Studies on Sulphur Amino Acid Metabolising Enzymes of *Trichomonas Vaginalis*

Gordon Goodall

Division of Infection and Immunity
&
Protein Crystallography Group
University of Glasgow
Glasgow G12 8QQ
Scotland UK

This thesis is presented in submission for the degree of

Doctor of Philosophy in the Faculty of Science.

January 2001

Summary

Methionine γ-lyase (MGL) from *Trichomonas vaginalis* has been crystallised and the structure solved by molecular replacement at 2.2 Å. The enzyme has a similar overall secondary structure arrangement as the search model, cystathionine β-lyase from *Escherichia coli*, but differs in the active site. In addition to the holoenzyme, the structure of the enzyme in complex with the acetylenic suicide inhibitor L-propargylglycine has also been solved. This has revealed the mechanism of inhibition by this compound. It has also provided insights into the catalytic mechanism of the enzyme and allowed the postulation of a scheme for its action. Particularly, this predicts core roles for active site tyrosine and cysteine residues (111 and 113, respectively).

To test the hypothesised mechanism of catalysis, seven site-directed mutants were produced. The activities of the mutants were determined and the structure solved for one of the most interesting. The resulting data confirm the mode of action of L-propargylglycine and also reveal a proton relay which activates the important active site cysteine which mediates much of the enzyme activity. This mechanism is substantially different from that proposed for other members of the same family of enzymes and explains the substrate specificity of MGL. The structure also reveals a potential role for the enzyme. The active site cysteine is positioned ideally to form an internal persulphide with a sulphur atom released from the substrate. This implies a possible role for the enzyme in iron-sulphur cluster formation in an analogous fashion to the enzyme, cysteine desulphurase.

A cDNA clone, isolated from a *T. vaginalis* library, with high level of homology to cysteine synthase has been subcloned and recombinant protein expressed in *E. coli*. The purified recombinant enzyme possesses both cysteine synthase activity and the "activated serine sulphydrase" activity previously described in the parasite. Southern blot analysis suggests the presence of just a single copy of the gene in the parasite, with other closely related homologues also being present. Antibodies raised against the purified enzyme revealed that there are at least two isoenzymes present in parasite lysates. Cultivation of the parasite in medium supplemented with L-cysteine has revealed reduced expression of the enzyme. This is the first demonstration of part of the *de novo* synthesis of cysteine in the *T. vaginalis*.

THESIS CONTENTS

LIST OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

CHAPTER 4

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

ABBREVIATIONS

AAT	Aspartate Aminotransferase
ADP	Adenosine 5'-Diphosphate
$AmmSO_4$	Ammonium Sulphate
AMPSO	3-[(1,1-Dimethyl-2-hydroxyethyl0amino]-2-hydroxy-propanesulfonic acid
ATP	Adenosine 5'-Triphosphate
BIS-TRIS	bis[2-hydroxyethyl]iminotris[hydroxmethyl]methane
BTP	1,3-bis[tris(Hydroxmethyl)methylamino]propane
CAPSO	3-[Cyclohexylamino]-2-hydroxy-1-propanesulfonic acid
CBL	Cystathionine β-Lyase
CBS	Cystathionine β-Synthase
CGL	Cystathionine γ-Lyase
CGS	Cystathionine γ-Synthase
CH_3SH	Methanethiol
DLS	Dynamic Light Scattering
DTT	Dithiothreitol
EDTA	[Ethylenedinitrilo]tetraacetic acid
EPPS	N-[2-Hydroxyethyl]piperazine-N'[3-propanesulfonic acid]
Fe-S	Iron:sulphur core
HEPES	N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]
H_2S	Hydrogen Sulphide
IPTG	Isopropyl-β-D-thiogalactopyranoside
k_{cat}	Enzyme Turnover Number
Km	Michaelis Constant
LB_{AMP}	Luria Bertani Broth + Ampicillin
$LB_{AMP/KAN}$	Luria Bertani Broth + Ampicillin + Kanamycin
LDH	Lactate Dehydrogenase
MBTH	Methylbenzoylthiolhydrazone
MDM	Modified Diamonds Medium
MES	2-[N-Morpholino]ethanesulfonic acid
Methyl Viologen	1,1'-Dimethyl-4,4'-bipyridinium

MGL	Methionine γ-Lyase
MOPS	3-[N-Morpholino]propanesulfonic acid
MR	Molecular Replacement
NCS	Non-Crystallographic Symmetry
NH_4^+	Ammonia
Ni-NTA	Nickel-Nitrilotriacetic Acid
OAHSS	O-Acyl Homoserine Sulphydrylase
OAS	O-Acetyl-L-Serine
OASS	O-Acyl Serine Sulphydrylase
OD	Optical Density
ORF	Open Reading Frame
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PFOR	Pyruvate:ferredoxin oxidoreductase
PIPES	Piperazine-N,N'-bis[2-ethanesulfonic acid]
PLP	Pyridoxal 5'-Phosphate
RACE	Rapid Amplification of cDNA Ends
SAH	S-Adenosyl-Homocysteine
SAM	S-Adenosyl Methionine
SAT	Serine Acetyl Transferase
SCL	Selenocysteine Lyase
S.D.	Standard Deviation
SDM	Site-Directed Mutagenesis
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
TCA	Trichloro Acetic Acid
TES	N-tris[Hydroxymethyl]methyl-2-aminoethanesulfonic acid
TRIS	tris[Hydroxymethyl]aminomethane
TWEEN	Polyoxyethylenesorbitan
V_{max}	Enzyme Maximum Rate
X-GAL	5-bromo-4-chloro-3-indol-β-D-galactopyranoside

DECLARATION

I declare that the work recorded in this thesis is entirely my own, unless otherwise stated and that it is of my own composition. No part of this work has been submitted for any other degree.

The results given in table 3.1 were obtained by Dr. AJ Lapthorn. Figures 5.18 and 5.19 were produced by Dr. GD Westrop.

ACKNOWLEDGEMENTS

I would firstly like to acknowledge the contribution of both of my supervisors, Prof. Graham Coombs and Dr. Adrian Lapthorn. Thanks to all the people from North lab, protein crystallography and briefly WUMP who helped me during my PhD.

There are too many people and too many reasons for me to thank everybody individually. Thanks for help, patience, favours, laughter and countless other things and if you believe you deserve it, then you have it. Even so I do have to say personal thanks to Angus and Ali and they know why and especially to Jude for reminding me.

I would like to dedicate this thesis to my parents and family who have given me a place to call home and perspective whenever I was close to losing it. To Mum and Dad especially, for always supporting and never fussing - thanks for everything. Finally to Dad, for being the man I hope to grow up to be one day.

CHAPTER 1. INTRODUCTION

1.1 Biology of Trichomonads

Trichomonads are flagellated parasitic protozoa that infect a large variety of organisms. The life cycle is generally uncomplicated with only a single stage identified. True cysts have been seen only in species parasitic in reptiles and amphibians although pseudocysts are found more commonly in other species. There is no known sexual process and reproduction is by binary fission. It is generally accepted that only three species infect humans all of which are highly site specific. *Trichomonas vaginalis*, the subject of this study, infects the urogenitary tract of males and females, *Trichomonas tenax* occurs in the oral cavity and *Pentatrichomonas hominis* occurs in the large intestine. Only *T. vaginalis* is considered pathogenic. The taxonomic lineage for this parasite is: *Eukaryota*, *Parabasalidae*, *Trichomonadida*, *Trichomonas vaginalis*.

1.2 Cell Biology of *T. vaginalis*

T. vaginalis in free culture is between 10 to 20μm in length and ovoid in shape. During interactions with host tissues or cultured cells, however, the parasite exhibits irregular forms (Heath, 1981). Figure 1.1 illustrates the basic cellular structure of the parasite.

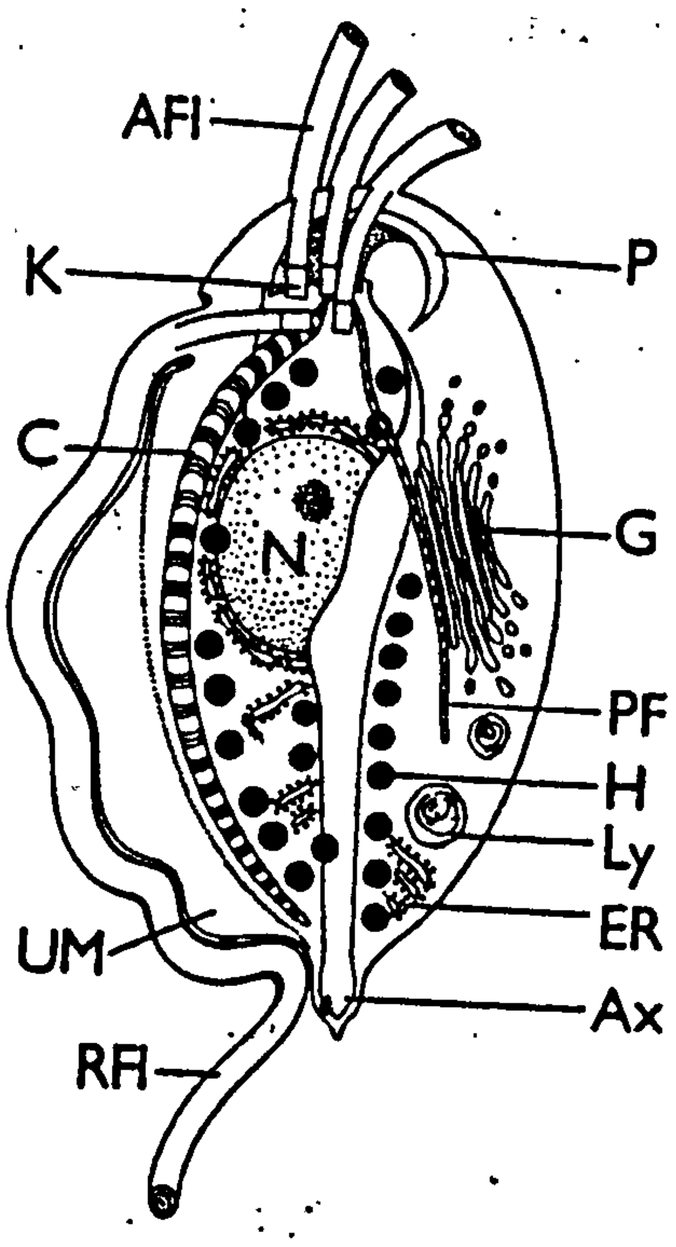

Figure 1.1: Schematic diagram of trichomonad cell. AFl - anterior flagella, K - kinetosomes (basal bodies of flagella), P - pelta, C - costa, N - nucleus, G - parabasal apparatus, PF - parabasal fibril, H - hydrogenosome, Ly - secondary lysosome, ER - rough endoplasmic reticulum, Ax - axostyle, UM - undulating membrane, RFl - recurrent flagellum. Figure reproduced from Kulda *et al.*, 1986.

Equipped with five flagella the parasite is capable of rapid movement. Four flagella are located anteriorly with the fifth attached along the length of the organism to form an undulating membrane. The parasite is also able to exhibit amoeboid movement when in contact with epithelial cells forming pseudpodia for attachment and motility. The cell has a complex cytoskeleton. In particular an axostyle formed of parallel microtubules stretches the length of the organism and protrudes posteriorly and a striated costa extends below the undulating membrane. There is a modified Golgi complex which is integrated with root fibrils to form the parabasal apparatus. There is a large nucleus and the cytoplasm also contains vacuoles, vesicles, ribosomes and glycogen granules. There is a

conspicuous lack of mitochondria, glycosomes or peroxisomes within the cytoplasm and instead the parasite possesses hydrogenosomes (Lindmark and Müller, 1973). These organelles are peculiar to fungi and anaerobic protozoa and play a pivotal role in metabolism within the parasite.

1.3 Hydrogenosomes

Hydrogenosomes of *T. vaginalis* are double membrane-bound spherical organelles which, unlike mitochondria, lack cristae. Around 0.5 - 1.0 μm in diameter there are typically 30 to 40 in each cell. They are normally located near the axostyle and costa away from the cell membrane. Their name is derived from the hydrogen evolved from them due to activity of their marker enzyme hydrogenase (Müller, 1993). Hydrogenosomes were first described in trichomonads (Lindmark and Müller, 1973) and have since been found exclusively in protists and lower fungi without mitochondria which inhabit anaerobic or microaerophilic environments (Snyers *et al.*, 1982; van Bruggen *et al.*, 1983; Yarlett *et al.*, 1986; Paul, *et al.*, 1990; Fenchel and Finlay, 1991). Also unlike mitochondria, it is largely accepted now that *T. vaginalis* hydrogenosomes lack DNA (Clemens and Johnson, 2000).

1.4 Evolutionary Position of *T. vaginalis*

In addition to the interest in the parasite due to its human pathogenicity and prevalence, it is also studied because of its possible evolutionary lineage. It is argued that it appears to be a representative of a very early divergence from the main line of eukaryotic evolution. This postulation is based upon phylogenetic trees derived from small subunit ribosomal RNA (Sogin *et al.*, 1996), the

possession of 70S ribosomes (rather than 80S) and the lack of mitochondria and possession of hydrogenosomes. Evidence based upon the sequences of heat shock protein (HSP) genes has also been given (Bui *et al.*, 1996; Germot *et al.*, 1996). Arguments ensue from the question as to whether hydrogenosomes are degenerate mitochondria (Cavalier Smith, 1987), have evolved alongside mitochondria from a common progenitor organelle or arisen completely independently via a separate event (Whately *et al.*, 1981). The situation is far from clear though and perhaps the only fact that is definite, is that, at the moment, nobody knows (Embley and Hirt, 1998).

1.5 Clinical Aspects of *T. vaginalis* Infection

Reports on the manifestations and frequency of infection are notoriously difficult to interpret at times due to the source of the information. For example, a distorted view of the prevalence of the disease can be given if the data are collected from an STD (Sexually Transmitted Disease) clinic or gynaecologist. Patients presenting to a gynaecologist are likely to be symptomatic and seeking treatment whereas those at an STD clinic may be attending because of a symptomatic partner or by contact tracing. Thus although certain facts are apparent, the true pathological burden of the organism is very difficult to quantify.

1.5.1 Diagnosis

Diagnosis of trichomonosis is complicated by the broad spectra of symptoms exhibited and the similarities to vaginitis resulting from other sources, such as bacterial vaginosis. The most successful method for positive diagnosis of trichomonad infection is wet mounted microscopic examination and cultivation,

with detection levels of 95% achievable (Kreiger *et al.* 1988). Despite the fact that this is a time consuming and labour intensive process, efforts to design a system based upon *T. vaginalis* nucleic acids, such as PCR based systems (Riley *et al.*, 1991; Paterson *et al.*, 1999) and fluorescent in-situ hybridisation (FISH) (Muresu *et al.*, 1994) have gone largely unnoticed.

1.5.2 Pathology

Common manifestations of the disease in females are a vaginal discharge and vulvovaginal irritation at frequencies of up to 75% and 25%, respectively. The classical disease symptoms are described as a frothy, malodorous, green discharge but this is seen in few symptomatic women. Dysuria is often present in sufferers and less common is abdominal pain. The majority of males appear to be asymptomatic carriers of the disease, a state which can last up to four months (Kreiger *et al.* 1993), but complain most frequently of mild urethritis. An added concern is that trichomonosis appears to be a risk factor for HIV transmission, through the effects the parasite has on host immune mechanisms (Laga *et al.*, 1993; Martin *et al.*, 1999).

1.5.3 Trichomonosis Occurrence

Trichomonosis is the most prevalent non-viral sexually transmitted disease in the world (McLellan *et al.*, 1982). The most recent estimates from the WHO predicted an estimated 170 million new cases for 1995. The distribution of the disease varies significantly with the local population. Known risk factors include multiple sexual partners, history of sexually transmitted disease, non use of contraceptives, black race and co-existent infection with *Neisseria gonorroheae*

(Cossack, 1989). Transmission of the disease is only via sexual intercourse with no documented cases of infection by any other means despite evidence that the parasite remains viable for short periods on moist objects (Whittington, 1957; Burch *et al.*, 1959). Infection with *T. vaginalis* varies widely. An estimated 25-50% of females are asymptomatic but up to a third of patients develop symptoms within six months (Rein, 1990).

1.5.4 Treatment

The treatment of trichomonosis is essential to avoid further transmission to sexual partners and complications arising from the infection. Metronidazole is the drug treatment of choice, usually administered in a single oral dose. One of the 5-nitroimidazole group of compounds, others such as ornidazole and tinidazole are also used occasionally. Metronidazole exerts its effect via the reduction of its nitro group by the pyruvate:ferredoxin oxidoreductase of the parasite and the concomitant production of cytotoxic intermediates (Edwards, 1993). Other treatments have a much lower degree of efficiency and are generally more difficult to administer and thus not generally considered. However, although there is no direct evidence for any carcinogenic or teratogenic properties of the toxic intermediates of metronidazole, use of the drug is avoided during pregnancy.

1.5.5 Drug Resistance

The extent to which drug resistance is a problem in clinical terms is difficult to quantify. In 1989 the Centre for Disease Control in the United States estimated that 5% of *T. vaginalis* isolates exhibited some form of resistance to metronidazole. There have been reports of at least 100 metronidazole resistant

isolates in the United States of America (Müller *et al.*, 1980; Kellock and O'Mahoney, 1996; Land and Johnson, 1999) and a substantial number in Europe as well (Heyworth *et al.*, 1980; Kulda *et al.*, 1982; Meri *et al.*, 2000). The consequences of widespread resistance to the 5-nitroimidazoles are serious and they highlight the need for further research into the organism and possible drug targets.

1.5.6 Immunity to *T. vaginalis*

There is no protective immunity from prior infection with the parasite, although *T. vaginalis* specific antibodies have been detected in patients (Coyne *et al.*, 1975; Cogne *et al.*, 1985). However these observations are complicated by results showing the immunoprecipitation of large numbers of *T. vaginalis* proteins by sera from infected and non-infected patients (Davis-Hayman *et al.*, 2000). There is also evidence of non-specific host immune responses being activated, such as neutrophils and complement. However, the persistent infection that typically occurs in females highlights the inadequacies of these mechanisms. It has been possible to induce a protective response in mice after subcutaneous immunisation with whole parasites and subsequent intravaginal challenge (Abraham *et al.*, 1996).

1.6 Biochemistry of *T. vaginalis*

1.6.1 *In vitro* growth

Studies on all aspects of the biochemistry and molecular biology of the organism have relied upon successful cultivation *in vitro* to provide adequate starting

material. The parasite is thought of as an aerotolerant anaerobe and yet it grows optimally in low amounts of oxygen (Paget and Lloyd, 1990). It has been demonstrated by many different investigators that the organism requires a wide range of nutrients. These include carbohydrates (Read, 1957), fatty acids (Roitman *et al.*, 1978), purines (Heyworth *et al.*, 1982), pyrimidines (Wang and Cleng, 1984), iron (Gorrell, 1985), vitamins (Hollander and Leggett, 1985) and amino acids (Rowe and Lowe, 1986). These nutritional requirements are met *in vivo* by the vagina and environment (Huggins and Petri, 1981) but interestingly the parasite has been shown to be able to phagocytose different cell types including mammalian cells and bacteria (Francioli *et* al., 1983; Rendón-Maldonado *et al.*, 1998). Whether this ability is used primarily to supplement nutrition or as a defence mechanism is unknown.

The most widely used types of media for all trichomonad cultivation are undefined. They each have a similar composition and contain elements such as peptone, yeast extract and serum (for example, Diamond, 1957). A number of semi-defined media have been proposed for various species and a defined medium for *T. vaginalis* cultures also exists (Linstead, 1981).

1.6.2 Carbohydrate Metabolism

T. vaginalis has a fermentative metabolism in common with other anaerobic protozoa such as *Giardia lamblia* and *Entamoeba histolytica*. The major end products are a mixture of acids (e.g. acetate, succinate and lactate) and alcohols (e.g. ethanol) (Müller, 1988). The parasite is thought to be heavily reliant upon glycolysis for the production of ATP with the end products able to maintain redox

balance and generate further energy via additional pathways. For reviews of the parasites metabolism the reader is directed to Müller, 1988 and Coombs and Müller, 1995.

Of particular note is the fate of pyruvate produced by glycolysis which via oxidative decarboxylation is catabolised to acetyl Co-A by the enzyme pyruvate:ferredoxin oxidoreductase (PFOR)(Williams *et al.*, 1987). This enzyme utilises ferredoxin as its electron acceptor and the reduced ferredoxin is then oxidised by hydrogenase. All three central proteins in *T. vaginalis* energy production possess iron:sulphur (Fe-S) cores.

1.6.3 Amino Acid Metabolism

Although it is assumed that *T. vaginalis* relies heavily upon glycolysis for energy production, there is evidence that it can utilise other substrates and in particular amino acids. The following sections of this introduction considers relevant aspects of amino acid metabolism in the parasite and in particular sulphur containing amino acids, the main consideration for this study.

1.6.3.1 Proteinases

The presence of high levels of proteinase activity with the parasite, both intra- and extracellularly, has been well documented (for review see North, 1991). The enzymes may digest extracellular host polypeptides, generating free amino acids, for transport into the cell via amino acid transporters or alternatively act upon phagocytosed material. However, this does not exclude other roles for

proteinases in disruption of the host immune system and modulating cytoadherence

1.6.3.2 Aminotransferase Activities

Aminotransferase enzymes, as their name implies, are able to transfer the amino group from a donor molecule to an acceptor. This process is intimately linked with the synthesis and degradation of amino acids and generally constitutes the last and first steps respectively. There are a considerable number of aminotransferase activities in *T. vaginalis*. Those active towards aspartate and aromatic, branched chain and ω-amino acids have all been demonstrated (Lowe and Rowe, 1986).

1.6.3.3 Amino Acids in Energy Production

There is direct evidence that the degradation of arginine can be utilised to produce ATP at a ratio of 1:1. Arginine serves as a rich source of nitrogen and medium used to culture trichomonads is considerably depleted of this amino acid (Lockwood and Coombs, 1991a). The excretion of large amounts of alanine, an end product of glucose metabolism, may account for this need for nitrogen. The breakdown of arginine catalysed by arginine deaminase and catabolic ornithine transcarbamoylase leads to the production of ornithine which can then be used for polyamine biosynthesis. The subsequent action of carbamate kinase yields ATP and ammonia and CO_2 as by-products (Linstead and Cranshaw, 1983). This pathway is summarised in figure 1.2 below.

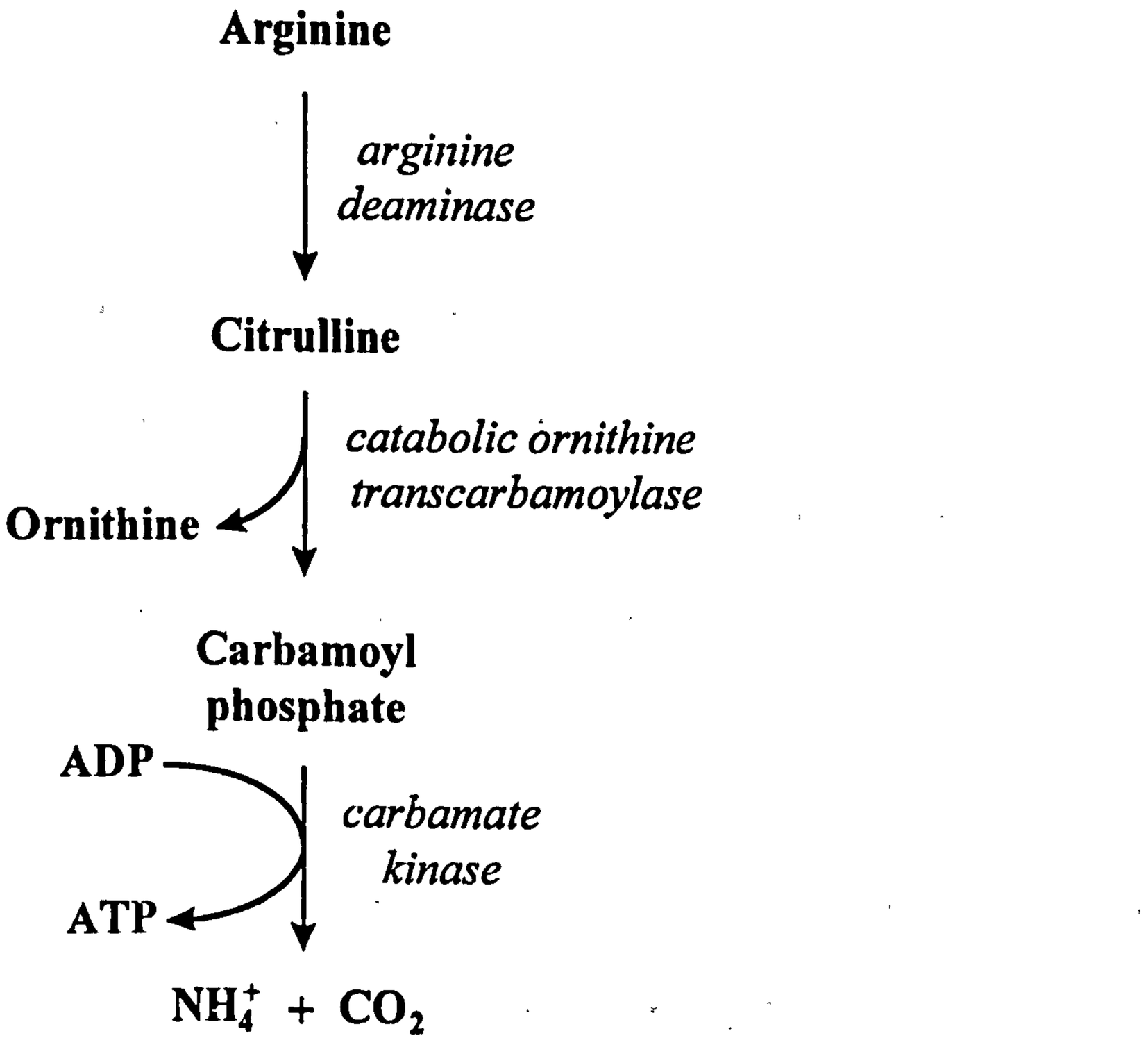

Figure 1.2: Pathway leading to production of ATP from the catabolism of arginine

The aminotransferase activities may also play a part in energy metabolism. It has been observed that PFOR is able to utilise a number of α-keto acids in addition to pyruvate in its core role in energy production (Williams *et al.*, 1987). Thus if glucose levels are low the parasite may be able to switch energy production to the use of amino acids. There have also been described alternative 2-keto acid oxidoreductases able to utilise these deaminated amino acid carbon skeletons (Brown *et al.*, 1999).

1.7 Sulphur Amino Acids

The enzymes investigated in this study from *T. vaginalis* act upon sulphur containing amino acids. The following part of this chapter considers the use of

sulphur in biological systems and then specifically in amino acids. Finally specific enzymes are reviewed.

1.7.1 Biological Roles of Sulphur

It is far beyond the scope of this work to review properly this subject but it is worth mentioning some of the enormous number of functions that sulphur- containing compounds are involved in (for a comprehensive summary the reader should see Mitchell, 1996).

The most familiar biologically relevant sulphur-containing compounds are methionine and cysteine, two amino acids that commonly occur in proteins. Cysteine especially plays important catalytic roles in enzymes and because of its ability to dimerise to form cystine is also crucial in protein and tissue stability. N-Formylmethionine is encoded as the first amino acid in polypeptide chains although often subsequently cleaved. It is also used as a single carbon donor and this is perhaps its most important role. S-Adenosyl-L-methionine (SAM) is synthesised with ATP and can then be used in methylation reactions or in the production of polyamines. The synthesis of glutathione indicates another key role of sulphur, that of a reducing agent and scavenger of toxic radicals. Other sulphur-containing compounds such as thioredoxin, glutathione, metallothionein also play varied critical roles in a wide range of processes. Non-amino acid sulphur is also used in proteins in Fe-S core formation. These use a combination of cysteine residues within the protein and also sulphur atoms to coordinate iron atoms. This may be of particular interest in this study because of the three proteins with Fe-S cores having key roles in *T. vaginalis* metabolism. There are in

addition a large number of biologically active molecules which are useful because of the chemistry of sulphur, these include Coenzyme A (CoA), lipoic acid and thiamine (vitamin B_1).

1.7.2 Sulphur Amino Acid Metabolism

There are certain features of sulphur amino acid metabolism shared amongst all organisms, other aspects apply to only a few species. To understand the roles of enzymes involved in both synthesis and degradation of sulphur containing amino acids in *T. vaginalis* it is necessary first to examine the available data for other organisms.

Methionine and cysteine, the two sulphur containing amino acids found in proteins, are synthesised and interconverted predominantly by three mechanisms. The existence and relevance of any individual step in these is organism dependent and the major differences will be highlighted here. The three pathways are: sulphydrylation, transsulphuration and the methionine cycle. Sulphydrylation is the condensation of sulphide and a carbon backbone to produce a C-S compound. Transsulphuration is the two step conversion of cysteine to homocysteine via cystathionine. Reverse transsulphuration unsurprisingly converts homocysteine to cysteine, again via cystathionine. The enzymes used in each direction of transsulphuration are related, but the presence of those catalysing one direction does not indicate the presence (or absence) of the other. The methionine cycle begins with the methylation of homocysteine and then generally incorporation of methionine into proteins or production of SAM. SAM is converted in two steps back to homocysteine.

An additional point worth considering is the origin of the sulphur atom. The assimilation of inorganic sulphur is commonly used to provide the sulphide necessary for sulphydrylation and so the *de novo* synthesis of cysteine and homocysteine. Briefly, inorganic sulphate is usually taken up from the extracellular environment, phosphorylated twice using ATP and then reduced sequentially to sulphite and finally sulphide. This then provides an active sulphur molecule ready to be used. It would be wrong to imply that this is the only source of sulphur for incorporation into amino acids. However in many microorganisms and plants this may provide the most abundant source *in vivo*.

This brief overview is an oversimplification but suffices as a basic introduction to the subject. Figures 1.3, 1.4 and 1.5 summarise the steps referred to above and indicate the key enzymes. Figure 1.6 shows the chemical structures of the molecules involved for reference. The function, mechanism and existence of these enzymes is considered in detail below.

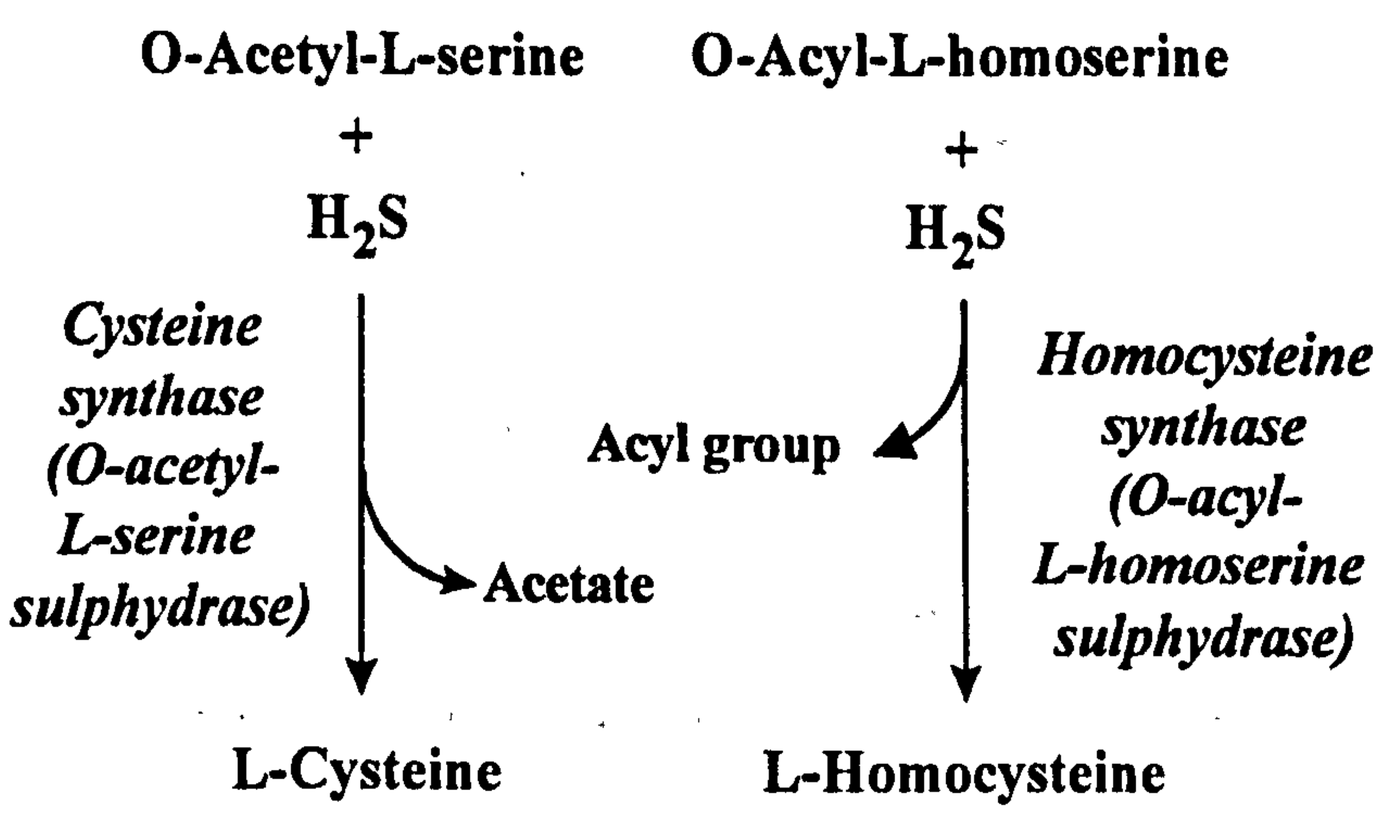

Figure 1.3: Mechanisms of sulphydrylation to produce cysteine and homocysteine. The acyl group attached to L-homoserine for the production of homocysteine varies (details are given in the following sections) , the production of cysteine always uses acetate.

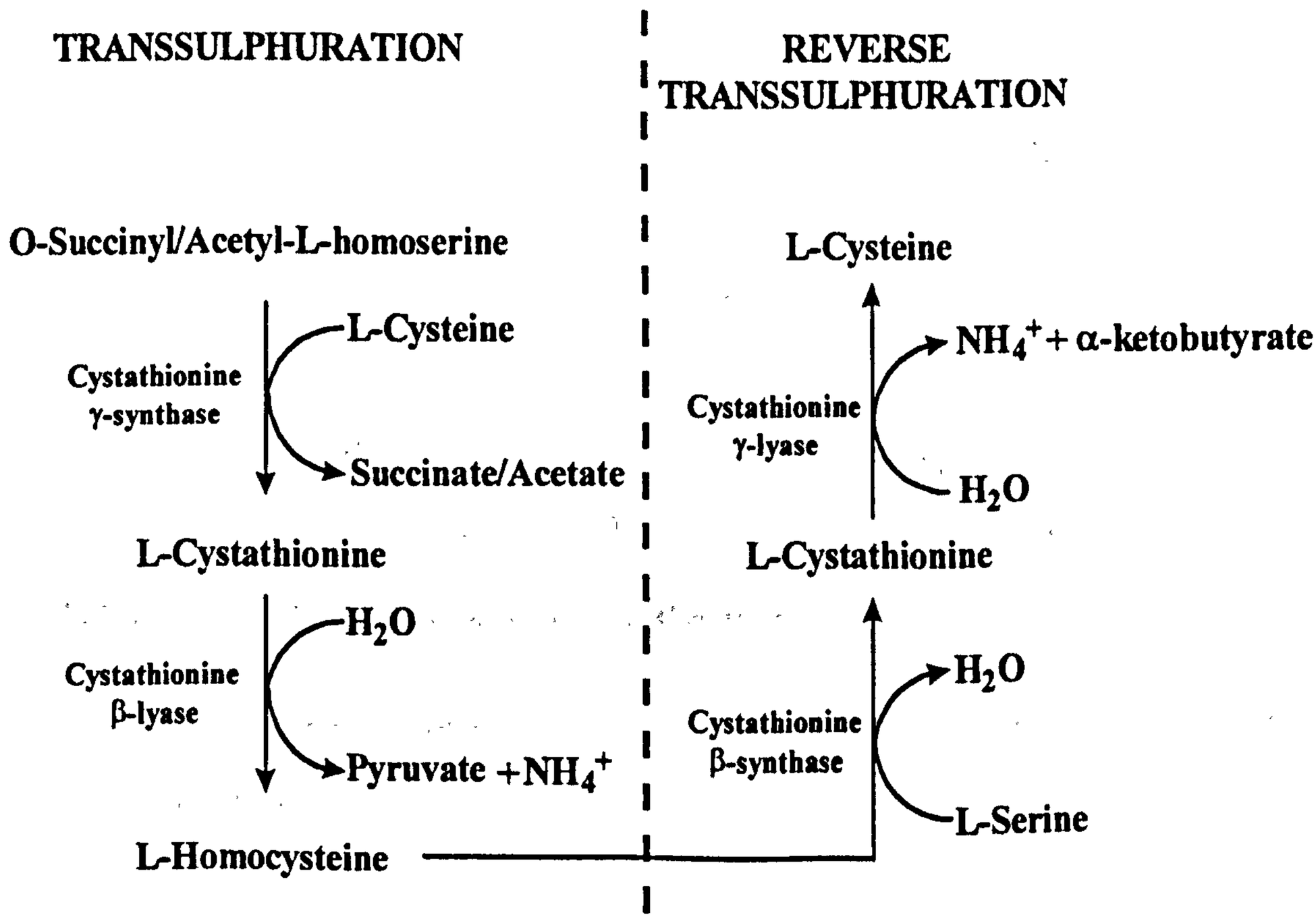

Figure 1.4: The interconversion of L-cysteine to L-homocysteine in both directions. The left side of the dashed line shows transsulphuration. Note cysteine is shown as a substrate for cystathionine γ-synthase as it provides the thiol group for homocysteine. Reverse transsulphuration is shown on the right hand side of the dashed line.

THE METHIONINE CYCLE

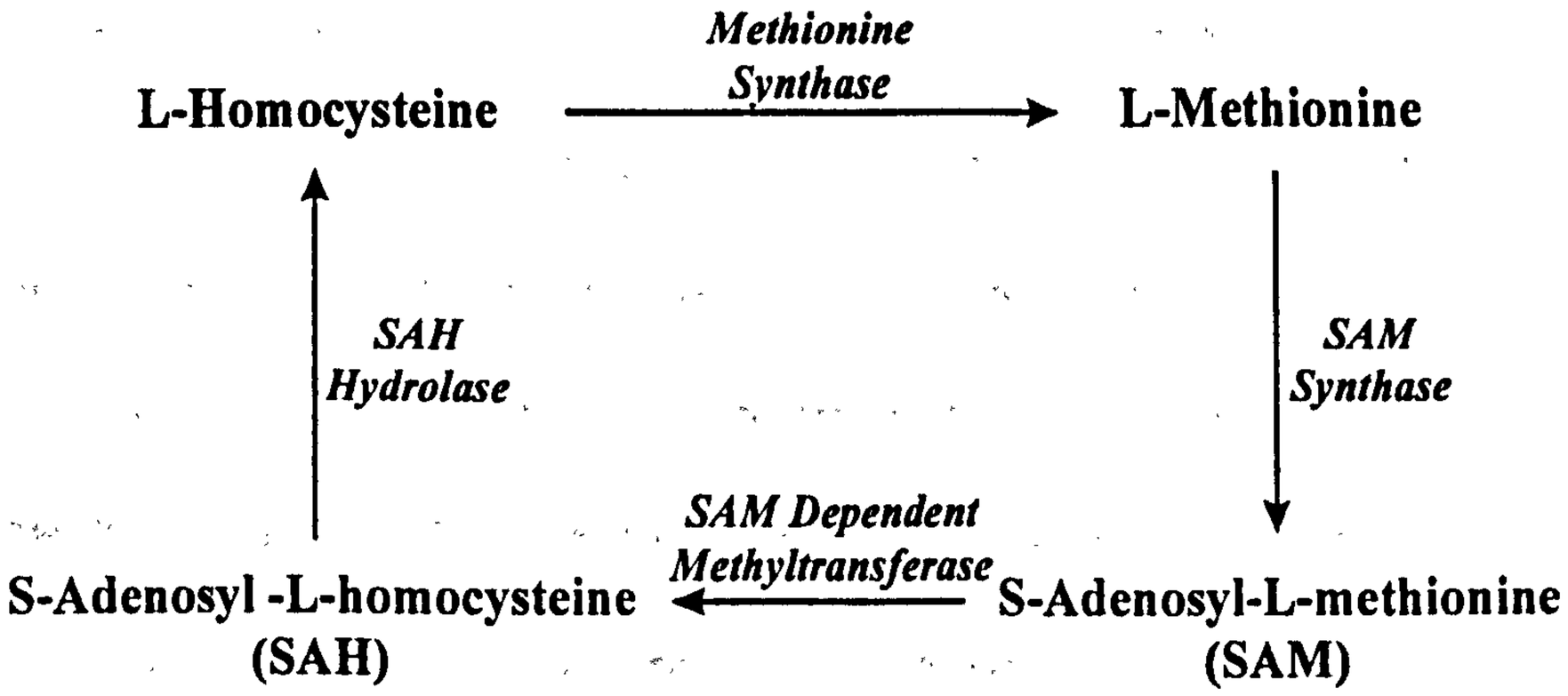

Figure 1.5: The methionine cycle showing the production of methionine from homocysteine, production of SAM and recycling back to homocysteine.

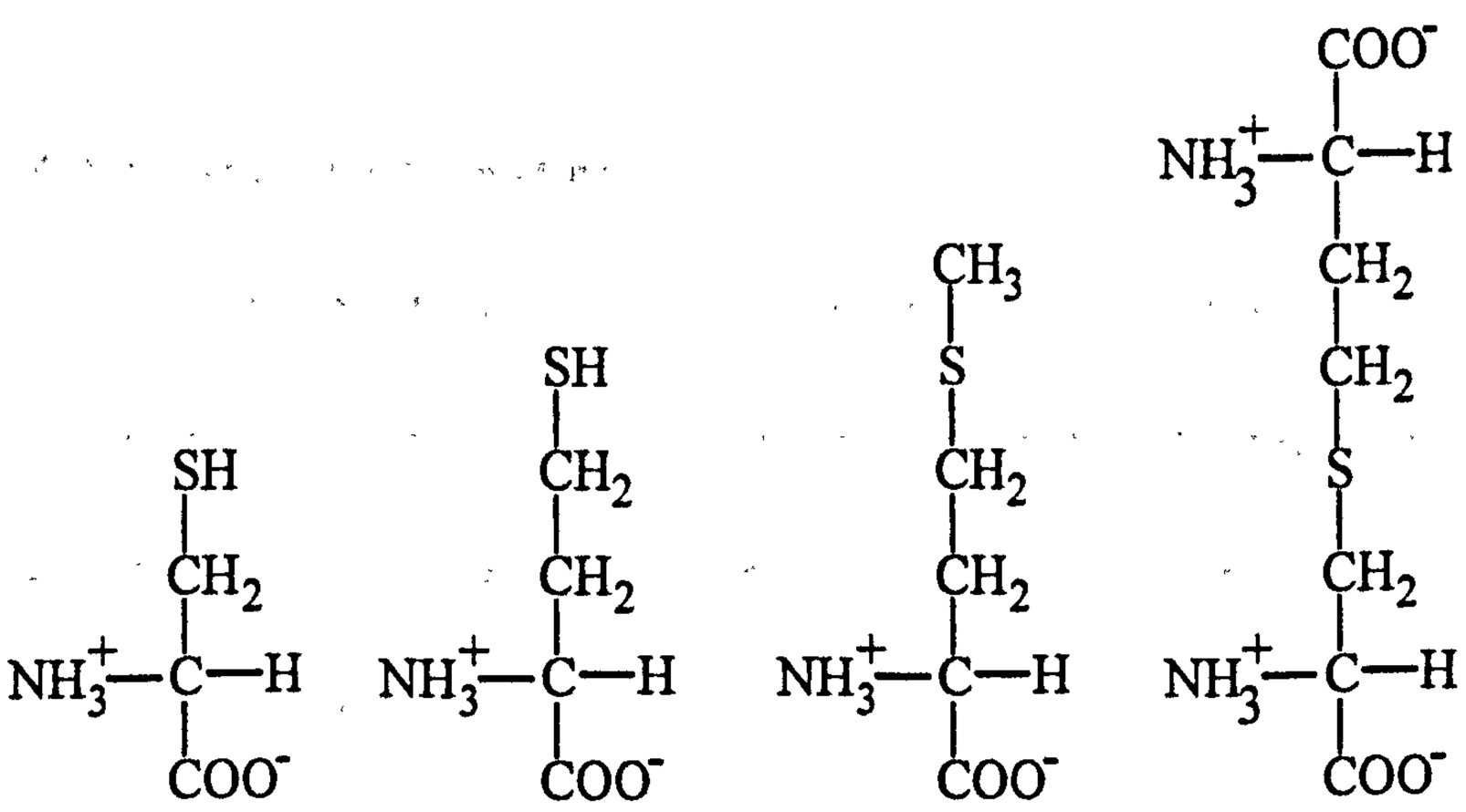

Figure 1.6: Structures of the two commonly occurring amino acids in proteins L-cysteine and L-methionine and their intermediates in metabolism L-homocysteine and L-cystathionine.

1.7.2.1 Sulphydrylation

Sulphydrylation is the name given to the process of synthesising cysteine and homocysteine from sulphide and a carbon backbone. The enzymes used for this

are O-acetyl-L-serine sulphydrase (OASS or cysteine synthase) and an O-acyl-homoserine sulphydrase (OAHSS or homocysteine synthase), respectively. The acyl group of the O- substituted molecule is replaced by sulphide in both cases. In cysteine synthesis this acyl group is acetate but for homocysteine it can be acetate, succinate or phosphate. *In vitro* it has been shown that both enzymes are able to utilise other sulphur donors, perhaps most notably methanethiol (CH_3SH) (Flavin and Slaughter, 1967). This would lead to the direct synthesis of methionine with OAHSS (a process not shown in figure 1.3). This reaction occurs much more slowly than that with H_2S and its relevance *in vivo* is unknown. Both of these enzymes are PLP (Pyridoxal 5'-Phosphate)-dependent and the use of this cofactor begins a common theme in sulphur amino acid metabolism.

1.7.2.2 Transsulphuration

The two step conversion between cysteine and homocysteine, in either direction, is referred to as the transsulphuration or reverse transsulphuration pathway, respectively. All four enzymes associated with this are closely related PLP-dependent enzymes.

In the formation of homocysteine, cysteine acts much like H_2S in that it provides the sulphur moiety. It is condensed with homoserine by the action of cystathionine γ-synthase (CGS) to form cystathionine. The molecule then undergoes cleavage by cystathionine β-lyase (CBL) to release homocysteine, pyruvate and ammonia. The reverse transsulphuration pathway utilises two very similar but distinct enzymes to those used in the opposite direction. Here homocysteine acts as the sulphur donor in the conversion to cysteine. It is

condensed with serine, catalysed by cystathionine β-synthase (CBS), to produce cystathionine prior to cleavage by cystathionine γ-lyase (CGL) to release cysteine, α-ketobutyrate and ammonia. CBS has been referred to as serine sulphydrase and has been shown to catalyse the production of cysteine from serine and H_2S *in vitro*. This activity is low though and not thought to be relevant *in vivo*.

The reactions catalysed by these enzymes are very similar and they each utilise PLP as a cofactor. Both pairs of enzymes (synthases and lyases) are sufficiently specific, however, that they are able to carry out reactions on opposite sides of the same atom without appreciably catalysing the other. Although *in vitro* it is possible to demonstrate these "side reactions", they are at very low rates and thought to be insignificant *in vivo*. By possessing this high degree of specificity a fruitless cycle of synthesis by the first enzyme and cleavage at the same point by the other is avoided. This is worth commenting on here both because of the elegance and specificity of the mechanism which will be discussed later, and also because of the large number of reactions that these enzymes are able to catalyse potentially *in vivo*.

1.7.2.3 The Methionine Cycle

Methionine is predominantly synthesised by the methylation of homocysteine by methionine synthase. As mentioned, OAHSS can utilise methanethiol as a sulphur donor in sulphydrylation to produce methionine directly but this is unlikely to be significant. The synthesis of methionine from homocysteine constitutes the first step in the methionine cycle. A critical role for methionine in biology is its part in the formation of SAM produced from methionine and ATP by SAM synthase.

Fulfilling one of its roles as a single carbon donor SAM donates its methyl group via a SAM dependent methyltransferase. The remaining molecule is S-adenosyl homocysteine (SAH). This is hydrolysed by SAH hydrolyase to adenosine and homocysteine, which completes the cycle. Homocysteine can then either be remethylated to again produce methionine or undergo reverse transsulphuration to produce cysteine. Thus homocysteine sits at a possible branch point regulating the levels of methionine and cysteine in the cell.

Not shown in the figures above to avoid confusion is the action of MGL. This enzyme is able to catabolise methionine, homocysteine and cysteine to varying degrees. It is thus able to deplete the cellular pool of these amino acids.

The presence or absence of any of these enzymes and processes in any parasites has not been addressed fully yet, nor are there specific details on the enzymes themselves except for the use of PLP by some. In the following sections their distribution will be considered followed by a detailed examination of published data on cysteine synthase and MGL.

1.7.3 Sulphur Amino Acids in Bacteria

Many bacteria have the ability to synthesise both cysteine and homocysteine through sulphydrylation and can assimilate inorganic sulphate to produce sulphide. The two reduction steps used to convert sulphate to sulphide utilise thioredoxin and NADPH (Siegel *et al.*, 1973;Tsang, 1983). However it has been shown for *E. coli* that other sulphur containing compounds such as sulphite,

thiosulphate, taurine and glutathione can act as the sole sulphur source (Kredich, 1996).

The synthesis of cysteine via sulphydrylation by cysteine synthase is well characterised in bacteria. The enzyme is PLP-dependent and uses OAS as the carbon backbone, synthesised from serine and acetylCoA by serine acetyltransferase (SAT). There are frequently two isoforms of cysteine synthase present and one in particular is found to copurify with SAT, a complex sometimes referred to as cysteine synthetase (Kredich *et al.*, 1969). The structure of an uncomplexed cysteine synthase from *S. typhimurium* has been published (Burkhard *et al.*, 1998). Synthesis of homocysteine by sulphydrylation occurs only in non-enteric bacteria. The identity of the acyl group associated with the homoserine carbon backbone varies. It is a succinate group in *P. aeruginosa* (Foglino *et al.*, 1995) and an acetate in *Brevibacterium* (Brush and Paulus, 1973), *Bacillus* (Brush and Paulus, 1971), *Pseudomonas syringae* (Andersen *et al.*, 1998). There is evidence that homocysteine synthase can utilise thiols other than hydrogen sulphide to produce other Cγ-Sδ-R compounds. For example methanethiol (CH_3SH), which leads to the production of methionine directly (Flavin and Slaughter, 1967). This direct sulphydrylation is slow and the relevance of this *in vivo* is unknown at this time.

The sequence databases contain many examples of the four enzymes of transsulphuration and there are numerous examples of purification and characterisation of the enzymes. Also the structures of the enzymes responsible for transsulphuration in *E. coli* have been solved (Clausen *et al.*, 1996, Clausen *et*

al., 1998). They have been proposed as possible drug targets because the bacteria lack the reverse transsulphuration pathway and the enzymes are not present in mammals. This lack of the reverse pathway occurs in all enteric bacteria but others such as *P. aeruginosa* possesses both the forward and reverse pathway and thus can grow equally well on either cysteine or methionine as the sole source of sulphur (Foglino *et al.*, 1995). The closely related bacteria *P. putida* lacks the key enzymes of reverse transsulphuration and transsulphuration is not very active. Perhaps to compensate for this apparently poor ability to interconvert cysteine and homocysteine, the organism is able to synthesise both of these through sulphydrylation (Vermeij and Kertesz, 1999). Interestingly the bacteria is one of the few organisms identified so far to possess MGL. It has been proposed that it may provide an alternative to reverse transsulphuration with methanethiol converted to a sulphonate and then sulphate before being reduced in he normal manner to provide sulphide (Vermeij and Kertesz, 1999).

The differences between species of bacteria alone highlights the variations that exist in sulphur amino acid metabolism and the difficulties in making generalised statements.

1.7.4 Plant Sulphur Amino Acid Metabolism

Plants are also able to assimilate inorganic sulphur for incorporation into amino acids in much the same way as bacteria (for a recent review see Ravanel *et al.*, 1998). The ability of plants to accomplish this is of undoubted importance to all mammals, with the possible exception of ruminants, as this provides the essential

source of dietary sulphur required in the form of methionine. One difference in this mechanism between plants and bacteria is the utilisation of glutathione for sulphate reduction by plants (Tsang and Schiff, 1976a) and thioredoxin in bacteria (Tsang and Schiff, 1976b).

In plants, as in bacteria, the synthesis of L-cysteine is achieved by condensation of sulphide and OAS by OASS. This enzyme has been studied thoroughly in plants and found to be similar to the enzyme in bacteria. However, whether there exists a bienzyme complex of SAT and OASS in all plants is less clear. The published evidence seems contrary in some cases. For example, purification of cysteine synthase from rape leaves (Masada *et al.*, 1975) and spinach leaves (Warrilow and Hawkesford, 1998) make no mention of co-purification with SAT. Others have shown the purification of both enzymes in a tight complex, for example in spinach chloroplasts (Droux *et al.*, 1992), *Allium tuberosum* (Nakamura and Tamura, 1990) and water melon (Saito *et al.*, 1995). Analysis using the yeast two-hybrid system has proposed association of the enzymes from *Arabidopsis thaliana in vivo* (Bogdanova and Hell, 1997) but later work using *A. thaliana* proteins expressed recombinantly in *E. coli* have claimed this is not the case (Droux *et al.*, 1998). In common with bacterial cysteine synthase, multiple isoforms are found in plants. These can be classified into distinct isoforms which are intercellularly localised into plastids, mitochondria and cytosol (Lunn *et al.*, 1990; Rolland *et al.*, 1992; Yamaguchi *et al.*, 1998). Four species have been cloned from rice (Nakamura *et al.*, 1999) and at least three from *A. thaliana* (Hell *et al.*, 1994; Barroso *et al.*, 1995; Hesse and Altmann, 1995; Nakamura *et al.*, 1997). Although there are obvious organelle distributions for the enzymes (i.e.

chloroplast,) in plants there are also multiple isoforms present in a single cellular location.

In plants cystathionine is produced by the condensation of cysteine with O-phosphohomoserine rather than an acetate or succinate substituted molecule (Datko, *et al.*, 1974). Homocysteine is generated in an identical fashion to bacteria with the α,β elimination of cystathionine by CBL (also producing pyruvate and ammonia). The production of homocysteine by direct sulphydrylation *in vivo* is unlikely in plants. CGS has been shown *in vitro* to be able to utilise sodium sulphide as the sulphur donor to produce homocysteine (Ravanel *et al.*, 1995). However it has been calculated that this would be only able to provide 3% of the homocysteine requirements (MacNicol *et al.*, 1981). There is thought to be no reverse transsulphuration pathway in plants as no CBS or CGL genes or proteins have been identified.

Plants are also unusual in the high concentration of uncommonly occurring sulphur amino acids present. These substituted cysteine and methionine compounds such as S-methylcysteine, homomethionine and various sulphoxide derivatives have largely unknown functions.

1.7.5 Yeast and Fungi Sulphur Amino Acid Metabolism

It had been thought for some time that *S. cerevisiae* was able to synthesise cysteine by two distinct pathways. Either via reverse transsulphuration from homocysteine and also by a dual action O-acetylserine sulphydrylase/O-acetylhomoserine sulphydrylase (Yamagata *et al.*, 1974). This latter mechanism

is now thought not to occur for the production of cysteine *in vivo* (Ono *et al.*, 1999). Instead only the sulphydrylation of O-acetyl-L-homoserine is of physiological significance. The OAHSS responsible for this has also been characterised in *S. lipolytica* (Morzycka and Paszewski, 1979), *Aspergillus nidulans* (Paszewski and Grabski, 1974) and *Neurosporra crassa* (Piotrowska *et al.*, 1980; Jacobson and Metzenburg, 1977). Thus it would appear that cysteine synthesis does not occur through direct sulphydrylation but from reverse transsulphuration only. OAHSS has also been shown *in vitro* to produce methionine and ethionine when sulphide was substituted for the appropriate thiols (Yamagata, 1971). However the relevance of the abilities of this enzyme *in vivo* is unknown at this time.

The enzymes required for reverse transsulphuration have been found in these organisms. CBS from *S. cerevisiae* has been purified (Ono *et al.*, 1994) and CGL cloned and expressed recombinantly in *E. coli* (Yamagata *et al.*, 1993). This would allow for the production of cysteine from homocysteine. Conversion of cysteine back to homocysteine is also possible as CBL and CGS have also been identified. There is no evidence of MGL in yeast or fungi. For a comprehensive review on sulphur amino acid metabolism in *S. cerevisiae* the reader is directed to Thomas and Surdin-Kerjan (1997).

1.7.6 Sulphur Amino Acids in Mammals

The most obvious and key difference in the metabolism of sulphur containing amino acids in mammals to lower organisms is the absence of sulphydrylation. Methionine but not cysteine is commonly referred to as an essential amino acid in

mammalian nutrition. This is due to the lack of transsulphuration (conversion of cysteine to homocysteine) but the presence of the reverse transsulphuration pathway. Thus mammals can produce sufficient cysteine for requirements provided the dietary source of methionine is adequate. This is synthesised in the same manner as described for bacteria utilising CBS and CGL. For a summary see Griffith, 1987.

1.7.7 Parasite Sulphur Amino Acid Metabolism

As with many other aspects of their biology, the knowledge of the processes of sulphur amino acid metabolism in parasites is much scarcer than for other organisms. For a review see Walker and Barrett, 1997. The details of sulphur amino acid metabolism in parasites other than *T. vaginalis* will be considered in this section.

There is evidence for direct sulphydrylation in *Entamoeba*. Two cysteine synthase genes have been cloned and partially characterised from both *E. histolytica* (Nozaki *et al.*, 1998) and *E. dispar* (Nozaki *et al.*, 1999). There are no other published accounts of the purification or cloning of enzymes capable of sulphydrylation in parasites although data on activities attributed to CBS have been published. This activity termed "activated serine sulphydrase" has been recorded in *E. histolytica* and *Leishmania* spp. (Thong and Coombs, 1985b) and also species of nematodes (Walker and Barrett, 1991). The enzyme responsible for the "activated serine sulphydrase" activity has also been partially characterised from a free-living nematode source. Described as a novel CBS enzyme it is able to catalyse the production of cystathionine from homocysteine

and serine or cysteine (Walker *et al.*, 1992; Papadopolous *et al.*, 1996). There is also evidence of an additional CBS in nematodes. Cystathionine synthesis and catabolism activity has also been detected in *Hymenolepis diminuta* (Gomez-Bautista and Barrett, 1988) and *Fasciola hepatica* (Bankov *et al.*, 1996). The treatment of African trypanosomes with metabolic inhibitors and incubation with excess L-methionine have both shown increased intracellular levels of reverse transsulphuration products (cystathionine and cysteine) (Bacchi, *et al.*, 1995; Goldberg *et al.*, 2000). There does therefore exist part of the transsulphuration pathway in some parasites but the complete pathway has not been described.

The role of SAM in one carbon donor reactions is likely to exist in a number of parasites. Several enzymes involved in the metabolism of SAM have been characterised. This includes SAM decarboxylase from *Onchocerca volvulus* (Da'Darra *et al.*, 1996), *Trypanosoma cruzi* (Persson *et al.*, 1998) and a bifunctional SAM decarboxylase/ornithine decarboxylase from *P. falciparum* (Rathaur and Walter, 1987; Müller *et al.*, 2000). Also SAH hydrolyase has been identified in both *Leishmania donovani* (Henderson *et al.*, 1992) and *Plasmodium falciparum* (Creedon *et al.*, 1994). In addition to this part of the methionine cycle, methionine γ-lyase has also been shown in whole parasite lysates of *E. histolytica* (Lockwood and Coombs, 1991b).

1.7.8 Sulphur Amino Acid Metabolism in *T. vaginalis*

The synthesis and fate of sulphur amino acids in *T. vaginalis* is the central area of investigation of this project. Several processes and enzymes had been

characterised from the parasite prior to this study and these will be discussed here in detail.

As with other organisms methionine is required for the synthesis of SAM by the parasite. Radiolabelled methionine has been shown to be incorporated into living trichomonads (Thong, 1985) The radiolabelled carbon was detected in SAM and also nucleic acids and lipids presumably denoting its role as one carbon donor. The parasite has been shown to possess methionine adenosyl transferase (Yarlett, 1988) and SAM methyltransferases (Thong *et al.*, 1987). S-Adenosyl homocysteine hydrolyase activity has also been detected (Thong *et al.*, 1985) and the gene cloned and expressed (Bagnara *et al.*, 1996; Minotto *et al.*, 1998). Although the presence of methionine synthase has not been confirmed it would complete the cycle of production of methionine from homocysteine and eventual conversion back to homocysteine.

No method of direct sulphydrylation has yet been published for *T. vaginalis* either for the production of cysteine or homocysteine. The presence of an "activated" serine sulphydrylase in species of trichomonads has been shown in crude lysates (Thong and Coombs, 1985b) but not characterised. The production of hydrogen sulphide from this enzyme is stimulated by the presence of β-mercaptoethanol, a situation also found in nematodes (Walker and Barrett, 1991). In *T. vaginalis* this activity has been shown to be attributable to several isoforms (Thong and Coombs, 1987).

CGL activity has been reported in *T. vaginalis* lysates (Thong, 1985). This was at very low levels however in a crude cell preparation and has since not been investigated. Methionine γ-lyase has been purified and cloned from the parasite and as this is a major focus of this study it will be covered in detail in the next chapter.

1.8 Pyridoxal 5'-Phosphate (PLP) Dependent Enzymes

PLP-dependent enzymes are heavily involved in amino acid metabolism in general and that of sulphur amino acids in particular. The following sections of this introduction give a brief overview of these enzymes followed by a summary of those involved in this study.

Pyridoxal 5'-phosphate, a derivative of pyridoxal or vitamin B6, acts as a cofactor in a very large number of enzymes. These PLP-dependent enzymes are ubiquitous in eukaryotes and prokaryotes and often play central roles in amino acid metabolism. The unique chemistry of the cofactor, which enables it to act as an electron "sink", conveys its versatility. The most commonly utilised substrates are amino acids and when conjugated to PLP by their amine group the α-carbon is labilised. That is, there is the potential to break any of the four bonds around the Cα. For this reason PLP-dependent enzymes are able to catalyse transamination, decarboxylation, racemisation, elimination and replacement reactions. Not only can changes take place at the Cα but also along the amino acid side chain as well at the β and γ positions. The specific architecture of the

enzymes active site dictates the catalytic event promoting the occurrence of the desired reaction, whilst excluding almost entirely all others.

There are a very large number of sequences available for PLP-dependent enzymes in the databases, a situation increased dramatically by the recent genome sequencing projects. This has only recently been complemented with structural information on these enzymes. Classification has been difficult up until this time because of the low sequence homology even between related enzymes.

A simple and largely accurate classification has been suggested in which these enzymes can be broadly subdivided into three classes, α, β and γ (Alexander *et al*, 1994). This was based upon the chemical characteristics of the reaction, corresponding to the position at which they act upon their amino acid substrates side chain, and correlated with sequence data. From sequence analysis it was predicted that the α and γ families were more closely related to each other than the β family. Some PLP-dependent enzymes though do not fit this classification and it is thus not comprehensive as it stands. A classification based upon structural data has also been proposed (Grishin *et al.*, 1995). At this point five different folds were observed and through sequence analysis it was predicted that further unknown folds existed. A recent review of PLP-dependent enzymes presents the most current attempt to summarise the available data and classify the enzymes (Jansonius, 1998).

1.8.1 Three Dimensional Structures of PLP-Dependent Enzymes

The first structure of a PLP-dependent enzyme to be solved was that of aspartate aminotransferase (AAT) in 1980 (Ford *et al.*, 1980). This enzyme has since been used as a model system and is the most extensively studied of all such enzymes. Currently, a search of the structural database reveals a total of 36 sets of atomic coordinates for AAT. These include enzyme-complexes with various compounds, enzymes from different organisms and numerous mutants. The total number of sets of coordinates for all PLP-dependent enzymes, complexes, mutants, etc. now exceeds 120. This is an even more remarkable number when one considers that the second enzyme structure to be solved, that of tryptophan synthase (TRPS), was not published until 1988 (Hyde *et al.*, 1988). Since then a rapidly accelerating number of new structures have been solved (following the general trend of numbers of structures being solved) and with this the opportunity to assign each to an evolutionary family. Many PLP-dependent enzymes are multimers and normally active as dimers of at least 80 kDa. The large size of these molecules have undoubtedly caused many problems in the past in solving structures. Table 1.1 lists the different enzymes for which the three-dimensional structures have been solved.

Enzyme	Source	Reference
Aspartate aminotransferase	Chicken mitochondria	Ford *et al.*, 1980
	Chicken cytosol	Malashkevich *et al.*, 1995
	Pig cytosol	Malashkevich *et al.*, 1995
	E. coli	Okamoto *et al.*, 1994
	S. cerevisiae	Jeffrey *et al.*, 1998
	Thermus thermophilus	Nakai, *et al.*, 1999
Tryptophan synthase	*Salmonella typhimurium*	Hyde *et al.*, 1988
ω-Amino acid:pyruvate aminotransferase	*Pseudomonas* sp.	Watanabe *et al.*, 1989
Glycogen phosphorylase	Rabbit cytosol	Martin *et al.*, 1990
Dialkylglycine decarboxylase	*Pseudomonas capacia*	Toney *et al.*, 1993
Tyrosine phenol-lyase	*Citrobacter freundii*	Antson *et al.*, 1993
	Erwinia herbicola	Pletnev *et al.*, 1997
D-Amino acid aminotransferase	*Bacillus* sp.	Sugio *et al.*, 1995
Ornithine decarboxylase	*Lactobacillus* 30a	Momany *et al.*, 1995
	Trypanasoma brucei	Grishin *et al.*, 1999
	Mouse	Kern *et al.*, 1999
Cystathionine β-lyase	*E. coli*	Clausen *et al.*, 1996
Alanine racemase	*Bacillus stearothermophilus*	Shaw *et al.*, 1997
Glutamate-1-semialdehyde aminomutase	*Synechoccus*	Hennig *et al.*, 1997
O-Acetylserine sulphydrylase	*S. typhimurium*	Burkhard *et al.*, 1998
8-Amino-7-oxononanoate synthase	*E. coli*	Alexeev *et al.*, 1998
Aromatic amino acid aminotransferase	*Paracoccus denitrificans*	Okamoto *et al.*, 1998
	Trypanasoma cruzi	Blankenfeldt *et al.*, 1999
	E. coli	3TAT

Cystathionine γ-synthase	*E. coli*	Clausen *et al.*, 1998
	Nicotiana tabacum	Steegborn *et al.*, 1999
Ornithine aminotransferase	*Homo sapiens*	Shen *et al.*, 1998
Serine hydroxymethyltransferase	*Homo sapiens*	Renwick *et al.*, 1998
	Rabbit cytosol	Scarsdale *et al.*, 1999
	E. coli	Scarsdale *et al.*, 2000
Threonine deaminase	*E. coli*	Gallagher *et al.*, 1998
Tryptophan indole-lyase	*Proteus vulgaris*	Isupov *et al.*, 1998
AHBA synthase	*Amycolatopsis mediterranei*	Eads *et al.*, 1999
1-Aminocyclopropane-1-carboxylate synthase	Apple	Capitani *et al.*, 1999
Anthranilate synthase	*Sulfolobus solfataricus*	Knochel *et al.*, 1999
Diaminopelargonic acid synthase	*E. coli*	Kack *et al.*, 1999
Phosphoserine aminotransferase	*E. coli*	Hester *et al.*, 1999
	Bacillus sp.	1BT4
CsdB (SCL)	*E. coli*	Fujii *et al.*, 2000
Cystalysin	*Treponema denticola*	Krupka *et al.*, 2000
Cystine lyase (C-DES)	*Synechocystis*	Clausen *et al.*, 2000a
MalY	*E. coli*	Clausen *et al.*, 2000b
NifS	*Thermotoga maritima*	Kaiser *et al.*, 2000

Table 1.1: List of structures published with PLP cofactor. Only the first example of any structure reported is referenced. Where there is no publication available on a deposited set of coordinates the PDB accession code is given.

At the beginning of this study very few of those listed above were in the database (though not necessarily available). CBL was the only representative of the proposed γ-family and since that time only two other examples have been published, those of CGS (Clausen *et al.*, 1998; Steegborn *et al.*, 1999). CBL was

also the only example of a PLP-dependent enzyme acting upon a sulphur containing amino acid. Since that time though four structures have been published that act upon cysteine (or cystine), CsdB (Fuji *et al.*, 2000), cystalysin (Krupka *et al.*, 2000), C-DES (Clausen *et al.*, 2000) and NifS (Kaiser *et al.*, 2000), in addition to CGS.

