Amino acid *N*-acyltransferases in Human Metabolism

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Daniel Schulke

aus

Gelsenkirchen

Bonn 2022

Angefertigt mit der Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Jörn Oliver Saß
2. Gutachter: Prof. Dr. Michael Gütschow

Tag der Promotion: 03.11.2022
Erscheinungsjahr: 2022

Die vorliegende Arbeit wurde im Zeitraum von Mai 2016 bis Mai 2021 unter Anleitung von Prof. Dr. Saß im Fachbereich Angewandte Naturwissenschaften an der Hochschule Bonn-Rhein-Sieg angefertigt.

Table of content

1. Introduction

1.1. Glycine conjugation and coenzyme A homeostasis in the liver

In humans, several detoxification enzyme systems convert various hydrophobic xenobiotics to more hydrophilic conjugates. This detoxification includes the conjugation to endobiotics like sulfate, glucuronate or glycine. Glycine conjugation is one of numerous metabolic pathways that conjugates xenobiotics and was one of the first identified detoxification pathways in humans [151]. The central topic of this project is the glycine conjugation of benzoic acid resulting in the formation of hippuric acid, a reaction catalyzed by human glycine *N*-acyltransferase (GLYAT, EC. 2.3.1.13). The related enzymes GLYATL1, GLYATL2 and GLYATL3 (EC 2.3.1.13 applies to all) catalyze the same conjugation with other substrates. However, a potential involvement in regulation of metabolic diseases is hypothesized for GLYAT and GLYATL1. The identification of the GLYAT product hippuric acid as an amide conjugate between glycine and benzoic acid has been identified in 1953 [133] and the enzyme has been isolated from rabbit and bovine liver in 1954 [134]. Glycine conjugation results in less toxic compounds resembling other detoxification pathways. Compared to glucuronidation, the range of substrates for glycine conjugation is very limited, which may explain relatively scarce research in this field. Notwithstanding, glycine conjugation plays a major role in maintaining coenzyme A (coA) homeostasis in the liver [6]. Acyl-coA esters are key molecules in various anabolic and catabolic pathways.

Adequate levels of free coA within the cells are important, as various biological processes depend on maintenance of the sensible biological equilibrium of free and esterified coA. Disturbances in coA metabolism or changes in coA and acyl-coA levels may have serious consequences on the whole metabolism network [112]. So-called CASTOR diseases (coenzyme A sequestration, toxicity, or redistribution) were suspected to occur in various hereditary and acquired conditions, in which degradation of organic acyl-coA esters was impaired [112]. Resulting accumulation of coA esters and reduction of acetyl-coA and free coA will quickly trigger reaction cascades leading to certain diseases. The depletion of free coA is one of the most frequent consequences of acyl-coA accumulation with severe influences on energy metabolism. However, information on coA degradation in mammals is scarce. Nonetheless, a limitation of coA levels probably correlates with limiting energy metabolism [177]. The same study has introduced the concept of important metabolic re-arrangements that maintain the coA pool within the cell and are thus critical for normal mitochondrial function (e.g., fatty acid oxidation and urea cycle). The key problem therein was the shifted ratio between esterified and free coA to esterified coA. The L-carnitine buffers acyl groups of acyl-coA molecules illustrating its role for maintenance of free coA levels. The human GLYAT regulates the maintenance of the coA level in metabolic diseases [6]. The coA pool

influences various mitochondrial functions including gluconeogenesis, fatty acid oxidation and tricarboxylic acid and urea cycle. In addition, the valproate metabolism might be impaired by coA sequestration. The anti-epileptic drug valproate is metabolized to valproyl-coA, which is not a substrate for glycine conjugation. Hence, it may accumulate in the liver and deplete free coA possibly causing hepatic steatosis [55].

1.2. Glycine conjugation in urea cycle defects & organic acidemias

Glycine conjugation is a part of human biotransformation that facilitates excretion of xenobiotic substances from the diet or drugs. In general, human biotransformation can be differentiated into phase I and phase II reactions. Phase I includes oxidation reactions with cytochrome P450 enzymes (CYPs), reduction and hydrolysis reactions to introduce polar and reactive moieties into the molecules [83]. Phase II reactions for product excretion are divided into glucuronidation, methylation, acetylation, sulfation, glutathione conjugation and amino acid conjugation [77]. Amino acid conjugation and the involved enzymes are the main topics within the present study.

Beyond its role in detoxification, GLYAT also plays a central role in maintaining coA homeostasis [8, 112] as described above. It also possesses supportive functions in urea cycle defects or organic acidemias. In certain inborn errors of metabolism, e.g., in various amino acid catabolism defects and in selected fatty acid oxidation defects, acyl-coA esters accumulate. They are subsequently conjugated by GLYAT to the corresponding less-toxic acylglycines that are often less toxic than their parent compounds. The conjugation of coA-esters with glycine results in water-soluble acylglycines lowering their toxicity [6].

For several disorders, like isovaleric acidemia (IVA; isovaleryl-coA dehydrogenase deficiency) or propionic acidemia (PA, propionyl-coA carboxylase deficiency), coA-glycine conjugates have been identified [8]. In the leucine degradation defect IVA, the glycine administration represents one of the treatment options in addition to dietary protein restriction. The administration of L-carnitine represented an alternative [102, 128]. However, in a more recent study an IVA patient has been simultaneously treated with glycine and L-carnitine, which has resulted in a gradual increase of urinary acylcarnitine levels [28]. Isovaleryl-coA can also be conjugated to alanine and sarcosine [96], whereas alanine represents another substrate of GLYAT. Rapid accumulation of coA esters resulting from IVA and other metabolic disorders in early childhood [117] may cause acute encephalopathy [30]. Relatively small amounts of propionylglycine have been detected in the urine of PA patients [6, 8]. As described for bovine GLYAT, propionyl-coA is conjugated at a much lower rate compared with isovaleryl-coA [8, 116]. A similar comparison has not yet been reported for the human enzyme. Impaired phase II detoxification has been described with adverse reactions to pharmaceutical drugs and may be involved in pathogenesis of complex multifactorial

diseases, e.g., cancer [118, 152]. The genetic differences in phase I and phase II coding enzymes have been identified as crucial factors in defining cancer susceptibility and the carcinogenic power of environmental chemicals. Polymorphisms in the *CYP1A1* (aryl hydrocarbon hydroxylase) and *NAT2* (arylamine *N*-acetyltransferase) genes have been determined as biomarkers for genetic susceptibility to cancer [118]. The depletion of free coA is a major cause of problems in CASTOR diseases, which were described in Section 1.1. Due to accumulation of acyl-coAs the depletion of free L-carnitine is induced, confirmed by acyl-carnitines in blood and urine. Conjugation of exceeding acyl-coAs regulates and normalizes free coA levels and carnitines can also be restored [8, 112]. The metabolism of other amino acids is consequently disordered due the reduced nitrogen efflux and the accumulation of ammonia [49, 144]. To stabilize the nitrogen homeostasis, patients are treated with a low-protein diet combined with compounds that create alternative pathways for nitrogen excretion [63]. The therapy in patients with urea cycle defects often comprises benzoate as an ammonia scavenger, while phenylbutyrate, a prodrug of phenylacetate, and glycerol phenylbutyrate are used for the conjugation of L-glutamine [63] (Figure 1.1.).

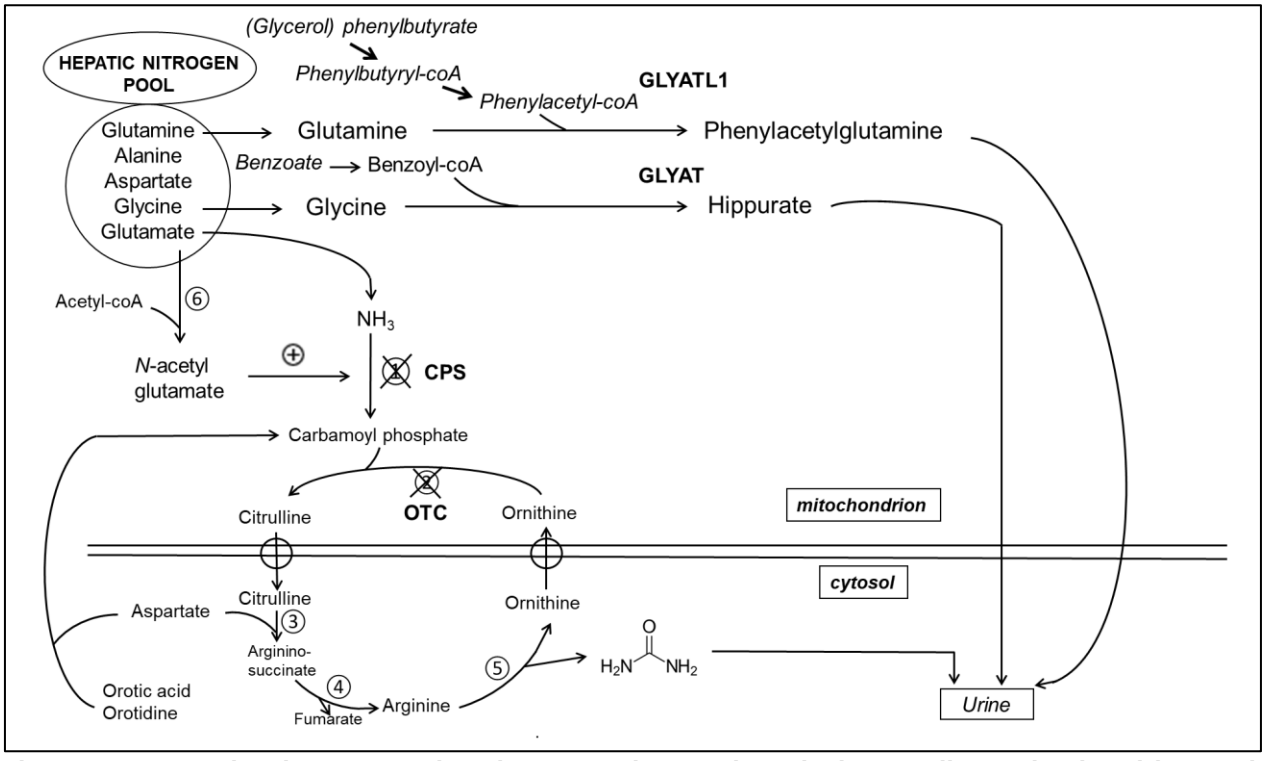


Figure 1.1: Hepatic nitrogen pool and urea cycle reactions in human liver mitochondrion and cytosol. Several amino acids are converted to urea via the Krebs-Henseleit cycle, trivial named as urea cycle. The prodrug *N*-acetyl glutamate enters the cycle via conversion to carbamoyl phosphate. To circumvent urea cycle the amino acid glycine can be converted to hippurate via benzoyl-coA, while glutamine can be converted to phenylacetylglutamine via phenylacetyl-coA conjugation. Hippurate and phenylacetyl-glutamine can pass the human body via urine circumventing the urea cycle (1: Carbamoyl phosphate synthetase I, 2: Ornithine transcarbamoylase, 3: Arginosuccinate synthetase, 4: Arginosuccinase, 5: Arginase).

Alternative pathways for nitrogen excretion catalyzed by GLYAT and GLYATL1 are shown in the upper part of Figure 1.1. Waste nitrogen levels are reduced via benzoylglycine excretion upon administration of sodium benzoate. For each mole of applied sodium benzoate, one mole of nitrogen is removed [49]. On the other side, phenylbutyrate is given for conjugation with glutamine to form phenylacetylglutamine. However, phenylbutyrate replaced phenylacetate because the latter has an unpleasant, clinging odor. For each given mole of phenylbutyrate, two moles of nitrogen are excreted [49]. Both compounds increase the removal of waste nitrogen by formation of water-soluble products excreted via urine. Nonetheless, two new phenylbutyrate metabolite categories (glucuronides, phenylbutyrate β-oxidation side products) have been identified in the urine of humans and rats [20]. There is a potential interference between the metabolism of phenylbutyrate and dietary carbohydrates and fatty acids. The nitrogen-excreting potential of phenylbutyrate might be diminished by ingestion of carbohydrates or lipids. Brunengraber and colleagues suggested the administration of phenylbutyrate between the meals for the therapy of urea cycle defects and underlined the toxicity of phenylbutyrate metabolism side products [20].

Carbamoylphosphate synthase (CPS) I deficiency and ornithine transcarbamoylase (OTC) deficiency are the most common urea cycle defects [49]. However, these defects can be treated with sodium benzoate and glycerol phenylbutyrate [144], which are substrates of GLYAT and GLYATL1. For dietary treatment, patients are retained in a low-protein nutrition and treated with sodium benzoate to conjugate glycine to form hippurate. As described above, glycerol phenylbutyrate is administered to conjugate glutamine to phenylacetylglutamine. Both conjugations decrease nitrogen levels in the human body by providing alternative pathways of nitrogen excretion [49, 63, 144].

In summary, sodium benzoate and sodium phenylbutyrate can be used intravenously for the treatment of hyperammonaemia episodes [144]. This was supplemented by the novel drug glycerol phenylbutyrate and the development of novel therapeutic avenues (gene therapy). Oral treatment options for urea cycle defects are administrations of sodium benzoate, glycerol phenylbutyrate, L-arginine, L-citrulline and N-carbamylglutamate [63]. Human GLYAT and GLYATL1 fulfill central tasks in urea cycle disorders and nitrogen homeostasis of the human body. The urea cycle is the major pathway for waste nitrogen that is produced within amino acid degradation in mammals. Brusilow has identified phenylacetylglutamine as waste nitrogen product in patients with urea cycle defects [21]. Treatment with phenylacetate or phenylbutyrate (sodium phenylbutyrate or glycerol phenylbutyrate) has resulted in normal levels of glutamine indicating that conjugation with glutamine and administration of phenylbutyrate can serve as an alternative nitrogen excretion pathway.

1.3. Interindividual variation in glycine conjugation

Interindividual variation in *GLYAT* gene influences the glycine conjugation. Variations in glycine conjugation has been observed for isovaleric acidemia (IVA) [39]. A variation in clinical outcome of IVA and susceptibility to glycine supplementation was identified for a South African population, where all patients have been homozygous for the same mutation in isovaleryl-coA dehydrogenase. It was assumed that interindividual variation in *GLYAT* might have influenced that. The interindividual variation in glycine conjugation may influence therapies of isovaleric acidemia and hyperammonaemia. It was suggested that GLYAT activity partly contributed to this, however that warrants further research [6].

Furthermore, significant interindividual variation in the glycine conjugation rate of xenobiotics has been described [6, 160]. For example, Campbell et al. have described that interindividual changes in glycine conjugation influence the excretion of industrial solvent m-xylene (2-methylhippuric acid) and salicyluric acid. As both substances are metabolized via the glycine conjugation pathway, a mutual inhibition on the formation of respective glycine conjugates has been observed [25]. The competition between the acyl-coA synthetase and GLYAT shed light on the considerable role of interindividual glycine conjugation in xenobiotic detoxification.

The first indications for decreased glycine conjugation in hepatitis patients have been described in form of impaired hepatic β-oxidation followed by lower ATP availability for ligation of benzoate to free coA [6, 90]. It has also been demonstrated that there is a significant interindividual variation in the rate of glycine conjugation for different xenobiotics. For instance, it has been shown that there was a high variation in the capacity of hippurate synthesis from the glycine conjugation with benzoate with a slight decrease of capacity for the elderly within human liver samples [147]. Dorne et al. have calculated a variation coefficient of 15-24 % in a large group of subjects for hippurate and salicylurate formation, respectively [41]. It has been demonstrated in 1976 that glycine administration is a therapy option in the management of ketoacidotic episodes in IVA [92].

The conjugation with glycine is an important metabolic pathway for drugs e.g., with carboxylic groups or salicyluric acid [98, 147]. Temellini et al. have demonstrated that glycine conjugation in the human liver and kidney is subjected to an interindividual variability and that is normally distributed within one population [147]. It is often not clear, whether variation in the rate of glycine conjugation has resulted from differences in the xenobiotic/medium-chain fatty acid: coA ligases (ACSMs) or GLYAT activities [6]. The rate of glycine conjugation might be influenced by liver diseases, as studies with higher variation in glycine conjugation in liver disease patients have indicated [41]. The compound p-aminobenzoic acid, which is conjugated with glycine in the liver,

has served as indicator for liver function. However, liver disease patients are indirectly manifested by interindividual rates of glycine conjugation.

Benzoate treatment in high doses can lead to reduced plasma glycine levels (e.g., for reduction of seizures) and increases alertness in patients suffering from the glycine cleavage system defect nonketotic hyperglycinemia (NKH) [157]. Defects in the glycine cleavage system critically affect the benzoate metabolism. The so-called glycine-index highlights the sensitive relation between the benzoate dose and glycine availability. In a more recent publication, van Hove et al. have described the reduction of glycine plasma levels upon administration of sodium benzoate in patients with attenuated NKH and severe NKH [158]. However, patients with attenuated NKH have required lower doses of benzoate than severe NKH patients. Because sodium benzoate has an effect in the liver and kidney, it is not clear, to what extent it reduces brain or cerebrospinal fluid glycine. Several side effects of benzoate and its potential toxicity when using >2.5 mmol/L have been described. Nonetheless, that can be regulated by withholding benzoate, glycine administration or hemodialysis [158].

In summary, glycine conjugation represents a detoxification pathway that increased the water solubility of organic acids to facilitate urinary excretion. Interindividual susceptibilities to glycine administration were also observed within the treatment of isovaleric acidemia [39]. The studies on interindividual variation in glycine conjugation report variations in the entire glycine conjugation pathway including coA ligation in the step before [155]. Next to other factors, e.g., coA and ATP availability and ACSM (mitochondrial xenobiotic/medium-chain fatty acid: coA ligases) activity, the individual GLYAT activity, which may also partly account for the glycine conjugation rate, is crucial for that process.

1.4. Sequence variations of Amino acid *N*-acyltransferase genes

1.4.1. Human *GLYAT* gene

The human *GLYAT* gene is located on chromosome 11 at position 11q12 spanning over 23,000 bp containing six exons, separated by five introns [6]. Within the gene 182 coding SNPs have been reported up to now [179] The gene possesses six transcript variants, 226 orthologues and four paralogues [180].

In this project the transcript variant 1 (Transcript-ID: ENST00000344743) (NM_201648.2) was analyzed as all references reported in the literature are based on this variant. The gene product encodes GLYAT isoform a with a size of 34 kDa consisting of 296 amino acids. A splice variant retaining part of intron five and lacking exon six, encodes for shorter isoform b with 163 amino acids (NM_005838.3).

1.4.2. Identification of human *GLYAT* polymorphisms in the literature

Only two studies have elucidated cohorts regarding allele frequencies of SNPs in the *GLYAT* gene [101, 175]. In a small cohort of a Japanese and French Caucasian individuals the c.467A>G variant of *GLYAT* has been the dominant allele with frequencies of 85 % in the Japanese and 97 % in the French Caucasian population.

For a Japanese cohort five SNPs have been identified by resequencing the entire coding region and exon-intron junctions of 95 individuals. One of the identified polymorphisms has resulted in an amino acid exchange (p.(Arg131His)) in Japanese individuals. However, this polymorphism has been absent in the Caucasian cohort [175]. They identified three other polymorphisms (7527T>A, 21289G>A, 21364A>G) resulting in amino acid exchanges, whereas at least one is predicted to change the enzyme activity.

For a French Caucasian cohort of 55 individuals seven SNPs of *GLYAT* have been identified including two in the 5´-flanking region of the gene, two in intron five and three missense mutations in the exons two, five and six. Analyses of the p.(Ser17Thr) and p.(Arg199Cys) missense mutations in predicted secondary structures have suggested impairments of GLYAT protein activity [101]. Polymorphism investigations have been executed using PCR-SSCP (polymerase chain reaction-single strand conformation polymorphism) and sequencing strategies including bioinformatical approaches. In this way even effects of protein secondary structures have been elucidated. Distribution of sequence variation allowed characterization of four different alleles in addition to the reference allele. Due to the suspicious high allele frequency of c.467A>G, this variant may have been regarded as the real *GLYAT* wild-type [101], which has been corroborated by enormous high enzyme activity of p.(Asn156Ser) variant [129, 152, 154].

1.4.3. Earlier studies on enzyme characterizations of human GLYAT

Enzyme characterization of the GLYAT enzyme has been performed after overexpression of the gene in the *E. coli* Origami 2(DE3) strain and HEK293 cells. Detailed enzyme studies, especially with an overexpressed variant from eukaryotic cells, are missing so far. Only one publication is known, where *GLYAT* wild-type and six sequence variants have been overexpressed in *E. coli* and enzymatically analyzed [152]. Furthermore, the GLYAT enzyme has been investigated in a native isolated form from liver and kidney of several mammals [8, 60, 84, 85, 91, 108].

Suitability of the E. coli Origami 2(DE3) for the overexpression of amino acid N-acyltransferases

E. coli Origami 2(DE3) is a special strain, which is optimized for the overexpression of proteins containing disulfide bonds, which sensitively stabilizes protein structure and are present in 21 % of all known protein structures [82]. The human GLYAT, GLYATL1 and GLYATL2 have seven, six and seven cysteines in their polypeptide sequence, respectively. The natural *E. coli* contains

thioredoxin and glutaredoxin system that reduce disulfide bonds in bacteria cytoplasm, thus destroying tertiary protein structure [124]. This strain observes mutations in the *thioredoxin* and *glutathione reductase* genes [22] enabling sufficient folding of heterologous proteins via more oxidative cytoplasm [13, 174]. Beyond that, a higher bioactivity for an overexpressed target protein (bovine fibroblast growth factor BbFGF) has been confirmed [174]. The *E. coli* Origami 2(DE3) possesses impaired *thioredoxin* and *glutaredoxin reductase* genes. Those, in combination with co-expressed GroES/GroEL [17], the *Trx*-tag fusion [33] and the compatible pET32 plasmid system [174] cover the certain compulsory requirements for the overexpression of eukaryotic GLYAT in a prokaryotic system. It has been used several times for the overexpression of GLYAT [129, 152, 154]. On the other side, mutations within these genes resulted in a decreased growing rate of *E. coli* Origami 2(DE3) in contrast to other *E. coli* strains [22, 124] - a problematic challenge within the project, which will be described more precisely (Section 4.4.). Notwithstanding, the mutations in *thioredoxin* and *glutathione reductase* genes optimize the tertiary structure of the proteins as it is described above.

The *N*-terminal fusion with a thioredoxin tag strongly increases the expression level and solubility of target proteins, in particular due to its solvent accessible C-terminus [33]. The co-expression of chaperone system GroES/GroEL contributes to a sufficient target protein overexpression and folding. It has been demonstrated for the pig liver esterase [17] and ensured within the present study by the kindly provided pGro7 plasmid.

One disadvantage of *E. coli* is the limited PTM system for overexpressed proteins [37]. PTMs are necessary for functionality of proteins described by altering structure/function relationships, complex formations, enzyme catalysis and other biomolecule interactions [23]. Authentic PMTs may not be guaranteed by *E. coli* as it has no cell organelles, enzyme components and intracellular compartmentalization. However, there are several studies that have confirmed a reduced PTM pattern for *E. coli* proteins: chemotaxis protein CheY needed to be acetylated and phosphorylated to influence the interaction partner FliM from the flagellar motor inducing a change of direction within flagellar rotation which is an example for a prokaryotic acetylation and phosphorylation network [23].

Acetylation research in bacteria is an evolving process and new acetylated proteins are continually identified [26], such as in many antibiotic producing bacteria like *Saccharopolyspora erythraea,* where several acetylated proteins are involved in the biosynthesis of the commercially used antibiotic erythromycin [26, 93]. Beyond that, it has been demonstrated that prokaryotes “own eukaryotic-type” phosphorylation proteins such as serine/threonine and tyrosine kinases in addition to the well characterized 2-component-regulatory-systems with phosphor relay from

sensor histidine kinase to acceptor aspartic acid binding to promotor region of target gene. This has been exemplary represented by AcuA protein system from *Bacillus subtilis* [56].

1.4.4. Human *GLYATL1* gene

The human *GLYATL1* gene is located on chromosome 11q12.1 containing 14 exons [181]. It has 21 transcripts (splice variants), 150 orthologues and four paralogues [182]. The transcript variant 4 (Transcript-ID: ENST00000317391) (NM_001220496.2) was part of the present study. It encodes for the protein isoform 1 having a size of 302 amino acids. The protein has two main isoforms (isoform 1: 302 amino acids, isoform 2: 333 amino acids) with 19 remaining, smaller isoforms, which represents splice variants and other truncated, non-functional forms of the protein. For the whole gene 368 coding SNPs have been detected [183]. So far, no systematic analyses or clinically relevant polymorphisms are known for this gene.

1.4.5. Human *GLYATL2* gene

The gene of human *GLYATL2* is located on chromosome 11q12.1 as well consisting of nine exons [184]. It has three transcripts, 167 orthologues and four paralogues [185]. In the present study the transcript variant 1 (Transcript-ID: ENST00000287275) (NM_145016) was studied. It encodes for a protein with a size of 294 amino acids. The remaining transcript variants are shorter, non-functional forms of the protein. For this gene 308 coding SNPs were counted from databases [186]. So far, no systematic comparison of these polymorphisms or clinical relevance of the latter are known for this gene.

1.4.6. Human *GLYATL3* gene

For the human *GLYATL3* gene, which is located on chromosome six, two transcripts are known so far [187]. The gene possesses 305 coding SNPs [188] demonstrating the high demand on research for the future. The *GLYATL3* was not part of the present study.

1.5. Amino acid conjugation in Human Metabolism

Amino acid *N*-acyltransferases catalyze amino acid conjugation of xenobiotic carboxylic acids in the human body, a poorly characterized pathway [90]. One reason for that may be the small number and limited diversity of substrates for amino acid conjugation compared with the cytochromes P450 (CYP) or uridine-5´diphosphate glucuronosyltransferase (UGT). These are the so-called premier drug-metabolizing enzymes with high variety of known substrates. Even if CYP or UGT drug metabolizing routes have been characterized more precisely, amino acid conjugation is the predominant route of salicylic acid metabolization with salicyluric acid which is responsible for 75 % of the Aspirin® excretion in the urine [90]. Amino acid conjugation is the second step of xenobiotic elimination followed by initial formation of xenobiotic coA-ester catalyzed by ATP dependent ACSMs enzymes. Two human liver medium-chain ligases belong to that group and

have been isolated and characterized in 1999 [159]. HXM-A (48 kDa) and HXM-B (49 kDa) exhibited high activity towards various xenobiotic substrates, like arylacetic carboxylic acids e.g., aromatic acids such as benzoate, propionate, hexanoate and octanoate [90]. Nonetheless, the substrate specificities of the other ACSMs have been poorly characterized. Followed by coupling xenobiotics to coA via HXM-A/B ligases as a second step, the transfer of acyl group to amino acid catalyzed by amino acid *N*-acyltransferases occurs. Due to limitation of amino acid conjugation to the mitochondria formation of xenobiotic-coAs, which are resistant against hydrolysis and not preferably amino acid conjugated (e.g., valproic acid), may sequester intermitochondrial coA and perturb β-oxidation. Nevertheless, amino acid conjugation in humans is essential regarding treatment of NKH and hyperammonaemia [64, 157].

Recently, the Uniprot database published protein structure predictions for the human amino acid *N*-acyltransferases (Figure 1.2.).

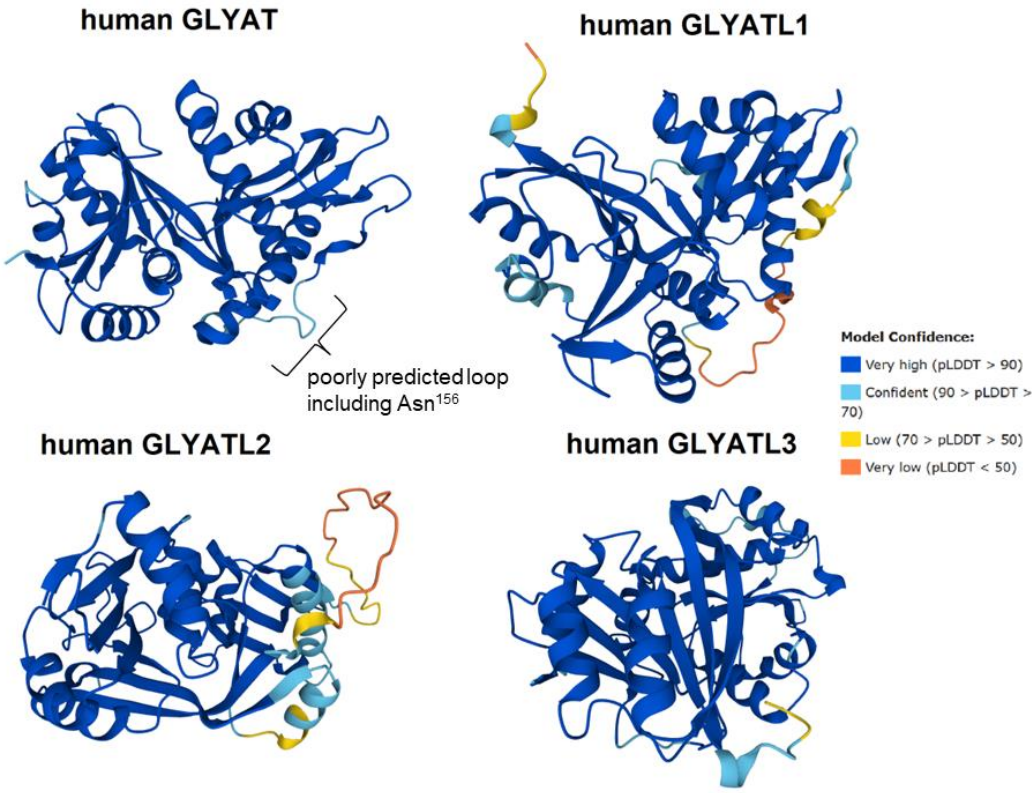


Figure 1.2: Protein structure predicitions on human GLYAT [189], GLYATL1 [190}, GLYATL2 [191] and GLYATL3 [192]. The models were produced by the AlphaFold software, which specifies a per-residue confidence score (pLDDT) between 0-100.

The amino acid Asn^{156} is localized on a poorly predicted loop from Lys^{159} to Met^{167} in the human GLYAT [153], which is shown in the lower right corner of the GLYAT molecule (Figure 1.2.).

Indeed, for the human GLYATL3 structure information were shown, although only one study on the mouse GLYATL3 is published so far [79].

1.5.1. GLYAT (Glycine *N*-acyltransferase)

The human GLYAT is a monomeric phase II detoxification enzyme and has been first isolated in 1954 from rabbit and bovine liver [134]. Following purification from human liver mitochondria [169], it has been demonstrated that (at least) 2 enzymes exist: one was specific for glycine using butyryl-coA and benzoyl-coA as donors, whereas the other preferably conjugates phenylacetyl-coA or indoleacetyl-coA with glutamine. The second enzyme has been described later as GLYATL1 [178]. The *N*-acyltransferase selectivity via glycine conjugates within the urine has been shown [90]. GLYAT has been identified in bovine and human liver and kidney mitochondria [108, 116, 156] and have reached its maximum around the 18th month of human life [109]. It has a size of 34 kDa with a polypeptide chain of 296 amino acids [189] and detoxifies preferably dietary benzoates by conjugation with glycine (Figure 1.3.). Those benzoates are mainly found in plant materials [6, 172].

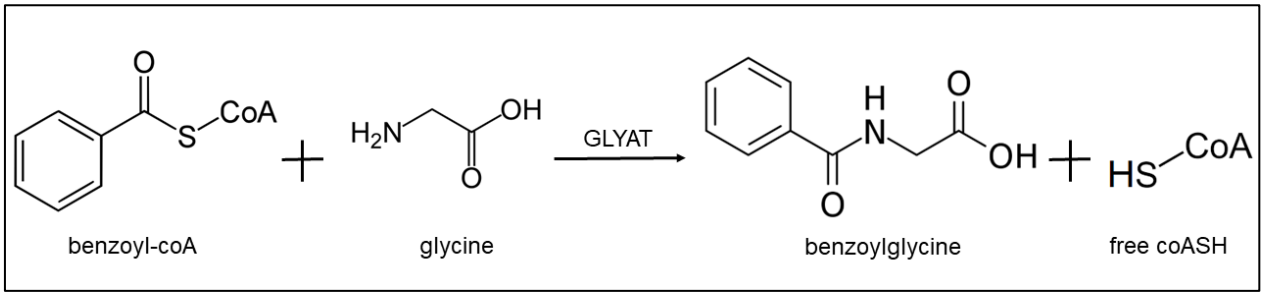


Figure 1.3: GLYAT catalyzes the amino acid conjugation of glycine with benzoyl-coA resulting in benzoylglycine and free coASH.

While the substrate width of the human enzyme is restricted, the bovine GLYAT conjugates isovaleryl-coA, propionyl-coA, acetyl-coA, salicyl-coA and tiglyl-coA with glycine [8]. Salicyl-coA is an intermediate of the degradation of acetylsalicylic acid, which is the active substance of the Aspirin® tablets. The major route of salicylic acid metabolism (followed by hydrolysis of acetylsalicylic acid) is glycine conjugation, whereas glucuronide conjugation represents the other metabolic route [74]. Nonetheless, a small amount of salicylate undergoes aromatic hydroxylation yielding gentisic (2,5-dihydroxybenzoic) acid, which is also conjugated with glycine to form gentisuric acid.

Kinetic studies with GLYAT have confirmed that the reaction mechanism is described as sequential reaction with acyl-coA substrate binding first to the enzyme, followed by an addition of glycine, before dissociation of free coA and releasing the peptide product as a last step [90, 116]. So far, the characterization of GLYAT kinetics has usually been performed using a colorimetric assay based on DTNB reacting with free thiol groups [46]. No X-ray structure information has been

reported for human GLYAT so far. The Gcn5-related *N*-acyltransferases (GNAT) enzyme superfamily includes GLYAT due to high homology to the amino acid *N*-acyltransferases and acyl-coAs and glycine as substrates [44, 161]. Numerous genome sequencing efforts of the last decades coupled with bioinformatic approaches have revealed above 10,000 members of GNAT family in all areas of life. First and most popular members of GNAT enzymes are the bacterial aminoglycoside acetyltransferases forcing resistance against the antibiotics micin and kanamycin [161]. Furthermore, the well-described histone *N*-acetyltransferases belong to the GNAT enzymes. The known reaction mechanism of these enzymes may help to understand GLYAT reaction mechanism.

The K_M value of human GLYAT for benzoyl-coA has been reported between 13 µmol/L and 57.9 mmol/L [8, 60, 84, 85, 90, 108, 152, 156]. The v_{max} published for human GLYAT wild-type were also highly variable as summarized by van der Sluis [152]. v_{max} has been reported between 543 and 17,100 nmol/(min*mg). Notwithstanding, the summarized reports describe different source materials and different isolation procedures, which were used for determination of the kinetic properties.

The active center of GLYAT has been predicted to be located between amino acid 128 and 178 of the primary structure [107]. For the bovine GLYAT, a 3D-structure modelling has been performed [5]: they exhibited a highly conserved Glu^{226} residue in a structural superposition of SNAT, SSAT and Esa1 structures (other GNAT enzymes). The catalytic residue side chains bind acetyl-coA to SNAT binding pocket.

Badenhorst and colleagues have revealed that Glu^{226} of bovine GLYAT coincides spatially with the catalytic residue His^{120} of serotonin *N*-acyltransferase [5]. However, this residue has been highly conserved in the top 40 GLYAT homologs obtained by databank search, which underlined its functional relevance [5]. Similarities have been identified in studies with other GNAT enzymes with known catalytic mechanisms. Based on this relevance of Glu^{226}, a corresponding missense mutant p.(Glu226Gln) and pH dependency testing has been performed by Badenhorst and colleagues. At a pH 8.0 the activity of the mutant was decreased as expected, which was compensated by an increased pH. Catalytic mechanism with Glu^{226}, as base catalyst for glycine, revealed p.(Glu226Gln) activity in a pH dependent fashion (maximum activity at pH 9.6). Thus, Glu^{226} facilitates the glycine deprotonation and hence nucleophilic attack on the acyl-coA substrate.

Since there are various studies on the human and the bovine GLYAT, one publication has treated the overexpression of the mouse GLYAT, which preferred glycine and benzoyl-coA as substrates accordingly [38].

1.5.2. GLYATL1 (Glutamine *N*-phenylacetyltransferase)

The L-glutamine conjugation activity has been detected in human liver mitochondrial fractions [169]. The Glycine *N*-Acyltransferase like protein 1 (GLYATL1) enzyme have been described later [178]. Webster et al. have revealed phenylacetyl-coA and glutamine as preferred substrates for the human glutamine *N*-acetyltransferase (Figure 1.4.).

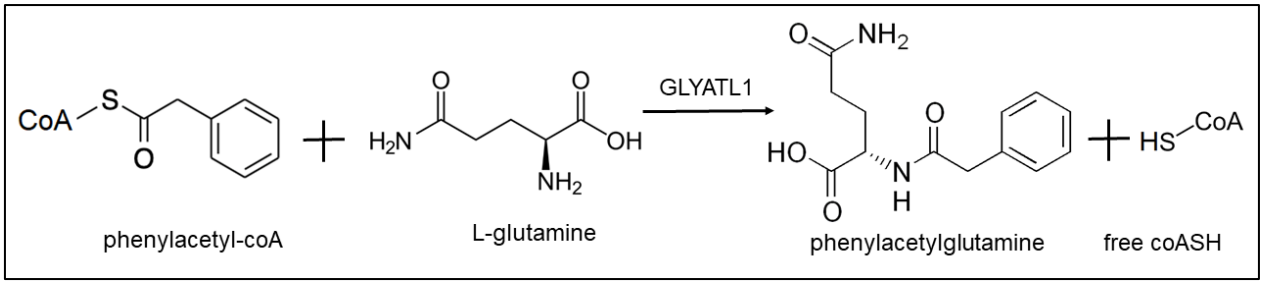


Figure 1.4: GLYATL1 catalyzes amino acid conjugation of L-glutamine with phenylacetyl-coA resulting in phenylacetylglutamine and free coASH.

Human *GLYATL1* encodes for two protein isoforms: isoform 1 with a size of 35 kDa (302 aa) and isoform 2 with a size of 38 kDa (333 aa) [190]. Both isoforms are produced by alternative splicing. Isoform 1 has been chosen as the canonical sequence and observes a shorter *N*-terminus compared with isoform 2. There are no differences known regarding localization. Due to 100 % identity in polypeptide chain from Leu^{56} until the C-terminus there are no differences in the active center. According to the Transcript-ID: ENST00000317391 and the comparison with literature data [178], the present study treats the isoform 1.

Human GLYATL1 expression has been highly detected in liver, while weak detection was shown in kidney and stomach. Zhang et al. have identified cytosolic localization of GLYATL1 in COS-7 cells [178]. However, in another study glutamine *N*-acyltransferase activity has been identified in mitochondrial fractions from liver of rhesus monkey and man [169]. In opposition, it has been also suggested to be localized within the ER [107]. In the database [190] the enzyme is declared as mitochondrial.

Interestingly, GLYATL1 has been supposed to activate HSE (heat shock elements) signaling pathway dose-dependently as demonstrated in HEK293T cells with reporter gen studies of luciferase activity [178]. So far, no correlation has been demonstrated between the announced detoxification activity with stress-regulatory pathway of heat shock response. The kinetic properties of human GLYATL1 have been also analyzed [85] and a sequential ordered reaction mechanism for the bovine enzyme has been already described [116]. The K_D value for phenylacetyl-coA has been proposed to be 14 µM, while K_M for glutamine it has been reported with 120 +/- 45 mM [sic!] [85]. V_{max} has been reported on 1.5 µmol/min*mg for the human enzyme. The salt sensitivity of the enzyme has been described in addition, which was demonstrated by

activity decrease of 30 % by 80 mM KCl [116]. Glutamine, which serves as substrate for GLYATL1, critically accumulated in inborn errors of urea cycle associated with hyperammonaemia [49] as previously described (Section 1.2.).

Nonetheless, recent studies ascertained that GLYAT and GLYATL1 possess functions beyond detoxification. Human GLYAT was downregulated in human breast cancer tissue [150], whereas GLYATL1 expression level was reduced in hepatocellular carcinoma [62]. The same finding was described for the human GLYAT eight years before [107].

Moreover, the amino acid *N*-acyltransferases may also contribute to the regulation of several metabolic diseases e.g., isovaleric acidemia [92] and urea cycle defects [144]. In particular, the polymorphisms of genes are worth to be investigated to get a more fundamental understanding of the enzyme function involved in these diseases. The effects of certain deficiencies, of GLYAT, GLYATL1 or GLYATL2, may shed new light on the treatment and therapy options of certain inborn errors of metabolism.

1.5.3. GLYATL2 (Glycine *N*-acyltransferase like protein 2)

The human GLYATL2 synthesizes various *N*-acylglycines via conjugation of long chain fatty acid coA-esters with glycine (Figure 1.5.).

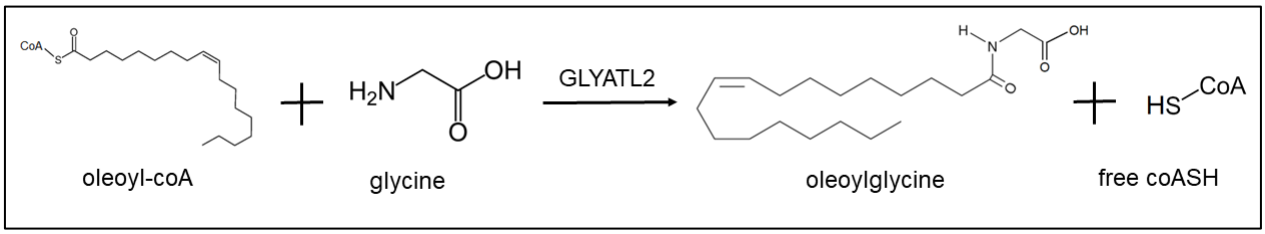


Figure 1.5: GLYATL2 catalyzes amino acid conjugation of glycine with oleoyl-coA resulting in oleoylglycine and free coASH.

The recombinantly overexpressed GLYATL2 enzyme efficiently conjugates oleoyl-coA, arachidonoyl-coA and other medium- and long-chain acyl-coAs to glycine with special preference for oleoyl-coA and glycine [165]. The Lys^{19} residue is highly conserved and critical for the enzyme since it must be deacetylated for the full enzyme function [166]. The enzyme has been predicted to be localized to the ER with high mRNA expression in salivary gland, trachea, spinal cord, and skin fibroblasts.

1.5.3.1. *N*-acylglycines and CB1 receptors

This expression pattern in combination with high levels of *N*-acylglycines in skin and lung [19] may indicate a certain role of GLYATL2 in barrier function or immune response [165]. Another interesting aspect has been illustrated by chemical similarity of *N*-acylglycines to endocannabinoids without activation of endocannabinoid receptors. Endocannabinoid receptors had been separated into two groups, while group one (CB1) has been mainly located within

neocortex, hippocampus, basal ganglia, cerebellum, and brainstem [68], whereas group two (CB2) has been mainly localized to microglia cells of macrophage system of CNS [105]. However, CB1 have been identified in many olfactory brain areas of rodents and involvements in specific odor-related processes [148]. A structural similarity of endocannabinoids and *N*-acyl glycines, which are GLYATL2 products, have been described [165] that might highlight a possible GLYATL2 attachement to the brain, especially to olfactory areas. *N*-arachidonoylglycine has been the first identified *N*-acylglycine in the rat brain [73]. This endogenous substance has been found in brain, spinal cord, small intestine, and kidney. It has been also suggested to be a ligand for G-protein coupled receptors such as GPR72 or GPR92 [165]. Next to *N*-arachidonoylglycine, the *N*-oleoylglycine regulates fundamental neurochemical processes in rats such as increasing hypothermia and decreasing locomotion in rats [27]. Due to the emerging *N*-acylglycines as a potent family of lipid signaling molecules, GLYATL2 may obtain roles moreover than detoxification of xenobiotics.

1.5.3.2. GLYATL2 protein structure and activity

The human GLYATL2 enzyme consists of 294 amino acids representing a protein of 34 kDa [191]. The human GLYATL2 has been hypothesized to be localized in the ER [107]. The protein contains a C-terminal dilysine motif (Lys-Lys-Tyr-Cys), which has been identified as ER retention signal [176]. The potency of di-lysine motifs suggests the mediated ER localization. There is a high distribution of these ER signaling motifs for membrane proteins type I and III [176]. However, the GLYATL2 protein is lacking transmembrane domains [165].

Kinetic studies were performed using oleoyl-coA and glycine as substrates. The K_M for formation of *N*-oleoylglycine was 4.4 µM with a v_{max} of 933 nmol/min*mg [165]. Above 50 µM of oleoyl-coA GLYATL2 has demonstrated a slight substrate inhibition. Oleoyl- and arachidonoyl-coA were the preferred substrates for human GLYATL2 [164]. The bovine enzyme showed conjugation activity for alanine, serine and glutamic acid [156]. *N*-oleoylglycine, the main product of GLYATL2 catalysis, was first identified in mouse neuroblastoma N18TG2 cells [111] with proposal of involvement of oleamide biosynthesis, a sleep-inducing component. GLYATL2 has been thus proposed to be involved in oleamide biosynthesis [165], which produces *N*-oleoylglycine as intermediate substrate. As expected, the highest mRNA expression of GLYATL2 has been detected in salivary gland, not in parotid gland, indicating the very high expression of peptidylglycine α-amidating enzyme (PAM), which is an important precursor in oleamide biosynthesis [34].

1.5.3.3. Critical Lys^{19} residue of GLYATL2

The Lys^{19} is highly conserved and catalytically important for the GLYATL2 enzyme [166]. The p.(Lys19Gln) and p.(Lys19Arg) mutations has resulted in 50-80 % decreased enzyme activity, which highlights the relevance of that residue for the enzyme function. LC-MS/MS studies have shown deacetylation of Lys^{19} in GLYATL2 wild-type indicating requirement of deacytelation for fully protein function. The acetylation of Lys^{19} has resulted in an inhibitory effect on the enzyme. However, the treatment with nicotinamide (inhibitor of sirtuin deactelyases, which cause acetylated enzymes) resulted in the acetylation of GLYATL2, which was thus inhibited [166].

The crystal structure of human GLYAT proteins has not yet been solved, which complicated to predict, where Lys^{19} is located and how it affects the 3D structure of the protein or substrate binding. The human GLYATL2 produces *N*-acylglycines, which furthermore regulates lipid signalling pathways [166].

A proteomic survey has identified a huge amount of lysine acetylated non-nuclear proteins in human and mouse models, whereas 2 of the 195 identified enzymes were member of human glycine *N*-acyltransferase proteins [88].

1.5.4. GLYATL3 (Glycine *N*-acyltransferase-like protein 3)

The human GLYATL3 is a long-chain specific glycine conjugating enzyme with possible involvement in fatty acid oxidation or biosynthetic pathways of *N*-acylglycines. Like GLYATL2, which prefers oleoyl-coA as main substrate, GLYATL3 conjugates primary fatty acid amides (PFAMs) and long-chain fatty acids [79]. For checking that, the *mGLYATL3* gene has been knocked down by siRNA in $N_{18}TG_2$ mouse cells, which produces fatty acid amides including long-chain *N*-acylglycines and PFAMs [111]. Knock-down of *mGLYATL3* has resulted in decreased level of these compounds supporting the biosynthetic pathway of fatty acid by acyl-coA synthetase followed by glycine conjugation of corresponding acyl-coA. Hence, the knock-down of *mGLYATL3* has confirmed a role of *N*-acylglycines as intermediates in biosynthesis of long-chain PFAMs [110]. GLYATL3 has been predicted to be localized within mitochondria and has a size of 32.7 kDa [192]. Human GLYAT has 35 % sequence similarity to GLYATL3, which underlines sequential differences. So far, no studies have been reported on human GLYATL3 protein. The GLYATL3 enzyme was not part of the present study.

1.6. Aims of the thesis

Mutations of amino acid *N*-acyltransferase genes may affect enzyme activity. The first aim in the present study was to conceptualize the prokaryotic overexpression system and to establish functional assays for GLYAT and related enzymes in the Rheinbach laboratory. Selected novel sequence variants without functional assessments should be analyzed regarding kinetic properties to provide a basis for assessing their potential pharmacogenetic impact. Due to a lack of data on overexpression and characterization of human GLYATL1 and GLYATL2, they should be overexpressed and functionally tested accordingly.

As no systematic, eukaryotic overexpression studies on human GLYAT, GLYATL1 and GLYATL2 exist thus far, the second aim was to attempt the overexpression in a human-derived cell line (HEK293) in order to more closely resemble the properties of human enzymes. The activity data of the sequence variants may shed light on the therapy of urea cycle defects or certain organic acidemias. No patients with a deficiency in amino acid *N*-acyltransferases are known so far. They would deserve special consideration, when suffering from a CASTOR-related disease or urea cycle defects.

The literature data on the intracellular localization were not consistent for GLYAT and GLYATL1. Thus, the third aim was to specify that by confocal laser scanning microscopy with HEK293 cells.

The function of the human GLYATL2 goes beyond detoxification and no data on the expression of *GLYATL2* in human organs were available. Therefore, that should be clarified by gene expression studies formulated in a fourth aim.

Overall, the present study should contribute to a holistic perspective on the amino acid *N*-acyltransferases leading to a more precise comprehension of their role in the human body. Moreover, a contribution to individualized therapies of certain urea cycle defects or isovaleric acidemia may be delivered. Sequence variation analyses and genotyping in amino acid conjugation pathway in patients with metabolic diseases might optimize and individualize therapies. Hence, the variabilities in metabolic pattern and clinical courses can supply valuable input for diagnosis.

2. Material

2.1. Technical equipment

The devices and laboratory equipment, which were used in the present study are listed in Table 1.

Table 1: Used devices and manufacturers.

Device	Manufacturer
Autoclave Varioklav® Zyclondampf	HP Labortechnik GmbH, Oberschleißheim, Germany
Centrifuge 5804 R	Eppendorf, Hamburg, Germany
Confocal Microscope Type: A1	Nikon, Minato, Japan
Cooling thermal Shaker Touch	VWR, Radnor, PA, USA
Flake Ice Maker Icematic D100	Icematic Deutschland , Meerbusch, Germany
FPLC system	Pharmacia, Uppsala, Sweden
Gas chromatograph 7890 A	Agilent, Waldbronn, Germany
5975C mass-sensitive detector	Agilent, Waldbronn, Germany
IKA® Shaker KS 4000 I control	IKA, Staufen, Germany
MH15 Rotilabo Magnetic Stirrer with Heater	Carl Roth, Karlsruhe, Germany
Milli-Q deionized water system	Merck Millipore, Darmstadt, Germany
Mini PROTEAN® Tetra Cell Electrophoresis cell	Bio-Rad, Hercules, USA
Microwave, Daewoo Electronics, KOR-63°5	Daewoo Electronics, Gangnam, South Korea
Mikro 20 Centrifuge	Hettich Zentrifugen, Tuttlingen, Germany
Mr. Frosty Freezing Container	Thermo Fisher, Waltham, MA, USA
My ECL Imager TM	Thermo Fisher, Waltham, MA, USA
Molecular Imager® ChemiDoc™ XRS	Bio-Rad, Hercules, CA, USA
NanoVUE	GE Healthcare, Chicago, IL, USA
NanoVue Plus™ spectrophotometer	Biochrom US, Holliston, MA,USA
Neubauer cell counting chamber	VWR, Radnor, PA,USA
Plate reader EON	BioTek, Bad Friedrichshall, Germany
pH-Fix 4.5- 10.0, 92120	Macherey-Nagel, Düren, Germany
qTower3	Analytik Jena, Jena, Germany
Power supply nanoPAC-300	Carl Roth, Karlsruhe, Germany
Roller mixer SRT2	Stuart Equipments, Staffordshire, UK
Sterilisator	Binder GmbH, Tuttlingen, Germany
Test -tube-rotator 34528	Snijders Labs, AR Tilburg, Netherlands
Trans-Blot Turbo Transfer System	Bio-Rad, Hercules, CA, USA
Ultrasonicator HD 2070	Bandelin electronic GmbH & Co. KG, Berlin, Germany
UV/Visible Spectrophotometer Ultrospec 2100 Pro	Biochrom, Berlin, Germany
Vortex-Genie 2	Scientific Industries, Bohemia, NY, USA
Water bath HAAKE SWB25	Thermo Fisher Scientific, Waltham, MA, USA

2.2. Chemicals, kits and oligonucleotides

Table 2 shows the chemicals used for production of buffers and solutions.

Table 2: Used chemicals for preparation of buffers and solutions.

Chemical	Article/lot number	Manufacturer
0.4 %(v/v) trypan blue solution, sterile filtered	T8154, lot: RNBD9396	Sigma Aldrich, St. Louis, MO, USA
Agar-agar	6494.1	Carl Roth, Karlsruhe, Germany
Agarose standard	3810.2	Carl Roth, Karlsruhe, Germany
Albumin Fraktion V	3854.2	Carl Roth, Karlsruhe, Germany
Ammonium persulfate	9592.3	Carl Roth, Karlsruhe, Germany
Ampicillin	K029.1	Carl Roth, Karlsruhe, Germany
Bacto-tryptone	8952.3	Carl Roth, Karlsruhe, Germany
Benzoyl-coA (used as Li^+ and Na^+ salt)	Lot: LP-04 (Li^+) /DS-09 (Na^+)	CoALA Biosciences, Austin, TX, USA
Blue S'Green qPCR	331416	Biozym, Hessisch Oldendorf, Germany
Bovine serum albumin	0052.2	Carl Roth, Karlsruhe, Germany
Bradford solution	K015.1, lot: 128269654	Carl Roth, Karlsruhe, Germany
Calcium chloride	CN92.2	Carl Roth, Karlsruhe, Germany
Chloramphenicol	3886.2	Carl Roth, Karlsruhe, Germany
Coenzyme A trilithium salt	C3019	Sigma Aldrich, St. Louis, MO, USA
Coomassie Brillant Blue G250	1.15444.0025	Merck Millipore, Darmstadt, Germany
Copper sulfate pentahydrate ($Cu(II)SO_4 * 5\ H_2O$)	C1297	Sigma Aldrich, St. Louis, MO, USA
Dimethyl sulfoxide	1.02952.1000	Merck Millipore, Darmstadt, Germany
Disodium hydrogen phosphate dodecahydrate ($Na_2HPO_4.12H_2O$)	T106.1	Carl Roth, Karlsruhe, Germany
Dithiothreitol	6908.3	Carl Roth, Karlsruhe, Germany
dNTP (10mM)	18427-013	Thermo Fisher, Waltham, MA, USA

DTNB (Ellmans reagent)	6334.2, lot: 515236273	Carl Roth, Karlsruhe, Germany
Ethanol 96 %	P075.4	Carl Roth, Karlsruhe, Germany
Ethidiumbromide	E1510	Sigma Aldrich, St. Louis, MO, USA
Ethylenediaminetetraacetic acid (EDTA.Na_2.$2H_2O$)	VS15701	USB Chemicals, Mundra, India
Fetal Bovine Serum (FBS)	10270-106	Thermo Fisher, Waltham, MA, USA
Folin & Ciocalteu's Phenol Reagent	1.09001.0500	Merck Millipore, Darmstadt, Germany
Formamide	N6749.3	Carl Roth, Karlsruhe, Germany
GeneRuler 100bp DNA Ladder	SM0241	Thermo Fisher, Waltham, MA, USA
GeneRuler 1kb DNA Ladder	SM0311	Thermo Fisher, Waltham, MA, USA
Geneticin disulfate (G-418)	2039.3, lot: 329287284	Carl Roth, Karlsruhe, Germany
Glycerol	7530.4	Carl Roth, Karlsruhe, Germany
Glycine	3187.3	Carl Roth, Karlsruhe, Germany
Hydrogen peroxide (30 %)	CP26.5	Carl Roth, Karlsruhe, Germany
Imidazole	X998.3	Carl Roth, Karlsruhe, Germany
IPTG	2316.3	Carl Roth, Karlsruhe, Germany
Kanamycine sulfate	T832.2	Carl Roth, Karlsruhe, Germany
L-Arabinose	5118.1	Carl Roth, Karlsruhe, Germany
L-cysteine	1693.1	Carl Roth, Karlsruhe, Germany
L-glutamine	1743.2	Carl Roth, Karlsruhe, Germany
Methanol	4627.5	Carl Roth, Karlsruhe, Germany
$MgCl_2$ x $6H_2O$	2189.1	Thermo Fisher, Waltham, MA, USA
Milk Powder	T145.3	Carl Roth, Karlsruhe, Germany
Ni-TED 2000 column material	M&N: 745200.30, lot: 1709/001, DALEX: 1-01-04-003, lot: ST-191119	Macherey & Nagel, Düren, Germany, DALEX Biotech. GmbH, Rheinbach, Germany
Oligo dT primers	58063	Thermo Fisher, Waltham, MA, USA
para-formaldehyde	818715	Sigma Aldrich, St. Louis, MO, USA
Penicillin/Streptomycin	15140-122	Thermo Fisher, Waltham, MA, USA
Phenylacetyl-coA Na^+ salt	lot: BP-02	CoALA Biosciences, Austin, TX, USA
Phosphate buffered saline (PBS)	1108.1	Carl Roth, Karlsruhe, Germany
Potassium dihydrogen phosphate (KH_2PO_4)	3904.1	Carl Roth, Karlsruhe, Germany
Potassium sodium tartrate tetrahydrate	1.08087.0500	Merck Millipore, Darmstadt, Germany

RNAseOUT™	100000840	Thermo Fisher, Waltham, MA, USA
Rotiphorese Gel 30 (37.5:1) (Acrylamide/Bisacrylamide solution)	3029.1	Carl Roth, Karlsruhe, Germany
Silver nitrate ($AgNO_3$)	209139	Merck Millipore, Darmstadt, Germany
Sodium carbonate (Na_2CO_3)	1.06392.5000	Merck Millipore, Darmstadt, Germany
Sodium chloride (NaCl)	3957.1	Carl Roth, Karlsruhe, Germany
Sodium dihydrogen phosphate (NaH_2PO_4)	1.06346.1000	Merck Millipore, Darmstadt, Germany
Sodium dihydrogen phosphate monohydrate ($NaH_2PO_4.H_2O$)	1.06346.1000	Merck Millipore, Darmstadt, Germany
Sodium dodecyl sulfate (SDS)	8.17034.1000	Merck Millipore, Darmstadt, Germany
Sodium hydroxide	6771.3	Carl Roth, Karlsruhe, Germany
Sodium hydroxide (NaOH)	71690	Carl Roth, Karlsruhe, Germany
Sodium thiosulfate pentahydrate ($Na_2S_2O_3$ * 5 H_2O)	1.86516.0500, lot: K26945316001	Merck Millipore, Darmstadt, Germany
Sodium thiosulfate pentahydrate ($Na_2S_2O_3.5H_2O$)	1.86516.0500	Merck Millipore, Darmstadt, Germany
Tetramethylethylenediamine	2367.3	Carl Roth, Karlsruhe, Germany
Tris	5429.2	Carl Roth, Karlsruhe, Germany
Triton X-100	T8787	Sigma Aldrich, St. Louis, MO, USA
Trypsin-EDTA (10X)	15400-054	Thermo Fisher, Waltham, MA, USA
Tween 20	P1379	Sigma Aldrich, St. Louis, MO, USA
Yeast extract	2363.1	Carl Roth, Karlsruhe, Germany
β-mercaptoethanol	8.05740.0250	Merck Millipore, Darmstadt, Germany

Table 3: List of reagent kits used.

Kit	**Article number/Lot number**	**Manufacturer**
4-20 % SDS pre-casted gels	4561096	Bio-Rad, Hercules, CA, USA
GeneJET gel-extraction kit	K0691	Thermo Fisher, Waltham, MA, USA
GeneJET Plasmid Miniprep Kit	K0503	Thermo Fisher, Waltham, MA, USA
innuPREP Plasmid Mini Kit	845-KS-5041010	Analytik Jena, Jena, Germany
Lipofectamine P3000 Transfection reagent	L3000-008	Thermo Fisher, Waltham, MA, USA
NucleoSpin Plasmid, Mini kit for plasmid DANN	740588.10	Macherey & Nagel, Düren, Germany
QIAquick® Gel extraction kit	28706	Qiagen, Hilden, Germany
QIAquick® PCR purification kit	28106	Qiagen, Hilden, Germany
QuickChange Lightning Site-Directed Mutagenesis Kit	210519	Agilent Technologies, Santa Clara, CA, USA
ROTI®Prep Plasmid MINI kit	HP29.1	Carl Roth, Karlsruhe, Germany
SuperSignal™ West Femto Maximum Sensitivity Substrate	34095	Thermo Fisher, Waltham, MA, USA
TransBlot® Turbo™ mini-size PVDF membrane	BR20190814	Bio-Rad, Hercules, CA, USA
Xfect transfection reagent kit	631318, lot: 13074791	Clontech Laboratories Inc., Kusatsu, Japan

Table 4 provides an overview about oligonucleotides applied in this thesis. The pimers for site-directed mutagenesis were designed using the instruction of the manufacturer [193]. Other primers for detection of recombinant genes were designed using name specific reference sequences on NCBI website [194] and the *primer blast* tool [195].

Table 4: Used oligonucleotides for PCR and mutagenesis in this thesis.

Oligonucleotide	**Sequence (5´→ 3´)**
GLYAT_exon3_for	ATGGAACTGTCTTTCACATAAACC
GLYAT_exon4_rev	GCTGTTTCCAGTTGATGAGTTC
GLYAT_cds 3´end_for	ACCAGACTGGAGAGATGAGAA
pEGFP-N_rev	GCTTGCCGTAGGTGGCATC
GLYAT_pET32a_for	CTGGAGAAATCCTTGAGGAAGAG
GLYAT_pET32a_rev	ATTTCTGGGCGTGGGAATAG

CMV_MCS_for	CCCACTGCTTACTGGCTTATC
GLYAT_exon1_rev	GAGGCTCTTCCTCAAGGATTTC
GLYAT_exon4_for	AGATCCCCAAAACTGTCAGGAA
GLYAT_exon5_rev	TTCAGGCTAGGCTGTGAACT
GLYATL1_ex4_for	ATCCCTGAGTCCCTGAAGGT
GLYATL1_ex5_rev	GGAGTCTCTGTTTCCAGTTTACG
GLYATL1_21_for	CCATAAGCTGCTGGCCCTAT
GLYATL1_893_rev	TTCTGTGGGTAGCAAGTCCA
GLYAT c.112G>T_for	GCCTTCAGATTGAATGGATTTCAATGGTTTATGTGAAAGACAGTT
GLYAT c.112G>T_rev	AACTGTCTTTCACATAAACCATTGAAATCCATTCAATCTGAAGGC
GLYAT c.169G>T_for	CTCCTGAGGGCAGACAAACACTGTATTAAAATCAGGC
GLYAT c.169G>T_rev	GCCTGATTTTAATACAGTGTTTGTCTGCCCTCAGGAG
GLYAT c.182A>T_for	GTCATATCCTGCTCCAGAGGGCAGACAACCA
GLYAT c.182A>T_rev	TGGTTGTCTGCCCTCTGGAGCAGGATATGAC
GLYAT c.194T>C_for	ATCAAGGTCATCTGTCGTATCCTGCTCCTGAGGG
GLYAT c.194T>C_rev	CCCTCAGGAGCAGGATACGACAGATGACCTTGAT
GLYAT c.301C>T_for	CTGTGAACTTTGAATCTGTAAATACTGTTTCCAGTTGATGAGTTCTG
GLYAT c.301C>T_rev	CAGAACTCATCAACTGGAAACAGTATTTACAGATTCAAAGTTCACAG
GLYAT c.691G>A_for	CCGGCAAGGTGCCTGTCATTCTCATCTCTCC
GLYAT c.691G>A_rev	GGAGAGATGAGAATGACAGGCACCTTGCCGG
GLYATL1 c.187G>A_for	TCCGGCCTCAAAAGCAGAAGATGACTGATGACATG
GLYATL1 c.187G>A_rev	CATGTCATCAGTCATCTTCTGCTTTTGAGGCCGGA
GLYATL1 c.259G>T_for	TCTCCAAAGAGCCTCAAAAATCAGAATAAGTTTTGAAAAATTGTGAG
GLYATL1 c.259G>T_rev	CTCACAATTTTTCAAAACTTATTCTGATTTTTGAGGCTCTTTGGAGA
GLYATL1 c.373G>C_for	CATTTTCAAAGTCAGTGAAACTAGAGCATTCGAGAGCACT
GLYATL1 c.373G>C_rev	AGTGCTCTCGAATGCTCTAGTTTCACTGACTTTGAAAATG
GLYATL1 c.670G>A_for	GTCTCATGGGTAACCATGAACCCTTCTTGTGAAGTAG
GLYATL1 c.670G>A_rev	CTACTTCACAAGAAGGGTTCATGGTTACCCATGAGAC
GLYATL2 c.43A>T_for	CTCTCAGAAGCTGCAGATTCTGTATTAATCCTTAGAAAAGAG
GLYATL2 c.43A>T_rev	CTCTTTTCTAAGGATTAATACAGAATCTGCAGCTTCTGAGAG
GLYATL2 c.55A>C_for	GCTGCAGATTCTGTATAAATCCTTAGAACAGAGCATCCCTGAAT
GLYATL2 c.55A>C_rev	ATTCAGGGATGCTCTGTTCTAAGGATTTATACAGAATCTGCAGC
GLYATL2 c.251A>T_for	CACCAAAGCTCCTGACATATTAGAGGAAGTCCTGT
GLYATL2 c.251A>T_rev	ACAGGACTTCCTCTAATATGTCAGGAGCTTTGGTG

GLYATL2 c.326G>T_for	CCAAGGTTGCCAAGAGGTCTTGGATGAAGCAATAA
GLYATL2 c.326G>T_rev	TTATTGCTTCATCCAAGACCTCTTGGCAACCTTGG
GLYATL2 c.550A>G_for	AACACTGGGCCTTTGGGGAAAATGAGAGGAGCTTG
GLYATL2 c.550A>G_rev	CAAGCTCCTCTCATTTTCCCCAAAGGCCCAGTGTT
GLYATL2 c.567G>T_for	GGGAAAAATGAGAGGAGCTTTAAATATATTGAACGC TGCCT
GLYATL2 c.567G>T_rev	AGGCAGCGTTCAATATATTTAAAGCTCCTCTCATTT TTCCC
GLYATL2 307_for	ATCCAAGGTTGCCAAGAGGG
GLYATL2 552_rev	TTTCCCAAAGGCCCAGTGTT
Human S18 rRNA forward	CGACGACCCATTCGAACGTCT
Human S18 rRNA reverse	CTCTCCGGAATCGAACCCTGA
RP2_for_2	TAGCTGCCTTTCGTGGATCTC
RP2_rev_2	GACTGAGCAACTCTGCTCTCT
ASPA_exon6_for	CTCAGGTGATCCACCCAACT
ASPA_exon6_rev	GCCAGCACAGGTTACACTGA

The same primers were used for the mutagenesis of *GLYAT* wild-type and the very frequent c.467A>G variant. All listed oligonucleotides were synthesized and obtained from Thermo Fisher (Waltham, MA, USA). Table 5 shows the restriction enzymes used in the present study.

Table 5: Used enzymes for cloning and test restrictions.

Enzyme	**Article number, specific activity**	**Manufacturer**
Cut smart buffer	B7204S	New England Biolabs, Ipswich, MA, USA
EcoRV	ER0305, 10 U/µL	Thermo Fisher, Waltham, MA, USA
HindIII	ER0502, 10 U/µL	Thermo Fisher, Waltham, MA, USA
Klenow fragment	EP0051, lot: 00594554	Thermo Fisher, Waltham, MA, USA
NheI	ER0975, 10 U/µL	Thermo Fisher, Waltham, MA, USA
NotI	ER0595, 10 U/µL	Thermo Fisher, Waltham, MA, USA
rSAP	M6371S, 1 U/mL	New England Biolabs, Ipswich, MA, USA
SacI	ER1135, 10 U/µL	Thermo Fisher, Waltham, MA, USA

2.3. Buffers, solutions and media

Following tables show used buffers (Table 6), solutions, standards (Table 7) and media for the cell cultivation (Table 8).

Table 6: Self-made buffers.

Buffer	**Composition**
10 x MOPS buffer	200 mmol/L MOPS, 50 mmol/L sodium-acetate, 20 mmol/L EDTA, pH 7.0
10 x SDS running buffer	250 mmol/L Tris-base, 1.92 M glycine, pH 8.3, 1 %(w/v) SDS
10 x Towbin transfer buffer	250 mmol/L Tris-base, 1.92 M glycine, pH 8.3
4x Laemmli sample buffer	200 mmol/L Tris-HCl (pH 6.8), 8 % (w/v) SDS, 40 % (v/v) glycerol, 20 % (v/v) β-mercaptoethanol, 0.2 % (w/v) bromphenol blue
50 x TAE buffer	40 mmol/L Tris-base,2 mmol/L EDTA (from 0.5 M EDTA pH 8.0) , 20 mmol/L acetic acid

Coomassie Brillant Blue G250 destaining solution	70 % (v/v) H_2O, 20 % (v/v) Methanol, 10 % (v/v) acetic acid
Coomassie Brillant Blue G250 staining solution	0.1 % (w/v) Coomassie Brillant Blue G250 in 70 % (v/v) H_2O, 20 % (v/v) Methanol, 10 % (v/v) acetic acid
Dialysis buffer	25 mmol/L Tris-HCl, pH 8.0
E. coli Origami 2(DE3) disruption buffer	50 mmol/L NaH_2PO_4, 300 mmol/L NaCl, pH 8.0, 10 % (v/v) glycerol, 1 % (v/v) Triton X-100
Elution buffer	50 mmol/L NaH_2PO_4, 300 mmol/L NaCl, pH 8.0, 250 mmol/L imidazole
GLYAT activity assay buffer	500 mmol/L Tris-acetate, pH 8.0, 2 M glycine
GLYATL1 activity assay buffer	500 mmol/L Tris-acetate, pH 8.0, 0.3 M L-glutamine
HEK293 disruption buffer	50 mmol/L NaH_2PO_4, pH 8.0, 0.1 % (v/v) Triton X-100
LEW buffer	50 mmol/L NaH_2PO_4, 300 mmol/L NaCl, pH 8.0
Phusion polymerase	20 mmol/L Tris/HCl pH 7.9, 100 mmol/L KCl, 1 mmol/L DTT, 0.1 mmol/L EDTA, 0.1 % (v/v) Triton X-100, 30 % (v/v) glycerin
RNA loading buffer	100 µL formamide, 40 µL para-formaldehyde, 20 µL 10x MOPS buffer, 1 µL 10 mg/mL ethidiumbromide
Solution A (Lowry)	2 % (w/v) sodium carbonate in 0.1M sodium hydroxide
Solution B (Lowry)	2 % (w/v) potassium sodium tartrate tetrahydrate in deioinized water
Solution C (Lowry)	1 % (w/v) copper(II) sulfate pentahydrate in deioinized water
Stop solution (cDNA synthesis)	25 mmol/L EDTA, pH 8.0
Taq polymerase	20 mmol/L Tris/HCl pH 7.9, 100 mmol/L KCl, 1 mmol/L DTT, 0.1 mmol/L EDTA, 0.1 % (v/v) Triton X-100, 30 % (v/v) glycerin

Table 7: Used solutions and standards.

Buffer/solution	**Article/lot number**	**Company**
1 kbp DNA ladder	Y014.1	Carl Roth, Karlsruhe, Germany
10 x B1 *Taq* polymerase buffer (-$MgCl_2$, +KCl)	18067017	Thermo Fisher, Waltham, MA, USA
100 bp DNA ladder	SM0241	Thermo Fisher, Waltham, MA, USA
4´6-diamidino-2-phenylindole (DAPI)	D9542, lot: 026M4030V	Sigma Aldrich, St. Louis, MO, USA
5 x FS buffer	P/N Y02321, lot: 1712019	Invitrogen, Carlsbad, CA, USA
5 x HF *Phusion* polymerase buffer	F521L	Thermo Fisher, Waltham, MA, USA
5´-azacytidine	A2385	50 mmol/L stock solution (12 mg powder were dissolved in 1 mL sterile DMSO)
6x loading dye buffer	R0611	Thermo Fisher, Waltham, MA, USA
Ambion™ DNAseI	AM22222, lot: 00587496	Thermo Fisher, Waltham, MA, USA
Ambion™ TRIzol	15596026	Thermo Fisher, Waltham, MA, USA

Blue S´Green qPCR 2x mix	331416, lot: B088617901	Biozym, Hessisch Oldendorf, Germany
Fluoromount-G™ Mounting Medium	00-4958-02	Thermo Fisher, Waltham, MA, USA
Folin-Ciolteu solution	1090010500	Merck Millipore, Darmstadt, Germany
*HotFIRE*pol® DNA polymerase	Lot: HF1470	Medibena Life Science and Diagnostic solutions, Vienna, Austria
Human liver fractions	098H0610.C (Lot nr. Cytosole: 1610027, Lot nr. Homogenate: 1510072, Lot nr. Mitochondria: 1210126)	Tebu-bio, Offenbach, Germany
Human RNA samples (liver, testis, sceletal muscle, kidney, heart, brain)	-	BioChain, Newark, CA, USA
MG-132	474790, lot: 2975583	Sigma Aldrich, St. Louis, MO, USA
Miotracker Orange CMTMRos (200 µM in DMSO)	M-7510, lot: 2642-6	Invitrogen, Life Technologies, Carlsbad, CA, USA
MuLV reverse transcriptase	P/N 100023379, lot: 1912605	Applied Biosystems, Foster City, CA, USA
PageRuler™ Prestained Protein Ladder	26616	Thermo Fisher, Waltham, MA, USA
Phusion polymerase	-	production described in Section 3.5.10.
Poly-D-lysine	A-003-E, lot: 60601-1	Merck Millipore, Darmstadt, Germany
Protease Inhibitor Cocktail	P8340	Sigma Aldrich, St. Louis, MO, USA
Proteinase K	-	Carl Roth, Karlsruhe, Germany
Random Hexamer Primers (50 µM)	N8080127, lot: 1831820	Invitrogen, Carlsbad, CA, USA
RIPA buffer	R0278	Sigma Aldrich, St. Louis, MO, USA
RNAse	EN0531	Thermo Fisher, Waltham, MA, USA
SuperScript™ IV buffer	LT-02241, lot: 00576287	Invitrogen, Carlsbad, CA, USA
SuperScript™ IV reverse transcriptase	LT-02241, lot: 00708125	Invitrogen, Carlsbad, CA, USA
SuperSignal™ West Femto Maximum Sensitivity Substrate	34096	Thermo Fisher, Waltham, MA, USA
T4 DNA Ligase	15224-017, lot: 1867121	Invitrogen, Carlsbad, CA, USA
Taq polymerase	-	production described in Section 3.5.10.

Table 8: Used media, compositions and corresponding experiments.

Medium	Composition/substitution	Approach	Source/ reference
DMEM medium (commercial)	+10 % (v/v) FCS, +1 % (v/v) Penicillin/Streptomycin	Cultivation of HEK293 cells	Thermo Fisher, Waltham, MA, Art.no: 22320-022
GLYAT/L1 overexpression medium for bacteria (self-made)	2 % (w/v) bacto tryptone, 1.25 % (w/v) yeast extract, 0.625 % (w/v) sodium chloride, 0.5 % (w/v) Na_2HPO_4 * $12H_2O$, 0.1 % (w/v) KH_2PO_4, 0.2 % (v/v) glucose	Protein overexpression in *E. coli* Origami 2(DE3)	[148]
LB medium (self-made)	1 % (w/v) bacto tryptone, 0.5 % (w/v) yeast extract, 1 % (w/v) sodium chloride	Cultivation and plasmid increasing of *E. coli* DH10B	[12]
NZY^+ broth medium (self-made)	1 % (w/v) NZ amine (casein hydrolysate), 0.5 % (w/v) yeast extract, 0.5 % NaCl, pH 7.5 After autoclaving: 1.25 %(v/v) 1 M $MgCl_2$ 1.25 %(v/v) 1 M $MgSO_4$ 1 % (v/v) 2 M glucose	Transformation of *E. coli* XL10 gold	Agilent Technologies, Santa Clara, CA, USA
Opti-MEM medium (commercial)	Not substituted	Transfection of HEK293 cells	Thermo Fisher, Waltham, MA, USA, Art.no: 31985-070, Thermo Fisher
SOC medium	2 % (w/v) bacto tryptone, 0.5 % (w/v) yeast extract, 0.06 % (w/v) sodium chloride, 0.02 % (w/v) potassium chloride, 0.22 % (w/v) $MgCl_2$, 0.12 % (w/v) $MgSO_4$, 0.2 % (v/v) glucose	Transformation of *E. coli* Origami 2(DE3)	Cold Spring Harbor Protocols [219]

The self-made media for cultivation of bacterial cells were autoclaved for 20 min at 121 °C and 2 bars and handled under sterile conditions. The media for eukaryotic cells were sterilized by the company accordingly.

2.4. Plastic and glass wares

Table 9: Used plastic and glass wares for handling and keeping solutions and media.

Name	Article number	Manufacturer
Cryoconservation tube	73.380.992	Sarstedt, Nümbrecht, Germany
Eppendorf Tubes 1.5 mL	72.690.001	Sarstedt, Nümbrecht, Germany
Falcon Tubes (15 mL)	62.554.002	Sarstedt, Nümbrecht, Germany
Falcon Tubes (50 mL)	62.547.004	Sarstedt, Nümbrecht, Germany
Pasteur Pipettes	European: 612-1702	VWR, Langenfeld, Germany
Petri Dishes (10 cm diameter)	82.1472.001	Sarstedt, Nümbrecht, Germany
Pipette tips	0-10 µL: 613-0364 20-200 µL: 70.760.002 100-1000 µL: 70.762	VWR, Langenfeld, Germany Sarstedt, Nümbrecht, Germany
Serological pipettes 5 mL	86.1253.001	Sarstedt, Nümbrecht, Germany
Serological pipettes 10 mL	86.1254.001	Sarstedt, Nümbrecht, Germany
TC plate 6-well	83.3920.005	Sarstedt, Nümbrecht, Germany
96-well plates	82.1581	Sarstedt, Nümbrecht, Germany

2.5. Vectors and constructs

Following vectors were used as constructs for cloning (Section 3.4.4.) or for overexpression.

Table 10: Used vectors for cloning and overexpression of target genes.

Vector	Resistance	Reference
pcDNA3.1(+)	Amp, Neo	Genscript, Piscataway Township, NJ, USA
pET32a(+)	Amp	Genscript, Piscataway Township, NJ, USA
pET30Taq	Kan	Prof. Reinscheid (Bonn Rhein-Sieg University of Applied Sciences)
pETPhusion	Kan	Prof. Reinscheid (Bonn Rhein-Sieg University of Applied Sciences)
pGro7	Cam	Dr. Ondrej Kaplan, Charles University Prague, Czech Republic
pUC18	Amp	Agilent Technologies, Santa Clara, CA, USA

The vector maps are shown in the appendix (Suppl. Fig. 1-7). The pcDNA3.1(+) and pET32a(+) vectors were used for overexpression and for re-cloning of *GLYAT* and related genes using single cutters of corresponding MCS. The re-cloning of mutagenized *GLYAT, GLYATL1* and *GLYATL2* genes from pcDNA3.1(+) to pET32a(+) enables the bacterial overexpression of target genes is described in Section 3.4.4.

2.6. Prokaryotic cell systems

Table 11 provides an overview of used prokaryotic cell systems and corresponding genotype and reference information. The genotype information was obtained using database [198]. For genotype information on *E. coli* Origami 2(DE3) the homepage of Merck Millipore was used [196].

Table 11: Prokaryotic cell systems used for cloning and overexpression of target genes.

Cells	Genotype	Source
E. coli BL21 (DE3) pLysS pETTaq	B F$^-$ *ompT gal dcm lon hsdS$_B$*($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S) pLysS [*T7p20 ori*$_{p15A}$](Cm^R)	Prof. Dr. Reinscheid, Bonn Rhein-Sieg University of Applied Sciences, Rheinbach, Germany
E. coli BL21 (DE3) pLysS pETPhusion	B F$^-$ *ompT gal dcm lon hsdS$_B$*($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S) pLysS [*T7p20 ori*$_{p15A}$](Cm^R)	Prof. Dr. Reinscheid, Bonn Rhein-Sieg University of Applied Sciences, Rheinbach, Germany
E. coli DH10B	K-12 F$^-$ Δ(*ara-leu*)*7697*[Δ(*rapA'-cra'*)] Δ(*lac*)*X74*[Δ('*yahH-mhpE*)] duplication(514341-627601)[*nmpC-gltI*] *galK16 galE15* e14$^-$(*icd*^{WT} *mcrA*) φ80d*lacZΔM15 recA1 relA1 endA1* Tn*10.10 nupG rpsL150*(Str^R) *rph*$^+$ *spoT1 Δ(mrr-hsdRMS-mcrBC)* λ$^-$ Missense(*dnaA glmS glyQ lpxK mreC murA*) Nonsense(*chiA gatZ fhuA*? *yigA ygcG*) Frameshift(*flhC mglA fruB*)	Prof. Dr. Reinscheid, Bonn Rhein-Sieg University of Applied Sciences, Rheinbach, Germany
E. coli Origami^{TM} 2(DE3) Singles^{TM} competent cells (article number: 71408, lot: 2735141)	*Δ(ara-leu)7697 ΔlacX74 ΔphoA Pvu*II *phoR araD139 ahpC 36sse galK rpsL* F′[*lac*$^+$ *lacI*^{q} *pro*] (DE3) *gor522*::Tn*10 trxB* (Str^R, Tet^R)	Merck Millipore, Darmstadt, Germany
E. coli XL10 gold	endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Occurred Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 tet^R F'[proAB lacI^{q}ZΔM15 Tn10(Tet^R Amy Cm^R)]	Agilent Technologies, Santa Clara, CA, USA

2.7. Eukaryotic cell systems

Table 12: Eukaryotic cell systems used for overexpression of target genes.

Cell system	Reference
HepG2	Laboratory stock, Bonn Rhein-Sieg University of Applied Sciences, Rheinbach, Germany
HEK293	Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Braunschweig, Germany (originally transformed with Adenovirus type 5 DNA [59])
HeLa [135]	Laboratory stock, Bonn Rhein-Sieg University of Applied Sciences, Rheinbach, Germany

2.8. Antibodies

Table 13: Antibodies used for the immunoblot detection and immunofluorescence of target proteins.

Antibody	Article Number	Lot Number	Manufacturer
Anti-α-tubulin, monoclonal, mouse	ab7291	GR310199-2	Abcam, Cambridge, UK
Anti-GFP, polyclonal, rabbit	GTX113617	40058	Genetex, Irvine, CA USA
Anti-GLYAT, polyclonal, rabbit	Ab86102	GR-44559-3	Abcam, Cambridge, UK
Anti-GLYAT, polyclonal, rabbit	PA5-48504	RL2316912, RL2354376	Thermo Fisher, Waltham, MA, USA
Anti-GLYATL1, polyclonal, rabbit	GTX106956	40002	Genetex, Irvine, CA, USA
Anti-GLYATL2, polyclonal, rabbit	Orb157196	E10846	Biozol, Eching, Germany
Anti-GAPDH, polyclonal, rabbit	GTX110118	40352	Genetex, Irvine, CA, USA
Anti-6x-His-tag, polyclonal, rabbit	GTX115045	40863	Genetex, Irvine, CA, USA
Anti-mouse IgG, ECL™, Horseradish Peroxidase	NA931V	9597364	GE Healthcare, Chicago, IL, USA
Anti-rabbit IgG, ECL™, Horseradish Peroxidase	NA934V	13997044	GE Healthcare, Chicago, IL, USA
Anti-TOMM20 monoclonal, mouse	sc-17764	F1417	Santa Cruz Biotechnology, Inc., Dallas, TX, USA

2.9. Software

The following softwares were used for data analysis and the illustration of results:

Table 14: Software used for *in-silico* data analysis.

Software	Manufacturer	URL
Gen5 software	BioTEK, Winooski, VT, USA	https://www.biotek.com (last accessed: 01.06.2022)
Image J	Java, Oracle Corporation, Redwood City, CA, USA	https://imagej.nih.gov/ij/ (last accessed: 01.06.2022)
Microscoft Office	Microscoft, Redmond, WA, USA	https://products.office.com/de-de/home?rtc=1 (last accessed: 01.06.2022)
qPCR Soft 3.4.	Analytik Jena, Jena, Germany	https://www.analytik-jena.de (last accessed: 01.06.2022)
SnapGene Viewer 5.1	GSL Biotech LLC, San Diego, CA, USA	https://www.snapgene.com/ (last accessed: 01.06.2022)

2.10. Online Tools

The online prediction tools (Table 15) were used for *in silico* analysis and protein predictions in this thesis.

Table 15: Online tools used for *in-silico* data preparation.

