Structural and functional studies of human methionine adenosyltransferases

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by Ben Murray

August 2015

Abstract

S-Adenosylmethionine (SAMe) is the principal methyl donor of the cell and is synthesized via an ATP-driven process by methionine adenosyltransferase (MAT) enzymes. It is tightly linked with cell proliferation in liver and colon cancer. In humans, there are three genes, *MAT1A*, *MAT2A* and *MAT2B*, which encode MAT enzymes. *MAT2A* and *MAT2B* transcribe MATα2 and MATβ enzyme subunits, respectively, with catalytic and regulatory roles. The MATα2β complex is expressed in nearly all tissues and is thought to be essential in providing the necessary SAMe flux for methylation of DNA and various proteins including histones. In human hepatocellular carcinoma *MAT2A* and *MAT2B* genes are upregulated, highlighting the importance of the MATα2β complex in liver disease. The individual subunits have been structurally characterized but the nature of the complex has remained elusive despite its existence having been postulated for more than 20 years and the observation that MATβ is often co-localized with MATα2. Though SAMe can be produced by $MAT(\alpha2)_4$ alone, this thesis shows that the k_{cat} of the MATα2β complex is three- to four fold higher depending on the variants of MAT that participate in complex formation. Using X-ray crystallography and solution X-ray scattering, the first structures are provided of this 258 kDa functional complex both in crystals and solution with an unexpected stoichiometry of 4α2 and 2βV2 subunits. It is demonstrated that the N-terminal regulates the activity of the complex and it is shown that complex formation takes place surprisingly via the C-terminal of MATβV2 that buries itself in a tunnel created at the interface of the $MAT(\alpha2)_2$. The structural data suggest a unique mechanism of regulation and provide a gateway for structure-based drug design in anticancer therapies.

Crystal structures of human MATα2 containing various bound ligands providing a 'structural movie' of the catalytic steps are also presented. High to atomic-resolution structures reveal the structural elements of the enzyme involved in the utilization of substrates, methionine and adenosine, and the formation of the product SAMe. MAT enzymes are also able to produce S-adenosylethionine (SAE) from substrate ethionine. Ethionine, an S-ethyl analogue of the amino acid methionine, is known to induce steatosis and pancreatitis. It is shown here that S-adenosylethionine occupies the active site in a manner similar to SAMe, confirming that ethionine also binds and utilizes the same catalytic site to form the product SAE. Through by gel filtration and small angle x-ray scattering (SAXS) it is shown that the catalytic MAT enzymes exist in multiple oligomeric populations in solution, whilst the regulatory MAT subunit, MATβ, is monomeric. In view of these data and recent crystallographic structures of the MAT enzyme complex, a rationalization of nomenclature for the MAT enzymes and their complexes has been produced.

Acknowledgements

The work of this thesis used the facilities of both the University of Liverpool, UK, and CIC bioGUNE, Bilbao, Spain and was funded by the Biotechnology and Biological Sciences Research Council (BBSRC), the National Institutes of Health (NIH) and the Plan Nacional I+D. I would like to thank the staff and management of Diamond, SOLEIL and ALBA for the provision of the crystallographic facilities at their synchrotron centres.

I would like to thank my four supervisors for their continued support over the past four years, Prof Samar Hasnian, Dr Svetalana Antonyuk, Dr Jose Mato and Dr Adriana Rojas who all added different aspects to my PhD experience and to all who I am extremely grateful. I highly appreciate the assistance of the members of Lab D who, whilst I have been at Liverpool, have supported and entertained me and for this I thank you all. Thank you to Dr Gaz Wright for collecting my first SAXS data and to Dr Fiona Bailey for her assistance in the design and running of DSF experiments and to Dr Changye Sun for his programme to analyse the DSF data.

Thank you to all my friends of the IIB who often provided needed escape from PhD life in the forms of, the pub, dinners, gym classes, partying, or coffee breaks. This especially applies to Jade, my long suffering plant PhD friend, who has made my time in Liverpool so enjoyable and whose friendship I will have for a long time to come. To Eamon and Pam, two friends who dabbled in research during their medical degrees, thank you for all your support throughout the four years especially with pizza, vodka, partying, coffee breaks and dauphinoise potatoes.

Seeing that I spent two years in the Basque country I thought I would write In Spanish. En primer lugar me gustaría agradecer a mi directora de tesis, Adriana Rojas, que sin ella no habría realizado esta tesis. Gracias a ella, he aprendido mucho acerca de lo que

significa hacer ciencia de calidad y estoy agradecido por el alto nivel de dedicación que ha invertido en mí.

Un enorme agradecimiento a Alberto, que siempre me escuchó y apoyo en lo posibles problemas que vivi en Espana y que me hizo ver la solucion. Mi tesis sería más pobre sin su esfuerzo y dedicación. Gracias por hacer mi vida más fácil con ayuda y apoyo y realmente disfruté trabajar contigo. Gracias por conducir en cada viaje al ESRF y sin ninguna queja al respecto!

Por supuesto, todos los miembros pasados y presentes de Rx y RMN que me han proporcionaron con gran asesoramiento técnico, especialmente la Dr. María Lucas y también el Dr. Sebastiaan Van Liempd por su gran ayuda con los experimentos cinéticos, que, sin él no habría sido posible.

Me gustaría agradecer a mis amigos: June, Iker, Maria (mi abuela), Maria, Ander, Igor, Guillermo, Neka, Georgina, Alberto, Bea, y Bego, que siempre me dieron conversación, me hicieron sentir bienvenido, tradujeron al ingles y se hicieron cargo de mi durante mi estacia en Bilbao. Eskerrik asko chicos por todo, especialmente la introducción a los kalimotxo, pinxos y la cultura del país vasco.

Finally I would like to thank my family for all their support and interest in my work and for visiting me whilst I lived in Bilbao.

Table of Contents

List of Figures

List of Tables

Abbreviations

5,10-MTHF	5,10- methyltetrahydrofolate
5-MTHF	5-methyltetrahydrofolate
AMP-PNP	5'-Adenylylimidodiphosphate
ATP	Adenosine triphosphate
BHMT	Betaine Hcy methlytransferase
CBS	Cystathionine β-synthase
CSE	Cystathionase
dcSAMe	Decarboxylated SAMe
DTT	Dithiothreitol
GNMT	Glycine N-methyltransferase
GSH	Glutathione
HCC	Hepatocellular carcinoma
Hcy	Homocysteine
HGF	Hepatocyte growth factor
iNOS	Inducible NO synthase
MAT	Methionine adenosyltransferase
MTA	5'-methylthioadenosine
MTAP	MTA phosphorylase
MTHFR	Methylenetetrahydrofolate reductase
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NO	Nitric oxide
NSD	Normalised spatial discrepancy
PEG	Polyethylene glycol
Pi	Orthophosphate
PPi	Pyrophosphate
PPPi	Triple phosphate
R.M.S	Root-mean-square
ROS	Reactive oxygen species
rpm	Revolutions per minute

SAE	S-adenosylethionine
SAH	S-adenosyl-L-homocysteine
SAMe	S-adenosylmethionine
SAXS	Small angle X-ray scattering
SDR	Short-chain dehydrogenase/reductase
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SPD	Spermidine
SPM	Spermine
SRE	SAMe radical enzyme
TCEP	Tris(2-carboxyethyl)phosphine
THF	Tetrahydrofolate

Chapter I. Introduction

1.1 Introduction to SAMe

Giulio Cantoni, in 1952, discovered S-Adenosylmethionine (SAMe) by demonstrating that adenosine triphosphate (ATP) and methionine were able to react to form a new molecule, which he identified and named SAMe. He showed that SAMe was able to transfer its methyl group to guanidinoacetic acid and nicotinamide to form creatine and N-methylnicotinamide, respectively (Cantoni 1952). Since the 1950s, SAMe has been extensively studied and has been implicated in thousands of different methylation reactions and to date SAMe competes with ATP as the most widely used enzyme substrate (Cantoni 1952; Loenen 2006).

SAMe (Fig 1.1) emerged as the foremost donor of methyl groups in methylation reactions early in evolution and many molecules contain methyl groups, such as DNA, RNA, proteins and sugars, that have been derived from SAMe. SAMe is synthesised from ATP and the amino acid L-methionine in a two-step SN2 reaction where the triple phosphate moiety is cleaved from ATP as SAMe is being formed. SAMe is a typical sulphonium salt that contains a sulphur with three single covalent bonds and, therefore, an overall positive charge (Bentley 2005). The sulphonium ion can transfer its methyl group to a vast range of nucleophilic atoms through a typical SN_2 substitution reaction resulting in the release of S-adenosyl-L-homocysteine (SAH) (Fig 1.1) (Hegazi et al. 1976; Scavetta et al. 2000). In methylation reactions, dependent on SAMe, the nucleophile can be a nitrogen atom, as in the N-methylation reaction of the amino group of glycine methylation; a carbon atom as in DNA methylation of C5-cytosine; an oxygen atom of a hydroxyl moiety; or a sulphur atom in the methylation of thiopurines (Lin et al. 2002).

There are two chiral centres in SAMe, a chiral carbon at the α amino position and the other chiral is a sulphur at the sulphonium centre, therefore, SAMe has two diastereomeric forms. With respect to the sulphonium ion the forms are (S,S)-SAMe and (R,S)-SAMe, the former is the only enzymatically synthesised form, which is used in almost all methylation reactions (Beaudouin et al. 1993; Bentley 2005; Cannon et al. 2002; de la Haba et al. 1959). Under physiological conditions (i.e pH 7.5 and 37 °C) the sulphonium ion of (S,S)-SAMe can racemize to form its R counterpart, which is biologically inactive (Hoffman 1986), and this accounts for 1.5 % and 3 % of total SAMe levels in mouse liver and rat brain, respectively (Beaudouin et al. 1993; Hoffman 1986).

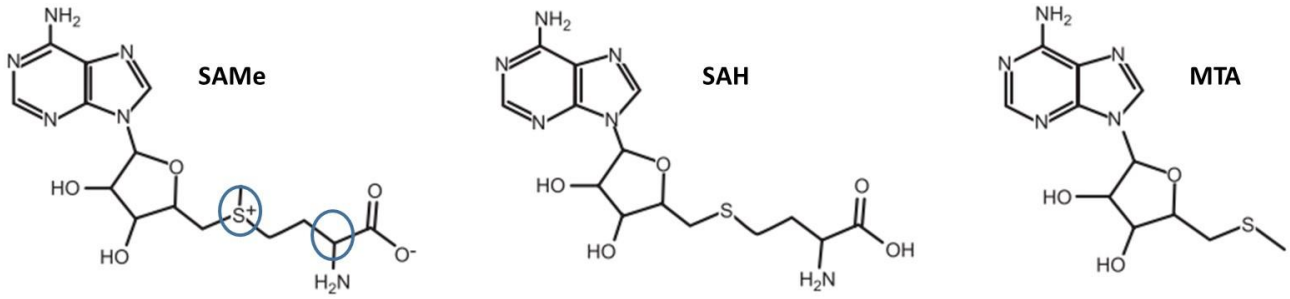

Figure 1.1 Structure of S-adenosyl-L-methionine and its derivatives. The Structure of S-adenosyl-L-methionine (SAMe), S-adenosyl-L-homocysteine (SAH) and 5'-methylthioadenosine (MTA). S-adenosyl-L-homocysteine is formed when the methyl group from the sulphonium ion of SAMe is transferred. 5'-methylthioadenosine is formed when SAMe is spontaneously broken down into 5'-methylthioadenosine and homoserine lactone. The two chiral centres in SAMe are shown by blue rings. Adapted from Lu & Mato, 2012.

There are two spontaneous reactions that SAMe can undergo within the cell: the first is the production of 5'-methylthioadenosine (MTA) (Figure 1.1) and homoserine lactone through an intramolecular cleavage and the second spontaneous reaction is where the hydrolysis of SAMe at the glycosidic bond produces adenine and S-pentosylmethionine (Hoffman 1986; Wu et al. 1983). These two reactions contribute to the control of the level of (S,S)-SAMe within the cell and stops the accumulation of the biologically inactive form. Non-mammalian systems contain certain methyltransferases including Mht-1 (S-methylmethionine-homocysteine methlytransferase) and Sam4 (SAMe-homocysteine methyltransferase) that are able to process (R,S)-SAMe and control levels of this racemic form (Vinci & Clarke 2007). Mammalian systems have to cope with levels of (R,S)-SAMe through other means, as they lack homologs to Mht-1 and Sam4 (Vinci & Clarke 2010).

1.2 SAMe Biosynthesis

This production of SAMe is catalysed by the enzyme methionine adenosyltransferase (MAT, EC 2.5.6.1) (Kotb et al. 1997). With the exception of certain parasites and infectious agents, that obtain SAMe from their hosts (Merali & Clarkson 2004; Stephens et al. 1998), all living cells express MAT and so the *mat* gene is essential to sustaining life (Glass et al. 2006). In mammals, all cells studied to date express MAT with the greatest levels of expression found in the liver (Finkelstein 1990) and so not surprisingly, individuals with hepatic MAT deficiency present with hypermethioninemia, as methionine is unable to be cleared from the liver (Ubagai et al. 1995). Compared to control individuals that contain ~35 μM methionine, patients with hepatic MAT deficiency show levels of methionine of ~1300 μM (Ubagai et al. 1995). SAMe synthesis has been studied using MAT from *E.coli* and rat liver (Sanchez del Pino et al. 2000; Markham & Pajares 2010) and will be described in Section 1.8. SAMe is tightly linked to four metabolic pathways: a) transmethylation, b) transsulphuration, c) polyamine synthesis and d) as a mediator for radical chemical reactions.

1.2.1 Transmethylation

Although SAMe is made in the cytosol of all mammalian cells, it is the liver where it is produced the most, where half of the daily intake of methionine is metabolised, and where 85 % of methylation reactions take place (Mato & Lu 2007; Mudd & Poole 1975). Due to the nature of SAMe, it is able to donate its methyl group to a vast array of acceptor molecules, which are all catalysed by methyl transferases and in fact there are over 200 SAMe-dependent methyl transferase proteins (Petrossian & Clarke 2011). Methyl transferases transfer the methyl group from SAMe to oxygens, nitrogens, and carbon atoms on proteins, DNA, small molecules, lipids and carbohydrates. The main methyl

transferase in the liver, which accounts for 1 % of liver cytosolic protein, is glycine N-methyltransferase (GNMT) (Lu et al. 2000; Stipanuk 2004). *GNMT* knockout mice develop hepatocellular carcinoma and steatohepatitis, whilst human patients with genetic GNMT deficiency have elevated plasma methionine and SAMe levels and also develop liver injury and hepatomegaly (Augoustides-Savvopoulou et al. 2003; Mudd et al. 2001).

SAMe-dependent methylation reactions cause the generation of SAH, which is a potent inhibitor of methylation reactions as its inhibition constant (K_i) is in the same micromolar range as the Michaelis constant (K_m) of most SAMe-dependent methyl transferases (Chiang 1998; Chiang et al. 1996). An important metabolic gauge in controlling methylation reactions *in vivo* is the ratio between SAMe and SAH whereby an increase in this ratio increases the methylation capacity of the cell (Finkelstein 2007). Levels of SAH are regulated tightly by SAH hydrolase that through a reversible cleavage reaction turns SAH into homocysteine and adenosine (Figure 1.2) (Della Ragione & Pegg 1983). SAH synthesis is favoured over hydrolysis due to the thermodynamics of SAH hydrolase (Finkelstein 2007). *In vivo*, adenosine and homocysteine can be rapidly removed allowing for the hydrolysis of SAH which is an essential process to prevent the accumulation of SAH (Hoffman et al. 1980; Mato et al. 1997). Any impairment in homocysteine catabolism will lead to SAH accumulation and this results in the inhibition of methyl transferases. homocysteine can be remethylated in the liver by two different enzymes: betaine homocysteine methlytransferase, which requires betaine, and methionine synthase, an enzyme that requires 5-methlytetrahydrofolate which is converted to tetrahydrofolate once it donates its methyl group (Figure 1.2). 5-methlytetrahydrofolate production is catalysed by methylenetetrahydrofolate reductase that converts 5,10-methlytetrahydrofolate to 5- methlytetrahydrofolate.

1.2.2 Transsulphuration

Homocysteine can also be processed into cysteine via the transulphuration pathway that links SAMe synthesis to cysteine production (Figure 1.2). Here, cystathionine β-synthase converts homocysteine to cystathionine before cystathionase converts it into cysteine in a two-step enzymatic process and both reactions require vitamin B_6 (Figure 1.2). Cysteine is an important precursor for the formation of glutathione, which is a powerful antioxidant. In the liver the transsulphuration pathway is highly active highlighting SAMe as an important precursor for glutathione (Lu 2009). Betaine homocysteine methlytransferase expression is limited to liver and kidney whilst both MAT and methionine synthase enzymes, which are important as SAMe inhibits methylenetetrahydrofolate reductase and activates cystathionine β-synthase, are expressed ubiquitously (Mato et al. 1997; Prudova et al. 2006). This allows the level of SAMe to control the flow of homocysteine processing, whether it is remethylated or if it is funnelled into the transsulphuraton pathway when levels of homocysteine are high. Mice which have methylenetetrahydrofolate reductase or betaine homocysteine methlytransferase deficiencies have increased levels of homocysteine and SAH and reduced SAMe (Finkelstein 2006).

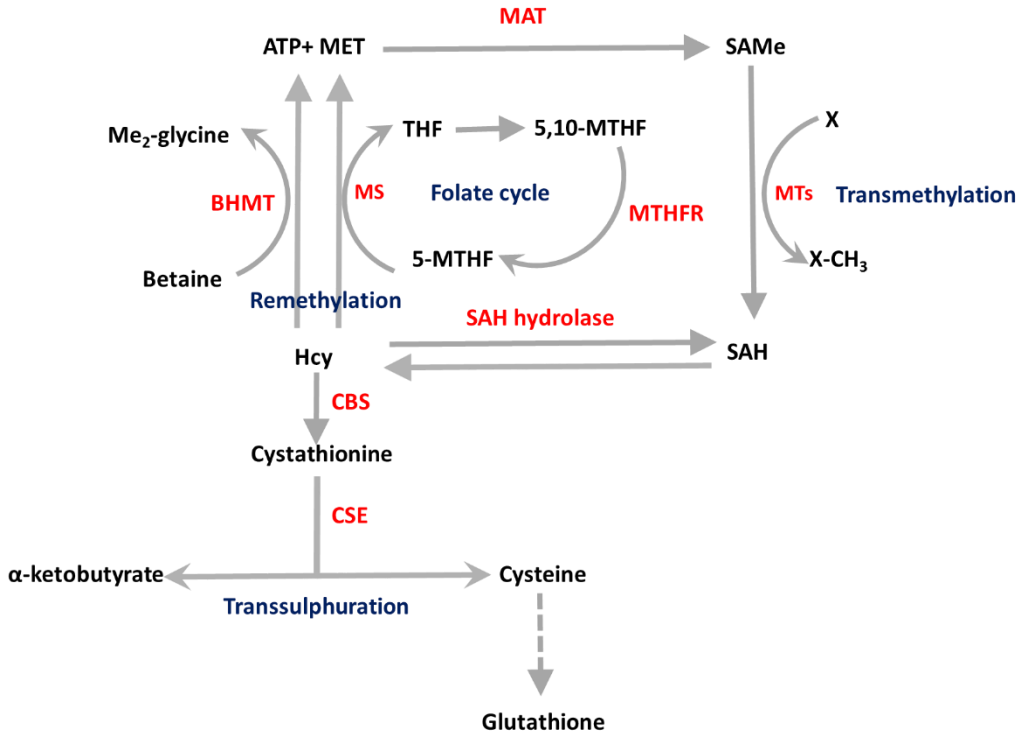

Figure 1.2. Methionine metabolism. Methionine adenosyltransferase (MAT) catalyses the conversion of methionine and ATP into SAMe. In transmethylation reactions, SAMe donates its methyl group to a large variety of acceptor molecules (X) in reactions catalysed by methyl transferases (MTs). SAMe-dependent methylation reactions yield SAMe as a by-product. SAMe cellular content is regulated by the enzyme SAH hydrolase which in a reversible reaction cleaves SAH into adenosine and homocysteine (Hcy). Remethylation of homocysteine to form methionine occurs by two enzymes: methionine synthase (MS), which requires normal levels of folate and vitamin B_{12} and betaine homocysteine methyltransferase (BHMT), which requires betaine. MS-catalysed homocysteine remethylation requires 5-methyltetrahydrofolate (5-MTHF), which is derived from 5,10-methylenetetrahydrofolate (5,10-MTHF) in a reaction catalysed by methylenetetrahydrofolate reductase (MTHFR). 5-MTHF is then converted to tetrahydrofolate (THF), as it donates its methyl group and THF is converted to 5,10-MTHF and this completes the folate cycle. The transsulphuration pathway links SAMe to cysteine biosynthesis. Homocysteine is converted to cysteine via a two-step enzymatic process catalysed firstly by cystathionine-β-synthase (CBS) and then cystathionase (CSE), both requiring vitamin B6. α-Ketobutyrate, the other product of cystathionine cleavage, is further metabolised by the mitochondria through the Kreb's cycle.

1.2.3 Polyamine synthesis

SAMe is also involved in the synthesis of polyamines which are low-molecular weight molecules that are positively charged and ubiquitously found in all living cells (Perez-Leal & Merali 2012). Many biological processes, including transcription, translation, cell growth and apoptosis, require polyamines highlighting their crucial role within the cell. In order to produce polyamines from SAMe, SAMe must first be decarboxylated by SAMe decarboxylase to produce decarboxylated SAMe. Spermidine and spermine are the predominant polyamines in mammalian cells and are produced by the sequential addition of aminopropyl groups to decarboxylated SAMe using the enzymes spermine synthase and spermidine synthase. These reactions produce 5'-methylthioadenosine as a by-product (Fig. 1.1), which interestingly is able to inhibit both spermine synthase and spermidine synthase with a K_i of 0.3 μM and 2-10 μM, respectively (Hibasami et al. 1980; Pajula et al. 1979; Pajula & Raina 1979; Raina et al. 1982). Fortunately for the cell this inhibition has little effect as 5'-methylthioadenosine is rapidly metabolised, through the methionine salvage pathway, which enables SAMe to be regenerated (Pegg & Williams-Ashman 1969). Here, 5'-methylthioadenosine is cleaved to give rise to adenine and 5-methylthioribose-1-phosphate by the enzyme 5'-methylthioadenosine phosphorylase and the latter product is further processed into methionine, which can be used by MAT enzymes to give SAMe (Figure 1.3).

1.2.4 SAMe radical reactions

Another function of SAMe is that it is able to mediate radical chemical reactions in organisms that are grown anaerobically (Knappe & Schmitt 1976). SAMe mediates radical chemical reactions for over 2800 radical enzymes that contain a [4Fe-4S] cluster,

which SAMe binds to and this results in a reductive cleavage that generates [4Fe-4S]-methionine and 5'-deoxy-adenosyl 5'-radical (Figure 1.3) (Frey et al. 2008). The latter molecule is able to remove a hydrogen atom from proteins, RNA, DNA or small molecules initiating a radical mechanism (Frey et al. 2008). An example of a SAMe radical enzyme is the interferon-inducible antiviral protein, viperin, whose expression is used as a predictive response in the treatment of chronic hepatitis C infection (Sarasin-Filipowicz et al. 2008).

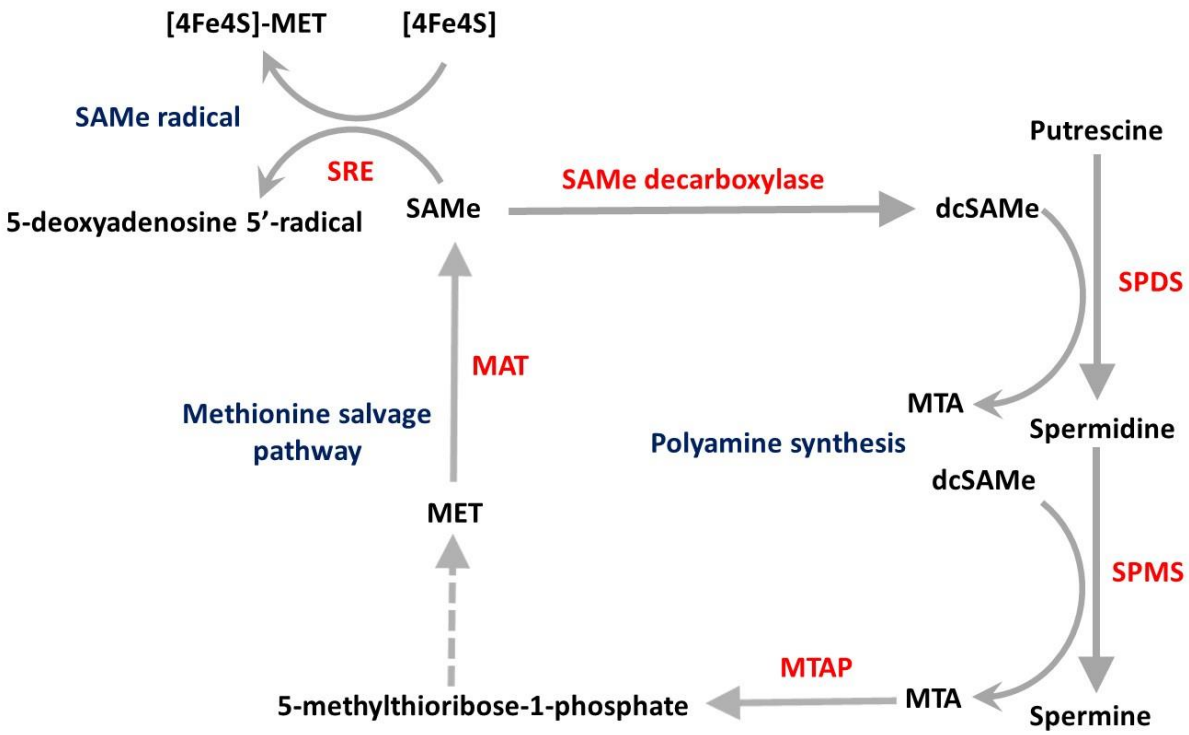

Figure 1.3. Polyamine synthesis, methionine salvage pathway, and SAMe radical reactions. To synthesise polyamines, SAMe is decarboxylated, by the enzyme SAMe decarboxylase, to form decarboxylated SAMe (dcSAMe). Common polyamines in mammalian cells are spermidine (SPD) and spermine (SPM). These polyamines are made by sequential addition of aminopropyl groups from dcSAMe, giving rise to 5'-methylthioadenosine (MTA) as a by-product. SPD synthase (SPDS) catalyses the transfer of the first aminopropyl group from dcSAMe to putrescine to form SPD and MTA. SPM synthase (SPMS) catalyses the transfer of the second aminopropyl group to SPD to form SPM and a second molecule of MTA. MTA is metabolised through the methionine salvage pathway to regenerate SAMe. In this pathway cleavage of MTA by the enzyme MTA phosphorylase (MTAP) occurs producing adenine and 5-methylthioribose-1-phosphate, which is further metabolised to methionine and adenine to AMP. SAMe may be also converted to 5'-deoxyadenosyl 5'-radical by SAMe radical enzymes (SRE) which have a CX_3CX_2C motif forming a characteristic [4Fe-4S] cluster. This [4Fe-4S] cluster binds SAMe catalysing its reductive cleavage to generate [4Fe-4S]-methionine and a 5'-deoxyadenosyl 5'-radical.

1.3 Effect of SAMe in the cell

An effect that SAMe has within the cell is that it regulates cellular proliferation. Using models of partial hepatectomy, mechanisms of how levels of SAMe induces cellular pathways for the regeneration of the liver have been studied and elucidated. The effect SAMe has on proliferation is cell specific, for example during lymphocyte activation intracellular levels of SAMe increase five-fold (Halim et al. 1999) and this is related to the need of polyamines for the activation of lymphocytes. In liver, this is not the case and levels of SAMe reflect the state of cell, whether it is proliferating or quiescent and thus having low or high levels of SAMe respectively (Martínez-Chantar et al. 2003; Cai et al. 1998). During partial hepatectomy, the hepatic SAMe level dramatically drops, which coincides with induction of early response genes and DNA synthesis (Huang et al. 1998). The *MAT1A* knock out mouse model, displays increased basal hepatocyte proliferation, but impaired regeneration following partial hepatectomy (Chen et al. 2004).

The molecular mechanisms of how SAMe inhibits growth in hepatocytes has been deduced. SAMe inhibits both hepatocyte growth factor, which is a mitogenic growth factor essential for liver generation (Latasa et al. 2001) and leptin (Ramani et al. 2008). Both hepatocyte growth factor and leptin have been shown to be needed for liver cell proliferation by activating *MAT2A* transcription (Latasa et al. 2001; Pañeda et al. 2002; Ramani et al. 2008). In order for the mitogenic effects of hepatocyte growth factor to take effect, hepatic parenchymal cells need to be primed to be able to react to hepatocyte growth factor which is accomplished by tumour necrosis factor (TNF)-α and interleukin-6 (Cressman et al. 1996; Yamada et al. 1997). Both TNF-α and interleukin-6 are inflammatory cytokines that promote the activation of STAT-3, NFκB, activator protein 1 and CCAAT/enhancer binding protein β, which leads to the activation of genes in

hepatocytes and subsequent DNA synthesis and cell proliferation (Varela-Rey et al. 2011).

One of the genes that is upregulated is inducible nitric oxide synthase (iNOS) and increased levels of nitric oxide is essential for liver repair and for mice lacking this enzyme they display impaired liver regeneration (Rai et al. 1998) and interestingly, hepatocyte growth factor requires iNOS activity and a drop in the levels of SAMe to function (García-Trevijano et al. 2002). Moreover, nitric oxide is a specific inhibitor of MATα1 but not MATα2 and is able to nitrosylate residues causing the inactivation of MATα1 (Lu & Mato 2012). Taken together these observations paint a picture whereby following an attack on the liver, such as a partial hepatectomy, an influx of inflammatory cytokines induces the expression of iNOS and allows for the increased production of nitric oxide. The elevated levels of nitric oxide cause the inactivation of MATα1 facilitating a fall in SAMe and the knock-on effect of releasing inhibition of growth factors. Lower levels of SAMe causes reduction in the methylation and activation of phosphatase 2A, which reduces the dephosphorylation and deactivation of serine/threonine kinase 11 and AMP-activated protein kinase that are kinases stimulated by hepatocyte growth factor (Vázquez et al. 2010). This is supported by the *MAT1A* knock out mouse phenotype that does not display decreased SAMe levels following partial hepatectomy and which are resistant to the mitogenic effects of HGF (Chen et al. 2004).

1.4 Methionine adenosyltransferases-Overview of genes and proteins.

In mammals there are two genes, *MAT1A* and *MAT2A*, which produce the catalytic MAT proteins, MATα1 and MATα2 (Mudd 1962). The *MAT1A* gene encodes a 395 amino acid catalytic subunit (MATα1) that shares a 84 % sequence homology with MATα2 (Kotb et al. 1997) and is predominantly expressed in the adult liver only (Gil et al. 1996) within parenchymal cells, liver endothelial, hepatocytes and Kupffer cells (Shimizu-Saito et al. 1997). MATα1 can exist in two functional forms, as a dimer (MATIII) or a tetramer (MATI) (Kotb et al. 1997). The *MAT2A* gene encodes a 395 amino acid catalytic subunit (MATα2) expressed in all extrahepatic tissue including erythrocytes, lymphocytes, kidney, brain, and fetal liver and is expressed in very low levels in the liver of adults (Kotb et al. 1997; Mato et al. 1997; Kotb & Kredich 1985; Gil et al. 1996). Using phenyl sepharose columns the protein products of *MAT1A* and *MAT2A* were separated with different hydrophobic characteristics into three peaks which were named MATI, MATII and MATIII for the order of elution (Sullivan & Hoffman 1983).

MAT promoters have been cloned sequenced and studied giving insights into expression patterns of these genes (Hiroki et al. 1997; LeGros et al. 2001; Alvarez et al. 1997; Mao et al. 1998; Zeng et al. 2000). In the developing liver *MAT1A* expression increases progressively during gestation and, after birth, it increases 10-fold and peaks at 10 days. A slight decrease occurs after 10 days to a level that is maintained for the adult life (Gil et al. 1996). On the other hand, *MAT2A* expression decreases towards birth and continues to do so postnatally until minimal levels are reached in the adult (Gil et al. 1996). Under normal cellular conditions MATα2 does not contribute notably to the metabolism of hepatic methionine compared to MATα1. The developmental pattern of MAT enzymes is closely related to both albumin and α-fetoprotein for hepatocyte differentiation (Gil et

al. 1996) and expression of *MAT1A* can be used as a marker of liver differentiation (Mato et al. 1997). The *MAT1A* promoter contains two CpG sites that in are hypermethylated in extrahepatic tissue and in fetal liver and this becomes unmethylated, allowing transcription in adult liver (Torres, López-Rodas, et al. 2000; Torres, Avila, et al. 2000). Furthermore, histones that are associated with *MAT1A* promoter in the liver are acetylated ~15 fold higher than those in the kidney (Torres, Avila, et al. 2000). Looking at the promoter for *MAT2A* in different cell types no differences were found between that of the liver, where low levels of expression are observed, and other tissues where the gene is transcribed, suggesting that methylation of *MAT2A* promoters seems to be uniform regardless of tissue type (M. Avila et al. 2000). During hepatocellular carcinoma the human *MAT2A* promoter becomes hypomethylated and this allows for increased expression (Latasa et al. 2001).

The *MAT2B* gene encodes four isoforms MATβV1, MATβV2, MATβV2a and MATβV2b with the first two being the dominant ones. MATβV1 and MATβV2 share very high sequence identity, but differ in their number of N-terminal residues (Yang et al. 2008). MATβV2a and MATβV2b are expressed at very low levels compared to MATβV2 and have not been studied in detail. The two dominant forms are differentially expressed in normal tissues with MATβV1 expressed in prostate, lung, fetal liver, brain, thyroid and the adrenal gland, which is opposed to MATβV2 that is expressed in skeletal muscle and heart. Both isoforms are expressed in the kidney and thymus (Yang et al. 2008). MATα2 and MATβ are able to interact to form the MATαβ complexes and binding of the regulatory subunit, MATβ, regulates the activity of MATα2 by lowering the K_i for SAMe and the K_m for methionine (Halim et al. 1999; LeGros et al. 2000). The oligomeric form of these MATαβ complexes remains in question and several groups have tried to

work out the composition of the complexes. Two models, a tetrameric ($MAT(\alpha2\beta)_2$) (Kotb & Kredich 1985; Mitsui et al. 1988; Langkamp-henken et al. 1994) and trimeric ($MAT(\alpha2)_2\beta V1$) (González et al. 2012) assemblies have been proposed. For this thesis the nomenclature shown in Figure 1.4 will be used to describe MAT proteins.

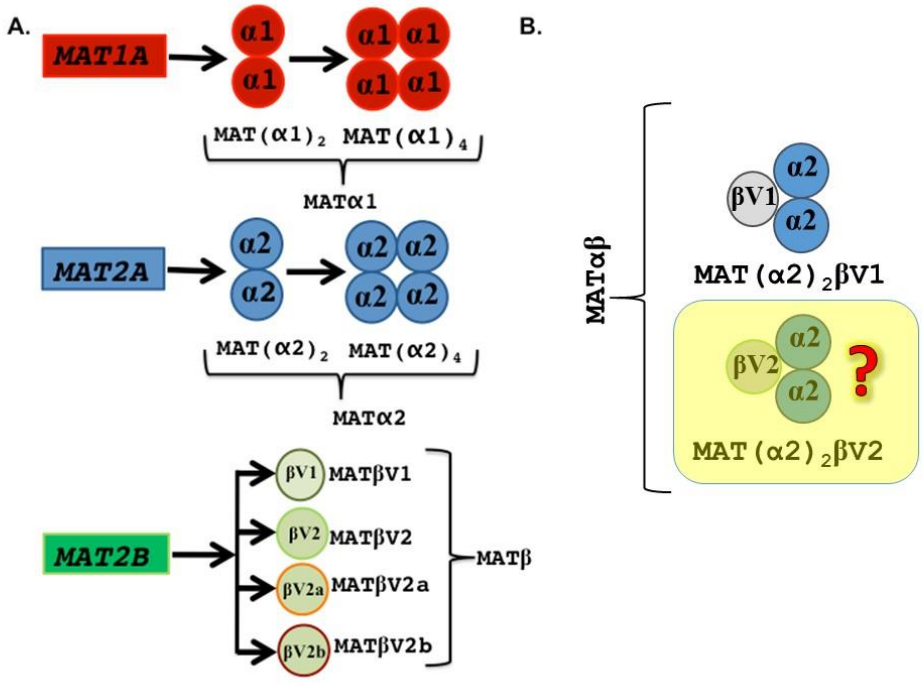

Figure 1.4. Nomenclature of mammalian MATs. A. The mammalian genes *MAT1A*, *MAT2A* and *MAT2B* encode the catalytic subunits MATα1, MATα2 and the regulatory subunit MATβ respectively. MATα1 and MATα2 subunits can be organised as dimers and tetramers. On the other hand, *MAT2B* encodes for four isoforms, MATβV1, MATβV2, MATβV2a, MATβV2b. **B.** Proposed model of MATαβ complexes. Experimental evidence from Gonzalez *et al.* 2012 suggested that a MATαβ complex is composed of two MATα2 and one MATβV1.

1.5 Post translational modifications of MATs.

Estimates of MATα1 subunit size, as determined by SDS-PAGE indicated a molecular weight of 48 kDa (Cabrero et al. 1988), which was slightly larger than the predicted weight from sequence of 45 kDa. This observation led to the suggestion that MATα1 is modified post-translationally and further analysis of the MATα1 sequence revealed phosphorylation sites for several kinases, including PKC, which has been explored *in vitro* (Pajares et al. 1994). MATα2 was purified from human chronic lymphocytic leukemia cells to apparent homogeneity, but showed three bands on SDS-PAGE, at sizes of 53 kDa, 51 kDa and 38 kDa with the latter corresponding to MATβ (Kotb & Kredich 1985) giving further evidence that MAT proteins are subjected to post translational modification.

Nitric oxide and reactive oxygen species (ROS) are able to inactivate MATα1 through either the S-nitrosylation or oxidation of Cys120. A single MATα1 monomer contains 10 cysteine residues and mutagenesis studies have revealed that replacing cysteine with serine (C120S) yields a MATα1 that cannot be deactivated by NO or ROS (Avila et al. 1997; Sánchez-Góngora et al. 1997). Cys120 is one of the residues that makes up the gating loop, the flexible loop that facilitates access to the active site, and modification of this single cysteine residue is sufficient to inactivate MATα1 (Pérez-Mato et al. 1999; Sánchez-Góngora et al. 1997). MATα2 contains a glycine residue at position 120 rather than a cysteine and is not effected by NO or ROS. Basic and acidic residues (Asp354, Arg356 and Arg362), surround Cys120 and reduce the acid dissociation constant (pKa) of the thiol group belonging to cysteine enabling it to become more nucleophilic (Stamler et al. 1997) that can be S-nitrosylated. Modifications of these surrounding residues

reduces the ability of nitric oxide to S-nitrosylate MATα1 (Pérez Mato et al. 2001), and these residues are conserved in MATα2. Mutating glycine 120 to a cysteine causes MATα2 to become sensitive to nitric oxide and ROS induced inhibition (Castro et al. 1999).

Aligning multiple sequences of MATs from bacteria, yeast, plants, flies, mouse, rat and humans reveals that the Cys120 is a feature of liver specific enzymes whilst other tissues favour different amino acids and in particular, glycine. Inactivation of MATα1 by nitric oxide or ROS can be reversed by glutathione (Avila et al. 1997; Sánchez-Góngora et al. 1997; Corrales et al. 1999), at millimolar concentrations which differ depending on the oligomeric form of MATα1 ($MAT(\alpha1)_4$= 3 mM, $MAT(\alpha1)_4$= 25 mM). Normal hepatic glutathione levels are between 5-10 mM which may contribute to the selective loss of $MAT(\alpha1)_4$ in liver diseases (Corrales et al. 1990). When glutathione synthase is inhibited a reduction in the SAMe level and increased MATα1 activation is observed (Corrales et al. 1999). Taken together these mechanisms of MAT regulation may be the cause of hepatic MATα1 inactivation, in conditions of oxidative stress, such as septic shock, liver cirrhosis and hepatitis B or in conditions of increased nitric oxide synthesis like during hypoxia and liver regeneration (Mato et al. 1997; Avila et al. 1997; Avila et al. 1998; Martin-Duce et al. 1988; Lu et al. 2000). It has been shown that MATα1 can be phosphorylated by protein kinase C (PKC) but this modification does not change the kinetic properties of MATα1 and the role of this phosphorylation in the regulation of MAT remains unclear (Pajares et al. 1994). Whether MATα2 is regulated by phosphorylation is unknown.

In recent years, it has been speculated that MAT proteins provide a continuous source of nuclear SAMe, as it has been reported that both MATα1 and MATα2 can be located in the nucleus (Reytor et al. 2009; Katoh et al. 2011). To support this further, accumulation of MATα1 in the nucleus coincides with higher levels of histone H3K27 trimethylation, which is a known epigenetic gene repression mechanism (Reytor et al. 2009). MATβ variants are also localised to the nucleus and can stabilise proteins such as HuR, the RNA binding protein. Katoh *et al* reported that MATαβ complex resides in the nucleus where it serves as a transcriptional corepressor complex with MafK and that this complex provides SAMe needed for methylation of DNA and histones, participating in chromatin based regulation.

It has been shown that, in liver cancer, MATα2 can be acetylated and directed for degradation by ubiquitylation. This occurs when folate levels are insufficient enabling P300 to acetylate MATα2 at Lys81, which in turn promotes UBR4-mediated ubiquitylation, leading to its proteasomal degradation and reduced cell growth. Conversely, when folate levels are sufficient MATα2 is stabilised by deacetylation via HDAC3 which inhibits ubiquitylation and in turn allows MATα2 to accumulate and enhances tumour growth (Yang et al. 2015).

1.6 MAT associated diseases.

The changes in both MAT activity and expression have been linked through numerous studies, to many types of disease. These changes correspond to altered levels of SAMe and, therefore, affect the many reactions and pathways that this molecule is involved in. Given this, a direct relationship between MAT and the development of a disease can be tricky to establish. One of the most well established MAT related diseases is in cancers such as hepatocellular carcinoma and colon cancer. Cancer is a major cause of death in the first world whose basis depends on both environmental and genetic factors that cause anomalous cell growth through dysregulation of the cell cycle. Furthermore, a property of human cancer cells is that they depend on exogenous methionine for growth whereas in a healthy cell homocysteine is metabolised (Halpern et al. 1974). Studying the methionine cycle has revealed roles in carcinogenesis of several diseases and the MAT family of proteins plays a role in driving the development of hepatocellular carcinoma.

Data from several studies indicates that a reduction in *MAT1A* mRNA, protein and activity (Simile et al. 2005; Huang et al. 1999; Liang et al. 2005; Okada et al. 1979) coincides with the increase in the same parameters in *MAT2A* (Huang et al. 1999; Horikawa et al. 1993; Liau et al. 1979). Other studies have shown that different types of tumour cell lines and tumours (Liau et al. 1979; Abe & Tsukada 1981) have shown rapid growth, increased levels of MATα2, decreased MATα1 and moreover, this expression pattern has been characterised in liver diseases such as HCC where the expression patterns of MAT proteins reverts back to a fetal form (Table 1.1).

Table 1.1. Summary of MAT expression in liver in healthy and disease states (Mato & Lu 2007).

MAT gene	Protein	Isoform	Expression pattern in healthy adult liver	Expression during HCC
MAT1A	MATα1	$MAT(\alpha1)_2$	High	Low
		$MAT(\alpha1)_4$	High	Low
MAT2A	MATα2	$MAT(\alpha2)_2$	Low	High
		$MAT(\alpha2)_4$	Low	High
MAT2B	MATβ	MATβV1	Low	High
		MATβV2	Low	High
		MATβV2a	Low	Low
		MATβV2b	Low	Low

Global hypomethylation seems to be a common property of several diseases where *MAT1A* expression and SAMe levels are altered (Yang et al. 2001). The role of MAT proteins in HCC has been studied for many years and the role that SAMe plays in cancer development was confirmed when the *MAT1A* knockout mouse (Martínez-Chantar et al. 2002) showed that a chronic deficiency in SAMe lead to the spontaneous appearance of HCC. Interestingly, treatment with SAMe has been shown to be efficient in treating HCC (Garcea et al. 1989) as it inhibits the mitogenic effect of HGF by blocking NFκB activation of cytokine-stimulated hepatocytes (Majano et al. 2001). If antisense *MAT2A* RNA is used in normal or *MAT2A* overexpressing hepatoma cells the result is cell death (Cai et al. 1998). The expression of the regulatory subunit MATβ provides a growth advantage to human hepatoma cells, and the expression of these proteins is common in cancer and readily associated with hepatic dysfunction (Martínez-Chantar et al. 2003).

Ethanol plays an interesting role in the development of liver disease. Combining a high fat diet and intragastric feeding of ethanol in rats increases the mRNA levels of both *MAT1A* and *MAT2A*, but only the protein levels of MATα2 rose noticeably (Lu et al. 2000). Ethanol induction of *MAT2A* is accomplished by the stimulation of TNF-α, which

causes NFκB and AP-1 to bind to the *MAT2A* promoter (Yang et al. 2001). A reduction in the level of SAMe has been observed during the ethanol feeding of rats, baboons, micropigs and in patients who suffer from ethanol hepatitis (Lu et al. 2000; Villanueva & Halsted 2004; Lieber et al. 1990; Lee et al. 2004). However a reduction in SAMe does not always translate to same level of reduction in SAH and differences in the levels of SAH have been observed between these models. Overall there is a net reduction in the SAMe/SAH ratio that correlates with a decrease in global DNA methylation whilst there is an increase in both c-Myc expression and genome-wide DNA strand breaks (Lu et al. 2000).

Alcoholic liver cirrhosis can be a precursor for liver cancer and samples from alcoholic hepatitis and cirrhotic patients saw a reduction in levels of *MAT1A* RNA but no induction of *MAT2A* (M Avila et al. 2000; Lee et al. 2004). Samples from cirrhotic patients (84 %) showed expression of MATβ, but the effect of this protein has not been explored (Martínez-Chantar et al. 2003). Liver damage due to both short and long term ethanol consumption has been linked to a reduced supply of oxygen to the liver and such sustained hypoxia induces the down regulation of *MAT1A* expression in both liver cirrhosis and long term ethanol consumption (Arteel et al. 1997; Ji et al. 1982; Tsukamoto & Xi 1989). This damage may be irreversible as methylation reactions and glutathione-dependent detoxification capacities are impaired and also by changing levels of glutathione and nitic oxide which can activate or deactivate MATα1, respectively (Avila et al. 1998; Corrales et al. 1999).

Transmethylation pathways have roles in several neurodegenerative conditions such as Parkinson's and Alzheimer's disease (Gomes Trolin et al. 1995; Gomes Trolin 1998; Surtees et al. 1991; Cheng et al. 1997) and also in psychiatric disorders such as dementia and schizophrenia (Regland et al. 1995). Studies in mice, monkeys and pigs have linked the development of myelopathies, ataxia, subacute combined degeneration of the spinal cord and peripheral neuropathy to defects in MAT (Lee et al. 1992; Scott 1992; Scott et al. 1994). Although an association with MATs and Parkinson's diseases exist no link between the levels of SAMe and/or SAH with the duration or age of Parkinson's has been described. There are elevated levels of MAT in erythrocytes of patients who suffer from Parkinson's and schizophrenia but they also have lower levels of SAMe and SAH in the blood (Cheng et al. 1997; Smythies 2012).

1.7 Structural basis of MAT proteins.

MAT enzymes are an exceptionally well conserved family of enzymes found throughout all walks of life and comparing sequences from *E.coli*, *Mycoplasma genitalium*, *Saccharomyces cerevisiae*, *Arabidopsis thaliana*, mouse, rat and human highlights their similarity (Mato et al. 1997). *E.coli* and human MAT share a 59 % sequence identity (Alvarez et al. 1991).

Crystal structures of various MAT enzymes have been published for rat MATα1 (González et al. 2000), *E.coli* MAT (eMAT) (Komoto et al. 2004; Fu et al. 1996; Takusagawa et al. 1996) and human MATα1 and MATα2 (Shafqat et al. 2013). Due to the high sequence identity, the overall topology is the same for each protein. A MAT monomer consists of three repeats of beta-alpha-beta-alpha-beta secondary structures, which are related to each other by pseudo threefold symmetry (Fig. 1.5A) and each monomer contains three domains: N-terminal, central and C-terminal domains. Pairs of subunits form dimers and each dimer forms a tight complex that creates two active sites using residues from each monomer (Fig 1.5B). MAT dimers can oligomerise to form a tetrameric state at right angles to their central tip (Fig 1.5C). The hydrophobic interface between MAT monomers is much greater than the polar interface between pairs of dimers forming the tetramer. A common feature of catalytic MAT proteins is that they share a disordered loop segment that becomes well-ordered on the binding of substrate or products, which was termed the gating loop. This loop was proposed to act as a flexible gate allowing access to the active site and this notion has been supported by kinetic and spectroscopic studies (Taylor et al. 1996; Taylor et al. 2002).

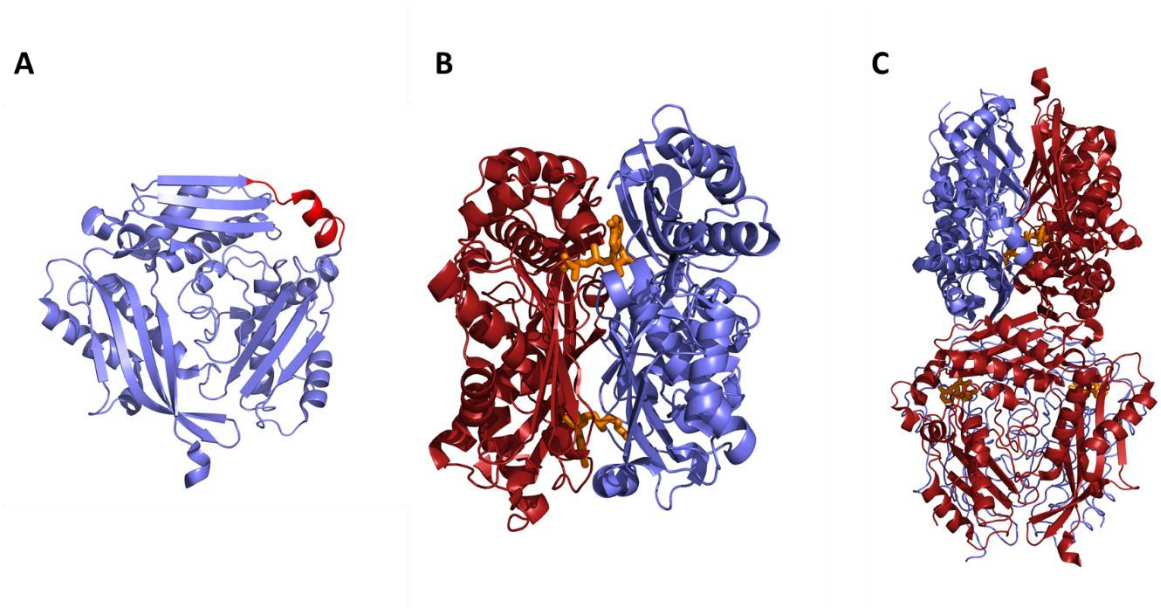

Figure 1.5. Structural organisation of MAT proteins. A. A monomer of MATα1 (PDB=2OBV) with the gating loop highlighted in red. **B**. $MAT(\alpha1)_2$ in slate and dark red showing the formation of the active enzyme. Two active sites are created upon the dimerisation of MAT monomers taking residues from each monomer. The product SAMe can be seen as orange sticks. **C**. $MAT(\alpha1)_4$ forms from two MAT dimers. Both dimer and tetramer are functional enzymes, whilst the monomeric form is not functional as it contains no complete active site.

The structurally unrelated regulatory subunit MATβ has two domains: a core Rossmann-like domain and a second extension domain (Fig 1.6A). The Rossman-like domain consists of a seven-stranded β-sheet that is sandwiched between six α-helices and it contains conserved signature sequences common for a short-chain dehydrogenase/reductase. These include, a co-factor binding domain rich in glycine residues ($G^{35}ATG^{38}LLG^{41}$) and a catalytic triad important for substrate activation ($S^{136}…Y^{159}GKTK^{163}$) (Oppermann et al. 2003; Shafqat et al. 2013). The second extension domain contains α-helices and double-stranded β-sheet that makes up a nicotinamide adenine dinucleotide phosphate (NADP) binding site consisting of the conserved GxxGxxG motif. The adenosyl phosphate end of NADP is anchored to MATβ by the L1 loop that contains a sequence that confers specificity for NADP over nicotinamide adenine dinucleotide (NAD) in many short-chain dehydrogenase/reductases ($G^{59}xRR^{62}$) (Fig 1.6B). Gly59 provides the space for the phosphate moiety whilst Arg62 supports the phosphate oxygen's through hydrogen bonding (Shafqat et al. 2013).

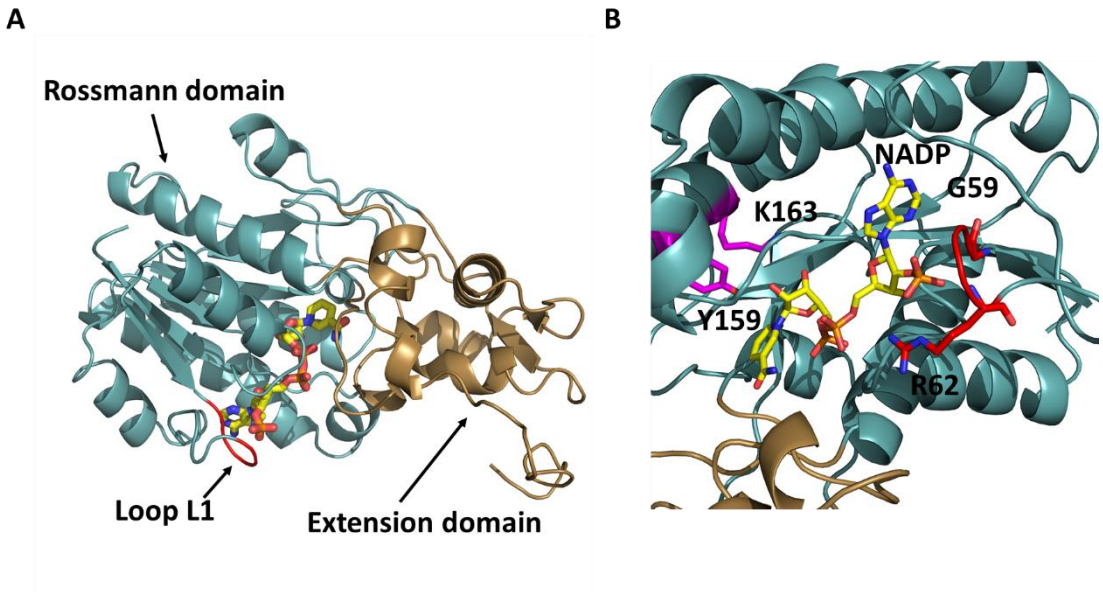

Figure 1.6. Crystal structure of MATβ. A. MATβ has two domains: The Rossmann domain in teal and the extension domain in bronze. The L1 loop, a loop needed to anchor NADP, is highlighted in red and a molecule of NADP can be seen as sticks. **B**. A closer view of NADP binding shows the G^{59}xRR^{62} motif in red which confers specificity of NADP, whilst in purple the residues Y^{159}GKTK^{163} are shown as sticks and these are important for substrate activation.

1.8 Enzymatic reaction of MAT

SAMe synthesis follows an SN_2 catalytic mechanism (Parry et al. 1982; Markham et al. 1987) where the reaction is initiated through a nucleophilic attack by the sulphur atom of methionine on the C5' atom of ATP which produces the intermediate triple phosphate (PPPi). The hydrolysis of the triple phosphate into pyrophosphate (PPi) and orthophosphate (Pi) occurs prior to the release of all three products, (SAMe, PPi and Pi) from the enzyme (Fig 1.7).

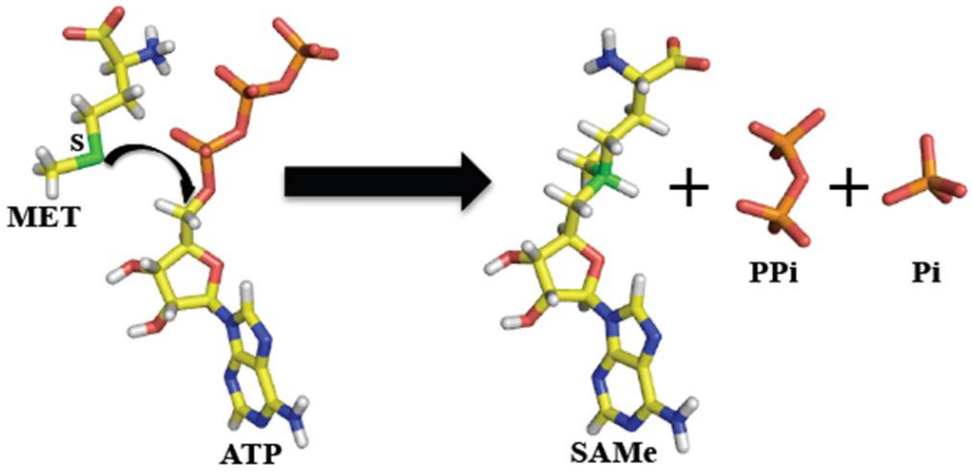

Figure 1.7. Schematic of SAMe synthesis. Methionine (MET) and ATP react together to produce SAMe and a triple phosphate (PPPi) which is then broken down into pyrophosphate (PPi) and orthophosphate (Pi). SAMe synthesis follows an SN_2 catalytic mechanism which is initiated through a nucleophilic attack by the sulphur atom of methionine on the C5' atom of ATP. This produces the intermediate triple phosphate that is broken down into pyrophosphate and orthophosphate.

Komoto *et al* (2004) proposed a detailed mechanism of actions based on several structures of eMAT. The substrates ATP (here AMP-PNP) and methionine enter the active site and the flexible gating loop becomes well-ordered closing the active site. The NE_2 of His14

forms a hydrogen bond with O5′ of AMP-PNP (O5′-NE_2=2.85 Å), and the ND_1 of His14 is surrounded by the two amide groups of Asp16 and Lys17 within hydrogen bond distance (ND_1-N[Asp16]=3.05 Å, ND_1-N[Lys17]=3.69 Å) (Fig 1.8A). The imidazole ring of His14 can be polarized so that the proton on NE_2 is readily released to O5′ of ATP (Fig 1.8B). The C5′-O5′ bond cleavage occurs following a molecular vibration allowing the amide groups of Asp16 and Lys17 to form strong hydrogen bonds with ND_1 of His14. This leads to a series of electron movements taking place (Fig 1.8C) resulting in the cleavage of the ATP C5′-O5′ bond forming the triple phosphate coproduct (Fig 1.8D). The negatively charged ND_1 of His14 is stabilised by the surrounding basic main chain amide groups of Asp16 and Lys17.

The positively polarized C5′ moves toward the electron-rich SD atom of the bound methionine by changing the ribose ring conformation from C4′-exo to C3′-endo. A nucleophilic attack on the C5′ to form SAM and displace triple phosphate occurs due to the SD atom. It must be noted that the C5′-O5′ bond cleavage and the C5′-SD bond formation occurs in a concerted manner, consistent with a SN_2 reaction. This proposed catalytic mechanism is supported with previous mutational studies which showed when H14N mutation was down a >1000-fold reduction in V_{max} was observed. The V_{max} of the mutant K165M only diminished by 2.5-fold, indicating that interactions with Lys165 are less important in catalysis (Taylor & Markham 2000). SAMe in crystal structures have an S-configuration at the SD atom, and the naturally occurring SAMe has the S-configuration as well (Cornforth et al. 1977; Stolowitz & Minch 1981), supporting the proposed catalytic mechanism. The triple phosphate is then proposed to be broken down into pyro and orthophosphate providing the energy to dislodge the gating loop and facilitate product release. The mechanism of hydrolysis of the triple phosphate

intermediate appears to be similar to other metal ion dependent ATPases and phosphatases. However, it still remains a question which water is activated to react with the triple phosphate and thus split the polyphosphate chain into PPi and Pi (Markham & Pajares 2010).

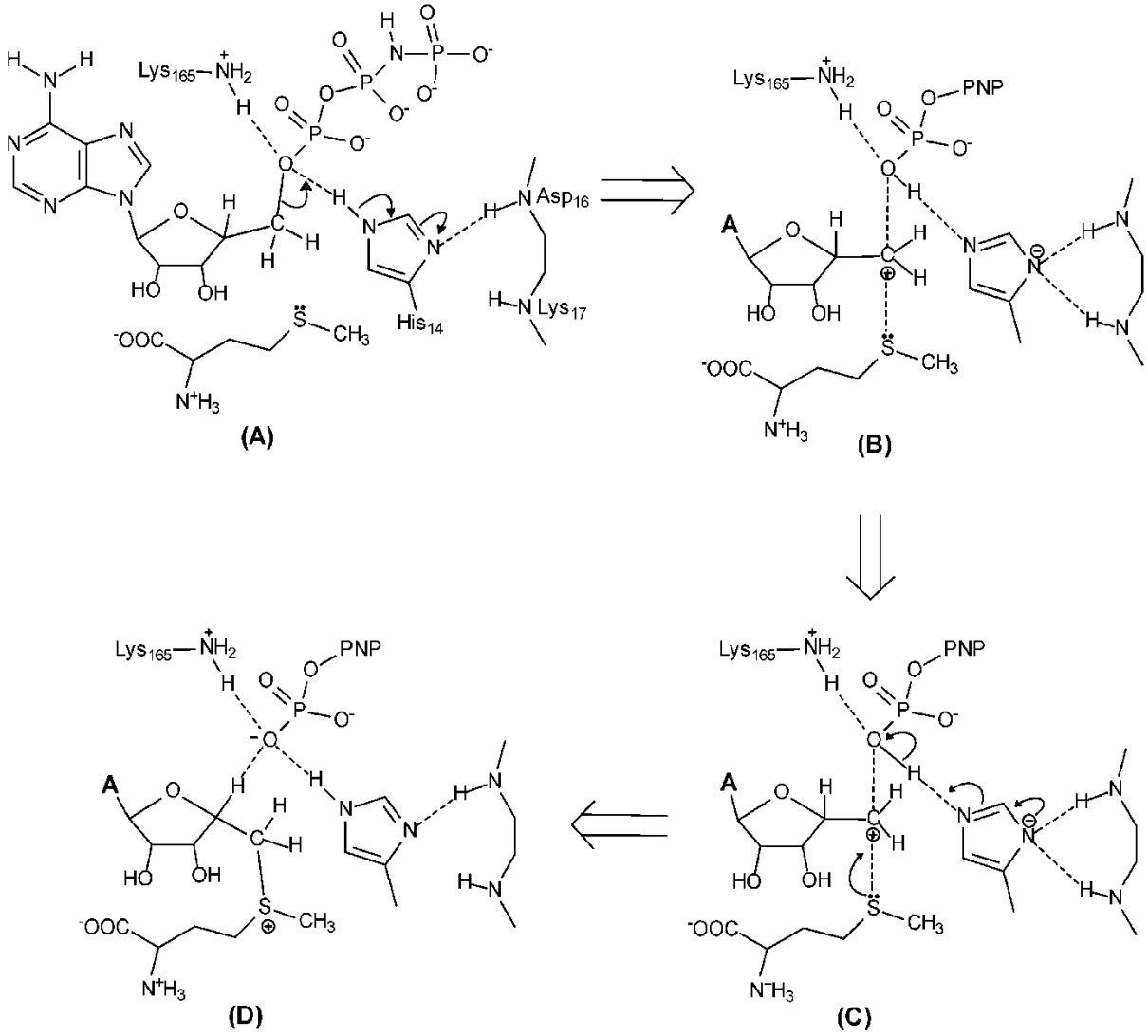

Figure 1.8. Proposed catalytic mechanism for SAMe formation. Curved arrows indicate the movements of electron pairs. Possible polar interactions are illustrated by dotted lines. The "+" and "-" indicate positive and negative charges, respectively. Although the mechanism is illustrated in a sequential fashion for clarity, the C5′-O5′ bond cleavage and the C5′-SD bond formation occur in a concerted manner, i.e., the proposed SAM formation is a SN_2 reaction. Figure taken from Komoto *et al* 2004.

It has been shown that MAT enzymes do not just metabolise methionine, but depending on species, can process many methionine analogues (Wang et al. 2014). Ethionine, an S-ethly analogue of the amino acid methionine, is known to induce steatosis and pancreatitis in rat (Rao et al. 1982). MAT enzymes are able to produce S-adenosylethionine (SAE) from ethionine (Smith & Salmon 1965), which in turn acts as a competitor of methionine and as a result, insufficient levels of SAMe is produced within the cell causing disruption of SAMe-dependent processes such as nicotinamide catabolism (Skordi et al. 2007). SAE can be used to ethylate targets by donating its ethyl group in direct competition with methylation, which results in abnormal alkylation of nucleic acids (Orenstein & Marsh 1968).

1.9 Aim of this research project

As HCC development is strongly associated with the increased expression of both MATα2 and MATβ providing a proliferative advantage for the cell, the aim of this PhD thesis is to determine the interface interaction between MATα2 and MATβ at the molecular level by solving the structure of the MATαβ complex using X-ray crystallography. Once the structure has been solved interaction interfaces between these molecules will be studied and then structure based mutants that would change the nature of the wild-type interactions will be designed. This will allow for the identification of the key residues needed to maintain the complexes.

Chapter II. An introduction to crystallography and the theory of x-ray diffraction.

2.1 An introduction to protein crystallography

Macromolecular X-ray crystallography is one of the most utilised techniques available to determine the three-dimensional structure of proteins and nucleic acids. There are a large number of books that cover the subject of protein crystallography so in this thesis an overview of protein crystallisation and the analysis of X-ray diffraction experiments will be given (Drenth 2007; Rupp 2009; Rhodes 2006).

