

Investigation of the bi-substrate kinetics of a recombinant human glycine N-acyltransferase with a known non-synonymous single nucleotide polymorphism (N156S)

Vida Ungerer
22155635

Dissertation submitted in partial fulfilment of the requirements for the degree *Magister Scientiae* in *Biochemistry* at the Potchefstroom Campus of the North-West University

Supervisor: Dr R van der Sluis

Co-supervisor: Mr E Erasmus

November 2015

“There is a single light of science, and to brighten it anywhere is to brighten it everywhere”

Isaac Asimov

Acknowledgements

I am sincerely grateful for the opportunity to have been able to complete this M.Sc and I am thankful of those who contributed to the completion of this great feat.

I would like to thank my supervisor, Dr. Rencia van der Sluis, for providing me with the freedom and confidence to direct my study as I saw fit and for her expert guidance. I greatly benefited from her keen scientific insight and her knack for problem solving.

I sincerely thank my co-supervisor, Mr. Lardus Erasmus. He taught me how to ask questions, how to read between the lines and to be consistently kind. I truly appreciated his indispensable wisdom and guidance.

To Prof. Albie van Dijk, thank you for awakening in me a thirst for knowledge and for encouraging me to look ahead even in the toughest of times.

I would also like to mention Ms. Carla Nortje and Mr. Peet Jansen van Rensburg, who were extremely reliable sources of practical scientific knowledge regarding the HPLC analyses; and Prof. Francois van der Westhuizen for guidance regarding the interpretation of the kinetics results.

Many thanks go to my dear friends, family and colleagues for their encouragement and friendship over the years.

I express my utmost gratitude to my parents (Martin and Ina) and sister (Marna) for their love and support. Thank you for enabling and encouraging me to pursue further studies; and for providing me with the opportunity to pursue a career in science.

Special thanks to my dearest Abel Bronkhorst for his unconditional love, support and willingness to endure with me throughout the years.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged.

Abstract

Biotransformation is fast becoming a buzzword of our time. This is due to the ever increasing exposure of humans to xenobiotic substances. The detoxification of these xenobiotics occur via various detoxifying mechanisms in which these substances are converted to more hydrophilic forms, to ease their excretion from the body, and thus also to avoid the possibly toxic accumulation of these compounds in the body. One such detoxifying mechanism is the glycine conjugation pathway. The focus of this study was on one of the enzymes involved in this pathway responsible for the detoxification of metabolites by conjugation to glycine, the glycine N-acyltransferase (GLYAT) enzyme. Interindividual variation has been observed in the glycine conjugation pathway, which may involve various contributing factors. One such factor is the genetic variation in the *GLYAT* gene. Single nucleotide polymorphisms (SNPs) occurring in the exon regions of the *GLYAT* gene have been associated with altered enzyme activities. This study focused on the N156S human GLYAT. In previous studies, this variant was found to have the highest enzyme activity as well as the highest allele frequency, and is now regarded as the wild-type variant of human GLYAT. Thus, making this variant exceptionally relevant for investigations. In this study a bacterially codon optimised N156S human GLYAT construct was generated by means of site-directed mutagenesis and the protein expressed with an N-terminal histidine tag for affinity purification. Throughout this process, some optimisation experiments were also conducted. These included the optimisation of the protein expression, extraction and storage conditions. The enzyme kinetic studies were then conducted with the purified recombinant N156S human GLYAT in the presence of two varying substrates (benzoyl-coenzyme A and glycine) to be able to characterise the bi-substrate kinetic parameters of this human GLYAT variant. The K_m value were $49 \pm 13 \mu\text{M}$ for benzoyl-coenzyme A and $20 \pm 4 \text{ mM}$ for glycine. These findings correlated with the values found in the literature which reported K_m values for benzoyl-coenzyme A ranging from 6 to 67 μM ; and from 6.4 to 26.6 mM for glycine. The kinetic model we proposed, the random order sequential mechanism, which is a ternary complex mechanism, also agreed to what had previously been described for GLYAT enzymes. The N156S human GLYAT enzyme activity was further characterised by quantifying the amount of product, hippuric acid (HA) formed in the presence of varying

substrate concentrations. This was done using an HPLC-MS/MS method with a stable isotope as internal standard. The findings suggested that increasing amounts of HA was formed when the glycine substrate concentrations were increased, and benzoyl-coenzyme A concentrations were kept constant. This effect, however, was only noticed up to a certain glycine concentration, after which the HA formation slowed down and reached a plateau. These findings suggested that the GLYAT enzyme had evolved with a limited rate of detoxification, possibly to avoid glycine depletion. Glycine is used for the production of creatine, bile salts, porphyrins, collagen, elastin, glutathione, as well as other proteins, and the depletion of glycine stores may have serious implications.

Further characterisation of this and other variants of GLYAT is however necessary, in order to fully understand the catalytic mechanisms of this enzyme. Once enough knowledge of these variants and their enzymatic properties are known, it may then become possible to attempt to design a GLYAT with altered substrate specificity, which may be useful for the treatment of organic acidemias.

Keywords: Glycine N-acyltransferase; biotransformation; glycine conjugation; single nucleotide polymorphisms; protein expression; enzyme kinetics; HPLC-MS/MS.

Uittreksel

Biotransformasie is vinnig besig om 'n modewoord van ons tyd te word. Dit is te danke aan die toenemende blootstelling van mense aan xenobiotiese stowwe. Die ontgiftiging van hierdie xenobiotiese middels word deur verskeie detoksifiserende meganismes uitgevoer. Hierdie stowwe word omgeskakel na meer hidrofiliese vorms, om hul uitscheidingsvermoë te verbeter, en sodanig ook die moontlikheid van 'n toksiese opeenhoping van hierdie verbindings in die liggaam te voorkom. Een so 'n detoksifiserende meganisme is die glisien-konjugeringsweg. Die fokus van dié studie was op een van die ensieme betrokke in hierdie weg, die glisien N-asieltransferase (GLIAT) ensiem, wat hoofsaaklik verantwoordelik is vir die ontgiftiging van metaboliete deur konjugasie met glisien. Interindividuele variasie is waargeneem in die glisien vervoegingspad, wat verskeie bydraende faktore kan betrek. Een so 'n faktor is die genetiese variasie van die *GLIAT* geen. Enkel nukleotied polimorfismes (ENPs) wat in die eksons van die *GLIAT* geen voorkom, hou verband met veranderde ensiemaktiwiteite. Hierdie studie fokus op die menslike N156S GLIAT. In vorige studies, was daar bepaal dat hierdie variant die hoogste ensiemaktiwiteit asook die hoogste alleelfrekwensie het, en word deesdae beskou as die wilde-tipe variant van menslike GLIAT. Dus is hierdie variant besonder relevant vir ondersoeke. In hierdie studie is 'n bakteriële kodongeoptimiseerde N156S menslike GLIAT konstrueer gebruik wat deur middel van 'n polimerase kettingreaksie (PKR) geproduseer is. Die proteïen was uitgedruk met 'n N-terminale histidien herkenningspunt wat gebruik was vir affiniteitssuiwering. Gedurende hierdie proses was 'n paar optimiseringseksperimente ook uitgevoer. Dit sluit in die optimalisering van die proteïenuitdrukking, ekstraksie en ook stortingstoestand. Die ensiemkinetika studies was uitgevoer met die gesuiwerde rekombinante N156S menslike GLIAT in die teenwoordigheid van twee verskillende substrate nl. bensoïel-koënsiem A en glisien; om sodoende die twee-substraat kinetiese parameters van hierdie menslike GLIAT variant te bepaal. Die K_m -waardes vir bensoïel-koënsiem A was $49 \pm 13 \mu\text{M}$ en $20 \pm 4 \text{ mM}$ vir glisien. Hierdie bevindinge stem ooreen met die waardes wat in die literatuur beskryf word (bensoïel-koënsiem A waardes het gewissel tussen 6 en $67 \mu\text{M}$; en tussen 6,4 en 26,6 mM vir glisien). Die kinetiese model wat ons voorgestel het, die ewekansige opeenvolgende meganisme, wat 'n drieledige kompleks meganisme is, het ook ooreengekom met wat

voorheen beskryf is vir GLIAT ensieme. Die N156S menslike GLIAT ensiemaktiwiteit was verder ondersoek deur die hoeveelheid produk, hippuursuur (HS), wat gevorm word in die teenwoordigheid van verskillende substraatkonsentrasies te kwantifiseer. Dit was gedoen met behulp van 'n HPLC-MS/MS metode, in samewerking met 'n stabiele isotoop wat as 'n interne standaard opgetree het. Die bevindinge het getoon dat daar groter hoeveelhede HS gevorm word soos wat die glisien substraatkonsentrasies verhoog word, terwyl die bensoëlskoënsiem A konsentrasies konstant gehou word. Hierdie effek was egter slegs opgemerk tot op 'n sekere glisienskonsentrasie, waarna die HS vorming vertraag en 'n plato bereik word. Hierdie bevindinge het voorgestel dat die GLIAT ensiem dalk ontwikkel het met 'n beperkte mate van ontgiftiging, moontlik om glisien uitputting te voorkom. Glisien word gebruik vir die produksie van kreatien, galsoute, porfiriene, kollageen, elastien, glutatioon, sowel as ander proteïene, en die uitputting daarvan kan tot ernstige implikasies lei.

Verdere karakterisering van hierdie en ander variante van GLIAT is egter noodsaaklik, om ten volle te kan verstaan hoe die katalitiese meganismes van hierdie ensiem werk. Sodra daar genoeg inligting aangaande hierdie variante en hul ensiematiese eienskappe bekend is, sal dit moontlik wees om 'n GLIAT ensiem met veranderde substraat spesifisiteit te vervaardig, wat nuttig sal wees vir verwickelinge ten opsigte van die behandeling van organiese suur asidemies.

Slutelwoorde: Glisien N-asieltransferase; biotransformasie; glisien konjugering; enkele nukleotied polimorfismes; proteïen uitdrukking; ensiemkinetika; HPLC-MS/MS.

Table of contents

ACKNOWLEDGEMENTS.....	ii
ABSTRACT.....	iii
UITTREKSEL.....	v
LIST OF TABLES	xii
LIST OF FIGURES	xv
LIST OF EQUATIONS.....	xviii
LIST OF ABBREVIATIONS.....	xix
CHAPTER 1: INTRODUCTION	1
1.1 Introduction.....	1
1.2 Dissertation outline.....	3
1.2.1 Chapter 2: Literature review.....	3
1.2.2 Chapter 3: The construction and expression of the N156S recombinant human GLYAT variant	3
1.2.3 Chapter 4: Bi-substrate kinetic analysis of the recombinant N156S hGLYAT variant 4	
1.2.4 Chapter 5: HPLC-ESI-MS/MS quantification of hippuric acid using a stable isotope 4	
1.2.5 Chapter 6: Final conclusions and future prospects	4
1.2.6 References	4
1.2.7 Appendices.....	5
CHAPTER 2: LITERATURE REVIEW	6
2. 1 Introduction.....	6

2.2	Drug metabolism	6
2.3	Biotransformation	7
2.3.1	Biotransforming enzymes	8
2.4	Phase I biotransformation.....	9
2.5	Phase II biotransformation.....	10
2.6	Phase 0 biotransformation.....	12
2.7	Phase III biotransformation.....	13
2.8	Amino Acid Conjugation.....	13
2.9	Glycine conjugation.....	14
2.9.1	An introduction to glycine conjugation	14
2.9.2	The glycine conjugation pathway	16
2.9.3	Interindividual variation in glycine conjugation	17
2.10	Molecular characteristics of GLYAT	22
2.11	The enzymatic properties of GLYAT	22
2.11.1	Enzymatic reaction.....	22
2.11.2	Substrate specificity	23
2.12	Site-directed mutagenesis.....	28
2.13	General principles of recombinant protein expression	29
2.13.1	Co-expression of chaperone proteins.....	31
2.14	Measurement of HA formed from recombinant GLYAT activity	32
2.15	Stable isotope quantification	33
2.16	Problem statement	34

2.17	Research aims and objectives	35
------	------------------------------------	----

CHAPTER 3: THE CONSTRUCTION AND EXPRESSION OF THE N156S RECOMBINANT

HUMAN GLYAT VARIANT 38

3.1	Introduction.....	38
3.2	Materials and methods	41
3.2.1	Preparation of stock solutions	41
3.2.2	Competent cells	41
3.2.3	WThGLYAT	41
3.2.4	coWThGLYAT.....	42
3.2.5	Transformation of chemically competent <i>E. coli</i> cells.....	42
3.2.6	Plasmid extraction (Midi-preparation)	43
3.2.7	Agarose gels	45
3.2.8	Sodium dodecyl sulfate polyacrylamide gels.....	45
3.2.9	Restriction endonuclease digests of plasmids.....	46
3.2.10	Gel extraction.....	48
3.2.11	DNA quantification and purity determination.....	49
3.2.12	Ligation reaction	50
3.2.13	Transformation of chemically competent <i>E. coli</i> cells.....	50
3.2.14	Provisional verification of colonies of transformed bacteria	51
3.2.15	DNA sequence determination and data analysis.....	51
3.2.16	Generation of the N156S hGLYAT variant by Site-directed mutagenesis	52
3.2.17	Transformation of chemically competent <i>E. coli</i> cells with modified amplicons	

3.2.18	Colony selection.....	57
3.2.19	DNA sequence determination and data analysis.....	58
3.2.20	Co-transformation of chemically competent Origami cells.....	58
3.2.21	Expression of coN156ShGLYAT and N156ShGLYAT proteins.....	59
3.2.22	Extraction of expressed coN156ShGLYAT and N156ShGLYAT proteins	60
3.2.23	Histidine-tagged protein purification	61
3.2.24	hGLYAT enzyme activity assays.....	63
3.2.25	Long term storage of purified enzyme	65
3.2.26	Determination of protein concentration	65
3.3	Results and discussion.....	67
3.3.1	Extraction of pET-32a(+) and pMA-T plasmids	67
3.2.1	Restriction endonuclease digests of plasmids.....	68
3.3.2	Verification of gel extraction	70
3.3.3	Colony screening after transformation.....	71
3.3.4	DNA sequence determination and data analysis.....	71
3.3.5	Generation of the coN156ShGLYAT variant.....	74
3.3.6	DNA sequence determination of coN156ShGLYAT variant	75
3.3.7	Optimisation of protein extraction	78
3.3.8	Recombinant N156S GLYAT enzyme activity assays.....	83
3.3.9	Long term storage of purified recombinant N156ShGLYAT variant	85
3.3.10	Determination of protein concentration	86
3.4	Summary	89

CHAPTER 4: BI-SUBSTRATE KINETIC ANALYSIS OF THE RECOMBINANT N156S HGLYAT

VARIANT	92
4.1 Introduction.....	92
4.2 Materials and methods	93
4.2.1 Reagents, standards and solutions	93
4.2.2 Preparation of standard stock solutions.....	93
4.2.3 Benzoyl-coenzyme A standard curve.....	93
4.2.4 Dilution of the purified N156S hGLYAT protein.....	95
4.2.5 Kinetic assays of the N156S hGLYAT variant.....	97
4.2.6 Data analysis for calculation of kinetic parameters	99
4.3 Results and discussion.....	101
4.3.1 Benzoyl-coenzyme A standard curve.....	101
4.3.2 Enzyme assay with dilution range of the purified N156S hGLYAT protein.....	102
4.3.3 Calculation of bi-substrate kinetic parameters of the recombinant N156S hGLYAT variant and selection of an enzyme kinetic model	104
4.3.4 Kinetic model of human GLYAT	116
4.4 Summary	120

CHAPTER 5: HPLC-ESI-MS/MS QUANTIFICATION OF HIPPURIC ACID USING A STABLE

ISOTOPE	122
5.1 Introduction.....	122
5.2 Materials and methods	123
5.2.1 Reagents, standards and solutions	123
5.2.2 Preparation of standard stock solutions.....	123

5.2.3	Preparation of mobile phases	123
5.2.4	Organic solvent extraction	123
5.2.5	Method development and optimisation	125
5.2.6	HPLC-ESI-MS/MS specifications	127
5.2.7	Quantification with stable isotope	128
5.2.8	Statistical data analysis	130
5.3	Results and discussion.....	130
5.3.1	Introduction	130
5.3.2	Linearity	130
5.3.3	Organic solvent extraction.....	132
5.3.4	Quantification with HPLC-ESI-MS/MS method.....	132
5.4	Summary	138
CHAPTER 6: FINAL CONCLUSIONS AND FUTURE PROSPECTS		140
6.1.	Introduction.....	140
6.2.	Conclusions.....	141
6.2.1	Bi-substrate kinetics of a recombinant N156ShGLYAT	141
6.2.2	HPLC-ESI-MS/MS quantification of hippuric acid using a stable isotope	143
6.3	Future recommendations	143
REFERENCES		147
APPENDIX A		162
APPENDIX B.....		166

List of Tables

Table 2.1: A summary of the K_m values (benzoyl-coenzyme A) obtained for recombinant human GLYAT variants.

Table 2.2: The K_m values for acyl-coenzyme A substrates and glycine for human GLYAT.

Table 2.3: K_m values for glycine and alanine; and benzoyl-coenzyme A for human GLYAT.

Table 2.4: Summary of some examples of conditions in which the alternative amino acid substrates would be utilised by GLYAT.

Table 3.1: Abbreviations created to refer to the different forms of human *GLYAT* genes.

Table 3.2: The restriction enzymes and their buffers that were used for this study.

Table 3.3: Oligonucleotide primers used in this study for Site-directed mutagenesis to generate the codon optimised N156S hGLYAT variant.

Table 3.4: Oligonucleotide primers used in a previous study, which were used for the positive control reaction to generate the coN156ShGLYAT variant via site-directed mutagenesis.

Table 3.5: Cycling conditions for the site-directed mutagenesis reactions.

Table 3.6: Cycling conditions for the control mutagenesis reactions.

Table 3.7: A summary of the concentrations of the N156S hGLYAT protein as determined by the Qubit Fluorometer method.

Table 3.8: A summary of the concentrations of the N156S hGLYAT as determined by the BCA assay.

Table 4.1: Plate layout for the enzyme reactions for the benzoyl-coenzyme A standard curve.

Table 4.2: Plate layout for the enzyme reactions prepared for the enzyme dilution experiment with high and low benzoyl-coenzyme A concentrations.

Table 4.3: The plate layout of the enzyme reactions used for the bi-substrate kinetics of the N156S hGLYAT variant, indicating the final substrate concentrations present in each reaction mixture.

Table 4.4: Raw data of absorbance values obtained from the plate reader over 20 minutes of the enzyme reaction with 60 μ M benzoyl-coenzyme A.

Table 4.5: Initial velocities (V_0) calculated from the change in A_{412} (nm) values over time (6 minutes).

Table 4.6: Processed data set of enzyme assay results used for calculation of kinetic parameters using SigmaPlot with the Enzyme Kinetics module.

Table 4.7: Kinetic parameters obtained by the random ordered mechanism and ordered mechanism models for the N156S hGLYAT enzyme using benzoyl-coenzyme A and glycine as substrates.

Table 4.7: Kinetic parameters of benzoyl-coenzyme A and glycine for hGLYAT enzyme as reported in the literature.

Table 5.1: Optimised MRM parameters for the quantification of HA and HA-d5 in positive ionisation mode.

Table 5.2: Gradient of the mobile phase composition over the time course of the method.

Table 5.3: The plate layout of the enzyme reactions indicating the substrate concentrations added to each reaction mixture.

Table 5.4: The linear ranges and the correlation coefficient values (r^2) for both calibration curves.

Table 5.5: Final HA concentrations, average HA concentrations, and %CV values.

List of Figures

Figure 2.1: A schematic of the liver biotransformation processes, with their respective reaction mechanisms and supportive nutrients.

Figure 2.2: Metabolic pathway of glycine conjugation.

Figure 2.3: Results presented by van der Sluis *et al.* (2013) showing the relative enzyme activities of selected variants of recombinant human GLYAT.

Figure 2.4: Substrate specificity of the GLYAT reaction.

Figure 2.5: Schematic showing the components of a triple quadrupole mass spectrometer system operated in MRM mode.

Figure 2.6: The structural formulae of A) hippuric acid, and B) hippuric acid-d5.

Figure 3.1: Schematic describing the overall layout of the methods described in this part of the study.

Figure 3.2: Vector map of the pMA-T vector containing the synthetic codon optimised *hGLYAT* gene.

Figure 3.3: Schematic representation of site-directed mutagenesis.

Figure 3.4: Vector map of the pGro7 chaperone plasmid.

Figure 3.5: Vector map of the pET-32a(+) plasmid.

Figure 3.6: Agarose gel (1%) of extracted pET-32a(+)/WThGLYAT and pMA-T/coWThGLYAT.

Figure 3.7: Agarose gel (1%) of the *Bam*HI and *Hind*III restriction enzyme digestions.

Figure 3.8: Agarose gel (1%) of the gel extracted pET-32a(+) vector and the coWThGLYAT insert.

Figure 3.9: Agarose gel (1%) of eight samples of successfully subcloned pET-32a(+)/coWThGLYAT.

Figure 3.10: The aligned sequences of the coWThGLYAT insert, the hGLYAT insert, and the sequences generated by the forward and reverse primers.

Figure 3.11: Agarose gel (1%) of the PCR amplification products of pET-32a(+)/coWThGLYAT using various annealing temperatures.

Figure 3.12: The aligned sequences of the coWThGLYAT insert, and the sequences generated by the forward and reverse primers.

Figure 3.13: The exploded view of the aligned sequences of the coWThGLYAT insert, and the sequences generated by the forward and reverse primers.

Figure 3.14A: SDS-PAGE (12.5%) analyses of the N156S hGLYAT extracted by means of the BugBuster Protein Extraction Reagent.

Figure 3.14B: SDS-PAGE (12.5%) analyses of the coN156ShGLYAT extracted by means of the BugBuster Protein Extraction Reagent.

Figure 3.15A: SDS-PAGE (12.5%) analyses of the N156S hGLYAT extracted by means of the Lysis buffer method.

Figure 3.15B: SDS-PAGE (12.5%) analyses of the coN156S hGLYAT extracted by means of the Lysis buffer method.

Figure 3.16: The enzyme assay plots of the purified mutant recombinant enzymes extracted with the BugBuster Reagent.

Figure 3.17: The enzyme assay plots of the purified mutant recombinant enzymes extracted with the Lysis buffer method.

Figure 3.18: The enzyme activity plots of the purified N156S recombinant enzymes from each of the time points at which long term storage was investigated.

Figure 3.19: Standard curve of BSA, used for concentration determination with BCA assay.

Figure 4.1: Graphic representation of the activity plots of enzymes with different dilution factors.

Figure 4.2: A plot illustrating the absorbance (412 nm) values over time.

Figure 4.3: Standard curve of absorbance (412 nm) versus benzoyl-coenzyme A concentrations.

Figure 4.4: The enzyme activity plots of the purified N156ShGLYAT recombinant enzymes with different dilution factors.

Figure 4.5: The enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration (60 μ M).

Figure 4.6: The enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration (20 μ M).

Figure 4.7: The enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration (40 μ M).

Figure 4.8: The enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration (60 μ M).

Figure 4.9: The enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration (80 μ M).

Figure 4.10: Random order sequential mechanism Lineweaver-Burk plots to determine the K_m values of benzoyl-coenzyme A and glycine for the purified N156S hGLYAT recombinant enzymes.

Figure 4.11: Ordered sequential mechanism Lineweaver-Burk plots to determine the K_m values of benzoyl-coenzyme A and glycine for the purified N156S hGLYAT recombinant enzymes.

Figure 4.12: Ping-pong Bi-Bi mechanism Lineweaver-Burk plots to determine the K_m values of benzoyl-coenzyme A and glycine for the purified N156S hGLYAT recombinant enzymes.

Figure 4.13: Lineweaver-Burk plots of the kinetic parameters of the bovine liver GLYAT (A) and recombinant bovine GLYAT (B) enzymes, respectively.

Figure 5.1: Calibration curve of HA with HA-d5.

Figure 5.2: A column chart of the final HA concentrations obtained at varying glycine levels.

Figure 5.3: A column chart of the final HA concentrations obtained at varying benzoyl-CoA levels.

List of Equations

Equation 3.1: The equation used to calculate the ratios of insert to vector for ligation reactions.

Equation 3.2: The equation used to calculate the ratios of insert to vector for ligation reactions, with the values substituted as is pertained to this study.

Equation 4.1: The equation used to calculate the initial velocity of enzyme reactions.

Equation 4.2: Formula for conversion from initial velocity (absorbance/min) to absorbance/min/ μg .

Equation 4.3: Formula for conversion from absorbance/min/ μg to $\mu\text{mol/min}/\mu\text{g}$.

Equation 4.4: Formula for conversion from $\mu\text{mol/min}/\mu\text{g}$ to nmol/min/mg.

Equation 5.1: Linear formula of the HA with HA-d5 calibration curve.

Equation 5.2: The linear equation with the substituted values for final HA concentration determination.

Equation 5.3: Equation for calculating coefficient of variation (% CV).

List of Abbreviations

Symbols

°C degrees Celsius

% percentage

Abbreviations

µg microgram

µl microliter

µM micromolar

A adenine

A412 absorbance at 412 nm

Abs absorbance

ATP adenosine triphosphate

BCA bicinchoninic acid

BSA bovine serum albumin

C carbon

CAF Central Analytical Facility

CID collisionally-induced dissociation

cm centimeter

co codon optimised

CoA coenzyme A

CV coefficient of variation

CYP450 cytochrome P450

Da	dalton
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTNB	5,5'-dithio-bis[-2-nitrobenzoic acid]
DTT	dithiothreitol
ESI	electrospray-ionisation
FA	formic acid
g	gram
<i>g</i>	gravitational force
G	guanine
GLYAT	glycine N-acyl transferase
Gly	glycine
GNAT	Gcn5-related N-acyltransferase
H ₂ O	water
HA	hippuric acid
HCl	hydrochloric acid
HPLC	high-performance liquid chromatography
HPLC-MS/MS	high-performance liquid chromatography tandem mass spectrometry
K	potassium
kb	kilobases
K _m	Michaelis constant
L/min	liter per minute
LC	liquid chromatography
m/z	mass-to-charge-ratio
M	molar

mg	milligram
min	minute
ml	milliliter
mm	millimeter
mM	millimolar
MRM	multiple reaction monitoring
MS	mass spectrometry
NWU	North-West University
ng	nanogram
PCR	polymerase chain reaction
Q	quadrupole
rpm	revolutions per minute
r^2	coefficient of determination
SD	standard deviation
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel-electrophoresis
sec	seconds
SNP	single nucleotide polymorphism
T	thymine
T_m	melting temperature
V	volts
v/v	volume per total volume
V_0	initial velocity value
WT	wild-type
w/v	weight per total volume

Chapter 1: Introduction

1.1 Introduction

The metabolism of drugs and xenobiotics has taken the world of pharmacological science by storm, pertaining significantly to therapeutics and toxicology (Bachmann & Bickel 1986; Badenhorst *et al.*, 2013). Drug and xenobiotic metabolism involves a series of detoxification processes which are collectively known as the biotransformation system. The discovery of glycine conjugation was the starting point for studies on drug metabolism. The excretion of hippuric acid after ingestion of benzoic acid was discovered in 1841 by Alexander Ure (Ure, 1841). In 1845, it was demonstrated that hippuric acid was in fact an amide conjugate between glycine and benzoic acid, making this the first conjugation reaction to be discovered (Knights *et al.*, 2007).

Detoxification is an indispensable physiological process, since it decreases the toxicity of compounds that are not catabolised (Liska, 1998). Such compounds may include endogenous metabolites (eg. steroid hormones) and exogenous toxins (eg. compounds in food or industrial chemicals) (Campbell *et al.*, 1988). The detoxification system is divided into four phases. The first of these is phase 0 which describes the absorption of lipophilic drugs and other xenobiotic substances into cells located mainly in the gastrointestinal tract (Döring & Petzinger, 2014). Phase I detoxification describes the activation of these metabolites by the addition of functional groups. These activated compounds are then rapidly acted on by phase II detoxification systems, which serve to conjugate the activated functional groups of the compounds (Jancova *et al.*, 2010). Phase III detoxification involves the transportation of the detoxified, hydrophilic compounds across the cellular membranes, for excretion (Omiecinski *et al.*, 2011). The purpose of these conjugation reactions is to convert various endogenous and xenobiotic metabolites to more hydrophilic conjugates that can be excreted in the urine and bile (Knights *et al.*, 2007; Knights & Miners, 2012; Oates & West, 2006). The resulting conjugates are often less toxic than the parent

compound (Badenhorst *et al.*, 2013; Jakoby & Ziegler, 1990; Liska, 1998).

Glycine conjugation prevents the accumulation of benzoic acid in the mitochondrial matrix by forming a less lipophilic conjugate, hippuric acid, that can be easily transported out of the mitochondria (Badenhorst *et al.*, 2014). Nowadays, our exposure to benzoic acid, salicylate, solvents, and drugs that are metabolised to acyl-coenzyme A intermediates is ever-increasing and this is placing significant pressure on the glycine conjugation pathway. Therefore, the consequences of interindividual variation in the glycine conjugation pathway may become more noteworthy as more xenobiotic organic acids are encountered in the future (Knights *et al.*, 2007). Glycine N-acyltransferase (GLYAT (EC 2.3.1.13)), is one of the enzymes involved in the glycine conjugation metabolism in the phase II detoxification pathway described above, and is responsible for the detoxification of a wide range of metabolites originating from inter alia environmental and drug related toxins. GLYAT is a member of the Gcn5-related N-acetyltransferase (GNAT) superfamily of N-acyltransferases which uses an acyl-coenzyme A and glycine as substrates (Schachter & Taggart 1954; Vetting *et al.*, 2005).

Variations in the glycine conjugation pathway could influence liver cancer, musculoskeletal development, and mitochondrial energy metabolism. Thus, there exists an obvious need for characterisation of the enzyme and to consequently be able to interpret the interindividual variability in glycine conjugation capacity by considering the kinetic parameters, enzymatic mechanisms, and structures of the enzymes. This will lead us to better understand these enzymes and to treat metabolic diseases and toxicity more effectively (Loots *et al.*, 2005).

In order to define interindividual variability on a molecular basis, it will be useful to evaluate the kinetic properties of the conjugating enzyme GLYAT itself. Approximately 1424 single nucleotide polymorphisms (SNPs) have already been identified within the *GLYAT* gene of which 99 are non-synonymous (<http://www.ncbi.nlm.nih.gov/Database/> - November 2015). It was confirmed that SNP variations in the human *GLYAT* gene influenced the kinetic

properties of the enzyme, it was suggested that substrate specificity may be a key factor to aid in attempting to explain some of the interindividual variation found in glycine conjugation capacity. This finding is particularly significant to the treatment of some metabolic disorders as well as the metabolism of xenobiotics such as aspirin and the industrial solvent xylene (van der Sluis *et al.*, 2013).

Predicting the degree to which glycine conjugation variation affects the outcomes of organic acidemias, is complicated by the absence of a characterised relationship between GLYAT variants and their substrate specificities. Badenhorst *et al.* (2013) mentioned that the substrate selectivity of GLYAT and its variants needs to be further characterised, since organic acids can be toxic if the corresponding acyl-coenzyme A is not a substrate for GLYAT.

1.2 Dissertation outline

1.2.1 Chapter 2: Literature review

This chapter consists of a thorough literature review which focuses on topics such as biotransformation, the glycine conjugation pathway, and the glycine N-acyltransferase enzyme and its properties. The problem statement, aims and objectives for this study are also included in this chapter.

1.2.2 Chapter 3: The construction and expression of the N156S recombinant human GLYAT variant

Chapter 3 describes the generation of a biologically active N156S variant of human GLYAT via site-directed mutagenesis and the optimisation of the expression and storage conditions of this enzyme. The results for this section of the study are included along with a discussion and a summary of the content presented in the chapter.

1.2.3 Chapter 4: Bi-substrate kinetic analysis of the recombinant N156S hGLYAT variant

This chapter describes the determination of the bi-substrate kinetic parameters and the kinetic reaction mechanism for the N156S GLYAT variant. Detailed descriptions of the methods developed for this section of the study are provided along with the results obtained and a thorough discussion of these results. A short summary is also included at the end of this chapter.

1.2.4 Chapter 5: HPLC-ESI-MS/MS quantification of hippuric acid using a stable isotope

Chapter 5 describes the utilisation of an HPLC-MS/MS method for the identification and quantification of HA formed by the enzymatic reaction of GLYAT with different substrate concentrations. This chapter includes the methods and results of this study, and a comprehensive discussion of these results. A summary of the work in this chapter is also included.

A peer reviewed scientific paper is currently being prepared for publication regarding the work presented in this chapter.

1.2.5 Chapter 6: Final conclusions and future prospects

The final chapter discusses the final conclusions made from the results obtained in this study. Future prospects for this study are also included in this chapter.

1.2.6 References

The literature sources used in this study are listed in this section. The references are

provided according to the guidelines specified in the NWU manual for postgraduate studies.

1.2.7 Appendices

The reagents and materials used in this study, along with their suppliers and catalogue numbers are listed in Appendix A. The human GLYAT protein sequence is included in Appendix B.

Chapter 2: Literature review

2.1 Introduction

This chapter consists of a review of the literature relevant to this study. This overview will discuss the following topics: (i) drug metabolism in general; (ii) biotransformation; (iii) general role of biotransforming enzymes; (iv) the four phases of biotransformation with their respective reaction mechanisms and supportive nutrients; (v) a deeper look at phase II biotransformation, pertaining specifically to amino acid conjugation reactions; (vi) a more focused look at amino acid conjugations, concentrating on the glycine conjugation pathway and its enzymes, which includes glycine N-acyltransferase, the enzyme being investigated in this study; (vii) the molecular characteristics, substrate specificity and kinetic properties of the glycine N-acyltransferase enzyme; (viii) site directed mutagenesis; (ix) the general principles of recombinant protein expression and expression with chaperones; and finally (x) the implementation of a high performance liquid chromatography tandem mass spectrometry method for quantification of metabolites using a stable isotope. This chapter will also include the problem statement along with the research aims and objectives of this study.

2.2 Drug metabolism

Most major drug metabolism pathways were first discovered and verified in the 19th century. The discovery and investigation of these pathways played an important role in the development of disciplines such as general biochemistry, physiology and particularly pharmacology (Conti & Bickel, 1977). As early as 1859, Rudolf Buchheim, the founder of modern pharmacology, stated the following: “in order to understand the actions of drugs it is an absolute necessity to have knowledge of the transformations they undergo in the body. It is obvious that we must not judge drugs according to the form and amount

administered, but rather according to the form and amount which actually is eliciting the action" (Buchheim, 1859). It thus became increasingly clear that these pathways were worth investigating.

Nowadays we are constantly and unavoidably being challenged by foreign chemicals, or xenobiotic compounds. These compounds may be manufactured or naturally occurring chemicals, which can be metabolised to acyl-coenzyme A intermediates, benzoic acid, salicylic acid, solvents, pesticides, pollutants, industrial chemicals, secondary plant metabolites, and toxins produced by plants and animals (Parkinson, 2011). The metabolism of these substances fulfills a vital role in a living system, since these pathways are adapted for elimination of the toxic compounds from our bodies (Jakoby & Ziegler, 1990). Drug metabolism, or xenobiotic metabolism, describes the biochemical modification of toxic compounds or xenobiotics in living organisms. The drug metabolism pathways found in mammals, are highly evolved systems which involve specialised enzymatic systems through which the substances are converted into forms in which they are generally more readily excreted.

2.3 Biotransformation

Many xenobiotics are usually easily absorbed through the lungs, skin, or gastrointestinal tract. This trait is attributed to their lipophilicity, which allows them to be readily reabsorbed, but this property also serves as an obstacle for their elimination from the body. Consequently, in order for these xenobiotic compounds to be successfully eliminated, they must first be converted to hydrophilic chemicals. This occurs by means of a process called biotransformation, which primarily occurs in the liver and is accomplished by a limited number of biotransforming enzymes with broad substrate specificities (Parkinson, 2011; Jancova *et al.*, 2010; Wilcox *et al.*, 1999; Bonafé *et al.*, 2000; Nortje *et al.*, 2015). This process is regarded as the principal mechanism for the maintenance of homeostasis during exposure to xenobiotics. Without biotransformation, the lipophilic xenobiotics would not be

efficiently excreted from the body which could result in the buildup of toxic chemicals, eventually overwhelming and possibly killing an organism (Parkinson, 2011). The biotransformation system contains multiple enzyme-mediated reactions responsible for minimising the toxicity of xenobiotic compounds (Liska, 1998). Figure 2.1 is a schematic representation of the liver biotransformation system indicating its general role. It also includes the phase I and II biotransformation, as well as the more recently discovered phases 0 and III pathways. The schematic also shows the main reaction mechanisms and supportive nutrients involved with these steps.

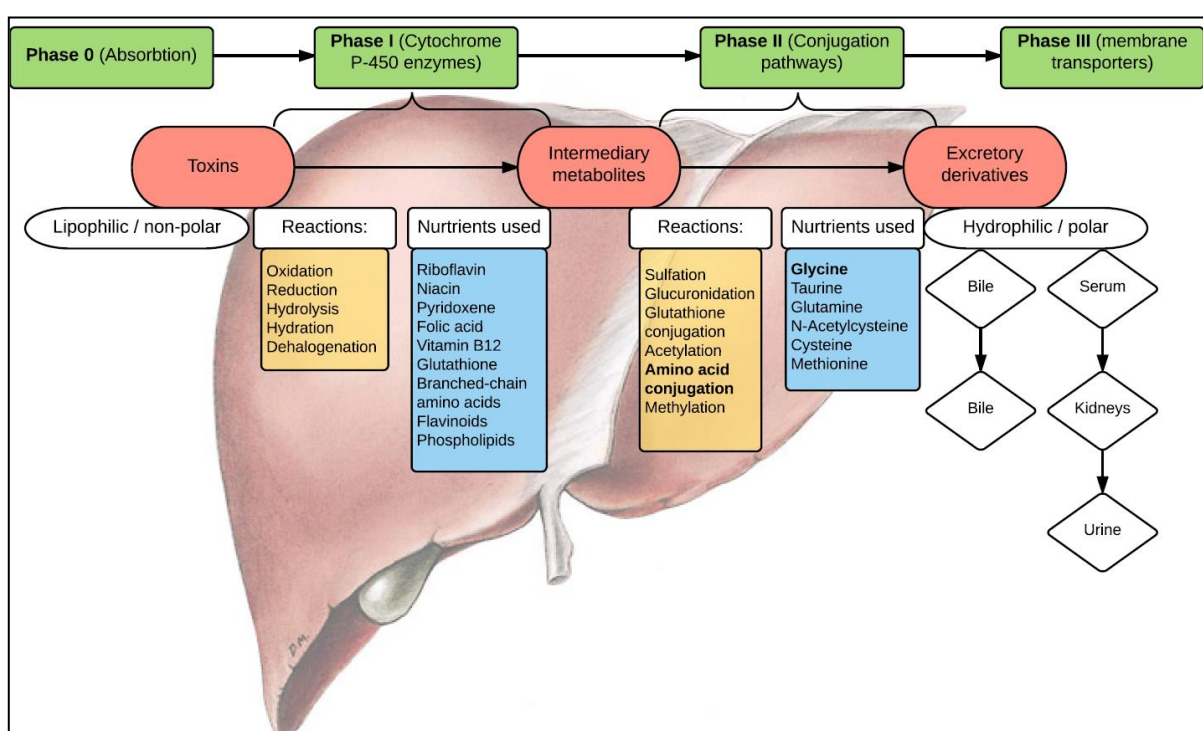


Figure 2.1: A schematic of the liver biotransformation processes, with their respective reaction mechanisms and supportive nutrients (Adapted from Liska, 1998).

2.3.1 Biotransforming enzymes

The enzymes participating in the metabolism and the detoxification of xenobiotic compounds have been a topic of much investigation. Remarkable discoveries relating to

their biotransformation roles have been made since the ground-breaking handbook of R.T. Williams was published in the mid-twentieth century (Bachmann & Bickel, 1986; Williams, 1959). These discoveries encompassed their amino acid sequences, their polymorphisms, genetic regulation, their cellular locale, their three-dimensional structures, their preferred substrates, their enzyme kinetics, their inhibition mechanisms and kinetics, and their allosteric regulation (Bachmann & Bickel, 1986).

The biotransforming enzymes are responsible for protecting the body from potentially harmful xenobiotics introduced into the body from the environment by biotransformation. They are particularly responsible for the metabolism, detoxification and elimination of these xenobiotics or exogenous substances (Xu *et al.*, 2005; Meyer, 1996). The reactions catalysed by the xenobiotic biotransforming enzymes are divided into three groups, called phase I, phase II and the latest addition, phase III (Parkinson, 2011; Nortje *et al.*, 2015). These xenobiotic biotransforming enzymes are also categorised as either phase I or II metabolising enzymes, or phase III transporters, according to their substrates, functional roles, and sequences in their respective metabolic pathways (Meyer, 1996; Rushmore & Kong, 2002; Xu *et al.*, 2005; Yamamoto & Williams, 1971).

Biotransformation traditionally involved the sequential biotransformation steps called phase I and phase II (Williams, 1959). The “phase I” and “phase II” terms were originally introduced in 1959 by Williams, and referred to the metabolism of acetanilide and phenacetin, oxidation, reduction and hydrolysis reactions for phase I, and synthesis reactions for phase II (Williams, 1959; Döring & Petzinger, 2014).

2.4 Phase I biotransformation

Phase I biotransformation is commonly described as the activation phase. The activation of compounds occurs via a variety of reactions including oxidation, hydroxylation, dehalogenation, desulfation, deamination, and reduction reactions (Parkinson, 2011; Porter

& Coon, 1991). These reactions expose a functional group ($-\text{OH}$, $-\text{NH}_2$, $-\text{SH}$ or $-\text{COOH}$), and typically result in only a slight increase in the hydrophilicity of the compounds. The phase I metabolising enzymes consist primarily of the cytochrome P450s (CYP450), monoamine oxidases and esterases flavin-containing monooxygenases, peroxidases, and alcohol dehydrogenases (Penner *et al.*, 2010; Xu *et al.*, 2005). The CYP450 enzymes, found mostly in the liver, gastrointestinal tract, and kidneys have attracted much attention and more than 150 isoforms of these enzymes have been characterised, some of which are capable of catalysing multiple reactions. Many inducers, substrates, and inhibitors of the CYP450 enzymes include a range of xenobiotics, which are consequently converted into activated intermediates, with a small increase in hydrophilicity (Xu *et al.*, 2005; Porter & Coon, 1991; Nortje *et al.*, 2015).

2.5 Phase II biotransformation

Phase II biotransformation, described as the conjugation phase, plays a significant role in the conversion of compounds into more hydrophilic forms (Nortje *et al.*, 2015; Xu *et al.*, 2005). Phase II biotransformation includes reactions such as glucuronidation, acetylation, sulfation, glutathione conjugation, methylation, and amino acid conjugation. During these reactions, compounds are conjugated to the activated functional groups of compounds (Jancova *et al.*, 2010). The primary purpose of phase II biotransformation is to carry out conjugating reactions. These reactions are catalysed by the phase II biotransformation enzymes, by conjugating endogenous and xenobiotic compounds to deliver more easily excretable metabolites, and also to metabolically inactivate pharmacologically active substances (Jancova *et al.*, 2010). The substrates for phase II conjugation reactions consist of the activated intermediates delivered from phase I as well as active pharmacological substances (Nortje *et al.*, 2015).

The cofactors for these conjugating reactions react with the functional groups present on the xenobiotic compounds, or the groups that were introduced/exposed during the phase I

biotransformation reactions (Parkinson, 2011). These reactions greatly increase the hydrophilicity of the endogenous and xenobiotic compounds. This results in the enhancement of their excretory capabilities, since these compounds are not likely to diffuse back into the hydrophobic interiors of phospholipid bilayers (Badenhorst *et al.*, 2014).

A detoxification effect is consequently achieved (Xu *et al.*, 2005; Parkinson, 2011). Although, phase II reactions are usually regarded as detoxifying, in certain situations, the conjugates formed from the conjugation with phase II enzymes could be activated metabolites which may mediate adverse effects and an increase in toxicity (Chen *et al.*, 2000; Kong *et al.*, 2000; Rushmore & Kong, 2002; Hinson & Forkert, 1995). An example of such an effect is the conjugates being used as carriers for possibly carcinogenic substances in the activation of polycyclic aromatic hydrocarbons, benzylic alcohols, hydroxamic acid, aromatic hydroxylamines, and nitroalkanes by sulphotransferases (Jancova *et al.*, 2010; Glatt, 2000).

In the years since the discovery of the biotransformation pathways, pharmacology studies have generally not focused as much on the phase II enzymes in comparison with the phase I CYP450 enzymes, since these enzymes did not display as many drug interactions as their phase I counterparts. It has since been proven, however, that a decrease in the metabolising capacity of the phase II enzymes, caused the accumulation of harmful compounds associated with clinical drugs, thus resulting in the manifestation of toxic effects (Jancova *et al.*, 2010).