Online Tool	Provider	URL
Acetylation (PAIL: Prediction of Acetylation of internal lysines)	Li A, Xue Y, Jin C, Wang M, Yao X. Prediction of Nε-acetylation on internal lysines implemented in Bayesian Discriminant Method Biochem Biophys Res Commun 2006;350:818-24.	http://bdmpail.biocuckoo.org/ (last accessed: 28.04.2022)
Enzyme database Brenda	Technische Universität Braunschweig, Germany	https://www.brenda-enzymes.org/ (last accessed: 28.04.2022)
Glycosylation (NetOGlyc 4.0 sever)	DTU Bioinformatics, Denmark	https://services.healthtech.dtu.dk/service.php?NetOGlyc-4.0 (last accessed: 28.04.2022)
Mutation Taster	Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster 2: mutation prediction for the deep-sequencing age Nat Methods 2014;11:361-2	http://www.mutationtaster.org/ (last accessed: 15.01.2022)
Nucleotide Blast (last accessed, 21.09.2020)	National Center for Biotechnology Information (NCBI), Bethesda, MD, USA	https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_SPEC=GeoBlast&PAGE_TYPE=BlastSearchLink (last accessed: 28.04.2022)
Phosphorylation (NetPhos3.1. server)	DTU Bioinformatics, Denmark	https://services.healthtech.dtu.dk/service.php?NetPhos-3.1 (last accessed: 28.04.2022)
Polyphen II (Polymorphism Phenotyping v2)	Ivan Adzhubei, Harvard Medical School, USA	http://genetics.bwh.harvard.edu/pph2/ (last accessed: 28.04.2022)
Provean (Protein Variation Effect Analyzer)	J. Craig Venter Institute, Rockville, USA	http://provean.jcvi.org/index.php (last accessed: 28.04.2022)
SIFT (Sorting Intolerant From Tolerant)	Bioinformatics Institute, Singapore	https://sift.bii.a-star.edu.sg/ (last accessed: 28.04.2022)
TargetP-2.0	[4]	https://services.healthtech.dtu.dk/service.php?TargetP-2.0 (last accessed: 28.04.2022)
Expasy Peptide Cutter	Swiss Institute of Bioinformatics, Lausanne, Switzerland	https://web.expasy.org/peptide_cutter/ (last accessed: 28.04.2022)

3. Methods

3.1. *In silico* analyses of sequence variants of amino acid *N*-acyltransferases

The sequence variants of human *GLYAT, GLYATL1* and *GLYATL2* were selected from Exac database (now gnomAD browser) [197] with particular focus on high allele frequency and damaging impact on enzyme level determined via online prediction tools (Table 16). Most of the sequence variants show allele frequencies below 1 % regarding highest population minor allele frequency (MAF) – they were named as “sequence variants”, whereas variants with frequencies above 1 % are defined as polymorphisms. However, this is an arbitrary value defined by scientists [136] and the Nature journal [213]. Table 16 shows all selected sequence variants within the present study.

Table 16: Sequence variants of human *GLYAT, GLYATL1* and *GLYATL2* with amino acid exchanges, rs numbers and highest population MAF.

Mutation	amino acid exchange	original triplet	mutated triplet	rs number	Highest population MAF [%]
c.112G>T (GLYAT)	p.(Glu38*)	GGA	TGA	rs746093728	< 0.01
c.169G>T (GLYAT)	p.(Val57Phe)	GTT	TTT	rs765808204	< 0.01
c.182A>T (GLYAT)	p.(Gln61Leu)	CAG	CTG	rs777343795	12
c.194T>C (GLYAT)	p.(Met65Thr)	ATG	ACG	rs145971997	< 0.01
c.301C>T (GLYAT)	p.(His101Tyr)	CAT	TAT	rs748514292	< 0.01
c.691G>A (GLYAT)	p.(Ala231Thr)	GCA	ACA	rs768718220	< 0.01
c.467A>G (GLYAT)	p.(Asn156Ser)	AAT	AGT	rs675815	95
c.595C>T (GLYAT)	p.(Arg199Cys)	CGC	TGC	rs138125182	< 0.01
c.187G>A (GLYATL1)	p.(Glu63Lys)	GAG	AAG	rs866520606	-
c.259G>T (GLYATL1)	p.(Glu118*)	GAA	TAA	rs771289750	< 0.01
c.373G>C (GLYATL1)	p.(Val125Leu)	GTA	CTA	rs374359364	< 0.01
c.670G>A (GLYATL1)	p.(Asp255Asn)	GAC	AAC	rs145756584	< 0.01
c.43A>T (GLYATL2)	p.(Lys15*)	AAA	TAA	rs776582095	< 0.01
c.55A>C (GLYATL2)	p.(Lys19Gln)	AAG	CAG	[164]	-
c.251A>T (GLYATL2)	p.(Lys84Ile)	AAA	ATA	rs76354468	0.06

c.326G>T (GLYATL2)	p.(Gly109Val)	GGC	GTC	rs753500768	< 0.01
c.550A>G (GLYATL2)	p.(Lys184Glu)	AAA	GAA	rs755639436	< 0.01
c.567G>T (GLYATL2)	p.(Leu189Phe)	TTG	TTT	rs61729332	0.03

3.2. Screening for sequence variants of amino acid *N*-acyltransferases

The sequence variants of human *GLYAT* (NM_201648.2), *GLYATL1* (NM_001220496.2) and *GLYATL2* (NM_145016) were chosen from the online database gnomAD [197] The GLYAT "wild-type" in this thesis refers to the canonical human reference sequence. The sequence variants have been selected referring to online predictions and, if available, publication data. For example, the *GLYAT* sequence variant c.182A>T was reported in the literature [153] and characterized as very frequent in a South African Caucasian cohort (12 %). This underlines the need for a functional assessment of this variant. Selection criteria were publications describing the variants, preferably high allele frequencies and negative predictions on enzyme level (Table 17).

Table 17: Chosen sequence variants for human genes *GLYAT, GLYATL1 and GLYATL2* with allele frequencies based on gnomAD database [197] (GRCh38) and predicted influences on enzyme level.

Mutation/HGVS names	Amino acid exchange	rs number	Allele frequency	Exon / cDNA position	Provean	SIFT	Polyphen	Mutation taster
11-58715393-C-A c.112G>T *(GLYAT)*	p.(Gly38Arg) (G38*)	rs746093728	0.0000068	Exon 3 (31 of 108); cDNA: 254	Deleterious (score: -6,43)	Damaging (score: 0.002)	Probably damaging (score: 0.997)	Disease causing
11-58715336-C-A c.169G>T *(GLYAT)*	p.(Val57Phe) (V57F)	rs765808204	0.00002	Exon 3 (90 of 108), cDNA: 311	Deleterious (score: -3.55)	Deleterious (score: 0)	Probably damaging (score: 0.954)	Prediction polymorphism
11-58715323-T-A c.182A>T (GLYAT)	*p.(Gln61Leu) (Q61L)*	*rs777343795*	*12% (South Africa)*	Exon 3 (100 of 108); cDNA:324	Deleterious (score: 0)	Deleterious (score: 0)	Benign (score: 0.247)	-
11-58712882-A-G c.194T>C *(GLYAT)*	p.(Met65Thr) (M65T)	rs145971997	0.0008	Exon 4 (5 of 127); cDNA: 336	Deleterious (score: -5,8)	Deleterious (score: 0)	Probably damaging (score: 0.995)	-
11-58712775-G-A c.301C>T *(GLYAT)*	p.(His101Tyr) (H101Y)	rs748514292	0.000046	Exon 4 (111 of 127); cDNA: 443	Deleterious (score: -4.82)	Tolerated (score: 0.426)	Probably damaging (score: 1)	Disease causing
11-58710611-T-C c.467A>G *(GLYAT)*	p.(Asn156Ser) (N156S)	rs675815 (van der Sluis., 2013)	0.9726	Exon 5 (151 of 172); cDNA: 609	Neutral (score: -0,16)	Tolerated (score: 1)	Benign (score: 0)	-
11-58710062-G-A c.595C>T *(GLYAT)*	p.(Arg199Cys) (R199C)	rs138125182 (van der Sluis., 2013)	0.00036	Exon 6 (106 of 1105); cDNA: 737	Deleterious (score: -6,43)	Damaging (score: 0,023)	Probably damaging (score: 0.984)	-
11:58477439 C / T	*p.(Ala231Thr) (A231T)*	*rs768718220*	*0.0000083*	Exon 6 (203 of 1105); cDNA: 833	Deleterious	Deleterious	Probably damaging	Prediction

c.691G>A (GLYAT) (GRCh37)					(score: -2.85)	(score: 0)	(score: 0.95)	polymorphism
11-58954770-G-A c.187G>A *(GLYATL1)*	p.(Glu63Lys) (E63K)	rs866520606	0.0000066	Exon 5 (1 of 127); cDNA: 330	Deleterious (score: 0)	Deleterious (score: 0,02)	Probably damaging (score: 0.931)	Prediction polymorphism
11:58722315 G / T c.259G>T (GLYATL1) (GRCh37)	*p.(Glu87*) (E87*)*	*rs771289750*	*0.0000083*	Exon 5 (73 of 127); cDNA: 402	NA	NA	NA	-
11:58722708 G / C c.373G>C (GLYATL1) (GRCh37)	*p.(Val125Leu) (V125L)*	*rs374359364*	*0.0000577*	Exon 6 (60 of 178); cDNA: 516	Deleterious (score: -2,65)	Damaging (score: 0,03)	Probably damaging (score: 1)	Prediction polymorphism
11-58955788-G-A c.670G>A *(GLYATL1)*	p.(Asp255Asn) (D255N)	rs145756584	0.000039	Exon 7 (179 of 1434); cDNA: 813	Deleterious (score: -3,73)	Damaging (score: 0,001)	Probably damaging (score: 0.994)	Prediction polymorphism
11-58839570-T-A c.43A>T *(GLYATL2)*	p.(Lys15*) (K15*)	rs776582095	0.0000066	Exon 2 (83 of 118), cDNA: 434	NA	NA	NA	-
11-58839570-T-A *c. 55 A>C (GLYATL2)*	*p.(Lys19Gln) (K19Q)*	*Variant of Waluk, 2012*	*Variants of Waluk, 2012*	Variants of Waluk, 2012	Variant of Waluk, 2012	Variant Waluk, 2012	Variant of Waluk, 2012	Variant of Waluk, 2012
11-58837333-T-A c.251A>T *(GLYATL2)*	p.(Lys84Ile) (K84I)	rs76354468	0.02088	Exon 4 (65 of 127); cDNA: 642	Deleterious (score: -4,99)	Damaging (score: 0)	Possibly damaging (score: 0.954)	Protein features might be affected
11:58604638 C / A c.326G>T (GLYATL2) (GRCh37)	*p.(Gly109Val) (G109V)*	*rs753500768*	*0.0000166*	Exon 5 (13 of 163); cDNA: 717	Deleterious (score: -2.99)	Deleterious (score: 0,02)	Possibly damaging (score: 0.812)	Prediction polymorphism
11-588-34764-T-C c.550A>G *(GLYATL2)*	p.(Lys184Glu) (K184E)	rs755639436	0.0000066	Exon 6(14 of 773); cDNA: 941	Neutral (score: -0,86)	Damaging (score: 0,01)	Possibly damaging (score: 0.579)	Prediction polymorphism (protein features might be affected)
11-58834747-C-A c.567G>T *(GLYATL2)*	p.(Leu189Phe) (L189F)	rs61729332	0.005001	Exon 6 (91 of 773); cDNA: 958	Deleterious (score: -2,61)	Deleterious (score: 0,02)	Probably damaging (score: 1)	Protein features might be affected

For each gene one stop mutation variant was chosen as a negative control for activity tests and immunoblot analyses. Since no polymorphism of *GLYATL1* and *GLYATL2* have been reported in the literature, they were chosen according to their allele frequencies and potential effects predicted in the data bases (Table 17).

3.3. Microbiological Methods

3.3.1. Production of chemical competent *E. coli* cells

To produce chemical competent *E. coli* DH10B and Origami 2(DE3) cells, the $CaCl_2$ method was used to modify the cells, which should consequently be able to channel in foreign plasmid DNA [36]. The Ca^{2+} ions bind phospholipids of the bacterial cell membrane, thus enhancing immobility of the membrane enabling incorporation of foreign DNA.

As a first step, the pre-culture from a saved glycerol stock (Section 3.5.2.) was inoculated 1:100 in 5 mL LB medium without antibiotic. After cultivation at 37 C o/n and 185 rpm, the 100 mL main culture was inoculated on a fixed OD_{600} of 0.05 for *E. coli* DH10B. Due to very slow growth of *E. coli* Origami 2(DE3) this strain was inoculated to a fixed OD_{600} of 0.1. Afterwards the culture was cultivated until an OD_{600} of 0.4-0.6, which represents the exponential growth phase. Subsequently, the cells were sedimented via centrifugation (3,300 g, 10 min, 4 °C). The obtained cell sediments were resuspended in respectively 20 mL pre-cooled 100 mM $CaCl_2$ solution and incubated on ice in a 4 °C chamber for 30 min. The cells were harvested under the same centrifugation conditions. The sediments were gathered in respectively 1.5 mL cooled 100 mM $CaCl_2$ solution supplemented with 15 % (v/v) glycerol. The cell suspensions were portioned to 200 µL aliquots and shock frozen at -154 °C and stored at -80 °C until transformation (Section 3.3.2.).

3.3.2. Transformation of chemical competent *E. coli* cells (DH10B + Origami 2(DE3) strain)

The transformation procedure started with thawing chemically competent bacteria on ice. The specific DNA amount was transformed depending on copy level of the used plasmid. For high copy plasmids 100 ng DNA were transformed at least, while 300-1000 ng were applied for low copy plasmids. The cells were incubated on ice for 15 min after DNA addition, whereas 5 min were sufficient for the Origami 2(DE3) strain. The heat shock at 42 °C was executed 45 s for DH10B and 30 s for Origami 2(DE3). For both strains, the heatshock was performed in a water bath to ensure optimal and rapid temperature transfer. After incubation on ice for 2 min, 800 µL LB medium for DH10B were added to the cells and they were incubated for 60 min at 37 °C on 185 rpm. For the Origami 2(DE3) strain 800 µL SOC medium were added. After this incubation phase, cells were harvested (5,000 g, 3 min, RT) and erased on LB/SOC plates with the specific antibiotic according to the antibiotic resistance cassette on the plasmid.

3.3.2.1. Calculation of transformation efficiency

To calculate the transformation efficiency the counted colonies from the plate were divided through µg transformed DNA resulting in transformants per µg DNA:

$$\text{Transformation efficiency (TE): } \frac{\text{counted colonies}}{\mu\text{g transformed DNA}}$$

3.4. Molecular biological methods

3.4.1. Mutagenesis of *GLYAT, GLYATL1* and *GLYATL2*

The inserted genes of human *GLYAT, GLYATL1* and *GLYATL2* wild-types were modified using the QuikChange Lightnining Site-Directed Mutagenesis Kit for PCR and *DpnI* treatment [193]. For the transformation into *E. coli* XL10 gold cells the QuikChange XL Site-Directed Mutagenesis Kit from the same manufacturer was used. The procedures were performed according to instructions from the manufacturer without the X-gal control. The sequencing reaction replaced the X-gal control as verification for mutagenesis implementation. It was performed via Sanger sequencing of the isolated plasmids [131] at LGC Genomics (Berlin, Germany). The mutagenesis-PCR with specific QuikChange Lightning polymerase was performed (Table 18). The polymerase is derived from *PfuUltra* high-fidelity DNA polymerase for highest fidelity of mutagenic primer-directed replication for both plasmid strands.

Table 18: Reaction scheme of mutagenesis.

Component	**Amount**
Plasmid	50 ng
Primer 1	125 ng
Primer 2	125 ng
10x reaction buffer	5 µL
10 mmol/L dNTP	1 µL
QuickSolution reagent	1.5 µL
deionized water	fill up to 50 µL
Reaction	
T [°C]	**Duration**
9	2 min
9	20 s
60	10 s
68	3.5 min
68	5 min
4	∞

18 x

Following the PCR, the 50 µL of solution with amplified DNA were treated with 2 µL *DpnI* restriction enzyme to separate parental, methylated DNA from newly synthesized mutated DNA. Only the remade, mutated DNA was transformed into *E. coli* XL10 gold ultracompetent cells to ensure the selection of positive transformed clones. Aliquots of 45 µL competent cells were treated with 2 µL β-mercaptoethanol initiating membrane disruption of the cells. Following application of 2 µL *DpnI*-treated PCR sample to the cells and 30 min incubation on ice, the heat shock was performed for 30 s at 42 °C in a water bath. After another incubation on ice for 2 min and adding 500 µL NZY$^+$ broth medium cells were incubated at 37 °C and 185 rpm for 1.5 h. As next step cell suspension was splitted on an LB agar plate with corresponding antibiotic according to 3 square crossing out. The plates were incubated at 37 °C o/n. On the next day, the grown clones were inoculated in 5 mL LB medium at 37 °C o/n and 185 rpm. After o/n inoculation plasmids were isolated with

GeneJET plasmid Miniprep kit (Thermo Fisher). Isolated plasmids were re-transformed into *E. coli* DH10B with subsequent inoculation and another plasmid isolation ensuring higher plasmid amounts and cleaner material compared to isolated material from *E. coli* XL10 gold. For control of the correctly mutated insert, plasmid samples from *E. coli* DH10B preparation were sended to LGC Genomics for Sanger sequencing [131]. Corresponding electropherograms and alignments, which were generated on NCBI online web tool “Nucleotide Blast” [195] were done for each sequence variant separately.

3.4.2. Plasmid Isolation from *E. coli* cells

For the isolation of a high copy plasmid from bacterial cells, a 5 mL pre-culture, produced from 50 µL glycerol stock in 5 mL LB medium (stored at -80 °C), was inoculated at 37 °C o/n and 185 rpm. After harvesting the cell sediment on the next day (4,000 g, 4°C, 5 min) the cells were resuspended in 250 µL resuspension buffer of plasmid isolation kit K0503 (Thermo Fisher, Waltham, MA, USA). This kit was chosen after comparing approach of several DNA isolation kits (Section 3.14.). In the event of a low copy plasmid, a higher amount of culture volume (up to 50 mL) was used. The plasmid isolation procedure was done according to manufacturer´s constructions with following lab-own specifications: After resuspension of the cells in 250 µL resuspension buffer 5 µL RNAse A were added to the cell suspension to enhance RNA degradation, which can possibly disturb transfections. After adding neutralization buffer to the solution and gently shaking (for effective separation of chromosomal DNA and proteins from soluble plasmid DNA) it was incubated on ice for 10 min. To optimize DNA yield from purification columns, the plasmid containing supernatants were mixed in 1:1 ratio with 96 % ethanol. Afterwards they were added to the column to precipitate DNA before and use ethanol filter efficacy [125]. To elute the DNA in the last step 50 °C pre-warmed deioinized water was used instead of the elution buffer that was provided within the kit.

3.4.3. Determination of DNA and RNA concentrations

A NanoVUE spectrophotometer was used for the determination of DNA and RNA concentrations. After adjustment of the reference against DNA/RNA solvent (deionized water), the concentration and the purity factor OD_{260}/OD_{280} were determined.

3.4.4. Cloning strategy of researched sequence variants

The selected sequence variants (Section 3.2.) were re-cloned from eukaryotic overexpression vector pcDNA3.1(+) into the bacterial overexpression vector pET32a(+). Therefore, the *GLYAT*-containing pcDNA3.1(+) constructs were treated with the restriction enzymes *NheI* and *NotI* to open the vector (Figure 3.1., step 1). After restriction digestion the DNA samples were analyzed using agarose electrophoresis (Section 3.4.6.) and the upcoming *GLYAT* fragment of about 900

bp was visualized on the gel. After cutting the fragment from the gel with a scalpel (Section 3.4.7.), the DNA was eluted using a gel extraction kit [199]. To ligate the inserts into the pET32a(+) vector, the *NheI*-cutting site was converted into a blunt-end site via filling up the nucleotides using the Klenow fragment (Figure 3.1., step 2). Likewise, the pET32a(+) was treated with *NotI* for fitting sticky-ends and *EcoRV* restriction using the blunt-end modification. The ligation of modified pET32a(+) vector and insert resulted in the final construct (Figure 3.1., step 3).

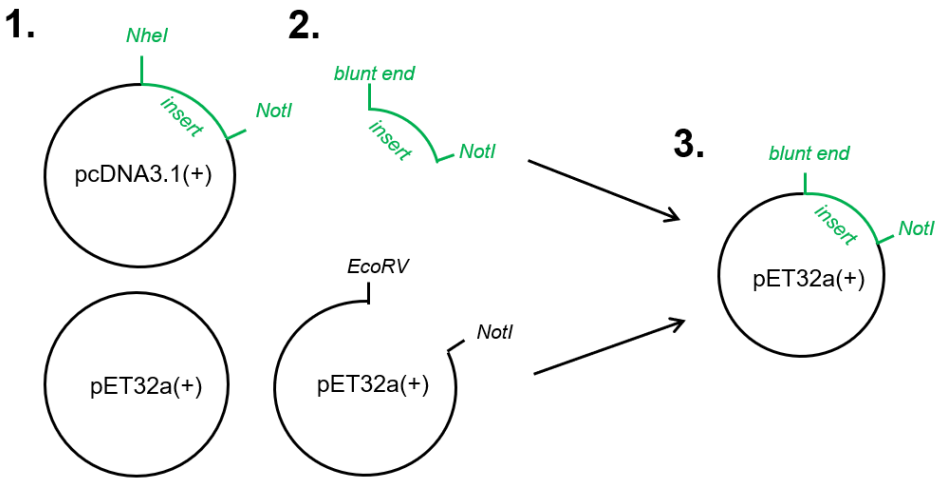


Figure 3.1: Cloning strategy for transfer of target genes (`insert´) from pcDNA3.1(+) vector to pET32a(+) vector for bacterial overexpression. The insert genes (*GLYAT, GLYATL1* and *GLYATL2*) were transferred from pcDNA3.1(+) to pET32a(+) vector after mutagenesis leadin to overexpression in *E. coli*.

3.4.5. Restriction and dephosphorylation of plasmid DNA

The isolated plasmids were treated with restriction enzymes in a 20 µL (for test restrictions) or 50 µL (for preparative restriction samples, described in 3.4.7.) sample (Table 19).

Table 19: Preparative and test restriction samples.

50 µL sample: preparative gel	20 µL sample: test restriction
1 µg DNA	0.5 µg DNA
5 µL 10x Tango buffer	2 µL 10x Tango buffer
1-2 µL restriction enzyme	0.5 µL restriction enzyme
ad. 50 µL deioinized water	ad. 50 µL deioinized water

The samples were incubated for 2 h at 37 °C before they were frozen at -20 °C to inactivate all restriction enzymes. After the freezing step the fragmented DNA was loaded on 1.5 % (w/v) agarose gel for electrophoresis (Section 3.4.6.). The vector was dephosphorylated using 5 µL rSAP in combination with 6 µL of corresponding 10x cut smart buffer to avoid religation. After incubation for 30 min at 37 °C, phosphatase was inhibited at 65 °C for 5 min. The vector material was cleaned from the restriction enzymes by preparation with the QIAquick PCR purification kit [220] (Section 2.2.) to prepare for the ligation step (Section 3.4.8.).

3.4.6. Agarose Gel electrophoresis

The agarose gel electrophoresis procedure was performed for the verification of DNA size and purity. Depending on the expected fragment size of the DNA the percentage of the agarose in the gel was adjusted between 1 % and 1.5 % (w/v). 1 % was used for separation of bigger fragments up to 10 kbp, while 1.5 % was chosen for the separation of smaller fragments below 1 kbp in. 3 µL 10 mg/mL ethidium bromide was added to each gel (50 mL) to stain DNA. After electrophoresis for 30 min at 120 V the DNA was visualized with a Molecular imager® ChemiDocTM XRS+ system (Bio-Rad, Hercules, CA, USA) using the DNA intercalator ethidium bromide. The excitation sources were trans-UV and epi-white lights (302 nm standard, 365 optional). The ethidium bromide enables visualization of DNA by absorbance maxima at 300 and 360 nm and the emission maximum at 590 nm.

3.4.7. Gel-extraction of DNA

The extraction of DNA bands from a preparative restriction sample (Section 3.4.5.) was performed via a scalpel on a preparative DNA cutting table. The DNA sample was divided into a big (90 %) and a small proportion (10 %). To avoid UV radiation a small aliquot of DNA was loaded separately on the gel and marked with the scalpel. This served as orientation for cutting the signal resulted from the bigger aliquot. The extraction from the gel was performed after weighing the gel piece with a Gel extraction kit (QIAquick Gel Extraction kit) according to the instruction manual of the manufacturer [199].

3.4.8. Ligation of DNA

The pET32a(+) vector treated with *NotI* and *EcoRV* and prepared mutated inserts were combined in a ligation sample (Table 20). The restricted plasmids and compatible inserts were ligated with T4-DNA-Ligase.

Table 20: Ligation sample constitution.

NotI/EcoRV treated pET32a(+)	0.5 µg
NotI/EcoRV treated DNA insert	1.5 µg
20 mmol/L ATP	2 µL
5x T4 DNA ligase buffer	4 µL
T4 DNA ligase	1 µL
ad. deioinized water	to 20 µL

The ligation samples were incubated at 4 °C for 24 h. 10 µL of ligation sample were used for the transformation of *E. coli* DH10B (Section 3.3.2.).

3.4.9. PCR

The Polymerase chain reaction (PCR) was utilized for the amplification of DNA fragments (Table 21). It is divided into a denaturation, annealing and extension phase. Section 3.4.1. describes the

procedure for modified DNA fragments obtained by mutagenesis while Section 3.6.2 demonstrates the gDNA verification to confirm transfection in HEK293 cells.

Table 21: gDNA verification PCR

Component	**Volumes**
deionized water	18 µL
10x taq buffer	2.5 µL
10 mmol/L dNTPs	0.5 µL
25 mmol/L $MgCl_2$	1.5 µL
DMSO	0.75 µL
10 µmol/L primer	0.5 µL
Taq polymerase	0.25 µL
cDNA sample	0.5 µL
Reaction	
T [°C]	**Duration**
95	3 min
62	30 s
62	30 s } 35x
72	30 s
72	7 min
4	∞

For further usage of PCR-amplified DNA fragments or digested vector material (Section 3.4.5.) the QIAquick PCR Purification kit [220] was used. tThe agarose gel electrophoresis (Section 3.4.6.) was performed to check the PCR amplicons.

3.5. Biochemical methods

3.5.1. Overexpression of recombinant GLYAT and GLYATL1 in *E. coli* Origami 2(DE3)

The available literature [152, 154] was used as an experimental template for the overexpression of the human GLYAT and GLYATL1. The pre-culture containing the corresponding plasmid was incubated o/n (37 °C, 185 rpm) in the overexpression medium. The main culture was then cultivated further (37 °C, 185 rpm) to an OD_{600} of 0.4 - 0.6. When that was achieved, the temperature was shifted to 28 C for 24 h (185 rpm). The cells were then harvested (4,000 g, 4°C, 15 min). If protein purification was not performed immediately afterwards, the cell sediments were frozen at -80 °C. A maximum of six months of storage was not exceeded. The pET32a(+) empty vector was used as negative control for the overexpression.

3.5.2. Cultivation and storage of bacteria

The *E. coli* Origami 2(DE3) cells were cultivated under standard conditions at 37 °C and 185 rpm. They were incubated at 37 C o/n on a plate and stored for maximum two months in 4 °C for saving the cells as colonies or after ligation (Section 3.4.8.). The o/n cultivated pre-cultures were mixed 1:1 with 50 % (v/v) glycerol (0.45 µm sterile filtered) and subsequently frozen at -80 °C to produce glycerol stocks for longer storage of the cells. The glycerol stocks were used maximum 3 times before they were renewed.

3.5.3. Bacterial cell disruption

The frozen sediments were resuspended in 1/10 pellet volume of bacterial disruption buffer (Section 2.3.) for the initiation of cell disruption. After incubation at RT for 10 min the cells were destroyed via sonication (Bandelin Sonoplus, 4/6 x30 s, 9 pulses, 10 s latency, 15 % amplitude), which depends on the culture size. The lysates were centrifuged (14,000 g, 4 C, 25 min) to get rid of the unsoluble cell debris. The obtained supernatant was immediately used for protein purification (Section 3.5.4.).

3.5.4. Metal-Affinity chromatography of recombinant GLYAT(L1) (FPLC and spin-columns)

For the selective elution of His-tagged target protein using FPLC system the Ni-TED column was filled with 3 g Ni-TED column material (first: Macherey & Nagel, later: DALEX Biotech. GmbH). Afterwards, it was rinsed with 30 mL deionized water and with 60 mL LEW buffer. After application of the target protein containing supernatant (treated with 0.45 µm filter caps) on the column with 1 mL/min, the column was washed again with 2 mL/min for 30 min. The target protein was eluted with 250 mmol/L imidazole in the elution buffer. The elution fractions (á 1 mL) were collected and the OD_{280} of each fraction was measured manually using a standard UV/Vis photometer. The photometer was set to zero at 280 nm using a cuvette with elution buffer.

To purify the target proteins (wild-type + sequence variants) in parallel series, 0.5 g Ni-TED material (DALEX Biotech. GmbH) were applied to the spin-columns, which were then treated with 100 mmol/L $NiSO_4$ for 15 min in end-over-end mixing at 4 °C. After washing with deionized water and subsequent centrifugation (1,000 g, 1 min) LEW buffer was applied for 60 min on the column in end-over-end mixing at 4 °C. The filtered lysates (0.45 µm) were applied on the column in 500 µL portions with 10 min incubation in end-over-end mixing. Washing was performed with 500 µL LEW buffer twice, while washing buffer was removed via centrifugation for the same conditions. Elution was performed with 500 µL elution buffer consisting of 250 mmol/L imidazole dissolved in LEW buffer and collected fractions were then pooled and saved for the dialysis (Section 3.5.5.).

3.5.5. Dialysis of FPLC isolated protein

The protein containing elution fractions (FPLC purified) with absorbances >0.1 were pooled into one tube. A dialysis tube (12,000 – 14,000 MWCO (molecular weight cut-off)) was prepared via boiling in deioinized water. The pooled protein fraction was transferred into the dialysis tube and desalted at 4 °C against 25 mmol/L Tris-HCl, pH 8.0 during stirring. After 1 h the dialysis buffer was renewed, and the dialysis was performed at 4 °C o/n during magnetic stirring.

3.5.6. Pre-concentration of FPLC isolated protein

The desalted protein fraction (FPLC purified) was pre-concentrated after dialysis with a 10 kDa MWCO spin concentrator if concentration was lower than 0.1 mg/mL. The concentrator was

prepared with 25 mmol/L Tris-HCl, pH 8.0 and centrifugation at 13,000 g for 5 min at 4 °C before the protein sample was applied and the centrifugation steps were performed for 10 min each. The next amount was added to the spin tube after short resuspension of the protein sample. The volume of the sample was reduced around 80 %. Only FPLC-prepared samples were pre-concentrated, whereas spin-column samples did not undergo pre-concentration.

3.5.7. Overexpression of recombinant *GLYAT* in HEK293 cells

The genes of interest, *GLYAT, GLYATL1* and *GLYATL2,* and the corresponding sequence variants (Section 3.1.), were cloned into the pcDNA3.1(+) vector and transfected into HEK293 cells for analysis and biochemical characterization of overexpressed enzymes. Therefore, the cells were transfected with *Lipofectamine 3000* reagent (Section 3.5.8.1.) or *Xfect* reagent (Section 3.5.8.2.).

3.5.8. Cultivation and transfection of HEK293 cells

HEK293 cells were cultivated by default in low glucose DMEM medium supplemented with 10 % (v/v) FCS and 1 % (v/v) Penicillin Streptomycin antibiotic at temperature of 37 °C and 5 % CO_2 gas supply for stabilization of medium pH. For cell splitting the medium was removed and cells were washed and resuspended in PBS, before they were seeded in 1:10 or 1:20 ratio on fresh cultivation dishes. The harvest of cell sediments for following enzyme activity assay was performed via washing the cells twice in cold PBS and freezing immediately at -80 °C. For cryo-conservation of the cells of a Petri dish (10 cm diameter), they were removed using a cell scraper, resuspended in PBS, centrifuged (560 rcf, RT, 5 min) and then resuspended in 1 mL freezing medium (70 % DMEM medium, 20 % FCS, 10 % DMSO). The empty vector pcDNA3.1(+) was used as negative control for overexpression.

3.5.8.1. Transfection of HEK293 cells using *Lipofectamine 3000*

The cells were seeded in 1:20 ratio in 6 well plates for ensuring 70-90 % confluency for transfection. On the next day the culture medium was changed before 2 solutions were prepared for transfection simultaneously: solution A (250 µL opti-MEM medium, 5 µL *Lipofectamine 3000* detergent) and solution B (250 µL opti-MEM medium, 10 µL P3000 helping plasmid, 5 µg plasmid DNA). Solution B was added to solution A and the mixture was incubated for 15 min at RT. 250 µL of this transfection solution (mixture) was added to the cells dropwise. After gently swirling, the cells were incubated for 48 h at 37 °C and 5 % CO_2 for transient transfection. For stable transfection a further incubation for 3-4 weeks was performed with mandatory adding of 500 µg/mL G-418 to the cells for selection of positive clones (*Lipofectamine 3000* user guide, [216]).

3.5.8.2. Transfection of HEK293 cells using *Xfect* reagent

The cells were seeded 1:40 one day prior to transfection for ensuring 50-70 % confluency on the next day [200]. 2 solutions were prepared for the transfection: solution A containing 5 µg plasmid DNA filled up to 100 µL with *Xfect* transfection buffer and solution B (incl. 1.5 µL *Xfect* polymer) filled up to 100 µL with *Xfect* buffer. After vigorously mixing, combining and RT incubation for 10 min 100 µL of nanoparticle complex solution was added dropwise to the cells. The plates were gently rocked and incubated for 48 h at 37 °C and 5 % CO_2 for transient transfection. An incubation for 3-4 weeks with 500 µg/mL G-418 in the medium stable transfection and selection of the cells was achieved.

3.5.9. Homogenization of HEK 293 cells

HEK293 cells were homogenized for the extraction of proteins from the cell bodies. To reach that, the cells were thawed on ice where the cells were kept for the whole procedure. Afterwards, the protein specific disruption buffer was used for resuspension of the cell sediments. The cell bodies were destroyed via shear forces using sonification (3x, 3s, amplitude 15 %). After disruption the cell debris was removed via centrifugation at 13,000 g for 15 min and 4 °C. The supernatants were used for Western blot (Section 3.10.) and enzyme analyses (Section 3.11.).

3.5.10. Overexpression and purification of *Taq*- and *Phusion* polymerase from *E. coli* BL21 (DE3) pLysS

The *E. coli* strains BL21 (DE3) pLysS pETTaq and BL21 (DE3) pLysS pETPhusion (Prof. Reinscheid, in-house) were used to produce *Taq*- and *Phusion* polymerase (Prof. Reinscheid, in-house) [31]. The colonies from the supplied transformation plates were cultivated in 5 ml LB medium supplemented with 50 µg/mL kanamycine and 15 µg/mL chloramphenicol. The pre-cultures were cultivated o/n (37 °C, 185 rpm). Likewise, the main cultures were started at OD_{600} of 0.05 and cultivated until exponential growth phase. The protein induction was initiated via adding 0.5 mmol/L IPTG to the cell suspension. After 3 h of induction, the cells were harvested by centrifugation (4,300 g, 20 min, 4 °C) and stored at -80 °C.

To start the purification, the *E. coli* cell sediments were thawed using a water bath for 45 min at 75 °C. The cells were disrupted subsequently by sonication for 15 s at 60 % amplitude (6x). The soluble proteins were collected via ultra-centrifugation (32,000 g, 40 min, 4 °C). The FPLC system (Section 3.5.4.) was used for the isolation of the polymerases. The purification pipe was treated for 30 min with LEW buffer before the obtained supernatant was added with 1 mL/min to the column filled with Ni-TED 2000 material. Likewise, the elution pipe was prepared with elution buffer consisting of LEW buffer supplemented with 250 mmol/L imidazole. After the lysate was added to the column the washing was started with 2 mL/min for 30 min using LEW buffer. The elution was performed with 20 mL elution buffer separated in to 1 mL elution fractions. The absorbances of

the protein containing fractions were measured photometrically at 280 nm in quartz cuvettes. The samples with high absorbances were combined and transferred into dialysis tubing (12,000-14,000 MWCO) desalted against 25 mmol/L Tris-HCl, pH 8.0 at 4 °C o/n under stirring. On the next day, the protein sample was transferred into tube and suspended solids were removed via centrifugation (5 min, 1,000 g, 4 °C). 30 % (v/v) glycerol was added to the sample as freezing protection (-80 °C).

3.5.10.1. Quality control of isolated *Taq*- and *Phusion* polymerase via PCR

The polymerase activities were verified by test PCR (Table 22).

Table 22: PCR for *Phusion* and *Taq* polymerase activity testing

PCR sample *(Phusion* polymerase*)*	
Component	**Amount**
gDNA of ASPA lymphocytes	4 µL (1 µg)
10 µmol/L ASPA_exon6_fw	0.5 µL
10 µmol/L ASPA_exon6_rev	0.5 µL
25 mmol/L $MgCl_2$	3 µL
10 mmol/L dNTPs	1 µL
5xHF *Phusion* buffer	10 µL
Phusion polymerase (self-made)	0.25 µL
deioinized water	fill up 50 µL
PCR sample (*Taq* polymerase)	
Component	**Amount**
gDNA of ASPA lymphocytes	4 µL (1 µg)
10 µmol/L ASPA_exon6_fw	0.5 µL
10 µmol/L ASPA_exon6_rev	0.5 µL
25 mmol/L $MgCl_2$	3 µL
10 mmol/L dNTPs	1 µL
10xHF *Phusion* buffer	5 µL
Taq polymerase (self-made)	0.25 µL
deioinized water	fill up to 50 µL

A positive control of 0.25 µL commercially available *HotFire Taq* polymerase was used instead of the self-made polymerase. As negative control a sample without gDNA was produced. The in-house produced polymerases were used un-diluted and as 1:50 and 1:100 dilutions. The PCR was performed afterwards (Table 23).

Table 23: PCR reactions of HotFire Taq, Taq and Phusion polymerases.

PCR reaction (*HotFire Taq* polymerase)		
T [°C]	**Duration**	
95	15 min	
95	30 s	35x
58	45 s	
72	45 s	
72	10 min	
4	∞	
PCR reactions (*Taq* and *Phusion* polymerases)		
T [°C]	**Duration**	

94	1 min	
94	30 s	30x
58	45 s	
72	60 s	
72	2 min	
4	∞	

The amplified PCR fragments were analyzed via agarose gel electrophoresis (Section 3.4.6.) with 1.5 % (w/v) agarose for 45 min at 120 V.

3.6. Verification of *GLYAT* transfection success in HEK293 cells

The transfection success of the HEK293 cells was checked via isolation of genomic DNA (Section 3.6.1.) as well as via isolation of RNA with subsequent RT-qPCR (Section 3.6.3. and 3.6.4.).

3.6.1. Isolation of gDNA from HEK293 cells

The HEK293 cells of a 10 cm diameter cultivation dish (Petri dish) were resuspended in 480 µL DNA lysis buffer supplemented with 20 µL Proteinase K for the isolation of whole genomic DNA. The cell suspensions were vigorously mixed at 56 °C o/n and 1000 rpm. Afterwards, the solutions were centrifuged (13,000 g, RT, 10 min) to sediment the unsoluble cell debris. The supernatant was transferred into a new tube and 1:1 volume of 100 % isopropanol was added. The tubes were inverted 6-8 times carefully. The samples were centrifuged (13,000 g, RT, 30 min) to sediment chromosomal DNA. The supernatant was discarded, and the DNA sediment was washed with 500 µL 75 % (v/v) ethanol. A centrifugation step of 10 min and 13,000 g was performed before the ethanol washing step was repeated. Subsequently, the DNA cell sediments were dried at RT for 30 min before 100 µL TE buffer was added to resuspend the pellet. This was supported again by vigorously mixing at 1,000 g and 56 °C for 30 min. The concentration of gDNA was measured using the NanoVUE device (Section 3.4.3.).

3.6.2. Checking chromosomal integration of recombinant gene via PCR

The PCR was performed to verify integration of gene of interest into chromosomal DNA of the host cells via homologous recombination (Table 24).

Table 24: PCR for verification of chromosomal integration.

Component	Volumes
10x*Taq* buffer	2.5 µL
25 mmol/L $MgCl_2$	1.5 µL
10 mmol/L dNTPs	0.5 µL
DMSO	0.75 µL
10 µmol/L fw/rev primer	0.5 µL
Self-made *Taq* polymerase	0.25 µL
deioinized water	fill up to 25 µL

Reaction		
T [°C]	**Duration**	
95	5 min	
95	30 s	34x
62	30 s	
72	1 min	
72	7 min	
4	∞	

3.6.3. Isolation of RNA from HEK293 cells using TRIzol Chloroform extraction

The cultivated HEK293 cells in 6 well plates were first resuspended in 500 µL TRIzol. 100 µL chloroform were added to the samples to reach phase separation between nucleic acids and proteins. After vigorously mixing the samples were incubated for 15 min at RT. The samples were then centrifuged (13,000 g, 4 °C, 15 min). The phase separation demonstrated a white middle phase of RNA, which was then carefully transferred into a new tube. 1:1 volume of 100 % (v/v) isopropanol was enclosed after short mixing. The samples were frozen at -20 °C for 1 h to precipitate the RNA. After centrifugation for 30 min at 13,000 g and 4 °C the RNA sediment was washed in 200 µL 75 % (v/v) ethanol. Subsequent vigorously mixing and repeating ethanol washing step were performed. The remaining ethanol was removed via drying the sample at RT and RNA was dissolved in deionized water with mixing at 1,000 g and 56 °C for 30 min. The RNA concentration was determined as described in Section 3.4.3.

3.6.4. Verification of RNA purity via agarose gel electrophoresis

An agarose gel electrophoresis was performed for checking the RNA quality and purity. Therefore, 1.5 g agarose was boiled in 135 mL deionized water and 15 mL 10 x MOPS buffer were added afterwards. The electrophoresis buffer was poured into a special RNA analysis gel chamber. This chamber was treated in 3 % (v/v) hydrogen peroxide solution o/n to destroy the contaminating RNAses. The RNA loading buffer was prepared subsequently (Table 6). The samples were prepared as follows: 7 µL isolated RNA solution, 5 µL deionized water and 4 µL RNA loading buffer before shaking in a thermal shaker for 15 min at 65 °C and 1,000 rpm. Thereafter, they were applied to the gel and the electrophoresis was executed in 1 x MOPS buffer to 130 V for 40 min. The results were checked via visualization with a Molecular imager® ChemiDocTM XRS+ system (Section 3.4.6.).

3.6.5. Reverse transcription – *MuLV* transcriptase vs. SuperScript™ IV

If the isolated RNA (Section 3.6.3.) had an adequate amount without impurities, it was used for cDNA synthesis either with *MuLV* reverse transcriptase or SuperScript™ IV.

The *MuLV* reverse transcriptase was used due to compatibility with bacterial and eukaryotic RNAs and the more lucrative price. Conversely, the more expensive SuperScript™ IV enzyme is characterized by several advantages: it is an engineered transcriptase (based on *MulV*

transcriptase) for superior performance even for challenging RNA samples. It has a higher processivity, which results in higher efficiency, shorter reaction times, increased thermostability, higher resistance to inhibitors and increased sensitivity. It is furthermore optimized for strong RNA secondary structures for longer cDNA and higher yields [57, 67]. In a first preparation step, 2 µg RNA were treated with 2 µL of *DNAseI* (corresponds to 2 U) with 2 µL 10 x reaction buffer in a 20 µL whole sample volume (filled up with deioinized water). After incubation at RT for 20 min 1.3 µL STOP solution (25 mmol/L EDTA, pH 8.0) was added to the samples and they were incubated in a PCR thermal cycler for 10 min at 70 °C to inhibit the *DNAseI*. The cDNA synthesis using *MuLV* reverse transcriptase was initiated via adding 1 µL random hexamer primers (50 µmol/L) and 1 µL dNTPs (10 mmol/L) to 10 µL of the *DNAseI* reaction sample in a whole volume of 13 µL filled up with deioinized water. Subsequent incubation for 5 min at 65 °C, rapid cooling on ice and following 10 min at 23 °C were executed for primer annealing in PCR thermal cycler. To the annealing samples 4 µL 5 x FS buffer, 1 µL RNAse out and 2 µL 100 mmol/L DTT were added before the samples were spreaded to a 50:50 ratio for the preparation of negative controls. One half of each sample was treated with 1 µL deioinized water, whereas on the other half 1 µL of *MuLV* reverse transcriptase enzyme was added. This separation should later allow the verification of reverse transcriptase specific amplification of DNA without expecting signals in the water samples. Last incubation phase for cDNA synthesis was executed in the PCR thermal cycler to 42 °C for 60 min followed by 95 °C for 5 min to inhibit *MuLV* enzyme reaction. The produced cDNA was stored at -20°C until further usage.

The SuperScript™ IV enzyme protocol differs from the *MuLV* protocol: After the DNAseI digestion the primer annealing was initiated by adding 1 µL 50 µmol/L oligo d(T) primer, 1 µL 10 mmol/L dNTPs and 1 µL deionized water to 10 µL DNAseI treated RNA sample. Following briefly mixing of the sample, an incubation at 65 °C for 5 min was performed before quick incubation on ice for 1 min. Subsequently, the reverse transcriptase reaction was started by adding 4 µL 5 x SSIV buffer, 1 µL 100 mmol/L DTT, 1 µL RNAse out and at least 1 µL SuperScript™ IV enzyme (200 U/µL) to the sample. The 20 µL sample was then incubated at 55 °C for 10 min followed by 80 °C incubation for 10 min. The produced cDNA was frozen at -20 °C for further usage in semi-quantitative PCR or RT-qPCR (Section 3.6.6.).

3.6.6. Checking mRNA by semi-quantitative PCR and RT-qPCR

To detect recombinant genes incorporated into gDNA via transfection (Section 3.5.8.) a semi-quantitative PCR was performed (Table 25).

Table 25: Semi-quantiative PCR for mRNA checking.

Component	**Volumes**	
cDNA	0.5 µL	
10 µmol/L primer_fw	0.5 µL	
10 µmol/L primer_rev	0.5 µL	
10xB1 Taq polymerase buffer	2.5 µL	
10 mmol/L dNTPs	0.5 µL	
25 mmol/L $MgCl_2$	1.5 µL	
DMSO	0.75 µL	
Taq polymerase (self-made)	0.5 µL	
deionized water	fill up to 25 µL	
Reaction		
T [°C]	**Duration**	
95	3 min	
95	30 s	} 35x
62 (flexible)	30 s	
72	30 s	
72	7 min	
4	∞	

The primers were designed for amplicons of about 200 bp with exon-intron transitions. Because only the endogenous genes have introns, merely recombinant gene without introns was able to be detected. The fragment would be highly above 2,000 bp with introns and therefore too big to be produced in the semi-quantitative PCR. The idea was to differentiate between endogenous and recombinant genes by primers, which bind very near to exon-intron transitions (Figure 3.2.).

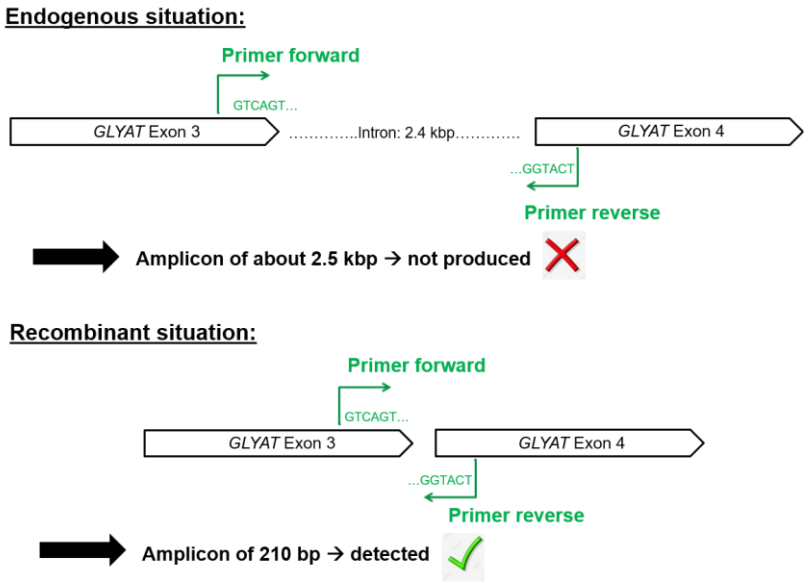


Figure 3.2: Theory of semi-quantitative PCR and RT-qPCR. To distinguish between endogenous genes and recombinant genes primers were designed near to exon-exon transitions. In endogenous case products will not be produced, because the intron avoids the amplicon production. The specific, short fragment can only be produced in the recombinant situation.

The annealing temperature was adapted to the primer specific %GC-content and melting temperature Tm. The elongation period (72 °C, 30 s) was adjusted to the velocity of used polymerase and amplicon sizes. Mainly the standard *Taq* polymerase was used possessing a velocity of 100 nucleotides per second. Hence, 30 s elongation period is sufficient in most cases for 100-500 bp sized amplicon. The RT-qPCR was performed accordingly using Blue S´Green 2x mix solution as fluorescence marker for generated melting products of cDNA. It was used as RT-qPCR mastermix providing dNTPs, polymerase and further required reagents for PCR.

3.7. Verification of GLYAT transformation in *E. coli* Origami 2(DE3) via semi quantitative PCR and colony PCR

Similar to HEK293 transfection validation (Section 3.6.), the transformation of *E. coli* Origami 2(DE3) was proofed by semi-quantitative PCR and colony PCR. The semi-quantitaive PCR was performed with isolated cDNA to check the mRNA of recombinant GLYAT, whereas the colony PCR checked the transformed clones by picking colonies from the agar plates for subsequent usage in PCR.

3.7.1. Semi-quantitative PCR of GLYAT transformation in *E. coli* Origami 2(DE3)

The PCR sample was prepared to 25 µL total volume (Table 26).

Table 26: Semi-quantitative PCR of GLYAT transformation in *E. coli* Origami 2(DE3)

Component	**Volume**	
cDNA	0.5 µL	
10 µmol/L primer_fw (GLYAT_pET32a_fw)	0.5 µL	
10 µmol/L primer_rev (GLYAT_pET32a_rev)	0.5 µL	
10xB1 Taq polymerase buffer	2.5 µL	
10 mmol/L dNTPs	0.5 µL	
25 mmol/L $MgCl_2$	1.5 µL	
DMSO	0.75 µL	
Taq polymerase (self-made)	0.5 µL	
deionized water	fill up to 25 µL	
Reaction		
T [°C]	**Duration**	
95	3 min	
95	30 s	32x
62	30 s	
72	1.5 min	
72	7 min	
4	∞	

The produced amplicons were checked via agarose gel electrophoresis (Section 3.4.6.).

3.7.2. Colony PCR of transformed *E. coli* Origami 2(DE3) clones from agar plates

The colony PCR was performed to distinguish between overexpressed proteins of different clones.

Table 27: Colony PCR of transformed *E. coli* Origami 2(DE3)

Component	**Volumes**
10 µmol/L primer_fw (GLYAT_pET32a_fw)	1 µL

10 µmol/L primer_rev (GLYAT_pET32a_rev)	1 µL	
10xB1 Taq polymerase buffer	2.5 µL	
10 mmol/L dNTPs	0.5 µL	
25 mmol/L $MgCl_2$	1.5 µL	
DMSO	0.75 µL	
Taq polymerase (self-made)	0.25 µL	
deionized water	fill up to 25 µL	
single colony	1x	
Reaction		
T [°C]	**Duration**	
95	5 min	
95	30 s	} 33x
62	30 s	
72	1 min	
72	7 min	
4	∞	

The amplicons were checked via agarose gel electrophoresis (Section 3.4.6.). All *GLYATL1* sequence variants were proofed via colony PCR to verify effective transformation of *GLYATL1* variants into *E. coli* Origami 2(DE3) and to check activity measurements (Section 4.6.6.). The motivation behind that was the wild-type resembling enzyme activity of certain sequence variants (including stop mutation p.(Glu87*)). However, this observation was not detected for the GLYAT series (Section 4.6.5.), and the clones were not checked via colony PCR.

3.8. Protein quantitation

Three protein quantitations were performed depending on the expected protein amount and on the used protein solution: Warburg-Christian approach [168], Lowry [103] and Bradford assay [18]. The Warburg-Christian approach was used for the rapid determination of purified protein samples and for the estimation of protein availability. The Bradford assay was performed to accurately determine lower protein concentrations (0.01 – 1 mg/mL) from purified proteins. The Lowry assay, which is independent from protein structure, was mainly performed for the determination of total protein concentration in cell homogenates.

3.8.1. Warburg-Christian method

The protein concentration was determined quickly with the Warburg-Christian method based on the absorptivity of aromatic amino acids at 280 nm including nucleic acid corrections at 260 nm [168]. For this purpose, the protein solution was diluted 1:30 in 0.1 M potassium phosphate buffer, pH 8.0. The absorbance values A_{260} and A_{280} were determined, and the approximate protein concentration was calculated using the following equation:

$$concentration\ [mg/mL] = (1.55 \times A_{280}) - (0.76 \times A_{260}) \times 30$$

3.8.2. Lowry assay

The quantitation of proteins in cell homogenates was performed via Lowry assay [103]. This procedure combines the known Biuret reaction [121], where Cu^{2+} ions react with the peptide bonds of the protein in alkaline solution. Hence, protein-Cu^{+}-complexes are formed, followed by the oxidation of aromatic protein residues. Resulting blue staining of the solution is determined and represents the protein concentration in a colorimetric assay at 500 nm. However, Lowry et al. suggested the detection at 750 nm (for 5-25 µg/mL protein solutions) and 500 nm for stronger solutions [103]. The total protein content of homogenates is mostly expected higher than 25 µg/mL, hence the detection at 500 nm was performed.

The Lowry standard curve samples (performed with a 0.2 mg/mL BSA stock solution) and the unknown protein samples were prepared accordingly in 4,8 and 12 µL filled up to 100 µL. After adding 500 µL solution ABC (100 parts solution A, 1 part solution B, 1 part solution C) and subsequent vigorously mixing, the samples were incubated for 30 min at RT. 100 µL 1:4.3 diluted Folin-Ciocalteau solution were added to each sample, which was then again vigorously mixed. After another incubation at RT for 30 min absorbance was measured at 500 nm.

3.8.3. Bradford assay

The Bradford assay [18] was used to determine low protein concentrations e.g., of purified proteins. It is based on the reaction of the dye Coomassie Brillant Blue G-250 with the hydrophobic residues of a protein. Absorption maximum is shifted from 475 nm to 595 nm within this reaction. The assay was started with a BSA standard curve from 0 – 150 µg/mL, which is the linear range of the assay. Therefore, 30-50 µL of standard dilutions and of analytes were treated with 600 µL Bradford dilution, which was prepared in deionized water from commercially available 5x solution. After incubation of the samples with added Bradford solution for 5 min at RT, a color change was detected photometrically via measuring the absorbance at 595 nm. The unknown protein concentration was calculated by usage of the standard curve.

3.9. SDS-PAGE and gel stainings

The SDS-PAGE is a qualitative procedure to separate charged proteins according to their size in an electric field [94]. The method was used for the visualization of GLYAT, GLYATL1 and GLYATL2 proteins and as preparation step for immunoblotting (Section 3.10.). The SDS gels were either Bio-Rad gels as 4-20 % (v/v) gradient gels or in-house produced gels (Table 28).

Table 28: Composition of separation (15 % (v/v)) and stacking gel

Separation gel		**Stacking gel**	
Component	**Volume**	**Component**	**Volume**
deionized water	1.1 mL	deionized water	0.68 mL
30 % acrylamide mix	2.5 mL	30 % acrylamide mix	0.17 mL
1.5 M Tris-HCl (pH 8.8)	1.3 mL	1.5 M Tris-HCl (pH 6.8)	0.13 mL
10 % SDS (w/v)	50 µL	10 % SDS (w/v)	0.1 mL
10 % APS (w/v)	50 µL	10 % APS (w/v)	0.1 mL
TEMED	2 µL	TEMED	1 µL

The protein analyte solutions were treated with 4xSDS Laemmli buffer [201] containing bromophenol blue for staining the running front, glycerol for weighing the proteins in an aqueous system and β-mercaptoethanol for the reduction of disulfide bond bridges. The gel and the buffer system (1xSDS running buffer) contain SDS to ensure a negatively charged surface of the analyzed proteins. They were separated at 200 V for 50 min in the Laemmli buffer system. If not otherwise specified, the SDS gels were self-made and 15 % (v/v). In some cases, especially for the purification by FPLC, (Section 4.5.1.) the premanufactured gels (Bio-Rad) were used for SDS-PAGE.

3.9.1. Classic Coomassie staining of SDS gels

The gels were stained in Coomassie Brillant Blue G250 staining solution at RT o/n on a shaker and destained for 3 *30 min in Coomassie destaining solution. They were kept at RT in deionized water before scanning. The detection range of the Coomassie staining is ≈ 100 ng per spot [171].

3.9.2. Silver staining of SDS gels

To detect the lower concentrated proteins with silver staining as a more sensitive method – the detection limit is below 1 ng protein per spot [171] – the gels were first sensitized by incubating in 0.005 % (w/v) $Na_2S_2O_3$ for 35 min followed by incubation in 0.1 % (w/v) $AgNO_3$ in the dark. The gels were rinsed in deionized water twice for 5 min and developed by adding 0.036 % (v/v) para-formaldehyde in 2 % (w/v) Na_2CO_3. The development was stopped after 2 min by treating the gel with 50 mmol/L EDTA (35 min), before washing in deionized water for 5 min and scanning.

3.10. Western blot

The Trans-Blot Turbo Transfer System was used for the transfer of proteins from SDS gel (Section 3.9.) on a PVDF membrane. For that purpose, a blotting sandwich was prepared: 7 filter stacks, PVDF membrane, activated in methanol for 20 s and stored in 1 x Towbin buffer, SDS gel and 7 filter stacks. The semi-dry blotting cassette performed the transfer for 15 min in 2.5 A and 25 V. Thereafter, the membrane was blocked for 1 h in 5 % milk in 1x-TBS-T, buffer, where 0.05 % (v/v) Tween 20 was added, before the specific antibody was applied with subsequent incubation at 4 °C o/n under rolling. The Western blot procedures within this thesis were adjusted to 15 min blotting time with 2.5 A and 25 V as standard conditions (Suppl. Fig. 12). Unless otherwise described, the

blots were incubated for 1 h in 5 % skim milk in PBS while shaking at 4 °C. The second antibody was mouse or rabbit anti-HRP, 1:10,000 depending on the species of the first antibody (mouse, rabbit).

3.11. Enzyme activity assays with Ellman´s reagent

The activity of coA releasing enzymes was determined using the Ellman´s assay [46]. The reaction is based on the reaction of Ellman reagent (DTNB (5,5´-dithio-bis-(2-nitrobenzoic-acid)), with free thiol groups in solution forming mixed disulfide and the dianion of 2-nitro-5-thiobenzoicacid (TNB^2), which is a yellow-colored product detectable at 412 nm. The absorption maximum within visible range enables the quantitation of coA and calculation of enzyme activity. This approach was chosen for activity quantitation due to high specificity for thiol groups, the high molar extinction coefficient and rapid reaction time. The free coA reacts with the DTNB to form the yellow-colored product TNB^{2-} (Figure 3.3.).

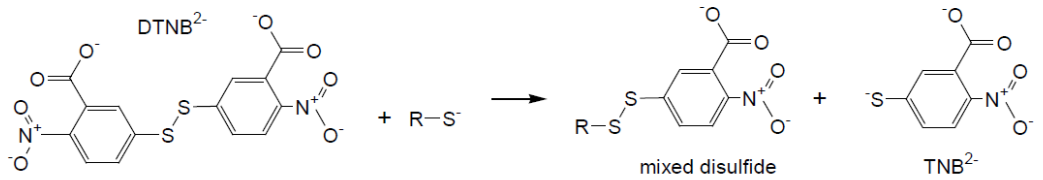


Figure 3.3: Reduction of Ellman´s reagent.

The amino acid *N*-acyltransferases produce SH-containing free coA via conjugation of glycine with an acyl residue from a coA ester. The color change to yellow correlates with the enzyme activity and was exploited for quantitation.

3.11.1. Calibration of Ellman´s assays with L-cysteine standard

To calibrate the activity assay, a standard curve with L-cysteine, an amino acid with a defined amount of thiol groups, was measured in a range from 0 to 0.6 mmol/L [202]. The linear range of the activity referred to the used coA substrate concentration was determined using a L-cysteine standard curve. The reaction was performed by using reaction buffer consisting of 0.1 M sodium phosphate, pH 8.0 containing 1 mM EDTA. The Ellman´s reagent was prepared in 1 mmol/L stock solution followed by 1:10 dilution in the assay leading to 100 µmol/L working concentration. The standard samples A-G were prepared (Table 29).

Table 29: Preparation of standard samples for Ellman´s assay calibration.

Standard	Reaction buffer volume [mL]	L-cysteine amount	Final concentration of L-cysteine [mM]
A	100	26.34 mg	1.5
B	5	25 mL of Standard A	1.25
C	10	20 mL of Standard A	1
D	15	15 mL of Standard A	0.75
E	20	10 mL of Standard A	0.5

F	25	5 mL of Standard A	0.25
G	30	-	0

The test tubes containing 1.25 mL reaction buffer and 25 µL Ellman´s assay stock solution were prepared according to Table 29. 125 µL of each standard solution (A-G) were added respectively. After vigorously mixing and incubation at RT for 15 min, the absorbances at 412 nm were detected. The absorbance values were plotted against L-cysteine concentrations [218]. A L-cysteine standard curve and varying concentrations of DTNB and Tris-acetate were used to calibrate and validate the assay.

3.11.2. Coenzyme A and L-cysteine calibration curves

CoA and L-cysteine were compared for the standardization of the assay since they own defined SH groups. Due to the high price of coA compared to L-cysteine both components were tested here. Therefore, the absorbance values resulting from TNB^{2-} were plotted against concentrations of coA respectively L-cysteine (Figure 3.4.).

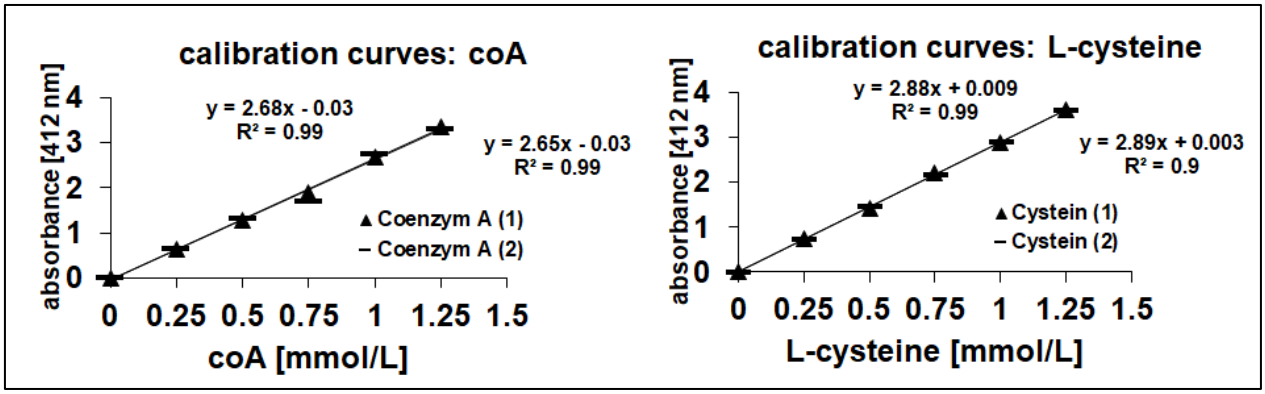


Figure 3.4: Ellman assay calibration curves with coA and L-cysteine. (n = 2, independent samples) For calibration of Ellman assay 0 – 1.5 mmol/l coA respectively L-cysteine were compared regarding their absorbance changes when treated with 0.2 mmol/L DTNB. The procedure was performed according to the instructions of the manufacturer [218].

CoA and L-cysteine resulted in the same absorbance values and curve courses of TNB^{2-} at 412 nm. The following calibration curves were performed only with L-cysteine due to the lower price of the latter. Using 1 mmol/L DTNB, instead of 0.2 mmol/L before, the curves showed no linearity (Figure 3.5.).

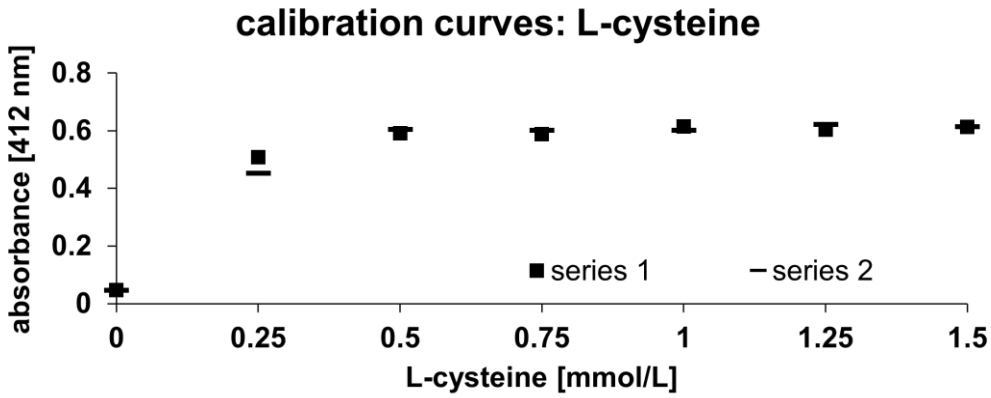


Figure 3.5: Ellman assay calibration curves with increasing concentrations of L-cysteine (0 – 1.5 mmol/L). (n = 2, independent samples) The experiment was performed as shown before (Figure 3.4.), but 1 mmol/L DTNB was used instead of 0.2 mmol/L. Usage of 1 mmol/L DTNB resulted in loss of linearity from 0.5 mmol/L L-cysteine indicated by flattening curve.

A loss of linearity is detected with concentrations higher than 0.6 mmol/L (Figure 3.5.). The L-cysteine calibration curve with 0.1 mmol/L DTNB in each reaction tube was performed to determine the linear range of Ellman´s assay (originally used at 0.67 mM, [46]) (Figure 3.6.) equal to the studies from van der Sluis et al. [152]. To detect the linear range of L-cysteine absorbances, the concentrations from 0-0.6 mmol/L were used. This approach considers and covers the working concentrations of used benzoyl-coA (200 µmol/L) by van der Sluis and colleagues [152].

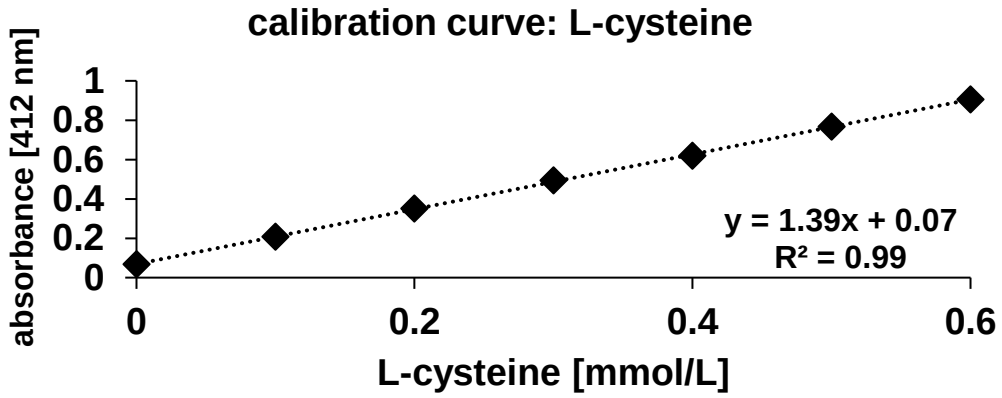


Figure 3.6: L-cysteine calibration curve for standardization and verification of linear range of Ellman´s assay. For detection of linear assay range L-cysteine was prepared in increasing concentrations from 0-0.6 mmol/L in combination with 0.1 mmol/L DTNB. The absorbance was measured at 412 nm.

A linear range of Ellman´s assay from 0 to 0.6 mmol/L L-cysteine is confirmed (Figure 3.6.). The limit for linearity is thus established and used by default for the absorbance detections of free coA

in GLYAT and GLYATL1 activity assay. Nonetheless, a decrease of linearity from 0.7 mmol/L L-cysteine was consequently observed (Suppl. Fig. 24).

3.11.3. Calibration of Ellman´s assay with different concentrations of DTNB

Different DTNB amounts were screened by analyzing 25 and 50 µg of pET32a(+) and GLYAT wild-type homogenates from *E. coli* Origami 2(DE3) overexpression (Figure 3.7.).The aim was to establish and calibrate the enzyme activity assay and to support the L-cysteine calibration. Thus, a useful concentration and quantitation limit of DTNB reagent were defined.

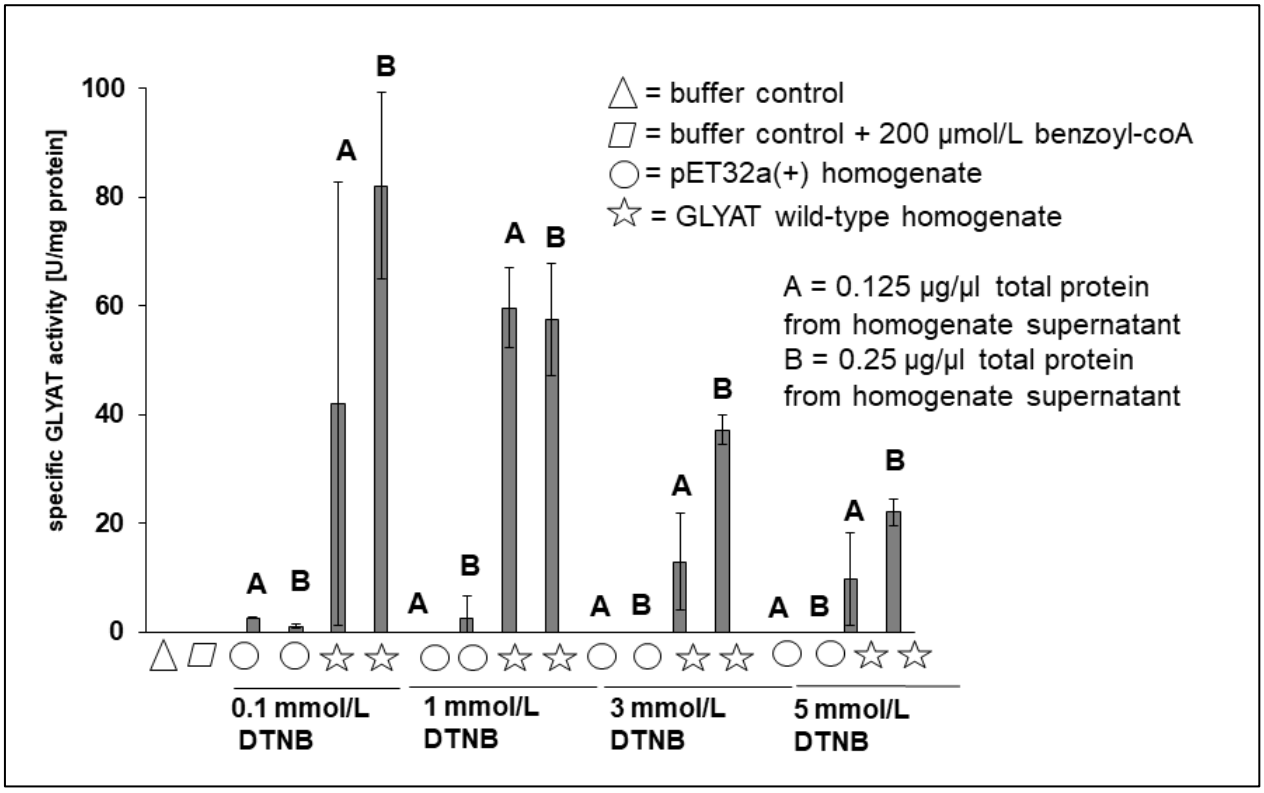


Figure 3.7: Specific GLYAT activities [U/mg protein] of 25 µg and 50 µg of empty vector and GLYAT vector of bacterial overexpression for increasing DTNB concentrations (0.1, 1, 3, 5 mmol/L) using 200 µmol/L benzoyl-coA (n=3, technical replicates). Data from left to right: 1, 2: n.d., 3: 2.6 ± 0.1 U/mg protein, 4: 1.1 ± 0.3 U/mg protein, 5: 42 ± 40.8 U/mg protein, 6: 82.1 ± 17.2 U/mg protein, 7: n.d., 8: 2.6 ± 4.1 U/mg protein, 9: 59.7 ± 7.3 U/mg protein, 10: 57.5 ± 10.4 U/mg protein, 11, 12: n.d., 13: 12.9 ± 8.9 U/mg protein, 14: 37.2 ± 2.7 U/mg protein, 15,16: n.d., 17: 9.8 ± 8.5 U/mg protein, 18: 22.1 ± 2.5 U/mg protein (n = 3, technical replicates). As controls one sample was prepared without substrate and homogenate ("buffer control") and another sample without homogenate substituted with 200 µmol/L benzoyl-coA.

The specific enzyme activity is decreased with increasing DTNB concentrations. The usual concentration of DTNB in the standard GLYAT and GLYATL1 activity assays was 0.1 mmol/L, because the specific activity was the highest (Figure 3.7.). This matches the literature studies [152] where also 0.1 mM DTNB was used in combination to 200 µmol/L of benzoyl-coA.

3.11.4. Standard GLYAT activity assay

The activity of GLYAT enzyme was quantified with the Ellman reagent. This component reacts with thiols, which are released by GLYAT in form of free coA, and forms the yellow-colored product TNB^{2-}, which can be photometrically detected at 412 nm. The samples were prepared according to Table 30.

Table 30: Pipetting scheme for standard GLYAT activity assay

Component	Volumes
deioinized water	139.5 µL
500 mmol/L Tris-acetate, pH 8.0/ 2 M glycine	20 µL
1 mmol/L DTNB	20 µL
Homogenate (protein)	up to 10 µL (1.5 µg/30 µg)

After background measurement at 412 nm and 37 °C for 10 min 200 µmol/L benzoyl-coA was added to check the maximum activity. To determine the kinetic properties 0, 10, 20, 30, 50, 75, 100, 125, 150, 175 and 200 µM benzoyl-coA were added. The glycine concentration was selected based on the literature data [107, 152]. In a more recent publication, the glycine concentration was varied from 1 – 200 mmol/L [129]. However, the choice of 200 mmol/L glycine represents the maximum concentration published in the literature and a multiple of the K_{Mapp} (glycine) of GLYAT [107, 129]. The maximum concentration of benzoyl-coA (200 µmol/L) was chosen due to literature data that confirm the usage of <200 µmol/L as maximum concentration for GLYAT activity assay [107, 129, 152, 154]. Values below 200 µmol/L were determined as K_M in these studies (benzoyl-coA).

3.11.5. Standard GLYATL1 activity assay

The GLYATL1 activity assay was performed according to the same procedure as described previously (Section 3.11.4.) except for the substrate and the used concentrations. Phenylacetyl-coA was used as second substrate next to L-glutamine [85]. Phenyacetyl-coA was used at 100 µmol/L as a multiple of the K_M, which had been described at 35 µM [sic] [169]. L-glutamine was used in a fixed concentration of 150 mmol/L, which was suggested by Webster et al., 1976, Table III.

Table 31: Pipetting scheme for standard GLYATL1 activity assay

Component	Volumes
deioinized water	50 µL
500 mmol/L Tris-acetate, pH 8.0	20 µL
0.3 M L-glutamine	100 µL
1 mmol/L DTNB	20 µL
Homogenate (protein)	up to 10 µL (1.5 µg/30 µg)

3.12. Confocal laser scanning microscopy with HEK293 cells

For localization studies on human GLYAT and GLYATL1, the cells were transfected (Section 3.5.8.1.) with CeGFP_pcDNA3.1(+) and corresponding fusion protein vector (Suppl. Fig. 7). Afterwards they were cultivated according to Section 3.5.8.

First the cells were fixed on objective slides, which were incubated before in 1:20 diluted 1 mg/mL poly-D-lysine for 1 h. The cells were either treated using Mitotracker Orange CMTMRos or TOMM20 antibody to label the mitochondria. The cell suspensions were splitted in 1:10 ratio on cover glasses and cultivated for 2 days under standard cell culture conditions (37 °C, 5 % CO_2) in 6 well plates. Afterwards, the cells were incubated for 15 min with 1 µL of 200 µmol/L Mitotracker solution, which was followed by fixation of the cells. In addition, the cells were washed in warmed PBS 3 times and fixed in 4 % para-formaldehyde for 15 min at 37 °C without CO_2 application. After washing the cells 3 times in warmed PBS, 1:5,000 DAPI solution was applied on the cells for 15 min in darkness at RT to stain nucleic acids. In the following, the cells were embedded in 50 µL Fluoromount-G™ Mounting medium on object carriers and kept in darkness one night for drying. For eGFP-fusion protein expressing cells, the preparation was completed at this point. In case of antibody staining of the aim protein (with GLYAT antibody) or mitochondria (with TOMM20 antibody), the cells were prepared for antibody admission. Therefore, they were incubated for 15 min in 1.5 % (v/v) Triton X-100 in PBS at RT under continuous mixing. After washing the cells 3 times in PBS they were blocked for 30 min in 10 % BSA (w/v) dissolved in PBS at RT under gently mixing. Antibody dilutions (1:100, 1:250, 1:500) were prepared in 5 % BSA filled up to 100 µL with deioinized water. After blocking incubation, the cover slips were put upside-down on parafilm in a petri-dish humid chamber, where drops of diluted antibody were applied before. They were incubated at 4 °C o/n. On the next day, the cell cover slips were washed in warmed PBS 3 times, before secondary antibody anti-goat, 568 nm was applied for 1 h in darkness. The cells were again washed 3 times in PBS, before 1:5,000 DAPI dilution was applied for 15 min in darkness at RT to stain the nuclei. After last washing of the cells in PBS cover slips were prepared upside down on object carriers, which were treated with 50 µL of Fluromount-G™ medium. The cells were kept at RT o/n in darkness for drying. The localization studies of human GLYAT and GLYATL1 in HEK293 cells were performed with the confocal laser scanning microscope (Type A1, Nikon, Minato, Japan). The used filters were: blue (405 nm) to detect DAPI, green (488 nm) to detect eGFP and red (561 nm) to detect mitochondria.

3.13. GC-MS and the verification of GLYAT reaction product

The GC-MS reaction samples were prepared for the verification of benzoylglycine (hippuric acid) production of human GLYAT (Table 32).

Table 32: Composition of sample for GC-MS to verify production of benzoylglycine (hippuric acid)

component	concentration
Tris-acetate, pH 8.0	50 mmol/L
Glycine	200 mmol/L
Trx-His-GLYAT purified	30 µg
Benzoyl-coA	200 µmol/L

The volumes were adapted to 1.2 mL with Tris-acetate buffer. The incubation was performed in a thermal shaker for 5 min at 37 °C followed by 10 min incubation at the same conditions after benzoyl-coA application. One sample was prepared without glycine, one without benzoyl-coA and one without the GLYAT enzyme. This step was performed as a part of the thesis, whereas the GC-MS treatment was performed by cooperative researchers (cooperation with Dr. M. Fernando, Universitätsklinikum Freiburg). The methylation approach followed the description by Lehnert [97]. The internal standard for methylation was 0.04 µmol/L isopropyl malonic acid. The injection to the gas-chromatograph (type: 7890A) and molecule separation with mass detection (5975C mass-sensitive detector) were accomplished for the derivatization. The volumes of the samples were adjusted to 1 mL with deionized water containing 0.04 µmol/L internal standard isopropyl malonic acid. The quantitation of hippuric acid was performed in relation to the signals of internal standard.

3.14. Comparison of plasmid isolation with commercially available kits

To investigate suitable plasmid DNA isolation conditions, several commercially available isolation kits were compared as preparing step in this thesis and for further economical orderings of the working group. For this purpose, the essential kit parameters, like the DNA yield, the price and the working duration were compared. DNA isolation kits and corresponding companies for comparison were as follows: innuPREP DNA mini kit (Analytik Jena, Jena, Germany), NucleoSpin DNA mini kit, (Macherey & Nagel, Düren, Germany) and GeneJET plasmid mini kit (Thermo Fisher, Waltham, MA, USA). Small test kits were inquired at the companies and the instructions suggested by the manufacturers were accomplished for the 2 samples pcDNA3.1(+) and pET32a(+) empty vectors. Apart from this one, an additional step was included into all protocols, named as “Miraprep” method replacing the usual mini or midi preparations of plasmids [125]. The crucial change in the method compared to the usual procedure is the combination of ethanol precipitation with subsequent spin column purification. The reduction of time without increasing costs, but increasing plasmid material, were the main advatanges.

3.15. Semi-quantitative PCR and RT-qPCR for analysis of *GLYATL2* expression in human tissues

The commercially available RNA aliquots of human tissues (liver, testis, skeletal muscle, kidney, heart, brain) were used for the cDNA synthesis using *MuLV* reverse transcriptase (Section 3.6.5.).

The semi quantitative PCR (Table 33) was performed as a preliminary step to check the cDNA and the qualitative verification of human *GLYATL2* mRNA.

Table 33: Sample composition of semi-quantitative PCR and RT-qPCR for *GLYATL2* tissue expression

Semi-quantitative PCR sample	
Component	**Volume**
deioinized water	18 µL
10 x Taq buffer	2.5 µL
10 mmol/L dNTPs	0.5 µL
25 mmol/L $MgCl_2$	1.5 µL
DMSO	0.75 µL
10 µmol/L primer each (*GLYATL2*_307_for, GLYATL2_552_rev)	0.5 µL
Taq polymerase	0.25 µL
1:4 diluted cDNA	0.5 µL
RT-qPCR sample	
Component	**Volume**
Blue S´Green qPCR 2 x mix	10 µL
10 µmol/L Primer (fw/rev)	0.8 µL
deioinized water	6.4 µL
cDNA (1:4)	2 µL

The program for the semi-quantitative PCR equals the one of RT-qPCR (Table 34) without the melting step. The cDNA was used in 1:4 dilution for the subsequent RT-qPCR experiment. As a negative control, 1:1,000 dilution of pcDNA3.1(+) empty plasmid was used. Positive control consisted of 1:1,000 dilution of *GLYATL2*_pcDNA3.1(+). The RT-qPCR program was performed with the qTower3 device (Analytik Jena, Jena, Germany).

Table 34: RT-qPCR reaction of *GLYATL2* tissue expression investigation

RT-qPCR reaction		
T [°C]	**duration**	
95	2 min	
95	5 s	} 40x
62	30 s	
50	15 min	
Melting	15 s	

4. Results

4.1. Pre-works

Pre-work studies were performed to get used to the laboratory and infrastructure and to produce resources for the experiments. Several preparative steps were performed in the beginning to optimize the experimental techniques.

Following this, the*Taq* and *Phusion* polymerase were isolated (Suppl. Fig. 8, 10) and functionally tested (Suppl. Fig. 9, 11) as an initiating step according to the interdisciplinary script (Section 3.5.10.) from Prof. Dr. Reinscheid (Hochschule Bonn-Rhein-Sieg) and a M.Sc. thesis [31]. The aim was to create resources for the experiments of the present study: a standard PCR polymerase (*Taq*) and a special polymerase with proof-reading ability for sequencing of PCR products (*Phusion*).

Furthermore, the plasmid isolation for transformations and transfections needed to be established by comparison of different kits from selected companies (Suppl. Fig. 13-16). The goal was to identify the most cost-effective, easy to handle, time efficient and thorough kit with the highest profit. Hence, several protocols for pcDNA3.1(+) vector preparation were executed following the manufacturer´s instructions with an additiona modified EtOH precipitation leading to more efficient plasmid isolation by pre-concentration [125]. The GeneJET Plasmid Miniprep Kit (Table 3) has proven to be the most efficient and economical.

Beyond that, the most suitable Western blot conditions for an efficient GLYAT transfer were studied and adapted for the GLYAT “wild-type” with the canonical human reference sequence NM_201648.2 (Suppl. Fig. 12). The blotting conditions with 25 V, 2.5 A and 15 min were most appropiate for the efficient protein transfer and therefore selected as standard conditions.

4.1.1. Native GLYAT expression in HepG2, HeLa and fibroblasts

Human GLYAT, as the main analyzed enzyme in the present study, was supposed to have different native expression patterns in the cell lines HEK293, HEK293T, HeLa and HepG2. Those were compared to identify the most feasible overexpression system with low level of disturbing native expression.

The endogenous expression of human GLYAT in mammalian cells was poorly characterized in the literature. Only one study has addressed this topic in a hepatocellular carcinoma cell line [107]. The GLYAT expression was markedly reduced that highlights the important function in xenobiotic metabolism of liver cells with suppression in human hepatocarcinoma cell lines.

Thus, native GLYAT expression had to be elucidated due to influences on recombinant GLYAT overexpression and possible falsification of activity. Thereby, the human cell lines HEK293, HEK293T, HeLa, HepG2 and fibroblasts were investigated as pre-working step to find feasible

human-derived overexpression system. The pcDNA3.1(+) vector served as negative control in each transfection series. However, *E. coli* lacks native GLYAT due to absence of cell organelles such as mitochondria and their related enzymes.

The analysis of native *GLYAT* expression with untransfected eukaryotic cell lines HEK293, HEK293T, HepG2 and HeLa was analyzed (Figure 4.1.). In addition, transfected HEK293 cells with pcDNA3.1(+) empty vector (negative control) and GLYAT-pcDNA3.1(+) (positive control) were analyzed.

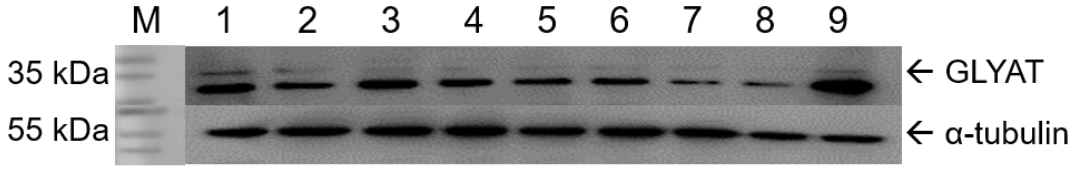


Figure 4.1.: Western blot of 25 µg native HEK293, HEK293T, HepG2 and HeLa total protein supernatants and of transfected HEK293 supernatant with pcDNA3.1(+) and GLYAT-pcDNA3.1(+) (12 % (w/v) SDS gel)). 1: HEK293 (1), 2: HEK293 (2), 3: HEK293T (1), 4: HEK293T (2), 5: HepG2 (1), 6: HepG2 (2), 7: HeLa, 8: HEK293 + pcDNA3.1(+), 9: HEK293 + GLYAT wild-type. First antibodies were anti-GLYAT PA5-48504, 1:1,000 and anti-tubulin, ab7291, 1:5,000.

