 25 ml YPD media supplemented with 50 mM citrate buffer (pH 4) as a pre-culture and incubated at 28 °C for 15 hours on a rotary shaker (180 rpm). When a concentration of 10^8 cells/ml was reached, cells were harvested via centrifugation for 10 min at 4000 *g* in Sovall RC 5B plus centrifuge. Cells were rinsed twice with TE buffer (10 mM Tris, 1 mM EDTA, pH 8), after which they were resuspended in 0.1 M LiAc (pH 6.0) to a final concentration of $5x10^7$ cfu/ml. The cells were transferred to a sterile 50 ml flask and incubated at 28 °C for 1 hour with gentle shaking. The cells were then centrifuged for 5 min at 2500 x*g*. The competent cells were resuspended in 0.1 M LiAc (pH 6) to a final concentration of $5x10^8$ cells/ml, before being aliquoted in 100 µL aliquots in 2 ml micro-centrifuge tubes. The 5 µL carrier DNA (10 mg/ml salmon sperm) and 10 µL of vector devoid of bacterial DNA were added to the 100 µL competent cells with gentle mixing. The transformation mixture comprised of *Not*1 linearized expression vector (about 0.5 µg) was incubated in a 28 °C water bath for 15 min. After the incubation, 700 µL of 40% PEG 4000 in 0.1 M LiAc (pH 6.0) was added and incubated for an hour at 28°C with gentle shaking. Cells were heat shocked for 10 min at 39 °C before 1.2 ml of 0.1 M LiAc (pH 6) was added. The cells were then centrifuged at 2500 x*g* for 5 min. Cells were resuspended in 500 µL LiAc and 200 µL plated on YNB casa agar plates. The plates were incubated for up to three weeks at 28 ^{o}C.

3.2.8. Genomic DNA isolation

The total genomic DNA was extracted from *Y. lipoytica* or *T. versicolar* as previously described (Labuschagne and Albertyn, 2007). Briefly, a single colony was inoculated into 5 ml YPD broth and was incubated overnight at 28°C with shaking at 250 rpm. Cells from 5 ml culture were harvested by centrifugation at maximum speed of 12000 rpm. The cells were re-suspended in 500 µL lysis buffer (100 mM Tris-HCl pH 8.0, 50 mM EDTA, 1% SDS-PAGE gel) and 200 µL glass beads (425–600µm) and vigorously vortexed for 4 min followed by cooling on ice for 2 min. Ammonium acetate 7 M (pH 7.0) was added, tubes were immersed in 65 °C water bath for 5 min and quickly placed on ice for 5 min. Chloroform 500 µL was added, vortexed followed by centrifugation (10000 rpm) for 2 min at 4ºC. The supernatant was transferred into a new 1.5 ml Eppendorf tube and the DNA was precipitated by adding 750 µL isopropanol. This was followed by centrifugation for 2 min at 10 000 rpm at 4°C. The pellet was washed with ice-cold 70 % ethanol, dried and dissolved in 100 µL TE buffer (10 mM Tris-HCl, 1mM EDTA, pH 8.0).

3.2.9. Construction of expression vectors

3.2.9.1 Cloning of the TEF, HP4D and POX2 promoter elements

The pKOV96 plasmid contains the TEF promoter (herewith denoted pKOV96-TEF) element flanked by *Cla*1 and *Bam*HI/*Acc*651 restriction sites (Labuschagne and Albertyn, 2007). To create the pKOV96 plasmid carrying the HP4D promoter, the pINA1291 plasmid was digested with *Cla*1 and *Bam*H1. The ~550 bp fragment containing the HP4D promoter was ligated into the pKOV96 backbone plasmid restricted with *Cla*1 and *Bam*H1. The resultant plasmid was denoted pKOV96-HP4D. Similar strategy was followed to create pKOV96-POX2, except that the POX2 promoter was obtained through restriction digestion of JMP62. The resultant plasmids pKOV96-HP4D, and pKOV96-POX2 contained the HP4D and POX2 promoter elements flanked by *Cla*1 and *Bam*H1/Acc651 restriction sites. The pKOV96-TEF, pKOV96-HP4D, and pKOV96-POX2 were used in further subcloning experiments to introduce the *T. versicolor* laccase (LACC) and the *Y. lipolytica* extracellular lipase (LIP2) secretion signals and the mature CPO gene fragments.

3.2.9.2 Cloning of secretion signals

The primer pair LACCSP-F (forward primer) and LACCSP-R (reverse primer) (Table 3.3) was used to amplify a DNA fragment encoding the LACC secretion signal from *T. versicolor* DNA. The primer pair LIP2SP-F (forward primer) and LIP2SP-R (reverse primer) was used to amplify a DNA fragment encoding LIP2 secretion signal from the *Y. lipolytica* genomic DNA. The forward primers and reverse primers have *Bam*HI and *Hind*II/*Acc*651 restriction sites respectively, to facilitate subsequent subcloning. The PCR products were cloned into pGemT-Easy according to the manufacturer's instruction to create pGEMT-Easy vectors carrying LACC and LIP2 secretion signals, respectively denoted pGemT-LACC and pGemT-LIP2. The LACC and LIP2 secretion signals were released from pGemT-LACC and pGemT-LIP2 by restriction digestion with *Bam*H1 and *Acc*651 followed by ligation into pKOV96-TEF, pKOV96-HP4D, or pKOV96-POX2 digested with *Bam*H1 and *Acc*651. The subcloning experiments resulted in the cloning of the LACC secretion signals under the TEF promoter (pKOV96-TEF-LACC), the HP4D promoter (pKOV96-HP4D-LACC) or the POX2 promoter (pKOV96-POX2-LACC. The LIP2 secretion signal was similarly cloned under the TEF promoter (pKOV96-TEF-LIP2), the HP4D promoter (pKOV96-HP4D-LIP2), and the POX2 promoter (pKOV96-POX2-LIP2). The subcloned LACC and LIP2 secretion signal DNA fragments are flanked by the *Bam*H1 and *Hind*III/Acc651/*Avr*II/*Eco*RI restriction sequences at the 5´ and 3´, respectively.

3.2.9.3 Construction of CPO expression cassettes

The gene encoding the complete CPO gene was custom synthesized by Geneart (GmbH, Regensburg, Germany) as a codon optimized construct for expression in *Y. lipolytica*. The plasmid carrying the complete CPO gene was denoted pGA4-CPO. The primer pair *Hind*IIICPO-F and CPO*Avr*II-R (Table 3.3) was used to amplify the DNA encoding the mature CPO gene. The resultant CPO PCR product was cloned into the pGemT-Easy vector to create pGemT-CPO. The CPO gene fragment is flanked by *Hind*III and *Avr*II restriction sites introduced as part of forward and reverse primers, respectively. The mature CPO gene fragment was released by *Hind*III and *Avr*II restriction digestion. The *Hind*III/*Avr*II restricted CPO was ligated into the following plasmids cut with *Hind*III and *Avr*II: pKOV96-TEF-LACC, pKOV96-HP4D-LACC,

pKOV96-POX2-LACC, pKOV96-TEF-LIP2, pKOV96-HP4D-LIP2, and pKOV96-POX2-LIP2 respectively to create pKOV96-TEF-LACC-CPO, pKOV96-HP4D-LACC-CPO, pKOV96-POX2-LACC-CPO, pKOV96-TEF-LIP2-CPO, pKOV96-HP4D-LIP2-CPO, and pKOV96-POX2-LIP2-CPO. The construction of expression cassettes in pKOV96 plasmids derivatives places the CPO for expression under the various promoters (TEF, HP4D, POX2), allowing options for the use of the LACC or LIP2 secretion signal. The pKOV96 contains the non-defective ura3d1 allele marker for selection of *Y. lipolytica* transformed with single copies of the gene expression cassette. The plasmids for multiple copy integration into *Y. lipolytica* genome were constructed by ligating the *Cla*1/*Eco*R1 expression cassettes obtained from the pKOV96-promoter-secretion signal-CPO constructs into pINA1293 digested with similar enzymes. The derived pINA1293 based multicopy expression plasmids are summarized in Table 3.4. The transformation of *Y. lipolytica* Po1f strain was created using *Not*1 linearized pINA1293 promoter-secretion signal-CPO constructs. The control strain was done by transforming *Y. lipolytica* Po1f with the linearized empty pINA1293 plasmid.

Table 3.4: Description of derived expression plasmid vectors used in the study

Vectors	Description
pINA1293-TEF-LACC-CPO	Expression vector carrying the zeta target site, ura3d4 multiple copy selection sequence, TEF promoter and *Trametes versicolar* laccase secretion signal peptide, the CPO ORF in-frame with the sequence encoding the 6xHis at the 3 prime end.
pINA1293-TEF-LIP2-CPO	Expression vector carrying the zeta target site, ura3d4 multiple copy selection sequence, TEF promoter and *Y. lipolytica* lip2 secretion signal peptide, CPO ORF in-frame with the sequence encoding the 6xHis at the 3 prime end.
pINA1293-HP4D-LACC-CPO	Expression vector carrying the zeta target site, ura3d4 multiple copy selection sequence, HP4D promoter and *Trametes versicolar* laccase secretion signal peptide, CPO ORF in-frame with the sequence encoding the 6xHis at the 3 prime end.
pINA1293-HP4D-LIP2-CPO	Expression vector carrying the zeta target site, ura3d4 multiple copy selection sequence, HP4D promoter and *Y. lipolytica* lip2 secretion signal peptide, the CPO ORF in-frame with the sequence encoding the 6xHis at the 3 prime end.
pINA1293-POX2-LACC-CPO	Expression vector carrying the zeta target site, ura3d4 multiple copy selection sequence, POX2 promoter and *Trametes versicolar* laccase secretion signal peptide, the CPO ORF in-frame with the sequence encoding the 6xHis at the 3 prime end.
pINA1293-POX2-LIP2-CPO	Expression vector carrying the zeta target site, ura3d4 multiple copy selection sequence, POX2 promoter, *Y. lipolytica* lip2 secretion signal peptide, the CPO ORF in-frame with the sequence encoding the 6xHis at the 3 prime end.

3.2.10. Production of recombinant CPO in *Yarrowia lipolytica* yeast

Yarrowia lipolytica Po1f yeast transformants were grown overnight in YPD medium. The overnight grown culture was used to inoculate 50 ml YPD, 50 mM citrate buffer, pH 4.0. The induction medium for the *Y. lipolytica* strains expressed with expression cassettes containing the POX2 promoter contained YPD, 50 mM citrate buffer, pH 4.0 and olive oil (2.5% v/v). The media were supplemented with 0.1 mM hemin. The cultures were incubated at 28°C at 250 rpm for 6 days in baffled flasks and samples were collected at different times of growth for CPO enzyme activity assays and Western blot analysis. The culture broth of the *Y. lipolytica* Po1f transformants was centrifuged at 5000 rpm for 10 min to separate the cells from the supernatant. The supernant was used to prepare the protein crude fraction for SDS-PAGE gels and Western blot analysis.

3.2.11. Purification of CPO using Ni-NTA affinity chromatography

The culture broth of the *Y. lipolytica* Po1f transformants was centrifuged at 5000 rpm for 10 min, and the resultant supernatant was adjusted to 300 mM NaCl in 10 mM sodium phosphate buffer pH 8.0 and allowed to mix with a slurry of Ni-NTA resin (Protino, Machinery-Nagel, Germany) pre-equilibrated with 300 mM NaCl, 10 mM imidazole in 10 mM sodium phosphate buffer pH 8.0. The Ni-NTA resin and supernatant proteins were mixed at 4 ^{o}C overnight at low speed stirring to allow protein-resin interaction to take place. The Ni-NTA resin: protein mixture was then loaded onto the XK (20/40) column packed with Ni-NTA column (Protino, Machinery-Nagel, Germany) and allowed to settle against gravity. The unbound proteins were removed by washing the column with 50 ml of 300 mM NaCl, 10 mM imidazole, 10 mM phosphate buffer pH 8.0. The elution of bound proteins was effected using the equilibration buffer containing 250 mM imidazole or 250 mM EDTA. The eluted protein fractions were analysed on 12.5% SDS-PAGE gels.

3.2.12. SDS-PAGE analysis

Proteins were separated by SDS-PAGE gel using a 12 % gel prepared as described by Laemmli (1970). The protein aliquots of 20 µL from the crude protein fractions were mixed with 10 µL of 2X SDS-PAGE sample buffer (1.5 M Tris-Cl pH 6.8, 20% SDS, 30% glycerol, 10% β-mercaptoethanol and 1.8 mg bromophenol blue). The protein samples were denatured by boiling for 2 min, and then loaded on SDS-PAGE gel for electrophoresis using the Mini-PROTEAN system (Bio-Rad, Johannesburg, South Africa). The SDS-PAGE gels were stained with Coomassie staining solution I (10% v/v acetic acid, 0.1% w/v Coomassie Brilliant Blue G, 25% v/v isopropanol) for 45 min and thereafter stained with Coomassie staining solution II (10% v/v acetic acid, 0.003% w/v Coomassie Brilliant Blue G, 10% v/v isopropanol). The stain solution was removed from the gels using the destaining solution (10% v/v acetic acid, 5% isopropanol).

3.2.13. Detection of CPO expression using Western Blotting

The proteins resolved on 12% SDS-PAGE gel were transferred to a nitrocellulose membrane by electrophoresis at a constant voltage of 70 V in ice cold CAPS buffer for 1 hour using the mini trans-blot electrophoresis apparatus (Bio-Rad, Johannesburg, South Africa). The membrane was blocked for 1 hour at room temperature in blocking buffer [10 mM Tris pH 7.5, 150 mM sodium chloride pH 7.0, and 50 mM sodium phosphate containing 3% (w/v) dried skimmed milk powder and 1% Tween 20]. The membrane was then incubated for 2 hours at room temperature with poly-histidine antibody conjugated with horseradish peroxidase (Abcam, USA) (1:1000 dilutions) in blocking buffer. The development of the signal was performed using TMB from Sigma Aldrich (Madison, USA).

3.2.14. Peptide mapping of CPO

The protein bands obtained from the CPO culture supernatant separated by SDS-PAGE were excised from the gel. The excised protein bands were diced into small pieces, and the pieces were submerged in 100 µL of 25 mM NH_4HCO_3 50% Acetonitrile (ACN) and vortexed. The supernatant was extracted and discarded. The gel pieces were completely dried by placing them in a Speed Vac. The 25 µl of 10 mM

DTT in 25 mM NH_4HCO_3 was added to the dried gel pieces. The tube was vortexed and spun briefly, and the mixture was incubated at 56°C for an hour. The supernatant was decanted and 25 µL of 55 mM iodoacetamide was introduced to the gel pieces. The mixture was vortexed and spun briefly, with the reaction allowed to proceed for 45 min at 25 °C. The supernatant was removed and the gel pieces were washed with 100 µL of NH_4HCO_3 followed by vortexing for 10 min and spinning. The supernatant was discarded and the gel pieces were dehydrated with 100 µL of 25 mM NH_4HCO_3 in 50 % ACN, the mixture was vortexed for 5 min followed by a brief spin. The gel pieces were dried completely by a Speed Vac. The dried gel pieces were treated with trypsin which was added to just barely cover them. The volume of the gel pieces was estimated and used as a reference to add 3X the volume of the trypsin solution. The gel pieces were rehydrated on ice for 10 min, spun and 25 mM NH_4HCO_3 was added to cover the mixture which was incubated at 37 °C overnight. The digested solution (aqueous extraction) was transferred into a clean 0.5 ml Eppendorf tube. Thirty microlitres of 50% ACN 5% formic acid was added to the reaction mixture, vortexed for 20-30 min and sonicated for 5 min. The mixture was completely dried in a speed Vac. The dried sample was cleaned using C18 stage tips and subjected to LC-MS at the CSIR (Biosciences, Pretoria, South Africa) using a Dionex Ultimate 3000 RSLC (Dionex, Voisins-le-Bretonneux, France) system coupled to a QSTAR ELITE mass spectrometer (Thermo Fisher, Waltham, MA, USA) via a nanoelectrospray ion source. The peptides were ionised by electrospray ionisation and the mass of the ions were measured using the mass spectrometer. The MS scans were acquired from *m/z* 400 - 1500 and the most intense fragment ions identified using the Paragon search engine (AB Sciex) (Shilov *et al.* 2007).

3.2.15. The CPO enzyme spectrophotometric activity assays

The enzyme assay for detecting the halogenating activity of CPO was performed according to the method described by Manoj and Hager (2006), which makes use of thionin as a coloured substrate. The assay was carried out at 21˚C in total volume of 3 ml reaction mixture consisting of the following: 300 µmol potassium phosphate buffer (pH 2.75), 60 µmol potassium chloride, 6 µmol of hydrogen peroxide and 120-150 nmol of thionin. The absorbance of the resulting solution was measured at 598 nm against

the corresponding reagent blank. To the reaction mixture extracellular protein crude extract was added to initiate the reaction and appropriate diluted commercial *C. fumago* CPO was used as the positive control and for comparison. The rate of decrease of absorbance at 598 nm was then recorded in the time interval of 1 min at 22˚C.

3.3. RESULTS

3.3.1. Cloning and sequence analysis of *Caldariomyces fumago* CPO

The CPO gene of *C. fumago* was custom synthesized and optimized for codon adaptation and G/C content for expression in *Y. lipolytica* (Figure 3.4). The other parameters that were optimized included elimination of cryptic splice sites, avoidance of stable RNA secondary structures and addition of stabilizing sequence elements. The Codon Adaptation Index scores of the wild type and optimized mature *C. fumago* CPO genes were 0.69 and 0.80 respectively. The wild type gene G/C content was increased from 43% to 50 % through this codon optimization. The molecular mass of the cloned *C. fumago* CPO genes was predicted as 40.5 kDa using the EXPASY protein analysis tool (Gasteiger *et al.,* 2003). The first 21 amino acid sequence occupying the *N*-terminal region of the *C. fumago* CPO gene is recognized as a signal peptide with the putative cleavage site positioned between the Gln and the Glu at positions 21 and 22 (Sundaramoorthy *et al.,* 1995). The cleavage of the signal peptide at this point releases a mature protein of 352 amino acid residues with the predicted relative molecular mass of 38.3 kDa and pI of 4.72. The synthetic CPO gene was constructed such that the mature CPO gene fragment is in-frame with a carboxyl terminal sequence coding for 6X His residues (Figure 3.4). The pI and MW of the mature CPO protein with the 6X His tag was calculated as 5.1 and 39.1 kDa, respectively. The motifs essential for CPO catalysis are also depicted in Figure 3.4.

The most important features of the primary CPO polypeptide includes amino acid residues Arg26–Asn37 which encompasses the cysteine thiolate heme ligand (Cys29) and provides a rigid scaffolding for the iron–sulphur interaction; Cys79 and Cys87 which are involved in disulphide bond formation (Sundaramoorthy *et al.,* 1995) and Glu183 which is postulated to function on distal side of the heme prosthetic group as an acid-base catalyst (Yi *et al*., 2003). Furthermore, the polypeptide is also comprised of the *N*-glycosylation sites at positions Asn12, 93, and 216 (Sundaramoorthy *et al.,* 1995). It has been suggested in the literature that the active mature CPO lacks the last 52 amino acid residues due to proteolytic processing with the processing occurring downstream of the dibasic (KR) processing site (Sundaramoorthy *et al.,* 1995, Conesa *et al.,* 2001b). In this study, the 52 aa carboxyl terminal propetide was included as part of the CPO protein fragment for expression and possible processing in *Y. lipolytica* yeast (Figure 3.4).

3.3.2. Construction of CPO expression and secretion vectors

The mature CPO gene, including the sequences encoding the C-terminal propetide and the His tag was cloned under the control of the HP4D, POX2 or TEF promoter with the LACC or LIP2 as signals directing CPO secretion to the extracellular (Figure 3.5). The expression cassettes were cloned in pINA1293 plasmid backbone which contains the *zeta* elements to enable random integration within *Y. lipolytica* yeast, and the ura3d4 allele for selection of transformants containing multiple copies of expression cassettes (Figure 3.5). The resultant plasmids were denoted pINA1293-HP4D-LIP2-CPO, pINA1293-HP4D-LACC-CPO, pINA1293-TEF-LIP2-CPO, pINA1293-TEF-LACC-CPO, pINA1293-POX2-LIP2-CPO, pINA1293-HP4D-POX2-CPO (Figure 3.6).

atgttctctaaggtcctgcccttcgtgggagccgtggccgccctgccccactctgtgcga
M F S K V L P F V G A V A A L P H S V R
cag▼gagcccggctctggcatcggctacccctacgacaacaacaccctgccctacgtggct
Q E P G S G I G Y P Y D **N** N T L P Y V A
cccggacccaccgactctcgagccccctgtcccgccctgaacgccctggccaaccacggc
P G P T D S **R A P *C* P A L N A L A N** H G
tacatcccccacgacggccgagccatctccagagagaccctgcagaacgccttcctgaac
Y I P H D G R A I S R E T L Q N A F L N
cacatgggcatcgccaactctgtgatcgagctggccctgaccaacgccttcgtggtgtgt
H M G I A N S V I E L A L T N A F V V ***C***
gagtacgtgaccggctctgactgtggcgactctctggtgaacctgaccctgctggccgag
E Y V T G S D ***C*** G D S L V **N** L T L L A E
ccccacgccttcgagcacgaccactcgttctctcgaaaggactacaagcagggcgtggcc
P H A F E H D H S F S R K D Y K Q G V A
aactctaacgacttcatcgacaaccgaaacttcgacgccgagaccttccagacctctctg
N S N D F I D N R N F D A E T F Q T S L
gacgtggtggccggcaagacccacttcgactacgccgacatgaacgagatccgactgcag
D V V A G K T H F D Y A D M N E I R L Q
cgagagtctctgtctaacgagctggacttccccggctggttcaccgagtctaagcccatc
R E S L S N E L D F P G W F T E S K P I
cagaacgtggagtctggcttcatcttcgccctggtgtctgacttcaacctgcccgacaac
Q N V **E** S G F I F A L V S D F N L P D N
gacgagaacccccctggtgcgaatcgactggtggaagtactggttcaccaacgagtctttc
D E N P L V R I D W W K Y W F T **N** E S F
ccttaccacctgggctggcacccccctcccctgcccgagagatcgagttcgtgacctct
P Y H L G W H P P S P A R E I E F V T S
gcctcttctgccgtgctggccgcctctgtgacctctacccccctcttctctgccctctggc
A S S A V L A A S V T S T P S S L P S G
gccatcggacccggcgctgaggccgtgcccctgtctttcgcctctaccatgacccccttt
A I G P G A E A V P L S F A S T M T P F
ctgctggccaccaacgccccctactacgcccaggaccccaccctgggccccaacgacaag
L L A T N A P Y Y A Q D P T L G P N D **K**
cga▼gaggccgctcccgccgctaccacctctatggccgtgttcaagaacccctacctggag
R E A A P A A T T S M A V F K N P Y L E
gccatcggcacccaggacatcaagaaccagcaggcctacgtgtcctctaaggctgccgcc
A I G T Q D I K N Q Q A Y V S S K A A A
atggcctctgccatggctgccaacaaggcccgaaacctgcatcatcatcatcatcattaa
M A S A M A A N K A R N L H H H H H H –

Figure 3.4. Custom synthesised complete *C. fumago* CPO gene with 6X His sequence at the 3´ end. The underlined N-terminal sequence is the native secretion signal and the underlined C-terminal sequence is the 52 aa propeptide. The double underlined amino acid sequence Arg26-Asn37 form the catalytic domain of the CPO protein and include the Cys29 ligand (bold and italics). The (▼) indicates the cleavage site of the native CPO secretion signal and of the 52 aa propeptide at the N- and C- termini, respectively.Other marked amino acids are Asn 12, 93, 216 for *N*-glycosylation site (bold and underlined), Glu 183 (bold and spotted underlined), Cys 79 and 87 for disulphide bond formation (double underlined and italics).

