# Cloning and expression of recombinant bovine glycine N-acyltransferase from bovine liver

By

**M Jooste**

**Dissertation submitted in partial fulfilment of the requirements for the degree *Master of Science in Biochemistry* at the Potchefstroom campus of the North-West University**

**Supervisor: Prof. A.A. van Dijk**

**August, 2011**

## SUMMARY

The importance and social focus on detoxification and improving a person's lifestyle has increased tremendously over the years. This is due to increasing levels of pollution and the resulting increase of foreign chemicals (toxins) introduced into our bodies. Detoxification is the process of clearing toxins from the body by turning them into other components that are not harmful to the body. Detoxification has become a popular topic and it is well established to the point where detoxification profiling is becoming available as a commercial service. Detoxification is a multiple reaction system that is divided into four phases: Phase 0, Phase I (Functionalization), Phase II (Conjugation) and Phase III (Elimination).

The focus of this MSc project was on glycine N-acyltransferase (GLYAT) which is a transferase enzyme involved in Phase II detoxification, specifically glycine conjugation. GLYAT is responsible to detoxify benzoic acid, a xenobiotic widely used in food preservatives to form a glycine conjugate known as hippuric acid (benzoyl glycine). The GLYAT conjugation reaction is also important for the metabolism of drugs such as Aspirin (Levy, 1965). More important for this study, is the fact that glycine supplementation is commonly used in the treatment of specific inborn errors of metabolism (IEM). When some metabolic pathways are defective, accumulating substrates that form can conjugate with glycine and excreted, thereby lessening toxicity in the body.

Although humans are closer to mice than bovine on a genetic level, bovine and humans have sufficient DNA sequence similarity to map the human genome almost entirely to the genome of the bovine. In this study, the differences and similarities of the molecular characteristics of bovine GLYAT (bGLYAT) and the human family of GLYAT enzymes were compared. The goal of this study was to clone bGLYAT and generate a recombinant bGLYAT. The recombinant enzyme can then be used for further studies to compare the enzymatic characteristics and species specificity of human GLYAT to bovine GLYAT.

I used bacteria to express bGLYAT. Bacteria are simple organisms and lack the ability to fold some recombinant eukaryotic proteins correctly. Molecular chaperones have been demonstrated to be involved in protein folding. Co-expression of chaperones such as trigger factor, GroEL-GroES and DnaK-DnaJ-GrpE with the target protein in bacteria may enable efficient folding of expressed proteins and enhances the production of active enzymes.

In Chapter 2 I described how I investigated whether the chaperone trigger factor (TF) can assist with protein folding of recombinant bGLYAT. I prepared cDNA from bovine liver total RNA. The open reading frame (ORF) of bGLYAT was amplified using gene specific primers with addition of NdeI and SalI sites in a two step RT-PCR procedure. bGLYAT was cloned into pColdTF and sequenced. The plasmid containing bGLYAT in the correct reading frame was then introduced into Origami™ cells. pColdTF is a vector that expresses trigger factor (56 kDa) as a fusion protein with recombinant proteins. bGLYAT with fusion protein (88 kDa) was over-expressed and was soluble. However, the recombinant bGLYAT expressed as a fusion protein with TF did not have any enzyme activity. It was speculated that the fusion protein might be affecting the enzyme activity. After removal of the fusion protein by means of thrombin cleavage, bGLYAT (36 kDa) still did not have any enzyme activity.

In Chapter 3 I described how I investigated whether bGLYAT (without any fusion proteins) co-expressed with chaperones would be soluble and enzymatically active. Sequencing results confirmed cloning of the DNA encoding bGLYAT ORF into pColdIII in the correct orientation and reading frame. bGLYAT was co-expressed with TF alone, or TF together with GroEL-GroES. In each case, the bGLYAT was insoluble according to SDS-PAGE results. Although bGLYAT could not be purified from other proteins, because it did not contain a tag, or seen to be soluble on SDS-PAGE, a GLYAT enzyme activity test was performed on various fractions of the bacterially expressed protein lysates. Both soluble fractions of bGLYAT co-expressed with TF alone or TF together with GroEL-GroES had GLYAT enzyme activity. This is the first report of an enzymatically active recombinant bGLYAT. The effect of pH on the solubility of bGLYAT was briefly investigated but the results were not conclusive.

In future studies, the bGLYAT should be expressed as a protein with a His or GST tag for purification and Western blotting performed on bacterial lysate fractions. The

recombinant bGLYAT can also be used for future site-directed mutagenesis studies, thereby analysing the impact of SNPs on enzyme activity and amino acid specificity.

# OPSOMMING

Die sosiale belangrikheid en fokus van detoksifikasie en die doel om menslike gesondheid te verbeter het geweldig toegeneem oor die jare. Dit is as gevolg van die toenemende vlakke van besoedeling in die omgewing en toksiese stowwe waaraan mense blootgestel word. Detoksifikasie is die proses in die liggaam om van toksiese stowwe in die liggaam ontslae te raak deur hulle in nie-skadelike komponente te verander. Detoksifikasie is ‘n baie populêr en is goed geïmplimenteer tot die punt waar ‘n persoon se detoksifikasie profiel verkrygbaar kan wees as ‘n diens. Detoksifikasie is bekend as a multi-ensiemreaksie-sisteem en word verdeel in vier fases: Fase 0, Fase I (Funksionalisering), Fase II (Konjugering) en Fase III (Eliminering).

Die fokus van hierdie MSc projek was op glisien-N-asieltransferase (GLYAT) wat ‘n transferase ensiem is wat betrokke is by Fase II detoksifikasie, spesifiek glisienkonjugering. GLYAT is verantwoordelik om benzoësuur te detoksifiseer. Benzoësuur is ‘n xenobiotika (eksogene toksien) wat in verskeie voedsel produkte gebruik word as preserveermiddel. GLYAT konjugeer die benzoësuur met glisien om hippuursuur (benzoïelglisien) te vorm wat maklik uit die liggaam verwyder kan word. Die GLYAT konjugasie reaksie is ook belangrik vir die metabolisering van Aspiriene (Levy, 1965). Meer belangrik vir hierdie studie is die feit dat glisien supplementering ‘n baie bekende behandeling is vir sekere aangebore metabolise siektes (IEM). Wanneer ‘n metaboliese weg defektief is, kan sommige van die akkumulerende substrate met glisien konjugeer en uitgeskei word om die toksiese vlak daarvan in die liggaam verlaag.

Alhoewel dit bekend is dat mense op genetiese vlak nader is aan die muis as die bees, het die mens en bees genoegsame DNS volgordes wat ooreenstem met mekaar om die twee genome amper heeltemal met mekaar te vergelyk. In hierdie studie is die ooreenkomste en verskille van die molekulêre karakteristieke van bees GLYAT en die familie van die mens GLYAT ensieme met mekaar vergelyk. Die doel van die studie was om die bees GLYAT te kloneer en ‘n rekombinante bees GLYAT ensiem te genereer. Hierdie rekombinante ensiem kan dan in verdere studies gebruik word om die ensiemkarakteristieke en spesie spesifisiteit van mens GLYAT met bees GLYAT te vergelyk.

Ek het bakterieë gebruik om die rekombinante bees GLYAT uit te druk. Bakterieë is eenvoudige organismes maar het nie die vermoë om sommige rekombinante eukariotiese proteïene korrek te vou in hul tersiêre struktuur nie. Molekulêre chaperone is betrokke by die vouing en vervoer van proteïene. Deur chaperone Trigger Faktor (TF), GroEL-GroES en DnaK-DnaJ-GrpE saam met rekombinante proteïene uit te druk in bakterieë kan korrekte vouing bevoordeel word en verhoog die kans om aktiewe ensieme te produseer.

In Hoofstuk 2 het ek beskryf hoe ek ondersoek het of TF kan help met die vouing van die rekombinante proteïen bees GLYAT. Ek het kDNS voorberei vanaf beeslewer se totale RNS. Die oop leesraam van bees GLYAT is vermeerder deur geen-spesifieke voorvoerders te gebruik met additionele NdeI and SalI restriksie-ensiem plekke in ’n twee-stap omgekeerde transkripsie polimerase ketting reaksie. Bees GLYAT is gekloneer in pColdTF en die nukleïensuur volgorde daarvan bepaal. Die plasmied wat die bees GLYAT bevat (wat in die korrekte leesraam was) was daarna in Origami™ selle getransformeer. pColdTF is ‘n vektor wat TF (56 kDa) uitdruk as a fusie proteïen. Bees GLYAT met die fussie proteïen (88 kDa) was uitgedruk en was oplosbaar. Die rekombinante bGLYAT met fussie proteïen TF het geen ensiemaktiwiteit gehad nie. Daar is gespekuleer dat die fussie protein dalk die ensiemaktiwiteit beïnvloed. Na die verwydering van die fussie proteïen deur middel van trombien vertering, was bGLYAT (36 kDa) steeds nie aktief nie.

In Hoofstuk 3 het ek beskryf hoe ek ondersoek het of bGLYAT (sonder enige fussie proteïen) uitgedruk word saam met molekulêre chaperones en of bGLYAT dan sal oplosbaar en ensiematies aktief sal wees. Nukleïensuur volgorde bepaling het bevestig dat die DNS wat die oop leesraam van bGLYAT bevat suksesvol gesubkloneer is in pColdIII in die korrekte oriëntasie en oop leesraam. bGLYAT is uitgedruk saam met TF alleen, of met TF met GroEL-GroES onderskeidelik. In albei gevalle was bGLYAT onoplosbaar volgens SDS-PAGE. Alhoewel bGLYAT nie gesuiwer kon word van ander proteïene met behulp van affiniteitschromatografie met HIS of GST sterte nie, of op ‘n SDS-PAGE in die oplosbare fraksie gesien kon word nie, is die ensiemaktiwiteit nogtans bepaal in verskeie bakteriële lisaat fraksies. Beide bGLYAT uitgedruk saam met TF alleen, of met sowel TF en GroEL-GroES onderskeidelik, het ensiemaktiwiteit getoon. Dit was die eerste bewys van

ensiematies aktiewe rekombinante bGLYAT. Die effek van pH op die oplosbaarheid van bGLYAT was vlugtig ondersoek, maar die resultate was onbeslis.

Vir toekomstige studies kan bGLYAT uitgedruk word met ‘n His of GST stert vir die suiwering van bGLYAT. Westelike klad moet ook gedoen word op verskeie bakteriële fraksies. Die rekombinante bGLYAT kan ook gebruik vir vir verdere spesifieke volgorde mutasie analises om sodoende die impak daarvan op ensiemaktiweit te ondersoek.

Hierdie werk word opgedra aan my man Jacques.

# ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to the following persons and institutions for their contributions and support towards the completion of this study:

My supervisor Prof A.A. van Dijk for her help, guidance and commitment

Retha Potgieter for the ordering of products

Rencia van der Sluis, Lizelle Zandberg and Jeanine Labuschagne for being my mentors and friends in and out of the laboratory

National Research Foundation (NRF) and North-West University bursaries

Funding of NRF FA/DST NRF BioPAD BPP0007

My parents for all their support and most importantly my saviour God.

# DECLARATION

I declare that the dissertation for the degree of Master of Natural Sciences (MSc) at the North-West University: Potchefstroom Campus hereby submitted, has not been submitted by me for a degree at this or another University, that it is my own work in design and execution and that all material contained herein has been duly acknowledged.

…… ……..

**Maritza Jooste** **Date**

# LIST OF FIGURES

# LIST OF TABLES

# TABLE OF CONTENTS

# CHAPTER 1

# LITERATURE REVIEW

## 1.1 Introduction

Today, we are facing serious problems due to increasing levels of foreign chemicals (toxins) in our body. Toxins are substances that cause harm to living organisms. They originate both within (endogenous) or outside (exogenous) the organism. Exogenous toxins, also known as xenobiotics are introduced into our bodies by eating fast food meals and preserved food products, inhaling polluted air, drug use and habits we have adopted such as smoking (Liska, 1998). In our daily routine, we are constantly reminded of these toxins by the media, since companies exploit detoxification as a way to market their products. Popular consumer products include foot pads, 'detox' diets, performance enhancers, green teas and many more. By improving one's lifestyle habits, one can reduce the exposure to xenobiotics.

## 1.2 Introduction to detoxification

Detoxification, also referred to as biotransformation, is the process of clearing toxins from the body or transforming them into other components that are not harmful to the body. Humans and animals have the ability to minimize the potential of damage caused by toxins. They have complex detoxification enzyme systems found mainly in the liver and kidney as well as in some other organs. After detoxification, biotransformed toxins are excreted via urine or bile (Liska, 1998). Knowledge of the metabolism of xenobiotics is basic to a rational understanding of various research fields such as pharmacology, therapeutics, toxicology, management of cancer and drug addiction. All these areas involve administration of, or exposure to, xenobiotics (Murray *et al*., 2006).

Hippuric acid was one of the first metabolites identified. In 1773 it was established that hippuric acid was the result of conjugation of glycine with benzoic acid (Keller,

1842). In 1831 Professor Wohler hypothesized that animals can transform xenobiotics that they have consumed. An experiment was done on the passage of benzoic acid in dog urine. When benzoic acid was ingested, it was converted to hippuric acid. This statement was not proven until Keller performed a test where he ingested benzoic acid, collected his urine and showed the direct relationship of ingesting benzoic acid and excreting hippuric acid (Keller, 1842). The work of Keller is at the root of this study because I aimed to investigate the enzyme responsible for the second step of metabolizing benzoic acid to hippuric acid. This enzyme is known as glycine N-acyltransferase (E.C. 2.3.1.13). The abbreviation of glycine N-acyltransferase is GLYAT.

In 1947 R.T. Williams defined the field of detoxification. He proposed that a non-reactive compound such as a toxin accumulates in the body and can be biotransformed and eliminated in two phases. The first phase is known as functionalization during which oxygen is used to form a reactive site on the toxin. The second phase is known as conjugation, which results in addition of a water soluble group to the reactive site in order for the toxin to be excreted via urine (Williams, 1947).

## 1.3 Inborn errors of metabolism

### 1.3.1 The basic understanding of inborn errors of metabolism

The totality of chemical reactions in a living organism is known as the metabolism. When genetic disorders cause an alteration of the metabolism, it is termed an inborn error of metabolism (IEM). In 1908 Sir Archibald Garrod studied patients with rare disorders such as alkaptonuria, albunism, cystonurie and others. He established that these conditions were genetically determined. He was the first to refer to these conditions as inborn errors of metabolism (Garrod *et al,.* 1923). There are more than 2000 metabolic enzymes known in humans. Defects in more than 600 of them are known to be able to cause IEMs (Baric *et al.,* 2001). Metabolic diseases are caused by gene defects responsible for the anabolism or catabolism of proteins, carbohydrates and lipids. Figure 1 illustrates the pathogenic mechanism of an IEM due to an enzymatic blockage.

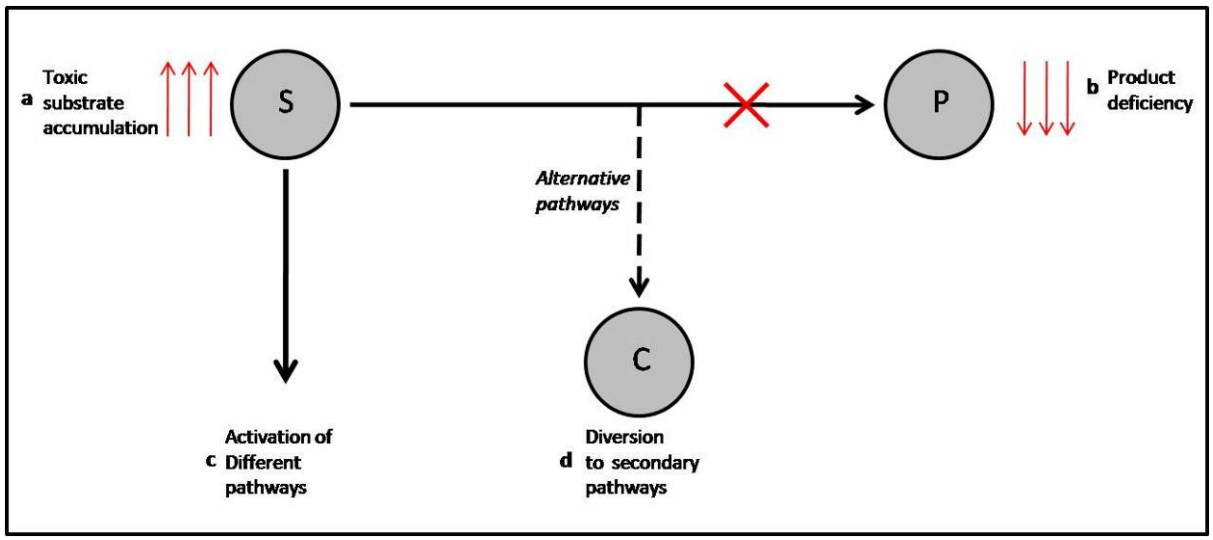



**Figure 1: Schematic illustration of the pathogenic metabolism due to an inborn error of metabolism.** The blocked metabolic pathway (due to a defective enzyme indicated with the red cross) results in: (a) direct toxicity of the accumulating upstream substrate (S); (b) deficiency of the downstream products (P); (c) activation of alternative pathway and (d) diversion of the metabolic flux to secondary pathways and alternative metabolite (C) production (Taken from Lanpher *et al.,* 2006).

When a metabolic pathway is blocked, the upstream substrate accumulates, resulting in toxicity while a deficiency of the downstream products is caused. This pathway blockage can sometimes activate an alternative pathway. The metabolic flux can then be diverted into a secondary pathway forming an alternative metabolite (Lanpher *et al.,* 2006). The accumulated substrate can also be referred to as an endogenous toxin and needs to be detoxified and eliminated from the body to avoid illness.

When a patient is suffering from an IEM, the rapid accumulation of metabolites generally exceeds the rate of detoxification. The detoxification system becomes overloaded as toxic endogenous metabolites accumulate. Abnormalities due to IEMs differ in neonates, children and adults. Symptoms will manifest once a certain concentration of the toxic metabolite is reached (Beaudet *et al.,* 1995). These symptoms may include nausea, lethargy, failure to thrive, dimorphic features, abnormalities of the hair, skin and or skeleton, abnormal odour, abnormal muscle tone, mental retardation or even death (Saudubray, *et al*., 1995 and Murray *et al*, 2006).

### 1.3.2 Management and therapy of inborn errors of metabolism

There are extreme patient-to-patient variations in symptoms for each diagnosed IEM. Therefore, each patient's management and therapy of the IEM should be personalized according to symptoms and severity thereof. General therapy includes dietary supplementation or restriction and enhancing detoxification. Dietary restriction of the metabolite upstream (substrate) of the block has become a central management strategy for many IEMs. For example, in the case of phenylkenonuria (PKU), the substrate phenylalanine accumulates in the body due to the phenylalanine hydroxylase deficiency. This enzyme is necessary to metabolize the amino acid phenylalanine to the amino acid tyrosine. As management of this IEM, phenylalanine is therefore restricted from the diet. Alternatively, the product tyrosine must be supplemented (Murray *et al*., 2006).

Another way of treating some IEMs is by means of enzyme replacement therapy. In the case of Gaucher disease Type 1, replacement of the defective enzyme has been successfully used for treatment (Pastores & Barnett*,* 2005). Enzyme replacement therapy on some patience with certain IEM has been evaluated for more than ten years. In some cases, a degree of immune response developed for these proteins and the enzyme replacement therapy could be used anymore.

Another therapeutic approach, and the one most important for this study, is the treatment of IEMs by means of glycine and carnitine supplementation. Phase II glycine and carnitine conjugation reactions are the major Phase II detoxification systems operational in some IEMs. Due to exhaustion of these substrates caused by certain IEM, secondary glycine and carnitine deficiency arises as the body tries to conjugate toxins to glycine and carnitine for excretion. Glycine (a non-essential amino acid) stores are depleted and glycine supplementation is therefore important for the treatment of the IEMs. The excretion of glycine and carnitine conjugates increases significantly after glycine and carnitine therapy (Tsai *et al.,* 1989; Rolland *et al.,* 1985; Baumgartner*,* 2003). In general, treatment of patients with IEMs involves enhancement of detoxification as well as continuous monitoring of the clinical symptoms of the patients.

Glycine therapy is used to treat organic acidemias such as isovaleric-acidemia (Ozand & Gascon, 1991). Patients suffering from organic acidemia normally excrete monocarboxylic acids with the corresponding N-acylglycine conjugate in the urine (Ozand & Gascon, 1991). A patient suffering from isovaleric-acidemia usually responds very well to glycine supplementation and excretes isovalerylglycine in urine (Tanaka & Isselbacher, 1967). However, too much glycine can be harmful to the body and can cause hyperammonemias and non-ketotic hyperglycinemia that manifest in neurological problems. These diseases can be treated with benzoate therapy in order to utilize the excess glycine with GLYAT (Mawal & Qureshi, 1994). As mentioned before, benzoate is metabolized by GLYAT by means of glycine conjugation, forming benzoyl-glycine. Glycine N-acyltransferase (GLYAT), the enzyme that was the focus in this study has no known associated IEM. It is investigated due to the fact that it has a beneficial effect in the treatment of IEMs. Glycine therapy is essential for better Phase II detoxification reactions. A fine balance of glycine supplementation is needed to do glycine therapy. More information of detoxification and GLYAT characteristics will be given later in this chapter.

### 1.4 Detoxification

At the moment we know detoxification as a multiple reaction system that is divided into four phases. Phase 0, Phase I (functionalization), Phase II (conjugation) and Phase III (elimination) (Nakata et al., 2006). Figure 2 (Liska, 1998) explains detoxification in more detail.

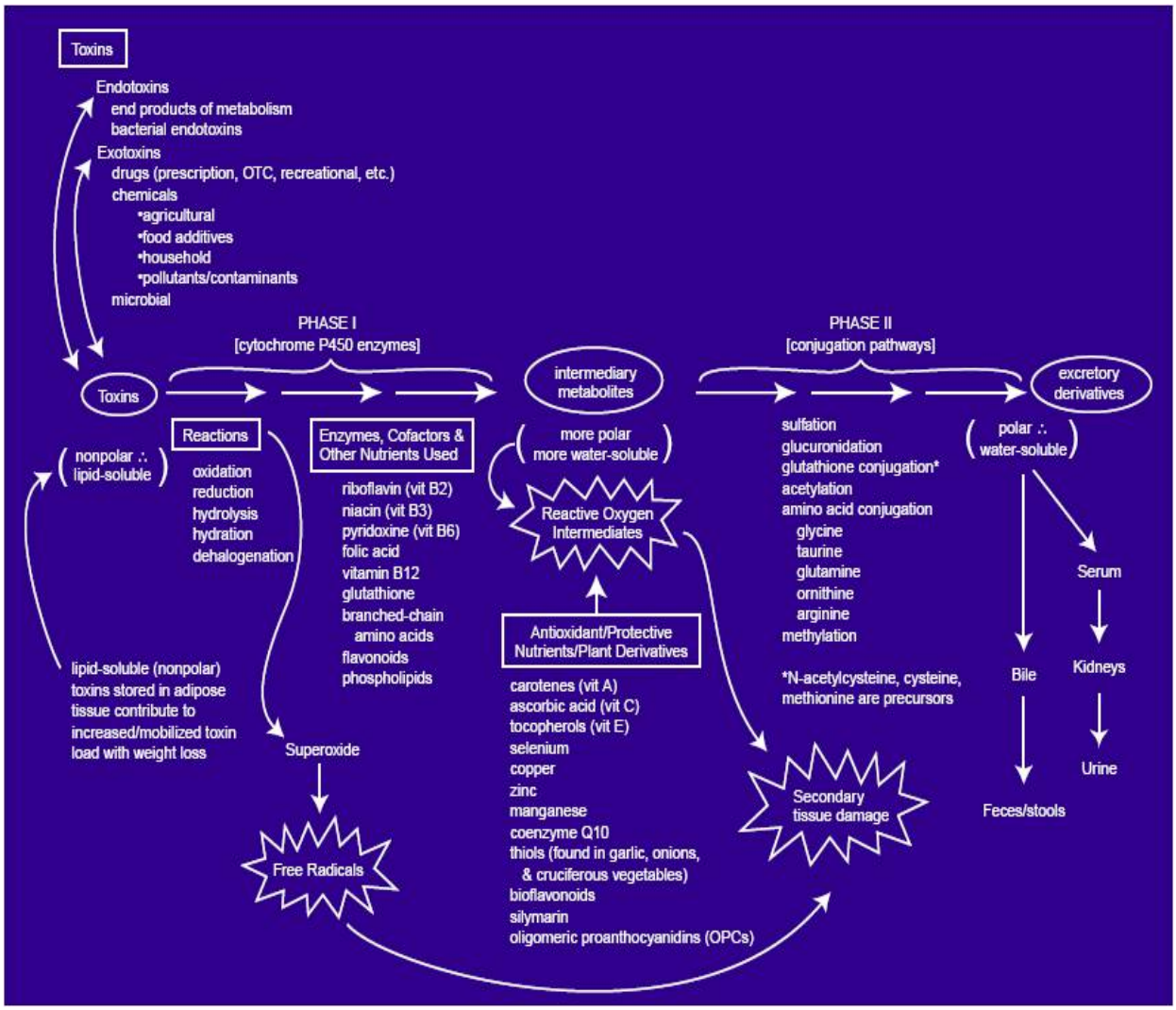



**Figure 2: Schematic illustration of the liver detoxification pathway.** Endotoxins, exotoxins and lipid-soluble molecules will first enter the Phase I detoxification pathway. The intermediary molecule formed is more polar and can therefore enter the Phase II (conjugation) pathway. The product formed is polar (water soluble) and the body is able to excrete the excretory derivatives via urine or bile. Reactions, enzymes, co-factors and other nutrients for the two phases are listed. The intermediatary metabolite can also react with reactive oxygen intermediates and can cause secondary tissue damage. Antioxidants, protective nutrients as well as plant derivatives can equalize the reactive oxygen intermediate and thereby eliminate the risk of secondary tissue damage (Liska, 1998).

Exogenous toxins enter the body from e.g. food products or endogenous toxins are present due to end products of metabolism. Non-polar toxins can also be stored in adipose (fatty) tissue, seeing that they are hydrophobic. Phase I detoxification will be the first to start detoxification of these toxins. Figure 2 lists the reactions involved in Phase II detoxification as well as enzymes, co-factors and other nutrients involved. After Phase I detoxification, reactive intermediary metabolites are present. Phase II will detoxify the intermediary metabolites further by making them more polar. The

reactions involved are shown in Figure 2 (Liska, 1998). Phase III is a membrane transport system that eliminates Phase II metabolites from the cells (Nakata et al., 2006). The water soluble excretory derivatives are eliminated from the blood via the kidney in urine, or via stools with bile (Liska, 1998).

### 1.4.1 Phase I detoxification: Functionalization

Phase I reactions involve mainly hydrolysis but also reduction, oxidation, hydration and dehalogenation (Murray *et al*., 2006). These reactions expose or introduce a functional group (-OH, -$NH_2$, -SH or –COOH) on the xenobiotic which leads to a small increase in hydrophilicity of the compound (Parkinson, 2003). The functional groups provide the 'handle' for Phase II conjugation species to react with. Usage of oxygen and NADH (a co-factor) to add a reactive group (hydroxyl radical) is a typical Phase I reaction (Parkinson, 2003). Pharmaceuticals, chemical pollutants and endogenous molecules such as steroids are primarily metabolized by the Phase I biotransformation (Liska, 1998).

The most popular and thoroughly study enzyme group of the Phase I detoxification pathway is the cytochrome P450 enzyme system. These enzymes usually start the Phase I detoxification process. Cytochrome P450 enzymes are found in virtually all tissue, but are most abundant in the hepatic endoplasmic reticulum (microsomes) (Parkinson, 2003). Among the Phase I biotransforming systems, the cytochrome P450s are the most prominent in terms of catalytic versatility and the number of xenobiotics it detoxifies or activates (Vermeulen, 1996). Approximately 50% of the common pharmaceutical drugs humans take every day are metabolized by isoforms of the cytochrome P450s. These enzymes also act on various carcinogens and pollutants. It was shown that an abnormality in CYP17 enzyme function plays an essential role in the development of some cancers seeing that the protein is the key enzyme in steroid hormone biosynthesis. This makes the development of selective inhibitors for human CYP17 an important task. A clinical advance in the P450 research is inducers of cytochrome P450 enzymes that increase the rate of xenobiotic biotransformation (Parkinson, 1995). Another clinical advance is the work of Pechurskaya *et al.* (2008) who engineered, expressed and purified a recombinant 'soluble' membrane-bound human Cytochrome P45017*alpha.*

After Phase I detoxification many substrates and/or their metabolites are chemically reactive and their continued presence may lead to toxicity (Parkinson, 1995). This is due to oxidants like superoxide ($O_2^-$) that are formed in cells by the cytochrome P450 enzymes during metabolization. They are referred to reactive oxygen species (ROS) with destructive effects by the free-radical chain reactions where an unpaired electron is constantly transferred. The ROS such as $OH^-$ can react with proteins, nucleic acids and lipids and can harm the cell and cause secondary tissue damage (Murray *et al.,* 2006). Failure to further metabolize the ROS by Phase II, has been shown to increases the risk of diseases such as cancer and Parkinson's disease (Kawajiri *et al*., 1990; Bandmann *et al.,* 1997).

In order for the detoxification enzyme system to function optimally, there should be a fine balance between the Phase I and Phase II reactions. Phase II reactions are typically much faster than Phase I detoxification reactions in order to avoid high levels of reactive intermediates, which can cause damage to DNA, RNA and proteins. The availability of the co-factor such as glycine in the GLYAT reaction may also be the rate-limiting factor in patients with specific IEMs. In some Phase II pathways glycine may be depleted and it prevents the conjugation of glycine with the accumulating substrate. A promoted Phase II detoxification pathway leads to better detoxification and will maintain the balance between the two phases (Oesch et al., 1999).

### 1.4.2 Phase II detoxification: Conjugation

As stated above, Phase II detoxification is a very important follow-up reaction for Phase I detoxification in order to reduce the risk of tissue damage and certain diseases. Refer to Figure 2 for the mechanism on how secondary tissue damage is caused (Liska, 1998). The different enzymes involved in Phase II detoxification are classified as transferase enzymes which perform the transfer of a co-factor to functional groups present or introduced by Phase I on the compounds. Multiple conjugation reactions are involved in Phase II detoxification such as glucuronidation, sulphation, acetylation, methylation and conjugation with glutathione and amino acids (Reviewed by Vermeulen, 1996).

In Section 1.3.2 the importance of glycine therapy for IEMs was discussed. GLYAT is responsible for the amino acid (glycine) conjugation reaction of the Phase II detoxification pathway and utilizes glycine for detoxification of benzoic acid, but is also to a lesser extent a xenobiotic. Xenobiotics such as benzoic acid contain an aromatic hydroxylamine group. These groups are possible substrates for amino acid conjugation with glycine (Parkinson, 1995). This transfer reaction is activated with coenzyme A to produce acyl-CoA (CoA-SH) thioesters. The coenzyme A serves as a recycle shuttle. Coenzyme A transports substrates from the point of generation to the point of utilization (Parkinson, 1995). The acyl-CoA is then conjugated with the amino group of glycine to form amide-linked conjugates. Substrates that undergo these conjugation steps are benzoic acid, phenylacetic acid and salicylic acid (Mawal & Qureshi, 1994). More details about the reaction will be given in Section 1.5, Figure 6.

In this study, I aimed to express a recombinant detoxification enzyme called glycine N-acyltransferase, a Phase II detoxification enzyme. If successful, we plan to use it initially for basic structure-function studies. In follow-up projects, we will start to investigate how a recombinant GLYAT can be developed into a therapeutic drug for IEMs.

### 1.4.3 Phase III detoxification: Elimination

When ingesting toxins, they are transported from the stomach and intestines by enterocytes to the liver where the majority of detoxification reactions take place. After Phase I and II detoxification, the toxins are more hydrophilic. The body is thus able to eliminate the toxins by means of urine and bile in Phase III as illustrated in Figure 3.

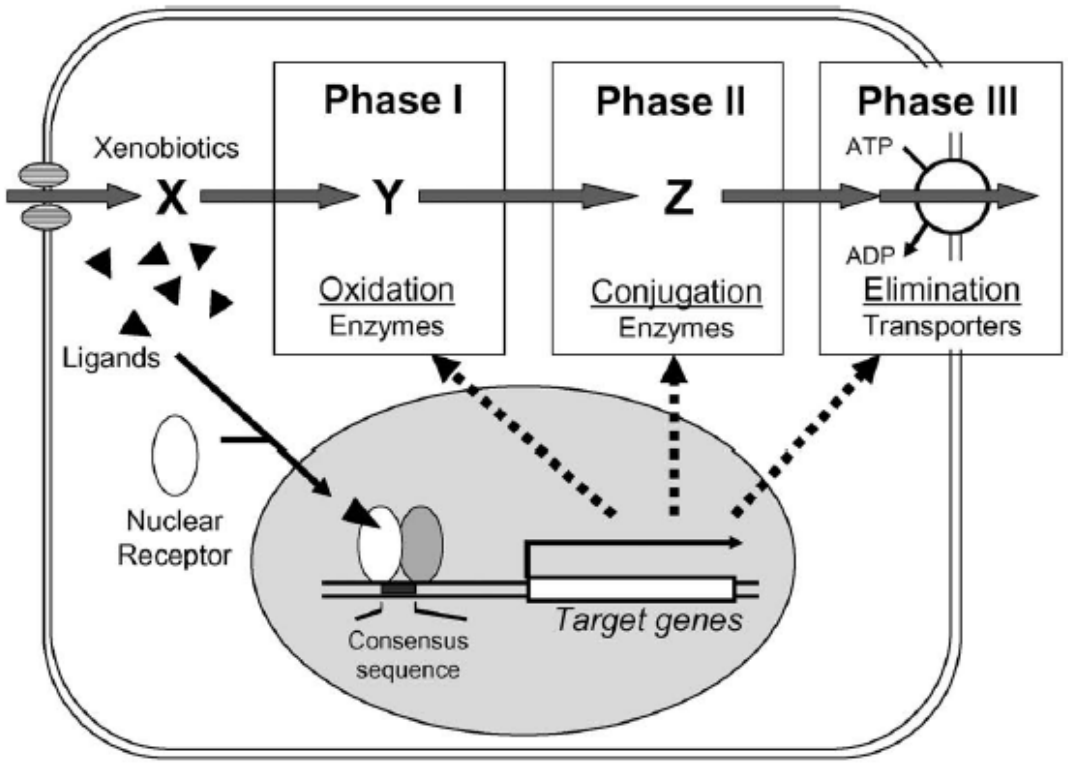



**Figure 3: Schematic illustration of Phase I, II and III drug metabolizing systems and the role of nuclear receptors in induction of enzymes and transporters in those systems**. The Phase I system involves the P450 (CYP) or flavin mixed-function oxidation that catalysis oxidation reactions of xenobiotics (X). The Phase II system deals with the conjugation of Phase I products (Y) to form a variety of conjugating metabolites (Z). Transporters in the Phase III system eliminate Phase II metabolites from cells. Hydrophobic ligands (xenobiotics and/or endogenous metabolites) and nuclear receptors are critically involved in the induction or down-regulation of various enzymes and drug transporters (target genes) involved in Phase I, II and II drug metabolizing systems (Ishikawa, 1992).

Figure 3 is an illustration of the typical drug metabolizing reactions and the cascade effect of the nuclear receptor that is involved in the induction or down regulation of various genes expressed involved in Phase I, II and III of the drug. If detoxification metabolites accumulate in the cell, it can lead to a decrease in Phase II detoxification activity.In 1992, Ishikawa described Phase III detoxification (Figure 3). It is a membrane transport system that eliminates Phase II metabolites from the cells using an ATP-dependent export pump (Ishikawa, 1992). There are more than 40 different human ATP-binding cassette (ABC) transporter genes involved in the transport of xenobiotics and metabolites. These pumps decrease the intracellular concentration of xenobiotics and are the final Phase of detoxification (Chin *et al.,* 1993*).*

As seen in Figure 4, the anti-porter activity in the intestine appears to be co-regulated with intestinal Phase I Cyp3A4 enzymes (Wacher *et al.,* 1995).

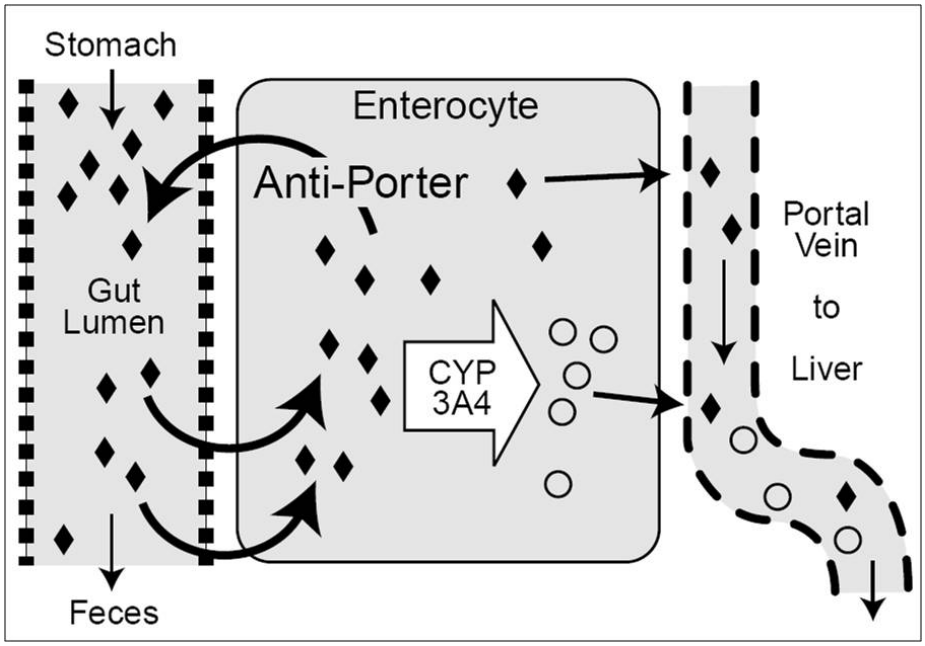



**Figure 4: Schematic illustration of the anti-porter function in Phase III detoxification pathway.** Toxins are transported from the stomach and intestines by enterocytes to the liver where the majority of detoxification reactions take place**.** The anti-porter activity in the intestine appears to be co-regulated with intestinal Phase I Cyp3A4 enzyme (Liska, 1998).

According to Liska (1998) this suggests that the anti-porter activity may support and promote detoxification. Because the xenobiotic is pumped out of the cell into the intestinal lumen for excretion before it is taken up into circulation for detoxification suggests that another cycle of detoxification might be possible (Liska, 1998).

### 1.5 Glycine N-acyltransferase

The remainder of this literature review will be devoted to amino acid conjugation of the Phase II detoxification system. The focus of this MSc project is on glycine N-acyltransferase which is a transferase enzyme involved in Phase II detoxification, specifically glycine conjugation. It was also mentioned before that that there are no GLYAT associated IEMs known, but glycine supplementation is commonly used in the treatment of IEMs. Another reason for investigating GLYAT is because the xenobiotic benzoic acid is widely used in food preservatives and we are constantly exposed to this xenobiotic in our diet. GLYAT is responsible to detoxify benzoic acid

into hippuric acid for excretion (Keller, 1842). This will be explained in more detail below.

Schachter and Taggart were the first scientists to begin research on glycine N-acytransferase in the early 1950s. They used pig and bovine liver for their studies (Schachter & Taggart, 1953). The systematic name of GLYAT is acyl-CoA: glycine N-acyltransferase. In the literature there is some confusion of nomenclature and GLYAT is known by various synonyms such as acyltransferase glycine; glycine acyltransferase: glycine-N-acylase and glycine N-acyltransferase (Brenda, 22 July 2007). The information was obtained from an internet database called Brenda: www.brenda-uni-koeln.de)

The enzyme of interest, GLYAT, is part of the GCN5-related N-acetyltransferase superfamily, also known as GNAT**.** All acetyltransferases use a common acetyl donor known as acetyl coenzyme A. This family of enzymes is the largest superfamily with over 10,000 members (Dyda *et al*., 2000). The first member of GNAT was identified in bacteria (Davis & Wright*,* 1997). Subsequently, some sequence homology was exhibited to a class of eukaryotes, the first being yeast in 1992 (Berger *et al.,* 1992). Despite the diversity in GNAT substrate specificities and little overall primary sequence homology due to environmental adaptation within the GNAT superfamily, the basic structure of the GNAT fold is well conserved in different organisms. The GNAT fold contains four amino acid motifs of approximately 100 - 120 amino acid residues that consist of an N-terminal strand followed by various beta sheets and alpha helices. Differences between GNAT fold structures are generally confined to the immediate N-terminus of exon 6, with much greater variation in the C terminus. The GNAT family of enzymes has a remarkable similarity in protein topology and mode of co-enzyme A binding. This may reflect a conserved catalytic mechanism of the GNAT fold (Dyda *et al.,* 2000). The GNAT fold has two universal functions: to bind the pantetheine arm of acyl-CoA and to hydrogen bond to and polarize the carbonyl of the thioester. The conservation of the 15 different GNAT fold structures is shown in Figure 5.

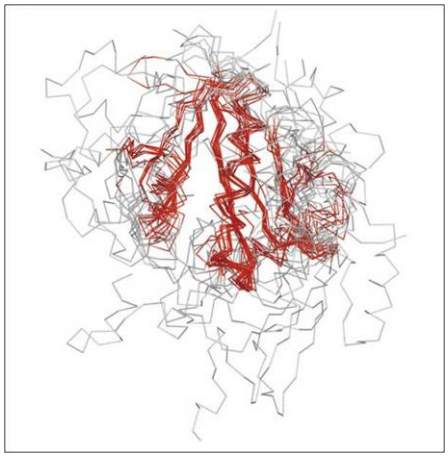


**Figure 5: Superposition of 15 GNAT structures** revealing that the topology of the GNAT fold containing 100 - 120 amino acid residues of an N-terminal strand followed by various beta sheets and alpha helices has nearly universal structure conservation. Differences between GNAT fold structures are generally confined to the immediate N-terminus, with much greater variation in the C-terminus (Vetting *et al.,* 2005).

In addition to the xenobiotic benzoic acid, GLYAT also detoxifies some short, medium, straight- or branched-chain acyl CoA substrates (Gregersen *et al.,* 1986). This reaction occurs mostly in the matrix space of the mitochondria of the liver and kidney and to a smaller extent in other tissues (Mawal & Qureshi, 1994). The detoxification of benzoic acid is shown in Figure 6.

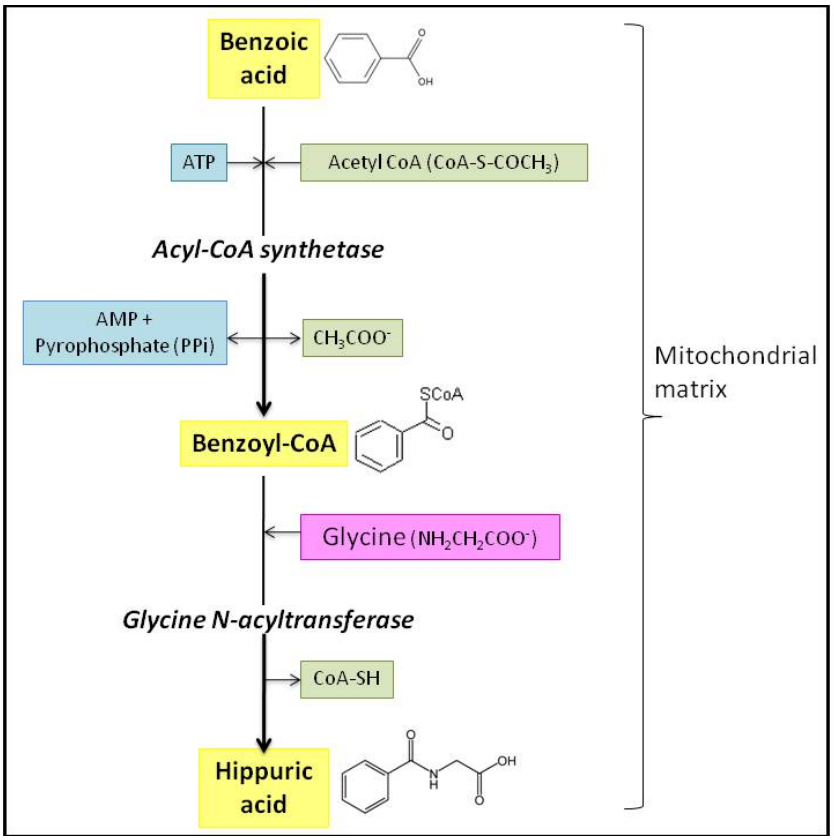



**Figure 6: Phase II detoxification pathway of the xenobiotic, benzoic acid, with glycine to form hippuric acid by means of two enzymes.** 1) acyl-coA synthetase to form benzoyl-CoA and 2) glycine N-acyltransferase to produce hippuric acid from benzoyl-CoA (Adapted from Parkinson, 1995).

The first step in benzoic acid detoxification is the activation of benzoic acid with coenzyme A in the presence of ATP. This is followed by the release of AMP and pyrophosphate and formation of benzoyl-CoA (Vessey *et al*, 1999). The enzyme responsible for this first step is Acyl-CoA-synthetase. Benzoyl-CoA is then conjugated with glycine in a two-step, acyl group transfer reaction to form benzoyl glycine (hippuric acid). Glycine acts as the acyl acceptor and acyl-CoA is subsequently released (Mawal & Qureshi, 1994). The acyl-CoA substrate (benzoyl-CoA) binds to glycine N-acyltransferase first, then glycine binds before the CoA leaves. The peptide dissociates last (Nandi *et al.,* 1979). The joining of the carboxylic acid to the amino nitrogen of the glycine is a sequential mechanism (Van der Westhuizen *et al,* 2000).

## 1.6 Comparison of the molecular and enzymatic characteristics of human and bovine Glycine N-acyltransferase

Although humans are closer to mice than bovine, bovine and humans have sufficient DNA sequence similarity to enable us to map the human genome almost entirely to the genome of the bovine (Wind *et al.,* 2005). Five GLYAT enzymes are of interest for this study namely bovine GLYAT (bGLYAT), human GLYAT isoform A (hGLYATa) and B (hGLYATb) as well as hGLYAT-like 1 (hGLYATL1) and hGLYAT-like 2 (hGLYATL2) (Nandi *et al.,* 1979; Vessey & Lau*,* 1996; Mawal & Qureshi, 1994; Zhang *et al.,* 2007; Waluk at al., 2010). The molecular and enzymatic characteristics of each will be explained in more detail in the next Section.

### 1.6.1 Molecular characteristics of glycine N-acyltransferase

The molecular characteristic of the gene, transcript and protein of these GLYAT enzymes mentioned above will be summarized as follows: Differences and similarities of the molecular characteristics of bovine GLYAT and the human GLYAT family known as hGLYAT(a) and hGLYAT(b) as well as hGLYATL1 and hGLYATL2 is shown in Table 1.

**Table 1: Differences and similarities of the molecular characteristics of bovine GLYAT and the human GLYAT family**

| | [a & b]*bGLYAT* | [b]*hGLYAT(a)* | *hGLYAT(b)* | [d]*hGLYATL1* | [e]*hGLYATL2* |
|---|---|---|---|---|---|
| **Chromosome** | 15 | 11 | 11 | 11 | 11 |
| **Gene length** | 9.55 Kb | 23.21 Kb | 21.89 Kb | 13.74 Kb | 10.72 |
| **Exons** | 5 | 6 | 5 | 7 | 6 |
| **Transcript length** | 984 bases | 2052 bases | 1147 bases | 2018 bases | 1636 bases |
| **GenBank Transcript number** | NM 177513 | NM 201648 | NM 005838 | NM 080661 | NM 145016 |
| **Translation length** | 295 res | 296 res | 163 res | 302 res | 294 res |
| **GenBank Translation number** | NP 803479 | NP 964011 | NP 005829 | NP 542392 | NP 659453 |
| **Molecular weight** | 33.77 kDa | 33.90 kDa | 18.51 kDa | 35.1 kDa | 34.28 kDa |

[a]Nandi *et al.,* 1979; [a]Vessey *& Lau,* 1996; [a-b]Mawal & Qureshi, 1994; [d]Zhang *et al.,* 2007; [e]Waluk at al., 2010 and Ensembl Genome Browser, 7 January 2010. The data was obtained from an internet database called Ensembl Genome Browser: www.ensemble.org.