1.8.2 Methionine γ-Lyase (MGL)

In *T. vaginalis* the breakdown of L-methionine into ammonia, α-ketobutyrate and methanethiol has been shown to be catalysed by the PLP-dependent enzyme methionine γ-lyase (MGL) (Thong and Coombs, 1987, Lockwood and Coombs, 1991b). The same enzyme is also responsible for the even more rapid breakdown of L-homocysteine to ammonia, α-ketobutyrate and hydrogen sulphide, formerly described as homocysteine desulphurase (Thong and Coombs, 1985). The enzyme has been purified from the parasite (Lockwood and Coombs, 1991b) and two genes cloned and expressed in a prokaryotic system (McKie *et al.*, 1998). As mentioned previously the enzyme is also present in *P. putida* and has been purified from the bacteria (Esaki and Soda, 1987) and two genes cloned in an analogous fashion to *T. vaginalis* (Inoue *et al.*, 1995; Hori *et al.*, 1996). Other sources from which the enzyme has been purified include *Aeromonas* sp. (Nakayama *et al.*, 1984), *Brevibacterium linens* (Dias and Weimer, 1998) and *Clostridium sporogenes* (Kreis and Hession, 1973). A further report of purification of MGL from *T. vaginalis* has incorrectly claimed isolation of a further enzyme to the two already published (Han *et al.*, 1998). The differences in this are very minor (two amino acid changes) and presumably strain specific or PCR based mutations. Activities for MGL have also been reported in a number of soil

bacteria, *E. histolytica* and rumen ciliates (Lockwood and Coombs, 1991b). There is no mammalian homologue of the enzyme and it is apparently absent from yeast and other eukaryotes as well. Trace levels of homocysteine desulphurase activity in mammals have been attributed to the action of cystathionine γ-lyase and thought to be of little significance (Cooper, 1983).

Characterisation of the enzyme has revealed it to be active as a multimer. Individual subunits have an M_r of 40-44 000 and the active enzyme has an estimated M_r of >150 000. This suggests a tetrameric enzyme and the activity of recombinant enzyme from *P. putida* and *T. vaginalis* suggests that this is likely to be a homotetramer. Whether this is the case *in vivo* has not been shown. PLP is present at a ratio of 1 mole per mole of enzyme subunit in each case. Relative levels of activity differ between sources but the enzyme is able to catalyse α,γ-elimination reactions on a range of C^{γ}-S^{δ} containing amino acids such as methionine, ethionine and homocysteine. To a lesser extent, similar activities have been recorded against C^{β}-S^{γ} substrates such as cysteine too.

The *in vivo* role of the enzyme is unknown at present although there is evidence that the α-keto acids produced may be utilised as an alternate energy source (Lockwood and Coombs, 1991b). This has also been suggested in *P. putida* because of its gene arrangement within an operon coding for part of an α-keto acid dehydrogenase (Inoue *et al.*, 1997). The specific role of the large amounts of low molecular weight thiols produced is unknown but is at least partly responsible for the characteristic odour produced by *T. vaginalis* when grown *in vitro*.

1.8.3 Cysteine Synthase

As already mentioned cysteine synthase or OASS is present in many bacterial species, plants and also recently discovered in *Entamoeba*. It catalyses the condensation of O-acetyl-L-serine and sulphide to produce L-cysteine and acetate (for a recent review see Tai and Cook, 2000). The enzyme has been purified from a large number of species and also cloned and expressed recombinantly. It is active as a homodimer and the structure of the enzyme from *S. typhimurium* has been solved (Burkhard *et al.*, 1998). As with MGL and the various cysteine desulphurase enzymes, it is a PLP-dependent enzyme. Unlike these others, though, it belongs to the β-family of PLP-dependent enzymes with a fold similar to tryptophan synthase (Hyde *et al.*, 1988).

For both plant and bacterial enzymes there is evidence for cysteine synthase existing as a complex with serine acetyl transferase (SAT), the previous enzyme in the cysteine biosynthetic pathway. The evidence available seems to suggest that in appropriate circumstances certain isoforms of the enzyme associate with SAT to form the "cysteine synthetase" complex. *In vivo* the association of the plant enzyme is dependent on the concentration of both OAS and sulphide (Droux *et al.*, 1998) and SAT forms high molecular weight complexes in the absence of cysteine synthase.

1.8.4 Cysteine Desulphurase

Cysteine desulphurase is a PLP-dependent enzyme that catalyses the removal of sulphur from cysteine to leave alanine. It is also able to utilise selenocysteine as substrate (Lacourciere and Stadtman, 1998). Note that this is a different reaction

from that of MGL which further reduces the carbon backbone to pyruvate and ammonia in the lysis of cysteine. The enzyme is alternatively referred to as NifS in reference to the gene product from *Azotobacter vinelandii* (Zheng *et al.*, 1993). Homologues have been identified in *E. coli* (Mihara *et al.*, 1997), *S. cerevisiae* (Kolman and Söll, 1993), *Caenorhabditis elegans* (Wilson *et al.*, 1994) and mice (Nakai *et al.*, 1998) amongst others. There does also exist an analogue first identified in pigs highly specific for selenocysteine termed selenocysteine lyase (SCL)(Esaki *et al.*, 1982). A homologue for this enzyme has been identified in *E. coli* termed CsdB (Mihara *et al.*, 1999). NifS has been shown to catalyse the formation of FeS clusters in nitrogenase to create a functional enzyme (Zheng *et al.*, 1993). It does this by forming an internal persulphide complex between the released substrate sulphur and a conserved cysteine residue within the enzyme (Zheng *et al.*, 1994). It is then believed that this would then be passed onto an unknown carrier molecule prior to use in FeS cluster formation or repair. There is, in addition, a different enzyme that can direct FeS cluster assembly. This enzyme isolated from *Synechocystis* has been named C-DES and has a different substrate profile and proposed mode of action to NifS proteins. Instead of forming an internal enzyme persulphide complex, C-DES preferentially utilises cystine (dimerised cysteine) and releases pyruvate, ammonia and cysteine persulphide (Lang and Kessler, 1999). This cysteine persulphide is then presumed to be released and act as a sulphur atom donor in an analogous fashion to the enzyme bound persulphide of NifS enzymes. Recently the structures of several of these enzymes have been solved and published: a NifS-like protein from *Thermotoga maritima* (Kaiser *et al.*, 2000), CsdB (a selenocysteine lyase) from *E. coli* (Fujii *et al.*, 2000) and C-DES from *Synechocystis* (Clausen *et al.*, 2000). All three

enzymes are members of the aminotransferase family, structurally similar to aspartate aminotransferase.

Another different cysteine lyase enzyme from *Treponema denticola* has also been characterised (Chu *et al.*, 1997). This enzyme is more similar to MGL in that it also catalyses an α,β-elimination reaction catabolising cysteine to pyruvate, ammonia and H_2S. Another interesting observation is that it is proposed to be at least partly responsible for the haemolytic properties of the organism. A feature also found in *T. vaginalis* pathology. Recently the structure of this enzyme has been reported (Krupka *et al.*, 2000) and confirmed its structural similarity to other PLP-dependent aminotransferases. The active site is more structurally similar to CBL and CGS reflecting the reaction similarity.

1.9 Aims of The Project

The overall aim of this study was to further understand sulphur amino acid metabolism in *T. vaginalis*. In particular the roles of MGL and cysteine synthase.

To do this three practical goals were set:

1. Crystallise and solve the structure of MGL
2. By site-directed mutagenesis analyse the mechanism of MGL and relate this to potential roles *in vivo*
3. Produce and purify recombinant cysteine synthase cloned from *T. vaginalis* and investigate its role *in vivo*

CHAPTER 2. MATERIALS AND METHODS

2.1 Crystallisation and Structure Solution Materials and Methods

2.1.1 MGL Purification

Enzyme for use in crystallisation was purified as published (McKie *et al.*, 1998) with minor revisions as indicated in section 3.2.1.

2.1.2 Enzyme Sample Preparation

Freshly prepared MGL1 and 2 for crystallisation was concentrated using a 10 kDa cut-off Microsep™ centrifugal concentrator (Flowgen) and simultaneously washed to remove glycerol and replace the buffers. Firstly the enzyme sample was concentrated to a volume of approximately 0.25 ml. 3 ml of 100 mM HEPES pH 7.5, 250 μM PLP, 1 mM DTT was added and the sample concentrated by centrifugation until the volume of the retained sample was approximately 0.25 ml. This process was repeated three times to ensure removal of glycerol, salts, etc. After concentration of the sample to a suitable volume, protein concentration was determined and the sample resuspended in the same buffer to 6 mg/ml.

2.1.3 Crystallisation Solutions

Sparse matrix screens were prepared in house from published recipes, e.g. Magic 50 (Jancarik and Kim, 1991). All solutions were filtered through a 0.22 μm syringe filter to reduce microbial contamination.

2.1.4 Sitting Drop Crystallisation

24 well sitting drop plates were used for crystallisations (NBS biologicals, UK). 1 ml of the appropriate crystallisation solution was added to the reservoir. For each condition, 1.5 μl enzyme solution was added to 1.5 μl reservoir solution and mixed by pipetting. Wells were sealed using transparent adhesive tape.

2.1.5 Dynamic Light Scattering

Monodispersity of the concentrated enzyme solution was examined by dynamic light scattering using a Dynapro-801 (Protein Solutions, UK). Enzyme was concentrated to >10 mg/ml in 25 mM HEPES pH 7.5 in an identical fashion to that described in section 2.1.2. Enzyme was then resuspended at a concentration of 1 mg/ml in the desired solution. Samples of 200 μl were loaded into the instrument by syringe through a 0.1 μm filter.

2.1.6 Data Processing and Scaling

All data were processed with DENZO and scaled using SCALEPACK (Otwinowski and Minor, 1997).

2.1.7 Molecular Replacement

Molecular replacement was performed using the programme AMoRe (Navaza,1994) with data from 10 to 4.5 Å and the structure of CBL (Clausen *et al.*, 1996) used as a search model.

2.1.8 Refinement

Refinement of all MGL1 structures in this study was carried out using REFMAC (Murshdov *et al.*, 1997) using the Free R-factor (Brünger, 1992) to monitor progress. 5% of randomly assigned reflections were selected for the Free R set and excluded from all stages of refinement. A bulk-solvent correction (Kostrewa, 1999) was used in later stages of refinement. Geometric restraints were calculated using PROTIN, new entries for PLP and the PLP-propargylglycine complex were produced using MAKEDIC. Maps were calculated from the output of REFMAC using the CCP4 programme FFT. Waters were assigned initially using the programme ARP (Lamzin and Wilson, 1993).

2.1.9 Model Building

All model building was done using the programme QUANTA (Molecular Simulations Inc., UK). Early rounds of model building fitted only one monomer and advantage was taken of the 2-fold NCS in the asymmetric unit. The maps produced by refinement were then averaged over a single monomer using the CCP4 programme MAPROT and the process repeated. This was performed until density was visible for all main chain atoms and subsequently monomers were checked individually.

2.2 Molecular Biology Materials and Methods

2.2.1 Bacterial Maintenance, Cultivation and Harvesting

Escherichia coli strains used in this study were XL1-Blue (Stratagene) and JM109 (Promega) for cloning and M15pREP4 (Qiagen) and BL21DE3 (Novagen) for expression of recombinant protein. Stocks of bacterial strains were maintained by

combining an overnight culture 1:1 with a 40% glycerol and 2% peptone solution and stored at -70 °C. Overnight bacterial cultures were always produced by picking a single discrete colony and using this to inoculate an aliquot of LB. Colonies were produced by streaking out a sample of bacterial stock onto a solid LB-agar plate with appropriate antibiotics. Bacteria were harvested by centrifugation for 5 minutes at 13 000 rpm in a benchtop microfuge or, for large scale cultures, 15 minutes at 4000 g. Antibiotics were used at the following concentrations: ampicillin 100 μg/ml, kanamycin 25 μg/ml.

2.2.2 Plasmid Purification

Plasmid DNA was routinely isolated from a 3 ml overnight culture using either Wizard™ mini prep kits (Promega) or Qiaspin columns (Qiagen). For automated sequencing, Qiaspin columns were used and plasmid resuspended in double distilled deionised water (dddH_2O).

2.2.3 Polymerase Chain Reaction (PCR) Amplification

PCR amplifications of DNA fragments which engineered restriction sites onto the ends each contained 5 ng template DNA, oligonucleotide primers at 1 mM, 0.1 mM dNTPs, 0.75 mM $MgCl_2$, reaction buffer as supplied by the manufacturer and 5 U of DNA polymerase. Vent™ DNA polymerase (New England Biolabs) was used because of its proof-reading ability to minimise mutations inserted during amplification. *Taq* DNA polymerase (Promega) was used for routine amplifications. *Pfu* polymerase (Stratagene) was used for site-directed mutagenesis experiments. Individual protocols including thermal cycling parameters will be referred to in the text as required. All PCR was carried out on a Perkin Elmer GeneAmp PCR System 2400.

2.2.4 Isolation of *T. vaginalis* Genomic DNA

Genomic DNA (gDNA) was purified from *T. vaginalis* using a phenol/chloroform extraction. Parasites in log phase were harvested by centrifugation at 4000 g for 15 minutes at 4 °C and washed twice in pre-chilled 1 x PBS. The resulting cell pellet was resuspended in extraction buffer (50 mM TRIS-HCl pH 8.0, 50 mM EDTA, 100 mM NaCl) with the addition of 1% SDS and 4 mg proteinase K in a final volume of 2 ml. This was then incubated at 37 °C for 2 hours under gentle agitation. The lysate was combined 1:1 with phenol and mixed gently for 5 minutes before centrifugation at 1500 g for 15 minutes. The upper, DNA containing aqueous phase, was carefully removed using a wide bore pipette and the remaining organic solvent discarded. The aqueous phase was then mixed gently with a phenol/chloroform solution (1:1 v/v) and centrifuged and separated as before. Finally the aqueous phase was combined 1:1 with chloroform and mixed gently before centrifugation as before. High molecular weight genomic DNA was then precipitated by adding 0.1 volumes of 3 M sodium acetate pH 5.2 and 2.5 volumes of 100% ethanol. The visible gDNA was spooled using a pipette tip to obtain the largest possible fragments and avoid co-purification of RNA. The precipitated gDNA was resuspended in TE buffer and incubated at 37 °C overnight with 1 mg/ml proteinase K before repeating the phenol/chloroform extraction procedure and ethanol precipitation. High levels of EDTA and proteinase K were deliberately used in the extraction buffer to minimise nuclease degradation of gDNA. The second proteinase K treatment removed contaminating proteins to aid in the complete enzymatic digest of the DNA for Southern analysis.

2.2.5 Restriction Endonuclease Digest of DNA

Plasmid DNA digests were routinely done using an excess of enzyme to ensure a rapid complete digest. Generally up to 200 ng of plasmid DNA was digested in a total volume of 10 µl using 5 to 10 units of restriction enzyme for 1 to 2 hours at 37 °C. Restriction enzymes were obtained from various commercial companies and used with compatible buffers as supplied. *T. vaginalis* gDNA was digested in a total volume of 40 µl with 40 units of restriction enzyme. 30 units were added initially and incubated for 3 hours at 37 °C before the addition of a further 10 units and continued incubation for another hour.

2.2.6 DNA Ligations

Ligation of PCR products into pGEM-T vector (Promega) were performed according to the manufacturers instructions with ligase and buffer as supplied. PCR products were purified after separation on 1 % agarose gel and DNA concentration estimated from known DNA standards. Sub-cloning of DNA fragments into restriction digested plasmid was performed using commercially obtained T4 DNA ligase (New England Biolabs or Promega) and buffer as supplied. Ligations were performed in a minimal volume, generally 10 µl, with varying ratios of insert to vector at either 16 °C or 4 °C overnight.

2.2.7 Competent Cells

Chemical competent cells were commercially supplied XL-1 Blue (Stratagene) or JM109 (Promega), efficiencies were 1 $x10^9$ and 1 $x\ 10^7$ respectively. Electrocompetent cells M15p(Rep4) (Qiagen) and BL21 DE3 (Novagen) were produced by the method of Tung and Chow (1995).

2.2.8 Transformation of Competent *E. coli* with Plasmid DNA

The protocol for transformation of chemical competent cells was identical for either type of cells with one small exception. For both, a sample of the ligation mix (usually 1 or 2 μl from a 10 μl sample) was added to the competent cells in a prechilled eppendorf tube and left on ice for 30 minutes. The tubes were then transferred rapidly to a water bath at 42 °C for exactly 45 seconds before replacing on ice for a further 2 minutes. 1 ml of LB, pre-warmed to 37 °C, was then added and incubated at 37 °C for 1 hour with shaking at 200 rpm. 100 μl of this was then spread onto an LB-agar plate containing the appropriate antibiotic. XL1-Blue cells followed exactly the same protocol with the addition of β-mercaptoethanol with the ligation mix as outlined in the supplied protocol. Where colour selection of positive clones was required, the plates were also supplemented with 60 μl of 100 mM IPTG and 40 μl of 20 mg/ml X-gal immediately prior to adding the cells. For routine transformation of intact circular plasmid, high efficiency competent cells were not used. Instead, 1 ml of an overnight culture of XL1-Blue cells was added to 9 ml LB and incubated at 37 °C for 1 hour with shaking at 200 rpm. A 1 ml aliquot was removed, the cells harvested by centrifugation in a bench top microfuge at 6000 rpm for 5 minutes and washed 2 times in 1 x PBS. The pellet was then resuspended in 100 μl ice cold 20 mM $CaCl_2$. 100 ng of plasmid was added to this and left on ice for 30 minutes before heat shocking at 42 °C. To this 400 μl LB, pre-warmed to 37 °C, was added and 100 μl immediately spread onto LB-agar plates with the appropriate antibiotic.

2.2.9 Agarose Gels

For routine analysis of DNA, agarose gels were prepared by melting 1% powdered agarose (w/v) in 0.5 x TBE. The solution would be allowed to cool to hand hot temperature before pouring. If the DNA was to be excised and purified from the gel then an identical procedure was carried out but 1 x TAE replaced the TBE buffer. Samples were loaded onto the gels in the presence of gel loading buffer containing glycerol and bromophenol blue dye. To visualise DNA after running the gels, they were stained in a solution of the relevant buffer and 0.5 μg/ml ethidium bromide. Stained gels were viewed under ultraviolet illumination and the image recorded using an Appligene Imager (Appligene Oncor). Long wave UV was used if the DNA was to be excised or used subsequently for transfer to a membrane to minimise damage to the DNA.

2.2.10 DNA Extraction from Agarose Gels.

DNA fragments were excised from TAE/agarose gels with a clean scalpel blade and isolated using a Qiaquick gel extraction kit (Qiagen) or GenElute column (Sigma). Despite manufacturer's claims, total DNA recovery was variable and it was essential that subsequent to the removal of contaminating agarose the exact concentration of DNA was confirmed by running a sample on a gel with a DNA marker of known quantities.

2.2.11 Automated DNA Sequencing

Purified plasmid for sequencing was resuspended and sent to one of the following commercial laboratories for analysis: PNACL (Leicester, UK), MWG (Milton Keynes, UK), MBSU (Glasgow, UK).

2.2.12 Oligonucleotide Primers for Cloning and Site Directed Mutagenesis

Gene specific primers were obtained from commercial suppliers. All were supplied desalted but not subject to further purification. The sequence of the primers used in this study are indicated in Table 2.1.

Oligonucleotide Name	Oligonucleotide Sequence
OL281	5'-AGATCTGAGGACTAAGTCGAGAGCCT-3'
OL282	5'-CCATGGGTGGCCACGCTATCGACCCAACACATA-3'
R58Kf	5'-CCGGCTACATCTACACAAAGCTCGGCAACCCAACAG-3'
R58Kr	5'-CTGTTGGGTTGCCGAGCTTTGTGTAGATGTAGCCGG-3'
R58Mf	5'-CCGGCTACATCTACACAATGCTCGGCAACCCAACAG-3'
R58Mr	5'-CTGTTGGGTTGCCGAGCATTGTGTAGATGTAGCCGG-3'
L59Ff	5'-GGCTACATCTACACACGTTTCGGCAACCCAACAG-3'
L59Fr	5'-CTGTTGGGTTGCCGAAACGTGTGTAGATGTAGCC-3'
Y111Ff	5'-CCGATGAGTGCCTTTTTGGCTGCACACATGC-3'
Y111Fr	5'-GCATGTGTGCAGCCAAAAAGGCACTCATCGG-3'
C113Pf	5'-CCGATGAGTGCCTTTATGGCCCCACACATGCTCTC-3'
C113Pr	5'-GAGAGCATGTGTGGGGCCATAAAGGCACTCATCGG-3'
K209Sf	5'-GTCCACTCTGCAACAAGCTACATCAACGGCC-3'
K209Sr	5'-GGCCGTTGATGTAGCTTGTTGCAGAGTGGAC-3'
K238Af	5'-CGTATGGTTGGTATCGCGGATATCACAGGATCTG-3'
K238Ar	5'-CAGATCCTGTGATATCCGCGATACCAACCATACG-3'
M13	5'-GTAAAACGACGGCCAGT-3'
T3	5'-AATTAACCCTCACTAAAGGG-3'
OASS 1	5'-CATATGATCTACGACAACATCCTCGAGAC-3'
OASS 2	5'-GAAGAAATCTGGGCCCAG-3'
OASS 3	5'-GGTTCTTGCCAACACCCC-3'
OASS 4	5'-CCGTGTCTACCGCGGTAGGCATTCC-3'
OASS 5	5'-CTCACTCCCGCAAAGTTAGCTTGCGGC-3'
OASS 7	5'-GCAGTGTCGACTTCTGTGTCGAAGAGCTTTG-3'

Table 2.1 Oligonucleotides used for sequencing and mutagenesis

2.2.13 Radioactive Labelling of Nucleotide Probes

DNA probes for Southern blot hybridisations were radiolabelled using the Prime It™ kit (Stratagene) following the supplied protocol. Briefly, 25 ng of template DNA was incubated in the presence of radioactively labelled dATP (^{32}P-3000 mCi/mmol, NEN Dupont). Random hexamer primers and Exo(-)Klenow DNA polymerase were used to incorporate the radioactive nucleotides into the probe. Probes were denatured by boiling for 10 minutes prior to incubation with the membrane.

2.2.14 Southern Blotting

T. vaginalis gDNA was digested with an appropriate restriction enzyme as detailed in section 2.2.5 and resolved on a 0.7% agarose gel. The gel was stained with ethidium bromide to confirm adequate digestion of the DNA and then photographed before further treatment prior to transfer to Hybond N membrane (Amersham Pharmacia). The gel underwent depurination in 0.25 M HCl for 10 minutes before rinsing briefly in dddH_2O. This was followed by denaturation (1.5 M NaCl, 0.5 M NaOH) and neutralisation (1.5 M NaCl, 0.5 M TRIS-HCl pH 7.5) for 30 minutes each. The gel was rinsed briefly in dddH_2O after both steps before transfer by capillary action as described in Sambrook *et al.*, 1989. After overnight transfer in 20 x SSC the gel was restained with ethidium bromide to confirm complete transfer of DNA. The transferred DNA was cross-linked using a Spectrolinker™ XL-1000 crosslinker (Stratagene) before the filters were used for hybridisation with radioactively labelled probes. The membranes were prehybridised for 4 hours in rotating glass tubes with sealed ends at 65 °C with 10 ml hybridisation solution (0.5 M sodium phosphate pH 7.2, 7% SDS, 1 mg denatured and sheared herring sperm DNA). Denatured radioactively labelled

probe was then added directly to this solution and incubated at 65 °C overnight with continuous rotation. The hybridisation solution was then decanted and the membrane washed. Increased stringency washes were performed at 65 °C with rotation for 30 minutes. Firstly in 2 x SSC, 0.1% SDS, then 1 x SSC, 0.1% SDS and finally 0.1 x SSC, 0.1% SDS. The blot was then wrapped in saran wrap and exposed to autoradiographic film (Fuji RX) at -70 °C for as long as necessary (typically 12-36 hours). Film was developed automatically using a Kodak Xomat film processor or manually using Kodak reagents.

2.2.15 Site Directed Mutagenesis of MGL1

Mutagenesis of specific residues was performed using mismatched oligonucleotide primers following the Quickchange (Stratagene) protocol. Briefly, plasmid gl14 which contains the *mgl*1 ORF in expression vector pQE60 (Qiagen) was used as a template and complementary oligonucleotides were used to mutate appropriate nucleotides changing the amino acid sequence. The remaining plasmid sequence was then synthesised using the thermostable DNA polymerase *pfu*, which possesses an exonuclease activity and has a reduced likelihood of incorporating errors into the DNA strand. After 25 cycles of DNA synthesis the original plasmid template was removed from the mutated DNA by digestion with *Dpn* I. This restriction enzyme is a methylated DNA specific endonuclease and as such only digested the template plasmid. The remaining nicked circular mutated plasmid was then transformed into competent *E. coli* and selected on LB agar plates containing ampicillin. Verification of the mutant was by automated DNA sequencing and restriction digest analysis where appropriate.

2.2.16 Production of New *mgl2* Construct

Plasmid gl15 encoding *mgl2* was used as a template to add restriction sites to the 5' and 3' ends by PCR using primers OL281 and OL282 (Table2.1) as detailed in section 2.2.3. This added an *Nco*I site to the 5' end of the ORF whilst replacing the missing residues and a *Bgl*II site at the 3' end. This fragment was then excised from the plasmid and ligated into precut pQE60 as described for *mgl*I (McKie, 1997). To produce mutant MGL2$^{(C116G)}$, a *Sal*I/*Sph*I 350 bp fragment was removed from the original construct and ligated into the new one. This circumvented the need to repeat the site-directed mutagenesis previously required.

2.2.17 Cysteine Synthase 5'-RACE

Isolation of the 5' untranslated region (5'-UTR) of *T. vaginalis* cysteine synthase was attempted using a 5'-RACE system kit (Life Technologies). Total RNA isolated from the parasite was used as the template for the reaction. Experiments were carried out according to the manufacturers instructions. Briefly, cDNA was transcribed from mRNA using reverse transcriptase and an oligo dT primer. A poly-dC tail was added to the isolated cDNA and this used as a template for PCR with a supplied anchor primer (complementary to the dC tail) and a cysteine synthase gene specific primer (OASS 4). Amplification of the product from this reaction was attempted with another gene specific primer (OASS 5) and a supplied anchor primer.

2.3 Biochemical Materials and Methods

2.3.1 Protein Concentration Estimation

Protein concentration was routinely determined with the Pierce BCA system using the supplied microtitre plate protocol. Bovine Serum Albumin or γ-globulin were used as protein standards over a range of 0.1 to 2.0 mg/ml.

2.3.2 Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Protein analysis and molecular weight estimation was done using discontinuous SDS-PAGE as described by Laemmli, 1971. Acrylamide resolving gels of between 7.5 and 12.5% (w/v), 0.75 mm thick were cast and run using a Bio-Rad Mini Protean II dual slab cell as per manufacturer's instructions. Molecular weight markers (10 kDa protein ladder, Gibco) were run routinely to estimate the apparent molecular weight of sample proteins.

Proteins were revealed by staining with Coomassie blue dye. Gels were immersed for 30 minutes or more in staining solution (40%(v/v) methanol, 7%(v/v) acetic acid, 0.1% (w/v) Coomassie Blue R250) and then destained with 40%(v/v) methanol, 7% (v/v) acetic acid until background staining was at a minimum. Gels were either photographed or stored by drying. To dry, gels were soaked overnight in a 4% glycerol (v/v) destain solution and then dried overnight between 2 cellulose sheets pre-soaked in dddH_2O.

2.3.3 Native Gel Electrophoresis

Native gels were cast and used exactly as for SDS-PAGE but with the omission of SDS from gel and buffers. Reducing agents such as DTT or β-mercaptoethanol were also omitted from loading buffer.

2.3.4 Western Blotting

Transfer of proteins to Hybond-ECL membrane (Amersham Pharmacia) from acrylamide gels was achieved as described by Towbin *et al.*, 1979. A Bio-Rad mini transblot cell was used following the manufacturer's instructions.

2.3.4.1 Membrane Staining

To verify successful transfer of proteins to the membrane they were stained with a Ponceau S solution. This also enabled the membrane to be cut accurately if required between protein lanes and molecular weight marker positions could be marked on the membrane. The membranes were immersed in the Ponceau stain for a few minutes and rinsed with $dddH_2O$ to remove background staining.

2.3.4.2 Prehybridisation of Membranes - Blocking

Blocking of non-specific binding is essential in Western blotting to obtain a clear result with minimal background. The blocking procedure prevents binding of immunoglobulins (both primary and secondary) to non-specific areas of low affinity on the membrane, leaving only the high affinity binding sites which would be preferentially bound. Before adding the primary antibody the membrane was immersed for 4 hours at room temperature with gentle agitation in blocking solution (5% (w/v) skimmed milk powder, 0.2% TWEEN 20, 1 x TRIS buffered saline (TBS) (20 mM TRIS-HCl, 150 mM NaCl, pH 7.6).

2.3.4.3 Primary Antibody Binding

After prehybridisation, the solution was removed and replaced by hybridisation solution (1% (w/v) skimmed milk powder, 0.1% TWEEN 20, 1 x TBS) and primary antibody or pre-immune serum added. This was then incubated overnight at 4 °C with gentle agitation. Subsequent to incubation with the primary antibody, membranes were washed 3 times for 30 minutes with 1 x TBS plus 1% (w/v) skimmed milk powder at room temperature.

2.3.4.4 Secondary Antibody Binding

The secondary antibody used for detection of anti-cysteine synthase primary antibody was a horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (Scottish Antibody Production Unit, SAPU). After washing the antibody was hybridised to the membrane in 1 x TBS plus 1% (w/v) skimmed milk powder at room temperature for 2 hours. After incubation the membrane was washed 3 times for 30 minutes in 1 x TBS plus 1% (w/v) skimmed milk powder at room temperature and finally in 1 x TBS for 10 minutes.

2.3.4.5 Western Blot Developing

Detection of HRP activity was by enhanced chemiluminescence (ECL) using the ECL kit (Amersham Pharmacia) as supplied. Essentially the two solutions provided are combined 1:1 in an appropriate volume and incubated with the membrane with agitation for 2 minutes. The luminol solution in the kit is cleaved by the action of HRP and photons of light are emitted. This should only occur at sites of secondary antibody binding to the primary antibody and thus specific for the presence of the protein antigen desired. Film was developed automatically using a Kodak Xomat film processor or manually using Kodak reagents.

2.3.5 MGL Purification

Recombinant MGL for biochemical analysis was purified as previously published (McKie *et al.*, 1998).

2.3.6 Homocysteine Desulphurase Assays

The $\alpha\gamma$ elimination of homocysteine and subsequent detection of H_2S production was by the published method of Thong and Coombs (1985a). This assay was routinely used for the measurement of enzyme activity as it was the most reliable, simple and sensitive of those available.

Other MGL assays investigated during this study are indicated in the relevant chapters.

2.3.7 Lysis of Small Samples of *E. coli*

When performing test expressions in *E. coli*, samples removed to monitor over-expression of recombinant proteins were too small to be conveniently lysed by sonication. Cells were instead lysed by enzyme and detergent action. 15 ml samples of cultured bacteria were removed as appropriate and harvested by centrifugation at 3200 g for 15 minutes at 4 °C. Pellets were washed in PBS and placed at -20 °C until required. Lysis was performed by resuspending the pellets in 0.5 ml of 2 mg/ml lysozyme, 0.1% Triton X-100, 25 mM TRIS pH 8.5 and incubating overnight at 37 °C with agitation. The lysate was then spun at 13 000 rpm in a benchtop microfuge for 25 mins to remove intact cells and insoluble debris. 40 µl of the supernatant was then used for analysis by SDS-PAGE.

2.3.8 Purification of Recombinant Cysteine Synthase

The purification of cysteine synthase is accomplished in a very similar manner to that used for MGL. The development of the method used for expression and purification is discussed in detail in section 5.2.3.

2.3.9 Cysteine Synthase Assays

Cysteine synthase activity was assayed under the following conditions. Buffer and dddH_2O were deoxygenated by bubbling with N_2 gas for 15 minutes prior to use. In a total volume of 0.5 ml, enzyme was added to 100 mM sodium phosphate buffer pH 7.3, 3 mM sodium sulphide, 10 mM O-acetyl-L-serine, 1 mM DTT to initiate the reaction. This was mixed briefly and incubated at 37 °C for 5 minutes in a water bath. The reaction was terminated by adding 150 µl of the assay mix to 350 µl of acidic ninhydrin solution (1.3% (w/v) ninhydrin in 1:4 concentrated hydrochloric acid:glacial acetic acid) (Gaitonde, 1967) in screw capped eppendorf tubes and mixed thoroughly. To complete the colourimetric reaction eppendorf tubes were boiled for 10 minutes in a boiling water bath before cooling on ice for 30 minutes. The samples were then diluted by adding 0.7 ml 100% ethanol and the absorbance measured at 560 nm. The quantification of the amount of cysteine produced was by calibrating against a standard curve of known concentrations of L-cysteine.

2.3.10 "Activated Serine Sulphydrase" Assays

These assays were carried out essentially as described previously (Thong and Coombs, 1985b). Enzyme was added to a total volume of 1 ml 100 mM sodium phosphate pH 7.3, 330 µM lead acetate, 3.3 mM β-mercaptoethanol and L-

cysteine to initiate the reaction. The production of lead sulphide from the reaction of H_2S with lead acetate was followed at 360 nm.

2.4 General Materials and Methods

2.4.1 Parasite Line

A single line of *T. vaginalis* was used in this study throughout. Clone G3 is a metronidazole sensitive line used for many years in this laboratory and also used for the random cDNA library sequencing project carried out by collaborators.

2.4.2 Cultivation of *T. vaginalis in vitro*

The parasite was cultured axenically in Modified Diamonds Medium (MDM) (Diamond, 1957), supplemented with 10% (v/v) heat inactivated horse serum (Life Technologies)(Complete MDM), at 37 °C. MDM is comprised predominantly of peptone (1% (w/v)), yeast extract (0.5% (w/v)) and maltose (0.5% (w/v)) buffered with potassium salts and supplemented with ascorbic acid and iron sulphate. Cultures were checked by microscopical examination for contamination before use.

To produce parasites for DNA purification, they were inoculated into fresh complete MDM and grown under anaerobic conditions until they reached late log growth phase, producing cells at a density of $1\text{-}2x10^6$ cells/ml.

Parasites cultured under different environmental pressures were grown in complete MDM supplemented with the relevant compound. 25 ml cultures were prepared in tightly capped universals, except for those exposed to aerobic growth

conditions which were grown in 500 ml tissue culture flasks in a normally aspirated incubator. For growth without ascorbate, complete MDM was prepared identically with the exception of the omission of ascorbic acid. All cultures were initiated with $1x10^5$ parasites per ml and grown for 18 hours before dividing into 5 ml aliquots and harvesting.

2.4.3 Cell Counts

Parasites were fixed with 4% paraformaldehyde and counted under phase contrast microscopy using an improved Neubauer hemocytometer.

2.4.4 Harvesting Parasites

T. vaginalis were harvested by centrifugation at 2300 g for 15 minutes at 4 °C and washed twice with ice cold 0.25 M sucrose. Cells for RNA analysis were resuspended immediately in 1 ml Trizol (Life Technologies) and stored at -70 °C until required. Others were stored as pellets at -70 °C until required.

2.4.5 Parasite Lysis

Crude lysates of parasites were obtained in one of two ways. For small numbers of cells (less than $5x10^7$) the washed cell pellet was resuspended in 0.25 M sucrose plus 0.25 % (v/v) Triton X-100. Cell debris was removed by centrifugation at 13 000 rpm in a benchtop microfuge for 5 minutes. For larger numbers of cells typically resulting from a 500 ml culture, the cells were lysed using a tissue homogeniser. Parasites were resuspended at approximately $2x10^8$ cells/ml in 1 x PBS or 0.25 M sucrose. Lysis of cells was checked by microscopical examination. Cell debris was removed by centrifugation at 13 000

g for 30 minutes. To reduce proteolytic degradation 10 μM leupeptin and 10 μM E64, both cysteine proteinase inhibitors, were added to the lysates.

CHAPTER 3. MGL1 STRUCTURE - HOLOENZYME AND INHIBITOR COMPLEX

3.1 INTRODUCTION

This chapter deals with the result of completing the primary aim of this study, obtaining the 3-dimensional crystallographic structure of MGL. Results of the processes carried out to achieve this and a discussion of these findings are given.

The first step to accomplish in the structural determination of MGL was the growth of suitably diffracting crystals. This is often a limiting factor in the structural studies of proteins. The situation can be made more difficult if the protein is post-translationally modified or must be purified from a native source if not available recombinantly. This was not the case for MGL and the production of large amounts of active, highly pure recombinant enzyme meant that it was possible to make many attempts to produce crystals. Preliminary trials were conducted at the beginning of the project to ascertain the likelihood of succeeding in this and the results are detailed in section 3.2.4 (this work was carried out by Dr. A J Lapthorn).

It was envisaged that the large number of PLP-dependent enzyme structures solved and published would aid in solving the structure of MGL using the technique of molecular replacement. Failing this, heavy atom derivatives would allow for isomorphous replacement. It was briefly considered at the beginning of the project to produce selenium derivatives of the enzyme for MAD (Multiple

Anomalous Diffraction) experiments using selenocysteine or selenomethionine but the activity of the enzyme towards both of these compounds prevented this approach.

At the beginning of this project only a single structure with significant sequence homology and predicted to be in the same family as MGL (Alexander *et al.*, 1994) had been published, that of cystathionine β-lyase (CBL) (Clausen *et al.*, 1996). The coordinates for this structure were unavailable at the time though and this prevented detailed analysis. More recently there have been two further structures published of the closely related enzyme cystathionine γ-synthase (CGS) (Clausen *et al.*, 1998; Steegborn *et al.*, 1999). A mechanism of action has been proposed for CBL (Clausen *et al.*, 1997) but this was thought likely to differ significantly from that of MGL because of the apparently important cysteine residue of MGL (McKie *et al.*, 1998). Differences in substrate preference between the enzymes, MGL from *T. vaginalis* is unable to act upon cystathionine, also suggested the active site would be dissimilar and thus possibly the mechanism. Solving the structure of MGL would enable us to propose a mechanism and assign roles to specific residues within the active site.

The role of an active site cysteine in MGL was difficult to define without structural information as the enzymes from *T. vaginalis* and *Pseudomonas. putida* were not completely inactivated by either chemical modification or mutagenesis of this residue (Nakayama 1988a; McKie *et al.*, 1998). The activity though was severely curtailed and this pointed to an important if not essential function. MGL is inactivated by the suicide inhibitor L-propargylglycine and this has been shown to

target either a cysteine or tyrosine in the active site (Johnston *et al.*, 1979). This also implicated the cysteine as playing a role in catalysis. The crystal structure would hopefully elucidate the role of this residue and confirm the inhibition by propargylglycine.

3.2 RESULTS

3.2.1 MGL purification

MGL1 and MGL2 were purified as published (McKie *et al.*, 1998) with some revisions. As enzyme was to be used for crystallisation, it was considered critical to obtain as close to a 100% homogenous enzyme preparation as possible. Elution by a linear gradient of 0-500 mM imidazole produced predominantly homogenous enzyme, but some contaminating *E. coli* proteins still co-eluted. A sample of each of the fractions collected was ran on 7.5% SDS-PAGE and coomassie blue stained. This generally revealed low level but detectable contamination of MGL with proteins of higher and lower molecular weights.

In an attempt to avoid this contamination and produce a purer preparation, up to 50 mM imidazole was included in the sonication and wash buffers to prevent *E. coli* proteins with low affinity from binding to the column. This was less successful than had been expected, giving no obvious improvements and causing the premature elution of enzyme. It is possible that this was due to low levels of imidazole competing for Ni-NTA binding with the hexahistidine tag of MGL. Although these concentrations are not high enough to elute the tagged protein

directly, during the loading of the crude lysate onto the column a proportion of the recombinant enzyme is washed away as it is continually displaced and passes further down the column until it is eventually eluted with the flow through buffer. In support of this suggestion is the observation that subsequent elution of the recombinant enzyme occurs over a larger volume. Moreover the MGL bound to the column can usually be seen as a band at the top of the column in a standard purification (its yellow colour visible due to the PLP cofactor). Adding 50 mM imidazole to the buffers results in the yellow band of enzyme becoming more diffuse than the normal tight band.

Therefore in a different attempt to remove the last remaining *E. coli* protein contaminants effectively, 75 mM imidazole was applied to the column prior to elution by the linear imidazole gradient. This caused the elution of a number of proteins while having a negligible effect upon the bound MGL and proved a highly successful method for removing trace contaminants and obtaining a high yield of very pure enzyme. Figures 3.1 and 3.2 show the FPLC purification profile of the recombinant enzyme and the subsequent analysis on coomassie stained SDS-PAGE, respectively.

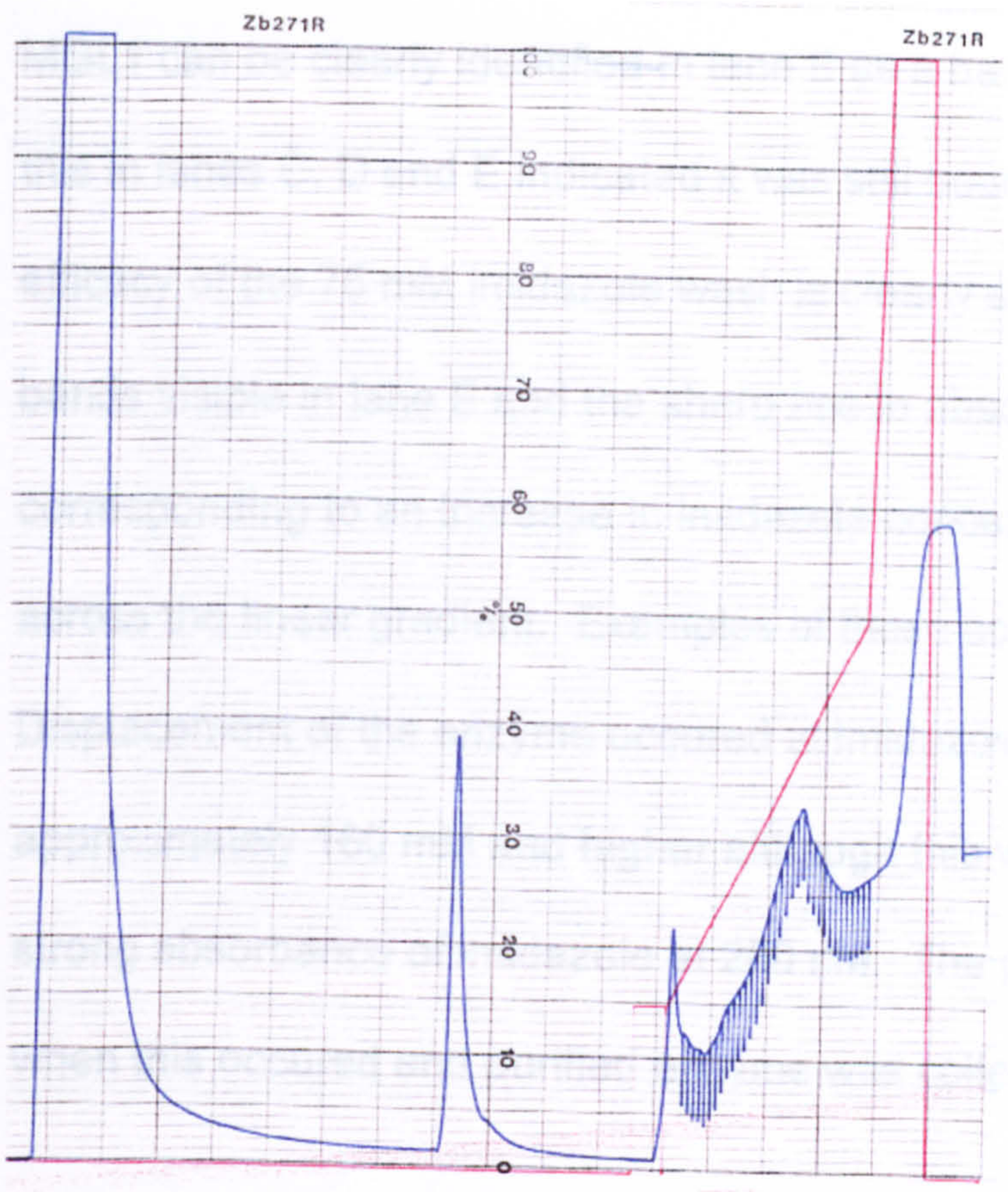

Figure 3.1: FPLC profile of MGL purification. The blue line represents the eluted sample absorbance at a wavelength of 280 nm. The upper limit of detection is 2.0 AU. Fractions collected are indicated on the figure by vertical downward strokes of the blue line. Red line represents the percentage of 500 mM imidazole containing buffer applied to the column.

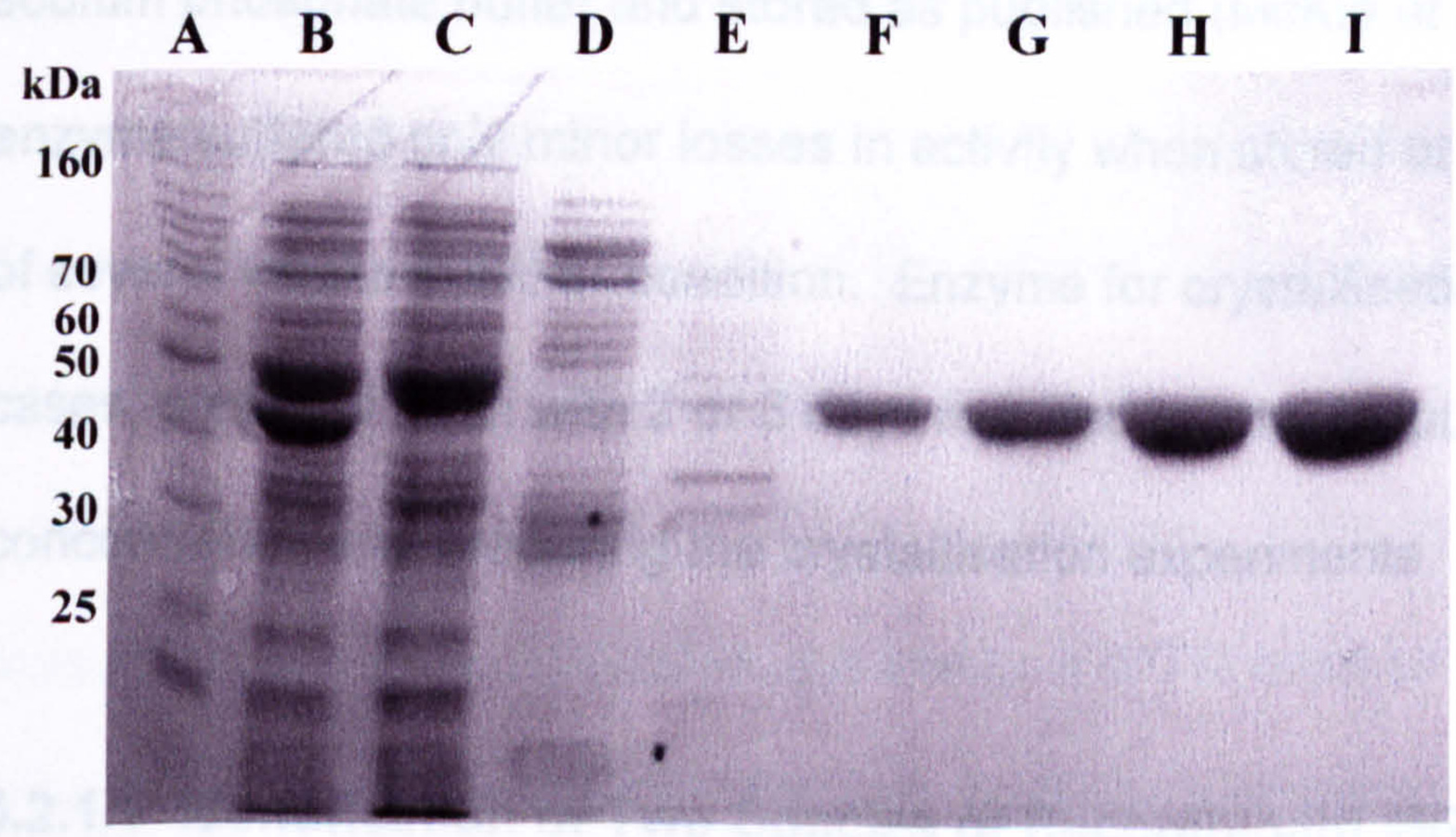

Figure 3.2: SDS-PAGE of a typical MGL1 purification. 10% Polyacrylamide gel stained with Coomassie blue. Lanes: A - 10 kDa protein ladder; B - Crude lysate; C - Column loading/flow through; D - First wash; E - 75 mM Imidazole wash; F to I - Fractions eluted by linear 155 to 185 mM imidazole gradient.

MGL1 can be clearly identified in lane B as a band at 43 kDa. The absence of this in lanes C, D and E indicated it was still bound to the Ni-NTA matrix. The efficacy of the 75 mM imidazole wash is clearly seen by the large number of bands visible in lane E and the sharp rise in absorbance at 280 nm. 1 ml fractions corresponding to an increase in imidazole concentration of 10 mM were collected across the linear gradient. Examples of those containing MGL1 are shown above. Displacement of the enzyme occured at imidazole concentrations of approximately 160 mM and higher although this was partially obscured by the strong absorbance of imidazole at 280 nm. The gradient was not halted though when this occured and purified enzyme was collected in a number of fractions.

Routinely, the MGL1 containing fractions from the column were identified by their yellow colour in solution and checked using SDS-PAGE. The purest fractions were pooled and either concentrated ready for crystallisation, or dialysed into sodium phosphate buffer and stored as published (McKie *et al.*, 1998). The enzyme suffered only minor losses in activity when stored at –20 °C over a period of several weeks in either condition. Enzyme for crystallisation was, in most cases, prepared fresh with 2 or 3 days required in total for purification, concentration and preparing the crystallisation experiments.

3.2.1.1 Identification of Two Species of Recombinant MGL1

Initially purification of MGL1 was monitored by analysing samples using 7.5% polyacrylamide SDS-PAGE. In order to detect contaminants more easily a significant amount of highly pure MGL1 (> 5 μgs) was loaded on each occasion. This ran as a large 43 kDa band with only a single species visible. However,

running the same samples on a 10% gel showed two very close but distinct bands differing by approximately 1 kDa (Figure 3.3).

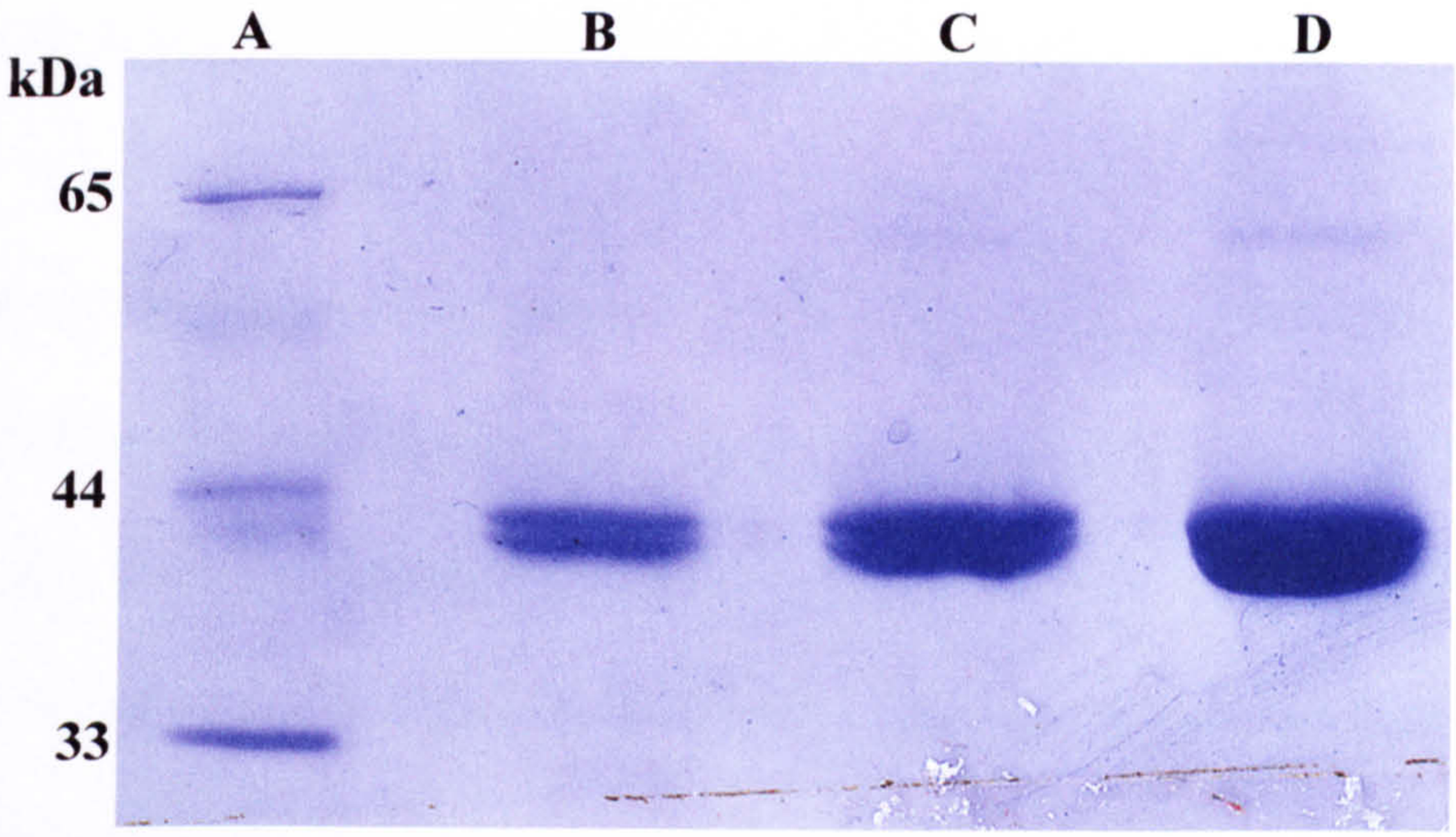

Figure 3.3: Coomassie blue-stained 10% SDS-PAGE of purified MGL1. Lanes: A - protein molecular weight markers; B to D - Increasing amounts of purified MGL1.

This was not noticed for some time and was suspected of causing the failure of some of the crystallisation trials. As can be seen in figure 3.3, both bands stained to a very similar degree and were estimated to be present in an approximately 1:1 ratio. The figure shows clearly the difficulty in identifying two bands in lane D compared with lane B. The large amount of protein loaded in lane D, though, does confirm the very high purity of the sample. It was assumed that both bands were recombinant MGL1 for various reasons including the appearance of two bands in a native substrate-PAGE gel (not shown). In an attempt to avoid the production of 2 distinct proteins, the levels of IPTG used to induce MGL1 expression were reduced from the very high concentration used of 2 mM to 0.1 mM. This had a positive effect, greatly reducing but not totally eliminating the smaller of the two bands (the second MGL1 band is not visible in figure 3.2 lanes F to I). The modification did, however, coincide with successful crystallisation of

the enzyme. Thus no further attempts were made to eliminate the second protein from the preparation. The cause of this second band and the effect of reducing IPTG levels will be considered in the discussion section of this chapter.

3.2.2 Protein Preparation

MGL was prepared for crystallisation as described in section 2.1.2. No problems were encountered for the concentration of MGL1 samples but it was apparent that there was some protein losses when concentrating MGL2 and this was attributed to binding to the Microsep™ membrane. Concentrators were soaked in 10% glycerol overnight at 4 °C before use in order to try and alleviate this problem as much as possible. Although not quantified the recovery of MGL2 after this was significantly better.

3.2.3 Analysis of MGL Monodispersity by Dynamic Light Scattering (DLS)

It has been proposed that that the monodispersity of a crystallisation sample can be correlated to the likelihood that crystals will form (Malkin *et al.*, 1994; Veesler *et al.*, 1994). Polydispersity can indicate the presence of impurities, isoforms or multiple aggregation states in the protein sample. A sample which is composed of a single species is theoretically more likely to produce crystals.

Monodispersity of MGL was measured as detailed in section 2.1.5. Several factors may effect the monodispersity of a solution, the most straightforward to examine being the pH or the presence of salts. The effect of pH was examined by diluting the enzyme sample to 1 mg/ml in 100 mM of relevant buffer. Buffers were selected to cover a wide range from pH 4.6 to pH 10.0. The results were

equivalent for both enzymes showing that it was possible to obtain a monodisperse solution by altering the pH to below 6.5 or above 8.0. Both enzymes aggregated heavily at pH 7.5 as was immediately evident from the difficulty in injecting the sample through a 0.1 μm filter. That part of the solution which did pass through the filter showed species of estimated molecular weights of greater than 1000 kDa confirming a large aggregate of protein. This was the case for both MGL1 and MGL2.

In combination with pH, the effect of DTT and additional PLP were also examined. However these had no apparent effect on the solution regardless of the buffer and pH.

3.2.4 Crystallisation of MGL1

Initial attempts to crystallise MGL1 preceded the DLS investigation and utilised a number of "sparse-matrix" screens (Jancarik and Kim, 1991). Here the enzyme had been stored at -20 °C in 100 mM HEPES pH 7.5, 40 μM PLP, 30 μM DTT, 40% glycerol and used at 6 mg/ml in the same conditions minus the glycerol. The positive results are summarised in the table below:

Precipitant	Buffer (0.1 M)	Salt	Result
30% PEG 4k	TRIS pH 8.5	0.2 M Li_2SO_4	Many thin rods
30% PEG 4k	TRIS pH 8.5	0.2 M NaCl	Many tiny rods
3 M $AmmSO_4$	Citrate pH 5.6	0.2 M Li_2SO_4	Thin plate like crystals
20% PEG 8k	HEPES pH 7.5	0.2 M Li_2SO_4	Irregular shapes
20% PEG 8k	TRIS pH 8.5	0.2 M $NaKPO_4$	Very fine granular
20% PEG 8k	TRIS pH 8.5	0.2 M NaCl	Very fine granular
30% Me2k PEG	Citrate pH 5.6		Small granular crystals
10% PEG 4k	Citrate pH 5.6	0.1 M Li_2SO_4	Feathered needles
20% PEG 4k	Citrate pH 5.6	0.2 M $AmmSO_4$	Long needles
30% Me2k PEG	MOPS pH 6.5		Irregular needles
25% PEG 4k	HEPES pH 7.5		Small rectangular
30% PEG 4k	TRIS pH 8.5		Many tiny rods
30% PEG 4k	TRIS pH 8.5	0.2 M $CaCl_2$	Few tiny rods
25% PEG 4k	PIPES pH 6.8		Small rods

Table 3.1: Results of initial crystallisation trials of MGL1. N.B. This work was done by Dr. AJ Lapthorn.

Thus PEG was seemingly the precipitant of choice with a range of suitable buffers at varying pH and with different salts. The first comprehensive attempts at crystallising MGL1 encompassed these conditions. After a number of trials it was evident that a combination of PEG 4k, 0.1 M citrate pH 5.0 - 6.0 and various salts gave the most consistent and promising results. The crystals resulting from these conditions varied from small cubic shaped (approx. 0.05 mm x 0.05 mm) and small rods (0.1 mm) to long fine needles (up to 0.5 mm). However none of these were suitable for structural analysis.

At this time the data on the monodispersity of the solution were obtained from the DLS analysis and the buffer was changed to 50 mM imidazole pH 6.5. Additionally, because of reports that the PLP cofactor can be unstable in some enzymes giving rise to a population of apoenzyme over time, the concentration of the cofactor was increased to 250 μM and the DTT concentration raised to 1 mM which is more commonly used than the previous 30 μM. These new conditions considerably changed the morphology of the resulting crystals. In the same conditions as had previously been tried, the crystals produced were much larger (0.15 mm x 0.075 mm) but were hollow. This suggested that their growth rate was too rapid and so the same conditions were repeated at 4 °C and with 4% glycerol in the solution to retard growth. Neither of these alternatives succeeded in producing large crystals without a hollow centre.

In parallel to the DLS analysis, duplicate sparse-matrix screens had been set up with the enzyme inactivated by the inhibitor DL-propargylglycine (5 mM). L-propargylglycine is a covalent, suicide inhibitor of several PLP-dependent enzymes, including MGL (Johnston *et al.*, 1979). It was hoped that the inclusion of this would produce a more rigid structure leading to larger crystals. Crystallisations were set up identically as previously described but with the addition of 5 mM DL-propargylglycine to the enzyme solution. The screen gave similar results as previously but with the addition of large, bipyramidal crystals (0.3 x 0.2 mm) in conditions of 3.0 M $AmmSO_4$, 0.1 M citrate buffer pH 5.6, 0.2 M Li_2SO_4. The enzyme at this time was concentrated in HEPES buffer indicating the monodispersity of the enzyme was not critical in allowing large crystal formation. After optimisation via linear screens the exact concentration of ammonium

sulphate required was found to be 2.3 M. The process of optimisation also highlighted an additional point, freezing of MGL1 reduced its ability to form crystals. This appears to be a relatively common problem with protein crystallisation and the most successful trials were consistently those using freshly purified enzyme no more than 2 or 3 days old and stored at 4 °C prior to use. The native enzyme also crystallised in exactly the same conditions as the inhibitor complex, once all of the important factors had been identified.

Thus in summary: crystals were grown at 20 °C by combining 1.5 μl of 6 mg/ml MGL1 in 0.1 M HEPES pH 7.5, 250 μM PLP, 1 mM DTT with 1.5 μl of precipitant comprising 2.3 M $AmmSO_4$, 0.2 M Li_2SO_4 and 0.1 M citrate buffer pH 5.6. Single crystals formed over 6 weeks and achieved approximate dimensions of 0.6 mm x 0.4 mm x 0.4 mm (See figure 3.4). The gross morphology of both types was identical and both diffracted anisotropically to 1.6 Å.

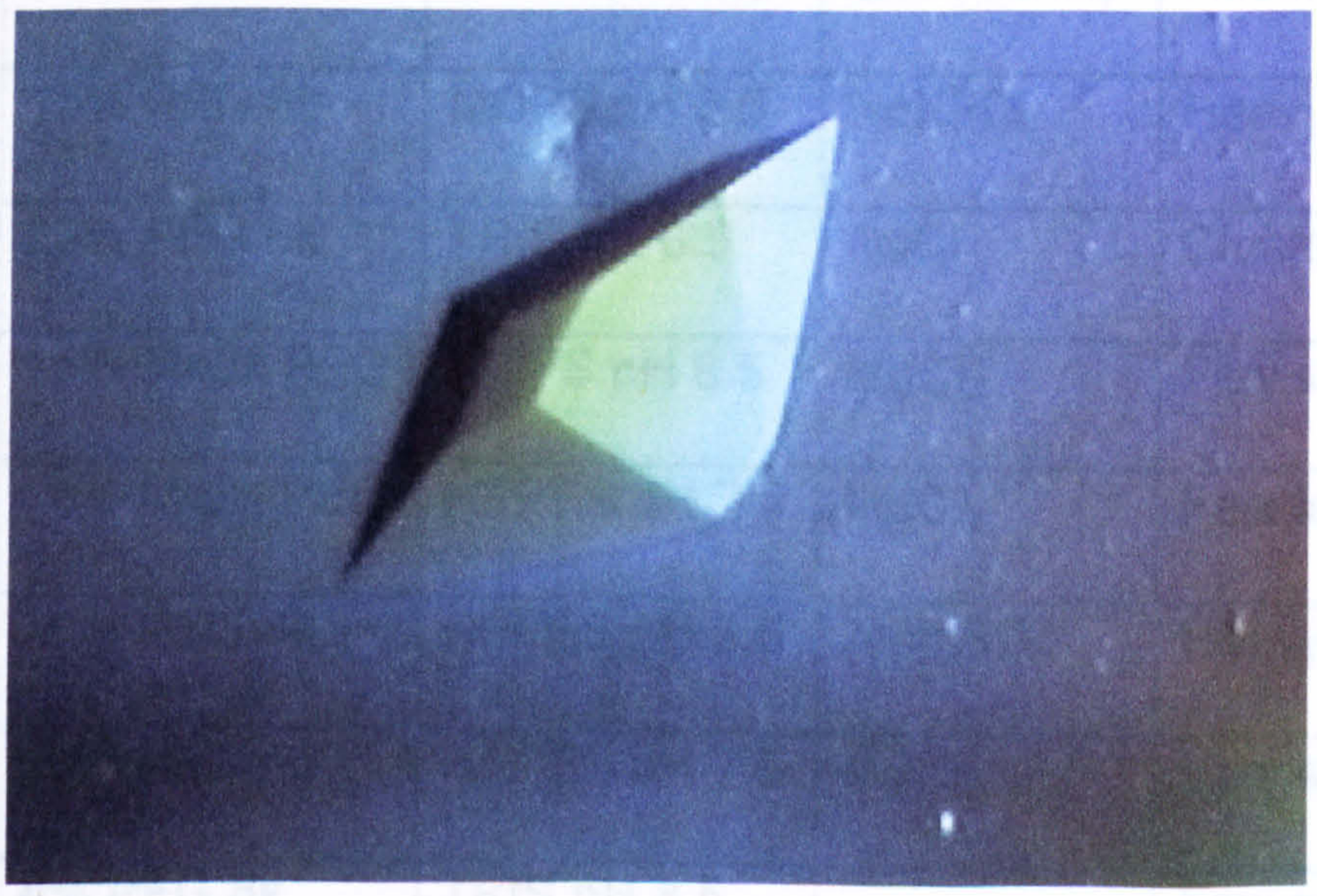

Figure 3.4: Single orthorhombic crystal of MGL1 in complex with L-propargylglycine grown as detailed above. Crystal shown had dimensions of approximately 0.6 x 0.4 x 0.4 mm.

Smaller crystals were obtained with ammonium sulphate concentrations of 2.4 to 2.5 M but these had rounded edges between faces rather than the sharp edges (as shown in figure 3.4 above). Note also the yellow colour of the crystals caused by the cofactor.

3.2.5 Crystallisation of MGL2

The first attempts at crystallising MGL2 were performed identically to those for MGL1. The protein was concentrated to 6 mg/ml and resuspended in 0.1 M HEPES pH 7.5, 40 μM PLP, 30 μM DTT but the results from the screens gave no promising leads. After DLS analysis the buffer was changed to 20 mM imidazole buffer pH 6.5, 250 μM PLP, 1 mM DTT. Using this new buffer produced some small crystals, the best results are listed below:

Precipitant	Buffer (0.1 M)	Salt (0.2 M)	Result
15% PEG 8k	MOPS pH 6.5	$NaKPO_4$	Heavy crystalline precipitate
20% PEG 8k	TRIS pH 8.5	NaCl	Clusters of very small needles
15% Me2k PEG	MOPS pH 6.5	$MgCl_2$	Few small rods
15% Me2k PEG	HEPES pH 7.5	$MgCl_2$	Few feathered rods
20% PEG 4k	HEPES pH 7.5	$AmmSO_4$	Heavy crystalline precipitate
30% PEG 400	HEPES pH 7.5	$MgCl_2$	Very small cubic crystals
8% PEG 8k	TRIS pH 8.5		Few clusters of needles
15% PEG 4k	TRIS pH 8.5	$MgCl_2$	Few small rods

Table 3.2: Most successful results of attempts to crystallise MGL2

It was not possible to improve upon these results further and so a new expression construct of MGL2 was produced in an attempt to improve the prospects for producing crystals. The previous expression construct had the hexahistidine tag at the N-terminus and also 5 residues missing from the original sequence which were replaced by 8 residues from the cloning vector (McKie, 1997). It was considered possible that the incorrect residues and N-terminal tag were preventing crystallisation in some manner and it was hoped that replacing these would lead to the production of useful crystals. Therefore the new construct was designed to remove the incorrect residues and replace with the correct sequence (see section 2.2.16 for method). The expression vector was also changed to position the hexahistidine tag at the C-terminus rather than the N-terminus as already used for MGL1.

This new MGL2 was examined for the retention of activity in various buffers at both 4 °C and 20 °C. It was decided to avoid using sodium phosphate buffer in crystallisations because of the likelihood of producing phosphate crystals. Thus a suitable alternative buffer had to be chosen. The previous enzyme had been shown to be most stable in sodium phosphate buffer at pH 7.5 and this pH was used as a starting point for the selection of others (McKie, 1997). The buffers chosen to be examined were: BIS TRIS pH 6.5, BTP pH 7.5, HEPES pH 7.5, PIPES pH 7.0 and TES pH 7.5. All were used at 0.1 M with 80 μM PLP and 60 μM DTT added. Enzyme was expressed and purified as already published. Pooled enzyme was then dialysed overnight at 4 °C into the relevant buffer and stored at a concentration of 1 mg/ml at either 4 °C or 20 °C. The enzyme was then assayed with DL-homocysteine as detailed (section 2.3.6) over two weeks.

The results of these assays are shown in figure 3.5 and 3.6.