Of the stages of solving the structure of a protein, protein crystallisation is the least understood part and remains, fundamentally, a trial and error process. The crystallisation process is an empiric method that depends on several physicochemical parameters that are not yet completely understood. The presence of impurities and other unknown factors plays a role in the formation of protein crystals. In general the purer the protein, i.e. the higher the homogeneity, the increased likelihood of a crystal growing and also obtaining crystals that diffract to a higher resolution. For this reason the requirement of protein purity is more stringent for a crystallographer than those of a biochemist who would be happy, for example with a purity of an enzyme that retains a catalytic activity despite high impurities.

2.2 Crystallisation of a protein

Crystallisation of protein involves several steps and requires bringing the macromolecule to supersaturation. The protein is dissolved in a solvent, usually a water-buffer solution of an organic solvent, before being precipitated by an organic compound or salt. The protein mixture must be brought to supersaturation whereby small aggregates, termed protein nuclei, are grown. The nuclei spontaneously form at a high level of

supersaturation which allows for the growth of crystals to occour. For optimal crystal growth once the protein nuclei has formed a reduction in the supersaturation is needed in order to produce larger single crystals suitable to x-ray diffraction (Fig. 2.1). If the protein or precipitant concentrations are too high amorphous precipitation forms whilst if the solution is too undersaturated, the drop can be devoid and clear of all precipitation. In these scenarios the likelihood of crystal formation is extremely minimal (Dessau & Modis 2011; Chayen & Saridakis 2008).

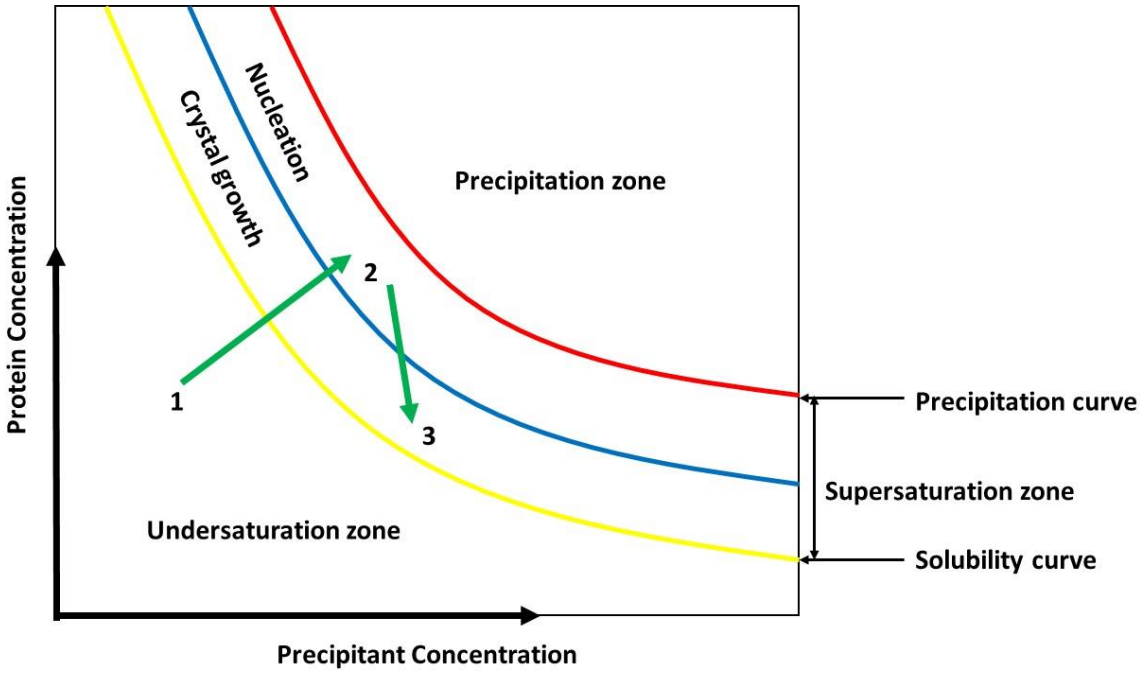

Figure 2.1 Overview of crystal growth. Initially the protein concentration is too low for crystal formation and lies in the undersaturation zone [1]. As the protein concentration increases, the solution becomes supersaturated and nucleation can occur [2]. As protein is used to form nuclei the protein concentration in the drop decreases allowing for the growth of the crystals [3]. If the protein becomes too concentrated it enters the precipitation zone where it precipitates out of solution.

2.3 Methods of protein crystallisation

The vapour diffusion method is one of the most widely used techniques for protein crystallisation where a mixture of protein/precipitant is placed placed on a siliconised cover slip. The drop is allowed to equilibrate in a closed vessel containing a high concentrated precipitant reservoir which causes the concentration of the protein to be gradually increased by diffusion. Volatile species, including water or organic molecules, diffuse from the drop to the reservoir until the concentration of precipitant is the same in both the drop and reservoir. There are two variations of the vapour diffusion method depending on the way the protein/precipitant drop is placed: hanging drop and sitting drop (Fig. 2.2), but both follow the same principles. During this progress, protein nuclei begin to form as the drop enters into a super-saturation state, surpassing an energetic barrier, due to diffusion causing the gradual increase in protein concentration.

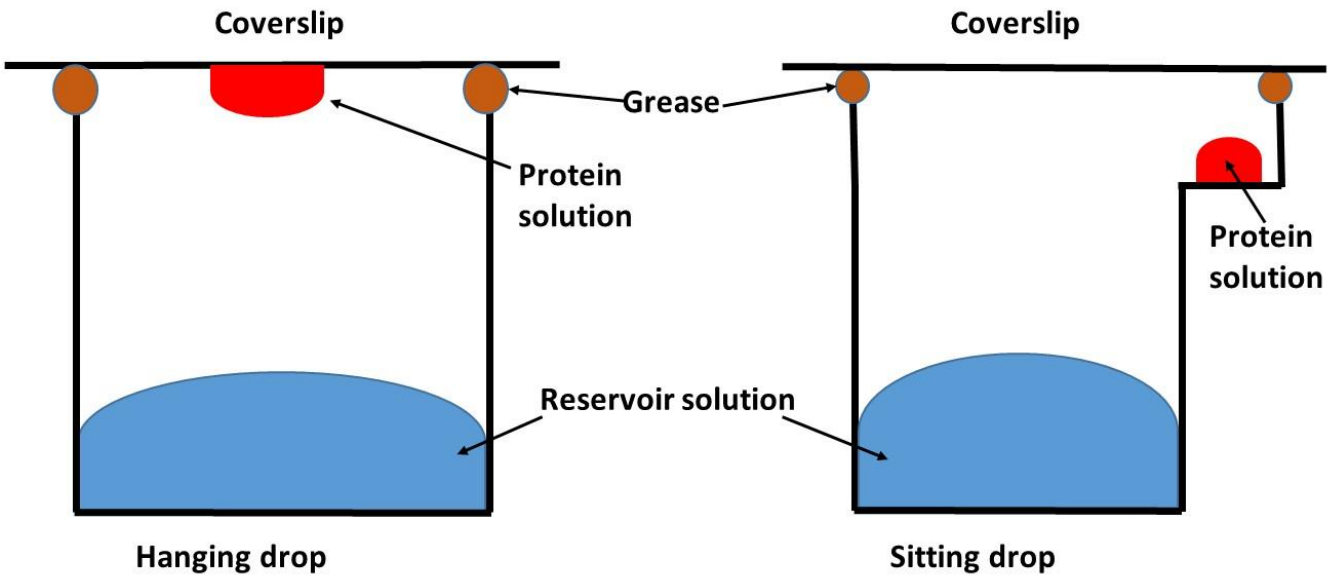

Figure 2.2. Vapour diffusion methods for crystallisation. Hanging drop (left) and sitting drop (right) methods contain a reservoir solution (blue) and protein/solution mix (red). Diffusion occur between the reservoir and the drop concentrating the protein within the drop allowing for supersaturation to be reached.

2.4 Crystallisation screening

In general, the search for suitable crystallisation conditions is done in two steps. First, is to use the wide range of commercially available crystallisation screen which contains a wide range of precipitant and buffer solutions to identify hits. The second step is to take the most promising conditions and vary the conditions that the crystals originally formed in changing the concentration of the protein and/or precipitant, pH, temperature or change the presence of additives such as cations, anions, organic compounds and detergents. The need to optimise an initial hit allows for the growth of better diffracting crystals.

2.5 Crystal lattices and symmetry

A crystal is a solid composed of atoms arranged in an orderly repeated way. Any crystal, organic or inorganic, is made up of unit cells, which can be considered as the ordered-translational repeat along three dimensional space. The unit cell itself is defined by three axis (a, b and c) and by three angles (α, β and γ) and the geometric arrangement of the axis and angles gives rise to the seven crystal systems: Triclinic, monoclinic, orthorhombic, trigonal, hexagonal, tetragonal and cubic. Regardless of which system the crystal is formed by each unit cell contains a smaller motif called the asymmetric unit (Fig 2.3). The asymmetric unit is defined by a combination of 32 point group symmetry operations, 14 Bravais lattices (or lattice translations), several rotational operations and screw axis symmetry operations. There are a total of 230 space groups, but for proteins, which are chiral molecules, only rotational and screw symmetry operations are allowed limiting the number of point groups to 11 and the total number of space groups to 65. Unlike salt crystals which have low solvent content and strong electrostatic bonds, protein crystal interactions use hydrogen bonding and as a result have increased fragility.

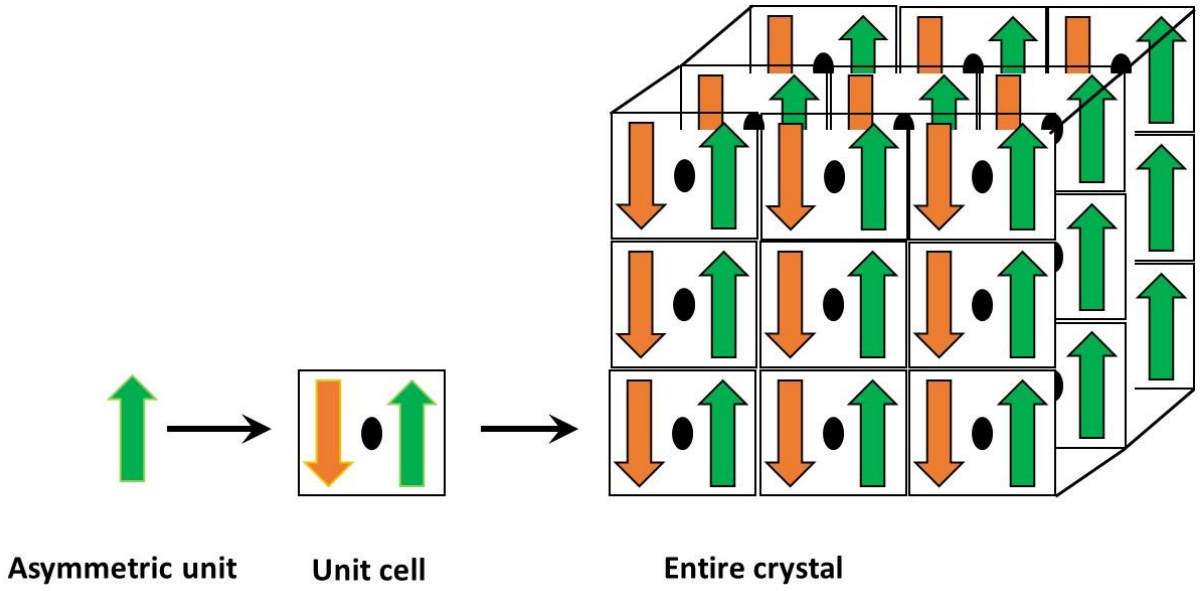

Figure 2.3. Formation of the asymmetric unit. The asymmetric unit green arrow sits within the unit cell after crystal symmetry has been applied (black dot) to give rise to a new orientation of the asymmetric unit (orange arrow). The entire crystal is made up of many repeats of the unit cell.

2.6 X-ray radiation

X-ray radiation, produced from an X-ray generator, is monochromatic. An X-ray generator contains a vacuum tube known as the X-ray tube which houses within it a metal filament that is heated to produce electrons. These electrons are then accelerated in vacuum by a high electric field (20-60 kV) towards a metal target called the anode. The anode collects the electrons and is typically made out of copper which allows for heat, generated by the collision to evacuate. When electrons collide with the target, 1 % of the resulting energy is emitted as X-rays whilst the remaining energy is dissipated as heat.

In synchrotron radiation sources, electrons that are produced from an electron gun are accelerated in a linear accelerator (linac) before being transmitted to a circular accelerator, known as a booster synchrotron, where they are accelerated further to reach an energy level of 6 GeV. These highly relativistic electrons are then stored in a storage ring which allows them to travel along a circular path at a constant energy, for many hours. The electrons are moving at speeds close to that of light within the storage ring and when their paths are altered by a magnetic field, through devices such as wiggler magnets, undulators and wave length shifters, they emit X-rays.

2.7 X-ray structure analysis

When a wavefront of X-rays interacts with a sample it will either be transmitted, in which case it continues along its original direction, or it will interact with an atom and the electrons in that atom will re-emit that X-ray radiation, normally without changing the wavelength, as a spherical wavefront. All the atoms in the path of the X-ray beam scatter X-rays and peaks form when scattered X-rays, constructively interfere. Constructive

interference occurs when two X-ray waves, with phases separated by an integer number of wavelengths, add to make a new wave which has a larger amplitude. When two parallel X-rays from a coherent source scatter from two adjacent planes their path difference must be an integer number of wavelengths for constructive interference to occur (Fig 2.4). For systems which are not ordered, most of these scattered waves interfere destructively and results in lower scattered intensities. However, in ordered systems, such as in crystalline structures the scattered waves interfere constructively as the atoms are placed in a periodic and defined pattern.

The noble laureates, William and Lawrence Bragg defined the diffraction by crystal lattices in 1913 (Bragg & Bragg 1913). They found that the co-operative scattering of X-rays could only occour when the X-ray, of wavelength λ, is incident onto the surface of a crystal, and the angle of incidence, θ, will reflect back with a same angle of scattering, θ. When this occurs and, when the path difference, d is equal to an integer, n, of wavelengths, constructive interferences occurs. (Equation 1, Figure 2.4).

$$n\lambda = 2d \sin\theta$$

Equation 1.

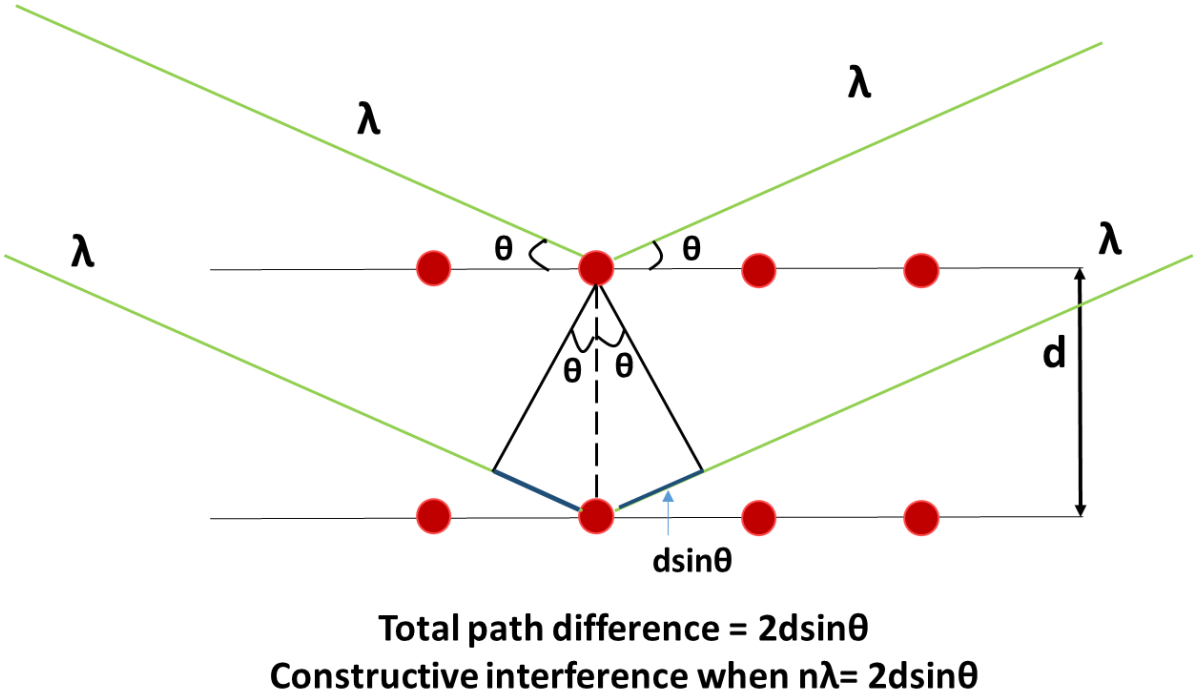

Figure 2.4. Graphical representation of Bragg's law. Constructive interference between two diffracted rays (green lines) only occurs if the path length difference (d) between them corresponds to a multiple of the wavelength, λ.

2.8 Calculating electron density

A waveform can be represented using a combination of both the sine and cosine components of the wave producing a complex number [*cos 2π (hx) + i sin 2π (hx)*] which in turn can be expressed as an exponential, $e^{2\pi i(hx)}$. The exponential contains the frequency of the wave (*h*), an imaginary component (*i*) and is equal to the phase (α_h). In terms of amplitude (F), regardless how complicated a wave is it can be defined by the sum of the simpler wave functions:

$$f(x) = \sum_{h=0}^{n} F_h \, e^{2\pi i(hx)}$$

Equation 2.

In order to solve the structure of a protein using x-ray crystallography a specific function that calculates the three-dimensional electron density of the molecules or $\rho(x,y,z)$ must be determined. To accomplish this the contribution of the frequencies (*h,k,l*) from waves scattering in different directions of space (*x,y,z*) must be considered to account for the three dimensional sum of waves (Equation 3).

$$f(x,y,z) = \sum_{h} \sum_{k} \sum_{l} F_h \, e^{2\pi i(hx+ky+lz)}$$

Equation 3.

For any function, *f(x)*, there is another reciprocal function called the Fourier transform, *F(h)*, which instead of terms of time *(x)* it is expressed in terms of frequency *(h)*. A Fourier transform can be applied to a sum of waves or Fourier series to obtain the contributions of the different waves (Equation 4), which is an approximation that relates the information in a diffraction pattern (reciprocal space) with the contents in the unit cell (real space).

$$F(h,k,l) = \iiint_{xyz} f(x,y,z)e^{2\pi i(hx+ky+lz)}\ dx\,dy\ dz$$

Equation 4.

Reflections, which make up the diffraction pattern, are generated by an x-ray wave that impacts the detector and each reflection represents an arriving wave that is the constructive sum of the waves diffracted by all scatterers in the unit cell. Therefore every atom in the unit cell contributes to every reflection in the diffraction pattern. The atomic structure factor, f_{hkl}, (Equation 5) is the contribution of one atom to the reflection *(h,k,l)* and the Fourier series or sum that describes the specific wave is called the structure factor of the reflection or F_{hkl} (Equation 6). For every atomic structure factor the scattering factor of a particular atom f_{hkl} has to be considered as a sphere of electron density. Therefore, the F_{hkl} is considered the Fourier transform of the electron density, *p(x,y,z)* and vice versa (Equation 6).

$$f_{hkl} = f_j e^{2\pi i(hx_j ky_j lz_j)}$$

Equation 5.

$$F_{hkl} = \iiint_{xyz} p(x,y,z)e^{2\pi i(hx+ky+lz)}\ dx\,dy\ dz$$

Equation 6

To summarise, an x-ray diffraction pattern contains a set of intensities that are associated with structural information called structure factors, F_{hkl} and by using the Fourier

transform the electron density, *p(x,y,z)*, can be related to each point in the crystal and whose structure factor correspond to each reflection (*h,k,l*) in the diffraction pattern.

2.9 Indexing, integration and scaling

Once a data set from a crystal has been obtained, which has different angular reflective snapshots, data processing can occur. Firstly, each reflection gets Miller indexes *(h,k,l)* assigned to it and a list of possible lattices is suggested. In order to analyse the data the correct lattice must be chosen so intense processing of the unit cell parameters and crystal mosaicity is done to select the lattice. The next step evaluates the intensities of the reflections in a process called integration. Here, a long list of indexed reflections containing their individual intensities I*(h,k,l)* and the standard error statistics σI*(h,k,l)* are defined. The huge amount of data here needs to be reduced, which is achieved by combining partial identical reflections from multiple frames collected at different angles and normalising the data onto an absolute scale; this is scaling.

Equation 6 relates the reflections that are collected with the electron density in the crystal and this requires not only the modules of the structure factor, $|F_{hkl}|$, but the phase α(*h,k,l*) of the wave. Unfortunately, the data from the reflections does not provide direct information about the phase of the wave, but only information pertaining to the structure factor modules, calculated from the reflection intensities. This pitfall of x-ray crystallography is known as the phase problem and several different methodologies have been developed to solve it, allowing for the calculation of the electron density and, finally, resolve the structure. In this thesis molecular replacement was used to overcome the phase problem.

2.10 Molecular replacement

Molecular replacement is the process of solving the phase problem for an unknown structure by placing the atomic model for a related, known structure in the unit cell of the unknown structure in such a way as to best reproduce the observed structure factors. The known model, once placed, may be used to calculate phases which, in combination with the observed structure factors for the unknown structure, allows for the model to be rebuilt and refined and a model to be generated. To do this a calculation involving a six dimensional search over all possible orientations and translations of the known model in the unit cell of the unknown structure has to occur. The six dimensional search is split into two parts. The first part is a three dimensional search over all possible rotations to determine the orientation. After this is done, another 3 dimensional search occurs over all possible translations to work out the position of the orientated model.

Using the search model diffraction data are generated, with a large P1 unit cell and this is termed F_c. The Patterson map is calculated for both the observed crystal data and the model (P_{obs} and P_c). The Patterson map represents all atom-atom vectors, translated to the origin including the vectors within the molecule and vectors between symmetry-related molecules (Fig. 2.5). The P_c is rotated against the P_{obs} to find the rotation matrix with the maximum correlation and from this the structure factors are recalculated (F_{model}) for a P1 unit cell of the same cell dimensions as the crystal unit cell. This rotated model is then translated to every position in the crystal unit cell and also accounts for the models related by space group symmetry to obtain the F_{calc}. The R-factor between F_{obs} and F_{calc} is determined and if this less than 50 % then the solution is correct.

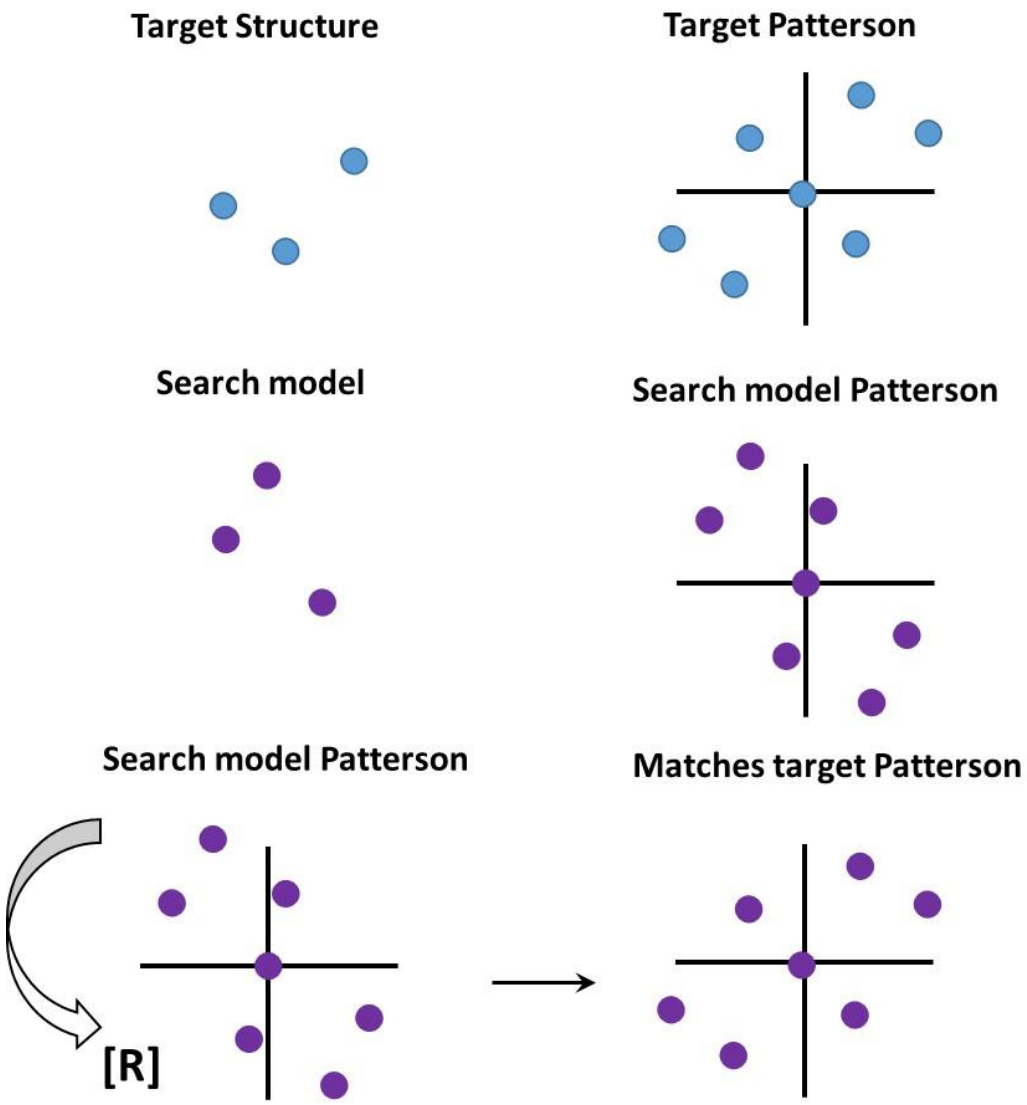

Figure 2.5. Schematic of the rotation function of molecular replacement. Patterson maps are generated for both the target data and the search model. A rotation matrix is applied to the search model patterson to align it with the target Patterson.

In summary, molecular replacement can be thought of as two parts, a rotation function and a translation function. The schematic in Fig 2.6 shows that in order to obtain the solution for the target molecule, A', the search molecule, A, must be rotated by the rotation matrix, [R], before a translational vector is applied, T (Fig 2.6).

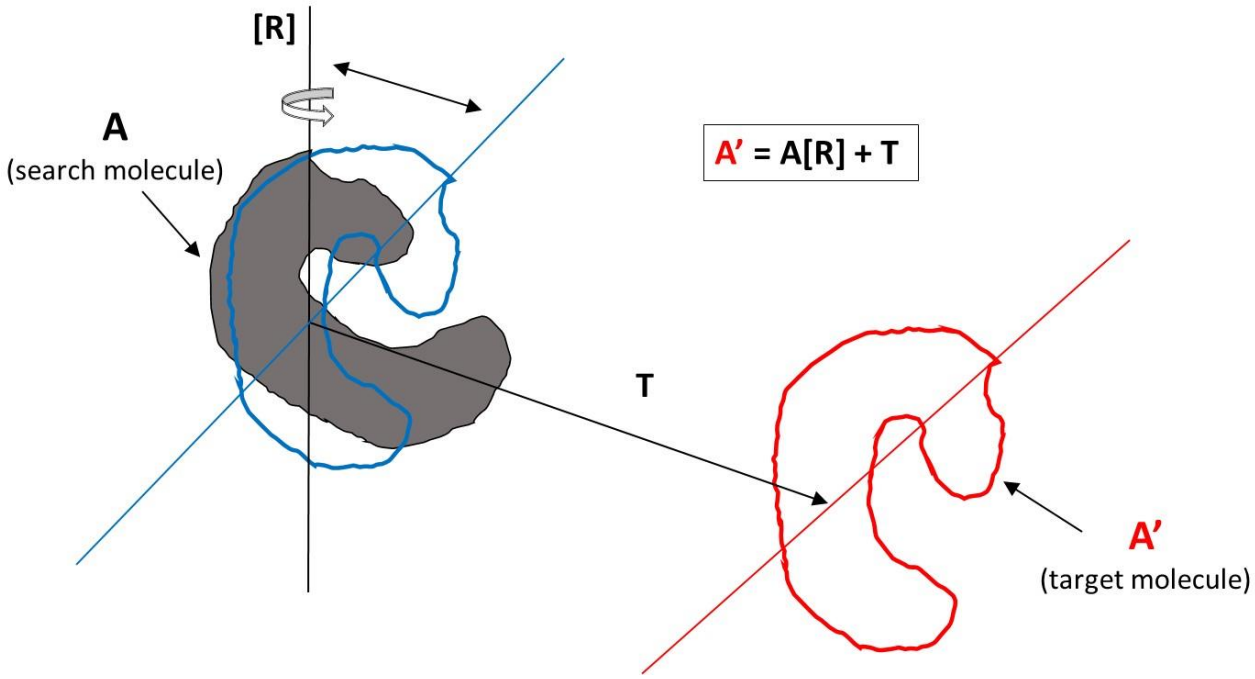

Figure 2.6. A simple view of Molecular replacement. Firstly the search model, A, is rotated by the rotation matrix [R], followed by a translation, T, to give the solution of the target molecule A'. The movement of the search model to the target model is given by the equation A'=A[R]+T.

2.11 Refinement and validation

Once a solution has been produced from molecular replacement the model has to be refined and evaluated for inaccuracies. When refining stereochemical parameters bond distance, bond angle and side chain rotamers are taken into account. The Ramachandran plot evaluates the Φ and Ψ angles of the peptide bonds within the molecule allowing for the identification of improper main chain rotations. Other molecules such as water, compounds and ions that were present during the protein purification or the crystallisation may also be present in the structure and must be built into the density. All atoms within the model must be checked for interatomic distances to ensure that the molecules are not too far or too near from other atoms. R-value is the measure of the quality of the atomic model obtained from the crystallographic data. When solving the structure of a protein, you build an atomic model and then calculate a simulated diffraction pattern based on that model. The R-value measures how well the simulated diffraction pattern matches the experimentally-observed diffraction pattern. A random set of atoms will give an R-value of about 63 %, whereas a perfect fit would have a value of 0 %. Values are resolution dependent and for a 2 Å structure a typical R-value would be about 20 %. Some bias can come from the starting model, which could affect the R-value so another parameter, R-free, is used in combination with R-value. When calculating the R-free and before refinement begins, about 10% of the experimental observations are removed from the data set and the remaining 90 % is used for model refinement. The calculated R-free value looks at how well the model predicts the 10% that were not used in refinement. For an ideal model, that is not over-interpreting the data, the R-free will be similar to the R-value and typically, it is 3-5 % higher than the R-factor (Kleywegt & Jones 1997). Validation of protein structures can be achieved through the online sever MolProbity (Davis et al. 2007), which provides a summary of any errors in stereochemistry or interatomic clashes.

Chapter III. Small angle x-ray scattering

3.1 Overview of Small angle x-ray scattering

Small angle x-ray scattering (SAXS) is a powerful structural biology technique that is used to investigate the oligomeric state of proteins and macromolecular complexes. It can also be used to quantitatively characterised kinetic reactions and assess the flexibility of intrinsically unfolded proteins(Koch et al. 2003; Putnam et al. 2007). SAXS shares both similarities and differences with x-ray crystallography, both are based on the elastic scattering of photons by molecules. For a crystal, where the macromolecules are regularly positioned the scattering waves interfere and produce diffraction patterns which can yield high resolution structures. For SAXS, the molecules are free to move and have random orientations within the solution, which allows for the magnitudes of the interatomic distances to be calculated and a low resolution shape of the molecules in solution to be reconstructed.

3.2 Developments in SAXS

An advancement in SAXS technology is the combination of SAXS with size exclusion chromatography (SEC), which allows for the separation of different populations of a protein. This is particularly useful ensuring that a protein sample is monodispersed and homogeneous before data are collected, and also allows for the separation of different oligomeric states. The SAXS-SEC method has been set up in the SWING beamline at the Soleil synchrotron, France. The beamline couples data collection with a high performance liquid chromatography (HPLC) (David & Pérez 2009).

3.3 SAXS theory

Scattered photons are collected on a 2D detector but as the scattering is isotropic, meaning that the x-rays scatter with equal efficiency into all possible direction, the images collected are radially averaged to obtain the scattered intensity, *I*, as a function of the momentum transfer *s*:

$$s = \frac{4\pi \sin \theta}{\lambda}$$

Equation 9

The wavelength here is represented by λ of the incident beam, and θ is half the angle between the incident beam and the scattered radiation. From Braggs law (Equation 1), distance probed is inversely proportional to the momentum transfer.

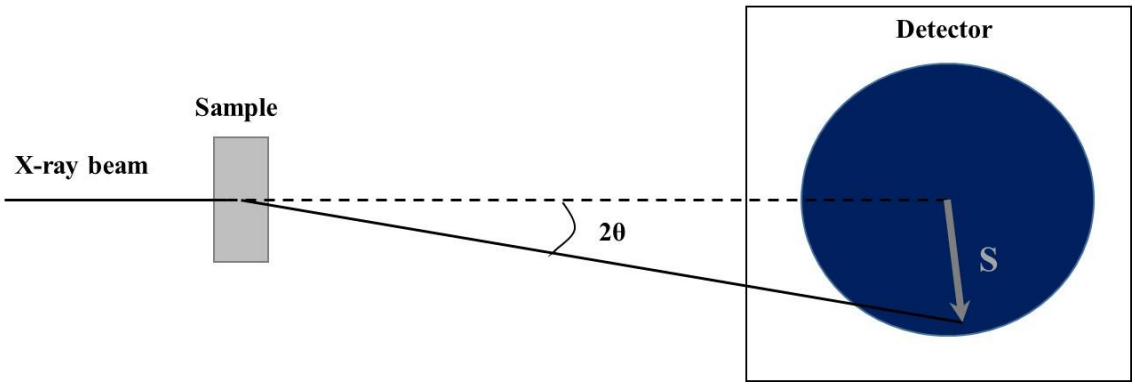

Figure 3.1. Schematic of a small angle x-ray scattering experiment. An X-ray beam passes through the sample and is diffracted which is recorded on the detector**.**

During a SAXS experiment, scattering data are collected for both the dissolved particles and the buffer that the particles are dissolved in separately and later the latter is subtracted from the former. The signal difference between these two data sets is proportional to the number of particles in the solution and also it is proportional to the difference in the electron density between the particle $p(r)$ and the solvent p_s. Unfortunately, for biological systems the biological molecules contain light atoms such that the particle density is only

slightly higher than that of the buffer. To overcome this SAXS instrumentation has been optimized to have the lowest possible background. For a monodisperse system of identical particles that are non-interacting, the corrected background scattering intensity I(s) is proportional to the scattering from a single particle that is averaged over all possible orientations. This can be expressed as:

$$I(s) = \langle \mathrm{I(s)} \rangle_{\Omega} = \langle A(s)A^{*}(s) \rangle_{\Omega}$$

Equation 10

Here, the scattering amplitude A(s) is a Fourier transformation of the excess scattering length and the solid angle in the reciprocal space of the scattering vector is denoted by Ω.

3.4 Overall particle characterisation

The Guinier approximation is a fast method of providing information on size, oligomeric state, interactions and overall quality of the data. The Guinier analysis (Guinier 1939) is the most straightforward method to determine both the radius of gyration, R_g, and the forward scattering $I(0)$. For a monodisperse system the Guinier equation:

$$I(s) = I(0) \cdot \exp\left(-\frac{1}{3} {R_g}^2 s^2\right)$$

Equation 11

The approximation can be used to calculate both parameters and for a monodisperse sample the gradient of the line and very small s values gives the R_g, and the intercept with the ordinate yields the intensity of forward scattering. The radius of gyration of a macromolecule is defined as the root mean squared distance of scattering particles that are weighted about its centre of mass and can be used to infer differences in the compactness between proteins. The linearity of the plot provides an indicator to sample

quality as a non-linear plot may show poor buffer subtraction, interparticle interaction or polydisperisty. The region that is linear is only such for a small region. Calculating the scattering at zero angle *I*(0), which once normalised to the solute concentration, enables an estimate of molecular mass to be calculated as *I*(0) is proportional to the molecular mass.

Another way to calculate the molecular weight is to use the Porod volume (Porod 1982). This volume is the hydrated volume of the molecule and can be calculated if a uniform electron density is assumed. The Porod volume can be calculated from the Porod invariant, Q:

$$Q = \int_0^\infty s^2 \, I(s) \cdot ds$$

Equation 12

$$V_p = 2\pi^2 \, I(0)/Q$$

Equation 13

Using Equation 13 the volume of the protein can be calculated for which the molecular weight is in kilodaltons, and this value is typically 0.6 times the Porod volume. Porod calculated that Q is the direct product of excess electron scattering of the particle and regardless of the x-ray wavelength, data should have the same constant with the same sample at the same concentration.

The P(r) function or pair-distance distribution function describes the paired-set of all distances between points within an object. In SAXS, the P(r) function is used to describe

the paired-set of distances between all of the electrons within the macromolecular structure. From this, the maximum distance between two particles can be calculated, D_{max}, which gives the largest length of the particle. The P(r) is related to I(s) by the integral:

$$I(s) = \int_0^{D_{max}} P(r) \frac{sin\ s \cdot r}{\sin s \cdot r} dr$$

Equation 14

3.5 *Ab Initio* modelling

Ab initio modelling methods allow one to build three-dimensional, low resolution models from 1 D scattering profile. There are a number of programs that can be used within the ATSAS suite (Petoukhov et al. 2012) that are able to do this, but for this thesis DAMMIN and DAMMIF were used (Petoukhov et al. 2012). DAMMIN used an algorithm that represents the shape of the particle using densely packed beads in a constrained search volume. The sphere usually has the diameter of the D_{max} previously calculated, and using a simulated annealing algorithm the beads are moved from a random configuration to an optimum configuration that minimises the difference between the collected scattering data and predicted scattering pattern from the model (Svergun 1999). *Ab initio* methods do not provide unique solutions and if you run the programme multiple times, numerous model will be generated. Thus, multiple runs of DAMMIN can be performed, which produce many models. These models can be aligned using the programme SUPCOMB, which identifies the most typical model in the ensemble and aligns them (Kozin & Svergun 2001). The aligned models can be constructed into a single averaged shape, conserving the most persistent features using the programme Damaver, which can be used to average all the models.

DAMMIF is a faster version of DAMMIN that uses an unconstrained volume that is able to grow during the simulated annealing if needed, which overcomes effects caused by a slight underestimation of D_{max} (Franke & Svergun 2009). DAMMIF also incorporates other programmes including Damaver and SUPCOMB which is automated and allows for multiple runs of DAMMIN, aligned, scored and averaged before a final model is produced. A limitation of these two methods is that they make the assumption of a uniform electron density within the particle, which limits the resolution of the bead model. So, this means that the scattering curve can only be fitted to a restricted range ($s = 0.3$ $Å^{-1}$) (Blanchet & Svergun 2013). Another programme, Crysol, can be used to compare the 1D scattering curve of a data collection to a known structure by generating a SAXS intensity profiles from atomic coordinates while considering the hydration shell surrounding the target sample (Barberato et al. 1995).

Chapter IV. Materials and Methods.

4.1 Materials

4.1.1 Bacterial strains

Strains that were used in this work are summarised in Table 4.1 below. For expression of MATα1, MATα2 and MATβV2 in all vectors, BL21 (DE3) cells were used whilst for MATβV1 BL21 (HC41) were used for expression. DH5α and XL1-Blue were used for routine cloning and mutagenesis work.

Table 4.1 Bacterial strains used in this work

Name	Supplier	Genotype
BL21 (DE3)	Novagen	*E. coli* B F- *dcm ompT hsdS*(r_B- m_B-) *gal* λ(DE3)
BL21 (HC41)	Lucigen	*E.coli F ompT hsdS B (rB -mB-) gal dcm (DE3)*
DH5α		*E. coli* Δ(*lac*ZYA–*arg*F)U169 *deo*R *rec*A1 *end*A1 *hsd*R17(r_K^+, m_K^+) *sup*E44 λ– *thi*-1 *gyr*A96 relA1 F- *pro*AB+ *lac*1qZ Δ M15 ssf::Tn5 [Km^r]
XL1-Blue	Stragene corp.	*E. coli rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 lac [F′ *pro*AB *lac*Iq ZΔM15 Tn*10* (Tet^r)]

4.1.2 Bacterial culture

Growth of bacterial strains was achieved using Lysogeny broth (LB) medium for all proteins. Ten mL overnight cultures were grown in 50 mL falcon tubes in the presence of antibiotics, either ampicillin at 100 μg/mL or kanamycin at 50 μg/mL. Cultures were grown at 37 °C at 250 rpm.

4.1.3 DNAs

MATα1 and MATα2 (pNIC vector) constructs were kindly provided by structural genomic consortium, Oxford and the MATβV1 and MATβV2 constructs were from the laboratory of Dr Shelly Lu (University of California Los Angeles). MATβV1 construct was subcloned into the pHIS2 parallel vector. MATα2, MATβV1 and MATβV2 was also

subcloned into pET-28M-SUMO3-GP vector. GIT1 DNA was purchased from addgene (Plasmid #15226).

4.1.4 Plasmids

Plasmids containing oligo(Histidine) or oligo(Sumo) tags were used. These vectors contained multiple restriction sites and a sequence encoding a protease cut site which allows the cleavage of the tag sequences. Information regarding this can be found in Table 4.2.

Table 4.2. Summary of Plasmids used in this work

Full Name	Abbreviated Name	Size (bp)	Cleavage	Resistance	Source
pET-28M-SUMO3-GP	pET_SUMO	6268	SENP2	Kanamycin	EMBL
pET-28a(+)	pET28	5396	Thrombin	Kanamycin	Addgene
pHIS2 parallel	His-Par	7207	TEV	Ampicillin	(Sheffield et al. 1999)
pNIC28-Bsa4	pNIC	5353	TEV	Kanamycin	Structural genomics consortium
pEGFP-GIT1	pEGFP	6900	-	Kanamycin	Addgene

4.1.5 Oligonucleotides

The oligonucleotides used in this work are listed below in Table 4.3. Primers were designed based on DNA sequences of interest and checked for secondary structure using the multiple primer analyser from Life Technologies.

Table 4.3. Oligonucleotides used for cloning of recombinant plasmids and mutagenesis. Primer names describe the primer type (forward or reverse). Restriction site (RS) indicates sites used for cloning.

Name	Vector	Restriction Site (RS)	Oligonucleotide sequence
MATβV1W320Stop Forward	His_Par		GGAATCAAAGAATCACTTTAGCCTTTCCTCATTGACAAG
MATβV1W320Stop Rev	His_Par		CTTGTCAATGAGGAAAGGCTAAAGTGATTCTTTGATTCC
MATβV2W309Stop Forward	Pet28		GGAATCAAAGAATCACTTTAGCCTTTCCTCATTGACAAG
MATβV2W309Stop Reverse	Pet28		CTTGTCAATGAGGAAAGGCTAAAGTGATTCTTTGATTCC
MATα2Δ19 Forward	pET_SUMO	Age1	TTTTTACCGGTGGATTCACCTCAGAGTCGGTCGGGGAA
MATα2Δ19 Reverse	pET_SUMO	XhoI	TTTTTCTCGAGTCAATATTTAAGCTTCTTGGGCACTTCCC
MATβΔ16 Forward	His_Par		TTTTTCCATGGGACTCTCTATACACTTTGTTCCCGGG
MATβΔ16 Reverse	His_Par		TTTTTCTCGAGCTAATGAAAGACCGTTTG
MATα2D134A Forward	PNIC		CATTGGTGCTGGAGCCCAGGGCTTAATG
MATα2D134A Reverse	PNIC		CATTAAGCCCTGGGCTCCAGCACCAATG
MATα1C120A Forward	PNIC		CAGATATTGCCCAGGCCGTCCATCTGGAC

MATα1C120A Reverse	PNIC		GTCCAGATGGACGGCCTGGGCAATATCTG
MATα1C120G Forward	PNIC		CAGATATTGCCCAGGCCGTCCATCTGGAC
MATα1C120G Reverse	PNIC		GTCCAGATGGACGCCCTGGGCAATATCTG
MATβR328/317stop Forward	His_Par/Pet28		CATTGACAAGAGATGGTGACAAACGGTCTTTC
MATβR328/317stop Reverse	His_Par/Pet28		GAAAGACCGTTTGTCACCATCTCTTGTCAATG
MATβT330/319stop Forward	His_Par/Pet28		GACAAGAGATGGAGACAATAGGTCTTTCATTAGCTCG
MATβT330/319stop Reverse	His_Par/Pet28		CGAGCTAATGAAAGACCTATTGTCTCCATCTCTTGTC
MATβH331/322stop Forward	His_Par/Pet28		GGAGACAAACGGTCTTTTAATAGCTCGAGCACCACC
MATβH331/322 stop Reverse	His_Par/Pet28		GGTGGTGCTCGAGCTATTAAAAGACCGTTTGTCTCC
MATα2SUMO Forward	pET_SUMO	Age1	TTTTTACCGGTGGAATGAACGGACAGCTCAACGGC
MATα2SUMO Reverse	pET_SUMO	XhoI	TTTTTCTCGAGTCAATATTTAAGCTTCTTGGGCACTTCCC
MATβV1SUMO Forward	pET_SUMO	Age1	TTTTTACCGGTGGAATGGTGGGGCGGGAGAAAGAGCT
MATβV1SUMO Reverse	pET_SUMO	XhoI	TTTTTCTCGAGTCACTAATGAAAGACCGTTTGTCTCCATCTCTTG
MATβV2SUMO Forward	pET_SUMO	Age1	TTTTTACCGGTGGAATGCCTGAAATGCCAGAGGACATGGAG
MATβV2SUMO Reverse	pET_SUMO	XhoI	TTTTTCTCGAGTCACTAATGAAAGACCGTTTGTCTCCATCTCTTG
GIT1 Forward	pET_SUMO	XhoI	TTTTTACCGGTGGAATGTCCCGAAAGGGGCCGCGAGCG
GIT1 Reverse	pET_SUMO	Not1	TTTTTGCGGCCGCTCACTGCTTCTTCTCTCGGGTGGTGATGGTGACC

4.2 Methods

4.2.1 DNA amplification

Polymerase chain reaction (PCR) was used to amplify DNA sequences was achieved by the using Phusion® High-Fidelity DNA polymerase (NEB) and primers specified in Table 4.4. Reaction mixture consisted of the following:

Table 4.4. Composition of PCR reaction.

	Volume (µL)
Template DNA (1 µg)	1
Forward primer 10 mM	2
Reverse primer 10 mM	2
dNTP	0.5
Phusion buffer	5
Phusion enzyme	0.5
DMSO	1.5
ddH_20	12.5

4.2.2 Agarose gel electrophoresis

Amplified DNA fragments and plasmid DNA were analysed by 1.5 % (w/v) agarose gel electrophoresis. Agarose D1 Low electroendosmosis was dissolved in TAE (Tris-acetate-EDTA) 1x buffer and 0.25 µg/ml SYBR® Safe DNA gel stain was added to the solution. DNA samples were mixed with the loading buffer (Thermo) in a 5:1 (v/v) ratio and GeneRuler™ 1 kb DNA ladder (Thermo) was used as molecular weight marker. Electrophoresis was run in TAE 1x buffer.

4.2.3 Colony PCR

The presence of the recombinant plasmids was detected using colony PCR. One colony was suspended into a tube containing Master Mix reaction (Thermo):

Table 4.5. Composition of colony PCR

	Volume (µL)
Master Mix	10
Forward primer 10 mM	0.5
Reverse primer 10 mM	0.5
Water	9

This was then subjected to PCR following using the following protocol:

Table 4.6. PCR cycling conditions for amplification of MATα2, MATβV1 and MATβV2 sequences.

	Stage	Temperature (^{0}C)	Time (s)
	Initial Denaturation	95	600
25 cycles	Denaturation	95	30
	Annealing	50-58*	30
	Elongation	72	60
	Final Elongation	72	600

*Annealing temperature varied depending on melting temperature of primers that were used.

The success of the ligation process was determined by a band appearing with the size of the inserted DNA after being run on agarose gel electrophoresis

4.2.4 Transformations

DNA was either transformed into XL1Blue or DH5α strains. Fifty µl of competent cells was mixed with 50 ng of DNA for 30 min on ice. Then, the mixture was subjected to a heatshock at 42 °C for 45 s. After 5 min incubation on ice, 800 µl of LB medium was added and cells were incubated at 37 °C for 1 hour. Transformed cells were spread on an agar plate supplemented with the appropriate antibiotic (ampicillin at 100 µg/mL or kanamycin at 50 µg/mL) and incubated overnight at 37 °C.

4.2.5 DNA sequencing

DNA sequencing of the recombinant plasmids was performed by the DNA Sequencing Service of the Centro Nacional de Investigaciones Oncológicas (CNIO, Madrid) and compared to the expected sequence using ClustalW server (Thompson et al. 1994). For pET28 vectors standard primers were used for sequencing as provided by CNIO, but for other vectors sequencing primers were designed and produced (Life Tech), and can be found in table 4.7.

Table 4.7. List of sequencing primers from 5'-3'

Name	Sequence
pNIC_forward	TGTGAGCGGATAACAATTCC
pNIC_reverse	AGCAGCCAACTCAGCTTCC
pET_Sumo_forward	GAGGACACCATCGACGTGTT
pET_Sumo_reverse	TTCACTTGTAGAGCTCGTCCAT
His_parallel_forward	CGTCCGGCGTAGAGGATCG
His_parallel_reverse	CTCAAGACCCGTTTAGAGGC

4.2.6 Site-directed mutagenesis

Point mutation mutants of MAT proteins were produced using site-directed mutagenesis and reaction mix is shown in Table 4.4. Forward and reverse primers are summarised in Table 4.3. Phusion® High-Fidelity DNA Polymerase (NEB) PCR protocol was followed. PCR products were subjected to Dpn1 (NEB) digest to remove template DNA for 2 hours at 37 °C then 20 mins at 80 °C before PCR product was transformed into DH5α cells and plated on antibiotic selective agar plates and incubated overnight at 37 °C.

4.2.7 Cloning of MATα2, MATβV1 and MATβV2 genes into pETSUMO

MATα2, MATβV1 and MATβV2 DNA were amplified by polymerase chain reaction (PCR) using primers MATα2SUMO Forward & Reverse, MATβV1SUMO Forward & Reverse and MATβV2SUMO Forward & Reverse (Table 4.3) with conditions described

in Table 4.8. MATα2_PNIC, MATβV1_HisParallel and MATβV2_pET28^{+} vectors were used as template DNA and followed the following cycling conditions:

Table 4.8. PCR cycling conditions for amplification of MATα2, MATβV1 and MATβV2 sequences used in cloning.

	Stage	Temperature (^{0}C)	Time (s)
	Initial Denaturation	95	120
25 cycles	Denaturation	95	10
	Annealing	61	30
	Elongation	72	30
	Final Elongation	72	600

Amplified DNA was cleaned using Wizard® SV Gel and PCR Clean-Up System (Promega). pET_SUMO vector and Amplified DNA sequences were digested overnight at 37 ^{0}C using AgeI and XhoI (NEB) restriction enzymes. Digested pET_SUMO vector was cleaned and dephosphorylated for 1 hour at 37 ^{0}C using Antarctic Phosphatase (NEB). Fragment MATα2 (1185 bp), MATβV1 (1002 bp) and MATβV2 (969 bp) and digested pET_SUMO (5516 bp) were extracted from a 1x TAE agarose gel after electrophoresis at 100 V for 50 minutes in 1 x TAE buffer followed by purification using the Wizard® SV Gel and PCR Clean-Up System. Purified amplified DNA was ligated for 1 hour at 22 ^{0}C into AgeI-XhoI digested pET_SUMO vector using T4 DNA ligase (Fermentos), transformed into DH5α strain *E.coli*, plated on to agar media with added kanamycin (50 μg/mL) and left at 37 ^{0}C overnight. Colony PCR was used to test for the presence of the desired insert using the pET_SUMO_forward and pET_SUMO_reverse sequencing primers and PCR master mix (Thermo). The following PCR cycling conditions were used (Table 4.9):

Table 4.9. Colony PCR cycling conditions for amplification of MATα2, MATβV1 and MATβV2 sequences

	Stage	Temperature (^{0}C)	Time (s)
	Initial Denaturation	95	120
30 cycles	Denaturation	95	30
	Annealing	50	30
	Elongation	72	60
	Final Elongation	72	600

Samples were run on a 1x TAE agarose gel at 100 V for 50 minutes in 1 x TAE buffer. Colonies containing the MATα2, MATβV1 and MATβV2 DNA were grown overnight in selective LB media at 37 ^{0}C and plasmids were extracted by Wizard SV Minipreps DNA Purification System (Promega). DNA was sent to and sequenced by CNIO.

4.2.8 Cloning of GIT1 into pETSUMO vector

Human GIT1 gene was purchased from Addgene (15526) in the pEGFP vector. Cloning of GIT1 into pETSUMO vector was the same as stated in section 4.2.7, but with the use of GIT1_forward and GIT1_reverse as the primers and the vector, pETSUMO, and amplified insert were digested using the restriction enzymes Xho1 and Not1.

4.2.9 Production of TEV protease

TEV expression was carried out in *E. coli* BL21 (DE3). Cells were grown in LB medium at 37 °C to A_{600}= 0.6-0.8 at which protein expression was induced by the addition of 1 mM isopropyl β-D-1 IPTG (GoldBio), at 20 °C overnight. Cultures were centrifuged at 6000x g for 20 minutes at 4 °C and the pellet was taken and was flash frozen using liquid nitrogen before being stored at -80 °C. Purification of TEV protein was carried out following the protocol of Alsarraf *et al* (Alsarraf et al. 2011).

4.2.10 Production of SENP2 Protease

The expression of SENP2 was the same as TEV protease (section 4,2,9). Cell pellets were lysed at 4°C using high-pressure homogenisation at 27 Kpsi (Constant System, Ltd, UK) in lysis buffer (20 mM Tris-HCl pH 8.0, 1 mM β-mercaptoethanol (β-ME), 350 mM NaCl, 10 mM imidazole, 10 μg/ml DNAse, 20 μg/ml lysozyme) and the cell homogenate was clarified by centrifugation at 20,000x g for 40 minutes. Clarified supernatant was loaded into nickel resin equilibrated with lysis buffer. Column was washed with lysis buffer and then by wash buffer (20 mM Tris-HCl pH 8.0, 350 mM NaCl, 1 mM β-ME, and 20 mM imidazole). Proteins were eluted with the elution buffer (20 mM Tris-HCl pH 8.0, 350 mM NaCl, 1 mM β-ME, and 400 mM imidazole). SENP2 was then loaded into an ion-exchange chromatography column, HiTrap Q HP column (GE Healthcare), that was pre-equilibrated with 20 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM β-ME, then the purification was performed using an isocratic gradient from 0.05 M to 1 M NaCl. Fractions containing SENP2 were collected and 10 % (v/v) glycerol was added before protein was flash frozen at -80 °C.

4.2.11 Purification of MATα1, MATα2 (pNIC vector), MATβV1 (His-Parallel) and MATβV2 (pET28)

The expression of MATβV1 was carried out in *E. coli* BL21 (HC41) and the expression of other three proteins in *E. coli* BL21 (DE3) strain. Cells were grown in LB medium at 37 °C to A_{600}= 0.6-0.8 at which protein expression was induced by the addition of 1 mM isopropyl β-D-1 IPTG (GoldBio), at 20 °C overnight. Cultures were centrifuged at 6000x g for 20 minutes at 4 °C and the pellet was taken and flash frozen using liquid nitrogen before being stored at -80 °C.

Cell pellets were lysed at 4°C using high-pressure homogenization at 27 Kpsi (Constant System, Ltd, UK) in lysis buffer (500 mM NaCl, 5% (v/v) glycerol, 5 mM Imidazole, 10 mM β-ME) and cell homogenate was clarified by centrifugation at 20,000x g for 40 minutes. Clarified supernatant was loaded into nickel resin equilibrated with lysis buffer. Column was washed with lysis buffer and then by wash buffer (500 mM NaCl, 5% (v/v) glycerol, 30 mM Imidazole and 10 mM β-ME). Proteins were eluted with the elution buffer (500 mM NaCl, 250 mM Imidazol, 10mM β-ME) and the Histag was then cleaved from MATα2, MATα1 and MATβV1 by incubation overnight with tobacco etch virus (TEV) protease and with thrombin in the case of MATβV2**.** MATα2, MATα1 and MATβV2 were then loaded into an ion-exchange chromatography column, HiTrap Q HP (GE Healthcare), and MATβV1 in a HiTrap S HP, that were pre-equilibrated with Buffer A (20 mM Tris-HCl pH 8.0, 50 mM NaCl, 5 mM β-ME), then the purification was performed using an isocratic gradient from 0.05 M to 1 M NaCl. Selected fractions of MATα2 and MATα1 were then concentrated and loaded onto a HiLoad 16/60 Superdex 200 gel filtration column (GE Healthcare) and fractions of MATβV2 and MATβV1 were loaded onto HiLoad 16/60 Superdex 75. Finally, fractions containing pure protein were pooled and stored at -80 ºC. Buffer used in each purification step was 25 mM HEPES pH 7.5 unless otherwise states.

4.2.12 Purification of MATα2, MATβV1 and MATβV2 (pET_SUMO)

The expression, induction, lysis of cell pellets and affinity chromatography of the pET_SUMO constructs was the same as described above. After the protein was eluted from the nickel resin, 100 μg/mL SENP2 enzyme was added and the protein was left at

room temperature for one hour with agitation before further overnight incubation at 4 ^{0}C and dialyzed into 25 mM HEPES pH 7.5, 50 mM NaCl and 10 mM β-ME. Proteins were then loaded into an ion-exchange chromatography column, HiTrap Q HP column (GE Healthcare), that was pre-equilibrated with Buffer A (25 mM HEPES pH 7.5, 25 mM NaCl, 5 mM β-ME), then the purification was performed using an isocratic gradient from 0.05 M to 1 M NaCl. Selected fractions of MATα2 and MATα1 were then concentrated and loaded into a HiLoad 16/60 Superdex 200 Gel Filtration column (GE Healthcare) and fractions of MATβV2 and MATβV1 into HiLoad 16/60 Superdex 75. Finally, fractions containing pure protein were pooled and stored at -80 ºC.

4.2.13 Assembly of MAT protein complexes

In order to assemble the complexes, either MATα1 or MATα2 were incubated together with MATβV1 or MATβV2 for 1 hour at 4 ^{0}C in 50 mM HEPES buffer pH 7.5 containing 10 mM $MgCl_2$, 50 mM KCl, ± 300–500 mM NADP. The complex was then loaded onto a Superdex 200 10/300 column and eluted with buffer consisting of 200 mM NaCl, 25 mM HEPES pH 7.5, 1 mM $MgCl_2$, 5mM KCl, 1 mM TCEP [tris(2-carboxyethyl)phosphine]. Concentrations of MAT protein varied to form the MATαβ complexes,, but were always in a ratio of MATα:MATβ protein of 1:2.

4.2.14 Purification of GIT1 proteins

Growth, induction expression, centrifugation, lysis and nickel affinity chromatography of GIT1 and GIT1 mutants was the same as described in section 4.2.12. GIT1 proteins were incubated with 100 μg/mL SENP2 for 1 hour at room temperature before further

incubation overnight at 4 °C. Proteins were then loaded into an ion-exchange chromatography column, HiTrap Q HP (GE Healthcare) that was pre-equilibrated with Buffer A (25 mM Tris pH 8.0, 50 mM NaCl, 10 mM β-ME), then the purification was performed using an isocratic gradient from 0.05 M to 1 M NaCl. Selected fractions of GIT1_R413stop were then concentrated and loaded into a HiLoad 16/60 Superdex 200 gel filtration column (GE Healthcare).

4.2.15 Polyacrylamide gel electrophoresis (PAGE)

Analysis of the proteins was done by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) in 10-12% (w/v) acrylamide gels. Samples were mixed with Laemmli loading buffer 2x (Bio-Rad) supplemented with 10 mM β-ME and heated at 90 °C for 10 min. Broad range (Bio-Rad) pre-stained standard was used as a molecular weight markers. Electrophoresis was carried out at 150 V with standard running buffer (0.025 M Tris, 0.192M glycine, 0.1% (w/v) SDS, pH 7.5).

4.2.16 Native-PAGE

NativePAGE™ Novex® 4-16 % Bis-Tris gels were purchased from Invitrogen and the protocol used was the same as the one described by the manufacturer.

4.2.17 Staining of gels

Gels were stained for 1 hour in a solution containing 0.1% (w/v) Coomassie Brilliant Blue R-250, 40 % (v/v) methanol and 10 % (v/v) glacial acetic acid before background staining was eliminated by washing with 40 % (v/v) methanol 10 % (v/v) glacial acid solution.

4.2.18 Activity assays

Stock solutions of MATα2, MAT$(\alpha2)_4(\beta V1)_2$ and MAT$(\alpha2)_4(\beta V2)_2$ were diluted in dilution buffer (10 mM HEPES pH 7.5, 500 mM NaCl, 5% (v/v) glycerol and 0.5 mM TCEP) to concentrations of 250 nM, 125 nM and 62.5 nM, respectively. The reaction buffer consisted of 25 mM HEPES pH 7.5, 5 mM $MgCl_2$, 5 mM KCl. The methionine solutions (25-1000 μM) were prepared in dilution buffer. For measurements with NADP, stock 0.5 mM NADP, prepared in water was added to the reaction buffer.

Reactions were started and final concentrations were achieved upon the addition of 40 μL 5 mM ATP, diluted in dilution buffer. Each sample was pre-incubated for 15 minutes with methionine and reaction buffer prior to the addition of ATP. Final concentrations of proteins were 50 nM, 25 nM and 12.5 nM for MATα2, MAT$(\alpha2)_4(\beta V1)_2$ and MAT$(\alpha2)_4(\beta V2)_2$ respectively. The reaction mixtures were thermostated and agitated (37 °C, 1400 rpm). After 10 minutes the reactions were quenched with 75 % (v/v) acetonitrile and 1.2 % (v/v) formic acid in water. To ensure all reactions had been stopped, samples were shaken (4 °C, 1400 rpm) before centrifugation (14,000 rpm) and transferred to 96-well plates for ultra-pure liquid chromatography mass spectrometry (UPLC-MS) analysis (Van Liempd et al. 2013).

Samples were injected into an ACQUITY UPLC system in a randomised order and the detection and quantification of SAMe and methionine was carried out. Briefly, upon injection, polar metabolites bind to the UPLC-column and then are eluted in order of increasing polarity. Elutes are subjected to time of flight mass spectroscopy to determine identity. Ten point calibration curves with exponentially spaced concentrations were used

for quantification of samples. The rate of SAMe formation was calculated (fmole/s/nmol MAT) for each substrate concentration. Prism 4.0 (www.graphpad.com) software was used to fit the enzyme kinetic data with the Michaealis-Menten equation in order to obtain values for k_{cat}, V_{max} and K_m.

4.2.19 Differential scanning fluorimetry (DSF)

MATα2 was purified as described in section 4.2.11. Differential scanning fluorimetry was performed in a 96-well PCR plate using an Mx3005p RT (real-time)-PCR machine (Stratagene). Each well (25 µl) consisted of MATα2 protein (0.2-1.2 mg/mL) diluted in 25 mM HEPES pH 7.5, 5 mM $MgCl_2$, 5 mM KCl (section 4.2.18) and Sypro Orange (Invitrogen). Wells also contained added compounds including SAMe (1-10 mM), PEG 600 (2.5-20 % (w/v), and imidazole (1-100 mM). Fluorescence intensities were measured from 25 °C to 96 °C with a ramp rate of 1° C/min. Sypro Orange dye (Invitrogen) was diluted as described in the Invitrogen thermal stability assay protocol. For the analysis of the data, each of the melting curves was normalized. This was done by taking the minimum value of each curve and subtracting it from all the points on that curve followed by dividing the melting curve by its maximum value. This normalization enables each melting curve to be compared to each other. The first derivative of the melting curve was calculated which is used to give the melting temperature. The melting temperature is the temperature that corresponds to the maximum point on of the first derivative curve (Sun et al. 2015).

4.2.20 Circular Dichroism

Circular dichroism data were recorded on a J-815 circular dichroism spectrometer (Jasco, UK) equipped with thermoelectric temperature control. Samples of MATβV1 and

MATβV2 and the mutant variants were diluted to 1-5 μM in 0.1 M phosphate buffer pH 7.0 and placed in a 2-mm cell. All circular dichroism wavelength scans were performed at a constant temperature of 25 ºC between 200 and 280 nm with a scanning speed of 50 nm ⁄ min and a data pitch of 1 nm. Data analysis was done using Origin and JASCO Spectra Manager integrated software. Raw data were converted to Ellipticity*10^{-3} (deg cm^2 $dmol^{-1}$) using Microsoft Excel.

4.2.21 Isothermal titration calorimetry (ITC)

In order to examine the association constant (K_a) of MATαβ complexes, MATα2, MATα1, MATβV1, MATβV2 and any MATmutant variants were buffer-exchanged by gel filtration on a Superdex 200 10/300 GL column equilibrated with 200 mM NaCl, 20 mM HEPES pH 7.5 buffer before ITC analysis. Subsequently, MATβ isoforms and mutants were injected into MATα2 or MATα1 solution in aliquots of 10 or 20 μl, respectively. The concentration of MAT proteins were: MATα2=10.0 μM, MATβV1=51.2 μM, MATβV2=36.5 μM, MATβV1Δ320 and MATβV2Δ309= 100 μM. For reactions using MATα1, 45.1 μM of MATα1 and 236.8 μM of MATβV1 or MATβV2 were used. To verify the interaction of NADP with different MATβ isoforms, each protein was buffer-exchanged by gel filtration on a Superdex 200 10/300 GL column equilibrated in 5 mM $MgCl_2$, 5 mM KCl, 20 mM HEPES pH 7.5, 200 mM NaCl before ITC analysis. NADP was also diluted in the same buffer. NADP was injected into MATβV2 in aliquots of 15 μl and into MATβV1 in aliquots of 10 μl. For measurements with NADP, proteins were buffer exchanged into 200 mM NaCl, 20 mM HEPES pH 7.5. The concentration of MAT proteins were: MATβV1=29.8 μM, MATβV2=30.0 μM, $MAT(\alpha2)_4(\beta V2)_2$=10.2 μM. 500 μM NADP was titrated against the MATβ proteins whilst 155 μM NADP was titrated against the $MAT(\alpha2)_4(\beta V2)_2$ complex. All ITC

measurements were carried out at 25 °C on a VP-ITC Microcalorimeter (MicroCal/GE Healthcare). The ITC data were processed using Origin software (OriginLab Corp., USA).