The expression of endogenous GLYAT, visualized by the upper band, was very decreased in all samples. Due to the percentage of this SDS gel the band separation was hardly present. Therefore, it was increased to 15 % (w/v) in the following sections (if not otherwise specified). The native expression of GLYAT protein seemed to be very low in all analyzed cell lines. To validate that, the specific activity of GLYAT was determined for these supernatants (Figure 4.2).

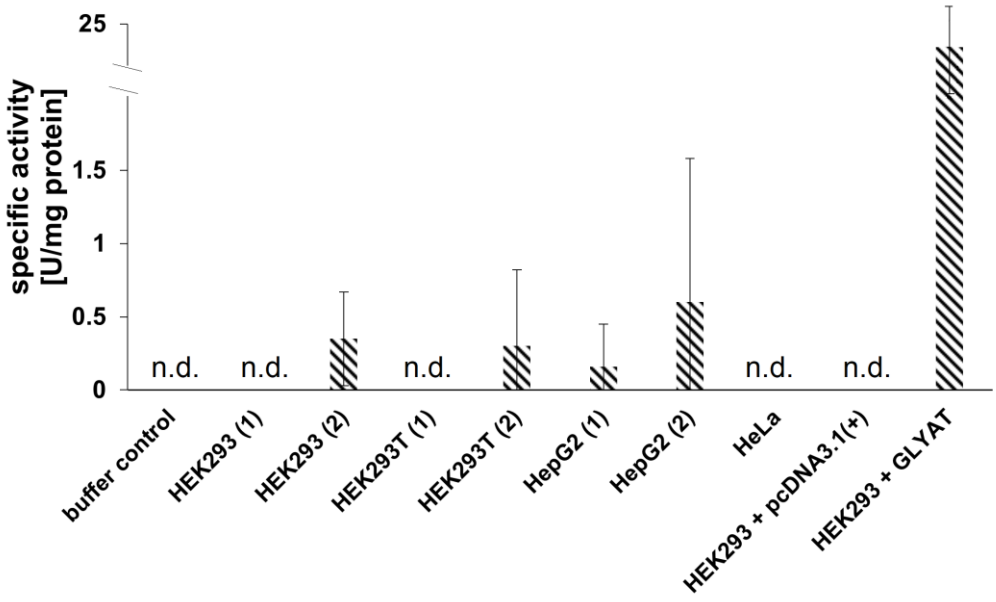


Figure 4.2: Maximum GLYAT activities of cell supernatants using 200 µmol/L benzoyl-coA (n=3, technical replicates). Transfected GLYAT in HEK293 cells showed 23 ± 4 U/mg protein activity and the remaining supernatants have activities below 1 U/mg protein.

Benzoyl-coA was used at 200 µmol/L based on literature data [152] and the enzyme database Brenda [203]. 200 µmol/L is a multiple of the published K_M of the enzyme and a reaction saturation is considered for this benzoyl-coA concentration. The expression of native GLYAT was decreased but still detectable (Figure 4.1.)., which was corroborated as only minor activities below 1 U/mg protein were detected for the endogenous GLYAT (Figure 4.2.). The positive control of transfected HEK293 cells with GLYAT wild-type showed higher expression on the blot (Figure 4.1.). The GLYAT native expression was remarkably lower compared to recombinant GLYAT in transfected HEK293 cells.

In another overexpression series the GLYAT signal intensities of the Western blot were quantified (Figure 4.3.).

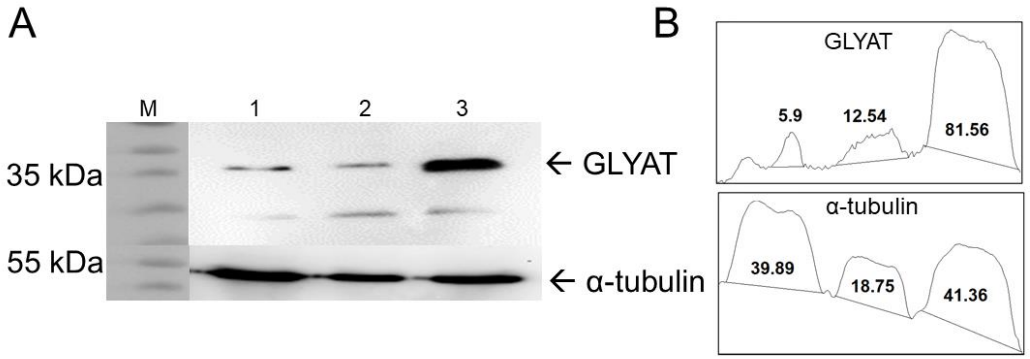


Figure 4.3: A: Western blot of 30 µg total protein supernatant of transient transfected HEK293 cells with recombinant GLYAT and peak intensities (ImageJ) (B). M: PageRuler™ Prestained Protein Ladder, 1: HEK293, native, 2: HEK293 + pcDNA3.1(+), 3: HEK293 + GLYAT wild-type. First antibodies were anti-GLYAT PA5-48504, 1:1,000 and anti-tubulin ab7291, 1:5,000. GLYAT/ α-tubulin signal ratios: sample 1: 0.15, sample 2: 0.67, sample 3: 1.97.

The overexpression of recombinant GLYAT resulted in 13x higher protein yield compared with the native GLYAT (sample 1). This agrees with the previous activity study (Figure 4.2.) where native expression constituted only a small proportion (below 10 %) of recombinant GLYAT. The α-tubulin detection looked very uniform on the blot. The better separation of GLYAT signals here was due to the higher percentage of the acrylamide in this gel (15 %) in comparison to Figure 4.2. However, the GLYAT sample (sample 3) was the predominant signal, which was corroborated by the GLYAT/α-tubulin signal ratios in addition.

4.2. Mutagenesis of *GLYAT*, *GLYATL1* and *GLYATL2* genes and confirmation via Sanger sequencing

The sequence variants of human amino acid *N*-acyltransferases were screened for high allele frequency and enzyme deficiency predictions and published data in the literature (Section 3.2.). All sequence variants (Table 17) were produced successfully, which was confirmed by sequencing (Suppl. Fig. 17-19).

4.3. Subcloning of mutagenic *GLYAT, GLYATL1* and *GLYATL2* inserts from pcDNA3.1(+) to pET32a(+) vector

The modified DNA inserts in pcDNA3.1(+) vector were transferred into the pET32a(+), which contains bacterial overexpression tools to guarantee an efficient overexpression of target proteins. Enabled by inducible *lac*-operon, the fusion proteins are equipped with *N*-terminal *Trx*-tag to enhance the solubility. The pBR322 represents bacterial origin of replication allowing fast propagation of plasmid molecules within bacterial cells. The *His*-tag enables the purification of target protein via Ni^{2+}-affinity chromatography. Followed by the restriction enzyme digests (*NheI*, *NotI*) and agarose gel electrophoresis, the inserts were extracted from the gels (Figure 4.4.).

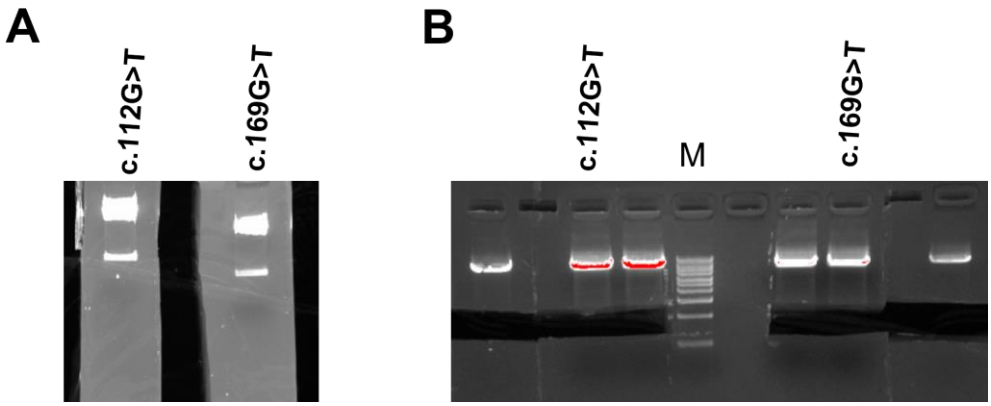


Figure 4.4: 1 % (w/v) agarose gel checking strips (UV treated) (A) and preparative gel (not UV treated) (B) of *GLYAT* mutated inserts in pcDNA3.1(+) vector. M: 1 kbp DNA ladder. The small gel strip (left side) was treated with UV light to detect the band level. The bigger strips (right side, not UV treated) were cutted and the DNA fragment was eluted from the gel using QIAquick Gel Extraction kit [199]. The cutted GLYAT inserts (908 bp) were not treated with UV light to avoid unspecific DNA damage.

The same procedure was conducted for the *GLYAT, GLYATL1* and *GLYATL2* inserts that were extracted from the pcDNA3.1(+) vector (Suppl. Fig. 20-23). The pET32a(+) vector was digested using *NotI* and *EcoRV* (Figure 4.5.) according to the cloning strategy (Section 3.4.4.).

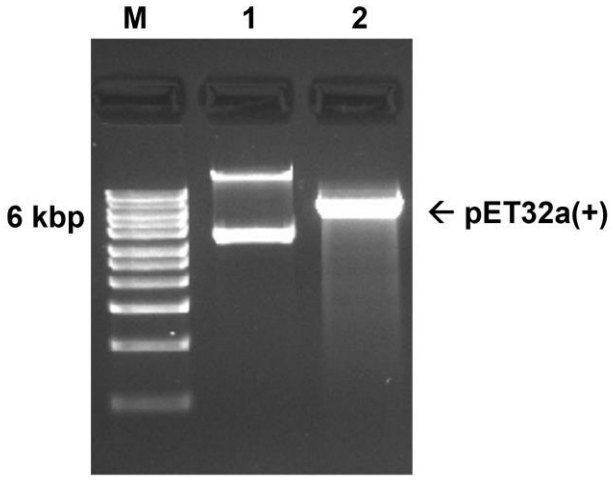


Figure 4.5: 1 % (w/v) agarose gel of 0.5 µg pET32a(+) vector (1) compared with the digested vector (2). M: 1 kbp DNA ladder, 1: 100 ng pET32a(+) untreated, 2: pET32a(+) digested with *NotI/EcoRV.*

The sufficiently digested pET32a(+) (sample 2) was the basis for the ligations with the other inserts (Figure 4.4., Suppl. Fig. 20-23), which was performed after dephosphorylation with rSAP (Section 3.4.5.). For comparison, the untreated sample demonstrated the bigger, open circle plasmid conformation with the more compact and a smaller signal resulted from the supercoiled form (Figure 4.5., sample 1).

4.4. Overexpression of GLYAT and GLYATL1 in *E. coli* Origami 2(DE3)

For the overexpression of GLYAT and GLYATL1 the studies of van der Sluis et al. [152, 154] were used to validate the system and for comparison. The used *E. coli* Origami 2(DE3) strain (Section 1.4.4.), which is transformed with pET32a(+) vector, covers the occurred requirements for the overexpression of the eukaryotic GLYAT in a prokaryotic system [17, 22, 33, 174].

To start the overexpression experiments, a double transformation of *E. coli* Origami 2(DE3) had to be conceptualized at first. After simultaneous transformation of both plasmids, pGro7 and pET32a(+), the cells did not grow. After getting familiar with the known slow growth of this *E. coli* strain [124], the cells were transformed with the plasmids one by one, which consequently resulted in a low number of colonies (3 and 98 transformants/µg DNA in an exemplary series). However, the *E. coli* DH10B strain resulted in intensively higher colony numbers per plate [43]. The *E. coli* Origami 2(DE3) colonies were then used for the protein induction using T7 promotor of pET32a(+) plasmid (Section 3.5.1.).

The pGro7 plasmid was transformed at first, because it was required for all other transformations. Followed by the production of fresh competent cells of those pGro7 transformed cells, the transformation with pET32a(+) plasmid containing human *N*-acyltransferase genes was

conducted. However, the work with *E. coli* Origami 2(DE3) as non-established strain with special requirements turned out to be very protracting and time-consuming.

The “GLYAT wild-type” always refers to the “Trx-His-GLYAT” fusion protein. The same applies for the GLYATL1 wild-type and the sequence variants of both enzymes.

The SDS gel (Figure 4.6. A) shows the bacterial overexpression approach of human GLYAT wild-type. An additional immunoblot analysis (Figure 4.6 B) demonstrates sufficient GLYAT and GLYATL1 wild-type overexpression in *E. coli* Origami 2(DE3) accordingly.

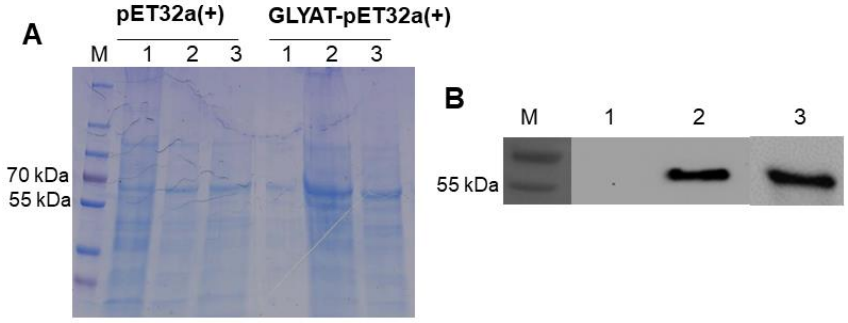


Figure 4.6: Overexpression test of human GLYAT (30 µg total protein supernatant) in *E. coli* Origami 2(DE3) by Coomassie staining (A) and Western blot (B). A: M: PageRuler™ Prestained Protein Ladder, 1: pre-induction fraction, 2: post-induction fraction, 3: supernatant fraction (after cell disruption). B: Overexpression test of human GLYAT (2) and GLYATL1 (3) in *E. coli* Origami 2(DE3) by comparison with pET32a(+) empty vector (1). First antibody: anti-GLYAT, PA5-48504, 1:1,000.

The signal between 55 and 70 kDa was increased for GLYAT wild-type samples, hence GLYAT overexpression was performed appropriately. The overexpression of human GLYAT and GLYATL1 Trx-His fusion proteins was confirmed by the strong, specific protein signal at 56 kDa on the Western blot (Figure 4.6. B). The protein has a size of 56 kDa, because GLYAT/L1 is *N*-terminally fused to *Trx*- and His-*tag* to optimize protein solubility and enabling purification. The size represents the wild-types and sequence variants of GLYAT and GLYATL1 and will be detected on each following gel. GLYAT and GLYATL1 wild-type (Section 4.4., 4.5.) will be used synonymously to Trx-His-GLYAT and Trx-His-GLYATL1 fusion protein and serve as positive controls in the overexpression series. The pET32a(+) empty vector was used as negative control. The verification was observed via SDS PAGE and protein staining (Coomassie) or Western blot as Figure 4.7. shows for an exemplary series.

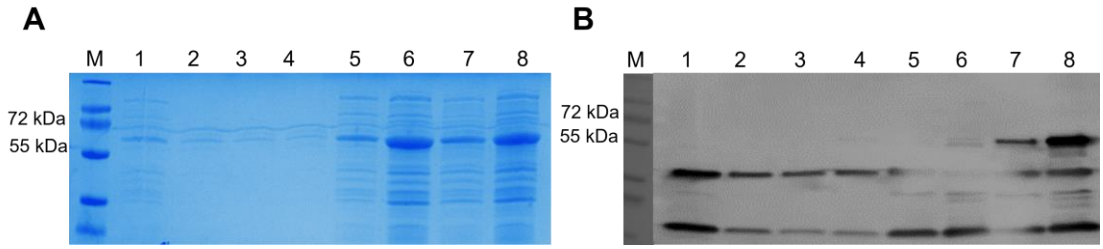


Figure 4.7: Coomassie stained SDS PAGE (A) and Western blot (B) of 15 µg *E. coli* Origami 2(DE3) total protein supernatants overexpressing GLYAT sequence variants. M: PageRuler™ Prestained Protein Ladder, 1: *E. coli* Origami 2(DE3) + pET32a(+) pre induction, 2: *E. coli* Origami 2(DE3) + GLYAT p.(Gly38*) pre induction, 3: *E. coli* Origami 2(DE3) + GLYAT p.(Gln61Leu) pre induction, 4: *E. coli* Origami 2(DE3) + GLYAT p.(Ala231Thr) pre induction, 5: *E. coli* Origami 2(DE3) + pET32a(+) post induction, 6: *E. coli* Origami 2(DE3) + GLYAT p.(Gly38*) post induction, 7: *E. coli* Origami 2(DE3) + GLYAT p.(Gln61Leu) post induction, 8: *E. coli* Origami 2(DE3) + GLYAT p.(Ala231Thr) post induction. First antibody was anti-GLYAT PA5-48504, 1:1,000.

The sequence variants p.(Gln61Leu) and p.(Ala231Thr) were expressed after induction (Figure 4.7., A, samples 7-8). The target protein Trx-His-GLYAT with 56 kDa might be induced sufficiently, but due to simultaneous overexpression of the 60 kDa chaperonin GroEL (encoded by pGro7 plasmid) the Coomassie staining was not conclusive and misses the ability to distinguish between overexpressed Trx-His-GLYAT (56 kDa) and chaperonin GroEL (60 kDa). Only the sequence variants p.(Gln61Leu) and p.(Ala231Thr) were sufficiently overexpressed in the first series (Figure 4.7.) and thus used for FPLC purification (Section 4.5.1.). As expected, for pET32a(+) and p.(Gly38*) no target protein signal appeared. The remaining sequence variants p.(Val57Phe), p.(Met65Thr) and p.(His101Tyr) showed no cell growth or had not been overexpressed (data not shown) and were therefore not isolated via FPLC. The isolation of all sequence variants successively was not performed, because the FPLC system allowed only one protein isolation. The growth of cells seemed to be no evidence for sufficient target protein overexpression since certain GLYAT sequence variants were not detected on immunoblot, although the cells grew. If the cells grew, their homogenates were used for immunoblotting. Although each cultivated *E. coli* strain was transformed with both plasmids (pGro7, pET32a(+) derivative), they did not necessarily express the target proteins, which cause false positive clones. Hence, Western blot controls before each purification were compulsory and further overexpression validations were performed only by Western blot.

The characterization of target proteins is rather comparable when isolation is performed simultaneously. Therefore, an overexpression test was performed with all GLYAT sequence variants in a smaller scale as preparation for spin column purification (Section 4.5.2.).

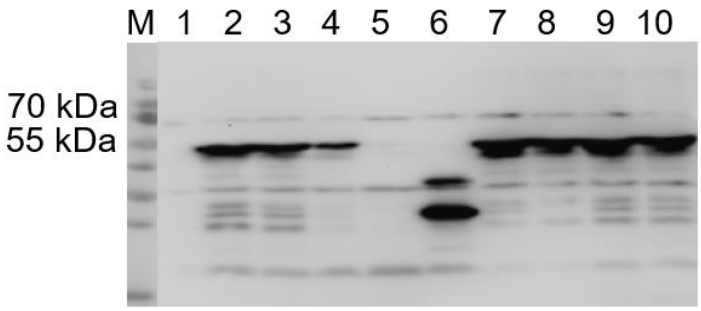


Figure 4.8: Western blot of 20 µg total protein supernatants of *E. coli* Origami 2(DE3) of another overexpression series of GLYAT and sequence variants. M: PageRuler™ Prestained Protein Ladder, 1-10: *E. coli* Origami 2(DE3) homogenates; 1: pET32a(+), 2: GLYAT wild-type, 3: GLYAT p.(Asn156Ser), 4: GLYAT p.(Arg199Cys), 5: GLYAT p.(Gly38*), 6: GLYAT p.(Val57Phe), 7: GLYAT p.(Gln61Leu), 8: GLYAT p.(Met65Thr), 9: GLYAT p.(His101Tyr), 10: GLYAT p.(Ala231Thr). First antibody was anti-GLYAT PA5-48504, 1:1,000.

Thereby, all sequence variants were sufficiently overexpressed except p.(Gly38*) (a mutant with an early stop codon) and p.(Val57Phe). A lower expression intensity was detected for the sequence variant p.(Arg199Cys). The p.(Val57Phe) sequence variant revealed two unspecific bands at 25 and 35 kDa, which probably represented degradation products. After the successful validation of overexpressions, the spin-column purifications were performed (Section 4.5.2.).
In summary, the overexpression of GLYAT sequence variants worked for all chosen variants (Section 3.2.) except for p.(Val57Phe), which did not work in other series as well (data not shown).

Published information on GLYATL1 overexpression conditions were scarce, which made a systematic testing compulsory (Figure 4.9.).

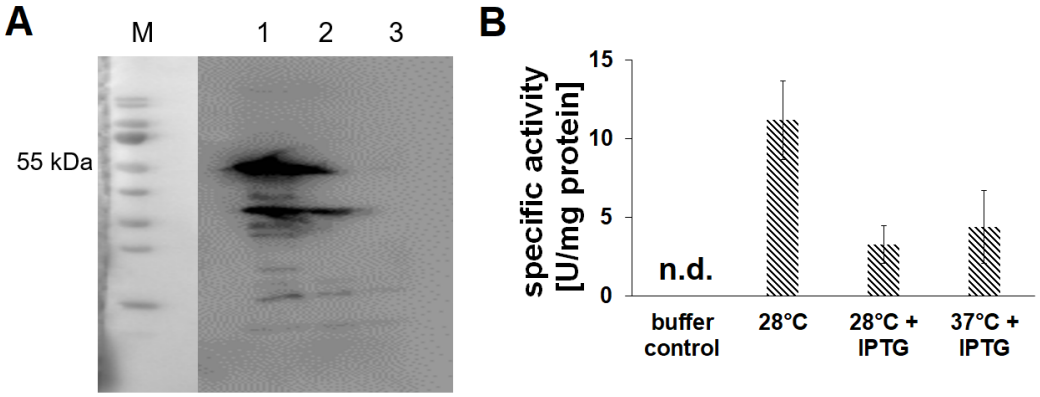


Figure 4.9: A: GLYATL1 overexpression Western blot of *E. coli* Origami 2(DE3) (12 µg total protein supernatant) (A) and maximum enzyme activities under different induction conditions (B): 1: *E. coli* Origami 2(DE3) + GLYATL1 (28 °C), 2: *E. coli* Origami 2(DE3) + GLYATL1 (28 °C + IPTG), 3: *E. coli* Origami 2(DE3) + GLYATL1 (37 °C + IPTG). B: Determination of maximum GLYATL1 activity using 100 µmol/L phenylacetyl-coA for induction conditions named above (n = 6). B: buffer control: 0, 28 °C: 11.2 ± 2.5 U/mg protein, 28 °C + IPTG: 3.3 ± 1.2 U/mg protein, 37 °C + IPTG: 4.4 ± 2.3 U/mg protein.

The GLYATL1 overexpression was most efficient for 28 °C induction indicated by the highest enzyme activity and strongest signal on the Western blot (Figure 4.9.).

Hence, for overexpression of all sequence variants of GLYAT and GLYATL1 the same induction and overexpression conditions (induction temperature 28 °C, without IPTG) were selected.

The GLYATL1 sequence variants p.(Glu63Lys), p.(Val125Leu) and p.(Asp255Asn) were sufficiently overexpressed and thus used for spin-column purifications (Section 4.5.4.).

4.5. Purifications of GLYAT and GLYATL1 from *E. coli* Origami 2(DE3)

4.5.1. FPLC purification of GLYAT wild-type and sequence variants

The first purifications were performed by using the FPLC system (Section 3.5.4.). The elution profiles of the purifications and Coomassie stained SDS gels of human GLYAT wild-type are shown (Figures 4.10., 4.11.).

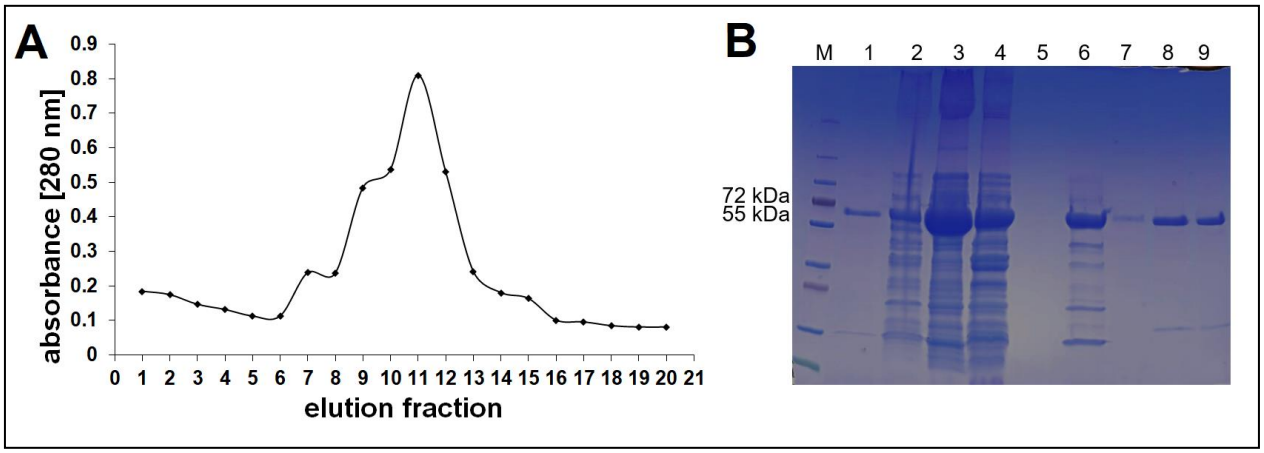


Figure 4.10: Elution chromatogram (A) and SDS-PAGE (B) of GLYAT wild-type purification via FPLC. B: M: PageRuler™ Prestained Protein Ladder, 1: *E. coli* Origami 2(DE3) + GLYAT wild-type pre induction, 2: *E. coli* Origami 2(DE3) + GLYAT wild-type post induction, 3: homogenate supernatant, 4: homogenate pellet, 5: flow-through, 6: wash fraction 1, 7: wash fraction 2, 8: elution, 9: desalted protein.

The purification of Trx-His-GLYAT protein (56 kDa) was suitable (Figure 4.10.). The elution chromatogram demonstrated a prominent elution peak in the fractions 6-16, which were united afterwards. A pure and distinct protein signal at 56 kDa was visible after dialysis (sample 9, Figure 4.10. B). Nevertheless, a high protein loss was detected in the lysate-pellet and wash fraction 1. After pre-concentration, 0.5 mL of 0.4 mg/mL protein solution remained. Thus, 0.2 mg target protein were isolated.

The GLYAT wild-type was purified in a 2nd series (Figure 4.11.) to generate a biological replicate.

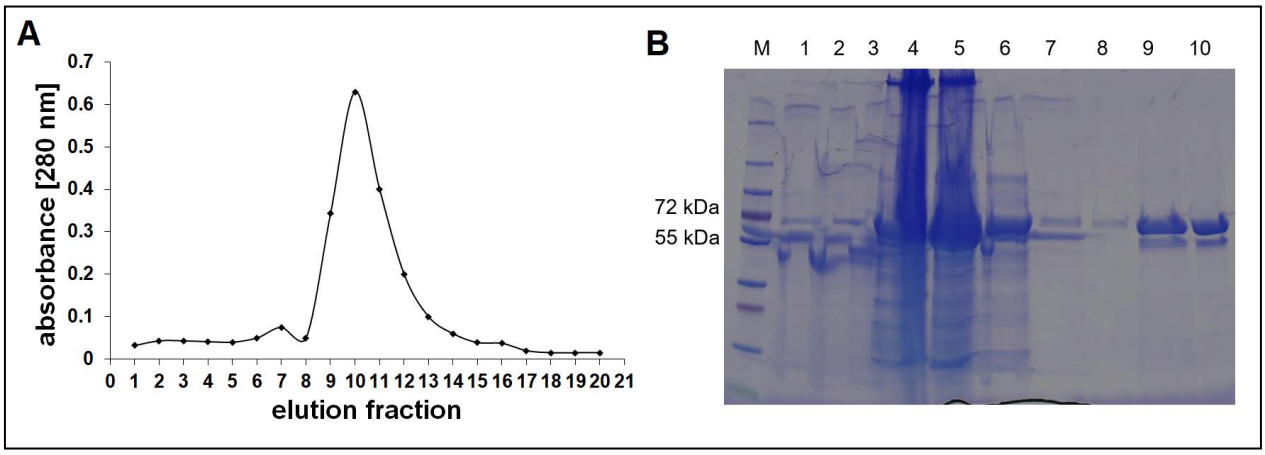


Figure 4.11: Elution chromatogram (A) and SDS-PAGE (B) of the second GLYAT wild-type FPLC-purification. A: Elution fractions 8-16 were combined and transferred into dialysis tubing; B: M: PageRuler™ Prestained Protein Ladder, 1: *E. coli* Origami 2(DE3) + GLYAT wild-type pre induction, 2: *E. coli* Origami 2(DE3) + GLYAT wild-type post induction, 3: homogenate supernatant, 4: homogenate pellet, 5: flow-through, 6: wash fraction 1, 7: wash fraction 2, 8: elution fraction 5, 9: pooled elution fraction, 10: desalted protein.

The purification was reproduced successfully as the comparison with Figure 4.10. indicates. Thereby, 5 mg total protein were isolated here resulting from 5 mL of 1 mg/mL protein solution.

The same procedure was performed for the GLYAT sequence variants p.(Asn156Ser), p.(Arg199Cys), p.(Gln61Leu) and p.(Ala231Thr) (Suppl. Fig. 25-28).

4.5.2. Spin-column purification of GLYAT wild-type and sequence variants

To save time and costs, the GLYAT wild-type and sequence variants were purified using spin-columns accordingly (Section 3.5.4.). The main advantage was the parallel purification for all target proteins. The procedure was first validated for the GLYAT wild-type (Figure 4.12.).

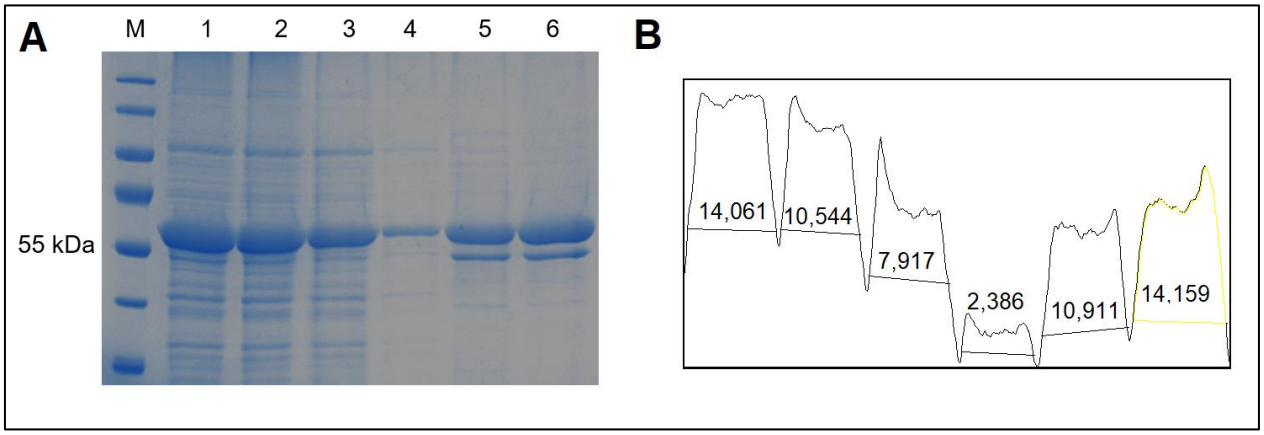


Figure 4.12.: Coomassie stained SDS PAGE of purified human GLYAT samples (A) using spin-columns and ImageJ analysis (B). M: PageRuler™ Prestained Protein Ladder, 1: supernatant fraction 2: flow-through fraction, 3: wash fraction 1, 4: wash fraction 2, 5: elution fraction 1, 6: elution fraction 2.

Compared to the purification with a FPLC system (Section 4.5.1.), the isolations with spin columns also lead to a suitable degree of purification. The ImageJ analysis indicated that the peaks of the eluted proteins (areas: 10,911, 14,159) resemble those of the supernatant and flow-through fractions (areas: 14,061, 10,544). That underlines a sufficient extraction of target protein from the supernatant. However, the gel demonstrated 2 signals with nearly the same size. The target protein Trx-His-GLYAT appeared in the upper signal, which indicates a size nearly above 55 kDa. The contaminating protein background was stronger in the elution fractions compared to FPLC preparations (Section 4.5.1.) due to less used column material within the spin-columns.
The purification with spin-columns was performed for the sequence variants p.(Gln61Leu) and p.(Asn156Ser) in parallel and the purification was subsequently analyzed by a Coomassie stained SDS gel and Western blot (Figure 4.13.).

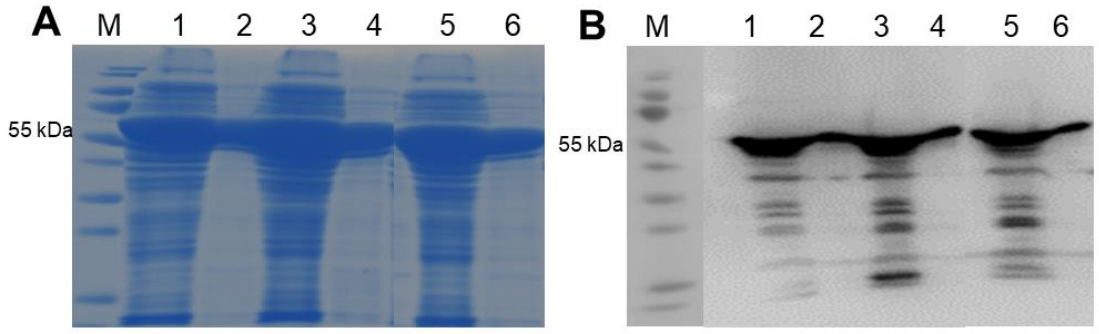


Figure 4.13: Coomassie stained SDS PAGE (A) and Western blot (B) of spin-column prepared proteins: GLYAT wild-type, p.(Asn156Ser) and p.(Gln61Leu). For the homogenates 15 µg total protein were loaded; for the purified proteins 3 µg total protein were applied. M: PageRuler™ Prestained Protein Ladder, 1: *E. coli* Origami 2(DE3) GLYAT wild-type homogenate, 2: GLYAT wild-type purified, 3: GLYAT p.(Asn156Ser) homogenate, 4: GLYAT p.(Asn156Ser) purified, 5: GLYAT p.(Gln61Leu) homogenate, 6: GLYAT p.(Gln61Leu) purified. First antibody was anti-GLYAT PA5-48504, 1:1,000.

The blot demonstrated that GLYAT wild-type and sequence variants p.(Asn156Ser) and p.(Gln61Leu) were overexpressed and purified appropriately, while the comparison of 15 µg homogenate protein and 3 µg of purified protein signal show only a intensity decrease of about 20 %. The purity grade was sufficient and feasible for activity measurements, which was confirmed by Coomassie staining (Figure 4.13 A). However, the target protein might be overlaid here due to parallel expression of the GroEL chaperonin (60 kDa).

All generated GLYAT sequence variants were analyzed and compared to wild-type (Western blot Figure 4.14.) in another series.

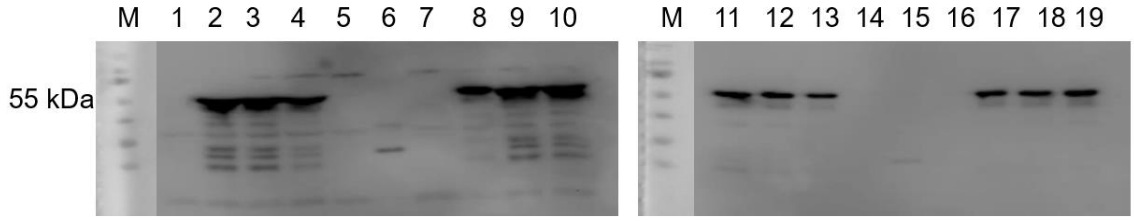


Figure 4.14: Western blot of purified GLYAT wild-type and sequence variants. 20 µg total protein were loaded for the homogenates (1-10) and 1.5 µg total protein were applied for purified proteins (11-19). M: PageRuler™ Prestained Protein Ladder, 1: pET32a(+), 2: GLYAT wild-type, 3: GLYAT p.(Asn156Ser), 4: GLYAT p.(Arg199Cys), 5: GLYAT p.(Gly38*), 6: GLYAT p.(Val57Phe), 7: GLYAT p.(Gln61Leu), 8: GLYAT p.(Met65Thr), 9: GLYAT p.(His101Tyr), 10: GLYAT p.(Ala231Thr), 11: GLYAT wild-type, 12: GLYAT p.(Asn156Ser), 13: GLYAT p.(Arg199Cys), 14: GLYAT p.(Gly38*), 15: GLYAT p.(Val57Phe), 16: GLYAT p.(Gln61Leu), 17: GLYAT p.(Met65Thr), 18: GLYAT p.(His101Tyr), 19: GLYAT p.(Ala231Thr). First antibody was anti-GLYAT PA5-48504, 1:1,000.

All sequence variants including GLYAT wild-type were successfully overexpressed with the same expression intensities. However, p.(Val57Phe) and p.(Gln61Leu) represented exceptions. An effective purification was confirmed due to comparable signal strengthes between homogenates (samples 1-10) and purified proteins (samples 11-19). Furthermore, decreasing background signals of samples 1-10 compared to samples from 11-19 are proofs for sufficient purifications.
Due to the absent overexpression of sequence variant p.(Gln61Leu), the whole procedure was repeated with two clones of each sequence variant (Figure 4.15.). The absence of target protein for the p.(Gln61Leu) variant cannot be related to the used PA5-48504 GLYAT antibody. The immunogen is a synthetic peptide, which binds to a central region of human GLYAT (amino acid 171-199) [212]. Hence, only the slightly reduced signal of p.(Arg199Cys) can be explained thereby. Because p.(Val57Phe) was also not overexpressed in the previous studies (Figure 4.14.) the production of this variant was not repeated.

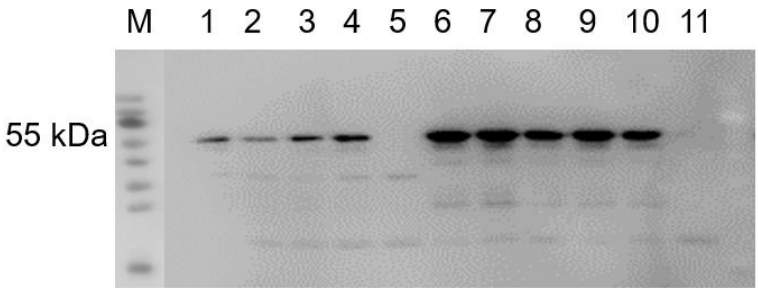


Figure 4.15: Western blot of purified GLYAT wild-type and sequence variants p.(Asn156Ser) and p.(Gln61Leu). 1.5 µg total protein loaded for the homogenates were compared to 1.5 µg total protein loaded of the purified proteins. M: PageRuler™ Prestained Protein Ladder, 1-5: *E. coli* Origami 2(DE3) homogenates; 1: GLYAT wild-type (clone 1), 2: GLYAT p.(Asn156Ser) (clone 1), 3: GLYAT p.(Asn156Ser) (clone 2), 4: GLYAT p.(Gln61Leu) (clone 1), 5: GLYAT p.(Gln61Leu) (clone 2), 6-11: purified proteins; 6: GLYAT wild-type (clone 1), 7: GLYAT wild-type (clone 2), 8: GLYAT p.(Asn156Ser) (clone 1), 9: GLYAT p.(Asn156Ser) (clone 2), 10: GLYAT p.(Gln61Leu) (clone 1), 11: GLYAT p.(Gln61Leu) (clone 2). First antibody was antil-GLYAT PA5-48504, 1:1,000.

Likewise, one clone of p.(Gln61Leu) was expressed (sample 4), whereas the other clone showed the opposite (sample 5). The same amount of homogenate and purified protein were loaded (1.5 µg) and the signal intensity increased. In addition, this demonstrates an efficient concentration by purification of target proteins.

The Figure 4.16. is a summary of the 2nd spin-column series and the clone 1 of p.(Gln61Leu) of the 3rd spin-column purification series with the aim to show all expressed clones in one Western blot.

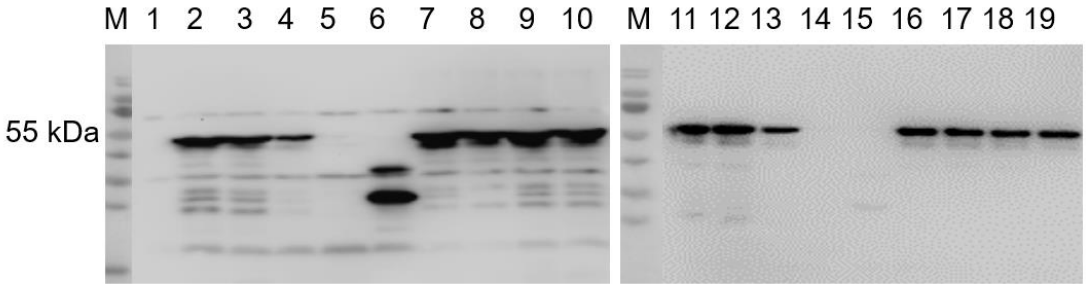


Figure 4.16: Western blot of purified GLYAT wild-type and sequence variants (summary of Figure 4.14. and 4.15.). 20 µg total protein were loaded for the homogenates (1-10), whereas 1.5 µg total protein were applied for purified proteins M: PageRuler™ Prestained Protein Ladder, 1: pET32a(+), 2: GLYAT, 3: GLYAT p.(Asn156Ser), 4: GLYAT p.(Arg199Cys), 5: GLYAT p.(Gly38*), 6: GLYAT p.(Val57Phe), 7: GLYAT p.(Gln61Leu), 8: GLYAT p.(Met65Thr), 9: GLYAT p.(His101Tyr), 10: GLYAT p.(Ala231Thr), 11: GLYAT wild-type, 12: GLYAT p.(Asn156Ser), 13: GLYAT p.(Arg199Cys), 14: GLYAT p.(Gly38*), 15: GLYAT p.(Val57Phe), 16: GLYAT p.(Gln61Leu), 17: GLYAT p.(Met65Thr), 18: GLYAT p.(His101Tyr), 19: GLYAT p.(Ala231Thr). First antibody was anti-GLYAT PA5-48504, 1:1,000.

The Western blots (Figures 4.14.-16.) indicated that all GLYAT sequence variants were overexpressed at least one time (except p.(Val57Phe)), which was a prerequisite for activity determinations. In a related study, the GLYAT wild-type and p.(Asn156Ser) were purified twice to check the effects of a protease inhibitor cocktail (PIC) from Sigma (Figure 4.17.).

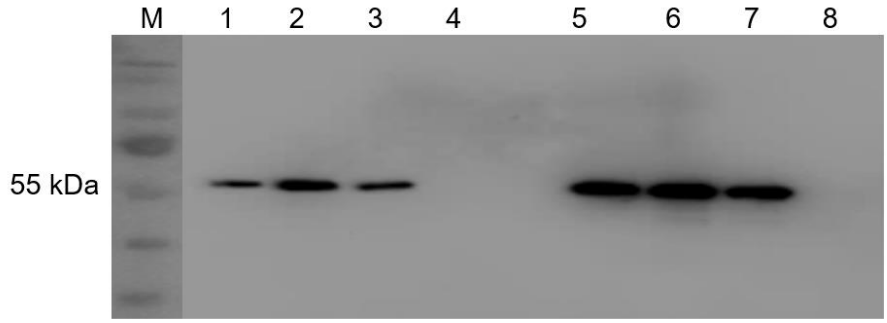


Figure 4.17: Western blot of homogenates and purified GLYAT wild-type and sequence variants p.(Asn156Ser) and p.(Gln61Leu) from *E. coli* Origami 2(DE3). 1.5 µg protein were loaded in each lane. M: PageRuler™ Prestained Protein Ladder, 1: *E. coli* Origami 2(DE3) + GLYAT wild-type homogenate, 2: *E. coli* Origami 2(DE3) + GLYAT p.(Asn156Ser) homogenate , 3: *E. coli* Origami 2(DE3) + GLYAT p.(Asn156Ser) homogenate (treated with PIC), 4: *E. coli* Origami 2(DE3) + GLYAT p.(Gln61Leu) homogenate, 5: purified GLYAT wild-type, 6: purified GLYAT p.(Asn156Ser), 7: purified GLYAT p.(Asn156Ser) (treated with PIC), 8: purified GLYAT p.(Gln61Leu). First antibody was anti-GLYAT PA5-48504, 1:1,000.

The Western blot demonstrated repetitively that the purification of the GLYAT variants was successful. The signal intensities of the loaded purified proteins was increased in comparison to the homogenates (same amounts of total loaded protein: 1.5 µg). Likewise, the GLYAT sequence variant p.(Gln61Leu) was repeatedly not expressed within this series (compared to Figure 4.15.) assuming false positive clone on the *E. coli* Origami 2(DE3) transformation plate. However, the remaining sequence variants and the wild-type were sufficiently overexpressed. No difference was obvious between p.(Asn156Ser) without treatment and with treatment of PIC.

Only the overexpressed proteins (Figures 4.13. – 4.17.) were used for the determinations of the enzyme activities (Section 4.6.5.).

4.5.3. FPLC purification of GLYATL1 wild-type protein

The purification of Trx-His-GLYATL1 wild-type protein was performed 3 times and the approach with the highest target protein yield is shown in Figure 4.18.

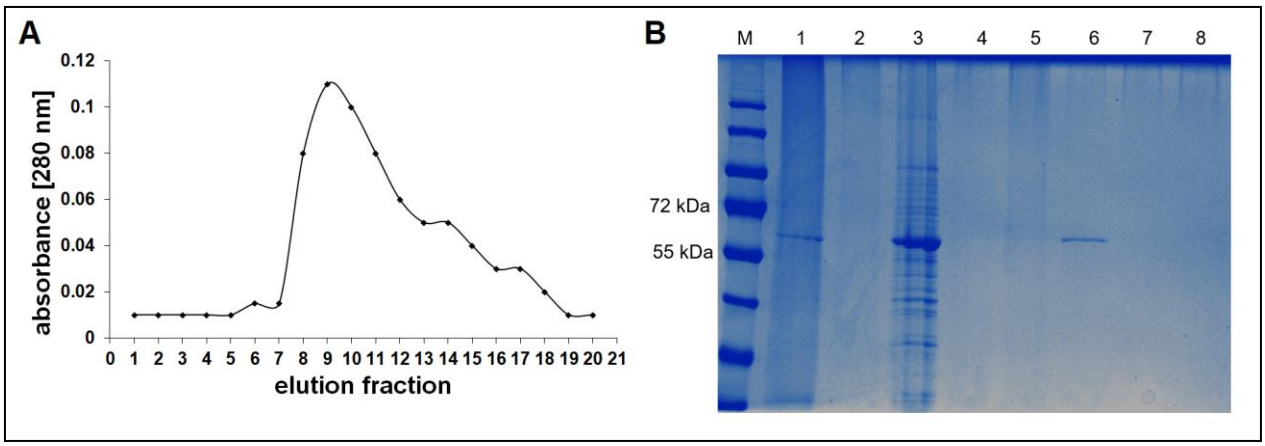


Figure 4.18: Elution chromatogram (A) and SDS-PAGE (B) of human GLYATL1 purification via FPLC. A: Elution chromatogram. Elution fractions 8-15 were combined and transferred into dialysis tubing; B: M: PageRuler™ Prestained Protein Ladder, 1: *E. coli* Origami 2(DE3) + GLYAT wild-type homogenate supernatant, 2: flow-through, 3: wash fraction 1, 4: wash fraction 2, 5: elution fraction 3, 6: pooled elution fractions, 7: desalted protein, 8: pre-concentrated protein.

In the elution fraction (sample 6) a remainder of target protein was released but main part got lust after washing as sample 3 indicates (Figure 4.18 B). The weak protein signal in elution fraction (sample 6) is corroborated by the low absorbance maximum at 0.12 (Figure 4.18 A). The absence of the GLYATL1 protein was confirmed by an enzyme activity test of the desalted fraction in addition (data not shown). Due to complete loss of target protein after dialysis (sample 7), the FPLC-purification was not applied for the GLYATL1 wild-type and it was proceeded using spin columns (Section 4.6.5.) having the advantage of parallel analysis.

4.5.4. Spin-column purification of GLYATL1 wild-type and sequence variants

The Figure 4.19. demonstrates the results of the GLYATL1 overexpression series.

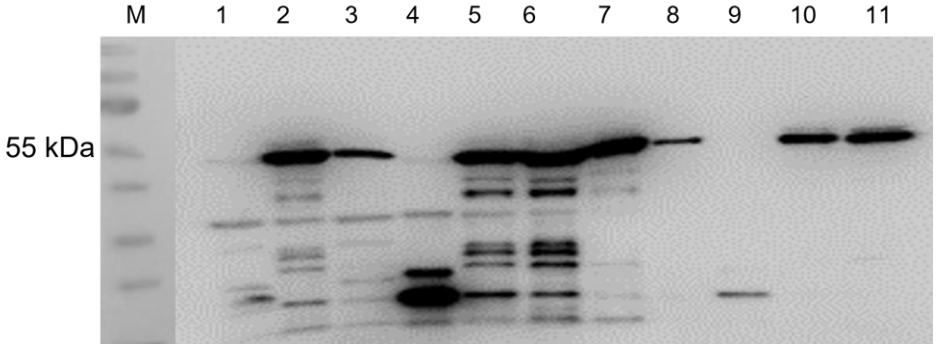


Figure 4.19: Western blot of purified GLYATL1 wild-type and sequence variants (1). 15 µg total protein homogenate of *E. coli* Origami 2(DE3) (samples 1-6) were compared to 3 µg purified protein (samples 7-11). 1: pET32a(+), 2: GLYATL1 wild-type, 3: p.(Glu67Lys), 4: p.(Glu87*), 5: p.(Val125Leu), 6: p.(Asp255Asn), 7: GLYATL1 wild-type, 8: GLYATL1 p.(Glu63Lys), 9: GLYATL1 p.(Glu87*), 10: p.(Val125Leu), 11: GLYATL1 p.(Asp255Asn). First antibody was anti-GLYATL1 GTX106956, 1:1,000.

The comparison between 15 µg total protein homogenate (sample 1-6) to 3 µg of purified proteins (samples 7-11) demonstrated nearly equal signal intensities, which confirmed a sufficient protein

purification. However, the p.(Glu63Lys) sequence variant showed a weaker protein signal (sample 8). As expected, p.(Glu87*) yielded no target protein signal at 56 kDa.
In another purification series, the p.(Glu63Lys) was not expressed (data not shown). To describe that in more detail, the whole purification was repeated (Figure 4.20.).

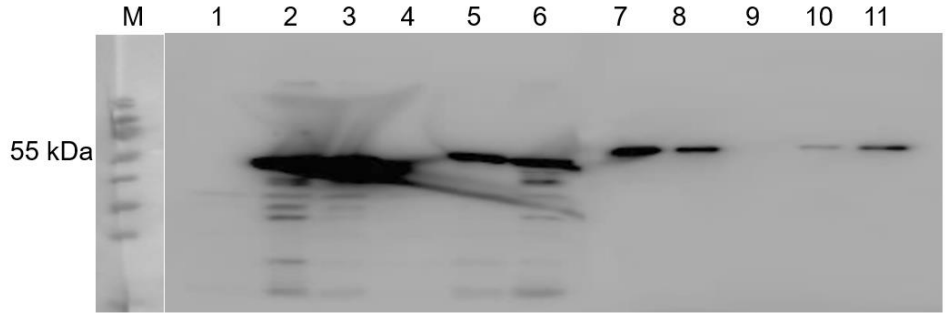


Figure 4.20: Western blot of purified GLYATL1 wild-type and sequence variants (2). M: PageRuler™ Prestained Protein Ladder, 1-7: *E. coli* Origami 2(DE3) homogenates (20 µg); 8-11: purified proteins (3 µg); 1: pET32a(+), 2: GLYATL1 wild-type, 3: GLYATL1 p.(Glu63Lys), 4: GLYATL1 p.(Glu87*), 5: GLYATL1 p.(Val125Leu), 6: GLYATL1 p.(Asp255Asn), 7: GLYATL1 wild-type purified, 8: GLYATL1 p.(Glu63Lys), 9: GLYATL1 p.(Glu87*), 10: GLYATL1 p.(Val125Leu), 11: GLYATL1 p.(Asp255Asn). First antibody was anti-GLYATL1 GTX106956, 1:1,000.

The Western blot demonstrated a sufficient purification of the GLYATL1 wild-type and sequence variants. Nonetheless, the eluted protein amount of p.(Val125Leu) variant seemed to be limited.
To extend these analyses and to investigate the purity grade of GLYATL1 wild-type and sequence variant p.(Glu63Lys) more precisely, both homogenates and purified proteins were compared via juxtaposition of Coomassie staining, silver staining and Western blot (Figure 4.21.). Moreover, different sensitivity ranges of the procedures were compared.

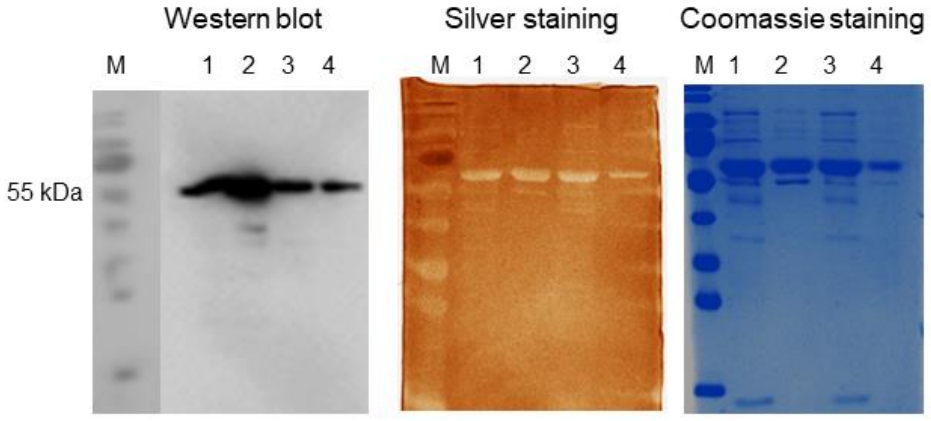


Figure 4.21: Western blot, silver and Coomassie staining of SDS PAGE of *E. coli* Origami 2(DE3) homogenates and purified proteins. 1.5 µg total protein were loaded in each fraction. 1: GLYATL1 wild-type, 2: GLYATL1 wild-type purified, 3: p.(Glu63Lys), 4: p.(Glu63Lys) purified. First antibody was anti-GLYATL1 GTX106956, 1:1,000.

The GLYATL1 wild-type homogenate and purified protein demonstrated more intensive signal bands compared with p.(Glu63Lys) samples (Figure 4.21.). As the silver and Coomassie staining

indicated, the background is reduced for the purified proteins. The silver staining method, which should be more sensitive did not show any advantages compared with the Coomassie staining due to very similar signals.

4.5.5. FPLC purification of the GLYATL2 wild-type

The purification of Trx-His-GLYATL2 was outsourced to a BSc thesis [164]. The preparational experiments and the co-supervision of the enzyme analyses were performed by Daniel Schulke. However, one representative purification experiment is shown here by SDS PAGE and Western blot (Figure 4.22.). The aim was to connect to the previous work on GLYATL2 subcloning (Section 4.3.) by demonstrating the purification outcome.

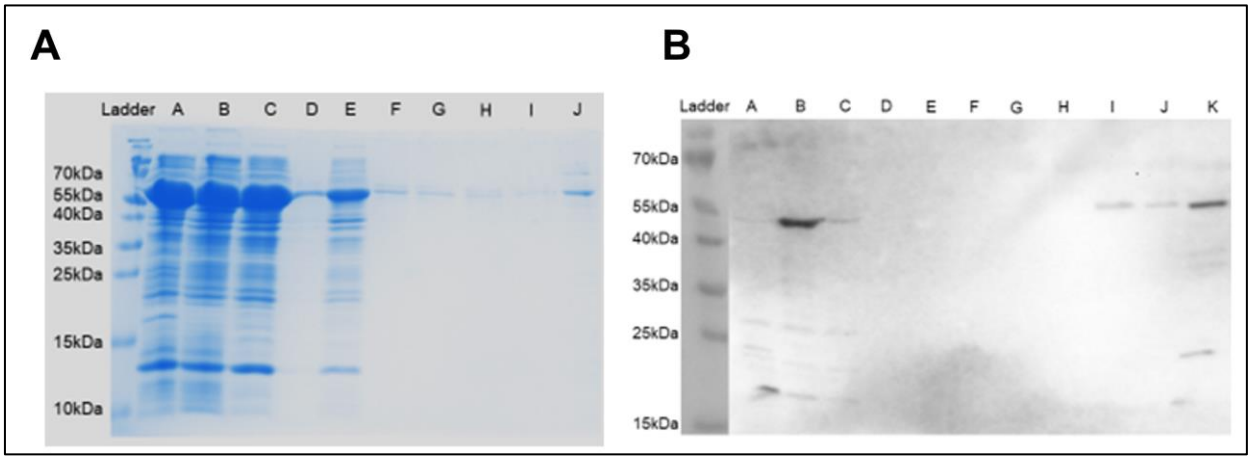


Figure 4.22: Coomassie stained SDS PAGE (A) and Western blot (B) of GLYATL2 wild-type purification using FPLC [164]. A: A: homogenate empty vector, B: homogenate GLYATL2 wild-type, C: flow-through, D: 25 mM Tris-HCl (Wb: wash fraction 1), E: wash fraction 1 (Wb: wash fraction 2), F: wash fraction 2 (Wb: wash fraction 3), G: wash fraction 3 (Wb: wash fraction 4), H: wash fraction 4 (Wb: wash fraction 5), I: wash fraction 5 (Wb: pre-dialysis), J: purified GLYATL2 wild-type (Wb: post-dialysis), K (Wb): purified GLYATL2 wild-type. Empty vector homogenate (hom.) and 25mM Tris-HCl, pH 8.0 (dialysis buffer) were applied as controls, although the buffer control was spilled with flow through (FT) while loading the gel. 20 µg total protein were loaded for homogenates, whereas 15µL of the other samples were mixed with 4x Laemmli and applied onto the gel. Hence, 18 µg FF, 4.5 µg WF I and 0.2 µg pure GLYATL2 were loaded. B: Either 20 µg protein (homogenates) or 12 µL of the sample (FT, WF, purified GLYATL2 pre-and post dialysis and purified, concentrated GLYATL2 wt (1.8 µg)) were applied. Signals were detected using anti-GLYATL2 antibody (PA5-43275, Thermo Fisher).

Western blot analysis of overexpressed proteins purified by FPLC resulted in fewer bands (Figure 4.22. B) than analysis of such samples on the Coomassie stained gel (Figure 4.22. A).

The whole purification procedure of the GLYATL2 wild-type seems to have worked, however there was still a loss of target protein indicated by band intensity of GLYATL2 homogenate (sample B) compared with pre-concentrated GLYATL2 protein (sample K). There was no loss of target protein as wash fractions D-H (Figure 4.22. B) clearly demonstrate. The dialysis worked sufficiently due to no loss of GLYATL2 target protein (samples I and J). It must be considered that the loaded protein amount for homogenate (20 µg) was 10x higher compared with purified GLYATL2 (1.8 µg).

The blot also yielded weak unspecific protein signals at 22 and 17 kDa. The corresponding enzyme activity tests were performed in the co-supervised B.Sc. thesis [164].

4.6. Enzyme activity tests of purified GLYAT wild-type and sequence variants

The GLYAT activity determinations were performed based on the validation of linear range of L-cysteine absorbances (Section 3.11.)

After successful purification of human GLYAT as Trx-His fusion protein via FPLC (Figure 4.10.) the maximum specific activity was determined 2x using 200 µmol/L benzoyl-coA (Section 3.11.4) with a specific enzyme activity of 1,547 U/mg protein (ẋ).

The maximum enzyme activity of the other purification of Trx-His-GLYAT wild-type (Figure 4.11.) was determined (Figure 4.23.).

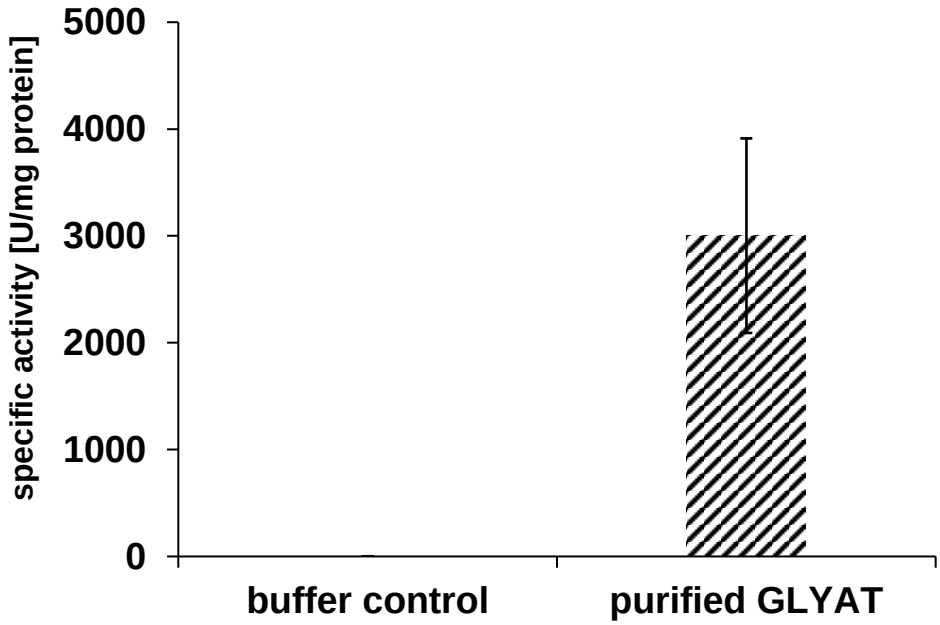


Figure 4.23: Maximum specific activity of the FPLC-purified GLYAT wild-type sample (Figure 4.11.) with 200 µmol/L benzoyl-coA (n=4, technical replicates). Four different protein amounts (5, 10, 20 and 40 µg) were used for maximum activity detection of the same sample. Enzyme activity was detected at 3,002 ± 910 U/mg protein. Variation coefficient: 30 %.

The diagram underlines that specific GLYAT activity here (3,002 ± 910 U/mg) was higher compared to the previous series (1,547 U/mg protein). However, the standard deviation of 901 U/mg protein and the calculated variation coefficient of 30 % demonstrate a considerable data variation.

4.6.1. Intra- and inter-assay variation of FPLC purified human GLYAT

The enzyme activities of the FPLC purified GLYAT wild-type (Figure 4.24.) and the sequence variant p.(Gln61Leu) (Figure 4.25.) were detected 6x to examine the intra-assay variation.

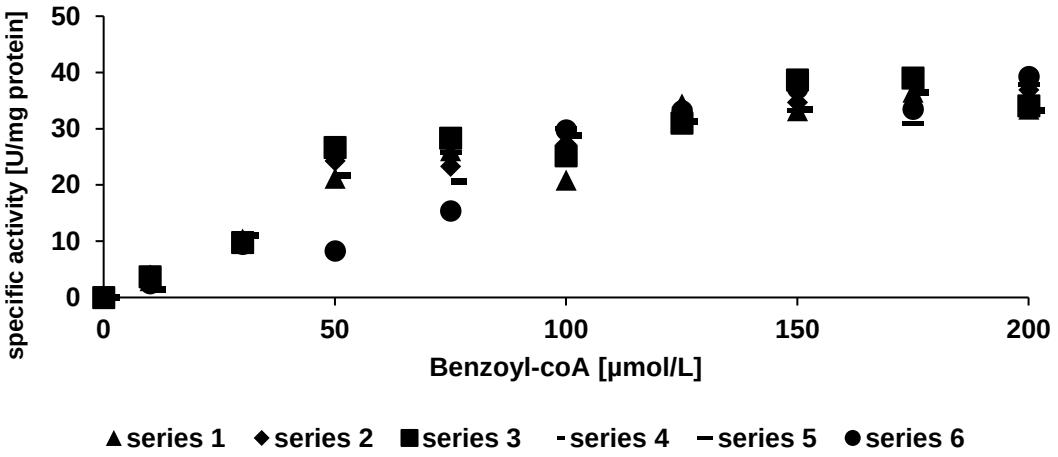


Figure 4.24: Intra-assay variation study of 1.5 µg purified GLYAT wild-type by measuring kinetic curve 6x with the same sample (n=6, technical replicates). Benzoyl-coA concentrations from 0 to 200 µmol/L were applied and specific activities were calculated; v_{max} and K_M values were calculated using Excel Solver tool. Variations coefficient for v_{max} is 8 %, whereas for K_M 19 % were calculated.

With v_{max} of 53.3 ± 4.1 U/mg protein and K_M of 87.6 ± 16.5 µmol/L the purified Trx-His-GLYAT showed the expected curve regressions. The variation coefficients of 8 % (v_{max}) and 19 % (K_M) describe the data variation within the intra-assay series.

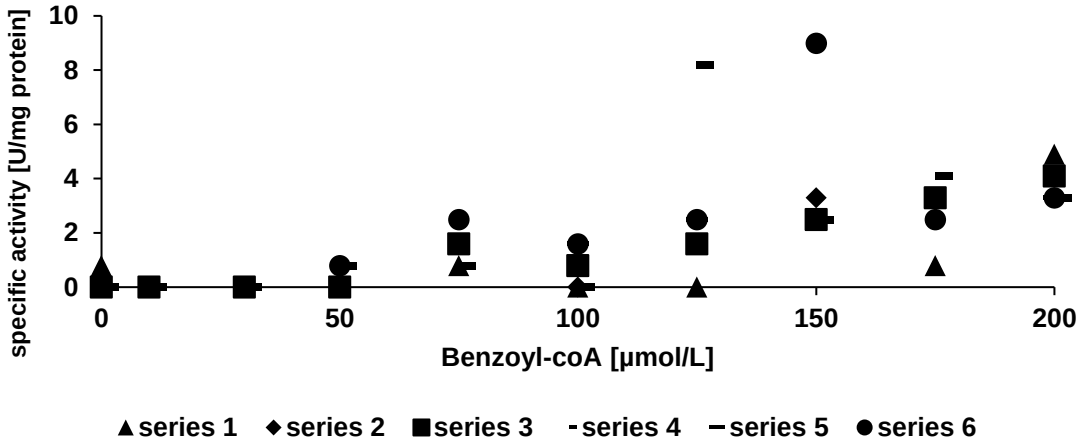


Figure 4.25: Intra-assay variation study of purified GLYAT sequence variant p.(Gln61Leu) by measuring kinetic curve 6x with the same sample (n=6, technical replicates). Benzoyl-coA concentrations from 0 to 200 µmol/L were applied and specific activities were calculated; v_{max} and K_M could not be determined due to missing saturation of the curves.

No kinetic parameters were calculated for p.(Gln61Leu) due to absent saturation of the graph. However, the lower activity for p.(Gln61Leu) compared to the wild-type was a consistent finding.

The Figure 4.26. demonstrates the inter-assay variation across 6 series measured with the GLYAT wild-type in different samples on different days.

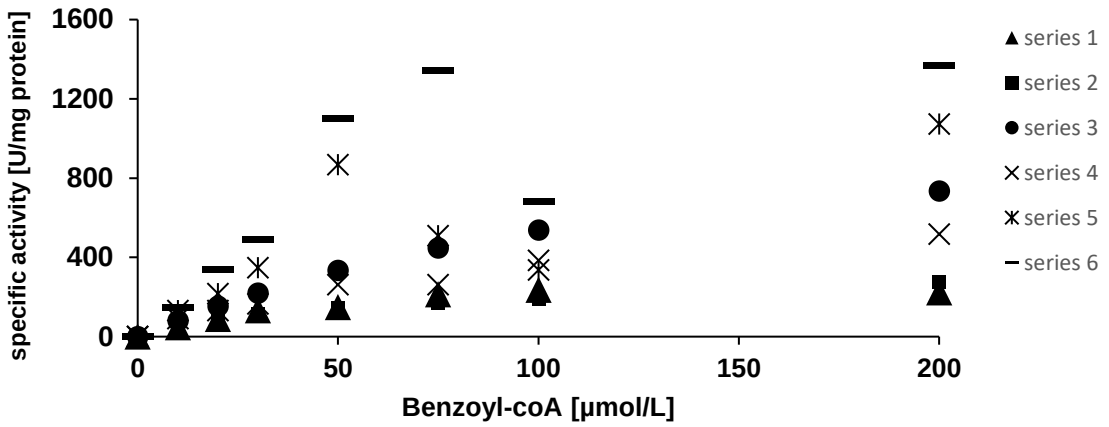


Figure 4.26.: Inter-assay variation study of 6 different purified Trx-His-GLYAT wild-type samples (FPLC) across different days (n=6, biological replicate). Benzoyl-coA concentrations from 0 to 200 µmol/L were applied and specific activities were calculated. V_{max} was 948 ± 554 U/mg protein (variation coefficient: 58 %) and K_M: 85 ± 36 µmol/L (variation coefficient: 42 %).

A higher data variation of the GLYAT wild-type was obvious compared with the wild-type intra-assay study (Figure 4.24.). This was confirmed by the higher variation coefficients calculated here (v_{max}: 58 %, K_M: 42 %).

4.6.2. Long-time stability tests of purified GLYAT wild-type

To investigate the long-time stability of the purified GLYAT wild-type, maximum activity measurements were performed during the entire working time using aliquots of the same stock solutions (Figure 4.27.).

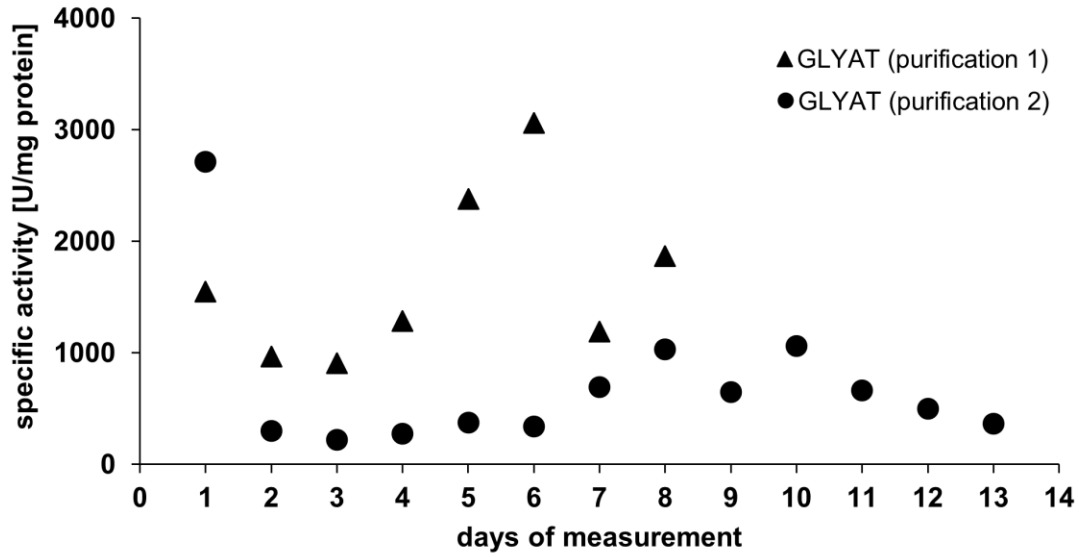


Figure 4.27: Long time stability analysis of the two FPLC-purified human GLYAT enzyme preparations (Figures 4.10. and 4.11.). All specific activities were obtained using 200 µmol/L benzoyl-coA in standard GLYAT activity assay. Triangle: data points of GLYAT activity (purification 1), circle: data points of GLYAT activity (purification 2). Variation coefficient 1st purification: 45.6 %, variation coefficient 2nd purification: 94 %, variation coefficient 2nd purification (without the 1st value): 52.7 %.

The data sets of the 2 purifications showed remarkable variations (Figure 4.27.). Notwithstanding, there was no trend of protein degradation identifiable. While the first enzyme preparation started with an activity of 1,547 U/mg protein on the first day, variations of 40 % over the time could be detected. However, on the last days of measurements there seemed to be a level off around 1,300 -1,500 U/mg protein activity. The 2nd enzyme preparation started with an activity of 2,713 U/mg protein. The activity was decreased to only 4-11 % residual activity (variation coefficient: 52.7 % without the 1st value). It leveled off at a stable value over the time, which is around 700 U/mg protein. If the first measurement of the 2nd series is excluded, the remarkable data variation is reduced to a variation coefficient of 52.7 % – the very high activity of 2,713 U/mg protein may have resulted from a technical measurement failure on this day.

The determined variation coefficients expand the one of 30 % calculated in Figure 4.23. However, the 30 % were calculated with the same sample on the same day and these samples here were measured over a long time on different days. Thus, the values were only conditionally comparable.

4.6.3. Enzyme activity of FPLC purified GLYAT wild-type compared to sequence variants

Followed by the stability analyses and the determination of maximum activity, kinetic studies with increasing benzoyl-coA concentrations were performed (Figures 4.28.-31.).

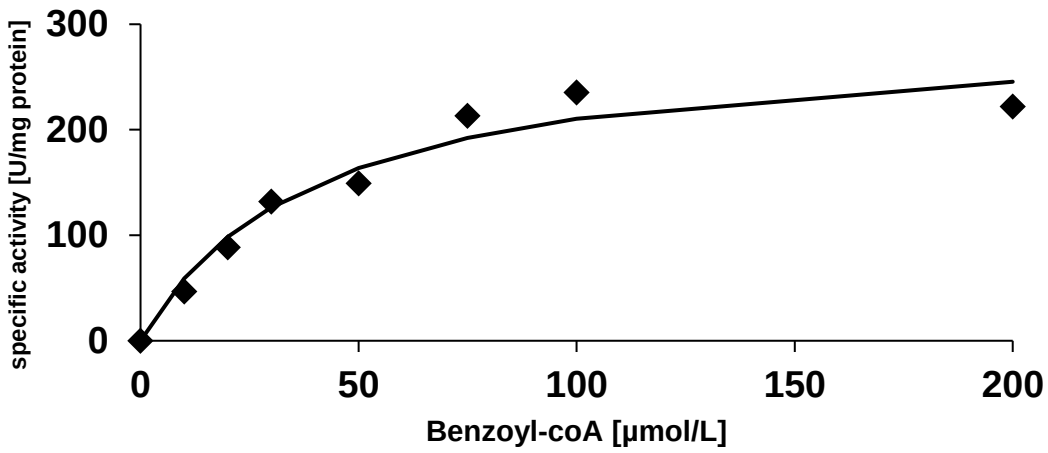


Figure 4.28: Kinetic study of FPLC purified GLYAT wild-type (Figure 4.11.). Benzoyl-coA concentrations were applied from 0 – 200 µmol/L. v_{max} and K_M were determined using the Excel Solver tool. V_{max} = 294 U/mg protein, K_M (Benzoyl-coA) = 39.8 µmol/L.

The purified GLYAT wild-type enzyme (Figure 4.11.) demonstrated an expected curve with typical Michaelis-Menten properties. The substrate saturation started at 100 µmol/L benzoyl-coA and is also detectable at 200 µmol/L. The calculated v_{max} of 294 U/mg protein and K_M of 39.8 µmol/L (Excel Solver tool) matches the curve regression.

3x determinations of enzyme activities from the FPLC purified GLYAT wild-type and the sequence variants p.(Gln61Leu) and p.(Asn156Ser) are shown below (Figures 4.29.-31.).

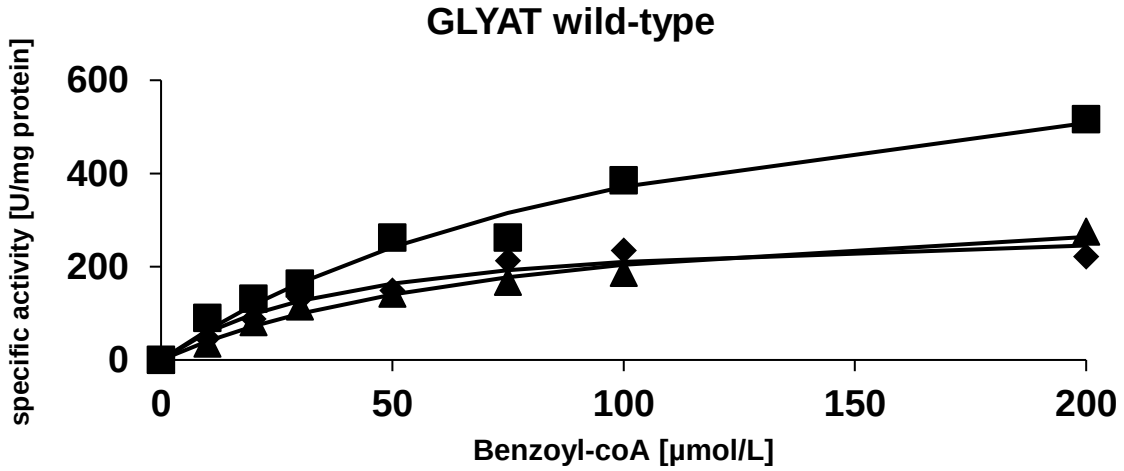


Figure 4.29: Kinetic study of the purified GLYAT wild-type (n=3). 3x determination of enzyme activity of Trx-His-GLYAT wild-type (same sample, different days of measurements) lead to following results: ♦: v_{max}: 294 U/mg protein, K_M: 39.8 µmol/L; ▲: v_{max}: 373 U/mg protein, K_M: 82.6 µmol/L, ■: v_{max}: 803 U/mg protein, K_M: 116 µmol/L. Variation coefficients: v_{max} = 56 %, K_M = 48 %.

The kinetic parameters varied substantially with v_{max} of 490 ± 274 U/mg protein and K_M of 79.5 ± 38.2 µmol/L. However, the data observed a Michaelis Menten kinetic of GLYAT enzyme. Notwithstanding, the variation coefficients of 56 % (v_{max}) and 48 % (K_M) strongly resembled the range of the values determined in Section 4.6.1. (Figure 4.26.).

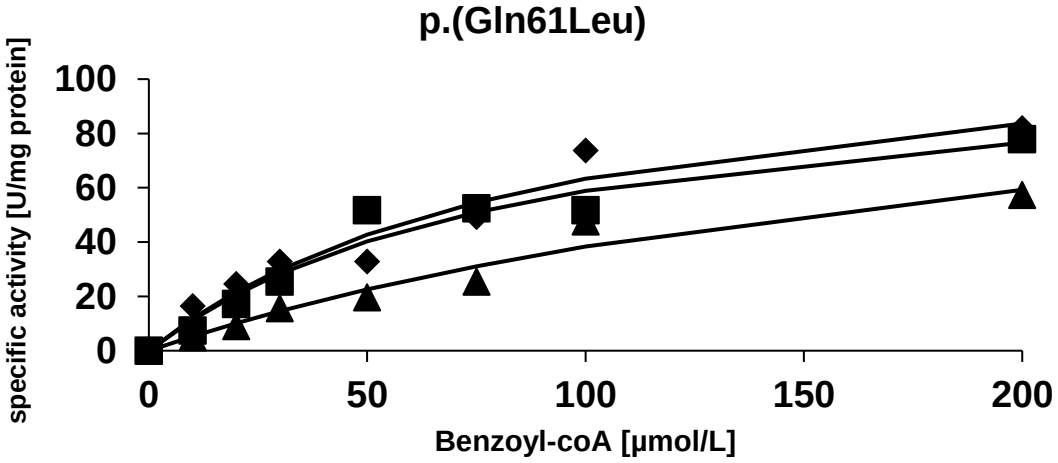


Figure 4.30: Kinetic study of the purified GLYAT p.(Gln61Leu) (n=3). 3x determination of enzyme activity of GLYAT p.(Gln61Leu) (same sample, different days of measurements) lead to following results: ♦: v_{max}: 123 U/mg protein, K_M: 93.9 µmol/L; ▲: v_{max}: 130 U/mg protein, K_M: 237 µmol/L, ■: v_{max}: 110 U/mg protein, K_M: 86.1 µmol/L. Variation coefficients: v_{max} = 8 %, K_M = 61 %.

These curves revealed a v_{max} of 121 ± 10 U/mg protein and K_M of 139 ± 85 µmol/L. With only 1/5 of maximum wild-type activity the activity of p.(Gln61Leu) sequence variant was substantially decreased while K_M seems to be 3x higher The variation coefficient of v_{max} (8 %) falls within the range determined in Section 4.6.1. (Figure 4.26.) but the K_M variation of 61 % exceed the 42 % in Section 4.6.1.

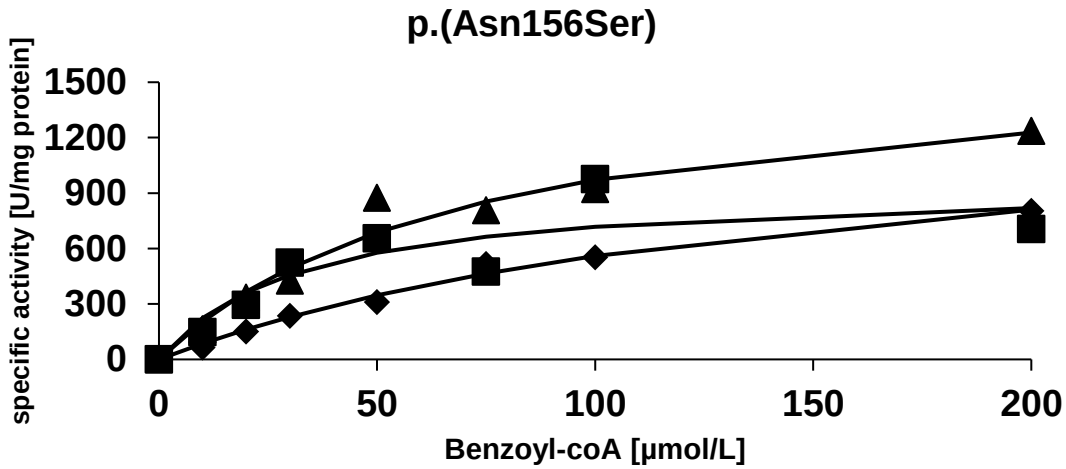


Figure 4.31: Kinetic study of the purified GLYAT p.(Asn156Ser) (n=3). 3x determination of enzyme activity of GLYAT p.(Asn156Ser) (same sample, different days of measurements) lead to following results: ♦: v_{max}: 1,462 U/mg protein, K_M : 161 µmol/L; ▲: v_{max}: 1,663 U/mg protein, K_M : 71 µmol/L, ■: v_{max}: 953 U/mg protein, K_M: 32.6 µmol/L. Variation coefficients: v_{max} = 27 %, K_M = 75 %.

The p.(Asn156Ser) variant showed a higher enzyme activity of 1,359 ± 366 U/mg protein compared to wild-type and K_M of 88 ± 66 µmol/L, which resembled the wild-type. Similar to p.(Gln61Leu), the variation coefficient of v_{max} falls within the determined range (Figure 4.26.) and an increase is detected for K_M.

These results (Figures 4.29. – 4.31.) were published [137].

The Figure 4.32. summarizes the maximum specific activities of the homogenates, which were intended for further FPLC purifications with subsequent enzyme activity determinations.

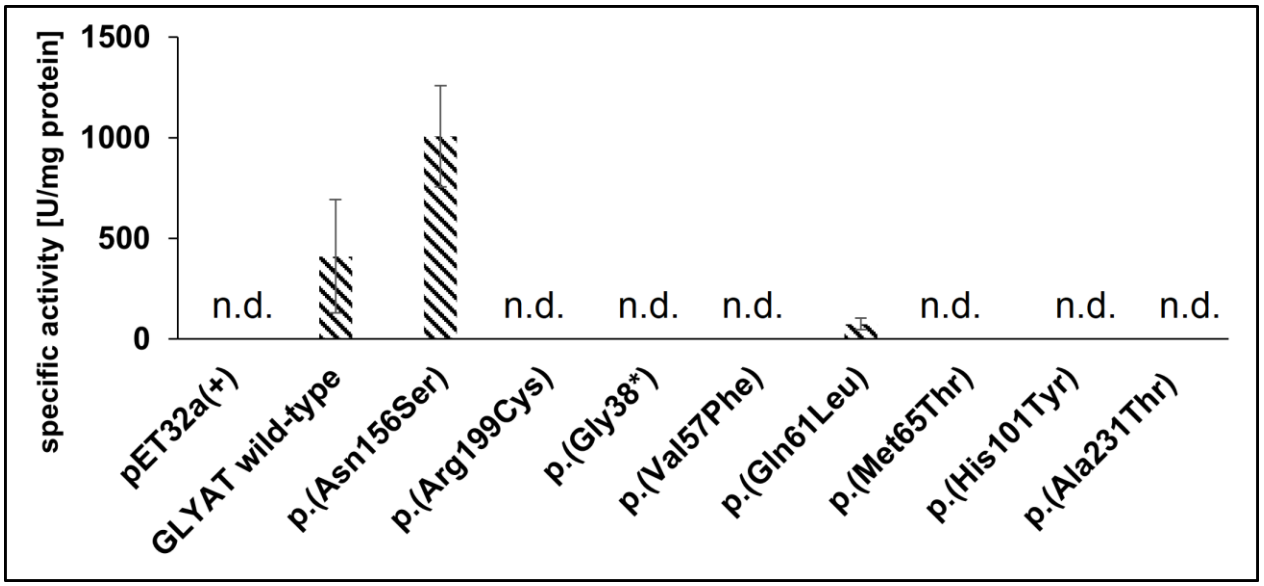


Figure 4.32: Maximum activity determinations of GLYAT wild-type and sequence variants using 200 µmol/L benzoyl-coA (n=3, biological replicates). For GLYAT wild-type 417 ± 150 U/mg protein were determined; p.(Asn156Ser) shows 1,008 ± 251 U/mg protein and p.(Gln61Leu) 75 ± 28 U/mg protein total GLYAT activity for 200 µmol/L benzoyl-coA. Remaining sequence variants and negative controls pET32a(+) and p.(Gly38*) were inactive.