The phase II metabolising enzymes or conjugating enzymes are mostly transferases, and include enzymes from many superfamilies including sulfotransferases (SULT) (Jancova *et al.*, 2010; Banoglu, 2000; Weinshilboum *et al.*, 1997), DT-diaphorase or NAD(P)H:quinone oxidoreductase (NQO) or NAD(P)H: menadione reductase (NMO) (Jaiswal, 1994; Kong *et al.*, 2001), glutathione S-transferases (GST) (Moscow & Dixon, 1993; Tew & Ronai 1999; Schilter *et al.*, 1993), UDP-glucuronosyltransferases (UGT) (Innocenti *et al.*, 2002; King *et al.*, 1999;

Tukey & Strassburg 2000), epoxide hydrolases (EPH) (Hinson & Forkert, 1995), and N-acetyltransferases (NAT) (Vatsis *et al.*, 1995).

The traditional two-phase explanation of biotransformation was the subject of much research, but this concept was still lacking, since these phases did not account for how the drugs and other xenobiotics entered the cells and how the hydrophilic metabolites left the cells for excretion from the body (Williams, 1959; Döring & Petzinger, 2014). The terms phase 0 and phase III were consequently introduced to describe the additional steps responsible for the absorption of the lipophilic xenobiotic compounds and elimination of the hydrophilic metabolites by means of carrier-mediated efflux, respectively (Döring & Petzinger, 2014; Ishikawa, 1992).

2.6 Phase 0 biotransformation

As mentioned in the previous section, the original biotransformation theory was inadequate in addressing the absorption of the xenobiotic compounds into the body. Xenobiotics are ingested along with over 25 tons of food over a lifetime, which are all passed through the gastrointestinal tract. Thus, it is here where the body first makes contact with these substances (Liska, 1998). It could therefore be assumed that the gastrointestinal tract was bound to develop complex biochemical systems to deal with this enormous load of exogenous compounds. The gastrointestinal tract was later described as the second major site for detoxification in the body, the liver being the primary detoxifying organ (Parkinson, 2011; Jancova *et al.*, 2010; Wilcox *et al.*, 1999; Bonafé *et al.*, 2000). The detoxification reactions occurring in the gastrointestinal tract, precedes the better known phase I and II biotransformation reactions, hence the term phase 0 biotransformation which was formulated to describe these reactions (Liska, 1998).

Phase 0 specifically describes the absorption of lipophilic drugs and other xenobiotic substances into cells located primarily in the gastrointestinal tract. The term “absorptive

carriers”, pertaining specifically to the luminal membranes in the gut, corresponds to phase 0 transporters. These phase 0 uptake transporters are regarded as imperative for the pharmacodynamics of the drugs that interact with intracellular targets, seeing that they are responsible for determining the selectivity of the cell entrance and intracellular concentration (Döring & Petzinger, 2014).

2.7 Phase III biotransformation

Phase III biotransformation refers to the reactions involving active membrane transporters that serve as transporters for drugs and other xenobiotic compounds across cellular membranes (Omiecinski *et al.*, 2011). The discovery of this part of the detoxification system only dates back to 1976, making it a relatively new research field, in contrast with the well established history of the phase I and phase II biotransformation pathways, which was developed as early as the 1800’s (Omiecinski *et al.*, 2011; Conti & Bickel, 1977). A discovery made by Juliano and Ling (1976) regarding a 170-kDa carbohydrate complex which modulates drug permeability in the extracellular membranes of Chinese hamster ovary (CHO) cells, lead to the discovery of what are now called the ATP-binding cassette (ABC) family of drug transporters. Other key xenobiotic and drug transporters with similar reaction mechanisms include the organic cation and anion transporters (SLC22A superfamily) and the organic anion-transporting polypeptides (SLCO superfamily) (Hagenbuch 2010). Later on Ishikawa (1992) coined the term “phase III” biotransformation to describe the reactions of these transporters (Ishikawa, 1992). Shortly after this, Zimniak *et al.* (1993) described this phase as the next sequential step for the drug derivatives and metabolites awaiting elimination, after having passed through phases I and II biotransformation.

2.8 Amino Acid Conjugation

As mentioned previously, amino acid conjugation reactions form part of phase II

biotransformation (Jancova *et al.*, 2010; Parkinson, 2011). Figure 2.1 shows the approximate location of these conjugation reactions with regard to the overall detoxification system. Amino acid conjugation is physiologically essential for the detoxification and subsequent elimination of various xenobiotic compounds and endogenous metabolites which may include drugs, dietary components, and food additives (Webster *et al.*, 1976) .

There exist two primary pathways by which xenobiotic compounds are conjugated with amino acids. The first pathway, or activation step, occurs mainly in mitochondria and involves xenobiotic compounds holding a carboxylic acid group conjugating with the amino group of amino acids including glycine, taurine, and glutamine (Parkinson, 2011). This pathway entails the activation of the xenobiotic substance by conjugation with coenzyme A, which produces an acyl-coenzyme A thioester. This compound then reacts with the amino group found on the amino acids to form an amide linkage. This reaction is dependent on ATP and is catalysed by acyl-coenzyme A synthetases (ATP-dependent acid:coenzyme A ligases). The second pathway is one in which the xenobiotic compound with an aromatic hydroxylamine (N-hydroxy aromatic amine) conjugates with the carboxylic acid group of amino acids such as serine and proline (Parkinson, 2011). This pathway is catalysed by acyl-coenzyme A:amino acid N-acyltransferases, which transfers the acyl moiety of the xenobiotic compound to the amino group of the acceptor amino acid (Parkinson, 2011; Webster *et al.*, 1976). The amino acid conjugates of xenobiotic compounds are primarily eliminated in urine.

2.9 Glycine conjugation

2.9.1 An introduction to glycine conjugation

Benzoic acid was the first compound to have its biological footprint traced in the human body. In 1801, it was proposed that urine may contain a substance similar to benzoic acid, and was only confirmed about three decades later when hippuric acid was discovered in the

urine of horses (Conti & Bickel, 1977). In 1841, Ure demonstrated the presence of hippuric acid in urine following the ingestion of benzoic acid (Ure, 1841), and in the next year this discovery was corroborated by Keller (Keller, 1842). In 1845, Dessaignes showed that hippuric acid formed as a product of benzoic acid and glycine conjugation (Conti & Bickel, 1977). These discoveries marked a new point of departure in the investigation of biotransformation. Indeed, it set the stage for researchers to elucidate most of the major drug metabolism pathways before the turn of the 19th century. This included conjugation to sulfate, acetate, and glucuronic acid, as well as oxidation and reduction, and hydrolysis reactions. Moreover, as a result, many of the enzymes involved in the former reactions have since been characterised (Conti & Bickel, 1977; Coon, 2005; Gamage, 2005; Guillemette, 2003; Oates & West, 2006; Ritter, 2000; Rodriguez-Antona *et al.*, 2010; Weinshilboum *et al.*, 1997).

Because of their impact on metabolism and predisposition to adverse drug effects, these biotransformation reactions are still receiving a great amount of attention in the pharmacology and toxicology research communities (Gamage, 2005; Guillemette, 2003; Ritter, 2000; Rodriguez-Antona *et al.*, 2010). Research on glycine conjugation has, however, enjoyed much less scrutiny after its discovery. This is probably due to the fact that a very small number of pharmaceutical drugs are conjugated to glycine (Badenhorst *et al.*, 2013; Knights & Miners, 2012; Knights *et al.*, 2007). However, in the last couple of years interest in glycine conjugation has been revitalised, since it has become clear that glycine conjugation may actually be a very important metabolic pathway. This is based on observations that the glycine conjugation pathway overlaps with and influences the metabolic pathways of ATP, coenzyme A, glycine and other versatile metabolites. It has recently been stated that glycine conjugation, which is thought of as the “poor cousin” of the drug metabolism family, may now have “inherited a fortune” (Beyoğlu & Idle, 2012; Beyoğlu *et al.*, 2012).

2.9.2 The glycine conjugation pathway

In the mitochondrial matrix of cells in the mammalian liver and kidney, a number of xenobiotic organic acids are conjugated to glycine, which facilitates the urinary excretion of these compounds (Knights & Miners, 2012; Knights *et al.*, 2007). In most mammals, hippuric acid is the most abundant amino acid conjugate in urine (Lees *et al.*, 2013), and is formed in two steps (see Figure 2.2). First, a mitochondrial ATP dependent acid:coenzyme A ligase, otherwise known as HXM-A (E.C. 6.2.1.2), activates benzoic acid by converting it to benzoyl-coenzyme A (Knights, 1998; Knights & Drogemuller, 2000; Schachter & Taggart, 1953; Vessey *et al.*, 1999; Vessey *et al.*, 2003). Second, the GLYAT enzyme binds benzoyl-coenzyme A and converts it to hippuric acid and coenzyme A via the acylation of glycine (Schachter & Taggart, 1953; Schachter & Taggart, 1954).

The main substrates for glycine conjugation that occur naturally include salicylic acid, benzoic acid, 3- and 4-hydroxybenzoic acid, and 4-aminobenzoic acid. These substrates do not, however, make up the majority of glycine conjugation substrates. The main source for these substrates is the metabolites of dietary polyphenols such as flavonols and hydroxycinnamates that are produced by the gut microbes (Knights & Miners, 2012; Rechner *et al.*, 2002). Dietary polyphenols are then converted by the gut microbiota into aromatic compounds including phenylpropionate, cinnamate, and benzoic acid (Bravo, 1998; Jenner *et al.*, 2005; Rechner *et al.*, 2002; Tsao, 2010). These aromatic acids are activated by ATP dependent acid:coenzyme A ligases to acyl-coenzyme A thioesters such as phenylpropionyl-coenzyme A. Phenylpropionyl-coenzyme A is then converted to cinnamoyl-coenzyme A by medium chain acyl-coenzyme A dehydrogenase, making up the first step of the β -oxidation cycle. Cinnamoyl-coenzyme A is oxidised to benzoyl-coenzyme A and acetyl-coenzyme A. The benzoyl-coenzyme A is conjugated to glycine and converted to hippuric acid by GLYAT. This results in hippuric acid being the main urinary metabolite of phenylpropionate catabolism and polyphenol metabolism (Dakin, 1908; Rechner *et al.*, 2002).

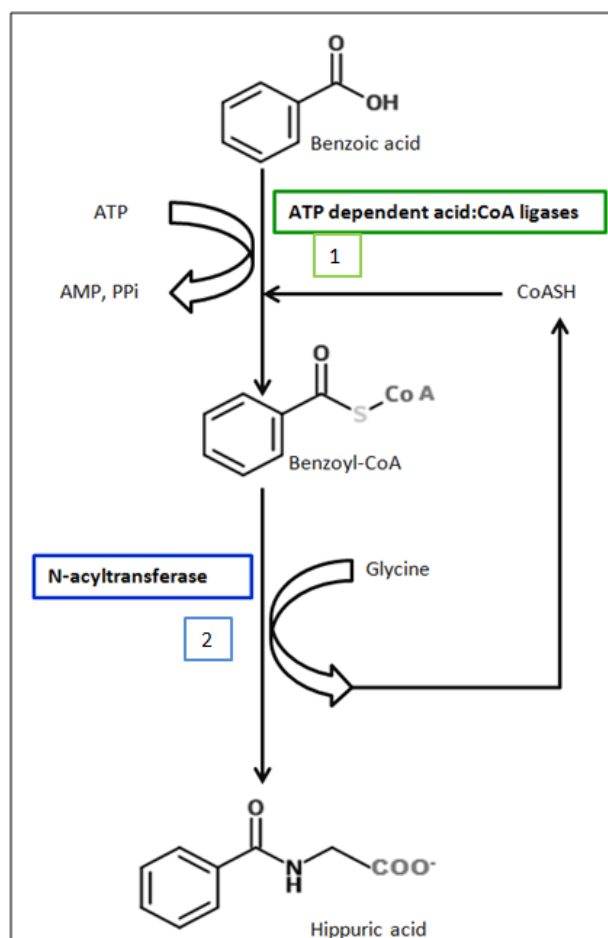


Figure 2.2: Metabolic pathway of glycine conjugation. Benzoyl-coenzyme A is converted to hippuric acid by glycine N-acyltransferase. The green box indicates the first step of glycine conjugation, catalysed by the ATP dependent acid:coenzyme A ligases. The blue box indicates the second step of glycine conjugation, catalysed by the N-acyltransferase enzymes. Adapted from Badenhorst *et al.*, 2014).

2.9.3 Interindividual variation in glycine conjugation

The major metabolite of benzoic acid is hippuric acid. In some cases in which large doses of dietary polyphenols are ingested, other metabolites such as benzoylglucuronide and benzoylcarnitine may also be formed (Bray *et al.*, 1951; Sakuma, 1991). This formation of

secondary metabolites is dependent on the dose of benzoic as well as the capacity of the glycine conjugation system (Bray *et al.*, 1951; Saltzman & Caraway, 1953). Significant interindividual variation of urinary hippuric acid levels in humans has also been reported. According to the Human Metabolome Database, urinary hippuric acid concentrations can range from 27.92 to 932.66 mmol/mmol creatinine (Lees *et al.*, 2013; Wijeyesekera *et al.*, 2012). This variation can be attributed to the factors influencing the glycine conjugation pathway (discussed in Section 2.9.2) or other factors that do not necessarily reflect on the glycine conjugation capacity of an individual. It is important, however, to differentiate between this variation in hippuric acid concentrations excreted, and variation in the rate of glycine conjugation itself. Patients with very low urinary hippuric acid levels might still demonstrate a normal rate of hepatic glycine conjugation when challenged with an oral dose of benzoic acid (Williams *et al.*, 2010). This is noticed especially in patients with Crohn's disease during which significant alterations of the gut microbiome is present, resulting in decreased fermentation of dietary polyphenols and consequently reduced production of phenylpropionate (Badenhorst *et al.*, 2014). The polyphenol fermentation products formed by the gut microbiota depend on the composition of the gut microbiome, the type and quantity of food consumed, as well as its total passage time through the gastrointestinal tract (Bravo, 1998; Fedotcheva *et al.*, 2008; Rechner *et al.*, 2002). The use of antibiotics may cause suppression of this microbial activity, which can also result in decreased production of phenylpropionate and therefore also lower urinary hippuric acid excretion, but does not at all reflect on the glycine conjugation capacity (Lees *et al.*, 2013).

There are several factors that may influence the overall rate or capacity of glycine conjugation pathway. These include: (i) the availability of ATP, coenzyme A, and glycine; (ii) genetic variations and expression of the enzymes catalyzing this pathway; (iii) duons. Short summaries of how these factors influence glycine conjugation are described below.

2.9.3.1 Availability of ATP, coenzyme A, and glycine

As previously mentioned the synthesis of benzoyl-coenzyme A from benzoic acid describes

the first step in the glycine conjugation pathway (Figure 2.2). Unacylated coenzyme A is needed for this reaction to take place, and it can therefore be assumed that the availability of coenzyme A can directly limit the rate of glycine conjugation (Gregus *et al.*, 1996). This step is also dependent on two molecules of ATP per molecule of benzoic acid which is hydrolysed to AMP (Schachter & Taggart, 1953; Schachter & Taggart, 1954). The unavailability of ATP will, thus, significantly limit the glycine conjugation rate by restricting the formation of benzoyl-coenzyme A (Gregus *et al.*, 1996). Glycine availability is another factor that can severely influence the tempo of hippuric acid synthesis (Figure 2.2) (Beliveau & Brusilow, 1987; De Vries *et al.*, 1948; Levy, 1979). Therefore, even though it is known that hippuric acid synthesis in humans is saturable even with large doses of benzoic acid, the co-administration of glycine will considerably boost the rate of hippuric acid formation (Knights & Miners, 2012; Levy, 1979).

2.9.3.2 Genetic variation in HXM-A and GLYAT

The influence of genetic variation on the glycine conjugation pathway has not been studied comprehensively (Badenhorst *et al.*, 2014). HXM-A, a ligase enzyme active in the first step of glycine conjugation is encoded for by the *ACSM2B* gene. Further investigations regarding the effect of genetic variation on the enzymatic parameters of the enzyme still needs to be conducted (Boomgaarden *et al.*, 2009).

Investigations of the genetic variation in the *GLYAT* gene encoding for the GLYAT enzyme, which participates in the second glycine conjugation reaction, have been done. Thus far, it was concluded that genetic variation had an influence on the enzyme activity of recombinant human GLYAT enzymes that were bacterially expressed (van der Sluis *et al.*, 2013). Van der Sluis and colleagues studied the enzyme activities of these recombinant enzymes with six non-synonymous SNPs (N156S; R131H; F168L; R199C; S17T; and K16N). The kinetic parameters of these variants were determined (for benzoyl-coenzyme A) and compared to those determined for the wild-type enzyme. It was found that the K16N and S17T variants had K_m values similar to the wild-type enzyme while the N156S variant had an

increased K_m value. The F168L variant showed a decreased enzyme activity. Both the R199C and E227Q variants of recombinant human GLYAT were virtually inactive, having less than 5% of the activity of the wild-type. Figure 2.3 shows the relative enzyme activities determined for these recombinant human GLYAT variants, while Table 2.1 summarises the K_m values determined for each variant.

In a second study by van der Sluis *et al.* (2015) which investigated the genetic variation of the *GLYAT* gene, it was found that the N156S polymorphism had the highest homozygous genotype frequency of 89.9%, followed by S17T (4.6%) and R131H (0.1%). Thus, the human GLYAT variant (N156S) that demonstrated the highest enzyme activity was also the variant with the highest allele frequency. On the other hand, the variants that displayed negative effects on the enzymatic activity were found to occur extremely rarely (van der Sluis *et al.*, 2015). It was consequently suggested that the N156S GLYAT variant be considered the “wild-type” allele, since it had a much higher enzyme activity and allele frequency in comparison to the wild-type variant encoded by the reference sequence (NM_201648.2) (van der Sluis *et al.*, 2013; Badenhorst *et al.*, 2013).

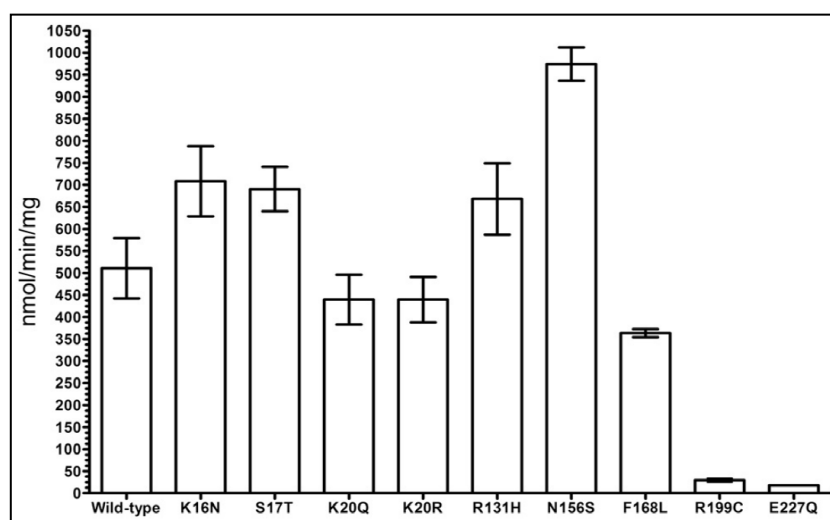


Figure 2.3: Results presented by van der Sluis *et al.* (2013) showing the relative enzyme activities of selected variants of recombinant human GLYAT. Error bars indicate the mean $\pm$ standard deviation for triplicate assays.

Table 2.1: A summary of the K_m values (benzoyl-coenzyme A) obtained for recombinant human GLYAT variants.

Variant	K_m (benzoyl-coenzyme A) (μ M)
Wild-type	24 $\pm$ 3
K16N	21 $\pm$ 1
S17T	28 $\pm$ 5
R131H	71 $\pm$ 11
N156S	38 $\pm$ 4
F168L	53 $\pm$ 6

Values obtained from van der Sluis *et al.* (2013).

2.9.3.3 *In vivo* expression levels of HXM-A and GLYAT

The processes regarding gene expression *in vivo* are extremely complex, and fluctuations in protein levels may be ascribed to various factors (Pedraza & Paulsson, 2008). Even though the exact mechanisms regarding the induction of the ligase and transferase enzymes are not well understood, it is known that they are inducible (Badenhorst *et al.*, 2014). Studies in which rats were pre-treated with salicylate for six days, showed increases in the synthesis of hippuric acid, salicylylglycine, and 4-aminohippuric acid of their liver and kidney homogenates (Irljala, 1972). Salicylate pre-treatment of humans also revealed increased rates of salicylylglycine formation (Furst *et al.*, 1977). It was also recently found that dietary restrictions of rats increased the expression of GLYAT in their livers, since an increase in urinary acylglycine excretion was observed (Wen *et al.*, 2013). Based on studies conducted with hepatocellular carcinoma specimens, it was found that GLYAT expression is regulated transcriptionally (Matsuo *et al.*, 2012).

2.10 Molecular characteristics of GLYAT

As alluded to in Section 2.9.2, GLYAT is a phase II detoxification enzyme that facilitates the detoxification and subsequent bile and urinary excretion of several toxic organic acids, via glycine conjugation (Nandi *et al.*, 1979; Schachter & Taggart, 1954). These typically include benzoic acid, methyl-benzoic acid, salicylic acid, and many other endogenous metabolites. Although the exact structure of GLYAT has not yet been elucidated, the literature suggests a molecular size of approximately 27 kDa to 30 kDa (Nandi *et al.*, 1979; van der Westhuizen *et al.*, 2000). The human gene encoding GLYAT is located on chromosome 11, at position 11q12. The transcript is approximately 23 000 base pairs long, and contains six exons. It also contains 1424 known single nucleotide polymorphisms (SNPs), of which 99 are non-synonymous (www.ensembl.org, November 2015).

The open reading frame (ORF) of the human *GLYAT* gene, XXX bp in size, is highly conserved and does not display great genetic diversity (van der Sluis *et al.*, 2013; van der Sluis *et al.*, 2015). Two haplotypes (S_{156} and $T_{17}S_{156}$) are found in all populations and occur at relatively high frequencies (~70% and 20%, respectively). The $S_{156}C_{199}$ haplotype has an extremely low frequency (0.05%), and it has been found that this haplotype only has <5% residual GLYAT enzyme activity. The $S_{156}H_{131}$ haplotypes also displays a very low allele frequency, which correlates well with its weak substrate affinity (van der Sluis *et al.*, 2013). The fact that GLYAT is so well conserved indicates its importance in the glycine conjugation pathway, and the importance of this pathway in the overall detoxification pathway.

2.11 The enzymatic properties of GLYAT

2.11.1 Enzymatic reaction

The phase II biotransformation pathway consists of a series of conjugation reactions, catalysed by biotransforming enzymes, which cooperate to decrease the toxicity of xenobiotic compounds and to make them more soluble for excretion via the urine and bile.

One such enzyme is GLYAT, which serves to detoxify a wide range of endogenous and xenobiotic compounds. These substances include benzoic acid (found in fruits, vegetables, medicines, and as a preservative in foods and drinks), salicylic acid (a metabolite of aspirin), methyl-benzoic acid (a metabolite of toluene) and several endogenous metabolites which include propionic and isovaleric acid (Bartlett & Gompertz, 1974; Fenton *et al.*, 2001; Kølvrå & Gregersen, 1986; Nandi *et al.*, 1979; van der Westhuizen *et al.*, 2000; Sweetman & Williams, 2001; Tanaka & Isselbacher, 1967; Temellini *et al.*, 1993). The variety of endogenous substances detoxified by this reaction is demonstrated by the extensive range of acylglycines found in the urine of those with organic acid metabolism defects (Bartlett & Gompertz, 1974; Ogier de Baulny & Saudubray, 2002; Sweetman & Williams, 2001; Tanaka & Isselbacher, 1967). No defect has yet been described for GLYAT, but as mentioned earlier in Section 2.9.3, significant interindividual variation occurs in glycine conjugation capacity (Temellini *et al.*, 1993).

2.11.2 Substrate specificity

The substrate specificity of the GLYAT enzyme has been a topic of quite a bit of research in the past few decades. The K_m and V_{max} values of GLYAT have been determined for an assortment of substrates, of which some are summarised in Table 2.2. It was noticed, however, that the values reported in the literature were not very comparable between studies, which significantly complicated the interpretation of these kinetic parameters, since no consensus exists around the values reported (Table 2.2). This variation can be attributed to differences in the reaction temperatures used, quality of enzyme preparations used, different buffer or buffer concentrations used, the presence or absence of potassium ions in the reaction mixtures, differences in the quality of substrates used, as well as genetic variation (Bartlett & Gompertz, 1974; Gregersen *et al.*, 1986; Kelley & Vessey, 1994; Kølvrå & Gregersen, 1986; Mawal & Qureshi, 1994; Nandi *et al.*, 1979; Schachter & Taggart, 1954; van der Westhuizen *et al.*, 2000). Most kinetic parameters reported in the literature were determined using liver samples; it was thus not known which variants were characterised.

Table 2.2: The K_m values for acyl-coenzyme A substrates and glycine for GLYAT isolated from liver samples.

Acyl-coenzyme A substrates	K_m for acyl-coenzyme A (μM)	Ref.*	K_m for glycine (mM)	Ref.*
Butyryl-coenzyme A	2400 ± 800	1	970 ± 210	1
Isobutyryl-coenzyme A	5580 ± 1440	1	2880 ± 1190	1
Isovaleryl-coenzyme A	672 ± 164	1	523 ± 206	1
	124	3		
Hexanoyl-coenzyme A	2680 ± 770	1	1150 ± 210	1
Octanoyl-coenzyme A	322 ± 87	1	770 ± 110	1
	198	3		
Decanoyl-coenzyme A	2408 ± 887	1	690 ± 210	1
Benzoyl-coenzyme A	67 ± 5	2	6.5 ± 1	2
	57.9	4	6.4	4
	13			

*1. Gregersen *et al.*, 1986; 2. Kelley & Vessey, 1994; 3. Mawal & Qureshi, 1994; 4. van der Westhuizen *et al.*, 2000).

2.11.2.1 The acyl donor substrate

A variety of aromatic or substituted aromatic groups, as well as a range of short and medium chain aliphatic acyl groups can be transferred by GLYAT. This enzyme, however, has its preferences regarding substrate choice and seems to be for four or five carbon and benzoyl groups. While its affinity for shorter chain substrates such as propionyl-coenzyme A (C3) and acetyl-coenzyme A (C2) is much lower (Nandi *et al.*, 1979). It has been demonstrated that most mammal GLYAT orthologs (isolated from the livers and kidneys of rats, rabbits, sheep, rhesus monkeys, cows, and humans) also prefer benzoyl-coenzyme A as a substrate (Gregersen *et al.*, 1986; Kelley & Vessey, 1990; Kelley & Vessey, 1994; Kølvrå & Gregersen, 1986; Mawal & Qureshi, 1994; Nandi *et al.*, 1979; Schachter & Taggart, 1954;

Webster *et al.*, 1976) There exist many alternative acyl donor substrates, such as hexanoyl-coenzyme A, and isovaleryl-coenzyme A, 4-aminobenzoyl-coenzyme A and salicylyl-coenzyme A, but they are much less efficient than benzoyl-coenzyme A (Badenhorst *et al.*, 2012; Bartlett & Gompertz, 1974; Kølvrå & Gregersen, 1986; Mawal & Qureshi, 1994; Nandi *et al.*, 1979). Figure 2.4 represents the order of substrate preference of the GLYAT enzyme. The acyl groups shown in the figure are arranged according to the preference of GLYAT.

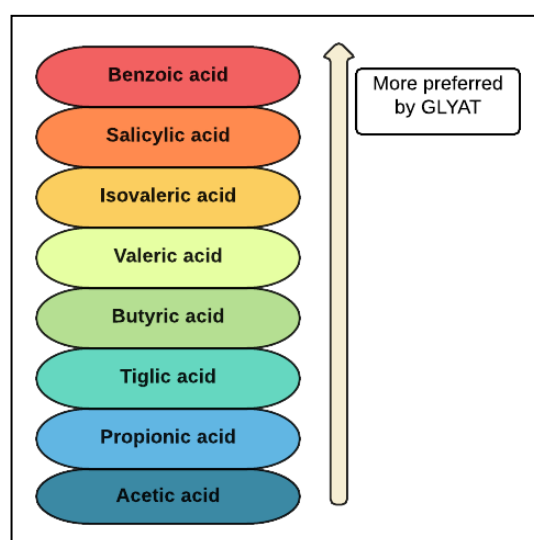


Figure 2.4: Substrate specificity of the GLYAT enzyme. The acyl groups shown in the figure are arranged according to the preference of GLYAT. The arrow indicates the direction of ascending preference. Benzoyl-coenzyme A is thus regarded as the best substrate while acetyl-coenzyme A is regarded as the poorest substrate (Nandi *et al.*, 1979).

2.11.2.2 The amino acid substrate (acyl acceptor substrate)

The preferred amino acid or acyl acceptor substrate for human GLYAT is glycine (Gregersen *et al.*, 1986; van der Westhuizen *et al.*, 2000). The GLYAT enzyme can also utilise alanine and glutamate, but at significantly lower rates. Table 2.3 summarises the kinetic parameters for the different amino acids (with benzoyl-coenzyme A as the acyl-donor) for GLYAT (van der Westhuizen *et al.*, 2000). Although many other acyl-acceptor substrates exist, such as the

amino acids asparagine, glutamate, alanine, serine, glutamine, and threonine, GLYAT strongly favours glycine. In addition, these other amino acids can only be used at low rates, which appear to be irrelevant under normal physiological conditions (van der Westhuizen *et al.*, 2000; Webster *et al.*, 1976).

Table 2.3: K_m values for glycine and alanine; and benzoyl-coenzyme A for human GLYAT.

Amino acid conjugate	K_m for benzoyl-coenzyme A (μM)	K_m for amino acid (mM)
Benzoyglycine	13	6.4
Benzoylalanine	15.2	997

Values from van der Westhuizen *et al.* (2000) used.

Under normal conditions the use of alternative amino acids such as alanine is not physiologically significant, since their formation is so slow. When metabolic stress is present in the body, for example persons suffering of organic acidemias, the alternative amino acid conjugations can become more significant, since an overload of acyl-coenzyme A and limiting amounts of glycine are present in the body of such patients (Beliveau & Brusilow, 1987). Examples supporting this observation are summarised in Table 2.4.

Table 2.4: Summary of some examples of conditions in which the alternative amino acid substrates would be utilised by GLYAT.

Defects	Conjugates formed
Propionic acidemia	Propionylserine Propionylglutamate Propionylglycine
Urea cycle defects	Benzoylalanine Benzoylglycine
Isovaleric acidemia	Isovalerylalanine Isovalerylglutamate Isovalerylglycine etc.

Compiled using (Fenton *et al.*, 2001; Lehnert, 1983; Loots *et al.*, 2005; van der Westhuizen *et al.*, 2000)

2.11.2.3 Organic acidemias

As mentioned previously, GLYAT is responsible for the detoxification of many toxic organic acids by means of conjugation to glycine, which results in less toxic acylglycines which can be easily excreted (Bartlett & Gompertz, 1974; Bonafé *et al.*, 2000; Dercksen *et al.*, 2012; Nandi *et al.*, 1979; Schachter & Taggart, 1954; Tanaka & Isselbacher, 1967). The efficiency of the detoxification of these accumulating acyl-coenzyme A metabolites determines the clinical outcomes of organic acidemias (Sweetman & Williams, 2001; Tanaka & Isselbacher, 1967).

It was found that a relationship exists between the substrate selectivity of GLYAT and the acylglycines excreted in the urine of patients with organic acidemias (Bartlett & Gompertz, 1974). One such example is isovaleric acidemia, in which isovaleryl-coenzyme A accumulates. Isovalerylglycine is excreted in large amounts in the urine of these patients,

since isovaleryl-coenzyme A is a good substrate for GLYAT (as seen in Figure 2.4). The isovaleryl-coenzyme A can thus be conjugated with glycine and consequently detoxified. Glycine supplementation is often regarded as a successful treatment for this disease, as it aids this detoxification process (Krieger & Tanaka, 1976; Sweetman & Williams, 2001; Tanaka & Isselbacher, 1967). However, in patients with propionic acidemia, a disease in which propionic acid accumulates, only limited amounts of propionic acid are excreted as propionylglycine (Fenton *et al.*, 2001). This is because propionic acid is not a good substrate for GLYAT (Figure 2.4). Thus, even with glycine supplementation, these compounds are not very efficiently detoxified by GLYAT. A treatment for propionic acidemia is yet to be developed (Fenton *et al.*, 2001). A possible solution would be to find a variant of the GLYAT enzyme that is capable of efficiently detoxifying propionic acid. It is thus of utmost value to investigate the substrate specificities of GLYAT enzymes with different SNPs, to evaluate their affinity for propionic acids. Then it would be possible to 'design' a recombinant therapeutic GLYAT enzyme with this property to treat propionic acidemia, in combination with glycine supplementation. When the substrate binding and specificities of GLYAT variants are more fully characterised, it may even be possible to manipulate GLYAT enzymes to use other substrates efficiently and, eventually, to treat other organic acidemias.

2.12 Site-directed mutagenesis

Site-directed mutagenesis is a molecular technique used to create intentional changes to the DNA sequence of a specific gene of interest. Typically, when a new nucleotide is introduced into the coding sequence, it changes one amino acid of the translated protein. In the context of protein engineering, this technique can provide useful information about the importance of one or more amino acids (Carter, 1986). For instance, if it is thought that a certain amino acid residue is functionally important, one can presume that substituting it with another will either alter or terminate enzyme function. Then, by generating a mutant recombinant enzyme, as described in Chapter 3, this hypothesis can be explored by investigating its biochemical properties. When assessing the effects of a certain mutation, it is imperative to keep in mind that an amino acid residue in a protein can have both a

functional and structural role. Thus, when a protein loses function due to a mutation, it can be ascribed to altered catalysis or it may be a consequence of altered protein folding, solubility and stability. To steer clear of making a mistake, a few guidelines on proper mutation design and subsequent analyses should be followed:

a) Only when possible, amino acids should be replaced by residues that have a similar structure, but are chemically distinct. As such, it can be said with more confidence that altered activity is a consequence of the loss of the specific chemistry of the residue and not due to a change of protein structure or stability.

b) In contrast, by replacing the residue to another that is rather chemically similar but structurally distinct, the abovementioned hypothesis can be tested. This should then confirm that not the structure, but the chemical nature of the amino acid essentially affects catalysis.

c) Finally, it will help to alter the assay conditions in such a way that the effect of the mutation is augmented. If these guideline are implemented, site-directed mutagenesis will be the most powerful approach one can follow to elucidate an enzyme reaction mechanism

2.13 General principles of recombinant protein expression

In the case where mutant GLYAT proteins are needed for biochemical characterisation, it is unrealistic to expect that these proteins be extracted and purified from human liver samples. The question was raised: how to obtain these proteins in a recombinant form? (Rosano & Ceccarelli, 2014). There are many answers to this question, but two stand out prominently. First, mammalian expression systems, such as CHO cells, have been used to express several complex proteins with considerable success. However, there are some substantial drawbacks. In terms of growth medium and culture conditions, this system is

relatively expensive. In addition, it is impractical and sometimes inconvenient due to slow growth, low densities and contamination.

An alternative that has become quite popular is the gram negative bacterium *Escherichia coli* (*E. coli*) expression system. Many soluble and biologically active proteins have been produced with this method (Sørensen & Mortensen, 2005). This system is superior to the mammalian expression system in three ways. (i) It has extremely fast growth kinetics. In optimal environmental conditions it can reach a doubling time of 20 min (Sezonov *et al.*, 2007). (ii) High cell densities can be achieved rather easily. (iii) The components needed to make rich complex media are readily available and inexpensive. (iv) Transformation with exogenous DNA is not difficult and very fast.

For this reason, in an effort to enhance its efficiency, bacterial expression systems have been the recipients of profuse experimental attention (Baneyx, 1999). However, several complex proteins, especially mammalian proteins, have not yet been successfully expressed using this host. Usually, transcription and translation levels are not an obstacle, but the expressed protein is often biologically inactive due to insolubility and incorrect folding (Baneyx, 1999; Sørensen & Mortensen, 2005). This drawback to bacterial expression is due to the lack of post-translational modification of the expressed protein. This problem can be moderated by the engineering of cell culture conditions and cellular physiology (Hannig & Makrides, 1998). A less complicated strategy, which appears to be more generally successful, involves the co-expression of the target protein with molecular chaperones (Nishihara *et al.*, 1998; Nishihara *et al.*, 2000; Walter & Buchner, 2002). This is discussed in more detail in the following section.

In previous studies done in our laboratory, human GLYAT has been successfully expressed in a soluble, enzymatically active form. The expression conditions included the use of co-expressed molecular chaperones. This protocol was optimised by van der Sluis *et al.* (2013) and is summarised briefly. An expression medium was prepared using 2% bacto-tryptone

(w/v), 1.25% yeast extract (w/v), 0.625% NaCl (w/v), 0.5% Na₂HPO₄·12H₂O (w/v), 0.1% KH₂PO₄ (w/v), and 0.2% (w/v) glucose. The desired construct is used to inoculate 10 ml of the expression medium. The cultures are incubated overnight at 37 °C with shaking at 180 rpm. The cells are then harvested by centrifugation (4000 x g, 5 minutes, 4 °C). The harvested cells are resuspended in 100 ml of the same expression medium and if necessary, the cultures are further incubated at 37 °C until the optical density values of the cultures are <0.5. For expression, the cultures are then incubated at 28 °C for 24 hours with shaking at 200 rpm (van der Sluis *et al.*, 2013). The molecular chaperones are induced by adding 2 mg/ml L-Arabinose to the medium before use. No isopropyl β-D-1-thiogalactopyranoside (IPTG) is added due to the proteins expressing through 'leaky' expression.

2.13.1 Co-expression of chaperone proteins

The three dimensional structure of a protein is determined by its primary amino acid sequence. Molecular chaperones are proteins that assist in the folding and unfolding or assembly and disassembly of other proteins. The underlying chemistry of chaperone action is complex and will not be discussed here. It is only necessary to mention that they counteract premature folding, or facilitate the unfolding or refolding of poorly folded proteins (Baneyx, 1999; Nishihara *et al.*, 1998; Nishihara *et al.*, 2000; Walter & Buchner, 2002). The company Takara designed a set of plasmids that contain different combinations of various chaperone proteins, and is commercially available. For this study, the Takara pGTf2 plasmid, containing the GroEL, GroES, and trigger factor (TF) chaperones, will be used. TF prevents premature folding of new peptides by binding to ribosomes and associating with hydrophobic amino acid stretches during translation. GroEL unfolds misfolded proteins by binding to them. This complex then readily associates with a heptameric GroES complex, which results in the release of the protein into the barrel formed by GroEL. In this space, the protein can fold without risk of aggregation. The correctly folded protein is then liberated from the chaperone complex via the action of ATP hydrolysis (Nishihara *et al.*, 2000; Walter & Buchner, 2002).

2.14 Measurement of HA formed from recombinant GLYAT activity

Previously used methods for estimation of HA formation only involved colorimetric analyses in which the release of coenzyme A, a product of the GLYAT reaction, was detected by the reaction of its free sulfhydryl group with the 5,5'-dithiobis (2-nitrobenzoic acid) (DTNB) colouring reagent added to the reaction mixtures (described in Section 3.2.24) (Dietz & Rubinstein, 1972; Kølvråa & Gregersen, 1986). Thus, with this method HA formation could only be assumed, and not quantified.

Using liquid chromatography in combination with tandem mass spectrometry ensures a method with high sensitivity and specificity, since the compounds of interest are separated chromatographically prior to MS/MS analysis (Nordström *et al.*, 2004). In this study the MS/MS system was employed in MRM mode. The method used for the HA quantification with the HPLC-ESI-MS/MS in this study was previously developed and validated in our laboratory by Nortje *et al.*, (2015).

As illustrated in Figure 2.5, the MS/MS system used consists of three quadrupoles (QQQ), each with a specific purpose in the analysis. When using the MS/MS system in MRM mode, the sample is first ionised by the electrospray ionisation source (ESI) to produce ions. These ions enter the Q1 cell (MS1) where the ions are sorted according to their mass-to-charge (m/z) ratios and the ion of choice (precursor ion) is selected. This ion is then transmitted to the Q2 (collision cell chamber) where collisionally-induced dissociation (CID) takes place causing the ion sample to be fragmented. After being fragmented, the sample is sent to the Q3 cell (MS2) where the fragmented ions are again sorted according to their mass-to-charge (m/z) ratios and a particular ion (product ion) is selected for detection. The detector produces a peak corresponding to each analyte which is then compared to the peak of the internal standard quantitative purposes (Banta-Wright & Steiner, 2004; Boja & Rodriguez, 2011; Cohen Freue & Borchers, 2012).

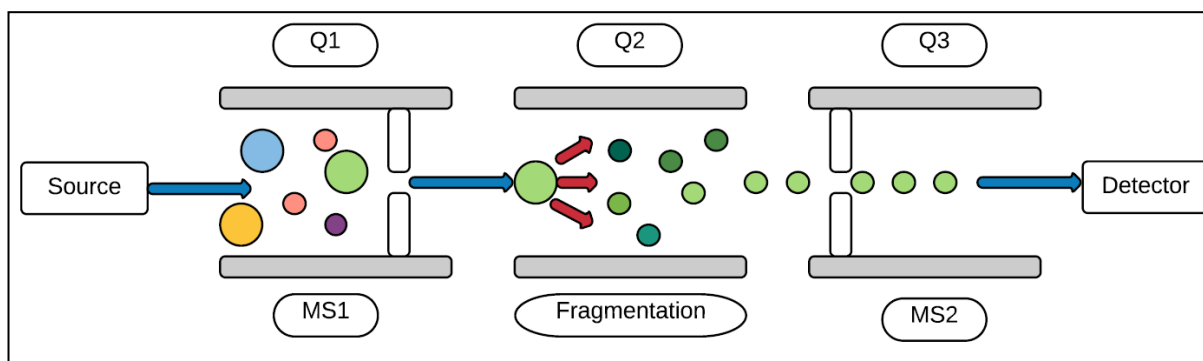


Figure 2.5: Schematic showing the components of a triple quadrupole mass spectrometer system operated in MRM mode. Q1 and Q3 represent the mass spectrometers for precursor and product ion selection. Q2 (collision cell) produces the fragment ions via CID. The green precursor ion is selected in Q1, fragmented in Q2 and the product ion (analyte of interest) is selected in Q3 and sent to the detector (Figure adapted from Boja & Rodriguez, 2011).

2.15 Stable isotope quantification

To increase the accuracy and repeatability of quantitative measurements, stable isotope-labeled analogues of the analytes of interest can be used as internal standards. These stable, isotopically labeled internal standards are chemically identical to the analytes to be measured and have similar molecular weights. For this study, a stable isotope for hippuric acid, hippuric acid-d5, was employed to serve as the internal standard (Figure 2.6). Early in the analyses, samples can be spiked with known concentrations of stable isotopes. The structural similarities will account for any sample losses during sample preparation, matrix effects as well as ion suppression or enhancement during analyses. Quantitative analysis is done by comparing the peak areas corresponding with the analyte and the peak area corresponding with the internal standard. These ratios are then used for further quantification of samples with unknown concentrations (Lebedev, 2013).

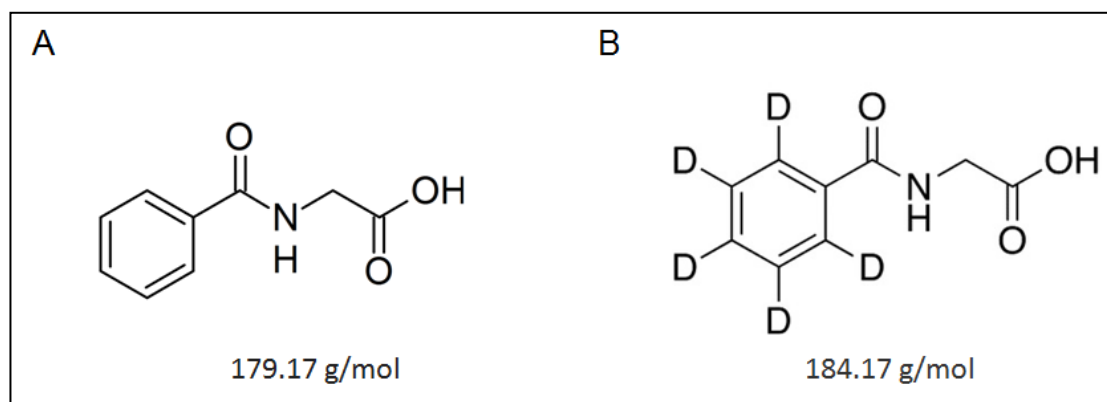


Figure 2.6: The structural formulae of A) hippuric acid, and B) hippuric acid-d₅. This figure shows the structural similarities between the analytes of interest (hippuric acid) and the stable isotope-labeled analogue (hippuric acid-d₅) along with their respective molecular weights.