4.2.22 Crystallisation and data collection

Crystals of the MAT$(\alpha 2)_4(\beta V2)_2$ complex appeared at 25 °C within 1–2 days in drops containing 2 µl MAT$(\alpha 2)_4(\beta V2)_2$ complex at 5.8 mg ml^{-1} mixed with 1 µl precipitant solution of 100 mM MES/imidazole buffer pH 6.5, 10% (v/v) ethylene glycol, 20% (w/v) PEG 8000. Before crystallization the MAT$(\alpha 2)_4(\beta V2)_2$ complex was incubated with its product SAMe (1 mM), or its substrate ATP (1 mM) or AMP-PNP (250 µM) and MET (1 mM). These compounds were added to the precipitant and cryo solution. Cryo solution consisted of well condition with 20 % (v/v) ethylene glycol.

For structures described in chapter 6 that contained only MATα2, MATα2 and MATβV2H323stop proteins were incubated together for 1 h at 4 ^{0}C in 50 mM HEPES buffer pH 7.5, 10 mM $MgCl_2$, 50 mM KCl, 100 µM AMP-PNP (Sigma), 10 mM DTT with either 10 mM methionine (Sigma) or 10 mM ethionine (Sigma). Each of these two complexes was then loaded onto a Superdex 200 10/300 column (GE Healthcare) and eluted using 25 mM HEPES pH 7.5, 200 mM NaCl, 1 mM $MgCl_2$, 5 mM KCl 1 mM TCEP. Before crystallization the MAT$(\alpha 2)_4(\beta V2H323stop)_2$, that was subjected to gel filtration with AMP-PNP and methionine was incubated with AMP-PNP (250 µM) or selenium-methionine (250 µM) and these resulted in the crystals SAMe+ADO+MET+PPNP bound and PPNP bound, respectively. Crystallization of MAT$(\alpha 2)_4(\beta V2H323stop)_2$, that was subjected to gel filtration with AMP-PNP and ethionine was incubated with ethionine (250 µM) and resulted in the SAE bound structure. For all crystallisation, protein concentration was 5.8 mg/mL.

SAMe+ADO+MET+PPNP bound and PPNP bound crystals appeared at 25 ^{0}C within 2-3 days in drops containing 1 µL of 0.1 M HEPES pH 7.5, 30 % (w/v) PEG 600 with 1 µL of MATα2. For SAE, crystals appeared at 18 ^{0}C within 3-4 days in drops containing 800 nL of 0.1 M Imidazole pH 7.0, 50 % (w/v) MPD with 800 nL of MATα2. Crystals were cryoprotected in the reservoir solution and flash frozen in liquid nitrogen and for the SAE bound crystal, 20 % (v/v) ethylene glycol was added for the cryoprotectant.

Different data sets were collected at the PROXIMA1, XALOC, BM29, I04 and I24 beamlines at SOLEIL (St Aubin, France), ALBA (Barcelona, Spain), ESRF (Grenoble) and Diamond (Oxford, England) synchrotron centres. Data reduction was carried out with the HKL-2000 (Otwinowski & Minor 1997) and XDS programs (Kabsch 2010). The phases were calculated with Phaser (McCoy et al. 2007) using MATα2 (PDB entry 2P02) and MATβ (PDB entry 2YDX, (Shafqat et al. 2013) as search models for molecular replacement. Model building and refinement were performed using Coot (Emsley & Cowtan 2004) PHENIX (Adams et al. 2010) and REFMAC (Murshudov et al. 2011). Data collection and refinement statistics are in table 6.3.

4.2.23 Small angle X-ray scattering data collection and processing

$MAT(\alpha2)_4(\beta V2)_2$ and $MAT(\alpha2)_4(\beta V\Delta16)_2$ complexes was assembled as described in section 4.2.13 before SAXS data were collected on the SWING beamline at the SOLEIL synchrotron, using the HPLC-integrated SAXS setup (David & Pérez 2009) with a two-dimensional AVIEX CCD detector over an angular range q_{min} = 0.01 Å^{-1} to q_{max} = 0.5 Å^{-1}. 80 µl of the $MAT(\alpha2)_4(\beta V2)_2$ or the $MAT(\alpha2)_4(\beta V\Delta16)_2$ complex at 5 mg/ml was loaded onto a pre-equilibrated Shodex KW-402.5-4F 150 kDa SEC column, 250 frames

of SAXS data were taken over the course of protein elution. MATα1, MATα2 and GIT1_R413stop, proteins were buffer-exchanged using a Superdex 200 10/300 column into 25 mM HEPES pH 7.5, 200 mM NaCl, 1 mM DTT and addition of 5 % (v/v) glycerol before being concentrated to 7.7 mg/mL, 8.7 mg/mL and 10 mg/mL for MATα1, MATα2 and GIT1_R413stop respectively. Data averaging and reduction, which included preliminary calculations of R_g and $I(0)$ were carried out with the Foxtrot suite, developed at SOLEIL for the SWING beamline. Further analyses were performed with the ATSAS suite (Petoukhov et al. 2012) this included Primus (Konarev et al. 2003) for determination of radius of gyration, which fulfilled the condition $qR_{(g)}<1.3$. Particle distance distribution function $p(r)$ analysis was carried out with GNOM refining the D_{max} value based on the Guinier (R_g) values. In order to generate an *ab initio* model twenty runs of DAMMIN (Svergun 1999) were performed without the use of symmetry restraints, and after averaging and filtering a model was produced. SUPCOMB (Kozin & Svergun 2001) was used to compare ab initio model with crystallographic data whilst CRYSOL (Barberato et al. 1995) compared the 1D scattering cure with the theoretical scatter of the $MAT(\alpha2)_4(\beta V2)_2$ complex.

Static SAXS data sets were collected at BM29 at ESRF at 5.5 mg/mL and 3.0 mg/mL for MATβV1 whilst MATβV2 was collected at 6.0 mg/mL and 3.0 mg/mL. For data collections with NADP, MATβ proteins were incubated with NADP for 1 hour before gel filtration was used to remove excess NADP. Samples were then concentrated and frozen. MATβV1+NADP data sets were collected at 2.0 mg/mL and 6.0 mg/mL whilst MATβV2+NADP were collected at 3.0 mg/mL and 6.0 mg/mL. Data averaging, reduction and merging of different concentrations were carried out using Primus (Konarev et al. 2003) Data collection was carried out at 20 °C. Further analyses were

performed with the ATSAS suite (Petoukhov et al. 2012) and ScÅtter. Molecular weights were determined using SAXS MoW (Fischer et al. 2009).

Chapter V. Structural and biochemical analysis of MATα2β complexes.

5.1 Results

5.1.1 Production of MATα & MATβ proteins.

In order to obtain protein with a high enough purity for crystallisation multiple chromatography methods were needed to purify MATα2, MATβV1 and MATβV2 which is described in section 4.2.11. MATα2 eluted from a Hitrap Q HP column with a maximum conductivity at ~25 mS/cm (Fig 5.1A). The predicted isoelectric point of MATβV1 is 6.9, yet it did not bind to a Q HP column despite having a isoelectric point lower the pH of the buffer (pH 7.5) so a S HP column was used and this was found to bind MATβV1 (Fig 5.1B), which eluted at ~17 mS/cm. MATβV2 bound a Q HP column and was eluted from this at ~21 mS/cm (Fig 5.1C). All three proteins, after gel filtration, gave good purity as shown by SDS-PAGE (Fig 5.1D-F) and showed a good agreement with predicted molecular weight. Purification of MATα2, MATβV1 and MATβV2 produced different yields of purified protein. MATα2 gave 5 mg/L, MATβV1 gave 1.4 mg/L and MATβV2 produced 2.5 mg/L.

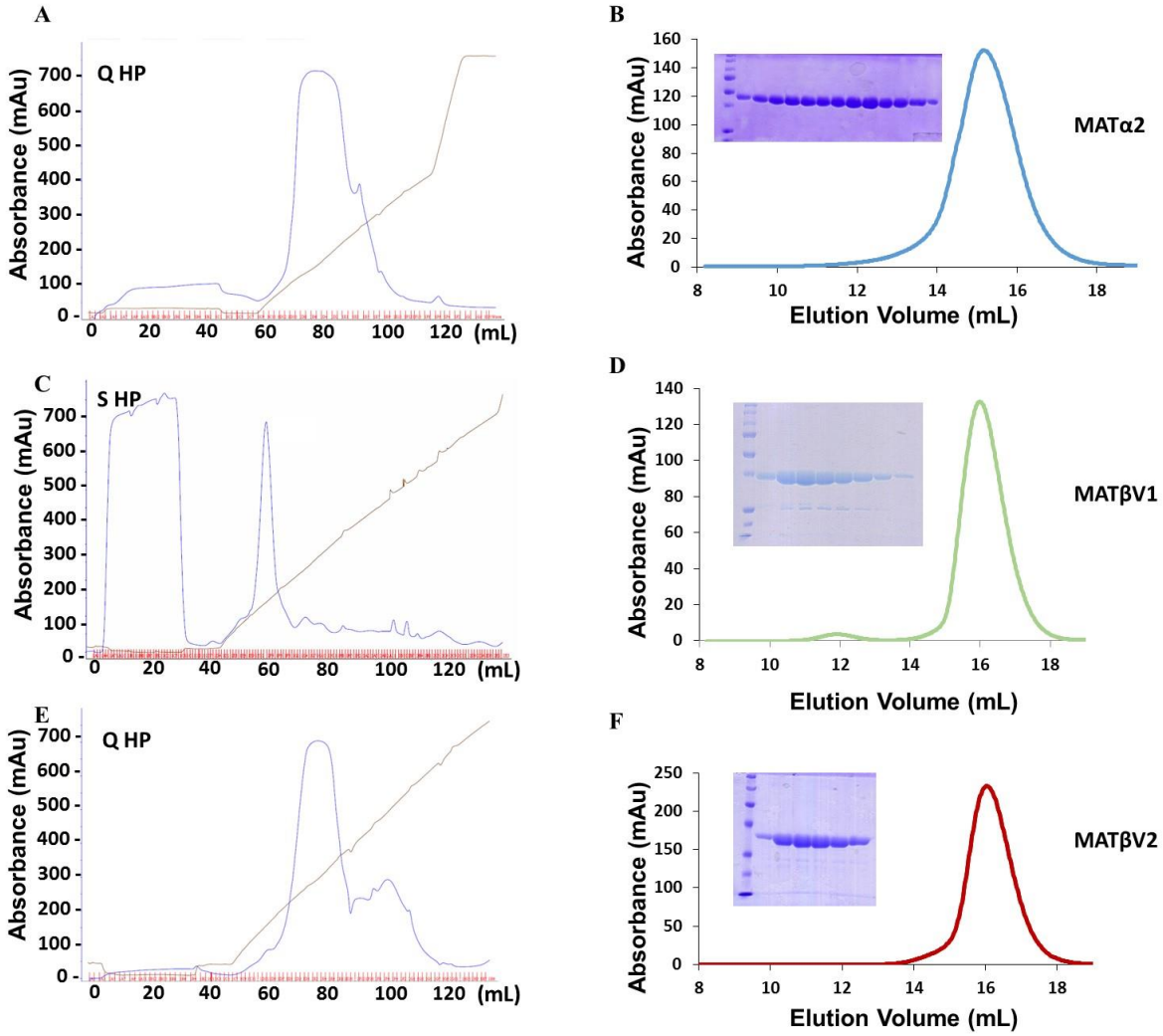

Figure 5.1 Purification of MATα2, MATβV1 and MATβV2. A. Q HP ion-exchange elution profile of MATα2 showing absorbance at 280 nm in blue. **B**. Gel filtration profile (blue) and the SDS-PAGE of purified MATα2 from fractions of the elution (12.5-17.5 mL). **C**. S HP ion-exchange spectrum of MATβV1 showing absorbance at 280 nm in blue. **D**. Gel filtration profile (green) and the SDS-PAGE of purified MATβV1 from fractions of the elution (14.0-18.0 mL). **E**. S HP ion-exchange spectrum of MATβV2 showing absorbance at 280 nm in blue. **F**. Gel filtration profile (red) and the SDS-PAGE of purified MATβV2 from fractions of the elution (14.0-18.0 mL). All panels show absorbance at 280 nm which was used as an indication of the presence of protein. A Superdex 200 10/300 column was used for gel filtration.

5.1.2 MATαβ complex formation and characterisation.

Upon incubating MATα2 with either MATβV1 or MATβV2 for 1 hour the formation of the MATαβ complexes; $MAT(\alpha2)_4(\beta V1)_2$ (green) and $MAT(\alpha2)_4(\beta V2)_2$ (red) were observed (Fig 5.2A). Gel filtration elution peaks of both complexes were at a different volume (12.0 mL) to that of MATα2 (13.5 mL) suggesting a larger molecular weight. Fractions of both complexes were run on SDS-PAGE and showed that both proteins were present in the fractions (Fig 5.2B) and also on native-PAGE which showed the complexes have a much larger molecular weight than either of the subunits alone (Fig 5.2C). The molecular weight of the complexes according to native-PAGE corresponded to ~130 kDa. Initially, to make the MATαβ complexes equal ratios of MATα2 to MATβ were incubated together, which resulted in two peaks appearing on gel filtration, one corresponding to the complex at ~12.0 mL whilst the other peak, which appeared around ~15.0 mL was MATβ. The ratio of MATα2:MATβ was then changed to 2:1 and this caused the second peak, MATβ, to disappear leaving the complex peak only.

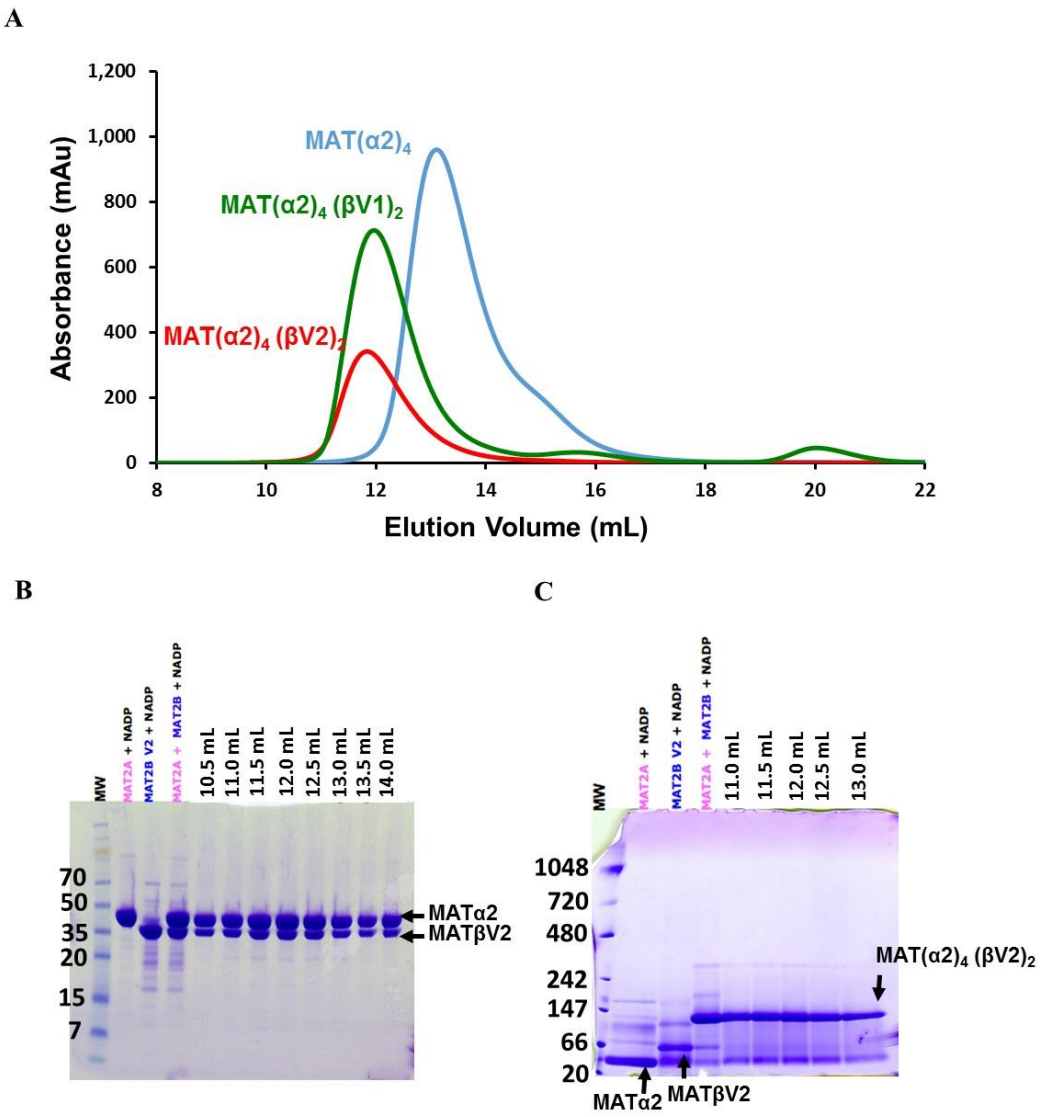

Figure 5.2. Formation of the MATαβ complexes. Purified MATα2 protein was incubated with either purified MATβV1 or MATβV2 (section 4.2.11) before being injected into a Superdex 200 10/300 column (section 4.2.13). **A**. Gel filtration profiles of 250 µM MATα2 (blue) and MATαβ complexes (green & red). $MAT(\alpha2)_4(\beta V2)_1$ consisted of 140 µM MATα2 and 70 µM of MATβV1. $MAT(\alpha2)_4(\beta V2)_2$ consisted of 100 µM MATα2 and 50 µM of MATβV2. The shift in the MATαβ complexes, shown by a lower elution volume, suggests that the MATαβ complexes had a greater molecular size compared to MATα2. Absorbance was measured at 280 nm which was used as an indication of the presence of protein. **B**. SDS-PAGE of selected fractions from the elution of $MAT(\alpha2)_4(\beta V2)_2$ showing that $MAT(\alpha2)_4(\beta V2)_2$ complexes contain two proteins. **C**. Native-PAGE showing that MATαβ complexes have a greater molecular weight (~147 kDa) than individual proteins (MATα2= ~45 kDa, MATβV2= ~66 kDa). Arrows indicate specific proteins and/or protein complexes.

5.1.3 Crystallisation and data collection, processing of MAT(α2)$_4$(βV2)$_2$.

For crystallisation, both MATα2βV1 and MATα2βV2 were concentrated to 10 mg/mL and the PCT^{TM} PreCrystallization test was done, which caused heavy precipitation of the protein. As a result, the concentration of the complexes was reduced to 6 mg/mL and the PCT^{TM} PreCrystallization test was done again and this concentration was found suitable for crystallisation trials. Numerous commercial screens were used in 96 well screening plates which allowed for two drops to be set per condition, one for MATα2βV1 and the other for MATα2βV2. After 2-3 days crystals appeared at 18 °C in a 100:100 nL drop (Fig 5.3A) in 0.1 M MES/imidazole pH 7.5, 20 % (w/v) PEG 8000, 10 % (v/v) ethylene glycol and 0.1 M carboxylic acid mix (Na-formate; NH_4-acetate; Na^3-citrate; NaK-tartrate (racemic); Na-oxamate) for the MATαβV2 complex.

These crystals were taken to a synchrotron source (section 4.2.22) for testing and diffracted to 8.5 Å (Fig 5.3B) revealing that they were protein crystals but they did not diffract to a suitable resolution to solve the structure. Matthews's coefficient was calculated based on cell dimensions of the crystal (Table 5.1) and gave a probable solution size of 258 kDa. Refinement screens were set up in order to grow bigger crystals and involved screening around the initial hit by varying the components of the screen, changing the temperature, size of drop and ratio of protein to well condition. Varying the amount of PEG 6000 or changing this for another PEG variant resulted in no crystal growth or greater precipitation. Larger crystals were able to be attained at 25 °C in a 2:1 µL drop of protein to well condition (Fig 5.3C). Before these crystals were tested a range of cryo protective solution were tested, on the home X-ray source at CIC bioGUNE, for which 20 % (v/v) additional ethylene glycol was selected. These crystals showed varied

morphology and often a cracked appearance was seen (Fig 5.3D), which caused diffraction with poorly defined spots. For crystals which contained a non-cracked morphology several data sets were collected ranging from 2.35-3.3 Å. To note no crystals of MATαβV1 complex were ever seen.

Table 5.1. Calculation of Matthews coefficient from a crystal of $MAT(\alpha 2)_4(\beta V2)_2$ that diffracted to 8.5 Å.

Model	Prob (N) for resolution	Prob (N) overall	Matthews coefficient	Solvent content (%)	Molecular weight (kDa)
1	0.015	0.016	4.930	75.100	129.000
2	0.981	0.980	2.470	50.100	258.000
3	0.005	0.005	1.640	25.200	387.000

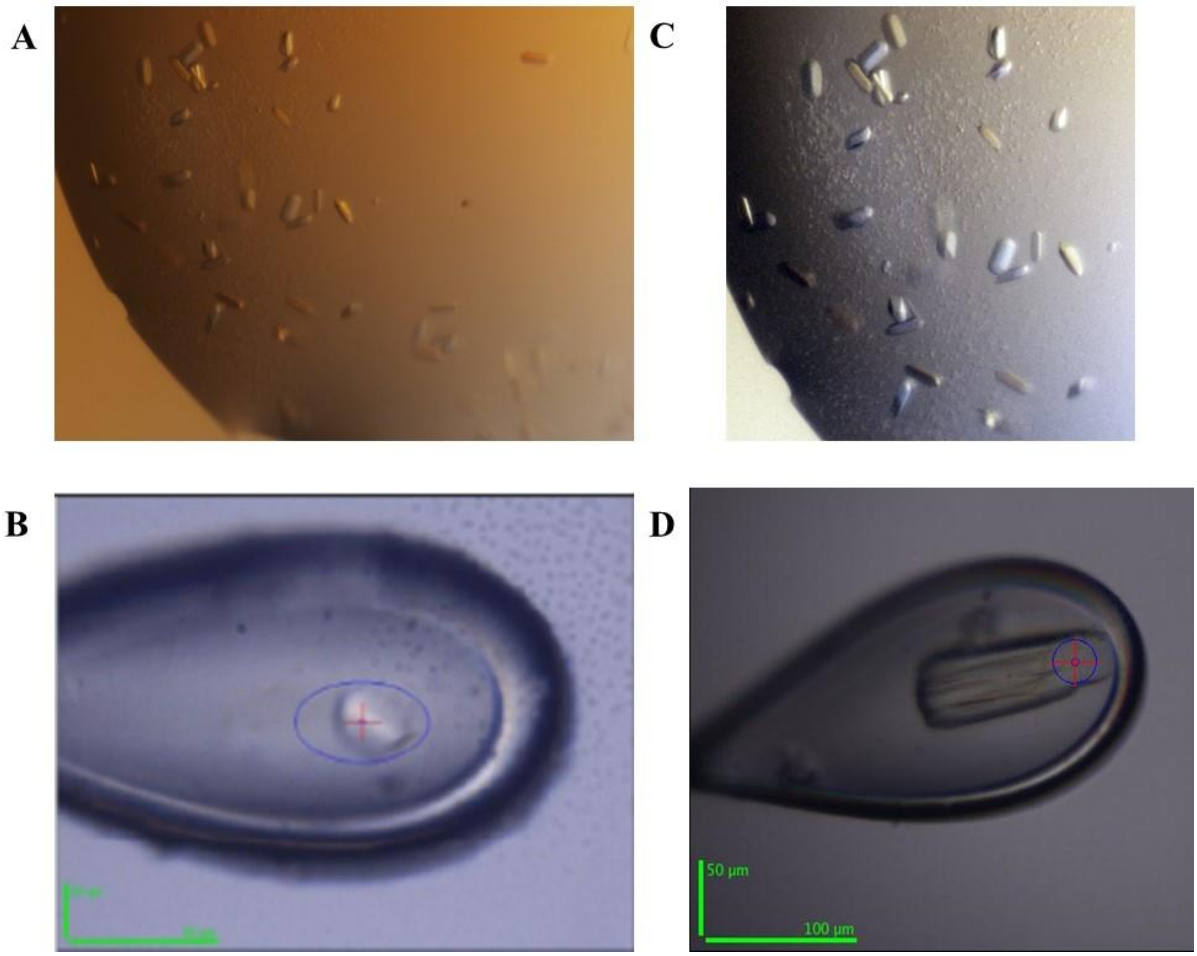

Figure 5.3. Protein crystals of MAT(α2)$_4$(βV2)$_2$. A. Initial screening hit from Morpheus crystal screen (Molecular Dimensions) at 18 °C with 200:200 nL protein: well condition size drop.. **B**. Protein crystal in loop from initial hit (30 µm by 10 µm). **C**. Refinement screen produced larger MAT(α2)$_4$(βV2)$_2$ crystals of drop size 1:1 µL protein: well condition at 25 °C. The refinement screen condition was 100 mM MES/imidazole buffer pH 6.5, 10% (v/v) ethylene glycol, 20% (v/v) PEG 8000. **D**. Larger crystals (100 µm by 30 µm) often gave a different morphology with a cracked appearance taken during data collection at PROXIMA1, SOLEIL.

5.1.4 Structural determination of the MATαβV2 complex.

Using molecular replacement with structures of MATα2 and MATβ (2P02 and 2YDX) (Shafqat et al. 2013) the crystal structure of MATαβ could be solved. Several crystallographic structures of the 258 kDa MATα2βV2 complex were solved with different ligands of resolutions ranging from 2.35 to 3.3 Å, which all revealed that the complex consists of a 185 kDa MATα2 tetramer, flanked by two MATβV2 subunits of 36.5 kDa each (Fig 5.4) giving the $MAT(\alpha2)_4(\beta V2)_2$ complex. The complex has four active sites located at the interface of the MATα2 dimers, two which contain ligands bound within the active site and two which do not contain ligands. The N-terminus of all four subunits of the MATα2 tetramer were disordered for residues 1-15. Also, the two MATα2 monomers which had empty active sites showed no density for the 'gating loop' (residues 113-131). The flexibility of this loop induces two different conformations of the catalytic subunit: opened and closed. The loop is disordered in the open conformation causing the entrance to the active site to be opened. The entrance to the active site is blocked in the closed conformation and the gating loop becomes well ordered. Each MATβV2 interacts with MATα2 through their C-terminal region flanking the tetrameric catalytic core between tunnels created by the N-domain and central domain of two MATα2 monomers. In addition, the N-terminal loop (residues 1-13) of both MATβV2 in the complex, which are orientated to the same side as the open active sites, are disordered. The Proteins, Interfaces, Structures and Assemblies (PISA) server was used to calculate if the hexameric, $MAT(\alpha2)_4(\beta V2)_2$, was the most energetically stable form (Table 5.2). The hexamer is more energetically stable (ΔG_{int}= -62 kcal/mol) then a trimeric $MAT(\alpha2)_2(\beta V2)_2$ (ΔG_{int}= -26.3 kcal/mol). Refinement statistics of the $MAT(\alpha2)_4(\beta V2)_2$ structures can be found in Table 5.3.

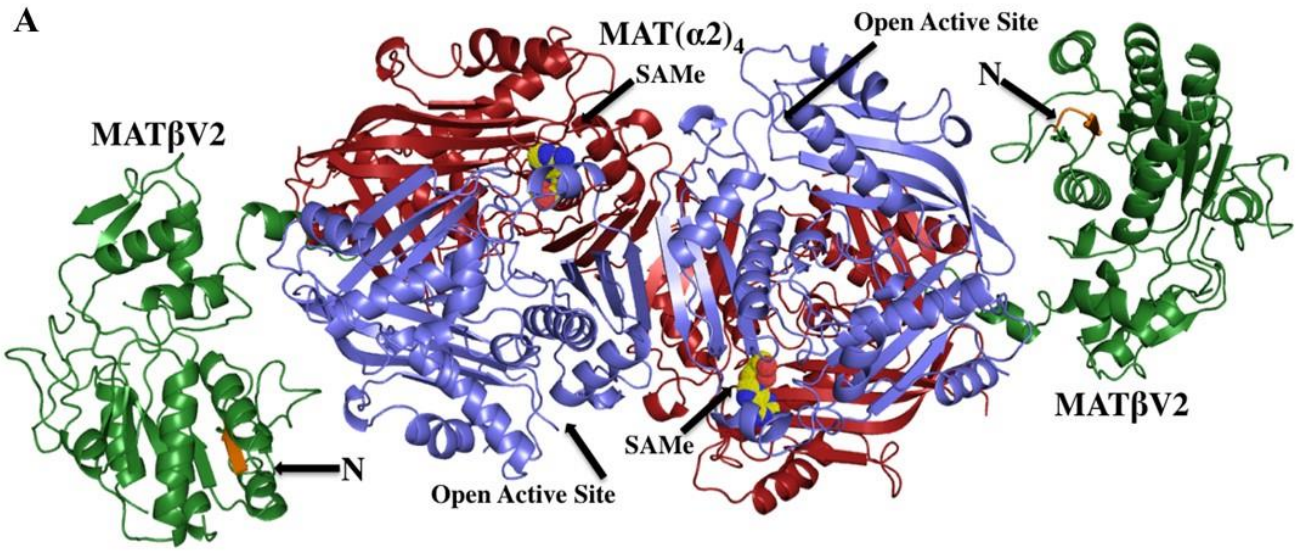

Figure 5.4. Structure of MAT(α2)$_4$(βV2)$_2$ complex. A. Crystallographic structure of MAT(α2)$_4$(βV2)$_2$ complex (protein data bank accession number 4NDN). Two MATα2 monomers with visible gating loops are coloured in slate whereas the other two MATα2 monomers are coloured in red forming a central tetramer. MATβV2 monomers are coloured in green with the first visible residues at the N-terminal in orange. The reaction product (SAMe) is represented by spheres.

Table 5.2 Analysis of protein interface of MAT(α2)$_4$(βV2)$_2$ complex ranked in order of stability of quaternary structures as determined by Proteins, Interfaces, Structures and Assemblies (PISA) server. ΔG_{int} represents the standard free energy of binding and ΔG_{diss} is the standard free energy of dissociation. Chains A-D are MATα2 and Chains E-F are MATβV2.

Rank	Composition	Chains	Stable	ΔG_{int} (kcal/mol)	ΔG_{diss} (kcal/mol)
1	$(\alpha2)_4(\beta)_2$	ABCDEF	Yes	-62.0	1.1
2	$(\alpha2)_2(\beta)_1$	ABF	Yes	-26.8	1.4
3	$(\alpha2)_2(\beta)_1$	CDE	Yes	-27.8	1.0

Table 5.3. Data collection and refinement statistics of MAT$(\alpha 2)_4(\beta V2)_2$ crystals. *Highest resolution shell is shown in parenthesis.

	SAMe-bound	ADO-bound	PPNP-bound
Data collection			
Wavelength	0.98	0.978	0.979
Detector	Pilatus 6M	Pilatus 6M	Pilatus 6M
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	72.44, 115.72, 298.45	72.09, 116.57, 299.48	72.14, 122.18, 298.42
Resolution (Å)	50-2.6 (2.69-2.6) *	108.63-3.3 (3.48-3.3) *	50-2.35 (2.43-2.35) *
R_{merge} (%)	14.4 (59.4)	13.9 (49.4)	10.1 (70.4)
$I/\sigma I$	10.12 (1.75)	6.8 (2.7)	17.3 (1.6)
Completeness (%)	98.7 (90.1)	99.9 (99.9)	98.0 (83.2)
Redundancy	5.6 (4.5)	3.3 (3.5)	7.5 (5.2)
Refinement			
Resolution (Å)	2.6	3.3	2.35
No. reflections	76590	38907	108283
R_{work}/ R_{free}	21.8/27.9	17.6/26.0	21.1/25.1
No. atoms			
Protein	15061	16370	16463
Ligand/ion	54 / 1	38 / 1	54 / 4
Water	108	37	498
B-factors (Å^2)			
Protein	76.64	71.19	58.4
Ligand/ions	SAMe / Mg 68.16 / 60.62	ADO / Mg 97.95/49.14	SAMe /PNPP /Mg 47.73 / 63.48 /39.57
Waters	56.13	30.67	48.39
Ramachandran Statistics			
Residues in Preferred Regions	1920 (94%)	1863 (90%)	2001 (96%)
Residues in Allowed regions	108 (5%)	185 (9%)	76 (4%)
Outliers	10(0.5%)	20 (1%)	8 (0.38%)
R.m.s deviations			
Bond lengths (Å)	0.006	0.008	0.005
Bond angles (º)	1.091	1.265	1.026

5.1.5 Comparison of ligand bound MAT(α2)$_4$(βV2)$_2$ structures.

In order to probe the catalytic reaction of MAT(α2)$_4$(βV2)$_2$ complex different ligands were pre incubated with the complex before crystallisation. After incubation of MAT(α2)$_4$(βV2)$_2$ complex, with its product SAMe, its substrate MET, ATP or AMP-PNP (non-hydrolysing *ATP* analog), the presence of SAMe, adenosine or PPNP were clearly observed in different crystals at the active site (Fig 5.5A-C) as supported by the corresponding omit maps (Fig 5.5D-F). In all of the structures the adenine group makes a π-π stacking with Phe250 of MATα2, supporting the hypothesis that the substrate (ATP) and product (SAMe) can occupy the active site in similar orientations. The adenine group of adenosine and SAMe is stabilised by hydrogen bonding between OH of Ser247 side chain and the main chain oxygen of Arg249. Gln113 forms a hydrogen bond with its NE_2 to the terminal oxygen of SAMe. The structures show that two of the four active sites are occupied by SAMe or adenosine whereas the other two are empty thus providing details of structural differences that accompany SAMe formation by comparison of empty and occupied sites.

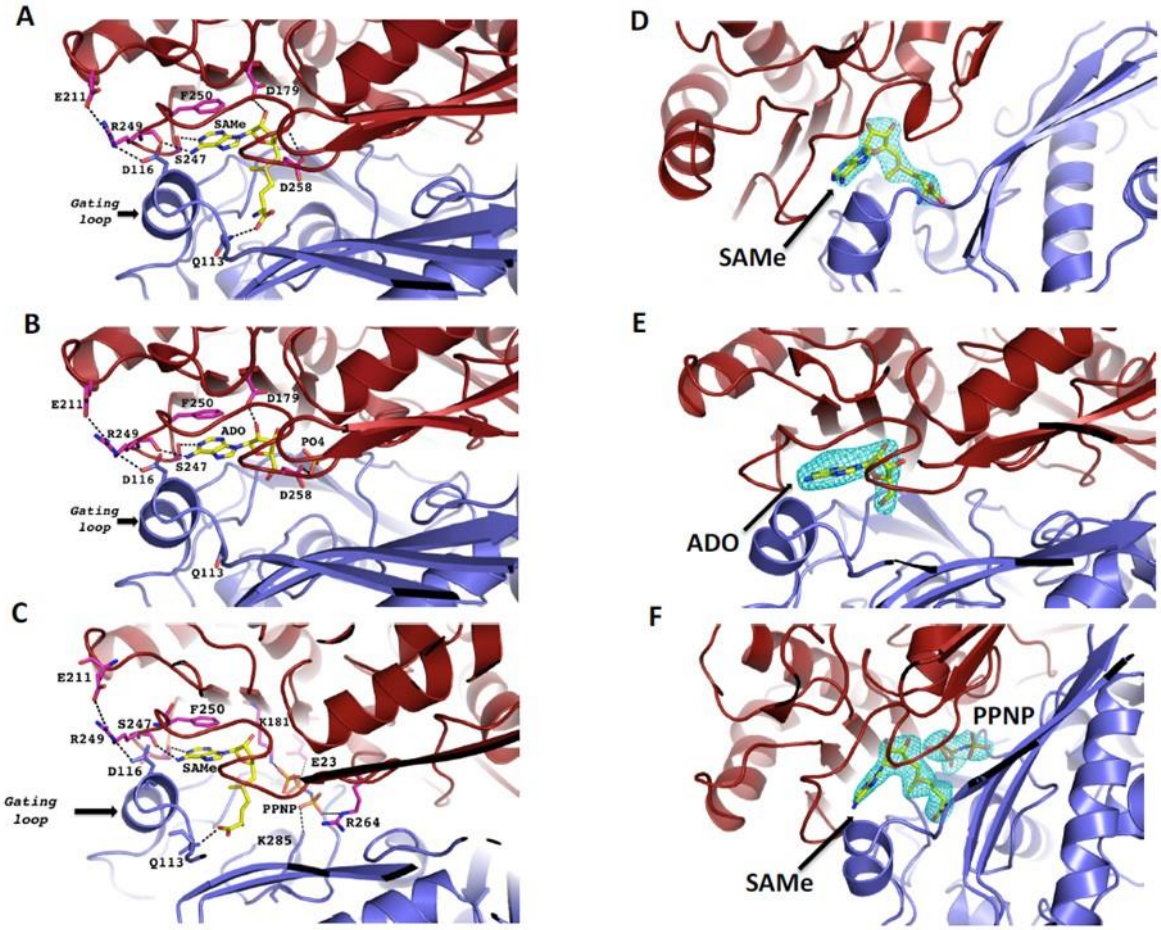

Figure 5.5. Active site comparison of MAT(α2)$_4$(βV2)$_2$ crystals. MATα2 monomer with visible gating loop is coloured in slate whereas MATα2 monomer with disordered loops is coloured in red. Important hydrogen bonds are shown as dotted lines. **A**. Stick representation of the bound product, SAMe, at the active site. **B**. Crystal co-crystallised and soaked with only ATP shows adenosine molecule and PO_4 at the active site, as a result of the ATPase and tripolphosphatase activity of MATα2. **C**. Crystal co-crystallised and soaked with ANPPNP and methionine shows SAMe and PPNP. **D**. Omit (Fo-Fc) electron density map and stick representation of bound molecule, the map is contoured at the 2.5 σ level around the SAMe molecule **E**. around the adenosine molecule and **F**. around SAMe and PPNP. Details of MATαβ complex crystallisation are in section 4.2.21.

A comparison of SAMe-bound MAT$(\alpha 2)_4(\beta V2)_2$ complex with SAMe-bound MAT$(\alpha 2)_4$ (PDB:2P02) (Shafqat et al. 2013) shows that in the absence of MAT$(\beta V2)_2$ the four active sites of MAT$(\alpha 2)_4$ are in a closed conformation (Fig 5.6A). The complex formation with MAT$(\beta V2)_2$, causes an asymmetry in MAT$(\alpha 2)_4$, in which two sites are found in an open state while the other two are in closed conformation. In the absence of SAMe, the apo-structures exhibit no density for the gating loop indicating its flexible nature (Fig 5.6B). However in MAT$(\alpha 2)_4(\beta V2)_2$ the open active sites show two additional flexible loops, near the inserted MATβV2 C-terminus (Fig 5.6C).

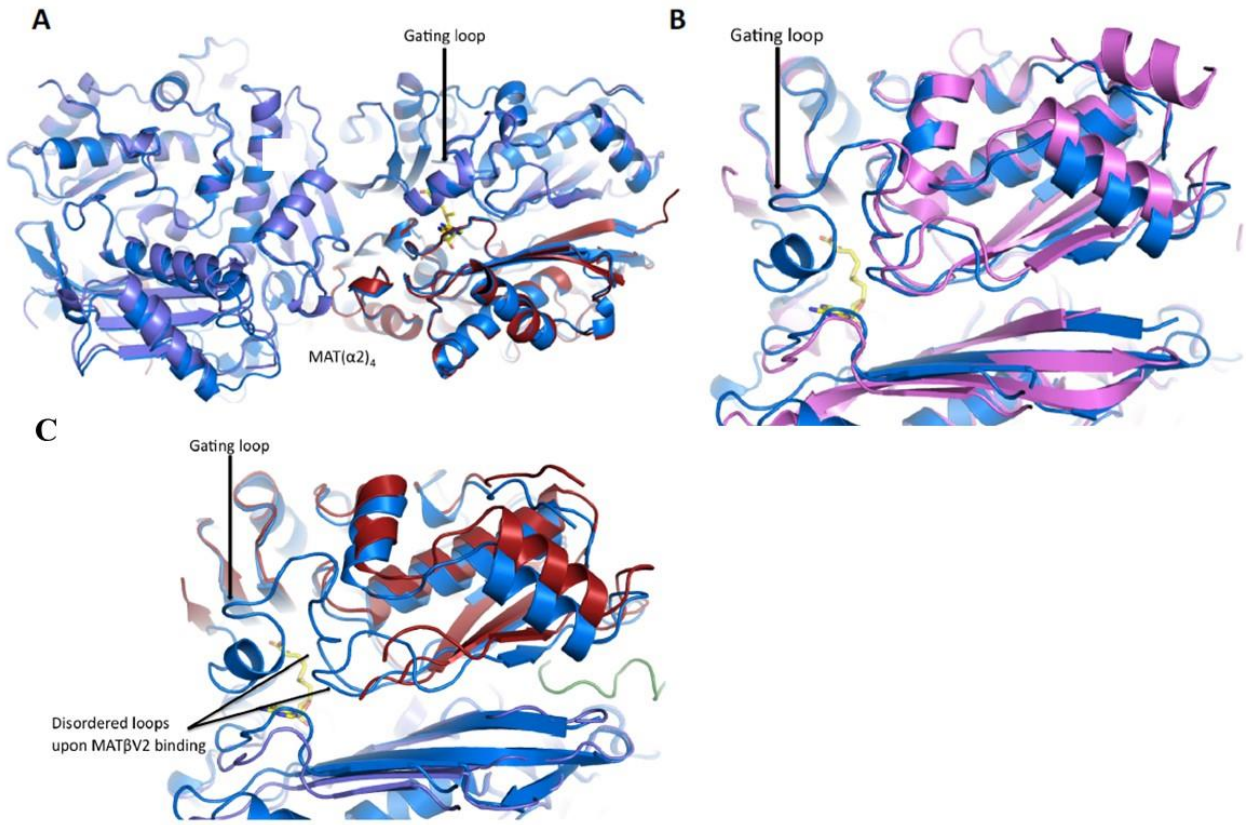

Figure 5.6. Comparison of MAT(α2)$_4$(βV2)$_2$ complex with MAT(α2)$_4$. A. Superposition of SAMe-bound MAT(α2)$_4$ (PDB:2P02 in blue) and SAMe-bound MAT(α2)$_4$ (slate, red) from the complex MAT(α2)$_4$(βV2)$_2$. In the closed state both structures are similar. **B**. Superposition of apo-MAT(α2)$_2$ from *Burkholderia pseudomallei* (PDB:3IML in pink) with the SAMe-bound MAT(α2)$_2$ (PDB:2P02 in blue), in the absence of SAMe only the gating loop is disordered. **C**. Superposition of SAMe-bound MAT(α2)$_2$ (PDB:2P02 in blue) with MAT(α2)$_2$ after MATβV2 binding, the open state in this case show two additional flexible loops, near the inserted MATβV2 C-terminus.

5.1.6 Small angle x-ray scattering of MATα2β complexes.

The oligomeric state of the crystallographic structure of $MAT(\alpha2)_4(\beta V2)_2$ is different to the suggested tetrameric form $[MAT(\alpha2)_2(\beta)_2]$ (Kotb & Kredich 1985) or the recently proposed model in which MATαβ was assumed to be a trimer $[MAT(\alpha2)_2(\beta)_1]$ (González et al. 2012). In order to confirm that the crystallographic oligomer is representative of the solution state and is thus physiologically relevant, small-angle X-ray scattering (SAXS) experiments using the HPLC-integrated SAXS set-up at the SOLEIL synchrotron were performed. The scattering curve of the complex in solution obtained by SAXS is in agreement with the theoretical curve calculated from our 2.35 Å crystal structure (X^2 = 3.2) (Fig. 5.7A), confirming that the MATα2β complex in solution does indeed have the same composition and overall conformation as observed in our crystal structures, i.e. $MAT(\alpha2)_4(\beta V2)_2$ (Fig. 5.7B). Statistics for SAXS modelling can be found in Table 5.4.

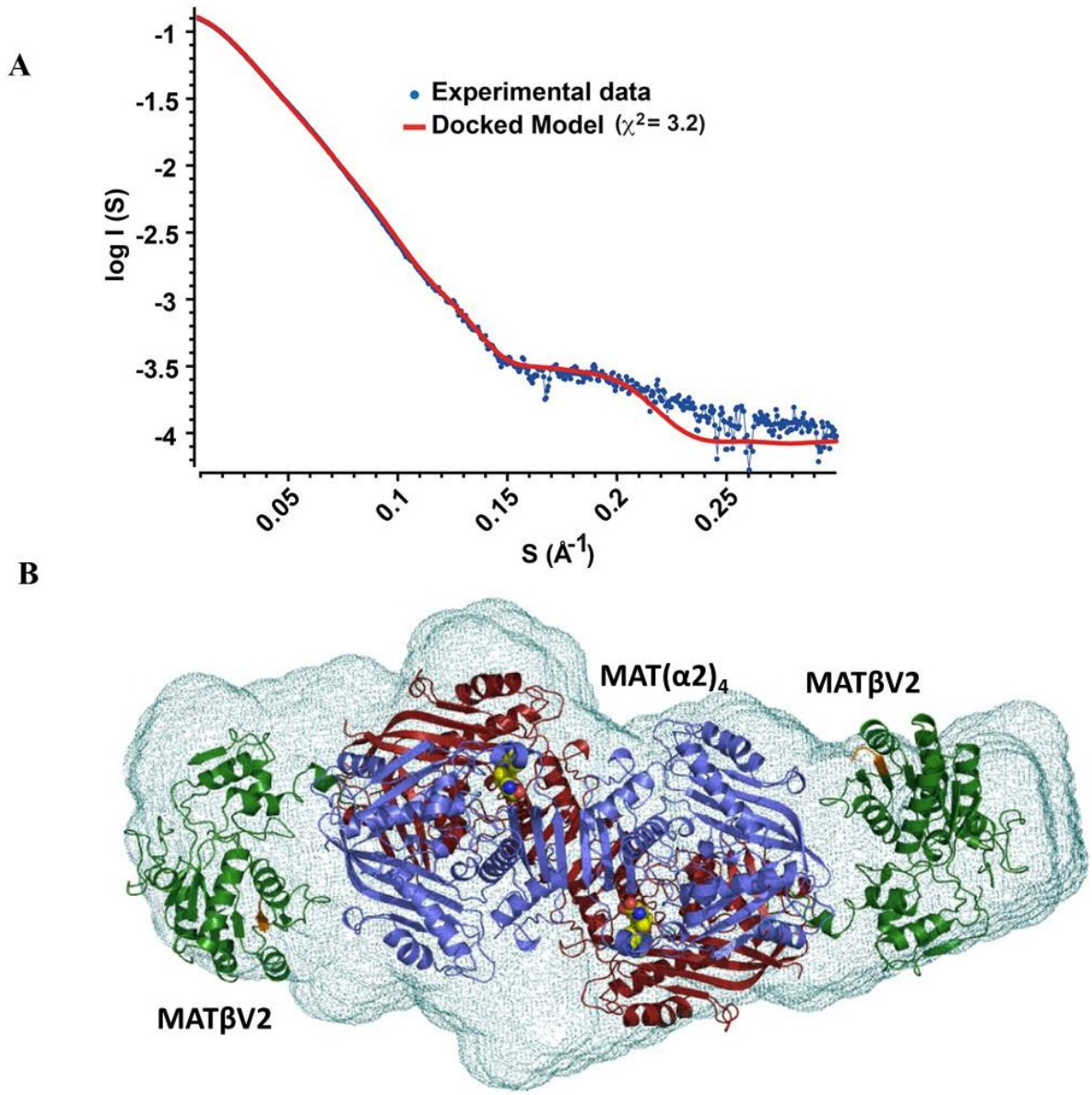

Figure 5.7. Conformation of oligomeric state of MAT(α2)$_4$(βV2)$_2$. A. Small-angle X-ray scattering of MAT(α2)$_4$(βV2)$_2$ complex (5.0 mg/mL), with the experimental spectrum (blue) shown with a simulated fit (red) obtained from the crystal structure (q_{max} = 0.30 Å^{-1}). The radius of gyration, R_g=50.1 ± 0.05 Å, was estimated from the low angle scattering region by a Guinier plot (15-39). The distance distribution function, P(r), with maximum linear dimension D_{max} = 187 Å. **B**. The ab initio shape reconstruction of MAT(α2)$_4$(βV2)$_2$ by DAMMIN using P1 symmetry, showing a good agreement between the predicted molecular shape (light green mesh) and the crystal structure (R_g=50.7) (cartoon), after alignment by SUPCOMB (NSD=0.93). The MAT(α2)$_4$(βV2)$_2$ consists of a tetramer of MATα2 proteins displayed in slate and red flanked by two MATβV2 proteins shown in orange. The figure was generated using Pymol (Schrödinger, LLC 2010).

As the Chi squared value of the predicted scattering of the MAT$(\alpha 2)_4(\beta V2)_2$ structure versus the 1D scattering data that was collected was considered to be slightly high, which was attributed to the flexibility of the MATβV2 N-terminus, a mutant of MATβV1 was created which lacked the first 16 residues (MATβΔ16). This mutant was used to assemble the MAT$(\alpha 2)_4(\beta \Delta 16)_2$ complex and then used to perform SAXS experiments. Comparing the experimental data to the docked model (4NDN) gave a closer likeness, given by a lower X^2 value (1.7), than the wild type MAT$(\alpha 2)_4(\beta V2)_2$ complex (Fig 5.8A). Maximum particle size (D_{max}) of MAT$(\alpha 2)_4(\beta \Delta 16)_2$ (18.7 nm) is simialar to the D_{max} calculated from the crystallographic model (19.6 nm) (Fig 5.8B). The solution scattering of the MAT$(\alpha 2)_4(\beta \Delta 16)_2$ complex, again suggests that it has the same composition and overall conformation as observed in the MAT$(\alpha 2)_4(\beta V2)_2$ crystal structure (Fig 5.8C).

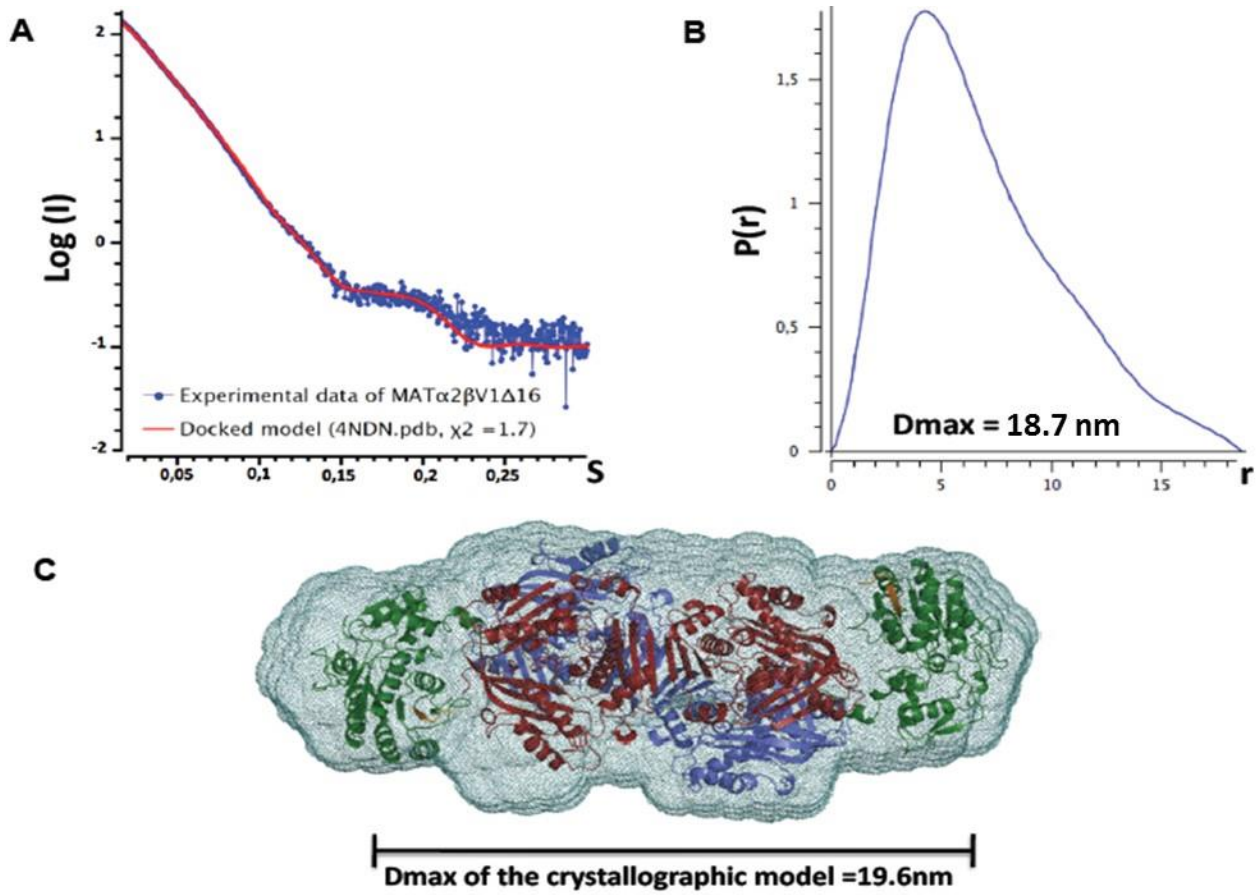

Figure 5.8 Solution scattering of MAT(α2)$_4$(βV1Δ16)$_2$. A. Small-angle X-ray scattering of MAT(α2)$_4$(βVΔ16)$_2$ complex (5.0 mg/mL), with experimental spectrum (blue) is shown with a simulated fit (red) obtained from the crystal structure (q_{max} = 0.30 Å^{-1}). The radius of gyration, R_g=53.34 ± 0.15 Å, was estimated from the low angle scattering region by a Guinier plot (25-35). **B**. The distance distribution function, P(r), with maximum linear dimension D_{max} = 18.70 nm. **C**. The *ab initio* shape reconstruction of MAT(α2)$_4$(βV1Δ16)$_2$ by DAMMIN using P1 symmetry, showing a good agreement between the predicted molecular shape (light green mesh) and the crystal structure (R_g=50.70) (cartoon), after alignment by SUPCOMB (NSD=0.93). The MAT(α2)$_4$(βV2)$_2$ consists of a tetramer of MATα2 proteins displayed in slate and red flanked by two MATβV2 proteins shown in orange. The figure was generated using Pymol (Schrödinger, LLC 2010).

Table 5.4. Summary of small angle X-ray scattering (SAXS) processing of $MAT(\alpha 2)_4(\beta V2)_2$ and $MAT(\alpha 2)_4(\beta V1\Delta 16)_2$.

SAXS Process	**$MAT(\alpha 2)_4(\beta V2)_2$**	**$MAT(\alpha 2)_4(\beta V1\Delta 16)_2$**
Guinier		
q range ($Å^{-1}$)	0-0.3	0-0.3
Guinier range	15-39	24-35
Rg (Å)	50.1 ± 0.05	53.34 ± 0.15
P(r)		
Maximum particle distance (D_{max}) (nm)	18.70	18.67
I(0)	$144.2 \pm 8.424x10^{-3}$	$175.83 \pm 8.79x10^{-2}$
Porod volume (nm^3)	310.51	341.65
Dammin		
Number of runs	10	10
Maxium q ($Å^{-1}$)	0.3	0.3
Dammin refinement		
Chi squared (X^2)	1.45	1.5
Crysol (4NDN.pdb)		
Chi squared (X^2)	3.208	1.720
Supcomb		
Normalized spatial discrepancy (NSD)	0.931	0.922

5.1.7 Stabilisation of MATβ binding with the MAT(α2)$_2$ cavity.

One of the key findings from determining the structure of the MAT(α2)$_4$(βV2)$_2$ complex is how MATβ interacts with MATα2. The structure reveals that MATβV2 interacts with MATα2 through the insertion of the C-terminal tail of the beta subunit into a cavity created at the interface of the MATα2 dimer (Fig 5.9A). Residues Lys315 to His323 of MATβV2 establish extensive hydrophobic and polar interactions with side chains of both MATα2 monomers. A list of interactions can be found in Table 5.5 highlighting all the hydrogen bonding between MATα2, MATβ and waters. In the complex, the usually disordered tail of the MATβV2 C-terminal folds into a helical structure, and within MAT(α2)$_2$ binding cavity the interaction generates a dilation of the cleft without any change in the orientation of the side chains consistent with a 'lock & key' mechanism. Comparing MATβ crystallised in the absence of MATα2 (2YDY) with MATβ from the MAT(α2)$_4$(βV2)$_2$ complex shows that only a small structural rearrangement in the C-terminus occurs to facilitate binding and that the majority of the MATβ structure remains unchanged (Fig 5.9B).

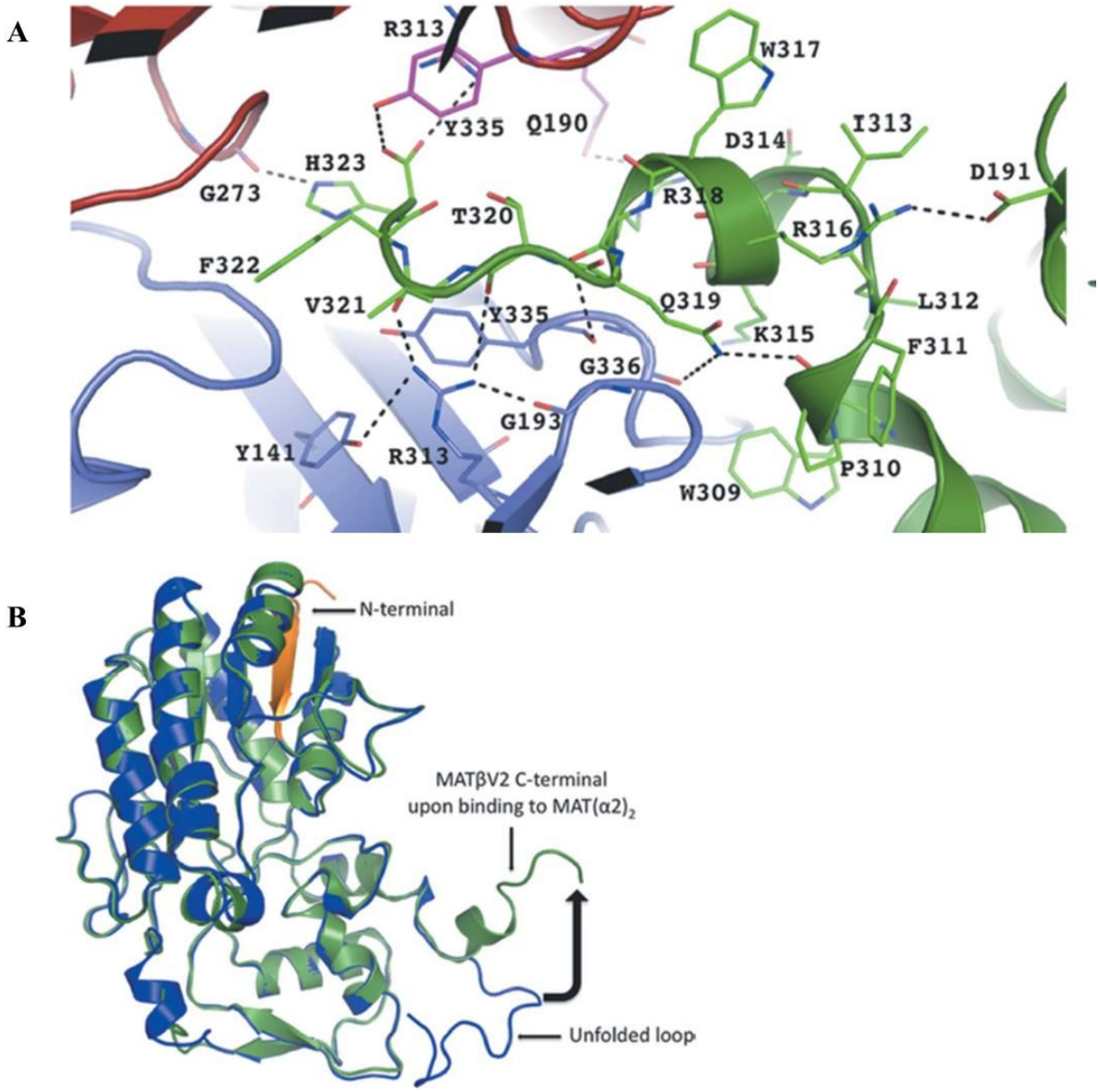

Figure 5.9 MATβV2 C-terminal interactions with MATα2. A. Each MATβV2 subunit interacts with the core by inserting its C-terminus (green) in a tunnel created at the interface between two MATα2 subunits (slate, red). The residues involved in the interaction are in stick representation (hydrogen bonds in black dotted lines). **B**. Superposition of MATβ (PDB: 2YDY) in blue with MATβV2 from the complex with MATα2 (green). The black arrow indicates the conformational change of the unfolded C-terminal loop.

Table 5.5. List of hydrogen bonding stabilising the interaction of MATβ with MATα2.

MATβV2 Residue	Atom	Interacting partner	Interacting Residue	Interacting Atom	Distance (Å)
Lys315	O	MATβ	Trp317	N	3.26
	O	MATβ	Arg318	N	2.88
Arg316	N	MATβ	Leu312	O	2.93
	N	MATβ	Ile313	O	3.26
	O	MATβ	Arg318	N	3.25
	NH_1	MATβ	Asp191	OD_1	2.65
	NH_1	MATβ	Phe311	O	2.87
	NH_2	MATβ	Asp191	OD_1	3.14
Tyr317	N	MATβ	Ile313	O	3.14
Arg318	O	Water			3.01
	O	MATβ	Thr320	N	2.84
	OE_1	Water			2.76
	NE_2	MATα2	Arg192	O	3.07
	NE_2	MATα2	Gly336	O	2.99
	NE_2	MATβ	Pro310	O	2.99
Thr320	OG_1	MATβ	Phe322	O	2.68
	O	MATα2	Arg313	NH_2	2.77
Val321	O	MATα2	Arg313	NH_1	3.00
	O	MATα2	Arg313	NH_2	3.20
Phe322	N	Water			2.96
	O	Water			3.08
His323	N	Water			3.21
	ND_1	Water			3.07
	NE_2	MATα2	Gly273	O	2.80
	O	MATα2	Arg313	NH_2	2.35
	O	Water			2.84
	OXT	MATα2	Arg313	NH_2	3.27
	OXT	MATα2	Tyr335	OH	2.99

5.1.8 Validation of C-terminal interaction site.

To establish that the C-terminal tail is indeed the key region of the interaction a truncated mutants lacking the last 15 residues at the C-terminus in both variants of MATβ, MATβV1 and MATβV2 were generated. These deletions precluded the assembly of the complex in solution (Fig 5.10A). Also, these deletions maintained the secondary structure as in the wild type MATβ isoforms, as confirmed by Circular Dichroism (CD) spectra (Fig 5.10B). The circular dichroism spectra of MATβV1ΔW320 is different from the other circular dichroism spectrums suggesting, although the MATβV1ΔW320 is folded, that it is different from the wild type protein. The interaction was also confirmed with the use of ITC, which showed that wild type MATβ isoforms interact with MATα2, but the mutant isoforms, MATβV1ΔW320 and MATβV2ΔW309 produced no heat release when titrated against MATα2 and, therefore, did not interact (Fig 5.11, Table 5.6). Taken together, these data show that MATβV1 and MATβV2 interact with MAT$(\alpha 2)_2$ through the insertion of the C-terminal tail of the beta subunit. The ITC also supported the ratio of MATα2:MATβ being 2:1 which is consistent with the crystal structure, gel filtration and SAXS experiments.

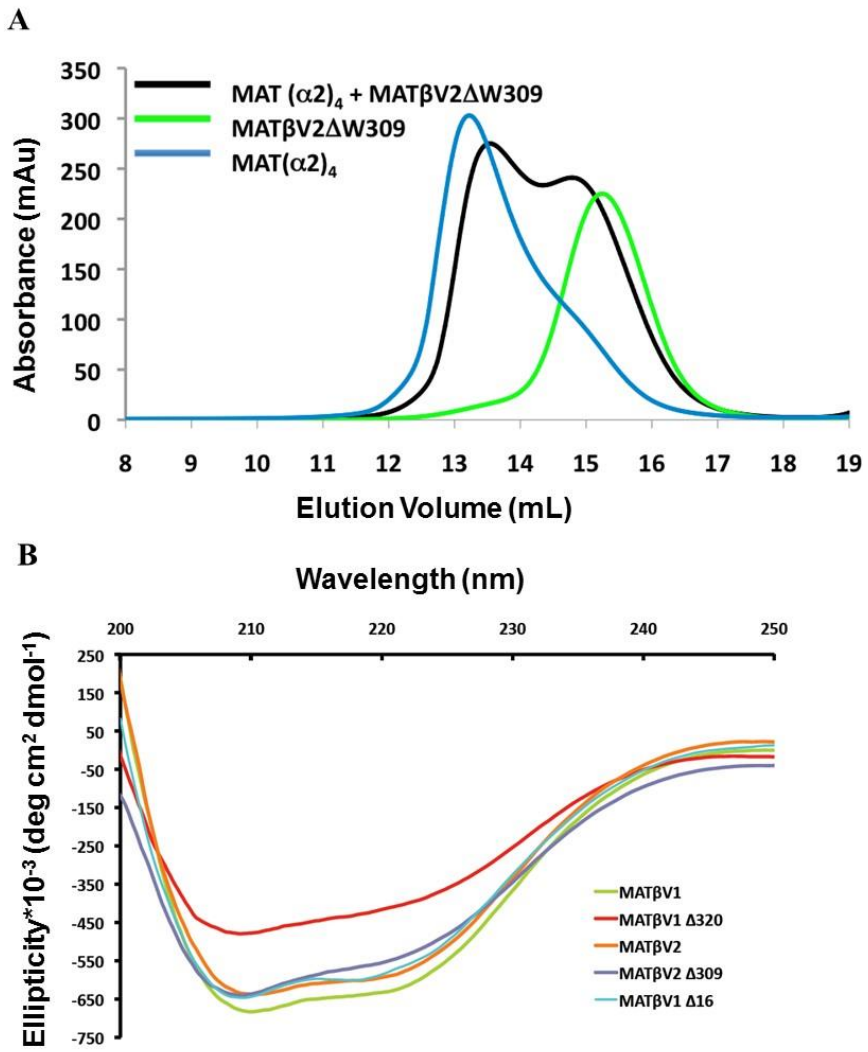

Figure 5.10 Conformation of MATβV2 C-terminal interaction with MATα2. A. Gel filtration profiles for MATα2+MATβV2ΔW309 (black), MATβV2ΔW309 (light green) and MAT$(\alpha2)_4$ in blue. 100 μM of MATα2 and 50 μM of MATβV2Δ309 were used in complex formation, whilst 100 μM of MATα2 and 50 μM of MATβV2 were used as controls. Absorbance at 280 nm was used to measure for the presence of proteins. For complex formation MATα2 was incubated with both MATβ variants prior to being loaded on a Superdex 200 10/300 column (section 4.2.13). **B**. Circular dichroism spectra of MATβV1 (green), MATβV2 (orange), MATβV1Δ320 (red), MATβV2Δ309 (violet) and MATβV1Δ16 (blue). Proteins were diluted to 1-5 μM in 0.1 M phosphate buffer pH 7.0**.** Comparison between the circular dichroism spectra of the wild type and the mutants shows that the secondary structure is conserved after the deletion of the C-terminus (MATβV1Δ320 and MATβV2Δ309) or the N-terminus (MATβV1Δ16). Data is displayed as ellipticity*10-3 (section 4.2.20).