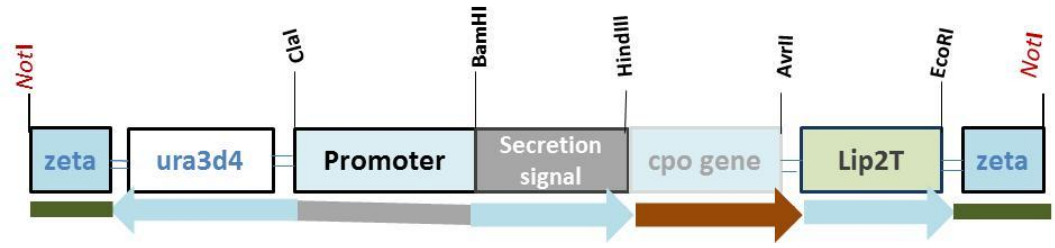

Figure 3.5: Schematic diagram of the expression cassette integrated into *Y. lipolytica* Po1f genome. The linearized vector consisting of *zeta* for random integration in *Y. lipolytica* genome, the promoter (TEF, HP4D or POX2), secretion signal (LIP2 or LACC), mature CPO ORF and LIP2 terminator sequence. The schematic diagram is not drawn on scale.

The presence of the promoter, the secretion signal and the mature CPO encoding DNA fragments in the pINA1293 based expression vectors was confirmed by restriction analysis using *Cla*I, *BamH*I, *Hind*III and *EcoR*I. The DNA fragments of about 100 bp (LACC secretion signal), 500 bp (HP4D promoter), 1100 bp (mature CPO), and ~4196 bp (pINA1293 backbone) were expected and obtained upon restriction digestion of pINA1293-HP4D-LACC-CPO (Figure 3.7, lane 1A). The restriction analysis of pINA1293-TEF-LACC-CPO resulted in DNA fragments of 100 bp (LACC), 360 bp (TEF promoter), 1100 bp (mature CPO), and ~4196 bp (pINA1293 plasmid backbone) (Figure 3.7, lane 2A). The restriction analysis of pINA1293-POX2-LACC-CPO resulted in DNA fragments of 100 bp (LACC), 1200 bp (POX2 promoter), 1100 bp (mature CPO), and ~4196 bp (pINA1293 backbone plasmid) (Figure 3.7, lane 3A). The DNA fragments of about 110 bp (LIP2 secretion signal), 500 bp (HP4D promoter), 1100 bp (mature CPO), and ~4196 bp (pINA1293 backbone) were expected and obtained upon restriction digestion of pINA1293-HP4D-LIP2-CPO (Figure 3.7, lane 1B). The restriction analysis of pINA1293-TEF-LIP2-CPO resulted in DNA fragments of 110 bp (LIP2), 360 bp (TEF promoter), 1100 bp (mature CPO), and ~4196 bp (pINA1293 plasmid backbone) (Figure 3.7, lane 2B). The restriction analysis of pINA1293-POX2-LIP2-CPO resulted in DNA fragments of 110 bp (LIP2), 1200 bp (POX2 promoter), 1100 bp (mature CPO), and ~4196 bp (pINA1293 backbone plasmid) (Figure 3.7, lane 3B). The DNA sequences of the different components were verified by sequencing at different stages of subcloning.

The *Y. lipolytica* expression vectors carry the ura3d4 allelle which retained only 6 bp upstream from the URA3 start codon, and is unable to confer Ura+ phenotype as a single copy, but could promote the amplification of the vector copy number in multiple integrations (Le Dall *et al*., 1994; Pignede *et al*., 2000; Nicaud *et al.*, 2002). The ura3d4 defective selection marker has been used successfully for selection of multiple copy integrants in heterologous protein production (Pignede *et al*., 2000; Nicaud *et al.*, 2002). The promoters with different mechanisms of transcriptional activation were selected in this study. The TEF is a constitutive promoter based on the translational elongation factor TEF (Muller *et al.*, 1998). The POX2 promoter is an inducible promoter, activated by the presence of fatty acid substances such as olive oil or oleic acids in growth media (Juretzek *et al.,* 2000). The HP4D promoter is defined as a quasi-constitutive promoter, exhibiting high transcriptional activity relative to the growth phase of the *Y. lipolytica* yeast (Madzak *et al.,* 2000). An endogenous (LIP2) and a foreign (*T. versicolor* LACC) secretion signals known to direct heterologous protein production in *Y. lipolytica* yeast were selected. The LIP2 secretion signal has been shown to be the most efficient sequence in directing the secretory expression of fully glycosylated heterologous proteins in *Y. lipolytica* (Moon *et al.*, 2013). The heterologous production of laccase from the fungus *T. versicolor* yielded relatively higher enzyme activity in the culture medium of *Y. lipolytica* when expressed from its native secretion signal, indicating potentially enhanced efficiency of the LACC as protein secretion signal, at least for fungal derived complex proteins (Jolivalt *et al.*, 2005).

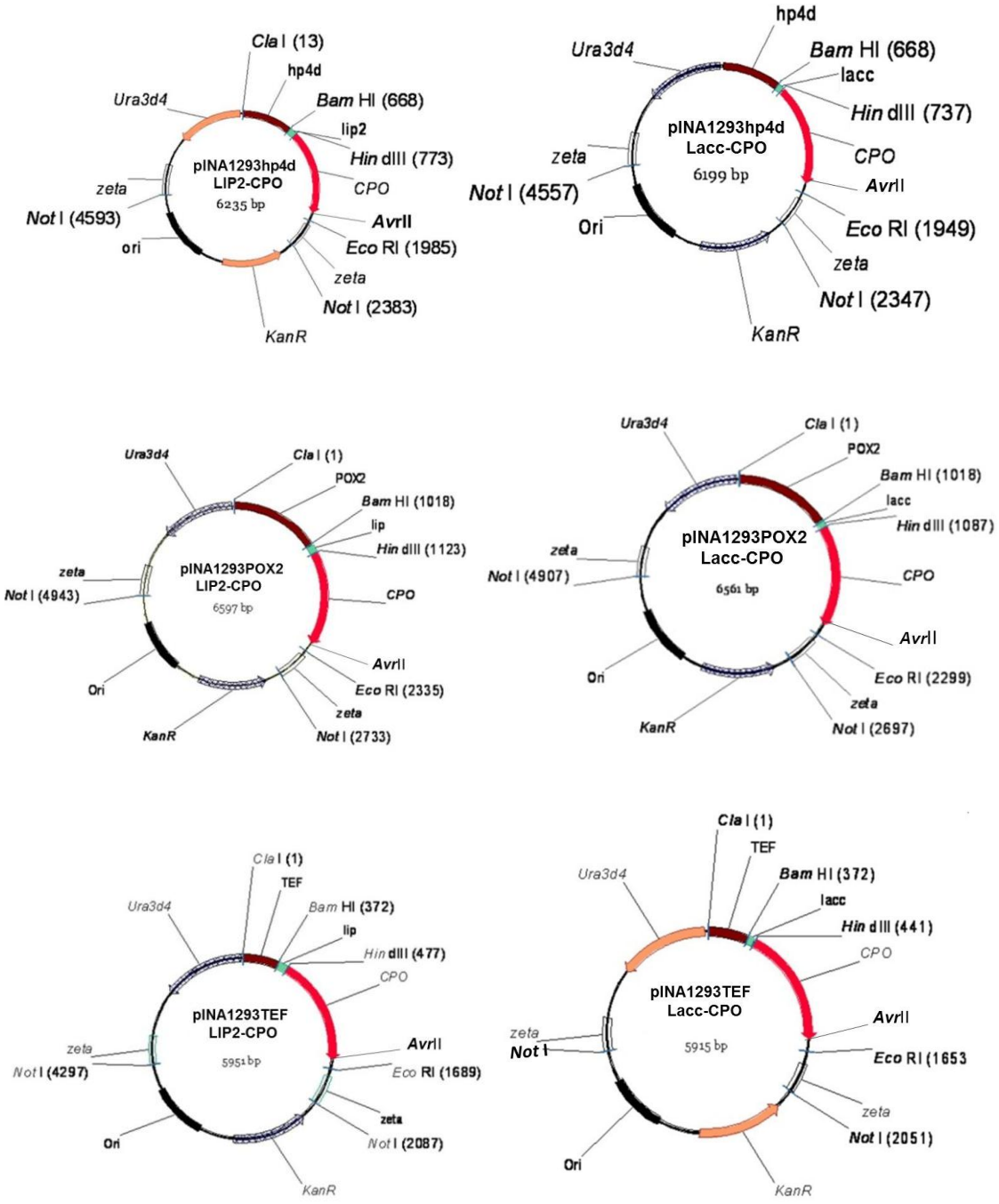

Figure 3.6 Schematic diagram showing pINA1293 vectors used for random multiple copy integration in *Y. lipolytica*.

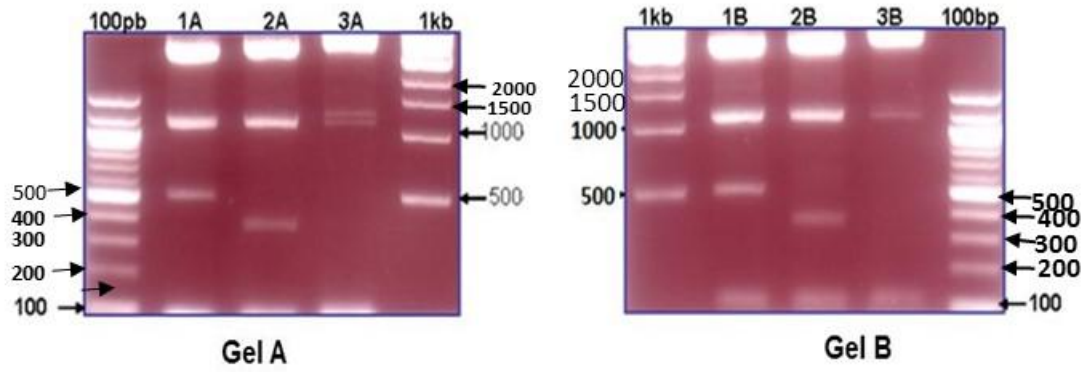

Figure 3.7: Restriction analyses of the six pINA1293 expression vectors as digested with *Cla*I, *BamH*I, *Hind*III and *Eco*R1 to confirm the presence of promoter, secretion signal sequence and mature CPO gene. Agarose gel **A** (1%) stained with ethidium bromide showing, pINA1293-HP4D-LACC-CPO (lane 1A), pINA1293-TEF-LACC-CPO (lane 2A), pINA1293-POX2-LACC-CPO (lane 3A). Agarose gel **B** (1%) showing, pINA1293-HP4D-LIP2-CPO (lane 1B), pINA1293-TEF-LIP2-CPO (lane 2B), pINA1293-HP4D-POX2-LIP2-CPO (lane 3B).

3.3.3. *Yarrowia lipolytica* transformation and screening of transformants

The *Y. lipolytica* Po1f strain was transformed using *Not*I linearized expression cassettes (Figure 3.5). The linearization with *Not*I ensures the removal from the expression cassettes of plasmid DNA moieties required for replication in *E. coli* and antibiotic resistance. The transformants of the *Y. lipolytica* Po1f were selected for Ura+ prototrophy in YNB agar medium supplemented with casamino acids. A total number of 12 ura^+ prototrophic transformants were obtained (Table 3.5). The primer pair *Hind*IIICPO-F and CPO*Avr*II-R targeting the CPO gene carried in the expression cassettes were used in the PCR amplification to confirm the integration of the expression cassettes within the Ura+ prototrophic transformants of the *Y. lipolytica* yeast genome. The PCR product of approximately 1100 bp was expected and obtained with seven of the twelve transformants (Figure 3.8). Numerous attempts to amplify the target CPO from the other 5 ura+ prototrophic transformants of the *Y. lipolytica* did not yield any desired PCR product. The genomic DNA of the PCR confirmed transformants was digested with *Eco*RI and subjected to Southern blot analysis using the CPO gene fragment as the probe. The Southern blot analyses confirmed only 4 ura+ prototrophs

to have integrated CPO gene fragments within the genome (Table 3.5, Figure 3.9). It was interesting to note that no integration was obtained from transformations with the expression cassettes containing the TEF promoter despite several attempts of transformation. This pointed to a possible toxic effect of the CPO gene or gene product to the *Y. lipolytica* host given the nature of the constitutively expressed TEF promoter. Transformants were obtained with pINA1294-POX2-LIP2-CPO, pINA1294-HP4D-LIP2-CPO and pINA1294-HP4D-LACC-CPO (Table 3.5, Figure 3.9). The *Y. lipolytica* transformants confirmed by Southern blot to be carrying the expression cassettes were denoted Yl-4, Yl-11, Yl-12, and Yl-13 for pINA1293-POX2-LIP2-CPO (lane 4), pINA1293-HP4D-LIP2-CPO (lane 11), pINA1293-HP4D-LACC-CPO (lane 12) and pINA1293-HP4D-LACC-CPO (lane 13), respectively.

Table 3.5. Number of *Y. lipolytica* Po1f transformants obtained after transformation with the respective expression cassettes.

Expression cassette	**Transformants obtained**	**PCR confirmed transformants**	**Southern blot confirmation**
pINA1293-TEF-LACC-CPO	0	0	0
pINA1293-TEF-LIP2-CPO	0	0	0
pINA1293-HP4D-LACC-CPO	2	2	2
pINA1293-HP4D-LIP2-CPO	4	2	1
pINA1293-POX2-LACC-CPO	2	2	0
pINA1293-POX2-LIP2-CPO	4	1	1
TOTAL	**12**	**7**	**4**

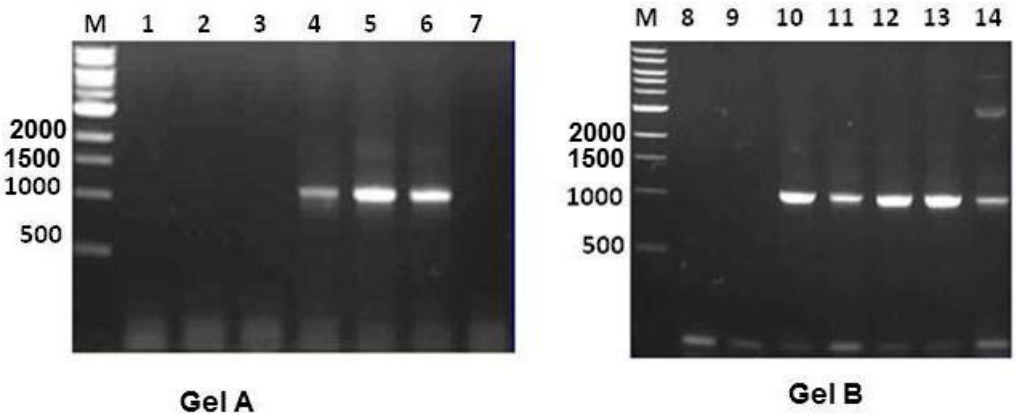

Figure 3.8. PCR amplification of CPO gene from *Y. lipolytica* Po1f transformants. The ethidium stained (0.8%) agarose gel showing, preparations from *Y. lipolytica* Po1f strains transformed with expression cassette pINA1293-POX2-LIP2-CPO (Gel A lane 1, 2, 3 and 4), pINA1293-POX2-Lacc-CPO (Gel A lane 5 and 6), and *Y. lipolytica* Po1f strain as negative control (lane 7). B. *Y. lipolytica* Po1f strains transformed with pINA1293-HP4D-LIP2-CPO (Gel B lane 8, 9, 10 and 11), pINA1293-HP4D-Lacc-CPO (Gel B lane 12 and 13), and the CPO positive control amplified from pGA-CPO (Gel B lane 14). M is the 1kb DNA ladder.

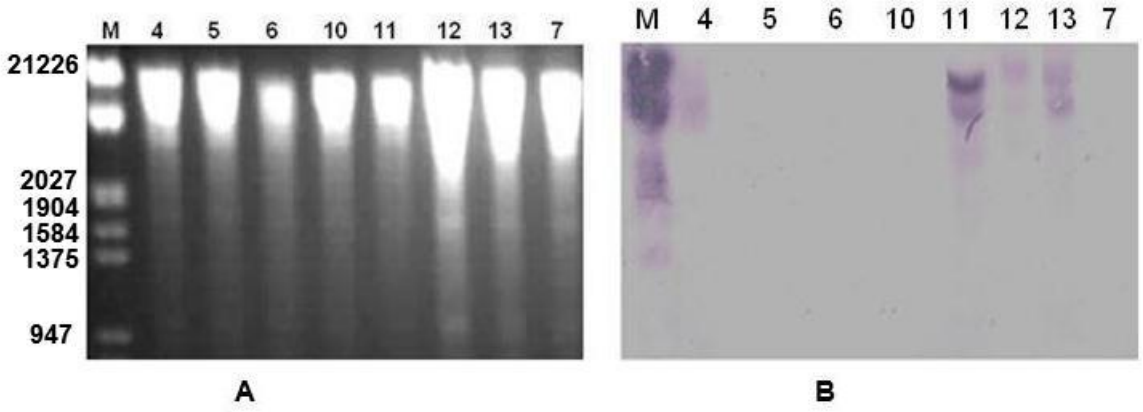

Figure 3.9: Southern blotting analysis of *Y. lipolytica* Po1f transformants. Gel A: Genomic DNA isolated from *Y. lipolytica* Po1f transformants digested with *EcoR*I. Gel B: Hybridization of Gel A with CPO DNA probe. The expression cassettes used to transform the *Y. lipolytica* are as follows: pINA1293-POX2-LIP2-CPO (lane 4), pINA1293-POX2-Lacc-CPO lane 5 and 6, pINA1293-HP4D-LIP2-CPO lane 10 and 1, pINA1293-HP4D-Lacc-CPO lane 12 and 13, *Y. lipolytica* Po1f transformed with pINA1293 as negative control (lane 7).

The ura3d4 is a defective multiple copy selection marker which should result in multiple integrations in the *Y. lipolytica* yeast genome (Juretzek *et al.,* 2001). The Southern blot analysis revealed potentially one integrated copy for the expression cassette pINA1293-POX2-LIP2-CPO (Figure 3.9B, lane 4), 3 copies integrated for pINA1293-HP4D-LIP2-CPO (Figure 3.9B, lane 11), and two copies each for pINA1293-HP4D-Lac-CPO (Figure 3.9B, lanes 12 and 13), judging by the number of distinct bands obtained after hybridisation with the CPO DNA probe. The *EcoR*I restriction site used in restriction digestion of genomic DNA for Southern blot does not cut within the targeted CPO gene. The hybridization bands were therefore expected to be genetically distinct due to the different location sites of the expression cassette and the number of copies integrated into the genome. It has been reported in literature that the *Y. lipolytica* yeast transformed with the Ura3d4 allele stabilizes to the average copy number of 10±13 copies integrated within the genome (Juretzek *et al.,* 2001). The number of copies integrated could not be determined with certainity due to the smearing of the hybridization events in Southern blotting (Figure 3.9).

The *Y. lipolytica* was not explored for CPO production from monocopy expression systems due to the reported amplification and deamplification processes limiting and stabilizing Ura3d4 transformants to the average copy number of 10±13 copies (Juretzek *et al.,* 2001). It could be argued that the stabilisation of the copy number in ura3d4 transformants is a balancing act by the yeast cell between maintaining the copy number sufficient for complementation of the uracil auxotrophy and relieving the yeast cell from the metabolic burden imposed by increased demand of protein synthesis, processing and secretion to the extracellular. It has been reported in literature that upon transfer to alkaline extracellular protease (AEP) induction medium, *Y. lipolytica* strain 773-2 (50 integrated copies of XPR2), grew for at least 10 hours before AEP production began, and then growth rate decreased before increasing again; by then, cells had lost copies of XPR2 (Le Dall *et al*., 1994). The slowing of growth following AEP induction suggested that the increased secretory pathway cargo load was affecting cell growth and that such a system had potential for secretion stress studies (Ogrydziak and Nicaud, 2012). In accordance with this argument, transforming the cells with multiple copy integrant expression cassettes, the yeast cell would stabilize to copy

numbers adequate to balance and maintain uracil prototrophy and metabolic demand imposed by heterologous CPO expression.

3.3.4. Expression of CPO in *Y. lipolytica* and protein purification

The ura+ prototrophic *Y. lipolytica* transformants confirmed by Southern blot to have integrated within their genomes the expression cassette were denoted Yl-4 (for pINA1293-POX2-LIP2-CPO), Yl-11 (for pINA1293-HP4D-LIP2-CPO), and Yl-12, and Yl-13 for the pINA1293-HP4D-Lac-CPO integrants (Figure 3.9). The YL-11, Yl-12 and Yl-13 were grown in YPD medium while the Yl-4 strain was grown in YPD medium containing oleic acid for induction of the gene under the POX2 promoter. The media was supplemented with hemin which has been reported to enhance the production of recombinant CPO when expressed in *A. niger* (Conesa *et al*., 2001a). The culture supernatant and cell lysate containing crude protein fractions were analysed by SDS-PAGE gel to detect possible overexpression of the CPO protein. The SDS-PAGE gel showed two prominent bands of sizes approximating 38-42 kDa (Figure 3.10). The two bands suspected to be recombinant CPO were excised from SDS-PAGE and subjected to peptide mapping by protease digestion and LC-MS/MS analysis. The peptide map analyses did not reveal peptide sequence corresponding to the CPO protein (data not shown). It is noteworthy that the protein band patterns of the control (Po1f) differed markedly from the experimental samples (4, 11, 12 and 13) “containing” the heterologous CPO (Figure 3.10). These points to a probable switch in expression of a different set of genes by the *Y. lipolytica* host in the presence of the heterologous CPO gene or protein product. This might be true if the CPO is indeed toxic to the host cell as suggested by the failure to obtain recombinant clones when the *Y. lipolytica* cells was transformed with constructs of CPO under the TEF promoter. Under conditions of toxicity, the yeast would most likely respond through expression of proteins that would function to alleviate the metabolic burden or stress. Yeast cells have been reported to develop metabolic strategies to enhance long-term survival under conditions of metabolic stress which included increasing the expression and activity of stress responsive proteins (Kim *et al*., 2014).

The recombinant CPO expression cassette was designed such that it contained sequence encoding 6X His tag at the carboxyl terminus. The protein samples were subjected to Western blot analysis using anti-histidine polyclonal antibodies to detect the 6XHis tag attached at the C-terminal end of the CPO protein. The protein bands of approximately 38-42 kDa were detected in extracellular protein fractions of the Yl-4, Yl-11, Yl-12, and Yl-13 following Western blot analyse (Figure 3.10). The protein extract from the control strain *Y. lipolytica* Po1f transformed with empty pINA1293 expression cassette also showed a reactive bands, albeit comparatively faint in signal intensity, and at position of ~55 kDa which is the size outside the expected 38-42 kDa (Figure 3.11). The results suggested expression of full length His-tagged CPO protein (inclusive of the carboxyl terminal 52 aa propeptide) in the extracellular medium of the Yl-4, Yl-11, Yl-12, and Yl-13.