The sequence variants p.(Arg199Cys), p.(Met65Thr), p.(His101Tyr) and p.(Ala231Thr) were overexpressed but no enzyme activity was detected (Figure 4.32.). Hence, no FPLC purification and kinetic measurements were performed for them.

The kinetic parameters of FPLC purified GLYAT wild-type enzyme and sequence variants are summarized by Table 35.

Table 35: Kinetic analyses of FPLC-purified GLYAT wild-type and sequence variants.

GLYAT sequence variant	v_{max} [U/mg protein] (Excel Solver)	K_M [µmol/L] (Excel Solver)
Wild-type	490 ± 274	79.2 ± 38.3
p.(Asn156Ser)	1359 ± 366	88.4 ± 66.3
p.(Arg199Cys)	cannot be determined	cannot be determined
p.(Gly38*)	cannot be determined	cannot be determined
p.(Val57Phe)	cannot be determined	cannot be determined
p.(Gln61Leu)	121 ± 10.3	139 ± 85.2
p.(Met65Thr)	cannot be determined	cannot be determined
p.(His101Tyr)	cannot be determined	cannot be determined
p.(Ala231Thr)	cannot be determined	cannot be determined

The determinations of v_{max} and K_M were only performed with successfully overexpressed and active GLYAT wild-type and sequence variants p.(Asn156Ser) and p.(Gln61Leu) (Figures 4.29.-4.31.). Nonetheless, if no overexpression has taken place, enzyme analyses were not performed. The enzyme activity of pET32a(+) empty and p.(Gly38*) sequence were measured as controls and showed no activity (Figure 4.32.).

To summarize, the sequence variant p.(Asn156Ser) revealed 3x higher activity compared to wild-type, whereas K_M values resemble each other (Figure 4.29., 31, Table 35). Conversely, the sequence variant p.(Gln61Leu) (Figure 4.30.) was 5x less active than GLYAT wild-type. However, the K_M value appeared to be 2-fold higher but the data variation (variation coefficient of 48 % for K_M of GLYAT wild-type) came into effect here. Nevertheless, the tending higher K_M of p.(Gln61Leu) confirmed lower substrate affinity indicating decreased enzyme activity.

Furthermore, all other sequence variants listed in the Table 35 showed no enzyme activity confirming results from the online prediction tools (Section 4.2.) where missense mutation should lead to inactive enzyme. All studies performed in this experimental row (Figure 4.29.-4.31., Table 35) were performed with the same biological materials measured on different days.

These results might confirm that the p.(Asn156Ser) variant is the more probable GLYAT wild-type as it was supposed in the literature [137, 154].

The spin-column purifications (Section 4.6.6.) were chosen later due to better comparability between the wild-type and sequence variants, which were extracted and purified in parallel (technical replicates) (which was not the case for the FPLC purified proteins).

4.6.4. Glycine dependency of FPLC purified GLYAT wild-type

To analyze the glycine dependency of GLYAT, the specific activity of the FPLC purified enzyme (Figure 4.10.) was measured without (-) and with (+) glycine (Figure 4.33.).

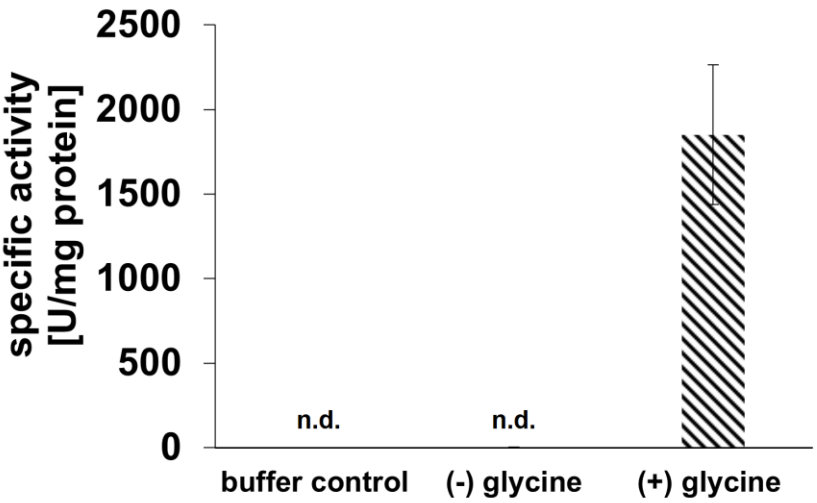


Figure 4.33: GLYAT activity assay of 1.5 µg purified GLYAT using 200 µmol/L benzoyl-coA analyzing buffer control sample without enzyme, sample without glycine (-) and with glycine (+) (n=3, technical replicates). Buffer control shows no detected signal at all, whereas purified GLYAT without glycine displays a specific activity of 2 ± 1 U/mg protein. Purified GLYAT with glycine demonstrates a specific activity of 1,851 ± 414 U/mg protein (n=3, technical replicates). Variation coefficient is 22 %.

The human GLYAT enzyme is glycine dependent as Figure 4.33 confirms. The detected activity of GLYAT was 1,851 ± 414 U/mg protein when glycine is present. A merely minor activity was

detected without glycine. The variation coefficient of 22 % lies within the tolerable range of 30 % (Section 4.6.1.). Beyond that, the unspecific degradation of benzoyl-coA can be excluded due to observed activity signal of glycine and benzoyl-coA in combination. Conversely, the benzoyl-coA degradation without GLYAT catalysis can be ruled out since there is no activity in the sample without glycine. The unspecific benzoyl-coA degradation would be visible in the sample without glycine as benzoyl-coA was supplied there and coA release would be detected then.

Beyond that, the glycine dependency was confirmed via a GC/MS experiment (cooperation with Dr. M. Fernando, Universitätsklinikum Freiburg) in addition (Figure 4.34.). The samples were prepared within this project (Section 3.13.) and subsequently analyzed in Freiburg.

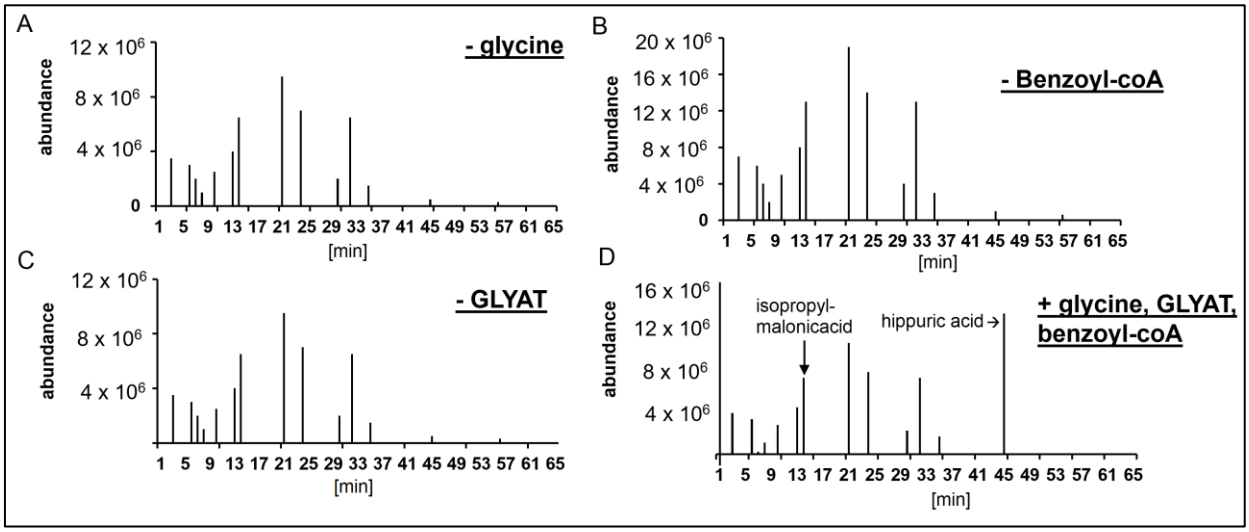


Figure 4.34: GC/MS chromatograms of the purified GLYAT wild-type. The sample volume was 1.2 mL. Incubation time was 5 min background measurement followed by 10 min reaction time after benzoyl-coA addition. The internal standard isopropylmalonicacid was added at a concentration of 0.04 mmol/L and the peak was 6 x 10^6 after 14 min. The peak for hippuric acid was double compared to the standard (12 x 10^6). Hence, the concentration of hippuric acid is 0.08 mmol/L. The absolute amount of hippuric acid in the sample is 96 nmol.

The GLYAT enzyme requires benzoyl-coA and glycine for the reaction to produce benzoylglycine (hippuric acid) as the abundance peak in the sample supplemented with glycine, GLYAT and benzoyl-coA after 45 min indicated (Figure 4.34. D).

Thereby, both series (Figure 4.33., 4.34.) corroborate the production of benzoylglycine and the glycine dependency of human GLYAT.

4.6.5. Enzyme activity of purified GLYAT wild-type and sequence variants

The present section treats the spin-column purified GLYAT in contrast to the FPLC purified enzyme (Section 4.6.3.).

The Figures 4.35.-4.40. demonstrate the kinetic studies of GLYAT wild-type and the sequence variants p.(Asn156Ser) and p.(Gln61Leu).

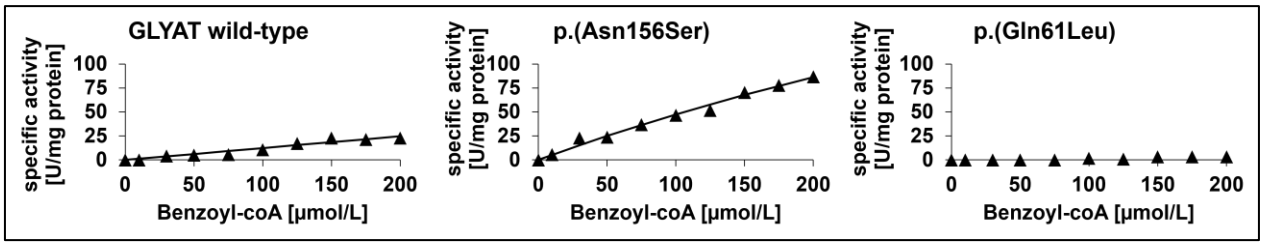


Figure 4.35: Enzyme kinetics of 1.5 µg purified GLYAT wild-type and sequence variants p.(Asn156Ser) and p.(Gln61Leu) (SDS PAGE & Western blot: Figure 4.13). GLYAT wild-type: v_{max}: cannot be determined, K_M: cannot be determined, p.(Asn156Ser): v_{max}: 477 U/mg protein, K_M: 908 µmol/L, p.(Gln61Leu): v_{max}: cannot be determined, K_M: cannot be determined.

The p.(Asn156Ser) variant showed a 3x higher activity compared with wild-type, which matches the results observed for the FPLC purified samples (Section 4.6.3.). The p.(Gln61Leu) variant revealed merely 10 % of wild-type activity. No saturation of activity values was observed for GLYAT wild-type and therefore K_M and v_{max} calculations were impossible. The activity of p.(Gln61Leu) was minor without correlation with the substrate concentration. Nevertheless, a tendency of remarkable higher p.(Asn156Ser) activity compared to the wild-type was detectable.
To test the GLYAT stability and for reproduction of kinetic results, the same measurement was performed again 6 months later with the same samples as above (Figure 4.36.).

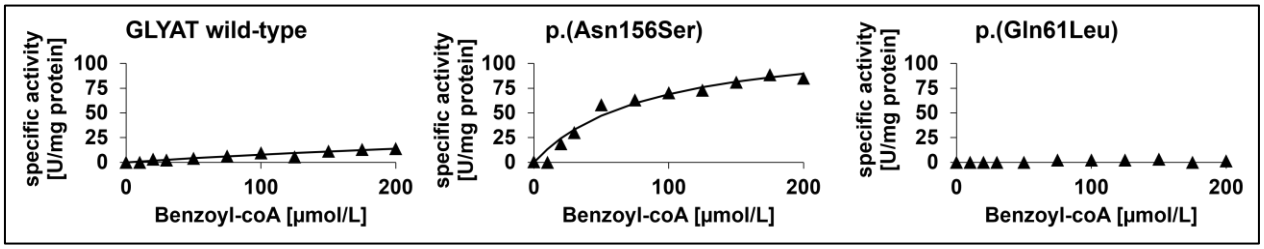


Figure 4.36: Enzyme kinetics curves of 1.5 µg purified GLYAT wild-type and sequence variants p.(Asn156Ser) and p.(Gln61Leu) six month after purification. (SDS PAGE & Western blot: Figure 4.13.) GLYAT wild-type: v_{max}: 54.5 U/mg protein, K_M: 593 µmol/L, p.(Asn156Ser): v_{max}: 128 U/mg protein, K_M: 85.4 µmol/L, p.(Gln61Leu): v_{max}: cannot be determined, K_M: cannot be determined.

While the GLYAT wild-type demonstrated a v_{max} of 54.5 U/mg protein and a K_M of 593 µmol/L the p.(Asn156Ser) variant was repeatedly more active (v_{max}: 128 U/mg protein) and showed the lower K_M (85.4 µmol/L). The p.(Gln61Leu) variant showed minor activity confirming the previous series (Figure 4.35.).
The same study was performed using all generated sequence variants of human GLYAT (Section 4.2., Figure 4.14.) in comparison to the wild-type. The whole kinetic curves were measured for the expressed (Figure 4.14.) and active GLYAT variants (Figure 4.37.). All activities below 5 U/mg were considered as inactive due to minor background activities.

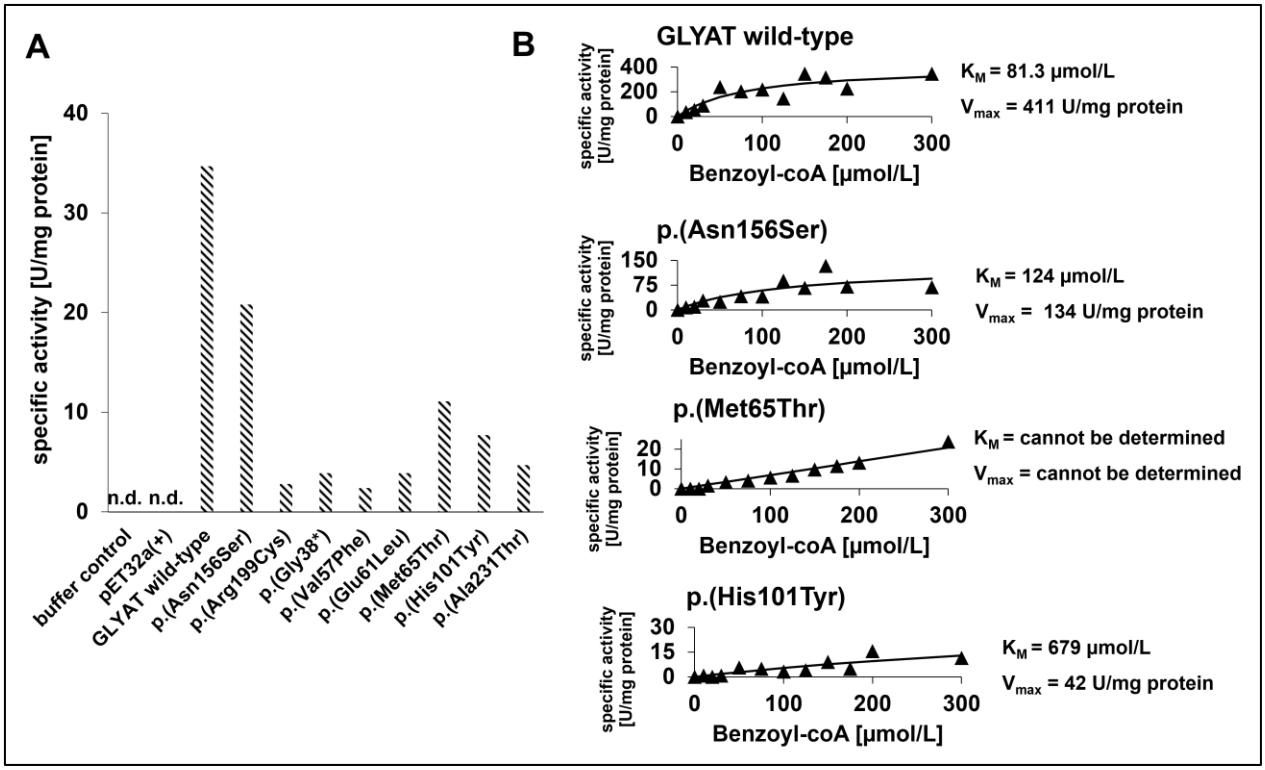


Figure 4.37: Maximum enzyme activity of the *E. coli* Origami 2(DE3) homogenates (A) and enzyme kinetics of the active GLYAT variants (B) (Western blot: Figure 4.14.). 1.5 µg GLYAT variant was measured with benzoyl-coA concentrations from 0 to 300 µmol/L.

The homogenates demonstrated that only the GLYAT wild-type and the sequence variants p.(Asn156Ser), p.(Met65Thr) and p.(His101Tyr) observed sufficiently high activities (Figure 4.37. A) for the determination of the Michaelis-Menten curves (Figure 4.37. B). While the wild-type showed a remarkable high activity of 411 U/mg protein and K_M of 81 µmol/L, the p.(Asn156Ser) demonstrated 134 U/mg protein and K_M of 127 µmol/L. That contrasts the previous data with the more active p.(Asn156Ser) variant (Figure 4.37.). For the sequence variant p.(Met65Thr) a linear correlation of the graph was observed, hence no kinetic parameters could be calculated. However, for p.(His101Tyr) a curve fitting was possible but very high K_M of 679 µmol/L indicated that the saturation of the curve was not sufficient as well for a clear Michaelis Menten course. Nonetheless, a v_{max} of 42 U/mg protein highlighted the lower activity compared to wild-type and p.(Asn156Ser) sequence variant (corroborated by the very low substrate affinity expressed by very high K_M of 679 µmol/L). However, this expression series (Figure 4.37.) confirmed the date from Figure 4.32. with markedly reduced enzyme activities of the sequence variants p.(Met65Thr), p.(His101Tyr) and p.(Ala231Thr). Due to the insufficient expression of p.(Gln61Leu) variant (Figure 4.14.) the whole procedure was performed by another series and the kinetics of the purified enzymes were determined (Figure 4.38.).

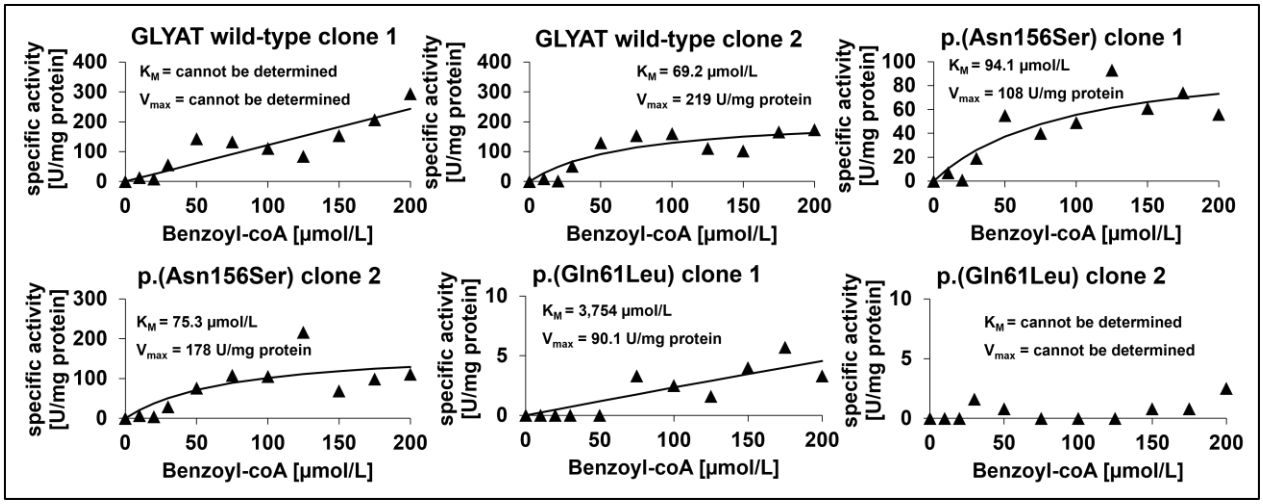


Figure 4.38: Enzyme kinetics of active GLYAT variants (1) (Western blot: Figure 4.15.). 1.5 µg GLYAT variant was measured with benzoyl-coA concentrations from 0 to 200 µmol/L. Excel Solver tool yielded K_M and v_{max} values shown in the figure.

Due to varying data it was difficult to determine kinetic parameters. However, a tendency was obvious: the 2nd clone of GLYAT wild-type showed the more reliable results with v_{max}: 219 U/mg protein and K_M: 69.2 µmol/L, whereas the p.(Gln61Leu) demonstrated merely minor activity, which increased the difficulties of kinetic descriptions. A remarkable variation between the wild-type clones was obvious, which was decreased for the p.(Asn156Ser) clones. However, the second clone of p.(Gln61Leu) did not yield in target protein expression and therefore the observed "activity" is only minor.

Figure 4.39. demonstrates another series, whereas one sample of p.(Asn156Ser) was treated with protease inhibitor cocktail (PIC).

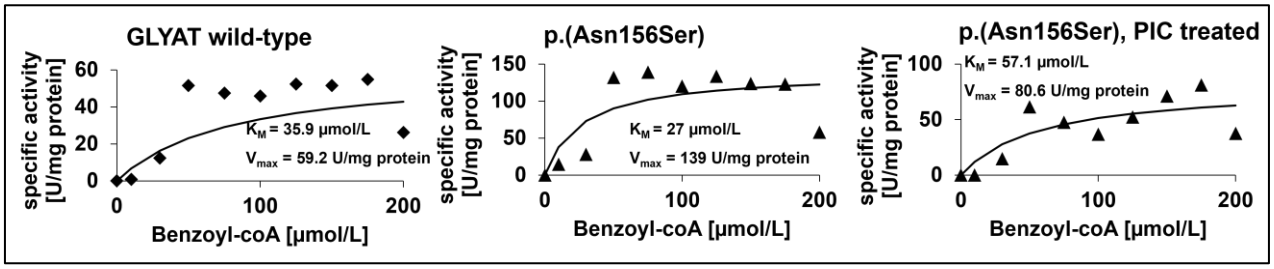


Figure 4.39: Enzyme kinetics of active GLYAT variants (2) (Western blot: Figure 4.16.). 1.5 µg GLYAT variant was measured with benzoyl-coA concentrations from 0 to 200 µmol/L. Excel Solver tool yielded K_M and v_{max} values shown in the figure.

There were no worth mentioning differences between the enzyme activities of the p.(Asn156Ser) samples without and with PIC treatment. Notwithstanding, the trend for the more active p.(Asn156Ser) variant compared to the wild-type was confirmed repeatedly.

In summary, the trend for the 3x higher active p.(Asn156Ser) sequence variant compared to GLYAT wild-type was only detectable in the series of the Figures 4.36. and 4.39. It was not

confirmed in the other series (Figures 4.37., 4.38.) where the wild-type activity resembled the one of p.(Asn156Ser). Figures 4.36. – 4.38. confirmed an absent activity of p.(Gln61Leu) variant, which resembled the findings of the FPLC samples (Section 4.6.3.). There were repetitious absent overexpressions for the p.(Gln61Leu) variant.

To summarize the inter-assay variation, the spin-column purified enzymes observed similar v_{max} values for GLYAT wild-type (186 ± 168 U/mg protein) and p.(Asn156Ser) variant (178 ± 136 U/mg protein), whereas the inter-assay variation observed high data variations (variation coefficient v_{max} wild-type: 90 %, variation coefficient v_{max} p.(Asn156Ser): 76 %). However, the p.(Gln61Leu) sequence variant demonstrated tending lower v_{max} value (47.5 U/mg protein) compared to wild-type and p.(Asn156Ser) and a quite higher K_M value tendency, which might underline lower activity and substrate affinity. The problematic overexpression of this variant complicated the comparison. The purifications were only comparable within the series not between different series due to the high inter-assay variations.

4.6.6. Enzyme activity of purified GLYATL1 wild-type and sequence variants

The first FPLC-purified GLYATL1 wild-type (Figure 4.18.) resulted in enzyme inactivity detected in the dialysis fraction. This meets the expectations because no target protein signal was observed after dialysis (Figure 4.18. B).

The activities of the spin-column purified GLYATL1 wild-type are shown by Figure 4.40.

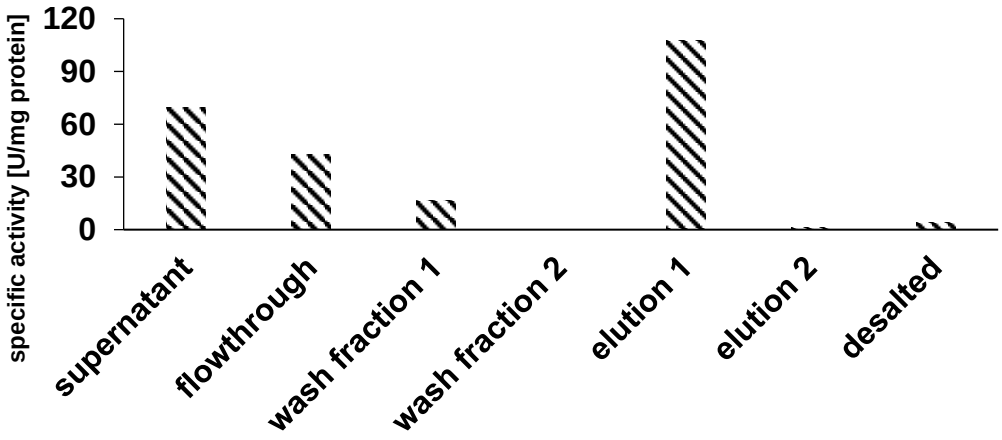


Figure 4.40: Maximum activity determination of the FPLC purified GLYATL1 wild-type using 100 µmol/L phenylacetyl-coA. Supernatant: 69.7 U/mg protein, flow-through: 43 U/mg protein, wash fraction 1: 16.5 U/mg protein, wash fraction 2: -, elution 1: 107.5 U/mg protein, elution 2: 1 U/mg protein, desalted: 4 U/mg protein.

While the activity of the supernatant fraction comprised all activity data of all present coA releasing proteins (69.7 U/mg protein), the GLYATL1 activity was specifically detected within the elution fraction 1 (107.5 U/mg protein). After desalting via dialysis merely 4 U/mg protein specific activity remained, which meets the observations from Figure 4.18.

Notwithstanding, for a better comparison of the spin-column purified GLYATL1 variants with the wild-type, the latter was analyzed by an inter-assay study with resembling kinetic properties (Figure 4.41.).

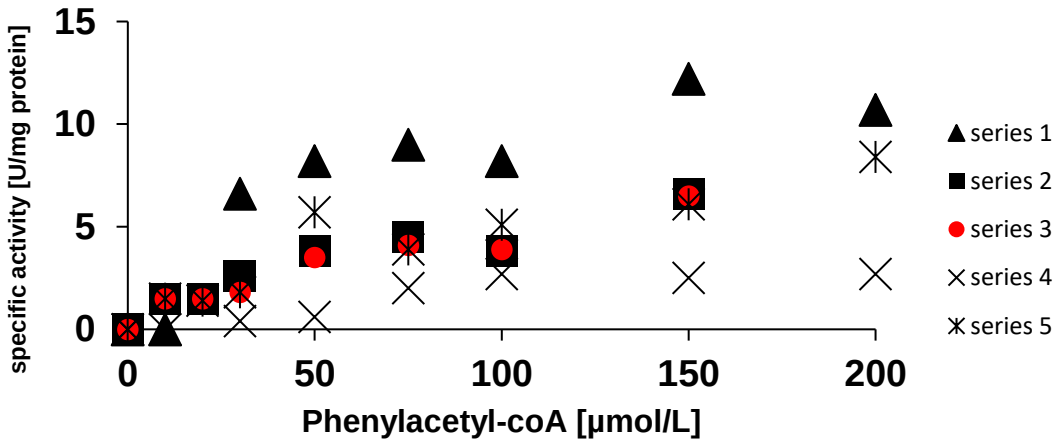


Figure 4.41.: Inter-assay enzyme activity of spin-column purified GLYATL1 wild-type across 5 different overexpression series (n=5, biological replicates). Kinetic data of the GLYATL1 wild-type was calculated as follows: v_{max}: 10.3 ± 3.8 U/mg protein (variation coefficient: 37 %), K_M: 96.3 ± 33.9 µmol/L (variation coefficient: 35 %).

The GLYATL1-wild-type was measured as control by default in the series with the sequence variants.

The first kinetic series of the GLYATL1 wild-type in comparison to the sequence variants is shown by Figure 4.42.

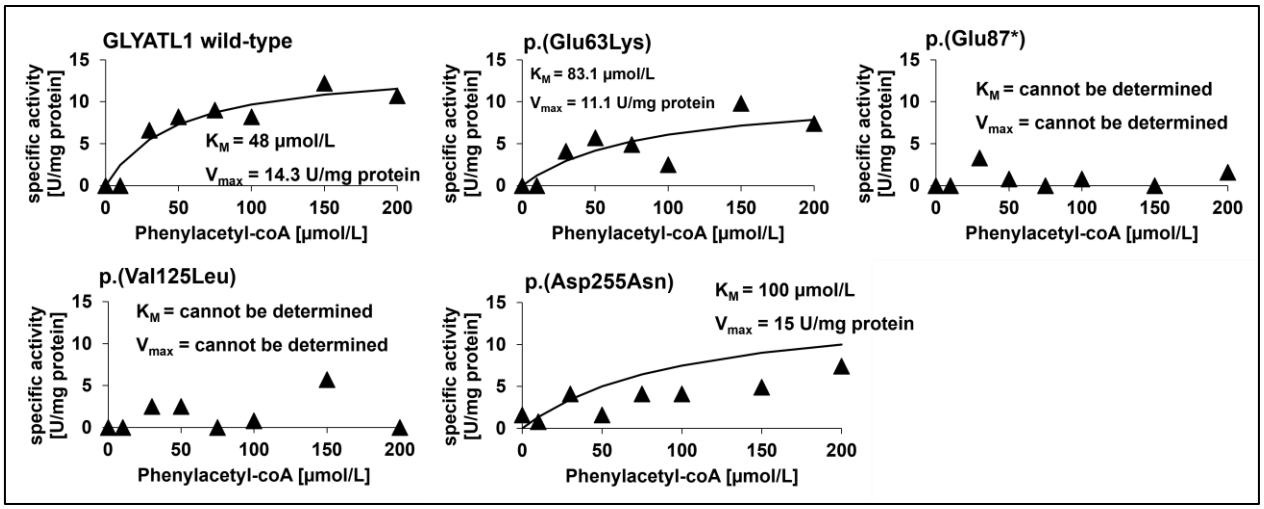


Figure 4.42: Enzyme kinetics of 1.5 µg purified GLYATL1 wild-type, p.(Glu63Lys), p.(Glu87*), p.(Val125Leu) and p.(Asp255Asn) sequence variants (Western blot: Figure 4.19.). Excel Solver tool results of kinetic parameters are shown in the figure**.**

As demonstrated thereby, the GLYATL1 wild-type showed highest activity (v_{max}: 14.3 U/mg protein) followed by the sequence variant p.(Glu63Lys), which demonstrates similar values. The

p.(Glu87*) and p.(Val125Leu) variants showed only minor activities (near 0 U/mg protein) and the Michaelis Menten kinetics for them could not be detected. The substrate turnover of the p.(Asp255Asn) variant was not saturated by usage of 200 µmol/L phenylacetyl-coA. The variant reached a v_{max} of 15 U/mg protein with a K_M of 100 µmol/L, which resembled the wild-type parameters. This variant will be of special consideration in the discussion due to unexpected, possibly sigmoidal curve progression, which was also obvious for the GLYAT wild-type (Figure 4.38.). The p.(Val125Leu) variant was repeated in another series (data not shown), which resulted in a v_{max} of 27 U/mg protein and K_M : 55.3 µmol/L.

For the purpose of reproduction, the enzyme activities were repeatedly measured (Figure 4.43.).

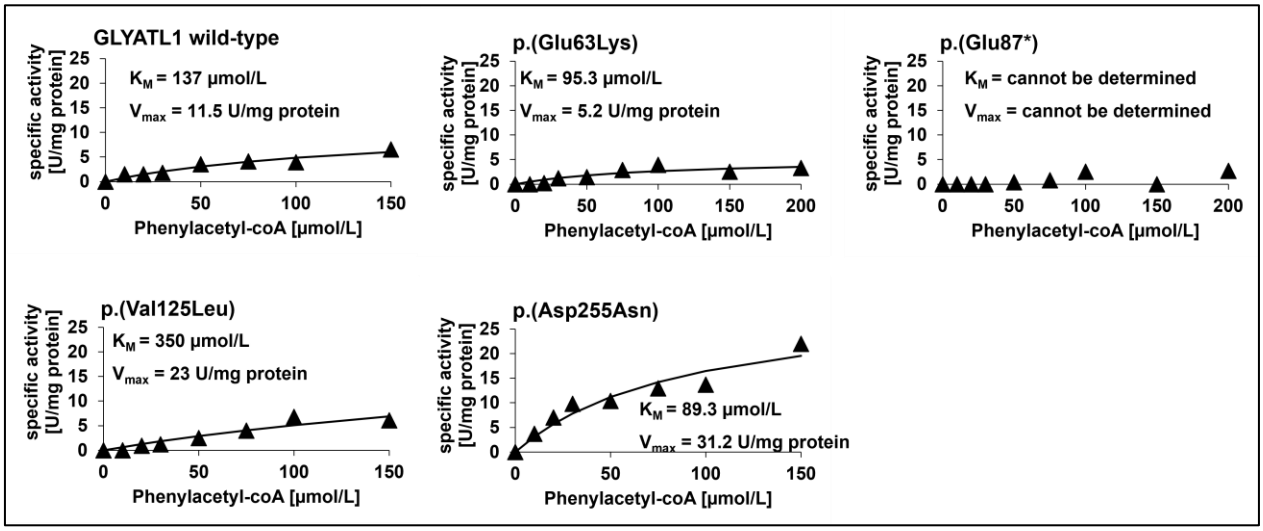


Figure 4.43: Enzyme kinetics of GLYATL1 wild-type, p.(Glu63Lys), p.(Glu87*), p.(Val125Leu) and p.(Asp255Asn) (Western blot: Figure 4.20.). Excel Solver tool results of kinetic parameters are shown in the figure**.**

The GLYATL1 wild-type and p.(Glu63Lys) activity resembled the results of the detection before (Figure 4.42.). The p.(Glu87*) variant showed low activity again activity but kinetic parameters could not be detected due to unsaturation of the curve. The p.(Val125Leu) sequence variant showed wild-type comparable activity again, whereas p.(Asp255Asn) seems to be 3x more active than wild-type.

In another series, the following kinetic parameters were detected: GLYATL1 wild-type: v_{max}: 12.1 U/mg protein, K_M: 117 µmol/L, p.(Val125Leu): v_{max}: 19.6 U/mg protein, K_M: 82.5 µmol/L, p.(Asp255Asn): v_{max}: 15.6 U/mg protein, K_M: 46.8 µmol/L (data not shown). However, in the p.(Glu63Lys) variant was not overexpressed in that series.

The unexpected enzyme activities, especially of the p.(Gly87*) (data not shown) and p.(Asp255Asn) variant, required colony PCR studies (Suppl. Fig. 59, 60), which analyzes the transformation of *E. coli* Origami 2(DE3) cells. With the help of this colony PCR, a distinction can

be made between non-specific signal noise in the activity detection (without transformation) and the specific signal of the variant (with transformation). As the appendix highlights, the transformation of *E. coli* Origami 2(DE3) with the GLYATL1 sequence variants was successful. Table 36 provides an overview on the kinetic parameters of the studies on human GLYATL1.

Table 36: Kinetic parameters (v_{max}, $K_{M)}$ of spin-column purified GLYATL1 wild-type and sequence variants (n=3)

Sequence variant	v_{max} [U/mg protein]	K_M (phenylacetyl-coA) [µmol/L]
GLYATL1 wild-type (n = 5)	10.3 ± 3.8	96.3 ± 33.9
p.(Glu63Lys)	11.1 / 5.2	83.1 / 95.3
p.(Glu87*)	cannot be determined	cannot be determined
p.(Val125Leu)	32.6 ± 19.7	153 ± 172
p.(Asp255Asn)	20.6 ± 9.2	78.7 ± 28.2

To summarize that part, the purified GLYATL1 wild-type resulted in v_{max} of 10.3 ± 3.8 U/mg protein and K_M (phenylacetyl-coA) of 96.3 ± 33.9 µmol/L across 5 different series (Figure 4.41.).

The sequence variants strongly resembled the values of the wild-type except for the stop mutant p.(Glu87*). Nonetheless, this activity assay showed very high inter-assay variation and much lower v_{max} values were observed compared to GLYAT activity assay (Section 4.6.5.). The p.(Val125Leu) variant demonstrated a very suspicious deviation in the K_M value (153 ± 172 µmol/L) across the measured series. The p.(Glu63Lys) was only expressed in two series that showed once wild-type similar (v_{max}: 11.1 U/mg protein) and once merely half of the wild-type enzyme activity (v_{max}: 5.2 U/mg protein).

4.7. Prediction of PTMs of human GLYAT, GLYATL1 and GLYATL2

The PTM patterns of human *N*-acyltransferases GLYAT, GLYATL1 and GLYATL2 were determined using online prediction tools (Section 2.10.; Suppl. Fig. 29-37). The overexpression of the enzymes in the human-derived HEK293 cells was attempted to ensure these predicted PTMs. A summary of the *O*-glycosylation, phosphorylation, and acetylation predictions for GLYAT, GLYATL1 and GLYATL2 is shown by Table 37.

Table 37: Predictions of PTMs (*O*-glycosylation, phosphorylation, acetylation) of human enzymes GLYAT, GLYATL1 and GLYATL2.

Modification / Enzyme	*O*-glycosylation	phosphorylation	acetylation
Human GLYAT	Ser^{154}	Ser^{17}, Ser^{25}, Thr^{31}, Thr^{66}, Tyr^{72}, Thr^{75}, Ser^{92}, Ser^{107}, Ser^{110}, Ser^{123}, Thr^{130}, Thr^{139}, Thr^{144}, Ser^{151}, Ser^{154}, Ser^{171}, Ser^{172}, Ser^{193}, Thr^{215}, Thr^{225}, Thr^{244}, Ser^{249}, Ser^{263}, Tyr^{267}, Ser^{268}, Ser^{276}	Lys$^{16, 20, 27, 44, 122, 125, 141, 149, 151, 159, 161, 169, 253, 256}$
Human GLYATL1	Ser^{145}, Ser^{151}, Thr^{155}, Thr^{166}	Ser^{20}, Ser^{24}, Tyr^{28}, Ser^{30}, Ser^{48}, Tyr^{52}, Thr^{65}, Ser^{70}, Tyr^{71}, Ser^{79}, Ser^{85}, Ser^{109}, Thr^{118}, Ser^{120}, Ser^{122}, Ser^{128}, Ser^{144}, Ser^{147}, Ser^{151}, Thr^{155}, Ser^{164}, Tyr^{171}, Ser^{177}, Tyr^{233}, Ser^{265}	Lys$^{9, 15, 61, 84, 121, 124, 146, 148, 170, 187, 190, 199, 253}$
Human GLYATL2	Ser^{144}, Ser^{145}	Ser^{7}, Ser^{16}, Ser^{20}, Ser^{24}, Tyr^{52}, Thr^{57}, Tyr^{71}, Thr^{72}, Thr^{74}, Ser^{91}, Ser^{96}, Thr^{119}, Ser^{120}, Ser^{144}, Ser^{145}, Ser^{164}, Ser^{188}, Tyr^{191}, Ser^{213}, Thr^{228}, Ser^{250}, Thr^{289}	Lys$^{15, 61, 80, 84, 116, 121, 139, 140, 142, 148, 184, 190, 231, 287, 292}$

Human GLYAT might possess several PTMs – while serine at position 154 is predicted to be *O*-glycosylated, several lysines should be acetylated. Numerous serines, threonines and tyrosines are phosphorylated (Table 37). Considering that, many of these modifications are located within the active center of the enzyme – amino acid 128-178 [107] – they might play essential roles in executing enzyme function.

Regarding human GLYATL1, two serines at positions 145 and 151 and threonines at positions 155 and 166 are predicted to be *O*-glycosylated. The active center was determined between amino acid 128 and 180 [107]. Hence, all predicted *O*-glycosylations fall into the active center (glutamine binding region) of the enzyme. Like the GLYAT enzyme, several serine, threonine, and tyrosine residues might be phosphorylated with prediction signals above critical threshold. Numerous lysines were predicted to be acetylated. However, 3 of them, at positions 146, 148 and 170, are located within the active center of the enzyme.

Two serines of the human GLYATL2 protein at positions 144 and 145 are possibly *O*-glycosylated, whereas predominantly serines were predicted to be phosphorylated. Moreover, several lysines might be acetylated However, a critical lysine residue at position 19 was supposed to be

deacetylated to ensure enzyme function [166]. The predictions confirmed that by absence of acetyl residue for this lysine.

4.8. Localization studies of human GLYAT and GLYATL1

4.8.1. Localization studies with overexpressed human GLYAT/L1-eGFP in HEK293 cells

The intracellular localization of human GLYAT and GLYATL1 was determined by confocal laser scanning microscopy in HEK293 cells. To achieve that, the fluorescence marker eGFP was used. HEK293 cells were transfected with CeGFP_pcDNA3.1(+) and the corresponding fusion protein vectors (Section 3.12.). After preparation of the cells for the microscopy, the images were produced (Figure 4.44., 4.45.).

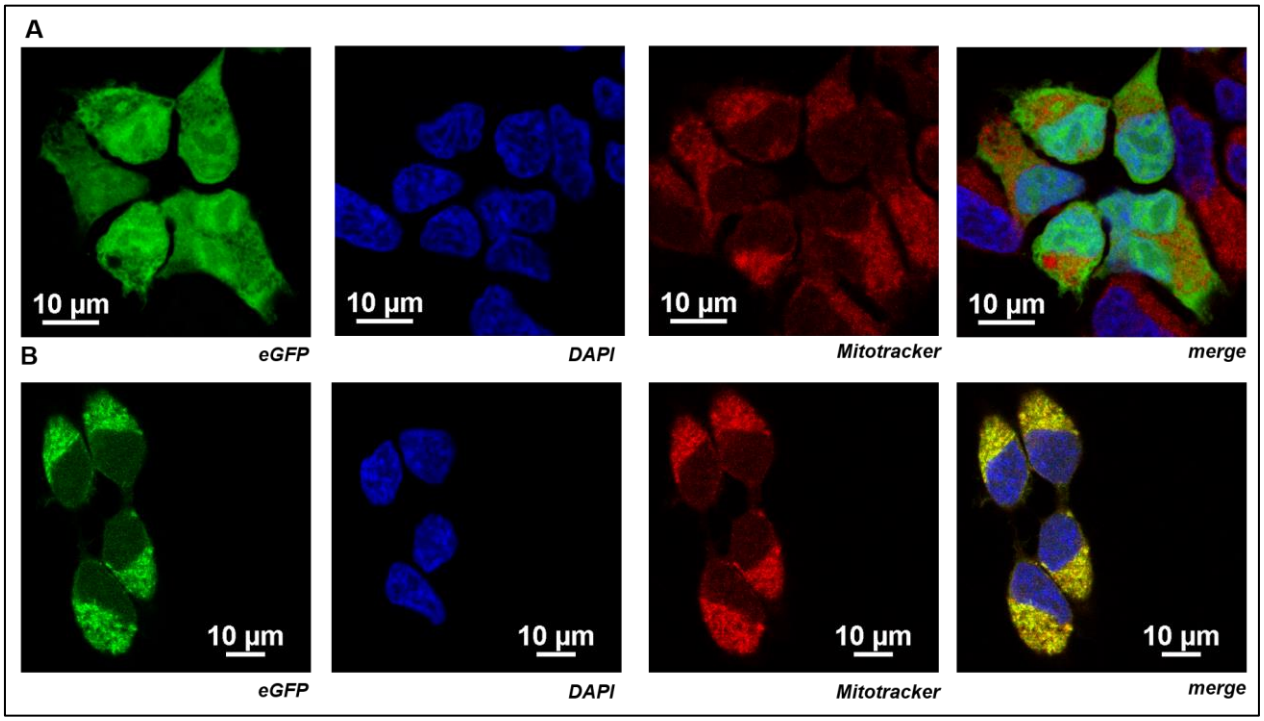


Figure 4.44: Confocal laser scanning microscopy images of stably transfected HEK293 cells with eGFP-pcDNA3.1(+) (A) and GLYAT-eGFP fusion protein (B). Cells were monitored by eGFP (488 nm), DAPI (405 nm) and Mitotracker (561 nm) filter to stain eGFP protein, nuclei and mitochondria, respectively. Mitochondria are identified as small intracellular circle structures highlighted by overlay between GFP green and Mitotracker Orange CMTMRos (Invitrogen, Thermo Fisher) signals.

The free eGFP was detected in whole cell bodies, mainly in the cytosole (Figure 4.44. A). In contrast to that, the GLYAT-eGFP was only noticed in small intracellular circle structures, which may represent the mitochondria. This was confirmed by the overlay with Mitotracker Orange CMTMRos (Figure 4.44. B, merge), which resulted in a publication [137].

In the same way, the GLYATL1 localization was determined using stably transfected HEK293 cells and confocal laser scanning microscopy (Figure 4.45.).

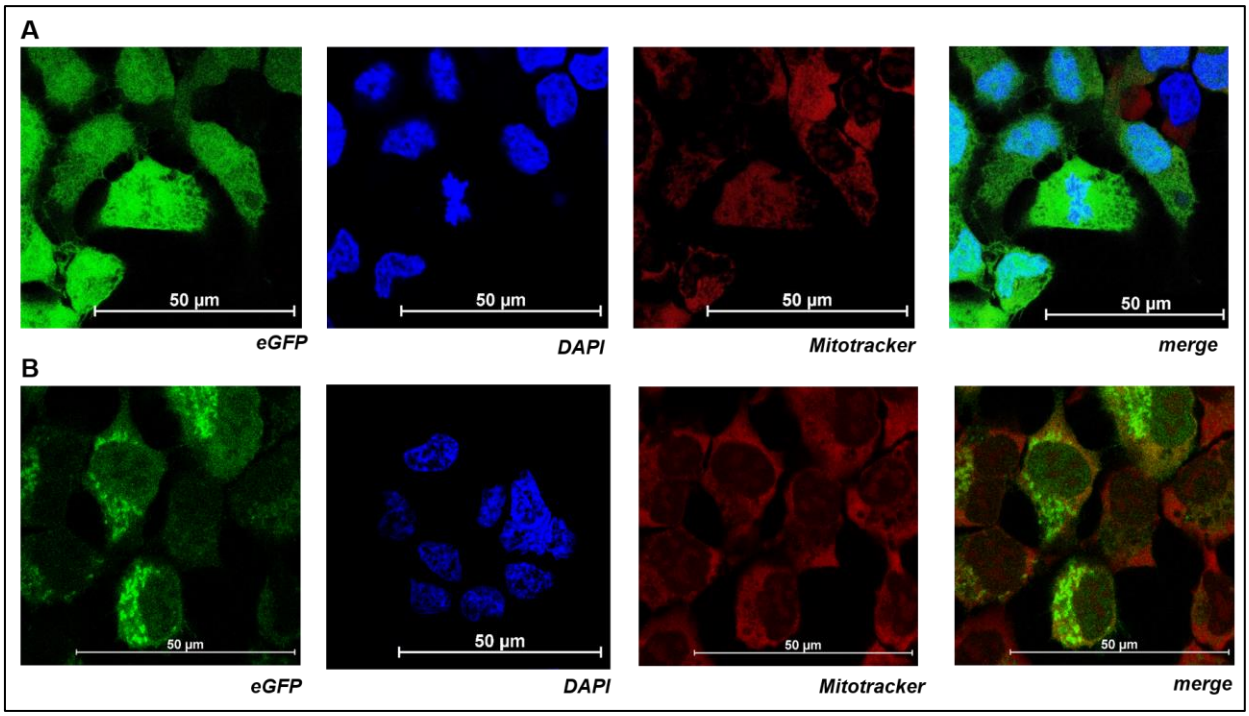


Figure 4.45: Confocal laser scanning microscopy images of stably transfected HEK293 cells with eGFP-pcDNA3.1(+) (A) and GLYATL1-eGFP fusion protein (B). Cells were monitored by eGFP (488 nm), DAPI (405 nm) and Mitotracker (561 nm) filter to stain eGFP protein, nuclei and mitochondria, respectively. Mitochondria are identified as small intracellular circle structures highlighted by overlay between GFP green and Mitotracker Orange CMTMRos (Invitrogen, Thermo Fisher) signals.

GLYATL1 seemed to be localized within mitochondria due to overlaying signals of GLYATL1-eGFP and Mitotracker signals (Figure 4.45. B, merge).

To corroborate that, the GLYAT and GLYATL1 localization was investigated further using TOMM20 antibody (Figure 4.46.-4.48.). TOMM20 represents human translocase protein of outer mitochondrial membrane encoded by the *TOMM20* gene.

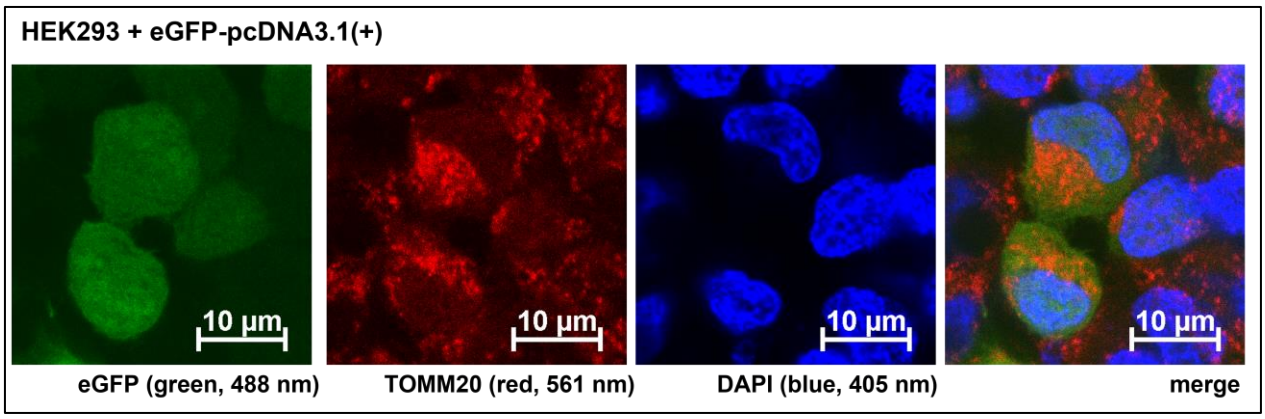


Figure 4.46: Confocal laser scanning microscopy images of stably transfected HEK293 cells with eGFP-pcDNA3.1(+) using GFP, TOMM20 and DAPI filter. Cells were monitored by eGFP (488 nm), DAPI (405 nm) and TOMM20 (561 nm) filter to stain eGFP protein, nuclei and mitochondria, respectively. Free eGFP is distributed in the whole cell body of HEK293 cells. GFP filter demonstrates equal green staining in the whole cell body, whereas TOMM20 antibody highlights the mitochondria. DAPI stains the DNA in the nuclei.

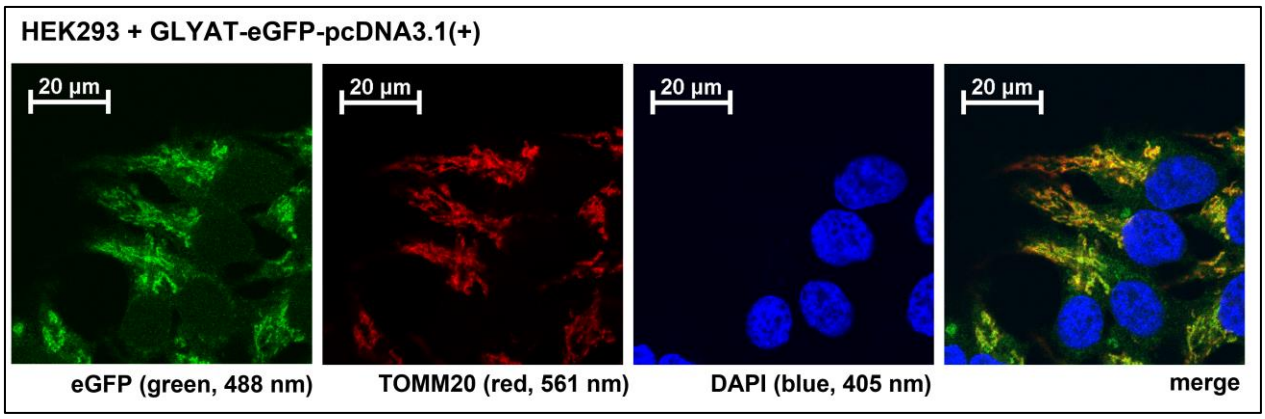


Figure 4.47: Confocal laser scanning microscopy images of stably transfected HEK293 cells with GLYAT_pcDNA3.1(+) using GFP, TOMM20 and DAPI filter. Cells were monitored by eGFP (488 nm), DAPI (405 nm) and TOMM20 (561 nm) filter to stain eGFP protein, nuclei and mitochondria, respectively Human GLYAT is localized within mitochondria of HEK293 cells. GFP filter shows clear small intracellular circle structures, which overlay with TOMM20 signals to yellow signal (merge picture). DAPI stains the DNA in the nuclei.

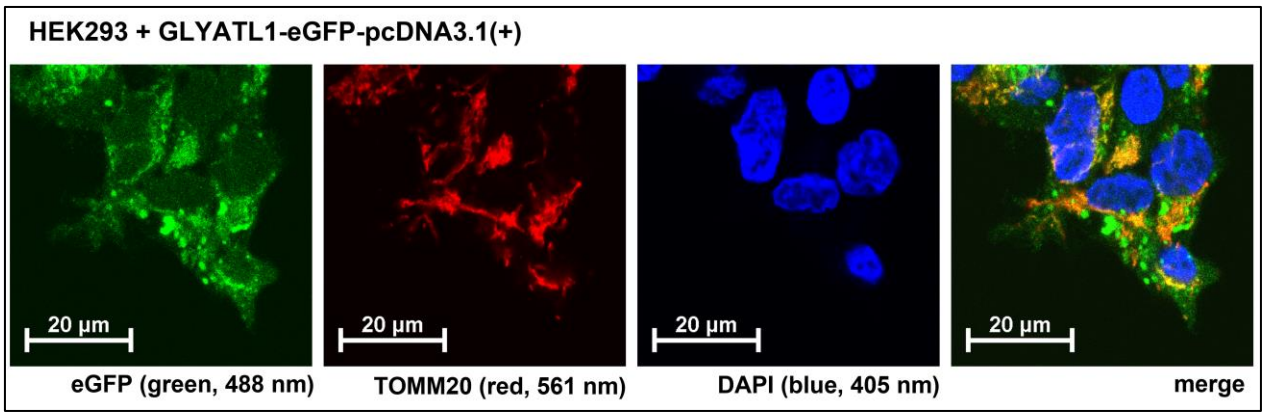


Figure 4.48: Confocal laser scanning microscopy images of stably transfected HEK293 cells with GLYATL1_pcDNA3.1(+) using GFP, TOMM20 and DAPI filter. Cells were monitored by eGFP (488 nm), DAPI (405 nm) and TOMM20 (561 nm) filter to stain eGFP protein, nuclei and mitochondria, respectively Human GLYATL1 is localized within mitochondria of HEK293 cells. GFP filter shows clear small intracellular circle structures, which overlay with TOMM20 signals to yellow signal (merge picture). Similar to GLYAT-eGFP images (Figure 4.47.) the small intracellular structures are mainly located in the outside of the cell. DAPI stains the DNA in the nuclei.

The free eGFP was located in the whole cell body, whereas GLYAT and GLYATL1 seemed to appear as small intracellular circle structures in the same areas as TOMM20 (Figures 4.47., 48.). This is concluded due to yellow overlaying signals in the merge images. The free eGFP was distributed in the whole cell bodies (Figure 4.46.). Therefore, the mitochondrial localization of human GLYAT and GLYATL1 was detected confirming previous studies with the Mitotracker staining (Figures 4.44., 4.45.).

Immunoblot studies with cell homogenates of HEK293 transfected with eGFP and GLYAT-eGFP were performed (Figure 4.49.) to corroborate the microscopy results (Figure 4.44.-.4.48.).

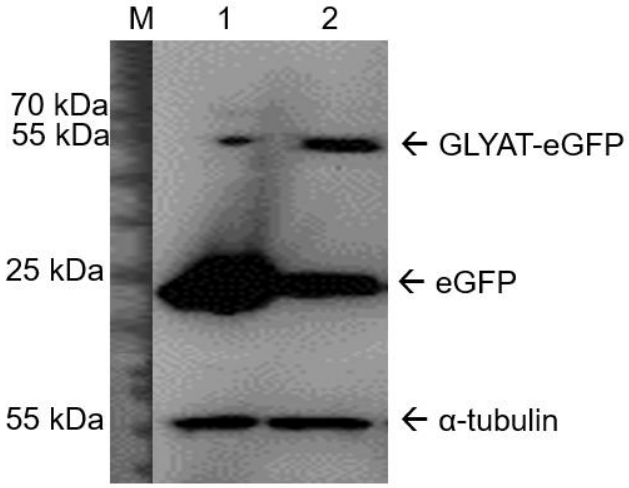


Figure 4.49: Western blot of HEK293 total protein supernatants transfected with eGFP_pcDNA3.1(+) (1) and GLYAT-eGFP_pcDNA3.1(+) (2). M: PageRuler™ Prestained Protein Ladder. 20 µg homogenate protein were loaded on the gel; anti-GFP GTX113617 (Genetex) was applied in 1:5,000 dilution to detect 60 kDa GLYAT-eGFP and 26 kDa free eGFP. Anti-tubulin ab7291 (Abcam), 1:5,000 was applied after rehybridization step to detect 50 kDa α-tubulin.

The GLYAT-eGFP was detected at 60 kDa, whereas free GFP was mainly detected in HEK293 cell supernatant transfected with eGFP only. However, the same Western blot was performed with GLYATL1-eGFP, which did not appear on the blot (data not shown). Hence, the confocal microscopy studies of transfected HEK293 cells were verified merely for the GLYAT-eGFP fusion protein.

4.8.2. Localization studies with native GLYAT/GLYATL1 in human liver samples

Followed by the studies with the overexpressed GLYAT-eGFP and GLYATL1-eGFP proteins, the native human GLYAT and GLYATL1 were analyzed by immunoblotting of human liver fractions (Figure 4.50.).

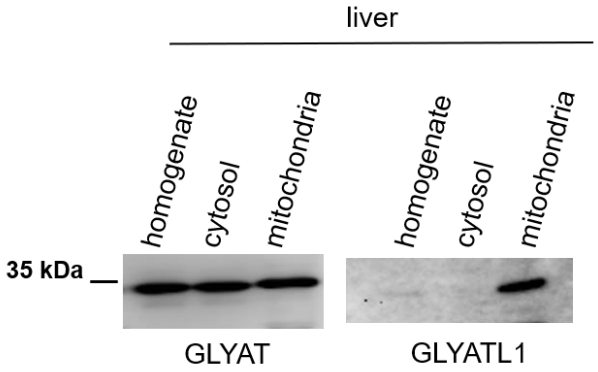


Figure 4.50: Western blot of 30 µg total protein supernatants of human liver homogenate, cytosol and mitochondria. First antibodies were anti-GLYAT PA5-48504, 1:1,000 and anti-GLYATL1, GTX106956, 1:1,000.

The human GLYAT was detected in all liver fractions with no significant differences. This was also confirmed by GLYAT activity measurement (Figure 4.51.).

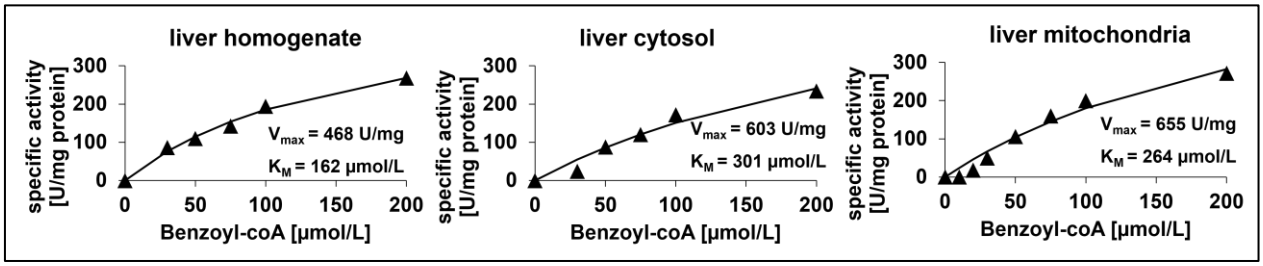


Figure 4.51: Enzyme kinetics of human liver fractions (homogenate, cytosol and mitochondria).

The GLYAT activity measurements (Figure 4.51.) showed no severe differences in GLYAT activity, which correlates with the Western blot (Figure 4.50.). However, the blot and the activity measurements were normalized to the total protein amounts without considering the fact that the homogenate fraction is higher protein-enriched than the cytosol fraction. Conversely, GLYATL1 was mainly detected in the mitochondrial fraction (Figure 4.50.). Nonetheless, a very low amount of GLYATL1 protein was furthermore detected in the homogenate fraction where mitochondrial fraction was isolated from.

In summary, the GLYAT and GLYATL1 enzymes are clearly located within mitochondria since confocal laser scanning microscopy and Western blot studies revealed that (Section 4.8.). A supporting Western blot study with human liver fractions was not convincing for GLYAT as the protein (Figure 4.50.) and the activity (Figure 4.51.) were detected in all liver fractions. However, that might be related to the applied sample amount, which was related to the total protein amount without fraction specificities. Notwithstanding, GLYATL1 could be clearly detected in the mitochondrial liver fraction.

The mitochondrial localization was predicted by the TargetP-2.0 tool (Section 2.10.) which resulted in 18 %, 8 % and 2 % likelihood for a *N*-terminal mitochondrial transfer peptide of GLYAT, GLYATL1 and GLYATL2, respectively (Suppl. Fig. 38-40). Due to the observations of the present study (Figure 4.44. – 4.51.), the further work in HEK293 cells were performed using the mitochondria-enriched supernatants after cell disruption.

4.9. Overexpression of recombinant GLYAT in human-derived cells and optimization approaches

The human GLYAT was overexpressed in HEK293 cells accordingly to compare with bacterial data. The bacterial system *E. coli* Origami 2(DE3) may have several limitations in expression of more complex proteins due to the absent ability for authentic PTMs [37].

Therefore, the HEK293 cells were transfected using *Xfect* transfection reagent and the ab86102 anti-GLYAT antibody from Abcam (Section 4.9.1.).

4.9.1. First overexpression tests in HEK293, HeLa and HepG2 cells

The results of the first overexpression series of human GLYAT in HEK293 cells are demonstrated (Figure 4.52.).

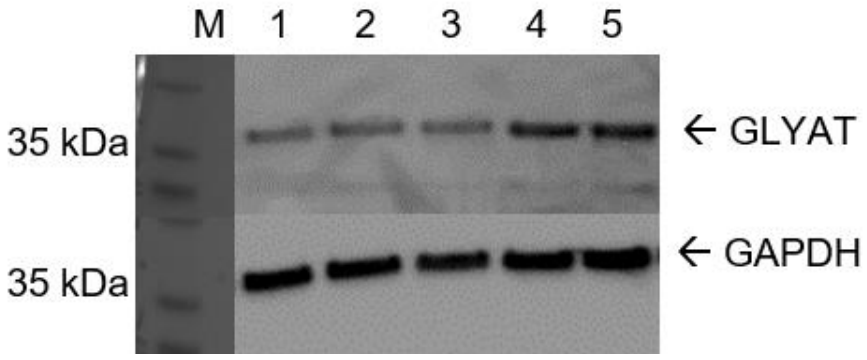


Figure 4.52.: Western blot (SDS gel) of 10 µg total protein supernatants of stably transfected HEK293 cells with pcDNA3.1(+) and GLYAT_pcDNA3.1(+) (2 transfections using *Xfect* reagent). M: PageRuler™ Prestained Protein Ladder, 1: HEK293, native (untransfected), 2: HEK293 + pcDNA3.1(+) (transfection 1), 3: HEK293 + GLYAT wild-type (transfection 1), 4: HEK293 + pcDNA3.1(+) (transfection 2), 5: HEK293 + GLYAT wild-type (transfection 2). First antibodies were anti-GLYAT ab86102 (Abcam), 1 µg/ml and anti-GAPDH GTX100118 (Genetex), 1:5,000.

The stably transfected HEK293 cells with pcDNA3.1(+) and GLYAT-pcDNA3.1(+) (samples 2-5) did not show an overexpression of GLYAT target protein at 34 kDa by using *Xfect* transfection reagent. No remarkable target overexpressed protein was detected in several other series (data not shown). It was also assumed that the protein may have been degraded in the stably transfected HEK293 cells and therefore not visible. However, under these conditions (*Xfect* reagent, Abcam antibody) the same results were observed for the transient transfection.

The cells were disrupted using the HEK293 disruption buffer (Section 2.3.). Another buffer system (RIPA buffer) was checked to compare the efficacy between the used buffer and a commercial RIPA buffer. No differences were detected by using the latter to disrupt transient transfected cells (data not shown).

Following the search of the recombinant GLYAT in HEK293 cells, the protein was suggested to be within in the unsoluble cell debris after cell disruption. To investigate that, the corresponding pellet was compared with the soluble supernatant on a Western blot. However, there was no target protein signal within the unsoluble pellet fraction (data not shown).

In another series performed with stably transfected HEK293 cells the free eGFP and the fusion proteins were analyzed (Figure 4.53.).

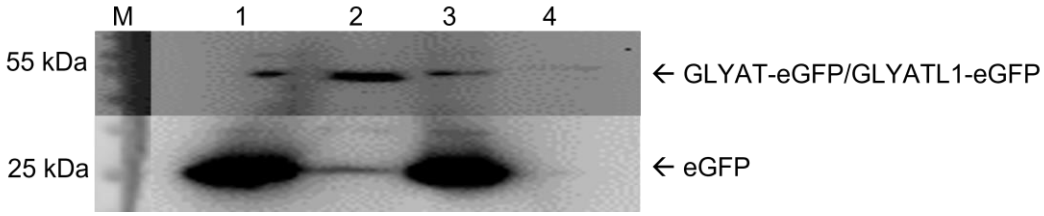


Figure 4.53: Western blot of stably transfected HEK293 cell homogenates with CeGFP-pcDNA3.1(+) and GLYAT-eGFP and GLYATL1-eGFP. M: PageRuler™ Prestained Protein Ladder, 1: HEK293 + eGFP-pcDNA3.1(+), 2: HEK293 + GLYAT-eGFP, 3: HEK293 + eGFP-pcDNA3.1(+), 4: HEK293 + GLYATL1-eGFP. 20 µg total protein homogenate were loaded. First antibody was anti-GFP 113617 (Genetex), 1:5,000.

While the free eGFP (26.9 kDa) was clearly visible in the transfected HEK293 cells (sample 1, 3), the GLYAT-eGFP is detected in the corresponding transfection cell batch (sample 2). However, the GLYATL1-eGFP could not be detected by eGFP antibody, which was repeatedly shown in another series (data not shown).

In summary., neither recombinant GLYAT nor GLYAT-eGFP fusion protein was detected in any of the used systems. However, the stably transfected HEK293 cells observed target protein signals in the confocal laser scanning microscopy studies (Section 4.8.).

Due to the missing overexpression of recombinant GLYAT in the transient (data not shown) and stably transfected HEK293 cells (Figure 4.52.), several approaches for improving were perfomed

(Section 4.9.2. – 4.9.4.). Notwithstanding, these experiments were not performed with GLYAT-eGFP since the priority was on recombinant GLYAT and the planned activity measurements.

The recombinant human GLYAT in HEK293 cells was assumed to be spontaneously degraded by the proteasome directly after production. Hence, the tripeptide aldehyde MG-132, an effective membrane-permeable proteasome inhibitor [114], was used to inhibit the GLYAT degradation via the proteasome. MG-132 was used for the treatment of transfected HEK293 cells for 4 h under cell culture conditions with 10 µmol/L MG-132. The results of these Western blots revealed no noticeable differences between DMSO and MG-132 treated cells (Suppl. Fig. 42). There was no obvious effect, which suggested a degradation of GLYAT without involvement of the proteasome in HEK293 or HEK293T cells.

The 5´-azacytidine is an effective *CMV* promotor methylation inhibitor [29], which enhances the *CMV* driven protein overexpression. Due to the absence of recombinant GLYAT overexpression in HEK293 and HeLa cells, it was used to turn off the promotor methylation, thus initiating *GLYAT* overexpression. *Xfect* transfected HeLa cells were treated for 2 days with 3 µmol/L of 5´-azacytidine. As another step, the *Xfect* transfected HeLa and HEK293 cells were treated with 50 µmol/L 5´-azacytidine in addition. The 5´-azacytidine did not enhance the GLYAT overexpression: neither in HeLa cells nor in HEK293 cells (Suppl. Fig. 41).

The linearization of pcDNA3.1(+) vector may enhance the target protein expression [204] - it decreases the likelihood of vector integration in a way disrupting the gene of interest or other expression required features. Hence, the linearization with the *NheI* enzyme should “save” genes of interest and thus optimizes the mammalian overexpression. *NheI* was selected due to the cost effiency and the presence of an intersection in the *MCS* of the vector. However, the protection effect of *NheI* usage was not detected by Western blot (data not shown), which demonstrated that linearization was not enhancing mammalian overexpression of GLYAT.

4.9.2. Antibody comparison

The accuracy of GLYAT antibody ab86102 (Abcam) and PA5-48504 antibody (Thermo Fisher) for the detection in transfected HEK293T cells was checked (Figure 4.54.). The HEK293T cells stably express the SV40 large T antigen, which can bind to SV40 enhancers of expression vectors (*CMV* of the pcDNA3.1(+) vector) to increase target protein production [106]. For this reason, they were preferred here to the HEK293 cells.

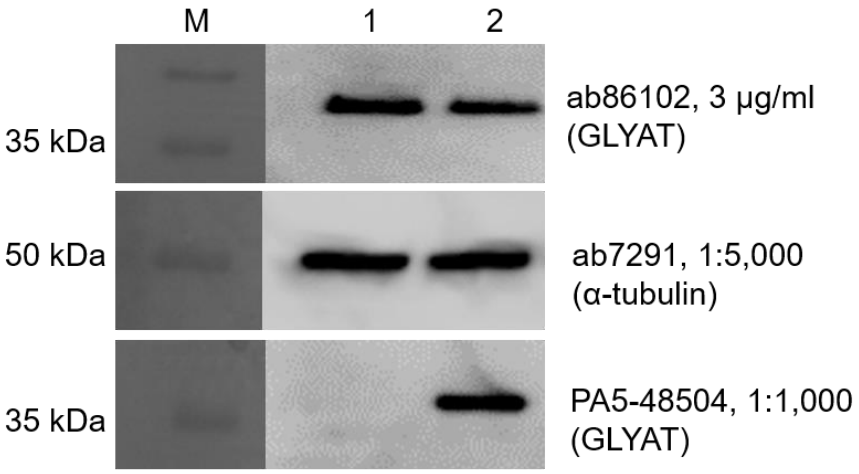


Figure 4.54: Western blot of total protein supernatants of GLYAT transfected HEK293T cells. 1: HEK293T + pcDNA3.1(+), 2: HEK293T + GLYAT wild-type. 30 µg total protein homogenate were loaded and analyzed on Western Blot. First GLYAT antibody ab86102 (Abcam) was applied (3 µg/µL). As loading control anti-tubulin ab7291, 1:5,000 (Abcam) was used

As confirmed by this Western blot, the recombinant GLYAT detection in transfected HEK293T cells is only sufficient when using PA5-48504 anti-GLYAT (Thermo Fisher). The GLYAT antibody from Abcam does not seem to work, because only unspecific signals with the same intensities for pcDNA3.1(+) and GLYAT wild-type were observed. The same observation – a clear, specific signal for recombinant GLYAT – was later found for HEK293 cells accordingly (data not shown). HEK293 cells were chosen as the main working system for future experiments (Section 4.10.) since the former experiments were performed with them (Section 4.9.1.) and the comparability is thus more sufficient.

4.9.3. Comparison of *Xfect* and *Lipofectamine 3000* transfections

The transfection reagents *Xfect* and *Lipofectamine 3000* were compared to check and to enhance GLYAT transfection efficiency. 4 independent transfection series of HEK293 and HEK293T cells using *Xfect* and *Lipofectamine 3000* were analyzed concerning expression (Figure 4.55.) and enzyme activities (Figure 4.56.).

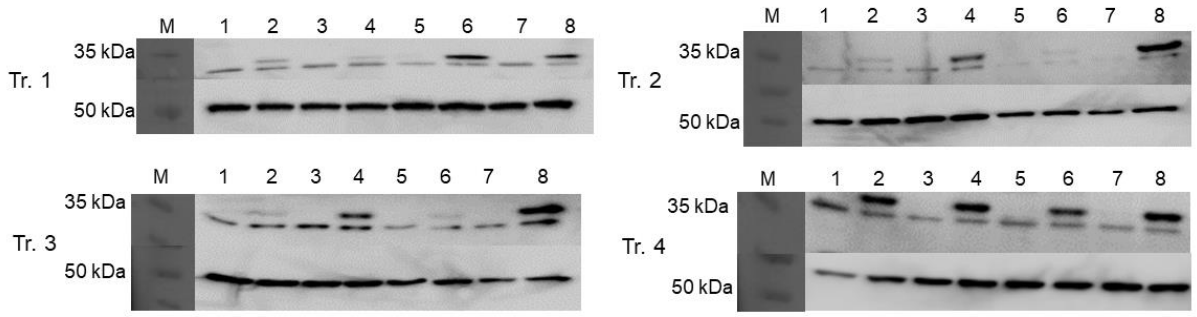


Figure 4.55: Western blots of total protein supernatants of HEK293 and HEK293T transfection series using *Xfect* and *Lipofectamine 3000*. Upper part of each Western blot: GLYAT detection, lower part of each Western blot: α-tubulin detection. M: PageRuler™ Prestained Protein Ladder, 1: HEK293 + pcDNA3.1(+) *Xfect*, 2: HEK293 + GLYAT wild-type *Xfect*, 3: HEK293 + pcDNA3.1(+) *Lipofectamine*, 4: HEK293 + GLYAT wild-type, *Lipofectamine*, 5: HEK293T + pcDNA3.1(+) *Xfect*, 6: HEK293T + GLYAT wild-type *Xfect*, 7: HEK293T + pcDNA3.1(+) Lipofectamine, 8: HEK293T + GLYAT wild-type, *Lipofectamine*. First antibodies were anti-GLYAT PA5-48504, 1:1,000 and anti-tubulin ab7291, 1:5,000 (Tr.= transfection).

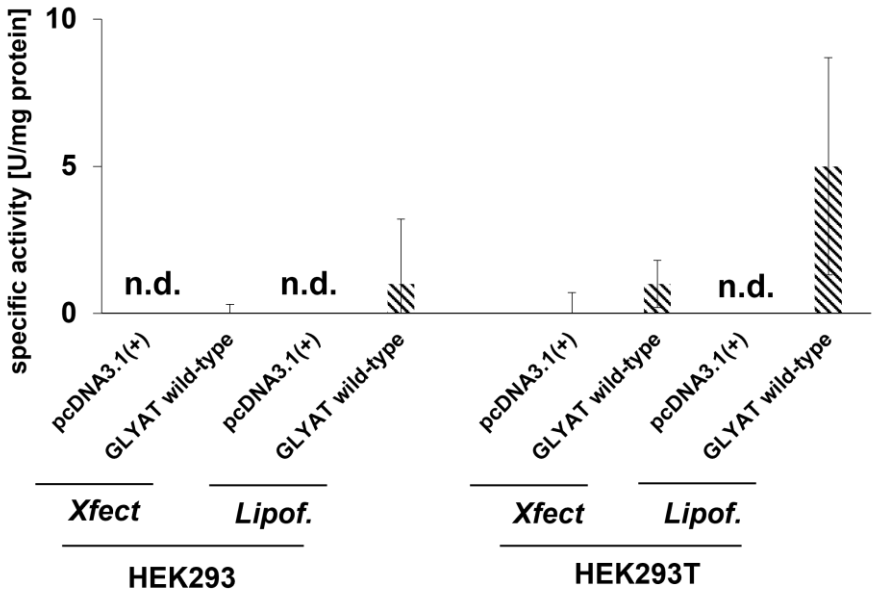


Figure 4.56: Specific enzyme activities (200 µmol/L benzoyl-coA) of recombinant GLYAT detected in respectively 2 series of transiently transfected HEK293 and HEK293T cells. Series 1-4: HEK293 + pcDNA3.1(+), Xfect: 0 + 0 U/mg protein, HEK293 + GLYAT wild-type, Xfect: 0 + 0.3 U/mg protein, HEK293 + pcDNA3.1(+), Lipofectamine: 0 + 0 U/mg protein, HEK293 + GLYAT wild-type, Lipofectamine: 1 + 2.2 U/mg protein, HEK293T + pcDNA3.1(+), Xfect: 0 + 0.7 U/mg protein, HEK293T + GLYAT wild-type, Xfect: 1 + 0.8 U/mg protein, HEK293T + pcDNA3.1(+), Lipofectamine: 0 + 0 U/mg protein, HEK293T + GLYAT wild-type, Lipofectamine: 5 + 3.7 U/mg protein. (n = 4, biological replicates of 4 transfection series).

There were no severe differences between transfected HEK293 and HEK293T cells with *Lipofectamine 3000* or *Xfect* (Figure 4.55., 4.56.). However, the transfected HEK293 cells revealed lower GLYAT target protein signals in the transfection series 1 and 2 compared to HEK293T (Figure 4.55.). That observation was confirmed in the transfection series 3 and 4. The enzyme activities (Figure 4.56.) correlated with the Western blots (Figure 4.55.). The highest enzyme

activity was observed for the HEK293T cells transfected with *Lipofectamine 3000*, but the high standard deviation relativizes that. Nonetheless, a slight trend for the *Lipofectamine 3000* reagent was obvious, which was confirmed by a parallel expression study of free eGFP in HEK293 cells (data not shown).

If not otherwise noted, the Western blot studies in the following Sections 4.10.- 4.11. are based on *Lipofectamine 3000* transfected HEK293 cells.

4.9.4. Influences on GLYAT activity in HEK293 homogenates

Due to the absent GLYAT activity of the transfected HEK293 cells (data not shown), the present section elucidates possible reasons. The purified and active GLYAT enzyme (Section 4.5., Figure 4.10.) was treated with the used buffers and homogenates for investigation of buffer (component) or homogenate influences on GLYAT activity (Figure 4.57.).

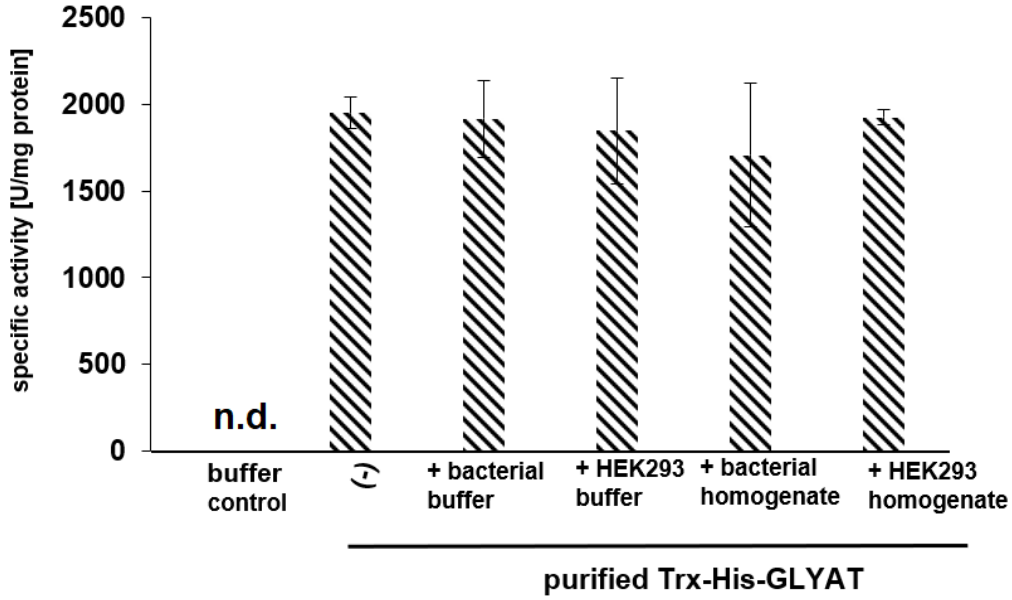


Figure 4.57: Maximum activities of purified GLYAT enzyme without additives, with *E. coli* Origami 2(DE3) disruption buffer, HEK293 disruption buffer, *E. coli* Origami 2(DE3) homogenate and HEK293 homogenate (n=3). Buffer control without purified GLYAT demonstrated no activity; activities of treated GLYAT proteins were as follows: (-): purified GLYAT: 1951 ± 92 U/mg protein, purified GLYAT (+ bacterial disruption buffer): 1916 ± 221 U/mg protein; purified GLYAT (+ HEK293 disruption buffer): 1846 ± 302 U/mg protein; purified GLYAT (+ bacterial homogenate): 1707 ± 416 U/mg protein; purified GLYAT (+ HEK293 homogenate): 1925 ± 45 U/mg protein. Variation coefficients are as follows: untreated GLYAT: 5 %, + bacterial buffer treated: 12 %, + HEK293 buffer treated: 16 %, + bacterial homogenate: 24 %, + HEK293 homogenate: 2 %.

The specific activity of the purified GLYAT enzyme indicated no remarkable change when *E. coli* or HEK293 disruption buffer was added. The same was obvious for added bacterial or HEK293 homogenate (Figure 4.57.). The activity variations were within the usual assay variation of 30 % (Section 4.6., Figure 4.23.). The standard deviation was higher compared to the others for the added bacterial homogenate (here: 416 U/mg protein, variation coefficient: 24 %), most likely due

to the high background signals of bacterial thioesterases [16]. Low variation coefficients confirm the inefficient impairment of buffers and homogenates on GLYAT activity.

4.10. Overexpression and enzyme activity assays of GLYAT in HEK293 cells

The human GLYAT wild-type was overexpressed in parallel with the selected sequence variants (Section 4.2.) after establishing the most suitable conditions for eukaryotic overexpression (Section 4.9.). The Figures 4.58. – 4.61. demonstrate the results of enzyme activities and Western blot findings.