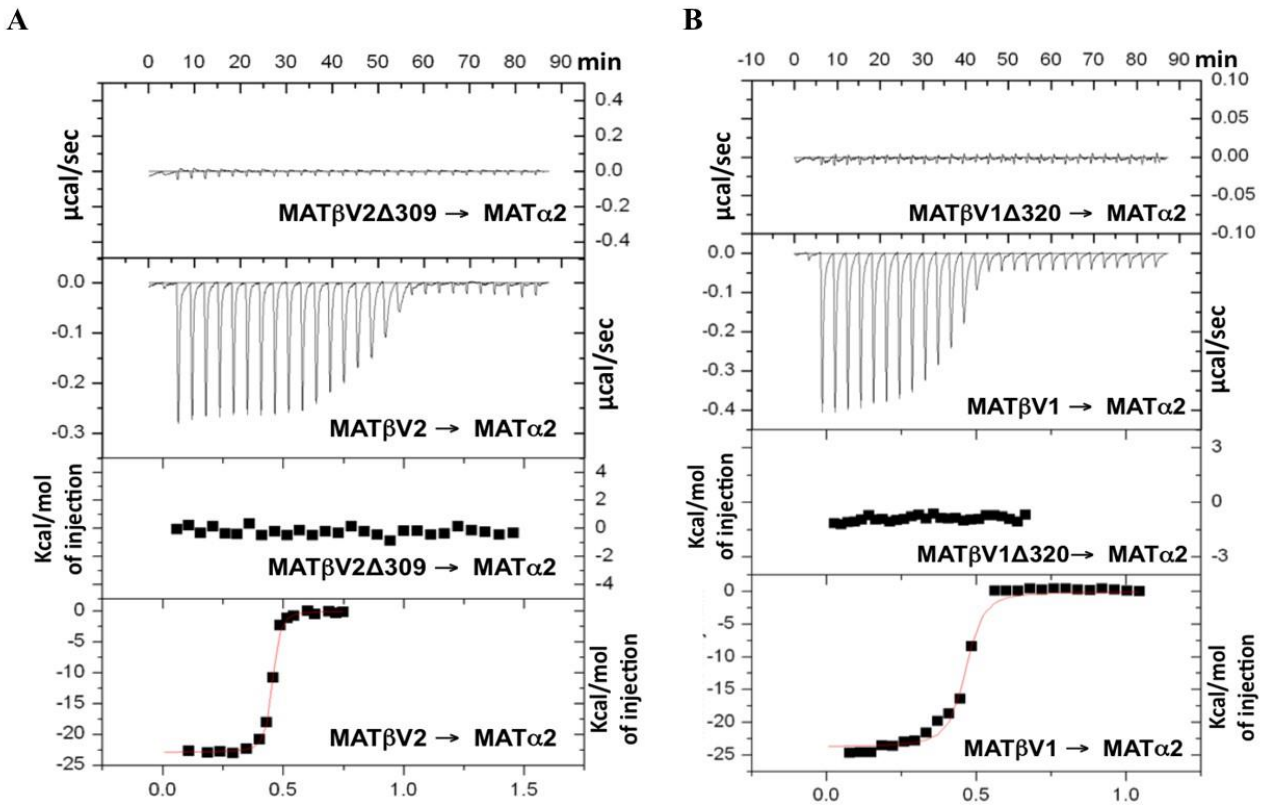

Figure 5.11. Interaction MATα2 and MATβ. A. Isothermal titration calorimetry of MATα2 with MATβV2 showing the interaction of these two proteins. The top graphs represent the differential heat released during the titration of MATβV2ΔW309 or MATβV2 with MATα2. The bottom graphs represent the fitted binding isotherms. **B**. As in **A** Isothermal titration calorimetry of MATα2 with MATβV1ΔW320 or MATβV1. Proteins were buffer exchanged into 200 mM NaCl, 20 mM HEPES pH 7.5. The concentration of MAT proteins were: MATα2=10.0 µM, MATβV1=51.2 µM, MATβV2= 36.5 µM, MATβV1Δ320 and MATβV2Δ309= 100 µM.

Table 5.6. Thermodynamic parameters of MATαβ complex formation. K_a is calculated from as $1/K_d$ and is the binding constant (M^{-1}). The K_d represents the affinity of the reaction (M). ΔH binding enthalpy of the interaction in kcal/mol. N is the stoichiometry indicating the ratio of the injected macromolecule-cell macromolecule binding.

	MATβV2 → MATα2	MATβV1 → MATα2
[]	36.5μM → 10μM	51.2μM → 10μM
$K_a(M^{-1})$	$(1.14 \pm 0.17) \times 10^8$	$(2.90 \pm 0.82) \times 10^7$
$K_d(M)$	$(8.77 \pm 1.34) \times 10^{-9}$	$(3.44 \pm 0.98) \times 10^{-8}$
ΔH kcal/mol	$-2.28 \times 10^4 \pm 188.4$	$-2.37 \times 10^4 \pm 367.6$
N	0.44 ± 0.001	0.44 ± 0.004

5.1.9 MATβ binding motif in MATαβ complexes.

To narrow down the interaction site further a more extensive mutational analysis of the MATβ C-terminus was conducted. Several different MATβ constructs were made by introducing stop codon mutations in different positions in the C-terminus of MATβ. Six mutants were produced and purified (section 4.2.6, 4.2.11. & table 4.3): MATβV1ΔR329, MATβV1ΔT331, MATβV1ΔH334, MATβV2ΔR318, MATβV2ΔT320, and MATβV2ΔH323. These were then incubated with MATα2 and passed through gel filtration to evaluate formation of the complexes. The gel filtration experiments showed that the minimum motif required for the formation of $MAT(\alpha2)_4(\beta V2)_2$ complex comprises of four residues at the end of the C-terminal of MATβV2 ($THR_{320}VAL_{321}PHE_{322}HIS_{323}$). This was also the same for MATα2βV1 complex which also had an absence of the complex peak when MATα2 was incubated with MATβV1T331stop ($THR_{331}VAL_{332}PHE_{333}HIS_{334}$) (Fig 5.13). The deletion of the terminal histidine residue was insufficient to preclude the formation of the complex with either MATβV1 or MATβV2.

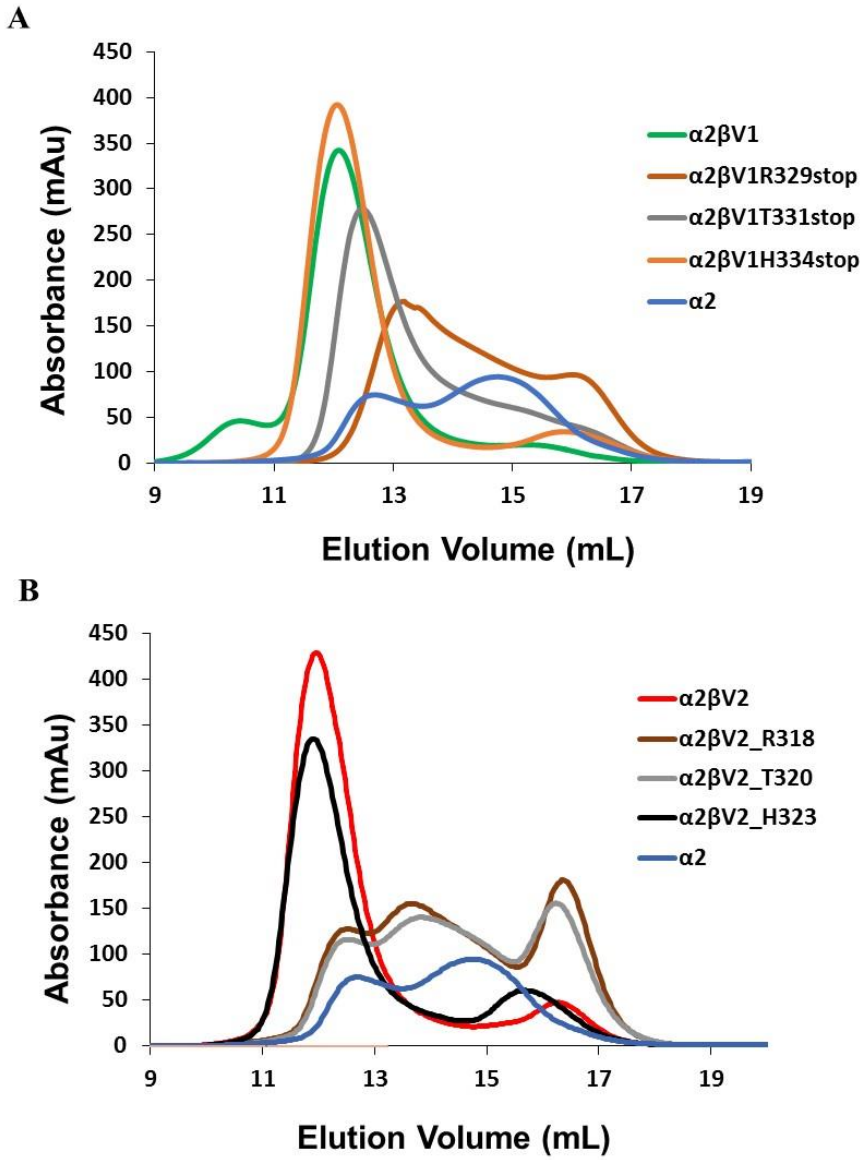

Figure 5.12. Mutational analysis into MATβ binding of MATα2. Gel filtration spectrums of various C-terminal beta mutants mixed with MATα2 for 1 hour prior to gel filtration on a Superdex 200 10/300 in 25 mM HEPES pH 7.5, 200 mM NaCl, 1 mM $MgCl_2$, 5 mM KCl, 1 mM TCEP (section 4.2.13). For complex formation, 100 μM of MATα proteins and 50 μM of MATβ were used. **A**. Mutant MATβV1 mutants showed that deleting the last four residues (MATβV1T331stop) is enough to prevent MATαβ complex formation. **B**. As in A, but using MATβ2 mutants. Absorbance was measured at 280 nm which was used as an indication of the presence of protein.

5.1.10 Rotation of MATβ bound to MATα2.

Interestingly, the tunnel created at the interface of the MATα2 dimer has a symmetry that allows two possible conformations of the C-terminal MATβV2. In both orientations the MATβV2 tail is stabilised by hydrogen bonds and hydrophobic interactions of the same residues of the opposing MATα2 dimer (Fig 5.13A). The interchangeability of these two orientations could allow certain rotational flexibility of the MATβV2 subunit enabling MATβ to interact with MATα2 on both sides of the dimer and both active sites. To note the N-terminal of MATβV2 face the direction of the MATα2 active site and are at a suitable distance that the disordered MATβV2 terminus could interact with the empty active site. At a first glance it did not seem that MATβ was close to the active site of MATα2, but upon examination of the interaction MATβV2, His323 interacts with Gly273 that is part of a loop that is directly connected to the MATα2 active site (Fig 5.13B).

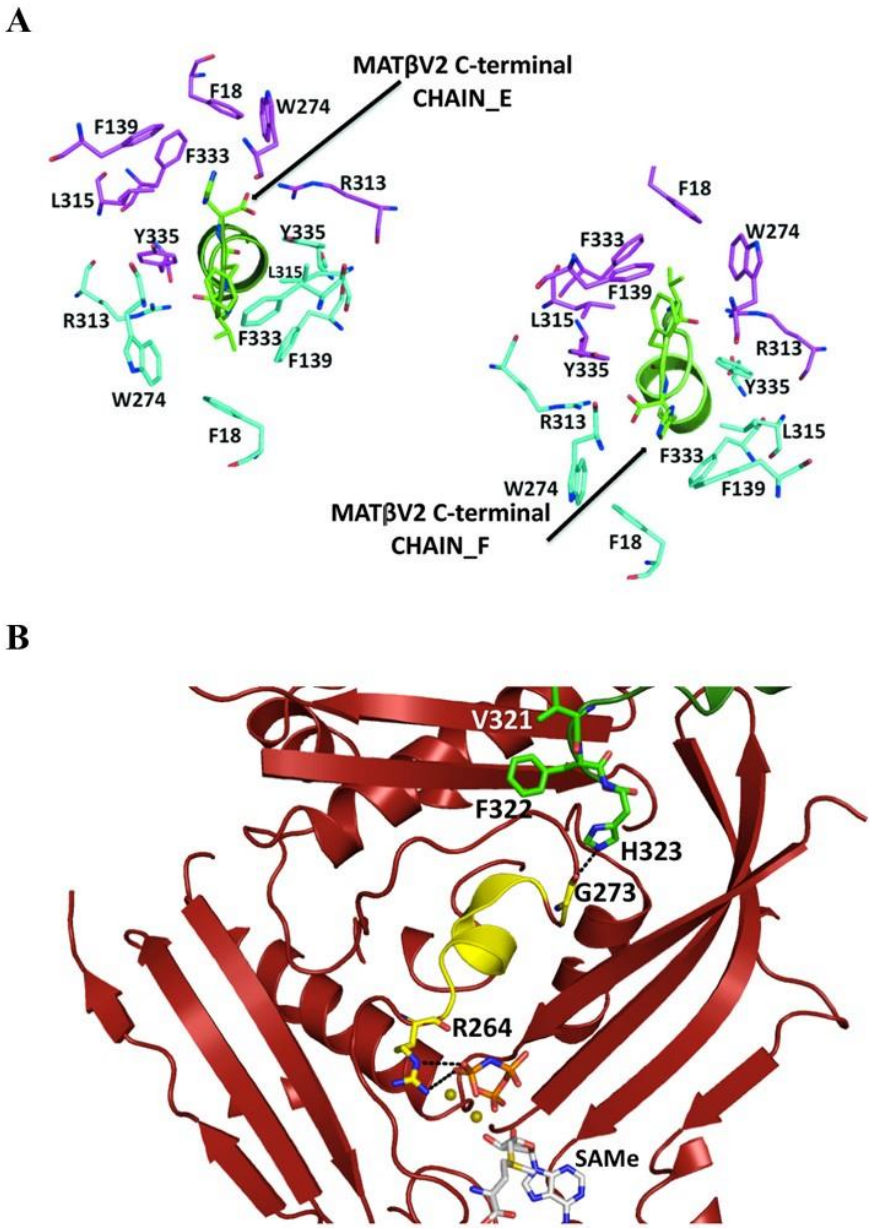

Figure 5.13. Close view of the tunnel created at the interface of the MATα2 dimer. **A**. Stick representation of the MATα2 residues at the dimer interface, side chains involved in the interaction with MATβV2 (chain_E in green) are coloured in cyan, the symmetry residues are shown in magenta. Side chains that are involved in the interaction with MATβV2 (chain_F in green) are coloured in magenta, the symmetry residues are shown in cyan. Note that between the conformation represented in A and B there is a twofold symmetry. **B**. Cartoon representation of the MATα2 monomer from the $MAT(\alpha2)_4(\beta V2)_2$ complex (4NDN), showing the loop that connects the active site with the buried tail of MATβV2 that is highlighted in yellow. The MATβ C-terminal is shown in green.

5.1.11 Enzymatic activity of MATαβ complexes.

The observation that the MATβV2 N-terminus is close to the active site of MATα2 raises the question of whether the N-terminus of MATβ could regulate the gating loop of MATα2 and affect the kinetic activity of MATα2. If this was the case, the presence of MATβ should affect their enzymatic activities and also differences should be observed between isoforms of MATβ due to the varying lengths of the N-terminus of MATα2βV1 and MATα2βV2. Thus, the activity of MATα2, $MAT(\alpha2)_4(\beta V1)_2$, $MAT(\alpha2)_4(\beta V2)_2$ and $MAT(\alpha2)_4(\beta V1\Delta16)_2$ complexes were compared. Ideal conditions for the assay were set up whereby the concentration of protein was tested to ensure that substrate would remain in excess by the end of the reaction. Different buffers were tested to see if the MATα2 could produce SAMe and it was found that in HEPES pH 7.5, PIPES pH 7.5 and PBS pH 7.5 SAMe could be formed (Fig 5.14A-C). When 10 mM DTT was added to the reaction buffer SAMe formation dropped. As the proteins were purified in HEPES pH 7.5 it was decided to use this buffer to perform the experiment (Fig 5.14D).

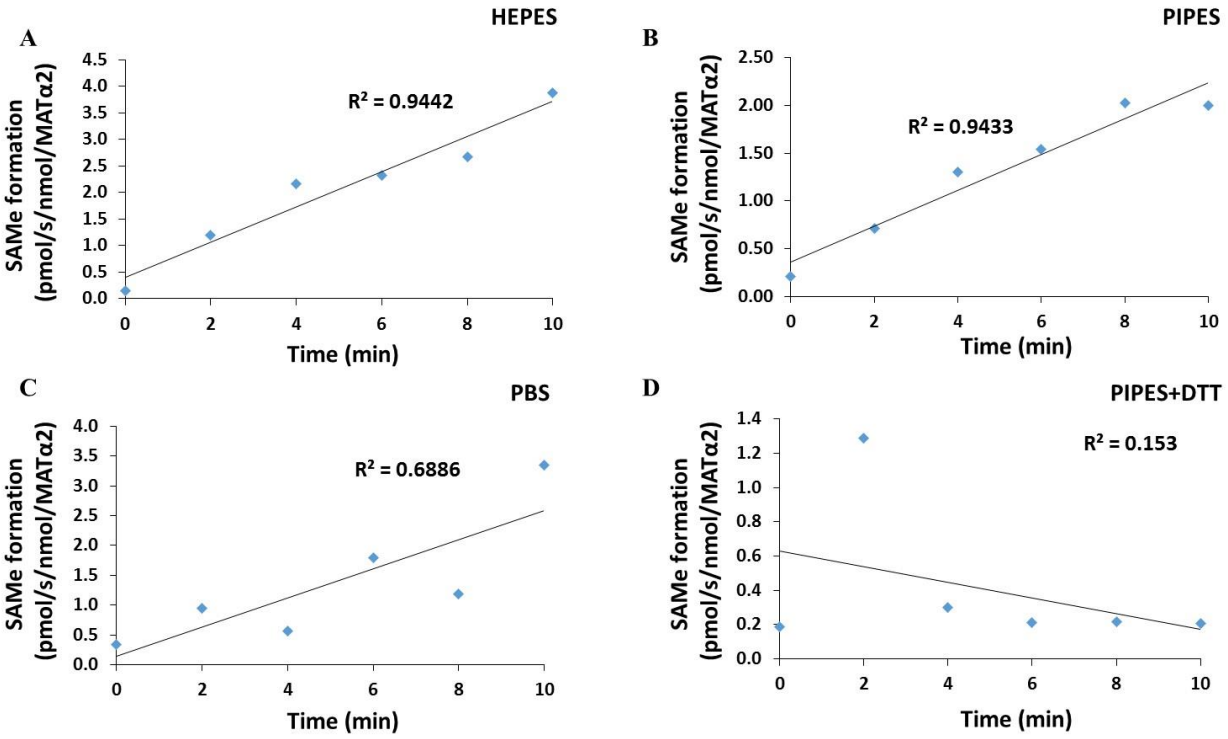

Figure 5.14. The effect of different reaction buffers on the formation of SAMe by MATα2. 100 nM of MATα2 was pre-incubated with 25 mM buffer (HEPES, PIPES, PIPES and DTT, or PBS), 5 mM KCl, and 5 mM $MgCl_2$.To start the reaction 200 µM methionine was added and reactions were quenched with 75 % (v/v) acetonitrile and 1.2 % (v/v) formic acid in water over a 10 min course. SAMe formation was quantified using UPLC-MS/MS (Van Liempd et al. 2013). **A**. HEPES pH 7,5. **B**. PIPES pH 7.5. **C**. PBS pH 7.5. **D**. PIPES pH 7.5 +DTT.

Notably, the presence of either MATβV1 or MATβV2 increased the k_{cat} of MAT$(\alpha2)_4$ 4 or 3 times respectively without altering the K_m for methionine. Additionally, the k_{cat} of MAT$(\alpha2)_4$ was 30 % higher in presence of MATβV1 than with MATβV2, thus emphasising that differences at the N-terminus affect the activity of MAT$(\alpha2)_4$. Furthermore, the utilisation of the MAT$(\alpha2)_4(\beta V1\Delta16)_2$ complex, in which MATβ has the shortest N-terminus, enabled it to be possible to see if the activity of this complex would return to the same level as in MAT$(\alpha2)_4$ alone. The MAT$(\alpha2)_4(\beta V1\Delta16)_2$ complex also has the lowest k_{cat} of the complexes which was still higher than the catalytic subunit alone (Fig 5.15, Table 5.7). Therefore, increasing the N-terminal length correlates with an increase in k_{cat} of the complex activity, confirming the role of N-terminal in the regulation of the activity. It also suggests that presence of MATβ has a positive effect on the activity of MAT$(\alpha2)_4$ and this must be due to the C-terminal interaction.

As a result of the cloning process several extra residues we added to the N-termini of MATβV1 and MATβV2. In order to rule out that these residues were causing the observed kinetic differences the MAT proteins, MATα2, MATβV1 and MATβV2 were sub cloned (section 4.2.7) into the PET_SUMO vector system which produced proteins with no extra residues. The same trend was observed with these proteins as with those produced in the other vector system when the kinetic tests were performed.

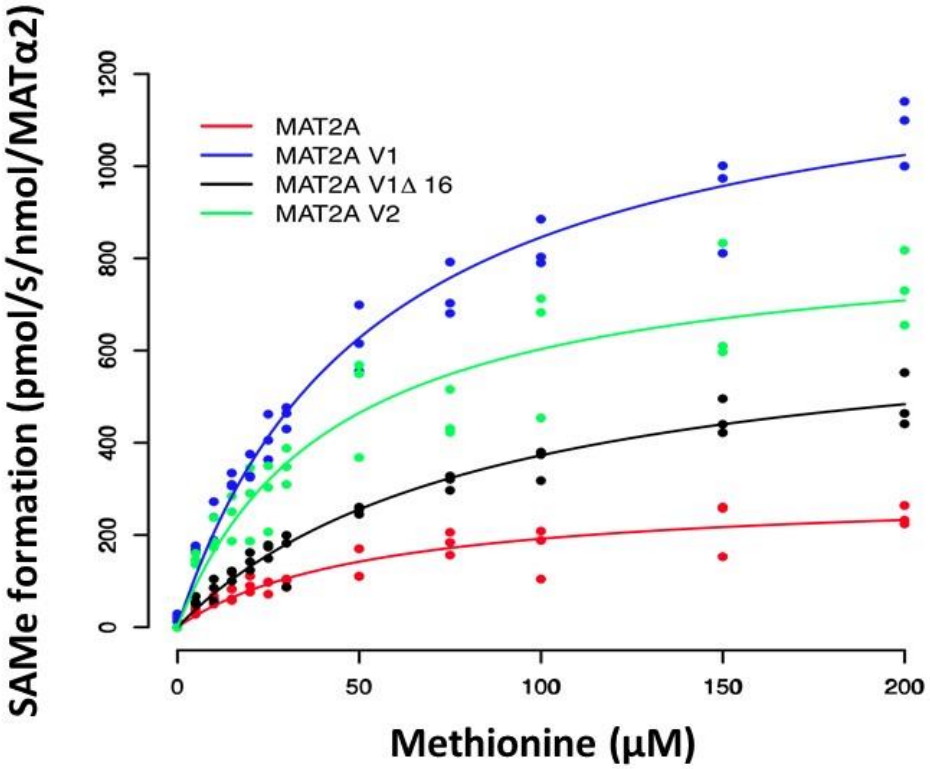

Figure 5.15 Kinetic reactions of MATα2 and it's complexes. ATP was added to $MAT(\alpha2)_4$, $MAT(\alpha2)_4(\beta V1\Delta16)_2$, $MAT(\alpha2)_4(\beta V2)_2$ and $MAT(\alpha2)_4(\beta V1)_2$ pre-incubated with 5 μM to 200 μM of methionine then SAMe formation was quantified using UPLC-MS/MS (section 4.2.20). It shows that the k_{cat} of MATα2β complexes are higher depending on the variant of MATβ. Triplicate data points were collected and are displayed as circles. $MAT(\alpha2)_4$ as red circles, $MAT(\alpha2)_4(\beta V1\Delta16)_2$ as black circles $MAT(\alpha2)_4(\beta V2)_2$ as green circles and $MAT(\alpha2)_4(\beta V1)_2$ as blue circles. SAMe formation is displayed as pmol/s/nmol of protein.

Table 5.7. Kinetic properties of MATα2β complexes as determined by kinetic assay (section 4.2.18 & fig 5.15). k_{cat} is the turnover number and K_m is defined as substrate concentration at half the maximum velocity.

	$MAT(\alpha2)_4$	$MAT(\alpha2)_4(\beta V1\Delta16)_2$	$MAT(\alpha2)_4(\beta V2)_2$	$MAT(\alpha2)_4(\beta V1)_2$
k_{cat}	295. ± 23.7	688. ± 38.4	859 ± 57.1	1300 ± 47.8
K_m	53.8 ± 10.5	84.8 ± 9.93	42.5 ± 7.38	53.5 ± 4.79

5.1.12 Investigation into the effect that NADP has on MAT proteins.

Previously, it has been shown that NADP binds tightly to a conserved glycine-rich GXXGXXG motif ($G_{24}ATG_{27}LLG_{30}$) at the N-terminal domain of MATβ (Shafqat et al. 2013). Mutants lacking residues involved in NADP binding of the β subunit have been shown to be able to form a complex with MATα2 supporting the idea that NADP is not needed to form the complex (González et al. 2012). In our case, co-crystallisation and soaking experiments with an excess of NADP yielded crystals of the $MAT(\alpha2)_4(\beta V2)_2$ complex without NADP. Also, incubating the proteins with NADP prior to gel filtration did not yield a shift in the elution profile or produced MATαβ crystals that contained NADP (Fig 5.16A). No observable effect was seen on Native PAGE with bands for either MAT subunit or $MAT(\alpha2)_4(\beta V2)_2$ complex in the presence or absence of NADP (Fig 5.16B). This observation prompted us to evaluate whether NADP could interact in solution with the preformed complex. Using the kinetic assay that was set up the effect of NADP was examined on the activity of $MAT(\alpha2)_4(\beta V1)_2$ and showed that the k_{cat} in the presence and absence of NADP was not significantly different (Fig 5.16C, Table 5.8).

Table 5.8. Kinetic parameters of $MAT(\alpha2)_4(\beta V1)_2$ ± NADP as determined by kinetic assay (section 4.2.18 & fig 5.16). k_{cat} is the turnover number and K_m is defined as substrate concentration at half the maximum velocity.

	$MAT(\alpha2)_4(\beta V1)_2$	$MAT(\alpha2)_4(\beta V1)_2$+NADP
k_{cat}	2190 ± 145	1770±150
K_m	60.2±9.10	50.3± 10.6

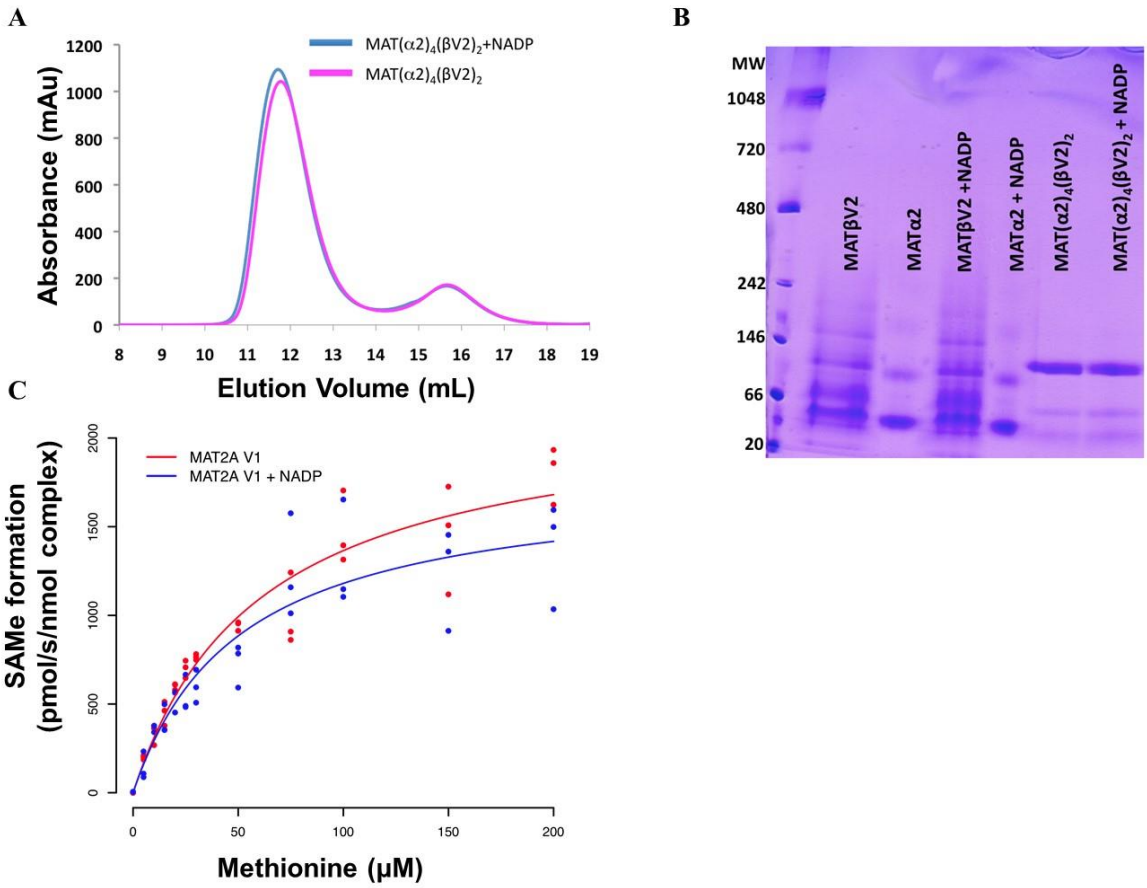

Figure 5.16. The effect of NADP on MATα2. A. Gel filtration of $MAT(\alpha2)_4(\beta V2)_2$ ((MATα2 200 μM and MATβV2 100 μM), complex with (blue) and without (magenta) NADP incubation (250 μM) and Native-PAGE of MAT subunits and $MAT(\alpha2)_4(\beta V2)_2$ complex with and without NADP (section 4.2.13). **B**. Native PAGE of MATα2 MATβV2 and $MAT(\alpha2)_4(\beta V2)_2$ complex in the presence and absence of 250 μM NADP. **C**. ATP was added to $MAT(\alpha2)_4(\beta V1)_2$ ± NADP pre-incubated with 5 μM to 200 μM of methionine before SAMe formation was quantified using UPLC-MS/MS (Van Liempd et al. 2013) (section 4.2.18). It shows that the k_{cat} of $MAT(\alpha2)_4(\beta V1)_2$ in the presence and absence of NADP are not significantly different.

ITC experiments confirmed that though NADP binds to both MATβ isoforms (Fig 5.17A-B) no interaction could be observed with the MAT$(\alpha 2)_4(\beta V2)_2$ complex (Fig 5.17C, Table 5.9). To note, NADP binds MATβ isoforms; MATβV1 and MATβV2 with very similar affinity and in a ratio of one NADP molecule to one MATβ. The overall structure of MATβV2 within the MAT$(\alpha 2)_4(\beta V2)_2$ complex has no significant structural changes as compared with the NADP-bound MATβ structure (RSMD=0.42 Å). The position of residues that are involved in NADP binding, such as Phe49, Arg51, Ser125 and Trp148, do not differ in the presence of MATα2 (Fig 5.18A). Modelling NADP in the active site of MATβV2 within the MAT$(\alpha 2)_4(\beta V2)_2$ complex shows that NADP could interact with MATβV2 (Fig 5.18B). Furthermore, the observation that MAT$(\alpha 2)_4$ does not block the binding pocket, suggests that the NADP could bind to MAT$(\alpha 2)_4(\beta V2)_2$ complex (Fig 5.18C).

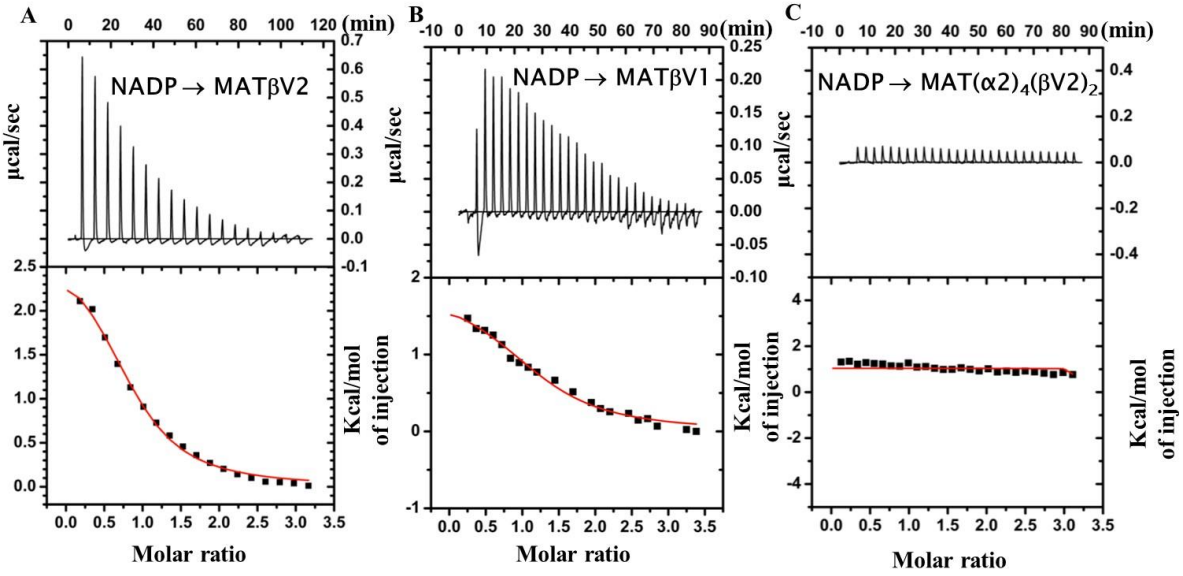

Figure 5.17. Interaction of MATα2β complexes with NADP. A,B. Isothermal titration calorimetry of MATβV2, MATβV1 and $MAT(\alpha2)_4(\beta V2)_2$ complex with NADP. The top graphs represent the differential heat released during the titration of NADP with the different MAT subunits, the bottom graphs represent the fitted binding isotherms. Proteins were buffer exchanged into 200 mM NaCl, 20 mM HEPES pH 7.5. The concentration of MAT proteins were: MATβV1=29.8 µM, MATβV2=30.0 µM, $MAT(\alpha2)_4(\beta V2)_2$=10.2 µM. 500 µM NADP was titrated against the MATβ proteins whilst 155 µM NADP was titrated against the $MAT(\alpha2)_4(\beta V2)_2$ complex.

Table 5.9. Thermodynamic parameters of NADP binding to MATβ isoforms from ITC experiments. K_a is calculated by taking $1/K_d$ and is the binding constant (M^{-1}). The K_d represents the affinity of the reaction (M). ΔH binding enthalpy of the interaction. N is the stoichiometry indicating the ratio of the injected ligand-cell macromolecule binding.

	NADP → MATβV2	NADP → MATβV1	NADP →$MAT(\alpha2)_4(\beta V2)_2$
[]	500µM → 30.0µM	500µM → 29.8µM	155µM→ 10.2µM
$K_a(M^{-1})$	(1.12 ± 0.130) $X10^5$	(1.17 ± 0.220) $X10^5$	--
$K_d(M)$	(8.93 ± 1.07) X 10^{-6}	(8.55 ± 1.60) $X10^{-6}$	--
ΔH	2780 ±117	1870 ± 128	--
N	0.850 ± 0.0260	1.23 ± 0.0600	--

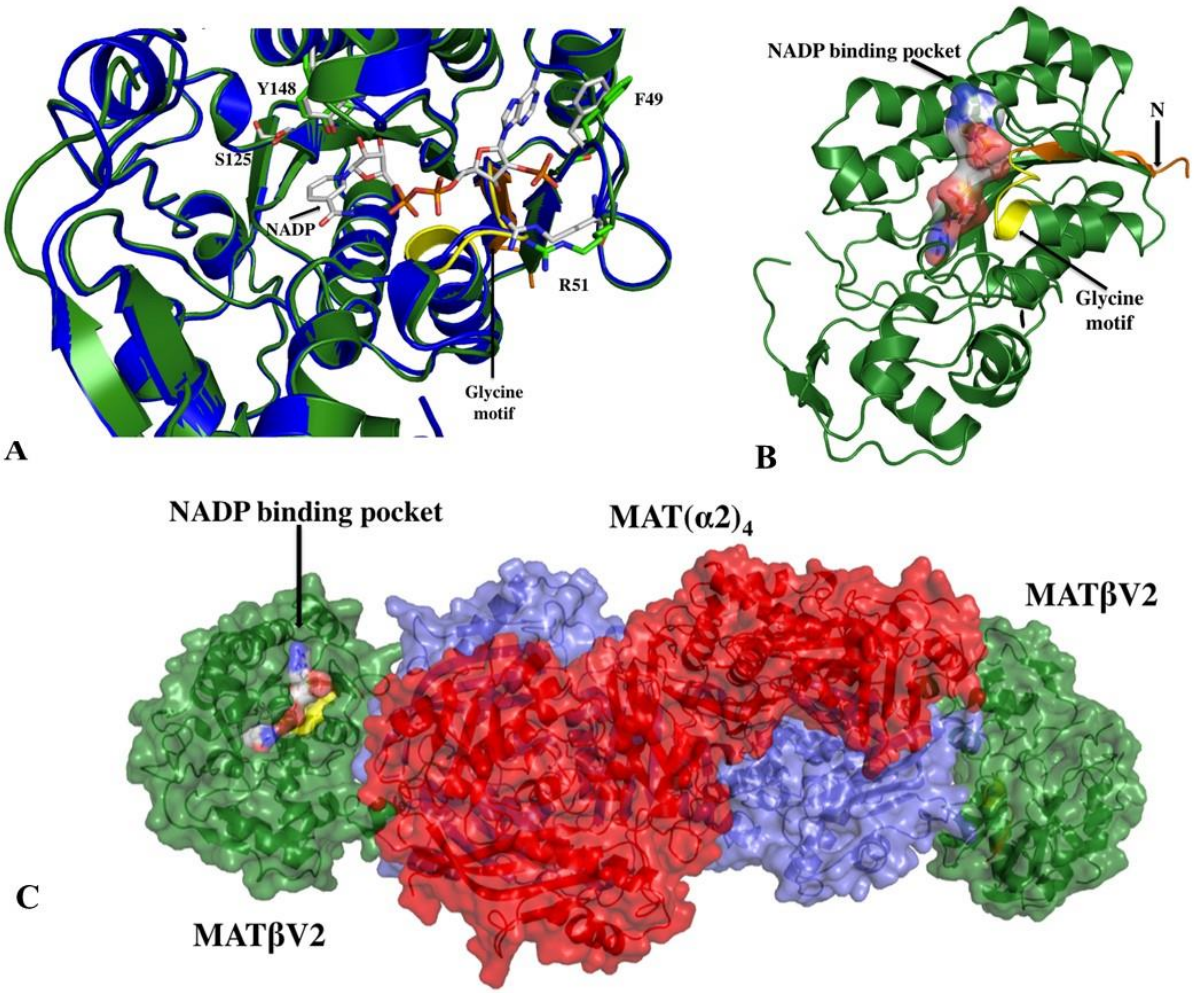

Figure 5.18. NADP binding site. A. Superposition of the NADP-bound MATβ structure (PDB: 2YDX in blue) over MATβV2 as it is in the $MAT(\alpha2)_4(\beta V2)_2$ complex (in green). Residues directly involved in the interaction with NADP are labelled and shown as sticks. **B**. Surface representation of the model of a NADP molecule within the binding pocket of MATβV2 from the $MAT(\alpha2)_4(\beta V2)_2$ complex showing that interaction of MATβ with MATα2 does not prevent NADP from binding. **C.** The $MAT(\alpha2)_4(\beta V2)_2$ complex shown as the $MAT(\alpha2)_4$ (in red and slate) and MATβV2 in green which does not block the access of NADP to MATβV2 (in green).

5.1.13. The effect of resveratrol binding to the MATαβ complex.

Resveratol, a molecule found in the crystal structure of MATβ (2YDX), was able to interact with MATβ, as reported so by Shafqat *et al*, so we wanted to test whether this molecule had any influence of the MATαβ complex. When resveratrol was titrated against the $MAT(\alpha 2)_4(\beta V2)_2$ no interaction was seen (Fig 5.19A) and further to this gel filtration experiments revealed no shift in the elution volume of the protein with and without resveratrol, which supported that resveratrol does not interact with the MATαβ complexes (Fig 5.19B).

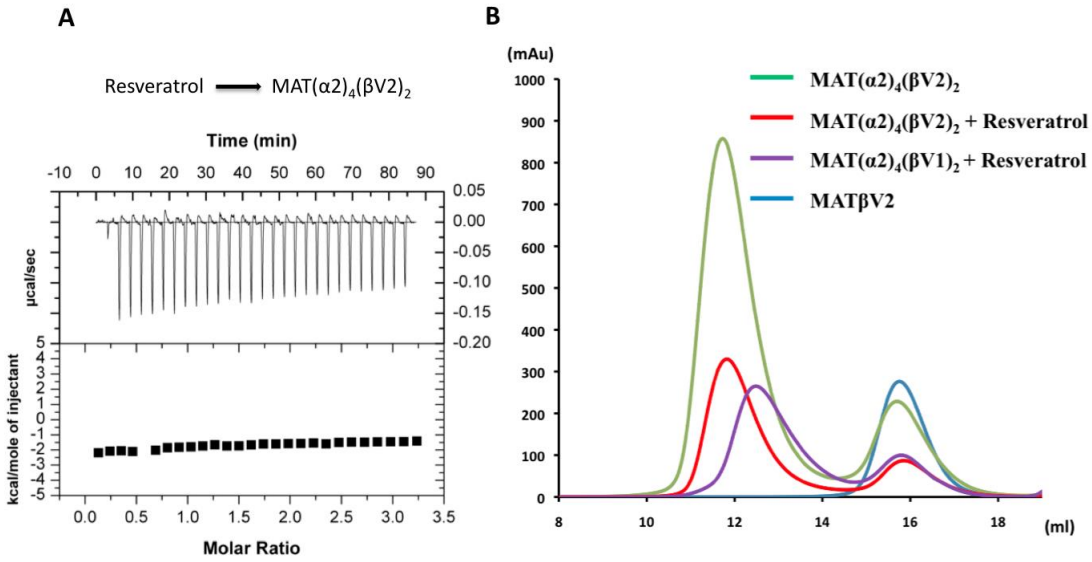

Figure 5.19. **Interaction of MATαβ complexes with resveratrol**. **A.** Isothermal titration calorimetry of MAT$(\alpha 2)_4(\beta V2)_2$ complex with resveratrol. The top graphs represent the differential heat released during the titration of resveratrol to MAT$(\alpha 2)_4(\beta V2)_2$ complex. 500 µM of resveratrol was titrated against 150 µM of the MAT$(\alpha 2)_4(\beta V2)_2$ complex. The bottom graphs represent the fitted binding isotherms. **B**. Chromatograms of the gel filtration performed after the incubation of MATβV1 (violet), MATβV2 (red) with resveratrol before addition of MATα2, the chromatograms shows that resveratrol does not influence MATα2β complexes formation as they eluted at the same volume of MAT$(\alpha 2)_4(\beta V2)_2$ complex without resveratrol (green).

5.2 Discussion

Initially when trying to form the MATαβ complexes a ratio of 1:1 MATα2:MATβ was used which resulted in the formation of the complexes, but also a second peak, corresponding to the elution profile of MATβ (~16 mL) was present. The literature at the time suggested that MATαβ complexes were a tetramer of two MATα2 to two MATβ [$MAT(\alpha2)_2(\beta)_2$] (Kotb & Kredich 1985). This prompted the revaluation of the oligomeric state of the MATαβ complexes, so using ITC it was observed that MATα2 interacts with MATβ in a 2:1 ratio (Fig 5.11), which was consistent with the model proposed by (González et al. 2012) ($MAT(\alpha2)_2(\beta)_1$) and what was seen using gel filtration upon changing the ratio of MATα2: MATβ to 2:1. The initial hit for $MAT(\alpha2)_4(\beta V2)_2$ did not allow for solving the structure, but did give insights into the oligomeric state as Matthews coeffcient analysis suggested a molecule of 258 kDa, which would be akin to size of a hexamer (Table 5.1). This predicition also had 50 % solvent content within the crystal, which is around the average for most protein crystals. Once the crystallographic strucuture was solved (2.35-3.3 Å) it revealed the complex to be a hexamer (Fig 5.4).

A pitfall of crystallography is that structure, resolved through this technique, may not represent the real biological interface and therefore not represent what is happening in solution. The contact between the $MAT(\alpha2)_2$ dimers to form $MAT(\alpha2)_4$ may be a result of crystal contacts that only exist due to the packing of the protein molecules in the crystal lattice (Carugo 1997). Small angle x-ray scattering was used to examine the oligomeric state of MATαβ in solution which supported a hexameric form (Fig 5.7, Table 5.4). Moreover, scattering data from $MAT(\alpha2)_4(\beta V1\Delta16)_2$ gave a much better fit to the crystallographic structure (4NDN) given that the disordered N-terminal were not seen in the crystal structure, but they would have contributed to the scattering of the wild type

complex (Fig 5.8, Table 5.4). Also, using the PISA server it suggested that a hexamer would be the most energetically favourable state (Table 5.2). Taken together it seems that the hexamer, although different from current literature was the correct state for MATαβ complexes. The simple schematic in figure 5.20 could be used to visualise the MATαβ complexes which are split into $MAT(\alpha2)_4(\beta V1)_2$ and $MAT(\alpha2)_4(\beta V2)_2$ depending on which MATβ isoform is involved.

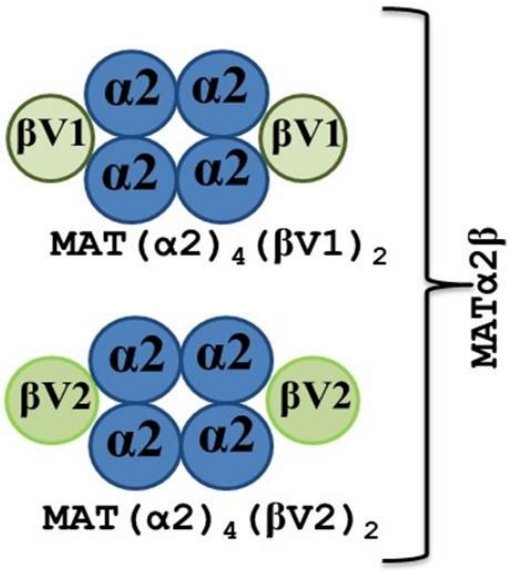

Figure 5.20. Schematic drawing of the MATαβ complexes. The MATαβ complexes are hetero-hexamers consisting of a tetrameric core of MATα2 flanked by two copies of the regulatory protein MATβ. The crystal structure of the MAT$(\alpha2)_4(\beta V2)_2$ has been solved through x-ray crystallography and solution scattering data support the idea that MATαβ complexes are hexameric**.**

Before solving the structure of $MAT(\alpha2)_4(\beta V2)_2$ there was little insight into the interaction site of MATβ with MATα2. Gonzalez *et al.* proposed, through ClusPro modelling, a trimer ($MAT(\alpha2)_2(\beta)_1$) whereby the MATβ subunit binds the central domain, close to the active site, through several different regions of MATβ incuding the NADP binding site and C-terminal tail. From this model the binding of MATβ would prevent oligomerisation of MATα2 into a tetramer. Our model revealed that MATβ interacts with MATα2 through the C-terminus which gains secondary structure and this interaction site is located in the interface of the MATα2 N-domain and C-domain (Fig 5.9), which is on the opposite side to the suggested trimeric model (González et al. 2012). Generation of the two stop mutants MATβV2ΔW309 and MATβV1ΔW320, through gel filtration (Fig 5.10) and ITC (Fig 5.11, Table 5.6), showed that the deletion of the C-terminal prevented the formation of the complex and therefore implies that this was the area of interaction. To note these stop mutants retained their secondary structure as shown through circular dichroism (Fig 5.10) and also the method of purification, though yielding less protein, was the same as in the wild types.

In order to probe the interaction site in further detail several mutants were produced with the hope of finding the minimum motif required to prevent the binding of MATβ. The deletion of the last four residues was enough to inhibit complex formation (Fig 5.12) and this motif could be the site of a small inhibitor to block the binding of MATβ *in vivo*. This motif interacts with the loop that recognises the tripolyphosphate of the ATP at the active site (Fig 5.13), suggesting a possible allosteric mechanism that could be responsible for the observed increase of the k_{cat} of $MAT(\alpha2)_4(\beta V1\Delta16)_2$ in comparison with the catalytic subunit alone (Fig 5.15, Table 5.7).

One of the surprising features of the MAT$(\alpha 2)_4(\beta V2)_2$ complex is that there is an asymmetry in the active sites, two which are in an apo state and two which are in a holo state (Fig 5.4). This is a new feature for MAT enzymes as all the structures of MAT enzymes deposited in the protein data bank have either all active site empty or all occupied (Komoto et al. 2004; Takusagawa et al. 1996; González et al. 2000; González et al. 2012; Shafqat et al. 2013). The presence of MATβ has a profound effect on MATα2 not only effecting the active site occupancy but also the k_{cat}. Kinetic experiments showed that the presence of MATβ increased the activity of the enzyme compared to MATα2 (Fig 5.16, Table 5.7), but from the crystal structure it still remains unclear on how this occurs. The position and orientation of MATβ indicates that the N-terminal, which is disordered, is in close proximity to the active site such that it could interact with it or influence it. It is doubtful that MATβ directly affects the enzymatic mechanism of the MATα2, as SAMe synthesis is a fundamental process preserved throughout the kingdoms. It could play a role in increasing the turnover and release of the product from the active site either through interacting with the gating loop or through MATβ rotation destabilising products within the active site causing a quicker release of products.

A comparison of the active site in two different states reveals that loops flanking the empty catalytic pockets are disordered, in particular the 'gating loop' (residues 113-131) that has been proposed to act as a dynamic lid controlling the access to the active site (Komoto et al. 2004). In the case of the catalytic subunit MATα1, S-nitrosylation of residue C121 in the 'gating loop' promotes the inactivation of the enzyme (Pérez-Mato et al. 1999) so interaction with the gating loop by MATβ could explain the reasons for higher SAMe production in MATαβ complexes.

Although MATβ contains a NADP binding domain no interaction between MATαβ complexes and NADP was observed by either gel filtration (Fig 5.16) or ITC experiments where NADP was titrated against the MATαβ complexes and this produced no heat reaction (Fig 5.17, Table 5.9). This shows that NADP is not needed as a co-factor for the formation of the MATαβ complexes nor is it needed for SAMe synthesis. This is in support of work done by Gonzalez *et al.* 2012, who mutated the NADP binding pocket, preventing NADP binding, and then used these mutant MATβs to form MATαβ complexes. We wanted to confirm that the recombinant MATβ was able to be bound by NADP, so, NADP was titrated against MATβ and interaction with both isoforms was observed with both isoforms showing similar affinity for NADP (Fig 5.17, Table 5.9), which is not surprising given the high sequence similarity (94 %) between these proteins and that the NADP binding pocket is conserved. It also suggests that the N-termini of MATβ, the only part which differs in sequence, do not aid in NADP binding as titration of NADP against the N-terminal mutant, MATβΔ16 showed interaction. If NADP did interact with the MATαβ complexes, then significantly different kinetic properties should be observed in the presence and absence of NADP which they do not (Fig 5.16, Table 5.8).

The reason why NADP does not bind the MATαβ complex remains unclear. From the crystallographic structure of $MAT(\alpha 2)_4(\beta V2)_2$ there is no reason why NADP should not bind, as the complex formation does not block the binding pocket (Fig 5.18). The dynamic movement of MATβ rotating around MATα2 could prevent binding of NADP, if this movement caused a structural change in MATβ by compression of the binding site inhibiting NADP binding. Furthermore, the observation that $MAT(\alpha 2)_4$ does not block the binding pocket, suggests that the NADP could bind to $MAT(\alpha 2)_4(\beta V2)_2$ complex in

the context of the interaction with other proteins such as HuR, GIT1, MEK, ERK, MafK (Xia et al. 2010; Peng et al. 2014; Katoh et al. 2011). In fact, there is evidence suggesting that MAT enzymes are involved in regulation of many pathways, some of which are chromatin-based and some may be independent of SAMe (Katoh et al. 2011).

Whereas resveratrol bound to MATβV2, as shown using differential scanning fluorimetry by Shafqat *et al* we observed no appreciable interaction between MATα2βV2 and resveratrol by ITC (Fig 5.19A). In addition, we observed that the presence of resveratrol had no effect in the complex formation, as observed by gel filtration experiments (Fig 5.19B). The interaction of resveratrol with MATβ might be relevant for functions other than the formation or activity of the MATα2β complexes. Further to this Yang *et al.* reported that resveratrol induced the expression of MATβV1 and MATβV2 *in vivo* by increasing transcription and stabilisation of both, mRNA and protein and proposed that resveratrol enhanced the interaction with other proteins that can bind MATβ such as SIRT1 and HuR (Yang et al. 2013).

Chapter VI. Structural insights into MATα and MATβ proteins.

6.1 Results

6.1.1 The effects of sodium chloride concentration on the oligomeric states of MAT proteins as determined by gel filtration

Human MATα1 has been known to exist in two oligomeric states, as a dimer or a tetramer, but little information was known what drives MATα1 proteins to change oligomeric state. Also, little was known regarding the oligomeric state of MATα2 so it was interesting to see whether MATα2 had mulitple oligomeric states like MATα1. The concentration of sodium chloride was varied to see if simply changing this in the buffer would have an effect on the oligomeric state of MAT proteins. Upon dialyzing MATα1 and MATα2 into different concentrations of NaCl and injecting into a gel filtration column both proteins exhibit two peaks the first of ~13 mL and the second of ~15.4 mL (Fig 6.1A, B). Buffer exchanging the proteins into different concentrations of sodium chloride had no effect on the equilibrium between the two peaks. The regulatory proteins MATβV1 and MATβV2 could only be eluted as single peaks (~16 mL) with molecular mass close to their predicted molecular mass (Fig 6.2A, Table 6.1). A calibration curve of known proteins (Fig 6.2B) was used to determine the mass of MAT proteins based on the elution volume (Table 6.1). For MATα1 and MATα2 the second peak (Fig 6.1A, B) corresponded to a molecular mass close to a monomeric species (~45 kDa) whilst the first peak correspond to a larger mass of ~124 kDa. This larger weight would suggest a trimeric state, but based on the crystallographic structure monomeric and trimeric states will not be stable in solution. As gel filtration is not the best technique for estimating the mass of non-globular proteins, the only clear information that can be inferred from MATα chromatograms is the presence of different oligomeric states. For MATβV1 and MATβV2 the elution volumes corresponded to a molecular weight that was around ~32 kDa, which is close to their estimated molecular weight of ~37 kDa.

Table 6.1. Molecular weight estimates of MAT proteins based on elution volume of gel filtration. Molecular weights were calculated using the elution volume to work out phase distribution coefficient (K_{av}) and then using this value in the equation K_{av}=-0.2760Log (Mw)+1.774. Proteins were injected into a Superdex 200 10/300 column with a buffer containing 25 mM HEPES pH 7.5, 200 mM NaCl and 1 mM TCEP.

Protein	Elution volume (mL)	Molecular weight (kDa) from elution volume	Estimated molecular weight (kDa) from sequence
MAT$(\alpha1)_2$	15.37	43.07	87.00
MAT$(\alpha1)_4$	13.36	123.2	172.0
MAT$(\alpha2)_2$	15.48	45.48	90.00
MAT$(\alpha2)_4$	13.34	124.4	180.0
MATβV1	16.06	32.79	37.50
MATβV2	16.09	32.31	36.40

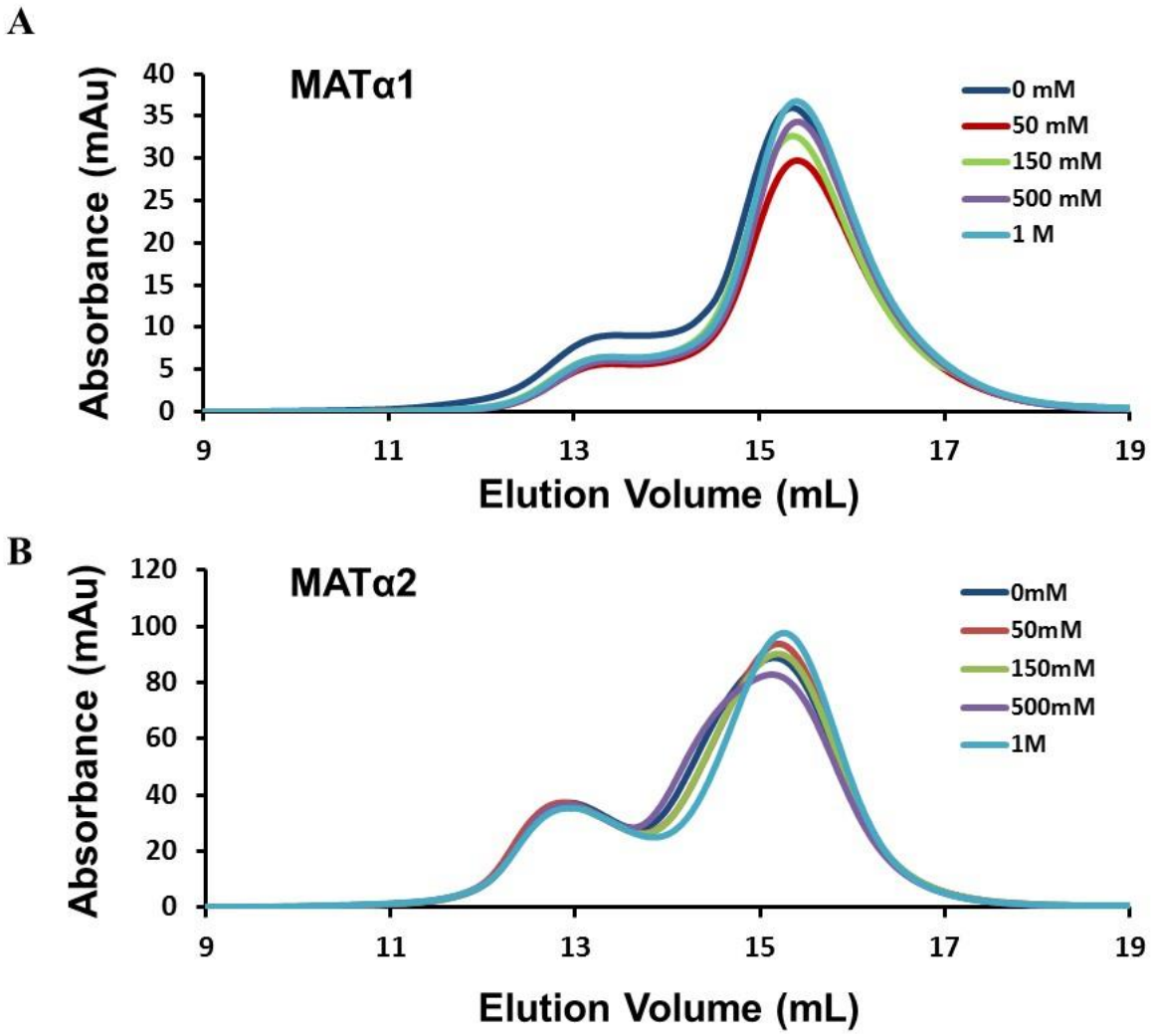

Figure 6.1 Sodium chloride concentration-dependence of MAT oligomeric state. **A**. MATα1 elution profiles, colour coded according to concentration of used NaCl. Absorbance was measured at 280 nm which was used as an indication of the presence of protein. Two peaks correspond to tetramer and dimer from left to right. Elution buffer 25 mM HEPES pH 7.5, 0-1 M NaCl and 1 mM TCEP **B**. As in A, but with MATα2. 100 µM of MATα protein were used for each NaCl concentration.

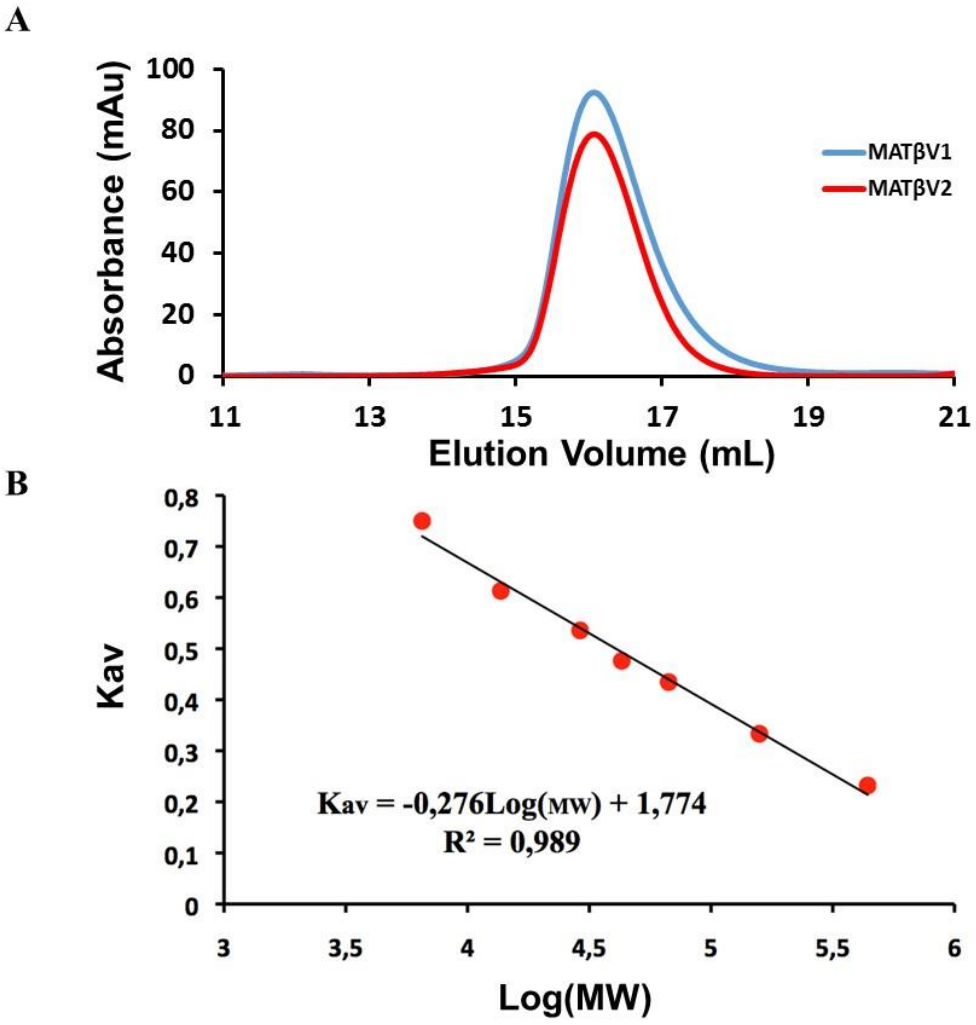

Figure 6.2. A. MATβ proteins elute as single peaks. **A**. Both MATβV1 (in blue) and MATβV2 (in red) elute as one peak. Elution buffer HEPES pH 7.5, 200 mM NaCl and 1 mM TCEP. **B**. Calibration curve made using known protein standards shown as red dots on the graph. From right to left: ferritin (440 kDa), gamma-globulin (158 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), ribonuclease (13.7 kDa) and aprotinin (6.5 kDa). Phase distribution coefficient, K_{av}, was calculated based on elution volume of protein standards and plotted against Log of their molecular weight.

6.1.2 The effects of protein concentration on the oligomeric states of MAT proteins as determined by gel filtration.

Changing the concentration of either MATα1 or MATα2 had an apparent effect of the equilibrium of the two populations. For MATα1 a lower concentration (0.3-1.5 mg/mL) caused the majority of the protein to be in the second population (~15.4 mL), but as the concentration increased the proportion of the larger oligomeric state (~13 mL) also increased. Regardless of concentration of MATα1, the smaller oligomeric state is in greater proportion (Fig 6.3A). Taking multiple fractions of MATα1 at different protein concentrations from either the higher molecular weight peak and the lower molecular peak and then running them on a native gel showed that each peak splits into two populations highlighting the dynamic nature of MATα1 (Fig 6.3B).

