Structural and interactional studies of the cobalamin-dependent enzyme human methionine synthase

Douglas Ferreira

St Catherine's College

Supervisors: Dr Wyatt Yue, Dr Jon Elkins, Dr Jesse Coker

A thesis submitted for the degree of *Doctor of Philosophy* in Clinical Medicine at the Centre for Medicines Discovery Nuffield Department of Medicine, University of Oxford

Michaelmas term 2024

Declaration

I declare that no parts of this research or thesis have been submitted for another award, degree or diploma at any other university or learning institution. This thesis contains no other person's work, except where stated in text.

Acknowledgements

I would like to thank my supervisors, Wyatt Yue, Jon Elkins, and Jesse Coker, for all their support and guidance throughout my studies, always bringing their expertise to the development of this project. In particular, to Wyatt, who has been the best mentor a student could have and entrusted me with this challenging and rewarding project. Thank you also to the past and current members of the Yue Group, who have been amazing colleagues and friends, in particular to Thomas McCorvie, who has greatly helped and guided me through the cryo-EM world since the beginning of my DPhil. Thank you also to the Nuffield Department of Medicine for funding the DPhil work and to the members of the Centre for Medicines Discovery. Also, thank you to all the staff and students at Newcastle University who have welcomed and supported me during most of my DPhil.

I also thank Carol and Barbara for their unwavering support throughout my scientific career. You believed I could be here even before I did. Thank you to all the friends I have made throughout this journey in the UK, especially those from Brazil and Latin America, for always making me feel at home. Thank you to Nina, my friend for more than a decade who shares a similar journey that has allowed us to always meet and comfort each other in the toughest and happiest times, and to Fernanda for the support and many hours of calls during lonely times. Thank you to Julián for being the most supportive companion, best friend, and greatest partner I could ever ask even across the ocean, Si estás conmigo, todo es más fácil, todo es mejor.

Aos meus pais, Denise e Ernani, que sempre me incentivaram a continuar estudando e me deram, e continuam dando, todo o suporte necessário para que eu chegasse até aqui. Obrigado pelo exemplo de pessoas e por toda a dedicação que tiveram ao longo de toda a

minha vida. Aos meus irmãos, Rafael e Alexandre, por sempre serem companheiros e amigos, sempre presentes. Amo todos vocês.

Abstract

The enzyme methionine synthase (MTR) is a modular five-domain protein that utilises cobalamin (Cbl) as a cofactor to perform a methyl transfer from methyltetrahydrofolate (MeTHF) to homocysteine (Hcy), producing tetrahydrofolate (THF) and methionine (Met). The eventual oxidation of Cbl leads to the inactivation of MTR, which requires an electron from its partner protein methionine synthase reductase (MTRR) and a methyl group from S-adenosylmethionine (SAM) for its recovery. Deficiencies in MTR are related to developmental disorders, megaloblastic anaemia, and other health consequences in humans. Most of the structural and biophysical data on MTR are derived from its bacterial orthologue, which shows that it is a very dynamic protein, with different conformations acquired throughout its catalytic and recovery cycle. To address the lack of structural data on human MTR, and on studies focusing on MTR interactions, cryo-EM, biophysical assays, and AlphaFold were employed in this study. Using a recombinant full-length human MTR, novel high-resolution structures of the protein were acquired using cryo-EM. The structures of *apo* and Cbl-loaded MTR revealed the position of the conserved residue Tyr1177 and provided novel insights into the reactivation cycle of the human protein, revealing differences from its bacterial orthologue. Furthermore, with the use of AlphaFold Multimer coupled with biophysical assays, the complex formation of MTR with MTRR and MMADHC, which delivers Cbl to *apo* MTR, was studied. The models revealed a novel interaction interface in the MTR-MTRR complex, which supports the data suggesting the formation of a higher-order complex. The results presented herein provide answers to unsolved questions on the structural and interactional aspects of MTR, broadening the knowledge of the biology of this protein.

Abbreviations

ABCD4 - Adenosine triphosphate-binding cassette subfamily D member 4

AD - Activation domain

AF - AlphaFold

ATP - Adenosine triphosphate

B_{12} - Vitamin B_{12}

Cbl - Cobalamin

CD - Connecting Domain

CMD - Centre for Medicines Discovery

CNCbl - Cyanocobalamin

CNN - Convolutional neural networks

CO_2 - Carbon dioxide

CRISPR - Clustered regularly interspaced short palindromic repeats

CTF - Contrast transfer function

CV - Column Volume

CYPOR - Cytochrome P450 reductase

DMBI - 5,6-dimethhylbenzimidazole

DNA - Deoxyribonucleic acid

DSF - Differential Scanning Fluorimetry

DTT - Dithiothreitol

EM - Electron Microscopy

FAD - Flavin adenine dinucleotide

FADBD - Flavin adenine dinucleotide-binding domain

FMN - Flavin mononucleotide

FMNBD - Flavin mononucleotide-binding domain

Fol - Folate

FolBD - Folate-binding domain

FSC - Fourier shell correlation

FT - Flow-through

GF - Gel filtration

GPU - Graphics processing units

GSFSC - Golden Standard fourier shell correlation

GST - Glutathione S-transferase

GTP - Guanosine-5'-triphosphate

H_2 - Hydrogen

H_2O_2 - Hydrogen peroxide

HC - Haptocorrin

Hcy - Homocysteine

HcyBD - Homocysteine-binding domain

HEK293 - Human embryonic kidney cell line

HEPES - 2-(4-(2-hydroxyethyl)-1-piperazinyl)-ethanesulfonic acid

HF - High-fidelity

LMBD1 - Lipocalin-1 interacting membrane receptor domain-containing protein 1

MBP - Myelin basic protein

MDRP1 - Multidrug resistance protein 1

Met - Methionine

MetCbl - Methylcobalamin

MeTHF - Methyltetrahydrofolate

MMAA - Methylmalonic aciduria type A protein

MMAB - Methylmalonic aciduria Type B protein

MMACHC - Methylmalonic aciduria and homocystinuria type C protein

MMADHC - Methylmalonic aciduria and homocystinuria type D protein

MMUT - Methylmalonyl-CoA mutase

MS – Mass spectrometry

MSA - Multiple sequence alignments

MTR – Methionine synthase

MTRR - Methionine synthase reductase

MW – Molecular weight

N_2 - Nitrogen

NADPH - Nicotinamide adenine dinucleotide (reduced)

NADPHBD - Nicotinamide adenine dinucleotide (reduced)-binding domain

NEB - New England Biolabs

NS - Negative staining

NSBL - Newcastle University structure biology laboratory

NUPPA - Newcastle University protein and proteome analysis

O_2 - Oxygen

OHCbl - Hydroxocobalamin

PABA - Para-aminobenzoic acid

PAGE - Polyacrylamide gel electrophoresis

PAM - Protospacer adjacent motif

PBS - Phosphate-buffered saline

PCR - Polymerase chain reaction

PDB - Protein data bank

PEEK - Polyether ether ketone

PNK – Polynucleotide kinase

QTOF - Quadrupole time-of-flight

RMSD - Root mean square deviation

RNA - Ribonucleic acid

SAH - S-Adenosyl-L-homocysteine

SAM - S-Adenosyl methionine

SAXS - Small-angle X-ray scattering

SDS - Sodium dodecyl fulfate

SOMO - singly occupied molecular orbital

SPR - Surface plasmon resonance

TC - Transcobalamin

TCEP - Tris(2-carboxyethyl)phosphine

TEV - Tobacco etch virus

UV - Ultraviolet

YSBL - York structural biology laboratory

µL - Microliter

µM - Micromolar

Contents

List of Figures

List of Tables

1. Introduction

1.1 Vitamin B_{12}: a cofactor with 90 years of history

2024 is the 90th anniversary of the Nobel Prize in Physiology or Medicine, shared by George Whipple, George Minot, and William Murphy for their research on the cure for pernicious anaemia by a substance available in the raw liver, which was later found to be vitamin B_{12} (Cobalamin). It is also the 60th anniversary of the Nobel Prize in Chemistry for Dorothy Hodgkin's work on the crystal structure of various molecules, including vitamin B_{12}.

Vitamin B_{12}, or Cobalamin (Cbl), is one of the eight essential B vitamins in animals. Through being an enzyme cofactor, Cbl is necessary for DNA synthesis, cell replication, one-carbon metabolism, and normal erythropoiesis (Butola, Kute et al. 2020). After the first purification of Cbl in the 1940s by Shorb, Folker, and Todd, its structure was solved by Dorothy Hodgkin in 1955 (Hodgkin, Kamper et al. 1956, Hodgkin, Kamper et al. 1957). Often described as nature's most beautiful cofactor, Cbl has the most complex structure of all B vitamins (Figure 1.1A). Cbl is composed of a corrin ring containing a cobalt atom, which is positioned in the centre of the ring by four nitrogen bonds from pyrrole groups. All four pyrrole groups are in the same plane, connected to each other by a C-CH_3 and a C-H. Another ligand to cobalt is an N from the 5,6-dimethhylbenzimidazole (DMBI), located straight down from the cobalt and connecting back to the structure by a five-carbon sugar.

The interaction of the DMBI tail with the lower axial of the cobalt is called the 'base-on' state (Figure 1.1B), and it is the form of Cbl in solution. The active site of cobalt is situated right above (upper axial) the corrin ring. It can connect to different types of ligands, forming different Cbl types, including Cyanocobalamin (CNCbl), Methylcobalamin (MetCbl), Adenosylcobalamin (AdoCbl), and Hydroxycobalamin (OHCbl) (Banerjee 1999).

The cobalt metal (Co) in Cbl is usually in the oxidation state +3 (cob[III]alamin) and 6-coordinated with the four nitrogen in the corrin ring, the DMBI tail on the lower axis, and the ligand in the upper axis. MetCbl and AdoCbl are examples of Cbl in the Co +3 oxidation state and can be reduced to +2 or +1 via homolytic or heterolytic cleavage of the Co-C bond, respectively (Park and Brunold 2013). These changes in the oxidation state happen with the decrease of the coordination of Co, from six (cob[III]alamin) to five (cob[II]alamin) or four (cob[I]alamin) (Perry and Marques 2004).

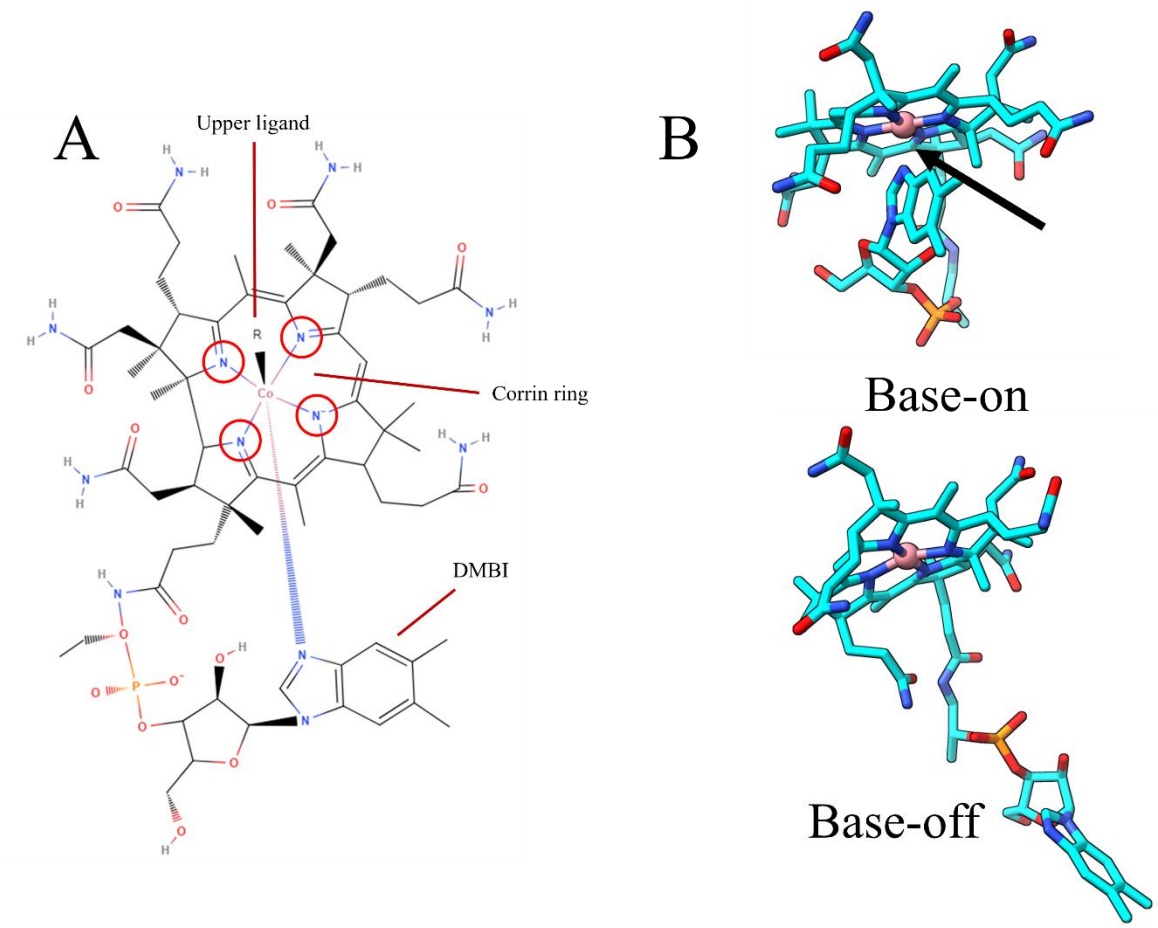

Figure 1.1: Cobalamin structure. A. Chemical structure of Cbl, indicating the localisation of the upper axis (R), corrin ring, and DMBI, covalently linked to the central Co atom. The four nitrogen bonds from pyrrole groups are identified by red circles. **B.** Cbl molecule in solution, with the DMBI group interacting with the lower ligand of Co (shown by a black arrow), in the "Base-on" state, and the "Base-off" state, found only when Cbl is bound to certain proteins.

Cbl is essential for humans and some plants as a vitamin, but they do not have the machinery to synthesise it. Cbl is only produced by some archaea and bacteria (Schneider

and Stroinski 1987), and humans acquire Cbl mainly through diet. It is knowsn that raw liver from pork, chicken, or beef contains high amounts of Cbl (Gille and Schmid 2015). Another significant source of Cbl is the milk from ruminants (cow, goat, and sheep) and its fermented form (cheese or yoghurt) (Raynal-Ljutovac, Lagriffoul et al. 2008). Although the concentration of Cbl per gram in milk is lower than in meat, milk and dairy products are consumed in high amounts in many populations (Matte, Britten et al. 2014), sufficient to meet the minimum intake requirement for Cbl (around 2.4 μg/day for adults). Besides animal sources, plants can also contain significant levels of Cbl, although most do not produce or require Cbl for normal functioning (Roth, Lawrence et al. 1996). Overall, the acquisition of Cbl by humans from food intake is mainly due to a symbiotic interaction between bacteria and ruminants or fish (Watanabe and Bito 2018). Still, Cbl can also be acquired via supplementation, which some countries require for people with vegan or vegetarian diets (Craig and Mangels 2009). The most used form of Cbl supplementation is CNCbl because it is more stable, cheap, and safe (Obeid, Fedosov et al. 2015), but it can also be done with MetCbl or OHCbl (Adams, Ross et al. 1971).

1.2 The complex network of cobalamin absorption, processing, and delivery

Cobalamin (Cbl) metabolism requires a complex multi-protein network to efficiently process, store, and deliver this essential vitamin to the two end-user enzymes in humans, namely Methionine Synthase (MTR) and Methylmalonyl-CoA Mutase (MMUT), as their cofactor. Although several proteins and enzymes involved in Cbl metabolism have been identified, significant gaps remain in our understanding of this sophisticated process.

1.2.1 Absorption and storage of Cbl

The absorption of Cbl begins with ingesting Cbl-containing food, where Cbl is typically bound to enzymes or carrier proteins. These carriers are degraded in the stomach,

releasing Cbl, which binds to haptocorrin (HC), a Cbl-binding protein in saliva that protects Cbl from degradation (Fedosov, Fedosova et al. 2007). In the duodenum, HC is degraded, and the neutralisation of the acidic pH facilitates the binding of Cbl to the gastric intrinsic factor (IF) produced by the gastric parietal cells (Quadros 2010). The IF-Cbl complex binds to the cubam receptor, which mediates the uptake of the IF-Cbl complex in the distal ileum via endocytosis (Tanner, Aminoff et al. 2003). In the lysosome, IF is degraded, and Cbl exits the cell via the multidrug resistance protein 1 (MDRP1) into the bloodstream, where it binds to transcobalamin (TC) (Beedholm-Ebsen, van de Wetering et al. 2010), the blood carrier for Cbl responsible for its cellular delivery (Nielsen, Rasmussen et al. 2012). Most Cbl is stored in the liver and can be excreted in the bile, undergoing enterohepatic circulation (Green, Jacobsen et al. 1982). Most Cbl in the bloodstream is bound to HC (70-90%), with only 10-30% bound to TC (Stabler 2020). Nonetheless, deficiency in TC can cause severe megaloblastic anaemia, methylmalonic aciduria, and homocystinuria, while deficiency in HC typically only results in lower serum levels of Cbl (Huemer and Baumgartner 2019).

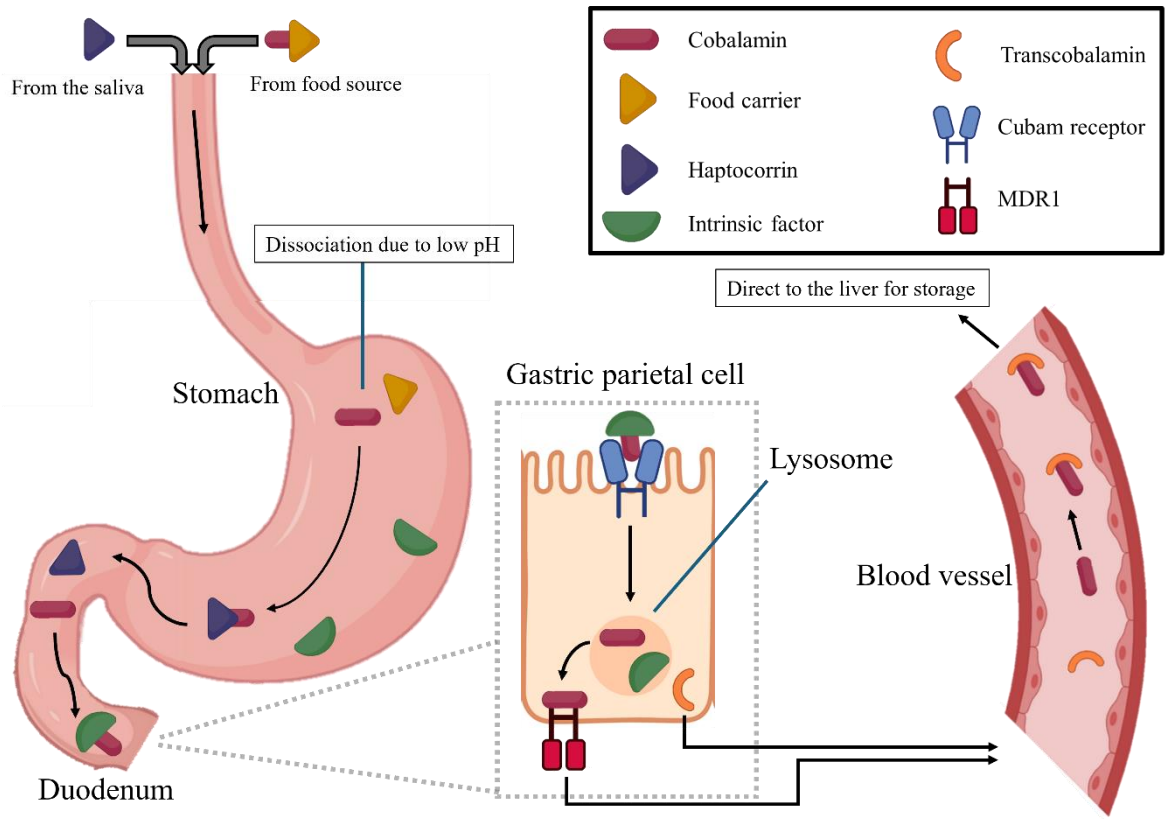

Figure 1.2: Absorption and circulation of Cbl. Following the intake of Cbl from a food source, it is released from its carrier in the stomach and binds to haptocorrin, which protects it from the acidic pH. In the duodenum, haptocorrin is degraded due to the pH change, and Cbl binds to intrinsic factor, which allows Cbl to be internalised by the gastric parietal cells via the cubam receptor. In the lysosome of the cell, the intrinsic factor is degraded, and Cbl exits the cell via MDR1 into the bloodstream, where it binds to transcobalamin. Illustration created with BioRendeR and Microsoft PowerPoint.

1.2.2 Intracellular processing and transport of Cbl to the end-user proteins

The Cbl carrier transcobalamin (TC) is responsible for transporting Cbl to the cells. Cells absorb TC-Cbl via the CD320 receptor, which is present in virtually all cell types. Knock-out studies on the CD320 receptor have suggested that other receptors for TC-Cbl may exist but have not been identified yet (Lai, Nakayama et al. 2013). After endocytosis, Cbl proceeds through the intracellular branch of its metabolism (Figure 1.3). In the lysosome, TC is degraded via proteasomal degradation, and Cbl is transported to the cytosol via two membrane proteins: lipocalin-1 interacting membrane receptor domain-containing protein 1 (LMBD1) and adenosine triphosphate-binding cassette subfamily D member 4 (ABCD4) (Rosenblatt, Hosack et al. 1985, Coelho, Kim et al. 2012). The protein ABCD4 can transport Cbl out of the lysosome alone *in vitro* (Kitai, Kawaguchi et al. 2021), but

LMBD1 is required to recruit ABCD4 to the lysosome and protect it from degradation (Fettelschoss, Burda et al. 2017).

After entering the cytosol, Cbl binds to the protein MMACHC (Methylmalonic Aciduria and Homocystinuria Type C Protein), which is known to interact with both ABCD4 and LMBD1 (Deme, Hancock et al. 2014). MMACHC can interact with any type of Cbl, such as MetCbl, AdoCbl, and CNCbl (Koutmos, Gherasim et al. 2011). Cbl binds to MMACHC in the 'base-off' state, meaning that the DMBI tail is not interacting with the corrin ring, as shown in Figure 1.1B. This would disrupt the interaction of the DMBI tail with the lower axis of Cbl, which leads to the weakening of the upper axial ligand of cobalt. Following this binding, MMACHC can process Cbl by removing the upper axial group through dealkylation (for MetCbl or AdoCbl) (Kim, Hannibal et al. 2009) or decyanation (for CNCbl), generating cob[II]alamin (Kim, Gherasim et al. 2008). The dealkylation and decyanation processes performed by MMACHC require glutathione, which interacts with the conserved residues Arg161, Arg206, and Arg230, acting as an electron donor (Ruetz, Shanmuganathan et al. 2017).

After being processed by MMACHC, Cbl in the form of cob[II]alamin is delivered to either methionine synthase (MTR) in the cytosol or methylmalonyl-CoA mutase (MMUT) in the mitochondria via the protein MMADHC (Methylmalonic Aciduria and Homocystinuria Type D Protein) (Suormala, Baumgartner et al. 2004). Upon forming a heterodimer with MMACHC, MMADHC coordinates its residue Cys261 with the upper axis of cob[II]alamin through a Co-S bond, forming thiolato-cob[III]alamin (Li, Mascarenhas et al. 2020). MMADHC can interact with Cbl and deliver it to MTR *in vitro*, but the availability of MMACHC-free Cbl in cells is very low. Therefore, the interaction and processing of any Cbl by MMACHC is necessary for MMADHC binding to Cbl(Mascarenhas, Guha et al. 2023). The exact mechanism of delivery of Cbl to either MTR in the cytosol or MMUT in

the mitochondria is not yet fully elucidated. It is suggested that the heterodimer MMACHC-MMADHC interacts with MTR for cytosolic delivery (Bassila, Ghemrawi et al. 2017) or is directed to the mitochondria due to the mitochondrial leader sequence on the N-terminus of MMADHC (Stucki, Coelho et al. 2012).

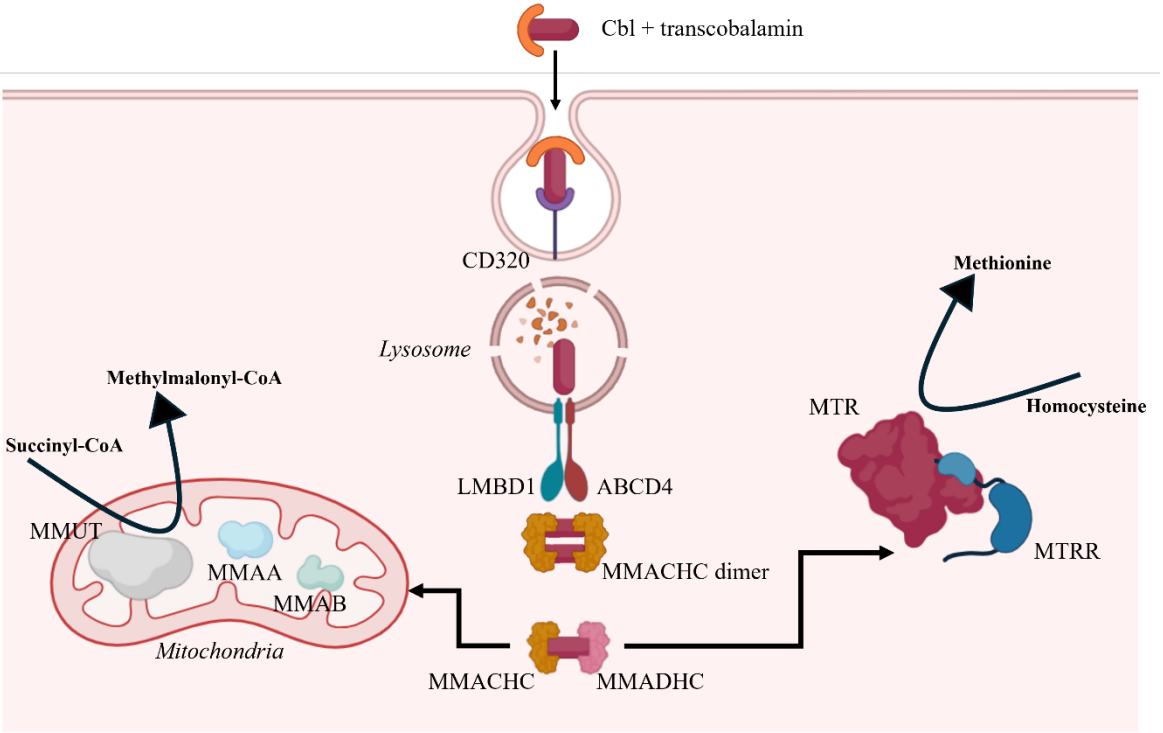

Figure 1.3: Intracellular metabolism of Cbl. The complex TC-Cbl enters the cells via endocytosis through the CD320 receptor, which is present in all cell types. In the lysosome, TC is degraded, and Cbl exits into the cytosol through LMBD1 and ABCD4, immediately binding to a dimer of MMACHC. After MMACHC removes the upper axial ligand of Cbl, one protomer of the MMACHC homodimer will dissociate, and MMADHC will form a heterodimer with the Cbl-carrier MMACHC. MMADHC interacts with Cbl via the coordination of its Cys261 sulphur with Cbl's upper axial ligand, and the protein complex MMACHC-MMADHC is responsible for delivering Cbl to either MMUT in the mitochondria or MTR in the cytosol. Illustration created with BioRendeR and Microsoft PowerPoint.

1.2.3 Biology of the end-user proteins of cobalamin: MMUT and MTR

In the mitochondria, the end-user protein MMUT utilises AdoCbl as a co-factor and needs two partner proteins to process and deliver Cbl from MMACHC-MMADHC. After MMADHC carrying Cbl enters the mitochondria, ATP-cobalamin adenosyltransferase (also known as Methylmalonic Aciduria Type B Protein, MMAB) processes Cbl in an ATP-dependent reaction (Mascarenhas, Ruetz et al. 2020). MMAB receives Cbl from MMACHC-

MMADHC in the form of four-coordinate cob[II]alamin (base-off, no upper ligand), and then converts it to cob[I]alamin and transfers the 5'-deoxyadenosine group from ATP, forming AdoCbl (Padovani, Labunska et al. 2008). AdoCbl is transferred to MMUT from MMAB in a reaction facilitated by the protein MMAA (Methylmalonic Aciduria Type A Protein), creating a possible three-way complex (Plessl, Bürer et al. 2017). MMAA promotes the replacement of the inactive Cbl of MMUT to AdoCbl in the presence of the hydrolysis of GTP, restoring MMUT catalytic activity (Mori and Toraya 1999).

MMUT through its AdoCbl cofactor catalyses the isomerisation of methylmalonyl-CoA, derived from the metabolism of branched-chain amino acids (methionine, valine, threonine, and isoleucine), cholesterol, and odd-chain fatty acids, to form succinyl-CoA (EC 5.4.99.2) (Forny, Bonilla et al. 2023). During the MMUT catalytic cycle, AdoCbl (Co in the 3+ state) can be oxidised into the inactive cob[II]alamin (Co in 2+ state), which is off-loaded onto MMAB with the aid of MMAA. MMAB then converts cob[II]alamin into functional AdoCbl and transfers it back to MMUT, allowing the restart of its catalytic cycle (Ruetz, Campanello et al. 2019). In this manner, MMAA functions as a GTP-dependent chaperone to MMUT, assisting in the maintenance of its activity by restoring the co-factor to AdoCbl.

MTR is the only other human protein that uses Cbl as a cofactor (in the form of MeCbl) and is located in the cytosol. Unlike MMUT, which requires two intermediate proteins in the mitochondria to receive and process Cbl, the MMACHC-MMADHC heterodimer can deliver Cbl directly to MTR, in a process that involves two adjacent Cys residues of MMADHC. During this process, MMADHC, which coordinates its Cys261 with the upper axial ligand of Cbl forming thiolato-cob[III]alamin, interacts with MTR. Upon interaction, MMADHC's the S_N2 attacks Cys262 to the Co-S bond transfers an electron pair onto Cbl, displacing and depositing it onto MTR as cob[I]alamin (Mascarenhas, Guha et al. 2023). The highly reactive cob[I]alamin is then methylated by S-adenosyl methionine

(SAM) in the activation domain of MTR, generating the active form of Cbl used by MTR, MetCbl.

MTR (EC 2.1.1.13) catalyses the methylation of homocysteine (Hcy) to methionine (Met) and the demethylation of 5-methyltetrahydrofolate (MeTHF) to tetrahydrofolate (THF), with methylcobalamin (MetCbl) serving as a mediator for the methyl transfer reactions. Due to its involvement with the abovementioned substrates and products, MTR is positioned between two critical metabolic pathways—the Folate and SAM cycles—acting as a crucial link between them (Matthews, Koutmos et al. 2008). Like AdoCbl, MetCbl also undergoes different oxidation states and can be oxidised into its inactive form, cob[II]alamin (Wolthers and Scrutton 2009). The reactivation of cob[II]alamin into MetCbl occurs in situ, in the presence of SAM and an electron, which is provided by another protein, methionine synthase reductase (MTRR) (Olteanu and Banerjee 2001). After this reactivation process known as reductive methylation, MTR can restart its catalytic cycle with a functional MetCbl.

1.3 Disorders of intracellular cobalamin metabolism

Inherited Mutations in the genes encoding proteins involved in intracellular cobalamin metabolism can cause a variety of rare metabolic disorders, with the resulting phenotypes and ages of onset depending on the specific proteins affected. Any part of the Cbl metabolism can be impacted, potentially leading to conditions such as methylmalonic acidemia (MMA), homocystinuria (HC), or combined methylmalonic acidemia and homocystinuria (MMA-HC) (Figure 1.4) (Huemer and Baumgartner 2019). Understanding the structure of proteins involved in disorders of intracellular cobalamin metabolism, and how they interact, is crucial for uncovering the structural basis of these disorders. Such conditions are often caused by alterations in specific protein residues, which can disrupt

function and lead to metabolic imbalances. Mutations affecting distinct genes involved in Cbl metabolism define the "complementation group" to which these mutations belong, with each group corresponding to a specific gene. These complementation groups are named *cblA* through *cblJ*, where the prefix "cbl" is followed by a capital letter representing the affected gene.

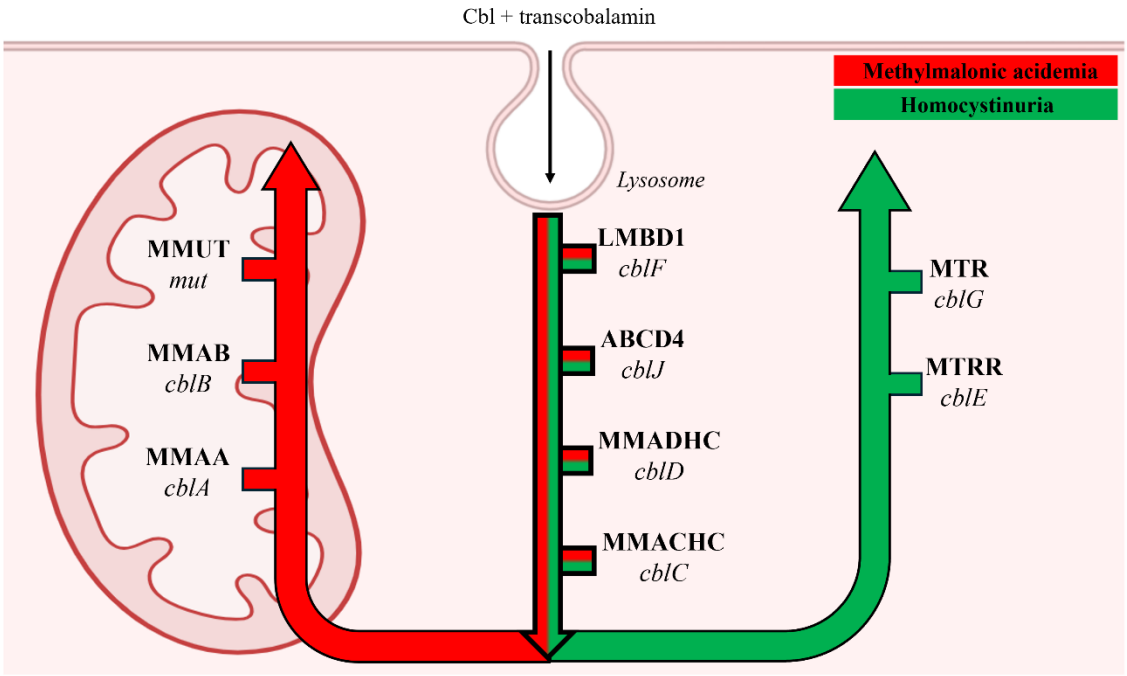

Figure 1.4: Gene mutations within the intracellular cobalamin metabolism, the disorder caused, and the biochemical phenotype. The map of the proteins related to the intracellular metabolism of Cbl, with the complementation group (identified as *mut* or *Cbl+identification letter*). The biochemical phenotypes are colour-coded: green for homocystinuria, red for methylmalonic acidemia, and red and green for the combination of both. Illustration created with BioRendeR and Microsoft PowerPoint.

Methylmalonic acidemia is characterised by the elevation of methylmalonic acid levels in the blood and urine. It occurs due to the failure to convert methylmalonyl-CoA into succinyl-CoA, the process catalysed by MMUT, which utilises AdoCbl as a cofactor (Manoli, Sloan et al. 2016). Homocystinuria is caused by high levels of homocysteine in the blood and urine, which MTR converts to methionine using MetCbl as a cofactor (Watkins, Ru et al. 2002). This section summarises the disorders of intracellular cobalamin metabolism and the genes involved that have been identified so far.

1.3.1 Methylmalonic acidemia: AdoCbl deficiency

Four complementation groups are involved in methylmalonic acidemia, caused either by the impaired delivery of Cbl to the mitochondria or by defects in the activity of MMUT. These complementation groups *cblA, cblB, mut,* and *cblD* refer to mutations of genes that encode MMAA, MMAB, MMUT, and MMADHC, respectively (Manoli, Sloan et al. 2016). Since MMAA is responsible for adding the adenosyl group to Cbl, while MMAB modulates the transfer of Cbl to MMUT (Section 1.2.2), any impairment in these proteins due to their gene mutations will lead to a lack of cofactor for MMUT, resulting in the accumulation of methylmalonyl-CoA (Padovani, Labunska et al. 2008).

Specific mutations in the *cblD* complementation group, of the gene encoding MMADHC, can cause methylmalonic acidemia (Coelho, Suormala et al. 2008). The N-terminus of MMADHC contains a mitochondrial leader sequence, which directs the MMACHC-MMADHC heterodimer, containing Cbl, to the mitochondria (Stucki, Coelho et al. 2012) (Section 1.2.2). Mutations in the 5' region of the *MMADHC* gene impair the delivery of Cbl to the mitochondria. However, they do not affect its delivery to MTR, causing disruption only in the mitochondrial branch (Froese and Gravel 2010).

The complementation group *mut* is characterised by mutations in the *MUT* gene, which encodes MMUT (Lempp, Suormala et al. 2007). This group is characterised by defects in the mitochondrial end-user protein, leading to the accumulation of the MMUT substrate methylmalonyl-CoA. Depending on the gene region affected by the mutation, the complementary group can be classified as mut^{-} when some residual MMUT activity is detected or mut^{0} when no activity is detectable (Ledley and Rosenblatt 1997).

1.3.2 Combined methylmalonic acidemia and homocystinuria: AdoCbl and MetCbl deficiency

Combined methylmalonic acidemia and homocystinuria (MMA-HC) is caused by defects occurring upstream of the division of Cbl metabolism into cytosolic and mitochondrial branches. The *cblF*, *cblJ*, *cblC*, and *cblD* complementation group of disorders are associated with MMA-HC due to defects in LMBD1, ABCD4, MMACHC, and MMADHC, respectively.

Since LMBD1 and ABCD4 are responsible for transporting Cbl out of the lysosome (Kitai, Kawaguchi et al. 2021) (Section 1.2.2), mutations in these proteins render Cbl unavailable for use by either MMUT or MTR. Furthermore, when the mutation occurs in the 3' region or the middle of the *MMADHC* gene (encoding the C-terminus of the protein), the delivery of Cbl is impaired in both the mitochondrial and cytosolic branches, preventing MMUT and MTR from receiving their cofactor (Froese and Gravel 2010).

The *cblC* defect (due to mutations in the *MMACHC* gene) is the most prevalent disorder of intracellular Cbl metabolism. It is caused by MMACHC's inability to process Cbl before it is delivered to its end-user proteins (Lerner-Ellis, Anastasio et al. 2009). In addition to c*blC* defect, *cblX* and *cblX-like* disorders also cause MMA-HC due to mutations in the genes *HCFC1* (which causes c*blX*) and *THAP11* or *ZNF143* (which causes c*blX-like* disorders). These genes encode proteins that lead to the downregulation of MMACHC expression (Yu, Sloan et al. 2013, Pupavac, Watkins et al. 2016, Quintana, Yu et al. 2017), therefore mutations lead to decrease of MMACHC proteins in the cell.

1.3.3 Homocystinuria: MetCbl deficiency

Defects in proteins within the cytosolic branch of Cbl metabolism cause homocystinuria. Depending on whether the affected gene is *MMADHC*, *MTRR*, or *MTR*, the

condition can be classified as c*blD*, c*blE*, or *cblG* complementation groups, respectively. The *CblG* defect is characterised by low or no activity of MTR due to its impaired function (Watkins, Ru et al. 2002), (Drummond, Huang et al. 1993). Impairments in MTRR, whose primary role is reactivating MTR after every 200-1000 cycles due to Cbl oxidation, also cause homocystinuria (c*blE* defect), as inactivated MTR cannot regain its activity without MTRR (Watkin and Rosenblatt 1989). Regarding the *CblD* defect, it is suggested that mutations in the 3' region of the MMADHC gene may cause homocystinuria alone without impairing the delivery of Cbl to the mitochondrial branch (Froese and Gravel 2010).

1.4 The cytosolic end-user protein Methionine Synthase (MTR)

Methionine Synthase (full name: 5-methyltetrahydrofolate-homocysteine methyltransferase, EC 2.1.1.13) is an enzyme whose bacterial orthologue (known as MetH in *E. coli*) has been extensively studied to gain insights into the structure and function of this dynamic protein (Guéant, Guéant-Rodriguez et al. 2022). The *E. coli* and human MTR enzymes share 55% sequence identity, and their domains are conserved. As a result, the bacterial enzyme serves as a valuable model for understanding the function of the human protein due to their high degree of identity and structural similarity.

MTR is positioned at the convergence of the folate and SAM cycles, playing a crucial role in preventing the accumulation of Hcy, which, in excess, can lead to homocystinuria and other toxic effects. MTR recycles Hcy into Met by utilising MeTHF as a methyl donor. The Met produced by this reaction is not only essential as a protein precursor but also as a substrate for the biosynthesis SAM, a key methyl donor in numerous cellular processes (Broderick, Duffus et al. 2014).

SAM is a primary methyl donor, not only for the reactivation of MTR but also in a variety of biochemical processes, including DNA and RNA methylation, thereby regulating

gene expression (Evans, Huddler et al. 2002). The production of Met by MTR ensures a continuous supply of SAM, which is vital for maintaining cellular physiology. During these processes, SAM is converted to S-adenosylhomocysteine (SAH), which is subsequently hydrolysed into Hcy and adenosine. Hcy is then remethylated to Met, regenerating SAM. Because SAM is synthesised directly from Met, Hcy methylation cannot be carried out by SAM-dependent methyltransferases, as this would lead to a futile cycle (Salbaum and Kappen 2012). Instead, MTR carries out this methylation, using a methyl group from MeTHF, thereby linking the SAM cycle to the folate cycle via the methionine synthase reaction (Matthews, Koutmos et al. 2008).

This section explores the catalytic and reactivation cycles of MTR, as well as the conformational changes the enzyme undergoes during these processes.

1.4.1 The anatomy and function of MTR

The human MTR is a 140 kDa enzyme (1265 residues) composed of five modular domains catalysing three distinct methylation reactions. The first domain at the N-terminus is the homocysteine binding domain (HcyBD, aa Leu18-Val338), followed by the folate binding domain (FolBD, aa Val372-Thr652), which together form the N-terminal half of the protein. The C-terminal half of MTR comprises the Cap domain (aa Gln662-Asn762), the cobalamin binding domain (CblBD, aa Gln772-Glu907), and the activation domain (AD, aa Ser923-Asp1265), which contains a binding site for S-adenosylmethionine (SAM) (Evans, Huddler et al. 2004). The main linkers are located between the FolBD and the Cap (aa Gln653-Ile661, Linker 1) and between the CblBD and AD (aa Glu908-Glu922, Linker 2).

The flexibility of MTR allows it to perform its three methylation reactions through different conformational changes between the CblBD with the HcyBD, FolBD, or AD. These include the conversion of homocysteine (Hcy) into methionine (Met) in the HcyBD via the

methyl group of MetCbl, the remethylation of Cbl to MetCbl using the methyl group from methyltetrahydrofolate (MeTHF) in the FolBD, which releases tetrahydrofolate (THF), and the remethylation of the non-productive oxidised cob[II]alamin via SAM in the activation domain (AD) and MTRR (Jarrett, Huang et al. 1998) (Figure 1.5).

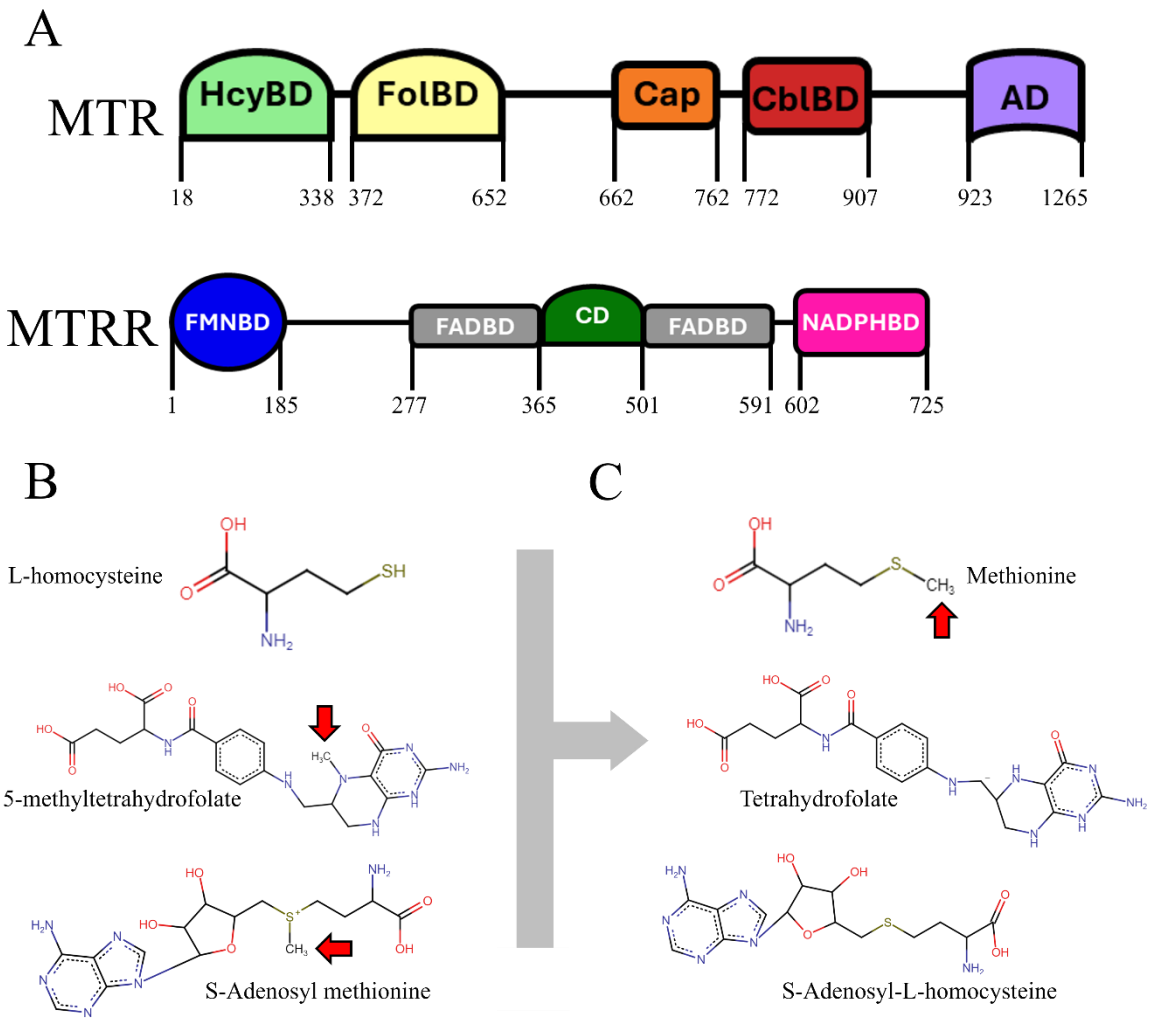

Figure 1.5: MTR domains division and its substrates and products. A. Domains division of the human MTR and MTRR, with the approximate residues' boundaries. **B**. List of substrates used by MTR and C. their respective products. The methyl groups are identified by red arrows.

The overall activity of MTR can be divided into two cycles: the catalytic and reactivation cycles. The catalytic cycle is responsible for the methylation of Hcy and

remethylation of Cbl via MeTHF. The reactivation cycle occurs via an interaction with MTRR for the reductive methylation of cob[II]alamin (Jarrett, Goulding et al. 1997).

1.4.2 The catalytic cycle of MTR in methionine synthesis

During the various steps of the catalytic cycle (Figure 1.6), MTR undergoes different conformational changes to facilitate the interaction between the cobalamin binding domain (CblBD) and both the homocysteine binding domain (HcyBD) and folate binding domain (FolBD). MetCbl, with a Co oxidation state of +3 (cob[III]alamin), is reduced to cob[I]alamin after the methyl group from its upper axial position is transferred to homocysteine (Hcy) (Figure 1.6, step A). The CblBD then rapidly interacts with the FolBD, allowing the highly reactive cob[I]alamin to be remethylated by methyltetrahydrofolate (MeTHF) (Figure 1.6, step B). The newly generated MetCbl can restart the cycle (Bandarian, Pattridge et al. 2002). Throughout the entire catalytic cycle, the lower axial ligand of the cobalt ion in Cbl interacts with His785 of MTR, referred to as the 'His-on' state. The His-on state is made possible because, when MMACHC interacts with Cbl, it changes Cbl from the Base-on to the Base-off state (Figure 1.1B), freeing the lower axial position of the cobalt ion (Co) from its previous interaction.

The CblBD can continue alternating between the HcyBD and the FolBD for approximately 200-1000 cycles. However, due to the high instability of cob[I]alamin, it is easily oxidised into its non-productive form, cob[II]alamin (Jarrett, Huang et al. 1998). When Cbl is oxidised to cob[II]alamin, the CblBD interacts with the Cap domain, placing Cbl in a protective "Cap-on" state (Drennan, Huang et al. 1994, Jarrett, Hoover et al. 1998) (Figure 1.6, step C). In this Cap-on state, MTR is inactive and must undergo the reactivation cycle for the reductive methylation of Cbl to restore its activity.

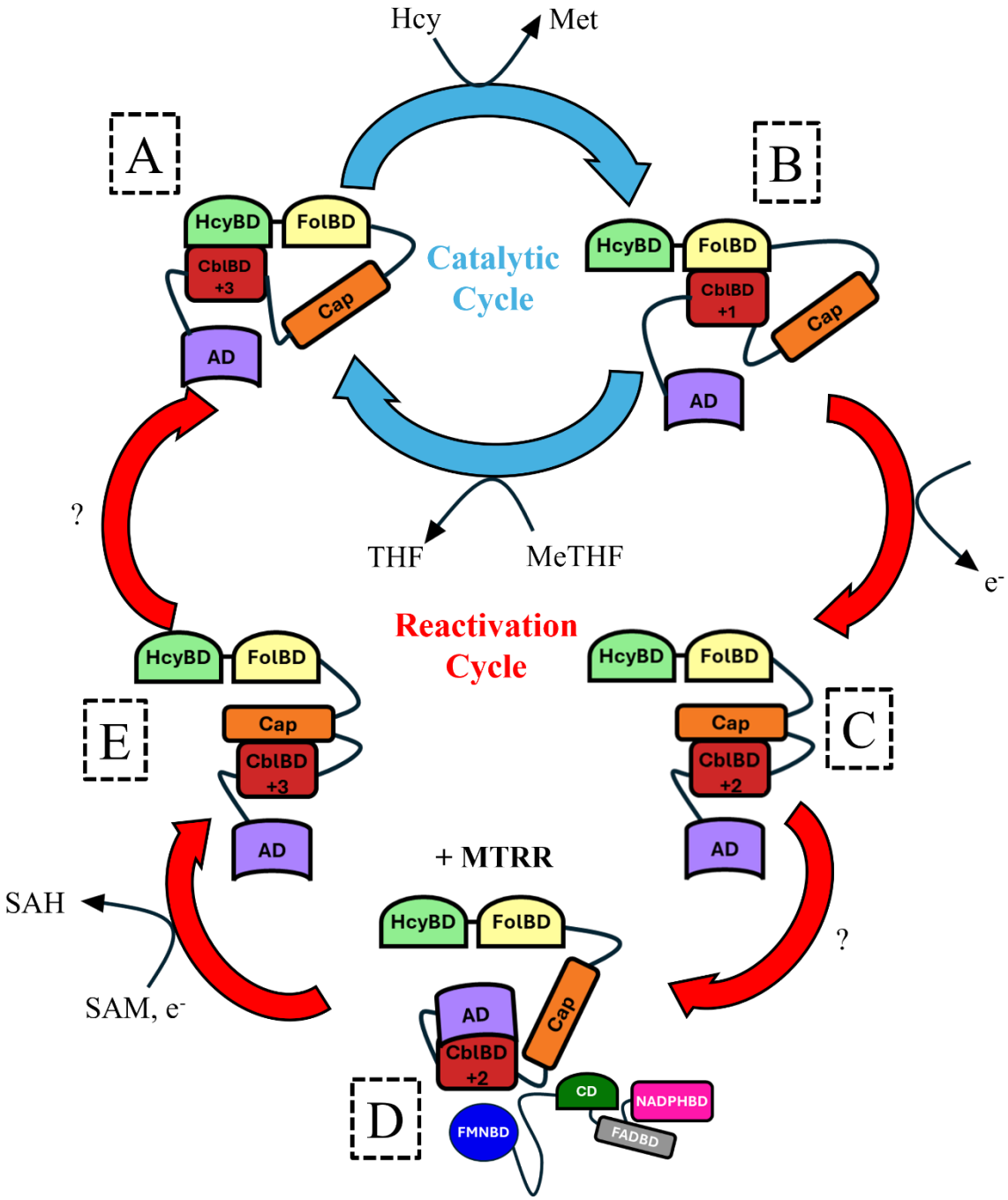

Figure 1.6: Scheme of the conformational changes of MTR during the catalytic and reactivation Cycles. **A.** The catalytic cycle involves the methylation of Hcy into Met through the interaction of the CblBD with the HcyBD. **B.** Cbl is remethylated by MeTHF, releasing THF through the interaction of the CblBD with the FolBD. **C.** When Cbl is oxidised to cob[II]alamin, MTR transitions to the Cap-on state in the reactivation cycle, and upon an unknown trigger, it goes to the activation conformation. **D**. In the activation conformation, the CblBD interacts with the AD for reductive methylation with the help of methionine synthase reductase (MTRR). The electron provided by MTRR and the SAM from the AD convert cob[II]alamin back to MetCbl, and MTR enters a resting state with the Cap domain protecting the CblBD. **E.** Upon an unknown trigger, MTR switches back to the catalytic cycle from the resting state.

1.4.3 Process and reactivation dynamics of MTR: reactivation cycle

To initiate the reactivation cycle, the Cap-on inactive MTR must undergo a conformational change, though the exact trigger for this change remains unknown. This 'activation conformation' is characterised by the interaction of CblBD with the AD (Figure 1.6 step D), which exposes cob[II]alamin to the solvent (Watkins, Wang et al. 2023). In this conformation, MTR can interact with its partner enzyme, methionine synthase reductase (MTRR). The FMN binding domain (FMNBD) of MTRR is connected to its other domains—FAD binding domain (FADBD), connecting domain (CD), and NADPH binding domain (NADPHBD)—by a long hinge (Wolthers and Scrutton 2007), enabling FMNBD to interact with the AD of MTR (Zhang, Liu et al. 2020).

In the reactivation conformation, cob[II]alamin alternates between the His-on (Figure 1.7, B) and His-off states(Figure 1.7, C), predominantly remaining in the His-on state (Jarrett, Hoover et al. 1998). In bacterial MTR, the recovery of Cbl is performed by a generic flavodoxin, which shifts the cobalamin population mostly to the His-off state, resulting in a five-coordinate species with a water molecule interacting with the upper axial position of cobalamin (Figure 1.7, C). When SAM binds to the activation domain (AD), the repositioning of Tyr1139 in bacterial MTR displaces the water molecule from cobalamin, lengthening the Co-O bond and forming a four-coordinate cob[II]alamin species (Figure 1.7, D). This four-coordinate form has a higher reduction potential, allowing cobalamin to be reduced by the electron from the partner protein (Figure 1.7, E). While this mechanism in human MTR is not fully elucidated, it is proposed that cobalamin must also be in a four-coordinate His-off state to receive an electron from the FMN binding domain (FMNBD) of MTRR (Figure 1.7 D) (Koutmos, Datta et al. 2009).

Upon receiving an electron, cob[II]alamin is reduced to cob[I]alamin (Figure 1.7 E), which is highly reactive and can easily be remethylated to MetCbl by SAM (Figure 1.7, F),

positioned in its binding pocket near the upper axial of the Co atom (Bandarian, Pattridge et al. 2002). MetCbl then reverts to the His-on state (Figure 1.7 G) before undergoing another conformational change back to the Cap-on state (Figure 1.7 H) (Watkins, Wang et al. 2023). The exact mechanisms by which MetCbl transitions to the His-on state or what triggers the conformational change to the Cap-on state remain unknown.

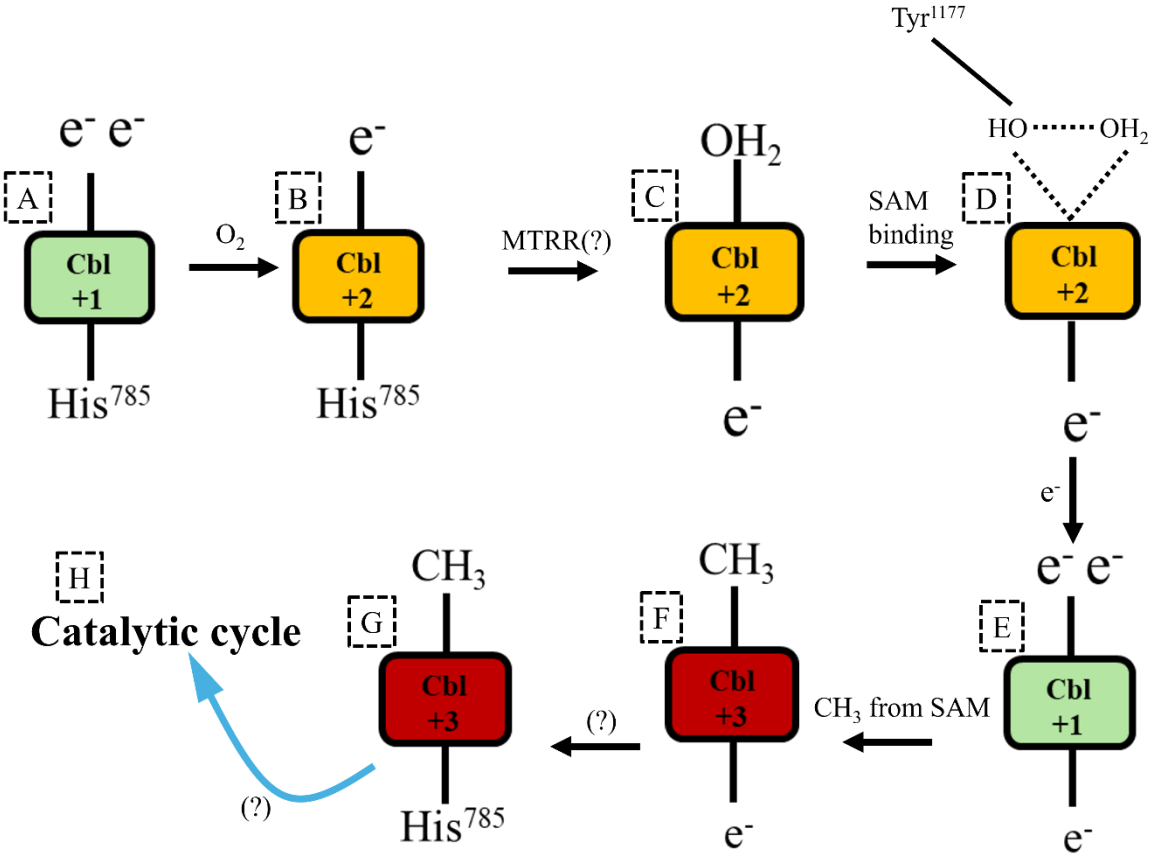

Figure 1.7: When cob[I]alamin is oxidised to cob[II]alamin, MTR undergoes reactivation to regenerate MetCbl. Upon the oxidation of His-on cob[I]alamin (**A**), the transition of cob[II]alamin His-on (**B**) to cob[II]alamin His-off (**C**), is triggered by flavodoxin binding in bacterial MTR. The five-coordinate cob[II]alamin His-off (**C**) then transitions to a four-coordinate cob[II]alamin His-off (**D**) due to the lengthening of the bond between the upper axial cobalt (Co) and a water molecule, a shift induced by SAM binding to the activation domain (AD). The four-coordinate cob[II]alamin His-off (**D**) then receives an electron from the partner protein, converting it back to cob[I]alamin His-off (**E**). This cob[I]alamin His-off is subsequently remethylated by SAM, forming MetCbl in the His-off state (**F**). The mechanisms underlying the transition of MetCbl from the His-off to MetCbl His-on (**G**), as well as how this transition induces the conformational change to the resting state for the catalytic cycle (**H**), remain unknown in both human and bacterial MTR.

1.5 Key protein-protein interactions in MTR

As shown in section 1.1, The intracellular metabolism of Cbl is characterised by a network of protein-protein interactions essential for its processing, delivery, and utilisation. Beyond the complex catalytic and reactivation cycles described above, the MTR protein is integrated into this network of interactions, which are crucial for its proper function (Bassila, Ghemrawi et al. 2017). This section explores the two key interactions involved in delivering Cbl to MTR and maintaining its activity, from a protein structure perspective.

1.5.1 Interaction of MTR with MMACHC and MMADHC for Cbl delivery

As introduced in section 1.2.2, The delivery of Cbl to *apo* MTR occurs by interacting with the heterodimer formed between MMADHC and MMACHC (Figure 1.8) (Deme, Miousse et al. 2012, Gherasim, Hannibal et al. 2013). MMADHC interacts with MMACHC after Cbl processing (Gherasim, Hannibal et al. 2013, Froese, Kopec et al. 2015), and defects in either protein impair MTR activity due to a lack of MetCbl (Huemer, Diodato et al. 2017). The interaction between MMACHC and MMADHC is stabilised by the MMADHC loop Thr182-Thr189 and Cys261. The Thr182-Thr189 loop interacts with MMACHC and the acetamide side-chain of the Cbl corrin ring, while Cys261 coordinates with the upper axial ligand of Cbl (Li, Mascarenhas et al. 2020, McCorvie, Ferreira et al. 2023).

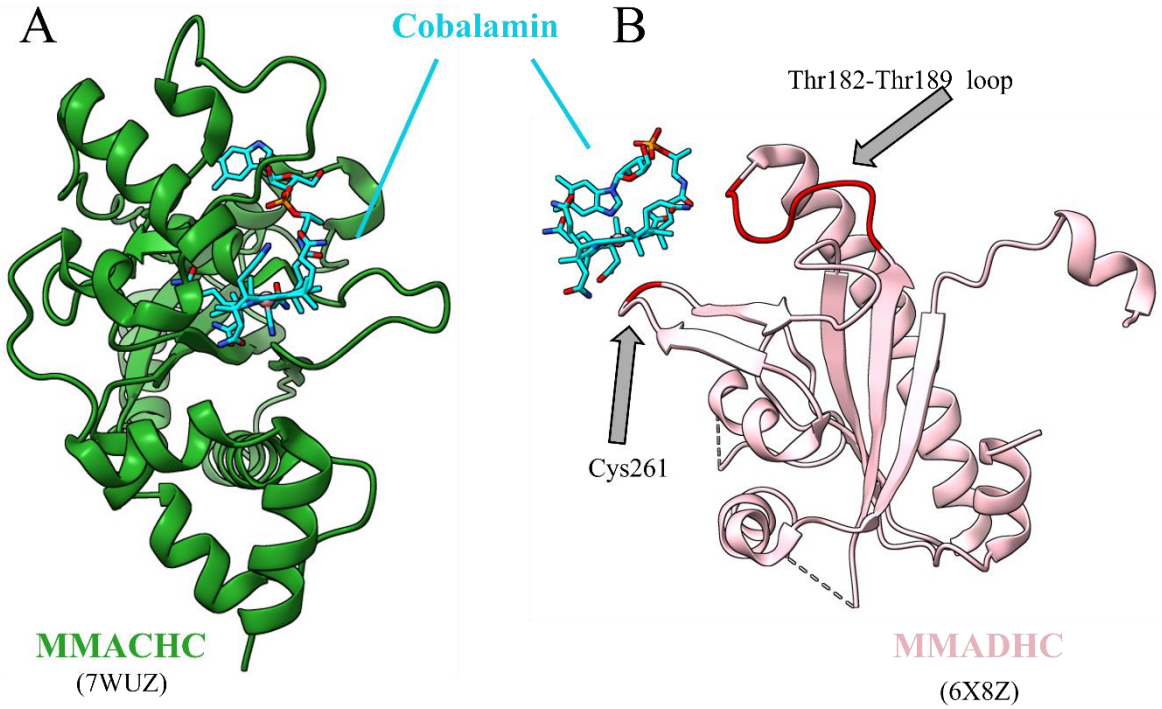

Figure 1.8: Crystal structure of MMACHC and MMADHC bound to Cbl. A. The crystal structure of human MMACHC (PDB: 7WUZ) bound to Cbl (cyan). **B**. The interaction of human MMADHC with Cbl as determined by X-ray crystallography (PDB: 6X8Z), showing the Cbl and the specific regions of MMADHC (loop Thr182-Thr189 and Cys261 in red) that interact with it. There are currently no structural data available for the MMACHC-MMADHC heterodimer.

Although Cbl-loaded MMADHC can carry and deliver Cbl to MTR without interacting with MMACHC *in vitro*, this is unlikely to occur *in vivo*, as most of the Cbl in the cytosol interacts with MMACHC (Mascarenhas, Guha et al. 2023). The MMACHC-MMADHC heterodimer needs to interact with MTR to deposit Cbl into the CblBD of MTR, which is expected to occur when MTR is in the activation conformation, the state observed in *apo* bacterial MTR (Mendoza, Purchal et al. 2023). This interaction likely forms at least a three-way complex composed of MMACHC, MMADHC, and MTR, and it was demonstrated by Duolink proximity ligation assays in fibroblasts (Bassila, Ghemrawi et al. 2017).

While evidence supports an interaction between the MMACHC-MMADHC heterodimer and MTR, the exact interface and duration of this complex remain unclear. Once

MTR receives Cbl in the cob[I]alamin form, it is methylated to MetCbl using SAM and can start the catalytic cycle (Mascarenhas, Guha et al. 2023). Further studies are needed to determine how transient this interaction is and which regions of those proteins interact for the Cbl delivery.

1.5.2 MTR and MTRR interaction integral to mechanisms of enzyme reactivation

As stated in Section 1.4.3, MTRR is the enzyme responsible for reactivating inactive MTR by transferring an electron necessary for the reductive methylation of cob[II]alamin during the MTR reactivation cycle (Olteanu and Banerjee 2001) (Figure 1.9). The FMN binding domain (FMNBD) of MTRR interacts with MTR through the AD to facilitate electron transfer (Wolthers, Basran et al. 2003). Some evidence suggests that MTRR's role extends beyond electron transfer. *In vitro* MTRR can stabilise apo MTR and enhance Cbl incorporation, implying that MTR and MTRR may interact during the Cbl transfer by the MMACHC-MMADHC heterodimer, potentially forming a four-way complex, as suggested by DuoLink results in fibroblasts (Bassila, Ghemrawi et al. 2017).