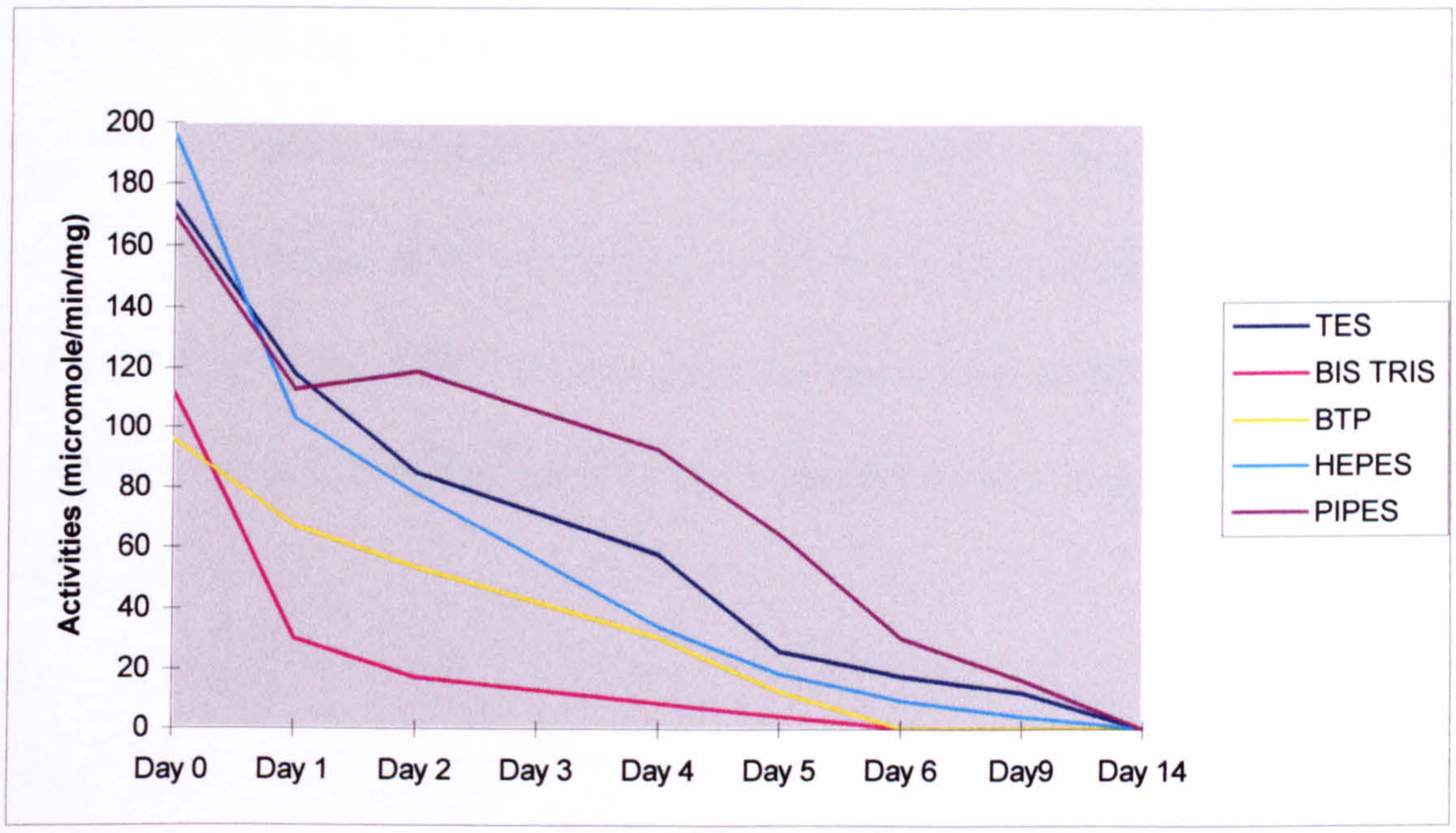

Figure 3.5: $\alpha\gamma$-elimination activities of MGL2 toward DL-homocysteine after storage in different buffers at 20 °C.

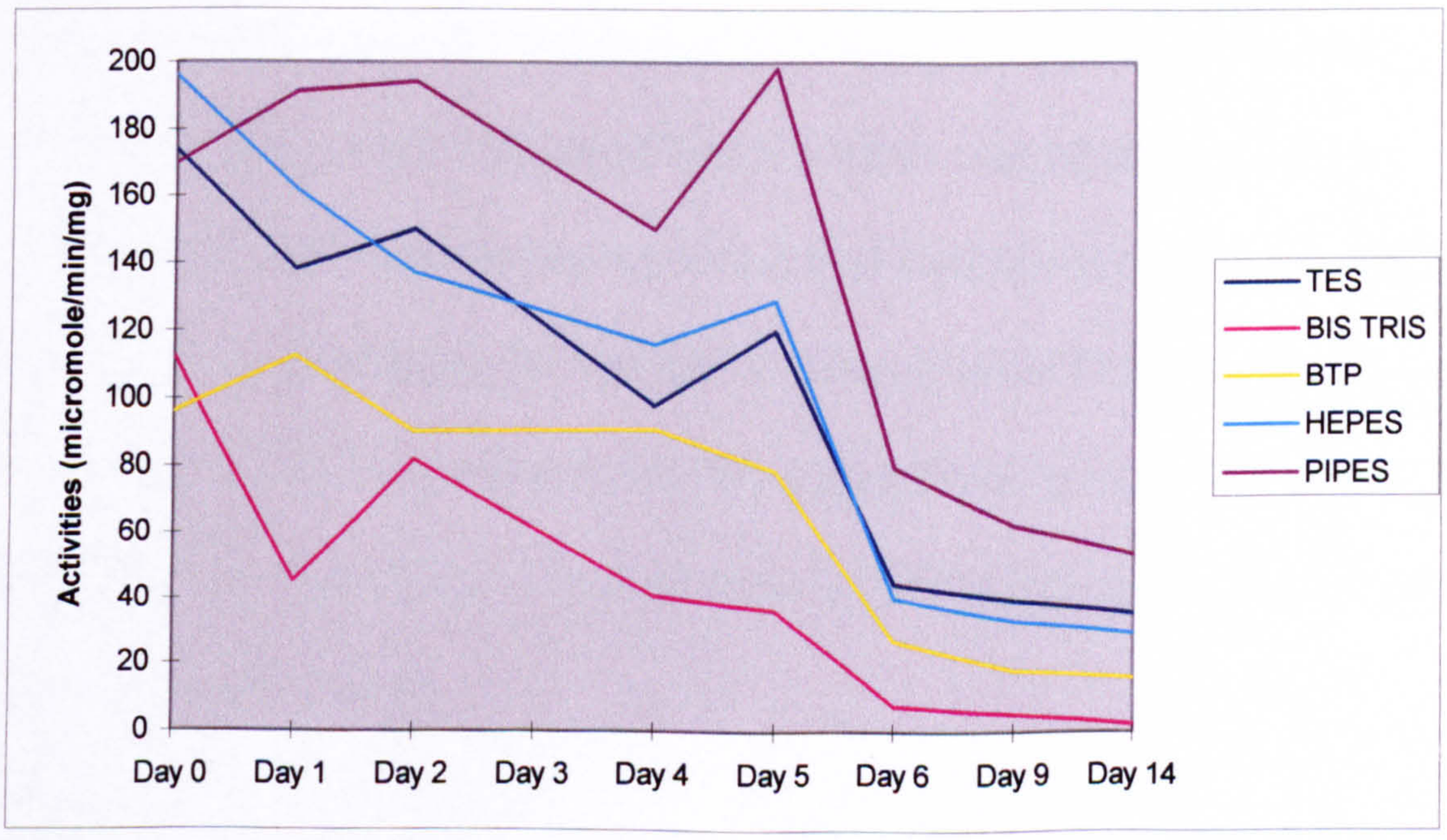

Figure 3.6: $\alpha\gamma$-elimination activities of MGL2 toward DL-homocysteine after storage in different buffers at 4 °C

The results clearly showed that there is no measurable activity after 2 weeks when the enzyme is kept at 20 °C. The results also showed that the enzyme lost considerable activity when stored in BTP and BIS-TRIS buffers regardless of the

temperature. Much better retention of activity was exhibited in PIPES, HEPES and TES. The highest values recorded were in PIPES and so this buffer was selected for use in crystallisation. However, despite these changes, and attempts at co-crystallisation with propargylglycine, it was not possible to produce suitably diffracting crystals of MGL2 during the course of this study. However the attempts made were not exhaustive and further experimentation may yield good quality crystals.

3.2.6 Structure Solution

This section of the chapter describes the results of the structure solution process for native and inhibitor complex MGL1. As stated in the aims of this project the intention was to solve the structure of MGL and from this learn something of its function *in vivo*. It was envisaged that if suitable crystals were grown then the technique of molecular replacement would be used to facilitate the process rapidly and easily. This assumption was due to the number of PLP-dependent enzyme structures already deposited in the PDB (albeit of the α-and β-families) and the anticipation of the release of the coordinates of CBL (Clausen *et al.*, 1996), a member of the γ-family.

3.2.6.1 Data Collection and Processing

Complete 2.18 Å datasets were collected from cryocooled single crystals of both the holoenzyme and the enzyme-inhibitor complex at station 9.5 at the SRS, CLRC Daresbury Laboratory at a wavelength of 1.1 Å on a 30 cm MAR image plate. A summary of the statistics for each is presented below:

Structure		Holoenzyme	Inhibitor Complex
Resolution Range (Å)		25.0-2.18	25.0-2.18
Oscillation Range (°)		0.5	0.5
Temperature (K)		100	100
Unique Reflections		49696	48724
Completeness	Overall (25.0 - 2.18)	98.6	97.6
	Highest (2.55 - 2.18)	97.2	77.4
Redundancy		2.3	3.6
R_{merge}		4.5	6.7

Table 3.3: Data collection and processing statistics for MGL1 holoenzyme and inhibitor complex

As mentioned previously these crystals exhibited diffraction to 1.6 Å but with differing levels of anisotropism. This was seen subsequent to collecting complete 2.18 Å datasets. It was not possible to collect higher resolution data because of a lack of time. Decreasing the wavelength of the X-ray beam away from its optimum of 1.1 Å would have resulted in a much lower intensity beam and therefore a significant increase in the time required for each frame.

Both holoenzyme and inhibitor complex crystals contain a dimer in the asymmetric unit and belong to the same spacegroup and possessed identical dimensions, that is:

Spacegroup	a	b	c	α	β	γ
$P3_112$	88.26 Å	88.26 Å	217.85 Å	90°	90°	120°

Table 3.4: Unit cell parameters for crystals of holoenzyme and inhibitor complex MGL1.

Attempts were also made at this time to soak crystals in solutions containing heavy metal atoms. The lack of time and small number of crystals available allowed only a few attempts. 1 mM mercury acetate and 1 mM mercury chloride were used and crystals were soaked for between 1 and 30 minutes. The diffraction exhibited by these crystals after exposure to both compounds was significantly poorer than that shown for untreated crystals. It was decided not to collect data from these crystals and whilst waiting for the growth of larger crystals a molecular replacement solution was sought.

3.2.6.2 Structure Solution by Molecular Replacement

It was envisaged that it would be possible to find a Molecular Replacement (MR) solution for MGL1 using one of the numerous PLP-dependent enzyme structures available. No γ-family enzyme structure was initially accessible from the database and so attempts were made using members of the α and β families. These all proved unsuccessful whether using the whole polypeptide chain or selected fragments. After the eventual release of the coordinates for CBL from *E. coli* it was possible to use a search model with a much higher degree of sequence similarity and presumed structural homology. Comparing the sequence of MGL1 with those used to find a MR solution shows that the α- and γ-families of PLP-dependent enzymes may share a similar fold but sequence identity is quite low (Table 3.5).

Enzyme	Sequence Identity	Sequence Similarity
Cystathionine β-lyase (Clausen *et al.*, 1996)	28.8%	36.9%
Tryptophanase (Isupov *et al.*, 1998)	15.2%	41.4%
Tyrosine phenol lyase (Antson *et al.*, 1993)	14.6%	36.9%
Aspartate aminotransferase (Okamoto *et al.*, 1994)	11.4%	37.6%
Ornithine aminotransferase (Shen *et al.*, 1998)	15.2%	32.3%
D-amino acid transferase (Sugio *et al.*, 1995)	9.4%	16.8%

Table 3.5: Sequence comparison of enzymes used to find a MR solution with MGL1.

The first attempts at MR with AMoRe (Navaza, 1994) using inhibitor data from 25 to 2.2 Å and the entire CBL molecule as a search model gave no clear solution. Subsequently the CBL model was truncated to more closely resemble MGL1. Insertions in the primary sequence alignment were removed and where individual residues differed they were changed to either alanine or glycine. See figure 3.7 for an alignment of the two sequences. This again gave no single solution but by using data to only 4.5 Å and the truncated CBL model a best solution with an R-factor and correlation coefficient of 47.3% and 36.0%, respectively was obtained. The packing was verified visually using Setor (Evans, 1993). This provided a usable map with which to start modelling the structure of the inhibitor complex.

The solution for the holoenzyme structure was solved rapidly by using the phases from the inhibitor complex model directly without the need to use molecular replacement. This does result in greater model bias to be carried over from the

inhibitor from the active site in early stages of refinement. The Free R-factor set of reflections set aside for monitoring the progress of refinement were identical for both datasets.

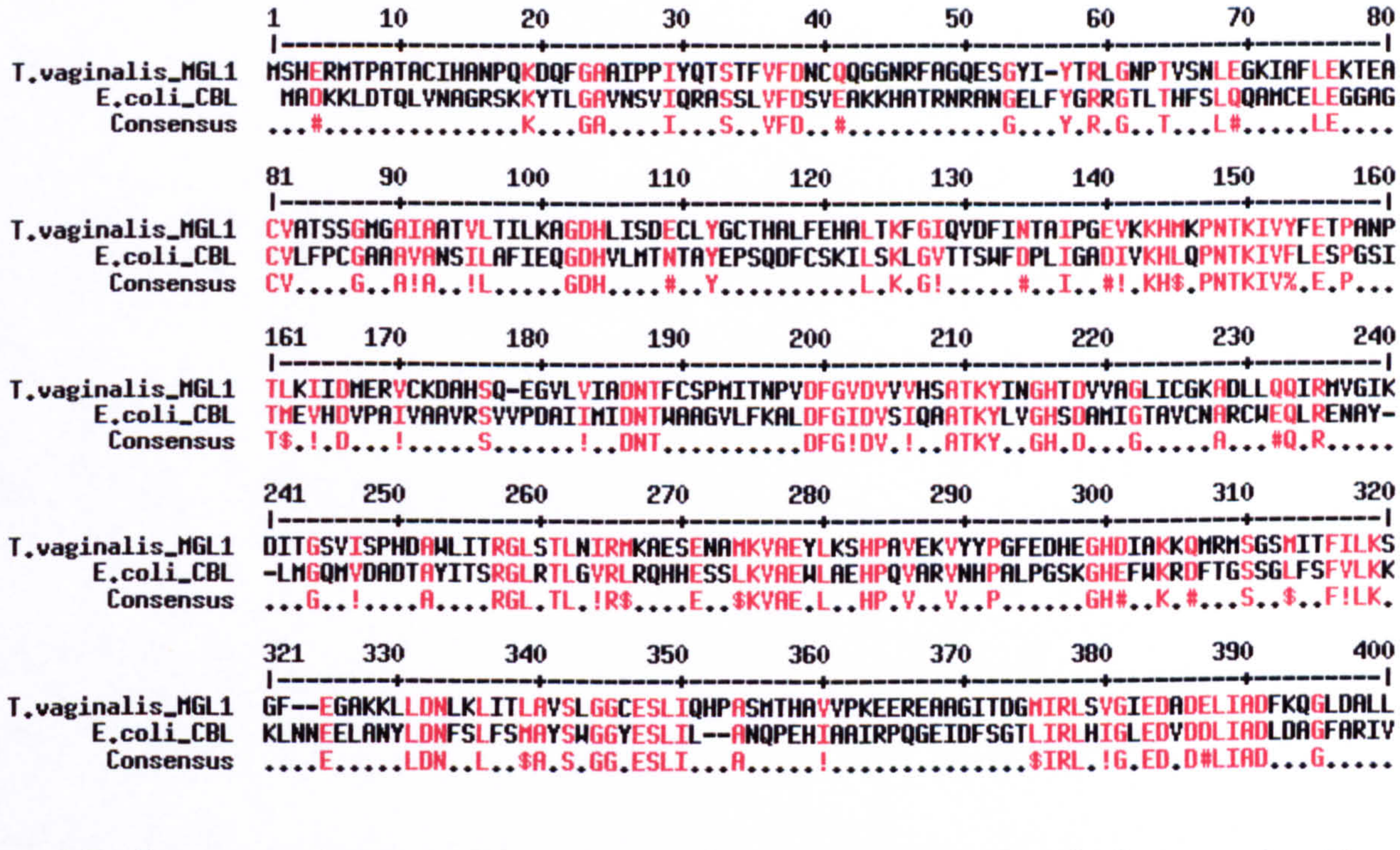

Figure 3.7: Sequence alignment of *T. vaginalis* MGL1 and CBL from *E. coli*. Conserved residues and similarities are highlighted in red. The consensus sequence is given in the third line.

3.2.6.3 Refinement of MGL1

Refinement of MGL1 was performed as outlined in section 2.1.8. Table 3.6 and 3.7 lists the improvements in the R and R_{free} factors for the native and inhibitor complex structures after successive rounds of refinement using all data from 40 to 2.18 Å.

Round of Refinement	R-factor	R_{free}
1. After AMoRe	54.9	53.5
2. Main chain 14% incomplete	36.5	40.6
3. Main chain complete, 352 waters added	29.7	34.4
4. Bulk solvent correction added	20.8	25.5
5. Poor waters removed, NCS restraints relaxed	18.9	24.3
6. 2 glycerols identified	16.3	20.2
7. Final round	14.7	19.0

Table 3.6: MGL1 inhibitor complex refinement progression.

Round of Refinement	R-factor	R_{free}
1. Using phases direct from inhibitor complex	22.0	27.0
2. Removed poor atoms	20.1	26.3
3. 456 waters, cofactor and bulk solvent correction added	16.3	21.2
4. Final round	16.2	21.2

Table 3.7: MGL1 holoenzyme refinement progression.

3.2.6.4 MGL1 Model Building

The structure of CBL itself was used to guide placement of residues using the molecular replacement map calculated from the MGL1 inhibitor complex data. Areas of large divergence between the two structures lead to poor electron density and there were several breaks in the main chain where it was impossible to assign Cα positions as well as numerous incomplete side chains. In particular

the following regions were not visible 13-22, 41-55, 227-241, 294-296, 316-325 and 348-370 (MGL1 numbering). Advantage was taken of the 2-fold NCS in the asymmetric unit, allowing a single monomer to be constructed and then transformed into a dimer to complete the asymmetric unit. The assumption was that during early stages of refinement, differences between the two monomers would be insignificant compared to any errors and poor quality of the electron density map. Thus the maps produced by refinement were averaged over a single monomer and the modelling process repeated. Until density was visible for the entire main chain and waters added this was repeated and then subsequently each atomic position checked individually.

3.2.7 MGL1 Structure

The following part of this chapter describes the crystal structure of MGL1 and the inhibitor complex with L-propargylglycine. The large size of the structure prevents a detailed description of all residues and their apparent roles, instead those with key roles and/or positions will be mentioned as appropriate. In the multimeric structure of the enzyme, residues from an adjacent monomer involved in an interaction will be marked by an asterisk (*).

3.2.7.1 MGL1 Monomer

The MGL1 monomer can be divided into three domains: 1) An extended loop N-terminal domain; 2) a PLP-binding domain; and 3) a C-terminal domain (Figure 3.8). The enzyme shares the same overall fold as the other published γ-family members and is closely related to the α-family. Both families possess the PLP-dependent transferase fold, the γ-family have in addition an N-terminal domain.

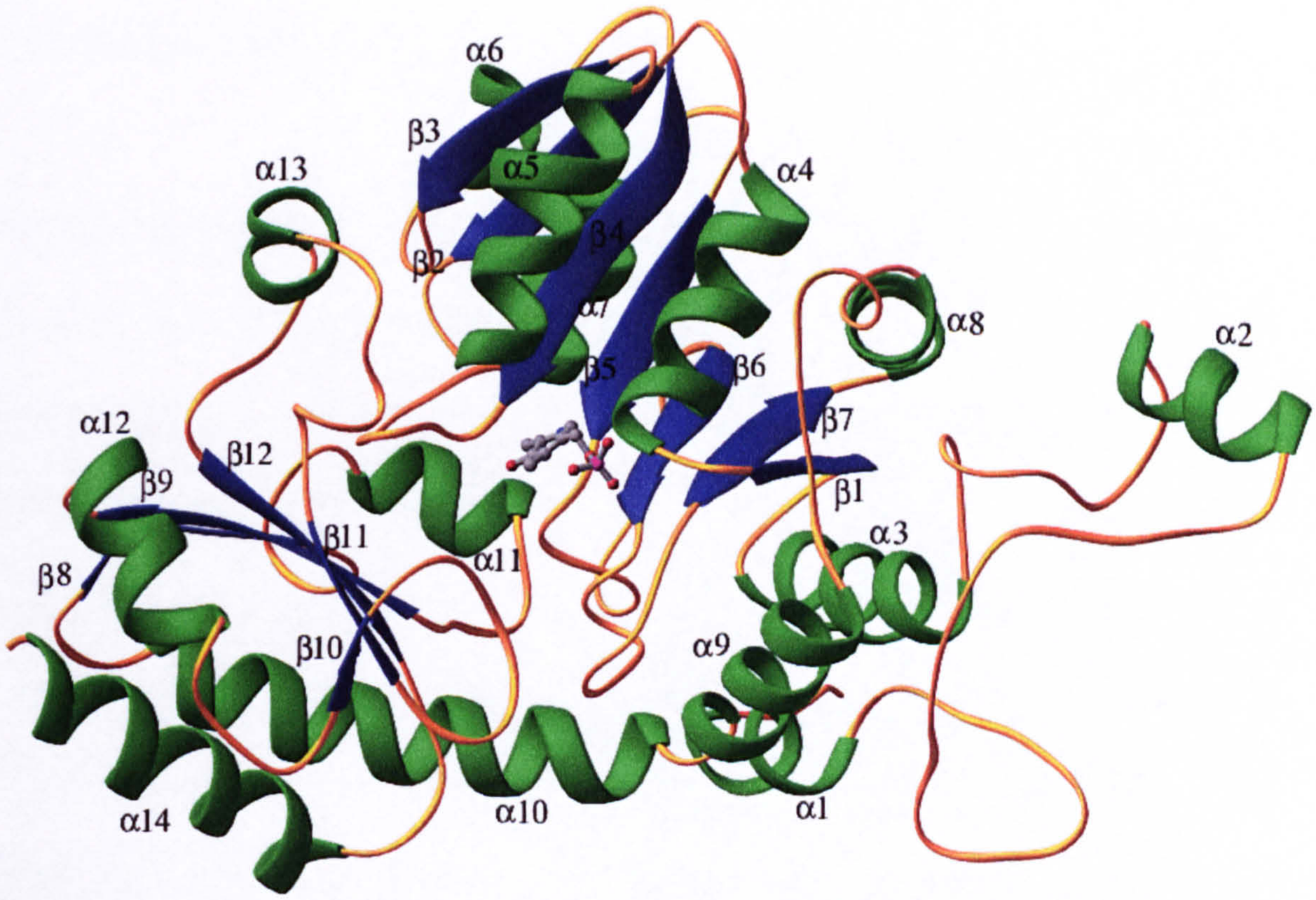

Figure 3.8: Ribbon diagram of the MGL1 monomer. α-helices are coloured green, β-strands are coloured blue and numbered individually. The PLP cofactor is represented as a ball and stick model and coloured by atom type. Figure produced using the programmes RIBBONS (Carson, 1997) and THE GIMP (http://www.gimp.org).

The N-terminal domain of MGL1 (residues 4-60), comprises helices α1 and α2 and a fairly open loop structure. Comparing other γ-family members reveals that this is the most varied domain, in both sequence and length. CBL differs having a short α-helix and β-strand between the two helices (Clausen *et al.*, 1996) and CGS has no defined secondary structure elements and a more open conformation (Clausen *et al.*, 1998; Schoenbrook *et al.*, 1999). Helix α1 of MGL1 packs against helices α9 and α3 of the PLP-binding domain. Residues 22-35 interact to form

part of the tetramer interface, while helix α2 forms part of the opening to the active site and is critical in mediating monomer-monomer interactions.

The PLP-binding domain (residues 61-258) is the largest domain with a core seven stranded, mainly parallel, β-sheet surrounded by seven α-helices. The strands are arranged in the order A, G, F, E, D, B, C, with G in an antiparallel direction. The cofactor is bound at the edge of the monomer, between strands β E, βF and helices α4, α5, where it would be accessible to solvent. However, the interactions of the C-terminal domain and the N-terminal domain of the adjacent monomer in the catalytic dimer bury the cofactor and form the active site. Figure 3.8 shows quite clearly the distorted helix α5. The regular geometry is distorted by His120 and Ala121, which are close to Lys238* and Asp239*. These residues are implicated in the enzyme mechanism and shall be discussed further below.

The C-terminal domain (residues 259-396) is composed of a five strand β-sheet with five α-helices. The strands are in a mainly anti-parallel arrangement in the order H, I, L, K, J. The domain is dominated by helix α10. This has a large kink induced by Thr259 which connects to helix α9 of the PLP-binding domain. Together helices α9 and α10 extend for almost the entire length of the monomer. Helix α10 is situated on the outer edge of the domain, between helix α14 and the loops at the β-sheet of the C-terminal domain. The loop between strand βK and helix α13 (residues 348-357) is critical in sculpting the entrance and restricting access to the active site.

3.2.7.2 Disulphide Bridges in MGL1

There are no intact disulphide bridges visible within the structure of MGL1. The enzyme has not been indicated as being located extracellularly so this is perhaps as expected. The crystallisation solution also contained high concentrations of DTT (1 mM) and thus may also be expected to reduce any bonds. Examination of the structure reveals two cysteine residues with the potential to form a disulphide under appropriate conditions. Cys80 and Cys223 have their Sγ 4.1 Å apart in both structures but would be able to move closer by rotation around the Cβ. Figure 3.9 shows the relative positions to each other and the position of Cys80 on strand βA and Cys223 on βG. The strands are positioned in an antiparallel direction. It can be seen from an alignment of MGL sequences (Figure 3.10) that none of the residues shown in figure 3.9 are conserved including the cysteines except for MGL1 and 2 from *T. vaginalis*. Neither residue is involved in intersubunit interactions or catalysis in anyway and it is difficult to see how the formation of a disulphide bridge would change this. One possibility is that disulphide formation may shift helix α4 enough to disrupt either tetramer binding or cofactor binding via Gly86 and Met87.

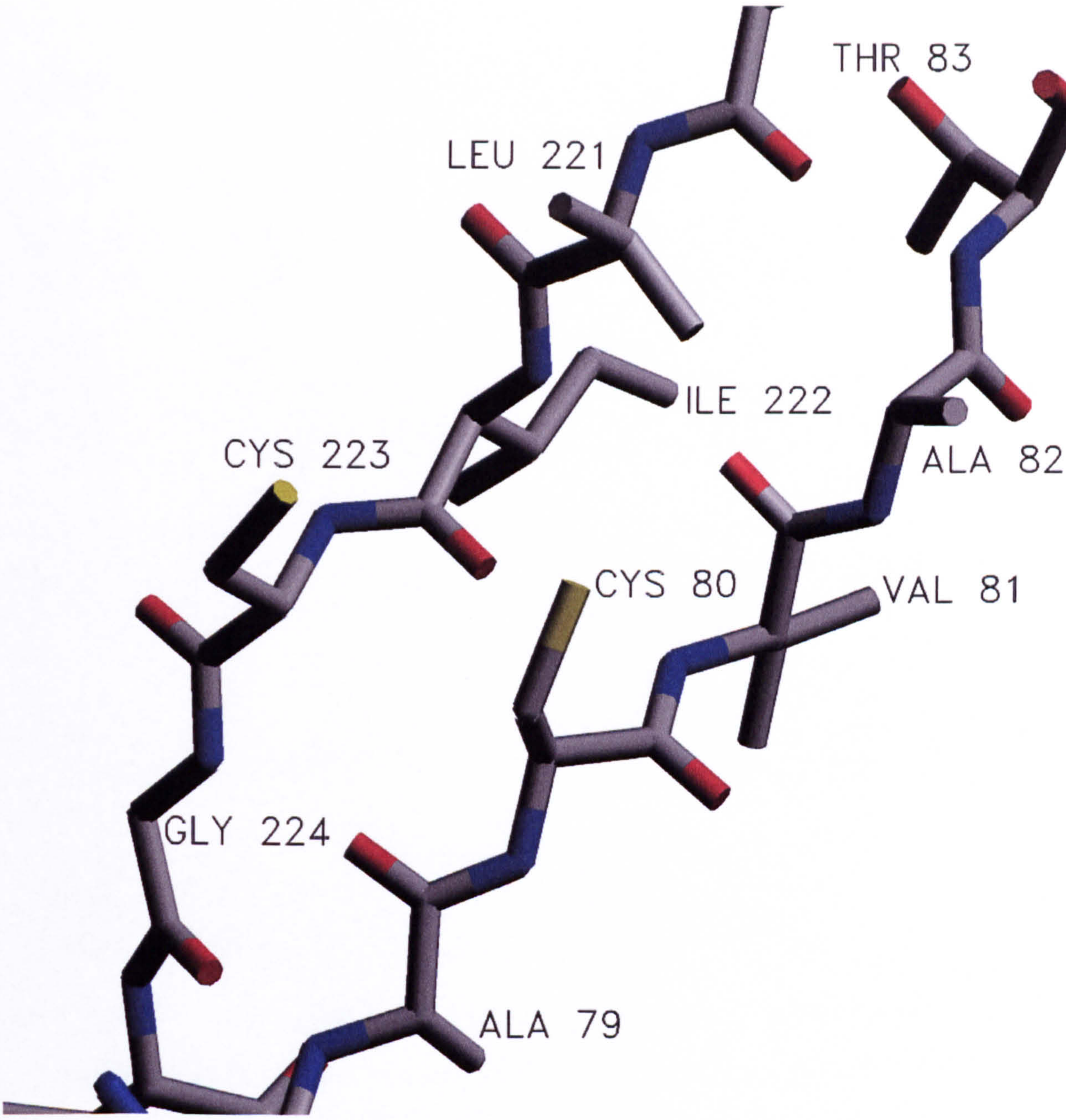

Figure 3.9: Arrangement of Cys80 and Cys223 in MGL1. Residues are coloured according to atom type and labelled. The distance between Sγ atoms is 4.1 Å in both holoenzyme and inhibitor complex structures and identical between monomers.

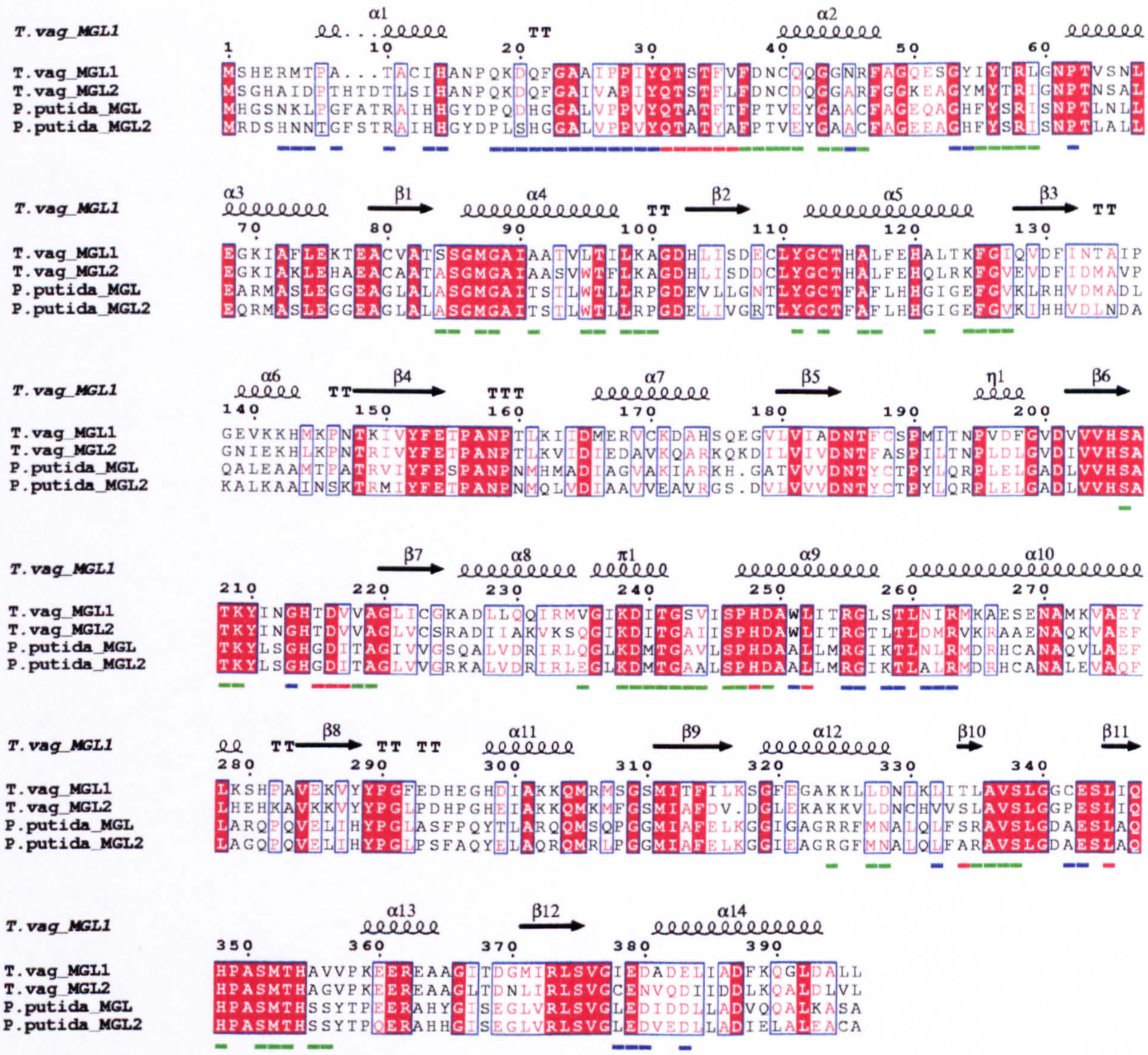

Figure 3.10: Alignment of MGL sequences from *T. vaginalis* and *P. putida*. 100% conserved residues are in red and 75% conserved in blue. Secondary structure elements in MGL1 as denoted by DSSP (Kabsch and Sander, 1983) are indicated above the sequences. Residues involved in subunit contacts are underlined in colour, blue for monomer-monomer interactions, green for dimer-dimer and red if involved in both. The alignment was carried out using Multalin (Corpet, 1988)

3.2.7.3 *cis*-Prolines in MGL1

There are two *cis*-prolines in the structure of MGL1, Pro156 and Pro159. Both are rigidly conserved in MGL sequences identified so far and also CGS but only Pro156 is conserved in CBL. Neither residue is involved in intersubunit contacts but they have an indirect role in cofactor binding. Both residues are positioned in type IV β-turns (residues 154-157 and 158-161) between strand βD and helix α7.

This positions Asn158 in the optimal site for making hydrogen bonds with O3' of the cofactor and Arg373. The role of O3' is important in the catalytic mechanism proposed here and Arg373 is conserved in all PLP-dependent enzymes that utilise amino acids as substrates. There are two conserved *cis*-prolines in AAT although not in the same arrangement as MGL1. The two residues are numbered 138 and 195 for the *E. coli* enzyme but are positioned within van der Waals contact with each other. Pro195 is analogous to Pro159 of MGL1, coordinating the adjacent asparagine residue (Asn194 in AAT). However mutagenesis studies have revealed that neither of these residues are critical for structure or activity (Birolo *et al.*, 1999). They have been postulated to play a subtle role in optimising catalysis but are substitutable for alanine with very little deleterious effect.

3.2.7.4 Intersubunit Interactions

There are extensive contacts made between adjacent monomers in the catalytic dimer with residues from the adjacent subunit being essential for activity, throughout the text residues from the neighbouring subunit involved in interactions are marked with an asterisk (*). The MGL1 homotetramer is constructed by symmetric association of two catalytic dimers. Analysis of the subunit interactions was performed using the protein-protein interaction server (Jones and Thornton, 1995). Residues involved in intersubunit interactions are indicated in figure 3.10 in an alignment with other MGL sequences. Those residues at the monomer-monomer interface are underlined in blue, those in the tetramer interface green and those involved in both are underlined in red. The most significant are mentioned individually below.

3.2.7.4.1 Monomer-Monomer in the Catalytic Dimer

The association of the two monomers in the asymmetric unit buries approximately 2600 $Å^2$ of solvent accessible surface in an area of approximately 67 Å by 39 Å in both structures. The interface formed is predominantly non-polar but does include 36 intersubunit hydrogen bonds. The symmetry between the two dimers can be clearly seen in figure 3.11. The dimer is formed predominantly by close association of the PLP-binding domains, which comprises 51% of the interface. The "clamp" of the N-terminal domain contributes 31% of the interface surface area, interacting with both the PLP-binding and C-terminal domains of the adjacent monomer.

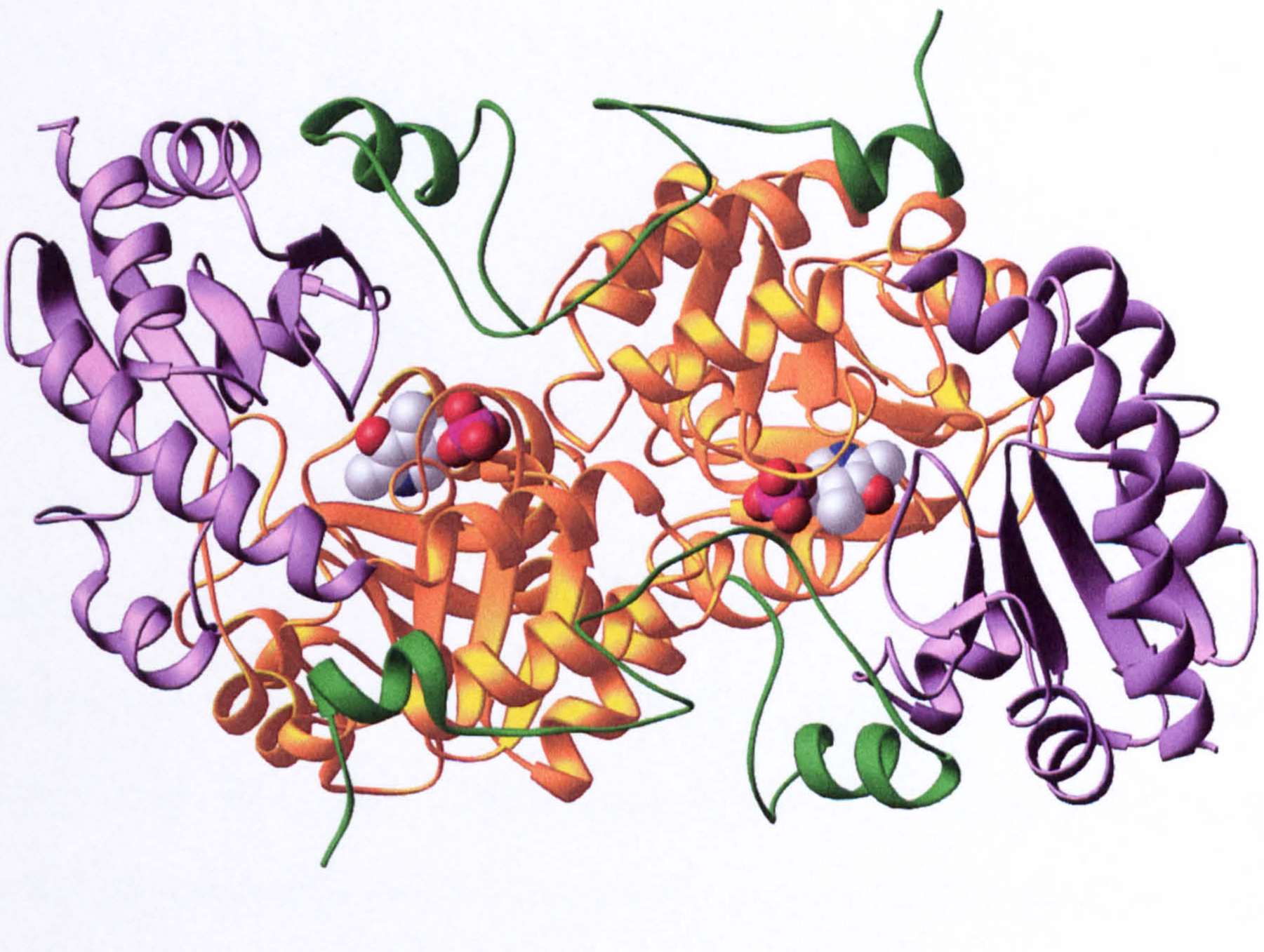

Figure 3.11: MGL1 dimer. The three domains are highlighted in colour: green = N-terminal; gold = PLP-binding; purple = C-terminal. The PLP cofactor is represented as a spacefilling model. The figure was produced using RIBBONS.

There are a number of residues playing a significant role in dimer formation; Gln31 forms a hydrogen bond between its amide nitrogen and the carbonyl oxygen of Thr215* and in addition also hydrogen bonds via waters 55, 84 and 31 to the amide nitrogen and side chain hydroxyl of Thr215*. In this region, also close to the dimer surface involved in tetramer formation, these three waters mediate interactions between Ser33 and residues Thr208*, Thr215*, Ser338* and Leu339. Phe37 in concert with Phe47, Ile55, Leu59, Val337* and Met352* forms a hydrophobic patch along the active site channel leading to the cofactor. Phe37, Phe47, Val337 and Met352 are all highly conserved in MGL enzymes with Ile55 and Leu59 being substituted only for hydrophobic residues. Another hydrophobic patch formed by the two monomer interactions involves residues Leu117, Leu122, Leu95, Ile240 and Leu95* and Ile240* and a stacking interaction between Phe125 from both subunits.

3.2.7.4.2 Tetramer formation

The MGL1 tetramer is constructed by symmetric association of two catalytic dimers, as shown in figure 3.12. The approximate dimensions of the interface between the two dimers is 64 Å by 46 Å and buries around 4100 $Å^2$ of solvent-accessible surface, which is comparable to other γ-family enzymes. The tetramer is formed predominantly by hydrophobic interactions which comprise 70% of the total buried surface area but also includes 33 intersubunit hydrogen bonds. The N-terminal region, despite being the smallest domain of the enzyme, is responsible for 55% of the interface. This highlights its key role in all intersubunit interactions as well as active site formation. The PLP-binding and C-terminal domains play less central roles and form 20% and 25% of the interface,

respectively. Residues 25-33 form one β-strand (not identified by DSSP) in a two strand antiparallel β-sheet at the tetramer interface. The second strand is comprised of the same residues from the diagonally opposite subunit.

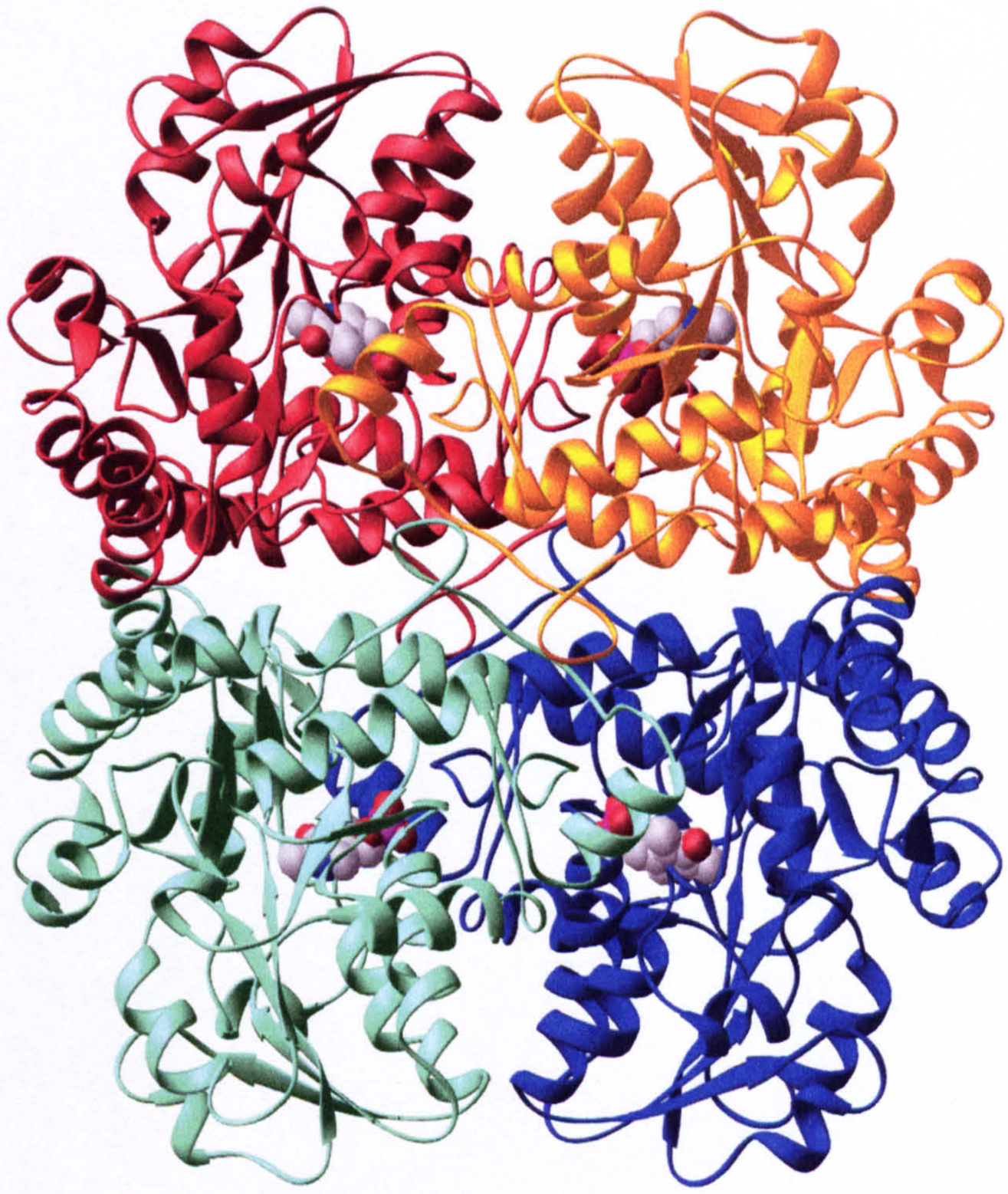

Figure 3.12: Ribbon diagram of MGL1 tetramer. Monomers are coloured individually in red, gold, blue and green. The cofactor is shown as a spacefilling model coloured according to atom type. The figure was produced using RIBBONS.

For the most part residues involved in tetramer formation are not conserved between MGL, CBL and CGS. There are equivalent regions involved but few of these are identical at the primary structure level.

A crucial interaction that is conserved is the intersubunit salt bridge made by Arg255 which lies antiparallel to its symmetry partner in the adjacent dimer and is coordinated by Asp216* (Figure 3.13). Alongside this is His214 which is also conserved and with His214* coordinates a water molecule. Trp251 on the other side of the salt bridge is not conserved. The sum of these interactions account for approximately 17% of the interface.

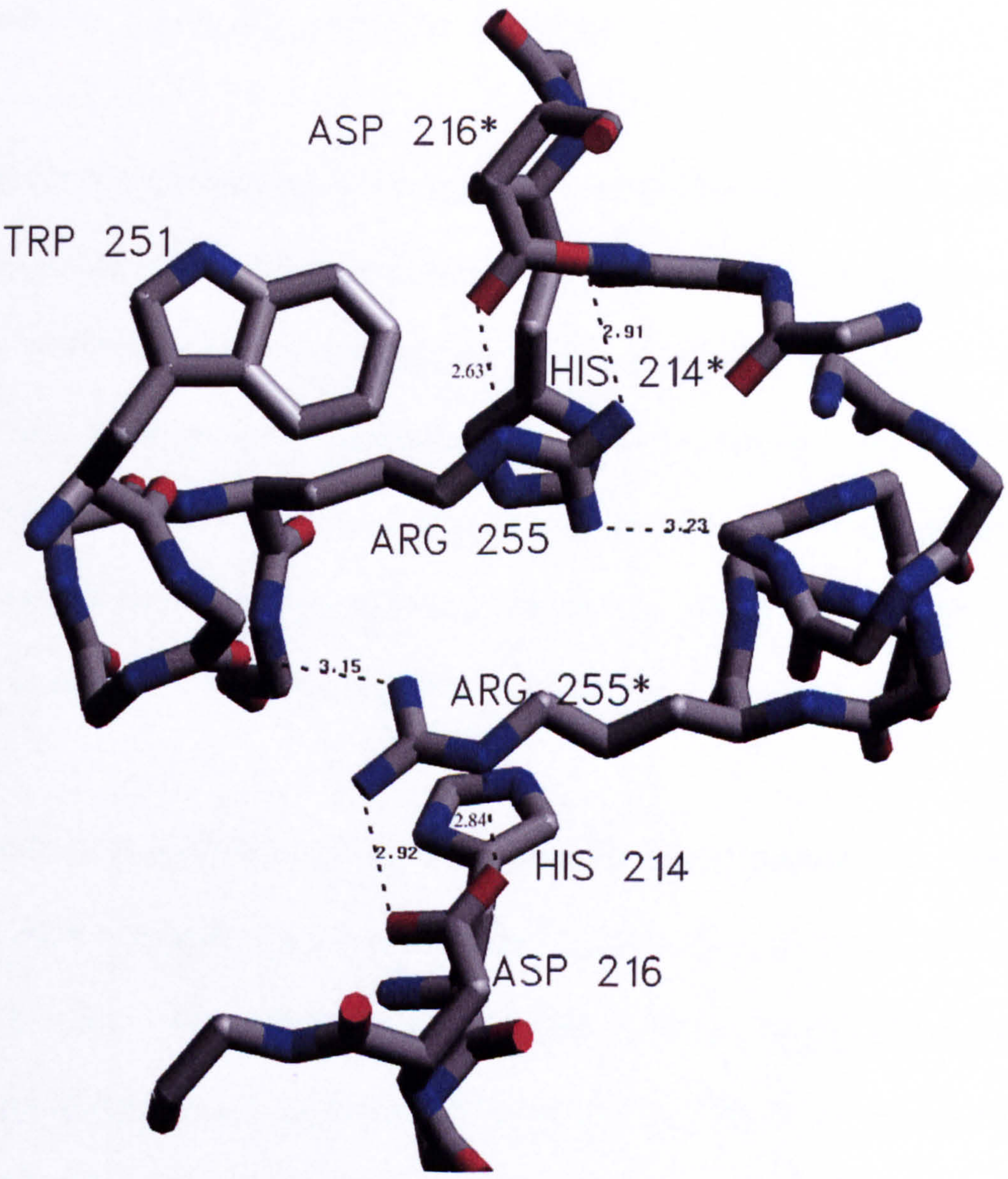

Figure 3.13: Salt bridge interaction between catalytic dimers. Atoms are coloured according to type, distances between key interactions are marked in Angstroms and denoted by dashed lines. Residues from adjacent dimer are indicated by *. Figure produced using Setor.

A large hydrophobic patch of residues from the N-terminal domain is also key in MGL1 and although these residues are not conserved in CBL and CGS there are analogous interactions. In MGL1, Phe22 is angled roughly perpendicular to the plane of Tyr30, Phe35, Tyr54 and Pro62 in an area which makes up approximately 12.5% of the total interface area. Tyr30 is also involved in an interaction on its opposite side with Pro28 and Ile29 to form another hydrophobic patch. Ile29 is conserved between CBL, CGS and MGL and is positioned adjacent to its symmetry partner in the opposite dimer.

3.2.8 Comparison of Holoenzyme and Inhibitor Complex Structures

There are no major conformational differences between monomers in the catalytic dimer of either model nor between the structures of the holoenzyme and inhibitor complex. Superimposing monomers onto each other gives an RMS deviation of 0.360 Å and 0.326 Å for main chain atoms in the holoenzyme and inhibitor complex respectively. Superimposing the dimers of the two structures gives an RMS deviation for main chain atoms of 0.251 Å.

However, a comparison of main chain atom B-factors between the holoenzyme and inhibitor complex models does reveal higher values for the holoenzyme (Figure 3.14). After allowing for a systematic variation, the most significant variation occurs for the residues 38-50 and 350-370. Both of these regions are in the entrance to the active site. Flexibility for this part of the structure is likely to be important to enable the active site to accommodate bulky substrates such as S-adenosyl methionine (Lockwood and Coombs, 1991b).

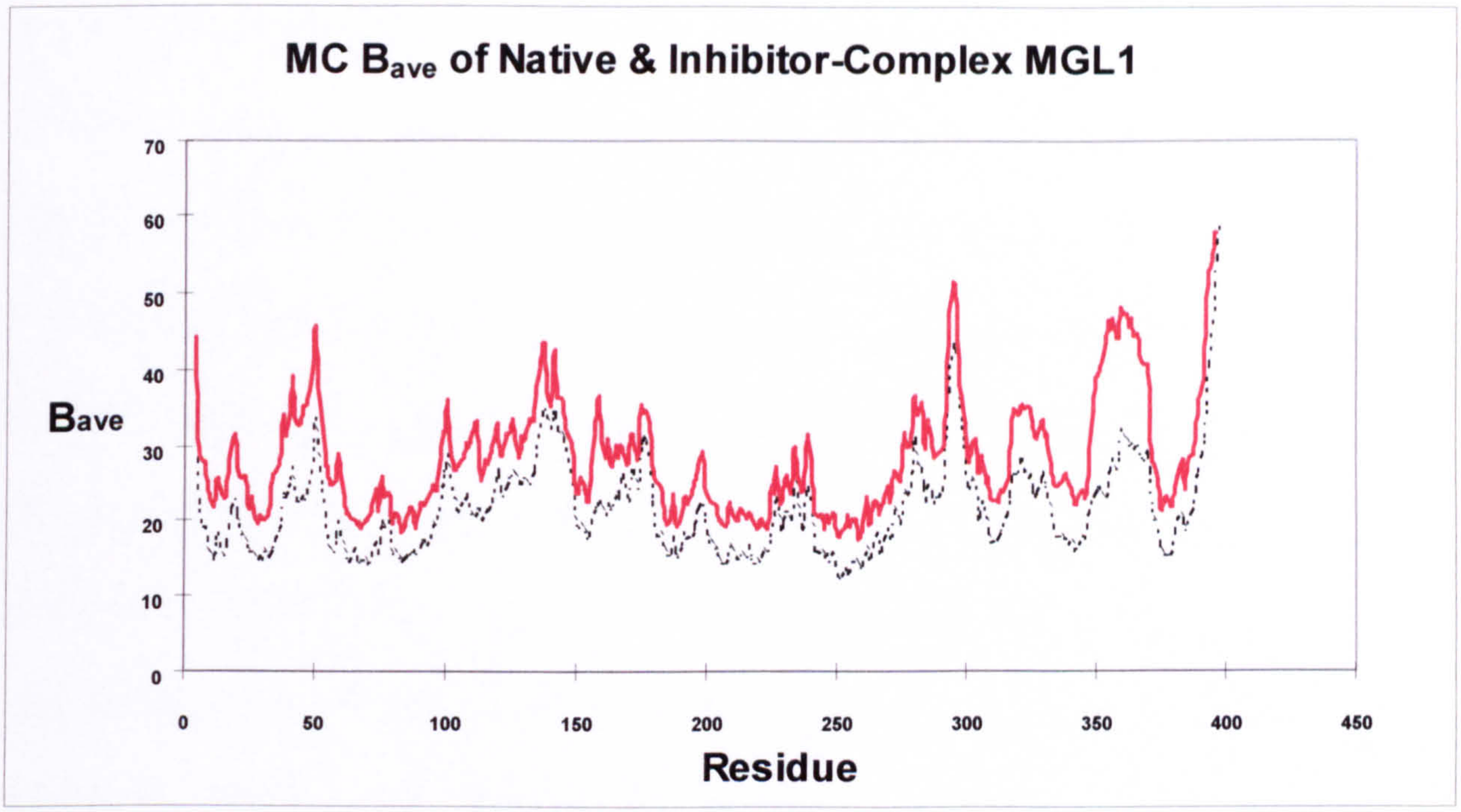

Figure 3.14: Comparison of atomic B-factors for holoenzyme and inhibitor complex. The native model values are indicated by the bold solid line, the inhibitor complex values by the dashed line.

3.2.9 Active Site Structure and Implications for the Catalytic Mechanism

The following description of the structure details the architecture in the active site around the cofactor and the residues implicated in the enzyme mechanism. The presence of a suicide inhibitor covalently bound allows a good estimate of the likely position of a bound substrate especially when studied in conjunction with a native enzyme structure.

3.2.9.1 Cofactor Binding

The active site is located at the subunit interface in the catalytic dimer with the cofactor held in place by a network of hydrogen bonds characteristic of PLP-dependent enzymes. All of the close contacts between the protein and cofactor are shown schematically in figure 3.15, including those involving the inhibitor L-

propargylglycine. Differences between the two structures are minimal as previously stated. In addition, Van der Waals contacts between the pyridine ring, Tyr111 and Ser206 also exist, restricting any vertical movement and holding the ring approximately parallel to the plane of Tyr111. As described for aspartate aminotransferase (Hayashi *et al.*, 1990), the positioning of the aromatic side chain should increase the electron sink character of the cofactor. From the second monomer, Arg58* forms a salt bridge to the phosphate group of the cofactor whilst also being close (3.1 Å) to the OH group of Tyr111. In addition Tyr56* hydrogen bonds to the PLP phosphate and Lys209 when the enzyme is in the external aldimine form, as seen in the inhibitor complex. In the holoenzyme Lys209 is covalently attached to the cofactor via a Schiff base. Also in the holoenzyme structure there is a sulphate ion positioned in the active site adjacent to Arg373 instead of the inhibitor carboxylate group.

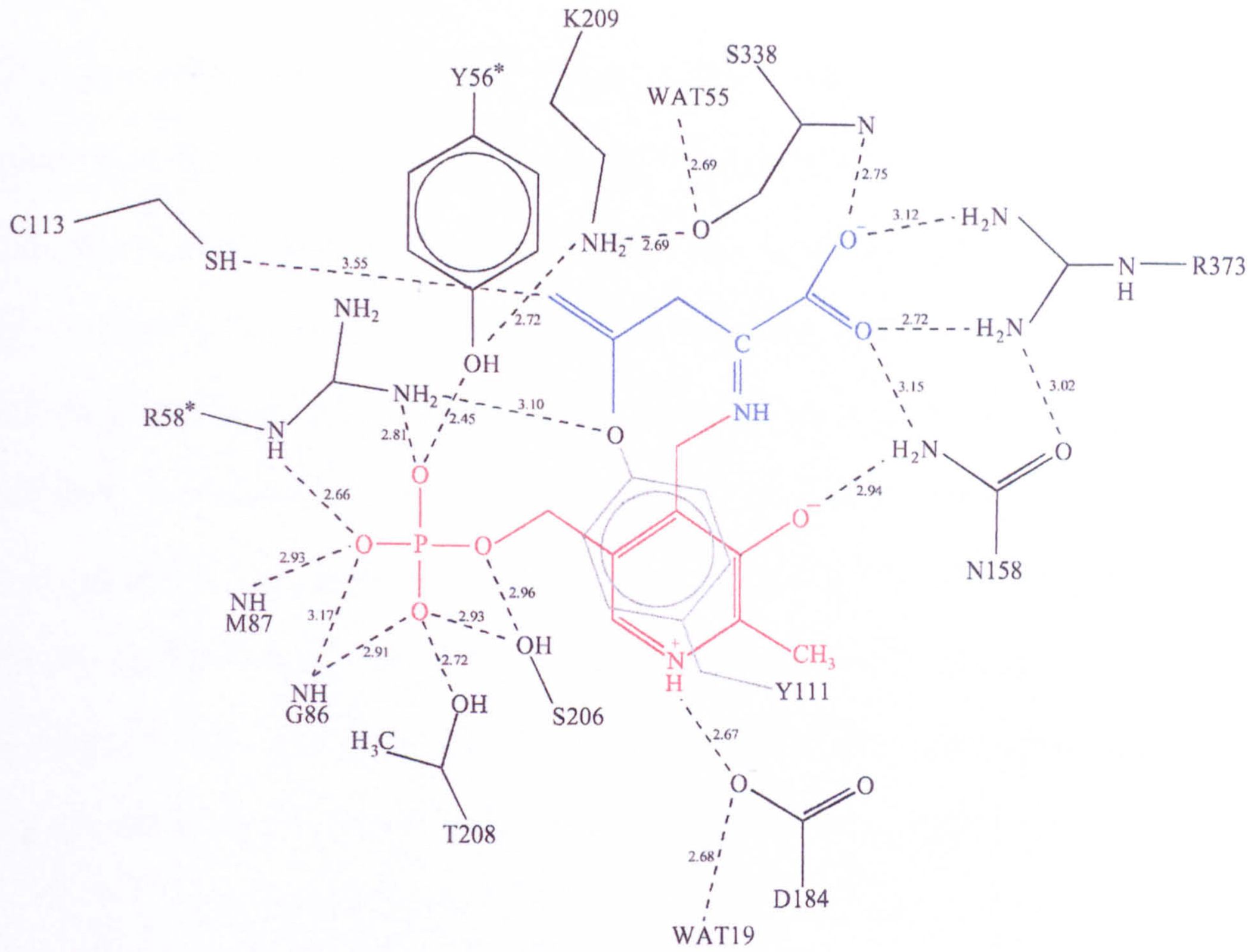

Figure 3.15: Schematic representation of the L-propargylglycine inactivated MGL1 active site. Cofactor atoms are coloured red, the covalently attached inhibitor molecule is shown in blue and protein atoms and water molecules are labelled in black. Non-covalently attached atoms are indicated by dotted lines and distances are shown in Angstroms.

3.2.9.2 Substrate Binding

Substrate binding site in the active site is primarily through the Schiff base link to the cofactor and also the coordination of the carboxylate group by Arg373. This residue is conserved in all PLP-dependent enzymes that have amino acids as substrate.

3.2.9.3 Substrate Access

Entrance to the active site is via a channel approximately 15 Å long from the position of the substrate Cα to the solvent, running along the dimer interface. Consequently residues from both monomers are involved. The cofactor is angled at ~90 ° to the channel and the inhibitor is positioned with the side chain pointing outwards as shown in figure 3.16. It can be seen that Phe47* and Leu59*, although not playing an active part in catalysis, are in this channel to the active site and form a hydrophobic patch along with Val337 and Ile55* (not shown). Also crucial are Cys113 and Lys238* which are positioned on the opposite side relative to His354*. These two residues and Asp239* appear to be functionally important but are not directly involved in the catalytic mechanism proposed.

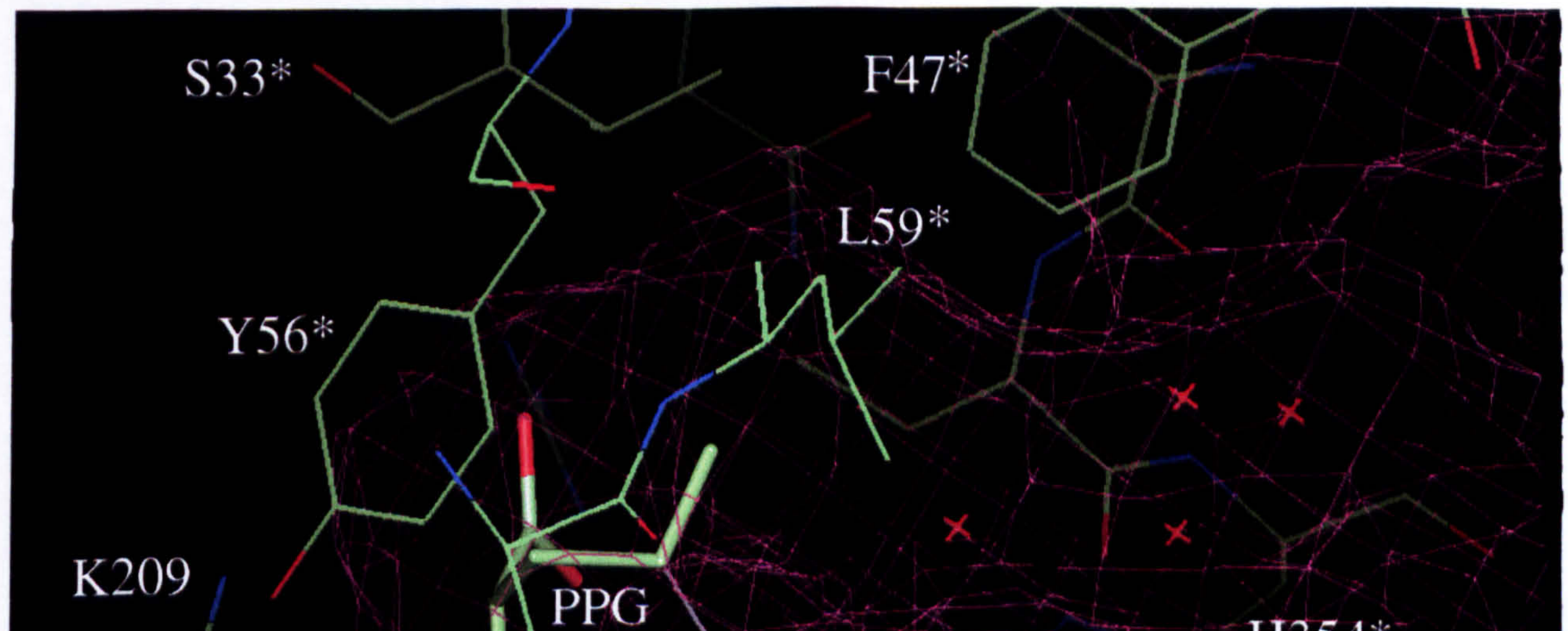

Figure 3.16: Diagram of the channel leading to the active site in the MGL1-inhibitor complex. The solvent accessible area is shown by pink lines. Water molecules in the inhibitor complex are indicated by red crosses. Key residues involved in the channel formation are indicated, with residues involved in catalysis also marked. Solvent accessible surface was calculated by the programme SURFNET using the inhibitor structure model without cofactor and inhibitor atoms. The figure was produced using QUANTA.

3.2.9.4 Covalent Attachment of L-Propargylglycine

The attachment of this suicide inhibitor of several PLP-dependent enzymes is unambiguous in the electron density. The earliest maps in the inhibitor complex model showed continuous density between the hydroxyl group of Tyr111 and the cofactor consistent with the binding of a substrate/inhibitor to the cofactor and enzyme. Fitting a putative propargylglycine molecule was simple following the predicted mechanism from (Washtien and Abeles, 1977). The active site nucleophile was thought to be either a cysteine or tyrosine residue, and although activity data from mutagenesis and chemical modifications implicated a cysteine, the model showed clearly a covalent link to Tyr 111. Thus propargylglycine is held in place by covalent attachment to Tyr111 and also via aldimine linkage to the cofactor. The distance from the Tyr111 side chain oxygen atom to the inhibitor Cβ in the final model is 1.8 Å. Arg373 plays its usual role in coordinating the carboxylate group of the substrate and in this case the inhibitor.

3.2.9.5 Comparison of the Holoenzyme and Inhibitor Complex Active Site

As has been previously stated there are no gross structural changes between the two crystal structures and this allows us to make certain assumptions concerning

binding of the substrate and residues involved in catalysis. However the active sites differ subtly and this provides more information on the likely mechanism.

A sulphate ion is substituted for the carboxylate group of the inhibitor in the holoenzyme structure and the Schiff base attachment of Lys209 to the cofactor replaces the attachment of the inhibitor amine group. Superimposing the two structures shows essentially no movement of the cofactor nor changes in the pucker of the ring. The position of Tyr111 shows minimal movement, the most significant factor is the poor electron density for this residue (Figure 3.17). As mentioned there is a systematic increase in B-factors for all of the holoenzyme, however the electron density for the cofactors in each model is similarly clear. The poor density exhibited by this residue adjacent to the cofactor can be taken to be because of mobility, a situation prevented in the inhibitor complex because of a covalent bond.

