Chapter 3: Elucidation of the catalytic mechanism of bovine GLYAT

3.1 Introduction

In this chapter the catalytic mechanism of bovine GLYAT, which is likely to be representative of the GLYAT family of enzymes, was investigated using the recombinant bovine GLYAT expression system described in Chapter 2. Knowledge of the enzyme's catalytic mechanism is important for several reasons. Understanding the mechanism may enable the engineering of enzymes with higher catalytic efficiencies and altered substrate specificities. The observation that there is inter-species variation in GLYAT activity suggests that it may be possible to adjust the enzyme's activity up or down along a smooth continuum by systematically manipulating the appropriate amino acid residues. Furthermore, the catalytic machinery is by definition proximal to the binding site, together forming the active site. Understanding more about the catalytic mechanism, and the parts of the protein sequence responsible for it, can thus benefit our understanding of the substrate binding mechanism of GLYAT, which is one of our long term goals.

The approach was to use biochemical arguments to propose a catalytic mechanism, after which a putative catalytic residue was identified using bioinformatics, and experimentally analysed using site-directed mutagenesis.

3.1.1 Principles for investigation of catalytic mechanisms

In this section the available literature is used to predict the kind of catalytic mechanism that might be employed by GLYAT, based on biochemical arguments and the conservation of catalytic mechanism in the GNAT superfamily of enzymes, of which GLYAT is a member (ENSEMBL, January 2009)).

The properties of an enzyme are ultimately dependent on its three dimensional structure, which is directly related to its amino acid sequence or primary structure. These properties, including the catalytic mechanism, should then in principle be deducible from the amino acid sequence of an enzyme, provided enough information is available. The information that may be relevant to prediction of a catalytic mechanism includes biochemical data, structural data and bioinformatic comparisons. A catalytic mechanism, after being proposed based on the available information, must subsequently be experimentally confirmed or rejected. Complicated biochemical studies may yield more insight into the reaction mechanism itself, while site-directed mutagenesis can be used to determine which amino acid residues are involved in the mechanism.

In predicting a reaction mechanism, the factors to consider are diverse, and vary from protein to protein. Over time, some excellent guidelines have been established to aid the prediction of catalytic mechanisms (Palmer, 2001). The reaction kinetics and pH dependence of a reaction can give immediate insight into the mechanism employed. Inhibition by substrates, products and other inhibitors can reveal information about the sequence of events in the reaction mechanism, and even about which residues may be present in the active site. Other information, including reaction conditions, the number of enzyme subunits and, most importantly, the mechanisms of related proteins, can also be used in proposing a catalytic mechanism for an enzyme.

Since evolution is a "tinkerer and not an inventor," a new protein function will usually be adapted from mechanisms that are already in available (Khersonsky et al., 2006, Todd et al., 2001). Therefore, families and even superfamilies of enzymes often share common catalytic mechanisms and means of substrate binding. Thus, by looking at the known mechanisms of enzymes related to the enzyme in question (based on sequence homology), one can form a clear picture of the kind of mechanism that may be employed by the enzyme under investigation. For this reason, the reactions and catalytic mechanisms of the GNAT superfamily of N-acyltransferases was briefly reviewed in Section 1.5.

To summarise, the GNAT superfamily of acyltransferases employs a conserved mode of substrate binding and a conserved catalytic strategy (Dyda et al., 2000, Vetting et al.,

2005). Binding of the acyl coenzyme A substrate usually precedes binding of the amine substrate, after which direct acyl transfer takes place. It appears that some form of general acid-base catalysis is usually involved. General base catalysts (aspartate, glutamate, and histidine residues) are used to remove protons from the amine, and general acid catalysts (tyrosine and cysteine residues) are used to donate protons to leaving thiolate ions. The general mechanism for the superfamily is depicted in Figure 1.18, and is central to the discussions in this chapter.

3.2 Molecular biology and biochemistry relevant to the GLYAT reaction mechanism

Based on the conservation of catalytic mechanism in the GNAT superfamily, it appears that GLYAT might also have a compulsory order ternary complex mechanism with direct acyl transfer and some form of general acid-base catalysis. In the next few sections the available biochemical data is briefly discussed, to see whether this proposal can be substantiated or not. In what follows, the terminology relevant directly to GLYAT will be used. Thus the "amine or acyl acceptor" will be referred to as glycine and the "acyl donor substrate" will be referred to as benzoyl coenzyme A.

3.2.1 Reaction kinetics

Secondary or double reciprocal plots are used to deduce kinetic parameters of enzymes from reaction velocities. For two-substrate reactions, the pattern of these plots can reveal important information about the reaction mechanism. If the plots of 1/Vo against 1/[S1] for a number of [S2] values are parallel, it implies that a ping-pong mechanism is employed. However, when the lines are not parallel, but converge at some point to the left of the vertical (y) axis when extrapolated, a ternary complex mechanism is implicated (Palmer, 2001). The secondary kinetic plots of the human and bovine GLYAT enzymes indicate that a compulsory order ternary complex mechanism is employed. In accord with the rest of the GNAT superfamily, the benzoyl coenzyme A binds before glycine. Direct acyl transfer takes

place and the products are released as the rate limiting step (Nandi et al., 1979, van der Westhuizen et al., 2000).

3.2.2 pH dependence of the GLYAT reaction

The pH dependence of the bovine and human GLYAT enzymes were reviewed in Section 1.4.2. The activities of both enzymes are low at low pH, but the reaction rate increases rapidly from pH 6 to pH 7.5, until a plateau stretching to beyond pH 10 is reached (Mawal & Qureshi, 1994, Nandi et al., 1979). This indicates that a catalytic base residue is likely involved in catalysis, but not a catalytic acid residue, because the rate would then again decrease at higher pH values. This means that the catalytic residue is likely any one of glutamate, histidine or aspartate.

3.2.3 Inhibition of the GLYAT reaction by divalent cations

GLYAT activity is inhibited by high, non-physiological concentrations of divalent metal ions, including zinc, nickel and magnesium. The inhibitory effect is concentration dependent and is more pronounced for nickel and zinc than for magnesium (Nandi et al., 1979). The inhibitory effect may be due to a change in enzyme structure induced by the high divalent cation concentration. For example, a number of enzymes are known to be allosterically regulated by magnesium ions. Another possibility, supported by the high concentrations required for inhibition, is that the metal ions can form coordinate complexes with groups that are critical to catalysis. Whether the function of the residue is to remove a proton, or to form an electrostatic interaction, formation of such a complex would effectively render it non-functional. The observations that high concentrations of divalent cations are needed for inhibition and that magnesium is a weaker inhibitor than nickel or zinc suggests that the inhibition is of this nature rather than regulatory. It may then be speculated that a residue capable of forming a coordinate bond with a divalent cation is involved in catalysis. This residue can be either one of glutamate, histidine, aspartate or cysteine.

3.2.4 Insensitivity of GLYAT to sulfhydryl reagents

The observation that GLYAT is insensitive to inhibition by several compounds that react with sulfhydryl groups suggests that no cysteine residue is involved in the catalytic mechanism (Nandi et al., 1979). This is further support for a direct transfer mechanism as opposed to an acyl-enzyme intermediate (ping-pong) mechanism.

3.2.5 Experimental approach for elucidation of the GLYAT catalytic mechanism

The above review of the GLYAT reaction suggests that the enzyme employs a compulsory order, ternary complex mechanism with direct acyl transfer, as is known for other GNAT enzymes. It appears that a catalytic base residue is involved in the catalysis, either histidine, glutamate, or aspartate. No general acid catalyst appears to be involved. The remainder of this chapter is concerned with identification of the putative general base residue employed in this mechanism. Bioinformatics was used to predict the functionally important residue and site-directed mutagenesis was used to experimentally assess its catalytic importance. The original bioinformatic analyses were performed on the human and bovine GLYAT sequences, and the experimental analysis was performed using the recombinant bovine GLYAT described in Chapter 2.

3.3 Materials, methods and resources

3.3.1 BLAST searches and ClustalX alignments

The BLAST algorithm (www.ncbi.nlm.nih.gov) was used to gather sequences with significant homology to human GLYAT from the GenBank database. The human GLYAT amino acid sequence (gi 11038137) was submitted, and the default search parameters used. The top 100 sequence homologues identified were downloaded in FASTA format. For multiple alignment of the GLYAT sequences, ClustalX 2.0.10 was used (Larkin et al., 2007). The Gonnet 250 protein alignment matrix and the slow-accurate algorithm were employed, using a gap opening penalty of 10 and a gap extension penalty of 0.1.

3.3.2 GenTHREADER and FUGUE predictions

The human and bovine GLYAT amino acid sequences (gi 11038137 and gi 29135315 were submitted to the GenTHREADER (bioinf.cs.ucl.ac.uk/psipred) and FUGUE (tardis.nibio.go.jp/fugue) servers to search for potential structural homologues (Jones 1999, Shi et al., 2001). The PSIPRED server (bioinf.cs.ucl.ac.uk/psipred) was used t predict the secondary structures of human and bovine GLYAT (McGuffin et al., 2000).

3.3.3 Molecular modelling

Molecular modelling was performed using the alignments generated by GenTHREADE For generation of the protein backbone of the model, MODELLER 9.3 (salilab.org/modelle was used. SCWRL 3.0 (dunbrack.fccc.edu/SCWRL3) was used for modelling of amino ac side chains (Canutescu et al., 2003). Model optimisation was performed using Accelrys D 2.0. Superpositions of protein structures were generated using the MatchMaker algorith of the UCSF Chimera software package (version 1, www.cgl.ucsf.edu/chimera). T software was also used to generate molecular visualisations of the model and structu alignments (Pettersen et al., 2004).

3.3.4 Download of molecular coordinates as PDB files

The protein structures were downloaded as PDB files from the Protein Data Ba (www.rcsb.org) for use in modelling and structural comparisons (Berman et al., 2003). protein structures containing bound coenzyme A or acetyl-coenzyme A, Ligplots w downloaded from the PDBSum web page (www.ebi.ac.uk/pdbsum) for identificatior binding site residues (Laskowski et al., 2005).

3.3.5 Site-directed mutagenesis

The mutants investigated in this study were generated using the megaprimer metho site-directed mutagenesis (Aiyar & Leis, 1993). The desired amino acid chang

converted into the required DNA sequence change and an oligonucleotide designed that contains this altered codon approximately in its middle, with about ten unaltered bases on both sides of the altered base. The primer is then used together with one that binds to the nearest end of the coding sequence in a PCR amplification (mPCR-1). This amplifies a part of the coding sequence while introducing the desired mutation, as depicted in Figure 3.1 (a and b). The resulting amplicon contains the desired mutation, and is then used as a primer in a second amplification (mPCR-2). In this amplification the other outer primer for the coding sequence is used together with the megaprimer, to amplify the whole coding sequence. Because the megaprimer contains the desired mutation, the amplicon produced in mPCR-2 will also contain the desired mutation, as shown in Figure 3.1 (c and d).