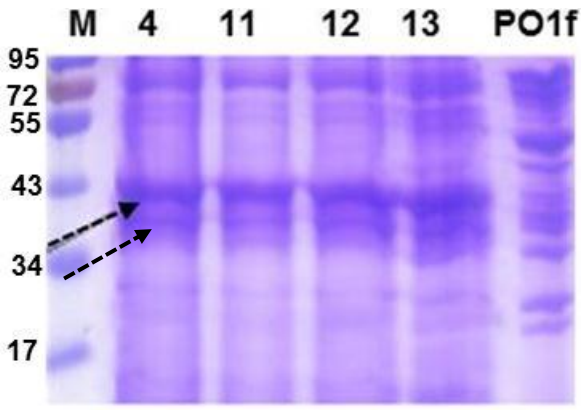

Figure 3.10. SDS-PAGE analysis of extracellular protein fraction from shake cultures. The Yl-4, Yl-11, Yl-12, and Yl-13 (Lanes 4, 11, 12, and 13) were transformed with pINA1293-POX2-LIP2-CPO, pINA1293-HP4D-LIP2-CPO, pINA1293-HP4D-Lac-CPO and pINA1293-HP4D-Lac-CPO, respectively. The protein samples were separated on 12% SDS-polyacrylamide gel and visualised with Coomassie Brilliant Blue stain. Analysis showing 10 µL of protein fractions, M is the prestained molecular weight marker, Po1f is the *Y. lipolytica* control strain transformed with pINA1293. The arrows point to protein bands that were excised for peptide mapping using LC-MS/MS analysis.

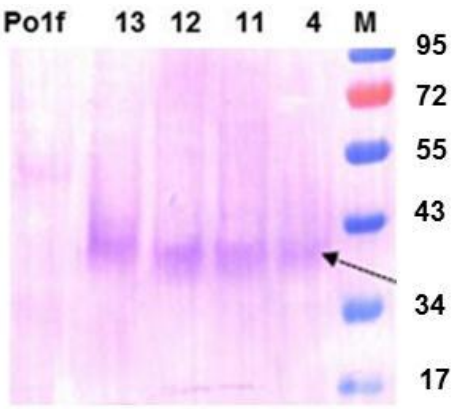

Figure 3.11. Western blotting analysis of the extracellular protein fraction using anti-Histidine polyclonal antibody. The analysis was done using 10 µL extracellular protein extract of *Y. lipolytica* Po1f transformed with the following expression cassettes pINA1293-POX2-LIP2-CPO, pINA1293-HP4D-LIP2-CPO, pINA1293-HP4D-Lac-CPO and pINA1293-HP4D-Lac-CPO respectively to create the following strains yeast strains Yl-4, Yl-11, Yl-12, and Yl-13 (Lanes 4, 11, 12, and 13). M represents the prestained molecular weight marker standard. The arrow is pointing at the protein bands detected by the Western blot using anti Histdine polyclonal antibody.

3.3.5. Purification by Ni-NTA

The purification of the recombinant CPO proteins using Ni-NTA affinity chromatography was considered due to the incorporated sequence encoding 6X His residues at the C-terminus of the protein. The extracellular crude protein extract was mixed with the Ni-NTA resin under conditions that allow binding of His tagged protein to the matrix resin, followed by the elution of bound proteins using increased concentrations of imidazole or EDTA. The eluted protein fractions were analysed on SDS-PAGE gel for protein bands. However, no protein bands could be detected from the elution fractions, despite the addition of EDTA which promote protein elution by stripping the resin of the Ni^{2+} ions from the affinity column (Petty, 1996). The results indicated that the His-tagged CPO protein could have failed to bind to the Ni-NTA column, and this might be due to the His tail being buried in the protein and being inaccessible to the resin.

3.3.6. Catalytic activity

The CPO was assayed for catalytic activity by measuring its chlorination ability using thionin substrate with commercial *C. fumago* CPO as positive control. The method detects and assays the halogenating activity of hemeperoxidases using the coloured substrate, thionin (Manoj and Hager, 2006). In the presence of suitable amounts of peroxide and chloride, chloroperoxidase chlorinates thionin and bleaches the intense colour of the substrate resulting in the decrease in absorbance (Manoj and Hager, 2006). The post cultivation extracellular protein extracts of the Yl-4, Yl-11, Yl-12, and Yl-13 showed no chloroperoxidase activity in contrast to the commercial *C. fumago* CPO (Figure 3.12). The results indicated expression of inactive recombinant CPO by the *Y. lipolytica* yeast.

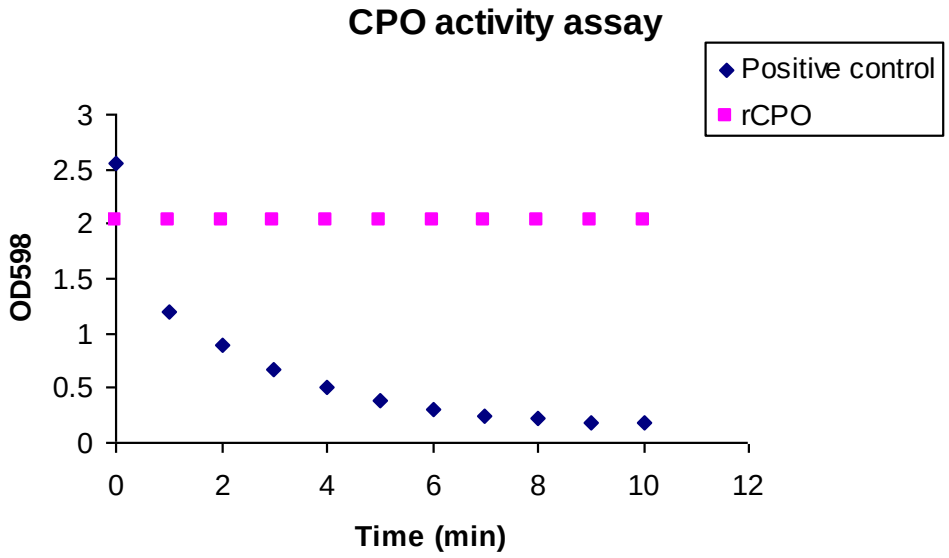

Figure 3.12. Chlorination activity assay of CPO. The recombinant CPO extract is represented by the symbol (■) and the commercial native *C. fumago* CPO the positive control is represented by the symbol (♦).

3.4. DISCUSSION AND CONCLUSION

This study describes the attempts to produce with multicopy expression systems, the recombinant CPO gene in *Y. lipolytica* under the control of the constitutive TEF promoter, quasi-constitutive HP4D promoter, and the inducible POX2 promoter using LACC or LIP2 secretion signals derived from *T.vesicolor* laccase and the *Y. lipolytica* extracellular lipase, respectively. The PCR screening and Southern blot analyses of *Y. lipolytica* isolates transformed with the expression cassettes confirmed successful integration of the CPO gene fragment contained in pINA1293-POX2-LIP2-CPO, pINA1293-HP4D-LIP2-CPO, pINA1293-HP4D-Lac-CPO and pINA1293-HP4D-Lac-CPO to yield *Y. lipolytica* strains Yl-4, Yl-11, Yl-12, and Yl-13. The transformation of the *Y. lipolytica* with the multicopy expression cassettes of CPO failed to yield a yeast isolate producing active CPO. Heterologously expressed proteins undergo a wide range of assortments within the host secretory pathway before being released to the extracellular. In yeast, the protein secretory pathway is a highly-integrated, multi-organelle system, and dysfunction can happen at many levels including at post-transcription/translation level, during Golgi vesicle sorting, as a result of high protein degradation rates, or oxidative stress (Tyo *et al*., 2012). The failure and low expression of foreign genes in *Y. lipolytica* could be attributed to factors like the genetic design of the construct such as codon bias optimization, the use of an inappropriate signal peptides, inadequate translation initiation codon environment and growth medium composition (Gasmi *et al*., 2011). The codon usage of the recombinant CPO gene was optimized for expression in *Y. lipolytica*, and as such codon biases is not considered a factor in the lack of active CPO.

The Western blot studies revealed successful transcription and translation of a 38-42-kDa His-tagged protein in growth cultures of the selected transformed *Y. lipolytica* yeast strains, suggesting that the pH of the growth medium did not affect HP4D and POX2 promoter activities. The protein was presumed to be the cloned recombinant CPO protein based on the determined and expected protein size and reactivity to poly-histidine antibodies. The protein was however not highly expressed to be distinct and intense to allow selective excision from SDS-PAGE for peptide mapping. The data pointed to a successful expression of the CPO protein in *Y. lipolytica*, albeit at a low levels and in an inactive form.

The CPO could have been produced at only low levels due to culture conditions that are unfavourable to induce expression of active CPO. In microbial fermentations, the engineering aspects and nutritional requirements including concentrations in the growth medium, fermentation parameters, pH, temperature, aeration and agitation, are essential for efficient heterologous protein production (Masurekar, 2008). In this study, the *Y. lipolytica* strains were grown at the yeast's optimal growth temperature of 28 °C, with agitation to ensure efficient aeration of this strictly aerobic yeast. The pH of the growth medium has an important influence on the function of promoters, and this might affect heterologous protein production in *Y. lipolytica* (Blanchin-Roland, 2008). The inducible XPR2 promoter is active only at pH above 6 and its full induction requires high levels of peptones in the culture medium (Juretzek *et al*., 2000; Madzak *et al*., 2000; Madzak *et al*., 2004). The HP4D promoter employed in this study has been derived from the XPR2 promoter transcriptional elements. However, it has been rendered quasi-constitutive and the gene expression has become growth-phase dependent, but independent of pH, carbon and nitrogen sources (Madzak *et al*., 2000). Although there are no studies detailing the effect of pH on the transcriptional activity of the acyl-CoA oxidases POX2 promoter, it is known to be induced by fatty acids (Juretzek *et al.*, 2000), an indication of its activity at acidic growth conditions. Indeed, the medium for large scale production of heterologous proteins by *Y. lipolytica* under the POX2 promoter have been designed with final pH values ranging between 2.7-5.5 (Nicaud *et al*., 2002; Gasmi *et al.*, 2011). However, literature on the application of POX2 promoter in growth media buffered at pH 6.8 have been reported (Juretzek *et al.*, 2000). The results suggested that the use of HP4D and POX2 promoter could not be the reason for low level expression of CPO due to the effect of pH on transcriptional activity of the promoter elements.

While the HP4D promoter can be used in virtually any medium, the drawback has been described to be its limited potential when a heterologous protein toxic to *Y. lipolytica* is being expressed (Gasmi *et al*., 2011). The inducible property of the POX2 promoter is considered to be the option when expressing toxic proteins where gene expression remains repressed by glucose or glycerol in the medium, and upon exhaustion of glucose or glycerol, the fatty acids contained in the media serve as carbon source and

inducer for expression of toxic genes, thus allowing sufficient accumulation of the yeast cell biomass (Gasmi *et al*., 2011). The HP4D and the POX2 promoters have been compared for their capacity to produce extracellular lipase using the XPR2 or LIP2 secretion signals, and were found to be complementary options as they resulted in similar protein production levels (Nicaud *et al*., 2002). In this study both the HP4D and POX2 promoters were used either with LACC or LIP2 secretion signals, and still very little CPO was produced by the transformed *Y. lipolytica* strains. The use of both the HP4D and POX2 promoter, was aimed at the separation of the growth phase and the CPO production phase to eliminate any possible lack of CPO activity as a result of toxicity of the CPO protein to the cells, or as a result of growth conditions affecting the transcriptional activity of the promoter driving CPO expression.

The cultivation conditions could affect the catalytic properties and function of CPO due to pH fluctuations or as a result of extended periods of cultivation that promote protein destabilization. The heterologous production of CPO enzyme was studied in the growth media at pH 4.0 in light of CPO being reported to be optimally active in acidic pH values (Sundaramoorthy *et al.,* 1995; Li *et al*., 2011; Bayramoglu *et al*., 2011). The CPO protein is indeed an acidic protein, with the calculated pI of 4.7. The protein has been reported to be affected when the growth medium pH rises above 6.5 (Carmichael *et al*., 1986). In this study, the recombinant *Y. lipolytica* strains were grown in buffered growth medium at pH 4.0, and culture samples were withdrawn daily for 6 days for enzyme assays using undisrupted cell culture to ensure timely collection of active CPO. In addition, the production of CPO in its native *C. fumago* host has been reported to survive 14 day period of cultivation and pH fluctuations ranging between pH 4-6.5 (Carmichael *et al*., 1986). Proper precautions appear to have been made to ensure production and maintenance of active CPO in the *Y. lipolytica* during cultivation.

The secretory expression of heterologous proteins in yeast is often subjected to several bottlenecks that affect yield (Idiris *et al*., 2010). One of the main limiting factors in heterologous recombinant protein production is the failure of the foreign polypeptides to be processed to their native conformation in host cells (Gasser *et al.,* 2008). Indeed,

the secreted CPO is a protein requiring several post-translational modifications, i.e. *N*- and *O*-glycosylation, prosthetic group incorporation, cleavage of *N*-terminal and *C*-terminal sequences and disulphide bond formation (Kenigsberg *et al*. 1987). The expression of recombinant active CPO in *E. coli* showed that glycosylation is not a mandatory requirement for refolding and activity of this enzyme (Zong *et al*., 1995). Similarly, the properties of recombinant CPO enzyme expressed in *A. niger* were not affected by overglycosylation (Conesa *et al*., 2001a). It is therefore assumed that the glycosylation of the CPO expressed in *Y. lipolytica* strains is most probably not the cause for the lack of CPO enzyme activity. The heme is reported as essential for the catalytic property of the CPO protein (Conesa *et al*., 2001a). The growth medium that was used to produce CPO was supplemented with hemin, and no CPO activity was observed in culture medium suggesting that heme is not the causal effect for the absence of CPO activity.

The N-terminal processing of the recombinant CPO protein would entail effective cleavage of the LACC or LIP2 secretion signal upon translocating the protein to the *Y. lipolytica* extracellular fluid. The LIP2 and the LACC secretion signals have been selected due to their reported superiority in targeting heterologous proteins for extracellular secretion (Jolivalt *et al*., 2005; Moon *et al*., 2013). The Western blot analyses of extracellular proteins produced in *Y. lipolytica* revealed His-tagged proteins of 38-42 kDa, the size corresponding to that of the recombinant CPO. This suggested that the LACC and LIP2 secretion signal peptides were effectively recognized and processed by the *Y. lipolytica* secretory pathway. However, it was not possible to determine the N-terminal sequence of the extracellular proteins to confirm successful processing of the N-terminal sequence of the recombinant CPO protein given that the purification of the CPO protein was not successful. The CPO protein has a disulphide bond that is critical for its structure and function (Sundaramoorthy *et al.,* 1995). While the *Y. lipolytica* yeast has been reported to successfully express proteins requiring disulphide bond formation (Swennen *et al*., 2002; Theerachat *et al*., 2012) it was not possible to establish in this study if CPO inactivity is related to misfolding due to improper disulphide bond formation. The fusion of the 6X His tag to the C-terminus of *Rhodobacter sphaeroides* YedY was found to be detrimental to the enzyme's

reductase activity and resulted in an eight-fold decrease in catalytic efficiency (Sabaty *et al*., 2013). When a cleavable His-tag was fused to the N-terminus of the mature enzyme in the absence of the signal sequence, YedY was expressed and folded with its cofactor. The study underscored the risk of using a tagged C-terminus with certain enzymes, and presented an alternative strategy to express signal sequence-containing enzymes with an N-terminal tag (Sabaty et al., 2013). This study did not investigate the effect of the presence and location of the 6X His-tag on CPO expression and maturation into functional enzyme.

It is worthy to note the suggestion that the active mature native CPO lacks at the carboxyl terminal the last 52 amino acid sequence, due to proteolytic processing that occurs downstream of the dibasic Lys-Arg (KR) processing site (Conesa *et al*., 2001b). The KR motif resembles the recognition sequence for cleavage by Golgi-located endoproteases encoded by the KEX2 gene in the *S. cerevisiae* yeast (Fuller *et al*., 1989). In *Y. lipolytica*, the protein is encoded by the XPR6 gene (Enderlin and Ogrydziak, 1994). The homologue of the KEX2 and the XPR6 genes in *A. niger* is the *pcl*A gene; the product of this gene has been implicated in the processing of carboxyl terminal propetides (Conesa *et al*., 2001b, Jalving *et al*., 2000; Punt *et al*., 2003). In the absence of the *pcl*A gene, the mutant *A. niger* produces inactive full length CPO inclusive of the 52 amino acid long carboxyl terminal propetide. It was concluded by Conesa *et al*., (2001b) that the carboxyl terminal propetide in CPO acts like a chaperone, being required to ensure correct CPO protein maturation, while its removal at a later stage during secretion is required to yield an active CPO enzyme. In this study, the His tag tail was placed downstream to the 52 aa long carboxyl terminal propeptide, and its detection using poly anti-His antibodies during Western blot studies and at position corresponding to full length CPO protein suggested the failure by the *Y. lipolytica* to process through proteolytic cleavage, the 52 aa propeptide, resulting in an inactive recombinant CPO protein. The *Y. lipolytica* yeast has been reported to possess the XPR6 sequence, which codes for KR dibasic processing endoprotease (Enderlin and Ogrydziak, 1994). The encoded Kex2-like protease is 42% identical to the ORF encoded by the KexB protein from *A. niger* with endoprotease activity (Jalving *et al*., 2000). It is apparent that despite the similarities between the endoproteases, the *Y. lipolytica* Kex2-like endoprotease failed to process the recombinant CPO propeptide, at least under the cultivation and protein production conditions described in this study.

In conclusion, the expression of active CPO in *Y. lipolytica* yeast failed despite using codon optimized gene sequences, transcriptionally highly active constitutive, quasi-constitutive and inducible promoters, utilization of secretion signals known to be efficient, and employment of cultivation conditions established to enhance heterologous protein production and CPO. All attempts reported in literature on the expression of active CPO have not been successful, apart of the fungal *A. niger* expression system. The *A. niger* host was found to contain the *pcl*A gene encoding endoprotease activity capable of processing carboxyl terminal propetides. The CPO gene expressed in *Y. lipolytica* was found to contain unprocessed carboxyl terminal propeptide, most likely rendering the recombinant CPO inactive. This postulation calls for future attempts in heterologous expression of CPO enzyme in *Y. lipolytica* to include development of platforms co-expressing compatible endoproteases to enable efficient processing of the chaperone-like 52 amino acid long carboxyl terminal propetide. The *A. niger pcl*A gene would be worth considering as a co-expression partner during CPO expression in *Y. lipolytica* yeast given its reported endoprotease activity in processing the protein's propeptide. To this end, *Y. lipolytica* molecular and protein expression tools exist to enable co-expression of the CPO gene with the gene encoding endoprotease as intracellular, extracellular protein or displayed on the yeast cell surface. It would be imperative for future studies aimed at successful expression of CPO in *Y. lipolytica* to take into account the presence and location of the 6X His tag, as the tag might have detrimental effect on the activity or maturation of the protein into a functional enzyme.

CHAPTER 4

Cloning and expression of human Granulocyte Colony Stimulating Factor in *Yarrowia lipolytica* yeast

4.1. INTRODUCTION

The human granulocyte colony stimulating factor (hG-CSF) is a glycoprotein used to regulate the production, differentiation and survival of neutrophils, bone marrow-associated cells, and macrophages (Metcalf, 1985). It is a hematopoietic growth factor required for the proliferation and differentiation of hematopoietic precursors of neutrophil granulocytes (Xu *et al*., 2000). The other therapeutic applications include the treatment of myelodysplastics and myelolosuppression syndromes resulting from chemotherapy and bone marrow transplantation (Xu *et al*., 2000), aplastic anemia (Dedhar *et al*., 1988; Liesechke *et al*., 1992 Dale *et al*., 2002; Kroger and Zander, 2002), neurodegenerative diseases (Schneider *et al*., 2005); myocardial infarction (Oh *et al*., 2006), cerebral ischemia (Lu *et al*., 2006), inflammation (Lieschke *et al*., 1994; Anderlini *et al*., 1996), maintenance of steady state hematopoiesis and skeletal muscle regeneration therapy (Hara *et al*., 2011).

The hG-CSF is a single chain polypeptide with a total molecular size of 18.8 kDa and is composed of 174 amino acid residues (Faraji *et al*., 2010). The hG-CSF glycoprotein contains two intra molecular disulfide bonds between Cys36-Cys42 and Cys64–Cys74 which are necessary for proper folding and biological activity (Wang *et al*., 2005; Abolghasemi *et al*., 2010; Basu *et al*., 2002; Faraji *et al*., 2010). The naturally occurring hG-CSF cytokine is an *O*-linked glycoprotein (Carter *et al*., 2004). There are two recombinant human granulocyte colony-stimulating factors that are clinically available, a glycosylated (lenograstim) and a non-glycosylated (neupogen) form (Pedrazzoli *et al*., 1996; Bönig *et al*., 2001). Accumulated evidence suggests that lenograstim the glycosylated form of hG-CSF possesses greater bioactivity and confers advantages in terms of *in vitro* stability to temperature, pH and degradation by proteases (Nissen, 1993, Nohynek *et al*., 1996; Querol *et al*, 1999; Bönig *et al*., 2001).On the other hand an *in vivo* study on children with chemotherapy-induced neutropenia suggested that in terms of clinical effects on neutropenia, the two G-CSF preparations have identical activity (Bönig *et al*., 2001).

The prokaryotic host *E. coli* has been used as a host for the expression of the recombinant hG-CSF which is commercially available for therapeutic use as non-glycosylated hG-CSF (Wang *et al*., 1995; Querol *et al*, 1999; Jeong *et al*., 2001; Srinivasa *et al*., 2009). However, high level production of recombinant hG-CSF in *E. coli* is hampered by formation of insoluble inclusion bodies and the lack of proper metabolic machinery to effect desired glycosylation (Khasa *et al*., 2011). Novel strategies have been developed to compensate the lack of human glycosylation pathways in *E. coli* (DeFrees *et al*., 2006). The process involves post production enzymatic *N*-Acetylgalactosamine (GalNAc) glycosylation at specific serine and threonine residues in the hG-CSF proteins expressed without glycosylation in *E. coli*, followed by enzymatic transfer of sialic acid conjugated with Polyethylene glycol (PEG) to the introduced GalNAc residues (DeFrees *et al*., 2006).

As alternatives to *E. coli* based expression systems, there have been attempts to express recombinant hG-CSF in yeast-based host systems. However, the use of the *P. pastoris* AOXI (alcohol oxidase1) promoter resulted in extracellular recombinant hG-CSF accumulating as aggregates in the culture (Lasnik *et al*., 2001) and optimal production required the use of surfactants (Apte-Deshpande *et al*., 2009). The most efficient way to obtain monomeric, biologically active protein was complete denaturation by guanidine-HCl or urea and subsequent renaturation (Lasnik *et al*., 2001). The general drawback with the use of methanol induced AOX1 promoter is that the by-products of methanol have been reported to be toxic to yeast cells, thus requiring stringent control and monitoring of the fermentation processes in order to optimize recombinant protein yields (Bawa and Darby, 2012). The yeast *Saccharomyces cerevisisae* has also been considered for expression of recombinant hG-CSF (Bae *et al*., 1999). However the protein existed as large multimers in the culture broth due to strong hydrophobic interactions, thus requiring monomerization (Bae *et al*., 1999). The different glycosylation pathways between human and yeast has hampered the use of yeast for production of therapeutic proteins. Indeed, the expression of hG-CSF in the *P. pastoris* yeast resulted in diffused protein bands being attained on SDS-PAGE at the range of 28-35 kDa indicating differential glycosylation

(Bhatacharya *et al*., 2007). However, recent studies on hG-CSF expressed in glycoengineered *P. pastoris* argued for the use of this yeast as a platform for therapeutic recombinant hG-CSF production on the basis of equivalent *O*-linked glycosylation site at Thr134 on both the recombinant and endogenous hG-CSF (Gong *et al*., 2014).