For MATα2, concentration also has a great impact on the equilibrium between the two populations. At a low concentration (0.3-1.5 mg/mL) the smaller oligomeric population is in favour, but as the concentration is increased this changes. At 3 mg/mL the amount of the higher oligomeric state increases so that the two states are almost equal with a slight favour for the higher oligomeric state. Above 6 mg/mL the second population decreases and the higher oligomeric state is the majority of the protein (Fig 6.4A). According to native gel, MATα2 proteins, just like MATα1, exist in a dynamic equilibrium and show multiple bands and only at 1.5 mg/mL (Lane 2) does it seem to mainly run as a single population (Fig 6.4B)

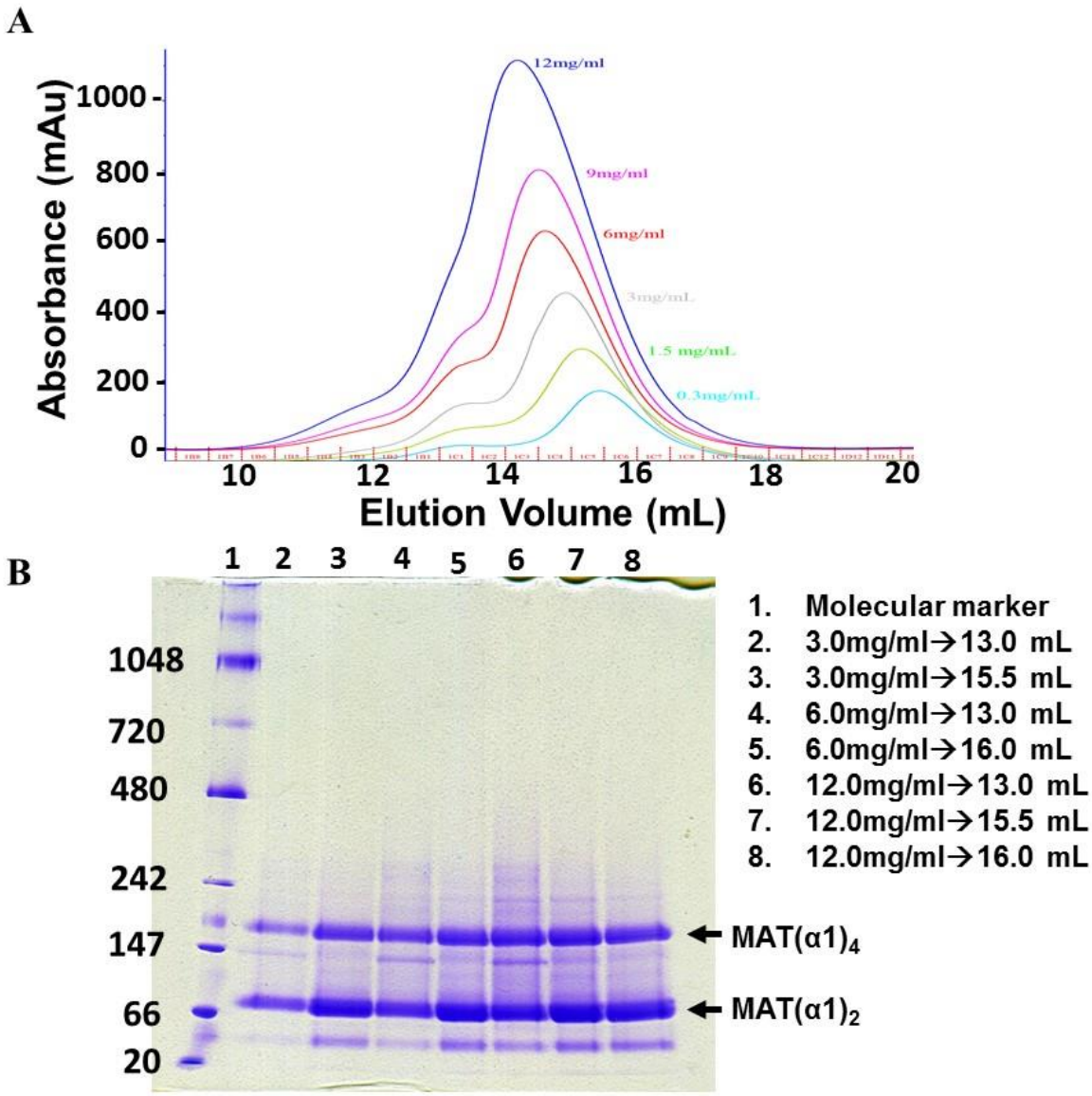

Figure 6.3 Protein concentration dependence of Methionine Adenosyltransferases on oligomeric state. A. MATα1 elution profiles, colour coded according to concentration of protein used. Absorbance was measured at 280 nm which was used as an indication of the presence of protein. Two peaks correspond to tetramer and dimer from left to right. Elution buffer 25 mM HEPES pH 7.5, 200 mM NaCl and 1 mM TCEP. **B**. Native gel of MATα1 at different protein concentrations. C1 refers to the first peak whilst C3 and C4 refer to the second peak. Although protein fractions were taken from different peaks in the gel filtration two bands can be seen indicating the dynamic nature of MATα1.

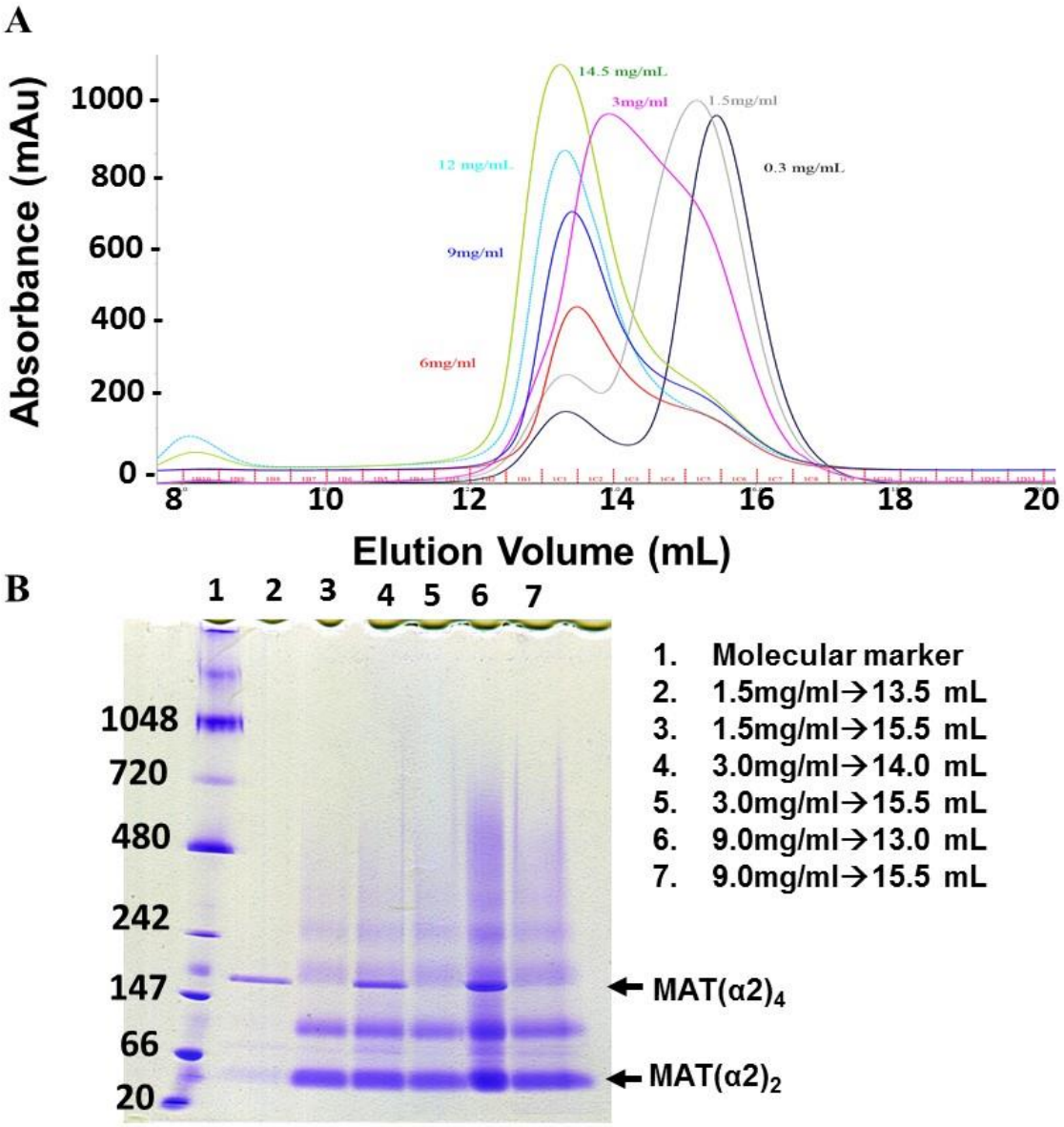

Figure 6.4 Protein concentration dependence of Methionine Adenosyltransferases on oligomeric state. **A**. MATα2 elution profiles, colour coded according to concentration of protein used. Absorbance was measured at 280 nm which was used as an indication of the presence of protein. Two peaks correspond to tetramer and dimer from left to right. Elution buffer 25 mM HEPES pH 7.5, 200 mM NaCl and 1 mM TCEP. **B**. Native gel of MATα2 at different protein concentration. Although protein fractions were taken from different peaks in the gel filtration multiple bands can be seen indicating the dynamic nature of MATα2.

6.1.3 Solution scattering of MAT proteins.

In order to determine oligomeric states of MAT proteins small angle X-ray scattering was performed and structural parameters are presented in Table 6.2. The experiment suggested that MATα1 exists in dimeric (~90 kDa) and tetrameric (~162 kDa) forms in close agreement with predicted molecular weight (93 kDa and 164 kDa, respectively), as expected, different from molecular weight determined by gel filtration. MATα2 only showed one population during SAXS experimentation which corresponded to a tetrameric state (162 kDa). Alignment of SAXS DAMMIN models with crystallographic structures of human MATα1 and MATα2 show a good fit (Fig 6.5A-C). Extra space can be seen for the MATα1 dimer, MATα1 tetramer and MATα2 tetramer which, is due to the flexible nature of the N-terminus that is disordered in the crystal structures.

Both MATβ isoforms, MATβV1 and MATβV2 data were collected through static-SAXS experiments, in the presence and absence of the co-factor NADP, and the proteins appear to be monomers, which is consistent with gel filtration experiments (Table 6.2). Although NADP does not cause a change in oligomeric state it does create a tighter protein as the maximum particle distance (D_{max}) decreases for both MATβ isoforms. Aligning the DAMMIN model to the crystal structure of human MATβ (2YDY) gave a good agreement, but clearly indicating the presence of C-terminal region and the N-terminal region that had not been observed in the high resolution crystallographic structure of MATβ (Fig 6.6). Modelling both the N and C-terminals of MATβ using CORAL produced a model that satisfied the density obtained from SAXS data (Fig. 6.7A) and produced a predicted one dimensional (1D) scatter that shows an excellent agreement with the SAXS data of MATβV2 (χ^2=3.3) (Fig. 6.7B). The C-terminus becomes ordered

during interaction with MATα2 (section 5.1.7. Fig 5.9), but why the N-terminus is not observed in any crystal structure remains an open and intriguing question.

Table 6.2 Structural parameters of MAT proteins as determined by Small angle X-ray scattering. Radius of gyration (Rg), maximum particle distance (D_{max}) and I(0) were calculated using Primus and Scatter whilst molecular weight was calculated using SAXS MoW server (Fischer et al, 2010).

	Rg (Å)	D_{max} (Å)	I(o)	Expected Molecular weight (kDa)	Calculated Molecular weight (kDa)
MATα1 (dimer)	29.6	89.0	$4.54x10^{-2}$	~86.0	93.0
MATα1 (tetramer)	38.5	118	$3.62x10^{-2}$	~172	164
MATα2 (tetramer)	37.7	116	$2.41x10^{-1}$	~180	162
MATβV1	30.0	94.0	$5.57x10^{-1}$	~37.5	41.9
MATβV1+NADP	26.0	92.0	$2.90x10^{-1}$	~37.5	35.1
MATβV2	27.9	90.0	$4.53x10^{-1}$	~36.4	37.2
MATβV2+NADP	25.2	88.0	$2.93x10^{-1}$	~36.4	36.0

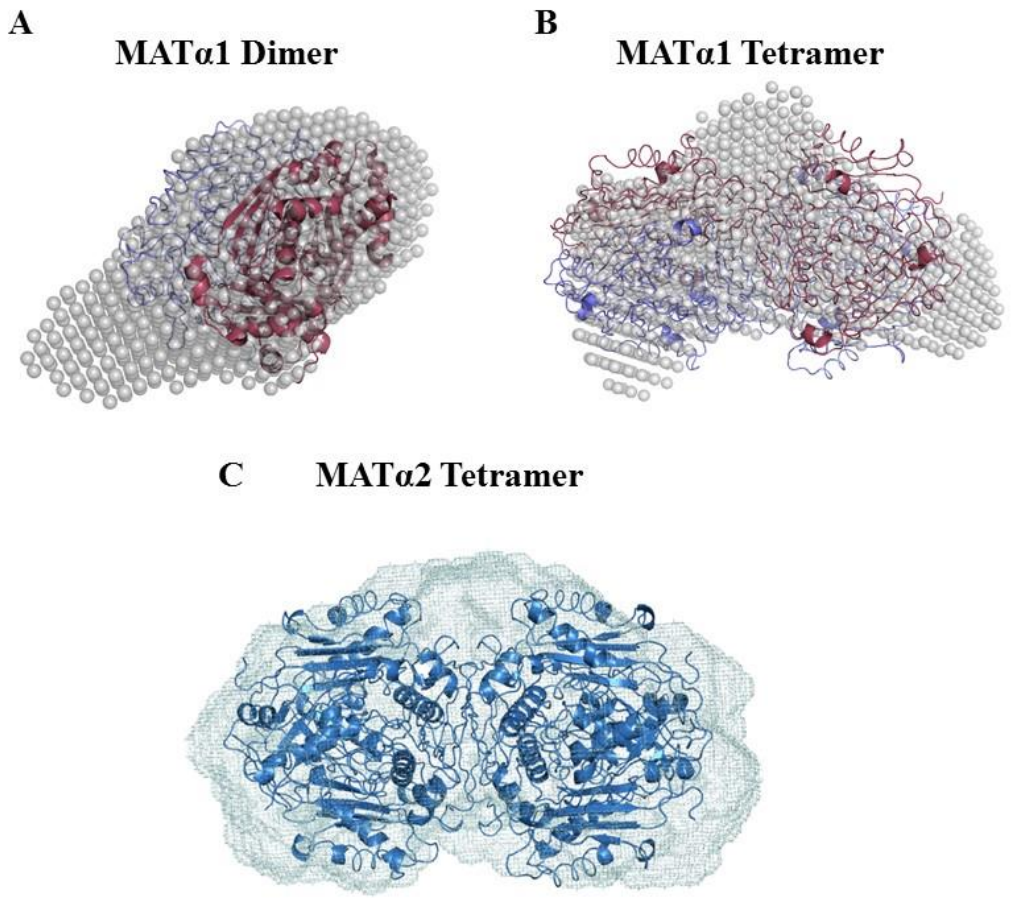

Figure 6.5. Alignment of crystallographic structures of MATα proteins with SEC-SAXS data. Using the P(r) of the solution scattering data, twenty runs of DAMMIN were performed before models were aligned and averaged. The average model was then subjected to a further run of DAMMIN before being aligned using SUPCOMB against the crystal structure of MATα1 (20BV) or MATα2 (2P02) (section 4.2.23). Coot was used to generate the tetrameric form of both MATα isoforms. **A**. MATα1 dimer. **B**. MATα1 tetramer. **C**. MATα2 tetramer. Extra space can be seen in the models where the N-termini of the MAT proteins are not present in the crystal structure, but they would have contributed to the overall scattering and, therefore, the DAMMIN model. Proteins were purified as described in section 4.2.11. Protein concentrations and data analysis is described in section 4.2.23.

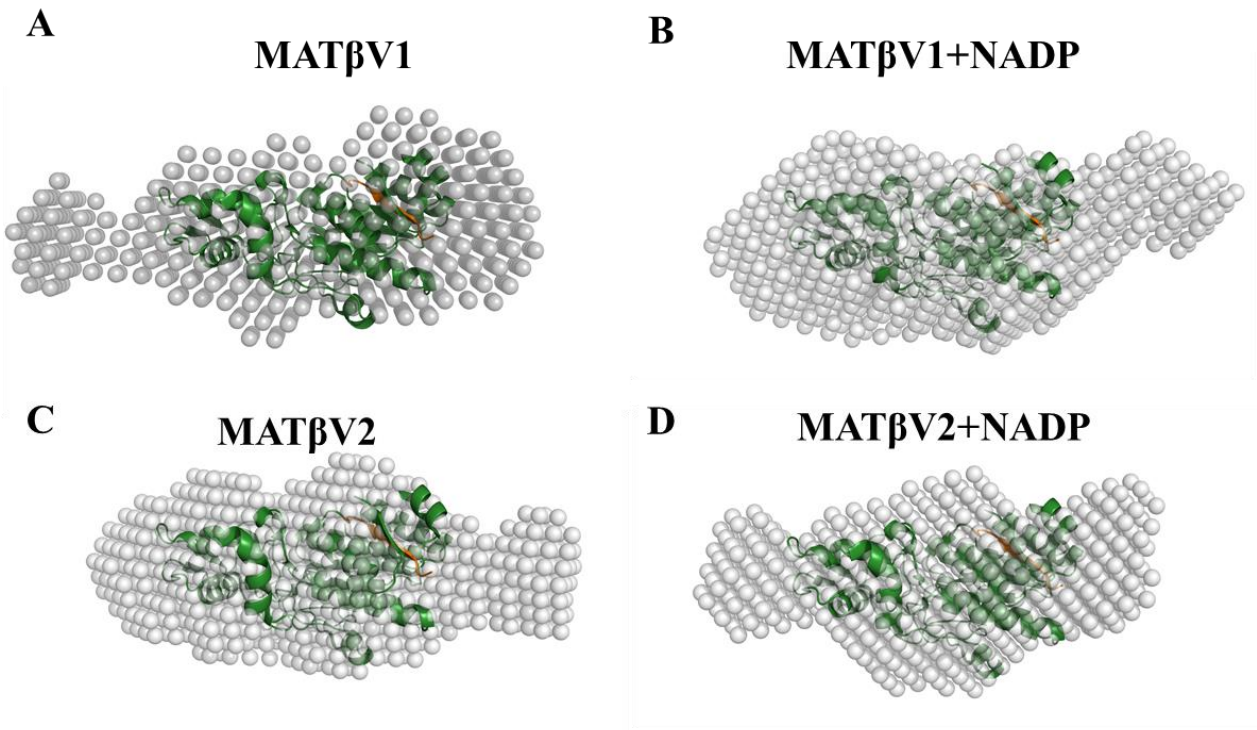

Figure 6.6. Alignment of crystallographic structure of MATβ proteins with Static SAXS data. Using the P(r) of the solution scattering data, twenty runs of DAMMIN were performed before models were aligned and averaged. The average model was then subjected to a further run of DAMMIN before being aligned using SUPCOMB against the crystal structure of MATβ (2YDY) (section 4.2.23). **A**. MATβV1 dimer. **B**. MATβV1+NADP. **C**. MATβV2. **D**. MATβV2+NADP. Extra space can be seen in the models where the N-termini of the MAT proteins are not present in the crystal structure, but they would have contributed to the overall scattering and, therefore, the DAMMIN model. Proteins were purified as described in section 4.2.11. Protein concentrations and data analysis is described in section 4.2.23.

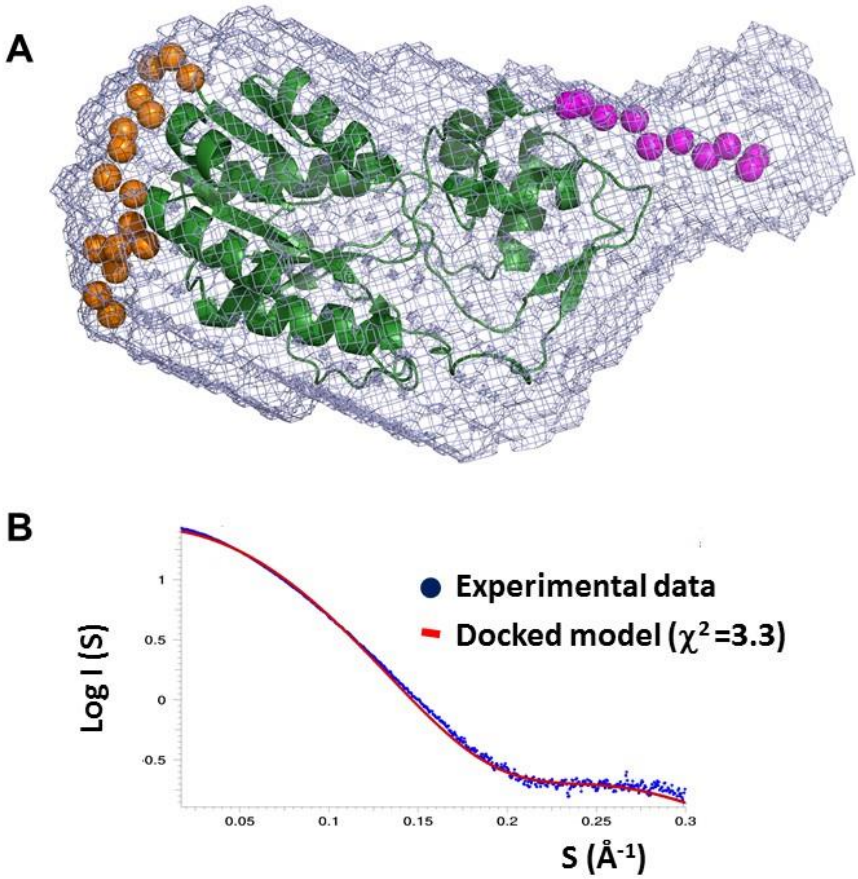

Figure 6.7. SAXS modelling of MATβV2. A. The *ab initio* shape reconstruction of MATβV2 by DAMMIN using P1 symmetry (light blue mesh), showing the fit with the crystallographic structure of the monomer (PDB:2YDY), the orange balls correspond to the flexible N-terminus whilst the purple balls are the modelled C-terminal. Extra residues were modelled with CORAL(Petoukhov et al. 2012) (section 4.2.23). **B**. Experimental spectrum (blue) of the small angle scattering of MATβ and the simulated fit (red) obtained from the modified crystal structure (PDB=2YDY) that contained both the C and N-terminal regions modelled using CORAL (q_{max} = 0.30 Å^{-1}) (χ^2 =3.3). q_{max} is the maximum scattering angle that data was measured.

6.1.4 MATβV1 and MATβV2 are monomers that do not interact *in vitro*.

Using MATβ proteins that contained a N-terminal histidine tag, pull down experiments were performed whereby either MATβV1_HIS or MATβV2_HIS proteins were immoblised on nickel resin and the other isoform wild type protein containing no histidine tag was passed through the resin (Fig 6.8). This was done in order to evaluate whether or not the different isoforms of MATβ interact with each other. MATβV1_HIS and MATβV2 do not interact (lanes 2-6) as MATβV2 was eluted from the nickel resin (lane 5) whilst the immoblised MATβV1_HIS remained (lane 6). For MATβV2_HIS versus MATβV1 both proteins ran at the same molecular weight (lanes 7 & 9) so from SDS gel it was not possible to tell if they eluted as separate proteins (lanes 11-15).

Gel filtration experiments were performed using the wild type MATβ isoforms to see if a shift in elution would occour upon mixing the isoforms and this was done in the presence and absence of NADP. It was clear that MATβV1 and MATβV2 do not interact, as both proteins eluted at ~16 mL and that the MATβV1+ MATβV2 peak is just a combination of the two individual proteins with the same elution volume (Fig 6.9A). The same pattern can be observed in the presence of NADP where binding of NADP does not cause interaction between the two isoforms. Excess NADP can be seen at 20.5 mL (Fig 6.9B). Running the peak of MATβV1+MATβV2 on SDS gel shows that in the absence of NADP (lanes 6-10) or presence (lanes 10-13) that both proteins are present in similar levels (Fig 6.10).

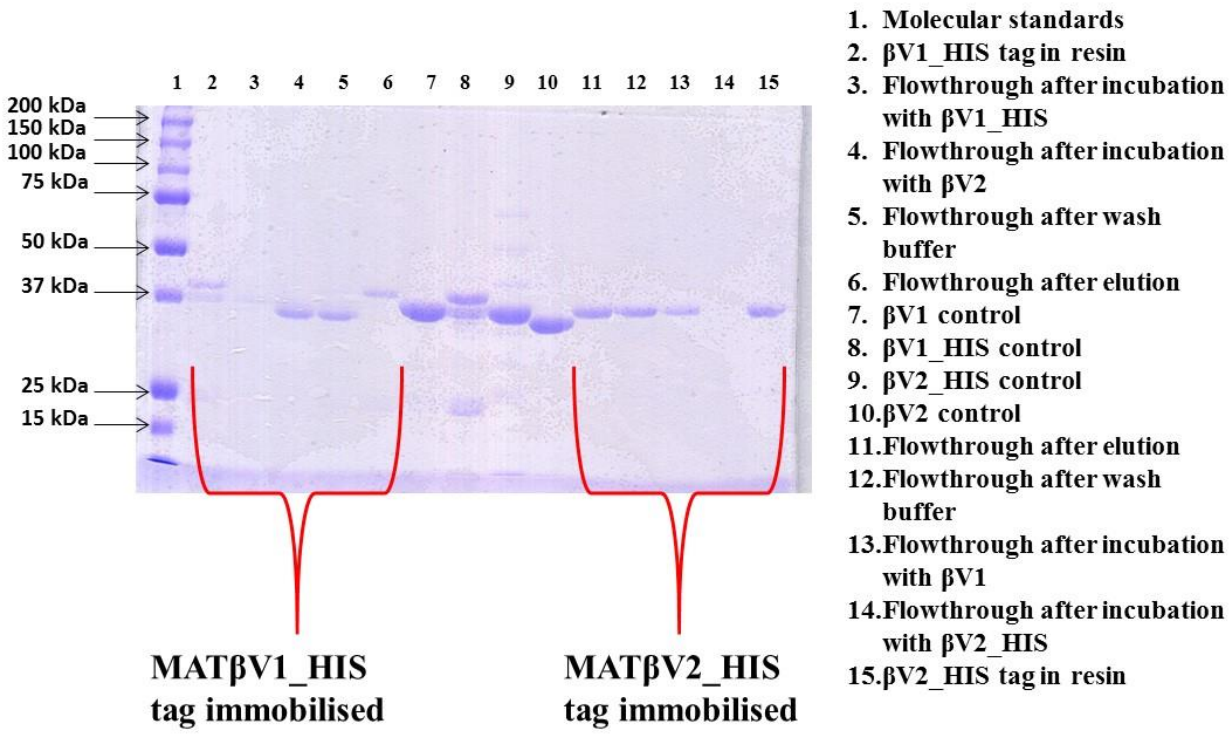

Figure 6.8. Pull down experiments showing that MATβ isoforms do not interact. MATβV1_HIS or MATβV2_HIS proteins (50 μM) were incubated with nickel resin, that had be equilibrated with 25 mM HEPES pH 7.5, for 1 hour before being loaded into 2 mL flow column. Next, the opposing wild type isoform (50 μM) was passed through the column and the flow-through was collected. A low imidazole (25 mM HEPES pH 7.5, 40 mM imidazole) wash was applied to remove no specific binding. An elution buffer containing 25 mM HEPES pH 7.5 and 250 mM imidazole was applied to the column and the flow-through was collected. Lanes 2-6 show MATβV1_HIS versus MATβV2. Lanes 7-10 show control proteins. Lanes 11-15 show MATβV2_HIS versus MATβV1. The flow-through at each step was collected in order to verify if the protein had bound or simply passed through the nickel resin.

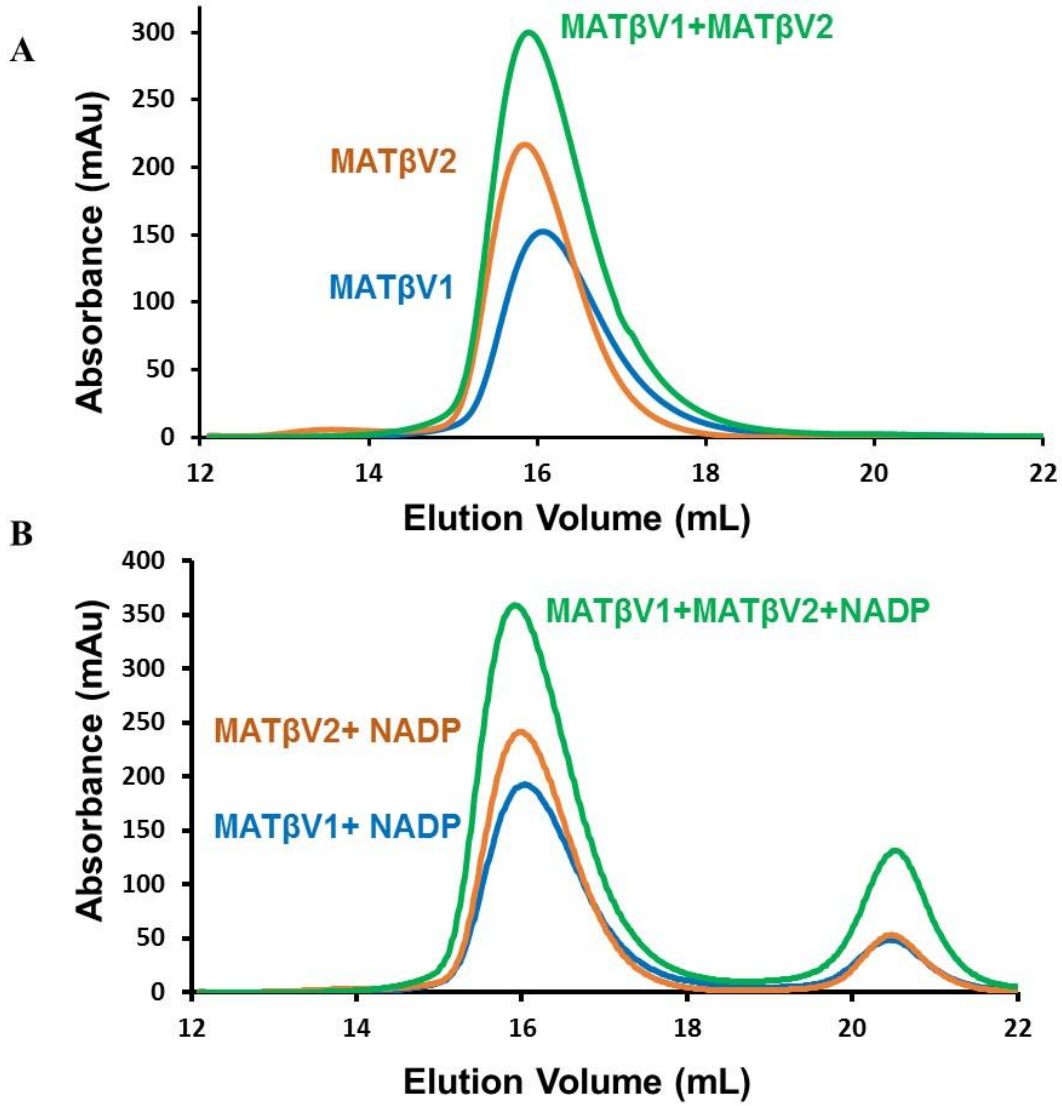

Figure 6.9. MATβ proteins do not interact and elute as single peaks. A. MATβV1 (in blue), MATβV2 (in orange) and MATβ1+MATβV2 (green) elute as one peak. Elution buffer was 25 mM HEPES pH 7.5, 200 mM NaCl and 1 mM TCEP. **B**. As in A but in the presence of 250 µM NADP. 100 µM of each protein was applied to the column. 50 µM of each protein was incubated together for 1 hour at 4 °C, prior to gel filtration (See section 4.2.13).

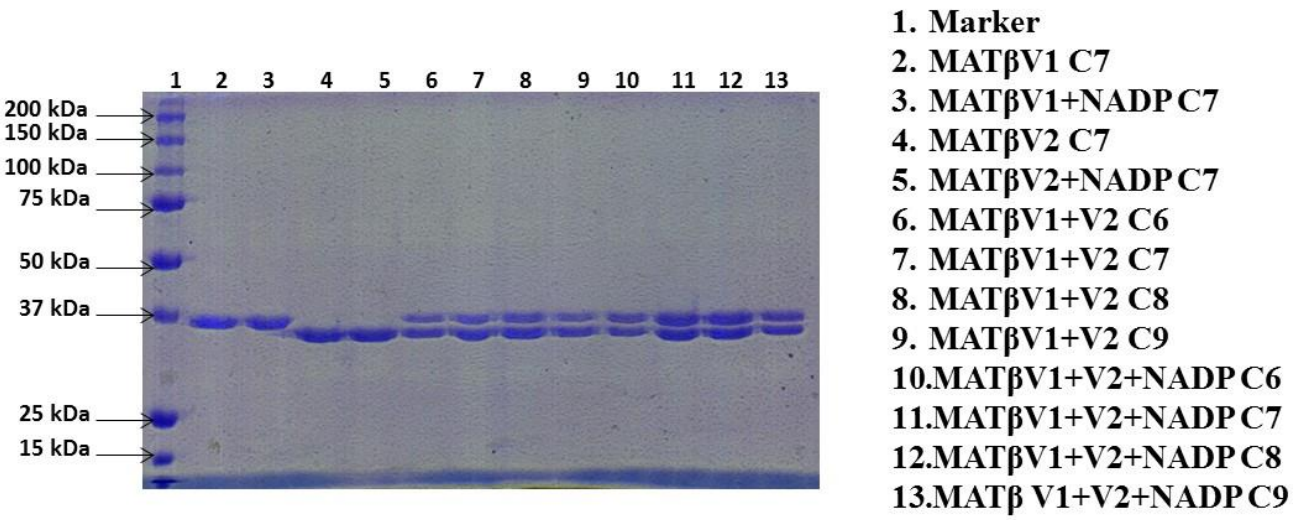

Figure 6.10. SDS page of fractions from gel filtration, of incubating MATβV1 with MATβV2 in the presence and absence of NADP. MATβV1 (100 μM) and MATβV2 (100 μM) were mixed together for one hour at 4°C before being injected into a gel filtration Superdex 200 10/300 in the presence and absence of NADP. C6-C9 refer to the fractions of the gel filtration and have the elution volume of 15-18 mL (Fig 6.9).

6.1.5 Novel position of methionine in MATα2: from Substrate binding to formation of SAMe.

The atomic resolution structure (SAMe+ADO+MET+PPNP bound) at 1.1 Å (Table 6.3), which had been incubated with AMP-PNP and methionine shows multiple occupancies of ligands in different subunits including substrates and products within the active site: Methionine (occupancy 0.3), adenosine (occupancy 0.5), SAMe (occupancy 0.5) and PPNP (occupancy 0.8) can be seen. The omit map clearly shows the substrate methionine in a novel position which sits within the active site so that its functional group directed away from its final place as part of the product SAMe (Fig 6.11A). In order to observe the mixed active site in a clearer way two snapshots of intermediate reactions were created (Fig 6.11B,C). The main chain of methionine is stabilized through a host of hydrogen bonding between residues of different MATα2 monomers (grey and dark grey), which include OD1 of Glu70 (2.8 Å), Asp258 (2.7 Å), NE2 of Gln113 (2.9 Å) and water molecules. The position of adenosine is maintained through π-π stacking between the purine base adenine and Phe250, as well as several hydrogen bonds with the OG of Ser247 (2.8 Å), main chain oxygen of Arg249 (2.9 Å) and waters (Fig 6.11B). One water forms a hydrogen bond with another water and the O4 of adenosine, and occupies the position where later the sulphur of SAMe lies (Fig 6.11B).

The substrate methionine forms a hydrogen bond via its nitrogen with OD1 atoms of both Glu70 (2.8 Å) and Asp258 (2.7 Å) and via its two terminal oxygens with NE2 atom of Gln113 (3.0 Å) and water ligands (Fig 6.11B). The triphosphate can be modeled precisely due to high quality of the omit map at atomic resolution (Fig 6.11, 6.13A). The PPNP is secured in place through multiple interactions with residues His29, Lys181, Lys265, Ap291 and Ala281, waters, two hexacoordinated magnesium ions and a pentacoordinated

potassium ion (Fig 6.12). The high resolution of this structure allows for the observation of three snapshots of the reaction: the hydrolysis of the AMP-PNP into adenosine and PPNP, and the orientation of methionine before SAMe formation (Fig 6.11B-C) is observed whilst, the formation of SAMe and the placement of PPNP before its hydrolysis into pyrophosphate (PPi) and orthophosphate (Pi) (Fig 6.11D) is seen. SAMe lies within the active site interacting with the same residues that secure the methionine main chain and the adenosine molecule, identical to those in the first step, whilst the SD atom occupies the position previously occupied by a water molecule.

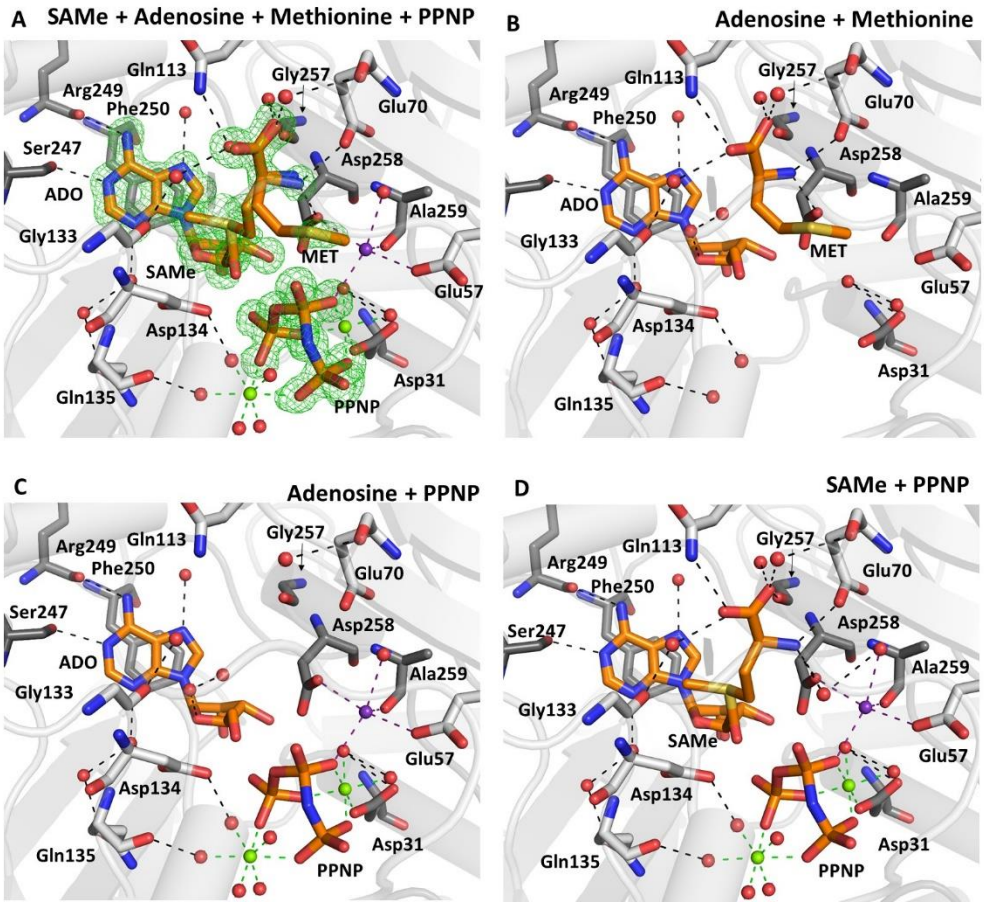

Figure 6.11 SAMe synthesis in human methionine Adenosyltranferases. **A**. Active site of SAMe+ADO+MET+PPNP bound MATα2 showing low occupancy substrate methionine (0.3), adenosine (0.5), SAMe (0.5) and PPNP (0.8). **B**. Novel position of substrate methionine, adenosine and PPNP all hydrogen bonded to MATα2 residues and water molecules. **C**. Adenosine and PPNP within the active site. **D**. Product SAMe and PPNP shown hydrogen bonded by the same residues that secure both substrates. Omit map (Fo-Fc) is coloured green around and contoured at 3σ level. Two MATα2 monomers form a dimer coloured grey and dark grey. Substrates and residues are shown as sticks, colour corded according to the nature of the atom, where blue corresponds to nitrogen, red to oxygen. Hydrogen bonds are shown as black dotted lines, green and purple dotted lines show magnesium and potassium ions coordination, respectively. Water molecules, magnesium and potassium ions are represented as red, green and purple spheres respectively.

6.1.6 Coordination of Ions within the MATα2 active site.

The high resolution data has enabled the identification of ions and their co-ordination within the active site (Fig 6.12A). Both magnesium ions are hexacoordinated whilst the single potassium ion is pentacoordinated. The first magnesium (Fig 6.12B) interacts with four water molecules and with two atoms, O2A and O3G, of the PPNP moiety. The second magnesium ion interacts with two waters, OD2 of Asp31 and three atoms of PPNP, O1A, O2B and O2G (Fig 6.12C). The potassium ion interacts with a single water, residues from different MATα2 monomers such as OE2 of Glu57, OD1 Asp258 and O of Ala259 and also the O2B atom of PPNP (Fig 6.12D). All magnesium bonds are between 1.9-2.3 Å whilst the potassium bonds are 2.7-2.8 Å (Harding 2006).

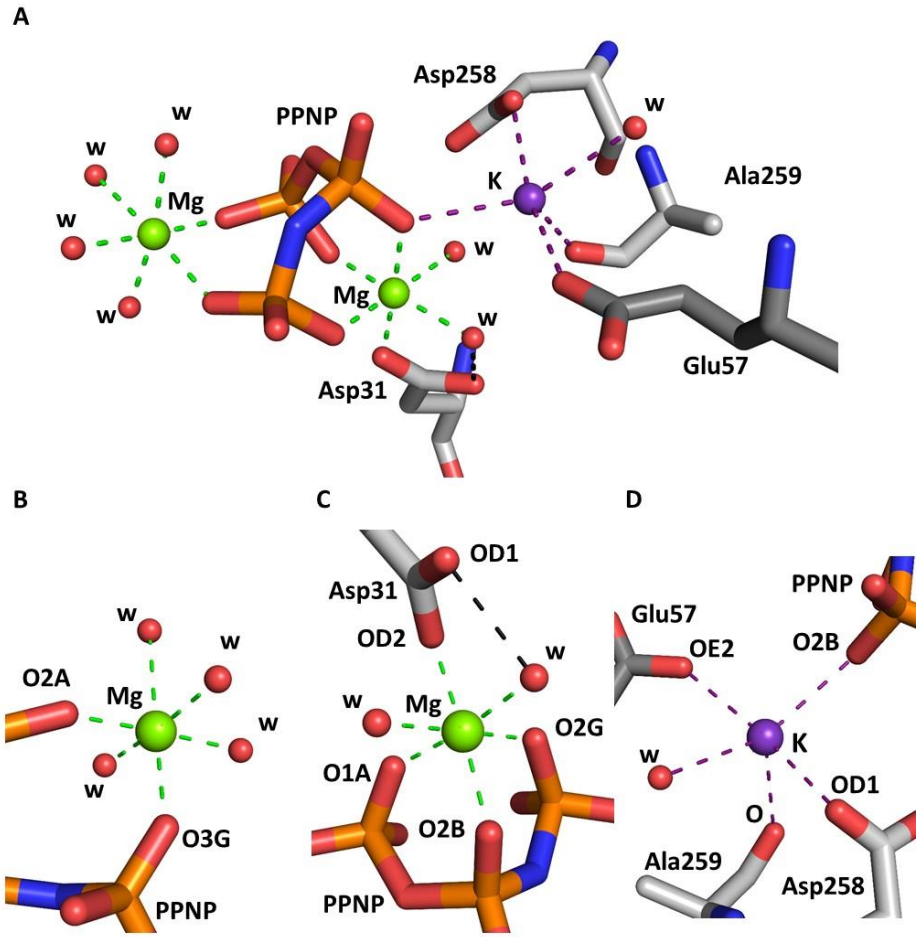

Figure 6.12. Magnesium and potassium ions coordination in ligand bound MATα2. **A**. Active site of MATα2 (SAMe+ADO+MET+PPNP) showing two magnesium ions (green) and one potassium ion (purple). **B**. Closer view of the first magnesium ion, hexacoordinated by 4 waters, and 2 oxygen atoms, O2A and O3G of PPNP. **C**. Closer view of the second magnesium ion, hexacoordinated by two waters, and 4 oxygen atoms, OD2 Asp31, O1A, O2B and O2G of PPNP **D**. Potassium ion is hexacoordinated by water and 3 oxygen atoms originating from PPNP (O2B) and residues Asp258 (OD1) and Ala259 (O). PPNP and MATα2 residues are shown as sticks colour coded according to atom type (C- in grey, O- in red), waters are shown as red balls. Coordination bonds for magnesium ions are shown as green dotted lines and for potassium as purple dotted lines.

6.1.7 Mode of PPNP binding in the absence of the regulatory subunit.

Re-examination of the PPNP conformation in the MATα2βV2 complex structure (4NDN) using this group from the 1.1Å SAMe+ADO+MET+PPNP structural model suggests that the PPNP group was previously not correctly positioned, as lower resolution MATα2β structures did not allow the identification of magnesium ions. Positioning of these phosphates and the three metal ions from the atomic resolution structure (Fig 6.13A) into the lower resolution electron density map of the MATα2β complexes gives a much improved electron density fit (Fig 6.13B). The conformation of the triple phosphate is preserved in MATα2βV2, SAMe+ADO+MET+PPNP and in the AMP-PNP from *E.coli* MAT (1P7L) in spite of cleavage from the adenosine moiety (Fig. 6.13C).

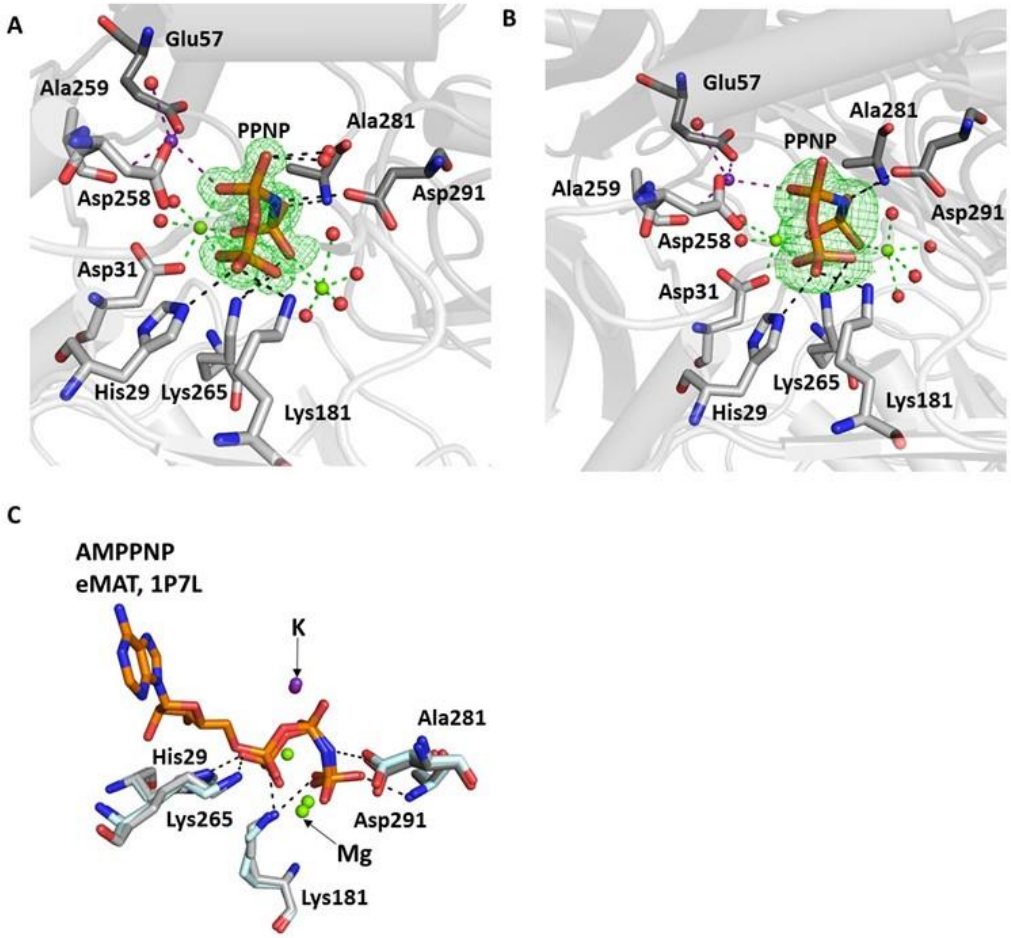

Figure 6.13 Comparison of phosphate moiety conformations in human MAT enzymes. **A**. Omit (F_o-F_c) electron density map, contoured at 3 sigma level around the PPNP from SAMe+ADO+MET+PPNP MATα2 structure. **B**. Coordination of PPNP from re-refined MAT$(\alpha 2)_4(\beta V2)_2$ (PDB 4NDN) showing that position of PPNP is the same as in SAMe+ADO+MET+PPNP. Omit map and colour cording is as in A. **C**. Superposition of AMP-PNP from eMAT (PDB accession number 1P7L) and PPNP from SAMe+ADO+MET+PPNP MATα2 shown in stick representation. Residues from MATα2 monomers shown as sticks in grey or dark grey with hydrogen bonds are black dotted lines with magnesium ions as green balls and potassium in purple. Coordination of magnesium and potassium ions are shown as green and purple dotted lines respectively. All phosphate moieties are shown as sticks, colour corded by atom: P-orange, O-red, N-blue.

6.1.8 PPNP does not provide the energy to open the gating loop.

In the structure of PPNP bound MATα2 it has a disordered gating loop. The complex $MAT(\alpha2)_4(\beta V2)_2$ (4NDN) contains 4 active sites, two of these reveal ordered gating loops while the other two do not show the gating loop, presumably due to its highly flexible disordered nature (Murray et al. 2014). The two ordered sites contain SAMe, PPNP, Mg^{2+} and K^+, whist the disordered sites only contain an ethylene glycol molecule from the cryoprotectant. A comparison of three structures (PPNP bound, $MAT(\alpha2)_4(\beta V2)_2$ and SAMe+ADO+MET+PPNP bound) shows that gate is open when the active site is occupied by PPNP. In contrast the gate is shut like a lid when the active site is occupied by SAMe or adenosine. Comparison of the SAMe+ADO+MET+PPNP structure with holo re-refined MATα2 from $MAT(\alpha2)_4(\beta V2)_2$ complex (Fig. 6.14A, B) shows that the residues in the active site do not differ in position and the mechanism of the active site and the catalytic subunit is preserved in the absence of the regulatory subunit MATβ. The gating loop in the (PPNP bound) (Fig. 6.14C) and apo subunit of $MAT(\alpha2)_4(\beta V2)_2$ complex (Fig. 6.14D) are invisible, suggesting that the opening of the gate to release product is not driven by hydrolysis of the triple phosphate. Gln113 is the only residue of the gating loop to interact with SAMe directly through its NE2 atom forming a hydrogen bond with the O of SAMe. Ser114 interacts with a water molecule that also interacts with SAMe (Fig 6.15). What drives the opening of the gate and release of product thus remains an open question.

In SAMe+ADO+MET+PPNP structure OE1 atom of Glu70 interacts with the nitrogen atoms of Lys61 (NZ) and Ala259 (N), while its OE2 oxygen is hydrogen bonded to the substrate methionine or SAMe (2.8 Å). Glu70 maintains the same position in all holo

active sites (Fig. 6.14A-C). In apo MATα2 subunit of MAT$(\alpha 2)_4(\beta V2)_2$ complex (Fig. 6.14D) there are significant movements of the residues. For example Lys61 adopts a different conformation compared to those in all other MATs. The change in lysine conformation also enables movement of other residues in the active site, for example, Gly257 and Asp258 both move away from the active site. Oxygen atom, OD1 Asp258 can no longer participate in the interaction with the potassium ion coordinated by O2B of PPNP, as the distance between the side chain and potassium ion increases to 4.1 Å. Also, the movement of Asp258 enables it to interact with N of Ala259. The oxygen, OE1 of Glu70 is unable to interact with Lys61 or Ala259, but could do so with the N of Asp258, which alters its conformation as a result of Lys61 movement.

In the MATα2 structures, SAMe+ADO+MET+PPNP bound (Fig. 6.14A) the potassium ion is present and the interaction with Asp258 and PPNP occurs. In the PPNP bound MATα2 (Fig 6.14C) the PPNP adopts a different position whereby the central phosphate P_B has rotated onto the side. The orientation of the residues around the PPNP remain in the same position as in the holo forms and interaction with a magnesium ion occurs but interaction with the second magnesium and potassium is lost. Due to the position of Lys61 and the coordination through the magnesium ion, the PPNP is able to occupy the active site whilst the gating loop is disordered, i.e., when it is no longer in the closed position. Furthermore, there are movements in residues that interact with the potassium ion, Glu57 and Ala259 and the PPNP moiety. Three further residues, His29, Lys181 and Asp291, which interact with the PPNP, show movement in the absence of PPNP in the empty active site.

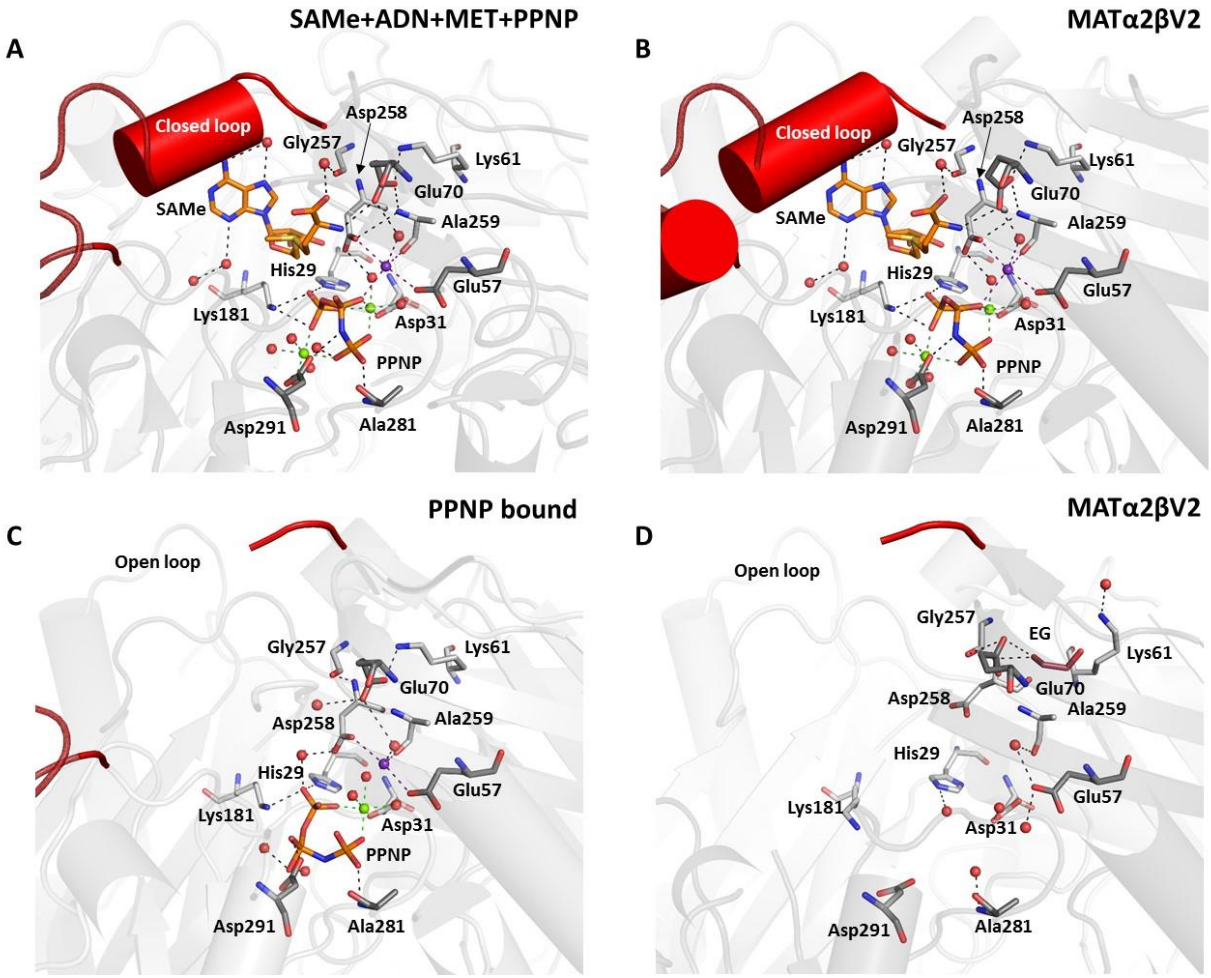

Figure 6.14 Comparison of the holo and apo active sites in MATα2 and MATαβ. A. Holo SAMe+ADO+MET+PPNP MATα2 active site. **B**. Holo MATαβ (PDB accession number 4NDN) active site. **C.** Holo PPNP bound MATα2 active site with disordered gating loop. **D**. Apo MATαβ (PDB accession number 4NDN). Gating loop is highlighted in red. Two MATα2 monomers form a dimer coloured grey and dark grey with residues of interest shown as sticks, SAMe can be seen in orange, PPNP as orange sticks, with magnesium ions as green, potassium ion in purple and water molecules as red balls. Hydrogen bonds are shown as black dotted lines, green dotted lines show interaction with magnesium ions whist purple dotted lines show interaction with potassium ion.

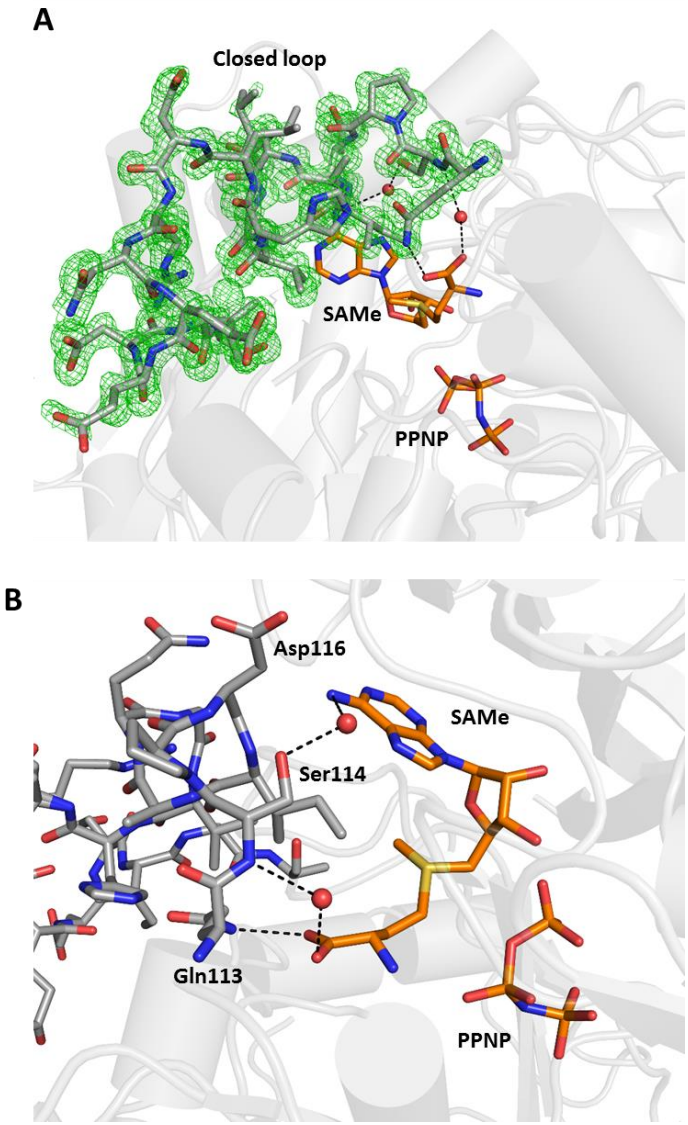

Figure 6.15. Ordered gating loop of holo MATα2. A. Omit map (Fo-Fc) is coloured green around the gating loop from SAMe+ADO+MET+PPNP bound MATα2 structure and contoured at 3σ level. **B**. Interactions between the gating loop and the product SAMe. Gating loop residues are shown in stick representation. Two MATα2 monomers form a dimer coloured in cartoon representation. Substrates and residues are shown as sticks, colour corded according to the nature of the atom, where blue corresponds to nitrogen, red to oxygen. Hydrogen bonds are shown as black dotted lines whilst water molecules are represented as red spheres.

6.1.9 Catalytic steps captured for human MAT.

Using the following structures, SAMe+ADO+MET+PPNP, PPNP bound, MATα2βV2 complex and MAT from *E.coli* (eMAT, 1P7L) a clearer picture of SAMe synthesis can be provided from structural snapshots starting from the substrate methionine. In the initial position the functional group of methionine (SH) forms hydrophobic interactions with the phosphate groups of ATP (here AMP-PNP, 1P7L) and the nitrogen on the methionine is at a suitable distance to form two strong hydrogen bonds with OD1 of Asp258 (2.7 Å), and OD1 of Glu70 (2.8 Å) (Fig 6.16A). The conformation of AMP-PNP is stabilized by two magnesium ions, one of which interacts with O1A and O2B of the triple phosphate (PPNP) whilst the other interacts with its O2A, O2G and O3G. In order for the sulphur atom of methionine to bond with the C5' atom of the AMP-PNP, the methionine has to move its side chain from its initial position towards the C5' of AMP-PNP (Fig 6.16B) while, the nitrogen of methionine still interacts with OD1 of Asp258 (2.7 Å), and OD1 of Glu70 (2.8 Å) to aid in its stabilisation and rotation.

Once the methionine is in position its electron rich sulphur reacts with the positively polarized C5' of AMP-PNP in a nucleophilic substitution to produce adenosine and PPNP (Fig 6.16C) before the final SAMe and PPNP (Fig 6.16D) is synthesized. MATα2 (PPNP bound) shows a pre-apo state that after the product has been formed, SAMe, it is able to leave the active site leaving the triple phosphate moiety within the active site that still contains a magnesium and potassium ions (Fig. 6.16E) The loss of one of the magnesium ions causes a re-orientation of the PPNP moiety losing the interaction with the potassium ion. In order for the active site to cycle and be ready to receive the substrates again, the triple phosphate must be broken down before exiting the active site. The triphosphatase

activity of MATα2 causes the breakdown of the triple phosphate, which results in the loss of all ion coordination. Movement of Glu57, Glu70, Asp258 and Ala259 occurs after all products have left the active site (Fig. 6.16F) (4NDN) all of which are able to hydrogen bond to water molecules.

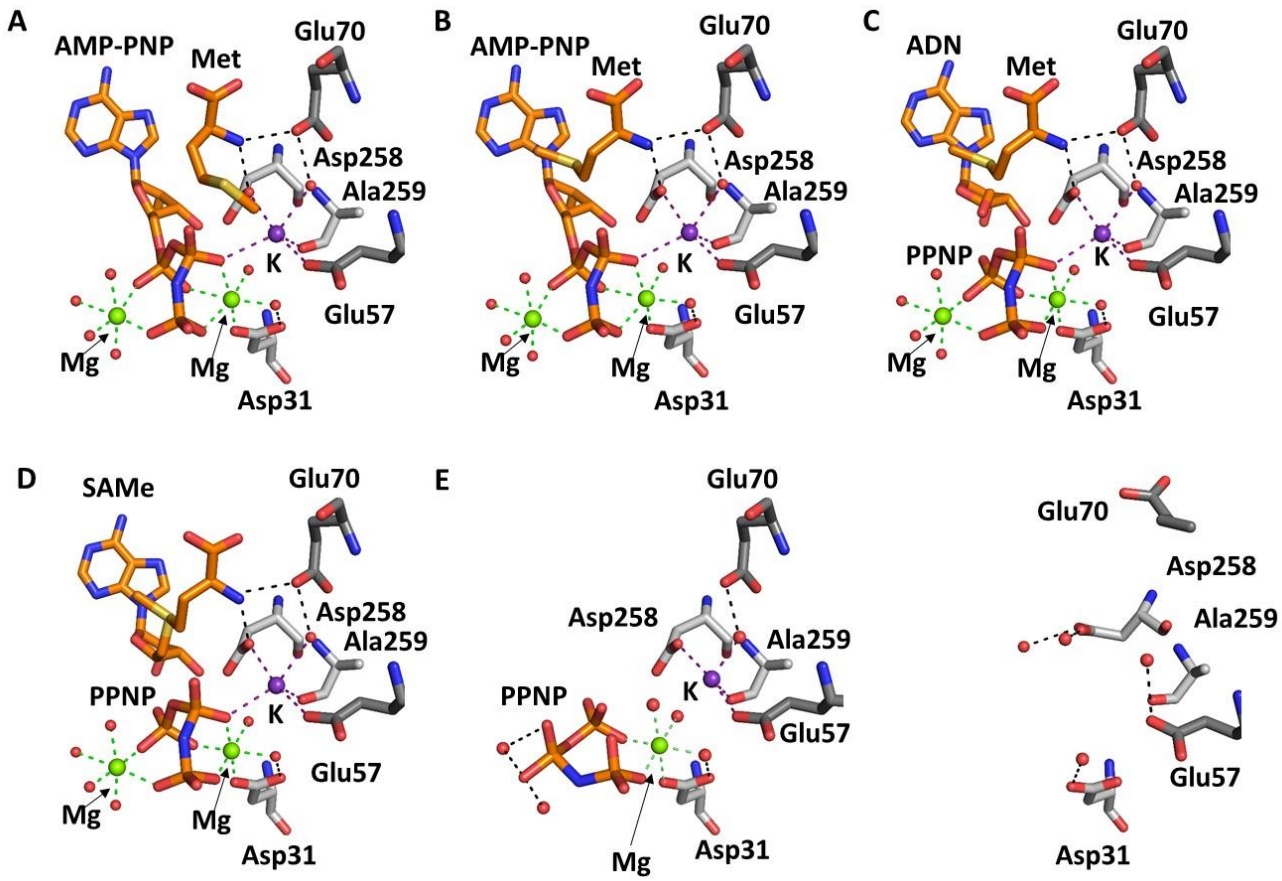

Figure 6.16. Insights into SAMe synthesis in human methionine adenosyltransferase. **A**. Methionine from SAMe+ADO+MET+PPNP and ATP analog AMP-PNP drawn as ineMAT 1P7L. The methionine is in a position where its nitrogen is hydrogen bonded to Asp258 and Glu70. **B**. Movement of substrate methionine as it moves towards C5' bond of AMP-PNP (ligands modelled from 1P7L). **C**. Cleavage of AMP-PNP producing ADO and PPNP from SAMe+ADO+MET+PPNP with methionine (1P7L). **D**. Formation of SAMe from ADO and methionine with PPNP present (SAMe+ADO+MET+PPNP). **E**. Active site after product SAMe has left the active site leaving PPNP (PPNP bound). **F**. On breakdown of the triple phosphate, Glu70 and Asp258 change conformations and magnesium ions no longer coordinate the PPNP. Substrates and residues are shown as sticks with hydrogen bonds shown as black dotted lines. Green and purple dotted lines show magnesium and potassium ions coordination, respectively. Water molecules, magnesium and potassium ions are represented as red, green and purple spheres respectively.