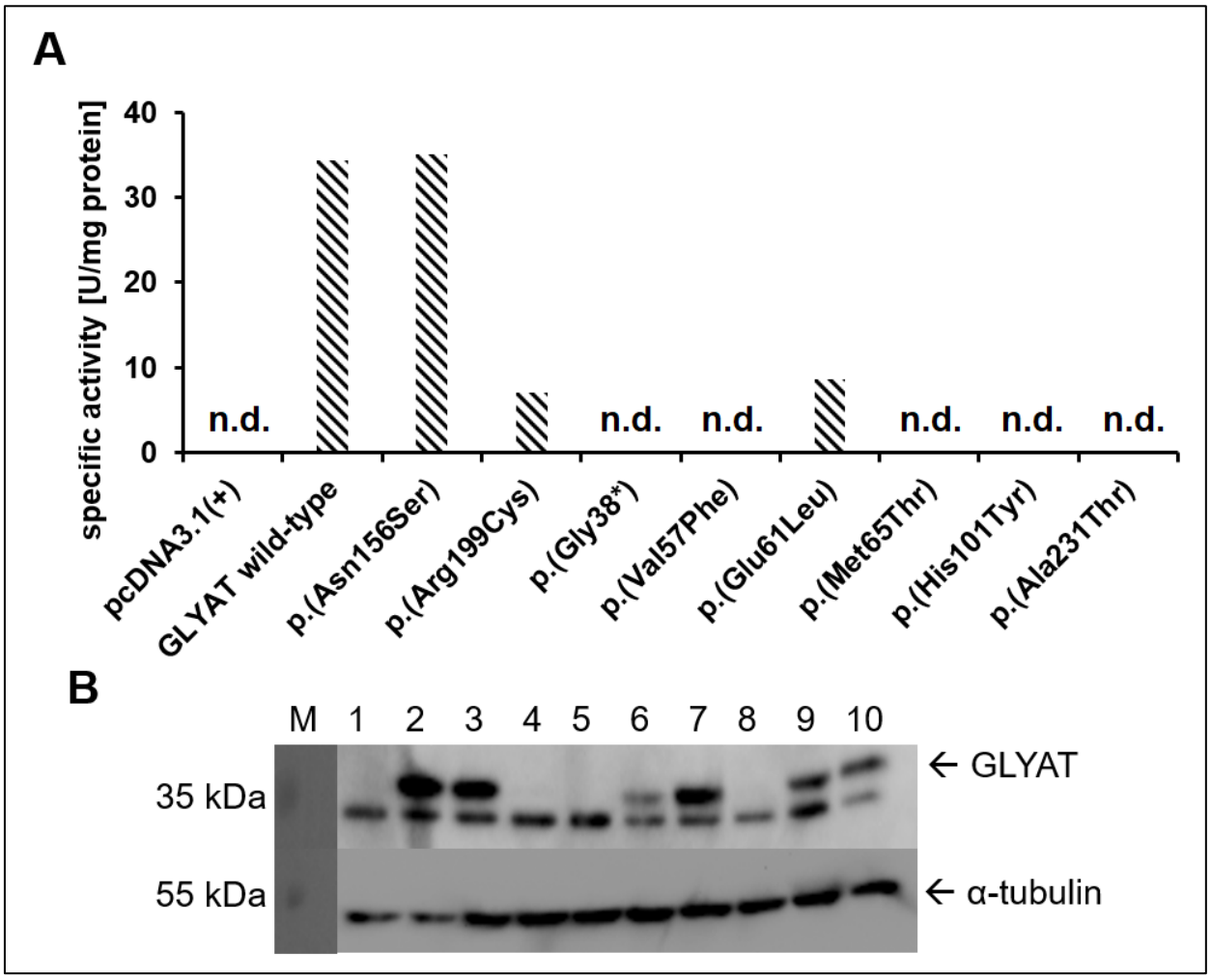


Figure 4.58: Maximum GLYAT activity of transient transfected HEK293 cell homogenates using 200 µmol/L benzoyl-coA (A) and Western blot (12 % SDS gel) of 15 µg total protein from homogenate supernatants of transfected HEK293 cell homogenates (B). 1: pcDNA3.1(+), 2: GLYAT wild-type, 3: GLYAT p.(Asn156Ser), 4: GLYAT p.(Arg199Cys), 5: GLYAT p.(Gly38*), 6: GLYAT p.(Val57Phe), 7: GLYAT p.(Glu61Leu), 8: GLYAT p.(Met65Thr), 9: GLYAT p.(His101Tyr), 10: GLYAT p.(Ala231Thr). Anti-GLYAT PA5-48504, 1:1,000 and anti-tubulin ab7291, 1:5,000 were used as first antibodies.

As shown thereby, the enzyme activities of human GLYAT wild-type were determined using a benzoyl-coA concentration of 200 µmol/L to detect maximum specific activities.

No protein bands were observed for the GLYAT sequence variants p.(Arg199Cys), p.(Gly38*) and p.(Met65Thr), which was demonstrated in additional series as well (data not shown). To clarify the absent overexpression of the named sequence variants, the gDNA of the transfected cells were analyzed (Suppl. Fig. 43-45). Furthermore, mRNA analyses (Suppl. Fig. 47-53) should check whether the problem is on transcriptional level. The level of inefficient overexpression – transfectional, transcriptional or translational – should be identified by these studies.

While GLYAT wild-type and p.(Asn156Ser) variant showed similar activities, the p.(Arg199Cys) and p.(Glu61Leu) variant revealed merely 20 % of wild-type activity. The measured activity is invalid for p.(Arg199Cys) and p.(Met65Thr) due to the absence of the proteins (Figure 4.58. B). All the other sequence variants were completely inactive (in other series as well, data not shown), which is of special interest for the overexpressed variants p.(Val57Phe), p.(His101Tyr) and p.(Ala231Thr). This study confirms the bacterial data (section 4.6.) where wild-type and p.(Asn156Ser) variant demonstrated similar activity with slightly larger activity of the mutant. The p.(Gln61Leu) variant showed only small proportion of the wild-type activity, which resembles the bacterial data.

The activities and protein overexpression of another series are shown by Figure 4.59.

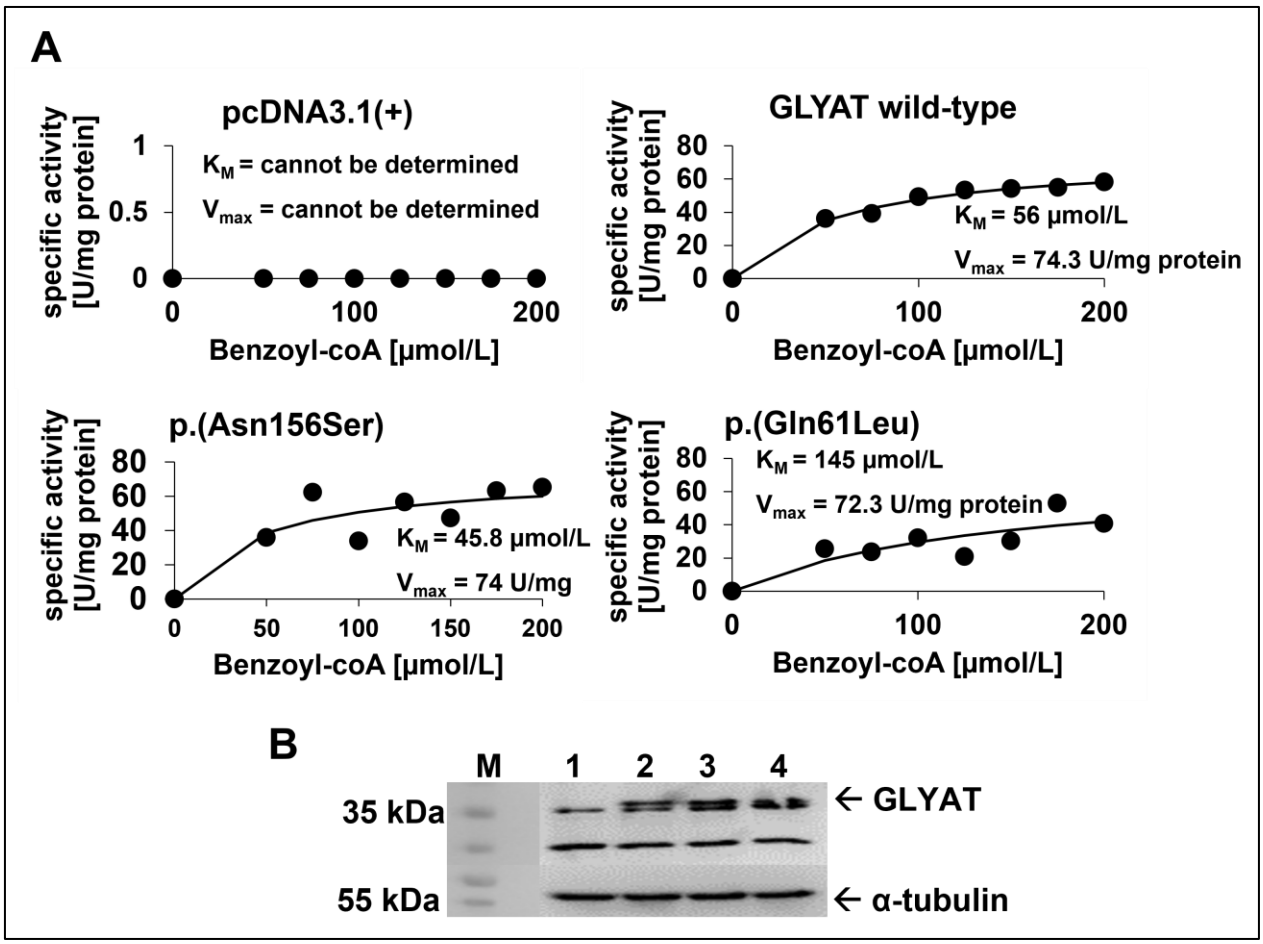


Figure 4.59: Enzyme kinetics of total protein supernatants of transiently transfected HEK293 cells (A) and corresponding Western blot (12 % SDS gel) of 15 µg supernatant (B). B: Anti-GLYAT PA5-48504, 1:1,000 and anti-tubulin ab7291, 1:5,000 were used as first antibodies.

The kinetic analyses of GLYAT wild-type and sequence variants resulted in similar v_{max} values for all three enzymes. However, the p.(Gln61Leu) sequence variant yielded slightly lower activity. At least the K_M calculation via Excel Solver tool for p.(Gln61Leu) showed 145 µmol/L, the highest obtained value indicating lowest substrate affinity. Nevertheless, the v_{max} data remained the same with lower value for p.(Gln61Leu) but K_M showed no prominent differences.

Figure 4.60. summarizes the results of a repdroducing HEK293 series

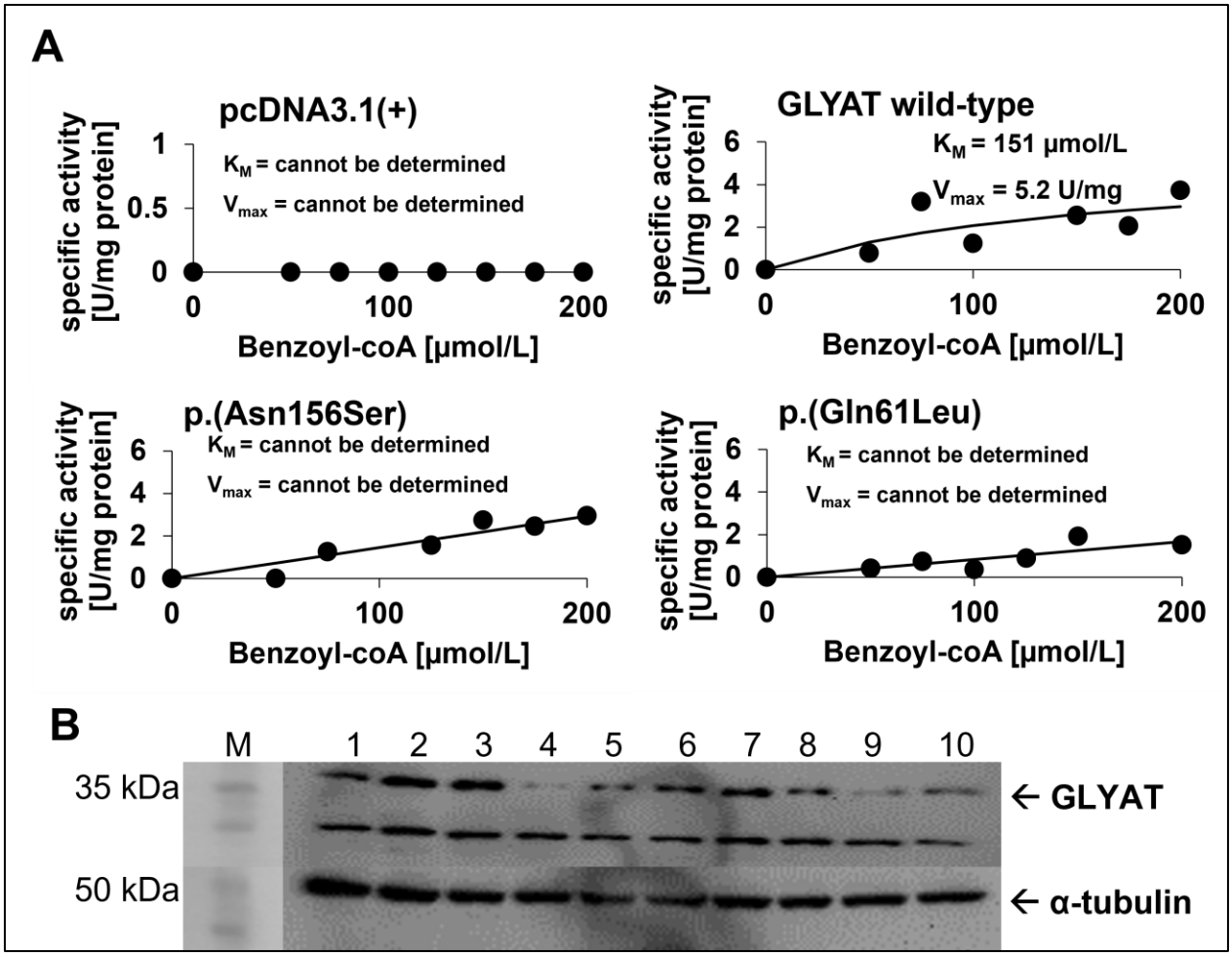


Figure 4.60: Enzyme kinetics of transiently transfected HEK293 cells with pcDNA3.1(+), GLYAT wild-type and sequence variants p.(Asn156Ser) and p.(Gln61Leu) (A) and corresponding Western blot (B) (30 µg total protein from homogenate supernatant). 1: pcDNA3.1(+), 2: GLYAT wild-type, 3: GLYAT p.(Asn156Ser), 4: GLYAT p.(Arg199Cys), 5: GLYAT p.(Gly38*), 6: GLYAT p.(Val57Phe), 7: GLYAT p.(Gln61Leu), 8: GLYAT p.(Met65Thr), 9: GLYAT p.(His101Tyr), 10: GLYAT p.(Ala231Thr). Anti-GLYAT PA5-48504, 1:1,000 and anti-tubulin ab7291, 1:5,000 were used as first antibodies.

The percentage of the used SDS gels showed influences here due to better signal separation in the higher percentage (15 %) gel. However, the enzyme activities differed remarkably from those observed in Figure 4.59. Thereby, a high inter-assay variation became obvious. Both sequence variants demonstrated activities, which resemble the wild-type activity. The v_{max} values could not be calculated via Excel Solver tool due to high variation and no saturation of the curves.

Nonetheless, an enzyme activity was observed in all studies thus far (Figures 4.58. – 4.60.) except for the pcDNA3.1(+) control. Another reproduction series is demonstrated by Figure 4.61.

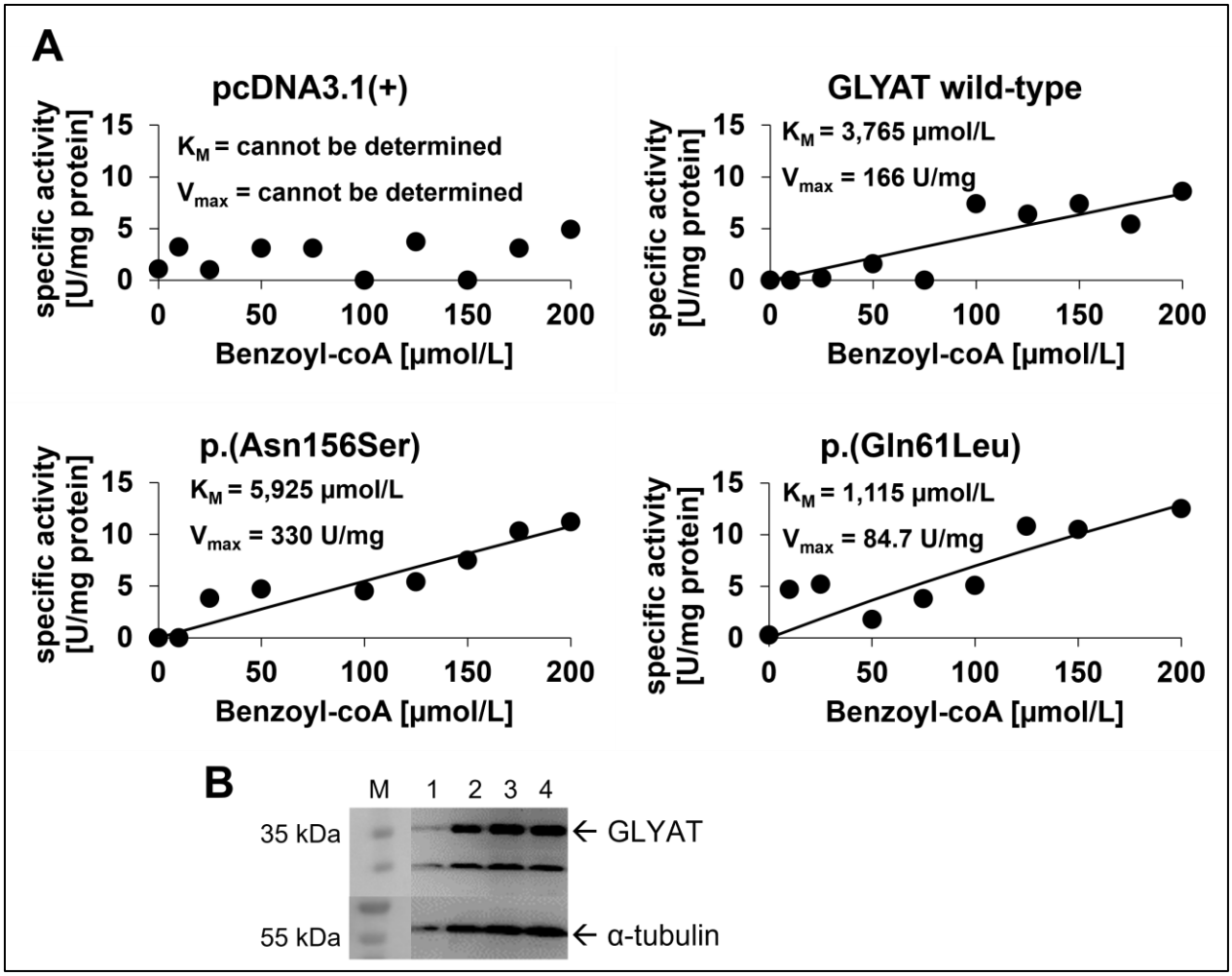


Figure 4.61: Enzyme kinetics (A) and corresponding Western blot (B) of 15 µg total protein of homogenate supernatant from transiently transfected HEK293 cell homogenates with GLYAT wild-type and sequence variants. 1: pcDNA3.1(+), 2: GLYAT wild-type, 3: GLYAT p.(Asn156Ser), 4: GLYAT p.(Gln61Leu). Anti-GLYAT PA5-48504, 1:1,000 and anti-tubulin ab7291, 1:5,000 were used as first antibodies.

Derived from that, the GLYAT p.(Asn156Ser) obtained the highest activity (v_{max}: 330 U/mg protein) followed by wild-type with 166 U/mg protein and p.(Gln61Leu) with only 84.7 U/mg protein. K_M values were difficult to evaluate due to completely unrealistic values (above 1000 µmol/L).

Overall, the series of transfected HEK293 cells with GLYAT (Figures 4.58. – 4.61.) demonstrated that GLYAT overexpression worked but the enzyme activities differ clearly. However, in all series an enzyme activity for GLYAT wild-type and sequence variants p.(Gln61Leu) and p.(Asn156Ser) was detected but the quantitative differences between the activities are difficult to identify. Table 38 summarizes the results of the kinetic parameters for GLYAT wild-type and sequence variants overexpressed in HEK293 cells.

Table 38: Kinetic parameters for GLYAT wild-type and sequence variants in HEK293 cells (mean ± standard deviation; n=5)

Sequence variant	v_{max} [U/mg protein]	K_M [µmol/L]
GLYAT wild-type	95.6 ± 69.6	1389 ± 1751
p.(Gln61Leu)	94.7 ± 28.7	534 ± 513
p.(Asn156Ser)	144 ± 163	1996 ± 3408

To sum this up, the activity of human GLYAT wild-type was determined in HEK293 cell homogenates (v_{max}: 95.6 ± 69.6 U/mg protein) but it was indistinguishable from the sequence variants p.(Gln61Leu) and p.(Asn156Ser). Thereby, high activity variations in crude cell homogenates were detected, which underline the difficulties to derive statements from that. However, the qualitative confirmation of the GLYAT overexpression could thus be shown.

4.11. Overexpression and enzyme activity assays of GLYATL1 in HEK293 cells

For overexpression of human GLYATL1 in HEK293 cells, the same approach as in the previous section was performed using transiently transfected HEK293 cells.

The Figures 4.62.-4.64. demonstrate the GLYATL1 activities using different phenylacetyl-coA concentrations in combination with overexpression validation by Western blot.

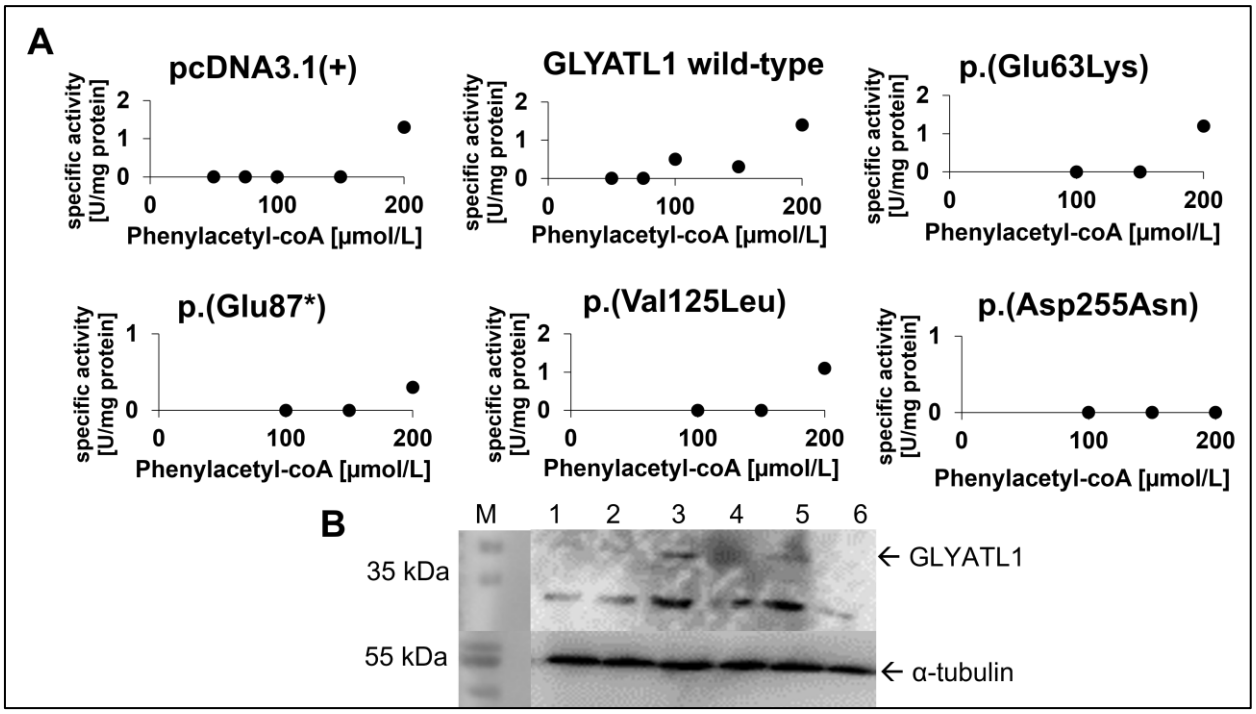


Figure 4.62: **Enzyme kinetics (A) and Western blot (B) of 30 µg total protein from homogenate supernatant of transient transfected HEK293 cells**. For the sequence variants p.(Glu63Lys), p.(Glu87*), p.(Val125Leu) and p.(Asp255Asn) only 100, 150 and 200 µM phenylacetyl-coA concentrations were determined due to inactivity of the protein. For pcDNA3.1(+) and GLYATL1 wild-type additionally 50 and 75 µM were measured. B: Western blot of 30 µg total protein homogenate of transiently transfected HEK293 cells with GLYATL1 wild-type and sequence variants. M: PageRuler™ Prestained Protein Ladder, 1: pcDNA3.1(+), 2: GLYATL1 wild-type, 3: p.(Glu63Lys), 4: p.(Glu87*), 5: p.(Val125Leu), 6: p.(Asp255Asn). Anti-GLYATL1 GTX106956, 1:1,000 and anti-tubulin ab7291, 1:5,000 were used as first antibodies.

No Michaelis-Menten concentration dependent protein activity was detected for any of the protein variants (Figure 4.62.). Hence, no K_M or v_{max} calculation was performed. To save material for the assay, the lower phenylacetyl-coA concentrations (<100 µmol/L) were not applied for the sequence variants. The Western blot (Figure 4.62. B) only showed very weak signals for GLYATL1 wild-type and sequence variant p.(Val125Leu), whereas p.(Glu63Lys) variant demonstrated the most intensive target protein band.

Stably transfected HEK293 of another series were used for the determination of GLYATL1 activity and overexpression validation by Western blot (Figure 4.63.).

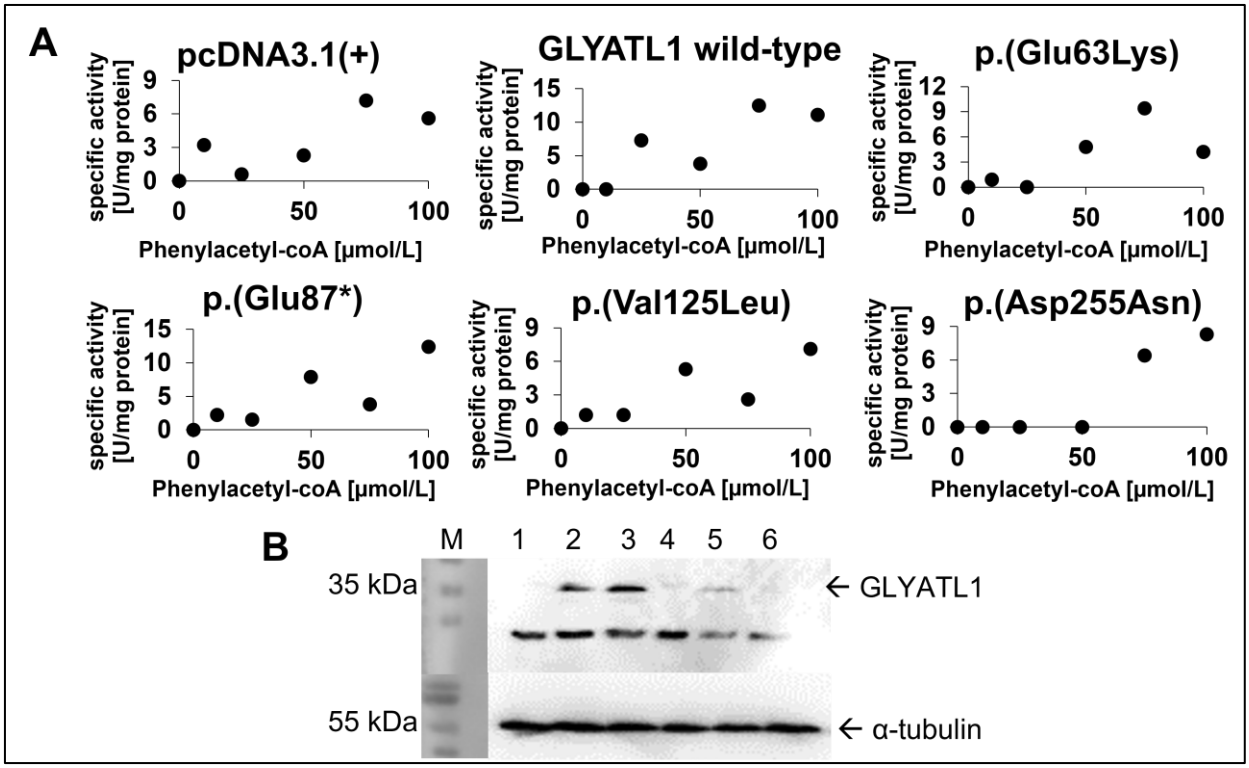


Figure 4.63: A: Enzyme kinetics (A) and Western blot (B) of stably transfected HEK293 cells (30 µg total protein from homogenate supernatant) with GLYATL1 wild-type and sequence variants. G. B: Western blot of 30 µg total protein homogenate of stably transfected HEK293 cells with GLYATL1 wild-type and sequence variants. 1: pcDNA3.1(+), 2: GLYATL1 wild-type, 3: p.(Glu63Lys), 4: p.(Glu87*), 5: p.(Val125Leu), 6: p.(Asp255Asn). SDS PAGE and Western blot were performed according to standard conditions. Anti-GLYATL1 GTX106956, 1:1,000 and anti-tubulin ab7291, 1:5,000 were used as first antibodies.

No considerable difference between pcDNA3.1(+) empty vector and GLYATL1 wild-type was detected, which was corroborated by wild-type comparable activity of stop mutant p.(Glu87*). Hence, the previously detected insufficient overexpression of GLYATL1 in HEK293 cells (referred to enzyme activity testing) was confirmed second time with independent biological replicates.

However, another transfected HEK293 cell series showed maximum GLYATL1 activity using 100 µmol/L phenylacetyl-coA - and Western blot for overexpression check (Figure 4.64.).

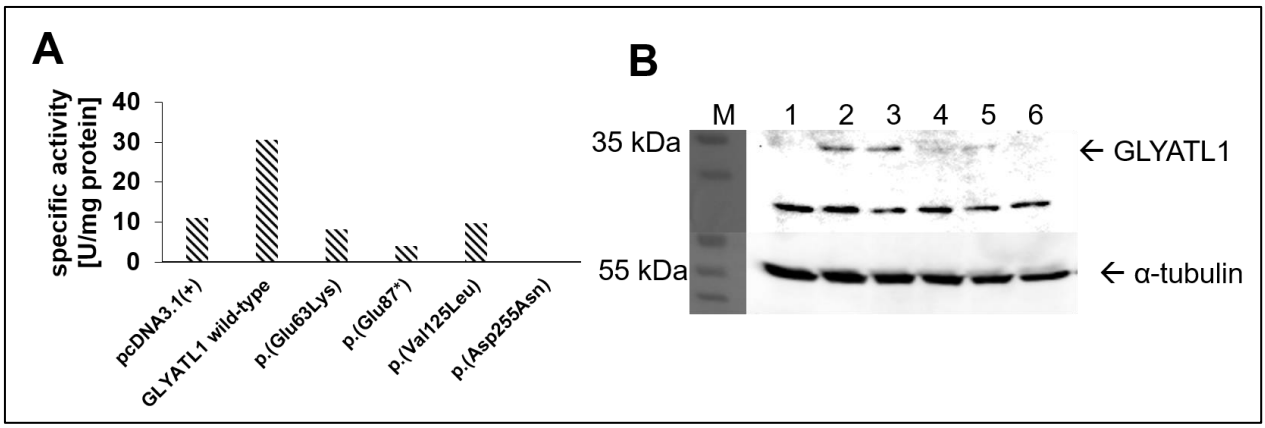


Figure 4.64: Maximum GLYATL1 activity of stably transfected GLYATL1 wild-type and sequence variants at 100 µmol/L phenylacetyl-coA (A) and Western blot (B) of 30 µg total protein from homogenate supernatant. 1: pcDNA3.1(+), 2: GLYATL1 wild-type, 3: p.(Glu63Lys), 4: p.(Glu87*), 5: p.(Val125Leu), 6: p.(Asp255Asn). For all sequence variants except p.(Asp255Asn) GLYATL1 activity signal was detected. For wild-type the activity was 30.7 U/mg protein, while for empty vector pcDNA3.1(+) the unspecific activity was at 11 U/mg protein. Sequence variants show following activity values: p.(Glu63Lys): 8.2 U/mg protein, p.(Gly87*): 4.1 U/mg protein, p.(Val125Leu): 9.8 U/mg protein, p.(Asp255Asn): -. B: Anti-GLYATL1 GTX106956, 1:1,000 and anti-tubulin ab7291, 1:5,000 were used as first antibodies.

The Western blot demonstrated that only GLYATL1 wild-type and sequence variants p.(Glu63Lys) and p.(Val125Leu) showed target protein signals confirming sufficient stable overexpression in HEK293 cells. However, the sequence variants p.(Glu87*) and p.(Asp255Asn) showed no target protein signal at all comparable with same study with transient transfected HEK293 cells (Figure 4.64. B). The specific activity for the GLYATL1 wild-type was detected with 30.7 U/mg protein, whereas for all other variants only 1/3 of that activity was detected (Figure 4.64.).

The Western blots (Figure 4.62. -4.64.) demonstrated reproducible signals for the GLYATL1 wild-type and p.(Glu63Lys). Nonetheless, the activity detections between the series were not comparable and thus no kinetic properties can be determined.

Nevertheless, the overexpression and the qualitative protein detection of recombinant GLYAT and GLYATL1 in HEK293 cells were sufficient (confirmed by the Western blot studies in sections 4.10. and 4.11.).

The gDNA and mRNA analyses from stably transfected HEK293 cells were performed to find an explanation for the absent upregulated enzyme activities. There might be a problem in the transfection (Suppl. Fig. 43-46) or the transcription (Suppl. Fig. 47-55). As the supplemental data indicate, these assumptions could be denied.

4.12. Investigation of *GLYATL2* expression in human tissues via semi-quantitative PCR and RT-qPCR

Only very limited data are available regarding GLYATL2 expression in human tissues. However, one study treated that and localized a *GLYATL2* mRNA expression mainly in salivary gland, trachea, spinal cord and skin fibroblasts [165].

The RT-qPCR for detection of specific *GLYATL2* mRNA expression needed to be established and adapted to laboratory specific conditions–- it was examined in the available human tissue RNA of liver, testis, skeletal muscle, kidney, heart and brain. Thereby, the produced mRNA was qualitatively checked by semi-quantitative PCR in a first series (Figure 4.65.).

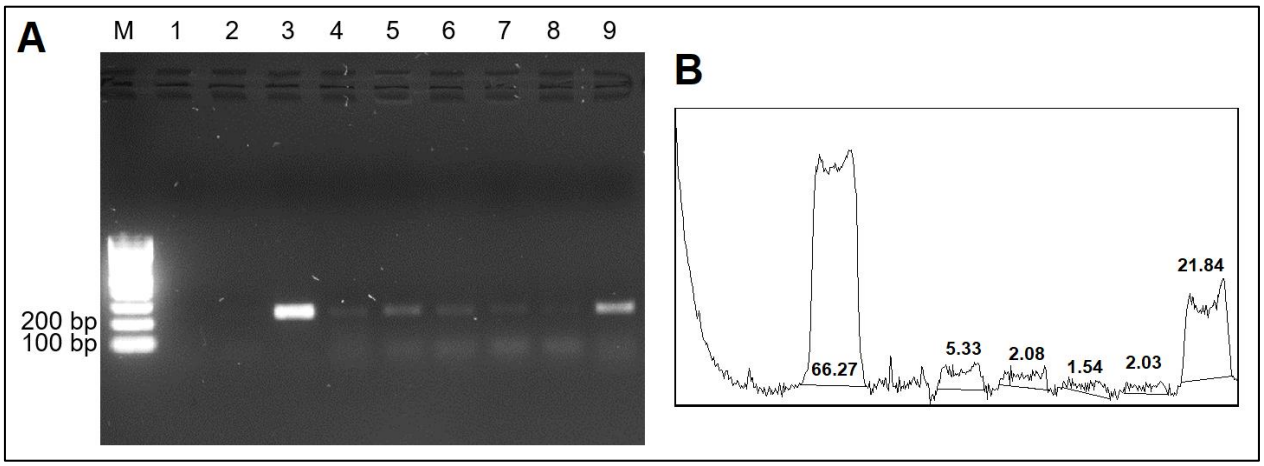


Figure 4.65: A: 1.5 % (w/v) agarose gel and ImageJ peak percentages (B) showing amplicons of GLYATL2 mRNA produced via semi-quantitative PCR. M: GeneRuler 100 bp DNA ladder, 1: deionized water control, 2: pcDNA3.1(+) 1:1,000, 3: GLYATL2_pcDNA3.1(+) 1:1,000, 4: cDNA liver, 5: cDNA testis, 6: cDNA skeletal muscle, 7: cDNA kidney, 8: cDNA heart, 9: cDNA brain. GLYATL2 amplicon produced with primers GLYATL2_302_for and GLYATL2_552_rev has a size of 246 bp

The deionized water and pcDNA3.1(+) as negative control showed no amplified DNA signal, whereas GLYATL2_pcDNA3.1(+) positive control demonstrated the expected amplicon at 246 bp (66.27 % signal intensity). Hence, the general semi-quantitative approach for detecting human *GLYATL2* mRNA was functional. Notably an upregulated *GLYATL2* expression in brain (21.84 %) and testis (5.33 %) was obvious. However, very weak *GLYATL2* signals are even detected for liver, skeletal muscle and kidney leading to assumption that these signals might be unspecific. Nonetheless, the brain signal seemed to be very specific as it reached the intensity level of GLYATL2-pcDNA3.1(+) positive control (66.27 %).

For the quantitation of relative mRNA amounts RT-qPCR was first attempted by using the same *GLYATL2* primers and human *S18* rRNA primers, which resulted in undefined signal pattern in *S18* rRNA control amplicons (data not shown). Afterwards the *S18* rRNA approach was sorted out and *RNA polymerase II* (RP2) gene, which is quite feasible as internal tissue control gene was then used instead of *S18* rRNA for tissue expression control [126] (Figure 4.66.).

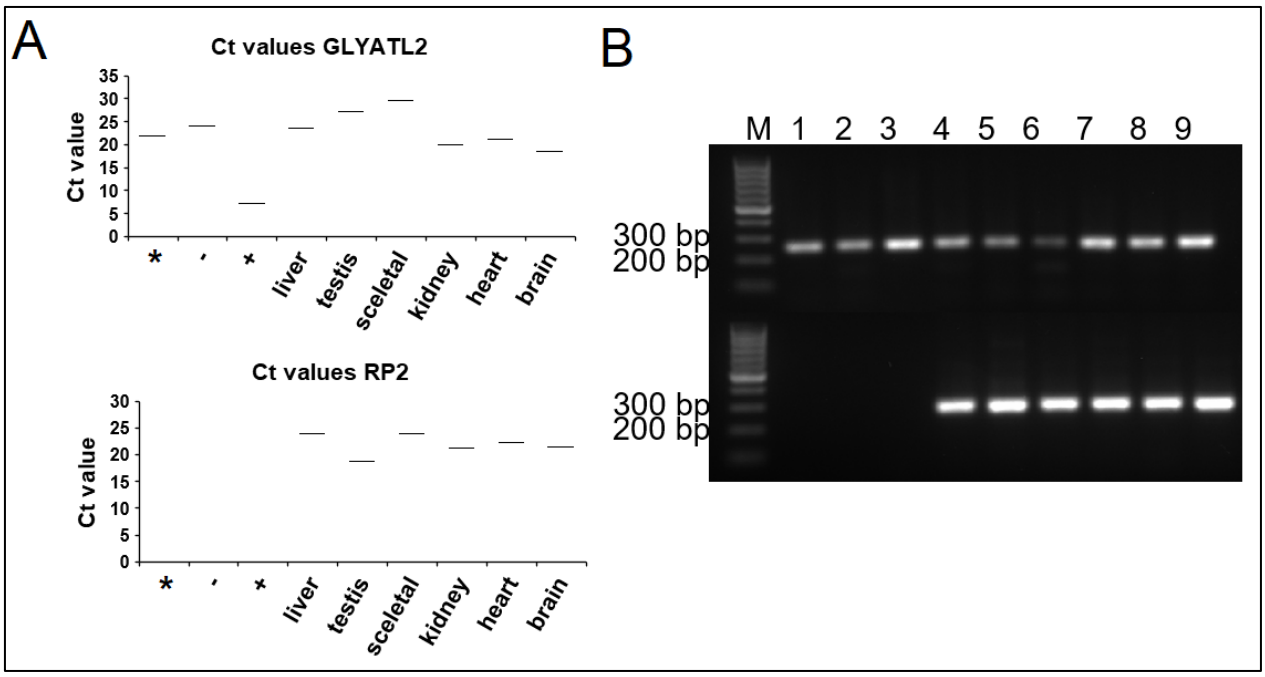


Figure 4.66: Ct values (A), 1.5 % (w/v) agarose gel (B) and ImageJ peak percentages of GLYATL2 (C) and RP2 (D) of tissue expression RT-qPCR. * = deionized water, - = pcDNA3.1(+), 1:1,000, + = GLYATL2-pcDNA3.1(+), 1:1,000. B: 1: deionized water control, 2: pcDNA3.1(+) 1:1000, 3: GLYATL2-pcDNA3.1(+) 1:1000, 4: cDNA liver, 5: cDNA testis, 6: cDNA skeletal muscle, 7: cDNA kidney, 8: cDNA heart, 9: cDNA brain. M: GeneRuler 100 bp DNA ladder. GLYATL2 amplicon produced with primers GLYATL2_302_for and GLYATL2_552_rev has a size of 246 bp, whereas in the lower part of the picture human RNA Polymerase II was used as control with an amplicon size of 266 bp

Regarding internal control gene *RP2* qPCR was adequate due to equal amplicon signals (variation of signal intensity from 14.95 % to 18.34 % only). However, for *GLYATL2* detection there was still unspecific signal for deionized water (7.38 %) and pcDNA3.1(+) control (6.85 %) remaining while GLYATL2 signal observed 13.23 % signal intensity. To reduce unspecific signal another approach was the treatment of all experimental materials for 45 min under UV hood leading to the same disturbing unspecific signal as shown before.

Since there were unspecific amplicons observed different annealing temperatures (70 °C, 72 °C, 74 °C) were tested (Suppl. Fig. 61).

Corresponding to that, an annealing temperature of 71 °C was chosen for further studies, because no signal is detected in the non-template control deionized water and pcDNA3.1(+) and positive control of GLYATL2 plasmid was still functional (Suppl. Fig.62).

71 °C annealing reduced unspecific formation of GLYATL2 amplicon in deionized water and pcDNA3.1(+) negative controls (Suppl. Fig. 62). However, it seemed to decrease cDNA signals in general due to only brain cDNA remaining signal here (7.84 % signal intensity). RP2 signals did not differ strongly from each other (except outlier of cDNA liver with 11.83 % signal intensity). Due to lack of signals for most of the cDNAs (even for the positive control salivary gland [166] several

optimizing steps were performed (Suppl. Fig. 63). None of them revealed the expected result of an upregulated *GLYATL2* amplicon.

Nonetheless, the differences between signal intensities were attributed to varying expression patterns and not to cDNA concentrations as they were measured and equalized before.

Derived from all the *GLYATL2* tissue expression RT-qPCR experiments no remarkable endogenous expression of *GLYATL2* in any of the tissues or the cells was detected. However, the RT-qPCR itself was functional concluded from the used positive and negative controls.

To summarize, the GLYATL2 tissue expression could not detect the protein with reliable data in a specific tissue.

Summary of the results

As shown in the results section, the human GLYAT and GLYATL1 were overexpressed in *E. coli* and, with limitations, in the human-derived HEK293 cells. The wild-types and certain sequence variants, which had not previously been characterized on the enzyme level were produced and characterized within the work for the presented thesis. However, the GLYAT sequence variants p.(Val57Phe) and p.(Met65Thr) did not show protein signals after overexpression. Due to low expression rates and activity signals of most GLYAT and GLYATL1 sequence variants in HEK293 cells, the major focus was on the bacterial system.

The GLYAT wild-type and sequence variants p.(Gln61Leu) and p.(Asn156Ser) were purified by FPLC and kinetic properties could be determined with reliable data (Section 4.6.3.). The glycine dependency of the purified GLYAT wild-type could be confirmed (Section 4.6.4.). Nonetheless, the inter-assay comparison of the spin-column purified enzymes was high (Section 4.6.5.), which was not the case for the intra-assay comparison (Section 4.6.1.).

The FPLC purified GLYATL1 wild-type could not be analyzed completely (Section 4.5.4.), whereas the spin-column purified wild-type delivered reproducible data (Section 4.6.6.). Nonetheless, high inter-assay variation was observed for the sequence variants and the p.(Glu63Lys) could not be expressed in different series (Section 4.6.6.).

While the recombinant GLYAT and GLYATL1 could be overexpressed in HEK293 cells, their kinetic properties delivered no stable data (Section 4.10., 4.11.). However, the GLYAT wild-type and the p.(Gln61Leu) and p.(Asn156Ser) sequence variants could be overexpressed in the same way as the GLYATL1 wild-type and its p.(Glu63Lys) and p.(Val125Leu) sequence variants. Notwithstanding, the gDNA (Suppl. Fig. 43-46) and mRNA analyses (Suppl. Fig. 47-55) showed that the problem of low expression levels lies in the level of protein translation, although the 5´-azacytidine and MG-132 did not present the expected results (Suppl. Fig. 41, 42).

The intracellular localization studies on GLYAT and GLYATL1 (Section 4.8.) confirmed their mitochondrial localization. Thus, the mitochondria-enriched supernatants were used for the analyses of the recombinant enzymes in HEK293 cells (Section 4.10., 4.11.). However, these activity studies did not result in reliable data.

While the preparatory molecular work was performed as part of this thesis work, the enzyme analyses were outsourced to a bachelor thesis [164] that was co-supervised by the author. The GLYATL2 tissue expression studies did not confirm the literature data (Section 4.12.), although the same experimental set-up as described in the literature was used.

5. Discussion

The human enzymes GLYAT, GLYATL1 and GLYATL2 were intended to be characterized regarding their kinetic properties. Therefore, frequent sequence variants were analyzed in addition to elucidate the influence on enzyme activity. These mutants and the wild-types were overexpressed in a bacterial and human overexpression system. The sequence variations may represent pharmacogenetic peculiarities with therapeutic consequences in several metabolic diseases, for instance urea cycle defects or isovaleric acidemia. The latter was very prevalent in South Africa [39], where the GLYAT p.(Gln61Leu) sequence variation was also very frequent due to a founder effect [153]. Hence, this variant was overexpressed and analyzed in a bacterial and human-derived system. Since the p.(Asn156Ser) sequence variant was suggested to be the authentic GLYAT wild-type [129, 152], it was analyzed in parallel to the canonical GLYAT wild-type. It should be investigated, to what extent gene variants influence the enzyme activities. This may have consequences for drug metabolism and personalized medicine.

Due to the lack of data on the overexpression of the GLYATL1 and GLYATL2 wild-types and sequence variants, the sequence variants were analyzed in databases and produced by mutagenesis. The kinetic properties were determined for GLYAT and GLYATL1. The GLYATL2 enzyme analysis was outsourced to a bachelor thesis [164]. The intracellular localization of human GLYAT and GLYATL1 should be addressed as the literature data were contradictory.

5.1. Overexpression and PTMs of human amino acid *N*-acyltransferases

The HEK293 system did not turn out as suitable for the overexpression of recombinant GLYAT and GLYATL1. However, a qualitative confirmation of protein overexpression of the wild-types and certain sequence variants was functional (Sections 4.10., 4.11.).

The ability of *E. coli* to produce authentic PTMs is limited [37]. Therefore, it was not capable to guarantee the occurred PTMs that are predicted for GLYAT, GLYATL1 and GLYATL2 (Section 4.7.). Notwithstanding, it is not clarified, to what extent human *N*-acyltransferases require the predicted PTMs to fulfill 100 % of their activity. Modifications, especially within the active center, may contribute to enzyme activity but, the reaction mechanism of the human enzyme remains unclear due to unknown X-ray data. Conversely, the reaction mechanism for the bovine GLYAT was described at the active center [5].

Next to the wild-type enzymes, several sequence variants with negative predictions for enzyme activity yielded sufficient target protein yields in *E. coli* (Section 4.4). Nonetheless, certain sequence variants were not overexpressed due to false positive clones e.g., p.(Gln61Leu) sequence variant (Figure 4.15.). The selection was functional by chloramphenicol towards the pGro7 plasmid and in a second transformation via ampicillin towards the pET32a(+) plasmid. The

false positive clones might have been present for several reasons: the antibiotic concentration might have been too low as it was adapted to *E. coli* Origami 2(DE3). Conversely, ligation problems with re-ligands produced without insert (albeit dephosphorylation via rSAP was performed) might have occurred. Indeed, false positive clones for the selection by ampicillin have been described [52]. Thus, Western blot analyses were required to ensure the presence of target protein before starting enzyme activity tests for the bacterial studies.

5.2. Sequence variations of human *N*-acyltransferases and influences on enzyme activities

5.2.1. Adenine or guanine at nucleotide position 467?– real wild-type of human *GLYAT*

The GLYAT c.467A>G (p.(Asn156Ser)) sequence variant was previously described with 85 % in a Japanese and with 97 % allele frequency in a French Caucasian cohort [101, 175]. Based on this, the enzyme studies described the p.(Asn156Ser) variant as the real wild-type, which might be incorrectly annotated in gene databases (NM_201648.2) [152, 154]. Thus, these activity studies were re-examined. The increased p.(Asn156Ser) activity was indicated by the present study and another publication [129, 137]. Notwithstanding, the data of the present study were published earlier [137] than the data from Rohwer and associates [129].

The canonical human GLYAT reference sequence (NM_201648.2) possesses an adenosine at position 467 of the coding sequence and was reported as the original enzyme wild-type [206]. That is corroborated by numerous other publications [107, 152, 154]. The amino acid Asn^{156} is localized on a poorly predicted loop from Lys^{159} to Met^{167}, which complicates the interpretations on active center and enzyme activity. Notwithstanding, in this project all generated sequence variants shown in Section 4.2. were produced based on the canonical human GLYAT wild-type and for the more probable wild-type c.467A<G in addition.

However, the sequence variants based on c.467G>A were only produced on a plasmid level yet and remained uncharacterized on an enzyme level due to lost time in investigations of the original GLYAT wild-type and corresponding sequence variants.

5.2.2. The selection of frequent enzyme variants for mutagenesis

The human GLYAT and GLYATL1 sequence variants were selected according to their allele frequency and online prediction tools (Section 3.2.). A high number of low-frequency alleles were observed in the *GLYAT* gene by Tajima´s neutrality test [129], which complicates identification of frequent variants. Variants undergo negative selection during evolution complicating identification of alleles with frequencies > 0.2 %. Those rare variants were selected due to their relevance on cohort analyses (high allele frequencies) increasing the chances of GLYAT deficient patient material in the future. Most of the analyzed sequence variants (Section 3.2., Table 17) exhibit cohort frequencies below 1 % and are thereby not defined as polymorphisms. However, a

consensus definition of ´polymorphism´ is missing thus far and the usage of this term seems to be subjective [81, 136]. Van der Sluis and associates evaluated certain frequent GLYAT variants p.(Lys16Asn), p.(Ser17Thr), p.(Lys20Gln), p.(Lys20Arg), p.(Arg131His), p.(Asn156Ser), p.(Phe168Leu), p.(Arg199Cys) and p.(Glu227Gln), whereas p.(Ser17Thr) and p.(Asn156Ser) exhibited allele frequencies of about 1 %, thus titled as polymorphisms [136].

The p.(Asn156Ser) and p.(Arg199Cys) variants were chosen for reproduction and as controls. However, according to data bank analysis [197] the polymorphisms were difficult to identify. Only the p.(Gln61Leu) variant that was described with 12% and prevalent within the South African cohort [153] was found and, due to lacking functional assessment, examined in more detail.

193 missense variants of GLYAT were identified, whereas only two of them, i.e. p.(Ser17Thr) and p.(Asn156Ser), showed allele frequencies higher than 0.5 % [129]. The remaining variants showed allele frequencies of 0.2 % - 0.0004 %. The p.(Asn156Ser) variant also had the highest homozygous genotype frequency of 90.3 %, whereas six other variants were found as homozygotes [129]. As GLYAT maintain a sensible balance of coA levels within mitochondria and might be essential for life (Section 1.4.3.), the main part of the deleterious sequence variants appeared at very low frequencies. The GLYAT sequence variant p.(Val57Phe) remained unexpressed in all series in *E. coli* (Section 4.5., 4.6.) and in HEK293 cells (section 4.10.). However, the association between missense mutations and diseases caused by protein instability or conformational changes were reported [143]. The introduction of the aromatic residue by Phe might contribute to destabilization and fast degradation of target protein that might have resulted in absent overexpression.

The GLYAT sequence variant p.(Asn156Ser) was overexpressed and extracted from *E.coli* Origami 2(DE3) under the influence of a protease inhibitor cocktail (PIC) (Figure 4.17.). No difference between the untreated and PIC-treated sample was obvious. Thus, the present proteases in the *E.coli* periplasm [17] did not affect the overexpressed GLYAT sequence variant.

The impact of the GLYAT sequence variations on the therapy of certain metabolic diseases will be discussed in Section 5.5.

5.2.2.1. Special case of GLYAT p.(Arg199Cys)

The GLYAT sequence variant p.(Arg199Cys) was not appropriately overexpressed (*E. coli*: Figure 4.8., HEK293: Figure 4.58.), which complicated the interpretation of enzyme activity studies on this variant. The decreased enzyme activity (*E. coli*: Figure 4.37., HEK293: Figure 4.58.) can be justified by missense mutation or, more likely, by reduced presence of the target protein. Nonetheless, the p.(Arg199Cys) variant with equal expression pattern compared to wild-type but with only low total enzyme activity was identified [152]. A more recent study described the

alteration of α-loop-α motif by p.(Arg199Cys) mutation, which is crucial for substrate binding in the Gcn5-related *N*-acetyltransferases [44, 129, 161]. The mutation itself may not serve as a reason for the impaired overexpression here as it was successful in the study by van der Sluis [152]. However, as some false positive clones were observed for the p.(Gln61Leu) sequence variant (Figures 4.14., 4.15.) this might have also been the case for the p.(Arg199Cys) variant.

The amino acids Arg^{199} and Asp^{48} stabilize central α-loop-α-structure [152], which is important for substrate binding in the GNAT enzyme superfamily [44] including GLYAT. The highly conserved loop provides the five amide backbone hydrogen bonds to the pyrophosphate group of coA [161]. A reduced expression of the p.(Arg199Cys) variant was demonstrated (Figures 4.8., 4.16., 4.58.), which might be related to the instable α-loop-α-structure.

5.2.3. Is the glycine conjugation pathway essential for life?

Due to absent reports of glycine conjugation defects and the strong conservation of the *GLYAT* gene, the investigation of sequence variations is of considerable interest. The coding regions of *GLYAT* in Afrikaner Caucasian individuals have been analyzed and several founder mutations were identified [153]. It was obvious that c.182A>T sequence variant of *GLYAT* attracted attention due to suspicious high allele frequency of 12.3 % in Afrikaner Caucasian population. This variant might have become prevalent in this population due to a founder effect. This was one of the reasons for detailed investigation of this variant in the present study.

As the GLYAT enzyme conjugates isovaleryl-coA to glycine in the event of IVA [39], the glycine therapy optimizes life quality by reduction of disease symptoms e.g., poor feeding, vomiting, seizures, and metabolic acidosis within the first two weeks of life [163].

Due to the missing reports on glycine conjugation defects and the high conservation and activity of GLYAT in hepatic metabolism, the glycine conjugation seems to be essential for survival [7, 129, 153]. Genetic knowledge for this key detoxification enzyme is insufficient and various pathologic conditions such as diabetes, obesity, gut dysbiosis, autism, schizophrenia, depression, hepatitis, and cancer are correlated with abnormal high levels of hippurate in the urine [95]. Amino acid conjugation processes are also seen as homeostatic and neuroregulatory mechanisms [15] due to influences on body store amino acid regulation. Amino acids like glycine, glutamate or arginine are key neurotransmitters in the central nervous system (CNS). The Janus nature of the amino acid, which represents conjugation agents and neurotransmitters was described [15]. The excretion of glycine and glutamine as their aromatic acid conjugates in the urine may cause psychosomatic consequences because they are trafficked from the brain and thus influence the CNS. Several drugs replacing benzoic acid may perturbate glycine homeostasis e.g., cancer chemotherapeutic agent ifosfamide [14]. Enhancement of the production of aldehyde metabolites

depletes ATP production and sequesters NADH in the mitochondrial matrix. This in turn, inhibits the glycine deportation system and induces *de novo* glycine synthesis. These phenomenona may result in hyperglycinemia and encephalopathic effects [14]. The glycine deportation system in the mitochondrial matrix is essential for glycine homeostasis and is upregulated by benzoic acid administration. Thus, congenital urea cycle enzymopathies and NKH might be treated efficiently by benzoic acid [158].

5.3. Localization of human *N*-acyltransferases GLYAT and GLYATL1

5.3.1. Localization prediction of human *N*-acyltransferases

To investigate the localizations of human *N*-acyltransferases GLYAT, GLYATL1 and GLYATL2, the prediction tool TargetP-2.0 (Section 2.10.) was used. According to this tool, the three enzymes might not possess *N*-terminal mitochondrial signal peptides (Suppl. Fig 38-40) as the likelihood for that were only 18 % (GLYAT), 8 % (GLYATL1) and 2 % (GLYATL2). However, for GLYATL2 it seemed to be senseful as it was described as a protein, which was supposed to be located in the ER [165]. Human GLYAT was suggested to be localized within mitochondria [107], but there was conflicting information on GLYATL1 since it was detected in the cytosol [178] as well as in the ER [107]. The experiments on localization (Section 4.8.) served to verify the predictions and the literature data.

5.3.2. Mitochondrial localization of endogenous human GLYAT and GLYAT-eGFP

As confirmed by the confocal laser scanning microscopy images (Section 4.8.), the eGFP-fused human GLYAT enzyme was clearly localized within mitochondria [137], which confirmed the findings by Matsuo et al. [107]. A Western blot study underlined the sufficient expression of the GLYAT-eGFP (Figure 4.49.). The results addressed questions from previous studies on that part [60, 91, 107] and provided a clear result on it. However, the used methods differ: Kølvraa & Gregersen detected Glycine *N*-acylase activity in the mitochondrial fraction of a rat liver [91], whereas Gregersen et al. confirmed the same for human liver mitochondria [60]. Matsuo et al. identified mitochondrial human GLYAT via immunostaining of HeLa cells and Western blot [107]. A Western blot study with human liver homogenate fractions did not show the endogenous GLYAT exclusively within the mitochondrial fraction, because it was detected in each fraction with the equal signal intensities (Figure 4.50.). Nonetheless, according to the previous confocal laser scanning microscopy studies, the human GLYAT should be localized within mitochondria only (Figure 4.44., 4.47.). Since the homogenate fraction contains the mitochondrial proportion accordingly, human GLYAT would also be expected therein. Nevertheless, an equal intensity in each fraction did not meet the expectations, especially not for the cytosol. The used GLYAT antibody PA5-48504 is polycloncal, which may resulted in unspecific cross-reactions with proteins

of the same size in the protein-enriched homogenate and in the cytosol fraction. This might have overlayed the expected human GLYAT signal. Another aspect might have been the cross-contamination of the fractions delivered from the Tebu-bio company. The cytosolic fractions contain many unique drug-metabolizing enzymes, mainly sulfotransferases, *N*-acetyltransferases, and glutathione S-transferases, that are recommended as test systems for phase II reactions of test compounds. However, as the samples contain these various marker enzymes, they might have interfered with the polycloncal GLYAT antibody (Figure 4.50.).

The enzyme activity of human GLYAT in liver fractions (homogenate, cytosol, mitochondria) was investigated accordingly (Figure 4.51.). Nonetheless, as human GLYAT was expected to be localized only in the mitochondria, it was confusing to have nearly similar activity of human GLYAT in cytosole (v_{max}: 603 U/mg) compared with mitochondria (v_{max}: 655 U/mg). It must be considered that the same total amounts of total protein were used for the activity determination in the cytosol and mitochondria. 99 % of the synthesized proteins in the cytosol are transported into mitochondria [119], which demonstrates the protein-accumulation in the mitochondria compared to the cytosol. Hence, the activity measurements were difficult to compare as further coA releasing enzymes might be present in the mitochondria. Most probably, the GLYAT activity in the mitochondria is thus higher, but overlapped through the contaminating enzymes. The factor of difference in protein amounts might be subtracted here. Possible cross-contaminations between cytosolic and mitochondrial fractions activity cannot be 100 % excluded.

Indeed, the GLYAT enzyme is expected to be localized within mitochondria as it is a considerable contributor of human liver detoxification, which takes place in the liver. The liver has a high mitochondrial density – 500 to 4,000 mitochondria per hepatocyte are expected in humans [47]. They are involved in the hepatic metabolism of carbohydrates, lipids and proteins and allow ammonia detoxification through the urea cycle [49, 144]. Thereby, GLYAT is expected to be frequent in the liver mitochondria. Another confirmation of the mitochondrial localization of GLYAT was shown by detection of the enzyme in mitochondrial granular structures of noncancerous cells of hepatocellular carcinoma specimen [107]. However, that complemented their identification of GLYAT in the mitochondria of HeLa cells.

5.3.3. Localization of human GLYATL1 – mitochondrial or cytosolic?

As described previously, the human GLYATL1 was predicted to be localized within mitochondria by one study, which identified glutamine *N*-acyltransferase activity in human mitochondria [169]. However, another study located the enzyme within the cytosol of COS-7 cells [178]. In a further study the enzyme was suggested to be localized in the ER [107]. This has prompted us to re-examine this question.

Nonetheless, in the present study GLYATL1 was localized within the mitochondria of HEK293 cells determined via confocal laser scanning microscopy (Figure 4.45.). The overexpressed GLYATL1-eGFP could not be detected by Western blot (data not shown). One reason could be the low expression rate of GLYATL1-eGFP, which was sufficient for detection in confocal laser scanning microscopy, but not sufficient for Western blot detection. The Western blot method can detect picogram levels of a protein in a sample. Notwithstanding, that depends on the separation efficiency of the proteins in the SDS gel and specificity of the antibody-antigen interaction [58]. The used SuperSignal™ West Femto Maximum Sensitivity Substrate can detect proteins in low femtogram amounts as published by the company [214]. The missing signal of GLYATL1-eGFP in the Western blot might be related to a technical problem with the blot or an antibody dysfunctionality.

However, the comparison between the Western blot and the confocal laser scanning microscopy strongly depended on the used amounts of target protein. While for Western blot a defined amount of supplied µg total protein was used, the exact amount detected in confocal laser scanning microscopy could only be estimated. Another point might be the rapid proteasomal degradation or wrong localization. Nonetheless, the C-terminal eGFP fusions (as applied in this case) should less influence native localization of fusion protein compared to *N*-terminal eGFP fusion [120].

Nevertheless, the endogenous GLYATL1 was identified in human liver samples by Western blot (Figure 4.51.), which confirms the microscopy results (Figure 4.45., 4.48.).

Due to confirmed mitochondrial localization of human GLYAT and GLYATL1, the subsequent enzyme activity assays with HEK293 cells (Sections 4.10., 4.11.) were performed with the mitochondria-enriched homogenate fractions.

5.4. Quantitation of the amino acid *N*-acyltransferases activity

5.4.1. Suitability of the Ellman´s assay

Referring to the enzyme activity studies from the *Xfect* and *Lipofectamine* comparison (Section 4.9.3.), strong data variations were obvious since high standard deviations revealed. The imprecisions of DTNB assay used for crude cell extracts may play a major role, e.g., due to the constitutively expressed thioesterase 1, which facilitates the S-acylation consequently releasing free coA [72]. Furthermore, the DTNB assay is not recommended for crude cell homogenates due to the present glutathione reductase and peroxidase, which also reduce DTNB [207]. The GLYAT transfection was sufficient due to the prominent band signals of recombinant GLYAT (e.g., Figure 4.58.). Comparing HEK293 and HEK293T cell transfections, a slightly increased GLYAT overexpression is detected for HEK293T cells (Figure 4.55., Tr. 2). The published higher transfection efficiency for HEK293T cells, due to stable expression of SV40 large T antigen to

increase target protein production [106], compared to HEK293 cells was confirmed thereby. No essential differences were obvious through the comparison of *Xfect* and *Lipofectamine 3000*. However, the HEK293T cells transfected with *Lipofectamine 3000* showed a slight trend of a more efficient GLYAT overexpression in the Western blot (sample 8, Figure 4.55.) and enzyme activity (Figure 4.56.). Nonetheless, this effect was relativized by the high standard deviation.

Therefore, optimization approaches of eukaryotic overexpression were performed by regulation of *CMV* promotor (5´-azacytidine) or proteasomal degradation (MG-132). To investigate that more precisely, the *CMV* promotor silencing was elucidated by treatment with 5´-azacytidine. This compound is a *CMV* promotor methylation inhibitor [29] that can enhance the *CMV* driven protein overexpression. The aim was to turn off the methylation and thereby initiate the GLYAT overexpression. As the data revealed, no difference between 5´-azacytidine treated and untreated HEK293 cells occurred (Suppl. Fig. 41). Hence, the low protein expression might not be dependent on *CMV* methylation and is justified elsewere. MG-132, which is a proteasomal inhibitor decreases the proteasomal activity [114], thus protecting the GLYAT target protein from rapid degradation. However, when comparing DMSO and MG-132 treated GLYAT transfected HEK293 cells no severe difference on Western blot is obvious, which also applied to the p.(Asn156Ser) variant (Suupl. Fig. 42). Only for the sequence variant p.(Val57Phe) the MG-132 effect seemed to be obvious, confirmed by an increased protein signal. Albeit for this variant there seems to be an effect, it was not longer haunted due to the low activity signals for the similar expressed GLYAT wild-type.

To summarize, the investigation of post-translational protein impairments 5´-azacytidine and MG-132 did not show the expected effects and the low expression rate seems to have another cause. However, severe disadvantages of this method are published and contribute to high background activities within the measurements. Especially HEK293 homogenates exhibited high background activities (Section 4.10., 4.11.), which resulted partially in confusing results: the sequence variant p.(Gln61Leu) showed only cloudy activity data without regression due to high background signals (Section 4.10.), which resemble the GLYAT wild-type activity. Notwithstanding, the GLYATL1 wild-type seemed to be inactive at phenylacetyl-coA concentrations below 100 µmol/L (Section 4.11.). These observations complicated the analysis and the identification of correlating graphs. High background activities, supplemented by low expression rates of GLYATL1 in HEK293 cells (Section 4.11.), resulted in data clouds within activity measurements. Thereby, Michaelis Menten kinetic could not be detected. The assay is not recommended for activity determinations in crude cell homogenates due to interfering substances like free thiol groups and thioesterases, which contributes to the coA release [40]. For instance, the acyl protein thioesterase 1 is continuously

expressed in HEK293 cells [69] contributing to S-acylation. As already mentioned above, the acyl-group transfer is catalyzed thereby releasing free coA, which is present in crude cell homogenates [72] and thus interferes with GLYAT reaction.

The pH crticically influences the activity assay because it is designed for a pH of 8.0 [46]. Hence, tiny variations might have essential impacts on the measurement. The activity optimum of the bovine GLYAT was determined at pH 7.5 [5]. In the present study, a pH of 8.0 was maintained due to the literature specifications [107, 152]. The molar extinction coefficient of TNB is stable at pH range from 7.6 to 8.6 [127]. Beyond that, the temperature sensitivity of DTNB played a leading role: due to thermochromic effects of DTNB [127] constant temperature of 37 °C was maintained. However, pipetting steps before were performed at RT, thereby constant 37 °C were only kept during measuring process in the plate reader. Supporting that, the light sensitivity of DTNB was described [48]. Thereby, yellow photodegradation products generated during measurements were problematic. This is of special importance for endpoint determinations, which were not performed within this project. Nonetheless, the DTNB assay determines relative amounts of generated product related to released coA. Thus, it represents an indirect enzyme activity assay.

The spontaneous re-oxidation of SH-groups directly influences protein structure and activity [154] representing another issue of the assay. DTNB may also inhibit target enzyme, as it was shown for the carnitine acetyltransferase from skeletal muscle of the Arabian camel [3]. The enzyme activity was reduced by 50 % by using 0.7 mmol/L DTNB for 2 min. Although this effect was not described for bovine GLYAT [116], it was chosen for the human GLYAT.

Nevertheless, an overall high data variation within DTNB assay was obvious (Section 4.6., 4.10., 4.11.) and possible reasons might be of technical nature as well: time of substrate adding, cell disruption and foam development. The time point of benzoyl-coA adding and equal start of the reaction in each tube were very critical for the activity assay and resulting data. Accordingly, it was not possible to apply the substrate into all wells simultaneously due to frozen aliquots of benzoyl-coA in reaction vessels. A multi-channel pipet was not suitable due to impossibility of sucking the solution from different reaction vessels. The substrate solutions were frozen in small proportions in reaction tubes which impeded the transfer into another vessel or the work with a multi-channel pipet. Moreover, it was not available in the laboratory during these measurements. Hence, small time delays from well to well were unavoidable. However, cell disruption and foam development did influence the assay in a critical level. Foam leads to protein denaturation by molecule desorption at the gas-liquid interface [32]. Therefore, the uncontrollable foam development during cell disruption via needle treatment in sonication might explain the partially destroyed protein molecules that resulted in restricted activity.

Nonetheless, the GLYATL1 enzyme seems to be very sensitive within Ellman´s assay: monovalent cations such as K^+, Na^+, Rb^+, Li^+ or NH_4^+ are required for enzyme activity [116]. Nonetheless, GLYAT and GLYATL1 enzyme are inhibited by high concentrations of Mg^{2+}, Ni^{2+} or Zn^{2+} but GLYATL1 might be more prone to inhibition factors causing lower enzyme activities (Section 4.6.6., 4.11.). Indeed, several factors influence the DTNB assay (Figure 5.1.).

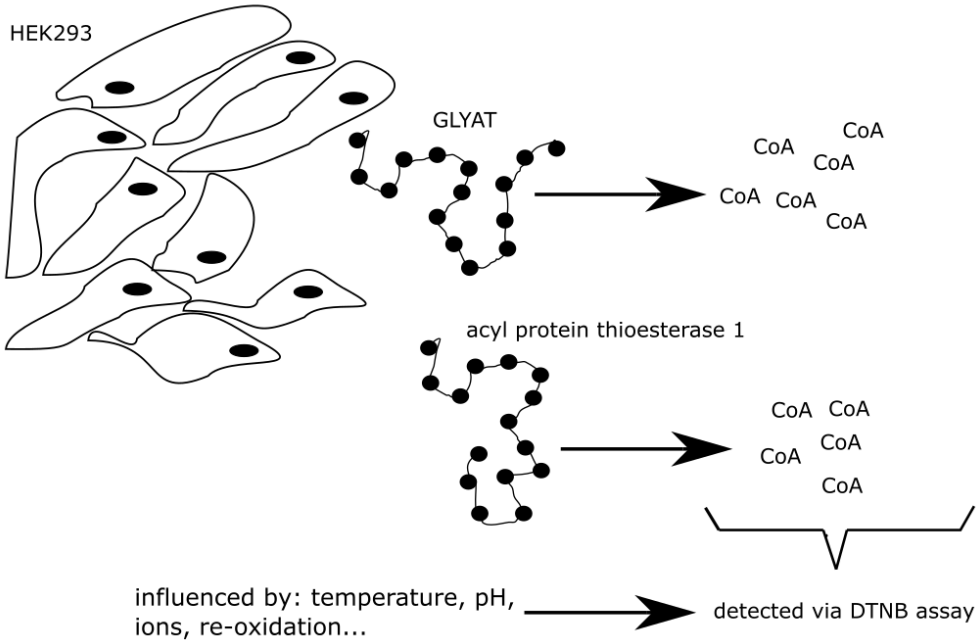


Figure 5.1: DTNB assay in HEK293 cell homogenates: HEK293 cells release coA via GLYAT and acyl protein thioesterase 1. The DTNB measurement is also affected by several other factors.

Regardless of this, the DTNB assay is well-established and was selected due to several reasons. As it was used for activity determination of human GLYAT several times before [107, 129, 152], it was performed due to re-examination and validity confirmation. Notwithstanding, the simplest reason for the assay is the color reaction with released coA groups [46], because they are released by human GLYAT and GLYATL1. Furthermore, the DTNB assay was easy to handle and cost-effective serving as a first test to check human GLYAT activity. Moreover, numerous studies underlined the overall relevance and distribution of this cost-effective, quick, and reliable assay [5, 40, 107, 152, 169].

For a prospective direct enzyme activity with determination of absolute hippuric acid amounts the HPLC-ESI-MS/MS would be a suitable alternative [154]. Another option might be the quantitation of the target protein activity via chromogenic or fluorogenic peptide substrates. Via these analyses kinetic mechanisms can be described in more detail as it was done for the therapeutically relevant cathepsine inhibitors [54]: these approaches might be useful for further research on GLYAT

reaction mechanisms or inhibitor studies since they might arise kinetic properties and protein-ligand recognition processes.

5.4.2. Activities of human GLYAT wild-type and sequence variants purified from *E. coli*

The FPLC purified human GLYAT wild-type demonstrated a v_{max} of 490 ± 274 U/mg protein and a K_M (Benzoyl-coA) of 79 ± 38 µmol/L. The investigation of FPLC purified GLYATL1 was cancelled due to enzyme loss after dialysis and focusing on the spin-column purification thereafter.

The GLYAT wild-type, isolated with spin-columns, exhibited v_{max} of 186 ± 168 U/mg protein, whereas high data variation between the series are obvious (Section 4.6.5.). The series were only reliable within one series (intra), whereas the interassay comparison was meaningless. As a representative series, the GLYAT wild-type demonstrated a v_{max} of 411 U/mg protein and K_M: 81.3 µmol/L (Figure 4.37.), which resemble the values obtained through FPLC purification. Overall, the v_{max} values of the spin-column purified enzymes were lower than the FPLC samples. The K_M values were comparable.

The present study demonstrated 100 x higher activities of human GLYAT than van der Sluis and colleagues, although their approach was very similar [152]. However, the dialysis procedure in the present study was different as it was performed in a tubing at 4 °C o/n under stirring, while van der Sluis used a desalting column. Nevertheless, this difference may not serve as explanation for 100 x higher enzyme activities.

There is large variation in kinetic data for human GLYAT wild-type as it was summarized by van der Sluis et al. [152] and in a more recent study [129].

The kinetic parameters of human GLYAT were extracted from literature data (Table 39).

Table 39: Molecular mass and measured kinetic parameters of the canonical GLYAT wild-type gathered from literature

	PARAMETER	SOURCE
PROTEIN SIZE (KILODALTON)	27	[156]
	30	[108]
	30.5	[85]
	33.9	NM_201648.2 reference
	34	[152]
K_M (BENZOYL-COA) [µMOL/L]	13	[156]
	67 +/- 5	[85]
	209	[107]
	57,900	[108]
	24 +/- 3	[152]
V_{MAX} [U/MG PROTEIN]	543 +/- 21	[156]
	807	[107]
	17,100	[108]
	730 +/- 30	[152]

The K_M value for benzoyl-coA is reported between 13 µmol/L and 57.9 mmol/L [8, 60, 84, 85, 90, 108, 152, 155]. However, the K_M value of 57.9 mmol/L [108] as upper limit was very suspicious and already identified as literal error. The determined K_M (Benzoyl-coA) of 79 µmol/L falls into the indicated range from literature. Maximum activity values published for human GLYAT wild-type are also highly variable as summarized by van der Sluis [152]. The v_{max} lies between 543 and 17,100 nmol/(min*mg). Nonetheless, our K_M value of 79 µmol/L is only half of the observed value by Matsuo et al. (209 µmol/L) [107], while our v_{max} of 490 U/mg is critically higher than theirs with merely 0.8 U/mg protein. Our procedure resembles that of Matsuo et al. but they used the different *E. coli* strain BL21 (DE3). For instance, the GLYAT activity was detected in bovine liver mitochondria with v_{max} of 32 U/mg, which is lower than ours by factor 100 and K_M of 15 µmol/L [5], which is also decreased compared to our value. Reported kinetic parameters for GLYAT prepared from bovine kidneys were K_M of 110 µmol/L and v_{max} of 17 U/mg protein [84], which strongly refers to our K_M value (79 µmol/L) but differs from our v_{max} by factor 29. Beyond that, our K_M value show high similarity to the one of GLYAT characterized from Rhesus monkey liver mitochondria with values of K_M: 67 µmol/L and highest determined activity of 110 nmol/(min*mL) [85]. While van der Sluis et al. considered their own data obtained with wild-type human GLYAT purified from the *E. coli* strain Origami 2(DE3) on the lower end of the literature data (K_M: 24 µmol/L, v_{max}: 0.73 U/mg protein), which are often based on the enzyme isolated from human liver [109], we report much higher activities and affinities than they did. Nevertheless, the reported v_{max} deviations (0.543 – 17.1 U/mg protein [107] – both for GLYAT prepared from human liver sample) are similar to quite high K_M deviations and the present study provides a new upper limit of v_{max} value (1359 ± 366 U/mg protein) for GLYAT p.(Asn156Ser)).

Nevertheless, the different enzyme preparations and experimental conditions may result in varying kinetic parameters. Beyond that, the quality of the used benzoyl-coA influenced the enzyme activity. The more instable the molecule is the lower would be the detected GLYAT activity. In the present study, benzoyl-coA from CoAlaBIO was used and the stability was tested via absorption spectra (data not shown) before activity measurements. This in turn, van der Sluis and Matsuo et al. did not provide any information about the company or stability of their coA ester.

To sum up, in most cases the determined K_M is comparable with most of the published data [84, 85, 152], while the present v_{max} exceeded the range in the literature. However, in agreement with the findings by van der Sluis et al. [152], we observed a considerable increase of the activity of the frequently occurring GLYAT variant p.(Asn156Ser) compared with the canonical wild-type. The present study corroborates the previous suggestion that the S_{156} haplotype should be considered as the real “wild-type” human GLYAT [101, 153].

The GLYAT variant p.(Gln61Leu), which accounts for about 12 % of the alleles in a Caucasian Afrikaner cohort from South Africa [153] was also analyzed. This variant had not undergone functional assessment before. The thesis revealed a decreased specific GLYAT activity (Sections 4.6.3., 4.6.5.) [137], which was a major new recognition. It suggested that this variant may play a role in phase II reactions in affected individuals, possibly resulting in pharmacogenetic peculiarities. It needs special consideration in South Africa where the isovaleric acidemia is very frequent [39]. This organic acidemia can be effectively treated with glycine and L-carnitine in combination [28, 102]. The rather high frequency of this sequence variant underlines that GLYAT sequence variants may indeed be of clinical relevance, although reported GLYAT sequence variants with functional consequences are mostly very rare [129, 152, 153]. Thereby, GLYAT studies going beyond this thesis and the publication [137] are indicated.

Certain sequence variants of GLYAT were analyzed in this project and negative prediction on enzyme activity was confirmed for GLYAT sequence variants p.(Gln61Leu), p.(His101Tyr) and p.(Ala231Thr) (Section 4.6.3., 4.6.5.), while the active center of human GLYAT was predicted to be located between aa 128 and 178 [107]. It is obvious that none of the sequence variants is present within this region. However, Matsuo and associates did not explain how they determined the active center. Critical amino acid exchanges, like polar glutamine to non-polar leucine or non-polar alanine to more polar threonine, are outside the active center and somehow influences GLYAT activity possibly by interaction with the active center motif.

The Michaelis Menten parameters (K_M, v_{max}) were calculated using the Excel Solver tool, which is applied for non-linear regressions. The tool was already used in 1995 for the fitting of non-linear regressions [167] and recommended for fitting of Michaelis-Menten curves.

There are numerous linear regression methods (Lineweaver-Burk [100], Eadie-Hofstee [45, 70] Hanes-Woolf [65]) to calculate v_{max} and K_M values from a Michaelis-Menten curve. The Lineweaver-Burk plot has the big disadvantage to be inaccurate for small substrate concentrations as it is described in a publication [89].

Figure 5.2. demonstrates the analysis of former presented data (with Excel Solver fitting) of the GLYAT sequence variant p.(Gln61Leu) (Figure 4.30.) in comparison to Lineweaver Burk, Eadie-Hofstee and Hanes Woolf plot.

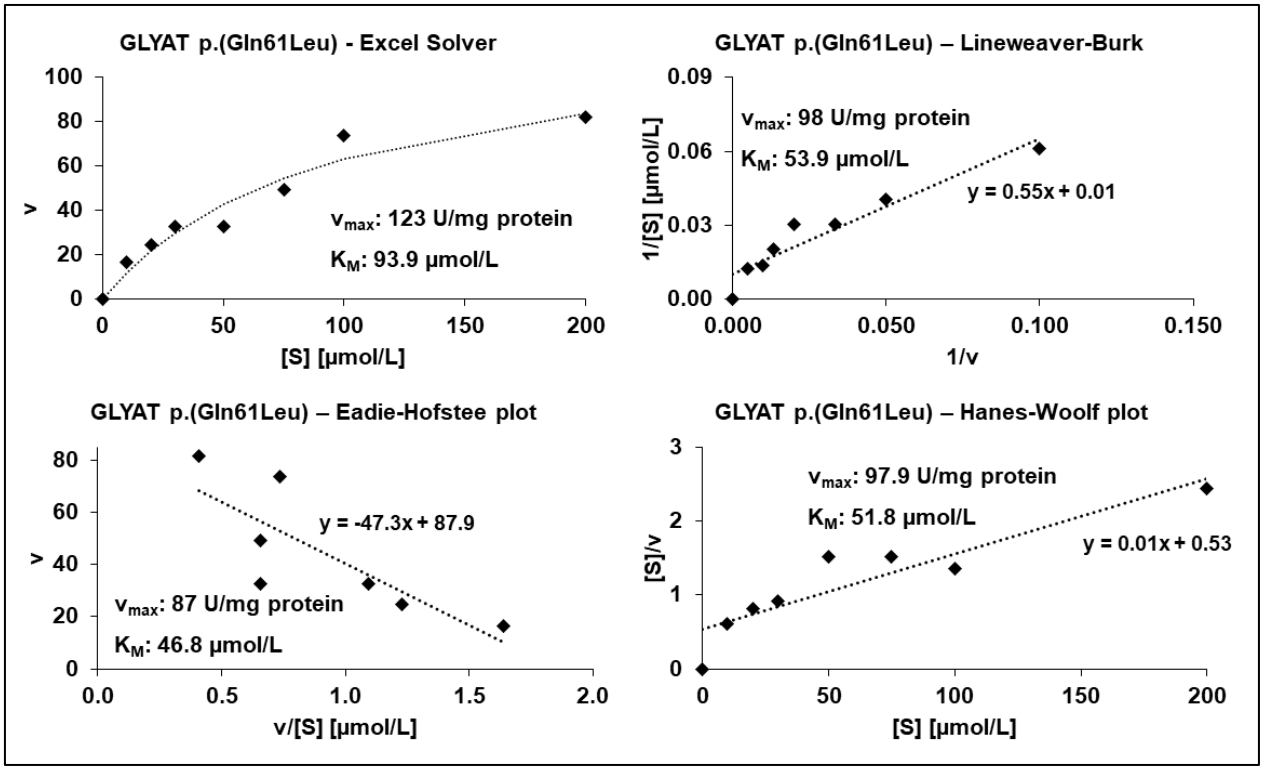


Figure 5.2: Comparison of kinetic analysis of purified GLYAT sequence variant p.(Gln61Leu) by using Excel Solver, Lineweaver-Burk, Eadie-Hofstee and Hanes-Woolf plots.

The observed kinetic parameters (v_{max}, K_M) of the applied linear regression methods resemble each other. However, there is a slight deviation in v_{max} and a stronger deviation in K_M in comparison to the Excel Solver result.

The Eadie-Hofstee method has the disadvantage of inaccuracy in the v_{max} determination as the v is part of both axes. The error increases with v/[S]. Since v enters at both coordinates, all deviations converge to the origin. For the linear function of the Hanes-Woolf graph the slope and ordinate intercept can be read directly from the function equation. However, the Hanes-Woolf plot has disadvantages, e.g., pipetting inaccuracies since the substrate concentration [S] is part of both axes. Errors in [S]/v are a far better approximation of the errors in v. Due to an unbiased spread of the measurement points along the [S]-axis, the result is in principle less biased by individual outliers. Nonetheless, since dependent and independent variables are mixed, the data optimization by linear regression is only conditional useful.

The Excel Solver tool stands out from the linearization fittings because it is a suitable model describing inherently non-linear data. The linearization process may introduce data distorsions, e.g., in standard deviations of logarithmically scaled data related to analyte concentrations [165]. Likewise, there is application in the current research since it has been used for providing a mathematical model for the estimation of water quality parameters [10].

5.4.2.1. Compatibility of glycine conjugation defects with life

No GLYAT deficiency patients are known so far and no studies on impacts of such GLYAT impairments are published. Likewise, it is not clarified whether these enzymes are directly connected to any metabolic disease. Only supportive functions in selected errors of metabolism are known thus far [8, 49, 143].

As described previously, certain *GLYAT* polymorphisms are known. However, selected deficiencies of the enzyme might not have been described yet and no phenotype has been connected to them. This leads to the assumption that they might be incompatible with life. It was also suggested that the glycine conjugation might be essential for life since a related metabolic defect has yet to be detected [129]. Due to the small number of drugs conjugated with glycine the interest in glycine conjugation faded shortly after its discovery [7]. Nonetheless, as van der Sluis et al. [153] and this thesis argued, the glycine conjugation appears to be an essential pathway without which humans are not viable. Due to its high conservation and the role in hepatic metabolism, the *GLYAT* gene was supposed to be essential for survival. Hence, the genetic variations in human amino acid *N*-acyltransferases and their consequences are of considerable interest.

Next, for the related enzymes *GLYATL1* and *GLYATL2* several SNPs with predicted protein damages are known without associated phenotypes [206]. In fact, a complete defect of *GLYAT* has yet to be reported. Table 40 demonstrates the numbers of homozygotes and heterozygotes calculated according to the Hardy-Weinberg principle [66, 170].