The other systems such as the baculovirus-based or the stably transformed insect cell systems using *Drosophila* (S2) or *Spodoptera frugiperda* (SF9) and Chinese Harmster Ovary (CHO) derived cell lines have gained widespread use for routine production of some recombinant proteins. However, insect cell line-based platforms can be cumbersome to implement and do not correctly reconstitute complex mammalian *N*-glycans containing galactose or sialic acid residues (Kost *et al*., 2005). Furthermore, CHO-based expression systems may not be ideal for therapeutic protein production due to the addition of terminal galactose-α-1,3-galactose (α-Gal) epitopes during *N*-glycosylation of recombinant glycoproteins (Bosques *et al*., 2010). The α-Gal antigen is reportedly responsible for allergic hypersensitivities that could lead to adverse clinical events (Bosques *et al*., 2010).

The yeast *Y. lipolytica* is a well-known prolific producer of several endogenous proteins. This yeast has also been extensively used as a host to produce a plethora of heterologous proteins (Nicaud *et al*., 2002; Madzak *et al*., 2004). It grows at a high density on low-cost medium with extensively established fermentation processes and secretes high yields of extracellular proteins (Barth and Gaillardin, 1997). In addition, this yeast has the ability to perform post-translational processing similar to mammalian systems apart from divergent glycosylation pathways (Madzak *et al*., 2004, De Pourcq *et al*., 2012). The *Y. lipolytica* is currently the subject of glycoengineering studies with the ultimate objective of developing it as an efficient expression system for the production of glycoproteins with humanized glycans (De Pourcq *et al*., 2012). Although the “humanization” of the *Y. lipolytica* glycosylation pathway is not yet completed, however, the genetics of the glycosylation pathways have been studied well enough (Barnay-Verdier *et al*., 2004; Park, *et al*., 2011; De Pourcq *et al*., 2012) to enable immediate development of the yeast as host for production of at least the non-glycosylated form of hG-CSF. The objective of this study was therefore to explore the

possibility of expressing and secreting functional recombinant hG-CSF in the *Y. lipolytica*. At the time of conception of this study, there was no report in literature of any attempts on the expression of hG-CSF in the *Y. lipolytica*.

4.2. MATERIAL AND METHODS

4.2.1. General chemicals, reagents, kits and enzymes.

The sources of general chemicals, reagents, kits and enzymes used in this study are as described in Section 3.2.1, Chapter 3. The ifosfamide (Holoxan®) and Neupogen (Roche, Germany) were obtained from commercial sources.

4.2.2. Plasmids, microbial strains and cultivation conditions

The plasmids, microbial strains and cultivation conditions are as described in Section 3.2.2, Chapter 3. The gene encoding the mature hG-CSF was designed based on the hG-CSF mRNA (Genebank Accession No. M17706) and custom synthesized by Geneart (GmbH, Regensburg, Germany). The plasmid carrying the custom synthesized hG-CSF preceded by the *Yarowia lipolytica* LIP2 secretion signal was denoted pGA-GCSF.

4.2.3. Recombinant DNA techniques

DNA restriction digestion, ligation, DNA sequencing, plasmid DNA isolation, genomic DNA isolation, transformation of *E. coli* and *Y. lipolytica*, agarose gel electrophoresis and PCR reaction mixtures were as described in Section 3.2.3, Chapter 3. The PCR was done using Taq DNA Polymerase DNA polymerase using mature hG-CSF specific oligonucleotide primer pair hG-CSF-F (5'-CTGGGACCCGCCTCTTCTCTGCCCCAG TCTTTC-3') and hG-CSF-R (5'-GGGCTGGGCCAGGTGTCGCAGCACTCGGTAAGA-3') as forward and reverse primers, respectively.

4.2.4. Construction of expression vectors

The pG4-GCSF gene encoding the complete hG-CSF mature gene was custom synthesised with the 6XHis tag at the C-terminus (Figure 4.1). To construct pINA1293-LIP2-GCSF, the pG4-GCSF plasmid was digested with *Hind*III and *Avr*II to release the DNA fragment encoding the LIP2-hG-CSF-6XHis open reading frame. The *Hind*III-*Avr*II fragment was cloned into pINA1293 digested with similar enzymes to create pINA1293-LIP2-GCSF.

4.2.5. *Yarrowia lipolytica* transformation and identification of transformants

The constructed pINA1293-LIP2-GCSF was linearized with the *Not*I restriction enzyme to remove the bacterial plasmid moiety. The resultant LIP2-hG-CSF DNA expression cassette was used to transform *Y. lipolytica* Po1f as described in section 3.2.3, Chapter 3. The Ura^+ prototroph transformants were subjected to genomic DNA isolation and PCR screening using the hG-CSF gene-specific primer pair of hG-CSF-F and hG-CSF-R.

4.2.6. Expression of hG-CSF

A single colony of *Y. lipolytica* containing the Ura^+ prototroph transformant of hG-CSF expression cassette was used to inoculate 5 ml of YPD which served as a pre-inoculum to seed 250 ml YPD media in 500 ml Erlenmeyer shake flasks. The cultures were grown at 28°C with shaking at 250 rpm. Sample aliquots were collected at different times of cultivation to evaluate hG-CSF protein accumulation in the culture extracellular fluid.

4.2.7. Purification and SDS-PAGE analysis of recombinant hG-CSF

The Ni-NTA affinity chromatography purification of recombinant hG-CSF was done as described in Section 3.2.11, Chapter 3. The elution of bound His tagged protein was performed using 250 mM imidazole in 10 mM sodium phosphate buffer, pH 8.0. The SDS-PAGE analysis of the crude and of the purified hG-CSF was performed as described in Sections 3.2.12, Chapter 3. The purified hG-CSF was used for biological

assays. Two purified preparations of hG-CSF namely; hG-CSF-1 and hG-CSF-7 wereprepared and tested.

4.2.8. Peptide mapping of the recombinant hG-CSF

The protein band on SDS-PAGE suspected to be over-expressed hG-CSF was excised for peptide mapping using LC-MS/MS analysis as described in Section 3.2.14, Chapter 3.

4.2.9. *In vivo* hG-CSF biological assay

The *in vivo* biological assay of hG-CSF using mice was done at the University of Pretoria Biomedical Research Centre (UPBRC, Onderstepoort, South Africa). A total of 24 inbred, male SPF Balb-c mice that met the inclusion criteria (healthy mice, seven days acclimatization period within the UPBRC) were used in this study. Animals were housed under conditions of optimum light, temperature and humidity (12 h light-dark cycle, 22-25°C, and 41-58% humidity) with food and water provided. The number of animals used was the minimum necessary to demonstrate the consistent effect, and procedures used were approved by the University of Pretoria Animal Ethics Committee. Animals were subdivided into 4 groups: the recombinant hG-CSF-1, hG-CSF-7, the positive control non-glycosylated (neupogen) hG-CSF and negative control groups (6 mice per group) in a fully randomized order (Table 4.1). The mice were uniquely identified by means of ear notches with numbers. The study made use of the purified His-tagged hG-CSF-1 and hG-CSF-7 preparations, which are purified recombinant hG-CSF derived from independent *Y. lipolytica* transformations.

To induce neutropenia, on day zero all mice received 0.5 ml Ifosfamide (Holoxan®) at a concentration of 8.6 mg/ml per animal. Ifosfamide is a drug known to cause a severe decrease in the number of blood cells in the bone marrow. Animals in the three treated groups [(GCSF-1, GCSF-7 and the non-glycosylated GCSF (Neupogen) positive control] received multiple injections of 0.5 ml consist of 32 µg/ml in 0.1% Bovine serum albumin (BSA) in Tris-Cl buffer solution from day 1 to 4. The control group (group 4) was injected with vehicle (0.1% BSA in Tris-Cl) in the same schedule of administration. Six hours after the last injection, peripheral blood was collected from the orbital venus sinus. Smears were prepared and stained by May-Grűnwald Giemsa method (Ludwig

and Frick, 1987). The white blood cells were counted and expressed as total number of neutrophils. The complete report on the in vivo biological assay of the hG-CSF is available as **Appendix 1**.

Table 4.1. Allocation of mice into treatment groups

Animal nr	Weight (g)	Group	Average group weight	Group Std Deviation
2	21.8	1	23.4	1.21
4	23.3	1		
6	24.4	1		
11	24.9	1		
15	23.9	1		
20	22.3	1		
5	22.1	2	23.2	1.19
8	24.0	2		
12	23.9	2		
17	24.6	2		
23	22.7	2		
24	21.6	2		
1	23.2	3	23.2	1.16
7	24.0	3		
9	23.3	3		
16	22.2	3		
19	21.5	3		
22	24.7	3		
3	23.7	4	23.3	1.57
10	24.2	4		
13	25.2	4		
14	23.1	4		
18	22.7	4		
21	20.6	4		

Group 1: GCSF-1 protein, group 2: GCSF-7 protein, group 3: positive control, group 4 negative control

4.3. RESULTS

4.3.1. Codon optimization and synthesis of hG-CSF gene.

This study describes the expression of the gene encoding the mature recombinant hG-CSF in *Y. lipolytica* under control of the quasi-constitutive HP4D promoter and in the pINA1293 multicopy integration vector. The mature hG-CSF gene was cloned for extracellular secretion and easy purification by creation of in-frame fusions of LIP2 secretion signal and inclusion of a fragment encoding the 6X His tag sequence at the N- and C-termini, respectively (Figure 4.1). The fused ORF was custom synthesized and codon optimized to eliminate codon bias. The other parameters that were optimized to enhance optimal gene expression included elimination of cryptic splice sites, avoidance of stable RNA secondary structures and addition of stabilizing sequence elements that may affect heterologous gene transcription in the *Y. lipolytica* host.

*Hind*III
atg***aagctt***tccaccatccttttcacagcctgcgctaccctggctgccgccctcccttcc
M K L S T I L F T A C A T L A A A L P S
cccatcactccttctgaggccgcagttctgcagaagcgattcacccccctgggccctgcc
P I T P S E A A V L Q K R F T P L G P A
agctccctgccccagagcttcctgctcaagtgcttagagcaagtgaggaagatccagggc
S S L P Q S F L L K ***C*** L E Q V R **L** I Q G
gatggcgcagcgctccaggagaagctgtgtgccacctacaagctgtgccaccccgaggag
D G A A L Q E K L **C** A T Y K L **C** H P E E
ctggtgctgctcggacactctctgggcatcccctgggctcccctgagcagctgccccagc
L V L L G H S L G I P W A P L S S **C** P S
caggccctgcagctggcaggctgcttgagccaactccatagcggccttttcctctaccag
Q A L Q L A G **C** L S Q L H S G L F L Y Q
gggctcctgcaggccctggaagggatctcccccgagttgggtcccaccttggacacactg
G L L Q A L E G I S P E L G P T L D T L
cagctggacgtcgccgactttgccaccaccatctggcagcagatggaagaactgggaatg
Q L D V A D F A T T I W Q Q M E E L G M
gcccctgccctgcagcccacccagggtgccatgccggccttcgcctctgctttccagcgc
A P A L Q P T Q G A M P A F A S A F Q R
cgggcaggaggggtcctagttgcctcccatctgcagagcttcctggaggtgtcgtaccgc
R A G G V L V A S H L Q S F L E V S Y R
gttctacgccaccttgcccagccccaccaccaccaccaccactgataa***cctagg***
V L R H L A Q P H H H H H H - - *Avr*II

Figure 4.1. Nucleotide and protein sequences of hG-CSF codon optimised for *Y. lipolytica*. The 27 aa long sequence of the lip2p secretion signal upstream of the mature gene is underlined. The amino acids cys36-cys42, cys64-cys74 for disulphide bond formation (bold and thick underlined) the free cys17 (bold, italics and thick underlined) and leucine at position 23 (double underlined and bold) are also illustrated. The *Hind*III and *Avr*II restriction enzymes (sites in italics and underlined) were used in subcloning of the DNA fragments into pINA1293 plasmid.

4.3.2. Construction of the expression vector

The mature hG-CSF gene, including the sequences encoding the N-terminal LIP2 secretion signal and C-terminal 6XHis tag was cloned under the control of the HP4D promoter of pINA1293 to create the plasmid denoted pINA1293-LIP2-GCSF. The succesful cloning of the hG-CSF expression cassette in pINA1293 was confirmed by restriction analysis of the pINA1293-LIP2-GCSF plasmid (Figure 4.2). The restriction analysis using *Hind*III and *Avr*II resulted in the expected DNA fragments of about 600 and ~5000 bp, corresponding to the Lip2-hG-CSF and pINA1293 backbone fragments, respectively.

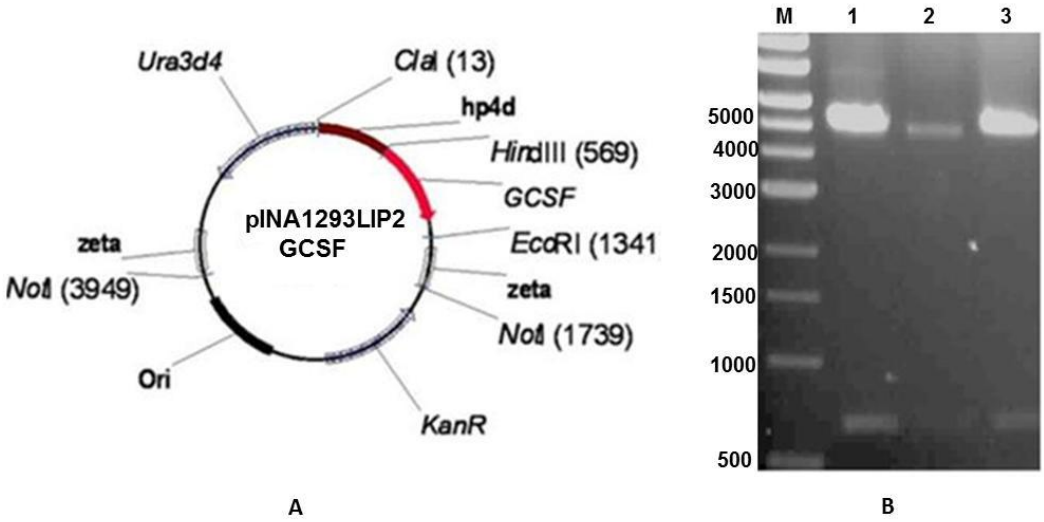

Figure 4.2: A. Schematic map of the pINA1293-LIP2-GCSF expression plasmid used to transform the *Y. lipolytica* Po1f strain. (B) Agarose gel (1%) showing restriction analysis of the pINA1293-LIP2-GCSF using *Hind*III and *Avr*II restriction enzymes.

4.3.3. Expression cassette integration in *Yarrowia lipolytica*

The transformation of *Y. lipolytica* Po1f with pINA1293-LIP2-GCSF expression cassette resulted in the selection on YNB casa plates of numerous Ura+ prototrophs. Three clones were selected and subjected to PCR screening using the gene specific primer pair GCSF-F and GCSF-R to confirm integration of the cassette within the Ura+ prototrophs of the *Y. lipolytica* Po1f strain. A PCR fragment of 500 bp was expected and obtained from two of the clones (Figure 4.3). The positively screened clones represented in Figure 4.3 as Lanes 1 and 3 were denoted GCSF-1 and GCSF-7.

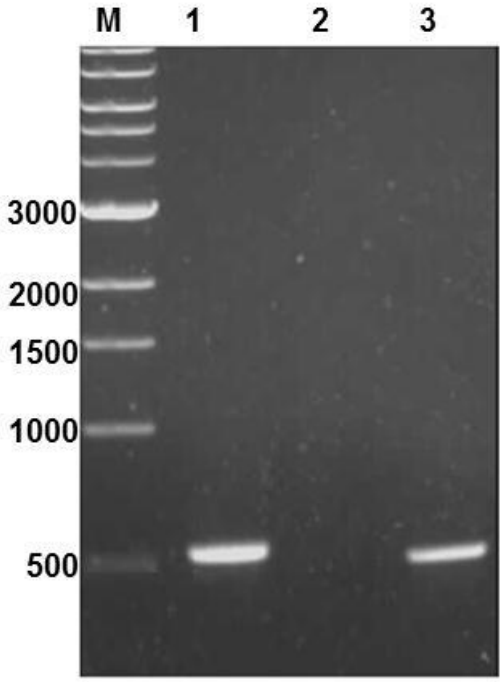

Figure 4.3: A 1% agarose gel electrophoresis showing PCR screening result of the *Y. lipolytica* Po1f Ura+ prototrophs transformed with hG-CSF expression cassette. Lane 1 to 3 represent the different Ura+ prototrophs of *Y. lipolytica* Po1f transformed with the hG-CSF expression cassette. The positively screened clones represented by Lanes 1 and 3 were denoted GCSF-1 and GCSF-7. Lane M represents the DNA molecular weight marker.

4.3.4. Expression, purification and confirmation of hG-CSF by peptide mapping.

The *Y. lipolytica* Po1f GCSF-1 and GCSF-7 clones were cultured in YPD medium in shake flask and aliquot samples were withdrawn for hG-CSF expression analysis on SDS-PAGE. The hG-CSF-7 protein expression analysis data is presented in Figure 4.4. The SDS-PAGE gel analysis revealed the presence of a prominent band of approximately 17-18 kDa, the protein band was suspected to be the recombinant hG-CSF protein (Figure 4.4). The band was excised digested with trypsin and subjected to LC-MS/MS analysis. The list of matching peptides is shown in Table 4.2. The data confirmed that the ~17 kDa protein secreted to the extracellular was indeed of the hG-CSF. The recombinant hG-CSF was constructed such that the ORF contained a 6XHis tag at the C-terminus, and this was to facilitate protein purification with Ni-NTA affinity chromatography. The bound protein was eluted to a single pure protein band of 17-18 kDa using increased concentrations of imidazole as analysed by SDS-PAGE (Figure 4.4). The M_r corresponded with the predicted molecular mass of the cloned hG-CSF gene with 6XHis tag which was predicted by the EXPASY protein analysis tool (Gasteiger *et al.,* 2003) to be 18.5 kDa.

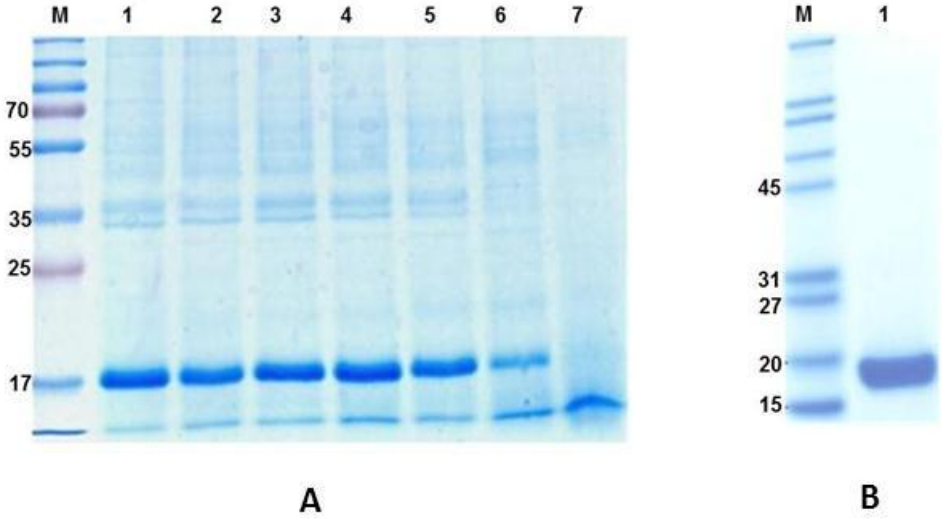

Figure 4.4: Extracellular protein profile of culture supernatants of *Y. lipolytica* GCSF7. The SDS-PAGE gel (12 %) is showing accumulation of the protein with samples harvested at different intervals in hours: 168 (lane 1), 144 (lane 2), 120 (lane 3), 96 (lane 4), 72 (lane 5), 48 (lane 6), 24 (lane 7). M is the protein molecular weight marker used as a standard**. (B):** Protein profile of GCSF-7 purified using Ni-NTA matrix resin. SDS-PAGE gel (12 %) showing eluted GCSF-7 (lane 1) M=Protein molecular mass marker used as standard.

Table 4.2. LC-MS/MS analysis of hG-GCSF after trypsinolysis.

Peptide sequence of recombinant hG-CSF protein	Positions on mature hG-CSF sequence
TPLGPASSLPQSFLLK	1-16
LCHPEELVLLGH	41-51
CLEQVRKIQGDGAALQEK	17-34
QGAMPAFASAFQRRAGGVLV	147-166

4.3.5. Recombinant GCSF in vivo bioactivity Assay

To determine the *in vivo* activity of the recombinant hG-CSF expressed in *Y. lipolytica* GCSF-1 and GCSF-7 strains, SPF Bulb/C male mice model were used. The mean neutrophil counts per treatment group are presented in Figure 4.5. In summary, GCSF-7 induced marginal increase of 7.75%, while GCSF-1 was the same as the negative control. However, the marginal increase was not statistically significant to conclude on succesful expression of active recombinant hG-CSF.

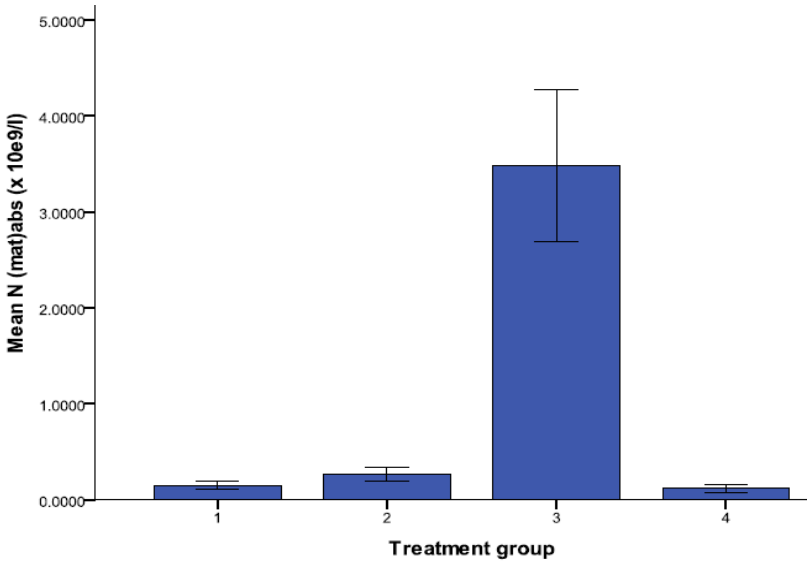

Figure 4.5: Mean neutrophil counts for each mice treatment group. Treatment group 1: GCSF-1 protein, group 2: GCSF-7 protein, group 3: commercial recombinant hG-CSF neupogen as positive control group 4: negative control.

4.4. DISCUSSION AND CONCLUSION

The expression cassette comprising of Ura3d4 auxotrophic marker, HP4D as promoter, LIP2 secretion signal, codon optimized mature hG-CSF gene, the terminator of LIP2 gene, and zeta sequences was used to generate *Y. lipolytica* Po1f Ura^+ prototrophic strains secreting recombinant hG-CSF protein. The recombinant hG-CSF contained 6XHis tag at the C-terminus and was purified by Ni-NTA affinity chromatography to homogeneity as judged by the single protein band with M_r of about 18 kDa obtained on SDS-PAGE analysis. The protein was confirmed to be hG-CSF by peptide mapping using LC-MS/MS analysis.