6.1.10 Incorporation of non-native methionine analogs.

The S-adenosylethionine structure is the first non-archaea MAT enzyme structure to contain a nonnative reaction product, S-adenosylethionine. Overall there was no structural rearrangement in the main chain and secondary structure is maintained ((RMSD) 0.185 Å) (Fig 6.17A), when compared to SAMe+ADO+MET+PPNP. S-adenosylethionine occupies the active site in a similar way to SAMe, there is little movement of the residues involved in the stabilization of both SAMe and S-adenosylethionine, including residues Ser247, Arg249 and Asp258 (Fig 6.17B). The extra CH_3 group is able to be accommodated within the active site without causing clashes with residues near the methyl/ethyl group (Fig 6.17B) specifically Ile117, Ile322, Gly133 and Asp134. Omit map shows clear positions for all atoms of S-adenosylethionine molecule (Fig 6.17C).

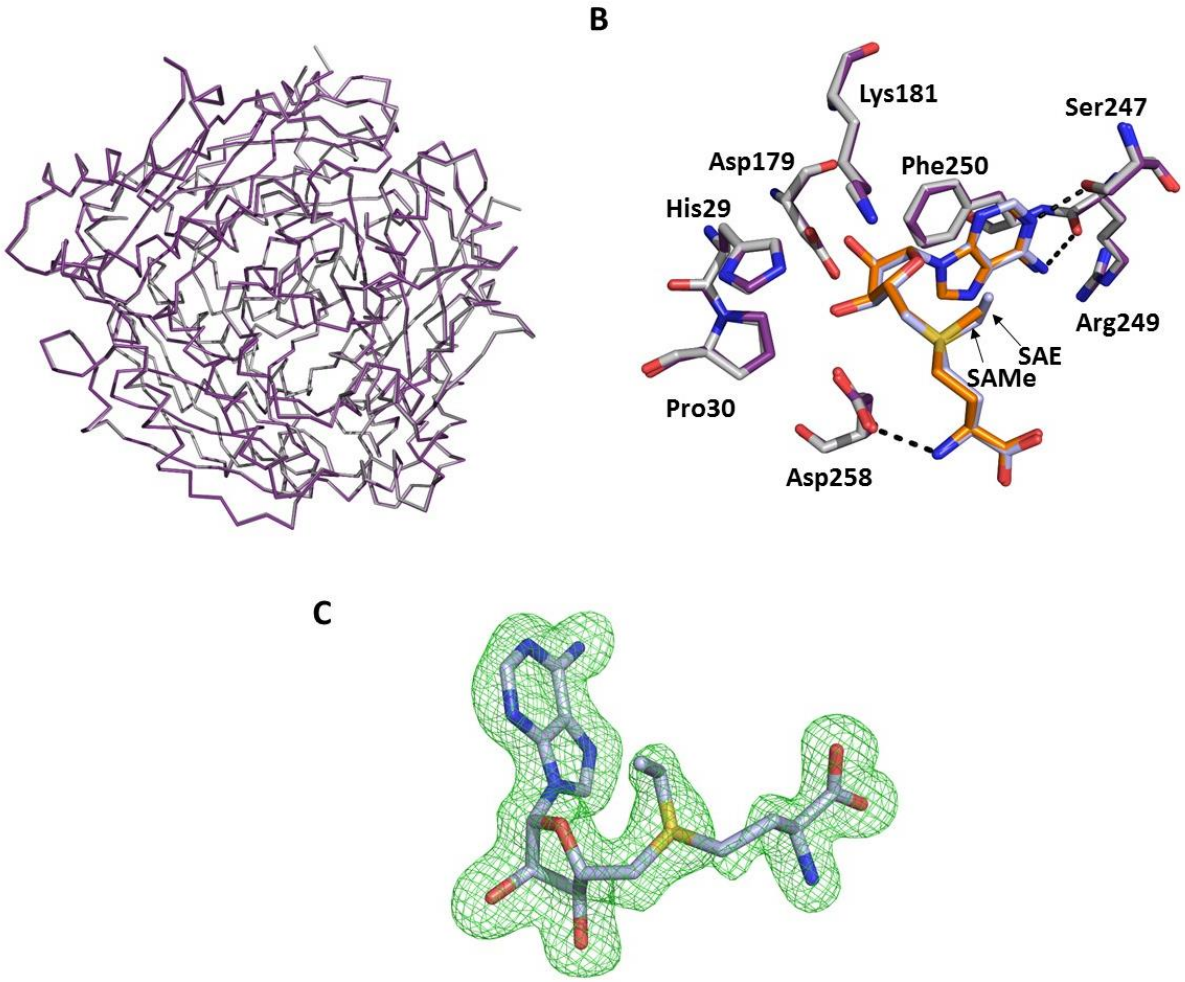

Figure 6.17. Structure of S-adenosylethionine bound MATα2. **A**. Superposition of MAT$(\alpha 2)_2$ with S-adenosylethionine bound (in purple) and SAMe bound (in silver) displayed as lines showing no change in folding (0.185 Å RMSD). **B**. Superposition of SAMe (green) and S-adenosylethionine (light blue) within the active site of MATα2 showing no change in position of surrounding residues. Residues from SAMe bound structure are shown as silver sticks whilst S-adenosylethionine bound as purple sticks. Hydrogen bonds are shown as black dotted lines. **C.** Omit map (Fo-Fc) electron density map and stick representation of SAE from MATα2, the map is contoured at 3 σ level around S-adenosylethionine molecule.

Table 6.3. Data collection and refinement statistics of MATα2 structures. Values in parentheses are for the highest resolution shell.

	SAME+ADO+MET+PPNP bound	SAE bound	PPNP bound
Data collection Wavelength Å	0.92	0.92	0.92
Detector	Pilatus	Pilatus	Pilatus
Space group	I222	I222	I222
Unit-cell dimensions (a,b,c) (Å)	67.92, 94.07, 117.22	68.39, 94.39, 117.39	66.11, 95.33, 117.51
Resolution (Å)	50-1.10	58.7-2.0	74.04-2.33
Rmerge % (last shell)	8.9 (9.1)	10.3 (8.14)	13.6 (7.46)
I/σ (last shell)	24.5 (3.4)	21.3 (3.4)	20.7 (3.9)
Completeness (%)	99.1 (99.0)	99.5 (97.5)	99.9 (99.9)
Redundancy	12.3 (9.7)	13.3 (12.5)	13.0 (13.7)
No. of reflections	176512	33901	16225
Rwork/Rfree	12.9/15.2	13.4/17.7	14.4/20.1
No. of atoms			
Protein	3290	2972	2876
Ligand/ion	68/3	41/3	13/3
water	552	265	170
B factor (Å^2)			
Protein	11.2	20.0	29.1
Ligands/ions	SAMe/ADO/MET/PPNP/ Mg/K 7.5/13.3/14.5/10.3/8.5/10	SAE/PPNP/Mg/K 22.5/22/21/27	PPNP/Mg/K 32.8/28.5/62
Waters	24.0	29.1	32.4
Ramachandran statistics			
Residues in preferred regions	413 (97%)	369 (97%)	358 (98%)
Residues in allowed regions	14 (3%)	12 (3%)	9 (2%)
Outliers	0 (0%)	0 (0%)	0 (0%)
R.M.S deviations			
Bond length (Å)	0.002	0.002	0.014
Bond angles (°)	1.14	1.94	1.66

6.1.11 Binding of molecules into the MATβ binding tunnel of MATα2.

Crystallisation of MATα2 has not only provided insights into how the synthesis of SAMe may proceed, but has also revealed that molecules, such as PEG and imidazole, are able to bind in the tunnel that holds the C-terminal of MATβ. Depending on crystallisation conditions different ligands were seen within this pocket. Crystallising with 0.1 M imidazole or 20 % PEG600 resulted in these molecules co-crystallising with MATα2 (Fig 6.18A, B). Both ligands sit within the beta binding pocket, hydrogen bonded to water molecules. SAMe and PPNP are shown to highlight the position of the active site of MATα2.

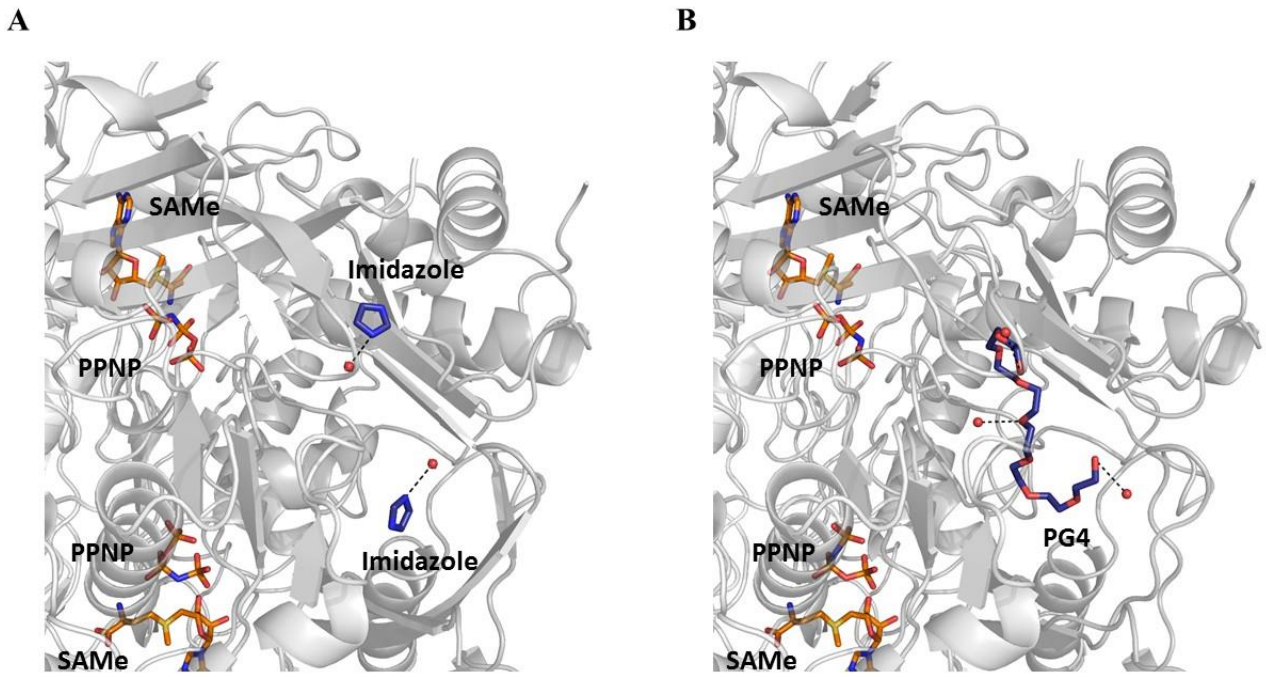

Figure 6.18 Binding of ligands to the MATβ binding tunnel of MATα2. A. When MATα2 was crystallised with 0.1 imidazole pH 7.0, 50 % (w/v) MPD and two imidazole molecules sit within the tunnel. **B**. When MATα2 was crystallised in 0.1 M HEPES pH 7.5, 20 % (v/v) PEG600 and a PEG (PG4) molecule sit within the tunnel. Residues from MATα2 monomers shown as sticks in grey or dark grey with hydrogen bonds are black dotted lines, ligands are shown as stick coloured by atom.

6.1.12 Exploitation of differential scanning fluorimetry (DSF) for analysis of MATα2 ligand binding.

Differential scanning fluorimetry was used to examine whether the presence of ligands affected the stabilisation of MATα2. A range of MATα2 concentrations were tested to see if protein concentration affected the melting temperature (T_m) of the protein. As the protein concentration increased the amount of fluorescence increased (Fig 6.19A). Small differences in melting temperature were observed with changing concentrations, but this was not in a concentration dependent manner (Fig 6.19B) and the average melting temperature over the range of MATα2 concentrations was 46.0 °C. No change in the shape of the unfolding protein was seen over any of the protein concentrations. Fixing the protein concentration at 0.4 mg/mL and examining the effect of pre incubation with an increasing range of the enzyme product, SAMe, showed a stabilisation of MATα2 (Fig 6.19C). Compared to the control a slight shift can be seen with the addition of SAMe. The melting temperature of MATα2 in the presence of SAMe increased and concentrations of 2 mM and above caused similar increases but this was again not in a dose dependent manner (Fig 6.19D). Incubation with 10 mM SAMe provided the greatest thermal shift (+1.50 °C).

The crystal structures of MATα2, presented in section 6.1.11, showed that compounds that were part of the crystallisation conditions could bind to the MATα2 protein. The binding of both imidazole and PEG600 were investigated to see if they changed the thermal stability of MATα2 using differential scanning fluorimetry. Interestingly, incubation with imidazole caused a concentration dependant destabilisation which is clearly visible at 100 mM imidazole (Fig 6.20A). Compared to the control (47.0 °C, no imidazole) clear changes in the melting temperature can be seen from 20 mM imidazole

(46.3 °C). Decrease in the melting temperature suggests that imidazole caused MATα2 to become destabilised (Fig 6.20B). This effect is not seen in the crystal structure as the temperature of crystallisation was only 25 °C even though the concentration of imidazole in crystallisation was 100 mM. Given the protective effect of SAMe it was tested to see whether the presence of 5 mM SAMe could stabilise MATα2 in the presence of imidazole. SAMe does protect MATα2 against the effects of imidazole but a destabilisation still occurs for a concentration over 50 mM imidazole (Fig 6.20C). A stabilisation occurs when a concentration of 1-20 mM imidazole is used but this is due to the presence of SAMe but beyond 20 mM a clear destabilisation effect can be seen (Fig 6.20D). When PEG600 was incubated with MATα2 a concentration dependent stabilisation effect was observed (Fig 6.21A) with a clear shift in the position of the melting curve of all concentrations of PEG600 compared to the control (Fig 6.21A, brown). The melting temperature of MATα2 increased in the presence of all PEG600 concentrations, but the greatest stabilisation was observed with 20 % (w/v) PEG600 (+5.95 °C) compared to the control (Fig 6.21B). This effect was further enhanced by the addition of 5 mM SAMe where a clear shift in melting temperature is seen (Fig. 6.21C). The addition of 5 mM SAMe provided an additional stabilisation effect for the MATα2 which was the greatest at 20 % (w/v) PEG600 (+7.15 °C) (Fig 6.21D). The stabilisation effect that SAMe provides on the recombinant MATα2 protein is clear, either on its own or in the presence of imidazole, and PEG600. Imidazole caused a concentration dependent destabilisation of MATα2 which cannot be recovered with the addition of 5 mM SAMe (Fig 6.22A), whilst PEG600 caused a concentration dependent stabilisation of MATα2 that was further enhanced by the presence of SAMe (6.22B).

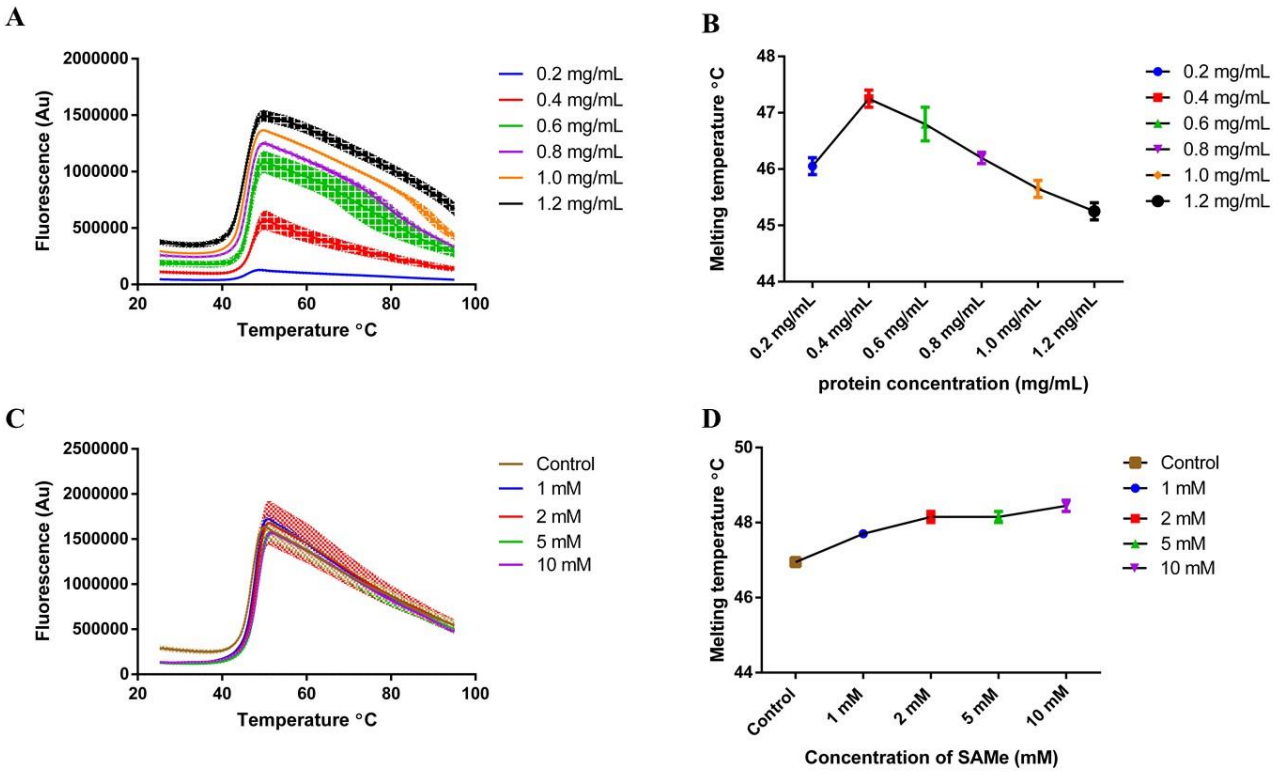

Figure 6.19. Thermal stability assay of MATα2. **A**. Melting curves of MATα2 for a range of protein concentrations (0.2-2.0 mg/mL). **B**. Melting temperatures (°C) of MATα2 over a range of protein concentrations (±range, n=2). At 0.4 mg/mL MATα2 the melting temperature is greatest and there are small variations over the range of protein that were tested. **C**. Melting curves of MATα2 that was pre-incubated with SAMe (1-10 mM) prior to the thermal stability assay. **D**. Melting temperatures (°C) of MATα2 +SAMe over a range of protein concentrations (±range, n=2). SAMe increases the melting temperature of MATα2. A thermal ramping procedure (0.3 °C/min over a range of 25 °C to 94 °C) was employed to denature MATα2 and mean melting temperature (°C) values (±range) are shown for each condition. Section 4.2.19 describes the differential scanning fluorimetry method. The melting temperature was calculated from the first derivative curve of the melting curve and was defined as the maximum peak of the first derivative curve (Sun et al. 2015). Figures were produced using GraphPad.

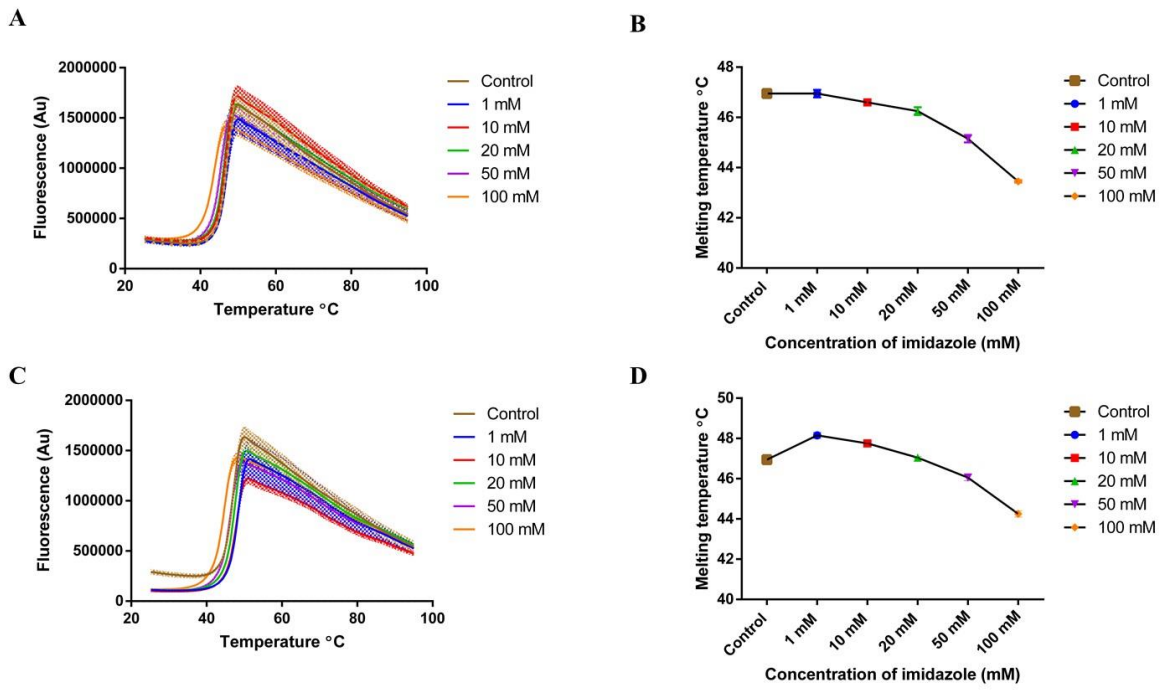

Figure 6.20. Thermal stability assay of MATα2 with Imidazole. **A**. Melting curves of MATα2 that was pre-incubated with imidazole (1-200 mM). **B**. Melting temperatures (°C) of MATα2 pre-incubated with imidazole (1-200 mM) (±range, n=2). **C**. Melting curves of MATα2 that was pre-incubated with imidazole (1-100 mM) and SAMe (5 mM). A destabilising effect can be seen with increasing imidazole concentrations **D**. Melting temperatures (°C) of MATα2 pre-incubated with imidazole (1-200 mM) and SAMe (5 mM) (±range). A destabilising effect can be seen with increasing imidazole concentrations which is slightly decreased by the presence of SAMe. A thermal ramping procedure (0.3 °C/min over a range of 25 °C to 94 °C) was employed to denature MATα2 and mean melting temperature (°C) (±range, n=2) are shown for each condition. Section 4.2.19 describes the differential scanning fluorimetry method. The melting temperature was calculated from the first derivative curve of the melting curve and was defined as the maximum peak of the first derivative curve (Sun et al. 2015). Figures were produced using GraphPad.

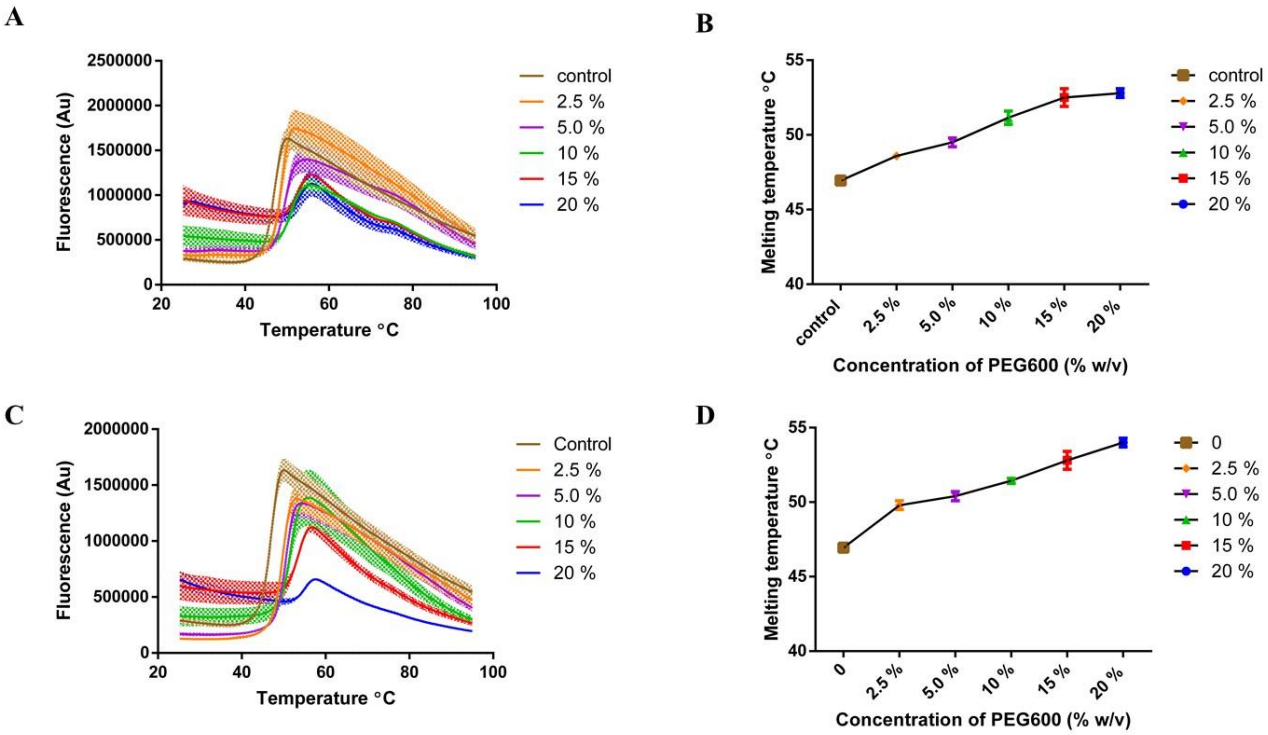

Figure 6.21. Thermal stability assay of MATα2 with PEG600. **A**. Melting curves of MATα2 that was pre-incubated with PEG600 (0-20 % w/v). **B**. Melting temperatures (°C) of MATα2 pre-incubated with PEG600 (0-20 % w/v) (±range, n=2). **C**. Melting curves of MATα2 that was pre-incubated with PEG600 (0-20 % w/v) and SAMe (5 mM). A stabilising effect can be seen with increasing PEG600 concentration **D**. Melting temperatures (°C) of MATα2 pre-incubated with PEG600 (0-20 % w/v) and SAMe (5 mM) (±range, n=2). A destabilising effect can be seen with increasing PEG600 concentration which is further enhanced by the presence of SAMe. A thermal ramping procedure (0.3 °C/min over a range of 25 °C to 94 °C) was employed to denature MATα2 and mean T_m values (±range) are shown for each condition. Section 4.2.19 describes the differential scanning fluorimetry method. The melting temperature was calculated from the first derivative curve of the melting curve and was defined as the maximum peak of the first derivative curve (Sun et al. 2015). Figures were produced using GraphPad.

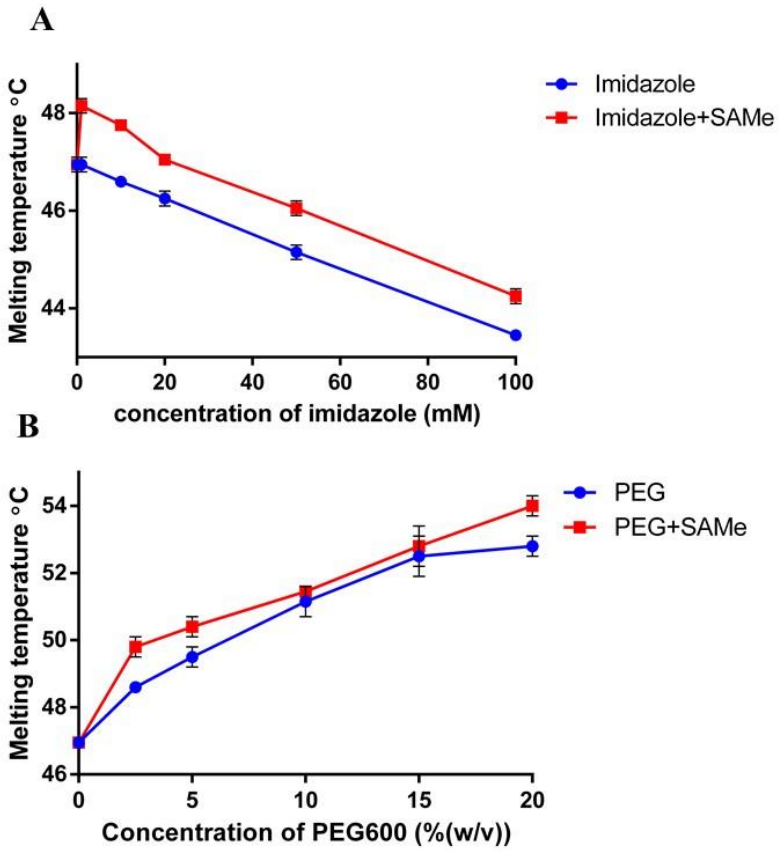

Figure 6.22. Thermal stability of MATα2 in the presence of SAMe, PEG600, and imidazole. **A**. Imidazole (1-100 mM) decreases the stability of MATα2 which is partially protected by incubation with SAMe (5 mM). **B**. PEG600 (2.5-20 % w/v) provided a stabilising effect for MATα2 that is concentration dependent. This can be further enhanced by SAMe (5 mM). Data is displayed with range (n=2). Section 4.2.19 describes the differential scanning fluorimetry method. The melting temperature was calculated from the first derivative curve of the melting curve and was defined as the maximum peak of the first derivative curve (Sun et al. 2015). Figures were produced using GraphPad.

6.2 Discussion

Even though there was a great effort to make a consensus nomenclature for mammalian MAT enzymes by Kotb *et al* in 1997 there still remains discussion regarding the oligomeric states of human MAT enzymes. MATα1 has long been regarded as being able to associate into both a dimer and tetramer (Varela-Rey et al. 2011), whilst on the other hand MATα2 has not been well characterized and no agreement has been reached as to the oligomeric state of this protein. In order to function as an enzyme, MATα2 must be at least a dimer and when MATαβ was purified from human chronic lymphocytic cells it was suggested that this was the case (Kotb & Kredich 1985). The crystallographic structure of MATα2 solved by Shafqat *et al.* shows that it is a dimer of dimers and in our previous publication we showed that MATα2 can function as a tetramer in the context of the $MAT(\alpha2)_4(\beta V2)_2$ complex (Murray et al. 2014). Using the PISA (Krissinel & Henrick 2007) server the protein assembly of the MATα2 structures, containing SAMe+ADN+MET+PPNP bound and PPNP bound were checked, which suggested that both dimers and tetramers are energetically favourable. For the reasons above MATα2 was used to investigate whether, in solution, it can assemble into dimers and/or tetramers.

Just like MATα1, gel filtration experiments of MATα2 showed multiple peaks (Fig 6.1A, B) but when molecular mass was determined using a standard curve of known protein standards these peaks did not correspond to dimers or tetramers. Gel filtration experiments suggested that the first peak would correspond to a trimeric species of ~124 kDa, which seemed very unlikely, as in order to function two MATα monomers must dimerize to create functional active sites. Due to the equilibrium between oligomeric states for both MATα1 and MATα2, SAXS-Size exclusion chromatography (SEC) was chosen to ensure that both peaks were separated enough to collect reliable data. For

MATα1 the first peak corresponded to a tetramer with a molecular weight of ~160 kDa whilst the second peak gave a molecular mass of ~93 kDa (Table 6.2). For MATα2 only data corresponding to a tetrameric mass were collected, as the protein became unstable from purification to experimentation and precipitated, limiting the amount for data collection. Further data collections would have allowed for the dimeric population to be measured. Considering the high sequence identity between MATα1 and MATα2 (84 %) it is not difficult to see that MATα2 can oligomerise in a similar way.

Of the two human MATβ structures deposited in the PDB, one is a monomer (2YDY) and the other is a dimer (2YDX) (Shafqat et al. 2013). Using both SAXS and gel filtration MATβV1 and MATβV2 correspond to monomers (Fig 6.2A 6.6A-D, Table 6.1 & Table 6.2). NADP binding reduced the flexibility of the MATβ proteins, as overall maximum particle distance (D_{max}) of the protein decreased compared to MATβ without NADP (Table 6.2). Although 2YDX is shown as a dimer, this dimerisation may only exist due to the co-crystallization ligand, resveratrol, which binds at the dimer interface. Resveratrol is a naturally occurring polyphenol commonly found in grapes, peanuts, and other plant sources, which has been reported to possess both chemotherapeutic and chemopreventive activities in certain types of cancer including liver cancer (Mbimba et al. 2012). When resveratrol was incubated with MATβ isoforms no shift was observed in the UV spectrum of gel filtration suggesting no oligomerisation occurred in solution. In fact, we investigated what function this molecule had in the context of the MATα2β complexes (section 5.1.13) and saw no interaction between the complex and resveratrol. It was reported that resveratrol induces the expression of *MAT2B* by multiple mechanisms and leads to the stabilisation of the protein, though resveratrol was not directly responsible for the stabilisation of MATβ proteins (Yang et al. 2013). The results from ITC and gel

filtration experiments showed that resveratrol does not bind to MATβ. Comparing the crystallographic structure of MATβ to the solution X-ray scattering data provided a good fit, which could be further improved by taking the missing N and C-terminal residues in the crystal structure into account (Fig. 6.6). Furthermore MATβ isoforms do not interact to form heterodimers, as both pull down and gel filtration experiments suggested that they do not interact with each other (Fig 6.8 & 6.9). The schematic in figure 6.23 provides a summary of the oligomeric states of the MAT family of proteins whereby the two catalytic isoforms, MATα1 and MATα2, can form both dimers and tetramers. The regulatory subunit, MATβ, shows no evidence of forming higher oligomeric forms and has only been seen as monomer.

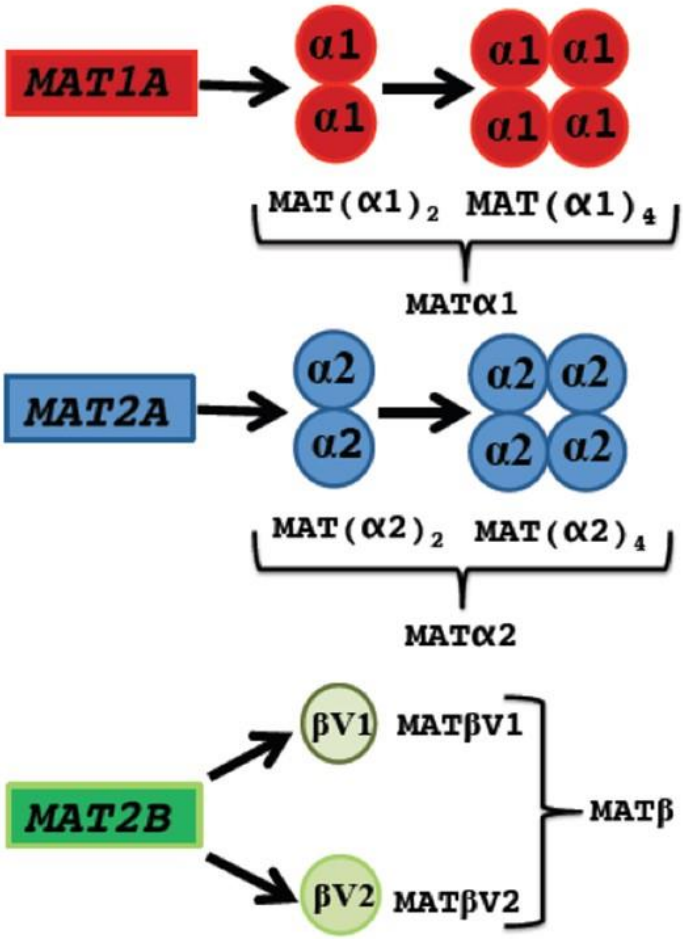

Figure 6.23. Summary of the oligomeric states of MATα1, MATα2 and MATβ proteins. The mammalian genes *MAT1A*, *MAT2A* and *MAT2B* encode the catalytic subunits MATα1, MATα2 and the regulatory subunit MATβ respectively. MATα1 and MATα2 subunits can be organised as dimers and tetramers and can be referred to as $MAT(\alpha1)_2$, $MAT(\alpha1)_4$, $MAT(\alpha2)_2$ and $MAT(\alpha2)_4$. The regulatory subunits MATβV1 and MATβV2 are monomeric in solution.

Despite a plethora of MAT structures that have been published (González et al. 2003; Komoto et al. 2004; Fua et al. 1996; Takusagawa et al. 1996; González et al. 2000) or deposited in the protein data bank, to date, only one human MATα2 (Shafqat et al. 2013) (1.2 Å resolution) and three MAT$(\alpha 2)_4(\beta V2)_2$ complexes (Murray et al. 2014) (2.35 to 3.3 Å resolution) have been published. Mechanisms of how MAT enzymes function have been described for *E.coli* MAT (González et al. 2003) which is the accepted mechanism for MAT enzymes. Here, several high resolution crystallographic structures are presented that support the *E.coli* MAT's reaction mechanism, but defining the movements of methionine within the active site during catalysis from its initial position within the active to its final position as part of SAMe. We see that methionine initially occupies the active site (Fig 6.11B, 6.16A) in a different manner to *E.coli* MAT (1P7L) (Fig 6.15B). A direct attack by the sulphur of methionine on the C5' atom of the ATP molecule was proposed to form SAMe (Parry et al. 1982), which was confirmed with kinetic isotope studies (Markham et al. 1987). With SAMe+ADO+MET+PPNP the initial position (Fig 6.11B) of methionine can be seen which is stabilized through hydrogen bonds between the main chain nitrogen of methionine with Asp258 and Glu70, conserved in mammals and *E.coli* (Fig 6.24) and through the terminal oxygen and Gln113. The distance between the methionine functional group and PPNP is large enough that would allow it to rotate into a position to attack the C5' atom of ATP without encountering steric hindrance.

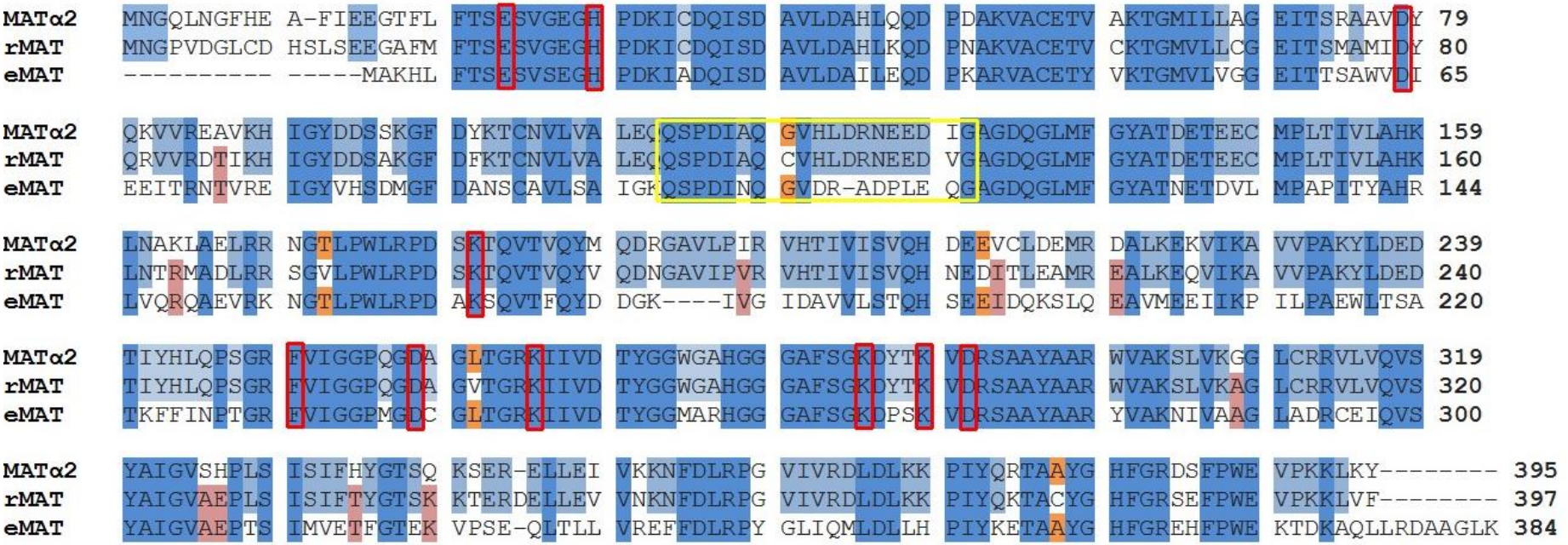

Figure 6.24 Sequence alignment of MATα2 from *homo sapiens*, rMAT from *rattus norvegicus* and eMAT from *E.coli*. Fully conserved residues between three sequences are shown in blue, conservation only between human and rat are in light blue, only between human and *E.coli* in orange and in pink for those conserved between *E.coli* and rat. Red boxes show residues involved in catalysis. Gating loop is highlighted by yellow box.

PPNP is a non-hydrolysable triple phosphate moiety that is particularly useful when studying MAT enzymes, as it cannot be cleaved to pyro and orthophosphate. The high resolution structure SAMe+ADO+MET+PPNP has shown not only the position of the PPNP, but also positions of magnesium and potassium ions within the active site (Fig 6.12 & 6.13) which were previously difficult to model. Interestingly, the positions of the three ions within the active site share high similarity to those in *E.coli* (Fig 6.13C). The high resolution data (SAMe+ADO+MET+PPNP bound) clearly show that all three active site ions are highly coordinated either hexacoordinated or pentacoordinated (Fig 6.12). Orientation of the PPNP is the same as in AMP-PNP from *E.coli*, which demonstrates that as the AMP-PNP is cleaved, the newly formed PPNP does not cause a change in orientation or change the residues that were initially anchoring the AMP-PNP in the active site. Refining $MAT(\alpha 2)_4(\beta V2)_2$ complex (4NDN) using the high resolution position of the PPNP satisfies electron density to a greater extent. PPNP is stabilized through interactions with magnesium and potassium ions, as well as a number of residues (Fig 6.12 & 6.13). These interaction would become destabilised upon the cleavage of the triple phosphate to pyro and orthophosphate, allowing them to exit the active site. Only a small structural movement within the region close to the active site is needed to aid in the anchoring of ATP initially and stabilization of the cleaved triple phosphate moiety in order to allow the triple phosphate to be broken down into pyro and orthophosphate. It is difficult to say whether this pre-apo structure (Fig 6.16E) is set up for hydrolysis, because the mechanisms of how MAT enzymes hydrolysize the triple phosphate remains unclear. From a recent MAT review by Pajares & Markham in 2012 they summarize this: "The mechanism of hydrolysis of the triple phosphate intermediate appears to be similar to that of the myriad of metal ion dependent ATPases and other phosphatases. However the chemical means by which water is activated to react with the triple phosphate and thus split the triple phosphate chain into pyrophosphate and orthophosphate is unknown".

The observation that the MATα2 active site can be occupied but have the gating loop disordered is new in human MAT enzymes and gives evidence for the exiting order of products from the active site (Fig 6.14). SAMe exits the active site upon the gating loop becoming disordered followed by the cleaved triple phosphate products, which causes residues (Fig. 6.14) such as Lys61, Glu70 and Lys181 to return to their unbound (apo) positions. The non-hydrolysable PPNP remains in the active site and position of the highlighted residues remains the same as if the active site was occupied and gating loop remains ordered and shut. An interesting observation is that in the PPNP bound MATα2 the position of the PPNP is different and this may represent a pre-apo structure whereby the central phosphate moves causing a change in ion coordination and allowing for the triple phosphatase activity of MATα2 to occur or if it could just represent the effect on having a non-hydrolysable PPNP in the active site and that its movement may only occur due to non-hydrolysable nature. The data suggests that ATP (PPNP) binding will change the conformation of Glu70 from open (Fig 6.14D) to substrate bound (Fig 6.14A-C), ready for methionine binding. From a comparison of the SAMe+ADO+MET+PPNP bound MATα2 to MAT$(\alpha 2)_4(\beta V2)_2$ complex it is clear that the presence of MATβ does not directly affect the catalytic site (Figure 16.4A, B). It was demonstrated before that recombinant MATαβ complexes have greater activity compared to MATα2 alone (Murray et al. 2014) yet it remains unclear as to how MATα2 activity is regulated by MATβ.

Although it has been known that human MATα2 can use ethionine as a substrate for many years (Smith & Salmon 1965) and the effect that this has within the cell (Rao et al. 1982), no human MATα2 structure has been published with this substrate or its product, S-adenosylethionine. Recently, Wang *et al* published the structure of sMAT (Wang et al. 2014)

from *Sulfolobus solfataricus* containing SAE within the active site and suggested that moderate clashes would occur between the ethyl group and Ile117 and Ile322 of human MATα2. In the human MATα2 SAE structure, containing S-adenosylethionine, this is not the case and the ethyl group is over 3 Å away from the surrounding residues. S-adenosylethionine, has the same orientation as SAMe (Fig 6.17B) with very little noticeable movements.

These results show that SAMe can leave the active site without the breakdown of the PPNP which has been presumed to provide the energy for opening of the gating loop. It is clear that the catalytic mechanism of SAMe formation in MATα2 is retained in the absence of the regulatory subunit MATβ as residues involved in catalysis are in similar positions (Fig 6.14). Both of these unexpected findings have wider implications on the formation and release of SAMe. The utilisation of the same catalytic site for substrate ethylene for production of S-adenosylethionine opens therapeutic possibilities for diseases connected to SAMe and SAE utilisation.

Differential scanning fluorimetry provides one of the most simple, yet effective, screening method to identify low-molecular-mass ligands that bind and 'protect' purified, folded, proteins from thermal denaturation as detected by fluorescence of a probe compound (Desantis et al. 2012; Mashalidis et al. 2013; Niesen et al. 2007). Reduction in the amount of MATα2β complex has shown a decrease in the rate of growth in human liver cancer cells (Wang et al. 2008) so targeting this complex could provide a treatment of various forms of liver diseases. The structure of MATαβ revealed a binding cavity, which is where MATβ interacts with MATα2 (Murray et al. 2014), and more recently crystal structures have revealed that other molecules, including imidazole and PEG600 can bind within this cavity (Section 6.1.11). Screening small molecules that could bind to MATα2, specifically the MATβ binding cavity

could provide a starting point in rational drug design and using a rapid and inexpensive screening method, such as differential scanning fluorimetry, seemed a suitable methodology to use.

Given that MATα2 exists in a dimer/tetramer equilibrium and before molecules were screened against it the effect of concentration of MATα2 was examined. This was done to see whether an effect of the melting temperature or the shape of the curve, suggesting different unfolding rates, would be observed. Neither of these were seen when the protein concentration was varied. Small changes in melting temperature were observed but this was not in a concentration dependent manner (Fig 6.19A). It is known that binding within the MATα2 active site causes several disordered loops to become well ordered (Shafqat et al. 2013; Murray et al. 2014) and that if the active site is mutated causing the inactivation of the protein then it cannot function then the cell dies which raised the question whether SAMe, the product, could be pre-bound to MATα2 before exposure to compounds. When searching for a compound to bind MATα2 we wanted to try and reduce the likelihood of compounds binding within the active site. Incubation with SAMe was used to try to lock the active site and prevent other molecules from binding in hope of reducing the number of hits that bind to MATα2. SAMe was able to stabilise MATα2 over a range of concentrations (2-10 mM) (Fig 6.19C). Knowing that ligands such as imidazole and PEG600 were able to bind to MATα2 the effect that they had on this protein were examined. Imidazole had a destabilising effect on MATα2 that could be protected against (up to 20 mM) by the addition of 5 mM SAMe (Fig 6.20). Imidazole is a small and planar molecule, which may be able to fit in and interact with sites on MATα2 monomer or between the tetrameric interfaces of the higher oligomeric state that was not revealed by the crystal structure which caused destabilization effect. PEG600 on the other hand had a stabilising effect on

MATα2 which may be due to a higher affinity it may have with MATα2 compared to imidazole (Fig 6.21).

Chapter VII. Interacting partners of MATβ: MATα1 and GIT1

7.1 Results-MATα1 and MATβ

7.1.1 Interaction of MATβ isoforms with MATα1.

MATα1 is the liver specific isoform of MAT in humans and looking at the structure of the $MAT(\alpha2)_4(\beta V2)_2$ complex it is difficult to understand why, *in vivo*, MATβ has so far been found to only interact with MATα2 (Martínez-Chantar et al. 2003; Kotb & Kredich 1985; Huang et al. 1998), even though MATα1 and MATα2 have a sequence identity of 84 % and share the same fold. The differences between these two isoforms are not restricted to a particular part of the protein but rather spread out throughout the sequence, (Fig 7.1). All catalytic residue and those residues involved with binding MATβ (green) are conserved, suggesting that the catalytic mechanism is the same for MATα1 and MATα2. The largest stretch of amino acid substitutions are within the N-terminus of the MATα proteins. The observation that the residues involved in MATβ binding are conserved raised the question if MATβ could bind to MATα1. For MATβ proteins, the two major isoforms, MATβV1 and MATβV2 share over 94 % sequence identity and the interaction region with MATα2 is conserved and the only differences arise in the N-terminal region of MATβ (Fig 7.2).

Looking at the surface charge representation of the MAT catalytic subunits no major differences can account for MATα1 apparent lack of interaction with MATβ (Fig 7.3A-B). The area of interaction with MATβ (Fig 7.3C) as shown by sequence alignment shows that overall fold of the MATβ interaction area is equal for both MATα1 and MATα2.

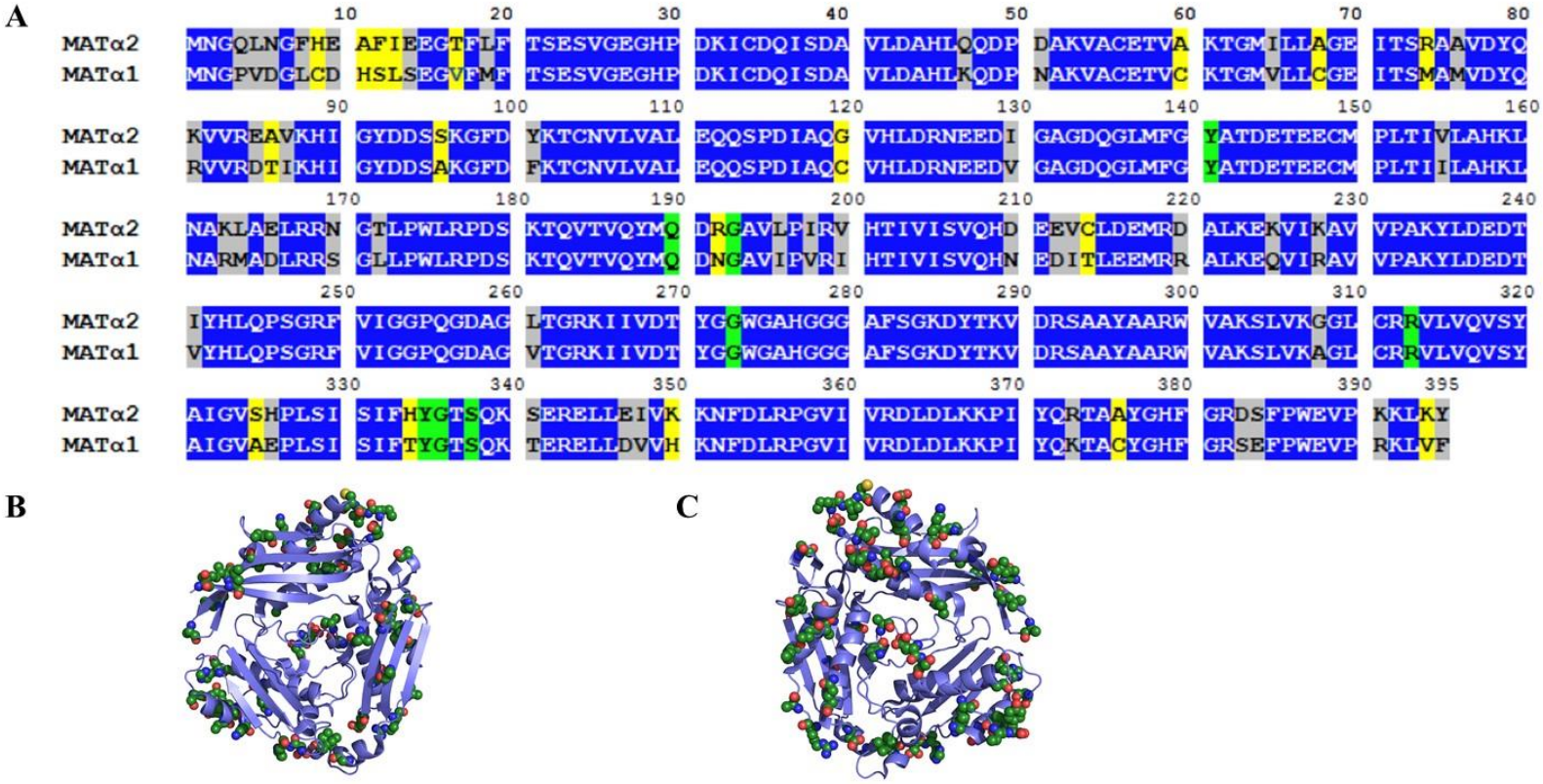

Figure 7.1 Sequence alignment of the two catalytic human MATα subunits. **A**. In blue are residues that are the same in both isoforms. In green are the residues of MATα2 that interact with MATβV2, which are also equal in MATα1. In grey are the residues whose side chains are similar in biochemical properties and in yellow the residues whose side chains are different in biochemical properties. Proteins were aligned using clustalW server (Thompson et al. 1994). **B**. Position of all residues that a different between MATα1 and MATα2 located on the structure of MATα2 (2P02). The view of the MATα2 proteins shows the side used to form the MAT$(\alpha 2)_2$. **C**. As in B but rotated 180° around and the view of the MATα2 proteins shows the side that would be exposed to the cell. MATα2 is shown in cartoon form in slate, whilst highlighted amino acids are shown as spheres.

Figure 7.2. Sequence alignment of MATβV1 and MATβV2. A sequence comparison of the regulatory MAT proteins, MATβV1 and MATβV2. Residues that interact with MATα dimer are highlighted by a blue line and the NADP binding motif is also shown as a green box. Proteins were aligned using clustalW server (Thompson et al. 1994).

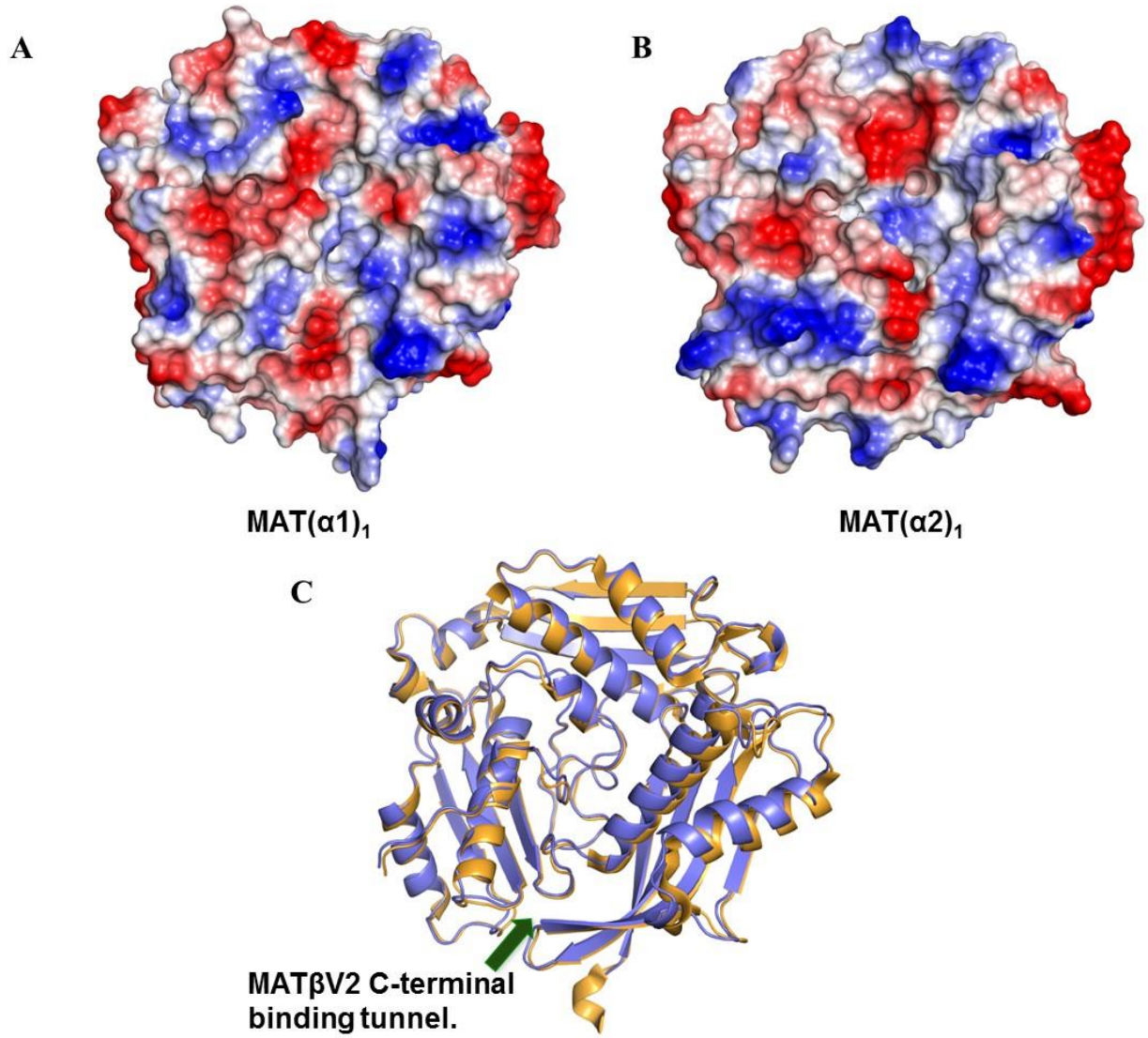

Figure 7.3. Structural comparison of MATα proteins. Electrostatic surface representation of **A**. MATα1 (PDB: 20BV) and **B**. MATα2 (PDB: 2P02) shows different charge distribution at the surface of both proteins. Negative charge is shown in red colour, positive in blue. The surface potential has been calculated with Pymol/APBS (Schrödinger, LLC 2010). **C**. Superposition of the MATα1 monomers (PDB: 2OBV in orange) and MATα2 from the $MAT(\alpha2)_4(\beta V2)_2$ complex (slate, PDB 4NDN), the green arrow represents the region where MATβ C-terminal interacts.

7.1.2 MATα1 interacts with MATβ isoforms

MATα1 was incubated with MATβV1 for one hour at 4 °C (Section 4.2.13) before injection into a gel filtration column and evidence for complex formation can be seen (Fig 7.4A) with a peak around 12 mL (orange) which is typical for MATαβ complexes. UV spectra for MATβV1 (green) and MATα1 (purple) eluted at different volumes to the complex suggesting different masses. Excess MATβV1 that was no incorporated into the $MAT(\alpha1)_4(\beta V1)_2$ complex can be seen in orange at 16 mL. Running fractions of the $MAT(\alpha1)_4(\beta V1)_2$ complex on SDS gel show that both proteins are present in similar amounts for the complex peak (lanes 3-8). Lanes 9-15 show both proteins are in different levels suggesting that some of the MATα1 and MATβV1 protein did not form a complex (Fig 7.4B).

Incubation with the MATβV2 yielded different results compared to MATβV1. The majority of the two proteins, MATα1 and MATβV2, did not form a complex (blue) and were eluted at similar volumes to the elution profiles of MATα1 (~14.2 mL, purple) and MATβV2 (~16 mL, red) (Fig 7.5A) alone. A small peak appeared at 12 mL which is typical of MATαβ complexes although this represents a very small proportion of the proteins. Analysing fractions from this gel filtration, 13 mL-17 mL, on SDS page shows that the proteins here are not interacting and do not co-elute as the maximum level of protein is different for each protein (Fig 7.5B).

To confirm interaction of MATα1 and MATβV1 ITC was performed (Fig 7.6A) which showed that the two proteins interacted and Table 7.1 summarises the thermodynamic properties of

their interaction. Compared to MAT$(\alpha 2)_4(\beta V1)_2$ complexes (K_d= (3.38 ±0.98)X 10^{-8}, Table 5.6) MAT$(\alpha 1)_4(\beta V1)_2$ had a much weaker interaction together (K_d= (2.96 ±0.21)X 10^{-6}, Table 7.1). A strong interaction by gel filtration between MATβV2 with MATα1 was not seen (Fig. 7.5A), indicating a role of the N-terminus of MATβ in providing the stability to MATα1β complex. Using the truncated version of MATβV1, lacking the first 16 amino acids (MATβV1Δ16), complex formation was tested and showed that this mutant produced much smaller amount of complex with MATα1 (Fig 7.6B). Similar results were seen using MATβ1Δ16, as when using MATβV2, with most of the two proteins not forming a complex and eluting separately. A small population of protein could be seen at 12 mL indicating the formation of the MAT$(\alpha 2)_4(\beta V1\Delta 16)_2$ complex.

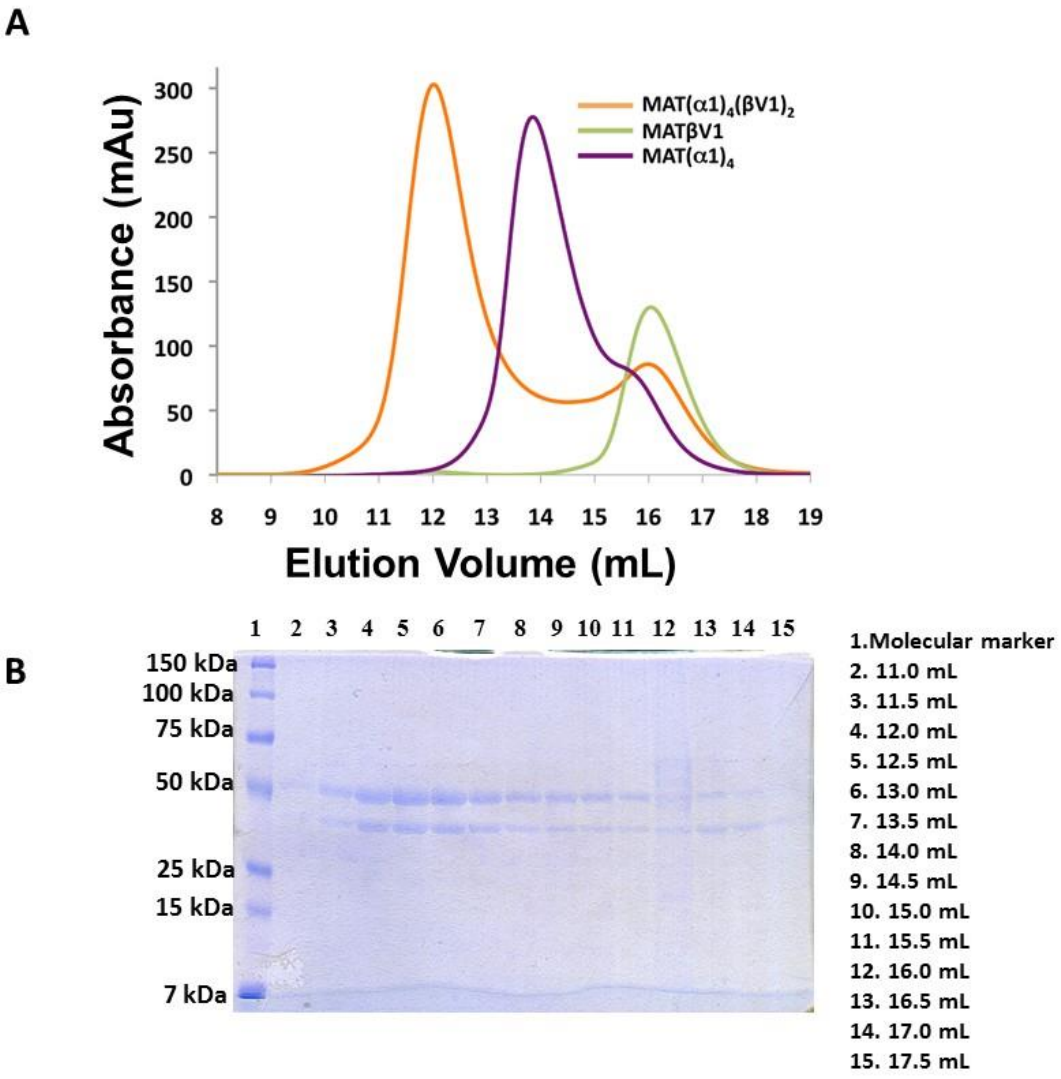

Figure 7.4. Formation of MAT(α1)$_4$(βV1)$_2$ complex analysed by gel filtration chromatography. **A**. Elution profiles of MATα1, MATβV1 and MAT(α1)$_4$(βV1)$_2$. 100 µM of MATα1 was incubated with 50 µM of MATβV1 for one hour at 4 °C (section 4.2.13) prior to injection onto a Superdex 200 10/300. For the controls, 100 µM of MATα1 and 50 µM MATβV1 were used. The elution buffer was 25 mM HEPES pH 7.5, 200 mM NaCl and 1 mM TCEP. Absorbance at 280 nm was used to measure for the presence of proteins. **B**. SDS PAGE of fractions of MAT(α1)$_4$(βV1)$_2$ from 11.0 mL to 17.5 mL.

