International Journal of
Molecular Sciences

MDPI

Article

Functional Characterisation of Three Glycine *N*-Acyltransferase Variants and the Effect on Glycine Conjugation to Benzoyl–CoA

Johann M. Rohwer [1], Chantelle Schutte [2] and Rencia van der Sluis [2,*]

[1] Laboratory for Molecular Systems Biology, Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland, Stellenbosch 7602, South Africa; jr@sun.ac.za
[2] Focus Area for Human Metabolomics, North-West University, Private Bag X6001, Potchefstroom 2520, South Africa; schutte.tella@gmail.com
* Correspondence: 21224919@nwu.ac.za; Tel.: +27-18-299-2068

Abstract: The glycine conjugation pathway in humans is involved in the metabolism of natural substrates and the detoxification of xenobiotics. The interactions between the various substrates in this pathway and their competition for the pathway enzymes are currently unknown. The pathway consists of a mitochondrial xenobiotic/medium-chain fatty acid: coenzyme A (CoA) ligase (ACSM2B) and glycine *N*-acyltransferase (GLYAT). The catalytic mechanism and substrate specificity of both of these enzymes have not been thoroughly characterised. In this study, the level of evolutionary conservation of GLYAT missense variants and haplotypes were analysed. From these data, haplotype variants were selected (156Asn > Ser, [17Ser > Thr,156Asn > Ser] and [156Asn > Ser,199Arg > Cys]) in order to characterise the kinetic mechanism of the enzyme over a wide range of substrate concentrations. The 156Asn > Ser haplotype has the highest frequency and the highest relative enzyme activity in all populations studied, and hence was used as the reference in this study. Cooperative substrate binding was observed, and the kinetic data were fitted to a two-substrate Hill equation. The coding region of the *GLYAT* gene was found to be highly conserved and the rare 156Asn > Ser,199Arg > Cys variant negatively affected the relative enzyme activity. Even though the 156Asn > Ser,199Arg > Cys variant had a higher affinity for benzoyl-CoA ($s_{0.5,benz}$ = 61.2 μM), k_{cat} was reduced to 9.8% of the most abundant haplotype 156Asn > Ser ($s_{0.5,benz}$ = 96.6 μM), while the activity of 17Ser > Thr,156Asn > Ser ($s_{0.5,benz}$ = 118 μM) was 73% of 156Asn > Ser. The in vitro kinetic analyses of the effect of the 156Asn > Ser,199Arg > Cys variant on human GLYAT enzyme activity indicated that individuals with this haplotype might have a decreased ability to metabolise benzoate when compared to individuals with the 156Asn > Ser variant. Furthermore, the accumulation of acyl-CoA intermediates can inhibit ACSM2B leading to a reduction in mitochondrial energy production.

Keywords: glycine conjugation; glycine *N*-acyltransferase (GLYAT); benzoate; hippurate; coenzyme A

Citation: Rohwer, J.M.; Schutte, C.; van der Sluis, R. Functional Characterisation of Three Glycine *N*-Acyltransferase Variants and the Effect on Glycine Conjugation to Benzoyl–CoA. *Int. J. Mol. Sci.* **2021**, *22*, 3129. https://doi.org/10.3390/ijms22063129

Academic Editor: Narimantas K. Cenas

Received: 16 February 2021
Accepted: 15 March 2021
Published: 18 March 2021

1. Introduction

The glycine conjugation pathway is a two-step enzymatic reaction responsible for the metabolism/detoxification of natural substrates from (i) food (for example., salicylate [1], dietary polyphenols and medium-chain fatty acids (MCFAs)), (ii) xenobiotics (for example benzoate), and (iii) metabolites produced from organic acidaemia [2–4] (Figure 1). Benzoate and salicylate are activated to an acyl-coenzyme A (CoA) by the mitochondrial xenobiotic/medium-chain fatty acid: CoA ligases (ACSM2B,EC 6.2.1.2) [5,6] and subsequently conjugated to glycine by glycine *N*-acyltransferase (GLYAT, EC 2.3.1.13) to form hippuric acid and salicyluric acid, respectively [3,7,8]. Gut microorganisms produce benzoyl–CoA, which is a substrate for glycine conjugation, from dietary polyphenols [9]. MCFAs, for example, caprylic acid, are activated by ACSM2B ligase in the liver before entering the mitochondrial beta-oxidation cycle [10]. The overall rate of glycine conjugation can be influenced by the availability of cofactors, variations in the *ACSM2B* and *GLYAT* genes, and a difference in expression levels of ACSM2B and GLYAT [11–14].

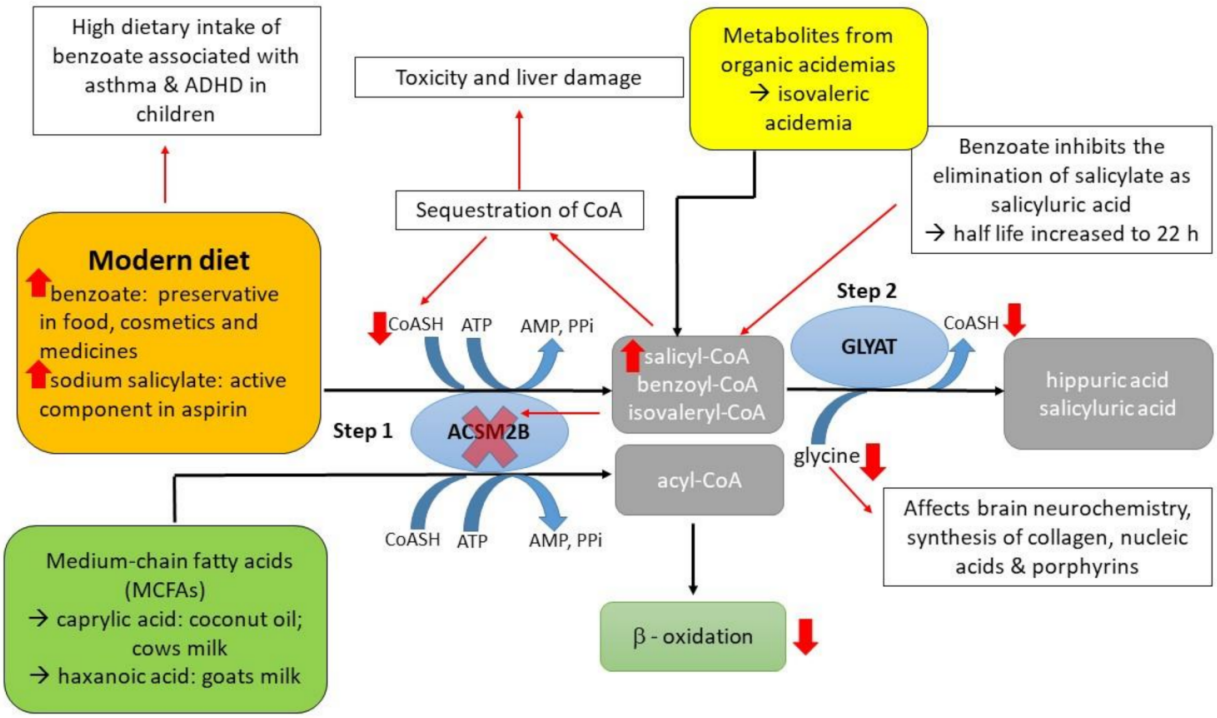

Figure 1. A schematic representation depicting the effect of a high dietary intake of benzoate on the glycine conjugation pathway. Benzoate and salicylate are activated to an acyl-CoA by the mitochondrial xenobiotic/medium-chain fatty acid: coenzyme A (CoA) ligase (ACSM2B) and subsequently conjugated to glycine by glycine *N*-acyltransferase (GLYAT). A high dietary intake of benzoate can lead to a decrease in available glycine. The acyl-CoAs can no longer be conjugated to glycine by GLYAT resulting in an increase of the acyl-CoA intermediates and sequestration of free CoA. The acyl-CoA intermediates can inhibit ACSM2B leading to a reduction in mitochondrial energy production. Metabolites from organic acidaemia, for example isovaleryl-CoA from isovaleric acidaemia, provide an additional detoxification load to the glycine conjugation pathway.

All individuals will need to metabolise benzoate and salicylate from natural sources (for example, benzoate and salicylate present in berries and milk products [1]), and depending on whether an individual's diet is high in polyphenols, benzoate will also be formed by the gut microbes [15]. However, the exposure of humans to these compounds is increasing [16,17] since benzoate is widely used as a preservative in food and pharmaceuticals [18]. Benzoic acid consumption has been linked to adverse effects such as diarrhoea, metabolic acidosis, tremors, and childhood hyperactivity syndrome [19,20]. These symptoms might be the result of low levels of glycine as glycine conjugation is needed to detoxify the high levels of benzoate. The glycine shortage results in a reduction in creatinine, glutamine, urea, and uric acid production (Figure 1) [1]. A shortage of glycine can, furthermore, increase the accumulation of acyl-CoA intermediates. Accumulation of xenobiotic-CoA esters results in the sequestration of CoA and the inhibition of the acid: CoA ligases [21]. If the ACSM2B ligase is inhibited by the acyl-CoA intermediates, ACSM2B can no longer activate MCFAs for use during beta-oxidation. The glycine conjugation pathway, therefore, plays a fundamental role in the homeostatic energy balance within the mitochondria by preventing coenzyme A (CoASH) sequestration. Although hippuric acid and benzoic acid have a similar water solubility value, glycine conjugation decreases the toxicity of benzoate by forming less lipophilic conjugates that can be more readily transported out of the mitochondria [22]. Studies have shown that benzoic acid inhibits the elimination of salicyluric acid, and that the accumulation of salicyl-CoA may result in toxicity and liver damage [23,24].