2.16 Problem statement

In previous studies conducted in our laboratory only single substrate kinetic experiments were performed for six human GLYAT variants using benzoyl-coenzyme A as the only varying substrate (van der Sluis *et al.*, 2013). These non-synonymous SNP variants included the R199C, S17T, K16N; R131H; N156S; and F168L variants. This method, however, did not investigate the bi-substrate kinetic parameters for the enzymes. Thus, for this study, we aimed to determine the bi-substrate kinetics of the N156S variant. From the literature it is clear that benzoyl-coenzyme A and glycine are the preferred substrates for GLYAT, and will therefore be used in this study. The N156S variant has both the highest allele frequency and relative enzyme activity and is therefore considered to be the wild-type enzyme. Once a protocol has been established to determine the bi-substrate kinetics of the N156S variant, the remaining variants will then also be investigated.

It was hypothesised that the expression of recombinant GLYAT results in the exhaustion of coenzyme A and glycine in *E. coli* cells. This could consequently inhibit the growth of the *E. coli* cells. Therefore, as a way to prevent this issue, glycine supplementation in the growth medium of the cells was considered. Han *et al.* (2002), found that the addition of glycine to

the medium, improved the growth of *E. coli* cells, and they named it the 'glycine effect' (Han *et al.*, 2002).

The exact mechanism by which the glycine supplementation functions in the improvement of growth was not clear. Han *et al.* (2002) hypothesised that it could be ascribed to the formation of tetrahydrofolate bound one-carbon units (CH₂-THF) and/or CO₂ from glycine via the 'glycine cleavage system'. This system is induced by extracellular threonine and glycine (Bell & Turner, 1976; Ghrist & Stauffer, 1995).

The colorimetric reaction with DTNB is generally used to detect the glycine dependent release of coenzyme A (Kølvraa & Gregersen, 1986). This method, though sufficient for determining enzyme activity and estimation of kinetic parameters, cannot be used for the accurate quantification of HA formed during the human GLYAT reaction in the presence of varying substrate concentrations. We aim to address this void by employing an HPLC-MS/MS method for HA quantification, previously developed and validated, for urinary HA measurements, in our laboratory by Nortje *et al.* (2015). This method would be further improved by making use of a stable isotope of HA (HA-d5) as an internal standard, which would account for matrix effects and losses during sample preparation (Lebedev, 2013).

2.17 Research aims and objectives

Since this study consisted of three major 'legs', there were three main aims for this study. Each of these aims was allocated a chapter in this dissertation. These aims and objectives associated with each chapter are listed below.

Chapter 3

Aim: To generate a recombinant human GLYAT enzyme with a known non-synonymous SNP (N156S) by means of site-directed mutagenesis.

Objectives:

- i. To clone a codon optimised human wild-type GLYAT into a pET-32a(+) vector with an N-terminal histidine tag.
- ii. To generate the codon optimised N156S GLYAT variant via site-directed mutagenesis.
- iii. To optimise the expression medium to increase the yield of biologically active enzyme.
- iv. To purify the enzyme using nickel affinity chromatography.
- v. To optimise the storage conditions for the purified human GLYAT without a reduction in enzymatic activity.

Chapter 4

Aim: To investigate the bi-substrate kinetics of the purified N156S human GLYAT variant using benzoyl-coenzyme A and glycine as substrates.

Objectives:

- i. To determine the optimal concentration of purified N156S GLYAT protein to add to the reactions, in order to be able to identify linear regions over time.
- ii. To prepare reaction mixtures containing purified N156S GLYAT protein with varying substrate concentrations.
- iii. To identify the linear regions of the N156S GLYAT variant to calculate the initial velocities.
- iv. To plot Lineweaver-Burk plots of the kinetic data and to calculate the kinetic

parameters for both benzoyl-coenzyme A and glycine using nonlinear regression.

Chapter 5

Aim: To accurately quantify the amount of hippuric acid formed from the N156S human GLYAT variant reaction using varying benzoyl-coenzyme A and glycine concentrations by means of HPLC-MS/MS.

Objectives:

- i. To prepare reaction mixtures containing purified N156S human GLYAT protein with varying substrate concentrations.
- ii. To optimise the extraction of HA from the reactions using a liquid-liquid extraction with ethyl acetate.
- iii. To separate the samples using HPLC and to quantify HA accurately using tandem MS, run in MRM-mode, and a stable isotope, hippuric acid-d5 (HA-d5).

Chapter 3: The construction and expression of the N156S recombinant human GLYAT variant

3.1 Introduction

The aim of this part of the study was to further optimise the expression and storage conditions of the enzymatically active N156S recombinant human GLYAT (hGLYAT) enzyme. The first objective was to clone a codon optimised human wild-type GLYAT into a pET-32a(+) vector with an N-terminal histidine tag. The second objective was to generate the N156S hGLYAT variant via site-directed mutagenesis. The third objective was to optimise the expression medium to increase the yield of biologically active enzyme and to then purify the enzyme using nickel affinity chromatography. The fourth objective was to optimise the storage conditions for the purified hGLYAT without a reduction in enzymatic activity, as the long term storage of enzymes was of significant value to studying the properties of the enzymes. To assess whether the N156S hGLYAT variant remained enzymatically active, enzyme assays were performed with saturating substrate concentrations (200 mM glycine and 100 μ M benzoyl coenzyme A). A schematic summary of the basic layout of this part of the study is shown in Figure 3.1.

To simplify the flow of the text, abbreviations were created for the different constructs of the human *GLYAT* gene used in this section of the study. These abbreviated forms are summarised in Table 3.1, and will be used from here on throughout the text.

Table 3.1: Abbreviations created to refer to the different constructs of human *GLYAT* genes.

Human GLYAT	Abbreviation	Construct origin
Codon optimised wild-type human GLYAT	coWThGLYAT	Generated in this study
Codon optimised N156S human GLYAT	coN156ShGLYAT	Generated in this study
Wild-type human GLYAT	WThGLYAT	Available from previous study
N156S human GLYAT	N156ShGLYAT	Available from previous study

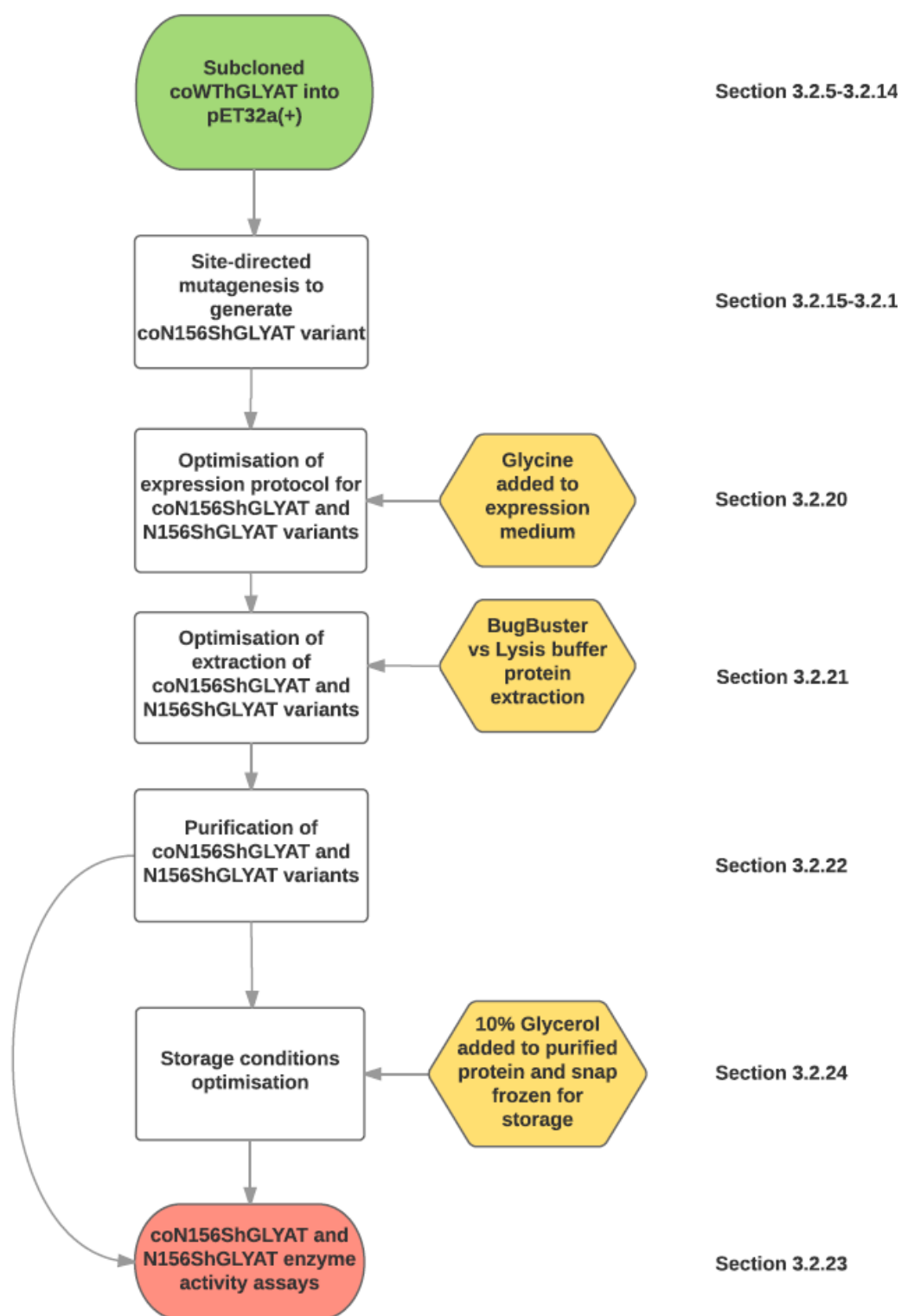


Figure 3.1: Schematic describing the overall layout of the methods described in this part of the study. The section(s) covering each step is included on the right. The yellow boxes indicate the sections of the study that were optimised.

3.2 Materials and methods

For the clarity of presentation the reagents and materials used for this study, as well as their suppliers and catalogue numbers are listed in Appendix A, and are not included in the text.

3.2.1 Preparation of stock solutions

All stock solutions were prepared with molecular biology grade water. Glycerol stock (80%, v/v) was prepared by adding glycerol to deionised water. The stock was sterilised by autoclaving for 1 hour. The stock solution was stored at room temperature. A 1 M glycine stock solution was prepared. The solution was filter sterilised using a sterile 0.8/0.2 µm Acrodisc PF Syringe Filter. Benzoyl-coenzyme A stocks were prepared in water to a final concentration of 10 mM, aliquoted and frozen at -20°C.

3.2.2 Competent cells

Commercial *E. coli* chemically competent cells and *Origami* cells with chaperone plasmids (already available in laboratory) were used in this study for cloning and expression purposes respectively (manufacturers are included in appendix A).

3.2.3 WThGLYAT

The synthesised hGLYAT (Gen Bank ID: NM_201648. 2) dsDNA fragment was purchased from GeneArt (Piscataway, NJ) for a previous study. The purchased hGLYAT sequence was cloned into a pET-32a(+) (Novagen) expression vector with *Bam*HI and *Hind*III as cloning sites. For long term storage, the vector was transformed into chemically competent JM109 cells and glycerol stocks were stored at -80°C.

Glycerol stocks of N156ShGLYAT and WThGLYAT variants in *Origami* cells were already available in the laboratory.

3.2.4 coWThGLYAT

A DNA fragment representing the codon optimised coding sequence of WThGLYAT was purchased from GeneArt. The sequence was codon optimised for bacterial expression, specifically for *Escherichia coli*. The fragment, cloned into a pMA-T vector, with SfiI and SfiI as cloning sites was purchased (shown in Figure 3.2).

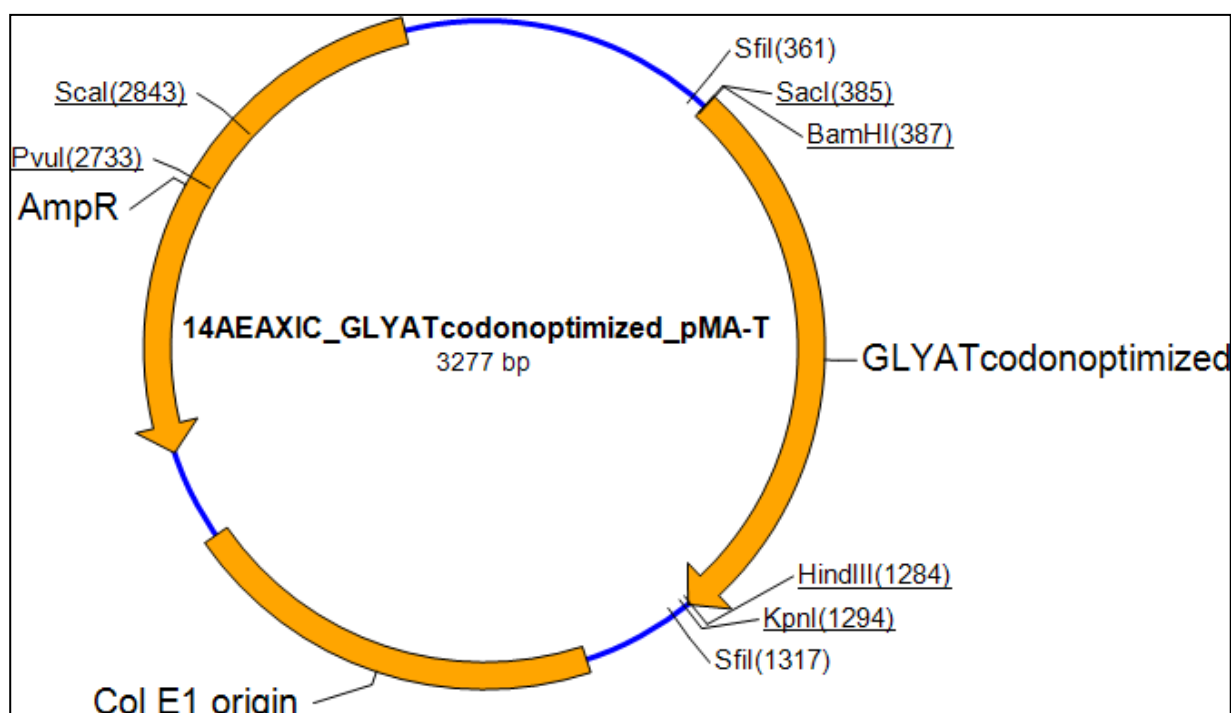


Figure 3.2: Vector map of the pMA-T vector containing the synthetic coWThGLYAT gene.

3.2.5 Transformation of chemically competent *E. coli* cells

The transformation protocol followed in this study was adapted from the method described in the transformation protocol for high efficiency ($>1 \times 10^9$ cfu/ μ g) cells. Chemically competent *E. coli* cells were used for the transformation protocol. The *E. coli* cells were thawed completely on wet ice for approximately 20 minutes, after which 40 μ l of the cells were added to a chilled culture tube along with 4 μ l of the plasmid DNA containing the

coWThGLYAT gene. The mixture was carefully stirred and incubated on ice for 30 minutes. The cells were heat shocked by placing them in a water bath set to 42°C for 45 seconds. The cells were then incubated on ice for another 2 minutes. The culture tube was removed from the ice and 960 µl SOC recovery medium was added to the culture tube. The tubes were then incubated for 1 hour at 37 °C with vigorous shaking (250 rpm).

Nutrient agar plates containing ampicillin (50 µg/ml) were plated with a dilution range of the transformed cells and were then incubated overnight at 37°C. A single colony from the overnight plates was inoculated in 100 ml Luria-Bertani (LB) medium (1% tryptone (w/v), 0.5% yeast extract, (w/v), 1% NaCl (w/v)), fortified with 50 µg/ml ampicillin, in a 250 ml Erlenmeyer flask and incubated overnight, shaking at 37°C at 200 rpm. To allow for long term storage, glycerol stocks of the transformed cells were prepared by mixing 800 µl of the cells with 200 µl sterile 80% (v/v) glycerol in a microcentrifuge tube. The stocks were then stored at -80 °C.

3.2.6 Plasmid extraction (Midi-preparation)

The PureYield Plasmid Midiprep System (Promega) was used to extract the plasmids from the *E. coli* cells into which the plasmids were previously transformed. This system purifies the plasmid DNA using silica-membrane column technology. The pET-32a(+)/WThGLYAT and pMA-T/coWThGLYAT plasmids were extracted.

Two bacterial cell cultures were grown overnight. A 100 ml *E. coli* bacterial cell culture containing the pMA-T/coWThGLYAT as well as a 100 ml *Origami* bacterial cell culture containing the pET-32a(+)/WThGLYAT was prepared in LB medium with 50 µg/ml ampicillin, for selection. The cultures were incubated at 37°C at 200 rpm for approximately 18 hours, or until an O.D.₆₀₀ reading of 2-4 was reached. The cultures were then centrifuged (5000 x g, 10 minutes, room temperature) and the supernatants were discarded before the tubes were drained on paper towels to rid the pellets of excess media.

The cell pellets were resuspended in 3 ml Cell Resuspension Solution by pipetting the cell pellet up and down. A further 3 ml Cell Lysis Solution was added to the cells, and mixed by inverting the tubes several times and incubating them at room temperature for 3 minutes. The lysed cells were then mixed with 5 ml Neutralization Solution and mixed by inverting the tubes for 10 minutes. The lysates were then centrifuged ($15\,000 \times g$, 15 minutes, room temperature) to pellet the cellular debris. PureYield Clearing Columns were then used for further debris removal.

Column stacks were assembled by placing PureYield Clearing Columns onto PureYield Binding Columns and placed onto a vacuum manifold. The cleared lysates were then poured into the PureYield Clearing Columns and vacuum was administered until all the liquid had passed through both columns. The PureYield Clearing Columns removed the remaining cellular debris, while the plasmid DNA bound to the binding membranes of the PureYield Binding Columns.

The PureYield Binding Columns were then filled with 5 ml Endotoxin Removal Wash, and vacuum was applied to drain the solution through the columns. The Endotoxin Removal Wash is specifically designed to rid plasmid DNA of endotoxin contaminants, RNA, and proteins, thus providing pure plasmid DNA yields, suitable for downstream applications. The columns were then washed by addition of 20 ml Column Wash Solution and applying vacuum to draw the solution through the columns. To dry the binding membranes, a vacuum was applied to the emptied columns for an additional minute, or until the binding membranes appeared dry and the ethanol odour had become undetectable.

Eluator Vacuum Elution Devices were used for the elution of the plasmid DNA by vacuum. The devices were assembled by capping the device with the provided tube and the PureYield Binding Columns were then fitted into the top of the devices. For final elution, 500 μ l Nuclease-Free Water was added to the binding columns and left for 2 minutes before applying maximum vacuum for approximately 1 minute or until all the liquid had passed through the columns and into the microcentrifuge tubes.

3.2.7 Agarose gels

Agarose gel electrophoresis was used for the separation and visualisation of DNA samples. The DNA samples were separated on 1 % (w,v) agarose gels, with 1 x TAE buffer (40 mM Tris, 20 mM Acetic acid, and 1 mM EDTA, pH 8.5). The electrophoresis runs were performed at a constant voltage of 90V for 70 minutes, unless otherwise specified. The Syngene G:BOX F3 Fluorescence gel documentation system was used in conjunction with the GeneSys image acquisition software program (Syngene, Cambridge, UK). Ethidium bromide (EtBr) was used for visualisation of the DNA fragments. The O'GeneRuler 1kb DNA Ladder Mix (100-10,000 bp) was used as a size marker.

3.2.8 Sodium dodecyl sulfate polyacrylamide gels

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was used to monitor protein expression and purification steps. The protocol used in this study was adapted from the method described by Sambrook and Russell (2001) and Laemmli (1970).

The running gels (12.5%) were prepared by mixing 6.25 ml of the monomer solution (30% acrylamide and 2.7% bisacrylamide), 3.75 ml of the 4X running gel buffer (1.5 M TrisCl, pH 8.8), 150 μ l of 10% (w,v) SDS, 100 μ l of 10% (w,v) ammonium persulfate and 4.75 ml nuclease-free H₂O in an Erlenmeyer flask. Prior to use, TEMED was added to a final concentration of 0.05% (v,v). This running gel solution was then poured between the vertical glass plates of the gel casting apparatus, up to approximately 3 cm from the top of the glass plates, and overlaid with H₂O. The gel was allowed to set at room temperature for approximately 1 hour (or until an interface between the gel solution and the water layer could be seen). The water layer was then absorbed by sliding filter paper between the plates.

The stacking gels (4%) were prepared by adding 940 μ l of the monomer solution (30% acrylamide and 2.7% bisacrylamide), 1.75 ml of the 4X stacking gel buffer (500m M TrisCl,

pH 6.8), 70 µl of 10% SDS, 35 µl of 10% ammonium persulfate and 4.3 ml nuclease-free H₂O in an Erlenmeyer flask. Prior to use, TEMED was added to a final concentration of 0.1%. This stacking gel solution was then poured between the vertical glass plates on top of the running gel up to the notch of the glass plates, and a comb was inserted to form the wells. The stacking gel was allowed to set for 30 minutes.

Protein samples were prepared by mixing 15 µl with an equal amount of 4X DualColor Protein Loading Buffer, 2 µl DTT (reducing agent) and 13 µl nuclease-free water. The samples were boiled for 5 minutes at 98°C, after which they were immediately used for the SDS-PAGE analysis. Unless otherwise specified, 15 µl was loaded onto the gel for electrophoresis. PageRuler Plus Prestained Protein Ladder (10 kDa to 250 kDa) was used as a molecular size marker.

The gel was then placed in the gel chamber with 1X Tris-glycine-SDS (TGS) buffer (25 mM Tris, 192 mM glycine and 0.1% SDS, pH 8.6). For electrophoresis, a constant current of 30 mA was applied for 60 minutes. Thereafter, the gel was removed from the glass plates and, for the visualisation of the protein bands, placed into a container with Coomassie gel staining solution for approximately 2 hours, with shaking at 50 rpm. For the destaining of the gel, it was then removed from the Coomassie gel staining solution and rinsed with water to remove excess stain. The gel was resubmerged in a container filled with destain solution (50% (v/v) methanol, 10% (v/v) acetic acid) with gentle shaking until the blue stain was removed from the gel. The gel images were digitised by scanning.

3.2.9 Restriction endonuclease digests of plasmids

To be able to isolate the coWThGLYAT insert and subclone it into the pET-32a(+) plasmid, the pET-32a(+) and pMA-T plasmids were separated from the WThGLYAT and coWThGLYAT inserts using restriction enzymes. After gel electrophoresis was done for verification of the extracted plasmids, the plasmids were double digested with *Bam*HI and *Hind*III restriction enzymes to excise the WThGLYAT and coWThGLYAT inserts from the respective plasmids.

Table 3.2 describes briefly the restriction enzymes and conditions used for this study. The restriction reactions were set up according to the recommended conditions as specified by the Fermentas DoubleDigest software program.

Table 3.2: The restriction enzymes and their buffers that were used for this study.

Restriction enzyme	Recognition sequence (5' → 3')	Buffer	Incubation temperature (°C)
BamHI	G↓GA TCC	Buffer BamHI	37
HindIII	A↓AG CTT	Buffer BamHI	37

The DoubleDigest tool is designed to easily be able to determine the optimal reaction conditions for each double digest reaction. This program accurately estimates the optimal incubation conditions; enzyme concentrations, enzymatic reaction buffer, and whether further additives are needed to set up a successful double digest reaction without star activity. The following conditions were calculated for double digestion with *Bam*HI and *Hind*III: (i) To use 1X Buffer *Bam*HI as the enzymatic reaction buffer, a 2-fold excess of *Hind*III was recommended and (ii) the optimal incubation temperature was determined as 37°C (Table 3.2).

According to these guidelines, the reactions were then set up for both the pET-32a(+)/WThGLYAT and the pMA-T/coWThGLYAT plasmids in two 1.5 ml microcentrifuge tubes respectively: 45 µl plasmid DNA (100 ng), 11 µl 1X *Bam*HI reaction buffer, 4 µl *Bam*HI restriction enzyme, 8 µl *Hind*III restriction enzyme, and 42 µl ddH₂O to a final volume of 110 µl. The reactions were then incubated overnight in a temperature controlled water bath at 37°C. Agarose gel electrophoresis analyses were then performed to separate the digested plasmids from the inserts. An agarose gel was prepared as described in Section 3.2.7.

3.2.10 Gel extraction

Agarose gel electrophoresis was used, as described above, for the separation of the DNA inserts from the digested plasmids. The aim of this separation was to be able to finally purify the pET-32a(+) plasmid and the coWThGLYAT insert from the gel, and then to subclone the coWThGLYAT insert into the pET-32a(+) vector. To extract the bands, containing the desired fragments from the agarose gel, the sections of the gel containing the bands were excised with a sterile blade and placed into pre-weighed microcentrifuge tubes. The DNA was then recovered using spin column technology.

The NucleoSpin Gel and PCR Clean-Up kit (Macherey-Nagel, Düren, Germany; #1502/001) was used for the gel extraction, according to the manufacturer's gel extraction instructions. Briefly, the gel slices were weighed and 200 µl buffer (NTI) per 100 mg of gel was added to each excised fragment. The samples were then incubated for 10 minutes at 50°C and vortexed briefly every 2 minutes to solubilise the agarose. This solution was then transferred to a spin column and centrifuged (11 000 x *g*, 30 seconds, room temperature) to allow the solution to flow through the silica membrane. The silica membrane binds DNA in the presence of the NTI buffer which contains chaotropic salts. The membrane was then washed with 700 µl of buffer NT by centrifugation (11 000 x *g*, 30 seconds, room temperature). The wash step was carried out twice to ensure minimal carry-over of chaotropic salts. The membrane was then dried by centrifugation (11 000 x *g*, 2 minutes, room temperature). The DNA remained bound to the membranes and flow-through was discarded after the initial binding, wash and drying steps. For DNA elution, the membranes were wet with 15 µl Nuclease-Free water, incubated at room temperature for 2 minutes and centrifuged (11 000 x *g*, 1 minute, room temperature). The elution step was repeated, to ensure maximum recovery.

Electrophoretic analyses were performed to verify the extraction of the pET-32a(+) plasmid and coWThGLYAT inserts from the agarose gel. An agarose gel was prepared as described in Section 3.2.7.

3.2.11 DNA quantification and purity determination

Before the ligation step could be performed, it was first important to accurately determine the concentration and purity of the extracted coWThGLYAT insert and linearised pET-32a(+) plasmid, seeing that ligation is very dependent on the insert to vector ratios.

DNA concentration was determined by means of spectrophotometry using the NanoDrop ID-1000 spectrophotometer system (Thermo Scientific). This method relies on the principle of the Beer Lambert Law in which the concentration of a molecule can be calculated from the amount of light it absorbs. For double-stranded DNA, a change of one absorbance unit at 260 nm corresponds to a concentration of 50 µg/ml. The DNA purity was also determined with the above mentioned system by assessing the A260/A280 ratio. For a pure DNA sample, an A260/A280 value of approximately 1.8 is expected. This ratio can reveal whether a sample is contaminated with proteins, as the proteins will absorb strongly at 280 nm, thus causing the ratio to fall below the desired 1.8 value.

After the concentration of the vector and insert samples were determined, calculations were done to determine the amount of vector and insert to add to the ligation reactions. Equation 3.1 was used for these calculations. Equation 3.2 shows the values substituted as it pertained to this study.

Equation 3.1: The equation used to calculate the ratios of insert to vector for ligation reactions.

$$\text{Insert (ng)} = \left(\frac{\text{Vector (ng)} \times \text{Kb size of insert}}{\text{Kb size of vector}} \right) \times \text{molar ratio of } \left(\frac{\text{insert}}{\text{vector}} \right)$$

Equation 3.2: The equation used to calculate the ratios of insert to vector for ligation reactions, with the values substituted as it pertained to this study.

$$\text{Insert (ng)} = \left(\frac{\text{Vector (ng)} \times 0.9 \text{ Kb}}{5.9 \text{ Kb}} \right) \times \left(\frac{3}{1} \right)$$

3.2.12 Ligation reaction

The coWThGLYAT insert and linearised pET-32a(+) vector were then ligated using the Rapid DNA Ligation Kit (Thermo Scientific). After the concentration and purity of the inserts and plasmid DNA were determined and the necessary ratio calculations were done, the ligation reactions could be set up according to the specifications outlined in the kit manual. The reactions were set up as follows: 7 µl of the linearised pET-32a(+) vector, 1 µl WThGLYAT insert, 4 µl 5X Rapid ligation buffer, 1 µl T4 DNA ligase and 7 µl nuclease-free water was added to a microcentrifuge tube and briefly vortex mixed and centrifuged to collect the drops. The mixture was then incubated for 5 minutes at 22°C and stored at 4°C until used for transformation.

A self-ligation control reaction was also set up, as follows: 7 µl of the linearised vector DNA, 4 µl 5X Rapid ligation buffer, 1 µl T4 DNA ligase and 8 µl nuclease-free water was added to a microcentrifuge tube. The mixture was vortex mixed and centrifuged, incubated and stored at 4°C until used for transformation.

3.2.13 Transformation of chemically competent *E. coli* cells

The ligation products were then transformed into chemically competent *E. coli* cells. Three transformation reactions were performed, namely: (i) transformation of the ligation reaction with the pET-32a(+)/coWThGLYAT vector, (ii) a self-ligation control reaction, and (iii) the pMA-T/coWThGLYAT vector which served as the transformation control. The transformation protocol described in Section 3.2.5 was followed. Nutrient agar plates were prepared with 50 µg/ml ampicillin for selection, and 100 µl of the transformation reactions were plated out onto the agar plates in triplicate, with the exception of the self-ligation transformation reaction, of which 200 µl was plated out in triplicate. The plates were incubated overnight at 37°C.

3.2.14 Provisional verification of colonies of transformed bacteria

Ten colonies from the transformed cells were randomly chosen and screened to verify the presence of the desired coWThGLYAT insert using agarose gel electrophoresis. Erlenmeyer flasks containing 10 ml LB medium (containing 50 µg/ml ampicillin) were each inoculated with one of the chosen colonies and incubated overnight at 37°C, with shaking at 200 rpm. Prior to the plasmid extraction, glycerol stocks were prepared of each of the transformed cells for long term storage. This was done by mixing 800 µl of the cells with 200 µl sterile 80% (v,v) glycerol in a microcentrifuge tube and placing the stocks in -80°C for storage. The plasmid DNA was then isolated from each of the overnight cultures by means of the PureYield Plasmid Midiprep System as described in Section 3.2.6.

An agarose gel electrophoresis analysis was performed. Two 1% (v,v) agarose gels were prepared as described in Section 3.2.7. The first five samples were loaded onto one gel and the further five samples on a second gel. From these gel results, it could roughly be estimated whether the samples contained the plasmid and the desired insert, by analysing the estimated size of the fragments in comparison with the bands from the control reactions.

3.2.15 DNA sequence determination and data analysis

Plasmids corresponding to the estimated size were then sent to the Central Analytical Facility (CAF) of the University of Stellenbosch for DNA sequence determination, using the Sanger method. This was done to confirm that the *coN156ShGLYAT* gene was successfully inserted into the pET-32(a)+ plasmid in the correct orientation and without any additional mutations being introduced.

Before sequencing the plasmid DNA, the concentration and purity of each of the isolated plasmids were determined using a NanoDrop spectrophotometer system. DNA concentrations of more than 100 ng/µl were required in order for the samples to be sent for

Sanger sequencing. Initially, some of the samples' DNA concentrations were slightly below this value, and a concentration step was carried out with a SpeedVac concentrator (Cat. No. ISS110FS-115) (Thermo Scientific, Waltham, MA, USA).

The DNA sequencing chromatograms provided by the CAF were analysed using FinchTV (version 1.40) (www.geospiza.com/finchtv). ClustalX was used for the alignment of the sequences to reference sequences.

3.2.16 Generation of the N156S hGLYAT variant by Site-directed mutagenesis

For this study, the N156S variant of the *coWThGLYAT* had to be generated. The N156S SNP was introduced into the *coWThGLYAT* gene in the pET-32a(+) plasmid. This alteration caused an amino acid change from an asparagine to a serine residue, and was achieved by site-directed mutagenesis in which a point mutation was introduced in the insert by amplification of the plasmid DNA through polymerase chain reaction (PCR). This method involves using HPLC purified primers, one of which contains a mismatched base which differs from the wild-type sequence. This base is then incorporated into the PCR product (Hemsley *et al.*, 1989). This method is advantageous, since mutations can be introduced at any desired site. The Phusion Site-Directed Mutagenesis kit was used for this step.

3.2.16.1 Primer design

In this study, two primers, one of which contained the desired mutation, were designed to anneal 'back to back' to the target plasmid with the 5' ends apposing and the 3' ends directed for extension in opposite directions around the plasmid (see Figure 3.3). Table 3.3 shows the primers that were employed. The primers were purchased from Integrated DNA Technologies (IDT, Coralville, Iowa).

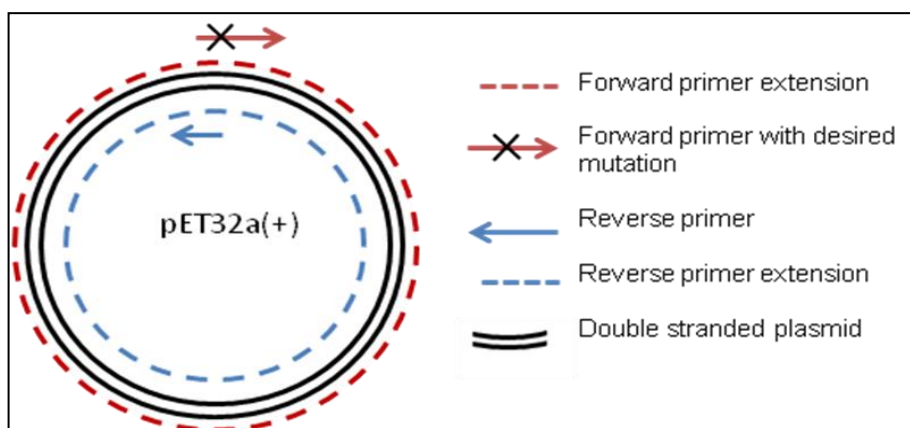


Figure 3.3: Schematic representation of site-directed mutagenesis. In this figure, a plasmid is shown along with two primers, one of which contains the mutation. When the primers anneal back to back to the target plasmid, a linearised plasmid, containing the desired point mutation, is formed.

Table 3.3: Oligonucleotide primers used in this study for Site-directed mutagenesis to generate the coN156ShGLYAT variant.

Name of primer	Sequence (5' → 3')	Length (bp)	Melting temperature (T _m)
hGLYATApF	CTG AGC CCG <u>AGT</u> GGT GGT AAA C	22	69.1 °C
hGLYATApR	AAT TTT GCT TTT CAG CAG AAA CG	23	65.4 °C

3.2.16.2 Polymerase chain reaction (PCR) reaction mixture

The PCR reaction mixtures were set up according to the guidelines provided in the Phusion Site-Directed Mutagenesis kit: In PCR tubes, 23.8 µl nuclease-free water, 10 µl 5X Phusion HF Buffer, 4 µl dNTP's (final concentration 200 µM each), 2.5 µl of 10 µM forward primer (final concentration 0.5 µM), 2.5 µl of 10 µM reverse primer (final concentration 0.5 µM), 3.7 µl of 67 ng/µl plasmid DNA (final concentration 5 ng/µl), 1 µl Phusion Hot Start II DNA Polymerase, and 2.5 µl DMSO (to a final concentration of 5%).

To serve as a positive control, a pET-32a(+)/WThGLYAT vector was used, since a site-directed mutagenesis protocol for this variant had previously been developed in our lab by van der Sluis *et al.* (2013). The primers used were previously designed by Dr. R. van der Sluis and were purchased from IDT (Coralville, Iowa). Table 3.4 shows the primers that were used for this control reaction.

Table 3.4: Oligonucleotide primers used in a previous study, which were used for the positive control reaction to generate the coN156ShGLYAT variant via site-directed mutagenesis.

Name of primer	Sequence (5' → 3')	Length (bp)	Melting temperature (T _m)
N156SFor	AAG ATT TTA TCT CCC AGT GGT GGC AAA CCC AAG	33	77 °C
N156SRev	TGA TTT CAG CAG GAA AGG AGT	21	60 °C

The reaction mixture for the positive control was set up in a PCR tube as follows: 22.1 µl nuclease-free water, 10 µl 5X Phusion HF Buffer, 4 µl dNTP's, 2.5 µl of 10 µM forward primer (final concentration 0.5 µM), 2.5 µl of 10 µM reverse primer (final concentration 0.5 µM), 5.4 µl of 46 ng/µl plasmid DNA (final concentration 5 ng/µl), 1 µl Phusion Hot Start II DNA Polymerase, and 2.5 µl DMSO.

A negative control reaction was also set up, identical to the reactions described above, but in which the plasmid DNA was omitted and replaced with an additional 5.4 µl of nuclease-free water.

3.2.16.3 Optimisation of PCR conditions

After the preparation of the PCR reaction samples, the PCR reaction could be set up along the recommended guidelines provided in the Phusion Site-Directed Mutagenesis kit manual.

Typically, the optimal annealing temperature varies between reactions. It was therefore necessary for some optimisation regarding the annealing temperatures of this specific protocol. For this purpose, additional reaction mixtures were prepared and exposed to different annealing temperatures during the PCR amplification. The annealing temperatures were varied between 65-72°C for five test reactions, thus exposing each of the reaction mixtures to annealing temperatures of 65.2°C, 67.2°C, 69.0°C, 70.5°C, and 71.6°C respectively. Table 3.5 shows the specific cycling conditions that were employed for the PCR reactions. An Eppendorf thermal cycler was used to perform thermal cycling.

Table 3.5: Cycling conditions for the site-directed mutagenesis reactions.

Cycle step	Temperature	Time	Number of cycles
Initial denaturation	98°C	1 min	1
Denaturation	98°C	10 sec	35
Annealing	65-72°C	30 sec	
Extension	72°C	60 sec	
Final extention	72°C	2 min	1
Hold	4°C	∞	∞

It was, however, not necessary for optimisation of the annealing temperatures of the control reactions, as the positive control reaction's annealing temperature was already determined as 72°C. In the case of the negative control, an annealing temperature of 72°C was also used. Table 3.6 shows the specific cycling conditions that were employed for the

control PCR reactions.

An agarose gel electrophoresis analysis was then performed to verify the amplification of the mutant plasmid. An agarose gel was prepared as described in Section 3.2.7.

Table 3.6: Cycling conditions for the control mutagenesis reactions.

Cycle step	Temperature	Time	Number of cycles
Initial denaturation	98°C	1 min	1
Denaturation	98°C	10 sec	35
Annealing	72°C	30 sec	
Extension	72°C	60 sec	
Final extention	72°C	2 min	1
Hold	4°C	∞	∞

3.2.16.4 Gel extraction of amplified mutant plasmid DNA

After electrophoresis identified the sample with the most optimal annealing temperature, the remaining 40 μ l of that amplified plasmid DNA (mixed with 8 μ l loading dye) was loaded onto an agarose gel once again. The aim of this step was to purify the amplified pET-32a(+)/coN156ShGLYAT plasmids, and then to extract the bands containing the desired fragments from the agarose gel. The section of the gel containing the band of the amplicon was then excised with a sterile blade and placed into pre-weighed microcentrifuge tubes. The plasmid DNA was then recovered using the NucleoSpin Gel and PCR Clean-Up kit as described in Section 3.2.10.

3.2.16.5 Recircularisation of mutated PCR products

Before the modified amplicons could be transformed, the linearised plasmids had to be ligated. For this recircularisation, the Rapid DNA Ligation Kit described in section 3.2.12 was again used with some adjustments. The concentration of the mutation containing amplicons first had to be determined before the ligation reaction could be set up. This was done with the NanoDrop Spectrophotometer as described in Section 3.2.11.

The reactions were then set up as follows: 8 µl of the amplicon DNA (total amount 50 ng), 10 µl 5X Rapid ligation buffer, 1 µl T4 DNA ligase and 31 µl nuclease-free water was added to a microcentrifuge tube and briefly vortexed and centrifuged to collect the drops. The mixture was then incubated for 5 minutes at 22°C and stored at 4°C until used for transformation.

3.2.17 Transformation of chemically competent *E. coli* cells with modified amplicons

The recircularised plasmids were then transformed into chemically competent *E. coli* cells. Two transformation reactions were performed, one for the recircularised plasmid and one with the linearised plasmid which served as the transformation control. Nutrient agar plates were prepared with ampicillin to a final concentration of 50 µg/ml. The transformation protocol as described in Section 3.2.5 was used. After the transformation, 100 µl of each of the transformation reactions were plated out onto three nutrient agar plates, and incubated overnight at 37°C.

3.2.18 Colony selection

Erlenmeyer flasks containing 50 ml of LB medium (containing 100 µg/ml ampicillin) were each inoculated with one of ten colonies that were picked from the overnight plates. The cultures were incubated overnight at 37°C with shaking at 200 rpm. The plasmid DNA was then isolated from each of the ten overnight cultures using the PureYield Plasmid Midiprep

System as described in Section 3.2.6.

3.2.19 DNA sequence determination and data analysis

The ten samples were then Sanger sequenced to confirm that the pET-32a(+) plasmid contained the *coWThGLYAT* gene in the correct orientation and that the N156S SNP had been successfully introduced. Samples were sent to the Central Analytical Facility (CAF) of the University of Stellenbosch for Sanger sequencing. The DNA sequencing chromatograms provided by the CAF were analysed using FinchTV (version 1.40) (www.geospiza.com/finchtv). ClustalX was used for the alignment of the sequences to reference sequences.

3.2.20 Co-transformation of chemically competent *Origami* cells

After confirmation of the introduction of the N156S mutation in the hGLYAT codon optimised insert, the purified plasmids were then transformed into the chemically competent *Origami* cells containing the pGro7 chaperone plasmid (Figure 3.4) for co-expression of the GroEL-GroES chaperone team (Takara, Madison, WI).

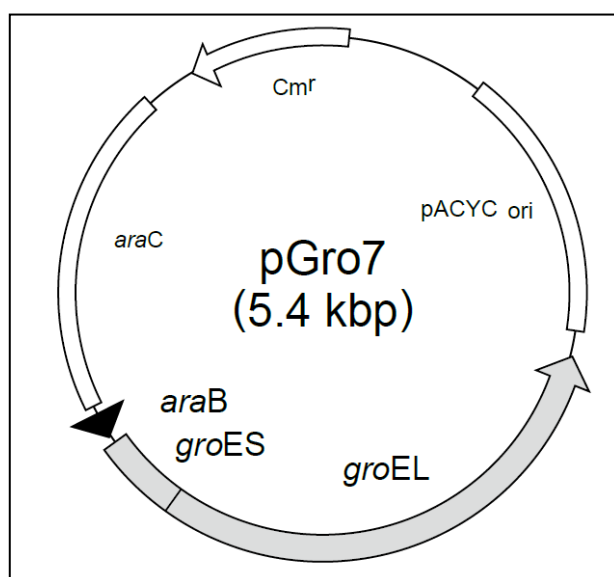


Figure 3.4: Vector map of the pGro7 chaperone plasmid.

Briefly, the competent *Origami* cells, already containing the pGro7 plasmids, were thawed completely on wet ice for approximately 20 minutes, after which 200 µl of the cells were added to a chilled culture tube. 100 ng of the plasmid DNA (1 µl) was then added to the cells and stirred with a pipet tip before incubating on ice for 1 hour. 800 µl transformation storage buffer (TSBG) (1.6% Peptone (w/v), 1.0% Yeast Extract (w/v), 0.5% NaCl (w/v), 10.0% PEG (w/v), 5.0% DMSO (w/v), 1M MgCl₂, 1M MgSO₄, 20mM glucose) was then added to the cells, after which the samples were placed in an incubator at 37°C, with shaking at 225 rpm for 1 hour.

Nutrient agar plates were prepared with ampicillin added to a final concentration of 50 µg/ml and chloramphenicol added to a final concentration of 20 µg/ml. After the transformation, 100 µl of each of the transformed cells were plated out onto the plates. The plates were incubated overnight at 37°C. In a 250 ml Erlenmeyer flask, 100 ml LB medium (with 50 µg/ml ampicillin and 20 µg/ml chloramphenicol for selection) was then inoculated with a single colony from the overnight plates containing the co-transformed colonies and incubated overnight at 37°C with vigorous shaking at 225 rpm.

To allow for long term storage, glycerol stocks were prepared from the transformed cells by mixing 800 µl of the cells with 200 µl sterile 80% (v/v) glycerol in a microcentrifuge tube. The stocks were then stored at -80 °C. These were then used for inoculation of cultures for the expression steps.