Table 40: Hetero- and homozygotes number of human GLYAT sequence variants determined from gnomAD database (GRCh38) [197] and the Hardy-Weinberg principle

Variant ID and amino acid exchange	rs number	Allele number	Number of heterozygous variant carriers	Number of homozygous variant carriers (calculated, expected)	Number of homozygous variant carriers (gnomAD browser)
11-58715393-C-A (p.(Gly38*))	rs746093728	146,802	1	1.2×10^{-8} (≙ 0)	0
11-58715336-C-A (p.(Val57Phe))	rs765808204	152,100	2	1.5×10^{-7} (≙ 0)	0
11-58712882-A-G (p.(Met65Thr))	rs145971997	152,160	41	4.5×10^{-7} (≙ 0)	0
11-58712775-G-A (p.(His101Tyr))	rs748514292	152,134	7	1.2×10^{-8} (≙ 0)	0
11-58710611-T-C (p.(Asn156Ser))	rs675815	152,168	1,760	0.462 (≙ 70,302)	72,095
11-58710062-G-A p.(Arg199Cys)	rs138125182	152,182	25	2.8×10^{-7} (≙ 0)	0
11-58477439-C-T p.(Ala231Thr)(only GRCh37)	rs768718220	250,936	1	0	0

Two example calculations for the p.(His101Tyr) and p.(Asn156Ser) sequence variants are shown below:

Hardy-Weinberg principle: $p^2 + 2pq + q^2 = 1$

(p^2: homozygous dominant (AA), pq: heterozygous (Aa), q^2: homozygous recessive (aa))

Determination of homozygotes on the basis of the heterozygotes (p.(His101Tyr)):

Heterozygotes frequency = $2p*(1-p) \triangleq 2pq$

Heterozygotes of p.(His101Tyr) (extracted from gnomAD browser [197]): 7

- ➔ *Heterozygotes frequency (2pq) (Aa): 7:152,134 = 4.6×10^{-5}*
- ➔ *Homozygotes frequency (p^2) (AA): 152,134-7 = 152,127 ➔ 1521,27:152,134 = 0.999953987*
- ➔ *Calculation of homozygotes frequency (q^2) (aa):*
- ➔ $p^2 + 2pq + q^2 = 1$
- ➔ $0.999953987 + 4.6 \times 10^{-5} + q^2 = 1$
- ➔ $q^2 = 1.2 \times 10^{-8}$ (= 0.000000012 ($\triangleq$ 0))

Determination of homozygotes on the basis of the heterozygotes (p.(Asn156Ser)):

- ➔ *Heterozygotes frequency = $2p*(1-p) \triangleq 2pq$*
- ➔ *Heterozygotes of p.(Asn156Ser) (extracted from gnomAD browser [197]): 1,760*
- ➔ *Heterozygotes frequency (2pq) (Aa): 1,760:152,168 = 0.012*
- ➔ *Homozygotes frequency (p^2) (AA): 152,168 – 72,095 = 80,073 -> 80,073:152,168 = 0.526*
- ➔ *Calculation of homozygotes frequency (q^2) (aa):*
- ➔ $p^2 + 2pq + q^2 = 1$
- ➔ $0.526 + 0.012 + q^2 = 1$
- ➔ *q^2 = 0.462 -> 152,168 x 0.462 = 70,302 homozygote individuals (aa)*

The calculations were performed to verify, if there are less homozygotes than expected. If this would be the case, the homozygosity would have severe disadvantages and be possibly incompatible with life. However, due to the low number of heterozygous individuals, this would not be expected. Notwithstanding, this law is a theoretical model, which do not match 100 % to the reality, because mutation, natural selection, non-random mating, genetic drift and gene flow cause deviations [1].

It is noticed that the determined number of homozygous corresponds to the extracted numbers from the gnomAD database. Nonetheless, the sequence variants p.(Met65Thr), p.(His101Tyr) and p.(Asn156Ser) are exceptional. Most of the variants are very rare and thus no homozygous

individuals are expected. Notwithstanding, the low allele frequencies of the sequence variants (from the databases) matches the low numbers that were calculated.

Regardless of this, the possible incompatibility with life warrants further research. So far, no genetic defect of glycine conjugation in humans has been reported in the literature. That matches the fact that *GLYAT* gene is well conserved and the deleterious mutations occure at low frequencies [129]. To further investigate the essential role of glycine conjugation pathway an *GLYAT* knockdown *in vivo* mouse model could be produced. The essential nature of this pathway might also explain why no metabolic defect has been described.

The human GLYAT was downregulated in hepatocellular carcinoma cells [107]. A more recent study described the association of GLYAT to breast cancer [150]. GLYAT was identified as antioncogene associated with malignant clinicopathological features that enhances breast cancer metastasis in mice. As the study by Tian et al. is the only one treating GLYAT in breast cancer patients, further research is required to illustrate this phenomenon.

5.4.3. Glycine dependency and reaction mechanism of wild-type human GLYAT

The human GLYAT requires glycine for exhibition of full activity (Figure 4.33.). GLYAT needs both substrates, glycine and benzoyl-coA, for execution of the entire enzyme function. Hence, the GLYAT activity was confirmed and spontaneous degradation of benzoyl-coA was excluded as a relevant contributor of the release of free coA. The glycine was used at 200 mmol/L (fixed concetration). Various publications that performed the same assay with the same enzyme [107, 152] served as an orientation. While van der Sluis et al. determined only the K_M (benzoyl-coA) (24 µM) in the study from 2013 [152], Matsuo and colleagues determined a K_M of 26.6 mmol/L for glycine [107]. Nevertheless, van der Sluis used 200 mmol/L as a fixed glycine concentration in the assay. In previous studies the glycine concentration was fixed to 50 mmol/L [85].

Thus, the used concentration of glycine (200 mmol/L) exceeded the published concentrations and represents a multiple of the determined K_M value.

The description of reaction mechanism of human GLYAT has changed over the time in the literature: a sequential 2-substrate mechanism was described in five studies [90, 107, 116, 155] ((1) acyl-coA binds (2) glycine binds (3) coA released (4) peptide product released) (Figure 5.3.). Another publication exhibited a non-sequential reaction mechanism (ping-pong mechanism) by kinetic studies [154] – after acyl-coA binds, coA is released directly before glycine is added and peptide product is released (Figure 5.3.).

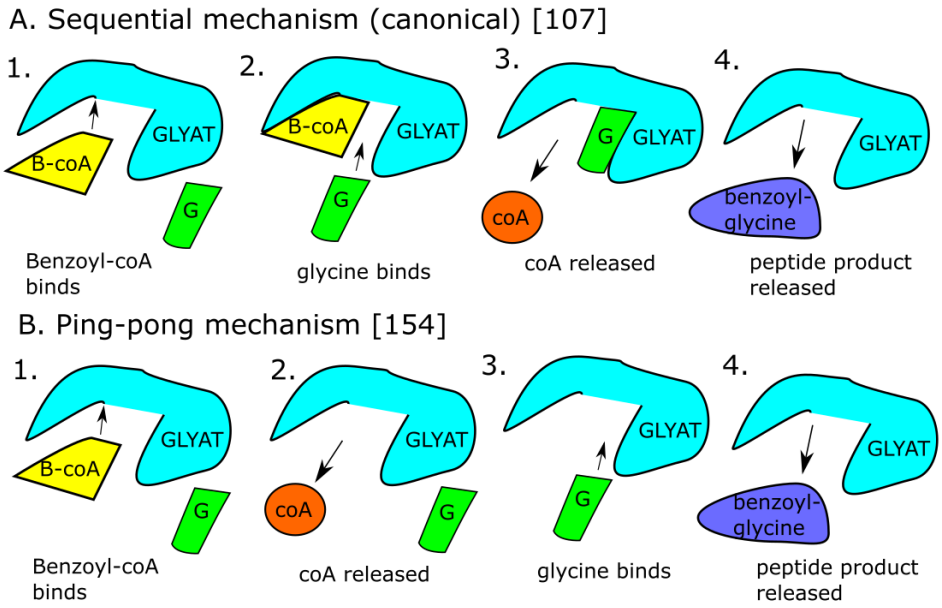


Figure 5.3: Hypothesized GLYAT reaction mechanisms. A: Sequential mechanism: First the two substrates, Benzoyl-coA (B-coA) and glycine (gly) bind to GLYAT before the products, coA and hippuric acid, are released. B: Ping-pong mechanism: Benzoyl-coA binds and coA is released first before glycine binds and peptide product, hippuric acid, is released (postulated reaction mechanisms by [107, 154]).

This canonical reaction mechanism was expanded by HPLC-ESI-MS/MS method [154]: they selectively identify and precisely quantify formed hippuric acid in the presence of varying substrate concentrations. The main difference to the usual DTNB assay, which is only an indirect activity quantitation, is the absolute quantitation of formed hippuric acid. However, they were able to justify a ping-pong reaction mechanism of human GLYAT. The present thesis demonstrated that glycine and benzoyl-coA are both required for enzyme activity of GLYAT (Section 4.6.4.). In absence of glycine, coA cannot be produced (Figure 4.33.) and in absence of benzoyl-coA and glycine, the hippuric acid (benzoylglycine) cannot be produced (Figure 4.34). This might rather be an indication for the sequential mechanism since both substrates are necessary for product synthesis.

Conversely, the kinetic determined by van der Sluis [154] fit the Hill plot, which is a sign for allosteric sigmoidal enzyme kinetics. Indeed, a few results of the present study for the GLYAT wild-type demonstrated sigmoidal curve courses in approach (Figure 4.38.). These courses were not identified in the other series, where rather Michaelis Menten saturations were measured. The sigmoidal course was only identified for the purified GLYAT wild-type and thus seems to be exceptional. However, this study might confirm the allosteric sigmoidal enzyme kinetic module for GLYAT described by van der Sluis [154]. Sulfotransferases follow the Michaelis Menten kinetic when applying narrow ranges of substrate concentrations. Nonetheless, when these concentration ranges were expanded, they demonstrated allosteric sigmoidal kinetics accordingly [154]. Another

very predominant example for allosteric sigmoidal kinetics is hemoglobin [11]. Due to the unknown X-ray structure of human GLYAT, it is difficult to derive the subunit interactions and their influences on the kinetic model. Nonetheless, a structure prediction model was published by the Uniprot protein database (Figure 1.2.).

Sigmoidal course was also obtained for increasing glycine concentrations and fixed benzoyl-coA, which confirmed that glycine binding enhances benzoyl-coA binding (very high affinity of GLYAT-glycine interaction) [154]. Moreover, glycine availability limits the reaction of GLYAT. Mechanistic kinetic cooperativity was firstly identified by Ferdinand, who demonstrated non-hyperbolic rates for the phosphofructokinase, which is another 2-substrate enzyme [50]. This enzyme kinetic is therefore alternatively titled as Ferdinand mechanism. Thereby, substrates may activate or inhibit enzyme activity. For instance, van der Sluis and associates demonstrated a substrate activation with sigmoidal plot curves, when they increased glycine concentration with constant benzoyl-coA concentration. The higher the glycine concentration was, the more hippuric acid was formed. Conversely, a substrate inhibition was identified with maximum passing plot when benzoyl-coA concentration was increased and glycine kept constant. That was independent from benzoyl-coA concentration and the plot always reached the plateau at constant hippuric acid forming rate [153]. In their study, a positive kinetic cooperativity, which was demonstrated by sigmoidal curves, was confirmed for varying benzoyl-coA concentrations. Conversely, for varying glycine concentrations, the same was observed for concentrations below 5 mM – for higher concentrations the curve matched the Michaelis-Menten equation. The more recent study from 2021 identified the kinetic cooperativity as preferred reaction mechanism of human GLYAT [129]. One substrate binds to the enzyme first and the enzyme shows preference for one of the substrates.

In summary, in the present study sigmoidal curve course was predicted to be exceptional and the Michaelis Menten courses were identified as the usual and expected result.

5.4.4. Activity of human GLYATL1 wild-type and sequence variants purified from *E. coli*

The enzyme activity data of the purified GLYATL1 are scarce in the literature, although few studies treated the enzyme from mitochondrial fractions of rhesus monkey and man [169] and from primate liver mitochondria [85]. The most suitable overexpression conditions for human GLYATL1 in *E. coli* were determined at an induction temperature of 28 °C without IPTG supplementation (Figure 4.7.). Thereby, the IPTG addition seemed to impair the overexpression of the target protein, which was confirmed in addition by an induction profiling for several recombinant target proteins [115]. The elution fraction 1 of FPLC purified GLYATL1 resulted in 107.5 U/mg protein (Figure 4.40.) but, due to the loss of the enzyme after dialysis (Figure 4.18) and better comparison

to the sequence variants through parallel processing, the isolated GLYATL1 via FPLC was not analyzed further.

While the spin-column purified GLYATL1 wild-type showed v_{max} of 10.3 ± 3.8 U/mg protein (Section 4.6.6.) the values in the literature were decreased: Webster et al. determined glutamine transferase activity of 0.05 U/mg protein [169], while Kelley & Vessey demonstrated a value of about 0.07 U/mg protein [85]. They distinguished between mitochondrial glycine and glutamine *N*-acyltransferase activities that showed an increased activity for the glycine conjugation. However, they used partially purified enzyme from rhesus monkey [169] or human mitochondrial lysate [85], which complicates the comparison to the purified sample from overexpression in *E. coli*.

The determined K_M (phenylacetyl-coA) value of 96.3 ± 33.9 µmol/L (Section 4.6.6.) for the purified enzyme resembled that obtained by Webster et al. for the glutamine *N*-acyltransferase activity (35 µmol/L for phenylacetyl-coA) and the one from Kelley & Vessey (12 µmol/L for phenylacetyl-coA). The conditions in the activity assay were similar to Webster et al. as they used 150 mmol/L glutamine. Notwithstanding, the high standard deviation of 33.9 µmol/L (variation coefficient: 35 %) underlined the K_M sensitivity to variations, especially for low substrate concentrations. In the present study, a fixed concentration of L-glutamine was used (150 mmol/L). This was chosen based on literature data [169], which indicated a K_M (L-glutamine) of approximately 600 mmol/L. Nonetheless, they described a 2-3x variation of that value and have used a L-glutamine concentration of 150 mmol/L, which served as an orientation.

The GLYATL1 sequence variants p.(Glu63Lys), p.(Val125Leu) and p.(Asp255Asn) demonstrated kinetic properties, which resemble the wild-type (Section 4.6.6.). A decreased enzyme activity would be expected according to the prediction tools (Section 3.2., Table 17), which is not confirmed by the activity measurements. In particular, the charge critical exchange of aspartate to asparagine (p.(Asp255Asn)) outside the active center (aa 128 – 180 [107]) may impaire the GLYATL1 activity by most likely conformational change of protein structure. Impaired protein stability caused by missense mutations was already described exemplary for receptor tyrosine kinases and cancer development [143]. Today, databases can deliver information on the association between SNPs of proteins and pathological conditions [24].

Only minor activities were observed for the GLYAT and GLYATL1 stop variants p.(Gly38*) and p.(Glu87*) in the *E. coli* homogenates (Figure 4.32.).The stop mutations result in a C-terminal truncated protein with loss of function because of lacking active center (human GLYAT: aa 128 – 178, human GLYATL1: aa 128 – 180) [107].

In summary, the v_{max} result for the purified GLYATL1 wild-type differed from the literature data, whereas the K_M value resembled the latter. This meet the expectations due to observed

differences in the purity grade of the purified enzyme (increased v_{max}, similar K_M). The predicted enzyme impairments for the GLYATL1 sequence variants (Section 3.2.) were not confirmed by enzyme activity determinations since they resemble the wild-type activity.

It was obvious that the activity of purified GLYAT is higher compared to the activity of purified GLYATL1. The direct comparison between GLYAT and GLYATL1 are absent for the purified proteins, but in human cell homogenates a few studies are present [84, 85, 169]. As described above, Webster et al. determined a glycine *N*-acyltransferase activity of 0.1 U/mg protein, whereas the glutamine *N*-acyltransferase activity is detected as 0.05 U/mg protein [169]. This meets the observations of the thesis as the determined v_{max} of purified human GLYAT is higher than GLYATL1. The same was concluded by Kelley & Vessey in 1994: they determined the highest activity of the aralkyl transferase (glycine conjugation) followed by the activity of the arylacetyl transferase (glutamine conjugation) (1 U/mg protein) [85].

A more recent study described the downregulation of human GLYATL1 in hepatocellular carcinoma [62]. The enzyme was supposed to be an independent prognostic factor for hepatocellular carcinoma patients. The promotor methylation level of *GLYATL1* was significantly higher in hepatocellular carcinoma tissue compared with other tissues indicating regulated decrease of expression. Interestingly, Guan et al. suggested that downregulation of *GLYATL1* in hepatocellular carcinoma is associated positively with xenobiotic metabolism and expression of genes involved in mitochondrial glutamine metabolism, like *SLC1A5*, *SLC1A11* etc. Indirectly, GLYATL1 catalyzes glutamine and affects its metabolism in mitochondria, thus inhibiting rapid proliferation of hepatocellular carcinoma. Thereby, glutamine, which is also a substrate of GLYATL1 [107] seems to be a product accordingly.

5.4.5. Activities of GLYAT and GLYATL1 in HEK293 cell homogenates

The enzyme activities of overexpressed human *N*-acyltransferases in bacterial and human system differed, as lower specific activities for the human system HEK293 are obvious (GLYAT: v_{max}: 95.6 ± 69.6 U/mg protein) (Table 38).

The enzyme activity tests in HEK293 cells were performed directly after cell disruption to avoid protein degradation in the homogenate supernatants.

High data variations were obvious due to interfering activities in crude cell homogenates (Section 5.4.). Notwithstanding, the sequence variants p.(Gln61Leu), p.(His101Tyr) and p.(Ala231Thr) indicated markedly reduced enzyme activities (Section 4.10.), which confirmed the prediction of enzyme activity (Section 3.2.). However, the sequence variants p.(Arg199Cys), p.(Val57Phe) and p.(Met65Thr) could not be overexpressed, which resembled the bacterial data. In eukaryotes, the

PTMs may influence the expression pattern of recombinant enzymes [24], which can be critically influenced by the introduction of mutations [143].

The GLYATL1 enzyme showed maximum activity of 30.7 U/mg protein for 100 µmol/L phenylacetyl-coA in HEK293 cell homogenates (Figure 4.64.). The K_M could not be detected due to an absent correlation between the activity and substrate concentration. Previously discussed problems of Ellman´s assay with crude cell homogenates (Section 5.4.) were of consequence here. The studied sequence variants p.(Glu63Lys), p.(Val125Leu) and p.(Asp255Asn) revealed no suspicious activity differences to the wild-type. Most probably the slight differences between variants got lost in the severe variations the assay.

The GLYATL1 activity assay with HEK293 homogenates (Section 4.11.) observed critically lower specific activities than the GLYAT assay (Section 4.10.). This might be due to different substrates, varying conjugation efficiencies and enzyme availabilities.

The v_{max} values of the FPLC and spin-column purified human GLYAT strongly exceed the ones observed from HEK293 homogenates. Various other proteins, which reduce DTNB are present in the homogenates, for instance acyl protein thioesterase 1 [69], while the purified proteins should contain nearly none of background proteins explaining the higher specific activity.

The *in-silico* PTM data (Section 4.7.) showed possible PTMs for all three enzymes (GLYAT, GLYATL1, GLYATL2) and laid the foundation for the overexpression in HEK293 cells. Usually, the HEK293 cells should provide authentically modified overexpressed target proteins [149] with highest possible activity. But, due to the low expression rates and high background activities of the Ellman assay within homogenates [40] (Section 4.10., 4.11.), those higher activities were not observed. For the bacterial homogenates, the maximum enzyme activities were also determined (Figure 4.32.) with much higher signals compared to HEK293 cells. Thioesterases with coA ester reducing activity are continuously expressed in *E. coli* [16] that might explain the high background activities in *E. coli* homogenates, which was also observed for the pET32a(+) empty homogenates (data not shown).

Interestingly, HEK293 cells can hardly compensate the high GLYAT background activity, while for bacteria it is much more obvious. Explained by lower expression level of human GLYAT and GLYATL1 in HEK293 cells this fact falls into calculation.

Because of these two main reasons, the low expression rates, and the high background activities in HEK293 cells, *E. coli* was selected as the main activity testing system. However, the data variations were indispensable high and the inter-assay variations was severely higher (Figure 4.26.) than the intra-assay comparison (Figure 4.24.). These data variations are strongly based on differences in growing behavior between different *E. coli* Origami 2(DE3) cultures. As described

in the literature [22, 124] and in Section 4.4., this *E. coli* strain exhibits unpredictable growing rates and sometimes it did not grow in selection medium even when transformed with resistance-delivering plasmid. The growing rates and cell conditions should be considered here because they might directly influence enzyme assembly and activity. The same was obvious for the GLYATL1 analyses with much lower specific activities and more severe data variations (Section 4.6.6.).

5.5. GLYAT and GLYATL1 in inborn errors of metabolism and other diseases

Impaired phase II detoxification enzymes are associated with a variety of diseases ranging from urea cycle defects, organic acidurias and even adverse reactions to drugs that might be involved in cancer [118]. Beyond that, it was shown that the GLYAT expression was downregulated in hepatocellular carcinoma specimens [107] followed by the same observation for GLYATL1 eight years later [62]. The challenge of scarce information on physiological implications of glycine conjugation is problematic as benzoate is widely used as a preservative in food [122, 129]. Nevertheless, the glycine conjugation pathway is essential for humans [7, 129] which highlights the relevance of identification of corresponding metabolic defects or patients. For instance, the salicylate metabolism needs to be studied further as it depends on an initial conjugation with coA and the subsequent acyl-transfer to glycine. The role of benzoate in food and associated recommendations for healthier food without benzoate might allow conclusions on risks of such dietary additives. IVA patients might observe benefits from those considerations and therapy adaptations.

As GLYAT and GLYATL1 conjugate coA-esters to amino acids, they provide alternative pathways of human body detoxification in inborn errors of metabolism. Thereby, they compensate the toxic effects of hyperammonaemia (within urea cycle defects) or ketoacidosis (within organic acidemias).

5.5.1. GLYAT and GLYATL1 in urea cycle defects

Enzyme deficiencies that are involved in the urea cycle might be associated with hyperammonaemia and increased glycine and glutamine concentrations in plasma [144]. Thus, increasing amino acid concentrations were treated with a low-protein diet for many years. However, recently compounds increasing waste nitrogen by alternative pathways were identified. Urea cycle disorders, i.e. CPS and OTC deficiency, inescapably lead to increasing ammonia concentration in the blood. Compounds that increase waste nitrogen removal, resulting from glycine and glutamine accumulations, are sodium benzoate and phenylbutyrate which represent the GLYAT and GLYATL1 substrate, respectively. Benzoate is conjugated with glycine to form benzoylglycine, which is rapidly excreted in urine [49, 144]. Conversely, the administration of phenylbutyrate, as a congener of phenylacetate with unpleasant and clinging odor, adversely

initiates conjugation with glutamine to form phenylacetylglutamine enabling the removal of nitrogen [21].

The GLYAT sequence variations impair enzyme function as demonstrated for p.(Gln61Leu), p.(Met65Thr), p.(His101Tyr) and p.(Ala231Thr) (Figure 4.32., 4.37.). Patients with these sequence variations are rather treated with phenylbutyrate as they cannot be treated with sodium benzoate for nitrogen excretion (Figure 5.4.). Patients occupying mutations in the *GLYATL1* gene, such as p.(Gly63Lys) (Figure 4.43., 4.64.), are otherwise recommended to be administered with sodium benzoate as they possibly cannot use phenylbutyrate (Figure 5.4.).

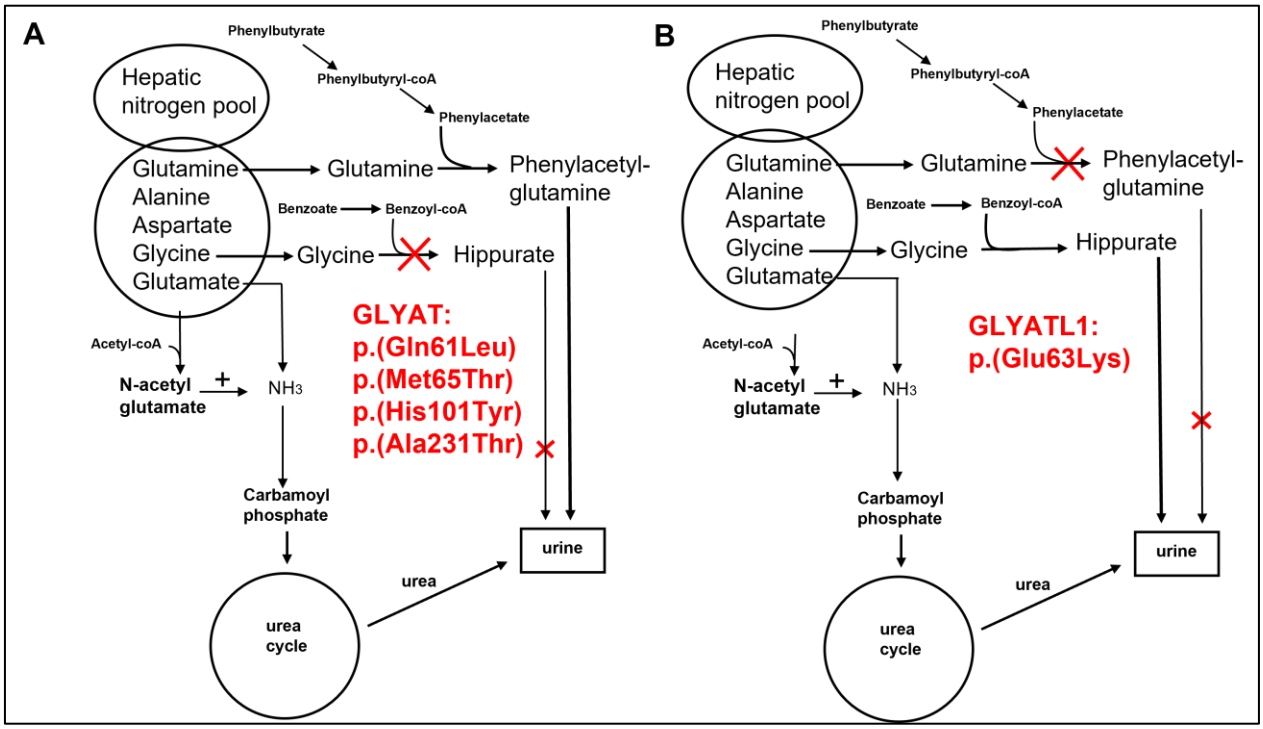


Figure 5.4: Urea cycle scheme and influence of GLYAT (A) and GLYATL1 polymorphisms (B). The GLYAT sequence variations p.(Gln61Leu), p.(Met65Thr), p.(His101Tyr) and p.(Ala231Thr) avoid alternative nitrogen excretion pathway via sodium benzoate administration. Thus, the therapy via phenylbutyrate is recommended. GLYATL1 sequence variation p.(Glu63Lys) avoid alternative pathway of phenylbutyrate metabolism, thus sodium benzoate therapy might be favoured.

As demonstrated by Figure 5.4. the red marked GLYAT/GLYATL1 sequence variations affect alternative nitrogen excretion in a way, that either benzoyl-coA or phenylacetate conversion is impaired. The corresponding other pathway is preferred thereby. The observed enzyme activity data for the remaining analyzed GLYAT ((p.Val57Phe)) and GLYATL1 sequence variants (p.(Val125Leu), p.(Asp255Asn)) were not reliable. The target enzyme was not expressed in case of p.(Val57Phe) sequence variant. Referring to the enzyme activity predictions (Table 17), the named sequence variants should lead to an enzyme deficiency. Notwithstanding, missense

mutations might critically influence the expression of proteins [143], which might explain the absence of target protein in these cases.

5.5.2. GLYAT and GLYATL1 in organic acidemias

Isovaleric acidemia (IVA)

Human GLYAT may also provide alternative pathways of detoxification in toxic acidosis as it conjugates toxic accumulating acyl-coAs with glycine [8]. The *N*-isovalerylglycine was first identified in human urine in 1967 [146]. For instance, IVA patients are to be treated with glycine initiating formation of excretable and less toxic isovalerylglycine [92]. IVA leads to accumulation of isovaleric acid as well as secondary metabolites, such as isovalerylglycine (through GLYAT conjugation), isovalerylcarnitine and 3-hydroxyisovaleric acid. One missense mutation, i.e. p.(Ala282Val), is particularly common in IVA patients, identified via newborn screening with mild metabolite elevations and without symptoms [163]. Further mutations in the gene of the responsible enzyme (isovaleryl-coA dehydrogenase) were already identified in 1991 [162]. The clinical outcome of this mutation and the effects on therapy remained unclear. IVA is one of the most common inborn errors of metabolism in South Africa, where 10 patients were found to be homozygous for p.(Gly123Arg) mutation within the isovaleryl-coA dehydrogenase gene [39]. Indeed, the very frequent GLYAT polymorphism p.(Gln61Leu) was identified in the South African cohort accordingly [153] with impaired enzyme activity [137]. The high prevalence of IVA in South Africa combined with their impaired and frequent p.(Gln61Leu) variant underline the high majority of glycine conjugation for this cohort from a pharmaceutical point of view. Nevertheless, the general pharmaceutical interest is low due to small number of drugs conjugated with glycine [7].

Another study demonstrated similar effects of L-carnitine administration for IVA patients compared with glycine therapy regarding the removal of isovaleryl-coA: in particular, L-carnitine decreased the plasma isovaleryl-coA level [128]. L-carnitine serves as an alternative therapy compound to glycine, when GLYAT is impaired via missense mutation (Figure 5.5.).

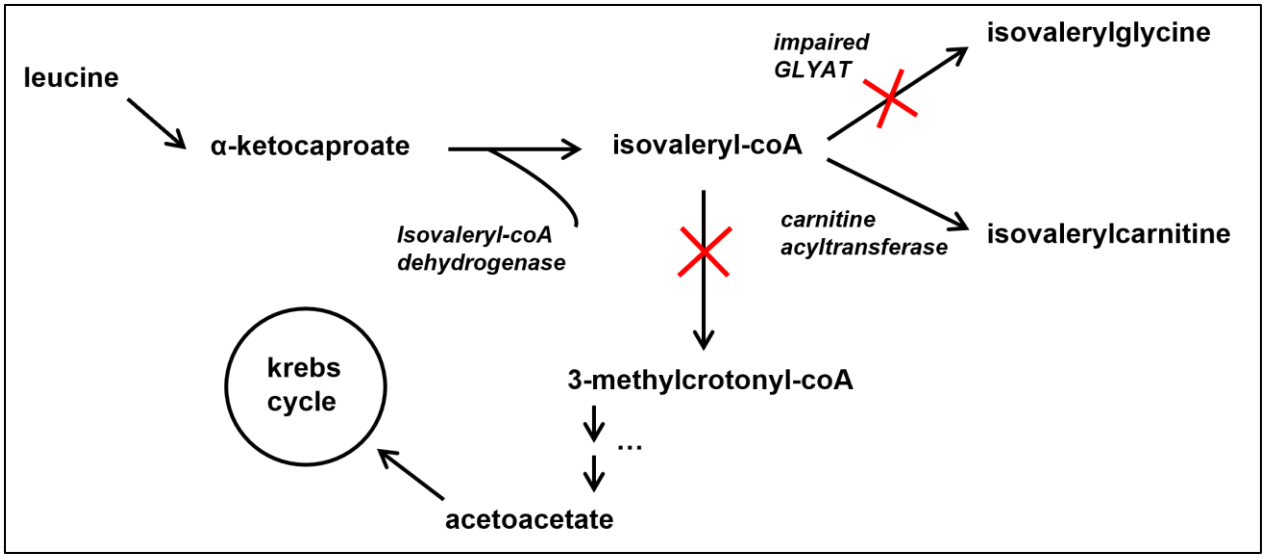


Figure 5.5: Impaired GLYAT affects the isovaleryl-coA conjugation with glycine. L-carnitine treatment initiates the conversion of isovaleryl-coA to isovalerylcarnitine by carnitine acyltransferase. IVA is characterized by deficiency of isovaleryl-coA dehydrogenase, thus blocking metabolic pathway of leucine degradation and accumulation of isovaleryl-coA. GLYAT impairing polymorphisms such as p.(Gln61Leu), p.(Met65Thr), p.(His101Tyr) or p.(Ala231Thr) avoids IVA therapy via glycine supplementation, because glycine cannot be metabolized. Those patients should rather consume carnitine to produce isovalerylcarnitine via carnitine acyltransferase.

L-carnitine treatment has to be considered with regards to GLYAT sequence variations p.(Gln61Leu), p.(Met65Thr), p.(His101Tyr) and p.(Ala231Thr), which inhibit the GLYAT enzyme activity (Section 4.6.5.). Thus, they might influence the IVA therapy. GLYAT deficient patients should be rather treated with L-carnitine (Figure 5.5.), while patients with active GLYAT enzyme can metabolize L-carnitine to detoxify accumulating isovaleric acid. However, a therapy with glycine and L-carnitine supplementation was suggested accordingly [102]. Isovalerylcarnitine was detected in the urine of isovaleric acidemia patients via GC/MS in 1984 and later also in human neutrophils [123, 128]. Nonetheless, it was demonstrated in 1984 that the administration of L-carnitine is more effective in reducing plasma isovaleryl-coA than glycine therapy [123]. In a more recent study, the urinary isovalerylglycine levels have been increased 2-fold more through glycine supplementation in comparison to L-carnitine or with both compounds [28].

Nonetheless, L-carnitine is an effective treatment option for IVA patients, a dose of 100 mg/kg body weight had been shown to increase excretion of isovalerylcarnitine in urine. The combination of carnitine with glycine treatment maximizes the total isovaleryl-coA conjugate excretion during metabolic stress but not under stable conditions [53]. Nonetheless, L-carnitine may serve as an alternative for isovaleric acidemia treatment replacing glycine in the event of GLYAT deficiency.

Another example of organic acidemia is propionic acidemia (PA), i.e., autosomal recessive disorder within isoleucine degradation, where small amounts of toxic propionate is excreted as propionylglycine, which is catalyzed by GLYAT [9].
Propionic acidemia and methylmalonic acidemia (MMA) and their treatments were described in a recently published guideline [51]. In case of PA, tiglylglycine and propionylglycine is frequently detected in combination, which indicates a provision of alternative pathway for detoxification via GLYAT conjugation. Thus, the analyzed GLYAT sequence variants might influence the treatment of the propionic acidemia as propionylglycine has been identified in the urine of PA patients [60].

Nonketotic hyperglycinemia (NKH)

GLYAT might also be associated with nonketotic hyperglycinemia (NKH), a deficiency in the glycine cleavage system. Benzoate is administered for the treatment of this metabolic disorder [157] to reduce the plasma glycine level via GLYAT conjugation reaction. Thereby, the possibility of toxication via benzoate might occur due to GLYAT inactivity that is possibly triggered by sequence variations. In this case, benzoate accumulates in the human body, which results in toxicity. Known side-effects of excessive sodium benzoate concentrations are associated gastritis (500 – 750 mg/kg/day), L-carnitine loss, hypocalcemia and unpalatability [158]. For ketogenic diet which decreases glycine levels, the sodium benzoate level should also be reduced accordingly to avoid benzoate toxicity [35]. Hence, GLYAT sequence variations must be considered regarding toxicity, when benzoate is administered for NKH patients. However, for attenuated NKH, sodium benzoate used in combination with Dextrometorphan has an improving neurocognitive outcome with a decrease of propensity of seizures (an indication of NKH) [158]. Beyond that, a connection between glycine *N*-acyltransferase activity and NKH was identified in rats [104].

Methylacetoacetyl-coA thiolase deficiency (MATD)

Another example of an autosomal recessive disorder in the isoleucine degradation pathway is the 2-methylacetoacetyl-coA thiolase deficiency (MATD) caused by a mutation in the *ACAT1* gene. This mitochondrial enzyme catalyzes the conversion from acetoacetyl-coA to acetyl-coA in ketone body utilization [132]. The enzyme is moreover involved in the conversion of methylacetoacetyl-coA to propionyl-coA in isoleucine catabolism [61]. In the isoleucine degradation, the enzyme 2-methyl branched chain acyl-coA dehydrogenase (ACADSB) catalyzes the conversion of 2-metyhlbutyryl-coA to tiglyl-coA, which is also one step in isoleucine degradation. However, in the event of a ACADSB deficiency, the human GLYAT may also convert tiglyl-coA into tiglylglycine to decrease the level of toxic accumulating tiglyl-coA. However, tiglyl-coA was described as the most reactive substrate with a v_{max} of 33.3 µmol/min*mg for the bovine GLYAT [8] and needs to be

checked for the human enzyme. As an alternative tiglylcarnitine can also be produced by the corresponding L-carnitine transferase, the same alternative, which is valid for IVA as described above [39]. The main recommendation for patients is an isoleucine reduced diet and avoidance of fat excess to prevent metabolic decompensations. The majority of MATD patients is supplemented with L-carnitine to reduce the level of toxic accumulating coA esters. Due to the conversion of tiglyl-coA to tiglylglycine (as described above), glycine supplementation would also be an alternative therapy option.

So far, GLYAT or GLYATL1 deficient patients have not yet been reported and the impacts of these deficiencies have not been investigated. Functional consequences of *in vitro* studies of a highly prevalent p.(Gln61Leu) variant suggested pharmacogenetic implications. The search for partial GLYAT deficiency, possibly in certain heterozygotes, requires further study. Probably, due to their role in compensation of urea cycle defects [49] or certain organic acidemias as described above.

5.5.3. GLYAT and GLYATL1 in other diseases

The mitochondrial glycine synthetic pathway is strongly correlated to the proliferation of cancer cells [76]. Glycine metabolism contributes to mortality of breast cell cancer proliferation shown by expression of mitochondrial glycine synthesizing enzymes SHMT2, MTHFD2 and MTHFD1L. Summarizing the results by Jain and associates, the glycine metabolism represents metabolic vulnerability in proliferating cancer cells that can be targeted for therapeutic benefits. From another point of view, GLYAT is involved in hepatocellular carcinoma as is it strongly downregulated in cancerous rat cells shown by Western blot and RT-qPCR study [107] Thus, a relation of GLYAT to human liver cancer or other liver diseases is impressively confirmed. In more recent studies, the roles of human GLYAT and GLYATL1 were investigated in the context of cancer [62, 150]. Human GLYAT was decreased in human breast cancer tissues and a GLYAT knockdown augmented the cancer cell proliferation *in vivo* and *in vitro* [150]. However, that phenomenon was not confirmed when GLYAT was overexpressed in transfected cells. Nonetheless, the downregulation of the GLYAT expression in human breast cancer seemed to be correlated with epithelial-mesenchymal transition via the PI3K/AKT/Snail pathway and might be associated with histological grade, tumor TNM stage (Ki-67 status) [150]. In another recent study, the GLYATL1 enzyme was downregulated in human hepatocellular carcinoma cells [62], which first shed light on the prognostic value of GLYATL1 in cancer and hypothesizes potential regulatory mechanisms underlying liver cancer development.

GLYAT was also suggested as a biomarker for liver damages and cancer [107], which was planned to be checked in the present study in addition. Due to time consuming expression studies and related enzyme activity tests, this side project was cancelled. However, our results indicated

that certain GLYAT polymorphisms, i.e. p.(Glu61Leu) or p.(Ala231Thr), decreased enzyme activity. Especially the prevalent p.(Glu61Leu) polymorphism (12 % frequency in South Africa) demonstrated the unsuitability of GLYAT to work as biomarker for liver damages when it is not functional.

Moreover, it has been reported that the carbonic acid 4-aminobenzoate correlated with hepatic reserves in patients with hepatitis [6, 42]. Thus, the glycine conjugation rate might function as liver test. However, GLYAT may also serve as liver function marker due to its release into blood plasma in the event of liver damages. Thus, detected GLYAT activity in blood plasma would correlate with liver damages.

5.6. Tissue expression of human *GLYATL2* – other functions than detoxification?

Human GLYATL2 conjugates medium- and long-chain acyl-coA esters to glycine forming *N*-oleoylglycine and *N*-arachidonoylglycine. The activity is regulated by acetylation of Lys^{19} as mutation of this residue to arginine or glutamine resulted in 80 % activity decrease [162]. The GLYATL2 products, i.e. *N*-oleoylglycine and *N*-arachidonoylglycine, are structurally related to endocannabinoids and other signaling molecules that regulate functions like pain perception, body temperature and anti-inflammatory properties [73, 110, 166]. Thus, GLYATL2 might be involved in processes beyond detoxification.

5.6.1. Increased abundance of GLYATL2 in testis and brain – *N*-oleoylglycine as signaling molecule?

The mRNA signal of human GLYATL2 seemed to be upregulated within testis and brain demonstrated by semi-quantitative PCR (Figure 4.65.).

GLYATL2 reaction product *N*-oleoylglycine was suggested to be involved in the biosynthesis of oleamide, a primary fatty acid amide [34, 166]. It has cannabinoid-like features and contributes to mechanisms that induce sleep [27]. It was isolated from cerebrospinal fluid of sleep-deprived cats and identified as natural constituent of human cerebrospinal fluid. Moreover, the other preferred GLYATL2 substrate, arachidonic acid [165], is present in bovine and rat brain identified by ion trap mass spectrometry analysis [73]. Therein, it suppresses tonic inflammatory pain. Considering that, GLYATL2 might also be involved in n the synthesis of brain signaling molecules. For instance, the acyl-coA synthetase forming arachidonoyl-coA from arachidonic acid and coA was detected in the brain, platelets, and aorta [73, 113, 173]. Followed by that arachidonoyl-coA conjugating enzyme (to glycine), GLYATL2, is most likely in proximity to that. Beyond that, structural and functional similarity of *N*-oleoylglycine and *N*-arachidonoylglycine to endocannabinoids with receptors that are present in the brain [86] would corroborate possible GLYATL2 appearance within the brain. In a more recent study *N*-arachidonoylglycine was identified as efficacious agonist against G-protein

coupled receptor 18 in human microglia [110]. The once determined upregulation of GLYATL2 mRNA in the brain (Figure 4.66.) could be explained by that.

For human testis a very weak mRNA signal was detected (Figure 4.65.). Interestingly Waluk et al. did not detect any mRNA signal within testis [165], whereas no testis signal came up in the RT-qPCR study (Figure 4.66.). Due to no present studies suggesting possible GLYATL2 expression within testis, the obvious, but very tiny (5.33 % signal intensity), testis signal might be unspecific. However, in the subsequent RT-qPCR studies (Suppl. Fig. 61-63) no major difference between brain signal and the other tested mRNAs became obvious. Hence, the determined upregulation in the semi-quantitative PCR study before could not be confirmed. The section emphasizes that RT-qPCR experimental set-up worked sufficiently.

Thereby, the detected GLYATL2 mRNA tissue expression results (Section 4.12.) contradict the study by Waluk and associates, who identified intensive GLYATL2 mRNA signals were identified in salivary gland (mixed), trachea, spinal cord, and skin fibroblasts [165]. The trachea and spinal cord mRNA were not investigated, but the salivary gland and fibroblast signals (Suppl. Fig. 63) did not show any mRNA-upregulation compared to the controls. However, there are differences between the present study and the publication from Waluk et al. [165] regarding the provider of the human tissue RNA samples and the method of cDNA synthesis. Discrepancies between the specimen donors cannot be excluded because Waluk and associates did not provide information on that.

Beyond these differences, they used another SYBR green mastermix and did not publish the RT-qPCR protocol, which decreases the chance of correct reproduction. Nonetheless, Waluk and co-workers also tested mRNA of human parotid gland with very low mRNA expression.

In summary, these differences in the methods between the study from Waluk and the present thesis may serve as explanation for contradictory data.

According to the literature, human GLYATL2 is mainly expressed in skin fibroblasts and salivary gland. In our studies, only very slight up-regulation of *GLYATL2* mRNA within testis and brain was detected. Hence, the literature data could not be confirmed thereby.

5.7. Error discussion of low expression levels and activities of GLYAT and GLYATL1 in HEK293 cells

5.7.1. gDNA and mRNA analyses of transfected HEK293 and HeLa cells

The human GLYAT and GLYATL1 enzymes were not sufficiently overexpressed in HEK293 cells (Sections 4.10., 4.11.). Compared with *E. coli* Origami 2(DE3), the HEK293 cells revealed remarkably lower expression signals and enzyme activities. Beyond the inefficiently expressed wild-types, several sequence variants showed only proportions of wild-type expression. This did not meet the expectations, because missense mutations usually not interfere with expression

pattern of target enzyme. However, when aspartic acid is changed against asparagine (GLYATL1 p.(Asp255Asn)), the expression might be influenced due to charge difference possibly leading to altered folding patterns (Section 5.1.). Indeed, other studies described how missense mutations may influence target protein expression or folding [24, 143], which is more precisely described in Section 5.1.

Nevertheless, gDNA and mRNA were checked for recombinant GLYAT and GLYATL1 in HEK293 cells (Suppl. Fig. 43-55). For the transfected GLYAT wild-type in HEK293 and HEK293T cells, corresponding gDNA signals were clearly shown (Suppl. Fig. 43-46), which demonstrated a sufficient transfection. The agarose gels underlined the expected recombinant GLYAT signal on gDNA level for the wild-type and all produced sequence variants except p.(Met65Thr). For this variant a bigger fragment of 500 bp was detected when amplified with GLYAT primers of exon 3/exon 4 transition. The same was obvious for mRNA analyses, which is demonstrated in the appendix (Suppl. Fig. 47-53.), however, only when using these primers for exon3/exon 4 transition. Conversely, primers for amplification of CMV_MCS/exon 1 transition demonstrated p.(Met65Thr) amplicons resembling that of wild-type (Suppl. Fig. 53). Hence, the bigger fragment seemed to be a primer specific artifact and the missense mutation seemed to be correctly introduced. Indeed, the mutagenesis was correctly performed but the insert appeared as exceptional artifact, which makes the activity results of this variant useless.

Beyond that, recombinant GLYATL1 gDNA signals appeared in expected sizes demonstrated by Suppl. Fig. 46.

To answer the question of transfection functionality, the recombinant GLYAT and GLYATL1 fulfilled the requirements on nucleic acid levels by showing the expected amplicons.

For the investigation of transcriptional level, the mRNAs of recombinant *GLYAT* and *GLYATL1* genes were elucidated (Suppl. Fig. 47-55). The Suppl. Fig. 47-50 demonstrate a sufficient mRNA production of recombinant GLYAT stably transfected only for HEK293 but not for HeLa cells. The transfection efficiency might not be a critical factor here due to high rate for HeLa cells compared to HEK293 cells [71]. HeLa cells (with higher transfection rate) demonstrated lower expression of GLYAT target protein, whereas it seemed to be higher for HEK293 cells, which are predicted to have the lower transfection rate. Lower expression was also observed for GLYAT-eGFP (Suppl. Fig. 43, 44 (mRNA: Suppl. Fig. 46) due to the size of the inserts and associated difficulty of plasmid transfection. Nonetheless, other transfection methods, such as electroporation, were suggested for larger plasmids [217]. The same results were observed for HEK293 and HeLa cells in the quantitative RT-qPCR studies. The mRNA ratios with similarities to wild-type ratio were furthermore obtained in the studies with sequence variants (Suppl. Fig. 52). Nonetheless, a

problematic fragment size of p.(Met65Thr), which repeatedly appeared, was circumvented by using different primers for recombinant *GLYAT* (Suppl. Fig. 53). Thereby amplification of exon 1 was declared as unproblematic, whereas amplification of region between exon 3 and 4 repeatedly leaded to unspecific amplicon of 500 bp. Thus, the bigger fragment seemed to be an unspecific artifact resulting from exon 3 and 4 transition.

Nonetheless, the presence of recombinant mRNA for GLYATL1 wild-type and analyzed sequence variants was clearly demonstrated by Suppl. Fig. 54 and 55.

To summarize, gDNA and mRNA of transfected GLYAT and GLYATL1 were confirmed in HEK293 cells and HeLa cells. The problematic low protein expression seems to be neither referred to a transfectional (gDNA) nor to a transcriptional (mRNA) issue. The low expression is most likely referred to the translational level discussed in Section 5.7.2.

5.7.2. Translational issues of transfected HEK293 cells?

The Western blot studies demonstrated signals for overexpressed GLYAT (Section 4.10.) and GLYATL1 (Section 4.11.). However, the expression rates were too low for convincing and reliable enzyme activity data. The data strongly depend on transfection series and amounts of used total protein. For instance, Section 4.10. showed very low activity data in one series (Figure 4.60., GLYAT wild-type v_{max}: 5.2 U/mg protein), while most other series exhibited higher activities (Figure 4.61., GLYAT wild-type v_{max}: 166 U/mg protein). Thus, the reproducibility of activity data seems to be invalid demonstrating high inter-assay variation.

The GLYATL1 activity data within HEK293 cells (Section 4.11.) exhibited comparable insights and stable kinetic data could not be derived. The assay was performed equally to GLYAT studies but with glutamine replacing glycine and phenylacetyl-coA instead of benzoyl-coA. Mainly due to the other used substrates for GLYATL1 activity test, lower conversion and activities were expected. Nonetheless, the glutamine was somehow difficult to dissolve and is quite more instable compared to glycine. At 37 °C glutamine is degraded over the time (7 % per day) [75], which decreases GLYATL1 activity due to lower amounts of available substrate present in the assay set-up. For fast providing of glutamine for activity tests, the solution was briefly heated at 50 °C to ensure fast dissolving and the stability might have suffered from that process.

Nonetheless, possible reasons might also be fast degradation due to unusability of the protein within in the cell or instability of the protein due to cleavage sensitive sequences. Indeed, protein sequences of GLYAT, GLYATL1 and GLYATL2 demonstrated cleavage sites for huge variety of proteases [208].

However, GLYAT and GLYATL1 might be impaired by folding or solubility problems, when overexpressed in HEK293 cells, which could explain the low expression rates (Section 4.10.,

4.11.). These challenges were described in Section 5.1. for corresponding sequence variants. A *CMV* promotor driven inhibition of expression was not confirmed, while proteasomal degradation was also excluded. The effects of the *CMV* promotor methylation inhibitor 5´-azacytidine and the proteasomal inhibitor MG-132 were discussed in Section 5.4. The reason for low expression levels in HEK293 cells might lie within unspecific cleavage and subsequent degradation of the proteins. Nevertheless, the problematic DTNB assay for homogenates due to parallel expressed coA-releasing enzymes [69, 72] came into account for the HEK293 measurements accordingly (Section 5.4.). The high background activities by constitutively expressed acyl thioesterases in HEK293 cells and the low activities, due to the low expression rates, complicated the interpretation of the kinetic properties.

One reason for the limited activities of the target enzymes might be the chosen vector system pcDNA3.1(+), although it is recommended to be used in HEK293 cells [149]. The *CMV* promotor of the pcDNA3.1(+) vector seemed not to enhance the GLYAT and GLYATL1 expression essentially, albeit it is widely used [78, 87, 138, 145]. However, other vectors containing a c-*myc* tag are suggested to increase target protein expression [80], which was also confirmed in our working group by overexpression of the human aspartoacylase 2 [2].

5.8. Overall assignment and motivation for the project

The present thesis treated the overexpression and characterization of amino acid *N*-acyltransferases. The main motivation behind the project was the incomplete research on this metabolic pathway as related pathways are more thoroughly researched. For example, the glucuronidation or sulfation have been studied in more details [77] than amino acid conjugations [5]. Correspondingly the amino acid conjugation was designated as the ´poor cousin´ of the drug metabolism family [7], because only a small number of drugs are conjugated with glycine. Today, the pathway becomes more relevant stepwise with special consideration on the regulation of the coA homeostasis in the liver [7]. At the start of the project, only one other research group had applied site-directed mutagenesis to investigate the influence of SNPs on recombinant GLYAT [152]. Notably, the reported investigations were limited to a bacterial overexpression, which does not ensure authentic PTMs. Therefore, the rather frequent GLYAT sequence variant p.(Gln61Leu), which observed 12% allele frequency in a South African cohort, but did not as yet undergo any functional assessment [152], was re-evaluated by the present study. The missense mutations of overexpressed human GLYAT may affect enzyme activity and consequently phase II detoxification. The high prevalence of p.(Gln61Leu) may suggest pharmacogenetic implications. Possible GLYAT deficiency and its analysis deserves consideration in CASTOR diseases or urea cycle defects that are to be treated with glycine [102, 112].

Beyond that, information on human GLYATL1 and GLYATL2 is scarce. This seems to be of special interest, because human GLYATL1 may be associated with the regulation of hepatocellular carcinoma [62] and GLYATL2 may have functions beyond detoxification in the liver [165, 166].

5.8.1. Holistical discussion of the project results

Regardless the occurring challenges in the beginning of the project (Section 4.1.), a few new insights and the clarification of contradictory literature data on the intracellular localization of the GLYATL1 enzyme were successful. The characterization of human GLYAT wild-type and the p.(Gln61Leu) and p.(Asn156Ser) sequence variants resulted in a first publication [137]. The p.(Gln61Leu) sequence variant was previously described as very frequent in a South African cohort due to a founder effect [153]. This variant possessed only a few percent of wild-type activity [137], which delivers new pharmacogenetic implications as isovaleric acidemia is very frequent in South Africa [39]. The intracellular localization of human GLYAT, which was supposed to be mitochondrial [107], was confirmed within this thesis. The widespread opinion that p.(Asn156Ser) represents the real wild-type of human GLYAT was repeatedly confirmed. Regardless of this, van der Sluis et al. analyzed the most frequent sequence variants of human GLYAT [152]. Due to the rarity of frequent GLYAT sequence variants, it might be a challenge to identify other diseases influenced by GLYAT polymorphisms. Likewise, it is difficult to find GLYAT deficient patient material. As the human enzyme was only overexpressed in prokaryotic systems thus far [107, 149] another part of the present study was the overexpression in HEK293 cells. However, the target protein yield was insufficient, which complicated the evaluation of enzyme activity data. Nevertheless, the protein overexpression was functional as target protein signals were detected. Nonetheless, it must be optimized in the future, e.g., by use of a more efficient vector system, for example a *myc*-tag vector. As the publications in the field of eukaryotic overexpression of human GLYAT are scarce, it would be of general interest to investigate enzymes equipped with authentic PTMs.

The human GLYATL1 was also overexpressed in the *E. coli* system. The present study re-evaluates the GLYATL1 overexpression as they are only limited in the literature [107, 178]. Due to the supportive function of GLYATL1 in specific urea cycle defects, which result in hyperammonaemia [144], selected GLYATL1 sequence variants were investigated accordingly. The overexpression of the wild-type and sequence variants was successful in the *E. coli* system. A publication was not derived from that, because the data variations were intolerable. The overexpression in HEK293 cells observed the same challenge as it was demonstrated for human GLYAT. Zhang et al. characterized the human GLYATL1 enzyme to be cytosolic [178], whereas Matsuo et al. observed a localization in the ER [107]. A previous study identified glutamine-

transferase activity in human mitochondria suggesting a mitochondrial localization of the enzyme [169]. Due to this contradiction, we re-examined the localization of human GLYATL1 in HEK293 cells, which resulted in a clear mitochondrial localization underlined by a Western blot study with human liver fractions.

Human GLYATL2 was investigated *in-silico* within this project and the downstream analyses of the enzyme were outsourced to a bachelor project [164]. Overexpression in *E. coli* Origami 2(DE3) and the functional assessment of the human enzyme wild-type and certain sequence variants was established therein. Indeed, human GLYATL2 undercut the other two enzymes regarding research presence. Therefore, and especially due to the assumed expanding functions, such as regulatory functions in inflammation or pain perception [166] the enzyme was elucidated sufficiently. The mRNA expression pattern that was described in the literature [165] could not be reproduced within the thesis, most likely due to granular differences in the applied methods. However, within the bachelor project the wild-type and five sequence variants were overexpressed in *E. coli.* The five sequence variants, i.e., p.(Lys19Gln), p.(Lys84Ile), p.(Gly109Val), p.(Lys184Glu) and p.(Leu189Phe), demonstrated decreased activity compared to the wild-type enzyme [164]. p.(Lys19Gln) was described in the literature as a decreased active enzyme [166], which was confirmed thereby. The other findings represent new research aspects as they were not published thus far.

Notwithstanding, as the results confirmed, GLYAT and GLYATL1 sequence variations impair the enzyme activity (Section 4.6.) and thus influence treatment options of certain inborn errors of metabolism (Section 5.5.). Urea cycle defects e.g., CPS or OTC deficiency, should be considered under the light of frequent GLYAT polymorphisms such as p.(Gln61Leu). The same applies for the treatment of the isovaleric acidemia, which might be rather treated with L-carnitine. The initial plan to investigate these influences on the enzyme activities of amino acid *N*-acyltransferases was fulfilled. However, further research using a more efficient overexpression system e.g., *c-myc* containing vectors, leading to more reliable activity data is required. As the human GLYAT and GLYATL1 were detected in the mitochondria, the third aim of the thesis was reached thereby. Nonetheless, the fourth aim of the GLYATL2 tissue expression analysis did not result in a specific tissue with upregulated GLYATL2 expression.

5.8.2. Outlook

Two publications have thus far resulted from the present thesis and the cornerstones for further projects have been set. Thus, mutagenesis of the described GLYAT sequence variants have been performed and they might be subcloned into a more efficient vector system to analyze the effects in human derived cell lines. Notwithstanding, the enzyme activities for GLYAT and GLYATL1 and

selected sequence variants were determined serving as a basis for follow-up and reproducing studies. The molecular biological pre-works (clonings) are completed in the bacterial system for all selected sequence variants of human GLYAT, GLYATL1 and GLYATL2. Other *E. coli* strains could possibly be suitable for further overexpressions of the enzymes. As the mutagenesis constructs also exist for the most probable GLYAT wild-type p.(Asn156Ser), enzyme activity studies with the latter would be recommendable. This might be supplemented by studies on GLYATL1 and GLYATL2 activities as they were not reproducible within the present study. Due to the partially confusing data in the DTNB activity assays and the described restrictions for eukaryotic homogenates, another activity measurement method might be suitable. A waiver on the coA measuring principle would be appropriate as this will interfere with unspecific releases in eukaryotic homogenates. The fluorescence-based assay, where the activity correlates with fluorescence signal of a specific probe, might be a senseful alternative. Conversely, the GFP fluorophoring of GLYAT substrates or products might be suitable. The intracellular localization studies on human GLYAT and GLYATL1 might be expanded by a subcellular fractionation experiment to further confirm their mitochondrial localization.

References

Websites are listed from reference no. 179 on

1. Abramovs N, Brass A, Tassabehji M. Hardy-Weinberg Equilibrium in the Large Scale Genomic Sequencing Era. Front Genet 2020;11:1-11, doi:10.3389/fgene.2020.00210

2. Aktar S. Bi-allelic co-expression of selected sequence variants in the *ASPA* gene for the study of compound heterozygosity. Master thesis, Bonn-Rhein-Sieg University of Applied Sciences, working group ´inborn errors of metabolism´, Rheinbach, supervisor Prof. Dr. Sass, 2020

3. Alhomida AS. Inhibition studies of the carnitine acetyltransferase from skeletal muscle of the camel (Camelus dromedarius) by sulfhydryl reagents and metal ions. Biochem Mol Bio Int 1996;39:923-31, doi: 10.1080/15216549600201072

4. Armenteros JJA, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, Nielsen H. Detecting sequence signals in targeting peptides using deep learning. Life Sci Alliance 2019;2:1-14, doi: 10.26508/lsa.201900429

5. Badenhorst CPS, Jooste M, van Dijk AA. Enzymatic characterization and elucidation of the catalytic mechanism of a recombinant bovine glycine *N* acyltransferase. Drug. Metab. Dispos. 2012;40:346-52, doi: 10.1124/dmd.111.041657

6. Badenhorst CPS, van der Sluis R, Erasmus E, van Dijk AA. Glycine conjugation: importance in metabolism, the role of glycine *N*-acyltransferase, and factors that influence interindividual variation. Drug Metab Toxicol 2013;9:1139-53, doi: 10.1517/17425255.2013.796929

7. Badenhorst CPS, Erasmus E, van der Sluis R, Nortje C, van Dijk AA. A new perspective on the importance of glycine conjugation in the metabolism of aromatic acids. Drug Metab Rev 2014;46:343-61, doi: 10.3109/03602532.2014.908903

8. Bartlett K, Gompertz D. The specificity of glycine *N*-acylase and acylglycine excretion in the organic acidemias. Biochem Med. 1974; 10;:15-23, doi: 10.1016/j.cca.2021.10.007

9. Baumgartner MR, Hörster F, Dionisi-Vici C, Haliloglu G, Karall D, Chapman KA, Huemer M, Hochuli M, Assoun M, Ballhausen D, Burlina A, Fowler B, Grünert SC, Grünewald S, Honzik T, Merinero B, Perez-Cerda C, Scholl-Bürgi S, Skovby F, Wijburg F, MacDonald A, Martinelli D, Sass JO, Valayannopoulus V, Chakrapani A. Proposed guidelines for the diagnosis and management of methylmalonic and propionic acidemia. Orphanet J Rare Dis 2014;9:1-36, doi: 10.1186/s13023-014-0130-8

10. Bayatvarkeshi M, Imteaz MA, Kisi O, Zarei M, Yaseen ZM. Application of M5 model tree optimized with Excel Solver Platform for water quality parameter estimation. Environ Sci Pollut Res Int 2021;28:7347-64, doi: 10.1007/s11356-020-11047-w

11. Bellelli A, Brunori M, Miele AE, Panetta G, Vallone B. The allosteric properties of hemoglobin: insights from natural and site-directed mutants. Curr Protein Pept Sci 2006;7:17-45, doi: 10.2174/138920306775474121

12. Bertani G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. J Bacteriol 1951;62:293–300, doi: 10.1128/jb.62.3.293-300.1951

13. Bessette PH, Aslund F, Beckwith J, Georgiou G. Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm. Proc Natl Acad Sci USA 1999;96:13703-8, doi: 10.1073/pnas.96.24.13703

14. Beyoglu D, Idle JR. The glycine deportation system and its pharmacological consequences. Pharmacol Therap 2012;135:151-67, doi: 10.1016/j.pharmthera.2012.05.003

15. Beyoglu D, Smith RL, Idle JR. Dog bites man or man bites dog? The enigma of the amino acid conjugations. Biochem Pharmac 2012;83:1331-39., doi: 10.1016/j.bcp.2011.12.031

16. Bonner WM, Bloch K. Purification and properties of fatty acyl thioesterase I from Escherichia coli. J Biol Chem 1972;25:3123-33, doi: 10.1016/S0021-9258(19)45222-8

17. Böttcher D, Brüsehaber E, Doderer K, Bornscheuer UT. Functional expression of the gamma-isoenzyme of pig liver carboxyl esterase in Escherichia coli. Appl Microbiol Biotechnol 2007;73:1282–89, doi: 10.1007/s00253-006-0585-1

18. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding Anal Biochem. 1976;72:248-5, doi: 10.1006/abio.1976.9999

19. Bradshaw HB,Rimmermann N, Hu SSJ, Burstein S, Walker M. Novel endogenous *N*-acyl glycines identification and characterization. Vitam Hom 2009;81:191-205, doi: 10.1016/S0083-6729(09)81008-X

20. Brunengraber H, Kasumov T, Brunengraber LL, Comte B, Puchowitz MA, Jobbins K, Thomas K, David F, Kinman R, Wehrli S, Dahms W, Kerr D, Nissim I. New secondary metabolites of phenylbutyrate in humans and rats. Drug Metab Dispos 2004;32;10-9, doi: 10.1124/dmd.32.1.10

21. Brusilow SW. Phenylacetylglutamine May Replace Urea as a Vehicle for Waste Nitrogen Excretion Pediatr Res 1991;29:147-50, doi: 10.1203/00006450-199102000-00009

22. Brüsehaber E, Schwiebs A, Schmidt M, Böttcher D, Bornscheuer UT. Production of pig liver esterase in batch fermentation of *E.coli* Origami. Appl Microbiol Biotechnol 2010;86:1337-44, doi: 10.1007/s00253-009-2392-y

23. Cain JA, Solis N, Cordwell SJ. Beyond gene expression: the impact of protein post-translational modifications in bacteria. J Proteomics 2014;97:265-86, doi: 10.1016/j.jprot.2013.08.012

24. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics 2006;22:2729-34, doi: 10.1093/bioinformatics/btl423

25. Campbell L, Wilson HK, Samuel AM, Gompertz D. Interactions of m-xylene and aspirin metabolism in man. Br J Ind Med 1988;45:127-32, doi: 10.1136/oem.45.2.127

26. Carabetta VJ Cristea IM. Regulation, Function and Detection of Protein Acetylation in Bacteria J Bacteriol 2017;199:1-15, doi: 10.1128/JB.00107-17

27. Chaturvedi, S, Driscoll, W.J, Elliot, BM, Faraday, MM, Grunberg, NE, Mueller, GP. In vivo evidence that *N*-oleoyl glycine acts independently of its conversion to oleamide. Prostaglandins Other Lipid Mediat. 2006;81,:136–49, doi: 10.1016/j.prostaglandins.2006.09.001

28. Chinen Y, Nakamura S, Tamashiro K, Sakamoto O, Tashiro K, Inokuchi T, Nakanishi K. Isovaleric acidemia: Therapeutic response to supplementation with glycine, L-carnitine, or both in combination and a 10 year follow-up case-study. Mol Genet metab Rep 2017;11:2-5, doi: 10.1016/j.ymgmr.2017.03.002

29. Christman JK. 5-azacytidine and 5-aza-2´-deoxycytidine as Inhibitors of DNA methylation: Mechanistic Studies and their Implication for Cancer Therapy. Oncogene 2002;21:5483-95, doi: 10.1038/sj.onc.1205699

30. Christopher, R, Sankaran, BP. An insight into the biochemistry of inborn errors of metabolism for a clinical neurologist. Ann Indian Acad Neurol, 2008;11:68-81, doi: 10.4103/0972-2327.41873

31. Cierniak P. Cloning of genes encoding thermostable DNA polymerases, subsequent protein expression, purification and functional characterization. Master thesis, Bonn-Rhein-Sieg University of Applied Sciences, supervisor Prof. Reinscheid, 2008

32. Clarkson JR, Cui ZF, Darton RC. Protein Denaturation in Foam. J Colloid Interface Sci 1999;215:323-32, doi: 10.1006/jcis.1999.6256

33. Corsini L, Hothorn M, Scheffzek K, Sattler M, Stier G. Thioredoxin as a fusion tag for carrier-driven crystallization. Protein Sci 2008;17:2070-9, doi: 10.1110/ps.037564.108

34. Cravatt, BJ, Prospero-Garcia, O, Siuzdak G, Gilula, NB, Henriksen, SJ, Boger, DL, Lerner RA. Chemical characterization of a family of brain lipids that induce sleep. Science 1995;268:1506–9, doi: 10.1126/science.7770779

35. Cusmai R, Martinelli D, Moaevero R, Dionisi Vici C, Vigevano F, Castana C, Elia M, Bernabei S, Bevivino E. Ketogenic diet in early myoclonic encephalopathy due to nonketotic hyperglycinemia. Eur J Paed Neurol 2012;16:509-13, doi: 10.1016/j.ejpn.2011.12.015

36. Dagert, M, Ehrlich, S. Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. Gene. 1979;6:23-8, doi: 10.1016/0378-1119(79)90082-9

37. Demain AL, Vaishnav P. Production of recombinant protein by microbes and higher organisms. Biotechnol Ad., 2009;27:297-306, doi: 10.1016/j.biotechadv.2009.01.008

38. Dempsey DR, Bond JD, Carpenter AM, Ospina SR, Merkler DJ. Expression, Purification and Characterization of Mouse Glycine *N*-Acyltransferase in *Escherichia coli*. Protein Exp Purif 2014;97:23-8, doi: 10.1016/j.pep.2014.02.007

39. Dercksen M, Duran M, Ijlst L, Miene LJ, Reinecke CJ, Ruiter JPN, Waterham HR, Wanders RJA. Clinical variability of isovaleric acidemia in a genetically homogeneous population. J Inherit Metab Dis 2012;35:1021-9, doi: 10.1007/s10545-012-9457-2

40. Dingova D, Leroy J, Check A, Garaj V, Krejci E, Hrabovska A. Optimal detection of cholinesterase activity in biological samples: modifications to the standard Ellman's assay. Anal Biochem 2014;462:67-75, doi: 10.1016/j.ab.2014.05.031

41. Dorne JLCM, Walton K, Renwick AG. Human variability for metabolic pathways with limited data (CYP2A6, CYP2C9, CYP2E1, ADH esterases, glycine and sulphate conjugation). Food Chem Toxicol 2004;42;397-421, doi: 10.1016/j.fct.2003.10.003

42. Duffy LF, Kerzner B, Seeff L, Barr SB, Soldin SJ. Preliminary assessment of glycine conjugation of para-aminobenzoic acid as a quantitative test of liver function. Clin Biochem 1995;28:527-30, doi: 10.1016/0009-9120(95)00036-9

43. Durfee T, Nelson R, Baldwin S, Plunkett G, Burland V, Mau B, Petrosino JF, Qin X, Muzny DM, Ayele M, Gibbs RA, Csörgo B, Pósfai B, Weinstock GM, Blattner FR. The complete Genome Sequence of *Escherichia coli* DH10B: Insights into the Biology of a Laboratory Workhorse. J Bacteriol 2008;190:2597-606, doi: 10.1128/JB.01695-07

44. Dyda F, Klein DC, Hickman AB. GCN5-Related *N*-Acetyltransferases: A Structural Overview Annu Rev Biophys Biomol Struct 2000;29:81-103, doi: 10.1146/annurev.biophys.29.1.81

45. Eadie GS. The Inhibition of Cholinesterase by Physostigmine and Prostigmine. J Biol Chem 1942;146:85-93, doi: 10.1016/s0021-9258(18)72452-6

46. Ellman GL. Tissue Sulfhydryl Groups. Arch Biochem Biophys 1959;82:70-7, doi: 10.1016/0003-9861(59)90090-6

47. Esposti DD, Hamelin J, Bosselut N, Saffroy R, Sebagh M, Pommier A, Martel C, Lemoine. A Mitochondrial Roles and Cytoprotection in Chronic Liver Injury. Biochem Res Int 2012;387626:1-16, doi: 10.1155/2012/387626

48. Eyer P, Worek F, Kiderlen D, Sinko G, Stuglin A, Simeon-Rudolf V, Reiner E. Molar absorption coefficients for the reduced Ellman´s reagent: reassessment. Anal Biochem 2003;312:224-7, doi: 10.1016/s0003-2697(02)00506-7

49. Feillet F, Leonard JV. Alternative pathway therapy for urea cycle disorders. J Inher Metab Dis 1998; 21:101-11, doi: 10.1023/a:1005365825875

50. Ferdinand W. The interpretation of non-hyperbolic rate curves for 2-substrate enzymes. A possible mechanism for phosphofructokinase. Biochem J. 1966;98:278-83, doi: 10.1042/bj0980278

51. Forny P, Hörster F, Ballhausen D, Chakrapani A, Chapman KA, Dionisi-Vici C, Dixon M, Grünert SC, Grunewald S, Haliloglu G, Hochuli M, Honzik T, Karall D, Martinelli D, Molema M, Sass JO, Scholl-Bürgi S, Tal G, Williams M, Huemer M, Baumgartner MR. Guidelines for the diagnosis and the management of

methylmalonic acidemia and propionic acidemia: first revision. J Inherit Metab Dis 2021;44:566-92, doi: 10.1002/jimd.12370

52. Fortun J, Martin-Davila P, Alvarez ME, Norman F, Sanchez-Sousa A, Gajate L, Barcena R, Nuno SJ, Moreno S. False-positive results of Aspergillus galactomannan antigenemia in liver transplant recipients. Transplantation 2009;87:256-60, doi: 10.1097/TP.0b013e31819288d5

53. Fries MH, Rinaldo P, Schmidt-Sommerfeld E, Jurecki E, Packman S. Isovaleric acidemia: response to a leucine road after 3 weeks of supplementation with glycine, L-carnitine, and combined glycine-carnitine therapy. J Pediatr 1996;129:449-52, doi: 10.1016/s0022-3476(96)70081-1

54. Frizler M, Yampolsky IV, Baranov MS, Stirnberg M, Gütschow M. Chemical introduction of the green fluorescence: imaging of cysteine cathepsins by an irreversibly locked GFP fluorophore. Org Biomol Chem. 2013;11:5913-21, doi: 10.1039/c3ob41341a

55. Fromenty B, Pessayre D. Inhibition of mitochondrial beta-oxidation as a mechanism of hepatotoxicity. Pharmacol Ther 1995;67:101-54, doi: 10.1016/0163-7258(95)00012-6

56. Gardner JG, Escalante-Semerena JC. Biochemical and Mutational Analyses of AcuA, the Acetyltransferase Enzyme That Controls the Activity of the Acetyl Coenzyme A Synthetase (AcsA) in *Bacillus subtilis.* J Bacteriol 2008;190:5132-36, doi: 10.1128/JB.00340-08

57. Gerard GF, Potter RJ, Smith MD, Rosenthal K, Dhariwal G, Lee J, Chatterjee DK. The role of template-primer in protection of reverse transcriptase from thermal inactivation. Nucleic Acid Res 2002;30:3118-29, doi: 10.1093/nar/gkf417

58. Ghosh R, Gilda JE, Gomes AV. The necessity of and strategies for improving confidence in the accuracy of Western blots. Expert Rev Proteomics 2014;11:549-60, doi: 10.1586/14789450.2014.939635

59. Graham, FL, Smiley, J, Russell, WC, & Nairn R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J Gen Virol 1977;36:59–74, doi: 10.1099/0022-1317-36-1-59

60. Gregersen N, Kolvraa S, Mortensen P. Acyl-CoA: Glycine *N*-Acyltransferase: *In vitro* studies on the glycine conjugation of straight- and branched chained acyl-CoA esters in human liver. Biochem Med Metabo Biol 1986;35:210-8, doi: 10.1016/0885-4505(86)90076-9

61. Grünert SC, Sass JO. 2-metyhlacetoacetyl-coenzyme A thiolase (beta-ketothiolase) deficiency: one disease – 2 pathways. Orphanet J Rare Dis 2020;15:1-7, doi: 10.1186/s13023-020-01357-0

62. Guan R, Hong W, Huang J, Peng T, Zhao Z, Lin Y, Yu M, Jian Z. The expression and prognostic value of GLYATL1 and its potential role in hepatocellular carcinoma. J Gastrointest Oncol 2020;11:1305-21, doi: 10.21037/jgo-20-186

63. Häberle J, Burlina A, Chakrapani A, Dixon M, Karall D, Lindner M, Mandel H, Martinelli D, Pintos-Morell D, Santer R, Skouma A, Servais A, Tal G, Rubio V, Huemer M, Dionisi-Vici C. Suggested guidelines for the diagnosis and management of urea cycle disorders. First revision J Inherit Metab Dis 2019;42:1192-1230, doi: 10.1002/jimd.12100

64. Hamosh A, Maher JF, Bellus GA, Rasmussen SA, Johnston MV. Long-term Use of High-Dose Benzoate and Dextrometorphan fort he treatment of nonketotic hyperglycinemia. J Pediatr. 1998;132:709-13, doi: 10.1016/s0022-3476(98)70365-8

65. Hanes CS. Studies on plant amylases. Biochem J 1932;26:1406-21, doi: 10.1042/bj0261406

66. Hardy GH. Mendelian proportions in a mixed population. Science 1908;28:49-50, doi: 10.1126/science.28.706.49

67. Harrison GP, Mayo MS, Hunter E, Lever AM. Pausing of reverse transcriptase on retroviral RNA templates is influenced by secondary structures both 5´and 3´of the catalytic site. Nucleic Acid Res 1998;26:3433-42, doi: 10.1093/nar/26.14.3433

68. Herkenham M, Lynn AB, Johnson MR, Melvin LS, deCosta BR, Rice KC. Characterization and localization of cannabinoid receptors in rat brain: a quantitative in vitro autoradiographic study. J Neurosci 1991;11:563–83, doi: 10.1523/JNEUROSCI.11-02-00563.1991

69. Hirano T, Kishi M, Sugimoto H, Taguchi R, Obinata H, Ohshima N, Tatei K, Izumi T. Thioesterase activity and subcellular localization of acylprotein thioesterase 1/lysophospholipase 1. Biochim Biophys Acta, 2009,1791,:797-805, doi: 10.1016/j.bbalip.2009.05.001

70. Hofstee BHJ. On the Evaluation of the Constants Vm and KM in enzyme reactions. Science 1952;26:329-31, doi: 10.1126/science.116.3013.329

71. Horibe T, Torisawa A, Akiyoshi R, Hatta-Ohashi Y, Suzuki H, Kawakami K. Transfection efficiency of normal and cancer cell lines and monitoring of promotor activity by single-cell bioluminescence imaging. Luminescence 2014;29:96-100, doi: 10.1002/bio.2508

72. Hornemann T. Palmitoylation and depalmitoylation defects. J Inherit Metab Dis 2015;38:179-86, doi: 10.1007/s10545-014-9753-0

73. Huang SM, Bisogno T, Petros TJ, Chang SY, Zavitsanos PA, Zipkin RE, Sivakumar R, Coop A, Maeda DY, De Petrocellis L, Burstein S, Di Marzo V, Walker JM. Identification of a new class of molecules, the arachidonoyl amino acids, and characterization of one member that inhibits pain. J Biol Chem 2001;276,:42639–44, doi: 10.1074/jbc.M107351200

74. Hutt AJ, Caldwell J, Smith RL. The metabolism of aspirin in man: a population study. Xenobiotica 1986;16;239-49, doi: 10.3109/00498258609043527

75. Jagusic M, Forcic D, Brgles M, Kutle L, Santak M, Jergovic M, Kotarski L, Bendelja K, Halassy B. Stability of Minimum Essential Medium functionality despite L-gluamine decomposition. Cytotechnology 2016;68:1171-83, doi: 10.1007/s10616-015-9875-8

76. Jain M, Nilsson R, Sharma S, Madhusudhan N, Kitami T, Souza AL, Kafri R, Kirschner MW, Clish CB, Mootha VK. Metabolite Profiling Identifies a Key Role for Glycine in Rapid Cancer Cell Proliferation. Science 2012;336:1040-4, doi: 10.1126/science.1218595

77. Jancova P, Anzenbacher P, Anzenbacherova E. Phase II drug metabolizing enzymes Biomed Pap Med 2000;154:103-16, doi: 10.5507/bp.2010.017

78. Janknecht R, Nordheim A. Affinity purification of histidine-tagged proteins transiently produced in HeLa cells. Gene 1992;121:321-4, doi: 10.1016/0378-1119(92)90137-e

79. Jeffries KR, Dempsey DR, Farrell EK, Anderson RL, Garbade GJ, Gurina TS, Grunhonjic I, Gunderson CA, Merkler DJ. Glycine *N*-acyltransferase-like 3 is responsible for long-chain *N*-acylglycine formation in N18TG2 cells. J Lipid Res 2016;57:781-90, doi: 10.1194/jlr.M062042

80. Ju X, Ren M, Chen K, Wang Q. Overexpression of c-Myc enhances recombinant protein production in High 5 cells after baculovirus infection. Z Naturforsch C J Biosci, 2018;73:147-51, doi: 10.1515/znc-2017-0076

81. Karki R, Pandya D, Elson RC, Ferlini C. Defining „mutation“ and „polymorphism“ in the era of personal genomics. BMC Med Genomics 2015;8:1-7, doi: 10.1186/s12920-015-0115-z

82. Karimi M, Ignasiak MT, Chan B, Croft AK, Radom L, Schiesser CH, Pattison DI, Davies MJ. Reactivity of disulfide bonds is markedly affected by structure and enviroment: implications for protein modification and stability. Nature 2016;6:1-12, doi: 10.1038/srep38572

83. Kebamo S, Tesema S, Geleta B. The Role of Biotransformation in Drug Discovery and Development J Drug Metab Toxicol 2015;6:1-13, doi: 10.4172/2157-7609.1000196

84. Kelley M, Vessey DA. Isolation and characterization of mitochondrial acyl-coA: glycine *N*-acyltransferases from kidney. J Biochem Toxicol 1993; 8:63-9, doi: 10.1002/jbt.2570080203

85. Kelley M, Vessey DA. Characterization of the acyl-coA: amino acid *N*-acyltransferases from primate liver mitochondria. J Biochem Toxicol. 1994;9:153-8, doi: 10.1002/jbt.2570090307

86. Kendall DA, Yudowski GA. Cannabinoid Receptors in the Central Nervous System: Their Signaling and Roles in Disease. Front Cell Neurosci 2007;10:1-10, doi: 10.3389/fncel.2016.00294

87. Khan KH. Gene Expression in Mammalian Cells and ist Applications. Adv Pharm Bull 2013;3:257-63, doi: 10.5681/apb.2013.042

88. Kim SC, Sprung R, Chen Y, Xu Y, Ball H, Pei J, Cheng T, Kho Y, Xiao H, Xiao L, Grishin NV, White M, Yang XJ, Zhao Y. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. Mol Cell 2006;23:607–18, doi: 10.1016/j.molcel.2006.06.026

89. Kocamemi BA, Cecen F. Kinetic analysis of the inhibitory effect of trichloroethylene (TCE) on nitrification in cometabolic degradation. Biodegradation 2007;18:71-81, doi: 10.1007/s10532-005-9037-3

90. Knights KM, Sykes MJ, Miners JO. Amino acid conjugation: contribution to the metabolism and toxicity of xenobiotic carboxylic acids. Expert Opin Drug Metab Toxicol 2007,3:159-68, doi: 10.1517/17425255.3.2.159

91. Kølvraa S, Gregersen N. Acyl-coA:glycine *N*-acyltransferase: organelle localization and affinity toward straight- and branched-chained acyl-coA esters in rat liver. Biochem Med Metab Biol 1986; 36:98-105, doi: 10.1016/0885-4505(86)90112-x

92. Krieger I, Tanaka K. Therapeutic effects of glycine in isovaleric acidemia. Pediatr Res 1976;10:25-9, doi: 10.1203/00006450-197601000-00005

93. Labeda DP. Transfer of the type strain of Streptomyces erythraeus (Waksman 1923) Waksman and Henrici 1948 to the Genus Saccharopolyspora Lacey and Goodfellow 1975 as Saccharopolyspora erythraea sp. nov., and designation of a neotype strain for Streptomyces erythraeus. Int J Syst Bacteriol 1987;37:19 –22, doi: :10.1099/00207713-37-1-19

94. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970;227:680-5, doi: 10.1038/227680a0

95. Lees HJ, Swann JR, Wilson ID, Nicholson JK, Holmes E. Hippurate: The Natural History of a Mammalian–Microbial Cometabolite. J Proteome Res 2013;12:1527-46, doi: 10.1021/pr300900b

96. Lehnert W. *N*-isovalerylalanine and *N*-isovalerylsarcosine: 2 new minor metabolites in isovaleric acidemia. Clin Chim Acta 1983;134:207-12, doi: 10.1016/0009-8981(83)90198-5

97. Lehnert W. Long-term results of selective screening for inborn errors of metabolism. Eur J Pediatr 1994;153;9-13, doi: 10.1007/BF02138770

98. Levy G. Pharmakokinetics of salicylate elimination in man. J Pharm Sci 1965;54:959-67, doi: 10.1002/jps.2600540703

99. Li A, Xue Y, Jin C, Wang M, Yao X. Prediction of *N*-acetylation on internal lysines implemented in Bayesian Discriminant Method. Biochem Biophys Res Commun 2006;350:818-24, doi: 10.1016/j.bbrc.2006.08.199

100. Lineweaver H, Burk D. The Determination of Enzyme Dissociation Constants. J Am Chem Soc 1934;56:658-66, doi: 10.1021/ja01318a036

101. Lino-Cardenas CL, Bourgine J, Cauffiez C, Allorge D, Lo-Guidice JM, Broly F, Chevalier D. Genetic polymorphisms of Glycine *N*-acyltransferase (GLYAT) in a French Caucasian population. Xenobiotica 2010;40:853-61, doi: 10.3109/00498254.2010.519407

102. Loots DT, Mienie MJ, Erasmus E. Amino-acid depletion induced by abnormal amino-acid conjugation and protein restriction in isovaleric acidemia. Eur J Clin Nutr 2007;61:1323-27, doi: 10.1038/sj.ejcn.1602648

103. Lowry OH, Rosebrough NR, Farr AL, Randall RJ. Protein Measurement with the Folin Phenol Reagent. J Biol Chem 1977;193:265-75, doi: 10.1016/s0021-9258(19)52451-6

104. MacDermot KD, Nelson W, Soutter V, Towne W, Schulman JD. Glycine and benzoate conjugation and glycine acyltransferase activity in the developing and adult rat: possible relationships to nonketotic hyperglycinemia. Dev Pharmacol Ther 1981;3:150-9, doi: 10.1159/000457436

105. Mackie K. Cannabinoid receptors:where they are and what they do. J. Neuroendocrinol. 2008; 20:10-4, doi: 10.1111/j.1365-2826.2008.01671.x

106. Mao Y, Yan R, Li A, Zhang Y, Li J, Du H, Chen B, Wei W, Zhang Y, Sumners C, Zheng H, Li H. Lentiviral vectors mediate long-term and high efficiency transgene expression in HEK293T cells. Int J Med Sci 2015;12:407-15, doi: 10.7150/ijms.11270

107. Matsuo M, Terai K, Kameda N, Matsumoto A, Kurokawa Y, Funase Y, Nishikawa K, Sugaya N, Hiruta N, Kishimoto T. Designation of enzyme activity of glycine *N*-acyltransferase family genes and depression of glycine *N*-acyltransferase in human hepatocellular carcinoma. Biochem Biophys Res Commun. 2012; 420: 901-6, doi: 10.1016/j.bbrc.2012.03.099

108. Mawal YR, Qureshi IA. Purification to homogeneity of mitochondrial acyl-coA: glycine *N*-acyltransferase from human liver. Biochem Biophys Res Commun. 1994; 205:1373-9, doi: 10.1006/bbrc.1994.2817

109. Mawal Y, Paradis K, Qureshi IA. Developmental profile of mitochondrial glycine *N*-acyltransferase in human liver. J Pediatr 1997;130:1003-7, doi: 10.1016/s0022-3476(97)70293-2

110. McHugh D, Hu SSJ, Rimmerman N, Juknat ., Vogel V, Walker JM, Bradshaw HB. *N*-arachidonoyl glycine, an abundant endogenous lipid, potently drives directed cellular migration through GPR18, the putative abnormal cannabidiol receptor. BMC Neurosci 2010;11:1-13, doi: 10.1186/1471-2202-11-44

111. Merkler, DJ, Chew GH, Gee AJ, Merkler KA, Sorondo JPO, Johnson ME. Oleic acid derived metabolites in mouse neuroblastoma N18TG2 cells. Biochemistry 2004;43:12667–74, doi: 10.1021/bi049529p

112. Mitchell GA, Gauthier N, Lesimple A, Wang SP, Mamer O, Qureshi I. Hereditary and Acquired Diseases of Acyl-Coenzyme A Metabolism Mol Genet Metab 2008;94:4-15, doi: 10.1016/j.ymgme.2007.12.005

113. Morisaki N, Saito Y, Kumagai A. Synthesis and metabolism of arachidonoyl- and eicosapentaenoyl-coA in rat aorta. Biochem Biophys Acta 1983;752:301-6, doi: 10.1016/0005-2760(83)90127-3

114. Mroczkiewicz M, Winkler K, Nowis D, Placha G, Golab J, Ostaszewski R. Studies of the Synthesis of All Stereoisomers of MG-132 Proteasome Inhibitors in the Tumor Targeting Approach. J Med Chem, 2010;53:1509-18, doi: 10.1021/jm901619n

115. Mühlmann M, Forsten E, Noack S, Büchs J. Optimizing recombinant protein expression via automated induction profiling in microtiter plates at different temperatures. Microb Cell Fact 2017;16:220, doi: 10.1186/s12934-017-0832-4

116. Nandi DL, Lucas SV, Webster LT. Benzoyl-Coenzyme A: Glycine *N*-Acyltransferase and Phenylacetyl-Coenzyme A: Glycine *N*-Acyltransferase from Bovine Liver Mitochondria. J Biol Chem 1979;254:7230-37, doi: doi.org/10.1016/S0021-9258(18)50309-4

117. Narayanan MP, Kannan V, Vinayan KP, Vasudevan DM. Diagnosis of Major Organic Acidurias in Children: 2 Years Experience at a Tertiary Care Centre. Ind J Clin Biochem, 2011;26:347-53, doi: 10.1007/s12291-011-0111-9

118. Nebert D, McKinnon R, Puga A. Human drug-metabolizing enzyme polymorphisms: effects on risk of toxicity and cancer. DNA Cell Biol 1996;15:273-80, doi: 10.1089/dna.1996.15.273

119. Nowicka U, Chroscicki P, Stroobants K, Sladowska M, Turek M, Uszczynska-Ratajczak B, Kundra R, Goral T, Perni M, Dobson CM, Vendruscolo M, Chacinska A. Cytosolic aggregation of mitochondrial proteins disrupt cellular homeostasis by stimulating the aggregation of other proteins. Elife 2021;10;1-27, doi: 10.7554/eLife.65484

120. Palmer E, Freeman T. Investigation into the use of C- and N-terminal GFP fusion proteins for subcellular localization studies using reverse transfection microarrays. Comp Func Genom 2004;5:342-53, doi: 10.1002/cfg.405

121. Piotrowski G. Eine neue Reaction auf Eiweisskörper und ihre näheren Abkömmlinge Proceedings of the Imperial Academy of Philosophies in Vienna, mathematical-natural sciences section 1857;24:335-7.