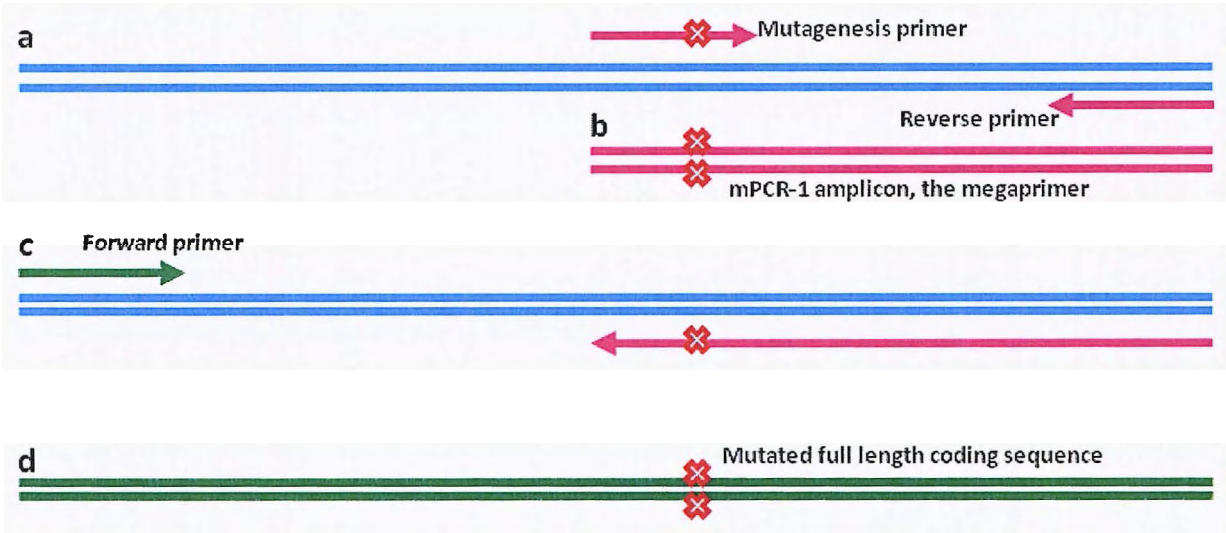

Figure 3.1 Megaprimer site-directed mutagenesis. a) PCR amplification (mPCR-1) using an oligonucleotide primer containing the mutation to be introduced; b) The megaprimer is the mutated amplicon of mPCR-1; c) PCR amplification using the megaprimer (mPCR-2) to generate a full length mutated coding sequence; d) The amplicon generated by mPCR-2 contains the desired mutation, indicated by the crosses.

The first PCR reaction (mPCR-1) was performed using the four mutagenesis primers listed in Table 3.1, and the general bovine GLYAT reverse primer XhoRev (Table 2.2). The primers were designed using DNAMAN version 4.13. The PCR reactions contained 1X Takara ExTaq buffer, 10 nmol of each dNTP, 25 pmol of each primer, approximately 50 ng of template DNA and 2 units of Takara ExTaq polymerase, in a final volume of 50 µl. The cycling conditions were 94 °C for 1 min, then 30 cycles of 94 °C for 30 seconds, 70 °C for

30 seconds, and 72 °C for 1 minute, followed by a final extension at 72 °C for 10 minutes. Thermal cycling was performed using an Eppendorf thermal cycler.

Table 3.1: Oligonucleotide primers used for site-directed mutagenesis

Primer name	Oligonucleotide sequence (5' → 3')	T_m (°C)	Length (bp)	Manufacturer
NdeFor	GCCGCATATGATGTTCCTGCTGC	66	23	Inqaba Biotech
XhoRev	CTTCTCGAGAGGCTCACAGTTCCACTGG	70	28	Inqaba Biotech
E226Q	CCAGACGGGACAGATGCGGATGG	76	23	Inqaba Biotech
E226H	CCAGACGGGACACATGCGGATGG	76	23	Inqaba Biotech
E226D	CCAGACGGGAGACATGCGGATGG	76	23	Inqaba Biotech
E226G	CCAGACGGGAGGCATGCGGATGG	78	23	Inqaba Biotech
CdomFor	GCCGCATATGATGTTCCTGCTGC	66	23	Inqaba Biotech

The amplicons of mPCR-1 were gel purified as described in Section 2.2.7, and used as the megaprimer in mPCR-2. The mPCR-2 reaction mixtures contained 1X Takara PCR buffer (1.5 mM Mg^{2+}), 10 nmol of each dNTP, 25 pmol of primer NdeFor (Table 2.2), approximately 50 ng of template DNA, 2 units of Takara ExTaq polymerase, 3% DMSO, and 250 pmol of the 240 bp megaprimer (a ten fold excess) in a final volume of 50 µl. The cycling conditions were 94 °C for 1 min, then 30 cycles of 94 °C for 45 seconds, 66 °C for 1 minute, and 72 °C for 1 minute, followed by a final extension of 10 minutes at 72 °C. The cooling rate for the annealing step was set to 30%. Thermal cycling was performed using an Eppendorf thermal cycler.

3.3.6 TA cloning of PCR amplicons, plasmid isolation and sequencing

The mPCR-2 amplicons were gel purified as described in Section 2.2.7 and then cloned into the TA cloning vector pTZ57R/T as described in Section 2.2.5. Colony screening for recombinant plasmid, plasmid isolation and plasmid sequencing were performed as described in Sections 2.2.11 to 2.2.14.

3.3.7 Sub-cloning into pColdIII-A expression vector

Mutant GLYAT sequences were excised from the recombinant pTZ57R/T vector, and sub-cloned from the TA cloning vector into pColdIII-A as described in Sections 2.2.8 to 2.2.14.

3.3.8 Protein expression and purification

The mutant proteins were expressed and purified as described in Sections 2.2.15 to 2.2.17.

3.3.9 GLYAT enzyme assays

Mutant enzymes were assayed at pH 8.0 as for the wild type, as described in Section 2.2.20. For determination of the pH dependence of the reaction, reaction buffers of different pH were set up. For pH values between 6.0 and 8.0, a 50 mM potassium phosphate buffer was used. A large volume of master mix containing the glycine, buffer, and DTNB was prepared, so that the pH could be accurately adjusted. The benzoyl-coenzyme A and enzyme were then mixed, placed in the wells of a 96-well plate and the master mix added to initiate the reaction.

3.4 Results and discussion

The overall aim of the work presented in this chapter was to determine the catalytic mechanism of the GLYAT reaction, and to identify the catalytic residues involved. A general catalytic mechanism was proposed, and bioinformatics used to predict a putative catalytic residue. This prediction was then experimentally validated by using site-directed mutagenesis to generate mutant coding sequences which could be expressed and characterised.

3.4.1 Prediction of a catalytic residue

Based on the extensive conservation of structure and reaction mechanism in the GNAT superfamily, and the relatedness of GLYAT to this superfamily, we hypothesised that human and bovine GLYAT might have similar structures and catalytic mechanisms. Molecular modelling of human and bovine GLYAT was thus performed, in an attempt to learn something about the possible structure and reaction mechanism of the enzyme. A molecular model can give insight into the substrate binding sites and can be superimposed with enzymes of known mechanism in order to reveal putative catalytic residues.

3.4.1.1 Molecular modelling

As a first approach, the human and bovine GLYAT protein sequences (gi 11038137 and gi 29135315) were submitted to the SWISS MODEL server to see if a model could be automatically generated. No results were obtained, since there were no structures with significant sequence homology to either human or bovine GLYAT in the Protein Data Bank at the time (January 2009).

Even when no template structure suitable for simple automated modelling is available, it may still be possible to obtain a model, albeit one of lower quality. Therefore, a more advanced search of the Protein Data Bank was performed using the FUGUE and GenTHREADER servers (Jones, 1999, Shi et al., 2001). These servers use more advanced fold recognition algorithms, which enables the identification of very distant homologues. By taking structural data into account and threading the query sequence against the potential homologous structures, an evolutionary relationship can be detected even when most of the sequence similarity has faded away.

Both servers returned similar lists of possible structural homologues from the Protein Data Bank. A sample of the results of the GenTHREADER prediction for human GLYAT are shown in Table 3.2. Most of the homologues identified align to the C-terminal domain of GLYAT, as shown in Figure 3.2. For this reason, the C-terminal halve of the GLYAT sequence was used to initiate another GenTHREADER search, but similar results were obtained (not shown). The best template structure returned by both servers, for both the

human and bovine GLYAT sequences, an uncharacterised protein from Drosophila melanogaster with PDB ID 1SQH (Table 3.2), was chosen for further modelling.

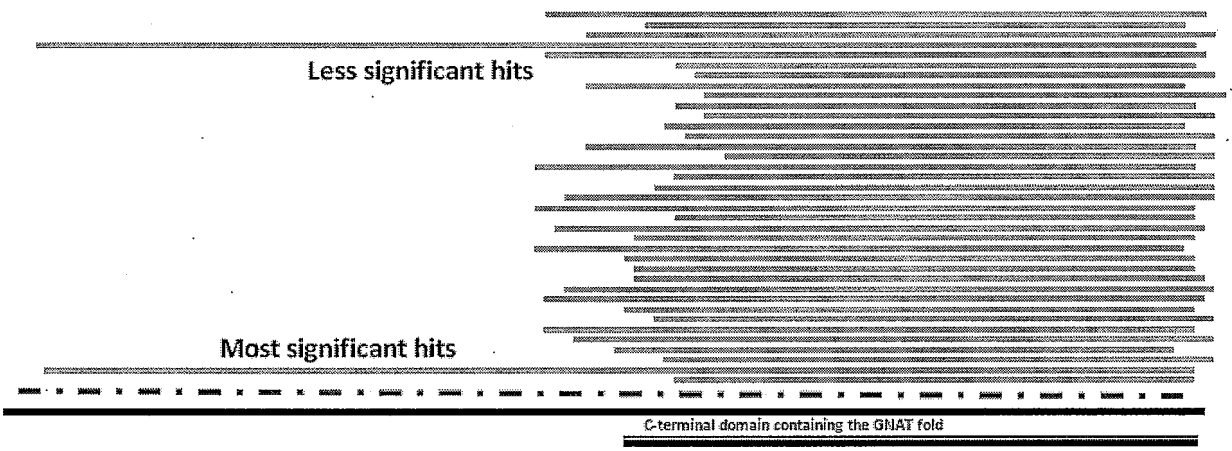

Figure 3.2 Alignment of the GenTHREADER results to human GLYAT. The diagram demonstrates the alignment of the GenTHREADER results to the C-terminal domain of the human GLYAT sequence (indicated by the double lines). The solid black line represents the human GLYAT sequence and the broken line represents the 1SQH sequence used for modelling. The grey lines represent the rest of the GenTHREADER results. The most significant results are represented by the lines closest to the human GLYAT line.

Table 3.2: Results of the GenTHREADER search for human GLYAT

PBB ID	Sequence identity (%)	Alignment score	Short description of structure
1SQH	11.5	57.972	Hypothetical protein from *D. melanogaster*
1MK4	7.6	56.862	Putative acetyltransferase from *B. subtilis*
1P0H	6.6	54.713	Mycothiol synthase from *M. tuberculosis**
2Q4V	8.3	54.609	Human diamine N-acetyltransferase*
1QSM	9.3	54.592	Histone acetyltransferase from *S. cerevisiae**
1Z4R	6.7	51.016	Human GCN5 histone acetyltransferase*
1CJW	8.4	48.151	Human serotonin N-acetyltransferase*

* Structures with coenzyme A, acetyl-coenzyme A or a bisubstrate analogue.

A model can only be as good as the template structure used, and only so if the alignment between template and query sequences is good. To check the validity of the alignments

generated by GenTHREADER, few simple analyses were performed on a multiple alignment between GLYAT and a few of the template structures. The multiple alignments were constructed by manually aligning the template structures to each other, based on their individual GenTHREADER alignments to GLYAT.Only the analyses using the human GLYAT sequence are shown here, but those for the bovine GLYAT were similar.