3.2.21 Expression of coN156ShGLYAT and N156ShGLYAT proteins

The recombinant hGLYAT proteins were expressed with an N-terminal His-tag to facilitate purification. The expression of the proteins was performed as described by van der Sluis *et al.* (2013). The protocol was summarised in Section 2.13. As mentioned, it was hypothesised that the expression of the recombinant hGLYAT resulted in the depletion of coenzyme A and glycine in the *E. coli* cells, which lead to very poor growth of the *E. coli* cells. The protocol was then optimised by the addition of 0.5% glycine to the expression medium

to boost the growth of the *E. coli* cells (Ghrist & Stauffer, 1995; Han *et al.*, 2002).

3.2.22 Extraction of expressed coN156ShGLYAT and N156ShGLYAT proteins

After the expression of the recombinant hGLYAT proteins, the proteins had to be extracted from the cultured cells. Two methods for protein extraction were compared. The first method described involves the BugBuster Protein Extraction Reagent. The other method, developed by van der Sluis *et al.* (2013) specifically for hGLYAT protein extraction, involves a lysis buffer.

3.2.22.1 BugBuster protein extraction

For protein extraction using the BugBuster Protein Extraction Reagent, the protocol described in the manufacturer's manual was used, with minor modifications. Briefly, 100 ml of the expression medium was used and the cells were harvested by centrifugation (2000 x g, 10 minutes, room temperature) in a pre-weighed 50 ml Falcon tube. After draining the cell pellet, the wet weight of the pellet was determined. For each gram of wet cell mass, 5 ml room temperature BugBuster Reagent was added to the tube. Lysonase Bioprocessing Reagent, a ready-to-use mix of Lysozyme Solution and Benzonase Nuclease, was then added to the cell solution (10 µl per gram wet cell paste) and mixed thoroughly by pipetting. A 100 µl sample was taken from the lysate as a total protein fraction. The lysate was then centrifuged (15 000 x g, 20 minutes, 4 °C) for the removal of the insoluble materials and cell debris. The supernatant was transferred to a fresh tube and saved for further SDS-PAGE analysis, which was used to confirm the successful extraction of the recombinant hGLYAT protein. The SDS-PAGE protocol is described in Section 3.2.8.

3.2.22.2 Lysis buffer protein extraction

The protein extraction method followed was adapted from van der Sluis *et al.* (2013). A 100 ml lysis buffer consisting of 50 mM NaH₂PO₄ and 300 mM NaCl was prepared.. The pH was then set to pH 8 and filter sterilised using a sterile 0.8/0.2 µm Acrodisc PF Syringe Filter.

After sterilisation 10% (v/v) Glycerol, 1% (v/v) Triton X-100 and 10ul Lysonase (for every 10 ml buffer) was added. After the 24 hour incubation period at 28°C, 100 ml of the expression medium was used and the cells were harvested by centrifugation (4000 x *g*, 15 minutes, room temperature). The pelleted cells were resuspended in 10 ml lysis buffer and incubated for 10 minutes at room temperature. The lysate was passed through a 22G needle five times for further cell disruption. The lysate was then cleared of cell debris by another centrifugation step (14 000 x *g*, 25 minutes, room temperature) and the supernatant, which served as the soluble fraction, was saved for further SDS-PAGE analyses. This step was once again included to serve as a verification step for the successful extraction of the expressed recombinant hGLYAT proteins. SDS-PAGE analyses were performed for verification of protein expression and extraction procedures. The SDS-PAGE gels and buffers were prepared as stated in Section 3.2.8.

3.2.23 Histidine-tagged protein purification

A DNA sequence encoding for six to nine histidine residues, is usually added in vectors for the expression of recombinant proteins. The expressed proteins then possess poly-His tags fused to their N- or C-termini, which then aids in the purification or detection of the recombinant protein, without the need for protein-specific probes or antibodies. Immobilised metal-affinity chromatography (IMAC) methods are generally applied for the purification of recombinant proteins containing affinity tags such as histidine-tags. In this study, nickel-affinity chromatography was employed for the purification purposes, as the chain of histidine residues have an affinity for the immobilised nickel ions (Bornhorst & Falke, 2000). For purification of the recombinant protein, a Protino Ni-TED 2000 column kit was used. The cleared lysate was passed through the column according to the protocol described in the product manual. A summarised version of the protocol is described below.

First, an equilibration step was performed by adding 4 ml LEW buffer (Macherey-Nagel Inc., Bethlehem, PA) to the column and then allowing the solution to pass completely through the column. The cleared lysate was then poured through the column for the binding step. A

wash step was then done with 2 X 5 ml LEW buffer, to rid the membrane of any contaminants. Finally, the proteins were eluted with 3 ml EB buffer (Macherey-Nagel Inc., Bethlehem, PA). From each of the above mentioned steps, 100 µl samples were again taken to assess the efficiency of the purification process and to investigate the presence of the expressed protein in these fractions by SDS-PAGE analysis. The expected sizes of the hGLYAT protein bands on the gels are approximately 37 kDa when calculating from the length of the sequence which is 903 bp. This results in an amino acid sequence of 301 amino acids, which is approximately 37 kDa in size. However, it is important to note that the recombinant GLYAT protein is expressed with an N-terminal fusion tag of approximately 29 kDa containing the Trx-tag followed by a 6X His-tag and an S-tag. This will result in the final size of the protein band seen on the gel, to be approximately 56 kDa. Figure 3.5 shows a vector map of the pET-32a(+) plasmid along with the sequence map indicating the sequences for the fusion tag. SDS-PAGE analyses were performed to assess the success of all the protein purification steps after expression. The SDS-PAGE gels were prepared as described in Section 3.2.8.

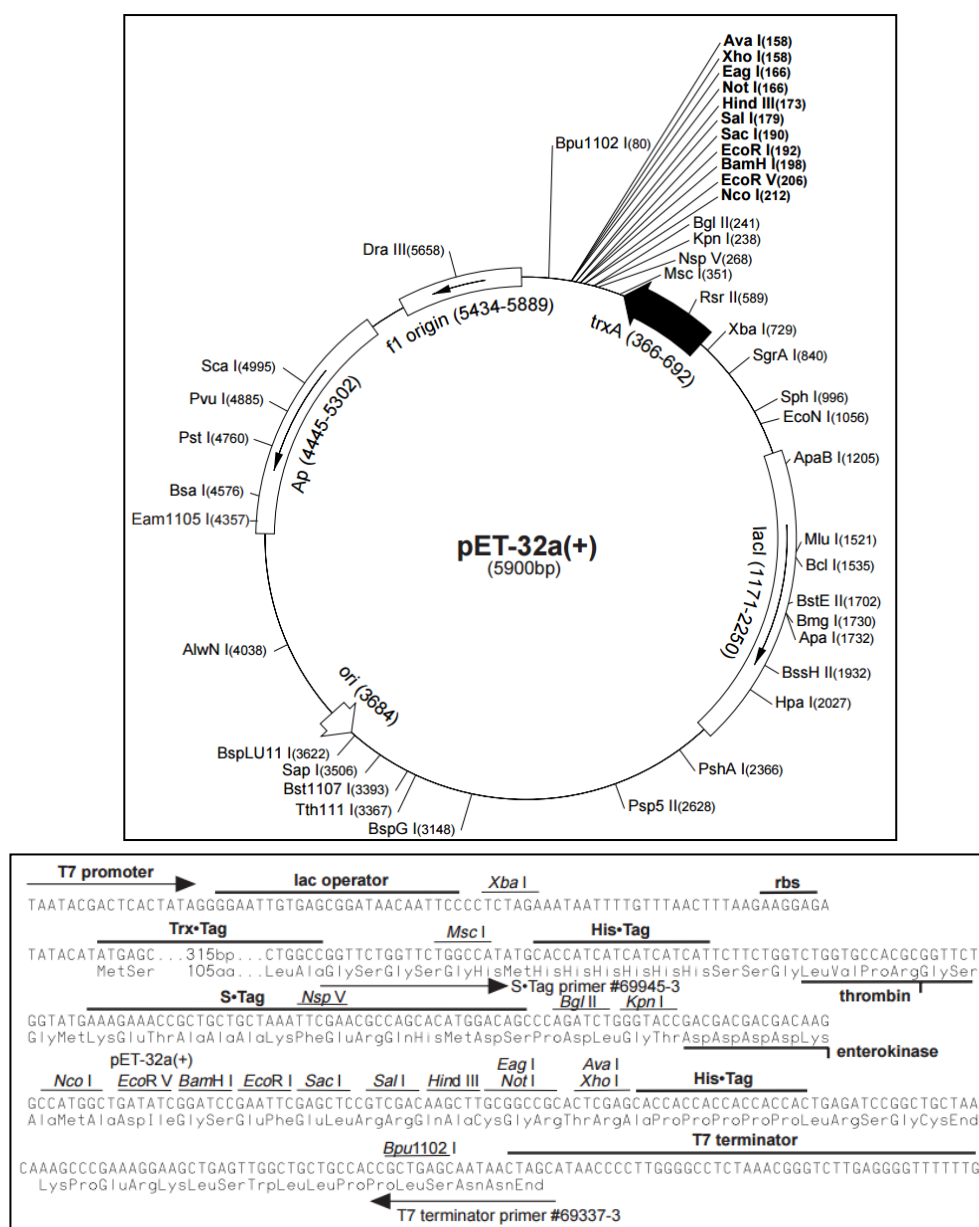


Figure 3.5: Vector map of the pET-32a(+) plasmid. The sequence map indicates the region containing the T7 promoter; the Trx-tag, 6X His-tag and S-tag; the T7 terminator site and the multiple cloning site.

3.2.24 hGLYAT enzyme activity assays

Enzyme activity assays were performed after each expression step, to verify hGLYAT activity. For this purpose, a colorimetric reaction was used to detect coenzyme A, a product of the

GLYAT reaction (Kølvraa & Gregersen, 1986). Into each reaction mixture, 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) was added for a colorimetric reaction. A colour change from colourless to deep yellow was indicative of GLYAT activity. This colour change is due to the reaction between the DTNB with the free sulfhydryl group of the coenzyme A (Dietz & Rubinstein, 1972). This reaction is also known as the Ellman reaction (Dietz & Rubinstein, 1972). The yellow compound formed was absorbed strongly at 412 nm and an increase in absorbance was thus attributed to formation of coenzyme A, and therefore the progression of the GYLAT reaction.

The expressed hGLYAT proteins were extracted using two different methods. The BugBuster Protein Extraction Reagent method was compared with the Lysis buffer protocol. The final eluates of the expressed and purified recombinant hGLYAT proteins extracted by these methods were used for enzyme assays, to compare their enzymatic activities. The activity of the coN156ShGLYAT enzyme was also compared to the N156ShGLYAT enzyme simultaneously.

For the enzyme assays, reaction mixtures were prepared in the form of master mixes for the sake of convenience and accuracy. The master mixes were composed of 0.1 mM DTNB, 25 mM TrisCl (pH 8), 100 μ M benzoyl-coenzyme A, 200 mM glycine, and filled with nuclease-free water to a final volume of 100 μ l. Negative control reactions were set up identically to the positive reactions, but with glycine omitted from the reaction mixture. Equal amounts (10 μ g) of the different enzyme samples were added to a 96-well plate, after which 100 μ l of the master mix was also added to each of the enzymes in the wells. The reaction mixtures and enzymes were then mixed by briefly shaking the plate by hand. For analysis a BioTek plate reader and accompanying Gen5 software were used. The reactions were allowed to proceed for 6 minutes at 37°C and absorbance measurements were made at 412 nm every 15 seconds.

3.2.25 Long term storage of purified enzyme

In previous enzyme extract preparations of the hGLYAT enzyme, almost all enzyme activity was lost after just one day due to its instability. It was therefore needed to express and purify proteins every time for every kinetic assay and other downstream experiments, which made it extremely time consuming and expensive to study the enzyme. There existed, therefore, a major need for a storage method that preserves significant enzyme activity. A protocol described by Dercksen *et al.* (2013) consisted of a method of snap freezing the proteins after mixing with glycerol. This method was tested with the hGLYAT enzyme by aliquoting 175 µl of the enzymes into small tubes, and adding 25 µl sterile 80% glycerol to the proteins to a final concentration of 10% (Dercksen *et al.*, 2013). These mixtures were then placed into liquid nitrogen to snap freeze, after which they were stored at -80°C.

Enzyme assays were performed in triplicate on the same batch of enzymes after 24 hours, after 5 days, and after 4 months, to assess whether the protein remained active when frozen in this way. The enzyme assays were done as described in Section 3.2.24.

3.2.26 Determination of protein concentration

After verifying the activity of the expressed proteins, it was necessary to determine their concentrations before further analysis. The Qubit® 2.0 Fluorometer (Invitrogen, Life Technologies) and the bicinchoninic acid (BCA) assay were compared for the quantification of the proteins to establish the most accurate and repeatable method.

3.2.26.1 Determining protein concentration with the Qubit 2.0 Fluorometer

The Qubit Protein Assay kit was purchased from Molecular Probes. This method relies on the use of specific dyes that bind to the targets proteins. After binding, the samples become fluorescent. The Qubit protein assay kit can measure protein concentration accurately between 12.5 µg/ml to 5 mg/ml.

Before starting the analyses, a working solution and a set of standards had to be prepared. These solutions were prepared according to the guidelines stated in the Qubit 2.0 Fluorometer protocol. The working solution consisted of 5 µl reagent and 995 µl buffer combined in a 1.5 ml microcentrifuge tube. The standards were prepared by mixing, in separate specialised Qubit tubes, 10 µl of each Standard with 190 µl of the previously prepared working solution. These standards were then used to calibrate the Qubit system prior to protein analyses.

The protein samples were prepared for the Qubit analyses by mixing 1 µl of each protein sample with 199 µl of the working solution in separate Qubit tubes. The tubes were then vortex mixed and incubated for 15 minutes at room temperature. After incubation, the samples' protein content was measured. The N156S hGLYAT protein concentrations were measured six times to evaluate the repeatability of the measurements. The average concentration, the standard deviation and the coefficient of variance (%CV) were calculated.

3.2.26.2 Determining protein concentration with the BCA assay

The BCA method used was adapted from the protocol as described by Sambrook and Russel (Sambrook & Russell, 2001) (Smith, 1985). In the BCA method a highly sensitive colorimetric change occurs when copper ions (Cu^{2+}) are reduced to the cuprous ion (Cu^{1+}) by protein, which then binds to bicinchoninic acid. The resulting BCA:copper complex then absorbs strongly at 562 nm, in a linear manner within the limits of quantification.

Prior to the measurement of the purified recombinant N156S hGLYAT protein concentrations, the samples had to be prepared for the BCA assay. The elution buffer in which the proteins are eluted and stored contains imidazole. For this assay imidazole is considered as an interfering substance and needed to be removed. This was done by dialysis of the samples through Slide-A-Lyzer MINI dialysis devices (Thermo Scientific, Waltham, MA, USA). These tubes contain cellulose membranes across which samples were dialysed into a Tris-HCl (pH 8) buffer.

A Slide-A-Lyzer MINI Dialysis Device was removed from its tube and wet with the dialysis buffer, Tris-HCl. Into the 50 ml Falcon tube, 44.5 ml of the same buffer was added and set aside. The buffer was then removed from the dialysis device by decanting and shaking the device without letting the membrane dry completely. The device was then filled with 0.2-2 ml of the sample and placed into the Falcon tube containing the buffer. The tube was capped and placed carefully into an incubator at room temperature with moderate shaking at 200 rpm. The samples were left to dialyse for 2 hours, and fresh dialysis buffer was then added and further incubated overnight. The device was then removed and the samples collected from the corner of the device using a pipette.

The bicinchoninic acid solution was mixed with a copper sulfate solution in a 50:1 (v/v) ratio and was used as the working solution. Bovine serum albumin (BSA) was used as the protein standard in concentrations between 1 and 10 µg/ml to form a standard curve. In a 96-well plate, 10 µl of these protein samples were added, followed by the addition of a further 200 µl of the working solution to each sample. The reactions were set up in triplicate. The plate was shaken to ensure mixing of the reactions and incubated for 20 minutes at 37 °C. The BioTek plate reader was used to measure the absorbance at 560 nm. The standard curve was drawn by the accompanying Gen5 software and the protein concentrations of the samples could be determined automatically.

3.3 Results and discussion

3.3.1 Extraction of pET-32a(+) and pMA-T plasmids

Midi-preparations were done to extract the pET-32a(+)/WThGLYAT plasmid from *E. coli* cells, and pMA-T/cohGLYAT from *E. coli* cells. After extraction, electrophoretic analysis was performed to verify the successful extraction of the correct plasmids. Figure 3.6 shows the bands formed by electrophoresis. The O'GeneRuler DNA Ladder was added to the first lane to serve as a size marker. For the pET-32a(+) plasmid and the hGLYAT coding sequence, the literature reports sizes of 5900 bp and 888 bp, respectively. Furthermore, for the pMA-

T/cohGLYAT vector, a size of 3277 bp is expected (as per manufacturers specifications) It is clear that the results obtained by this experiment are consistent with the literature, and that the correct plasmids have been extracted.

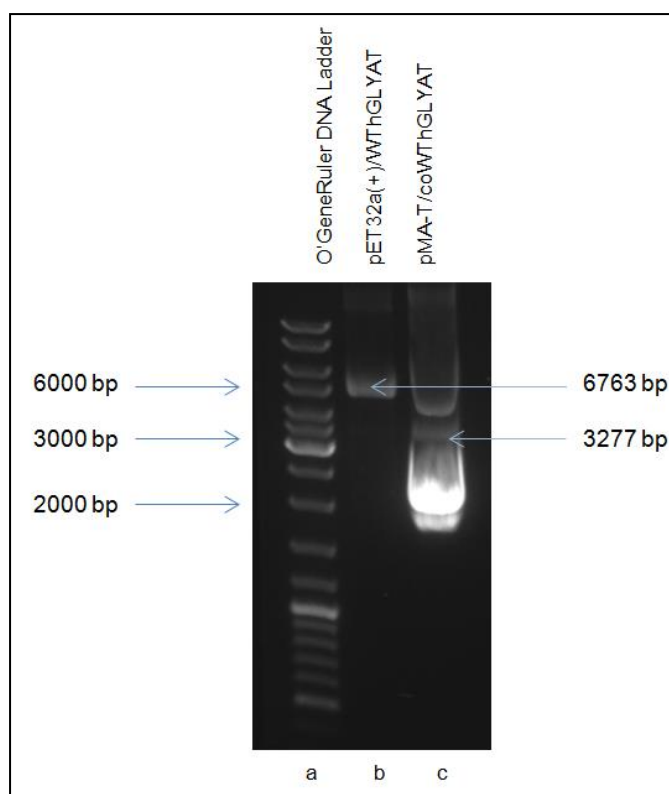


Figure 3.6: Agarose gel (1%) of extracted pET-32a(+)/WThGLYAT (b) and pMA-T/coWThGLYAT (c). O'GeneRuler DNA Ladder (a) served as a molecular weight marker.

3.2.1 Restriction endonuclease digests of plasmids

After plasmid extraction, the pET-32a(+)/WThGLYAT and pMA-T/coWThGLYAT plasmids were double digested with *Bam*HI and *Hind*III restriction enzymes in order to remove the hGLYAT inserts from the plasmids. The pET-32a(+) plasmids contained the WThGLYAT gene and the pMA-T plasmids contained the coWThGLYAT gene. The plasmids needed to be separated from their hGLYAT inserts, to be able to isolate the coWThGLYAT insert from its vector and subclone it into the pET-32a(+) plasmid. Agarose gel electrophoresis analyses were performed to separate the digested plasmids from their inserts (Figure 3.7). The

yellow boxes on Figure 3.7 indicate a 5900 bp band (lane b) and a 903 bp band (lane d). These bands correspond with the expected sizes for the digested linearised pET-32a(+) plasmid and the coWThGLYAT fragment. To serve as negative controls, the undigested plasmids (with their respective inserts) were also loaded onto the gel in separate wells (lanes c and e). Following this step, the pET-32a(+) vector and the coWThGLYAT insert were excised from the gel and gel purified.

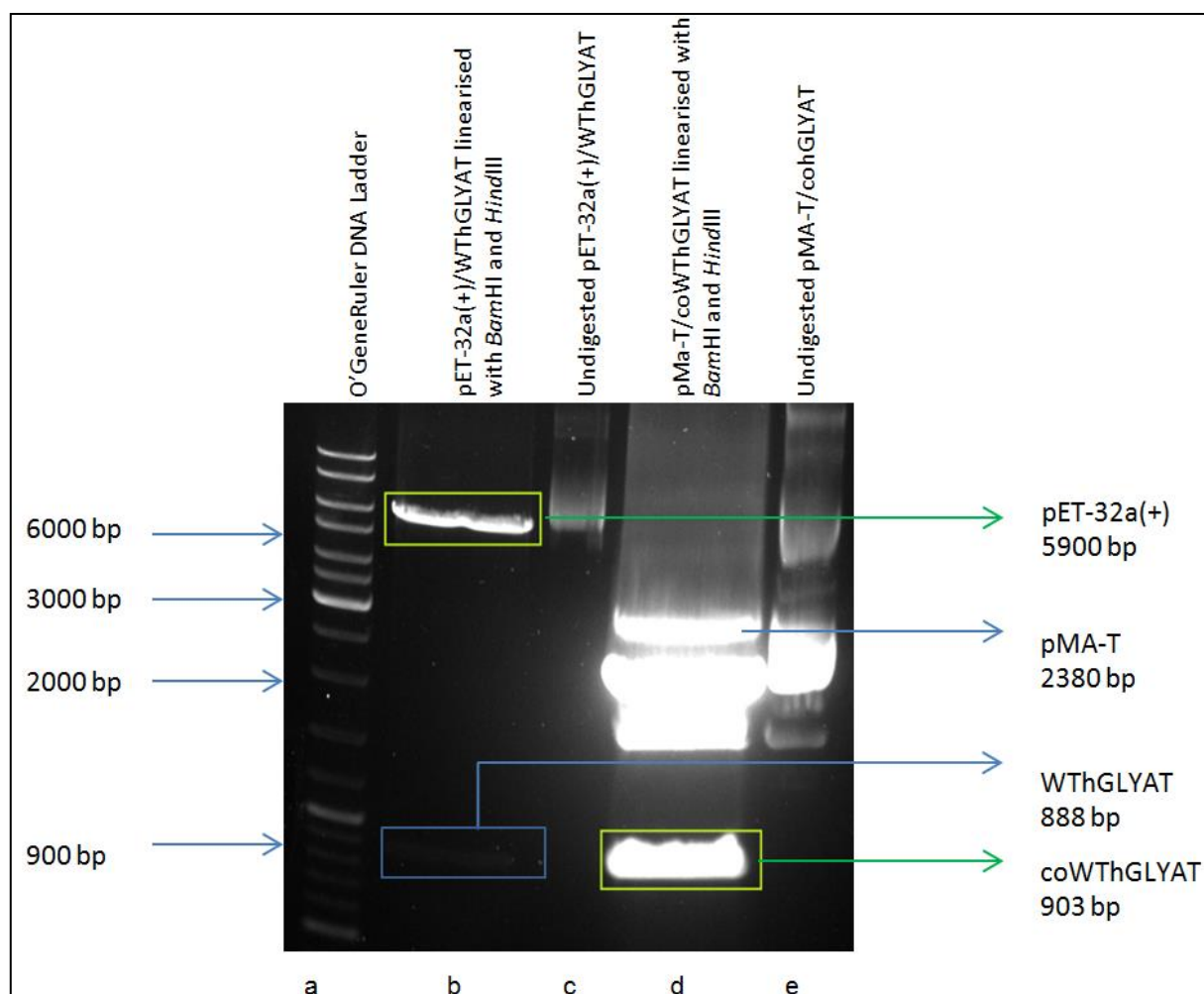


Figure 3.7: Agarose gel (1%) of the *Bam*HI and *Hind*III restriction enzyme digestions of the pET-32a(+) (b) and pMa-T (d) recombinants to gel purify the vector and insert fragments. The undigested pET-32a(+)/WThGLYAT (c) and pMa-T/coWThGLYAT (e) served as controls. O'GeneRuler DNA Ladder (a) served as a molecular weight marker.

3.3.2 Verification of gel extraction

After the digested pET-32a(+) vector and the coWThGLYAT insert were gel purified, another electrophoretic analysis was performed. This was done to verify that the purified fragments were in fact the vector and insert excised from the gel slices. To serve as a negative control, the undigested pET-32a(+)/WTGLYAT plasmid was also loaded into the second lane. Figure 3.8 shows the 5900 bp linearised pET-32a(+) plasmid that was extracted from the previous gel, and the 903 bp coWThGLYAT fragment that was also isolated from the gel. After verifying the successful isolation of the fragments, the coWThGLYAT was cloned into the *Bam*HI/*Hind*III sites of the pET-32a(+) vector. Ligation reactions were performed with the Rapid DNA Ligation Kit as described in Section 3.2.12 and transformed into chemically competent *E. coli* cells.

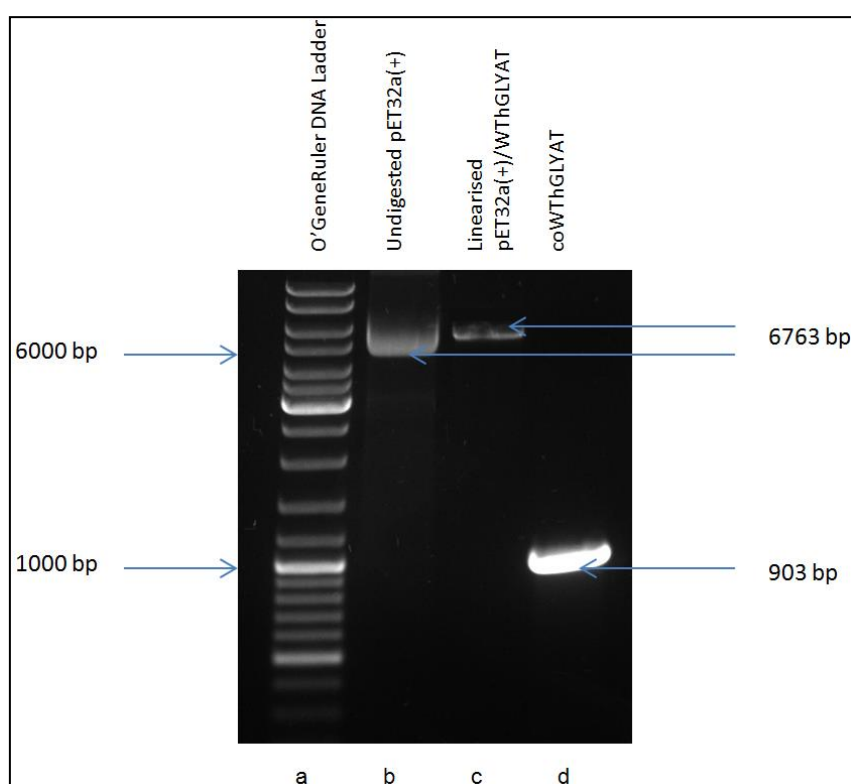


Figure 3.8: Agarose gel (1%) of the gel extracted linearised pET-32a(+)/WThGLYAT vector (c) and the coWThGLYAT insert (d). The undigested pET-32a(+)/WThGLYAT (b) served as a control. O'GeneRuler DNA Ladder (a) served as a molecular weight marker.

3.3.3 Colony screening after transformation

Following the ligation reactions and transformation, colonies were obtained. Ten of these colonies were then grown overnight in liquid culture (Section 3.2.13) and the plasmids were isolated the following day (Section 3.2.6). In order to verify the presence of the desired *coWThGLYAT* insert in the pET-32a(+) plasmid, an agarose gel electrophoresis analysis was performed. To serve as a control, the pET-32a(+) plasmid containing the *WThGLYAT* insert was loaded into the second lane. From these gels it could be estimated whether the samples contained the plasmid and the desired insert, by analysing the estimated size of the fragments. Figure 3.9 confirms the success of the subcloning in 8 out of the 10 sample mixtures. The ~6000 bp bands of the pET-32a(+)/*coWThGLYAT* vectors can be seen in lanes d-f in Figure 3.9A and lanes j-n in Figure 3.9B. Lanes c and g from Figure 3.9A indicated failed ligation steps, since the band sizes were smaller than expected, and could be ascribed to a contaminating vector.

3.3.4 DNA sequence determination and data analysis

The recombinant pET-32a(+)/*coWThGLYAT* vectors (colonies from Figure 3.9A, lanes d-f, and Figure B, lanes j-n) were then selected and sent for Sanger sequencing by the CAF of the University of Stellenbosch, South Africa, to confirm that the *coWThGLYAT* gene was successfully subcloned into the pET-32a(+) plasmid in the correct orientation. The sequencing chromatograms were analysed using FinchTV (version 1.40) (www.geospiza.com/finchtv). ClustalX was used for the alignment of the sequences to reference sequences.

The results of the sequences confirmed that the *coWThGLYAT* gene was correctly inserted into the pET-32a(+) plasmid. Figure 3.10 shows the aligned sequences of the *WThGLYAT* gene and *coWThGLYAT* gene, and the sequences generated by the forward and reverse primers of one of the sequenced recombinant vector samples.

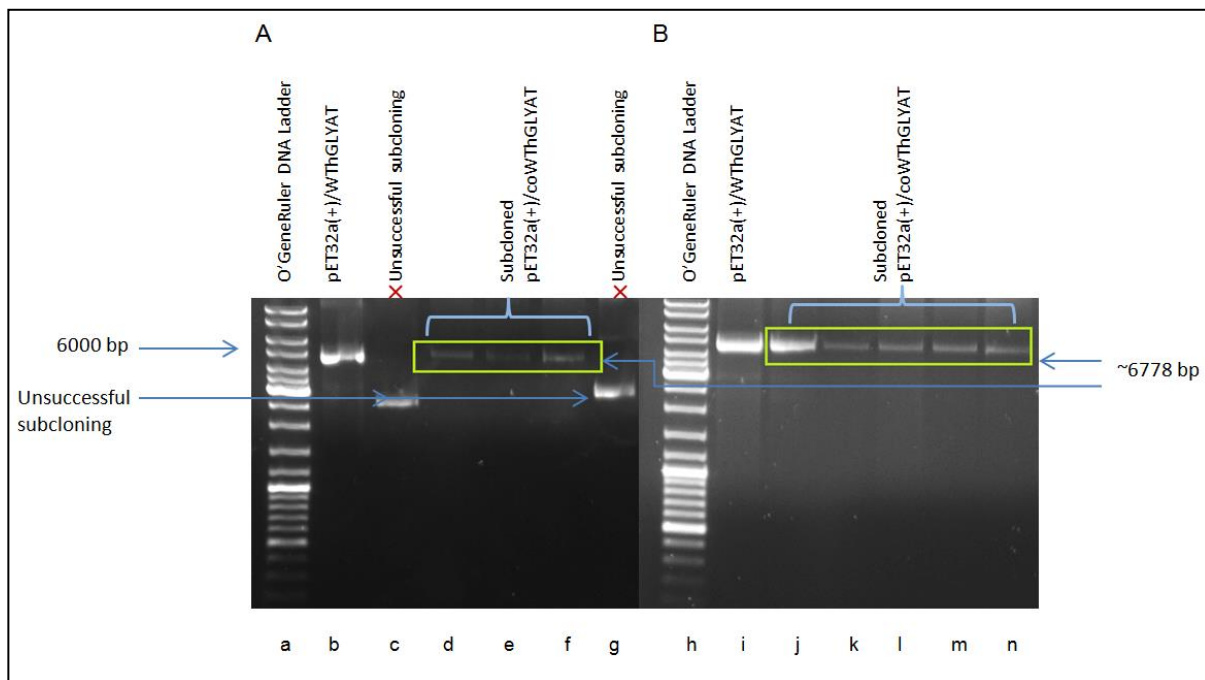


Figure 3.9: Agarose gel (1%) of eight samples of successfully subcloned pET-32a(+)/coWThGLYAT (Figure A, d-f, Figure B, j-n). Two samples indicated failed subcloning steps, because the bands obtained were smaller than expected and could be a contaminating vector (Figure A, c and g). The undigested pET-32a(+)/WThGLYAT (Figure A, b and Figure B, i) served as controls. O'GeneRuler DNA Ladder (Figure A, a and Figure B, h) was included as a molecular weight marker.

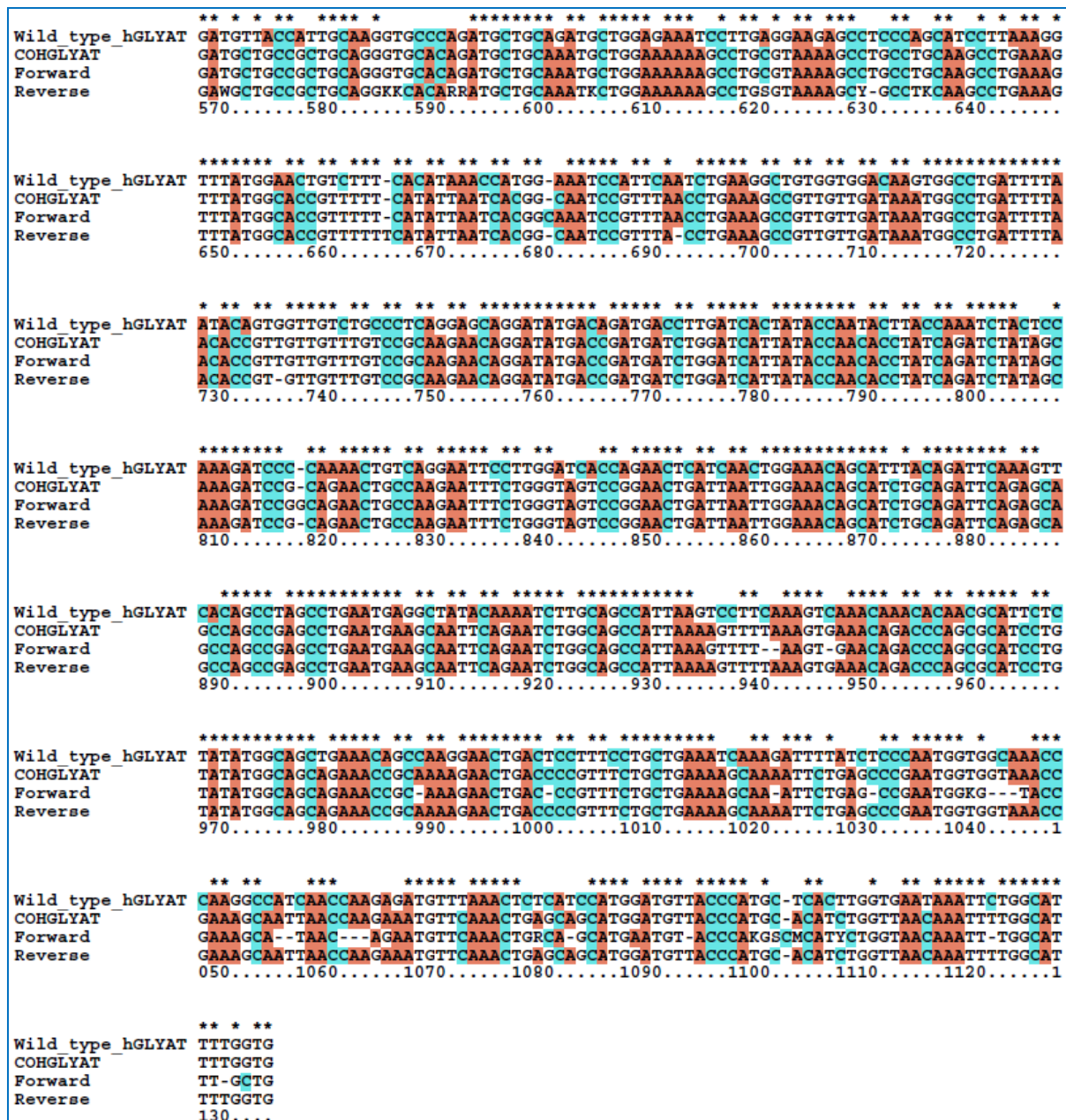


Figure 3.10: The aligned sequences of the *coWThGLYAT* insert, the *hGLYAT* insert, and the sequences generated by the forward and reverse primers. These sequences confirmed that the *coWThGLYAT* gene was successfully inserted to the pET-32a(+) vector instead of the *WThGLYAT* gene.

3.3.5 Generation of the coN156ShGLYAT variant

For this study, the N156S variant of the *coWThGLYAT* gene had to be generated. This single nucleotide polymorphism (SNP) was introduced into the *coWThGLYAT* gene in the pET-32a(+) plasmid by site-directed mutagenesis as described in Section 3.2.16. During this method, the coding sequence was amplified, while the desired SNP was introduced in the sequence.

3.3.5.1 Polymerase chain reaction (PCR) optimisation

Site-directed mutagenesis usually requires some optimisation of the PCR conditions. One such condition is the annealing temperature which typically varies between reactions. It was thus necessary to determine the optimal annealing temperature for this specific protocol. For this purpose, several additional reaction mixtures were prepared and exposed to different annealing temperatures during the PCR amplification cycles. The annealing temperatures were varied between 65-72°C for five test reactions (65.2°C, 67.2°C, 69.0°C, 70.5°C, and 71.6°C respectively). Electrophoretic analysis was then performed to verify the amplification of the vector with the modified hGLYAT insert. By analysing the size of the fragments on the gels, it could be estimated whether PCR reactions were successful. However, this step just served as provisional verification, since absolute verification of the success of the site-directed mutagenesis process, Sanger sequencing was done.

As shown in Figure 3.11 (lane e) an annealing temperature of 65.2°C was found to be optimal, as it resulted in the formation of a band of roughly 6778 bp when run on an agarose gel. This band was assumed to represent the pET-32a(+)/coN156ShGLYAT plasmid. The undigested pET-32a(+)/WThGLYAT plasmid was loaded into lane c. To serve as a negative control, no DNA was added to the PCR reaction, which was loaded onto the gel in lane d. To serve as a positive control, PCR products of pET-32a(+)/N156ShGLYAT was added to lane j.

Some non-specific bands were also formed in lanes f-i (Figure 3.11). These non-specific products could have been ascribed to various causes involving the PCR components. These include (i) non-optimal Mg^{2+} concentrations, (ii) too high dNTP concentrations, (iii) DNA contamination/carry over, (iv) mispriming caused by template secondary structure, (v) too high primer concentrations (Bio-Rad PCT Troubleshooting). Other causes related to the temperatures and cycling times include (i) excessive cycling, (ii) too slow ramp speed, (iii) and inaccurate T_m . Since these non-specific bands were only formed in the reactions with higher T_m values, it was assumed that too high T_m values were the cause of these bands.

3.3.6 DNA sequence determination of coN156ShGLYAT variant

As mentioned previously, the samples were Sanger sequenced by the CAF of the University of Stellenbosch, South Africa, to confirm the introduction of the N156S mutation into the *coWThGLYAT* gene and that the gene was successfully subcloned into the pET-32a(+) plasmid in the correct orientation.

The sequencing results confirmed that in the *coWThGLYAT* gene an asparagines codon (AAT) was successfully changed to a serine codon (AGT) at nucleotide position 156 by site-directed mutagenesis. Figures 3.12 and 3.13 show the aligned sequences of the reference *coWThGLYAT* gene, and the sequences generated by the forward and reverse primers. The dark blue box indicated in Figure 3.12 shows the position where the SNP was introduced in the mutation containing sequences, as opposed to the reference *coWThGLYAT* insert. Figure 3.13 shows an exploded view of the position on the *coWThGLYAT* gene where the SNP was introduced, resulting in the coN156ShGLYAT variant.

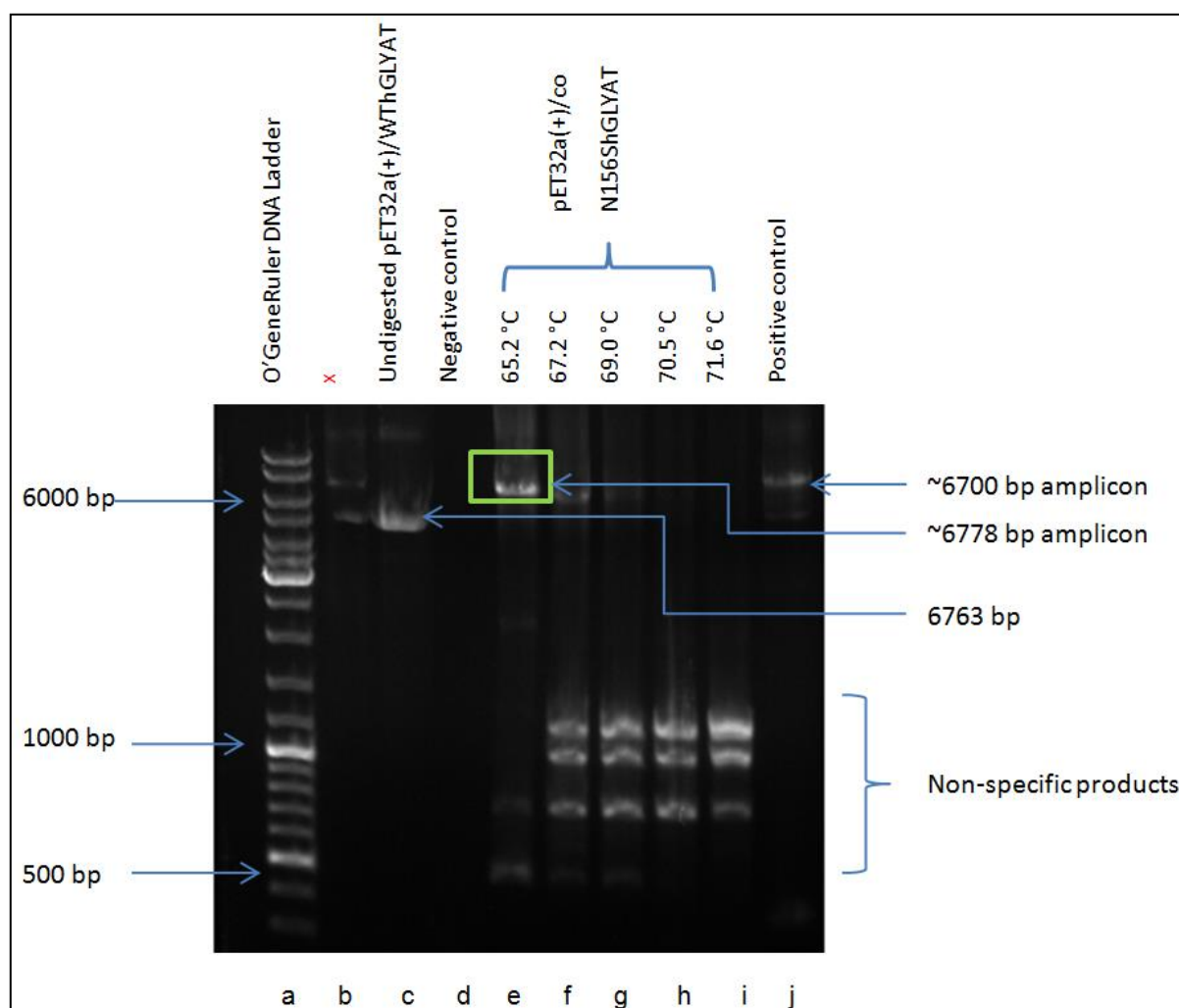


Figure 3.11: Agarose gel (1%) of the PCR amplification products of pET-32a(+)/coN156ShGLYAT using various annealing temperatures. The undigested pET-32a(+)/WThGLYAT plasmid (c) and the negative control reaction with no DNA added (d). PCR products of pET-32a(+)/coN156ShGLYAT using various annealing temperatures ranging from 65.2 to 71.6 °C (e-i). The positive control with the PCR product of pET-32a(+)/N156ShGLYAT was added (j). The green box (e) indicates the optimal annealing temperature (65.2°C) since a band of the correct size can be seen. O'GeneRuler DNA Ladder (a) served as a molecular weight marker.