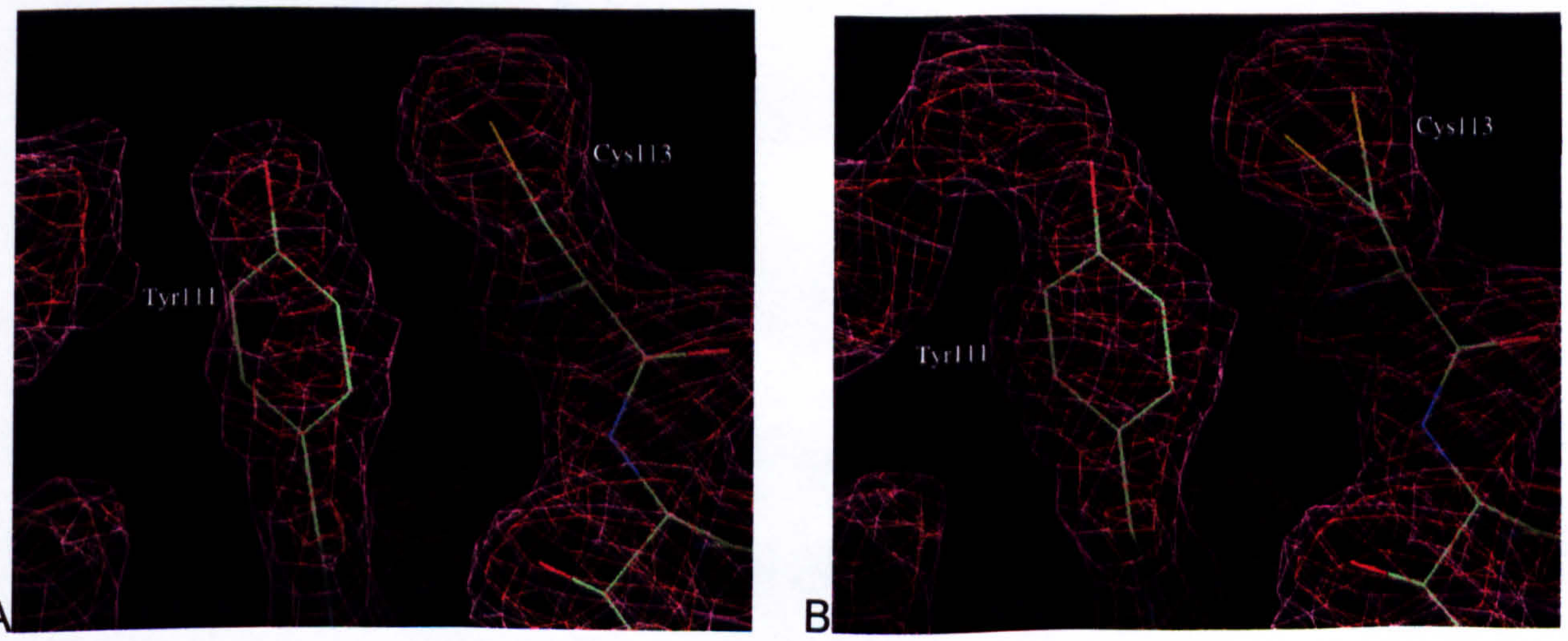

Figure 3.17: Comparison of active site electron density. In the native (A) and inhibitor complex (B) structures. 2 $F_O - F_C$ electron density maps contoured at 1 and 2 σ are shown in orange and pink respectively. Inhibitor molecule atoms have been omitted from B for clarity. The continuous electron density for the covalent bond between Tyr111 and propargylglycine is clearly visible. The figure was produced using QUANTA and the GIMP.

3.2.9.6 The Role of Cys113

It has been shown previously that Cys113 is very important for MGL activity. The mutant MGL1$^{(C113G)}$ was shown to have a greatly reduced activity towards all substrates, though still exhibiting measurable activity (McKie *et al.*, 1998). Chemical modification studies on MGL from *P. putida* have also indicated the importance of an unidentified cysteine residue in catalysis, but confirmed that it is not essential (Nakayama *et al.*, 1988a).

In the inhibitor complex Cys113 has an obvious dual conformation with an angle between the two Sγ of 33.4 °. This places the Sγ 3.5 Å from the Cδ atom of L-propargylglycine and 3.15 Å from -OH of Tyr111 in one conformation and 2.8 Å from the Nε of Lys238* which is 2.8 Å from the Oδ of Asp239* in the other. The same residue in the holoenzyme model does not appear to be in such a geometric arrangement and has only a single conformation despite appearing similar in the figure above. Continually through refinement any attempts at arranging Cys113 as a dual conformation were rejected because of negative density in the difference maps. It is probable that a limited range of movement of Cys113 exists in the native structure but it is certainly more pronounced in the inhibitor complex. Perhaps most interestingly in the holoenzyme are the differences in the active sites in which the distances between atoms are not equal. The distance between Cys113 Sγ and Tyr111 -OH is 3.66 Å and 4.17 Å, Sγ to Lys238* Nε is 3.71 Å and 4.13 Å and Lys238* Nε to Asp239* Oδ 2.65 Å and 2.98 Å. Thus one active site has a more relaxed arrangement with the distances between all atoms larger than the other.

The distances between the Sγ of Cys113 and the Cδ of propargylglycine and Sδ of homocysteine are likely to be very similar. The 3.55 Å distance seen in the inhibitor complex structure suggests that this is too great for a direct interaction between Cys113 and the substrate in catalysis unless a considerable conformational change occurs. The sequence information on MGL sequences (Figure 3.10) indicates that both Asp239 and Lys238 are conserved. The distances in the inhibitor complex suggest that a charge relay between the three residues would result in the Cys113 Sγ being ionised. The influence of this charge upon other atoms in the active site and therefore activity may be two-fold. The distance between the Sγ and Tyr111 OH is close enough to influence the charge upon this and with the added effect of Arg58* could conceivably produce a phenolate ion. This will be discussed further in the proposed catalytic mechanism. The other effect of the Sγ may well be directly upon the substrate.

The action of MGL is a predominantly Cγ-S and to a lesser extent Cβ-S lyase. Homocysteine is the preferred substrate for the enzyme with a published activity almost 40 times more than that for methionine (McKie *et al.*, 1998). The structure of homocysteine is identical to methionine but lacks the terminal methyl group of the amino acid side chain. The much higher activity towards homocysteine relative to methionine may be due to the "shielding" effect this methyl group has upon Cys113 Sγ. The much lower activity towards cysteine can be postulated to reflect the greater distance between Cys113 Sγ and the substrate Sγ. That is, the nucleophilic attack by Tyr111 is not aided by the action of Cys113 Sγ to anywhere near the same degree. This idea is further supported by the substrate preferences of the $MGL1^{C113G}$ mutant. It shows a greater relative loss of activity

for αγ-elimination reactions, i.e. 9.3% remaining activity toward homocysteine and 39% activity toward cysteine compared to wild type enzyme (McKie *et al.*, 1998).

Another possible role for the residue is suggested by the action of a related enzyme. There is a cysteine analogous to Cys113 in the active site of cysteine desulphurase, the enzyme encoded by the *nif*S gene of A. vinelandii (Zheng *et al.*, 1993), and its homologues. As mentioned in section 1.8.4 these enzymes belong to the α-family of PLP-dependent enzymes and catalyse the removal of sulphur from cysteine by mechanisms different from that proposed here. However the analogous cysteine residue has been shown to form an internal enzyme bound persulphide complex after cleavage of the substrate SH group (Zheng *et al.*, 1993). The sulphur produced is then ideally placed to be donated for Fe-S cluster formation in the case of NifS. It is the generation of a persulphide group that is key here and the controlled delivery of sulphur. A comparison of MGL1 with the structures of cysteine desulphurase enzymes and homologues follows in this chapter.

3.2.10 Predicted Catalytic Mechanism for MGL

Although MGL is classified in the same family of enzymes as CBL and CGS and possesses a same overall fold, it has considerably different substrate preferences. It catalyses the breakdown of homocysteine and methionine at a much more rapid rate than CBL or CGL but has no activity towards cystathionine. This is surprising given its published activity toward S-adenosyl L-methionine (Lockwood and Coombs, 1991b), a much larger substrate, suggesting that this is not entirely due to steric restrictions on substrate binding. It is the proposal of this

study that MGL has a different catalytic mechanism (Figure 3.18) from that proposed for CBL and CGS (Clausen *et al.*, 1996, 1997, 1998; Steegborn *et al.*, 1999).

In PLP-dependent enzymes, a lysine residue (Lys209 in MGL1) is "essential" for activity. However, it has been observed that mutated PLP-dependent enzymes lacking the analogous lysine do retain trace levels of activity (Watabe *et al.*, 1999). This residual activity is at very low levels though and for most PLP-dependent enzymes removal of this residue produces an inactive enzyme. In agreement with the published mechanisms for CBL and CGS it is proposed that in MGL1 Lys209 catalyses deprotonation of the substrate Cα and generation of the quinonoid intermediate. The second part of the mechanism involves nucleophilic attack at the β or γ atom of the substrate but in MGL1 Lys209 does not appear to be optimally positioned to perform this function. There are few other residues in the active site as suitable alternatives except for Tyr111. The phenol side chain of this residue seems to be in precisely the correct orientation for interaction with the substrate at this point.

Tyr111 has been suggested to exist as a phenolate ion in CGS (Clausen *et al.*, 1998) and to have a role in transaldimination (Steegborn *et al.*, 1999), i.e. the conversion between enzyme-cofactor and substrate-cofactor complexes. The residue is indeed likely to be in a deprotonated form under the influence of Cys113 and the combined action of Arg58* and the phosphate group of the cofactor. If this is accepted as the case upon binding of the substrate then the side chain must be protonated prior to cleavage of the substrate γ-δ bond. This is

because the enamine complex generated is highly reactive and would covalently bond to the Tyr111 in an analogous fashion to the action of L-propargylglycine. Acting as the active site nucleophile upon the substrate removes this problem and the hydroxyl group seems optimally positioned for this. Addressing the proposal that this residue is instead involved in transaldimination, this function could be carried out instead by the cofactor itself. It is likely that the O3' of the cofactor is deprotonated and this could mediate a proton transfer between the amine group of the incoming substrate and the Nε of Lys209 and generation of the external aldimine (substrate-cofactor) complex.

Thus we are able to postulate an enzyme mechanism based upon the structure of MGL1 and the known biochemical evidence (figure 3.18):

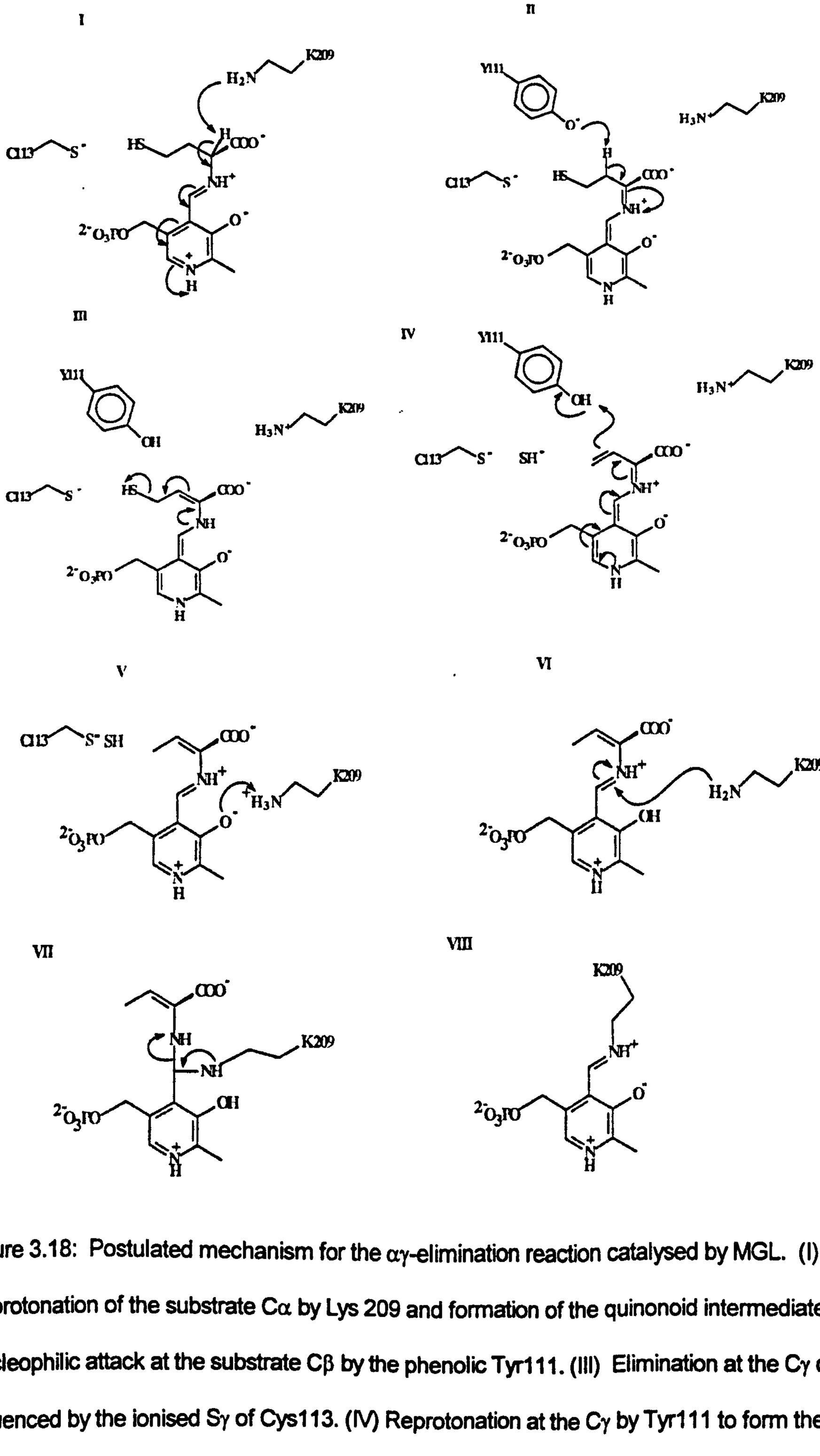

Figure 3.18: Postulated mechanism for the αγ-elimination reaction catalysed by MGL. (I) Deprotonation of the substrate Cα by Lys 209 and formation of the quinonoid intermediate. (II) Nucleophilic attack at the substrate Cβ by the phenolic Tyr111. (III) Elimination at the Cγ of SH influenced by the ionised Sγ of Cys113. (IV) Reprotonation at the Cγ by Tyr111 to form the enamine intermediate. (V) Deprotonation of Lys209 by the cofactor O3. (VI-VIII) Regeneration of the internal aldimine and displacement of the enamine which via hydrolysis generates α-ketobutyrate and

ammonia. This production of an α-keto acid is not catalysed by the enzyme and is presumed to occur in free solution away from the active site.

3.2.11 Comparison of other PLP-dependent enzyme structures

It is beyond the scope of this study to perform a comprehensive comparison of MGL with published structures of PLP-dependent enzymes. However where specific points on the structure of MGL1 have been raised and other structures show analogous situations these are mentioned here.

The structure of CBL from *E. coli* is similar enough that it provides a molecular replacement solution and this reflects the overall secondary structure arrangement. CGS is also similar and yet all three catalyse different reactions reflecting their different active site arrangements. It is apparent as well that the opening to the active site is different to MGL for both of these enzymes. The size of cystathionine as a substrate or product suggests this but the ability of MGL to catabolise SAM does not fit perfectly with this argument.

Despite the close similarity in substrate for MGL and the various "cysteine desulphurase" structures published there is little homology seemingly beyond that normally seen for these types of PLP-dependent enzymes. As previously mentioned there are now three published structures of enzymes able to catalyse the removal of sulphur from L-cysteine and a related enzyme whose preferred substrate is L-cystine.

The structure of C-DES from *Synechocystis* is perhaps unsurprisingly the least informative (Clausen *et al.*, 2000). It does not possess the active site cysteine so strongly implicated in the action of cysteine desulphurase and has a binding site specific for cystine, a situation more analogous to CBL.

The published structure of a NifS-like protein from *T. maritima* (NifS) is similar in structure to MGL, being a member of the α-family of PLP-dependent enzymes (Kaiser *et al.*, 2000). It shows more similarity overall to the structure of phosphoserine aminotransferase (Hester *et al.*, 1999) than CBL. However, the active site appears to be more analogous to the situation found in CBL. Thus His99 in NifS is in an analogous position to Tyr111 in MGL. The predicted active site cysteine residue in NifS (Cys324) is not visible in the structure, being located on a presumably mobile loop. It is impossible therefore to compare the positions of this and Cys113 in MGL. The mechanism predicted for NifS gives Cys324 a direct role in a nucleophilic attack upon the substrate. This is plainly not the case in MGL as already mentioned with the distance being too great. Further evidence in support of this mechanism in NifS is provided by an inhibitor complex structure utilising allylglycine (AG). The inhibition of NifS by AG is thought to be by alkylation of the active site cysteine and in the inhibitor complex neither Cys324 nor AG are visible. This implies that AG is bound to Cys324 and therefore that Cys324 must move close to the substrate at some point. Allylglycine is not an irreversible inhibitor of MGL though, again suggesting a different mechanism of action. It must also be remembered that unlike MGL the products of the reaction with NifS are alanine and sulphide, the formation of α-keto acid and ammonia do not occur.

The structure of CsdB from *E. coli* (Fujii *et al.*, 2000) is similar to NifS from *T. maritima* and therefore to a degree also to MGL. The active site arrangement is much more like that already mentioned for NifS. In CsdB, His123 mimics Tyr111 in MGL in much the same way as His99 in NifS from *T. maritima*. However the putative active site cysteine (Cys364) is visible in CsdB. The lower resolution of the structure (2.8 Å) prevents accurate measurements of interatomic distances but the Cys364 Sγ to hypothetical substrate Seγ is around 5 Å, clearly too far to make a direct interaction. However the authors still speculate on the ability of some residues to deprotonate Cys364 enabling it to perform a nucleophilic attack upon the substrate. Consideration though is given to an alternative method of action in which the cysteine is not directly involved in catalysis. This involves a deprotonation of the selenohydryl group and subsequent elimination in a fashion similar to that proposed here for MGL. The residue predicted though is not His123 (analogous to Tyr111 in MGL1) but His55. Additionally there has been site-directed mutagenesis performed upon Cys364 (mutating it to an alanine) which has confirmed a role for the residue in cysteine but not selenocysteine degradation (Fujii *et al.*, 2000). Although no activity data has been shown this seems to agree with the loss of activity toward homocysteine shown in MGL (McKie *et al.*,1998).

The crystal structure of cystalysin from *T. denticola* (Krupka *et al.*, 2000) is perhaps the most promising in terms of hoping to learn something that may help in the analysis of the mechanism in MGL. It catalyses an α, β-elimination reaction, has a similar overall structure and an analogous residue to Tyr111 in MGL (Tyr123 in cystalysin). Comparing the active sites though reveals no analogous

residue to Cys113 nor Arg58, both of which are implicated in the action of MGL. A very similar mechanism to that of CBL and MalY is predicted in which Lys238 (analogous to Lys209 in MGL) first deprotonates the Cα and then protonates the Sγ of the substrate to cause γ-elimination.

3.3 DISCUSSION

The results of initial crystallisation screens of MGL1 were promising and indicated that it was likely that suitable crystals for structure determination could be grown. Ultimately however these results proved to be unhelpful in the search for optimal growth conditions. The results of DLS although indicating a polydisperse solution of enzyme in the selected buffer also proved to be misleading in finally growing large crystals. The most significant factors involved in growing strongly diffracting, large crystals was the preparation of enzyme and concentration of $AmmSO_4$ used as precipitate. The use of freshly purified, unfrozen, recombinant enzyme produced crystals but this ability was lost if the enzyme sample was combined with glycerol and stored at -20 °C. Although no apparent loss of activity is detectable there is undoubtedly a detrimental change, a feature also of purified glutathione S-transferases in this laboratory (R Thom, personal communication). By varying the $AmmSO_4$ concentration by 100 mM also abolished crystal formation, 2.4 M produced an abundance of much smaller crystals and 2.2 M failed to produce any crystals at all. The use of alternative salts to Li_2SO_4 in crystallisation was not investigated.

After apparent optimisation of the purification to prevent even the minor contamination that previously existed it was highly surprising to discover a second apparent MGL1 species. A native polyacrylamide gel stained for homocysteine desulphurase activity showed two bands which indicated that there were indeed two MGLs. The fact that both proteins co-eluted from the column with imidazole also implied that they were both histidine tagged MGL. Contamination of the expression culture was ruled out by picking a single colony from an agar plate for use in each expression and purification. Proteolysis was unlikely because of the presence of only two species at apparently equal concentrations and C-terminal degradation could be ruled out because of both proteins ability to still bind strongly to the Ni-NTA column. What seemed most likely was the production of two forms during expression. This seemed entirely plausible with the presence of a methionine residue at position 6. Why lowering the concentration of IPTG added to the cultures should reduce the amount of this truncated MGL1 is not clear. It may be that a consequence of expressing this eukaryotic gene in *E. coli* is the recognition of the 2nd methionine as a transcriptional start codon as readily as the first. The cells used for expression of MGL1 contain an additional plasmid, pREP4, which encodes additional copies of the *lac* repressor protein. It may be that high concentrations of IPTG are enough to fully dissociate all repressor proteins. Whereas lower concentrations may result in only partial derepression and, for an unknown reason, produce predominantly the full length transcript.

Although small crystals were grown of MGL2 and linear screens performed around these conditions, with and without the use of propargylglycine, no substantial improvements could be made. The production of an alternative

expression construct was also unsuccessful in improving upon these results. At this point the success of growing large well diffracting crystals of MGL1 and the work involved in solving the structure lead to the abandonment of any further attempts to crystallise MGL2.

The failure to crystallise MGL2 was mediated by the success with MGL1 but a good deal of useful information could have been gained by comparing the two structures. Specifically, in examining the fine detail of the active site and correlating this with the small but measurable differences in activity. Given more time (and less success in crystallising MGL1) further attempts could have been made. The process of crystallisation is poorly understood at least in terms of being able to predict the correct conditions required or the reasons for success with any in particular. Searching for the appropriate recipe in a largely empirical manner therefore allows for an almost inexhaustible number of variations. For MGL2 the next logical progression could have been temperature variations, a different set of sparse-matrix screens (there are always more to try) and variations in protein concentration and drop size. However, using the small crystals already obtained to seed other drops would probably be the first course to follow. The conclusion is clear though, although it is impossible to predicted how to, or even whether a protein will crystallise, there remains many more options to try if required.

The results of refinement confirm the observation that the crystals of inhibitor complex MGL1 diffracted better than the native enzyme. It was apparent from comparing the diffraction patterns of both datasets that the native enzyme crystals

were more anisotropic with the familiar "rugby ball" pattern of reflections much more obvious. The maps for the native enzyme were always inferior, a reflection of the lower redundancy in the data despite being apparently more complete in the highest resolution shell. It is possible that the better data for the inhibitor complex is a result of the enzyme being held in a fixed conformation. The variation in B-factors for the two structures gives further support to this theory.

The structure of MGL confirms its relationship to CBL and CGS in the postulated γ -family of PLP-dependent enzymes (Alexander *et al.*, 1994). The secondary structure elements in the two largest domains are very similar and the N-terminal domain in the three enzymes plays a comparable role in subunit interactions. Despite this, although residues Tyr56 and Arg58 are conserved in these enzymes the short N-terminal region shows considerable sequence variability.

The attraction of finding a change in the structure resulting from disulphide formation is the generation of a negative feedback scenario. It is possible to imagine a situation in which the lack of products from the enzyme mechanism leads to inactive enzyme (via the disulphide bridge). This is because a possible role for MGL *in vivo* may be in the production of sulphide for cysteine formation by direct sulphydration or as an alternative to the reverse transulphuration pathway in converting methionine and homocysteine to cysteine. Thus inactivation of MGL leads to lower cellular cysteine levels because of insufficient sulphide which in turn leads to increased transcription and translation of the *mgl* genes. The possible roles that MGL may play *in vivo* will be discussed in the final chapter of this thesis.

It has been reported that there is an asymmetry of the B-factors in residues in the entrance to the active sites of CGS (Clausen *et al.*, 1998). This is said to show evidence of non-equivalency in the active sites and supports the evidence from other studies on mechanism-based inactivators (Abeles and Walsh, 1973; Silverman and Abeles, 1977; Johnston *et al.*, 1979). The structure of MGL1 does not exhibit this variation in B-factors but there are subtle differences between monomers in the catalytic dimer. Although small the differences in arrangements of Tyr111, Cys113, Lys238 and Asp239 in the holoenzyme structure point to the possible existence of non-equivalent active sites.

The covalent bond between Tyr111 and propargylglycine confirms the previous biochemical evidence, that inhibition was caused by binding to an active site cysteine or tyrosine (Johnston *et al.*, 1979). It may have been thought however that Cys113 was the probable target in MGL1 based on the mutagenesis evidence. Even without the inhibitor complex structure it would be easily seen now that this cannot be the case and the positioning of Cys113 precludes any direct interaction with either mechanism-based inhibitor or substrate.

The main disagreement between the mechanism postulated for MGL and that for CBL, CGS and cystalysin is the action of Lys209. In MGL1 we believe that the residue is positioned poorly for interaction at the substrate β or γ position and that Tyr111 is more likely to fulfill this role. Cys113 has been confirmed as playing no direct role in catalysis. Both Tyr111 and Lys209 are conserved in MGL, CBL and CGS sequences and are likely to have essential roles. The next chapter in this

thesis details the investigation into the catalytic mechanism of MGL and attempts to define key residues and their roles.

CHAPTER 4. SITE-DIRECTED MUTAGENESIS OF MGL1 AND SEQUENCE ANALYSIS OF MGL ENZYMES

4.1 INTRODUCTION

The structure of MGL1 presented in chapter 3 allows us to postulate roles for individual residues both in the physical structure of the enzyme and also in catalysis. Taking into account this, in conjunction with published data on PLP-dependent enzymes in general and MGL specifically, facilitated the proposal of a catalytic mechanism for MGL. In this, we have putatively identified residues responsible for the substrate specificity of the enzyme and involved directly in catalysis. To test these hypotheses, site-directed mutagenesis (SDM) was carried out to obtain modified proteins for analysis. The results are presented in this chapter. In addition, the primary amino acid sequences of MGLs were compared to provide more information on apparent key residues and aid in the identification of the enzyme in various organisms.

Site-directed mutagenesis had been used previously on *T. vaginalis* MGL to identify a role in catalysis for Cys113 (Cys116 in MGL2) (McKie *et al.*, 1998). This showed that the residue was important for activity, but not essential. Chemical modification and random mutagenesis of the gene were techniques that could have been used to extend this study, but it was decided that a targeted approach was more appropriate. Mutagenesis is a powerful tool for determining the role of functional groups in proteins and has been extensively used in a variety of PLP-dependent enzymes. For example, it has been used to show the essential role of

the active site lysine residue, which forms the Schiff base with the cofactor and has a role in catalysis (Rege *et al.*, 1996; Hunter and Ferreira, 1999). There are also a large number of mutants for which the structures have been solved and are now deposited in the protein structure database. The structures have highlighted a consideration when mutating residues within a protein, what effect may the change have upon the overall structure? It would therefore be ideal to accompany enzymatic data on mutant enzymes with structural data. Attempts to obtain the structures of mutants produced in this study are also detailed here.

4.1.1 Site-Directed Mutagenesis of MGL1

The decision on which residues to perform mutagenesis upon was based predominantly on the structure of MGL1 presented in chapter 3. It had been envisaged that if the structural studies had been delayed that these decisions would have been based upon sequence analysis and comparisons with other related enzymes. This would have been less desirable not only because of the difficulty in identifying key targets but also because of the difficulty in interpreting results based solely upon *in vitro* assay data. The proposed roles of the residues selected were discussed in chapter 3, but mutations planned are listed and explained here as they form a central part of this chapter.

4.1.2 Residues Selected for Mutagenesis

The residues selected for mutagenesis and the changes introduced are listed in Table 4.1:

Original Residue	Proposed Role	Mutant
Arg58	Cofactor binding and catalysis	Met & Lys
Leu59	Substrate access	Phe
Tyr111	Cofactor binding and catalysis	Phe
Cys113	Indirectly in catalysis	Pro & Ser
Lys209	Cofactor binding and catalysis	Ser
Lys238	Catalysis	Ala

Table 4.1 Residues selected for site-directed mutagenesis and their predicted roles.

4.1.2.1 Arginine 58

A primary function of Arg58 appears to be the coordination of the PLP cofactor in the active site. Together with Tyr56 it belongs to the opposite chain of the catalytic dimer when compared with the other residues that directly interact with the cofactor. As can be seen in figure 3.15 in the previous chapter, Arg58 is within hydrogen bond distance of OP2 and OP3 of the cofactor and also 3.1 Å from the side chain hydroxyl of Tyr111 in the holoenzyme and inhibitor complex. With the proposed role in catalysis of Tyr111 reliant upon its deprotonation, Arg58 seemed well positioned to modify the pKa of Tyr111 together with a role in positioning the cofactor. It was envisaged that substituting methionine would retain the architecture of the active site because of its similar size, but would remove the charge interactions. Lysine was also selected as a means of partially removing the positive charge. It was hoped that this would bind the cofactor adequately but be less effective in deprotonating the Tyr111 hydroxyl group.

4.1.2.2 Leucine 59

This residue is not implicated in the catalytic mechanism at all but it is in the entrance to the active site. It was envisaged that the substrate preference of the enzyme could be altered by manipulating the size of this entrance. By substituting a bulkier residue, the enzyme activity toward larger compounds such as S-adenosyl-L-methionine, which had previously been reported as substrate (Lockwood and Coombs, 1991b), may be reduced. The decision was taken to mutate this residue to a phenylalanine, a substitution that would place the large phenyl ring of the side chain directly into the "channel". The prediction was that the smaller molecules, e.g. L-cysteine, would be the best substrates.

4.1.2.3 Tyrosine 111

The catalytic mechanism for MGL proposed in this study predicts a critical role for this residue. After a proton abstraction by Lys209 at the substrate Cα, the β or γ-elimination is catalysed by the nucleophilic attack from Tyr111. The phenol ring of Tyr111 also coordinates the pyridine ring of the cofactor, holding it approximately parallel by van der Waals interactions. This interaction by an aromatic group, alternatively substituted for by phenylalanine or histidine in other PLP-dependent enzymes, increases the electron sink character of the cofactor (Hayashi *et al.*, 1990). The inhibitor complex structure also shows that the hydroxyl group of this residue is unequivocally targeted by propargylglycine in its action as a suicide inhibitor. Changing this residue to any other would be expected to greatly curtail catalytic activity, or remove it completely, and also prevent covalent attachment of propargylglycine and its action as a suicide inhibitor. The most reasonable change would be to histidine or phenylalanine,

thereby removing the active hydroxyl group but retaining the aromatic ring. Another possible change would be to alanine and then subsequently attempting to rescue any lost activity with phenol. Phenylalanine was selected for this study as the residue most likely to retain the structure of the enzyme whilst removing the active group and so allowing the testing of our hypothesis.

4.1.2.4 Cysteine 113

The possible role of Cys113 in the enzyme mechanism is one of the most interesting features. Previous mutagenesis and chemical modifications have shown that the thiol group is important and its loss has a detrimental affect upon the catalytic activity (Nakayama *et al.*, 1988a; McKie *et al.*, 1998). It was something of a surprise therefore to find in the MGL1 structure that the residue was too distant from the presumed position of the substrate to be directly involved in catalysis. Neither does the cysteine appear to coordinate the cofactor. There are several possible reasons for the drop in activity seen upon modification of this residue:

1/ Chemical modification may partially block access to the catalytic centre of the active site;

2/ A cysteine to glycine change would allow Tyr111 (only two residues away) to move a significant distance from the optimal position for its proposed role in a nucleophilic attack upon the substrate;

3/ A non-catalytic interaction between the thiol group of Cys113 and the substrate is lost.

Through mutagenesis it was intended to substitute a serine which would maintain all steric conditions (with the exception of the Sγ being larger than the corresponding Oγ) but remove the thiol group. It was also chosen to mutate Cys113 to a proline. This was selected to mimic the active site structure of CBL.

4.1.2.5 Lysine 209

An essential lysine residue analogous to Lys209 in MGL is present in all PLP-dependent enzymes studied so far. It is via the side chain amine group that the Schiff base linkage is formed between cofactor and enzyme (the so called "internal aldimine" (Christen and Metzler, 1985)). The binding of substrate, release of products and regeneration of the starting complex is catalysed by this residue. Where substrates are amino acids, it catalyses a deprotonation at the Cα. It was envisaged that a mutant MGL1 lacking this residue would be inactive and this would confirm its essential role in catalysis. More importantly, perhaps it would then allow enzyme substrate-complexes to be produced and their structures determined. This would allow examination of the binding of different substrates and identify reasons for differences in specific activities. The substitution chosen was serine, although it could have equally been another non-charged, non-aromatic residue such as alanine or methionine.

4.1.2.6 Lysine 238

The reason for mutating this amino acid was related to the desire to discover the role of Cys113 in the enzyme mechanism. The method of catalysis postulated from the crystal structures implies that the Sγ of Cys113 is ionised and Lys238* in

combination with Asp239* may be required to produce this. Thus the hope was that mutating Lys238 may prevent this ionisation and so the activity would fall to a level reflecting the contribution of the cysteine alone. This possibly being similar to that measured in the MGL1$^{(C113G)}$ mutant. Lys238 is also located in the channel leading to the cofactor, but makes no direct interactions with the adjacent monomer. Thus any change would not be predicted to interfere with subunit association as water would fill any spaces left by the missing side chain. To fulfill the aims for this mutation, alanine was chosen as the alternative residue. It is small, non-polar and would allow water to enter into any empty space and so be unlikely to distort the structure.

4.2 RESULTS

4.2.1 Site-Directed Mutagenesis of MGL1

This section of this study details the results of construction, production and analysis of mutant recombinant MGL1 enzymes.

SDM was carried out using the Quickchange™ kit (Stratagene) as detailed in section 2.2.15. An overview of the technique is shown in figure 4.1. The design of complementary mismatched oligonucleotide primers followed the guidelines in the manufacturers protocol and are listed in section 2.2.12.

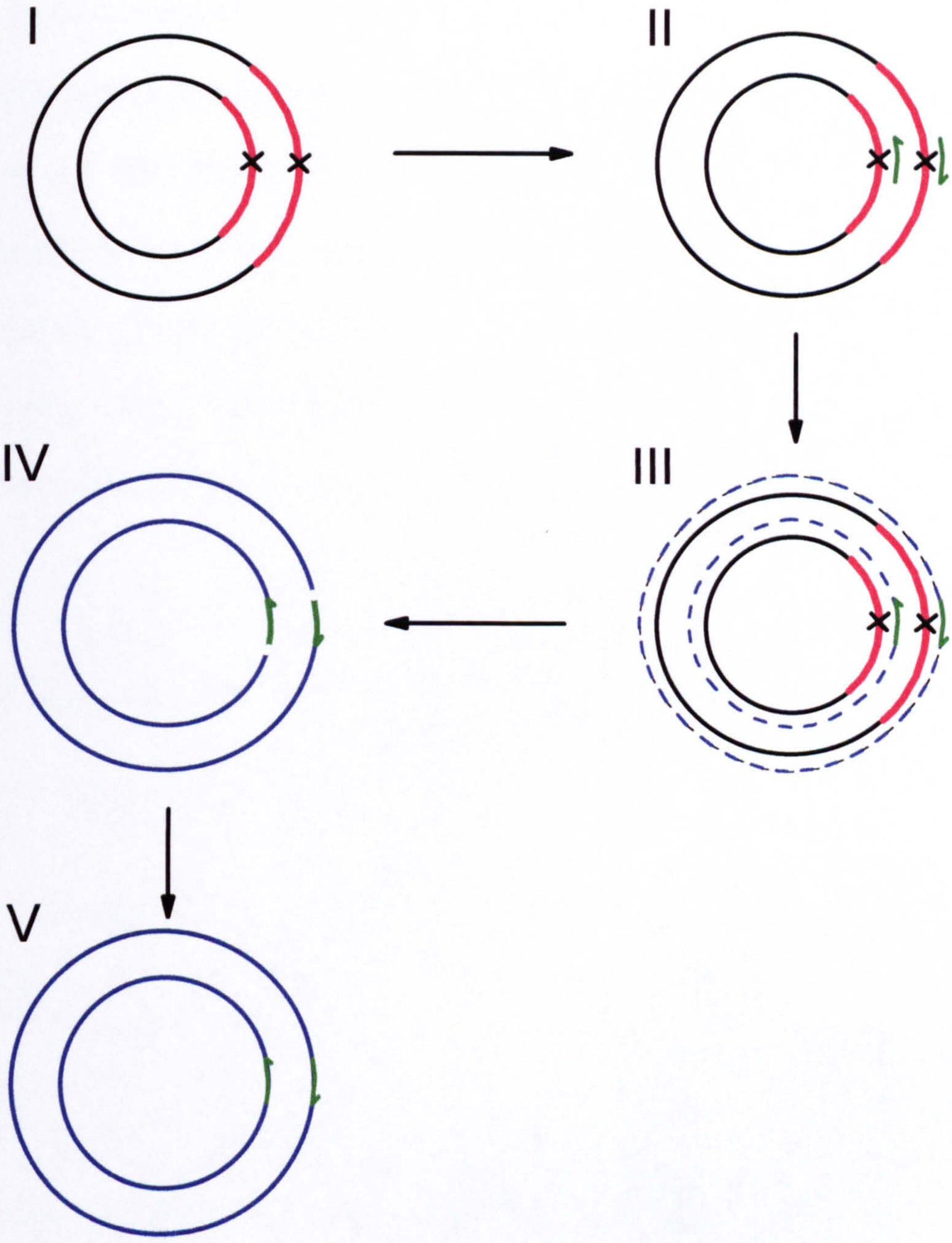

Figure 4.1: Overview of site-directed mutagenesis. I - Isolate plasmid with gene target for mutagenesis; II - Denature plasmid and anneal complementary oligonucleotide primers containing the desired mutation; III - Extend both DNA strands using thermostable proof-reading DNA polymerase, incorporating mutagenic oligonucleotide primer. Repeat steps II and III for 20-25 cycles. IV - Incubate amplification products with *Dpn*I restriction enzyme to cleave parental methylated DNA template. Leaves only mutated, circular nicked plasmid; V - Transform mutated plasmid into *E. coli* where cellular machinery repairs nicks and produces intact circular plasmid.

Transformation of competent *E. coli* with the products from the mutagenesis reactions gave a varied number of colonies, from less than ten to more than

several hundred. In the first instance three colonies from each plate were picked and grown up overnight for mini-prep plasmid purification. Successful production of mutants L59F, R58K and R58M was selected for by restriction digest and then confirmed by DNA sequencing. For all other mutants, DNA sequencing alone was used to identify positive mutations. These three mutants however, resulted in the same change to the restriction map of the plasmid by destroying a *Bsm*AI recognition site. The changes are listed in Table 4.2.

Mutant	Sequence
Parental Plasmid	TACACAC**GTCTC**GGCAAC ATGTGTG**CAGAG**CCGTTG Y T R L G N
R58K	TACACAAAGCTCGGCAAC ATGTGTTTCGAGCCGTTG Y T K L G N
R58M	TACACAATGCTCGGCAAC ATGTGTTACGAGCCGTTG Y T M L G N
L59F	TACACACGTTTCGGCAAC ATGTGTGCAAAGCCGTTG Y T R F G N

Table 4.2: Sites of mutagenesis of MGL1 R58K, R58M and L59F mutants. The *Bsm*AI restriction site is marked in bold text in the parental plasmid template sequence. Changes made to nucleotide and amino acid sequences are highlighted in red in the mutants.

Several mutations were successfully isolated and identified rapidly but, surprisingly, sequencing showed a number of unexpected non-positive mutation results. These were in addition to the expected levels of negative clone background which results from template plasmid being uncut by *Dpn*I and transformed. There were also a number of clones which were mutated but did not encode the specific mutations designed into the oligonucleotide primers. The

strategy used to produce the mutants ensures that the site of mutagenesis and the upstream and downstream flanking sequence is defined by the oligonucleotide primers. Thus it was very surprising to have a number of mutants which encoded changes within these areas not designed into the primers. This problem was most evident in the designed Tyr111 to Phe and Cys113 to Ser mutations. A positive $MGL1^{(Y111F)}$ mutant was isolated only after sequencing twelve transformants. $MGL1^{(C113S)}$ was difficult to synthesise for an unknown reason and resulted in the lowest number of transformants. After repeating the mutagenesis and transformation step and obtaining no positive clones, this mutant was abandoned as there had by then been success in producing seven other mutants.

4.2.2 Expression and Purification of Mutants

Confirmed mutants were transformed into *E. coli* strain M15 pREP4, for expression in an identical fashion to the parental plasmid. Successful transformants were selected by growing on $LB_{AMP/KAN}$ agar plates. Single colonies were used to prepare an overnight starter culture and expression and purification were performed identically to that already detailed for unmutated MGL1. All mutants were produced as soluble enzyme and the recombinant proteins were isolated to a similar purity as the parent MGL1.

4.2.3 Enzyme Assays

This section of the thesis describes the data obtained from the enzymatic analysis of the site-directed mutants produced of MGL1. There are three previously used and published assays for MGL:

1/ a stopped assay to detect α-keto acid production,

2/ a continuous, coupled assay measuring the production of α-keto acid,

3/ a continuous assay measuring the production of hydrogen sulphide.

The results and problems from each of these will be dealt with in turn.

4.2.3.1 Results of Stopped Assay Using Methylbenzoylthiolhydrazone (MBTH) to Determine α-Keto Acid Production

This detection of α-keto acid products is from a published method (Soda, 1968). It relies upon the reaction of MBTH specifically with the α-keto acids to produce an azine which has a measurable absorbance at a wavelength of 320 nm. The *in vitro* assay is thus incubation of enzyme with substrate followed by termination of the reaction and finally reaction of the products with MBTH. Calibration is by a standard curve produced using commercially available α-ketobutyrate.

The previously used conditions were a total reaction volume of 1 ml, 100 mM buffer with varying levels of substrate and enzyme maintained at 37 °C for 5 minutes. The assay was stopped by the addition of 125 µl of 12.5% (w/v) trichloroacetic acid (TCA) and the sample centrifuged to remove precipitates. 1ml of this was then added to 1 ml of 1 M acetate buffer pH 5.0 and 0.4 ml 0.1%

MBTH and heated at 65 °C for 30 minutes before measuring the absorbance at 320 nm. The high concentrations of acetate buffer were used to neutralise the effects of the TCA used in stopping the assay and maintaining an optimum pH for the reaction of MBTH with α-keto acids.

It was observed initially that the results from the assays were quite variable between replicates and also on any given date. These observations also included the apparent detection of activity towards compounds not shown previously to be substrates of MGL1. This was thought to be highly significant at first, until it became apparent that the results were not repeatable and then the accuracy of the assay came under question. These apparent errors and uncertainties were compounded by the low catalytic activities exhibited by the MGL1 mutants.

To obtain more reliable data, the components of the assay were investigated further. The most surprising finding was that the MBTH reagent reacted strongly with the acetate buffer giving high absorbance readings around 305-325 nm. This caused obvious concerns as the reaction was thought to be specific for α-keto acids. Evidently acetate (CH_3COOH) was able to cross react strongly and there was concern that the TCA used to halt the assay and even the carboxylate group of the substrate could also react. Investigations with a number of different acids showed varying degrees of reactivity and whilst the absorbance peaks were not precisely at 320 nm there was a high level of absorbance at this wavelength.

The decision was therefore taken at this point to not use this assay system further but to utilise one of the other assays available to us.

4.2.3.2 Continuous Coupled Assays Detecting α-Keto Acids

An alternative assay system previously used to study MGL has used lactate dehydrogenase (LDH) as a coupling enzyme to monitor the production of pyruvate from L-cysteine and O-acetyl-L-serine (McKie, 1997). The conversion of NADH to NAD^+ by the enzyme as pyruvate is catabolised is measured continuously at 340 nm. Several isoforms of LDH are expressed in cells with differing substrate specificities. The enzymes are not totally specific for pyruvate but also act upon larger α-keto acids including α-ketobutyrate. It was hoped that this could then be used to detect αγ- in addition to αβ-elimination reactions by MGL.

Assays were attempted in an identical manner to that already published for MGL. Unfortunately it was not possible to obtain reliable data for any of the mutants with this assay. In control reactions using pyruvate, the LDH enzyme was shown to be active. The activity of mutant MGL1 was also confirmed using the H_2S trapping assay detailed in the next section of this chapter. However the activity of these enzymes could not be coupled to the action of LDH. Because both enzymes exhibited activity separately it is assumed that the lack of success was due to the very low levels of α-keto acids produced.

The decision was then taken not to pursue the reasons for the failure of this assay but instead to use the H_2S trapping assay.

4.2.3.3 Hydrogen Sulphide Trapping Assays

These assays were performed as previously described for homocysteine desulphurase (Thong and Coombs, 1985b). This continuous assay utilises the reaction of hydrogen sulphide with lead acetate which forms lead sulphide. The lead sulphide forms a brown precipitate which in the timespan of the assay remains in suspension and can be measured at a wavelength of 360 nm. Hydrogen sulphide is one of the products of the $\alpha\gamma$-elimination of homocysteine and $\alpha\beta$-elimination of cysteine. Thus this assay was used for measuring activity of MGL1 mutants on these substrates. It is also used in detecting activity in native PAGE (Thong and Coombs, 1987). Enzyme resolved on a gel can then be visualised by incubating in a solution containing cysteine or homocysteine and lead acetate. The action of the enzyme causes the lead sulphide precipitate to accumulate in a band denoting the position of the enzyme in the gel.

4.2.4 Analysis of Mutated MGL1 Activities

4.2.4.1 MGL1 Arg58 Mutants

The results of mutating Arg58 to lysine or methionine could not be predicted with confidence. The residue is strictly conserved in other γ-family PLP-dependent enzymes but a role in catalysis, although proposed from the structure of MGL1, was not definitely established.

Both mutants (R58K and R58M) were overexpressed successfully in *E. coli* although neither were yellow in solution as seen for MGL1. This presented no

problems in purification though and highly homogenous preparations were produced.

Neither mutant exhibited normal Michaelis-Menten behaviour and, in fact, it was difficult initially to detect any activity at all for either enzyme. Activity was finally only seen at or below certain concentrations of substrate. These values were dependent upon the amount of enzyme being used and indicated likely substrate inhibition. When detectable, both enzymes exhibited high levels of activity at very low levels of substrate. By increasing the substrate concentration whilst maintaining the same amounts of enzyme, the linear rate observed would increase until a substrate concentration was reached when no activity was detectable. With normal substrate inhibition the activity would be expected to decrease more gradually. What was observed however, was an intermediate concentration where a lag phase of several minutes lasted until rapid linear activity was measured until the substrate was exhausted. An example of the differences between the two situations is given below in figure 4.2.

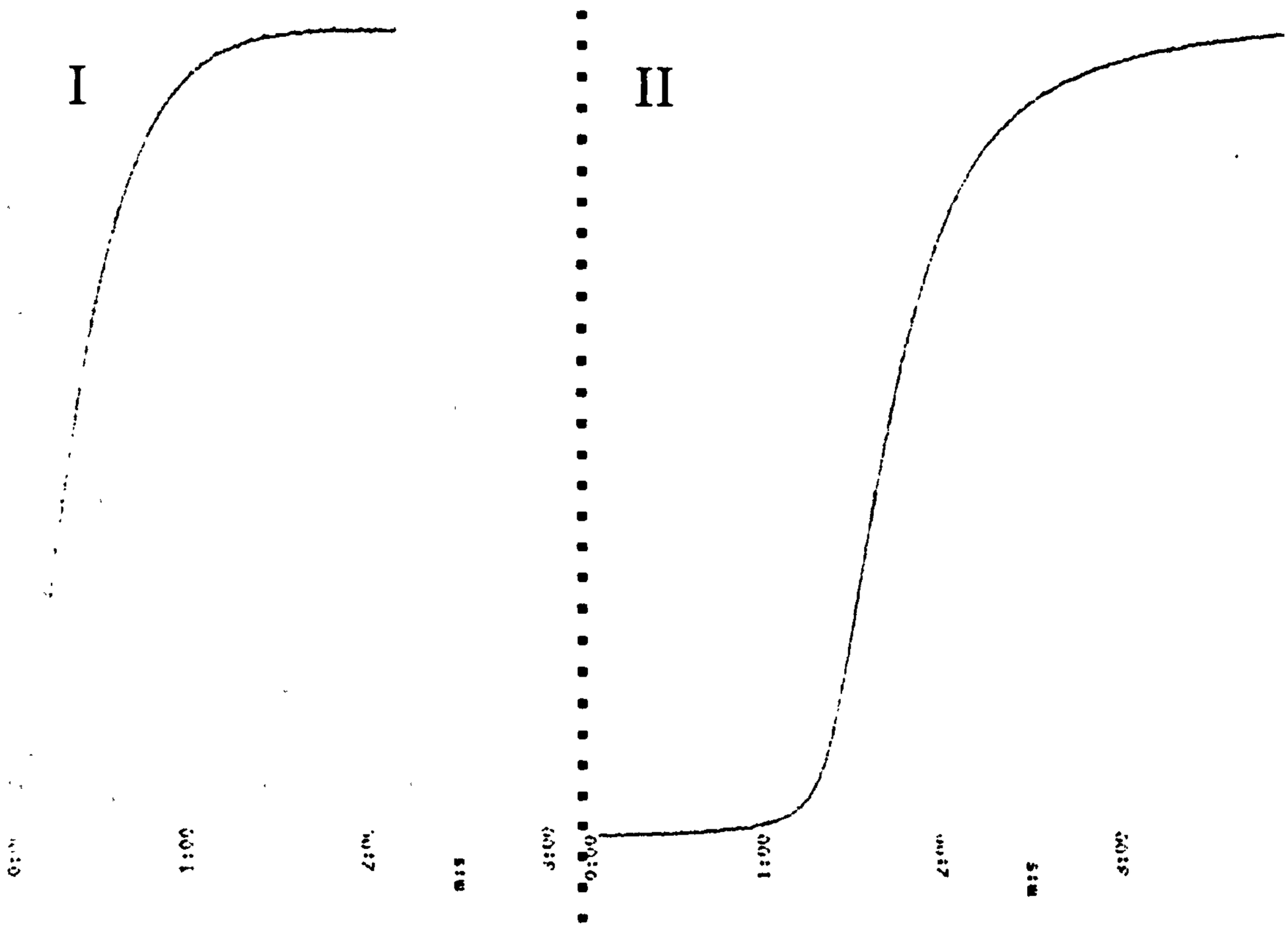

Figure 4.2: Spectrophotometer traces of MGL1$^{(R58K)}$ assays. I - DL-homocysteine concentration of 2.5 mM; II - DL-homocysteine concentration of 5 mM. X-Axis denotes time of reaction in minutes. Y-Axis denotes change in absorbance measured at 360 nm from approximately 0 to 1.0 absorbance units. Both assays gave approximately equal upper limits of absorbance, no scaling has been applied to the outputs.

The narrow "window" of substrate concentrations required for activity made it difficult to make estimates for Km and Vmax values. This was also compounded by the difficulty to accurately measure an initial rate because of the rapid depletion of substrate. Neither enzyme appeared to be exhibiting Michaelis-Menten kinetics and so attempts to define Km and Vmax values were not attempted.

Compounding problems in assaying MGL1$^{(R58K)}$ was the change in pH profile and buffer preferences of the mutant enzyme. Activity was measurable between pH

4.3 and 5.6 with acetate buffer and similar ones such as formate and succinate. However no activity was seen with buffers such as citrate, pthalate or imidazole at these pHs. The concentration of acetate buffers also had a critical effect on activity. At concentrations of 50 mM there was no lag in absorbance increase but at higher concentrations of buffer and the same concentrations of substrate and enzyme there was a distinct lag. This lag lengthened as the buffer concentration increased until no activity was detected at all at concentrations greater than 200 mM. The rates measured after these lags were also lower than those with no lag. It therefore seemed possible that acetate was acting as an activator at low concentrations and an inhibitor at high concentrations. This idea had to be discarded when it was found that equally high activity with no lag phase was measured using the sulphonic acid buffer MES. Although MGL1$^{(R58M)}$ showed a similar substrate inhibition profile as MGL1$^{(R58K)}$, it did not exhibit this strange buffer activation/inhibition and was highly active at pH 6.5 in 100 mM imidazole. Due to the difficulties in interpreting data from these mutants further investigations into the pH profile was not conducted.

Estimations of specific activity for both mutants with DL-homocysteine were made. The assay conditions are probably sub-optimal though and the true values may be higher.

Mutant	Specific Activity (μmol/min/mg protein)
R58K	308 ± 28 (9)
R58M	187 ± 22 (5)

Table 4.3: Estimated specific activities for MGL1$^{(R58K)}$ and $^{(R58M)}$ with DL-homocysteine. Activities given are means ± S.D. with the number of experiments given in parenthesis.

4.2.4.2 MGL1 L59F Mutant

This mutant exhibited the typical expression and purification profile seen for parent MGL1. The recombinant enzyme was coloured yellow, clearly seen binding to the column and after elution, and similar quantities of enzyme were purified (~ 10-15 mg of purified enzyme per litre of culture).

Unlike MGL1$^{(R58K)}$ and $^{(R58M)}$, there was no immediate problem in performing assays for this mutant. Firstly the enzyme was assayed with a range of buffers at varying pH values to find optimal conditions for calculating activities. A broad screen of pH values showed higher activity for the enzyme at pH values above that optimal for MGL1. The enzyme was then assayed in a variety of buffers above pH 7.0, in 0.2 pH unit increments, using sodium phosphate, TRIS and HEPES (data not shown). The results showed an optimum activity in 100 mM sodium phosphate pH 8.0 and so this was used for all further assays with this mutant. This determination of pH and buffer optimums was carried out for all mutants prior to attempting to calculate rate constants and specific activities.

To examine the kinetic properties of the mutant enzyme, a fixed amount of enzyme was assayed with varying concentrations of DL-homocysteine. The results of these assays are shown in figure 4.3.

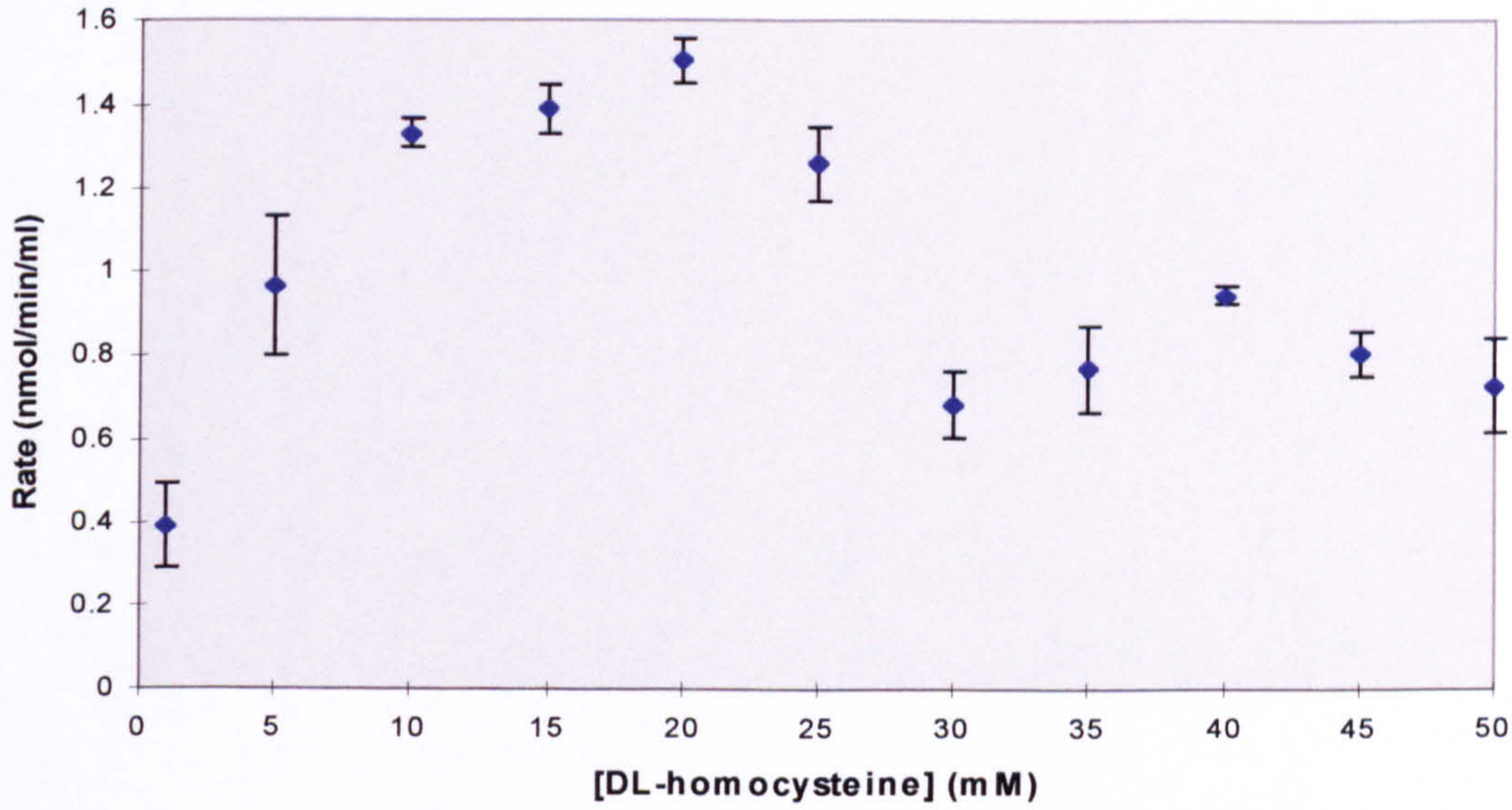

Figure 4.3: MGL1$^{(L59F)}$ reaction rate with DL-homocysteine. Data points are mean values, error bars are S.D. with n=3.

A plot such as this typically indicates substrate inhibition at high concentrations but can also be caused by other conditions such as the assay conditions being altered. The solution was checked to ensure the pH of the assay had not changed before or after the reaction. There was no evidence that this had occurred. Plotting the data obtained at substrate concentrations up to 25 mM shows a clear hyperbolic relationship between reaction rate and substrate concentration (figure 4.4).

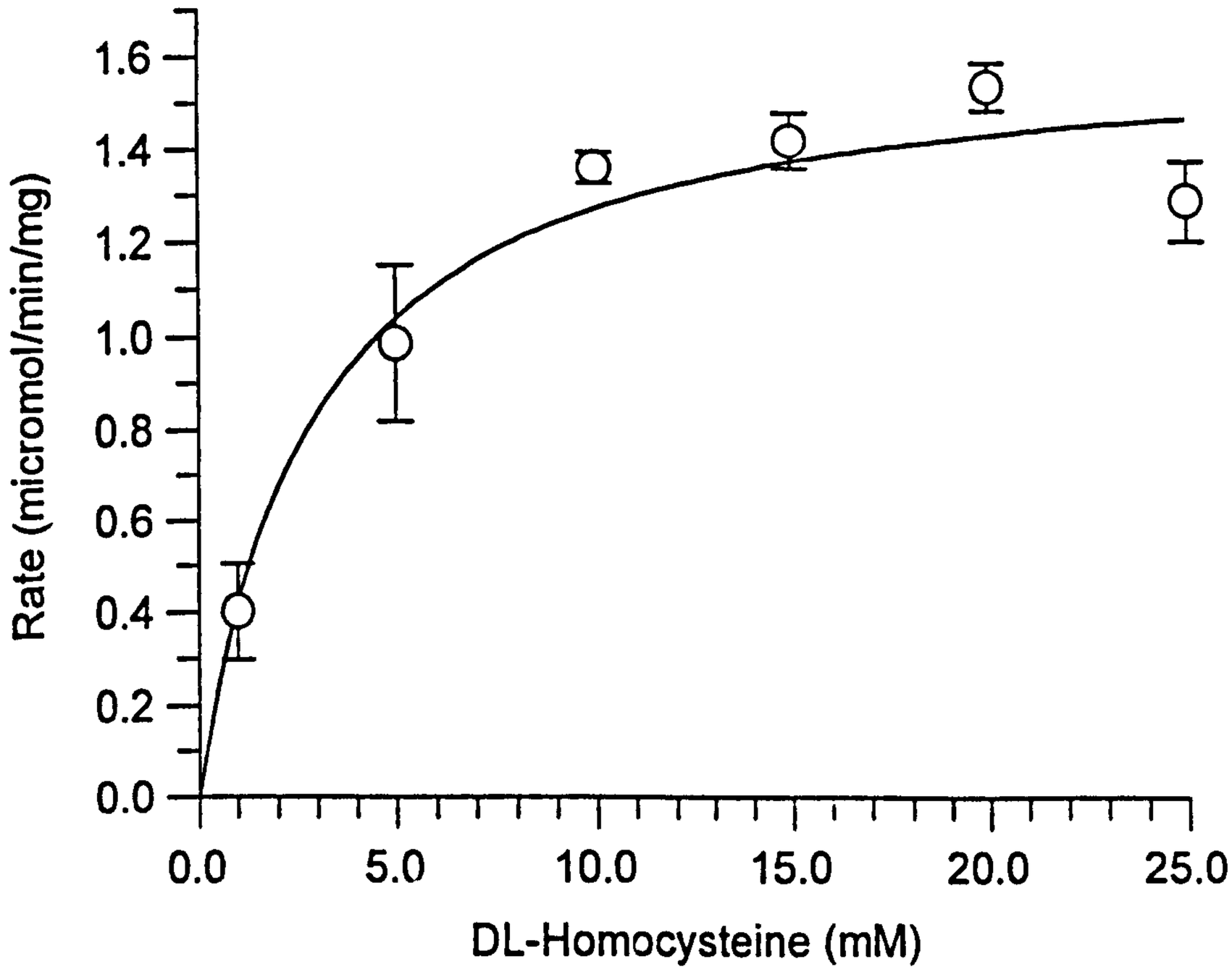

Figure 4.4: Saturation curve of MGL1$^{(L59F)}$ with DL-homocysteine up to 25 mM. Data points are mean values, error bars are S.D. with n=3.

A Hanes plot was then constructed to calculate Km and Vmax. Unfortunately this leaves only a few data points but these are enough to produce what appears to be a good linear plot (figure 4.5).

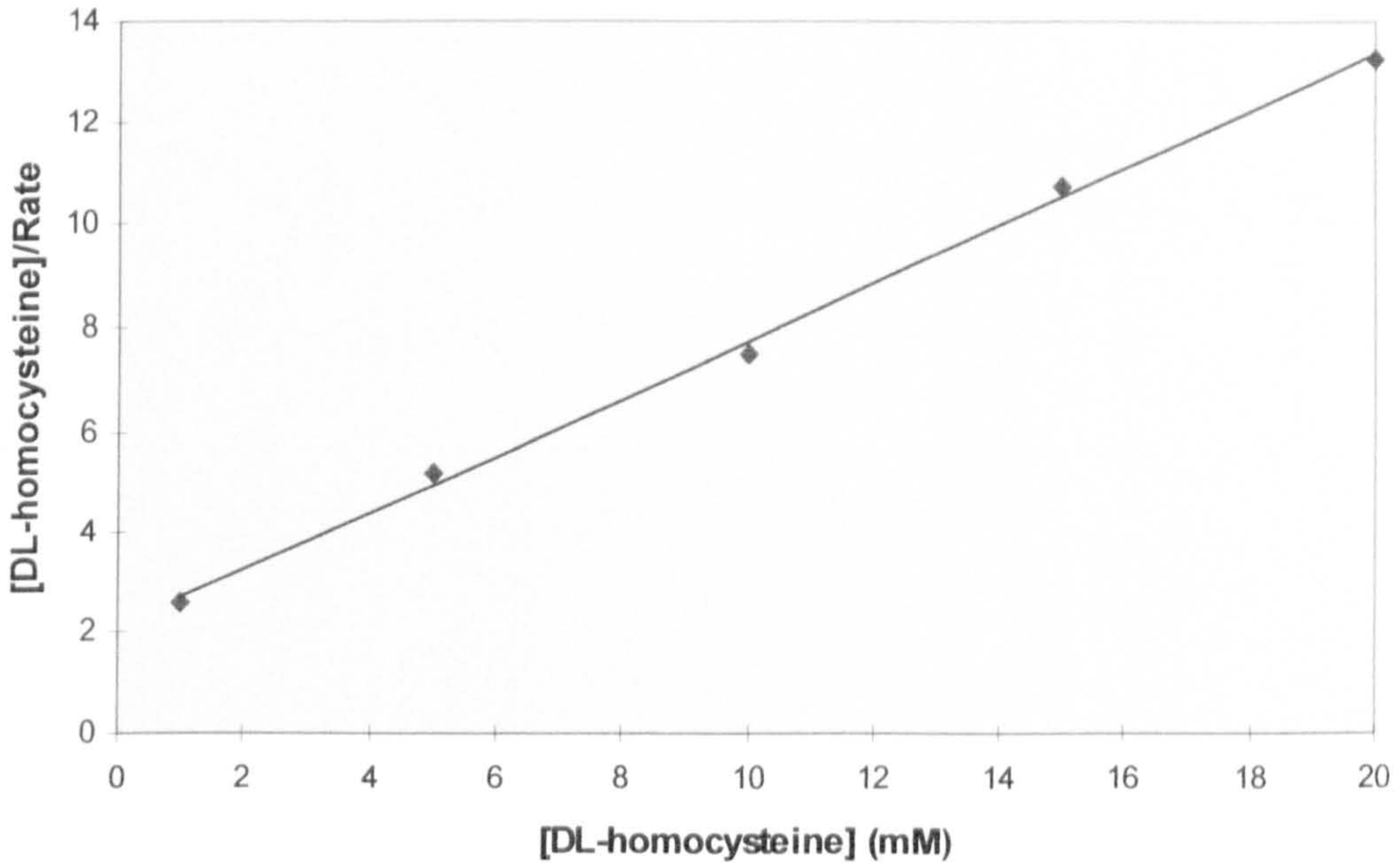

Figure 4.5: Hanes plot of homocysteine catabolism by MGL1$^{(L59F)}$

This Hanes plot produces a plot of [s]/v against [s] which is linear with a slope of 1/Vmax and an *x*-axis intercept of -Km. Knowing these values we can then go on to calculate the specific activity of the enzyme with a good estimate of the Km:

Specific Activity (μmol/min/mg protein)	8.26 (± 1.18)
Km (mM)	3.7
Vmax (μmol/min/mg protein)	2.0
k_{cat} (s^{-1})	1.5
k_{cat}/Km (M^{-1} s^{-1})	400

Table 4.4: Kinetic data for MGL1$^{(L59F)}$ with DL-homocysteine as substrate. Numbers in parenthesis are S.D. with n=5.

An identical process was carried out with L-cysteine as that detailed above for DL-homocysteine, with the exception of determining a pH optimum. A fixed quantity of enzyme was assayed with varying concentrations of L-cysteine. A plot of the results is shown below in figure 4.6.

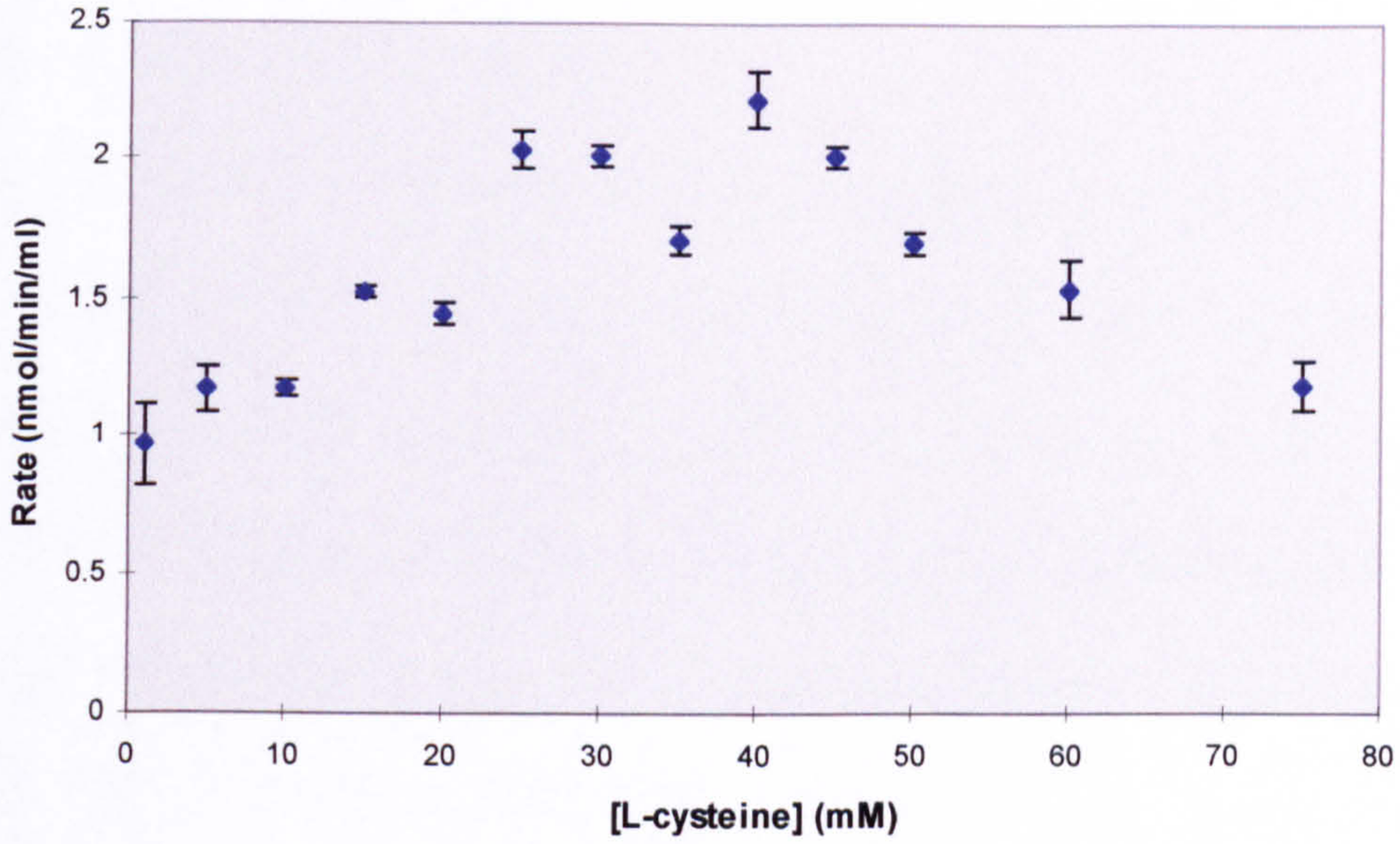

Figure 4.6: MGL1$^{(L59F)}$ reaction rate with L-cysteine. Data points are mean values, error bars are S.D. with n=3.