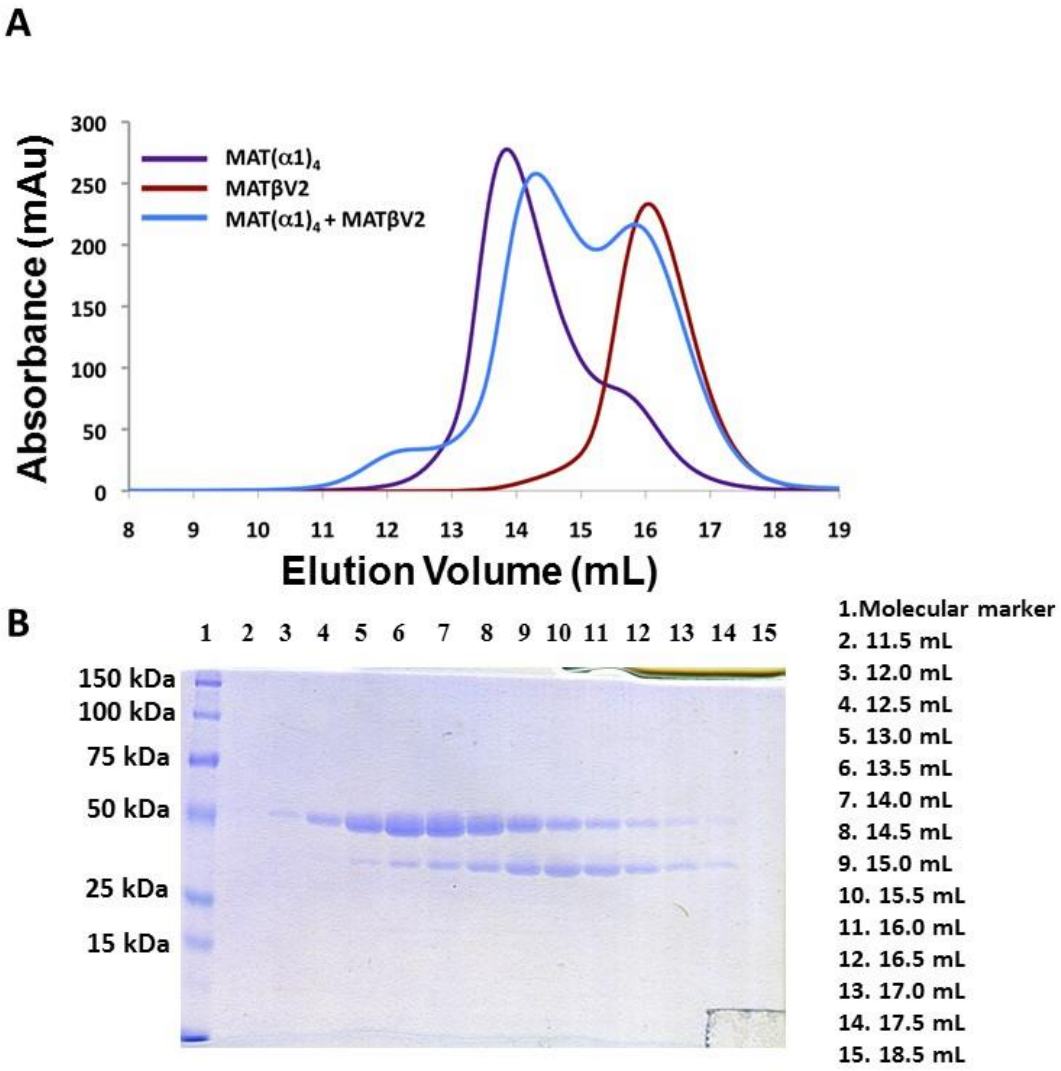

Figure 7.5. Formation of MAT(α1)$_4$(βV2)$_2$ complex analysed by gel filtration chromatography. **A**. Elution profiles of MATα1, MATβV2 and MAT(α1)$_4$(βV2)$_2$. 100 μM of MATα1 was incubated with 50 μM of MATβV2 for one hour at 4 °C (section 4.2.13) prior to injection onto a Superdex 200 10/300. For the controls, 100 μM of MATα1 and 50 μM MATβV2 were used. The elution buffer was 25 mM HEPES pH 7.5, 200 mM NaCl and 1 mM TCEP. Absorbance at 280 nm was used to measure for the presence of proteins. **B**. SDS gel of fractions of MAT(α1)$_4$(βV1)$_2$ from 11.5 mL to 18.5 mL.

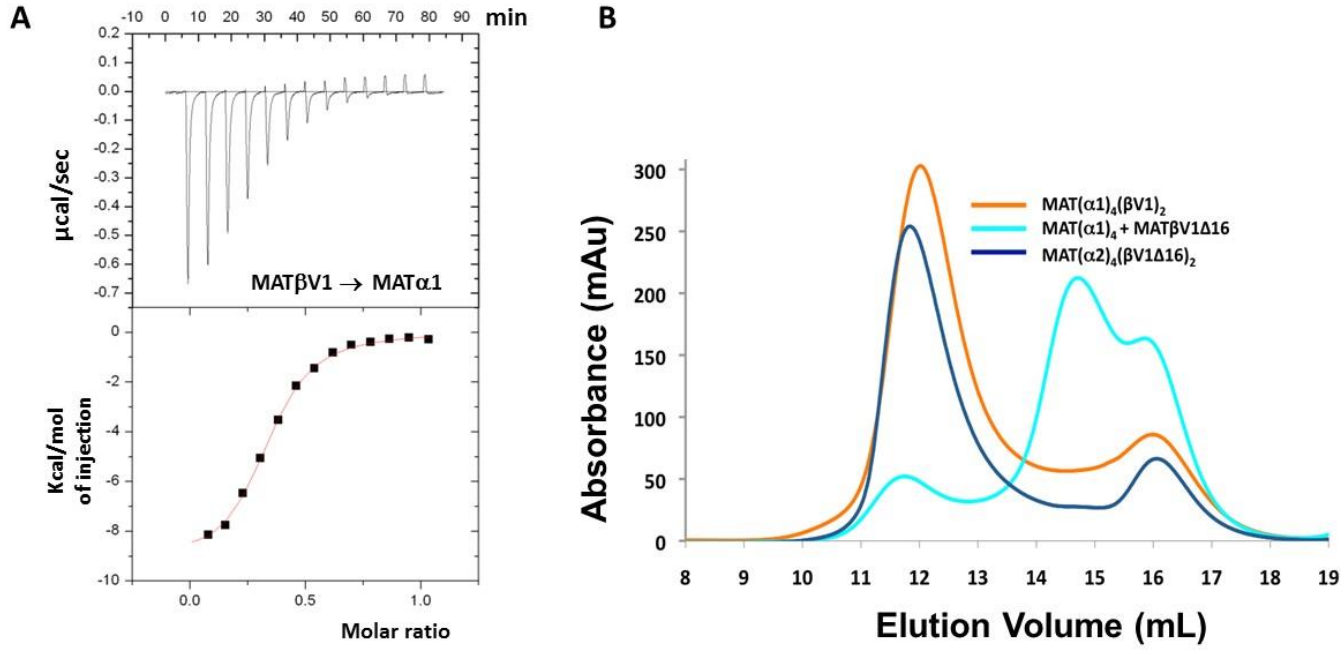

Figure 7.6. Interaction of MATα1, MATβV1 and MATβV1Δ16. A. Isothermal titration calorimetry of MATα1 with MATβV1. The top graph represent the differential heat released during the titration of MATβV1 (236.8 µM) with MATα1 (45.1 µM). The bottom graph represents the fitted binding isotherm (section 4.2.21). **B**. Formation of $MAT(\alpha1)_4(\beta V1\Delta16)_2$ complex analysed by gel filtration chromatography. Elution profiles of MATα1, MATβV1Δ16 and $MAT(\alpha1)_4(\beta V1\Delta16)_2$. 100 µM of MATα1 was incubated with 50 µM of MATβV1Δ16 for one hour at 4 °C (section 4.2.13) prior to injection onto a Superdex 200 10/300. For the controls, 100 µM of MATα1 and 50 µM MATβV1Δ16 were used. The elution buffer was 25 mM HEPES pH 7.5, 200 mM NaCl and 1 mM TCEP. Absorbance at 280 nm was used to measure for the presence of proteins.

Table 7.1 Thermodynamic parameters of MAT$(\alpha 1)_4(\beta V1)_2$ complex formation. K_a is calculated from as $1/K_d$ and is the binding constant (M^{-1}). The K_d represents the affinity of the reaction (M). ΔH binding enthalpy of the interaction in kcal/mol. N is the stoichiometry indicating the ratio of the injected macromolecule-cell macromolecule binding.

	MATβV1 → MATα1
[]	236.8 μM → 45.1 μM
$K_a(M^{-1})$	$(3.38 \pm 0.24) \times 10^5$
$K_d(M)$	$(2.96 \pm 0.21) \times 10^{-6}$
ΔH kcal/mol	-9301 ± 140.9
N	0.325 ± 0.004

7.1.3 Mutations in the gating loop increase formation of MATα1βV2 complexes.

The N-terminus of MATβ plays a role in the regulation of MATα2 increasing the overall turnover of the enzyme (Fig 5.15), but it is not clear how this occurs. One possibility is that the N-terminus is able to interact with the gating loop of the enzyme increasing the rate of turnover. Comparing the gating loop of MATα1 and MATα2 only one residue differs, which is Cys120 in MATα1 and Gly120 in MATα2. Two mutants were produced, MATα1C120A and MATα1C120G to see if this point mutation was sufficient enough to allow formation of a stable complex with MATβV2. It was also examined if the presence of a reducing agent, TCEP affected the complex formation involving MATα1.

In the presence of the reducing agent TCEP, the two mutants of MATα1 (orange and grey) were able to form complexes with MATβV1 (Fig 7.7A) and showed no difference compared to wild-type MATα1 (blue). For MATβV2, the MATα1 mutants did not cause the majority of the proteins to form a complex, but rather MATα1 and MATβV2 remained as separate proteins (Fig 7.7B). The MATα1C120A mutant (red) produced a higher level of complex than either the wild-type MATα1 (blue) or the MATα1C120G mutant (green). After dialysis overnight to remove the reducing agent TCEP, a similar pattern was observed with regards to formation of MATα1β complexes. For MATβV1 both mutants were able to form MATα1β complexes just like the wild-type (Fig 7.8A). However, for MATβ2 the majority of the proteins eluted as separate proteins with a small population of protein eluting as a complex. Again the MATα1C120A mutant produced a larger proportion of complex than the wild-type or the MATα1C120G mutant (Fig 7.8B).

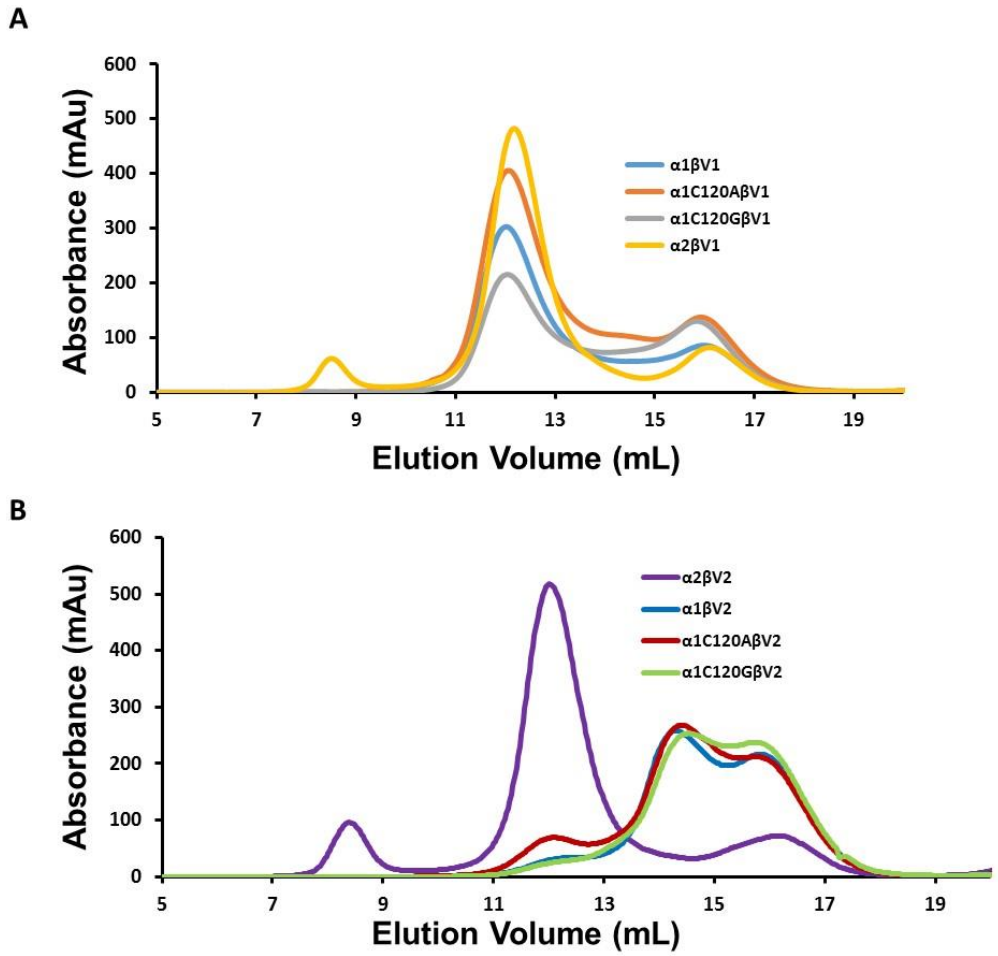

Figure 7.7. The effect of the presence of the reducing agent, TCEP, on complex formation of wild type and mutant MATα1β complexes. **A**. Elution profiles of MAT$(\alpha1)_4(\beta V1)_2$, MAT$(\alpha1C120A)_4(\beta V1)_2$, MAT$(\alpha1C120G)_4(\beta V1)_2$ and MAT$(\alpha2)_4(\beta V1)_2$. **B**. Elution profiles of MAT$(\alpha1)_4(\beta V2)_2$, MAT$(\alpha1C120A)_4(\beta V2)_2$, MAT$(\alpha1C120G)_4(\beta V2)_2$ and MAT$(\alpha2)_4(\beta V2)_2$ For all reactions, 100 μM of the catalytic MATα protein (MATα1, MATα1C120A, MATαC120G or MATα2) were incubated with 50 μM of either, MATβV1 or MATβV2 for one hour at 4 °C (section 4.2.13) prior to injection onto a Superdex 200 10/300. The elution buffer was 25 mM HEPES pH 7.5, 200 mM NaCl and 1 mM TCEP. Absorbance at 280 nm was used to measure for the presence of proteins.

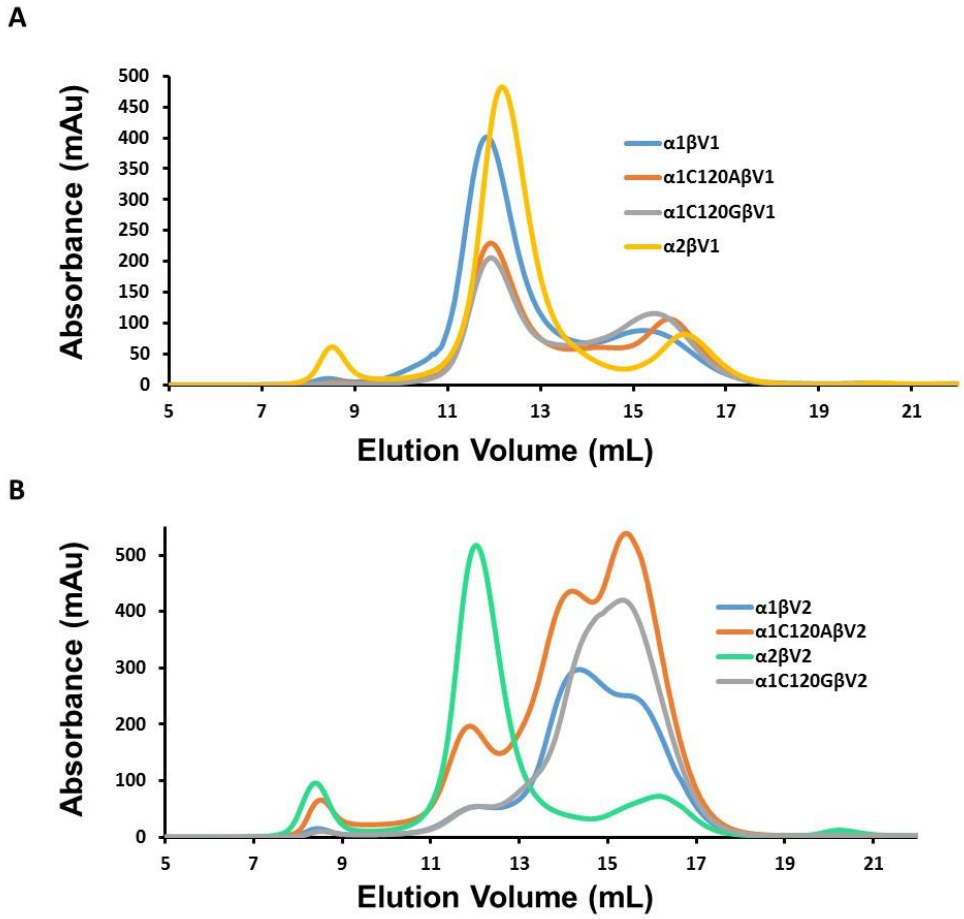

Figure 7.8. The effect of the absence of the reducing agent, TCEP, on complex formation of wild type and mutant MATα1β complexes. **A**. Elution profiles of MAT$(\alpha1)_4(\beta V1)_2$, MAT$(\alpha1C120A)_4(\beta V1)_2$, MAT$(\alpha1C120G)_4(\beta V1)_2$ and MAT$(\alpha2)_4(\beta V1)_2$. **B**. Elution profiles of MAT$(\alpha1)_4(\beta V2)_2$, MAT$(\alpha1C120A)_4(\beta V2)_2$, MAT$(\alpha1C120G)_4(\beta V2)_2$ and MAT$(\alpha2)_4(\beta V2)_2$ Proteins were dialysed overnight into 25 mM HEPES pH 7.5, 200 mM NaCl to remove the reducing agent TCEP. For all reactions, 100 μM of the catalytic MATα protein (MATα1, MATα1C120A, MATαC120G or MATα2) were incubated with 50 μM of either, MATβV1 or MATβV2 for one hour at 4 °C (section 4.2.13) prior to injection onto a Superdex 200 10/300. The elution buffer was 25 mM HEPES pH 7.5, 200 mM NaCl. Absorbance at 280 nm was used to measure for the presence of proteins.

7.1.4 Attempted crystallisation of MAT(α1)$_4$(βV1)$_2$

In order to try and crystallise MAT(α1)$_4$(βV1)$_2$ crystallisation trials were set up using a number of commercial screens. Several different conditions produced crystals including 0.1 M sodium HEPES pH 7.5, 0.06 M $MgCl_2$+ $CaCl_2$, 37.5% MPD, PEG 1K and PEG3350 (Fig 7.9A), 0.05 M HEPES Sodium pH 7.0, 0.005 M magnesium chloride hexahydrate and 25 % (v/v) PEG550 (Fig 7.9B), 0.01 M zinc chloride, 0.1 M MES pH 6.0 and 20 % (v/v) PEG 6000 (Fig 7.9C). Although multiple crystals were produced all were of MATα1 alone and not in complex with MATβV1. For the aforementioned crystals a protein concentration of 6.0 mg/mL was used and crystals were grown at 25 °C and appeared after 1 day.

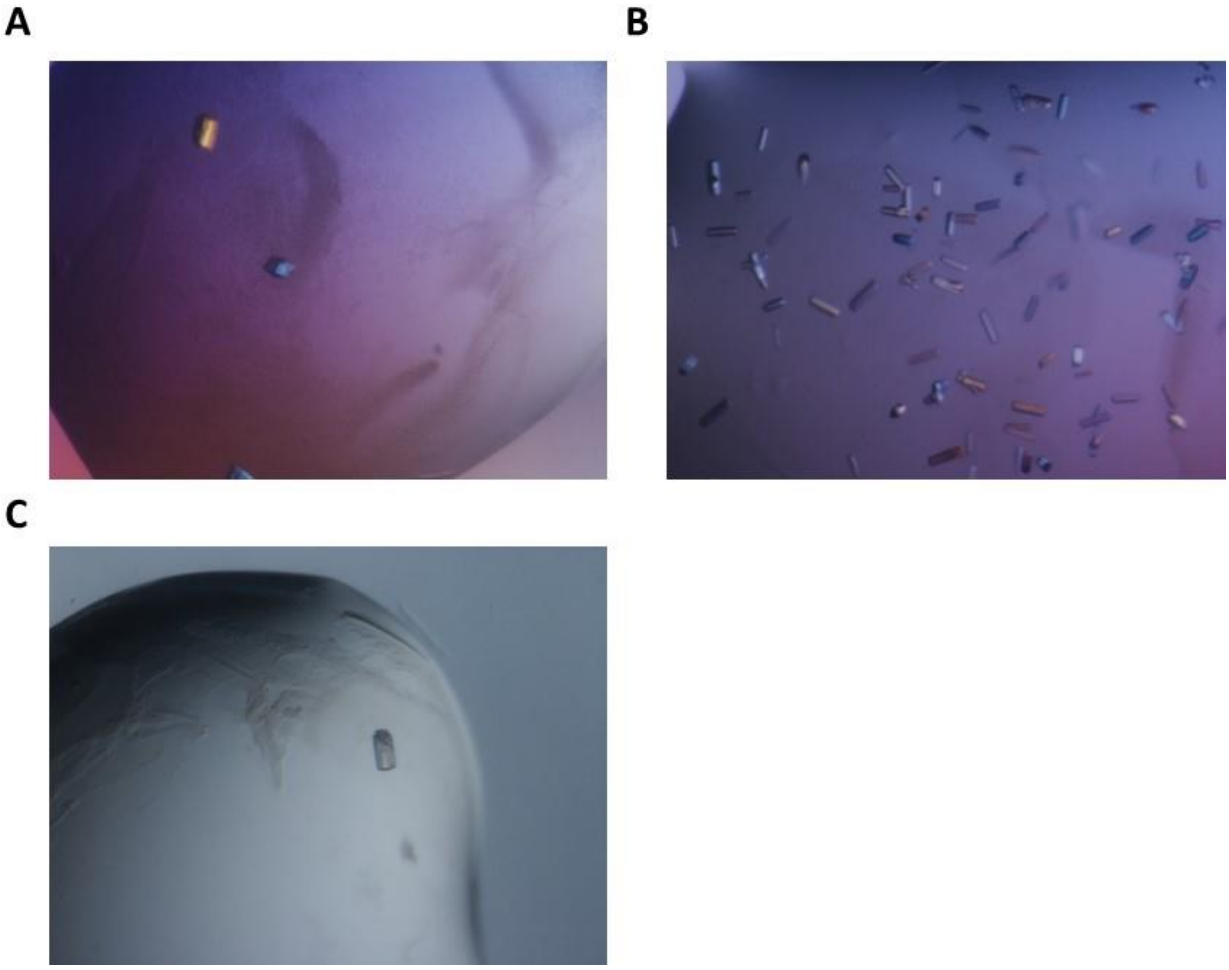

Figure 7.9. Pictures of MATα1 crystals that appeared in wells contains of MAT(α1)$_4$(βV2)$_1$. A. 0.1 M Sodium HEPES pH 7.5, 0.06 M $MgCl_2$+ $CaCl_2$, 37.5% MPD, PEG 1K and PEG3350. **B**. 0.05 M HEPES Sodium pH 7.0, 0.005 M magnesium chloride hexahydrate and 25 % (v/v) PEG550. **C**. 0.01 M zinc chloride, 0.1 M MES pH 6.0 and 20 % (v/v) PEG 6000.

7.2 Discussion

Looking at the structure of the MAT$(\alpha2)_4(\beta V2)_2$ complex it is difficult to understand why, *in vivo*, MATβ has so far been found to only interact with MATα2, even though MATα1 and MATα2 have a sequence identity of 84 % and share the same folding. A plausible explanation may be that in *in vivo* when the expression of *MAT1A* is switched on *MAT2A* and *MAT2B* expression is switched off and that MATα1 and MATβ are not expressed in sufficient amounts to interact. In adult hepatocytes *MAT1A* is highly expressed and the expression of *MAT2A* and *MAT2B* is low. In contrast during hepatocellular carcinoma, where the expression of *MAT2A* and *MAT2B* is turned on, as well as the expression of other proteins that interact with MATβ such as HuR, GIT1, MEK, ERK or MafK (Xia et al. 2010; Peng et al. 2014; Katoh et al. 2011), *MAT1A* expression is low or absent (Lu & Mato 2012). Accordingly, the formation of the MAT$(\alpha1)_4(\beta V1)_2$ complex upon incubation of MATβV1 with MATα1 was observed through gel filtration and ITC (Fig. 7.4; 7.6A; Table 7.1). The interaction between MATα1 and MATβV1 is a lot weaker than the interaction of MATα2 with MATβV1 suggesting the differences between MATα1 and MATα2 are sufficient to affect the formation of the MATαβ complexes. Also, this would support the lack of MATα1β complexes *in vivo* if MATβ preferentially binds MATα2. Remarkably, we did not observe a strong interaction by gel filtration between MATβV2 with MATα1 (Fig. 7.5), indicating a role of the N-terminus of MATβ in providing the stability of the MATα1β complex. MATβV1 and MATβV2 share high sequence identity (95 %) and only differ in the N-terminal region (Fig 7.2), and within this region there is no consensus or similarity between the isoforms. MATβV1 has a longer N-terminal region (21 residues) than MATβV2 (10 residues).

To confirm this hypothesis, a truncated version of MATβV1 lacking the first 16 amino acids (MATβV1Δ16) was generated. This mutant produces much smaller amount of complex with MATα1, supporting the hypothesis that the MATβ N-terminus is indeed important for the formation of stable MATα1β complexes (Fig. 7.6B). This is different from the interaction with MATα2, where the deletion of the MATβV1 terminus did not reduce the formation of the MATαβ complex.

Closer examination of the sequence of human catalytic MATs revealed that the differences between these two proteins were spread out throughout the whole sequence (Fig 7.1A). Mapping these differences onto the crystal structure of MATα2 (2P02) reveals that these mutations are spread out over the protein, with the majority of them being on the exposed surface rather than the surface that is used to form the dimer (Fig 7.1B-C).Interestingly, a residue in the gating loop, at position 120, is different between the proteins being a cysteine in MATα1 and glycine in MATα2. The structure of these two amino acids differ with cysteine having a larger functional group containing a methyl and a thiol compared to glycine, whose sidechain is a hydrogen. As it was shown that the N-terminus has a role in the activity of MATαβ it was hypothesised that the N-terminal could have a direct role in the interaction with the gating loop and if this was the case then mutating residue C120 of MATα1 may stabilise the formation of the complex. Mutating this cysteine residue to glycine, as it is in MATα2, did not have an effect on the amount of complex that was formed with either MATβ isoform (Fig 7.7). When Cys120 was mutated to an Ala120 the amount of $MAT(\alpha1)_4(\beta V2)_2$ formed was increased (Fig 7.7B) supporting the idea that the MATβ N-terminus interacts with the gating

loop promoting stability of the complexes and that the small differences in the lengths of the MATβ isoforms is sufficient to effect the formation of the MATαβ.

As the $MAT(\alpha1)_4(\beta V1)_2$ complex was able to be formed via gel filtration crystallisation of this protein complex was attempted. A wide range of commercial crystallisation screens were available to test but in order to try and reduce the number that were initially tried poly-ethylene glycol (PEG) based screens were tried first. Analysis of all the available crystal structures of protein-protein complexes by Radaev and Sun revealed that protein-protein complexes seem to favour PEG based solutions. In fact 71 % of protein-protein complexes crystallised have PEG based conditions (Radaev & Sun 2002; Radaev et al. 2006). Also, through ITC it was shown that MATα1 had a lower affinity for MATβV1 compared to the affinity MATα2 had for MATβV1 or MATβV2 (Fig 5.11, 7.6A, Table 5.6 & 7.1) and PEG based screens seemed a suitable choice as lower affinity complexes seem to favour PEGs over, salts and organic solvents.

Several conditions, which all contained various PEGs, produced protein crystals but upon testing them all were revealed to be of MATα1 only. After this, other screens which were not PEG based were tried and did not result in crystals, also protein concentration and temperature were varied but no crystals containing the complex $MAT(\alpha1)_4(\beta V1)_2$ were ever obtained. From the structure of the $MAT(\alpha2)_4(\beta V2)_2$ structures the N-terminal region of MATβV2 was disordered highlighting the flexible nature and also, to note, no crystals of $MAT(\alpha2)_4(\beta V1)_2$ were obtained suggesting that the moving N-terminus of the MATβV1 may be effecting the

likelihood of crystal formation. Crystallisation of MAT$(\alpha 1)_4(\beta V2)_2$ was not tried as only a small amount of protein formed complex upon incubation together showing a very unstable and weak interaction. It was interpreted that this protein complex would have destabilised in solution and not crystallised as a complex.

7.3 Results-Interaction of GIT1 and MATβ

7.3.1 Purification of full length GIT1 and analysis by trypsin digest.

Full length human GIT1 was cloned into the pet_SUMO vector as described in section 4.2.8 and was expressed in *E.coli*. Full length GIT1 was produced and was present in the soluble fraction after cell lysis. It was purified by nickel affinity chromatography, but appeared smaller than the predicted molecular weight of 107 kDa (with a sumo and histidine tag) and after incubation with SENP2, the protease for the SUMO tag, caused no change in migration on SDS-PAGE (Fig 7.10A). Different concentrations of full length GIT1 were run on SDS gel and selected bands were analysed for protein identification by peptide mass fingerprinting, which was carried out by the proteomic platform of CIC bioGUNE. Two bands, highlighted by red boxes were tested and both corresponded to GIT1 protein but of different lengths. Band 1, which showed low levels was of full length GIT1 whilst band 2 corresponded to GIT1 which had degradation from the C-terminus (661 amino acids).

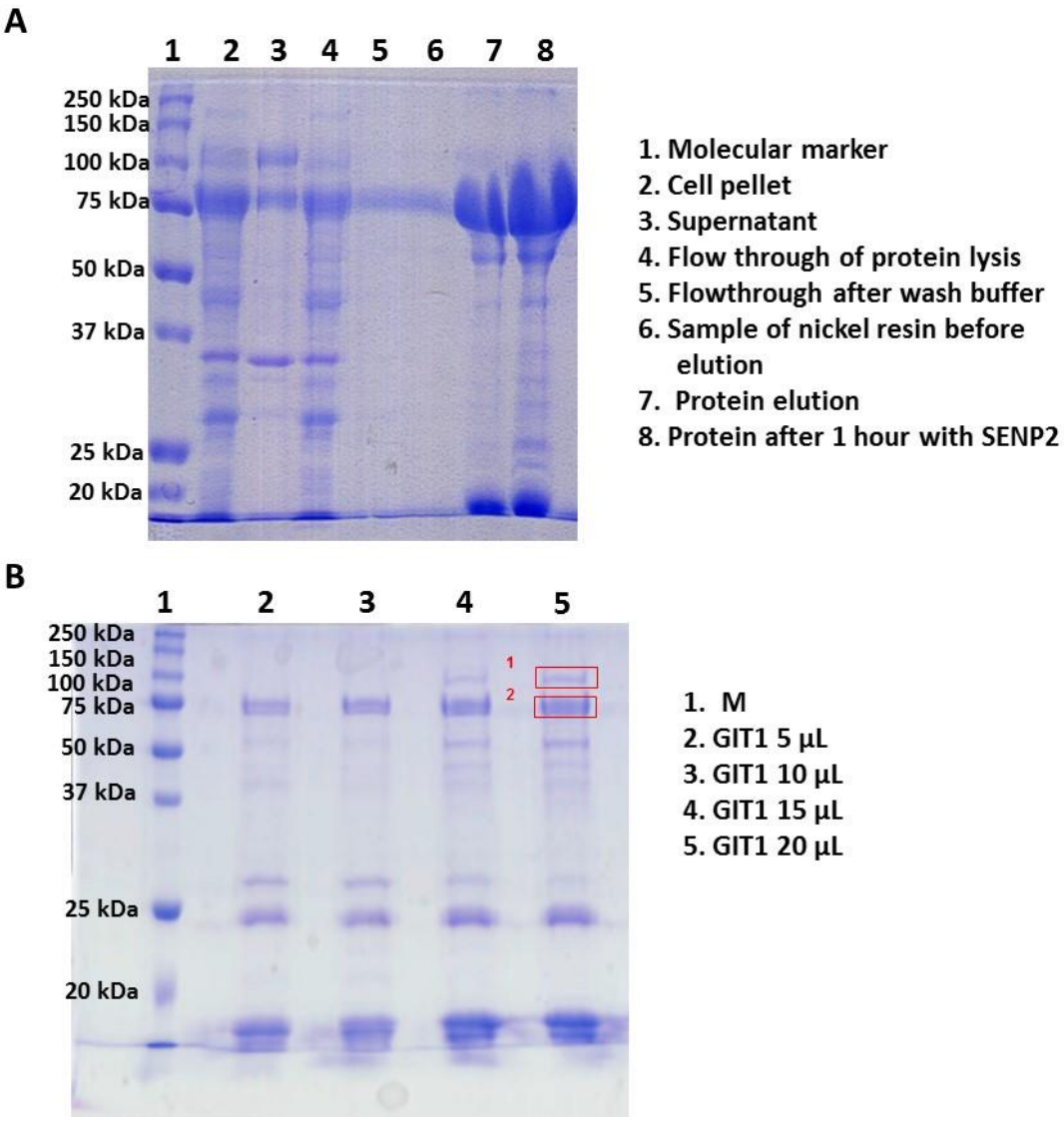

Figure 7.10. SDS gels of full length GIT1 protein. A. SDS-PAGE showing cell lysis and purification by nickel affinity chromotography of full-length human GIT1. **B**. SDS-PAGE showing GIT1 purified after incubation with the protease SENP2 loaded with different volumes (5-20 µL). Red boxes show bands that were analysed by peptide mass fingerprinting by the proteomic platform at CIC bioGUNE. Band 1 corresponded to the full length GIT1 (761 amino acids) whilst band 2 corresponded to a 661 amino acid protein that had degraded from the c-terminal.

7.3.2 Production and purification of GIT1_P491stop and GIT1_R413stop.

From analysis by peptide mass fingerprinting and looking at the predicted domain stuctures of full length GIT1 (Fig 1.9) two mutants were produced in which a stop codon was introduced after a certain residue: GIT1_P491stop and GIT1_R413stop. Again, like full length GIT1 both shorter constructs were expressed in *E.coli* and were soluble upon cell lysis. GIT1_P491stop after incubation with SENP2 shows a shift in size upon SDS-PAGE suggesting that the SUMO tag has been removed (Fig 7.11A), but after overnight dialysis and injection into an ion-exchange coloumn GIT1_P491stop had degraded which may have occurred due to a host protease and as a result it could not be seen on the SDS-PAGE (Fig 7.11B, C). After overnight dialysis and injection into an ion-exchange coloumn GIT1_R413stop showed to be stable and produced a narrow peak (Fig 7.12A). When fractions corresponding to this peak were run on a SDS-PAGEl it showed the presence of GIT1_R413stop protein (Fig 7.12B). These fractions were then concentrated and injected into gel filtration coloumn and produced one peak which corresponded to a molecular weight of ~49 kDa (Fig 7.12C, D), which was the expected mass of GIT1R413_stop.

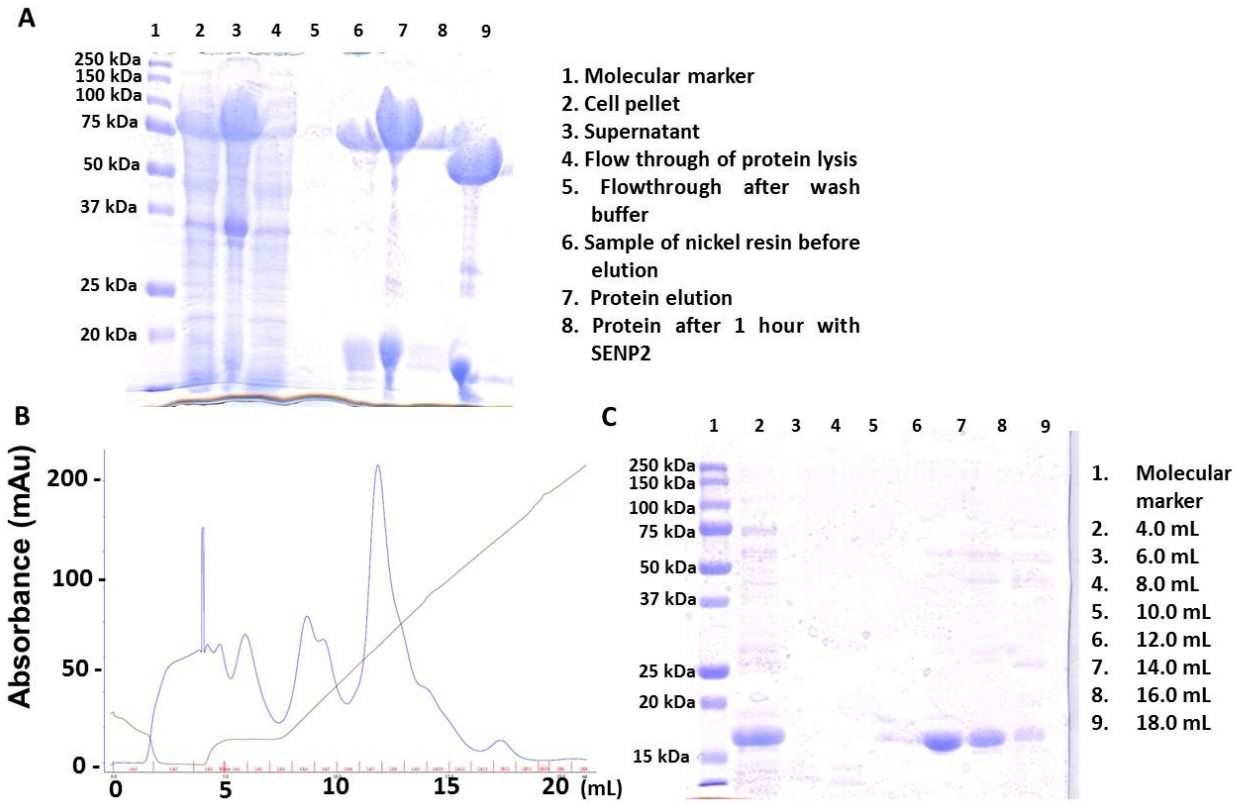

Figure 7.11. Purification of GIT1_P491stop. **A.** SDS-PAGE of cell lysis, nickel affinity chromotography and SENP2 digest of GIT1_P491stop (section 4.2.14). **B**. UV trace of ion-exchange of GIT1_P491stop after overnight dialysis into 25 mM Tris pH 8.0, 50 mM NaCl, 10 mM BME whereby GIT1_P491stop does not seem to present. **C**. SDS-PAGEof fractions from ion exhange showing that GIT1_P491stop is no longer present as no band around 56 kDa can be seen..

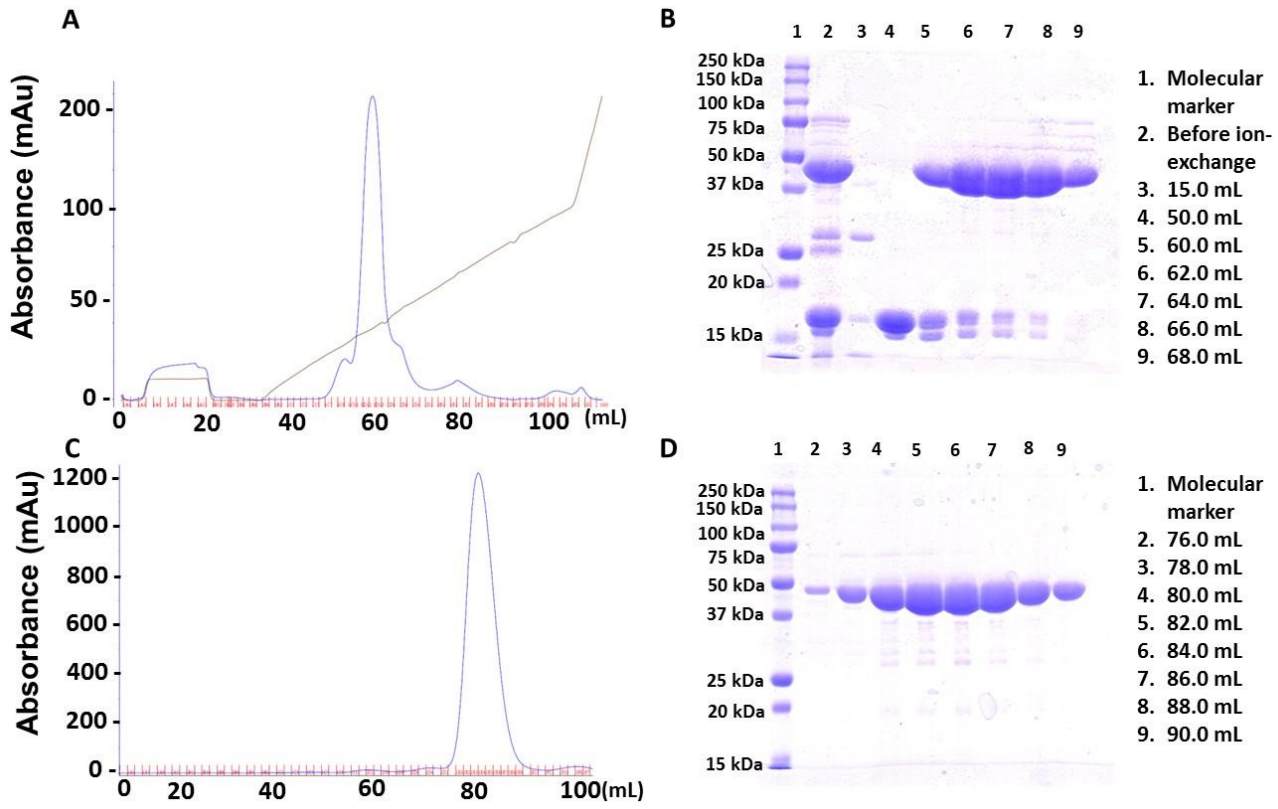

Figure 7.12. Purification of GIT1_R413stop. A. UV trace of ion-exchange of GIT1_R413stop after overnight dialysis into 25 mM Tris pH 8.0, 50 mM NaCl, 10 mM BME. **B**. SDS-PAGE of GIT1_R413stop after ion-exchange showing the presence of the protein in fractions 60.0-68.0 mL. **C**. UV trace from thegel filtation of GIT1_R413stop from fractions 60.0-68.0 mL that were combined and injected into a superdex 200 10/60.tThe gel filtration of GIT1_R413stop showed one peak only. **D**. Corresponding SDS-PAGE of fractions from 76.0-90.0 mL showing purified GIT1_R413stop.Details of purification can be found in section 4.2.14.

7.3.3 Pull down experiments with His_tagged MATβ proteins.

The N-terminal histidine tagged MATβ proteins were used to perform pulldown experiments. Briefly, either MATβV1 or MATβV2 were immoblised by nickel resin and then five times excess of GIT1_R413stop was incubated with the resin for one hour. A low imidazole buffer (40 mM imidazole, buffer 2) was applied to the resin to remove nonspecific binding. As MATβ proteins bind NADP this experiment was performed in the presence and absence of NADP. For MATβV1 most of the GIT1_R413stop is removed from the resin upon wash with buffer 2 (Fig 7.13A, lane 3) and MATβV1 is removed upon washing with 250 mM imidazole (Fig 7.13A, lane 5). This was the same in the presence of NADP (Fig 7.13A, lanes, 6-9). A small band can be seen above MATβV1 (Fig 7.13A Lanes 5 & 9) suggesting the presence of GIT1_R413stop, which maybe interact with MATβV1. To see if GIT1_R413stop could stay in the resin with the absence of MATβ a control whereby GIT1_R413stop was loaded in the absence of MATβ and the elution profile can be view in lanes 10-13. This shows that a small amount of GIT1_R413stop is eluted with 250 mM Imidazole and that this suggests that MATβV1 and GIT_R413stop do not interact. His_MATβV1 is an unstable protein and this is visible by multiple bands (Fig 7.13A, lane 14) which makes it harder to state if the two proteins are interacting. The experiment with MATβV1 resulted in the same outcome to that of MATβV2 (Fig 7.13B).

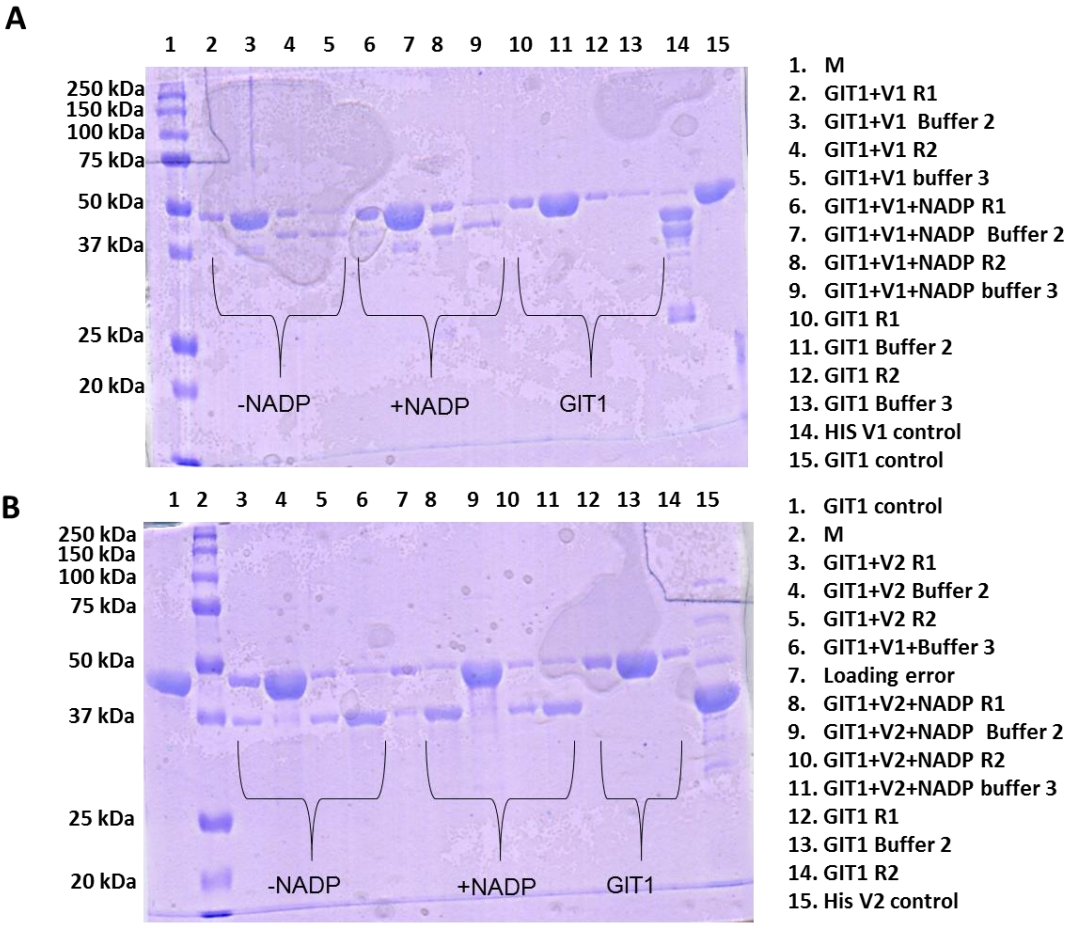

Figure 7.13. SDS gels of pull down experiments with N-terminally histidine tagged MATβ proteins and GIT1_R413stop. Note GIT1 refers to GIT1_R413stop. Experiments were performed in the presence and absence of NADP. A. SDS-PAGE of Pull down of GIT1_R413stop with His-MATβV1 with and with out NADP. B. As in A but with MATβV2. Briefly, 50 μM of His-MATβV1 or MATβV2 ± 100 μM NADP were incubated with nickel resin, that had be equilibrated with 25 mM HEPES pH 7.5, for 1 hour before being loaded into 2 mL flow column. Next, the GIT1_413Stop (50 μM) was passed through the column and the flow-through was collected. A low imidazole (25 mM HEPES pH 7.5, 40 mM imidazole) wash was applied to remove no specific binding. An elution buffer containing 25 mM HEPES pH 7.5 and 250 mM imidazole was applied to the column and the flow-through was collected.

7.3.4 Complex formation studies of GIT1_R413 and MATβ proteins.

As pulldown experiments were inconclusive another approach was taken to see if GIT_R413stop was able to interact with MATβ. GIT1_R413stop was pre-incubated with either MATβV1 or MATβV2 with and without NADP and then injected into a gel filtration column. For MATβV1, no interaction with GIT1_R413stop was seen, as the elution profile of the mixture of proteins (blue) was equal to the individual proteins (purple and green; Fig 7.14A). MATβV1 eluted at 15.5 mL which is typical for this protein whilst GIT1_R413stop, which is larger than MATβV1 eluted at ~15 mL. The results were the same in the presence of NADP, there was no shift in the peak (red) and excess NADP eluted at ~20 mL (Fig 7.14A). Fractions of the peaks were run on SDS-PAGE (Fig 7.14B), and show that the two proteins are not interacting as the maximum concentration of each protein are in different fractions. The same pattern was observed with MATβV2 (Fig 7.15).

GIT1 is a GTPase activating protein, to see if binding GMPPNP to GIT1_R413stop would enable interaction with MATβ, GIT1_R413stop was incubated overnight with EDTA and GMPPNP. Before the gel filtration experiments $MgCl_2$ was added. The incubation with GMPPNP had no effect on the formation of a complex with either MATβV1 or MATβV2 (Fig 7.16 & Fig 7.17). With both isoforms of MATβ the proteins eluted in the same volume regardless on the presence of GIT_R413stop. Excess GMPPNP eluted at 21 mL slightly after NADP. Again, running the peak of the GIT1_R413stop and MATβ on SDS PAGE shows that both proteins elute at different volumes with the peak of each protein eluting in separate fractions (Fig 7.16B & Fig 7.17B) once again indicating the lack of the formation of a complex.

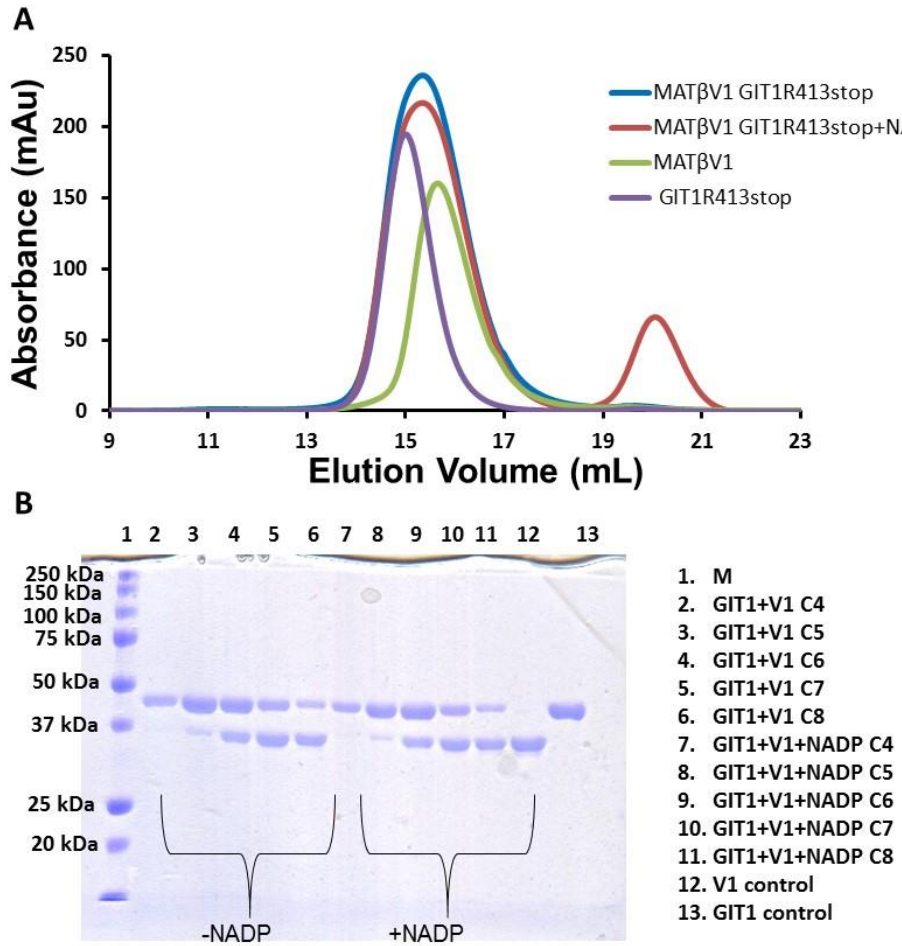

Figure 7.14. Interaction studies between MATβV1 and GIT1_R413stop. **A**. Gel filtration profiles of MATβV1, GIT1_R413stop, MATβV1 and GIT1_R413stop in the presence and absence of 100 μM NADP. Proteins were incubated for 1 hour at 4 °C, prior to gel filtration in 25 mM HEPES pH 7.5, 200 mM NaCl, 10 mM MgCl2, 10 mM KC l± 100 μM NADP.No shift was observed when MATβV1 and GIT1_R413stop were incubated together suggesting no complex has formed. **B**. SDS-PAGE of fractions of the MATβV1+GIT1_R413stop showing that the proteins eluted at different volumes. GIT1 refers to GIT1_R413stop whist V1 refers to MATβV1. C4-C8 refers to 14.5-16.5 mL in fractions of 0.5 mL. All protein concentrations were 50 μM.

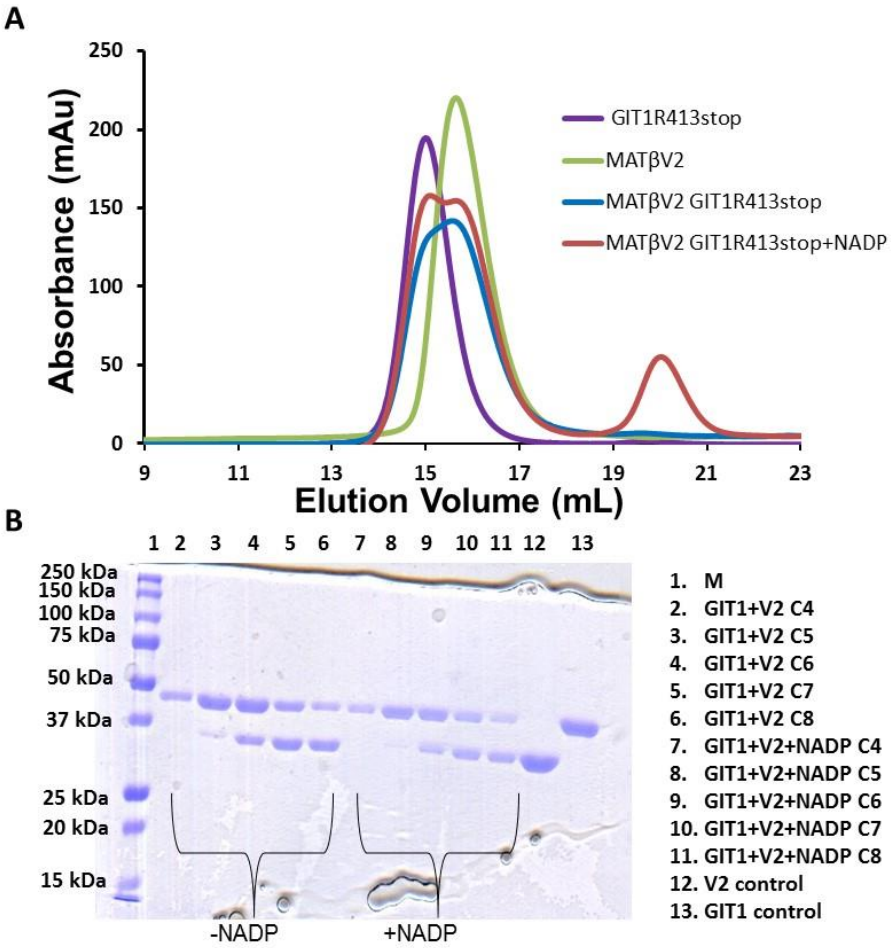

Figure 7.15. Interaction studies between MATβV2 and GIT1_R413stop. **A**. Gel filtration profiles of MATβV2, GIT1_R413stop, MATβV2 and GIT1_R413stop in the presence and absence of 100 μM NADP. Proteins were incubated for 1 hour at 4 °C, prior to gel filtration in 25 mM HEPES pH 7.5, 200 mM NaCl, 10 mM MgCl2, 10 mM KC l± 100 μM NADP. No shift was observed when MATβV2 and GIT1_R413stop were incubated together suggesting no complex had formed. **B**. SDS-PAGE of fractions of the MATβV2+GIT1_R413stop showing that the proteins eluted at different volumes. GIT1 refers to GIT1_R413stop whist V2 refers to MATβV2. C4-C8 refers to 14.5-16.5 mL in fractions of 0.5 mL. All protein concentrations were 50 μM.

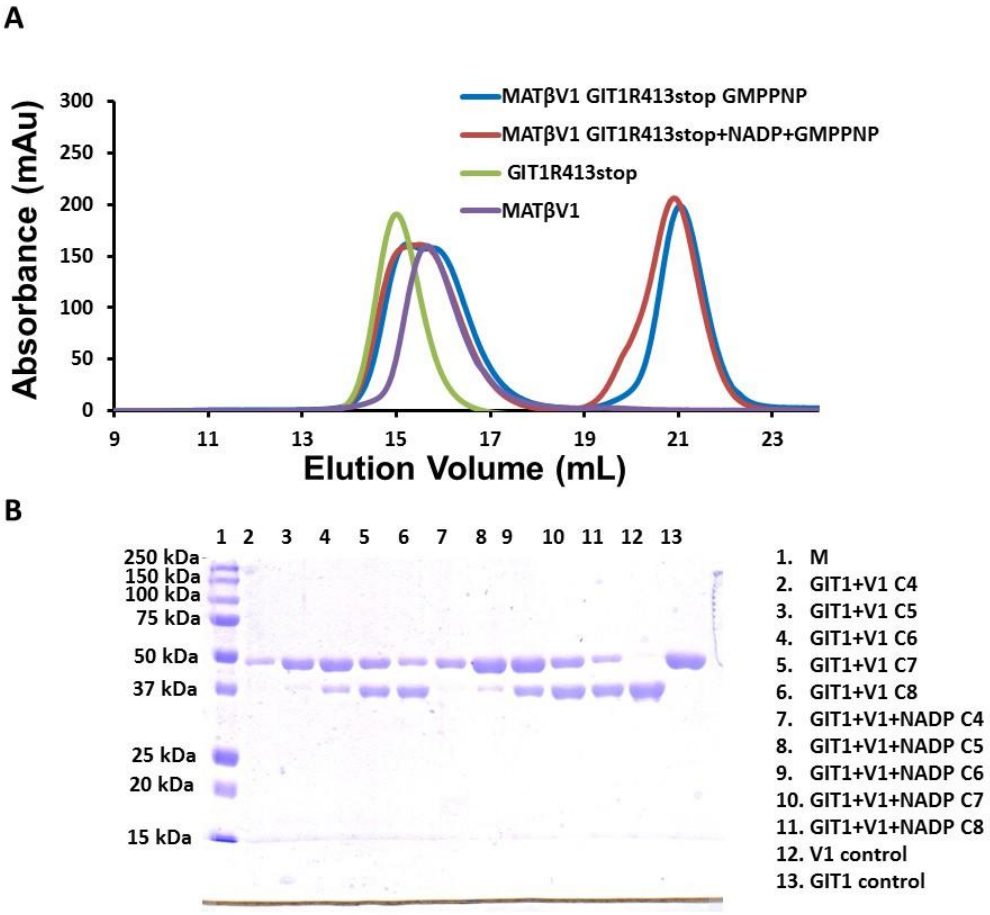

Figure 7.16. Interaction studies between MATβV1 and GIT1_R413stop. **A**. Gel filtration profiles of MATβV1, GIT1_R413stop, MATβV1 and GIT1_R413stop in the presence and absence of 100 μM NADP and the presence of 200 μM GMPPNP. Proteins were incubated for 1 hour at 4 °C, prior to gel filtration in 25 mM HEPES pH 7.5, 200 mM NaCl, 10 mM MgCl2, 10 mM KCl ± 100 μM NADP and/or 200 μM GMPPNP No shift was observed when MATβV1 and GIT1_R413stop are incubated together suggesting no complex has formed. **B**. SDS-PAGE of fractions of the MATβV1+GIT1_R413stop showing that the proteins eluted at different volumes. GIT1 refers to GIT1_R413stop whist V1 refers to MATβV1. Note all fractions that were run on the SDS-PAGE above contained GMPPNP. C4-C8 refers to 14.5-16.5 mL in fractions of 0.5 mL. All protein concentration were 50 μM.

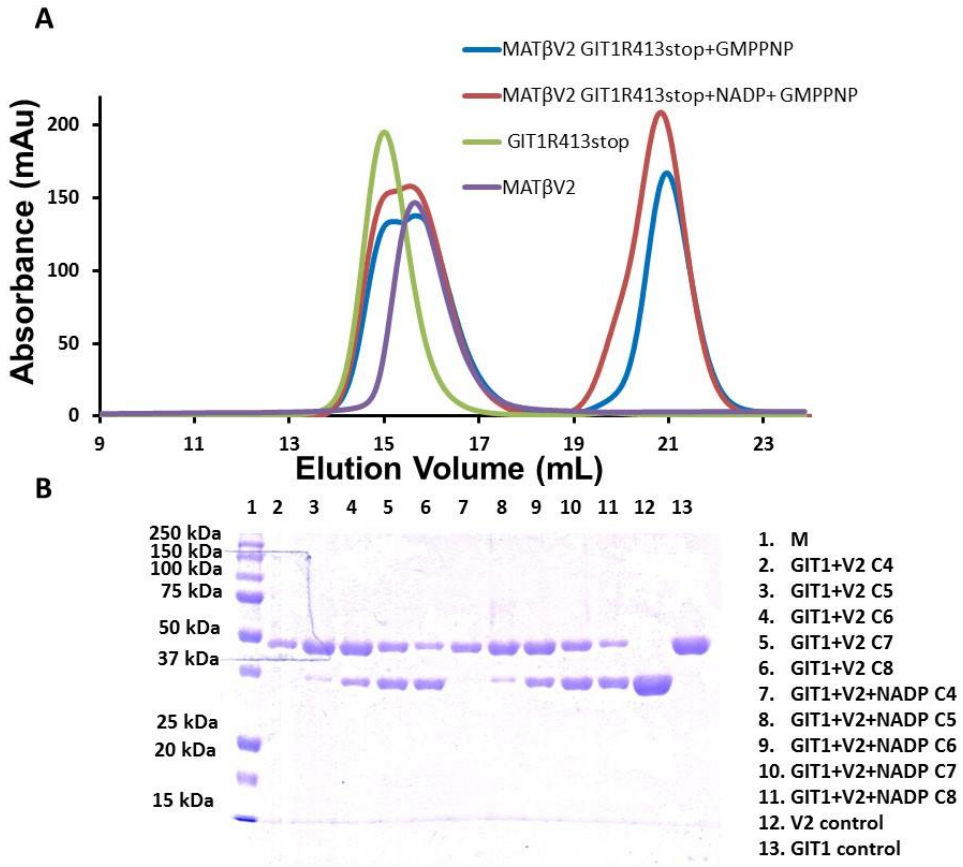

Figure 7.17. Interaction studies between MATβV2 and GIT1_R413stop. **A**. Gel filtration profiles of MATβV2, GIT1_R413stop, MATβV2 and GIT1_R413stop in the presence and absence of 100 μM NADP and the presence of 200 μM GMPPNP. Proteins were incubated for 1 hour at 4 °C, prior to gel filtration in 25 mM HEPES pH 7.5, 200 mM NaCl, 10 mM MgCl2, 10 mM KC l± 100 μM NADP and/or 200 μM GMPPNP. No shift was observed when MATβV2 and GIT1_R413stop are incubated together suggesting no complex has formed. **B**. SDS-PAGE of fractions of the MATβV2+GIT1_R413stop showing that the proteins eluted at different volumes. GIT1 refers to GIT1_R413stop whist V2 refers to MATβV2. Note all fractions that were run on the SDS gel above contained GMPPNP. C4-C8 refers to 14.5-16.5 mL in fractions of 0.5 mL. All protein concentrations were 50 μM.

7.3.5 Structural studies of GIT1_R413stop

As GIT1_R413stop was stable and appeared homogeneous in solution HPLC-SAXS was performed in order to determine the size of this 3 domain mutant. During data collection the radius of gyration remained stable and only a single peak was observed indicating a monomeric species (Fig 7.18A). The data collected was averaged from the frames 105-125 and produced a scattering curve, which was discrete up till a scattering angle of 0.4 $Å^{-1}$ (Fig 7.18A). After processing the scattering data to generate a DAMMIN envelope a low resolution model was created (Fig 7.18C-D). Using the scattering curve of GIT1_R413stop a predicted molecular weight was produced of 47.5 kDa which is very close to the molecular weight of GIT1_R413stop of 49 kDa.