During the course of this study, Gasmi *et al.*, (2012) reported on the production and characterization of extracellularly secreted recombinant hG-CSF in the *Y. lipolytica* yeast under the POX2 promoter in single and double copy expression systems using modified LIP2 sequence as secretion signal. The key finding of their study was that the hG-CSF was secreted as a nonglycosylated form of 14.5 kDa in addition to a trail of hyperglycosylated forms of hG-CSF with M_r varying from 17 kDa to more than 60 kDa. The expression of hG-CSF in the *P. pastoris* yeast has also been reported to result in diffused protein bands ranging between 28-35 kDa, and this was explained as differential glycosylation (Bhatacharya *et al*., 2007). The 17 kDa protein fragment obtained using the *Y. lipolytica* expression system was determined to be more active than the hyperglycosylated forms. However, the stability of this protein was found to be affected at pH 3 and 4 most probably as a result of proteolytic degradation. However, the hyperglycosylated forms were more stable at pH 3-4 (Gasmi *et al*., 2012). It is interesting to note that the prominently expressed protein in this study was the 17-18 kDa protein band as judged by SDS-PAGE analysis and recovery after protein purification using the His tag Ni-NTA affinity chromatography. In contrast to the Gasmi *et al.* (2012) recombinant hG-SCF, the hG-CSF described in this study contained the His tag at the C-terminus, and its biological activity was found to be very low in comparison with the commercially available hG-CSF. It has not been determined if the presence of the His tag at the C-terminus of the hG-CSF expressed in this study is affecting its biological activity.

The methods of assay of the Gasmi *et al.,* (2012) recombinant hG-SCF is different from the method employed in this study. This study made use of *in vivo* biological activity method of assay using the SPF Bulb/C mice model with the biological activity being expressed as neutrophil count percentage induced by recombinant hG-CSF derived from *Y. lipolytica* against the neutrophil count as induced by commercial hG-CSF neupogen. The hG-CSF constructed by Gasmi *et al*., (2012) was tested for biological activity using the TF-1 cell line. The comparative analysis of the hG-CSF described in this study with the commercially available hG-CSF, pointed to lack of biological activity. The possible explanation for the lack of biological activity could be attributed to the production conditions, the protein expression was done at pH 4.0, and according to Gasmi *et al*., (2012) hG-CSF function is affected at this pH affect. The comparative biological activity assay of the Gasmi *et al*., (2012) against any commercial hG-CSF was not reported. The C-terminal His tag could also have contributed to the lack of biological activity.

In conclusion, collectively, the studies on hG-CSF production highlighted the potential application of the *Y. lipolytica* yeast as platform for the recombinant production of this protein. The factors to be considered in future studies on the production of hG-CSF protein would entail the design of appropriate cultivation conditions to enhance the stability of hG-CSF during production, removal of the His tag, and the use of *Y. lipolytica* strains deleted for the yeast glycosylation pathway for the production of at least the nonglycosylated form of hG-CSF.

CHAPTER 5

SUMMARY AND GENERAL CONCLUSIONS

The *C. fumago* chloroperoxidase (CPO) and human Granulocyte-Colony Stimulating Factor (hG-CSF) proteins require post-translation glycosylation modifications, contain intra molecular disulfide bonds for proper folding and biological activity, and have presented challenges when heterologously expressed in yeast, bacterial, insect cells and mammalian expression systems. The constraints associated with recombinant production of the proteins included toxicity to the host cell, inappropriate post-translational modifications, low production yields, improper protein folding, and lack of biological activity. The objective of this study was to explore the versatility of *Y. lipolytica* as host for expression and extracellular secretion of these two commercially significant proteins. The *Y. lipolytica* yeast was selected for its inherent ability to secrete a variety of proteins, its reported cotranslational translocation protein secretion pathway similar to that of mammalian systems, low overglycosylation, high protein secretion efficiency, good protein product yield, and heterologous protein production performance reproducibility. The codon optimized synthetic CPO and hG-CSF genes were cloned under the quasi constitutive growth dependent HP4D promoter or the inducible POX2 promoter in multicopy zeta based integrative expression systems using the *Trimetes vesicolor* laccase (LACC) or *Y. lipolytica* extracellular lipase (LIP2) secretion signals to target protein expression to the extracellular.

The CPO protein was expressed as an inactive extracellularly located protein judging by Western blot analysis using poly-Histidine antibody. The antibody detected the His tag fused to the C-terminus of the recombinant protein. The detection of this His tag suggested the failure by the *Y. lipolytica* to process through proteolytic cleavage, the 52 amino acid propeptide located at the C-terminal end of the protein. The 52 amino acid propeptide acts like a chaperone, being required to ensure correct CPO protein maturation, and its removal at a later stage during CPO secretion is required to yield an active CPO enzyme.

The recombinant hG-CSF protein was successfully expressed in *Y. lipolytica* and purified by Ni-NTA affinity chromatography to homogeneity as judged by the single protein band with relative molecular mass of about 18 kDa. The protein was confirmed to be hG-CSF by peptide mapping using LC-MS/MS analysis. The biological activity of the expressed hG-CSF was found to be very low in comparison with the commercially available nonglycosylated hG-CSF. It remains to be determined if the His tag fused to the C-terminal end of the recombinant protein affected the biological activity of the recombinant hG-CSF. The possible explanation for the low biological activity could be the cultivation conditions that did not take into account pH ranges at which hG-CSF retains stability.

The study demonstrated the potential application of the *Y. lipolytica* yeast as platform for production of recombinant CPO and hG-CSF proteins. The proteins were successfully expressed and secreted to the extracellular. However, the SDS-PAGE analysis revealed that the CPO on the extracellular was less pronounced as compared to hG-CSF protein. Most importantly, this study provided insights on the requirements that have to be considered when developing the *Y. lipolytica* yeast as production systems for recombinant CPO and hG-CSF. The future attempts in heterologous expression of CPO enzyme in *Y. lipolytica* need to include development of platforms to co-express compatible endoprotease genes to enable efficient processing of the chaperonic 52 amino acid long carboxyl terminal propetide. The factors to be considered in future studies on the production of hG-CSF protein would entail the design of appropriate cultivation conditions to enhance the stability of hG-CSF during production, and the use of *Y. lipolytica* strains deleted for the yeast glycosylation pathway.

OPSOMMING

Die *C. fumago* chloroperoksidase (CPO) en menslike Granulosiet-kolonie stimulerende faktor (hG-CSF) proteïene benodig post-translasie glikosileringveranderinge en bevat intra-molekulêre disulfiedbindings vir 'n behoorlike vouïng en biologiese aktiwiteit. Dit bied dus uitdagings wanneer dit heteroloog uitgedruk word in gis, bakteriële, insek en soogdier uitdrukkingsisteme. Die beperkinge wat verband hou met rekombinante produksie van die proteïen sluit in toksisiteit vir die gasheersel, onvanpaste post-translasie modifikasies, lae produksie opbrengste, onbehoorlike proteïenvouing en 'n gebrek aan biologiese aktiwiteit. Die doel van hierdie studie was om die veelsydigheid van *Y. lipolytica* as gasheer vir die uitdrukking en ekstrasellulêre uitskeiding van hierdie twee kommersieel belangrike proteïene te ondersoek. Die *Y. lipolytica* gis is gekies vir sy inherente vermoë om 'n verskeidenheid proteïene uit te skei, sy gerapporteerde ko-translokasie proteïen uitskeidingsweg, soortgelyk aan dié van ander soogdier sisteme, lae oor-glikosilering, hoë proteïen uitskeiding doeltreffendheid, goeie proteïen produk opbrengs en reproduseerbare hoë heteroloë proteïen produksie. Die kodon geöptimiseerde, gesintetiseerde CPO en hG-CSF gene is gekloneer onder die kwasi konstitutiewe, groei-afhanklike hp4d promotor of die induseerbare POX2 promotor multikopie zeta gebaseerde geïntegreerde uitdrukkingstelsels met behulp van die Trimetes vesicolor laccase (LACC) of *Y. lipolytica* ekstrasellulêre lipase (LIP2) uitskeidingseine om proteïen uitdrukking na die ekstrasellulêre medium te teiken.

Die CPO proteïen is uitgedruk as 'n onaktiewe ekstrasellulêre proteïen volgens die Western-klad analise gedoen met poli-Histidien teenliggaampies. Die teenliggaampies bespeur die His-gemerkte C-terminus van die rekombinante proteïen. Die teenwoordigheid van die His-merker dui daarop dat *Y. lipolytica* nie instaat was tot proteolitiese splyting van die 52 aminosuur propeptide geleë op die C-terminus van die proteïen. Die 52 aminosuur propeptide tree op as 'n chaperone, en is noodsaaklik vir korrekte prosessering van die CPO proteïen. Verwydering op 'n later stadium tydens CPO uitskeiding is noodsaaklik om 'n aktiewe CPO ensiem te verkry.

Die rekombinante hG-CSF proteïen is suksesvol uitgedruk in *Y. lipolytica* en gesuiwer deur Ni-NTA affiniteitschromatografie on 'n suiwer proteïen te lewer te oordeel aan die

enkele proteïen-band met relatiewe molekulêre massa van ongeveer 18 kDa. Die proteïen is geïdentifiseer as hG-CSF deur peptied kartering met behulp van LC-MS / MS analise. Die biologiese aktiwiteit van die uitgedrukte hG-CSF was baie laag in vergelyking met die kommersieel beskikbare nie-geglikosideerde hG-CSF. Dit moet nog bepaal word of die His-merker aan die C-termius van die rekombinante proteïen die biologiese aktiwiteit van die rekombinante hG-CSF benadeel. 'n Ander moontlike verklaring vir die lae biologiese aktiwiteit kan wees dat die pH waarby proteïen produksie plaasgevind het nie die effek van pH op die stabilitiet van hG-CSF in ag geneem het nie.

Die studie toon die potensiaal van die *Y. lipolytica* gis as platform vir die produksie van rekombinante CPO en hG-CSF proteïene. Die proteïene is suksesvol uitgedruk en uitgeskei na die ekstrasellulêre medium. SDS-PAGE analise het getoon dat die CPO teen laer vlakke uitgedruk is as die hG-CSF proteïen. Hierdie studie het belangrike insigte gelewer oor die vereistes wat in ag geneem moet word tydens die ontwikkeling van die *Y. lipolytica* gis as produksiestelsels vir rekombinante CPO en hG-CSF. Toekomstige pogings om CPO ensiem in *Y. lipolytica* uit te druk moet versoenbaar wees met die verwydering van die 52 aminosuur C-terminale propeptied deur endoproteases. Die faktore wat in ag geneem moet word in toekomstige studies op die produksie van hG-CSF proteïen moet insluit produksie toestande wat stabiliteit van die hG-CSF proteïen sal verseker en die gebruik van *Y. lipolytica* stamme waarvan die gis glikosideringsweg geïnaktiveer is.

CHAPTER 6

LIST OF REFERENCES

Ageitos JM, Vallejo JA, Veiga-Crespo P, Villa TG. 2011. Oily yeasts as oleaginous cell factories. *Appl Microbiol Biotechnol*. 90, 1219-1227.

Abolghasemi DS, Babaeipour V, Modif MR, Divsalar A, Faraji F. 2010. An efficient purification method for high recovery of recombinant human granulocyte colony-stimulating factor from *Escherichia coli*. *International J. Environ. Sci. Dev*. 1 & 2.

Allain EA, Hager LP, Deng L, Jacobsen EN.1993. Highly enantioselective epoxidation of disubstituted alkenes with hydrogen peroxide catalyzed by chloroperoxidase. *J. Am. Chem. Soc*.115, 4415–4416.

Anderlini P, Przepiorka D, Champlin R, Körbling M. 1996. Biologic and clinical effects of granulocyte colony-stimulating factor in normal individuals. *Blood*. 88, 2819-2825.

Andersson LA, Dawson JH.1991. EXAFS spectroscopy of hemecontaining oxygenases and peroxidases. *Struct Bond*. 64, 1–40.

Apte-Deshpande A, Somani S, Mandal G, Soorapaneni S, Padmanabhan S. 2009. Over expression and analysis of O-glycosylated recombinant human granulocyte colony stimulating factor in *Pichia pastoris* using Agilent 2100 Bioanalyzer. *J Biotechnol.* 143, 44-50.

Bae CS, Yang DS, Chang KR, Seong BL, Lee J. 1998. Enhanced secretion of human granulocyte colony-stimulating factor directed by a novel hybrid fusion peptide from recombinant *Saccharomyces cerevisiae* at high cell concentration. *Biotechnol Bioeng*. 57, 600-609.

Bae CS, Yang DS., Lee J, Park YH. 1999. Improved process for production of recombinant yeast-derived monomeric human G-CSF. *Appl Microbiol Biotechnol.* 52, 338-344.

Bahrami A, Shojaosadati S A, Khalilzadeh R, Mohammadian J, Farahani EV., Masoumian M.R. 2009. Prevention of human granulocyte colony-stimulating

factor protein aggregation in recombinant *Pichia pastoris* fed-batch fermentation using additives. *Biotechnol Appl Biochem*.52, 141-148.

Bankar AV, Kumar AR, Zinjarde SS. 2009. Environmental and industrial applications of *Yarrowia lipolytica*. *Appl Microbiol Biotechnol.* 84, 847-865.

Barnay-Verdier S, Boisrame A, Beckerich JM. 2004. Identification and characterization of two alpha-1, 6-mannosyltransferases, Anl1p and Och1p, in the yeast *Yarrowia lipolytica*. *Microbiology* 150, 2185–2195.

Barth G, Gaillardin C. 1996. *Yarrowia lipolytica*. In Nonconventional Yeasts in Biotechnology, Wolf K (ed). *Springer-Verlag: Berlin, Heidelberg*. 313–388.

Barth G, Gaillardin C. 1997. Physiology and genetics of the dimorphic fungus *Yarrowia lipolytica*. *FEMS Microbiol Rev*. 19, 219–237.

Basu S, Dunn A, Ward A. 2002. G-CSF: function and modes of action. *Int J Mol Med*. 10, 3-10.

Bawa Z, Darby RA. 2012. Optimising *Pichia pastoris* induction. *Methods Mol Biol*. 866, 181-190.

Bayramoglu G, Altintas B, Yilmaz M, Arica MY. 2011. Immobilization of chloroperoxidase onto highly hydrophilic polyethylene chains via bio-conjugation: catalytic properties and stabilities. *Bioresour Technol*. 102, 475-482.

Beckerich JM, Boisrame-Baudevin A, Gaillardin C. 1998. *Yarrowia lipolytica*: a model organism for protein secretion studies. *Int Microbiol* 1, 123–130.

Bhatacharya P, Pandey G, Mukherjee KJ. 2007. Production and purification of recombinant human granulocyte-macrophage colony stimulating factor (GM-CSF) from high cell density cultures of *Pichia pastoris*. *Bioprocess Biosyst Eng*. 30, 305-312.

Beopoulos A, Cescut J, Haddouche R, Uribelarrea JL, Molina-Jouve C, Nicaud JM. 2009. *Yarrowia lipolytica* as a model for bio-oil production. *Prog Lipid Res*. 48, 375-387.

Beopoulos A, Desfougeres T, Sabirova J, Zinjarde S, Neuvéglise C, Nicaud JM. 2010. The hydrocarbon-degrading oleaginous yeast *Yarrowia lipolytica*. *In Handbook of Hydrocarbon and Lipid Microbiology*. 2111-2121.

Beopoulos A, Haddouche R, Kabran P, Dulermo T, Chardot T, Nicaud JM. 2012. Identification and characterization of DGA2, an acyltransferase of the DGAT1 acyl-CoA:diacylglycerol acyltransferase family in the oleaginous yeast *Yarrowia lipolytica*. New insights into the storage lipid metabolism of oleaginous yeasts. *Appl Microbiol Biotechnol*. 93, 1523-1537.

Beopoulos A, Mrozova Z, Thevenieau F, Le Dall MT, Hapala I, Papanikolaou S, Chardot T, Nicaud JM. 2008. Control of lipid accumulation in the yeast *Yarrowia lipolytica*. *Appl Environ Microbiol*. 74, 7779-7789.

Beopoulos A, Verbeke J, Bordes F, Guicherd M, Bressy M, Marty A, Nicaud JM. 2014. Metabolic engineering for ricinoleic acid production in the oleaginous yeast *Yarrowia lipolytica*. Appl *Microbiol Biotechnol.* 98, 251-262.

Bergès H, Joseph-Liauzun E, Fayet O. 1996. Combined effects of the signal sequence and the major chaperone proteins on the export of human cytokines in *Escherichia coli*. *Appl Environ Microbiol*. 62**,** 55-60.

Bhattacharya P, Pandey G, Srivastava P, Mukherjee KJ. 2007. Production and purification of recombinant human granulocyte-macrophage colony stimulating factor (GM-CSF) from high cell density cultures of *Pichia pastoris. Bioprocess and Biosystem Eng.* 30, 305-312.

Blanchin-Roland S, Cordero Otero RR, Gaillardin C. 1994. Two upstream activation sequences control the expression of the XPR2 gene in the yeast *Yarrowia lipolytica*. *Mol Cell Biol*. 14, 327-338.

Blanchin-Roland S, Da Costa G, Gaillardin C. 2008. Ambient pH signalling in the yeast *Yarrowia lipolytica* involves YlRim23p/PalC, which interacts with Snf7p/Vps32p, but does not require the long C terminus of YlRim9p/PalI. *Microbiology.* 154, 1668-1676.

Blanke SR, Hager LP. 1988. Identification of the fifth axial heme ligand of chloroperoxidase. *J Biol Chem.* 263, 18739-18743.

Bönig H, Silbermann S, Weller S, Kirschke R, Körholz D, Janssen G, Göbel U, Nürnberger W. 2001. Glycosylated vs non-glycosylated granulocyte colony-stimulating factor (G-CSF) results of a prospective randomised monocentre study. *Bone Marrow Transplant*. 28, 259-264.

Boonvitthya N, Bozonnet S, Burapatana V, O'Donohue MJ, Chulalaksananukul W. 2013. Comparison of the heterologous expression of *Trichoderma reesei* endoglucanase II and cellobiohydrolase II in the yeasts *Pichia pastoris* and *Yarrowia lipolytica*. *Mol Biotechnol.* 54, 158-169.

Borders F, Fudalej F, Dossat V, Nicaud JM, Marty A. 2007. A new recombinant protein expression system for high through put screening in the yeast *Yarrowia lipolyitica*. *J. Microbiol*. *Methods*. 70, 493-502.

Boisramé A, Chasles M, Babour A, Beckerich JM, Gaillardin C. 2002. Sbh1p, a subunit of the Sec61 translocon, interacts with the chaperone calnexin in the yeast *Yarrowia lipolytica*. *J Cell Sci.* 115, 4947-4956.

Bosques CJ, Collins BE, Meador JW 3rd, Sarvaiya H, Murphy JL, Dellorusso G, Bulik DA, Hsu IH, Washburn N, Sipsey SF, Myette JR, Raman R, Shriver Z, Sasisekharan R, Venkataraman G. 2010. Chinese hamster ovary cells can produce galactose-α-1,3-galactose antigens on proteins. *Nat Biotechnol*. 28, 1153-1156.

Braun A, Geier M, Bühler B, Schmid A, Mauersberger S, Glieder A. 2012. Steroid biotransformations in biphasic systems with *Yarrowia lipolytica* expressing human liver cytochrome P450 genes. *Microb Cell Fact*. 11, 106. doi: 10.1186/1475-2859-11-106.

Buchhaupt M, Ehrich K, Hüttmann S, Guder J, Schrader J. 2011. Over-expression of chloroperoxidase in *Caldariomyces fumago*. *Biotechnol Lett*. 33, 2225-2231.

Carmichael RD, Jones A, Pickard MA. 1986. Semicontinuous and Continuous Production of Chloroperoxidase by *Caldariomyces fumago* Immobilized in k-Carrageenan. *Appl Environ Microbiol*. 51, 276-280.

Carter CRD, Whitmore KM, Thorpe R. 2004. The significance of carbohydrates on G-CSF: differential sensitivity of GCSF's to human neutrophil elastase degradation. *J. Leukoc. Biol*. 75, 515–522.

Casarégola S, Feynerol C, Diez M, Fournier P, Gaillardin C. 1997. Genomic organization of the yeast *Yarrowia lipolytica. Chromosoma*. 106, 380-390.

Casarégola S, Neuvéglise C, Lépingle A, Bon E, Feynerol C, Artiguenave F, Wincker P, Gaillardin C. 2000. Genomic exploration of the hemiascomycetous yeasts: 17. *Yarrowia lipolytica*. *FEBS Lett*. 487, 95-100.

Casella L, Gullotti M, Ghezzi R, Poli S, Beringhelli T, Colonna S, Carrea G. 1992. Mechanism of enantioselective oxygenation of sulfides catalyzed by chloroperoxidase and horseradish peroxidase. Spectral studies and characterization of enzyme-substrate complexes. *Biochemistry*. 31, 9451-9459.

Celik E, Calık P. 2012. Production of recombinant proteins by yeast cells. *Biotechnol Adv.* 30, 1108-1118.

Chi ZM, Liu G, Zhao S, Li J, Peng Y. 2010. Marine yeasts as biocontrol agents and producers of bio-products. *Appl Microbiol Biotechnol*. 86, 1227-1241.

Colonna S, Gaggero N, Manfredi A, Casella L, Gullotti M, Carrea G, Pasta P. 1990. Enantioselective oxidations of sulfides catalyzed by chloroperoxidase. *Biochemistry.* 29, 10465-10468.

Conesa A, van De Velde F, van Rantwijk F, Sheldon RA, van Den Hondel CA, Punt PJ. 2001 (A). Expression of the *Caldariomyces fumago* chloroperoxidase in and characterization of the recombinant enzyme. *J Biol Chem.* 276, 17635-17640.

Conesa A, Weelink G, van den Hondel CA, Punt PJ. 2001 (B). C-terminal propeptide of the *Caldariomyces fumago* chloroperoxidase: an intramolecular chaperone? *FEBS Lett.* 503, 117-120.

Corbett MD, Chipko BR, Baden DG. 1978. Chloroperoxidasecatalysed oxidation of 4-chloroaniline to 4-chloronitrosobenze. *Biochem J*. 175, 353–360.

Cordero Otero R, Gaillardin C. 1996. Efficient selection of hygromycin-B-resistant *Yarrowia lipolytica* transformants. *Appl Microbiol Biotechnol*. 46, 143-148.

Dale DC. 2002. Colony-stimulating factors for the management of neutropenia in cancer patients. *Drugs*.1, 1-15.

Darvishi F, Destain J, Nahvi I, Thonart P, Zarkesh-Esfahani H. 2011. High-level production of extracellular lipase by *Yarrowia lipolytica* mutants from methyl oleate. *N Biotechnol*. 28, 756-760.

Das KM, Banerjee S, Shekhar N, Damodaran K, Nair R, Somani S, Raiker VP, Jain S, Padmanabhan S. 2011. Cloning, Soluble Expression and Purification of High Yield Recombinant hGMCSF in *Escherichia coli*. *Int J Mol Sci*. 12, 2064-2076.

De Pourcq K, Tiels P, Van Hecke A, Geysens S, Vervecken W, Callewaert N. 2012. Engineering *Yarrowia lipolytica* to produce glycoproteins homogeneously modified with the universal Man3GlcNAc2 N-glycan core. *PLoS One*. 7, 39976. doi: 10.1371/journal.pone.0039976.