Unlike the situation seen for DL-homocysteine, with this mutant there is no clear hyperbolic curve even at low substrate concentrations. It seems unlikely that a better curve could have been produced even had the range of concentrations been extended lower. Thus for some unknown reason the action of this enzyme upon L-cysteine *in vitro* does not seem to adhere to the Michaelis-Menten equation. The measured specific activity for the enzyme at a substrate concentration of 40 mM is given below.

Specific Activity (μmol/min/mg protein)	1.75 (± 0.34)

Table 4.5: Kinetic data for MGL1$^{(L59F)}$ with L-cysteine as substrate. Numbers in parenthesis are S.D. with n=7.

4.2.4.3 MGL1 Y111F Mutant

Over-expression of this mutant in *E. coli* produced a colourless recombinant enzyme, purified in an identical manner as parental MGL1. From the mechanism proposed for MGL it was not expected that this mutant would possess any activity, or at best only very low levels. This was apparent immediately because of the large amounts of enzyme required before any activity could be detected. Once it was confirmed that the enzyme was catalytically active, it was assayed to find the pH optimum and buffer preferences. Screening a range of buffers and pHs gave highest activity with 100 mM TRIS at pH 8.8. This was then followed up by trying to calculate kinetic parameters for homocysteine and cysteine.

Assays were carried out as before with varying substrate concentrations over a broad range whilst maintaining the same amount of enzyme. The result of this is shown below in figure 4.7.

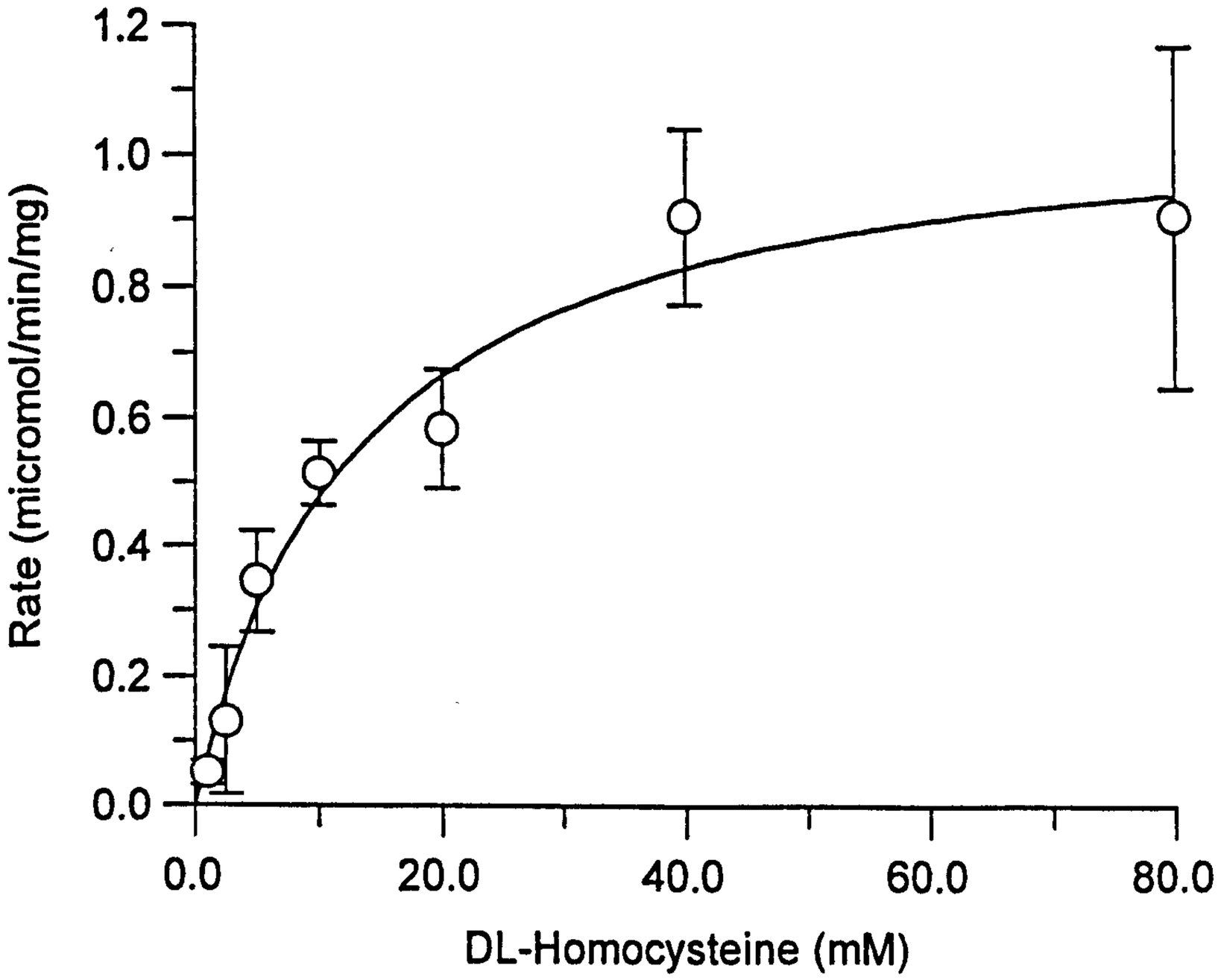

Figure 4.7: MGL1$^{(Y111F)}$ reaction rate with DL-homocysteine. Data points are mean values, error bars are S.D. with n=3.

Saturating concentrations of substrate were reached and Michaelis-Menten conditions apply for this enzyme. The data points on the graph conform to the hyperbolic curve expected and there is no evidence of substrate inhibition at high concentrations. A Hanes plot was then constructed to obtain a linear relationship and define Km and Vmax values (Figure 4.8).

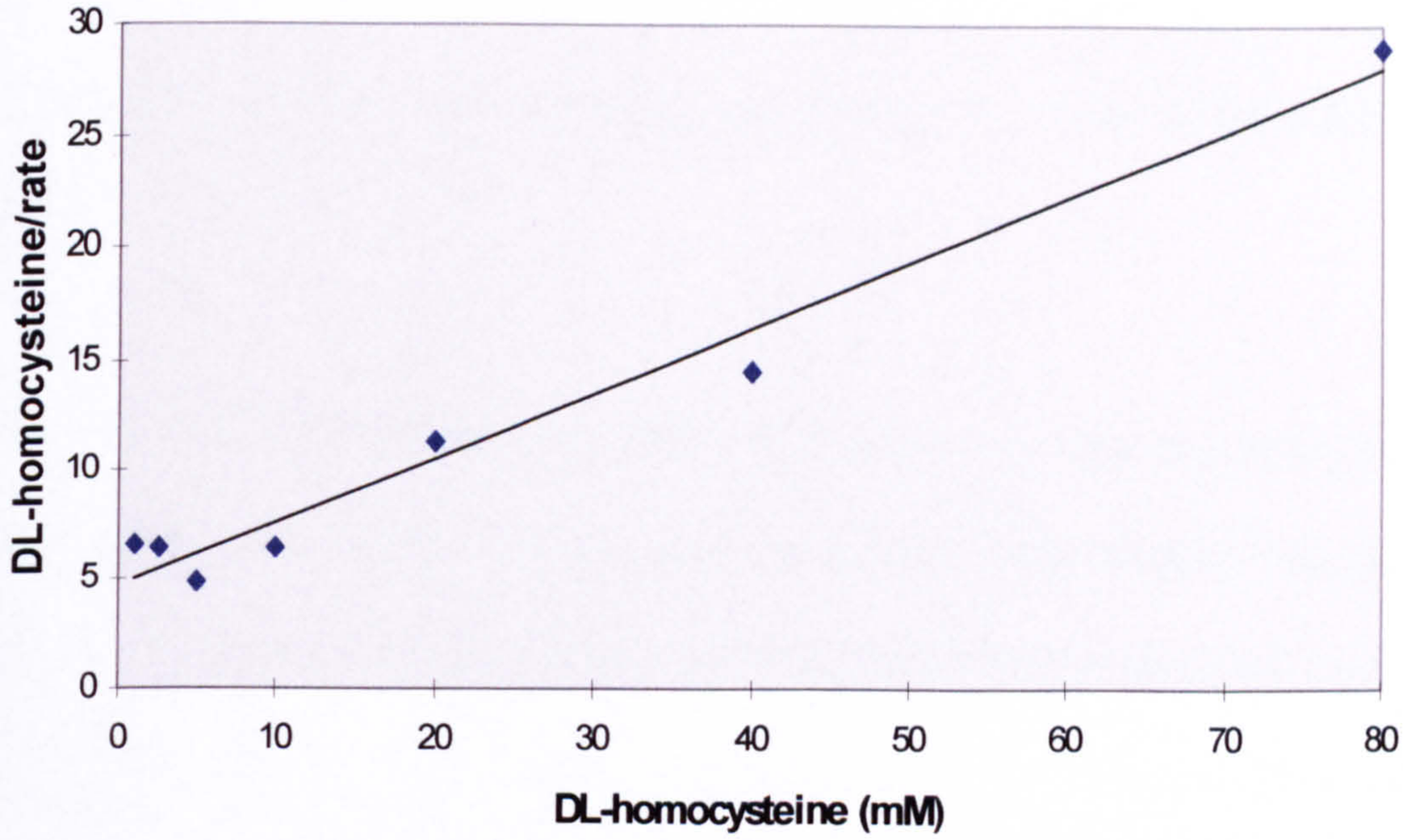

Figure 4.8: Hanes plot of DL-homocysteine catabolism by MGL1$^{(Y111F)}$.

Again this plot shows the non-linearity of the data, largely distorted by the two lowest substrate points. This is only a distortion, however from the perceived ideal and rather than discard experimental data they were included in the calculation of kinetic parameters. If these values were omitted, the Km and Vmax would be lower than the values given below (Table 4.6).

Specific Activity (μmol/min/mg protein)	0.99 (± 0.31)
Km (mM)	15.7
Vmax (μmol/min/mg protein)	1.0
k_{cat} (s^{-1})	0.8
k_{cat}/Km (M^{-1} s^{-1})	50

Table 4.6: Kinetic data for MGL1$^{(Y111F)}$ with DL-homocysteine as substrate. Numbers in parenthesis are S.D. with n=8.

No activity could be measured using L-cysteine as a substrate with this mutant. It is likely that any trace activity remaining is below the limits of detection for this assay.

Another reason for constructing this mutant was to confirm the action of L-propargylglycine. In the inhibitor-complex structure presented in the previous chapter there is continuous electron density between the hydroxyl group of the Tyr111 side chain and the C_γ of the propargylglycine. Removing the hydroxyl group was postulated to abolish activity for the enzyme and prevent suicide inhibition by the inhibitor. To test this, purified enzyme was incubated with inhibitor for 5 minutes at room temperature prior to assaying and also added to the assays immediately before the enzyme. Both these experiments gave identical results. The level of inhibition exhibited was dependent upon the concentration of DL-propargylglycine added. Enzyme preincubated with inhibitor had the same levels of activity as that with inhibitor added at the beginning of the assay. Assays were performed with increasing concentrations of inhibitor added immediately prior to enzyme. The table below shows the loss in activity. Each assay contained 20 mM DL-homocysteine with identical amounts of enzyme and was carried out in triplicate.

DL-propargylglycine concentration (mM)	Rate (nmol/min/ml)	Percentage Inhibition
0	3.73 (± 0.21)	0
0.5	3.50 (± 0.52)	7%
5.0	3.16 (± 0.27)	15%
12.5	2.81 (± 0.26)	25%
25.0	2.57 (± 0.14)	31%

Table 4.7: Inhibition of $MGL1^{(Y111F)}$ by DL-propargylglycine. Numbers in parenthesis are S.D. with n=3.

As a control the same amount of non-mutated MGL1 was assayed and inhibition was 100% at all concentrations of propargylglycine. The data show a non-suicide inhibition confirming the role of Tyr111.

4.2.4.4 MGL1 C113P Mutant

This mutant was overexpressed successfully in *E. coli* but exhibited no yellow colour in solution as seen for parental MGL1. This presented no problems in purification though and highly homogenous preparations were produced.

After screening a range of buffers and pH, this mutant was found to have the highest detectable activity in 100 mM sodium phosphate, pH 7.0, although the activity was very low. All further assays of this enzyme were performed under these conditions.

The saturation curve with DL-homocysteine is similar to that seen for MGL1$^{(Y111F)}$. No substrate inhibition is seen even at the highest concentrations, but the data from the lowest substrate concentrations did not produce an ideal hyperbolic curve. Again it is impossible to say what the cause of this is, due to the very low activity exhibited.

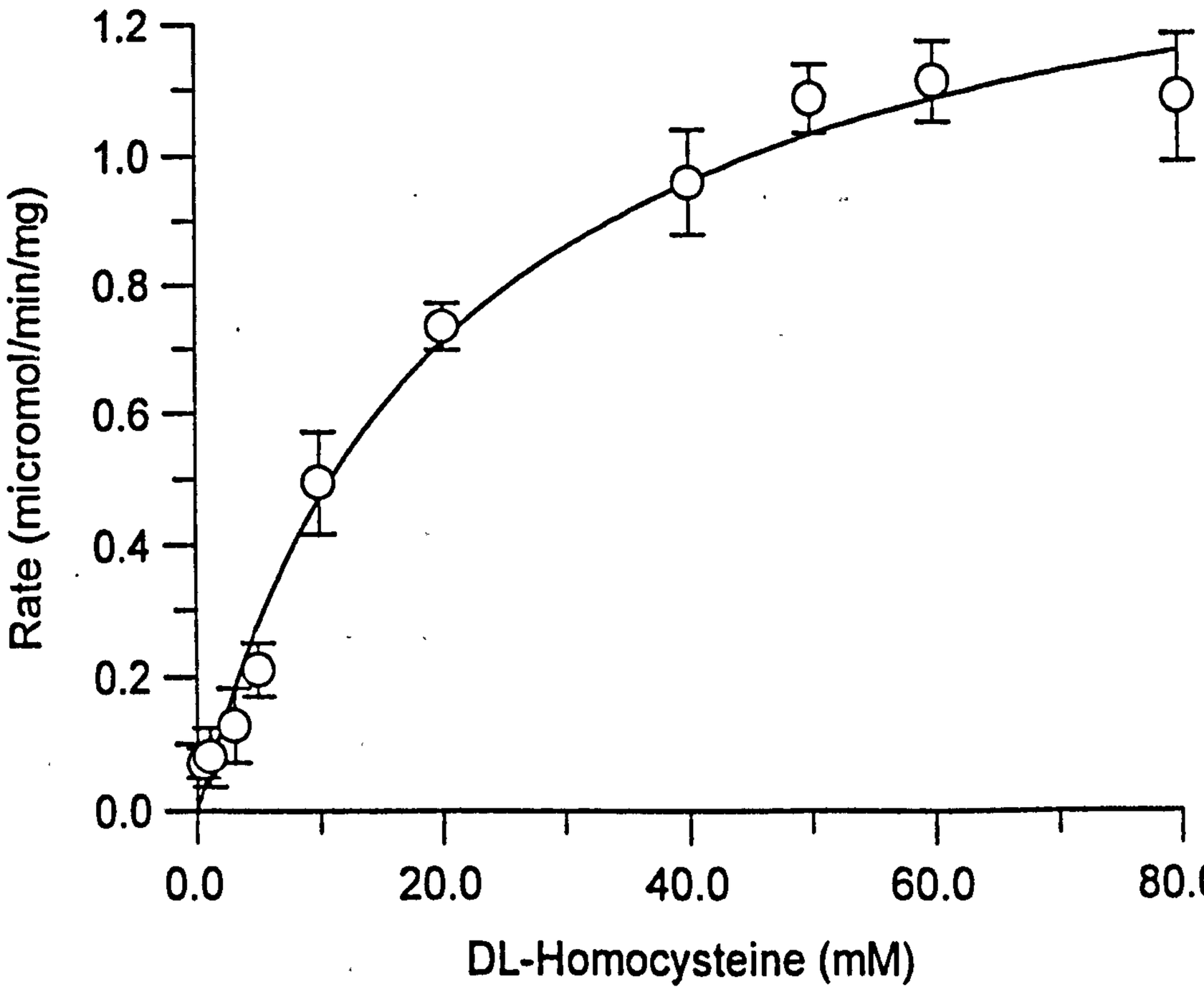

Figure 4.9: MGL1$^{(C113P)}$ reaction rate with DL-homocysteine. Data points are mean values, error bars are S.D. with n=3.

The data for this mutant enzyme appeared to adhere well to standard Michaelis-Menten kinetics. This was confirmed by producing a Hanes plot (Figure 4.10) where the only points not accurately fitting a linear plot were those at low substrate concentrations.

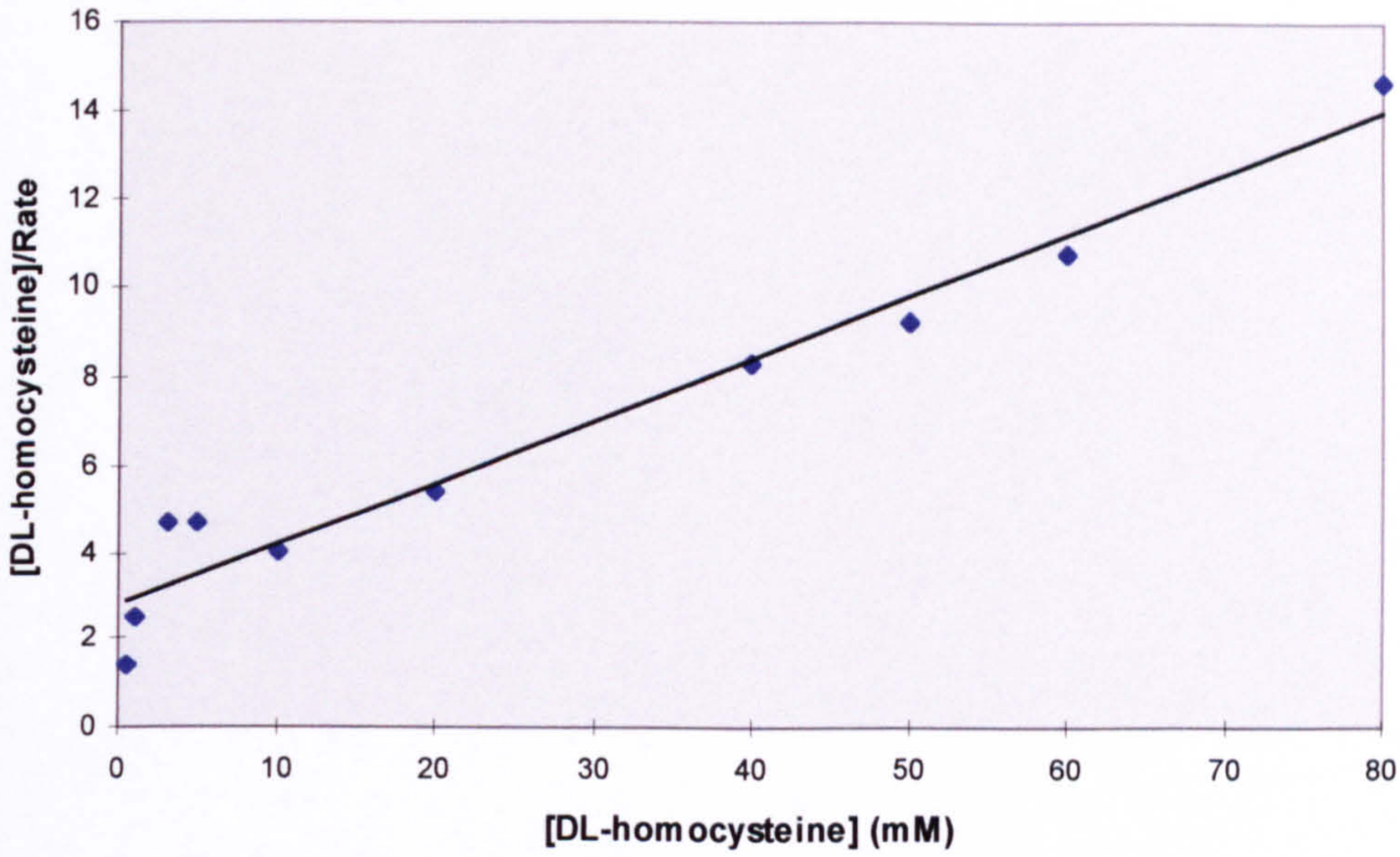

Figure 4.10: Hanes plot of DL-homocysteine catabolism by MGL1$^{(C113P)}$.

The plot above was used to calculate values for Km and Vmax values (Table 4.8). Using this information the specific activity of the enzyme was then measured and k_{cat} values calculated.

Specific Activity (μmol/min/mg protein)	0.88 (± 0.24)
Km (mM)	20.1
Vmax (μmol/min/mg protein)	1.0
k_{cat} (s^{-1})	0.7
k_{cat}/Km (M^{-1} s^{-1})	40

Table 4.8: Kinetic data for MGL1$^{(C113P)}$ with DL-homocysteine as substrate. Numbers in parenthesis are S.D. with n=4.

The results of assaying MGL1$^{(C113P)}$ cysteine catabolism are shown below in figure 4.11.

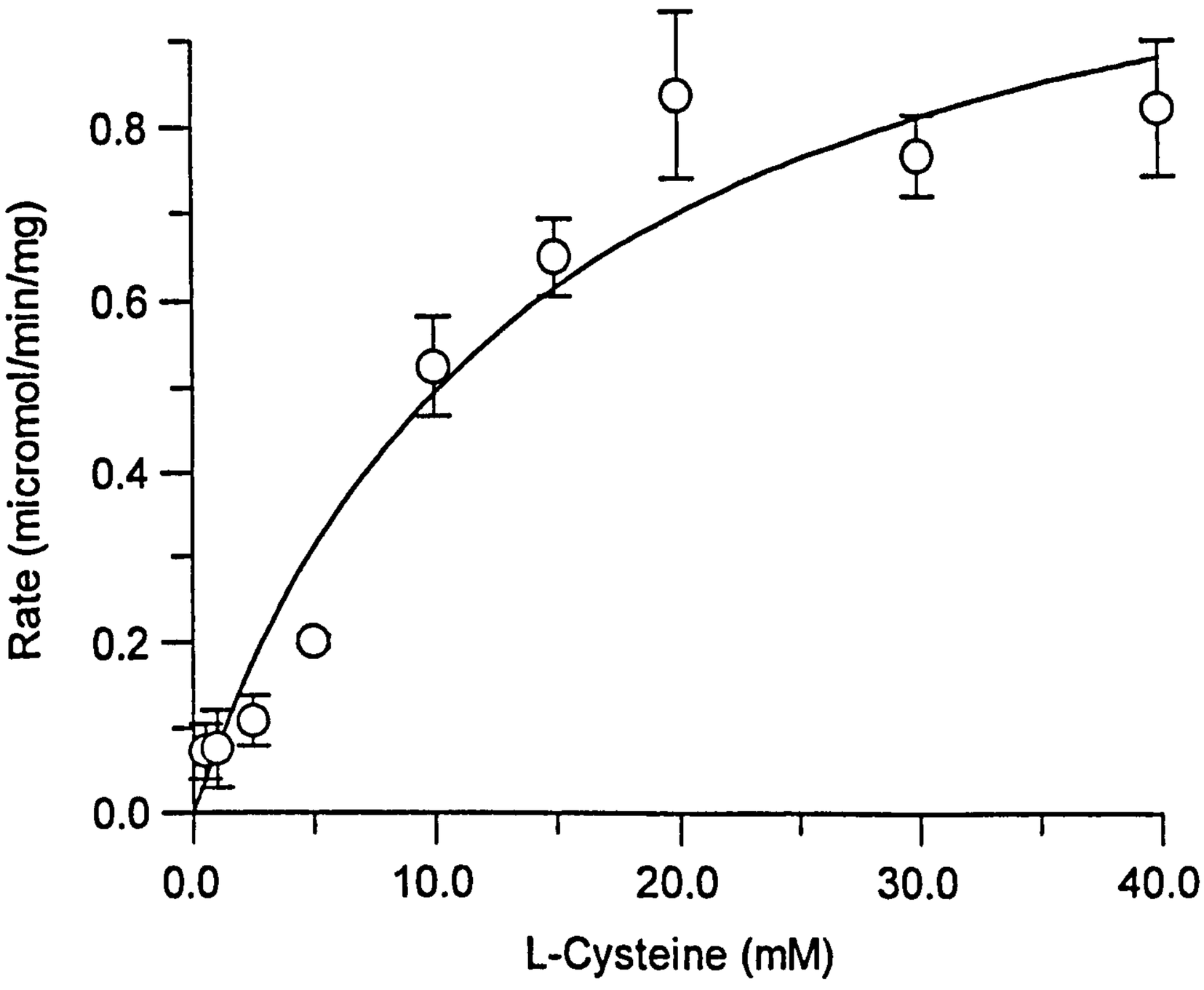

Figure 4.11: MGL1$^{(C113P)}$ reaction rate with L-cysteine. Data points are mean values, error bars are S.D. with n=3.

The kinetics exhibited by this mutant against L-cysteine are very similar to those seen with homocysteine. No evidence of substrate inhibition exists and Michaelis-Menten kinetics are apparently observed. As previously a Hanes plot was constructed to calculate estimates of Km and Vmax (Figure 4.12).

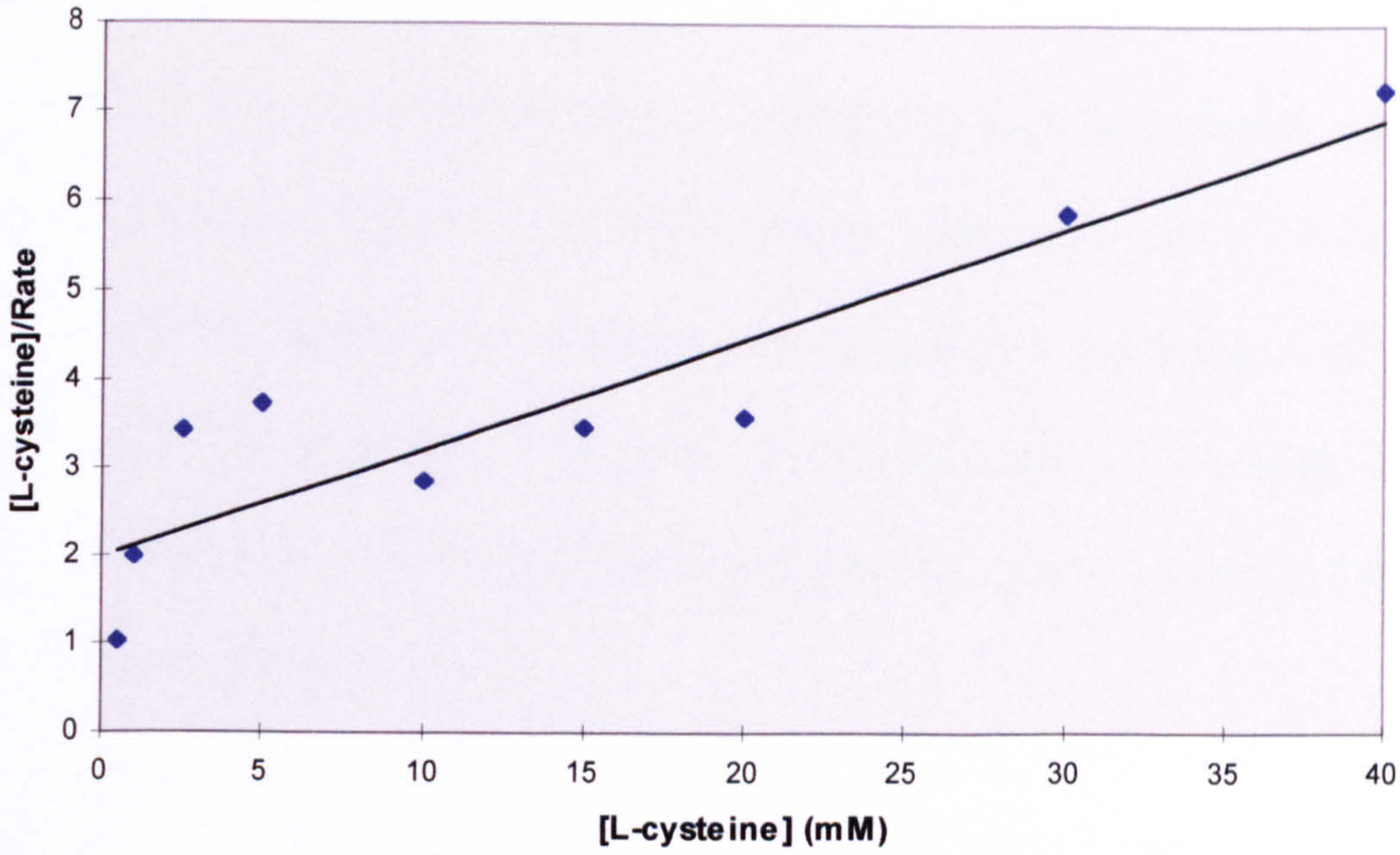

Figure 4.12: Hanes plot of L-cysteine catabolism by MGL1$^{(C113P)}$.

This plot was also very similar to that observed against homocysteine with least adherence to a linear relationship seen at low substrate concentrations. Values calculated from the plot are given below (table 4.9).

Specific Activity (μmol/min/mg protein)	1.20 (± 0.17)
Km (mM)	16.2
Vmax (μmol/min/mg)	1.1
k_{cat} (s^{-1})	0.9
k_{cat}/Km (M^{-1} s^{-1})	50

Table 4.9: Kinetic data for MGL1$^{(C113P)}$ with L-cysteine as substrate. Numbers in parenthesis are S.D. with n=3.

4.2.4.5 MGL1 K209S Mutant

This mutant was overexpressed in *E. coli* and purified successfully as a colourless soluble protein. Despite using large amounts of enzyme with either homocysteine or cysteine as substrate, no catalytic activity could be measured. Although this was the expected outcome, it was checked thoroughly to ensure the inactivity of the enzyme before attempting to prepare enzyme-substrate complexes for crystallographic analysis.

4.2.4.6 MGL1 K238A Mutant

Expression and purification of MGL1$^{(K238A)}$ was carried out identically as for unmutated MGL1. It was noticeable that the *E. coli* lysate loaded onto the Ni-NTA column was distinctly orange rather than the yellow colour normally seen. This was attributable to the mutant enzyme which, after elution from the column, retained the same colour in solution. The preparation retained this colour in the presence or absence of exogenous PLP or DTT.

The enzyme was assayed in an identical manner to all other mutants. It was determined first that activity was highest in 100 mM sodium phosphate, pH 8.0. Then a fixed amount of the enzyme was assayed with varying concentrations of substrate (DL-homocysteine and L-cysteine).

Figure 4.13 clearly shows a relationship between activity and substrate concentration consistent with Michaelis-Menten kinetics.

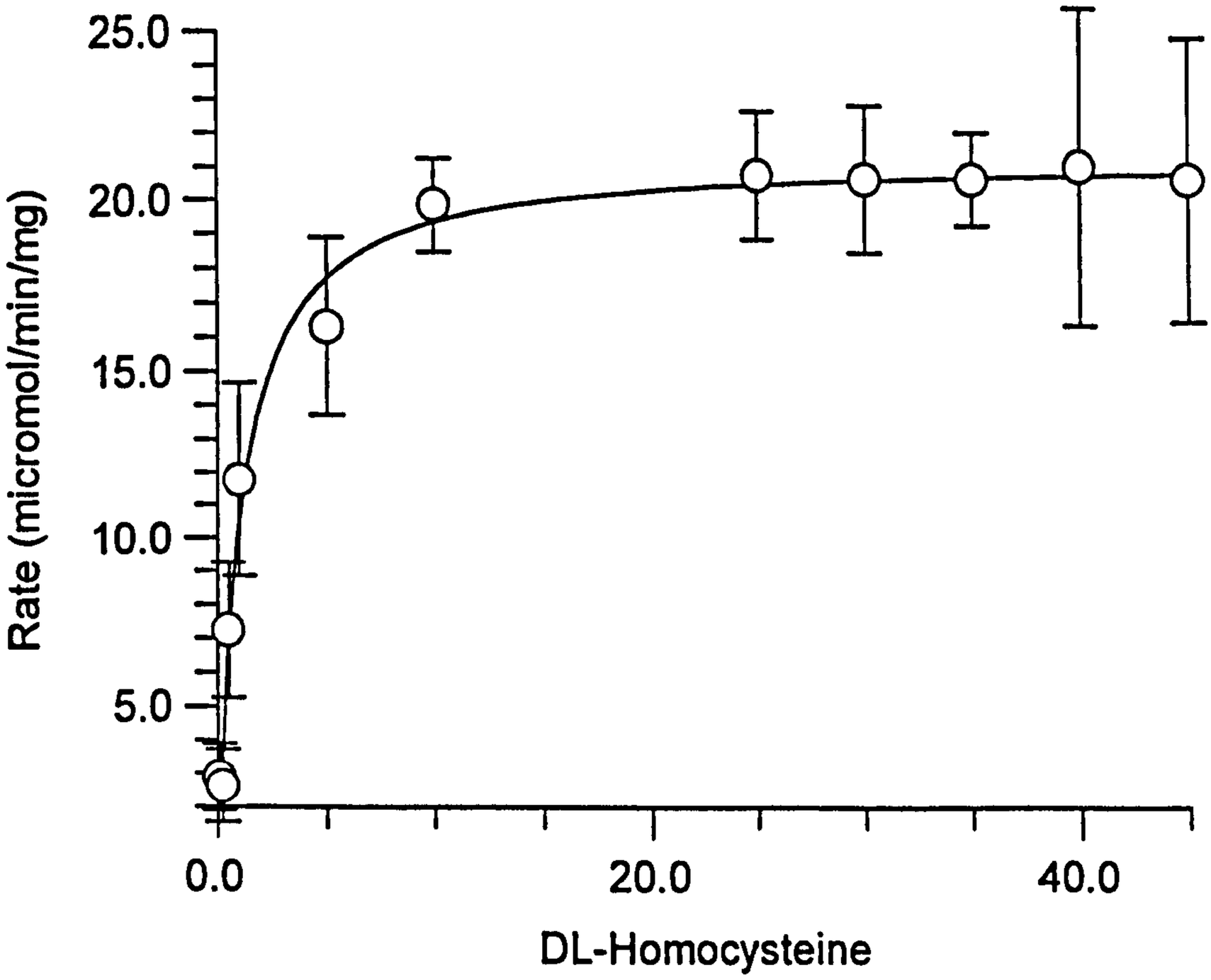

Figure 4.13: MGL1$^{(K238A)}$ reaction rate with DL-homocysteine. Data points are mean values, error bars are S.D. with n=3.

The curve shown in the figure above seems to correspond to Michaelis-Menten type kinetics well. It is immediately apparent that the rates at low substrate concentrations are very rapid and saturating levels of substrate are reached at ~ 10 mM, suggesting a low Km value. The data were expected to produce a good linear plot to calculate good estimates of Km and Vmax values (Figure 4.14).

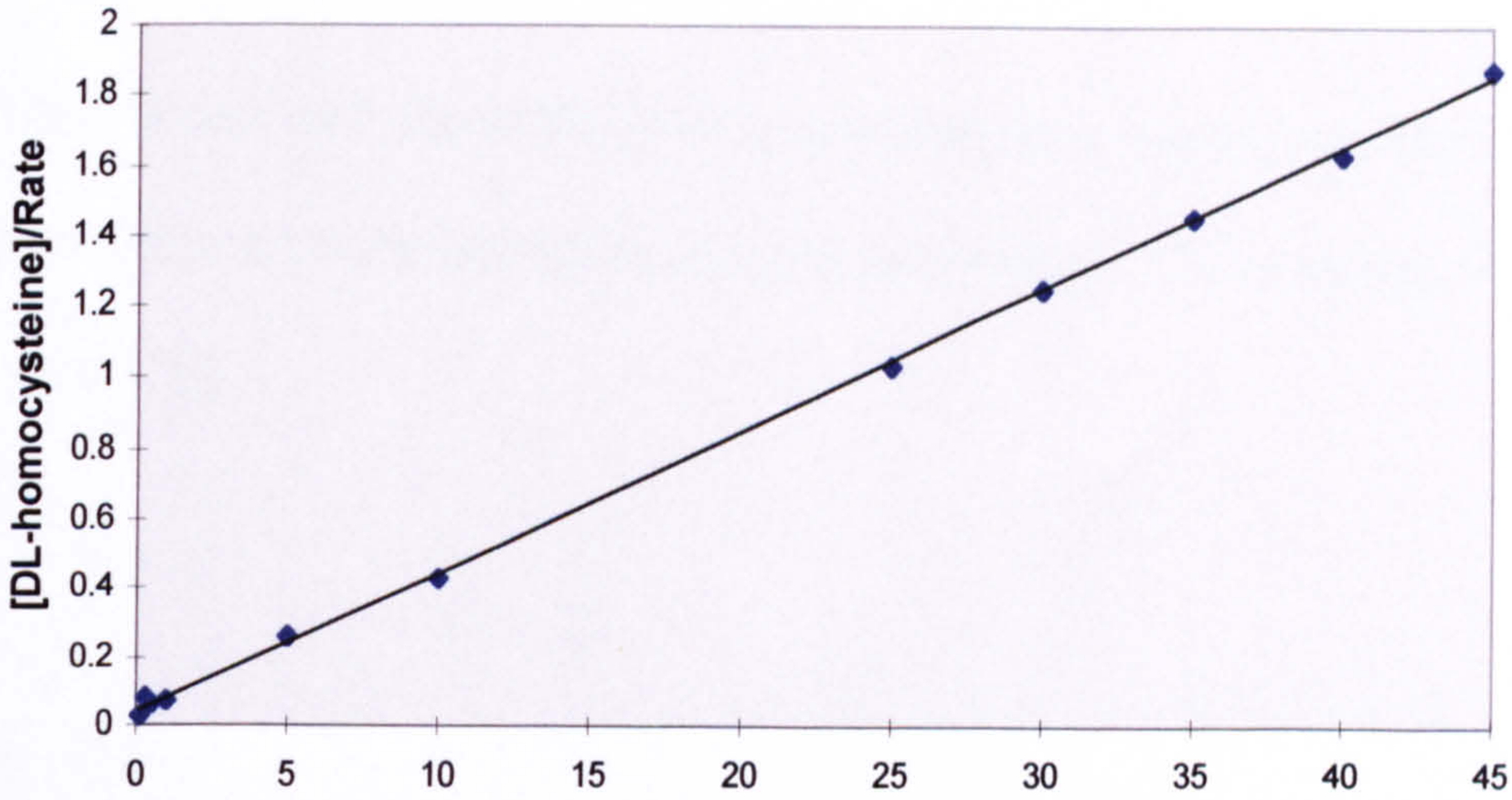

Figure 4.14: Hanes plot of DL-homocysteine catabolism by MGL1$^{(K238A)}$

As thought, the data from these experiments produced a clear linear relationship. The cluster of data points at the lowest substrate concentrations did not show the disagreement with the line of best fit as seen for MGL1$^{(C113P)}$ seen above. They did fix the y-axis intersection clearly and calculation of a Km did produce a very low figure as thought. The calculated figures for Km and Vmax from this graph and specific activity from other experiments are shown below (Table 4.10).

Specific Activity (μmol/min/mg protein)	39.2 (± 2.7)
Km (mM)	1.1
Vmax (μmol/min/mg protein)	21.4
k_{cat} (s^{-1})	16
k_{cat}/Km (M^{-1} s^{-1})	15000

Table 4.10: Kinetic data for MGL1$^{(K238A)}$ with DL-homocysteine as substrate. Numbers in parenthesis are S.D. with n=8.

The process was repeated with L-cysteine as substrate to find the highest activity toward this and the enzymes kinetic parameters. The results are shown below in figure 4.15.

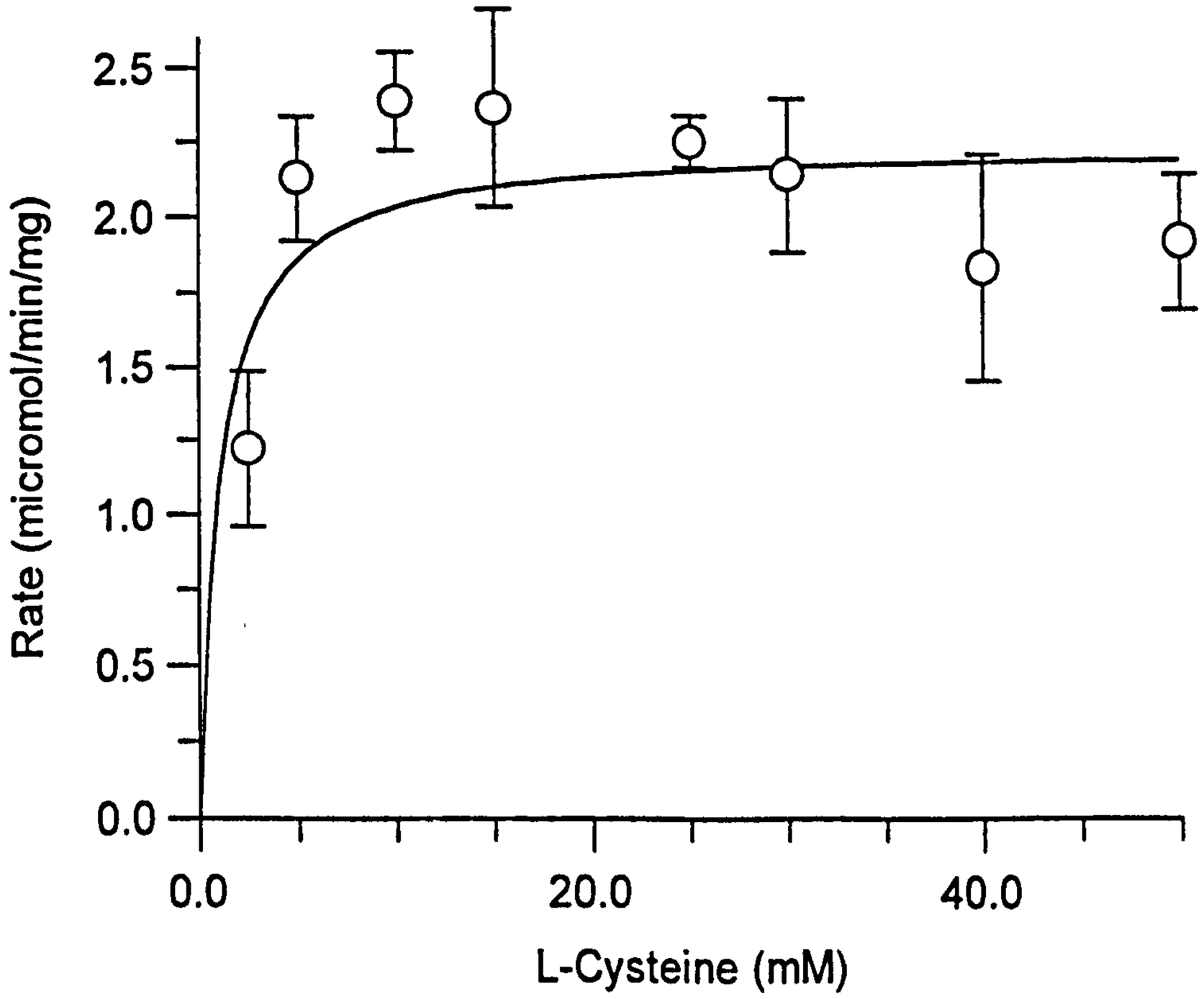

Figure 4.15: MGL1$^{(K238A)}$ reaction rate with L-cysteine. Data points are mean values, error bars are S.D. with n=3.

The data points suggested that there may have been some substrate inhibition at high concentrations. However this apparent non-adherence to a hyperbolic curve may also be explained by the errors observed in the rate measurements. Thus the data was truncated to produce a Hanes plot to calculate Km and Vmax values (Figure 4.16).

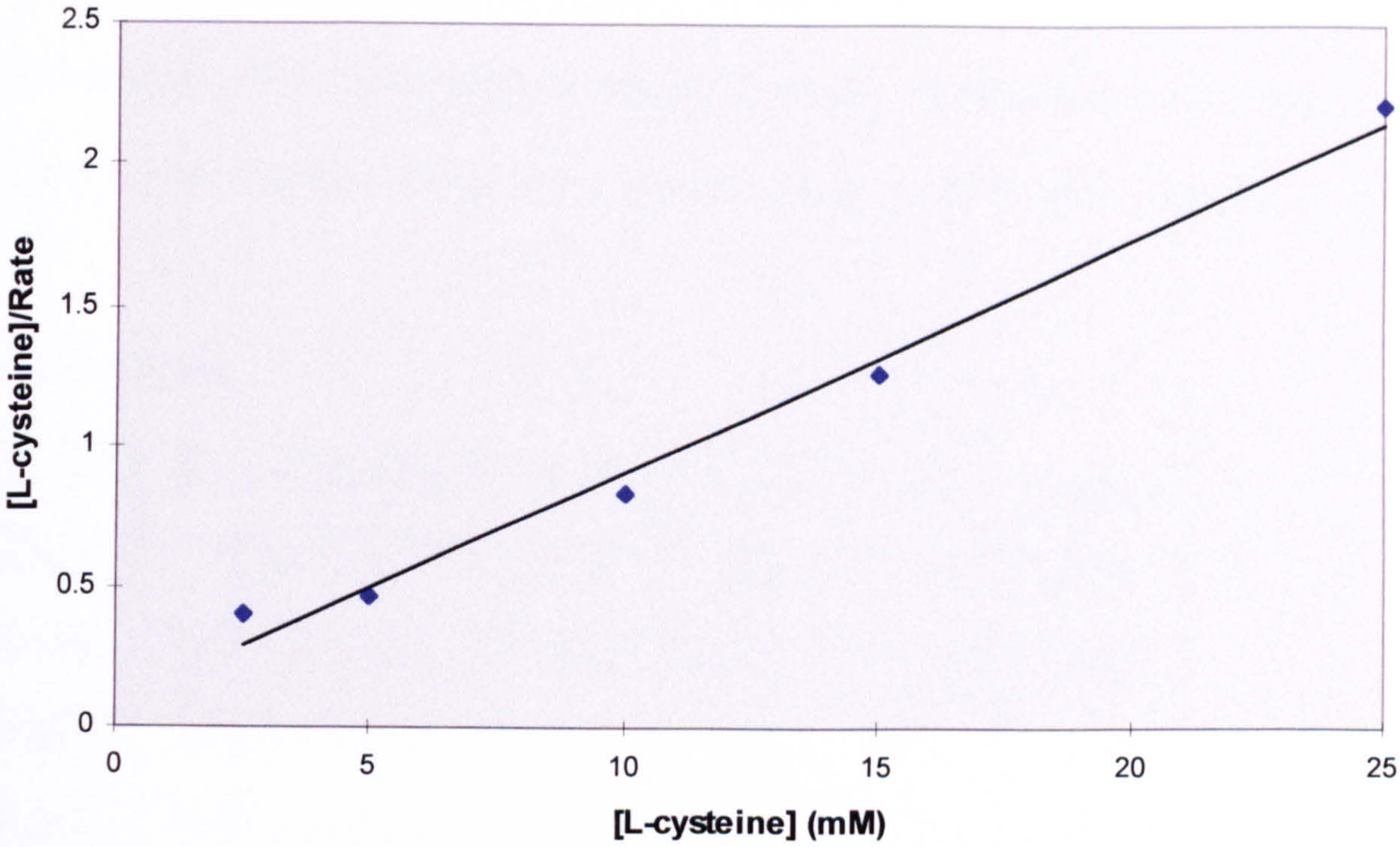

Figure 4.16: Hanes plot of L-cysteine catabolism by MGL1$^{(K238A)}$.

The linear plot was distorted by the data obtained at substrate concentrations of 25 mM and higher. The low number of data points used to plot a line of best fit though is more of a concern in terms of calculating Km and Vmax values from the graph. Despite this, the values calculated from the plot are given below (Table 4.11).

Specific Activity (μmol/min/mg)	3.27 (± 0.21)
Km (mM)	1.1
Vmax (μmol/min/mg)	2.9
k_{cat} (s^{-1})	2.2
k_{cat}/Km ($M^{-1} s^{-1}$)	2000

Table 4.11: Kinetic data for MGL1$^{(K238A)}$ with L-cysteine as substrate. Numbers in parenthesis are S.D. with n=6.

4.2.4.7 Summary of Assay Results for MGL1 Mutants

The following tables summarise the results from the previous sections of the chapter presenting the results of the enzyme assays of the mutant MGL1s.

Homocysteine

Enzyme	Specific Activity	Km	Vmax	k_{cat}	k_{cat}/Km
MGL1	370 ± 11	12.2	256	197	16100
C113G	34.5 ± 3.2 (9.3)	15.2	42	32.3	2100
R58K	308 ± 28 (83)	nd	nd	nd	nd
R58M	187 ± 22 (51)	nd	nd	nd	nd
L59F	8.26 ± 1.2 (2.2)	3.7	2.0	1.5	400
Y111F	0.99 ± 0.31 (0.3)	15.7	1.0	0.8	50
C113P	0.88 ± 0.24 (0.2)	20.1	1.0	0.7	40
K238A	39.2 ± 2.7 (11)	1.1	21.4	16.4	15000

Cysteine

Enzyme	Specific Activity	Km	Vmax	k_{cat}	k_{cat}/Km
MGL1	6.02 ± 0.63	8.5	14.9	11.5	1350
C113G	2.33 ± 0.35 (39)	9.7	4.6	3.5	360
L59F	1.75 ± 0.34 (29)	nd	nd	nd	nd
C113P	1.20 ± 0.17 (20)	16.2	1.1	0.9	50
K238A	3.27 ± 0.21 (54)	1.1	2.9	2.2	2000

Table 4.12: Summary of kinetic parameters for MGL1 mutants. Units: specific activity = μmol/min/mg protein, Km = mM, Vmax = μmol/min/mg, k_{cat} = s^{-1}, k_{cat}/Km = M^{-1} s^{-1}. The numbers in parenthesis are the percentage activities compared to unmutated MGL1 activity for the same

substrate. Data for MGL1 and mutant C113G were taken from McKie *et al.*, 1998. nd = not determined for reasons explained in the text.

Enzyme	Buffer/pH Optimum
MGL1	Imidazole pH 6.5
C113G	Imidazole pH 6.5
R58K	Acetate or MES pH 5.0 to 5.5
R58M	Imidazole pH 6.5[a]
L59F	Sodium Phosphate pH 8.0
Y111F	TRIS pH 8.8
C113P	Sodium Phosphate pH 7.0
K238A	Sodium Phosphate pH 8.0

Table 4.13: Buffer and pH optimum of MGL1 and mutants. [a] this enzyme was not examined thoroughly for pH optimum.

4.2.5 Structural Investigations of MGL1 Mutants

The effects of mutagenesis may be simply the removal of an active group, e.g. a thiol, from the catalytic centre of an enzyme or they may be more subtle, inferring a conformational change of some degree upon the enzyme. This effect is obviously more pronounced when substituting amino acids with very different properties, especially if these occur deep within the structure rather than on the surface. Perhaps the best way of interpreting any changes of activity in mutant enzymes is to have a high resolution structure to accompany enzymatic data. This is an ambitious and possibly unrealistic desire, but within the scope of this

project attempts were made to achieve this. The results are detailed in the following sections.

4.2.5.1 Crystallisation of MGL1 Mutants

Mutant enzymes were purified and prepared as detailed for MGL1 in chapter 2 of this thesis. In all cases enzyme for crystallisation was concentrated to 6mg/ml in 100 mM HEPES pH 7.5, 250 μM PLP, 1 mM DTT. Sitting drop vapour diffusion crystallisations were set up under identical conditions as for MGL1, i.e. 2.3 M $AmmSO_4$, 0.1 M NaCitrate pH 5.6, 0.2 M Li_2SO_4. The results of these experiments are summarised in the table below:

Mutant	Result
R58K, R58M, L59F, Y111F, C113G	No crystals produced
C113P	Small bipyramidal crystals (0.15 mm x 0.05 mm x 0.05 mm). Also co-crystallised with 5 mM DL-allylglycine
K209S	Large colourless bipyramidal crystals (0.6 mm x 0.4 mm x 0.4 mm)
K238A	Large orange bipyramidal crystals (0.6 mm x 0.4 mm x 0.4 mm)

Table 4.14: Results of crystallisations of MGL1 mutant enzymes

MGL1$^{(C113G)}$ previously produced in this laboratory was available at the beginning of the study. Attempting to find suitable crystallisation conditions it was screened in the same manner as MGL1 but yielded no promising results. Similar attempts were not performed with the mutants produced in this study due to time constraints.

4.2.5.2 Data Collection and Processing

Crystals of MGL1^{K238A}, MGL1^{K209S} and MGL1^{C113P} (holoenzyme and co-crystallised with DL-allylglycine) were taken to the ESRF synchrotron facility Grenoble, France for data collection. A summary of the diffraction exhibited by the crystals is given below:

Crystal	Highest Resolution
K238A	1.7 Å
K209S	1.6 Å
C113P - holoenzyme	5.0 Å
C113P - co-crystallised	5.0 Å

Table 4.15: Summary of the diffraction limits exhibited by mutant MGL1 crystals.

No data was collected for MGL1$^{(C113P)}$ because of the low resolution diffraction exhibited. Indexing of the initial frames of both MGL1$^{(K209S)}$ and MGL1$^{(K238A)}$ indicated the same space group as unmutated enzyme, P 3_112, and similar unit cell dimensions. A strategy was then computed to collect complete datasets for both crystals.

In an attempt to produce enzyme-substrate complexes, crystals of MGL1$^{(K209S)}$ were soaked with homocysteine and cysteine. Unfortunately the results of these soaks was a total loss of diffraction by the crystals. The gross morphology of the crystals did not alter during the soaking and they remained intact for several minutes. The loss of diffraction was concluded to be due to a conformational change. No further attempts were made to produce complexes with these crystals.

Only the data from MGL1$^{(K238A)}$ was processed and analysed further, a summary of the statistics is presented below. All data were processed with DENZO and scaled using SCALEPACK (Otwinowski and Minor, 1997).

Resolution		35 - 1.8 Å
Temperature		100K
Oscillation Range		0.5°
Completeness	Overall (40.0 – 1.82)	98.6%
	Highest (1.91 – 1.82)	84.7%
Unique Reflections		90109
R_{merge}		5.6

Table 4.16: Data collection and processing statistics for MGL1$^{(K238A)}$.

4.2.5.3 Structure Solution of MGL1$^{(K238A)}$

The solution of this mutant was relatively straightforward due to the similarity between it and native MGL1. Phases were obtained directly from the refined

holoenzyme structure and the same subset of reflections used for Free R-factor determination extended to include the higher resolution data.

4.2.5.4 Model Building and Refinement of MGL1$^{(K238A)}$

Model building of the mutant was carried out with QUANTA using the coordinates of MGL1 holoenzyme to trace the first maps. These maps were of a very high quality due to the similarity of the mutant to the refined holoenzyme structure, from which the phase information was obtained. No breaks in the electron density for main chain atoms were evident and there was no need for averaging of the maps. In the first round of model building waters were included from the holoenzyme structure and individually checked. Poorly defined atoms were removed and extra waters added, the glycerol molecule and sulphates were clearly visible in the first maps. Refinement was subsequently performed with REFMAC using a bulk solvent correction and maps created using the programme FFT from the CCP4 suite. Geometric restraints were relaxed after the second round of refinement. A summary of the refinement process and statistics is given below (Table 4.17):

Round of Refinement	R-factor	R_{free}
1. Phases taken from holoenzyme model	22.1	26.6
2. Few changes necessary, extra waters added	21.3	25.3
3. Poor waters removed, minor adjustments only	20.2	23.4
4. Final round	19.8	23.4

Table 4.17: Summary of refinement of MGL1$^{(K238A)}$ structure.

4.2.5.5 MGL1$^{(K238A)}$ Structure

The following section details the structure of the mutant MGL1$^{(K238A)}$ at 1.7 Å. The crystallisation and indexing of data from the crystals implied that the structure was likely to be very similar to the native enzyme. Purification of the enzyme revealed that instead of the yellow colour associated with the parental enzyme, purified mutant enzyme was orange and this was clearly visible in the crystals. This phenomenon was not attributable to a change in pH of the solution as conditions were identical in each case, and as seen in the native structure, Lys238 does not make a direct interaction with the cofactor.

4.2.5.5.1 Secondary Structure

The secondary structure of the mutant enzyme is, as expected, identical to the native enzyme. The two monomers in the asymmetric unit do not differ appreciably as is seen in the holoenzyme and inhibitor complex structures. An RMS deviation of only 0.294 Å for main chain atoms was seen when superimposing the two subunits. Similarly, superimposing the dimers of the mutant structure with the holoenzyme and inhibitor complex structures gave an RMS deviation of only 0.235 Å and 0.262 Å respectively.

4.2.5.5.2 Active Site Areas

The active site of the mutant is also largely similar to the holoenzyme structure. There is a clear difference seen at the site of the lysine to alanine change at residue 238. There is no electron density past the position of the Cβ and the space vacated is not clearly occupied by any other atoms. Figure 4.17 shows the electron density in the region around Ala238.

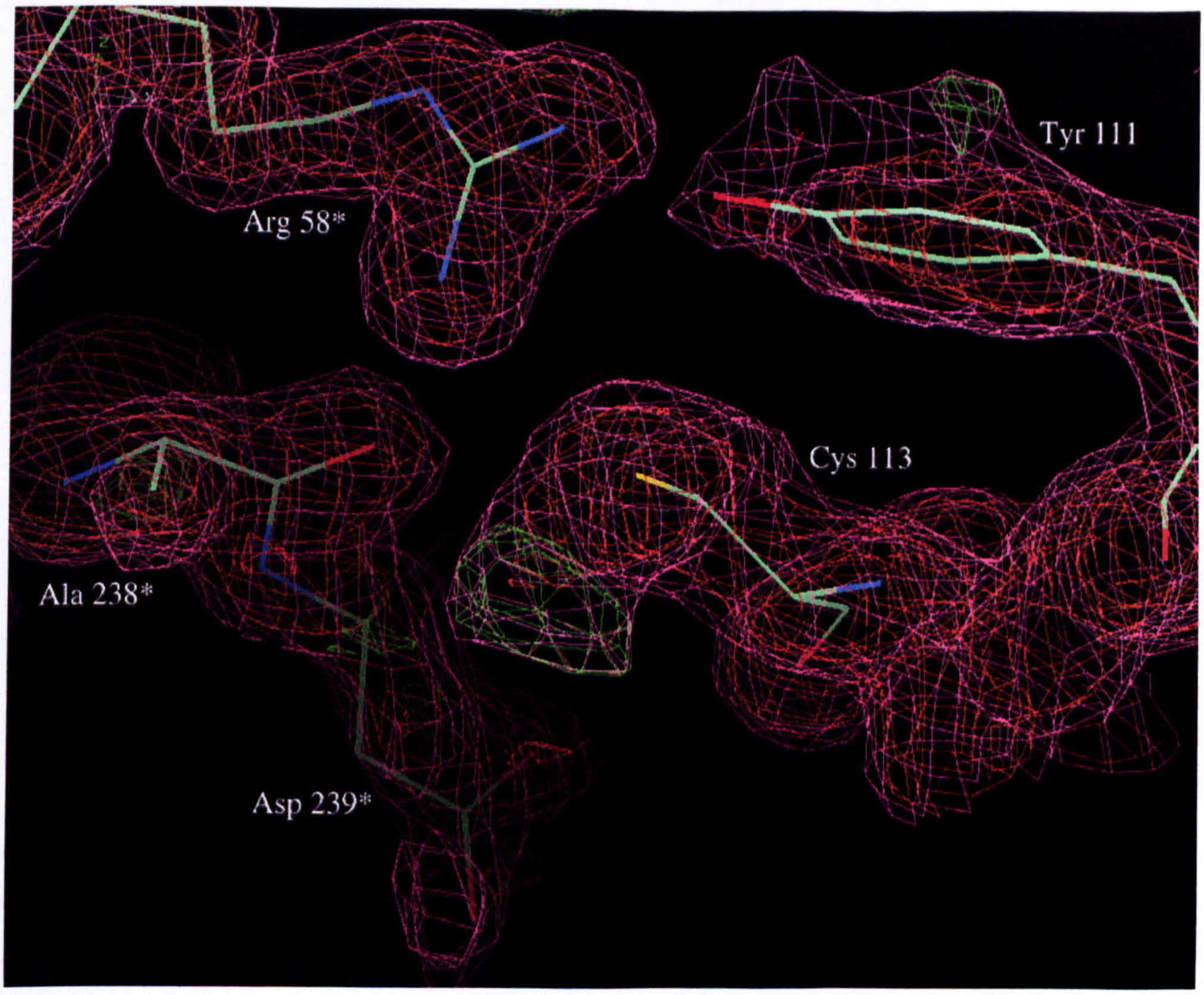

Figure 4.17: $2F_O - F_C$ and electron density maps of the area around Ala238 in the MGL1$^{(K238A)}$ structure. The $2F_O - F_C$ map is contoured at one and two σ in pink and orange, respectively. In the $F_O - F_C$ map positive density is indicated in green, negative density in red. Only Arg58, Tyr111*, C113*, Ala238 and Asp239 are shown in the figure for clarity.

There is weak density nearby to the Cβ of Ala238 which did not successfully refine as either water molecules or an alternative conformation for neighbouring Asp239. Both of these have been omitted from the final model, although it is possible that both may actually be present at low occupancy. The electron density for the side chain of Asp239 is significantly poorer than for neighbouring residues and is reflected in the high average B-factor of 38.2. This reinforces the idea that the side chain may possess a dual conformation.

Also evident in the mutant structure is the change in the electron density for Tyr111. There is a small amount of positive density above the plane of the phenol ring between the side chain and the cofactor pyridine ring. Although not clear in figure 4.17, this density is consistent with the ring rotating ~90° to position itself parallel with the cofactor. Attempts were made to refine a dual conformation of Tyr111 but were unsuccessful. A change such as this would force a movement of the cofactor as it is within Van der Waals distance of the phenol side chain.

The most intriguing observation in the active site though is the density around Cys113. At all stages of refinement, there was positive density adjacent to the Sγ. Unlike the maps for the holoenzyme and inhibitor complex this was not attributable to a dual conformation. Instead the unaccounted for density in the mutant structure extends in the same direction from the Sγ, resulting in an angle of ~120° between Cβ-Sγ-X, where X is the unknown atom. The distance appears to be approximately 1.7 Å from Sγ to X indicating a covalent bond between the two atoms.

It was supposed that the orange colour exhibited by this mutant was due to a shift in the position of the cofactor relative to the residues coordinating it. To test this monomers of the mutant, holoenzyme and inhibitor complex structures were aligned using LSQKAB and the models superimposed. Figure 4.18 shows the three cofactors aligned in this way.

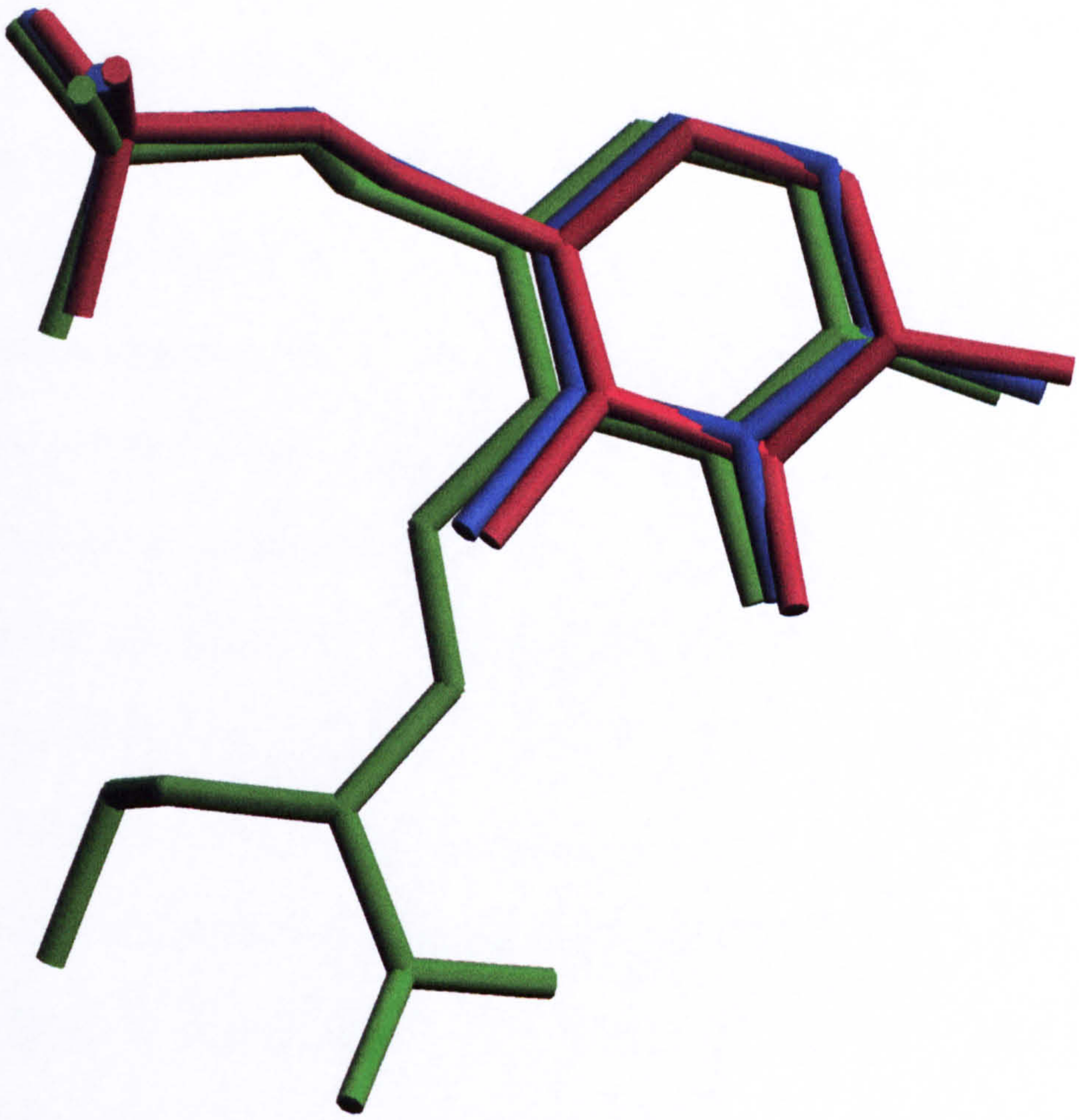

Figure 4.18: Superposition of the cofactors from the three MGL1 structures. Green - PLP + propargylglycine, Blue - holoenzyme structure, Red - K238A mutant structure. The figure was produced using SETOR.

There is very little movement of the cofactor in the MGL1$^{(K238A)}$ structure in comparison with the others. The cofactors were aligned through their position relative to the polypeptide chains but it was evident that there were no significant movements of any residue directly interacting with the cofactor.

4.2.6 MGL From Other organisms

MGL activity has been detected in a number of organisms. Much of these data are from total cell extracts and are therefore a little unreliable as related enzymes such as cystathionine γ-lyase (CGL) are able to catalyse some of the same reactions as MGL, albeit at much lower levels. At present there are only four *mgl* genes in the database that have been confirmed biochemically to encode MGL enzymes, two each from *T. vaginalis* and *P. putida* (Inoue *et al.*, 1995; Hori *et al.*, 1996; McKie *et al.*, 1998). These show a high degree of homology to other members of the γ-family making identification of putative MGL's difficult. Indeed the cloning of *mgl* from *T. vaginalis* was carried out using degenerate oligonucleotides based upon CGL sequences. As mentioned previously, in both the bacterial and protozoal enzymes studied biochemically so far a cysteine residue (Cys113, *T. vaginalis* MGL1 numbering) is important for activity and 100% conserved (Nakayama *et al.*, 1988; McKie *et al.*, 1998). This residue is not found in other γ-family enzymes, which was suggestive that the MGL mechanism differed from those of other PLP-dependent enzymes such as cystathionine synthases and cystathionine lyases. Moreover, the cysteine residue can be used as a marker in identifying putative MGL's. Searching the available data released from the various sequencing projects has revealed possible MGL enzymes in just a few organisms: *Porphyromonas gingivalis, Bacillus anthracis, Treponema denticola, Caulobacter crescentus* and *Shewanella putrifaciens* (accessible at http://www.tigr.org). An alignment of these sequences shows their extensive similarities (Figure 4.19).