In the middle 1970s, Webster and his colleagues investigated the enzymatic properties and differences between human and bovine GLYAT (Webster *et al*., 1976). Glycine N-acyltransferase consists of a single polypeptide chain (Nandi *et al*., 1979). Bovines have 60 chromosomes, of which one pair is sex chromosomes. The

bovine GLYAT gene is found on chromosome 15 and consists of five exons. The length of the gene is 9.55 kb. The nucleic acid sequence of mRNA encoding bovine GLYAT was determined by Vessey and Lau in 1996. The transcript length of bGLYAT is 984 bases with an open reading frame (ORF) encoding a polypeptide of 295 amino acid residues (Vessey & Lau, 1996). In 1979 Nandi determined the molecular weight of bovine GLYAT at 33 kDa, whereas Kelley and Vessey (1992) and Van der Westhuizen (2000) reported it to be 33.5 kDa and 36 kDa respectively (Nandi *et al.,* 1979; Kelley & Vessey, 1992; Van der Westhuizen *et al.,* 2000).

Humans have 46 chromosomes, of which one pair is sex chromosomes. Human GLYAT found on chromosome 11 has two splice variants; isoform A (hGLYATa) and B (hGLYATb). Isoform A consists of six exons with the gene length of 23.21 kb. The transcript length is 2052 bases and the translation length of the ORF is 296 amino acid residues. Isoform B has five exons with a gene length of 21.89 kb. The transcript length is 1147 bases and the translation length of the protein 163 amino acids. The difference between the two isoforms is that hGLYATb does not have the sixth exon like hGLYATa, thereby missing amino acids residues 183 to 296. It is therefore a splice variant of the mRNA. Mawal and Qureshi (1994) and Van der Westhuizen (2000) noted that the human GLYAT has a molecular weight of 30 kDa and 27 kDa respectively for isoform A (Mawal & Qureshi, 1994; Van der Westhuizen *et al.,* 2000).

In 2007, Zhang and his colleagues found a novel gene on chromosome 11 encoding a human GLYAT family member known as hGLYATL1. hGLYATL1 was expressed in mammalian cells (HEK293T cells). The hGLYATL1 gene consists of seven exons (Zhang *et al.,* 2007). The transcript length of hGLYATL1 is 2018 bases which encodes an ORF of 302 amino acid residues. The molecular weight is 33 kDa. The hGLYATL1 protein was found to activate the transcriptional activities of the heat shock element (HSE) signaling pathway. This suggests a potential role of this protein in HSE mediated transcriptional regulation. Zhang and his colleagues (2007) mapped the hGLYATL1 gene to the human chromosome 11q12.1 with two other members of the GLYAT family (Figure 7).

Recently, Waluk and colleagues (2010) identified human GLYAT-like 2 (hGLYATL2), a member of the human GLYAT family. This enzyme is found in the endoplasmic reticulum and produces N-acyl glycines. Interest in glycine conjugation of fatty acids (N-acyl glycines) has recently been renewed. N–acyl glycines are now emerging as a potent lipid signaling molecule and activities include antinociceptive, anti-inflammatory and antiproliferative effects (Waluk *et al.,* 2010). The hGLYATL2 is also located on chromosome 11 and has 6 exons, with a transcript length of 1636 bases. It encodes 294 aa and encodes a protein of 34 kDa. GLYATL2 was expressed in bacterial (BL21(DE3)pLysS) cells, purified and was enzymatically active. In 2010 Waluk *et al*. identified GLYATL2 as a transferase that produces N-acyl glycines in humans (Waluk et al., 2010).

Zhang and his colleagues mapped the *hGLYATL1* gene to the human chromosome 11q12.1. The gene *GLYAT and GLYATL2* is situated adjacent to the location (Figure 7).

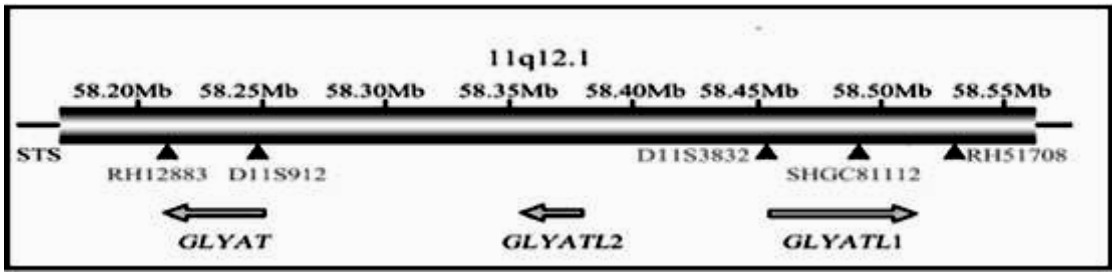



**Figure 7: The chromosomal localization and genomic organization of the genes *GLYAT*, *GLYATL1* and *GLYATL2* on human chromosome 11q12.1. The grey hollow arrows represent the transcriptional direction** (Zhang *et al.,* 2007).

There are no other genes between the three *GLYAT* genes on human chromosome 11q12.1 known (Figure 7). This suggests that the gene family might have some other unidentified members besides the *GLYAT, hGLYATL1, hGLYATL2* or that most of the members belonging to this family were cluster distributed in the 11q12.1 segment (Zhang *et al.,* 2007).

As mentioned in Section 1.5, Figure 5, the structure of the GNAT fold is generally confined to the immediate N-terminus, with much greater variation in the C-terminus.

In Figure 8, the amino acid sequence human and bovine GLYAT was aligned to hGLYAT isoform A and B, as well as hGLYATL1. The hGLYATL2 was not aligned.

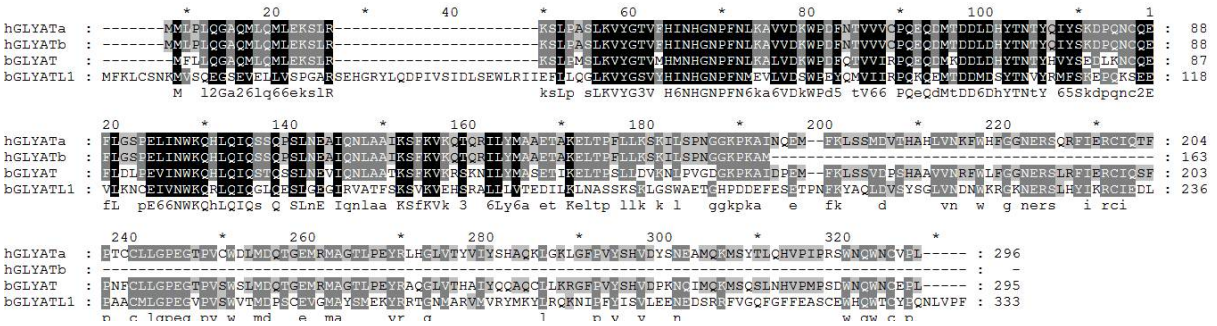


**Figure 8: Comparison of different amino acid sequences of** GLYAT obtained from the National Centre of Biotechnology Information website, NCBI. Comparing GLYAT of bovine (NM 177513), and human isoform A (NM 201648), B (NM 005838) and hGLYATL1 (NM 080661).

The amino acid residues in highlighted black are the same in all the amino acid sequences. Amino acid residues highlighted in shaded grey is partially the same for the sequences. Figure 8 shows that the N-terminus is more conserved and further along the sequence, much more variation on can be seen towards the C-terminus. With the analysis of the amino acid sequence done by Van der Westhuizen in his PhD study, human and bovine GLYAT do not reveal recognizable mitochondria target signal. It is though that mitochondria targeting may occur via the cytoplasm through a non-conservative pathway (Van der Westhuizen *et al*., 2000). There are in total 11 nucleic variations known in the bovine GLYAT. Seven are known as synonymous single nucleotide polymorphisms (SNPs) that do not affect the amino acid sequence. The other four SNPs are non-synonymous and result in an amino acid change in the encoded protein. The positions of this variation occur between amino acid residues 126 and 176. (Ensemble Genome Browser, August 2009). The human GLYAT SNPs will not be discussed as the focus is on bovine GLYAT.

### 1.6.2 Enzymatic characteristics and species specificity of glycine N-acyltransferase

The focus of this literature study will now shift to the enzymatic function and species specificities of human and bovine GLYAT. The pH for optimal GLYAT activity for both bovine and human GLYAT is 8.4, but can tolerate pH increase up to 8.6 before starting to lose enzyme activity (Nandi *et al.,* 1979; Webster *et al*., 1976).

In general, enzymes are specific both for the type of reaction that they catalyze and for a single substrate or a small set of closely related substrates. GLYAT shows species specificity with respect to the acyl and amino acid conjugation *(*Nandi *et al.,* 1979). There are some similarities and differences between bovine and human GLYAT enzyme activity and substrate specificity. Nandi *et al* (1979) reported that bovine amino acid:N-acyl transferase utilizes glycine as the preferred co-substrate, but also glutamine and asparagine, albeit the latter conjugation occurs at much slower rates. Van der Westhuizen *et al.* (2000) used the DTNB reduction enzyme activity assay according to Mawal and Qureshi (1994) in order to determine the rate of benzoyl-amino acid conjugation with bovine and human GLYAT (Mawal & Qureshi, 1994).

Even though the amino acid utilization of glycine N-acyltransferase with glycine, glutamine and asparagine has been investigated before, Van der Westhuizen *et al.* (2000) investigated the utilization of alanine, glutamic acid and serine as amino acid substrates for glycine N-acyltransferase as well. Benzoyl-CoA was the acyl donor in the kinetic study as it is the best characterized acyl-CoA substrate for GLYAT. The results of the kinetic parameters of bovine and human GLYAT for the formation of benzoyl-amino acids are indicated in Table 2.

**Table 2: Kinetic parameters of bovine and human GLYAT for the formation of benzoyl-amino acids**

| | *Km-$_{benzoyl\text{-}CoA}$(mM)* | | *Km-$_{amino\ acid\ (}$mM)* | | *Vmax (µmol/min/mg)* | |
|---|---|---|---|---|---|---|
| **Benzoyl-amino acid** | **Bovine** | **Human** | **Bovine** | **Human** | **Bovine** | **Human** |
| **Benzoyl-glycine** | 160 | 13 | **6.2** | **6.4** | 28.2 | 0.7 |
| **Benzoyl-glutamine** | 105 | - | 353 | - | 2.3 | - |
| **Benzoyl-asparagine** | 157 | - | 129 | - | 6.7 | - |
| **Benzoyl-alanin** | 41 | 15.2 | 1573 | 997 | 1.6 | 0.3 |
| **Benzoyl-glutamic acid** | 998 | nd | 1148 | nd | 0.1 | nd |

nd: Not detectable because of extremely low amount of conjugates (Van der Westhuizen *et al.,* 2000)

Km values in bold indicate preferred amino acid for benzoyl-CoA conjugation of glycine N-acyltransferase.

Van der Westhuizen and co-workers (2000) found that bovine GLYAT can utilize glycine, asparagine, glutamine as well as alanine, glutamic acid and serine (which is not mentioned in Table 2. Bovine GLYAT catalyze the conjugation of benzoyl-CoA

with selected amino acids with Km values as follows: glycine ($Km_{Gly}$ = 6.2 mM), asparagine ($Km_{Asp}$ = 129 mM), glutamine ($Km_{Gln}$ = 353 mM), alanine ($Km_{Ala}$ = 1573 mM), glutamic acid ($Km_{Glu}$ = 1148 mM). Only the benzoyl conjugation of glycine, alanine and glutamine were detected for human GLYAT. For human GLYAT, onjugation only occurred with glycine ($Km_{Gly}$ = 6.4 mM) and alanine ($Km_{Ala}$ = 997 mM) (Van der Westhuizen *et al.,* 2000). The values could not accurately be determined for serine for both human and bovine GLYAT, as well as bGLYAT or benzoyl-glutamic acid of hGLYAT and are, therefore, not listed in Table 2. The amount of these conjugates was extremely low and barely detectable at low substrate concentrations.

Bovine GLYAT is more specific for glycine conjugation compared to human GLYAT (Van der Westhuizen *et al*, 2000). A low Km value means that GLYAT has a high affinity for the substrate. The Km and Vmax-values were determined from linear primary and secondary rate plots (Van der Westhuizen *et al.,* 2000; results not shown). Glycine is the fastest utilized amino acid substrate for humans and bovines with a very low conjugation rate compared to other amino acids. This is indicated by the lower Km values and higher reaction velocities compared to the other amino acids. Alanine has very low conjugation rates with benzoyl-CoA relative to glycine (Van der Westhuizen *et al.,* 2000). According to Van der Westhuizen *et al* (2000), when any of the other amino acids were present, the enzyme activity of GLYAT was not inhibited (Van der Westhuizen *et al.,* 2000).

Reports from scientists over the years on the methods used for testing the enzyme activity of GLYAT vary greatly. For instance, James and Bend (1978) developed a radiochemical assay for GLYAT activity in rats and rabbits *in vivo* (James & Bend, 1978), whereas Mawal and Qureshi (1994) and Van der Westhuizen and co-workers (2000) used the GLYAT enzyme activity assay based on the method developed by Kolvraa and Gregersen in 1986 (Mawal & Qureshi, 1994; Van der Westhuizen *et al*., 2000; Kolvraa & Gregersen, 1986). It is, therefore, not really possible to accurately compare the reports in the literature of the enzyme activity of bovine GLYAT with human GLYAT.

Various ethical considerations have to be taken into account when doing research on humans and human tissue. Animal samples such as bovine tissue are more readily available and can be obtained from any abattoir. It is therefore useful to have an

animal model for research on human GLYAT. Bovine GLYAT is the most thoroughly studied mammalian GLYAT enzyme to date. Studies have also been done on other species such as monkeys (Webster *et al.,* 1976) and rats and rabbits were also investigated (James & Bend, 1978; Kolvraa & Gregersen, 1986). The bovine GLYAT is well compared to the human GLYAT with regards to molecular and enzyme characteristics as mentioned in the previous sections in order to get insight into the detoxification process that GLYAT is involved in.

## 1.7 Detoxification profiling

Over the years, an increased interest in genetic regulation involved in drug metabolising enzymes and drug transporters gave an understanding of the molecular mechanism of drug response and toxic events (Nakata *et al,* 2006). Detoxification has become a popular topic and it is well established to the point where detoxification profiling is becoming available as a commercial service. An example of a detoxification company is Genovations. They do a Detoxi Genomic Profile to identify gene variations that may affect one's ability to detoxify specific toxins, medication or even food. Polymorphisms such as single nucleotide popymorphisms (SNPs) in the gene coding for a particular enzyme can increase, or even more commonly decrease the activity of the enzyme. For instance, some SNPs in specific CYP enzymes can cause increases in Phase I detoxification activity. If Phase II detoxification reactions are not capable to process the Phase I compound, this may lead to toxins accumulating in the body. On the other hand, a decrease in Phase II detoxification activity may lead to lower capability to eliminate specific pharmaceutical drugs from the body. Genovations developed a personalized treatment strategy to match each individual's detoxification in order to live a better and healthier life. The Centrum for Human Metabolomics of the subject group Biochemistry, School for Physical and Chemical Science at the North-West University is also already providing a pilot biochemical and oxidative stress profiling service and recommending lifestyle and diet changes to improve the quality of life of people. This is achieved by doing loading tests with caffeine or aspirin to investigate the efficacy of the Phase I and II detoxification of individuals. Some of these individuals also participate in a pilot genetic profiling system.

## 1.8 Problem Formulation

With the increasing levels of toxins that we are exposed to every day such as food components, preservatives, pharmaceuticals and environmental factors, we need a better understanding of how these substances are detoxified in the human body. Detoxification has become a popular well known topic affecting quality of life of a person, or as a general therapy in the management of IEMs. The preservative benzoic acid is a toxin we are daily exposed to in the food component we eat. GLYAT is a Phase II detoxification enzyme responsible to detoxify benzoic acid. Investigating this enzyme might give insight into detoxification.

Previously the staff of the Biochemistry Department at the North-West University had initiated research projects on human and bovine GLYAT. Molecular characteristics and the enzyme function were compared (Van der Westhuizen *et al.,* 2000). This was done on an enzyme purified from bovine liver. We would now like to further investigate the structure and function relationship of GLYAT between individuals and species. For this we need to know the genetic variation and generate a recombinant GLYAT as a backbone to eventually probe the specificity of amino acid residues by site-directed mutagenesis. Thereby, we can start to analyse the impact of individual SNPs on enzyme activity and amino acid specificity.

Bovine GLYAT is the most studied GLYAT. There is a small difference between the enzyme activity and substrate specificity of human and bovine GLYAT. There is no recombinant bovine GLYAT or human GLYAT isoform A and B available. So far, the only recombinant GLYAT enzymes are human hGLYATL1 (Zhang *et al*., 2007) and GLYATL2 (Waluk *et al.,* 2010). hGLYATL1 was expressed in mammalian cells (HEK293T cells). GLYATL2 was expressed in bacterial (BL21(DE3)pLysS) cells, purified and was enzymatically active.

## 1.9 Objectives of this investigation

1. To prepare the mRNA encoding Glycine N-acyltransferase from bovine liver;
2. To generate a recombinant bGLYAT;
3. To verify if the recombinant bGLYAT exhibits enzyme activity.

# CHAPTER 2

# CLONING OF THE mRNA ENCODING GLYCINE N-ACYL-TRANSFERASE FROM BOVINE LIVER AND EXPRESSING RECOMBINANT bGLYAT WITH THIRD GENERATION EXPRESSION VECTOR pCOLDTF

## 2.1 Introduction

In our laboratory we have a focus on glycine N-acyltransferase. Several pilot projects aimed at generating a recombinant GLYAT which is enzymatically active have been done under the supervision of Prof AA van Dijk at the North-West University, Biochemistry Division. The first generation bacterial expression systems were investigated by Mr DA Grundling. He used the pET32a+ and pGEX vectors. Mr Grundling found that the recombinant human GLYAT was over-expressed using both vectors, but it was insoluble and had no GLYAT enzyme activity (personal communication). Problems such as aggregation and lack of biological function of the recombinant protein are one of the biggest obstacles in the field of biotechnology. This is usually due to the lack of post-transcriptional modification such as glycosylation and phosphorylation. In this study I investigated whether the use of third generation expression system would yield soluble, enzymatically active bovine glycine N-acyltransferase.

In the three decades since the advent of recombinant DNA technology, over-production of eukaryotic proteins through various expression systems has been explored in many ways. Prokaryotic expression systems, especially *Escherichia coli (E. coli),* are the preferred host for the expression of recombinant proteins simply because of their inexpensive carbon source requirements for growth and their ability to rapidly accumulate biomass. They are also amenable to high-cell density fermentation and require simple process scale-up technologies. The expression of proteins in bacteria can also be regulated. The most important use for recombinant protein expression in bacteria is for the production of proteins on an industrial scale such as the production of vaccines (Sahdev *et al.,* 2008).

Post-translational modification involved in protein folding is a complex and integrated process associated with chaperones. When a protein is mis-folded, it will be degraded by cellular proteases, a reaction called proteolysis. Aggregation of the proteins can also form inclusion bodies that are biologically inactive (Thomas *et al.,* 1997). Figure 9 shows a network of molecular chaperones assisting with *in vivo* protein folding which is an energy-dependent process.

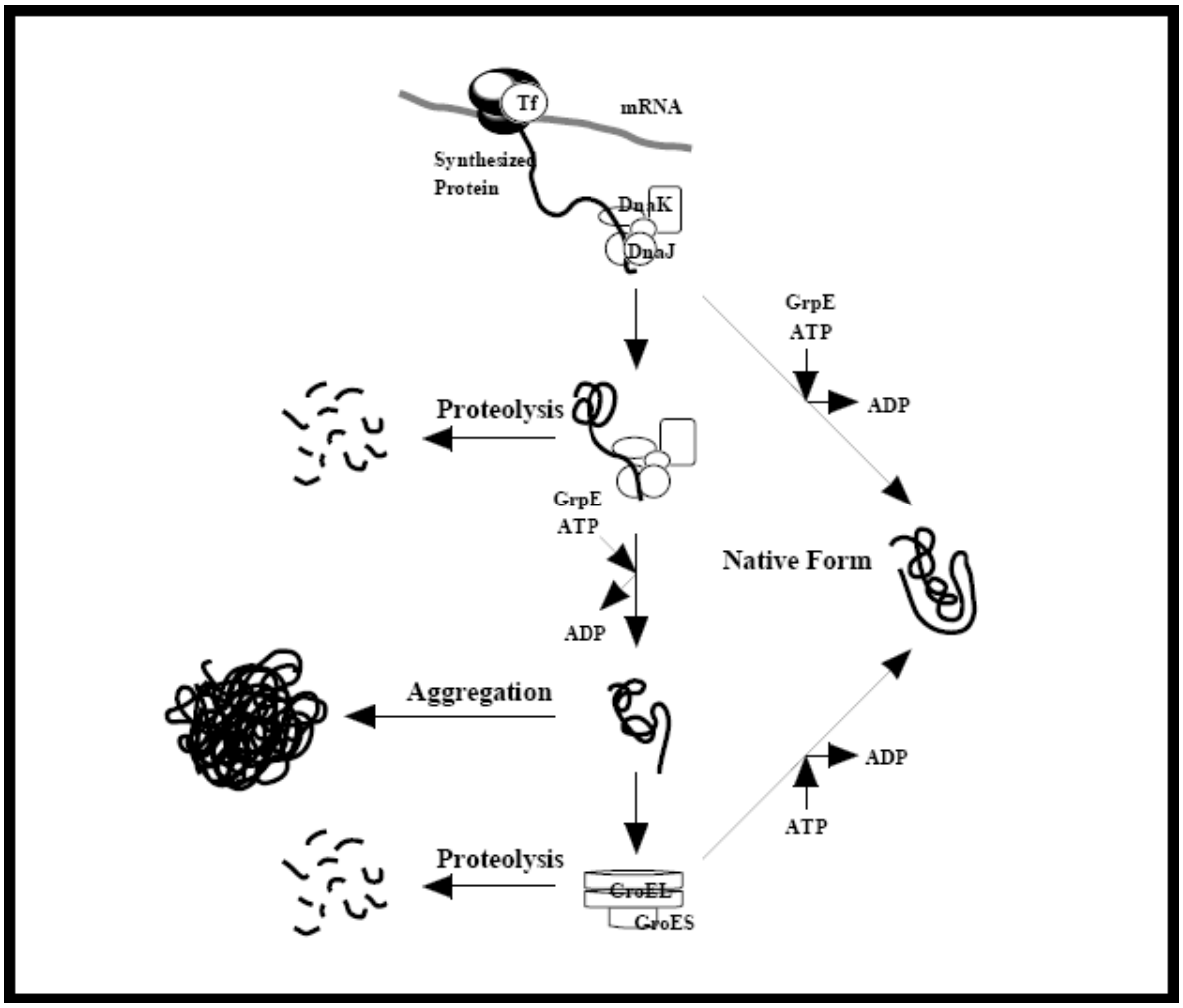



**Figure 9: Possible model for chaperone assisted protein folding in *E. coli.*** Molecular chaperones systems, such as DnaK-DnaJ-GrpE and GroEL-GroES suppress off-pathway aggregation or proteolysis reactions to facilitate proper folding (Thomas *et al.,* 1997).

The classes of folding modulators mediating this process are molecular chaperones systems, such as trigger factor (TF), DnaK-DnaJ-GrpE and GroEL-GroES. These systems suppress off-pathway aggregation reactions and facilitate proper folding. Purified folding modulators and artificial systems such as plasmids expressing these chaperones that mimic the mode of action have proven useful in improving the *in vitro* refolding yields of chemically denatured polypeptides (Thomas *et al.,* 1997).

The biggest problem in bacterial expression of recombinant eukaryotic proteins is the lack of post-translational modification and aggregation. Recombinant eukaryotic proteins expressed in *E. coli* often aggregate or degrade rapidly because of their inability to adapt their biologically active conformation due to problems in protein folding. This is often due to the fact that appropriate and /or enough chaperones are not available to help with the folding of the protein (Thomas *et al.,* 1997).

In 2000, Nishihara and colleagues investigated whether over-expression of trigger factor (TF) would prevent aggregation of recombinant proteins in *E. coli* by means of a series of different plasmid constructs (Nishihara *et al.,* 2000). Trigger factor is a chaperone-like 50 kDa protein (Stoller *et al.,* 1995). It acts from the start of protein synthesis to assist with protein folding by associating with the nascent protein and the 50 S ribosomal subunit (Valent *et al*., 1995). TF arches over the ribosomal exit tunnel accepting the nascent peptide chain in a protective void. The protein remains in this protective cradle until folding is completed. Trigger factor directs the nascent chain through the interior of TF in a sequential and length-determined manner (Merz *et al.,* 2008). Figure 10 shows a model of passage of nascent polypeptide chain through TF.

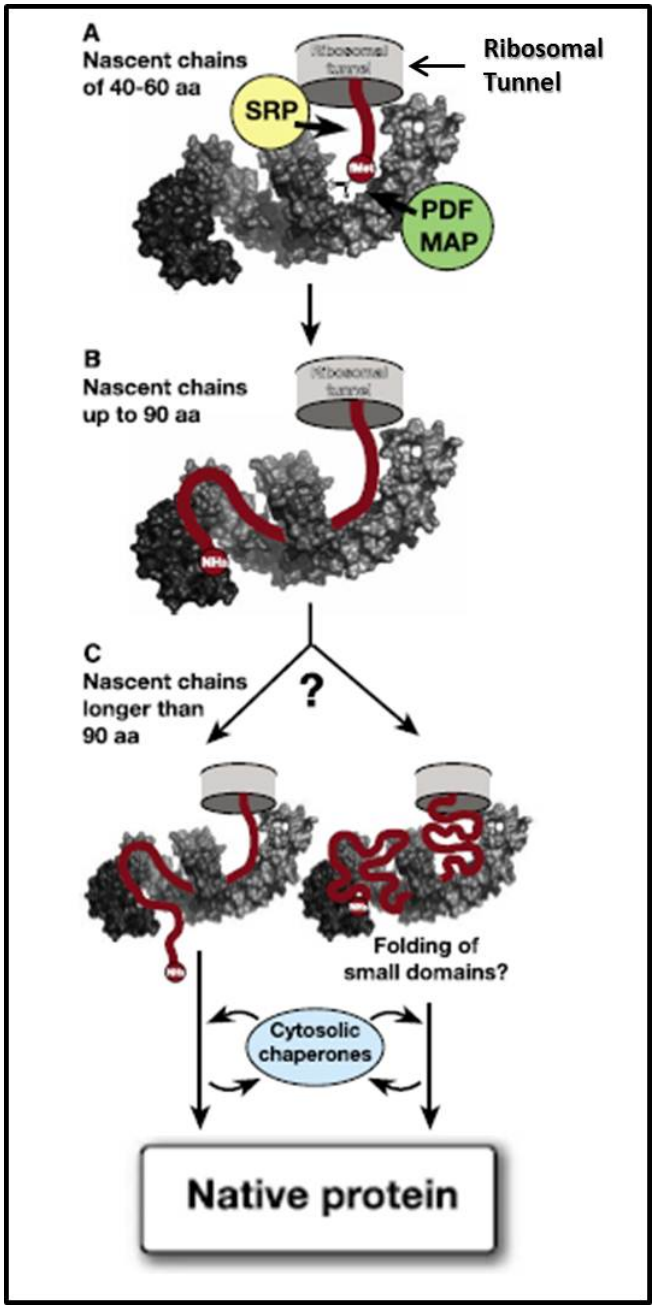



**Figure 10: Model of passage of nascent polypeptide chain through TF.** TF directs the nascent chains through its interior in a sequential and length-dependent manner. (A) Initially (with a length of 40–60 aa), the N terminus of the nascent chain slides along the N domain, where it might be accessible by means of the lateral openings from both sides for processing factors such as protein deformylase (PDF), N-terminal methionine by aminopeptidases (MAP) or signal recognition particle (SRP). (B) Up to a length of 90 aa, the nascent chain traverses through the C-terminal arms towards the PPIase domain and engages the entire interior. (C) Upon further elongation, the nascent chain might leave TF or, alternatively, may accumulate and perhaps fold in the interior of the TF chaperone. On demand, the folding of a subset of newly synthesized proteins is further assisted by cytosolic chaperones (Taken from Merz *et al.,* 2008).

The approach was to investigate whether co-expression of molecular chaperones such as TF and GroEL-GroES would facilitate protein folding and result in the production of an enzymatically active bovine GLYAT.

The commercial company, Takara collaborated with Professor Massayori Inouye (University of Medicine and Dentistry of New Jersey, USA) to develop pCold DNA vectors in an effort to solve problems with the expression of insoluble proteins. Figure 11 illustrates the ORF encoding TF that is incorporated into one of their cold-shock expression vectors called pColdTF (Takara, catalogue number 3365). In many cases, expression of genes using pColdTF resulted in soluble recombinant proteins.

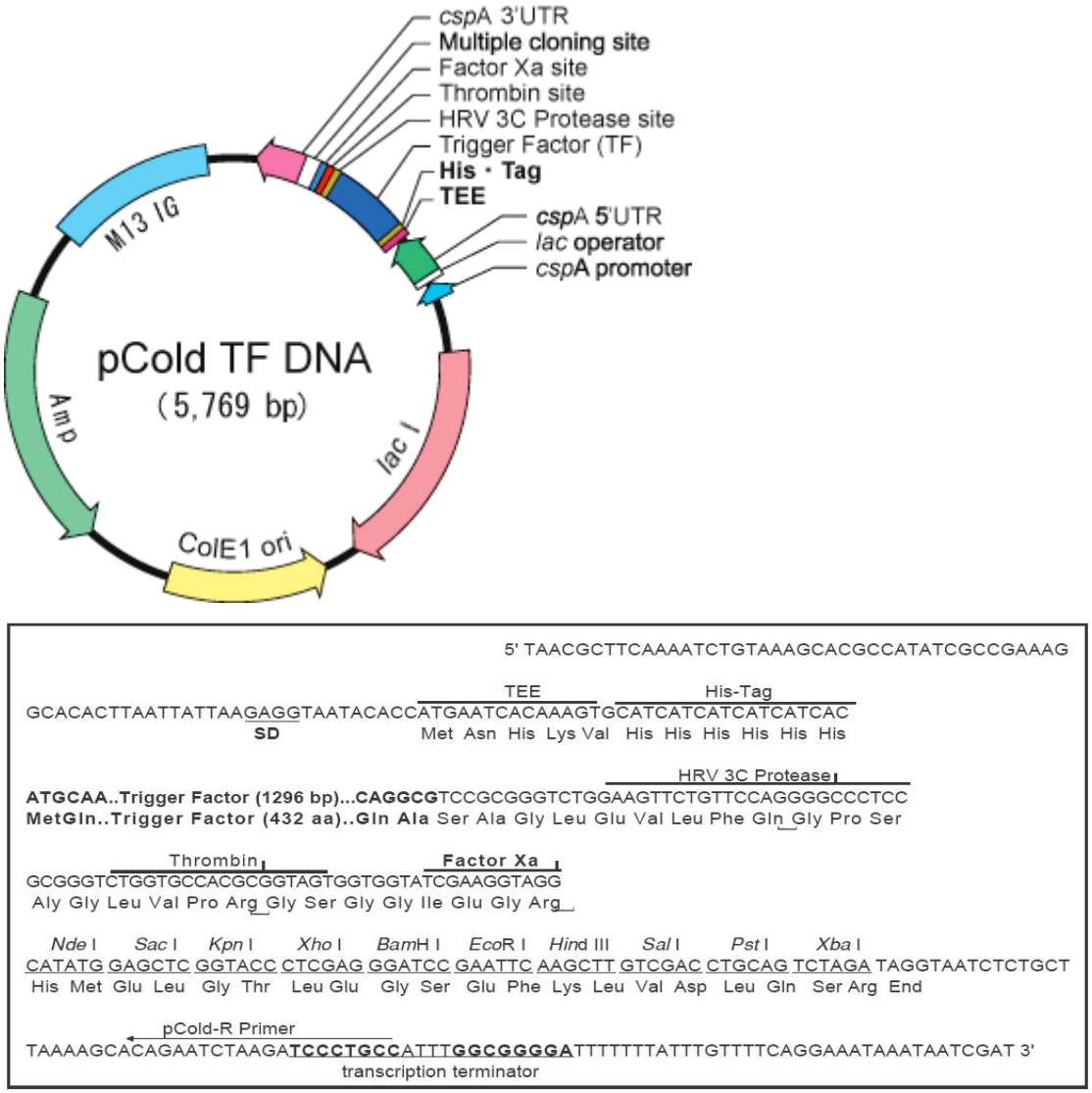



**Figure 11: Plasmid map and multiple cloning site of pColdTF.** pColdTF contains the multiple cloning site, TEE, His tag, three protease cleavage sites (HRV 3C Protease, Thrombin and Factor Xa), the 3' and 5' UTR *cspA* gene, *cspA* promotor and the *lacI* operator (obtained from Takara Bio Inc, 2009).

Cold-shock expression vectors are nowadays widely used for protein expression. They are based on using the promoter of the gene *csp*A, which is induced by low-temperatures in *E. coli* cultures to drive expression. At lower temperatures, the rate of native protein synthesis decreases which facilitates better folding. By incorporating the cspA promoter in the vector, one can induce the expression of a recombinant protein by lowering the temperature while expression of housekeeping genes slows down. The pCold vector also contains a *lac*I operator that functions to strictly regulate the expression of recombinant proteins by means of IPTG induction. IPTG is an

inducer of β-galactosidase activity in many bacteria. Functioning as a *lacI* analog, IPTG induces β-galactosidase activity by binding to and inhibiting the *lac* repressor. If the *E. coli* culture temperature is reduced, the bacterial growth is temporarily halted and *E. coli* protein expression decreases. Inhibiting the Lac repressor by IPTG allows for target gene expression. Recombinant proteins can then be over-expressed by up to 60%. The incorporation of the *lacI* operator and the cspA gene is an excellent component in bacterial expression because a wide range of *E. coli* subspecies can be used as hosts. The pCold vectors also contain a transcription enhancement element (TEE).

Features of pColdTF other than those stated above used for expression is the histidine tag fused to the recombinant protein and the protease cleavage sites for purification of recombinant proteins (Factor Xa, Thrombin and HRV 3C) for removal of the fusion protein after expression.

The experimental approach for the work presented in this chapter was as follows: The RNA in a sample of bovine liver was first stabilised with RNAlater™. The tissue was then homogenised to extract total RNA and cDNA was synthesised from total RNA extracted. Polymerase chain reaction (PCR) was used to generate a bGLYAT amplicon with specific bovine GLYAT primers. This amplican was then inserted into the pColdTF vector by means of directional cloning. Plasmids were isolated and sequenced to verify the open readfing frame of the gene encoding bGLYAT. The recombinant proteins were expressed by means of cold shock and IPTG induction from the pColdTF plasmid. The recombinant bGLYAT containing a histidine tag was purified with nickel column chromatography. The protein fraction was desalted, concentrated and then used in a GLYAT enzyme activity assay. The fusion protein was removed from bGLYAT by means of thrombin cleavage and bGLYAT was again desalted, concentrated and used in a GLYAT enzyme assay. A flow diagram of the experimental approach is seen in Figure 12.

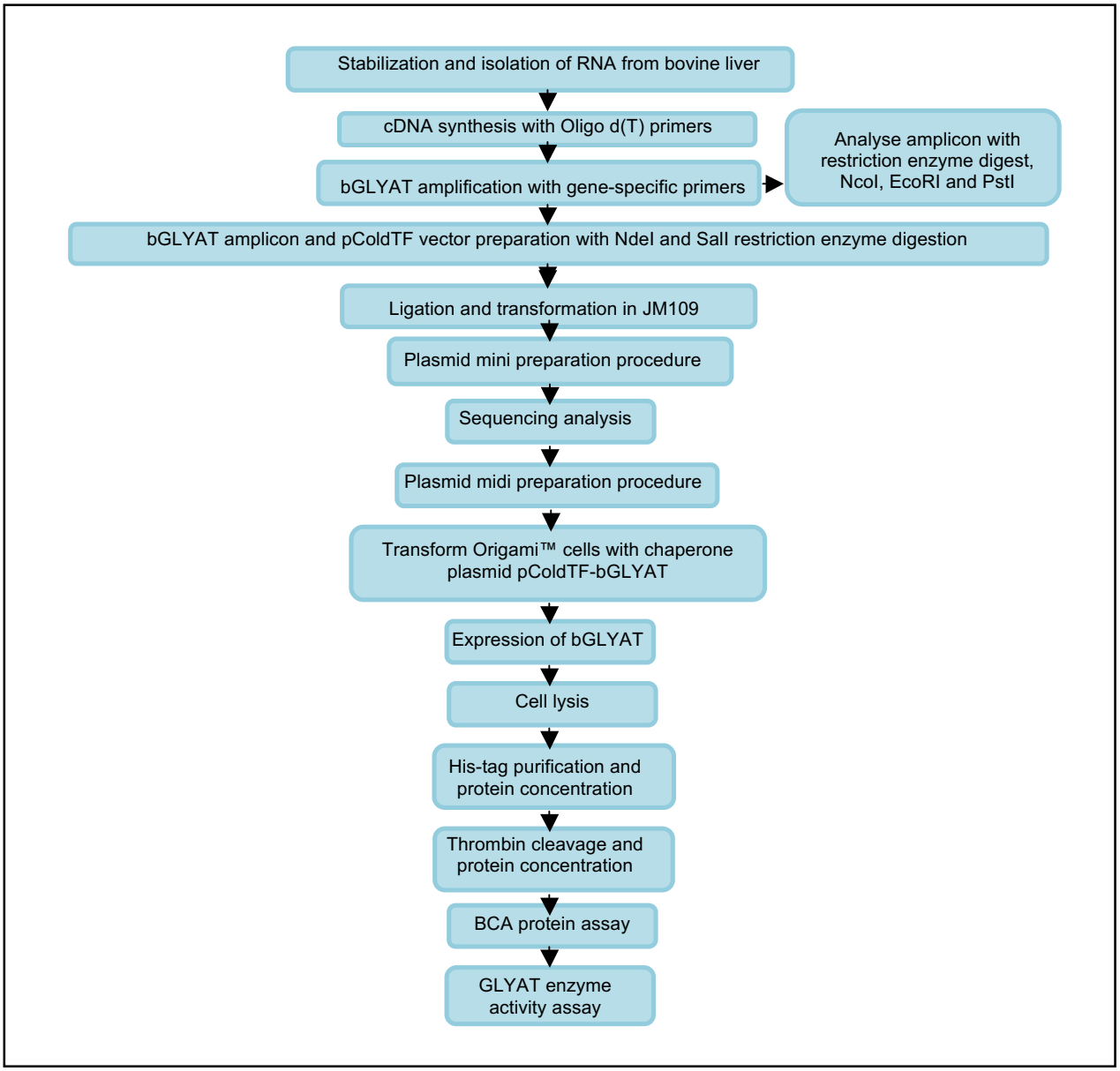



**Figure 12: Experimental approach in order to clone the gene encoding Glycine N-acyltransferase from bovine liver and to express the recombinant enzyme.**

## 2.2 Material and methods

### 2.2.1 Extraction of total RNA from bovine liver

Single-stranded RNA (ssRNA) is not very stable in biological samples because directly after harvesting, specific and non-specific RNA degradation occurs due to the presence of nucleases. RNases, nucleases specific for RNA, are abundant in all tissue samples and degrade ssRNA (Qiagen: RNAlater™™ Handbook, 2002). Therefore, RNA (including mRNA) needs to be preserved for use in further molecular biology applications such as cDNA synthesis (Garrett & Grisham, 2005).

A portion of bovine liver of about 500 g was obtained from the Potchefstroom Abattoir. Pieces of approximately 150 mg tissue were immediately immersed in 1.5 ml RNAlater™ Reagent and stored at –80°C. RNAlater™ is an aqueous tissue storage reagent that rapidly permeates most tissues. It eliminates the need to immediately process or freeze samples. Tissue size is critical for successful RNA stabilisation so that the reagent can penetrate the tissue to protect the cellular RNA. RNA stabilization was done according to the instructions of the manufacturer: RNAlater™ Handbook for stabilisation and protection of RNA in tissue (Qiagen, catalogue number 7021). Qiagen does not disclose the composition of RNAlater™ since it gives them a commercial competitive advantage.

Total RNA was isolated from bovine liver with the commercial RNeasy® Mini kit suitable for isolation of RNA from a wide variety of samples (Qiagen, catalogue number 74104) following the instructions of the manufacturers; “Isolation of total RNA from animal tissues”. The method is based on silica-gel-membrane technology combined with microspin technology. RNA binds to the silica beads at high salt concentrations and elutes at low salt concentrations. Up to 100 µg of RNA (longer than 200 bases) binds to the RNeasy® silica-gel membrane. The average yield obtained with the RNeasy® Mini kits with mouse liver tissue is 15 µg. Most RNAs smaller than 200 nucleotides (such as 5.8S rRNA, 5S rRNA, and tRNAs which together comprise 15–20% of total RNA) are excluded and washed away (RNeasy® Mini Handbook, 2003). Therefore, this kit provides enrichment for mRNA for downstream applications.

The RNAlater™ treated frozen bovine liver was thawed and 30 mg starting material was weighed and used for total RNA extraction. The liver was cut into smaller pieces before homogenization and lysis in the presence of a highly denaturing guanidine isothiocyanate (GITC)-containing RLT (lysis) buffer. The composition of the reagents is not available from the company. This buffer immediately inactivates RNases® to ensure isolation of intact RNA. The sample was homogenised in the lysis buffer, Rin a Heidolph bench top homogeniser for 90 sec in a 2 ml micro centrifuge tube to shear the genomic DNA and achieve sample viscosity. Addition of 70% ethanol created conditions that promote selective binding of RNA to the RNeasy® MinElute membrane. RNA binds to the column under high salt conditions while contaminants were efficiently washed away. RNA can be released from the column with a buffer containing a low concentration of salt or eluted in 50 µl water. All centrifugation steps were done for 15 seconds at 8000 x g in an Eppendorf bench top centrifuge.

### 2.2.2 Spectrometric analysis of nucleic acids

The Nanodrop® ND-1000 Spectrophotometer and software supplied with the equipment combines fibre optic technology and surface tension properties of liquid samples to allow micro-volume quantitation and analysis of 0.5 µl - 2.0 ml samples without the need for cuvettes or capillaries. The UV wavelength ranges from 220 nm to 750 nm. The Nanodrop® ND-1000 Spectrophotometer was used to measure the yield and the purity of the RNA and DNA preparations in 2 µl volumes. Optical density was measured at 260 nm and 280 nm. The absorption maximum of nucleic acid such as RNA and DNA is 260 nm. The $A_{260/280}$ nm ratio reflects the purity of the nucleic acid since nucleic acid preparations often contain protein contaminants. $A_{260/280}$ ratio of 2 implies pure nucleic acid samples which do not contain any proteins (Sambrook & Russel, 2005). For a ssRNA solution, one optical density at 260 nm ($OD_{260}$) corresponds to 40 ng/µl. The concentration of ssRNA can be calculated by the following equation (Sambrook & Russel, 2005).

$$\text{[single-stranded RNA]} = A_{260} \text{ X (40 ng/µl X dilution factor)}$$

**Equation 1: Equation to calculate concentration of single-stranded RNA**

For a double-stranded DNA (dsDNA) preparation, one $OD_{260}$ corresponds to 50 ng/µl. The following equation was used to determine the concentration of DNA (Sambrook & Russel, 2005).

**[double-stranded DNA] = $A_{260}$ X (50 ng/µl X dilution factor)**

**Equation 2: Equation to calculate concentration of double-stranded DNA**

The total yield of nucleic acids obtained was calculated by the following equation (Sambrook & Russel, 2005).

**Total yield = concentration (ng/µl X volume of sample (µl)**

**Equation 3: Equation to calculate the total yield of nucleic acids**

### 2.2.3 Gel electrophoresis

Since nucleic acids have a negatively charged sugar phosphate backbone, they migrate to the anode during electrophoresis when an electric current is applied. Agarose gel electrophoresis is the easiest and most common way of separating and analyzing nucleic acids due to their charge migration. The DNA and RNA are commonly visualised in the gel by adding ethidium bromide (EtBr) that binds strongly with the nucleic acid backbone by intercalating between the bases. EtBr is fluorescent and absorbs invisible UV light (Sambrook & Russel, 2005).

#### 2.2.3.1 Denaturing formaldehyde agarose gel electrophoresis

A 1.2% formaldehyde agarose gel (FA agarose gel) was made to visualise the integrity and size distribution of total RNA purified from bovine liver according to the instructions of Qiagen; Protocol for Formaldehyde Agarose Gel Electrophoresis. Formaldehyde (FA) is a denaturing reagent which relaxes the complex secondary structure of RNA. The respective 18 S and 28 S ribosomal RNA should appear as sharp bands on the stained gel under UV light. 28 S ribosomal RNA bands should appear at approximately twice the intensity of the 18 S ribosomal RNA band (Qiagen: RNeasy® Micro Handbook, 2003).

A 1.2% formaldehyde agarose gel was made as follows: agarose (Hispanagar; catalogue number D1LE / H111206) and FA running buffer with a final concentration of 2 mM MOPS 3-[N-morpholino] propanesulfonic acid) (Roche, catalogue number 11124684001), 0.5 mM sodium acetate (Fluka; catalogue number 71183) and 0.1 mM EDTA (Merck; catalogue number 3685) in 100 ml RNase free water at pH 7. The mixture was heated to dissolve the agarose and then cooled to a temperature comfortable to touch the mixture. 37% Formaldehyde and EtBr at a final concentration of 200 ng/ml was added. The gel was canted into a 6 cm x 10 cm x 0.5 cm tray and a comb with 10 wells was inserted into the gel. Prior to running the gel, the gel was equilibrated in 1 x FA gel running buffer with 37% formaldehyde for at least 30 min. RNA samples were prepared for FA gel electrophoresis by adding one volume of 5 x loading buffer according to Sambrook and Russel (2005) per four volumes of RNA sample. The RNA samples were heated for 10 minutes at 65ºC in a BioRad PowerPac Basic system thermo cycler to further denature the rRNA. After the

sample was chilled on ice, it was loaded onto the equilibrated gel. Gels were run at constant 70 V for 1 h in 1x FA running buffer. RNA in FA agarose gels was visualised with UV transillumination using Syngene a ChemiGenius Bio-Imaging Gel-documentation system and GeneSnap software (Syngene Vacutech, England).

#### 2.2.3.2 Agarose gel electrophoresis

Agarose gel electrophoretic analysis was done according to Sambrook and Russel (2005). Gels contained 1% agarose in 1 x TAE buffer (Tris-Acetic acid-EDTA, pH 8.2). The gel was canted into a 6 cm x 10 cm x 0.5 cm tray with 10 wells. O'geneRuler DNA ladder mix (Fermentas, catalogue number SM1173) was used in all agarose electrophoretic analysis as a DNA marker. Samples were loaded with 6 x gel loading buffer (Fermentas; catalogue number R0611). Gel electrophoresis was performed in 1 x TAE running buffer in a BioRad PowerPac Basic system at 70V for 1h. DNA in agarose gels was visualised with UV transillumination using Syngene a ChemiGenius Bio-Imaging Gel-documentation system and GeneSnap software (Syngene Vacutech, England).