Dedhar S, Gaboury L, Galloway P, Eaves C. 1988. Human granulocyte-macrophage colony-stimulating factor is a growth factor active on a variety of cell types of nonhemopoietic origin. *Proc Natl Acad Sci U S A*. 85, 9253-9257.

DeFrees S, Wang ZG, Xing R, Scott AE, Wang J, Zopf D, Gouty DL, Sjoberg ER, Panneerselvam K, Brinkman-Van der Linden EC, Bayer RJ, Tarp MA, Clausen H. 2006. GlycoPEGylation of recombinant therapeutic proteins produced in *Escherichia coli*. *Glycobiology*. 16, 833-843.

Delic M, Valli M, Graf AB, Pfeffer M, Mattanovich D, Gasser B. 2013. The secretory pathway: exploring yeast diversity. *FEMS Microbiol Rev*. 37, 872-914.

Doerge DR, Corbett MD. 1991. Peroxygenation mechanism for chloroperoxidase-catalyzed N-oxidation of arylamines. *Chem Res Toxicol*. 4, 556-560.

Domínguez A, Fermiñán E, Sánchez M, González FJ, Pérez-Campo FM, García S, Herrero AB, San Vicente A, Cabello J, Prado M, Iglesias FJ, Choupina A, Burguillo FJ, Fernández-Lago L, López MC. 1998. Non-conventional yeasts as hosts for heterologous protein production. *Int Microbiol.* 1, 131-142.

Dorr RT, 1993. Clinical properties of yeast-derived versus *Escherichia coli*-derived granulocyte-macrophage colony-stimulating factor. *Clin Therapeutics*.15, 19-29.

Emond S, Montanier C, Nicaud JM, Marty A, Monsan P, André I, Remaud-Siméon M. 2010. New efficient recombinant expression system to engineer *Candida antarctica* lipase B. *Appl Environ Microbiol.* 76, 2684-2687.

Enderlin CS, Ogrydziak DM. 1994. Cloning, nucleotide sequence and functions of XPR6, which codes for a dibasic processing endoprotease from the yeast *Yarrowia lipolytica*. *Yeast*. 10, 67-79.

Fabre E, Nicaud JM, Lopez MC, Gaillardin C. 1991. Role of the proregion in the production and secretion of the *Yarrowia lipolytica* alkaline extracellular protease. *J Biol Chem*. 266, 3782-3790.

Fabre E, TharaudC, Gaillardin C. 1992. Intracellular transit of a yeast protease is rescued by trans-complementation with its prodomain. *J. Biol. Chem*. 267, 15049–15055.

Faraji F, Modif MR, Babaeipour V, Divsalar A, Alboghasemi DS. The structural characterization of recombinant human granulocyte colony-stimulating factor 2010. International *J. Environ. Sci. Dev*. 1, 1.

Fickers P, Fudalej F, Nicaud JM, Destain J, Thonart P. 2005. Selection of new over-producing derivatives for the improvement of extracellular lipase production by the non-conventional yeast *Yarrowia lipolytica*. *J Biotechnol*. 115, 379-386.

Fickers P, Le Dall MT, Gaillardin C, Thonart P, Nicaud JM. 2003. New disruption cassettes for rapid gene disruption and marker rescue in the yeast *Yarrowia lipolytica*. *J Microbiol Methods*. 55, 727-737.

Fickers P, Marty A, Nicaud JM. 2011. The lipases from *Yarrowia lipolytica*: genetics, production, regulation, biochemical characterization and biotechnological applications. *Biotechnol Adv*. 29, 632-644.

Fournier P, Abbas A, Chasles M, Kudla B, Ogrydziak DM, Yaver D, Xuan JW, Peito A, Ribet AM, Feynerol C. 1993. Colocalization of centromeric and replicative functions on autonomously replicating sequences isolated from the yeast *Yarrowia lipolytica*. *Proc Natl Acad Sci U S A.* 90, 4912-4916.

Fukuda R. 2013. Metabolism of hydrophobic carbon sources and regulation of it in n-alkane-assimilating yeast *Yarrowia lipolytica*. *Biosci Biotechnol Biochem*. 77, 1149-1154.

Fuller RS, Brake AJ, Thorner J. 1989. Intracellular targeting and structural conservation of a prohormone-processing endoprotease. *Science*. 246, 482-486.

Galvagno MA, Iannone LJ, Bianchi J, Kronberg F, Rost E, Carstens MR, Cerrutti P. 2011. 2011. Optimization of biomass production of a mutant of *Yarrowia lipolytica* with an increased lipase activity using raw glycerol. *Rev Argent Microbiol.* 43, 218-225.

Gasmi N, Ayed A, Nicaud JM, Kallel H. 2011. Design of an efficient medium for heterologous protein production in *Yarrowia lipolytica*: case of human interferon alpha 2b. *Microb Cell Fact*. 10, 38.

Gasmi N, Fudalej F, Kallel H, Nicaud JM. 2011. A molecular approach to optimize hIFN alpha 2b expression and secretion in *Yarrowia lipolytica*. *Appl Microbiol Biotechnol*, 89, 109-119.

Gasmi N, Lassoued R, Ayed A, Tréton B, Chevret D, Nicaud JM, Kallel H. 2012. Production and characterization of human granulocyte-macrophage colony-stimulating factor (hGM-CSF) expressed in the oleaginous yeast *Yarrowia lipolytica*. *Appl Microbiol Biotechnol*. 96, 89-101.

Gasser B, Saloheimo M, Rinas U, Dragosits M, Rodríguez-Carmona E, Baumann K, Giuliani M, Parrilli E, Branduardi P, Lang C, Porro D, Ferrer P, Tutino ML, Mattanovich D, Villaverde A. 2008. Protein folding and conformational stress in

microbial cells producing recombinant proteins: a host comparative overview.*Microb Cell* Fact. 7, 11. doi: 10.1186/1475-2859-7-11.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*. 31, 3784-3788.

Geigert J, Neidleman SL, Dalietos DJ. 1983a. Novel haloperoxidase substrates. Alkynes and cyclopropanes. *J Biol Chem*. 258, 2273-2277.

Geigert J, Neidleman SL, Dalietos DJ, Dewitt SK. 1983b. Haloperoxidases: Enzymatic Synthesis of alpha,beta-Halohydrins from Gaseous Alkenes. *Appl Environ Microbiol*. 45, 366-374.

Geigert J, Neidleman SL, Dalietos DJ, Dewitt SK. 1983c. Novel haloperoxidase reaction: synthesis of dihalogenated products. *Appl Environ Microbiol*. 45, 1575-1581.

Gellissen G, Kunze G, Gaillardin C, Cregg JM, Berardi E, Veenhuis M, van der Klei I. 2005. New yeast expression platforms based on methylotrophic *Hansenula polymorpha* and *Pichia pastoris* and on dimorphic *Arxula adeninivorans* and *Yarrowia lipolytica* - a comparison. *EMS Yeast Res.* 5, 1079-1096.

Ghezelbash GR, Nahvi I, Emamzadeh R. 2014. Improvement of Erythrose Reductase Activity, Deletion of By-products and Statistical Media Optimization for Enhanced Erythritol Production from *Yarrowia lipolytica* Mutant 49. *Curr Microbiol.* 69, 149-157.

Gonçalves FA, Colen G, Takahashi JA. *Yarrowia lipolytica* and its multiple applications in the biotechnological industry. *Scientific World Journal*. 2014, 476207. doi: 10.1155/2014/476207.

Gong B, Burnina I, Lynaugh H, Li H. 2014. O-linked glycosylation analysis of recombinant human granulocyte colony-stimulating factor produced in glycoengineered *Pichia pastoris* by liquid chromatography and mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci*. 945-946, 135-140.

Green MT, Dawson JH, Gray HB. 2004. Oxoiron(IV) in chloroperoxidase compound II is basic: implications for P450 chemistry. *Science,* 304, 1653-1656.

Groenewald M, Boekhout T, Neuveglise C, Gaillardin C, van Dijck PW, Wyss M. 2014. *Yarrowia lipolytica*: Safety assessment of an oleaginous yeast with a great industrial potential. *Crit*. *Rev. Microbiol*. 40, 187–206.

Groves JT. 2006. High-valent iron in chemical and biological oxidations. *J Inorg Biochem*. 100, 434-447.

Guerfal M, Claes K, Knittelfelder O, De Rycke R, Kohlwein SD, Callewaert N. 2013. Enhanced membrane protein expression by engineering increased intracellular membrane production. *Microb Cell Fact*. 12, 122. doi: 10.1186/1475-2859-12-122.

Gustafsson C, Govindarajan S, Minshull J. 2004. Codon bias and heterologous protein expression. *Trends Biotechnol*. 22, 346-353.

Hager LP, Lakner FJ, Basavapathruni A. 1998. Chiral synthons via chloroperoxidase catalysis. *J. Mol. Cat. B: Enzymatic* 5, 95-101.

Hamsa PV, Chattoo BB. 1994. Cloning and growth-regulated expression of the gene encoding the hepatitis B virus middle surface antigen in *Yarrowia lipolytica*. *Gene*. 143, 165-170.

Hamsa PV, Kachroo P, Chattoo BB. 1998. Production and secretion of biologically active human epidermal growth factor in *Yarrowia lipolytica*. *Curr Genet*. 33, 231-237.

Hanahan D. 1985. Techniques for transformation of *Escherichia coli.* In DNA cloning: a practical approach. 1 (Glover, DM. ed) *Oxford University Press*. 109-135.

Hara M, Yuasa S, Shimoji K, Onizuka T, Hayashiji N, Ohno Y, Arai T, Hattori F, Kaneda R, Kimura K, Makino S, Sano M, Fukuda K. 2011. G-CSF influences mouse skeletal muscle development and regeneration by stimulating myoblast proliferation. *J Exp Med.* 208, 715-727.

Hasemann CA, Kurumbail RG, Boddupalli SS, Peterson JA, Deisenhofer J. 1995. Structure and function of cytochromes P450: a comparative analysis of three crystal structures. *Structure*. 3, 41-62.

Hatfield GW, Roth DA. 2007. Optimizing scaleup yield for protein production: Computationally Optimized DNA Assembly (CODA) and Translation Engineering. *Biotechnol Annu Rev*. 13, 27-42.

Hofrichter M, Ullrich R. 2006. Heme-thiolate haloperoxidases: versatile biocatalysts with biotechnological and environmental significance. *Appl Microbiol Biotechnol*. 71, 276-288.

Hollenberg PF, Hager LP. 1973. The P-450 nature of the carbon monoxide complex of ferrous chloroperoxidase. *J Biol Chem.* 248, 2630-2633.

Hu S, Dordick JS. 2002. Highly enantioselective oxidation of cis cyclopropylmethanols to corresponding aldehydes catalyzed by chloroperoxidase. *J Org Chem*. 67, 314-317.

Hu S, Hager LP. 1998. Unusual propargylic oxidations catalyzed by chloroperoxidase. *Biochem Biophys Res Commun*. 253, 544-546.

Hu S, Hager LP. 1999. Highly enantioselective propargylic hydroxylations catalyzed by chloroperoxidase. *J. Am. Chem. Soc*. 121, 872-873.

Idiris A, Tohda H, Kumagai H, Takegawa K. 2010. Engineering of protein secretion in yeast: strategies and impact on protein production. *Appl Microbiol Biotechnol*. 86, 403-417.

Jacobs PP, Inan M, Festjens N, Haustraete J, Van Hecke A, Contreras R, Meagher MM, Callewaert N. 2010. Fed-batch fermentation of GM-CSF-producing glycoengineered *Pichia pastoris* under controlled specific growth rate. *Microb Cell Fact*. 9, 93. doi: 10.1186/1475-2859-9-93.

Jacques N, Casaregola S. 2008. Safety assessment of dairy microorganisms: the hemiascomycetous yeasts. *Int J Food Microbiol*. 126, 321-326.

Jalving R1, van de Vondervoort PJ, Visser J, Schaap PJ. 2000. Characterization of the kexin-like maturase of *Aspergillus niger*. *Appl Environ Microbio*l. 66, 363-368.

Jeong KJ, Lee SY. 2001. Secretory production of human granulocyte colony-stimulating factor in *Escherichia coli*. *Protein Expr Purif*. 23, 311-318.

Jolivalt C, Madzak C, Brault A, Caminade E, Malosse C, Mougin C. 2005. Expression of laccase IIIb from the white-rot fungus Trametes versicolor in the yeast *Yarrowia lipolytica* for environmental applications. *Appl Microbiol Biotechnol.* 66, 450-456.

Juretzek T, Le Dall M, Mauersberger S, Gaillardin C, Barth G, Nicaud J. 2001. Vectors for gene expression and amplification in the yeast *Yarrowia lipolytica*. *Yeast*. 18, 97-113.

Kedderis GL, Koop DR, Hollenberg PF. 1980. N-Demethylation reactions catalyzed by chloroperoxidase. *J Biol Chem.* 255, 10174-10182.

Kedderis GL, Rickert DE, Pandey RN, Hollenberg PF. 1986. 18O studies of the peroxidase-catalyzed oxidation of N-methylcarbazole. Mechanisms of carbinolamine and carboxaldehyde formation. *J Biol Chem*. 261, 15910-15914.

Kenigsberg P, Fang GH, Hager LP. 1987. Post-translational modifications of chloroperoxidase from *Caldariomyces fumago*. *Arch Biochem Biophys*. 254, 409-415.

Kerscher S, Durstewitz G, Casaregola S, Gaillardin C, Brandt U. 2001. The complete mitochondrial genome of *Yarrowia lipolytica*. *Comp Funct Genomics*. 2, 80-90.

Khasa YP, Khushoo A, Tapryal S, Mukherjee KJ. 2011. Optimization of human granulocyte macrophage-colony stimulating factor (hGM-CSF) expression using asparaginase and xylanase gene's signal sequences in *Escherichia coli*. *Appl Biochem Biotechnol.* 165, 523-537.

Kim CK, Choi JH, Lee SB, Lee SM, Oh JW. 2014. Expression and purification of recombinant human granulocyte colony-stimulating factor in fed-batch culture of *Escherichia coli*. *Appl Biochem Biotechnol.* 172, 2425-2435.

Kim JW, Park TJ, Ryu DD, Kim JY. 2000. High cell density culture of *Yarrowia lipolytica* using a one-step feeding process. *Biotechnol Prog*. 16, 657-660.

Kim JY, Kim EJ, Lopez-Maury L, Bähler J, Roe JH. 2014. A metabolic strategy to enhance long-term survival by Phx1 through stationary phase-specific pyruvate decarboxylases in fission yeast. *Aging (Albany NY)*. 6, 587-601.

Kost TA, Condreay JP, Jarvis DL. 2005. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat Biotechnol*. 23, 567-575.

Kröger N, Zander AR. 2002 .Dose and schedule effect of G-GSF for stem cell mobilization in healthy donors for allogeneic transplantation. *Leuk Lymphoma*. 43, 1391-1394.

La Rotta CE, Bon EP. 2002. 4-chlorophenol degradation by chloroperoxidase from *Caldariomyces fumago*: formation of insoluble products. *Appl Biochem Biotechnol.* 98-100, 191-203.

Labuschagne M, Albertyn J. 2007. Cloning of an epoxide hydrolase-encoding gene from *Rhodotorula mucilaginosa* and functional expression in *Yarrowia lipolytica*. *Yeast.* 24, 69-78.

Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 227, 680-685.

Lakner FJ, Basavapathruni, A., and Hager, L.P. (1998) "Chiral Synthons Via Chloroperoxidase Catalysis. *J. Mol. Catalysis B:Enzymic*. 5, 1-4.

Lakner FJ, Hager LP. 1996. Chloroperoxidase as Enantioselective Epoxidation Catalyst: An Efficient Synthesis of (R)-(-)-Mevalonolactone. *J Org Chem*. 61, 3923-3925.

Laloi, M., Mac Carthy, J., Morandi, O., Gysler, C., Bucheli, P., 2002. Molecular and biochemical characterisation of two aspartic proteinases TcAP1 and TcAP2 from *Theobroma cacao* seeds*. Planta* 215, 754–762.

Lanza AM, Curran KA, Rey LG, Alper HS. 2014. A condition-specific codon optimization approach for improved heterologous gene expression in *Saccharomyces cerevisiae*. *BMC Syst Biol.* 8:33. doi: 10.1186/1752-0509-8-33.

Lasnik M.A, Porekar VG, Stalc A. 2001. Human granulocyte colony stimulating factor (hG-CSF) expressed by methylotrophic yeast *Pichia pastoris*. *Pflugers Arch*.; 442, 184-186.

Le Dall MT, Nicaud JM, Gaillardin C. 1994. Multiple-copy integration in the yeast *Yarrowia lipolytica. Curr Genet*. 26, 38-44.

Li C, Wang L, Jiang Y, Hu M, Li S, Zhai Q. 2011. Activity and stability of chloroperoxidase in the presence of small quantities of polysaccharides: a catalytically favourable conformation was induced. *Appl Biochem Biotechnol*. 165, 1691-1707.

Li H, d 'Anjou M. 2009. Pharmacological significance of glycosylation in therapeutic proteins. *Current Opinion in Biotechnology*. 20, 678–684.

Li JW, Liu Y, Li K, Song D, Ye QP. 2010. Cloning of a fusion gene containing hGM-CSF gene and gene encoding L protein of HBV and expression in *Pichia Pastoris*.*Hepatogastroenterology*. 57, 578-582.

Li PC, Zhang XG, Zhou L, Zeng Y. 2003. Gene optimization is necessary to express HPV type 6 L1 protein in the methylotrophic yeast *Pichia pastoris. Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi. 17,* 310-314*.*

Li Z, Moy A, Gomez SR, Franz AH, Lin-Cereghino J, Lin-Cereghino GP. 2010. An improved method for enhanced production and biological activity of human secretory leukocyte protease inhibitor (SLPI) in *Pichia pastoris. Biochem Biophys Res Commun.* 402, 519-524*.*

Lieschke GJ, Grail D, Hodgson G, Metcalf D, Stanley E, Cheers C, Fowler K J, Basu S, Zhan YF, Dunn AR. 1994. Mice lacking granulocyte colony-stimulating factor have chronic neutropenia, granulocyte and macrophage progenitor cell deficiency, and impaired neutrophil mobilization. *Blood*. 84, 1737-1746.

Lu CZ, Xiao BG. 2006. G-CSF and neuroprotection: a therapeutic perspective in cerebral ischaemia. *Biochem Soc Trans*. 34, 1327-1333.

Lu D, Yang C, Liu Z. 2012. How hydrophobicity and the glycosylation site of glycans affect protein folding and stability: a molecular dynamics simulation. *J Phys Chem B*. 116, 390-400.

Lü H, Zhao WM, Zheng Y, Wang H, Qi M, Yu XP. 2005. Analysis of synonymous codon usage bias in Chlamydia. *Acta Biochim Biophys Sin* (Shanghai). 37, 1-10.

Lu HS, Clogston CL, Narhi LO, Merewether LA, Pearl WR, Boone TC. 1992. Folding and oxidation of recombinant human granulocyte colony stimulating factor produced in *Escherichia coli*. *J Biol Chem*. 267, 8770-8777.

Ludwig G, Frick J. 1987. Ejaculate analysis (spermiogram). In Ludwig, G. and Frick, J. (eds) *Spermatology*. *Springer-Verlag, Berlin, Heidelberg*, 30–32.

Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. 2005. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*. 22, 249-270.

Madzak C, Beckerich JM. 2013. Heterologous Protein Expression and secretion in *Yarrowia Microbology Monographs.* 25, 1-17.

Madzak C, Gaillardin C, Beckerich JM. 2004. Heterologous protein expression and secretion in the non-conventional yeast *Yarrowia lipolytica*: a review. *J Biotechnol* 109, 63–81.

Madzak C, Otterbein L, Chamkha M, Moukha S, Asther M, Gaillardin C, Beckerich JM. 2005. Heterologous production of a laccase from the basidiomycete *Pycnoporus cinnabarinus* in the dimorphic yeast *Yarrowia lipolytica. FEMS Yeast Res*. 5, 635-646.

Madzak C, Treton B, Blanchin-Roland S. 2000. Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast *Yarrowia lipolytica*. *J Mol Microbiol Biotechnol*. 2, 207–216.

Maharajh D, Lalloo R, Görgens J. 2008. Effect of an exponential feeding regime on the production of Rhodotorula araucariae epoxide hydrolase in *Yarrowia lipolytica*. *Lett Appl Microbiol*. 47, 520-525.

Maharajh D, Roth R, Lalloo R, Simpson C, Mitra R, Görgens J, Ramchuran S. 2008. Multi-copy expression and fed-batch production of *Rhodotorula araucariae* epoxide hydrolase in *Yarrowia lipolytica*. *Appl Microbiol Biotechnol*. 79, 235-244.

Manoj KM, Hager LP. 2006. A colorimetric method for detection and quantification of chlorinating activity of hemeperoxidases. *Anal Biochem*. 348, 84-86.

Manoj, K. M., Lakner, F. J., & Hager, L. P. (2000). Epoxidation of indene by chloroperoxidase. *Journal of Molecular Catalysis B: Enzymatic*, *9*, 107-111.

Masurekar PS. 2008. Nutritional and engineering aspects of microbial process development. *Prog Drug Res*. 65, 293-328.

Matsuoka M, Matsubara M, Daidoh H, Imanaka T, Uchida K, Aiba S. 1993. Analysis of regions essential for the function of chromosomal replicator sequences from *Yarrowia lipolytica*. *Mol Gen Genet*. 237, 327-333.

Mattanovich D, Branduardi P, Dato L, Gasser B, Sauer M, Porro D. 2012. Recombinant protein production in yeasts. *Methods Mol Biol*. 824, 329-358.

Mauersberger S, Novikova, Ludmila A, Shkumatov, Vladimir M. 2013. Cytochrome P450 Expression in *Yarrowia lipolytica* and its use in steroid Biotransformations. *Microbiology Monographs.* 25, 171-226.

Max B, Salgado JM, Rodríguez N, Cortés S, Converti A, Domínguez JM. 2010. Biotechnological production of citric acid. *Braz J Microbiol*. 41, 862-875.

Metcalf D. 1985. The granulocyte-macrophage colony-stimulating factors. *Science* 229, 16–22.

Miller VP, Tschirret-Guth RA, Ortiz de Montellano PR. 1995. Chloroperoxidase-catalyzed benzylic hydroxylation*. Arch Biochem Biophys.* 319, 333-340.

Miyajima A, Otsu K, Schreurs J, Bond MW, Abrams J S, Arai K. 1986. Expression of murine and human granulocyte-macrophage colony-stimulating factors in *Saccharomyces cerevisiae*: mutagenesis of the potential glycosylation sites. *EMBO J.* 5, 1193-1197.

Moon HY, Van TL, Cheon SA, Choo J, Kim JY, Kang HA. 2013. Cell-surface expression of *Aspergillus saitoi*-derived functional α-1,2-mannosidase on *Yarrowia lipolytica* for glycan remodeling. *J Microbiol*. 51, 506-514.

Morgan JA, Lu Z, Clark DS. 2002. Toward the development of a biocatalytic system for oxidation of p-xylene to tetraphtalic acid: oxidation of 1,4-benzenedimethanol. *J Mol Catal, B* 18,147‑154.

Morris DR, Hager LP. 1966. Chloroperoxidase. I. Isolation and properties of the crystalline glycoprotein. *J Biol Chem*. 241, 1763-1768.