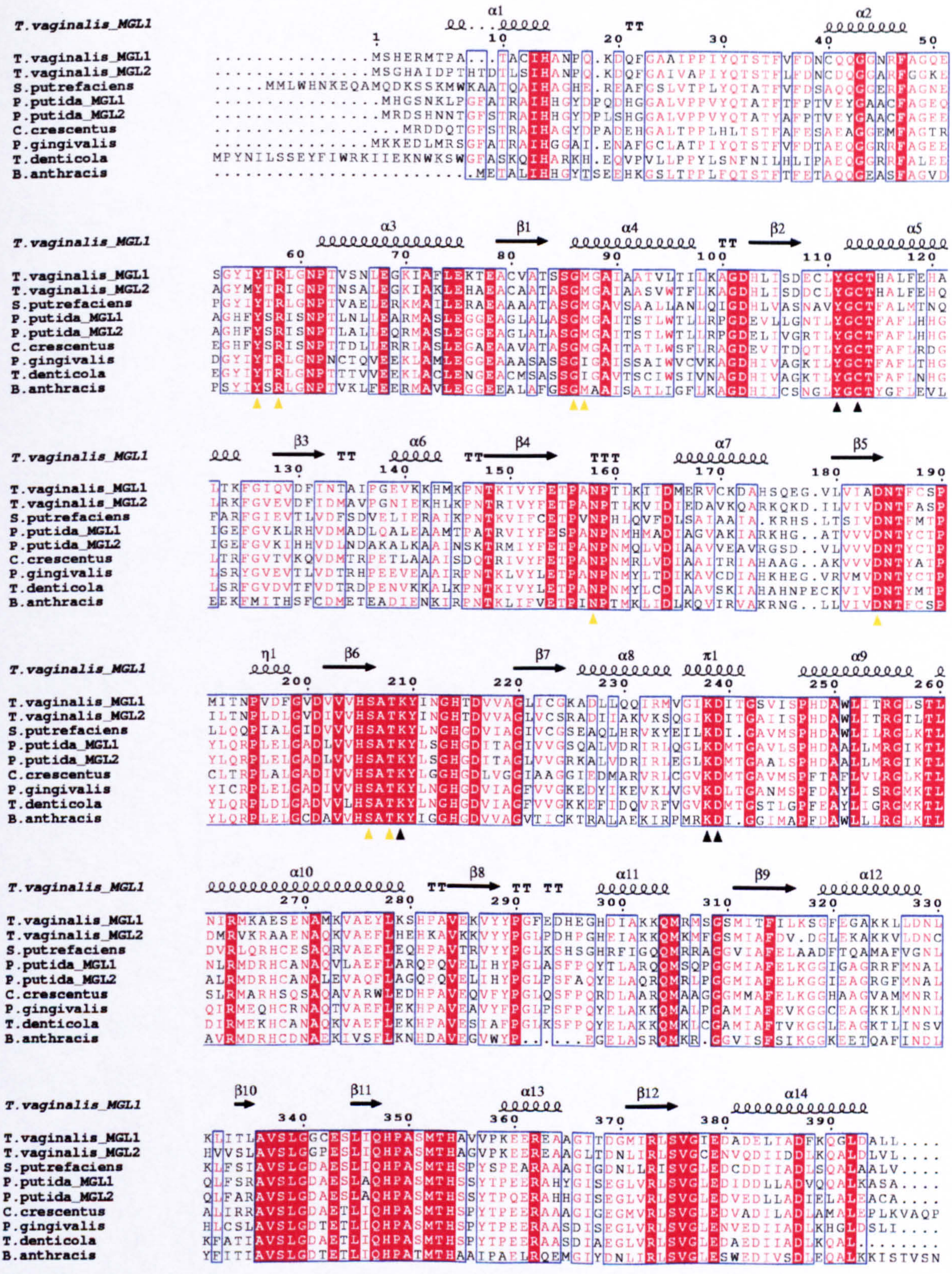

Figure 4.19: Sequence alignment of confirmed and putative MGL enzymes. Secondary structure elements of MGL1 are indicated. α-helices are represented as coils, β-strands as black arrows, each is numbered as for figure 3.6. Conserved residues are coloured in red and boxed. Black triangles indicate those residues proposed to have a role in the catalytic mechanism. Yellow triangles indicate residues coordinating the cofactor. The figure was generated by ESPript. All numbers refer to MGL1.

There are 87 conserved residues amongst all 9 sequences and the level of similarity exhibited to MGL1 ranges from 58% against the *B. anthracis* sequence to 85% against MGL2. With the exception of Met87, all residues involved in cofactor and substrate binding and those that appear to be involved in catalysis are conserved. Met87 is alternatively substituted for an isoleucine and the main chain amine group of the residue is involved in the interaction and so the side chain is less crucial. Significantly Lys238 and Asp239 are also conserved in these sequences but absent from other PLP-dependent enzymes. Thus in identifying possible MGL enzymes from the abundance of other PLP-dependent enzymes retrieved in a BLAST search of the available databases, Cys113, Lys238 and Asp239 are critical. Not a single instance was found of any one of these residues being present without the other two.

4.3 DISCUSSION

Results have been presented in this chapter on rationally designed site-directed mutagenesis of MGL1. This technique has been applied to many proteins to investigate roles of particular residues. It has also been successfully used on MGL from *T. vaginalis* to investigate the contribution of an active site cysteine to catalysis (McKie *et al.*, 1998). For this study mutations were selected based upon the structure of MGL1 and its comparison to CBL from *E. coli* (Clausen *et al.*, 1996). Eight mutants were designed, although one of these was not successfully produced, targeting a total of six different residues. Attempts were also made to solve the structures of these seven mutants, crystals were grown of three and one structure was solved to a resolution of 1.7 Å.

When searching for other MGL sequences in the databases the homology to other members of the γ-family of PLP-dependent enzymes was obvious. Identifying any of these positively as MGL though was difficult. To attempt to do this, we used information from the structure of MGL1 and the results of SDM which identified key residues in the active site. Additionally we utilised the structures published of CBL and CGS to highlight differences. With this information it allowed us to putatively identify ORFs from the genome sequencing projects that encode likely MGLs. In the active site of γ-family PLP-dependent enzymes Tyr55, Arg58, Tyr111 and Lys209 (MGL1 numbering) are usually present. However key residues in MGL1, Cys113, Lys238 and Asp239, are absent. By performing BLAST searches with MGL1 and examining those hits with high homology for the presence of these seven key residues, putative MGLs were identified. Frequently these organisms had apparent homologues of one or more cystathionine-utilising enzymes, but in each case these other enzymes lacked Cys113, Lys238 and Asp239. Only those sequences which had these three residues are thought to be MGL enzymes. Interestingly, no example was found in any of these organisms of two isozymes of MGL as is found in both *T. vaginalis* and *P. putida*.

The technique employed to produce mutants utilises oligonucleotide primers that flank the target site by 15 to 20 bases. It was therefore a surprise to isolate a large number of putative mutant plasmids whose sequences in these regions was not encoded by the primer or original clone. The number of these varied but in the case of the C113S mutant no correct mutations were isolated. The production of this mutant was abandoned because of the work involved in analysing the other

seven. It can only be concluded that the cause of these "random" changes were due to the oligonucleotides used. Their synthesis may not have been entirely specific, such that primer populations incorporating errors were supplied.

The catalytic mechanism proposed in the previous chapter gives a direct role for both Tyr111 and Lys209 in catalysis, with the activity being strongly mediated by the action of Cys113. Purified MGL1$^{(K209S)}$ was inactive according to the assays performed which is in agreement with this mechanism. However crystals of the enzyme no longer diffracted X-rays after adding substrate, which might suggest the retention of some activity. However, this may equally have been produced by a conformational change resulting from substrate entering the active site. The published structure of CGS from *E. coli* shows an asymmetry in the B-factors for residues in the N-terminal domain which is postulated to reflect a conformational change. This may be necessary for the active site of CGS which has to accommodate two substrates (L-cysteine and O-acyl-L-homoserine) but there is no evidence for a conformational change in MGL1 when the holoenzyme and inhibitor-complex structures are compared.

It was thought that the mutation MGL1$^{(L59F)}$ would produce an enzyme with changes in substrate specificity and a probable drop in activity. The calculated specific activities show a much larger fall in activity than expected, although this fall is much larger for homocysteine than cysteine (98% and 71% respectively), which was predicted. The k_{cat} and k_{cat}/Km values calculated for this mutant also agree with this, showing a much larger fall in turnover and affinity for homocysteine than cysteine. Because no structural data is available to

accompany these results we are unable to say exactly what is responsible for these changes. The residue is adjacent to Arg58, changes in which we have seen in this chapter result in significant alterations in enzyme activity. Thus the activity of the MGL1$^{(L59F)}$ mutant may be due to both a change in the active site and movement of Arg58.

The assay results of both MGL1$^{(R58K)}$ and $^{(R58M)}$ are difficult to explain. Both exhibit substrate inhibition but only the lysine mutant shows the activation/inhibition by acetate and similar buffers. The reason for substrate inhibition is unknown but it is possible to give some explanation for the lag in activity seen at intermediate substrate concentrations ([s]). With a low level of residual activity remaining during inhibition by high [s], as the substrate is depleted it eventually drops below inhibitory levels. This accounts for the activity seen after a few minutes delay.

Previously Cys113 has been shown to be important for activity and mutation to a glycine has resulted in the loss of ~90% of homocysteine catabolism. The structure of MGL1 shows that the distance from the $S\gamma$ of Cys113 to the substrate is too large for a direct interaction occur. The position of the cysteine only two residues from Tyr111 does allow for the possibility that the change in activity is due to movement of this residue. To further test the role of the cysteine, mutations to serine and proline were designed and also the Lys238 to alanine change because of its proposed interaction with Cys113. The loss in activity of MGL1$^{(C113P)}$ is very probably due to a change in the active site geometry. The glycine mutation at this position would allow a greater freedom of movement for the polypeptide chain because of the lack of steric hindrances caused by the

absence of a side chain. A proline mutation though would restrict the conformations permissible more than any other residue. The restricted geometry of the proline permits only certain phi/psi angles and these may well distort the neighbouring structure, displacing Tyr111. As can be seen in figure 3.4, an alignment of CBL and MGL1 shows the sequences to be quite different in a stretch of 18 residues (104 to 121) with only Tyr111 being conserved and one other similarity. Thus where a proline residue is optimal in CBL it has a severely detrimental effect in MGL1. The fact that this mutant crystallised in identical conditions to MGL1 probably indicates the overall structural changes are not large or at least do not affect crystal contacts. It is interesting to note however that this mutant has similar levels of activity toward homocysteine as $MGL1^{(Y111F)}$, but unlike this change, $MGL1^{(C113P)}$ exhibits activity toward cysteine.

The role of Lys238 in the mechanism of MGL1 was proposed to be through the combined influence of it and Asp239 to ionise Cys113. The fall in specific activity of this mutant, compared to the parental MGL1, by an order of magnitude is matched by the drop in Km and k_{cat} values. The k_{cat}/Km value however, which can be taken as a measure of the specificity of the enzyme for a substrate, is similar to the values for parental MGL1. This data alone would indicate the possibility of a mutation that resulted in no significant structural change as substrate affinity was not substantially affected. The fall in activity though points to the loss of a functional group important but not essential for catalysis. These conclusions are confirmed by the mutant enzyme structure. The only significant difference is the area around residue 238 and the side chain of Cys113. The electron density around the Sγ suggests the side chain is oxidised. With the proposed action of

Cys113 dependent upon its ionisation (by Lys238 and Asp239) oxidisation would prevent this. Thus the data obtained from this mutant agrees with our hypothesis on the role of Lys238 and Cys113 and the importance of the ionised state in the mechanism.

We are confident that the action of Cys113 on the substrate is crucial for optimum activity, but it is still difficult to explain. The distance from Cys113 Sγ to substrate β and γ atoms is ~ 6.5 and 5.0 Å respectively. The agreed mechanism of action for β- or γ-elimination requires a nucleophilic attack at the β atom, obviously too distant to be influenced by this residue directly. Reconsidering the mechanism we know that the "quinonoid intermediate" is formed after proton abstraction at Cα. The new distribution of π molecular orbitals produces a "vinylogous enamine". This structure produces a cofactor which is electron rich and basic and can either be protonated or push out a leaving group. However the thiol group of homocysteine (and cysteine) is a strong nucleophile and it is presumably the assistance here by Cys113 that produces the high levels of activity we see.

Cys113 and its proposed oxidation in MGL1$^{(K238A)}$ is likely to be responsible for the orange colour seen in the enzyme preparation, as the structure of the mutant reveals there is no movement of the cofactor. Although there appears to be no close contact between the cofactor and the side chain of Cys113 it is noteworthy that neither of the mutants MGL1$^{(C113P)}$ and MGL1$^{(C113G)}$ have any yellow colour in solution, as seen in unmutated MGL1. However the reason for this colour change has yet to be explained.

It is tempting to look for answers about Cys113 in the mechanism of NifS enzymes and their homologues. These enzymes all have proposed mechanisms with an essential active site cysteine forming a persulphide after catalysis of cysteine or selenocysteine (Zheng *et al.*, 1994). However they all also predict that the same residue takes part in a nucleophilic attack upon the substrate. Unless a large conformational change takes place in MGL1, for which there is no evidence in the structure, this is highly unlikely. The remaining activity in the MGL1$^{(C113G)}$ mutant also totally discounts the possibility of it acting as an essential active site nucleophile. Thus despite some outward similarities these mechanisms are fundamentally different from MGL1.

The structure of cystalysin from *Treponema denticola* (Krupka *et al.*, 2000) was also examined for help in defining the mechanism of MGL. This enzyme catalyses an $\alpha\beta$-elimination reaction on cysteine to release H_2S, pyruvate and ammonia rather than the production of alanine and enzyme bound persulphide by NifS-like enzymes. However in this structure there is no analogous active site cysteine. Catalysis is again predicted to be entirely due to the ubiquitous active site lysine although here it is under the influence of the PLP phosphate group, a water molecule and a tyrosine and alanine residue in the active site. It has to be concluded then that after a search of the available literature MGL has a novel method of producing a good leaving group from a strong nucleophile.

The inability to detect any activity with MGL1$^{(K209S)}$ is consistent with the finding in other mutated PLP-dependent enzymes where the residue is essential. In CBL and CGS it has been proposed that the analogous lysine to this is the sole

residue involved in catalysis (Clausen *et al.*, 1997, 1998). This has been considered for MGL1, but thought to be unlikely because of the distance from the amine group of the lysine side chain to the proposed position of the substrate Cβ. Instead it is proposed that Tyr111 would fulfill this role. MGL1$^{(Y111F)}$ retains 0.3% of its activity towards homocysteine compared with the unmutated enzyme and no measurable activity towards cysteine. Although very low, this is not predicted for the loss of an essential catalytic residue. Thus the question to be answered is can the very low levels of activity still exhibited by MGL1$^{(Y111F)}$ be explained in terms of the mechanism proposed here? The only way in which the mutant activity data can be made to fit our hypothesis is if there is an additional mechanism for the αγ-elimination of homocysteine. There appear to be only two options for this: the action of Lys209, or the action of Cys113. Perhaps it is possible that the trace activity retained by the mutant is in fact due to Lys209, as proposed by others. The unfavourable position seen for Lys209 in MGL1 may not be totally prohibitive for catalysis and may be able to deprotonate the substrate Cβ albeit at very low levels. The action of Cys113 in conjunction with the cofactor may also be enough to cleave the Cγ-Sδ bond. It is feasible that the action of the electron rich quinonoid intermediate and the ionised Sγ of Cys113 (approximately 3.5 Å from the Sδ of homocysteine) may break the Cγ-Sδ bond. If neither of these possibilities are correct then an alternative role for Tyr111 must be given which accounts for the loss of activity.

An alternative role for Tyr111 has already been proposed in which the phenolate side chain is critical for transaldimination (Clausen *et al.*, 1998). That is, the

conversion from enzyme-PLP (internal aldimine) to substrate-PLP (external aldimine) complex. This is the primary step in catalysis and involves the removal of a proton from the incoming substrate amine group and passing this to the displaced lysine amine (Lys209). It is this conjugate form of PLP-substrate that labilises the substrate Cα and allows the reaction to proceed. The role of Tyr111 in this still relies upon the negatively charged side chain hydroxyl and so loss of this should also prevent transaldimination. In MGL1 the proposed method for transaldimination is controlled by the cofactor itself. The hydroxyl group on C3 of the cofactor is generally accepted to be negatively charged. This should then be able to shuttle protons from the incoming substrate, four atoms along to the lysine amine, although we have no evidence for this.

Alternatively the activity data of the mutant can be explained in terms of the CBL/CGS mechanism proposed. That is, transaldimination by Tyr111 is abolished in the phenylalanine mutant and the remaining low activity levels are a result of sub-optimal transaldimination, possibly by the cofactor. The kinetic data agrees with this, producing a k_{cat}/Km much lower than that for unmutated MGL1 which indicates a much lower specificity for the substrate: Unfortunately the data we have does not allow us to rule out either hypothesis and so further investigation is undoubtedly needed to identify the true function (or functions) of Tyr111.

What is confirmed by the data is not only the importance of Tyr111 in activity but also its role in inhibition by propargylglycine. The action of the inhibitor on MGL, CGS and CGL has been studied biochemically and confirmed a covalent linkage

to the enzymes (Washtien and Abeles, 1977; Johnston *et al.*, 1979). Though clear and unambiguous in the inhibitor complex structure the verification of no inactivation of $MGL1^{(Y111F)}$ by the compound confirms the role of the side chain hydroxyl group as the site for covalent attachment.

CHAPTER 5: INVESTIGATION INTO *T. VAGINALIS* CYSTEINE SYNTHASE

5.1 INTRODUCTION

In vitro cultivation of *T. vaginalis* requires reducing agents for optimal growth. It has been shown that thiol blocking reagents act as potent anti-trichomonal compounds with cultured parasites (Gillin *et al.*, 1984). Lacking glutathione or an analogous molecule such as trypanothione (Ellis *et al.*, 1994), the parasite is thought to utilise cysteine as the primary thiol reductant. Despite this there is no published experimental evidence for cysteine synthesis by the parasite, either by direct sulphydration (the condensation of OAS and sulphide by cysteine synthase), or transulphuration (the conversion of homocysteine to cysteine by cystathionine β-synthase (CBS) and cystathionine γ-lyase (CGL)). Very low levels of CGL activity in crude lysates has been detected (Thong, 1985) and an "activated serine sulphydrase" activity described that has similarities to CBS (Thong and Coombs, 1985b). However neither of these have been shown to produce cysteine from homocysteine either *in vivo* or *in vitro*. The aim of this part of the study was to investigate whether the parasite has the ability to synthesise cysteine by direct sulphydrylation.

A single EST clone with substantial amino acid sequence homology to cysteine synthase was identified from *T. vaginalis*. It was discovered through single pass automated DNA sequencing of clones from a cDNA library of *T. vaginalis* strain G3. The library was produced using a ZAP-cDNA® synthesis kit (Stratagene)

from a total RNA preparation from the parasite. A large number of clones were randomly chosen and sequenced. A tBLASTn search of the databases was performed with the nucleotide sequences, whereby the cDNA sequence was converted into all 6 possible ORFs (3 forward and 3 reverse) and a BLAST search performed to find homologues. It was the result of this preliminary search which identified the clone as a possible cysteine synthase gene. In order to confirm the identity of the protein encoded by this gene it was decided first to verify the nucleotide sequence. If this confirmed that the gene possessed high homology to other cysteine synthase genes, then it was to be cloned into a prokaryotic expression vector to produce recombinant protein in *E. coli* for subsequent purification and characterisation. The expression vector selected to produce recombinant enzyme was chosen to append a hexahistidine tag to the C-terminus of the mature protein. It was hoped that this method would be able to repeat the success in expression and affinity purification of recombinant *T. vaginalis* MGL. When deciding upon the position of the hexahistidine tag, the published structure of O-acetyl serine sulfhydrylase (cysteine synthase) from *S. typhimurium* was examined. The model shows a disordered C-terminal end in the crystal structure (Burkhard *et al.*, 1998) implying that changes made to this end would have minimal effects upon the overall fold. In order to introduce the fewest changes into the cysteine synthase ORF and to position the hexahistidine tag at the C-terminus of the protein, the vector chosen was pET21a+ (Novagen). This would then allow affinity purification of the recombinant enzyme in an analogous fashion to MGL using a metal chelate chromatography method (McKie *et al.*, 1998). Basic biochemical characterisation of the enzyme could then be carried out and it would

be examined for "activated serine sulphydrylase" activity as previously identified in the parasite (Thong and Coombs, 1985b).

There are commonly multiple isoforms of cysteine synthase present in micro-organisms and plants. These may be either cytosolic or targeted to specific organelles. However only a single clone was discovered in the several hundred clones from *T. vaginalis* sequenced. It was envisaged that using either antibodies raised against the recombinant protein or nucleic acid hybridising techniques, any additional gene copies or homologues would be identified. Also with these specific probes, the role of the enzyme in the parasite could be investigated.

5.2 RESULTS

5.2.1 Sequencing and Genomic Analysis of Cysteine Synthase

5.2.1.1 Verification of the Putative Cysteine Synthase Gene

Clone 215 was provided by Dr R Hirt (Natural History Museum, London) and the results of preliminary sequencing showed homology to cysteine synthase genes in the databases. A BLAST search of the databases showed high levels of homology to identified cysteine synthase genes and low homology to cystathionine β-synthase and threonine deaminase, other members of the β-family of PLP-dependent enzymes. To confirm the sequence of the cDNA clone the plasmid insert was fully sequenced in both sense and antisense directions using automated DNA sequencing. Oligonucleotide primers designed against the

flanking plasmid sequence (M13 and T3) and internal gene specific primers (OASS2 and OASS3) were used as detailed in section 2.2.12. The sequence of the gene is given in figure 5.6.

The sequence of clone 215 showed the presence of the poly-A tail from the reverse transcribed mRNA. This confirmed the 3' end of the ORF was complete. After fully sequencing the gene in both directions an ORF of 897 nucleotides was evident. Comparison of the amino acid sequence with other cysteine synthase enzymes showed that the predicted length of 299 residues was consistent with that generally found in other organisms. A TAA stop codon was confirmed by repeated sequencing of the clone and assumed to be correct. Upstream of the putative start codon were an additional 15 nucleotides. These bases, shown below, were then taken to be the 5' untranslated region (5'-UTR) of the mRNA:

5' - CGAGTCGGCACGAGC - 3'

This very short UTR is consistent with the findings with other genes isolated from *T. vaginalis* (Quon *et al.*, 1994). However to properly identify the correct start codon and ensure a full length cDNA had been isolated, 5'-RACE was attempted. Experimental details are given in section 2.2.17. This was carried out in parallel to the production of an expression construct as specified in section 5.2.1.1. Attempts to isolate the 5'-UTR by 5'-RACE were unsuccessful though. Repeatedly, the final products of the reactions were a faint smear visible on an agarose gel. This corresponded to dsDNA products of various sizes up to approximately 1 kbp. An example of the results obtained is shown in figure 5.1.

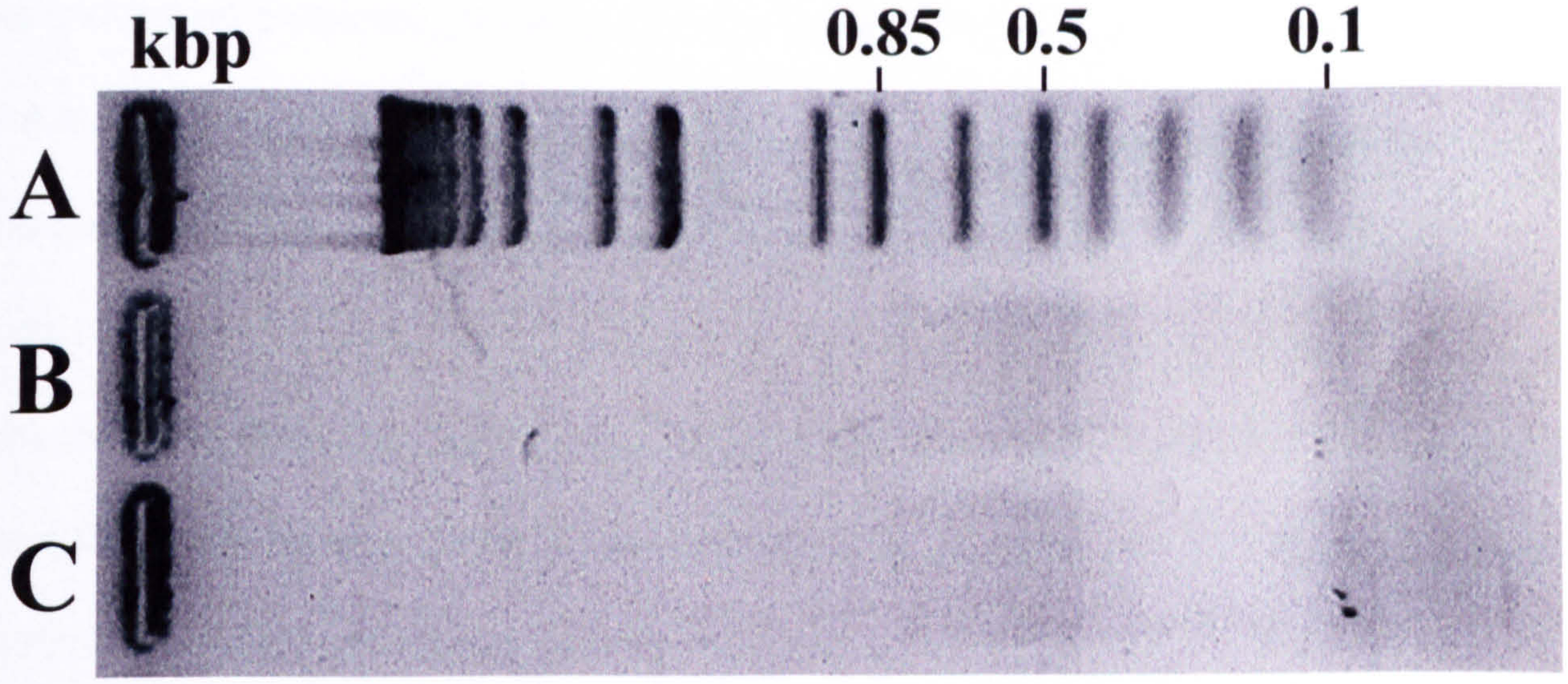

Figure 5.1. 5'-RACE products of cysteine synthase. Ethidium bromide stained 1% agarose gel showing the final DNA products. Lane A - DNA standards; Lanes B & C - Final PCR products of two separate reactions. DNA from 100 to 400 bp is partially obscured by the bromophenol blue dye present in the DNA loading buffer.

From the design of the gene specific oligonucleotide primers (OASS 4 and OASS 5), the expected size of the fragments produced was ≥ 350 bp. Control reactions indicated that the conversion of mRNA to cDNA and poly-dC tailing were successful and that failure was therefore due to PCR amplification.

5.2.1.2 Identification of the Copy Number of *T. vaginalis* Cysteine Synthase Gene

Southern blot analysis of gDNA with the cysteine synthase ORF was performed to identify the copy number of the gene, to determine if any homologues are present within the parasite genome and also to confirm the presence of the gene and rule out contamination by other organisms in the original cDNA synthesis step.

As mentioned previously there are often multiple isoforms of cysteine synthase present in an organism. Often one isoform is specifically associated with an organelle such as chloroplasts or mitochondria, with the possibility of multiple isoforms in the cytosol. It was therefore of some interest to examine the possibility that this may be the case in the parasite. It was hoped that there would be sufficient homology between any isoforms to allow a probe synthesised against the known cysteine synthase gene to be used to investigate this.

Initially restriction digests of *T. vaginalis* gDNA showed a number of distinct bands of various sizes when separated on an agarose gel. This was thought to be due to highly repetitive nucleotide sequences within the parasite genome (Wang and Wang, 1985). This banding had also been seen in the analysis of *mgl* (McKie, 1997). However by performing a second proteinase K treatment of the gDNA with the intention of removing bound proteins (for method see 2.2.4), subsequent digests produced a heterogeneous smear of DNA when separated on an agarose gel. This does not exclude the presence of highly repetitive sequences but suggests that incomplete digestion of the gDNA may have been at least partly responsible for the bands visible previously. When isolating gDNA, it was collected by spooling out the precipitated DNA with a pasteur pipette tip rather than by centrifugation. This reduced the amount of small fragments of sheared gDNA recovered which could potentially lead to spurious results.

Thus complete digests of 10 μg of gDNA were resolved by 0.7% agarose gel electrophoresis and transferred to Hybond N filters (Amersham) using the supplied protocol adapted from the method of Southern (1975). The transferred

DNA was UV cross-linked before the filters were used for hybridisation with radioactively labelled probes. The probe was synthesised and radioactively labelled by a random primer method using the Prime It™ kit (Stratagene) incorporating 50 μCi ^{32}P-dATP using a partial ORF of EST215 as the probe. This fragment was produced by performing an *Xho*I digest on the plasmid and gel purifying the resulting 900 bp fragment to be used. Only 20 bp of the gene sequence is omitted in this template and should not reduce the specificity of the probe. Hybridisation of the probe and subsequent washing of the filters was performed under high stringency conditions of 0.1 x SSC with 0.1% SDS at 65 °C. Figure 5.2 illustrates the result of Southern blot analysis of *T. vaginalis* gDNA digested with four different restriction enzymes.

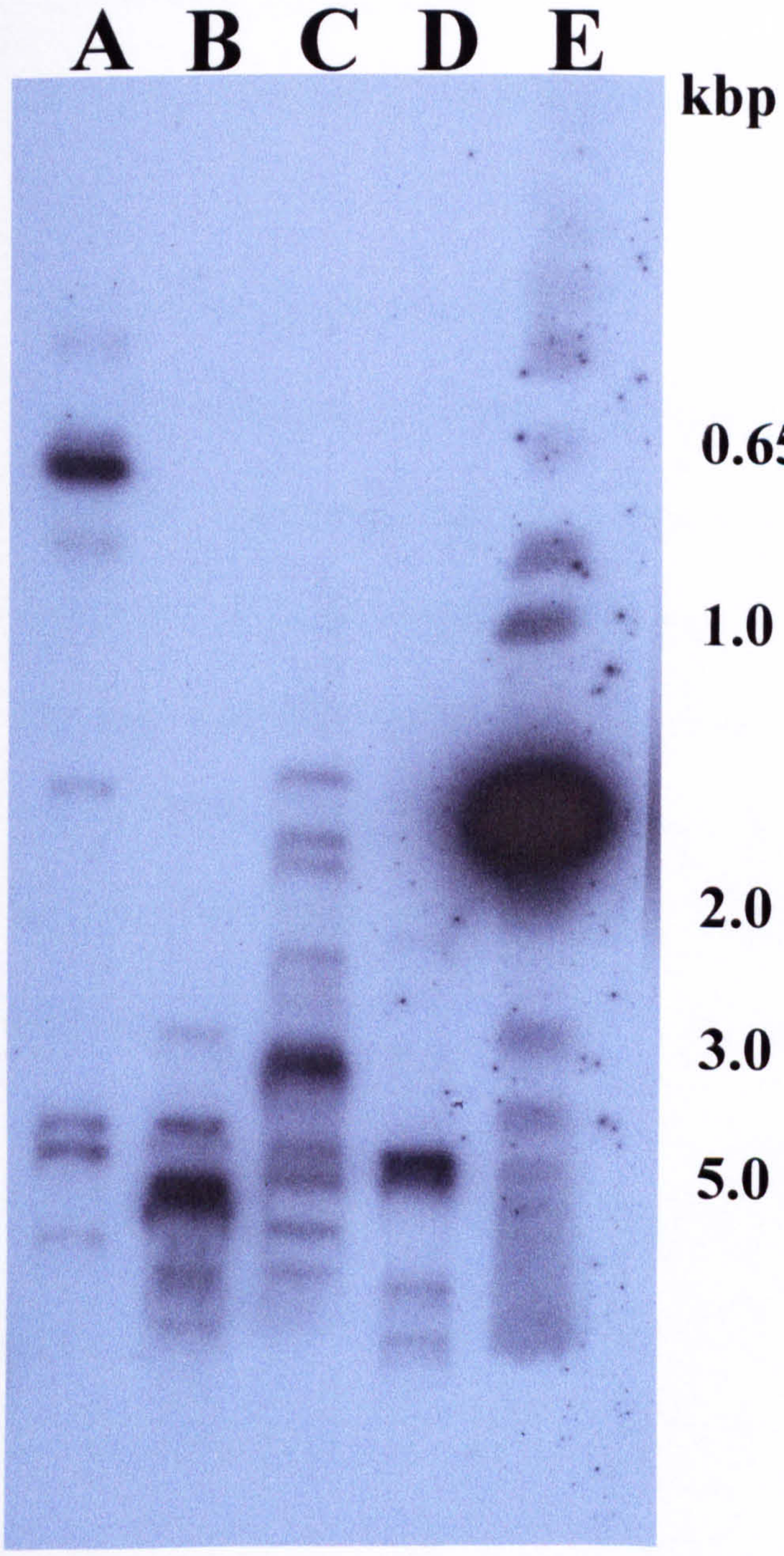

Figure 5.2: Southern blot of *T. vaginalis* genomic DNA hybridised with the 900 bp cysteine synthase specific probe. Lane A - *Hind*III digest; B - *Eco*RV digest; C - *Eco*RI digest; D - *Cla*I digest; E - DNA standards. Film was exposed for 72 hours at –70 °C.

Multiple bands are visible for each digest with a single band in each lane producing a stronger signal compared with the others. The digest by *Hind*III produces a strong signal corresponding to ~700 bp which would be obtained if the cDNA sequence was digested with the same enzyme. To identify if any of the

bands were non-specific or had arisen through incomplete digest of gDNA, the work was repeated. The result was identical banding as seen in the figure above.

5.2.1.3 Sequence Analysis

In order to identify the cDNA clone isolated from the parasite, the translated amino acid sequence was examined more closely for homology to other sequences in the database. The results of blast searches consistently indicated that the sequence was indeed a cysteine synthase. When aligned against a number of these the homology was clear (Figure 5.3). Table 5.1 gives the sequence identities and similarities of selected cysteine synthases to the *T. vaginalis* enzyme.

Organism	Identities (%)	Similarities (%)
Aquifex aoelicus	49	67
E. histolytica Type I	47	65
S. typhimurium Type B	45	64
Capsicum annuum	43	62
Spinicia oleracea	42	60
S. typhimurium Type A	42	57

Table 5.1: Sequence homologies of selected cysteine synthase enzymes with putative *T. vaginalis* cysteine synthase.

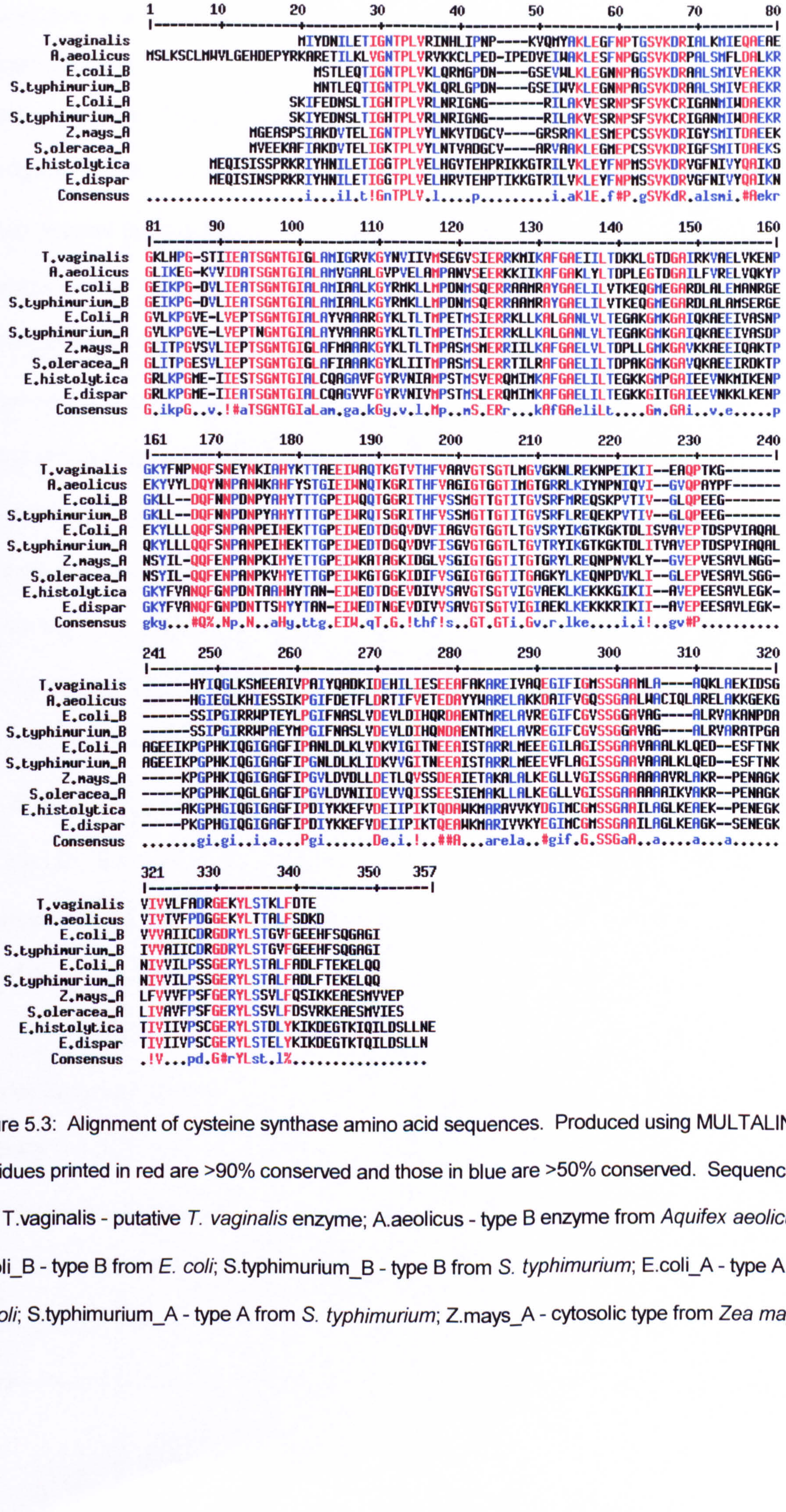

Figure 5.3: Alignment of cysteine synthase amino acid sequences. Produced using MULTALIN. Residues printed in red are >90% conserved and those in blue are >50% conserved. Sequences are: T.vaginalis - putative *T. vaginalis* enzyme; A.aeolicus - type B enzyme from *Aquifex aeolicus*; E.coli_B - type B from *E. coli*; S.typhimurium_B - type B from *S. typhimurium*; E.coli_A - type A from *E. coli*; S.typhimurium_A - type A from *S. typhimurium*; Z.mays_A - cytosolic type from *Zea mays*

(maize); S.oleracea_A - cytosolic type from *S. oleracea* (spinach); E.histolytica - type 1 from *E. histolytica*; E,dispar - type 1 from *Entamoeba dispar.*

The high degree of similarity and identity between the sequences is evident from the figure, despite the diversity of the organisms. The large number of identities would prevent the prediction of essential or important residues by sequence analysis alone. However the crystal structure of the type A enzyme from *S. typhimurium* is available for us to study (Burkhard *et al.*, 1998). As with all PLP-dependent enzymes, the cofactor is bound to the enzyme through the side chain amine group of a lysine residue. This is Lys41 in *S. typhimurium* and is unsurprisingly 100% conserved. Also 100% conserved are a number of residues involved in cofactor binding: Asn71, Thr177, Thr180 and Ser272 (*S. typhimurium* numbering). Although no substrate or inhibitor complex structures were produced the authors postulated a model for O-acetyl-L-serine binding. Of the residues predicted to interact with the substrate during the first part of the reaction five are conserved amongst cysteine synthase enzymes: Thr68, Thr72, Gln142, Phe143 and His152 (*S. typhimurium* numbering). This limited comparison with the structure would indicate that the same functional groups are present in the active site of the *T. vaginalis* enzyme.

In order to further analyse the sequence of the cysteine synthase from *T. vaginalis*, it was decided to produce a graphic comparison of a large number of cysteine synthase sequences. It was hoped that the results of this would provide more information on the relative relationships between enzymes. Such information may be helpful in postulating an evolutionary lineage for the enzyme, or more likely and helpful in this case, identify functional homology. A number of

cysteine synthase sequences were first obtained from the databases. The programme MULTALIN was then used first to produce a sequence alignment and then a plot of the relative similarities (Figure 5.4).

The plot produced by this alignment programme is a display of the "hierarchical clustering" used to produce the final alignment (Corpet, 1988). The programme uses this technique to produce accurate pairwise alignments of multiple sequences which cannot be achieved directly. Thus the plot gives an accurate representation of the similarities between sequences but there is no consideration given to the codon changes necessary for a specific residue change. Therefore the final product does not represent a phylogenetic analysis of the *T. vaginalis* enzyme but a more direct comparison of the amino acid sequences.

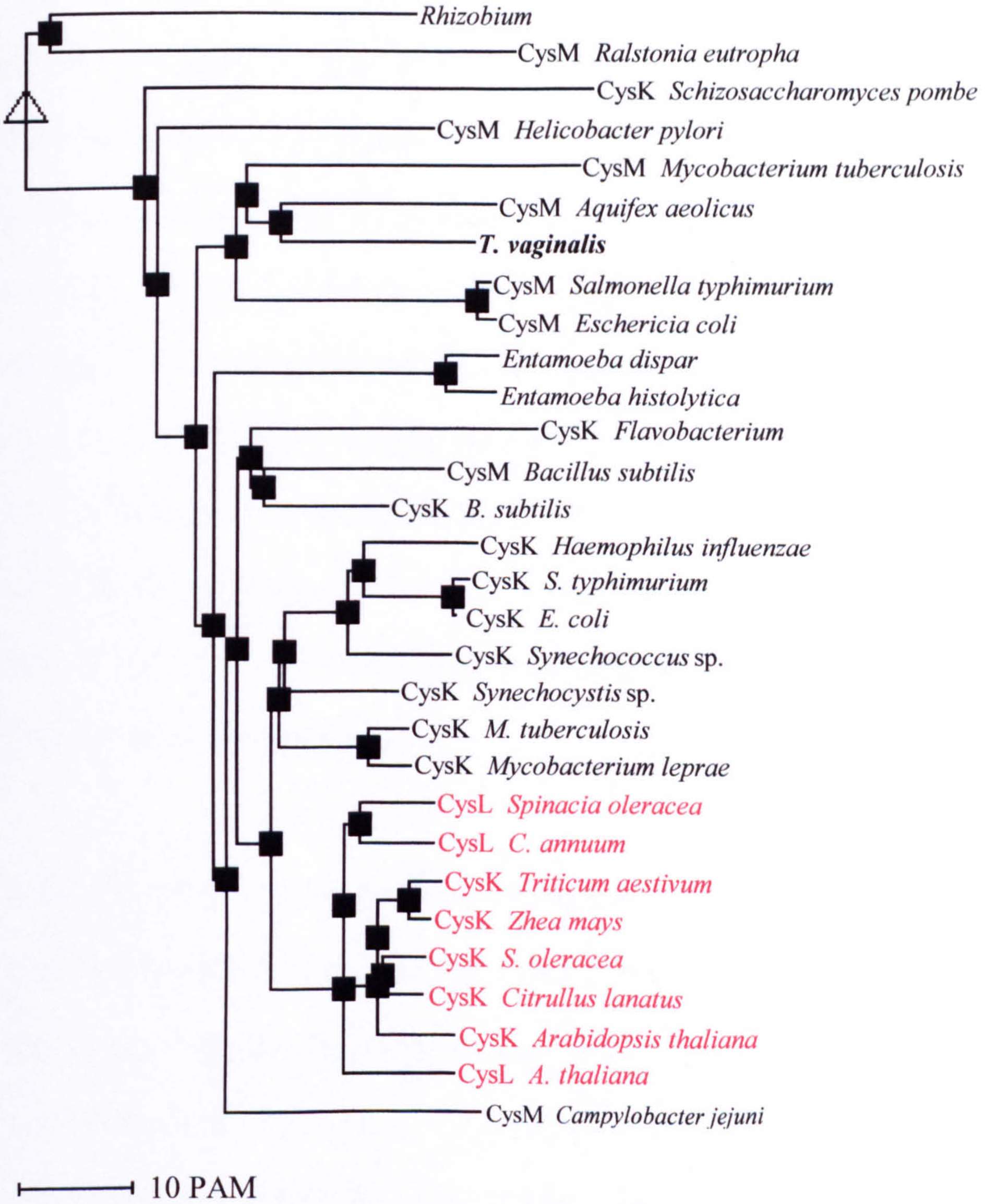

Figure 5.4: Hierarchical clustering of cysteine synthase amino acid sequences. Figure plotted using MULTALIN. Cys K - plant cytosolic or type A, Cys L - plant organelle localised, Cys M - type B (cytosolic). Enzyme nomenclature is taken from database submissions. The position of the *T. vaginalis* sequence is highlighted in bold text. All plant sequences are shown in red.

The simple analysis of a reasonable, but not exhaustive, sample of cysteine synthase enzymes from a number of organisms shows that despite the obvious sequence homology overall, the trichomonad enzyme does differ significantly from

most other eukaryotic and prokaryotic examples. Figure 5.4 shows the clear subdivision of plant sequences from all others. Bacterial sequences are more obviously divisible into subspecies, type A and B. The sequences from both *Entamoeba* species are distinguishable from each other. However the two enzymes from both species do not differ significantly and so are not separated in the figure. The *T. vaginalis* sequence shows the least divergence from a type B enzyme from *A. aeolicus*, a hyperthermophilic bacteria (Deckert *et al*, 1998). It also seems to be more similar to type B rather than type A enzymes from *E. coli* and *S. typhimurium* among others. This is an interesting observation as the type B bacterial enzymes are proposed to be preferentially expressed under anaerobic conditions in these organisms.

5.2.1.4 Cysteine Synthase Subcloning into a Prokaryotic Expression Vector

To confirm that the enzyme encoded by the nucleotide sequence was a functional cysteine synthase enzyme, it was decided to clone the gene into a plasmid vector for expression in *E. coli*. Due to the low homology seen at the N-terminus of cysteine synthase sequences, it was not possible to align sequences sufficiently to be sure that the correct start codon was being used. However whilst 5'-RACE experiments were being conducted to confirm this, it was decided to continue with production of an expression construct.

A PCR based strategy was adopted to facilitate rapid and easy cloning of the gene into the pET21a+ expression vector. This involved adding appropriate restriction sites to the 5' and 3' end of the cysteine synthase ORF via specific oligonucleotide primers. After subcloning of this PCR product into a plasmid

vector, the ORF could then be excised by restriction endonuclease digest and the fragment re-ligated into the precut expression vector. Figure 5.5 summarises the procedure.

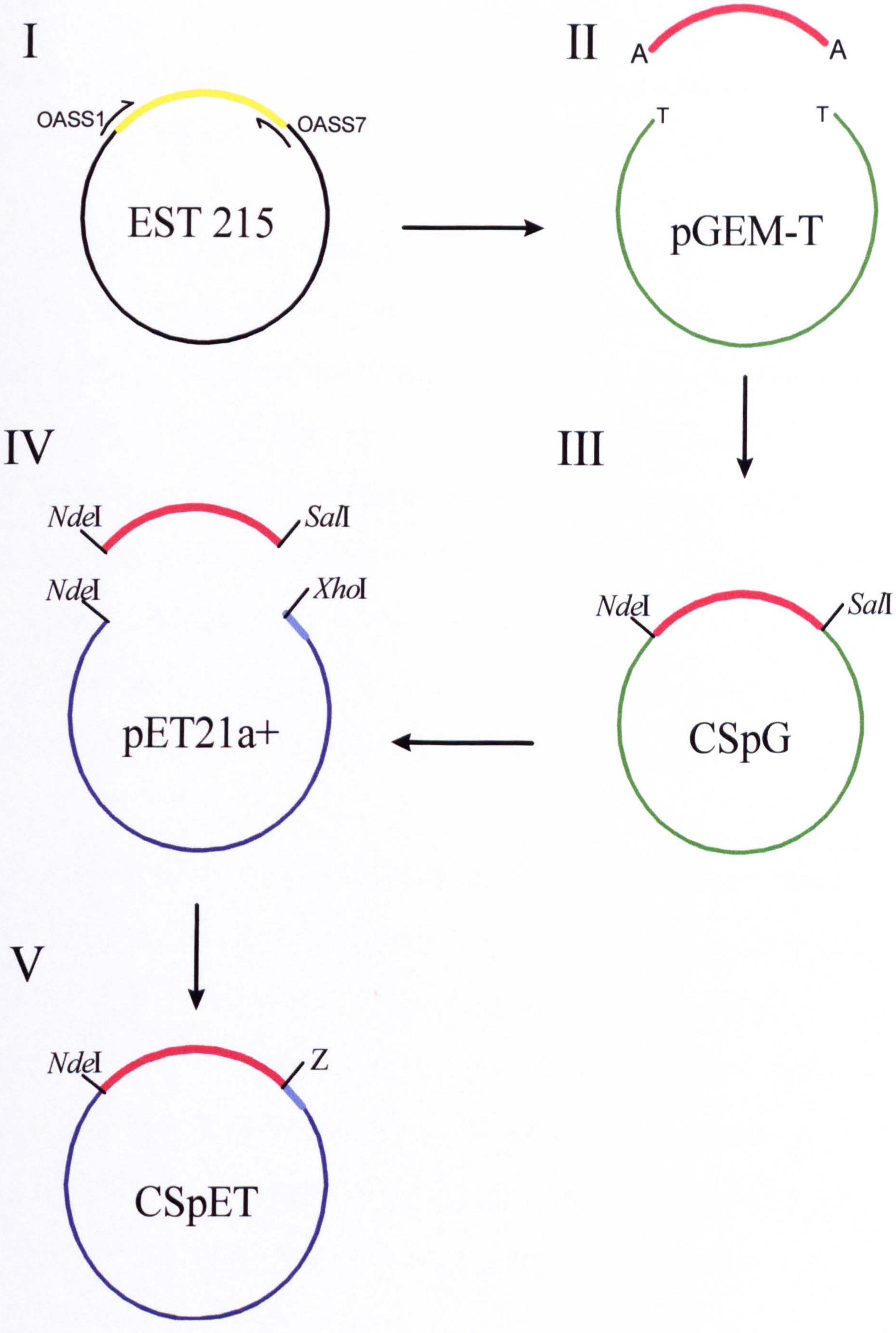

Figure 5.5: Construction of recombinant cysteine synthase expression vector. I - PCR amplification of cDNA clone EST 215; II - Ligation of PCR product into plasmid vector pGEM-T and construction of plasmid CSpG; III - *NdeI/SalI* double digest of plasmid CSpG to excise cysteine synthase ORF; IV - Ligation of fragment into pET21a+ vector predigested with *Nde*I and *Xho*I to produce expression plasmid CSpET; V - Transform plasmid CSpET into *E. coli* BL21DE3 for recombinant enzyme expression. Primers used and restriction sites are indicated. Z represents the chimeric *SalI/Xho*I restriction site, the location of the hexahistidine tag is marked in blue.

The supplied EST clone 215 was used as the template for PCR with *Taq* DNA polymerase using the oligonucleotide primers OASS1 and OASS7 as listed in section 2.2.12. Non-proof-reading polymerases such as *Taq* add an adenine base at the 3' end of the DNA product. The use of a T-tailed cloning vector makes the ligation of these PCR products easier because of a 5' complementary thymine overhang. However there is a higher likelihood of incorporating errors into the amplified fragment using *Taq* compared with thermostable polymerases which have a proof-reading ability. Therefore replicates were performed to ensure a copy without mutations incorporated was synthesised.

Four separate reactions were carried out and each was ligated into pGEM-T (Promega). Positive clones were selected after transformation into commercially supplied chemical competent XL1-Blue cells (Stratagene) by blue-white selection with X-GAL. Selected positives were confirmed by *NdeI/SalI* double restriction digest and automated DNA sequencing. The cloned PCR fragment was excised from purified plasmid by digestion with *Nde*I and *SalI* restriction enzymes and separated on a 1% agarose gel prior to gel extraction. This fragment was then ligated into *NdeI/Xho*I restricted pET21a+, to produce plasmid CSpET,

transformed into *E. coli* XL1-Blue cells and antibiotic resistant positive clones selected for by growing on LB_{AMP} plates.

The *Nde*I site at the 5' end regenerates the ATG start codon and, unlike the *Nco*I used in the cloning of *mgl*1, does not enforce any changes in the ORF (McKie *et al.*, 1998). A *Sal*I site was used at the 3' end of the gene to provide a compatible overhang to the linearised plasmid. This site was chosen in preference to an *Xho*I site because of an internal *Xho*I site 20 bp downstream of the start codon. The two restriction enzymes have different recognition sites but cleavage results in complementary overhangs. Had an *Xho*I site been used it would have required a partial *Nde*I/*Xho*I double digest to excise only the 899 bp fragment. There would have been difficulty in distinguishing between this and the smaller 879 bp fragment resulting from cleavage at the internal *Xho*I site. Ligating together these ends results in the loss of both the *Sal*I and *Xho*I sites but this was not deemed to be detrimental.

Thus the predicted expressed recombinant protein would possess two residues encoded by the chimeric *Sal*I/*Xho*I restriction site (aspartic acid and glutamic acid) and six histidine residues at the C-terminus in addition to the original sequence. Positive clones were sequenced to confirm no nucleotide errors had been incorporated during the cloning process before transformation of the plasmid into *E. coli* strain BL21DE3(Novagen) for expression. The complete nucleotide and predicted amino acid sequence of the cysteine synthase gene is shown in figure 5.6. The location of the restriction sites added in generation of the expression

construct are indicated. Amino acid differences between the original sequence and the final expression construct are shown in bold text.

```
      NdeI                        XhoI
         M  I  Y   D  N  I   L  E  T   I  G  N  T   P  L  V
1  CATATGATCT ACGACAACAT CCTCGAGACA ATCGGTAACA CACCACTCGT
   GTATACTAGA TGCTGTTGTA GGAGCTCTGT TAGCCATTGT GTGGTGAGCA

                                                    HindIII
     R  I  N   H  L  I  P   N  P  K   V  Q  M   Y  A  K
51 TCGCATCAAT CATCTCATCC CAAATCCAAA GGTCCAGATG TACGCCAAGC
   AGCGTAGTTA GTAGAGTAGG GTTTAGGTTT CCAGGTCTAC ATGCGGTTCG

   HindIII
   L  E  G  F   N  P  T   G  S  V  K   D  R  I   A  L  K
101 TTGAAGGTTT CAACCCAACA GGTTCCGTCA AGGATCGTAT CGCTCTTAAG
    AACTTCCAAA GTTGGGTTGT CCAAGGCAGT TCCTAGCATA GCGAGAATTC

    M  I  E  Q   A  E  A   E  G  K   L  H  P  G   S  T  I
151 ATGATTGAGC AGGCAGAAGC CGAAGGCAAG CTCCATCCAG GATCAACAAT
    TACTAACTCG TCCGTCTTCG GCTTCCGTTC GAGGTAGGTC CTAGTTGTTA

     I  E  A   T  S  G  N   T  G  I   G  L  A   M  I  G
201 CATCGAGGCT ACATCCGGAA ACACAGGTAT CGGCCTTGCT ATGATTGGCC
    GTAGCTCCGA TGTAGGCCTT TGTGTCCATA GCCGGAACGA TACTAACCGG

     R  V  K  G   Y  N  V   I  I  V  M   S  E  G   V  S  I
251 GTGTCAAGGG CTACAACGTC ATCATCGTTA TGAGTGAGGG CGTTTCAATC
    CACAGTTCCC GATGTTGCAG TAGTAGCAAT ACTCACTCCC GCAAAGTTAG

     E  R  R  K   M  I  K   A  F  G   A  E  I  I   L  T  D
301 GAACGCCGTA AGATGATCAA GGCCTTCGGT GCTGAAATCA TCCTTACAGA
    CTTGCGGCAT TCTACTAGTT CCGGAAGCCA CGACTTTAGT AGGAATGTCT

      K  K  L   G  T  D  G   A  I  R   K  V  A   E  L  V
351 CAAGAAGCTC GGCACAGATG GCGCCATCCG TAAGGTTGCT GAACTCGTTA
    GTTCTTCGAG CCGTGTCTAC CGCGGTAGGC ATTCCAACGA CTTGAGCAAT

     K  E  N  P   G  K  Y   F  N  P  N   Q  F  S   N  E  Y
401 AGGAGAACCC AGGCAAGTAC TTCAACCCTA ACCAGTTCTC TAACGAATAC
    TCCTCTTGGG TCCGTTCATG AAGTTGGGAT TGGTCAAGAG ATTGCTTATG

      N  K  I  A   H  Y  K   T  T  A   E  E  I  W   A  Q  T
451 AACAAGATCG CCCACTACAA GACAACAGCA GAAGAAATCT GGGCCCAGAC
    TTGTTCTAGC GGGTGATGTT CTGTTGTCGT CTTCTTTAGA CCCGGGTCTG

      K  G  T   V  T  H  F   V  A  A   V  G  T   S  G  T
501 AAAGGGAACA GTTACCCACT TCGTTGCTGC TGTCGGCACA TCCGGCACAC
    TTTCCCTTGT CAATGGGTGA AGCAACGACG ACAGCCGTGT AGGCCGTGTG

     L  M  G  V   G  K  N   L  R  E  K   N  P  E   I  K  I
551 TCATGGGTGT TGGCAAGAAC CTTCGTGAGA AGAACCCAGA AATCAAGATC
    AGTACCCACA ACCGTTCTTG GAAGCACTCT TCTTGGGTCT TTAGTTCTAG
```

```
      I  E  A  Q   P  T  K   G  H  Y   I  Q  G  L   K  S  M
601   ATTGAAGCTC AGCCAACAAA GGGCCACTAT ATCCAGGGCT TGAAGTCCAT
      TAACTTCGAG TCGGTTGTTT CCCGGTGATA TAGGTCCCGA ACTTCAGGTA

        E  E  A   I  V  P  A   I  Y  Q   A  D  K   I  D  E
651   GGAAGAAGCT ATTGTCCCAG CTATTTACCA GGCTGACAAG ATTGATGAGC
      CCTTCTTCGA TAACAGGGTC GATAAATGGT CCGACTGTTC TAACTACTCG

      H  I  L  I   E  S  E   E  A  F  A   K  A  R   E  I  V
701   ACATCCTCAT CGAGTCTGAA GAGGCATTCG CCAAGGCACG TGAAATCGTT
      TGTAGGAGTA GCTCAGACTT CTCCGTAAGC GGTTCCGTGC ACTTTAGCAA

       A  Q  E  G   I  F  I   G  M  S   S  G  A  A   M  L  A
751   GCTCAGGAAG GCATCTTCAT CGGCATGAGC TCTGGCGCTG CTATGCTCGC
      CGAGTCCTTC CGTAGAAGTA GCCGTACTCG AGACCGCGAC GATACGAGCG

                HindIII
        A  Q  K   L  A  E  K   I  D  S   G  V  I   V  V  L
801   TGCTCAGAAG CTTGCTGAGA AGATTGACAG CGGTGTCATT GTTGTTCTCT
      ACGAGTCTTC GAACGACTCT TCTAACTGTC GCCACAGTAA CAACAAGAGA

      F  A  D  R   G  E  K   Y  L  S  T   K  L  F   D  T  E
851   TCGCTGATCG CGGTGAGAAA TACCTTTCAA CAAAGCTCTT CGACACAGAA
      AGCGACTAGC GCCACTCTTT ATGGAAAGTT GTTTCGAGAA GCTGTGTCTT

      V  E  H  H   H  H  H   H  *
901   GTCGAGCACC ACCACCACCA CCACTGA
      CAGCTCGTGG TGGTGGTGGT GGTGACT
```

Figure 5.6: Nucleotide and single letter amino acid code sequence of plasmid CSpET. Changes introduced in cloning are shown in bold text.

5.2.2 Expression and Purification of Recombinant Cysteine Synthase

5.2.2.1 Test Induction and Expression of Recombinant Cysteine Synthase

Small scale test expressions of recombinant enzyme produced in *E. coli* were carried out before proceeding on to larger scale preparations. Single colonies of *E. coli* containing plasmid CSpET were obtained by streaking out cells onto ampicillin-containing LB agar plates and growing overnight at 37 °C. An overnight culture was produced by inoculating 10 ml LB with a single colony and growing overnight at 37 °C with shaking. 5 ml of this was added to 50 ml LB_{AMP} and grown at 37 °C with shaking until the OD_{600} reached 0.6-0.8. Recombinant enzyme

expression was induced by the addition of 0.2, 0.5, 1.0 and 2.0 mM IPTG. 5 ml samples were taken at hourly intervals from 0 to 4 hours post induction, harvested by centrifugation at 3200 g for 15 minutes and stored at -20 °C. Cells were lysed as detailed in section 2.3.7. The resulting supernatant was analysed under reducing conditions on 10% SDS-PAGE and stained with Coomassie blue (Figure 5.7).

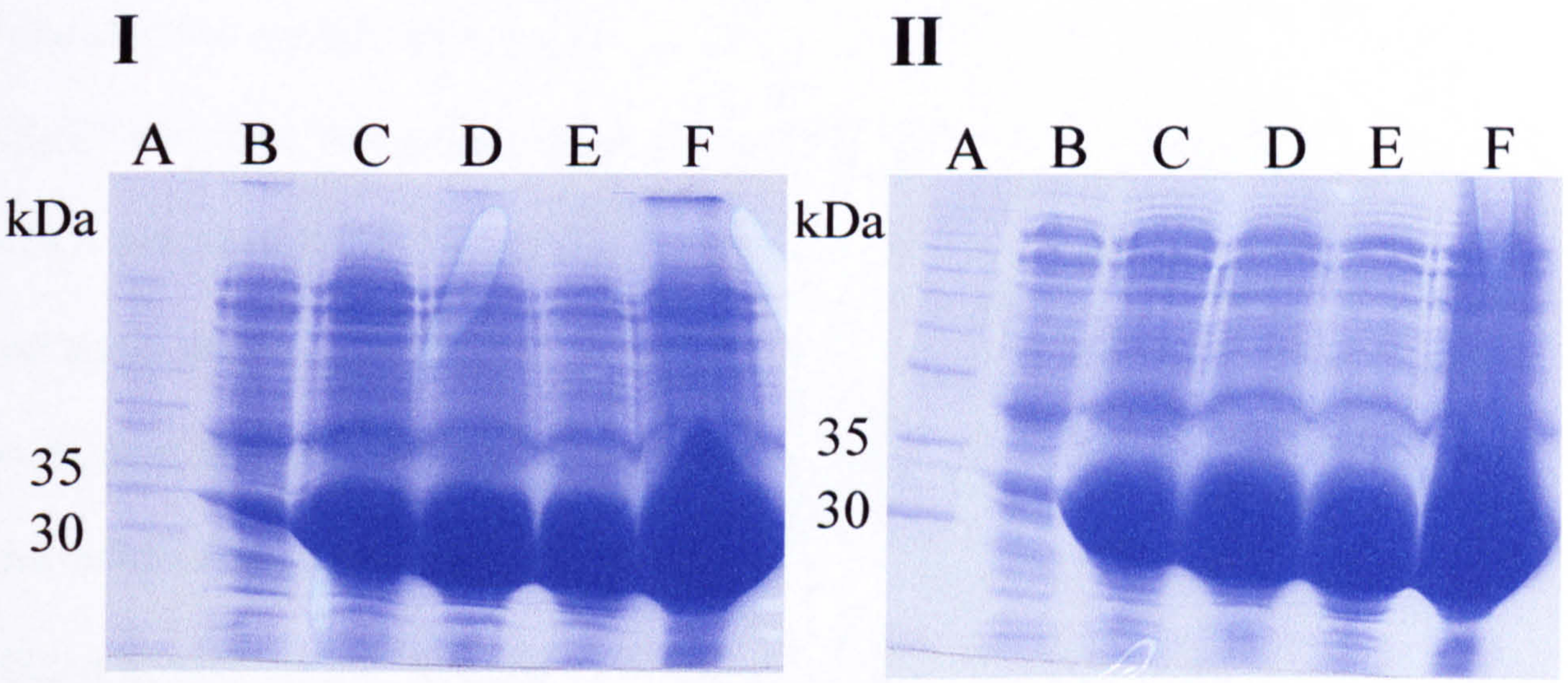

Figure 5.7: Coomassie blue stained SDS-PAGE of *E. coli* soluble proteins from test inductions of recombinant cysteine synthase expression. Image I - 0.5 mM IPTG, II - 2.0 mM IPTG. Lanes A - protein standards, B - 0 hours induction, C - 1 hour post-induction, D - 2 hours post-induction, E - 3 hours post-induction, F - 4 hours post-induction.

A prominent band of between 30 and 35 kDa was clearly visible in the non-precipitated portion of the bacterial lysate after induction. This situation was highly similar with all concentrations of IPTG (data not shown for 0.2 and 1.0 mM IPTG). There was no apparent increase in recombinant protein production after 3 hours of induction. The insoluble fraction of the *E. coli* lysate was not examined for recombinant protein due to the high levels of soluble protein evident.

The results of these experiments confirmed high level expression of a soluble protein of the expected size and provided a set of standard conditions to enable purification of the recombinant enzyme.

5.2.2.2 Routine Expression and Purification of Recombinant Cysteine Synthase

Following the conditions found above, a single *E. coli* colony containing plasmid CSpET was used to inoculate 50 ml LB_{AMP} to provide a starter culture. This was grown overnight at 37 °C with shaking and added the next day to 500 ml LB_{AMP} and again grown at 37 °C with shaking until reaching an OD_{600} of 0.6-0.8. The production of recombinant enzyme was induced by the addition of 0.5 mM IPTG. The culture was grown for a further 3 hours before harvesting by centrifugation at 4000 g at which point the cells were stored at -20 °C until required. The cell pellet was resuspended in ~5 ml 50 mM sodium phosphate pH 8.0 with 20 μM PLP and lysed by sonication at maximum amplitude with a Soniprep 1000. 10 rounds of sonication were performed to adequately lyse the cells. Each round was of 10, 1 second pulses with a 1 second pause between each pulse and a 1 minute pause between each round. The cytosolic fraction containing soluble enzyme was separated from intact cells and insoluble debris by centrifugation at 10 000 g for 30 minutes at 4 °C before clarification through a 0.22 μm syringe filter. The soluble crude lysate preparation was then loaded onto a pre-equilibrated Ni-NTA column and purified in an identical manner to that used for rMGL as detailed already in this thesis in section 3.2.1. The purification profile output is shown in

figure 5.8. Figure 5.9 shows the results of a routine preparation with samples of each eluate loaded onto SDS-PAGE for examination.

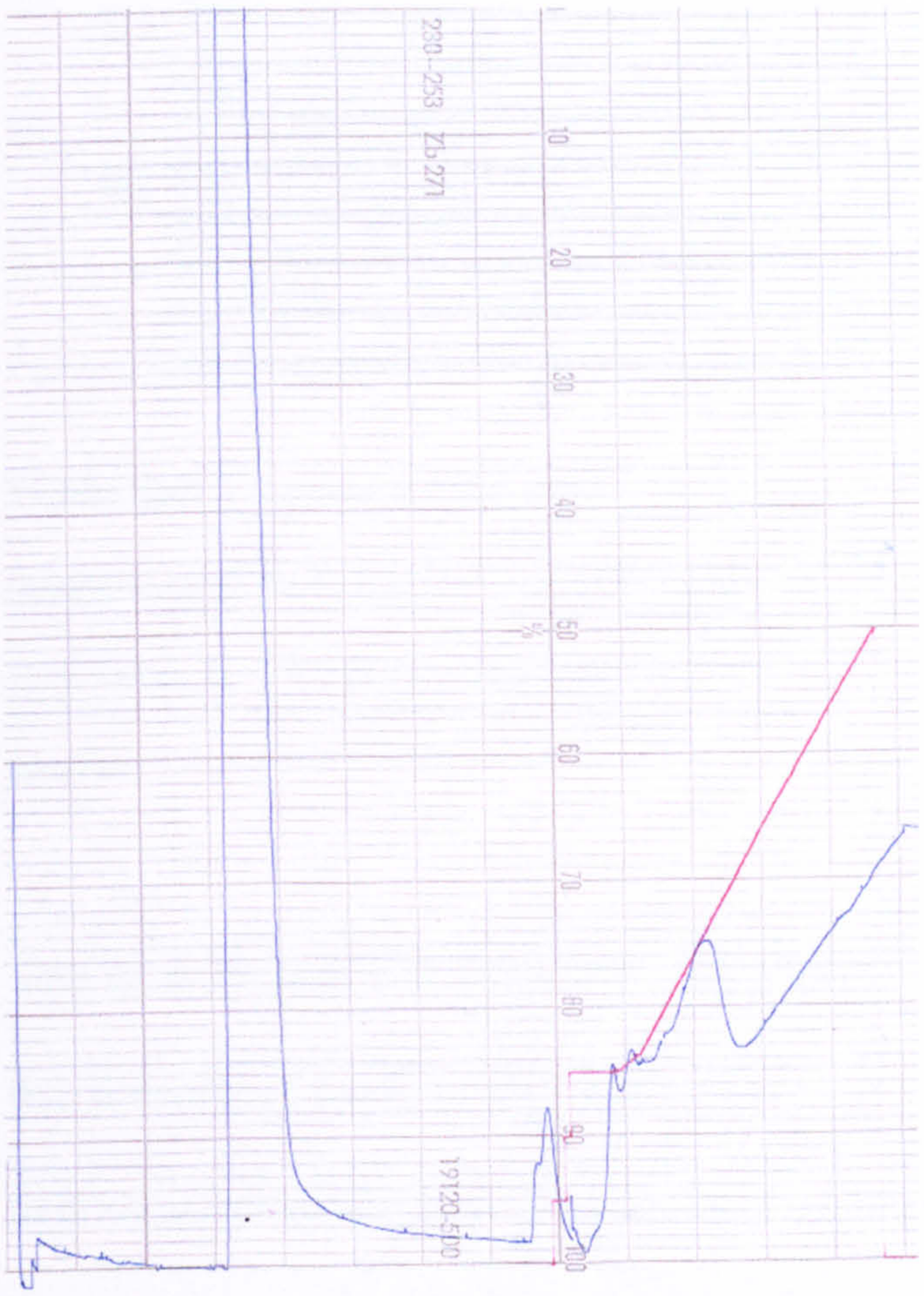

Figure 5.8: FPLC profile of recombinant cysteine synthase. The blue line represents the absorbance at a wavelength of 280 nm of the column eluate (limit of absorbance is 2.0 AU). The red line represents the proportion of imidazole containing buffer (500 mM) loaded.