Three rudimentary analyses were performed. In the first, the three best template structures (Table 3.2) were structurally superimposed using the MatchMaker algorithm of the UCSF Chimera software package. The residues that were used in the final iterations of the structural superposition (interpreted as the least structurally divergent residues) were identified and coloured on the multiple alignment. As shown in Figure 3.3, these residues align well, suggesting that this stretch of amino acids is functionally significant. This was confirmed by the next two analyses.

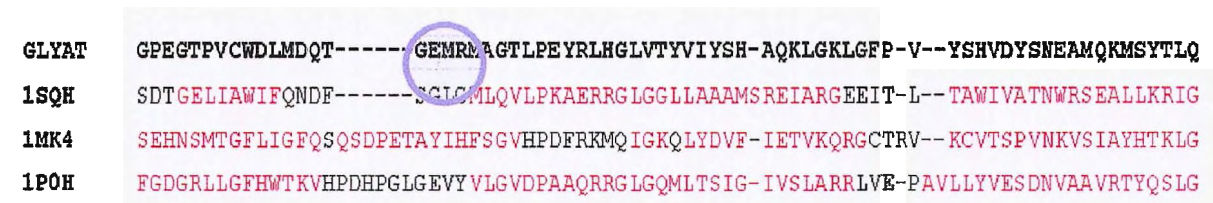

Figure 3.3 Analysis 1 of the GenTHREADER alignments. The diagram indicates the alignment of the sequences discussed in the text. Residues used in the final iteration of a structural alignment generated by the Chimera MatchMaker algorithm are coloured in red. The glutamate residue in the circle is the proposed catalytic glutamate residue.

The second analysis was comparison of secondary structure. The secondary structure of GLYAT was predicted using the PSIPRED server (McGuffin et al., 2000). The same multiple alignment as in the above analysis was used, and the secondary structure elements coloured (secondary structure assignments were performed using UCSF Chimera). The result is shown in Figure 3.4. The secondary structure elements of the three templates align well and correspond to the predicted secondary structure of GLYAT.

The third and most informative analysis involved identification of residues that participate in substrate binding. Ligplots, which indicate contacts between protein and substrate, were obtained for five of the template structures that have a molecule of coenzyme A or an analogue bound (Table 3.2). The residues involved in substrate binding were coloured on a

multiple alignment between GLYAT and the five template structures, which was constructed as described for the first analysis. The result is shown in Figure 3.5 and it can be seen that the residues involved in substrate binding align well between the different sequences. The residues coloured in this analysis were not explicitly considered in generating the GenTHREADER alignments, yet they align well in the multiple alignment. This provides strong support for the validity of the alignment and suggests that it is biologically accurate. The 1SQH structure was thus used as a template for generation of a molecular model.

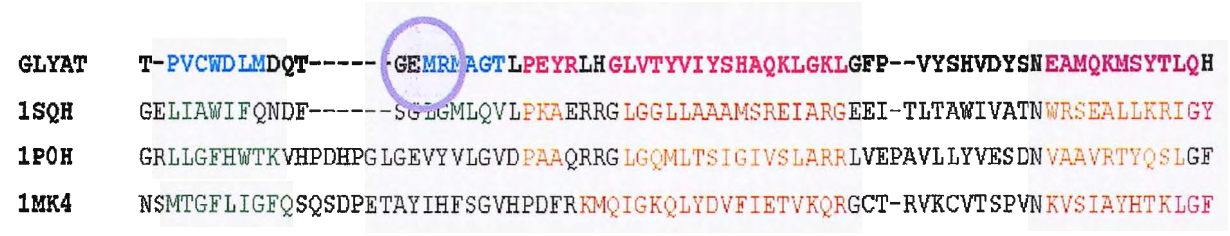

Figure 3.4 Analysis 2 of the GenTHREADER alignments. The diagram demonstrates the alignment of secondary structure elements of the GNAT enzymes discussed in the text. The predicted α-helices of GLYAT are coloured pink. The α-helices of the other enzymes are coloured orange. The predicted β-strands of the GLYAT sequence are coloured blue, and that of the other enzymes are coloured green. The glutamate residue in the circle is the proposed catalytic glutamate residue.

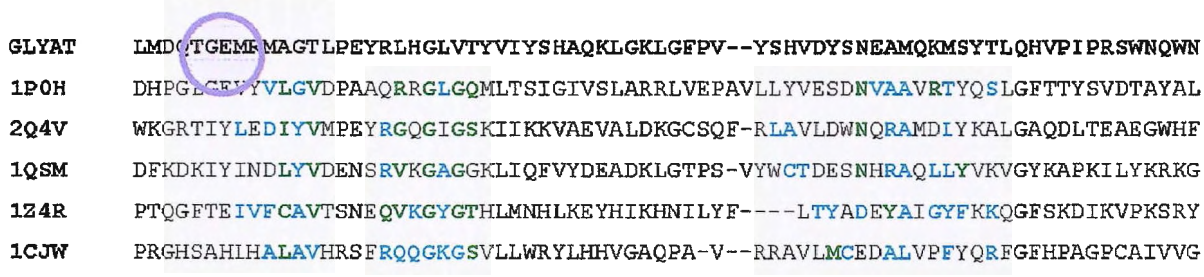

Figure 3.5 Analysis 3 of the GenTHREADER alignments. The diagram demonstrates the alignment of residues involved in substrate binding by several GNAT enzymes. The residues that are directly involved in forming coenzyme A binding sites on the enzymes are coloured in blue. The residues that make direct contact with the substrates are coloured in green. The glutamate residue in the circle is the proposed catalytic glutamate residue.

MODELLER v9.3 was used to generate models of human and bovine GLYAT. The template structure input was an atom file of template 1SQH (generated by manually deleting non protein atoms from the PDB file). The alignment input was the GenTHREADER alignment of the GLYAT sequence to the template sequence. After generation of the model backbone, SCWRL 3.0 was used to model the conformations of amino acid side chains.

Using UCSF Chimera, a molecule of coenzyme A was then superimposed on the models, by aligning the models with template structure 1P0H, which has a molecule of coenzyme A bound (Table 3.2). All of this template, except the coenzyme A, was then deleted, leaving only the model with a molecule of coenzyme A. The models were then optimised using the CHARMM force field algorithm of the Accelrys Discovery Studio 2.0 software package.

Due to the low sequence identities of 11.5% between the human GLYAT sequence and that of the template structure 1SQH, and 12.8% between the bovine GLYAT sequence and that of 1SQH, the models cannot be expected to be accurate to atomic resolution but must rather be regarded as a crude sketch of the GLYAT structure. For this reason the WHAT IF analyses (a set of tests to analyse the technical quality of molecular models) will not be discussed here. The alignment can simply not be good enough to be interested in such fine details (Rost, 1999, Vriend, 1990).

3.4.1.2 Identification of a putative catalytic residue by investigation of the model

Based on the considerable structural conservation in the GNAT superfamily, as discussed in Section 1.5.1 and demonstrated in Figure 1.11, it was thought that the GLYAT model could represent the actual GLYAT structure. The human and bovine GLYAT models were superimposed with the structures of other GNAT enzymes for which the catalytic mechanisms are known, including SNAT, SSAT, and Esa1 (Section 1.5.2). Figure 3.6 shows the bovine GLYAT model in superposition with the SNAT, SSAT and Esa1 enzymes, indicating catalytically important residues (He et al., 2003, Hegde et al., 2007, Scheibner et al., 2002).

Interestingly, the corresponding position on the GLYAT model is occupied by a glutamate residue, E226. This glutamate, as discussed in Section 3.1, is a suitable candidate for the role of catalytic base residue. To confirm this result, several GNAT enzymes for which the mechanisms are known were subsequently aligned to human and bovine GLYAT, as described in Section 3.4.1.1. On the alignment the catalytically important residues were coloured, and the result for bovine GLYAT is shown in Figure 3.7. The catalytic residues

aligned well, which substantiated the proposal that theresidue of bovine GLYAT could be catalytically significant. This was, however, not enough evidence to conclude that the identified residue is indeed catalytically important, as the low sequence identity limited the reliability of the alignments.

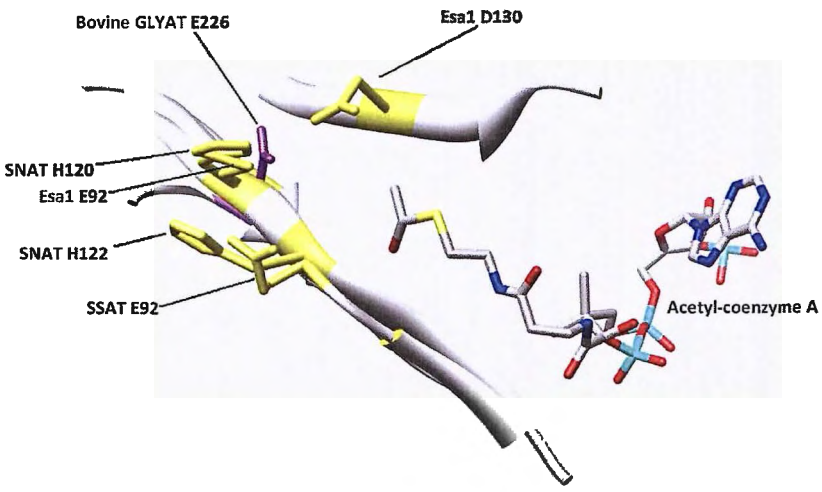

Figure 3.6 Structural alignment of the catalytic residues of three GNAT enzymes identifies a putative catalytic glutamate residue on the bovine GLYAT model. The figure demonstrates part of a structural superposition of the SNAT, SSAT and Esa1 structures with the bovine GLYAT model. The catalytic residues of the three GNAT enzymes are coloured in yellow. The purple residue is a putative catalytic glutamate residue on the bovine GLYAT model. Image generated using UCSF Chimera.

```
Bovine _____----------EMRMAGTLPEY----RLHGLVTYVIYSHAQKLGKLGFP-V--YS
2Q4V_____----------YLEDIYVMPEY----RGQGIGSKIIKKVAEVALDKGCSQF--RL
1CJW_____-ALHRPRGHSAHLHALAVHRSF----RQQGKGSVLLWRYLHHVGAQPA--V--RR
1P0H_____----------EVYVLGVDPAA----QRRGLGQMLTSIGIVSLARRLVE-PAVLL
2FIW_____----------HIDMLYVHPDY----VGRDVGTTLIDALEKLAGARGAL----IL
2Z4R_____----------EIVFCAVTSNE----QVKGYGTHLMNHLKEYHIKHNILYF--LT
1GHE_____----------EVQKLMVLPSA----RGRGLGRQLMDEVEQVAVKHKRGLL--HL
1RO5_____ ----------ELSRFAINSGQKGSLGFSDCTLEAMRALARYSLQNDIQTL--VT
Bovine _____----------EMRMAGTLPEY----RLHGLVTYVIYSHAQKLGKLGFP-V--YS
```

Figure 3.7 GNAT sequences showing alignment of multiple catalytic residues. Multiple alignment of several sequences of GNAT enzymes for which catalytic mechanisms have been reported to bovine GLYAT. The catalytic residues are coloured red, and the alignments were constructed as described in Section 3.4.1.1. The PDB identification codes of the proteins used are shown to the left. 2Q4V is human SSAT, 1CJW is human SNAT, 1P0H is mycothiol synthase from *M. tuberculosis,* 2FIW is cationic trypsin, 2Z4R is human GCN5, 1GHE is the tabtoxin resistance protein Esa1 and 1RO5 is acyl-homoserine lactone synthase.