To date, only biophysical data on the interaction between the FMNBD of MTRR and the AD of MTR are available, and these studies have primarily used recombinant isolated domains. Using enzymatic activity assays, it has been suggested that residues Arg2 and Arg3 of MTRR are important as they may increase the affinity between MTR and MTRR (Zhang, Liu et al. 2020), and interaction analyses by fluorescence binding assays also indicate that Lys987 and Lys1071 in the AD of MTR are involved in the interaction with the FMNBD of MTRR (Wolthers and Scrutton 2007). Further structural and biochemical studies involving the full-length forms of MTR and MTRR could provide more detailed insights into the interface of these protein interactions.

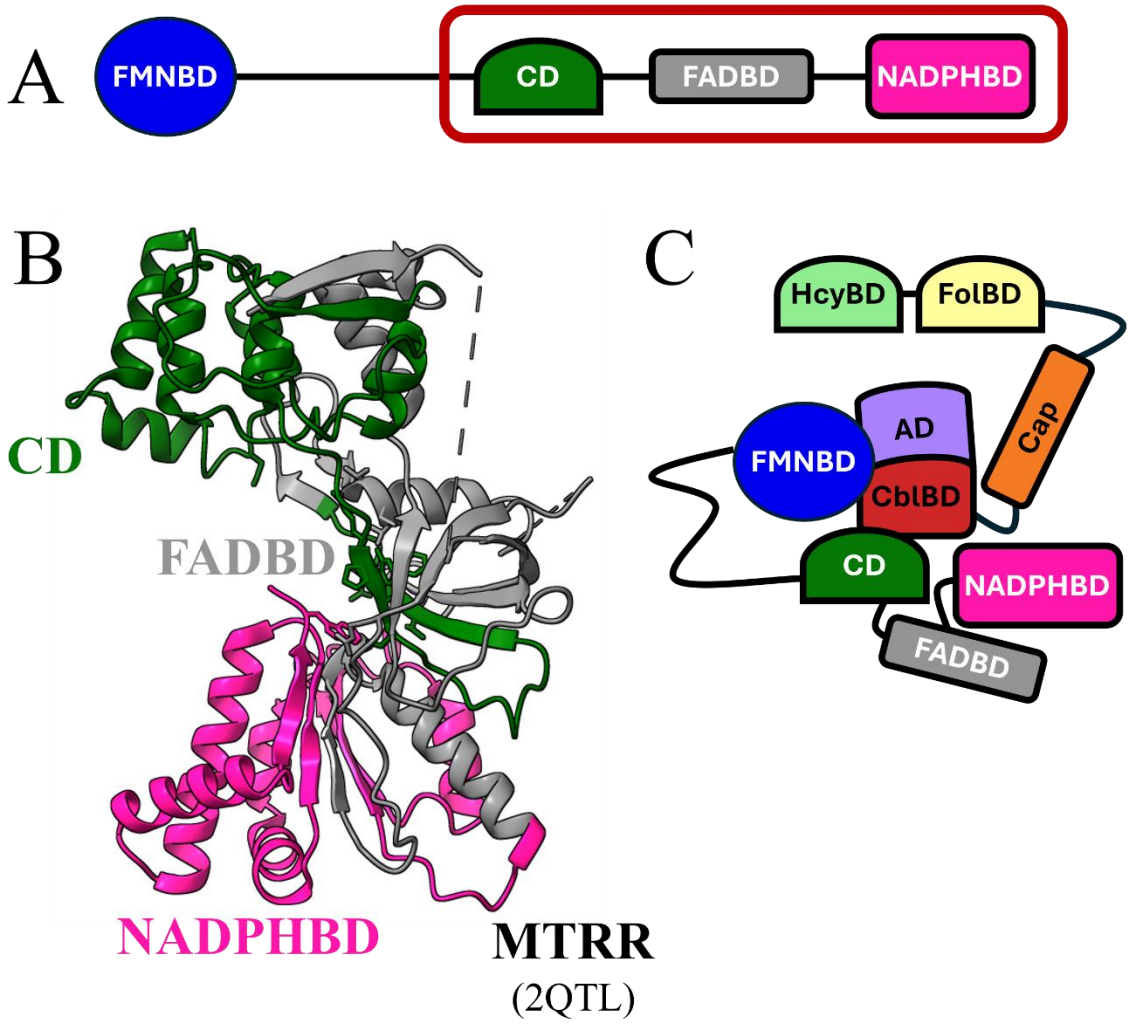

Figure 1.9: Scheme of MTRR Domains, Available Crystal Structure, and Proposed Domain Arrangements in the MTR-MTRR Complex. A. Schematic representation of MTRR domains, showing the FMN-binding domain (FMNBD – blue) at the N-terminus, followed by the connection domain (CD – green) after a long hinge, the FAD-binding domain (FADBD – grey), and the NADPH-binding domain (NADPHBD – pink). Red rectangle denotes the domains shown in B. **B**. Crystal structure of the CD, FADBD, and NADPHBD of MTRR (PDB: 2QTL). **C**. Proposed positions of MTRR domains during its interaction with MTR in the activation state, which is necessary for the reductive methylation of Cbl.

1.6 Architectural overview of Methionine Synthase

To date, 28 structures of MTR have been deposited in the Protein Data Bank. Two of these structures are from human MTR, while the remaining 26 are from bacterial orthologues of various species, including *Escherichia coli*, *Thermotoga maritima*, *Thermus thermophilus*, and *Thermus filiformis*. Due to the modular and flexible architecture of MTR

domains, a "divide-and-conquer" approach has been employed to solve the structures of MTR in smaller sections or isolated domains (Table 1.1).

Using truncated MTR constructs reduces the protein's flexibility, enhancing the probability of crystallisation, and enabling the acquisition of structural data via X-ray crystallography. The only structure not utilising X-ray is derived from a construct containing the HcyBD, FolBD, Cap, and CblBD of *T. filiformis* MTR, and was acquired by cryo-EM, a suitable technique for obtaining the structures of flexible proteins. This section explores the available MTR structures and highlights the structural insights gained from different regions of the protein.

Table 1.1: List of available structures of MTR, with their PDB ID, organism of origin, domains present in the structure, method of structure acquisition, ligands bound to the protein, and reference.

PDB ID	Organism	Domain(s)	Resolution (Å)	Method	Ligand	Reference
1BMT	*E. coli*	Cap, CblBD	3.00	X-ray	MetCbl	Drennan, 1994
1MSK	*E. coli*	AD	1.80	X-ray	SAM	Dixon, 1996
1K7Y	*E. coli*	Cap, CblBD, AD	3.00	X-ray	Cbl His-off	Bandarian, 2002
1K98			3.75			
1J6R	*T. maritima*	AD	2.30	X-ray	N/A	JSGC, 2002
1Q7M	*T. maritima*	HcyBD, FolBD	2.10	X-ray	N/A	Evans, 2004
1Q7Q			3.10			
1Q85			2.00			
1Q7Z			1.70			
1Q8A					Hcy	
1Q8J			1.90		Hcy, MeTHF	
2O2K	*H. sapiens*	AD	1.60	X-ray	N/A	Wolthers, 2007

3BOF	*T. maritima*	HcyBD, FolBD	1.70	X-ray	N/A	Koutmos, 2008
3BOL			1.85			
3BUL	*E. coli*	Cap, CblBD, AD	2.30	X-ray	Cbl His-off	Datta, 2008
3IVA	*E. coli*	Cap, CblBD, AD	2.70	X-ray	Cbl His-Off, AdoHcy	Koutmos, 2009
3IV9			3.25		Cbl His-On	
4CCZ	*H. sapiens*	HcyBD, FolBD	2.70	X-ray	THF	Vollmar, 2013
5VON	*T. thermophilus*	FolBD	2.10	X-ray	N/A	Yamada, 2018
5VOO			2.40		MeTHF	
5VOP			2.10			
6BM6	*E. coli*	AD	1.50	X-ray	AdoHcy	Fick, 2018
6BM5					SAM	
6BDY			1.51		Sinefungin	
8G3H	*T. filiformis*	HcyBD, FolBD, Cap, CblBD	3.60	Cryo-EM	Cbl	Watkins, 2023
8SSC	*T. thermophilus*	Full-length	2.75	X-ray	N/A	Mendoza, 2023
8SSD		Cap, CblBD, AD	2.4			
8SSE			3.5		Cbl His-off	

1.6.1 Insights into the N-half of Methionine Synthase: HcyBD and FolBD

The N-terminal half of MTR consists of the homocysteine binding domain (HcyBD) and the folate binding domain (FolBD). The structure of this region has been resolved for *T. maritima* MTR (Evans, Huddler et al. 2004, Koutmos, Pejchal et al. 2008) and human MTR (PDB 4CCZ, unpublished) (Figure 1.10). While the MTR proteins from these two organisms are structurally similar, the bacterial version is shorter, featuring a shorter FolBD compared to the human MTR.

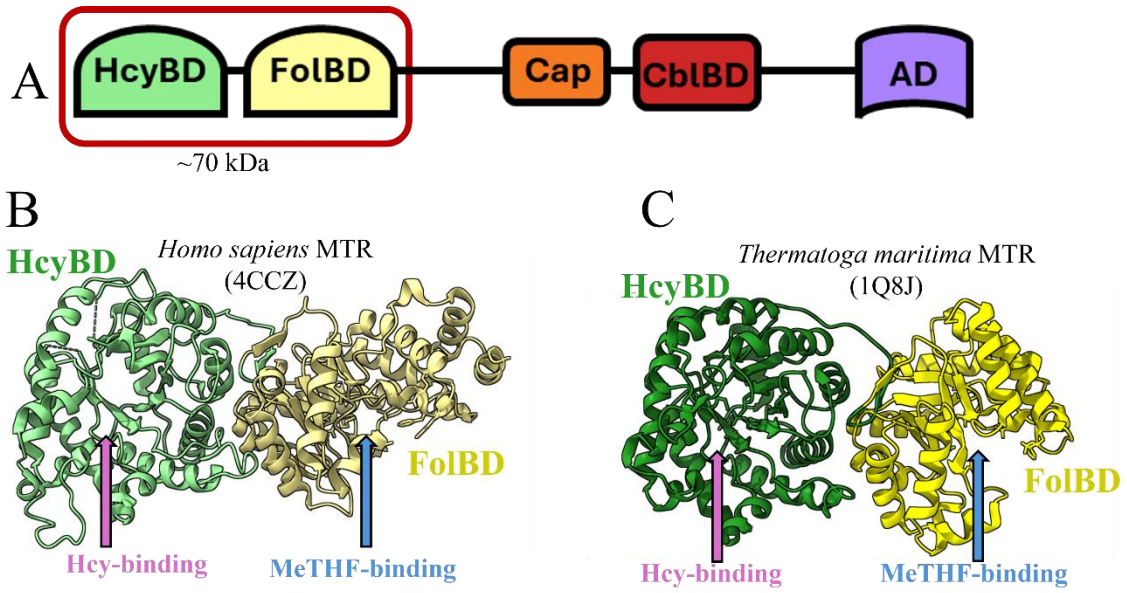

Figure 1.10: N-terminal Half of Human and Bacterial MTR. A. Domain Scheme of MTR, highlighting the HcyBD and FolBD which comprise the N-terminal half of MTR. Red rectangle denotes the domains shown in B and C. **B**. Crystal Structure of the human N-terminal half of MTR (PDB: 4CCZ), showing the HcyBD (green) and FolBD (yellow), along with the binding pockets for Hcy and MeTHF. **C**. Structure of the N-terminal Half of MTR from *T. maritima* (PDB: 1Q8J), displaying the same binding pockets and using the same colour scheme as in B.

The N-terminal halves of both organisms reveal that the HcyBD and FolBD are $(\alpha\beta)_8$ TIM-barrels, with limited flexibility between these two domains, indicating that they function as a rigid body (Evans, Huddler et al. 2004, Koutmos, Pejchal et al. 2008). The binding pockets for Hcy and MeTHF are located at the extremities of the N-terminal half, suggesting that the CblBD undergoes significant conformational changes to interact with these regions.

In the HcyBD, three conserved cysteine residues (Cys207, Cys272, and Cys273 in human MTR), along with a weak hydrogen bond from a conserved threonine (Thr195 in human MTR), coordinate the binding of Hcy with a Zn^{2+} ion, thereby activating the substrate (Banerjee, Frasca et al. 1990, Koutmos, Pejchal et al. 2008). Once activated, Hcy can accept the methyl group from MetCbl via an SN_2 mechanism, forming a thiol-methyl-cobalt intermediate (Kozlowski, Kuta et al. 2007, Alfonso-Prieto, Biarnés et al. 2010). The outcome of this reaction is the production of methionine (Met) and a 5-coordinate cob[I]alamin, as

the His785 residue from the CblBD coordinates with the lower axial position of the cobalt (Co) throughout the reaction (Banerjee, Frasca et al. 1990).

The interaction between the CblBD and FolBD for regenerating cob[I]alamin to become MetCbl occurs after the activation of MeTHF, as cob[I]alamin does not react with free MeTHF (Koutmos, Pejchal et al. 2008). The structures of the MeTHF-bound FolBD (PDB 1Q8J) reveal that MeTHF is secured in its binding pocket through a network of hydrogen bonds with conserved arginine residues (Evans, Huddler et al. 2004). These hydrogen bonds position the N5-methyl bond of the MeTHF pterin ring perpendicular to the binding barrel of the CobBD, oriented towards Cbl when the CblBD interacts with the FolBD. In methyltransferases, the activation of MeTHF typically occurs through the protonation of the N5 (Alonso, Cummins et al. 2009). In MTR, the identity of the proton donor for the cob[I]alamin-MeTHF interaction remains unknown (Jarrett, Choi et al. 1997). However, once MeTHF is activated, cob[I]alamin can proceed with the nucleophilic attack, leading to the demethylation of MeTHF, the regeneration of MetCbl, and the release of THF (Jarrett, Huang et al. 1998).

1.6.2 Insights into the C-half of Methionine Synthase: Cap, CblBD and AD domains

The C-terminal half of MTR consists of the Cap, CblBD, and AD domains. Structural studies of these three domains together for *E. coli* and *T. thermophilus* MTRs (Bandarian, Pattridge et al. 2002, Datta, Koutmos et al. 2008, Koutmos, Datta et al. 2009, Mendoza, Purchal et al. 2023) reveal the activation conformation, where the AD is positioned above the CblBD in a Cap-off state, leaving the Cbl exposed to the solvent (Figure 1.11). Currently, no structure is available for the complete C-terminal half of the human MTR; only the isolated AD has been crystallised (Wolthers, Toogood et al. 2007). The AD has also been crystallised for two bacterial MTR (Dixon, Huang et al. 1996, Fick, Clay et al. 2018).

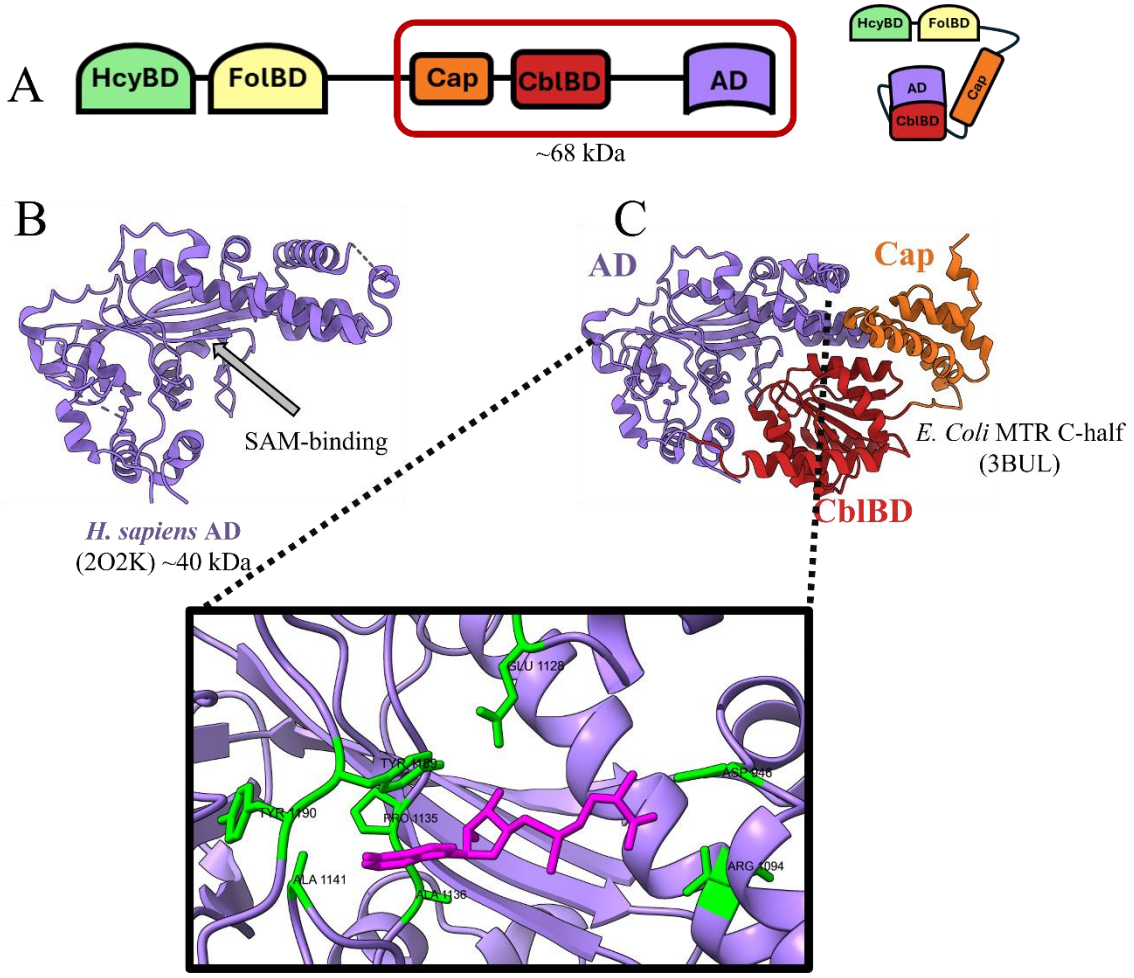

Figure 1.11: C-terminal Half of MTR. A. Domain Scheme of MTR, highlighting the Cap, CblBD, and AD, which comprise the C-terminal half of MTR. Red rectangle denotes the C-half of MTR. **B**. Crystal Structure of the human MTR AD (purple) (PDB: 2O2K), showing the binding pocket for SAM. **C**. Structure of the C-terminal Half of MTR from *E.coli* (PDB: 3BUL), displaying the Cap (orange), CblBD (red), and AD (purple) in the activation conformation, in which the cap is in the Cap-off state, and the AD is positioned above the CblBD. The inset indicates the residues (green) involved in the interaction with the SAM molecule (pink).

Different bacterial constructs have been used to obtain the C-half of MTR in the activation conformation, including one with the His759Gly substitution. The substituted histidine residue usually coordinates with the cobalt ion (Co) in the His-on state, and the His759Gly variant revealed the structure of the first His-off state for MTR (Bandarian, Pattridge et al. 2002). In the His-off state, Cbl is positioned closer to the SAM binding pocket, suggesting a movement of the cofactor that facilitates the reductive methylation process.

The isolated AD (Figure 1.11B) revealed a C-shaped, mixed α/β structure that holds the SAM molecule in its binding pocket through a hydrogen-bond network involving several conserved residues (Asp946, Arg1094, Glu1128, Pro1135, Ala1136, Ala1141, Tyr1189, and Tyr1190) (Dixon, Huang et al. 1996). The AD domains of human and *E. coli* MTR share a very similar conformation (sequence identity 49%, 63.5% similarity), including the SAM binding pocket.

Using disulfide-stabilised constructs, further insights were gained into the His-off state of Cbl in the activation conformation (Datta, Koutmos et al. 2008, Koutmos, Datta et al. 2009). These bacterial MTR structures demonstrated that His759 swings away from Cbl, forming 5-coordinate cob[II]alamin. Additionally, the structures revealed that Tyr1139 (Tyr1177 in human MTR) moves closer to Cbl, which is crucial for the reductive methylation of Cbl. The structure of the cob[III]alamin-bound C-terminal half of MTR revealed a His-on state, with Tyr1139 positioned further from Cbl, suggesting that the CblBD primarily drives the structural motions required for the reductive methylation of Cbl, as no significant movements were observed in the AD or Cbl (Koutmos, Datta et al. 2009). Collectively, these structural findings elucidated that upon SAM binding, Tyr1139 moves closer to the Co, lengthening its bond with a water molecule and resulting in a 4-coordinate Cbl, which can then be reduced and methylated Figure 1.12).

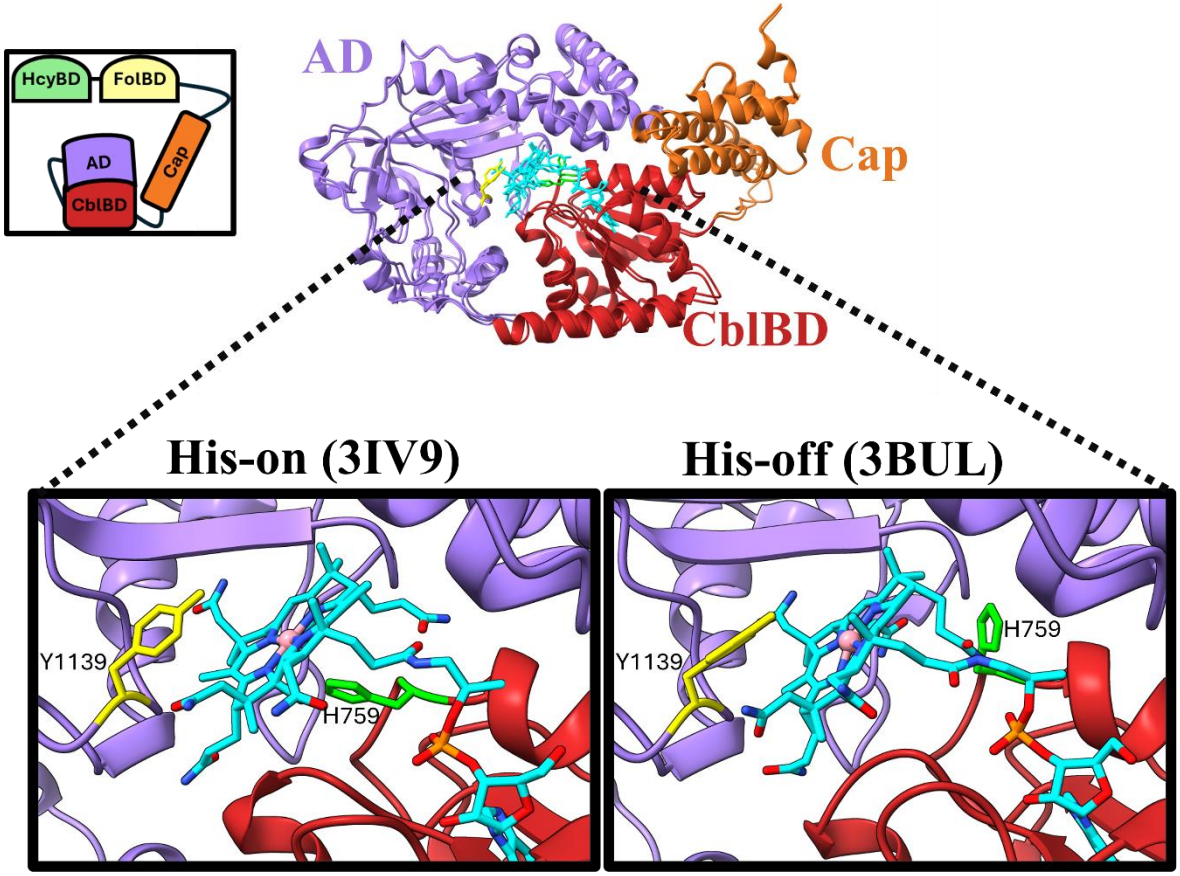

Figure 1.12: Position of His759 and Tyr1139 in *E. coli* MTR in the His-on and His-off states. The structure depicts the C-terminal half of MTR with the Cap (orange), CblBD (red), and AD (purple) in the activation conformation. The insets highlight the positions of Tyr1139 (yellow) and His759 (green) in the His-on state (left, PDB: 3IV9) and His-off state (right, PDB: 3BUL).

1.6.3 Cap and CblBD: Cap-on and Cap-off states

The first structure determined of MTR from any organism was the Cap-CblBD of *E. coli* with MetCbl in the Cap-on state (Drennan, Huang et al. 1994) (Figure 1.13). Although the initial intention was to crystallise the full-length protein, unintended in-cell proteolysis cleaved off the N-terminal half and the AD, resulting in the crystallisation of only the Cap-CblBD. In this structure, Cbl is in the Base-off His-on state, revealing for the first time the coordination of His759 with the lower axial position of Cbl and the dissociation of the DMBI tail from the corrin ring (Figure 1.12). The cap, formed by a four-helix bundle, is positioned on top of MetCbl, bound to the Rossmann-folded domain (CblBD). Conserved hydrophobic

residues (Phe708 and Leu715) are located at the end of the Cap, restricting solvent access to Cbl and protecting it from photolysis (Jarrett, Drennan et al. 1996). The cap-off state is characterised by the movement of the Cap away from the CblBD, and is found in the other conformation of MTR, in the catalytic cycle and in the activation conformation.

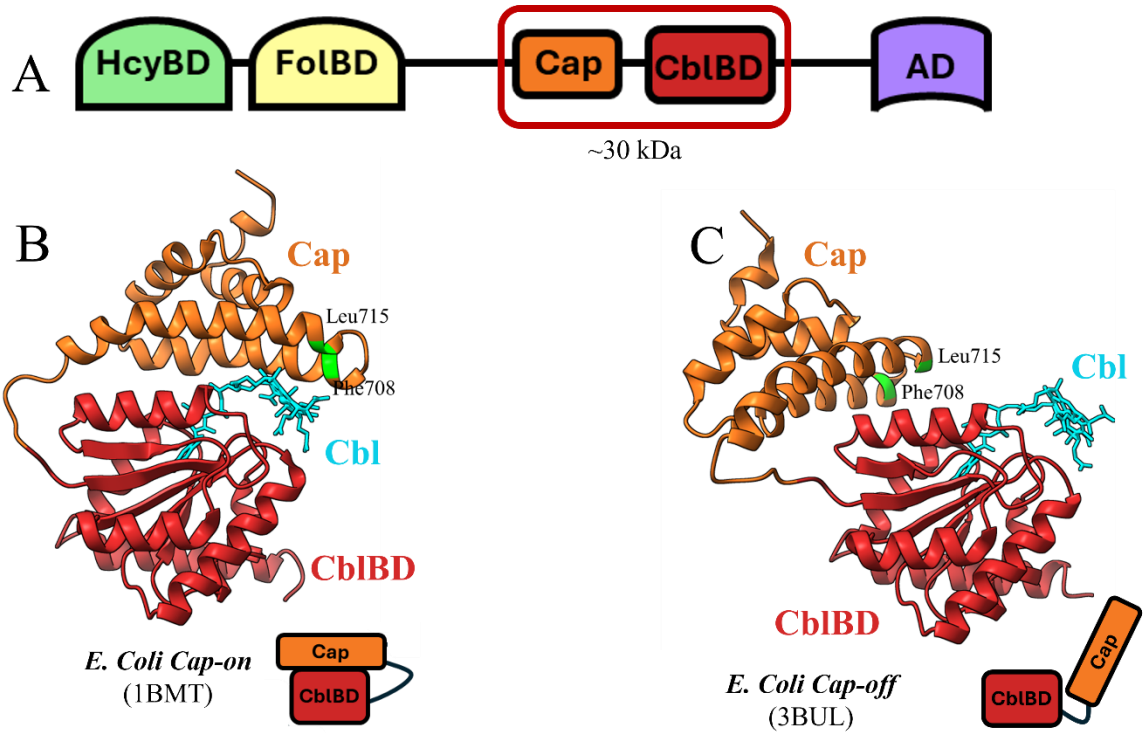

Figure 1.13: Cap-on and Cap-off conformations of MTR. A. Domain Scheme of MTR, highlighting the Cap and CblBD**.** Red rectangle denotes the domains shown in B and C. **B**. Crystal Structure of the E. coli Cap (orange) and CblBD (red) with Cbl (cyan) (PDB: 1BMT), showing the Cap-on conformation. **C**. Structure of the E. coli Cap and CblBD with Cbl (PDB: 3BUL), showing the Cap-off conformation. The conserved residues Phe708 and Leu715 are highlighted in green.

Another structure of the Cap-on conformation of bacterial MTR was obtained using Cryo-EM, making it the only available structure not acquired through X-ray crystallography so far (Watkins, Wang et al. 2023). This structure, derived from *T. filiformis* MTR, exhibits the Cap-on conformation in a manner similar to that of *E. coli* but with a novel interaction between the Cap with the FolBD and between the CblBD with the HcyBD (Figure 1.14). However, the residues involved in these interactions are not conserved across other organisms, suggesting that this interaction is organism-specific. This could explain why the

Cap-on structure of *E. coli* MTR (PDB: 1BMT) did not display any interactions with domains of the N-half.

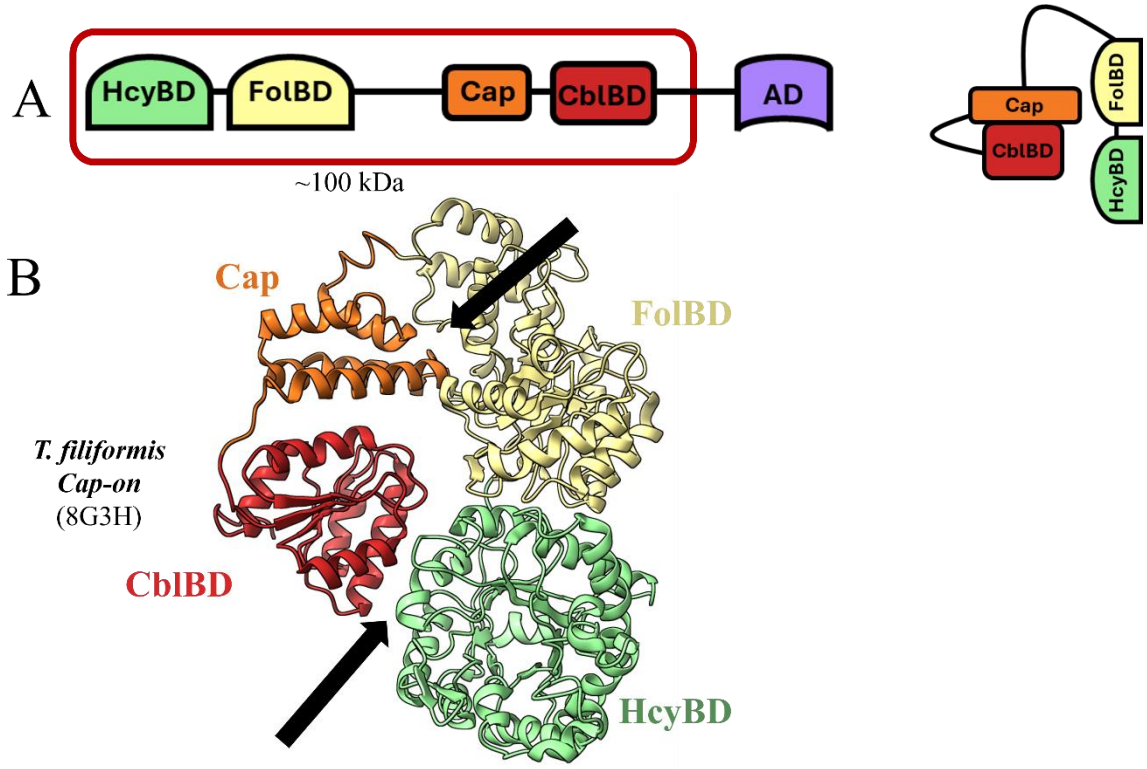

Figure 1.14: Cap-on state of *T. filiformis* MTR. A. Domain Scheme of MTR, highlighting the domains in the structure and a cartoon representing their arrangement in panel B. **B**. Cryo-EM structure of the *T. filiformis* HcyBD (green), FolBD (yellow), Cap (orange) and CblBD (red) (PDB: 8G3H), showing the Cap-on conformation. The black arrows indicate the interacting regions between the Cap and FolBD and CblBD with HcyBD.

Both structures of the Cap-on state (Figures 1.13 and 1.14) suggest that the protein is in a resting conformation, with the Cap protecting the Cbl cofactor before entering the catalytic or reactivation cycles. Using Small-angle X-ray scattering (SAXS) experiments, which reveals the overall conformation of the protein in solution, Watkins and colleagues (Watkins, Wang et al. 2023) demonstrated that *T. filiformis* MTR undergoes a conformational change from the resting conformation to the catalytic cycle or activation conformation upon MeTHF binding. This conformational change indicates that in *T. filiformis,* the trigger for MTR to leave the resting conformation is the binding of MeTHF. However, the specifics of

the conformational change from the resting state to the catalytic or reactivation cycles in other organisms remain unknown.

Additionally, these structures shown in Figures 1.13 and 1.14 highlight the significant conformational shift the Cap undergoes, transitioning from the Cap-on (resting conformation) to the Cap-off (activation conformation or during the catalytic cycle), where it rotates away from Cbl, exposing it for reductive methylation or the catalytic cycle.

1.6.4 The complete architecture of full-length MTR

Using the remarkably stable MTR from *T. thermophilus*, Mendoza and colleagues obtained the first structure of full-length MTR by x-ray crystallography (Mendoza, Purchal et al. 2023) (Figure 1.15). The structure reaffirmed the rigidity of the N-terminal half of MTR and highlighted the central role of the linkers between FolBD-Cap (Linker 1) and CblBD-AD (Linker 2) in modulating MTR conformations. It revealed that *apo* MTR is poised in the activation state, with the Cbl binding pocket exposed, requiring minimal rearrangements for Cbl binding. Additionally, soaking the *apo* MTR crystals with CNCbl reiterated the conformational changes in the conserved tyrosine (Tyr1132 in *T. thermophilus*, Tyr1177 in human) and histidine (His751 in *T. thermophilus*, His785 in human) during the reductive methylation of cob[II]alamin (Figure 1.12), which relies on the His-off state and the coordination of the tyrosine with the water molecule in the upper axial position of the Co ion.

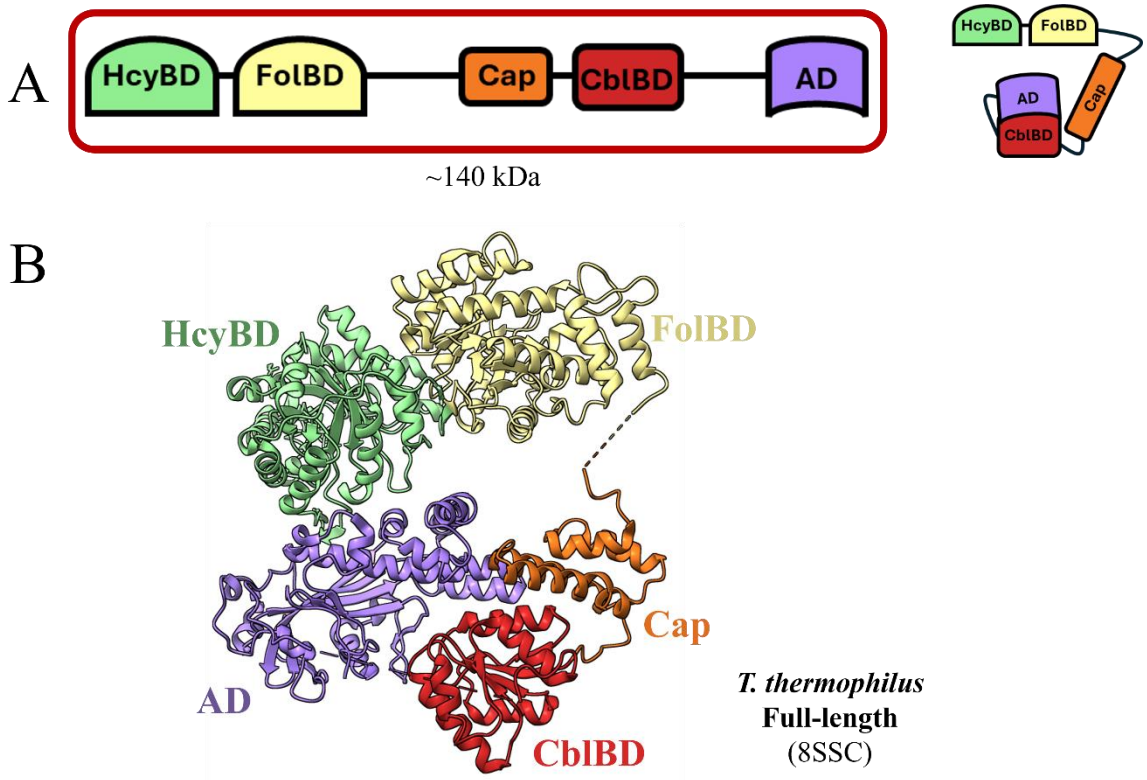

Figure 1.15: The full-length structure of MTR. A. Domain Scheme of MTR, highlighting all the domains and with a cartoon representing the activation conformation, in which the full-length *apo* bacterial MTR is positioned. **B**. Crystal Structure of the full-length bacterial MTR (PDB: 8SSC), displaying the HcyBD (green), FolBD (yellow), Cap (orange), CblBD (red), and the AD (purple). The AD is positioned on top of the CblBD, with the Cap away from the Cbl-binding pocket, indicating the activation state. Dotted line indicates a flexible loop that could not be modelled.

1.7 Cryo-electron microscopy (cryo-EM) for structural elucidation

Traditionally, X-ray crystallography has been the standard for determining the structure of different constructs of MTR. This technique has indeed provided valuable insights, particularly for the more rigid components of MTR, especially in its bacterial orthologue. However, certain limitations, such as the requirement for larger amounts of protein and the inherent flexibility of the protein, can make obtaining some protein structures challenging using crystallography. Alternatively, single-particle electron microscopy has also been employed to reconstruct macromolecular structures, and recent advances in this technology now allow for the determination of these structures at near-atomic resolution (1.2

Å) through cryo-electron microscopy (cryo-EM) (Nakane, Kotecha et al. 2020, Yip, Fischer et al. 2020, Glaeser, Nogales et al. 2021).

Cryo-EM involves using a transmission electron microscope to direct a beam of electrons into a sample that has been rapidly frozen to preserve its near-native state. As the beam passes through the sample, the scattered electrons generate images that can be reconstructed into a 3D model, revealing detailed information about the macromolecular structure (Faruqi and Henderson 2007, Bendory, Bartesaghi et al. 2020). The rapid advancement of cryo-EM technology enables its routine application nowadays and makes it an ideal approach for studying flexible proteins such as MTR. To date, a single structure of *T. filiformis* MTR has been described using cryo-EM, as shown in Figure 1.14, and reached a resolution of 3.6 Å. This section will look into the key steps of cryo-EM, from sample preparation to data acquisition, and finally, the data processing required to reconstruct the molecule's structure.

1.7.1 Key aspects of the sample preparation for Cryo-EM

Sample preparation for cryo-EM involves vitrifying the sample, which is loaded onto grids typically coated with a thin carbon or gold film. Unlike traditional electron microscopy, which uses stains and fixatives, cryo-EM employs liquid ethane to freeze the samples, preserving the protein in a near-native conformation. Proper grid preparation is crucial for generating a high-quality dataset that can yield a high-resolution structure (Bai, McMullan et al. 2015).

After loading the sample onto the grid, it is rapidly frozen in liquid ethane to vitrify it. This rapid freezing prevents the formation of ice crystals, preserving the sample in a glass-like state close to its native form. One of the main challenges during sample vitrification is controlling ice thickness, as it can significantly affect image quality because the ice can

reduce the contrast of the protein with the grid when too thick. Additionally, the air-water interface must be managed carefully, as it can lead to protein denaturation and aggregation. While managing ice thickness and the air-water interface can be challenging, the quality of the protein sample itself also plays a critical role in the final data quality. Therefore, ensuring the sample is intact and free from most contaminants is essential (Han, Avila-Sakar et al. 2023). For assessing the sample. negative staining electron microscopy can be a helpful preliminary step for quality control before using cryo-EM, which is more expensive and labour-intensive (Han, Avila-Sakar et al. 2023). Screening the sample with negative staining can help assess its homogeneity and determine the appropriate concentration of protein for cryo-EM (Henderson 2013).

The main differences between the two techniques lie in sample preparation, image contrast, and final resolution. In negative stain EM, a heavy metal stain is applied to increase contrast, producing a low-resolution representation of the protein structure. In contrast, cryo-EM does not require the use of a stain, resulting in lower sample contrast. However, this is compensated using more powerful microscopes, which enable the determination of high-resolution structures. Additionally, in cryo-EM, the sample is rapidly frozen in liquid ethane, preventing contact with air and preserving the native state of the sample, unlike in negative stain EM (Gonen 2021).

The grids used for EM consist of a mesh support with hundreds of micro-holes. Proper vitrification typically results in the ideal ice thickness within each hole, containing hundreds of protein particles. The grid material, hole size, and spacing can vary and are selected based on the sample type and results from prior sample screenings (Glaeser 2018). With a good-quality sample, the optimal grid type for the sample, and a reasonable ice thickness on the grid, it is ready to be loaded into the microscope for data collection.

1.7.2 Data collection parameters in Cryo-EM

Collecting small datasets from grids of the same sample can be a good strategy before a large-scale collection, guaranteeing the grid's quality. After choosing the appropriate grid for a larger collection, other parameters must be then considered before setting up the automated collection (Glaeser 2018).

The collection of datasets in cryo-EM is carried out at high acceleration voltages, typically between 100-300 keV, allowing the electron beam to penetrate the vitrified sample. The high voltage enhances the resolution of the micrographs by reducing diffraction effects (Cheng 2018). Modern cryo-EM microscopes are often equipped with autoloaders, which automatically load grids for imaging, facilitating the data collection of multiple grids and enabling overnight operation without human intervention (Glaeser 2016). These microscopes also use advanced detectors that improve the signal-to-noise ratio, capturing images as a series of frames (movies). This enables the correction of sample motion caused by the electron beam, ultimately enhancing image quality and increasing the final resolution (Li, Mooney et al. 2013).

The pixel size is the first parameter to consider in data collection, and it is directly related to the final resolution (i.e. 1.0 to 1.3 Å pixel size can lead to a final model at 2.5-3.0 Å resolution) (Feathers, Spoth et al. 2021). Smaller pixel sizes can lead to a higher resolution structure, but the number of particles per image decreases due to the higher magnification. Another parameter to consider is the total electron dose (electrons/Å^2), which is related to radiation damage to the sample, causing protein degradation and potentially leading to a lower resolution structure. Usually, the dose used is between 40-60 e^-/Å^2, balancing minimising contrast loss from the radiation with enhancing particle contrast (Langmore and Smith 1992, Peet, Henderson et al. 2019). Finally, the defocus range must be adjusted to optimise the trade-off between contrast and resolution. Lower defocus can lead to higher

resolution, but the contrast is reduced, which can be challenging for data analysis. Conversely, higher defocus can provide better contrast for particle detection, but the resolution is lower. Thus, the defocus range should be adjusted for each sample according to size, balancing contrast and resolution to generate micrographs that will produce high-resolution structures (Glaeser 2018).

1.7.3 Data analysis for 3D reconstruction

After data collection, the cryo-EM dataset containing hundreds to thousands of micrograph movies can be processed for the 3D reconstruction of the macromolecule using software packages, such as RELION (Scheres 2012) and cryoSPARC (Punjani, Rubinstein et al. 2017) (Figure 1.16).

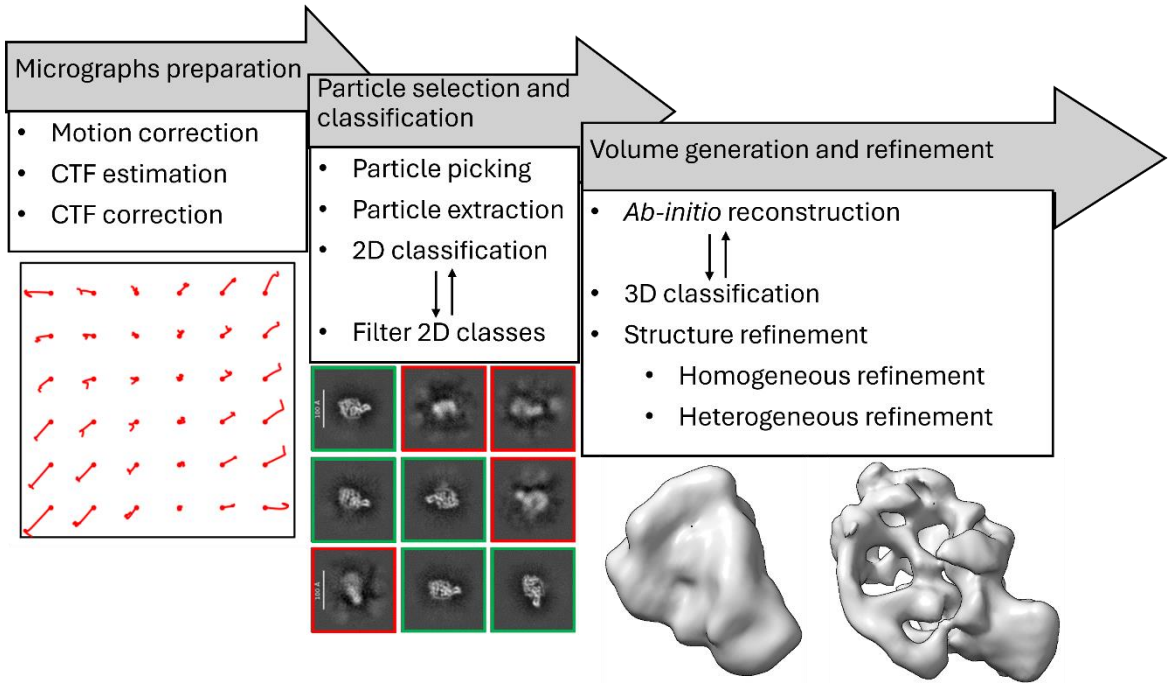

Figure 1.16: Overview of the usual processing pipeline for cryo-EM datasets. Following the collection of micrographs, they are first processed through motion correction and CTF estimation and correction. Once prepared, the particles can be selected in the micrographs using various tools. The particles are then extracted, classified, and filtered through several rounds of 2D classifications and classes selection. The best particles are used to generate *ab-initio* volumes, which can then be employed for 3D classification until a volume composed solely of appropriate particles is achieved. This volume is further refined to produce the final map of the target's structure.

The processing begins with the motion correction of the micrographs, which addresses the sample movement caused by the electron beam (Glaeser, Nogales et al. 2021). Movement during collection can reduce the quality of the images, leading to a lower resolution structure. The algorithms in data analysis software align the frames captured in a single image, restoring the original position of the particles, thereby enhancing the accuracy of the final models (Zheng, Palovcak et al. 2017). The next step is estimating and correcting the Contrast Transfer Function (CTF). CTF estimation calculates the experimental noise caused by the electron beam or the detector used during collection (Kirkland 1998, Zivanov, Nakane et al. 2020). After calculating the CTF, the micrographs are corrected according to the alterations induced during collection, restoring the particles' native image. Skipping CTF estimation and correction can lead to errors in the final 3D model, highlighting its importance at the beginning of data analysis.

After preparing the micrographs, the particles in each image can be picked for analysis. There are different methods for particle picking in cryoSPARC, including a blob picker, which selects any blobs found, including many blob-shaped ice particles that can be cleaned up downstream of the analysis (Zivanov, Nakane et al. 2020). More sophisticated algorithms, such as the template picker, are also available, using 2D templates from a reference structure or manually selected particles to screen the entire set of micrographs and select similar particles. Additionally, algorithms based on convolutional neural networks (CNNs) have been developed to pick particles from a small subset more accurately than the template picker, resulting in more particles selected and fewer false positives (Bepler, Morin et al. 2019, Wagner, Merino et al. 2019). In cryoSPARC, the TOPAZ picker (Bepler, Morin et al. 2019) is available for CNN-based particle picking.

Following particle picking, the particles are extracted from their respective micrographs in a box of user-defined size, usually twice as large as the expected particle

diameter (Glaeser, Nogales et al. 2021). These extracted particles can then be 2D-classified, aligning the particles and computing their correlations through several iterations (Wagner, Merino et al. 2019). Similar particles cluster together in the same class, and this tool is typically used as an initial clean-up step to remove false positives. The filtered 2D classes can be used to reconstruct a 3D volume of the protein (Reboul, Eager et al. 2018, Wagner, Merino et al. 2019). The *ab-initio* model generated from the 2D particles uses stochastic gradient descent to build the 3D model without a reference, solely utilising the 2D images provided (Punjani, Rubinstein et al. 2017, Zivanov, Nakane et al. 2018), thereby reducing model reconstruction bias. Generating two or more *ab-initio* models from a set of 2D images can be used as a filter for further particle 3D selection.

Once satisfied with the filtered particle set and the *ab-initio* structure it generates, the model can be refined through non-uniform refinement (cryoSPARC), which consists of CTF and defocus refinements and corrections for each particle composing the volume. CryoSPARC also refines regions of the volume that may undergo movement via local refinement, focusing on the most problematic areas to yield a better-resolution map (Zivanov, Nakane et al. 2020).

Finally, after the final map is generated and refined, the Fourier Shell Correlation (FSC) can be used to estimate the resolution of the reconstructed model. FSC is the correlation between two reconstructions (half-maps) built independently (Penczek 2020). The local resolution is also helpful for understanding the variability in resolution within the protein, revealing potentially more flexible areas.

1.8 AlphaFold: advances in protein structure prediction

While cryo-EM, like x-ray crystallography, provides valuable data and insights into experimental protein structures, computational approaches such as RoseTTAFold (Baek,

DiMaio et al. 2021) and AlphaFold2 (AF2) (Jumper, Evans et al. 2021) have been revolutionising protein structure prediction. These algorithms can often accurately model protein structures based solely on their amino acid sequences. The deep-learning algorithm of AF2 does not need a template for the structure prediction like the homology modelling due the analysis of the inherent structural propensity of the residues. Furthermore, the algorithm is trained from a large dataset of protein structures and sequences, which makes it understand complex patterns between residues in structures, also incorporating evolutionary information from multiple sequence alignments. Given the high success rates of these algorithms, it has become common practice to use experimental and modelling data in conjunction, for example including molecular replacement in cryo-EM map interpretation (Flower and Hurley 2021, Masrati, Landau et al. 2021).

An example of how AF2 can be utilised in combination with Cryo-EM data is demonstrated in the cryo-EM structure of *T. filiformis* MTR (Watkins, Wang et al. 2023). By exploring the AF2 database, which contains thousands of predicted MTR models from various organisms (ranging from prokaryotes to higher-order organisms), researchers discovered that the novel conformation of MTR—featuring interactions between the Cap:FolBD and CblBD:HcyBD (Figure 1.14)—had already been predicted by AF2 prior to its experimental observation. This highlights AF2's remarkable accuracy in predicting different conformations of multi-domain proteins.

1.8.1 AlphaFold Multimer: understanding multi-chain protein dynamics and interactions

Despite the power of AF2 in predicting monomeric proteins, there are limitations in multimer predictions, causing the generation of models without accurate inter-chain interactions. Hence, AlphaFold Multimer was developed in response to improving AF2 predictions for multi-chain protein complexes by adding neural network architectures and training it for modelling protein interactions. AF Multimer is specifically trained for

multimeric inputs, which enhances the accuracy of protein complex predictions compared to the standard AF2 without compromising its high intra-chain prediction accuracy (Evans, O'Neill et al. 2021). The accuracy of AF Multimer in predicting multi-domain proteins and protein complexes can generate multiple models representing different functional interactions and conformational states. These models can be compared with experimental data, providing a powerful tool for understanding protein dynamics (Evans, O'Neill et al. 2021). The advancements in this algorithm contribute to addressing longstanding questions in structural biology, potentially deepening the understanding of protein-protein interactions and their functions (Mirdita, Schütze et al. 2022).

This algorithm can be a valuable tool for MTR studies, as there are still gaps in understanding the protein-protein interactions involving MTR. The limited data on the complex formation of MTR with MMACHC, MMADHC, and MTRR (Bassila, Ghemrawi et al. 2017) can be enhanced by integrating experimental data and computational approaches.

1.8.2 AlphaMissense: the impact of mutations on protein structure

In addition to protein conformation and interactions, recent advances in AlphaFold enable the study of the structural consequences of gene mutations, in the form of missense changes (single amino acid substitution), on protein structure and protein-protein interactions (Cheng, Novati et al. 2023). AlphaMissense has emerged as a powerful tool for predicting the structural differences between protein variants and wild types, especially in the context of genetic disorders (Tordai, Torres et al. 2024).

The high accuracy of AF2 allows AlphaMissense to evaluate the structural modifications of proteins associated with disorders of intracellular cobalamin metabolism. Despite the intricate structure and conformational dynamics of MTR, its interactions with MMACHC, MMADHC, and MTRR involve key residues that remain unknown, some of

which have been reported in specific variants. The limited availability of structural data on the interactions between these proteins makes AlphaFold Missense a powerful tool for understanding the changes in protein variants.

The AlphaFold Missense algorithm can predict the pathogenicity of protein variants. Similar to AF2 and AF Multimer, it utilises a deep learning algorithm to assess the probability of a missense mutation affecting protein function. The algorithm takes into account evolutionary conservation and the biochemical properties of the protein within the context of its structure. By integrating these features, AlphaFold Missense can predict whether a protein variant is likely to be benign or pathogenic with high accuracy, providing insights into the role of specific residues in relation to overall protein activity (Frazer, Notin et al. 2021).

When combined with experimental data, AlphaMissense can significantly enhance the interpretation of specific mutations within cobalamin disorders, particularly those that may disrupt protein complex formation (Cheng, Novati et al. 2023).

1.9 Unsolved Questions in the Structure and Interactions of Methionine Synthase

1.9.1 How can the conformational of full-length human MTR be characterised to understand its dynamics?

As described in Section 1.4, the dynamic nature of MTR is crucial for the protein to undergo its catalytic and reactivation cycles, which involve interactions between the CblBD and other domains (Bandarian, Pattridge et al. 2002, Watkins, Wang et al. 2023). The inherent flexibility of MTR, coupled with the high reactivity of cob[I]alamin, presents significant challenges in obtaining structural data for the various conformations of MTR. Although it is hypothesised that MTR undergoes at least four distinct conformational changes depending on the exact interaction partner of the CblBD (CblBD:HcyBD,

CblBD:FolBD, CblBD:Cap, and CblBD:AD), only two of these conformations have been reported in bacterial MTR, and none observed in the human enzyme (Drennan, Huang et al. 1994, Bandarian, Pattridge et al. 2002, Datta, Koutmos et al. 2008, Koutmos, Datta et al. 2009, Mendoza, Purchal et al. 2023, Watkins, Wang et al. 2023). In addition to these conformational states, recombinant full-length human MTR production poses a substantial challenge, as its expression yields low amounts of protein, often insufficient for crystallography studies (Yamada, Gravel et al. 2006, Wolthers and Scrutton 2009, Zhang, Liu et al. 2020, Mascarenhas, Gouda et al. 2022). Consequently, the expression of isolated domains from bacterial MTR has been employed as the most effective strategy for acquiring the different conformations of the protein.

The approach taken in this thesis to study full-length human MTR involves expressing the protein in insect cell systems, followed by the use of cryo-electron microscopy (cryo-EM) to generate novel protein structures. Although our understanding of MTR's structure and conformational dynamics has expanded over the years, much of this knowledge is based on data obtained from bacterial MTR fragments. A significant gap remains in the structural characterisation of human MTR, particularly the full-length protein. **Taking advantage of the recent advancements in cryo-EM technology, which now make it more feasible to resolve the structures of flexible proteins like MTR, this thesis aims in addressing this gap in the understanding about this modular and dynamic protein.**

1.9.2 What are the structural basis of the protein-protein interactions involving MTR and other key proteins in the intracellular metabolism of cobalamin?

Another critical aspect of MTR biology is its interaction with other proteins involved in the intracellular metabolism of cobalamin. The complex network surrounding MTR includes proteins responsible for processing and delivering Cbl (MMACHC and MMADHC) as well as those involved in MTR's reactivation (MTRR) (Froese, Kopec et al.

2015, Mascarenhas, Guha et al. 2023). Although data exists on these interactions and the putative formation of a four-way complex (Bassila, Ghemrawi et al. 2017), the structural biology of these interactions remains largely unexplored. Specific gene mutations in the aforementioned essential proteins are known to cause various phenotypes associated with cobalamin metabolic disorders, possibly because they affect the protein interaction interfaces. These effects could be better understood through structural data.

This thesis focuses on MTR as the key component for studying these interactions, combining biophysical assays with a computational approach using AlphaFold Multimer to elucidate the basis of complex formation. Human MTR is involved in a unique network of protein-protein interactions, unlike bacterial MTR. These interactions are explored here using an advanced algorithm specifically designed to study protein interactions and enhance our understanding of the formation and dynamics of these complexes.

1.10 Thesis chapter layout

Chapter 2 outlines the methodologies employed to produce, purify, and conduct biophysical investigations on the full-length human MTR, focusing on its enzyme activity, thermal stability, and sample quality. It also details cryo-EM methods, from sample preparation to data analysis and protein model generation.

Chapter 3 provides detailed biophysical characterisation of MTR, covering its expression, purification, stability, ligand binding, and activity. This chapter reports that the recombinant full-length human MTR is expressed and purified in its *apo* form, capable of performing its catalytic cycle, homogeneous and monomeric, and suitable for structural studies.

Chapter 4 presents the grids preparation, data collection and processing. The generation of the EM map is also reported in this chapter.

Chapter 5 presents the model building of the structure of MTR determined by cryo-EM. The various samples provided insights into the conformation of the human protein in its *apo* state and upon MetCbl binding. High-resolution structures were obtained for the protein bound to MeTHF, OHCbl, and SAM, as well as in its MeTHF-bound form. These structures revealed conformations and domains of the human protein that have not been previously described.

Chapter 6 explores the protein-protein interactions involving MTR in intracellular cobalamin metabolism. Biophysical assays are used to study MTR's interaction with MMADHC and MTRR, while computational approaches predict the structural basis of these interactions. The results demonstrate that MMADHC and MTR can interact with MTR, forming a heterodimer and a three-way protein complex through different interfaces.

Chapter 7 summarises the outcomes of this thesis, discusses future directions for this research, and highlights its relevance to advancing the understanding of the complex protein MTR.

2. Methods

2.1 Recombinant protein production

2.1.1 Cloning

DNA fragments containing the full-length sequences of human MTR and MTRR were subcloned into pLIB vectors. Each sequence was amplified using primers containing each vector's overhang sequences.

Using the forward (MTR-f: CGGATCCATTAAGAAGGAGATATACTATGAAAACC CTGCGGGATGAG) and reverse (MTR-r: GATGGTGCGATTGGAAGTAGAGGTTCTC TGCGTCTGTATCATATCCCAAAATGG) primers for MTR, the sequence of the full-length protein was amplified using Q5® High-Fidelity DNA Polymerase (New England Biolabs). The sequence was cloned into a pLIB vector, adding a C-terminal TEV-cleavable His6-3xFLAG-tag, using NEBuilder® HiFi DNA Assembly (New England Biolabs). The reaction was possible because of the overhangs of each primer (shown in red), which are reverse-complementary to the pLIB vector used.

After cloning, the assembled plasmids were transformed into Mach1 *Escherichia coli* cells and plated on Luria Broth agar plate containing ampicillin at 50 µg/mL. Three colonies were selected, and the DNA was extracted and submitted for sequencing to confirm the correct integration of the sequence into the vector. The plasmid was named pLIB-MTR and stored at -20 °C.

For the full-length MTRR sequence, the forward (MTRR-f: CCGAGAACCTGTACTTC CAATCCATGGGCGCTGCGTCAGTG) and reverse (MTRR-r: CGACAAGCTTTATCCA CCTTTACTTCATGACCAAATATCCTGAAGGTAG) primers, also containing the overhangs for a pLIB vector. For MTRR, the vector added an N-terminal GST-His-6-tag.

The amplification and cloning process was performed as described above. After sequencing, the plasmid was named pLIB-MTRR and stored at -20 °C.

The pLIB vectors are commercially available and were designed with DNA-optimized linkers to allow more straightforward and more efficient cloning of multiple protein-coding sequences into a new construct called pBIG (Weissmann, Petzold et al. 2016). The primers in Table 2.1 were used to clone pLIB-MTR and pLIB-MTRR into a single plasmid for co-expression. The primer CasI_for and CasI_rev were used to amplify the sequence of MTR from the pLIB vector, and CasII_for and Casω_rev were used for MTRR. The assembly occurs because CasI_for and Casω_rev have overhang sequences to the co-expression plasmid pBIG1, and CasI_rev and CasII_for have complementary overhangs to themselves. This combination will create a single piece of DNA containing each sequence with its respective promoter and terminator.

Table 0: Names and sequences of primers used to assemble pLIB-MTR and pLIB-MTRR into a single co-expression plasmid.

Primer Name	Primer Sequence
CasI_for	AACGCTCTATGGTCTAAAGATTTAAATCGACCTACTCCGG AATATTAATAGATC
CasI_rev	AAACGTGCAATAGTATCCAGTTTATTTAAATGGTTATGATA GTTATTGCTCAGCG
CasII_for	AAACTGGATACTATTGCACGTTTAAATCGACCTACTCCGG AATATTAATAGATC
Casω_rev	AACCCCGATTGAGATATAGATTTATTTAAATGGTTATGATA GTTATTGCTCAGCG

The PCR reactions to amplify the sequences were performed using Q5® High-Fidelity DNA Polymerase (New England Biolabs). The amplicons generated from pLIB-MTR and pLIB-MTRR were incubated with SwaI-digested (New England Biolabs) pBIG1 plasmid on a 5:1 ratio. For the ligation of the complementary ends, NEBuilder® HiFi DNA Assembly (New England Biolabs) was used. After the recommended incubation, the plasmid was

transformed into Mach1 *E. coli* cells and plated on Luria Broth agar plate containing ampicillin at 50 µg/mL. A few colonies were selected for DNA extraction and submitted for sequencing. The correct plasmid was amplified, purified, named pBIG1a-MTR+MTRR and stored at -20 °C.

The assembly of MTR and MTRR into the pBIG1 plasmid kept the tags from each pLIB plasmid. For structural and interaction studies on the interaction between the proteins, the GST tag is not recommended due to its potential to form dimmers (Vargo, Nguyen et al. 2004). The GST tag from MTRR was removed from pBIG1a-MTR+MTRR with a single round of mutagenesis using a pair of primers (Mut-for: ATCCAATGCACCACCATCACCACCATCA, Mut-rev: GGT GGTGCATTGGATCCGCGCCCGA). These primers were designed on Benchling, using the pBIG1a-MTR+MTRR sequence as the template. They amplify the regions upstream and downstream of the GST tag and contain reverse-complementary sequences for the circularisation of the plasmid after the amplification.

The Mut-for and Mut-rev primers were used with Q5® High-Fidelity DNA Polymerase (New England Biolabs) to amplify the pBIG1a-MTR+MTRR, generating a single amplicon that was re-circularized using NEBuilder® HiFi DNA Assembly (New England Biolabs). The assembled plasmid was transformed into Mach1 *E. coli* cells and plated with ampicillin at 50 µg/mL. Six colonies were selected, the DNA was purified, and they were submitted for sequencing using the four primers shown in Table 2.1. The sequence of MTR and MTRR and the removal of the GST tag were confirmed, ensuring no single point mutations and that the N-terminal His-6-tag of MTRR was kept. The plasmid with the correct sequence was amplified, purified, named pBIG1a-MTR+MTRR(mut) and stored at -20°C.

2.1.2 Expression

The Sf9 (*Spodoptera frugiperda*) cells were maintained in Insect-XPRESS™ Protein-free Insect Cell Medium (Lonza) at 27 °C under rotation at 180 rpm. They were split every 3-4 days and diluted to $1x10^6$ cells/mL.

Each pLIB or pBIG plasmid was transposed into DH10Bac *E. coli* (ThermoFisher Scientific) by heat shock transformation to generate bacmids for expression on insect cells. After purification, the bacmids were transfected into a monolayer of Sf9 cells using Cellfectin II Reagent (Thermo Fisher Scientific), following the manufacture protocol. After 72 hours, the supernatant (baculovirus P1) was collected, centrifuged at 900 xg for 10 minutes and used to infect 50 mL of Sf9 for 72 hours.

After incubation, the cells were collected and centrifuged at 900 xg for 10 min. The supernatant containing the P2 baculovirus was used to infect 300 mL of Sf9 at 1x106 cells/mL. The P3 baculovirus was collected in the same way as mentioned above and used for large-scale expression.

For the large-scale expression, 600 mL of Sf9 at $1x10^6$ cells/mL were incubated for 72 hours with 10 mL of P3 baculovirus at 27 °C on rotation of 180 rpm. After incubation, cells were harvested at 900 xg for 20 minutes and resuspended into 15 mL of Lysis Buffer (50 mM HEPES, 500 mM NaCl, 5% glycerol, 20mM imidazole, 0.5 mM TCEP, and 0.1% Triton-X) per 600 mL cells. The pellet is frozen in liquid nitrogen and stored at -80 °C or immediately processed for protein purification.

2.1.3 Purification

The frozen pellets or freshly resuspended cells were sonicated for 20 minutes with 10 seconds on and 15 seconds off cycles. The lysate was clarified by centrifuging for 60 minutes at 75,000 xg and incubated under rotation at 4 °C with 100 µL of Ni-NTA resin pre-

equilibrated with Binding Buffer (50 mM HEPES, 500 mM NaCl, 5% glycerol, 20 mM imidazole, 0.5 mM TCEP) per 5 mL of lysate. It was then applied to a gravity-flow column. The resin was washed twice with 10 column volumes (CVs) of Binding Buffer, 5 CVs of Wash Buffer (50 mM HEPES, 500 mM NaCl, 5% glycerol, 40 mM imidazole, 0.5 mM TCEP), and eluted with 6x1 CV of Elution Buffer (50 mM HEPES, 500 mM NaCl, 5% glycerol, 250 mM imidazole, 0.5 mM TCEP). Samples of every purification step were analysed on SDS-PAGE. The samples containing the target protein(s) were pooled, concentrated and applied on a HiLoad 26/600 Superdex 200 pg chromatography column (Cytiva) with Gel Filtration (GF) Buffer (500 mM NaCl, 50 mM HEPES, 5% glycerol, 0.5 mM TCEP). The peaks from the gel filtration were analysed on SDS-PAGE, and the ones containing the target protein(s) were pooled, concentrated and frozen at -80 °C.

Unless otherwise stated, the same procedure was used for pLIB-MTR and pBIG1a-MTR+MTRR(mut).