A previous study, in a cohort of isovaleric acidaemia patients from South Africa, evaluated the effectiveness of glycine supplementation on the clinical outcome of these patients. Even though these patients were all homozygous for the same isovaleryl-CoA dehydrogenase mutation, variation in the responsiveness to glycine supplementation was observed [25].

Glycine supplementation is used as a treatment in these patients in order to conjugate the toxic accumulated isovaleryl-CoA to glycine by GLYAT forming isovalerylglycine [26] (Figure 1). It is hypothesised that missense variants in the *GLYAT* gene do not contribute significantly to the inter-individual glycine conjugation rates observed in these patients as the *GLYAT* gene was shown to be highly conserved in a small cohort of 1537 individuals [27]. Variation in the expression level of GLYAT may influence the glycine conjugation rate in humans, as was shown for rats, where diet affected the expression of GLYAT in the liver [28,29]. Furthermore, benzoyl–CoA is the preferred substrate for GLYAT, followed by salicyl–CoA and then isovaleryl–CoA [30]. Therefore, a diet high in benzoate can be one of the factors that better explain the inter-individual variation in responsiveness to glycine supplementation in these patients as the derived benzoyl–CoA will outcompete isovaleryl–CoA as a substrate. Currently, no studies are available that provide data on the interaction/competition of the various substrates involved in the glycine conjugation pathway.

It is extremely difficult to obtain fresh human liver samples in order to study the catalytic mechanism of GLYAT using a purified enzyme [14]. Previously determined GLYAT kinetic parameters, using either mitochondrial lysate preparations from liver tissue or purified recombinantly expressed GLYAT enzyme, vary considerably [28,30–34], with only three studies determining the K_m value for glycine [28,32,33]. These studies also assumed that GLYAT followed a Michaelis–Menten reaction mechanism and, as a consequence, reported a sequential two-substrate mechanism. A preliminary study, in which the recombinantly expressed purified 156Asn > Ser GLYAT variant was characterised, indicated that GLYAT exhibits mechanistic kinetic cooperativity and not a Michaelis–Menten reaction mechanism [35]. Most of the studies also did not report on whether the wildtype GLYAT or a variant was used in the analyses. This is important in order to compare the different studies as it has been shown that single nucleotide polymorphisms (SNPs) can alter the kinetic parameters of GLYAT [31,34].

The aim of this study was to analyse the genetic diversity and haplotype variation of the *GLYAT* gene using a larger population (125,748 exomes and 15,708 whole genomes) to determine the level of conservation in the worldwide population. Haplotypes occurring at low frequencies in a population are more likely to be deleterious and therefore associated with adverse detoxification. This analysis also allowed us to identify the haplotypes for further characterisation in terms of enzyme activity and mechanism. In an effort to address the limitations of previous studies, the bi-substrate (glycine and benzoyl–CoA) reaction kinetics of the purified 156Asn > Ser, 17Ser > Thr,156Asn > Ser, and 156Asn > Ser,199Arg > Cys GLYAT variants were determined over a wide range of substrate concentrations in order to determine the effect of genetic variants on the relative enzyme activity and kinetic parameters.

2. Results and Discussion

Understanding and quantifying the interaction and competition between various substrates or xenobiotics for detoxification by glycine conjugation requires information on differences in the enzyme activity and catalytic mechanism of GLYAT variants. Through the analyses of the variant data that are available on the gnomAD browser and Ensembl database, the allele frequencies in the worldwide population and the level of conservation of the GLYAT haplotypes could be determined. Identifying the haplotype frequencies also resulted in the enzymatic characterisation of relevant variants of the enzyme. Based on the determined haplotype frequencies, three haplotypes were chosen and characterised in terms of relative enzyme activity and catalytic mechanism in order to determine the effect of the missense variants on the enzyme activity.

2.1. Level of Conservation of the GLYAT Gene

The exome and whole genome data available on the gnomAD browser (gnomad.broadinstitute.org/; accessed on 10 September 2020) [36] and the haplotype data available on the Ensembl database (ensemble.org; accessed on 10 September 2020) [37] were analysed in order to determine the allelic variation and haplotype diversity in the worldwide population.

The gnomAD browser is the largest database that includes allele frequencies of variants located in protein-coding regions, as both exome and genome sequencing data from a wide variety of large-scale sequencing projects are combined. The allele frequencies of the missense variants found in GLYAT (ENST00000344743.3) were downloaded from the gnomAD browser [36] and analysed (Supplementary Table S1).

For *GLYAT*, 193 missense variants were identified. Of these, only two variants had an allele frequency > 0.5%, that is., Asn156Ser (94.87%) and Ser17Thr (19.67%), while the remaining variants were rare with allele frequencies ranging between 0.2% and 0.0004% (a rare variant is only found in one allele of one individual out of a total of 141,456 analysed). The Asn156Ser and Ser17Thr variants had the highest allele frequencies in all of the populations analysed (Table 1 and Figure 2). The same trend was observed in a smaller previous study that included data from the 1000 genomes and HapMap projects, in addition to data from 61 Caucasian Afrikaners, 4 Khoi-San, and 1 Bantu individual from South Africa [27]. The Asn156Ser missense variant also had the highest homozygous genotype frequency of 90.3%, followed by Ser17Thr (4.5%). Only six other variants were found as homozygotes, namely Arg131His (East Asian—0.005%), Arg131Cys (South Asian—0.002%), Met65Thr (Latino—0.001%), Thr73Ile (East Asian—0.001%), His101Tyr (South Asian—0.002%) and Thr244Met (Latino—0.001%).

Table 1. Population data used in this study.

Population	Exomes	Genomes	Total
African/African American	8 128	4 359	12,487
Latino	17,296	424	17,720
Ashkenazi Jewish	5 040	145	5 185
East Asian	9 197	780	9 977
Finnish	10,824	1 738	12,562
Non-Finnish European	56,885	7 718	64,603
South Asian	15,308	#	15,308
Other*	3 070	544	3 614
Female	57,787	6 967	64,754
Male	67,961	8 741	76,702
Total	125,748	15,708	138,632

31 South Asian samples were grouped with Other. * Individuals were classified as "other" if they did not unambiguously cluster with the major populations (that is i.e., African, African American, Latino, Ashkenazi Jewish, East Asian, Finnish, Non-Finnish European, South Asian) in a principal component analysis (PCA). (This is an extract of the population data available on the gnomAD browser).

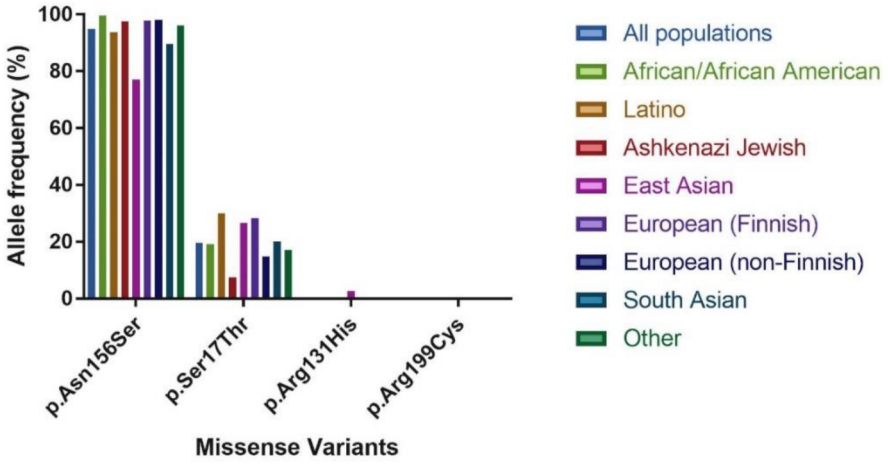

Figure 2. Comparison of the allele frequency between different populations of missense variants of interest to this study.

In the African/African American population, the Asn156Ser allele occurred at a frequency of 99.58%, and the Ser17Thr allele at 19.17%, with the rest of the variants occurring at a frequency below 0.07%. Several studies have shown that in African populations, the genetic diversity and discovery rate of novel variants is higher [38–41]. However, the expected high level of genetic diversity was not observed in the *GLYAT* gene, indicating that the gene is conserved even in genetically diverse individuals. In contrast, a lower allele frequency for Asn156Ser (77.16%) and slightly higher frequency for Ser17Thr (26.83%) and Arg131His (2.62%) were found in the East Asian population. Diet and the environment can be a strong driver of selection [42]. An example is the high prevalence of the slow acetylation phenotype in populations practicing farming and herding [43]. In other species, such as domestic cats, a reduced ability to metabolise several drugs and structurally related phenolic compounds has been observed [44–47] due to gene inactivation as a consequence of minimal exposure to plant-derived toxicants. It was also shown in rats that diet influences the expression level of GLYAT in the liver [28,29]. Whether diet played a role in the selection of the variants found in the East Asian population or if diet might affect the expression level of GLYAT in the liver of humans needs to be further investigated.