3.4.1.3 Conservation of the E226 residue of bovine GLYAT in other GLYAT sequences

If the glutamic acid residue identified by the modelling is indeed catalytically important, it is to be expected that the residue should be well conserved in the GLYAT family of enzymes. If not completely conserved, then at least the variations should be conservative, changing the glutamate for an aspartate or perhaps histidine, which would be capable of performing the same function. Unfortunately, this argument has a weak point, as the assumption must be made that all the available GLYAT sequences represent active enzymes (and not pseudo genes, for example). Further, the assumption that GLYAT is a crucial enzyme has to be made, as a sequence that is not crucial has no reason to be conserved. These assumptions cannot be assessed at present, but must be kept in mind when analysing the multiple alignments.

The bovine GLYAT sequence was used to initiate a BLAST search of the non redundant protein sequence library in GenBank. The top 100 homologous sequences identified by this BLAST search were downloaded in FASTA format and aligned using ClustalX (Larkin et al., 2007). As demonstrated in Figure 3.8, the glutamic acid residue in question is conserved in 38 of the top 40 results. The two sequences in which the residue is not present are the b-isoforms of human and chimpanzee GLYAT, which completely lack the C-terminal domain where this residue is found, and can thus be disregarded. This conservation is good, as would be expected of a catalytically important residue.

3.4.1.4 The catalytic function is likely located to the C-terminal domain of GLYAT

The C-terminal domain of GLYAT is thought to be responsible for substrate binding and catalysis. This is based on sequence homology (ENSEMBL, January 2009) of this part of the sequence to the GNAT superfamily. This proposal is more graphically demonstrated by looking at the GenTHREADER results obtained in Section 3.4.1.1. The majority of potential structural homologues obtained are members of the GNAT superfamily of enzymes. Most of these proteins are about half the size of GLYAT, between about 140 and 160 amino acids in length. All of these proteins align with the C-terminal domain of GLYAT, as shown in Figure 3.2. This is very strong evidence that the catalytic function resides in C-terminal domain of the protein. The function of the N-terminal domain is not understood at present.

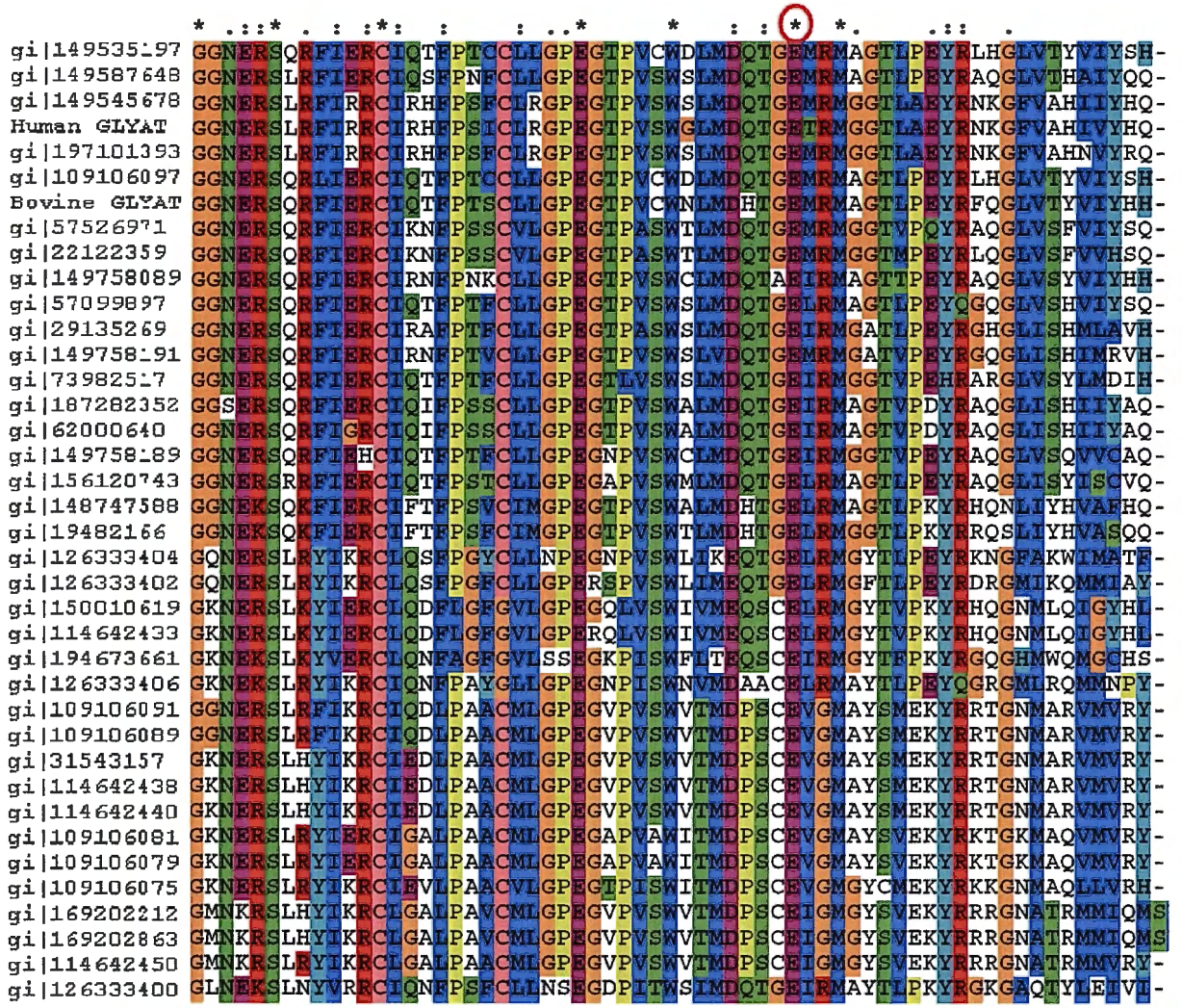

Figure 3.8 Multiple alignment of GLYAT sequences. A multiple alignment of sequences with significant sequence homology to bovine GLYAT, obtained by a BLAST search. Asterisks above the alignment indicate conservation of the residue at that position between all 38 aligned sequences. The circle indicates the proposed catalytic glutamate residue, which is conserved in all the aligned sequences. The GenBank accession numbers are shown to the left of the sequences.

Based on sequence conservation, only three possible candidates for the role of catalytic residue can be found in the C-terminal domain of bovine GLYAT. These are E212, D222 (this position is always aspartate or glutamate, so the residue chemistry is conserved), and E226, the proposed catalytic residue (Figure 3.8). Although it comes close to being a circular argument, the model can be inspected to see which of these residues are located sufficiently close to the active site to act as a catalyst. As shown in Figure 3.9, D222 and E226 are found in such positions. D222 is on another secondary structure element of the

protein, but in the three dimensional space of the model it is in the vicinity of the E226 residue and catalytic residues of other proteins (not shown).

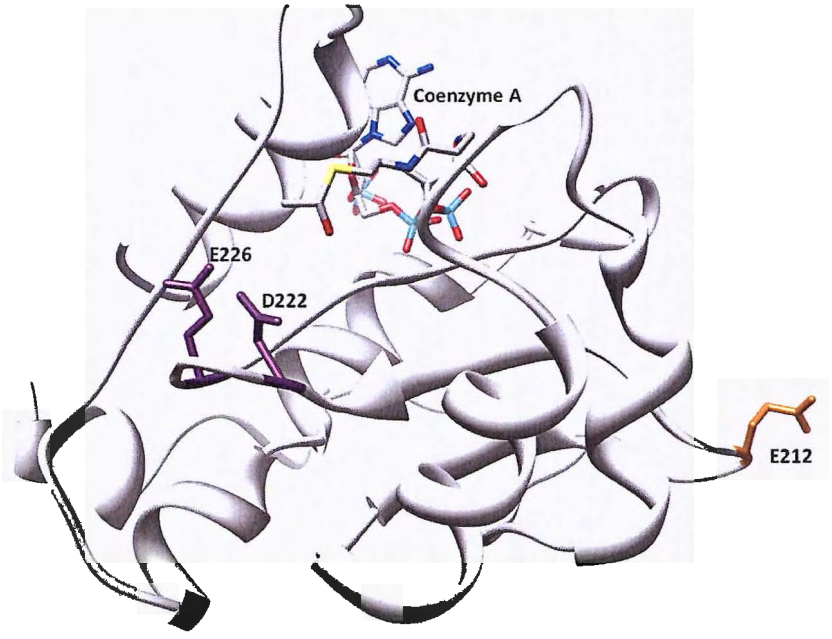

Figure 3.9 Three potentially catalytic residues that are well conserved between GLYAT sequences are shown on the bovine GLYAT model. The figure shows coenzyme A bound to the bovine GLYAT model, with three residues highlighted. The orange residue (E212) is situated on the surface of the model, away from the active site. The purple residues (E222 and E226) are situated close to the coenzyme A thioester bond.

It was thus suggested that E226 of bovine GLYAT could be the catalytic base residue, with D222 being an alternative possibility. Overall, this theoretical argument seemed to be substantial, but in the absence of a crystal structure of a GLYAT enzyme to confirm the proposal, site-directed mutagenesis was used to experimentally assess the validity of the argument. This is presented in the next section.

3.4.2 Experimental investigation of the importance of the bovine GLYAT E226 residue

As discussed in Section 3.2, the catalytic residue identified here is expected to be involved in acid-base catalysis, or more specifically, proton removal in the reaction mechanism. Using this information and following the guidelines discussed in Section 3.1.1, a set of mutants were designed to investigate the proposed reaction mechanism. The first choice

was to mutate the glutamate residue to a glutamine residue (E226Q). This amino acid is structurally similar to glutamate, but differs in that the carboxyl group is replaced by an amide group. This means that the residue will be incapable of performing any function dependent on its negative charge, including proton removal. If this mutant is non-functional, the effect can be attributed largely to the loss of the negative charge on the residue and not to amino acid structure. The second choice was to mutate the glutamate to a histidine residue (E226H), which is structurally and electrically distinct, but capable of proton removal. If this mutant were more active than the E226Q mutant, it would be strong support that the capacity of this residue to remove protons is important for catalysis. Two additional mutations would be to glycine (E226G) and aspartate (E226D). The argument for aspartate was the same as for histidine. Glycine was chosen because it completely removes the side chain (however, it also introduces flexibility which may be undesirable).

Furthermore, it would be interesting to know whether the C-terminal domain of the enzyme, where the catalytic machinery is thought to reside, can function independently of the N-terminal domain. For this reason an N-terminal truncation mutant was designed. By inspecting the secondary structure prediction of bovine GLYAT shown in Figure 3.10, the sequence "KAID" was chosen as the N-terminus of this protein. This includes all the secondary structure elements thought to comprise the GNAT fold (compare Figure 3.10 to Figure 1.12) and an extra few residues at the N-terminus, to ensure that no functional residues are cut off. This protein sequence was cloned and expressed along with the other mutants, as discussed below.

3.4.2.1 Generation of mutant GLYAT coding sequences using site-directed mutagenesis

In order to investigate the importance of the putative catalytic residue proposed, recombinant enzymes in which this residue was mutated had to be biochemically characterised. In order to generate the mutated GLYAT coding sequences, the megaprimer site-directed mutagenesis method described in Section 3.3.5 was used (Aiyar & Leis, 1993).