2.2 Biophysical Assays

2.2.1 Thermal stability by *Differential Scanning Fluorimetry (DSF)*

MTR protein samples were diluted to 0.4 mg/mL in 25 mM HEPES and 250 mM NaCl for a final volume of 18 µL per well in a white 96-well PCR plate. Then, 2 µL of 100x diluted SyproOrange dye (ThermoFisher S6650) were added to each sample. The plate was loaded into QuantStudio Real-Time PCR Systems (ThermoFisher Scientific), and the DSF was performed from 25 °C – 95 °C on a heating rate of 0.5 °C per minute. When any ligand was added, it was diluted in the same buffer (25 mM HEPES, 250 mM NaCl) and added to each well before adding the dye. Data were analysed in GraphPad Prism software, and the thermal melting temperatures (Tm's) were determined using a nonlinear Boltzmann equation.

2.2.2 Methionine synthase activity assay

MTR activity assay was adapted from the previously described cuvette-based non-radioactive assay (Drummond, Jarrett et al. 1995). The current assay was miniaturised for a U-bottom 96-well plate. The following reagents were added to each well in the mentioned order: 24.4 µL of 100 mM Phosphate Buffered Sali (PBS) pH 7.5, 2 µL of Dithiothreitol (DTT) 500 mM, 2 µL of S-adenosylmethionine (SAM) 3.8 mM, 2 µL of L-homocysteine (Hcy) 100mM, 5 µL of MTR 4 µM, and 4 µL of OHCbl 500 µM. The plate was then incubated for 5 minutes at 37 °C, allowing DTT and OHCbl to create a semi-anaerobic environment by reducing O_2 into H_2O_2. After this time, 2.4 µL of 5-methyltetrahydrofolate (MeTHF) 6.6 mM was added to initiate the catalytic cycle of MTR, thus producing tetrahydrofolate. After 10 minutes at 37 °C, the reaction was quenched by adding 10 µL of 5 N HCl in aqueous 60% formic acid and immediately heated at 80 °C for 10 minutes to convert the produced tetrahydrofolate to methenyltetrahydrofolate, which concentration can be determined using $\Delta\varepsilon 350 = 26{,}500\ M^{-1}\ cm^{-1}$. The data was analysed, and graphs were generated using RStudio.

2.2.3 Native Mass Spectrometry

Dr Rod Chalk performed native mass spectrometry at the Centre for Medicines Discovery (Oxford) to detect potential ligands on recombinant MTR. By dialysis, purified MTR was exchanged from the GF Buffer to 50 mM ammonium acetate, pH 6.5. The sample was loaded into Quantitative Time-of-Flight (Q-TOF) 6530 mass spectrometer attached to a standard electrospray ionisation source (Agilent Technologies), with 77 a constant flow rate of 6 µL/min, using a 1.0-ml gas-tight positive displacement syringe (Hamilton) and PEEK capillary tubing with an inner diameter of 0.005 inches.

The mass spectrometer was operated in positive ion mode, using the 1-GHz detector mode with a scan range of 100–20,000 m/z and a fragmentor voltage of 430 V. Mass spectrum was analysed using the MassHunter Qualitative Analysis software (Agilent Technologies).

2.2.4 Mass photometry

Mass photometry analysis of recombinant MTR was performed on a MassFluidix HC (Refeyn). The protein was diluted to 10 nM in 50 mM HEPES, and 10 µL of the sample was loaded into the MassFluidix HC chip, which can provide molecular mass distribution. The assay works by a rapid sample dilution through the chip, which measures the mass of the molecules based on light scattering. (Claasen, Kofinova et al. 2024). The built-in software DiscoverMP analysed the data and applied Gaussian fits to the histogram, adding the measured contrast, standard deviation (σ), and population counts.

2.3 Characterization of Cobalamin Binding and Transfer

2.3.1 Loading OH_2Cbl into MMADHC

Previous expressed and purified MMADHC construct lacking the N-terminal region (Q124-N296) was used to test the load of OH_2Cbl into the protein based on earlier work (Li, Mascarenhas et al. 2020). For the load, 100 µM of OH_2Cbl was mixed with 120 µM of MMADHC in KCl Buffer (100 mM HEPES, 150 mM KCl, 10% glycerol) for a final volume of 40 µL. The reaction was incubated at room temperature for 1 hour, and the UV-Vis spectrum was recorded every 3 minutes to check the reaction. The change in absorbance at 350 nm ($\Delta\varepsilon_{350\,nm}$ = 11.8 mM^{-1} cm^{-1}) was fitted to the single exponential equation $\Delta A_{350} = y_0 + a_1 \times e^{-k_{obs1}xt}$.

2.3.2 OH_2Cbl transfer from MMADHC to MTR

To detect the cobalamin transfer from MMADHC to MTR, previously OH_2Cbl-loaded MMADHC was concentrated by multiple rounds of centrifugation (4000 xg, 4 °C) in Amicon® concentrators with a 10 kDa molecular weight cut-off (Sigma Aldrich) followed by adding more KCl buffer to remove free OH_2Cbl. Then, 15 µM MMADHC-OH_2Cbl was incubated with 8.8 µM of MTR in a final volume of 100 µL of KCl buffer for 1 hour at room temperature. After this time, the sample was loaded into a Superose® 6 10/300 GL gel filtration column (Cytiva) equilibrated with KCl buffer at 4 °C at a flow rate of 0.2 mL/min. The elution profile was monitored at 280 and 350 nm to detect the proteins and cobalamin.

2.4 Negative stain electron microscopy

The negative stain electron microscopy was performed at the Electron Microscopy Research Services at Newcastle University. Copper grids with carbon coating (300 mesh, Quantifoil Micro Tools GmbH) were glow discharged for 1 minute. 10 µL of purified MTR at 40 nM was placed on the grids for 30 seconds. The excess sample was dried off using filter paper (Whatman), and the grid was placed on a drop of filtered 1.5 % uranyl acetate (Sigma-Aldrich) for 30 seconds. Grids were dried again using filter paper before visualisation. Grids were loaded and visualised under an HT7800 transmission electron microscope (Hitachi), and micrographs were acquired.

2.5 AlphaFold2 (AF2) and AlphaFold2-Multimer for protein structure prediction

AlphaFold2 (AF2) utilised a deep-learning algorithm to predict protein structures, combining information from homologous structures, amino acid sequences, and multiple sequence alignments (MSA) (Jumper, Evans et al. 2021). It currently has high accuracy and correctly predicts most single protein chains. AF2-Multimer was developed by training AF2

for multimeric protein predictions, increasing the correct predictions of protein complexes (Evans, O'Neill et al. 2021).

For either AF2 or AF2-multimer, the sequence of the desired protein(s)/domain(s) from the UniProt database was used to input AF2 using a computer cluster with potent GPU resources. The sequence generated MSA, identifying homologous sequences and capturing interactions between amino acid residue pairs. After refinements, it was converted into 3D coordinates, and a final structure was generated with confidence indicators. The generated models and all other structures were analysed on ChimeraX (Pettersen, Goddard et al. 2021).

2.6 Pulldown Assays

To assess the interaction between two or more proteins, the 3XFLAG tag on the N-terminal of MTR was used to immobilise the protein and act as bait for either MTRR or MMADHC. Briefly, 100 µL of ANTI-FLAG® M2 Affinity Gel (Merck) was equilibrated with KCl buffer, and 10 µg of either MTR or MTR-MTRR co-expressed complex without the GST tag (preys) was diluted for a final volume of 50 µL of the same buffer and added to the affinity gel and incubated at 4 °C for 1 hour. When another protein was added, 10 µg of it was diluted in KCl buffer to 50 µL, added to the gel and incubated for an extra hour at 4 °C.

The gel was added to a purification mini spin column (Neo Biotech) and centrifuged for 2 min at 900 xg to collect the flow-through. The affinity gel was washed twice with 500 µL of KCl buffer (10 CVs final) by centrifuging at the same speed. To elute the remaining proteins, 40 µL of reducing Blue Protein Loading Dye (New England Biolabs) was added to the gel, incubated for 1 minute and collected by centrifuging at 900 xg for 5 minutes. 10 µg of the prey proteins not containing the 3XFLAG tag were also incubated with the affinity gel for control. Samples were analysed using SDS-PAGE and Western Blot.

2.7 Cryo-EM

2.7.1 Cryo-EM sample preparation

The grids used were either Quantifoil™ Au 300mesh 1.2/1.3 (Quantifoil) or UltrAuFoil Holey Gold Au 300 1.2/1.3 (Quantifoil), previously glow discharged on a PELCO easiGlow™ (Ted Pella, Inc) for 60 seconds, under a 0.26 mBar pressure and 20 mA current. MTR samples were diluted to 0.5 mg/mL on Buffer A (25 mM HEPES, 250 mM NaCl), and 3 µL of the sample was vitrified by plunge-freezing in liquid nitrogen-cooled liquid ethane (Vitrobot Mark IV, Thermo Fisher Scientific). The Vitrobot was set to 4 °C and 100% humidity, and the blotting parameters were 3 sec for blot time, 10 sec for wait time, and 0 sec for drain time and blot force. When the protein was tested with any ligands, it was previously incubated with 50 µM of each ligand or 1 mM MeTHF diluted in Buffer A for 30 minutes at room temperature.

2.7.2 Cryo-EM data collection

Part of the data was collected using a 200 keV Glacios electron cryo-microscope (Thermo Scientific) with a Falcon 4 detector (pixel size 0.574 Å) and a total exposure dose of 50 $e^-/Å^2$ at the York Structural Biology Laboratory, University of York. Another part was collected on a 300 keV Krios cryogenic electron microscope (Thermo Scientific) with a Gatan K3 detector (pixel size 0.254 Å) equipped with the TFS Falcon 4i detector & Selectris X energy filters and a total exposure dose of 65.2 $e^-/Å^2$ at the electron Bio-Imaging Centre, Diamond Light Source.

2.7.3 Cryo-EM data analysis

The analysis of all datasets was conducted using cryoSPARC (Punjani, Rubinstein et al. 2017), one of the most common software for cryo-EM data processing. The micrographs

collected on Krios or Glacios (hundreds to thousands) were processed to generate a 3D protein reconstruction.

The first step for all data processing was motion correction, which processed the movies collected and corrected the particles' movement caused by the electron beam. Next was the estimation and correction of the contrast transfer function (CTF), which measured and described the optical abnormalities caused by the microscope optics, such as defocus and astigmatism, and corrected them, improving the resolution and image contrast, leading to 3D structures with higher resolution.

With the motion and CTF-corrected micrographs, the next step in the data analysis was to isolate particle images for single-particle analysis (SPA). First, it is necessary to find a balance between over-selecting particles and under-selecting them. The former can complicate the downstream processing, making it challenging to filter “good” particles from “bad particles,” and the latter can exclude important particles that could cover different protein orientations. The blob picker tool was used for this first particle picking, which picked theoretical blob particles from a user-defined diameter (80-150 Å was used for MTR data). After selection, the particles were extracted in a box size of 400 pixels for further processing.

To clean the dataset and get rid of particle stacks that were wrongly selected by the blob picker, rounds of 2D classifications were used. This classification aligns particles in 2 dimensions, correlating all images with randomly selected particles—the more particles in the class, the less noise and more detail in the class image. This step is ideal for getting a read of the “bad” particles, as it is possible to select the classes with the most protein images, which can even display some secondary structure depending on the particle numbers in the

class. Several rounds can be done, and different numbers of 2D classes can be created in each round depending on the number of particles remaining.

The selected particles from the 2D classifications can generate *ab initio* models without reference volumes, so no bias is added to the generated model. In the case of MTR, the N-terminal and C-terminal halves were shown separately, so a minimum of 3 *ab initio* models were generated: one for each half and the others for particles that did not fit in the reconstruction models of either half of MTR. Following the generation of *ab initio* volumes, a heterogeneous refinement classified the particles again and refined the volumes by doing a 3D classification of the particles.

The 2D classes used for the N- and C-terminal halves hetero refined volumes were then used for template picking, which, in the case of the datasets analysed, increased the number of particles corresponding to MTR. These particles were extracted, classified, and filtered as the ones from the blob picker. The particles from both rounds of picking were grouped, and the duplicates were excluded since they had the same box size. The grouped particles were used again for *ab-initio* models generation, which led to more detailed volumes because of the increase in particle number.

Different tools, such as Non-Uniform Refinement, can refine the generated volumes. This improves the reconstruction quality by focusing on the more disordered regions, enhancing the map's overall resolution. Global CTF refinement and local refinement can improve the map's resolution by correcting the CTF parameters of all particles and refining the alignment of specific map regions.

Most of the data processing used the abovementioned tools. If a different one was used for the data analysis, it will be specified and described in the appropriate section.

2.8 CRISPR/Cas9-Mediated Genome Editing of the MTR Gene in HEK293 Cells

2.8.1 CRISPR sgRNAs design and cloning

The plasmid eSpCas9(1.1) Dual_sgRNA (Addgene #86613) was used for the CRISPR assay. This plasmid expresses the Cas9 enzymes with the desired sgRNA and also targets the gene ATP1A1, which grants ouabain resistance to the cells when used with the appropriate donor plasmid (Agudelo, Duringer et al. 2017).

The sgRNAs were designed based on the MTR gene of HEK293 cells. First, PAM regions (NGG) were identified and located in potential areas to add a 3XFLAG tag to the endogenous MTR. Two complementary 20-nucleotide oligos (top and bottom) for each region were designed using Benchling as the sgRNAs sequences. These oligos were assembled and phosphorylated in a reaction containing 1 µL of each oligo (top and bottom) at 100 µM, 1 µL 10x T4 DNA Ligase Buffer (New England Biolabs), 1 µL T4 PNK (New England Biolabs), and 6 µL of double-distilled water. It was incubated at 37 °C for 30 minutes, then 95 °C for 10 minutes, and -0.1 °C/s decreased the temperature to 25 °C (Ran, Hsu et al. 2013).

120ng of Dual_sgRNA previously digested with the BbsI High-Fidelity enzyme (New England Biolabs) was incubated for 1 hour at room temperature with 2 µL of each pair of sgRNA diluted 1/50, 2 µL T4 DNA Ligation Buffer (New England Biolabs), 1 µL T4 DNA Ligase (New England Biolabs), and double-distilled water for a final volume of 20 µL. The reaction was then chemically transformed into Mach1 cells and plated overnight on Luria Broth agar plate containing ampicillin at 50 µg/mL.

The next day, the colonies were submitted to colony PCR to confirm the presence of the specific sgRNA. Briefly, each colony was mixed into 12.5 µL of Phusion® High-Fidelity PCR Master Mix with HF Buffer (New England Biolabs), 1.25 µL of hU6 forward primer

(GAGGGCCTATTTCCCATGATT), 1.25 µL of the bottom oligo for each sgRNA, and 10 µL of double-distilled water. The reaction followed the product guidelines using 55 °C for annealing temperature and 30 seconds for extension time. The presence of the amplified DNA was confirmed on a 1.5% agarose gel, and positive colonies were submitted to sequencing using the specific bottom oligo for each sgRNA.

2.8.2 sgRNA surveyor

To test the efficacy of each sgRNA cloned, $1x10^6$ HEK293T cells were plated on each well of a 24-well plate at 37 °C with 8% CO_2. After 24 hours, transfection was performed per sgRNA using Lipofectamine™ 3000 (Thermo Fisher Scientific), following the manufacturer's recommendation. After 72 hours, the medium was removed, and the gDNA of each well was extracted using 200 µL of DNAzol™ (Thermo Fisher Scientific).

Previously designed primers amplifying 500 nucleotides upstream and downstream of the PAM regions were used to amplify the potentially cleaved areas of the MTR gene and the same region on non-transfected cells. The amplicons were purified using the QIAquick PCR Purification Kit (Qiagen) for denaturation and annealing reactions. Briefly, 250 ng of each amplicon was mixed with 2 µL of 10x NEB Buffer 2 (New England Biolabs) and double-distilled water for a final volume of 19 µL. This reaction was incubated at 95 °C for 5 minutes, and the temperature was reduced to 85 °C at a rate of -2 °C per cycle in 5 cycles of 5 seconds each. It was then cooled to 25 °C at a rate of -1 °C per cycle in 60 cycles of 10 seconds each. After that, samples were treated with 1 µL of T7 endonuclease (New England Biolabs) for 30 minutes at 37 °C, followed by adding 5 µL of loading buffer to quench the reaction. All samples were loaded into a 1.5% agarose gel to compare the band integrity of the treated cells and their respective controls.

3. Biophysical Characterisation of MTR

3.1 Background: Methionine Synthase (MTR)

Recombinant expression and purification are well-established methods for studying proteins that exhibit low endogenous expression levels. These approaches enable the production of sufficient quantities of protein for various biophysical and structural characterisations.

Human methionine synthase (MTR) is a 140 kDa protein comprising five modular domains. Previous reports have detailed the purification of full-length recombinant human MTR using various affinity purification techniques following expression in a baculovirus-infected insect cell system. The most commonly employed method is cobalamin-affinity chromatography (Yamada, Gravel et al. 2006, Wolthers and Scrutton 2009, Zhang, Liu et al. 2020), which utilises an agarose resin bound to cobalamin (Cbl), for which MTR has affinity. Elution of Cbl-bound MTR is achieved by exposure to a 500-W tungsten lamp, facilitating the release of holo MTR. Although this method enables single-step purification due to the specificity of MTR for Cbl, it results in the loading of all MTR molecules with Cbl, thereby replacing *apo* MTR with Cbl-loaded MTR. This can potentially limit the range of ligand combinations available for biophysical assays. Another reported method involves the purification of MTR expressed with an N-terminal His-MBP (maltose-binding protein) fusion. This approach utilises amylose resin, followed by Ni-NTA chromatography after tag cleavage, and finally, size-exclusion chromatography (Mascarenhas, Gouda et al. 2022).

Recombinant MTR has been extensively studied through enzymatic activity assays. These assays, previously described for MTR from various species such as *E. coli*, *Thermus thermophilus*, and humans, (Goulding, Postigo et al. 1997, Mascarenhas, Gouda et al. 2022, Mendoza, Purchal et al. 2023) measure the methylation of homocysteine (Hcy) to

methionine (Met), derived from the demethylation of 5-methyltetrahydrofolate (MeTHF) to tetrahydrofolate (THF) during the catalytic cycle of MTR (Figure 1.6). The assay can detect the direct transfer of the methyl group from MeTHF to Hcy, forming Met (radioactive assay), or it can indirectly measure THF production (non-radioactive assay). These assays typically begin by mixing buffer, dithiothreitol (DTT), S-adenosylmethionine (SAM, required for reactivation of the inactive enzyme), Hcy, MTR enzyme, and hydroxocobalamin (OHCbl, a stable form of Cbl in the +3 oxidation state). The reaction is incubated for 5 minutes at 37 °C, after which the substrate MeTHF is added and incubated for a further 10 minutes at 37 °C. The protocol then diverges depending on the type of assay, which may be conducted either aerobically or anaerobically (Jarrett, Goulding et al. 1997) and may involve radioactive or non-radioactive ligands (Mascarenhas, Gouda et al. 2022).

The radioactive assay utilises [^{14}C]CH_3-H_4 folate, a MeTHF molecule containing a radioactive [^{14}C] methyl group. After 10 minutes of incubation, the reaction is quenched at 90 °C and loaded onto an AG 1-X8 (chloride form) column, a strong anion exchanger that binds both MeTHF and THF. The column is then washed, and the transfer of [^{14}C] from MeTHF to Met (present in the flow-through and wash fractions) is quantified by scintillation counting (Frasca, Banerjee et al. 1988, Chen, Chakraborty et al. 1995, Olteanu and Banerjee 2003, Wolthers and Scrutton 2007, Wolthers and Scrutton 2009, Mascarenhas, Gouda et al. 2022). The non-radioactive assay involves the addition of 5 N HCl in 60% aqueous formic acid, followed by heating, which converts THF to methenyltetrahydrofolate, a molecule with an absorbance at 350 nm (Drummond, Jarrett et al. 1995, Mascarenhas, Gouda et al. 2022).

Both radioactive and non-radioactive assays have been optimised for cuvette readings with a final reaction volume of 1 mL, typically under aerobic conditions. An anaerobic chamber is often employed to minimise the oxidation of cob[I]alamin to cob[II]alamin, which inactivates MTR. To study the recovery cycle of MTR, specifically the

activity of its partner protein methionine synthase reductase (MTRR), the use of an anaerobic chamber is usually employed, as the rate of MTR inactivation is slower under anaerobic conditions (Chen, Chakraborty et al. 1995, Olteanu and Banerjee 2003, Wolthers and Scrutton 2009). For MTR activity assays, however, aerobic conditions allow detection of active MTR even in the presence of oxygen.

Beyond enzymatic activity, the integrity of recombinant MTR can be assessed using protein quality assays. One key method is negative stain electron microscopy (EM), a technique that visualises proteins at low resolution using a staining reagent that enhances contrast. The stained sample is applied to a grid, and electrons passing through the grid produce images that allow the visualisation and analysis of protein samples, albeit without detailed structural information (Scarff, Fuller et al. 2018). Negative stain EM is commonly used as a quality control step before proceeding to more advanced techniques such as cryo-EM (Skiniotis and Southworth 2015, Glaeser, Nogales et al. 2021).

This chapter investigates the production and characterisation of recombinant full-length human MTR, evaluating its biophysical properties and sample quality for cryo-EM. While overexpression of MTR followed established literature protocols, a novel approach was employed for purification. Additionally, various assays were implemented to characterise the purified protein, including analysis of ligand binding, higher oligomeric states, homogeneity, and enzymatic activity.

3.2 Recombinant human MTR is expressed and purified using an insect cell system

*3.2.1 Design and construction of M*TR *expression plasm*ids

As outlined in Section 3.1, full-length human MTR has previously been expressed and purified from insect cells using either a maltose-binding protein fusion (Mascarenhas, Gouda et al. 2022) or a cobalamin-affinity column (Yamada, Gravel et al. 2006, Wolthers

and Scrutton 2009, Zhang, Liu et al. 2020). The MTR construct used in this study corresponds to the near full-length human protein (amino acids Lys16-Asp1265), incorporating a terminal tobacco etch virus (TEV)-cleavable His_{10}-3xFLAG fusion tag at the C-terminal, resulting in a protein of approximately 142 kDa, as detailed in Section 2.1. The recombinant protein includes all five functional domains: the homocysteine-binding domain (HcyBD), folate-binding domain (FolBD), Cap domain, cobalamin-binding domain (CblBD), and activation domain (AD).

3.2.2 Expression in insect cells

The pLIB vector containing the MTR sequence was first amplified in a monolayer of *Sf9* cells, followed by two rounds of amplification in suspension cultures to generate baculovirus at higher concentrations. Large-scale expression was carried out in *Sf9* cells at a density of 1×10^6 cells/mL, with 600 mL of culture in each 2 L glass Erlenmeyer flask. For each flask, 15 mL of baculovirus were added, and the cultures were incubated under agitation at 27°C for 72 hours. Each large-scale expression batch consisted of at least 12 flasks, and the cells were processed immediately following the incubation period with baculovirus, as described in Section 2.1.2.

3.2.3 Purification of human MTR

Recombinant MTR was purified using Ni-NTA affinity chromatography, followed by Superdex 200 size exclusion chromatography (Figure 3.1). Briefly, cells from 12 flasks, each containing 600 mL of *Sf9* culture, were harvested for the purification process. During the Ni-NTA step, substantial amounts of a protein with an approximate molecular weight of 140 kDa were detected across all analysed samples, particularly in the second elution. Given the strong affinity of the His_{10} tag on MTR for nickel, the majority of the recombinant protein

was expected to appear during the elution. However, a portion of the protein that did not bind to the resin was still present in the wash fractions and flow-through.

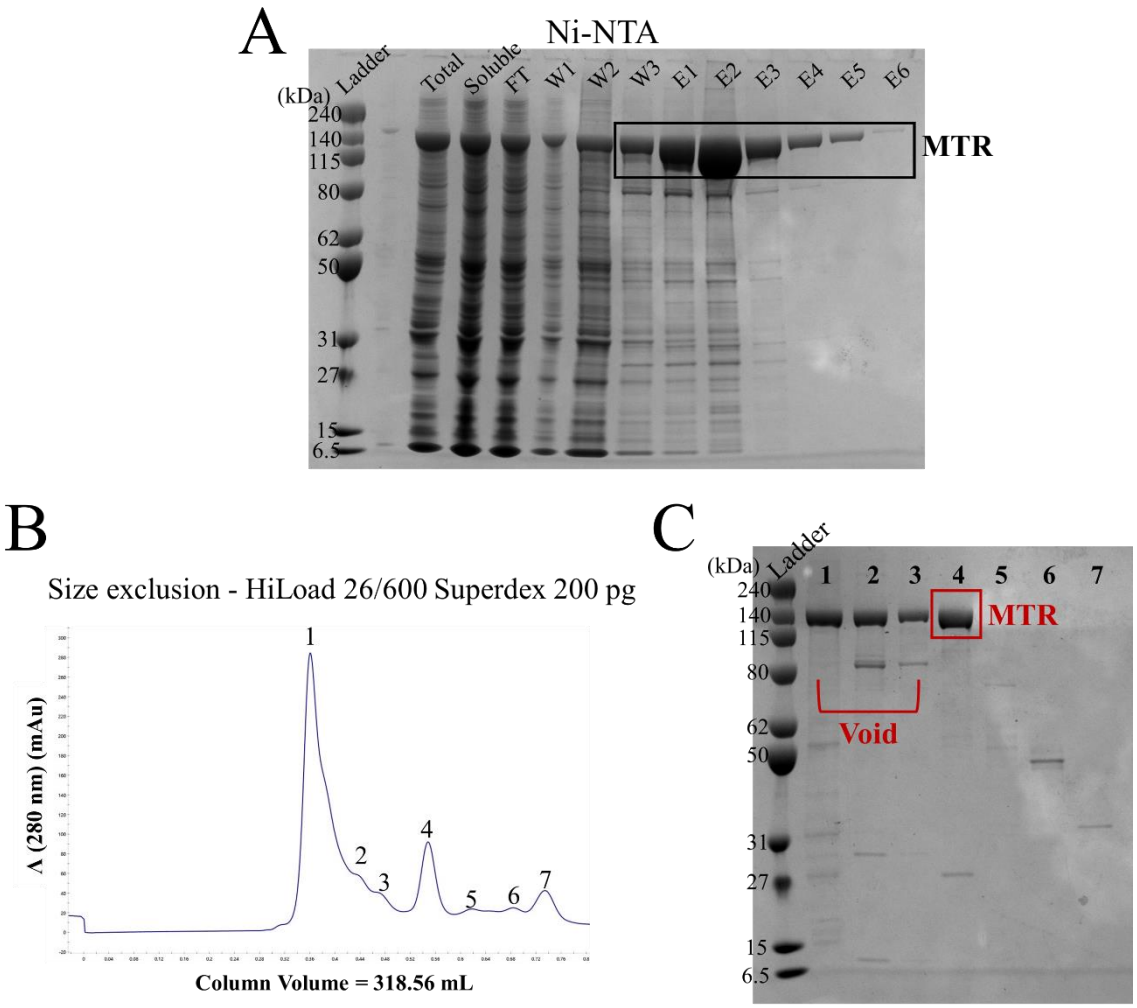

Figure 3.1: Purification of full-length MTR was carried out using Ni-NTA chromatography followed by Superdex 200 size exclusion chromatography. A. SDS-PAGE analysis of the Ni-NTA purification of recombinant MTR, showing the ladder, total sample, soluble fraction, flow-through (FT), washes (W1-3), and eluates (E1-6). The black square highlights the 142 kDa MTR band, and the samples pooled for size exclusion chromatography. **B.** Chromatogram from the size exclusion chromatography, with numbers 1-7 representing the different peaks in the sample. **C**. SDS-PAGE of the peaks identified in the chromatogram, showing MTR detected in the void volume and the MTR found in peak 4.

Although significant quantities of the recombinant protein were present in the elution fractions, contaminants were also observed in these samples. To remove these impurities, size exclusion chromatography was performed using a 26/600 Superdex 200 column (Figure 3.1B). The third wash fraction and all elution fractions were pooled and loaded into the

column for the second purification step. The full-length tagged MTR has a molecular weight (MW) of approximately 142 kDa, and the analysis of the size exclusion chromatogram peaks (Figure 3.1C) revealed the presence of a protein with this approximate MW in the initial four peaks. In addition to the MW, the abundance of this protein in the sample strongly suggests it is overexpressed recombinant MTR.

The majority of the MTR was found in the void (peak 1, Ve = 0.36), indicating potential aggregation or formation of higher-order oligomers (Duong-Ly and Gabelli 2014). The void also contained two additional "shoulders" (peaks 2 and 3, Ve = 0.42 and 0.44), which, upon SDS-PAGE analysis, revealed the presence of MTR along with some contaminants. The bands around ~80 kDa, observed in peaks 2 and 3, are likely contaminants interacting with MTR, such as endogenous MTRR from the insect cells, or represent degradation products of MTR. However, due to their proximity to the void, it is also possible that peaks 2 and 3 correspond to higher oligomeric forms of MTR as well.

Peak 4, which contained MTR (Ve = 0.54), had a lower absorbance at 280 nm compared to the void, indicating a reduced protein concentration. Nonetheless, this peak was not located in the void, and fewer contaminants were present in the sample. The elution of this protein after the void suggests it may represent properly folded and non-aggregated MTR. The other peaks (5-7) did not contain any protein of approximately 140 kDa, suggesting they consisted mainly of contaminant proteins that also interacted with the nickel during the initial purification step.

The MTR sample in the void fractions was not discarded, as the quality of the protein remained unknown, and this fraction represented over 70% of the total MTR obtained during purification. It is unclear whether the MTR in the void is monomeric or active, as the purification strategy employed here (Ni-NTA followed by size exclusion chromatography)

has not previously been applied to MTR. Earlier studies have demonstrated that heterologous expression of MTR produces predominantly *apo* MTR (Yamada, Yamada et al. 1999, Yamada, Gravel et al. 2006, Zhang, Liu et al. 2020), but during cobalamin-affinity purification, the apoenzyme is converted to the holoenzyme. Therefore, it is uncertain whether the absence of cobalamin-bound MTR resulted in the majority of the protein being eluted in the void, likely indicating aggregation.

Although most of the expressed MTR was eluted in the column's void volume, some protein could still be recovered in the subsequent peak (peak 4). The yield of purified MTR from size exclusion chromatography varied between purifications, but typically, 12 flasks of 600 mL Sf9 culture yielded approximately 1.5-2.0 mg of protein. In all purifications performed, the majority of the recombinant MTR was found in the void of the size exclusion chromatography, and the total recovery of MTR did not exceed 2 mg for 12 flasks. Consequently, each expression was conducted with at least 12 flasks, as smaller-scale expressions resulted in significantly lower yields of protein. Despite the modest recovery of recombinant MTR, the assays conducted with the protein do not require large quantities.

3.3 MeTHF increases MTR's melting temperature on Differential Scanning Fluorimetry (DSF)

Following the purification of recombinant MTR (Section 3.2), the protein present in peak 4 of the size exclusion chromatography (Figure 3.1C) was used to assess the impact of various ligands on its melting temperature. Due to the modular nature of MTR's domains, the protein can adopt different conformations depending on which domain the CblBD is interacting with, as depicted in Figure 1.6 (Datta, Koutmos et al. 2008). The conformational changes in MTR are primarily driven by the oxidation state of cobalamin (Cbl), which can

vary between +1, +2, or +3, and by the substrate or product bound to the protein (Jarrett, Huang et al. 1998).

Differential Scanning Fluorimetry (DSF) is a biophysical technique used to determine a protein's melting temperature (Tm), which indicates the temperature at which half of the protein in solution is unfolded. The assay relies on monitoring the fluorescence of a dye, such as Sypro Orange, that binds to the hydrophobic regions of proteins as they unfold.

During the DSF process, the temperature is gradually increased, causing the protein to denature and exposing its hydrophobic regions. This exposure leads to the binding of the fluorescent dye, resulting in an increase in the fluorescent signal. The determination of Tm under different conditions, such as in the presence of various ligands, can reveal protein stabilisation, indicated by an increase in Tm (Niesen, Berglund et al. 2007).

3.3.1 Thermal stability of MTR

Protein-ligand interactions can induce conformational changes in the protein, sometimes leading to stabilisation and, consequently, a higher melting temperature (Tm). Since DSF detects protein unfolding across a temperature gradient, this assay can be used to determine whether a ligand stabilises the protein (Niesen, Berglund et al. 2007).

To investigate how Hcy, Met, OHCbl, MeTHF, THF, and SAM affect the thermodynamic stability of MTR, the DSF assay was employed. By incubating MTR with each of these ligands and subjecting the samples to DSF, changes in MTR stability can be inferred from the corresponding Tm values. This assay provides an initial assessment of which ligands could be included in subsequent structural studies using cryo-EM.

To evaluate the effect of these ligands on MTR stability, 50 µM of each ligand was incubated with 0.4 mg/mL of protein and Sypro Orange in triplicate. Three wells containing

only the protein and dye were included to determine the Tm of the as-purified MTR for comparison. As described in Section 2.2.1, the temperature gradient ranged from 25 °C to 95 °C, with an increment of 0.5 °C per minute.

After the run was completed, the relative fluorescence of each sample was measured (Figure 3.2A), and the Tm was determined from the absolute fluorescence. The changes in Tm relative to the control (as-purified MTR) were then represented in a bar graph (Figure 3.2B).

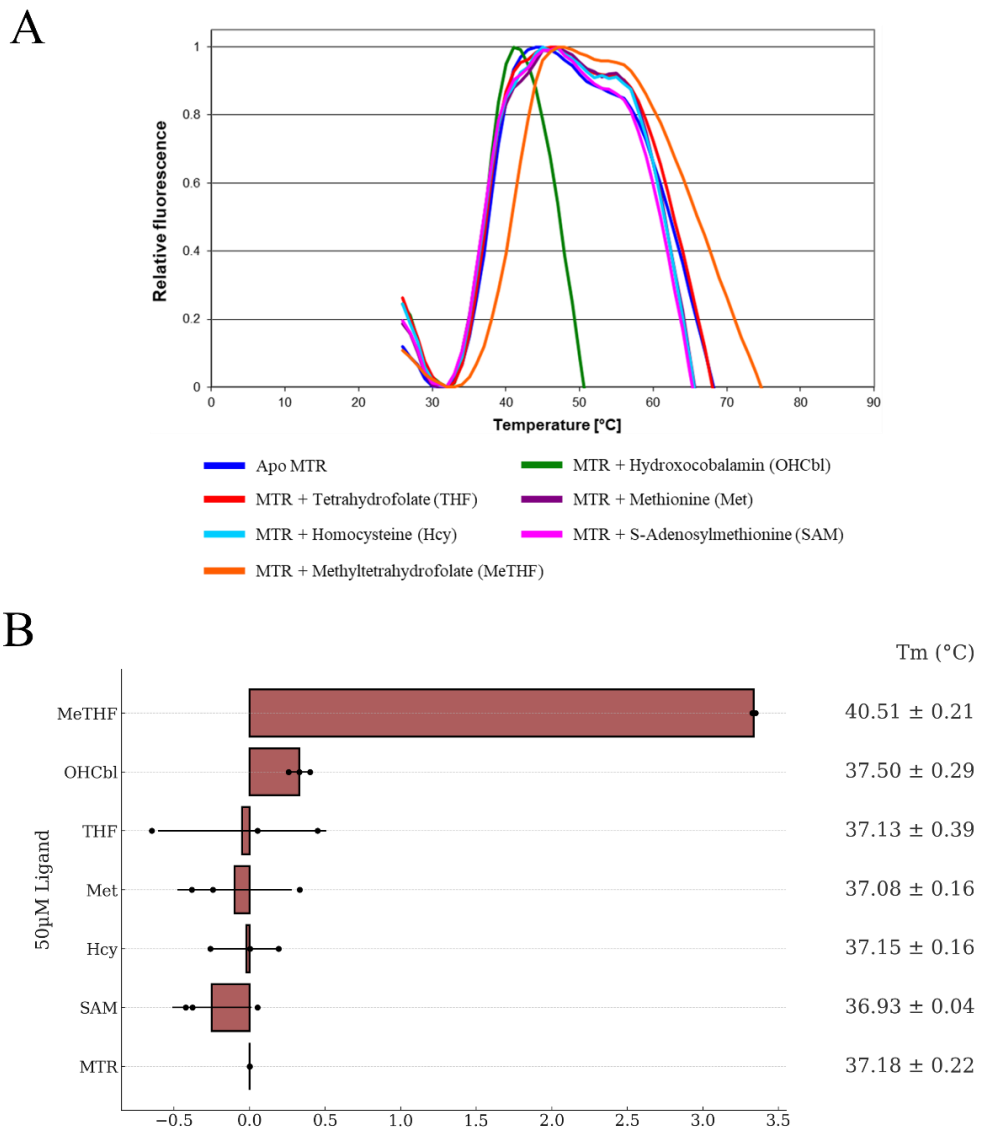

Figure 3.2: The effects of ligands on MTR stability were assessed using DSF. A. Melting curves display the relative fluorescence of as-purified MTR and MTR in the presence of various ligands. This data is representative of three independent samples. **B**. Changes in the melting temperature of MTR incubated with different ligands, compared to the as-purified protein. Error bars in B indicate the standard deviation of three technical replicates, while the black dots represent the ΔTm for each individual replicate.

The fluorescence curves presented in Figure 3.2A represent the relative fluorescence of the samples, which provides greater clarity and comparability than absolute fluorescence values, as the latter can be influenced by various factors such as background noise and baseline fluctuations (Moreau, Morin et al. 2012). These curves illustrate the relative fluorescence at each temperature increment, indicating that the rise in fluorescence is associated with protein unfolding. Thus, a delayed increase in fluorescence suggests that the protein is more stable. The eventual decrease in fluorescence, observed towards the end of the assay, signifies protein aggregation, which traps the dye and reduces fluorescence (McClure, Ahl et al. 2018).

The Tm, shown in Figure 3.2B, is determined by identifying the average baseline and plateau fluorescence values, and calculating the temperature at which the fluorescence reaches the midpoint between these two values. This Midpoint Method provides an estimate of the temperature at which approximately 50% of the protein is unfolded (Niesen, Berglund et al. 2007).

In the DSF runs, as-purified MTR at a concentration of 0.4 mg/mL exhibited a Tm of 37.17 ± 0.18 °C without any ligands. Considering that MTR is a human protein, a Tm of 37.1 °C is considered relatively low, given that it closely matches normal human body temperature, indicating instability in the as-purified protein. This value served as the baseline to assess changes in Tm (ΔTm) when each ligand was added in subsequent assays at concentrations of 50 μM (approximately 17-fold molar excess relative to the protein).

Among the ligands tested, only MeTHF caused a significant increase in ΔTm, raising the Tm by more than 3 °C compared to as-purified MTR. The other ligands did not exhibit any notable changes in Tm. However, the absence of a ΔTm change does not necessarily

imply a lack of ligand binding, as these ligands are known to interact with MTR during various stages of the catalytic or recovery cycles. Therefore, the lack of observed ΔTm for most ligands likely indicates that their binding does not significantly affect the protein's thermostability.

All tested ligands, except for OHCbl, displayed a subtle second peak in the DSF curves (Figure 3.2A), which was not considered when calculating the Tm. The appearance of a second peak might indicate multiple unfolding phases within the protein, potentially attributable to the multidomain structure of MTR (Augustijn, Mahapatra et al. 2019). The absence of a second peak in the MTR+OHCbl sample could be explained by conformational changes that did not alter the Tm of the protein but instead altered its unfolding behaviour. However, another explanation is that the second peak was masked by the absorbance of cobalamin, which has an absorption range between 500-580 nm, close to the emission wavelength of the Sypro Orange dye (586 nm), potentially interfering with the DSF signal.

3.4 Recombinant MTR is active on miniaturized activity assay

As introduced in Section 3.1, several enzymatic assays have been developed to measure the activity of recombinant MTR proteins (Frasca, Banerjee et al. 1988, Chen, Chakraborty et al. 1995, Drummond, Jarrett et al. 1995). To assess whether the recombinant human MTR generated in this study is active, the non-radioactive aerobic assay (Drummond, Jarrett et al. 1995) was modified for use in a 96-well plate format. This assay relies on derivatising THF with a formylating agent under acidic conditions to form methenyltetrahydrofolate, with absorbance measured at 350 nm. Previous assays used cuvettes and a final reaction volume of 1 mL. Miniaturising the assay enables simultaneous processing of multiple reactions and reduces the volume of reagents required, including recombinant MTR.

The principle of the assay followed the previously described protocol (Drummond, Jarrett et al. 1995), with slight modifications to all volumes and the final MeTHF concentration, as detailed in Section 2.2.2. Briefly, a reaction containing final concentrations of 25 µM DTT, 152 µM SAM, 5 mM Hcy, MTR at varying concentrations, and 50 µM OHCbl in a final volume of 37.6 µL PBS (100 mM) was incubated at 37 °C for 5 minutes. MeTHF was then added to a final concentration of 400 µM, bringing the total reaction volume to 40 µL. The reaction was incubated for an additional 10 minutes at 37 °C, then quenched with 10 µL of 5 N HCl in 60% aqueous formic acid, diluted to a final concentration of 1 N in a 50 µL reaction. Following a 10-minute incubation at 80 °C, the absorbance was measured at 350 nm.

The initial incubation without MeTHF allows DTT and OHCbl to reduce oxygen to H_2O_2, creating a semi-anaerobic environment and reducing the oxidation rate of Cbl during the assay. This environment eliminates the need for an anaerobic chamber, as Cbl oxidation leads to MTR inactivation. The second incubation, following the addition of MeTHF, allows MTR to perform its catalytic cycle, producing Met and converting MeTHF to THF, which is subsequently converted to methenyltetrahydrofolate in the final incubation, providing the assay readout.

The activity assay, summarised in Figure 3.3A, was optimised to determine the ideal final concentration of MTR (Figure 3.3B) and the optimal duration of the second incubation, during which MTR performs its catalytic cycle (Figure 3.3C). Each MTR concentration or timepoint was tested in triplicate, and the absorbance readings from each well were corrected by subtracting the average absorbance of two blank wells. The blank wells, which did not contain MTR, accounted for the intrinsic absorbance of the assay components.

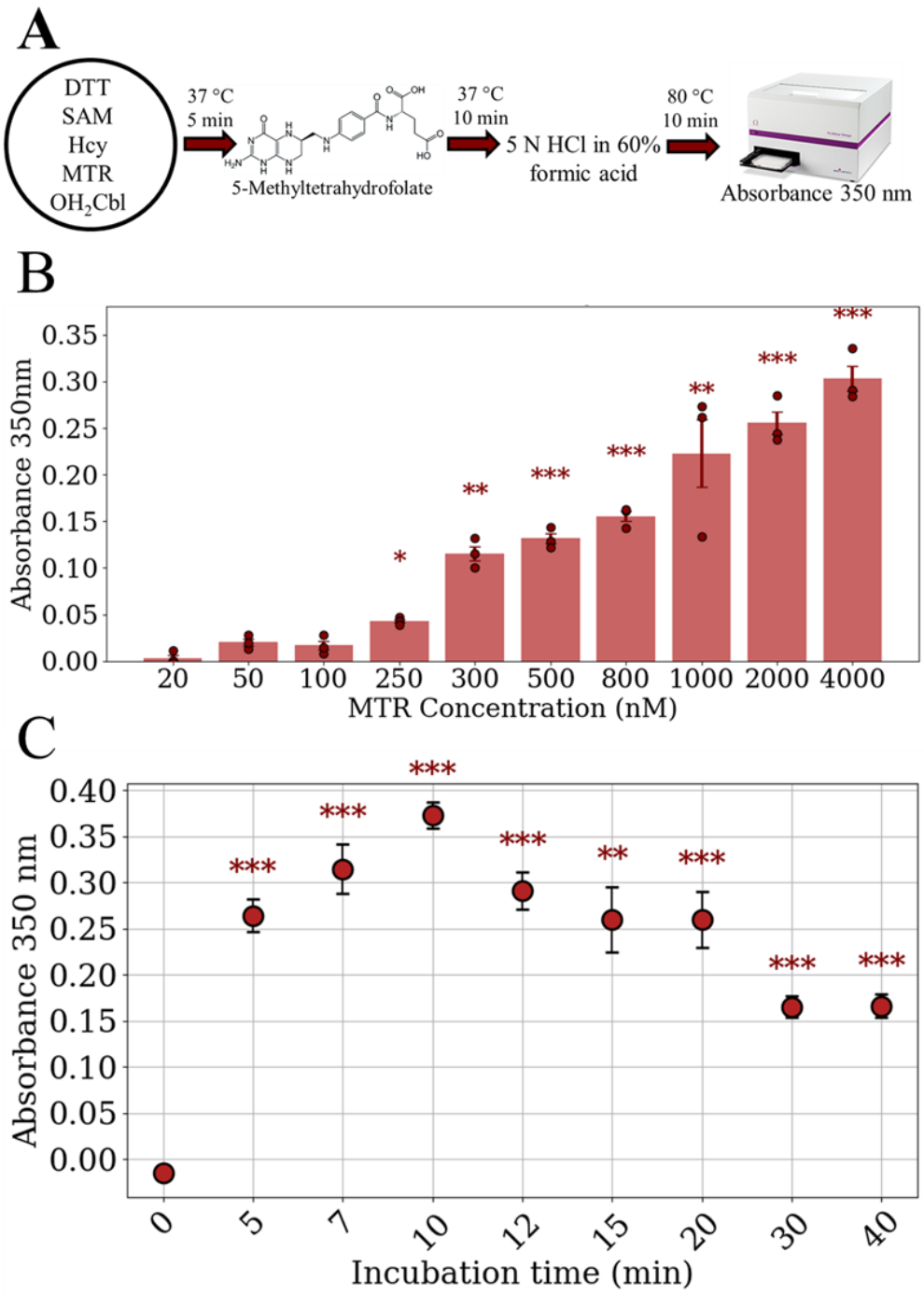

Figure 3.3: MTR activity assay. A. Schematic representation of the MTR activity assay. **B**. The activity assay was performed with varying concentrations of MTR, showing the absorbance at 350 nm for each concentration. **C**. The assay was conducted with different incubation times after the addition of MeTHF. Each time point represents the average reading from three technical replicates. For both B and C, error bars indicate the standard deviation. Statistical significance is denoted as *p > 0.01, **p > 0.001, ***p < 0.001 when compared to the 20 nM concentration (B) or 0 min incubation (C).

The 96-well plate assay allows for higher throughput measurements of enzymatic activity compared to previously published procedures, enabling multiple conditions to be

tested in parallel while using a smaller amount of MTR. An increasing concentration of MTR (10-2000 nM final concentration) was titrated to confirm that the recombinant protein activity is dose-dependent and to identify the optimal MTR concentration for future assays. As expected, a rise in absorbance was observed with increasing MTR concentrations (Figure 3.3B), demonstrating a clear correlation between MTR concentration and THF production.

In previous assays, a 10-minute incubation at 37 °C was used to allow MTR to complete its catalytic cycle. To optimise the incubation time for this assay, varying durations were tested post-MeTHF addition, using a final concentration of 500 nM MTR (Figure 3.3C). This assay was done to identify the incubation time that would provide the highest and most consistent readout. For each time point, three wells were used as triplicates for the assay, with two wells lacking the enzyme as blanks.

Different factors need to be considered when determining the optimal reaction time due to the nature of MTR, cobalamin, and the THF formed during the assay. It is well-known that after 200-1000 catalytic turnovers, cob[I]alamin can be oxidised to cob[II]alamin, a non-productive form of cobalamin (Fujii, Galivan et al. 1977). The oxidation rate increases over time due to the aerobic environment in which the assay is conducted, which can lead to higher rates of protein inactivation with longer incubation periods. Another important consideration is the stability of THF, the molecule measured to assess MTR activity. The half-life of THF at 37 °C is less than 30 minutes (Pollock and Kaufman 1978), and it can degrade further in the presence of oxygen (Zheng and Cantley 2019). As shown in Figure 3.3C, MTR activity, indicated by THF production, increases over time, reaching a peak at 10 minutes. After this point, a combination of MTR inactivation and THF degradation reduces the absorbance readings, compromising the correlation with protein activity.

The assay was repeated under anaerobic conditions to confirm the impact of oxygen on MTR inactivation and THF degradation. Before the assay, all reagents and the protein were incubated for 30 minutes with open lids in a Whitley A35 anaerobic workstation, using a gas mixture of 80% N_2, 10% CO_2, and 10% H_2. The MTR concentration and reagent amounts were kept constant (500 nM MTR final concentration), and the assay was conducted with incubation times of 10, 30, and 40 minutes. In parallel, the same incubation times were tested under aerobic conditions. All assays were performed in technical triplicates, with the absorbance values corrected by subtracting the average of the two blanks for each sample (Figure 3.4).

In aerobic conditions, MTR activity remained stable between 10 and 30 minutes but declined at 40 minutes. In contrast, the anaerobic assay showed increased THF production up to 40 minutes, confirming that oxygen interferes with the assay results by inactivating MTR and degrading THF. This indicates that while the assay can be reliably conducted in aerobic conditions for activity confirmation, the incubation time should not exceed 10 minutes to avoid interference from oxygen.

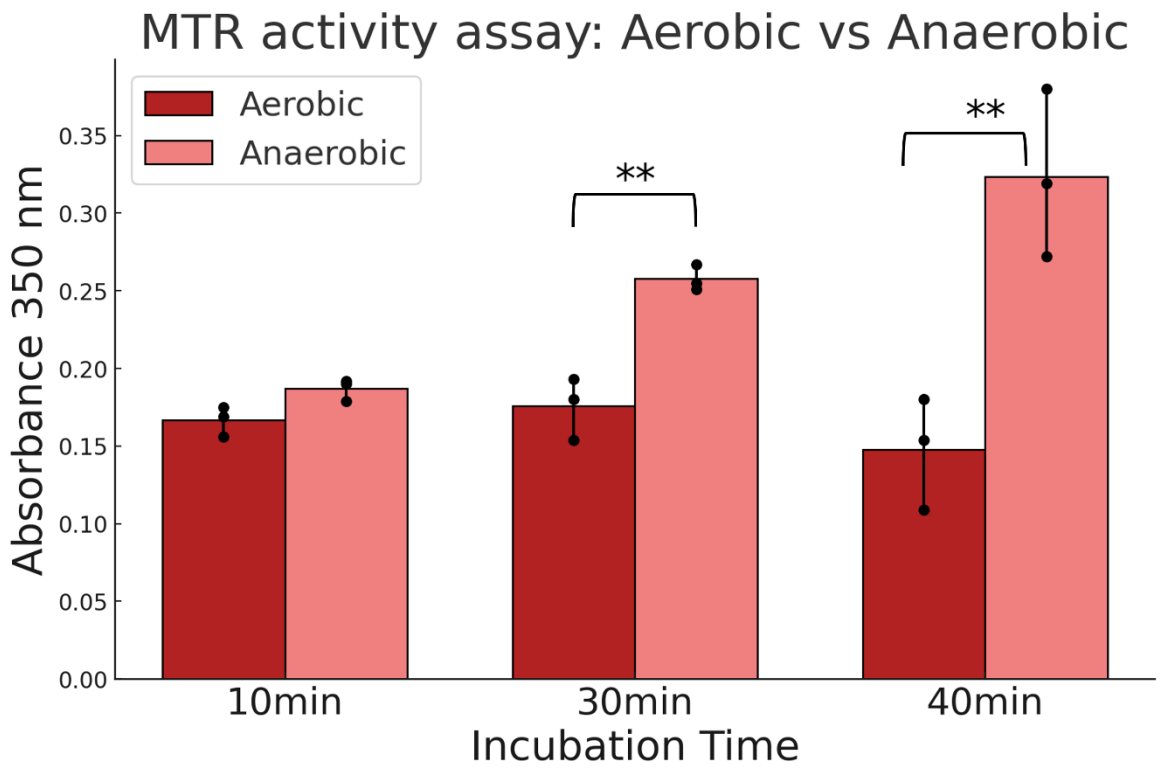

Figure 3.4: MTR activity assay in aerobic vs anaerobic conditions. Graph showing the difference in MTR activity measured in aerobic (dark red) and anaerobic (light red) environments at three different time points. Each bar represents the average of three technical replicates, with error bars indicating the standard deviation. Statistical significance is denoted by $p < 0.01$, comparing the aerobic and anaerobic readouts for each incubation time.

After determining the optimal protein concentration and reaction time for the assay, two purified MTR samples were tested in parallel to compare their activity. For this assay, the protein obtained from the void fraction (peak 1) of the size exclusion chromatography was tested alongside the MTR eluted in the subsequent peak 4 (Figure 3.1C). The results showed that the protein purified from the void lacked activity, while the MTR from the monomeric peak 4 exhibited the expected enzymatic activity (Figure 3.5). The experiment was conducted in triplicate and the data were blanked against the average reading of four wells containing no MTR.

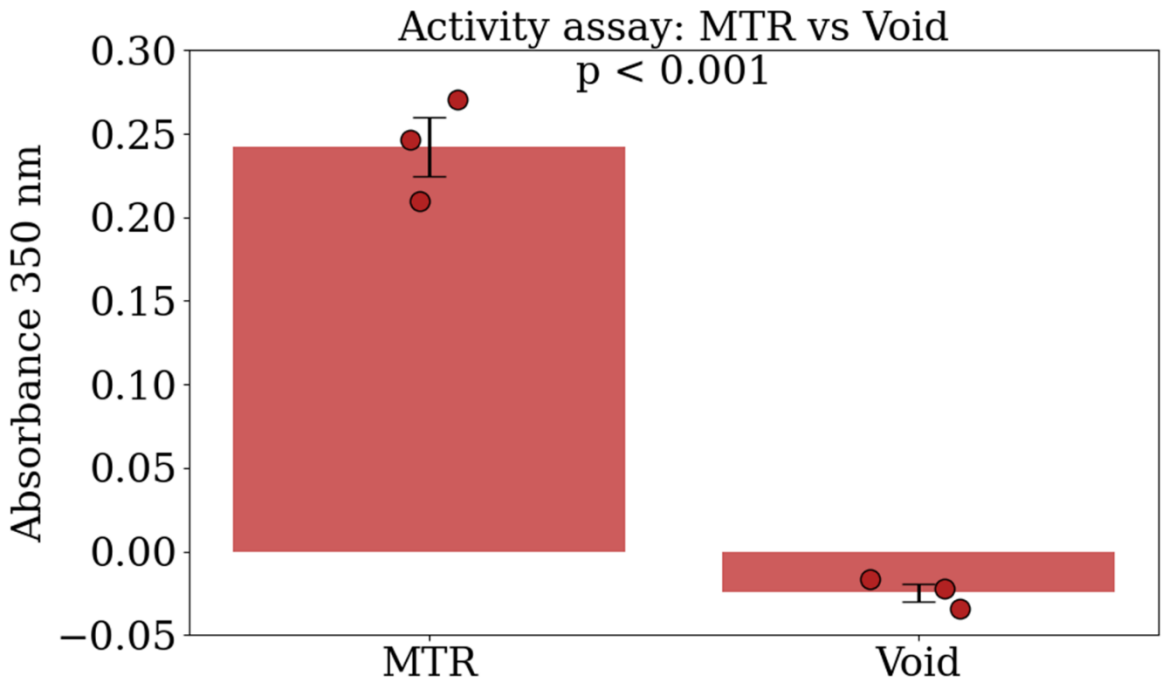

Figure 3.5: Activity assay of MTR purified from the void and monomeric peak of the size exclusion column. Each assay was performed in technical triplicates, with error bars indicating the standard deviation. The p-values displayed above the graphs represent the statistically significant difference between the two samples.

In conclusion, the recombinant human MTR expressed and isolated in this study is enzymatically active, though a significant portion of the expressed protein, eluted in the void of the size exclusion column, lacks activity. While oxygen can interfere with longer incubation times, it remains feasible to assess protein activity using the previously described 10-minute incubation.

3.5 MTR is homogeneous and monomeric

After confirming the activity of purified human MTR, the protein was subjected to native mass spectrometry (MS) and mass photometry (MP) to determine its molecular mass, oligomeric state, and any ligands bound to the protein. Dr. Rod Chalk from CMD University of Oxford conducted the native MS analysis to identify the oligomeric states of MTR and detect the presence of ligands. The MTR sample, at 2 mg/mL in 50 mM ammonium acetate (pH 6.5), was analysed as described in Section 2.2.3. The results showed that MTR has the

expected mass, confirmed by intact protein mass determination, where 1 µL of protein at 2 mg/mL was diluted in 59 µL of 0.1% formic acid and analysed using LC-QTOF. The protein mass was measured with an accuracy of ±0.5 Da and compared to the theoretical mass expected for the construct (Figure 3.6A).

The MS analysis also revealed that no ligands were detected, indicating that the purified protein exists as an apoenzyme. Furthermore, the charge radius analysis demonstrated that most of the enzyme was in a folded monomeric form (~5.5 x 10^2), with some unfolded (~2.4 x 10^2) and partially unfolded (~1 x 10^2) proteins, as well as folded dimers (~0.2 x 10^2) (Figure 3.6B).

Mass photometry (MP) was also employed to analyse the purified MTR using a MassFluidix HC (Refeyn) in a demonstration at Newcastle University (Section 2.2.4). Briefly, 10 µL of as-purified MTR at 10 nM in 50 mM HEPES was loaded into the MassFluidix HC chip and analysed. The light scattering of the molecules was assessed and the data analysed using the built-in software. Although native MS offers higher mass resolution than MP (den Boer, Lai et al. 2021), MP still successfully identified that approximately 70% of the purified protein molecules behaved as monomers and were of the expected size (Figure 3.6C). The 137 kDa protein identified in the graph corresponds to MTR, with a slightly lower molecular weight than MS, likely due to the lower resolution of MP, which is expected to slightly underestimate the mass compared to the theoretical 142 kDa. Additionally, some smaller proteins of approximately 55 kDa were detected, representing about 27% of the total molecules. These smaller proteins are likely degradation products of MTR or protein contaminants that were not completely separated during the purification process.

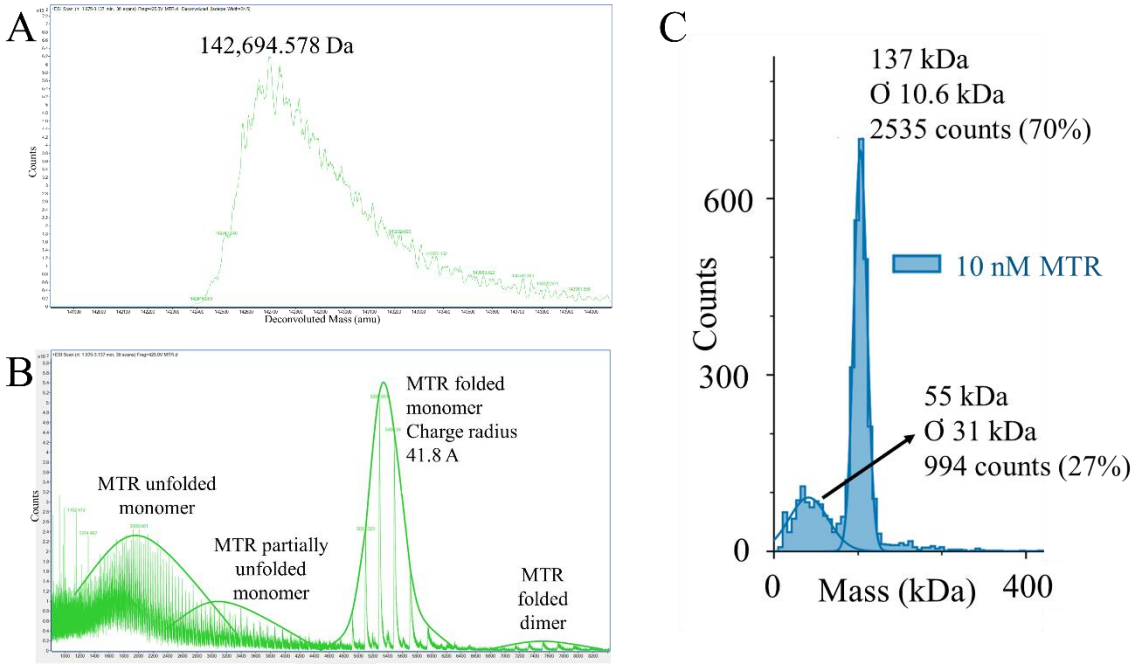

Figure 3.6: Native mass spectrometry and Mass Photometry analysis of MTR. A. Native mass spectrometry of as-purified MTR, displaying the mass of the denatured protein in a graph of count vs deconvoluted mass (amu). **B**. Native MS spectrum (count vs mass-to-charge (m/z)) of MTR, showing the calculated oligomeric states for each peak and the charge radius for the folded monomer. **C**. Histogram of the mass photometry (MP) signal distribution of MTR, with solid lines representing the Gaussian fits.

The presence of other states of MTR or contaminants, as shown by MS and MP, is not a concern, as the majority of the sample consists of the folded monomer, which is the expected state of MTR. This minor proportion of other molecules will not affect the subsequent steps in studying MTR, which involve using the protein for cryo-EM. The cryo-EM technique allows for the selection of single particles during dataset analysis (Section 2.7.3), enabling the exclusion of contaminants without compromising the quality of the data for the target protein.

Both native MS and MP proved to be effective strategies for sample screening prior to cryo-EM. These methods facilitate the identification of the protein's oligomeric state, with MP offering particularly fast results and requiring only minimal amounts of protein. MS is also highly efficient at detecting ligands bound to the sample, which can be valuable for

confirming ligand binding before preparing the protein for cryo-EM analysis (Olinares, Kang et al. 2021).

3.6 Negative staining reveals low-resolution structure features of MTR

Before cryo-EM grid preparation, MTR was analysed using negative staining electron microscopy (EM), which enables visualisation of proteins at low resolution. While samples may behave differently in cryo-EM due to vitrification, negative stain serves as a valuable tool to assess sample quality, including concentration and heterogeneity, prior to cryo-EM, which is a more expensive and time-consuming technique (Glaeser, Nogales et al. 2021).

Both the MTR monomer sample and the void sample were examined by NS to compare their characteristics and further investigate the lack of enzymatic activity in the void sample, as it was hypothesised that the MTR in the void might be degraded or aggregated. Carbon-coated grids were loaded with 10 µL of sample at 40 nM and stained with 1.5% uranyl acetate, as described in Section 2.4. Images were acquired using an HT7800 transmission electron microscope (Hitachi) at 60,000x magnification at the Electron Microscopy Research Services, Newcastle University (Figure 3.7A).

Electron micrographs of the void sample (Figure 3.7B) revealed aggregated proteins, characterised by larger particles with irregular shapes and variable sizes. In contrast, the MTR sample, identified as monomeric (Figure 3.7C), contained fewer aggregates and contaminants, with most particles appearing as monomeric, homogeneous MTR.

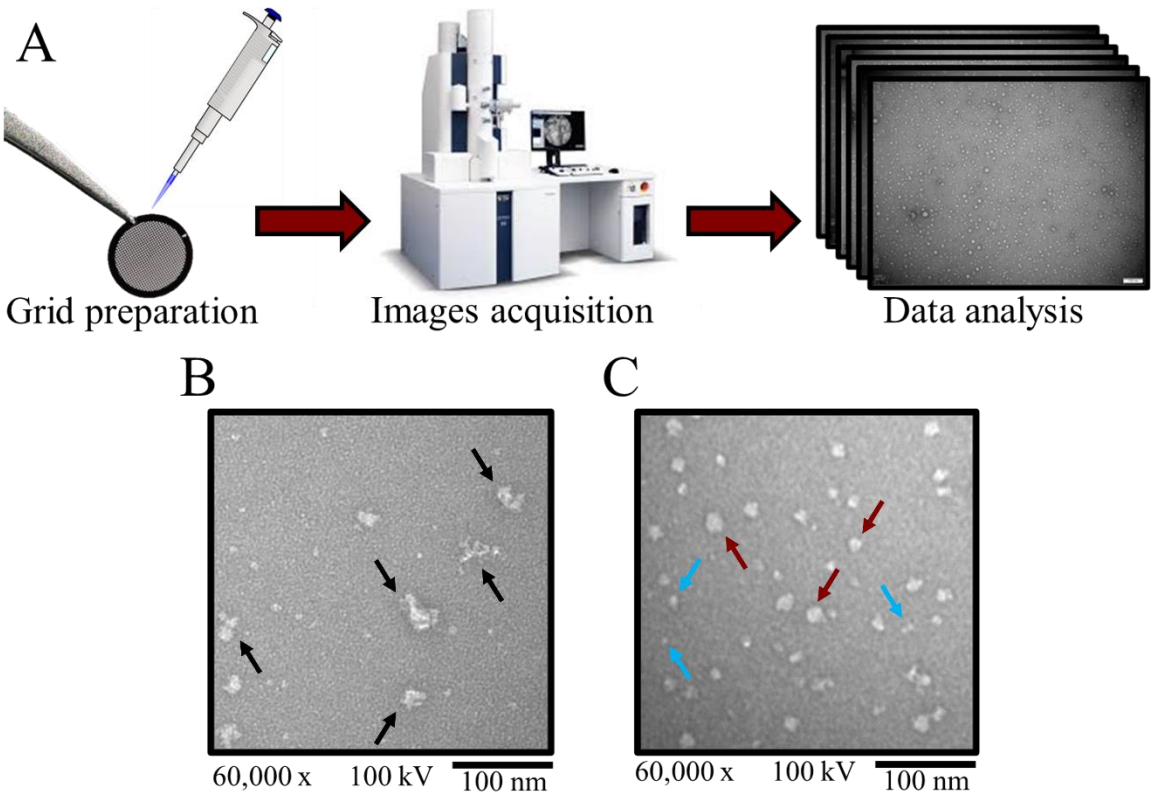

Figure 3.7: Negative Stain workflow and images of void and MTR samples. A. Diagram illustrating the negative stain EM workflow, where samples are applied to grids, images are captured using a transmission electron microscope, and the resulting images are analysed. **B**. Representative micrograph of the void protein, with black arrows indicating aggregated proteins. **C**. Representative micrograph of an MTR sample, with red arrows highlighting probable MTR molecules and blue arrows indicating degradation products or smaller contaminant proteins.

The difference in quality between these two samples correlates with the lack of activity in the void fraction. The quality assessments using both the activity assay and NS demonstrated that most of the MTR purified from the void is unsuitable for biophysical or structural studies due to its poor quality. However, as indicated by NS, the MTR eluted later in the size exclusion chromatography is capable of performing its catalytic cycle, is monomeric, and is homogeneous, making it suitable for further structural characterisation using cryo-EM.

3.7 Discussion

The data presented in this chapter display the efforts to express recombinantly and purify full-length human MTR. The insect-cell-expressed protein is purified in a novel two-step process. This purification is the first two-step process that yields to *apo* MTR, while the other purification of recombinant MTR uses either Cbl to purify the protein, leading to most of MTR being in the holoenzyme form (Yamada, Gravel et al. 2006, Wolthers and Scrutton 2009, Zhang, Liu et al. 2020), or uses a three-step purification process which includes the use of MBP-fused tag which has to be latter removed since it is a long tag (~42 kDa) (Mascarenhas, Gouda et al. 2022).