The 1000 Genomes haplotype data [37,48] for *GLYAT* are summarised in Supplementary Table S2. The haplotype frequencies were analysed across the 26 populations to find the haplotypes with the highest frequency. In total, 25 haplotypes were reported of which four have a frequency >0.5% [156Asn > Ser (69.9%); 17Ser > Thr,156Asn > Ser (21.5%); REF (7.15%); 131Arg > His,156Asn > Ser (0.52%)]. This study and previous studies [27,49] clearly show that the 156Asn > Ser variant should be regarded as the reference sequence due to the high allele frequency identified in all populations studied. The rare haplotype frequencies ranged from 0.22 to 0.02%. The 156Asn > Ser haplotype had the highest frequency in all the populations, but notably, the frequency was lower in the East Asian (51.6%) and the South Asian (65.6%) populations when compared to the African (79.5%) and European (80.8%) populations. Of the 25 haplotypes, 17 were predicted by the sorting intolerant from tolerant (SIFT) [50] and polymorphism phenotyping (PolyPhen) [51] algorithm tools, to have a deleterious effect on protein function. Only two rare haplotypes [73Thr > Ile (0.18%) and 17Ser > Thr (0.02%)] were not found in combination with the 156Asn > Ser variant.

The Tajima's D value was calculated using MEGA X [52] to determine whether the human *GLYAT* gene is evolving neutrally (Table 2). The highly negative Tajima's D value of –2.13 indicates that variants located within the *GLYAT* gene are under negative selection. This is supported by the large number of low-frequency alleles observed in all of the populations analysed in this study.

Table 2. Results from Tajima's neutrality test.

Gene	Number of Sequences (m)	Number of Segregating Sites (S)	Nucleotide Diversity (π)	Tajima Test Statistic (D)
GLYAT	25	21	0.007748	−2.13

Comparative phylogenetic studies between apes and humans clarify the patterns of evolutionary change in the human lineage [53–55]. To construct the phylogenetic tree, the *GLYAT* haplotype sequence data were used to perform maximum likelihood fits to determine the best amino acid substitution model to use [56]. The model predicted to have the best fit, was the Jones–Thornton–Taylor (JTT) model [57] with discrete gamma rate categories (+G). In order to determine the evolution of the *GLYAT* gene since the human and chimpanzee split, phylogenetic analyses were subsequently performed (Figure 3). The robustness of the tree was assessed using 500 bootstrap replicates. For *GLYAT*, the human/chimpanzee/gorilla/bonobo/orangutan clade had good bootstrap support of 100%, while ancestral nodes within this clade were poorly supported with values ranging from 23–63% (Figure 3) [58]. These ancestral nodes were made up of human haplotypes with very low haplotype frequencies (<0.6%). The phylogenetic analyses further indicate that the *GLYAT* gene is conserved across all population groups. The *Pan troglodytes*, *Pan*

paniscus, Pongo abelii, and *Gorilla gorilla* haplotypes are found within the human clade which indicates that relatively few changes have occurred within the *GLYAT* gene since the chimpanzee and human split.

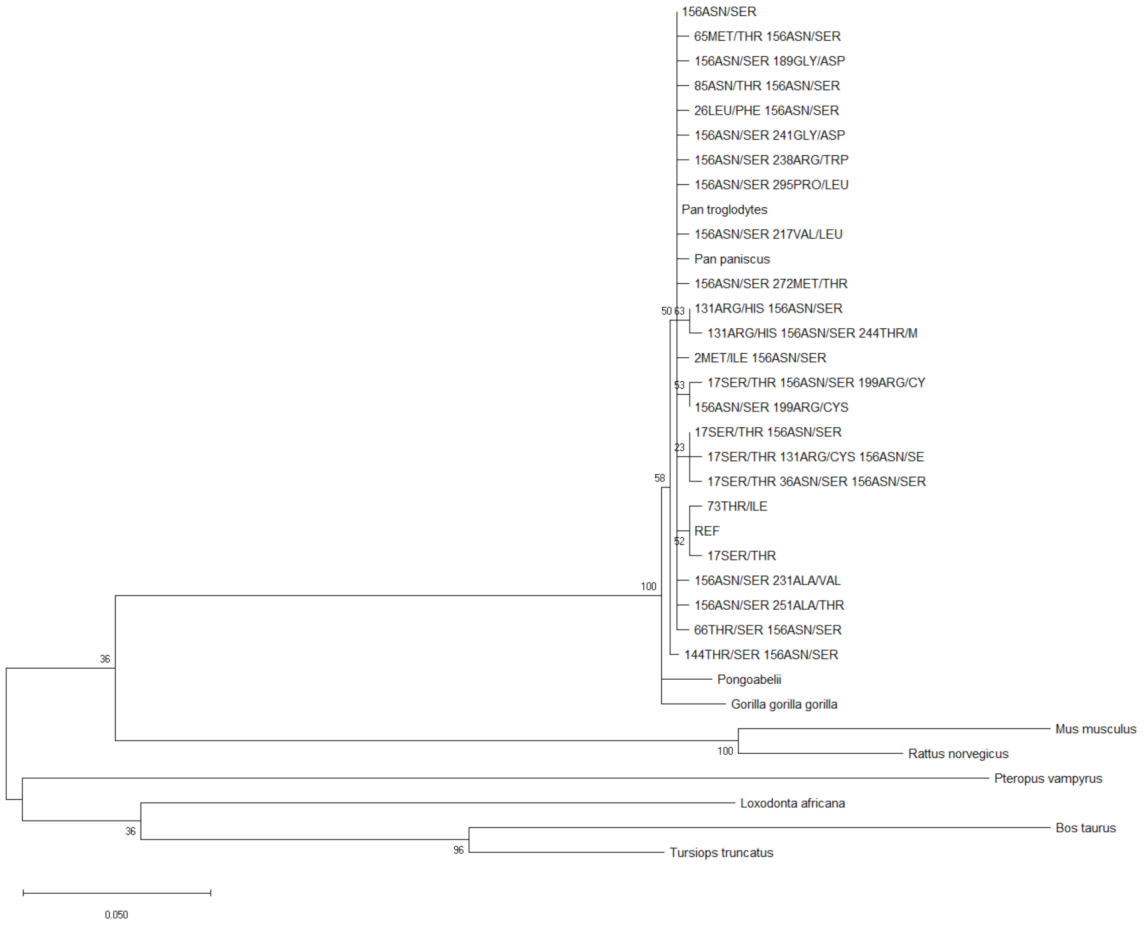

Figure 3. Molecular Phylogenetic analysis of the GLYAT haplotypes using the maximum likelihood method. The evolutionary history was inferred by using the maximum likelihood method based on the Jones–Thornton–Taylor (JTT) matrix-based model [57]. The bootstrap consensus tree inferred from 500 replicates [59] is taken to represent the evolutionary history of the taxa analysed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbour-Join (NJ) and Bio NJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (five categories; +*G*, parameter = 2.7686). The analysis involved 34 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 296 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [52].

The allelic and haplotype diversity analyses, together with the negative Tajima's D value, all point to the fact that a large number of rare variants/haplotypes are found in the worldwide population. The phylogenetic analyses (Figure 3) suggested that the coding regions of the *GLYAT* gene are well conserved through evolution. Therefore, even in a large diverse population of 141,456 individuals, the *GLYAT* gene was shown to be highly conserved. This confirmed the results of the previous smaller study where only 1537 individuals were included [27].

2.2. Relative Enzyme Activity and Catalytic Parameters

We selected two of the four haplotypes with a frequency > 0.5% [156Asn > Ser (69.9%); 17Ser > Thr,156Asn > Ser (21.5%)] and one rare haplotype [156Asn > Ser,199Arg > Cys

(0.02%)] for further kinetic analysis. The recombinantly expressed and purified 17Ser > Thr and 199Arg > Cys variants were previously characterised [31]. It was shown that the 17Ser > Thr variant had activity comparable to that of the wild-type enzyme (listed as "REF" in Supplementary Table S2) and the 199Arg > Cys mutation had less than 5% activity of that of the wild-type enzyme. The 17Ser > Thr,156Asn > Ser and 156Asn > Ser,199Arg > Cys haplotypes were both predicted by the SIFT [50] and PolyPhen [51] algorithm tools to have a deleterious effect on protein function (Supplementary Table S2).

To determine if there are indeed differences in catalytic activity between the GLYAT haplotype variants, we initially compared the relative enzyme activities of the three haplotypes at one specific set of substrate concentrations (Figure 4). There were significant differences between the three haplotypes, with 156Asn > Ser,199Arg > Cys showing only 12.3% of the activity of 156Asn > Ser, and 17Ser > Thr,156Asn > Ser exhibiting an intermediate level of activity (49.4%). When comparing the enzyme activity of the 199Arg > Cys mutation on its own [31] with that of the 156Asn > Ser,199Arg > Cys haplotype, an increase in enzyme activity of 5% was observed. This indicates how important it is to characterise haplotypes rather than SNPs, especially in the field of pharmacokinetics and pharmacogenomics. The most recent study [34] which compared the enzyme activity of the 61Gln > Leu variant with that of the wild-type, found that the 61Gln > Leu variant showed a decrease in specific activity when compared to the wild-type. It is important to note that the 61Gln > Leu mutation has only been identified in two haplotypes in the South African Afrikaner population that is, 61Gln > Leu,156Asn > Ser and 17Ser > Thr, 61Gln > Leu,156Asn > Ser [27], and therefore, the activity might be affected by the other SNPs in these haplotypes. The crystal structure of GLYAT is not available but based on a molecular model, the 199Arg > Cys mutation alters a highly conserved Arg in an α-loop- α motif which is important for substrate binding in the Gcn5-related *N*-acetyltransferases (GNAT) superfamily [60,61]. Moreover, 156Asn > Ser is on a poorly predicted loop from Lys159 to Met167, making interpretation of its role in enzyme activity difficult. The relative enzyme activity of the haplotypes also correlated with the PolyPhen and SIFT predictions, in addition to the haplotype frequency, with rarer haplotypes being less active.