122. Piper JD, Piper PW. Benzoate and Sorbate Salts: A systematic review of the potential hazards of these invaluable preservatives and the expanding spectrum of clinical uses for sodium benzoate. Comprehensive Reviews 2017;16:868-80, doi: 10.1111/1541-4337.12284

123. Pontremoli S, Melloni E, Michetti M, Sparatore B, Salamino F, Siliprandi N, Horecker BL. Isovalerylcarnitine is a specific activator of calpain of human neutrophils. Biochem Biophys Res Commun 1987;148:1189-95, doi: 10.1016/s0006-291x(87)80258-9

124. Prinz WA., Aslund F, Holmgren A, Beckwith J. The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the Escherichia coli cytoplasm. J Biol Chem 1997;272:15661-7, doi: 10.1074/jbc.272.25.15661

125. Pronobis MI, Deuitch N, Pfeifer M. The Miraprep: A Protocol that Uses a Miniprep Kit and Provides Maxiprep Yields. PloS One. 2016;11:1-12, doi: 10.1371/journal.pone.0160509

126. Radonić A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A. Guideline to reference gene selection for quantitative Real-time PCR. Biochem Biophys Res Commun 2003;313:856-62, doi: 10.1016/j.bbrc.2003.11.177

127. Riddles PW, Blakeley RL, Zerner B. Ellman´s reagent: 5´5´-dithiobis-2-nitrobenzoic acid: a reexamination. Anal Biochem 1979;94:75-81, doi: 10.1016/0003-2697(79)90792-9

128. Roe CR, Millington DS, Maltby DA, Kahler SG, Bohan TP. L-carnitine therapy in isovaleric acidemia. J Clin Invest 1984;74:2290-5, doi: 10.1172/JCI111657

129. Rohwer JM, Schutte C, van der Sluis R. Functional Characterisation of 3 Glycine *N*-Acyltransferase variants and the effect on glycine conjugation to benzoyl-coA. Int J Mol Sci 2021;22:1-17, doi: 10.3390/ijms22063129

130. Sakuma T. Alteration of urinary carnitine profile induced by benzoate administration. Arch Dis Child 1991;66:873-5, doi: 10.1136/adc.66.7.873

131. Sanger F, Nicklen S, Coulson AR. DNA-sequencing with chain-terminating inhibitors. Proc Natl Acad Sci 1977;74:5463-67, doi: 10.1073/pnas.74.12.5463

132. Sass JO, Fukao T, Mitchell GA. Inborn errors of ketone body metabolism and transport: An Update for the Clinic and for clinical laboratories. J Inb Err Metab 2018;6:1-7, doi: 10.1177/2326409818771101

133. Schachter D, Taggart JV. Benzoyl Coenzyme A and hippurate synthesis. J Biol Chem 1953;203:925-34, doi: 10.1016/S0021-9258(19)52362-6

134. Schachter D, Taggart JV. Glycine *N*-Acylase: Purification and Properties. J Biol Chem 1954;208;263-75, doi: 10.1016/S0021-9258(18)65643-1

135. Scherer WF, Syverton JT, Gey GO. Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. J Exp Med 1953;97:695–710, doi: 10.1084/jem.97.5.695

136. Schildgen V, Schildgen O. How is a molecular polymorphism defined? Cancer 2013;119:1, doi: 10.1002/cncr.27966

137. Schulke D, Sass JO. Frequent sequence variants of human glycine *N*-acyltransferase (*GLYAT*) and inborn errors of metabolism. Biochimie 2021;183:30-4, doi: 10.1016/j.biochi.2021.02.002

138. Schulke D, Hollenbeck JC, Klaas L, Sass JO. Funktionale Gen-Analytik. Angeborenen Stoffwechselstörungen auf der Spur. GIT-Laborfachzeitschrift 2022;4:31-3, doi: 10.1002/was.000600301

139. Schulke D, Sass JO. Frequent sequence variants of human glycine *N*-acyltransferase and inborn errors of metabolism. Poster Abstract P19 of the 34th Annual Conference of the Arbeitsgemeinschaft für Pädiatrische Stoffwechselstörungen (APS). Monatsschr Kinderheilkd 2022;170(Suppl1) S.14, doi: 10.1007/s0012-021-01291-w

140. Schulke D, Sass JO. Human glycine *N*-acyltransferase (GLYAT) and inborn errors of metabolism. Digital Poster at the 14th International Congress of Inborn Errors of Metabolism (ICIEM), Sydney, Australia (online); https://www.iciem2021.com.au/cms/wp-content/uploads/2021/11/ICIEM-2021-Congress-Poster-List.pdf

141. Schulke D, Sass JO. Intracellular Localization of Glycine *N*-Acyltransferase-like protein 1 (GLYATL1), a phase 2 enzyme relevant in urea cycle disorders. Poster Abstract of the SSIEM Annual Symposium, J Inher Metab Dis 2022, eingereicht.

142. Shi B, Xue M, Wang Y, Wang Y, Li D, Zhao X, Li X. An improved method for increasing the efficiency of gene transfection and transduction. Int J Physiol Pathophysiol Pharmacol 2018;10:95-104.

143. Stefl S, Nishi H, Petukh M, Panchenko AR, Alexov E. Molecular mechanism of disease-causing missense mutations. J Mol Biol 2013;425:3919-36, doi: 10.1016/j.jmb.2013.07.014

144. Summar ML, Ah Mew N. Inborn errors of metabolism with hyperammonaemia: urea cycle defects and related disorders. Pediatr Clin North Am 2018;65:231-46, doi: 10.1016/j.pcl.2017.11.004

145. Symens N, Rejmann J, Lucas B, Demeester J, Smedt SCD, Remaut K. Noncoding DNA in lipofection of HeLa cells-a few insights. Mol Pharm 2013;10:1070-9, doi: 10.1021/mp300569j

146. Tanaka K, Isselbacher KJ. The isolation and identification of *N*-isovalerylglycine from urine of patients with isovaleric acidemia. J Biol Chem 1967;242:2966-72, doi: 10.1016/S0021-9258(18)99599-2

147. Temellini A, Mogavero S, Giulianotti PC, Pietrabissa A, Mosca F, Pacifici GM. Conjugation of benzoic acid with glycine in human liver and kidney: a study on the interindividual variability. Xenobiotica 1993;23:1427-33.

148. Terral G, Marsicano G, Grandes P, Soria-Gomez E. Cannabinoid Control of Olfactory Processes: The *Where* Matters. Genes (Basel) 2020;11:1-8, doi: 10.3109/00498259309059451

149. Thomas P, Smart TG. HEK293 cell line: A vehicle for the expression of recombinant proteins. J Pharmacol Toxicol Methods 2005;51:187-200, doi: 10.1016/j.vascn.2004.08.014

150. Tian X, Wu L, Jiang M, Zhang Z, Wu R, Miao J, Liu C, Gao S. Downregulation of GLYAT Facilitates Tumor Growth and Metastasis and Poor Clinical Outcomes Through the PI3K/AKT/Snail Pathway in Human Breast Cancer. Front Ontol 2021;11:1-12, doi: 10.3389/fonc.2021.641399

151. Ure A. On Gouty Concretions, With a New Method of Treatment Med Chir Trans 1841;24:30-5, doi: 10.1177/095952874102400105

152. Van der Sluis R, Badenhorst CPS, van der Westhuizen F, van Dijk A. Characterisation of the influence of genetic variations on the enzyme activity of a recombinant human glycine *N*-acyltransferase. Gene 2013;515;447-53, doi: 10.1016/j.gene.2012.12.003

153. van der Sluis R, Badenhorst CPS, Erasmus E, van Dyk E, van der Westhuizen FH, van Dijk A.. A Conservation of the coding regions of the glycine *N*-acyltransferase gene further suggests that glycine conjugation is an essential detoxification pathway. Gene 2015; 571:126-34, doi: 10.1016/j.gene.2015.06.081

154. van der Sluis R, Ungerer V, Nortje C, van Dijk AA, Erasmus E. New insights into the catalytic mechanism of human glycine *N*-acyltransferase. J Biochem Mol Toxicol 2017;10:1-10, doi: 10.1002/jbt.21963

155. van der Sluis R. Analyses of the genetic diversity and protein expression variation of the acyl: CoA medium-chain ligases, ACSM2A and ACSM2B Mol Genet Genomics 2018;293:1279-92, doi: 10.1007/s00438-018-1460-3

156. van der Westhuizen FH, Pretorius PJ, Erasmus E. The utilization of alanine, glutamic acid, and serine as amino acid substrates for glycine *N*-acyltransferase. J Biochem Mol Toxicol. 2000; 14:102-9, doi: 10.1002/(sici)1099-0461(2000)14:2<102::aid-jbt6>3.0.co;2-h

157. van Hove JLK, Kerckhove KV, Hennermann JB, Mahieu V, Declercq P, Mertens S, De Becker M, Kishnani PS, Jaeken J. Benzoate Treatment and the Glycine Index in Nonketotic Hyperglycinemia. J Inherit Metab Dis 2005;28:651-63, doi: 10.1007/s10545-005-0033-x

158. van Hove JLK, Couglhin C, Swanson M, Hennermann JB, Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Gripp KW, Mirzaa GM, Amemiya A. Nonketotic hyperglycinemia. GeneReviews 2019;1-18, Bookshelf ID: NBK1357

159. Vessey DA, Kelley M, Warren RS. Characterization of the CoA ligases of human liver mitochondria catalyzing the activation of short- and medium-chain fatty acids and xenobiotic carboxylic acids. Biochim Biophys Acta 1999;1428:455-62, doi: 10.1016/s0304-4165(99)00088-4

160. Vest MF, Salzberg R. Conjugation reactions in the newborn infant: the metabolism of para-aminobenzoic acid. Arch Dis Childh 1965;40:97-105, doi: 10.1136/adc.40.209.97

161. Vetting MW, de Carvalho LPS, Yu M, Hegde SS, Magnet S, Roderick SL, Blanchard JS. Structure and functions of the GNAT superfamily of acetyltransferases Arch Biochem Biophys 2005;433:212-26, doi: 10.1016/j.abb.2004.09.003

162. Vockley J, Varimoo B, Tanaka K. Molecular characterization of 4 different classes of mutations in the isovaleryl-coA dehydrogenase gene responsible for isovaleric acidemia. Am J Hum Genet 1991;49;147-57.

163. Vockley J, Ensenauer R. Isovaleric acidemia: new aspects of genetic and phenotypic heterogeneity. Am J Med Genet C Semin Med Genet 2006;142:95-103, doi: 10.1002/ajmg.c.30089

164. von Stülpnagel I. Partial Characterisation of recombinant human glycine *N*-acyltransferase like protein 2 and the influence of sequence variation on enzyme activity. Bachelor thesis, Bonn-Rhein-Sieg University of Applied Sciences, working group ´inborn errors of metabolism´,Rheinbach, supervisor Prof. Dr. Sass, 2020

165. Waluk DP, Schultz N, Hunt MC. Identification of glycine *N*-acyltransferase-like protein 2 (GLYATL2) as a transferase that produces *N*-acyl glycines in humans. FASEB 2010;24:2795-803, doi: 10.1096/fj.09-148551

166. Waluk DP, Sucharski F, Sipos L, Silberring J, Hunt MC. Reversible Lysine Acetylation Regulates Activity of Human Glycine *N*-Acyltransferase-like 2 (hGLYATL2). J Biol Chem 2012;287:16158-67, doi: 10.1074/jbc.M112.347260

167. Walsh S, Diamond D. Non-linear curve fitting using Microsoft Excel Solver Talanta 1995;42:561-72, doi: 10.1016/0039-9140(95)01446-i

168. Warburg O, Christian W. Isolierung und Kristallisation des Gärungsferments Enolase. Biochem. Z. 1941;310:384-421, doi: 10.1007/BF01482279

169. Webster LT, Siddiqui UA, Lucas SV, Strong JM, Mieyal JJ. Identification of separate acyl-coA: glycine and acyl-coA: L-glutamine *N*-acyltransferase activities in mitochondrial fractions from liver of rhesus monkey and man. J Biol Chem 1976;251:3352-8.

170. Weinberg W. Über den Nachweis der Vererbung beim Menschen. Jahreshefte des Vereins Vaterländische Naturkunde in Württemberg 1908;64:369-82.

171. Weiss W, Weiland F, Görg A. Protein Detection and Quantitation Technologies for Gel-Based Proteome Analysis. Proteomics, 2009;564:59-82, doi: 10.1007/978-1-60761-157-8_4

172. Williams HRT, Cox IJ, Walker DG, Cobbold JFL, Taylor-Robinson SD, Marshall SE, Orchard TA. Differences in gut microbial metabolism are responsible for reduced Hippurate synthesis in Crohn´s disease. BMC Gastroenterol 2010;10:1-10, doi: 10.1186/1471-230X-10-108

173. Wilson DB, Prescott SM, Majerus PW. Discovery of an arachidonoyl coenzyme A synthetase in human platelets. J Biol Chem 1982;257:3510-5.

174. Xiong S, Wang YF, Ren XR, Li B, Zhang MY, Luo Y, Zhang L, Xie QL, Su KY. Solubility of disulfide-bonded proteins in the cytoplasm of Escherichia coli and its “oxidizing” mutant. World J Gastroenterol 2005;11:1077-82, doi: 10.3748/wjg.v11.i7.1077

175. Yamamoto A, Nonen S, Fukuda T, Yamazaki H, Azuma J. Genetic polymorphisms of glycine *N*-acyltransferase in Japanese individuals. Drug Metab Pharmacokinet. 2009;24:114-7, doi: 10.2133/dmpk.24.114

176. Zerangue N, Malan MJ, Fried SR, Dazin PF, Jan YN, Jan LY, Schwappach B. Analysis of endoplasmic reticulum trafficking signals by combinatorial screening in mammalian cells. Proc Natl Acad Sci U. S. A.2001;98:2431–36, doi: 10.1073/pnas.051630198

177. Zhang YM, Chonhan S, Virga KG, Stevens RD, Ilkayeva OR, Wenner BR. Bain JR, Newgard CB, Lee RE, Rock CO, Jackowski S. Chemical Knockout of Pantothenate Kinase Reveals the Metabolic and Genetic Program Responsible for Hepatic Coenzyme A Homeostasis. Chem Biol 2007;14:291-302, doi: 10.1016/j.chembiol.2007.01.013

178. Zhang H, Lang Q, Li J, Zhong Z, Xie F, Ye G, Wan B, Yu L. Molecular Cloning and Characterization of a Novel Human Glycine-*N*-acyltransferase Gene *GLYATL*1, Which Activates Transcriptional Activity of HSE Pathway. Int J Mol Sci 2007;8:433-44, doi: 10.3390/I8050433

179 https://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=10249, last accessed: 23.01.2022

180 https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000149124;r=11:58640426-58731974, last accessed: 23.01.2022

181 https://www.ncbi.nlm.nih.gov/gene/?term=ENST00000317391, last accessed: 23.01.2022

182 https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000166840;r=11:58905398-59043527, last accessed: 23.01.2022

183 https://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=92292, last accessed: 23.01.2022

184 https://www.ncbi.nlm.nih.gov/gene/219970, last accessed: 23.01.2022

185 https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000156689;r=11:58834065-58904215, last accessed: 04.05.2022

186 https://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=219970, last accessed: 23.01.2022

187 https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000203972;r=6:49499923-49528078), last accessed: 23.01.2022

188 https://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=389396, last accessed: 23.01.2022

189 https://www.uniprot.org/uniprot/Q6IB77, last accessed: 04.05.2022

190 https://www.uniprot.org/uniprot/Q969I3, last accessed: 04.05.2022

191 https://www.uniprot.org/uniprot/Q8WU03, last accessed: 04.05.2022

192 https://www.uniprot.org/uniprot/Q5SZD4, last accessed: 04.01.2022

193 https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiLl96qgrX3AhVsSvEDHcomBs0QFnoECBoQAQ&url=https%3A%2F%2Fwww.agilent.com%2Fcs%2Flibrary%2Fusermanuals%2FPublic%2F210518.pdf&usg=AOvVaw2hvvCA9IBPY_S39e6PPSB5, last accessed: 27.04.2022

194 https://www.ncbi.nlm.nih.gov/, last accessed: 04.05.2022

195 https://blast.ncbi.nlm.nih.gov/Blast.cgi?BLAST_SPEC=blast2seq&LINK_LOC=align2seq&PAGE_TYPE=BlastSearch, last accessed: 28.04.2022

196 https://www.merckmillipore.com/DE/de/product/Origami-2DE3pLysS-Competent-Cells-Novagen,EMD_BIO-71346?ReferrerURL=https%3A%2F%2Fwww.google.de%2F&gclid=CjwKCAjw9qiTBhBbEiwAp-GE0WEgsDWmSqlg2BvvxtZWoekIcjNJKJHge9VrxPAGt3fN08Tr-qEAZRoCkI0QAvD_BwE, last accessed: 28.04.2022

197 https://gnomad.broadinstitute.org/, last accessed: 28.04.2022

198 https://www.ncbi.nlm.nih.gov/gap/, last accessed: 04.05.2022

199 https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/dna-clean-up/qiaquick-gel-extraction-kit/, last accessed: 28.04.2022

https://www.google.de/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjwpN6D_7b3AhXFCOwKHc0hB6YQFnoECAcQAQ&url=https%3A%2F%2Fwww.takarabio.com%2Fdocuments%2FUser%2520Manual%2FXfect%2520Transfection%2520Reagent%2520Protocol%2FXfect%2520Transfection%2520Reagent%2520Protocol-At-AGlance_103012.pdf&usg=AOvVaw0pwiCw8X2HXJv7qTfKtjD6, last access: 28.04.2022

201 http://cshprotocols.cshlp.org/content/2007/2/pdb.rec10732.full?text_only=true, last accessed: 28.04.2022

202 https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0011216_Ellmans_Reag_UG.pdf, last accessed: 28.04.2022

203 https://www.brenda-enzymes.org/enzyme.php?ecno=2.3.1.13#SYNONYM, last accessed: 26.01.2022

204 https://www.thermofisher.com/order/catalog/product/V79020, last accessed: 03.05.2022

205 https://www.ncbi.nlm.nih.gov/tools/primer-blast/, last accessed: 04.05.2022

206 https://www.ensembl.org/index.html, last accessed: 04.05.2022

207 https://www.creativebiomart.net/assay-kit/pdf/?q=463845, last accessed: 03.05.2022

208 https://web.expasy.org/peptide_cutter/, last accessed: 04.05.2022

209 https://www.genscript.com/location.php?href=/gsfiles/vector-map/bacteria/pET-32a.pdf, last accessed: 23.01.2022

210 https://www.genscript.com/location.php?href=/gsfiles/vector-map/mammalian/pcDNA3.1-plus.pdf, last accessed: 23.01.2022

211 https://www.genscript.com/location.php?href=/gsfiles/vector-map/mammalian/pcDNA3.1-C-eGFP.pdf, last accessed: 23.01.2022

212 https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiSmPTWrub3AhUmS_EDHeinCJwQFnoECAYQAQ&url=https%3A%2F%2Fwww.fishersci.com%2Forder%2Fgenome-database%2FgeneratePdf%3FproductName%3DGLYAT%26assayType%3DPRANT%26detailed%3Dtrue%26productId%3DPA5-48504&usg=AOvVaw3Mpfvx-B1dXUwt7SRQE39L, last accessed: 17.05.2022

213 https://www.nature.com/scitable/definition/snp-295/, last accessed: 20.05.2022

214 https://www.thermofisher.com/order/catalog/product/34094, last accessed: 20.05.2022

215 https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjG7obdhvn3AhVzS_EDHXH0Cc0QFnoECBAQAQ&url=https%3A%2F%2Fwww.takarabio.com%2Fdocuments%2FUser%2520Manual%2F3340%2F3340_e.v1701Da.pdf&usg=AOvVaw3YT4nl03wI8_N6SX9SnJtI, last accessed: 24.05.2022

216 https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjnq_bU8_33AhV8lf0HHTWXB2kQFnoECAcQAQ&url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2Flipofectamine3000_protocol.pdf&usg=AOvVaw2oSOgxDl169WTLKiBDFX0q, last accessed: 26.05.2022

217 https://www.researchgate.net/post/Transfection-of-large-DNA-plasmids-into-HEK293T, last accessed: 27.05.2022

218 https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2FMAN0011216_Ellmans_Reag_UG.pdf, last accessed: 01.06.2022

219 http://cshprotocols.cshlp.org/content/2018/3/pdb.rec098863.full?rss=1, last accessed: 10.06.2022

220 https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/dna-clean-up/qiaquick-pcr-purification-kit/, last accessed: 16.06.2022

6. Registers

6.1. List of abbreviations

aa	amino acid
ACSMs	Xenobiotic/medium-chain fatty acid: CoA ligases (acyl-coA synthetases)
Ala	Alanine
AMP	Adenosine monophosphate
APS	Ammonium peroxide sulfate
Arg	Arginine
Asn	Asparagine
ATP	Adenosine triphosphate
bp	Basepair
BSA	Bovine serum albumin
cDNA	Complementary DNA
cm	centimeter
CMTMRos	Chloromethyltetramethylrosamine
CMV	Cytomegalovirus 5
CoA	Coenzyme A
CPS	Carbamoyl-phosphate synthase
Ct	Cycle threshold
CYP	Cytochromes P450
Cys	Cysteine
DMEM	Dulbecco´s Modified Eagle´s Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleotide acid
dNTP	Deoxyribonucleotide triphosphate
DpnI	*Diplococcus pneumoniae* G41
DTNB	5,5´-dithio-bis-2-nitrobenzoic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
e.g.	exempli gratia
EcoRV	*Escherichia coli* RY13
EDTA	Ethylenediamine tetraacetic acid
eGFP	enhanced green fluorescent protein
ER	endoplasmic reticulum
FBS	Fetal bovine serum
FCS	Fetal calf serum
FPLC	Fast protein liquid chromatography
FS	First strand
fw	Forward
g	Gravity
GC-MS	Gas chromatography-mass spectrometry
Gcn5	General control non-depressible 5
gDNA	Genomic DNA
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
GLYAT	Glycine *N*-acyltransferase
GLYATL1	Glutamine *N*-phenylacetyltransferase
GLYATL2	Glycine *N*-acyltransferase like protein 2
GNAT	Gcn5-related *N*-acyltransferase
GRCh	Genome Reference Consortium Human
h	hour/s
HEK293	a human embryonic kidney cell line
HEK293T	a human embryonic kidney cell line including T antigen
HeLa	Henrietta Lacks
HepG2	a hepatocellular carcinoma cell line
HindIII	*Haemophilus influenzae*
His	Histidine
HRP	Horseradish peroxidase
i.e.	id est
Ile	Isoleucine

IPTG	Isopropyl-β-D-thiogalactopyranoside
IVA	Isovaleric acidemia
kbp	Kilo base pairs
L	Liter
LB	Luria-Bertani
Leu	Leucine
LEW	Lysis equilibration washing
Lys	Lysine
M	Molar
MAF	Minor allelic frequency
MATD	Methylacetoacetyl-coA thiolase deficiency
MCS	Multiple cloning site
Met	Methionine
min	Minutes
mL	Milliliter
MMA	Methylmalonic acidemia
M-MuLV	Moloney Murine Leukemia Virus
MOPS	3-(*N*-morpholino)propanesulfonic acid
MTHFD1L	Mitochondrial C1-Tetrahydrofolate synthase
MTHFD2	Methylenetetrahydrofolate dehydrogenase/cyclohydrolase
MWCO	Molecular weight cut-off
n.d.	not detected
NheI	*Neisseria mucosa heidelbergensis*
NKH	Nonketotic hyperglycemia
NotI	*Nocardia otididis-caviarum*
NZY^+ broth	NZ amine (casein hydrolysate) and yeast
o/n	over night
OD	Optical density
opti-MEM	Optimized modified Eagle´s medium
OTC	Ornithine transcarbamoylase
PA	Propionic acidemia
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Phe	Phenylalanine
PIC	protease inhibitor cocktail
Pro	Proline
PTMs	posttranslational protein modifications
PVDF	polyvinylidene fluoride
rcf	Relative centrifugal force
rev	Reverse
RIPA	Radioimmuno precipitation assay
RNA	Rbonucleotide acid
rpm	Rounds per minute
rs	Reference SNP
rSAP	Recombinant shrimp alkaline phosphatase
RT	Room temperature
RT-qPCR	Reverse transcription quantitative real-time PCR
SacI	*Streptomyces achromogenes*
SDS	Sodium dodecyl sulfate
Ser	Serine
SHMT2	Serine hydroxymethyltransferase 2
SNP	Single nucleotide polymorphism
SOC	Super optimal broth
SSIV	SuperScript™ IV
Suppl. Fig.	Supplemental Figure
T	Temperature
TAE	Tris acetate EDTA
Taq	*Thermus aquaticus*
TBS	Tris buffered saline
TBS-T	Tris buffered saline and polysorbate 20
TC	Tissue culture
TE	Tris-EDTA

TED	Tris(carboxymethyl)ethylene diamine
TEMED	N, N, N´, N´-tetramethylethlyendiamine
Thr	Threonine
TOMM20	Translocase of the outer mitochondrial membrane complex subunit 20
TRI	Total RNA isolation
Trp	Tryptophane
Trx	Thioredoxin
Tyr	Tyrosine
UV	Ultraviolet
µL	Micro liter
V	Voltage
v	Volume
Val	Valine
w	Weight
X-gal	5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside
[digit]	Literature reference
[digit]	Website reference

6.2. List of figures

6.3. List of tables

7. Zusammenfassung

Die humanen Enzyme GLYAT (Glyzin *N*-Azyltransferase), GLYATL1 (Glutamin *N*-Phenylazetyltransferase) und GLYATL2 (Glyzin *N*-Azyltransferase ähnliches Protein 2) sind nicht nur bei der Entgiftung von Fremdstoffen über die menschliche Leber wichtig, sondern auch an der Eliminierung von Azylresten beteiligt, die bei einigen seltenen angeborenen Stoffwechseldefekten in Form ihrer Coenzym A (CoA)-Ester akkumulieren. Dies betrifft zum Beispiel Störungen im Abbau von verzweigtkettigen Aminosäuren, wie die Isovalerianazidämie oder Propionazidämie. Daneben unterstützen sie auch bei der Elimination von Ammonium, welches bei der Transaminierung von Aminosäuren anfällt und sich bei Harnstoffzyklusdefekten anreichert. Es wurden auch Sequenzvarianten der Enzyme untersucht, die Hinweise auf eingeschränkte Enzymaktivitäten liefern können, woraus sich potentiell Therapieanpassungen ableiten lassen.

Für die Überexpression und partielle biochemische Charakterisierung der Enzyme wurde ein modifizierter *E. coli* Stamm gewählt, der die Löslichkeit und die korrekte Faltung ermöglichen kann. Da posttranslationale Proteinmodifikationen in Bakterien sehr limitiert sind, wurde außerdem versucht, die Enzyme in menschlichen HEK293 Zellen zu überexprimieren. Neben der Charakterisierung über Immunoblots und Aktivitätsbestimmungen wurde auch die intrazelluläre Lokalisierung der Enzyme mittels GFP-Kopplung und konfokaler Laserscanning-Mikroskopie in transfizierten HEK293 Zellen durchgeführt. Für das GLYATL2-Enzym, für welches Aufgaben jenseits der Entgiftung und Stoffwechseldefekte postuliert wurden, sind die molekularbiologischen Vorarbeiten im Rahmen dieses Projekts durchgeführt worden - die Enzymaktivitätsbestimmungen wurden in eine mitbetreuten Bachelorarbeit ausgelagert.

Die Enzymaktivitätsbestimmungen mit aufgereinigtem rekombinanten Human-Enzym aus *E. coli* lieferten für GLYAT eine dreifach höhere Aktivität der Sequenzvariante p.(Asn156Ser), die als wahrscheinlich authentischer Wildtyp des Enzyms gelten sollte. Außerdem zeigte sich eine verringerte Aktivität der in Südafrika sehr häufigen GLYAT-Variante p.(Gln61Leu), was vor allem bei Behandlung der ebenfalls in Südafrika häufigen Isovalerianazidämie von Bedeutung sein könnte. Intrazellulär konnten GLYAT und GLYATL1 mitochondrial lokalisiert werden.

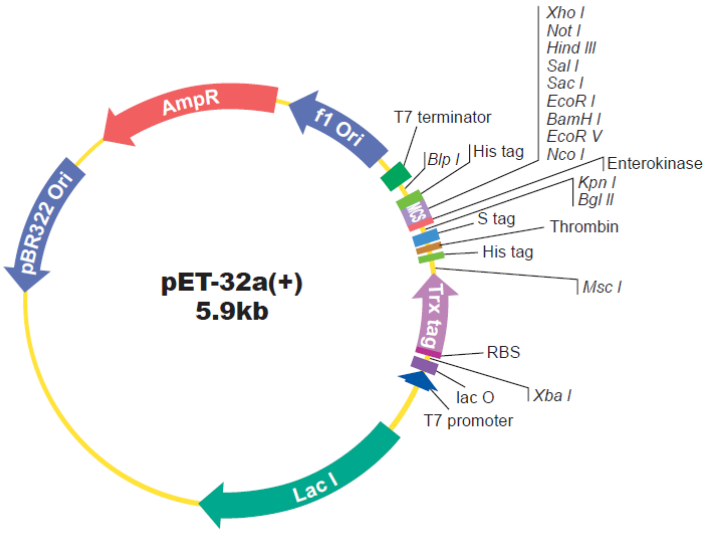
Wie die Analysen gezeigt haben, beeinflussen Sequenzvariationen von GLYAT und GLYATL1 deren Enzymaktivität. Als Beispiel ist die in Südafrika häufig vorkommende GLYAT Variante p.(Gln61Leu) zu nennen. Im Falle verminderter GLYAT-Aktivität könnten Patienten im Sinne einer individualisierten Therapie vermehrt mit L-Carnitin behandelt werden, da die Konjugation des toxischen Isovaleryl-CoA mit Glyzin durch die GLYAT-Sequenzvariation eingeschränkt ist. Aktivitätsmindernde Varianten, die in diesem Projekt identifiziert wurden, sind dabei von besonderer Relevanz, da sie die Behandlung gewisser Stoffwechseldefekte beeinflussen können.

8. Danksagung

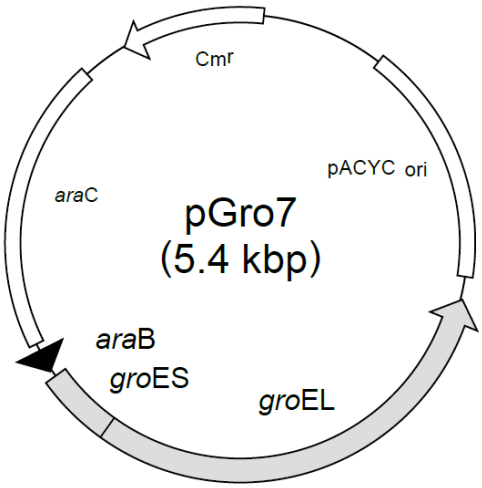
Einen besonderen Dank möchte ich an dieser Stelle Prof. Dr. Jörn Oliver Saß aussprechen, der mir die Gelegenheit gab, in seiner Forschungsgruppe eine Dissertation anzufertigen. Ich habe mich stets sehr gut betreut und in die Versuchsplanung einbezogen gefühlt, wobei ich ein hohes Maß an Selbstständigkeit genießen konnte.
Weiterhin danke ich auch Herrn Prof. Dr. Michael Gütschow von der Universität Bonn für die freundliche Übernahme des Korreferats und für die Möglichkeit, an Seminaren seiner Arbeitsgruppe teilzunehmen.
Für finanzielle Unterstützung danke ich der Heinz und Heide Dürr Stiftung (Berlin) und der Gottfried und Julia Bangerter-Rhyner-Stiftung (Basel), sowie dem Graduierteninstitut der Hochschule Bonn-Rhein-Sieg.

Des Weiteren danken möchte ich:
Prof. Dr. Dieter Reinscheid für die sehr zuvorkommende Einführung in die FPLC Anlage zum Start meines Projektes.
Prof. Dr. Hans Weier und Dr. Barbara Roitzheim für die sehr ansprechenden und ausführlichen molekulargenetischen Diskussionen.
Frau Lil Klaas, Dr. Anne Korwitz-Reichelt und Dr. Claudia Till für den freundlichen und sinnvollen Austausch innerhalb der Arbeitsgruppe, der mir stets weitergeholfen hat.
Dr. Chakravarthi Chintalapati und PD Dr. Renate Voit für die Diskussionen im Labor, die mein Projekt gut vorangetrieben haben.
Patrick Ottensmeyer, Christian Tonk, Patrick Babczyk und Glenn Theunissen für die freundliche Atmosphäre im Büro und den Austausch zu nicht immer wissenschaftlichen Themen.
Helga Kleinert und allen Auszubildenen für die unermüdliche Unterstützung im Bereich des Laborbedarfs.
Imke von Stülpnagel und Jana Hollenbeck, die ich einen Teil auf ihrem Weg zum Bachelorgrad begleiten durfte, sowie Despina Tsortouktzidis, Sonya Aktar, Niloufar Mosaddeghzadeh und Aghyad Al Kabbani, die während meiner Laborphase eine erfolgreiche Masterarbeit in unserer Arbeitsgruppe angefertigt haben.
Nicht zuletzt allen aktuellen Kolleg*innen der Arbeitsgruppe, des Fachbereichs und natürlich Migena und meiner Familie für die unermüdliche Unterstützung während der Promotion.

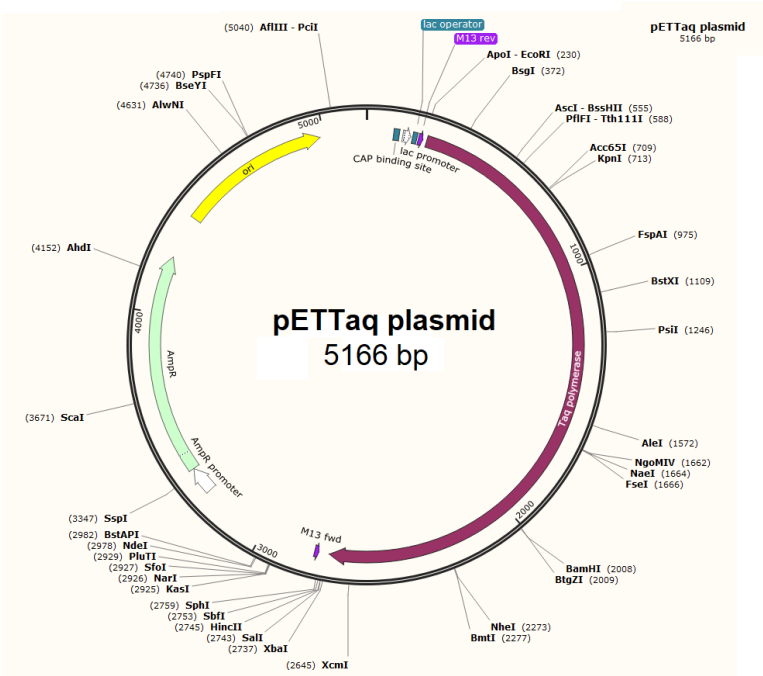
9. Appendix



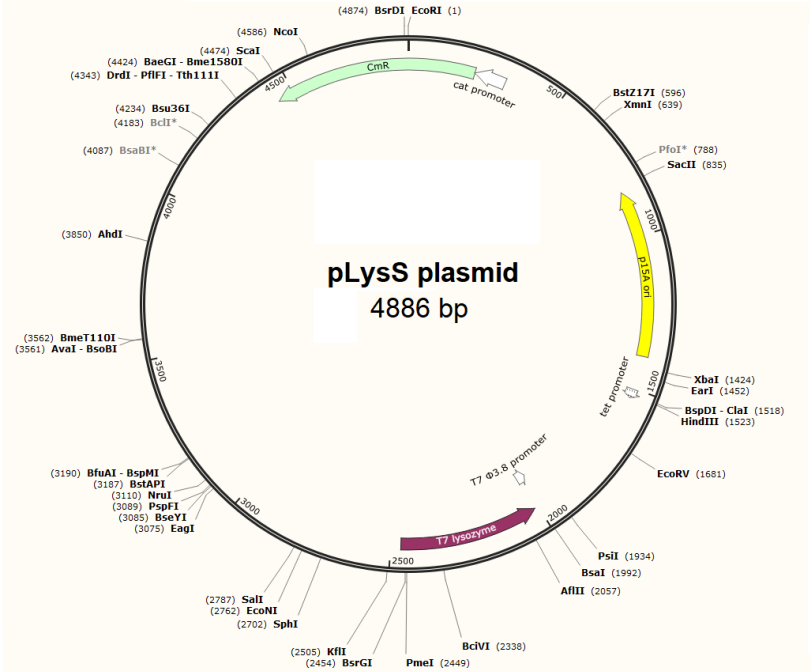
Supplemental Figure 1: pET32a(+) plasmid, ordered from Genscript, was used for overexpression of GLYAT and related genes in bacterial overexpression system *E. coli* Origami 2(DE3). The figure was provided by Genscript. [209]. The plasmid contains inducible T7 promotor of *lac-operon* for controlled induction of target protein, which is fused to *N*-terminal thioredoxin (Trx) *tag* for optimized solubility and stability of the target protein. An ampicillin resistance cassette is introduced for selection of the transformed clones. *N*- and *C*-terminal His-*tag* enables metal affinity chromatography of overexpressed fusion protein. Enterokinase site allows cleavage of Trx *tag* via specific protease on protein level.



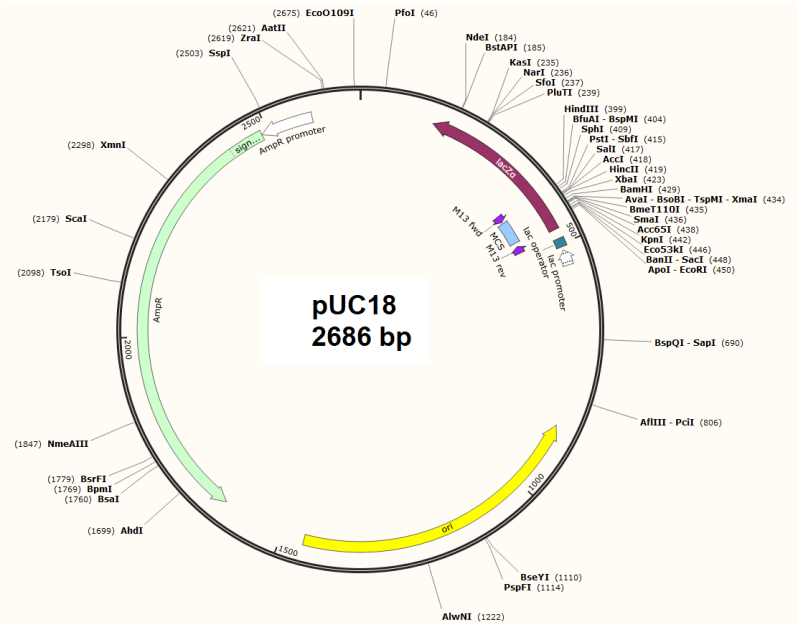
Supplemental Figure 2: pGro7 plasmid allowed co-expression of *E. coli* chaperone system GroES/GroEL for improving folding of target proteins in bacterial cells. pGro7 was co-transformed in *E. coli* Origami 2(DE3) overexpression of *GLYAT* and related genes to ensure optimal folding. The vector map was provided by Takara Bio [215]. Positive clones were selected exploiting Cloramphenicol resistance cassette Cm^r. Specific induction of chaperone system GroES/GroEL was perfomed via arabinose operon *araB* possible through the usage of 2 mg/ml L-arabinose in overexpression medium.



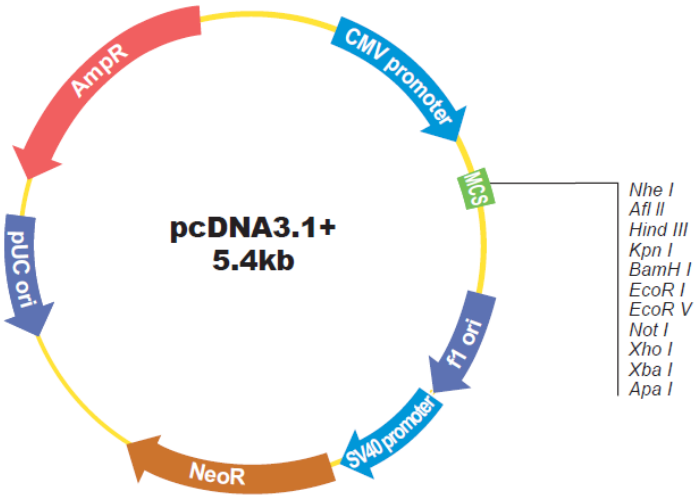
Supplemental Figure 3: pETTaq plasmid for overexpression of recombinant *Taq* DNA polymerase from *Thermus aquaticus* (SnapGene Viewer). For selection of positive clones Ampicillin resistance cassette was encoded. For the overexpression of *Phusion* DNA polymerase the same plasmid as shown here was used only differentiated by specific *Phusion* insert. The *Phusion* insert is characterized by a DNA-binding domain fused to *Pyrococcus*-like proof reading polymerase (*Pfu* polymerase).



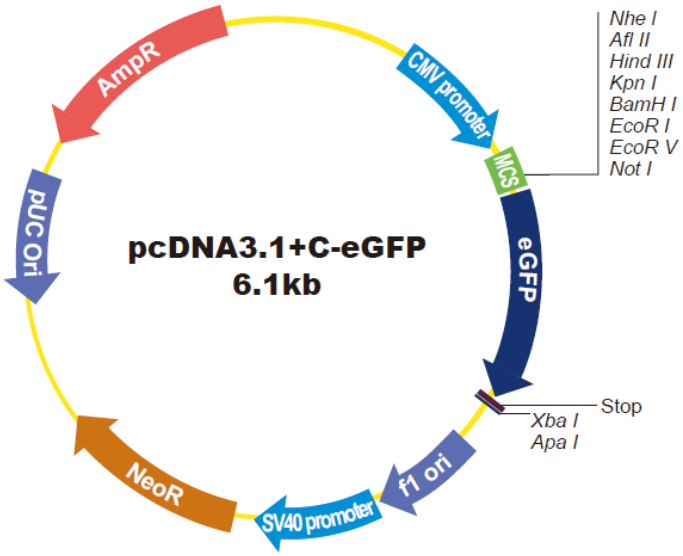
Supplemental Figure 4: pLysS plasmid additionally transformed into *E. coli* BL21 (D3) expressing T7 lysozyme to lower basal expression of target genes controlled by T7 promotor (SnapGene Viewer). The controllable T7 promotor is required for overexpression of toxic genes in *E. coli*. It is not interfering with IPTG induced overexpression of *Taq* and *Phusion* polymerase genes, hence ensuring specific and controlled overexpression here. *T7 lysozyme* gene express low levels of T7 lysozyme, the inhibitor of T7 RNA polymerase. A Chloramphenicol resistance cassette is included as open reading frame (*ORF*) for the selection of positive clones.



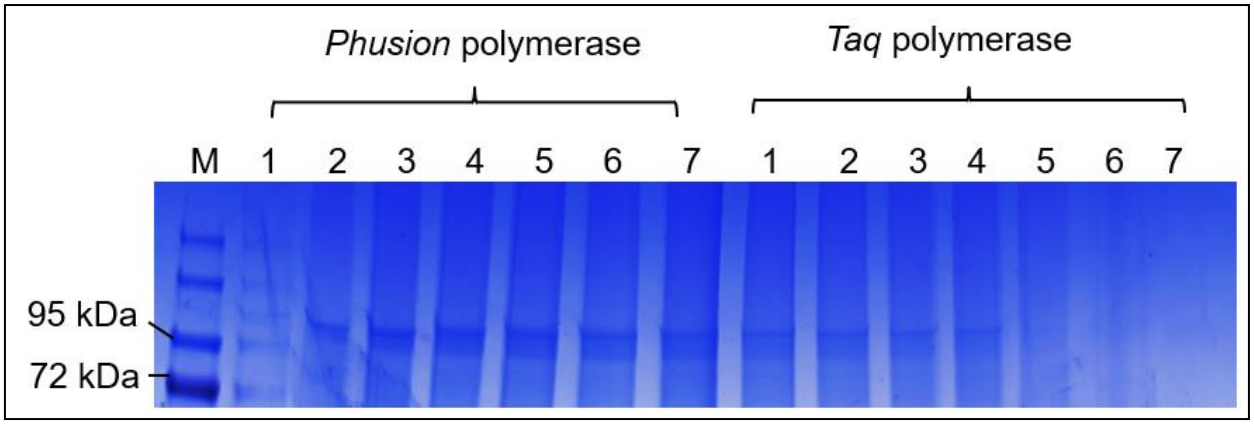
Supplemental Figure 5: pUC18 vector map showing transformation efficiency test plasmid (SnapGene Viewer). The pUC18 vector was used to determine transformation efficiency and competence of self-made competent cells *E. coli* DH10B. Amp^R allows selection for transformed clones.



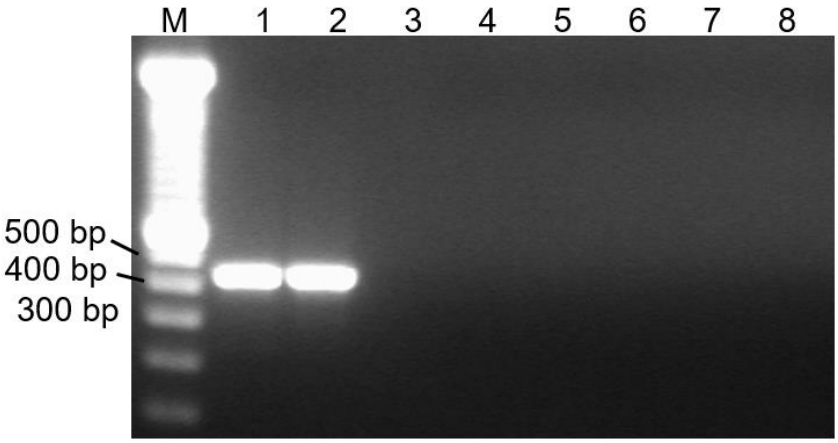
Supplemental Figure 6: pcDNA3.1(+) vector, ordered from Genscript, ensures overexpression of target genes in eukaryotic cell models e.g., HEK293, HeLa or HepG2 cells. The figure was provided by Genscript [210]. Vector is equipped with *CMV* promotor allowing specific and controlled overexpression of target genes in eukaryotic cells. Resistance cassette Amp^R permits selection in prokaryotes, while Neo^R ensures selection in eukaryotic cells using G-418 antibiotic. 2 origins of replication, *pUC ori* and *f1 ori*, enable replication in bacteria as well as in eukaryotic cells.



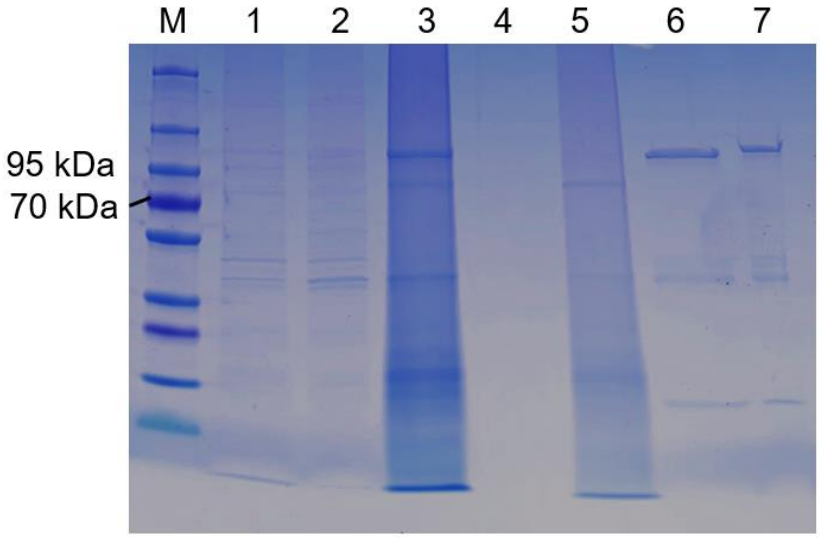
Supplemental Figure 7: pcDNA3.1+C-eGFP vector was ordered from Genscript and used for overexpression of CeGFP fused target proteins allowing confocal laser scanning microscopy approaches. The figure was provided by Genscript [211]. pcDNA3.1+C-eGFP plasmid equals pcDNA3.1(+) vector shown in figure 3, but additionally contains C-terminal *eGFP tag* used as fluorescence label for overexpressed target proteins. eGFP, which has a size of 26 kDa, marks target proteins revealing their intracellular localization.



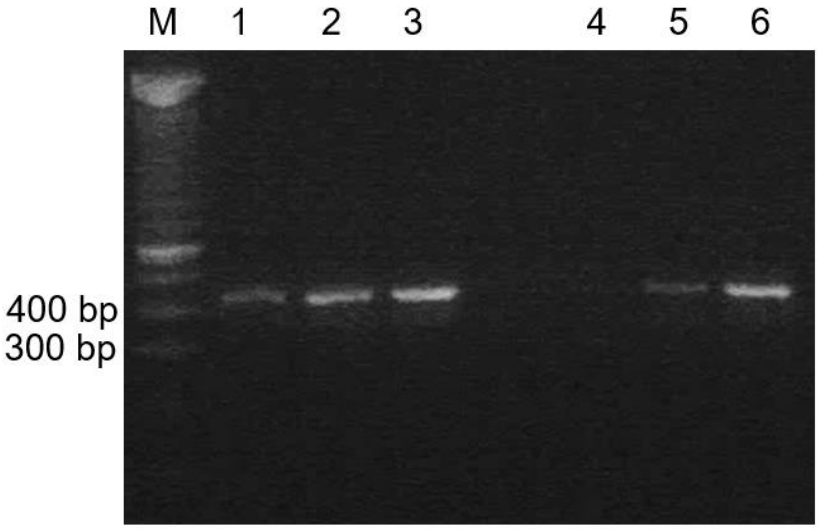
Supplemental Figure 8: Coomassie stained 12 % (v/v) SDS PAGE of purification of *Taq* and *Phusion* polymerase. M: PageRuler™ Prestained Protein Ladder, 1: fraction pre-induction, 2: fraction after induction, 3: fraction supernatant, 4: fraction flow-through, 5: wash fraction, 6: elution fraction, 7: desalted fraction. 20 µL of each fraction were loaded. *Taq* polymerase has a size of 94 kDa, while *Phusion* polymerase is a 95 kDa protein.



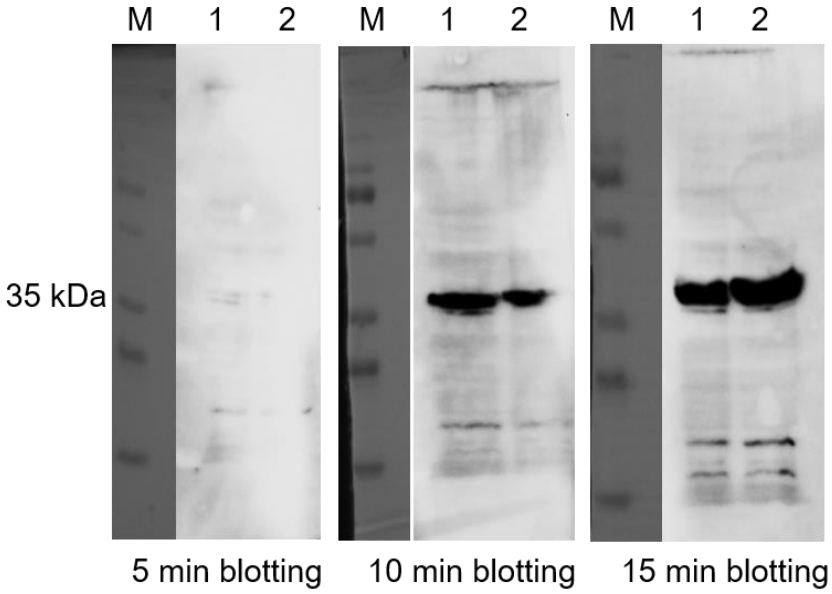
Supplemental Figure 9: Agarose gel (1.5 %(w/v)) of in-house produced *Phusion* and *Taq* polymerase undiluted, 1:50 and 1:100 diluted. M: GeneRuler 100 bp DNA ladder, 1: *Phusion* polymerase undiluted, once thawed, 2: Phusion polymerase undiluted, 3 times thawed, 3: *Taq* polymerase undiluted, once thawed, 4: *Taq* polymerase 1:50, once thawed, 5: *Taq* polymerase 1:100, once thawed, 6: *Taq* polymerase undiluted, 3 times thawed, 7: *Taq* polymerase 1:50, 3 times thawed, 8: *Taq* polymerase 1:100, 3 times thawed. The samples were loaded to 10 µL and gel incubated for 25 min at 120 V. Amplicon of *Aspartoacylase 2* (ASPA) gene exon 6 has a size of 424 bp.



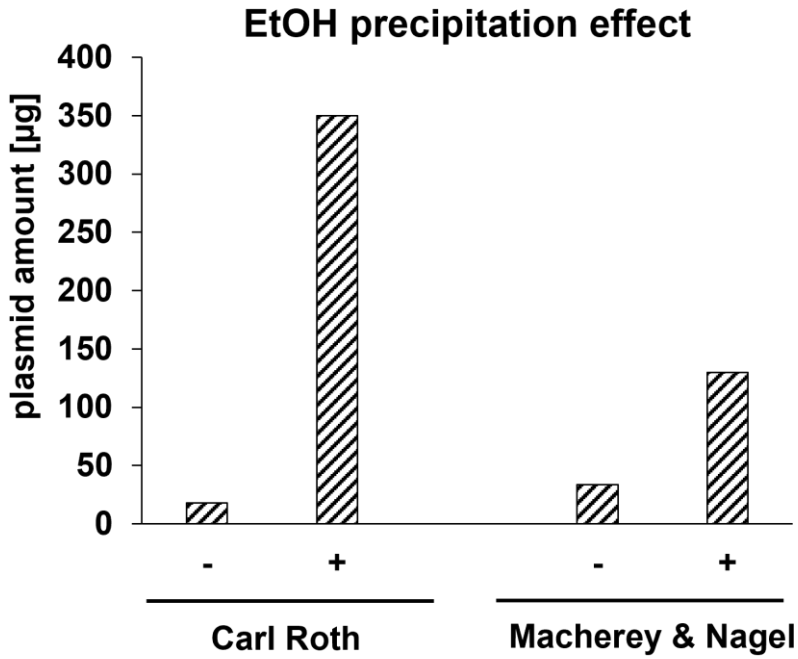
Supplemental Figure 10: Coomassie stained 12 %(v/v) SDS PAGE of the second purification of *Taq* polymerase. M: PageRuler™ Prestained Protein Ladder, 1: fraction pre-induction, 2: fraction after induction, 3: fraction supernatant, 4: fraction flow-through, 5: wash fraction, 6: elution fraction, 7: desalted fraction. 10 µL of each fraction were loaded on the gel. *Taq* polymerase has a size of 94 kDa visible in elution fraction and desalted protein fraction.



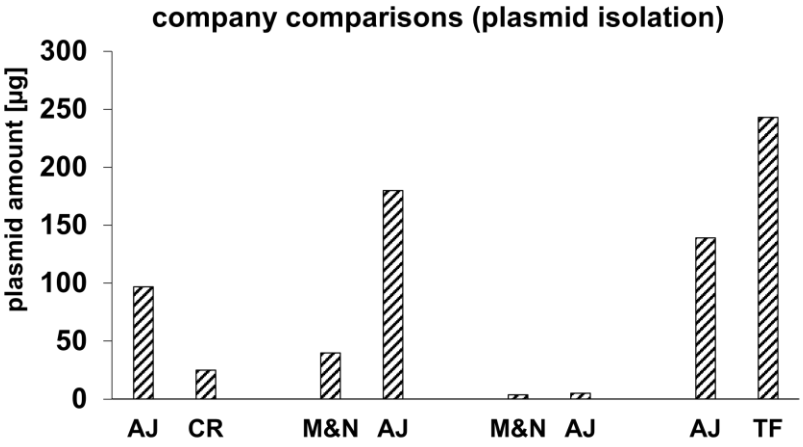
Supplemental Figure 11: 1.5 % (w/v) agarose gel of self-isolated *Taq* polymerase in 3 different amounts (0.25, 0.5, 1 µL), once and twice thawed. M: GeneRuler 100 bp DNA ladder, 1: 0.25 µL *Taq* polymerase, once thawed, 2: 0.5 µL *Taq* polymerase, once thawed, 3: 1 µL *Taq* polymerase, once thawed, 4: 0.25 µL *Taq* polymerase, twice thawed, 5: 0.5 µL *Taq* polymerase, twice thawed, 6: 1 µL *Taq* polymerase, twice thawed. Samples were loaded to 10 µL and gel incubated for 25 min at 120 V. Amplicon of exon six of Aspartoacylase 2 (ASPA) gene has a size of 424 bp.



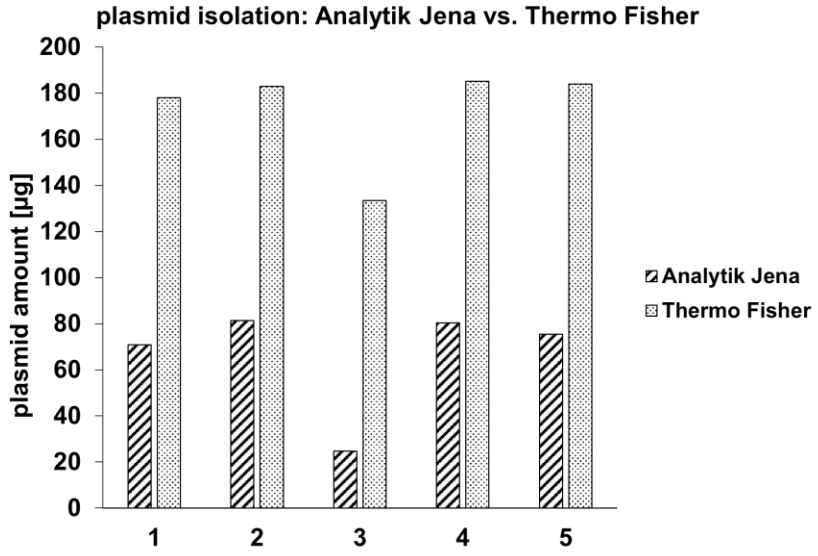
Supplemental Figure 12: Western blot of 30 µg of transfected HEK293 cells with human GLYAT wild-type and pcDNA3.1(+) empty vector as negative control. M: PageRuler™ Prestained Protein Ladder, 1: HEK293 + pcDNA3.1(+), 2: HEK293 + GLYAT wild-type.



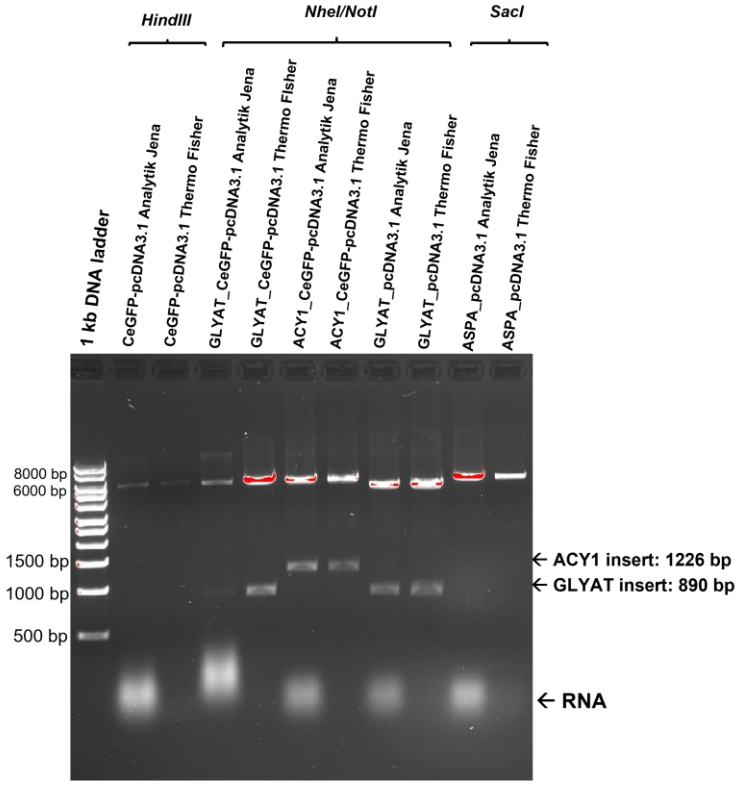
Supplemental Figure 13: EtOH precipitation [125] during plasmid isolation with Carl Roth and Macherey&Nagel kit in comparison. Adding EtOH precipitation before spin-column application leads to improvement of obtained plasmid amounts, in case of Carl Roth kit 18x, in Macherey & Nagel isolation 4x. (- = without EtOH precipitation, + = with EtOH precipitation).



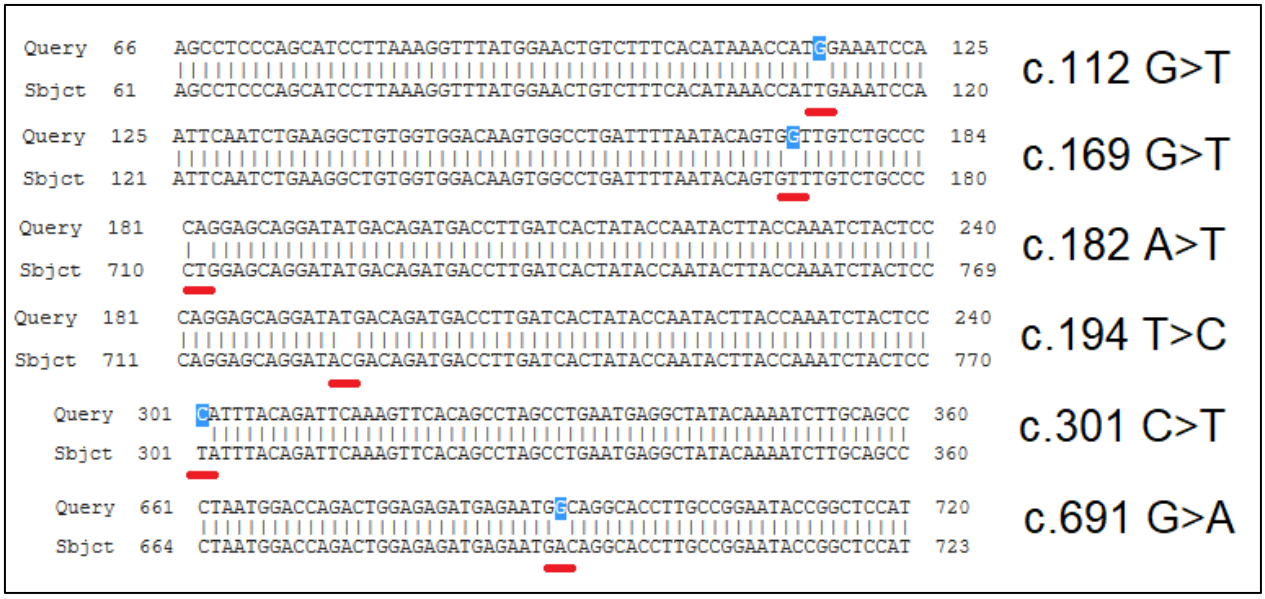
Supplemental Figure 14: Company comparisons of different plasmid isolation kits. AJ: Analytik Jena, CR: Carl Roth, M&N: Macherey & Nagel, TF: Thermo Fisher. For each experiment pcDNA3.1(+) vector was isolated in parallel with 2 different company kits. Thermo Fisher leaded tot he highest amount of isolated plasmid (243 µg).



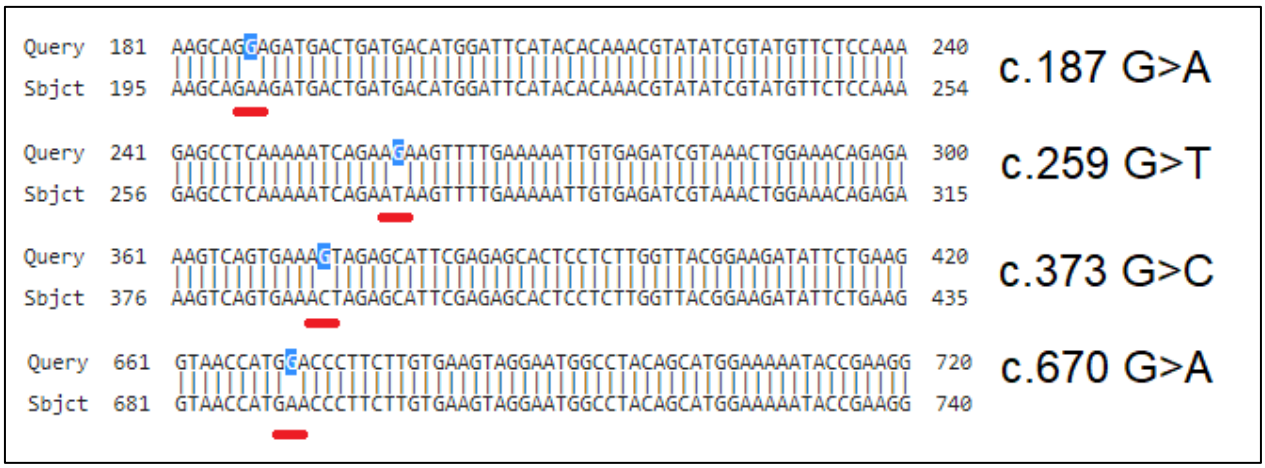
Supplemental Figure 15: Comparison of plasmid isolation kits of Analytik Jena and Thermo Fisher. Different plasmids were isolated in parallel with Analytik Jena and Thermo Fisher kit: 1: CeGFP-pcDNA3.1(+), 2: ACY1-eGFP-pcDNA3.1(+), 3: GLYAT-eGFP-pcDNA3.1(+), 4: ASPA-pcDNA3.1(+), 5: GLYAT-pcDNA3.1(+). In all cases Thermo Fisher leaded tot he higher amounts of isolated plasmid. Analytik Jena yielded in 66.6 ± 23.7 µg plasmid, whereas Thermo Fisher delivered 173 ± 22.1 µg total plasmid.



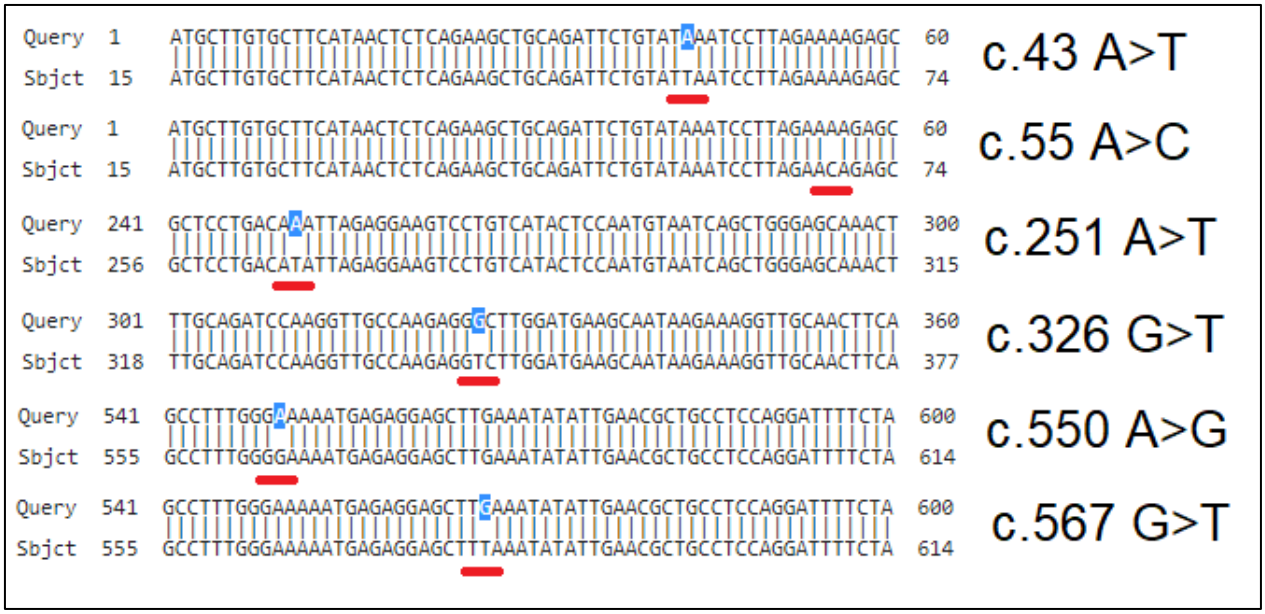
Supplemental Figure 16: Agarose gel (1.5 % (w/v)) as an addition to supplemental figure 15. Test restriction of all isolated constructs with Analytik Jena and Thermo Fisher plasmid isolation kit was carried out. Expected inserts were obtained in all cases, but in case of Analytik Jena isolation always big contaminating RNA signals are visible. Due to the more worktime and the more expensive kit this fact supports the decision for Thermo Fisher plasmid isolation kit for further plasmid isolations in the laboratory.



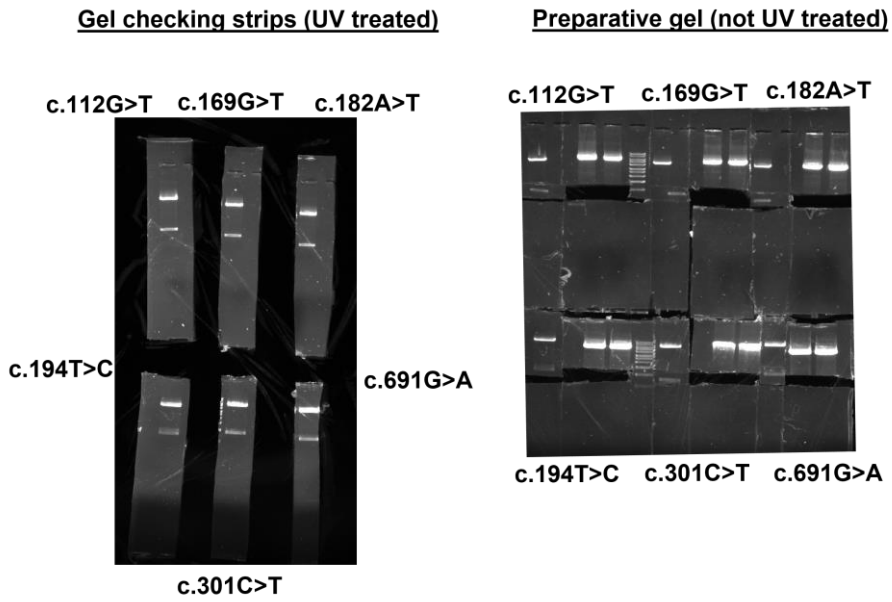
Supplemental Figure 17: Sequence alignment of produced *GLYAT* sequence variants in pcDNA3.1(+) vector. Exchanged nucleotides are labelled accordingly.



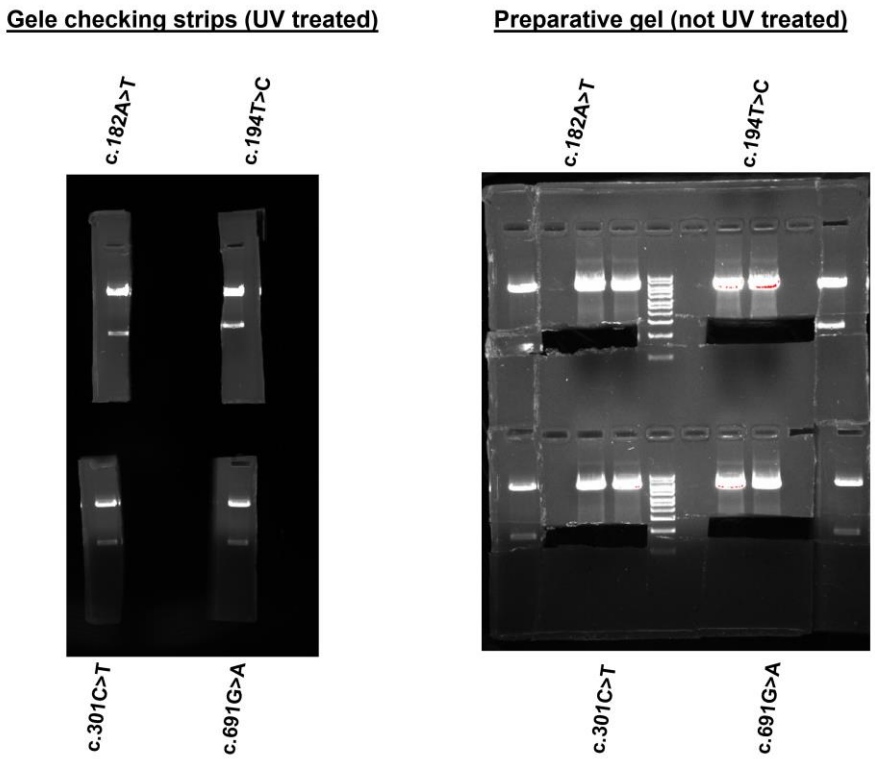
Supplemental Figure 18: Sequence alignment of produced *GLYATL1* sequence variants in pcDNA3.1(+) vector. Exchanged nucleotides are labelled accordingly.



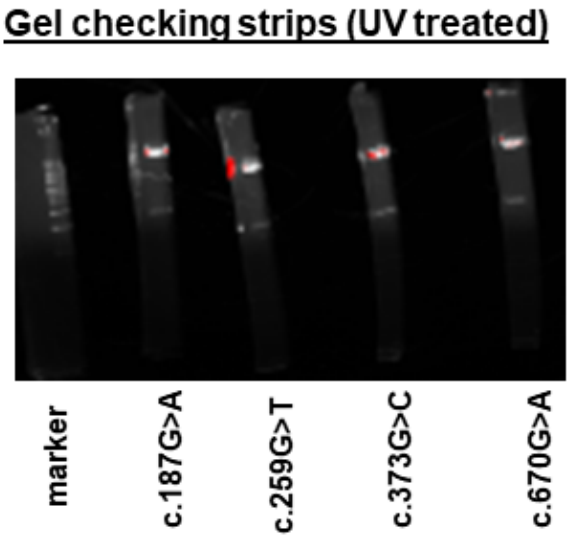
Supplemental Figure 19: Sequence alignment of produced *GLYATL2* sequence variants in pcDNA3.1(+) vector. Exchanged nucleotides are labelled accordingly.



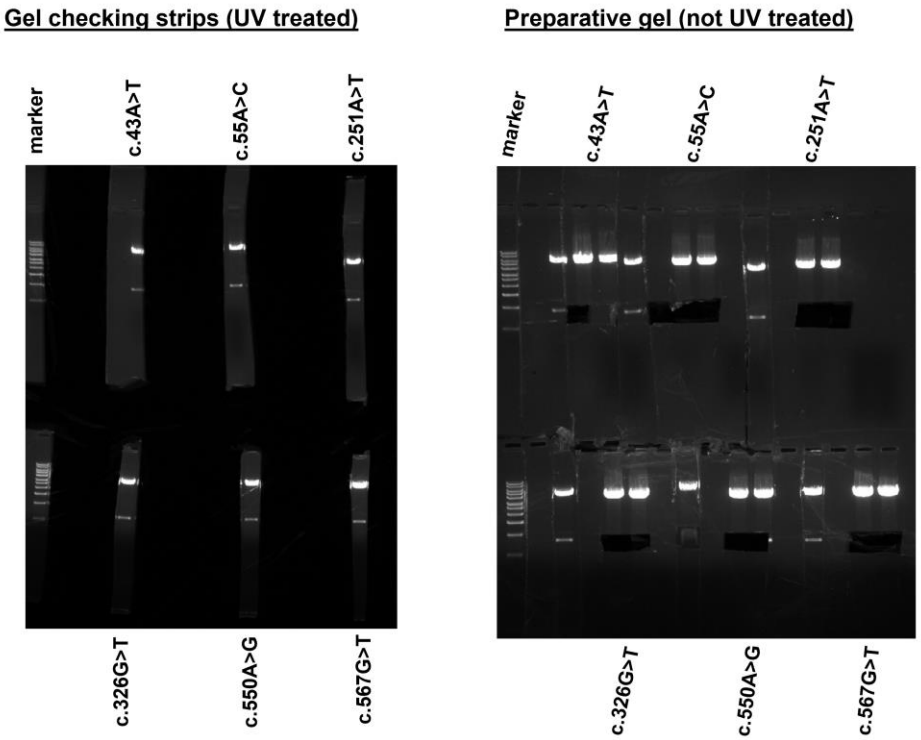
Supplemental Figure 20: Gel checking strips with UV treatment and preparative gel without UV treatment to save DNA from unspecific mutations. GLYAT c.467A>G mutagenic inserts are shown on the left side and the cutted insert gaps are visible on the right hand gel. The DNA was used for agarose gel purification by using the Qiagen kit (Section 2.2.) in the next step. Subsequently it was used for the ligation into the pET32a(+) vector (Section 3.4.8.).



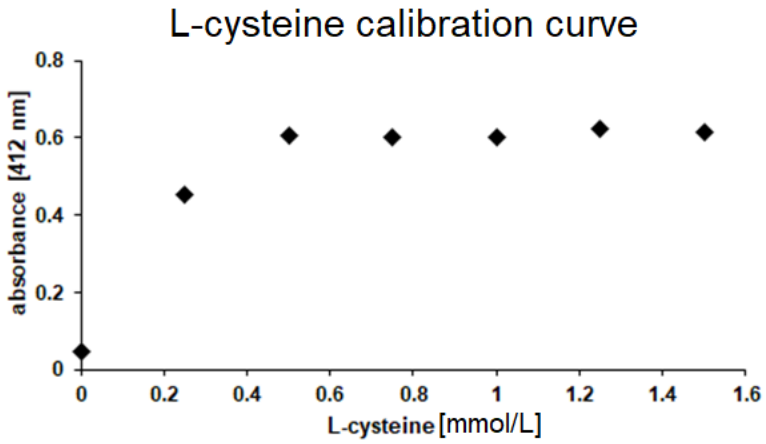
Supplemental Figure 21: Gel checking strips with UV treatment and preparative gel without UV treatment to save DNA from unspecific mutations. The *GLYAT* mutagenic inserts are shown on the left side and the cutted insert gaps are visible on the right side. The DNA was used for agarose gel purification by using the Qiagen kit (Section 2.2.) in the next step. Subsequently it was used for the ligation into the pET32a(+) vector (Section 3.4.8.).



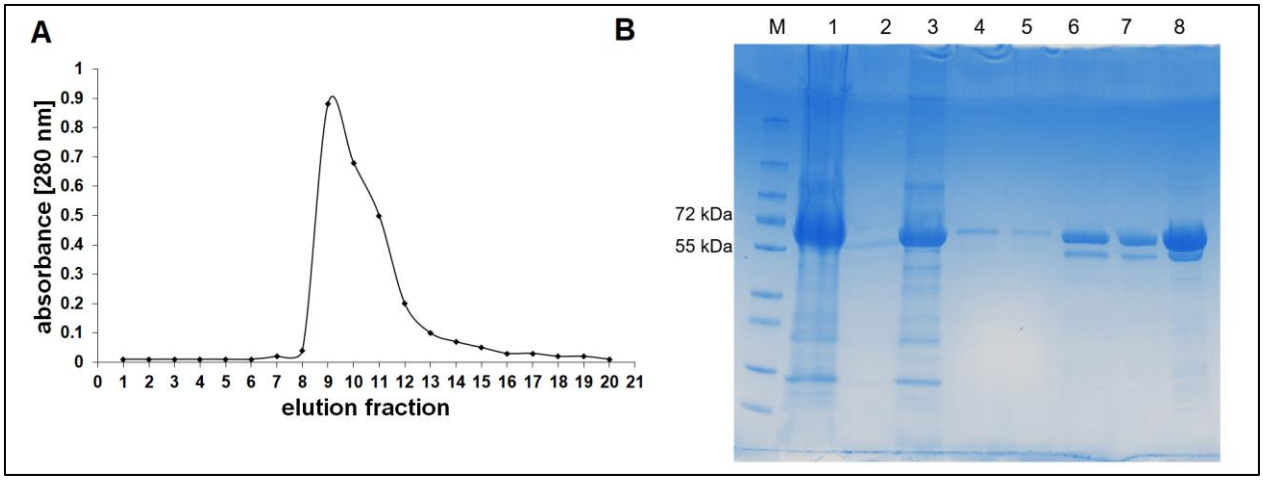
Supplemental Figure 22: Gel checking strips with UV treatment and preparative gel without UV treatment to save DNA from unspecific mutations. The *GLYATL1* mutagenic inserts are shown which were cutted for further cloning step (preparative gel part). The DNA was used for agarose gel purification by using the Qiagen kit (Section 2.2.). Subsequently it was used for the ligation into the pET32a(+) vector (Section 3.4.8.).



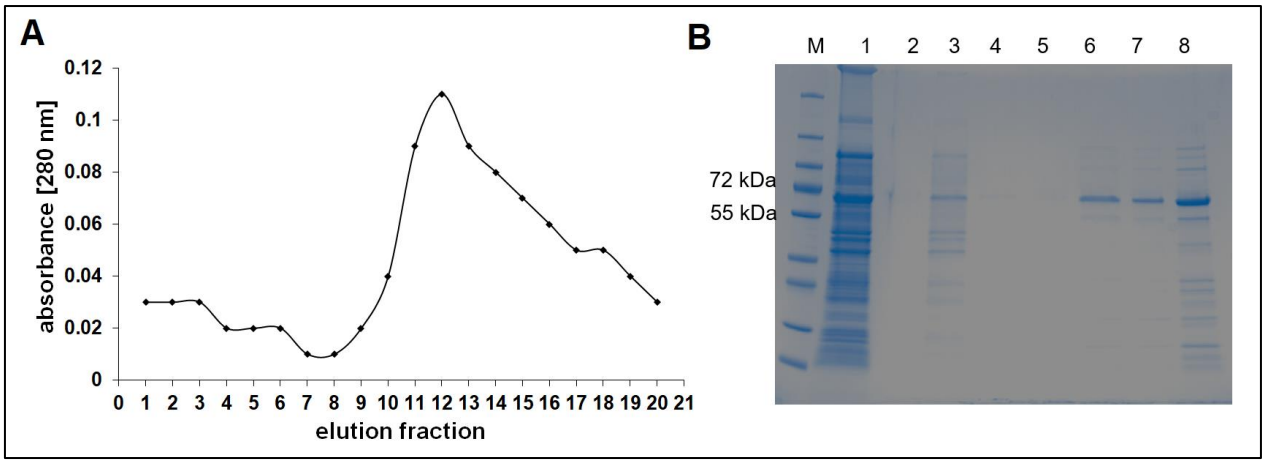
Supplemental Figure 23: Gel checking strips with UV treatment and preparative gel without UV treatment to save DNA from unspecific mutations. The *GLYATL2* mutagenic inserts are shown on the left side and on the right side the cutted insert gaps are visible. The DNA was used for agarose gel purification by using the Qiagen kit (Section 2.2.) in the next step. Subsequently it was used for the ligation into the pET32a(+) vector (Section 3.4.8.).