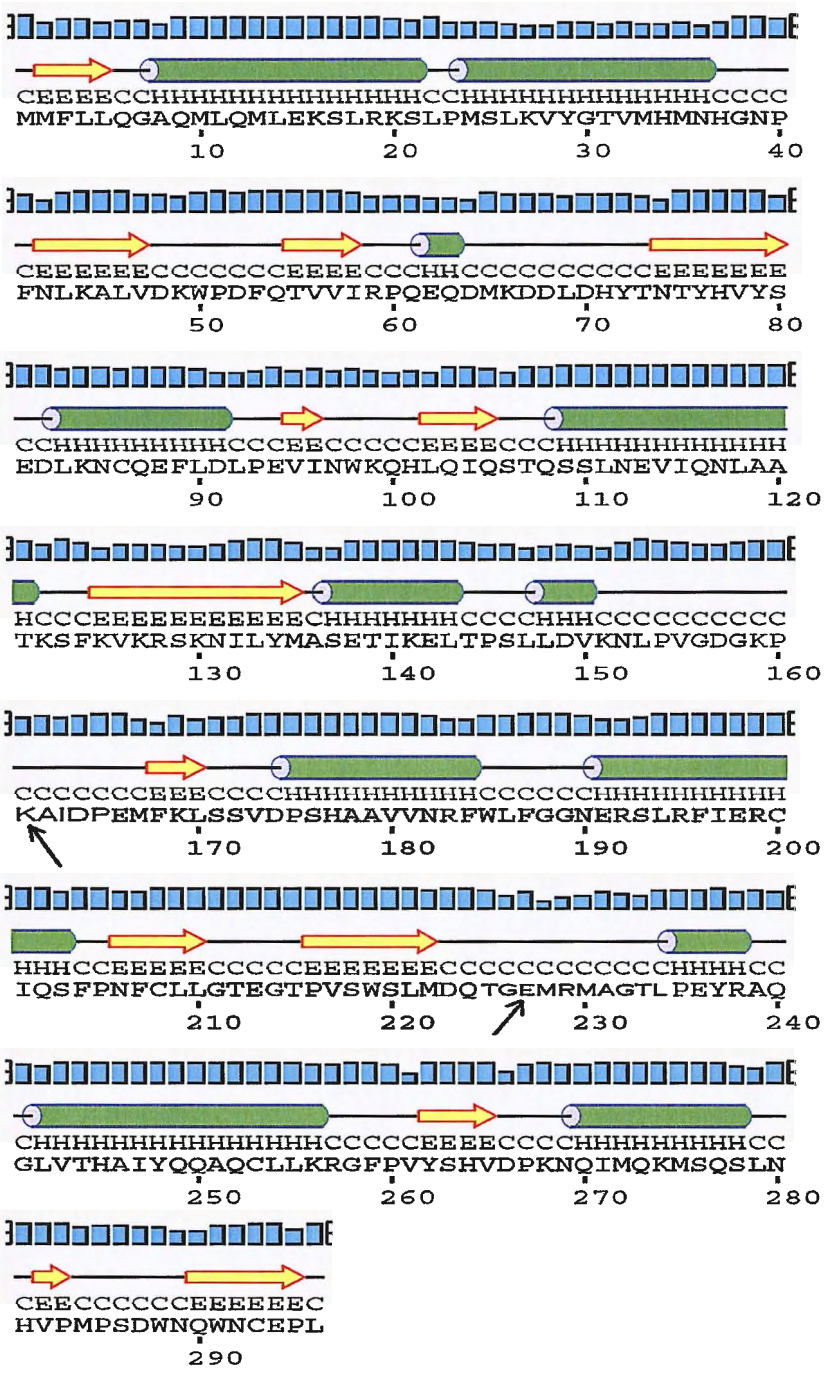

Figure 3.10 Predicted secondary structure of bovine GLYAT. The secondary structure of the bovine GLYAT sequence predicted by PSIPRED. Large arrows indicate β-strands and cylinders indicate α-helices. The bars above the prediction indicate the certainty of the prediction. The first small arrow indicates the lysine residue chosen as the N-terminus of the C-terminal domain mutant of bovine GLYAT. The second small arrow indicates the predicted catalytic glutamate residue.

The mutagenesis PCR-1 (mPCR-1) (Figure 3.1) was straightforward, and the products of amplification are shown on the agarose gel in Figure 3.11. Mutagenesis PCR 2 was more complicated, and some optimisation steps had to be performed in order to find the optimal

conditions described in Section 3.3.5 (results not shown). The amplification product appears on an agarose gel (Figure 3.12) to be approximately 1000 bp in size, using the DNA size ladder for comparison. This is in contrast to the expected amplicon size of 900 bp. As shown in Figure 3.12, the mPCR-2 amplifications were rather specific, with non-specific amplification being only faintly visible on the agarose gel. It can be seen on the agarose gel in Figure 3.12 that all the megaprimer was not consumed in the reaction (240 bp band). Since the amplification appeared to have been rather specific, it was suspected that the 1000 bp band on the gel may not reflect an amplification problem, but perhaps the formation of some unknown complex between the PCR amplicon and the surplus megaprimer. The PCR products were thus gel extracted and sequenced. Although the quality of the sequencing was poor in that the ends of the amplicons were not covered, the mutations had been introduced in the amplicons (results not shown). At this point I decided to use these amplicons to continue with cloning, knowing to return to this point if an error would be encountered later.

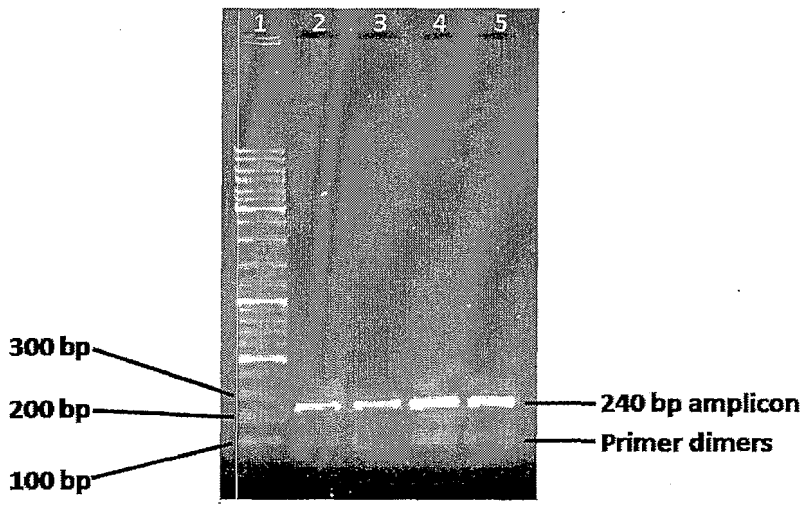

Figure 3.11 Amplicons of the mPCR-1 amplification. Agarose gel electrophoresis showing the amplicons of mPCR-1. Lane 1) 5 µl of O'GeneRuler DNA marker; 2) mPCR-1 for E226Q; 3) mPCR-1 for E226H; 4) mPCR-1 for E226D; and 5) mPCR-1 for E226G.

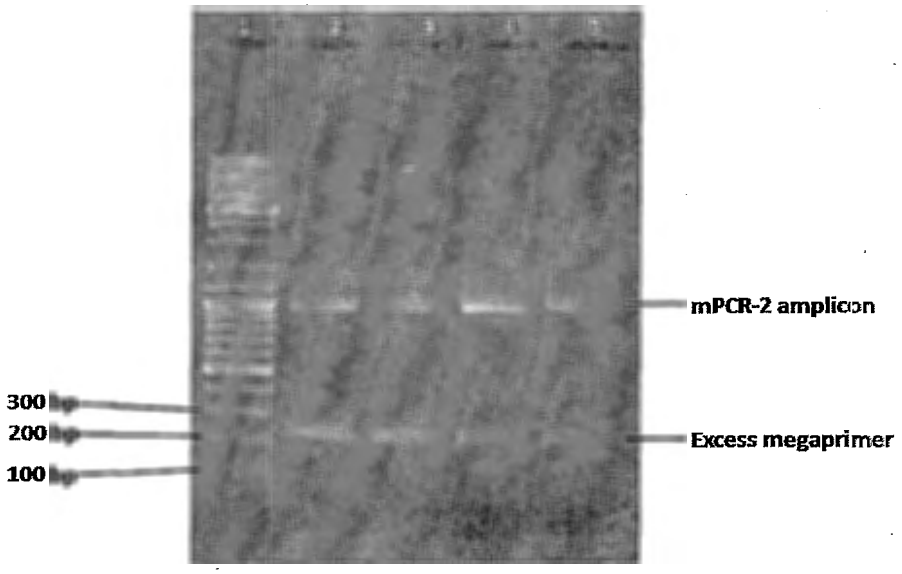

Figure 3.12 Amplicons of the mPCR-2 amplification. The products of the mPCR-2 amplification are shown. Lanes: 1) 5 µl of O'GeneRuler DNA marker; 2) product of the mPCR-2 amplification for the E226Q bovine GLYAT mutant; 3) product of the mPCR-2 amplification for the E226H bovine GLYAT mutant; 4) product of the mPCR-2 amplification for the E226D bovine GLYAT mutant; 5) product of the mPCR-2 amplification for the E226G bovine GLYAT mutant.

3.4.2.2 Cloning the mutant GLYAT amplicons into the pTZ57R/T TA cloning vector

The mPCR-2 amplicons were directly cloned into the TA cloning vector pTZ57R/T, to avoid problems with inefficient restriction enzyme digestion of the PCR products. Transformants were screened for the bovine GLYAT insert, as described in Section 2.2.11. Since no simple strategy for screening to distinguish mutated from wild type constructs was available, the desired mutants had to be identified by sequencing with M13 sequencing primers (Table 2.2).

Mutants E226Q and E226H were successfully generated on first attempt, as shown in Figure 3.13. The E226G and E226D mutations could not be generated. Despite the correct sequence being obtained on sequencing the PCR product, as described above, cloning of the E226D amplicon consistently yielded a sequence that is not mutated. This probably reflects severe cloning bias or mutant instability. The E226G mutation was successfully introduced, but a second mutation, N291S, was also consistently introduced. This mutation is located in the reverse primer sequence. The primer sequence does not contain this mutation, and the same primer (from the same tube) was used to do all other bovine

GLYAT amplifications, none of which contained this mutation. I have no explanation for this observation. The mutagenesis and cloning was repeated, starting with fresh template DNA and primers. Again two colonies of each of the E226D and E226G mutants were sequenced, yielding the same results. Other bacterial cell lines could be used for the cloning, which could solve the problem if mutant instability were the source of trouble. For example, ABLE cells from Stratagene are modified so that high copy number plasmids are propagated in these cells as low copy number plasmids. This reduces the influence of toxic gene products on the cells and stabilises plasmid constructs harbouring such genes. These unexpected results will be investigated further at a later stage. For the purposes of the present study, however, investigation of the E226H and E226Q mutants was sufficient.