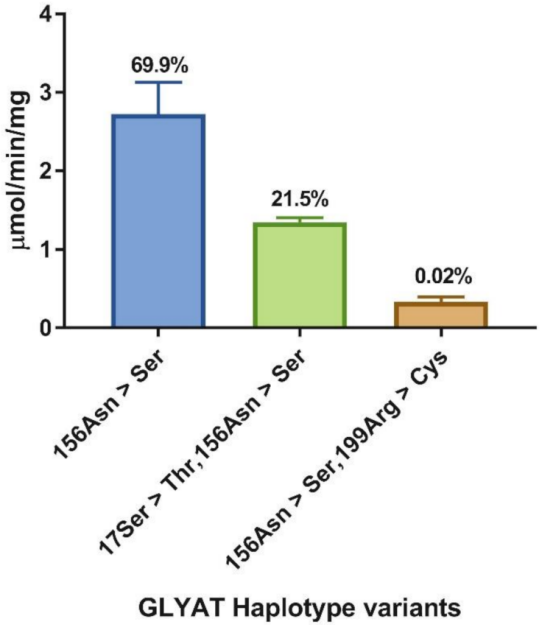

Figure 4. Relative enzyme activity of glycine *N*-acyltransferase (GLYAT) haplotypes 156Asn > Ser, 17Ser > Thr,156Asn > Ser and 156Asn > Ser,199Arg > Cys. Assays were performed in triplicate with 2 µg protein, 20 mM glycine, and 80 µM benzoyl–CoA. The standard deviation is shown by the error bars and the haplotype frequencies are indicated above each bar.

Because we observed such striking differences between the catalytic activities of the three GLYAT variants, the kinetic parameters of each of the haplotypes were characterised in greater detail to determine possible differences in the catalytic mechanism. Initial rates were determined for a range of concentrations of both substrates. Cooperative substrate binding was observed. To account for this, the data were fitted to a two-substrate Hill equation (see Section 3). Since the concentration of one substrate in two-substrate kinetic experiments will affect the values obtained for the kinetic parameters of the other substrate, we varied both substrate concentrations in a grid and performed a global fit using non-linear regression to obtain a single set of enzyme-kinetic parameters that best describe all the data.

The final fits for the three haplotypes were visualised with 3D-surface plots (Figure 5), showing the fit of the kinetic model to all of the experimental data. Data were processed and fitted as described in Section 3. Each data point represents the mean of triplicate measurements. The best model fit is indicated by the coloured surface. The cooperative kinetics can be clearly observed, especially for benzoyl–CoA as substrate. Because of the difficulty in visualising these data in three dimensions, especially the agreement between model and data, two-dimensional activity plots are included as well (Supplementary Figures S1–S6), showing rate against glycine concentration at each separate benzoyl–CoA concentration, and vice versa. Importantly, while each of these plots shows individual experimental data points and a model fit, it should be noted that the line indicates the global fit of the model to all the data, not only the data for a particular plot, which explains the discrepancies observed in some cases. For further illustration, Lineweaver–Burk plots of selected datasets are included as Supplementary Figures S7–S12; the cooperative response can be observed as a deviation from linearity in the fit. The kinetic parameters obtained from the fitting are summarised in Table 3. Parameters and standard errors were estimated from a global fit of initial rate data to the two-substrate Hill equation. Refer to Section 3 for details and definition of the kinetic parameters.

Table 3. Enzyme–kinetic parameters of the GLYAT variants.

Haplotype	V_f μmol min^{-1} mg protein^{-1}	k_{cat} s^{-1}	$s_{0.5,gly}$ mM	h_{gly}	$s_{0.5,benz}$ μM	h_{benz}
156Asn > Ser	0.85 ± 0.06	0.48 ± 0.03	23 ± 2	1.6 ± 0.1	97 ± 3	2.1 ± 0.1
17Ser > Thr,156Asn > Ser	0.62 ± 0.02	0.35 ± 0.01	29 ± 3	1.3 ± 0.1	118 ± 7	1.5 ± 0.1
156Asn > Ser,199Arg > Cys	0.083 ± 0.005	0.047 ± 0.003	30 ± 3	1.4 ± 0.1	61 ± 3	3.5 ± 0.5

A number of features of the kinetic dataset in Table 3 merit comment. First, the catalytic activity (k_{cat}) mirrored the trend observed in the initial activity study (Figure 4), with the haplotype variant 156Asn > Ser,199Arg > Cys displaying only 10% of the activity of 156Asn > Ser, while that of 17Ser > Thr,156Asn > Ser was 73% of 156Asn > Ser. The k_{cat} was also by far the most affected parameter by the sequence changes introduced in the haplotypes.

Secondly, the binding affinity was affected to a lesser extent, being slightly weaker with the half-saturation constant for glycine ($s_{0.5,gly}$) and increasing by 28% and 33% for 17Ser > Thr,156Asn > Ser and 156Asn > Ser,199Arg > Cys, respectively, when compared to the most abundant haplotype 156Asn > Ser. In contrast, the effects on the half-saturation constant for benzoyl–CoA ($s_{0.5,benz}$) were mixed, with 17Ser > Thr,156Asn>Ser showing an increase of 22%, while 156Asn > Ser,199Arg > Cys showed a decrease of 37% in this parameter. Overall, the binding affinities for both substrates were in the same range for all three GLYAT variants.

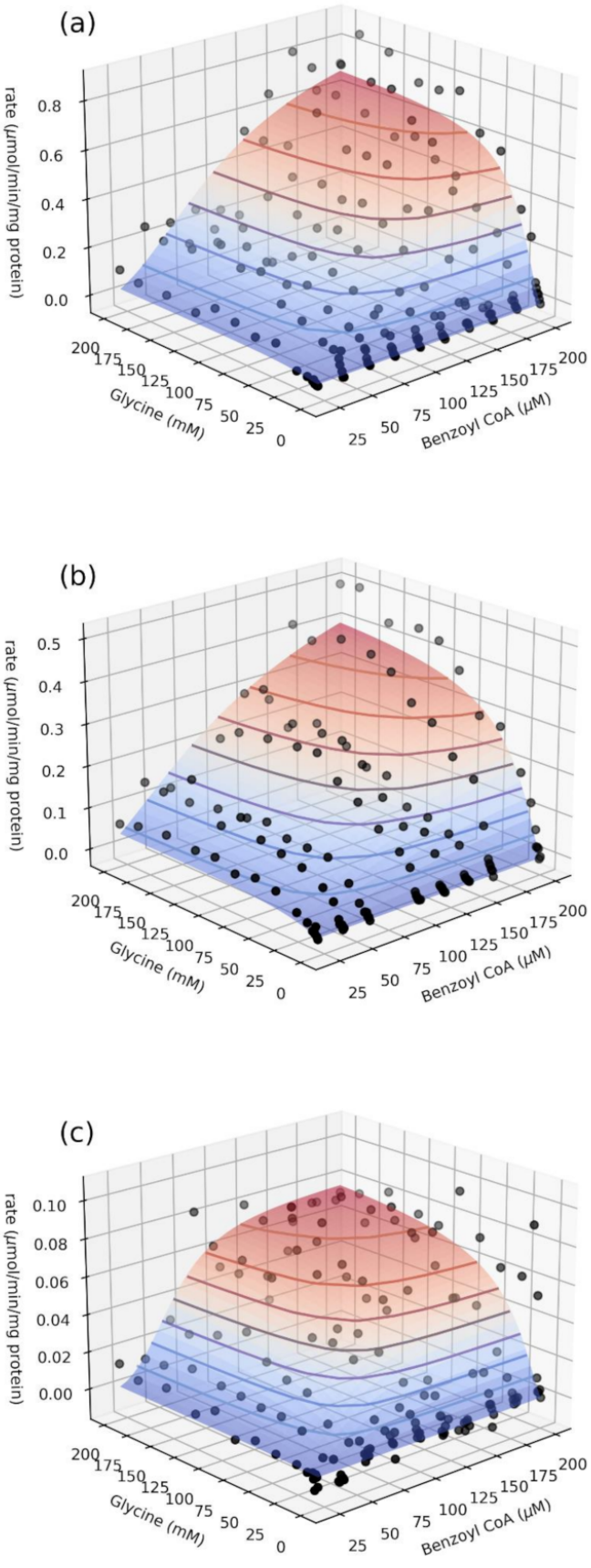

Figure 5. Global fit of initial rate data to the two-substrate Hill equation (see Section 3) for the haplotype variants 156Asn > Ser (**a**), 17Ser > Thr,156Asn > Ser (**b**), and 156Asn > Ser,199Arg > Cys (**c**).

Thirdly, both substrates exhibited cooperative binding for all of the haplotypes studied. The cooperativity was stronger for benzoyl–CoA (with Hill coefficients ranging between 1.5 and 3.5) than for glycine (with Hill coefficients ranging between 1.3 and 1.6). Moreover, all three haplotype variants showed similar cooperativity for glycine, while there were marked differences for benzoyl–CoA. In comparison to the 156Asn > Ser variant, 17Ser > Thr,156Asn > Ser showed decreased cooperativity, while 156Asn > Ser,199Arg > Cys showed markedly increased cooperativity.