In this study, after Ni-NTA purification facilitated by the histidine tag added to the C-terminal end of MTR, the eluate was further purified using size-exclusion chromatography. Unfortunately, most of the protein eluted in the void of the column and displayed no enzymatic activity (Figure 3.5). Moreover, negative stain analysis revealed that the protein in the void fraction was mostly aggregated, making it unsuitable for structural assays. This aggregation is likely related to the lower stability of the *apo* MTR form (Zhang, Liu et al. 2020). However, a portion of the MTR was eluted later, suggesting that this population may be suitable for further assays and structural data collection (Figure 3.1). Biophysical analysis confirmed that this MTR is in the *apo* state (Figure 3.6) and exhibits enzymatic activity (Figure 3.3). Negative stain EM analysis (Figure 3.7) indicated that the protein is homogeneous with virtually no aggregates, suggesting it is appropriate for further structural studies.

The loss of recombinant MTR due to aggregation is likely a result of the *apo* state of the enzyme. Although a significant portion of the protein was lost during purification and only a small yield was recovered, the fraction of MTR obtained was of high quality. In addition to the low levels of contaminants and confirmed enzymatic activity, the negative

stain EM analysis indicates that the sample is suitable for cryo-EM. Moreover, the *apo* state of the protein allows for a broader range of ligand combinations to be tested in structural studies. In contrast, using cobalamin-affinity chromatography would require MTR to be in the holoenzyme state, limiting the ability to study the apoenzyme or different ligand combinations.

Besides the integrity of the recombinant MTR. its thermal stability was also assessed with different ligands using DSF. Interestingly, the relative fluorescence detected during the assay (Figure 3.2A) revealed two distinct melting events. The presence of multiple peaks in DSF can be attributed to the multidomain nature of proteins, where different regions or domains exhibit different Tm values, leading to two or more peaks. In MTR, these two melting events may correspond to two segments of the protein corresponding to the N-terminal half (HcyBD and FolBD) and the C-terminal half (Cap, CblBD, and AD). This hypothesis is based on the conformation of *apo T. thermophilus* MTR, which revealed it is in the activation conformation, with these two distinct sections behaving as independent units due the flexible loop between them (Mendoza, Purchal et al. 2023). If human MTR adopts a similar configuration in the *apo* state, the two melting events could represent these N and C-terminal halves unfolding at slightly different temperatures.

The addition of OHCbl in the DSF assay caused the melting curve to become narrower, displaying a single peak rather than the elongated two-peak profile. This change could be due to a significant conformational change in the protein upon OHCbl binding. However, bacterial MTRs still exhibit the same conformation as the *apo* state (activation conformation) upon OHCbl binding (Bandarian, Pattridge et al. 2002, Datta, Koutmos et al. 2008, Koutmos, Pejchal et al. 2008, Mendoza, Purchal et al. 2023). A more likely explanation is the interference of OHCbl with the fluorescence reading. Cobalamin absorbs light at

wavelengths close to the excitation and emission wavelengths of the Sypro Orange dye, which can reduce the fluorescence signal and alter the curve profile.

Further analysis of the DSF data revealed that MeTHF was the only ligand to significantly increase MTR's Tm, by approximately 3.3°C. MeTHF is known to play a crucial role in initiating the primary turnover of methionine synthase in *E. coli*, triggering conformational changes that shift the enzyme from its resting state to the catalytic cycle (Banerjee, Frasca et al. 1990). Recent structural studies on *Thermus filiformis* methionine synthase confirmed that MeTHF induces conformational changes when bound to either cob[III]alamin or cob[I]alamin (Watkins, Wang et al. 2023). Watkins and colleagues demonstrated through small-angle X-ray scattering (SAXS) that MeTHF was the only ligand capable of inducing significant conformational changes in MTR, but in both cases, the prokaryotic enzyme was not in the *apo* state. The results from human MTR in this study suggest that MeTHF also stabilises human MTR but may not induce similar conformational changes due the possible different conformations upon MeTHF binding in the abovementioned cases, although this cannot be confirmed solely from DSF data.

Following thermostability testing, the enzymatic activity of the recombinant protein was also evaluated. While several MTR activity assays have been described, they typically use large volumes (1 mL), which can be protein-intensive and laborious, making it difficult to perform multiple reactions in parallel. This chapter describes a miniaturised MTR activity assay with a final volume of 50 µL, using a final MTR concentration of 500 nM. This setup allows for testing various assay variants in a single read without consuming large amounts of protein.

The assay demonstrated a positive correlation between protein concentration and activity (Figure 3.3B) and confirmed that incubation time is critical, as extended times can

reduce the measured activity due to THF degradation, as shown when the assay was performed in an anaerobic chamber (Figure 3.4). While the 96-well plate assay is sufficient to determine whether the protein is active and can reveal differences in activity based on concentration and incubation time, it is less suited for more detailed analytical measurements. For example, determining the Km for substrates could be challenging due to the low amounts of THF produced in a 50 μL reaction. Nevertheless, the assay is a valuable approach for quality control of recombinant protein before further biophysical and structural studies.

Overall, the two-step purification method employed in this study produced a high-quality sample of *apo* MTR that is active, monomeric, and homogeneous. This approach addresses the limitations of previous purification techniques for MTR, which often resulted in holoenzyme forms or required additional purification steps. The absence of bound ligands in the purified protein allows for a wider range of ligand combinations to be explored in cryo-EM studies, as discussed in Chapter 4. Additionally, the DSF assay identified MeTHF as a promising substrate for further structural investigation. These advancements lay the groundwork for future studies aimed at elucidating the structural features of MTR.

4. Cryo-EM data collection and processing

4.1 Background

Methionine Synthase (MTR) is a flexible, five-domain protein that undergoes various conformational changes depending on the oxidation state of cobalamin (Cbl) and the stages of the catalytic or reactivation cycles (Bandarian, Ludwig et al. 2003) (Section 1.4). The majority of MTR structures have been derived from bacterial orthologues in different species, such as *E. coli*, *T. maritima*, *T. thermophilus*, and *T. filiformis*, using a divide-and-conquer approach, in which protein domains are expressed recombinantly and purified for X-ray crystallography.

As outlined in Section 1.6 and Table 1.1, only two of the 28 available structures of MTR in the Protein Data Bank (PDB) are of the human protein, both determined by X-ray crystallography. Only one structure of an MTR orthologue encompassing all five domains has been obtained, using X-ray crystallography on the *T. thermophilus* protein (PDB: 8SSC) (Mendoza, Purchal et al. 2023). The protein was slightly truncated at the N- and C-termini to reduce its flexible loops, a common practice to enhance protein crystallisation (Mooij, Mitsiki et al. 2009) and employed in all MTR X-ray structures. Additionally, only a cryo-EM structure of MTR has been determined, which revealed the HcyBD, FolBD, Cap, and CblBD domains of *Thermus filiformis* methionine synthase (PDB: 8G3H, Section 1.6.3) (Watkins, Wang et al. 2023). This structure displays the Cap-on state of MTR, similar to that found in *E. coli* (Drennan, Huang et al. 1994), but with an unexpected interaction between the Cap, CblBD, HcyBD, and FolBD domains.

Cryo-EM is emerging as a promising approach for determining the structures of proteins and complexes, particularly those that are not amenable to crystallisation, such as full-length human MTR. Recent advancements in cryo-EM data analysis (Section 1.7) have

significantly improved the resolution of proteins smaller than 200 kDa (Wu and Lander 2020). Given MTR's size and flexibility, cryo-EM offers an excellent approach to studying its architecture and conformational changes.

In addition to experimental advances, algorithms for protein structure prediction based on machine learning, such as AlphaFold (AF) and AlphaFold2 (AF2), can now predict protein structures with remarkable accuracy (Section 1.8) (Jumper, Evans et al. 2021). AF2 can be a valuable tool for interpreting possible conformational states and understanding the dynamics of proteins, such as MTR. The further AF2 advancement now enables the prediction of protein complex structures, even in the presence of ligands, while maintaining its high accuracy in intra-domain predictions (Evans, O'Neill et al. 2021, Abramson, Adler et al. 2024).

The expression and purification of full-length human MTR in its *apo* form (Section 3.2) have significantly expanded the possibilities for structural studies. In this chapter, the combination of Cryo-EM data and AF2 predictions is utilised to reveal novel conformations and elucidate the structures of domains that have not yet been characterised for human MTR. To provide clarity and coherence, the subsections in this chapter will be presented in their chronological order of experiment timeline.

4.2 MTR on grids: initial screening

The initial dataset was collected using a 200 keV Glacios cryo-microscope equipped with a Falcon-IV detector at the York Structural Biology Laboratory (YSBL), as described in Section 2.7. The parameters selected were based on the protein's size (140 kDa) and the primary aim of assessing sample quality, rather than achieving a high-resolution structure at this point.

For sample preparation, 0.5 mg/mL of *apo* MTR was mixed with 1 mM MeTHF, 50 µM SAM, and 50 µM OHCbl in a buffer containing 250 mM NaCl and 50 mM HEPES. The reaction mixture was incubated for 30 minutes at 37 °C before grid application. The conformation that MTR would adopt with this tri-ligand combination (referred to thereafter as MTR+Ligands) was still unknown. The inclusion of MeTHF in the protein sample was based on Differential Scanning Fluorimetry (DSF) results, whereby this ligand produces the most significant increase in MTR's melting temperature (Section 3.3). The OHCbl and SAM ligands were added to stabilise the CblBD and AD.

The second sample consisted of 0.5 mg/mL of *apo* MTR in the same buffer, without the addition of any ligands. *Apo* MTR was shown to be monomeric and stable (Sections 3.3 and 3.5), making it a suitable candidate for generating a dataset that could reveal the protein's conformation prior to binding Cbl. Vitrification of the sample followed the parameters described in Chapter 2, using the Vitrobot. The selected parameters were the standard ones typically employed for sample preparation, with the possibility of adjustment based on the outcomes of data collection (e.g., if the ice on the grids is too thin or too thick). Two Quantifoil™ Au 300 mesh 1.2/1.3 grids were loaded with 3 µL of sample each.

Each of the four grids (two for each sample) was loaded into the Glacios, and 700 micrographs were collected from each sample for initial processing. This step was conducted to analyse the particles and assess grid quality before proceeding with large-scale dataset collection.

4.2.1 Sample preparation and data collection for MTR+Ligands

The dataset of MTR with MeTHF, SAM, and OHCbl (MTR+Ligands) was processed as described in Chapter 2. Using CryoSPARC, the micrographs were motion-corrected, and the contrast transfer function (CTF) was corrected. Particles were then picked using the "blob

picker" tool, with a diameter range of 80-150 Å, which selected over 379,000 particles. Most of the particles appeared as low-contrast blobs in a noisy background, which the blob-based particle picker algorithm utilised to select the particles. Given that the conformation of human MTR with these specific ligands is unknown, blob-picker is an effective strategy for initial particle identification.

The initial blob-picked particles for the MTR+Ligands sample were filtered, and two main particle classes were identified (Figure 4.1). The 2D classes presented some details corresponding to the two halves of MTR's five-domain architecture. The first set of 2D classes (Figure 4.1A) represents HcyBD and FolBD, collectively referred to as the N-terminal half (N-half), whereas the second set (Figure 4.1B) represents the Cap, CblBD, and AD, collectively referred to as the C-terminal half (C-half).

Throughout the analysis, 7,786 particles of the N-half and 22,962 particles representing the C-half were identified. The difference in particle numbers suggests a preferential orientation and/or partial denaturation, with fewer particles and 2D classes observed for the N-half (Figure 4.1A) than for the C-half (Figure 4.1B).

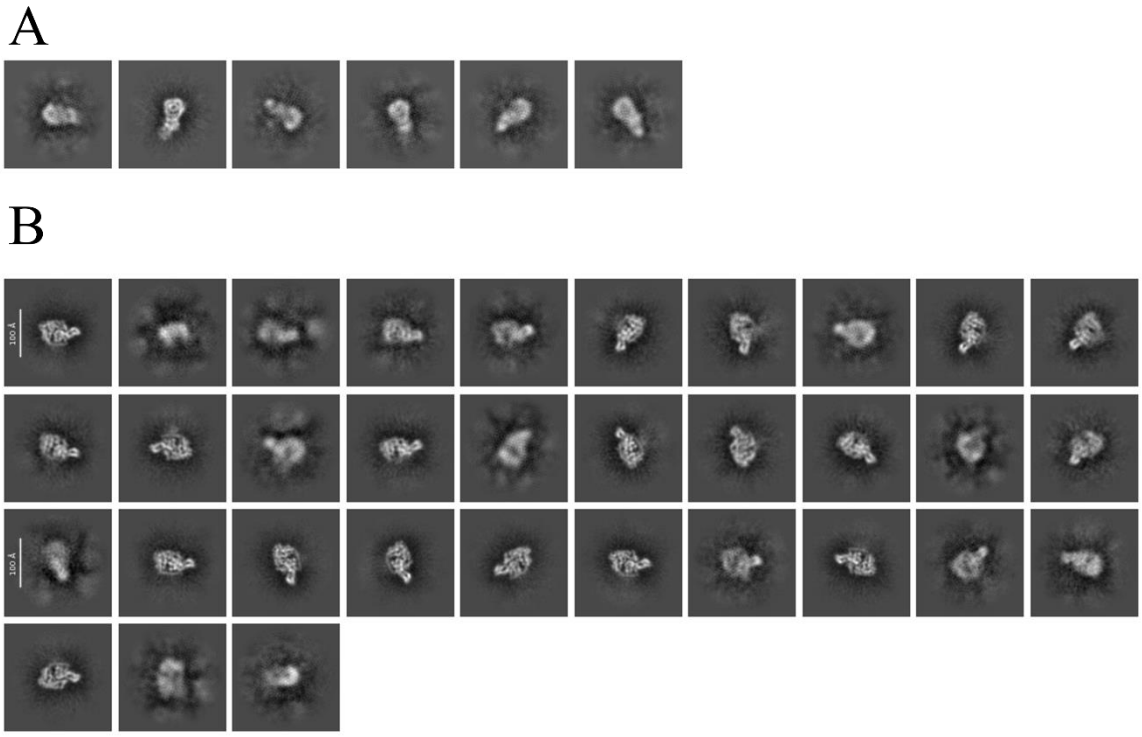

Figure 4.1: 2D classes of MTR with 1 mM MeTHF, 50 µM SAM, 50 µM OHCbl. A. 2D classes of particles representing the 7,786 particles of the N-half of MTR. **B**. 2D classes of 22,962 particles of the C-half of MTR.

To better understand the overall conformation(s) of the protein captured in this sample, one class from each half was selected for further analysis. The classes were chosen based on the clarity of protein details, absence of significant noise, and highest number of particles (N-half with 2,021 particles and C-half with 1,562 particles).

The AlphaFold predicted model of full-length human MTR was used to interpret the conformation observed in the 2D classes of MTR with ligands. This model shows the N-half as a rigid body, a finding supported by the previous crystal structure of the equivalent half (PDB: 4CCZ). The C-half is in the activation conformation, where the CblBD and AD are in close contact, a configuration suitable for the methyl transfer from SAM to cob[II]alamin after its reduction by MTRR (Bandarian, Pattridge et al. 2002) (State 'D' in figure 1.6).

When comparing the AlphaFold model with the 2D class of the N-half MTR (Figure 4.2), the overall shape of the particle from the cryo-EM dataset closely resembles the conformation of the N-half predicted by AlphaFold. This similarity includes the round shape of HcyBD and slightly elongated FolBD.

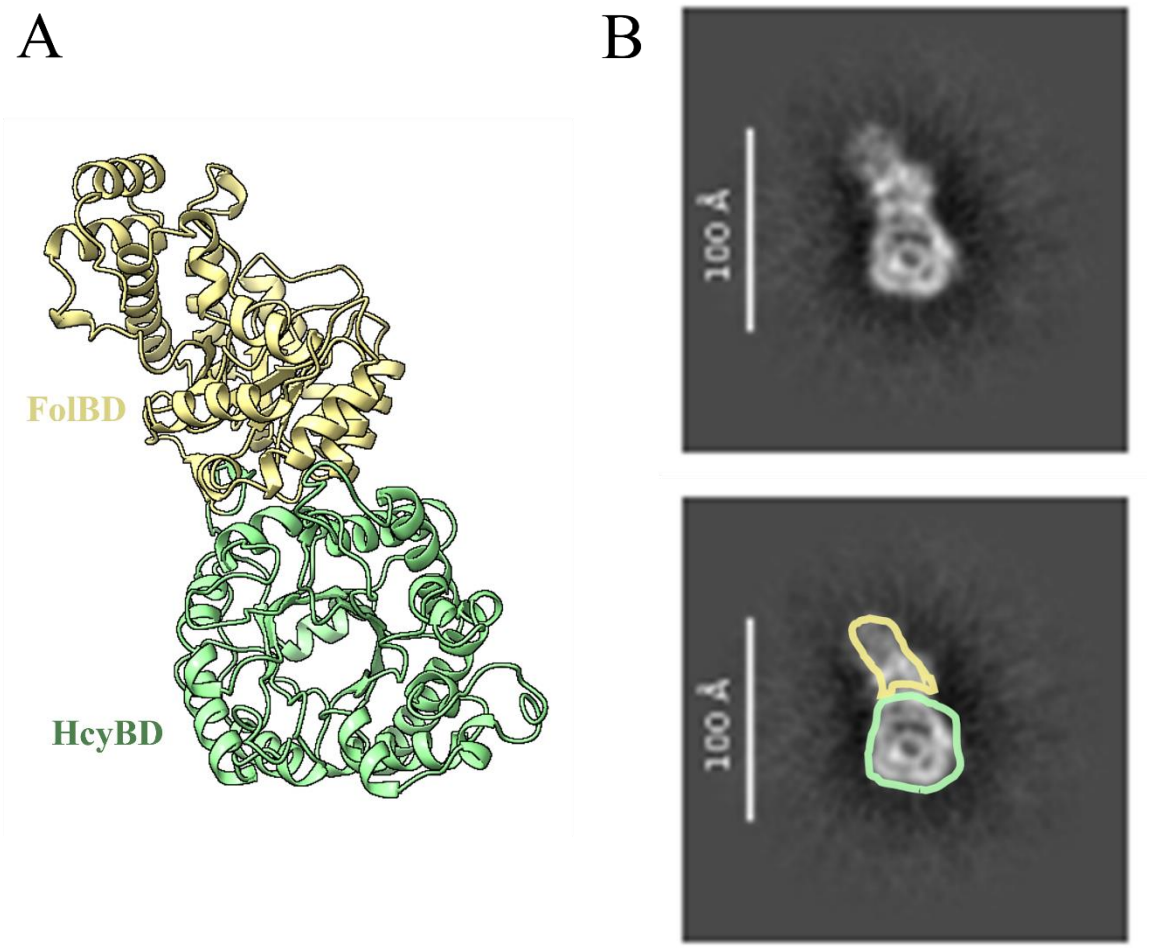

Figure 4.2: Comparison of the AlphaFold model of the HcyBD (green) and FolBD (yellow) of human MTR with the 2D class. A. AF model of the N-half of MTR. **B**. 2D class of the N-half from the MTR+Ligands dataset, with the domains marked with their respective colours.

The 2D class of the C-half of MTR was also compared with the AlphaFold model (Figure 4.3) showing good agreement as the AD and CblBD are in close contact under this activation conformation. This conformation has been captured multiple times in *E. coli* MTR crystal structures (Bandarian, Pattridge et al. 2002, Datta, Koutmos et al. 2008, Koutmos,

Datta et al. 2009) and is also the most prevalent conformation in the AlphaFold database for MTR from various organisms.

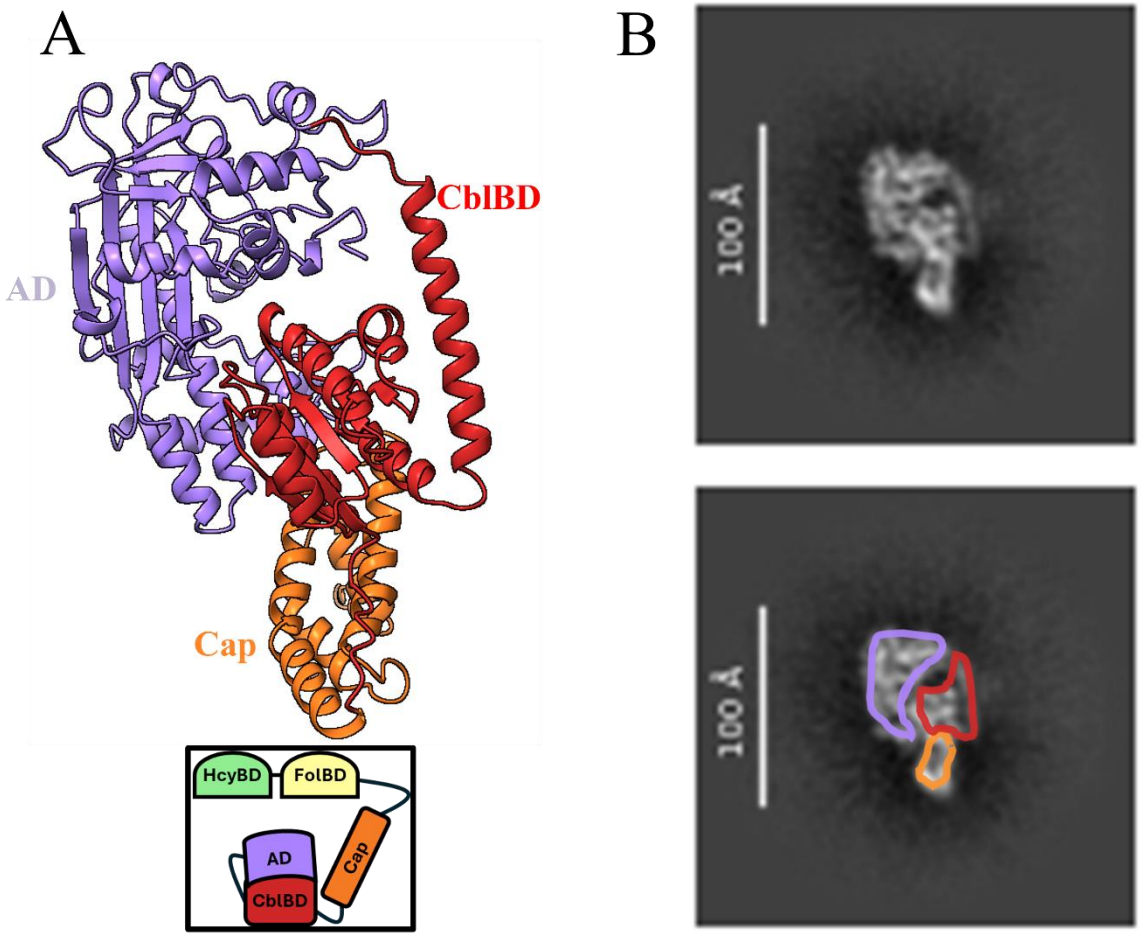

Figure 4.3: Comparison of the AlphaFold model of the Cap (orange), CblBD (red), and AD (purple) of human MTR with the 2D class of C-half. A. AF model of the C-half of MTR, and a cartoon representation of the activation conformation. **B**. 2D class of the C-half from the MTR+ligands grid, with the domains marked with their respective colours.

The 2D classes of both halves display a conformation similar to the AlphaFold model, indicating that human MTR adopts the activation conformation when the ligands MeTHF, SAM, and OHCbl are added to the sample. All experimental structures of the C-half of MTR and the full-length protein determined to date exhibit this domain arrangement.

Analysing this small dataset of MTR+ligands, it was not possible to identify particles containing both halves of MTR as an intact full-length protein. However, mass spectrometry

analysis of the recombinant protein (Figure 3.6) indicated that the purified protein remained intact up to the point of grid preparation. Additionally, negative stain EM data (Figure 3.7) suggests that the protein had not undergone proteolysis, at least on non-vitrified grids. The data therefore indicate that the highly flexible 17-residue linker between the FolBD and Cap (aa Gln653-Pro670), which connects the two halves of the protein, is responsible for the absence of particles containing all five domains of MTR. The various spatial positions that the protein halves can adopt on the grid make it impossible to generate 2D classes containing the entire MTR.

To further confirm that the two particle sets represent the N- and C-halves of MTR in the activation conformation, *ab initio* structures were generated to create volumes representing these particles without any prior bias. However, the number of particles and different views for each half of the protein was insufficient to create high-resolution structures, therefore only the overall shape of the structures could be obtained.

The *ab initio* volumes for the N and C-halves were then subjected to non-uniform refinement followed by local refinement. The refined volumes obtained for the N-half (Figure 4.4B) and the C-half (Figure 4.5B) adopt the same overall shape and dimensions of the AF predicted model (Figure 4.4A and 4.5A, respectively). The resolution of the N-half volume was calculated to be 8.77 Å, and 5.57 Å for the C-half, both by the Golden Standard Fourier Shell Correlation (GSFSC) method (Figure 4.4C and 4.5C, respectively).

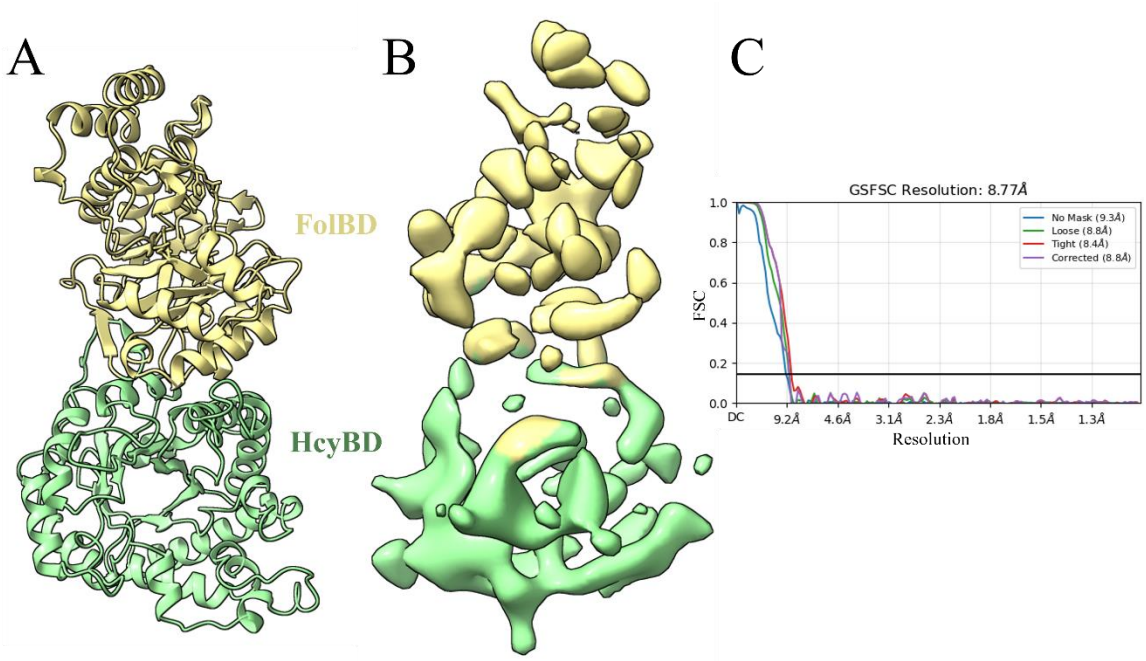

Figure 4.4: Comparison of the AlphaFold model of the N-half of human MTR with the 8.77 Å volume generated. A. AlphaFold model of HcyBD (green) and FolBD (yellow) of human MTR. **B**. Map obtained from MTR+Ligands cryo-EM dataset, displaying the volume with regions coloured according to the AF model. **C**. GSFSC curves for the unmasked, loose, tight and corrected volumes for the model in B.

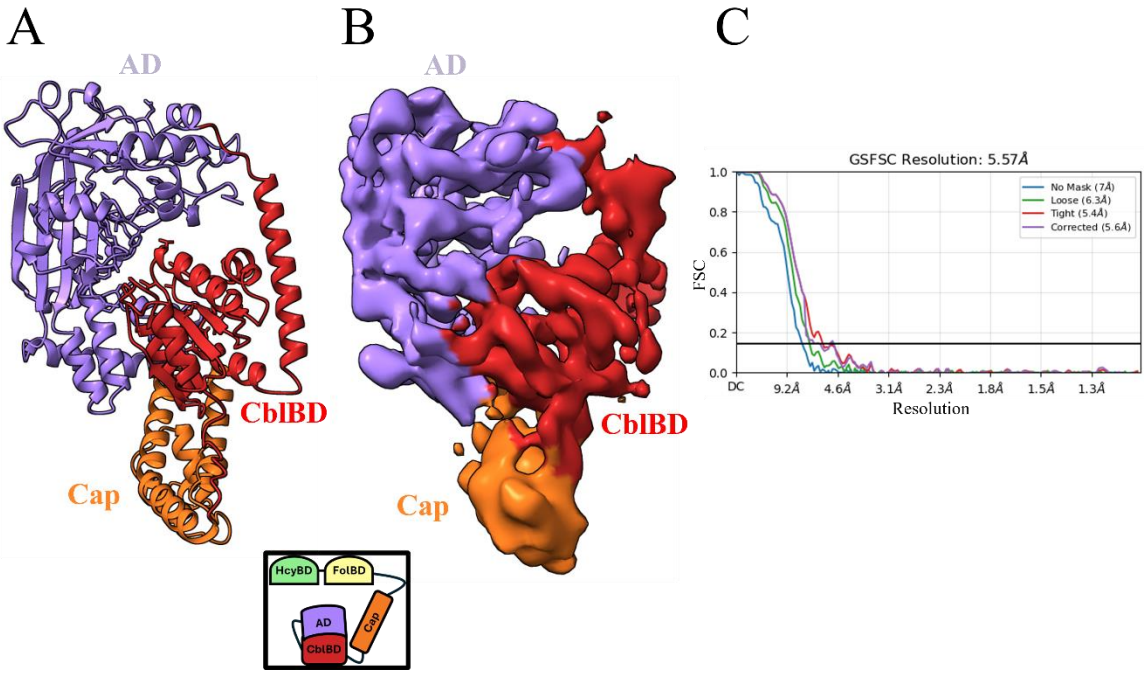

Figure 4.5: Comparison of the AlphaFold model of the C-half of human MTR with the 5.57 Å volume generated. A. AlphaFold model of Cap (orange), CblBD (red), and AD (purple) of human MTR. **B**. Map obtained from the MTR+ligands cryo-EM data displaying the volume with regions coloured according to the AF model. **C**. GSFSC curves for the unmasked, loose, tight and corrected volumes. The schematic cartoon in the inset illustrates the activation conformation of MTR.

The structures shown in Figures 4.4B and 4.5B were obtained from a relatively small number of particles (7,757 for the N-half and 22,962 for the C-half). Hence the maps did not exhibit any regions with high resolution, and only the overall shape of the halves could be confirmed through these maps. Consequently, a larger dataset collection is necessary at this point, as it could provide more particles for the determination of higher-resolution and better-quality maps, as demonstrated in later Sections 4.4 and 4.5.

4.2.2 Sample preparation and data collection for Apo MTR

The dataset of *apo* MTR was analysed similarly to the previous dataset of MTR+ligands. After using the blob-picker tool with a diameter of 80-150 Å, over 367,000 particles were selected. Several rounds of 2D classification were performed to clean up the particles. Most of the particles (approximately 83%) obtained during this process were junk particles, and only a small number (around 17%) displayed protein features. After three rounds of 2D classification, the remaining ~64,000 particles were used to create four *ab initio* volumes for further 3D classification. One of these volumes corresponded to the C-half in the activation state, as observed in the MTR+Ligands sample (Section 4.2.1). The number of particles obtained at this stage was 19,462.

No classes corresponding to the N-half were observed during the 2D classification, and none of the *ab initio* volumes corresponded to it. In the previous MTR+ligands dataset, there were nearly three times more C-half particles than N-half particles (23,000 vs 7,700). This suggests that the N-half views of MTR are less common than those of the C-half. In the *apo* MTR dataset, the N-half particles may be difficult to detect due to their even lower abundance in the micrographs, and the addition of ligands in the previous dataset might have increased the occurrence of N-half views compared to the *apo* sample. Given that structural and biophysical studies of the N-half of MTR have shown it behaves as a rigid body (Evans, Huddler et al. 2004), it is reasonable to infer that the absence of N-half particles in the *apo*

dataset is likely due to limited particle availability, and not any conformational changes. A larger dataset could potentially reveal more of the N-half particles.

Although no N-half particles were observed, 11 2D classes corresponding to the C-half of MTR were identified (Figure 4.6). These classes exhibited a similar appearance to those previously found for MTR+ligands (Figure 4.1B).

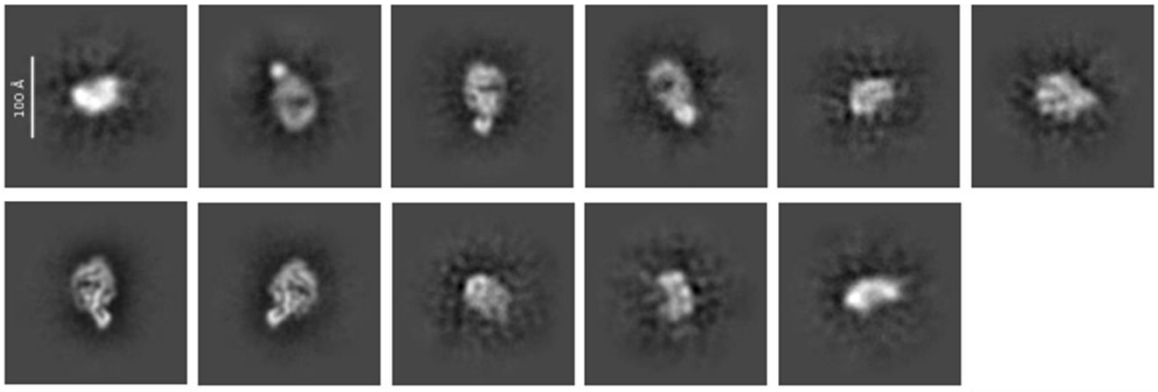

Figure 4.6: 2D classes of *apo* MTR. 2D classes of particles representing the 19,462 particles of the C-half MTR.

The class containing the majority of particles was analysed alongside the C-half of the AlphaFold model (Figure 4.7), revealing that it adopts the activation conformation, with the CblBD and AD in close interaction.

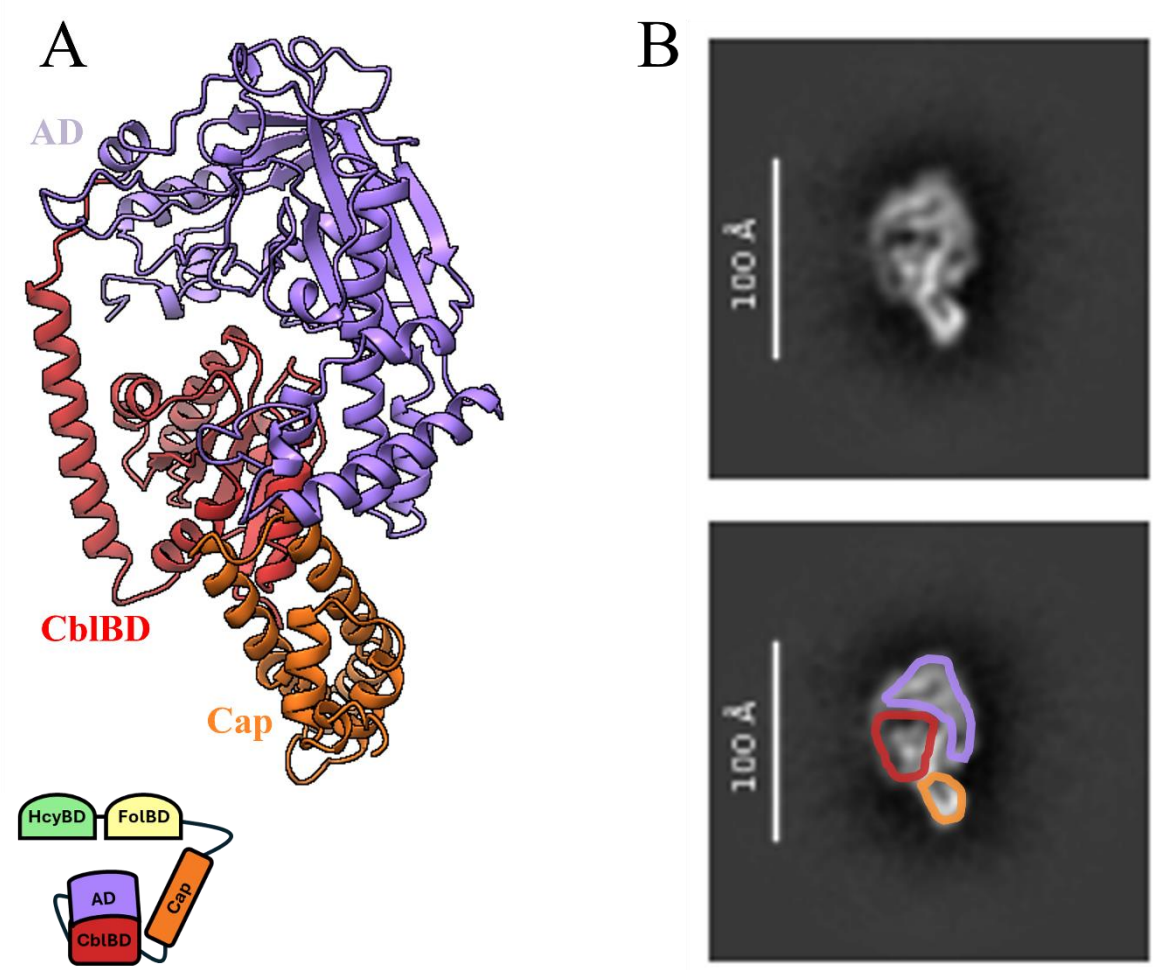

Figure 4.7: Comparison of the AlphaFold model of the Cap (orange), CblBD (red), and AD (purple) of human MTR with the 2D class of C-half from *apo* MTR dataset. A. AF model of the C-half of MTR and a schematic cartoon representation of the activation conformation of MTR. **B**. 2D class of the C-half from the *apo* MTR grid, with the domains marked with their respective colours.

The two datasets collected and analysed in this section demonstrated that human MTR behaves as two distinct halves, each approximately 70 kDa, and adopts the activation conformation on either the *apo* state or when incubated with MeTHF, OHCbl, and SAM. The next step to obtain a good density map of the full-length human MTR is to improve the quality of the grids before proceeding with large-scale dataset collection. This improvement can be achieved by changing the grid type, using freshly purified MTR, or altering the ligand combination used.

4.3 Grid and sample optimization: second collection

The second data collection trip at YSBL aimed to test a new grid type for MTR, hoping to increase the number of particles and obtain more varied views of the protein, as well as to analyse two new ligand combinations. Given the particle size, achieving a higher-resolution map for each of the two MTR halves could be challenging and may require larger datasets. Therefore, it is crucial to perform the collection on an optimised sample.

For testing the new grid type, 0.5 mg/mL of MTR was incubated with 1 mM MeTHF, 50 µM SAM, and 50 µM OHCbl (the same tri-ligand combination used in the first data collection) in 250 mM NaCl and 50 mM HEPES for 30 minutes at 37 °C. The same grid test (QuantiFoil vs UltrAuFoil) was performed for *apo* MTR, using 0.5 mg/mL of protein in the same buffer. Each sample was loaded onto two grids using the Vitrobot, following the same protocol described in Section 4.2. The first grid was Quantifoil™ Au 300 mesh 1.2/1.3, as used previously. The second grid was UltrAuFoil™ 1.2/1.3 300 Gold Mesh, entirely made from gold, which is reported to provide more stable support and facilitating the analysis of small molecules (Russo and Passmore 2016).

For the two new ligand combinations explored in this trip, one involved incubating 0.5 mg/mL MTR for 30 minutes at 37°C with 1 mM MeTHF. The rationale was to determine whether the addition of only MeTHF (without SAM and OHCbl) would significantly alter MTR's conformation, given that this ligand caused the most considerable shift in melting temperature (3.3 °C) in the DSF assay (Section 3.3.1). The sample was prepared following the standard protocol using UltrAuFoil™ 1.2/1.3 300 Gold Mesh grids.

The second new ligand combination involved incubating 0.5 mg/mL of MTR for 30 minutes at 37°C with 50 µM of MetCbl. Previously, the Cap-on state of *E. coli* MTR was obtained by adding MetCbl to MTR (Drennan, Huang et al. 1994), so MetCbl was applied

here to test whether human MTR would undergo similar conformational changes. It is known that *in vivo*, Cbl is delivered to human MTR by the MMACHC-MMADHC complex (Froese, Kopec et al. 2015, Mascarenhas, Guha et al. 2023), but *apo* MTR can be loaded with Cbl *in vitro* without the aid of partner proteins, as demonstrated in the activity assay performed in Section 3.4.

The workflow for these grids involved small collections of 500 micrographs each at the YSBL, using the Glacios with the same parameters as in the first collection in Section 4.2. After analysing the data, the most promising grids were selected for a more extensive dataset collection at the Electron Bio-Imaging Centre (eBIC) at Diamond Light Source.

During grid alignment prior to data collection, the grid containing MTR+MetCbl was not properly aligned, making it impossible to collect any micrographs. Nevertheless, The grids containing *apo* MTR and MTR+Ligands on Quantifoil or UltrAuFoil, and MTR+MeTHF on UltrAuFoil were screened, and 500 micrographs were collected from each.

Although the *apo* MTR dataset analysed in Section 4.2.2 yielded some 2D classes of the C-half of the protein, this time, the data analysis did not produce any 2D classes from either grid containing this sample. The poor quality of the grids on this occasion may be attributed to grids preparation, where certain conditions, such as ice thickness, cannot be fully controlled, resulting in micrographs where no particles are visible. The datasets containing MTR+Ligands on Quantifoil and UltrAufoil grids, and MTR+MeTHF on UltrAuFoil grids, were analysed, and the results are presented in Sections 4.3.1 and 4.3.2, respectively.

4.3.1 MTR+Ligands: Quantifoil vs UltrAuFoil grids

Five hundred micrographs of MTR+Ligands on Quantifoil and UltrAuFoil grids were collected to compare and determine the optimal grid type for a larger dataset collection. The same ligand combination used in Section 4.2.1 (1 mM MeTHF, 50 µM SAM, and 50 µM OHCbl) was employed, as the previous analysis demonstrated that high-quality 2D classes could be obtained even from a small-scale collection. Additionally, since this ligand combination had been previously tested, using it to assess the most suitable grid type for MTR was considered a sound strategy.

Data from both grid types of MTR+ligands were analysed using the same workflow as described in previous Section 4.2. For data from the Quantifoil grid, approximately 9,000 particles corresponding to the N-half and 10,000 corresponding to the C-half were identified. The classes of each MTR half are presented in Figure 4.8.

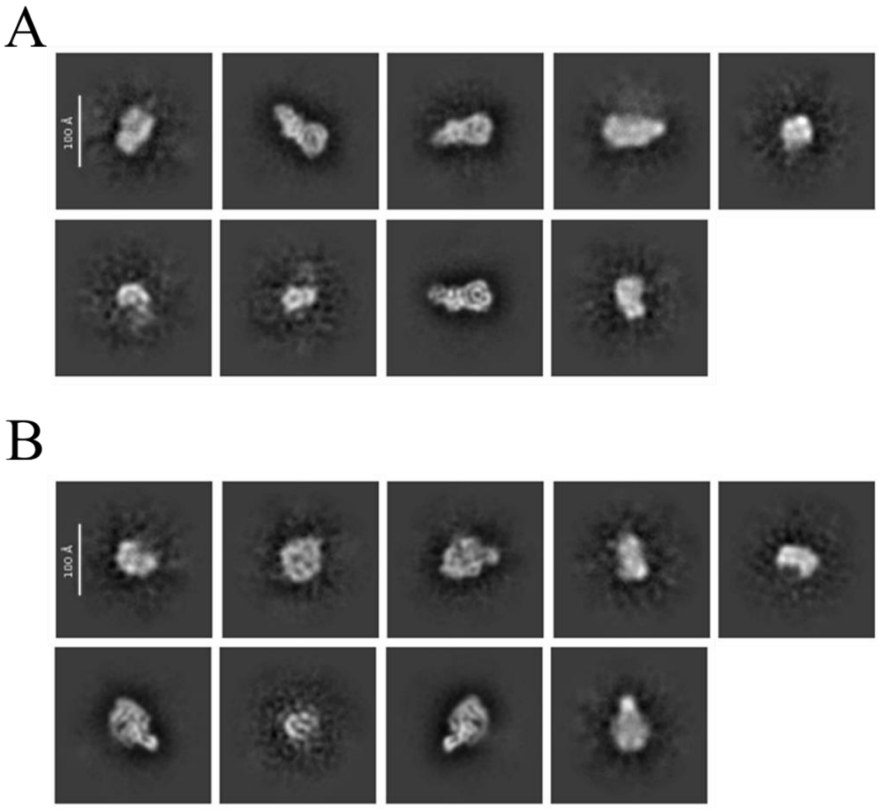

Figure 4.8: 2D classes of MTR+Ligands (with 1 mM MeTHF, 50µM SAM, and 50µM OHCbl) on Quantifoil grid. A. 2D classes of particles representing the 9,383 particles of the N-half of MTR. **B**. 2D classes of 10,171 particles of the C-half.

Although only 500 micrographs were obtained for this MTR+Ligands on Quantifoil, the number of N-half particles identified was higher than in the previous collection of 700 micrographs (~7,000 previously and ~9,000 in the current collection). The 2D classes revealed the same views as those observed before in Section 4.2, where the HcyBD and FolBD of the N-half are clearly identifiable by their overall shape (Figure 4.8A). However, the number of particles was still low with limited views, suggesting that a volume reconstruction would not achieve a resolution greater than 5 Å; therefore, the dataset was not analysed beyond 2D class identification.

In contrast, approximately 10,000 particles corresponding to the C-half were identified, which is less than half of the number found in the previous dataset (~22,900 particles), despite the current dataset containing 200 more micrographs. This discrepancy could be attributed to protein quality, as a different batch of recombinant protein was used, or to variations in ice thickness, potentially caused by human error during grid preparation (Klebl, Kay et al. 2022). Nevertheless, clear 2D classes of the C-half were obtained, confirming that the three domains are positioned in the activation conformation, with the AD and CblBD interacting, and the Cap positioned away from the CblBD. However, the low number of particles and limited views in the C-half again suggest that further processing would not yield a structure with high resolution.

The data of MTR+Ligands on the UltrAuFoil grid revealed approximately 11,493 particles for the N-half and around 24,700 for the C-half. The classes and particles were obtained in the same manner as those on the Quantifoil grids (Figure 4.8) and are shown in Figure 4.9.

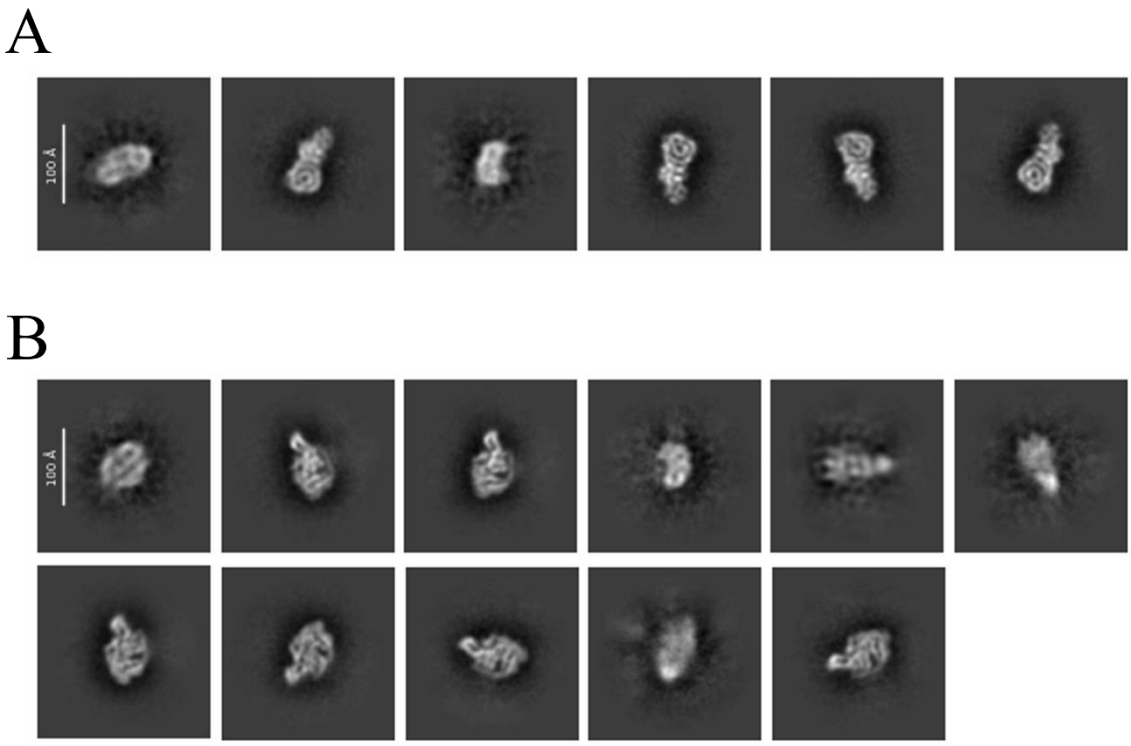

Figure 4.9: 2D classes of MTR+Ligands (with 1 mM MeTHF, 50µM SAM, and 50µM OHCbl) in UltrAuFoil grid. A. 2D classes of particles representing the particles of the 11,943 N-half MTR. **B**. 2D classes of 24,737 particles of the C-half.

Consistent with previous trend, the number of N-half particles was lower than that of the C-half in this dataset. Moreover, comparing the particles obtained from the two grid types revealed a significant increase in the number of MTR particles, particularly for the C-half, on the UltrAuFoil grid (from 10,000 to 24,000).

The screening of MTR+Ligands on two different grid types, prepared using the same sample batch, helped identify the optimal grid type that yields a greater number of particles. This is crucial for generating a higher-resolution 3D volume, particularly for the N-half of MTR, which is less abundant in the particle population compared to the C-half. Following this analysis, the UltrAuFoil grid with MTR+Ligands was preserved and shipped to the Electron Bio-Imaging Centre (eBIC) at Diamond Light Source for data collection on a Krios, with the aim of generating a high-resolution structure of both halves of the full-length human MTR.

4.3.2 New ligand combination: MTR+MeTHF

The new ligand combination to be tested, MTR+MeTHF, was applied on an UltrAuFoil Gold Mesh grid and vitrified following the standard protocol. 900 micrographs were collected and analysed to identify potential particles representing MTR.

After motion and CTF corrections, the particles were extracted using the blob-picker tool with a particle size range of 80-150 Å. The particles were then filtered through several rounds of 2D classification. During the first 2D classification, which contained more than 727,000 particles, it was already possible to identify classes corresponding to the N-half and C-half in the same orientation as previously observed. This strongly indicated that adding MeTHF to *apo* MTR does not alter the protein's conformation from the one observed in MTR+Ligands and MTR (Section 4.2) but suggests that it aids to the number of particles identified in the grid.

After several rounds of 2D classification, all particles were used to generate four *ab initio* volumes: one volume for each half of MTR and two volumes of junk particles that could not be classified as either half of the protein. These volumes and particles were then used for further 3D classification using heterogeneous refinement, and the particles corresponding to each half of MTR were again subjected to 2D classification. Demonstrative classes of each half are shown in Figure 4.10. In this dataset, the trend of having more C-half views than N-half views continued, with approximately 15,800 particles representing the N-half of MTR and around 51,400 representing the C-half.

The particles corresponding to each half of MTR (Figure 4.10) were used for *ab initio* reconstruction, generating a volume for each half of the protein. The 15,833 particles of the N-half produced a volume with a resolution of 3.79 Å. To better visualise the N-half, the final volume map was aligned with the AF-predicted model (Figure 4.11). The volume exhibits the overall shape expected for the N-half, with even some secondary structures visible in both the HcyBD and FolBD at this resolution.

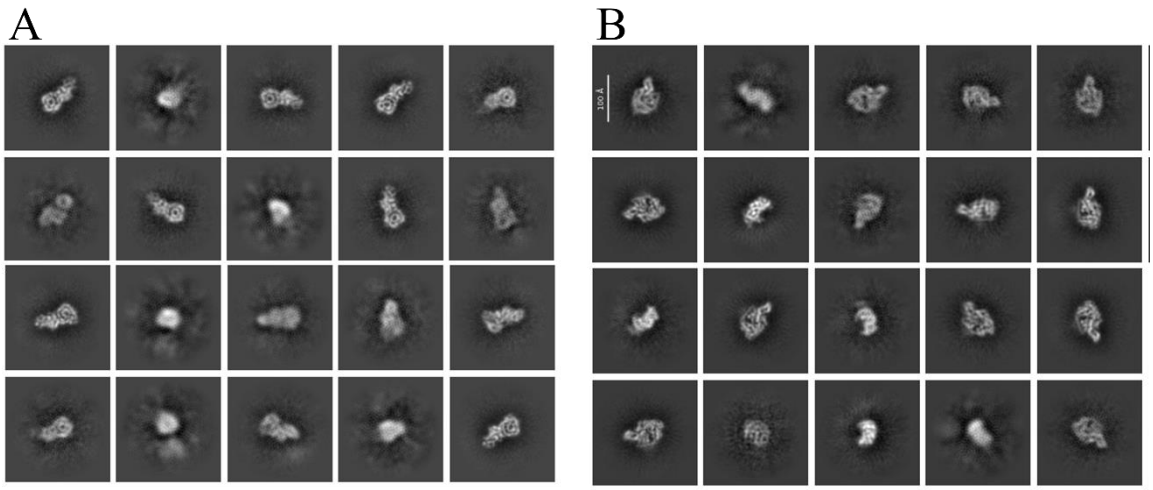

Figure 4.10: 2D classes of MTR with 1 mM MeTHF on UltrAuFoil grid. A. 2D classes of particles representing the 15,833 particles of the N-half of MTR. **B**. 2D classes of 51,496 particles of the C-half.

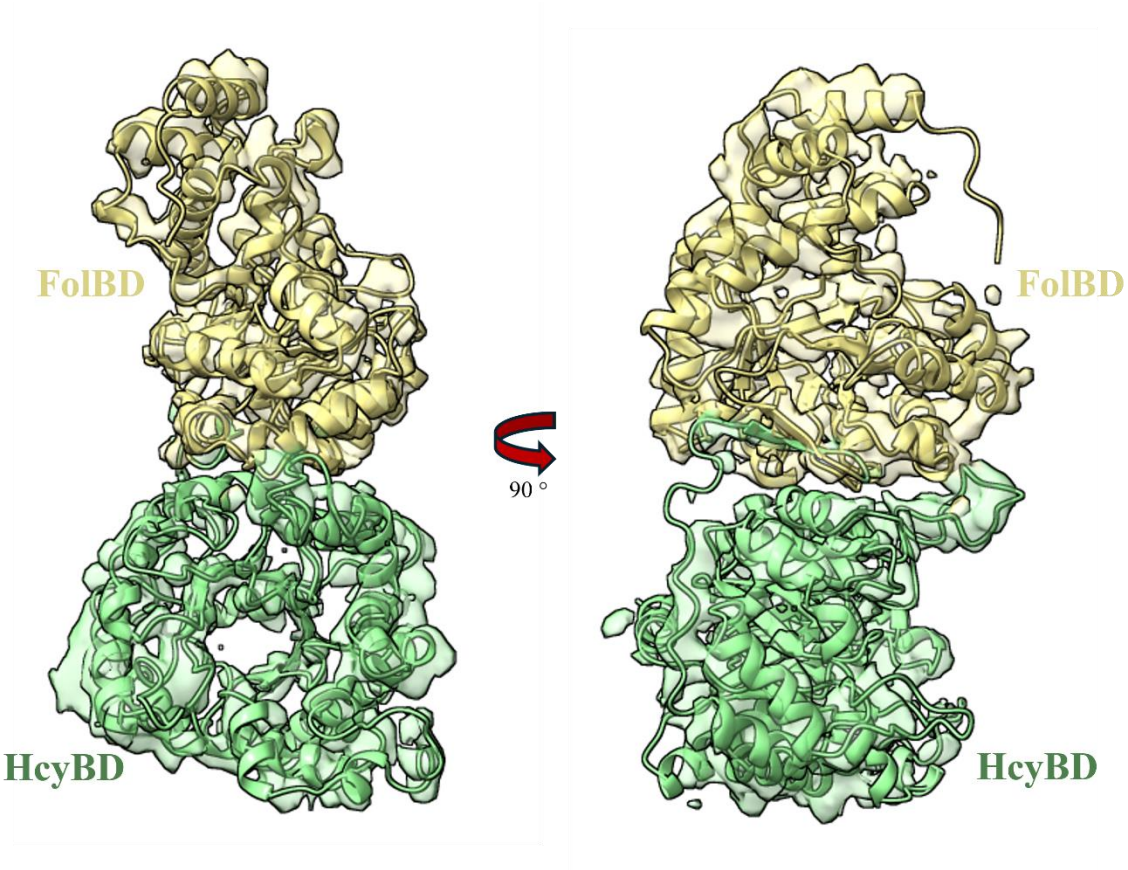

Figure 4.11: Superimposition of the 3.79 Å 3D volume with the AlphaFold model of the N-half of human MTR. Two different views of the superimposed models reveal some secondary structures in the cryo-EM volume.

The N-half volume generated from the MTR+MeTHF data shows an improvement compared to the dataset of the MTR+Ligands in the first collection (Figure 4.4). The higher number of particles and the inclusion of different views allowed for the reconstruction of a model with greater detail and reduced noise. Although some secondary structures are visible in both the HcyBD and FolBD, the map reconstruction still lacks sufficient detail. For instance, it remains challenging to confirm the presence of MeTHF in the FolBD due to poor electron density. While this data is promising, further refinement is necessary to obtain more detailed information, particularly regarding the ligand. Achieving this will require increasing the number of micrographs collected, which, in turn, will increase the number of particles

per class and provide a broader range of views, including some that may be rarer and not easily detected in a smaller dataset.

The 51,496 particles corresponding to the C-half MTR were processed, and the *ab initio* volume was further refined to a resolution of 3.49 Å. As with the N-half, the map was superimposed on the AlphaFold model for human MTR (Figure 4.12).

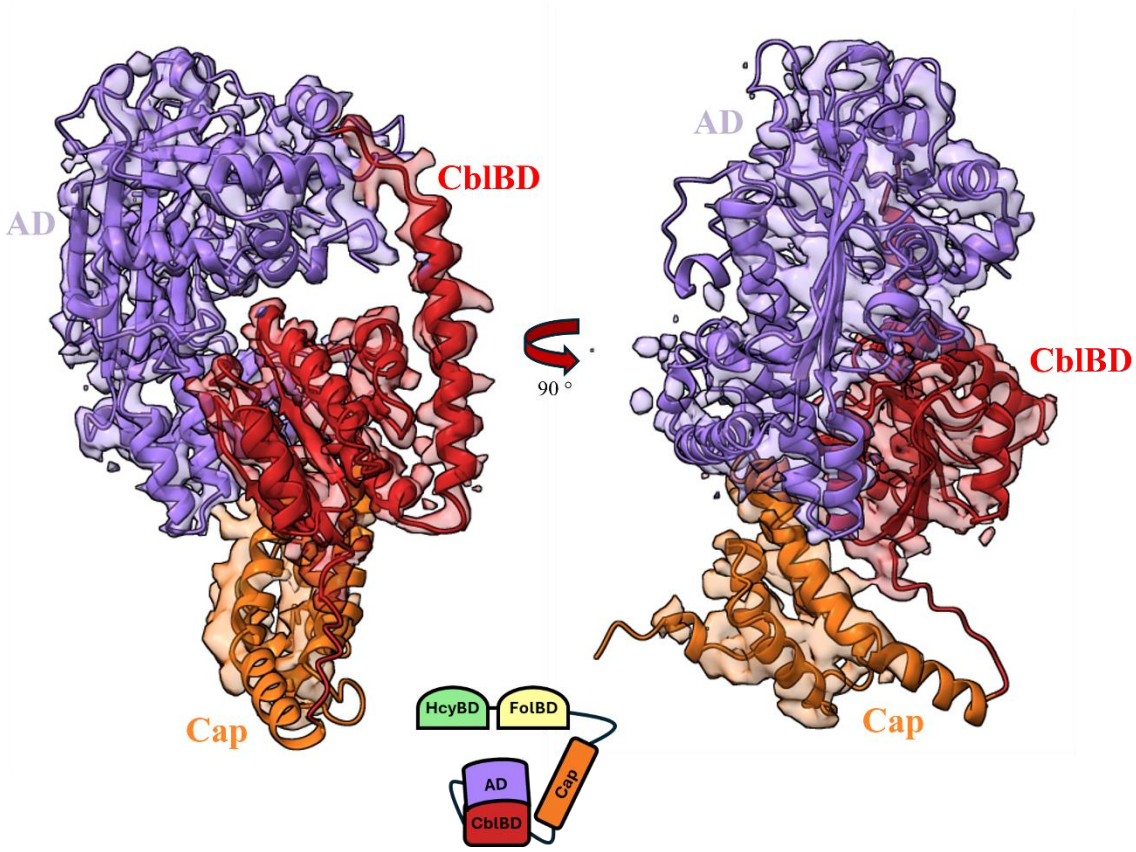

Figure 4.12: Superimposition of the 3.49 Å 3D volume with the AlphaFold model of the C-half of human MTR. Two different views of the superimposed models reveal some secondary structures in the cryo-EM volume. The schematic cartoon representation of the activation conformation of MTR is positioned between the two views.

Similar to the N-half, this map of the C-half reveals clear secondary structures, and its resolution could be further enhanced with additional particles and views. This volume also shows an improvement in resolution compared to the previous datasets from the initial screening (Section 4.2). Given its potential to generate a high-resolution volume with

additional micrographs, this grid (MTR+MeTHF) was subsequently sent to eBIC for a larger dataset collection, which will be described in Section 4.4.

4.4 Large Scale Dataset Collection to Obtain High Resolution Insights on MTR

This section reports on the large datasets collected at eBIC, using grids that were previously tested at YSBL (Table 4.1). All samples were loaded onto UltrAuFoil 1.2/1.3 300 Gold Mesh grids, identified as the best grid type from Section 4.3.

Table 4.1: Summary of the grids submitted for the third collection, at eBIC. The table includes the grid composition, the number of micrographs collected, and the outcome of the screening conducted during the previous second collection (Section 4.3).

Grid Composition	Micrographs Collected	Previous Screening Outcome
MTR+MetCbl	3,000	Grid did not align
MTR+MeTHF	5,000	Both halves of MTR in the activation conformation
MTR+Ligands (MeTHF, SAM, OHCbl)	7,000	Both halves of MTR in the activation conformation

Data were collected at eBIC using a 300 keV Krios cryogenic electron microscope equipped with a Gatan K3 detector (pixel size 0.254 Å). A total exposure dose of 65.2 $e^-/Å^2$ was used for the micrograph collection. The data were analysed using CryoSPARC, following the previous protocol with a few modifications that will be explained below.

The first grid, composed of MTR incubated for 30 minutes at 37°C with 50 μM of MetCbl (MTR+MetCbl), was prepared under the assumption that *apo* MTR is in the activation conformation, as observed during the analysis of the initial dataset (Section 4.2.2). In this conformation, the Cbl binding pocket in CblBD should be exposed to the solution, allowing for the binding of MetCbl (Mendoza, Purchal et al. 2023). Adding MetCbl could potentially induce a conformational change on MTR to the resting conformation (Cap-on),

as observed for the *E. coli* protein (Bandarian, Pattridge et al. 2002). This grid was not previously screened at YSBL due to alignment issues.

The second grid, containing MTR with 1 mM MeTHF (MTR+MeTHF), was previously screened at YSBL, and the data were analysed as described in Section 4.3. The protein is in the activation conformation, with a clear interaction between the CblBD and AD. The larger dataset collection at eBIC is expected to generate higher-resolution volumes. The key finding expected from the MTR+MeTHF grid is the C-half, which reveals MTR in the activation conformation necessary for the remethylation of oxidised Cbl (cob[II]alamin) (Bandarian, Pattridge et al. 2002) (State 'D' in figure 1.6). This conformation, where the AD and CblBD are interacting, has been demonstrated very recently for *T. thermophilus apo* MTR (Mendoza, Purchal et al. 2023), but has not yet been shown for the human enzyme.

The third grid, consisting of MTR incubated with 1 mM MeTHF, 50 μM SAM, and 50 μM OHCbl (MTR+Ligands), was also previously tested, yielding results similar to those found with MTR+MeTHF (Section 4.3.1). The N-half of MTR consistently exhibits the same HcyBD and FolBD arrangement as a single unit, but with fewer particles available compared to the C-half. In this third grid, the C-half is also in the activation conformation, with SAM and OHCbl expected to be bound to the AD and CblBD, respectively. Although there are structures of the C-half of bacterial MTR with cobalamin bound (Bandarian, Pattridge et al. 2002, Datta, Koutmos et al. 2008, Koutmos, Datta et al. 2009, Mendoza, Purchal et al. 2023), currently, only the structure of the isolated AD of the human protein is available (Wolthers, Toogood et al. 2007).

In summary, each of the selected grids for the large dataset collection offers the potential for different insights into human MTR conformation and the structure of each domain.

4.4.1 Processing of eBIC data from the MTR+MetCbl grid

The analysis workflow for this dataset is summarised in Figure 4.13. The micrographs were first subjected to motion correction and CTF correction before particle picking. The blob-picker tool was used twice, with different particle diameters. For the first pick, a diameter of 80-150 Å was selected, resulting in approximately 1.2 million particles, which were 2D classified into 200 classes (Classes A). The second pick utilised a particle diameter of 50-120 Å, identifying approximately 1.9 million particles. The second pick, with a smaller diameter, was conducted to detect smaller particles in the event of a conformational change in MTR. These particles underwent two rounds of 2D classification, with the best classes from each round being reclassified. Finally, the best classes from the 50-120 Å blob-picking were classified a third time (Classes B).