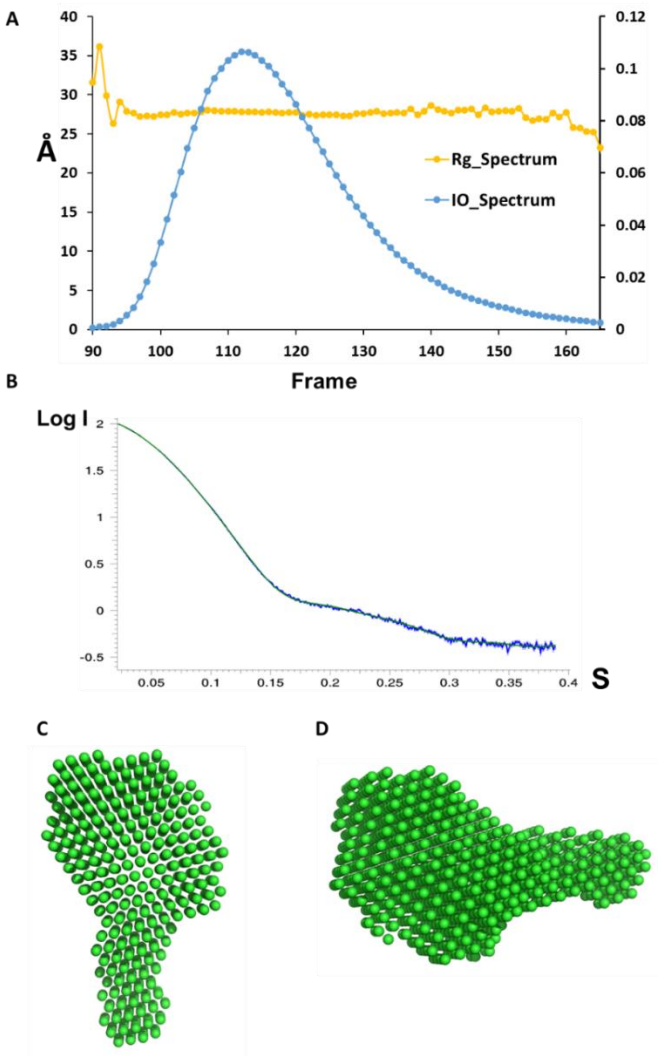

Figure 7.18. Solution scattering of GIT1_R413stop. **A**. Data collection of GIT1_R413stop showing I(0) and radius of gyration allowing for selection of frames to be used in data processing. Frames were chosen for data processing based upon the stability of the radius of gyration with frames that had a stable radius of gyration being used. **B**. Average 1D scattering curves of frames 105-125 for GIT1_R413stop up to 0.4 $Å^{-1}$ scattering angle. **C**. Average DAMMIN model of GIT1_R413stop. **D**. As in C but with a 90° rotation. Section 4.2.23 has further details on SAXS processing.

7.3.6 Homology modelling of GIT1_R413stop

There is no structure of the human GIT1 protein known and only two domains, the coiled coil and paxillin binding domain, have been individually solved. To find suitable models for the three domains of GIT1_R413stop (Fig 7.19A) the sequence of each domain was run in the server Homology Detection & Structure Prediction by HMM-HMM comparison (HHPRED) (Söding et al. 2005). Three structures were selected (Table 7.2) and used in combination with the CORAL programme of the SAXS suite(Petoukhov et al. 2012) . CORAL produced a model with which when compared to the scattering data shows close agreement (X^2=2.26, Fig 7.19B). Linker regions are modelled as spheres which tries to account for the residues between the domains and each ball being a protein Cα (Fig 7.19C). Alignment with the envelope generated by DAMMIN using SUPCOMB does show a good fit for the domains but not for the linker regions (Fig 7.19D).

Table 7.2. List of protein structures used in the modelling of GIT1_R413stop.

Domain	PDB reference	Protein identity
GAP	2IQJ	GAP domain of SMAP1L (LOC64744) stromal membrane-associated protein 1-like
ANK	1N0Q	A designed ankyrin repeat protein with three identical consensus repeats.
SHD	2L7K	Protein CD1104.2 from *Clostridium difficile*. Uncharacterised protein.

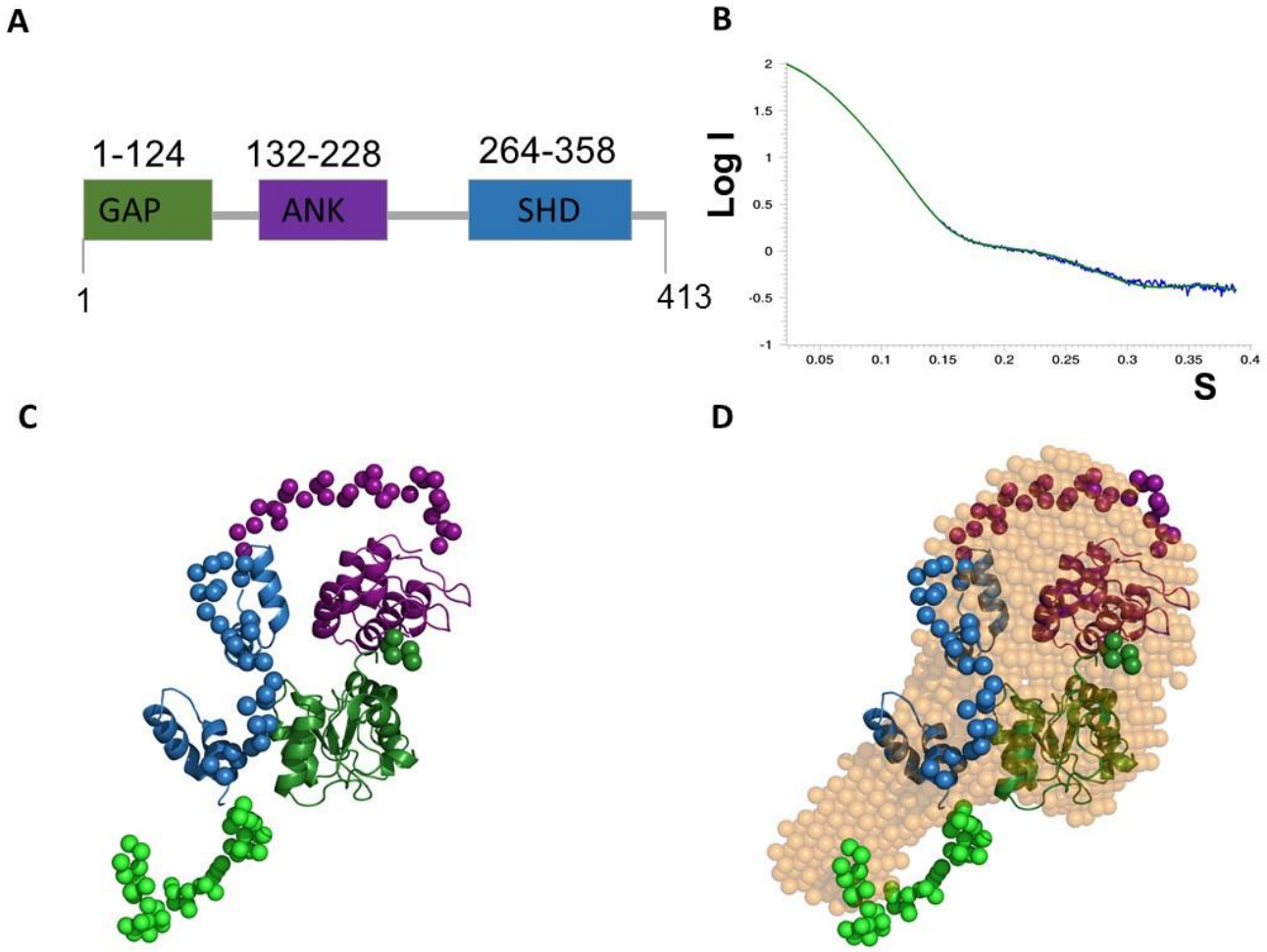

Figure 7.19. Modelling of GIT1_R413stop using predicted structural homologs of GIT1. **A**. Schematic of the domains of GIT1_R413stop showing the GTPase-activating domain (GAP), ankyrin domain (ANK) and the Spa2-homology domain (SHD). **B**. Comparative representation of 1D scattering data against the predicted 1D scattering curve produced from the CORAL model of GIT1_R413stop (X^2=2.26). **C**. CORAL model of GIT1_R413stop with domains shown as cartoons and linker regions shown as balls. Forest shows GAP domain, purple shows ANK domain and blue shows SHD domains. **D**. SUPCOMB alignment of DAMMIN model and CORAL model of GIT1_R413stop.

7.4 Discussion

MATβ proteins can interact with other proteins, such as HuR and are involved in pathways that are independent from SAMe including regulation of growth and apoptosis (Xia et al. 2010; Martínez-Chantar et al. 2003; Yang et al. 2008). G protein-coupled receptor kinase interacting ArfGAP 1 (GIT1) has been shown to interact with MATβ (Peng et al. 2014; Peng et al. 2015), but little information is known about the structure of GIT1 or the functional significance of this interaction.

G protein coupled receptors (GPCRs) are among the largest gene families in the mammalian genome which when activated by GPCR ligands initiate signal transduction through G proteins. One class of G proteins are the GPCR kinases (GRKs), which play a vital role for receptor phosphorylation and internalisation (Hoefen & Berk 2006). Proteins known as GRK-interacting proteins (GITs), a family of scaffolding proteins (Premont et al. 1998), can interact with GRKs. GITs can bind paxillin, a focal adhesion protein (Turner et al. 1999) and phospholipase Cγ (PLCγ) (Haendeler 2000) amongst others. The GIT family of proteins have a complex domain organisation and structure and different domains can bind to different proteins. The role of GITs appears to be in the control of cytoskeletal dynamics and membrane trafficking between the plasma membrane and recycling endosomes (Hoefen & Berk 2006).

Mammals and birds express two GITs, GIT1 and GIT2, both which are highly conserved across both mammals and avians. For example, human and mouse GIT1 have a 98% identity whilst human GIT1 and GIT2 have a 65% sequence identity. GITs can form homo and heteroligomers and are found at the cell periphery in focal complexes, focal adhesions on the basal surface of

the cell and throughout the cell at cytoplasmic structures including lamellipodium (Di Cesare et al. 2000; Loo et al. 2004; Manabe et al. 2002; Mazaki et al. 2001; Paris et al. 2003; Turner et al. 1999; West et al. 2001; Zhao et al. 2000). GIT1 has one form whilst GIT2 has at least 10 splice variants giving rise to up to 32 transcripts (Premont et al. 2000). The full length GIT proteins have an N-terminal ARF-GTPase activating (ARF-GAP) domain, three ankyrin (ANK) repeats, a Spa2-homology domain (SHD), a coiled-coil (CC) domain which is important for dimerisation and a paxillin-binding site (PBS) (Figure 7.20).

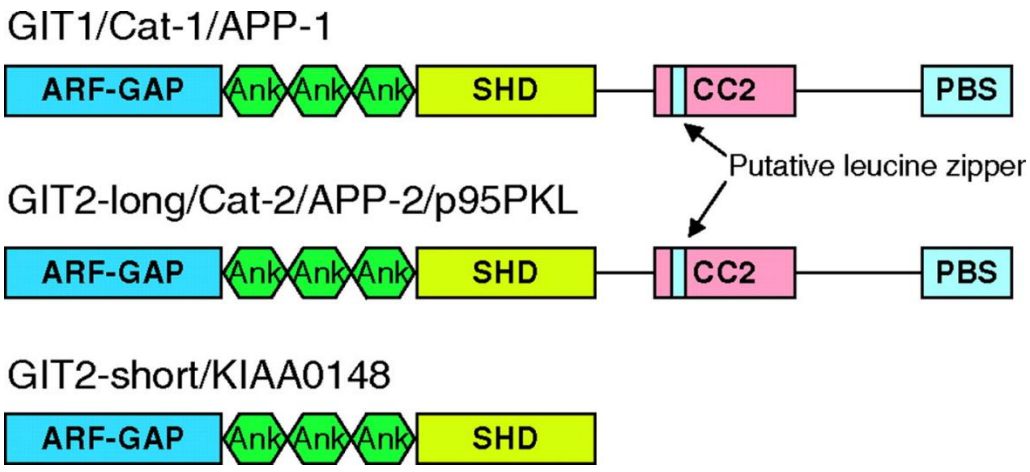

Figure 7.20 GIT1 and GIT2 domain structure. The full-length GIT proteins have an N-terminal ARF GTPase-activating protein (ARF- GAP) domain, three ankyrin (ANK) repeats, a Spa2-homology domain (SHD), a coiled-coil (CC) domain and a paxillin-binding site (PBS). The truncated GIT2-short protein lacks the CC domain and PBS. Figure taken from Hoefen & Berk 2006.

As MAT2β is known to interact with other proteins apart from MATα2 including GIT1 and RAS (Peng et al. 2014; Peng et al. 2015) producing recombinant GIT1 protein in order to characterise this interaction seemed like a logical step, but unfortunately there were problems in producing the full length protein. GIT1 full length expressed in bacterial system but the histidine tag was unable to be cleaved. Peptide mass fingerprinting revealed that the full length protein was unstable and a majority of the protein degraded from the C-terminal to around a 75 kDa protein. In the PET_Sumo vector there is an N-terminal histidine tag and SUMO tag which relies on specific folding to be cleaved by the protease SENP2. SENP2 binds to specific secondary structure to cleave so if the cleavage site has not folded correctly or it is not accessible, it cannot cleave. This was the case for full length GIT1. A possible reason for this is that full length GIT1 contains 5 domains, as shown in Fig 1.9, including a coiled coil region important for dimerisation and a C-terminal paxillin-binding domain, which both may prevent cleavage. Peptide mass fingerprinting of full length GIT1 after nickel affinity chromatography revealed degradation of the protein from the C-terminus. The two C-terminal stop mutants, GIT1_P491stop and GIT1_R413stop, which again like the wild-type overexpressed in *E.coli*, but unlike the wild type the tag was able to be removed. Both mutants lack the C-terminal paxillin-binding domain, which may have caused prevention of cleavage.

All experiments were done in both the presence and absence of NADP as MATβ contains a NADP binding domain (Shafqat et al. 2013). NADP is not needed in terms of the MATαβ complex as shown by ITC, kinetic and gel filtration (Fig 5.16; 5.17; Table 5.9) (Murray et al. 2014; Pajares et al. 2013) but it may be needed to form GIT1_MATβ complexes. The pull down experiments were inconclusive as a small band could be seen at the molecular mass of

GIT1_R413stop in the elution with His-MATβV1, but this band also appeared when the His-MATβV1 was run alone and also in the GIT1_R413stop control where MATβ was absence from the pull down. From this it could mean a small amount of GIT1_R413stop was pulled down, or it was just His-MATβV1 alone or even some GIT1_R413stop bound non-specifically to the nickel resin. Whether GIT1_R413stop and MATβ were able to interact remains a question. Another approach to answering this question was to use gel filtration, but this did not show evidence of complex formation. With GIT1_R413stop no complex was observed and incubation with NADP had no effect. It cannot be ruled out that NADP is needed to aid binding of MATβ to GIT1, as the truncated construct lacks two domains, which may be important for complex formation. The MATβ binding site may only be created upon oligomerisation of GIT1 via its coiled coil domain or it may bind to the paxillin-binding domain. Expression and purification of a stable full length GIT1may be able to help answer these questions.

SEC-SAXS showed that GIT1_R413stop was a homogeneous species as radius of gyration remained constant during data collection and the only one peak appeared on gel filtration (Fig 7.18A). As GIT1_R413stop lacks the dimerisation domain this is consistent with the protein being monomeric. The molecular mass generated from SAXS experimentation was 47 kDa, which is extremely close to the molecular weight of the protein at 49 kDa. HHPRED (Söding et al. 2005) allowed for the identification of structures that share structural homology with the predicted structure of GIT1. Having three protein domain structures enabled the possibility of modelling GIT1_R413stop against the SAXS data that were collected. Linker regions were added during the modelling to take into account the spaces between each domain. Each ball represents a protein Cα but they do not take into account any secondary structure that might

form making the model larger and more spaced out than it may be (Fig 7.19C). The selected protein structures do not share high sequence identity (20-30 %)to that of GIT1_R413stop, but they were selected as they shared high predicted structural homology. ANK domains mediate protein-protein interactions and show no consensus in binding motifs, but share high structural homology (Turner et al. 1999). The SHD domain is a tandem repeat present in orthologs of yeast Spa1 and Spa2 proteins (Arkowitz & Lowe 1997; Gehrung & Snyder 1990) but here, the two SHD domains were modelled using an uncharacterised protein from *Clostridium difficile* because it showed a better agreement with the target sequence.

Chapter VIII: Conclusions and future directions

8.1 New nomenclature for Methionine Adenosyltransferases

The research from this PhD thesis has led to the discovery of several new MAT complexes that had not been described previously as well as evidence supporting that the catalytic proteins have several oligomeric states. Kotb *et al* in 1997 gave a rational to nomenclature of MAT enzymes. This has been updated in view of new discoveries presented in this thesis and those recently reported by others (Murray et al. 2014; Yang et al. 2008). In mammals the three genes *MAT1A*, *MAT2A* and *MAT2B* encode the catalytic subunits MATα1, MATα2 and the regulatory subunit MATβ respectively. MATα1 and MATα2 subunits can self oligomerise to form either dimers or tetramers and can be simply and correctly be referred to referring to their oligomeric state. For example MATα2 can be referred as MAT$(\alpha2)_2$ and MAT$(\alpha2)_4$ to represent the dimeric and tetrameric forms of the same enzyme (Fig. 8.1A). On the other hand, *MAT2B* encodes four isoforms, MATβV1, MATβV2, MATβV2a and MATβV2b. Although only MATβV1 and MATβV2 have been studied in detail the other two isoforms MATβV2a and MATβV2b should be monomeric too, as MATβV2a shares a 94.7 % sequence identity with MATβV2, and MATβV2b is a truncated version of MATβV2a (Fig. 8.1A). The catalytic subunits can form complexes with the regulatory subunits forming hexameric complexes and giving rise to the MATαβ complexes. This group of complexes can be split into MATα1β and MATα2β complexes based on which catalytic subunit is involved. Both MATα1β and MATα2β complexes can be further divided by which MATβ they bind to give MAT$(\alpha1)_4(\beta V1)_2$, MAT$(\alpha1)_4(\beta V2)_2$, MAT$(\alpha2)_4(\beta V1)_2$ and MAT$(\alpha2)_4(\beta V2)_2$ (Fig. 8.1B). To date, no evidence exists for the formation of complexes involving MATβV2a and MATβV2b.

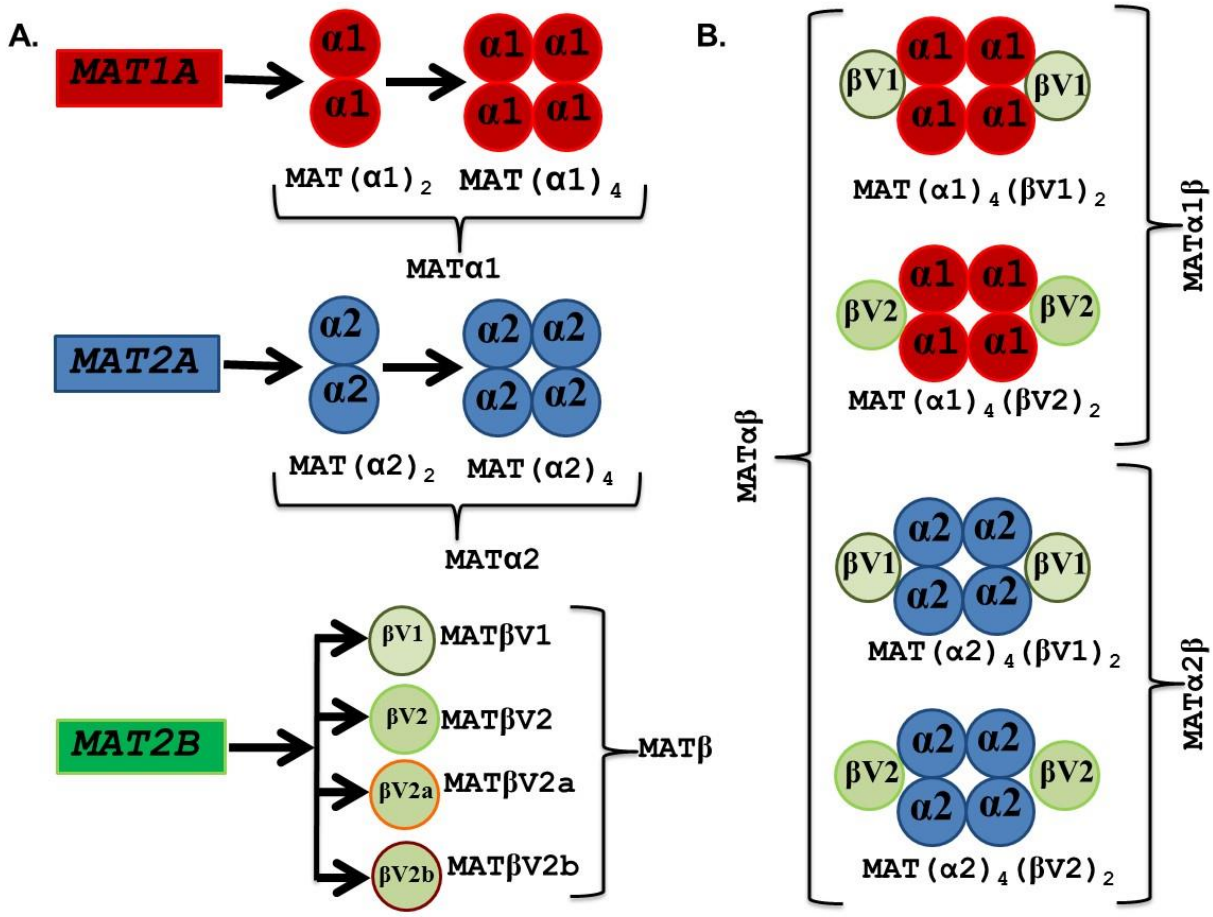

Figure 8.1. Consensus nomenclature of MAT proteins. **A**. The mammalian genes *MAT1A*, *MAT2A* and *MAT2B* encode the catalytic subunits MATα1, MATα2 and the regulatory subunit MATβ respectively. MATα1 and MATα2 subunits can be organised as dimers and tetramers. We propose to use MAT$(\alpha2)_2$ and MAT$(\alpha2)_4$ notation. On the other hand, *MAT2B* encodes for four isoforms, MATβV1, MATβV2, MATβV2a, MATβV2b. **B**. The catalytic subunits could form complexes with the regulatory subunits. The crystallographic structure of MAT$(\alpha2)_4(\beta V2)_2$ complex consists of a MATα2 tetramer flanked by two MATβV2 subunits. MAT$(\alpha2)_4(\beta V1)_2$ has been described *in vivo* (Yang et al. 2008). The other two MATαβ complexes, MAT$(\alpha1)_4(\beta V1)_2$ and MAT$(\alpha1)_4(\beta V2)_2$, have been described *in vitro* (Murray et al. 2014).

8.2 Future directions

8.2.1 What small molecules can bind to MATα2?

In sections 6.1.11 and 6.1.12 it was shown that small molecules are able to bind to MATα2 so using screening libraries and differential scanning fluorimetry (DSF), a wealth of small, cell permeable molecules can be tested against MATα2 to screen for binding. Initial hits would be screened further to ensure binding followed by crystallisation with MATα2. MATα2 could be crystallised to a very high resolution (~1 Å) as shown in section 6.1.5. Of the molecules that bind to the MATβ binding site of MATα2 the effect of these molecules in cell would be evaluated in terms of cell growth, formation of the MATαβ complex, levels of the product, SAMe, within the cell and protein expression levels. Given the high sequence identity between MATα1 and MATα2 successful candidates could be tested on MATα1 to check for binding to ensure that catalytic function is unaffected. Kinetic experiments can evaluate whether the rate of SAMe production is effected by the presence of the small molecules for both MATα isoforms. This would allow for the identification of molecules that both bind the MATβ binding cavity of MATα2 and which do not affect the overall kinetics of the enzymes, which then these molecules could be used as starting point for structure based drug design.

8.2.2 What are the interactomes of the MAT proteins?

There is increasing evidence that MAT proteins have a wider interactome than just other MAT proteins (Yang et al. 2013; Peng et al. 2014; Peng et al. 2015; Katoh et al. 2011). Recombinant MAT proteins that contain N-terminal tags could be produced that could be used as a pull down tool. These proteins will be immoblised in nickel resin and cell lysate from a range of cell

types, including disease and healthy cell lines can be passed through to look for binding partners. Mass spectrometry can be used to identify initial hits and then these can be confirmed by co-immunoprecipitation. Furthermore, the MAT proteins can be cloned into vectors that contain C-terminal tags and once purified subjected to the same procedure as the N-terminal tagged proteins. Once a selection of targets has been identified literature searches can be performed to evaluate the role of the binding targets within the cell to look for disease related proteins. Selected hits can be produced recombinantly, used to form complexes with their respective MAT protein and finally their structure determined.

8.2.3 Is there a back gate to the active site of MAT proteins?

A characteristic of MAT proteins is that the active site contains a 'gating loop' that closes the active site preventing products and substrates from leaving the active site during catalysis. However, during my PhD a number of data sets were collected which contained compounds within the active site that were only added during the crystal freezing process. Upon analysis, the structures contained a closed gating loop with products of SAMe synthesis and also molecules of compounds added after crystallisation e.g., glycerol. If the gate is closed how can these compounds enter? The MATβ binding tunnel of MATα2 provides a point of contact where the active site has access to the outside of the MATα2 protein, as there is a loop which connects the active and this channel. Mutation of this loop could reveal if products can enter or leave the active site through this loop. Mutants can be purified, crystallised and molecular modelling could be done to look if the size of this 'back gate' is changed and if so how it relates to the production of SAMe, which can be evaluated by LC-MS (Murray *et al* 2014).

8.2.4 What prompts the release of products from the active site?

It was previously believed that the breakdown of the triple phosphate provided the energy to open the closed gating loop, allowing the release of products. In section 6.1.8 it was shown that using a modified ATP analogue, containing a triple phosphate that is unable to be hydrolyzed, that it can sit within the active site without have a closed gating loop. Further modelling could be done to investigate how substrates are able to leave the active site and to see how the flexible gating loop is regulated. The high resolution allows for the placement of many waters within the active site and analysis may reveal if these waters are able to aid in the breakdown of the triphosphate, which would be consistent with other phosphatases.

8.2.5 Does the interaction site of GIT1 and MATβ lie in the coiled coil domain or the Paxillin binding domain?

In section 7.3 it was shown that GIT1, containing only 3 domains, is unable to interact with either of the MATβ proteins. In order to explore the site of interaction a full length GIT1 will need to be purified and subsequent complex formation with MATβ isoforms. Size-exclusion chromatography and isothermal titration calorimetry (ITC) will be used to access formation of the complexes. GIT1_MATβ complexes will be structurally characterised by crystallisation and SAXS experimentation. There is no structure of GIT1 published so either Single-wavelength anomalous diffraction (SAD) experimentation will need to be done on the GIT1 or if data of the GIT1_MATβ complex is collected then molecular replacement can be attempted using MATβ as a search model. As GIT1_R413stop has been purified previously crystallisation of this construct can be further tested. The GIT1_413stop plasmid will be transformed into a methionine auxotrophic strain of *E.coli* followed by expression in commercial SelenoMet

media. This will be done in order to replace methionine side chains with the selenomethionine which will allow for multi-wavelength anomalous scattering to be done to solve the structure without a search model (Hendrickson et al. 1990).

References

Abe, T. & Tsukada, K., 1981. S-adenosylmethionine synthetase isozymes in the liver of tumor-bearing mice. *Journal of Biochemistry*, 90(2), pp.571–574.

Adams, P.D. et al., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica. Section D, Biological crystallography*, 66(Pt 2), pp.213–21.

Alsarraf, H.M. a B. et al., 2011. Purification, crystallization and preliminary crystallographic studies of the TLDc domain of oxidation resistance protein 2 from zebrafish. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 67(10), pp.1253–1256.

Alvarez, L. et al., 1991. Analysis of the 5' non-coding region of rat liver S-adenosylmethionine synthetase mRNA and comparison of the Mr deduced from the cDNA sequence and the purified enzyme. *FEBS Letters*, 290(1-2), pp.142–146.

Alvarez, L. et al., 1997. Characterization of rat liver-specific methionine adenosyltransferase gene promoter. *The Journal of Biological Chemistry*, 272(36), pp.22875–22883.

Arkowitz, R.A. & Lowe, N., 1997. A Small Conserved Domain in the Yeast Spa2p Is Necessary and Sufficient for Its Polarized Localization. *The Journal of Cell Biology*, 138(1), pp.17–36.

Arteel, G.E. et al., 1997. Chronic enteral ethanol treatment causes hypoxia in rat liver tissue in vivo. *Hepatology*, 25(4), pp.920–926.

Augoustides-Savvopoulou, P. et al., 2003. Glycine N -methyltransferase deficiency: a new patient with a novel mutation. *Journal of Inherited Metabolic Disease*, 26, pp.745–759.

Avila, M. et al., 2000. Methyl groups in the regulation of MAT1A and MAT2A gene expression. In *Methionine Metabolism: Molecular Mechanisms and Clinical Implications*. pp. 91–104.

Avila, M. et al., 2000. Reduced mRNA abundance of the main enzymes involved in methionine metabolism in human liver cirrhosis and hepatocellular carcinoma. *Journal of Hepatology*, 33(6), pp.907–914.

Avila, M. a. et al., 1998. Regulation by hypoxia of methionine adenosyltransferase activity and gene expression in rat hepatocytes. *Gastroenterology*, 114(2), pp.364–371.

Avila, M.A. et al., 1997. Regulation of rat liver S-adenosylmethionine synthetase during septic shock: Role of nitric oxide. *Hepatology*, 25(2), pp.391–396.

Barberato, C. et al., 1995. CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography*, 28(6), pp.768–773.

Beaudouin, C. et al., 1993. The presence of (+)-S-adenosyl-L-methionine in the rat brain and its lack of effect on phenylethanolamine N-methyltransferase activity. *Journal of Neurochemistry*, 61, pp.928–935.

Bentley, R., 2005. Role of sulfur chirality in the chemical processes of biology. *Chemical Society Reviews*, 34, pp.609–624.

Blanchet, C.E. & Svergun, D.I., 2013. Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution. *Annual Review of Physical Chemistry*, 64, pp.37–54.

Bragg, W.H. & Bragg, W.L., 1913. The Reflection of X-rays by Crystals. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 88(605), pp.428–438.

Cabrero, C. et al., 1988. Specific loss of the high-molecular-weight form of S-adenosyl-L-methionine synthetase in human liver cirrhosis. *Hepatology*, 8(6), pp.1530–1534.

Cai, J. et al., 1998. Differential expression of methionine adenosyltransferase genes influences the rate of growth of human hepatocellular carcinoma cells. *Cancer Research*, 58(7), pp.1444–1450.

Cannon, L.M. et al., 2002. A stereospecific colorimetric assay for (S,S)-adenosylmethionine quantification based on thiopurine methyltransferase-catalyzed thiol methylation. *Analytical Biochemistry*, 308, pp.358–363.

Cantoni, G., 1952. The nature of the active methly donor formed enzymatically from L-methionine and adensoinetriphosphate. *Journal of the American Chemical Society*, 377(9), pp.7–8.

Carugo, O., 1997. Protein-protein crystal-packing contacts. *Protien Science*, 6, pp.2261–2263.

Castro, C. et al., 1999. Creation of a functional S-nitrosylation site in vitro by single point mutation. *FEBS Letters*, 459(3), pp.319–322.

Di Cesare, A. et al., 2000. p95-APP1 links membrane transport to Rac-mediated reorganization of actin. *Nature Cell Biology*, 2(8), pp.521–530.

Chayen, N.E. & Saridakis, E., 2008. Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods*, 5(2), pp.147–153.

Chen, L. et al., 2004. Impaired liver regeneration in mice lacking methionine adenosyltransferase 1A. *The FASEB Journal*, 18(7), pp.914–916.

Cheng, H. et al., 1997. Levels of L-methionine S-adenosyltransferase activity in erythrocytes and concentrations of S-adenosylmethionine and S-adenosylhomocysteine in whole blood of patients with Parkinson's disease. *Experimental Neurology*, 145(2 Pt 1), pp.580–585.

Chiang, P. et al., 1996. S-Adenosylmethionine and methylation. *FASEB J*, 10, pp.471–480.

Chiang, P.K., 1998. Biological effects of inhibitors of S-adenosylhomocysteine hydrolase. *Pharmacology and Therapeutics*, 77(2), pp.115–134.

Cornforth, J.W. et al., 1977. Determination of the absolute configuration at the sulfonium center of S-adenosylmethionine. Correlation with the absolute configuration of the diastereomeric S-carboxymethyl-(S)-methionine salts. *Journal of the American Chemical Society*, 99, pp.7292–7300.

Corrales, F. et al., 1990. Inactivation and dissociation of s-adenosylmethionine synthetase by modification of sulfhydryl groups and its possile occurrence in cirrhosis. *Hepatology*, 11(2), pp.216–222.

Corrales, F., Ruiz, F. & Mato, J.M., 1999. In vivo regulation by glutathione of methionine adenosyltransferase S-nitrosylation in rat liver. *Hepatology*, (5), pp.887–894.

Cressman, D.E. et al., 1996. Liver failure and defective hepatocyte regeneration in interleukin-6-deficient mice. *Science*, 274(5291), pp.1379–1383.

David, G. & Pérez, J., 2009. Combined sampler robot and high-performance liquid chromatography: A fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *Journal of Applied Crystallography*, 42, pp.892–900.

Davis, I.W. et al., 2007. MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research*, 35(13), pp.375–383.

Desantis, K. et al., 2012. Use of differential scanning fluorimetry as a high-throughput assay to identify nuclear receptor ligands. *Nuclear Receptor Signaling Atlas*, 10, pp.1–5.

Dessau, M. a & Modis, Y., 2011. Protein crystallization for X-ray crystallography. *Journal of visualized experiments : JoVE*, 9, pp.1–6.

Drenth, J., 2007. *Principles of Protein X-Ray Crystallography*, Springer-Verlag.

Emsley, P. & Cowtan, K., 2004. Coot: model-building tools for molecular graphics. *Acta Crystallographica. Section D, Biological Crystallography*, 60(Pt 12), pp.2126–32.

Finkelstein, J.D., 2006. Inborn Errors of Sulfur-Containing Amino Acid Metabolism 1. *The Journal of Nutrition*, 136, p.1750S–1754S.

Finkelstein, J.D., 2007. Metabolic regulatory properties of S-adenosylmethionine and S-adenosylhomocysteine. *Clinical Chemistry and Laboratory Medicine*, 45, pp.1694–1699.

Finkelstein, J.D., 1990. Methionine metabolism in mammals. *Journal of Nutritional Biochemistry*, 1, pp.228–237.

Fischer, H. et al., 2009. Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *Journal of Applied Crystallography*, 43(1), pp.101–109.

Franke, D. & Svergun, D.I., 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography*, 42, pp.342–346.

Frey, P. a, Hegeman, A.D. & Ruzicka, F.J., 2008. The Radical SAM Superfamily. *Critical Reviews in Biochemistry and Molecular Biology*, 43(July), pp.63–88.

Fu, Z. et al., 1996. Flexible loop in the structure of S-Adenosylmethionine synthetase crystallized in the tetragonal modification. *Journal of Biomolecular Structure and Dynamics*, 13(5), pp.727–739.

Fua, Z. et al., 1996. Flexible loop in the structure of S-adenosylmethionine synthetase crystallized in the tetragonal modification. *Journal of Biomolecular Structure and Dynamics*, 13(5), pp.727–39.

Garcea, R. et al., 1989. Inhibition of promotion and persistent nodule growth by S -adenosyl-L-methionine in rat liver carcinogenesis : role of remodeling and apoptosis. *Cancer Research*, 49, pp.1850–1856.

García-Trevijano, E.R. et al., 2002. NO sensitizes rat hepatocytes to proliferation by modifying S-adenosylmethionine levels. *Gastroenterology*, 122(5), pp.1355–1363.

Gehrung, S. & Snyder, M., 1990. The SPA2 gene of Saccharomyces cerevisiae is important for pheromone-induced morphogenesis and efficient mating. *The Journal of Cell Biology*, 111(4), pp.1451–1464.

Gil, B. et al., 1996. Differential expression pattern of S-adenosylmethionine synthetase isoenzymes during rat liver development. *Hepatology*, 24(4 I), pp.876–881.

Glass, J.I. et al., 2006. Essential genes of a minimal bacterium. *Proceedings of the National Academy of Sciences of the United States of America*, 103(2), pp.425–430.

Gomes Trolin, C., 1998. Erythrocyte and brain methionine adenosyltransferase activity in patients with schizophrenia. *Journal of Neural Transmission*, 105(10), pp.1293–1305.

Gomes Trolin, C., Regland, B. & Oreland, L., 1995. Decreased methionine adenosyltransferase activity in erythrocytes of patients with dementia disorders. *European Neuropsychopharmacology*, 5(2), pp.107–114.

González, B. et al., 2003. Crystal structures of methionine adenosyltransferase complexed with substrates and products reveal the methionine-ATP recognition and give insights into the catalytic mechanism. *Journal of Molecular Biology*, 331(2), pp.407–416.

González, B. et al., 2012. NADP+ Binding to the Regulatory Subunit of Methionine Adenosyltransferase II Increases Intersubunit Binding Affinity in the Hetero-Trimer. *PLoS ONE*, 7(11).

González, B. et al., 2000. The crystal structure of tetrameric methionine adenosyltransferase from rat liver reveals the methionine-binding site. *Journal of Molecular Biology*, 300(2), pp.363–75.

Guinier, A., 1939. *La Diffraction des rayons X aux tres petits angles: application a l'etude de phenomenes ultramicroscopiques.* Paris: Masson.

Haendeler, J., 2000. Angiotensin II-stimulated site specific phosphorylation of GIT1 determines the relative activation of PLC gamma and ERK1/2. *Circulation*, 102(18), p.15.

Halim, A.B. et al., 1999. Expression and functional interaction of the catalytic and regulatory subunits of human methionine adenosyltransferase in mammalian cells. *Journal of Biological Chemistry*, 274(42), pp.29720–29725.

Halpern, B.C. et al., 1974. The Effect of replacement of methionine by homocystine on survival of malignant and normal adult mammalian cells in culture. *Proceedings of the National Academy of Sciences of the United States of America*, 71(4), pp.1133–1136.

Harding, M.M., 2006. Small revisions to predicted distances around metal sites in proteins. *Acta Crystallographica Section D Biological Crystallography*, 62(6), pp.678–682.

Hegazi, M., Borchardt, R.T. & Schowen, R.L., 1976. SN2-like transition state for methyl transfer catalyzed by catechol-O-methyl-transferase. *Journal of the American Chemical Society*, 98, pp.3048–3049.

Hendrickson, W.A., Horton, J.R. & LeMaster, D.M., 1990. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *The EMBO Journal*, 9(5), pp.1665–1672.

Hibasami, H. et al., 1980. Studies of inhibition of rat spermidine synthase and spermine synthase. *The Biochemical Journal*, 187, pp.419–428.

Hiroki, T., Horikawa, S. & Tsukada, K., 1997. Structure of the Rat Methionine Adenosyltransferase 2A Gene and its Promoter. *European Journal of Biochemistry*, 250(3), pp.653–660.

Hoefen, R.J. & Berk, B.C., 2006. The multifunctional GIT family of proteins. *Journal of Cell Science*, 119, pp.1469–1475.

Hoffman, D. et al., 1980. S-Adenosylmethionnine and S-Adenosylhomocysteine Metabolism in Isoated Rat Liver. *Journal of Biological Chemistry*, 255(22), pp.10822–10827.

Hoffman, J.L., 1986. Chromatographic analysis of the chiral and covalent instability of S-adenosyl-L-methionine. *Biochemistry*, 25, pp.4444–4449.

Horikawa, S. et al., 1993. Immunohistochemical analysis of rat S-adenosylmethionine synthetase isozymes in developmental liver. *FEBS Letters*, 330(3), pp.307–311.

Huang, Z.Z. et al., 1998. Changes in methionine adenosyltransferase during liver regeneration in the rat. *The American Journal of Physiology*, 275(1 Pt 1), pp.G14–G21.

Huang, Z.-Z. et al., 1999. Differential effect of thioacetamide on hepatic methionine adenosyltransferase expression in the rat. *Hepatology*, 29(5), pp.1471–1478.

Ji, S. et al., 1982. Periportal and pericentral pyridine nucleotide fluorescence from the surface of the perfused liver: evaluation of the hypothesis that chronic treatment with ethanol produces pericentral hypoxia. *Proceedings of the National Academy of Sciences of the United States of America*, 79(17), pp.5415–5419.

Kabsch, W., 2010. XDS. *Acta Crystallographica Section D*, 66(Pt 2), pp.125–32.

Katoh, Y. et al., 2011. Methionine adenosyltransferase II serves as a transcriptional corepressor of Maf oncoprotein. *Molecular cell*, 41(5), pp.554–66.

Kleywegt, G.J. & Jones, T.A., 1997. Model building and refinement practice. *Methods in Enzymology*, 277(1995), pp.208–230.

Knappe, J. & Schmitt, T., 1976. A Novel reaction of S-Adenosyl-L-Methionine correlated with the activation of pyruvate formate-lyase. *Biochemical and Biophysical Research Communications*, 71(4), pp.1110–1117.

Koch, M.H., Vachette, P. & Svergun, D.I., 2003. *Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution.*,

Komoto, J. et al., 2004. Crystal structure of the S -adenosylmethionine synthetase ternary complex : a novel catalytic mechanism of S -adenosylmethionine synthesis from ATP and Met. *Biochemistry*, 43, pp.1821–1831.

Konarev, P. V. et al., 2003. PRIMUS : a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography*, 36(5), pp.1277–1282.

Kotb, M. et al., 1997. Consensus nomenclature for the mammalian methionine adenosyltransferase genes and gene products. *Trends in Genetics*, 13(2), pp.51–52.

Kotb, M. & Kredich, N.M., 1985. S-Adenosylmethionine synthetase from human lymphocytes. Purification and characterization. *Journal of Biological Chemistry*, 260(84), pp.3923–3930.

Kozin, M.B. & Svergun, D.I., 2001. Automated matching of high- and low-resolution structural models research papers. *Journal of Applied Crystallography*, 34(1), pp.33–41.

Krissinel, E. & Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. *Journal of Biological Chemistry*, 43(372), pp.774–797.

De la Haba, G. et al., 1959. S-Adenosylmethionine: The Relation of Configuration at the Sulfonium Center to Enzymatic Reactivity. *Journal of the American Chemical Society*, 81(1958), pp.3975–3980.

Langkamp-henken, B., Geller, A.M. & Legros, H.L., 1994. Characterization of distinct forms of methionine adenosyltransferase in nucleated, and mature human erythrocytes and erythroleukemic cells. *Biochimica et Biopysica Acta*, 1201, pp.397–404.

Latasa, M.U. et al., 2001. Hepatocyte growth factor induces MAT2A expression and histone acetylation in rat hepatocytes: role in liver regeneration. *The FASEB Journal*, 15(7), pp.1248–1250.

Lee, C.-C., Surtees, R. & Duchen, L.W., 1992. Distal motor axonopathy and central nervous system myelin vacuolation caused by cycloleucine, an inhibitor of methionine adenosyltransferase. *Brain*, 115(3), pp.935–955.

Lee, T.D. et al., 2004. Abnormal Hepatic Methionine and Glutathione Metabolism in Patients With Alcoholic Hepatitis. *Alcoholism: Clinical and Experimental Research*, 28(1), pp.173–181.

LeGros, H.L. et al., 2000. Cloning, Expression, and Functional Characterization of the Regulatory Subunit of Human Methionine Adenosyltransferase (MAT II). *Journal of Biological Chemistry*, 275(4), pp.2359–2366.

LeGros, L. et al., 2001. Regulation of the Human MAT2B Gene Encoding the Regulatory 2B Subunit of Methionine Adenosyltransferase, MAT II. *Journal of Biological Chemistry*, 276(27), pp.24918–24924.

Liang, C.R.M.Y. et al., 2005. Proteome analysis of human hepatocellular carcinoma tissues by two-dimensional difference gel electrophoresis and mass spectrometry. *Proteomics*, 5(8), pp.2258–2271.

Liau, M.C., Chang, C.F. & Becker, F.F., 1979. Alteration of S-adenosylmethionine synthetases during chemical hepatocarcinogenesis and in resulting carcinomas. *Cancer Research*, 39(6 I), pp.2113–2119.

Lieber, C.S. et al., 1990. S-Adenosyl-L-methionine attenuates alcohol-induced liver injury in the baboon. *Hepatology*, 11(2), pp.165–172.

Van Liempd, S. et al., 2013. A fast method for the quantitation of key metabolites of the methionine pathway in liver tissue by high-resolution mass spectrometry and

hydrophilic interaction ultra-performance liquid chromatography. *Analytical and Bioanalytical Chemistry*, 405, pp.5301–5310.

Lin, S. et al., 2002. A novel S-adenosyl-L-methionine:arsenic(III) methyltransferase from rat liver cytosol. *Journal of Biological Chemistry*, 277(13), pp.10795–10803.

Loenen, A., 2006. S-adenosylmethionine: jack of all trades and master of everything. *Biochemical Society Transactions*, 34(2), pp.330–333.

Loo, T.-H. et al., 2004. GIT1 Activates p21-Activated Kinase through a Mechanism Independent of p21 Binding. *Molecular and Cellular Biology*, 24(9), pp.3849–3859.

Lu, S.C. et al., 2000. Changes in methionine adenosyltransferase and S-adenosylmethionine homeostasis in alcoholic rat liver. *American journal of physiology. Gastrointestinal and liver physiology*, 279(1), pp.G178–G185.

Lu, S.C., 2009. Regulation of glutathione synthesis. *Molecular Aspects of Medicine*, 30(1-2), pp.42–59.

Lu, S.C. & Mato, J.M., 2012. S-adenosylmethionine in liver health, injury, and cancer. *Physiological reviews*, 92(4), pp.1515–42.

Majano, P.L. et al., 2001. S-Adenosylmethionine modulates inducible nitric oxide synthase gene expression in rat liver and isolated hepatocytes. *Journal of hepatology*, 35(6), pp.692–699.

Manabe, R. et al., 2002. GIT1 functions in a motile, multi-molecular signaling complex that regulates protrusive activity and cell migration. *Journal of cell science*, 115(Pt 7), pp.1497–1510.

Mao, Z. et al., 1998. Cloning and functional characterization of the 5'-flanking region of human methionine adenosyltransferase 2A gene. *Biochemical and Biophysical Research Communications*, 248(3), pp.479–484.

Markham, G.D. & Pajares, M. a., 2010. Struture-function relationships in methionine adenosyltransferases. *Cellular and Molecular Life Sciences*, 66(4), pp.636–648.

Markham, G.D., Parkinq, D.W. & Schramm, V.L., 1987. A kinetic isotope effect study and transition state analysis of the S-adenosylmethionine synthetase reaction. *The Journal of Biological Chemistry*, 262(12), pp.5609–5615.

Martin-Duce, A. et al., 1988. S-Adenosylmethionine synthetase and phospholipid methyl-transferase are inhibited in human cirrhosis. *Hepatology*, 8, pp.65–68.

Martínez-Chantar, M.L. et al., 2003. Methionine adenosyltransferase II β subunit gene expression provides a proliferative advantage in human hepatoma. *Gastroenterology*, 124(4), pp.940–948.

Martínez-Chantar, M.L. et al., 2002. Spontaneous oxidative stress and liver tumors in mice lacking methionine adenosyltransferase 1A. *The FASEB Journal*, 16(10), pp.1292–1294.

Mashalidis, E.H. et al., 2013. A three-stage biophysical screening cascade for fragment-based drug discovery. *Nature Protocols*, 8(11), pp.2309–2324.

Mato, J.M. et al., 1997. S-adenosylmethionine synthesis: molecular mechanisms and clinical implications. *Pharmacology & Therapeutics*, 73(3), pp.265–280.

Mato, J.M. & Lu, S.C., 2007. Role of S-adenosyl-L-methionine in liver health and injury. *Hepatology*, 45(5), pp.1306–1312.

Mazaki, Y. et al., 2001. An ADP-ribosylation factor GTPase-activating protein Git2-short/KIAA0148 is involved in subcellular localization of paxillin and actin cytoskeletal organization. *Molecular Biology of the Cell*, 12(3), pp.645–662.

Mbimba, T. et al., 2012. Alteration of hepatic proinflammatory cytokines is involved in the resveratrol-mediated chemoprevention of chemically-induced hepatocarcinogenesis. *Current Pharmaceutical Biotechnology*, 13(1), pp.229–234.

McCoy, A.J. et al., 2007. Phaser crystallographic software. *Journal of Applied Crystallography*, 40(Pt 4), pp.658–674.

Merali, S. & Clarkson, A.B., 2004. S-Adenosylmethionine and pneumocystis. *FEMS Microbiology Letters*, 237(2), pp.179–186.

Mitsui, K., Teraoka, H. & Tsukada, K., 1988. Complete purification and immunochemical analysis of S-Adenosylmethionine synthetase from bovine b rain. *Journal of Biological Chemistry*, 263(23), pp.11211–11216.

Mudd, S.H., 1962. Activation of methionine for transmethylation. *The Journal of Biological Chemistry*, 237(4), pp.1372–1376.

Mudd, S.H. et al., 2001. Glycine N-methyltransferase deficiency: A novel inborn error causing persistent isolated hypermethioninaemia. *Journal of Inherited Metabolic Disease*, 24(4), pp.448–464.

Mudd, S.H. & Poole, J.R., 1975. Labile methyl balances for normal humans on various dietary regimens. *Metabolism: Clinical and Experimental*, 24(6), pp.721–735.

Murray, B. et al., 2014. Structure and function study of the complex that synthesizes S-adenosylmethionine. *International Union of Crystallography*, 1(Pt 4), pp.240–9.

Murshudov, G.N. et al., 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica. Section D, Biological Crystallography*, 67(Pt 4), pp.355–67.

Niesen, F.H., Berglund, H. & Vedadi, M., 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*, 2(9), pp.2212–2221.

Okada, G. et al., 1979. Differential effects of dimethylsulfoxide on S-adenosylmethionine synthetase from rat liver and hepatoma. *FEBS Letters*, 106(1), pp.25–28.

Oppermann, U. et al., 2003. Short-chain dehydrogenases/reductases (SDR): The 2002 update. *Chemico-Biological Interactions*, 143-144, pp.247–253.

Orenstein, J.M. & Marsh, W.H., 1968. Incorporation in vivo of methionine and ethionine into and the methylation and ethylation of rat liver nuclear proteins. *Biochemical Journal*, 109(4), pp.697–699.

Otwinowski, Z. & Minor, W., 1997. Processing of X-Ray diffraction data collected in oscillation mode. *Methods in Enzymology*, 276, pp.306–315.

Pajares, M. a et al., 1994. Protein kinase C phosphorylation of rat liver S-adenosylmethionine synthetase: dissociation and production of an active monomer. *The Biochemical Journal*, 303 (Pt 3(1 994), pp.949–955.

Pajares, M. a., Álvarez, L. & Pérez-Sala, D., 2013. How are mammalian methionine adenosyltransferases regulated in the liver? A focus on redox stress. *FEBS Letters*, 587, pp.1711–1716.

Pajula, R.L. & Raina, A., 1979. Methylthioadenosine, a potent inhibitor of spermine synthase from bovine brain. *FEBS Letters*, 99(2), pp.343–345.

Pajula, R.-L., Raina, A. & Eloranta, T., 1979. Polyamine Synthesis in Mammalian Tissues. *European Journal of Biochemistry*, 101(2), pp.619–626.

Pañeda, C. et al., 2002. Liver cell proliferation requires methionine adenosyltransferase 2A mRNA up-regulation. *Hepatology*, 35(6), pp.1381–1391.

Paris, S. et al., 2003. Leucine-zipper-mediated homo- and hetero-dimerization of GIT family p95-ARF GTPase-activating protein, PIX-, paxillin-interacting proteins 1 and 2. *Biochemical Journal*, 372(Pt 2), pp.391–398.

Parry, R.J. et al., 1982. Studies of enzyme stereochemistry. Elucidation of the stereochemistry of S-adenosylmethionine formation by yeast methionine adenosyltransferase. *Journal of the American Chemical Society*, 104(3), pp.871–872.

Pegg, A. & Williams-Ashman, H.G., 1969. Phosphate-stimulated breakdown of 5'-methylthioadenosine by rat ventral prostate. *The Biochemical journal*, 115(2), pp.241–247.

Peng, H. et al., 2014. Methionine adenosyltransferase 2B- GIT1 interplay activates MEK1-ERK1/2 to induce growth in human liver and colon cancer. *Hepatology*, 57(6), pp.2299–2313.

Peng, H. et al., 2015. Methionine adenosyltransferase 2B-GIT1 complex Serves as a scaffold to regulate Ras / Raf / MEK1 / 2 activity in human liver and colon cancer cells. *The American Journal of Pathology*, 185(4), pp.1135–1144.

Pérez Mato, I. et al., 2001. Biochemical basis for the dominant inheritance of hypermethioninemia associated with the R264H mutation of the MAT1A gene. A monomeric methionine adenosyltransferase with tripolyphosphatase activity. *Journal of Biological Chemistry*, 276(17), pp.13803–13809.

Perez-Leal, O. & Merali, S., 2012. Regulation of polyamine metabolism by translational control. *Amino Acids*, 42(2-3), pp.611–617.

Pérez-Mato, I. et al., 1999. Methionine adenosyltransferase S-nitrosylation is regulated by the basic and acidic amino acids surrounding the target thiol. *Journal of Biological Chemistry*, 274(24), pp.17075–17079.

Petoukhov, M. V. et al., 2012. New developments in the ATSAS program package for small-angle scattering data analysis. *Journal of Applied Crystallography*, 45(2), pp.342–350.

Petrossian, T.C. & Clarke, S.G., 2011. Uncovering the human methyltransferasome. *Molecular & Cellular proteomics*, 10(1), pp.25–37.

Porod, G., 1982. *Small angle X-ray Scattering* Glatter., Academic Press.

Premont, R.T. et al., 1998. beta2-Adrenergic receptor regulation by GIT1, a G protein-coupled receptor kinase-associated ADP ribosylation factor GTPase-activating protein. *Proceedings of the National Academy of Sciences of the United States of America*, 95(24), pp.14082–14087.

Premont, R.T. et al., 2000. The GIT family of ADP-ribosylation factor GTPase-activating. *Biochemistry*, 275(29), pp.22373–22380.

Prudova, A. et al., 2006. S-adenosylmethionine stabilizes cystathionine beta-synthase and modulates redox capacity. *Proceedings of the National Academy of Sciences of the United States of America*, 103(17), pp.6489–6494.

Putnam, C.D. et al., 2007. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics*, 40(3), pp.191–285.

Radaev, S., Li, S. & Sun, P.D., 2006. research papers A survey of protein – protein complex crystallizations research papers. *Acta Crystallographica Section D*, 62, pp.605–612.

Radaev, S. & Sun, P.D., 2002. research papers Crystallization of protein ± protein complexes research papers. *Applied Crystallogaphy*, 35, pp.674–676.

Della Ragione, F. & Pegg, A.E., 1983. Effect of analogues of 5'-methylthioadenosine on cellular metabolism. Inactivation of S-adenosylhomocysteine hydrolase by 5'-isobutylthioadenosine. *Biochemical Journal*, 210(2), pp.429–435.

Rai, R.M. et al., 1998. Impaired liver regeneration in inducible nitric oxide synthasedeficient mice. *Proceedings of the National Academy of Sciences of the United States of America*, 95(23), pp.13829–13834.

Raina, A., Tuomi, K. & Pajula, R.-L., 1982. Inhibition of the synthesis of polyamines and macromolecules by 5′-methylthioadenosine and 5′-alkylthiotubercidins in BHK21 cells. *Biochemical Journal*, 204(3), pp.697–703.

Ramani, K. et al., 2008. Leptin's mitogenic effect in human liver cancer cells requires induction of both methionine adenosyltransferase 2A and 2β. *Hepatology*, 47(2), pp.521–531.

Rao, K.N., Eagon, P.K. & Lombardi, B., 1982. Acute Hemorrhagic Pancreatic Necrosis in Mice Induction in Male Mice Treated With Estradiol. *American Journal of Pathology*, 109, pp.8–14.

Regland, B. et al., 1995. Homocysteinemia is a common feature of schizophrenia. *Journal of Neural Transmission - General Section*, 100(2), pp.165–169.

Reytor, E. et al., 2009. Conformational signals in the C-terminal domain of methionine adenosyltransferase I/III determine its nucleocytoplasmic distribution. *The FASEB Journal*, 23(10), pp.3347–3360.

Rhodes, G., 2006. *Crystallography made crystal clear: a guide for users of macromolecular models*, Elsevier Ltd.

Rupp, B., 2009. *Biomolecular crystallography: principles, practice, and application to structural biology*, Garland science.

Sanchez del Pino, M.M., Corrales, F.J. & Mato, J.M., 2000. Hysteretic behavior of methionine adenosyltransferase III: Methionine switches between two conformations of the enzyme with different specific activity. *Journal of Biological Chemistry*, 275(31), pp.23476–23482.

Sánchez-Góngora, E. et al., 1997. Interaction of liver methionine adenosyltransferase with hydroxyl radical. *The FASEB journal*, 11(12), pp.1013–1019.

Sarasin-Filipowicz, M. et al., 2008. Interferon signaling and treatment outcome in chronic hepatitis C. *Proceedings of the National Academy of Sciences of the United States of America*, 105(19), pp.7034–7039.

Scavetta, R.D. et al., 2000. Structure of RsrI methyltransferase, a member of the N6-adenine beta class of DNA methyltransferases. *Nucleic Acids Research*, 28(20), pp.3950–3961.

Schrödinger, LLC, 2010. *The {PyMOL} Molecular Graphics System, Version~1.3r1*,

Scott, J.M. et al., 1994. Effects of the disruption of transmethylation in the central nervous system: an animal model. *Acta Neurologica Scandinavica. Supplementum*, 154, pp.27–31.

Scott, J.M., 1992. Folate-vitamin B12 interrelationships in the central nervous system. *The Proceedings of the Nutrition Society*, 51(2), pp.219–224.

Shafqat, N. et al., 2013. Insight into S-adenosylmethionine biosynthesis from the crystal structures of the human methionine adenosyltransferase catalytic and regulatory subunits. *The Biochemical Journal*, 452(1), pp.27–36.

Sheffield, P., Garrard, S. & Derewenda, Z., 1999. Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expression and Purification*, 15, pp.34–39.

Shimizu-Saito, K. et al., 1997. Differential expression of S-adenosylmethionine synthetase isozymes in different cell types of rat liver. *Hepatology*, 26(2), pp.424–431.

Simile, M.M. et al., 2005. Chemopreventive N-(4-hydroxyphenyl)retinamide (fenretinide) targets deregulated NF-kB and Mat1A gene in the early stages of rat liver carcinogenesis. *Carcinogenesis*, 26(2), pp.417–427.

Skordi, E. et al., 2007. Analysis of time-related metabolic fluctuations induced by ethionine in the rat research articles. *Journal of Proteomic Research*, 6, pp.4572–4581.

Smith, R.C. & Salmon, W.D., 1965. Formation of S-adenosylethionine by ethionine-treated rats. *Archives of Biochemistry and Biophysics*, 111(1), pp.191–196.

Smythies, J., 2012. The role of abnormalities related to the one carbon cycle in depression and schizophrenia. *Neuroscience and Medicine*, 03(01), pp.101–106.

Söding, J., Biegert, A. & Lupas, A.N., 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research*, 33, pp.244–248.

Stamler, J.S. et al., 1997. (S)NO signals: Translocation, regulation, and a consensus motif. *Neuron*, 18(5), pp.691–696.

Stephens, R.S. et al., 1998. Genome sequence of an obligate intracellular pathogen of humans: chlamydia trachomatis. *Science*, 282(5389), pp.754–759.

Stipanuk, M.H., 2004. Sulfur amino acid metabolism: pathways for production and removal of homocysteine and cysteine. *Annual Review of Nutrition*, 24, pp.539–577.

Stolowitz, M.L. & Minch, M.J., 1981. S-adenosyl-L- methionine and S-adenosyl-L-homocysteine, an NMR study. *Journal of the American Chemical Society*, 103(2), pp.4–8.

Sullivan, D.M. & Hoffman, J.L., 1983. Fractionation and kinetic properties of rat liver and kidney methionine adenosyltransferase isozymes. *Biochemistry*, 22, pp.1636–1641.

Sun, C. et al., 2015. SimpleDSFviewer: a tool to analyse and view differential scanning fluorimetry data for characterising protein thermal stability and interactions. *PeerJ PrePrints*, 3, p.e1937.

Surtees, R., Leonard, J. & Austin, S., 1991. Association of demyelination with deficiency of cerebrospinal-fluid S-adenosylmethionine in inborn errors of methyl-transfer pathway. *The Lancet*, 338(8782-8783), pp.1550–1554.

Svergun, D.I., 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical Journal*, 76(6), pp.2879–2886.

Takusagawa, F., Kamitori, S. & Markham, G.D., 1996. Structure and function of S-adenosylmethionine synthetase: crystal structures of S-adenosylmethionine synthetase with ADP, BrADP, and PPi at 2.8 Å resolution,. *Biochemistry*, 35(8), pp.2586–2596.

Taylor, J.., Takusagawaa, F. & Markham, G.D., 1996. A chimeric active site loop of S-adenoylmethionine synthetase. *The FASEB Journal*, 10, p.970.

Taylor, J.C. & Markham, G.D., 2000. The bifunctional active site of S -adenosylmethionine synthetase. *Journal of Biological Chemistry*, 275(6), pp.4060–4065.

Taylor, J.C., Takusagawa, F. & Markham, G.D., 2002. The active site Loop of S -adenosylmethionine synthetase modulates catalytic. *Biochemistry*, 41, pp.9358–9369.

Thompson, J.D., Higgins, D.G. & Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22), pp.4673–4680.

Torres, L., López-Rodas, G., et al., 2000. DNA methylation and histone acetylation of rat methionine adenosyltransferase 1A and 2A genes is tissue-specific. *International Journal of Biochemistry and Cell Biology*, 32(4), pp.397–404.

Torres, L., Avila, M. a, et al., 2000. Liver-specific methionine adenosyltransferase MAT1A gene expression is associated with a specific pattern of promoter methylation and histone acetylation: implications for MAT1A silencing during transformation. *The FASEB journal*, 14(1), pp.95–102.

Tsukamoto, H. & Xi, X.P., 1989. Incomplete compensation of enhanced hepatic oxygen consumption in rats with alcoholic centrilobular liver necrosis. *Hepatology*, 9(2), pp.302–306.

Turner, C.E. et al., 1999. Paxillin LD4 motif binds PAK and PIX through a novel 95-kD ankyrin repeat, ARF-GAP protein: A role in cytoskeletal remodeling. *Journal of Cell Biology*, 145(4), pp.851–863.

Ubagai, T. et al., 1995. Molecular mechanisms of an inborn error of methionine pathway methionine adenosyltransferase deficiency. *Journal of Clinical Investigation*, 96(4), pp.1943–1947.

Varela-Rey, M. et al., 2011. Alcohol, DNA Methylation and Cancer. *Alcohol Research*, 35(1), pp.25–36.

Vázquez, M. et al., 2010. Evidence for an LKB1/AMPK/eNOS cascade regulated by HGF, S-adenosylmethionine and NO in hepatocyte proliferation. *Hepatology*, 49(2), pp.608–617.

Villanueva, J.A. & Halsted, C.H., 2004. Hepatic transmethylation reactions in micropigs with alcoholic liver disease. *Hepatology*, 39(5), pp.1303–1310.

Vinci, C.R. & Clarke, S.G., 2007. Recognition of age-damaged (R,S)-adenosyl-L-methionine by two methyltransferases in the yeast Saccharomyces cerevisiae. *Journal of Biological Chemistry*, 282(12), pp.8604–8612.

Vinci, C.R. & Clarke, S.G., 2010. Yeast, plants, worms, and flies use a methyltransferase to metabolize age-damaged (R,S)-AdoMet, but what do mammals do? *Rejuvenation Research*, 13(2-3), pp.362–364.

Wang, F. et al., 2014. Understanding molecular recognition of promiscuity of thermophilic methionine adenosyltransferase sMAT from Sulfolobus solfataricus. *The FEBS Journal*, 281(18), pp.4224–39.

Wang, Q. et al., 2008. Inhibition of hepatocelluar carcinoma MAT2A and MAT2beta gene expressions by single and dual small interfering RNA. *Experimental & Clinical Cancer Research*, 9, pp.1–9.

West, K. a. et al., 2001. The LD4 motif of paxillin regulates cell spreading and motility through an interaction with paxillin kinase linker (PKL). *Journal of Cell Biology*, 154(1), pp.161–176.

Wu, S.E. et al., 1983. Chiral instability at sulfur of S-adenosylmethionine. *Biochemistry*, 22(12), pp.2828–2832.

Xia, M. et al., 2010. Novel function and intracellular localization of methionine adenosyltransferase 2β splicing variants. *Journal of Biological Chemistry*, 285(26), pp.20015–20021.

Yamada, Y. et al., 1997. Initiation of liver growth by tumor necrosis factor: Deficient liver regeneration in mice lacking type I tumor necrosis factor receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 94(4), pp.1441–1446.

Yang, H. et al., 2008. Expression pattern, regulation, and functions of methionine adenosyltransferase 2beta splicing variants in hepatoma cells. *Gastroenterology*, 134(1), pp.281–91.

Yang, H. et al., 2013. Methionine adenosyltransferase 2B, HuR, and sirtuin 1 protein cross-talk impacts on the effect of resveratrol on apoptosis and growth in liver cancer cells. *Journal of Biological Chemistry*, 288(32), pp.23161–23170.

Yang, H. et al., 2001. Role of promoter methylation in increased methionine adenosyltransferase 2A expression in human liver cancer. *American Journal of Physiology. Gastrointestinal and Liver Physiology*, 280(2), pp.G184–G190.

Yang, H.-B. et al., 2015. Acetylation of MAT IIα represses tumour cell growth and is decreased in human hepatocellular cancer. *Nature Communications*, 6, p.6973.

Zeng, Z. et al., 2000. Cloning and functional characterization of the 5'-flanking region of human methionine adenosyltransferase 1A gene. *Biochemical Journal*, 346(Pt 2), pp.475–482.

Zhao, Z.S. et al., 2000. Coupling of PAK-interacting exchange factor PIX to GIT1 promotes focal complex disassembly. *Molecular and Cellular Biology*, 20(17), pp.6354–6363.