Since all the known mammalian orthologs of GLYAT, including human (studied here), bovine, and chimpanzee, are monomeric enzymes [7,30,32], the cooperative kinetic responses observed merit further discussion. Ferdinand [62] already showed more than 50 years ago that bi-substrate monomeric enzymes can under certain circumstances exhibit sigmoidal kinetics. This so-called kinetic cooperativity can specifically occur if the two pathways in which, say, substrate A or substrate B binds to the enzyme first, both do occur, but the enzyme shows a preference for one of these pathways. More recently [63], kinetic cooperativity in human glucokinase has been shown to result from the unliganded enzyme existing in two states—a ground state and an activated state. If the rate of interconversion between these states is of the same order as the catalytic rate constant, sigmoidal kinetic responses towards glucose concentration can be observed. More generally, cooperative kinetic responses may also be the result of 'allokairy' [64], that is, information changes being transmitted through time, such as an enzyme in an active state after turnover relaxing back to an inactive state over time. The exact mechanism by which GLYAT exhibits sigmoidal kinetics still needs to be elucidated.

Taking all these data into account, the results presented in Table 3 suggest that changes in activity between the three haplotype variants are predominantly due to *V*-effects (that is changes in k_{cat}), with substrate binding playing a lesser role. The only exception to this might be the binding of benzoyl–CoA to the 156Asn > Ser,199Arg > Cys variant, in which a decreased $s_{0.5}$ and significantly increased *h*-value suggest a more potent response as compared to 156Asn > Ser. However, with the tenfold lower k_{cat}, this increased potency is unlikely to offer significant advantages due to an increase in the population of non-productive binding conformations [65]. Allosteric regulation of certain enzymes is an evolutionary mechanism of adaptation for the selection of specific substrates because the enzyme specificity for substrates controls metabolic flow by sorting metabolites into distinct paths [66]. The glycine conjugation pathway maintains a delicate balance in CoA levels within the mitochondria [21] and this might explain why deleterious variants such as the 156Asn > Ser,199Arg > Cys haplotype are maintained at very low frequencies in the population.

2.3. Comparison of GLYAT Kinetic Parameters to Literature Values

Human GLYAT has been kinetically characterised in previous studies, both in our laboratory and by other investigators. We, therefore, compared the kinetic parameters obtained in this study to previous values from the literature (Table 4). In terms of substrate affinity, previous studies reported apparent Michaelis constants and did not investigate cooperative effects, which makes a direct comparison with our values difficult. Nevertheless, we list the half-saturation constants obtained in this study together with the other K_{Mapp}-values, reasoning that at least a semi-quantitative comparison is justified on the grounds that firstly the reported Michaelis constants are apparent (and not true) values, and secondly, both K_{Mapp} and $s_{0.5}$ are operationally defined as the substrate concentration giving a rate of half V_{max} at saturating levels of the second substrate.

Overall, both the maximal activity and the affinity parameters agreed well with published literature, with our reported values falling in the ranges reported in previous studies (Table 4). It should be noted that two studies [30,34] reported in Table 4 might contain calculation errors as explained in the footnote. Importantly, though, the differences in the kinetic mechanism for the different haplotypes (Table 3) have not been reported previously.

Table 4. Kinetic parameters of benzoyl–CoA and glycine for GLYAT as reported in the literature.

Parameters	Values	Recombinant Variant/Isolated from Liver	Reference
K_{Mapp} (benzoyl-CoA) (μM)	13	Purified GLYAT from human liver	[33]
	28 ± 5	Purified recombinant 17Ser > Thr variant	[31]
	38 ± 4	Purified recombinant 156Asn > Ser variant	[31]
$s_{0.5}$	61 ± 3	Purified recombinant 156Asn > Ser,199Arg > Cys variant	This study (bi-substrate Hill)
K_{Mapp} (benzoyl-CoA) (μM)	67 ± 5	Partially purified GLYAT from human liver	[32]
	79 ± 38	Purified recombinant wildtype	[34]
	88 ± 66	Purified recombinant 156Asn > Ser variant	[34]
$s_{0.5}$	97 ± 3	Purified recombinant 156Asn > Ser variant	This study (bi-substrate Hill)
	118 ± 7	Purified recombinant 17Ser > Thr,156Asn > Ser variant	This study (bi-substrate Hill)
K_{Mapp} (benzoyl-CoA) (μM)	139 ± 85	Purified recombinant L_{61} variant	[34]
	209	Purified recombinant 156Asn > Ser variant	[28]
	57900 *	Purified GLYAT from human liver	[30]
K_{Mapp} (glycine) (mM)	6.4	Purified GLYAT from human liver	[33]
	6.5 ± 1	Partially purified GLYAT from human liver	[32]
$s_{0.5}$	23 ± 2	Purified recombinant 156Asn > Ser variant	This study (bi-substrate Hill)
K_{Mapp} (glycine) (mM)	26.6	Purified recombinant 156Asn > Ser variant	[28]
$s_{0.5}$	29 ± 3	Purified recombinant 17Ser > Thr,156Asn > Ser variant	This study (bi-substrate Hill)
	30 ± 3	Purified recombinant 156Asn > Ser199Arg > Cys variant	This study (bi-substrate Hill)
V_{max} (nmol/min/mg)	83 ± 5	Purified recombinant 156Asn > Ser199Arg > Cys variant	This study (bi-substrate Hill)
	543 ± 21	Purified GLYAT from human liver	[33]
	620 ± 20	Purified recombinant 17Ser > Thr,156Asn > Ser variant	This study (bi-substrate Hill)
	665 ± 40	Purified recombinant 17Ser > Thr variant	[31]
	807	Recombinant 156Asn > Ser variant	[28]
	850 ± 60	Purified recombinant 156Asn > Ser variant	This study (bi-substrate Hill)
	1230	Purified recombinant 156Asn > Ser variant	[31]
	17,100 #	Purified GLYAT from human liver	[30]
	121,000 +	Purified recombinant L_{61} variant	[34]
	490,000 +	Purified recombinant wildtype	[34]
	1,359,000 +	Purified recombinant 156Asn > Ser variant	[34]

* The unit reported in the article is mM; therefore, this value was converted to μM. We do think that this is either a typing or a calculation error and should be 57.9 μM but because the enzyme assay conditions are not described in the article it is difficult to ascertain where the error lies. # This error also affected the V_{max} value. + These V_{max} values are too high to be correct and might also be a calculation error.

Previous substrate specificity studies performed for GLYAT [30,33] need to be repeated in the light of the fact that the measure for substrate specificity of non-cooperative enzymes is k_{cat}/K_m or its apparent value if other co-substrates are present [67], but for cooperative enzymes, the appropriate measure is $k_{cat}/s^h_{0.5}$ [68]. This is due to the fact that the degree of cooperativity affects the order of the specificity of the substrates. This effect needs to be taken into account especially when evaluating the substrate specificity of competing substrates in physiological conditions. For example, in the case of isovaleric acidaemia, it would be important to know whether benzoyl–CoA will outcompete isovaleryl–CoA as substrate and whether this has an effect on the observed differences in the effectiveness of using glycine supplementation as therapy in these patients [25].

It is very difficult to extrapolate the in vitro values to the in vivo environment where several substrates and pathways need to be considered. Previous studies have found that if

the amount of benzoate administered to individuals is increased, hippuric acid excretion will also increase to a maximum after which the excretion level will remain constant. The administration of glycine, on the other hand, resulted in a rapid increase in the hourly excretion rate of hippuric acid [69–71]. These studies are very old and the GLYAT haplotypes of the individuals analysed in these studies were not known at the time. Even though significant inter-individual variability in glycine conjugation capacity has been demonstrated after administering benzoate [72], this is probably due to the fact that it is very difficult to control for individual differences in diet during these studies. The metabolites of dietary polyphenols produced by microorganisms contribute the largest portion of the natural substrates that are metabolised by the glycine conjugation pathway [9]. Aspirin and benzoate have been used to characterise individual glycine conjugation capacity; however, adverse reactions, aspirin intolerance, and Reye's syndrome in children are substantial drawbacks. Therefore, the use of p-aminobenzoic acid (PABA) as an alternative glycine conjugation probe was investigated in a previous study. For the study, 10 human volunteers participated in a PABA challenge test, and p-aminohippuric acid (PAHA), p-acetamidobenzoic acid, and p-acetamidohippuric acid were quantified in urine samples. The *GLYAT* gene of the volunteers was also screened for two polymorphisms associated with normal (17Ser > Thr) and increased (156Asn > Ser) enzyme activity. Although all of the individuals were homozygous for the SNP that results in increased enzyme activity in vitro (156Asn > Ser), excretion of PAHA varied significantly (16–56%, hippurate ratio). The intricacies of PABA metabolism revealed possible limiting factors for the use of this probe substance for the targeted profiling of glycine conjugation [73]. A method to accurately quantify the benzoate detoxification ability of humans in vivo will aid greatly in the understanding of this pathway.

3. Materials and Methods

The reference transcript of *GLYAT* (NM_201648.3; ENST00000344743) was used to determine the allelic variation and haplotype frequencies.

3.1. Missense Variants Identified in GLYAT Using gnomAD

Exome and whole-genome sequencing data in the genome aggregation database v2 (gnomAD browser—gnomad.broadinstitute.org/; accessed on 10 September 2020) [36] were used to analyse the allele frequencies of the missense variants found in the GLYAT gene. The gnomAD v2 data set contains data from 125,748 exomes and 15,708 whole genomes, all mapped to the GRCh37/hg19 reference sequence. Only high quality genotypes were included in this dataset (GQ $\geq$ 20, DP $\geq$ 10, allele balance > 0.2 for heterozygote genotypes).