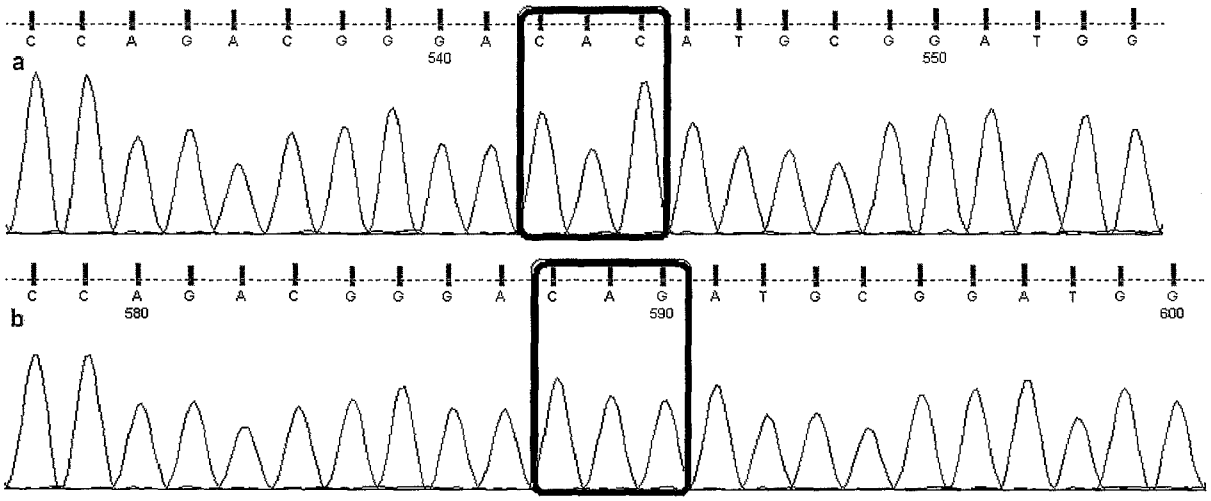

Figure 3.13 Sequencing electrophoretograms for the E226H and E226Q mutants. a) Part of the sequencing electrophoretogram for the pColdIII-E226H construct. The red block indicates the H226 codon. b) Part of the sequencing electrophoretogram for the pColdIII-E226Q construct. The red block indicates the Q226 codon.

3.4.2.3 Cloning of the E226Q, E226H, and C-domain coding sequences into pColdIII-A

The E226Q and E226H mutant coding sequences were excised from the pTZ57R/T vectors by digestion with NdeI and XhoI restriction enzymes, and sub-cloned into pColdIII-A, as described in Sections 2.2.6 to 2.2.14. The integrity of the plasmid constructs was analysed using Sanger sequencing of the entire coding sequences. The parts of the sequencing electrophoretograms corresponding to the E226Q and E226H mutant codons are shown in Figure 3.13 above.

The C-terminal domain mutant was generated as described in Section 3.3.5, using the CdomFor primer listed in Table 2.2. The amplification results are shown in Figure 3.14, which depicts an annealing temperature gradient from 50 to 60 °C. It is clear that the amplification is robust and not very much influenced by annealing temperature. This product was cleaned, digested with NdeI and XhoI, and cloned into pColdIII-A, as described in Sections 2.2.6 to 2.2.14. A positive colony was chosen and the plasmid DNA sequenced to confirm that the C-terminal domain sequence had been cloned without sequence aberrations (sequencing results not shown).

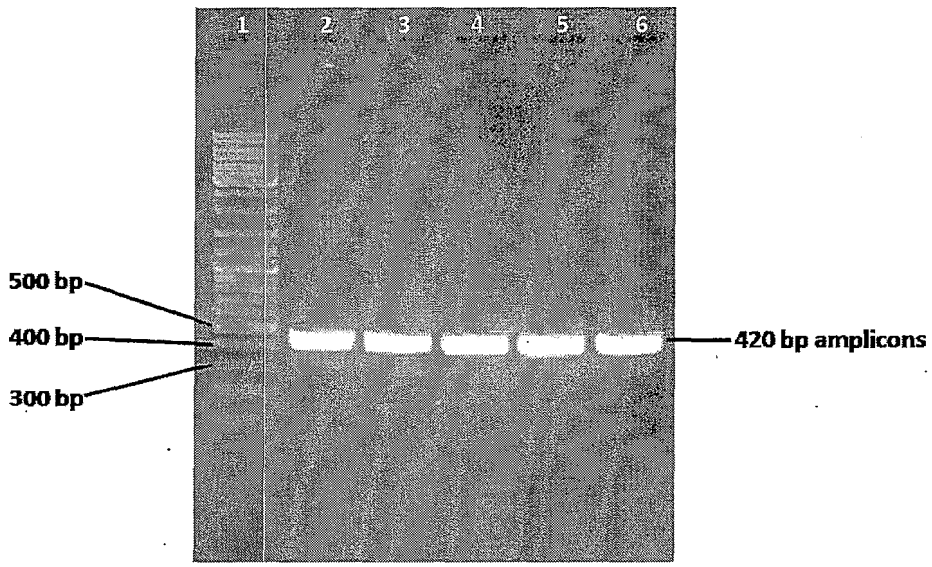

Figure 3.14 Temperature gradient for amplification of the bovine GLYAT C-domain coding sequence. Agarose gel electrophoresis showing the products of a PCR amplification of the coding sequence of the C-terminal domain of GLYAT, using different annealing temperatures. Lane 1) 5 µl of O'GeneRuler DNA marker; 2) 60 °C; 3) 64 °C; 4) 66 °C; 5) 68 °C; 6) 70 °C.

3.4.2.4 Expression, purification and enzyme assay of the C-terminal domain mutant

The C-terminal domain of bovine GLYAT was expressed using the same conditions as for the wild type protein, described in Section 2.2.15. As shown on the polyacrylamide gel in Figure 3.15 (lane 4), virtually all of the expressed protein was insoluble. However, with the wild type protein the soluble fraction must first be passed through a nickel-affinity purification column and concentrated, as described in Section 2.3.3, before the small

amount of soluble protein can be visualised on a gel. The soluble fraction of a C-terminal domain expression was thus subjected to histidine tag affinity chromatography, as described in Section 2.2.17. As shown in Figure 3.15 (lane 6), no soluble protein was visible on an SDS-PAGE gel, even after purification and concentration.

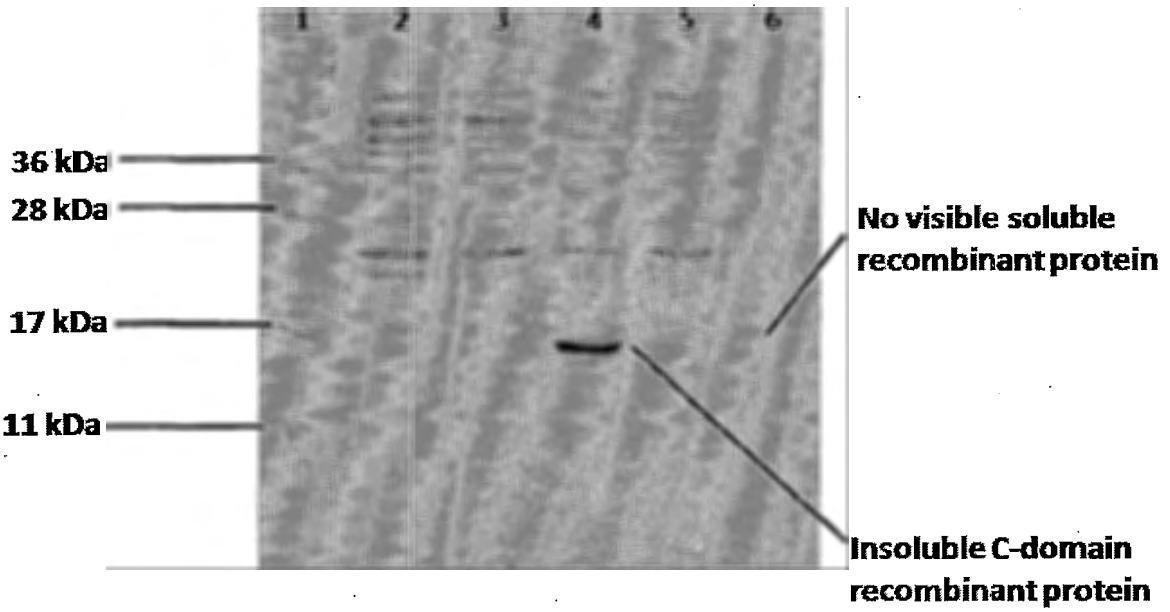

Figure 3.15 Expression and purification of the recombinant C-domain of bovine GLYAT. SDS-PAGE analysis of the expression and purification of the C-domain of bovine GLYAT. Lane 1) 5 µl of PageRuler protein marker; 2) un-induced culture, total fraction; 3) un-induced culture, soluble fraction; 4) induced culture, total fraction; 5) induced culture, soluble fraction; 6) after nickel affinity purification.

Enzyme activity assays on both the crude soluble protein fraction and the purified protein preparation were negative (result not shown). The complete insolubility of the protein may explain the lack of enzyme activity. Whatever the function of the N-terminal domain may be, it appears that it might be important for solubility and activity of the bovine GLYAT enzyme.

3.4.2.5 Expression and purification of the E226Q and E226H mutant GLYAT proteins

The E226H and E226Q bovine GLYAT mutants were expressed together with the wild type enzyme as described in Section 2.2.15, and purified as described in Section 2.2.17. SDS-PAGE analysis revealed that the levels of total mutant protein expressed was similar to that

obtained for the wild type enzyme (Figure 3.16, lanes 2, 5 and 8). At the time this expression was done, the strategy of adding 20 mM imidazole to the cell lysis and wash buffers had not yet been implemented. For this reason, the purified mutant protein preparations contained some non-specific proteins, as shown by SDS-PAGE analysis (Figure 3.16, lanes 4, 7 and 10).

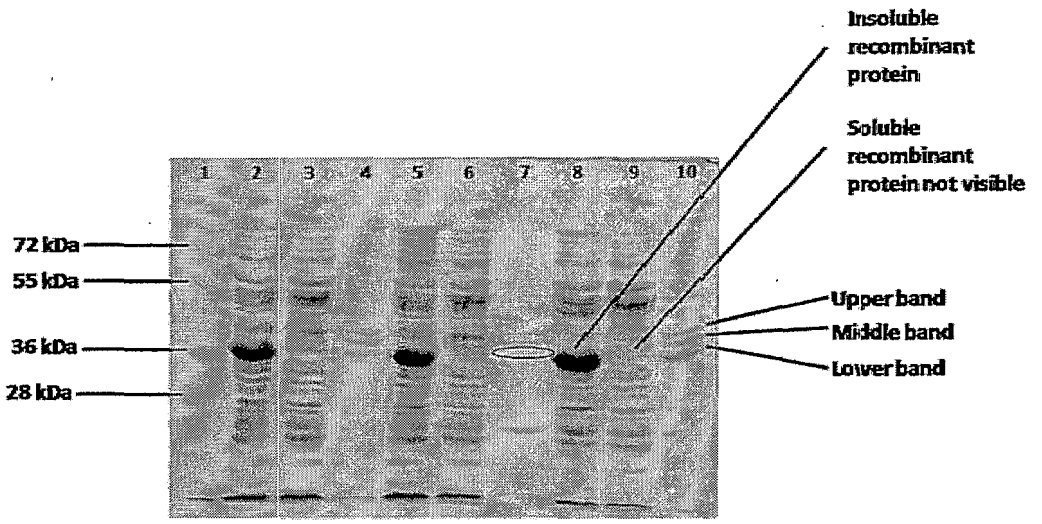

Figure 3.16 Expression and purification of the wild type and mutant recombinant bovine GLYAT enzymes. SDS-PAGE analysis of the expression and purification of the wild type and mutant recombinant bovine GLYAT enzymes. Lane 1) 5 µl of PageRuler protein marker; 2) wild type recombinant bovine GLYAT, total fraction; 3) wild type, soluble fraction; 4) wild type after purification; 5) E226H mutant recombinant bovine GLYAT, total fraction; 6) E226H, soluble fraction; 7) E226H after purification; 8) E226Q mutant recombinant bovine GLYAT, total fraction; 9) E226Q, soluble fraction; 10) E226Q after purification. The oval indicates the absence of the lower major protein band for the E226H mutant.