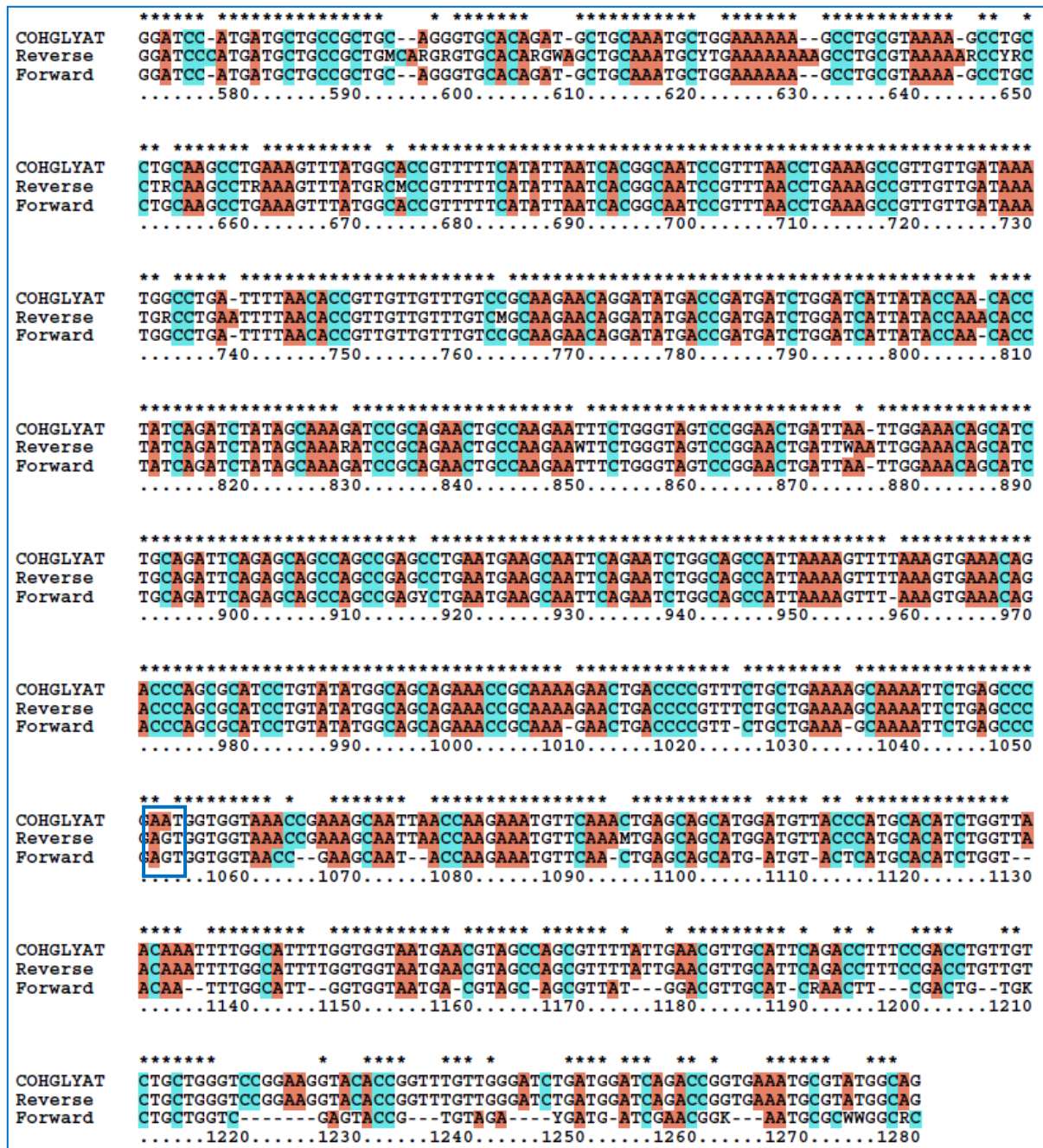


Figure 3.12: The aligned sequences of the coWThGLYAT insert, and the sequences generated by the forward and reverse primers. These sequences confirmed the introduction of the N156S SNP in the coWThGLYAT sequence resulting in the coN156ShGLYAT variant. The dark blue box shows the position where the SNP was introduced in the modified sequences, as opposed to the non-mutant original coWThGLYAT insert.

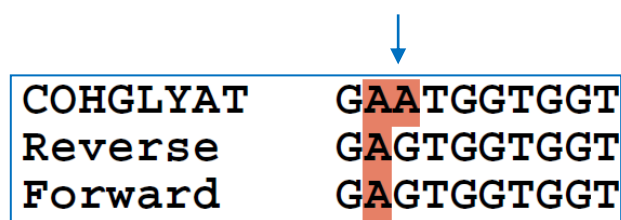


Figure 3.13: The exploded view of the aligned sequences of the coWThGLYAT insert, and the sequences generated by the forward and reverse primers. These sequences indicate the position where the SNP was introduced in the modified sequences (coN156ShGLYAT), as opposed to the non-mutant original codon optimised hGLYAT insert. The nucleotide change from AAT to AGT can be seen.

3.3.7 Optimisation of protein extraction

For the extraction of hGLYAT proteins, the BugBuster Protein Extraction Reagent method was compared with the Lysis buffer protocol (section 3.2.22). This was done to assess the efficiency of these methods to extract the hGLYAT proteins, and to evaluate its capacity to deliver enzymatically active proteins after extraction. The protein products obtained by both of these methods were purified with His-tag purification. SDS-PAGE analyses were performed throughout to confirm that the expressed recombinant hGLYAT proteins had been successfully extracted and purified. Finally enzyme activity assays were performed, as described in section 3.2.24.

The coN156ShGLYAT and N156ShGLYAT variants were compared to assess the protein yield, to thus determine whether the codon optimisation had a significant effect on the expression of the protein. The N156ShGLYAT variant was already available in the laboratory. SDS-PAGE analyses of the N156ShGLYAT and the coN156ShGLYAT proteins extracted by means of the BugBuster Protein Extraction Reagent are shown in Figures 3.14A and 3.14B, respectively. Figures 3.15A and 3.15B show the SDS-PAGE analyses of the N156ShGLYAT and the coN156ShGLYAT extracted by means of the Lysis buffer method, respectively. In all the gels, the first lane was loaded with 10 µl protein ladder (PageRuler Plus). The other lanes were loaded as stated in the figure descriptions. The gels show bands of approximately 56 kDa.

These band sizes correspond with those expected for the purified recombinant N156ShGLYAT proteins with an N-terminal His-tag. The expected sizes of the hGLYAT protein bands on the gels are approximately 37 kDa according to the estimated sizes according to the length of the amino acid sequence (903 bp; 301 amino acids; ~37 kDa). However, it is important to note that these sizes are influenced by the presence of the N-terminal fusion found on the pET-32a(+) plasmid. This tag is approximately 29 kDa in size and will contribute to the final size of the protein band seen on the gel, which would be 56 kDa.

In the final elution lane of Figure 3.14B a clear band corresponding with the hGLYAT cannot be seen. This could possibly be ascribed to over destaining. However, the fraction did show enzyme activity which indicates the presence of at least some protein. A slight difference in the yield of soluble proteins, from the final elutions of the different extraction methods (BugBuster vs Lysis buffer), could be noticed when looking at the SDS-PAGE gels. The proteins extracted with the Lysis buffer method resulted in a slightly thicker band, which indicates more protein being present in those fractions.

It was hypothesised that the expression of the recombinant hGLYAT results in the depletion of coenzyme A and glycine in the *E. coli* cells, which resulted in poor growth of the *E. coli* as well as low protein yields. To circumvent this issue, 0.5% glycine was added to the expression medium. A noticeable difference was observed between the yields of the coN156ShGLYAT and N156ShGLYAT proteins (as seen in Figure 3.14). It was consequently assumed that this could be ascribed to the coN156ShGLYAT using the added glycine more effectively, depleting the glycine levels prematurely, and resulting in poorer expression than its non-codon optimised N156ShGLYAT counterpart. However, further investigation is needed to ultimately confirm this theory. This can possibly be done by supplementing the glycine throughout the expression protocol in different concentrations, and assessing the yields.

It was therefore decided to continue further investigations of the N156S hGLYAT enzyme using the non-codon optimised N156S hGLYAT variant in combination with the Lysis buffer extraction method, since this combination yielded the most soluble proteins.

The Origami cells were co-transformed with the pGro7 plasmid which resulted in the co-expression of the GroEL-GroES chaperone proteins. These bands are also visible on the SDS-PAGE gels in lanes g and h on Figures 3.14 and 3.15. These bands corresponded to the sizes reported for the GroEL protein, which delivers a band of approximately 60 kDa and GroES, which delivers bands of around 10 kDa.

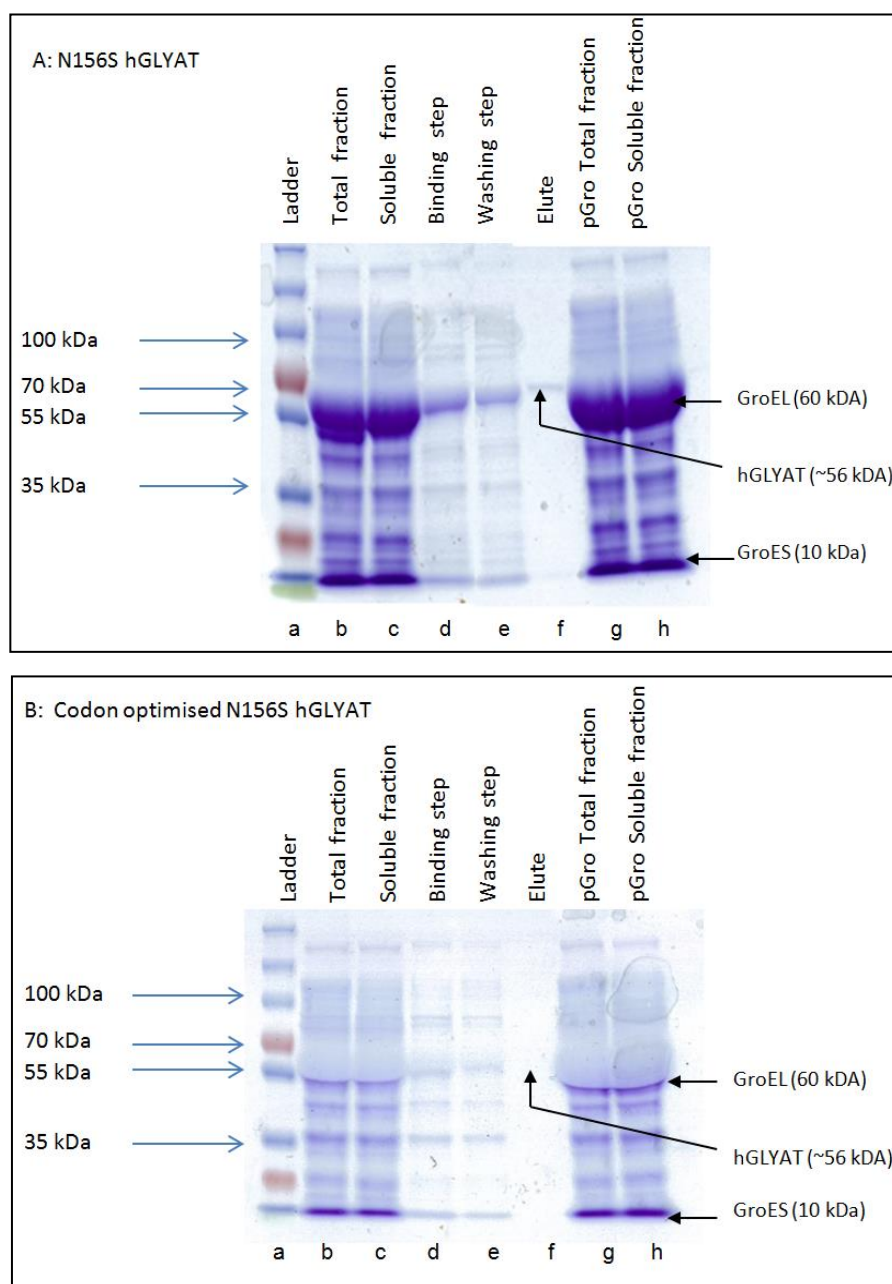
BugBuster method:

Figure 3.14: SDS-PAGE (12.5%) analyses of A) the N156ShGLYAT and B) the coN156ShGLYAT extracted by means of the BugBuster Protein Extraction Reagent (as described in the text). Total (b) and soluble fractions (c) were taken of the samples; fractions of binding step elute (d), washing step elute (e) and the final elute (f) were made between steps. The pGro7 chaperone total (g) and soluble fractions (h) were also included. The purified recombinant hGLYAT proteins are visible as bands of approximately 56 kDa. PageRuler protein marker (a) served as a molecular weight marker.

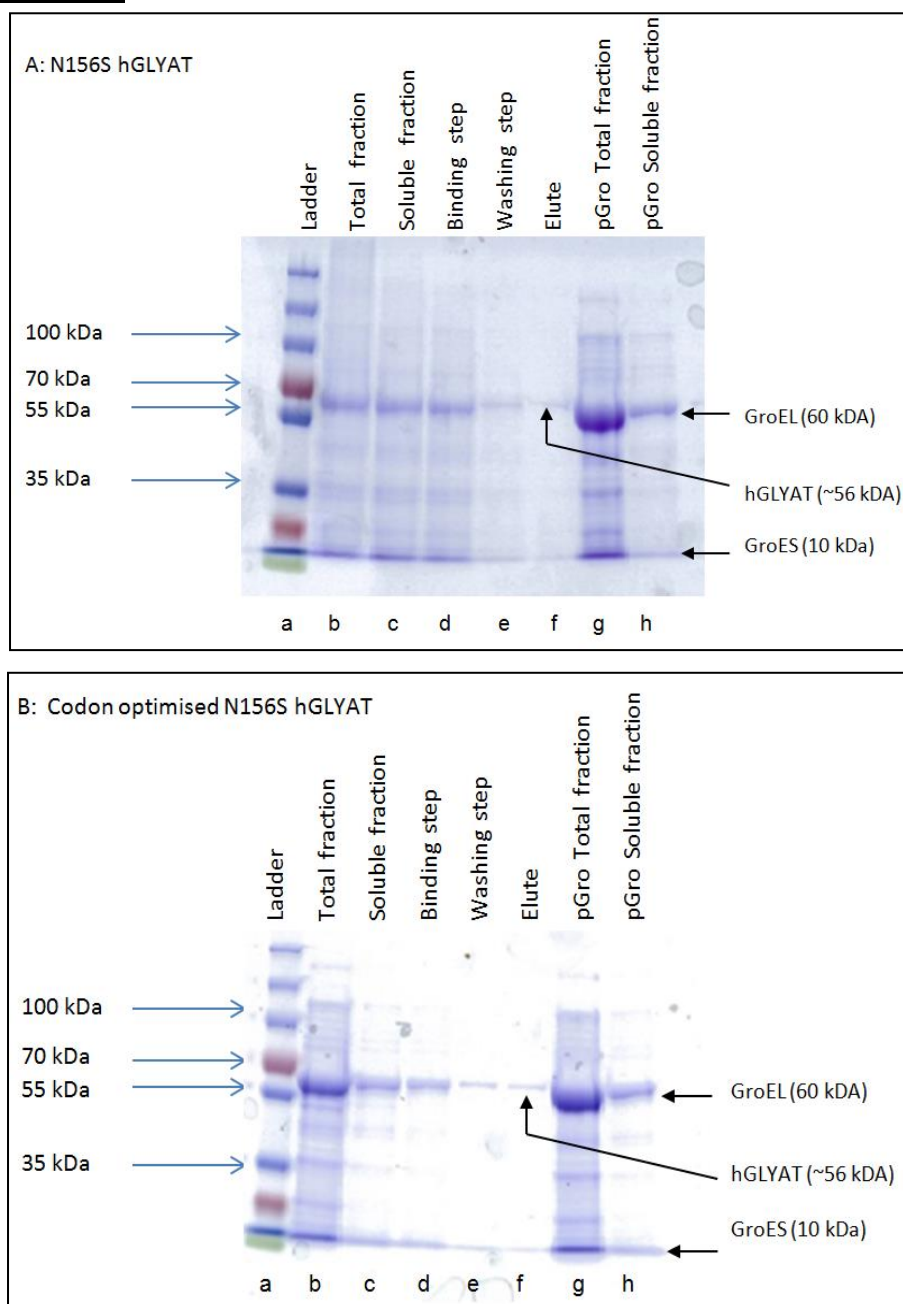
Lysis buffer method:

Figure 3.15: SDS-PAGE (12.5%) analyses of A) the N156ShGLYAT and B) the coN156S hGLYAT extracted by means of the Lysis buffer method (as described in the text). Total (b) and soluble fractions (c) were taken of the samples; fractions of binding step elute (d), washing step elute (e) and the final elute (f) were made between steps. The pGro7 chaperone total (g) and soluble fractions (h) were also included. The purified recombinant hGLYAT proteins are visible as bands of approximately 56 kDa. PageRuler protein marker (a) served as a molecular weight marker.

3.3.8 Recombinant N156S GLYAT enzyme activity assays

Enzyme activity assays were performed after each expression step to verify GLYAT activity. For this purpose, a colorimetric reaction was used to detect coenzyme A, a product of the GLYAT reaction. The absorbance levels of the reactions were measured at 412 nm and an increase in absorbance, compared to control reactions without glycine, was thus attributed to formation of coenzyme A, and therefore the progression of the GYLAT reaction.

As mentioned in section 3.2.22, proteins were extracted using two methods. The BugBuster Protein Extraction Reagent method was compared with the Lysis buffer protocol. Assuming that the hGLYAT enzyme relative contents are equal in the elution fractions, equal amounts (10µg) of the expressed and purified recombinant hGLYAT proteins extracted by these methods were used for enzyme assays, to compare their enzymatic activities. The activity of the coN156S hGLYAT enzyme was also compared to the N156ShGLYAT enzyme simultaneously.

Figures 3.16 and 3.17 shows the enzyme assay plots of the purified mutant recombinant enzymes extracted with the BugBuster Reagent and the Lysis buffer methods respectively. The graphs show the change in absorbance over time. The plot seen in figure 3.16 shows that the enzyme activities of the coN156ShGLYAT and the N156ShGLYAT enzymes extracted with the BugBuster reagent do not differ significantly. The N156ShGLYAT enzyme reached an absorbance value of 0.180 after 6 minutes and the coN156ShGLYAT enzyme reached an absorbance value of 0.166 after 6 minutes. Figure 3.17 shows that the enzyme activities of the coN156ShGLYAT and the N156ShGLYAT enzymes extracted with the lysis buffer method also do not differ significantly. The N156ShGLYAT enzyme reached an absorbance value of 0.194 after 6 minutes and the coN156ShGLYAT enzyme reached an absorbance value of 0.185 after 6 minutes. It can be concluded that the codon optimisation did not have a significant effect on the enzyme activity of the expressed proteins.

A difference in activity could however be noticed between the proteins extracted by the BugBuster reagent and the lysis buffer method. From the figures, it is clear that the absorbance values of the BugBuster extracted proteins were lower than those extracted by the lysis buffer, thus indicating that the lysis buffer method was a better choice of extraction. The lysis buffer method was used from this point on for further experiments.

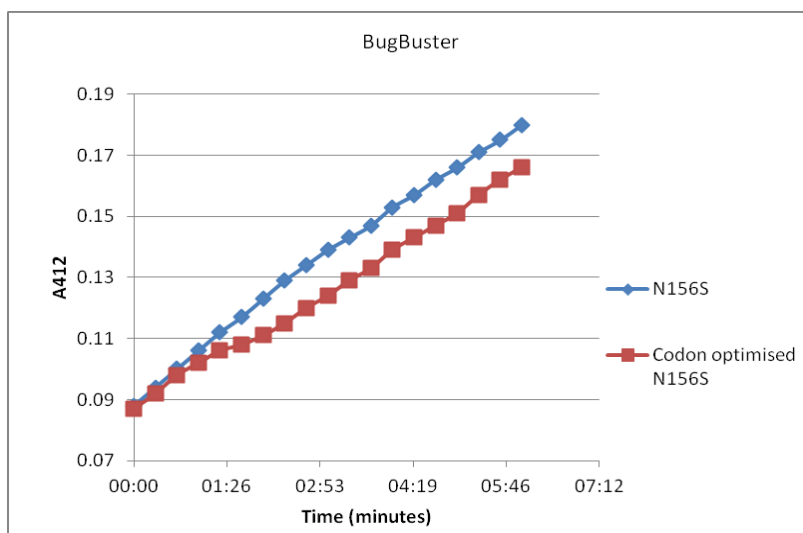


Figure 3.16: The enzyme assay plots of the purified mutant recombinant enzymes extracted with the BugBuster Reagent. The coN156ShGLYAT and the N156ShGLYAT enzymes' activity were compared. The graph shows the change in absorbance at 412 nm over six minutes.

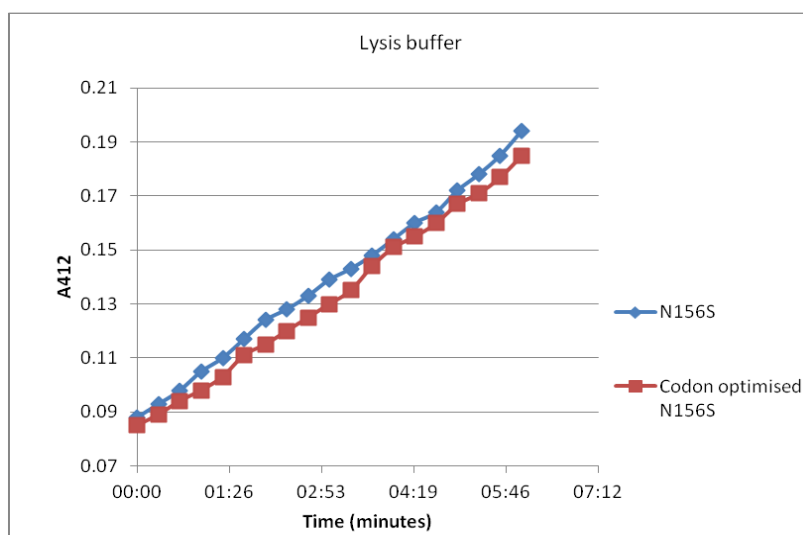


Figure 3.17: The enzyme assay plots of the purified mutant recombinant enzymes extracted with the Lysis buffer method. The coN156ShGLYAT and the N156S hGLYAT enzymes' activity were compared. The graph shows the change in absorbance at 412 nm over six minutes.

3.3.9 Long term storage of purified recombinant N156ShGLYAT variant

Optimal storage conditions for hGLYAT have not yet been investigated. Since this is lacking, fresh enzyme had to be prepared for every downstream experiment or several assays had to be performed in one day, severely complicating the work on hGLYAT. To circumvent this issue the long term storage of hGLYAT was investigated by using a protocol described by Dercksen *et al.* (2014) in which N-acetylglutamate synthase function was preserved. The protein was snap frozen after being mixed with 10% glycerol, followed by storage at -80°C.

After storage, enzyme assays were performed in triplicate on the same batch of enzymes after 24 hours, 5 days, and after 4 months to assess their activity. Figure 3.18 shows the enzyme activity plots measured at each of the above time points. These plots indicate a slight decrease in activity after the first freeze-thaw cycle after 24 hours of storage. After five days, no further decrease in activity was noticed, as the absorbance values were still comparable to those measured in the previous reading. After 4 months a slight decrease in activity was observed, but the enzyme, which was diluted 5-fold for previous assays, could just be diluted to a ratio of 1:3 instead and the absorbance values remained comparable to those measured earlier. Thus, the enzyme appeared to be stable after storage in 10% glycerol at -80°C.

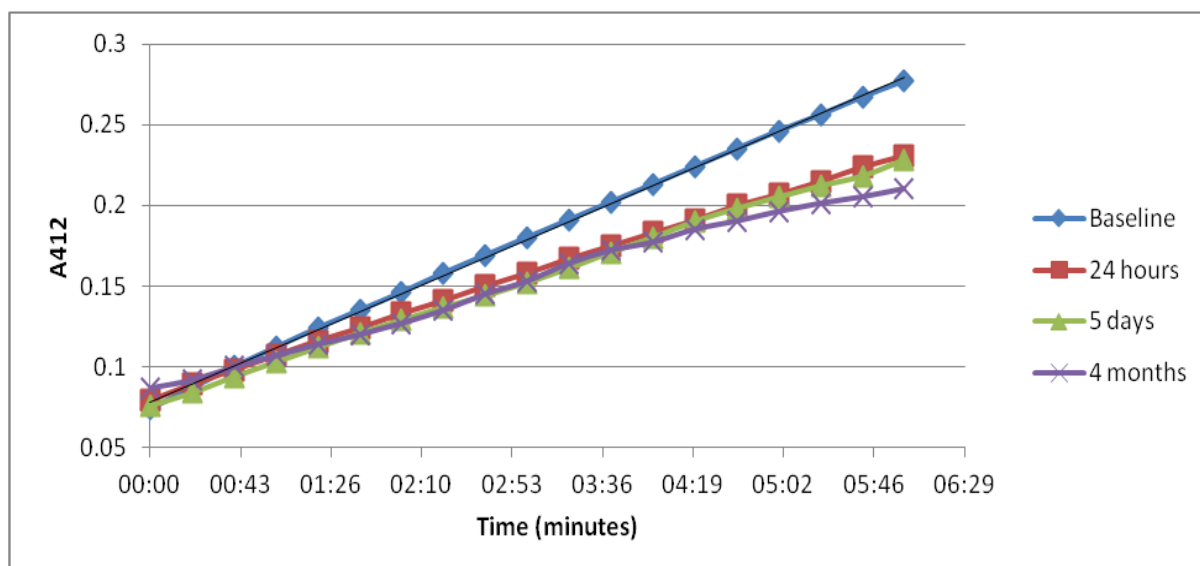


Figure 3.18: The enzyme activity plots of the purified N156S recombinant enzymes from each of the time points at which long term storage was investigated. Four time points were used, the first being at 0 minutes, which served as the baseline reading. The second measurement was done after 24 hours and the third after 5 days of storage and a final measurement was made after 4 months. The graph shows the change in absorbance at 412 nm over six minutes.

3.3.10 Determination of protein concentration

After verifying that the expressed N156S hGLYAT enzymes were active, the protein content was determined. The Qubit 2.0 Fluorometer and the BCA assay methods were employed for the quantification of the proteins to compare the methods in terms of accuracy and repeatability.

3.3.10.1 Determining protein concentration with the Qubit 2.0 Fluorometer

The Qubit 2.0 Fluorometer and the Qubit Protein Assay kit were used as described in Section 3.2.26.1. The working solutions and standards were prepared according to the guidelines stated in the protocol. The N156S hGLYAT protein concentrations were measured six times to evaluate the repeatability of the measurements. The average concentration, the standard deviation and the coefficient of variance (%CV) were calculated for the concentration measurements taken and are shown in Table 3.7. The standard deviation was

very small, indicating that the method demonstrated good repeatability. The precision (%CV) of the measurements was 1.38%, which was satisfactory, since the criterion of acceptance is generally within 15% of the %CV.

Table 3.7: A summary of the concentrations of the N156S hGLYAT protein as determined by the Qubit Fluorometer method. The average concentration, standard deviation and %CV of the values are also shown.

Measurement	Concentration ($\mu\text{g/ml}$)	Average concentration ($\mu\text{g/ml}$)	Standard deviation	%CV
1	586	598.66	8.26	1.38
2	596			
3	594			
4	608			
5	602			
6	606			

3.3.10.2 Determining protein concentration with the BCA assay

The BCA method was used as described in Section 3.2.26.2. Bovine serum albumin (BSA) was used as the protein standard. Standard concentrations of 0, 2, 4, 6, 8, and 10 $\mu\text{g/ml}$ BSA were prepared in triplicate. Prior to the measurement of the N156S hGLYAT protein concentrations the samples had to be prepared for the BCA assay. The elution buffer in which the proteins are eluted and stored contains imidazole. For this assay imidazole is considered as an interfering substance and needed to be removed from the samples. This was done by dialysis of the samples through Slide-A-Lyzer MINI dialysis devices with

cellulose membrane into a Tris-HCl (pH 8) buffer.

The BioTek plate reader was used to measure the absorbance at 560 nm and a standard curve was drawn by the accompanying Gen5 software. Figure 3.19 shows the standard curve generated from the absorbance values measured over the concentrations of the BSA standards. The r^2 value of the curve was 0.98. This curve was then used to determine the protein concentrations of the N156S hGLYAT protein samples automatically.

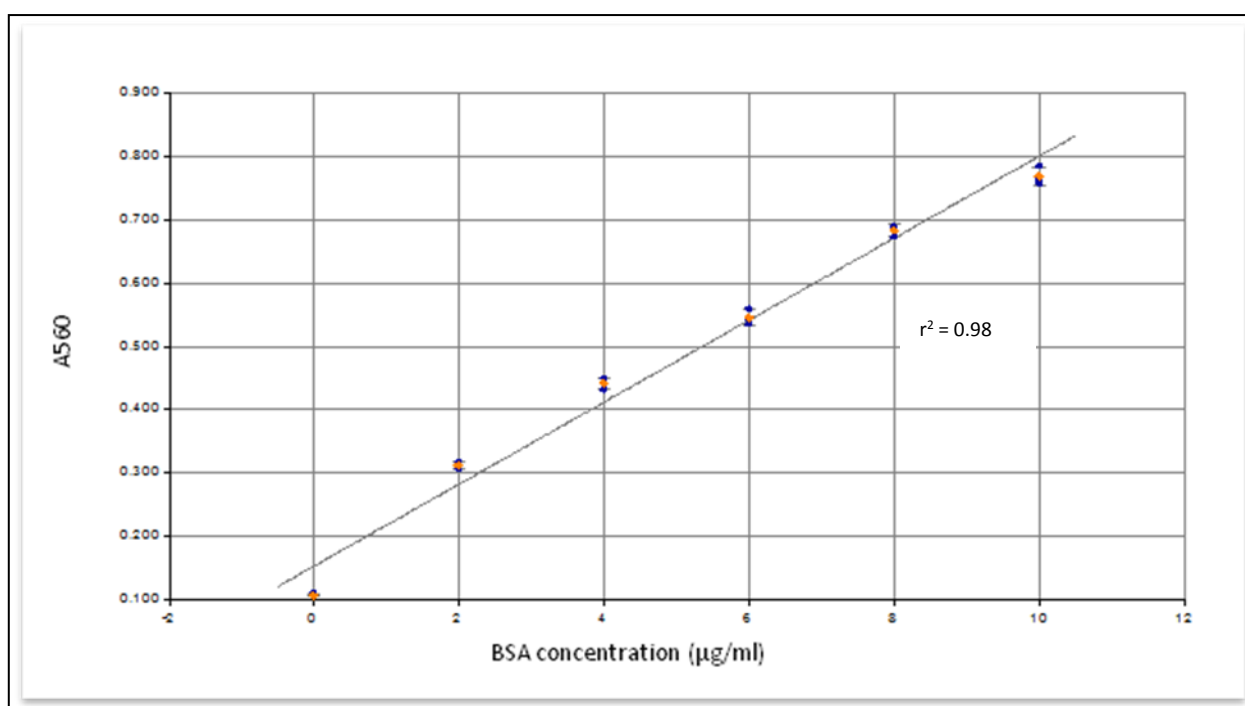


Figure 3.19: Standard curve of a concentration range (0-10 µg/ml) BSA. The plot was used for concentration determination with the BCA assay. The graph shows the absorbance values at 560 nm.

To evaluate the repeatability of the measurements, the protein content of the N156ShGLYAT was measured in triplicate. The concentrations measured in the samples are summarised in Table 3.8, along with the average concentration, the standard deviation and the %CV. The standard deviation was quite large, indicating that the method demonstrated poor repeatability. The precision (%CV) of the measurements was 22.47%, which was inadequate, since the criterion of acceptance is generally within 15% of the %CV.

Table 3.8: A summary of the concentrations of the N156ShGLYAT as determined by the BCA assay. The average concentration, standard deviation and %CV of the values are also shown.

Measurement	Concentration (µg/ml)	Average concentration (µg/ml)	Standard deviation	%CV
1	6.04	5.70	1.28	22.47
2	6.77			
3	4.28			

The concentrations determined with the BCA assay were almost 100-fold lower than those measured by the Qubit assay. This could be attributed to the loss of sample during the dialysis step. The dialysis was performed overnight and the hGLYAT protein is extremely unstable, which could have resulted in the degradation of the proteins in the sample. It was concluded that the Qubit Fluorometer protein concentration results were more repeatable and reliable and these values were consequently used for further experiments in this study.

3.4 Summary

The overall aim of this part of the study was to further optimise the expression and storage conditions of an enzymatically active recombinant N156ShGLYAT variant. To achieve this, a pET-32a(+)/WThGLYAT vector was digested with *Bam*HI and *Hind*III restriction enzymes and the fragments were gel purified. Another vector, the pMA-T/coWThGLYAT vector was also digested with *Bam*HI and *Hind*III and the fragments gel purified. The pET-32a(+) vector fragment and coWThGLYAT insert fragments were excised from the gel and purified. Following this, a ligation reaction was performed and the reactions were transformed into chemically competent *E. coli* cells. This construct was then Sanger sequenced and it was confirmed that the coWThGLYAT insert was inserted into the pET-32a(+) vector in the correct orientation.

The second objective was to generate the coN156ShGLYAT variant via site-directed mutagenesis. The construct was Sanger sequenced and it was confirmed that the cohGLYAT insert was inserted into the pET-32a(+) vector in the correct orientation and that it also contained the desired N156S mutation.

The third objective was to optimise the expression of hGLYAT to increase the yield of biologically active enzyme and to then purify the enzyme using nickel affinity chromatography. The expression protocol was improved by the addition of 0.5% glycine to the expression medium which resulted in enhanced growth of the *E. coli* cells, and it was assumed that it also increased protein yields. However, this theory needs to be further investigated by supplementing glycine throughout the expression protocol, in different concentrations, and assessing the protein yields. This finding, even though it has not been fully investigated, already improved the recombinant expression system previously described for hGLYAT, which could aid in the studying of hGLYAT, since more protein could be made in a single expression experiment, thus saving time as well as valuable reagents.

It was determined that the BugBuster protein extraction method delivered slightly less protein than the Lysis buffer method, when comparing the hGLYAT bands on their respective SDS-PAGE gels. It was also found that the coN156ShGLYAT enzyme did not show improved expression or higher enzyme activity when compared to the non-codon optimised N156ShGLYAT. It was therefore decided to continue further investigations of the N156S hGLYAT enzyme using the non-codon optimised N156S hGLYAT variant in combination with the Lysis buffer extraction method, since this combination delivered the most soluble proteins.

The fourth objective was to optimise the storage conditions for the purified hGLYAT, since the current storage options resulted in significant reductions in enzyme activity and a long term storage solution was desperately needed. To solve this problem, an alternative storage method was developed in which 10% glycerol was added to the purified proteins

prior to snap-freezing, followed by storage at -80°C . A final objective was to determine whether the N156S hGLYAT variant remained enzymatically active after long term storage. Enzyme assays were performed in the presence of abundant substrate concentrations (200 mM glycine and 200 μM benzoyl-coenzyme A) over a period of 4 months and it was found that the storage solution successfully preserved enzyme activity. This finding significantly simplified the study of the hGLYAT enzyme, since it was previously impossible to work with this enzyme for longer than a few days at a time.

One of the long term goals of this study is to develop a therapeutic recombinant hGLYAT enzyme with altered substrate specificity, which could be used for the treatment of specific organic acidemias, as previously described in Section 2.11.2.3. Studying the kinetic parameters and substrate specificities of all the hGLYAT variants will provide necessary information to allow us to be able to design a variant of the hGLYAT enzyme with altered properties and substrate specificities.

The findings presented in this study will aid in this ultimate goal, since the recombinant hGLYAT expression system has been optimised and improved, thus shortening the time to make the proteins. Another valuable finding was the storage solution for hGLYAT. This will notably simplify the study of the enzymes properties, since it will be possible to conduct a range of enzymatic characterisation experiments on a single batch of proteins, without introducing inter-batch variations in the experiments.

Chapter 4: Bi-substrate kinetic analysis of the recombinant N156S hGLYAT variant

4.1 Introduction

After determining the concentration and activity of the N156S hGLYAT variant, the next step was to characterise the bi-substrate kinetic parameters using benzoyl-coenzyme A and glycine as substrates. In previous studies conducted in our laboratory only single substrate kinetic experiments were performed for GLYAT using benzoyl-coenzyme A as the only varying substrate and glycine at a saturating concentration of 200 mM. This step was simplified considerably by the optimisation of the storage conditions for purified hGLYAT over longer periods of time (described in Section 3.3.10). The abovementioned parameters were determined by using the colorimetric reaction with DTNB as described in Section 3.2.24 to detect the glycine dependent release of coenzyme A (Kølvraa & Gregersen, 1986). Even though this method is not very specific, it is reliable, relatively simple and accurate enough to be able to estimate the kinetic parameters of the enzyme.

To reach the aim of this part of the study, reaction mixtures containing purified N156S hGLYAT protein were prepared in the form of master mixes, with varying substrate concentrations. Different glycine concentrations ranging from 2.5 mM to 20 mM were added to the master mixes. These mixtures were then added to a 96 well plate, where-in different concentrations of benzoyl-coenzyme A (20 μ M – 80 μ M) were already pipetted, to start the reactions. The 96 well plate was immediately placed into a BioTek plate reader for absorbance measurements at 412 nm over 20 minutes. The absorbance results were then exported from the plate reader and analysed to identify the linear regions of the N156S hGLYAT variant activity and to calculate the activity over the linear region. The data was processed with the SigmaPlot 12.0 (Systat software, Inc., San Jose, CA) software, with the Enzyme Kinetics module, to plot the Michaelis-Menten and Lineweaver-Burk plots, and to calculate the kinetic parameters using nonlinear regression.

4.2 Materials and methods

4.2.1 Reagents, standards and solutions

To simplify the flow of the text, all reagents, standards and solutions used are listed in Appendix A, along with the catalogue numbers and manufacturers.

4.2.2 Preparation of standard stock solutions

All stock solutions and master mixes were prepared with molecular biology grade water. A stock solution of glycine was prepared by dissolving it in water to a final concentration of 1M. A stock solution of the benzoyl-coenzyme was also prepared to a final concentration of 10mM by dissolving 5 mg of the benzoyl-coenzyme A in 575 μ l water. All stock solutions were aliquoted and stored at -20°C.

Since the GLYAT enzyme requires a pH of 8.0, the 1 M stock of Tris used for the enzyme reaction was adjusted to a pH of 8.0 with hydrochloric acid. The solution was filter sterilised by passing through a sterile 0.2 μ m Acrodisc PF Syringe Filter

4.2.3 Benzoyl-coenzyme A standard curve

A standard curve for benzoyl-coenzyme A conjugation was set up by doing an end point assay of seven different benzoyl-coenzyme A concentrations (0 μ M – 180 μ M). This was important to find the coefficient of determination, r^2 , to assess how well the data fits into the statistical model of linear regression. This value is a reflection of how effectively benzoyl-coenzyme A was converted to hippuric acid in the presence of an abundance of glycine and the GLYAT enzyme. The gradient of the plot was used in the final processing of the kinetic data, for conversions between absorbance values (A_{412} /ml/ μ g) to nmol/ml/ μ g values.

Benzoyl-coenzyme A concentrations were tested using a saturating glycine concentration of 200 mM in 200 μ l reactions. A diluted stock solution of 2 mM benzoyl-coenzyme A was prepared to minimise errors. Final concentrations of [10 μ M, 20 μ M, 40 μ M, 80 μ M, 120 μ M, and 180 μ M] benzoyl-coenzyme A were used. To correct for background activity detected by the plate reader, a control reaction was also included in which 0 μ M benzoyl-coenzyme A was added.

4.2.3.1 Preparation of master mix solutions for the benzoyl coenzyme A range

To simplify the enzyme assay and to minimise inter-reaction variability, a 180 μ l master mix solution was prepared for each reaction according to the basic composition described in Section 3.2.24. To summarise, the master mix solution consisted of 20 μ l of 0.1 mM DTNB (diluted in 1 M TrisHCl), 40 μ l glycine with a final concentration of 200 mM, 110 μ l water, and 10 μ l of the purified N156S hGLYAT protein (1.2 μ g protein). This master mix was then mixed thoroughly, pipetted into three rows of seven wells of a 96-well plate, covered, and incubated at 37 °C while the next step was performed. Into a separate 96-well plate, 20 μ l of each of the different benzoyl-coenzyme A concentrations were added in triplicate as shown in Table 4.1.

Table 4.1: Plate layout for the enzyme reactions for the benzoyl-coenzyme A standard curve.

Benzoyl-coenzyme A concentrations (μ M)							
	1	2	3	4	5	6	7
A	0	10	20	40	80	120	180
B	0	10	20	40	80	120	180
C	0	10	20	40	80	120	180

The reactions were initiated by rapidly adding, with a multichannel pipette, 180 μ l of the preheated master mix solution to the 96-well plate containing the benzoyl-coenzyme A. After the addition of the master mix, the plate was covered, gently shaken, and incubated for 25 minutes at 37 °C. The absorbance (at 412 nm) values were measured with the BioTek plate reader and the accompanying Gen5 software.

4.2.4 Dilution of the purified N156S hGLYAT protein

Prior to kinetic studies, it was important to determine the necessary amount of protein to add to each enzyme reaction which would result in a linear change of absorbance over time. A pilot study was performed with different dilutions of the protein stock solution in different enzyme reactions containing either 15 or 150 μ M benzoyl-coenzyme A with 200 mM glycine. This experiment was also necessary because the enzyme activities are so easily influenced by minor experimental differences such as handling and temperature variations. This dilution step is thus included to assess how fast the diluted protein reacts with the provided high and low substrate concentrations, and to ensure that the enzyme activity occurs at a rate that results in an activity plot indicating a clear linear region over a few minutes (approximately 5 minutes).

For example, if the enzyme is left undiluted, the protein content of the sample may be too high in the presence of high substrate concentrations, resulting in a plot from which a linear region over a few minutes cannot be chosen. If the sample is too diluted or the substrate concentration is too low, the activity may be too low to find a linear region from the plot, and further deductions from this data may thus not be reliable. Figure 4.1 illustrates these ideal and non-ideal cases.

In many cases the enzyme dilution test also enables the researcher to use the enzyme sparingly, because the ideal amount necessary for kinetic assays can be determined, thus avoiding too many repeat experiments with the valuable enzymes.

Five different dilutions of the purified N156S hGLYAT were prepared by further diluting the samples in buffer EB (10 mM Tris-Cl, pH 8.5). Enzyme:EB buffer dilutions of 1:4, 1:3, 1:2, 1:1 and an undiluted sample were prepared. A control reaction containing no enzyme was also included to correct for any background activity detected.

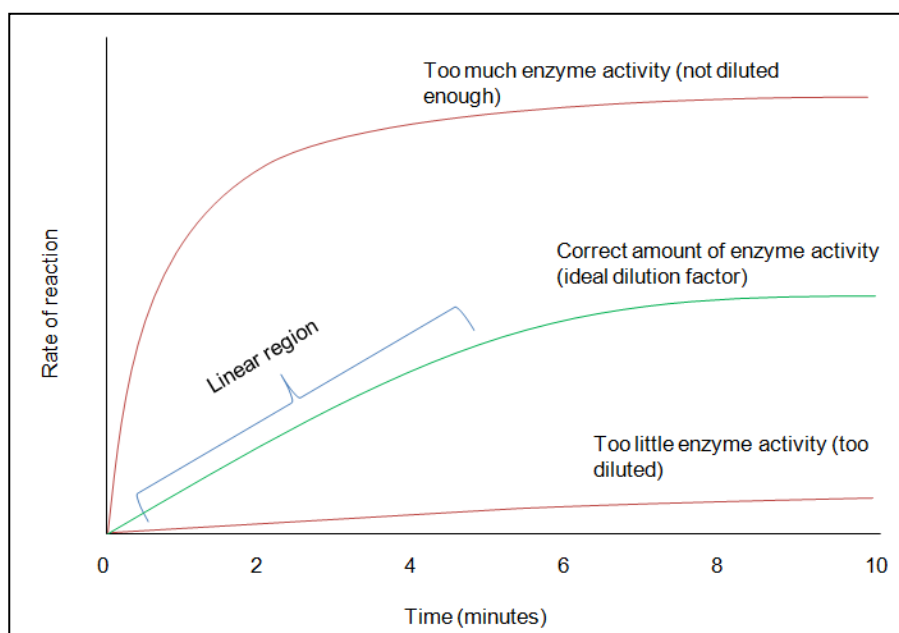


Figure 4.1: Graphic representation of the activity plots of enzymes with different dilution factors.

Each line represents a reaction with a certain amount of enzyme present. The red lines are indicative of too little or too high enzyme content, while the green line is representative of a reaction with the ideal amount of enzyme present, thus resulting in a linear region over a few minutes.

4.2.4.1 Preparation of master mix solution for enzyme dilution assay

To set up the reactions for testing the different dilution factors, two master mix solutions were prepared according to the composition described in Section 3.2.24, with some minor adjustments. The reaction mixtures were 100 μ l in volume and consisted of 0.01 mM DTNB (diluted in Tris-Cl), 20 μ l glycine with a final concentration of 20 mM, 30 μ l water, and 10 μ l of benzoyl-coenzyme A (15 μ M or 150 μ M). The master mixes were then mixed, pipetted

into two rows of six wells of a 96-well plate, covered, and incubated at 37 °C. Into a separate 96-well plate, 10 µl of each of the different enzyme dilutions were added in duplicate as indicated in Table 4.2.

The reactions were initiated by adding, with a multichannel pipette, 90 µl of the preheated master mix solution to the 96-well plate containing the N156S hGLYAT enzyme dilutions. After the addition of the master mix, the plate was covered, gently shaken, and incubated for 25 minutes at 37 °C. The absorbance (at 412 nm) values were measured with the BioTek plate reader and the accompanying Gen5 software.

Table 4.2: Plate layout for the enzyme reactions prepared for the enzyme dilution experiment with high and low benzoyl-coenzyme A concentrations.