#### 2.2.3.3 Polyacrylamide gel electrophoresis

Sodium dodecyl sulphate polyacrylamide gel electrophoretic (SDS-PAGE) analysis denatures and separates proteins according to their electrophoretic mobility. The rate of migration depends largely on the length of the polypeptide chain, molecular weight as well as post-translational modification (Sambrook & Russel, 2005). The SDS-PAGE used in this study was the discontinuous system developed by Laemmli (1970). A stacking gel was layered on top of the separating gel. Separating gels containing 8%, 10% and 12% polyacrylamide (as indicated in the text) were prepared to accommodate the size of protein to be analysed. The smaller the protein of interest to be separated from others, the higher the percentage. Each gel was made as described by Sambrook and Russel in 2005. The 2 x protein solvent buffer (PSB) which contained 0.5 mM Tris, 10% (w/v) SDS, 4% glycerol 20% ß-Mercaproethanol (ß-ME) was added to the protein in equal volumes (1:1). ß-ME is a reducing agent that helps with the denaturing of the proteins. Glycerol is denser than water and the proteins in the loading dye containing glycerol will not float out of the well causing loss or leaky results. Proteins were run in a TGS buffer consisting of 3% Tris-HCl pH7, 14.4% Glycine and 1% (w/v) SDS. After the gels were run for 1h at 130V, they were stained with Coomassie Brilliant Blue dye solution that consisted of 0.125%

Coomassie Brilliant Blue dye dye, 50% methanol and 10% acetic acid. Coomassie Brilliant Blue dye binds to proteins and makes them visible on the gel. The gels were destained overnight in a solution containing 5% acetic acid and 5% methanol. For all SDS-PAGE, the PageRuler™ Prestained Protein ladder Plus (Fermentas, catalogue number SM1811) was used. A 20 µl sample was loaded into each well containing 10 µg protein sample and 10 µl PSB, unless otherwise stated.

### 2.2.4 Reverse transcriptase

The messenger RNA (mRNA) encoding bovine GLYAT was converted to double stranded DNA (dsDNA) using a two-step approach. The first step was to convert all the mRNA present in the total RNA preparations (Section 2.2.1) from the bovine liver cells into single-stranded cDNA using a commercially available enzyme called Avian Myeloblastosis Virus AMV reverse transcriptase from SBS (catalogue number FAM-500 ) and Oligo $(dT)_{23}$ primers. Most eukaryotic mRNA contain a 3' poly-A tail to which the Oligo $(dT)_{23}$ primer can anneal. Total RNA extracted from bovine liver was used as template for a cDNA synthesis reaction. The reaction mixture consisted of 1 x AMV buffer (composition not revealed by manufacturer), dNTPs (10 mM each), 20 U RNasin, 100 pmol Oligo $(dT)_{23}$, 10 U AMV RT and 0.5 µg total bovine RNA [215 ng/ul]. Nuclease free water was added to a final reaction volume of 20 µl. The reaction was gently mixed and briefly centrifuged. A BioRad Thermo Cycler was used to incubate the reaction at 45˚C for 30 minutes. Enzyme activity was terminated by increasing the temperature to 94˚C for 5 minutes. This step was to degrade the enzymes. An aliquot of the bovine cDNA library was used in PCR reactions and the rest was stored at -20ºC.

### 2.2.5 Polymerase chain reaction

The second step in preparing dsDNA encoding bGLYAT was to selectively amplify the cDNA from the pool of cDNA derived from total RNA containing the mRNA. The PCR was performed directly after cDNA synthesis using 5 µl of the cDNA reaction mixture. Two specific bovine GLYAT primers of 23 bp each with a 57% GC content were designed based on the 5' - terminal sequences of the forward and reverse strand of bGLYAT (GenBank number NM177513). The forward primer was *bGLYAT-forN-NdeI* with the following sequence 5'-GCC GCA↓ TAT GAT GTT CCT GCT GC-3'. The reverse primer was *bGLYAT-revC-SalI* with the following sequence: 5'-CAT

G↓TC GAC TCA CAG AGG CTC AC-'3. These primers which also contained restriction enzymes NdeI and SalI for cloning (indicated by the underlined sequence and the arrow indicating the restriction site ↓) were used in a polymerase chain reaction (PCR). The primers were synthesised by Metabion. The concentration of the primer was 100 µM. The annealing time and temperature is dependent on the primer length and GC content. The annealing temperature of both primers was 66ºC.

For the polymerase chain reaction the DNA polymerase used was Phusion™. The PCR was carried out according to the procedure recommended by the manufacturers of Phusion™ High-Fidelity DNA Polymerase (Phusion™: Catalogue number F-530S from Finnzymes). Phusion™ High-Fidelity DNA Polymerase possesses a dsDNA binding domain of the *Pyrococcus*-like enzyme on the polymerase that increases the fidelity 10-fold (Figure 13).

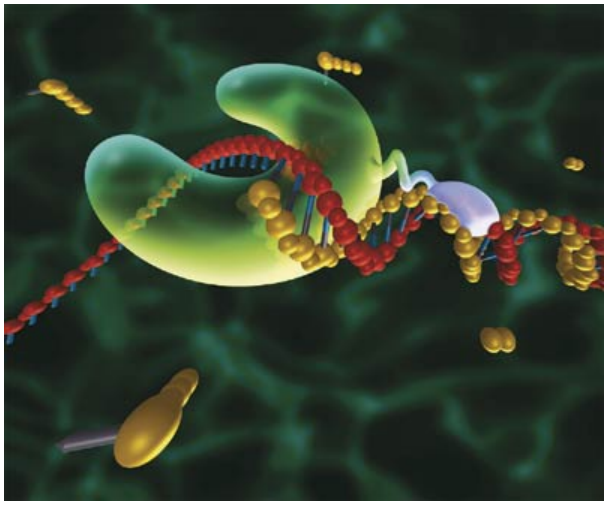


**Figure 13: The structure of Phusion™ High-Fidelity DNA Polymerases.** The double-strand DNA-binding domain (purple) is fused to a novel *Pyrococcus*-like enzyme (green) forming a unique high-performance polymerase - Phusion™ DNA polymerase (Finnzymes Oy).

The unique structure and characteristics of this polymerase makes Phusion™ DNA polymerase excellent for cloning. Its error rate is 50-fold lower than that of *Thermus aquaticus*, 6-fold lower than that of *Pyrococcus furiosus*. Phusion™ DNA polymerase has both 5' → 3' DNA polymerase activity and 3' → 5' exonuclease activity. The kit contained 100 U Phusion™ DNA Polymerase (2 U/µl), 5 x Phusion™ HF Buffer with 7.5 mM $MgCl_2$, DMSO (F515) and a 50 mM $MgCl_2$ solution.

The PCR mixture consisted of 1 x Phusion™ HF Buffer, 10 mM dNTP mix 0.5 µM bGLYAT-forN-NdeI primer, 0.5 µM bGLYAT-revC-SalI primer, 5 ng bovine cDNA, 3% DMSO and 1 U Phusion™ DNA polymerase. The final volume for the PCR was 50 µl. The reaction was incubated in a BioRad Thermo Cycler at an initial denaturing temperature of 98 °C for 30 seconds. The denaturing (overcoming the hydrogen bonds that keep dsDNA intact), annealing (primer attachment to the template DNA) and extension (synthesis of DNA) PCR steps were repeated for 30 cycles. Each cycle had a denaturing step at 98 °C for 10 seconds, an annealing step at 66 °C for 30 seconds and extension at 72 °C for 15 seconds. The final extension was carried out at 72 °C for 10 minutes. The PCR was then incubated at 4 °C for 30 minutes to reach the storage temperature and the reaction was removed anytime during this 30 minutes. The PCR negative control was set up by omitting the cDNA as template. The results were visualised by means of agarose gel electrophoresis (Section 2.2.3.2). The PCR amplicon was purified from primer dimers, polymerase enzymes, nucleotides and buffers by using a commercial kit available from Invitek and stored at 4ºC.

### 2.2.6 DNA clean up

A commercial spin column based kit from Invitek: MSB® Spin PCRapace; for ultrafast purification and concentration of PCR-fragments (Invitek, catalogue number 10202202) was used to purify some of the amplification reactions as indicated in the text from any enzymes, nucleotides, salt, buffer components, primers dimers and other DNA fragments smaller than 80 bp. This kit was also used to purify and concentrate DNA fragments from enzymatic reactions, like enzyme restriction digestion (Section 2.2.8). The DNA binds directly on the spin filter column based surface in the presence of the Binding Buffer supplied with the kit. After centrifugation of 3 minutes, the DNA was eluted with 40 µl nuclease free water. The DNA was either used for a downstream application, or stored at -20ºC for subsequent use.

### 2.2.7 Gel extraction of DNA fragments from an agarose gel

Agarose gel electrophoresis (Section 2.2.3.2) was performed to separate bovine GLYAT DNA amplicon from possible non-specific amplicons. Preparative wells of 1.5 cm each were prepared by taping three 0.5 cm wells (combs) together before pouring in the liquid form agarose. DNA samples were separated on a 1% agarose gel (see 2.2.3) and visualised under UV light. The DNA band of interest (compared to the molecular weight marker), was extracted and purified from the agarose gel according to the instructions of Qiagen with the QIAquick Gel Extraction Kit using a microcentrifuge (Qiagen, Catalogue number 28704). After solubilising the agarose with DNA, it was applied to the QIAquick spin column. QIAquick spin-columns are based on silica-gel membrane technology with selective binding properties. DNA adsorbs to the silica-membrane in the presence of high salt concentration and an optimum pH of $\leq$7. Impurities and contaminants such as primers, salt, enzymes, agarose dyes, other nucleotides, EtBr and detergents do not bind to the membrane and were washed away with a buffer passing through the column (Qiagen: QIAquick Spin Handbook, 2002*).* Pure DNA was recovered by elution with 50 µl $ddH_2O$. The DNA fragment was used for cloning and the rest stored at –20ºC.

### 2.2.8 Restriction enzyme digestion

Restriction enzymes are nucleases. They cleave the covalent bond of a sugar-phosphate backbone of double-stranded DNA molecules at a specific DNA sequence. The restriction enzymes were purchased from Fermentas (www.fermentas.com). Different restriction enzymes function optimally in buffers with different compositions, and the enzymes were therefore supplied with the appropriate buffer. Important factors in buffer compositions are pH as well as the type and concentration of the monovalent cations $K^+$ and $Na^+$. Restriction enzyme digestion was generally carried out on a standard protocol according to the company Fermentas. They recommend using 10 U enzyme per µg DNA. The general digestion reactions were set up as follows; 1 x buffer, 1 µg DNA, 10 units restriction enzyme in a total reaction volume of 20 µl. The reactions were mixed gently, centrifuged briefly and incubated at 37 °C for 3 hours to overnight. Restriction enzyme reactions were terminated as specified by the manufacturer. The compatibility of restriction enzymes to perform double digestion, as well as the ratio of enzymes needed per reaction was obtained on the Fermentas website (www.fermentas.com). The specific DNA fragment of interest was either purified by means of DNA cleanup (Section 2.2.6) or gel extraction (Section 2.2.7).

### 2.2.9 Ligation

DNA ligase is responsible for joining gaps that form in DNA during replication, DNA repair and recombination (Okazaki *et al.*, 1968). DNA ligase catalyses the formation of a phosphodiester bond between adjacent nucleotides with hydrolysis of ATP and inorganic phosphates. Ligation occurs at the nick in dsDNA or when joining cohesive- or blunt-ended dsDNA together (Higgins & Cozzarelli, 1989). The reactions were carried out as described in the Subcloning Notebook of Promega using T4 DNA ligase derived from T4 bacteriophage. The T4 DNA Ligase was obtained from Promega (catalogue number M1-809) and was supplied with the appropriate ligation buffer. Ligation reactions consisted of 3 units T4 DNA Ligase with a standard amount of 100 ng vector and 1 x ligation buffer (composition not revealed by the company) in a total reaction volume of 10 µl. The vector's molar ratio to the inserts' molar ratio used in the reaction was 3:1. To determine the amount of insert needed for ligation, the following equation (obtained from Promega Subcloning notebook) was used.

$$\frac{\text{ng of vector x kbsize of insert}}{\text{kbsize of vector}} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng of insert}$$

**Equation 4: Equation to calculate the amount of insert needed for successful ligation**

A negative control ligation reaction was set up by omitting the insert from the ligation reaction, thereby analysing self-ligation of the plasmid DNA. Ligation reactions were incubated at 24 °C for 3 hours (Promega Subcloning Handbook) and the clones sequence used for transformation of the vector into competent *E. coli* cells.

### 2.2.10 Transformation

#### 2.2.10.1 Preparation of competent cells

Various *E. coli* strains are extensively used as hosts for cloning and expression studies. Bacterial strains are easy to work with, cost effective and have a tremendously high multiplication rate as previously stated in Section 2.1. Before any plasmid can be introduced into a bacterial cell line, the cells have to be competent to take up the DNA. When cells are made competent, their membrane properties are modified to facilitate the uptake of plasmid DNA when they are exposed to temperature differences or chemical substances.

Two bacterial cells strains, namely JM109 for cloning and Origami™ for expression, were used to transform the plasmid of interest. JM109 cells are different from their wild type *E. coli* counterparts. They carry some mutations specifically engineered to help propagate the plasmid. JM109 cells and other cells such as DH5ά™ and XL-1 Blue have a mutation of the *rec*A gene (*rec*A1), a gene involved in recombination. This will limit plasmid recombination so that the plasmid insert is more stable and inactivates nucleases in order to purify higher quality plasmid. *E. coli* has the *Eco*K I restriction site to cleave foreign DNA and *Eco*K methylase to protect and mask host DNA recognition sequences. The JM109 strain has a mutation that knock-out the *Eco*K I restriction enzyme in order to inhibit the cleavage of foreign DNA, but the methylase is still intact to protect and mask host DNA recognition sequences (Promega Subcloning Handbook).

Origami™ cells have mutations in both the thioredoxin reductase (*trx*B) and glutathione reductase (*gor*) genes. This greatly enhances disulphate bond formation in the cytoplasm. Expression in Origami™ cells can yield in 10-fold more active proteins than in another host though overall expression levels is similar (Prinz *et al.,* 1997).

These bacterial strains were made competent in a modified protocol according to Chung & Muller (1988) that was adapted from Hanahan in 1985 (Hanahan, 1985). Luria-Butani (LB) medium was prepared consisting of 1% tryptone, 0.5% yeast and 1%NaCl. Agar plates were prepared that contained 15 g/L agar in LB media. The LB media and LB agar was autoclaved to sterilize it and the LB agar media was canted in 90mm Petri dishes. Frozen glycerol stock of JM109 and Origami$^{TM}$ cells that was stored at -80 °C were thawed, streaked on a LB agar plates and cultured overnight at 37 °C. One colony of about 2 – 3 mm was isolated of each strain with a sterile toothpick and inoculated in 5 ml LB media. The cells were cultured overnight at 37 °C, shaking at 200 rpm. Of the overnight culture, 1.5 ml was inoculated into 50 ml LB medium. The cells were grown to an early log phase at $OD_{600nm}$ of 0.5 at 37ºC, whilst shaking vigorously at 200 rpm. The cells were removed from the incubator and transferred to 50 ml centrifuge tubes and collected by centrifugation at 2500 x g at 4 °C for 10 minutes. The pellet was resuspended in 5 ml transformation and storage buffer (TSB) (Section 1.6% peptone, 1% yeast extract, 0.5% NaCl, 10% Polyethylene glycol, 5% DMSO, 0.1 M $MgCl_2$ and 0.1 M $MgSO_4$). TSB was sterilised with a syringe filter (PALL; catalogue number 4187) and the solution stored at 4 °C. Glycerol was added to the cells suspension to a final concentration of 15% glycerol. Microcentrifuge tubes (Section 1.5) were put on ice to cool down and 200 µl of the cell pellet suspension was stored in each tubes. After incubation on ice, the tubes were immediately chilled by immersion in liquid nitrogen. The frozen competent cells were then stored in liquid nitrogen until needed for transformation. All the LB media and agar plates used to make competent cells did not contain any antibiotics.

### 2.2.10.2 Transformation

The competent cells were transformed with different plasmids as follows: The entire aliquot of competent cells stored in liquid nitrogen (200 µl) was thawed on ice and gently mixed with 100 ng plasmid DNA. The ligation negative control (double

restriction enzyme digested plasmid with no insert) acted as the negative control in the transformation reaction and the original plasmid DNA (not digested) were used as positive controls in the transformation step. The cells and plasmid DNA were incubated on ice for 2 hours. Thereafter, 800 µl TSBG (TSB with 20% glucose) was added to the cells and incubated at 37°C, shaking at 225 rpm. Cells were plated out on LB plates containing 100 µl/ml ampicllin and incubated overnight at 37˚C at 200 rpm. The protocol was reported by Chung & Miller in 1988.

The transformation efficiency of the competent cells was also determined. Transformation efficiency is the amount of colonies obtained with 1 µg DNA transformed. Good competency yield is more than $10^6$ colonies per µg DNA for general cloning applications. The transformation efficiency was determined by adding 0.1 ng plasmid to 100 µl of the competent cells, incubated it for two hours and then adding 900 µl TSBG to the cells afterwards. 100 µl cells were streaked on a LB agar plate in duplicates containing 100 µl/ml ampicillin. The number of colonies was counted and the competency of the cells was calculated.

#### 2.2.10.3 Preparation of glycerol stocks

Bacterial strains were prepared for storage as follows: Cells were streaked on agar plates. An isolated colony was inoculated in 5 ml LB media with a sterile toothpick and cultured overnight at 37°C shaking at 200 rpm. The appropriate antibiotics was used when cells containing plasmids was propagated. Glycerol stocks were prepared by taking an aliquot of the saturated overnight culture and adding glycerol to a final concentration of 15% (Sumbrook and Russel, 2005). The glycerol stocks of bacterial cells were then stored at -80ºC.

### 2.2.11 Plasmid isolation

Plasmids were purified from small (5 – 10 ml) or large (50 ml) cultures. A random colony of bacterial cells on the LB plates was isolated from the plate by a sterile toothpick and inoculated in either 5 or 50 ml LB medium containing 100 µl/ml ampicillin. The cells were cultured overnight at 37ºC, shaking at 200 rpm. The plasmids were isolated by means of a plasmid mini preparation step (see Section 2.2.11.1) for small cultures and plasmid midi preparation procedure step (see Section 2.2.11.2) for larger cultures from the cells.

### 2.2.11.1 Plasmid mini preparation

A range of mini plasmid preparation kits are commercially available that allow extraction and purification of plasmid DNA from small bacterial cell cultures of 5 – 10 ml and remove proteins and other contaminants. Isolating plasmids using PeqLab: E.Z.N.A® Plasmid mini preparation Kit 1 (Classic-Line); (PeqLab; catalogue number 12-6942-02) was used with all the components supplied from the manufacturer. From an overnight 5 ml culture, 1.5 ml was centrifuged, the supernatant discarded and the pellet resuspended. The resuspended pellet was lysed by means of alkaline-SDS. The low pH of the cell suspension was then neutralized and loaded onto spin columns containing a HiBind® matrix that is based on silica-bead technology. The DNA binds to the HiBind® matrix under certain optimal high salt concentrations. DNA was washed and eluted in 50 µl sterile deionised water (Peqlab E.Z.N.A., 2006). An aliquot of the plasmid DNA was used to identify bacterial colonies containing plasmids that contain the DNA encoding bGLYAT. The rest of the purified plasmid was stored at 4ºC. The PeqLab mini preparation procedures for 5 ml overnight bacterial cultures generally yields 25 µg plasmid DNA.

### 2.2.11.2 Plasmid Midi preparation procedure

For extracting plasmids from 50 ml cultures, I used the PureYield™ Plasmid Midi preparation procedure system from Promega (catalogue number: A2495). The typical yield is about 800 µg high-quality endotoxin free plasmid DNA. This plasmid DNA is usually used in eukaryotic transfection and *in vitro* experiments. I did not need endotoxin free plasmids, but no other midi preparation procedure kit was available. It is a quick an easy way to purify large quantities of the plasmids DNA. Cells were collected by centrifugation in a bench top Hereus Multifuge 1 L-R at 1500 x g for 10 minutes in a swinging bucket rotor. The pellet was resuspended, the cells lysed and the pH neutralised. A PureYield™ clearing column was used to remove all the cell debris by collecting the bigger particle on top of the membrane while the DNA moved through the membrane. The flow-through from the clearing reactions was poured onto the PureYield™ binding column. The binding column was washed with a solution that contained isopropanol (to optimise the salt concentrations for binding of DNA onto the silica-membrane). The flow-through was discarded. The column was washed and the flow-through discarded. Another centrifugation step of 10 minutes ensured removal of ethanol. The plasmid DNA was eluted with 600 µl nuclease free water. The eluted plasmid DNA was transferred to a 1.5 ml micro centrifuge tube and

stored at -80ºC for further experiments. The concentrations were determined by means of Nanodrop Spectrometric analysis of nucleic acids (Section 2.2.2).

### 2.2.12 DNA sequencing analysis

Nucleotide sequencing of the recombinant plasmids was done by Inqaba Biotechnical Industries (Pty) Ltd. The plasmids were purified according to Section 2.2.11. The basic service for sequencing from Inqaba was used. Inqaba requires 100 ng/µl plasmid DNA of 5 kb plasmid size and an additional 20 ng/µl for every 1 kb in plasmid size. The results were obtained online from the Inqaba website (www.inqaba.com) and the free Vector NTI software from Invitrogen and Clustal was used to process and analyse the results.

### 2.2.13 Bacterial protein expression by pColdTF using Origami™ cells

A pipette tip scraping of the frozen Origami™ glycerol stock (Section 2.2.11.3) containing the plasmid of interest was inoculated in 5 ml LB media containing 100 µg/ml ampicillin and cultivated at 37 °C overnight, shaking at 200 rpm. The 5 ml overnight cell culture was then transferred to 100 ml LB containing 100 µg/ml ampicillin and again cultured at 37 °C until it reached a reading for optical density at 600 nm [$OD_{600nm}$] of 0.4 - 0.5. The culture was chilled to 15ºC for 30 minutes. For induction of expression of the target gene, 0.1 mM IPTG (Promega, catalogue number V395A) was added to the appropriate culture. IPTG (isoprpyl-β-D-thiogalactopyranoside) is an inducer of β-galactosidase activity in many bacteria. Functioning as a *lacI* analog, IPTG induces β-galactosidase activity by binding to and inhibiting the *lac* repressor, thereby favouring target protein expression. The cultures were incubated at 15°C for 24 hour shaking at 200 rpm.

### 2.2.14 Protein extraction from bacterial cells

The proteins expressed by the bacterial cells, especially the recombinant proteins usually need to be analysed. To enable this, a detergent known as BugBuster Protein Extraction Reagent (Novagen Catalogue number 70921-4) was used to gently disrupt the cell wall of the Origami™ cells. This step liberated proteins from the cells making it possible to purify the recombinant protein without denaturing it. The advantage of using BugBuster instead of mechanical methods such as a French Press or sonication is that it is quick, subtle and low-cost. BugBuster is fully compatible with

affinity purification resins such as nickel column chromatography. The BugBuster reagent is supplemented with Benzonase nuclease (Merck, catalogue number 70746-3) and rLysozyme™ solution (Novagen, catalogue number 71110-4). Benzonase nuclease reduces the viscosity of the lysate by digesting the chromosomal DNA. rLysozyme™ enhances the efficiency of the protein extraction by hydrolyzation of *N - acetylmuramide* linkages in cell walls.

A 50 ml overnight culture was harvested and collected by centrifugation in a cooled bench top Hereus Multifuge 1 L-R with a fixed angle rotor for 20 minutes at 1500 x g. Approximately 0.5 g wet cell pellet was obtained from 50 ml bacterial cultures. For each gram wet cell pellet, 5 ml BugBuster containing Benzonase nuclease and rLysozyme™ was used for resuspension of the pellet. The cell suspension was then incubated for 20 minutes shaking slowly at 80 rpm at room temperature. Of the resuspended pellet, 1.5 ml was used to separate the soluble proteins from the insoluble proteins and cell debris. The rest was used as the total fraction for future analysis. Separation of soluble and insoluble fractions was done by centrifugating the culture at 13000 x g for 20 minutes at 4 °C in an Eppendorf micro centrifuge. The supernatant was transferred to a 1.5 ml micro centrifuge tube and the pellet resuspended in 1.5 ml 1 x PBS. All three fractions (total, soluble and insoluble) were subsequently subjected to SDS-PAGE analysis. The rest of the protein fractions were stored at – 20 °C.

### 2.2.15 Purification of histidine tagged recombinant proteins

If a recombinant protein is expressed with a poly-histidine tag, (usually 6 His residues) the protein can be purified from other proteins that do not contain this histidine tag by means of nickel column chromatography.

The Protino® Ni-TED 2000 packed columns (Macherey Nagel, catalogue number 745120.25) enables fast and convenient purification of polyhistidine-tagged proteins. The columns are based on immobilized metal ion affinity chromatography and gravity flow. This column has the capacity to bind 5 mg histidine tagged proteins by means of dry silica-based resins pre-charged with $Ni^{2+}$ ions. This nickel ion has a single protein binding site for a histidine tag which minimizes non-specific binding of

contaminating proteins to the resin. The chelating group needed for binding of the nickel with histidine is based on TED (triscarboxymethyl ethylene diamine) which is a strong pentadene metal chelator. This ensures higher target protein purity. Tandem column purification was used as recommended by the Macherey Nagel for the purification of high concentrations of the recombinant proteins containing a C-terminal histidine tag. Tandem purification was used, which means that the loading of the sample from the first column dripped onto the next column. Thereafter, each step was repeated in order to wash and elute both columns as recommended by the manufacturer. Columns were equilibrated with LEW (lysis/equilibrium/wash) buffer supplied with the kit (details of the buffer are not supplied by the company) and the sample was loaded onto the first column. The flow through was directly loaded onto the second column. The columns were washed twice with LEW buffer. Thereafter the $Ni^{2+}$ bounded polyhistidine-tagged proteins were eluted in a three step elution step with elution buffer. The elution buffer contains 250 mM imidazole that binds to the immobilized $Ni^{2+}$ ions and competes with the polyhistidine-tagged (6 residues) proteins. This causes the proteins to be released from the resin in the presence of imidazole. The flow-through of each step was analysed with SDS-PAGE.

### 2.2.16 Concentration of proteins

Ultrafiltration devices are used to increase the concentration of proteins or to exchange buffers. Vivaspin 20 ml concentrators (Sartorius Stedium Bioteck; Catalogue number: VS2001) are disposable devices that concentrate and purify biological samples. This specific centrifugal device has a molecular weight cut off of 10,000 Da. Samples are concentrated in 10 to 30 minutes with a typical 95% recovery of proteins smaller than 10 kDa. Optimal flow conditions are achieved by the longitudinal twin vertical membrane orientation and channel concentration chambers. The centrifugal force pulls particles and solids away from the membrane to the bottom of the device. Macromolecules collect in an impermeable concentrate pocket integrally molded below the membrane surface, thereby eliminating the risk of filtration to dryness (Sartorius, 2007).

All the flow-through fractions of the elution steps in the Ni-TED column were pooled for concentration of the protein (Section 2.2.15). A 25 º fixed angle rotor was used for centrifugation in a bench top Hereus Multifuge 1 L-R. The molecular weight cut-off

(MWCO) was determined so that the cut off is at least 50% smaller than the molecular size of the recombinant protein. The maximum volume of 14 ml of the protein sample was centrifuged at 20ºC at 8000 x g for 18 minutes. After centrifugation, the flow through was discarded and 14 ml buffer added to the remaining solution above the membrane. The buffer contained 0.1 M KCl and 0.02 M Tris at pH 8 (Mawal & Qureshi, 1994). The protein extract was again centrifuged, the flow-through discarded and the device refilled with the buffer. This step was repeated three more times. The sample of approximately 500 µl was then recovered from the bottom of the concentration pocket with a micro pipette.

### 2.2.17 Bicinchoninic acid protein assay

Protein quantitation is important to verify success of the lysate step, to normalize multiple samples for storage or side-by-side comparison, as well as to determine the yield of protein obtained from expression. It is important to determine the protein concentration in a sample in order to quantify the results. The bicinchoninic acid (BCA) assay is based on a chemical principle similar to that of the Buiret and Lowry assay. Proteins react with $Cu^{2+}$ to produce $Cu^{+}$ and BCA, $Cu^{+}$ then chelates with BCA which results in a purple colour of the copper-BCA complex. The absorbance is then read at 560 nm and concentrations determined from a standard curve (Smith *et al.,* 1985).

A standard dilution series containing 0, 2, 4, 6, 8, and 10 µg / µl bovine serum albumin (BSA) was prepared in 10 µl volumes in a 96 well microtitre plate. A tenfold dilution series of recombinant protein was also prepared, each containing 0, 2, 4, 6, 8, and 10 µl of either protein in a final volume of 10 µl in the well. A solution containing bicinchoninic acid solution-BCA (Sigma; Catalogue number, B9643-1L) and copper II sulphate solution $CuSO_4.5H_20$ (Sigma; Cat no, C2284 – 25ML) was used as colorimetric detection solution to determine the concentration of the proteins. The ratio of BCA to copper II sulphate was 50:1. A 200 µl aliquot of this solution was added to each well. All samples were then incubated at 37ºC for 20 minutes. The absorbance was then read at 560 nm in a BioTek® microplate fluorescence (FL600) plate reader. The amount of protein in the sample was quantified using the standard curve from the bovine serum albumin (BSA) standard series.

### 2.2.18 Thrombin cleavage

Thrombin is an endoprotease and one of the most site-specific proteases. In the pColdTF expression system it is used for the proteolytic removal of the C-terminal fused His-tag peptide from the expressed recombinant target protein. Restriction grade Thrombin, (50 U at 1U/µl) (Novagen, catalogue number 69671-3) cleaves protein at the following amino acid sequence (LauValProArg↓GlySer) between the residue Arginine and Glycine. Thrombin is supplied with a dilution/storage buffer and a 10 x reaction buffer. One thrombin unit can cleave 1 mg of the protein when incubated in the standard digest buffer at 20ºC for 16 hours.

A pilot small scale optimization for thrombin cleavage was done by Rencia van der Sluis of our laboratory according to the recommendations of the manufacturer. The cleavage reaction contained 1 x thrombin cleavage buffer, 10 µg target protein, 1 µl diluted thrombin (0.02 U) in a total reaction volume of 50 µl. The reaction was incubated at 21ºC for 16 hours.

### 2.2.19 GLYAT enzyme activity reaction

GLYAT enzyme activity was determined using a method described by Kolvraa and Gregersen (1986) which is a colorimetric reaction for the detection of coenzyme A. In the GLYAT reaction coenzyme A is one of the products and is therefore amendable to colorimetric analysis using the chromogen DTNB. During the reaction the free thiol group of the liberated coenzyme A (glycine-dependent CoASH release from benzoyl-CoA) reacts with glycine and the yellow compound formed is strongly absorbed at 412 nm. The progress of the GLYAT reaction can be followed by monitoring the increase in absorbance at 412 nm. The reaction mixture consisted of 25 mM Tris.HCl (pH 8) buffer, 0.1 mM benzoyl-CoA as substrate, 0.1 mM DTNB which is the colorimetric substance and 200 mM glycine (pH 8) in a final volume of 100 µl (Mawal & Qureshi, 1994). The reaction was incubated at 37ºC over a period of 10 minutes using a UVICON XS spectrophotometer. The LabPower Junior software was used to analyse the glycine-dependent CoASH release from benzoyl-CoA at 412 nm. The reaction was initiated by the addition of the sample of which the protein content varied. A crude cytoplasm extract from bovine liver containing GLYAT was prepared by a fellow student (Mr. JHJ Fourie) was used as positive control according to Nandi

(1979) and Van der Westhuizen (2000). For each reaction a negative control was set up by omitting glycine from the reaction.

## 2.3 Results

### 2.3.1 Extraction of total RNA from bovine liver

In order to clone the mRNA encoding glycine N-acyltransferase from bovine liver, total RNA was extracted from the liver. The bovine liver was obtained from the Potchefstroom abattoir and the RNA was preserved with RNAlater™ to avoid any RNA degradation. The total RNA was isolated from the bovine liver by means of the RNAeasy® Mini kit (Qiagen) (Section 2.2.1). The RNA concentration was determined by spectrophtometric analysis for RNA (Section 2.2.2). RNA was analysed on a formaldehyde denaturing 1.2% agarose gel (Section 2.2.3.1). Figure 14 shows total RNA isolated from bovine liver.

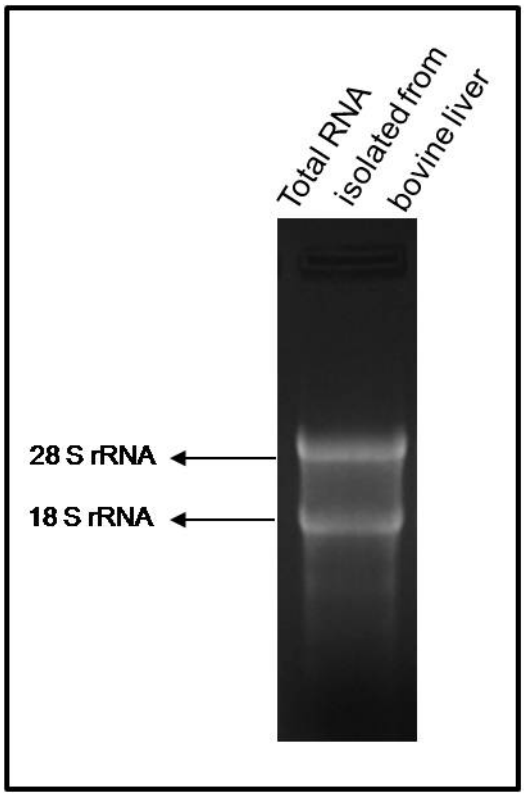



**Figure 14: Formaldehyde denaturing agarose gel electrophoretic analysis of total RNA isolated from bovine liver.** 4 µl of a 50 µl preparation was loaded with 1 µl loading dye. The 28 S rRNA and the 18 S rRNA band are indicated.

The 28 S and 18 S ribosomal RNA (rRNA) bands are visible. A RNA marker was not available. In preparations of very good quality total RNA the ratio of 28 S rRNA : 18 S rRNA is usually 2:1. Visual inspection of Figure 14 shows that the 28 S rRNA and 18 S rRNA is more or less present in the same ratio (1:1). Even though a good result should have had a 2:1 ratio, the RNA was used for further experiment never-the-less. A slight smear of RNA from the 28 S rRNA downwards is visible. This is most probably due to degradation of some RNA. The purity determined by measuring the absorbance at 260/280 nm on a Nanodrop spectrophotometer was 1.97 (Section 2.2.2, equation 1). This ratio is higher than 1.8 and indicates a pure nucleic acid sample without any protein contamination. The concentration of the RNA preparation was 492.7 ng/µl (Section 2.2.2, Equation 3). The total yield of RNA extracted from 10 mg bovine liver was 7 µg total RNA. According to the RNeasy® mini kit (Qiagen), the average yield of total RNA isolated from 5 mg of mouse or rat tissue is usually 15 µg. The yield achieved was 25% of what was expected. It is quite possible that I used less liver tissue than what I thought I weighed because the sample was already submerged in RNAlater™. The RNAlater™ might have contributed to the weight of the tissue.

### 2.3.2 Generation of a bovine cDNA library with reverse transcriptase and amplifying the bGLYAT cDNA by means of PCR and gene-specific primers

In order to generate dsDNA encoding bGLYAT, a two-step RT-PCR method was used. A cDNA library was prepared by a reverse transcriptase (RT) reaction using the Oligo$(dT)_{23}$ as primer and total RNA isolated from bovine liver as template. The cDNA library was prepared as described in Section 2.2.4. To amplify DNA encoding bGLYAT from the cDNA, gene-specific primers were designed as follows: The sequence of bGLYAT determined by Vessey and Lau (1996), GeneBank number, NM177513, was used to design gene-specific primers for the forward and reverse strand for amplification of DNA encoding bGLYAT (Figure 15). A 5’ NdeI and 3’ SalI restriction site (Section 2.2.5) for cloning in frame with TF in pColdTF were added to the primers respectively (Section 2.2.5) as seen in Figure 15.


**Figure 15: Reference DNA sequence of the bovine glycine N-acyltransferase (GLYAT), obtained from the mRNA sequence NM 177513 (Vessey and Lau, 1996).** The sequence highlighted in yellow is the open reading frame of bGLYAT. The forward primer (bGLYAT-forN-NdeI) and reverse primer (bGLYAT-revC-SalI) with incorporated restriction enzyme sites NdeI and SalI underlined and highlighted. The addition of the NdeI and SalI site in the primer sequence for directional cloning is marked in red.

bGLYAT gene specific primers were used to amplify the DNA encoding bGLYAT from the cDNA library using Phusion DNA polymerase. The PCR was described in Section 2.2.5. A PCR temperature gradient with annealing temperatures between 60ºC and 70ºC and 1ºC intervals was carried out (results not shown). The optimum annealing temperature for the PCR was 66ºC. Six PCR reactions were subsequently done in parallel in order to prepare enough bGLYAT amplicon for downstream use. The amplicons were visualised on a 1% agarose gel (Figure 16).

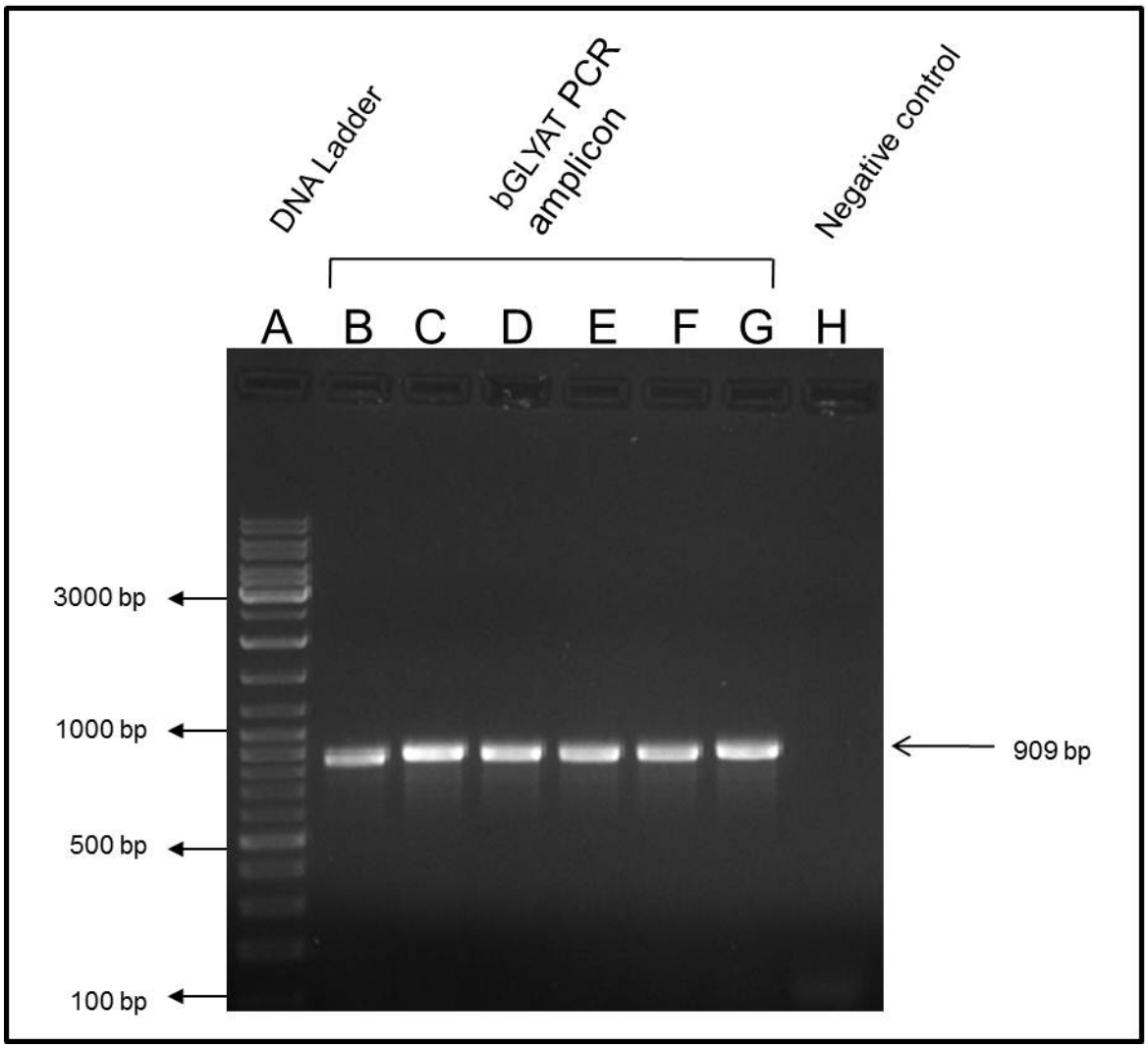



**FIGURE 16: Agarose gel electrophoretic analysis of amplicons obtained by PCR of bovine cDNA with bGLYAT specific primers to generate bGLYAT encoding amplicons.** Lanes: A) 5 µl O'geneRuler DNA ladder; B – G) 5 µl of 50 µl PCR reactions done in parallel six times; H) Negative control, 5 µl PCR reaction containing no bovine GLYAT cDNA as template.

The amplicon was at approximately 900 bp which compared well to the expected amplicon size which is 909 bp (Figure 16, lanes B - G). There was no non-specific amplification and also no amplification in the negative control (Figure 16, lane H) which indicates that there was no contamination in the reaction mixture. The next step was to pool three of the six parallel reactions (Figure 16, lanes B, C and D) and gel extract the DNA as explained in Section 2.2.7. This method also purified the DNA from components such as excess primers, buffers, enzymes and nucleotides. An aliquot of the amplicon was analysed after gel extraction (Figure 17).

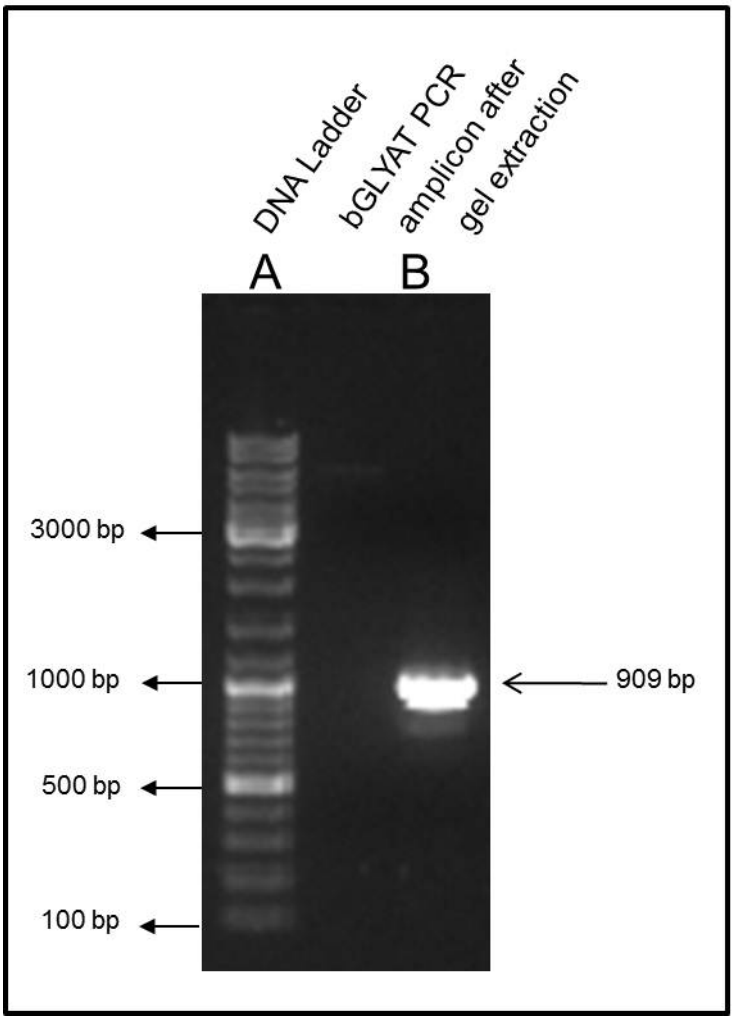



**Figure 17: Agarose electrophoretic analysis of the bGLYAT amplicon after gel extraction.** Lanes: A) 5µl O'geneRuler DNA ladder mix; B) 10 µl of a 50 µl preparation of bGLYAT amplicon after gel extraction.

A 900 bp amplicon similar to that shown in Figure 16 was obtained (Figure 17, lane B). The well was slightly overloaded. No other bands were visible in the amplicon preparation. The $A_{260/280}$ nm ratio of the purified DNA was 1.97. The concentration was 55 ng/ul which implied a yield of 2.75 µg DNA (2750 ng). The next step was to determine whether the 900 bp DNA amplicon was indeed bGLYAT DNA.

### 2.3.3 Putative bGLYAT amplicon digestion with NcoI and PstI

In order to determine whether that 900 bp DNA fragment extracted from the agarose contained the bGLYAT ORF, it was digested with NcoI (Fermentas, catalogue number ER0572) and PstI (Fermentas, catalogue number ER0611). These two restriction sites are not present in the pCold vector's multiple cloning site but is present in the bGLYAT ORF. NcoI appears once in the ORF of bGLYAT and PstI twice. The fragments were estimated during the planning phase using the bGLYAT sequence obtained from NCBI (NM177513) and the restriction enzyme site using VectorNTi software. The expected fragments of the bGLYAT amplicon digested with NcoI are 795 bp and 114 bp. The expected fragments for the bGLYAT amplicon digestion with PstI are 595 bp, 273 bp and 41 bp. The results are shown in Figure 18.

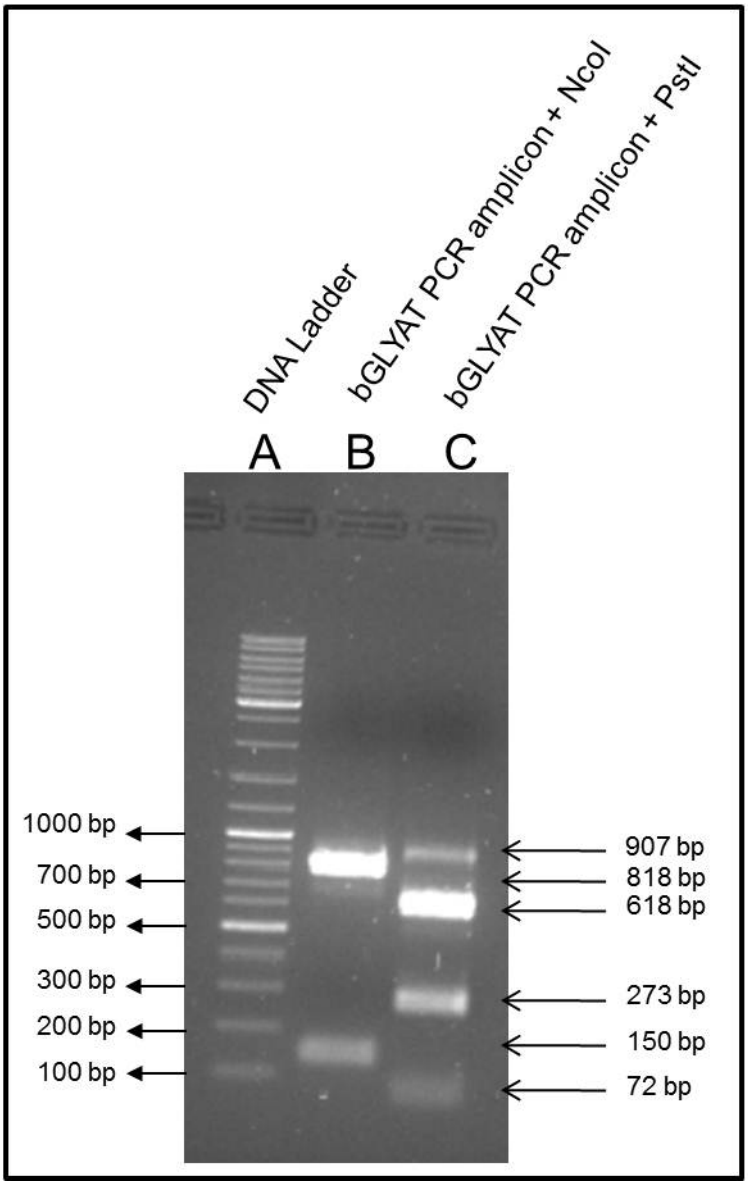



**Figure 18: Agarose electrophoretic analysis of NcoI and PstI restriction enzyme digestion of the 900 bp DNA amplicon generated by bGLYAT gene-specific primers.** Lanes: A) 5 µl O'geneRuler DNA ladder; B) 5 µl of a 20 µl reaction of 900 bp amplicon digested with NcoI; C) 5 µl of a 20 µl reaction of 900 bp amplicon digested with PstI. The sizes of the ladder DNA are indicated on the left and the calculated fractions of the bovine GLYAT amplicon on the right.