Müller S, Sandal T, Kamp-Hansen P, Dalbøge H. 1998. Comparison of expression systems in the yeasts *Saccharomyces cerevisiae*, *Hansenula polymorpha*, *Klyveromyces lactis*, *Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Cloning of two novel promoters from *Yarrowia lipolytica*. *Yeast*. 14, 1267-1283.

Nakajima R, Yamazaki I, Griffin BW. 1985. Spectra of chloroperoxidase compounds II and III. *Biochem Biophys Res Commun*. 128, 1-6.

Nasser MW, Pooja V, Abdin MZ, Jain SK. 2003. Evaluation of yeast as an Expression system. *Indian Journal of Biotechnol*. 2, 447-493.

Nevoigt E. 2008. Progress in metabolic engineering of *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev*. 72, 379-412.

Nicaud JM, Fournier P, La Bonnardière C, Chasles M, Gaillardin C. 1991. Use of ars18 based vectors to increase protein production in *Yarrowia lipolytica*. *J Biotechnol*. 19, 259-270.

Nicaud JM, Madzak C, van den Broek P, Gysler C, Duboc P, Niederberger P, Gaillardin C. 2002. Protein expression and secretion in the yeast *Yarrowia lipolytica*. *FEMS Yeast Res*. 2, 371-379.

Nicaud. JM. 2012. *Yarrowia lipolytica*. *Yeast*; 29, 409–418.

Nissen C. 1994. Glycosylation of recombinant human granulocyte colony stimulating factor: implications for stability and potency. *Eur J Cancer.* 3, S12-4.

Nohynek GJ, Plard JP, Wells MY, Zerial A, Roquet F. 1997. Comparison of the potency of glycosylated and nonglycosylated recombinant human granulocyte colony-stimulating factors in neutropenic and nonneutropenic CD rats. *Cancer Chemother Pharmacol*. 39, 259-266.

Nthangeni MB, Urban P, Pompon D, Smit MS, Nicaud JM. 2004. The use of *Yarrowia lipolytica* for the expression of human cytochrome P450 CYP1A1. *Yeast*. 21, 583-592.

Niu J, Yu G. 2004. Molecular structural characteristics governing biocatalytic chlorination of PAHs by chloroperoxidase from *Caldariomyces fumago*. *SAR QSAR Environ Res.* 15, 159-167.

Nuell MJ, Fang GH, Axley MJ, Kenigsberg P, Hager LP. 1988. Isolation and nucleotide sequence of the chloroperoxidase gene from *Caldariomyces fumago*. *J Bacteriol*. 170, 1007-1011.

Ogrydziak DM. 1993. Yeast extracellular proteases. *Crit Rev Biotechnol*.13, 1-55.

Ogrydziak DM, Mortimer RK. 1977. Genetics of Extracellular Protease Production in *SACCHAROMYCOPSIS LIPOLYTICA*. *Genetics.* 87, 621-632.

Ogrydziak DM, Nicaud JM. 2012. Characterization of *Yarrowia lipolytica* XPR2 multi-copy strains over-producing alkaline extracellular protease—a system for rapidly increasing secretory pathway cargo loads. *FEMS Yeast Res* 12, 938–948.

Oh-eda M, Hasegawa M, Hattori K, Kuboniwa H, Kojima T, Orita T, Tomonou K, Yamazaki T, Ochi N. 1990. O-linked sugar chain of human granulocyte colony-stimulating factor protects it against polymerization and denaturation allowing it to retain its biological activity. *J Biol Chem*. 265, 11432-11435.

Olaniran AO, Igbinosa EO. 2011 Chlorophenols and other related derivatives of environmental concern: properties, distribution and microbial degradation processes. *Chemosphere.* 83, 1297-1306.

Osborne RL, Raner GM, Hager LP, Dawson JH. 2006. C. fumago chloroperoxidase is also a dehaloperoxidase: oxidative dehalogenation of halophenols. *J Am Chem Soc*. 128, 1036-1037.

Outchkourov SN, Stiekema WJ, Jongsma MA. 2002. Optimization of the expression of equistatin in *Pichia pastoris*. *Protein Expr Purif*. 24, 18–24.

Palande AS, Kulkarni SV, León-Ramirez C, Campos-Góngora E, Ruiz-Herrera J, Deshpande MV. 2014. Dimorphism and hydrocarbon metabolism in *Yarrowia lipolytica* var. indica. *Arch Microbiol.* [Epub ahead of print]

Park JN, Song Y, Cheon SA, Kwon O, Oh DB, Jigami Y, Kim JY, Kang HA. 2011. Essential role of YlMPO1, a novel *Yarrowia lipolytica* homologue of *Saccharomyces cerevisiae* MNN4, in mannosylphosphorylation of N- and O-linked glycans. *Appl Environ Microbiol.* 77, 1187-1195.

Pedrazzoli P, Gibelli N, Pavesi L, Preti P, Piolini M, Bertolini F, Robustelli della Cuna G. 1996. Effects of glycosylated and non-glycosylated G-CSFs, alone and in combination with other cytokines, on the growth of human progenitor cells. *Anticancer Res*. 16, 1781-1785.

Pérez-Campo FM, Domínguez A. 2001. Factors affecting the morphogenetic switch in *Yarrowia lipolytica*. *Curr Microbiol*. 43, 429-433.

Petty KJ. 1996. "Metal-Chelate Affinity Chromatography. *Current Protocols in Protein Science*. 9-4.

Pickard MA, Hashimoto 1982. A. Isoenzymes of chloroperoxidase from *Caldariomyces fumago Can. J. of Microbiol*. 28, 1382-1388.

Pickard MA, Tenshuk A, Kadima, Robert D C. 1991. Chloroperoxidase, a peroxidase with potential. *Journal of Industrial Microbiology*. 7, 235-242.

Pignède G, Wang H, Fudalej F, Gaillardin C, Seman M, Nicaud JM. 2000. Characterization of an extracellular lipase encoded by LIP2 in *Yarrowia lipolytica*. *J Bacteriol*. 182, 2802-2810.

Pignède G, Wang HJ, Fudalej F, Seman M, Gaillardin C, Nicaud JM. 2000. Autocloning and amplification of LIP2 in *Yarrowia lipolytica*. *Appl Environ Microbiol*. 66, 3283-3289.

Porro D, Sauer M, Branduardi P, Mattanovich D. 2005. Recombinant protein production in yeasts. *Mol Biotechnol*. 31, 245-259.

Pozzuolo S, Breme U, Salis B, Taylor G, Tonon G, Orsini G. 2008. Efficient bacterial expression of fusion proteins and their selective processing by a recombinant Kex-1 protease. *Protein Expr Purif*. 59, 334-341.

Prokop Z, Oplustil F, DeFrank J, Damborský J. 2006. Enzymes fight chemical weapons. Quartley E, Alexandrov A, Mikucki M, Buckner F S, Hol W G, DeTitta G T, Phizicky E M, Grayhack EJ. 2009. Heterologous expression of L. major proteins in *Saccharomyces cerevisiae*: a test of solubility, purity, and gene recoding. *J Struct Funct Genomics*.10, 233-347.

Punt PJ, Drint-Kuijvenhoven A, Lokman BC, Spencer JA, Jeenes D, Archer DA, van den Hondel CA.2003. The role of the *Aspergillus niger* furin-type protease gene in processing of fungal proproteins and fusion proteins. Evidence for alternative processing of recombinant (fusion) proteins. *J Biotechnol*. 106, 23-32.Querol S, Cancelas JA, Amat L, Capmany G, Garcia J. 1999. Effect of glycosylation of recombinant human granulocytic colony-stimulating factor on expansion cultures of umbilical cord blood CD34+ cells. *Haematologica*. 84, 493-498.

Ramagoma F. 2011. Development of *Yarrowia lipolytica* for enhanced production of heterologous proteins. PhD thesis, University of the Free State.

Rao A, Pimprikar P, Bendigiri C, Kumar AR & Zinjarde S. 2011. Cloning and expression of a tyrosinase from *Aspergillus oryzae* in *Yarrowia lipolytica*: application in LDOPA biotransformation. *Appl Microbiol Biotechnol*. 92, 951–959.

Robert D, Carmichael A J, Michael A P. 1986. Semicontinuous and Continuous Production of Chloroperoxidase by *Caldariomyces fumago* Immobilized in k-Carrageenan. *Appl Environ Microbiol*. 51, 276–280.

Robinson GK, Jackman SA, Stratford J. 2000. Halocompounds. In: Rehm H-J, Reed G

Roepcke CB, Muench SB, Schulze H, Bachmann TT, Schmid RD, Hauer B. 2010. Analysis of phosphorothionate pesticides using a chloroperoxidase pretreatment and acetylcholinesterase biosensor detection. *J Agric Food Chem.* 58, 8748-8756.

Roth R, Moodley V, van Zyl P. 2009. Heterologous expression and optimized production of an *Aspergillus aculeatus* endo-1,4-beta-mannanase in *Yarrowia lipolytica*. *Mol Biotechnol*. 43, 112-20.

Rywińska A, Musiał I, Rymowicz W, Zarowska B, Boruczkowski T. 2012. Effect of agitation and aeration on the citric acid production by *Yarrowia lipolytica* grown on glycerol*. Prep Biochem Biotechnol.* 42, 279-291.

Sabaty M, Grosse S, Adryanczyk G, Boiry S, Biaso F, Arnoux P, Pignol D. 2013. Detrimental effect of the 6 His C-terminal tag on YedY enzymatic activity and influence of the TAT signal sequence on YedY synthesis. *BMC Biochem.* 14, 28.

Sabirova JS, Haddouche R, Van Bogaert I, Mulaa F, Verstraete W, Timmis K, Schmidt-Dannert C, Nicaud J, Soetaert W. 2011. The 'LipoYeasts' project: using the oleaginous yeast *Yarrowia lipolytica* in combination with specific bacterial genes for the bioconversion of lipids, fats and oils into high-value products. *Microb Biotechnol*. 4, 47-54.

Sae ASW, Cunningham BA. 1979. Isolation and properties of chloroperoxidase isoenzymes. *Phytochemistry* 18, 1785–1787.

Sambrook J, Fitsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold SpringHarbor Laboratory, *Cold Spring Harbor*, New York

Sanfilippo C, D'Antona N, Nicolosi G. 2004. Chloroperoxidase from *Caldariomyces fumago* is active in the presence of an ionic liquid as co-solvent. *Biotechnol Lett.* 26, 1815–1819.

Schmid-Berger N, Schmid B, Barth G. 1994. Ylt1, a highly repetitive retrotransposon in the genome of the dimorphic fungus *Yarrowia lipolytica*. *J Bacteriol.* 176, 2477-2482.

Schneider A, Krüger C, Steigleder T, Weber D, Pitzer C, Laage R, Aronowski J, Maurer MH, Gassler N, Mier W, Hasselblatt M, Kollmar R, Schwab S, Sommer C, Bach A, Kuhn HG, Schäbitz WR. The hematopoietic factor G-CSF is a neuronal ligand that counteracts programmed cell death and drives neurogenesis. *J Clin Invest.* 115, 2083-2098.

Seelbach K, Kragl U. 1997. Nanofiltration membranes for cofactor retention in continuous enzymatic synthesis. *Enzyme Microb Technol*. 20, 389–392.

Shilov IV, Seymour SL, Patel AA, Loboda A, Tang WH, Keating SP, Hunter CL, Nuwaysir LM, Schaeffer DA. (2007). The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Molecular & Cellular Proteomics*. 6, 1638-1655.

Shiningavamwe A, Obiero G, Albertyn J, Nicaud JM, Smit M. 2006. Heterologous expression of the benzoate para-hydroxylase encoding gene (CYP53B1) from *Rhodotorula minuta* by *Yarrowia lipolytica*. *Appl Microbiol Biotechnol*. 72, 323-329.

Sinclair G, Choy FY. 2002. Synonymous codon usage bias and the expression of human glucocerebrosidase in the methylotrophic yeast, *Pichia pastoris*. *Protein Expr Purif*. 2002 26, 96-105.

Skerker JM, Lucks JB, Arkin, AP. (2009). Evolution, ecology and the engineered organism: lessons for synthetic biology. *Genome biology*, *10*, 114.

Spencer JF, Ragout de Spencer AL, Laluce C. 2002. Non-conventional yeasts. *Appl Microbiol Biotechnol.* 58, 147-156.

Spencer JF, Spencer DM. 1983. Genetic improvement of industrial yeasts. *Annu Rev Microbiol*. 37, 121-142.

Srinivasa Babu K, Muthukumaran T, Antony A, Prem Singh Samuel SD, Balamurali M, Murugan V, Meenakshisundaram S. 2009. Single step intein-mediated purification of hGMCSF expressed in salt-inducible *Escherichia coli*. *Biotechnol Lett.* 31, 659-664.

Srinivasa Babu K, P AO ulicherla K K, Antony A , Meenakshisundaram S. 2014. Cloning and Expression of Recombinant Human GMCSF from *Pichia pastoris* GS115-A Progressive Strategy for Economic Production. *Am J Ther*. 2014 [Epub ahead of print].

Srinivasa Babu K, Antony A, Muthukumaran T, Meenakshisundaram S. 2008. Construction of intein-mediated hGMCSF expression vector and its purification in *Pichia pastoris. Protei Expr Purif.* 57, 201-205.

Sun W, Kadima TA, Pickard MA, Dunford HB. 1994. Catalase activity of chloroperoxidase and its interaction with peroxidase activity. *Biochem Cell Biol*. 72, 321-231.

Sundaramoorthy M, Terner J, Poulos TL. 1998. Stereochemistry of the chloroperoxidase active site: crystallographic and molecular-modeling studies. *Chem Biol*. 5, 461-473.

Sundaramoorthy M, Terner J, Poulos TL. 1995. The crystal structure of chloroperoxidase: a heme peroxidase--cytochrome P450 functional hybrid. *Structure.* 3, 1367-1377.

Swennen D, Paul MF, Vernis L, Beckerich JM, Fournier A, Gaillardin C. 2002. Secretion of active anti-Ras single-chain Fv antibody by the yeasts *Yarrowia lipolytica* and *Kluyveromyces lactis*. *Microbiology.* 148, 41-50.

Swennen D, Beckerich JM. 2007. *Yarrowia lipolytica* vesicle-mediated protein transport pathways. *BMC Evol Biol.* 7, 219.

Takayama H, Miyake Y, Nouso K., Ikeda F, Shiraha H, Takaki A, Kobashi H, Yamamoto K. 2011. Serum levels of plateletderived growth factor-BB and

vascular endothelial growth factor as prognostic factors for patients with fulminant hepatic failure. *J Gastroenterol Hepatol*. 26, 116-121.

Tharaud C, Ribet AM, Costes C, Gaillardin C. 1992. Secretion of human blood coagulation factor XIIIa by the yeast *Yarrowia lipolytica*. *Gene*. 121, 111-119.

Theerachat M, Emond S, Cambon E, Bordes F, Marty A, Nicaud JM, Chulalaksananukul W, Guieysse D, Remaud-Siméon M, Morel S. 2012. Engineering and production of laccase from *Trametes versicolor* in the yeast *Yarrowia lipolytica*. *Bioresour Technol.* 125, 267-274.

Thevenieau F, Le Dall MT, Nthangeni B, Mauersberger S, Marchal R, Nicaud JM. 2007. Characterization of *Yarrowia lipolytica* mutants affected in hydrophobic substrate utilization. *Fungal Genet Biol*. 44, 531-542.

Thomas JA, Morris DR, Hager LP. 1970. Chloroperoxidase. 8. Formation of peroxide and halide complexes and their relation to the mechanism of the halogenation reaction. *J Biol Chem*. 245, 3135-3142.

Tobe S, Takami, T, Ikeda S, Mitsugi K. 1976. Production and some enzymatic properties of alkaline protease of *Candida lipolytica*. *Agric. Biol. Chem*. 40, 1087–1092.

Tyo KE, Liu Z, Petranovic D, Nielsen J. 2012. Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress. *BMC Biol*. 16 1;10:16. doi: 10.1186/1741-7007-10-16.

Ullrich R, Nüske J, Scheibner K, Spantzel J, Hofrichter M. 2004. Novel haloperoxidase from the agaric basidiomycete *Agrocybe aegerita* oxidizes aryl alcohols and aldehydes. *Appl Environ Microbiol.* 70, 4575-4581.

van der Walt JP, von Arx JA. 1980. The yeast genus *Yarrowia* gen. nov. *Antonie Van Leeuwenhoek*. 46, 517-521.

van de Velde F, Bakker M, van Rantwijk F, Sheldon RA. 2001. Chloroperoxidase-catalyzed enantioselective oxidations in hydrophobic organic media. *Biotechnol Bioeng*. 72, 523-529.

Vázquez-Duhalt R, Ayala M, Márquez-Rocha FJ. 2001. Biocatalytic chlorination of aromatic hydrocarbons by chloroperoxidase of *Caldariomyces fumago*. *Phytochemistry*. 58, 929-33.

Vernis L, Abbas A, Chasles M, Gaillardin CM, Brun C, Huberman JA, Fournier P. 1997. An origin of replication and a centromere are both needed to establish a replicative plasmid in the yeast *Yarrowia lipolytica*. *Mol Cell Biol*. 17, 1995-2004.

Waché Y, Aguedo M, Nicaud JM, Belin JM. 2003. Catabolism of hydroxyacids and biotechnological production of lactones by *Yarrowia lipolytica*. *Appl Microbiol Biotechnol.* 61, 393-404.

Wang C, Wang L, Geng X. 2008. High recovery refolding of rhG-CSF from *Escherichia coli*, using urea gradient size exclusion chromatography. *Biotechnol Prog*. 24, 209-213.

Wang CZ, Liu JF, Geng XD. 2005. Refolding with Simultaneous Purification of recombinant human granulocyte colony-stimulating factor from *Escherichia coli* using strong anion exchange chromatography. *Chinese Chem. Letter*. 16, 389-392.

Wang H, Song HL, Wang Q, Qiu BS. 2013. Expression of glycoproteins bearing complex human-like glycans with galactose terminal in Hansenula polymorpha. *World J Microbiol Biotechnol.* 29, 447-58.

Wang J, Zou M, Huang B, Zhang M, Mao N. 1995. Molecular cloning and expression of human granulocyte-macrophage colony stimulating factor in *Escherichia coli* system. *Chin J Biotechnol*.11, 45-81.

Wannstedt C, Rotella D, Siuda JF. 1990. Chloroperoxidase mediated halogenation of phenols. *Bull Environ Contam Toxicol*. 44, 282-287.

Wilkinson B, Gilbert HF. 2004. Protein disulfide isomerase. *Biochim Biophys Acta*. 1699, 35-44.

Xu S, Höglund M, Hâkansson L., Venge P. 2000. Granulocyte colony-stimulating factor (G-CSF) induces the production of cytokines in vivo. *Br J Haematol*. 108, 848-853.

Yaipakdee P, Robertson LW. 2001. Enzymatic halogenation of flavanones and flavones. *Phytochemistry*. 57, 341-347.

Yang CH, Huang YC, Chen CY, Wen CY. 2010. Heterologous expression of *Thermobifida fusca* thermostable alpha-amylase in *Yarrowia lipolytica* and its application in boiling stable resistant sago starch preparation. *J Ind Microbiol Biotechnol*. 37, 953-960.

Yarrow D. 1972. Four new combinations in yeasts. *Antonie Van Leeuwenhoek*.38, 357-360.

Yi, Conesa A, Punt PJ, Hager LP. 2003. Examining the role of glutamic acid 183 in chloroperoxidase catalysis. *J Biol Chem*. 278, 13855-13859.

Yi X, Mroczko M, Manoj KM, Wang X, Hager LP. 1999. "Replacement of the proximal heme thiolate ligand in chloroperoxidase with a histidine residue." *Proc. Natl.Acad. Sci. USA* 96, 12412-12417.

Yoko-o T, Tsukahara K, Watanabe T, Hata-Sugi N, Yoshimatsu K, Nagasu T, Jigami Y. 2001. *Schizosaccharomyces pombe* och1 (+) encodes alpha-1, 6-mannosyltransferase that is involved in outer chain elongation of N-linked oligosaccharides. *FEBS Lett*. 489:75–80.

Yue L, Chi Z, Wang L, Liu J, Madzak C, Li J, Wang X. 2007. Construction of a new plasmid for surface display on cells of *Yarrowia lipolytica*. *J Microbiol Methods*. 72, 116-123.

Yuzbashev TV, Yuzbasheva EY, Vibornaya TV, Sobolevskaya TI, Laptev IA, Gavrikov AV, Sineoky SP. 2012. Production of recombinant *Rhizopus oryzae* lipase by the yeast *Yarrowia lipolytica* results in increased enzymatic thermostability. *Protein Expr Purif.* 82, 83-89.

Zaret, KS, and Sherman F. 1982. DNA sequence required for efficient transcription termination in yeast. *Cell*. 28, 563-573.

Zhang B, Chen H, Li M, Gu Z, Song Y, Ratledge C, Chen YQ, Zhang H, Chen W. 2013. Genetic engineering of *Yarrowia lipolytica* for enhanced production of trans-10, cis-12 conjugated linoleic acid. *Microb Cell Fact*. 12, 70.

Zhang B, Rong C, Chen H, Song Y, Zhang H, Chen W. 2012. De novo synthesis of trans-10, cis-12 conjugated linoleic acid in oleaginous yeast *Yarrowia lipolytica*. *Microb Cell Fact.* 11, 51.

Zhao X, Huo KK, Li YY. 2000. Synonymous codon usage in *Pichia pastoris*. *Sheng Wu Gong Cheng Xue Bao*.16, 308-311.

Zinjarde S, Apte M, Mohite P, Kumar AR. 2014. *Yarrowia lipolytica* and pollutants: Interactions and applications. 14, 9734-9750.

Zong Q, Osmulski PA, Hager LP. 1995. High-pressure-assisted reconstitution of recombinant chloroperoxidase. *Biochemistry*. 34, 12420-12425.

APPENDIX 1

Biological assays of recombinant hG-CSF expressed in *Yarrowia lipolytica*

Study Report

Confidential in Contract

Single dose effect ofrhG-CSF in comparison to the commercially available rhG-CSF (Neupogen)

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

University of Pretoria Biomedical Research Centre
University of Pretoria
South Africa

On behalf of

CSIR Biosciences Pretoria

Prepared by: Dr T Pulker
Date of issue: February 2011

TABLE OF CONTENTS

Contents

QualityStatement

The following study was undertaken at the research facilities of the University of Pretoria using an in-house developed quality control system. The facility is not ISO17025 or GLP accredited.