3.2. Haplotype Data Obtained from Ensembl

The 1000 Genomes [48] haplotype data for GLYAT were downloaded from the Ensembl GRCh38 Homo sapiens assembly [74]. The 1000 genomes dataset contains data for 2504 individuals from 26 populations.

3.3. Tajima's Test of Neutrality

To determine if the GLYAT gene is evolving randomly or under directional selection, the Tajima's *D* test was determined [75–77] using MEGA X [52].

3.4. Phylogenetic Analyses

The amino acid sequences for each GLYAT haplotype, in addition to a selection of orthologs (*Pan troglodytes, Pan paniscus, Pongo abelii, Gorilla gorilla gorilla, Mus musculus, Rattus norvegicus, Bos taurus, Tursiops truncates, Loxodonta africana, Pteropus vampyrus*, and *Danio rerio*), were downloaded from Ensembl [37]. The amino acid sequences were aligned using ClustalX v2.1 [78]. The amino acid substitution model [56] and the phylogenetic analyses were performed using MEGA X [52]. The robustness of the tree topology was evaluated using bootstrap analysis with a resampling size of 500 replicates.

3.5. Expression and Nickel-Affinity Purification of the Recombinant GLYAT Haplotypes

The 156Asn > Ser; 17Ser > Thr,156Asn > Ser and 156Asn > Ser,199Arg > Cys GLYAT recombinant proteins were expressed and purified, as previously reported [31], with 0.5% glycine added to the expression medium in this study. The purified proteins include an *N*-terminal fusion tag of approximately 29 kDa containing the Trx-tag followed by a 6X His-tag. For long-term storage of the purified enzyme preparations at −80 °C, glycerol was added to the purified proteins to a final concentration of 10% and then snap-frozen in liquid nitrogen. Protein concentrations were determined using the Qubit v2.0 Fluorimeter and the Qubit Protein Assay Kit (Thermo Fisher Scientific Inc., Waltham, MA, USA. Enzyme kinetic analyses were performed on stored aliquots from the same batch of purified protein. These aliquots were thawed on ice before use.

3.6. Bi-Substrate (Benzoyl–CoA and Glycine) Kinetic Analysis

To determine the bi-substrate kinetic parameters for each of the three GLYAT haplotype variants, enzyme assays were performed in which both substrate concentrations were varied simultaneously. The glycine concentrations were varied from 1–200 mM and the benzoyl–CoA concentrations from 20–200 µM in various combinations according to a grid. Enzyme activity was determined using a colorimetric assay that measures glycine-dependent release of CoA at 412 nm, in the presence of 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) [79]. Enzyme assays were 200 µL in volume and contained 25 mM Tris-acetate, pH 8.0, 100 µM DTNB, 2 µg of a particular recombinant GLYAT variant, glycine and benzoyl–CoA. Reactions were carried out at 37 °C in 96-well plates and monitored for 20-min at 40 s intervals using a BioTek plate reader and accompanying Gen5 software (BioTek, Winooski, VT, USA). Raw initial rate data for each of the combinations of substrate concentrations are provided in Supplementary Table S3. Activities from triplicate assays were calculated by linear regression of the A_{412} versus time data over the linear range of the time-course and expressed as µmol/min/mg protein.

The data were processed using the Python programming language, making use of Jupyter notebooks [80] (https://jupyter.org, accessed on 14 March 2021) and the numpy [81], scipy (https://www.scipy.org/, accessed on 14 March 2021), and pandas [82] libraries. Raw and processed data were visualised with the matplotlib plotting library [83]. Kinetic parameters were determined by non-linear regression, using the Python lmfit module [84] of the initial rate vs. substrate concentration data to the bi-substrate Hill equation:

$$v = \frac{k_{cat} \cdot e_T \cdot \left(\frac{G}{s_{0.5,g}}\right)^{h_g} \left(\frac{B}{s_{0.5,b}}\right)^{h_b}}{\left(1 + \left(\frac{G}{s_{0.5,g}}\right)^{h_g}\right)\left(1 + \left(\frac{B}{s_{0.5,b}}\right)^{h_b}\right)} \quad (1)$$

where k_{cat} is the catalytic rate constant (turnover number), e_T is the total enzyme concentration, G and B are the concentrations of glycine and benzoyl–CoA, respectively, $s_{0.5,g}$ and $s_{0.5,b}$ are the respective half-saturation constants, and h_g and h_b the respective Hill coefficients for glycine and benzoyl–CoA. Kinetic parameters were estimated from a global fit to all the data (with both substrates varied) simultaneously. Because excessive correlations were observed between the k_{cat} and other parameters, the fit was performed in two stages: initially, the k_{cat} was estimated from a fit of the uni-substrate Hill equation to the rate vs. glycine concentration data collected at the highest benzoyl–CoA concentration (200 µM). This value was subsequently fixed and the remaining parameters were estimated from a global fit of the bi-substrate Hill equation to all the data for both varied substrates.

4. Conclusions

Impaired phase II detoxification has been associated with adverse reactions to pharmaceutical drugs and may be involved in the pathogenesis of complex multifactorial diseases such as cancer [85,86]. In the case of glycine conjugation, it has been shown that GLYAT expression is transcriptionally down-regulated in hepatocellular carcinoma specimens [28].

Very little is still understood about the physiological implications of the impairment of glycine conjugation and the consequences of substrate interaction and competition in this pathway. This is especially relevant because the exposure of humans to benzoate is increasing due to the wide use of benzoate as a preservative [18]. This is reflected in the high levels of hippurate (up to 932.66 µmol/mmol creatinine) found in urine [87,88].

Glycine conjugation can be influenced by several factors, including the availability of ATP, CoA, and glycine, variants in the *ACSM2B* and *GLYAT* genes, and differential expression of ACSM2B and GLYAT. The present study and previous studies have shown that both the *ACSM2* [2] and the *GLYAT* genes [27] are highly conserved and that alleles, predicted to have a deleterious effect on enzyme function, are found at very low frequencies. The glycine conjugation pathway might be essential for life as a metabolic defect related to this pathway has yet to be identified. This hypothesis, however, needs to be tested by establishing a model system in which variants of ACSM2B and GLYAT can be co-expressed to investigate the effect of genetic variation on the detoxification ability and/or toxicity of the pathway. This will provide a better understanding of what the effect of genetic variants is on the in vivo glycine conjugation ability. Development of a model system in which the detoxification ability of the glycine conjugation pathway can be evaluated by measuring both the initial conjugation with CoA and the acyl-transfer to the amino acid will contribute to filling one of the biggest gaps in the literature, whereby previous studies characterised glycine conjugation as a one-step process. This is especially relevant for studies on the pharmacokinetic evaluation of salicylate metabolism.

If it could be established that the glycine conjugation pathway is overloaded by overconsumption of benzoate and salicylate, this could lead to recommendations for healthier choices when consuming food, in addition to a decreased use of benzoate as a food additive. It will then also be possible to provide genetic testing to identify individuals with variants affecting their glycine conjugation ability and to advise individuals on whether they might potentially be at risk from these dietary additives. This is particularly relevant to isovaleric acidaemia patients.

Supplementary Materials: The following are available online at https://www.mdpi.com/1422-0067/22/6/3129/s1.

Author Contributions: Conceptualisation, R.v.d.S.; data curation, J.M.R., C.S. and R.v.d.S.; formal analysis, J.M.R. and R.v.d.S.; funding acquisition, R.v.d.S.; investigation, J.M.R. and C.S.; methodology, J.M.R., C.S. and R.v.d.S.; project administration, R.v.d.S.; resources, R.v.d.S.; software, J.M.R.; supervision, R.v.d.S.; validation, J.M.R. and C.S.; visualisation, J.M.R. and R.v.d.S.; writing—original draft preparation, J.M.R. and R.v.d.S.; writing—review and editing, J.M.R., C.S. and R.v.d.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work is based on the research supported in part by the National Research Foundation of South Africa (NRF) for the grant, Unique Grant No 117890, awarded 7 December 2018. Any opinion, finding, conclusion, or recommendation expressed in this material is that of the authors and the NRF does not accept any liability in this regard. The funding body played no role in this study.

Institutional Review Board Statement: "The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Health Research Ethics Committee (HREC) of North-West University, ethics number NWU-00095-16-A1 (approved 2016-11-24).

Informed Consent Statement: Patient consent was not required as the data used in this study is available on public databases.