Figure 18 shows the restriction enzyme digestion of the 900 bp amplicon generated from bovine liver cDNA with bGLYAT specific primers. Digestion with NcoI resulted in two fragments, one of approximately 800 bp and another at 150 bp. This correlates with the expected sizes of 795 bp and 114 bp (Figure 18, lane B). When the 900 bp amplicon was digested with PstI, four fragments were visible at 900 bp, 600 bp, 300 bp and one less than 100 bp in size. The three smaller bands correlated with the expected results of 595 bp, 273 bp and 41 bp. A 900 bp band indicates partial

digestion of bGLYAT amplicon with PstI (Figure 18, lane C). Large DNA fragments such as the 800 bp fragment stain more intense than the smaller fragments seeing that EtBr intercalates into the sugar-phosphate backbone of DNA in an equimolar ratio. This is true for both Figure 18 lane B and C. The finding that the band at 900 bp is very light despite its size indicates that the only a small amount of the DNA was not digested with PstI. These restriction enzyme digests confirmed that the 900 bp DNA amplicon has the NcoI and PstI recognition sites present in the reference bGLYAT sequence. Actual DNA sequence confirmation was not done at this stage. Sequencing was only done later in the study after the cloning of the amplicon into pColdTF and pColdIII respectively was completed.

### 2.3.4 Cloning of the bGLYAT amplicon into the bacterial expression plasmid pColdTF

For the bGLYAT amplicon to be cloned into pColdTF, DNA fragments with compatible restriction enzyme termini had to be generated. Preparation of the insert, the bGLYAT amplicon, and the vector pColdTF, was done by double digesting each of them with NdeI and SalI. The bGLYAT amplicon was amplified with gene-specific primers that contained the restriction sites to ensure that the insert and plasmid had compatible overhangs. Restriction enzyme digestion was done as described in Section 2.2.8. After restriction digestion, the bGLYAT amplicon and the pColdTF plasmid DNA was purified from the smaller fragments by means of gel extraction (Section 2.2.7), thereby removing buffers, salts and enzymes.

As stated in previous results, the bGLYAT PCR amplicon is expected to be 909 bp. After digestion with NdeI and SalI, very small fragments of the DNA are removed from the 3’ and 5’ ends of the PCR amplicon. The expected fragment after digestion is 894 bp. When pColdTF is linearized with either NdeI or SalI, the plasmid is 5769 bp. After digestion with NdeI and SalI, 40 bp of the multiple cloning site will be removed. The plasmid should then be 5728 bp. The results are shown in Figure 19.

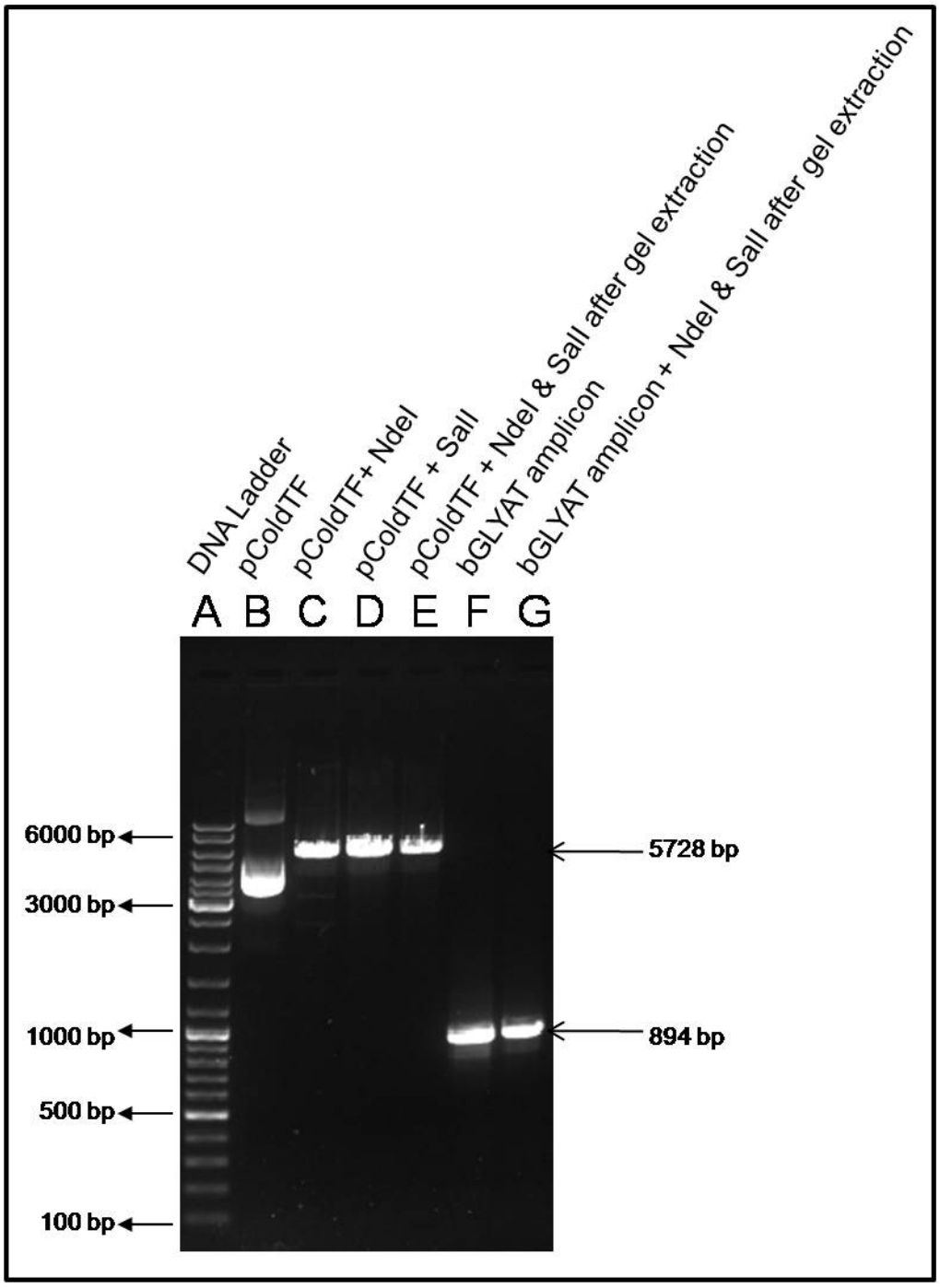



**Figure 19: Agarose gel electrophoretic analysis of NdeI and SalI digested pColdTF and the bGLYAT amplicon.** Lanes contain 10 µl sample with 2 µl loading dye: Lanes A) O'geneRuler DNA ladder; B) pColdTF undigested; C) pColdTF digested with NdeI; D) pColdTF digested with SalI; E) pColdTF digested with NdeI and SalI and gel extracted; F) bGLYAT amplicon undigested; G) bGLYAT amplicon digested with NdeI and SalI and gel extracted.

As seen in Figure 19 lane F, the bGLYAT amplicon is approximately 900 bp. Digestion of the pColdTF vector with either NdeI (Figure 19, lane C) or SalI (Figure 19, lane D) resulted in a linearized plasmid of about 6000 bp. This single digestion was also useful to determine whether the restriction enzyme were capable of cutting

the plasmid completely. The results were successful. The size difference between the single digested and double digested fragments of both pColdTF and the bGLYAT amplicon is not visible, seeing that the size of the DNA removed is 41 bp and 15 bp respectively which is too small and cannot be resolved on a 1% agarose gel. In lane B, the pColdTF circular plasmid can be seen in the supercoiled (bottom band) and relaxed open circle form (top band). There are no other bands visible on the gel except the 5728 bp pColdTF (Figure 19, lane E) and 894 bp PCR amplicon (Figure 19, lane G) after gel extraction. The insert and plasmid were successfully purified. The DNA concentration of the bGLYAT insert after NdeI and SalI digestion as well as gel extraction was 11.2 ng/µl. The A260/280 ratio was 2.08. The total yield of digested and gel extracted bGLYAT was 560 ng of DNA. The pColdTF concentration was 28.7 ng/µl and a 260/280 ratio of 1.83. The yield of the pColdTF plasmid was 1435 ng.

The ligation reaction of the prepared bGLYAT insert and pColdTF plasmid was set up as described in Section 2.2.9. The negative control contained no insert DNA and the positive control was the circular pColdTF plasmid. The plasmids were transfected into competent JM109 cells (Section 2.2.10.2) and streaked out on two LB agar plates containing ampicillin. Ligation of bGLYAT with pColdTF resulted in 80 colonies on one plate and 113 colonies on the other. There were four colonies on the negative control and over 500 on the positive control (results not shown). The transfection efficiency calculated as described in Section 2.2.10.2 was 5 x $10^5$ colonies per µg of DNA. An acceptable amount of colonies were obtained from the transformation. These colonies can be seen as possibly containing a plasmid of the amplicon encoding bGLYAT that was cloned into pColdTF. It is not possible for the pColdTF plasmid to ligate on itself after NdeI and SalI double digestion because NdeI and SalI do not have compatible ends. The colonies on the ligation control can be plasmids that did not digest completely or at all during restriction enzyme digestion with NdeI and SalI. This causes the plasmid to form a circular DNA molecule again and have ampicillin resistance, explaining the colonies on the negative control LB Amp plate.

Ten colonies were then selected with a sterile toothpick and streaked on a master LB Amp plate, sealed with parafilm and then stored at 4ºC. The toothpick was then used to inoculate cells still on the toothpick in LB Amp media. The cultures were incubated overnight shaking at 200 rpm at 37ºC. Plasmids were extracted from these cultures

using the mini plasmid preparation protocol described in Section 2.2.11.1. The plasmids were subjected to agarose gel electrophoresis and compared to empty pColdTF. Only 5 plasmids were visibly bigger in size compared to the empty pColdTF (results not shown). The plasmid DNA from these five colonies was then screened by restriction enzyme digestion to verify whether pColdTF contained the DNA encoding bGLYAT.

### 2.3.5 Colony screening of possible recombinant plasmids (pColdTF-bGLYAT)

To analyse whether cloning of the bGLYAT ORF into the pColdTF vector was successful, the 5 plasmids were analysed by means of restriction enzyme digestion. If the bGLYAT amplicon was successfully cloned into pColdTF, the new plasmid should be 6622 bp. In order to see the correct size distribution and not be confused with the super coiled and relaxed coil form of a plasmid, the plasmid should be linearized by restriction enzyme digestion. When the insert can be cut from the plasmid by NdeI/SalI double digestion, two DNA fragments of 5728 bp and 894 bp should be visible. The results are shown in Figure 20.

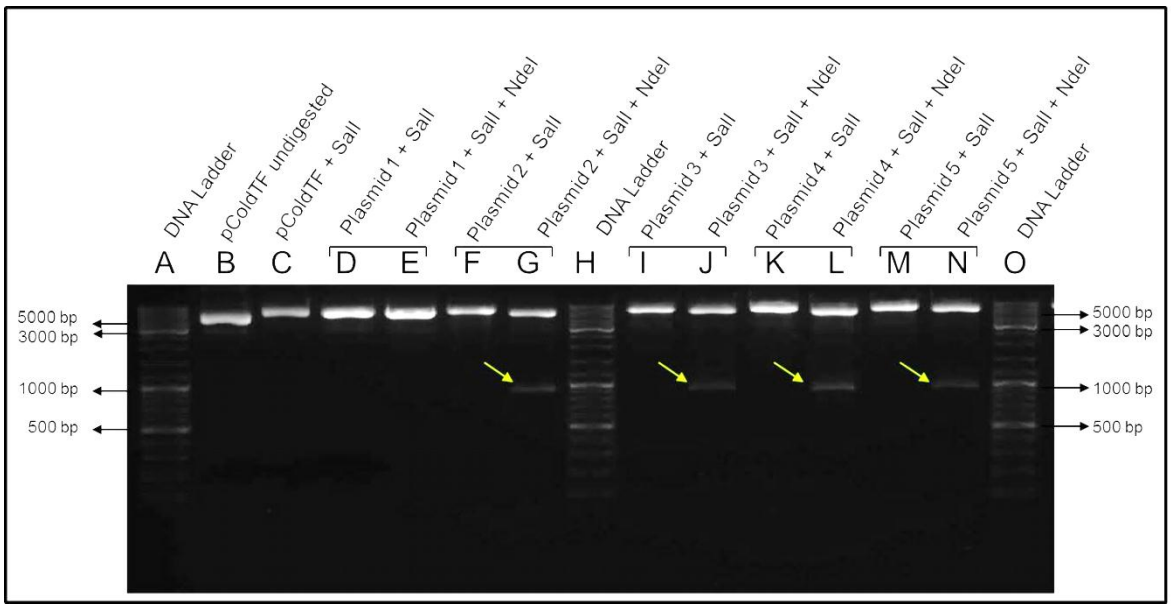



**Figure 20: Agarose gel electrophoretic analysis of restriction enzyme digestion of 5 plasmids selected after cloning the bGLYAT amplicon into pColdTF.** Lanes contain 10 µl sample with 2 µl loading dye: Lanes A, H, O) 5 µl O'geneRuler DNA ladder; B) pColdTF undigested; C) pColdTF digested with SalI; D, F, I, K, M) Respectively; plasmids from colony 1, 2, 3, 4 and 5 linearized with SalI; E, G, J, L, N) Respectively; plasmids from colony 1, 2, 3, 4 & 5 double digested with SalI and NdeI. Light yellow arrows indicate insert (approximately 900 bp) removed by NdeI and SalI digestion.

The five plasmids were linearized with SalI. DNA bands were visible at the position indicative of the expected 6622 bp for the plasmid of colony 2, 3, 4 and 5 (Figure 20, lane F, I, K, M). After double digest with NdeI and SalI, two fragments were visible at positions that could be 894 bp and 5728 bp (Figure 20, lane G, J, L, N). It was assumed that these plasmids did contain the described dsDNA encoding bGLYAT. Digestion with NdeI/SalI of plasmid 1 (Figure 20, lane E) did not result in a fragment at 894 bp. Therefore, it was assumed that pColdTF did not contain the bGLYAT amplicon. The size difference of 894 bp of the linearized plasmid and the plasmid where the insert was removed with NdeI/SalI double digestion is not very clear to see. This is because the gel was not run for long enough to ensure good separation of the bigger fragments. The larger fragment at approximately 6000 bp is more intense than the smaller fragment at 900 bp due to less amount of EtBr in the nucleotide backbone that could be visualised under UV light. The next step was to further confirm the correct cloning and reading frame of pColdTF-bGLYAT by nucleotide sequencing.

### 2.3.6 Nucleotide sequencing analysis of bGLYAT cloned into pColdTF

Inqaba Biotech Industries (Pty) Ltd was used for the nucleotide sequencing services according to their specifications in Section 2.2.12. Sequencing was performed on plasmid 2 and 4 to confirm the correct cloning and reading frame of two possible constructs of pColdTF-bGLYAT.

15 µl of each plasmid at 140 ng DNA per µl was submitted to Inqaba. The forward primer for sequencing was bGLYAT-forN-NdeI with the following sequence: 5'-GCCGCATATGATGTTCCTGCTGC-3'. The reverse sequencing primer was pColdTFRev with the following sequence: 5'-GGCAGGGATCTTAGATTGAG-3'. The chromatogram and nucleotide sequences of both the forward and reverse strands of plasmid 2 and 4 were obtained from the Inqaba website. Vector NTi software was used to align both the forward sequences obtained with the bGLYAT-forN-NdeI primers to the bGLYAT sequence obtained from NCBI (NM177513). The reason for only using the forward stands was because the trimmed sequence length according to Inqaba was longer than 894 bp. The bGLYAT amplicon insert is 894 bp. The sequence was therefore expected to include the primer sequence, the bGLYAT insert as well as a small part of the pColdTF sequence after the SalI restriction enzyme site in the multiple cloning region. This was adequate for this purpose. Results are shown in Figure 21.

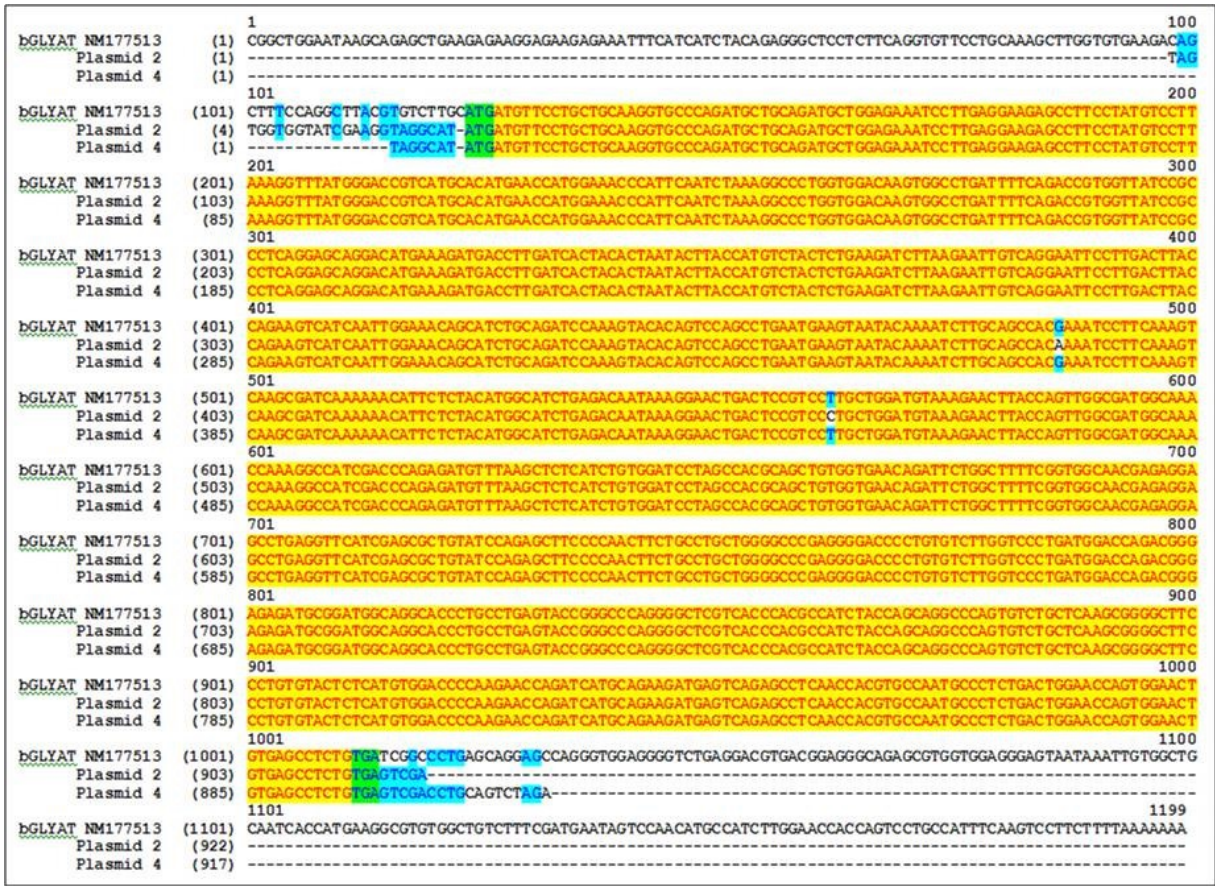



**Figure 21. Forward sequence alignment of the putative bGLYAT DNA of plasmid 2 and 4 to a GenBank bGLYAT reference NM177513.** The open reading frame of bGLYAT are highlighted in yellow. The start and stop codon of the ORF is highlighted in green. The blue sequence highlighted indicate a different nucleotide compared to the reference sequence NM177513.

Figure 21 shows the alignment of the complete open reading frame encoding bGLYAT from a reference sequence from GenBank (NM177513) and the forward primer sequence of plasmid 2 and plasmid 4 which I generated. The start and end of the bGLYAT reading frame is indicated in grey letters. (Methionine) and stop codon (Stop) is highlighted in green. The sequence are highlighted in yellow is identical to the reference sequence NM177512 from NCBI (Vessey *et al.,* 1996). Upstream of the start of the bGLYAT ORF and downstream of the end of the bGLYAT ORF, the nucleic acid sequence identical to that of pColdTF can be seen. This confirmed that the DNA encoding bGLYAT was cloned in the correct orientation into pColdTF and that the reading frame was intact. Plasmid 4 forward sequence was exactly the same as the reference sequence with regard to the open reading frame. Plasmid 2, however did have two nucleotide differences (485 A→G and 563 C→T). Seeing that only two plamids were sequenced, I could not determine if these differences were SNPs or just errors from the Phusion polymerase (Finnzymes). The nucleic acid

changes in the plasmid 2 sequence did, however, not change the amino acid it encodes. Nevertheless, plasmid 2 was discarded. The reverse compliment sequences of both Plasmid 2 and 4 were not aligned with the forward sequences, as only 200-300 base pairs were obtained. Seeing that the complete ORF was sequenced, I did not use the reverse sequences for analysis. This plasmid 4 will be referred to pColdTF-bGLYAT for the rest of the study.

A sample of plasmid 4 of the original bacterial colony containing pColdTF-bGLYAT on the master plate was inoculated in 50 ml LB Amp media and cultured overnight at 37ºC while shaking at 200 rpm. The cell culture was used to prepare stock of the pColdTF-bGLYAT by means of plasmid Midi prep according to Section 2.2.11.2. Some of the cells were used to make glycerol stock that was stored at -80ºC according to Section 2.2.10.3.

### 2.3.7 Bacterial expressing of bGLYAT in pColdTF using Origami™ cells

Origami™ cells were chosen as the host bacterial cell line to express bGLYAT and TF as a fusion protein or TF alone from pColdTF-bGLYAT and pColdTF respectively. The plasmids were transformed into Origami™ cells (Section 2.2.10.2). The cells were cultured to an optical density at 600 nm of 0.5. Expression of TF fusion protein or bGLYAT fused to TF was induced by 0.1 mM IPTG (Section 2.2.13). The cells were then collected and the proteins extracted with BugBuster reagent (Section 2.2.14). The extracted proteins were divided into soluble and insoluble fractions and analysed on 8% (w/v) SDS-PAGE (Section 2.2.3.3).

According to Nandi and co-workers (1979) bGLYAT is a 37kDa protein. Trigger factor (TF) is a 50 kDa chaperone-like protein (Stoller *et al.,* 1995). The expressed fusion protein of pColdTF starts at the TEE and includes the His-tag, TF, the three protease cleavage site sequences as well as the multiple cloning site. The total size of all these purification tags (tf fusion protein) and protein cleavage sites is 56.4 kDa. The size of the total bGLYAT protein with fusion protein was calculated with the following equation: [1kbDNA = 333 aa = 37 kDa]. The total aa sequence size of the fusion protein expressed according to the DNA sequence in pColdTF is 508 aa. This was then divided by 9 to calculate the size of the protein that was 56.4 kDa. Therefore, bGLYAT expressed as a protein with the fusion proteins and other elements is

expected to be 88 kDa. With IPTG induction, Origami™ transformed with pColdTF is expected to produce a 56 kDa protein. It is expected to see this 88 kDa protein in the protein fraction with IPTG induction of Origami™ transformed with pColdTF-bGLYAT. It is also expected that the Origami™ cells with no induction of IPTG, will not have an over-expressed protein. The total protein produced in the induced fraction should also be less than the uninduced fraction. Figure 22 shows the proteins that were expressed in Origami™ cells transformed with pColdTF in the presence and absence of IPTG induction.

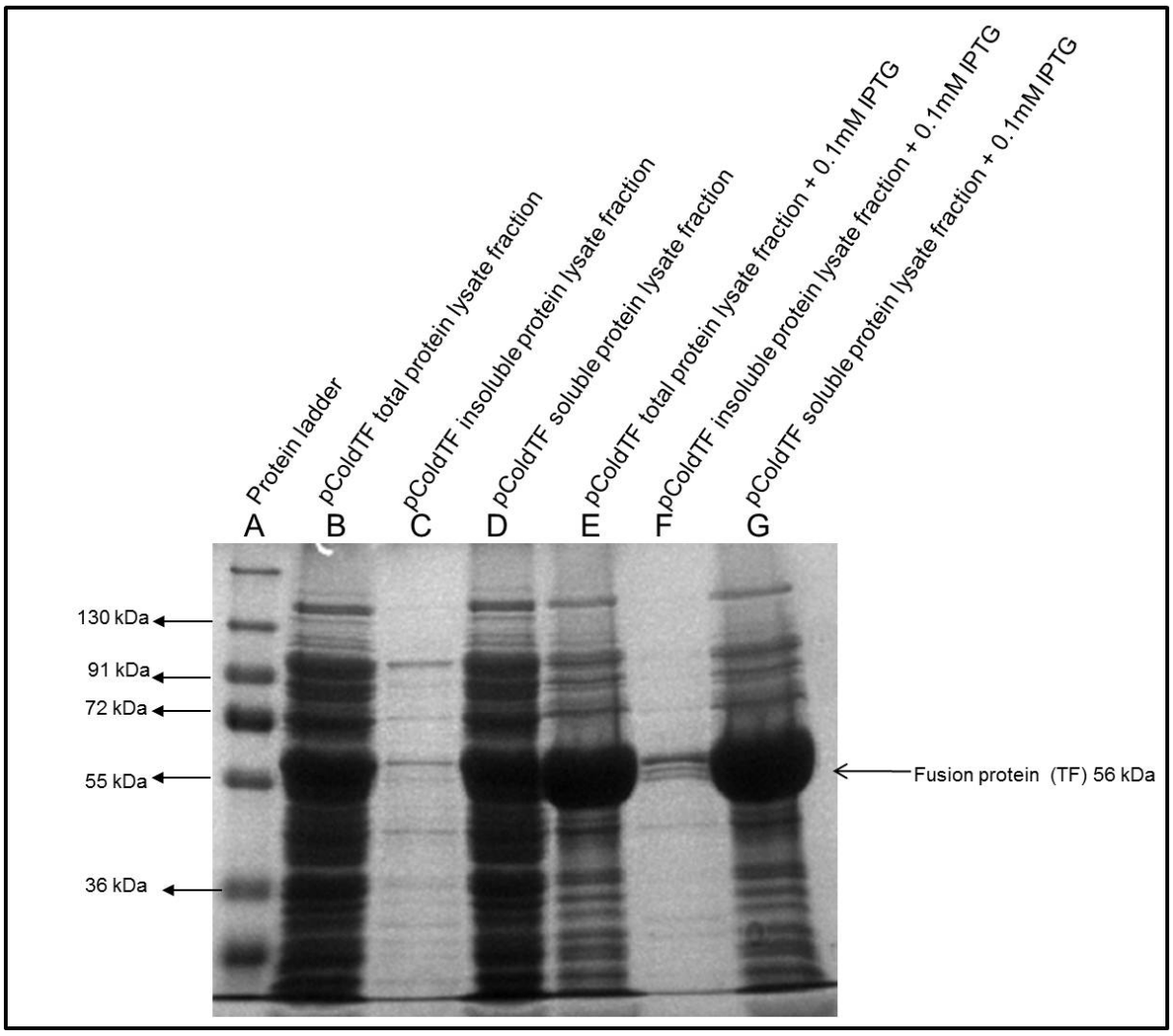



**Figure 22: SDS-PAGE analysis of protein expression of trigger factor (TF) from pColdTF in Origami™ cells with and without IPTG induction.** Lanes contain 20 µl sample consisting of 10 µl protein sample and 10 µl PSB. Lanes: A) 5 µl PageRuler™ Prestained Protein Ladder Plus; B) Uninduced, total fraction; C) Uninduced, insoluble fraction; D) Uninduced, soluble fraction; E) Induced with 0.1 mM IPTG, total fraction; F) Induced with 0.1 mM IPTG, insoluble fraction; G) Induced with 0.1 mM IPTG, soluble fraction.

In Figure 22, the total protein fraction that was not induced with IPTG (lane B) showed an over-expressed protein band at presumably 56 kDa and presumed to be the TF fusion protein. This was unexpected as the expression of recombinant proteins with fusion protein or fusion protein alone is strictly regulated by the *lac*I operator. IPTG induces β-galactosidase activity by binding to, and inhibiting the *lac* repressor. This can be due to leaky expression. With IPTG induction (Figure 22, lane E) the TF fusion protein at 56 kDa was highly expressed compared to the uninduced fraction (Figure 22, lane B). The total protein in the induced fraction in lane E was less than the uninduced fraction in lane B. When the protein fractions were separated into soluble (Figure 22, lane D & G) and insoluble (Figure 22, lane C & F) fractions, most of the proteins were soluble.

Figure 23 shows the expression of the bGLYAT with fusion protein in Origami™ cells transformed with pColdTF-bGLYAT with or without IPTG induction.

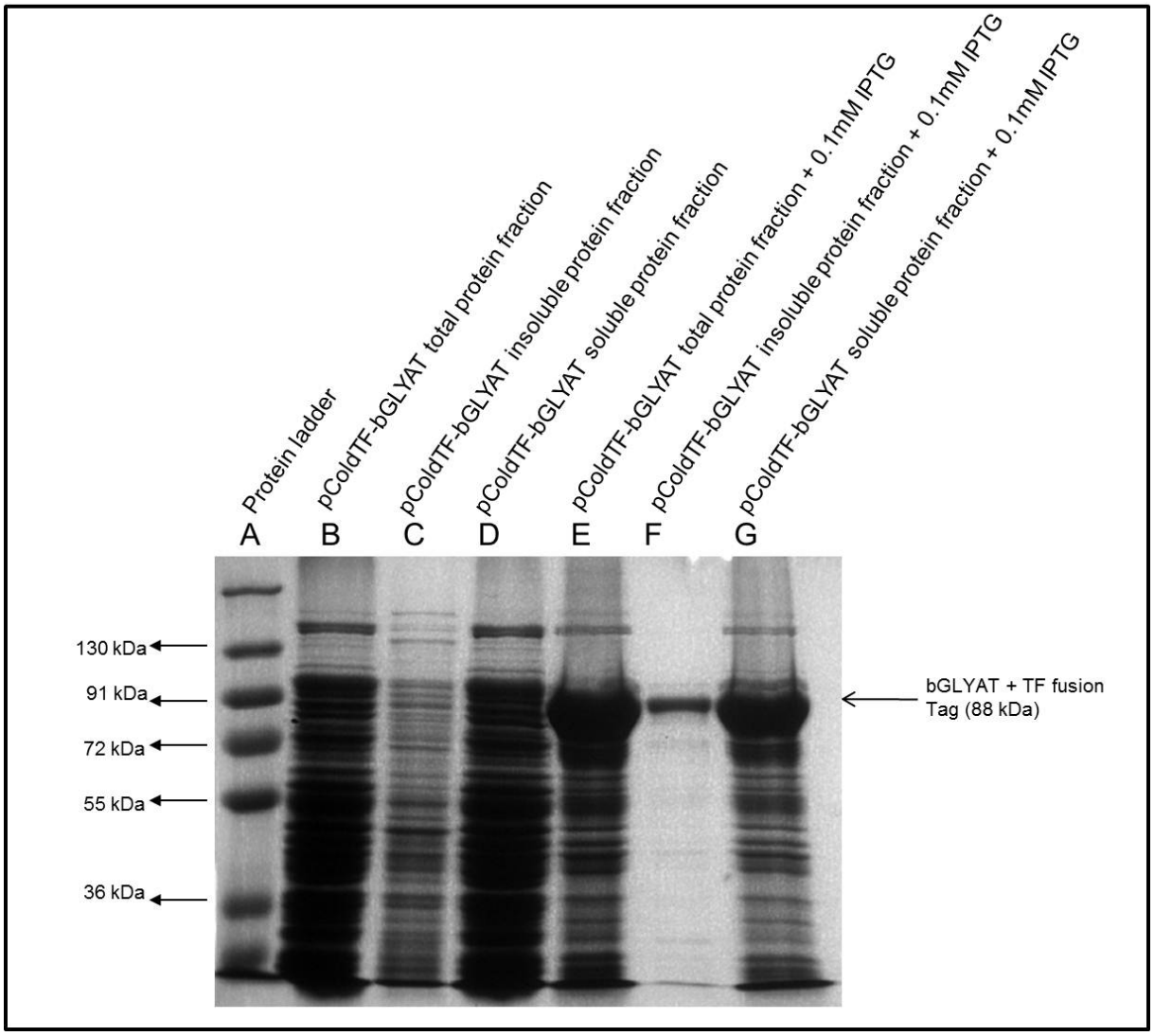



**Figure 23: SDS-PAGE analysis of protein expression of bGLYAT fused with trigger factor (TF) from pColdTF-bGLYAT in Origami™ cells with and without IPTG induction**. Lanes contain 20 µl sample consisting of 10 µl protein sample and 10 µl PSB. Lanes: A) 5 µl PageRuler™ Prestained Protein Ladder Plus; B) Uninduced, total fraction; C) Uninduced, insoluble fraction; D) Uninduced, soluble fraction; E) Induced with 0.1 mM IPTG, total fraction; F) Induced with 0.1 mM IPTG, insoluble fraction; G) Induced with 0.1 mM IPTG, soluble fraction.

In the total protein fraction that was induced with IPTG (Figure 23, lane E), presumably bGLYAT with fusion protein was over-expressed at 88 kDa compared to the uninduced culture (Figure 23, lane B). The bGLYAT with fusion protein is soluble (Figure 23, lane G). The total protein expression in the Origami™ cells was a bit less in the induced protein fraction (Figure 23, lane E) compared to the uninduced protein fraction (Figure 23, lane B). This is due to the cold shock step introduced by lowering

the temperature of the culture which caused the regular expression of *E. coli* housekeeping proteins to slow down, promoting better folding. Under IPTG induction, the *cspA* promoter (cold-shock protein) induced over-expression of the target protein due to IPTG inhibiting the *lac* repressor. A small amount of the recombinant bGLYAT with TF fusion protein (88 kDa) was partially produced as insoluble protein aggregates (Figure 23, lane F).

To summarise, under IPTG induction, expression of TF and associated fusion proteins is 56 kDa and soluble (Figure 22, lane G). bGLYAT with TF fusion protein is expressed to very high levels at 88kDa and also mostly soluble (Figure 23, lane G).

### 2.3.8 Purification and concentration of the bGLYAT fusion protein

Expression of a target protein in pColdTF allows purification with nickel column chromatography by means of the histidine tag which is expressed as a fusion with the protein. Therefore, the expressed bGLYAT-TF fusion contains a poly histidine tag. The soluble protein fraction (Figure 23, lane G) obtained from the Origami™ cells transformed with pColdTF-bGLYAT under 0.1 mM IPTG induction was loaded onto the Protino Ni-TED columns (Macherey Nagel) as described in Section 2.2.15. The expressed 88 kDa bGLYAT-TF fusion containing the his-tag was expected to bind to the nickel column. Expressed proteins such as housekeeping proteins that do not have a poly histidine tag would not have any affinity for nickel and would flow through the column and be washed away. Proteins that bound to the column were eluted with three consecutive 3 ml volumes of the elution buffer, containing imidazole. These three elution fractions were pooled together and the proteins were concentrated with VivaSpin Concentrators (Sartorius) as described in Section 2.2.16. It was expected that the combined eluted fraction should only contain the 88 kDa bGLYAT-TF fusion protein and that the recovery should be more that 90%. The proteins were analysed on 8% (w/v) SDS-PAGE as described in Section 2.2.3.3. Figure 24 shows the nickel affinity column purification and concentration of the 88 kDa bGLYAT-TF fusion protein.

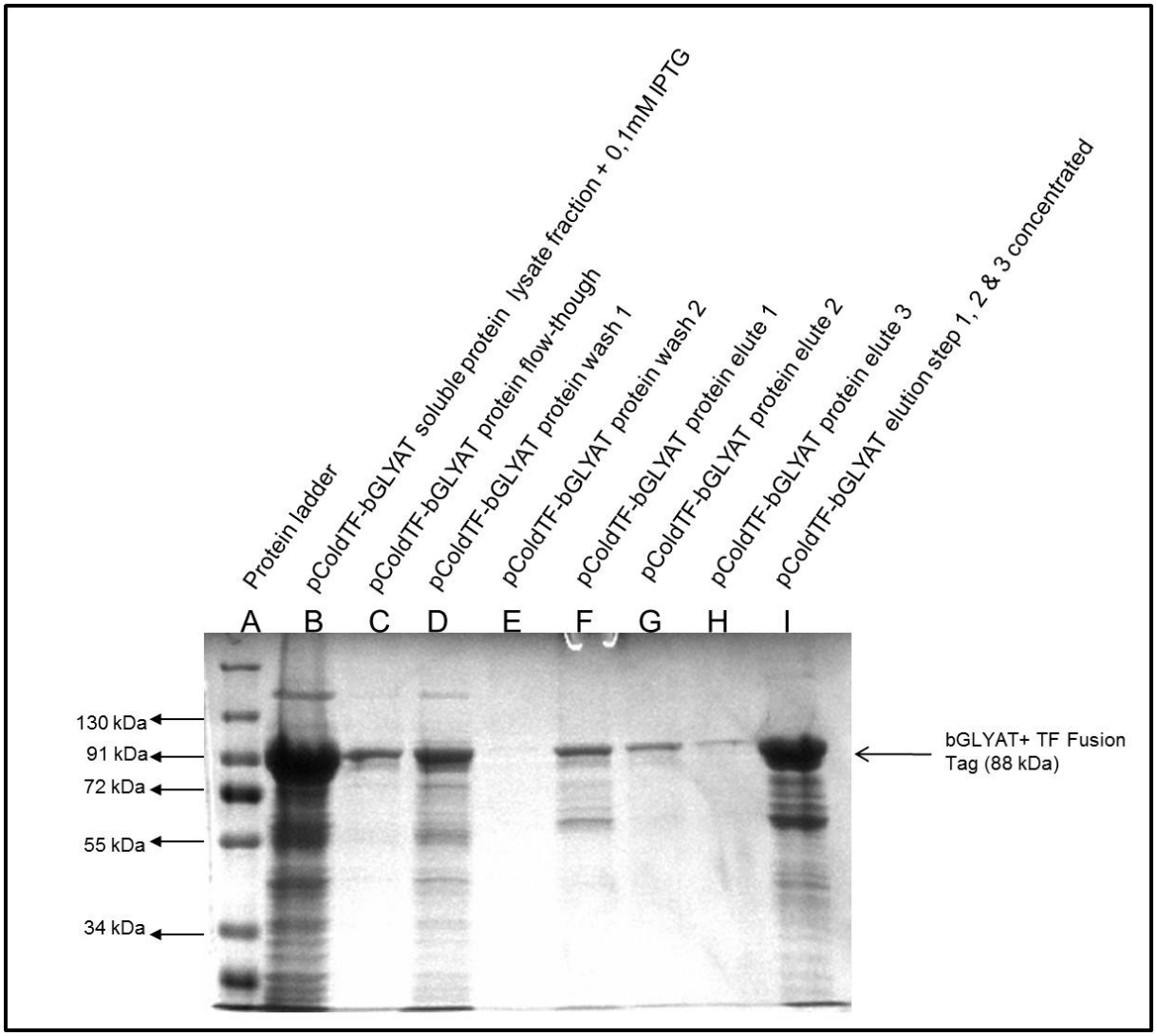



**Figure 24: SDS-PAGE analysis of purification and concentration of the soluble fraction of the IPTG induced protein fraction of bGLYAT fused with fusion protein (TF) expressed from pColdTF-bGLYAT.** Lanes contain 20 µl sample consisting of 10 µl protein sample and 10 µl PSB. Lanes: A) 5 µl PageRuler™ Prestained Protein Ladder Plus; B) Induced, soluble protein fraction; C) Flow-through; D) Wash step 1; E) Wash step 2; F) Elute step 1; G) Elute step 2; H) Elute step 3; I) Concentration of nickel affinity purified proteins containing a histidine fusion tag by combining elute sample 1, 2, and 3 (lanes F, G and H) and concentrations thereafter.

Figure 24, lane C showed that the 88 kDa bGLYAT-TF fusion protein band did not bind completely to the column and is found in the flow through. This could be due to overloading of the column. The protein binding capacity of the column was 5 mg protein. However, I did not quantify the protein content of the sample prior to loading it onto the column. Figure 24, lane D showed that some of the 88 kDa bGLYAT fusion protein was present in the first wash fraction, resulting in loss of the protein. All unbound proteins were washed away in the first wash step, and no proteins were visible in the second wash step (Figure 24, lane E). This indicated successful

washing of the bGLYAT fusion protein from the rest of the proteins expressed by Origami™ cells. Most proteins were eluted in the first elute step (Figure 24, lane F), less in the second step (Figure 24, lane G) and almost no proteins were visible in the third elution step (Figure 24, lane H). Imidazole therefore successfully replaced the his-tag proteins from the nickel affinity column for recovery. There are unknown protein bands smaller than the 88 kDa protein visible in the first elute (Figure 24, lane F). Since these proteins were not present in the second wash (Figure 24, lane D) but are visible after the first elution, it can be speculated that these proteins contain a histidine tag which enables them to bind to the nickel resin. Since the histidine tag is attached to the N-terminal, incompletely synthesized bGLYAT peptides should also be retained on the column. The only way to establish the origin of these smaller nickel binding peptides will be with additional analysis such as Western blotting. Due to the lack of a commercially available bGLYAT, Western blotting was not done. It was speculated that the bGLYAT-TF fusion protein was not fully translated. Figure 24, lane I illustrates the 88kDa bGLYAT-TF fusion after concentration of the three combined elution fractions. The protein was well recovered. This would indicate that if proteins were 90% recovered as expected, the 88 kDa band protein should be almost the same intensity as the purified and concentrated protein. Of the 1.5 ml extract that was loaded onto the column, the proteins were eluted in a total of 14 ml elution buffer. After concentration, the volume was 600 µl. The protein bands are almost the same intensity and concentration was thus successful. The next step was to remove the fusion tag from the recombinant bGLYAT.

### 2.3.9 Removal of the TF fusion protein from recombinant expressed bGLYAT in pColdTF

The fusion protein of the bGLYAT expressed contains a thrombin cleavage site. This protease was used in order to remove the fusion protein containing TF from bGLYAT (Section 2.2.1 – pColdTF vector map and multiple cloning site). After thrombin cleavage (Section 2.2.18), it was expected that the TEE, His-tag protein, TF protein, HRV 3C protease and thrombin cleavage site would be removed from the bGLYAT protein. TF is a 50 kDa protein. After thrombin cleavage, it was expected that the bGLYAT protein would be 34.7 kDa. The removed fusion protein was expected to be 53.3 kDa. The results (Figure 25) were analysed on 12% (w/v) SDS-PAGE as described in Section 2.2.3.3.

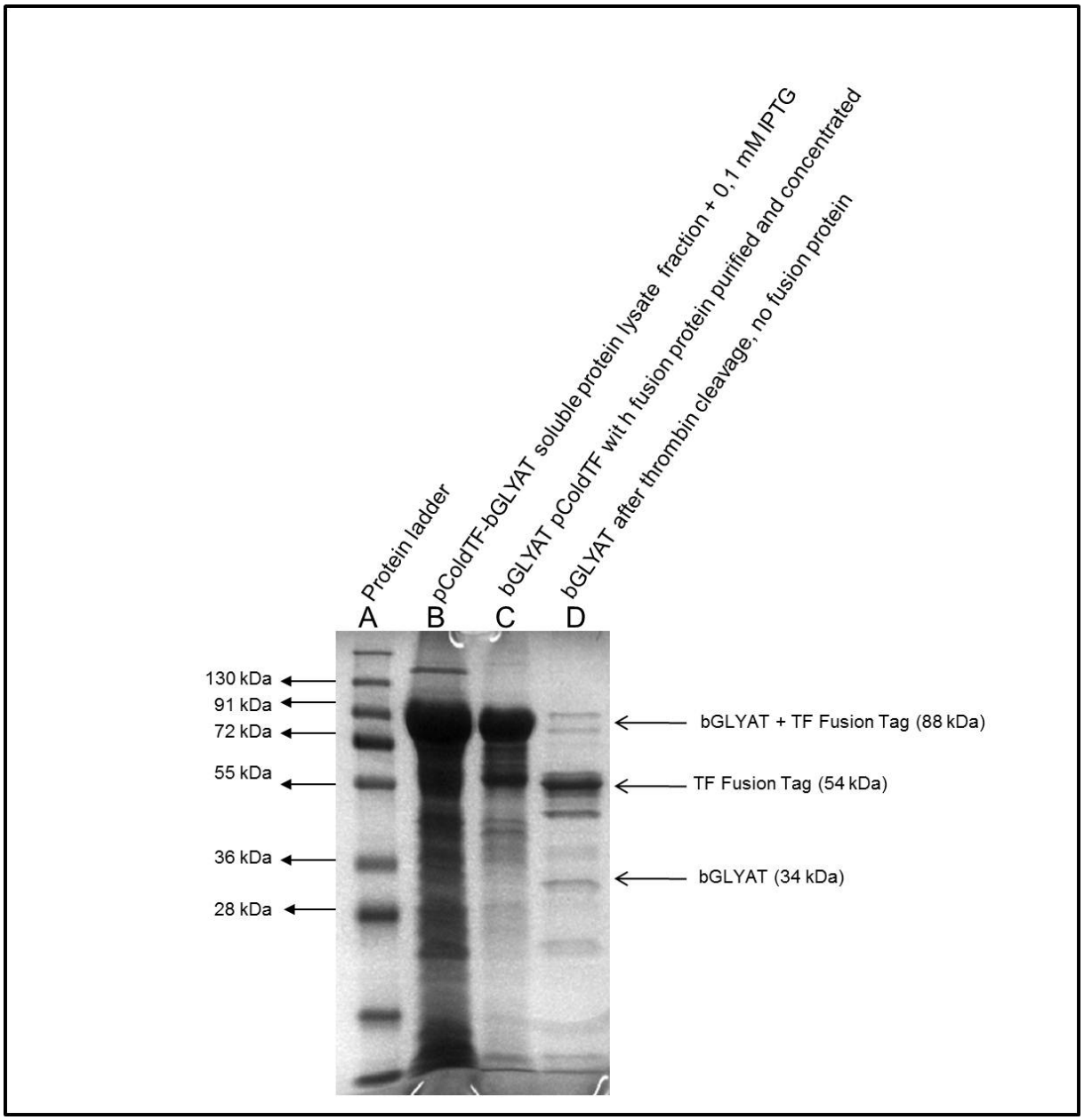



**Figure 25: SDS-PAGE analysis of expressed bGLYAT-TF fusion protein directly after harvesting the cells, purification with nickel affinity chromatography and concentration of the protein and then thrombin cleavage to remove the fusion protein.** Lanes contain 20 µl sample consisting of 10 µl protein sample and 10 µl PSB. Lanes: A) 5 µl PageRuler™ Prestained Protein Ladder Plus; B) Induced, soluble fraction; C) bGLYAT-TF fusion after nickel affinity purification and concentration; D) bGLYAT after thrombin cleavage.