High benzoyl-coenzyme A concentrations (150 µM)						
	1	2	3	4	5	6
A	0	1:4	1:3	1:2	1:1	Undiluted
B	0	1:4	1:3	1:2	1:1	Undiluted
Low benzoyl-coenzyme A concentrations (15 µM)						
	1	2	3	4	5	6
A	0	1:4	1:3	1:2	1:1	Undiluted
B	0	1:4	1:3	1:2	1:1	Undiluted

4.2.5 Kinetic assays of the N156S hGLYAT variant

To determine the bi-substrate kinetic parameters of the N156S hGLYAT enzyme, an enzyme assay was performed where both substrate concentrations were varied. The glycine

concentrations were varied from 2.5-20 mM at constant benzoyl-coenzyme A concentrations. The benzoyl-coenzyme A concentrations used were 20-80 μM at constant glycine concentrations. Table 4.3 summarises the plate layout of the enzyme reactions used for the bi-substrate kinetics of the N156S hGLYAT protein, indicating the final substrate concentrations present in each reaction mixture.

Table 4.3: The plate layout of the enzyme reactions used for the bi-substrate kinetics of the N156S hGLYAT variant, indicating the final substrate concentrations present in each reaction mixture.

Substrate concentrations		1	2	3	4	5	6	7	8	9	10	11	12
A	Benzoyl CoA (μM)	20	20	20	20	20	20	20	20	20	20	20	20
	Glycine (mM)	2.5	2.5	2.5	5	5	5	10	10	10	20	20	20
B	Benzoyl CoA (μM)	40	40	40	40	40	40	40	40	40	40	40	40
	Glycine (mM)	2.5	2.5	2.5	5	5	5	10	10	10	20	20	20
C	Benzoyl CoA (μM)	60	60	60	60	60	60	60	60	60	60	60	60
	Glycine (mM)	2.5	2.5	2.5	5	5	5	10	10	10	20	20	20
D	Benzoyl CoA (μM)	80	80	80	80	80	80	80	80	80	80	80	80
	Glycine (mM)	2.5	2.5	2.5	5	5	5	10	10	10	20	20	20

4.1.5.1 Preparation of master mix solutions for enzyme kinetics assay

To ensure accurate and reliable data, and to minimise inter-reaction variability, master mix solutions were once again prepared for each reaction. For this experiment, four separate master mixes were prepared, one for each of the glycine concentrations. These master mixes consisted of 0.01 mM DTNB (diluted in Tris-Cl), 40 μl of each glycine concentration

(final concentrations of 2.5, 5, 10, and 20 mM), 110 µl water, and 10 µl of the 1:4 diluted, purified N156S hGLYAT enzyme (1.2 µg). Each of the master mix solutions with different glycine concentrations was pipetted into 12 wells, the first 4 wells of a column (in triplicates) on a 96-well plate. The plate was then covered, and placed in an incubator at 37 °C while the second plate containing the benzoyl-coenzyme A was prepared. Into a separate 96-well plate, 20 µl of each of the different benzoyl coenzyme A concentrations were added to a row in all 12 wells horizontally as shown in Table 4.3.

The plates were taken to the plate reader before mixing, as the absorbance measurements needed to be taken as soon as possible after the initiation of the reactions. After setting up the method on the plate reader and preheating the reader to 37 °C, the reactions were started by rapidly adding, with the multichannel pipette, 180 µl of the preheated master mix solutions to the 96-well plate containing the human GLYAT dilutions. After the addition of the master mix, the plate was immediately placed into the BioTek plate reader with which absorbance (at 412 nm) was measured every 41 seconds for 20 minutes. The accompanying Gen5 software was used to process the data for further interpretation.

4.2.6 Data analysis for calculation of kinetic parameters

The A_{412} versus time data was exported to Microsoft Excel and plots were drawn to indicate the changes in absorbance over 20 minutes. Equation 4.1 was then used to calculate the initial velocities. Firstly, the linear regions were identified over 6 minutes and their slopes calculated. Figure 4.3 describes the relationship between the slope of the tangent and the initial velocity.

Equation 4.1: The equation used to calculate the initial velocity of enzyme reactions

$$\text{Initial velocity} = \frac{\Delta A_{412}}{\Delta \text{time (minutes)}}$$

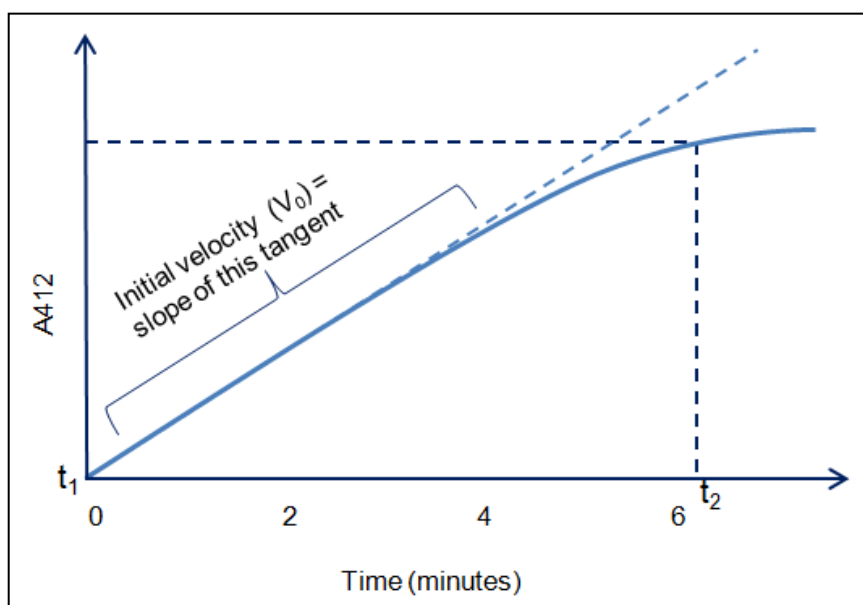


Figure 4.2: A plot illustrating the absorbance (412 nm) values over time. This plot shows how initial velocities are calculated from the original enzyme activity plots. The slope of the linear region (in this example up to approximately 6 minutes) is used to calculate the initial velocity. Adapted from Palmer & Bonner (2007).

To calculate the initial velocities the change in absorbance per minute was firstly converted to absorbance/min/ μg by dividing with the amount of protein added to each reaction (Equation 4.2). (Section 4.3.2 describes how it was determined that the optimal hGLYAT protein content for kinetic studies is 1.2 μg . For now, assume protein content as this value). These values were then multiplied with the gradient of the benzoyl-coenzyme A standard curve (0.00495) to convert values to $\mu\text{mol}/\text{min}/\mu\text{g}$ (Equation 4.3). The values were finally multiplied by 1 000 000 to convert to $\text{nmol}/\text{min}/\text{mg}$ (Equation 4.4).

Equation 4.2: Formula for conversion from initial velocity (absorbance/min) to absorbance/min/ μg .

$$\frac{\text{Absorbance}}{\text{min}} / \mu\text{g} = \frac{\text{Initial velocity}}{1.2 \mu\text{g}}$$

Equation 4.3: Formula for conversion from absorbance/min/ μ g to μ mol/min/ μ g.

$$\frac{\frac{\text{Absorbance}}{\text{min}}}{1.2 \mu\text{g}} \times 0.00495 = \frac{\mu\text{mol}}{\text{min}} / \mu\text{g}$$

Equation 4.4: Formula for conversion from μ mol/min/ μ g to nmol/min/mg.

$$\frac{\frac{\mu\text{mol}}{\text{min}}}{\mu\text{g}} \times 1\,000\,000 = \frac{\text{nmol}}{\text{min}} / \text{mg}$$

After the initial velocities were calculated, they were used to draw the Lineweaver-Burk (double reciprocal) plots. These plots were used for the estimation of the kinetic parameters. The SigmaPlot 12.0 software program was used in conjunction with its Enzyme Kinetics module to plot the experimentally determined values automatically, and to perform a linear regression analysis to calculate the estimated K_m and $V_{\max}$ values with their standard deviations. These values can be read from the Lineweaver-Burk plots from the x- and y-axes respectively.

4.3 Results and discussion

4.3.1 Benzoyl-coenzyme A standard curve

A standard curve for benzoyl-coenzyme A conjugation was set up by doing an end point assay of seven different benzoyl-coenzyme A concentrations (0 μ M – 180 μ M) with saturating glycine concentrations of 200 mM in 200 μ l reactions.

The reactions were prepared in a 96-well plate, in triplicate, and incubated for 25 minutes at 37 °C with gentle shaking. The end point absorbance (at 412 nm) values were measured with the Biotech plate reader and the accompanying Gen5 software. Figure 4.3 shows the plot of the absorbance (412 nm) values versus the benzoyl-coenzyme A concentrations,

which indicated the linearity of the coenzyme A release from the reactions. The coefficient of determination, r^2 , was determined as 0.9903. This indicated that the data fit the statistical model of linear regression almost perfectly. The linearity of the plot allowed for the gradient of the plot to be used in the final processing of the kinetic data.

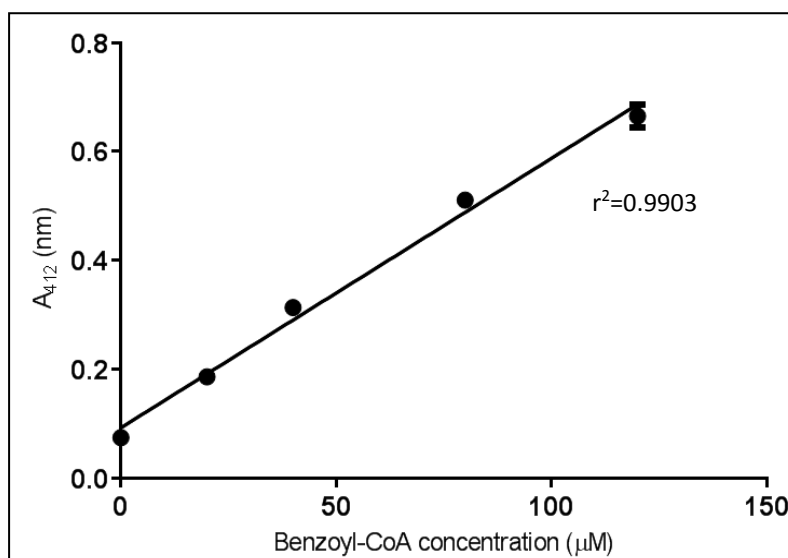


Figure 4.3: Standard curve of absorbance (412 nm) versus benzoyl-coenzyme A concentrations. The r^2 is 0.9903, indicating a near linear correlation.

4.3.2 Enzyme assay with dilution range of the purified N156S hGLYAT protein

The enzyme reactions were set up with different dilutions of enzyme present in the reactions. This step was necessary to determine the optimal enzyme dilution factor, which would result in enzyme activity plots indicating a clear linear region over at least a few minutes. A stock solution of purified N156S hGLYAT protein with a concentration of 598.66 μg/ml was used to prepare five different dilutions of the protein by mixing the samples with buffer EB. Enzyme:EB buffer dilution ratios of 1:4, 1:3, 1:2, 1:1 and an undiluted sample were prepared, as well as a control reaction containing no enzyme to correct for background activity. Thus, the dilutions had protein concentrations of 119.73, 149.67, 199.5, 299.33, and 598.66 μg/ml respectively.

The reactions were prepared as stated in Section 4.2.4.1 and kinetic absorbance values (at

412 nm) were measured over 20 minutes with the BioTek plate reader and the accompanying Gen5 software. Figure 4.4 shows the graph of the absorbance values of each of the dilutions plotted against time. The undiluted enzyme resulted in a plot from which a linear region over a few minutes could not be identified, as the enzyme activity was too high. The 1:1, 1:2, and 1:3 dilutions resulted in similar, but indicated that those protein concentrations were still too high. The 1:4 dilution of the purified N156S hGLYAT protein delivered a curve with a much more useable linear region over a period of up to 8 minutes. This dilution factor was chosen for further enzyme kinetic analyses conducted with the same batch of enzymes.

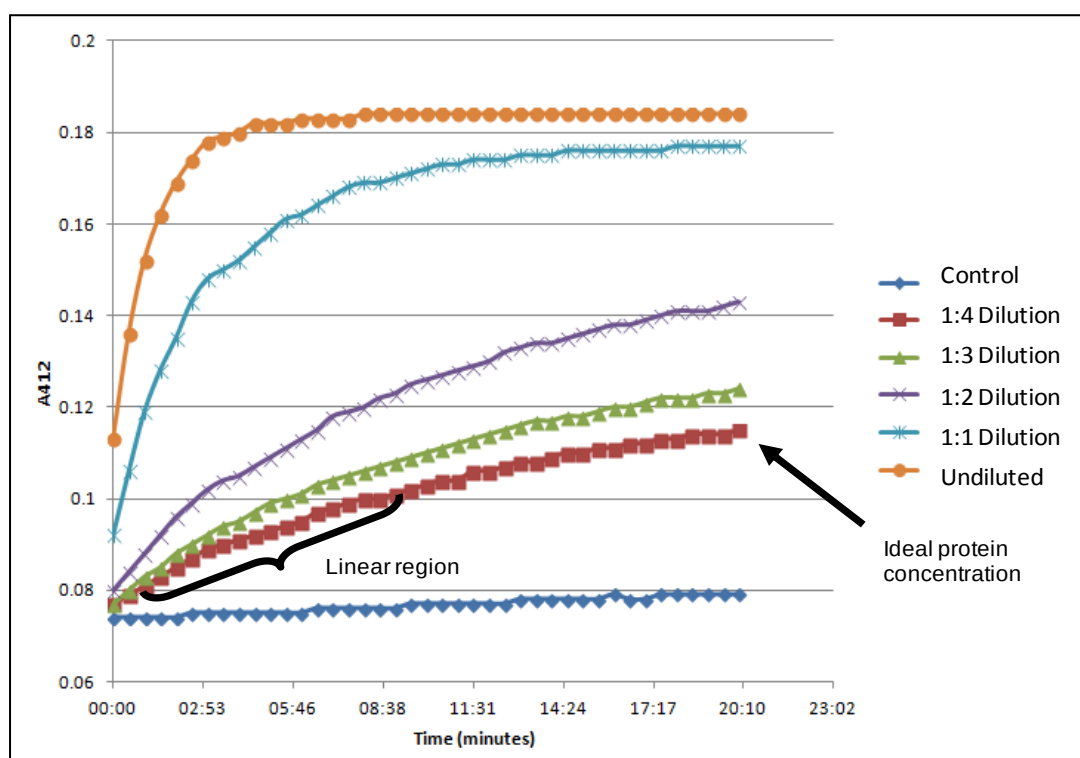


Figure 4.4: The enzyme activity plots of the purified N156S recombinant enzymes with different dilution factors. Five enzyme dilutions were tested including 1:1, 1:2, 1:3, 1:4, and 1:5 dilution factors. A negative control was included and served as the baseline reading. The graph shows the change in absorbance at 412 nm over 20 minutes. The red line (1:4 dilution) represents the dilution factor with the ideal amount of enzyme present as a clear linear region over several minutes can be seen.

4.3.3 Calculation of bi-substrate kinetic parameters of the recombinant N156S hGLYAT variant and selection of an enzyme kinetic model

To determine the bi-substrate kinetic parameters of the N156S hGLYAT, assays were performed with both glycine and benzoyl-coenzyme A concentrations being varied. The glycine concentrations were varied from 2.5-20 mM at constant benzoyl-coenzyme A concentrations and the benzoyl-coenzyme A concentrations were varied from 20-80 μ M at constant glycine concentrations. Table 4.3 summarised the plate layout of the enzyme reactions used for the bi-substrate kinetics of the N156S hGLYAT protein.

The reactions, prepared in triplicate, were initiated by addition of the substrate containing

master mix solutions, after which the 96-well plate was placed into the BioTek plate reader to measure the absorbance values (at 412 nm) every 41 seconds over 20 minutes. The accompanying Gen5 software was used to process the data for further interpretation. The raw results obtained from the plate reader software were plotted as the absorbance values over a time course of 20 minutes. To save space and to avoid redundancy, only one example was chosen to illustrate how the linear regions were calculated. Table 4.4 contains an example of the raw data from the enzyme reaction with 60 μ M benzoyl-coenzyme A from which plots were drawn to identify the time period over which a linear region is present. Figure 4.5 shows the enzyme activity plots of the purified N156S hGLYAT enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration of 60 μ M over 20 minutes. From these plots, linear regions could be identified for the reactions and linear regions of 6 minutes were chosen for this study. These linear regions are indicated in red, along with their respective r^2 values.

Table 4.4: Raw data of absorbance values of the enzyme reaction with 60 μ M benzoyl-coenzyme A obtained from the plate reader over 20 minutes. These values were plotted on graphs to identify the time period over which a linear region is notable. All measurements were done in triplicate (continued on the next page).

Time (minutes)	Absorbance values											
	Glycine (2.5 mM)			Glycine (5 mM)			Glycine (10 mM)			Glycine (20 mM)		
00:00	0.071	0.071	0.07	0.073	0.073	0.078	0.078	0.078	0.078	0.086	0.085	0.086
00:40	0.073	0.074	0.073	0.078	0.078	0.082	0.085	0.085	0.085	0.096	0.095	0.097
01:20	0.077	0.077	0.076	0.083	0.082	0.087	0.093	0.092	0.093	0.106	0.105	0.107
02:00	0.079	0.08	0.079	0.088	0.087	0.091	0.100	0.099	0.100	0.116	0.114	0.116
02:40	0.082	0.083	0.082	0.092	0.092	0.096	0.106	0.105	0.106	0.126	0.124	0.123
03:20	0.084	0.085	0.085	0.097	0.096	0.100	0.113	0.111	0.113	0.135	0.133	0.129
04:00	0.087	0.087	0.088	0.101	0.100	0.104	0.119	0.117	0.118	0.144	0.142	0.135
04:40	0.089	0.089	0.09	0.105	0.104	0.108	0.125	0.122	0.124	0.153	0.151	0.141
05:20	0.090	0.091	0.092	0.109	0.107	0.112	0.131	0.128	0.129	0.161	0.158	0.148
06:00	0.092	0.094	0.095	0.112	0.110	0.116	0.136	0.133	0.133	0.167	0.165	0.155

Table 4.4: (Continued)

Time (minutes)	Absorbance values											
	Glycine (2.5 mM)			Glycine (5 mM)			Glycine (10 mM)			Glycine (20 mM)		
06:40	0.093	0.096	0.097	0.116	0.114	0.119	0.140	0.138	0.138	0.175	0.172	0.161
07:20	0.094	0.098	0.099	0.119	0.117	0.121	0.143	0.142	0.143	0.182	0.177	0.168
08:00	0.095	0.100	0.1	0.122	0.120	0.124	0.147	0.146	0.148	0.188	0.182	0.174
08:40	0.096	0.103	0.102	0.126	0.124	0.127	0.152	0.151	0.154	0.194	0.187	0.181
09:20	0.097	0.105	0.105	0.129	0.126	0.13	0.156	0.155	0.158	0.200	0.19	0.187
10:00	0.098	0.107	0.107	0.132	0.129	0.133	0.160	0.159	0.161	0.205	0.195	0.192
10:40	0.100	0.109	0.108	0.135	0.131	0.135	0.164	0.163	0.164	0.21	0.199	0.198
11:20	0.102	0.110	0.110	0.138	0.134	0.139	0.168	0.167	0.168	0.214	0.203	0.203
12:00	0.104	0.112	0.112	0.14	0.136	0.142	0.172	0.170	0.173	0.219	0.209	0.208
12:40	0.106	0.114	0.114	0.143	0.139	0.143	0.176	0.174	0.177	0.224	0.215	0.213
13:20	0.107	0.115	0.115	0.145	0.141	0.146	0.18	0.179	0.181	0.228	0.220	0.218
14:00	0.108	0.117	0.117	0.147	0.144	0.150	0.183	0.182	0.184	0.233	0.226	0.222
14:40	0.109	0.119	0.118	0.150	0.146	0.153	0.186	0.185	0.187	0.237	0.230	0.226
15:20	0.110	0.120	0.120	0.152	0.149	0.155	0.190	0.187	0.191	0.241	0.235	0.230
16:00	0.112	0.121	0.12	0.154	0.151	0.156	0.193	0.190	0.194	0.245	0.239	0.234
16:40	0.113	0.123	0.121	0.156	0.153	0.159	0.196	0.193	0.197	0.249	0.244	0.237
17:20	0.114	0.124	0.123	0.158	0.156	0.161	0.199	0.195	0.199	0.253	0.247	0.241
18:00	0.116	0.125	0.124	0.160	0.158	0.163	0.202	0.198	0.203	0.257	0.251	0.244
18:40	0.116	0.126	0.125	0.162	0.160	0.165	0.205	0.201	0.205	0.260	0.254	0.247
19:20	0.117	0.128	0.127	0.164	0.162	0.166	0.207	0.204	0.207	0.263	0.257	0.250
20:00	0.118	0.129	0.127	0.167	0.163	0.168	0.210	0.207	0.210	0.266	0.260	0.253
Δ Absorbance	0.021	0.023	0.025	0.039	0.037	0.038	0.058	0.055	0.055	0.081	0.08	0.069

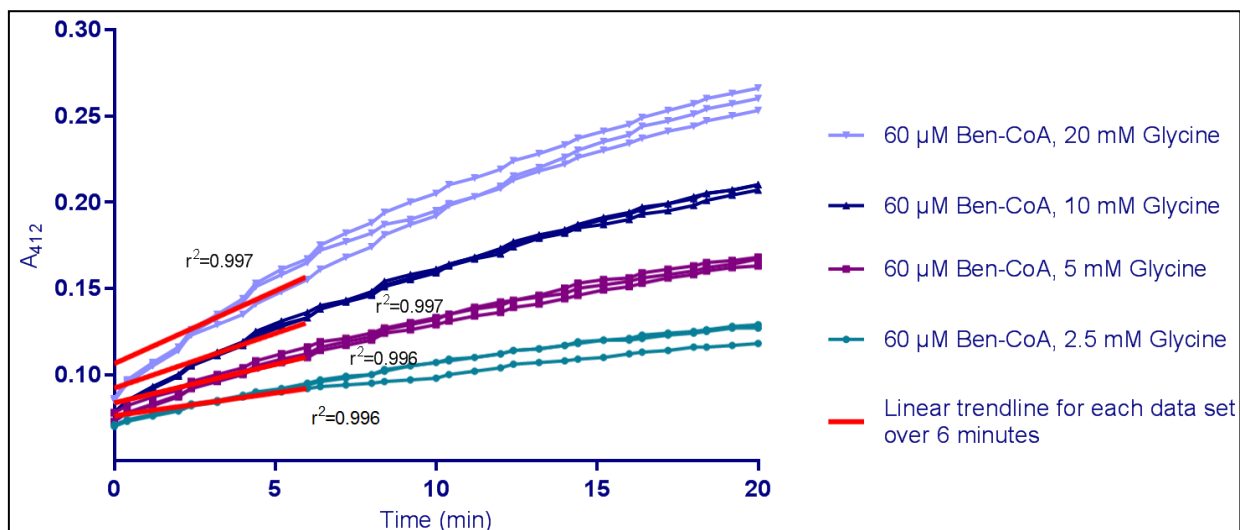


Figure 4.5: The enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration (60 μM). Reactions were measured in triplicate. The graph shows the change in absorbance at 412 nm over 20 minutes.

Figures 4.6-4.9 show the linear regions of the enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with varying glycine concentrations (2.5 – 20 mM) at the different benzoyl-coenzyme A concentrations (20, 40, 60, and 80 μM respectively) over 6 minutes. Reactions were measured in triplicate. The graphs show the changes in absorbance at 412 nm over six minutes for each reaction. The figure legends indicate the substrate concentrations used for each reaction.

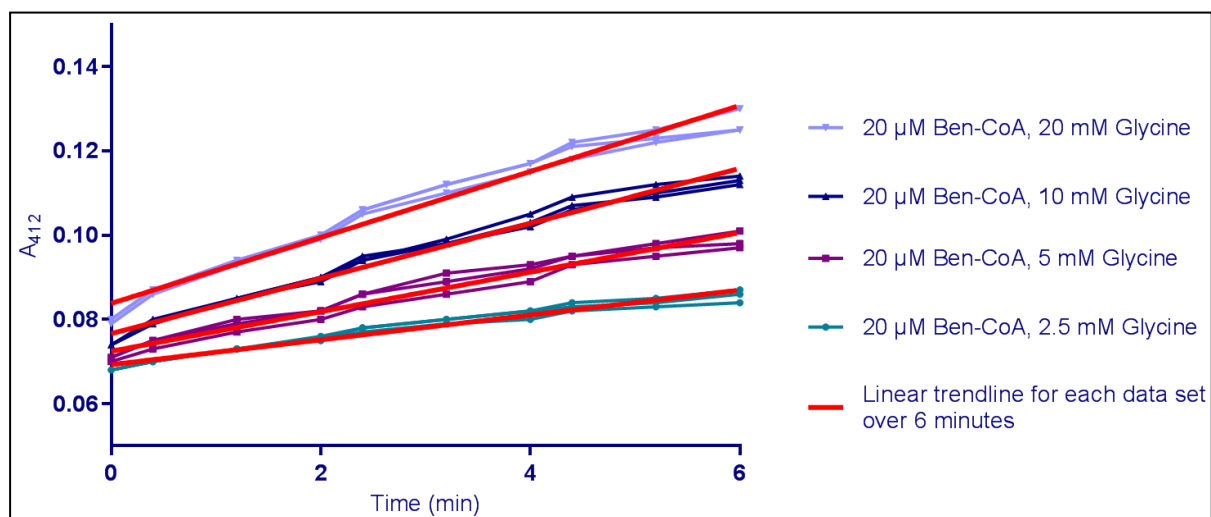


Figure 4.6: The enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration (20 μ M). Reactions were measured in triplicate. The graph shows the change in absorbance at 412 nm over six minutes. A clear linear region over the timeframe can be seen.

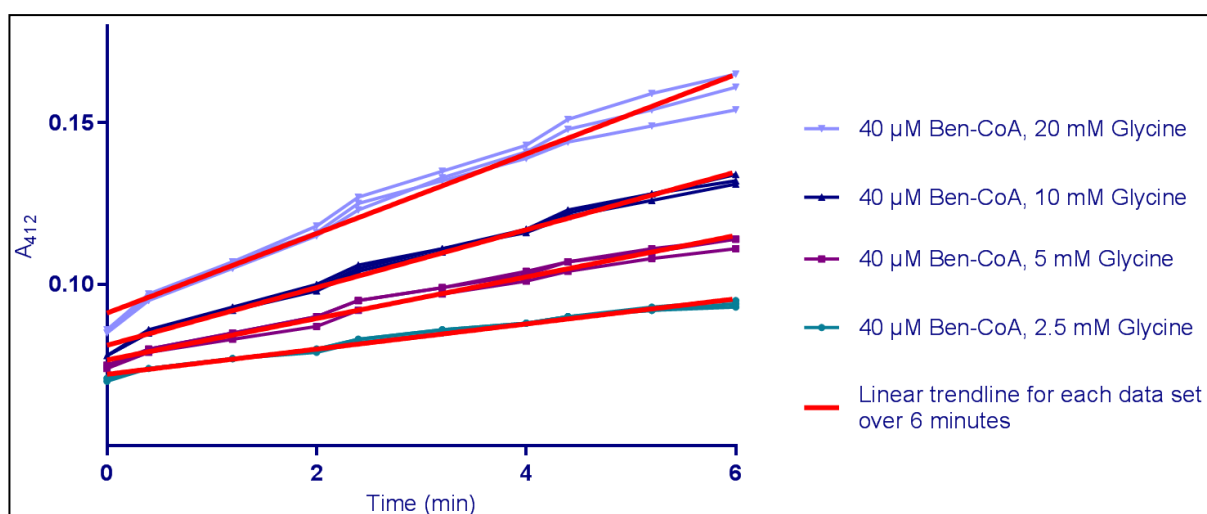


Figure 4.7: The enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration (40 μ M). Reactions were measured in triplicate. The graph shows the change in absorbance at 412 nm over six minutes. A clear linear region over the timeframe can be seen.

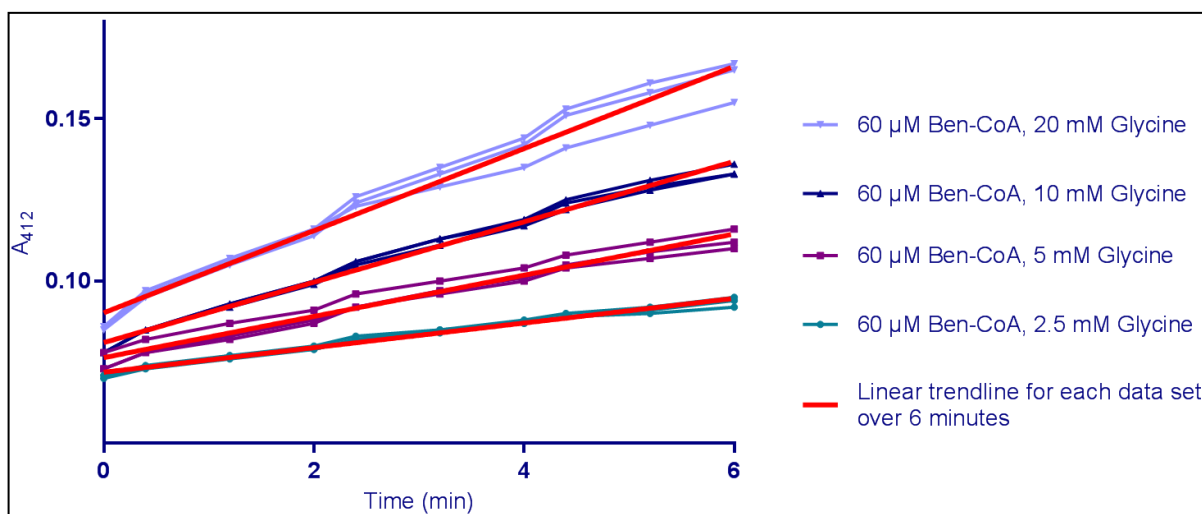


Figure 4.8: The enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration (60 μ M). Reactions were measured in triplicate. The graph shows the change in absorbance at 412 nm over six minutes. A clear linear region over the timeframe can be seen.

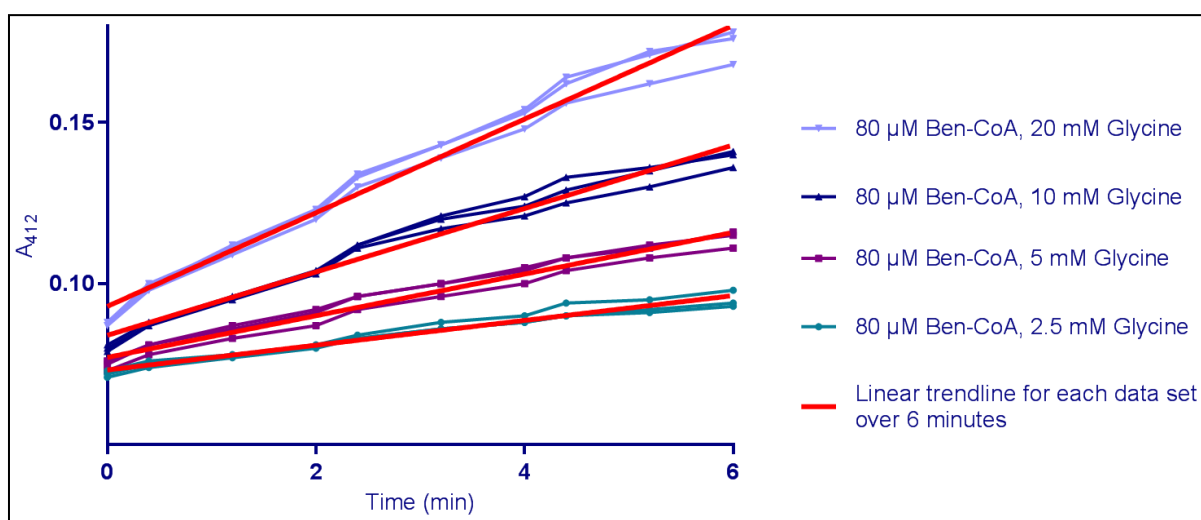


Figure 4.9: The enzyme activity plots of the purified N156S hGLYAT recombinant enzymes with different glycine concentrations (2.5 – 20 mM) at a constant benzoyl-coenzyme A concentration (80 μ M). Reactions were measured in triplicate. The graph shows the change in absorbance at 412 nm over six minutes. A clear linear region over the timeframe can be seen.

For the determination of the initial velocities, Equation 4.1, was used to calculate the change in absorbance over six minutes. Table 4.5 summarises the initial velocity (V_0) values determined for all reactions.

The initial velocities per μg protein were then calculated using Equation 4.2, in which the initial velocity values obtained from Equation 4.1 are divided by the amount of protein present in each reaction (determined as $1.2 \mu\text{g}$ in Section 4.3.2). This calculation then converts the values to absorbance/min/ μg . These values were then multiplied with the gradient of the benzoyl-coenzyme A standard curve (0.00495) to convert values to $\mu\text{mol/min}/\mu\text{g}$ as stated in Equation 4.3. Equation 4.4 was finally used to convert the values to the final unit of nmol/min/mg by multiplying with 1 000 000. These values, summarised in Table 4.6, were then used to calculate the kinetic parameters using the Enzyme Kinetics module of Sigma Plot.

Table 4.5: Initial velocities (V_0) calculated from the change in A_{412} values over 6 minutes. All measurements were done in triplicate. These values were calculated using Equation 4.1.

Benzoyl-coenzyme A (μM)	Glycine (mM)	Triplicate reactions initial velocities (V_0)		
		[ΔA_{412}^* (nm)/Time (6 minutes)]		
20	2.5	0.0027	0.003	0.0032
40		0.0037	0.0038	0.0042
60		0.0035	0.0038	0.0042
80		0.0035	0.0038	0.0042
20	5	0.0045	0.0045	0.005
40		0.0065	0.0062	0.0067
60		0.0065	0.0062	0.0063
80		0.0067	0.0063	0.0067
20	10	0.0067	0.0065	0.0063
40		0.0093	0.0088	0.009
60		0.0097	0.0092	0.0092
80		0.0098	0.0102	0.0095
20	20	0.0083	0.0077	0.0075
40		0.0132	0.0127	0.0113
60		0.0135	0.0133	0.0115
80		0.0147	0.015	0.0135

* ΔA_{412} – The change in absorbance at 412 nm.

Table 4.6: Processed data set of enzyme assay results used for calculation of kinetic parameters using SigmaPlot with the Enzyme Kinetics module. Values were obtained using Equations 4.2-4.4.

Benzoyl-coenzyme A (μM)	Glycine (mM)	Triplicate reactions values (nmol/min/mg)		
20	2.5	11	12.375	13.0625
40		15.125	15.8125	17.1875
60		14.4375	15.8125	17.1875
80		14.4375	15.8125	17.1875
20	5	18.5625	18.5625	20.625
40		26.8125	25.4375	27.5
60		26.8125	25.4375	26.125
80		27.5	26.125	27.5
20	10	27.5	26.8125	26.125
40		38.5	36.4375	37.125
60		39.875	37.8125	37.8125
80		40.5625	41.9375	39.1875
20	20	34.375	31.625	30.9375
40		54.3125	52.25	46.75
60		55.6875	55	47.4375
80		60.5	61.875	55.6875

Lineweaver-Burk (double reciprocal) plots were generated from these experimentally determined values with the SigmaPlot 12.0 software program using its Enzyme kinetics module. These plots were used for the automatic calculation of the K_m and V_{max} values for benzoyl-coenzyme A and glycine for the N156S hGLYAT variant using nonlinear regression analysis. These values were determined individually by plotting the constant glycine values with the varying benzoyl-coenzyme A values on a double reciprocal plot and vice versa on a

second double reciprocal plot. The software program could also plot the values according to different kinetic models, so that these models could be compared to one another. Three kinetic models, namely the random ordered mechanism, the ordered mechanism, and the ping-pong mechanism were compared in this study. The double reciprocal plots of both the benzoyl-coenzyme A and glycine values are shown in Figures 4.10, 4.11 and 4.12. These figures represent the random order kinetic model, the ordered kinetic model and the ping-pong bi-bi model, respectively. The figure legends indicate the substrate concentrations of the constant substrate used in each reaction.

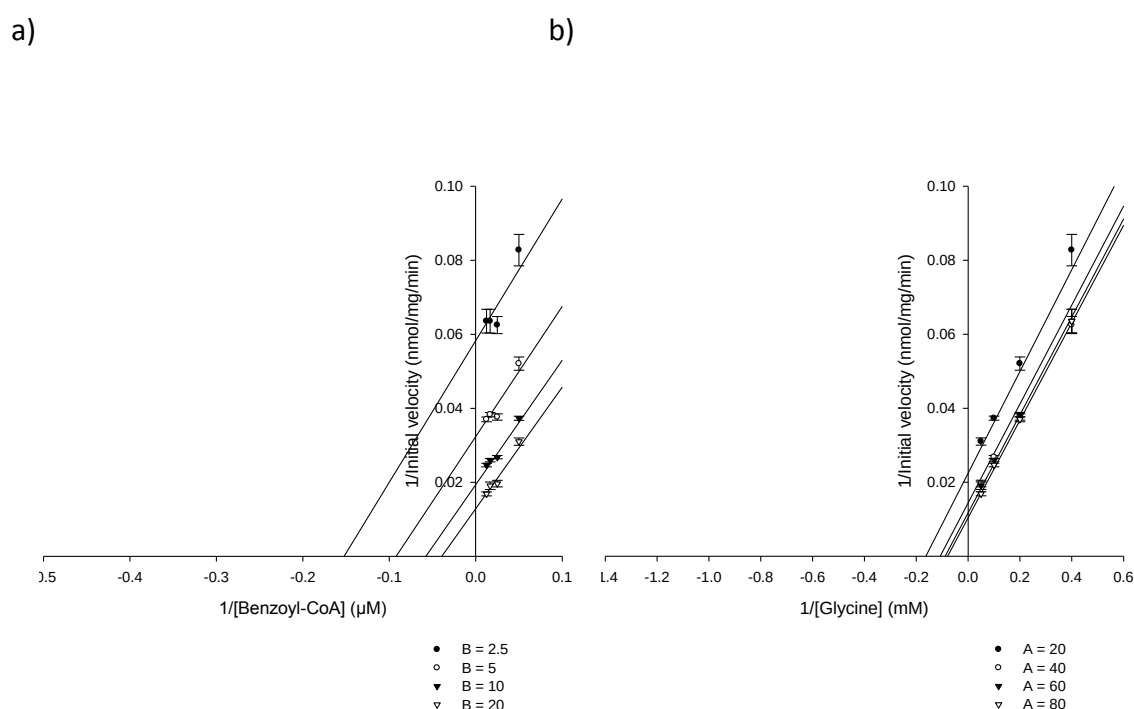


Figure 4.10: Random order sequential mechanism Lineweaver-Burk plots to determine the K_m values of benzoyl-coenzyme A and glycine for the purified N156S hGLYAT recombinant enzymes with different glycine (2.5 , 5, 10, and 20 mM) and benzoyl-coenzyme A (20, 40, 60, and 80 μ M) concentrations. The data points indicate average values $\pm$ standard deviation, with $n = 3$. a) For this plot, the glycine concentrations were kept constant, and benzoyl-coenzyme A concentrations were varied. This plot was used to calculate the K_m of benzoyl-coenzyme A. b) For this plot, benzoyl-coenzyme A concentrations were kept constant, and the glycine concentrations were varied. This plot was used to calculate the K_m of glycine.

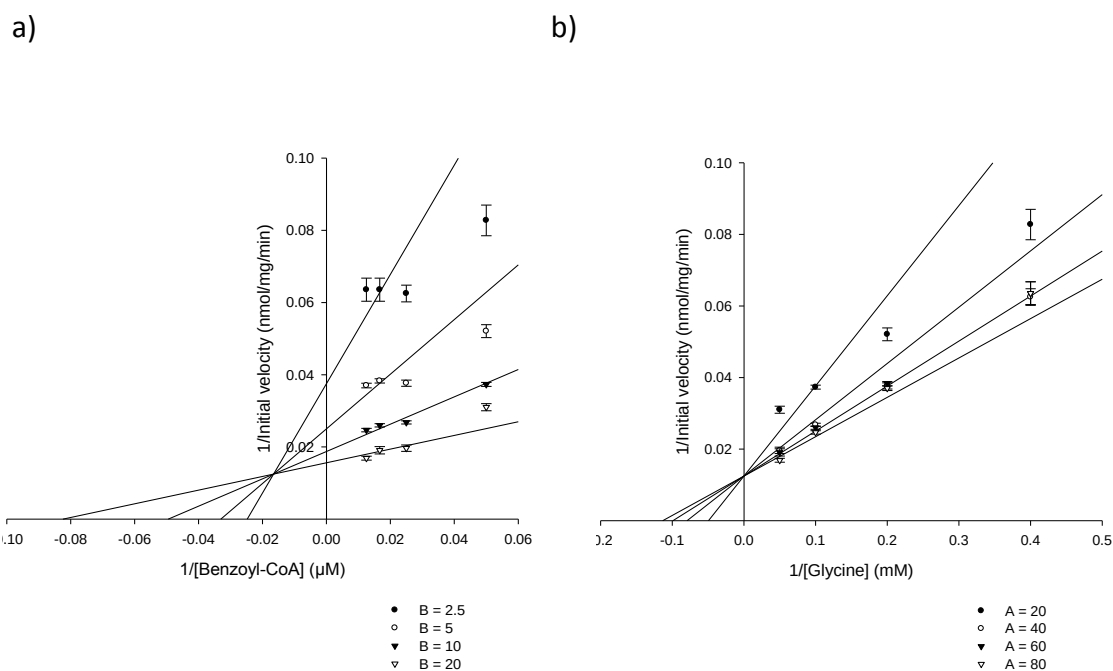


Figure 4.11: Ordered sequential mechanism Lineweaver-Burk plots to determine the K_m values of benzoyl-coenzyme A and glycine for the purified N156S hGLYAT recombinant enzymes with different glycine (2.5 , 5, 10, and 20 mM) and benzoyl-coenzyme A (20, 40, 60, and 80 μM) concentrations. The data points indicate average values $\pm$ standard deviation, with $n = 3$. a) For this plot, the glycine concentrations were kept constant, and benzoyl-coenzyme A concentrations were varied. This plot was used to calculate the K_m of benzoyl-coenzyme A. b) For this plot, benzoyl-coenzyme A concentrations were kept constant, and the glycine concentrations were varied. This plot was used to calculate the K_m of glycine.

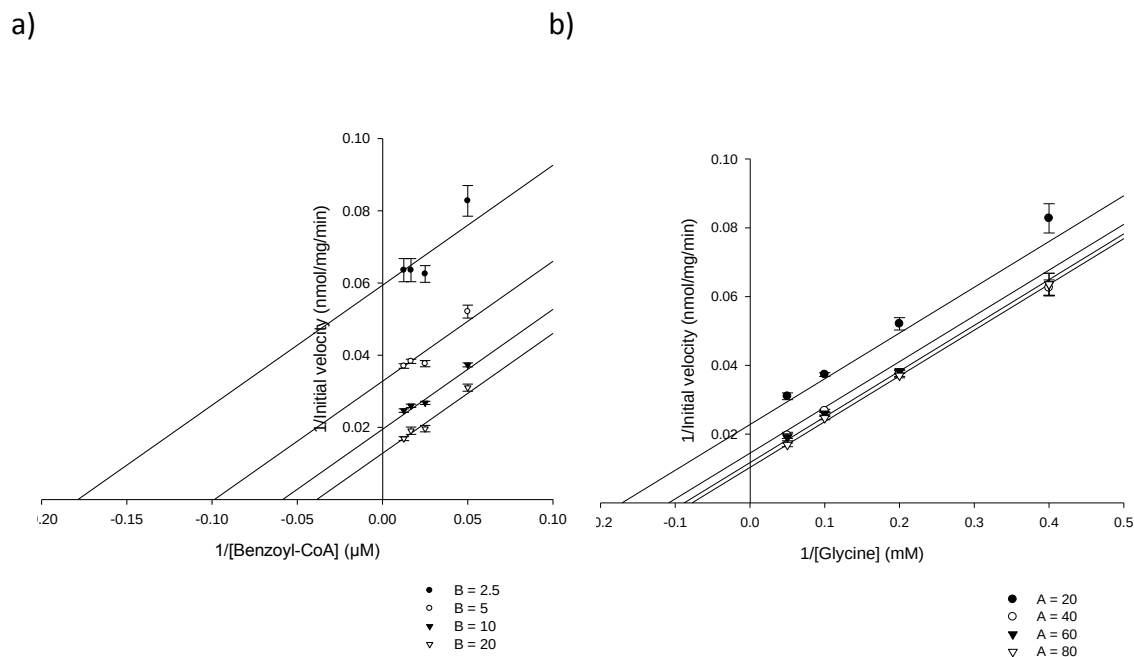


Figure 4.12: Ping-pong Bi-Bi mechanism Lineweaver-Burk plots to determine the K_m values of benzoyl-coenzyme A and glycine for the purified N156S hGLYAT recombinant enzymes with different glycine (2.5 , 5, 10, and 20 mM) and benzoyl-coenzyme A (20, 40, 60, and 80 μM) concentrations. The data points indicate average values $\pm$ standard deviation, with $n = 3$. a) For this plot, the glycine concentrations were kept constant, and benzoyl-coenzyme A concentrations were varied. This plot was used to calculate the K_m of benzoyl-coenzyme A. b) For this plot, benzoyl-coenzyme A concentrations were kept constant, and the glycine concentrations were varied. This plot was used to calculate the K_m of glycine.