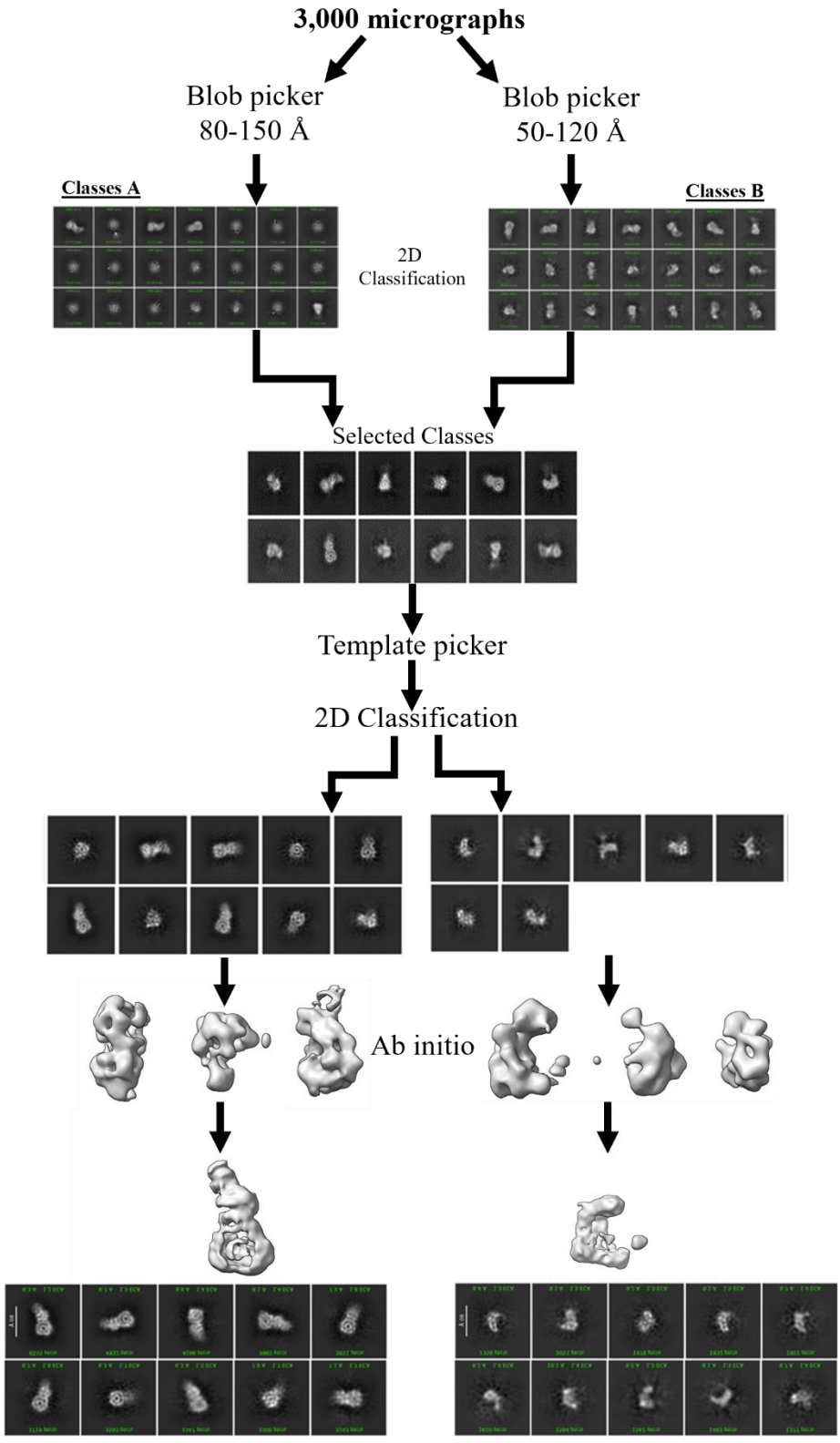

Figure 4.13: Overview of the data processing for MTR+MetCbl from the Krios dataset.

Six classes were chosen to illustrate each Class A and B, shown in Figures 4.14A and 4.14B, respectively. Together, these classes had around 55,000 particles in total in 12 classes.

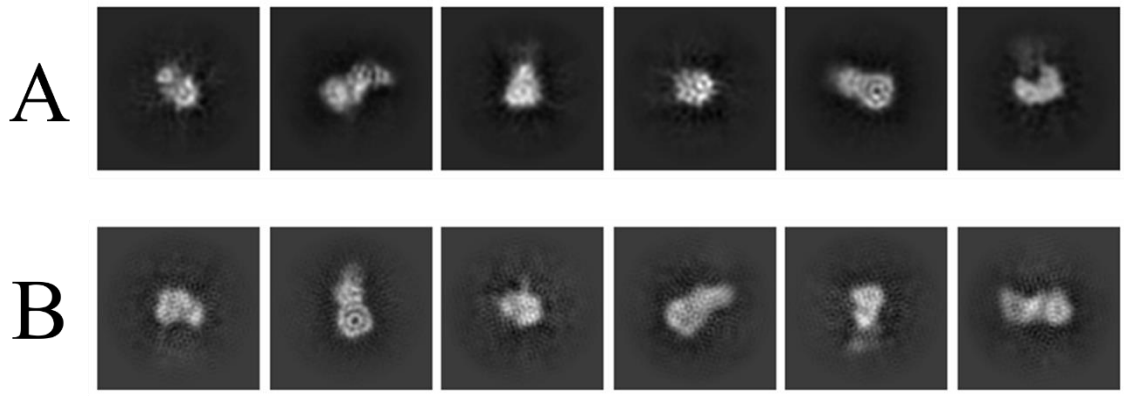

Figure 4.14: 2D classes from blob-picking particles. A. Classes selected from the blob-picking with particle sizes ranging from 80 to 150 Å. **B**. Classes selected from the blob-picking with particle sizes ranging from 50 to 120 Å.

In order to expand the number of particles found for this dataset, the classes shown in Figure 4.14 were subsequently used for a different type of particle picking, known as template picking, which can identify particles missed by the blob-picker tool. The templates used for this method can be derived either from a protein structure or from classes formed entirely by particles within the dataset. To minimise bias, only classes derived from the dataset were chosen to generate templates for use in template picking (Doerr 2016, Glaeser, Nogales et al. 2021). Similar to blob picking, template picking can result in false positives, leading to the inclusion of many undesirable "particles." The template picking can also identify particles of varying sizes to generate classes, allowing for the selection of a greater number of particles and encompassing a broader range of particle sizes. The template picker selected over 2.5 million particles, which were filtered through sequential 2D classifications. The classes became visible after three rounds of 2D classifications and could be further separated (Figure 4.15).

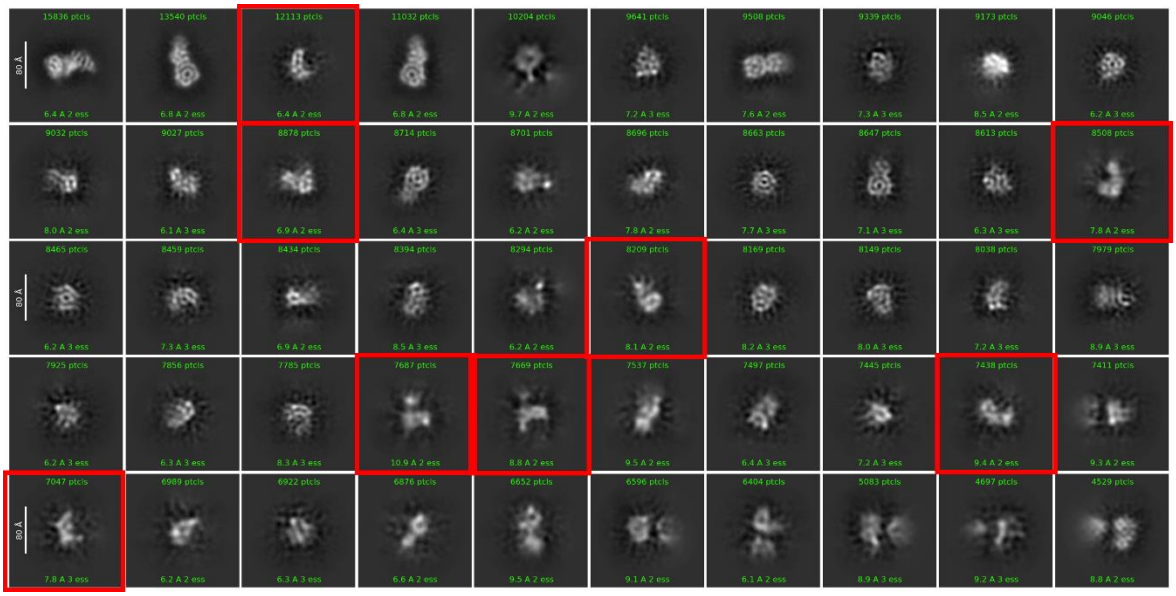

Figure 4.15: 2D classes of template-picking particles. After the particles were picked using the template-picking tool, three rounds of 2D classifications were performed, revealing known N-half classes. Novel 2D classes, that do not represent N-half and are not observed in the other datasets, are highlighted in red.

The N-half of MTR was identified, consistent with previous data. Surprisingly, no C-half particles were observed in this dataset, despite their higher prevalence compared to the N-half in all other samples. Additionally, novel 2D classes, formed by particles not seen in previous dataset analyses, were also identified and are highlighted in red in Figure 4.15.

To confirm the presence of the N-half, three *ab initio* volumes were generated using only the N-half classes displayed in Figure 4.15. The particles were then subjected to 3D classification based on these volumes, and the class with the largest number of particles, which also matched the overall shape of the N-half, was further refined using non-uniform and local refinement. The final volume of the N-half was generated using 40,880 particles, achieving a resolution of 6.30 Å. The particles used for this volume reconstruction were classified into nine classes, shown in Figure 4.16A, alongside the low-resolution volume (Figure 4.16B).

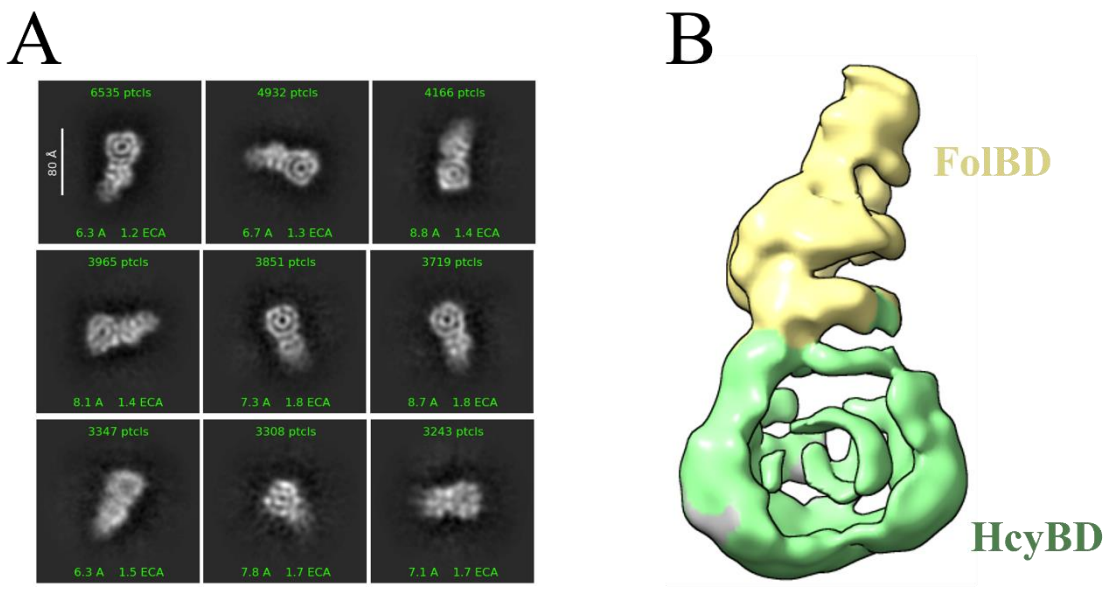

Figure 4.16: Summary of the N-half of MTR obtained from the eBIC dataset of MTR+MetCbl. A. 2D classes of the particles used to reconstruct the volume of the N-half. **B**. Low-resolution volume generated from the particles in panel A, showing the overall shape of the HcyBD (green) and FolBD (yellow).

The resolution of the N-half from this 3D reconstruction is lower than that of other datasets, particularly for a dataset containing 3,000 micrographs. The quantity and quality of selected particles likely contributed to this low resolution, possibly due to the ice thickness on the grid. The absence of classes representing the C-half of MTR could suggest a change in the protein's conformation, disrupting the interaction between the CblBD and AD and increasing the flexibility of the loop connecting these two domains.

To further extrapolate classes from Figure 4.15, those that did not correspond to the N-half and were identified in this dataset (highlighted in red) were selected to gain insights into the potential conformational changes that occurred when MTR was incubated with MetCbl. These selected classes were subjected to another round of 2D classification to further refine the particles. The classes derived from these 56,408 particles are presented in Figure 4.17.

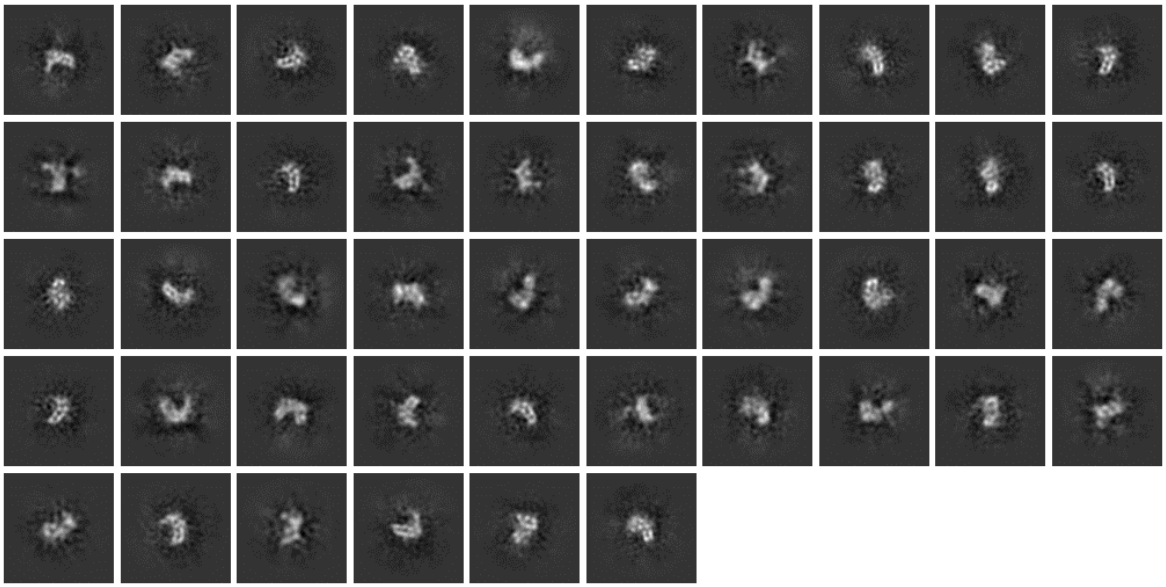

Figure 4.17: 2D classes of the 56,408 particles that did not correspond to the N-half, highlighted in red on Figure 4.15. The classes represent particles that did not fit in the classes or volume of MTR's N-half.

All particles shown in Figure 4.17 were used to construct three *ab initio* volumes, which were then used to further filter the particles and generate three refined volumes. These volumes were refined through non-uniform refinement, and one of them, comprising 17,920 particles, was found to be similar to the AD. To investigate whether these particles represent the AD, the AlphaFold model of this domain was used to generate 50 different 2D templates, representing various views of the AD as it would appear in micrographs. These classes generated from the AlphaFold model were then compared to those found in the MTR+MetCbl dataset (Figure 4.18). Some of the classes identified in the dataset were similar to those generated from the AD model, suggesting that these classes potentially correspond to the AD of MTR. Importantly, the identification of particles corresponding solely to the AD suggests that it is no longer interacting with the CblBD, as observed in all other datasets, implying a conformational change in MTR has happened due to MetCbl binding.

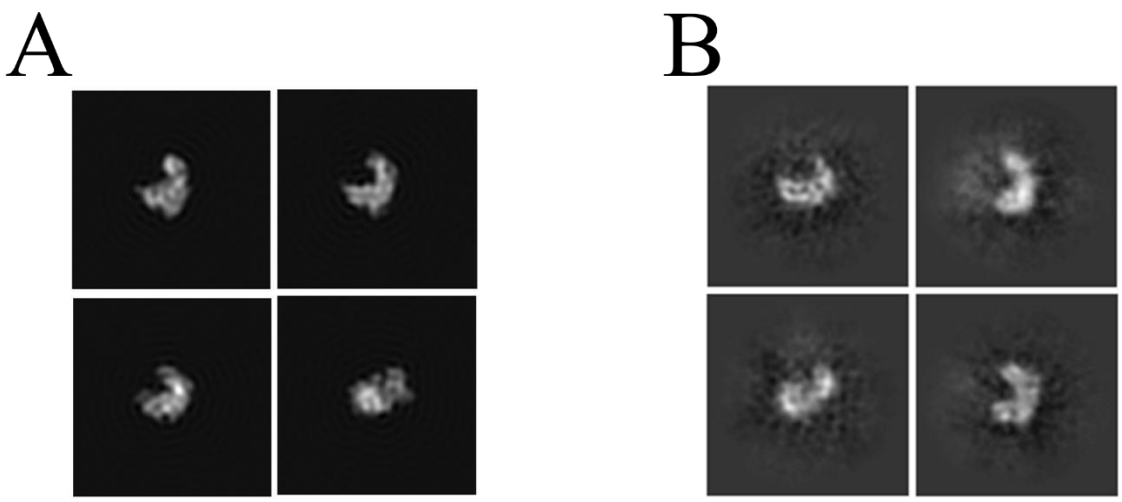

Figure 4.18: Investigating the possibility of 2D classes representing AD in MTR+MetCbl dataset. A. Examples of 2D classes generated using the AF model of the AD. **B**. Example of 2D classes of AD generated by the particles extracted from the dataset MTR+MetCbl.

The AD volume had a low resolution of 6.80 Å, with 17,920 particles. This volume and the AF-predicted model were aligned and displayed similar overall shape and size (Figure 4.19).

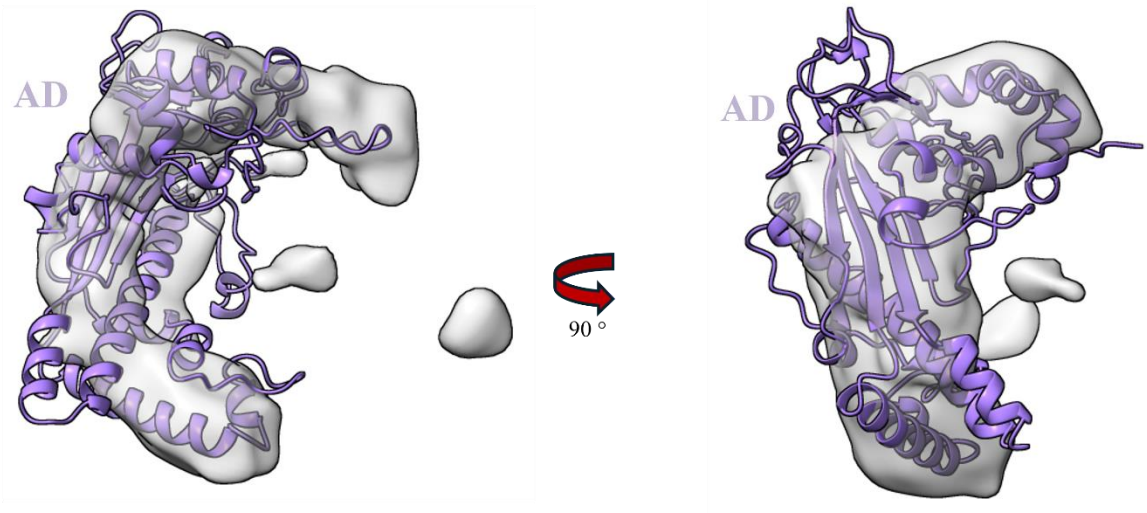

Figure 4.19: Superimposition of the 6.80 Å 3D volume of AD with the AlphaFold model. Two different views of the superimposed models reveal similar shape, confirming the presence of the isolated AD in the sample MTR + MetCbl.

In summary, the MTR+MetCbl grid resulted in 2D classes identifying two sections of MTR. The first was the N-half as a rigid body (Evans, Huddler et al. 2004), allowing the HcyBD and FolBD to be visualised together in all cryo-EM datasets collected thus far (Sections 4.2 and 4.3). The second was the AD alone, approximately 40 kDa. This is contrary to datasets displaying the activation conformation, such as *apo* MTR, in which the AD interacts with the CblBD, resulting in these domains being visualised together (Sections 4.2 and 4.3).

The observation of the isolated AD here indicates that MTR has undergone conformational changes, leaving the activation conformation upon MetCbl binding.

4.4.2 Processing of eBIC data from the MTR+MeTHF grid

Another dataset collected at eBIC was from the MTR+MeTHF grid which was previously screened (Section 4.3.2), and the analysis of the micrographs revealed MTR in the activation conformation, with particles corresponding to the N and C-half, where the AD and CblBD were in close proximity. Since the only ligand added to the sample was MeTHF, the C-half should be in the *apo* form.

The dataset of MTR+MeTHF, along with the subsequent dataset to be presented in Section 4.4.3 (MTR+Ligands—MeTHF, SAM, OHCbl), was analysed using an additional particle-picking tool in CryoSPARC, the TOPAZ particle picker (Bepler, Morin et al. 2019). This tool is particularly effective for large datasets, enabling the selection of a greater number of particles with fewer false positives, leading to a more accurate structure. In the MTR+MeTHF and MTR+Ligands (next section 4.4.3) datasets, particles were initially picked using blob-picking and template-picking, and these were then used to train the TOPAZ algorithm to identify similar particles throughout the micrographs.

Both the MTR+MeTHF and MTR+Ligands datasets were analysed using a similar approach (Figure 4.20). Both datasets had been previously analysed in smaller dataset collections (Section 4.3), so it was anticipated that MTR would be in the activation conformation and that the two halves of the enzyme would appear as separate classes.

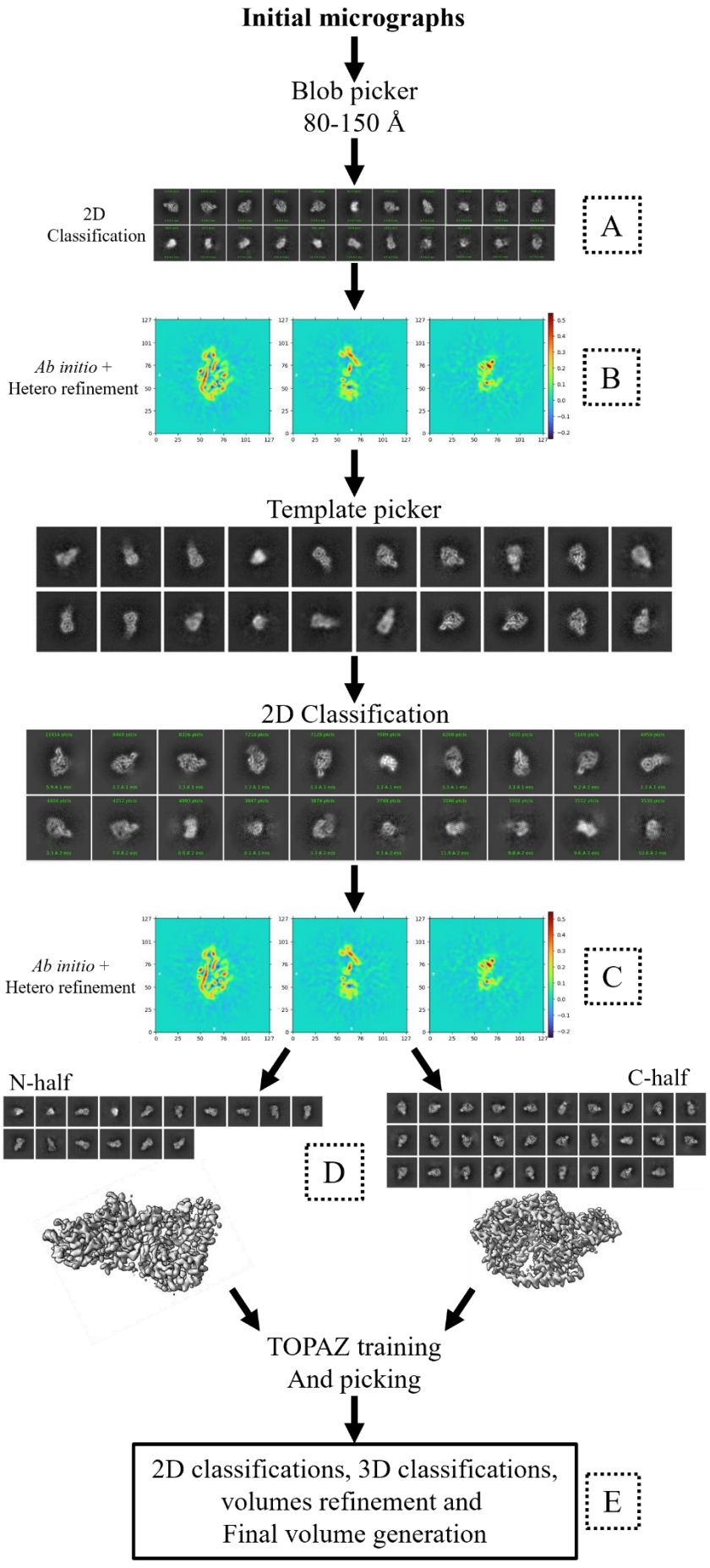
Initial micrographs
Blob picker
80-150 Å
2D
Classification
A
Ab initio +
Hetero refinement
B
Template picker
2D Classification
Ab initio +
Hetero refinement
C
N-half
C-half
D
TOPAZ training
And picking
2D classifications, 3D classifications,
volumes refinement and
Final volume generation
E

Figure 4.20: Overview of the data processing for MTR+MeTHF and MTR+Ligands from the eBIC Krios datasets. The particles were selected using a blob picker tool with a diameter of 80-150 Å, and after rounds of 2D classifications (**A**) and *ab-initio* volumes generation to find particles corresponding to MTR, the 3D-classified particles (**B**) were used to generate classes for template-picking. Those particles were then submitted to various rounds of 2D and 3D classification over *ab-initio* generated classes (**C**). After 3D-classifying the particles, the volumes for each half of MTR were obtained. Some of the particles of each half were used to train TOPAZ picking (**D**), and then it was used to pick particles on all micrographs. Finally, rounds of 2D and 3D classifications were performed (**E**), and each volume was refined to generate the final structure of each half.

For the MTR+MeTHF dataset, all 5,000 micrographs were motion-corrected and CTF-corrected. Approximately 1.6 million particles from blob picker were subjected to six rounds of 2D classification, which filtered out most of the junk particles, yielding a final set of approximately 138,000 particles. The 46 classes shown in Figure 4.21 correspond to both the N and C-half of MTR, with the majority of particles representing the C-half.

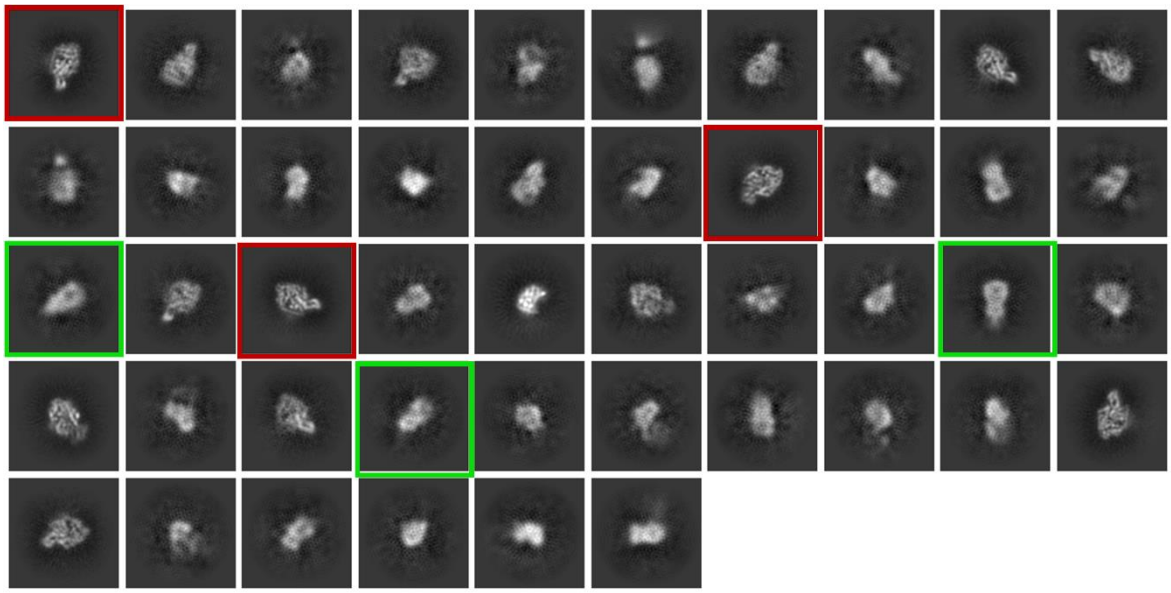

Figure 4.21: 2D classes of particles from blob picking. After six rounds of 2D classifications, the ~1.6 million particles were filtered to ~138,000, classified in 46 classes. Both N and C-half are present in the dataset. Some examples of the classes form the N-half are highlighted in green and some of the C-half are in red.

The selected classes were used to generate four *ab initio* volumes. These generated volumes were employed for 3D classification of the particles, and by performing 2D classification on the particles corresponding to each of the four volumes, it was determined that one volume represented the N-half, two represented the C-half, and the final volume

comprised junk particles. The best classes from all three volumes representing MTR were selected (Figure 4.22) and were then used for template particle picking.

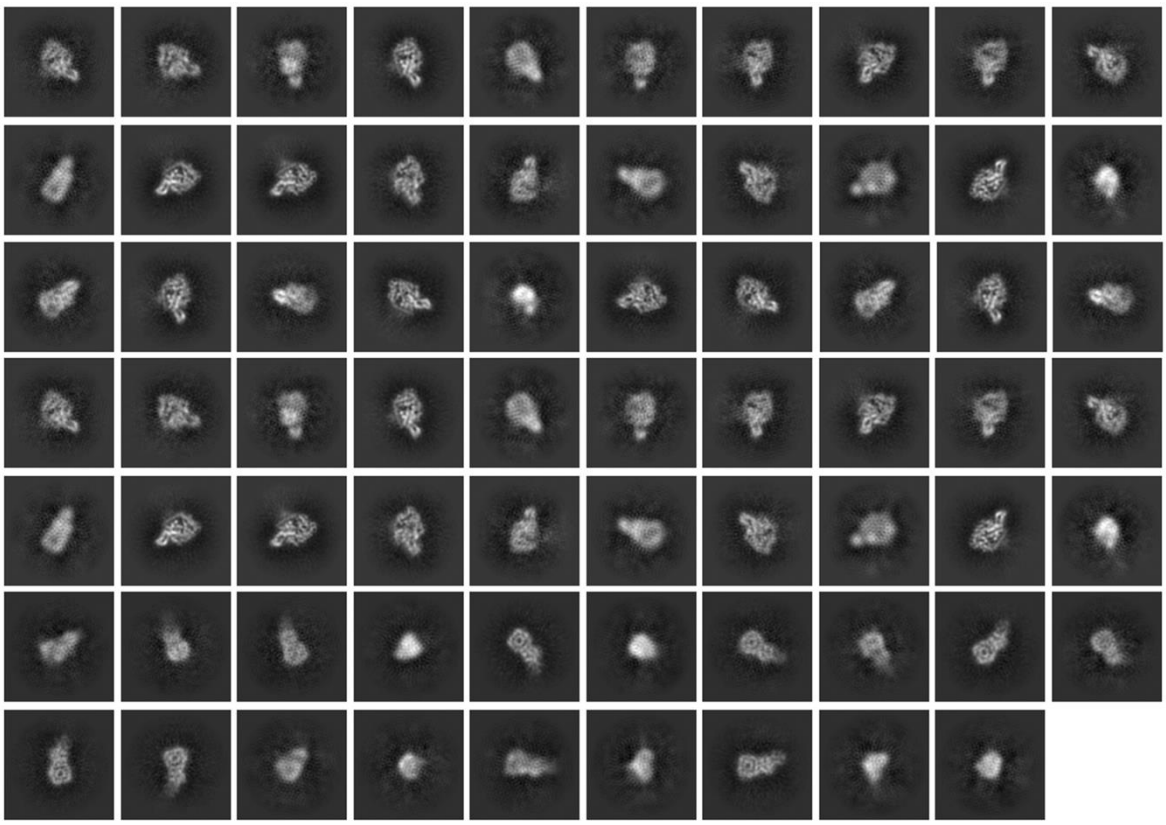

Figure 4.22: 2D classes used as template for template picking. All 69 classes were used for a new round of particle picking, using the template-picking tool.

The template picker identified approximately 4.5 million particles, which were then filtered through several rounds of 2D classification. Following this process, the particles were used to generate six *ab initio* volumes. One volume corresponded to the N-half, two corresponded to the C-half of MTR, and the remaining three were composed of junk particles. After further filtering the template-picked particle classes, each half was analysed independently.

Starting with the N-half, all 42,391 particles corresponding to its *ab initio* volume were used to generate two additional *ab initio* volumes. Following another round of 3D classification and further refinement, a structure with a resolution of 3.37 Å was obtained

from the volume composed of 30,847 particles. Figure 4.23 displays the 20 2D classes derived from the particles that contributed to this 3.37 Å volume.

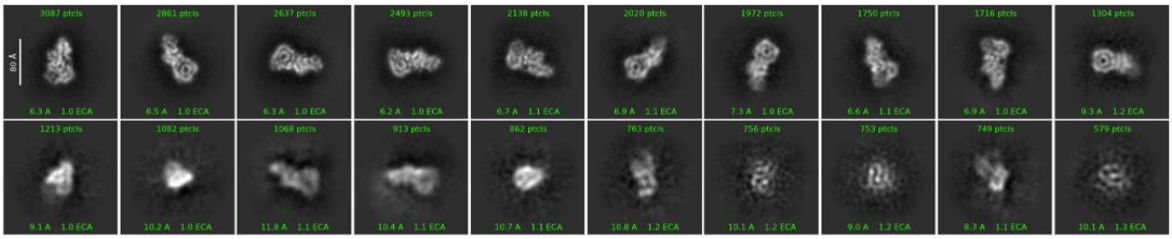

Figure 4.23: 2D classes of 30,847 particles corresponding to the N-half of MTR. The particles were used to train the TOPAZ picking tool for picking N-half particles.

To further improve the resolution of the obtained N-half volume, the 30,847 particles identified (Figure 4.23) were used to train the TOPAZ particle-picking tool, which was then employed to pick particles across all micrographs. The trained tool successfully identified approximately 213,000 particles, which were subsequently 2D-classified. Through several rounds of heterogeneous refinement, particles were filtered against the 3.37 Å volume and a previously generated junk volume. The resulting ~44,000 particles from the N-half were combined with the previous 30,847 particles, and after removing duplicates, a total of 51,788 particles were used to generate an N-half volume. This process yielded a final N-half volume of MTR+MeTHF at a resolution of 3.05 Å, revealing visible structural details of the protein (Figure 4.24).

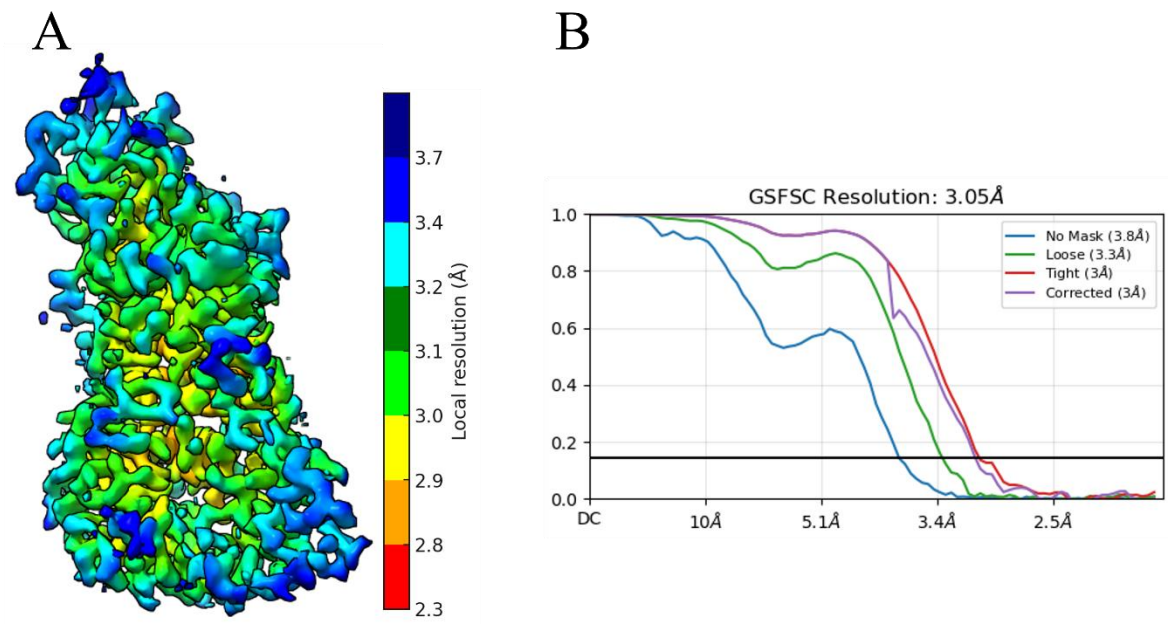

Figure 4.24: MTR N-half structure map from MTR+MeTHF grid. A. Local resolution of the N-half map of overall resolution of 3.05 Å. **B**. GSFSC curves for the unmasked, loose, tight and corrected volumes.

The particles corresponding to the C-half of this sample were next analysed following the same workflow. Initially, the two volumes corresponding to the C-half, which were *ab initio* generated from template picking, were selected, and the particles from these volumes were 2D classified and filtered by selecting the best classes. After particle filtering, these were used to generate three new *ab initio* reconstructions, followed by an additional round of 3D classification. Finally, another round of 2D classification was performed on the particles filtered into the volume corresponding to the C-half. The selected 85,512 particles were then used to build a volume that achieved a resolution of 3.18 Å. The particles used for this volume were divided into 20 classes for visualisation (Figure 4.25).

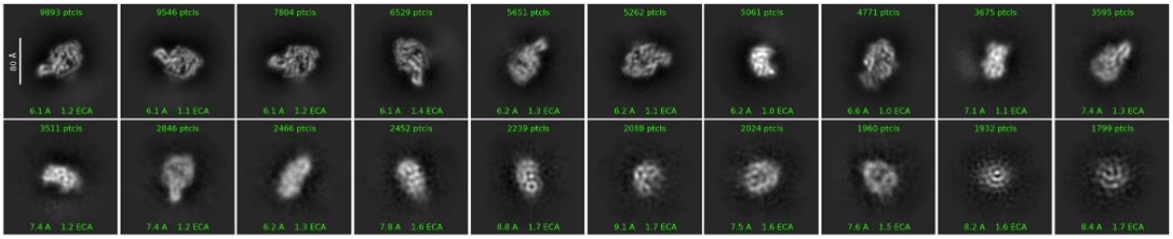

Figure 4.25: 2D classes of 85,512 particles corresponding to the C-half of MTR. The particles were used to train the TOPAZ picking tool for picking C-half particles.

These 85,512 particles were used to train the TOPAZ picking tool. After training, all the micrographs were screened for particles. Approximately 286,000 particles were picked and filtered through one round of 2D classification and seven rounds of 3D classification. Similar to the N-half analysis, the 3D classification was performed using the volume that generated the classes in Figure 4.25 and a previously generated junk volume. The 101,656 filtered particles identified by TOPAZ were then combined with the 85,512 particles from the template picking, and duplicates were removed. The resulting ~154,000 particles were used to generate two *ab initio* volumes, followed by further 3D classification. After additional refinements, a final volume corresponding to the C-half of MTR was generated at 2.82 Å resolution, using 100,343 particles.

In the 2.82 Å resolution map of the MTR C-half, there is no visible density for Cbl in the CblBD, indicating that the activation conformation can be adopted without the presence of the cofactor (Figure 4.26). Furthermore, clear secondary structures can be observed in the density map, which has a resolution high enough for detailed model building, which will be explored in the next chapter.

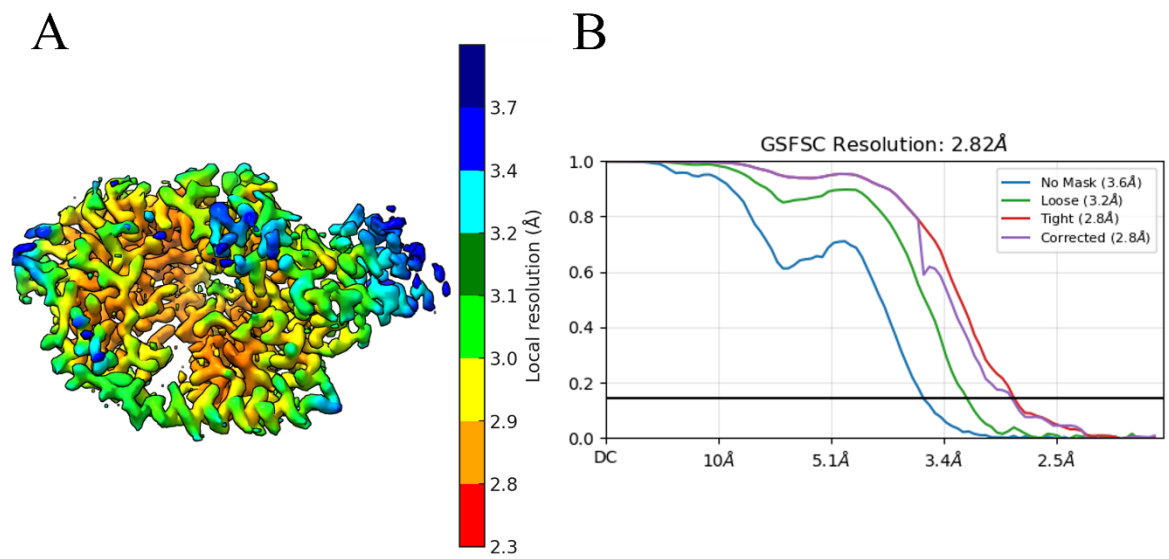

Figure 4.26: MTR C-half structure map from MTR+MeTHF grid. A. Local resolution of the C-half map of overall resolution of 2.82 Å. **B**. GSFSC curves for the unmasked, loose, tight and corrected volumes.

4.4.3 Processing of eBIC data from MTR+Ligands

In addition to MTR+MetCbl and MTR+MeTHF, a large dataset was also collected from the MTR+Ligands grid. This sample was prepared by incubating 0.5 mg/mL of MTR with 1 mM MeTHF, 50 µM SAM, and 50 µM OHCbl. The data analysis followed the previous protocol (Section 4.4.2, Figure 4.20).

After motion and CTF correction of the 7,000 collected micrographs, over 2.2 million particles were picked using the blob-picking tool with a diameter range of 80-150 Å. These particles were then filtered through multiple rounds of 2D classification, resulting in approximately 205,000 particles divided into 44 classes, representing both the N and C-halves of MTR (Figure 4.27).

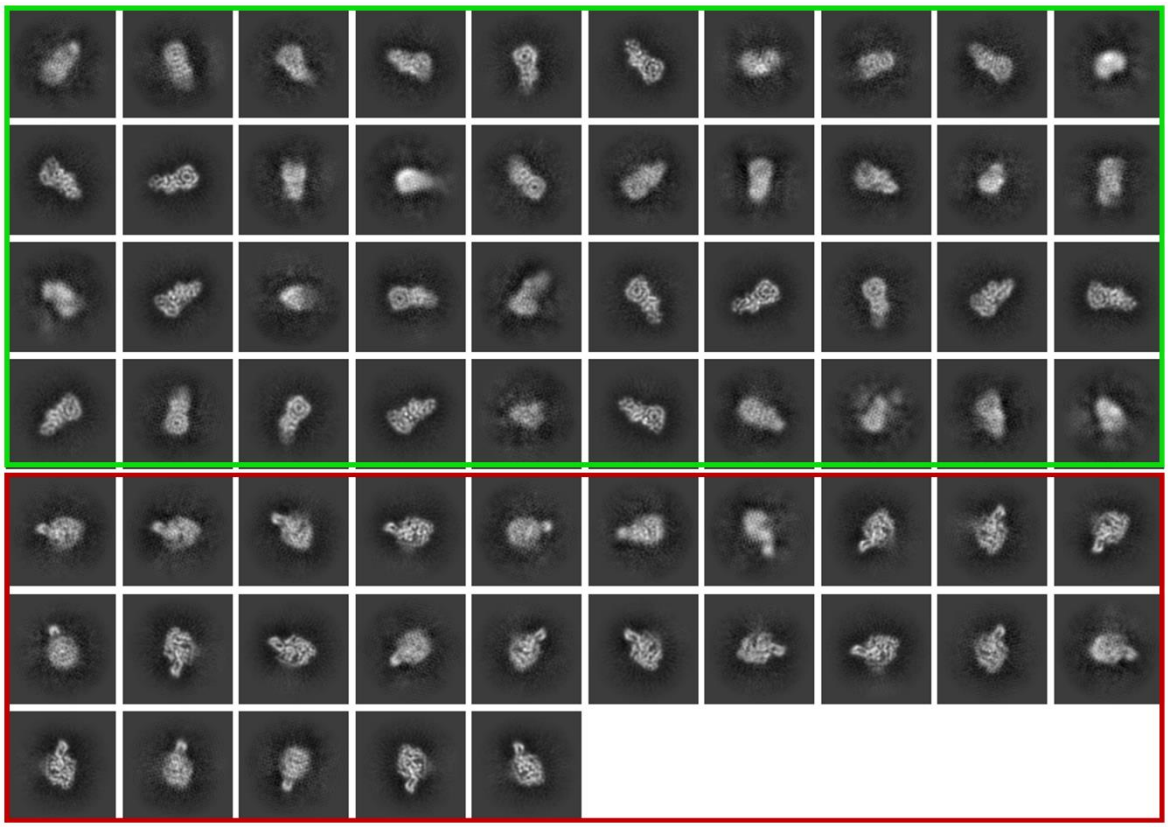

Figure 4.27: 2D classes of particles from blob picking. After three rounds of 2D classifications, the ~2.2 million particles were filtered to ~205,000, classified in 44 classes. Both N (green) and C-halves (red) are present in the dataset.

The particles were subsequently used to create four *ab initio* volumes, resulting in one volume for each half of the protein and two volumes composed of junk particles. Following a round of 3D classification using these volumes, 40 classes of the N-half (36,164 particles) and 25 classes of the C-half (35,791 particles) were selected for template picking, as displayed in Figure 4.28.

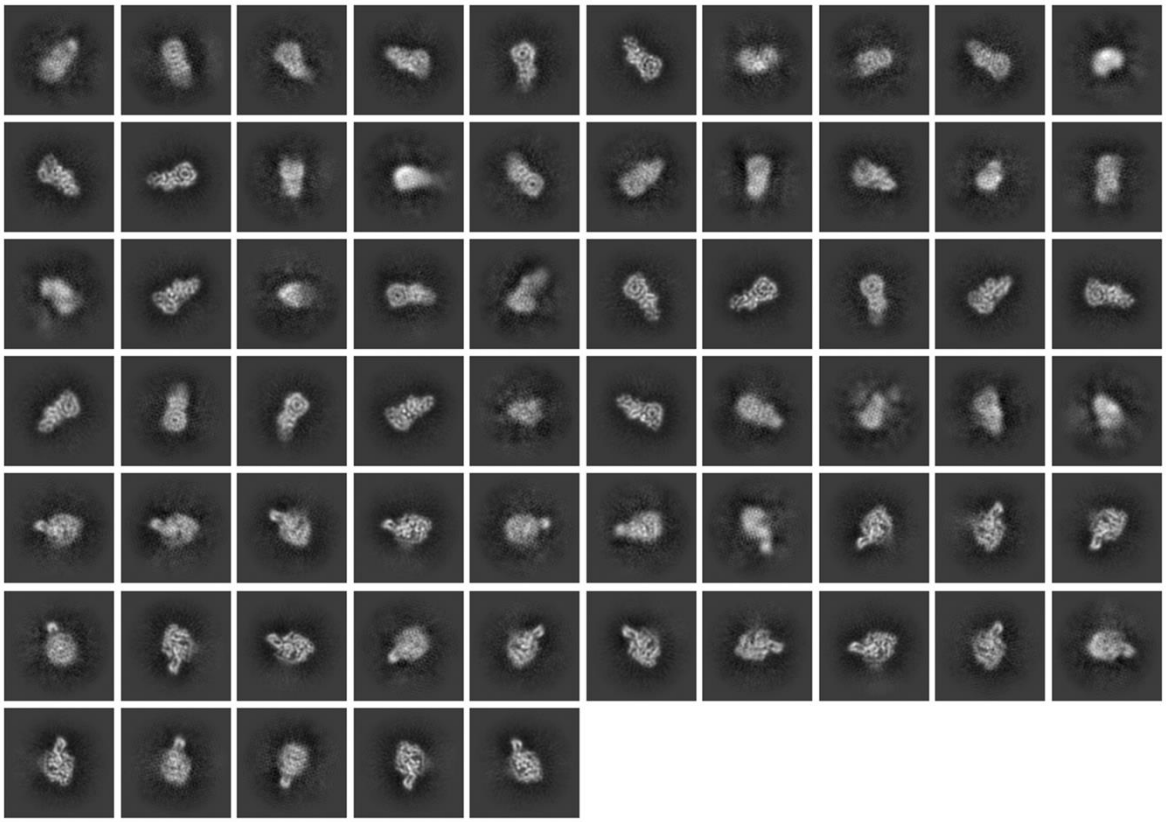

Figure 4.28: 2D classes used as template for template picking. All 65 classes were used for a new round of particle picking, using the template-picking tool.

Over 6.2 million particles were selected using template picking, and after several rounds of 2D classification, these particles were filtered down to approximately 372,000. These were then used to generate five *ab initio* volumes. Once again, one volume was obtained for each of the two halves, while the remaining three volumes corresponded to junk particles. All particles were subsequently 3D classified against these volumes.

After the 3D classification of the particles, each half of MTR was processed independently. Further refinement of the N-half using all 74,917 particles resulted in a 2.90 Å resolution volume map, with its corresponding classes represented in Figure 4.29. These particles were then used to train the TOPAZ picking tool, enabling the selection of N-half particles across all 7,000 micrographs.

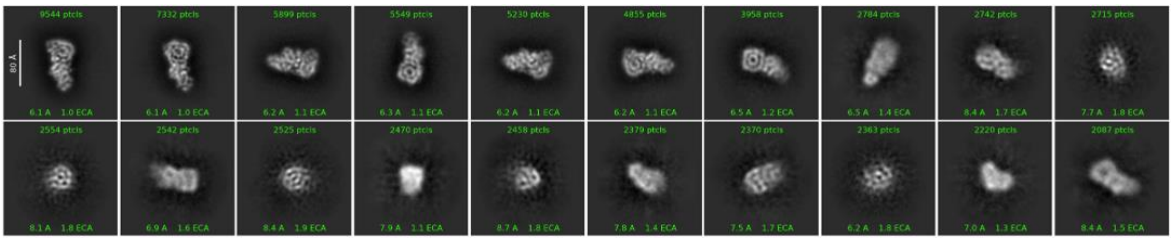

Figure 4.29: 2D classes of 74,917 particles corresponding to the N-half of MTR. The particles were used to train the TOPAZ picking tool for picking N-half particles.

The TOPAZ picking tool identified approximately 183,000 particles corresponding to the N-half of MTR. These particles were classified using the previously generated 2.90 Å volume and a junk volume as references. The filtered ~74,500 N-half particles from TOPAZ picking were combined with the previous ~74,900 particles from template picking, resulting in a total of ~101,000 particles after duplicates were removed. These particles were then used to create two *ab initio* volumes. Following multiple rounds of 3D classification, the volume corresponding to the N-half was further refined using non-uniform and local refinement tools, resulting in a structure formed from 76,292 particles at a resolution of 2.84 Å. This structure is expected to be identical to the N-half from the MTR+MeTHF dataset, as MeTHF is bound to the N-half of both samples. The resolution map and GSFSC curve for the MTR+Ligands N-half map are presented in Figure 4.30.

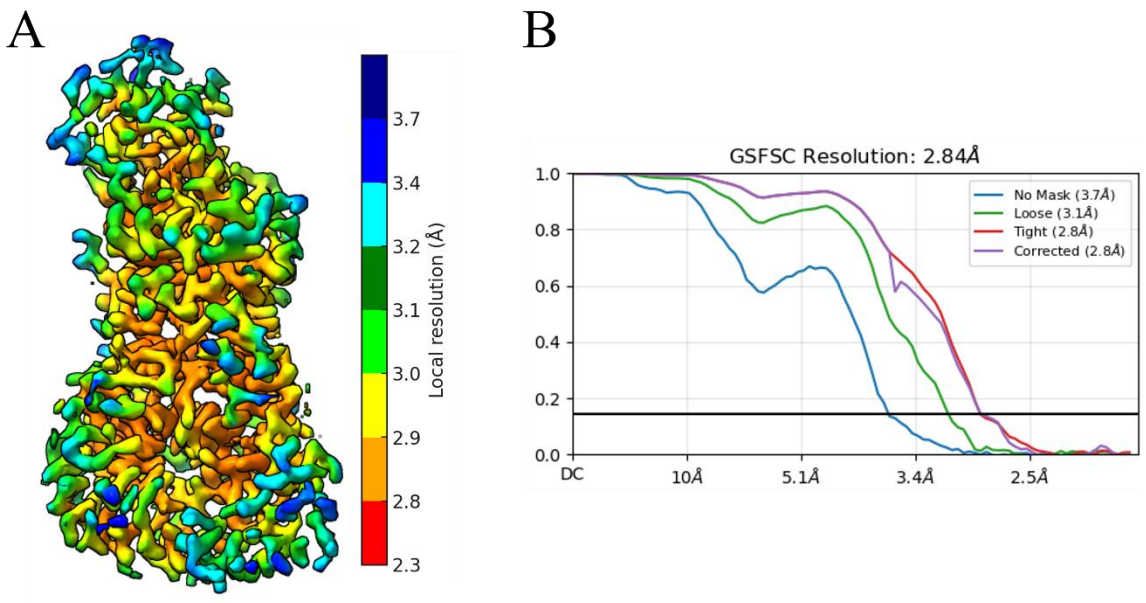

Figure 4.30: MTR N-half structure map of MTR+Ligands grid. A. Local resolution of the N-half map of overall resolution of 2.84 Å. **B**. GSFSC curves for the unmasked, loose, tight and corrected volumes.

For MTR C-half, the 69,281 particles identified through template picking were used to construct a map with a resolution of 3.12 Å. These particles (Figure 4.31) were subsequently employed to train the TOPAZ picking tool.

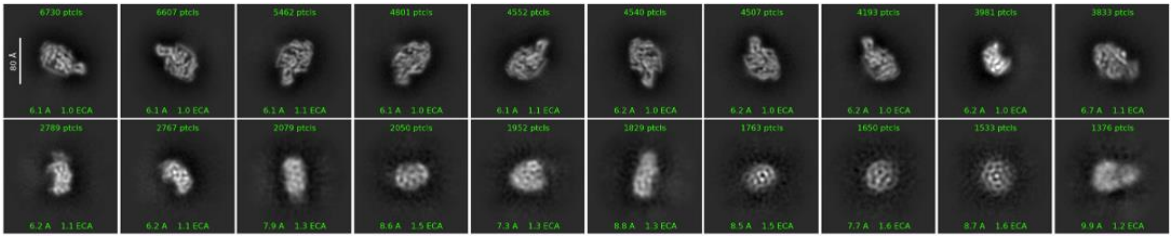

Figure 4.31: 2D classes of 69,281 particles corresponding to the C-half of MTR. The particles were used to train the TOPAZ picking tool for picking C-half particles.

The TOPAZ tool identified approximately 390,000 particles corresponding to the C-half of MTR, which were subjected to 3D classification as described, using the 3.12 Å volume of the C-half and a junk volume. After several rounds of classification, the filtered set of around 198,000 particles was pooled with the previous set of approximately 69,000, resulting in a final total of approximately 169,000 particles after removing duplicates. These particles were then used to generate two *ab initio* volumes. Following further 3D classifications and volume refinements, a map with a resolution of 2.79 Å was obtained, composed of 123,757 particles (Figure 4.32). As observed in the smaller dataset of this same grid (Sections 4.2.1 and 4.3.1), MTR is in the activation conformation, where the AD and CblBD are interacting.

To assess the potential flexibility of the C-half of MTR when bound with Cbl and SAM (MTR+Ligands grid), the 3D Variability Analysis (3DVA) tool in CryoSPARC was employed to explore the heterogeneity of some areas within the model (Punjani and Fleet 2023). The 3DVA was applied to this map due to the observed variability in the Cbl ligand and the region surroiunding it, as indicated by poorer resolution in that specific area (marked by the black arrow in Figure 4.32). The variability seein in the map resolution suggests potential flexibility and the existence of particle populations with different configurations in the Cbl-binding pocket area in the selected dataset.

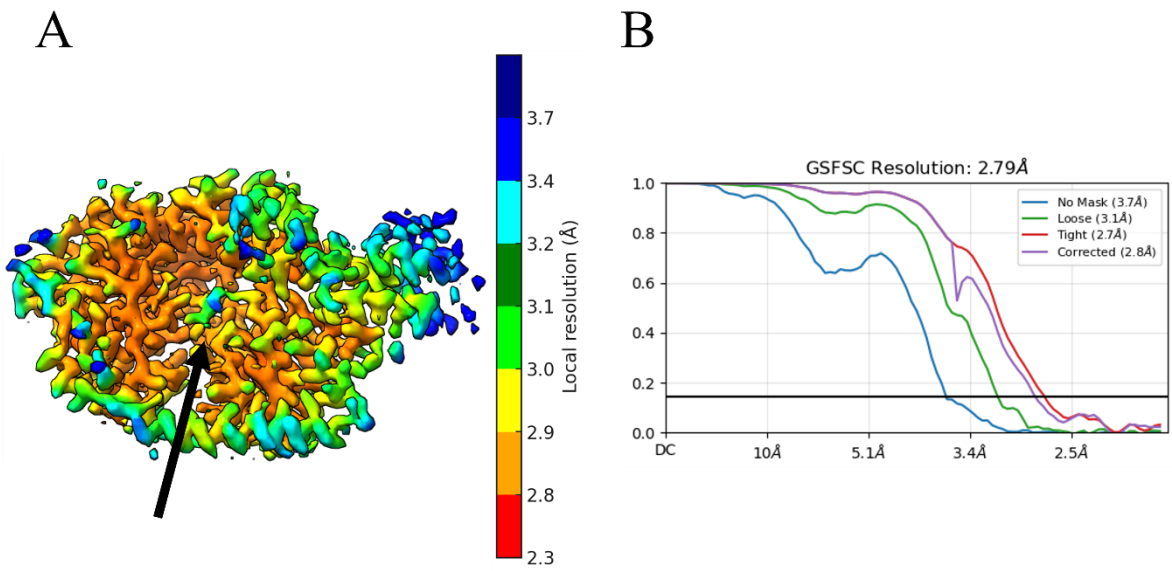

Figure 4.32: MTR C-half structure map derived from the MTR+Ligands grid. A. Local resolution of the C-half map of overall resolution of 2.79 Å. **B**. GSFSC curves for the unmasked, loose, tight and corrected volumes. The black arrow indicates the density corresponding to the Cbl ligand.

The 3DVA analysis revealed two primary volumes of C-half representing distinct conformational states, likely His-on and His-off states of MTR. The first volume, which was more prevalent, consisted of 97,889 particles reached a resolution of 2.82 Å (Figure 4.33A), while the second volume was composed of 25,594 particles and achieved a resolution of 3.01 Å (Figure 4.33C).

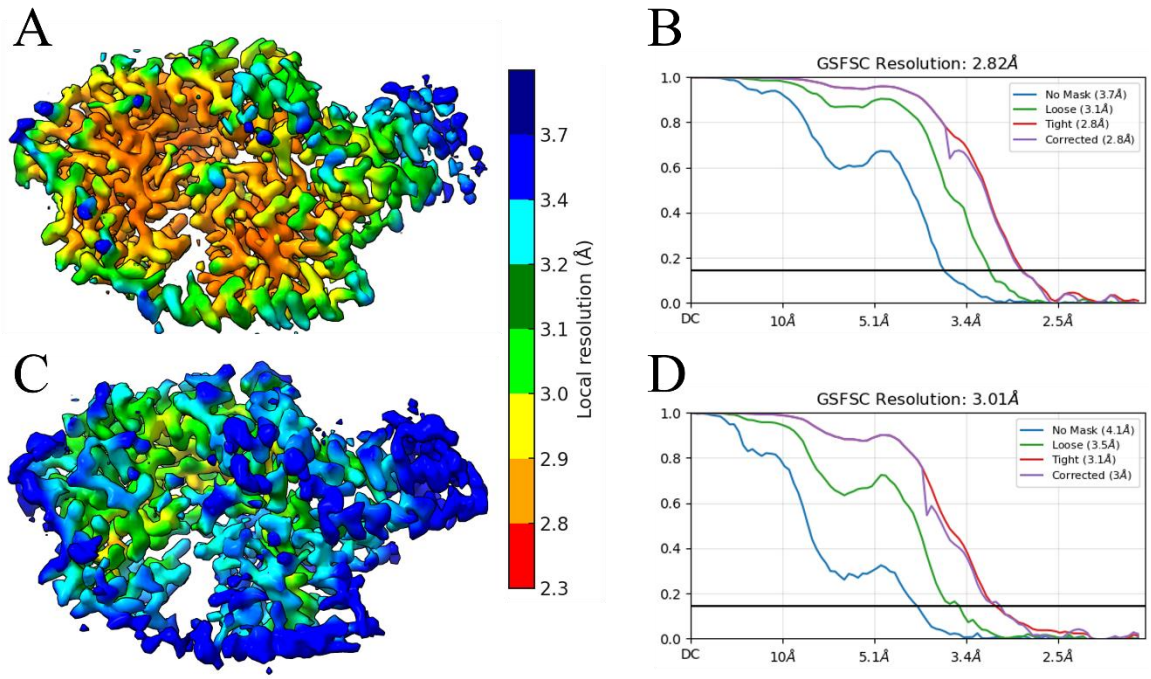

Figure 4.33: MTR C-half structure map derived from the MTR+Ligands grid after the 3DVA. A. C-half volume composed of 97,889 particles and a final resolution of 2.82 Å. **B**. GSFSC curvesof the map shown in A. **C**. C-half volume composed of 25,594 particles and a final resolution of 3.01 Å. **D**. GSFSC curves of the map shown in C.

The conformational difference between the two C-half maps will be further analysed through atomic model building but is likely related to the Cbl ligand and CblBD conformational changes, as Figure 4.32 shows that the centre of the map, where the Cbl-binding pocket is located, has lower local resolution, potentially indicating flexibility. All three maps were utilised for atomic model building, which is detailed in the next chapter.

4.5 Discussion

The use of cryo-EM, combined with access to a high-quality *apo* MTR sample, greatly enhances the possibilities for studying such a flexible protein. Cryo-EM evades the need to truncate the protein into sections to reduce its flexibility, and the availability of an *apo* sample facilitates the exploration of various ligand combinations without being constrained, for instance, to always having Cbl bound to the protein.

The activation conformation, in which the AD and CblBD are positioned to facilitate the reductive methylation of Cbl with the assistance of MTRR, has been captured multiple times in *E. coli* MTR crystal structures (Bandarian, Pattridge et al. 2002, Datta, Koutmos et al. 2008, Koutmos, Datta et al. 2009). This conformation is also the most prevalent form found in the AlphaFold database for MTR across various organisms. The present study is the first to visualise the C-terminal half of human MTR. Previous structural acquisitions of the C-terminal half were typically conducted using truncated proteins, which lacked the N-terminal half of the protein.

Although an intact full-length MTR was utilised in this study, analysis of the collected micrographs revealed that, in the cryo-EM grids, the protein is observed in two distinct halves: one comprising the HcyBD and FolBD (N-half), and the other containing the Cap, CblBD, and AD (C-half). These two sections are separated by a flexible linker (aa 653–661, Linker 1), which is likely crucial for the movements that allow the CblBD to interact with the domains of the N-terminal half. This flexible linker, coupled with the observation that the N-half behaves as a rigid body while the domains in the C-half interact and act as a single unit, likely explains why the protein is visualised as two separate halves.

The structure of full-length MTR from *T. thermophilus*, acquired by X-ray crystallography, also encountered a similar issue with the same Linker 1 connecting the two halves of the protein (Mendoza, Purchal et al. 2023). This linker, located between the FolBD and Cap (aa Ala612-Leu627 in *T. thermophilus*; Gln653-Pro661 in humans), along with another linker between the CblBD and AD (amino acids Ala839-Phe877 in *T. thermophilus*; Asp898-Ile956 in humans), were not modelled due to their mobility. This also suggests that these linkers are key for MTR across species to perform the extensive conformational changes required during catalytic and activation cycles.

Although the protein was observed as two separate units rather than as a single entity, the collection of a large dataset of MTR grids at eBIC (Section 4.4) provided sufficient particles and different orientations of MTR to generate high-density maps. Processing each half of MTR separately enabled the generation of five final maps for further analysis in Chapter 5.

The grid of MTR+MeTHF yielded two maps, one for each half of MTR. The N-half, displaying the HcyBD and FolBD with a bound MeTHF molecule, reached a resolution of 3.05 Å and was generated from 51,788 particles. The C-half, comprising the Cap, CblBD, and AD, achieved a resolution of 2.82 Å. Both maps were used for atomic model construction through rigid-body fitting using AlphaFold structures, which are detailed in the next chapter.

The MTR+Ligands grid yielded three final maps of MTR. The first map corresponds to the N-half, displaying the HcyBD and FolBD with a bound MeTHF molecule, at 2.84 Å resolution. The other two maps represent distinct states of the C-half in the activation conformation (AD and CblBD interaction with Cap-off), at resolution of 2.82 Å and 3.01 Å. Although SAM and Cbl were added to the sample, the density corresponding to SAM is not very evident, but the Cbl density is visible.

In summary, the five density maps generated for the N- and C-terminal halves of MTR will offer valuable insights into the structure of the human protein, for which limited data is currently available. These maps will facilitate the atomic reconstruction of MTR and lay the foundation for a deeper exploration of its conformation and activation mechanisms.

5. Modelling and analysis of MTR structures

5.1 Background

The maps generated from the MTR+MeTHF (Section 4.4.2) and MTR+Ligands (Section 4.4.3) datasets were used to construct models of the N and C-halves of MTR. A total of five models were built: two N-halves (one from each dataset) and three C-halves, corresponding to the *apo* state from the MTR+MeTHF dataset and the two states with Cbl from the MTR+Ligands dataset. The software packages ISOLDE and Phenix were used for molecular replacement, and the details of the maps and models used to generate them are provided in Tables 4.2 and 4.3. The bulit models were then analysed and compared with each other to generate insights into the catalytic mechanism of human MTR.

Table 5.1: Cryo-EM data collection, refinement, and validation statistics of MTR N-half from MTR+MeTHF and MTR+Ligands datasets.