In Figure 25, lane B, the soluble protein fraction of the 88 kDa bGLYAT-TF fusion was visible. The purification and concentration of the bGLYAT-TF fusion is seen in Figure 25, lane C. After thrombin cleavage (Figure 25, lane D), a distinct band was visible at presumably 53.3 kDa that correlated to the expected size of the fusion protein containing TF that was supposed to be removed from the bGLYAT protein.

There was also a band at presumably 36 kDa that could be bGLYAT protein. Another two protein bands at about 72 and 91 kDa were visible. The top band is expected to be bGLYAT with fusion protein at 88 kDa that was partially digested with thrombin. The lower band above the 70 kDa protein marker could be the partially expressed bGLYAT with fusion protein. This could also perhaps explain the unknown band smaller than 28 kDa.

The bovine GLYAT with the fusion protein removed was not analysed on SDS-PAGE to determine if the protein is soluble or not. If bGLYAT was not soluble, it would be suspected that bGLYAT would not have any enzyme activity.

### 2.3.10 GLYAT enzyme activity assay

In order to have a positive control for the recombinant bGLYAT enzyme activity assay, a crude cytoplasmic extract containing bovine GLYAT was prepared from bovine liver by Mr JHJ Fourie (2006). The result was analysed on 12% (w/v) SDS-PAGE (Section 2.2.3.3). According to Van der Westhuizen, *et al*. (2000) the molecular weight of bGLYAT is 36 kDa. The bovine liver crude cytoplasmic extract containing bGLYAT was visualised in Figure 26.

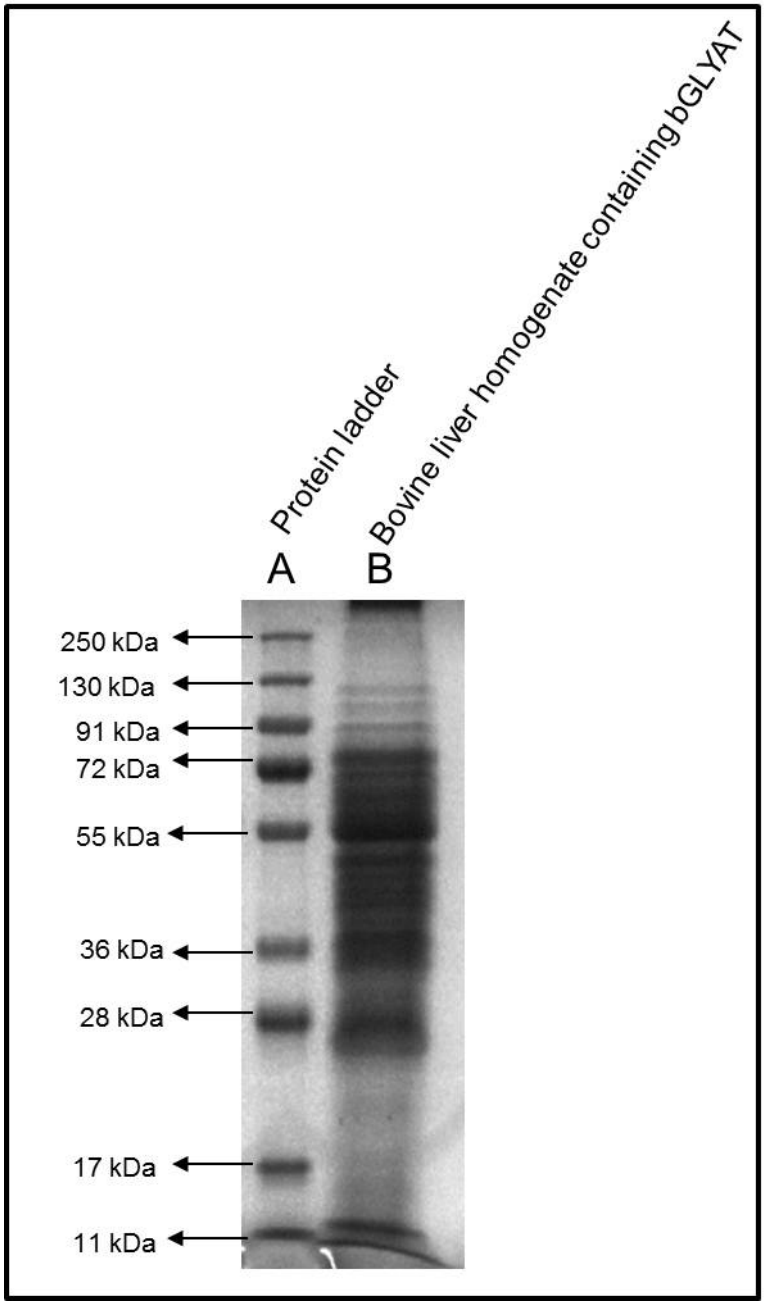



**Figure 26: SDS-PAGE analysis of a crude cytoplasmic extract from bovine liver containing bGLYAT.** Lanes: A) 5 µl PageRuler™ Prestained Protein Ladder Plus; B) Crude cytoplasmic extract of bovine liver containing bGLYAT prepared by Mr JHJ Fourie in 2006 used as control in all enzyme activity assays in this study.

In Figure 26, the molecular weight of the majority of proteins present in the bovine liver extract ranges from 25 kDa to 75 kDa.

### 2.3.10.1 GLYAT enzyme activity assay on a crude cytoplasmic extract of bovine liver

In order to have a positive control for GLYAT activity, a crude extract of cytoplasmic bovine liver containing bGLYAT was prepared by Mr JHJ Fourie, a fellow student.

The concentration of the preparation was determined by the BCA protein assay described in 2.2.17. The enzyme activity of GLYAT was measured as described by Kolvraa and Gregersen in 1986 (Section 2.2.19). In Figure 27, the GLYAT activity of the bovine liver crude cytoplasmic extract containing bGLYAT was measured using a standard dilution series (1:2, 1:5, 1:10, 1:100).

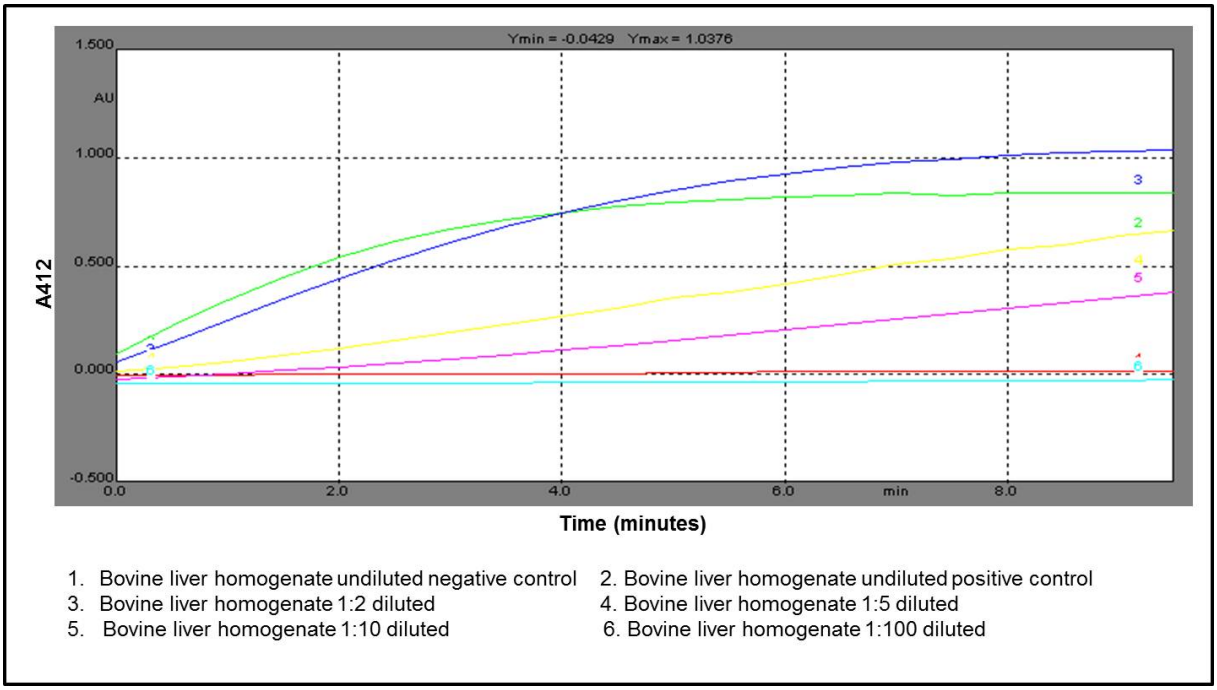



**Figure 27: GLYAT enzyme assay of a dilution series of the crude cytoplasmic extract of bovine liver containing bGLYAT.** Graph lines: 1) Bovine liver homogenate undiluted negative control by omitting glycine, red line. 2) Cytoplasmic extract of bovine liver not diluted, green line. 3) Cytoplasmic extract of bovine liver two times diluted, blue line. 4) Cytoplasmic extract of bovine liver five times diluted, yellow line. 5) Cytoplasmic extract of bovine liver ten times diluted, pink line. 6) Cytoplasmic extract of bovine liver hundred times diluted, turquoise line.

Figure 27 shows the GLYAT enzyme activity of the dilution range of the crude cytoplasmic extract of bovine liver containing bGLYAT. The negative control, (red line) did not show any enzyme activity. The positive control (green line) is the crude cytoplasmic extract of bovine liver and as expected, had GLYAT activity. The samples with the highest protein content had the highest GLYAT activity and decreased as the dilution increases (protein content decreased). Even though the undiluted form (green line) has a lower activity that the 1:2 diluted sample (blue line), the initial velocity of the undiluted reaction is still greater. Hereafter, GLYAT activity of the recombinant bGLYAT had to be tested and compared to the liver crude cytoplasmic extract containing bGLYAT.

### 2.3.10.2 GLYAT enzyme activity assay on nickel affinity purified bGLYAT expressed as a His-tagged fusion protein with TF

The recombinant bGLYAT-TF fusion protein encoded by pColdTF-bGLYAT was expressed in Origami™ bacterial cells. The soluble fraction of the bGLYAT-TF fusion protein was purified by nickel affinity chromatography and concentrated by membrane size exclusion spin chromatography. The GLYAT enzyme activity of the fraction explained above was determined as explained in Section 2.2.19. This will indicate if TF was able to fold the bGLYAT correctly and also whether any other protein except bGLYAT had GLYAT activity. The bovine liver cytoplasmic extract containing bGLYAT obtained from Mr JHJ Fourie was the positive control. It was expected that protein expressed from empty pColdTF should not have any GLYAT activity, seeing that GLYAT is not expressed. It was also hoped that bGLYAT with the TF fusion protein should have enzyme activity seeing that GLYAT was expressed as a soluble protein. Negative controls were set up by omitting glycine from the reaction.

The bGLYAT concentration in the liver crude cytoplasmic extract could not be determined as it does not have any tag for purification. The only way to compare the GLYAT activity in different samples was by means of total protein concentration determination. The BCA protein assay was performed according to Section 2.2.17. The bovine liver cytoplasmic extract had a protein concentration of 3.1 µg/µl. The soluble bacterial cell lysate concentration of proteins expressed from empty pColdTF was 4.8 µg/µl. The soluble bacterial cell lysate concentration of proteins expressed from pColdTF-bGLYAT was 5.4 µg/µl. After this fraction was purified and concentrated, the total protein concentration was 1.9 µg/µl. It is evident that protein concentration was lower after purification and concentration, although, in Figure 24 this step was adequate to retain enough putative bGLYAT-TF fusion protein. The results comparing GLYAT activity in expressed protein from empty pColdTF and pColdTF-bGLYAT is shown in Figure 28.

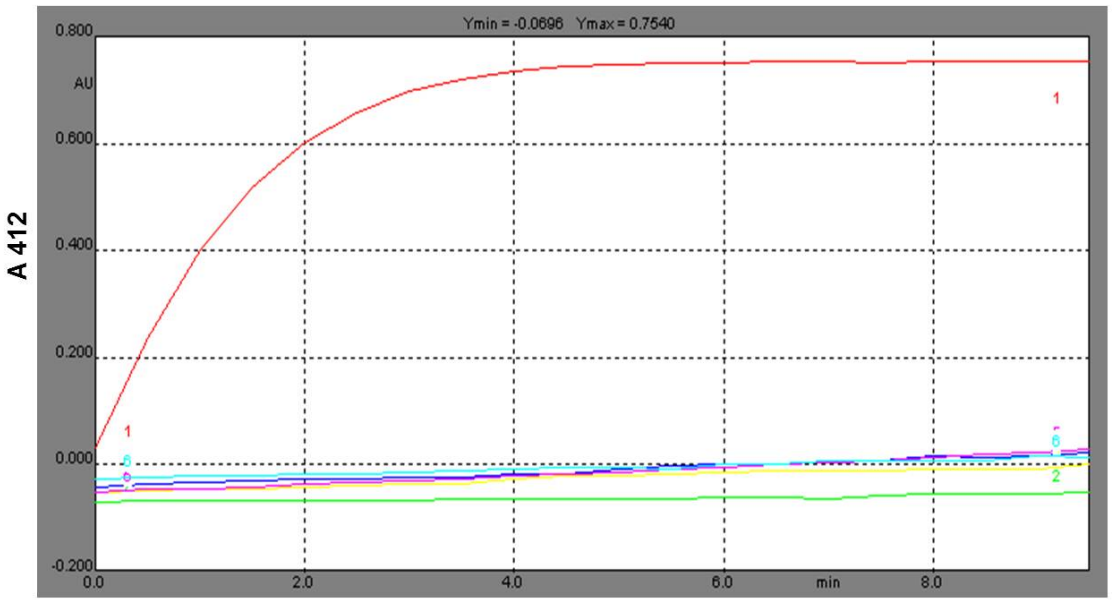



**Figure 28: GLYAT enzyme activity assay of the soluble fraction of an Origami™ cell lysate expressing bGLYAT-TF fusion protein.** Graph lines: 1) Bovine liver crude cytoplasmic extract as positive control; 2) Bovine liver crude cytoplasmic extract as negative control by omitting glycine; 3 – 6) Expressed proteins from soluble fraction of Origami™ cell lysates encoded by; 3) empty pColdTF and other proteins; 4) empty pColdTF and other proteins, omitting glycine, 5) pColdTF-bGLYAT expressing bGLYAT-TF fusion protein and other proteins; 6) pColdTF-bGLYAT expressing bGLYAT-TF fusion protein and other proteins, omitting glycine.

Figure 28 showed GLYAT enzyme activity in the positive control, which indicated that the assay was working. The negative controls (omitting glycine) were all negative which meant that GLYAT enzyme activity is glycine dependent as expected. No GLYAT activity was detected in the fraction containing soluble proteins expressed from Origami™ cells encoded by pColdTF or pColdTF-bGLYAT. Seeing that the bGLYAT-TF fusion protein was soluble, but did not have enzyme activity, and the TF fusion protein (56 kDa) is much bigger than bGLYAT (34.7 kDA), the active site of bGLYAT might be obstructed by the fusion protein. From the results, it is not possible to determine why the soluble recombinant bGLYAT-TF fusion protein does not have GLYAT activity. A logical next step in pursuit of an enzymatically active recombinant bGLYAT would be to remove the TF part by means of thrombin cleavage.

#### 2.3.10.3 Comparing GLYAT enzyme activity of bGLYAT with and without the TF fusion protein

To determine if recombinant bGLYAT was enzymatically active without the TF fusion protein, the fusion protein was removed from bGLYAT by means of thrombin cleavage (Section 2.2.18). The TF fusion protein was situated on the N-terminal. The His-tag would also be removed by thrombin cleavage. There was thus no way to purify the bGLYAT protein from the TF fusion protein. After thrombin cleavage was performed and the proteins concentrated with size exclusion spin chromatography (Section 2.2.16), the protein concentration determined by a BCA assay (Section 2.2.17) was 1 µg/µl. According to these results, approximately 13 µg proteins were used in the GLYAT enzyme activity, comparing the bovine liver crude cytoplasmic extract, to the His-tag purified bGLYAT-TF fusion protein and bGLYAT where the TF fusion protein was removed by thrombin cleavage. The result is shown in Figure 29.

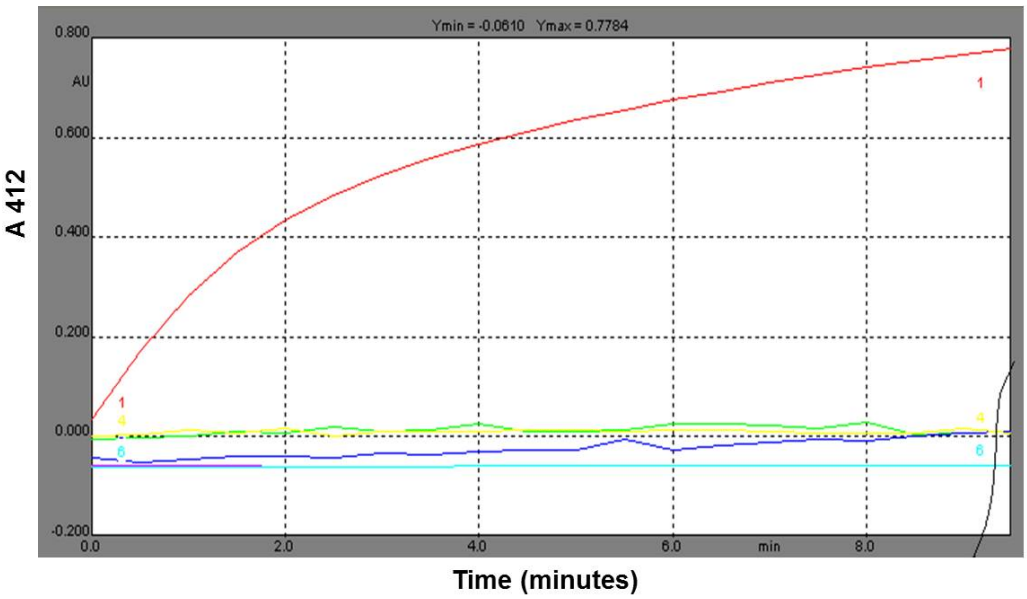



**Figure 29: GLYAT enzyme activity assay of the purified recombinant bGLYAT-TF fusion protein and the recombinant bGLYAT from which the TF fusion protein was removed;** Graph lines: GLYAT activity in; 1) Bovine liver crude cytoplasmic extract as positive control; 2) Bovine liver crude cytoplasmic extract as negative control by omitting glycine; 3) Purified recombinant bGLYAT from Origami™ cell lysate encoded by pColdTF-bGLYAT; 4) Purified recombinant bGLYAT from Origami™ cell lysate encoded by pColdTF-bGLYAT, omitting glycine; 5) Recombinant bGLYAT with fusion protein removed; 6) Recombinant bGLYAT with fusion protein removed, omitting glycine.

Figure 29 showed GLYAT enzyme activity in the positive control. The bovine liver crude cytoplasmic extract showed GLYAT activity, indicating that the assay worked. No GLYAT enzyme activity was detected in any of the other samples. bGLYAT-TF fusion protein was soluble, but did not have any enzyme activity. After removal of the N-terminal tag containing TF, bGLYAT still did not have any enzyme activity. In Section 2.3.9 it was mentioned that the recombinant bGLYAT was not tested on SDS-PAGE to determine whether bGLYAT was still soluble after the TF fusion protein was removed by thrombin cleavage. The lack of enzyme activity in the recombinant expressed bGLYAT with the removed TF fusion protein could be due to insufficient folding of the protein. TF is a chaperone-like protein, but might not have been able to correctly fold the bGLYAT expressed in Origami™ cells which resulted

in an inactive enzyme. New strategies had to be followed in order to correctly fold the recombinant bGLYAT protein for it to have GLYAT activity.

## 2.4 Summary

In order to generate a recombinant bovine GLYAT a bovine liver had to be obtained from the abattoir and total RNA extracted from a sample thereof. The messenger RNA (mRNA) included in the total RNA extracted served as the template for reverse transcriptase in order to prepare a cDNA library. Specific bGLYAT primers containing appropriate restrictions sites amplified GLYAT-specific cDNA from the pool of cDNA. Sequencing analysis confirmed that the bGLYAT PCR amplicon was successfully cloned into pColdTF in the correct orientation and open reading frame for expression.

bGLYAT was expressed as a 88 kDa fusion protein with TF and was mostly soluble. The soluble bGLYAT-TF fusion protein was purified with nickel column chromatography (by virtue of the His-tag) and concentrated by size exclusion spin chromatography. The recombinant bGLYAT did not have any GLYAT activity although it was a soluble fusion protein. It was speculated that TF might interfere with the bGLYAT active site.

After removal of the N-terminal tag containing TF, recombinant bGLYAT (34 kDa) still did not have any enzyme activity. An oversight in this work was that I did not test whether bGLYAT of which the tag was removed was still soluble. The results were inconclusive whether the TF obstructed the active site of recombinant bGLYAT, or that bGLYAT was indeed insoluble. The lack of enzyme activity in the recombinant expressed bGLYAT with the removed TF fusion protein could also be due to insufficient folding of the protein.

The next step was to express bGLYAT without a fusion protein, thereby eliminating the possibility of it affecting enzyme activity and co-expressing the chaperones instead. This approach might facilitate correct folding of bGLYAT and thereby produce a soluble enzymatically active bGLYAT.

# CHAPTER 3

# CLONING AND EXPRESSION OF bGLYAT USING pCOLDIII AND CO-EXPRESSION WITH THE CHAPERONES TF AND GroEL-GroES SEPARATELY

## 3.1 Introduction

As mentioned in Chapter 2, bacterial expression is the most cost effective and easy way to express recombinant proteins. The pCold vector system for expression of recombinant proteins in bacterial cells was explained in Chapter 2. Furthermore, in Chapter 2, the pColdTF vector was used successfully for the expression of recombinant bGLYAT as a soluble fusion protein with the TF chaperone. Unfortunately, neither the soluble bGLYAT-TF fusion protein nor the bGLYAT from which the TF fusion was removed had any enzyme activity. However, there is a range of pCold expression plasmids (Table 3) that contains various combinations of the TEE, His tag, Factor Xa cleavage site and TF.

**Table 3: Features of the Takara pCold expression vectors (Takara Bio Inc, 2009).**

| | TEE | His Tag | Factor Xa Cleavage Site | Trigger Factor |
|---|---|---|---|---|
| **pCold™I DNA** | O | O | O | X |
| **pCold™II DNA** | O | O | X | X |
| **pCold™III DNA** | O | X | X | X |
| **pCold™IV DNA** | X | X | X | X |
| **pCold™TF DNA** | O | O | O | O |

**TEE = Transcription enhancement element; His Tag = Histidine Tag; O = contain; X = does not contain**

In Chapter 2, I used pColdTF since it contains all the features mentioned in Table 4. In this Chapter I investigated the expression of bGLYAT using the pColdIII system. This system allows for co-expression of recombinant proteins with different

combinations of chaperones. pColdIII contains only the TEE, and not the TF, His tag and Factor XA cleavage site (Table 4). Another reason for choosing this specific vector is that the recombinant bGLYAT can be expressed without any additional fusion proteins. In Chapter 2 it was speculated that the TF fusion and His tag might somehow adversely affect the enzyme activity of bGLYAT. The plasmid map and multiple cloning site of pColdIII together with chaperone plasmid pTf16 from which TF is expressed and pG-Tf2 from which TF and GroEL-GroES are expressed is shown in Figure 30.

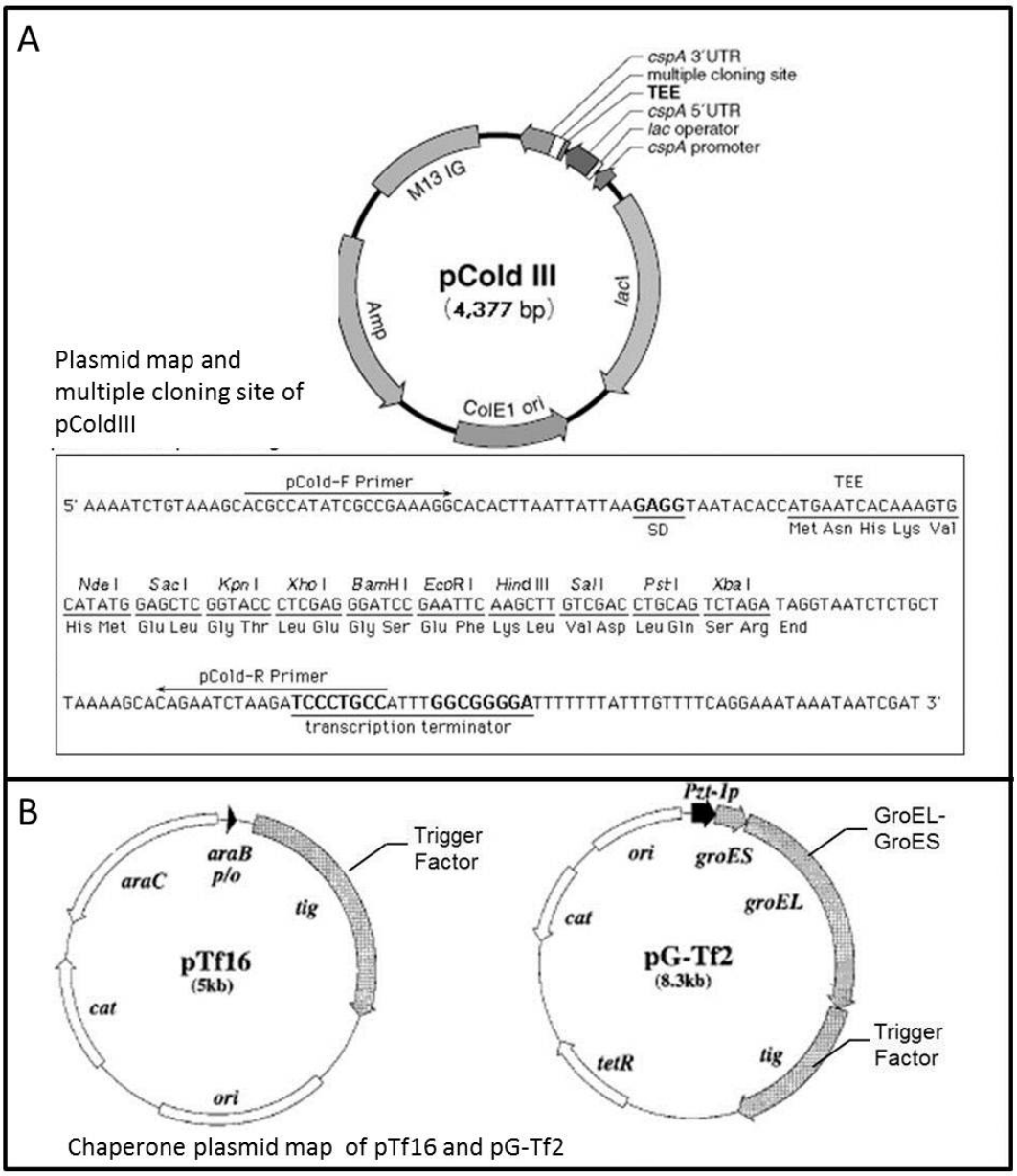



**Figure 30: A) Plasmid map and multiple cloning site of pColdIII; B) Map of plasmids for co-expression of chaperones pTf16 for expression of TF and pG-Tf2 for expression of TF and GroEL-GroES.**

Molecular chaperones have been demonstrated to be involved in the protein folding process. Co-expression of these chaperones with the target protein enable efficient folding of the expressed proteins. Over-expression of the molecular chaperones such as TF, GroEL-GroES and DnaK-DnaJ-GrpE facilitates protein folding and enhance the production of active enzymes. Refer to Figure 9 for illustrations of the possible model for chaperone assisted protein folding in *E. coli* (Thomas *et al.,* 1997). TF and GroEL-GroES have co-operative roles in assisting protein folding because TF binds to GroEL. TF and DnaK-DnaJ-GrpE have similar functions during protein folding because their function in folding nascent polypeptides overlaps partially (Nichihara, 2000). When the protein is not folded correctly, it is degraded. Trigger factor is expressed by pTf16 and TF and GroEL-GroES are expressed by pG-Tf2. In this chapter I present my results of experiments in which I used two vectors to investigate the effect of co-expression of the chaperones on recombinant bGLYAT expressed from pColdIII.

The experimental approach for the work presented in this chapter was as follows: The RNA in a sample of bovine liver was first stabilised. The tissue was then homogenised to extract total RNA. cDNA was synthesised from total RNA. The polymerase chain reaction (PCR) with specific bovine GLYAT primers was used to generate a bGLYAT amplicon which was then inserted into the pColdIII vector by means of directional cloning. Putative pColdIII-bGLYAT plasmid DNA was isolated and sequenced to confirm correct orientation and reading frame of bGLYAT in pColdIII. The recombinant bGLYAT protein was expressed by means of cold shock and IPTG induction. The recombinant protein was co-expressed with chaperones in order to investigate the folding effect of these chaperones on the enzyme activity of bGLYAT. TF from pTf16 was induced by L-arabinose and TF together with GroEL-GroES from pG-Tf2was induced by tetracycline. A GLYAT enzyme activity assay was performed on various protein bacterial lysates. The flow diagram with a summary of the experimental approach is presented in Figure 31.

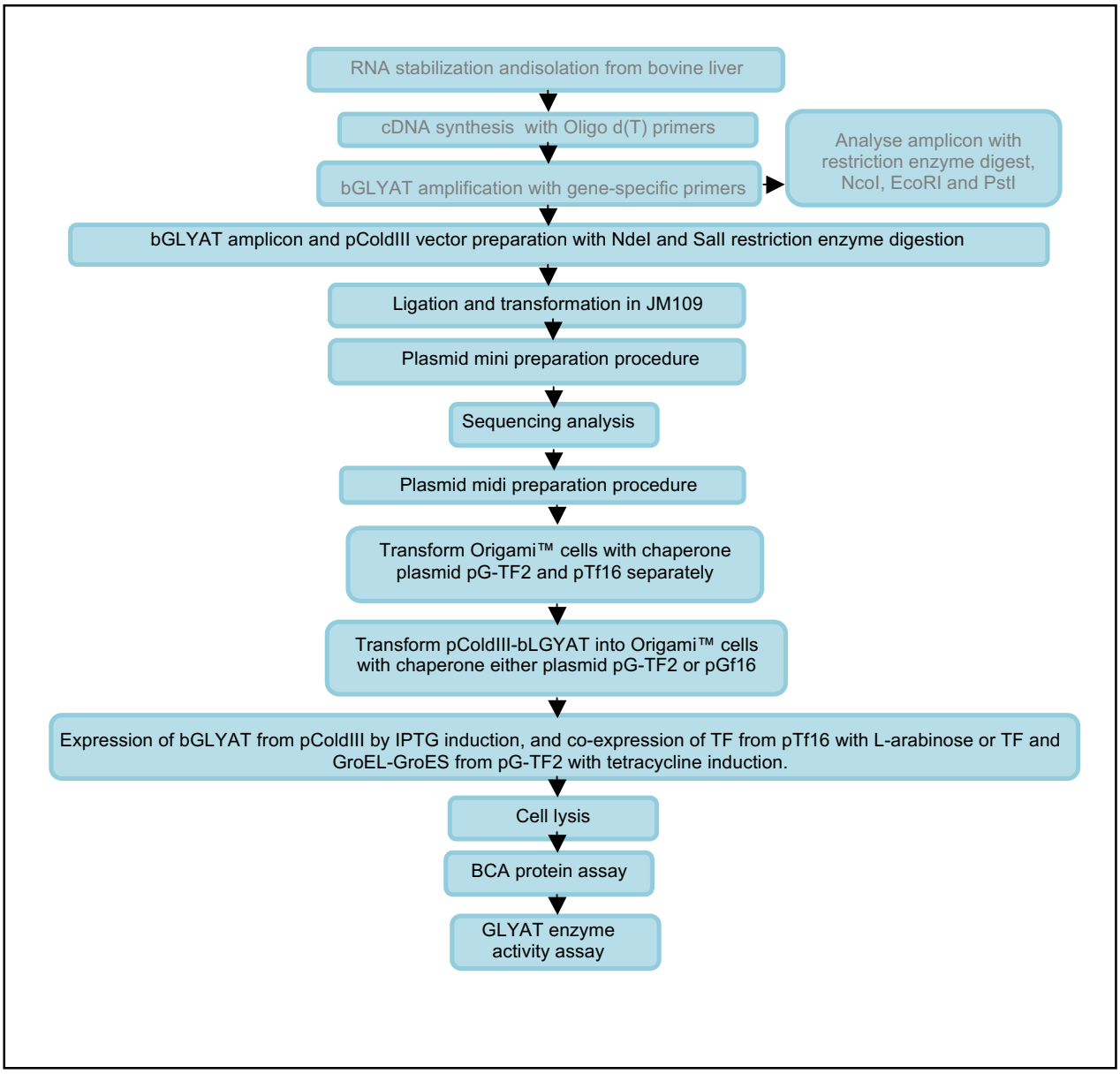



**Figure 31: Experimental approach for cloning the gene encoding glycine N-acyltransferase from bovine liver and co-express the recombinant bGLYAT with TF alone or TF together with GroEL-GroES.**

## 3.2 Material and methods

The experimental approach was to clone bGLYAT into pColdIII by means of directional cloning. Sequencing was done to confirm that the amplicon encoding bGLYAT was cloned into pColdIII in the correct orientation and that the open reading frame encoding bGLYAT remained intact. Origami™ cells were transformed with pColdIII-bGLYAT and pG-Tf2 or pTf16 separately. The recombinant bGLYAT was co-expressed with either TF alone or a combination of GroEL-GroES and TF as chaperones. The protein content of the bacterial cells was analysed by SDS-PAGE. The bGLYAT enzyme activity assay was carried out on lysates of bacteria cells and the protein fractions thereof. Lastly, it was investigated whether it is possible to solublize an insoluble aggregate of bGLYAT using different buffers.

### 3.2.1 Cloning of the amplicon encoding bGLYAT into pColdIII

The bGLYAT amplicon (generated as described in Chapter 2) and pColdIII plasmid were prepared for directional cloning by digesting with the restriction enzymes NdeI and SalI (Section 2.2.8) followed by gel extraction (Section 2.2.7). The ligation reaction of the prepared bGLYAT insert and pColdIII plasmid was set as described in Section 2.2.9. The plasmids were transformed into competent JM109 cells (Section 2.2.10.2). A bacterial colony containing a plasmid that might contain the bGLYAT insert was identified. The plasmid was isolated by means of a mini plasmid preparation procedure (Section 2.2.11.1). Restriction enzyme digestion (Section 2.2.8) with NdeI and SalI was used to determine whether the plasmid did contain the bGLYAT insert. The positive pColdIII-bGLYAT was sent for sequence analysis (Section 2.2.12).

### 3.2.2 Co-expression of bGLYAT and the chaperone TF and GroEL-GroES

The plasmids pTf16 encoding TF and pG-Tf2 encoding TF and GroEL-GroES were transformed separately into competent Origami™ cells. The cells were made competent as described in Section 2.2.10.1. All cultures of cells containing expression plasmids encoding the chaperones were propagated in medium containing 20 µg/µl chloramphenicol. Glycerol stocks were prepared for cells containing either pTF16 or pG-TF2 (Section 2.2.10.3).

The pColdIII-bGLYAT was then transformed into the competent Origami™ cells that already contained either pG-Tf2 or pTf16 (Section 2.2.10.1) in order to co-express bGLYAT either with TF alone or TF and GroEL-GroES (also refer to Section 2.2.13). Cells were plated out on LB agar plates. All cultures that contained both the pColdIII vector as well a chaperone plasmid were grown using media containing 20 µg/µl chloramphenicol and 100 µg/µl ampicllin.

Colonies were picked from the LB plates and grown overnight at 37ºC while shaking at 200 rpm in 5 ml LB media containing both antibiotics. These cultures were then transferred to 50 ml LB media with antibiotics. Cultures that contained chaperone plasmid pG-Tf2 were induced with 10 ng/ml tetracycline and cultures that contained chaperone plasmid pTf16 were induced with 4 mg/ml L-arabinose. Cells were cultured at 37ºC to an $OD_{600nm}$ of 0.5 and then chilled at 15ºC for 30 minutes.

Thereafter, the cultures were induced with 0.1 mM IPTG and incubated for 24 h at 15ºC shaking at 200 rpm. The proteins were extracted from the bacterial cells as described in Section 2.2.14. Bacterial cells were lysed and the protein fractions were analysed by SDS-PAGE (Section 2.2.3.2). A total protein fraction, soluble protein fraction and insoluble protein was analysed. The protein concentration was determined with the BCA assay explained in Section 2.2.17. The GLYAT enzyme activity assay was performed as described in Section 2.2.19.

### 3.2.3 Solubilisation of insoluble protein

Bovine GLYAT activity is low at low pH and the reaction rate increases rapidly from pH 6 to pH 7.5. At pH values above 10, a platau is reached. The rate does not decrease after the pH 10. This indicated the catalytic residue is a base such as glutamate, histidine and aspartate (Nandi *et al*., 1979; Mawal & Qureshi, 1994).

Four different pH buffers at pH 8, 9, 10 and 11 respectively were used to investigate whether it would be possible to solubilize insoluble aggregates of bGLYAT. The buffer N-Tris(hydromethyl)methyl-4-aminobutanesulfonic acid, is better known as TABS (Sigma-Aldrich; cat no T1302). TABS has a pH buffer range of 8.2 - 9.6. TABS buffer (0.1 M) was prepared at pH 8 and 9 respectively. The 3-(cyclohexyl amino)-1-propani sulfonic acid, is better known as CAPS buffer solution (Sigma-Aldrich; cat no C6070). CAPS has a useful pH range of 9.7 – 11.1. CAPS buffer (0.1 M) was

prepared at pH 10 and 11 respectively. The insoluble protein fraction in PBS of bGLYAT expressed from pColdIII co-expressed with either TF or TF together with GroEL-GroES respectively in Origami™ cell was pelleted by centrifugation at 13000 x g for 15 minutes at 4ºC. The supernatant was discarded. These pellets were resuspended in four different buffers at pH 8, 9, 10 and 11 respectively. The solution was mixed thoroughly by vortexing. The samples were again divided into soluble and insoluble fractions by means of centrifugation at 13000 x g for 15 minutes at 4ºC. The supernatant contained soluble proteins. The pellet was resuspended in PBS and represented the insoluble fraction. The proteins were analysed by SDS-PAGE as described in Section 2.2.3.3.

## 3.3 Results

### 3.3.1 Cloning of the amplicon encoding bGLYAT into the bacterial expression plasmid pColdIII

The remainder of the bGLYAT amplicon that was generated as described in Chapter 2 as well as the pColdIII plasmid were prepared for directional cloning by digestion with the restriction enzymes NdeI and SalI (Section 2.2.8) followed by gel extraction (Section 2.2.7). First the pColdIII vector was digested with NdeI and SalI separately to determine if both sites are cleavable. Then a double restriction digestion reaction with NdeI and SalI of both pColdIII and bGLYAT amplicon was performed followed by gel extraction. The bovine GLYAT amplicon was not digested with NdeI and SalI individually simply because the fragments removed with the restriction enzyme on the 3' or 5' site of the bGLYAT ORF would be too small (13 bp) to see on a 1% agarose gel. It was expected that the pColdIII plasmid linearized with NdeI or SalI would be 4377 bp. After double digestion of pColdIII with NdeI and SalI, the plasmid was expected to be 4336 bp. The bovine GLYAT PCR amplicon was calculated to be 907 bp. After double digestion of bGLYAT amplicon with NdeI and SalI, the amplicon was expected to be 894 bp. The results of the restriction enzyme digestion of the pColdIII and bGLYAT amplicon can be seen in Figure 32.

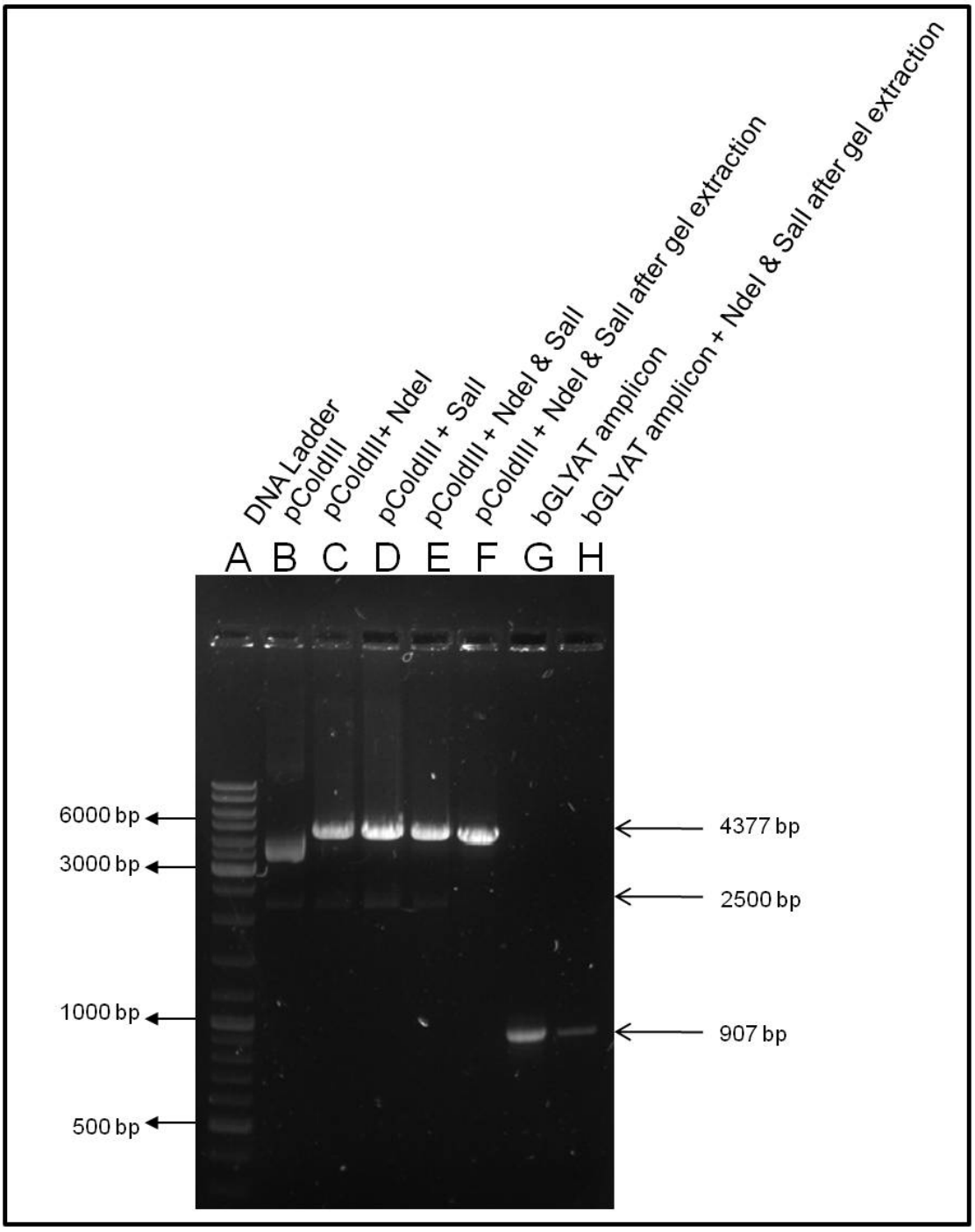



**Figure 32: Agarose gel electrophoretic analysis of NdeI and SalI digested pColdIII and the bGLYAT amplicon.** Lanes B-H contain 10 µl sample with 2 µl loading dye. Lanes: A) O'geneRuler DNA ladder; B) pColdIII undigested; C) pColdIII digested with NdeI; D) pColdIII digested with SalI; E) pColdIII digested with NdeI and SalI; F ) pColdIII digested with NdeI and SalI after gel extraction; G) bGLYAT amplicon undigested; H) bGLYAT amplicon digested with NdeI and SalI after gel extraction.

The undigested pColdIII plasmid DNA was predominantly in the supercoil form (Figure 32, lane B). NdeI and SalI digestion respectively resulted in a linearized fragment of pColdIII which was presumed to be 4377 bp (Figure 32, lanes C and D). With double digestion (Figure 32, lane E) it was not very clear that the fragment was

smaller, because the 41 bp removed was too small to be visible on a 1% agarose gel. In other words, it was difficult to separate the 4336 bp fragment form the 4377 bp fragment. In lane B, C, D and E there was a faint non-specific band at 2500 bp. This band could be due to contamination with some bacterial DNA. It was removed from the NdeI/SalI digested pColdIII with gel extraction as can be seen in (Figure 32, lane F). The 2500 bp non-specific band was not present in the bGLYAT samples. The bGLYAT amplicon migrated at about 900bp compared to the size ladder (Figure 32, lane F) which is very close to its actual size of 907 bp. Although after double digestion and gel extraction, the yield of the NdeI/SalI digested bGLYAT DNA was low, there was enough to continue with the cloning (Figure 32, lane H).

After preparation of both the pColdIII vector and bGLYAT insert for directional cloning, the fragments were ligated with each other as explained in Section 2.2.9. The plasmid construct was then transformed into competent JM109 cells (Section 2.2.10.2). Of the 200 µl cells that were plated on each LB Amp plate, colonies were detected as follows: There were three colonies on the ligation control, 81 possible recombinant pColdIII-bGLYAT containing colonies were obtained and more than 500 colonies grew on the control plate which contained the JM109 cells transformed with pColdIII. The transformation efficiency as calculated according to Section 2.2.10.1 was 5 x $10^5$ colonies per µg DNA. This is a low value compared to the standard good value of 1 x $10^9$, but served its purpose. The three colonies that were on the negative control could be due to undigested pColdIII plasmid, or self ligation of the vector on itself if one of the restriction enzyme digests was not complete. Random colonies were selected to perform mini plasmid preparations as described in Section 2.2.11.1. A series of plasmids were analysed on a 1% agarose gel (results not shown). Two plasmids (A and B) were selected, that were bigger than pColdIII. They were digested with NdeI and SalI. It was expected that if the vector contained the DNA encoding bGLYAT insert, it would be possible to cut the insert from the vector with the same restriction enzymes. The size of the vector is 4337 bp and the bGLYAT DNA is 894 bp. The result is shown in Figure 33.