Data Availability Statement: The data presented in this study are available in Supplementary Tables S1–S3, and Supplementary Figures S1–S6.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. del Olmo, A.; Calzada, J.; Nunez, M. Benzoic acid and its derivatives as naturally occurring compounds in foods and as additives: Uses, exposure, and controversy. *Crit. Rev. Food Sci. Nutr.* **2017**, *57*, 3084–3103. [CrossRef] [PubMed]
2. 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. Genom. MGG* **2018**, *293*, 1279–1292. [CrossRef] [PubMed]
3. Tanaka, K.; Isselbacher, K.J. The isolation and identification of *N*-isovalerylglycine from urine of patients with isovaleric acidemia. *J. Biol. Chem.* **1967**, *242*, 2966–2972. [CrossRef]
4. Bartlett, K.; Gompertz, D. The specificity of glycine-*N*-acylase and acylglycine excretion in the organicacidaemias. *Biochem. Med.* **1974**, *10*, 15–23. [CrossRef]
5. Killinberg, P.G.; Davidson, E.D.; Webster, L.T. Evidence for a medium-chain fatty acid: Coenzyme A ligase (adenosine monophosphate) that activates salicylate. *Mol. Pharmacol.* **1971**, *7*, 260–268.
6. Knights, K.M. Role of hepatic fatty acid: Coenzyme A ligases in the metabolism of xenobiotic carboxylic acids. *Clin. Exp. Pharmacol. Physiol.* **1998**, *25*, 776–782. [CrossRef]
7. Nandi, D.L.; Lucas, S.V.; Webster, L.T., Jr. Benzoyl-coenzyme A: Glycine *N*-acyltransferase and phenylacetyl-coenzyme A: Glycine *N*-acyltransferase from bovine liver mitochondria. Purification and characterization. *J. Biol. Chem.* **1979**, *254*, 7230–7237. [CrossRef]
8. Schachter, D.; Taggart, J.V. Glycine *N*-acylase: Purification and properties. *J. Biol. Chem.* **1954**, *208*, 263–275. [CrossRef]
9. Rechner, A.R.; Kuhnle, G.; Bremner, P.; Hubbard, G.P.; Moore, K.P.; Rice-Evans, C.A. The metabolic fate of dietary polyphenols in humans. *Free Radic. Biol. Med.* **2002**, *33*, 220–235. [CrossRef]
10. Lemarie, F.; Beauchamp, E.; Legrand, P.; Rioux, V. Revisiting the metabolism and physiological functions of caprylic acid (C8:0) with special focus on ghrelin octanoylation. *Biochimie* **2016**, *120*, 40–48. [CrossRef]
11. Gregus, Z.; Fekete, T.; Varga, F.; Klaassen, C.D. Dependence of glycine conjugation on availability of glycine: Role of the glycine cleavage system. *Xenobiotica Fate Foreign Compd. Biol. Syst.* **1993**, *23*, 141–153. [CrossRef] [PubMed]
12. Gregus, Z.; Fekete, T.; Varga, F.; Klaassen, C.D. Effect of valproic acid on glycine conjugation of benzoic acid. *J. Pharmacol. Exp. Ther.* **1993**, *267*, 1068–1075. [PubMed]
13. Knights, K.M.; Miners, J.O. *Amino Acid Conjugation: A Novel Route of Xenobiotic Carboxylic Acid Metabolism in Man*, 1st ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2012.
14. Knights, K.M.; Sykes, M.J.; Miners, J.O. Amino acid conjugation: Contribution to the metabolism and toxicity of xenobiotic carboxylic acids. *Expert Opin. Drug Metab. Toxicol.* **2007**, *3*, 159–168. [CrossRef] [PubMed]
15. Beyoglu, D.; Idle, J.R. The glycine deportation system and its pharmacological consequences. *Pharmacol. Ther.* **2012**, *135*, 151–167. [CrossRef] [PubMed]
16. Leth, T.; Christensen, T.; Larsen, I.K. Estimated intake of benzoic and sorbic acids in Denmark. *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess.* **2010**, *27*, 783–792. [CrossRef] [PubMed]
17. Mischek, D.; Krapfenbauer-Cermak, C. Exposure assessment of food preservatives (sulphites. benzoic and sorbic acid) in Austria. *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess.* **2012**, *29*, 371–382. [CrossRef]
18. Piper, P.W. Yeast superoxide dismutase mutants reveal a pro-oxidant action of weak organic acid food preservatives. *Free Radic. Biol. Med.* **1999**, *27*, 1219–1227. [CrossRef]
19. Eigenmann, P.A.; Haenggeli, C.A. Food colourings. preservatives, and hyperactivity. *Lancet* **2007**, *370*, 1524–1525. [CrossRef]
20. Piper, J.D.; Piper, P.W. 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. *Compr. Rev. Food Sci. Food Saf.* **2017**, *16*, 868–880. [CrossRef]
21. Badenhorst, C.P.; van der Sluis, R.; Erasmus, E.; van Dijk, A.A. Glycine conjugation: Importance in metabolism. the role of glycine *N*-acyltransferase, and factors that influence interindividual variation. *Expert Opin. Drug Metab. Toxicol.* **2013**, *9*, 1139–1153. [CrossRef] [PubMed]
22. Badenhorst, C.P.; Erasmus, E.; van der Sluis, R.; Nortje, C.; van Dijk, A.A. A new perspective on the importance of glycine conjugation in the metabolism of aromatic acids. *Drug Metab. Rev.* **2014**, *46*, 343–361. [CrossRef]
23. Levy, G.; Yaffe, S.J. Clinical implications of salicylate-induced liver damage. *Am. J. Dis. Child.* **1975**, *129*, 1385–1386. [CrossRef] [PubMed]
24. Zimmerman, H.J. Effects of aspirin and acetaminophen on the liver. *Arch. Intern. Med.* **1981**, *141*, 333–342. [CrossRef] [PubMed]
25. Dercksen, M.; Duran, M.; Ijlst, L.; Mienie, L.J.; Reinecke, C.J.; Ruiter, J.P.; Waterham, H.R.; Wanders, R.J. Clinical variability of isovaleric acidemia in a genetically homogeneous population. *J. Inherit. Metab. Dis.* **2012**, *35*, 1021–1029. [CrossRef] [PubMed]
26. Naglak, M.; Salvo, R.; Madsen, K.; Dembure, P.; Elsas, L. The Treatment of Isovaleric Acidemia with Glycine Supplement. *Pediatric Res.* **1988**, *24*, 9–13. [CrossRef]
27. van der Sluis, R.; Badenhorst, C.P.; Erasmus, E.; van Dyk, E.; van der Westhuizen, F.H.; 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*, 126–134. [CrossRef] [PubMed]
28. 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*, 901–906. [CrossRef]

29. Wen, H.; Yang, H.J.; An, Y.J.; Kim, J.M.; Lee, D.H.; Jin, X.; Park, S.W.; Min, K.J.; Park, S. Enhanced phase II detoxification contributes to beneficial effects of dietary restriction as revealed by multi-platform metabolomics studies. *Mol. Cell. Proteom. MCP* **2013**, *12*, 575–586. [CrossRef]
30. Mawal, Y.R.; Qureshi, I.A. Purification to homogeneity of mitochondrial acyl coa:glycine *N*-acyltransferase from human liver. *Biochem. Biophys. Res. Commun.* **1994**, *205*, 1373–1379. [CrossRef]
31. van der Sluis, R.; Badenhorst, C.P.; van der Westhuizen, F.H.; van Dijk, A.A. Characterisation of the influence of genetic variations on the enzyme activity of a recombinant human glycine *N*-acyltransferase. *Gene* **2013**, *515*, 447–453. [CrossRef]
32. Kelley, M.; Vessey, D. Characterization of the acyl-CoA: Amino acid *N*-acyltransferases from primate liver mitochondria. *J. Biochem. Toxicol.* **1994**, *9*, 153–158. [CrossRef] [PubMed]
33. van der Westhuizen, F.H.; Pretorius, P.J.; Erasmus, E. The utilization of alanine. glutamic acid, and serine as amino acid substrates for glycine *N*-acyltransferase. *J. Biochem. Mol. Toxicol.* **2000**, *14*, 102–109. [CrossRef]
34. Schulke, D.; Sass, J.O. Frequent sequence variants of human glycine *N*-acyltransferase (GLYAT) and inborn errors of metabolism. *Biochimie* **2021**. [CrossRef] [PubMed]
35. van der Sluis, R.; Ungerer, V.; Nortje, C.; van Dijk, A.A.; Erasmus, E. New insights into the catalytic mechanism of human glycine *N*-acyltransferase. *J. Biochem. Mol. Toxicol.* **2017**, *31*, e21963. [CrossRef] [PubMed]
36. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alföldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [CrossRef]
37. Yates, A.; Akanni, W.; Amode, M.R.; Barrell, D.; Billis, K.; Carvalho-Silva, D.; Cummins, C.; Clapham, P.; Fitzgerald, S.; Gil, L.; et al. Ensembl 2016. *Nucleic Acids Res.* **2016**, *44*, D710–D716. [CrossRef] [PubMed]
38. Ramsay, M. Africa: Continent of genome contrasts with implications for biomedical research and health. *FEBS Lett.* **2012**, *586*, 2813–2819. [CrossRef]
39. Altshuler, D.M.; Gibbs, R.A.; Peltonen, L.; Altshuler, D.M.; Gibbs, R.A.; Peltonen, L.; Dermitzakis, E.; Schaffner, S.F.; Yu, F.; Peltonen, L.; et al. Integrating common and rare genetic variation in diverse human populations. *Nature* **2010**, *467*, 52–58.
40. Buchanan, C.C.; Torstenson, E.S.; Bush, W.S.; Ritchie, M.D. A comparison of cataloged variation between International HapMap Consortium and 1000 Genomes Project data. *J. Am. Med. Inform. Assoc. JAMIA* **2012**, *19*, 289–294. [CrossRef]
41. Retshabile, G.; Mlotshwa, B.C.; Williams, L.; Mwesigwa, S.; Mboowa, G.; Huang, Z.; Rustagi, N.; Swaminathan, S.; Katagirya, E.; Kyobe, S.; et al. Whole-Exome Sequencing Reveals Uncaptured Variation and Distinct Ancestry in the Southern African Population of Botswana. *Am. J. Hum. Genet.* **2018**, *102*, 731–743. [CrossRef]
42. Vatsiou, A.I.; Bazin, E.; Gaggiotti, O.E. Changes in selective pressures associated with human population expansion may explain metabolic and immune related pathways enriched for signatures of positive selection. *BMC Genom.* **2016**, *17*, 504. [CrossRef]
43. Sabbagh, A.; Darlu, P.; Crouau-Roy, B.; Poloni, E.S. Arylamine *N*-acetyltransferase 2 (NAT2) genetic diversity and traditional subsistence: A worldwide population survey. *PLoS ONE* **2011**, *6*, e18507. [CrossRef]
44. Millburn, P. Factors Affecting Glucuronidation in vivo. *Biochem. Soc. Trans.* **1974**, *2*, 1182–1186. [CrossRef]
45. Capel, I.D.; French, M.R.; Millburn, P.; Smith, R.L.; Williams, R.T. The fate of (14C) phenol in various species. *Xenobiotica* **1972**, *2*, 25–34. [CrossRef]
46. Court, M.H.; Greenblatt, D.J. Molecular genetic basis for deficient acetaminophen glucuronidation by cats: UGT1A6 is a pseudogene. and evidence for reduced diversity of expressed hepatic UGT1A isoforms. *Pharmacogenetics* **2000**, *10*, 355–369. [CrossRef]
47. Court, M.H.; Greenblatt, D.J. Molecular basis for deficient acetaminophen glucuronidation in cats. An interspecies comparison of enzyme kinetics in liver microsomes. *Biochem. Pharmacol.* **1997**, *53*, 1041–1047. [CrossRef]
48. Auton, A.; Brooks, L.D.; Durbin, R.M.; Garrison, E.P.; Kang, H.M.; Korbel, J.O.; Marchini, J.L.; McCarthy, S.; McVean, G.A.; Abecasis, G.R. A global reference for human genetic variation. *Nature* **2015**, *526*, 68–74.
49. Cardenas, C.L.L.; Bourgine, J.; Cauffiez, C.; Allorge, D.; Lo-Guidice, J.M.; Broly, F.; Chevalier, D. Genetic polymorphisms of Glycine *N*-acyltransferase (GLYAT) in a French Caucasian population. *Xenobiotica* **2010**, *40*, 853–861. [CrossRef] [PubMed]
50. Sim, N.L.; Kumar, P.; Hu, J.; Henikoff, S.; Schneider, G.; Ng, P.C. SIFT web server: Predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* **2012**, *40*, W452–W457. [CrossRef] [PubMed]
51. Adzhubei, I.A.; Schmidt, S.; Peshkin, L.; Ramensky, V.E.; Gerasimova, A.; Bork, P.; Kondrashov, A.S.; Sunyaev, S.R. A method and server for predicting damaging missense mutations. *Nat. Methods* **2010**, *7*, 248–249. [CrossRef] [PubMed]
52. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef] [PubMed]
53. Sibley, C.G.; Ahlquist, J.E. The phylogeny of the hominoid primates. as indicated by DNA-DNA hybridization. *J. Mol. Evol.* **1984**, *20*, 2–15. [CrossRef]
54. Lovejoy, C.O. The origin of man. *Science* **1981**, *211*, 341–350. [CrossRef]
55. McGrew, W.C. In search of the last common ancestor: New findings on wild chimpanzees. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **2010**, *365*, 3267–3276. [CrossRef]
56. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood. evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739. [CrossRef]
57. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci. CABIOS* **1992**, *8*, 275–282. [CrossRef]