	N-half MTR+MeTHF	**N-half MTR+Ligands**
Data collection and processing		
Magnification	165,000	165,000
Voltage (kV)	300	300
Electron exposure (e^{-}/Å^2)	65.2	65.2
Defocus range (μm)	-0.6 to -2.0	-0.6 to -2.0
Pixel size (Å)	0.508	0.508
Symmetry imposed	N/A	N/A
Initial particle images (no.)	1,638,370	2,195,581
Final particle images (no.)	54,200	76,292
Map resolution (Å)	3.01	2.84
FSC threshold	0.143	0.143
Refinement		
Initial model used	AlphaFold2	AlphaFold2
Model resolution (Å)	3.2	2.9
FSC threshold	0.5	0.5
Model composition		
Nonhydrogen atoms	4925	4925
Protein residues	638	638
Ligands	THH	THH

***B* factors (Å^2)**		
Protein	47.24	48.21
Ligand	44.71	41.20
R.m.s. deviations		
Bond lengths (Å)	0.003	0.003
Bond angles (°)	0.590	0.600
Validation		
MolProbity score	1.59	1.77
Clashscore	6.61	7.32
Poor rotamers (%)	0.38	1.52
Ramachandran plot		
Favored (%)	96.54	96.54
Allowed (%)	3.46	3.46
Disallowed (%)	0.00	0.00

Table 5.2: Cryo-EM data collection, refinement, and validation statistics of MTR C-half models from MTR+MeTHF and MTR+Ligands datasets.

	C-half MTR+MeTHF	**C-half MTR+Ligands His-on**	**C-half MTR+Ligands His-off**
Data collection and processing			
Magnification	165,000	165,000	165,000
Voltage (kV)	300	300	300
Electron exposure (e$^-$ /Å^2)	65.2	65.2	65.2
Defocus range (μm)	-0.6 to -2.0	-0.6 to -2.0	-0.6 to -2.0
Pixel size (Å)	0.508	0.508	0.508
Symmetry imposed	N/A	N/A	N/A
Initial particle images (no.)	1,638,370	2,195,581	2,195,581
Final particle images (no.)	100,343	25,594	97,889
Map resolution (Å)	2.82	3.01	2.82
FSC threshold	0.143	0.143	0.143
Refinement			
Initial model used	AlphaFold2	AlphaFold2	AlphaFold2
Model resolution (Å)	3.0	3.3	2.9
FSC threshold	0.5	0.5	0.5
Model composition			

Nonhydrogen atoms	4739	4876	4830
Protein residues	596	601	596
Ligands	N/A	B12	B12
***B* factors (Å^2)**			
Protein	63.08	71.18	59.61
Ligand	N/A	64.35	48.34
R.m.s. deviations			
Bond lengths (Å)	0.003	0.003	0.002
Bond angles (°)	0.554	0.636	0.508
Validation			
MolProbity score	1.34	1.46	1.48
Clashscore	6.12	8.53	4.97
Poor rotamers (%)	0.98	0.59	1.97
Ramachandran plot			
Favored (%)	98.99	98.16	98.82
Allowed (%)	1.01	1.67	1.18
Disallowed (%)	0.00	0.17	0.00

5.2 Structures of N-half from MTR+MeTHF and MTR+Ligands dataset

5.2.1 Building the N-half model

Based on the N-half map of the MTR+MeTHF dataset, an atomic model was constructed using ISOLDE (Croll 2018), covering residues Thr17-Gly654. The model was generated by rigid-body fitting the AF2-predicted N-half of MTR, based on the crystal structure (PDB: 4CCZ), into the density map. The superimposition of the molecular model and cryo-EM map is presented in Figure 5.1A, along with the Fourier Shell Correlation (FSC) curve from the model construction using ISOLDE and Phenix (Terwilliger, Grosse-Kunstleve et al. 2008, Croll 2018) (Figure 5.1B). Similarly, an N-half model was constructed from the map of MTR+Ligands dataset (Figure 5.1C and D).

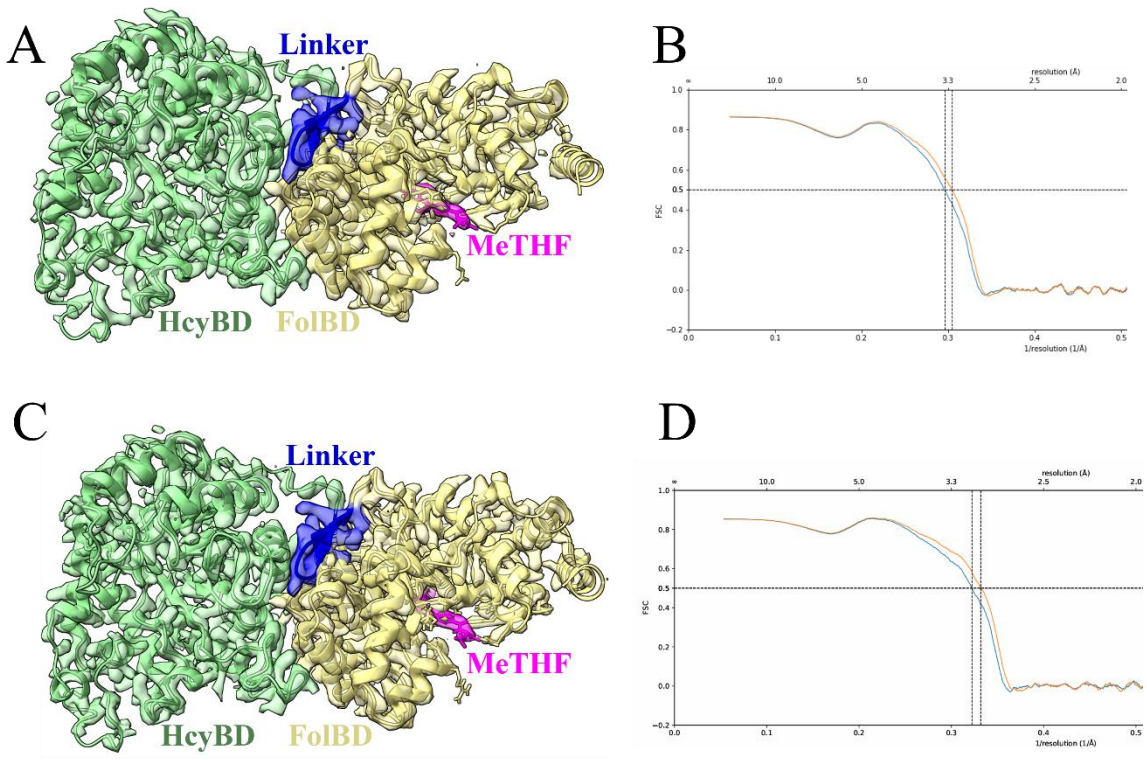

Figure 5.1: Superimposition of the cryo-EM N-half MTR maps and their respective atomic models. A. The model of N-half from MTR+MeTHF dataset highlights the HcyBD (green), FolBD (yellow), MeTHF ligand (pink), and the linker between the domains (blue). The density map reached a resolution of 3.05 Å. **B**. The unmasked and masked FSC curves for the atomic model, constructed using ISOLDE and further refined with Phenix. **C**. The model of N-half from MTR+Ligands dataset highlights the HcyBD, FolBD, MeTHF ligand, and the linker between the domains. The density map reached a resolution of 2.84 Å. **D.** The unmasked and masked FSC curves for the atomic model, constructed using ISOLDE and further refined with Phenix.

Comparison between the N-half models from MTR+MeTHF and MTR+Ligands datasets (Figure 5.2) show that although they were derived from different ligand combinations, both adopt nearly identical structures (RMSD 0.356 Å), with very mild differences in some loops on both domains. These subtle differences could be due to slight flexibility in the loops or, different resolution of the models (3.05 Å resolution for MTR+MeTHF, 2.84 Å resolution for MTR+Ligands), which could result in placing the residues of the loops in slightly different positions.

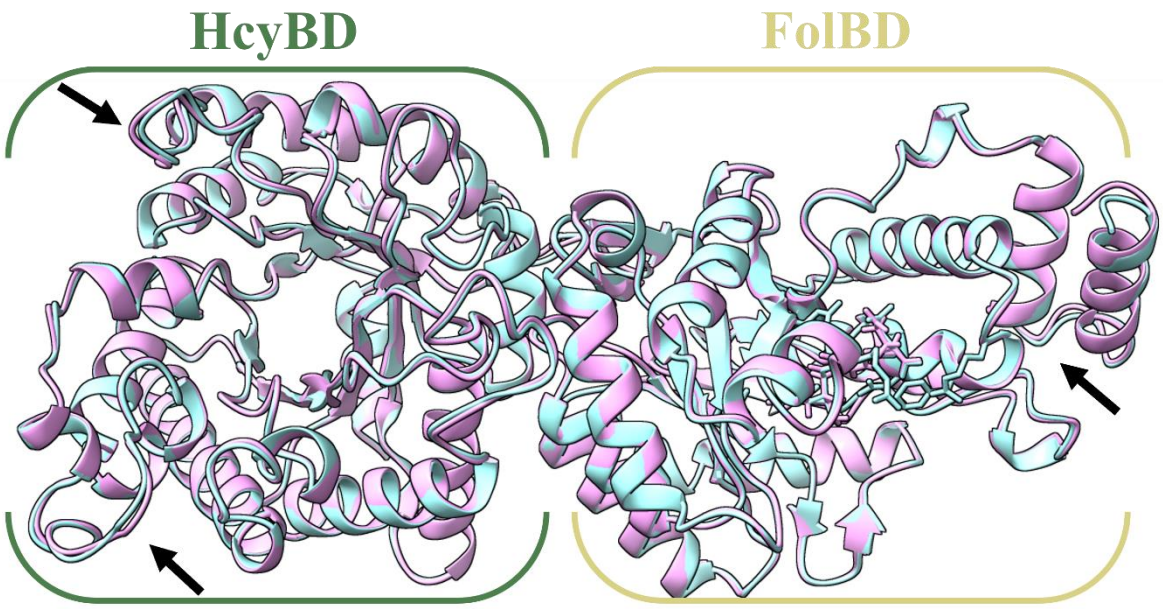

Figure 5.2: Superimposition of the atomic model derived from the MTR+Ligands cryo-EM dataset (pink) with the model derived from the MTR+MeTHF dataset (blue). There are some very discreet conformational differences in some loops of the structure (identified by black arrows). The root-mean-square deviation (RMSD) between the 636 pruned atom pairs is 0.356 Å.

In the following description of the N-half structure, model derived from the higher resolution MTR+Ligands map will be used.

5.2.2 MTR N-half structure and MeTHF binding

The N-half model built from the MTR+Ligands dataset was then superimposed with the crystal structure of human MTR bound with THF (PDB: 4CCZ) (Figure 5.3). The HcyBD and FolBD exhibited the expected folds, which consist of a (β/α)8 barrel, with their axes arranged perpendicularly. Furthermore, the linker between HcyBD and FolBD, an 11-residue β-hairpin buried between the domains (aa Met355-Ile365), is also visible and is known to enhance the stability of this half of MTR (Bandarian, Pattridge et al. 2002, Koutmos, Pejchal et al. 2008). This inter-domain stability was previously confirmed in *E. coli* and *T. maritima* MTR through tryptic digestion, where even 3% trypsin was unable to cleave any argenine or lysine residues within these domains (Jarrett, Huang et al. 1998). The refined model of the N-half MTR displayed no significant differences from the crystal structure (RMSD 0.694

Å), despite being bound to different ligands (THF in the crystal structure, MeTHF in the current cryo-EM structure).

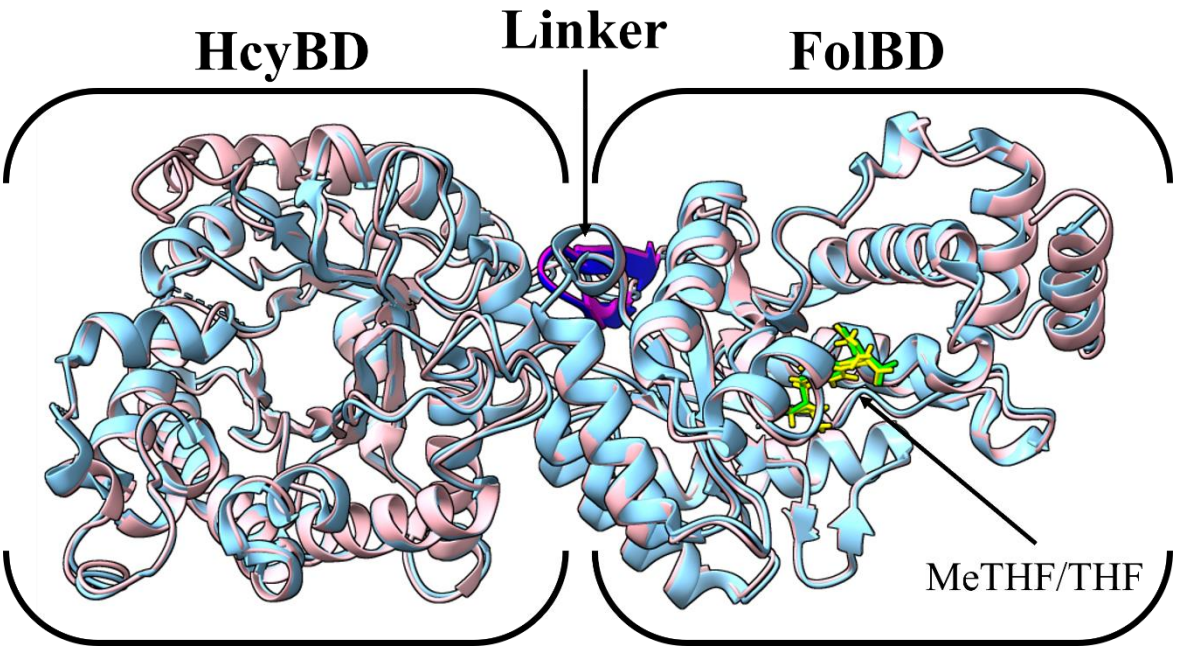

Figure 5.3: Superimposition of the atomic model derived from the MTR+Ligands cryo-EM dataset with the crystal structure of the human MTR N-half (PDB: 4CCZ). The cryo-EM derived atomic model is shown in pink, while the crystal structure is depicted in blue. The linkers between the domains are shown in dark blue for the crystal structure and in dark pink for the cryo-EM structure. The THF molecule bound to the crystal structure is shown in green and the MeTHF bound to the cryo-EM-derived structure is shown in yellow. The root-mean-square deviation (RMSD) between the 596 pruned atom pairs is 0.694 Å.

In the N-half map, there is clear density for the MeTHF ligand (Figure 5.1A and C). The MeTHF molecule is positioned within its binding pocket in the FolBD, as observed in previous crystal structures of *T. maritima* and *T. thermophilus* MTR bound to MeTHF (Evans, Huddler et al. 2004, Koutmos, Pejchal et al. 2008, Yamada and Koutmos 2018). Figure 5.4 illustrates the atomic model with MeTHF, providing a detailed view of its placement.

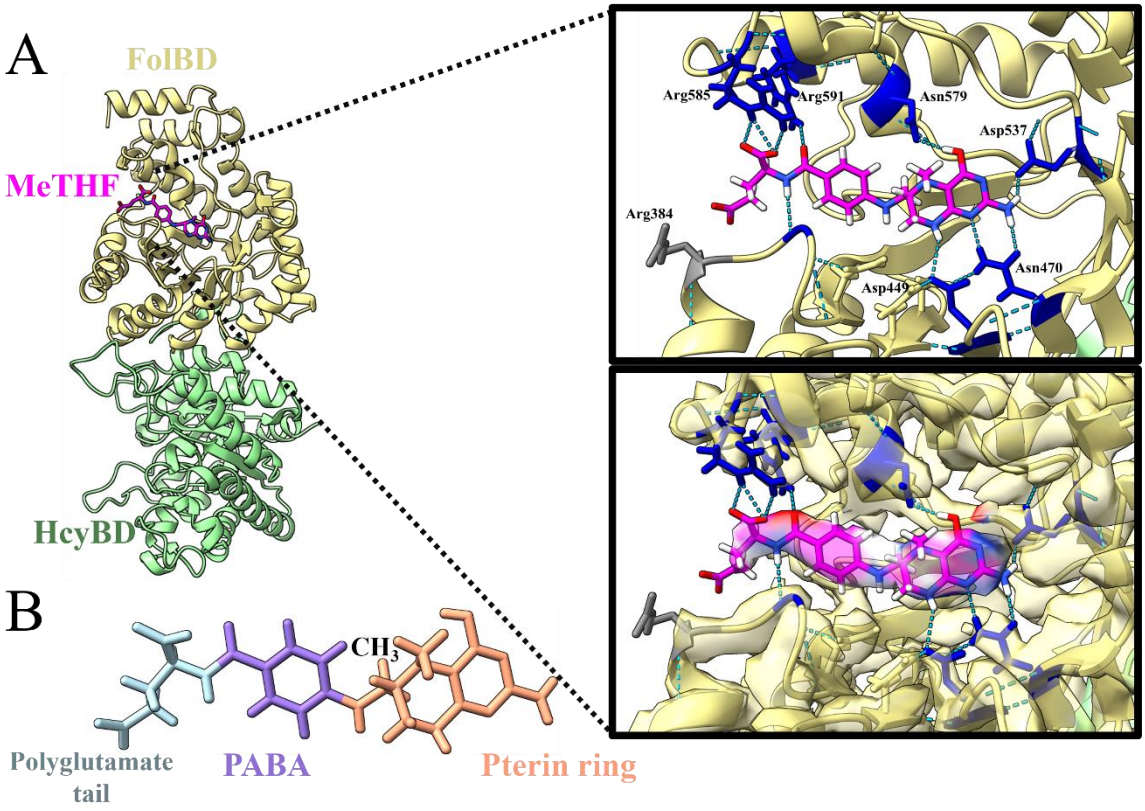

Figure 5.4: Molecular model of the N-half MTR from the MTR+Ligands dataset. A. The overall structure comprises the HcyBD (green) and the FolBD (yellow) with a molecule of MeTHF (pink). The detailed view highlights the residues interacting with the ligand in blue, while the residue observed to interact with the ligand in the crystal structure (PDB: 4CCZ) but not in this model (Arg384) is displayed in grey. It also displays the overlaying of the density map and model. **B**. Different regions of MeTHF molecule, indicating the pterin ring (light orange), p-aminobenzoic acid (PABA) side chain (purple), and the polyglutamate tail (light blue).

In the human crystal structure of N-half MTR (PDB: 4CCZ), the pterin ring of THF is stabilized by hydrogen bonds formed with four conserved residues: two aspartates (Asp449, Asp537) and two asparagines (Asn470, Asn579). These bonds are also observed in the *T. maritima* MTR bound to MeTHF (PDB: 1Q8J) (Evans, Huddler et al. 2004). The high resolution of MTR+Ligands dataset allowed for the visualisation of all four interactions between the pterin ring of MeTHF and Asp449, Asp537, Asn470, and Asn579, similar to the crystal structure of the human protein (PDB: 4CCZ) and the structure of *T. maritima* MTR (PDB: 1Q8J) (Evans, Huddler et al. 2004).

Additionally, the p-aminobenzoic acid (PABA) side chain of THF in the human crystal structure forms hydrogen bonds with the residues Gly382 (main-chain atoms),

Arg384 (main-chain atoms), and Arg591. In the cryo-EM model, the residues Gly382, Arg585, and Arg591 are shown to interact with the PABA side chain of MeTHF. The interaction with Arg585 shown in the cryo-EM model is not displayed in the crystal structure. Lastly, the interaction with Arg384, found in the crystal model, is not observed in this cryo-EM model.

Overall, the MTR+Ligands density map generated a model that allowed for the visualisation of the interaction of the ligand with the four conserved residues (Asn470, Asn579, Asp449, Asp537), which are known to interact with the folate ligand not only in the human protein but also in the bacterial orthologue and were not fully observed in the previous cryo-EM model (Evans, Huddler et al. 2004, Koutmos, Pejchal et al. 2008).

5.3 Structures of C-half from MTR+MeTHF and MTR+Ligands dataset

5.3.1 Building the C-half model from MTR+MeTHF dataset

Based on the C-half map from the MTR+MeTHF dataset (2.82 Å resolution), an atomic model was constructed using ISOLDE and Phenix (Terwilliger, Grosse-Kunstleve et al. 2008, Croll 2018), covering residues Pro670-Asp1265 (Figure 5.5). As there is no available structure for the human C-half of MTR, the model was built by rigid-body fitting the AlphaFold predicted structure into the cryo-EM density.

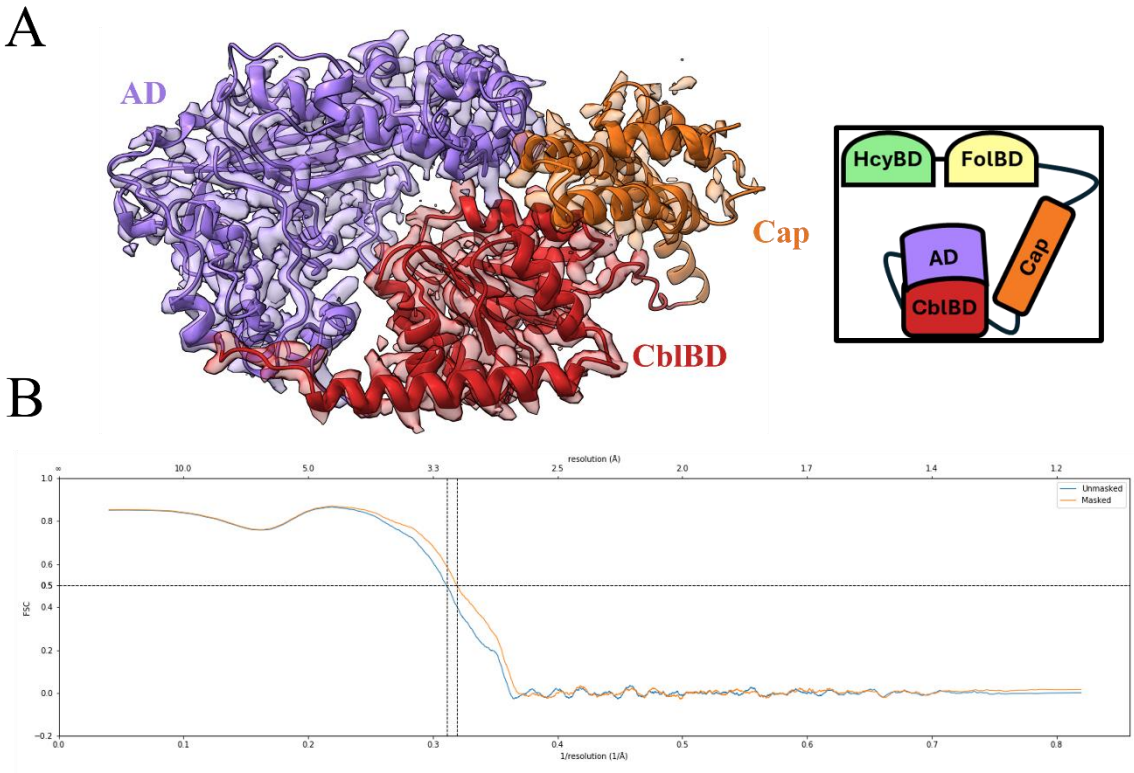

Figure 5.5: Superimposition of the cryo-EM map (2.82 Å resolution) and the atomic model of the C-half of MTR. A. The model highlights the Cap (orange), CblBD (red), and the AD (purple). The cryo-EM map reveals clear density for secondary structure elements and the overall shape of MTR in the activation conformation, with the CblBD and AD interacting with each other. **B**. The unmasked and masked FSC curves for the atomic model, constructed using ISOLDE and further refined with Phenix.

The C-half of MTR does not display any bound ligands and is positioned with the CblBD (aa Gln772-Glu907) interacting with the AD (aa Ser923-Asp1265), indicating that it is in the activation conformation. The Cap domain (aa Pro670-Asn762), formed by a four-helix bundle, is positioned away from the CblBD, referred to as the Cap-off state. The CblBD exhibits its characteristic Rossmann fold consisting of alternating parallel β-strands and α-helices. The Cbl binding pocket does not show any ligand electron density (Figure 4.5), confirming the absence of the cofactor bound to MTR. Finally, the AD, which features a mixed α/β fold and adopts a typical C-shape, is positioned above the CblBD, with no ligand (SAM) observed in its binding pocket.

The Cap domain of MTR is known to adopt either a Cap-on or Cap-off conformation, depending on whether it is interacting with the CblBD to protect the cobalamin (Cbl) or not (Bandarian, Pattridge et al. 2002). In the generated model here, the Cap is in the Cap-off state but exhibits significant flexibility, as evidenced by the lack of density in part of one of its four helices (helix α4) (Figure 5.6). The ligands Cbl and S-adenosylhomocysteine (SAH) were placed in the structure using AlphFill for better visualisation the lack of their density in the cryo-EM map. The absence of density for these ligands confirms that the C-terminal half of MTR in the current sample is in an *apo* form, and that the AD and CblBD can interact even in the absence of Cbl.

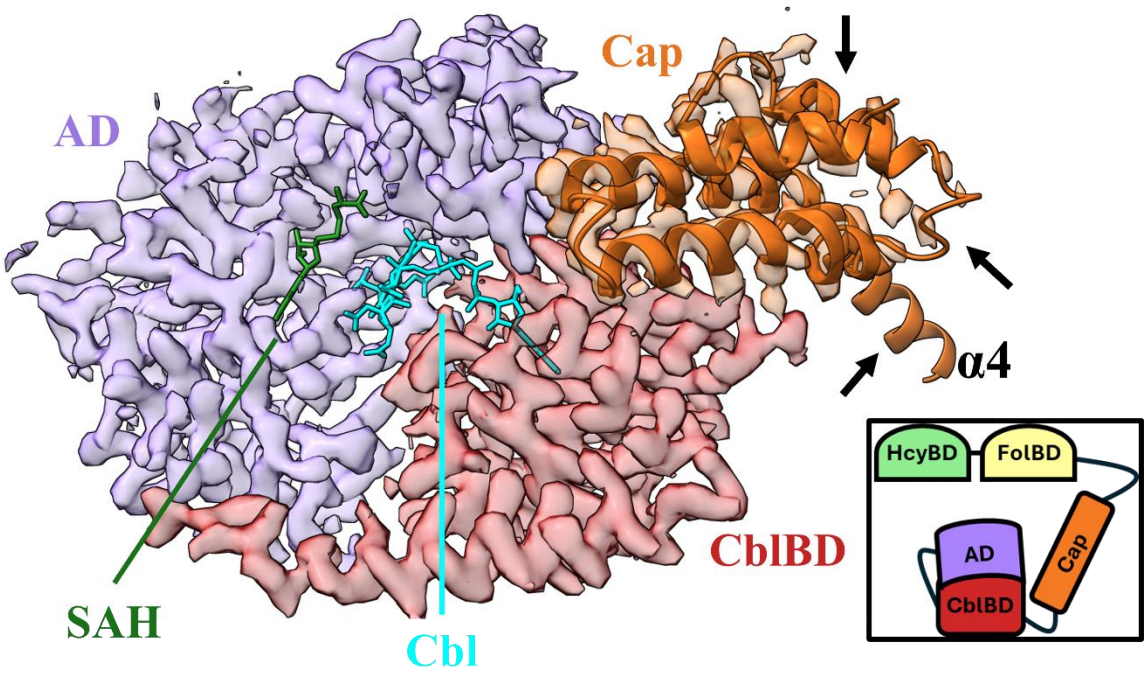

Figure 5.6: Superimposition of the cryo-EM map (2.82 Å resolution) and the atomic model of the C-half of MTR. A. The model highlights the Cap (orange), CblBD (red), the AD (purple). The ligands SAH (dark green), and Cbl (cyan) were superimposed in the model for visualisation of the lack of their densities. Black arrows indicate the regions of the Cap that did not display any density in the map, suggesting high flexibility in these areas. The molecular models of the CblBD and AD were removed for better visualisation of the density in those regions.

The residue His785 in human MTR is highly conserved across species. In addition to coordinating with the lower axial position of the cobalt atom (Base-off, His-on), it has

been shown in *E. coli* MTR to play a crucial role in the reactivation cycle of the protein (Drennan, Huang et al. 1994, Bandarian, Pattridge et al. 2002, Koutmos, Datta et al. 2009). A conservation analysis of MTR was conducted using ConSurf (Ashkenazy, Abadi et al. 2016), where 150 orthologues MTR sequences were aligned, and the variability of the residues was examined (Figure 5.7).

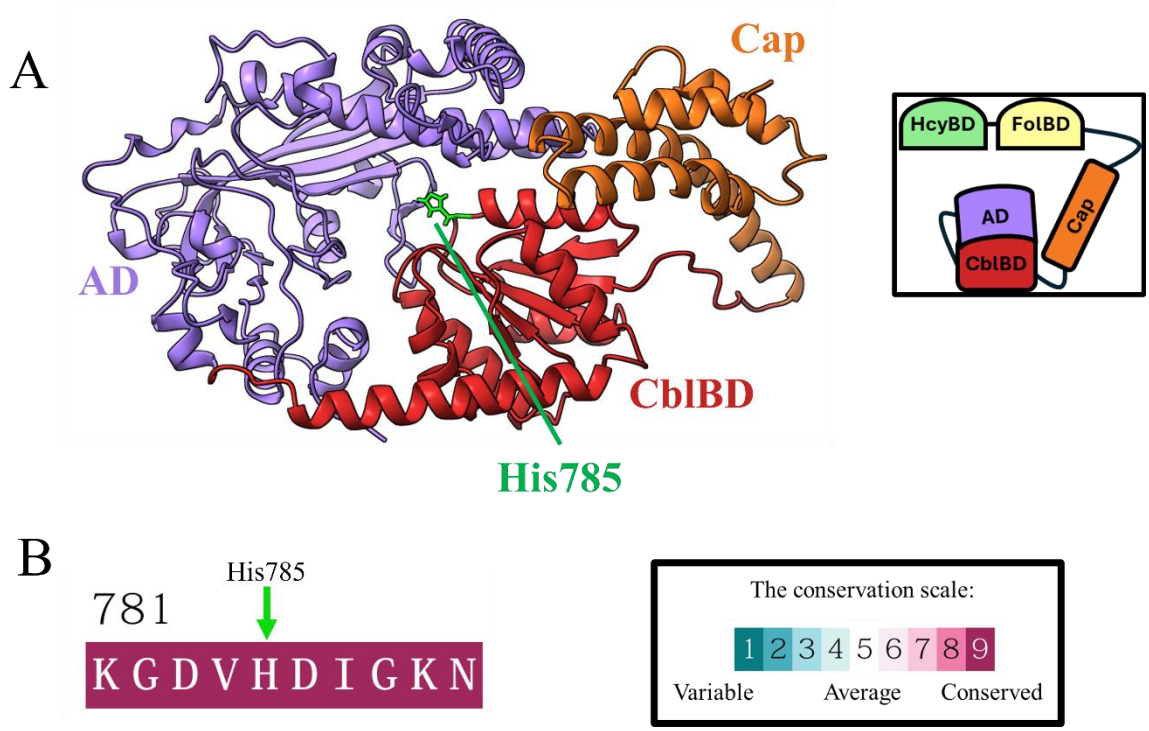

Figure 5.7: Position of His785 in Human MTR and its Conservation Analysis Performed on ConSurf. A. C-half of the MTR atomic model from the MTR+MeTHF dataset, showing the Cap (orange), CblBD (red), AD (purple), and His785 (green). The schematic cartoon illustrates the position of the domains within the model. **B.** Conservation analysis of the His785 region with a corresponding conservation scale.

The region of the His785 residue, in the CblBD, is highly conservated in MTR orthologues, and the residues in the region Lys781-Asn790 displayed no variation within the sequences analysed. The lack of variability of His785 is not surprising, as it is responsible for shifting Cbl between the His-on and His-off states (Section 1.4.3) by coordinating with the lower axis of the Co atom (Koutmos, Datta et al. 2009). This mechanism allows the change of coordination of Cbl to 4-coordinate, which is necessary for the reduction of cob[II]alamin into cob[I]alamin during the reactivation cycle. The *apo* structure of the C-

half from *T. thermophilus* MTR shows the conserved histidine (His761 in this species) positioned away from the docked Cbl cofactor (Mendoza, Purchal et al. 2023). To examine the position of His785 in the human *apo* C-half MTR, a molecule of Cbl was docked into its expected binding pocket in the structure using AlphaFill, and the structure from the bacterial *apo* C-half (PDB: 8SSD) was aligned with the human MTR structure (Figure 5.8). The histidine side-chains in both structures exhibit similar configuration, positioned away from the lower axis of the Co atom. In the human structure, His785 is approximately 6.0 Å from the Co atom, and therefore this distance means that it cannot coordinate with Cbl. The Loop1, which contains the His785 and is located between the β1 and α1, may need to undergo a reorientation to allow Cbl binding in the His-on state.

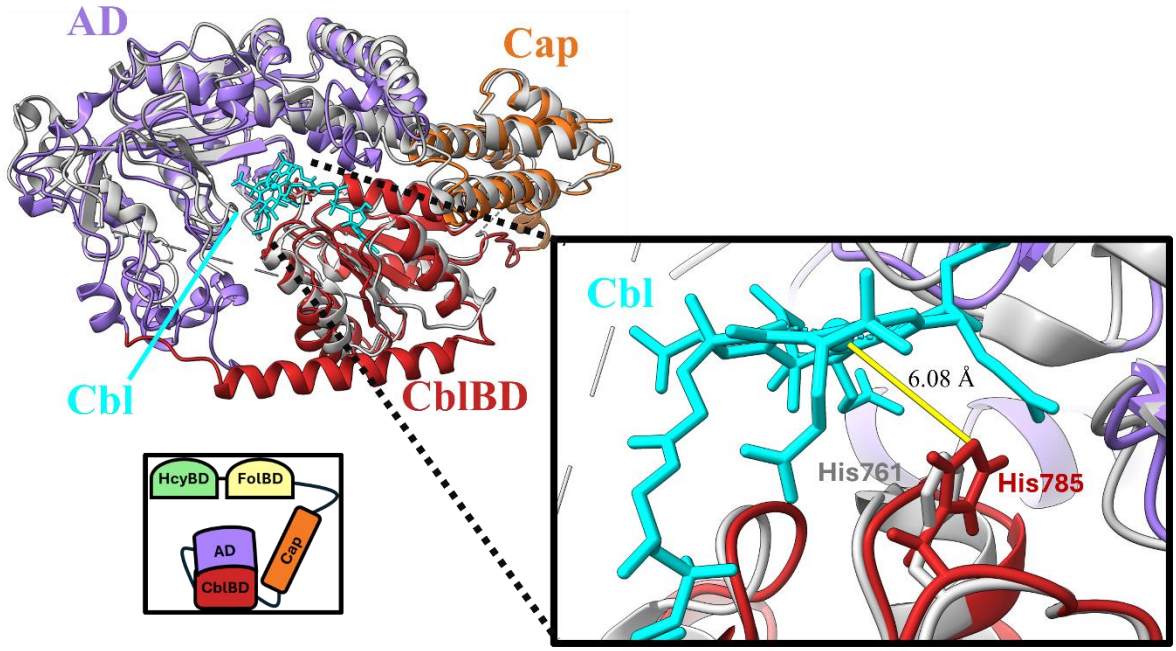

Figure 5.8: Superimposition of the apo C-half from T. thermophilus MTR (PDB: 8SSD) with the corresponding apo region of the human protein. The human domains are highlighted in orange (Cap), red (CblBD), and purple (AD), while the bacterial structure is shown in grey. The Cbl molecule, in cyan, was manually docked onto the structure using AlphaFill, to show its putative position. The zoomed-in area shows the position of the conserved histidines relative to the lower axial of the Co atom. The distance between His785 (red) in human MTR and the Co ion is 6.08 Å, suggesting that the CblBD of MTR may need to undergo conformational changes to achieve the His-on state.

In addition to the conserved histidine (human His785) in the CblBD, a conserved tyrosine in the AD (Tyr1177 in human MTR) plays a crucial role in the reactivation of cob[II]alamin (Section 1.4.3 State 'D' in figure 1.7). The same conservation analysis as shown in Figure 5.7B revealed that the residues in that region have a conservation greater than 80% in all sequences analysed, and the Tyr1177 showed a conservation of 100% (Figure 5.9).

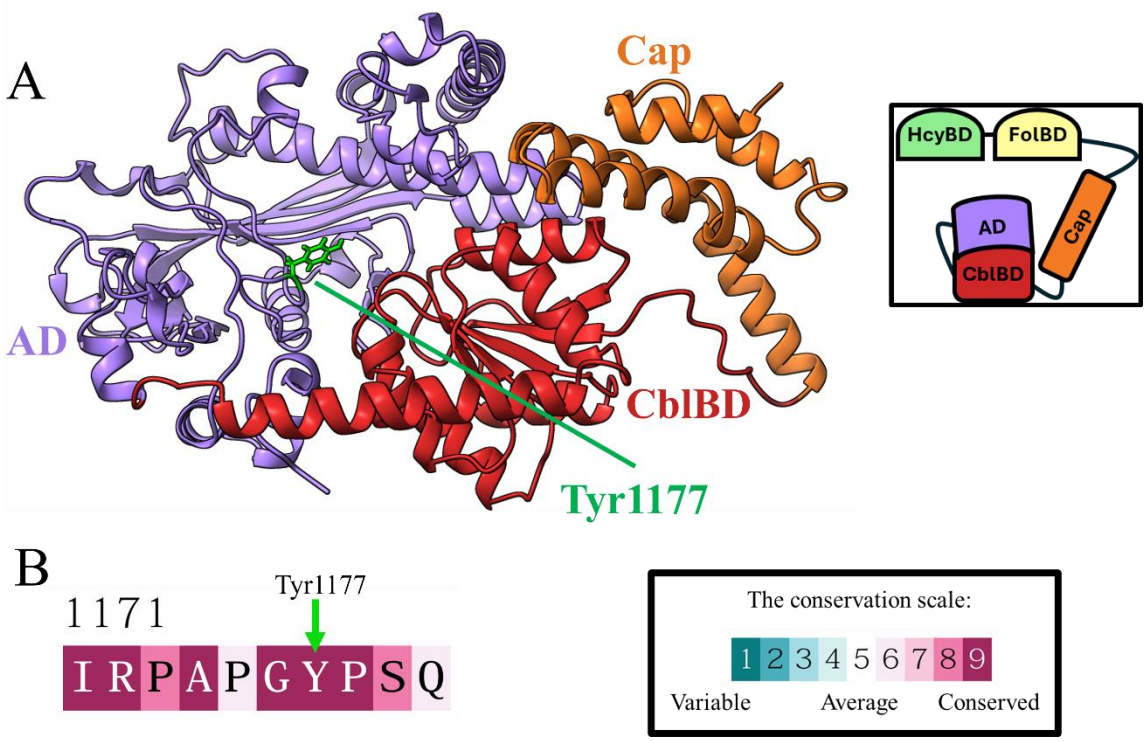

Figure 5.9: Position of Tyr1177 in Human MTR and its Conservation Analysis Performed on ConSurf. **A**. C-half of the MTR atomic model from the MTR+MeTHF dataset, showing the Cap (orange), CblBD (red), AD (purple), and Tyr1177 (green). The schematic cartoon illustrates the position of the domains within the model. **B**. Conservation analysis of the Tyr1177 region with a corresponding conservation scale.

During the reductive methylation of Cbl, MTR adopts the activation conformation, with the AD positioned above the CblBD (Bandarian, Pattridge et al. 2002). As described in Section 1.4.3, Tyr1177 coordinates with the water molecule located at the upper axial position of the Co atom in the 5-coordinated His-off cob[II]alamin, which lengthens the Co-O bond. This bond elongation shifts the Cbl into 4-coordinated His-off cob[II]alamin,

making it more receptive to an electron, facilitating its conversion into MetCbl (Koutmos, Datta et al. 2009).

It has been suggested that this tyrosine undergoes a positional change to move closer to cobalamin (Cbl) in its His-off state in *E. coli* MTR structure (Koutmos, Datta et al. 2009); however, due to the lack of structural data it remains uncertain whether this conformational change occurs in the human protein. To evaluate the position of Tyr1177 in the *apo* C-half structure of human MTR, two *T. thermophilus* MTR structures were used for comparison (Mendoza, Purchal et al. 2023) (Figure 5.10). The first structure, the *apo* C-half (PDB: 8SSD), shows the equivalent tyrosine (Tyr1132) oriented away (7.3 Å) from the theoretical position of the Cbl upper axial ligand. In the second structure, the Cbl-loaded C-half in the His-off state (PDB: 8SSE), the tyrosine is closer to the Cbl upper axial position (4.8 Å), potentially lengthening the Co-O bond.

Although the human MTR structure is also in the *apo* state, Tyr1177 is positioned similarly to its location in the Cbl-loaded structure, nearer to the cobalt ion (3.1 Å, Figure 5.10). This positioning of Tyr1177 suggests that the conformational change of the conserved tyrosine, shown in *T. thermophilus*, may not occur in human MTR, as the residue is well-defined in the density map with no apparent flexibility, which might otherwise indicate movement of this residue.

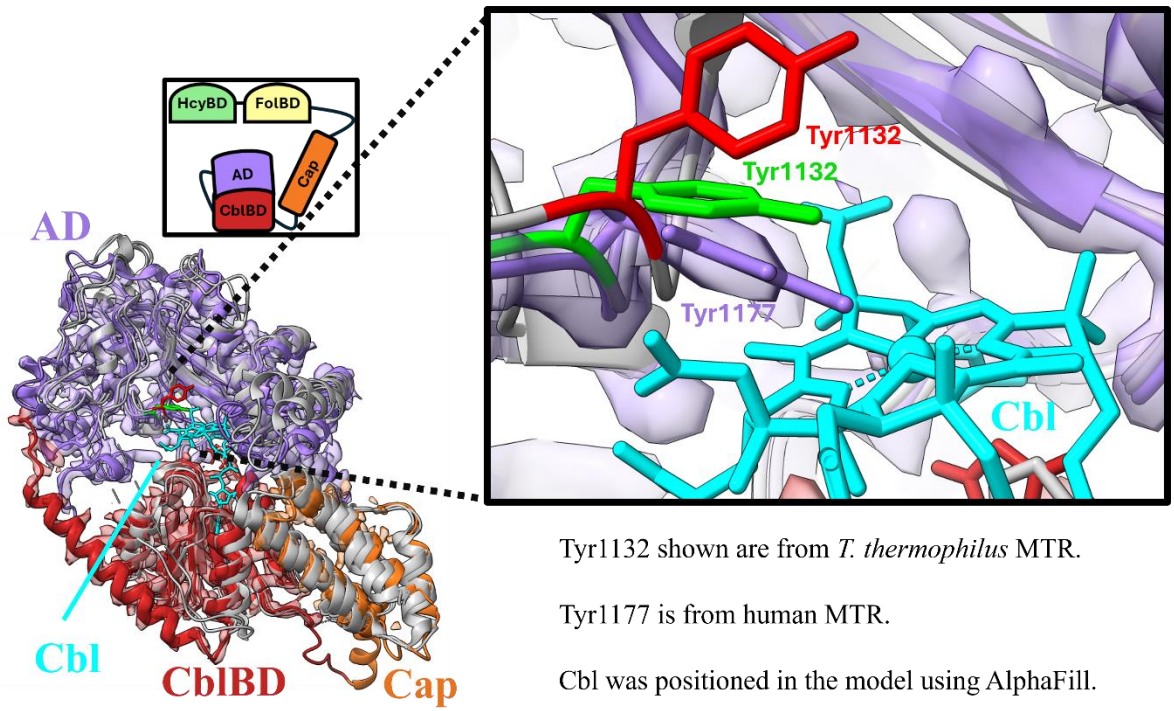

Figure 5.10: Superimposition of the T. thermophilus MTR apo C-half (PDB: 8SSD) and the Cbl-loaded His-off C-half (PDB: 8SSE) with the corresponding apo region of the human protein from the MTR+MeTHF dataset. The human domains are highlighted in orange (Cap), red (CblBD), and purple (AD), while the bacterial structures are shown in grey. The Cbl molecule, in cyan, was positioned using AlphaFill to illustrate its putative position. The zoomed-in area illustrates the positions of the conserved tyrosines relative to the cobalt atom. The human Tyr1177 is shown in purple, the tyrosine from the His-off structure in green, and the tyrosine from the apo bacterial MTR in red.

In summary, the structure of the human C-half MTR from MTR+MeTHF dataset revealed the Cap, CblBD, and AD positioned in the activation conformation, without the presence of any ligand, potentially displaying the same interdomain interactions as the MTR without added ligands would. Furthermore, the positioning of His785 shows that it is in the His-off position, indicating that it would need to undergo conformational changes to coordinate with the lower axial ligand of Cbl. Finally, Tyr1177 is positioned close to where Cbl would be, a position typically seen in bacterial MTR during the elongation of the Co-O bond, which is triggered by SAM binding (Koutmos, Datta et al. 2009). The position of Tyr1177 in the human MTR structure suggests that it may not need to undergo conformational changes for the reductive methylation of Cbl, as there is no SAM bound to

the AD to trigger its movement closer to Cbl. Further structural insights into the C-half of MTR can be gained by comparing this dataset with the Cbl-bound structures.

5.3.2 Building two C-half models from the MTR+Ligands dataset

The two C-half maps generated from the MTR+Ligands grids were used to construct atomic models. Both maps display MTR in the activation conformation, showing the interaction between the AD and CblBD, in a Cap-off state. There is density for the Cbl ligand in the expected pocket, and it was added during model construction. The density for SAM in both C-half maps was not fully evident, so it was not included during the model building.

After refinement with ISOLDE and Phenix, the models were reconstructed and classified. The model composed of ~97,000 particles represent MTR C-half with Cbl bound in the His-off stateat 2.82 Å, comprising the residues Pro670-Asp1265. The model built with ~25,000 particles represent MTR C-half with Cbl bound in the His-on state at 3.01 Å resolution, comprising residues Glu665-Asp1265. It is clear, based on the number of particles, that the His-off state is more prevalent than His-on. Both structures are analysed in the following subsections 5.3.2.1 and 5.3.2.2.

5.3.2.1 Structure of MTR C-half in the His-off state

Figure 5.11 shows the atomic model superimposed with the cryo-EM map, along with the FSC curve from the model building indicating its resolution cutoff at 2.82 Å. The ligand Cbl was positioned and modelled based on the density observed in the map.

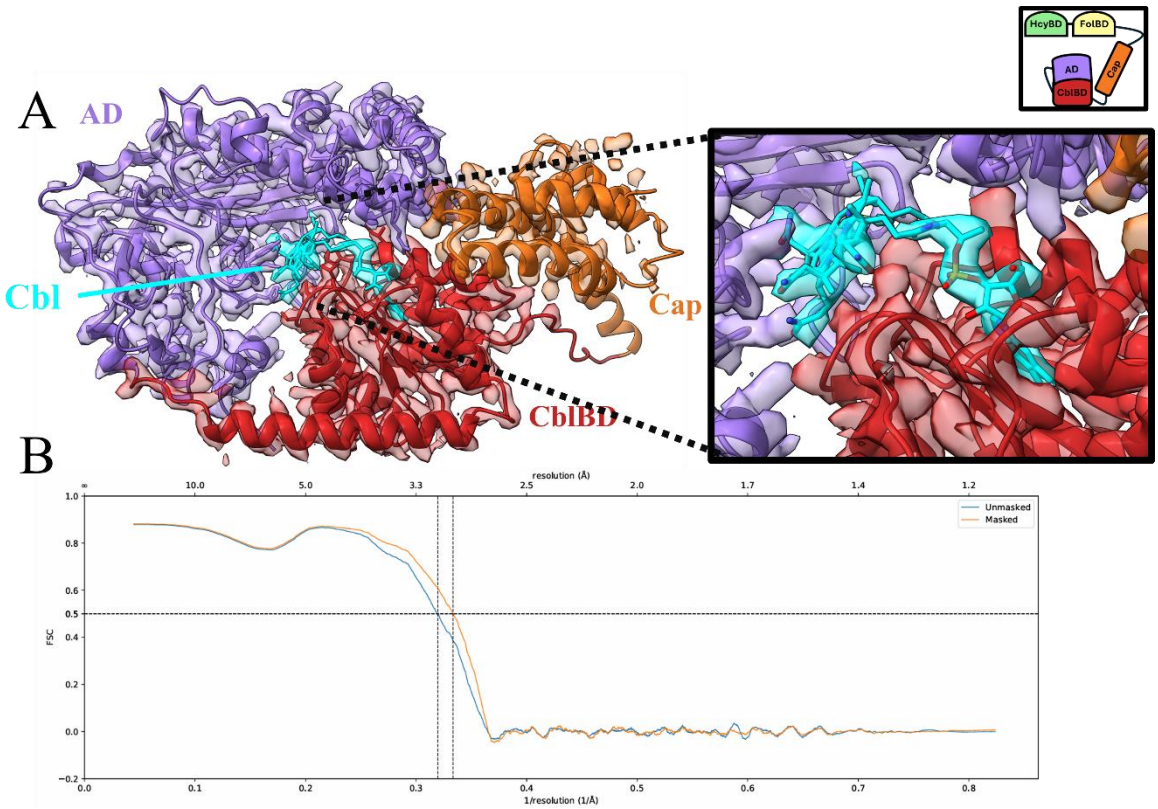

Figure 5.11: Superimposition of the cryo-EM map and the atomic model of the C-half of His-off MTR from MTR+Ligands dataset. A. The model (which turns out to be in the His-off state) highlights the Cap (orange), CblBD (red), the AD (purple), and Cbl (cyan). The cryo-EM map reveals clear density for secondary structure elements and the overall shape of MTR in the activation conformation, with the CblBD and AD interacting and the Cbl positioned on its binding pocket, and no clear density for SAM. The inset highlights the Cbl density superimposed with Cbl molecule. **B**. The unmasked and masked FSC curves for the atomic model (3.0 Å resolution), constructed using ISOLDE and further refined with Phenix.

Similar to the C-half model from the MTR+MeTHF dataset (Section 5.3.1), the density of the Cap in the C-half His-off state is not well defined. This can be attributed to the flexibility of the region, as it needs to undergo significant conformational changes when shifting between the Cap-on and Cap-off states. Additionally, Cbl is bound within the expected binding pocket, with the DMBI tail accommodated in the CblBD and the corrin ring close to the AD, as previously described for *E. coli*, *T. filiformis*, and *T. thermophilus* MTR structures (Bandarian, Pattridge et al. 2002, Datta, Koutmos et al. 2008, Koutmos, Datta et al. 2009, Mendoza, Purchal et al. 2023, Watkins, Wang et al. 2023). Figure 5.12 illustrates the conservation analysis of the residues in the CblBD, displaying high

conservation on residues surrounding Cbl. The residues still in the CblBD but further from the Cbl-binding pocket display less conservation in the analysis performed by ConSurf.

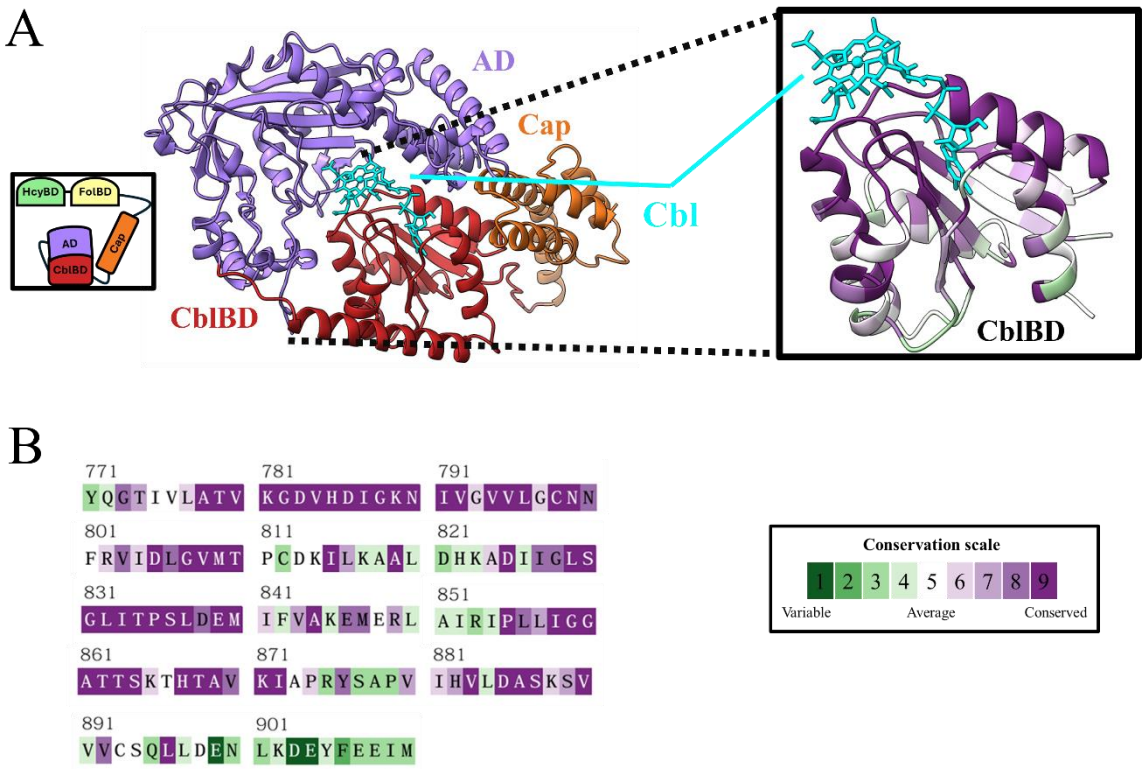

Figure 5.12: Analysis of the CblBD conservation. A. C-terminal half of MTR in the His-off state from the MTR+Ligands dataset, showing the position of the Cap (orange), CblBD (red), AD (purple), and Cbl (cyan). The inset displays the conservation of residues within the CblBD, coloured according to a conservation scale. **B**. Conservation analysis of the residues in the CblBD as shown in A.

The His-off state of this structure is confirmed by the position of His785 relative to Cbl (Figure 5.13). The histidine is oriented away from the Cbl molecule, at a distance of 6.1 Å from the Co atom, indicating that they are not coordinating. This state likely represents cob[II]alamin before the reductive methylation that generates MetCbl (State 'C' in figure 1.7, as it must be in the His-off state to undergo remethylation.

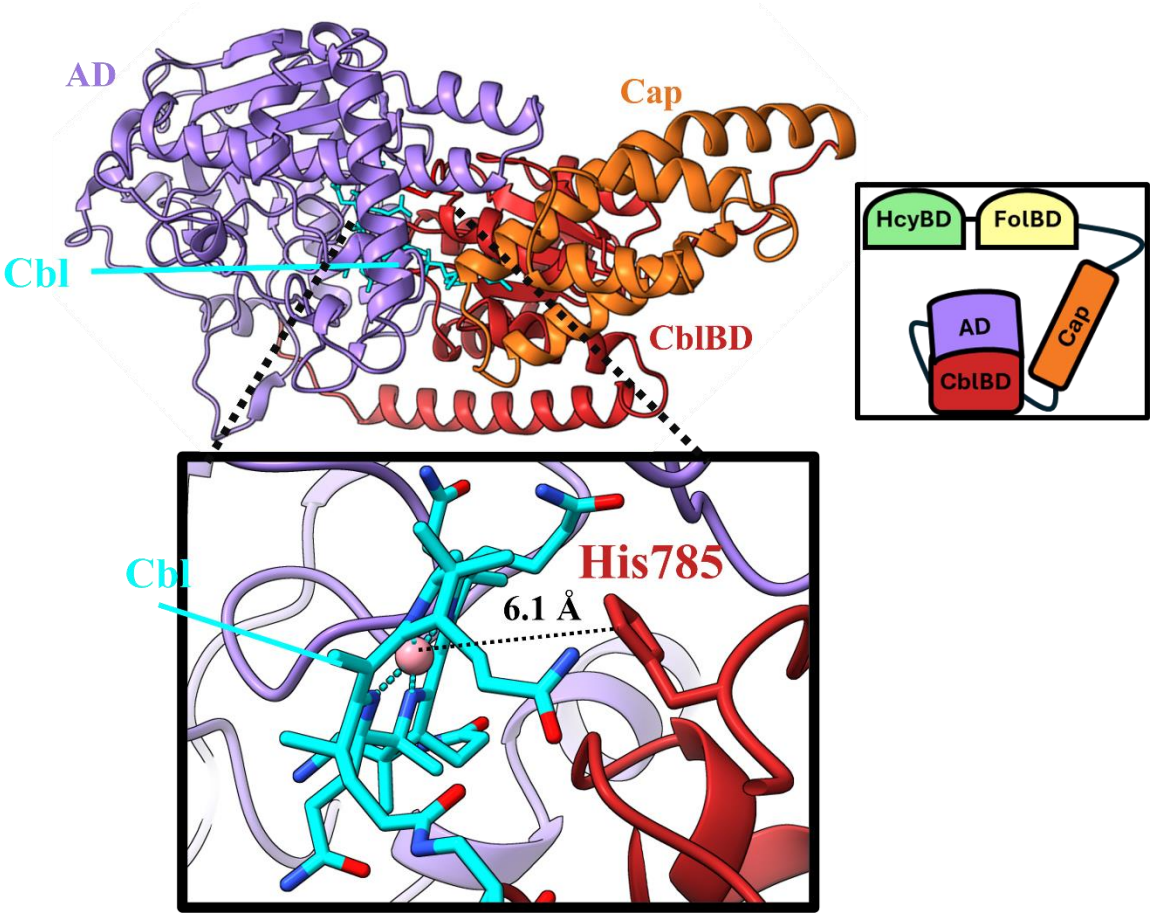

Figure 5.13: C-half of human MTR from the MTR+Ligands dataset in the His-off state. The structure displays the molecular model of the Cap (orange), CblBD (red), and AD (purple), with a zoomed-in area of His785 and Cbl, highlighting that the distance between them is 6.1 Å. Furthermore, a schematic diagram displaying the position of the domains in the activation conformation.

5.3.2.2 Structure of MTR C-half in the His-on state

The His-on, C-half model from MTR+Ligands was constructed similarly to the His-off state, based on a 3.01 Å resolution map (Figure 5.14)The ligand Cbl was also positioned and modelled based on the density observed in the map. And similar to the His-off C-half map, no density of SAM was observed.

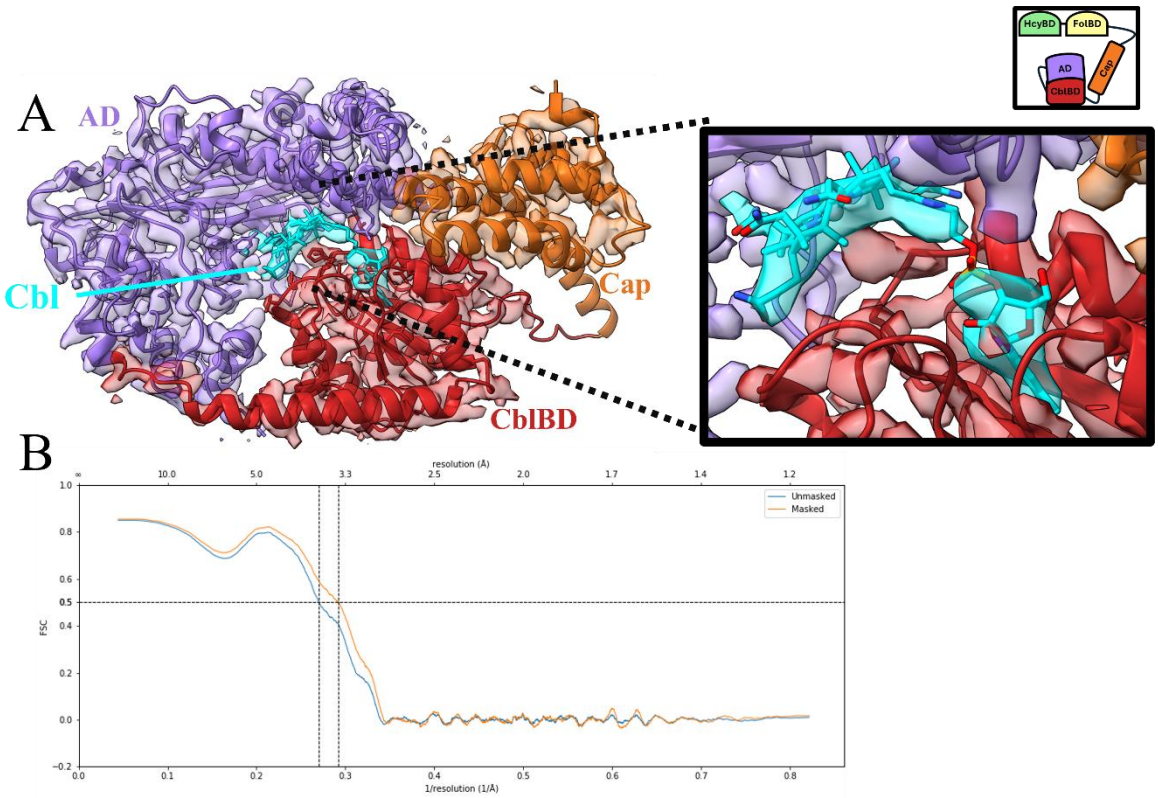

Figure 5.14: Superimposition of the cryo-EM map and the atomic model of the C-half of His-on MTR from MTR+Ligands dataset. A. The model (His-on state) highlights the Cap (orange), CblBD (red), the AD (purple), and Cbl (cyan). The cryo-EM map reveals clear density for secondary structure elements and the overall shape of MTR in the activation conformation, with the CblBD and AD interacting and the Cbl positioned on its binding pocket, and no clear density for SAM. The inset highlights the Cbl density superimposed with Cbl molecule. **B**. The unmasked and masked FSC curves for the atomic model, constructed using ISOLDE (3.4 Å resolution) and further refined with Phenix.

The C-half His-on MTR shown here is also in the activation conformation, with the AD positioned close to the CblBD, and the Cap in the Cap-off state.The Cap domain here followed the same behaviour observed in all other C-half structures, where its density is not well-defined, with some regions lacking clear definition, confirming flexibility of the Cap needed for its transition.

With structures now available for both the His-on and His-off states, as well as the *apo* C-terminal half of human MTR, the three structures were overlaid to examine the position of the conserved histidine. This analysis is crucial for understanding the role of histidine, which has been shown to coordinate with the lower axial position of the cobalt

atom and is essential for the reactivation cycle in bacterial MTR (Bandarian, Pattridge et al. 2002, Koutmos, Datta et al. 2009). While this function has been well-characterised in bacterial MTR, no structural data is currently available for the human protein to corroborate these findings.

The overlap of the three models for C-half (Figure 5.15) revealed that both the *apo* and the Cbl-bound His-off MTR structures share a similar positioning of His785, with the histidine oriented away from the lower axial position of Cbl at a distance of 6.7 Å from the Co ion (in agreement with the apo MTR structure of *T. thermophilus* (Mendoza, Purchal et al. 2023)). In contrast, in the His-on conformation, His785 is displaced ~4.6 Å from its position in the His-off state, now at a distance of ~3.5 Å from the Co ion of the Cbl cofactor. This shift in the histidine's position is primarily due to a movement in the Loop2, in which it is located.

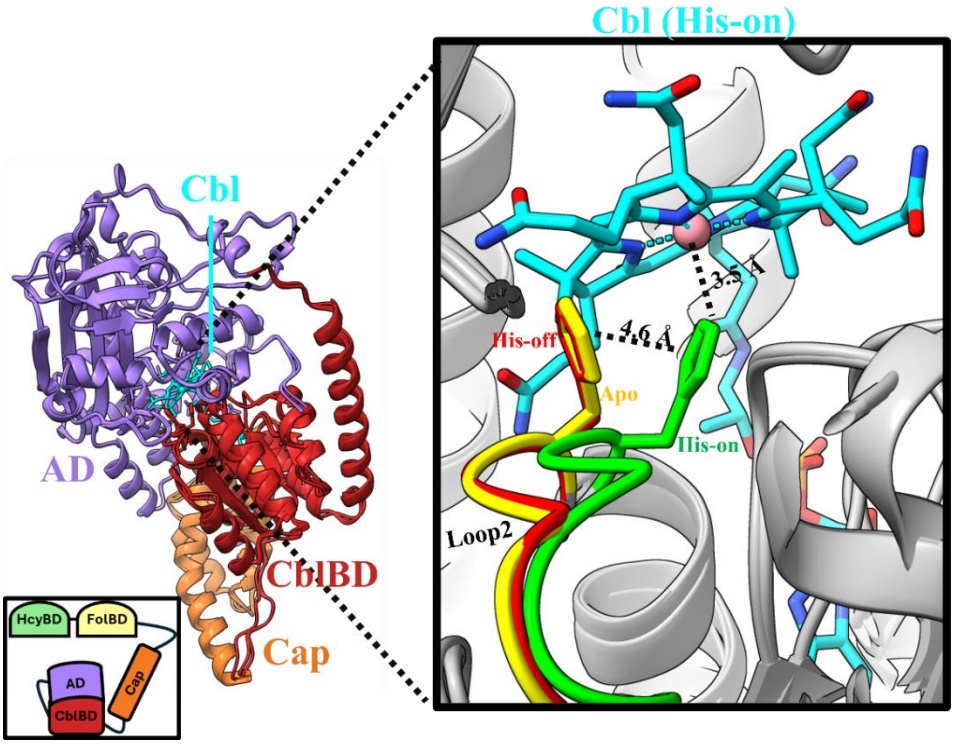

Figure 5.15: Superimposition of the *apo*, His-on, and His-off C-half models of human MTR. The C-half is in the activation conformation, displaying the Cap (orange), CblBD (red), AD (purple), and Cbl ligand (cyan). This view is related from that of Figure 5.14, by a rotation of 90° along X axis. The zoomed area shows the different positions of His785 in the His-off (red), *apo* (yellow), and His-on (green) structures, also indicating the distances of His785 in the His-on structure from the Cbl (3.5 Å) and from His785 in the His-off structure (4.6 Å). The Cbl molecule in the zoomed-in structure is from the His-on structure.

The conformational changes of Tyr1177 were also compared across the *apo*, His-on, and His-off structures of the C-half of human MTR. This tyrosine residue has been shown to participate in the reductive methylation of Cbl by lengthening the Co-O bond at the upper axial position of cob[II]alamin (Section 1.4.3). This elongation in the His-off Cbl causes it to become 4-coordinate, making it more prone to receive the electron needed to reduce it to cob[I]alamin, which is then remethylated by SAM (Koutmos, Datta et al. 2009).

The assessment of the Tyr1177 positions in the three structures was performed similarly to the analysis of His785. The three C-half structures were superimposed, and the position of Tyr1177 was analysed (Figure 5.16). There is virtually no difference in the position of the tyrosine across the three structures. When comparing the region of the AD surrounding Tyr1177 side-by-side with the Cbl bound in the His-on and His-off structures (Figure 5.16 inset), it is evident that, although the tyrosine side-chain has not changed its position, the His-off state shifts the Cbl closer to Tyr1177 by a distance of 2.8 Å. This shift likely causes the elongation of the Co-O bond, as required for the reductive methylation process, since the Tyr1177 is already positioned closer to the cofactor.