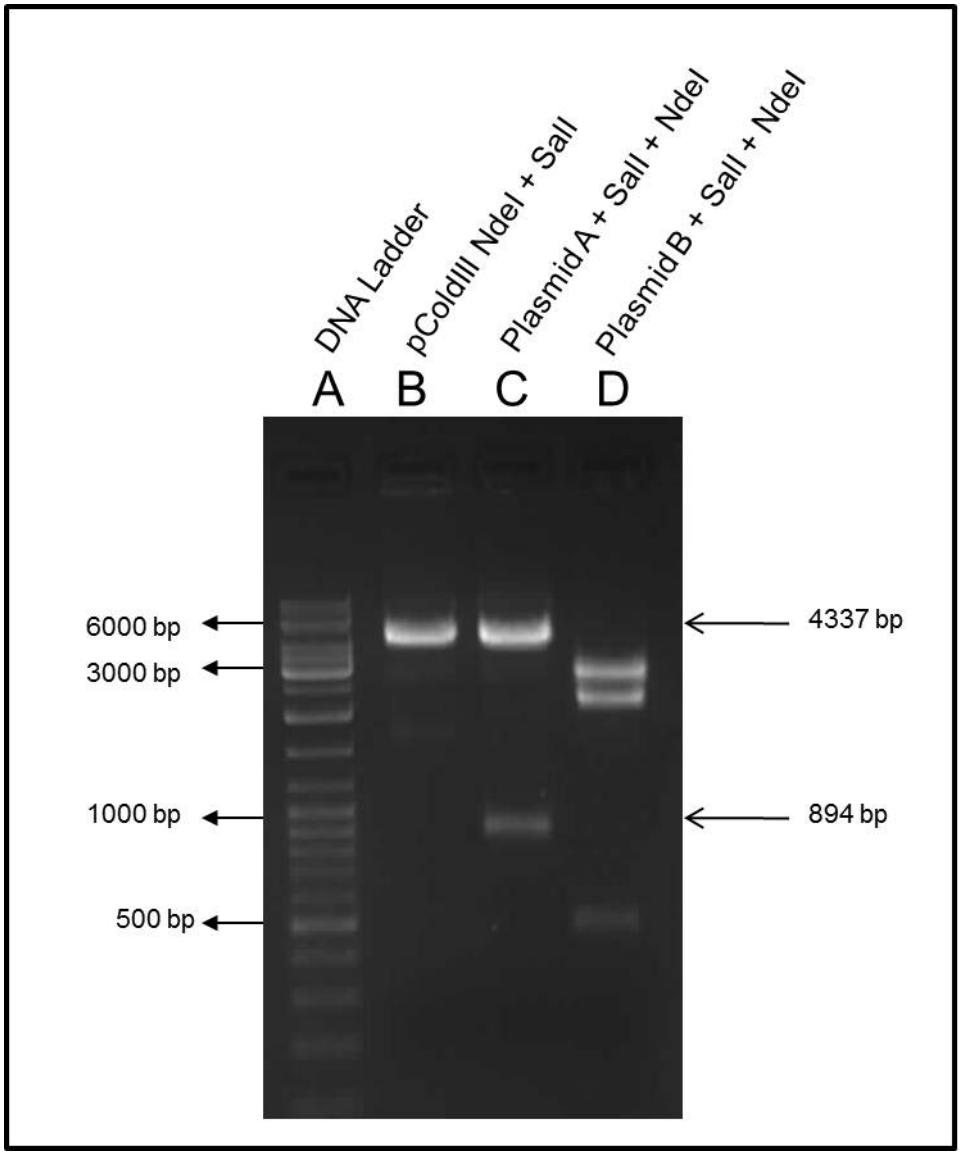



**Figure 33: Agarose gel electrophoretic analysis of restriction enzyme digestion of 2 plasmids obtained after cloning the bGLYAT amplicon into pColdIII.** Lanes B-D contains 10 µl sample with 2 µl loading dye. Lane A) 5 µl O'geneRuler DNA ladder; Lane B) pColdIII double digested with NdeI and SalI; Lane C) Plasmid A double digested with NdeI and SalI; D) Plasmid B double digested with NdeI and SalI.

Lane B in Figure 33 was the empty pColdIII double digested with restriction enzymes NdeI and SalI. The vector is approximately 4000bp that correlated to the expected results of 4337 bp. After NdeI and SalI digestion of plasmid A, two bands were visible (Figure 33, lane C). One was about 4337 bp which was the size of the vector, pColdIII. The other was presumed to be the bGLYAT insert at 894 bp (Figure 33, Lane C). Plasmid B gave completely different results from what was expected. SalI and NdeI double digestion of plasmid B resulted in three DNA bands at 3000bp, 2000bp and 500bp and is thus a 5500bp plasmid (Figure 33, lane D). This was

indeed close to the total size of the correct pColdIII-bGLYAT construct, but it was not further investigated and was discarded since it did not contain the bGLYAT 907 bp encoding insert. The next step was to determine if the construction of pColdIII-bGLYAT was successful. This was done by nucleotide sequencing analysis of plasmid A.

### 3.3.2 Nucleotide sequencing analysis of the 900 bp DNA fragment cloned into pColdIII

Restriction enzyme digestion identified plasmid A as a plasmid into which bGLYAT encoding DNA was cloned into pColdIII. A 50 ml LB culture containing plasmid A was prepared and propagated overnight. Plasmid DNA was isolated from it using a midi plasmid extraction kit (Section 2.2.11.2). Some of the plasmid DNA was sent for sequencing to Inqaba Biotech as described in Section 2.2.12. The rest of the plasmid A stock was aliquoted and stored at -80ºC. The forward primer bGLYAT-forN-NdeI with the sequence 5'-GCCGCATATGATGTTCCTGCTGC-3' was used with pColdIII reverse primer called pCold-R with the sequence 5'-CAGAATCTAAGATCCCTGCC3' to obtain the nucleotide sequence of the approximate 900bp bGLYAT encoding DNA fragment. The chromatogram and nucleotide sequences of sequenced strands of plasmid A were obtained from the Inqaba Biotec website. Vector NTi software was used to align both sequences to that of the GenBank bGLYAT reference sequence obtained from NCBI (NM177513). See Figure 34 for alignment.

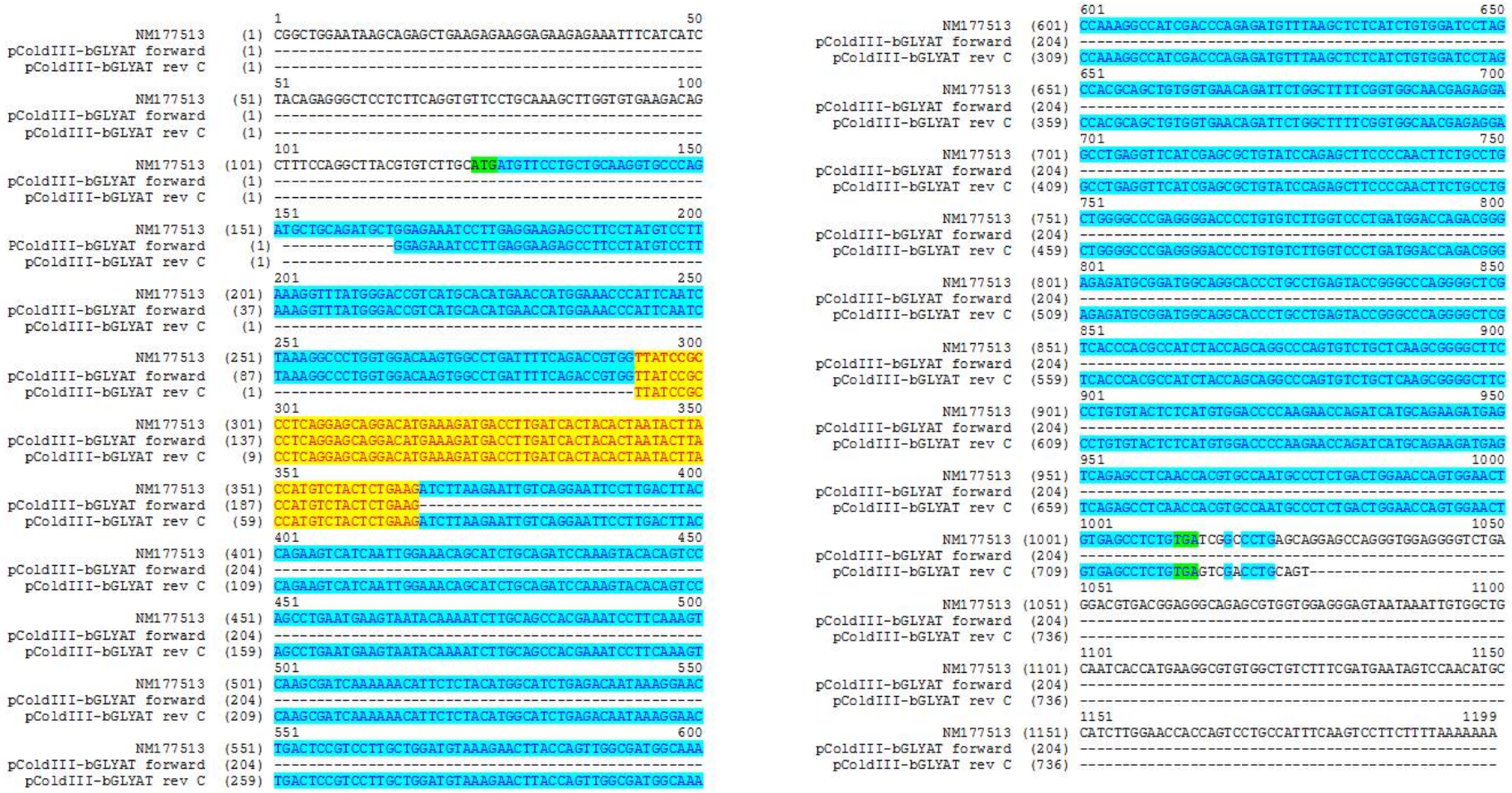



**Figure 34: Forward and reverse sequence alignment of the putative bGLYAT DNA of pColdIII-bGLYAT to a GenBank bGLYAT reference NM177513.** The open reading from of bGLYAT in the reference sequence (NM177513) starts at the green codon sequence highlighted and ends at the stop green highlighted codon sequence. The overlapping sequence of the forward and reverse primer is highlighted in yellow. The blue sequence highlighted indicates sequence similarity compared to the reference sequence NM177513.

.

Figure 34 shows the alignment of the open reading frame encoding bGLYAT from a reference sequence from GenBank (NM177513) and the forward and reverse complement primer sequence of the plasmid which I generated. The start and end of the bGLYAT reading frame is indicated in green (Methionine and stop codon). The sequence highlighted in yellow (75 nucleic acids) was where the forward and reverse sequence generated overlapped. This confirmed that the DNA encoding bGLYAT was cloned in the correct orientation into pColdIII and that the reading frame was intact. The sequences generated (forward and reverse) were identical to the reference sequence NM177512 from NCBI (Vessey & Lau., 1996). 3' after the C-terminal end of the bGLYAT ORF, the nucleotide sequence was identical to that of pColdIII. Although the first 41 nucleic acids of the 5' sequence of the recombinant bGLYAT ORF was not generated, plasmid A was still used for further experiments. Plasmid A will be referred to as pColdIII-bGLYAT for the rest of the study.

### 3.3.3 Co-expression of recombinant bGLYAT with chaperone TF alone as well as TF together with GroEL-GroES

Nishihara *et al*. (2000) studied the prevention of aggregation of over-expressed recombinant proteins in *E. coli* when co-expressed with trigger factor and other chaperones. Some recombinant proteins were expressed as soluble protein when co-expressed with trigger factor only. Other proteins were only soluble when co-expressed with trigger factor and GroEL-GroES. Their results strongly suggested that TF together with GroEL-GroES worked cooperatively in preventing aggregation of some proteins (Nishihara *et al.,* 2000). I therefore decided to express TF by means of pTf16 as a separate protein together with bGLYAT to help with the correct folding as opposed to a fusion protein as described in Chapter 2 that did not have any enzyme activity. Seeing that TF and GroEL-GroES seems co-operate to prevent aggregation of some proteins, plasmid pG-Tf2 was also used to co-express TF and GroEL-GroES with bGLYAT.

Competent Origami™ cells were prepared (Section 2.2.10.1) and transformed (Section 2.2.10.2) with plasmids as follows: First the cells were transformed with a plasmid containing DNA encoding the different chaperones (TF alone from pTf16 or TF together with GroEL-GroES from pG-Tf2) and these transformed cells were made competent again and then transformed with either pColdIII or pColdIII-bGLYAT. The

plasmid pTf16 expresses the 50 kDa trigger factor under the control of the L-arabinose inducible (*araB*) promoter (Nishihara *et al.,* 2000). The plasmid pG-Tf2 co-expresses trigger factor (50 kDa), GroEL (57 kDa) and GroES (10 kDa) from the tetracycline inducible promoter (*Pzt-1*) (Nishihara *et al.,* 2000). It was expected that recombinant bGLYAT (36 kDa) would be expressed from pColdIII-bGLYAT (Nandi *et al*., 1979). The expression of proteins was induced as described in Section 3.2.2. The proteins were extracted from the bacterial cells as described in Section 2.2.13 and analysed by SDS-PAGE (Section 2.2.3.2). A summary of the Origami™ cells with the plasmids and the proteins they encode is given in Table 4.

**Table 4: Origami™ cells transformed with certain plasmids and the proteins of interest that they encode**

| Plasmid | Expressed proteins | Molecular weight of expressed proteins |
|---|---|---|
| pTf16 + pColdIII | TF | 50 kDa |
| pTf16 + pColdIII-bGLYAT | TF + bGLYAT | 50 kDa + 36 kDa |
| pG-Tf2 + pColdIII | TF + GroEL + GroES | 50 kDa + 57 kDa + 10 kDa |
| pG-Tf2 + pColdIII-bGLYAT | TF + GroEL + GroES + bGLYAT | 50 kDa + 57 kDa + 10kDa + 36 kDa |

### 3.3.3.1 Co-expression of recombinant bovine GLYAT and trigger factor

In Chapter 2 it was shown that trigger factor did prevent aggregation of recombinant bGLYAT. However, recombinant bGLYAT did not have any enzyme activity when expressed with TF as a soluble fusion protein. After the TF fusion was removed by thrombin cleavage, recombinant bGLYAT still did not have any enzyme activity. In this Section, I investigated whether TF can prevent aggregation when co-expressed with bGLYAT as a separate protein and not a fusion protein and whether bGLYAT would then have enzyme activity. In order to investigate this, bGLYAT was cloned into pColdIII and co-expressed with TF from pTf16. The proteins expressed in the bacterial cells without any induction was compared to those expressed when expression of bGLYAT was induced by IPTG induction (0.1 mM) and TF was induced by L-arabinose induction (4 mg/ml). The total bacterial cell lysate containing the proteins expressed as well as the soluble and insoluble bacterial protein cell lysate fraction were prepared and subjected to SDS-PAGE to determine if the proteins of interest were expressed as soluble proteins or not as described in Section 3.2. The results are shown in Figure 35.

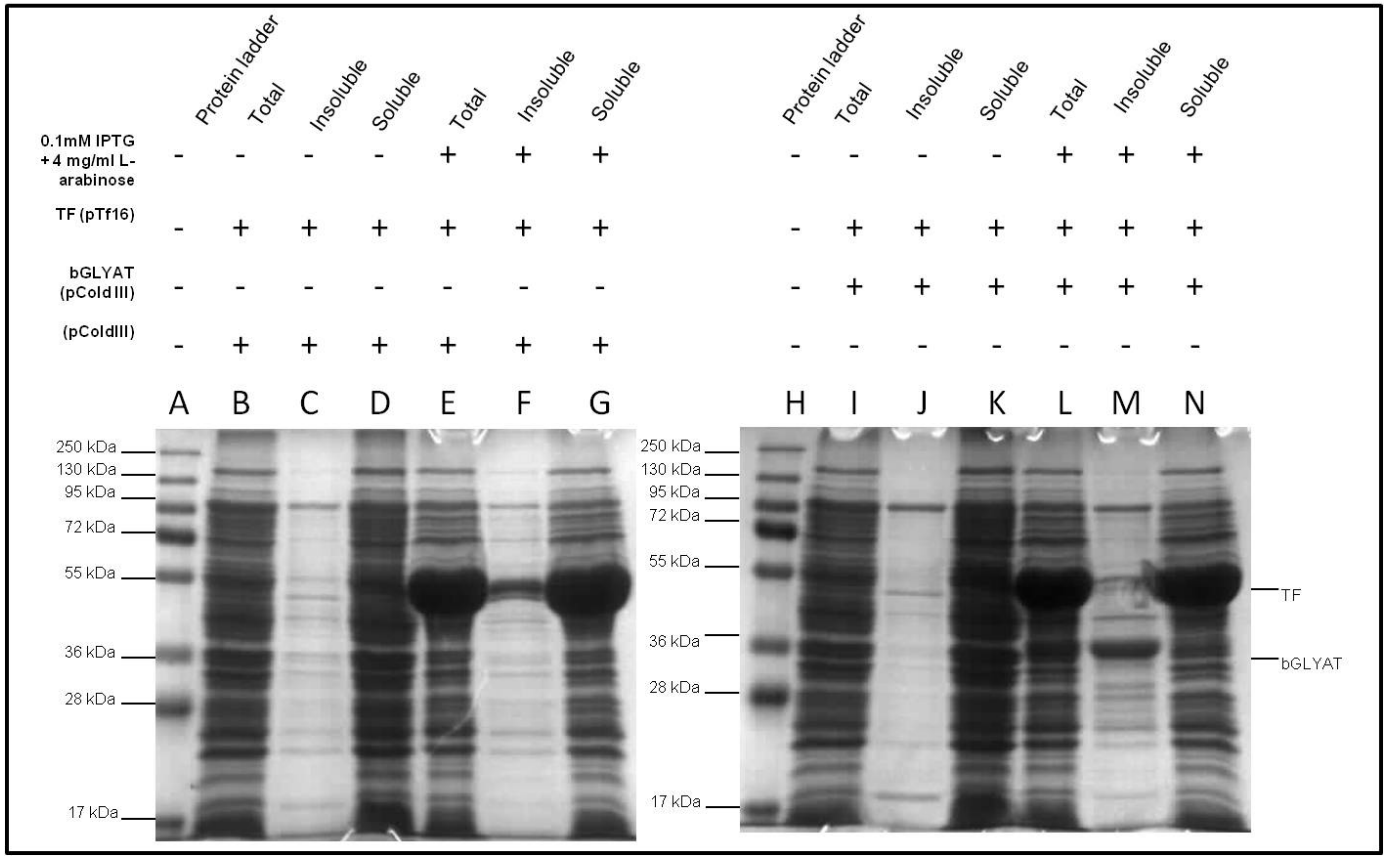



**FIGURE 35: SDS-PAGE of the co-expression of trigger factor from pTf16 and recombinant bovine GLYAT from pColdIII-bGLYA.** Lanes: A) 5 µl PageRuler™ Prestained Protein Ladder Plus; B) pColdIII and pTf16, total protein fraction; C) pColdIII and pTf16, insoluble protein fraction; D) pColdIII and pTf16, soluble fraction; E) pColdIII and pTf16 induced total protein fraction; F) pColdIII and pTf16 induced insoluble protein fraction; G) pColdIII and pTf16 induced soluble protein fraction; H) 5 µl PageRuler™ Prestained Protein Ladder Plus; I) pColdIII-bGLYAT and pTf16, total protein fraction; J) pColdIII-bGLYAT and pTf16, insoluble protein fraction; K) pColdIII-bGLYAT and pTf16, soluble fraction; L) pColdIII-bGLYAT and pTf16 induced total protein fraction; M) pColdII-bGLYAT and pTf16 induced insoluble protein fraction; N) pColdIII-bGLYAT and pTf16 induced soluble protein fraction. In lanes E, F, G, L, M and N, protein expression was induced with 0.1 mM IPTG and 4 mg/ml arabinose.

The majority of proteins as visualised by SDS-PAGE were soluble (Figure 35, lanes D, G, K and N compared with lanes C, F, E and M respectively). TF was over-expressed as a 50 kDa soluble protein when induced with 4 mg/ml arabinose (Figure 35, lanes G and N). bGLYAT was expressed as a 36 kDa protein upon induction with 0.1 mM IPTG (lane L) which correlated to the molecular weight reported by Nandi (1979). A protein band of approximately 36 kDa was visible in the insoluble fraction when GLYAT was expressed from pColdIII (Figure 35, lane M) and indicated that the bGLYAT was again insoluble.

#### 3.3.3.2 Co-expression of recombinant bovine GLYAT with trigger factor together with GroEL-GroES

When bGLYAT was co-expressed with TF, it was mostly insoluble (Figure 35, lane M). In this Section, I investigated whether it would be possible to prevent aggregation of recombinant bGLYAT by co-expressing it with TF and GroEL-GroES and whether bGLYAT would then have enzyme activity. In order to investigate this, bGLYAT cloned into ColdIII was co-expressed with TF and GroEL-GroES expressed from pG-TF2. The proteins that were expressed in the bacterial cells without any induction was compared to those expressed when expression of bGLYAT was induced by IPTG induction (0.1 mM) and TF with GroEL-GroES was induced by tetracycline induction (10 ng/ml). The total protein fraction as well as the soluble and insoluble proteins were prepared and subjected to SDS-PAGE to see of the proteins of interest were expressed as soluble proteins or not as described in Section 2.2.3.3. The results of co-expressing bGLYAT with TF as well as GroEL-GroES were essentially similar to the experiments where bGLYAT was co-expressed with TF alone (Figure 35). The results can be seen in Figure 36.

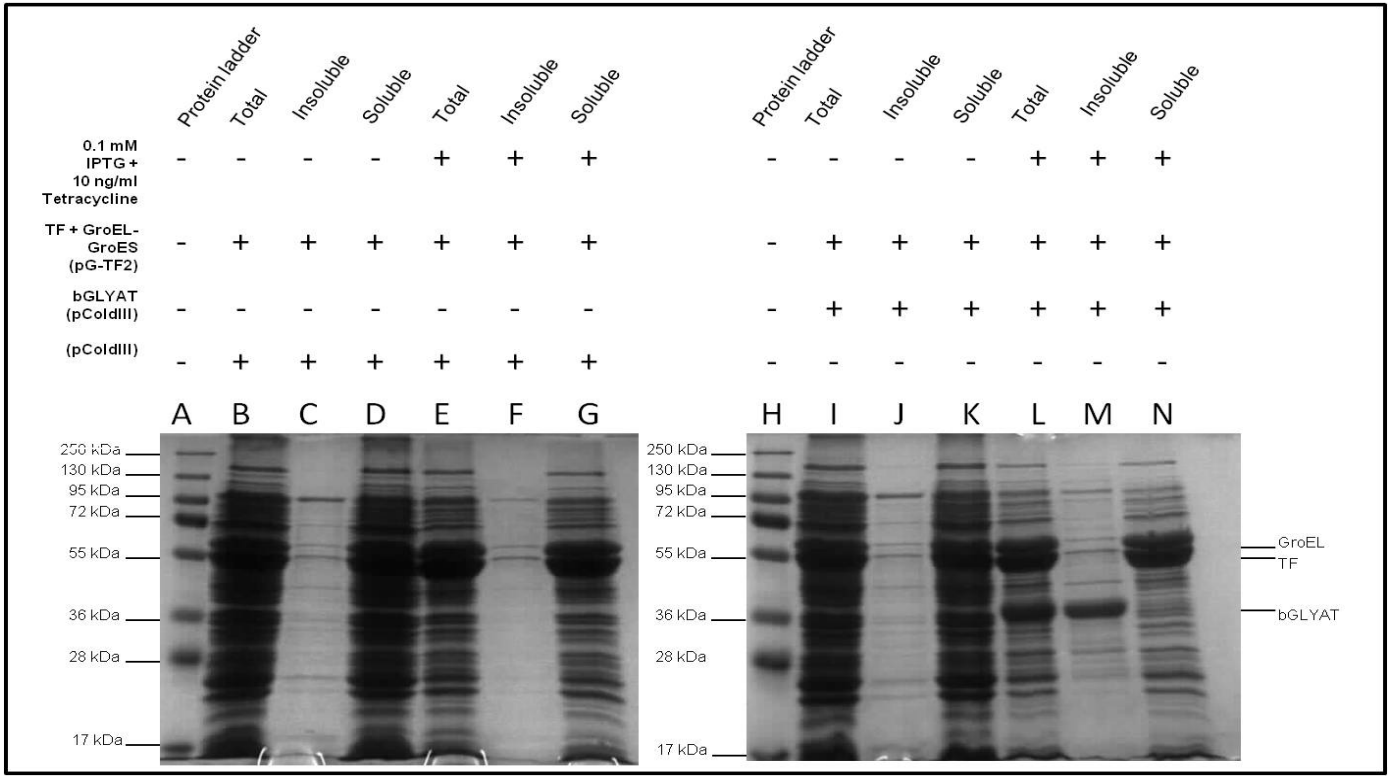



**FIGURE 36: SDS-PAGE of the co-expression of trigger factor and GroEL-GroES from pG-Tf2 and the bovine GLYAT from pColdIII-bGLYAT.** Lanes: A) 5 µl PageRuler™ Prestained Protein Ladder Plus; B) pColdIII and pG-Tf2, total protein fraction; C) pColdIII and pG-Tf2, insoluble protein fraction; D) pColdIII and pG-Tf2, soluble fraction; E) pColdIII and pG-Tf2 induced total protein fraction; F) pColdIII and pG-Tf2 induced insoluble protein fraction; G) pColdIII and pG-Tf2 induced soluble protein fraction; H) 5 µl PageRuler™ Prestained Protein Ladder Plus; I) pColdIII-bGLYAT and pG-Tf2, total protein fraction; J) pColdIII-bGLYAT and pG-Tf2, insoluble protein fraction; K) pColdIII-bGLYAT and pG-Tf2, soluble fraction; L) pColdIII-bGLYAT and pG-Tf2 induced total protein fraction; M) pColdIII-bGLYAT and pG-Tf2 induced insoluble protein fraction; N) pColdIII-bGLYAT and pG-Tf2 induced soluble protein fraction. In lanes E, F, G, L, M and N protein expression was induced with 0.1 mM IPTG and 10 ng/ml tetracycline.

The majority of all the proteins were soluble. In Figure 36, 50 kDa TF and 57 kDa GroEL were expressed as soluble proteins (lanes G and N) when induced with 10 ng/ml tetracycline. GroES, a 10 kDa protein was run off the SDS-PAGE gel because electrophoreses was done too long. bGLYAT (36 kDa) was again expressed as an insoluble protein (Figure 36, lane M) when induced with 0.1 mM IPTG. The 36 kDa molecular weight of the putative recombinant bGLYAT correlated to the 36 kDa reported by Nandi (1979) for bGLYAT purified from bovine liver.

### 3.3.4 GLYAT enzyme activity assay on different preparations of bacterially expressed bGLYAT

Although no significant amount of soluble recombinant bGLYAT could be detected by SDS-PAGE as visualized in Figures 35 and 36 lanes N, the GLYAT enzyme activity in the cell lysates containing recombinant bGLYAT was tested nevertheless. To analyse the enzyme activity of recombinant bGLYAT, the GLYAT enzyme activity assay described in Section 2.2.19 was performed. Three different enzyme activity experiments were performed as follows: 1) a dilution series of the soluble fraction of an Origami™ cell lysate which expressed TF, GroEL-GroES, bGLYAT and other proteins encoded by pColdIII, pColdIII-bGLYAT, pTF16 and pG-TF2 respectively; 2) GLYAT enzyme activity of the soluble fraction of an Origami™ cell lysate which expressed TF, bGLYAT and other proteins encoded by pColdIII, pColdIII-bGLYAT, pTF16 and pG-Tf2 respectively; 3) comparison of GLYAT enzyme activity in the total, insoluble and soluble fractions of an Origami™ cell lysate which expressed bGLYAT with TF or TF together with GroEL-GroES respectively.

To be able to compare the GLYAT enzyme activity in the soluble fractions of different preparations of bacterial lysates, a bicinchoninic acid (BCA)-based protein assay (Smith *et al.,* 1985) was performed. The protein concentration of the different lysates as well as the bovine liver crude cytoplasmic extract was determined and listed in Table 5.

**Table 5: Protein concentration of the soluble fraction of different bacterial lysates determined by a bicinchoninic acid (BCA)-based protein assay**

| Lysate | Protein concentration |
|---|---|
| 1. Bovine liver crude cytoplasmic extract containing bGLYAT | 2.6 µg/µl |
| 2. Soluble fraction of Origami™ cell lysate which expressed TF and other proteins encoded by pColdIII and pTF16 | 1.5 µg/µl |
| 3. Soluble fraction of Origami™ cell lysate which expressed TF, bGLYAT and other proteins encoded by pColdIII-bGLYAT and pTF16 | 2.2 µg/µl |
| 4. Soluble fraction of Origami™ cell lysate which expressed TF, GroEL-GroES and other proteins encoded by pColdIII and pG-Tf2 | 2.5 µg/µl |
| 5. Soluble fraction of Origami™ cell lysate which expressed TF, GroEL-GroES, bGLYAT and other proteins encoded by pColdIII-bGLYAT and pG-Tf2 | 2.7 µg/µl |

The bovine liver crude cytoplasmic extract containing bGLYAT was used as positive control in all the enzyme assays. It was prepared by a fellow student, Mr JHJ Fourie, and shown to have GLYAT enzyme activity. Six enzyme assays were performed at a time. For each sample analysed, a negative control was set up by omitting glycine in order to compensate for non-specific background caused by the increasing colour reaction that is not due to CoASH release. GLYAT enzyme activity was measured by monitoring the glycine-dependent CoASH release from benzoyl-CoA at 412 nm.

Although a BCA protein assay was done before the enzyme assays, quantitative analysis of the bGLYAT content was not possible at this point of our research because bGLYAT could not be purified from the other proteins by means of affinity chromatography or any other means. Thus, the exact protein concentration of bGLYAT in these preparations could not be determined. Two approaches to relatively quantify the GLYAT activity was using 13 µg total protein of each sample (except liver GLYAT of 2.6 µg). The second was to use a dilution series of the samples where 8 µl and less of the sample were used. This was a very rough quantification. For this assay I used 1 µl, 2 µl, 4 µl and 8 µl of soluble bacterial lysate, since 8 µl was the maximum volume available.

To at least relatively quantify the activity of recombinant bGLYAT between the different total fractions and the soluble fraction prepared from it, a dilution series of

the soluble bacterial lysate fraction of bGLYAT expressed with TF and TF with GroEL-GroES was performed. The results are shown in Figure 37.

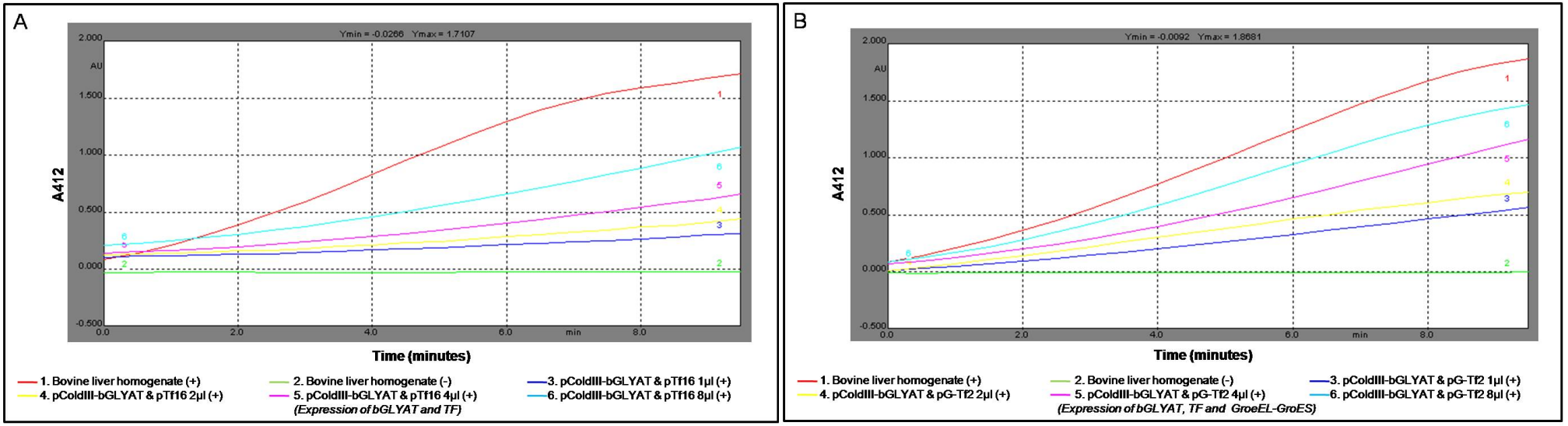



**Figure 37: Verification of the specificity of GLYAT activity by a dilution series of the soluble fraction of cell lysates containing of recombinant bGLYAT co-expressed with (A) trigger factor and (B) trigger factor and GroEL-GroES.** The graph shows the changes in absorbance at 412 nm over time, using different volumes of each bacterial lysate. (+) Reactions containing glycine, (-) Reactions not containing glycine that act as a negative control Graphs: A) Lysate of Origami™ cell which expressed TF, bGLYAT and other proteins encoded by pColdIII-bGLYAT and pTF16; B) Lysate of Origami™ cells which expressed TF, GroEL-GroES, bGLYAT and other proteins encoded by pColdIII-bGLYAT and pG-Tf2. Lines: A1/B1) 2.6 µg bovine liver crude cytoplasmic extract as positive control; A2/B2) Repeat of lane A1/B1, but omitting glycine from the reaction, negative control; A3/B3) 1 µl bacterial lysate; A4/B4) 2 µl bacterial lysate; A5/B5) 4 µl bacterial lysate; A6/B6) 8 µl bacterial lysate.

In Figure 37 GLYAT enzyme activity was detected in the soluble fraction of the bacterial cell lysate where in bGLYAT was co-expressed with TF alone (Figure 35, lane N) or TF together with GroEL-GroES (Figure 36, Lane N). Although bGLYAT was not visible in the soluble fraction with SDS-PAGE, this was the first result of GLYAT enzyme activity of a recombinant bovine GLYAT. In Figure 37 it can be seen that the activity of the lysates decreased the higher the dilution factor (lower protein concentration). Under these reaction conditions (Section 2.2.19), the initial reaction rate is proportional to enzyme concentration alone. The acyl acceptor, glycine, and the substrate, benzoyl-CoA, concentration is not the rate limiting factors, but rather the protein content. Therefore, the increase in enzyme activity with increasing volumes of lysate varied the enzyme concentrations and increased the release of CoASH which correlates with GLYAT activity. The reaction was performed at substrate concentrations significantly above the saturation level so that the initial reaction rate is zero order for the substrate.

The second series of enzyme assays was performed using a volume of bacterial cell lysate of the different samples that contained 13 µg protein except for the crude cytoplasmic bovine liver extract. A volume of bovine liver crude cytoplasmic extract containing 2.6 µg of protein was used. In pilot GLYAT enzyme assays (results not shown) the GLYAT activity of the bovine liver crude cytoplasmic extract was judged to be too high to detect low levels of enzyme activity (as seen in the dilution series) that was expected of the bacterial cell lysates expressing bGLYAT.

Three different protein preparations were tested in this experiment. They were: 1) the bovine liver crude cytoplasmic extract containing bGLYAT; 2) the soluble fraction of an Origami™ cell lysate in which TF and proteins encoded by pColdIII and pTF16 were expressed; and 3) the soluble fraction of an Origami™ cell lysate in which TF and bGLYAT and proteins encoded by pColdIII-bGLYAT and pTF16 were expressed. The results are shown in Figure 38 A.

A similar experiment was done to determine the effect of GroEL-GroES in conjunction with TF on the folding of bGLYAT and its enzyme activity. A series enzyme activity assays was performed in: 1) the bovine liver crude cytoplasmic extract containing bGLYAT; 2) the soluble fraction of an Origami™ cell lysate in which TF, GroES-GroEL and other proteins encoded by pColdIII and pG-Tf2 were

expressed; and 3) the soluble fraction of an Origami™ cell lysate in which TF, GroEL-GroES, bGLYAT and proteins encoded by pColdIII-bGLYAT and pG-Tf2 were expressed. The results are shown in Figure 38 B.

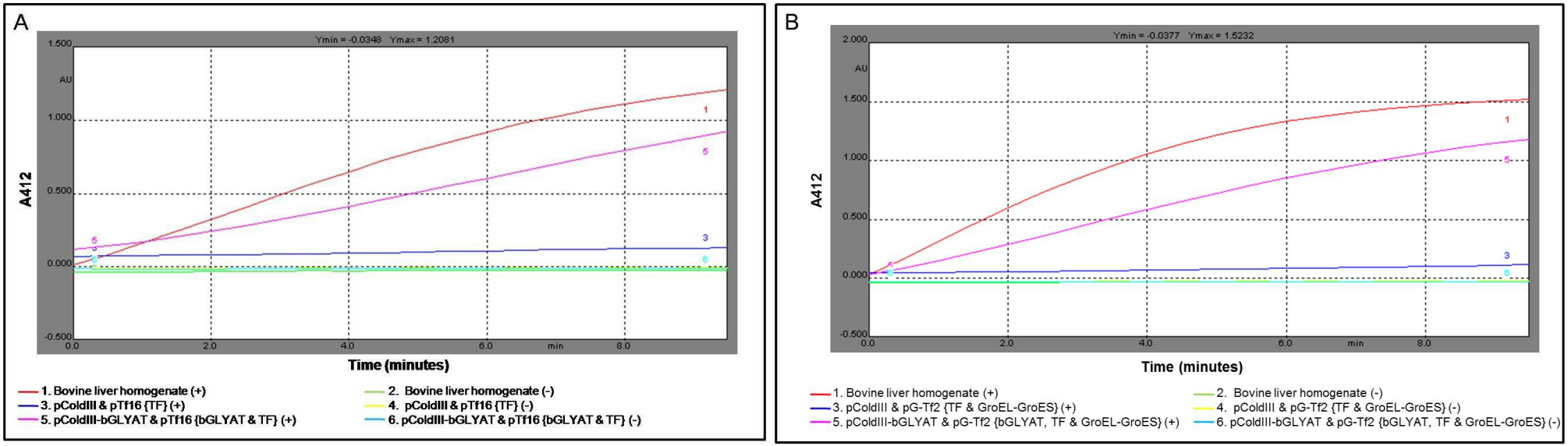



**Figure 38: GLYAT enzyme assay of the soluble fraction of an Origami™ cell lysate which co-expressed recombinant bGLYAT with (A) trigger factor; (B) trigger factor and GroEL-GroES with recombinant bGLYAT.** The graphs show the changes in absorbance at 412 nm over time. (+) Reactions containing glycine, (-) Reactions not containing glycine and acts as a negative control. Graph lines: A1, B1) 2.6 µg bovine liver crude cytoplasmic extract as positive control; A2, B2) Repeat of the reaction depicted by lane A1/B1, but omitting glycine from the reactions, negative control; A3) Soluble fraction of Origami™ cell lysate which expressed TF and proteins encoded by pColdIII; A4) Repeat of the reaction depicted by lane A3, but omitting glycine from the reaction, negative control; A5) Soluble fraction of Origami™ cell lysate which expressed TF and bGLYAT and other proteins encoded by pColdIII-bGLYAT and pTF16; A6) Repeat of reaction depicted by lane A5, but omitting glycine from the reaction, negative control; B3) Soluble fraction of Origami™ cell lysate which expressed TF and GroEL-GroES and other proteins encoded by pColdIII and pG-Tf2; B4) Repeat of reaction depicted by lane B3, but omitting glycine from the reaction, negative control; B5) Soluble fraction of Origami™ cell lysate which expressed TF, GroEL-GroES, bGLYAT and other proteins encoded by pColdIII-bGLYAT and pG-Tf2; B6) Repeat of reaction depicted in lane B5, but omitting glycine from the reaction, negative control,

GLYAT enzyme activity was detected in the bovine liver crude cytoplasmic extract (Figure 38, line A1/B1). This confirmed the results obtained by Mr JHJ Fourie. When glycine was omitted, there was no GLYAT activity as expected (Figure 38, line A2/B2). Similarly, there was no GLYAT activity in any of the other reactions that did not contain glycine (Figure 38, line A4/B4 and A6/B6). The soluble fraction of the Origami™ cell lysate in which TF and other proteins encoded by pColdIII were expressed did not show any GLYAT activity (Figure 38, line A3). It can therefore be also be concluded that proteins expressed by Origami™ cells and other proteins encoded in pColdIII or pTf16 did not have any GLYAT activity. GLYAT enzyme activity was detected in the soluble fraction (Figure 38, line A5) of the bacterial cell lysate wherein bGLYAT was co-expressed with TF alone. Although bGLYAT was not visible in the soluble fraction in SDS-PAGE (Figure 35, lane N), GLYAT activity was detected.

The same results were obtained from the experiment depicted in Figure 38 B. GLYAT activity was only detected in the bovine liver crude cytoplasmic extract (Figure 38, line B1) and in the soluble fraction of Origami™ lysate when bGLYAT was co-expressed with TF and GroEL-GroES (Figure 38, line B5).

The results depicted in Figure 38 indicated that co-expression of bGLYAT and either TF alone or TF together with GroEL-GroES resulted in bGLYAT which is enzymatically active. Therefore, chaperones TF and GroEL-GroES co-expressed with a recombinant bGLYAT facilitated correct protein folding that resulted in an enzymatically active recombinant protein. The enzyme activity of bGLYAT co-expressed with TF alone or of bGLYAT co-expressed with TF and GroEL-GroES cannot be quantitatively compared to each other.

Next, I also compared the GLYAT enzyme activity of the total, soluble and insoluble fractions of Origami™ cell lysates wherein bGLYAT was co-expressed with TF alone or with TF together with GroEL-GroES. The results are shown in Figure 39.

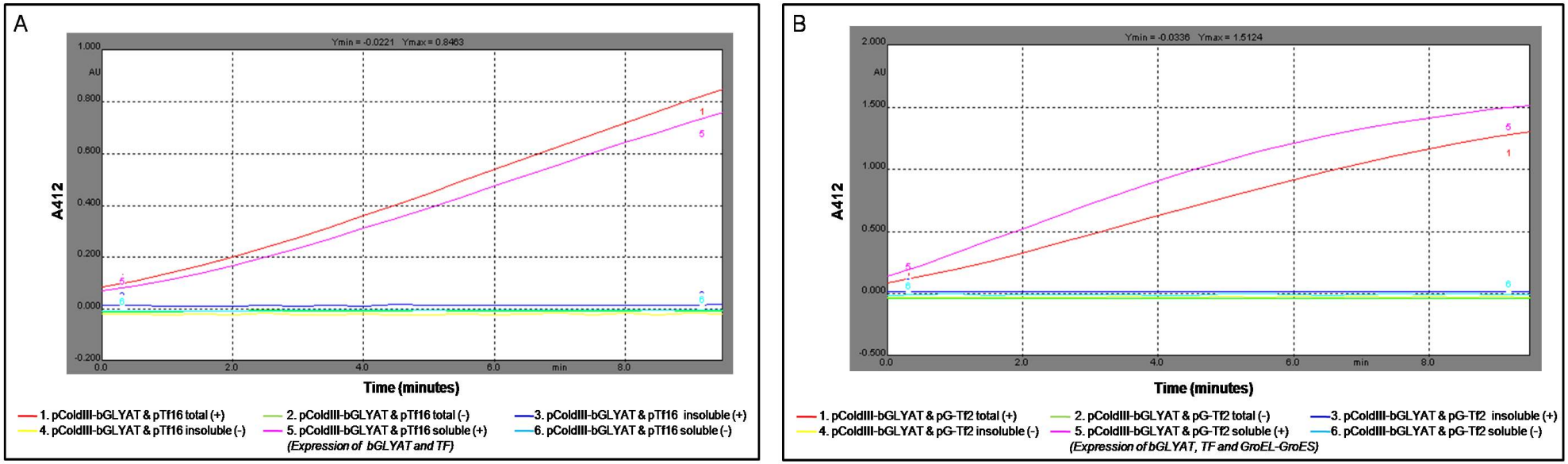



**Figure 39. GLYAT enzyme assay of the total, insoluble and soluble fractions of an Origami™ cell lysate which co-expressed recombinant bGLYAT with (A) trigger factor; or (B) trigger factor and GroEL-GroES.** The graphs show the changes in absorbance at 412 nm over time, using 6 µl from each bacterial lysate (+) Reactions containing glycine, (-) Reactions not containing glycine that act as negative controls. Graph lines: A) Lysate of Origami™ cell which expressed TF, bGLYAT and other proteins encoded by pColdIII-bGLYAT and pTF16; B) Lysate of Origami™ cells which expresses TF, GroEL-GroES, bGLYAT and other proteins encoded by pColdIII-bGLYAT and pG-Tf2. Lines 1) Total fraction; 2) Total fraction without glycine, negative control; 3) Insoluble fraction; 4) Insoluble fraction without glycine, negative control; 5) Soluble fraction; 6) Soluble fraction without glycine, negative control;

Both total fractions containing recombinant bGLYAT co-expressed with TF (Figure 39, line A1) and TF together with GroEL-GroES (Figure 39, line B1) as well as both the soluble fractions prepared from the recombinant total fraction did have enzyme activity (Figure 39, lanes A5 and B5). No enzyme activity was detected in any of the insoluble fractions (Figure 39, line A3 & B3).

The next step was to investigate whether more soluble recombinant bGLYAT could be obtained from the insoluble fraction by using different pH buffers.

### 3.3.5 The effect of pH on the solubility of the aggregated form of bGLYAT when co-expressed in bacteria with trigger factor alone, or trigger factor and GroEL-GroES

When bGLYAT was co-expressed with TF alone or TF with GroEL-GroES, the majority of the proteins were insoluble. The goal of this experiment was to investigate whether different pH buffers would be able to solubilise the recombinant bGLYAT aggregates (found in the insoluble fraction of the bacterial cell lysate), thereby increasing enzyme activity. The insoluble fraction in PBS (pH 7.4) was pelleted, and resuspended in four different pH buffers respectively and divided into soluble and insoluble fractions again (Section 3.2.3). It was expected that if the pH did increase solubility of the bGLYAT aggregate, the increased amount of bGLYAT would be seen in the soluble fraction. The results can be seen in Figure 40.

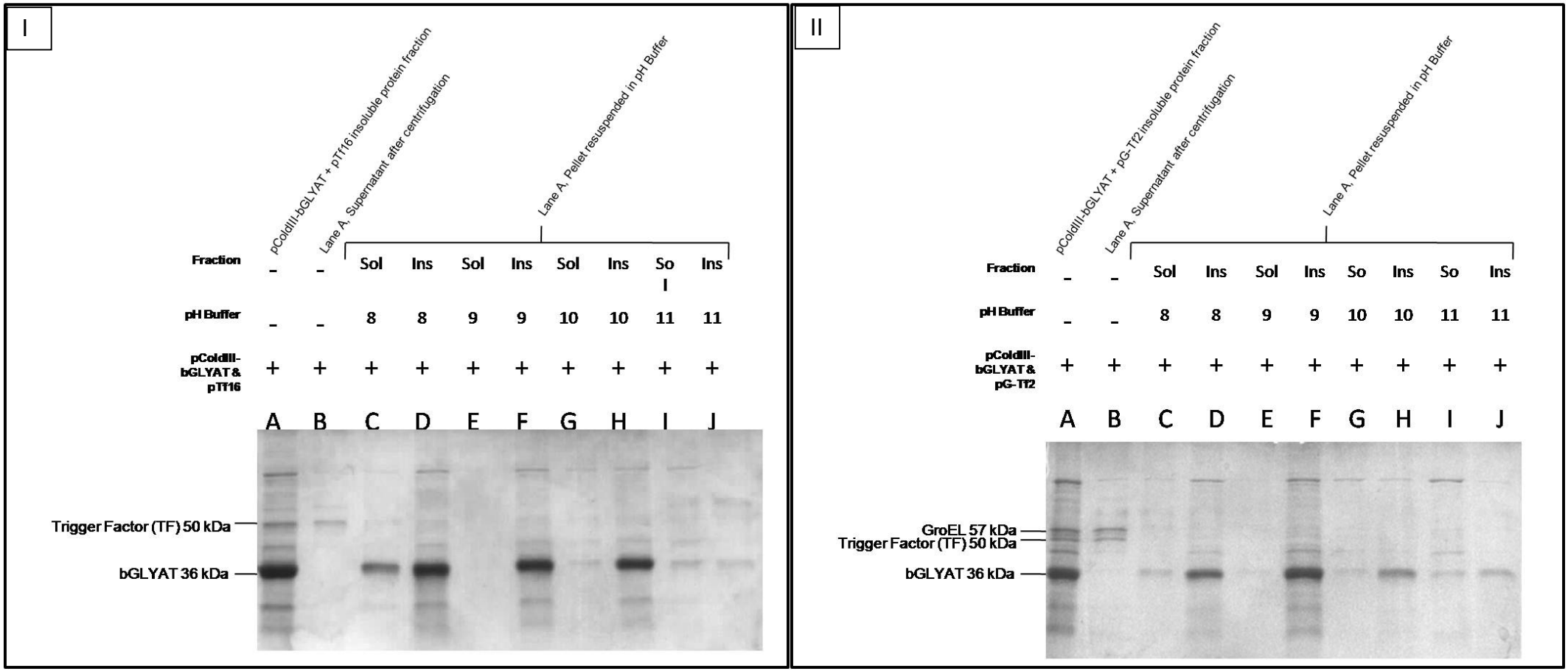



**FIGURE 40: SDS-PAGE analysis of treating insoluble bGLYAT co-expressed with (I) trigger factor or (II) trigger factor and GroEL-GroES with two different buffers at four different pH ranges.** Insoluble fraction of proteins expressed encoded by pColdII-bGLYAT and (I) pTf16 or (II) pG-Tf2; Lanes: A supernatant; Lane C, E, G & I) Lane A pellet resuspended in TABS pH 8 and 9, and CAPS pH 10, 11 respectively, soluble protein fractions; Lanes D, F, H & J) Lane A pellet resuspended in TABS pH 8 and 9, and CAPS pH 10 and 11 respectively, all insoluble protein fraction resuspended in PBS.