58. Taylor, D.J.; Piel, W.H. An assessment of accuracy. error, and conflict with support values from genome-scale phylogenetic data. *Mol. Biol. Evol.* **2004**, *21*, 1534–1537. [CrossRef]
59. Felsenstein, J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evol. Int. J. Org. Evol.* **1985**, *39*, 783–791. [CrossRef] [PubMed]
60. Dyda, F.; Klein, D.C.; Hickman, A.B. GCN5-related *N*-acetyltransferases: A structural overview. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 81–103. [CrossRef] [PubMed]
61. Vetting, M.W.; de Carvalho, L.P.S.; Yu, M.; Hegde, S.S.; Magnet, S.; Roderick, S.L.; Blanchard, J.S. Structure and functions of the GNAT superfamily of acetyltransferases. *Arch. Biochem. Biophys.* **2005**, *433*, 212–226. [CrossRef]
62. Ferdinand, W. The interpretation of non-hyperbolic rate curves for two-substrate enzymes. A possible mechanism for phospho-fructokinase. *Biochem. J.* **1966**, *98*, 278–283. [CrossRef]
63. Whittington, A.C.; Larion, M.; Bowler, J.M.; Ramsey, K.M.; Brüschweiler, R.; Miller, B.G. Dual allosteric activation mechanisms in monomeric human glucokinase. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11553–11558. [CrossRef]
64. Hilser, V.J.; Anderson, J.A.; Motlagh, H.N. Allostery vs. "allokairy". *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11430–11431. [CrossRef]
65. Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; Hadler, G.L., Ed.; W. H. Freeman and Company: New York, NY, USA, 1999.
66. Ebrecht, A.C.; Solamen, L.; Hill, B.L.; Iglesias, A.A.; Olsen, K.W.; Ballicora, M.A. Allosteric Control of Substrate Specificity of the Escherichia coli ADP-Glucose Pyrophosphorylase. *Front. Chem.* **2017**, *5*, 41. [CrossRef]
67. Cornish-Bowden, A. Enzyme specificity: Its meaning in the general case. *J. Theor. Biol.* **1984**, *108*, 451–457. [CrossRef]
68. Cornish-Bowden, A.; Cárdenas, M.L. Specificity of non-Michaelis-Menten enzymes: Necessary information for analyzing metabolic pathways. *J. Phys. Chem. B* **2010**, *114*, 16209–16213. [CrossRef] [PubMed]
69. Quick, A.J. The conjugation of benzoic acid in man. *J. Biol. Chem.* **1931**, *92*, 65–85. [CrossRef]
70. Levy, G. Pharmacokinetics of salicylate elimination in man. *J. Pharm. Sci.* **1965**, *54*, 959–967. [CrossRef] [PubMed]
71. Schachter, D. The chemical estimation of acyl glucuronides and its application to studies on the metabolism of benzoate and salicylate in man. *J. Clin. Investig.* **1957**, *36*, 297–302. [CrossRef] [PubMed]
72. Campbell, L.; Wilson, H.K.; Samuel, A.M.; Gompertz, D. Interactions of m-xylene and aspirin metabolism in man. *Br. J. Ind. Med.* **1988**, *45*, 127–132. [CrossRef] [PubMed]
73. Nortje, C.; van der Sluis, R.; van Dijk, A.A.; Erasmus, E. The Use of *p*-Aminobenzoic Acid as a Probe Substance for the Targeted Profiling of Glycine Conjugation. *J. Biochem. Mol. Toxicol.* **2016**, *30*, 136–147. [CrossRef]
74. Zerbino, D.R.; Achuthan, P.; Akanni, W.; Amode, M.R.; Barrell, D.; Bhai, J.; Billis, K.; Cummins, C.; Gall, A.; Girón, C.G.; et al. Ensembl 2018. *Nucleic Acids Res.* **2018**, *46*, D754–D761. [CrossRef]
75. Simonsen, K.L.; Churchill, G.A.; Aquadro, C.F. Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* **1995**, *141*, 413–429. [CrossRef]
76. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **1989**, *123*, 585–595. [CrossRef] [PubMed]
77. Fu, Y.X.; Li, W.H. Statistical tests of neutrality of mutations. *Genetics* **1993**, *133*, 693–709. [CrossRef] [PubMed]
78. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [CrossRef]
79. Kolvraa, 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. [CrossRef]
80. Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B.E.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J.B.; Grout, J.; Corlay, S.; et al. Jupyter Notebooks—A publishing format for reproducible computational workflows. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas*; Loizides, F., Scmidt, B., Eds.; IOS Press: Amsterdam, The Netherlands, 2016; pp. 87–90.
81. van der Walt, S.; Colbert, S.C.; Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Comput. Sci. Eng.* **2011**, *13*, 22–30. [CrossRef]
82. McKinney, W. Data Structures for Statistical Computing in Python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010.
83. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]
84. Newville, M.; Otten, R.; Nelson, A.; Ingargiola, A.; Stensitzki, T.; Allan, D.; Fox, A.; Carter, F.; Michał; Pustakhod, D.; et al. lmfit/lmfit-py 1.0.2 (Version 1.0.2).Zenodo. 2021. Available online: https://zenodo.org/record/4516651 (accessed on 1 January 2020).
85. Nebert, D.; McKinnon, R.; Puga, A. Human Drug-Metabolizing Enzyme Polymorphisms: Effects on Risk of Toxicity and Cancer. *DNA Cell Biol.* **1996**, *15*, 273–280. [CrossRef]
86. Wallig, M.A. Glucuronidation and susceptibility to chemical carcinogenesis. *Toxicol. Sci.* **2004**, *78*. [CrossRef] [PubMed]
87. Lees, H.J.; Swann, J.R.; Wilson, I.D.; Nicholson, J.K.; Holmes, E. Hippurate: The Natural History of a Mammalian-Microbial Cometabolite. *J. Proteome Res.* **2013**, *12*, 1527–1546. [CrossRef] [PubMed]
88. Tfouni, S.A.; Toledo, M.C. Estimates of the mean per capita daily intake of benzoic and sorbic acids in Brazil. *Food Addit. Contam.* **2002**, *19*, 647–654. [CrossRef] [PubMed]