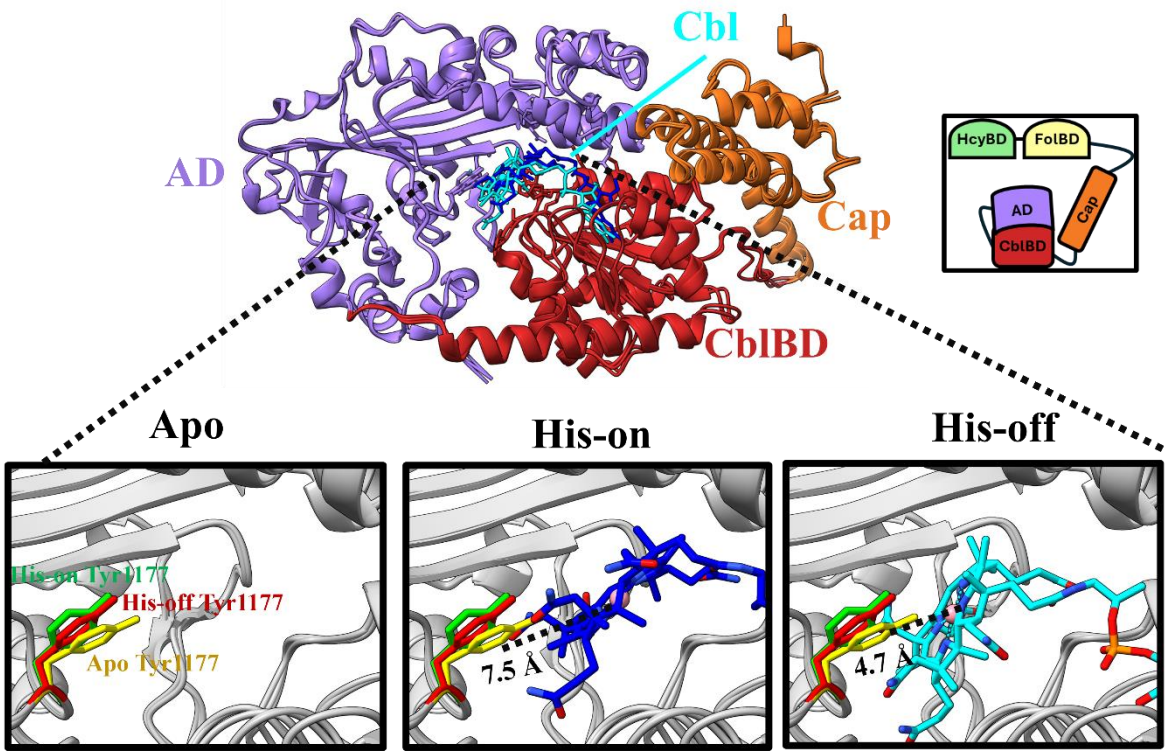

Figure 5.16: Superimposition of the *apo*, His-on, and His-off C-half structures of human MTR. The C-half is in the activation conformation, displaying the Cap (orange), CblBD (red), AD (purple), and Cbl ligand (cyan or dark blue). Inset: The zoomed areas show the different positions of Tyr1177 in the His-off (red), *apo* (yellow), and His-on (green) structures. The zoomed-in areas also represent the absence of Cbl (left panel), the Cbl bound in the His-on structure (dark blue, middle panel), and the Cbl bound in the His-off structure (cyan, right panel), showing the distance between the Cbl and its respective Tyr1177 from the same structure.

Since all three C-half structures display a similar position of Tyr1177, the structure of the His-on MTR was chosen to be aligned with two structures of *T. thermophilus* MTR (*apo* (PDB: 8SSD) and His-off (PDB: 8SSE) C-half) (Figure 5.17). This comparison between the human and bacterial MTR tyrosine aims to understand the conformational changes that the bacterial tyrosine undergoes and to compare them to the static position of this same residue in the human protein.

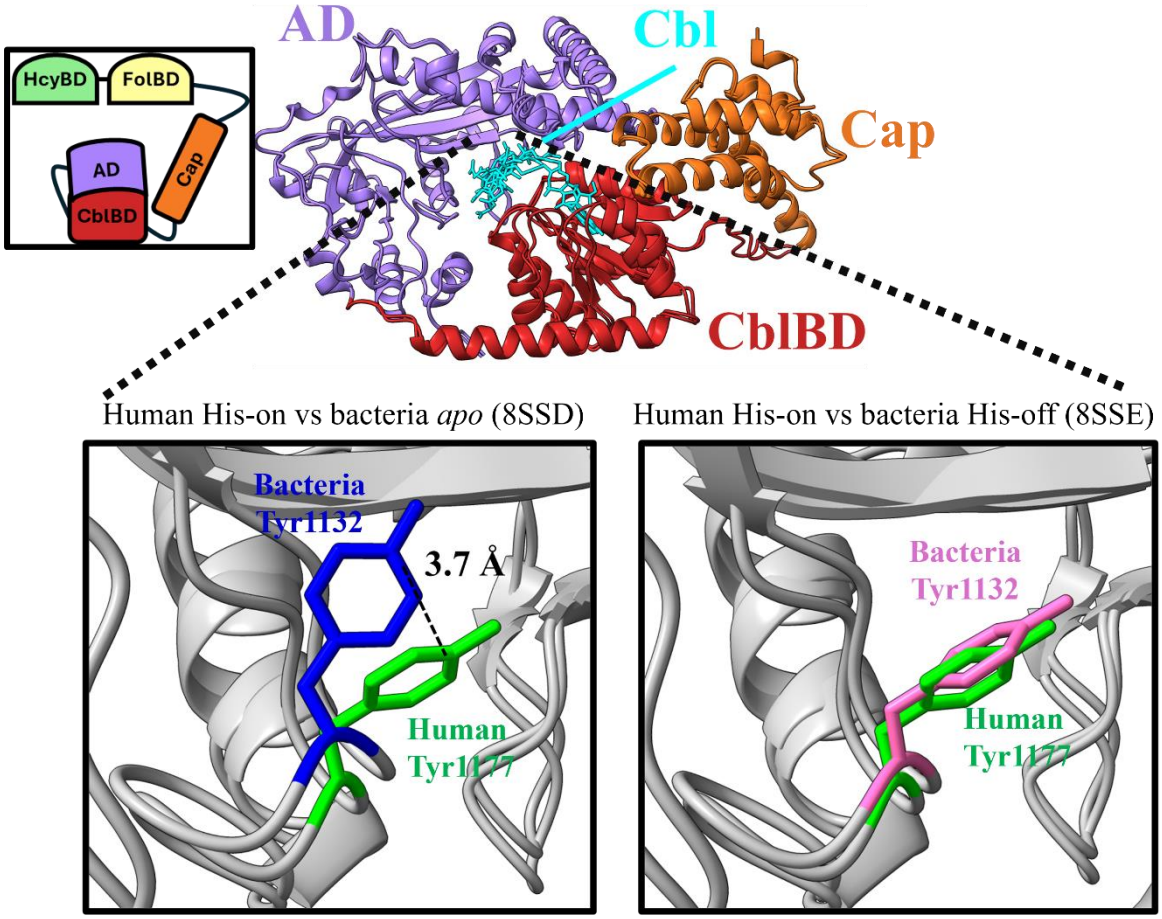

Figure 5.17: Comparison of Human His-on MTR and *T. thermophilus* MTR in *apo* and His-off states. The C-half is in the activation conformation, displaying the Cap (orange), CblBD (red), AD (purple), and Cbl ligand (cyan). Inset: The superimposition of the *apo* and His-off structures of *T. thermophilus* MTR (PDB: 8SSD for *apo*, and 8SSE for His-off) with the His-on state of human MTR focuses on the conserved tyrosine residue (Tyr1177 in human, Tyr1132 in *T. thermophilus*). In the bacterial *apo* structure, the tyrosine (blue) is 3.7 Å apart from the His-on tyrosine (green), which is positioned closer to Cbl-binding pocket. In the His-off bacterial structure, the tyrosine (pink) is positioned in the same space as in human His-on (green), as well as in human His-off and *apo* structures.

The superimposed structures (Figure 5.17) show that Tyr1177 in human MTR (whether in the *apo*, His-off, or His-on states) occupies a position similar to that of the His-off structure in bacterial MTR. In bacterial MTR's *apo* state, the tyrosine side-chain is rotated and positioned away by ~3.7 Å from the region where Cbl binds. The movement of tyrosine towards Cbl is the final step before cob[II]alamin receives an electron to become cob[I]alamin in *E. coli* (Koutmos, Datta et al. 2009) (State 'D' in figure 1.7). In human MTR, there is no difference in the Tyr1177 positioning between the *apo*, His-on, or His-off states, suggesting that this residue might not undergo any movement in this context. Furthermore, the trigger for the tyrosine to swing closer to Cbl in the case of bacterial MTR seems to be

the presence of SAM (Koutmos, Datta et al. 2009). In the case of human MTR, the position of Tyr1177 is likely to not change according to any ligand binding.

Although Tyr1177 did not show any positional differences in all three human structures compared (*apo*, His-on, and His-off), there is an apparent difference between the human His-on and His-off structures in the positioning of Cbl, as observed in Figure 5.16. To further understand the movement observed for Cbl, the human Cbl-bound His-on and His-off structures were superimposed (Figure 5.18).

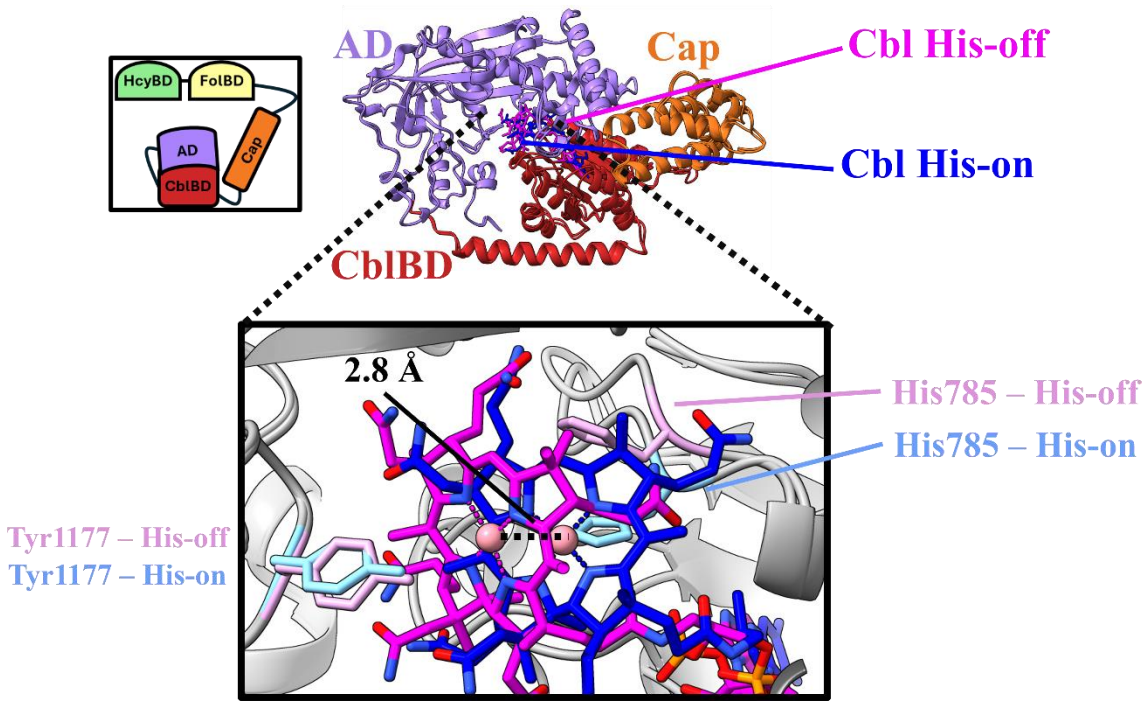

Figure 5.18: Comparison of Cbl position Between His-on and His-off states of human MTR. The superimposed structures of the His-on and His-off C-halves showing the Cap (orange), CblBD (red), and AD (purple), highlight the positions of Tyr1177 and His785 in each model, with light blue representing the His-on residues and light pink representing the His-off residues. The Cbl ligand from the His-off structure (pink) is positioned 2.8 Å closer to Tyr1177 compared to the Cbl ligand from the His-on structure (blue).

In the overlayed structures, the Cbl molecules from His-on and His-off states are 2.8 Å apart from each other. There is a shift in the Cbl ligand towards Tyr1177 in the His-off state, relative to that in the His-on state, likely caused by the dissociation of the Cbl with His785. This shift is most likely responsible for lengthening the Co-O bond at the upper axis

of the Co ion, which is necessary before the reduction of cob[II]alamin to cob[I]alamin (State 'D' in Figure 1.7). In other words, in the His-on conformation, the Cbl ligand is displaced away from the AD, compared to the His-off conformation.

The two distinct positions observed in Cbl between the His-on and His-off states likely represents the different steps in the process of reductive methylation of cob[II]alamin (Section 1.4.3). To gain a broader global understanding of the Cbl movement with respect to the protein, the entire C-half structures of human His-on and His-off MTR were compared (Figure 5.19), as the movement of the Cbl molecule is a result of shifts within the domain where it is located.

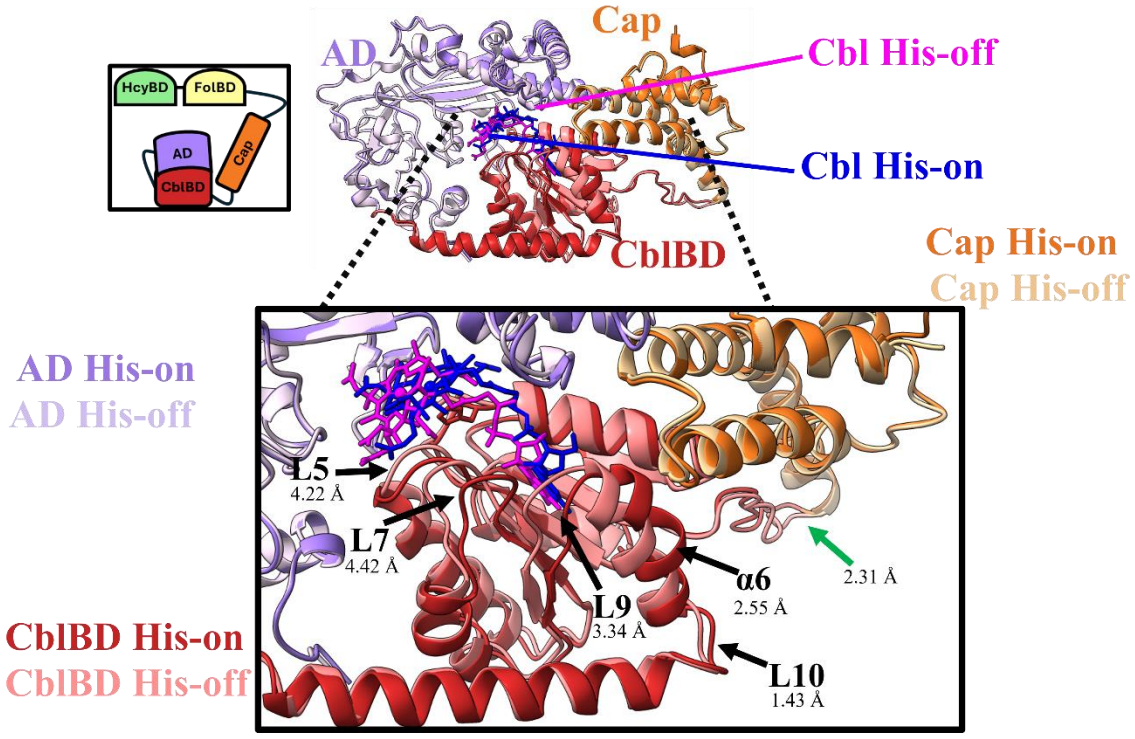

Figure 5.19: Superimposition of the His-on and His-off structures of human MTR. The Cap (orange), CblBD (red), and AD (purple) domains are shown for both His-on and His-off states, with the domains of the His-off structure displayed in pastel colours. The Cbl ligand is represented in blue for the His-on state and pink for the His-off state. The inset highlights that in the His-on structure, Cbl is positioned further from the AD, likely indicating a pre-resting conformation. Arrows indicate regions with significant conformational changes between the two structures, along with their corresponding RMSD values. The green arrow points to the loop connecting the Cap and CblBD.

The clear conformational changes in the MTR structure between the His-on and His-off states are primarily observed in the CblBD, as illustrated in Figure 5.19 arrows. In the His-off state, where the Cbl is poised for remethylation, the CblBD is positioned closer to the AD, facilitating the interaction of Cbl with Tyr1177 and SAM, which will donate the methyl group to Cbl. In turn, cob[III]alamin then adopts the His-on state, when the CblBD is positioned further from the AD, which also displaces the Cbl molecule. The observed changes in various loops and helices such as Loop5, Loop7, Loop9, Loop10, and α6 within the CblBD indicate the beginning of the disengagement between the CblBD and AD, which is required when MTR enters the resting state in order to start the catalytic cycle.

5.4 Discussion

The primary insights gained from the datasets will be discussed below, focusing on the conformation of *apo* MTR in a pre-Cbl loading state (Section 5.4.1), the conformational changes induced by MetCbl (Section 5.4.2), and the reactivation process of Cbl in human MTR (Section 5.4.3), which until now has only been understood in the context of its bacterial orthologue.

5.4.1 Pre-cobalamin conformation: conformational insights from apo MTR

The preparation of the grid containing only *apo* MTR required the purification of the protein as outlined in Chapter 3. This step was crucial because most existing purification protocols for MTR involve the addition of Cbl, which is utilised for protein immobilisation (Yamada, Gravel et al. 2006, Wolthers and Scrutton 2009, Zhang, Liu et al. 2020). The full-length human MTR used for this grid is devoid of Cbl or any other ligands, thereby allowing a broader range of ligand combinations to be studied, including the *apo* form of the protein, which is relevant because the conformational changes of MTR are influenced by the

oxidation state of Cbl and the binding of various substrates (Jarrett, Huang et al. 1998). In this study, the structure of the protein is examined in its bound and unbound form, the latter representing the protein before its interaction with other proteins involved in the Cbl delivery.

The data collection and analysis of the grid containing *apo* MTR (Section 4.4.2) resulted in 2D classes indicating just one half of the protein. However, due to the grid quality and the number of particles found, the dataset was insufficient for the structural reconstruction of these particles. Despite further attempts to collect a better dataset, these efforts were unsuccessful due grid quality. Nevertheless, the 2D classes generated from the initial attempt provided sufficient view into the conformation of human MTR in the *apo* state.

The 2D classes from the *apo* MTR dataset (Figure 4.6) revealed the C-half of the *apo* proteinin the activation conformation, where the AD and CblBD are positioned close to each other. No classes of the N-half domains (HcyBD and FolBD) were seen in this dataset likely due low number of particles found. The activation conformation in the C-half MTR has been described during the remethylation of Cbl in different bacterial MTR and as the *apo* state of the *T. thermophilus* protein (Bandarian, Pattridge et al. 2002, Datta, Koutmos et al. 2008, Koutmos, Datta et al. 2009, Mendoza, Purchal et al. 2023).

No particles in the dataset represented the entire protein, although biochemical assays described in Chapter 3 suggested that the protein is not cleaved (Figure 3.6). The behaviour of MTR as two independent units (N- and C-halves) is likely caused by the linker located between FolBD and Cap (Linker 1) (aa Gln653-Asp664). Linker1 (QGTGGKKVIQTD) contains three glycine residues, which are commonly found in loops, turns, and flexible regions of proteins due to their small side chains that promote increased backbone mobility.

This high flexibility likely allows each half of the protein to adopt different spatial positions, making it difficult to find enough particles where these two halves are in the same relative position to generate 2D classes. The rigid structure of HcyBD and FolBD, combined with the high flexibility of Linker 1, composed of 12 residues, suggests the large movements that CblBD undergoes between these two domains during the catalytic cycle (Evans, Huddler et al. 2004).

In the recent full-length MTR structure from *T. thermophilus* (PDB: 8SSC), Linker 1 was also identified as highly flexible, making it impossible to reconstruct its residues in the molecular model (Mendoza, Purchal et al. 2023). Altogether, although no 2D classes corresponding to the full-length protein were identified in this work, it is inferred that the protein is not cleaved, but the flexibility of Linker 1 is responsible for MTR behaving as two units in the grid.

Although the dataset from the *apo* MTR did not produce a high-resolution map, the grid containing MTR+MeTHF generated two density maps, each corresponding to one half of the protein (Section 4.4.2). The C-half of the MTR+MeTHF structure does not have any ligands bound and is positioned in the activation conformation, suggesting that the positioning and interaction of domains are similar to *apo* MTR. To analyze the inter-domain interactions that might stabilize this conformation, the model built from the C-terminal half of MTR+MeTHF (Section 5.3.1) was utilized.

The interaction analysis between the CblBD and the AD revealed at least five interactions (Figure 5.20), between Lys871, Gly782, Lys781, Asp786, and Asn790 from the CblBD with Asp1265, Thr1204, Glu1205, Lys1127, and Glu1115 from the AD. The residues involved in these interactions are not directly associated with the Cbl-binding pocket and are

likely to dissociate from the AD during the conformational change to the resting conformation, which is characterised by the movement of the Cap from Cap-off to Cap-on.

Also in the structure from the MTR+MeTHF dataset, one interaction was found between Lys741 of the Cap and Asp1120 of the AD, which possibly contributes to maintaining the activation conformation. Additionally, two residues from the CblBD, Asn790 and Glu768, interact with Lys717 and Arg703 from the Cap, further supporting the activation conformation, which is independent of the presence of Cbl. After the remethylation of Cbl, the movements in the CblBD are expected to disrupt most of these interactions, leading to the dissociation of the AD from the CblBD and allowing the Cap to swing over the Cbl, establishing new interactions with the CblBD (Drennan, Huang et al. 1994).

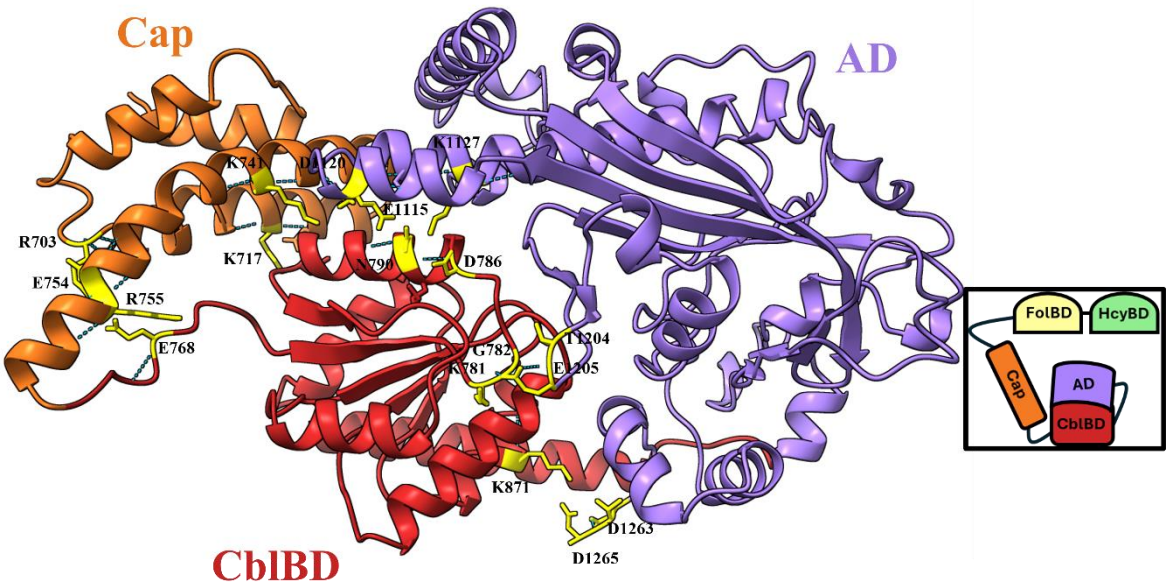

Figure 5.20: Inter-domain Interaction within the *apo* C-terminal Half of MTR. The figure illustrates the residues involved in inter-domain interactions, highlighted in yellow, between the Cap (orange), CblBD (red), and AD (purple). These interactions play a crucial role in maintaining the activation conformation even in the absence of any ligands. The schematic cartoon accompanying the figure represents this activation conformation.

In this conformation, human MTR interacts with the FMN-binding domain (FMNBD) of MTRR, which transfers an electron to Cbl to facilitate its remethylation

(Wolthers and Scrutton 2007). For FMNBD to transfer the electron to Cbl, it needs to be positioned close to the Cbl-binding pocket. Thus, in this conformation, Cbl is exposed to solvent, making it available to interact with FMNBD (Section 1.4.3 State 'D' in figure 1.7).

The *apo* MTR in the activation conformation has the empty Cbl-binding pocket available for interaction, enabling Cbl delivery by MMACHC-MMADHC. This is the likely conformation MTR adopts to facilitate the interaction required for Cbl delivery. The interaction interface of MTR with MMADHC is probably the same as that of MTR with FMNBD for Cbl reactivation, as MTR assumes the same conformation for both interactions, occurring in the same region (Cbl-binding pocket) (Figure 5.21).

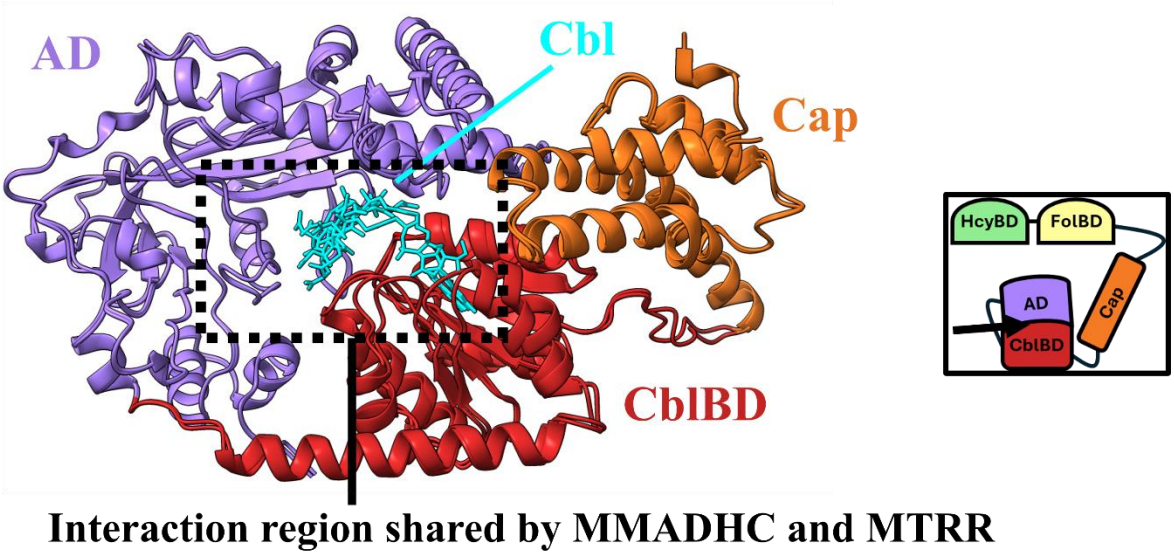

Figure 5.21: Interaction region on MTR. The C-half of MTR in the activation conformation is showed, with the dashed area indicating the interaction region shared by MMADHC and the FMN-binding domain of MTRR. The cartoon on the right illustrates the activation conformation of MTR, with the same interaction region highlighted by a black arrow.

The putative shared interaction region of MTR with MTRR and MMADHC could indicate that these proteins are unlikely to form a higher-order protein complex, as suggested by *in vivo* studies (Bassila, Ghemrawi et al. 2017), unless MTRR or MMADHC has an

alternative interaction region with MTR, which allows the simultaneously interaction. This aspect is explored in Chapter 6.

Furthermore, understanding the conformation of MTR during its interaction with MMACHC-MMADHC for Cbl acquisition can provide insights into the structural dynamics of MTR. Thus far, the interaction between AD and CblBD has only been described for the reactivation of Cbl in human MTR (Koutmos, Datta et al. 2009). According to the *apo* human MTR dataset, the activation conformation also represents a pre-cobalamin loading conformation (Section 1.4.3 State 'D' in figure 1.7), positioning MTR to interact not only with MTRR for reactivation but also with MMACHC-MMADHC for Cbl delivery.

The cryo-EM determination of the *apo* human MTR conformation paves the way for studying the Cbl loading mechanism onto MTR. This represents the first view into how human MTR is positioned in the absence of a cofactor or any other ligand.

5.4.2 Conformational changes induced by MetCbl: from activation conformation to resting conformation

The multi-modular nature and inter-domain flexibility of MTR allows it to possibly exist in at least four different conformations (Figure 5.22). Two of these conformations occur during the catalytic cycle (Figure 5.22A and B), the third conformation is the activation conformation (Figure 5.22D), and the final one is the resting conformation (Figure 5.22C), also called as Cap-on state (Bandarian, Ludwig et al. 2003).

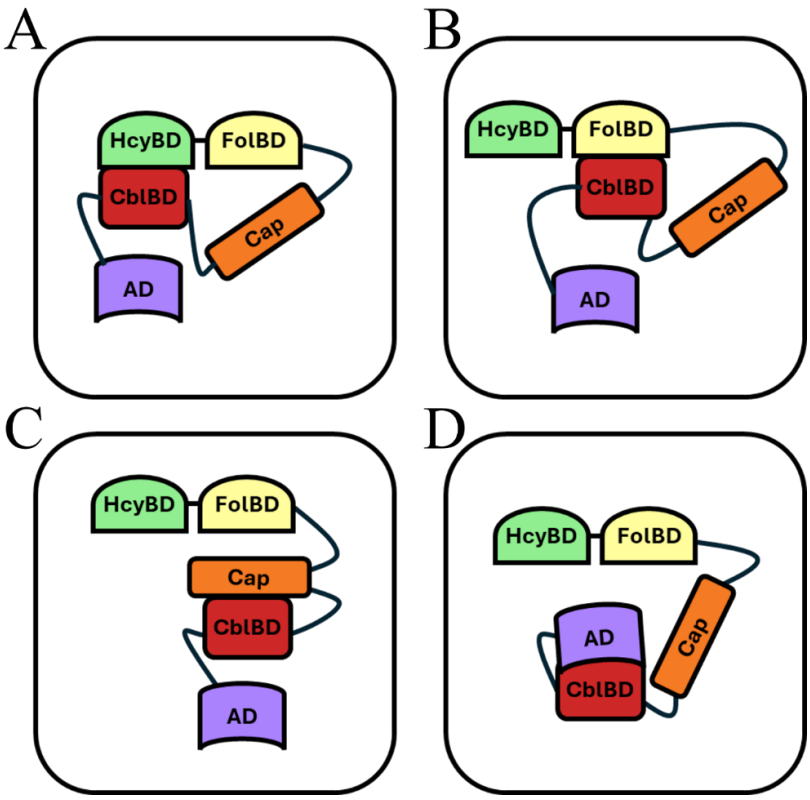

Figure 5.22: Different conformations of MTR based on CblBD interactions. A. CblBD interaction with the HcyBD for the methylation of Hcy to Met. **B**. Interaction of CblBD with FolBD for the remethylation of Cbl to MetCbl, converting MeTHF to THF. **C**. Cap-on state, demonstrating the interaction between CblBD and the Cap domain. **D**. Interaction of CblBD with the AD, facilitating Cbl remethylation; this is the same conformation observed for apo MTR. The conformations shown in A and B represent the catalytic cycle, C corresponds to the resting conformation, and D depicts the activation conformation.

The data collected from the MTR+MetCbl grid was insufficient to generate high-resolution density maps, despite the large number of micrographs collected (5,000 micrographs). Nevertheless, the particles found were sufficient to generate 2D classes that identified different part of MTR. The first part identified in the MTR+MetCbl dataset was the N-half, which was expected, as it behaves as a single unit and does not exhibit any flexibility between its domains.

Unexpectedly, the second part identified was the isolated AD. All particles representing this domain were confirmed by comparing the 2D classes (Figure 4.18) and generating a low-resolution volume that fits with the available human crystal structure of

this domain (PDB: 2O2K) (Figure 4.19). Apart from the AD, no other domains of the C-half were found, and no particles representing the activation conformation were identified in the dataset, despite their predominance over N-half particles in all other grids analysed.

The absence of C-half particles in the activation conformation, combined with the presence of the N-half and AD particles in the sample, suggests that upon the addition of MetCbl to full-length human MTR, the protein undergoes a conformational change from the activation conformation to the resting conformation. The third part of MTR, comprising the Cap and CblBD is likely to be in the Cap-on configuration, although it was not found.

The resting conformation (Figure 5.22D), also referred to as the Cap-on state, is characterised by the movement of the four-helix bundle Cap domain to the CblBD, protecting the Cbl from oxidation and photolysis (Jarrett, Drennan et al. 1996). There are only two structures available for this state, from *E. coli* and *T. filiformis* (PDB: 1BMT and 8G3H, respectively). The resting conformation is an intermediate state between the catalytic and reactivation cycles. It is suggested that when Cbl is oxidised during the catalytic cycle, MTR may transition to the resting conformation before entering the reactivation cycle. Additionally, it is proposed that after Cbl is remethylated in the reactivation cycle, MTR returns to the resting conformation before re-entering the catalytic cycle (Watkins, Wang et al. 2023).

There are two possible explanations for the lack of classes showing the interaction between the Cap and CblBD in the grid of MTR+MetCbl. The first is the small size of the Cap-on unit, which is only 30 kDa and may be difficult to identify in cryo-EM, potentially requiring an even larger dataset for detection. The second is the flexibility of Linkers 1 and 2, which flank the Cap and CblBD.

The flexibility of Linker 1 is consistent with the other cryo-EM data presented in this present chapter, where it causes the protein to appear as two separate units in the activation conformation. Linker 2, besides undergoing proteolysis in the *E. coli* Cap-on structure (Drennan, Huang et al. 1994), also exhibited high flexibility in the *T. filiformis* Cap-on structure, where the isolated AD appears as a shadow in the 2D classes (Watkins, Wang et al. 2023). The flexibility of both linkers can cause the human protein to be observed as three separate sections in cryo-EM when in the resting conformation (Figure 5.23). The absence of particles containing simultaneously the HcyBD, FolBD, Cap, and CblBD further supports the hypothesis that the resting conformation observed in *T. filiformis (*Figure 1.14) does not translate directly to human MTR.

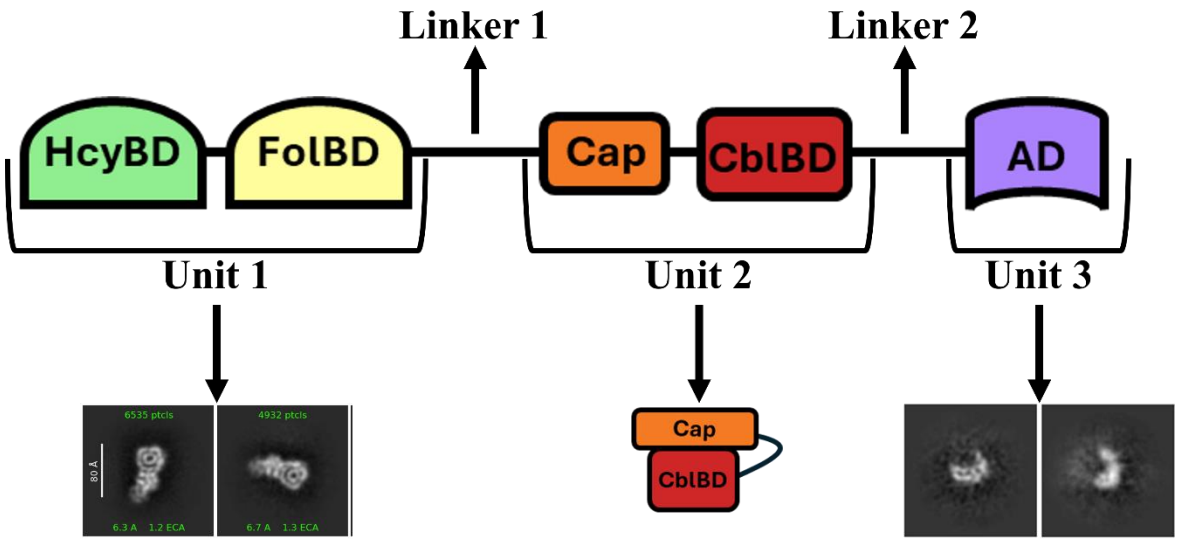

Figure 5.23: Different units of MTR in the resting conformation and their representation in cryo-EM upon the addition of MetCbl. Unit 1 (HcyBD and FolBD) is observed in the 2D classes as a round domain followed by an elongated one, exhibiting no flexibility between these domains. Unit 2, the Cap-on configuration, composed of the Cap and CblBD, was not detected in the dataset but is expected to illustrate the interaction between these two domains. Unit 3 is characterised by the isolated AD. The figure also identifies the locations of Linkers 1 and 2, which are responsible for dividing MTR into these distinct units.

The trigger for the conformational change from the activation conformation of MTR to the resting conformation, if any, has not been identified. It is known that during the remethylation of Cbl, the His-off 4-coordinate cob[II]alamin receives an electron from

MTRR and is then remethylated by SAM, generating MetCbl, which transitions from the His-off to the His-on state before a conformational change (Koutmos, Datta et al. 2009).The cryo-EM data (section 4.2.2) revealed that the addition of MetCbl alone to *apo* human MTR is sufficient to induce the conformational change to the resting conformation. This suggests that the trigger may be related to the change in the Cbl upper ligand to a methyl group, coupled with the His-on configuration of the cofactor, which has been shown to cause a movement of the CblBD away from the AD in the His-on C-half from MTR+Ligands (Section 4.5.4.2) (Koutmos, Datta et al. 2009).

The results herein suggest that upon MetCbl binding, MTR can undergo a conformational change from the activation state (*apo*) to the resting conformation, likely due to movements in the CblBD caused by the change in the cofactor's upper ligand to the methyl group and the His-on state. Furthermore, this data suggests that the resting conformation of MTR may be characterised by the division of MTR by three parts, separated by two flexible linkers (Figure 5.23), different from the two-part protein observed in the activation conformation. The first unit is the N-half of MTR, composed of the HcyBD and FolBD, as observed in crystal structures (PDB: 4CCZ). The second unit is the Cap and CblBD in the Cap-on state, similar to the conformation found in E. coli (PDB: 1BMT), and the third is the isolated AD, which dissociates from the CblBD during the conformational change, and has the same overall configuration as the crystal structure (PDB: 2O2K). These units exhibit high flexibility relative to each other, which is part of the modular nature of MTR. This flexibility could result in proteolysis or, as observed here, be responsible for the protein not being fully observed in the cryo-EM 2D classes

5.4.3 The C-half trio: domains structure and ligand-driven insights

Large dataset collections from the MTR+MeTHF and MTR+Ligands grids at eBIC yielded high-quality data and the generation of five density maps with approximately 3 Å

resolution. These maps were used to construct five models by rigid-fitting the human MTR AlphaFold model. Two MeTHF-bound N-half MTR models from each of the dataset and three models of the C-half, representing the *apo*, His-on, and His-off states in the activation conformation. The quality of the maps for the three C-half structures enabled the generation of detailed models that provide insights into the remethylation process of Cbl during the recovery cycle.

The reactivation of Cbl is a multi-step process, with insights derived from *E. coli* MTR, as no structural data for the human C-half MTR was previously available. In the activation conformation, MTR bound to cob[II]alamin shifts between His-on and His-off states, with most of the protein in the His-on state (Koutmos, Datta et al. 2009). Upon binding of the partner protein flavodoxin (MTRR for human), the population of MTR shifts to the His-off state. When SAM binds to the AD, the conserved tyrosine (Tyr1177 in human MTR) undergoes a conformational change, coordinating with the water molecule at the upper axial of Cbl. This movement lengthens the Co-O bond, making Cbl 4-coordinate and more susceptible to electron transfer from the partner protein, converting it to cob[I]alamin. The high reactivity of cob[I]alamin allows it to accept a methyl group from SAM, generating MetCbl, which then shifts back to the His-on state (Koutmos, Datta et al. 2009). The protein can then undergo conformational changes and transition into the resting conformation.

During reactivation, two residues are essential: the histidine in the CblBD that coordinates the His-on/off states (His785) and the tyrosine in the AD that coordinates with the upper axial of Cbl (Tyr1177). The structural models built for the C-half of human MTR allowed the study of these conserved residues.

The His-on state, where His785 coordinates with the lower axial position of the Co atom, persists throughout the catalytic cycle, only shifting to the His-off state during Cbl

remethylation. Comparing the position of His785 in the *apo*, His-on, and His-off structures (Figure 5.15) reveals that the residue is in the same position in the *apo* and His-off structures, swung away from the Cbl-binding pocket. In the His-on structure, His785 is positioned approximately 4.6 Å closer to Cbl compared to His-off or *apo*, facilitating the coordination.

Recent structural studies of *T. thermophilus* MTR in the *apo* state identified that the histidine is in the same position as the His-off structure (Mendoza, Purchal et al. 2023), similar to what is observed here in human MTR. The position of His785 (human MTR) away from the Cbl-binding pocket in the *apo* structure is advantageous for Cbl loading, as the flipped-out position of the residue opens the cleft for the DMBI tail of Cbl to be accommodated, facilitating its insertion.

While His785 exhibits similar behavior in MTR orthologues across species, the same is not observed for Tyr1177. In MTR, this residue moves to elongate the Co-O bond and facilitate the reduction of cob[II]alamin to cob[I]alamin during the reactivation cycle (Liptak, Datta et al. 2008). In the His-on state of *E. coli* MTR, the conserved tyrosine rotates and shifts further from Cbl, whereas in the His-off state, it is positioned closer to the upper axial position of Cbl, conformational change potentially triggered by SAM binding (Koutmos, Datta et al. 2009). However, the structure of *T. thermophilus* MTR showed that the tyrosine movement is not triggered by SAM, but it moves closer to Cbl in the His-off state than in the *apo* structure (Mendoza, Purchal et al. 2023). The structures obtained for human MTR show that Tyr1177 remains in the same position across all states (*apo*, His-on, and His-off) (Figure 5.16). The rigidity of Tyr1177 in human MTR suggests that, unlike in *E. coli* MTR, tyrosine movement is not triggered by SAM, and, in contrast to the bacterial structures, including *T. thermophilus*, there is no apparent position change in Tyr1177.

During the remethylation of Cbl, coordination with Tyr1177 at the upper axial position of Cbl is essential. The rigidity of Tyr1177 in different states of human MTR indicates that this coordination occurs through a mechanism other than tyrosine movement. Comparing the His-on and His-off structures reveals that the Cbl molecule moves closer to Tyr1177 in the His-off state, which could enable their coordination (Figure 5.18).

The Cbl-bound His-on and His-off structures of *E. coli* MTR (PDB: 3IV9 and 3IVA) show that despite movements in the CblBD, the position of Cbl does not change between states (Koutmos, Datta et al. 2009). In contrast, the human structures of the C-half show a 2.8 Å shift in Cbl away from Tyr1177 from the His-off to the His-on state (Figure 5.18). This movement of Cbl in human MTR, caused by shifts in the CblBD (Figure 5.19), likely drives the elongation of the Co-O bond, enabling the reduction of cob[II]alamin to cob[I]alamin. Additionally, the motion of the entire CblBD away from the AD in the His-on state is more significant in the human enzyme (calculated RMSD of 1.781 Å) than in the bacterial one (calculated RMSD of 1.359 Å) (Figure 5.24) (Koutmos, Datta et al. 2009).

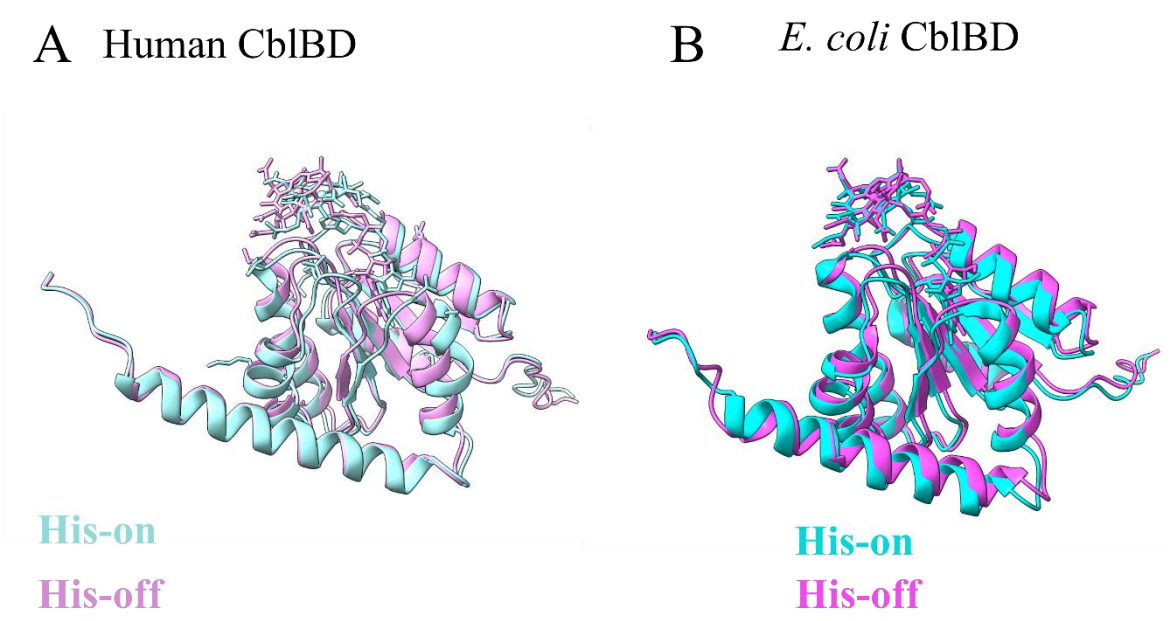

Figure 5.24: Compairson of CblBD movement in human and *E. coli* between the His-on and His-off states. A. The CblBD with Cbl bound in the His-on (light blue) and His-off (light pink) states from human MTR. The comparison of this domain in these two states yields an RMSD of 1.781 Å. **B**. The CblBD with Cbl bound in the His-on (blue) and His-off (pink) states from E. coli MTR. The comparison of this domain in these two states yields an RMSD of 1.359 Å. This comparison indicates that the CblBD of human MTR undergoes more significant conformational changes from the His-off to the His-on state than that of E. coli MTR.

Because of the CblBD movement, an analysis of interactions between Cbl, the CblBD, and the AD were performed to reveal differences between the His-on and His-off states in human MTR (Figure 5.25). In the His-off state, Cbl is positioned closer to the AD, allowing three potential interactions between the cofactor and the AD. Lys1186 and Ala1212 may interact with O39 and N45 of Cbl, and His1183 also could interact with O34 of Cbl and Thr834 of the CblBD. In the His-on state, with Cbl further from the AD, the interactions observed are the interaction between Thr834 of the CblBD and O34 of Cbl and the Ala1212 from the AD with N45 of Cbl (Figure 5.25). These differences, in particular losing interactions of Lys1186 and His1183, suggest that certain AD residues coordinate with Cbl during the His-off state, possibly helping secure the activation conformation. In the His-on state, as MTR transitions to the resting state, these interactions between the AD and Cbl disappear, indicating a dissociation of the AD. The His-on/off state transitions, and thus the

conformational changes in MTR to the resting conformation, are triggered by the oxidation state of Cbl (Fleischhacker and Matthews 2007) and the radical group in its upper ligand, as the OHCbl (cob[III]alamin) in the His-on state did not trigger the resting conformation, whereas MetCbl (also cob[III]alamin) did (Section 4.4.1).

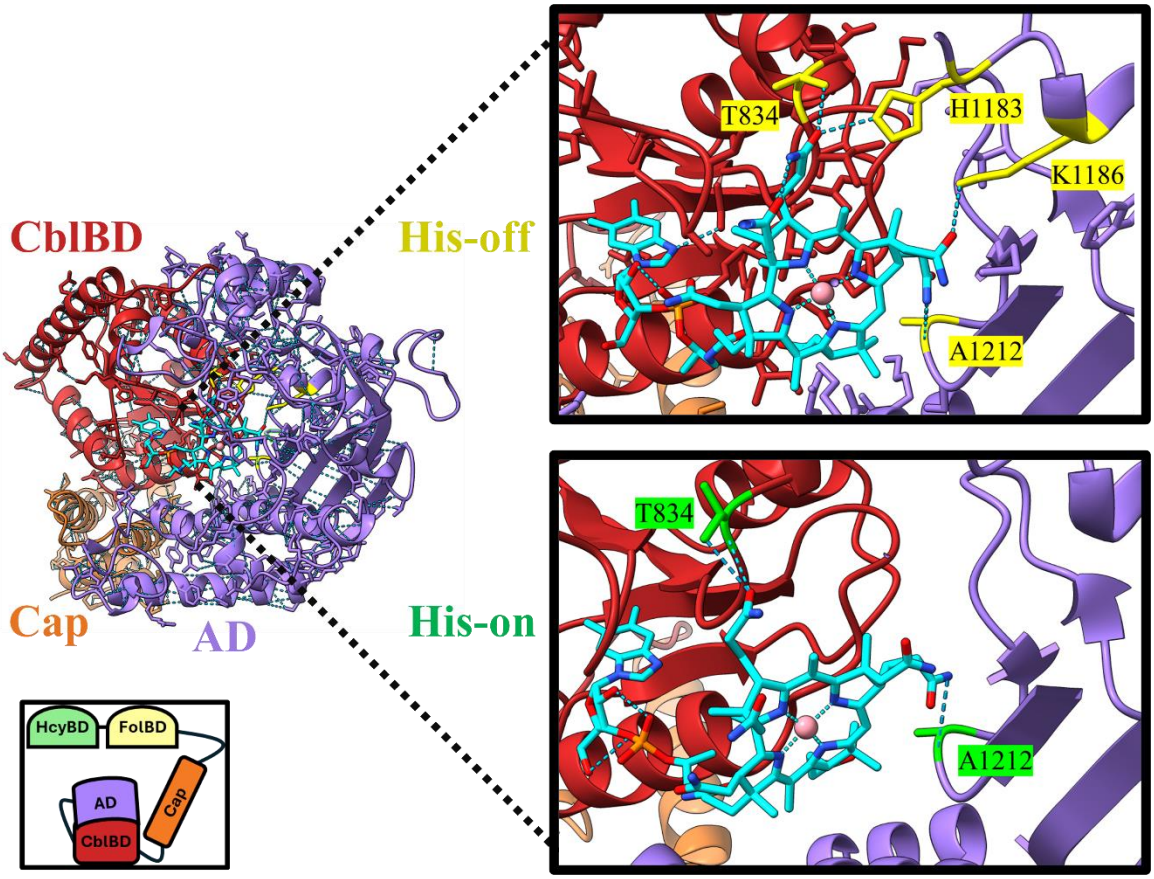

Figure 5.25: Analysis of Cbl interactions with the CblBD and AD in the His-off and His-on state in the human MTR. The figure illustrates the C-half MTR in the activation conformation, highlighting the Cap (orange), CblBD (red), AD (purple), and Cbl (cyan). The schematic cartoon represents this activation conformation. The insets focus on the Cbl molecule: the upper inset shows the interactions between Cbl and the CblBD and AD in the His-off state, with interacting residues highlighted in yellow. The lower inset displays the interactions of Cbl in the His-on state, with interacting residues highlighted in green.

These structural insights, derived from both the His-on and His-off state models, were obtained from a single dataset in which MTR was incubated with MeTHF, SAM, and OHCbl. The His-off state structure described in this chapter likely represents cob[II]alamin positioned for reductive methylation (Figure 5.26, state 'C'). This reductive methylation is not completed due to the lack of reducing agents (*in vivo* reducing agent is provided by the

FMNBD of MTRR). The His-on structure likely represents OHCbl-bound MTR, corresponding to molecules that were not oxidized to cob[II]alamin. Although OHCbl has oxidation state of +3, and binds to MTR in the His-on state (Fleischhacker and Matthews 2007), MTR cannot undergo further conformational changes from the His-on state to the resting conformation due to the absence of a methyl group on the upper axial position of Cbl, which is potentially one of the main triggers for transitioning conformations.

The findings on the structural differences between the His-on and His-off states of MTR suggest that the reactivation cycle of human MTR differs from that of *E. coli* MTR. The proposed reactivation cycle of human MTR is summarised in Figure 5.26.

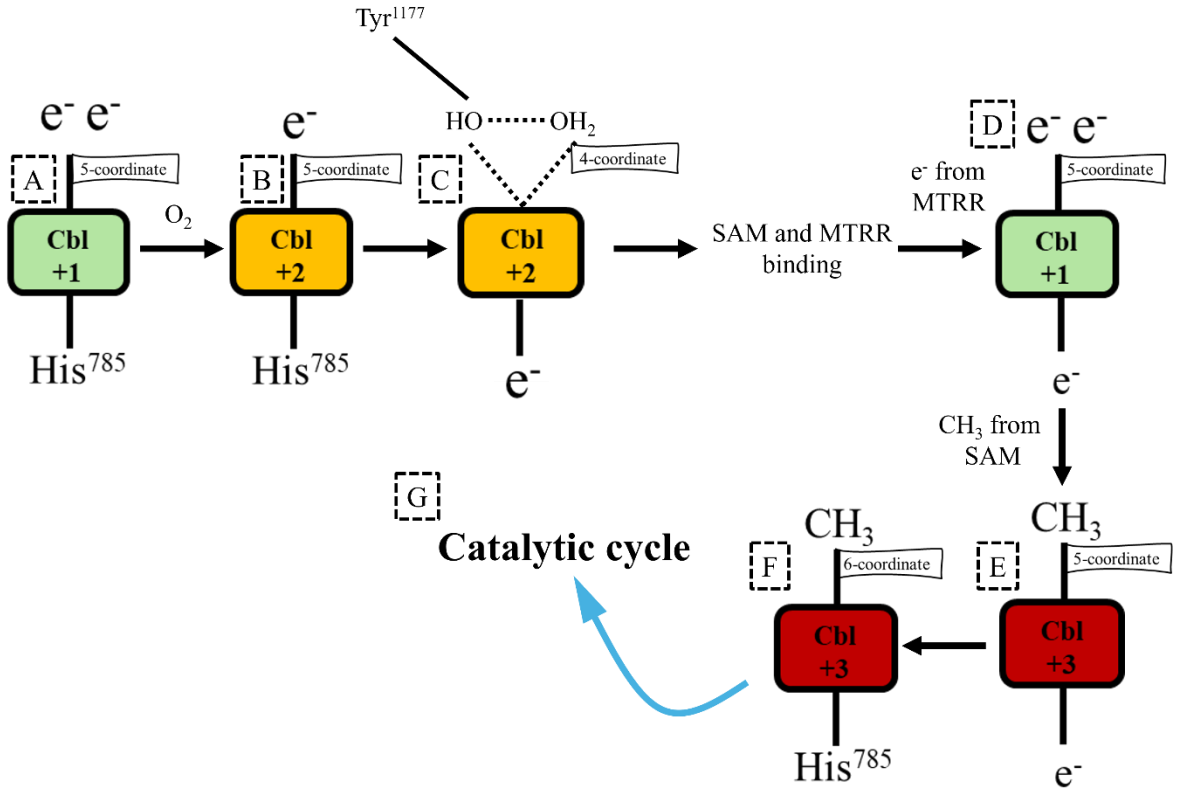

Figure 5.26: When cob[I]alamin is oxidised to cob[II]alamin, MTR undergoes reactivation to regenerate MetCbl. Upon the oxidation of cob[I]alamin (**A**) into cob[II]alamin in the His-on state (**B**), the reactivation cycle of MTR starts. The His-on cob[II]alamin undergoes a state change to the His-off due the oxidation state of Cbl. The His-off state induces a conformational change that shifts the Cbl next to Tyr1177, which will coordinate with the water molecule on Cbl upper axis, elongating it. The elongation will form a stable His-off 4-coordinate cob[II]alamin (**C**), which has a high reduction potential. Upon the binding of MTRR and SAM, the stabilised Cbl can be reduced by MTRR (**D**) and remethylated by SAM, generating MetCbl His-off (**E**). Due the oxidation state of MetCbl, the lower ligand will coordinate with the His785, forming a His-on MetCbl (**F**) and causing movements in the CblBD, desingaging it form the AD and allowing the conformational change to the resting state and catalytic cycle (**G**).

In *E. coli*, the reactivation of Cbl requires SAM binding to move the conserved tyrosine and elongate the Co-O bond (Koutmos, Datta et al. 2009). However, in human MTR, Tyr1177 is already positioned closer to Cbl, and the movement of Cbl caused by conformational changes between the His-on and His-off states is sufficient to convert cob[II]alamin into its 4-coordinate form (Figure 5.26, state 'C'). Although the *T. thermophilus* structure shows that tyrosine positioning is not dependent on SAM binding, a conformational change in the residue is still necessary to coordinate with Cbl (Mendoza, Purchal et al. 2023). The lack of tyrosine movement in human MTR also suggests that the

CblBD alone may be responsible for the conformational changes during Cbl reactivation, as its movements shift Cbl from a 5-coordinate to a 4-coordinate state (Figure 5.26, state 'C'), inducing reductive methylation.

Given the position of His785 and Cbl relative to Tyr1177, the model constructed for the His-off state likely represents the step before cob[II]alamin receives an electron from the FMNBD of MTRR (Figure 5.26, state 'C'), as Cbl is already 4-coordinated. The His-on state likely represents the activation conformation just prior to the transition to the resting state (Figure 5.26, state 'F'), which may not occur due to the absence of a methyl group on Cbl. The movement of the entire CblBD away from the AD in the His-on state suggests that these conformational changes will disengage the AD from CblBD and allow MTR to adopt the resting conformation, as observed in the loose of interdomains interactions in Figure 5.25.

Based on the position of Tyr1177 and movements in the CblBD, the model for the reactivation of Cbl in the human protein (Figure 5.26) may involve fewer steps than in bacterial MTR (Figure 1.7). In *E. coli* MTR, most cob[II]alamin is in the His-on state, with a smaller proportion in the His-off state, requiring the binding of the partner protein to shift to a predominantly His-off state (Hoover, Jarrett et al. 1997) (Figure 1.7, state 'C'). Upon SAM binding, the conserved tyrosine coordinates with the water molecule on the upper axial position of the His-off Cbl, generating the 4-coordinate Cbl (Figure 1.7, state 'D') (Koutmos, Datta et al. 2009). The 4-coordinate cob[II]alamin has a higher reduction potential than the 5-coordinate cob[II]alamin, making it more prone to reduction (Figure 1.7, state 'E') (Mera and Escalante-Semerena 2010). Cbl can then be remethylated by SAM in the AD, generating MetCbl (Figure 1.7, state 'F') and adopting a His-on MetCbl 6-coordinate state (Figure 1.7, state 'G') that precedes the conformational changes to the resting state (Figure 1.7, state H).

In the human protein, the Cbl-loaded C-half MTR is found between His-on and His-off states, with approximately 55% of the protein population in the His-off state (Mascarenhas, Guha et al. 2023). In the dataset collected here from the full-length protein, the particles in the His-off conformation represented over 78% of the population, with only 22% in the His-on state. This difference could be stochastic or possibly due to differences in constructs, as the full-length protein was used in this current work. The higher prevalence of the His-off state without the need for a partner protein, as shown for *E. coli*, could be explained by the static position of Tyr1177 in the human protein. The residue is positioned similarly to the structure in which it coordinates with the water molecule on the upper axis of Cbl (PDB: 3IVA). This suggests that the movement of Cbl towards the activation domain might cause the shift from 5-coordinate cob[II]alamin (Figure 5.26, state 'B') to 4-coordinate cob[II]alamin (Figure 5.26, state 'C').

When Cbl is in a 4-coordinate state, it exhibits a higher reduction potential than the 5-coordinate form by approximately 250 mV, making it more susceptible to reduction. Additionally, the cobalt ion in the 4-coordinate state is stabilised by its Singly Occupied Molecular Orbital (SOMO), as this reduces variability in its state transitions (Liptak, Fleischhacker et al. 2009). In summary, the 4-coordinate Cbl is not only more stable but also allows the reducing agent (MTRR for human protein) to approach tangentially or perpendicularly, potentially increasing the rate of Cbl reduction (Marcus and Sutin 1985). Therefore, when human MTR transitions to the His-off state, it may not readily revert to the His-on state, as observed in bacterial MTR. This is because, in the His-off state, Cbl is immediately converted to its stable 4-coordinate form, potentially hindering the transition back to the His-on state.

In *E. coli*, 4-coordinate Cbl is not stabilised before reductive methylation; the binding of the reducing agent (flavodoxin) and the methyl donor (SAM) causes Cbl to shift to the

His-off 4-coordinate state, which can then be reductively methylated. In humans, it appears that the His-off 4-coordinate state is stabilised by the transition from the His-on to His-off state only (Figure 5.26, state 'B' to 'C'), which is dictated by the oxidation state of Cbl. The stabilised Cbl His-off 4-coordinate form can then be reduced and methylated upon MTRR and SAM binding (Figure 5.26, state 'D'). An example of 4-coordinate-stabilised Cbl can be found in ATP-cobalamin adenosyltransferase (MMAB), a protein also involved in intracellular cobalamin metabolism. MMAB stabilises cob[II]alamin in a 4-coordinate state without requiring additional ligands, similar to human MTR (Gouda, Li et al. 2023). The stabilised Cbl is then converted to AdoCbl upon ATP binding. This suggests that human MTR may follow a similar pattern, where Cbl is stabilised in its most stable form solely through its interaction with the protein.

The structure of the C-half MTR from the full-length protein has finally allowed for the visualisation of the arrangement of the three domains during the reactivation cycle. The structural details, particularly concerning His785, Tyr1177, and Cbl, have provided valuable insights into the remethylation process in the human protein, which appears to differ from the bacterial process. Although much of the bacterial MTR biology can be translated to the human protein, this study demonstrates that there are nuances to the mechanism and regulation of the human MTR cycle.

6. Characterisation of MTR, MTRR, and MMADHC complex formation

6.1 Background

The proteins responsible for the processing and transport of cobalamin (Cbl) to the two client enzymes in humans—methionine synthase (MTR) in the cytosol and methylmalonyl-CoA mutase (MMUT) in the mitochondria—rely on multiple protein-protein interactions (McCorvie, Ferreira et al. 2023). Co-immunoprecipitation assays in fibroblasts have demonstrated proximity between MMACHC which receives and processes Cbl in the cytosol, MMADHC which transports and delivers Cbl, MTR, and MTRR, the latter being responsible for the reductive methylation of Cbl in MTR (Fofou-Caillierez, Mrabet et al. 2013, Bassila, Ghemrawi et al. 2017). These four proteins may form a supramolecular complex, given that MMACHC and MMADHC must interact to deliver Cbl to MTR, and that MTRR may have regulatory roles in MTR beyond its reactivation function.

This supramolecular complex is estimated to have a molecular weight of approximately 282 kDa assuming a 1:1:1:1 stoichiometry. Although some of these interactions are better characterised, such as MMACHC-MMADHC (Froese, Kopec et al. 2015) and MTR-MTRR (Zhang, Liu et al. 2020), as discussed in Section 1.5, other interactions lack evidence in structural and biophysical data. Notably, MMADHC and MTRR would both need to interact with MTR via the same interface, as discussed in Section 5.4.1. No structural data currently exist for any of the protein-protein interactions. All structural characterisations of these proteins to date have been conducted on the individual isolated proteins.

This chapter aims to provide further insights into the interactions between MMADHC, MTR and MTRR using recombinant human proteins and structure modelling. Protein interactions were analysed through a co-expression system, blue native PAGE, spectrometry studies, and pulldown assays. To model the formation of these complexes, AlphaFold2-Multimer was employed to predict the structural basis of these interactions and to support the biophysical data. The transfer of Cbl from MMADHC to MTR is not dependent on MMACHC *in vitro* (Mascarenhas, Guha et al. 2023); therefore, MMACHC was excluded from the interaction studies.

6.2 The complex MTR-MTRR is expressed in insect cell system

To investigate the formation of the MTR-MTRR complex, the plasmid pBIG1a-MTR+MTRR(mut) was constructed to co-express MTR and MTRR in baculovirus-infected insect cells. The plasmid encodes full-length MTR (aa Lys16-Asp1265) with a C-terminal TEV-cleavable His_{10}-FLAG tandem tag, and full-length MTRR (aa Met1-Ser725) with an N-terminal TEV-cleavable His_{10}-tag. Each protein was under the control of its own promoter and terminator within the same bacmid vector for expression in insect cells. The predicted molecular weights of the tagged proteins are 142 kDa for MTR and 80 kDa for MTRR.

After generating and amplifying the baculoviruses using a monolayer of insect cells, up to three further amplifications were carried out in *Sf9* cells in suspension at a concentration of $1x10^6$ cells/mL. This approach generated a high-titre virus, which enabled the infection of most of the *Sf9* cells in culture, resulting in higher yields of recombinant proteins. The cells expressing the proteins were harvested, and the purification steps were conducted as outlined in Section 2.1.3. A sample from each one of the purification steps were saved for SDS-PAGE for analysis.

6.2.1 Ni-NTA purification

The first affinity purification step utilised the histidine (His_{10}) tag present on both recombinant proteins. Ni-NTA resin was selected for this step due to its strong affinity for His-tagged proteins, allowing the simultaneous binding of MTR and MTRR. This approach increased the yield of the two purified proteins, as they could each bind to the resin independently, even if they were not interacting with each other.

SDS-PAGE of the Ni-NTA purification (Figure 6.1) displays samples from various stages: before and after cell lysis (Total, Lysate), supernatant of the centrifuged lysate (SN), resin flow-through (FT), washes (W1-W3), and eluates (E1-E5). Bands corresponding to the expected sizes of each protein are visible in the gel, most prominently in the first elution fraction (E1). The MTR band, running close to the 140 kDa molecular weight marker, appears consistently throughout the washes and all eluates, while the MTRR band (near the 80 kDa marker) is visible only in elutions E1 and E2. This suggests an excess of MTR relative to MTRR, indicating that the surplus MTR present in the affinity-purified sample needs to be separated later on from the MTR-MTRR complex and, less likely, from any free MTRR. The eluates were pooled and concentrated for further purification.