Different kinetic parameters were obtained when processing the kinetic data with different kinetic models. When the data points were plotted according to the random order kinetic model, the K_m values for benzoyl-coenzyme A and glycine were $49 \pm 13 \mu\text{M}$ and $20 \pm 4 \text{ mM}$, respectively. When the data points were plotted according to the ordered mechanism kinetic model, the K_m values for benzoyl-coenzyme A and glycine were $60 \pm 16 \mu\text{M}$ and $5 \pm 1 \text{ mM}$, respectively. When the data points were plotted according to the ping-pong mechanism kinetic model, the K_m values for benzoyl-coenzyme A and glycine were $54 \pm 8 \mu\text{M}$ and $22 \pm 3 \text{ mM}$, respectively. These kinetic parameters according to each of the models are summarised in Table 4.7.

Table 4.7: Kinetic parameters obtained by the different kinetic mechanisms for the N156S hGLYAT enzyme using benzoyl-coenzyme A and glycine as substrates.

Kinetic model	K_m benzoyl-coenzyme A (μM)	K_m glycine (mM)	Relative $V_{\max}$ (nmol/min)
Random ordered mechanism	49 $\pm$ 13	20 $\pm$ 4	156
Ordered mechanism	60 $\pm$ 16	5 $\pm$ 1	80
Ping-pong mechanism	54 $\pm$ 8	22 $\pm$ 3	162

4.3.4 Kinetic model of human GLYAT

In this section the available literature and our results are used to predict the type of kinetic model that might be employed by human GLYAT. The three dimensional structure of an enzyme, ultimately determines its properties.

The GNAT superfamily of acyltransferases, of which GLYAT is a member, employs a conserved mode of substrate binding and a conserved catalytic strategy (Dyda *et al.*, 2000). The binding of the acyl-donor substrate normally precedes the binding of the acyl acceptor substrate, after which direct acyl transfer takes place. It has been suspected that general base catalysts (glutamate, histidine, and aspartate residues) are used for the removal of protons from the amine, and that general acid catalysts (cysteine and tyrosine residues) are used for the donation of protons to the thiolate ions.

The kinetic model of a reaction can be experimentally determined by calculation of its kinetic parameters. The secondary or double reciprocal plots, generated from kinetic data

such as reaction velocities, are used for the calculation of kinetic parameters of enzymes. For bi-substrate reactions, the pattern displayed by these plots may reveal important information about the reaction mechanism. If the lines on these plots are parallel, it generally implies that a ping-pong mechanism is employed and when the lines converge at some point to the left of the vertical (y) axis when extrapolated, a ternary complex mechanism is generally implicated (Palmer and Bonner, 2007).

Previous studies done in our laboratory by Badenhorst *et al.* (2012) and van der Westhuizen *et al.* (2000) entailed the kinetic characterisation of normal and recombinant bovine liver, and human GLYAT enzymes, respectively. The results shown in Figure 4.13 are the Lineweaver-Burk plots of bovine liver GLYAT (A) and recombinant bovine GLYAT (B) enzymes that were generated by Badenhorst *et al.* (2012). When assessing the pattern displayed by these plots, Badenhorst *et al.* (2012) also concluded that a ternary-complex, base catalysed reaction mechanism was followed for bovine GLYAT. Van der Westhuizen *et al.* (2000) also stated that it was clear that GLYAT follows a sequential two-substrate mechanism.

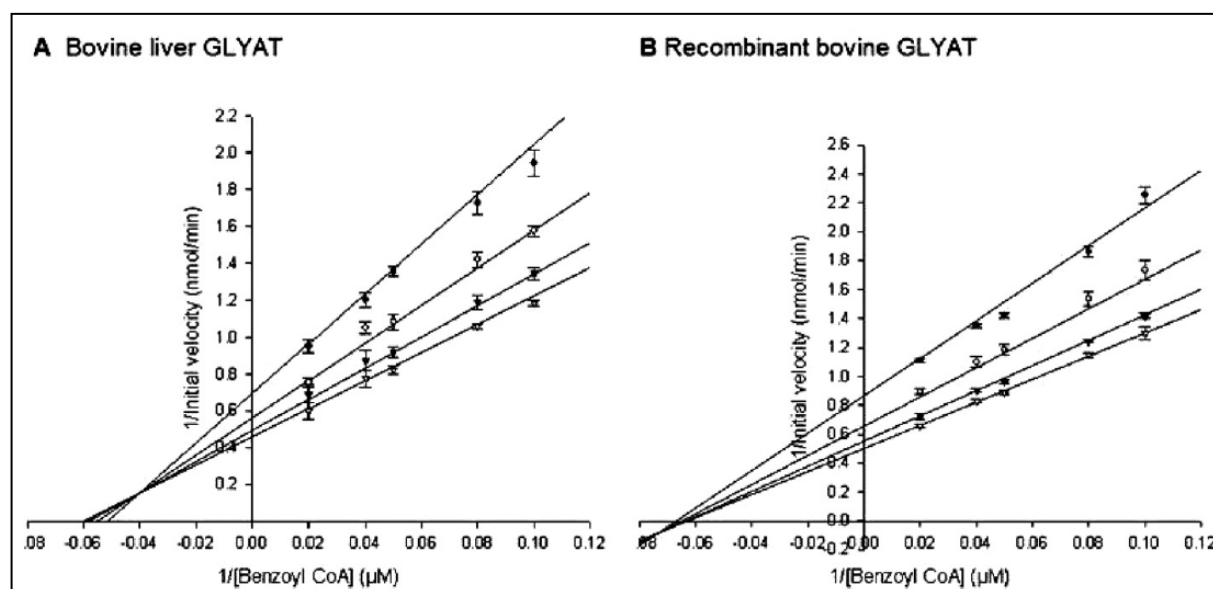


Figure 4.13: Lineweaver-Burk plots of the kinetic parameters of the bovine liver GLYAT (A) and recombinant bovine GLYAT (B) enzymes, respectively. Obtained with permission from Badenhorst *et al.* (2012).

Even though the lines of the Lineweaver-Burk plot representing the ping-pong bi-bi mechanism seem to fit the model (Figure 4.12), these findings are invalidated by the literature. As mentioned earlier, when considering the rest of the GNAT superfamily, it has been determined that the benzoyl-coenzyme A substrate binds before glycine. This indicates that a direct acyl transfer takes place and the products are released as the rate limiting step (Nandi *et al.*, 1979; van der Westhuizen *et al.*, 2000). Keeping in mind the high level conservation of the catalytic mechanism of the GNAT superfamily, it can be assumed that human GLYAT may also present with a compulsory order ternary complex mechanism with direct acyl transfer and some form of general acid-base catalysis.

More evidence supporting the theory of human GLYAT employing a direct transfer mechanism as opposed to an acyl:enzyme intermediate (ping-pong) mechanism exists. An example follows: when GLYAT reactions are spiked with compounds that react with sulfhydryl groups, GLYAT proves insensitive to these agents, since no inhibition is shown. This suggests that no cysteine residue is involved in the catalytic mechanism as one would see in an acyl:enzyme intermediate (ping-pong) mechanism (Nandi *et al.*, 1979).

The Lineweaver-Burk plots representing the random ordered and ordered kinetic models, revealed converging lines. This also eliminates the ping-pong mechanism, in which parallel lines would be produced as seen in Figure 4.12 (Cleland, 1963; Badenhorst *et al.*, 2012; Nandi *et al.*, 1979; van der Westhuizen *et al.*, 2000). Thus, it can be assumed that the GLYAT enzyme cannot be described as a ping-pong enzyme (Matsuo *et al.*, 2012; Nandi *et al.*, 1979).

All the kinetic experiments were performed in triplicate, which suggested that the parameters determined were reliable. The results of the bi-substrate kinetics in this study were in relative accord with the kinetic parameters as described in the literature (van der Sluis *et al.*, 2013; van der Westhuizen *et al.*, 2000; Kelley & Vessey, 1994; Mawal & Qureshi, 1994). Table 4.8 shows the K_m values for benzoyl-coenzyme A and glycine as

reported in the literature. The K_m value for benzoyl-coenzyme A reported in the literature varies between 6 μM and 67 μM (Table 4.8). Thus, the values of 49 ± 13 (36-62) μM obtained for the recombinant N156S GLYAT variant in this study, fall within the range described in the literature (Table 4.8). Great variation, however, exists in the literature regarding the values of kinetic parameters of GLYAT enzymes using glycine and benzoyl-coenzyme A as substrates (Bartlett & Gompertz, 1974; Gregersen *et al.*, 1986; Kelley & Vessey, 1994; Matsuo *et al.*, 2012; Mawal & Qureshi, 1994; van der Westhuizen *et al.*, 2000). For example, a previous study conducted in our laboratory by van der Sluis *et al.*, (2013), used the same N156S hGLYAT variant for their kinetic studies, but the resulting kinetic parameters differed. They obtained K_m values of 38 ± 4 (34-42) μM for benzoyl-coenzyme A for the recombinant N156S hGLYAT variant, while a value of 49 ± 13 (36-62) μM , were obtained in this study which overlaps with their results, but are still not identical. This difference is difficult to explain, but could be ascribed to experimental variables. These may include temperature variations, different expression levels (protein yield), protein purity, different levels of enzyme activity, handling errors, and slight differences in reaction master mix solutions. One factor worth considering is that the values obtained from the single substrate kinetics that they performed, could deliver different results from the bi-substrate kinetics done in this study. In this study, we also further optimised the expression conditions of the enzyme, which could have resulted in the formation of more active enzymes than the previous studies.

One limitation of the literature surrounding hGLYAT kinetics is that most previous kinetic studies regarding hGLYAT were performed with enzymes isolated from liver tissue samples, of which the variant analysed was not determined, except for van der Sluis *et al.* (2013), who also used the recombinant N156S hGLYAT variant for kinetic studies. Thus, the K_m values reported in the literature are not actually representative of those obtained from kinetic studies done with recombinant enzymes. Also, many literature sources do not report K_m values for glycine for the N156S hGLYAT variant, thus the values determined in this study could only be compared to a select few sources. The K_m values reported for glycine (using benzoyl-coenzyme A as acyl donor) ranged from 6.4 mM to 26.6 mM (Table

4.8). The value of 20 ± 4 (16-24) mM determined in this study, was at the high end of the reported range, but still comparable to the reported values (Table 4.8).

Table 4.8: Kinetic parameters of benzoyl-coenzyme A and glycine for hGLYAT enzyme as reported in the literature.

Source	Recombinant/liver enzymes and variant	K _m benzoyl-coenzyme A (μM)	K _m glycine (mM)
This study	Recombinant (N156S variant)	49 ± 13 (36-62)	20 ± 4 (16-24)
Van der Sluis <i>et al.</i> 2013	Recombinant (N156S variant)	38 ± 4 (34-42)	n/a
Matsuo <i>et al.</i> 2012	Recombinant	n/a	26.6
Mawal & Qureshi, 1994	Liver enzyme (unknown variant)	57.9	n/a
Kelley & Vessey, 1994	Liver enzyme (unknown variant)	67	n/a
Van der Westhuizen <i>et al.</i> 2000	Liver enzyme (unknown variant)	13	6.4
Webster <i>et al.</i> 1976	Liver enzyme (unknown variant)	6	20

4.4 Summary

The kinetic parameters for the N156S hGLYAT variant were determined for both benzoyl-coenzyme A and glycine. These values obtained were comparable and corresponded well to the human GLYAT kinetic parameters reported in the literature. The enzyme displayed significant activity which made kinetic studies relatively simple.

In conclusion, the aim of this chapter was reached by determining the kinetic parameters of the recombinant N156S hGLYAT variant using benzoyl-coenzyme A (20 μ M – 80 μ M) and glycine (2.5 mM – 20 mM) as substrates. The conditions for kinetic studies, such as optimal protein content, were optimised and implemented for determination of kinetic parameters using the SigmaPlot 12.0 software, along with the Enzyme Kinetics module for graphical K_m determinations. The kinetic parameters determined in this study correlated well with the values reported in the literature.

It was also concluded that even though the Lineweaver-Burk plots representing the ping-pong bi-bi mechanism seem to fit the model, these findings were inconsistent with the literature, since the GLYAT enzymes are insensitive to thiol-modifying reagents, and that GLYAT is homologous to the GNAT superfamily of acyltransferase, of which the members are known to utilise direct transfer mechanisms and not ping-pong mechanisms (Badenhorst *et al.*, 2012). Thus, it was concluded that the GLYAT enzyme cannot be described as a ping-pong enzyme (Matsuo *et al.*, 2012, Nandi *et al.*, 1979). The Lineweaver-Burk plots representing the random ordered and ordered kinetic models, revealed converging lines. Both models represented ternary-complex mechanisms, which was in accordance with the results previously published by van der Westhuizen *et al.* (2000); Badenhorst *et al.* (2012); and Nandi *et al.* (1979). The data points fit the random order sequential mechanism model better, and it was thus assumed to represent the reaction.

Chapter 5: HPLC-ESI-MS/MS quantification of hippuric acid using a stable isotope

5.1 Introduction

The methods described in Chapter 4 that were used for enzyme kinetics only relied on colorimetric analyses to verify and measure the enzyme activity of human GLYAT. Although these methods were sufficient for calculating kinetic parameters, the absolute amounts of HA formed during the human GLYAT activity, was yet to be determined. It was also not definitely known whether the measured colorimetric changes were a result of background absorption or a result of HA formation. An HPLC-ESI-MS/MS method was consequently utilised to be able to selectively identify and accurately quantify the amount of HA formed in the presence of varying substrate concentrations. A validated HPLC-MS/MS method developed by Nortje *et al.* (2015) was employed for the accurate quantification and identification of the HA in the reaction samples. Reaction mixtures were prepared as described in Section 4.2.5 and stopped with formic acid. A liquid-liquid extraction with ethyl acetate was used to purify HA from the reaction mixture and dried under nitrogen. The dried samples were then reconstituted in water for analysis. The samples were first separated on an HPLC column, after which they were detected with the tandem MS. Multiple reaction monitoring was used to detect only the precursor and product ions known to be associated with HA and its stable isotope, hippuric acid-d5 (HA-d5), which served as an internal standard. Accurate quantification of HA within all samples was done relative to HA-d5. The known amounts of HA-d5 within each sample were also used to compensate for any possible losses during the extraction process.

5.2 Materials and methods

5.2.1 Reagents, standards and solutions

Again, for the sake of clarity of presentation, the reagents and materials used for this part of the study, as well as their suppliers and catalogue numbers are listed in Appendix A, and are thus not included in the text.

5.2.2 Preparation of standard stock solutions

All stock solutions were prepared with HPLC-grade water, which was also used throughout this part of the study. A stock solution was prepared of HA in water to a final concentration of 1280 $\mu\text{mol/L}$. The HA-d5 stable isotope was used as an internal standard to compensate for losses during HA extraction method as well as for pipetting errors. A stock solution of the isotope was prepared to a final concentration of 6500 $\mu\text{mol/L}$. All stock solutions were stored at 4°C. A 1 M stock of Tris was prepared for enzyme reactions.

5.2.3 Preparation of mobile phases

The mobile phases for the HPLC instrument were prepared according to the compositions as indicated by Nortje *et al.* (2015). The solvents used were water containing 0.5% formic acid (mobile phase A) and 100% methanol (mobile phase B).

5.2.4 Organic solvent extraction

The liquid-liquid extraction method proposed by van der Westhuizen *et al.* (2000), was used for the extraction of HA. The method involved the use of an organic solvent, ethyl acetate, to achieve phase separations. The efficacy of the extraction method was investigated by testing its ability to extract the protein as well as the high concentrations of glycine present

in the enzyme reactions. Test enzyme reactions were set up as described in Section 4.2.5. The reactions proceeded for 20 minutes after which they were terminated by adding 20 μ l concentrated formic acid (van der Westhuizen *et al.*, 2000). The addition of formic acid also aided in protein precipitation which was necessary before HPLC-MS/MS analyses (Weiner *et al.*, 1972; van der Westhuizen *et al.*, 2000).

The enzyme reaction samples were then transferred to Kimax tubes before proceeding with further analyses. For the extraction of the benzoyl conjugates, an organic solvent extraction was performed using ethyl acetate. Approximately 2 ml ethyl acetate was added to the tubes, vortex-mixed for 1 minute and then centrifuged (2800 x g, 1 minute, room temperature) to separate the phases. The top aqueous layer was then removed and transferred to a clean Kimax tube. This procedure was repeated 3 times, and the aqueous layer, of each individual repetition, was transferred into the same final Kimax tube. The samples were then dried under nitrogen for approximately 15 minutes, or until the samples were reduced to approximately 1 ml and transferred to smaller HPLC vials. This transfer was necessary because it was noticed in preliminary studies that the larger vials resulted in significant sample losses. After careful transfer of the samples, they were dried further within the smaller HPLC vials. Prior to measurement, the dried samples were reconstituted with 200 μ L water. This solution was then injected into the HPLC-ESI-MS/MS system.

To test the recovery of HA after the extraction procedure, a known concentration level (16 μ M) of HA (in water) was selected and 200 μ l thereof aliquoted into two separate tubes. One of the samples was then put through the entire extraction procedure. Both samples were then injected into the HPLC-ESI-MS/MS system. The responses obtained from the samples were then compared to evaluate the recovery of the HA after the extraction.

5.2.5 Method development and optimisation

5.2.5.1 Mass spectrometry optimisation for HA

Since the method previously developed by Nortje *et al.* (2015) was optimised for the simultaneous detection of various phase 2 biotransformation metabolites along with HA directly from human urine, and not only HA in water, the MS conditions required for the detection and quantification of HA were optimised briefly before starting the analyses. The Masshunter optimizer software from Agilent (Santa Clara, CA, USA) was employed to optimise the data acquisition parameters to be used for multiple reaction monitoring (MRM), a highly selective setting of the MS. The optimisation program selected the optimal parameters required to form both the intact precursor ion of HA as well as a characteristic product ion. The optimised parameters included the selection of the most abundant precursor and product ions, collision energy, and fragmentor voltage.

From the 1280 μM HA stock solution, 10 μl was added to 90 μl water, to a final concentration of 128 μM . For the optimisation procedure the sample was directly infused into the MS. The mobile phases used during optimisation were methanol and water containing 0.1% formic acid. The optimisation was carried out in positive ionisation mode. The results acquired from the method optimisation are reported in Table 5.1.

5.2.5.2 Mass spectrometry optimisation for HA-d5

Before starting the HA quantification analyses, it was also necessary to optimise the MS conditions required for the accurate detection of the HA-d5 isotope, to verify whether the optimised conditions for HA-d5 isotope detection are similar to the conditions for HA. The Masshunter optimiser software from Agilent (Santa Clara, CA, USA) was once again employed to optimise the data acquisition parameters for multiple reaction monitoring (MRM) mode for HA-d5.

From the 650 μM HA-d5 stock, 5 μl was added to 500 μl water, to a final concentration of 6.44 μM . For the optimisation procedure the sample was infused directly into the MS. The mobile phases used during optimisation were methanol and water with 0.1% formic acid. The optimisation protocol was carried out in positive ionisation mode. The results acquired from the method optimisation are reported in Table 5.1.

Table 5.1: Optimised MRM parameters for the quantification of HA and HA-d5 in positive ionisation mode.

Compound	Molecular weight	Precursor ion (m/z)	Fragmentor voltage	Most abundant product ion (m/z)	Collision Energy (V)
HA	179.17	180.1	45	105.1	8
HA-d5	184.20	185.1	45	110.1	8

5.2.5.3 Calibration curve of HA with HA-d5

Range and linearity was assessed for HA and HA-d5 detection by preparing a calibration curve from a 1:10 dilution of the 1280 μM HA stock. A concentration range of 1 – 128 μM HA was prepared from this stock by means of serial dilution with water. All of the concentration levels within the range were analysed in triplicate and the acquired data was used to construct a calibration curve. The correlation coefficient (r^2) was used to evaluate the linearity of the calibration curve. It was also necessary to be able to generate a linear equation of the response and concentration ratios for later use. A 650 μM solution of the HA-d5 isotope was prepared separately by diluting the 6500 μM stock solution 1:10 in water. The 1 – 128 μM HA sample concentrations were then slightly diluted by the addition of 5 μl of the 650 μM HA-d5 stock solution to a final concentration of 6.44 μM . The adjusted concentration range was 0.99 - 126.733 μM HA, all containing 6.44 μM isotope. These

samples were then also analysed in triplicate and the acquired data was used to construct a calibration curve of HA response/HA-d5 response versus HA concentration/HA-d5 concentration. The correlation coefficient (r^2) was used to evaluate the linearity of the calibration curve.

5.2.6 HPLC-ESI-MS/MS specifications

An Agilent 6460 triple quadrupole mass spectrometer with a Jet Stream electrospray ionisation (ESI) source was used for the detection of the analytes (Santa Clara, CA, USA) in positive ionisation mode (Nortje *et al.*, 2015). An Agilent StableBond – Phenyl (2.1 mm× 100 mm, 1.8 μ M) column with a Phenomenex KrudKatcher Ultra in-line filter was used to chromatographically separate the analytes and the mobile phase composition is summarised in Table 5.3. The method used for the HA detection and quantification was adapted from the method described by Nortje *et al.* (2015). The source parameters of the MS were as follows: The gas temperature was 280 °C, a gas flow of 6 L/min, the nebulizer set at 45 psi, sheering gas at 350, sheering gas flow of 7 L/min, a capillary voltage of 3000 V, and a nozzle voltage of 1500 V. The mobile phase flow rate was 0.2 ml/min. Successful chromatographic separation was achieved with the following gradient mobile phase composition: 5% methanol for 1 minute, an increase to 31% for 6.5 minutes, an increase to 100% methanol from 7.5 to 9.5 minutes, isocratic for another 4 minutes until 13.5 minutes, and then back to 5% methanol for the last minute. A post-time of 7 minutes was set to allow the pressure and the composition of the mobile phases to stabilise between analyses. Thus, a total run time of 22 minutes was used. Table 5.2 summarises the gradient mobile phase composition. Injection volumes were 1 μ l and were coupled with a needle wash step. The fragmentor voltage was set at 45 V and the collision energy at 8 V for both HA and HA-d5.

Table 5.2: Gradient mobile phase composition.

Time (minutes)	Mobile phase composition (% methanol)
1	5%
7.5	31%
9.5	100%
13.5	100%
14.5	5%

5.2.7 Quantification with stable isotope

The next step in the study was to accurately quantify the amount of HA formed during the GLYAT enzyme reaction using the HPLC-ESI-MS/MS method described above. These measurements were coupled with the addition of a stable isotope of HA (HA-d5) to ensure absolute quantification with high repeatability and accuracy.

5.2.7.1 Enzyme reactions

For quantification of HA, enzyme reactions were set up in a 96-well plate as described in Section 4.2.5. The benzoyl-coenzyme A concentrations were 20, 40, 60, and 80 μM respectively and the glycine concentrations used were 2.5, 5, 10, and 20 mM. The final reaction mixtures were 200 μl in volume. To simplify the enzyme assay and to minimise inter-reaction variability, four master mix solutions were prepared for each glycine concentration. This master mix solution consisted of 20 μl TrisCl, 40 μl glycine (final concentrations of 2.5, 5, 10, and 20 mM), 110 μl water, and 10 μl of the purified N156S human GLYAT (1.2 μg). Each master mix was mixed thoroughly and 200 μl was transferred into the first four wells of a row of a 96-well plate, covered, and incubated at 37 $^{\circ}\text{C}$ while the next step was performed. Into a separate 96-well plate, 20 μl of each of the different benzoyl-coenzyme A concentrations were added. The master mixes were added to the

plate containing the benzoyl-coenzyme A, which resulted in the final substrate concentrations as shown in Table 5.4.

The reactions were allowed to proceed for 20 minutes and then stopped by adding 20 μ l concentrated formic acid. Before further analysis, the samples were spiked with the HA-d5 stable isotope to a final concentration of 6.44 μ M. The extraction method described in Section 5.2.4 was then followed to extract the HA and HA-d5 from the reaction samples. The samples were then analysed with the HPLC-ESI-MS/MS method with the parameters described in Section 5.2.6. The samples were analysed in triplicate and the precision/repeatability expressed as the coefficient of variation (%CV).

Table 5.4: The plate layout of the enzyme reactions indicating the substrate concentrations added to each reaction mixture.

	Substrate concentrations	1	2	3	4
A	Benzoyl-coenzyme A (μ M)	20	20	20	20
	Glycine (mM)	2.5	5	10	20
B	Benzoyl-coenzyme A (μ M)	40	40	40	40
	Glycine (mM)	2.5	5	10	20
C	Benzoyl-coenzyme A (μ M)	60	60	60	60
	Glycine (mM)	2.5	5	10	20
D	Benzoyl-coenzyme A (μ M)	80	80	80	80
	Glycine (mM)	2.5	5	10	20

5.2.8 Statistical data analysis

The raw mass spectrometric data was acquired using the onboard Agilent MassHunter Data Acquisition (B.04.00) and Quantitative Analysis (B.04.00) software. HA was quantified relative to the internal standard (HA-d5) using the calibration curve previously described in Section 5.2.5.3. GraphPad Prism was used for statistical analyses and data representation.

5.3 Results and discussion

5.3.1 Introduction

The aim of this part of the study was to employ an HPLC-ESI-MS/MS method to both identify and quantify HA formed by the enzyme activity of human GLYAT, in the presence of its substrates glycine and benzoyl-coenzyme A. This method ensured accurate quantification of the HA in the reaction samples, and extremely high selectivity. The method involved chromatographic separation on an Agilent ZORBAX StableBond Phenyl column, with methanol and acetonitrile as mobile phases and finally detection of the precursor and product ions using in MRM mode.

5.3.2 Linearity

The linearity of the HA detection was investigated with the HA-d5 stable isotope. The linearity was determined using a concentration range of 1 – 128 μM HA. The chosen range was decided upon after determining that the expected HA levels were 20 - 80 μM , or the equivalent of the amount of benzoyl-coenzyme A added to the reaction mixtures. Thus, a wider range of this expected range was used. The stable isotope was added to a final concentration of 6.44 μM . The calibration curves were linear within the selected calibration ranges. The calibration curve with HA and HA-d5 are illustrated in Figure 5.1. The linear ranges and the correlation coefficient values (r^2) for each calibration curve is reported in Table 5.5. The linear equation of the calibration curve of HA with HA-d5 was

later used for calculation of the final HA in the enzyme reaction samples.

Equation 5.1: Linear formula of the HA with HA-d5 calibration curve.

$$y = 1.1176x + 0.0733$$

Table 5.5: The linear ranges and the correlation coefficient values (r^2) for both calibration curves.

Compound(s)	Linear range (μM)	Correlation coefficient value (r^2)
HA	1 – 128 μM	0.9987
HA with HA-d5	0.99 – 126.733 μM 6.44 μM	0.9994

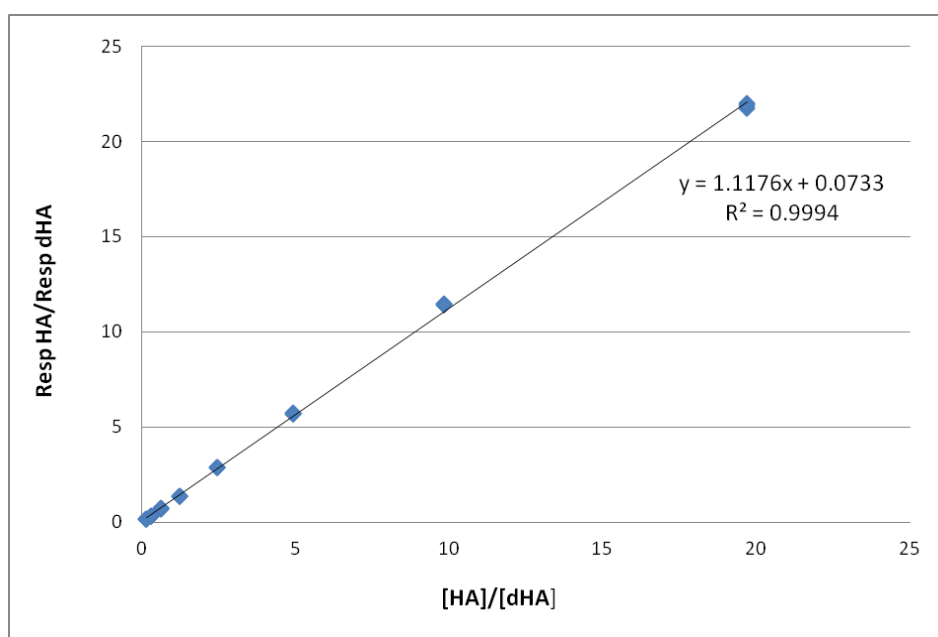


Figure 5.1: Calibration curve of HA with HA-d5. The calibration curve shows the linear range for HA with HA-d5 in water. Linearity was assessed using a concentration range of 0.99 – 126.733 μM of HA, spiked with 6.44 μM HA-d5 in each sample. The correlation coefficient value (r^2) for this plot was 0.9994.

5.3.3 Organic solvent extraction

A simple extraction of HA from the enzyme reactions was performed, followed by drying under nitrogen and finally resuspension into water. The recovery of using this method was evaluated and was found to be an extremely efficient method, as near 100% of the HA was retrieved after extraction.

5.3.4 Quantification with HPLC-ESI-MS/MS method

The HPLC-ESI-MS/MS method with the parameters described in Section 5.2.6 was implemented for the quantification of HA formed from the hGLYAT enzyme reactions that were set up, extracted, and spiked with the stable isotope as described in Section 5.2.7.1. The samples were analysed in triplicate. The final HA concentrations could be determined relative to the internal standard (HA-d5) using the equation of the linear calibration curve (shown by Equation 5.2).

Equation 5.2: The linear equation with the substituted values for final HA concentration determination.

$$\frac{\text{Response HA}}{\text{Response HAd5}} = 1.1176 \left(\frac{x}{\text{HAd5 concentration}} \right) + 0.0733$$

The final HA yield formed from N156S hGLYAT enzyme activity, with glycine and benzoyl-coenzyme A as substrates, were then calculated. The precision/repeatability was calculated in terms of the coefficients of variation (%CV) using Equation 5.3. The mean and standard deviation (SD) values for each triplicate sample measurement were calculated using the quantified amounts for each of the triplicates of each sample.

Equation 5.3: Equation for calculating coefficient of variation (% CV).

$$\%CV = \frac{\text{Standard deviation}}{\text{Mean}} \times 100$$

The concentrations of HA determined in the enzyme reaction samples were found to be 9.68 – 66.84 μM . The precision (%CV) of the triplicate measurements were between 0.22% and 1.36% for all samples. These values showed acceptable precision, as the criterion of acceptance is generally within 15%. According to the United States Food and Drug Administration guidelines for bioanalytical method validation (2001): “the precision determined at each concentration level should not exceed 15% of the coefficient of variation (CV)”. This demonstrated the repeatability and efficacy of this analytical method. The final HA concentrations, averages, and %CV values are shown in Table 5.6.

Table 5.6: Final HA concentrations, average HA concentrations, and %CV values (continued on the next page)..

[Benzoyl-CoA] (μM)	[Glycine] (mM)	Final [HA] (μM)	Average [HA] (μM)	%CV
20	2.5	9.8	9.6823994	0.91
		9.7		
		9.6		
20	5	15.0	14.900621	0.76
		14.9		
		14.8		
20	10	15.8	15.818116	0.51
		15.7		
		15.9		
20	20	15.5	15.456734	0.29
		15.4		
		15.5		
40	2.5	18.2	18.233930	0.37
		18.2		
		18.3		
40	5	29.4	29.572761	0.39
		29.7		
		29.6		

Table 5.6: (Continued)

[Benzoyl-CoA] (μM)	[Glycine] (mM)	Final [HA] (μM)	Average [HA] (μM)	%CV
40	10	31.8	32.311997	1.36
		32.5		
		32.7		
40	20	33.0	33.446127	1.17
		33.6		
		33.7		
60	2.5	24.4	24.503901	0.42
		24.6		
		24.5		
60	5	46.3	46.168682	0.69
		45.8		
		46.4		
60	10	52.4	51.957595	0.98
		52.1		
		51.4		
60	20	50.5	50.104959	1.34
		49.3		
		50.5		
80	2.5	25.5	25.341930	0.43
		25.3		
		25.3		
80	5	56.3	56.704273	1.11
		56.4		
		57.4		
80	10	67.6	67.769724	0.22
		67.8		
		67.9		
80	20	66.7	66.849286	0.23
		66.9		
		67.0		

For convenience and clarity of presentation, column charts were drawn from the above mentioned data. Figure 5.2 shows the final HA concentrations obtained from the N156S hGLYAT variant activity at varying glycine levels and constant benzoyl-coenzyme A levels, where each colour represents a different concentration level of benzoyl-coenzyme A.

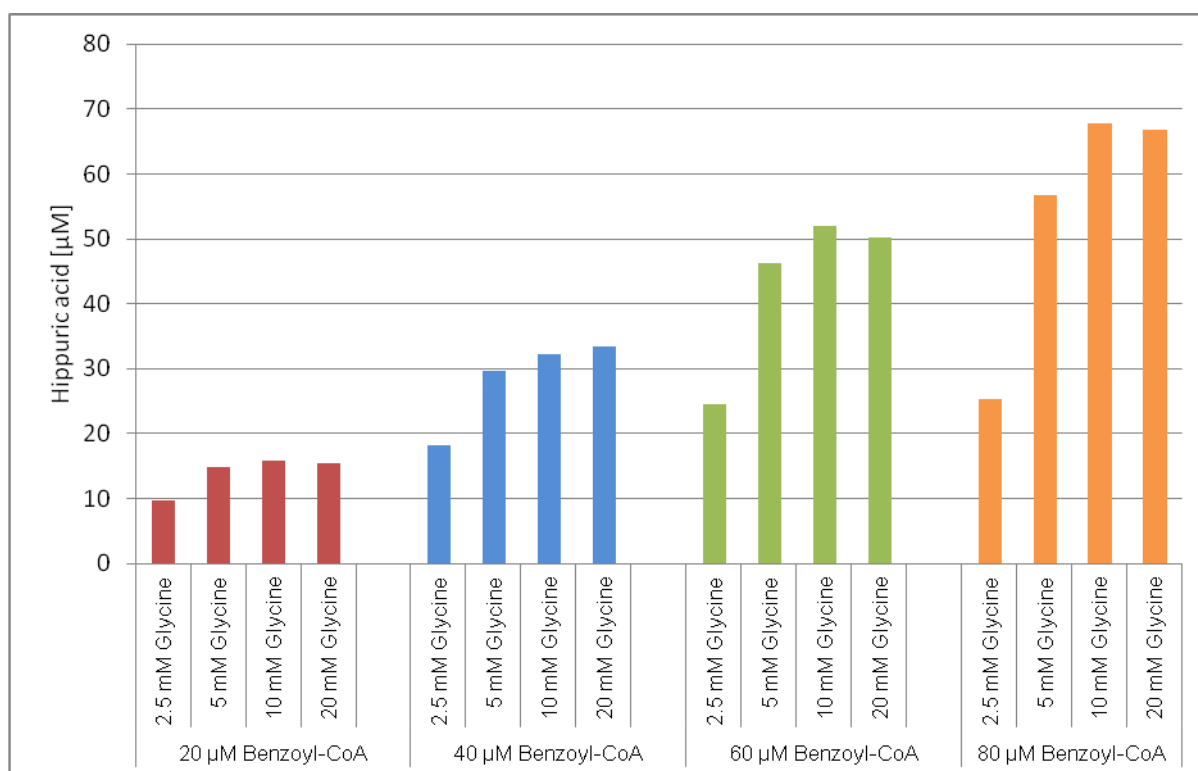


Figure 5.2: A column chart of the final HA concentrations obtained at varying glycine levels. Each colour represents a different constant concentration level of benzoyl-CoA (20 μM, 40 μM, 60 μM, 80 μM).

From Figure 5.2 it can be noticed that an increase in glycine concentrations in the enzyme reaction samples, results in an increase of HA formation by human GLYAT activity at constant benzoyl-coenzyme A concentrations. This increase, however, reaches a peak at 10 mM glycine, after which a further increase in glycine concentration no longer has a significant effect on the HA formation and the curve reaches a plateau.

On the other hand, it was also necessary to assess the role of benzoyl-coenzyme A variation in HA formation. For this purpose, the data used for the graph indicated in Figure 5.2 was inverted and re-plotted. Figure 5.3 shows the final HA concentrations obtained from the N156S hGLYAT variant activity at varying benzoyl- coenzyme A levels and constant glycine levels, where each colour represents a different concentration level of glycine.

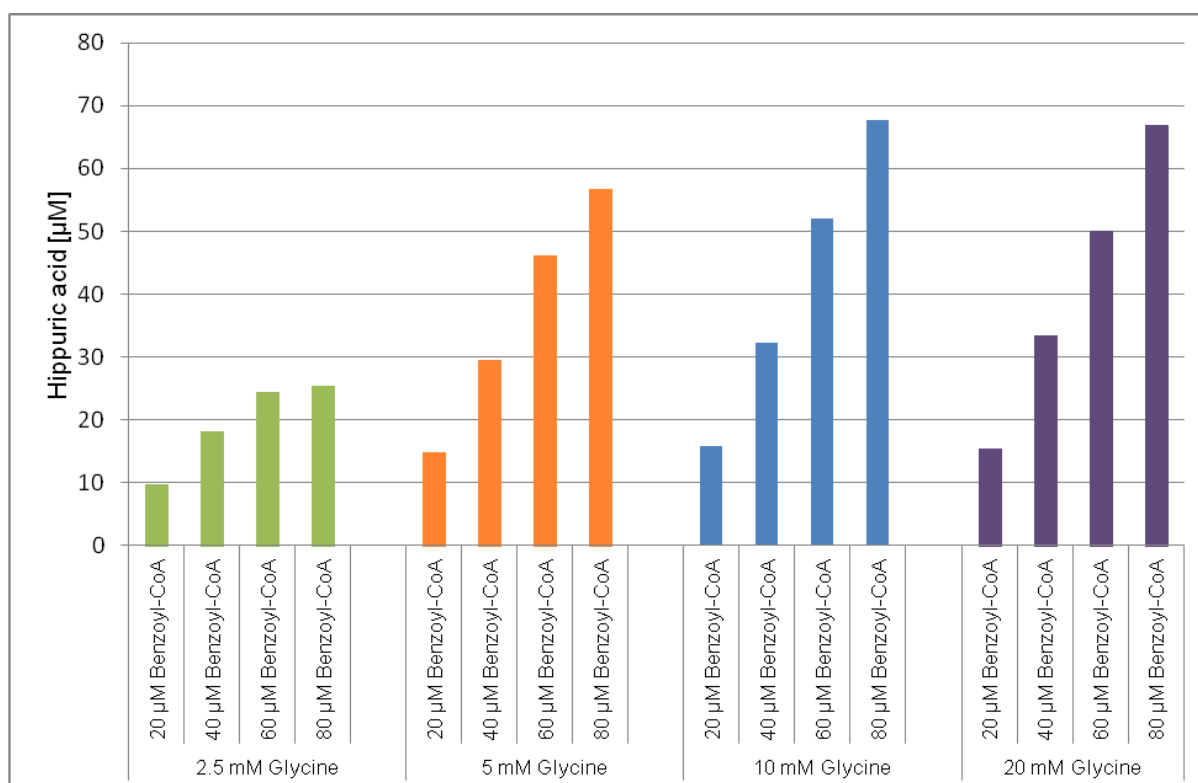


Figure 5.3: A column chart of the final HA concentrations obtained at varying benzoyl-CoA levels. Each colour represents a different constant concentration level of glycine (2.5 mM, 5 mM, 10 mM, and 20 mM).

From Figure 5.3 it can be deduced that an increase in benzoyl-coenzyme A concentrations in the enzyme reaction samples, results in an increase of HA formation by hGLYAT activity at constant glycine concentrations. The increase in benzoyl-coenzyme A concentration, different from that seen in glycine concentration increase, resulted in a near linear relationship with the HA formation, although not all of the benzoyl-coenzyme A is converted to HA.

As mentioned in Section 2.9.3.1, the availability of glycine is an important factor that can severely influence the tempo of hippuric acid synthesis (Figure 2.2) (Beliveau & Brusilow, 1987; De Vries *et al.*, 1948; Levy, 1979). Therefore, even though it is known that hippuric acid synthesis in humans is saturable even with large doses of benzoic acid, the co-administration of glycine can considerably boost the rate of hippuric acid formation (Knights & Miners, 2012; Levy, 1979). This effect can clearly be seen in Figure 5.3, in which the overall HA formation increases as the glycine concentrations were increased. The blue bars, representing 10 mM glycine appear much higher than the green bars, which represent 2.5 mM glycine. But, according to figure 5.2 there is a point where increasing glycine concentrations do not result in a further increase of HA formation, since N156ShGLYAT has a maximum rate at which it can detoxify xenobiotic acyl-coenzyme A compounds. This effect is ascribed to the V_{max} value of the N156ShGLYAT enzyme. We hypothesised that it is possible that the rate of detoxification of the enzyme may have evolved to be relatively limited, in order to avoid the complete depletion of glycine levels, which have important roles in many other pathways including the brain neurochemistry, the synthesis of collagen, nucleic acids, porphyrins, and other important metabolites (Badenhorst *et al.*, 2014).

It has already been shown that glycine availability is limiting in the developing human foetus, in young men, birds, and in milk-fed pigs (Jackson *et al.*, 1996; Wang *et al.*, 2013; Wu *et al.*, 2013). It has also been estimated that human adults have a shortage of about 10 grams of glycine per day. This glycine deficiency, and the resulting deficiency of the THF-C1 units, can influence the metabolism of glutathione, porphyrins, creatine, collagen, and nucleic acids (Brosnan *et al.*, 2011; Meléndez-Hevia *et al.*, 2009; Wang *et al.*, 2013). It has been argued that glycine shortage can be intensified by xenobiotic compounds, such as benzoic acid and other substrates for glycine conjugation (Au *et al.*, 2013; Badenhorst *et al.*, 2013; Jackson *et al.*, 1987).

5.4 Summary

Existing methods for measuring GLYAT enzyme activity rely on colorimetric analyses, in which DTNB is used to detect the glycine dependent release of coenzyme A (Kølvraa & Gregersen, 1986). The accurate quantification of HA formed during the human GLYAT activity, however, is not possible using this non-specific reaction since the measured colorimetric changes can be a result of other reactions or agents interfering with the colorimetric changes observed. The solution for this was to employ a much more selective technique such as HPLC-ESI-MS/MS detection to identify and accurately quantify the amount of HA formed. The HPLC-ESI-MS/MS method used for the HA quantification was previously developed and validated by Nortje *et al.* (2015).

Prior to extracting the HA using the method described in Section 5.2.6, the samples were spiked with the HA-d5 stable isotope to a final concentration of 6.44 μM to account for matrix effects and losses during sample preparation (Lebedev, 2013). The extracted samples were then analysed with the HPLC-ESI-MS/MS method described in Section 5.2.6 in MRM-mode, scanning only for the precursor and product ions representing HA and its stable isotope, HA-d5. The precision of the results was calculated in terms of the coefficient of variation (%CV) and the results showed excellent repeatability since the %CV values were between 0.22% and 1.36% for all samples measured in triplicate.

It was found that the amount of HA formed in the GLYAT reactions, displayed a linear relationship with the concentration of benzoyl-coenzyme A in the reactions, when glycine levels were kept constant. In the presence of increasing glycine concentrations and constant benzoyl-coenzyme A concentrations, however, it was noticed that HA formation still increased as expected, but did not show a linear relationship, since the HA formation reached a plateau after which the addition of higher glycine concentrations did not result in more HA formation. This limitation was ascribed to a possible evolutionary adaptation, in which the rate for glycine conjugation detoxification mechanisms is limited to avoid the depletion of *in vivo* glycine levels.

These findings may have an influence on the current approach for treatment of patients with isovaleric acidemia. These patients present with toxic accumulations of isovaleryl-coenzyme A and glycine supplementation is regarded as a successful treatment for this disease, seeing that the presence of more glycine may promote the glycine conjugation, and thus, the detoxification process (Krieger & Tanaka, 1976; Sweetman & Williams, 2001; Tanaka & Isselbacher, 1967). Our findings suggested that glycine supplementation will improve GLYAT activity, but only to a certain degree, after which it will not have a further impact on the detoxification capacity of the pathway. This may shed some light on the efficiency of glycine supplementation as a treatment for patients with severely high isovaleryl-coenzyme A accumulations.