_____

Prof V Naidoo
BVMCh, MSc(Vet), PhD
Director: UPBRC

_____

Dr T Pulker
BVSc
Study Site Investigator

PersonsInvolvedintheStudy

Study Sponsor: **Dr NtsaneMoleleki**
Yeast Expression Systems Research Group Leader
CSIR Biosciences
Building 20
MeiringNaude Rd
Brummeria, Pretoria
South Africa
Tel: +27 12 (0) 841 3981
Fax: +27 12 (0) 841 4790/3651
Email: NMoleleki@csir.co.za

Principal Investigator: **Dr TamsynPulker**
University of Pretoria Biomedical Research Centre
Faculty of Veterinary Science
University of Pretoria
Onderstepoort, 0110
South Africa
Tel: +27 (0)12 529 8524
Cell: +27 (0)82 368 0666
Fax: +27 (0)12 529 8321
Email: tamsyn.pulker@up.ac.za

Study Director &Veterinarian:
ProfVinnyNaidoo, BVMCh, MSc, PhD
University of Pretoria Biomedical Research Centre
Faculty of Veterinary Science
University of Pretoria
Onderstepoort, 0110
South Africa
Tel: +27 (0)12 529 8082
Cell: +27 (0)82 563 3271
Fax: +27 (0)12 529 8321
Email: vinny.naidoo@up.ac.za

Study Site: **University of Pretoria Biomedical Research Centre**
Faculty of Veterinary Science
University of Pretoria
Old Soutpan Rd (M35)
Onderstepoort, 0110
South Africa

Analytical Laboratory: **Clinical Pathology Section**
Department of Companion Animal Clinical Studies
Faculty of Veterinary Science
University of Pretoria
(Onderstepoort Veterinary Academic Hospital)
Onderstepoort, South Africa

Study Report Overview

Study title:	Single dose effect of rhG-CSF in comparison to the commercially available rhG-CSF (Neupogen)
Study objectives:	The primary objective wasto determine the single dose effect of the expressed G-CSF produced by the CSIR in comparison to the commercially available recombinant hG-CSF (Neupogen, Roche).
Study design:	Randomised group allocation, each animal having induction of neutropaenia by an intraperitoneal injection of Ifosfamide followed by daily single dose injections of the test and control products for four days. Euthanasia and white cell blood counts on the fourth day, six hours after treatment.
Trial animals:	24 SPF Balb/C male mice of 6 – 8 weeks of age
Test products:	GCSF 1 and GCSF 7, as supplied by the sponsor
Reference Product:	Neupogen, Roche
Dosage:	Group 1 – 16 µg GCSF 1 per mouse in bovine serum albumin and Tris-Cl Group 2 – 16 µg GCSF 7 per mouse in bovine serum albumin and Tris-Cl Group 3 – 16 µg Neupogen per mouse in bovine serum albumin and Tris-Cl Group 4 – 0.1 % Bovine serum albumin and Tris-Cl per mouse All groups received a constant volume of 0.5 ml / mouse IP
Blood samples:	Blood samples collected in heparin on the day of euthanasia by means of cardiocentesis for the determination of white blood cell counts
Statistical analysis:	Kruskal Wallis H Test and Mann-Whitney U Test to ascertain significance between groups.
Results:	No deaths were reported during the study.No significant differenceswere present in the neutrophil counts when the test group was compared to the negative control. In contrast the Neupogenpositive control produced a significant increase in the neutrophil count.
Conclusion:	The CSIR GSCF1 &GSCF7 failed to induce a significant increase in neutrophil count.

1. Preamble

Cancer therapy is a major clinical and pharmacological science, which involves the killing of neoplastic cells through the use of anti-neoplastic chemotherapeutics and/or radiation therapy. For the former, the majority of the drugs selected for use function by killing rapidly dividing cells. While this effect is beneficial in that tumour cells are rapidly dividing, the effect is equally as destructive against endogenous cells that have an equal of higher turnover rate. One such cell systems are the progenitor cells present in the bone marrow.

As a result, neutropaenia is a common complication following the use of many of the available anti-neoplastic agents. Since the circulating neutrophil population represents the body's major immune defence mechanism, oncology patients easily succumb to infectious agents. To prevent the latter, newer treatment regimens rely of the use of protein drugs that stimulate the granulocyte population to increase their rate of neutrophil production. The agent currently used is human granulocyte colony stimulating factor (hG-CSF), a cytokineproduced byrecombinantDNA technology (Neupogen, Roche).

The CSIR has recently developed their own hG-CSF, which they believe may serve as a generic agent to the current commercial product.

2. Study Objective

Determine the effect of a single dose of the expressed G-CSF (two forms) produced by the CSIR in comparison to the commercially available recombinant hG-CSF.

3. Study Designand Study Schedule

The study log of events is listed in table 1 below. (Refer to Appendix 1-1)

Table 1: Study log of Events

Day of Study	Date	Action
-17	05 October 10	Animals induced into study, weighed and ear notched (mouse 1 – 18)
-8	14 October 10	Daily monitoring and weighing (mouse 1 – 18)
-7	15 October 10	Ear notching of mouse 19 - 24 Daily monitoring and weighing (mouse 1 – 24)
-7 to -1	15 - 21 October 10	Acclimatisation period, Daily monitoring and weighing (mouse 1 – 24)
0	22 October 10	Randomisation into groups by weight Induction of neutropaenia, weighing, monitoring
1	23 October 10	Administration of test and reference products, weighing, monitoring
2	24 October 10	Administration of test and reference products, weighing, monitoring
3	25 October 10	Administration of test and reference products, weighing, monitoring
4	26 October 10	Administration of test and reference products, weighing, monitoring Euthanasia six hours post administration, terminal cardiocentesis for blood cell counts and collection of bone marrow (banked)

4. Study Animals

4.1 Animal Description

In total of twenty four inbred, male, SPFBalb/C mice that met the inclusion criteria and not the exclusion criteria, sourced from the UPBRC,were inducted into the study as perTable 2: Description of animals inducted into the study. The animals were uniquely identified by means of ear notches.(Refer to Appendix 1-2)

Table 2: Description of animals inducted into the study

Mouse No.	DOB	Weight (g)	Cage No.	Mouse No.	DOB	Weight (g)	Cage No.
1	25/08/2010	23.2	1	**13**	25/08/2010	25.2	5
2	25/08/2010	21.8	1	**14**	08/09/2010	23.1	4
3	25/08/2010	23.7	1	**15**	08/09/2010	23.9	4
4	25/08/2010	23.3	1	**16**	08/09/2010	22.2	4
5	25/08/2010	22.1	2	**17**	08/09/2010	24.6	4
6	25/08/2010	24.4	2	**18**	08/09/2010	22.7	4
7	25/08/2010	24.0	2	**19**	21/07/2010	21.5	6
8	25/08/2010	24.0	3	**20**	21/07/2010	22.3	6
9	25/08/2010	23.3	3	**21**	21/07/2010	20.6	6
10	25/08/2010	24.2	3	**22**	21/07/2010	24.7	7
11	25/08/2010	24.9	5	**23**	21/07/2010	22.7	7
12	25/08/2010	23.9	5	**24**	21/07/2010	21.6	7

4.1.1 Inclusion criteria

- Healthy mice bred for the study
- Seven days acclimatisation period within the UPBRC

4.1.2 Exclusion criteria

- Animals showing signs of disease prior to dosing
- Animals below the UPBRC expected normal weight pattern (Figure 1)

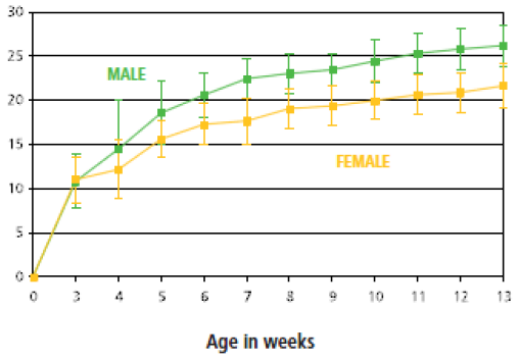

Figure1: Expected growth curve of Balb/C mice (Balb/C Information sheets from www.harlan.com). The y-axis represents the weight measured in grams

4.2 Housing, feed and water

The mice weregroup housed (3 to 5 per group) in Eurostandard type II open top cages (Tecniplast) in a temperature, humidity and particle (HEPA) controlled environment as described in Table 3. The mice received irradiated EPOL mouse pellets and autoclaved potable municipal water *ad libitum*. Autoclaved wood shavings and *Erogrottisteff*hay were provided as bedding material, with autoclaved tissues, egg carton, cardboard and mouse houses provided as enrichment. Cages werechanged weekly. (Refer to Appendix 1-3)

Table 3: Environmental conditions in the animal room

Condition	Range
Lighting	12 hour light-dark cycle, lights on at 06h00 and lights off at 18h00, on an automated timer
Room temperature	22 – 25°C
Relative humidity	41 – 58 %

4.3 Induction of disease

All the animals weighed between 20.6 – 25.2 grams at the start of the study. On day 0 neutropaenia was induced by means of an intraperitoneal injection of 0.5 ml Ifosfamide (Holoxan®) at a concentration of 8.6 mg/ml[1]. (Refer to Appendix 1-5)

4.4 Fate of the Animals

All animals were euthanasedby means of anIsoflurane overdoseupon completion of the study and the carcasseswere sent for incineration.

5. Treatments

5.1 Treatments and dose

Group 1: GSCF 1 protein at 0.5 ml / mouse of a 32µg/ml solution in 0.1% BSA and Tris-Cl
Group 2: GSCF 7 protein at 0.5 ml / mouse of a 32µg/ml solution in 0.1% BSA and Tris-Cl
Group 3 (Positive Control):Neupogen (filigrastim) at 0.5 ml / mouse of a 32 µg/ml solution in 0.1% BSA and Tris-Cl
Group 4 (Negative Control):0.5 ml Carrier substance /mouse (0.1% BSA in Tris-Cl)

5.2 Allocation of animals to treatment groups

Animals were assignedto treatment groups randomly (Table 4) by means of weight stratification onday 0. Starting from the lightest animal, the animals were placed into blocks of four. The animals in each block were then randomly assigned to one of the four groups by use of random numbers generated in Excel (Microsoft Office 2007).

Table 4: Allocation of animals into treatment groups

Animal nr	Weight (g)	Group	Average group weight	Group Std Deviation
2	21.8	1	23.4	1.21
4	23.3	1		
6	24.4	1		
11	24.9	1		
15	23.9	1		
20	22.3	1		
5	22.1	2	23.2	1.19
8	24.0	2		
12	23.9	2		
17	24.6	2		
23	22.7	2		
24	21.6	2		
1	23.2	3	23.2	1.16
7	24.0	3		
9	23.3	3		
16	22.2	3		
19	21.5	3		
22	24.7	3		
3	23.7	4	23.3	1.57
10	24.2	4		
13	25.2	4		
14	23.1	4		
18	22.7	4		
21	20.6	4		

Group 1: GSCF 1 protein group, Group 2: GSCF 7 protein group, Group 3: Positive Control, Group 4: Negative Control

5.3 Dosing

All treatments were administered to a set volume per animal of 0.5ml.All treatments were given before 08:00 am. The dose as stipulated in point 5.1 above was administered to each group of animals on days 1 – 4. (Refer to Appendix 1-5 and 1-7)

5.4 Drug inventory, storage and disposition

The drugs and chemicals used in this study are presented in Table 5: Drug Inventory. (Refer to Appendix 1-8)

Table 5: Drug Inventory

Treatment Agents	Company	Manufacture Date	Batch Number	Expiry Date	Day Administered
Ifosfamide (Holoxan)	Aventis	10/2008	22012A	09/2012	0
Sterile Water for injection	Kyron	Not supplied	91161	07/2013	0
Filigrastim (Neupogen)	Roche	10/2009	82019801	10/2011	1, 2, 3 & 4
GCSF 1	CSIR	Not supplied	Not supplied	Not supplied	1, 2, 3 & 4
GCSF 7	CSIR	Not supplied	Not supplied	Not supplied	1, 2, 3 & 4
Tris-Cl	CSIR	Not supplied	Not supplied	Not supplied	1, 2, 3 & 4
10% Bovine Serum Albumin (BSA)	CSIR	Not supplied	Not supplied	Not supplied	1, 2, 3 & 4

Holoxan was stored in a safe at room temperature, while the Neupogen and Tris-Cl were stored at 3°Cas per package instructions. The GSCF1, GCSF 7 and 10% BSA were stored in the -70°C Freezer.

- The test solution for groups 1 and 2 was prepared under a laminar flow bench on the 20 October 2010 and immediately aliquoted into four 5ml falcon tubes for storage at -70. (Refer to Appendix 1-8). This allowed drug to be available for each day of scheduled treatment in the absence of excessive defrosting and refreezing of all the test substance.
- One night prior to treatment one falcon tube was placed in a fridge and allowed to defrost.
- Group 4's solution was freshly made up in the laminar flow hood on each day of use(Appendix 1-8).

6. Methods

6.1. Animal Monitoring: General Observations

Animals were observed twice a day for changes in habitus, locomotion, respiratory pattern, and grooming for the duration of the study. Animals were weighed daily at the same time each day.

6.2. Adverse Reactions

Animals were observed twice a day for signs of adverse reactions.

6.3. Clinical Pathology

Mice wereeuthanased on day 4 of the study, six hours after the administration of treatments. The animals were first anaesthetised with Isofluraneby insufflation using the open drop method. Once pedal reflex was abolished, a cardiocentesis was performed and the maximal amount of blood was withdrawn. Blood was collected in syringes flushed with heparin and transferred to a micotainer® blood tube. These samples were subsequently transferred to the clinical pathology laboratory for analysis.

A full blood count consisting of red cell count (RCC), haematocrit (Ht), mean corpuscular volume (MCV), mean corpuscular haemoglobin (MCH), mean cell haemoglobin concentration (MCHC), red cell distribution width (RDW), white cell count (WCC), neutrophil percentage (Neuts %), immature neutrophil percentage (N(Immat %), lymphocyte percentage (Lymph %), monocyte percentage (mono %), eosinophil percentage (eos %), basophil percentage (baso %), large unidentified cells LUC %, absolute mature neutrophil count (N(mat) abs), absolute immature neutrophil count (N(imm) abs), absolute lymphocyte count (Lymp abs), absolute monocyte count (mono abs), absolute eosinophil count (eos abs), absolute basophile count (baso abs), large unidentified cell count (LUC), platelet count (Plt C), mean platelet volume (MPV) was performed using a multi-parameter, automated haematology analyser (ADVIA2 120 Haematology system, Siemens) at the Onderstepoort Veterinary Academic Hospital Clinical Pathology Laboratory according to their internal SOPs. Blood smears were made on the day the blood was collected, the smears where evaluated to confirm the machine results over the following few days by the Clinical Pathology laboratory.

6.4. Pathology

Long bones were harvested once death of the animal had been confirmed. The bone was spilt open and stored in 10% buffered formalin for future reference. One container per animal was stored at the UPBRC.

6.5. Statistical Procedures

All statistical analysis was undertaken in SPSS version 17. Normality of the data was ascertained with the Shapiro-Wilk test following natural logarithmic transformation. Due to the lack of normality for treatment group 3, significance between the groups was ascertained using a Kruskal Wallis H Test. Post-hoc comparison made use of a Mann-Whitney U test.

7. Results

7.1. Animal Monitoring: General Observations

No changes in habitus, locomotion, respiratory pattern, and grooming for the duration of the study were recorded (Refer to Appendix 1-4 and 1-5). The mean group weights are displayed in figure2.On the 23/10/2010 the scale was changed and the animals weighed again, resulting in the apparent drop in weight on the graph (Refer to Appendix 1-4).

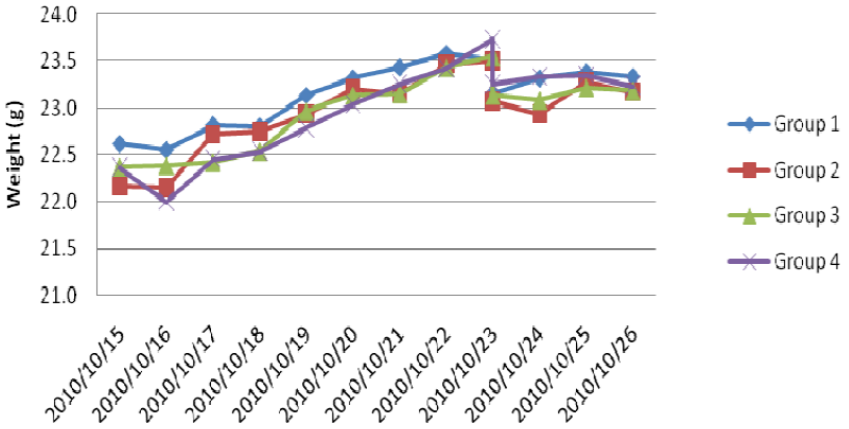

Figure 2: Mean Weights per treatment group from the time of acclimatisation

Group 1: GSCF 1 protein group, Group 2: GSCF 7 protein group, Group 3: Positive Control, Group 4: Negative Control

7.2. Adverse Reactions

On the 25/10/2010 Mouse 6 (from Group 1) had some blood at the injection site. The animal did not appear to be ill effected by this reaction. No further adverse reactions were noted for the duration of the study. (Refer to Appendix 1-6)

7.3. Clinical Pathology

The clinical pathology results are summarised per group in Table 6**Error! Reference source not found.**: Clinical Pathology Results. (Refer to Appendix 1-9)

7.3.1. General Haematology

A significant difference was present for the parameters MCV ($H(3)=8.867$, $p=0.031$), White cell count ($H(3)=8.566$, $p=0.036$), Neutrophil % ($H(3)=9.073$, $p=0.028$), Lymphocyte % ($H(3)=9.847$, $p=0.020$), Mature neutrophil count ($H(3)=9.658$, $p=0.022$) and absolute eosinophil count ($H(3)=14.543$, $p=0.002$). The immature neutrophil count was not evaluated as all groups had the same count of zero.

On post hoc analysis a statistical difference was found between group 1 and the negative control on MCV (p=0.036). The same difference was noted for group 2 MCV (p=0.01). A statistical difference between group 3 and 4 was found on WCC (p=0.025), Neut % (p=0.025), Lymp % (p=0.016), mature neutrophil count (p=0.025) and eosinophil count (p=0.008). The difference in MCV is not deemed to be relevant as the actual values are between 45.1 and 48. (Refer to Appendix 2 for the Statistical Data)

Table 6: Clinical Pathology Results

Parameter	Group 1		Group 2		Group 3		Group 4	
	Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM
RCC (x 10e12/l)	7.93	0.56	8.80	0.20	8.74	0.10	8.47	0.14
Ht (l/l)	0.39	0.02	0.40	0.01	0.41	0.01	0.40	0.00
MCV (Fl)*	46.40#	0.20	45.90#	0.30	46.30	0.30	47.20	0.30
MCH (g/dl RC)	15.20	0.10	15.20	0.10	15.30	0.00	15.50	0.10
MCHC (g/dl RC)	32.72	0.20	33.12	0.29	32.92	0.27	32.77	0.15
RDW (%)	13.57	0.61	13.70	0.52	13.92	0.62	16.15	1.67
WCC (x 10e9/l)*	1.22	0.22	1.32	0.11	5.31#	0.96	1.09	0.14
Neuts %*	12.62	2.30	19.45	4.19	56.95#	10.27	9.48	3.02
N (immat %)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lymph %*	70.57	4.28	65.30	5.85	32.17#	9.49	79.07	4.81
Mono %	1.57	0.51	1.97	0.69	1.30	0.55	1.38	0.53
Eos %	0.58	0.11	0.65	0.10	1.43	0.36	0.55	0.14
Baso %	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LUC %	16.05	2.66	12.62	2.77	8.78	1.68	9.50	1.80
N (mat)abs (x 10e9/l)*	0.15	0.04	0.27	0.07	3.48#	0.79	0.12	0.05
N (immat)abs (x 10e9/l)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lymph abs (x 10e9/l)	0.86	0.16	0.84	0.07	1.29	0.12	0.84	0.09
Mono abs (x10e9/l)	0.02	0.01	0.02	0.01	0.08	0.04	0.02	0.01
Eos abs (x 10e9/l)*	0.01	0.00	0.01	0.00	.09#	0.03	0.01	0.00
Baso abs (x 10e9/l)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LUC (x10e9/l)	0.19	0.04	0.17	0.05	0.42	0.10	0.11	0.03
Plt C (x 10e9/l)	698.00	51.00	733.00	52.00	513.00	89.00	697.00	64.00
MPV (fl)	8.77	0.23	8.73	0.19	9.68	0.79	9.00	0.14

Group 1: GSCF 1 protein group, Group 2: GSCF 7 protein group, Group 3: Positive Control, Group 4: Negative Control
* indicates significance evident on a Kruskal Wallis H Test
# indicates a significant difference detected on a Mann Whitney test when compared to group 4

7.3.2. Neutrophil Counts

The mean neutrophil counts per treatment are presented in Figure 3.
In summary:

- Negative Control (Group 4): The cytotoxic Ifosfamide had the desired effect as the negative control group became severely neutropaenic following administration.

- GCSF 1 (Group 1) induced a marginal increase in the neutrophil count at 4.36% of the effect of the positive control group. The increase was not statistically different.
- GCSF 7 (Group 2) performed in a similar manner to group 1, with only 7.75% of the effect of the positive control group. No statistical difference was present.
- Positive Control (Group 3) induced the expected increase in white cell count and neutrophil count. The increase in neutrophil count was 29 fold. The increase was significantly different to the negative control.

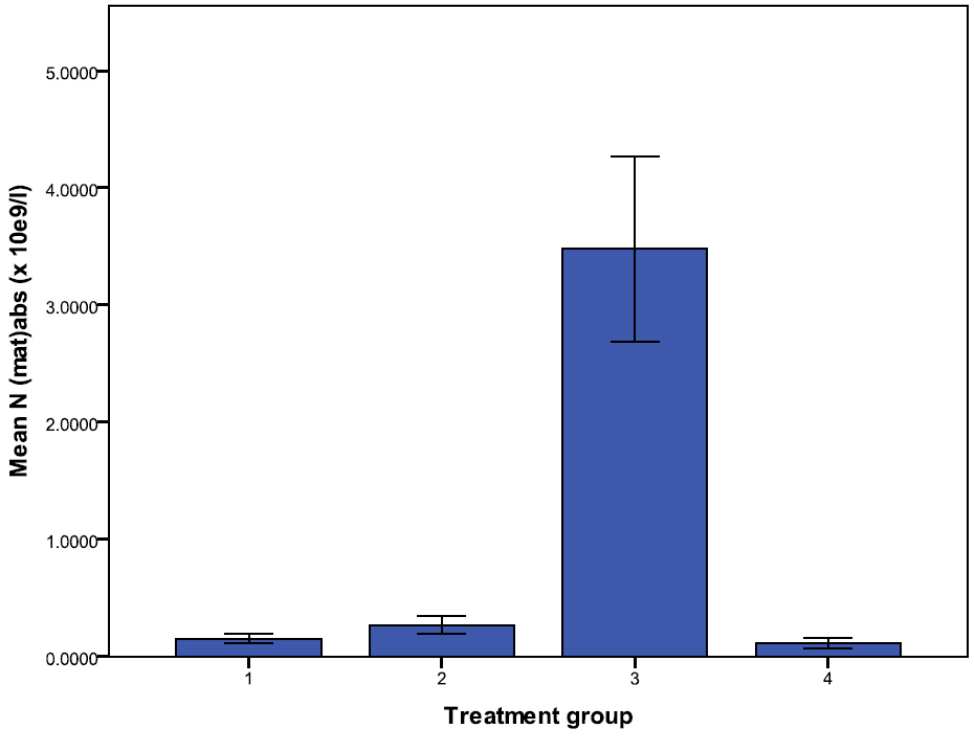

Figure 3: Mean Neutrophil counts for each treatment group. (Error bar indicates one standard deviation)*Group 1: GSCF 1 protein group, Group 2: GSCF 7 protein group, Group 3: Positive Control, Group 4: Negative Control*

8. Conclusion:

The GSCF1 & 7 from the CSIR failed to induce a significant increase in neutrophil countin this murine model.

9. References

1. Vanz,A., Renard, G., Palma, M., Chies, J., Dalmora, S., Basso, L., Santos D. Human granulocyte colony stimulating factor (hG-CSF): cloning, overexpression, purification and characterization. Microbial Cell Factories 2008, 7:13.