The purified mutant and wild type proteins were used in enzyme assays, performed in triplicate, where 2 µg of each protein was used, as described in Section 2.2.20. The result is reported in Figure 3.17. The wild type recombinant enzyme had significant activity and the E226Q mutant was about a third as active as the wild type. The E226H mutant had no enzyme activity. The catalytic model proposed in Section 3.1 suggested that the E226H mutant should be more enzymatically active than the E226Q mutant. This is not the result obtained, and the experiment was repeated to ensure that these results were not due to experimental error. However, the results were the same.

When the polyacrylamide gel in Figure 3.16 was re-examined, it became clear that the composition of the purified E226H protein preparation is different from the purified E226Q

and wild type proteins. In the region of the 36 kDa size marker, there are three prominent bands for the wild type recombinant protein (Figure 3.16, lane 4). The nature of these bands was discussed at length in Chapter 2. The top band was identified as a non-specific protein that co-purifies with GLYAT. When 20 mM imidazole was included in the cell lysis and wash buffers, this band was no longer found to co-purify with the recombinant bovine GLYAT. The middle or lower band must thus be responsible for enzyme activity. Whereas the E226Q mutant, like the wild type protein, has three protein bands in the 36 kDa region (lane 10), the E226H mutant has only two bands in this region (lane 7). The lower band is missing, perhaps due to instability or insolubility of the protein. It was argued in Section 2.3.3 that the lower band represents the recombinant bovine GLYAT. The lower band also appears to be the same size on the gel as the majority of the insolubly expressed recombinant protein (Figure 3.16, lanes 2, 5 and 8). Therefore, it is now thought that complete insolubility of the E226H mutant bovine GLYAT protein is responsible for the observed lack of enzymatic activity, explaining why a mutant that was expected to be active was observed to be inactive.

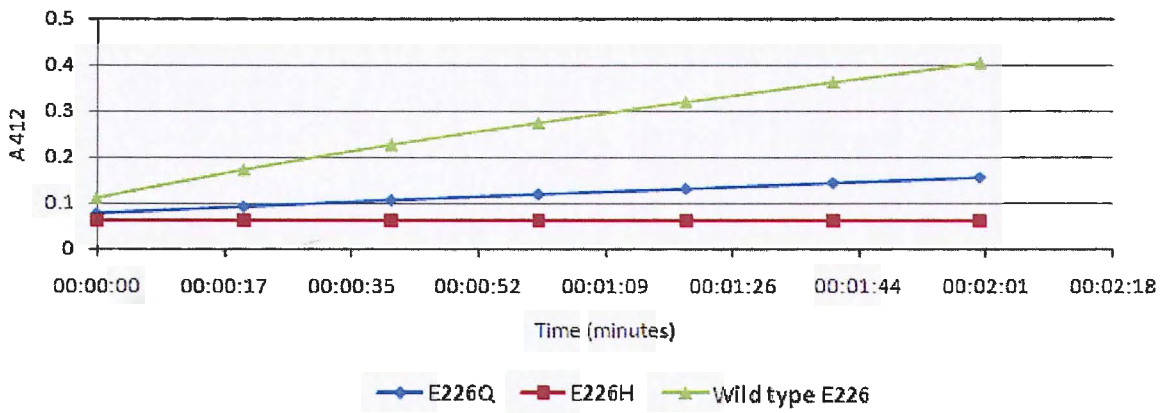

Figure 3.17 Standard assay of the wild type and mutant recombinant enzymes. The graph shows the change in absorbance at 412 nm with time. The wild type recombinant bovine GLYAT is compared to the E226Q and E226H mutant bovine GLYAT enzymes.

3.4.2.6 pH dependence of the enzyme activity of wild type and E226Q bovine GLYAT

As explained in Section 1.5.2.2, for direct nucleophilic acyl transfer to take place, the amine of glycine must be in the unprotonated form (Berndsen & Denu, 2005, Dyda et al., 2000, Palmer, 2001). Around physiological pH a primary amine will be protonated and thus

incapable of nucleophilic attack. Under the same conditions, however, the catalytic base residue will be deprotonated, and thus well positioned to remove a proton from the substrate, before releasing it to the bulk solvent. For this reason the reaction can occur rapidly at physiological pH, even though most glycine is present in the protonated form.

In the case of the bovine GLYAT E226Q mutant, the glutamate residue that is thought to be responsible for this proton removal is absent. Because it lacks the catalytic residue, it is expected that the mutant protein should be enzymatically inactive at low or physiological pH, where all glycine is protonated. If the activity of this enzyme is then dependent on the concentration of deprotonated glycine, it is expected that the activity will increase with increasing pH, as observed for several other GNAT enzymes, eventually approaching that of the wild type enzyme (Dyda et al., 2000, Vetting et al., 2005). This would be because a proton-removing catalyst is not as important under such high pH conditions and possession of one would give the wild type enzyme a less significant advantage (Dyda et al., 2000, Hegde et al., 2007, Scheibner et al., 2002, Vetting et al., 2005).

Therefore, by characterising the pH dependence of the mutant and wild type GLYAT enzyme reactions, the catalytic importance of residue E226 can be better assessed. If it is merely shown that the mutant protein is slightly inactive, as it is at pH 8.0 where the standard assay was carried out (Figure 3.17), it can not be concluded that the residue is catalytically important. The loss of enzyme activity of the mutant enzyme could also be explained by a change in solubility or structural stability of the enzyme. To irrevocably demonstrate that the loss of enzyme activity is the result of loss of a catalytic residue, it had to be demonstrated that the effect of the mutation is pH dependent.

The GLYAT assay was performed at several pH values from pH 6.0 to pH 8.0. The wild type and E226Q mutant recombinant enzymes were analysed, as well as the bovine liver enzyme which served as a control. To ensure that the enzyme is saturated with substrate at all the pH values, a vast excess of benzoyl-coenzyme A (300 µM) was used in the assay reactions (Palmer, 2001).

As shown in Figure 3.18, the enzyme activities of the wild type recombinant enzyme and the bovine liver enzyme correspond. The activities can be seen to increase gradually from

pH 6.0 to 7.0, after which it is constant up to pH 8.0 (Figure 3.18). This result corresponds to what has been reported in the literature. The pH dependence profile of the E226Q mutant GLYAT is very different. The enzyme is virtually inactive at pH 6.0, but the activity increases gradually with increasing pH, up to pH 8.0 (Figure 3.18). The effect of pH on the rate of the E226Q mutant is considerable. At pH 7.0 there is a difference in enzyme activity between the wild type and mutant enzymes of almost two orders of magnitude. At pH 8.0 the mutant enzyme becomes significantly active, about a third as active as the wild type enzyme. These results are as predicted by the catalytic model, and are strong support for it, as the observed pH dependence of the reaction is hard to explain in any other way (Dyda et al., 2000, Hegde et al., 2007, Hickman et al., 1999, Palmer, 2001, Scheibner et al., 2002, Vetting et al., 2005).

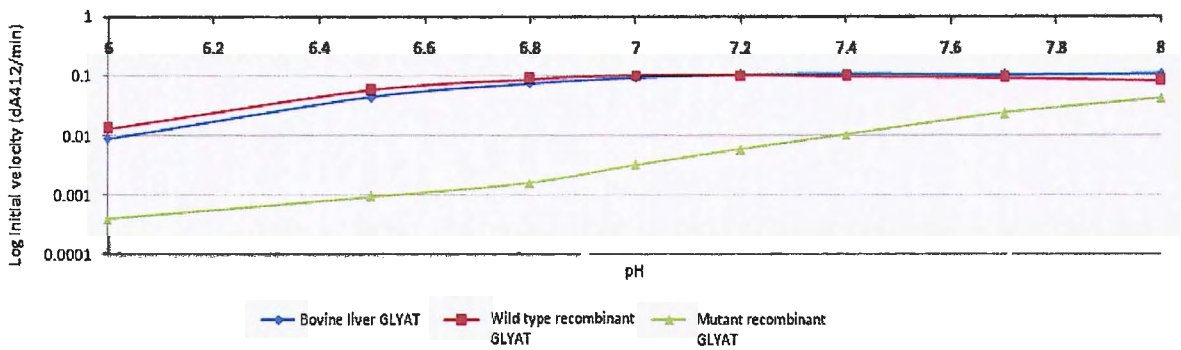

Figure 3.18 pH dependence of the activity of the wild type and E226Q enzymes. A logarithmic plot of initial velocity (dA412/min) against reaction pH for the bovine liver GLYAT, the wild type recombinant bovine GLYAT and the E226Q mutant bovine GLYAT.

3.4.2.7 Kinetic characterisation of the E226Q mutant

To further assess the properties of the E226Q mutant protein, its kinetic parameters were determined using benzoyl-coenzyme A and glycine as substrates. This was done as described in Section 2.2.22 of Chapter 2 for the wild type recombinant and bovine liver enzymes. The amount of protein added was determined in a pilot study (data not shown) such that the initial velocity at the maximal substrate concentrations (50 µM benzoyl-coenzyme A, and 20 mM glycine) was the same as that for the wild type enzyme, as discussed in Chapter 2 (a change of 0.24 absorbance units in four minutes). The purpose of this experiment was to determine the K_M values for the substrates, not to compare

reaction velocities. As was the case in Chapter 2, the fact that a homogenous protein preparation was not used makes comparison of maximal velocities pointless, as it is not known what proportion of the protein used is bovine GLYAT. Using equivalent amounts of protein for both the wild type and mutant recombinant proteins was also not feasible. Using such a small amount of mutant protein resulted in the enzyme reaction being immeasurably slow. However, for the present purposes of comparing K_M values, the approach of using more of the mutant enzyme was acceptable. The analysis was done at pH 8.0, the standard pH for GLYAT enzyme assays (Nandi et al., 1979). The experimental approach used in Section 2.2.22 of Chapter 2 was followed.

The kinetic parameters were calculated as described in Section 2.2.22, and the double reciprocal plot (Lineweaver-Burk plot) of the data obtained is shown in Figure 3.19. It can be seen that the dataset at 2.5 mM glycine is not reliable, as the standard deviation of the data points is too large. This is simply because the enzyme activity is very low under these conditions, and thus very difficult to measure accurately. The SigmaPlot software used to calculate the parameters takes this into consideration. The K_M values obtained for benzoyl-coenzyme A and glycine are reported in Table 3.3. The low enzyme activity of the mutant protein can be either because of loss of catalytic function, as proposed in Section 3.1, or it may be because the protein structure is altered in such a way as to increase the K_M values for the substrates. In that case, the protein may appear to be catalytically deficient simply because higher substrate concentrations are needed to reach a substantial velocity.

To investigate this, the K_M values for the wild type and E226Q recombinant GLYAT enzymes were compared. As shown in Table 3.3, the K_M value for benzoyl-coenzyme A is similar for the mutant protein and the wild type enzyme (18.1 ± 4.4 µM and 16.3 ± 1.2 µM, respectively). This means that inefficient binding of benzoyl-coenzyme A by the mutant enzyme is probably not the explanation of its inactivity relative to the wild type protein. The K_M value for glycine, however, is higher for the mutant GLYAT than for the wild type enzyme (7.0 ± 3.9 mM and 2.3 ± 0.3 mM, respectively). This may explain the lower activity of the mutant relative to the wild type enzyme to some extent. It is worth noting that in the case of serotonin N-acetyltransferase, the loss of activity upon mutation of the catalytic histidine residues (Section 1.5.2.1) is thought to be both because of loss of catalytic function and because the mutations increase the K_M value for serotonin, the acyl acceptor

substrate. Based on structural information this was explained by noting that the histidine residues, through a chain of water molecules on the surface of the protein, make contact with the serotonin amino group, and are thus involved in its binding to the protein. However, it must be considered that the increased K_M value may simply reflect that the enzyme can not use all of the available substrate, but only that which is already present in unprotonated form (Hickman et al., 1999, Scheibner et al., 2002). Levels of deprotonated substrate is in turn dependent on pH (Section 1.5.2.1).