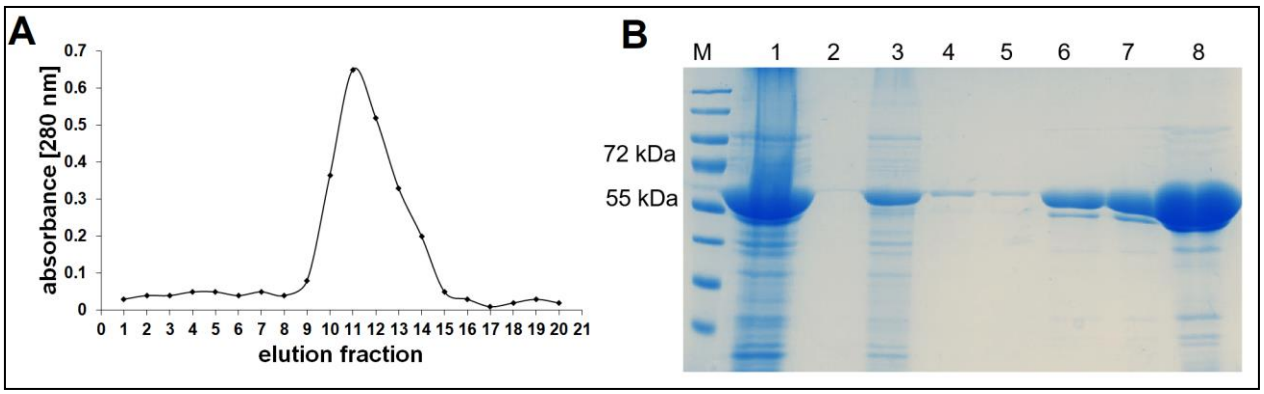
Supplemental Figure 24: L-cysteine standard curve with a concentration range of 0 to 1.5 mmol/L L-cysteine. Absorbance was detected at 412 nm. Decrease of linearity is obvious for concentrations higher than 0.7 mmol/L L-cysteine.



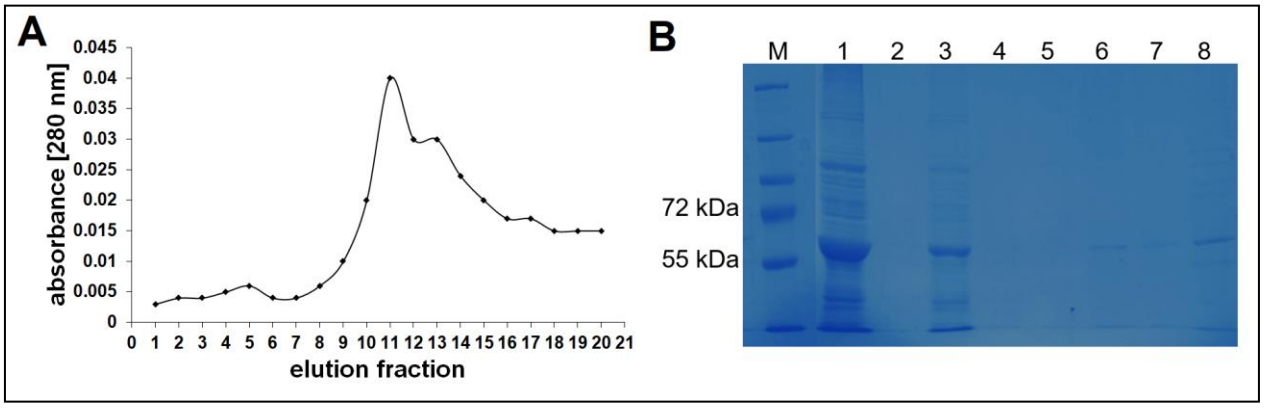
Supplemental Figure 25: Elution chromatogram (A) and SDS-PAGE (B) of GLYAT p.(Asn156Ser) purification. A: Elution fractions 7-15 were combined and transferred into dialysis tubing; B: M: PageRuler™ Prestained Protein Ladder, 1: *E. coli* Origami 2(DE3) homogenate supernatant, 2: flow-through, 3: wash fraction 1, 4: wash fraction 2, 5: elution fraction 3, 6: pooled elution fractions, 7: desalted protein, 8: pre-concentrated.



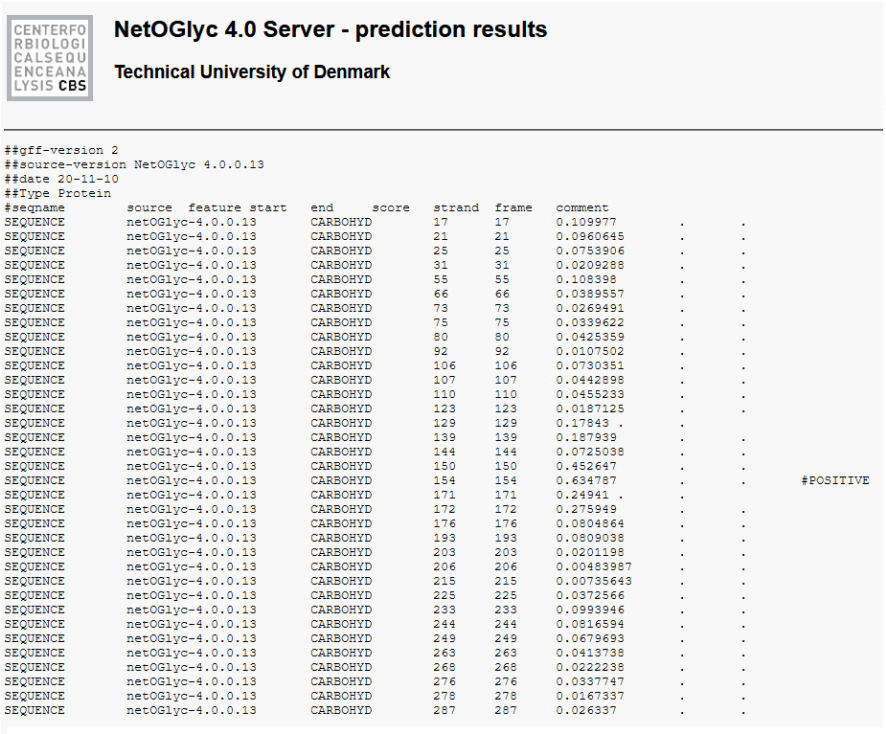
Supplemental Figure 26: Elution chromatogram (A) and SDS-PAGE (B) of GLYAT p.(Arg199Cys) purification. A: Elution fractions 9-16 were combined and transferred into dialysis tubing for dialysis; B: M: PageRuler™ Prestained Protein Ladder, 1: *E. coli* Origami 2(DE3) homogenate supernatant, 2: flow-through, 3: wash fraction 1, 4: wash fraction 2, 5: elution fraction 3, 6: pooled elution fractions, 7: desalted protein, 8: pre-concentrated protein.



Supplemental Figure 27: Elution chromatogram (A) and SDS-PAGE gel (B) of GLYAT p.(Gln61Leu) purification. A: Elution fractions 9-15 were combined and transferred into dialysis tubing; B: M: PageRuler™ Prestained Protein Ladder, 1: *E. coli* Origami 2(DE3) homogenate supernatant, 2: flow-through, 3: wash fraction 1, 4: wash fraction 2, 5: elution fraction 3, 6: pooled elution fractions, 7: desalted protein, 8: pre-concentrated.



Supplemental Figure 28: Elution chromatogram (A) and SDS-PAGE gel (B) of GLYAT p.(Ala231Thr) purification. A: Elution fractions 10-15 were combined and transferred into dialysis tubing; B: M: PageRuler™ Prestained Protein Ladder, 1: *E. coli* Origami 2(DE3) homogenate supernatant, 2: flow-through, 3: wash fraction 1, 4: wash fraction 2, 5: elution fraction 3, 6: pooled elution fractions, 7: desalted protein, 8: pre-concentrated. Each lane was loaded with 10 µg total protein sample of each collected fraction.

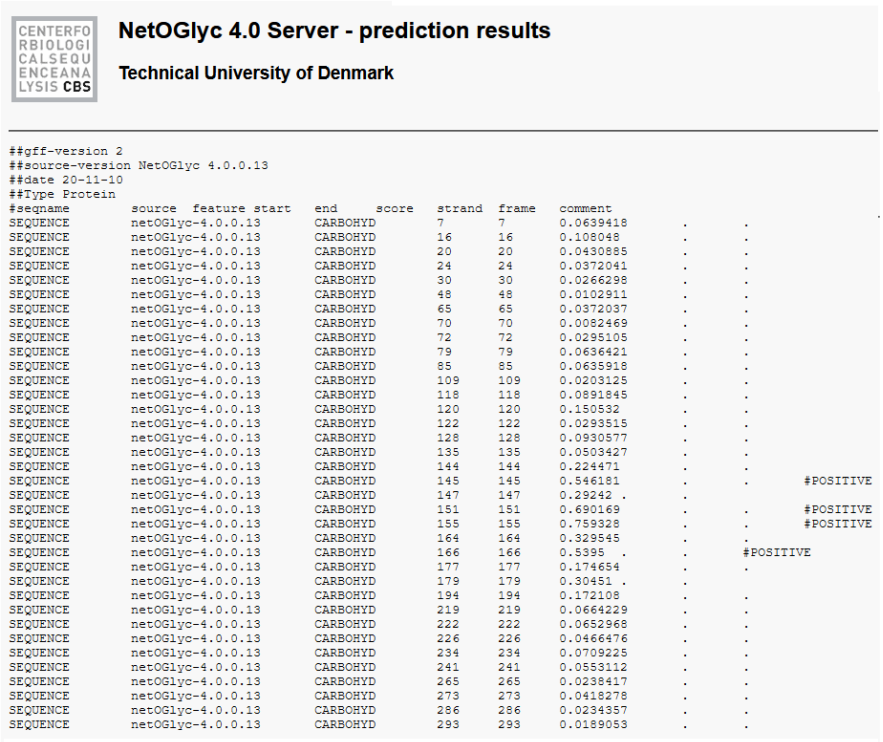


NetOGlyc 4.0 Server - prediction results

Technical University of Denmark

```
##gff-version 2
##source-version NetOGlyc 4.0.0.13
##date 20-11-10
##Type Protein
#seqname        source  feature start   end     score   strand  frame   comment
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        17      17      0.109977        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        21      21      0.0960645       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        25      25      0.0753906       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        31      31      0.0209288       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        55      55      0.108398        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        66      66      0.0389557       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        73      73      0.0269491       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        75      75      0.0339622       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        80      80      0.0425359       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        92      92      0.0107502       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        106     106     0.0730351       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        107     107     0.0442898       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        110     110     0.0455233       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        123     123     0.0187125       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        129     129     0.17843 .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        139     139     0.187939        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        144     144     0.0725038       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        150     150     0.452647        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        154     154     0.634787        .       .       #POSITIVE
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        171     171     0.24941 .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        172     172     0.275949        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        176     176     0.0804864       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        193     193     0.0809038       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        203     203     0.0201198       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        206     206     0.00483987      .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        215     215     0.00735643      .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        225     225     0.0372566       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        233     233     0.0993946       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        244     244     0.0816594       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        249     249     0.0679693       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        263     263     0.0413738       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        268     268     0.0222238       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        276     276     0.0337747       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        278     278     0.0167337       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        287     287     0.026337        .       .
```

Supplemental Figure 29: Predicted *O*-glycosylation pattern of human GLYAT. As predicted by NetOGlyc 4.0 Server online tool serine at position 154 is *O*-glycosylated.



NetOGlyc 4.0 Server - prediction results

Technical University of Denmark

```
##gff-version 2
##source-version NetOGlyc 4.0.0.13
##date 20-11-10
##Type Protein
#seqname        source  feature start   end     score   strand  frame   comment
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        7       7       0.0639418       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        16      16      0.108048        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        20      20      0.0430885       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        24      24      0.0372041       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        30      30      0.0266298       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        48      48      0.0102911       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        65      65      0.0372037       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        70      70      0.0082469       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        72      72      0.0295105       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        79      79      0.0636421       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        85      85      0.0635918       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        109     109     0.0203125       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        118     118     0.0891845       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        120     120     0.150532        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        122     122     0.0293515       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        128     128     0.0930577       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        135     135     0.0503427       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        144     144     0.224471        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        145     145     0.546181        .       .       #POSITIVE
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        147     147     0.29242 .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        151     151     0.690169        .       .       #POSITIVE
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        155     155     0.759328        .       .       #POSITIVE
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        164     164     0.329545        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        166     166     0.5395  .       .       #POSITIVE
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        177     177     0.174654        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        179     179     0.30451 .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        194     194     0.172108        .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        219     219     0.0664229       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        222     222     0.0652968       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        226     226     0.0466476       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        234     234     0.0709225       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        241     241     0.0553112       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        265     265     0.0238417       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        273     273     0.0418278       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        286     286     0.0234357       .       .
SEQUENCE        netOGlyc-4.0.0.13       CARBOHYD        293     293     0.0189053       .       .
```

Supplemental Figure 30: Predicted *O*-glycosylation pattern of human GLYATL1. As predicted by NetOGlyc 4.0 Server online tool serines at position 145 and 147 and threonine at position 166 are *O*-glycosylated**.**

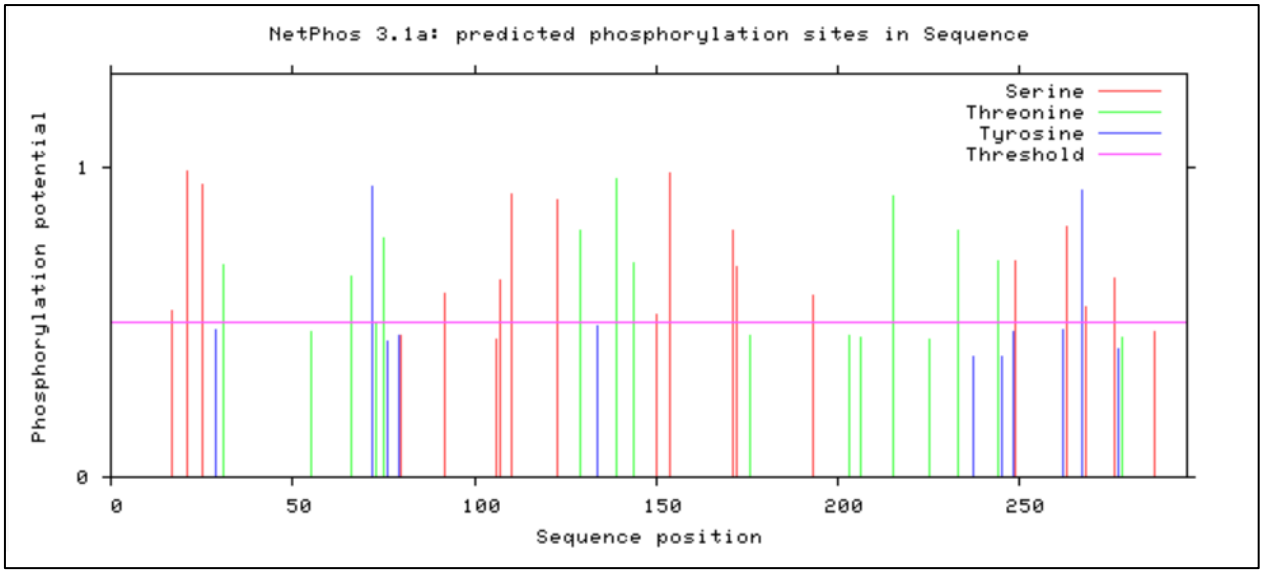
NetOGlyc 4.0 Server - prediction results

Technical University of Denmark

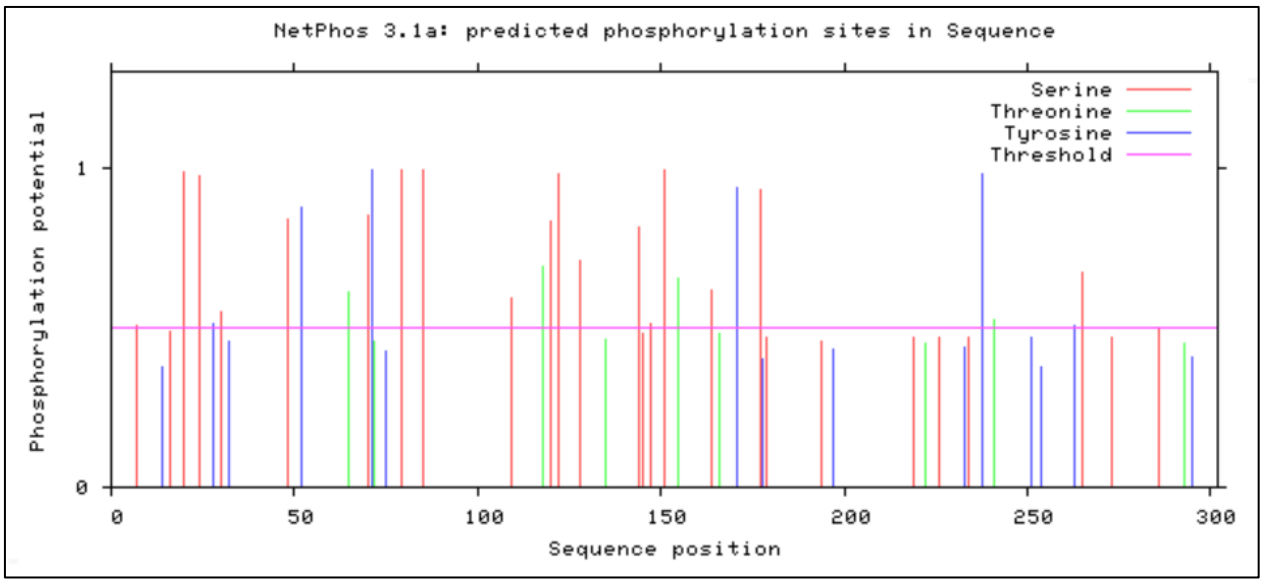
```
##gff-version 2
##source-version NetOGlyc 4.0.0.13
##date 20-11-10
##Type Protein
```

#seqname	source	feature	start	end	score	strand	frame	comment			
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	7	7	0.0899977	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	16	16	0.0840986	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	20	20	0.0237445	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	24	24	0.0279235	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	57	57	0.168232	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	72	72	0.0432401	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	74	74	0.0725414	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	79	79	0.114573	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	90	90	0.0175666	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	92	92	0.0284596	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	96	96	0.0881939	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	100	100	0.0217658	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	119	119	0.125738	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	120	120	0.233696	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	122	122	0.0599976	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	130	130	0.0199076	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	143	143	0.182924	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	144	144	0.713771	.	.	#POSITIVE			
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	145	145	0.735195	.	.	#POSITIVE			
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	164	164	0.0827031	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	171	171	0.198772	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	188	188	0.0394499	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	213	213	0.0232195	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	220	220	0.0128467	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	228	228	0.0740434	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	250	250	0.0245012	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	267	267	0.0197085	.	.				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	289	289	0.0574065	.	.				

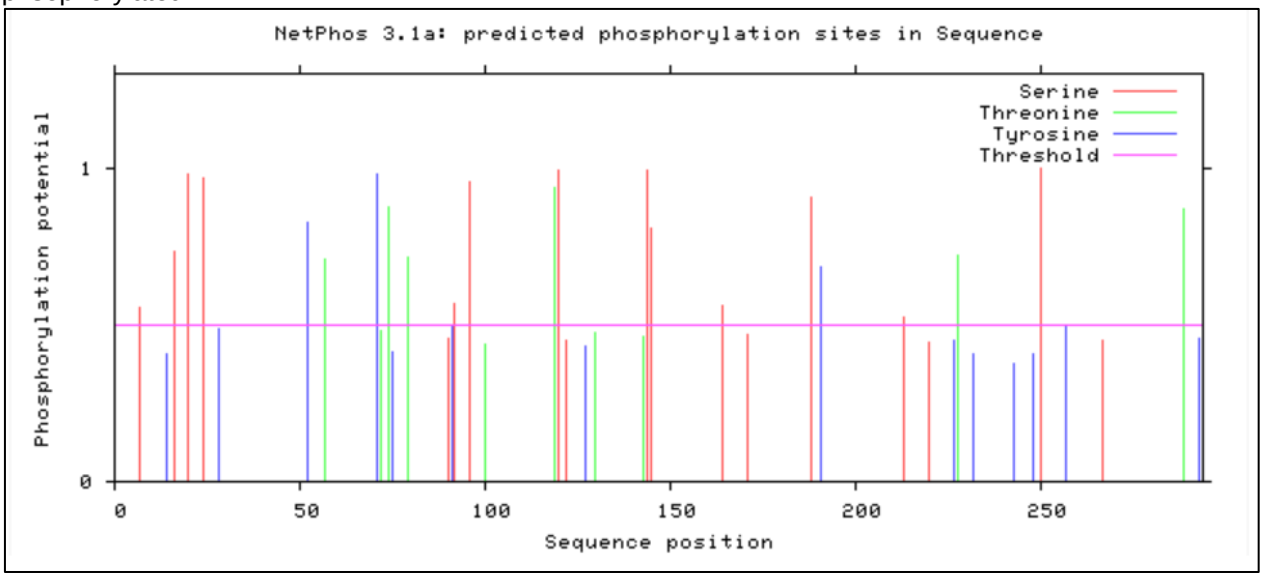
Supplemental Figure 31: Predicted *O*-glycosylation pattern of human GLYATL2. As predicted by NetOGlyc 4.0 Server online tool serines at position 144 and 145 are *O*-glycosylated.



Supplemental Figure 32: Predicted phosphorylation pattern of human GLYAT. As predicted by online tool NetPhos3.1. several serine, threonine and tyrosine residues of human GLYAT are phosphorylated.



Supplemental Figure 33: Predicted phosphorylation pattern of human GLYATL1. As predicted by online tool NetPhos3.1. several serine, threonine and tyrosine residues of human GLYATL1 are phsophorylated.



Supplemental Figure 34: Predicted phosphorylation pattern of human GLYATL2. As predicted by online tool NetPhos3.1. several serine, threonine and tyrosine residues of human GLYATL2 are phsophorylated.

※ PAIL: Prediction of Acetylation on Internal Lysines

Go back to PAIL prediction page

Predicted Acetylation sites:

Peptide	Position	Score	Threshold
MLQMLEKSLRKSL	16	0.36	0.2
LEKSLRKSLPASL	20	1.38	0.2
SLPASLKVYGTVF	27	0.28	0.2
GNPFNLKAVVDKW	44	0.28	0.2
QNLAAIKSFKVKQ	122	1.51	0.2
AAIKSFKVKQTQR	125	1.39	0.2
MAAETAKELTPFL	141	0.46	0.2
LTPFLLKSKILSP	149	0.43	0.2
PFLLKSKILSPNG	151	1.65	0.2
LSPNGGKPKAINQ	159	1.09	0.2
PNGGKPKAINQEM	161	0.78	0.2
INQEMFKLSSMDV	169	0.93	0.2
IYSHAQKLGKLGF	253	0.37	0.2
HAQKLGKLGFPVY	256	1.63	0.2

Supplemental Figure 35: Predicted *Nε*-acetylation pattern of human GLYAT using PAIL [99]. Peptide fragments, position, score and threshold of calculation are shown in the figure.

※ PAIL: Prediction of Acetylation on Internal Lysines

Go back to PAIL prediction page

Predicted Acetylation sites:

Peptide	Position	Score	Threshold
LLNNSHKLLALYK	9	0.24	0.2
KLLALYKSLARSI	15	0.37	0.2
VIIRPQKQEMTDD	61	0.64	0.2
FSKEPQKSEEVLK	84	0.52	0.2
RVATFSKSVKVEH	121	1.40	0.2
TFSKSVKVEHSRA	124	2.15	0.2
KLNASSKSKLGSW	146	1.35	0.2
NASSKSKLGSWAE	148	1.38	0.2
SETPNFKYAQLDV	170	0.29	0.2
LVNDNWKRGKNER	187	0.21	0.2
DNWKRGKNERSLH	190	1.54	0.2
RSLHYIKRCIEDL	199	0.30	0.2
VMVRYMKYLRQKN	253	0.70	0.2

Supplemental Figure 36: Predicted *Nε*-acetylation pattern of human GLYATL1 using PAIL [99]. The peptide fragments, position, score and threshold of calculation are shown in the figure.

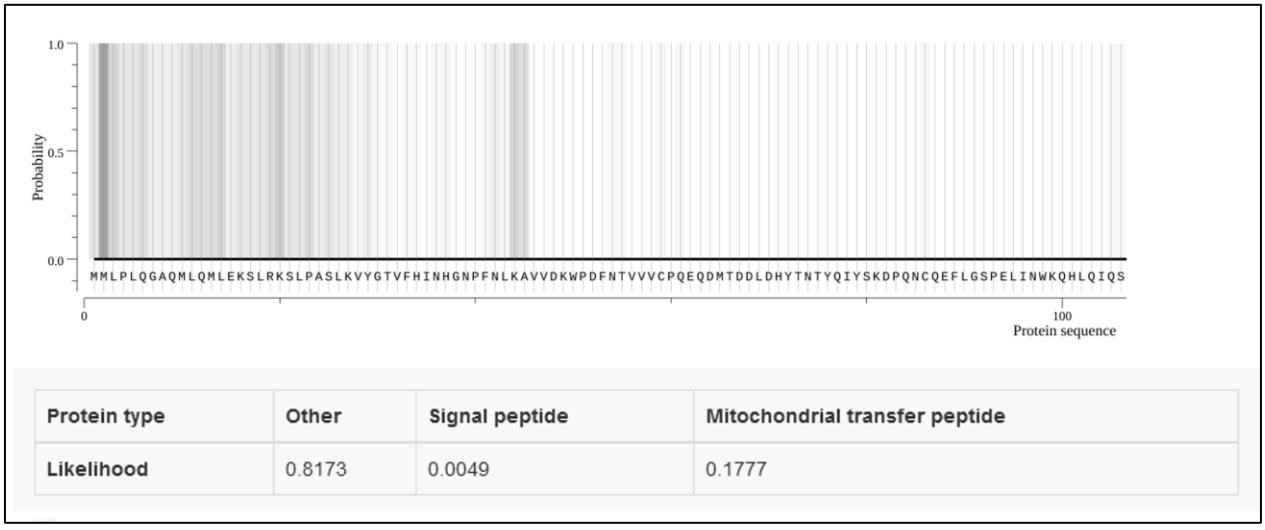
※ PAIL: Prediction of Acetylation on Internal Lysines

Go back to PAIL prediction page

Predicted Acetylation sites:

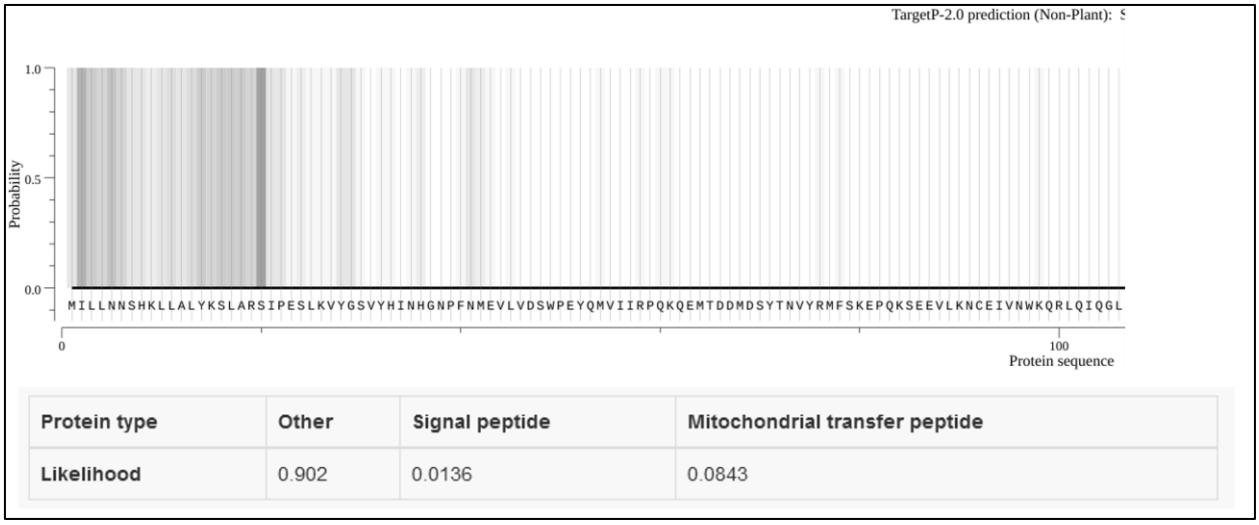
Peptide	Position	Score	Threshold
KLQILYKSLEKSI	15	0.35	0.2
VITRPQKQEMKDD	61	1.77	0.2
TYHIFTKAPDKLE	80	0.42	0.2
FTKAPDKLEEVLS	84	1.82	0.2
LDEAIRKVATSKS	116	0.30	0.2
RKVATSKSVQVDY	121	1.59	0.2
FIPELPKKHKTSS	139	0.76	0.2
IPELPKKHKTSSN	140	0.38	0.2
ELPKKHKTSSNDK	142	2.40	0.2
KTSSNDKMELFEV	148	0.75	0.2
EHWAFGKNERSLK	184	1.46	0.2
KNERSLKYIERCL	190	0.21	0.2
MGYTVPKYRHQGN	231	0.34	0.2
CGWHQWKCTPKKY	287	0.44	0.2
WKCTPKKYC****	292	7.39	0.2

Supplemental Figure 37: Predicted *Nε*-acetylation pattern of human GLYATL2 using PAIL [99]. The peptide fragments, position, score and threshold of calculation are shown in the figure.



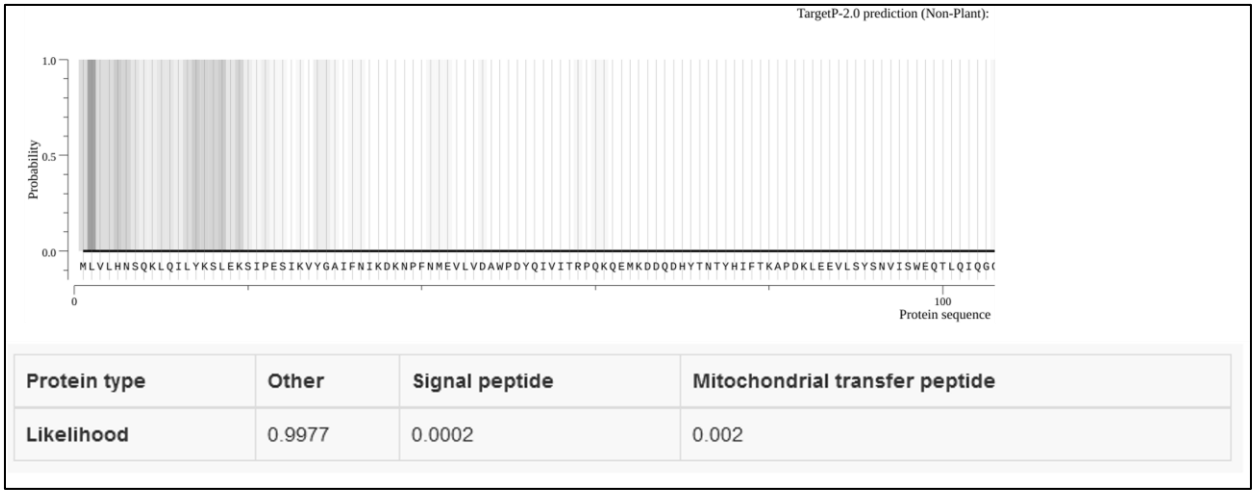
Protein type	Other	Signal peptide	Mitochondrial transfer peptide
Likelihood	0.8173	0.0049	0.1777

Supplemental Figure 38: Mitochondrial *N*-terminal signaling prediction of human GLYAT using the TargetP-2.0 online tool.



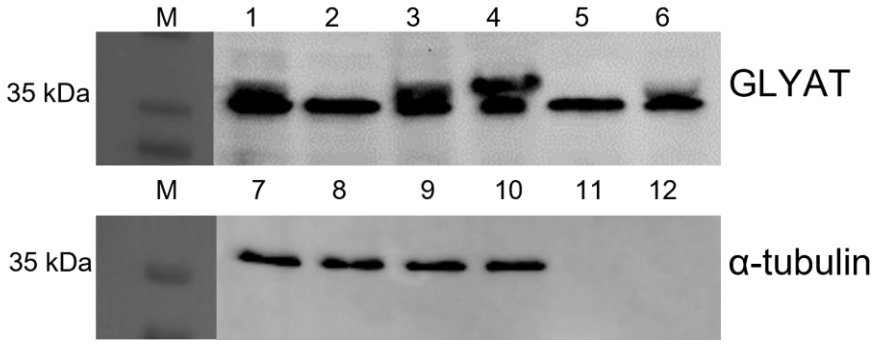
Protein type	Other	Signal peptide	Mitochondrial transfer peptide
Likelihood	0.902	0.0136	0.0843

Supplemental Figure 39: Mitochondrial *N*-terminal signaling prediction of human GLYATL1 using the TargetP-2.0 online tool.

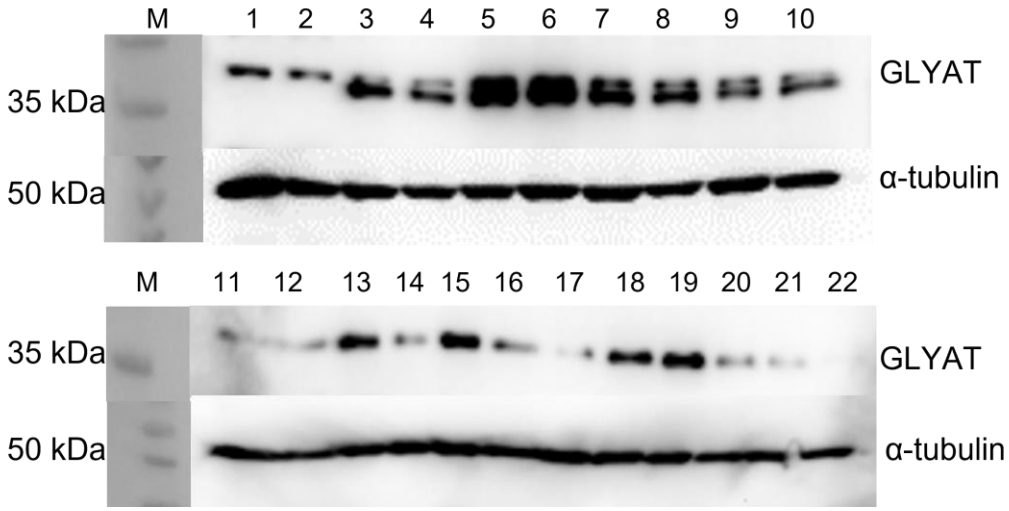


Protein type	Other	Signal peptide	Mitochondrial transfer peptide
Likelihood	0.9977	0.0002	0.002

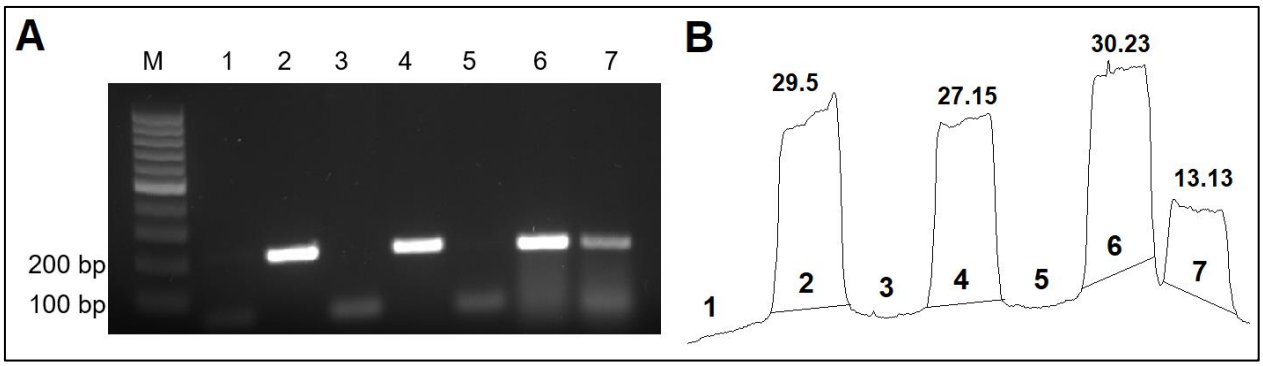
Supplemental Figure 40: Mitochondrial *N*-terminal signaling prediction of human GLYATL2 using the TargetP-2.0 online tool.



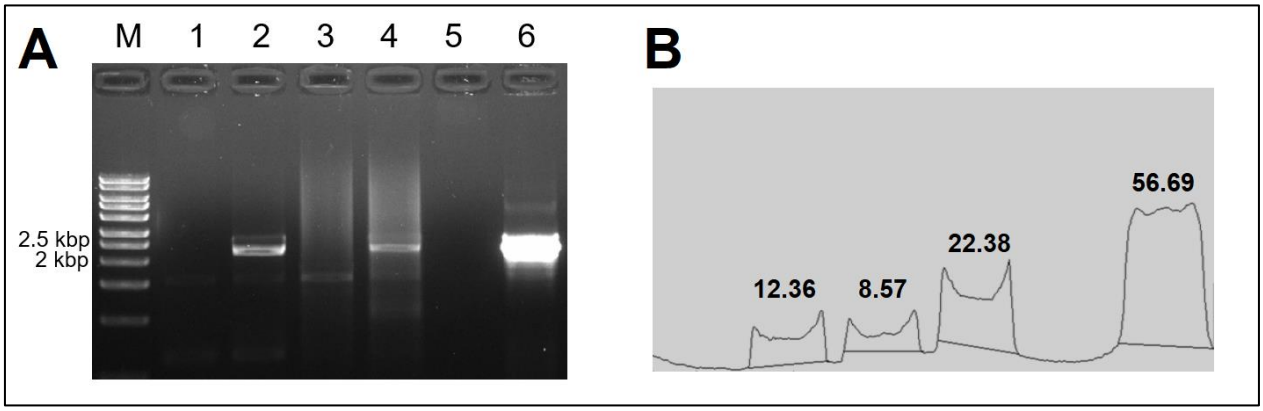
Supplemental Figure 41: Western blots of 5´-azacytidine treated, transiently transfected HeLa cell homogenates (samples 1-6) and HEK293 cell homogenates (samples 7-12). M: Page Ruler Prestained Protein ladder (Thermo Fisher), 1: HeLa + pcDNA3.1(+) – Aza, 2: HeLa + pcDNA3.1(+) + 1 day Aza, 3: HeLa + pcDNA3.1(+) + 2 days Aza, 4: HeLa + GLYAT wild-type – Aza, 5: HeLa + GLYAT wild-type + 1 day Aza, 6: HeLa + GLYAT wild-type + 2 days Aza, 7: HEK293 + pcDNA3.1(+) – Aza, 8: HEK293 + GLYAT wild-type – Aza, 9: HEK293 + pcDNA3.1(+) + 1 day Aza, 10: HEK293 + GLYAT wild-type + 1 day Aza, 11: HEK293 + pcDNA3.1(+) + 2 days Aza, 12: HEK293 + GLYAT wild-type + 2 days Aza. Anti-GLYAT ab86102 was used as first antibody in 3 µg/ml.



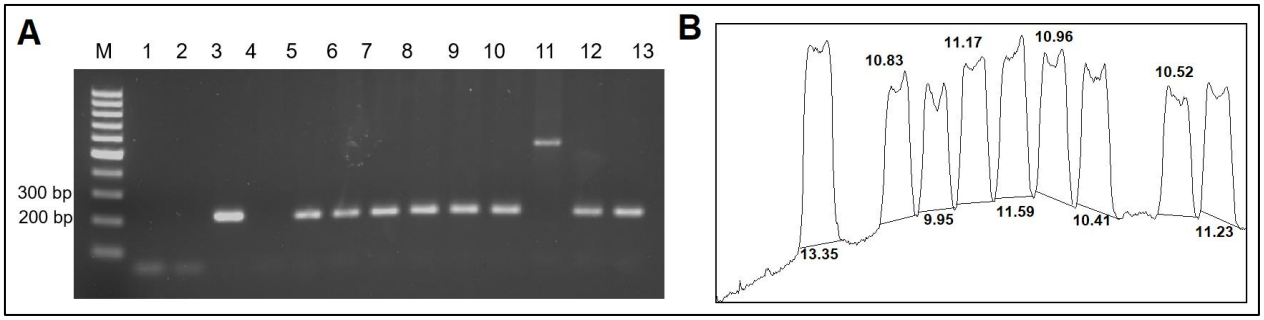
Supplemental Figure 42: **Western Blots of transfected HEK293 cell supernatants with GLYAT sequence variants (samples 1-22) treated with 10 µmol/L MG-132.** 1-22: transfected HEK293 cells; 1: pcDNA3.1(+) DMSO, 2: pcDNA3.1(+) MG-132, 3: GLYAT wild-type DMSO, 4: GLYAT wild-type MG-132, 5: GLYAT p.(Asn156Ser) DMSO, 6: GLYAT p.(Asn156Ser) MG-132, 7: GLYAT p.(Gln61Leu) DMSO, 8: GLYAT p.(Gln61Leu) MG-132, 9: GLYAT p.(His101Tyr) DMSO, 10: GLYAT p.(His101Tyr) MG-132, 11: GLYAT p.(Gly38*) DMSO, 12: GLYAT p.(Val57Phe) DMSO, 13: GLYAT p.(Gln61Leu) DMSO, 14: GLYAT p.(Met65Thr) DMSO, 15: GLYAT p.(His101Tyr) DMSO, 16: GLYAT p.(Ala231Thr) DMSO, 17: GLYAT p.(Gly38*) MG-132, 18: GLYAT p.(Val57Phe) MG-132, 19: GLYAT p.(Gln61Leu) MG-132, 20: GLYAT p.(Met65Thr) MG-132, 21: GLYAT p.(His101Tyr) MG-132, 22: GLYAT p.(Ala231Thr) MG-132. GLYAT antibody PA5-48504, 1:1,000 and anti-tubulin ab7291, 1:5,000 were used as first antibodies.



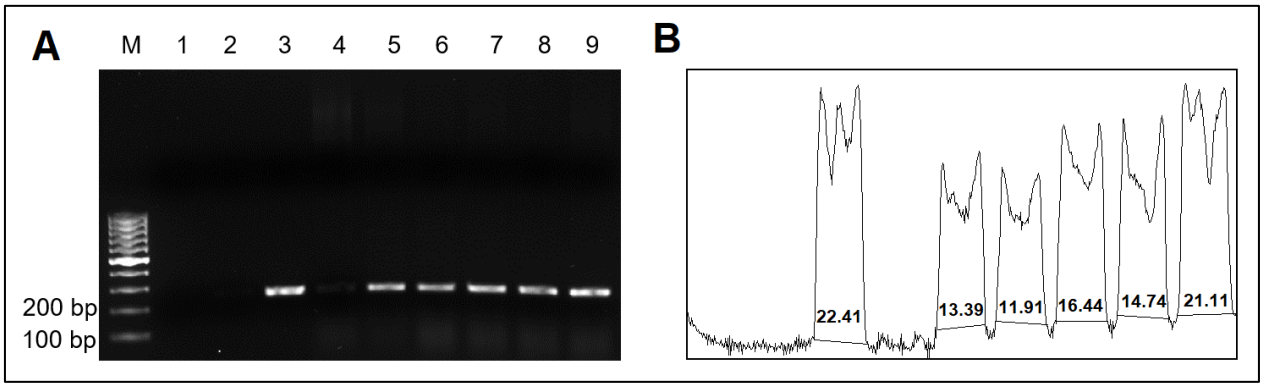
Supplemental Figure 43: 1.5 % (w/v) agarose gel electrophoresis of semi-quantitatively amplified recombinant GLYAT amplicons (gDNA) in stably transfected HEK293(T) cells (A) and ImageJ peak areas (%) (B). A: M: GeneRuler 100 bp DNA ladder, 1: pcDNA3.1(+) 1:1,000 (negative control), 2: GLYAT-pcDNA3.1(+), 1a:1,000 (positive control), 3: gDNA of HEK293 + pcDNA3.1(+), 4: gDNA of HEK293 + GLYAT-pcDNA3.1(+), 5: gDNA of HEK293T + pcDNA3.1(+), 6: gDNA of HEK293T + GLYAT_pcDNA3.1(+), 7: gDNA of HEK293T + GLYAT-eGFP_pcDNA3.1(+). Semi-quantitative PCR was performed according to standard protocol using following primers: GLYAT_ex3_for, GLYAT_ex4_rev. Resulting amplicon has a size of 210 bp. B: ImageJanalysis of observed DNA signals with calculated peak areas in %.



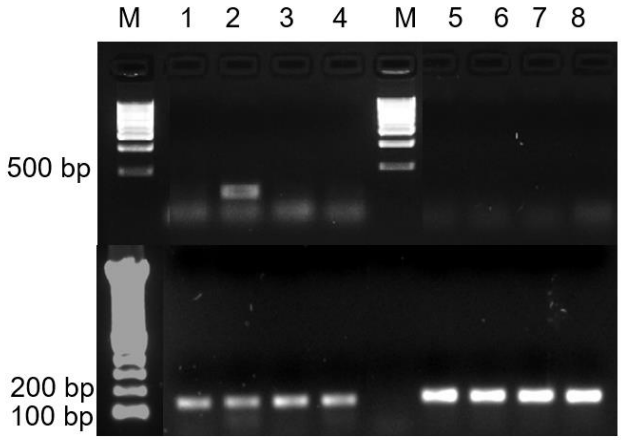
Supplemental Figure 44: 1.5 % (w/v) agarose gel with semi-quantitatively amplified gDNA samples of GLYAT-eGFP-transfected HeLa and HEK293 cells (A) and ImageJ peak areas (%) (B). M: GeneRuler 1 kbp DNA ladder, 1: cDNA HeLa + eGFP-pcDNA3.1(+), 2: cDNA HeLa + GLYAT-eGFP, 3: cDNA HEK293 + eGFP-pcDNA3.1(+), 4: cDNA HEK293 + GLYAT-eGFP, 5: eGFP-pcDNA3.1(+) plasmid (negative control), 6: GLYAT-eGFP plasmid (positive control). Isolated cDNA of transfected HeLa and HEK293 cells were amplified with semi-quantitive PCR using GLYAT-3´-end forward and neomycine-5´-end reverse primers resulting in amplicon of 2.1 kbp



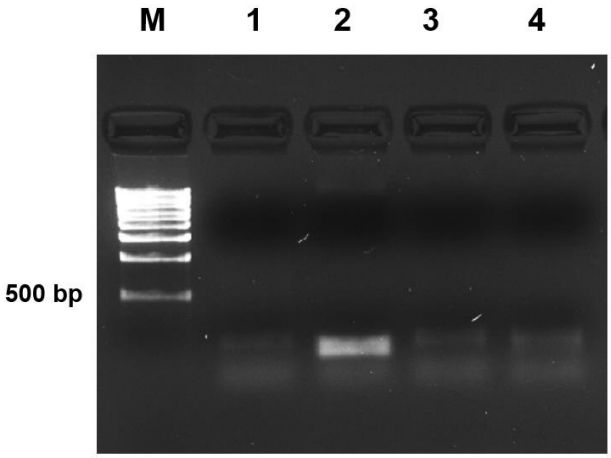
Supplemental Figure 45: A: 1.5 % (w/v) agarose gel and ImageJ peak areas (%) (B) of amplified gDNA with recombinant GLYAT semi-quantitative PCR of stable GLYAT transfected HEK293 cells. M: GeneRuler 100 bp DNA ladder, 1: deionized water control, 2: pcDNA3.1(+) 1:1,000, 3: GLYAT-pcDNA3.1(+) 1:1,000, 4: cDNA of HEK293 + pcDNA3.1(+), 5: cDNA of HEK293 + GLYAT wild-type, 6: cDNA of HEK293 + GLYAT c.467A>G, 7: cDNA of HEK293 + GLYAT c.595C>T, 8: cDNA of HEK293 + GLYAT c.112G>T, 9: cDNA of HEK293 + GLYAT c.169G>T, 10: cDNA of HEK293 + GLYAT c.182A>T, 11: cDNA of HEK293 + GLYAT c.194T>C, 12: cDNA of HEK293 + GLYAT c.301C>T, 13: cDNA of HEK293 + GLYAT c.691G>A. Primers used are GLYAT-exon3_forward and GLYAT-exon4_reverse resulting in 210 bp sized amplicon.



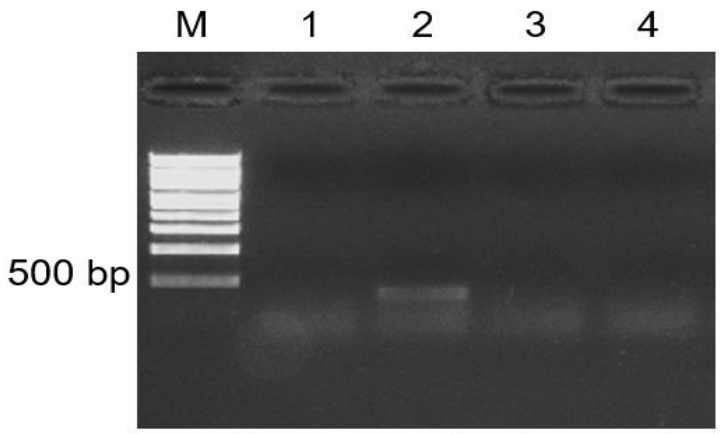
Supplemental Figure 46: 1.5 % (w/v) agarose gel of recombinant GLYATL1 gDNA amplified via semi-quantitative PCR and corresponding ImageJ peak percentages (B). The gel shows deionized water control (1), pcDNA3.1(+) 1:1,000 (2), GLYATL1_pcDNA3.1(+), 1:1000 (3) and gDNAs of stably transfected HEK293 cells: 4: HEK293 + pcDNA3.1(+), 5: HEK293 + GLYATL1 wild-type, 6: HEK293 + GLYATL1 c.187G>A, 7: HEK293 + GLYATL1 c.259G>T, 8: HEK293 + GLYATL1 c.373G>C, 9: HEK293 + GLYATL1 c.670G>A. gDNA PCR was performed according to standard protocol and annealing temperature was 60 °C. GLYATL1 amplicon with used primers GLYATL1_ex4_fw and GLYATL1_ex5_rev resulted in amplicon of 244 bp.



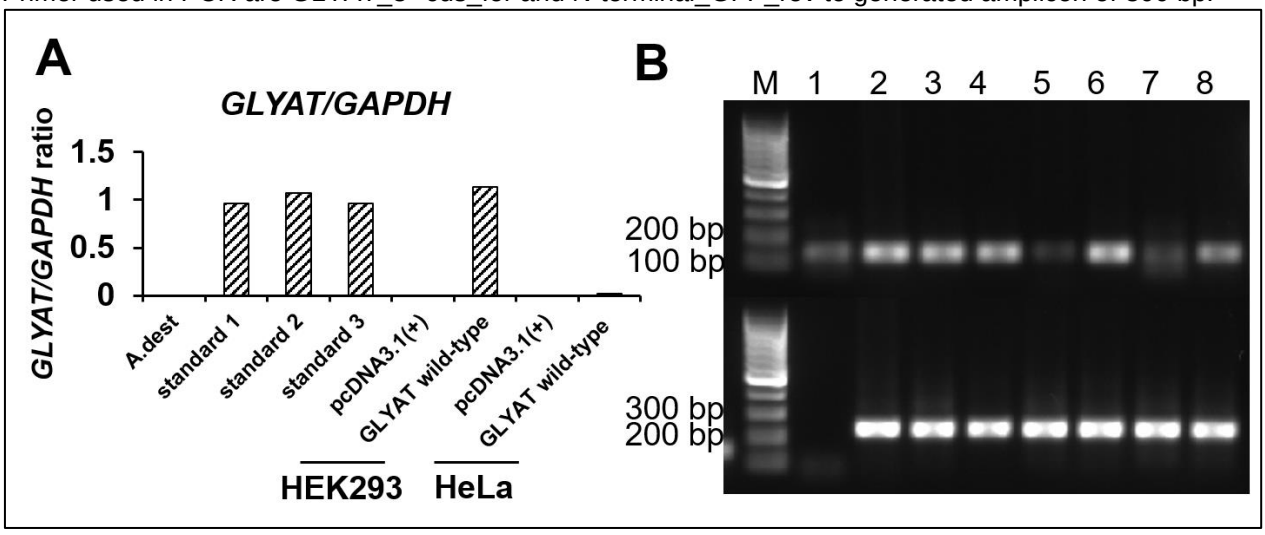
Supplemental Figure 47: 1.5 % (w/v) agarose gel of recombinant GLYAT and GLYAT-eGFP amplicons (upper part, 112 bp amplicon) and β-actin amplicons (lower part, 120 bp amplicon) amplified via semi-quantitative PCR using GLYAT_ex3_for, GLYAT_ex4_rev for recombinant GLYAT detection and β -actin qPCR primers; for GLYAT-eGFP detection GLYAT-3´-cds_for and N-terminal_GFP_rev primers were used. β-actin served as control for integrity of generated cDNA. M: GeneRuler 100 bp DNA ladder, 1: cDNA of HEK293 + pcDNA3.1(+), 2: cDNA of HEK293 + GLYAT-pcDNA3.1(+), 3: cDNA of HeLa + pcDNA3.1(+), 4: cDNA of HeLa + GLYAT_pcDNA3.1(+), 5: cDNA of HEK293 + eGFP_pcDNA3.1(+), 6: cDNA of HEK293 + GLYAT-eGFP, 7: cDNA of HeLa + eGFP_pcDNA3.1(+), 8: cDNA of HeLa + GLYAT-eGFP.



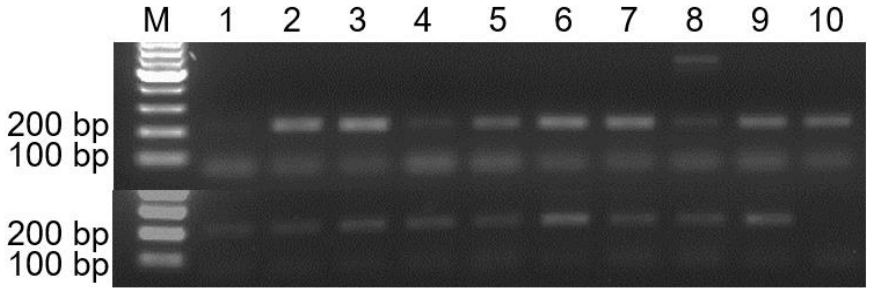
Supplemental Figure 48: 1.5 % (w/v) agarose gel of cDNA of transfected HEK293 and HeLa cells analyzed via semi-quantitative PCR. M: GeneRuler 100 bp DNA ladder, 1: cDNA of HEK293 + pcDNA3.1(+), 2: cDNA of HEK293 + GLYAT wild-type, 3: cDNA of HeLa + pcDNA3.1(+), 4: cDNA of HeLa + GLYAT wild-type. Primer used in PCR are CMV_MCS_for and GLYAT_exon1_rev to generated amplicon of 130 bp.



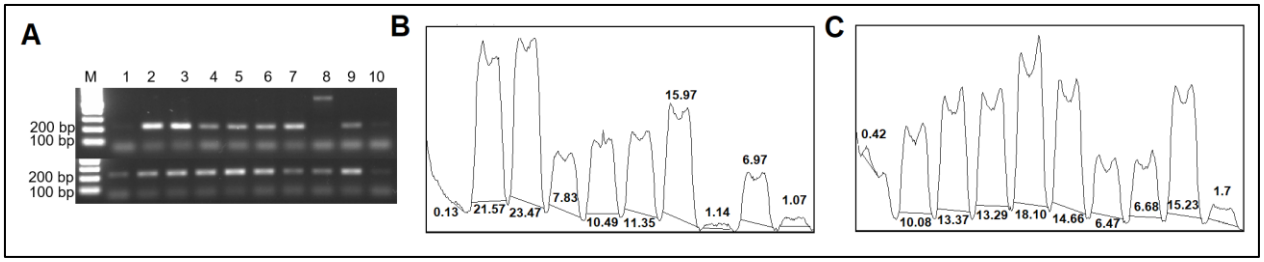
Supplemental Figure 49: 1.5 % (w/v) agarose gel of cDNA of transfected HEK293 and HeLa cells analyzed via semi-quantitative PCR. 1: cDNA of HEK293 + eGFP-pcDNA3.1(+), 2: cDNA of HEK293 + GLYATeGFP_pcDNA3.1(+), 3: cDNA of HeLa + eGFP_pcDNA3.1(+), 4: cDNA of HeLa + GLYAT-eGFP. Primer used in PCR are GLYAT_3´-cds_for and N-terminal_GFP_rev to generated amplicon of 300 bp.



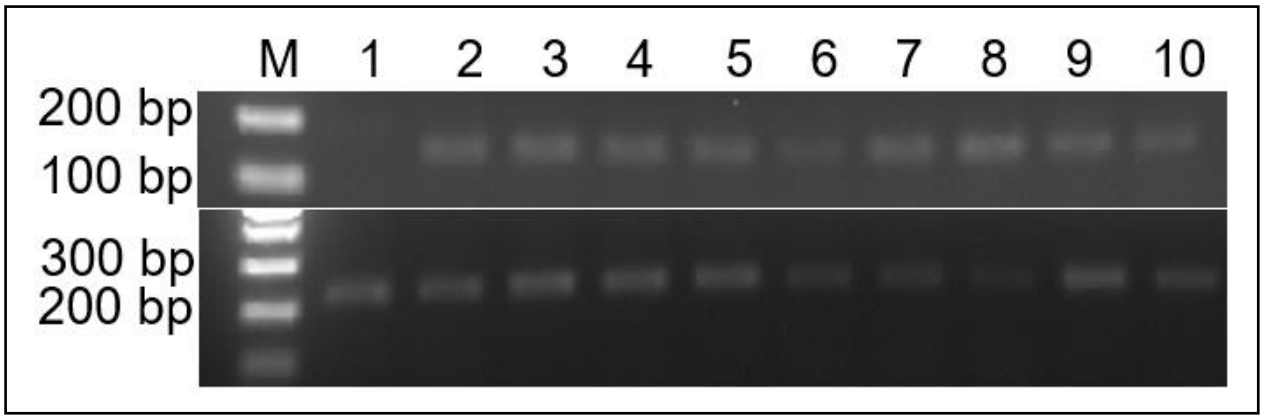
Supplemental Figure 50: GLYAT/GAPDH Ct value ratio (A) and 1.5 % (w/v) agarose gel (B) of cDNA amplified via RT-qPCR from stably transfected HEK293 and HeLa cells. M: GeneRuler 100 bp ladder, 1: non template control (Deioinized water), 2: standard 1 (cDNA HEK293 + GLYAT wild-type), 3: standard 2: 1:10 of standard 1, 4: standard 3: 1:10 of standard 2, 5: cDNA HEK293 + pcDNA3.1(+), 6: cDNA HEK293 + GLYAT wild-type, 7: cDNA HeLa + pcDNA3.1(+), 8: cDNA HeLa + GLYAT wild-type. GLYAT amplicon (130 bp) was amplified using CMV_MCS_fw and GLYAT_exon1_rev primers, while GAPDH (226 bp) was generated using GAPDH_RTqPCR primers.



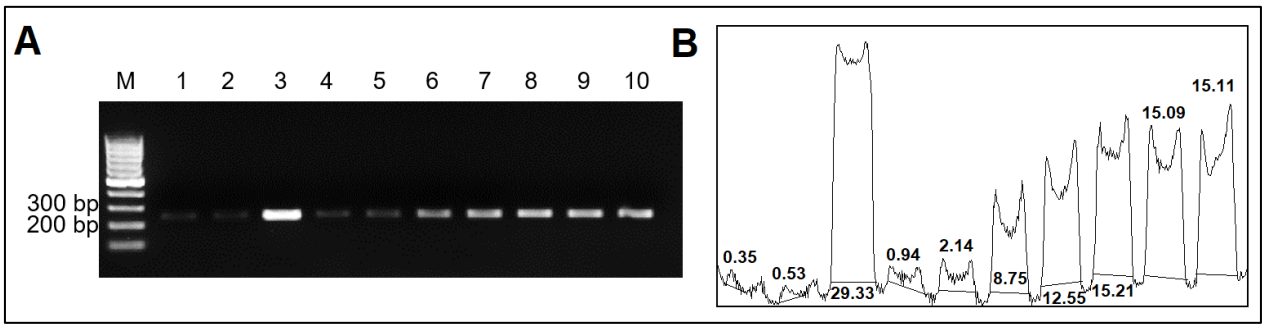
Supplemental Figure 51: 1.5 % (w/v) agarose gel with semi-quantitative PCR-amplified cDNAs of transfected HEK293 cells (upper part: GLYAT amplicon detection of 210 bp amplicon using GLYAT_ex3_fw and ex4_rev primers, lower part: GAPDH amplicon detection of 226 bp amplicon using GLYAT_ex3_fw and ex4_rev primers, lower part: GAPDH amplicon detection of 226 bp amplicon using GAPDH RTqPCR fw and rev primers). M: 100 bp DNA ladder (Thermo Fisher), 1: cDNA HEK293 + pcDNA3.1(+), 2: cDNA HEK293 + GLYAT wild-type, 3: cDNA HEK293 + GLYAT c.467A>G, 4: cDNA HEK293 + GLYAT c.595C>T, 5: cDNA of HEK293 + GLYAT c.112G>T, 6: cDNA of HEK293 + GLYAT c.169G>T, 7: cDNA of HEK293 + GLYAT c.182A>T, 8: cDNA of HEK293 + GLYAT c.194T>C, 9: cDNA of HEK293 + GLYAT c.301C>T, 10: cDNA of HEK293 + GLYAT c.691G>A.



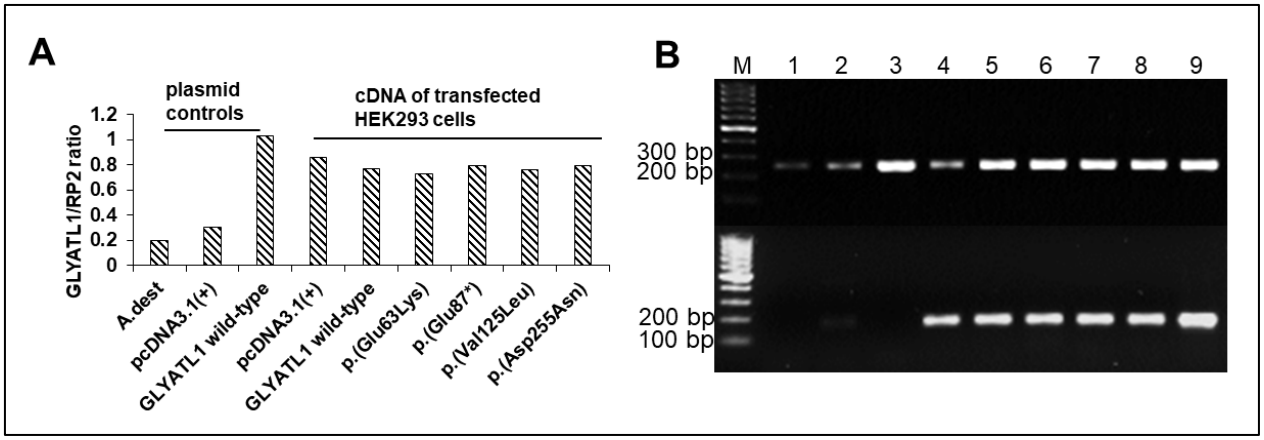
Supplemental Figure 52: 1.5 % (w/v) agarose gel with PCR-amplified cDNAs (now 60 °C annealing) of transfected HEK293 cells and corresponding ImageJ peak percentages of GLYAT (B) and GAPDH (C). (upper part of the gel: GLYAT amplicon detection of 210 bp using GLYAT_ex3_fw and ex4_rev primers, lower part: GAPDH amplicon detection of 226 bp using GAPDH RTqPCR fw and rev primers). M: 100 bp DNA ladder (Thermo Fisher), 1: cDNA HEK293 + pcDNA3.1(+), 2: cDNA HEK293 + GLYAT wild-type, 3: cDNA HEK293 + GLYAT c.467A>G, 4: cDNA HEK293 + GLYAT c.595C>T, 5: cDNA of HEK293 + GLYAT c.112G>T, 6: cDNA of HEK293 + GLYAT c.169G>T, 7: cDNA of HEK293 + GLYAT c.182A>T, 8: cDNA of HEK293 + GLYAT c.194T>C, 9: cDNA of HEK293 + GLYAT c.301C>T, 10: cDNA of HEK293 + GLYAT c.691G>A.



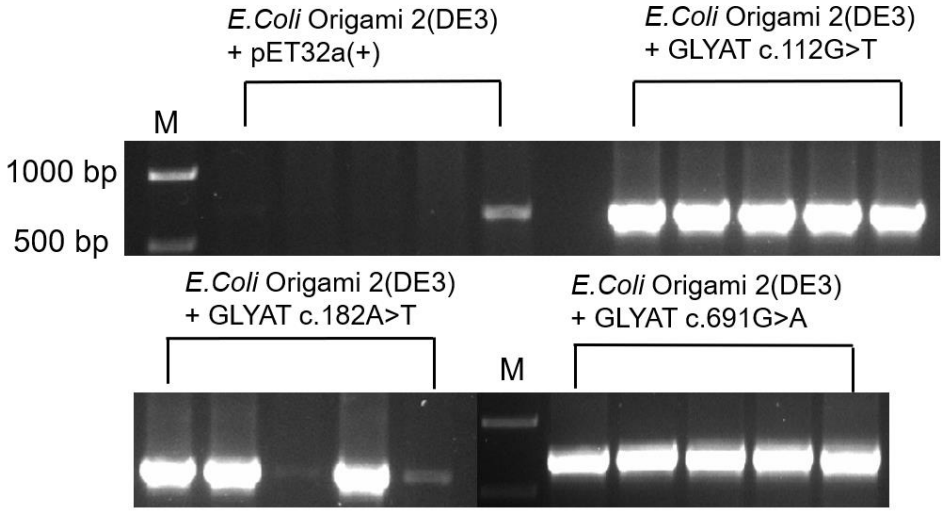
Supplemental Figure 53: 1.5 % (w/v) agarose gel with PCR-amplified cDNAs of transfected HEK293 cells (upper part: GLYAT amplicon detection of 130 bp using GLYAT_CMV_MCS and ex1_rev primers, lower part: GAPDH amplicon detection of 226 bp using GAPDH RTqPCR fw and rev primers). M: 100 bp DNA ladder (Thermo Fisher), 1: cDNA HEK293 + pcDNA3.1(+), 2: cDNA HEK293 + GLYAT wild-type, 3: cDNA HEK293 + GLYAT c.467A>G, 4: cDNA HEK293 + GLYAT c.595C>T, 5: cDNA of HEK293 + GLYAT c.112G>T, 6: cDNA of HEK293 + GLYAT c.169G>T, 7: cDNA of HEK293 + GLYAT c.182A>T, 8: cDNA of HEK293 + GLYAT c.194T>C, 9: cDNA of HEK293 + GLYAT c.301C>T, 10: cDNA of HEK293 + GLYATc.691G>A.



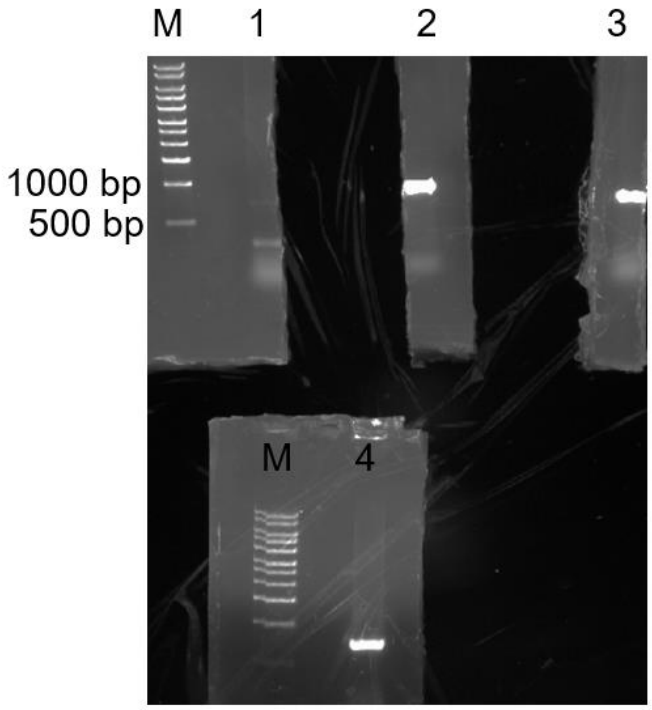
Supplemental Figure 54: 1.5 % (w/v) agarose gel of PCR and corresponding ImageJ peak percentages (B) of amplified cDNA of stably transfected HEK293 cells. M: GeneRuler 100 bp ladder, 1: deionized water control, 2: pcDNA3.1(+) 1:1,000, 3: GLYATL1_pcDNA3.1(+) 1:1,000, 4: GLYATL2_pcDNA3.1(+) 1:1,000, 5: cDNA of HEK293 + pcDNA3.1(+), 6: cDNA of HEK293 + GLYATL1 wild-type, 7: cDNA of HEK293 + GLYATL1 c.187G>A, 8: cDNA of HEK293 + GLYATL1 c.259G>T, 9: cDNA of HEK293 + GLYATL1 c.373G>C, 10: cDNA of HEK293 + GLYATL1 c.670G>A. Resulting GLYATL1 amplicon has a size of 244 bp.



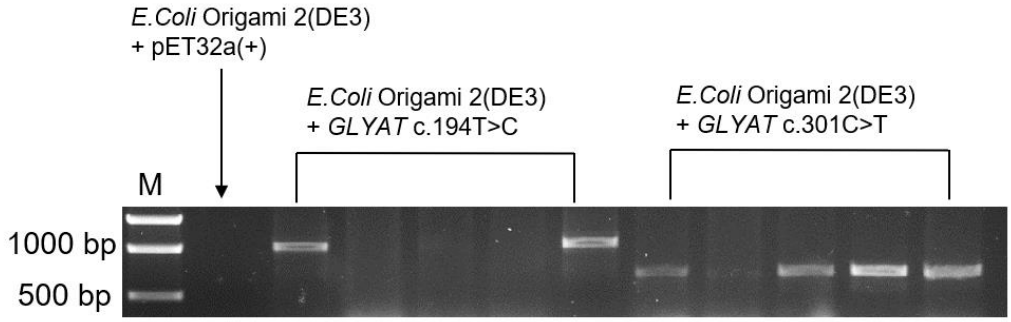
Supplemental Figure 55: GLYATL1/RP2 Ct value ratio (A), 1.5 % (w/v) agarose gel (B) and ImageJ peak percentages of GLYATL1 (C) and RP2 signals (D) of RT-qPCR results of cDNAs from stably transfected HEK293 cells. M: GeneRuler 100 bp ladder, 1: deionized water control, 2: pcDNA3.1(+) 1:1,000, 2: GLYATL1_pcDNA3.1(+) 1:1,000, 3: cDNA HEK293 + pcDNA3.1(+), 5: cDNA HEK293 + GLYATL1 wild-type, 6: cDNA HEK293 + GLYATL1 c.187G>A, 7: cDNA HEK293 + GLYATL1 c.259G>T, 8: cDNA HEK293 + GLYATL1 c.373G>C, 9: cDNA HEK293 + GLYATL1 c.670G>A. cDNAs were diluted 1:4 for qPCR preparation. GLYATL1 amplicon was produced with GLYATL1_ex4_for and GLYATL1_ex5_rev primer and has a size of 244 bp, whereas RP2 amplicon shows a size of 172 bp produced with RP2_for and RP2_rev primer. Gel was incubated for 25 min at 120 V.



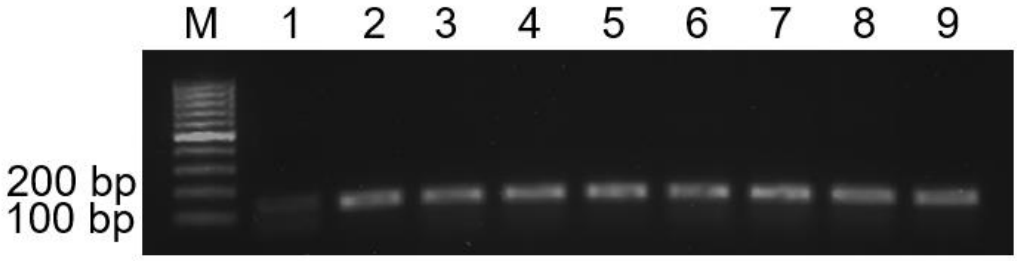
Supplemental Figure 56: 1.5% (w/v) agarose gel of colony PCR product of 5 clones of *E. coli* Origami 2(DE3) + pET32a(+), GLYAT c.112G>T, GLYAT c.182A>T and GLYAT c.691G>A respectively. M: GeneRuler 100 bp DNA ladder. Colony PCR was performed according to standard protocol with GLYAT_pET32a_for and GLYAT_pET32a_rev primers resulting in 720 bp sized amplicon.



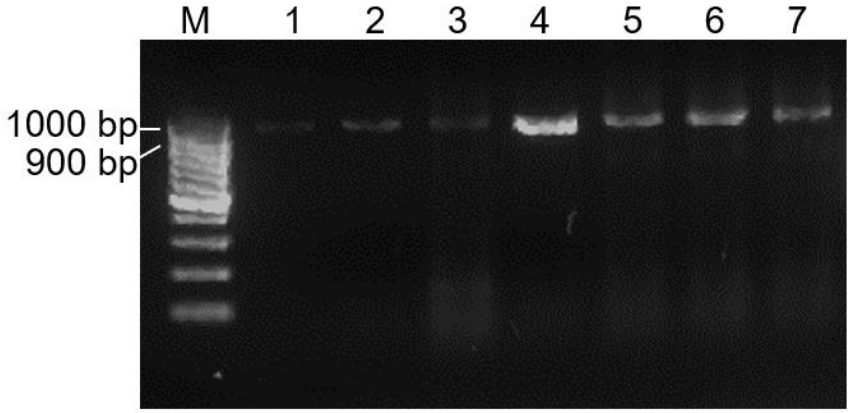
Supplemental Figure 57: Preparative agarose gel (1.5% (w/v)) of colony PCR with subsequent cutting of non-UV treated amplicons from the gel and sending to LGC for sequencing after gel elution (Section 3.4.7.). M: GeneRuler 100 bp DNA ladder. The resulting amplicon was 720 bp in size. Sequencing results were as expected.



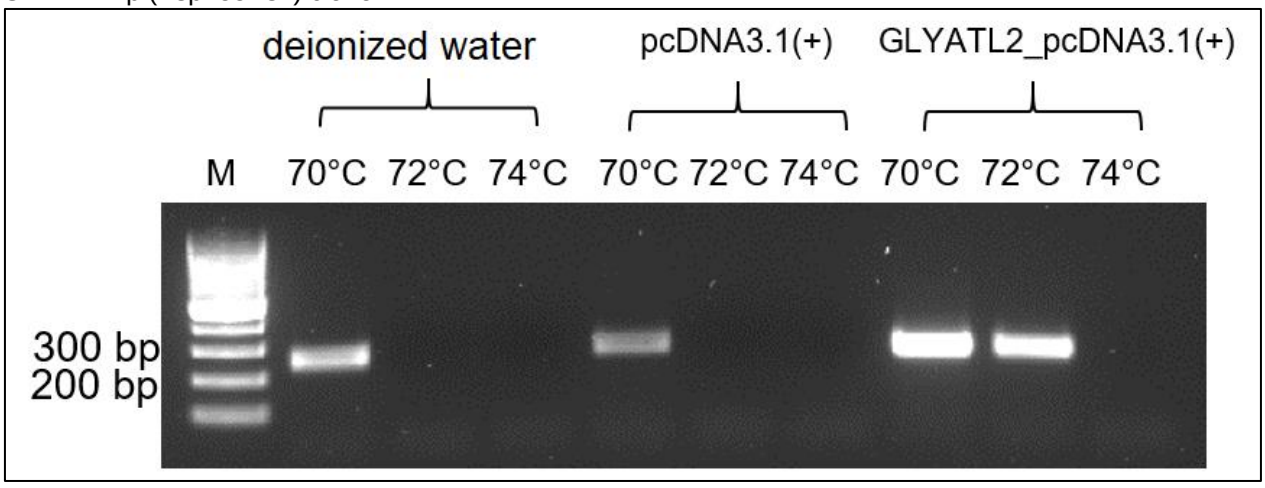
Supplemental Figure 58: 1.5% (w/v) agarose gel of colony PCR product of 5 clones of *E. coli* Origami 2(DE3) + pET32a(+) and GLYAT c.194T>C and c.301C>T. M: GeneRuler 100 bp DNA ladder. Colony PCR was performed according to standard protocol with GLYAT_pET32a_for and GLYAT_pET32a_rev primers resulting in 720 bp sized amplicon.



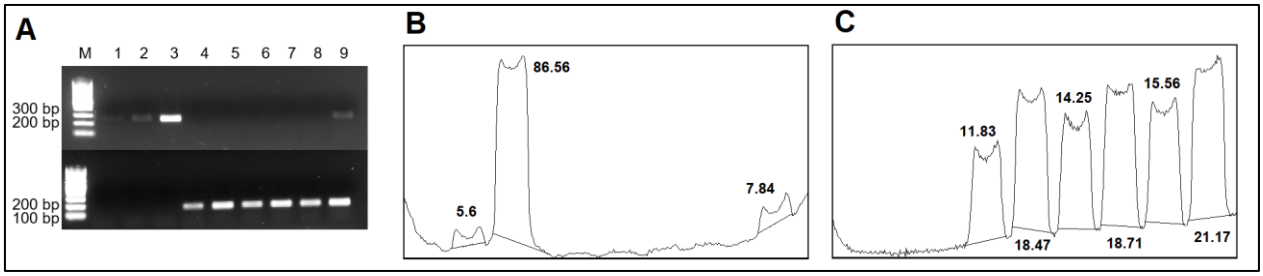
Supplemental Figure 59: Agarose gel electrophoresis (1.5% (w/v) gel) of GLYATL1 colony PCR performed with GLYATL1_bac_for and GLYATL1_bac_rev primers resulting in 137 bp amplicon. M: GeneRuler 100 bp DNA ladder, 1: deionized water control, 2: pET32a(+) clone 1, 3: pET32a(+): clone 2, 4: GLYATL1 wild-type clone 1, 5: GLYATL1 wild-type clone 2, 6: GLYATL1 wild-type clone 3, 7: GLYATL1 c.259G>T clone 1, 8: GLYATL1 c.259G>T clone 2, 9: GLYATL1 c.259G>T clone 3.



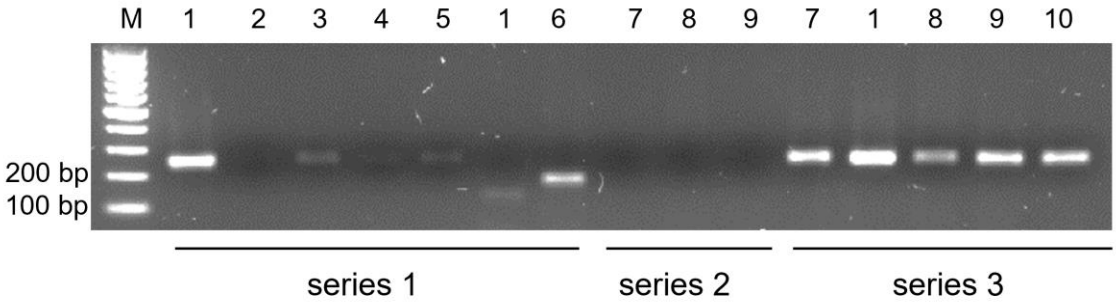
Supplemental Figure 60: Agarose gel electrophoresis (1.5% (w/v) gel) of GLYATL1 colony PCR performed with GLYATL1_21_for and GLYATL1_893_rev primers resulting in 873 bp amplicon. M: GeneRuler 100 bp DNA ladder, 1: deionized water control, 2: pET32a(+) clone 1, 3: GLYATL1 wild-type clone, 4: GLYATL1 p.(Glu63Lys) clone, 5: GLYATL1 p.(Gly87*) clone, 6: GLYATL1 p.(Val125Leu) clone, 7: GLYATL1 p.(Asp255Asn) clone.



Supplemental Figure 61: 1.5 % (w/v) agarose gel of GLYATL2 tissue expression RT-qPCR controls for annealing temperature check from 70 to 74 °C. M: GeneRuler 100 bp DNA ladder. RT-qPCR with annealing temperatures of 70 °C, 72 °C and 74 °C was performed using GLYATL2 qPCR primers (302_for, 552_rev) resulting in 246 bp amplicon



Supplemental Figure 62: A: 1.5 % (w/v) agarose gel and ImageJ peak percentages (B, C) of GLYATL2 tissue expression RT-qPCR (71 °C annealing). M: GeneRuler 100 bp DNA ladder, 1: deionized water control, 2: pcDNA3.1(+) 1:1000, 3: GLYATL2-pcDNA3.1(+) 1:1000, 4: cDNA liver, 5: cDNA testis, 6: cDNA skeletal muscle, 7: cDNA kidney, 8: cDNA heart, 9: cDNA brain. GLYATL2 amplicon produced with primers GLYATL2_302_for and GLYATL2_552_rev has a size of 246 bp, whereas in the lower part of the picture human RNA Polymerase II was used as control with an amplicon size of 172 bp



Supplemental Figure 63: 1.5 % (w/v) agarose gel of GLYATL2 tissue expression samples; RT-qPCR was performed with GLYATL2 primers and RP2 primers resulting in 246 bp and 172 bp amplicon respectively. Series 1: cDNA was undiluted, 1:2 and 1:4 diluted; series 2: fresh aliquot of Blue´s sybr green mix, annealing temperature: 71 °C, series 3: annealing temperature: 70 °C. M: GeneRuler 100 bp DNA ladder, 1: GLYATL2_pcDNA3.1(+), 1:1,000, 2: cDNA liver 1:4, 3: cDNA heart 1:4, 4: cDNA salivary gland 1:4, 5: cDNA F-a1 1:4, 6: cDNA skeletal muscle 1:4, 7: deioinized water, 8: cDNA skeletal muscle, 9: cDNA salivary gland, 10: cDNA F-a1.

Poster presentations

Schulke D, Sass JO. Frequent sequence variants of human glycine *N*-acyltransferase and inborn errors of metabolism. Poster Abstract P19 of the 34th Annual Conference of the Arbeitsgemeinschaft für Pädiatrische Stoffwechselstörungen (APS). Monatsschr Kinderheilkd 2022;170(Suppl1) S.14, doi: 10.1007/s0012-021-01291-w

Schulke D, Sass JO. Human glycine *N*-acyltransferase (GLYAT) and inborn errors of metabolism. Digital Poster at the 14th International Congress of Inborn Errors of Metabolism (ICIEM), Sydney, Australia (online); https://www.iciem2021.com.au/cms/wp-content/uploads/2021/11/ICIEM-2021-Congress-Poster-List.pdf

Schulke D, Sass JO. Intracellular Localization of Glycine *N*-Acyltransferase-like protein 1 (GLYATL1), a phase 2 enzyme relevant in urea cycle disorders. Poster Abstract of the SSIEM Annual Symposium, J Inher Metab Dis 2022, eingereicht.

Publications

Schulke D, Sass JO. Frequent sequence variants of human glycine N-acyltransferase (GLYAT) and inborn errors of metabolism. Biochimie 2021;183:30-4, doi: 10.1016/j.biochi.2021.02.002 (relevant for the dissertation)

Schulke D, Hollenbeck JC, Klaas L, Sass JO. Funktionale Gen-Analytik. Angeborenen Stoffwechselstörungen auf der Spur. GIT-Laborfachzeitschrift 2022;4:31-3, doi: 10.1002/was.000600301
(relevant for the dissertation)

10. Eidesstattliche Erklärung

Name	Schulke
Vorname	Daniel
Titel der Dissertation	Amino acid *N*-acyltransferases in Human Metabolism

Hiermit erkläre ich durch meine Unterschrift an Eides statt:

1. Die von mir eingereichte Dissertation habe ich selbstständig und ohne unzulässige fremde Hilfe verfasst. Hierbei habe ich weder Textstellen von Dritten oder aus eigenen Prüfungsarbeiten noch Grafiken oder sonstige Materialien ohne Kennzeichnung übernommen. Auszüge der Arbeit wurden von der Wort für Wort GmbH und Co. KG (Maternusstr. 4, 50678 Köln) kostenpflichtig orthographisch korrigiert.
2. Es sind ausschließlich die von mir angegebenen Quellen und Hilfsmittel verwendet worden.
3. Sämtliche wörtliche und nicht wörtliche Zitate aus anderen Werken sind gemäß den wissenschaftlichen Zitierregeln kenntlich gemacht.
4. Die von mir vorgelegte Arbeit ist bisher noch nicht oder an der nachstehend aufgeführten Stelle vollständig/auszugsweise veröffentlicht. (s. Anlage)
5. Die von mir vorgelegte Arbeit ist bisher noch in keiner Form als Bestandteil einer Prüfungs-/Qualifikationsleistung vorgelegt worden.
6. Ich bin nicht zweimal in einem Promotionsverfahren an einer deutschen Hochschule aufgrund einer Ablehnung der Dissertation oder nicht bestandener Prüfungsleistungen gescheitert.
7. Die von mir eingereichte Dissertation habe ich unter Beachtung der Grundsätze zur Sicherung guter wissenschaftlicher Praxis erstellt.
8. Über die Bedeutung und die strafrechtlichen Folgen einer falschen eidesstattlichen Erklärung gemäß § 156 StGB bin ich mir bewusst.
9. Meine Angaben entsprechen der Wahrheit und ich habe diese nach bestem Wissen und Gewissen gemacht.

Ort, Datum	Unterschrift der*des **Doktorand*in**