Figure 40 (I) showed the effect of pH on the insoluble fraction of Origami™ cell lysate which expressed TF, recombinant bGLYAT and other proteins encoded by pColdIII-bGLYAT and pTF16. Figure 40 (II) showed the effect of pH on the insoluble fraction of Origami™ cell lysate which expressed TF, GroEL-GroES, recombinant bGLYAT and other proteins encoded by pColdIII-bGLYAT and pG-Tf2.

When the soluble fraction at pH8 (Figure 40 I/II, lane C) was compared to pH 9, 10 and 11 (Figure 40 I/II, lane E, G, I respectively), it seems that bGLYAT was more soluble at pH 8 than any of the other pH values. At pH 11, no significant proteins can be seen in the soluble or the insoluble fractions. It can be that the proteins were unstable at this high pH and might have denatured. The results could not be quantified correctly as no means of purification for bGLYAT was available. The different pH ranges had little effect on the solubility and the results were not conclusive. The results were not used in further experiments and were fractions discarded.

Seeing that the pH had little effect in this experiment on the solubility of bGLYAT aggregates, other means of getting more recombinant enzymatically active bGLYAT should be attempted in future studies. Seeing that bGLYAT was not expressed with a phusion tag from pColdIII, in future studies, bGLYAT should be expressed with a small fusion tag like HIS in order to be able to purify the enzyme and correctly quantify it.

## 3.4 Summary

In Chapter 2 I explained how the DNA amplicon encoding the open reading frame of bGLYAT was obtained. For the work in this chapter I used the remainder of the DNA amplicon described in Section 2.3.2. The PCR amplicon was obtained as follows: The messenger RNA (mRNA) included in the total RNA extracted from bovine liver served as the template for reverse transcriptase (RT) in order to develop a bovine cDNA library. Gene specific bGLYAT primers containing appropriate restrictions enzyme sites were used to generate the bGLYAT amplified from the pool of cDNA.

The remainder of the PCR amplicon from Section 2.3.2 and the pColdIII vector was each double digested with NdeI and SalI restriction enzymes. The insert was cloned into the pColdIII vector by means of directional cloning. Sequencing results confirmed correct orientation and open reading frame for expression of bGLYAT in pColdIII. Protein expression of bGLYAT without a fusion protein was induced with IPTG and co-expressed with either TF alone (L-arabinose induction) or TF together with GroEL-GroES (tetracycline induction). According to SDS-PAGE analysis, no soluble bGLYAT was obtained when co-expressed with either TF alone or TF together with GroEL-GroES. According to the protein reference ladder, a band at approximately 36 kDa was visible in the insoluble fraction and presumed to be bGLYAT.

Even though bGLYAT seemed to be insoluble, the GLYAT enzyme activity assay was performed on the total, insoluble and soluble bacterial fractions of bGLYAT over-expressed with either TF alone, or TF together with GroEL-GroES. GLYAT enzyme activity was detected in the total and soluble bacterial cell lysate fractions, but not in the insoluble fraction where bGLYAT was over-expressed. This was the first finding of enzymatically active recombinant bGLYAT even though no visible, soluble bGLYAT on SDS-PAGE was detected. It is speculated that seeing that no enzyme activity was detected in the insoluble fraction, but indeed in the soluble fraction, that correct folding and solubility is very important for enzyme activity. The protein concentration could not be quantified as bGLYAT did not have a fusion protein and could not be purified. I was therefore not able to determine if TF was capable of folding bGLYAT correctly and thereby produced more soluble enzymatically active bGLYAT as to TF together with GroEL-GroES. The effect of pH on the solubility of

aggregated bGLYAT was briefly investigated. No increase in solubility of recombinant bGLYAT was obtained and the experiment was abandoned.

# CHAPTER 4

## CONCLUDING DISCUSSION

We are exposed to increasing levels of toxins every day. This includes food components, preservatives, pharmaceuticals and environmental factors. Detoxification is the process of clearing toxins from the body or transforming them into other components that are not harmful to the body. We need a better understanding of how these toxic substances are detoxified in the human body. Detoxification has become a popular topic for the general public. It is well established to the point where detoxification profiling is becoming readily available as a commercial service. At the moment we know detoxification as a multiple reaction system that is divided into four phases. Phase 0, Phase I (functionalization), Phase II (conjugation) and Phase III (elimination). The focus of this MSc project was on glycine N-acyltransferase which is part of Phase II detoxification, by performing glycine conjugation. The xenobiotic benzoic acid is widely used as a food preservative. We are constantly exposed to this xenobiotic in our diet. GLYAT is responsible to detoxify benzoic acid into hippuric acid for excretion by means of glycine conjugation. Glycine supplementation is commonly used in the treatment of groups of IEMs known as organic acidemias. No GLYAT associated inborn errors of metabolism (IEM) is known.

The Centrum for Human Metabolomics of the subject group Biochemistry, School for Physical and Chemical Science at the North-West University has done some research projects on human and bovine GLYAT in areas such as molecular characterization and the enzyme function (Van der Westhuizen, 1999). The goal of this study was to 1) prepare mRNA encoding glycine N-acyltransferase from bovine liver, 2) generate a recombinant bGLYAT and 3) verify if the recombinant bGLYAT exhibits enzyme activity. Since no success in previous studies was achieved in our laboratory with recombinant human GLYAT, we decided to do some more work on bovine GLYAT. Bovine GLYAT is quite similar to human GLYAT with regards to molecular and enzyme characteristics. This information was used to get insight into the detoxification process that GLYAT is involved in. Although humans are closer to mice than bovine, bovine and humans have sufficient DNA sequence similarity to

enable us to map the human genome almost entirely to the genome of the bovine. Studies have also been done on GLYAT of other species such as monkeys (Webster *et al.,* 1976), rats and rabbits (James & Bend, 1978; Kolvraa & Gregersen, 1986).

Since I completed my practical work in the laboratory in 2008, the open reading frame of two GLYAT-like enzymes were identified following the completion of the human genome project. These enzymes are hGLYATL1 (Zhang *et al*., 2007) and GLYATL2 (Waluk *et al.,* 2010). In 2007, Zhang and his colleagues found a novel gene encoded a GLYAT gene member known as hGLYATL1. hGLYATL1 was expressed in mammalian HEK293T cells but the enzyme activity was not tested (Zhang *et al.,* 2007). Subcellular localization of hGLYATL1 revealed that hGLYATL1 was distributed primarily in the cytoplasm (Zhang *et al.,* 2007). Recently, Waluk and colleagues (2010) identified human GLYAT like 2 (hGLYATL2), another member of the four putative glycine conjugation family of enzymes. hGLYATL2 was expressed in bacterial cells and was enzymatically active. Waluk and co-workers (2010) aligned the sequences of the four putative glycine conjugation family of enzymes and found that GLYATL2 has 40-51% sequence identity to each GLYAT. Subcellular localization of GLYATL2 revealed that GLYATL2 was distributed primarily in the endoplasmic reticulum of HepG2 cells GLYATL2 (Waluk *et al.,* 2010).

Since the discovery of GLYAT in 1842 (Keller, 1842), the focus of GLYAT moved to glycine supplementation therapy for IEMs. Today the focus has been renewed to glycine conjugation (N-acyl glycines), which is part of the lipid signalling family of medium and long chain fatty acids. N–acyl glycines are now emerging as a potent lipid signaling molecule and activities include antinociceptive, anti-inflammatory and antiproliferative effects (Waluk *et al.,* 2010). hGLYATL2 was identified in the salivary gland and respiratory track, as well as lung and skin. This may indicate that N-acyl glycines have a role in immune response.

The focus of this MSc was to use a third generation expression system that co-expresses chaperones and analyse whether they can assist with protein folding of recombinant bGLYAT. With previous studies done by Mr Japie Fourie and Danie Grundling in our laboratoy, bGLYAT was expressed as an insoluble and inactive protein. In this study I investigated whether the use of third a generation expression system would yield soluble, enzymatically active bovine glycine N-acyltransferase.

In Chapter 2 a recombinant bovine GLYAT was expressed using the following approach. The bovine liver was obtained from the abattoir and RNA extracted from a piece thereof. The messenger RNA (mRNA) included in the total extracted RNA served as the template for reverse transcriptase (RT) in order to prepare a cDNA library. Specific bGLYAT primers containing appropriate restriction enzyme sites amplified GLYAT cDNA from the pool of cDNA. The GLYAT amplicon was about 900 bp which compared to the transcripts length of bGLYAT which was 984 bases found by Vessey and Lau (1996). The DNA was successfully cloned into pColdTF. Sequencing analysis confirmed that the reading frame and orientation of the ORF encoding bGLYAT was correct.

In 1979 Nandi determined the molecular weight of bovine GLYAT to be 33 kDa, whereas Kelley and Vessey (1992) and Van der Westhuizen (2000) reported it to be 33.5 kDa and 36 kDa respectively (Nandi *et al.,* 1979; Kelley & Vessey, 1992; Van der Westhuizen *et al.,* 2000). Trigger factor (TF) is a 50 kDa protein and bGLYAT was expected to be about 36 kDa. According to SDS-PAGE, bGLYAT was over-expressed as a fusion protein with TF at the correct molecular weight of 88 kDa and was mostly soluble. The histidine tagged bGLYAT-TF fusion protein was purified with nickel column chromatography. GLYAT enzyme activity assays revealed that the recombinant bGLYAT expressed as a fusion protein with TF was inactive. It is speculated that the fusion protein was obstructing the active site of bGLYAT. The fusion tag was then removed by means of thrombin cleavage. After removal of the tag, the bGLYAT was about 34 kDa that correlated to the expected results of 36 kDa reported by Nandi (1979). Unfortunately, the recombinant bGLYAT again did not have any enzyme activity.

In Chapter 2, I explained how the DNA amplicon encoding the open reading frame of bGLYAT was obtained. bGLYAT was about 909bp that correlated to the expected results (Vessey & Lau*,* 1996). In Chapter 3 this amplicon was used for directional cloning into pColdIII. After appropriate restriction enzyme digestion of the bGLYAT PCR amplicon and the pColdIII vector, the insert was subcloned into the vector by means of directional cloning. Sequencing results confirmed that the orientation and open reading frame of the bGLYAT insert was correct for expression of bGLYAT. bGLYAT was co-expressed with either TF alone (encoded by pTf16) or TF together

with GroEL-GroES (encoded by pG-Tf2). GLYAT enzyme activity was detected in the total and soluble bacterial cell lysate fractions, but not in the insoluble fraction. This was the first time an enzymatically active recombinant bGLYAT was generated. Overexpression of TF alone and TF together with GroEL-GroES appeared to be sufficient to prevent complete aggregation of bGLYAT. Attempts to solubilise the insoluble recombinant bGLYAT failed.

In the future, the next step should be to express bGLYAT with a suitable fusion sequence for purification. This will enable us to quantify the concentration of bGLYAT correctly. This will also enable us to determine whether TF together with GroEL-GroES has a better ability to fold bGLYAT correctly as to TF alone. Western blotting should have been done to confirm the amount of bGLYAT in the soluble fraction, and is a shortcoming in the study due to unavailability of a bGLYAT antibody.

The recombinant bGLYAT can also be used for further investigation of genetic variation, to probe the specificity of amino acid residues by site-directed mutagenesis. Thereby we can start to analyse the impact of individual SNPs on enzyme activity and amino acid specificity.

Follow up investigations of this results by Mr CPS Badenhorst was successful and has lead to a manuscript (Appendix C) entitled “Enzymatic characterisation and elucidation of the catalytic mechanism of a recombinant bovine glycine N-acyltransferse” which has been submitted for publication to the peer-reviewed scientific journal Drug Metabolism and Disposition. I am one of the three co-authors.

# REFERENCES

APPLIED BIOSYSTEMS (England) RNA isolation: The basics. [Web:]. www.ambion.com [Date of access: Jan 2008]

BANDMANN, O., VAUGHAN, J. & HOLMANS, P. (1997) Association of slow acetylator genotype for N-acetyltransferase 2 with familial Parkinson's disease. Lancet. 350:1136-1139

BARIC, I., FUMIC. F. & HOFFMANN, G.F. (2001) Inborn errors of metabolism at the turn of the millennium. Croat Med J, 42: 379-83

BAUMGARTNER, M. (2003) 3 Methylcrotonyl-CoA carboxylase deficiency. Orphanet encyclopedia. July 2003.

BEAUDET. A.L., SCRIVER. C.R., SLY. W.S. & VALLE, D. (1995) Genetics, biochemisty and molecular basis of variant human phenotypes. In Scriver CR, Beaudet AL, Sly WS and Valle eds. The molecular basis of Inherited diseases, 7th edition, New York: McGraw-Hill, 53-228

BERGER, S.L., PINA, B., SILVERMAN, N., MARCUS, G.A., AGAPITE, J., REGIER, J.L., TRIESENBERG, S.J. & GUARENTE, L. (1992) Genetic isolation of ADA2: A potential transcriptional adaptor required for function of certain acidic activation domains Cell. 70;251-265

BRENDA. (2007) [Web;] www.brenda.uni-koeln.de (Date of access: 22 July 2007)

BUKAU, B., DEUERLING, E., PFUND, C. & CRAIG, E.A. (2007) Getting newly synthesizes proteins into shape. Cell. 101:119-122

CHIN, K.V., PASTAN, I. & GOTTESMAN, M.M. (1993) Function and regulation of the multidrug resistance gene. Adv Cancer Res. 60:157-180

CHUNG, C.T., NIEMELLA, S.L. & MILLER, R.H. (1989) One-step preparation of competent Eschericia coli: transformation and storage buffer of bacterial cella in the same solution. PNAS 89:2172-2175

DAVIS, J. & WRIGHT, G,D. (1997) Bacterial resistance to aminoglycoside antibiotic. Trends Microbiol. 5; 234-240

DYDA, F., KLEIN, D.C. & HICKMAN, A.B. (2000) GCN5-Related N-acyltransferases: A structural overview. Annu Rev Biomol Struct 29:81-103

ENSEMBLE GENOME BROWSER. (2011) [Web;] www.ensemble.org (Date of access: 8 August 2011)

FINNZYMES. (Ipswich) Phusion® High-Fidelity DNA Polymerase (2011) [Web;] www.finnzyme.fi [Date of access: Jan 2008]

GARRETT, R.H. & GRISHAM, C.M. (2005) Biochemistry 3rd ed. Thomson Books/Cole

GARROD, A.E., FROWDE, H., HODDERS & STOUGHTON. (1923) Inborn Errors of Metabolism. Second Edition

GATLEY, S.J., STANLEY, H & SHERRETT, A. (1977) The synthesis of hippurate from benzoate and glycine by rat liver mitochondria. Submitochondrial localization and kinetics. Biochem J. 166:39-47

GREGERSEN, N., KOLVRAA, S., & MORTENSEN, P.B. (1986) Acyl-CoA: glycine Nacyltransferas: organelle localization and affinity towards straight- and branched chain acyl-CoA esters in rat liver. Biochem Biomed Meta Biol 36:98-105

HANAHAN, D. (1985) IN: DNA Cloning, Volume 1, D. Glover, ed., IRL Press, LTS., London, 109

HIGGINS, N.P. & CORZZARELLI, R. (1989) Recombinant DNA Metholdology Wu, R., Grossmann, L. and Moldave, K., eds. Academic Press, Inc., San Diego, California.

INVITEK (Germany) MSB spin PCRapase, for purification of DNA fragments after PCR reactions, restriction digestion, ligation and cDNA synthesis. [Web;] www.invitek.eu (Date of access: Jan 2008)

ISHIKAWA, Y. (1992) The ATP-dependent gluthatione S-conjugate export pump Trends Biochem Sci. 463-468.

JAMES, M.O. & BEND, J.R. (1978) A radiochemical assay for glycine Nacyltransferase activity. Some properties of the enzyme inrat and rabbit. Biochem J. 172; 285-291

KAWAJIRI, K., NAKACHI, K., & IMAI, K. (1990) Identification of genetically high risk individuals to lung cancer by DNA polymorphisms of the cytochrome P450 A1 gene. FEBS Lett. 263:131-133

KELLER, W. (1842) On the conversion of benzoic acid into hippuric acid. Provincial Medical and Surgical Journal 4(13): 256–257

KELLEY, M. & VESSEY, D.A. (1990) The effects of ions on the conjugation of xenobiotic by the aralkyl-CoA and arylacetyl-CoA N-acyltransferases from bovine liver mitochondria. J Biochem Toxicol 5:125-135

KELLEY, M. & VESSEY, D.A. (1992) Structural comparison between the mitochondrial aralkyl-CoA and arylacetyl-CoA N-acyltransferases. Biochem J. 288:315-317

KOLVRAA, A.S. & GREGERSEN, N. (1986) Acyl-CoA: glycine Nacyltransferase: organelle localization and affinity toward straight and branchedchained acyl-CoA esters in rat liver. Biochem. Med. Metab. Biol. (36): 98-105

LAEMMLI, U.K. (1970) Cleavage of structural proteins during the assembly of the head of Bacteriophase T4. Nature. 227:680-685

LANPHER, B., BRUNETTI-PIERRI, N. & LEE, B. (2006) Inborn errors of metabolism: the flux from Mendelian to complex diseases. Nat Rev Genet. 7:449-60

LEVY, G. (1965) Pharmacokinetics of salicylate elimination in man. J. Pharm. Sci. 54:959-967

LISKA, D.J. (1998) The detoxification enzyme systems. Alternative Medicine Review. 3(3):187-198

MACHEREY-NAGEL (Germany) Purification of polyhistidine-tagged proteins; User manuel. Protine® Ni-TED 150 packed columns, Protine® Ni-TED 1000 packed columns, Protine® Ni-TED 2000 packed columns, Protine® Ni-TED resin. [Web;] www.mn-net.com (Date of access: Jan 2008)

MAWAL, Y.R. & QURESHI, I.A. (1994). Purification to homogeneity of mitochondrial acyl CoA:glycine N-acyltransferase from human liver. Biochem Biophys Res Commun. 205:1373-1379

MERZ, F., BOEHRINGER, D., SCHAFFITZEL, C., PREISSLER, S., HOFFMANN, A., MAIER, T., RUTKOWSKA, A., LOZZA, J., BAN, N., BUKAU, B. & DEUERLING, E. (2008) Molecular mechanism and structure of TF bound to the translational ribosome. The EMBO Journal. 27:1622-1632

MURRAY, R.K., GRANNER, D.K. & RODWELL, V.W. (2006) Harper's illustrated Biochemistry. 27th ed. New York : McCraw-Hill. 692 p.

NAKATA, K., TANAKA, Y., NAKANO, T., ADACHI, T., TANAKA, H., KUMINUMA, T. & ISHIKAWA, T. (2006) Nuclear Receptor mediated transcriptional regulation in Phase I, II and III xenobiotic metabolizing systems. Drug Metab Pharmacokinet. 21(6):437-457

NANDI, D.L., LUCAS, S.V. & WEBSTER, L.T. Jr. (1979) Benzoyl-coenzyme A:glycine N-acyltransferase and phenylacetyl-coenzyme A:glycine N-acyltransferase from bovine liver mitochondria. Purification and characterization. J Biol Chem. 254(15):7230-7237.

NISHIHARA, K.. KANEMORI, M., YANAGI, H., & YURA, T. (2000) Overexpression of TF prevents aggregation of recombinant proteins in Escherichia coli. Applied and Environmental Microbiology. 66(3):884-889

NOVAGEN (USA) Thrombin Kits. [Web;] www.merck-chemicals.com (Date of access: Jan 2008)

NOVAGEN (USA) BugBuster® protein extraction reagent. [Web;] www.merck-chemicals.com (Date of access: Jan 2008)

OKAZAKI, R., OKAZAKI, S., SUGIMOTO, K., KAINUMA, R., SUGINO, A. & IWATSUKI, R. (1968) *In vivo* Mechanism of DNA chain growth. Proc. Natl.acad. USA 59, 598

OZAND, P.T. & GASCON, G.G. (1991) Topical Review Article: Organic Acidurias: A Review. Part 1. J Child Neurol. 6:288-303

PARKINSON, A. (2003) In: Casarett & Doull's Essentials of Toxicology, Klaasen, C.D. and Wakins, J.B. McGRaw-Hill Companies, Inc.

PASTORES, G.M. & BARNETT, N.L. (2005) Current emerging therapies for the lysosomal storage disorders. Expert Opin Emerg Drugs. 10:891-902

PECHURSKAYA, T.A., LUKASHEVICH, O.P., GILEP, A.A. & Usanov, S.A. (2008) Engineering, Expression and Purification of 'Soluble' Human Cytochrome P45017alpha and its Functional Characteristics. 73(7):806-11

PEQLAB (Denmark) E.Z.N.A. ® Plasmid Mini preparation kit 1 (Classic Line); Instruction manual. [Web;] www.peqlab.de (Date of access: Jan 2008)

PRINZ, W.A., ASLUND, F., HOLMGREN, A. & BECKWITH, J (1997) J. Biol. Chem. 272, 15661-15667

PROMEGA (USA) Subcloning Notebook Guide BR152: Vital tools and techniques for transferring your insert between vectors.

PROMEGA (USA) PureYield™ Plasmid Midi preparation procedure system. [Web;] www.promega.com (Date of access: Jan 2008)

QIAGEN (Germany) RNAlater™ Handbook: RNAlater™ TissueProtect Tubes. RNAlater™ RNA Stabilization Reagent for stabilization and protection of RNA in Tisssue. [Web;] www.qiagen.com (Date of access: Jan 2008)

QIAGEN (Germany) QIAquick® Spin Handbook for QIAquick PCR purification kit. QIAquick nucleotide removal kit and QIAquick gel extraction kit. [Web;] www.qiagen.com (Date of access: Jan 2008)

QIAGEN (Germany) RNeasy® Micro Handbook: For Isolation of Total RNA from microdisected tissue, small amounts of tissue (e.g., fine needle aspirators), small amount of fibrous tissues (e.g., heart, muscle), small number of cells (e.g., FACS sorted cells) and for RNA cleanup and concentration. [Web;] www.qiagen.com (Date of access: Jan 2008)

ROLLAND, M.O., DIVRY, P., ZABOT, M.T., GUIBAUD, P., GOMEZ, S., LACHUAX, A. and LORAS. I. (1985) Isolated 3-methylcrotonyl-CoA carboxylase deficiency in a 16-month-old child. J Inher Metab Dis. 14:838-839

SAHDEV, S., KHATTAR, S.K & SAINI, K.S. (2008) Production of the active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol Cell Biology. 307:249-264

SAMBROOK, J. & RUSSEL, D.W. (2005) Molecular Cloning: A laboratory manual 3rd ed. Cold Spring Harbor. New York

SARTORIUS STEDIUM BIOTECH (2007) Vivaspin 6 and 20 ml: Technical data and operating instructions.

SAUDUBRAY, J. & CHARPENTIER. C. (1995) Clinical Phenotypes: Diagnosis/algoriths. In Scriver CR, Beaudet AL, Sly WS and Valle eds. The molecular basis of inherited diseases, 7th edition, New York: McGraw-Hill, 53-228

SCHACHTER, D. & TAGGARD, J.V. (1953) Glycine N-acylase: purification and properties. J Biol Chem. 208(1):263-275

SMITH, P.K., KROHN, R.I., HERMANSON, G.T., MALLIA, A.K., GARTNER, F.H., PROVENZANO, M.D., GOEKE, N.M., OLSEN, B.J. & KLENK, D.C. (1985) Measurement of protein using bicinchoninic acid. Anal Biochem. 150:75-85

STOLLER, G., RUCKNAGEL, K.P., NIERHAUS, K.N., SCHMID, F.X., FISCHER, G. & RAHFELS, J.U. (1995) A ribosome-associated peptidyl-propyl cis/trans isomerase identified as the TF. EMBO J. 14:4939-4948

TAKARA (Japan) pCold Vector [Web;] www.takara-bio.com (Date of access: Jan 2008)

TAKARA (Japan) Chaperone plasmid [Web;] www.takara-bio.com (Date of access: Jan 2008)

TANAKA, K. & ISSELBACHER, K. (1967) The isolation and identification of N-isovalerylglycine from urine of patients with Isovaleric acidemia. J Biol Chem.;242:2966-2972

THERMO SCIENTIFIC (2008) Thermo Scientific Technology Overview: NanoDrop Spectrophotometer and Fluorospectrophotometers. [Web:] www.nanodrop.com [Date of use: 28 Aug 2009]

THOMAS, J.G., AYLING, A. & BENEYX, F. (1997) Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from *E. coli.* To fold or to refold. Appl Biochem Biotechnol. 66(3):197-238

TSAI, M.Y., JOHNSON, D.D., SWEETMAN, L. & BERRY, S.A. (1989) Two siblings with biotin-resistant 3-methylcrotonyl-coenzyme A carboxylase deficiency. J Pediatr, 115, 110-3

VAN DER WESTHUIZEN, F.H., PRETORIUS, P.P. & ERASMUS, E. (2000) The utilization of alanine, glutamic acid and serine as amino acid substrates for Glycine N-acyltransferase. J Biochem Molecular Toxicology. 14:102-109

VALENT, Q.A., KENDELL, D.A., HIGH, S., KUSTERS, R., OUDEGA, B. & LUIRINK, J. (1995) Early events in the preprotein recognition in *E. coli*: interaction of SRP and TF with nascent polypeptides. EMBO J. 14:5494-5505

VERMEULEN, N.P.E. (1996) Role of metabolism in chemical toxicity. In: Ioannides C, ed. Cytochromes P450: Metabolic and Toxicological Aspects. Boca Raton, FL: CRC Press, Inc; 29-53.

VESSEY, D.A. & LAU, E. (1996) Determination of the sequence of aralkyl acyl-CoA:amino acid N-acyltransferase from bovine liver mitochondria. J. Biochem. Toxicol. 11:211-215

VETTING, M.W., DE CARVALHO, L.P.S., SUBRAY. M.Y., YU, M., HEDGE, S.S., MAGNET, S., RODERICK, S.L. & BLANHARD, J.S. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys. 433(1):212-226.

WACHTER, V.J., WU, C.U. & BENET, L.Z. (1995) Overlapping substrates specificities and tissue distribution of cytochrome P450 and P-glycoprotein: Implications for drug delivery and activity in cancer chemotherapy. Mol Carcinog; 13:129-134

WALUK, D.P., SCHULTZ, N. & HUNT, M.C. (2009) Human glycine N-acytransferase-like-1 is involved in the production of glycine conjugated fatty acids. Chem. Phys.Lipids 160S, S22-S24

WALUK, D.P., SCHULTZ, N. & HUNT, M.C. (2010) Identification of glycine Nacyltransferase-like 2 (GLYATL2) as a transferase that produces N-acyl glycines in humans. The FASEB journal 24(8):2795-803

WEBSTER, L.T., SIDDIQUI, U.A., LUCAS, S.V., STRONG, J.M. & MIEYAL, J.J (1976) Identification of separate acyl-CoA:glycine and acyl-CoA:L-glutamine N-acyltransferase activities in mitochondrial fractions from liver of rhesus monkeys and man. J Biol Chem. 251 (11): 3352-3358

WEBSTER, L.T. (1981) Benzoyl-CoA: amino acid and phenylacetic-CoA: amino acid N-acyltransferases. J Biol Chem. 77:301-308

WILLIAMS, R.T. (1947) Detoxi cation Mechanisms. John Wiley and Sons, Inc, New York.

WIND, A., LARKIN, D.M., GREEN, C.A., ELLIOT, J.S., OLMSTEAD, C.A., CHIU, R., SCHEIN, J.E., MARRA, M.A., WOMACK, J.E. & LEWIN HA. (2005) A high–resolution whole-genome cattle-human comparative map reveals details of mammalian chromosome evolution. Proc Natl Acad Aci USA 102:18526-18531

ZHANG, H., LANG, Q., Li, J., ZHONG, Z., XIE, F., YE, G., WAN, B. & YU, L. (2007). Molecular cloning and characterization of the novel human glycine Nacyltransferase gene GLYATL1, which activates transcriptional activity of HSE pathway. Int J Mol Sci. 8:433-444

# APPENDIX A

## LIST OF KITS, ENZYMES, REAGENTS, ANTIBIOTICS AND LADDERS

KITS:

| **Item & description:** | **Catalogue number** | **Manufacturer** |
|---|---|---|
| BugBuster® Protein Extraction Reagent | 70921-4 | Novagen |
| Chaperone Plasmid DNA Set | 3340 | Takara |
| QIA*quick*® Gel Extraction Kit | 28706 | Qiagen |
| MSB® Spin PCRapace Cleanup Kit | 10202202 | Invitek |
| pColdTF DNA Vector | 3365 | Takara |
| pColdIII DNA Vector | 3363 | Takara |
| PeqGOLD Plasmid Miniprep Kit I (Classic-Line) | 12.6942.01 | PEQLAB |
| Protino Ni-Ted Column Affinity Chromatography Kit | 745120.25 | Macherey-Nagel |
| PureYield™ Plasmid Midiprep System | A2492 | Promega |
| RNeasy® Micro Kit | 74104 | Qiagen |
| Vivaspin 6, 10,000 MWCO | VS0601 | Sartorius Stedim |
| Vivaspin 20, 10,000 MWCO | VS2001 | Sartorius Stedim |

Enzymes:

| **Item & description:** | **Catalogue number** | **Manufacturer** |
|---|---|---|
| BamHI | ER0051 | Fermentas |
| Benzonase Nuclease | 70746-3 | Merck |
| Cloned AMV Reverse Transcriptase | 12328-019 | Invitrogen |
| Complete Mini, EDTA-free Protease Inhibitor Cocktail | 11 836 170 001 | Roche |
| DNase | M6-109 | Promega |
| EcoRI | ER0271 | Fermentas |
| HindIII | ER0501 | Fermentas |
| Lysozyme | 71110-4 | Novagen |
| NcoI | ER0572 | Fermentas |

| | | |
|---|---|---|
| NdeI | ER0582 | Fermentas |
| Phusion High Fidelity DNA Polymerase | F530L | Finnzymes |
| PstI | ER0611 | Fermentas |
| SalI | ER0641 | Fermentas |
| Shrimp Alkaline Phosphatase (SAP) | EF0511 | Fermentas |
| T4 DNA Ligase | M1-809 | Promega |
| Thrombin | 69671-3 | Novagen |
| XhoI | ER0691 | Fermentas |

Reagents:

| **Item & description:** | **Catalogue number** | **Manufacturer** |
|---|---|---|
| Absolute ethanol | 1009832500 | Merck |
| Acrylamide | A2400 | Melford |
| Agarose | D1LE/ H111206 | Hispanagar |
| Ammonium Persulphate | 17874 | Pierce |
| Agar | HG00BX10.500 | Merck |
| Benzoyl-CoA | B1638 | Sigma-Aldrigh |
| β-Mercaptoethanol | 44143 | BDH |
| Bicinchonini acid solution (BCA) | B9643 | Sigma-Aldrigh |
| Bovine serum albumin (BSA) | 23209 | Pierce |
| Bromophenol blue | 064/0811B | BDH |
| CAPS (3-(Cyclohexylamino)-1-propane-sulfonic acid | C6070 | Sigma – Aldrich |
| Chloroform | C2432-500 | Sigma – Aldrich |
| Copper (II) sulphate Solution (CuSO4.5H2O) | C2284 | Sigma-Aldrigh |
| Dimethylsulphoxide (DMSO) | D2650 | Sigma-Aldrigh |
| dNTPs (25mM each) | DM105B/R1129 | Fermentas |
| DTNB (5,5'-Dithiobis 2-nitro-benzoic-acid) | D8130-10G | Sigma-Aldrigh |
| EDTA (Ethylenediaminetetraacetic acid) | 3685 | Merck |

| | | |
|---|---|---|
| Ethidium bromide | 200-271 | Roche |
| Formaldehyde | 47629 | Sigma – Aldrich |
| Glacial acetic acid | BB100017P | BDH |
| Glycerol | 1.4092.000 | Merck |
| Glycine | 1.04169.100 | Fermentas |
| HEPES (*4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid* ) | H3537 | Sigma-Aldrich |
| IPTG (Isopropyl β-D-1-thiogalactopyranoside) | V395A | Promega |
| Isopropanol | BB1022246L | BDH |
| Magnesium Chloride (MgCl) | 8.14733.0500 | Merck |
| Magnesium Sulphate | M.2643.500G | Sigma – Aldrich |
| Methanol | 11.06009.2500 | Merck |
| MOPS (3-(N-morpholino) propanesulfonic acid) | 11124684001 | Roche |
| Oligo (dT) 20 | 04387 | Sigma-Aldrich |
| PBS (Phosphate Buffered Saline) pellets | BR0014G | Oxoid |
| Potassium Chloride (KCl) | 60035 | Sigma – Aldrich |
| Sodium Acetate | 71183 | Fluka |
| Sodium Chloride (NaCl) | AB004936 | Merck |
| Sodium Dodecyl Sulphate (SDS) | L4390.1 | Sigma |
| Sodium Hydroxide Pellets (NaOH) | 1064980500 | Merck |
| TABS (N-tris(Hydroxymethyl)methyl-4-aminobutanesulfonic acid) | T1302 | Sigma – Aldrich |
| TEMED (Tetramethylethylenediamine) | 17919 | Pierce |
| Tris-base | 11814273001 | Roche |
| Tryptose | 1.10676 | Merck |
| Yeast Extract | BX 6 | Merck |

Antibiotics:

| **Item & description:** | **Catalogue number** | **Manufacturer** |
|---|---|---|
| Ampicillin | A8351 | Sigma – Aldrich |
| Chloramphenicol | C0378 | Sigma – Aldrich |
| L-arabinose | A3256 | Sigma – Aldrich |
| Tetracycline | T3258 | Sigma – Aldrich |

Ladders:

| **Item & description:** | **Catalogue number** | **Manufacturer** |
|---|---|---|
| GeneRuler™ DNA Ladder Mix | SM0339 | Fermentas |
| PageRuler™ Prestained Protein Ladder Plus | SM1811 | Fermentas |
| 6 x Loading dye (5x1 ml) | R0611 | Fermentas |

# APPENDIX B

## ABBREVIATIONS AND SYMBOLS

**A**

| | |
|---|---|
| α | Alpha |
| A, a | Adenine, a purine nucleotide |
| $A_{260}/A_{280}$ | Ratio of absorbance measured at 260 nm and 280 nm |
| ABC | ATP-binding cassettes |
| ADP | Adenosine diphosphate |
| Ala | Alanine |
| AMV | Avian myeblastosis virus |
| Arg | Arginine |
| Asn | Asparagine |
| Asp | Aspartic acid |
| ATP | Adenosine triphosphate |

**B**

| | |
|---|---|
| B | Beta |
| ß-ME | ß-Mercaptoethanol |
| BCA | Bicinchonicic acid |
| bGLYAT | Bovine glycine N-acyltransferase |
| bp | Base pairs (nucleotides) |
| BSA | Bovine serum albumin |

**C**

| | |
|---|---|
| C, c | Cytosine, a pyrimidine nucleotide |
| C, Cys | Cysteine |
| °C | Degree Celsius |
| cDNA | Complementary deoxyribonucleic acid |
| $cm^2$ | Square centimetre |
| CoA | Coenzyme A |
| CoA-SH | Acyl-Coenzyme A |
| $CO_2$ | Carbon dioxide |
| COOH | Carboxylic acid |
| $Cu^{2+}$ | Copper ion |
| $CuSO_4.5H_20$ | Copper II sulphate solution |

**D**

| | |
|---|---|
| $ddH_2O$ | Double distilled water |
| dNTP | Deoxynucleoside triphosphate |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| DNS | Deoksieribonukleïensuur |
| dsDNA | Double-stranded DNA |
| DTNB | - 5,5'-Dithio-Bis (2-Nitrobenzoic Acid) (Elman's reagent) |
| DTT | Dithiotreitol |

**E**

| | |
|---|---|
| EDTA | Ethylenediaminetetraacetic acid |
| e.g. | *exempli gratia*, for example |
| etc. | *et cetera*, and so forth |
| et al. | *et altera*, and others |
| EtOH | Ethanol |
| EtBr | Ethidium bromide |
| *E. coli* | Eschericia coli |

**F**

| | |
|---|---|
| For, fwd | Forward |
| FBS | Fetal bovine serum |
| FA | Formaldehyde |

**G**

| | |
|---|---|
| G, g | Guanine, a purine nucleotide |
| G | Gram (s) |
| x g | Gravitational field (g = 9,81 $m.s^{-2}$) |
| gDNA | Genomic deoxyribonucleic acid |
| GITC | Guanidine isothiocyanate |
| GLYAT | Glycine N-acyltransferase |
| hGLYATL1 | Human like 1 glycine N-acyltransferase |
| hGLYATL1 | Human like 2 glycine N-acyltransferase |
| Gln | Glutamine |
| Glu | Glutamic acid |
| Gly | Glycine |
| GNAT | GCN5-related N-acetyltransferase |

**H**

| | |
|---|---|
| $H^+$ | Hydrogen ion |
| $HCO_3^-$ | Bicarbonate ion |
| hGLYATa | Human GLYAT isoform a |
| hGLYATb | Human GLYAT isoform b |
| hGLYAT | Human Glycine N-acyltransferase |
| His | Histidine |
| His-tag | Histidine tag |
| $H_2O$ | Water |

**I**

| | |
|---|---|
| i.e. | *id est,* that is |
| IEM | Inborn errors of metabolism |
| Ile | IsoLaucine |
| IPTG | Isoprpyl-β-D-thiogalactopyranoside |

**J**

**K**

| | |
|---|---|
| Kb | Kilobase (1000 bp) |
| KCl | Potassium Chloride |
| Kg | Kilogram |
| $K^+$ | Potassium ion |
| Km | Michaelis Menton constant |

| | |
|---|---|
| kDa | Kilo Dalton |

**L**

| | |
|---|---|
| LB | Luria-Bertani |
| Lau | Laucine |
| Ltd. | Limited |
| Lys | Lysine |
| LEW | Lysis/Equilibrium/Wash buffer |

**M**

| | |
|---|---|
| M | Molar |
| ME | Mercaptoethanol |
| Met | Methionine |
| mg | Milligram (s) |
| $Mg^{2+}$ | Magnesium ion |
| $MgCl_2$ | Magnesium chloride |
| min | Minutes |
| ml | Millilitre (s) |
| mM | Millimolar (1 $x10^{-3}$) |
| MOPS | 3-[N-morpholino] propanesulfonic acid |
| mRNA | Messenger Ribonucleic acid |
| MWCO | Molecular weight cut-off |

**N**

| | |
|---|---|
| Na | Sodium |
| NaCl | Sodium chloride |
| $NAD^+$ | Nicotinamide adenine dinucleotide (oxidised) |
| NADH | Nicotinamide adenine dinucleotide (reduced) |
| $NADP^+$ | Nicotinamide adenine dinucleotide phosphate (oxidised) |
| NADPH | Nicotinamide adenine dinucleotide phosphate (reduced) |
| ng/µl | Nanogram per Microlitre |
| $NH_2$ | Amine group |
| Ni | Nickel |
| nm | Nanometre |
| Nr. | Number |
| NWU | North-West Univeristy |

**O**

| | |
|---|---|
| $O_2$ | Oxygen |
| $O_2^-$ | Superoxide |
| OD | Optical density |
| $OH^-$ | Hydroxy ion |
| $OH^\bullet$ | Hydrogen ion, reactive oxygen species |
| ORF | Open reading frame |
| Ω | Ohm |

**P**

| | |
|---|---|
| P | Phosphate |
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| pH | Percentage hydrogen |
| Phe | Phenylalanine |

| | |
|---|---|
| Pmol | Picomol |
| Pro | Proline |
| PSB | Protein solvent buffer |
| % | Percent |

**Q**

**R**

| | |
|---|---|
| RNA | Ribonucleic acid |
| RT | Reverse transcriptase |
| rRNA | Ribosomal ribonucleic acid |
| rev | Reverse |
| ROS | Reactive oxygen species |
| ref. seq | Reference sequence |
| Rpm | Revolutions per minute |
| ® | Registered |

**S**

| | |
|---|---|
| S.A. | South Africa |
| SDS | Sodium dodecyl sulphate |
| SDS-PAGE | Sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| Sec. | Seconds |
| Ser | Serine |
| SNP | Single nucleotide polymorphism |
| ssRNA | Single-stranded RNA |

**T**

| | |
|---|---|
| T, t | Thymine, apyrimidine nucleotide |
| TAE | Tris acetic acid ethylenediaminetetraacetic acid |
| TBS | Tris-buffered-saline |
| TEE | Transcription enhancement element |
| TF | TF |
| TGS | Tris-Glycine-SDS |
| Thr | Threonine |
| Tm | Calculated melting temperature |
| ™ | Trademark |
| Tris® | Tris (hydroxymethyl) aminomethane |
| tRNA | Transport ribonucleic acid |
| Trp | Tryptophan |
| TSB | Transformation and Storage Buffer |
| TSBG | TSB with 20% glucose |
| Tyr | Tyrosine |

**U**

| | |
|---|---|
| µ | Micro |
| µM | Micro molar |
| µl | Microliter (s) |
| µg | Microgram (s) |
| UV | Ultra violet |
| UTR | Untranslated region |
| µl/ml | Microliter per milliliter |
| U/µl | Units per microliter |

**V**

| | |
|---|---|
| Val | Valine |
| V | Volt |

**W, X, Y, Z**

**OTHER**

| | |
|---|---|
| 18 S | rRNA Ribosomal ribonucleic acid |
| 28 S | rRNA Ribosomal ribonucleic acid |

# APPENDIX C

**Enzymatic characterisation and elucidation of the catalytic mechanism of a recombinant bovine glycine N-acyltransferase**

DMD Fast Forward. Published on November 9, 2011 as doi:10.1124/dmd.111.041657



# Enzymatic characterisation and elucidation of the catalytic mechanism of a recombinant bovine glycine N-acyltransferase

Christoffel P.S. Badenhorst, Maritza Jooste, Alberdina A. van Dijk

Biochemistry Department, North-West University, Potchefstroom, South Africa (CPSB, MJ, AAvD)

**The catalytic mechanism of GLYAT**

Corresponding author:

A. A. Van Dijk

Biochemistry Department

North-West University

Private Bag X6001

Potchefstroom, 2520

South Africa

Tel: +27 (0)18 299-2317

Fax: +27 (0)18 299-2363

E-mail: Albie.VanDijk@nwu.ac.za

Number of pages: 27

Number of figures: 5

Number of Tables: 2

Number of references: 37

Number of words in Abstract: 225

Number of words in Introduction: 703

Number of words in Discussion: 1123

**Nonstandard abbreviations**

CASTOR, Coenzyme A Sequestration, Toxicity or Redistribution; DTNB, 5,5'-dithiobis(2-nitrobenzoic acid); GLYAT, Glycine N-acyltransferase; GNAT, Gcn5-related N-acyltransferase; IPTG, isopropyl-1-thio-β-D-galactopyranoside; $K_M$, Michaelis constant; SNP, Single nucleotide polymorphism; Vmax, maximum velocity.

**Abstract**

Glycine conjugation, a phase II detoxification process, is catalysed by glycine N-acyltransferase (GLYAT; E.C. 2.3.1.13). GLYAT detoxifies various xenobiotics, such as benzoic acid, and endogenous organic acids, such as isovaleric acid, which makes GLYAT important in the management of organic acidemias in humans. We cloned the open reading frame encoding the bovine orthologue of GLYAT from bovine liver mRNA into the bacterial expression vector pColdIII. The recombinant enzyme was expressed, partially purified and enzymatically characterised. Protein modelling was used to predict E226 of bovine GLYAT to be catalytically important. This was assessed by constructing an E226Q mutant and comparing its enzyme kinetics to that of the wild-type recombinant bovine GLYAT. The Michaelis constants for benzoyl-coenzyme A and glycine were determined and were similar for wild-type recombinant GLYAT, E226Q recombinant GLYAT, and GLYAT present in bovine liver. At pH 8.0, the E226Q mutant GLYAT had decreased activity, which could be compensated for by increasing the reaction pH. This suggested a catalytic mechanism in which E226 functions to deprotonate glycine, facilitating nucleophilic attack on the acyl-coenzyme A. The recombinant bovine GLYAT enzyme, combined with this new understanding of its active site and reaction mechanism, could be a powerful tool to investigate the functional significance of GLYAT sequence variations. Eventually, this should facilitate investigations into the impact of known and novel sequence variations in the human GLYAT gene.

**Introduction**

Humans and other animals are constantly exposed to various toxic substances. Homeostasis is maintained by the detoxification and elimination of these toxins. Detoxification is commonly divided into phase I and phase II detoxification enzyme systems (Liska, 1998). Phase I detoxification enzymes convert toxic compounds to secondary metabolites that are frequently more chemically reactive and thus more toxic than the original compounds. Phase II detoxification enzymes decrease the toxicity of the secondary metabolites by means of conjugation to soluble carrier molecules such as glycine or glutathione. The resulting conjugates are usually less toxic, more water-soluble and more readily excreted in urine than the secondary metabolites. In humans, impaired detoxification has been associated with diseases such as cancer and with adverse reactions to pharmaceutical drugs (Campbell et al., 1988, Liska, 1998, Wallig, 2004).

Glycine N-acyltransferase (GLYAT, E.C. 2.3.1.13) is a phase II detoxification enzyme found in the liver and kidney mitochondria of mammals (Schachter & Taggart, 1954). GLYAT is a member of the Gcn5-related N-acyltransferase superfamily (GNAT superfamily), one of the largest and most functionally diverse superfamilies of enzymes known (Vetting et al., 2005). GLYAT uses glycine and an acyl-coenzyme A as substrates, forming free coenzyme A and an acylglycine as products (Dyda et al., 2000, van der Westhuizen et al., 2000, Vetting et al., 2005). The acylglycines are less toxic than the unconjugated organic acids and are more readily excreted into urine by the kidneys (Duffy et al., 1995, Tanaka & Isselbacher, 1967). Examples of xenobiotic metabolites that are detoxified by conjugation to glycine include salicylic acid, benzoic acid, and methyl-benzoic acid, a metabolite of the industrial solvent xylene (Campbell et al., 1988, Duffy et al., 1995). Studies have shown that in humans, metabolism of aspirin to its glycine conjugate may be impaired by exposure to xylene, which is detoxified via the same pathway (Campbell et al., 1988). Organic acids of endogenous origin, including isovaleric acid, 3-methylcrotonic acid, tiglic acid and hexanoic acid, are also detoxified by conjugation to glycine. For this reason glycine conjugation is of key importance to the management of inherited organic acidemias in humans (Kolvraa & Gregersen, 1986, Ogier & Saudubray, 2002). For example, isovaleric acidemia is

treated by means of glycine supplementation, which enhances formation and excretion of isovalerylglycine (Tanaka et al., 1966, Tanaka & Isselbacher, 1967). Isovaleric acidemia is only one of a large group of human diseases known as CASTOR (*C*oenzyme *A* *s*equestration, *to*xicity or *r*edistribution) disorders (Mitchell et al., 2008), which represent a major segment of biochemical genetics. One of the primary mechanisms of pathogenesis in CASTOR disorders is depletion of free coenzyme A, which derails cellular metabolism. GLYAT converts the accumulated acyl-coenzyme As to acylglycines and free coenzyme A, restoring levels of free coenzyme A and carnitine (Sakuma, 1991). Significant inter-individual variation in glycine conjugation capacity has been demonstrated using human liver samples (Temellini et al., 1993). The basis for this variability is not understood, but genetic variations in the coding sequence may be a factor. Six non-synonymous SNPs have been identified in the open reading frame of the human GLYAT gene. However, it is not yet understood if, or how, these variations influence enzyme function (Lino Cardenas et al., 2010, Yamamoto et al., 2009).