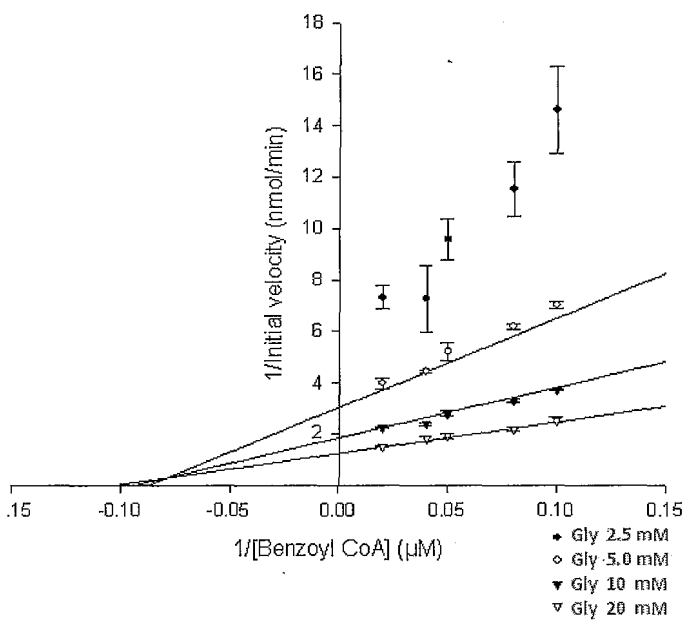

Figure 3.19 Lineweaver-Burk plots for the E226Q protein. The Lineweaver-Burk plots used to calculate the kinetic parameters of the E226Q mutant bovine GLYAT enzyme for benzoyl-coenzyme A and glycine. The data set for 2.5 mM glycine is not reliable, and a line is not drawn for this data set. The data points indicate average values ± standard deviation, with n = 3. Assays were performed at pH 8.0.

Table 3.3: Kinetic parameters for the wild type and E226Q mutant recombinant bovine GLYAT enzymes using benzoyl-coenzyme A and glycine as substrates

GLYAT enzyme	K_M benzoyl-coenzyme A (µM)	K_M glycine (mM)	Relative V_{max} (nmol/min)
E226Q recombinant GLYAT	18.05 ± 4.40	7.04 ± 3.92	1.52 ± 0.21
Wild type recombinant GLYAT	16.26 ± 1.23	2.33 ± 0.28	2.23 ± 0.07

Using high substrate concentrations of 100 µM benzoyl-coenzyme A and 200 mM glycine, the activity of the mutant enzyme appears to be approximately one third of that of the wild type enzyme. This is a benzoyl-coenzyme A concentration much higher than the K_M values of both the wild type and mutant enzymes (approximately 5X), and a glycine concentration several times higher than the K_M values of both enzymes (approximately 80X and 30X, respectively). These concentrations are thus likely to be near saturating for both enzymes, and the difference in activity can then not be completely explained by differences in the substrate binding constants of the proteins.

If this information is considered together with the pH dependence of the reaction, discussed in Section 3.4.2.6, it is clear that the low activity of the E226Q mutant bovine GLYAT enzyme, compared to the wild type recombinant enzyme, is largely because of loss of the catalytic glutamate residue. This probably reflects loss of the capacity to remove a proton in the reaction mechanism.

3.5 Conclusion and summary

The aim of the work presented in this chapter was to work out the catalytic mechanism of bovine GLYAT. In the case of acyltransferase enzymes there are just two fundamental types of mechanism that can be employed (Berndsen & Denu, 2005). These are the ping-pong mechanism and the ternary complex mechanism (Figure 3.20). Knowing that a ternary complex mechanism is employed by GLYAT (Nandi et al., 1979, van der Westhuizen et al., 2000) was the starting point for working out the detailed mechanism, since the general mechanism of direct acyl transfer is well understood (Berndsen & Denu, 2005, Dyda et al., 2000, Vetting et al., 2005). The hypothesis was that a conserved glutamic acid residue, E226 of the bovine GLYAT enzyme, is employed as a general base catalyst in the reaction mechanism (Figure 3.8). The proposed function for this residue in the reaction mechanism would be removal of a proton from glycine, making it nucleophilic and enabling it to attack the thioester bond of the acyl-coenzyme A.

The importance of the proposed catalytic glutamate residue was experimentally investigated using site-directed mutagenesis. The glutamate residue was mutated to a

glutamine residue (E226Q), and the mutant recombinant GLYAT enzyme subjected to biochemical characterisation. The enzyme activity of the E226Q mutant was found to be pH dependent. At pH 6.0, the enzyme is virtually inactive, but at pH 8.0 considerable activity is observed (Figure 3.18). This supports the proposal that the E226 residue acts as a proton removing catalyst in the reaction mechanism (Berndsen & Denu, 2005, Dyda et al., 2000, Hegde et al., 2007, Hickman et al., 1999, Palmer, 2001, Schachter & Taggart, 1954a, Scheibner et al., 2002, Vetting et al., 2005).

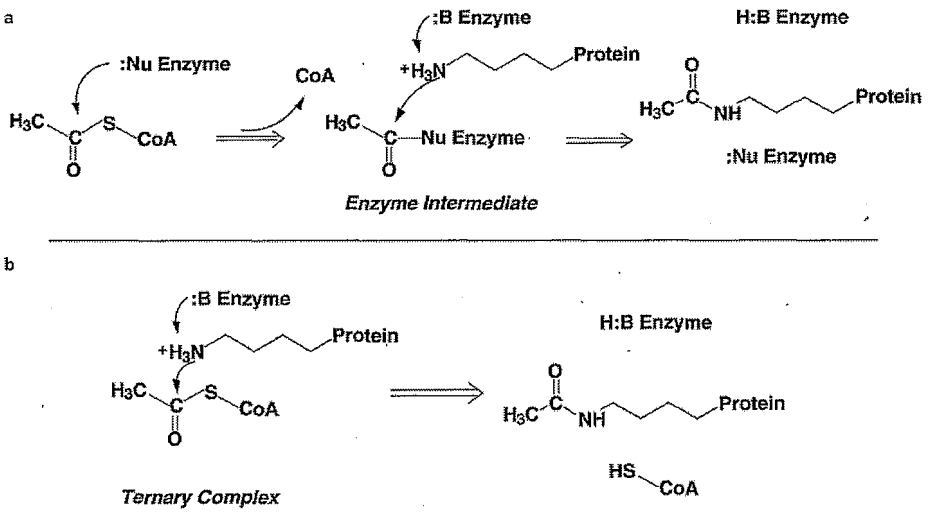

Figure 3.20 Mechanisms proposed for acyltransferase enzymes. a) the ping-pong reaction mechanism, involving an acyl-enzyme intermediate. b) the direct transfer mechanism, involving a ternary complex and direct nucleophilic attack. These mechanisms were discussed in Section 1.5.2.1 (Berndsen & Denu, 2005).

Kinetic characterisation of the mutant GLYAT enzyme showed that the K_M value for glycine is about 3X higher than for the wild type and the K_M value for benzoyl-coenzyme A is about the same as for the wild type enzyme (Table 3.3). These differences from the wild type enzyme are not sufficient to explain the low enzyme activity of the mutant GLYAT. When considered together with the pH dependence profile of the mutant enzyme, it can be concluded that the E226 residue is involved in the reaction mechanism as a proton removing catalyst.

The catalytic mechanism thus proposed is drawn in Figure 3.21. Although worked out for the bovine enzyme, it is thought to be representative of the whole family of GLYAT

enzymes. It is thought that most of the important details are included, but future work, including determination of the enzyme's structure, could further refine this model. For example, the proton removal step catalysed by the E226 residue is likely to involve a "proton wire," a chain of water molecules between the glutamate and the substrate. This proton conduit allows proton removal at a distance, and the phenomenon is common in the GNAT superfamily (Dyda et al., 2000, Vetting et al., 2005).

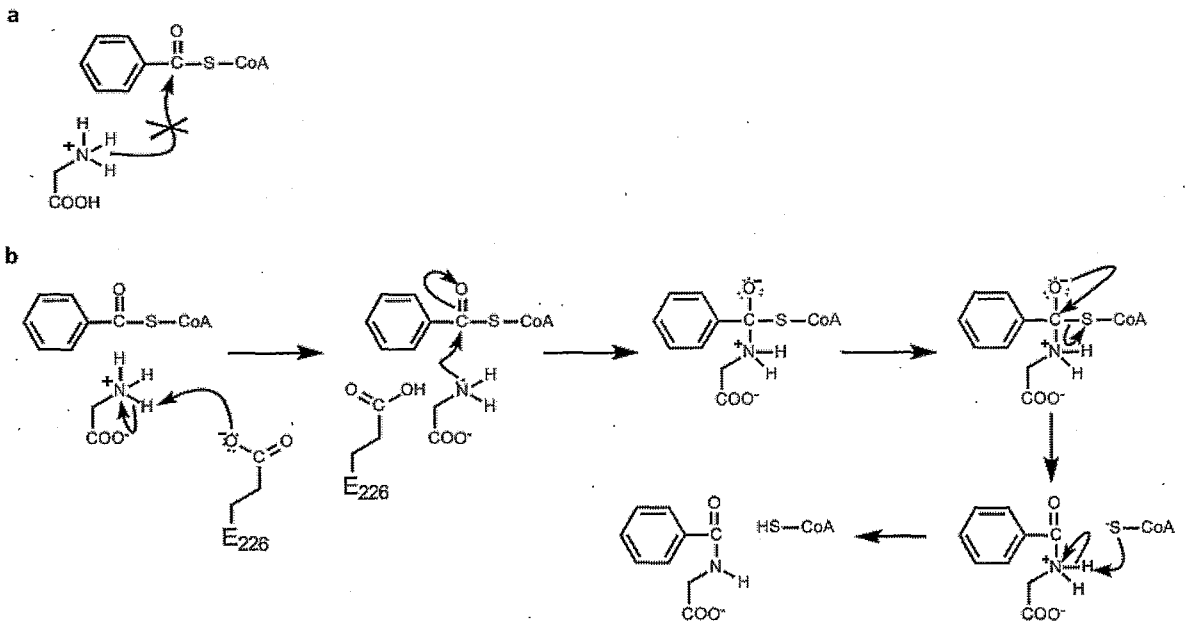

Figure 3.21 Proposed catalytic mechanism of the bovine GLYAT enzyme. The reaction mechanism proposed for bovine GLYAT is schematically represented. a) For nucleophilic attack to take place, the glycine amino group must be deprotonated; b) the proposed mechanism involves proton removal by E226 which enables nucleophilic attack on the thioester carbon atom. A tetrahedral intermediate is formed, following nucleophilic attack by the amino group of glycine on the thioester carbon atom. Finally the tetrahedral intermediate collapses, forming the peptide product and the thiolate form of coenzyme A, which is protonated before products leave the enzyme.

The catalytic mechanism for glycine N-acyltransferase has not been reported in the literature. We believe that this is the first time that the catalytic mechanism for a GLYAT enzyme has been worked out, and the work presented here will be published in a peer reviewed scientific publication. In addition to providing some new understanding of the molecular mechanisms of the bovine GLYAT enzyme, the work presented in this chapter serves to demonstrate the usefulness of the recombinant bovine GLYAT expression system as a tool for investigating the GLYAT enzyme.