Based on the work presented in this chapter, a peer reviewed scientific paper is currently being prepared for publication.

Chapter 6: Final conclusions and future prospects

6.1. Introduction

The biotransformation of xenobiotic compounds is becoming a topic of great interest, since we are constantly being exposed to such potentially toxic, chemicals virtually every day. One crucial step in this detoxification pathway is glycine conjugation. Glycine conjugation, which is considered one of the most fundamentally important mechanisms for maintaining homeostasis in animal physiology (Knights *et al.*, 2007). It was found that glycine conjugation plays a vital role in the maintenance of sufficient liver and kidney coenzyme A levels, and decreasing the inherent toxicity of numerous xenobiotic compounds by increasing their hydrophilicity.

As discussed in Section 2.9, a critical enzyme involved in this process is the GLYAT enzyme, which is mainly responsible for the detoxification of xenobiotic organic acids by conjugating them with glycine. One of the major substrates for GLYAT is benzoyl-coenzyme A, which is formed primarily by dietary polyphenols such as the flavonols and hydroxycinnamates that are produced by the gut microbes (Knights & Miners, 2012; Rechner *et al.*, 2002). Benzoyl-coenzyme A is conjugated to glycine and converted to hippuric acid by GLYAT. Hippuric acid is thus the main urinary metabolite of phenylpropionate catabolism and polyphenol metabolism (Dakin, 1908; Rechner *et al.*, 2002). The levels of hippuric acid excreted in urine have been associated with a variety of pathological conditions including diabetes, hepatitis, obesity, depression, cancer, schizophrenia, and gut dysbiosis (Lees *et al.*, 2013; Beyoğlu & Idle, 2012; Beyoğlu *et al.*, 2012).

It has previously been demonstrated that genetic variation in the *GLYAT* gene influences the enzyme activity of a recombinant GLYAT enzyme (van der Sluis *et al.*, 2013). It is thus of importance to study the properties of the different variants of the enzyme. The N156S

variant of GLYAT was studied here, since it is now regarded as the wild-type variant, due to the fact that it has both the highest allele frequency and also the highest enzyme activity (van der Sluis *et al.*, 2015). In pursuit of better characterisation of the glycine conjugation pathway, this study was aimed at investigating the bi-substrate enzyme kinetics of the N156ShGLYAT enzyme. Another aim was to accurately quantify the hippuric acid formed from N156ShGLYAT activity with glycine and benzoyl-coenzyme A substrates. These aims were successfully reached.

6.2. Conclusions

6.2.1 Bi-substrate kinetics of a recombinant N156ShGLYAT

The first objective of this part of the study was reached by cloning a coWThGLYAT insert into a pET-32a(+) vector with an N-terminal histidine tag. The coN156ShGLYAT variant was then generated via site-directed mutagenesis. The expression conditions for the enzymatically active recombinant hGLYAT were then further optimised by the addition of 0.5% glycine to the expression medium. Subsequently, the proteins were successfully purified using nickel affinity chromatography. The method for protein extraction was also optimised by comparing two methods, the BugBuster protein extraction method and the Lysis buffer method, of which the latter delivered slightly more soluble active protein. Upon further investigation, it was also found that the coN156ShGLYAT enzyme did not show improved expression or higher enzyme activity when compared to the non-codon optimised N156ShGLYAT. Subsequent investigations were thus conducted using the non-codon optimised N156S hGLYAT variant in combination with the Lysis buffer extraction method.

The establishment of proper storage conditions for the purified hGLYAT was another objective reached in this part of the study. This was necessary because the available storage options were poor and resulted in severe reductions in enzymatic activity. The alternative storage method developed in this study involved the addition of 10% glycerol to the purified proteins prior to snap-freezing and storage at -80°C. This solution also ensured that the

hGLYAT variant remained enzymatically active after long term storage. This finding simplified the study of the hGLYAT enzyme considerably, since it was now possible to work with a single batch of enzymes for extended periods, as opposed to only a few days at a time which may result in substantial inter-batch variations.

The next step in this study was to determine the bi-substrate kinetic parameters for the N156S hGLYAT variant. This aim was reached by varying the benzoyl-coenzyme A (20 μ M – 80 μ M) and glycine (2.5 mM – 20 mM) concentrations, and measuring the release of coenzyme A colorimetrically. The conditions for these kinetic studies, such as optimal protein content, were first optimised after which they were implemented for determination of the kinetic parameters. SigmaPlot 12.0 software in conjunction with its Enzyme Kinetics module, was implemented for the generation of double reciprocal plots, and consequently, for graphical K_m determinations. The K_m values determined for this enzyme were 49 ± 13 μ M for benzoyl-coenzyme A and 20 ± 4 mM for glycine. These kinetic parameters corresponded well with the values reported in the literature which described K_m values for benzoyl-coenzyme A ranging from 6 to 67 μ M; and from 6.4 to 26.6 mM for glycine. This suggested that the values determined in this study were correct and that this method may be used for future investigations regarding the purified N156S hGLYAT variant and the properties of the enzyme. It was also concluded that the Lineweaver-Burk plots represented ternary-complex mechanisms, which was in accordance with the results previously published by van der Westhuizen *et al.* (2000); Badenhorst *et al.* (2012); and Nandi *et al.* (1979). The data points fit the random order sequential mechanism model, and it was therefore assumed to represent the reaction. It was also concluded that the GLYAT enzyme cannot be described as a ping-pong enzyme, because according to the literature, GLYAT enzymes are insensitive to thiol-modifying reagents, and that they are homologous to the GNAT superfamily of acyltransferase, of which the members are known to utilise direct transfer mechanisms (Badenhorst *et al.*, 2012).

6.2.2 HPLC-ESI-MS/MS quantification of hippuric acid using a stable isotope

Previous methods for measuring GLYAT enzyme activity relied only on colorimetric analyses, in which DTNB is used to detect the glycine dependent release of coenzyme A (Kølvraa & Gregersen, 1986). The accurate quantification of the HA formed during the human GLYAT reaction, however, is yet to be determined. We addressed this void by successfully employing an HPLC-ESI-MS/MS method for HA quantification that was previously developed and validated in our laboratory by Nortje *et al.* (2015). The method was further improved by using a stable isotope of HA (HA-d5) as an internal standard, which could account for matrix effects and losses during sample preparation (Lebedev, 2013). Based on the results obtained with this method it was possible to quantify and correlate the amount of HA formed by N156ShGLYAT with the varying amounts of substrate present in the reactions. The precision of the results was calculated in terms of the coefficient of variation (%CV) and the results showed outstanding repeatability since the %CV values were between 0.22% and 1.36% for all samples measured in triplicate.

6.3 Future recommendations

Current research fails to interpret the interindividual variability in glycine conjugation capacity (Badenhorst *et al.*, 2013). By considering the kinetic parameters of the enzymes involved with biotransformation pathways, some light might be shed on this void. As discussed previously, it is known that genetic variation in the *hGLYAT* gene influences the enzyme activity of a recombinant hGLYAT. However, of the 99 known non-synonymous SNPs in the *hGLYAT* gene, only six have been investigated in terms of their single substrate kinetics (van der Sluis *et al.*, 2013). An additional dimension was added to this characterisation in this study, since the bi-substrate kinetic parameters of one of these variants, the N156ShGLYAT, were determined. A large void still exists, considering the large amount of uncharacterised SNP variants.

Another point to consider is that benzoic acid might not be the main organic acid substrate

for glycine conjugation, as it is currently assumed, for all of these variants. Thus, the substrate specificity of these hGLYAT variants also needs to be further characterised.

The influence of variation in the non-coding parts of the gene also needs to be investigated. When considering the effect of genetic variation on the glycine conjugation pathway, a relatively recent discovery might even further complicate the situation, namely the existence of “duons”, which will be discussed here briefly.

Transcription factors (TFs) are proteins that regulate transcription by binding to specified recognition sequences on the DNA (Stergachis *et al.*, 2013). A recent study found a new, previously hidden genetic code within the same sections of genes that encode proteins. It was found that approximately 15% of all human codons can be described as dual-use codons (“duons”). Until very recently, it was believed that the genetic code was used exclusively to encode for proteins and the introduction of duons caused a major paradigm shift. These sequences encode for two types of information: (i) the genetic code to produce proteins and (ii) the interpretation of the transcription factor–binding regulatory code to influence gene expression (Stergachis *et al.*, 2013).

Duons can have significant effects on gene expression, since it was proven that 17% of SNPs occurring within duons directly alter TF binding, which can result in altered gene transcription and ultimately the expression of the proteins (Stergachis *et al.*, 2013). The fact that the transcription factors bind to DNA within the protein-coding regions of the genome, has caused a significant bias in the usage of codons and thus the choice of amino acids, in such a way that is constrained by the binding motif of the transcription factors (Stergachis *et al.*, 2013). This finding also supports the concept that even without affecting the structure or function of a protein, SNPs located within protein-coding regions may lead to disease due to impaired TF binding (Plotkin & Kudla, 2011; Sauna & Kimchi-Sarfaty, 2011). Thus, when evaluating the impact of SNPs on protein function or interpreting disease mutation data, this entire new spectrum of codes within protein-coding regions should be considered

(Stergachis *et al.*, 2013). The investigation of these overlapping codes opens new vistas on the functional interpretation of variation in coding regions and makes it clear that the story of the genetic code has not yet run its course (Weatheritt & Babu, 2013).

Also, in order to conclusively prove the ternary-complex, base catalysed mechanism we suggested for N156ShGLYAT in this study, a crystal- or NMR structure for the GLYAT enzyme (and its variants) would have to be determined. This will provide much needed insights regarding the active sites of these enzymes and thus their catalytic mechanisms.

A long term goal of this study is to develop a therapeutic recombinant hGLYAT enzyme with altered substrate specificity for the treatment of specific organic acidemias (Badenhorst *et al.*, 2013). The next step in addressing this goal will be to explore the relationship between different variants of hGLYAT and the substrate specificity of the resulting hGLYAT enzymes. This could lead to further insights in developing a GLYAT enzyme with altered substrate specificity or increased activity for therapeutic purposes as mentioned in Section 2.11.2.3.

The impact of genetic variation in the *GLYAT* gene on the overall glycine conjugation rate *in vivo* is yet to be characterised (Badenhorst *et al.*, 2013; van der Sluis *et al.*, 2013). Another valuable research approach would be to develop an *in vivo* model, such as a GLYAT conditional knockout mouse model, which could be used to investigate the importance of GLYAT, as well as its genetic variations, in the glycine conjugation pathway. Seeing that the *GLYAT* gene is so well conserved (van der Sluis *et al.*, 2015), it is assumed that GLYAT must be of utmost importance in this pathway, as well as in the overall detoxification pathway. The nucleotide variation of other xenobiotic metabolising enzymes, such as the acid:coenzyme A ligases, needs to be further characterised to identify other variants that might play a role in xenobiotic response (Mortensen *et al.*, 2011).

It was mentioned earlier that our findings regarding the ability of GLYAT to utilise only a

certain amount of glycine concentrations, may have an influence on the current approach for treatment of patients with isovaleric acidemia. It was also mentioned that this effect may reflect on the efficiency of glycine supplementation as a treatment for patients with severely high isovaleryl-coenzyme A accumulations. One would also have to keep in mind that glycine availability is just one of the factors that contribute to the interindividual variations in glycine conjugation capacity (Section 2.9.3). Another factor at play here is genetic variation. It is known that SNPs in the *GLYAT* gene have an effect on the activity of GLYAT (van der Sluis *et al.*, 2013). Thus, the extent to which the glycine treatments will work will also depend on which GLYAT variant the patient possesses, and the kinetic parameters thereof. An improvement in the treatment of patients with organic acidemias would be to not only establish optimal glycine supplementation concentrations, but also to determine their inherent ability to detoxify such metabolites.

Another future prospect for this study would be to make use of HPLC-MS/MS methods to not only identify and quantify end point HA (product) formation, but to monitor the process over time and to detect the formation of the products. One may also use such a method to characterise the specific enzyme dynamics by looking at the disappearance of substrates, the formation and dissociation of enzyme-substrate/enzyme-product complexes, and product formation. This will also provide much needed insights on the exact mechanisms employed by the GLYAT enzyme and its variants.

Limitations of this study

Due to time constraints, the supplementation of the expression medium with glycine could not be fully investigated in this study. It is thus recommended that this optimisation step be further characterised by repeating the experiments with varying glycine concentrations, at different stages in the expression protocol, and to assess the effects on the cell growth and protein expression levels carefully. This characterisation would be especially important for the coN156ShGLYAT, which we hypothesised depleted glycine faster than the non-codon optimised N156ShGLYAT.

References

- Au, J. *et al.*, 2013. Whole-body protein turnover reveals the cost of detoxification of secondary metabolites in a vertebrate browser. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 183(7), pp.993–1003.
- Bachmann, C. & Bickel, M.H., 1986. History of drug metabolism: the first half of the 20th century. *Drug metabolism reviews*, 16(3), pp.185–253.
- Badenhorst, C.P.S. *et al.*, 2014. A new perspective on the importance of glycine conjugation in the metabolism of aromatic acids. *Drug metabolism reviews*, 6(3), pp.343–61.
- Badenhorst, C.P.S. *et al.*, 2013. Glycine conjugation: importance in metabolism, the role of glycine N-acyltransferase, and factors that influence interindividual variation. *Expert opinion on drug metabolism & toxicology*, 9(9), pp.1139–53.
- Badenhorst, C.P.S., Jooste, M. & van Dijk, A.A., 2012. Enzymatic Characterization and Elucidation of the Catalytic Mechanism of a Recombinant Bovine Glycine N-Acyltransferase. *Drug metabolism and disposition*, 40(2), pp.346–352.
- Baneyx, F., 1999. Recombinant protein expression in Escherichia coli. *Current Opinion in Biotechnology*, 10(5), pp.411–421.
- Banoglu, E., 2000. Current status of the cytosolic sulfotransferases in the metabolic activation of promutagens and procarcinogens. *Current drug metabolism*, 1(1), pp.1–30.
- Banta-Wright, S.A. & Steiner, R.D., 2004. Tandem Mass Spectrometry in Newborn Screening. *The Journal of Perinatal & Neonatal Nursing*, 18(1), pp.41–60.
- Bartlett, K. & Gompertz, D., 1974. The Specificity of Glycine-N-Acylase and Acylglycine Excretion in the Organicacidaemias. *Biomedical Medicine*, 10, pp.15–23.
- Beliveau, G.P. & Brusilow, S.W., 1987. Glycine availability limits maximum hippurate synthesis in growing rats. *The Journal of nutrition*, 117(1), pp.36–41.

- Bell, B.S.C. & Turner, J.M., 1976. Bacterial Catabolism of Threonine. *Biochemistry Journal*, 156, pp.449–458.
- Beyoğlu, D. & Idle, J.R., 2012. The glycine deportation system and its pharmacological consequences. *Pharmacology & therapeutics*, 135(2), pp.151–67.
- Beyoğlu, D., Smith, R.L. & Idle, J.R., 2012. Dog bites man or man bites dog? The enigma of the amino acid conjugations. *Biochemical pharmacology*, 83(10), pp.1331–9.
- Boja, E.S. & Rodriguez, H., 2011. The path to clinical proteomics research: Integration of proteomics, genomics, clinical laboratory and regulatory science. *Korean Journal of Laboratory Medicine*, 31(2), pp.61–71.
- Bonafé, L. *et al.*, 2000. Evaluation of urinary acylglycines by electrospray tandem mass spectrometry in mitochondrial energy metabolism defects and organic acidurias. *Molecular genetics and metabolism*, 69(4), pp.302–11.
- Boomgaarden, I. *et al.*, 2009. Comparative analyses of disease risk genes belonging to the acyl-CoA synthetase medium-chain (ACSM) family in human liver and cell lines. *Biochemical genetics*, 47(9-10), pp.739–48.
- Bornhorst, J.A. & Falke, J.J., 2000. Purification of proteins using polyhistidine affinity tags. *Methods Enzymology*, 326, pp.245–254.
- Bravo, L., 1998. Polyphenols : Chemistry , Dietary Sources , Metabolism , and Nutritional Significance. *Nutrition Reviews*, 56(11), pp.317–333.
- Bray, H.G., Thorpe, W. V & White, K., 1951. Kinetic studies of the metabolism of foreign organic compounds; the formation of benzoic acid from benzamide toluene, benzyl alcohol and benzaldehyde and its conjugation with glycine and glucuronic acid in the rabbit. *The Biochemical journal*, 48(1), pp.88–96.
- Brosnan, J.T., da Silva, R.P. & Brosnan, M.E., 2011. The metabolic burden of creatine synthesis. *Amino acids*, 40(5), pp.1325–1331.

- Buchheim, R., 1859. Lehrbuch der Arzneimittellehre. Voss, Leipzig, p. 19.
- Carter, P., 1986. Site-directed mutagenesis. *Biochemistry Journal*, 237, pp.1–7.
- Chen, C. *et al.*, 2000. Activation of antioxidant-response element (ARE), mitogen-activated protein kinases (MAPKs) and caspases by major green tea polyphenol components during cell survival and death. *Archives of pharmacal research*, 23(6), pp.605–12.
- Cleland, W.W., 1963. The kinetics of enzyme-catalyzed reactions with two or more substrates or products: III. Prediction of initial velocity and inhibition patterns by inspection. *Biochimica et Biophysica Acta*, 67, pp.188–196.
- Cohen Freue, G.V. & Borchers, C.H., 2012. Multiple Reaction Monitoring (MRM): Principles and Application to Coronary Artery Disease. *Circulation: Cardiovascular Genetics*, 5(3), pp.378–378.
- Conti, A. & Bickel, M.H., 1977. History of Drug Metabolism: Discoveries of the Major Pathways in the 19th Century *. *Drug metabolism reviews*, 6(1), pp.1–50.
- Coon, M.J., 2005. CYTOCHROME P450: Nature's Most Versatile Biological Catalyst. *Annual Review of Pharmacology and Toxicology*, 45(1), pp.1–25.
- Dakin, H.D., 1908. Comparative Studies of the Mode of Oxidation of Phenyl Derivatives of Fatty Acids by the Animal Organism and by Hydrogen Peroxide. *Journal of Biological Chemistry*, 277(15), pp.419–435.
- Dercksen, M. *et al.*, 2012. Clinical variability of isovaleric acidemia in a genetically homogeneous population. *Journal of inherited metabolic disease*, 35(6), pp.1021–9.
- Dercksen, M. *et al.*, 2013. Inhibition of N-acetylglutamate synthase by various monocarboxylic and dicarboxylic short-chain coenzyme A esters and the production of alternative glutamate esters. *Biochimica et biophysica acta*, 1842(12), pp.2510–2516.

- De Vries, A., Alexander, B. & Quamo, Y., 1948. Studies on amino acid metabolism. III. Plasma glycine concentration and hippuric acid formation following the ingestion of benzoate. *The Journal of clinical investigation*, 27(5), pp.665–668.
- Dietz, A.A. & Rubinstein, H.M., 1972. Laboratory note standardization of the ellman reaction. *Clinical Biochemistry*, 5, pp.136–138.
- Döring, B. & Petzinger, E., 2014. Phase 0 and phase III transport in various organs: combined concept of phases in xenobiotic transport and metabolism. *Drug metabolism reviews*, 46(3), pp.261–82.
- Dyda, F., Klein, D.C. & Hickman, A.B., 2000. GCN5-RELATED N-ACETYLTRANSFERASES: A Structural Overview. *Annual Review of Biophysics and Biomolecular Structure*, 29, pp.81–103.
- Fedotcheva, N.I. *et al.*, 2008. Toxic effects of microbial phenolic acids on the functions of mitochondria. *Toxicology letters*, 180(3), pp.182–8.
- Fenton, W.A., Gravel, R.A. & Rosenblatt, D.S., 2001. Disorders of Propionate and Methylmalonate Metabolism. *The Online Metabolic and Molecular Bases of Inherited Diseases*, pp.2165–2194.
- Food and Drug Administration Center for Drug Evaluation and Research (CDER). Guidance for Industry Bioanalytical Method Validation. www.fda.gov
- Furst, D.E., Gupta, N. & Paulus, H.E., 1977. Salicylate Metabolism in Twins. Evidence suggesting a genetic influence and induction of salicylurate formation. *The Journal of Clinical Investigation*, 60(July 1976), pp.32–42.
- Gamage, N., 2005. Human Sulfotransferases and Their Role in Chemical Metabolism. *Toxicological Sciences*, 90(1), pp.5–22.
- Ghrist, A.C. & Stauffer, G.V., 1995. The Escherichia coli glycine transport system and its role in the regulation of the glycine cleavage enzyme system. *Microbiology*, 141, pp.133–140.

- Glatt, H., 2000. Sulfotransferases in the bioactivation of xenobiotics. *Chemico-Biological Interactions*, 129(1-2), pp.141–170.
- Gregersen, N., Kølvrå, 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. *Biochemical Medicine and Metabolic Biology*, 35(2), pp.210–218.
- Gregus, Z., Fekete, T., Halaszi, E., & Klaassen, C.D. (1996). Does hepatic ATP depletion impair glycine conjugation in vivo? *Drug metabolism and disposition*, 24(12), pp.1347-1354.
- Guillemette, C., 2003. Pharmacogenomics of human UDP-glucuronosyltransferase enzymes. *The Pharmacogenomics Journal*, 3(3), pp.136–158.
- Hagenbuch, B., 2010. Drug uptake systems in liver and kidney: a historic perspective. *Clinical pharmacology and therapeutics*, 87(1), pp.39–47.
- Han, L. *et al.*, 2002. Effect of glycine on the cell yield and growth rate of *Escherichia coli*: evidence for cell-density-dependent glycine degradation as determined by ¹³C NMR spectroscopy. *Journal of Biotechnology*, 92, pp.237–249.
- Hannig, G. & Makrides, S.C., 1998. Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends in biotechnology*, 16(2), pp.54–60.
- Hemsley, A. *et al.*, 1989. A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Research*, 17(16), pp.6545–6551.
- Hinson, J.A. & Forkert, P.G., 1995. Phase II enzymes and bioactivation. *Canadian Journal of Physiology and Pharmacology*, 73, pp.1407-1413.
- Innocenti, F. *et al.*, 2002. Haplotype structure of the UDP-glucuronosyltransferase 1A1 promoter in different ethnic groups. *Pharmacogenetics*, 12(9), pp.725–33.
- Irjala K., 1972. Synthesis of p-aminohippuric, hippuric, and salicyluric acids in experimental animals and man. *Annales Academiae Scientiarum Fennicae*, 154, pp.1–40.

- Ishikawa, T., 1992. The ATP-dependent glutathione S-conjugate export pump. *Trends in biochemical sciences*, 17(11), pp.463–468.
- Jackson, A.A. *et al.*, 1996. Human and Clinical Nutrition urinary Excretion of 5-L-Oxoproline (Pyroglutamic Acid) Is Increased in Normal Adults Consuming Vegetarian or Low Protein Diets¹ *2. Journal of Nutrition*, 126(11), pp.2813–2822.
- Jackson, A.A. *et al.*, 1987. Urinary excretion of 5-oxoproline (pyroglutamic aciduria) as an index of glycine insufficiency in normal man. *British Journal of Nutrition*, 58(2), pp.207–214.
- Jaiswal, A.K., 1994. Jun and Fos regulation of NAD(P)H: quinone oxidoreductase gene expression. *Pharmacogenetics*, 4, pp.1-10.
- Jakoby, W.B. & Ziegler, D.M., 1990. The Enzymes. *Journal of Biological Chemistry*, 265(34), pp.409–430.
- Jancova, P., Anzenbacher, P. & Anzenbacherova, E., 2010. Phase II drug metabolizing enzymes. *Biomedical papers of the Medical Faculty of the University Palacký, Olomouc, Czechoslovakia*, 154(2), pp.103–16.
- Jenner, A.M., Rafter, J. & Halliwell, B., 2005. Human fecal water content of phenolics: The extent of colonic exposure to aromatic compounds. *Free Radical Biology and Medicine*, 38(6), pp.763–772.
- Juliano, R.L. & Ling, V., 1976. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochimica Biophysica Acta*, 455, pp.152–162.
- Keller, W., 1842. M. Keller on the Conversion of Benzoic into Hippuric Acid. *Provincial Medical Journal and Retrospect of the Medical Sciences*, 4(92), pp.256–257.
- Kelley, M. & Vessey, D., 1994. Characterization of the Acyl-CoA: Amino Acid N-Acyltransferases from Primate Liver Mitochondria. *Journal of Biochemical and Molecular Toxicology*, 9(3), pp.153–158.

- Kelley, M. & Vessey, D.A., 1990. The Effects of Ions on the Conjugation of Xenobiotics by the Aralkyl-CoA and Arylacetyl-CoA N-Acyltransferases from Bovine Liver Mitochondria. *Journal of Biochemical and Molecular Toxicology*, 5(2), pp.125–135.
- King, C.D. *et al.*, 1999. Expression of UDP-glucuronosyltransferases (UGTs) 2B7 and 1A6 in the human brain and identification of 5-hydroxytryptamine as a substrate. *Archives of biochemistry and biophysics*, 365(1), pp.156–62.
- Knights, K.M., 1998. ROLE OF HEPATIC FATTY ACID:COENZYME A LIGASES IN THE METABOLISM OF XENOBIOTIC CARBOXYLIC ACIDS. *Clinical and Experimental Pharmacology and Physiology*, 25, pp.776–782.
- Knights, K.M. & Drogemuller, C.J., 2000. Xenobiotic-CoA ligases: kinetic and molecular characterization. *Current drug metabolism*, 1(1), pp.49–66.
- Knights, K.M. & Miners, J.O., 2012. Amino Acid Conjugation : A Novel Route of Xenobiotic Carboxylic Acid Metabolism in Man. *Encyclopedia of Drug Metabolism and Interactions*, pp.1–16.
- Knights, K.M., Sykes, M.J. & Miners, J.O., 2007. Amino acid conjugation: contribution to the metabolism and toxicity of xenobiotic carboxylic acids. *Expert opinion on drug metabolism & toxicology*, 3(2), pp.159–168.
- Kølvraa, S. & Gregersen, N., 1986. Acyl-CoA:glycine N-acyltransferase: organelle localization and affinity toward straight- and branched-chained acyl-CoA esters in rat liver. *Biochemical medicine and metabolic biology*, 36(1), pp.98–105.
- Kong, A.N.T. *et al.*, 2001. Induction of xenobiotic enzymes by the MAP kinase pathway and the antioxidant or electrophile response element (ARE/EpRE). *Drug metabolism reviews*, 33(3-4), pp.255–71.
- Kong, A.N.T. *et al.*, 2000. Signal transduction events elicited by natural products: role of MAPK and caspase pathways in homeostatic response and induction of apoptosis. *Archives of pharmacal research*, 23(1), pp.1–16.

- Krieger, I. & Tanaka, a N.D.K. a Y., 1976. Therapeutic Effects of Glycine in Isovaleric Acidemia. *Pediatric research*, 10, pp.25–29.
- Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, pp.680-685.
- Lehnert, W., 1983. N-Isovalerylalanine and N-isovalerylsarcosine: two new minor metabolites in isovaleric acidemia. *Clinica Chimica Acta*, 134, pp.207-212.
- Lebedev, A.T., 2013. Environmental Mass Spectrometry Environmental Mass Spectrometry. *Annual review of analytical chemistry*, 6, pp.163–189.
- Lees, H.J. *et al.*, 2013. Hippurate: the natural history of a Mammalian-microbial cometabolite. *Journal of proteome research*, 12(4), pp.1527–46.
- Lehnert, W., 1983. N-Isovalerylalanine and N-isovalerylsarcosine: two new minor metabolites in isovaleric acidemia. *Clinica Chimica Acta*, 134, pp.207-212.
- Levy, G., 1978. Pharmacokinetics of salicylate in man. *Drug metabolism reviews*, 9(1), pp.3-19.
- Liska, D.J., 1998. The detoxification enzyme systems. *Alternative Medicine Review*, 3(3), pp.187–198.
- Loots, D.T., Erasmus, E. & Mienie, L.J., 2005. Identification of 19 new metabolites induced by abnormal amino acid conjugation in isovaleric acidemia. *Clinical chemistry*, 51(8), pp.1510–1512.
- Matsuo, M. *et al.*, 2012. Designation of enzyme activity of glycine-N-acyltransferase family genes and depression of glycine-N-acyltransferase in human hepatocellular carcinoma. *Biochemical and biophysical research communications*, 420(4), pp.901–6.
- Mawal, Y.R. & Qureshi, I. a, 1994. Purification to homogeneity of mitochondrial acyl coa:glycine n-acyltransferase from human liver. *Biochemical and biophysical research communications*, 205(2), pp.1373–9.

- Meléndez-Hevia, E. *et al.*, 2009. A weak link in metabolism: the metabolic capacity for glycine biosynthesis does not satisfy the need for collagen synthesis. *Journal of Biosciences*, 34(6), pp.853–872.
- Meyer, U.A., 1996. Overview of enzymes of drug metabolism. *Journal of Pharmacokinetics and Pharmacodynamics*, 24, pp.449–459.
- Mortensen, H.M. *et al.*, 2011. Characterization of genetic variation and natural selection at the arylamine N-acetyltransferase genes in global human populations. *Pharmacogenomics*, 12(11), pp.1545–58.
- Moscow, J.A. & Dixon, K.H., 1993. Glutathione-related enzymes, glutathione and multidrug resistance. *Cytotechnology*, 12, pp.155–170.
- Nandi, D.L., Lucas, S. V. & Webster, L.T., 1979. Benzoyl-Coenzyme A: Glycine N-acyltransferase and Phenylacetyl-Coenzyme A Mitochondria A: Glycine N-acyltransferase from Bovine Liver Mitochondria. *Journal of Biological Chemistry*, 254(15), pp.7230–7237.
- Nishihara, K. *et al.*, 1998. Chaperone coexpression plasmids: Differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. *Applied and Environmental Microbiology*, 64(5), pp.1694–1699.
- Nishihara, K. *et al.*, 2000. Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. *Applied and environmental microbiology*, 66(3), pp.884–9.
- Nordström, A. *et al.*, 2004. Derivatization for LC-electrospray ionization-MS: A tool for improving reversed-phase separation and ESI responses of bases, ribosides, and intact nucleotides. *Analytical Chemistry*, 76(10), pp.2869–2877.
- Nortje, C. *et al.*, 2015. The simultaneous detection and quantification of p-aminobenzoic acid and its phase 2 biotransformation metabolites in human urine using LC-MS/MS. *Bioanalysis*, 7(10), pp.1211–1224.

- Oates, P.S. & West, A.R., 2006. Heme in intestinal epithelial cell turnover, differentiation, detoxification, inflammation, carcinogenesis, absorption and motility. *World Journal of Gastroenterology*, 12(27), pp.4281–4295.
- Ogier de Baulny, H. & Saudubray, J.M., 2002. Branched-chain organic acidurias. *Seminars in Neonatology*, 7, pp.65–74.
- Omiecinski, C.J. *et al.*, 2011. Xenobiotic Metabolism, Disposition, and Regulation by Receptors: From Biochemical Phenomenon to Predictors of Major Toxicities. *Toxicological Sciences*, 120(Supplement 1), pp.S49–S75.
- Parkinson, A., 2011. Biotransformation of Xenobiotics. In *Casarett and Doull's Toxicology: The Basic Science of Poisons*, Edition: 8th. pp. 133–224.
- Palmer, T. & Bonner, P.L., 2007. ENZYMES, Biochemistry, Biotechnology, Clinical Chemistry: Horwood Publishing Limited, Chichester.
- Pedraza, J.M. & Paulsson, J., 2008. Effects of molecular memory and bursting on fluctuations in gene expression. *Science* (New York, N.Y.), 319(5861), pp.339–343.
- Penner, N. *et al.*, 2010. Quantitative determination of hippuric and benzoic acids in urine by LC-MS/MS using surrogate standards. *Journal of pharmaceutical and biomedical analysis*, 52(4), pp.534–43.
- Penner, N., Woodward, C. & Prakash, C., 2012. Metabolizing enzymes and biotransformation reactions. In: *ADME-Enabling Technologies in Drug Design and Development*. Zhang D, Surapaneni S (Eds.). Wiley, Hoboken, NJ, USA, pp.545–565.
- Plotkin, J.B. & Kudla, G., 2011. Synonymous but not the same: the causes and consequences of codon bias. *Nature Reviews Genetics*, 12(1), pp.32–42.
- Porter, T.D. & Coon, M.J., 1991. Cytochrome P-450, multiplicity of isoforms, substrates, and Catalytic and regulatory mechanisms*. *Nature*, pp.13469–13472.

- Rechner, A.R. *et al.*, 2002. The metabolic fate of dietary polyphenols in humans. *Free radical biology & medicine*, 33(2), pp.220–235.
- Ritter, J.K., 2000. Roles of glucuronidation and UDP-glucuronosyltransferases in xenobiotic bioactivation reactions. *Chemico-Biological Interactions*, 129(1-2), pp.171–193.
- Rodriguez-Antona, C. *et al.*, 2010. Molecular genetics and epigenetics of the cytochrome P450 gene family and its relevance for cancer risk and treatment. *Human Genetics*, 450(127), pp.1–17.
- Rosano, G.L. & Ceccarelli, E.A., 2014. Recombinant protein expression in Escherichia coli: Advances and challenges. *Frontiers in Microbiology*, 5(April), pp.1–17.
- Rushmore, T.H. & Kong, A.N.T., 2002. Pharmacogenomics, regulation and signaling pathways of phase I and II drug metabolizing enzymes. *Current drug metabolism*, 3(5), pp.481–490.
- Sakuma, T., 1991. Alteration of urinary carnitine profile induced by benzoate administration. *Archives of disease in childhood*, 66(7), pp.873–875.
- Saltzman, A. & Caraway, W.T., 1953. Cinnamic Acid As a Test Substance in the Evaluation of Liver Function. *Journal of Clinical Investigation*, 32(8), pp.711–719.
- Sambrook, J. & Russell, D.W., 2001. Molecular cloning : A laboratory manual, Third edition: Cold Spring Harbour Laboratory Press.
- Sauna, Z.E. & Kimchi-Sarfaty, C., 2011. Understanding the contribution of synonymous mutations to human disease. *Nature Reviews Genetics*, 12(10), pp.683–691.
- Schachter, D. & Taggart, J.V., 1953. Benzoyl coenzyme a and hippurate synthesis. *Journal of Biological Chemistry*, 203, pp.925–934.
- Schachter, D. & Taggart, J.V., 1954. Glycine n-acylase : purification and properties. *Journal of Biological Chemistry*, 208, pp.263–275.

- Schilter, B. *et al.*, 1993. Phase I and phase II xenobiotic reactions and metabolism of the food-borne carcinogen 2-amino-3,8- dimethylimidazo[4,5-f]quinoxaline in aggregating liver cell cultures. *Biochemical Pharmacology*, 45, pp. 1087-1096.
- Sezonov, G., Joseleau-Petit, D. & D'Ari, R., 2007. Escherichia coli Physiology in Luria-Bertani Broth. *Journal of Bacteriology*, 189(23), pp.8746–8749.
- Sørensen, H.P. & Mortensen, K.K., 2005. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microbial cell factories*, 4(1), pp.1–8.
- Stergachis, A.B. *et al.*, 2013. Protein Evolution. *Science*, 341(13), pp.1367–1372.
- Sweetman L. & Williams J. C., 2013. Part 9: Organic acids. Chapter 93: Branched chain organic acidurias. In: Vale D, Beaudet AL, Vogelstein B, *et al.*, eds. *The Online Metabolic and Molecular Bases of Inherited Diseases*. New York: McGraw Hill.
- Tanaka, K. & Isselbacher, K.J., 1967. The Isolation and Identification Urine of Patients with Isovaleric of Msovalerylglycine Acidemia *. *Journal of Biological Chemistry*, 242(12), pp.2966–2972.
- Temellini A. *et al.*, 1993. Conjugation of benzoic acid with glycine in human liver and kidney: A study on the interindividual variability. *Xenobiotica*, 23, pp.1427–1433.
- Tew, K.D. & Ronai, Z., 1999. GST function in drug and stress response. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy*, 2(3), pp.143–147.
- Tsao, R., 2010. Chemistry and biochemistry of dietary polyphenols. *Nutrients*, 2(12), pp.1231–1246.
- Tukey, R. H. and Strassburg, C. P., 2000. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annual Review of Pharmacology and Toxicology*, 40, pp.581-616.
- Ure, A., 1841. On gouty concretions, with a new method of treatment. *Medico-chirurgical transactions*, 24, pp.30-35

- Van der Sluis, R. *et al.*, 2013. Characterisation of the influence of genetic variations on the enzyme activity of a recombinant human glycine N-acyltransferase. *Gene*, 515(2), pp.447–53.
- Van der Sluis, R. *et al.*, 2015. Conservation of the coding regions of the glycine N-acyltransferase gene further suggests that glycine conjugation is an essential detoxification pathway. *Gene*, 571(1), pp.126–134.
- 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. *Journal of biochemical and molecular toxicology*, 14(2), pp.102–9.
- Vatsis, K.P. *et al.*, 1995. Nomenclature for N-acetyltransferases. *Pharmacogenetics*, 5, pp.1–17.
- Vessey, D.A. *et al.*, 2003. Isolation, sequencing, and expression of a cDNA for the HXM-A form of xenobiotic/medium-chain fatty acid:CoA ligase from human liver mitochondria. *Journal of biochemical and molecular toxicology*, 17(1), pp.1–6.
- Vessey, D.A., Kelley, M. & Warren, R.S., 1999. Characterization of the CoA ligases of human liver mitochondria catalyzing the activation of short- and medium-chain fatty acids and xenobiotic carboxylic acids. *Biochimica et biophysica acta*, 1428(2-3), pp.455–62.
- Walter, S. & Buchner, J., 2002. Molecular Chaperones – Cellular Machines for Protein Folding. *Angewandte Chemie International Edition*, 41, pp.1098–1113.
- Wang, W. *et al.*, 2013. Glycine metabolism in animals and humans: implications for nutrition and health. *Amino acids*, 45(3), pp.463–77.
- Weatheritt, R.J. & Babu, M.M., 2013. The Hidden Codes That Shape Protein Evolution. *Science*, 342(6164), pp.1325–1326.
- Webster, L.T. *et al.*, 1976. Identification of Separate Acyl-CoA : Glycine and Activities in Fractions from Liver of Rhesus Monkey and. *Journal of Biological Chemistry*, 251(11), pp.3352–3358.

- Weiner, A.M., Platt, T. & Weber, K., 1972. Amino-terminal Sequence Analysis of Proteins Purified on a Nanomole Scale by Gel Electrophoresis. *Journal of biochemical and molecular toxicology*, 247(10), pp.3242–3251.
- Weinshilboum, R.M. *et al.*, 1997. Sulfation and sulfotransferases 1: Sulfotransferase molecular biology: cDNAs and genes. *The FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 11, pp.3–14.
- Wen, H. *et al.*, 2013. Enhanced phase II detoxification contributes to beneficial effects of dietary restriction as revealed by multi-platform metabolomics studies. *Molecular & cellular proteomics: MCP*, 12(3), pp.575–86.
- Wijeyesekera, A. *et al.*, 2012. Quantitative UPLC-MS/MS analysis of the gut microbial co-metabolites phenylacetylglutamine, 4-cresyl sulphate and hippurate in human urine: INTERMAP Study. *Analytical methods: advancing methods and applications*, 4(1), pp.65–72.
- Wilcox, B.J. *et al.*, 1999. N-acylglycine amidation: Implications for the biosynthesis of fatty acid primary amides. *Biochemistry*, 38(11), pp.3235–3245.
- Williams, H.R.T. *et al.*, 2010. Differences in gut microbial metabolism are responsible for reduced hippurate synthesis in Crohn's disease. *BMC gastroenterology*, 10, p.108.
- Williams, R. T., 1959. Detoxification Mechanisms: The Metabolism of Detoxification of Drugs, Toxic Substances and Other Organic Compounds, 2nd ed., *John Wiley and Sons*, New York, pp.188-194.
- Wu, G. *et al.*, 2013. Dietary requirements of “nutritionally non-essential amino acids” by animals and humans. *Amino Acids*, 44(4), pp.1107–1113.
- Xu, C., Li, C.Y.T. & Kong, A.N.T., 2005. Induction of phase I, II and III drug metabolism/transport by xenobiotics. *Archives of pharmacal research*, 28(3), pp.249–268.

- Yamamoto, R.S. *et al.*, 1971. 8-Hydroxyquinoline: chronic toxicity and inhibitory effect on the carcinogenicity of N-2-fluorenylacetamide. *Toxicology and applied pharmacology*, 19(4), pp.687-698.
- Zimniak, P., Awasthi, S. & Awasthi, Y.C., 1993. Phase III detoxification system. *Trends in biochemical sciences*, 18(5), pp.164-165.

Appendix A

Item	Product number	Supplier
Acetonitrile	015-4	Honeywell Burdick & Jackson
Acrylamide	1.00209.1000	Merck
Agar	BX 10.500	Merck
Agarose	9201	Techcomp
Ammonium sulfate	A2939	Sigma
Ampicillin	A9393	Sigma
BamHI	ER0051	Fermentas
Benzoic acid	242381	Sigma
Benzoyl-coenzyme A	B1638	Sigma
Bicinchoninic acid solution	B9643	Sigma
Bovine serum albumin	775 827	Roche
Buffer BamHI (10X)	B57	Fermentas
Bug Buster	70584-4	Novagen
Chloramphenicol	442513	Supelco
Coomassie brilliant blue	1.12553.0025	Merck
Copper sulfate solution	C2284	Sigma
DMSO	K0181	Thermo Scientific
dNTP Mixture	4030	Takara

Item	Product number	Supplier
DTNB	D218200	Sigma
DTT	43816	Sigma
E. cloni chemically competent cells	60107	Lucigen
EDTA	3658	Fluka
Ethanol 100%	100,983.25	Merck
Ethidium bromide	160539	Sigma
Ethyl acetate	1.09623.2500	Merck
ExTaq	RR001	Takara
Formic acid	94318	Sigma
Glacial acetic acid	1.00063.2500	Merck
Glucose powder	G1400	Melford
Glycerol	49780	Fluka
Glycine	1.04169.100	Merck
HindIII	ER0501	Fermentas
Hippuric acid	112003	Sigma
Hippuric acid-d5 stable isotope	H356702	Toronto Research Chemicals
HPLC water	1.15333.250	Merck
Hydrochloric acid	SAAR3063040LP	Merck
Isopropanol	1.09634.2500	Merck
L-Arabinose	A3256	Sigma
Lysonase	71230	Novagen

Item	Product number	Supplier
Macherey Nagel Nucleospin II PCR cleanup and gel extraction kit	740 609.50	Macherey Nagel
Macherey Nagel Protino Ni-TED 2000 kit	745 120.25	Macherey Nagel
Magnesium chloride	8.14733.0500	Merck
Magnesium sulfate	M2643	Sigma
Methanol	494437	Sigma
N,N' Bisacrylamide	130672	Sigma
O'GeneRuler	#SM 1173	Fermentas
Origami cells	71408-3	Novagen
PageRuler Plus	26619	Thermo Scientific
Polyethylene glycol-4000	8.07490.1000	Merck
Potassium chloride	AB004936	Merck
Promega PureYield plasmid midiprep kit	C#A1220	Promega
Protein Loading buffer, 4 x Dual Colour	R 1 011	Fermentas
RNase A (solution or powder)	R 6,513	Sigma
SDS	L4390	Sigma
Sodium acetate	71183	Fluka
Sodium chloride	S0520	Melford
Sodium hydrogen carbonate	1.06329.0500	Merck
Sodium hydroxide	1.06498.0500	Merck
T4 DNA ligase	EL 0011	Fermentas
TEMED	1.1 0732-026	Merck

Item	Product number	Supplier
Tetracycline	268054	Sigma
Tris	B2005	Melford
Triton X- 100	1.08603.1000	Merck
Tryptose	1.10676.0500	Merck
Yeast extract	HG00BX6.500	Merck

Appendix B

Human GLYAT protein sequence

MMLPLQGAQMLQMLEKSLRKSLPASLKVYGTVFHINHGPNPFLKAVVDKWPDFNTVVVCPQEQDMT
DDLDHYTNTYQIYSKDPQNCQEFLGSPELINWKQHLQIQSSQPSLNEAIQNLAIAIKSFVKQTQRILYMA
AETAKELTPFLLKSKILSPNGGKPKAINQEMFKLSSMDVTHAHLVNKFWHFGGNERSQRFIERCIQTFPTC
CLLGPEGTPVCWDLMDQTGEMRMAGTLPEYRLHGLVTYVIYSHAQKLGLGFPVYSHVDYSNEAMQK
MSYTLQHVPIPRSWNQWNCVPL