Because of the large number of compounds metabolised by GLYAT, the human orthologue is of clinical interest. However, we have not yet been able to express an enzymatically active recombinant human GLYAT. Here we report the bacterial expression and enzymatic investigation of a recombinant bovine GLYAT. Bovine GLYAT is an enzyme expressed in bovine liver and kidney mitochondria, with a molecular mass reported to be between 33 kDa and 36 kDa (Nandi et al., 1979, van der Westhuizen et al., 2000). Investigations of human and bovine GLYAT have shown that these enzymes are similar in terms of molecular mass, reaction kinetics, and substrate specificity (Bartlett & Gompertz, 1974, Kelley & Vessey, 1993, van der Westhuizen et al., 2000). The recombinant bovine GLYAT, combined with molecular modelling and site-directed mutagenesis, was used to investigate the catalytic mechanism employed by GLYAT. The data suggested that residue $Glu^{226}$ of bovine GLYAT serves as a general base catalyst. The identification of this catalytic residue provides the first insights into the catalytic mechanism and active site location of the GLYAT enzymes.

## Methods

### *Sequence analysis and molecular modelling*

The bovine GLYAT amino acid sequence (Refseq accession number NP_803479) was submitted to the GenTHREADER server (Jones, 1999) for identification of potential structural homologues. An uncharacterised protein from *Drosophila melanogaster,* with PDB ID 1SQH, was identified as the best homologue and used for molecular modelling. Although this protein is only 13% identical to bovine GLYAT, it is structurally very similar to other GNAT enzymes, and was used as template structure because of the exceptional conservation of structure in the GNAT superfamily of acyltransferases, despite there being virtually no sequence similarity between some members in the superfamily (Vetting et al., 2005). The alignment generated by GenTHREADER was used with the structure of 1SQH as input for model generation using MODELLER 9.3 (Eswar et al., 2008). For side chain modelling SCWRL 3.0 was used (Wang et al., 2008). The molecular model was superimposed with the structures of serotonin N-acetyltransferase (PDB ID 1CJW), diamine N-acetyltransferase (PDB ID 2Q4V) and Esa1 (PDB ID 1GHE) using the matchmaker algorithm of UCSF Chimera (Pettersen et al., 2004). UCSF Chimera was used to generate images of the molecular model. The bovine GLYAT amino acid sequence was also submitted to a BLAST (www.ncbi.nlm.nih.gov) search and the homologues aligned using CLUSTALX 2.0.10.

### *Cloning of the bovine GLYAT open reading frame into the bacterial expression vector pColdIII*

Total RNA was isolated from bovine liver tissue using the RNeasy Mini Kit (QIAGEN, Valencia, CA, USA). Cloned AMV reverse transcriptase (Invitrogen, Carlsbad, CA, USA) was used to generate cDNA from the bovine liver mRNA. The open reading frame of bovine GLYAT was amplified from the cDNA using oligonucleotide primers (5'-GCC GCA TAT GAT GTT CCT GCT GC-3' and 5'-CAT CTC GAG TCA CAG AGG CTC AC-'3) that contained NdeI and XhoI restriction endonuclease recognition sites to facilitate cloning into pColdIII (Takara, Madison, WI, USA). Oligonucleotide primers were obtained from Inqaba Biotechnical Industries, Pretoria, South Africa. The pColdIII vector used was modified to encode a C-

terminal hexahistidine tag after the XhoI site. The recombinant plasmid was sequenced to confirm that bovine GLYAT had been cloned without any sequence aberrations.

***Construction of the E226Q mutant recombinant bovine GLYAT***

Site-directed mutagenesis using a mega-primer method (Aiyar & Leis, 1993) was used to generate the E226Q mutant coding sequence. In a first PCR the mutagenic oligonucleotide primer 5'-CCA GAC GGG ACA GAT GCG GAT GG-3' was used with the reverse primer 5'-CTT CTC GAG AGG CTC ACA GTT CCA CTG G-3' to generate a 240 bp amplicon. This amplicon was gel purified and used in a second PCR reaction with the forward primer 5'-GCC GCA TAT GAT GTT CCT GCT GC-3' to generate a full-length mutated GLYAT coding sequence. The mutated amplicon was gel purified, digested with NdeI and XhoI, and cloned into pColdIII.

***Expression and nickel-affinity purification of wild-type and E226Q recombinant bovine GLYAT***

The pColdIII-bovineGLYAT plasmid was introduced into Origami cells (Novagen, Madison, WI, USA) by electroporation. Expression of the recombinant GLYAT was carried out as follows. The cells from 50 ml overnight cultures in LB medium, containing 100 μg/ml ampicillin, were harvested by centrifugation at 4000 *g* for 5 minutes. The cells were resuspended in 200 ml of LB medium containing 50 μg/ml ampicillin. The cultures were gently shaken at 15 °C for one hour before IPTG (isopropyl-1-thio-β-D-galactopyranoside) was added to a final concentration of 0.5 mM. The cultures were incubated at 15 °C for 24 hours with vigorous shaking. Cells were then harvested by centrifugation at 4000 *g* for 20 minutes. The cell pellets were resuspended in 5 ml BugBuster protein extraction reagent (Novagen, Madison, WI, USA) containing 30 U/ml Lysozyme (Novagen, Madison, WI, USA) and 25 U/ml Benzonase nuclease (Novagen, Madison, WI, USA) followed by incubation at room temperature for 5 minutes. Insoluble material was removed by centrifugation at 12 000 *g* for 25 minutes at 4 °C. The cleared lysates were passed through Protino Ni-TED 2000 columns (Macherey-Nagel Inc., Bethlehem, PA, USA) equilibrated with buffer LEW (Macherey-Nagel Inc., Bethlehem, PA, USA). The columns were washed with 10 ml of

buffer LEW containing 20 mM imidazole. The bound protein was eluted from the columns in 9 ml of buffer EB (Macherey-Nagel Inc., Bethlehem, PA, USA), and added to Vivaspin 20 ultra-filtration devices (GE Healthcare, Björkgatan, Uppsala, Sweden). The proteins were concentrated to approximately 500 µl by centrifugation in a fixed angle centrifuge at 8000 *g* for 15 minutes. The proteins were then washed by adding 10 ml of 50 mM TrisHCl, pH 8.0, and repeating the centrifugation to again concentrate the solution to approximately 500 µl. Protein expression and purification was monitored by means of SDS-PAGE analyses and Coomassie brilliant blue staining (Laemmli, 1970).

***Preparation of an extract containing bovine liver mitochondrial GLYAT***

To prepare an extract containing bovine liver mitochondrial GLYAT, 100 g of liver tissue was homogenised in 400 ml of 0.13 M KCl. The homogenate was centrifuged for 10 minutes at 600 *g*. The supernatant was centrifuged again at 9 000 *g* for 10 minutes to isolate mitochondria, which were then lysed by three cycles of freezing and thawing in 10 ml of 0.13 M KCl. The lysates were clarified by centrifugation at 35 000 *g*, at 4 °C, for 2 hours. The GLYAT enzyme was further enriched from these lysates by collecting the fraction soluble between 40% and 60% ammonium sulfate. The precipitate was dissolved in 4 ml of 50 mM TrisHCl, pH 8.0, and dialysed overnight against 1000 ml of 50 mM TrisHCl, pH 8.0 (van der Westhuizen et al., 2000). This crude mitochondrial GLYAT preparation is from here on referred to as “bovine liver GLYAT”.

***Enzyme assays for determination of $K_M$ parameters***

Reaction mixtures were 400 µl in volume and contained, in addition to enzyme, 25 mM TrisHCl, pH 8.0, 0.1 mM 5,5’-dithiobis(2-nitrobenzoic acid) (DTNB), and varying concentrations of substrates (Kolvraa & Gregersen, 1986). The glycine concentration was varied from 2.5 mM to 20 mM, and the benzoyl-coenzyme A concentration was varied from 5 µM to 50 µM. The assays were carried out at 30 °C and the change in absorbance at 412 nm over the first four minutes was measured using a Uvicon XS spectrophotometer (van der Westhuizen et al., 2000). The absorbance change at 412 nm is the result of

reduction of DTNB by the liberated thiol group of coenzyme A, forming 2-nitro-5-thiobenzoate (NTB). NTB is a yellow species which absorbs at 412 nm with an extinction coefficient of 13.6 $mM^{-1}cm^{-1}$ (Ellman, 1959). To each reaction, 1 Unit of GLYAT activity was added. A unit of GLYAT activity was defined as a change in $A_{412}$ of 0.24 units, in four minutes, using cuvettes with a 1 cm light path, 20 mM glycine, and 100 μM benzoyl-coenzyme A. The amount of enzyme used per assay was defined in this way because the enzymes used were not purified to homogeneity. Although this precluded the determination of $V_{max}$ values, $K_M$ parameters could be determined using the partially purified enzyme preparations (Palmer, 2001, van der Westhuizen et al., 2000). All assays were performed in triplicate and the data analysed using SigmaPlot 11.0. To determine the apparent $K_M$ values for isovaleryl-, propionyl-, benzoyl- and octanoyl-coenzyme A (Sigma, St. Louis, MO, USA), the same conditions were used, except that the glycine concentration was fixed at 200 mM and the acyl-coenzyme A concentrations were varied from 10 μM to 600 μM (Nandi et al., 1979).

***Determining the pH dependence of the GLYAT enzymes***

Reaction mixtures were 100 μl in volume and consisted of 50 mM potassium phosphate buffer, 0.1 mM DTNB, 200 μM benzoyl-coenzyme A, and 200 mM glycine, at various pH values (Figure 4). Of the crude bovine liver GLYAT extract, 30 μg of protein was used per assay. An amount of protein that has activity comparable to the 30 μg of bovine liver GLYAT extract was used in each assay of the wild-type recombinant GLYAT. This was usually approximately 3 μg of protein. Because the proteins were not purified to homogeneity, SDS-PAGE analysis was generally used to determine the amount of E226Q recombinant GLYAT to use per assay. The mutant enzyme was diluted until the recombinant GLYAT band, as judged by SDS-PAGE, was of equal intensity to that in the wild-type recombinant GLYAT preparation. Because the same amount of enzyme is used in each assay in a particular experiment, pH is the only variable that influences activity. Reactions that contain all components except enzyme were also performed for control purposes. The reactions were monitored at 412 nm in 96-well plates on a BioTech plate reader and the accompanying Gen5 software. Measurements were made every 30 seconds for 6 minutes. Initial velocities were calculated and recorded as nM/min. Reactions were performed in triplicate

and the data plotted as nM/min against pH, using a logarithmic scale for the y-axis. Data was plotted using GraphPad Prism 4.02.

## Results

### *Prediction of the catalytic importance of Glu$^{226}$ of bovine GLYAT*

To investigate the catalytic mechanism of bovine GLYAT, a candidate for the catalytic residue had to be identified. A BLAST search using the bovine GLYAT amino acid sequence as query was performed. After removing sequences of significantly different length to bovine GLYAT, a multiple sequence alignment was performed (Figure 1A). The bovine GLYAT sequence was then used to construct a molecular model. By investigating the superposition of the GLYAT model with GNAT enzymes for which the catalytic residues are known, a putative catalytic residue was identified. The Glu$^{226}$ residue of bovine GLYAT coincided spatially (Figure 1B) with the catalytic residue, His$^{120}$, of serotonin N-acetyltransferase (Scheibner et al., 2002). Similar results were obtained when this superposition was done using other GNAT enzymes for which catalytic mechanisms are known (Figure 1B). The Glu$^{226}$ residue of bovine GLYAT was also conserved in the top 40 GLYAT homologues obtained by a BLAST search, suggesting the residue to be functionally significant (Figure 1A shows part of the multiple alignment).

### *Expression and purification of recombinant bovine GLYAT enzymes*

Both the wild-type and E226Q recombinant bovine GLYAT were expressed at high levels from the bacterial expression vector pColdIII. Most of the recombinant bovine GLYAT, both of the wild-type and the E226Q mutant, was insoluble (Figure 2A; GLYAT is indicated by an arrow). Soluble wild-type and E226Q recombinant GLYAT was obtained by means of nickel-affinity chromatography and ultra-filtration (Figure 2B). The lower bands in Figure 2B (indicated by the arrow) represent soluble recombinant bovine GLYAT enzymes and some co-purifying proteins. Expression of the recombinant bovine GLYAT enzymes with hexahistidine tags containing serine-glycine linkers of different lengths, to enhance the flexibility and accessibility of the tag, did not improve purification. Expression with the tags of different length did however result in size differences being observed for the lower bands in Figure 2B, indicating that these bands represent the recombinant bovine GLYAT enzymes (results not shown). The recombinant bovine

GLYAT proteins were not subjected to further purification, as the co-purified proteins did not seem to interfere with any subsequent investigations.

***Kinetic properties of the recombinant bovine GLYAT enzymes***

To characterise the wild-type and E226Q recombinant bovine GLYAT enzymes, Michaelis constants were determined using glycine and benzoyl-coenzyme A, and compared to those for GLYAT extracted from bovine liver mitochondria. Nonlinear regression was used to analyse the kinetic data, and Lineweaver-Burk plots were used to visually represent the data (Figure 3). The $K_M$ values for benzoyl-coenzyme A were similar for the bovine liver GLYAT and the recombinant bovine enzymes, at approximately 16 μM (Table 1). The $K_M$ values for glycine for bovine liver GLYAT and wild-type recombinant GLYAT were also similar, at approximately 2 mM. However, the $K_M$ value for glycine of the E226Q mutant was higher, at approximately 7 mM (Table 1). The apparent $K_M$ values (at 200 mM glycine) were also determined for bovine liver GLYAT and wild-type recombinant bovine GLYAT, using propionyl-coenzyme A, isovaleryl-coenzyme A, benzoyl-coenzyme A, and octanoyl-coenzyme A. These values were also comparable for the two enzymes (Table 2).

***The pH dependence of bovine GLYAT enzymes***

The catalytic importance of the $Glu^{226}$ residue of bovine GLYAT was investigated by determining the pH dependence of wild-type and E226Q mutant recombinant bovine GLYAT enzymes and comparing it to that of GLYAT extracted from bovine liver mitochondria. Increasing pH resulted in increased reaction rates for all three enzymes. Both the bovine liver GLYAT and wild-type recombinant bovine GLYAT enzymes had relatively low activity at pH 6.0 and activity increased with pH to a maximum at pH 7.5. As the pH increased further from 7.5 to 9.6, no significant increase in enzyme activity was observed (Figure 4). The activity of the E226Q mutant GLYAT enzyme increased as the pH was increased from 6.0 to 9.6. In a representative experiment the activity of the E226Q mutant bovine GLYAT was approximately 6 % of the activity of the wild-type recombinant bovine GLYAT, at pH 8.0. At higher pH values, the activity of the

mutant was increased significantly, with the E226Q mutant being approximately 111 % as active as the wild-type recombinant GLYAT, at pH 9.6 (Figure 4). This effect was consistently observed using different preparations of the enzymes. Because the recombinant bovine GLYAT enzymes were only partially purified, SDS-PAGE had to be used to compare the recombinant bovine GLYAT content of the wild-type and E226Q recombinant GLYAT preparations. This method worked well, but is not completely accurate, which explains the difference between the activities of the wild-type and E226Q mutant recombinant GLYAT at pH 9.6.

**Discussion**

The purpose of this study was to generate and enzymatically characterise a recombinant bovine GLYAT and to initiate investigations of the catalytic mechanism of the enzyme. It is important to understand the molecular and biochemical characteristics of the GLYAT family of enzymes, since these enzymes metabolise a wide range of endogenous and xenobiotic compounds, and may present as yet unknown targets for the therapeutic manipulation of inherited metabolic disorders or exposure to toxins. A recombinant bovine GLYAT was used in this study because it could be expressed in an enzymatically active form in *E. coli*. Our attempts to express enzymatically active human GLYAT in bacteria have thus far been unsuccessful. This perhaps correlates with the observation that GLYAT isolated from human liver is less active and less stable than GLYAT isolated from bovine liver (Mawal & Qureshi, 1994, Nandi et al., 1979, van der Westhuizen et al., 2000). However, the reaction kinetics and substrate specificity of the human and bovine GLYAT enzymes are similar, and we started investigations of the GLYAT domain architecture, catalytic mechanism, and functional residues, using the bovine orthologue (Bartlett & Gompertz, 1974, Kolvraa & Gregersen, 1986, Mawal & Qureshi, 1994, Nandi et al., 1979, Schachter & Taggart, 1954, van der Westhuizen et al., 2000).

Our results show that recombinant bovine GLYAT, expressed in *E. coli,* has enzymatic properties similar to those of GLYAT present in an extract of bovine liver mitochondria. The $K_M$ values for several acyl-coenzyme A substrates and glycine were determined and found to be similar for these two enzymes (Tables 1 and 2). In the literature, there is great variation in the $K_M$ values reported for bovine liver GLYAT. The $K_M$ values for benzoyl-coenzyme A range from 9 μM to 310 μM, and those for glycine from 2 mM to 15 mM (Bartlett & Gompertz, 1974, Gregersen et al., 1986, Kelley & Vessey, 1986, Kelley & Vessey, 1993, Kelley & Vessey, 1994, Kolvraa & Gregersen, 1986, Mawal & Qureshi, 1994, Nandi et al., 1979, Schachter & Taggart, 1954, van der Westhuizen et al., 2000). The $K_M$ values we determined for the bovine liver GLYAT and recombinant bovine GLYAT (approximately 16 μM for benzoyl-coenzyme A and 2 mM for glycine) fall within the range reported in the literature. The similarity of recombinant bovine GLYAT to bovine liver GLYAT, in terms of $K_M$ parameters, suggested that the recombinant enzyme could

be a valuable tool for investigation of the catalytic residues of bovine GLYAT. When glycine was omitted from the enzyme assays, no activity could be observed, confirming that the enzyme preparations were not contaminated with any proteins that non-specifically hydrolyse benzoyl-coenzyme A.

Based on the pH dependence of the human and bovine GLYAT enzymes, as reported in the literature, we anticipated that a general base catalyst is involved in the reaction (Mawal & Qureshi, 1994, Nandi et al., 1979, Schachter & Taggart, 1954). Our hypothesis, based on molecular modelling and the similarity of GLYAT to the GNAT superfamily of acyltransferases, was that the residue $Glu^{226}$ of bovine GLYAT served as a general base catalyst in the GLYAT reaction. To investigate this hypothesis, an E226Q mutant was expressed and enzymatically characterised. When similar amounts of wild-type and E226Q recombinant bovine GLYAT were assayed at pH 8.0, there was approximately a twentyfold difference in activity between the two enzymes. Similarity between the mutant and wild-type enzymes in terms of the $K_M$ values for glycine (approximately 7 mM and 2 mM, respectively) and benzoyl coenzyme A (approximately 18 μM and 16 μM, respectively) suggested that loss of substrate-binding ability of the mutant could not solely account for the lower activity of the mutant enzyme. This is because the assay mixture contained 200 μM benzoyl-coenzyme A and 200 mM glycine, concentrations much higher than the $K_M$ values of both wild-type and E226Q recombinant GLYAT, meaning that both substrates were saturating.

The loss of catalytic activity displayed by the E226Q mutant GLYAT is not sufficient evidence to conclude that our bioinformatic analyses and prediction of the catalytic importance of $Glu^{226}$ of bovine GLYAT are valid. This is because the mutation may simply have altered some structural component of the enzyme, lowering the catalytic rate. Since our objective was to demonstrate that E226 serves as a general base catalyst, the pH dependence of the mutant and wild-type enzymes was investigated. The wild-type recombinant bovine GLYAT and bovine liver GLYAT reached maximal activity at pH 7.5, and further increases in pH did not have a significant effect. However, the activity of the E226Q mutant increased significantly as pH increased from 6.0 to 9.6. At pH 9.6, the E226Q mutant had activity comparable to that of the wild-type recombinant bovine GLYAT and bovine liver GLYAT. When interpreted in light of the similarity of the kinetic parameters of the wild-type and mutant enzymes (suggesting that the mutant is

structurally intact), it was concluded that the $Glu^{226}$ residue is catalytically important in a pH dependent fashion, and is probably the catalytic base residue. This interpretation does not exclude the possibility that the $Glu^{226}$ residue acts in concert with another as yet unidentified residue to catalyse the deprotonation of glycine. What is clear is that the deprotonation of the amino group of glycine is important, as it is chemically impossible for the protonated amine to act as a nucleophile, and that $Glu^{226}$ seems to be involved in the process.

For acyl-transfer reactions, there is an alternative mechanism to the ternary-complex, direct transfer mechanism, commonly known as the ping-pong mechanism (Berndsen & Denu, 2005, Dyda et al., 2000). The literature supports our analyses that bovine GLYAT employs a general base catalysed, ternary-complex mechanism. First, the reaction kinetics support a ternary complex mechanism but not a ping-pong mechanism (Nandi et al., 1979, van der Westhuizen et al., 2000). Second, GLYAT is insensitive to thiol-modifying reagents, such as iodoacetamide and N-ethylmaleimide, which would inactivate the active site cysteine residue of a ping-pong enzyme (Nandi et al., 1979). Finally, GLYAT is homologous to the GNAT superfamily of acyltransferases, of which members studied to date all employ direct transfer mechanisms.

Based on our analyses, we propose the ternary-complex, base catalysed reaction mechanism, as depicted in Figure 5, for bovine GLYAT. Briefly, the glycine amino group is deprotonated by the $Glu^{226}$ residue in order to increase its nucleophilic character. The nucleophilic amine then attacks the thioester, forming a tetrahedral intermediate that collapses to form the coenzyme A and acylglycine products. We speculate that this mechanism should be conserved among the GLYAT enzymes of different species, based on the conservation of the E226 residue in the homologues of bovine GLYAT (Figure 1). If our interpretation is valid, this may provide the first insight into the active site of the human orthologue of bovine GLYAT, an enzyme of increasing clinical relevance. Conclusive evidence for the ternary-complex, base catalysed mechanism we propose for bovine GLYAT awaits the determination of a crystal- or NMR structure of a GLYAT enzyme, preferably with bound substrates or an inhibitor. Repeating the photoaffinity labelling of GLYAT, performed by Lau and co-workers, combined with mass spectrometric

identification of the labelled residues, would be another means of investigating the GLYAT active site (Lau et al., 1977).

**Acknowledgements**

We thank Professor Trevor Sewell for assistance with the molecular modelling and identification of the catalytic residue. We also thank Professor Francois van der Westhuizen for discussion of enzyme kinetic investigations, and Mrs Rencia van der Sluis for technical assistance. We thank Dr Frans O'Neill for critical reading of the manuscript.

**Author contributions**

*Participated in research design*: Badenhorst, Van Dijk, Jooste.

*Conducted experiments*: Badenhorst, Jooste.

*Performed data analysis*: Badenhorst.

*Wrote or contributed to the writing of the manuscript*: Badenhorst, Van Dijk.

References

Aiyar, A. & Leis, J. (1993). Modification of the Megaprimer Method of PCR Mutagenesis - Improved Amplification of the Final Product. *Biotechniques* **14**, 366.

Bartlett, K. & Gompertz, D. (1974). The specificity of glycine-N-acylase and acylglycine excretion in the organicacidaemias. *Biochem.Med.* **10**, 15-23.

Berndsen, C. E. & Denu, J. M. (2005). Assays for mechanistic investigations of protein/histone acetyltransferases. *Methods* **36**, 321-331.

Campbell, L., Wilson, H. K., Samuel, A. M. & Gompertz, D. (1988). Interactions of m-xylene and aspirin metabolism in man. *Br J Ind Med* **45**, 127-32.

Duffy, L. F., Kerzner, B., Seeff, L., Barr, S. B. & Soldin, S. J. (1995). Preliminary assessment of glycine conjugation of para-aminobenzoic acid as a quantitative test of liver function. *Clin.Biochem.* **28**, 527-530.

Dyda, F., Klein, D. C. & Hickman, A. B. (2000). GCN5-related N-acetyltransferases: a structural overview. *Annu.Rev.Biophys.Biomol.Struct.* **29**, 81-103.

Ellman, G. L. (1959). Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* **82**, 70-77.

Eswar, N., Eramian, D., Webb, B., Shen, M. Y. & Sali, A. (2008). Protein structure modeling with MODELLER. *Methods Mol.Biol.* **426**, 145-159.

Gregersen, N., Kolvraa, S. & Mortensen, P. B. (1986). Acyl-CoA: glycine N-acyltransferase: in vitro studies on the glycine conjugation of straight- and branched-chained acyl-CoA esters in human liver. *Biochem.Med.Metab Biol.* **35**, 210-218.

Jones, D. T. (1999). GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. *J.Mol.Biol.* **287**, 797-815.

Kelley, M. & Vessey, D. A. (1986). Interaction of 2,4-dichlorophenoxyacetate (2,4-D) and 2,4,5-trichlorophenoxyacetate (2,4,5-T) with the acyl-CoA: amino acid N-acyltransferase enzymes of bovine liver mitochondria. *Biochem.Pharmacol.* **35**, 289-295.

Kelley, M. & Vessey, D. A. (1993). Isolation and characterization of mitochondrial acyl-CoA: glycine N-acyltransferases from kidney. *J.Biochem.Toxicol.* **8**, 63-69.

Kelley, M. & Vessey, D. A. (1994). Characterization of the acyl-CoA:amino acid N-acyltransferases from primate liver mitochondria. *J.Biochem.Toxicol.* **9**, 153-158.

Kolvraa, S. & Gregersen, N. (1986). Acyl-CoA:glycine N-acyltransferase: organelle localization and affinity toward straight- and branched-chained acyl-CoA esters in rat liver. *Biochem.Med.Metab Biol.* **36**, 98-105.

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

Lau, E. P., Haley, B. E. & Barden, R. E. (1977). Photoaffinity labeling of acyl-coenzyme A:glycine N-acyltransferase with p-azidobenzoyl-coenzyme A. *Biochemistry* **16**, 2581-2585.

Lino Cardenas, C. L., Bourgine, J., Cauffiez, C., Allorge, D., Lo-Guidice, J. M., Broly, F. & Chevalier, D. (2010). Genetic polymorphisms of Glycine N-acyltransferase (GLYAT) in a French Caucasian population. *Xenobiotica* **40**, 853-61.

Liska, D. J. (1998). The detoxification enzyme systems. *Altern Med Rev* **3**, 187-98.

Mawal, Y. R. & Qureshi, I. A. (1994). Purification to homogeneity of mitochondrial acyl coa:glycine n-acyltransferase from human liver. *Biochem.Biophys.Res.Commun.* **205**, 1373-1379.

Mitchell, G. A., Gauthier, N., Lesimple, A., Wang, S. P., Mamer, O. & Qureshi, I. (2008). Hereditary and acquired diseases of acyl-coenzyme A metabolism. *Mol.Genet.Metab* **94**, 4-15.

Nandi, D. L., Lucas, S. V. & Webster, L. T., Jr. (1979). Benzoyl-coenzyme A:glycine N-acyltransferase and phenylacetyl-coenzyme A:glycine N-acyltransferase from bovine liver mitochondria. Purification and characterization. *J.Biol.Chem.* **254**, 7230-7237.

Ogier, d. B. & Saudubray, J. M. (2002). Branched-chain organic acidurias. *Semin.Neonatol.* **7**, 65-74.

Palmer, T. (2001). ENZYMES Biochemistry, Biotechnology, Clinical Chemistry. In *Horwood Series in Chemical Science*, 5 edn, pp. 402. Chichester: Horwood Publishing Limited.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J.Comput.Chem.* **25**, 1605-1612.

Sakuma, T. (1991). Alteration of urinary carnitine profile induced by benzoate administration. *Arch. Dis. Child.* **66**, 873-875.

Schachter, D. & Taggart, J. V. (1954). Glycine N-acylase: purification and properties. *J.Biol.Chem.* **208**, 263-275.

Scheibner, K. A., De Angelis, J., Burley, S. K. & Cole, P. A. (2002). Investigation of the roles of catalytic residues in serotonin N-acetyltransferase. *J.Biol.Chem.* **277**, 18118-18126.

Tanaka, K., Budd, M. A., Efron, M. L. & Isselbacher, K. J. (1966). Isovaleric acidemia: a new genetic defect of leucine metabolism. *Proc.Natl Acad.Sci.U.S.A* **56**, 236-242.

Tanaka, K. & Isselbacher, K. J. (1967). The isolation and identification of N-isovalerylglycine from urine of patients with isovaleric acidemia. *J.Biol.Chem.* **242**, 2966-2972.

Temellini, A., Mogavero, S., Giulianotti, P. C., Pietrabissa, A., Mosca, F. & Pacifici, G. M. (1993). Conjugation of benzoic acid with glycine in human liver and kidney: a study on the interindividual variability. *Xenobiotica* **23**, 1427-1433.

van der Westhuizen, F. H., Pretorius, P. J. & Erasmus, E. (2000). The utilization of alanine, glutamic acid, and serine as amino acid substrates for glycine N-acyltransferase. *J.Biochem.Mol.Toxicol.* **14**, 102-109.

Vetting, M. W., Carvalho Lp, S. d., Yu, M., Hegde, S. S., Magnet, S., Roderick, S. L. & Blanchard, J. S. (2005). Structure and functions of the GNAT superfamily of acetyltransferases. *Arch.Biochem.Biophys.* **433**, 212-226.

Wallig, M. A. (2004). Glucuronidation and susceptibility to chemical carcinogenesis. *Toxicological sciences*, 1-2.

Waluk, D. P., Schultz, N. & Hunt, M. C. (2009). Human glycine N-acyltransferase-like 1 is involved in the production of glycine conjugated fatty acids. *Chem. Phys. Lipids* **160S**, S22-S24.

Waluk, D. P., Schultz, N. & Hunt, M. C. (2010). Identification of glycine N-acyltransferase-like 2 (GLYATL2) as a transferase that produces N-acyl glycines in humans. *Faseb J* **24**, 2795-803.

Wang, Q., Canutescu, A. A. & Dunbrack, R. L., Jr. (2008). SCWRL and MolIDE: computer programs for side-chain conformation prediction and homology modeling. *Nat.Protoc.* **3**, 1832-1847.

Yamamoto, A., Nonen, S., Fukuda, T., Yamazaki, H. & Azuma, J. (2009). Genetic polymorphisms of glycine N-acyltransferase in Japanese individuals. *Drug Metab Pharmacokinet* **24**, 114-7.

**Footnotes**

The study was funded by contributions of BioPAD [BPP007] and the National Research Foundation [FA2005031700015].

A preliminary overview of this study was published as an abstract for the SSIEM 2010, in the Journal of Inherited Metabolic Disease (Badenhorst, CPS., Jooste, M., van Dijk, AA, 2010. Journal of Inherited Metabolic Disease, 33, S52).

Person to contact for reprint requests: Professor Albie van Dijk, Biochemistry Department, North-West University, Private Bag X6001, Potchefstroom, 2520, South Africa. E-mail: Albie.VanDijk@nwu.ac.za.

**Figure legends**

***Figure 1.*** Prediction of a putative catalytic glutamate residue for bovine GLYAT. A) Part of a multiple alignment of sequences with significant similarity to bovine GLYAT, demonstrating that the $Glu^{226}$ residue is conserved. Homologues are results of a BLAST search and GenBank accession numbers are shown to the left of the alignment. B) Part of a structural superposition of the structures of SNAT, SSAT and Esa1 with the bovine GLYAT model. The catalytic residue side chains of the GNAT enzymes are shown. The bovine GLYAT $Glu^{226}$ residue is indicated on both the multiple alignment and the molecular model. The small molecule to the right is acetyl-coenzyme A, a substrate bound to SNAT. This image was generated using UCSF Chimera.

***Figure 2.*** Bacterial expression and partial purification of the recombinant bovine GLYAT enzymes. A) SDS-PAGE analyses of the recombinant bovine GLYAT expression. Lanes: 1) PageRuler protein size marker; 2) Wild-type, total fraction; 3) E226Q, total fraction; 4) Wild-type, soluble fraction; 5) E226Q, soluble fraction. The arrow indicates the position of recombinant bovine GLYAT on the gel. The molecular mass of the recombinant bovine GLYAT enzymes is approximately 36.7 kDa. B) SDS-PAGE analyses of recombinant bovine GLYAT partially purified by nickel-affinity chromatography. Lanes: 1) PageRuler protein size marker; 2) Soluble, partially purified wild-type recombinant GLYAT; 3) Soluble, partially purified E226Q mutant recombinant GLYAT. Recombinant bovine GLYAT is indicated by the arrow, as some unidentified co-purifying proteins are also visible on the gel.

***Figure 3.*** Lineweaver-Burk plots used to visualise the kinetic parameters of the GLYAT enzymes. Benzoyl-coenzyme A and glycine were used as substrates. Benzoyl-coenzyme A concentrations were 10, 15, 25, 40 and 50 μM. The data points indicate average values ± standard deviation of triplicate assays. Plots were generated using SigmaPlot 11.0.

***Figure 4.*** The pH dependence of the bovine liver, wild-type and E226Q recombinant GLYAT enzymes. Representative initial velocity data (nM product formed per min) is plotted against pH on a logarithmic

scale. For each reaction, 30 μg of bovine liver GLYAT, or 3 μg of either recombinant enzyme, was used. Error bars indicate the mean ± standard deviation of triplicate assays. Data was plotted using GraphPad Prism 4.02.

***Figure 5.*** A schematic representation of the general base catalysed, ternary complex mechanism proposed for bovine GLYAT. A) For nucleophilic attack to take place, the glycine amino group must be deprotonated by $Glu^{226}$; B) A tetrahedral intermediate is formed, following nucleophilic attack by the amino group of glycine on the thioester carbonyl group. C) Finally, the tetrahedral intermediate collapses, forming the peptide product and coenzyme A. ChemDraw 10.0 was used to produce this schematic.

**Table 1** **$K_M$ values for wild-type and E226Q mutant recombinant bovine GLYAT and GLYAT from bovine liver (n=3)**

| GLYAT enzyme | $K_M$ benzoyl-coenzyme A (µM ± SD) | $K_M$ glycine (mM ± SD) |
|---|---|---|
| | | |
| Recombinant bovine GLYAT | 16 ± 1 | 2 ± 0.3 |
| E226Q recombinant GLYAT | 18 ± 4 | 7 ± 4 |
| GLYAT from bovine liver | 16 ± 3 | 1.6 ± 0.5 |

**Table 2 Apparent $K_M$ values for acyl-coenzyme A substrates of bovine liver and recombinant GLYAT enzymes (n=3)**

| | **$K_M$ for acyl-coenzyme A (µM ± SD)** | |
|---|---|---|
| **Acyl-coenzyme A** | **Bovine liver GLYAT** | **Recombinant GLYAT** |
| Benzoyl-coenzyme A | 18 ± 5 | 19 ± 2 |
| Octanoyl-coenzyme A | 70 ± 10 | 66 ± 8 |
| 3-methylcrotonyl-coenzyme A | 140 ± 24 | 123 ± 19 |
| Propionyl-coenzyme A | 184 ± 30 | 143 ± 23 |
| Isovaleryl-coenzyme A | 195 ± 32 | 127 ± 20 |

# Figure 1

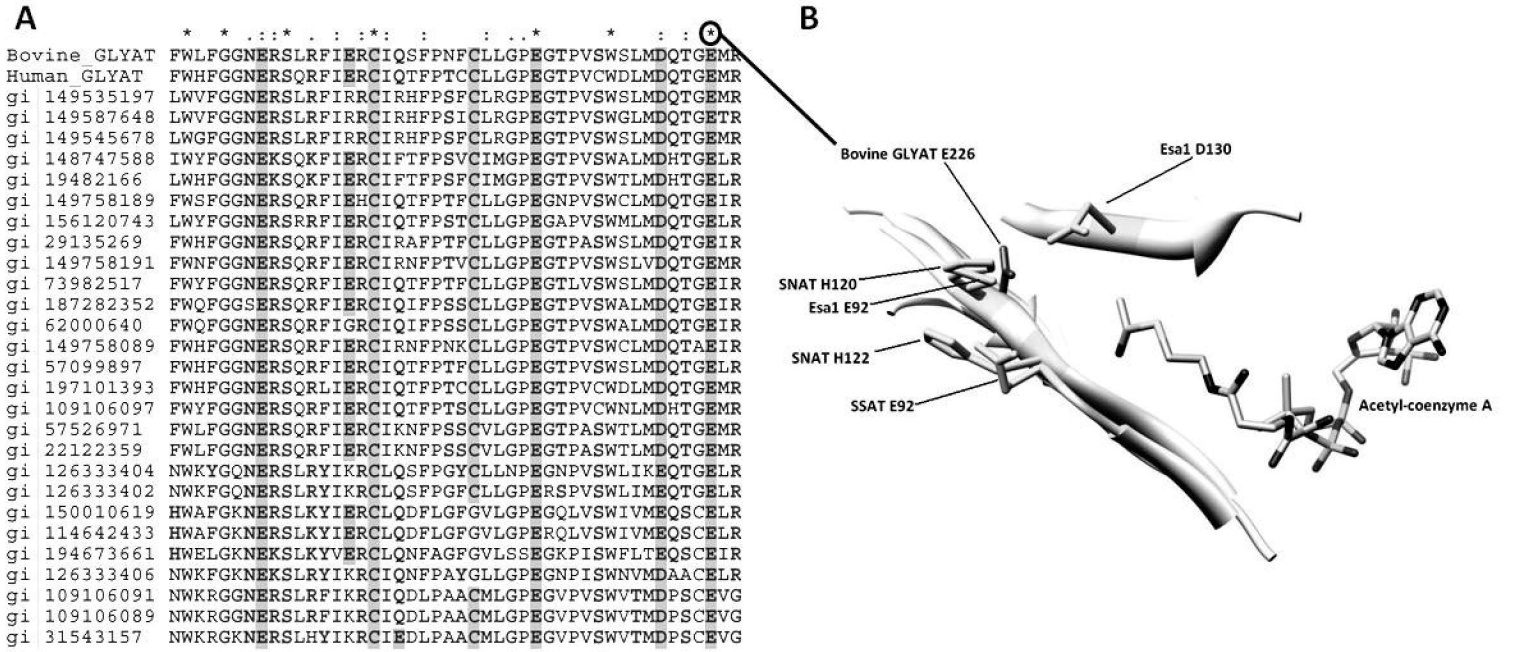
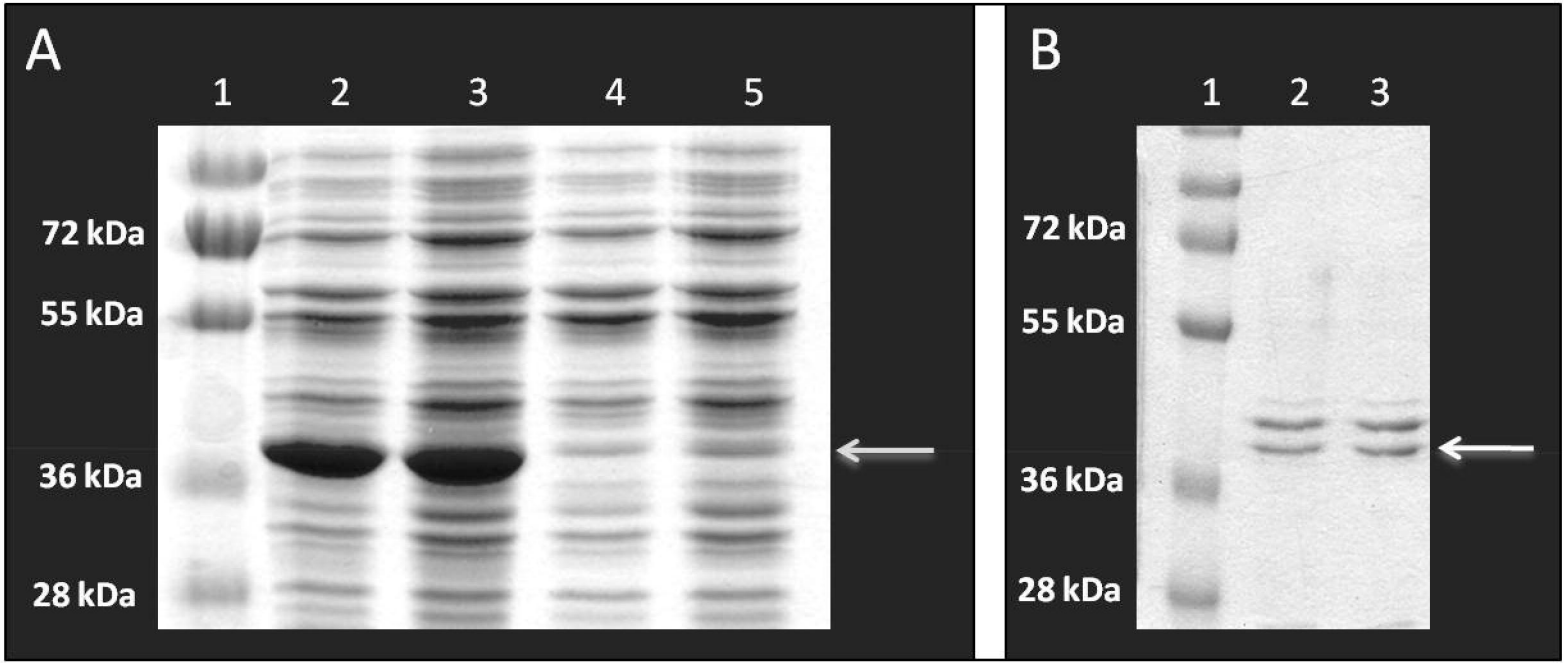
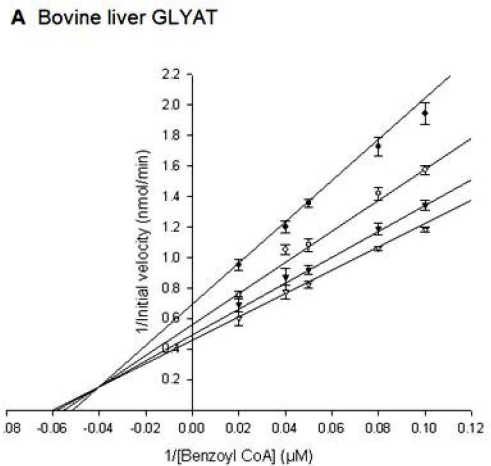
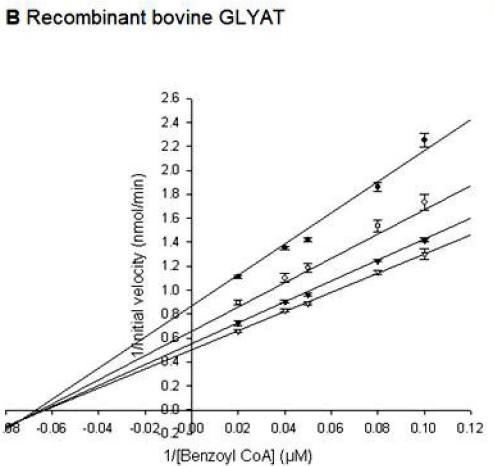

A
B
Bovine GLYAT E226
Esa1 D130
SNAT H120
Esa1 E92
SNAT H122
SSAT E92
Acetyl-coenzyme A

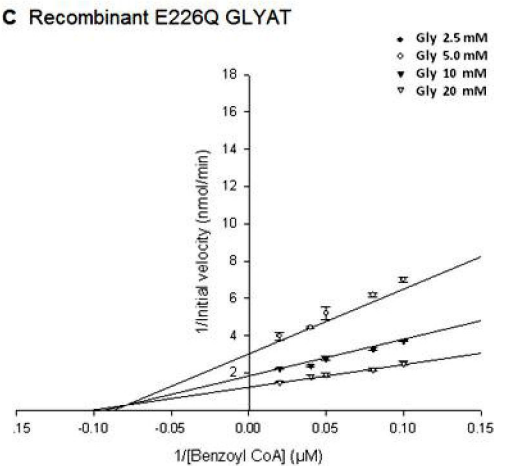
**Figure 2**


A
1
2
3
4
5
72 kDa
55 kDa
36 kDa
28 kDa
B
1
2
3
72 kDa
55 kDa
36 kDa
28 kDa

# Figure 3


A Bovine liver GLYAT
1/Initial velocity (nmol/min)
1/[Benzoyl CoA] (μM)


B Recombinant bovine GLYAT
1/Initial velocity (nmol/min)
1/[Benzoyl CoA] (μM)


C Recombinant E226Q GLYAT
Gly 2.5 mM
Gly 5.0 mM
Gly 10 mM
Gly 20 mM
1/Initial velocity (nmol/min)
1/[Benzoyl CoA] (μM)

# Figure 4

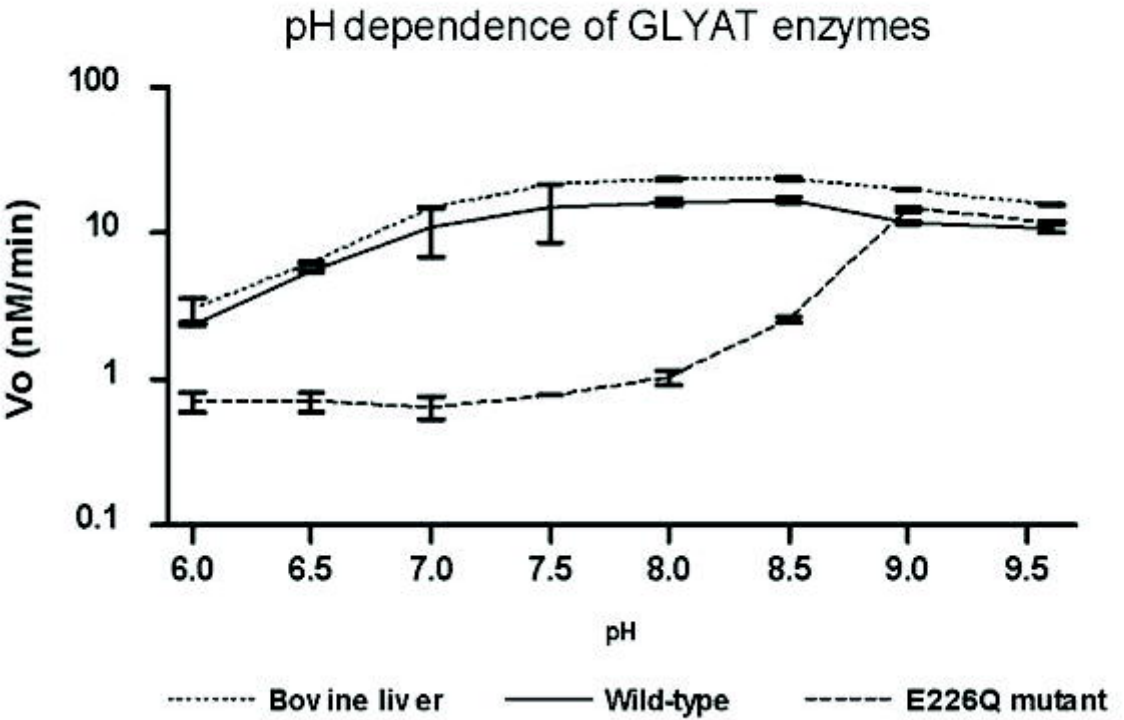

pH dependence of GLYAT enzymes
100
10
1
0.1
Vo (nM/min)
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
pH
Bovine liver
Wild-type
E226Q mutant

**Figure 5**