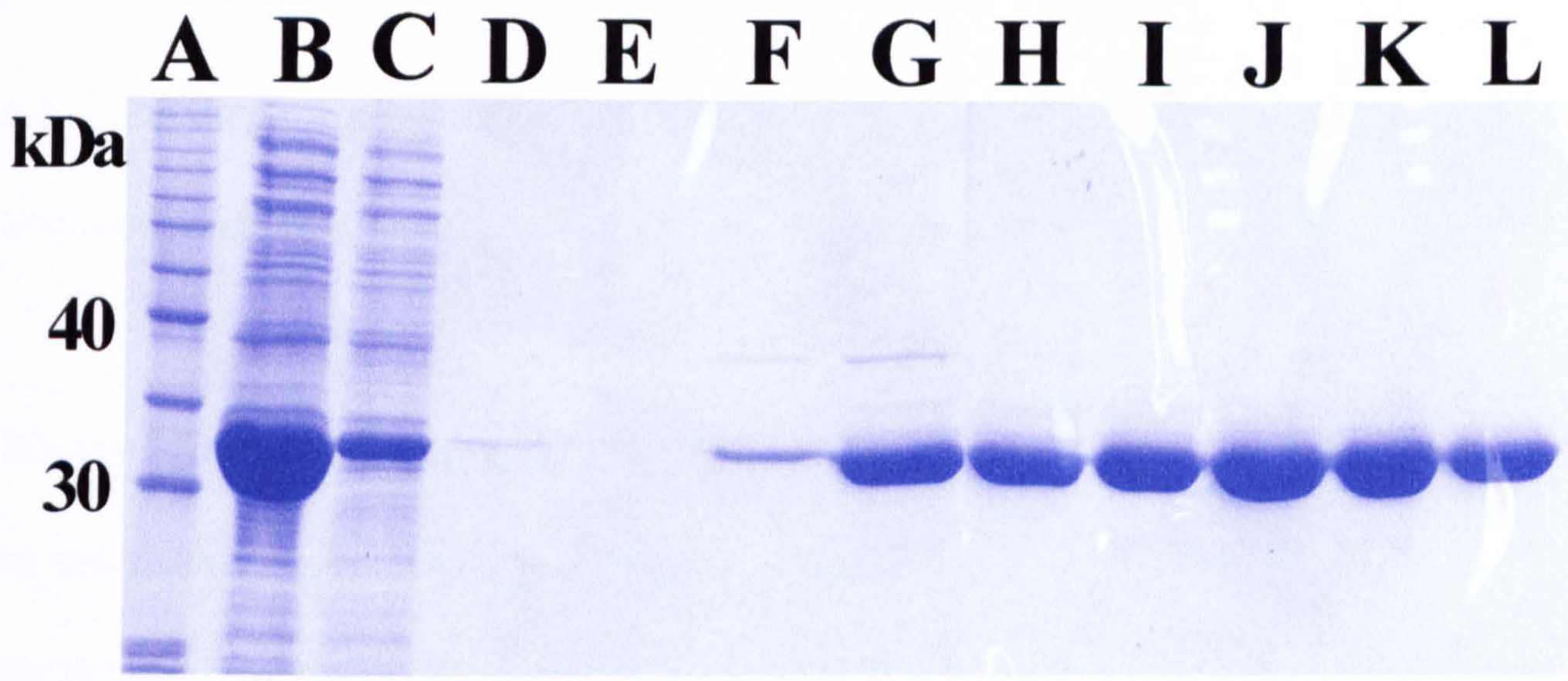

Figure 5.9: Coomassie blue stained 10 % SDS-PAGE of a standard purification of recombinant cysteine synthase. Lane A - 10 kDa ladder, B - *E. coli* supernatant, C - column flow through, D - first wash, E - 75 mM imidazole wash, F to L - successive 1 ml fractions from linear imidazole gradient.

The flow through collected when loading the column shows the great reduction in the prominent band at ~35 kDa when compared with the *E. coli* supernatant. The first wash step and the 75 mM imidazole wash step removed a large number of *E. coli* contaminating proteins. Although not easily seen in figure 5.9 this was most evident in figure 5.8 where it was shown as a sharp peak, partially obscured by the background absorbance of the imidazole containing buffer. Subsequent elution at apparent imidazole concentrations of between 90 and 130 mM produced a single major band which ran as an ~35 kDa protein in direct agreement with the predicted mass of the enzyme. Sample homogeneity was routinely checked by SDS-PAGE stained with Coomassie blue. The purest fractions were pooled and dialysed at 4 °C overnight, at a concentration of ~ 1mg/ml, into 100 mM sodium phosphate pH 7.5 with 300 mM NaCl. It has been noted for MGL that the presence of imidazole in the storage buffer has a

away. The action of imidazole as a chaotrophe is probably responsible for this and therefore the decision was taken to remove this after purification of cysteine synthase prior to storage.

This method of purification has been shown to be simple and highly effective in the preparation of recombinant MGL1. Approximately 10 mg of highly pure recombinant cysteine synthase was produced from a 500 ml culture. The cloning strategy used resulted in a recombinant protein with a small C-terminal extension which, unlike other larger constructs, is unlikely to interfere with the enzyme's activity and so should not need to be removed.

5.2.3 *In vitro* Assay of Recombinant Cysteine Synthase

Enzyme was assayed using the method detailed in section 2.3.8. The following sections describe the processes involved in finding the optimal conditions and the results of investigations into the kinetics.

5.2.3.1 Optimisation of Buffer and pH

The initial step in the optimisation of the cysteine synthase assay was to determine the buffer and pH preferences. The enzyme mechanism dictates the pH optimum but it is evident by the preferences for certain buffers in *in vitro* assays that the buffer molecule itself can have an effect. Thus the rate of the reaction was first checked using fixed levels of substrates while varying the pH between 5.5 and 9.5 with a variety of buffers. The pH optimum was apparently between 7.0 and 8.5 but it was also possible that the buffer itself played an important role (Figure 5.10).

between 7.0 and 8.5 but it was also possible that the buffer itself played an important role (Figure 5.10).

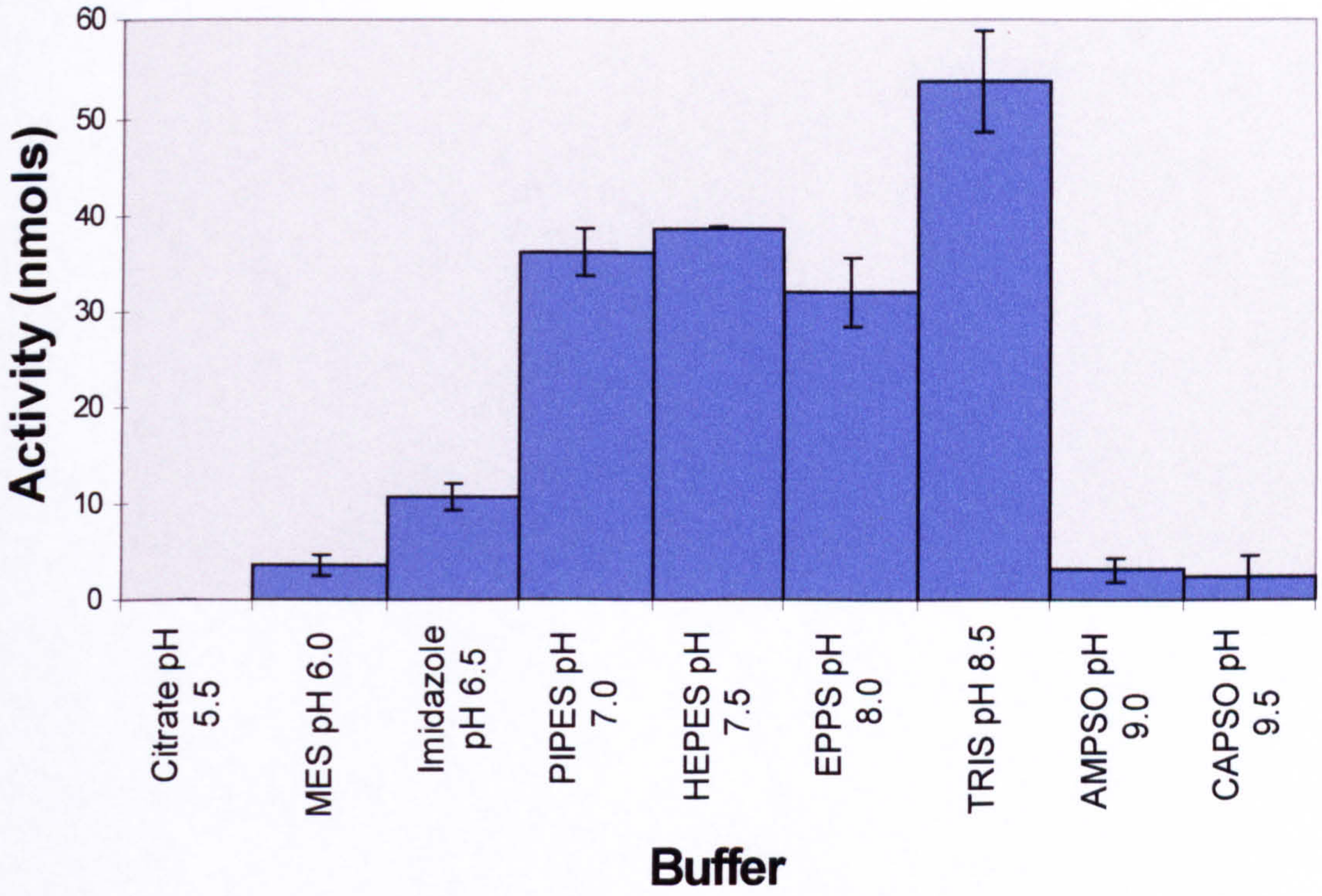

Figure 5.10: Apparent cysteine synthase activities at different pH. Buffer concentrations were 100 mM in all cases. Activities are given as nmoles of cysteine produced per reaction using identical amounts of enzyme for each assay. Activities are mean values, error bars are S.D. with n=3.

The apparent pH optimum in these experiments seemed to be around 8.5 with little activity below pH 7.0 or above 8.5. It was considered however that the buffers may affect the activity. Thus activities were assayed again using HEPES, TRIS and sodium phosphate buffers. The results showed that enzyme activity at equivalent pH was apparently lowest in HEPES with the highest activity measured in sodium phosphate (data not shown). This was then selected as the buffer of choice for the *in vitro* assay and a closer examination of the activity over a pH range undertaken. The results of these experiments showed an optimum pH with 100 mM sodium phosphate at pH 7.3 (Figure 5.11). No titration of the ionic

strength of the buffer was performed. In common with many other assays it was expected that 100 mM would provide sufficient buffering capacity and not interfere with the mechanism.

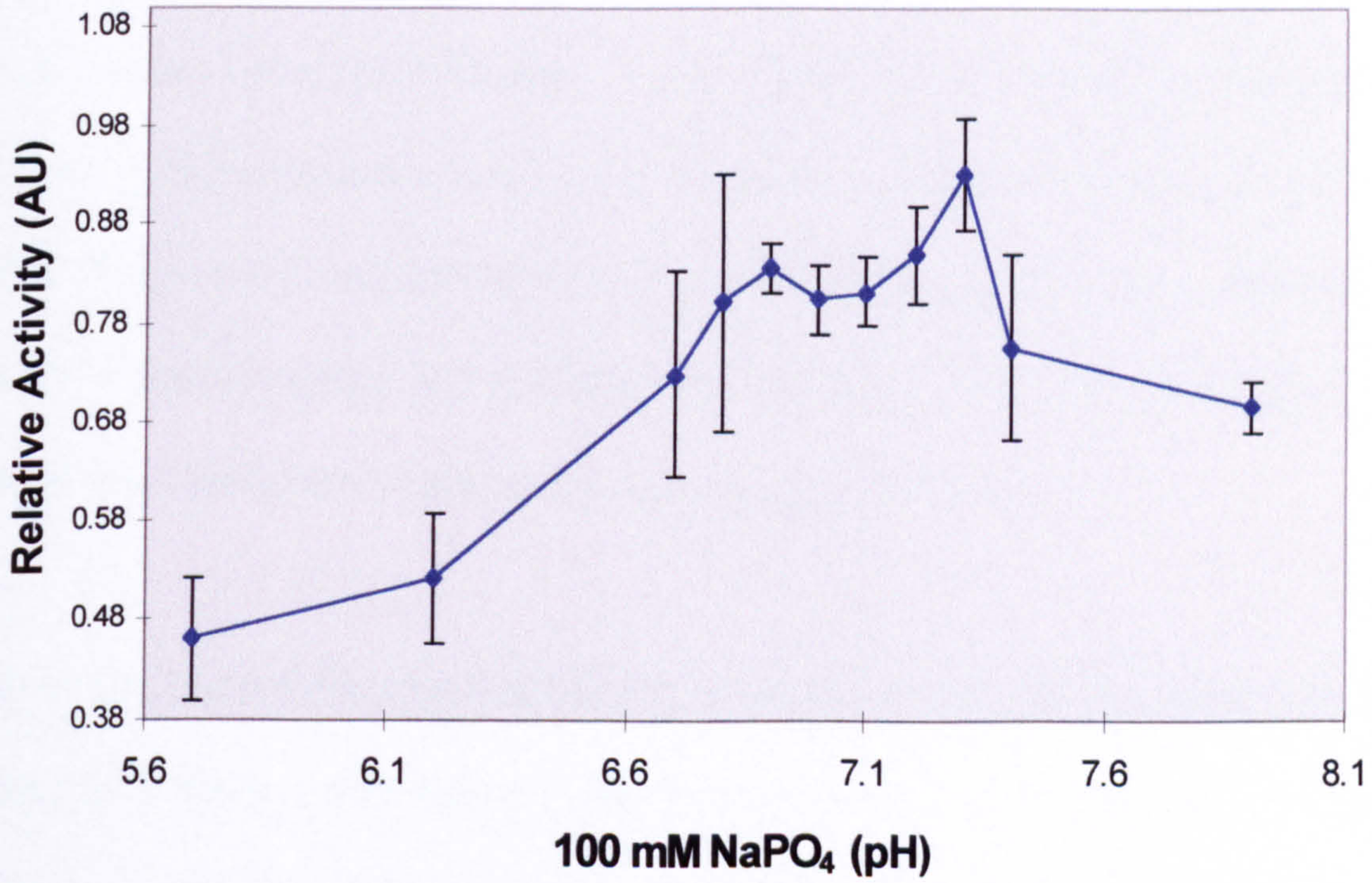

Figure 5.11: Cysteine synthase activities in sodium phosphate buffer at varying pH. Relative activities are given between samples with identical amounts of enzyme and substrate used in each assay. Activities are mean values, error bars are S.D. with n=3.

5.2.3.2 The Use of Thiols in Assaying Cysteine Synthase

Reducing agents are often added when assaying cysteine synthase *in vitro* with no explanation for their role (for example Warrilow and Hawkesford, 1998). DTT or β-mercaptoethanol are commonly used at concentrations up to 10 mM. However these compounds are also omitted from other published protocols (for example Masada *et al.*, 1975). These compounds may have a number of effects, e.g. involvement in the mechanism directly, stabilisation of the enzyme in some way or interaction with substrate or product. In the investigation it was decided to

such as a β-replacement with the thiol and OAS. This concern arises as both cysteine synthase and O-acyl homoserine sulphydrase enzymes have been shown to utilise other thiols *in vitro* such as methanethiol. Experiments were conducted with a variety of concentrations of DTT and it was found that initial results varied between replicates. It was apparent though that the presence of DTT did give an increase in the level of cysteine produced. Both substrates in the assay (Na_2S and O-acetyl-L-serine) were freshly prepared for each experiment and DTT was stored at -20 °C in single use aliquots. Thus the variability in results was unexplained simply by reagents deteriorating.

The role of a reducing agent is likely to be important to both the substrate and products of the enzyme reaction. Na_2S, the sulphur donor in the formation of cysteine, is readily oxidised and cysteine produced may dimerise to form cystine under the influence of oxygen in solution. It is assumed that DTT would be able to counteract both of these reactions. However the use of DTT in the assays still did not prevent considerable variation in the results. To further counteract the possible oxidation effects upon substrate and product, the buffer and water used in the assays was deoxygenated by "bubbling" with nitrogen gas for 15 minutes prior to use. This improved the assay greatly and gave much more repeatable results, but did not alleviate the need for exogenous thiol (as shown in Figure 5.12).

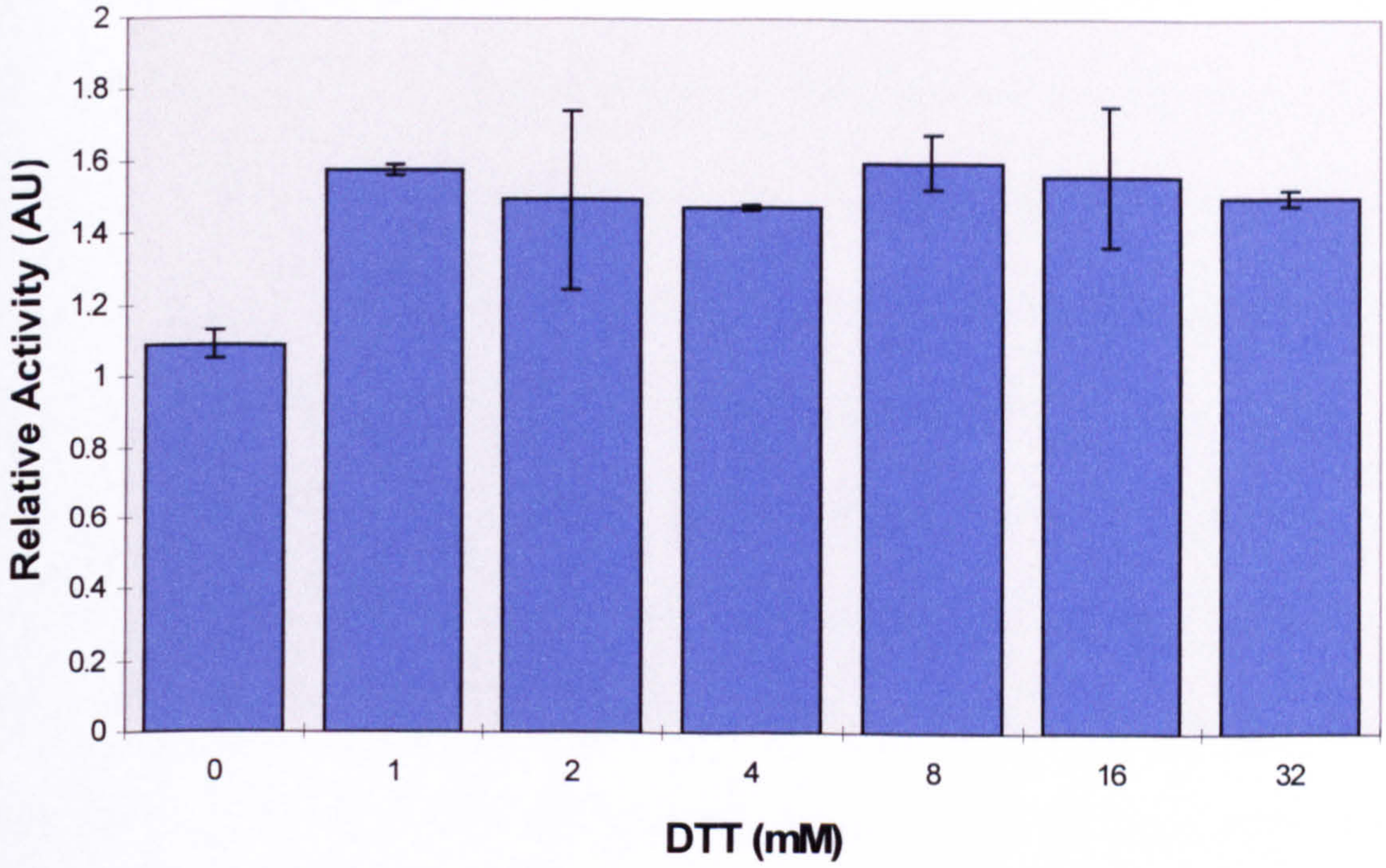

Figure 5.12: Effect of DTT on *in vitro* cysteine synthase activity. Activites are given as absorbance at 560 nm with identical levels of substrates and enzyme. Absorbance values are mean values, error bars are S.D. with n=3.

This experiment appeared to show that a concentration of 1 mM DTT was saturating for whatever its function was. The addition of higher concentrations seemed to have little further effect. There were no attempts at titrating the amount of DTT below 1 mM. At this point the standardised conditions were chosen for routine use to investigate the activity of the cysteine synthase.

5.2.3.3 "Activated Serine Sulphydrase" Activity

T. vaginalis has been shown to possess what has been termed "activated serine sulphydrase" activity due to its similarity to serine sulphydrase activity but denoting the substantial increase seen when β-mercaptoethanol (β-me) is added. The reaction produces H_2S from L-cysteine in the presence of β-me, however the

other products of the reaction were not examined. There are several possible mechanisms occurring: an αβ-elimination reaction analogous to that carried out by MGL, a NifS-like cysteine desulphurase, or a β-replacement reaction as used in the synthesis of cysteine. *T. vaginalis* does apparently possess two NifS homologues (J Tachezy, personal communication) but it is not possible to say at this time whether these are partly or wholly responsible for the "activated serine sulphydrase" activity seen. MGL does not exhibit an increase in activity toward cysteine in the presence of β-me (McKie, 1997). Neither is "activated serine sulphydrase" activity abolished by propargylglycine, a suicide inhibitor of MGL. This was investigated by substrate native PAGE as previously published (Thong and Coombs, 1987) and the results indicated that MGL is not responsible for the activity. Using the published method (see section 2.3.10), purified recombinant cysteine synthase exhibited no detectable cysteine desulphurase activity. However, in the presence of β-me, H_2S was produced at significant levels. A very brief investigation also confirmed the data from substrate native PAGE, the results are summarised in Table 5.2.

Specific activity	24.5 μmol/min/mg protein
Km (estimated)	12 mM
Vmax (estimated)	32 μmol/min/mg protein
Inhibition by DL-propargylglycine	Non-suicide/competitive
Inhibition by DL-allylglycine	Non-suicide/competitive

Table 5.2: "Activated serine sulphydrase" activity of recombinant cysteine synthase. Km and Vmax values were estimated from a limited number of datasets.

5.2.4 Analysis of Recombinant Cysteine Synthase Activity

Cysteine synthase from various sources has been investigated in great detail and the mechanism of action well characterised (for review see Tai and Cook, 2000). It was planned to examine the inhibitor profile and substrate preferences of the enzyme. The inhibitory profile was particularly interesting because of any differences that may exist between this and MGL. However, insufficient time was remaining in the project for a thorough characterisation of the recombinant cysteine synthase. The work carried out is detailed below.

5.2.4.1 Calculation of Km and Vmax for Cysteine Synthase Substrates

The primary substrates for cysteine synthase as mentioned are O-acetyl-L-serine (OAS) and a sulphur source. The *in vivo* sulphur group is likely to be hydrogen sulphide (H_2S) and the *in vitro* assay utilises commercially available sodium sulphide (Na_2S). Thus kinetic parameters were calculated for both compounds.

When setting out to calculate the kinetic parameters for the enzyme, preliminary experiments were performed to estimate the likely Km values for both substrates. This then allowed the design of experiments which would obtain all the data necessary to calculate values with a good degree of accuracy. These preliminary experiments carried out in parallel with optimisation of the assay conditions showed a likely Km of ~6 mM for OAS and ~1-2 mM for Na_2S. From these estimated values, assays were carried out for OAS and Na_2S which gave the results shown in figures 5.13 and 5.14, respectively. For the determination of OAS saturation, Na_2S was maintained at a concentration of 5 mM and for Na_2S saturation, OAS was maintained at 10 mM.

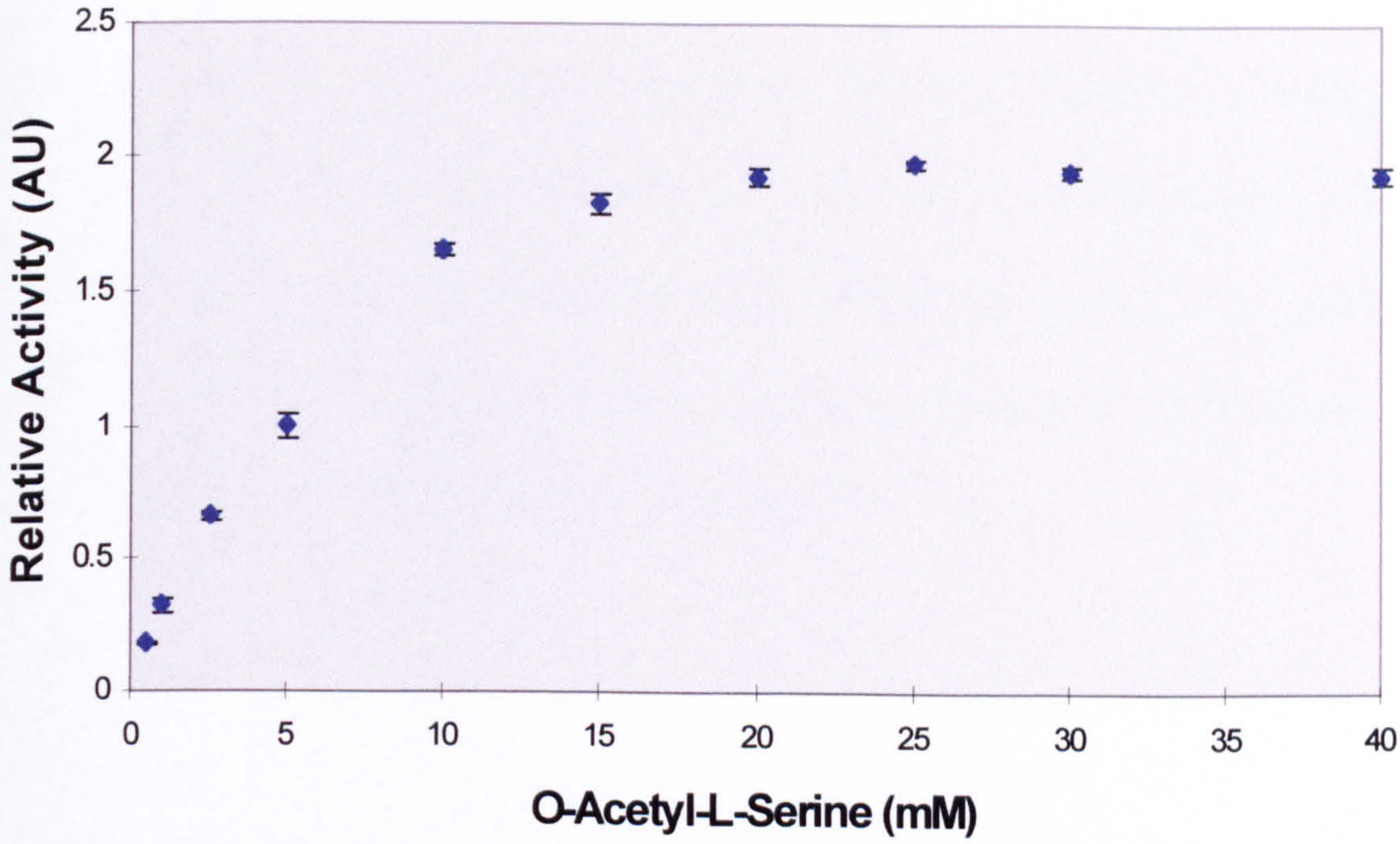

Figure 5.13: O-acetyl-L-serine saturation curve of recombinant cysteine synthase. Data points are mean values, error bars are S.D. with n=3.

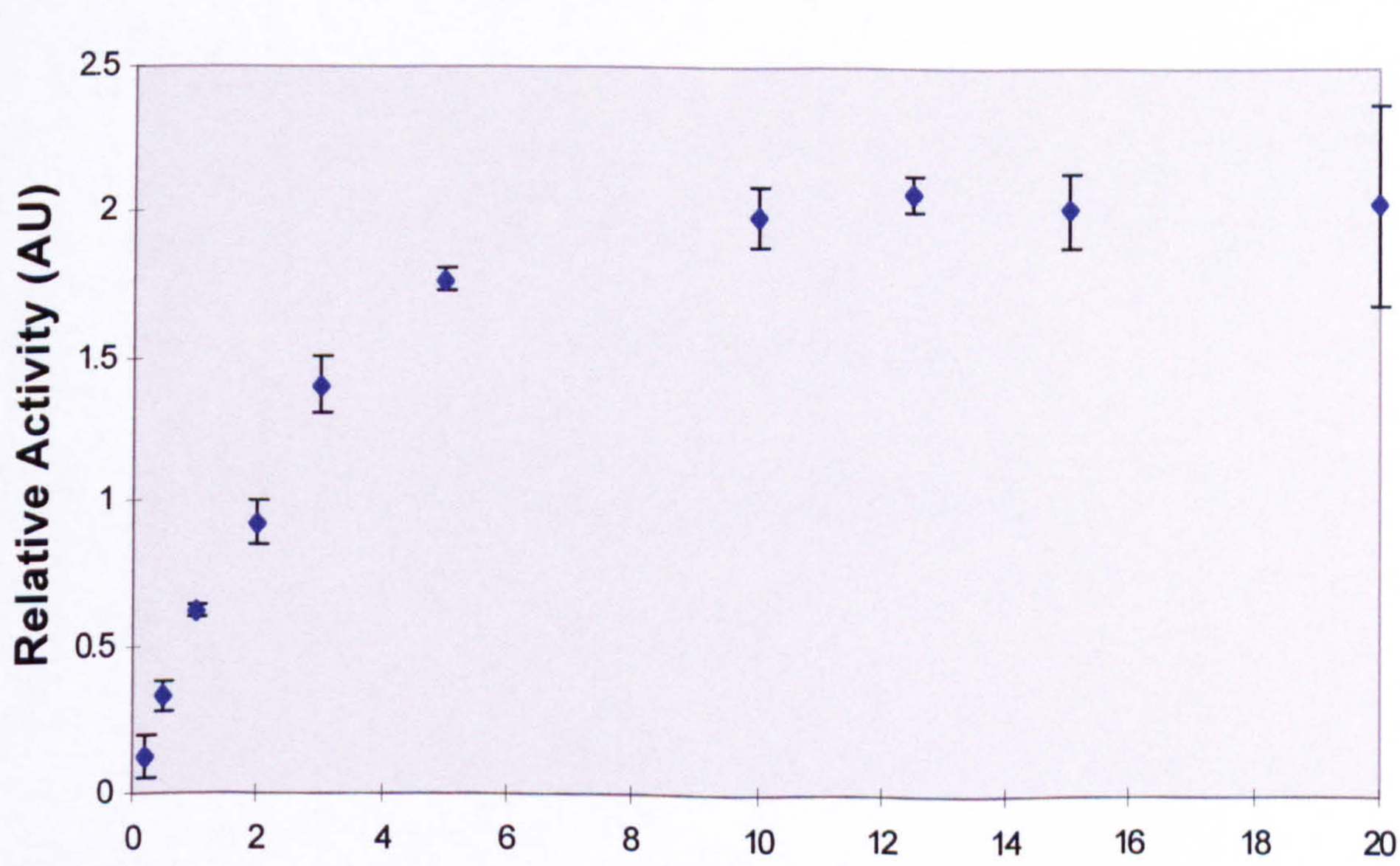

Figure 5.14: Na_2S saturation curve of recombinant cysteine synthase. Data points are mean values, error bars are S.D. with n=3.

Both results above show the hyperbolic curves commensurate with Michaelis-Menten kinetics. For both substrates it was intended to assay the enzyme in triplicate with a range of concentrations from 0.25 to 10 times the Km value. This should allow an accurate calculation of Km and Vmax values. The calculations of the Km and Vmax were performed using a Hanes plot as shown below in figures 5.15 and 5.16 for OAS and Na_2S respectively.

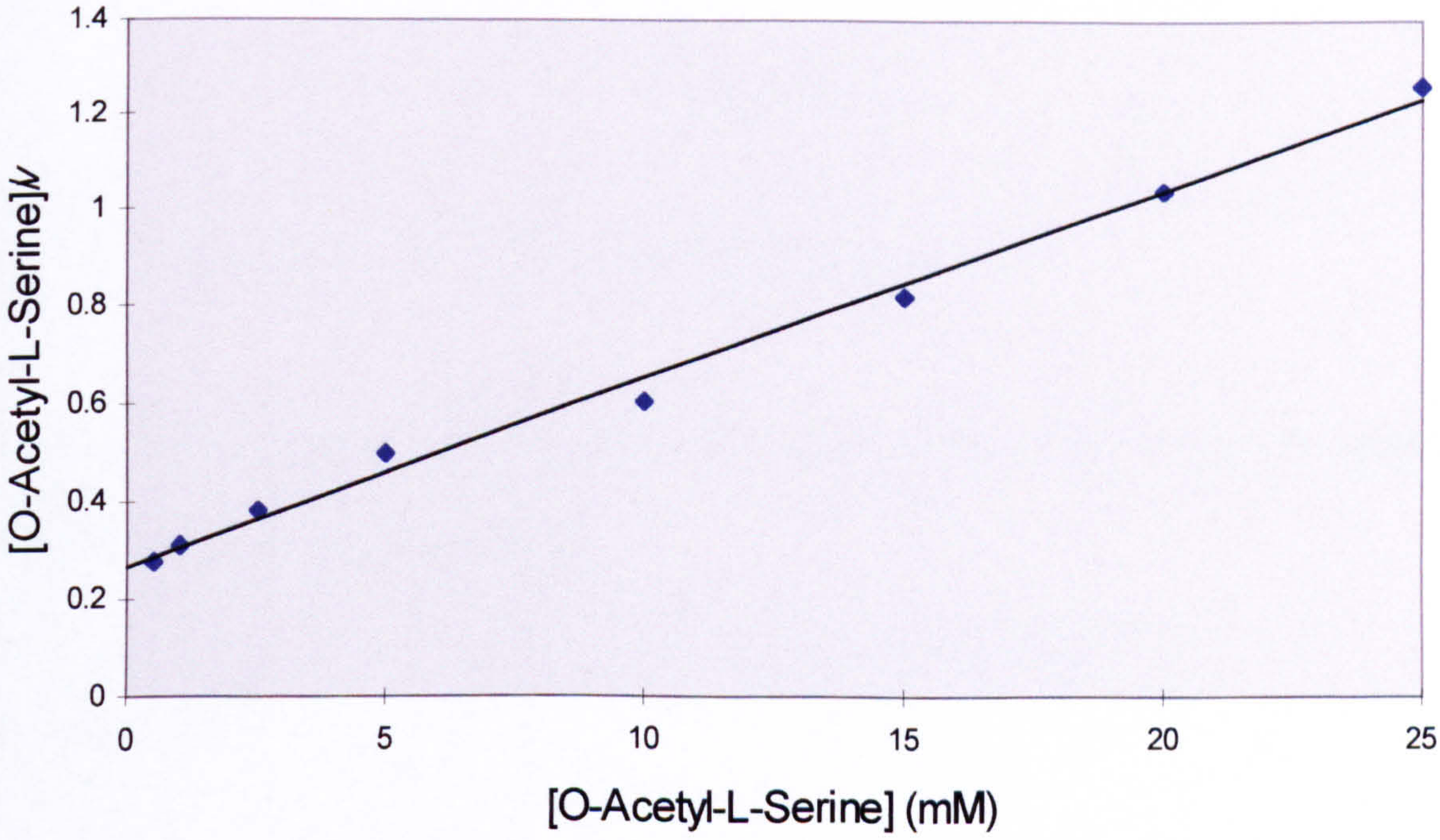

Figure 5.15: Hanes plot of OAS for recombinant cysteine synthase.

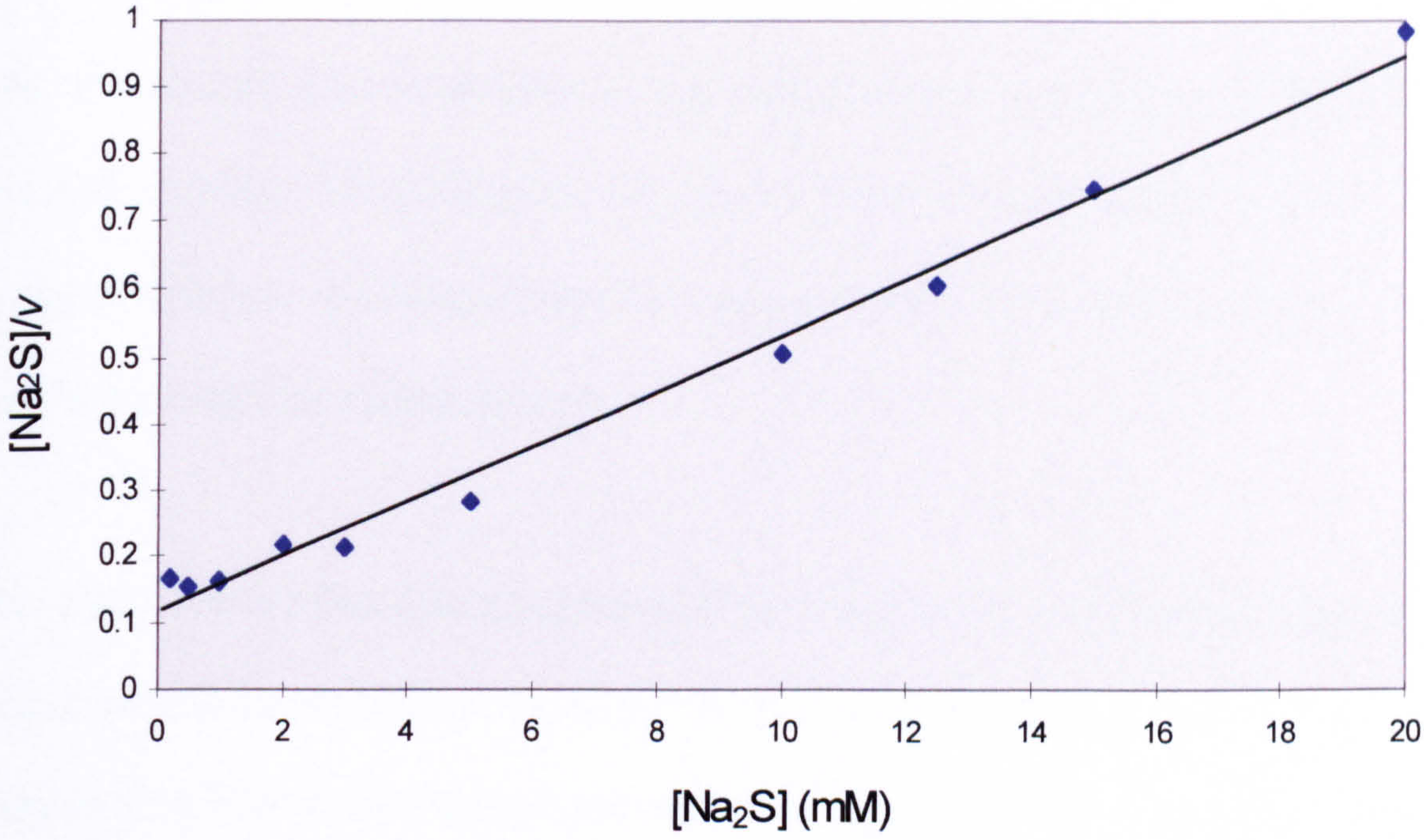

Figure 5.16: Hanes plot of Na_2S for recombinant cysteine synthase.

Thus after optimisation of the assay and investigation of the enzyme's kinetics the following results were calculated:

Specific Activity		42.7 ± 3.6 μmol/min/mg
O-Acetyl-L-Serine	Km	6.9 mM
	Vmax	61 μmol/min/mg protein
Na_2S	Km	2.8 mM
	Vmax	33 μmol/min/mg protein

Table 5.3: Kinetic parameters calculated for *in vitro* assays of recombinant cysteine synthase. Specific activity is given ± S.D. with n=11.

The enzyme was also assayed to determine if the production of cysteine using L-serine as the carbon backbone donor was possible. Assays were performed as detailed in section 2.3.9 with the substitution of L-serine for O-acetyl-L-serine.

Concentrations of Na_2S between 0.5 and 5 mM and L-serine between 1 and 20 mM respectively were examined for the production of L-cysteine with 25 μg of purified enzyme. The production of cysteine could not be detected with any of these conditions. No other O-acyl-L-serine derivatives were commercially available to test as carbon donors.

Recombinant enzyme was also assayed for inhibition by mechanism based PLP-dependent enzyme inhibitors allylglycine and propargylglycine. Both act as suicide inhibitors, irreversibly inactivating some enzymes. Assays were carried out as detailed in section 2.3.9 with varying concentrations of inhibitors. The results of the experiments are given in table 5.4.

Inhibitor Concentration	Rate (nmol/min/ml) + DL-allylglycine	Rate (nmol/min/ml) + DL-propargylglycine
0 mM	20.1 ± 1.4	21.0 ± 1.1
1.0 mM	18.6 ± 0.46 (7.5)	20.7 ± 0.36 (1.5)
5.0 mM	15.6 ± 0.98 (22)	17.5 ± 0.26 (17)
10.0 mM	12.9 ± 0.67 (36)	13.9 ± 0.25 (34)
25.0 mM	9.8 ± 0.9 (51)	11.2 ± 0.49 (47)

Table 5.4: Inhibition of cysteine synthase by allylglycine and propargyglycine. Rates are given ± S.D. with n=3. Numbers given in parenthesis are percentage activity remaining.

Neither inhibitor displayed total inactivation of the enzyme even at concentrations of 25 mM. Assuming the ratio of D to L forms of compounds is 1:1, there was still

12.5 μmoles of inhibitor present. The amount of enzyme used was approximately 20 nmoles of monomer per assay. This gives more than sufficient inhibitor for total inactivation.

5.2.5 Stabilisation Studies of Recombinant Cysteine Synthase

A recovery of around 20 mg of pure recombinant enzyme per litre of starter culture was easily obtainable. To ensure the enzyme used for these studies was in optimal condition, it was necessary to find suitable short and long term storage conditions. Following standard principles, a number of different conditions were investigated. It was clearly impossible to examine every possible condition when one considers the number of variables involves and therefore it was important to make some initial estimates of the most favourable storage conditions.

5.2.5.1 Temperature Stability

The conditions chosen reflected general principles of the best methods to store enzymes, published data on other cysteine synthase enzymes and also information gained from the previous work on MGL. Analysis began with temperature stability and then extended to buffer composition and additives. The enzyme was examined for retention of activity when stored at 4 °C, -20 °C and -70 °C with and without glycerol in identical buffer conditions. The results of this initial analysis are shown in table 5.5.

	Time (Days)					
	1	3	7	14	21	28
4 °C	100	108	102	85.6	84.4	79.6
4 °C, 40% glycerol	110	128	135	133	104	88.2
-20 °C	19.1	51.7	53.6	27.3	18.4	15.2
-20 °C, 40% glycerol	93.2	119	115	108	112	107
-70 °C	90.0	96.3	113	97.4	91.1	74.2
-70 °C, 40% glycerol	79.3	115	122	114	84.3	87.6

Table 5.5: Cysteine synthase activities recorded over time after storage under different conditions.

Data given are means from 2 experiments. Relative activities are given with results of storage at 4 °C for 1 day arbitrarily chosen as 100%.

Subsequent to purification, enzyme was dialysed overnight at 4 °C against 100 mM sodium phosphate buffer pH 7.5 with 300 mM NaCl and then combined with glycerol as appropriate. These experiments began the day following the purification of the recombinant enzyme and the activity of the enzyme stored at 4 °C was arbitrarily chosen as 100% and the relative levels to this value are shown above.

Glycerol is a very commonly used cryoprotectant for storing enzymes and it is not surprising that at -20 °C the enzyme without glycerol lost much of its activity quickly. This is almost certainly due to the formation of ice crystals and the repeated freeze-thaw action. These problems are avoided by the addition of 40% glycerol. However, the stored enzyme was not subject to the same losses in activity when stored at -70 °C without a cryoprotectant. This may be due to the rapid freezing at the lower temperatures moderating the damage done by ice

crystals. Due to time constraints, the long term effects of storage at -70 °C could not be investigated. The enzyme may behave as MGL2 does, losing more than 90% of it's activity after storage at -70 °C for longer than 12 months (unpublished data). At this point it was decided that for routine storage -20 °C with 40% glycerol added would be pursued as there was no evidence at this stage that it adversely affected the enzyme activity.

5.2.5.2 Selecting the Correct Buffer for Storage

The choice of buffer was made reflecting the published data on the activities of other cysteine synthase enzymes. Initially the pH range checked was 6.2-8.0 using the sulphonic acid derived buffers (Good *et al.*, 1966). PIPES and HEPES were used and an apparent optimum of 100 mM HEPES pH 7.2 was found. However the enzyme lost activity over a period of 2 to 3 weeks when stored at -20 °C in glycerol with these buffers. It is apparent that these buffers adversely affect the enzyme and are not suitable for storage for an unknown reason. This observation is not unusual though and highlights the need to make a suitably broad choice of buffers to sample. In the production of a new construct of MGL2, as detailed in section 2.2.16, similar experiments were carried out on the stability of the enzyme under different storage conditions (3.2.5). Here it was found that the sulphonic acid buffers were particularly favourable when compared to the large loss of activity exhibited when TRIS buffers were used. Thus the pH preferences indicated by the data were utilised, but a new buffer had to be found. Because enzyme used to monitor the effects of temperature on storage had retained greater activity while stored in sodium phosphate buffer, it was decided

to examine this. Due to a lack of time it was not possible to monitor the effect on activity of storage in a number of buffers. Instead enzyme was dialysed into 100 mM sodium phosphate pH 7.5, 300 mM NaCl, 100 μM PLP, 100 μM DTT plus glycerol to 40%. This enzyme was then used for the remainder of this work and showed no loss of activity over approximately 6 weeks.

5.2.5.3 Effects of Cofactor and Reducing Agents

The use of thiol reducing agents and cofactor to retain activity during storage was also considered. Previous work on MGL had shown that excess cofactor and reductant were required and it was decided that in an analogous fashion, cofactor and DTT would be added to cysteine synthase in order to prevent loss of activity. Thus 100 μM PLP and 100 μM DTT were added. Extensive investigations were not carried out on the merits of including these compounds because of insufficient time remaining in the project and also because initial results showed no detrimental effects of including them.

5.2.6 Western Blot Analysis of Cysteine Synthase

Polyclonal antibodies were raised in a rabbit against purified recombinant cysteine synthase. These were to be used in Western blot analysis as time permitted. Figure 5.17 shows the result of a Western blot using the antiserum against a parasite whole cell lysate and purified recombinant enzyme.

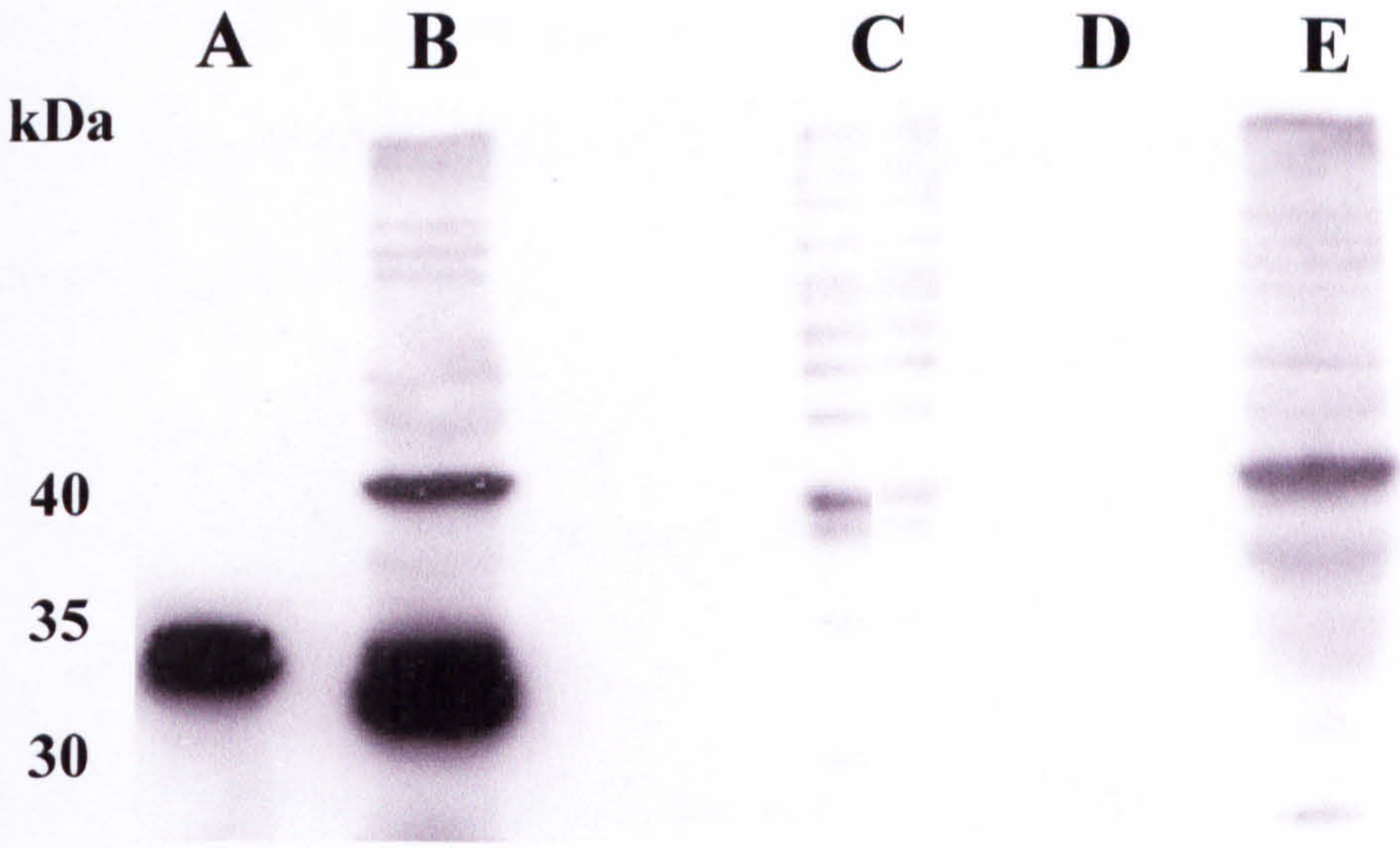

Figure 5.17: Western blot using purified cysteine synthase antiserum. Lane A - Purified recombinant enzyme probed with antiserum; B - *T. vaginalis* whole cell lysate probed with antiserum; C - 10 kDa protein standards; D - Purified recombinant enzyme probed with pre-immune sera; E - *T. vaginalis* whole cell lysate probed with pre-immune serum.

The polyclonal antiserum raised against purified recombinant cysteine synthase reacted as expected with the recombinant enzyme (Figure 5.17, lane A). This migrated to approximately 35 kDa in agreement with the predicted molecular mass of 33.8 kDa for a single recombinant enzyme monomer. There was no reaction of the pre-immune serum with the purified recombinant enzyme (lane D). The antiserum reacted with a large number of proteins in the parasite lysate (lane B) and many of these interactions are also seen in with the pre-immune serum (lane E). There were at least two additional bands detected by the immune serum (comparing lanes B and E). Assuming that the polyclonal antiserum is specific for cysteine synthase, this would indicate at least two cysteine synthase proteins or fragments present in *T. vaginalis*. The major band migrated slightly further than the recombinant enzyme (lane A) which is consistent with the predicted size of 32.7 kDa for the enzyme encoded by the isolated cDNA. However, there also

appeared to be a second, less intense band which was partially obscured by the strong signal of the major band. This appeared to be of the same size as the recombinant enzyme.

5.2.6.1 Expression of *T. vaginalis* Cysteine Synthase Under Different Growth Conditions

The parasite was cultured under different growth conditions and cysteine synthase mRNA and protein levels analysed to determine the effect of environmental factors and help elucidate the role of the enzyme in the parasite. Culturing of the parasite was performed as detailed in section 2.4.2. Parasites were harvested and washed as detailed in section 2.4.4 and cell pellets stored at -70 °C. Western blots were carried out as specified in section 2.3.4. The preparation and development of Western and Northern blots was done by Dr. G H Westrop.

To help define the role of cysteine synthase *in vivo* the parasite was cultured under various conditions. Medium was supplemented with amino acids to examine their effect on the expression of the enzyme. Amino acids chosen for this were L-cysteine, L-methionine, L-serine and DL-homocysteine. Previous work had shown a decrease in cysteine synthase activity in parasite lysates after growth in medium supplemented with 10 mM L-cysteine (P. Suchan, unpublished data). Therefore the concentration of amino acids used was 10 mM. Parasites were also cultured in conditions to mimic oxidative stress, which the parasites may experience as a result of host immune responses. Cultures were produced aerobically with exposure to normal environmental oxygen levels and in media

lacking the reducing agent ascorbate. Media was also supplemented with 0.2 mM hydrogen peroxide (H_2O_2) or 2 mM methyl viologen (MV) (a catalyst for the production of superoxide radicals). Concentrations of MV and H_2O_2 used, were decided upon by first titrating them to levels where they inhibited growth of the parasite by ~ 50% (data not shown). The parasite was also grown in the presence of 100 μM DL-propargylglycine. Previously this concentration of inhibitor has been shown to abolish MGL activity but also result in elevated levels of "activated serine sulphydrase" activity (Thong and Coombs, 1987). As the isolated cysteine synthase enzyme possesses this activity this constitutes some evidence that inhibition of MGL leads to upregulation of enzyme. The effect on growth of the parasite has been shown to be small at these concentrations of propargylglycine. The culturing of the parasites for these experiments is detailed in section 2.4.2. The numbers of parasites harvested from each of the conditions are listed below in table 5.6. These cells were used for protein and RNA isolation for subsequent Western and Northern blotting as described below.

Growth Conditions	Parasites (cells x 10^6/ml)
Control (Anaerobic)	1.44
Aerobic	3.94
Without Ascorbate	1.70
10 mM L-Cysteine	2.88
10 mM L-Methionine	1.36
10 mM L-Serine	2.80
10 mM DL-Homocysteine	1.80
100 μM DL-propargylglycine	0.52
0.2 mM H_2O_2	2.44
2.0 mM Methyl Viologen	1.40

Table 5.6: Parasite densities in modified medium. Numbers are counts from a single culture, cultures were performed in triplicate.

Western blots using whole cell lysates were performed using recombinant cysteine synthase antiserum. Equal volumes of lysate were loaded in each lane (Figure 5.18).

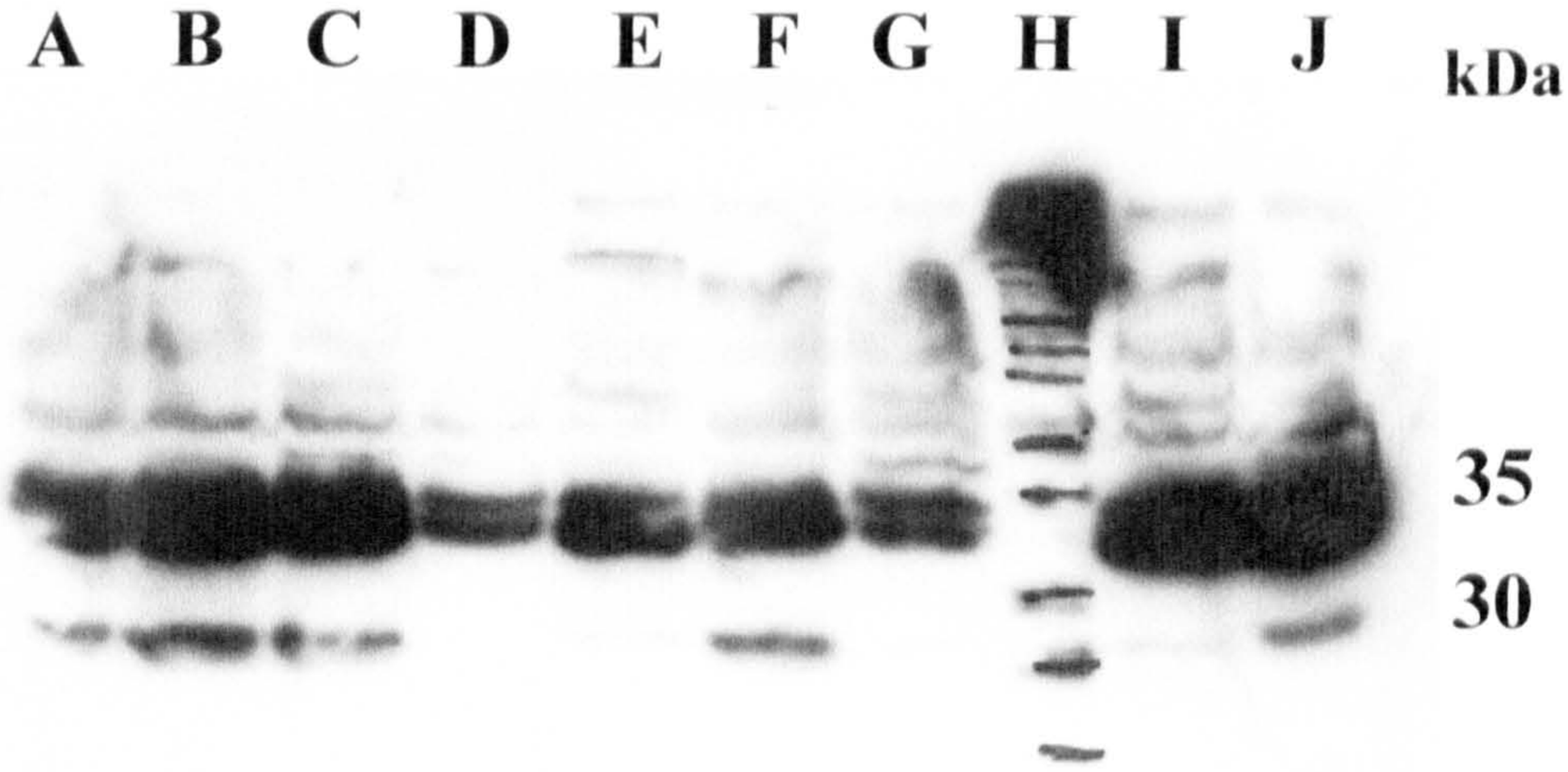

Figure 5.18: Western blot using parasites cultured under different conditions. Lanes A - Anaerobic, B - Aerobic, C- No ascorbate, D - L-Cysteine, E - L-Methionine, F - L-Serine, G - DL-Homocysteine, H - 10 kDa protein ladder, I - 0.2 mM H_2O_2, J - 2 mM MV.

Two bands were clearly visible in several lanes of the blot confirming the presence of at least two likely cysteine synthase enzymes. Due to the equal volume loading of lysates, rather than equal cell numbers, it was difficult to quantify any apparent changes in enzyme levels present. There was no apparent relative difference in the signal from cells grown in aerobic conditions and without ascorbate (lanes B and C) compared to anaerobically cultured parasites (lane A). Of the medium supplemented with amino acids the largest change appeared to be with L-cysteine where there was a significantly reduced signal (lane D). However each of lanes D to G had less intense bands than lane A after taking into account the cell densities. The parasites grown in the presence of H_2O_2 and MV (lanes I and J) both appeared to produce more enzyme as was evident from the increased intensity of bands.

To further examine the effects of culturing parasites under these different conditions, Northern blots were performed (Figure 5.19). Blots were hybridised with both a cysteine synthase specific probe and also an actin probe. Actin expression was assumed to be identical in all cells and could therefore be used to quantify RNA levels and thus loading.

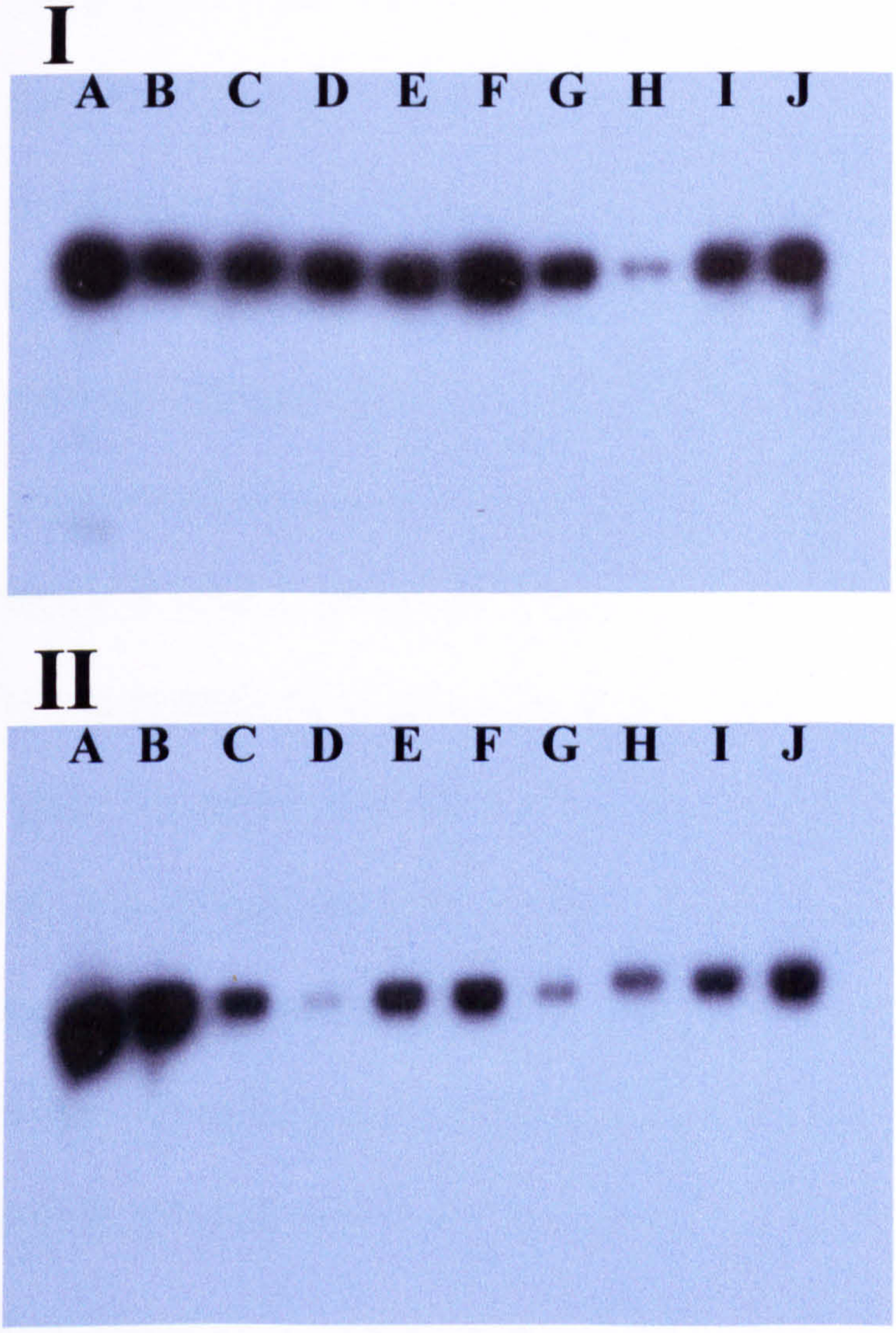

Figure 5.19: Northern blots of parasites grown under different conditions. Image I - Blot hybridised with actin probe, Image II - Blot hybridised with cysteine synthase probe. Lanes A - Anaerobic, B - Aerobic, C- No ascorbate, D - L-Cysteine, E - L-Methionine, F - L-Serine, G - DL-Homocysteine, H - 100 μM DL-Propargylglycine, I - 0.2 mM H_2O_2, J - 2 mM MV.

The hybridisation of an actin specific probe (Figure 5.19, I) showed similar levels of RNA in all lanes except H. These levels generally reflected the parasite numbers harvested from the cultures shown in table 5.6. The variation in lane H was proposed to be due to both the lower number of cells used and difficulty in isolating a sizable RNA sample (G H Westrop personal communication). A comparison of the signal intensities showed that relative to lane A there were a number of changes. The most dramatic were a fall in the signals in lanes D and G (L-cysteine and DL-homocysteine respectively) and an increase in the signal in lane H (propargylglycine). The changes observed in the other lanes were not as marked. Although the signal in both lanes I and J (H_2O_2 and MV) were much less pronounced than lane A, the level of cysteine synthase mRNA was apparently higher in lane J. Lanes E and F (methionine and serine) showed similar sized bands to each other, although again, these did seem smaller than the control lane. The effect of omitting ascorbate from the medium (lane C) was an apparent fall in the levels of cysteine synthase mRNA. The effect of cysteine synthase expression by culturing the parasites with exposure to normal atmospheric oxygen levels was difficult to distinguish from anaerobic growth. Surprisingly the cell growth was higher with oxygen present and enzyme levels seemed comparable.

5.3 DISCUSSION

A putative cysteine synthase gene from *T. vaginalis* was cloned into a prokaryotic expression vector and recombinant protein produced. The strategy adopted to do this, introduced eight additional amino acids into the mature enzyme not encoded by the original cDNA sequence. Though not ideal, it would be preferable to retain the original sequence intact, the presence of a hexahistidine tag allowed the use

of a single column chromatography step for purification. High levels of active, soluble enzyme were obtained easily which was apparently stable over a period of several weeks. The specific activity of the enzyme is below that published for many other cysteine synthase enzymes and the Km values for both substrates are higher. There is no reason though to believe that this is due to the method of expression or purification of the enzyme. With the possibility of several isoforms of enzyme present within the parasite these differences may be a reflection of a specific role.

The enzyme sequence is highly homologous to other cysteine synthases and important residues in the active site are conserved. The results of the alignment and clustering distinguished the enzyme clearly from plant species and indicated it to be more similar to bacterial type B enzymes. The type B enzymes from enteric bacteria are proposed to be expressed under anaerobic conditions and are able to utilise more reduced substrates such as thiosulphate (Nakamura *et al.*, 1983, 1984). This finding is consistent with the normal environmental growth conditions of *T. vaginalis in vivo*. However with this in mind, it may be that the lower activity exhibited by the trichomonad enzyme compared with others may be due to inappropriate assay conditions, for example the choice of sulphur donor. Deoxygenation and addition of reducing agents was highly effective in increasing activity but perhaps this is insufficient. Further biochemical investigation of the recombinant enzyme could centre around the substrate profile of the enzyme. Additionally the ability of the enzyme to utilise alternative sulphur donors may provide an insight on the likely physiological substrate.

The failure of 5' RACE experiments to confirm the validity of the proposed start codon did not hinder the production of an active recombinant enzyme. The short 5' UTR isolated with the cDNA clone was consistent with the findings from other genes from *T. vaginalis*. It does not however, possess the consensus sequence for the ubiquitous promoter element identified and analysed (Quon *et al.*, 1994; Liston and Johnson, 1999). Therefore there are concerns arising from the inability in confirming the entire gene sequence at the 5' end and thus the correct start codon. Not least of these is the possibility that a truncated enzyme may have biochemical characteristics considerably different from the full length enzyme *in vivo* (e.g. lower catalytic activity, changed substrate affinities).

The results of Southern blot analysis were consistent with the presence of a single copy of this cysteine synthase gene. The appearance of other lower intensity bands suggested, however, that *T. vaginalis* contains closely related homologues. The identical pattern seen in a repeated experiment makes non-specific interactions very unlikely, despite the obvious hybridisation to the DNA markers. This observation is a consequence of the large amount of DNA bound and is a common effect seen in this laboratory with a number of unrelated probes. The concern that the multiple bands may be present because of incomplete digestion of the gDNA prior to transfer was addressed. As mentioned in the text, efforts were made to ensure that this was not the case. The second proteinase K treatment of gDNA prior to the restriction digest abolished the discrete bands seen previously (not shown due to poor image reproduction). However the variation in the number of bands seen in the different lanes does still seem to point to the possibility that one or more of the digests is incomplete. Nevertheless

the data from the Southern and Western blot experiments indicated the likelihood of multiple cysteine synthase isozymes in *T. vaginalis*. The high stringency conditions of hybridisation and washing should have prevented any low affinity binding and implied that these proposed homologues of the cysteine synthase were likely to have considerable sequence similarity. It is unlikely that the probe hybridised with PLP-dependent enzymes from the α or γ families, as the sequences are typically quite diverse. Also the position of the bands did not correspond to those identified for *mgl*1 and *mgl*2 (McKie *et al.*, 1998) and so hybridisation to these can be excluded. The identity between these proposed cysteine synthase homologue sequences is likely to be greater than 70% as no cross reaction is seen between *mgl*1 and *mgl*2 which share this identity. As was described in section 5.2.1.3, the similarities between cysteine synthase sequences can be very high and so this allows for the possibility of very similar isoforms.

The polyclonal antisera produced against recombinant cysteine synthase reacts strongly with the purified enzyme and also with at least two proteins in the parasite lysates. The apparent size of the strongest signal in the parasite lysate corresponded to the predicted size of the native enzyme from the cDNA sequence. Additionally, although unclear, there was a second species of a slightly larger size also identifiable in the lysate. This was most obvious in figure 5.17 where two bands were clearly seen in lanes D, E and G. Assuming that there is no cross reactivity with other non-cysteine synthase enzymes, then this indicated the likelihood of at least one other isoform. It is also possible, however, that this band represented an immature form of the enzyme with a targeting signal

sequence still attached to the mature protein. The visualisation of two bands was then shown independently in Western blots using parasites grown under various conditions (Figure 5.18). In both blots there was a high level of background. This was a consequence of the large amount of immune sera used in primary antibody binding. Dot blots had been used to estimate the primary antibody dilution required, suggesting a dilution of 1:4000 was sufficient, but it is likely that a higher dilution is likely to be optimal. There was insufficient time remaining to titrate these levels.

The results of culturing parasites with different stimuli was quite successful. The parasites exhibited varying levels of expression which was evident in mRNA and protein levels. The Western blot results could be improved significantly by eliminating much of the background with lower amounts of immune sera. The success of this technique though should be extended to examine other proteins (specifically MGL) and other stimuli. The results and possible future work will be discussed in the final chapter.

During the course of this study, additional work by others identified cysteine synthase activity in parasite lysates confirming the presence of at least one functional enzyme. The work also showed that after subcellular fractionation, activity was localised to the cytosolic and hydrogenosomal fractions of the parasite (P Suchan, unpublished). Western blots have now confirmed the presence of at least two isoforms of cysteine synthase in *T. vaginalis*. It has also been seen previously that "activated serine sulphydrase" activity is attributable to multiple proteins (Thong and Coombs, 1987). Recombinant cysteine synthase

was shown to have this activity and thus presumably the enzyme accounts for two of the proteins responsible. Along with the continued attempts to identify the 5' UTR of the cysteine synthase gene it would be desirable to attempt to clone this second cysteine synthase gene in the future. Other work that could be carried out would include subcellular localisation of enzyme activity by immunomicroscopy. This may prove difficult though, if it is not possible to distinguish between the two enzymes using the polyclonal antisera.

CHAPTER 6 GENERAL DISCUSSION

6.1 INTRODUCTION

The overall aim of this study was to investigate sulphur amino acid metabolism in *T. vaginalis*. A variety of approaches were employed to achieve this. Discussions upon the experimental results obtained have been given in relevant chapters. In this chapter it is intended to relate these results and other data to the situation *in vivo*.

Prior to this study, with little, if any, experimental data available, it was possible to suggest a number of fates for the products of MGL's action and thus roles for the enzyme *in vivo*. These possible functions included:

1. The production of ATP from the α-keto acids, as detailed in section 1.6.3.3
2. The use of ammonia for combating the low pH of the vagina
3. Utilising sulphide for deoxygenation of the environment via conversion to sulphate
4. The use of free thiols as reducing agents or free radical scavengers in the cell or to cause damage to host cells as an immune evasion strategy or method to obtain extra nutrients.

There is some experimental evidence that the α-keto acid product of homocysteine and methionine catabolism, 2-oxobutyrate, can be utilised to generate ATP (GH Coombs, unpublished). The contribution of this energy generation pathway may be very different though for *in vivo* and *in vitro* growth. It

would seem most likely that the α-keto acid backbone can be used successfully to produce ATP when necessary, but is not essential and probably not utilised when more abundant energy sources, i.e. sugars such as glucose, are available.

There is, at present, no experimental evidence for the use of either NH_4^+ or H_2S to modify the extracellular environment beneficially. The parasite is sensitive to oxygen though (Paget and Lloyd, 1990) and so the possible use of H_2S in this way is still worthy of investigation.

The use of the thiols, released by the action of MGL, is probably fundamental to the enzyme's role in the parasite. The abundance of these thiols in cultured parasite preparations is immediately evident by the smell. Experiments using radiolabelled ^{35}S compounds have shown little specific incorporation of the isotope within the parasite (Thong, 1985). The measurement of thiols released by parasites cultured in complete MDM has revealed high quantities of CH_3SH released but no detectable H_2S (Thong *et al.*, 1987). CH_3SH levels were elevated when medium was supplemented with L-methionine but no H_2S was detected even when high levels of DL-homocysteine were added. Methionine could be substituted for compounds such as methyl-L-cysteine and L-ethionine and similar levels of CH_3SH were released. This implies that H_2S, the thiol product of MGL catabolism of homocysteine, does not escape the parasite cell. However, the detection of the two thiols was achieved differently, CH_3SH using DTNB (5,5'-dithio-bis-2-nitrobenzoic acid) and H_2S with lead acetate. It may be that the lack of any detectable of H_2S was due to the compound rapidly oxidising or dimerising. This may not have reacted with lead acetate to form lead sulphide and thus no

H_2S would have been detected. CH_3SH rapidly dimerises to form dimethyl sulphide ($CH_3S_2CH_3$) which loses its disulphide upon reaction with DTNB and is thus detected. Thiols such as H_2S and CH_3SH are both active reducing agents and toxic in high concentrations, but with no experimental evidence for their use in the parasite, possible functions remain speculative.

It has been well established through the use of irreversible inhibitors, that MGL is not essential for growth *in vitro* (Thong and Coombs, 1987). The growth rate is reduced at high concentrations of propargylglycine, but there are no morphological changes. However, at this point it is well worth considering the nature of propagating parasites in sterile cultures. Such parasites are provided with all the nutrients required for apparent optimum growth and simultaneously exposed to minimal stress. This is undoubtedly significantly different from the conditions found within the host's uro-genitary tract. For example, *in vivo* the parasite will be exposed to damaging host immune responses and the nutritional demands may not be so readily met. In this situation, MGL may play a critical role and therefore be suitable for investigation as a possible target for chemotherapy.

6.1.1 MGL in Iron-Sulphur (Fe-S) Cluster Production

With biochemical information and the structure of MGL1 available, we are able to postulate a catalytic mechanism for the enzyme (see section 3.2.10). The structure also revealed a possible role for the enzyme *in vivo*. The position of Cys113 in the active site of MGL1 is very similar to that of a cysteine residue found in NifS-like enzymes, whose role *in vivo* is the production and repair of Fe-S clusters. Iron-containing proteins play central roles in *T. vaginalis* metabolism (for

example, pyruvate:ferredoxin oxidoreductase, ferredoxin and hydrogenase), but there is as yet no experimental evidence on how Fe-S clusters are produced. There is an unconfirmed report that there are 2 *nifS* gene homologues in the parasite (J. Tachezy, personal communication), but MGL could nevertheless also have a role. As there are established *in vitro* assays for NifS enzymes measuring the production of functional Fe-S clusters (Zheng *et al.*, 1993), it should be straightforward to verify whether MGL has this ability.

6.1.2 Sulphydration and Transsulphuration in *T. vaginalis*

The results presented in chapter 5 of this thesis confirm the presence of a functional cysteine synthase enzyme in *T. vaginalis*. Despite the presence of direct sulphydration in the form of cysteine synthesis, there remains no firm evidence for a functional transsulphuration pathway. The previously reported low CGL activity in the parasite was detected in whole cell lysates (Thong, 1985) and no homogenous protein preparation or gene has been isolated. Indeed, the isolation of *mgl* genes utilised degenerate oligonucleotide PCR with primers designed to highly conserved regions of CGL sequences (McKie, 1997). As no CGL (or CGS, CBL or CBS) homologue was isolated using this approach or identified from the sequencing of cDNA clones, this points to the absence of these enzymes in the parasite.

This proposed lack of CGL in *T. vaginalis* (along with the presence of 2 *mgl* genes), shows a further similarity to the situation in *P. putida*. In this bacteria, MGL is postulated to have an alternative pathway to the standard reverse transsulphuration pathway (conversion of homocysteine to cysteine via

cystathionine) (Vermeij and Kertesz, 1999). In this hypothesis, the breakdown of homocysteine by MGL yields H_2S which is the substrate for cysteine synthase. In addition, CH_3SH from the catabolism of methionine is proposed to be converted by an unknown process to H_2S and so also to provide the thiol group for cysteine synthesis. The results of culturing parasites with propargylglycine described in this thesis are explainable in terms of this proposal. Inhibition of MGL leads to an increased expression of cysteine synthase, presumably in an attempt to compensate for the lack of thiol donors available and so maintain cellular cysteine levels. The hypothesis is also consistent with the results of culturing the parasites in excess cysteine; where there was a decrease in the expression of cysteine synthase. This may not only reflect L-cysteine specific feedback inhibition of cysteine synthase gene expression but also the overall concentration of thiols. There may be a demand for cysteine as a specific amino acid and also for thiol containing compounds as general reductants. These general reducing agents may include both low molecular weight thiols (e.g. H_2S, CH_3SH) and amino acids. The response to high levels of homocysteine was a fall in levels of cysteine synthase expression but not of the same magnitude as seen with excess cysteine. Thus exogenous homocysteine may have satisfied the requirements of the parasite for thiols but there still remained the need for cysteine for other purposes, such as protein synthesis.

It would be very interesting to examine the effects of the same conditions upon MGL levels. Experiments are planned to reproduce Western and Northern blots using MGL antibodies and nucleotide probes. The working hypothesis described would predict decreased levels of MGL expression when the parasites are

exposed to high levels of cysteine, due to a reduced demand for thiols in cysteine synthesis. The predicted response to elevated levels of methionine or homocysteine would be little or no change in MGL levels. It would be much more interesting to deplete the medium of one or more sulphur containing amino acids and examine the response then.

6.1.3 Thiols as Reducing Agents or Free Radical Scavengers

Despite the presence of an active cysteine synthase, there may be a role for compounds such as H_2S and CH_3SH as reducing agents for the parasite. It should be remembered that the parasite shows sensitivity to both intra- and extracellular thiol blocking reagents (define) which act as potent anti-trichomonad compounds in culture (Gillin et al., 1984). However from the observation that these thiols are mostly excreted from the cell they may be required extracellularly. It is possible that extracellular reducing agents may be required and the transport of small thiols across the cell membrane to satisfy this need is easier and more energetically efficient than excreting, for example, cysteine.

Determining the MGL levels under different growth conditions may also prove informative in indicating whether these low molecular weight thiols (H_2S and CH_3SH) function as reducing agents or free radical scavengers. Only when the parasite was cultured with propargylglycine did we see any dramatic upregulation of cysteine synthase activity. This was a little surprising, given that compounds such as H_2O_2 and methyl viologen were used to induce oxidative stress. It would be interesting to see if there was an upregulation of MGL under these conditions.

This may indicate that the thiol products of MGL are used as a protective response to damaging oxygen free radicals.

6.1.4 Low Molecular Weight Thiols in Haemolysis

Several examples of cysteine desulphurase enzymes have recently had their 3-dimensional structures solved and have already been discussed in chapter 3 of this thesis. One of these, cystalysin, has been shown to be responsible for haemolytic activity in the oral pathogen *Treponema denticola* (Chu et al., 1997; Krupka et al., 2000). It is unclear whether the protein itself or the H_2S it produces from the αβ-elimination of cysteine is responsible for this damage to host red blood cells. The toxicity of compounds such as H_2S and CH_3SH to mammalian cells has been documented (Beauchamp et al., 1984) and *T. vaginalis* is also haemolytic (Krieger *et al.*, 1983; Dailey *et al.*, 1990). The hypothesis that the action of MGL may be responsible for the haemolytic activity of *T. vaginalis* would appear to be simple to test. For instance, MGL activity can be inhibited by using propargylglycine and the haemolytic activity of the cells measured. However the cysteine desulphurase activity of cystalysin has been shown to be elevated in the presence of β-mercaptoethanol. This has similarities to the "activated serine sulphydrase activity" seen in *T. vaginalis* and although cysteine synthase has been shown to exhibit this activity, it is not certain that it is solely responsible. It has been shown that several enzymes can be resolved from parasite lysates with this activity, one or more of which may be cystalysin (Thong and Coombs, 1987). Thus, although culturing parasites in the presence of propargylglycine may eliminate haemolytic activity due to MGL, haemolysis may remain due to another,

unidentified factor. However, such an experiment would give some information upon whether or not MGL may be involved in haemolysis.

6.2 Future Work

A number of experiments that could be carried out immediately in attempts to answer specific questions have been mentioned already. MGL could be assayed for its ability to repair or construct functional Fe-S clusters and also to cause haemolysis. Parasite growth under various conditions could also be repeated and examined for its effect upon MGL expression. There are also a number of other stimuli that the parasites could be exposed to such as excess or reduced levels of iron, variations in pH or high levels of D-cysteine. It may also prove very useful to use a defined medium for culturing trichomonads (Linstead, 1981). The ability to easily omit medium components would be highly useful in examining the roles of enzymes in culture. The knowledge of exact levels of reducing agents, for example, could be exploited by examining which compounds could be substituted for others. With the probes for examining MGL and cysteine synthase expression, these experiments could prove highly informative.

In the longer term, the continued sequencing of the parasite EST library is likely to be particularly fruitful. The identification and characterisation of other genes may give further insights into the metabolism of sulphur amino acids. A continuing problem though is the inability to perform reverse genetics experiments upon the parasite.

References

Abeles, R.H. and Walsh, C.T. (1973) Acetylenic enzymic inactivators. Inactivation of γ-cystathionase, *in vitro* and *in vivo*, by propargylglycine. *Journal of the American Chemical Society* **95**, 6124-6125.

Abraham, M.C., Desjardins, M., Filion, L.G. and Garber G.E. (1996) Inducible immunity to *Trichomonas vaginalis* in a mouse model of vaginal infection. *Infection & Immunity* **64**, 3571-3575.

Alexander, F.W., Sandmeier, E., Mehta, P.K. and Christen, P. (1994) Evolutionary relationships among pyridoxal-5'-phosphate dependent enzymes. Regionspecific α, β and γ subgroups. *European Journal of Biochemistry* **219**, 953-960.

Alexeev, D., Alexeek, M., Baxter, R.L., Campopiano, D., Webster, S.P. and Sawyer, L. (1998) The crystal structure of 8-amino-7-oxononanoate synthase: a bacterial PLP-dependent, acyl-CoA-condensing enzyme. *Journal of Molecular Biology* **284**, 401-410.

Andersen, G.L., Beattie, G.A. and Lindow, S.E. (1998) Molecular characterization and sequence of a methionine biosynthetic locus from *Pseudomonas syringae*. *Journal of Bacteriology* **180**, 4497-4507.

Antson, A.A., Demidkina, T.V., Gollnick, P., Dauter, Z., Von Tersch, R.L., Long, J., Berezhnoy, S.N., Phillips, R.S., Harutyunyan, E.H. and Wilson, K.S. (1993) Three-dimensional structure of tyrosine phenol-lyase. *Biochemistry* **32**, 4195-4206.

Bacchi, C.J., Goldberg, B., Garofalo-Hannan, J., Rattendi, J. Lyte, P. and Yarlett, N. (1995) Fate of soluble methionine in African trypanosomes: effects of metabolic inhibitors. *Biochemical Journal* **309**, 737-743.

Bagnara, A.S., Tucker, V.E., Minotto, L., Howes, E.R., Gyung-Ae, K., Edwards, M.R. and Dawes, I.W. (1996) Molecular characterisation of adenosylhomocysteinase from *T. vaginalis*. *Molecular and Biochemical Parasitology* **81**, 1-11.

Bankov, I., Timanova, A. and Barrett, J. (1996) Methionine and cysteine metabolism in *Fasciola hepatica*. *International Journal for Parasitology* **26**, 1401-1404.

Barroso, C., Vega, J. M. and Gotor, C. (1995) A new member of the cytosolic *O*-acetylserine(thiol)lyase gene family in *Arabidopsis thaliana*. *FEBS Letters* **363**, 1-5.

Beauchamp, R.O., Bus, J.S., Popp, J.A., Boreiko, C.J. and Andjelkovich, D.A. (1984) A critical review of the literature on hydrogen sulfide toxicity. *Critical Reviews of Toxicology* **13**, 25-97.

Birolo, L., Malshkevich, V.N., Capitani, G., De Luca, F., Moretta, A., Jansonius, J.N. and Marino, G. (1999) Functional and structural analysis of *cis*-proline mutants of *Escherichia coli* aspartate aminotransferase. *Biochemistry*. **38**, 905-913.

Blankenfeldt, W., Nowicki, C., Hunter, G.R., Montemartini-Kalisz, M., Kalisz, H.M. and Hecht, H.J. (1999) Crystal structure of *Trypanasoma cruzi* tyrosine aminotransferase: substrate specificity is influenced by cofactor binding mode. *Protein Science* **8**, 2406-2411.

Bogdanova, N. and Hell, R. (1997) Cysteine synthesis in plants: protein-protein interactions of serine acetyltransferase from *Arabidopsis thaliana*. *The Plant Journal* **11**, 251-262.

Brown, D.M., Upcroft, J.A., Dodd, H.N., Chen, N. and Upcroft, P. (1999) Alternative 2-keto acid oxidoreductases in *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology* **98**, 203-214.

van Bruggen, J.J.A., Stumm, C.K. and Vogels G.D. (1983) Symbiosis of methanogenic bacteria and sapropelic protozoa. *Archives of Microbiology* **136**, 89-95.

Brünger, A. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.

Brush, A.W., and Paulus, H. (1971) O-acetylhomoserine sulphydrylase from *Bacillus subtilus*. *Biochemical and Biophysical Research Communications* **45**, 735-741.

Brush, A.W. and Paulus, H. (1973) O-acetylhomoserine sulphydrylase from *Brevibacterium flavum*. *Fedaration Proceedings* **32**, 63.

Bui, E.T.N., Bradley, P.J. and Johnson, P.J. (1996) A common evolutionary origin for mitochondria and hydrogenosomes. *Proceedings of the National Academy of Sciences, United States of America* **93**, 9651-9656.

Burch, T.A., Reis, C.W. and Reardon, L.V. (1959) Epidemiologic studies on human trichomoniasis. *American Journal of Tropical Medicine and Hygiene* **10**, 218-222.

Burkhard, P., Rao, G.S.J., Hohenester, E., Schnakerz, K.D., Cook, P.F. and Jansonius, J.N. (1998) Structure of *O*-acetylserine sulfhydrylase by X-ray crystallography at 2.2 Å. *Journal of Molecular Biology* **283**, 121-133.

Byrne, C.R., Monroe, R.S., Ward, K.A. and Kredich, N.M. (1988) DNA sequences of the *cysK* regions of *Salmonella typhimurium* and *Escherichia coli* and linkage of the *cysK* regions to *ptsH*. *Journal of Bacteriology* **190**, 3150-3157.

Capitani, G., Hohenester, E., Feng, L., Storici, P., Kirsch, J.F. and Jansonius, J.N. (1999) Structure of 1-aminocyclopropane-1-carboxylate synthase, a key enzyme in the biosynthesis of the plant hormone ethylene. *Journal of Molecular Biology* **294**, 745-760.

Carson, M. (1997) Ribbons. *Methods in Enzymology* **277**, 493-505.

Cavalier-Smith, T. (1987) The simultaneous symbiotic origin of mitochondria, chloroplasts and microbodies. *Annals of the New York Academy of Sciences* **503**, 55-72.

Christen, P. and Metzler, D.E. (1985) *Transaminases*, Wiley & Sons, New York.

Chu, L., Ebersole, J.L., Kurzban, G.P. and Holt, S.C. (1997) Cystalysin, a 46-kilodalton cysteine desulfhydrase from *Treponema denticola* with hemolytic and hemoxidative activities. *Infection and Immunity* **65**, 3231-3238.

Clausen, T., Huber, R. Laber, B., Pohlenz, H-D. and Messerschmidt, A. (1996) Crystal structure of the pyridoxal-5'-phosphate dependent cystathionine β-lyase from *Escherichia coli* at 1.83 Å. *Journal of Molecular Biology* **262**, 202-224.

Clausen, T., Laber, B. and Messerschmidt, A. (1997) Mode of action of cystathionine β-lyase. *Biological Chemistry* **378**, 321-326.

Clausen, T., Huber, R., Prade, L., Wahl, M.C. and Messerschmidt, A. (1998) Crystal structure of *Escherichia coli* cystathionine γ-synthase at 1.5 Å resolution. *EMBO Journal* **17**, 6827-6838.

Clausen, T., Kaiser, J.T., Steegborn, C, Huber, R. and Kessler, D. (2000a) Crystal structure of the cystine *C-S* lyase from *Synechocystis*: stabilisation of cysteine persulfide for FeS cluster biosynthesis. *Proceedings of the National Academy of Sciences, United States of America* **97**, 3856-3861.

Clausen, T., Schlegel, A., Peist, R., Schneider, E., Steegborn, C., Chang, Y.-S., Haase, A., Bourenkov, G.P., Bartunik, H.D. and Boos, W. (2000b) X-ray structure of MalY from *Escherichia coli*: A pyridoxal-5'-phosphate dependent enzyme acting as a modulator in *mal* gene expression. *EMBO Journal* **19**, 831-838.

Clemens, D.L. and Johnson, P.J. (2000) Failure to detect DNA in hydrogenosomes of *Trichomonas vaginalis* by nick translation and immunomicroscopy. *Molecular and Biochemical Parasitology* **106**, 307-313.

Cogne M., Brasseur, P. and Ballet, J.J. (1985) Detection and characterisation of serum antitrichomonal antibodies in urogenital trichomoniasis. *Journal of Clinical Microbiology* **21**, 588-592.

Coombs, G.H. and Müller, M. (1995) Energy metabolism in anaerobic protozoa. In Marr, J. J. and Müller, M. (eds) *Biochemistry and Molecular Biology of Parasites*. Academic Press, UK. pp. 33-47.

Cooper, A.J.L. (1983) Biochemistry of Sulphur Containing Amino Acids. *Annual Review of Biochemistry* **52**, 187-222.

Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research* **16**, 10881-10890.

Creedon, K.A., Rathod, P.K. and Wellems, T.E. (1994) *Plasmodium falciparum* S-adenosylhomocysteine hydrolase. CDNA identification, predicted protein

sequence and expression in *Escherichia coli*. *Journal of Biological Chemistry* **269**, 16364-16370.

Da'Darra, A.A., Henkle-Duhrsen, K. and Walter, R.D. (1996) A novel trans-spliced mRNA from *Onchocerca volvulus* encodes a functional S-adenosylmethionine decarboxylase. *Biochemical Journal* **320**, 519-530.

Dailey, D.C, Chang, T.H and Alderete, J.F. Characterization of *Trichomonas vaginalis* haemolysis. *Parasitology* **101**, 171–5.

Datko, A.H., Giovanelli, J. and Mudd, S.H. (1974) *O*-Phosphorylhomoserine as the physiological substrate for cystathionine γ-synthase. *Journal of Biological Chemistry* **249**, 1139-1155.

Davis-Hayman, S.R., Shah, P.H., Finley, R.W., Meade, J.C. and Lushbaugh, W.B. (2000) Antigenicity of *Trichomonas vaginalis* heat-shock proteins in human infections. *Parasitology Research* **86**, 115-120.

Deckert, G., Warren, P.V., Gaasterland, T., Young, W.G., Lenox, A.L., Graham, D.E., Overbeek, R., Snead, M.A., Keller, M., Aujay, M., Huber, R., Feldman, R.A., Short, J.M., Olsen, G.J. and Swanson R.V. (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**, 353-358.

Dias, B. and Weimer, B. (1998) Purification and characterization of L-methionine gamma-lyase from *brevibacterium linens* bl2. *Applied and Environmental Microbiology* **64**, 3327-3331.

Diamond, L. (1957) The establishment of various trichomonads of animals and man in axenic culture. *Journal of Parasitology* **43**, 488-490.

Droux, M., Martin, J., Sajus, P. and Douce, R. (1992) Purification and characterisation of O-acetylserine (thiol) lyase from spinach chloroplasts. *Archives of Biochemistry and Biophysics* **295**, 379-390.

Droux, M., Ruffet, M-L., Douce, R. and Job, D. (1998) Interactions between serine acetyltransferase and *O*-acetylserine (thiol) lyase in higher plants. Structural and kinetic properties of the free and bound enzymes. *European Journal of Biochemistry* **255**, 235-245.

Eads, J.C., Beeby, M., Scapin, G., Yu, T.W. and Floss, H.G. (1999) Crystal structure of 3-amino-5-hydroxybenzoic acid (AHBA) synthase. *Biochemistry* **38**, 9840-9849.

Edwards, D.I. (1993) Nitroimidazole drugs-action and resistance mechanisms. I. Mechanisms of action. *Journal of Antimicrobial Chemotherapy* **31**, 9-20.

Ellis, J.E., Yarlett, N., Cole, D., Humphreys, M.J. and Lloyd, D. (1994) Antioxidant defences in the microaerophilic protozoan *Trichomonas vaginalis*: comparison of metronidazole-resistant and sensitive strains. *Microbiology* **140**, 2489-2494.

Embley, T.M. and Hirt, R.P. (1998) Early branching eukaryotes? *Current Opinion in Genetics and Development* **8**, 624-629.

Esaki, N., Nakamura, T., Tanaka, H. and Soda, K. (1982) Selenocysteine lyase, a novel enzyme that specifically acts on selenocysteine. *Journal of Biological Chemistry*. **257**, 4386-4391.

Esaki, N. and Soda, K. (1987) L-methionine γ-lyase from *Pseudomonas putida* and *Aeromonas*. *Methods in Enzymology* **143**, 459-465.

Evans, S.V. (1993) SETOR - Hardware-lighted 3-dimensional solid model representations of macromolecules. *Journal of Molecular Graphics* **11**, 134-138, 127-128.

Fenchel, T and Finlay, B.J. (1991) The biology of free-living anaerobic ciliates. *European Journal of Protistology* **38**, 18-22.

Flavin, M. and Slaughter, C. (1967) Enzymatic synthesis of homocysteine or methionine directly from O-succinyl-homoserine. *Biochimica and Biophysica Acta* **132**, 400-405.

Foglino, M, Borne, F., Bally, M., Ball, G. and Patte, J.C. (1995) A direct sulfhydrylation pathway is used for methionine biosynthesis in *Pseudomonas aeruginosa*. *Microbiology* **141**, 431-439.

Ford, G.C., Eichele, G and Jansonius, J.N. (1980) Three dimensional structure of a pyridoxal-phosphate-dependent enzyme, mitochondrial aspartate aminotransferase. *Proceedings of the National Academy of Sciences, United States of America* **77**, 2559-2563.

Francioli, P., Shio, H., Roberts, R.B. and Müller, M. (1983) Phagocystosis and killing of *Neisseria gonorroheae* by *Trichomonas vaginalis*. *Journal of Infectious Diseases* **147**, 87-94.

Fujii, T., Maeda, M., Mihara, H., Kurihara, T., Esaki, N. and Hata, Y. (2000) Structure of a NifS homologue: X-ray structure analysis of CsdB an *Escherichia coli* counterpart of mammalian selenocysteine lyase. *Biochemistry* **39**, 1263-1273.

Gaitonde, M.K. (1967) A spectrophotometric method for the direct determination of cysteine in the presence of other naturally occuring amino acids. *Biochemical Journal* **104**, 627-633.

Gallagher, D.T., Gilliland, G.L., Xiao, G., Zondlo, J., Fisher, K.E., Chinchilla, D. and Eisenstein, E. (1998) Structure and control of pyridoxal phosphate dependent allosteric threonine deaminase. *Structure* **6**, 465-475.

Germot, A., Philippe, H. and Le Guayder, H. (1996) Presence of a mitochondrial-type HSP70 in *Trichomonas* suggests a very early mitochondrial endosymbiosis in eukaryotes. *Proceedings of the National Academy of Sciences, United States of America* **93**, 14614-14617.

Gillin, F.D., Reiner, D.S., Levy, R. B. and Henkart, P. A. (1984) Thiol groups on the surface of anaerobic parasitic protozoa. *Molecular and Biochemical Parasitology* **13**, 1-12.

Giovanelli, J. (1987) Sulfur amino acids of plants: an overview. *Methods in Enzymology* **143**, 419-426.

Goldberg, B., Rattendi, D., Lloyd, D., Yarlett, N. and Bacchi, C.J. (2000) Kinetics of methionine transport and metabolism by *Trypanosoma brucei brucei* and *Trypanosoma brucei rhodesiense*. *Archives of Biochemistry and Biophysics* **377**, 49-57.

Gomez-Bautista, M. and Barrett, J. (1988) Cysteine metabolism in the cestode *Hymenopolis diminuta*. *Parasitology* **97**, 149-159.

Good, N. E., Winget, G.D., Winter, W., Connolly, T.N., Izawa, S. and Singh, R.M. (1966) Hydrogen ion buffers for biological research. *Biochemistry* **5**, 467-477.

Gorrell, T.E. (1985) Effect of culture medium iron content on the biochemical composition and metabolism of *T. vaginalis in vitro*. *Journal of Bacteriology* **161**, 1228-1230.

Griffith, O. W. (1987) Mammalian sulfur amino acid metabolism: an overview. *Methods in Enzymology* **143**, 366-376.

Grishin, N.V., Phillips, M.A. and Goldsmith, E.J. (1995) Modelling of spatial structure of eukaryotic ornithine decarboxylase. *Protein Science* **4**, 1291-1304.

Grishin, N.V., Osterman, A.L., Brooks, H.B., Phillips, M.A. and Goldsmith, E.J. (1999) X-ray structure of ornithine decarboxylase from *Trypanosoma brucei*: the native structure and the structure in complex with α-difluoromethylornithine. *Biochemistry* **38**, 15714-15721.

Han, Q., Lenz, M., Tan, Y., Xu, M., Sun, X., Tan, X., Tan, X., Tang, T., Miljkovic, D. and Hoffman, R.M. (1998) High expression, purification and properties of recombinant homocysteine α,γ-lyase. *Protein Expression and Purification* **14**, 267-274.

Hayashi, H., Inoue, Y., Kuramitsu, S., Morino, Y. and Kagamiyama, H. (1990) Effects of replacement of tryptophan-140 by phenylalanine or glycine on the function of *Escherichia coli* aspartate aminotransferase. *Biochemical and Biophysical Research Communications* **167**, 407-412.

Hayashi, H. (1995) Pyridoxal enzymes: mechanistic diversity and uniformity. *Journal of Biochemistry* **118**, 463-473.

Heath, J.P. (1981) Behaviour and pathogenicity of *Trichomonas vaginalis* in epithelial cell cultures. A study by light and scanning electron microscopy. *British Journal of Venereal Disease* **57**, 106-117.

Hell, R., Bork, C., Bogdanova, N., Frolov, I. and Hauschild, R. (1994) Isolation and characterization of two cDNAs encoding for compartment specific isoforms of *O*-acetylserine(thiol)lyase from *Arabidopsis thaliana*. *FEBS Letters* **351**, 257-262.

Henderson, D.M., Hanson, S., Allen, T., Wilson, K., Coulter-Karis, D.E., Greenberg, M.L., Hershfield, M.S. and Ullman, B. (1992) Cloning of the

gene encoding *Leishmania donovani* S-adenosylhomocysteinehydrolase, a potential target for antiparasitic chemotherapy. *Molecular and Biochemical Parasitology* **53**, 169-183.

Hennig, M., Grimm, B., Contestabile, R., John, R.A. and Jansonius, J.N. (1997) Crystal structure of glutamate-1-semialdehyde aminomutase: an α^2- dimeric vitamin B_6-dependent enzyme with asymmetry in structure and active site reactivity. *Proceedings of the National Academy of Sciences, United States of America* **94**, 4866-4871.

Hesse, H. and Altmann, T. (1995) Molecular cloning of a cysteine synthase cDNA from *Arabidopsis thaliana*. *Plant Physiology* **108**, 851-852.

Hester, G., Stark, W., Moser, M., Kallen, J., Markovic-Houslay, Z. and Jansonius, J.N. (1999) Crystal structure of phopshoserine aminotransferase from *Escherichia coli* at 2.3 Å resolution: comparison of the unligated enzyme and a complex with α-methyl-L-glutamate. *Journal of Molecular Biology* **286**, 829-850.

Heyworth, R., Simpson, D., McNeillage, G.J., Robertson, D.H. and Young, H. (1980) Isolation of *Trichomonas vaginalis* resistant to metronidazole. *The Lancet* **2**, 476-478.

Heyworth, P.G., Gutteridge, W.E. and Ginger, C.D. (1982) Purine metabolism in *T. vaginalis*. *FEBS Letters* **141**, 106-110.

Hollander, D.H. and Leggett, N.C. (1985) Viamin B_{12} requirement for the growth of *T. vaginalis in vitro*. *Journal of Parasitology* **71**, 683-684.

Hori, H., Takabayashi, K., Orvi, L., Carson, D.A. and Nobori, T. (1996) Gene cloning and characterisation of *Pseudomonas putida* L-methionine-α-deamino-γ-mercaptomethane lyase. *Cancer Research* **56**, 2116-2122.

Huggins, G.R. and Petri, G. (1981) Vaginal odours and secretions. *Clinical Obstetrics and Gynaecology* **24**, 355-377.

Hunter, G.A. and Ferreira, G.C. (1999) Lysine-313 of 5-aminolevulinate synthase acts as a general base during formation of the quinonoid reaction intermediates. *Biochemistry* **38**, 3711-3718.

Hyde, C.C., Ahmed, S.A., Padlan, E.A., Wilson, M.E. amd Davies, D.R. (1988) Three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ multienzyme complex from *Salmonella typhimurium*. *Journal of Biological Chemistry* **263**, 17857-17871.

Inoue, H., Inagaki, K., Sugimoto, M., Esaki, N., Soda, K. and Tanaka, H. (1995) Structural-analysis of the l-methionine gamma-lyase gene from *Pseudomonas putida*. *Journal Of Biochemistry* **117**, 1120-1125.

Inoue, H., Inagaki, K., Eriguchi, S., Temura, T., Esaki, N., Soda, K. and Tanaka, H. (1997) Molecular characterisation of the *mde* operon involved in L-methionine catabolism of *Pseudomonas putida*. *Journal of Bacteriology* **179**, 3956-3962.

Isupov, M.N., Antson, A.A., Dodson, E.J., Dodson, G.G., Dementieva, I.S., Zakomirdina, L.N., Wilson, K.S., Dauter, Z., Lebedev, A.A. and Harutyunyan, E.H. (1998) Crystal structure of tryptophanase. *Journal of Molecular Biology* **276**, 603-623.

Jacobson, E.S. and Metzenburg, R.L. (1977) Control of aryl sulfatase in a serine auxotroph of *Neurosporra*. *Journal of Bacteriology* **130**, 1397-1398.

Jancarik, J and Kim, S-H. (1991) Sparse matrix sampling: a screening method for crystallisation of proteins. *Journal of Applied Crystallography* **24**, 409-411.

Jansonius, J.N. (1998) Structure, evolution and action of vitamin B_6-dependent enzymes. *Current Opinion in Structural Biology* **8**, 759-769.

Jeffrey, C.J., Barry, T., Doonan, S., Petsko, G.A. and Ringe, D. (1998) Crystal structure of *Saccaromyces cerevisiae* cytosolic aspartate aminotransferase. *Protein Science* **7**, 1380-1387.

Johnston, M., Jankowski, D., Marcotte, P., Tanaka, H., Esaki, N., Soda, K. and Walsh, C. (1979) Suicide inactivation of bacterial cystathionine γ-synthase and methionine γ-lyase during processing of L-propargylglycine. *Biochemistry* **18**, 4691-4701.

Jones, S. and Thornton, J. (1995) Protein-protein interactions: a review of protein dimer structures. *Progress in Biophysics and Molecular Biology* **63**, 13-20.

Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen bonded and geometrical features. *Biopolymers* **22**, 2577-2637.

Kack, H., Sandmark, J., Gibson, K.J., Scneider, G. and Lindqvist, Y. (1999) Crystal structure of diaminopelargonic acid synthase: evolutionary relationships between pyridoxal-5'-phosphate dependent enzymes. *Journal of Molecular Biology* **291**, 857-869.

Kaiser, J.T., Clausen, T., Bourenkow, G.P., Bartunik, H-D., Steinbacher, S. and Huber, R. (2000) Crystal structure of a NifS-like protein from *Thermotoga maritima*: implications for iron sulphur cluster assembly. *Journal of Molecular Biology* **297**, 451-464.

Kellock, D.J. and O'Mahoney, C.P. (1996) Sexually acquired metronidazole-resistant trichomoniasis in a lesbian couple. *Genitourinary Medicine* **72**, 60-61.

Kern, A.D., Oliveira, M.A., Coffino, P. and Hackert, M.L. (1999) Structure of mammalian ornithine decarboxylase at 1.6 Å resolution: stereochemical implications of PLP-dependent amino acid decarboxylases. *Structure Fold and Design* **7**, 567-579.

Knochel, T., Ivens, A., Hester, G., Gonzalez, A., Bauerle, R., Wilmanns, M., Kirschner, K. and Jansonius, J.N. (1999) The crystal structure of anthranilate synthase from *Solfolobus solfataricus*: functional implications. *Proceedings of the National Academy of Sciences, Unites States of America* **96**, 9479-9484.

Kolman, C. and Söll, D. (1993) Spl1-1, a *Saccharomyces cerevisiae* mutation affecting transfer-RNA splicing. *Journal of Bacteriology* **175**, 1433-1442.

Kredich, N.M., Becker, M.A. and Tomkins, G.M. (1969) Purification and characterisation of cysteine synthetase: a bifunctional protein complex,

from *Salmonella typhimurium*. *Journal of Biological Chemistry* **244**, 2428-2439.

Kredich, N.M. (1996) Biosynthesis of cysteine, pp514-527. *In* Neidhart, F.C., Curtiss, R., Ingraham, J.L., Lin, E.C.C., Low, B., Magasanik, B., Reznikoff, W.S., Riley, M., Schaccter, M. and Umbarger, H.E. (ed.), *Escherichia coli* and *Salmonella*, 2nd ed. ASM Press, Washington D.C.

Krieger, J.N, Poisson, M.A and Rein, M.F. Beta-haemolytic activity of *Trichomonas vaginalis* correlates with virulence. *Infection and Immunity* **41**, 1291–5.

Kreiger, J.N., Tam, M.R., Stevens, C.E., Nielsen, I.O., Hale, J., Kiviat, N.B. and Holmes, K.K. (1988) Diagnosis of trichomoniasis: comparison of conventional wet-mount examination with cytological studies, cultures, and monoclonal antibody staining of direct specimens. *Journal of the American Medical Association* **259**, 1223-1227.

Kreiger, J.N., Verdon, M., Siegel, N. and Holmes, K.K. (1993) Natural history of urogenital trichomoniasis in men. *Journal of Urology* **149**, 1455-1458.

Kreis, W. and Hession, C. (1973) Isolation and purification of L-methionine α-deamino γ-mercaptomethane-lyase (L-methioninase) from *Clostridium sporogenes*. *Cancer Research* **33**, 1862-1865.

Krupka, H.I., Huber, R., Holt, S.C. and Clausen, T. (2000) Crystal structure of cystalysin from *Treponema denticola*: a pyridoxal 5'-phosphate-dependent protein acting as a haemolytic enzyme. *EMBO Journal* **19**, 3168-3178.

Kulda, J., Vojtechovska, M., Tachezy, J., Demes, P. and Kunzova, E. (1982) Metronidazole resistance of *Trichomonas vaginalis* as a cause of treatment failure in trichomoniasis. *British Journal of Venereal Disease* **58**, 394-399.

Kulda, J., Nohýnková, E. and Ludvík, J. (1986) Basic structure and function of the trichomonad cell. *Acta Universitatis Carolinae - Biologica* **30**, 181-198.

Lacourciere, G.M. and Stadtman, T.C. (1998) The NIFS protein can function as a selenide delivery protein in the biosynthesis of selenophosphate. *Journal of Biological Chemistry* **273**, 30921-30926.

Laemmli, U.K. (1971) Cleavage of structural proteins during the assembly of the head of bacteriophage. *Nature* **227**, 680-685.

Laga, M., Manoko, A., Kivuvu, M. Malele, B., Tuliza, M., Nzila, N., Goeman, J., Behets, F., Batter, V., Alary, M., Heyward, W.L., Ryder, R.W. and Piot P (1993) Non ulcerative sexually transmitted diseases as risk factor for HIV-1 transmission in women: results from a cohort study. *AIDS* **7**, 95-102.

Lamzin, V.S. and Wilson, K.S. (1993) Automated refinement of protein models. *Acta Crystallogr*aphica **D49**, 129-147.

Land, K.M. and Johnson, P.J. (1999) Molecular basis of metronidazole resistance in pathogenic bacteria and protozoa. *Drug Resistance Update* **2**, 289-294.

Lang, T. and Kessler, D. (1999) Evidence for cysteine persulphide as reaction product of L-cyst(e)ine *C-S* lyase (C-DES) from *Synechocystis*. *Journal of Biological Chemistry* **274**, 189-195.

Lindmark, D.G. and Müller, M. (1973) Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate *Tritrichomonas foetus*, and its role in pyruvate metabolism. *Journal of Biological Chemistry* **248**, 7724-7728.

Linstead, D. (1981) New defined and semi-defined media for cultivation of the flagellate *Trichomonas vaginalis*. *Parasitology* **83**, 125-137.

Linstead, D. and Cranshaw, M. (1983) The pathway of arginine catabolism in the parasitic flagellate *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology* **8**, 241-252.

Liston, D.R. and Johnson, P.J. (1999) Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element. *Molecular and Cellular Biology* **19**, 2380-2388.

Lockwood, B. and Coombs, G.H. (1991a) Amino acid catabolism in anaerobic protists. In Coombs, G.H. and North, M.J. (eds) *Biochemical Protozoology*, pp. 113-122, Taylor and Francis, London, UK.

Lockwood, B. and Coombs, G.H. (1991b) Purification and characterisation of methionine γ-lyase from *Trichomonas vaginalis*. *Biochemistry Journal* **279**, 675-682.

Lossick, J.G. (1989) Epidemiology of urogenital trichomoniasis. In Honigberg, B.M. (Ed) *Trichomonads parasitic in humans*. pp311-323. New York: Springer Verlag.

Lowe, P.N. and Rowe, A.F. (1986) Aminotransferase activities in *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology* **21**, 65-74.

Lunn, J.E., Droux, M. Martin, J. and Drouce, R. (1990) Localisation of ATP sulfurylase and O-acetylserine(thiol)lyase in spinach leaves. *Plant Physiology* **94**, 1345-1352.

MacNicol, P.K., Datko, A.H., Giovanelli, J. and Mudd, S.H. (1981) Homocysteine biosynthesis in green plants - physiological importance of the transsulfuration pathway in *Lemna paucicostata*. *Plant Physiology* **68**, 619-625.

Malashkevich, V.N., Strokopytov, B.V., Borisov, V.V., Dauter, Z., Wilson, K.S. and Torchinsky, Y.M. (1995) Crystal-structure of the closed-form of chicken cytosolic aspartate-aminotransferase at 1.9 angstrom resolution. *Journal Of Molecular Biology* **247**, 111-124.

Malkin, A. and McPherson, A. (1994) Light-scattering investigations of nucleation processes and kinetics of crystallization in macromolecular systems. *Acta Crystallography* **D50**, 385-395.

Martin, H.L., Richardson, B.A., Nyange, P.M., Lavreys, L., Hillier, S.L., Chohan, B., Mandaliya, K., Ndinya-Achola, J.O., Bwayo, J and Kreiss J (1999) Vaginal lactobacilli, microbial flora, and risk of human immunodeficiency virus type 1 and sexually transmitted disease acquisition. *Journal of Infectious Diseases* **180**, 1863-1868.

Martin, J.L., Johnson, L.N. and Withers, S.G. (1990) Comparison of the binding of glucose and glucose-1-phosphate derivatives to T-state glycogen phosphorylase. *Biochemistry* **29**, 10745-10757.

Masada, M., Fukushima, K. and Tamura, G. (1975) Cysteine synthase from rape leaves. *Journal of Biochemistry* **77**, 1107-1115.

McKie, A.E. (1997) Molecular and Biochemical Characterisation of Methionine γ-lyase from *Trichomonas vaginalis*. PhD Thesis, University of Glasgow.

McKie, A.E., Edlind, T., Walker, J., Mottram, J.C. and Coombs, G.H. (1998) The primitive protozoon *Trichomonas vaginalis* contains two methionine γ-lyase genes that encode members of the γ-family of pyridoxal 5'-phosphate-dependent enzymes. *Journal of Biological Chemistry* **273**, 5549-5556.

McLellan, R,. Spence, M.R., Brockman, M., Rattel, L. and Smith, J.L. (1982) The clinical diagnosis of trichomoniasis. *Obstetrics and Gynaecology* **60**, 30-34.

Meri, T., Jokiranta, T.S., Suhonen, L. and Meri, S. (2000) Resistance of *Trichomonas vaginalis* to metronidazole: report of the first three cases from Finland and optimization of *in vitro* susceptibility testing under various oxygen concentrations. *Journal of Clinical Microbiology* **38**, 763-767.

Mihara, H., Kurihara, T., Yoshimura, T., Soda, K. and Esaki, N. (1997) Cysteine sulfinate desulfinase, a NIFS-like protein of *Escherichia coli* with selenocysteine lyase and cysteine desulfurase activities - Gene cloning, purification, and characterization of a novel pyridoxal enzyme. *Journal of Biological Chemistry* **272**, 22417-22424.

Mihara, H., Maeda, M., Fujii, T., Kurihara, T., Hata, Y. and Esaki, N. (1999) A *nif*S-like gene, *csdB*, encodes an *Escherichia coli* counterpart of mammalian selenocysteine lyase. *Journal of Biological Chemistry* **274**, 14768-14772.

Minotto, L., Ko, G.-A., Edwards, M.R. and Bagnara, A. (1998) *Trichomonas vaginalis*: Expression and characterisation of recombinant S-adenosylhomocysteinase. *Experimental Parasitology* **90**, 175-180.

Mitchell, S. (ed.) (1996) Biological interactions of sulfur compounds, Taylor and Francis, London UK.

Momany, C., Ernst, S., Ghosh, R., Chang, N.L. and Hackert, M.L. (1995) Crystallographic structure of a PLP-dependent ornithine decarboxylase from *Lactobacillus*-30a to 3.0 angstrom resolution. *Journal Of Molecular Biology* **252**, 643-655.

Morzycka, E. and Paszewski, A. (1979) Two pathways of cysteine biosynthesis in *Saccharomyces lipolytica*. *FEBS Letters* **101**, 97-100.

Müller, M., Meingassner, J.G., Miller, W.A. and Ledger, W.J. (1980) Three metronidazole resistant strains of *Trichomonas vaginalis* from the United States. *American Journal of Obstetrics and Gynaecology* **138**, 808-812.

Müller, M. (1988) Energy metabolism of protozoa without mitochondria. *Annual Review of Microbiology* **42**, 465-488.

Müller, M. (1993) The hydrogenosome. *Journal of General Microbiology* **139**, 28879-28889.

Müller, S., Da'dara, A., Lüersen, K., Wrenger, C., Das Gupta, R., Madhubala, R. and Walter, R.D. (2000) In the human malaria parasite *Plasmodium falciparum*, polyamines are synthesized by a bifunctional ornithine decarboxylase, *S*-adenosylmethionine decarboxylase. *Journal of Biological Chemistry* **275**, 8097-8102.

Muresu, R., Rubino, S., Rizzu, P., Baldini, A., Colombo, M. and Cappucinelli, P. (1994) A new method for identification of *Trichomonas vaginalis* by fluorescent DNA *in situ* hybridization. *Journal of Clinical Microbiology* **32**, 1018-1022.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum likelihood method. *Acta Crystallographica* **D53**, 240-255.

Nakai, Y., Yoshihara, Y., Hayashi, H. and Kagamiyama, H. (1998) cDNA cloning and characterization of mouse *nifS*-like protein, m-Nfs1: mitochondrial localization of eucaryotic NifS-like proteins. *FEBS Letters* **433**, 143-148.

Nakai, T., Okada, K., Akutsu, S., Miyahara, I., Kawaguchi, S., Kato, R., Kuramitsu, S. and Hirotsu, K. (1999) Structure of *Thermus thermophilus* HB8 aspartate aminotransferase and its complex with maleate. *Biochemistry* **38**, 2413-2419.

Nakamura, K., Kon, H., Iwahashi, H. and Eguchi, Y. (1983) Evidence that thiosulfate assimilation by *Salmonella typhimurium* is catalyzed by cysteine synthase B. *Journal of Bacteriology* **156**, 656-662.

Nakamura, K.H., Iwahashi, H. and Eguchi, Y. (1984) Enzymatic proof for the identity of the *S*-sulfocysteine synthase and cysteine synthase B of *Salmonella typhimurium*. *Journal of Bacteriology* **158**, 1122-1127.

Nakamura, K. and Tamura, G. (1990) Isolation of serine acetyl-transferase complexed with cysteine synthase from *Allium tuberosum*. *Agricultural Biological Chemistry* **54**, 649-656.

Nakamura, T., Koizumi, N. and Sano, H. (1997) Isolation of a novel cysteine synthase cDNA (AB003041) from *Arabidopsis thaliana*. *Plant Physiology* **114**, 747.

Nakamura, T., Yamaguchi, Y. and Sano, H. (1999) Four rice genes encoding cysteine synthase: isolation and differential responses to sulfur, nitrogen and light. *Gene* **229**, 155-161.

Nakayama, T., Esaki, N., Lee, W.-J., Tanaka, I., Tanaka, H. and Soda, K. (1984) Purification and properties of L-methionine γ-lyase from *Aeromonas* sp. *Agricultural and Biological Chemistry* **48**, 2367-2369.

Nakayama, T. Esaki, N. Tanaka, H and Soda, K. (1988a) Chemical modification of cysteine residues of L-methionine γ-lyase. *Agricultural and Biological Chemistry* **52**, 177-183.

Nakayama, T. Esaki, N. Tanaka, H and Soda, K. (1988b) Specific labelling of the essential cysteine residue of L-methionine gamma-lyase with a cofactor analog, n-(bromoacetyl)pyridoxamine phosphate. *Biochemistry* **27**, 1587-1591.

Navaza, J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallographica* **A50**, 157-163.

Nicholls, A., Sharp, K.A. and Honig, B. (1991) Protein folding and association - insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Structure Function and Genetics* **11**, 281-296.

North, M.J. (1991) Proteinases of trichomonads and *Giardia*. In Coombs, G.H. and North, M.J. (eds) *Biochemical Protozoology*, pp. 234-244, Taylor and Francis, London, UK.

Nozaki, T., Asai, T., Kobayashi, S., Ikegami, F., Noji, M., Saito,K. and Takeuchi, T. (1998) Molecular cloning and characterization of the genes encoding two isoforms of cysteine synthase in the enteric protozoan parasite *Entamoeba histolytica*. *Molecular and Biochemical Parasitology* **97**, 33-44.

Nozaki, T., Asai, T., Sanchez, L.B., Kobayashi, S., Nakazawa, M. and Takeuchi, T. (1999) Characterization of the gene encoding serine acetyltransferase, a regulated enzyme of cysteine biosynthesis from the protist parasites *Entamoeba histolytica* and *Entamoeba dispar* - Regulation and possible function of the cysteine biosynthetic pathway in entamoeba. *Journal of Biological Chemistry* **274**, 32445-32452.

Nozaki, T., Tokoro, M., Imada, M., Saito, Y., Abe, Y., Shigeta, Y. and Takeuchi, T. (2000) Cloning and biochemical characterisation of genes encoding two isozymes of cysteine synthase from *Entamoeba dispar*. *Molecular and Biochemical Parasitology* **107**, 129-133.

Okamoto, A., Higuchi, T., Hirotsu, K., Kuramitsu, S. and Kagamiyama, H. (1994) X-ray crystallographic study of pyridoxal 5'-phosphate-type aspartate aminotransferases from *Escherichia coli* in an open and closed form. *Journal of Biochemistry* **116**, 95-107.

Okamoto, A., Nakai, Y., Hayashi, H., Hirotsu, K. and Kagamiyama, H. (1998) Crystal structures of *Paracoccus denitrificans* aromatic amino acid aminotransferase: a substrate recognition site constructed by rearrangement of hydrogen bond network. *Journal of Molecular Biology* **280**, 443-458.

Ono, B-I., Kijima, K., Inoue, T., Miyoshi, S.I., Matsuda, A. and Shinoda, S. (1994) Purification and properties of *Saccharomyces cerevisiae* cystathionine β-synthase. *Yeast* **10**, 333-339.

Ono, B-I., Hazu, T., Yoshida, S., Kawato, T., Shinoda, S., Brzvwczy, J. and Paszewski, A. (1999) Cysteine biosynthesis in *Saccaromyces cerevisiae*: a new outlook on pathway and regulation. *Yeast* **15**, 1365-1375.

Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307-326.

Paget, T.A. and Lloyd, D. (1990) *Trichomonas vaginalis* requires traces of oxygen and high concentrations of carbon dioxide for optimal growth. *Molecular and Biochemical Parasitology* **41**, 65-72.

Papadopoulos, A., Walker, J. and Barrett, J. (1996) A novel cystathionine β-synthase from *Panagrellus redivivus*. *International Journal of Biochemistry and Cell Biology* **28**, 543-549.

Paszewski, A. and Granski, J. (1974) Regulation of S-amino acids biosynthesis in *Aspergillus nidulans*. Role of cysteine and/or homocysteine as regulatory effectors. *Molecular and General Genetics* **132**, 307-320.

Paterson, B.A., Garland, S.M., Bowden, F.J., Tabrizi, S.N. and Fairley, C.K. (1999) The diagnosis of *Trichmonas vaginalis*: new advances. *International Journal of STD & AIDS* **10**, 68-70.

Paul, R.G., Williams, A.G. and Butler, R.D. (1990) Hydrogenosomes in the rumen entodiniomorphid ciliate *Polyplastron multivesiculatum*. *Journal of General Microbiology* **136**, 1981-1989.

Persson, K., Aslund, L., Grahn, B., Hanke, J. and Heby, O. (1998) *Trypanosoma cruzi* has not lost its S-adenosylmethionine decarboxylase: characterisation of the gene and the encoded enzyme. *Biochemical Journal* **333**, 527-537.

Piotrowska, M., Kruszewska, A. and Paszewski, A. (1980) Homocysteine synthesis in *Neurosporra crassa*. *Acta Biochimica Poland* **27**, 395-403.

Pletnev, S.V., Antson, A.A., Sinitsyna, N.I., Dauter, Z., Isupov, M.N., Hurs, E.N., Faleev, N.G., Wilson, K.S., Dodson, G., Demidkina, T.V. and Arutyunyan, E.G. (1997) Crystallographic study of tyrosine phenol-lyase from *Erwinia herbicola*. *Crystallography Reports* **42**, 808-819.

Quon, D.V.K., Delgadillo, M.G., Khachi, A., Smale, S.T. and Johnson, P.J. (1994) Similarity between a ubiquitous promoter element in an ancient eukaryote and mammalian initiator elements. *Proceedings of the National Academy of Sciences, United States of America* **91**, 4579-4583.

Rathaur, S. and Walter, R.D. (1987) *Plasmodium falciparum*: *S*-adenosyl-L-methionine decarboxylase. *Experimental Parasitology* **63**, 227-232.

Rowe, A.F. and Lowe, P.N. (1986) Modulation of amino acid and 2-oxo-acid pools in *T. vaginalis* by aspartate aminotransferase inhibitors. *Molecular and Biochemical Parasitology* **21**, 17-24.

Saito, K., Yokoyama, H., Noji, M. and Murakoshi, I. (1995) Molecular cloning and characterisation of a plant serine acetyltransferase playing a regulatory role in cysteine biosynthesis from watermelon. *Journal of Biological Chemistry* **270**, 16321-16326.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular Cloning: a laboratory manual (2nd ed.). Cold Spring Harbor Laboratory Press.

Scarsdale, J.N., Kazanina, G., Radaev, S., Schirch, V. and Wright, H.T. (1999) Crystal structure of rabbit cytosolic serine hydroxymethyltransferase at 2.8 angstroms resolution: mechanistic implications. *Biochemistry* **38**, 8347-8358.

Scarsdale, J.N., Radaev, S., Kazanina, G., Schirch, V. and Wright, H.T. (2000) Crystal structure at 2.4 angstrom resolution of *E. coli* serine hydroxymethyltransferase in complex with glycine substrate and 5-formyl tetrahydrofolate. *Journal of Molecular Biology* **296**, 155-168.

Shaw, J.P., Petsko, G.A. and Ringe, D. (1997) Determination of the structure of alanine racemase from *Bacillus stearothermophilus* at 1.9 Å resolution. *Biochemistry* **36**, 1329-1342.

Shen, B.W., Hennig, M., Hohenester, E, Jansonius, J.N. and Schirmer, T. (1998) Crystal structure of human recombinant ornithine aminotransferase. *Journal of Molecular Biology* **277**, 81-102.

Siegel, L.M., Murphy, M.J. and Kamin, H. (1973) Reduced nicotinamide adenine dinucleotide phosphate-sulfite reductase of Enterobacteria. I. The *Escherichia coli* hemoflavoprotein: molecular parameters and prosthetic groups. *Journal of Biological Chemistry* **248**, 251-261.

Silverman, R.B. and Abeles, H.A. (1977) Mechanism of inactivation of γ-cystathionase by β,β,β-trifluoroalanine. *Biochemistry* **16**, 5515-5520.

Ravanel, S., Droux, M. and Douce, R. (1995) Methionine biosynthesis in higher plants. I. Purification and characterisation of cystathionine γ-synthase from spinach chloroplasts. *Archives of Biochemistry and Biophysics* **316**, 572-584.

Ravanel, S., Gakiére,B., Job, D. and Douce, R. (1998) The specific features of methionine biosynthesis and metabolism in plants. *Proceedings of the National Academy of Sciences, Unites States of America* **95**, 7805-7812.

Read, C.P. (1957) Comparative studies on the physiology of trichomonad protozoa. *Journal of Parasitology* **43**, 385-394.

Rege, V.D., Kredich, N.M., Tai, C-H., Karsten, W.E., Schnackerz, K.D. and Cook, P.F. (1996) A change in the internal aldimine lysine (K42) in *O*-acetylserine sulfhydrylase to alanine indicates its importance in transamination and as a general base catalyst. *Biochemistry* **35**, 13485-13493.

Rein, M.F. (1990) Clinical manifestations of urogenital trichomoniasis in women. In Honigberg, B.K. (ed.) *Trichomonads parasitic in humans*. New York: Springer-Verlag 225-234.

Rendón-Maldonado, J.G., Espinosa-Cantellano, M., González-Robles, A. and Martínez-Palomo, A. (1998) *Trichomonas vaginalis*: *In vitro* phagocytosis of lactobacilli, vaginal epithelial cells, leukocytes and erythrocytes. *Experimental Parasitology* **89**, 241-250.

Renwick, S.B., Snell, K. and Baumann, U. (1998) The crystal structure of human cytosolic serine hydroxymethyltransferase: a target for cancer chemotherapy. *Structure* **6**, 1105-1116.

Riley, D.E., Samadpour, M. and Krieger, J.N. (1991) Detection of variable DNA repeats in diverse eukaryotic microorganisms by a single set of polymerase chain-reaction primers. *Journal of Clinical Microbiology* **29**, 2746-2751.

Rolland, N, Droux, M. and Drouce, R. (1992) Subcellular distribution of *O*-acetylserine(thiol)lyase in cauliflower (*Brassica oleracea* L.) in fluorescence. *Plant Physiology* **98**, 927-935.

Snyers, L., Hellings, P., Bovy-Kessler, C. and Thines-Sempoux, D. (1982) Occurrence of hydrogenosomes in the rumen ciliates *Ophryoscolecidae*. *FEBS Letters* **137**, 35-39.

Soda, K. (1968) Micro-determination of D-amino acids and D-amino oxidase activity with 3-methyl-2-benzothiazolone hydrazone hydrochloride. *Analytical Biochemistry* **25**, 228-235.

Sogin, M.L., Silberman, J.D., Hinkle, G. and Morrison, H.G. (1996) Problems with molecular diversity in the eukarya. In Roberts, D.M., Sharp, P., Alderson, G. and Collins, M. (Eds.) *Evolution of Microbial Life*. Pp167-184. Cambridge University Press. Cambridge, UK.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503.

Steegborn, C., Messerschmidt, A., Laber, B., Streber, W., Huber, R. and Clausen, T. (1999) The crystal structure of cystathionine γ-synthase from *Nicotiana tabacum* reveals its substrate and reaction specificity. *Journal of Molecular Biology* **290**, 983-996.

Sugio, S., Petsko, G.A., Manning, J.M., Soda, K. and Ringe, D. (1995) Crystal structure of a D-amino acid aminotransferase: how the protein controls stereoselectivity. *Biochemistry* **34**, 9661-9669.

Tai, C.H. and Cook, P.F. (2000) O-acetylserine sulfhydrylase. *Advances in Enzymology & Related Areas of Molecular Biology*. **74**, 185-234.

Thomas, D. and Surdin-Kerjan, Y. (1997) Metabolism of sulfur amino acids in *Saccharomyces cerevisiae*. *Microbiology and Molecular Biology Reviews* **61**, 503-532.

Thong, K-W. (1985) The biochemistry of sulphur-containing amino acids in trichomonads. PhD thesis, University of Glasgow.

Thong, K-W and Coombs, G. H. (1985a) Homocysteine desulphurase activity in trichomonads. *IRCS Medical Science* **13**, 493-494.

Thong, K-W and Coombs, G.H. (1985b) L-serine sulphydrase activity in trichomonads. *IRCS Medical Sciencel* **13**, 495-496.

Thong, K-W., Coombs, G.H. and Sanderson, B.E. (1985) S-Adenosylhomcysteine hydrolase activity in *Trichomonas vaginalis* and other trichomonads. *Molecular and Biochemical Parasitology* **17**, 35-44.

Thong, K-W. and Coombs, G.H. (1987) *Trichomonas* species: Homocysteine desulphurase and serine sulphydrase activities. *Experimental Parasitology*. **63**, 143-151.

Thong, K-W., Coombs, G.H. and Sanderson, B.E. (1987) L-Methionine catabolism in trichomonads. *Molecular and Biochemical Parasitology* **73**, 193-198.

Toney, M.D., Hohenester, E., Cowan, S.W. and Jansonius, J.N. (1993) Dialkylglycine decarboxylase structure: bifunctional active site and alkali metal sites. *Science* **261**, 756-759.

Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proceedings of the National Academy of Sciences, United States of America* **76**, 4350-4354.

Tsang, M.L.-S. and Schiff, J.A. (1976b) Sulfate-reducing pathway in *Escherichia coli* involving bound intermediates. *Journal of Bacteriology* **125**, 923-933.

Tung, W.L and Chow, K-C. (1995) A modified medium for efficient electrotransformation of *E. coli*. *Trends in Genetics*. **11**, 128-129.

Veesler, S., Marcq, S., Lafont, S., Astier, J.P. and Boistelle, R. (1994) Influence of polydispersity on protein crystallization - a quasi-elastic light-scattering study applied to alpha-amylase. *Acta Crystallography* **D50**, 355-360.

Vermeij, P. and Kertesz, M.A. (1999) Pathways of assimilative sulfur metabolism in *Pseudomonas putida*. *Journal of Bacteriology*. **181**, 5833-5837.

Wagner, G. and Levin, R. (1984) Human vaginal pH and sexual arousal. *Fertility and Sterility* **41**, 389-394.

Walker, J. and Barrett, J. (1991) Cystathionine β-synthase and γ-cystathionase in helminths. *Parasitology Research* **77**, 709-713.

Walker, J., Barrett, J. and Thong, K-W. (1992) The identification of a variant form of cystathionine β-synthase in nematodes. *Experimental Parasitology* **75**, 415-424.

Walker, J. and Barrett, J. (1997) Parasite sulphur amino acid metabolism. *International Journal For Parasitology* **27**, 883-897.

Wang, A.L. and Wang, C.C. (1985) Isolation and characterization of DNA from *Tritrichomonas foetus* and *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology* **14**, 323-335.

Wang, C.L. and Cleng, H.W. (1984) Salvage of pyrimidine nucleotides by *T. vaginalis*. *Molecular and Biochemical Parasitology* **10**, 171-184.

Warrilow, A.G.S. and Hawkesford, M.J. (1998) Separation, subcellular location and influence of sulphur nutrition on isoforms of cysteine synthase in spinach. *Journal of Experimental Botany* **49**, 1625-1636.

Washtien,W. and Abeles, R.H. (1977) Mechanism of Inactivation of γ-Cystathionase by the Acetylenic Substrate Analogue Propargylglycine. *Biochemistry* **16**, 2485-2491.

Watabe, A., Kurokawa, Y., Yoshimura, T., Kurihara, T., Soda, K. and Esaki, N. (1999) Role of lysine 39 of alanine racemase from *Bacillus stearothermophilus* that binds pyridoxal 5'-phosphate. *Journal of Biological Chemistry* **274**, 4189-4194.

Watanabe, N., Sakabe, K., Sakabe, N., Higashi, T., Sasaki, K., Aibara, S., Morita, Y., Yonaha, K., Toyama, S. and Fukutani, H. (1989) Crystal structure analysis of ω-amino acid:pyruvate aminotransferase with a newly developed Weissenburg camera and an imaging plate using synchrotron radiation. *Journal of Biochemistry* **105**, 1-3.

Whateley, F.R. (1981) The establishment of mitochondria: *Paracoccus* and *Rhodopseudomonas*. *Annals of the New York Academy of Sciences* **361**, 330-340.

Whittington, J.M. (1957) Epidemiology of infections with *Trichomonas vaginalis* in the light of improved diagnostic methods. *American Journal of Tropical Medicine and Hygiene* **8**, 312-318.

Williams, K., Lowe, P.N. and Leadlay, P.F. (1987) Purification and characterisation of pyruvate:ferredoxin oxidoreductase from the anaerobic protozoan *Trichomonas vaginalis*. *Biochemical Journal.* **246**, 529-536.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., *et al.* (1994) 2.2 Mb of contiguous nucleotide-sequence from chromosome-iii of *C. elegans*. *Nature* **368**, 32-38.

Yamagata, S. (1971) Homocysteine synthase in various yeasts. *Journal of Biochemistry* **70**, 1035-1045.

Yamagata, S., Takeshima, K. and Naiki, N. (1974) Evidence for the identity of O-acetylserine sulfhydrylase with O-acetylhomoserine sulfhydrylase in yeast. *Journal of Biochemistry* **75**, 1221-1229.

Yamagata, S. (1993) Cloning and bacterial expression of the CYS3 gene encoding cystathionine γ-lyase of *Saccharomyces cerevisiae* and the physicochemical and enzymatic properties of the protein. *Journal of Bacteriology* **175**, 4800-4808.

Yamaguchi, T., Zhu, X. and Masada, M. (1998) Purification and characterisation of a novel cysteine synthase isozyme from spinach hydrated seeds. *Bioscience, Biotechnology and Biochemistry* **62**, 501-507.

Yarlett, N., Orpin, C.G., Munn, E.A., Yarlett, N.C. and Greenwood, C.A. (1986) Hydrogenosomes in the rumen fungus *Neocallimastix patriciarum*. *Biochemical Journal.* **236**, 729-739.

Yarlett, N. (1988) Polyamine biosynthesis and inhibition of *Trichomonas vaginalis*. *Parasitology Today* **4**, 357-360.

Zheng, L., White, R.H., Cash, V.L., Jack, R.F. and Dean, D.R. (1993) Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis. *Proceedings of the National Academy of Sciences, United States of America* **90**, 2754-2758.

Zheng, L., White, R.H., Cash, V.L. and Dean, D.R. (1994) Mechanism for the dusulfurization of L-cysteine catalyzed by the *nif*S gene product. *Biochemistry* **33**, 4714-4720.