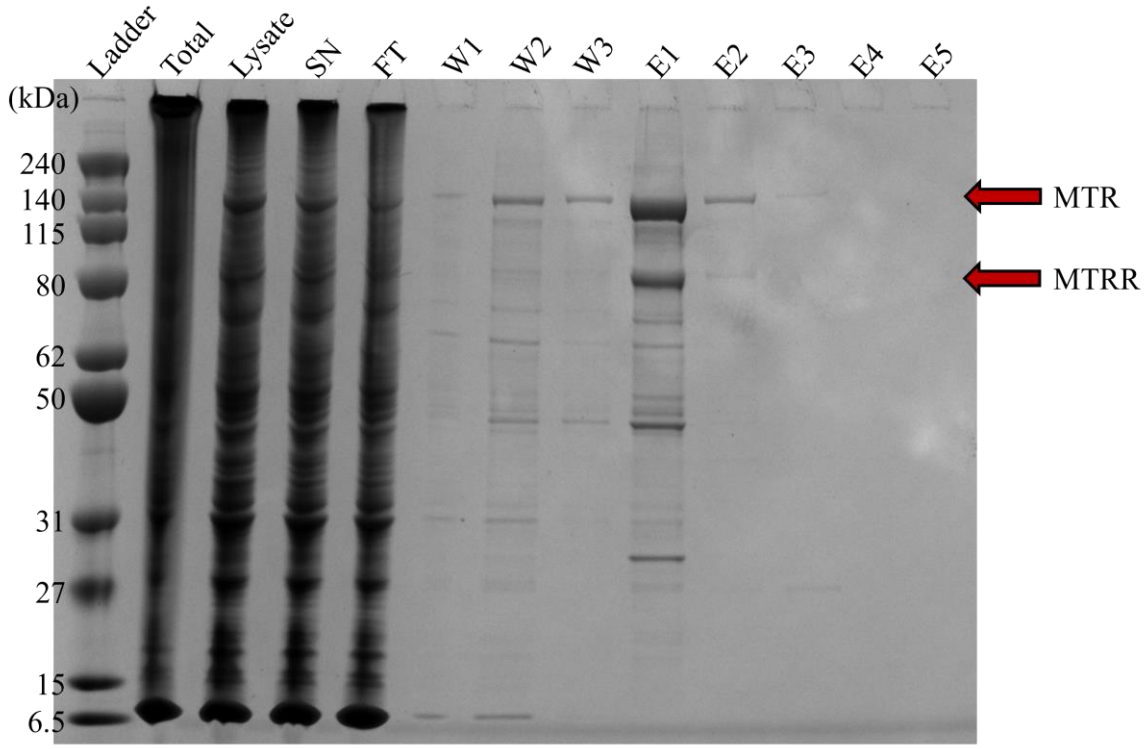

Figure 6.1: SDS-PAGE of MTR-MTRR complex purification using Ni-NTA resin. Each lane represents a step in the purification process. From left to right: protein standard ladder, total, lysate, supernatant (SN), flow-through (FT), washes 1-3 (W1-3), and eluates 1-5 (E1-5). The red arrows indicate the presence of protein in the corresponding size of MTR (142 kDa) and MTRR (80 kDa).

6.2.2 Gel filtration chromatography

Further purification of the MTR-MTRR complex (expected molecular weight 220 kDa) from excess MTR and other contaminants was carried out using gel filtration chromatography on a Superdex S200 26/60 column (Cytiva). Given the size difference between MTR (140 kDa) and MTRR (80 kDa), gel filtration was also capable of separating each isolated protein from the complex. The chromatography was performed with an elution buffer containing glycerol and high salt (500 mM NaCl), as described in Section 2.1.3, the same buffer used for the purification of isolated MTR. The glycerol and salt serve to stabilise the proteins and prevent aggregation, at the same time that it reduces the non-specific interactions between proteins, improving the gel filtration.

Five peaks were identified in the chromatogram (Figure 6.2A), which shows the absorbance of the elution fractions at 280 nm. Following a similar pattern to that observed during MTR purification (Section 3.2), the void fraction (Peak 1) displayed the highest absorbance of all peaks. Previous analysis revealed that the MTR present in the void fraction is aggregated (Section 3.5), and the high absorbance observed can also be attributed to the excess DNA (Boyes, Walker et al. 1988), likely due to the high concentration of baculovirus in the sample.

SDS-PAGE analysis (Figure 6.2B) revealed that the second peak (P2) contains both MTR and MTRR proteins in the same elution fraction, suggesting that these proteins may form a complex. Peaks P3 and P4 were composed solely of MTR, with P4 containing a greater amount of protein, as indicated by the band's intensity on the SDS-PAGE. Furthermore, the elution volume of P4 closely matches that obtained during MTR purification (Figure 3.1), indicating that the isolated MTR protein in P4 fraction is likely suitable for assays and structural analysis. The MTR protein eluted earlier in P3 is likely to correspond to the earlier peaks observed during MTR purification, where contaminant proteins were also detected on the SDS-PAGE (Figure 3.1). However, in this case, the lower amount of protein present likely prevented the contaminants from being observed. The final peak (P5) did not show any detectable protein on the gel and is most likely composed of smaller DNA fragments that migrated through the silica in the column.

No peak was observed containing only MTRR, suggesting that this recombinant protein is less stable than MTR, and its stabilisation may depend on complex formation with MTR. This instability of MTRR is expected, as previous studies have shown that MTRR could only be obtained with a glutathione S-transferase (GST) fusion added to the protein's N-terminus (Wolthers and Scrutton 2007, Bassila, Ghemrawi et al. 2017). The GST fusion is commonly used to increase the solubility of semi-soluble or insoluble proteins, thereby

enhancing the stability of recombinant proteins (Swaffield and Johnston 1996). While adding a GST fusion can improve protein stability and increase the final recombinant yield, this 26 kDa fusion protein is known to dimerise natively (Fabrini, De Luca et al. 2009), which could potentially interfere with interaction experiments between MTR and MTRR.

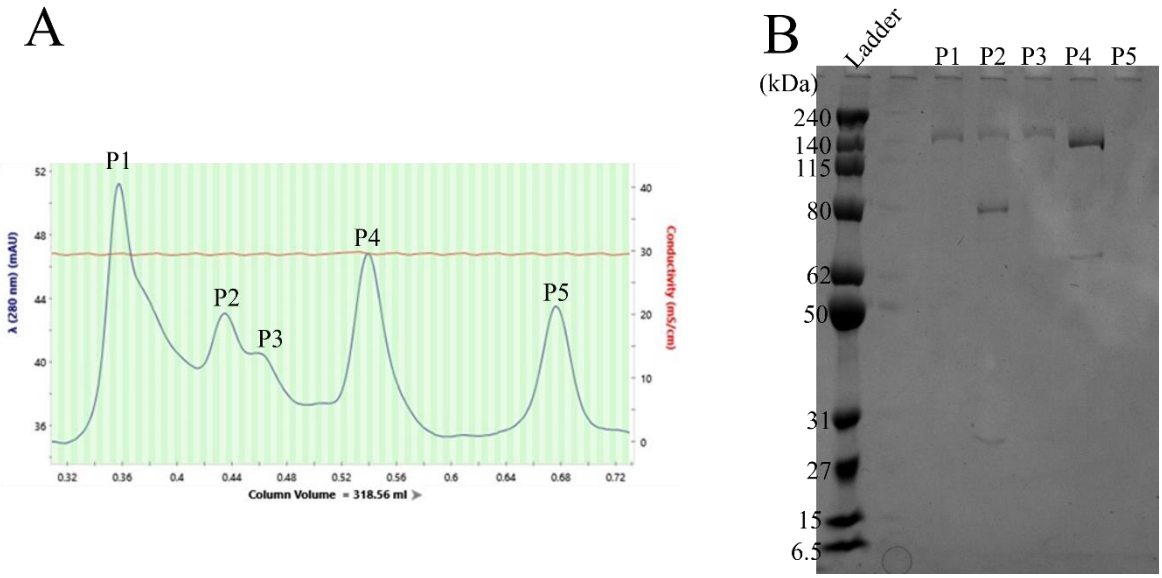

Figure 6.2: Gel filtration chromatography of the MTR-MTRR complex from the Ni-NTA purified sample. A. Chromatogram from the Superdex S200 26/60 column showing all five peaks observed under absorbance at 280 nm, labelled P1-P5. **B**. SDS-PAGE analysis of samples from each peak (P1-P5), indicating the presence of MTR in peaks 2-4, with both MTR and MTRR present in peak 2.

The sample corresponding to P2 was pooled, concentrated, and stored. Although the presence of both MTR and MTRR proteins in the gel filtration sample was confirmed by SDS-PAGE, their identities, as well as whether the two proteins are indeed interacting, will need to be confirmed in further studies.

6.3 The MTR-MTRR complex is confirmed by Blue Native PAGE and FLAG purification

After identifying from the SDS-PAGE that P2 from gel filtration contains two proteins with the appropriate molecular weights for MTR and MTRR (Figure 6.2B), both

bands were submitted to liquid chromatography-tandem mass spectrometry (LC-MS/MS) for protein identification. The Newcastle University Protein and Proteome Analysis (NUPPA) facility processed and analysed the protein bands. The gel bands were treated and the proteins extracted were incubated with trypsin and the peptides generated were analysed on a mass spectrometer coupled to the chromatography system. Data were processed using the MaxQuant software platform (v2.6.4) and further visualised and analysed in Perseus (version 2.1.2) (Tyanova, Temu et al. 2016).

The LC-MS/MS analysis confirmed the presence of MTR in the higher molecular weight gel band (140 kDa) and identified both MTR and MTRR in the lower gel band (80 kDa). The presence of MTR in the lower band is likely due to partial degradation of the protein, potentially at the Linker 1, which separates the N-half and C-half.

6.3.1 Blue native PAGE

To verify that the MTR and MTRR proteins present in the P2 sample form a direct interaction, the sample was analysed using Blue Native PAGE (Schägger and von Jagow 1991). This technique involves running a protein gel without denaturing agents (such as SDS), thereby preserving protein-protein interactions and providing insight into whether the proteins observed in SDS-PAGE are forming a complex. The addition of Coomassie Brilliant Blue dye in Blue Native PAGE imparts a uniform negative charge to the proteins in the sample, allowing them to be separated primarily by molecular weight while maintaining their native structure and interactions (Wittig, Braun et al. 2006).

Both MTR and MTR-MTRR samples were run side-by-side on the Blue Native PAGE, each at a concentration of 1 mg/mL (Figure 6.3). In the non-reducing gel, MTR appeared predominantly as a band corresponding to approximately 242 kDa, with a less intense band at 480 kDa. Although the molecular weight of MTR is 142 kDa, its migration

under native conditions reflects a higher molecular weight likely due to its non-globular conformation. The MTR-MTRR complex displayed a band consistent with a molecular weight of around 800 kDa, with no visible lower molecular weight bands corresponding to MTR or MTRR alone. The absence of lower bands suggests that most of the proteins in the sample are composed of the MTR-MTRR complex. The higher apparent MW in the gel might indicate a non-globular conformation of the complex, which aligns with the fact that both MTR and MTRR are flexible proteins. No MTRR was included as a separate sample for comparison because, up to this point, no construct for MTRR expression in insect cells has resulted in successful purification of the protein.

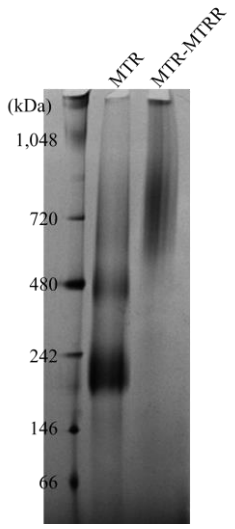

Figure 6.3: Blue Native PAGE of MTR and MTR-MTRR samples. Both recombinant MTR alone and the MTR-MTRR complex were analysed using Blue Native PAGE under non-denaturing conditions to confirm the interaction between MTR and MTRR. The complex sample (fraction P2 from Figure 6.2) appeared as a higher molecular weight band (i.e. lower mobility) in the gel compared to MTR alone, indicating an interaction between the two proteins. Representative SDS-PAGE from two independent experiments.

6.3.2 FLAG purification

To further confirm the interaction between MTR and MTRR, the FLAG tag present only on the MTR protein expressed from the pBIG1a-MTR+MTRR(mut) construct was utilised in a pull-down experiment. This experiment aimed to verify whether FLAG-tagged MTR could co-purify its interacting partner MTRR, which lacks the FLAG tag but possesses a His_{10} tag. The purified sample containing both MTR and MTRR (Figure 6.2, P2), obtained after Ni-NTA and gel filtration steps, was used in this assay. The presence of both MTR and

MTRR in the FLAG pull-down elution would indicate an interaction, as MTRR only contains the His_{10} tag.

The FLAG pull-down assay was performed as described in Section 2.6. In brief, 10 µg of the purified sample was diluted to a final volume of 100 µL in either KCl buffer (100 mM) or NaCl buffer (500 mM) and incubated with 100 µL of ANTI-FLAG® M2 Affinity Gel resin. The two different buffer conditions were tested to assess whether varying salt concentrations and types would influence the association/dissociation of MTR-MTRR complex.

The resin and protein sample were incubated at 4 °C for 2 hours, either stationary or under agitation, as agitation can sometimes cause dissociation of protein complexes (Jayaraman, Buck et al. 2014). After incubation, the resin was transferred to purification mini spin columns and centrifuged to remove the flow-through. The resin was then washed twice for each sample. To elute proteins bound to the resin, reducing Blue Protein Loading Dye was added, and after one minute of incubation, the eluates were collected by centrifugation. The flow-through, washes, and elutions from each test were analysed via SDS-PAGE (Figure 6.4).

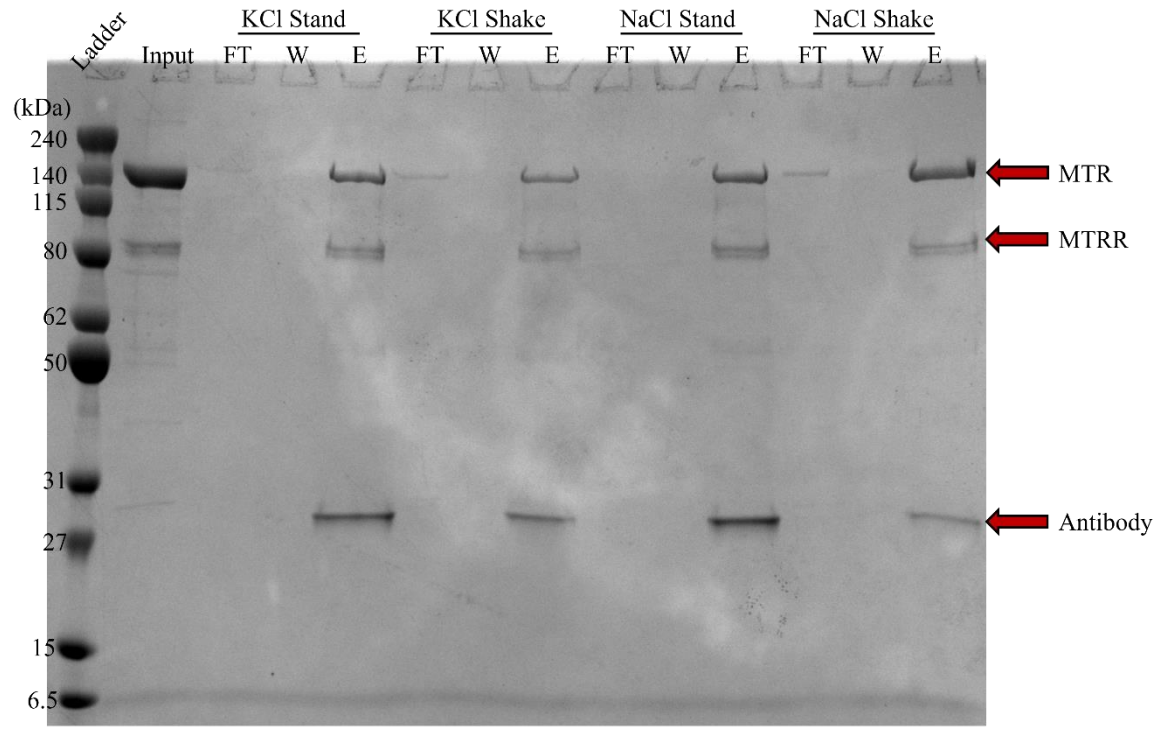

Figure 6.4: Pulldown of the MTR-MTRR complex using ANTI-FLAG® M2 Affinity Gel. The gel shows the protein ladder, the sample input used for the pulldown, and the flow-through (FT), wash (W), and elution (E) for each condition tested (KCl or NaCl buffer – incubated either standing or shaking). In all elution fractions, both MTR (140 kDa) and MTRR (80 kDa) are visible, indicating that they are interacting. The arrows denote the positions of MTR, MTRR, and antibody fractions. Representative SDS-PAGE from two independent experiments.

The ANTI-FLAG® M2 Affinity resin used for the pulldowns is composed of antibodies that specifically bind to the FLAG peptide (DYKDDDDK), which is incorporated only in MTR as per the construct. The presence of both MTR and MTRR in the elution lanes (Figure 6.4, lanes E) indicates that MTRR is interacting with MTR, which is bound to the resin via its FLAG tag. In all elutions, a band of approximately 30 kDa is visible, which is most likely due to the antibody chains from the resin. This occurs because the elution was performed using a denaturing buffer, which can reduce the disulfide bonds holding the M2 antibody chains together, causing them to appear in the SDS-PAGE. The NaCl and KCl buffers, although have different concentrations of salt (500 mM vs 100 mM), did not alter the elution of the MTR-MTRR complex.

The interaction between the FMN-binding domain (FMNBD) of MTRR and the C-terminal half of MTR has been demonstrated previously (Wolthers and Scrutton 2007);

however, the study here is the first to show that full-length MTR and MTRR can form a stable complex. This complex formation was initially confirmed by LC-MS/MS, where the identities of both proteins were verified in the purified sample. Blue Native PAGE further validated the interaction, as the sample containing both MTR and MTRR exhibited a higher molecular weight than MTR alone in its native state. Finally, the pulldown assay confirmed the interaction between the two proteins, as the resin used binds only the FLAG peptide, present on MTR, and the elution showed both MTR and MTRR, indicating a direct interaction.

Further structural studies are needed to fully understand this interaction. Unfortunately, the attempts to prepare cryo-EM samples of the MTR-MTRR complex were unsuccessful, likely due to the flexibility of both proteins, sample quality issues, or ice thickness. The particles identified in the micrographs (not shown) did not yield any clear 2D class averages, preventing further analysis.

6.4 OHCbl-loaded MMADHC fragment can interact and transfer OHCbl to MTR

After confirming the interaction between MTR and MTRR in a construct expressing both full-length proteins, the next step was to test the interaction between MTR and MMADHC. It is known that MMADHC dimerizes with MMACHC *in vivo* for Cbl delivery (McCorvie, Ferreira et al. 2023), but Cbl delivery to MTR can be performed *in vitro* by MMADHC alone (Li, Mascarenhas et al. 2020).

In this section, the transfer of Cbl from MMADHC to MTR was tested using recombinant proteins and a specific form of Cbl, Hydroxocobalamin (OHCbl). The recombinant MTR used was the full-length protein, while a truncated construct of MMADHC (Gln124-Asn296, $MMADHC_{\Delta N}$) was employed. This construct has an N-terminal His_6 tag and had been previously expressed in *E. coli* and purified by the group

(Froese, Kopec et al. 2015). The absence of the first 123 residues in the N-terminal makes the recombinant protein more stable, as these residues form a disordered region and are thought to constitute a mitochondrial leader sequence (Stucki, Coelho et al. 2012). Furthermore, the transfer of Cbl from a truncated MMADHC (Met116-Asn296) to the C-terminal half of MTR (Ile661-Asp1265) has been shown to take place (Mascarenhas, Guha et al. 2023), indicating that the N-terminal region of MMADHC is not essential for Cbl transfer to MTR.

The construct for $MMADHC_{\Delta N}$ was expressed in 2 L of *E. coli* and purified in a two-step process involving Ni-NTA purification, leveraging the His_6 tag in the protein's N-terminal region, followed by gel filtration using a Superdex 200 Increase 10/300 GL column. No Cbl was added during expression or purification, so the protein is assumed to be in its *apo* form, as confirmed by the experiments in this section. The protein aliquots were stored in 50 mM HEPES, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP at -80°C.

To test Cbl transfer from $MMADHC_{\Delta N}$ to MTR, $MMADHC_{\Delta N}$ was first incubated with OHCbl to form the $MMADHC_{\Delta N}$-thiolato-Cob[III]alamin complex (Section 6.4.1 below). OHCbl was chosen because it is a stable form of Cbl in the +3 oxidation state, making it ideal for this assay (Li, Mascarenhas et al. 2020, Mascarenhas, Gouda et al. 2022). The $MMADHC_{\Delta N}$-thiolato-Cob[III]alamin complex was then incubated with MTR and analysed by gel filtration to confirm the presence of Cbl-loaded MTR via absorbance (Section 6.4.2 below).

6.4.1 Cbl loading to $MMADHC_{\Delta N}$

To load Cbl onto $MMADHC_{\Delta N}$, excess $MMADHC_{\Delta N}$ was incubated with OHCbl in 150 mM KCl buffer for one hour at room temperature, and the UV-vis spectrum was recorded every 10 minutes. Cbl binding to MMADHC was assessed by comparing the λ_{max} values of

OHCbl and thiolato-Cob[III]alamin (the form of Cbl when bound to MMADHC) (Li, Mascarenhas et al. 2020). The plot of absorbance vs wavelength of the reaction at 0 minutes and 60 minutes (Figure 6.5) reveals a shift in λ_{max}, indicating a change in the wavelength at which maximum light absorption occurs. Since the optical properties of OHCbl and thiolato-Cob[III]alamin differ, the absorbance spectrum after 1 hour showed a shift in λ_{max} from 351 nm and 526 nm (OHCbl) to 334 nm, 372 nm, and 531 nm (thiolato-Cob[III]alamin).

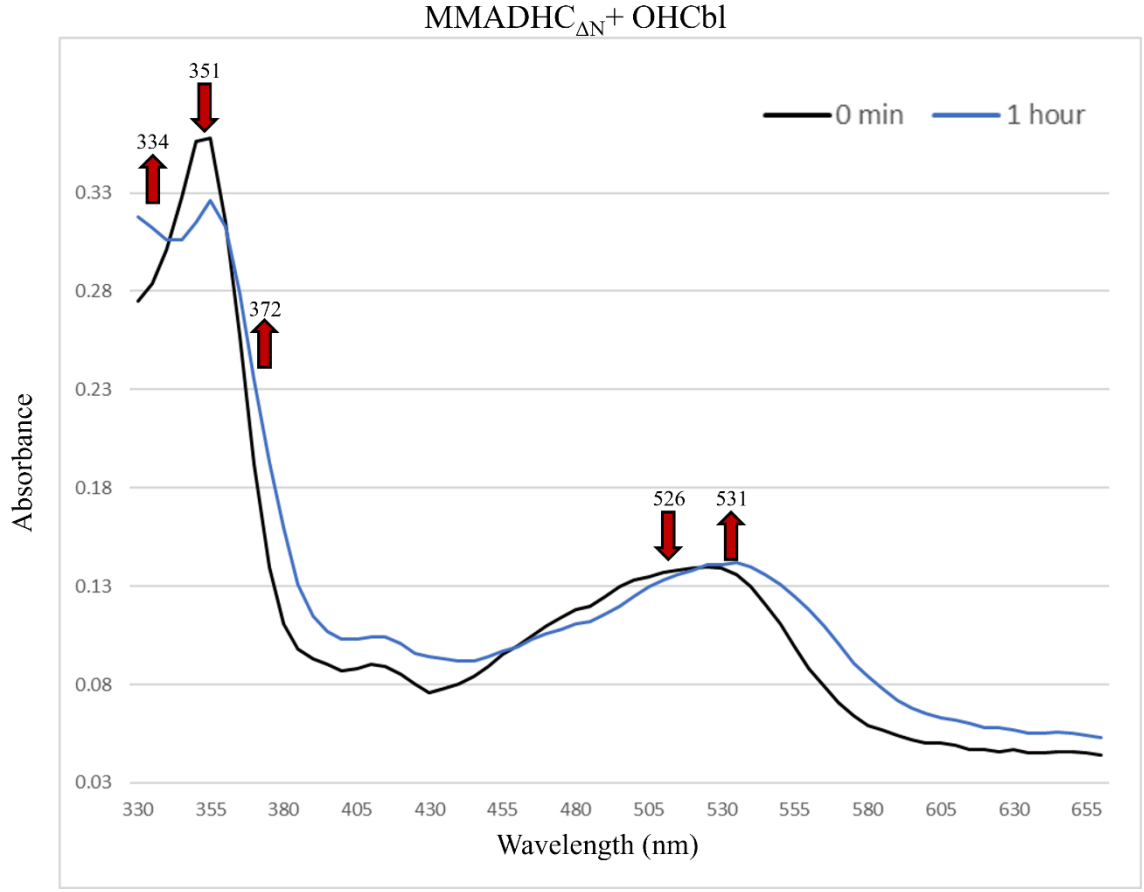

Figure 6.5: UV-vis spectrum changes after one hour incubation of MMADHC$_{\Delta N}$ with OHCbl. The red arrows show the increase or decrease of the specified wavelength that characterise the conversion of OHCbl to thiolato-Cob[III]alamin upon its binding to MMADHC$_{\Delta N}$. Representative graph from three independent experiments.

The UV-vis graph shows a change in the λ_{max}, specifically a decrease in the peaks at 351 nm and 526 nm, and an increase in the peaks at 334 nm, 372 nm, and 531 nm, confirming the binding of OHCbl to MMADHC$_{\Delta N}$. The excess OHCbl in the MMADHC$_{\Delta N}$ sample was then removed by several rounds of centrifugation using a 10 kDa filter microtube, followed by dilution of the reaction in KCl buffer. After five rounds, the flowthrough was no longer

red (from the Co ion, present in Cbl), indicating that no free OHCbl remained in the sample. The sample (referred to as MMADHC$_{\Delta N}$-Cbl) was then concentrated before incubation with MTR.

6.4.2 Cbl transfer from MMADHC$_{\Delta N}$ to MTR

For the Cbl transfer assay, MTR was incubated with equimolar MMADHC$_{\Delta N}$-Cbl for one hour at room temperature in KCl buffer. After the incubation, the sample was loaded onto a Superdex 200 Increase 10/300 column, and the absorbance was monitored at 280 nm and 350 nm to capture the absorbance of protein and Cbl, respectively. As controls, MTR alone, MMADHC$_{\Delta N}$ alone, MMADHC$_{\Delta N}$-Cbl, and OHCbl alone were tested separately at the same concentrations used in the MTR incubation with MMADHC$_{\Delta N}$-Cbl (Figure 6.6).

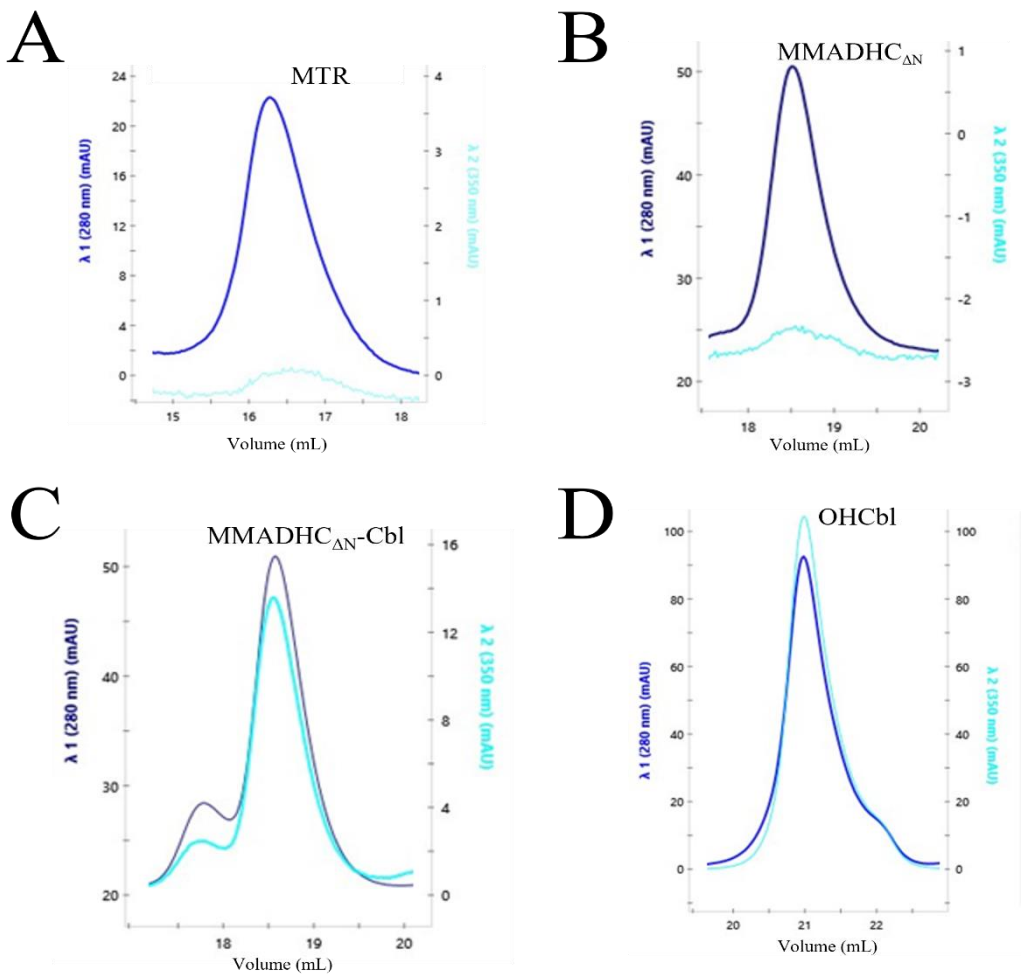

Figure 6.6: Controls were loaded into the Superdex 200 Increase 10/300 column, and absorbance readings were taken at 280 nm and 350 nm. The controls included: **A**. full-length MTR. **B**. MMADHC$_{\Delta N}$. **C**. MMADHC$_{\Delta N}$ incubated for one hour with OHCbl, followed by buffer exchange to remove excess OHCbl. **D**. OHCbl diluted in KCl buffer. The X-axis of each graph represents the column elution volume (mL) (total column volume = 23.56 mL), and the Y-axis shows the absorbance at either 280 nm (blue, left axis) or 350 nm (cyan, right axis). Representative chromatogras from two independent experiments for each condition.

The controls tested in the assay reveal that MTR (Figure 6.6A) and MMADHC$_{\Delta N}$ (Figure 6.6B) only exhibit peak maxima at 280 nm, typical of protein absorbance, without any absorbance at 350 nm, as expected due to the absence of Cbl. This is also supportive evidence that the as-purified MMADHC$_{\Delta N}$ is *apo*. When MMADHC$_{\Delta N}$ was incubated with OHCbl, a significant absorbance at 350 nm was observed (Figure 6.6C), indicating that MMADHC$_{\Delta N}$ had successfully bound Cbl, as shown previously in section 6.4.1. Both 280 nm and 350 nm absorbance peaks were observed for OHCbl alone (Figure 6.6D), which is

expected due to the presence of the corrin ring of Cbl, known to absorb at 280 nm (Salerno, Miller et al. 2020).

When MTR was incubated with Cbl-loaded MMADHC$_{\Delta N}$, the chromatogram profile displayed two 280 nm peaks, with elution volumes at 0.68 mL (corresponding to MTR) and 0.79 mL (corresponding to MMADHC$_{\Delta N}$) (Figure 6.7). Additionally, 350 nm absorbance peaks were observed at the same elution volumes for both MTR and MMADHC$_{\Delta N}$, indicating that both proteins were bound to Cbl. This suggests that an interaction between Cbl-loaded MMADHC$_{\Delta N}$ and MTR occurred, leading to successful Cbl transfer from MMADHC$_{\Delta N}$ to MTR.

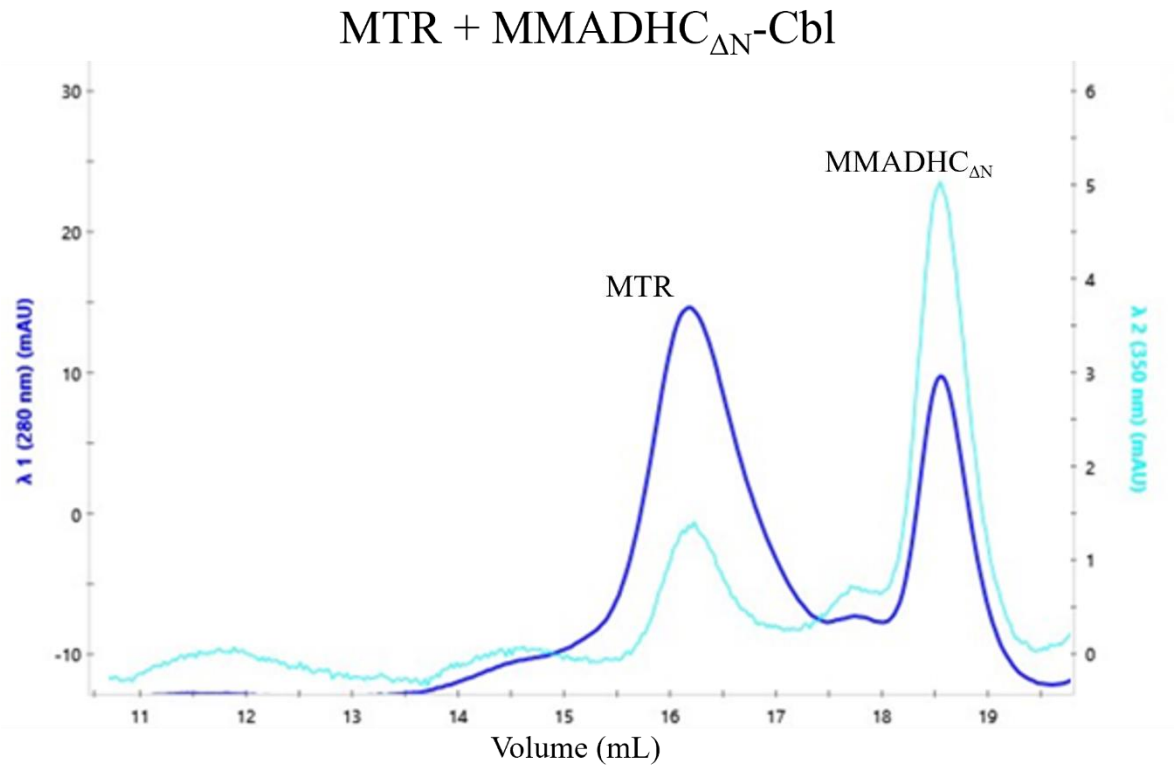

Figure 6.7: Chromatogram profile of MTR incubated with Cbl-loaded MMADHC$_{\Delta N}$ for one hour at room temperature. Absorbance at 280 nm (blue, left Y-axis) and 350 nm (cyan, right Y-axis) were recorded to detect protein and Cbl, respectively. The first peak corresponds to MTR, and the second peak represents MMADHC$_{\Delta N}$-Cbl. The X-axis shows the column elution volume (total column volume = 23.56 mL). Representative chromatogram from two independent experiments

When incubated with MMADHC$_{\Delta N}$-Cbl, the MTR peak eluted slightly earlier than the control MTR sample (elution volume 0.68 mL vs. 0.70 mL), suggesting that MTR might form a complex with MMADHC$_{\Delta N}$, which could be retained during the gel filtration run. However, despite this shift in elution volume, it remains unclear whether MTR and MMADHC$_{\Delta N}$ are still interacting during the chromatographic run or if the interaction is transient. Furthermore, the 350 nm absorbance peak, which is absent in the MTR control (Figure 6.6A), appears alongside the 280 nm peak corresponding to MTR after incubation with MMADHC$_{\Delta N}$-Cbl (Figure 6.7), suggesting successful Cbl transfer.

Although it was not possible to determine if the interaction between MMADHC$_{\Delta N}$ and MTR is transient or if they form a stable complex through the gel filtration studies, as shown for MTR-MTRR (Section 6.3), the chromatograms confirm that Cbl was successfully transferred from MMADHC$_{\Delta N}$ to MTR, a crucial interaction *in vivo* for the delivery of Cbl to *apo* MTR (Froese, Kopec et al. 2015, Mascarenhas, Guha et al. 2023).

6.5 Insights into MTR, MTRR, and MMADHC complex formation using AlphaFold2 Multimer

In previous sections, MTR was shown to form a stable complex with MTRR in a co-expression system (Section 6.3), and MMADHC$_{\Delta N}$ was demonstrated to transfer Cbl to MTR (Section 6.4). To further investigate the protein-protein interfaces of these two complexes (MTR-MTRR and MTR-MMADHC$_{\Delta N}$), AlphaFold2 Multimer (AF2 Multimer) was used to model these complexes. This tool extends the capabilities of the original AF2 algorithm, enabling the prediction of interaction models involving two or more proteins (Evans, O'Neill et al. 2021). Additionally, AF2 Multimer was used to model a three-way complex involving MTR, MTRR, and MMADHC. The sequence of each human full-length protein was used as input to generate interaction models for MTR with MTRR (Sections 6.5.1 and 6.5.2), MTR

with MMADHC (Section 6.5.3), and the three-way complex of MTR, MTRR, and MMADHC (Section 6.5.4).

The three-way complex model was based on the hypothesis that MTR, MTRR, and MMADHC may form a supramolecular complex (Bassila, Ghemrawi et al. 2017). As discussed in Chapter 5, MTR may interact with MTRR and MMADHC via a shared interface. Modelling this complex could provide insights into how the three proteins interact simultaneously. Understanding the molecular details of these interactions can reveal key residues involved in these protein-protein contacts and potentially shed light on the structural basis of mutations in these proteins that are associated with metabolic disorders.

6.5.1 Modelling the MTR-MTRR complex

MTR consists of five domains: HcyBD, FolBD, Cap, CblBD and AD, as introduced in Section 1.4.1. MTRR is composed of four domains: FMNBD, connected by a long hinge (residues Gln187-Tyr276) to the rest of the protein that contains FADBD, NADPHBD, and CD. The FMNBD of MTRR is known to interact with the AD of MTR, facilitating electron transfer to reduce Cbl to cob[I]alamin for remethylation (Wolthers, Basran et al. 2003, Wolthers, Lou et al. 2007, Wolthers and Scrutton 2007, Wolthers and Scrutton 2009). Additionally, MTRR has been shown to enhance MTR activity in vitro, suggesting a chaperoning role beyond electron transfer (Yamada, Gravel et al. 2006). Despite this proposed additional function, structural details of the interaction beyond the AD-FMNBD interface remain unknown.

Here, twenty-five AF2 Multimer models were generated using the full-length sequences of both human MTR (Q99707.2) and MTRR (Q9UBK8). The MTR-MTRR complex model (Figure 6.8) shows the expected interaction between the FMNBD of MTRR and the AD of MTR in the activation conformation, allowing electron transfer to Cbl. While

the interaction between FMNBD and AD has been previously demonstrated through biophysical assays (Zhang, Liu et al. 2020), the predicted model suggests additional possible interactions with the CblBD and Cbl.

Although the there is no experimental structure of MTRR FMNBD, insights from the NADPH-cytochrome P450 oxidoreductase (CYPOR) structure suggest the position of the FMN molecule within the FMNBD of MTRR. This positioning of the FMN suggests where it should be located within the interaction interface with MTR (Hamdane, Xia et al. 2009). Even though AF2 Multimer generate models that may not represent the most energetically favourable conformations, particularly in flexible regions and sidechains (Saldaño, Escobedo et al. 2022), the predicted models for MTR-MTRR complex provides valuable insights into possible interactions within the complex.

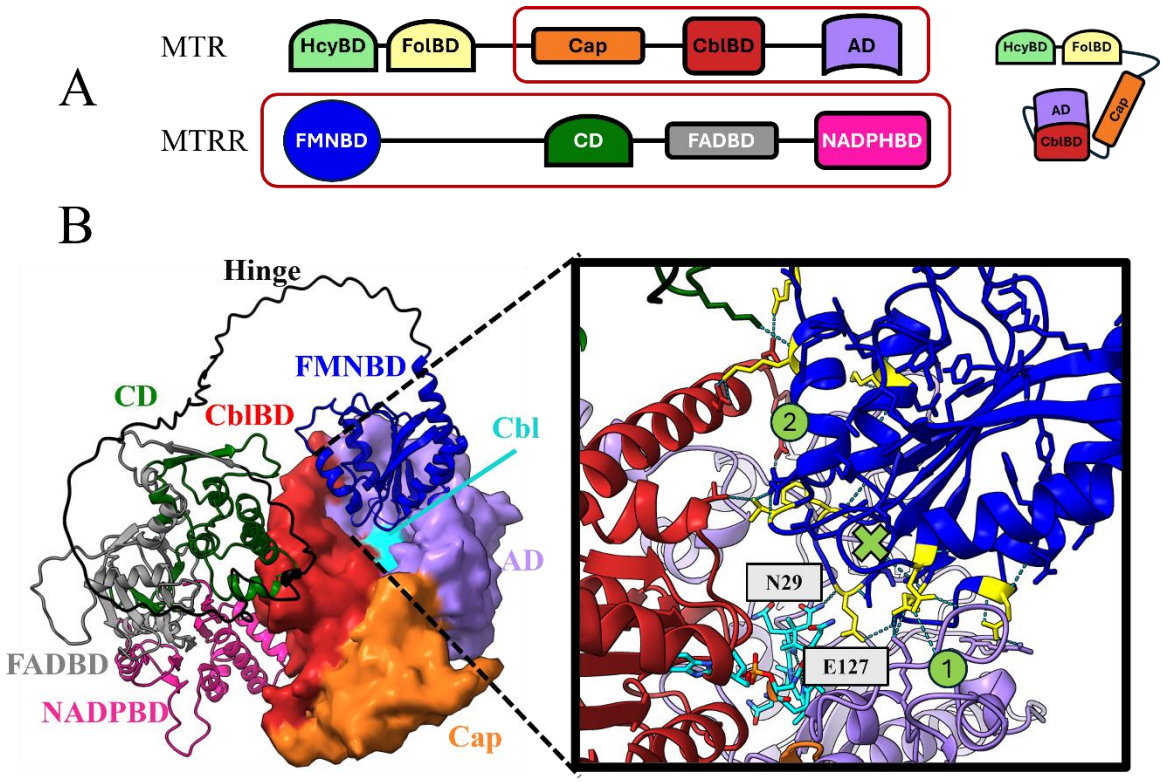

Figure 6.8: AlphaFold2 Multimer Structure Prediction of the MTR-MTRR Complex. A. Schematic representation of the MTR and MTRR domains used for model generation, with each domain color-coded to correspond to the model. A scheme of MTR domains in the activation conformation is also shown. Red rectangles represent the domains depicted in B. **B**. MTR-MTRR complex model, in which the HcyBD and FolBD domains of MTR are omitted for better visualization, but they functioned as a single unit in the predictions. Their positioning varied based on the position of the loop connecting them to the Cap. The Cap (orange), CblBD (red), AD (purple), and Cbl (cyan) of MTR are represented as a volume, while the FMNBD (blue), FADBD (grey), CD (green), and NADPHBD (pink) of MTRR are shown as cartoons. The inset highlights the interaction between the FMNBD of MTRR and MTR. The interacting residues are coloured yellow, and the green X represents the probable position of the FMN molecule based on the structure of the NADPH-cytochrome P450 oxidoreductase (CYPOR) (PDB: 6J79). The green ball marked with 1 shows the approximate position of the Lys987 and the one marked with 2 shows the position of Lys1071 in the AD of MTR. The predicted complex structure has an interface predicted template modeling (ipTM) score of 0.65.

The Cbl molecule was added to the predicted structure using AlphaFill, revealing that the FMN-binding pocket of the FMNBD is oriented toward Cbl (Figure 6.8, inset). The positioning of these two domains suggests that electron transfer from FMNBD to the Cbl ligand in MTR, may be feasible in this conformation. Furthermore, the potential interaction between Glu127 in the FMNBD and N29 of Cbl is also indicated. Although there is currently no evidence that Glu127 plays a crucial role in the interaction between FMNBD and MTR, it is the only residue in the model that interacts with the Cbl ligand. This suggests that it may

be important for stabilising this interaction, facilitating the electron transfer to Cbl. The FMNBD also appears to form multiple interactions with two key regions of the AD, as marked by the green circles in Figure 6.8.

Lys987 and Lys1071 of MTR have previously been shown to be important for MTR-MTRR interaction (Wolthers and Scrutton 2007). Although the model does not show direct interactions between Lys987 and Lys1071 of the MTR AD and the FMNBD, the regions where these residues are located display several other interactions in the predicted model (Figure 6.8, inset marked with number 1 and 2). The large number of interacting residues in the FMNBD (four residues in the region 1 and five residues in the region 2, indicated by the green balls in the inset of Figure 6.8), along with the positioning of the FMN molecule (marked as an X in the inset of Figure 6.8), suggest that the predicted placement of the FMNBD in interaction with MTR is likely accurate.

Across all twenty-five generated structures, the only significant conformational variation is in the relative position of the N-terminal half of MTR (HcyBD and FolBD) with respect to the C-terminal half, highlighting the flexibility of the loop connecting these regions (Figure 6.9). This high flexibility in the N-terminal half of MTR was also observed during data analysis of the micrographs of full-length MTR (Chapter 5).

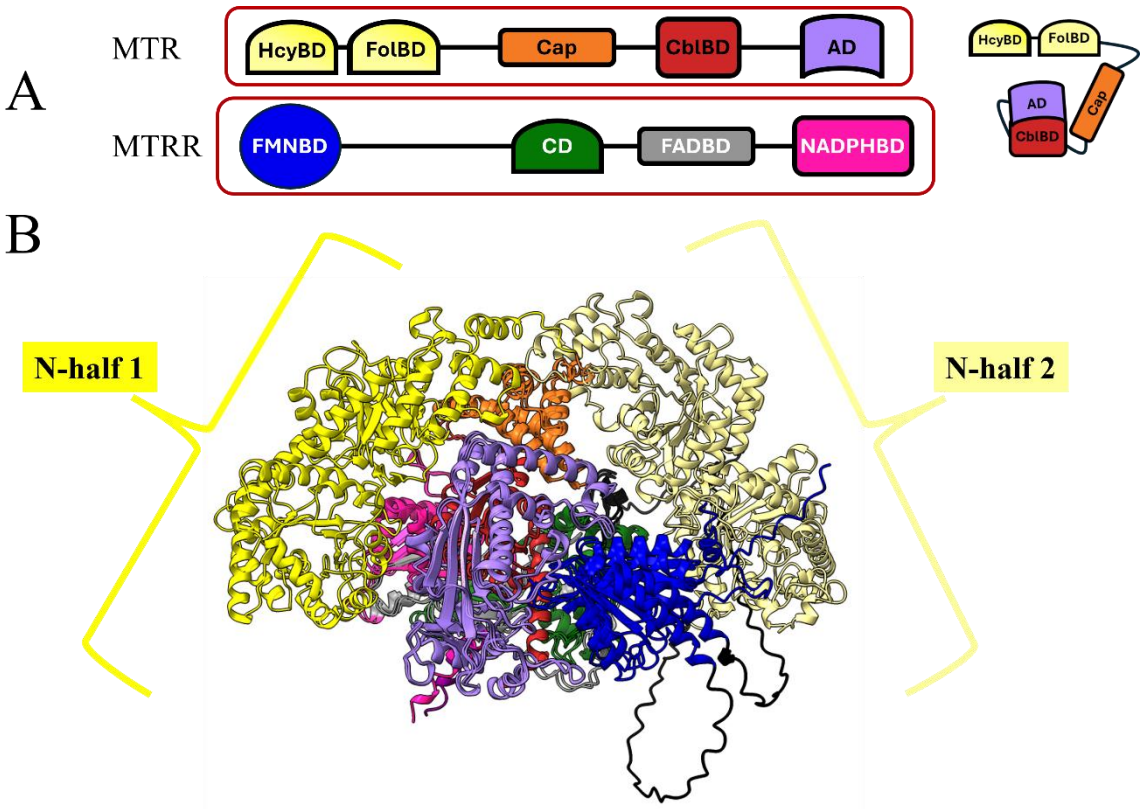

Figure 6.9: Comparison of the position of the N-half MTR between two models of the MTR-MTRR complex. A. Scheme of the domains displayed in the model, with their respective colours. A scheme of MTR domains in the activation conformation is also shown. Red rectangles represent the domains depicted in B. **B**. The two models were aligned according to the C-half domains of MTR, and the two different shades of yellow show two examples of positions the N-half MTR can acquire due the flexible loop connecting the N and C-half. The C-half MTR and MTRR have the same domain positions in both models.

Despite the flexibility of the N-half of MTR, all domains of MTRR (i.e., FMNBD, NADPHBD, FADBD, and CD) consistently assume the same relative position in all generated models. The FMNBD of MTRR is always positioned near the Cbl-binding pocket of MTR, which can be identified by the presence of the Cbl ligand in Figure 6.10. This positioning of the FMNBD is essential for electron transfer to Cbl. The other MTRR domains NADPHBD, FADBD, and CD) are consistently located in other regions of the CblBD of MTR. Specifically, the MTRR CD is always close to the long α-helix (α7) (Lys902-Glu926) in the CblBD of MTR, which connects the CblBD to the AD. The NADPHBD is close to the α3 (Cys812-Asp821) and α4 (Thr834-Arg849) of CblBD.

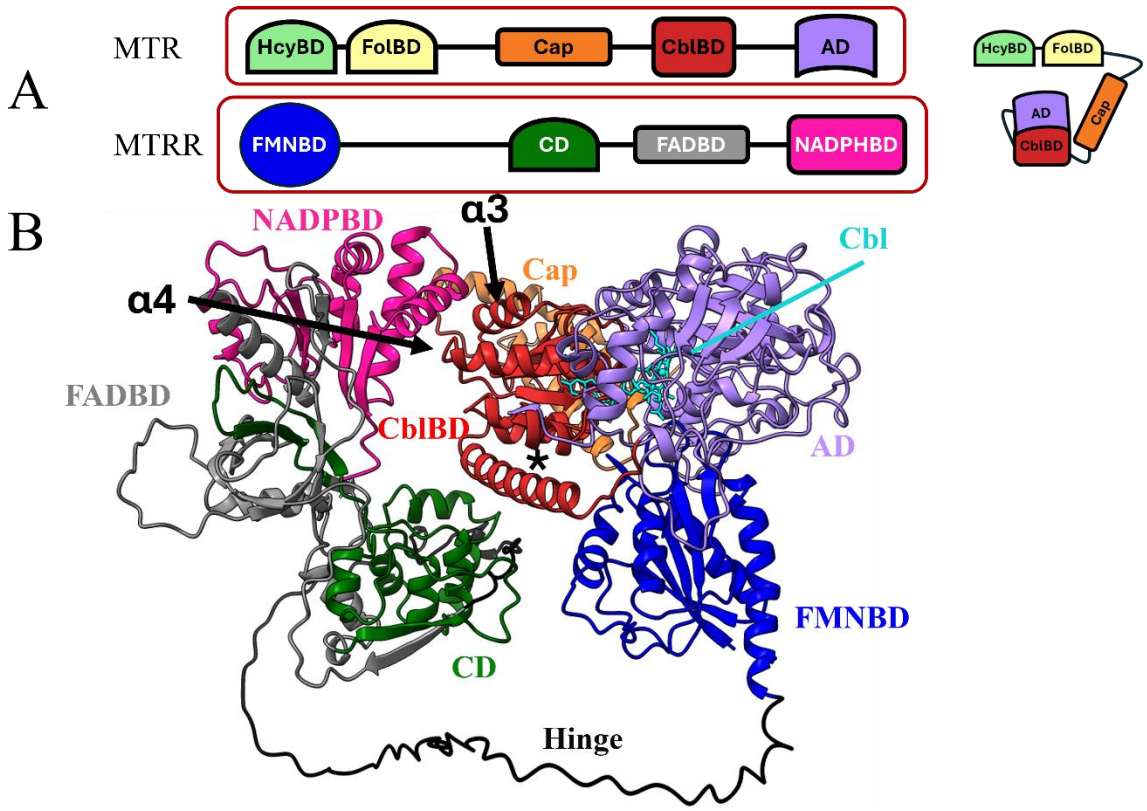

Figure 6.10: MTR-MTRR complex model focusing on the interaction of the CblBD of MTR with MTRR. **A**. The domains and colours of MTR and MTRR used to generate the models. A scheme of MTR domains in the activation conformation is also shown. Red rectangles represent the domains depicted in B. **B**. Model generated displaying the MTR domains Cap (orange), CblBD (red), and AD (purple) and Cbl (cyan). The HcyBD and FolBD were omitted for better visualization of the complex. The MTRR domains FMNBD (blue), FADBD (grey), CD (dark green), and NADPHBD (pink) are also shown. The long helix of CblBD (α7) is marked with a *, the helix α3 and α4 are also indicated in the figure, and are located next to the CD, FADBD, and NADPHBD.

The unexpected interaction between the MTR CblBD and MTRR outside of its FMNBD raises the question whether the MTR-MTRR complex might have an additional interaction interface beyond the AD-FMNBD interaction. To further investigate this novel interaction, new models of the MTR-MTRR complex were generated and are presented in the next section.

6.5.2 Modelling the MTR-MTRR$_{\Delta FMN}$ complex

To further investigate the novel interaction involving MTRR outside its FMNBD, with the CblBD of MTR, the FMNBD and hinge region were removed from the MTRR sequence, and a new complex prediction was generated using MTRR$_{\Delta FMN}$ (residues Asn237–

Ser725) and full-length MTR (Figure 6.11). By excluding these domains from the structure prediction, the bias towards the known interaction between the FMNBD of MTRR and AD of MTR was minimised. This approach allowed AF2 Multimer to reveal novel interactions between MTR and MTRR, independent of the AD-FMNBD interaction.

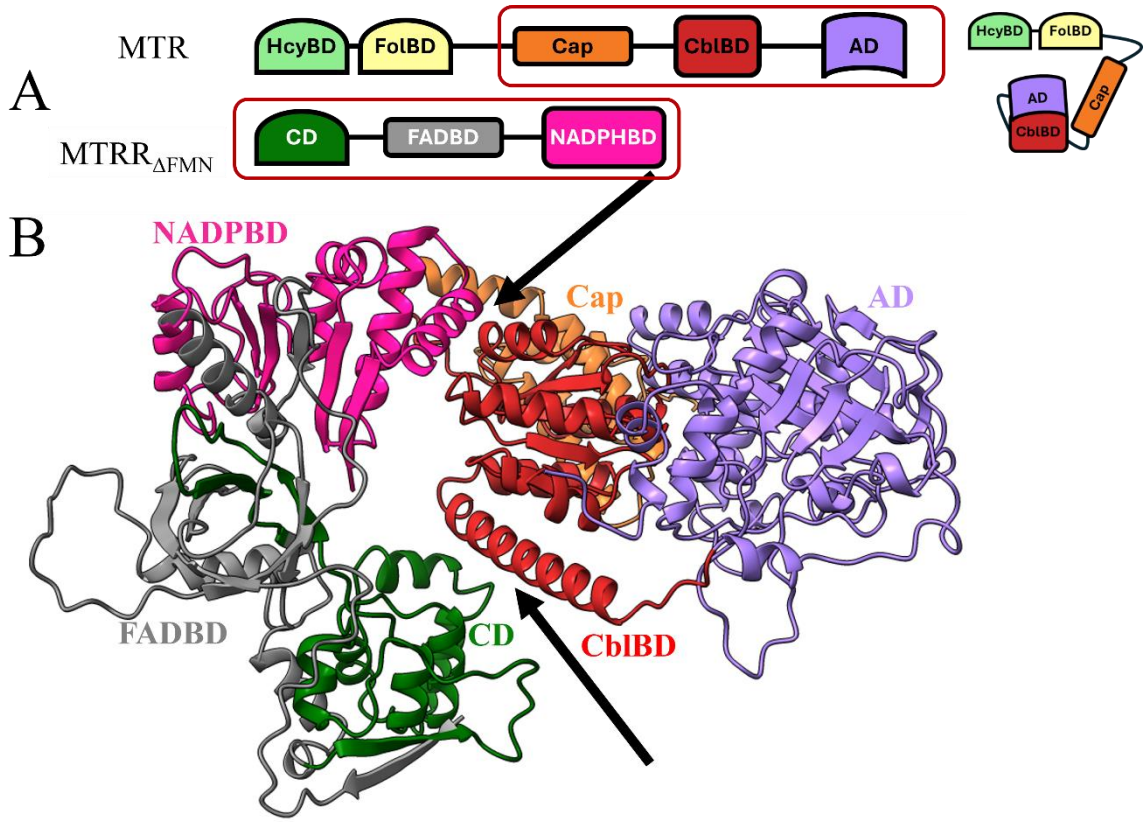

Figure 6.11: AlphaFold Multimer model of the MTR-MTRR$_{\Delta FMN}$ Complex A. The domains and corresponding colours of MTR and MTRR used for model generation are shown. A scheme of MTR domains in the activation conformation is also shown. Red rectangles represent the domains depicted in B. **B**. In this prediction, MTR adopts the activation conformation, with the CblBD (red) positioned close to the AD (purple) in the cap-off state (orange). Although the FMNBD and hinge are absent in MTRR$_{\Delta FMN}$, the remaining domains—FADBD (grey), CD (dark green), and NADPHBD (pink)—occupy the same positions in all predicted structures. The possible interactions between MTR CblBD and MTRR$_{\Delta FMN}$ through the CD and NADPHBD are marked with black arrows. The HcyBD and FolBD of MTR were omitted for clarity but behaved as a single unit in the predictions, assuming different positions depending on the orientation of the loop connecting them to the Cap. The predicted complex structure has an interface predicted template modeling (ipTM) score of 0.74.

The MTR-MTRR$_{\Delta FMN}$ model prediction revealed the same potential interactions between the CD of MTRR and the long helix of MTR, as well as between the NADPHBD and α3 and α4 helices of the CblBD, as seen in the MTR-MTRR model (previous Section

6.5.1). These interactions are indicated by arrows in Figure 6.11. Although the only previously described MTR-MTRR interaction involves AD and FMNBD (Wolthers, Basran et al. 2003, Wolthers and Scrutton 2007), all generated models here for MTR-MTRR and MTR-MTRR$_{\Delta FMN}$ displayed these possible new interactions involving the CD and NADPHBD of MTRR, with CblBD of MTR.

The known interaction between the AD and FMNBD occurs when MTR is in the activation conformation, where Cbl is accessible to receive an electron from the FMNBD during the reactivation cycle. The novel interactions, which do not involve the reactivation of Cbl, suggest that the MTR-MTRR complex might form independently of whether MTR is in the activation conformation. To better understand how the CD and NADPHBD of MTRR could interact with the CblBD of MTR, the PICKLUSTER software was employed (Genz, Mulvaney et al. 2023). This software can identify protein-protein interfaces in complex models generated by AF2 Multimer. The PICKLUSTER analysis of the MTR-MTRR$_{\Delta FMN}$ complex confirmed two main interaction sites of the CblBD with the CD and NADPHBD (Figure 6.12).

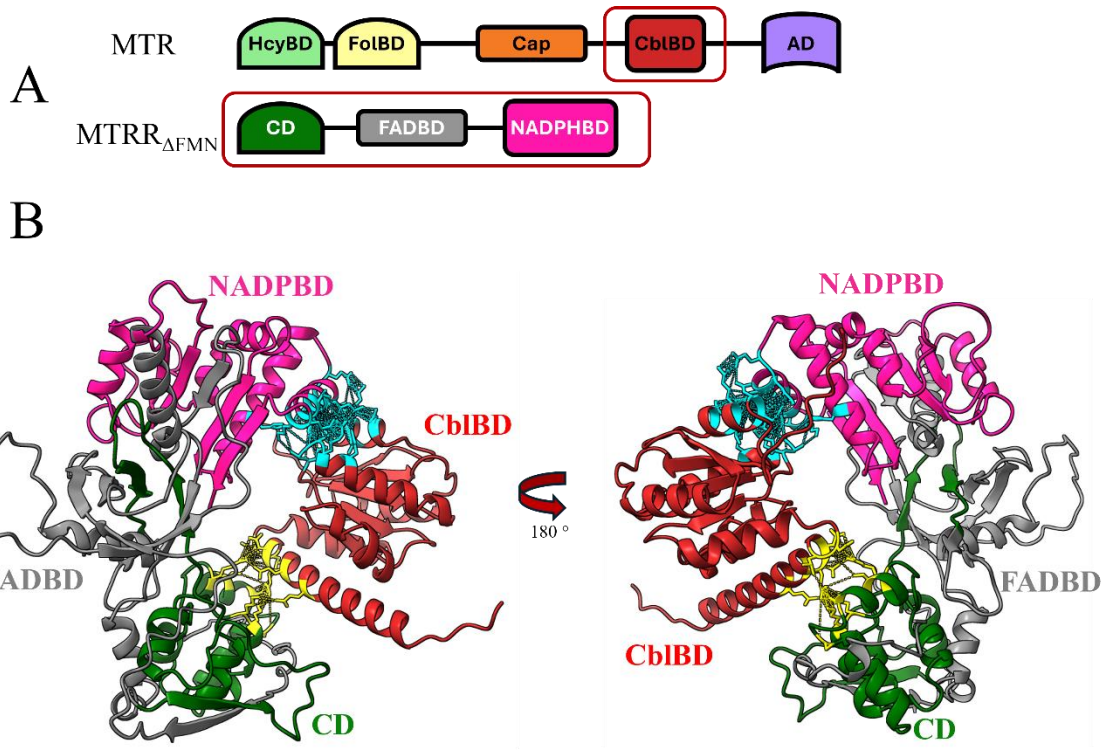

Figure 6.12: PICKLUSTER Interaction Analysis of the MTR-MTRR$_{\Delta FMN}$ Complex. A. The domains and their corresponding colours of MTR and MTRR$_{\Delta FMN}$ used for model generation are shown. A scheme of MTR domains in the activation conformation is also shown. Red rectangles represent the domains depicted in B. **B**. The PICKLUSTER software predicted two areas of interaction: one (yellow residues) between the CblBD (red) of MTR and the CD (green) of MTRR, and another (cyan residues) between the CblBD and the NADPHBD (pink) of MTRR. The FADBD of MTRR (grey) is also included for reference. All other domains of MTR have been omitted for clarity.

Further analysis of the interacting interface of the CD of MTRR (yellow in Figure 6.12) revealed the presence of three positively charged residues (Lys408, Lys409, and Arg445), while the long helix (α7) of the CblBD possesses negatively charged residues in the interacting region (Asp903, Glu904, Glu907, and Glu911). These residues may form salt bridges, stabilising the interaction. Moreover, the predicted model suggests the formation of hydrogen bonds and hydrophobic interactions in the same region.

The NADPHBD of MTRR also contains several charged residues (Glu702, Lys703, Lys708, Glu715), which may form salt bridges with charged residues located in the α3 and α4 helices of the CblBD (Asp813, Lys817, Glu846, and Arg849) (cyan in Figure 6.12). The presence of leucine residues in both the CblBD (Leu820) and NADPHBD (Leu704) further suggests possible hydrophobic interactions.

The electrostatic potential of the CblBD of MTR, and CD and NADPHBD of MTRR was analysed using ChimeraX to illustrate the global charges of these domains (Figure 6.13). This analysis allows the visualisation of residues with opposing charges in the interacting regions, providing strong evidence that the CblBD interacts electrostatically with the CD and NADPHBD of MTRR, which contributes to the stability of the MTR-MTRR complex, as discussed in Section 6.3.

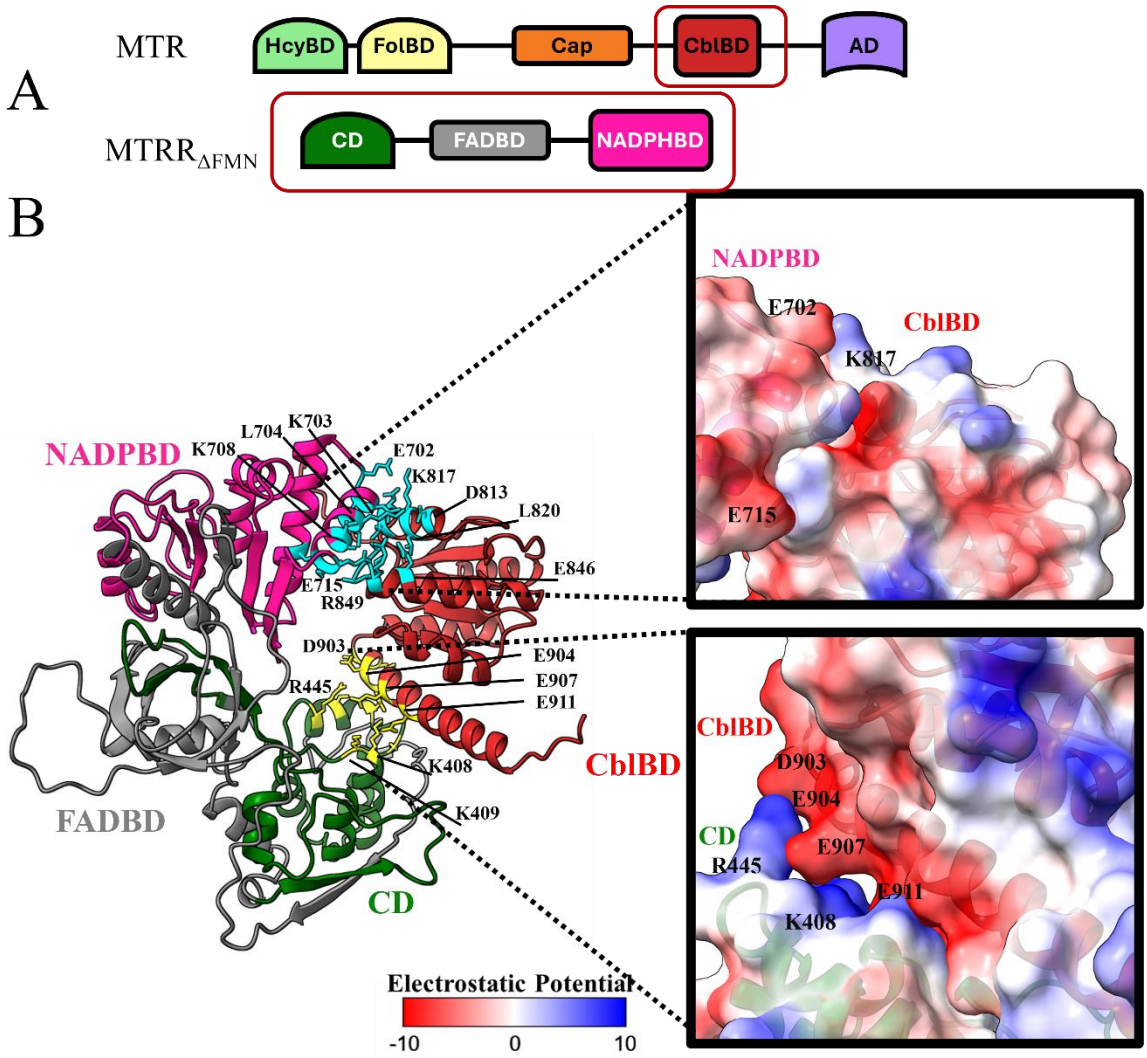

Figure 6.13: Electrostatic Potential of the Interacting Regions of MTR CblBD with MTRR$_{\Delta FMN}$ A. The domains and their corresponding colours of MTR and MTRR$_{\Delta FMN}$ used for model generation are shown. Red rectangles represent the domains depicted in B. **B**. The model of the MTR-MTRR$_{\Delta FMN}$ complex includes the CblBD (red) of MTR, along with the CD (green), FADBD (grey), and NADPHBD (pink) from MTRR$_{\Delta FMN}$. The interaction regions between the two proteins are highlighted in yellow and cyan, and the insets display the electrostatic potential of these interaction areas. All other domains of MTR have been omitted for clarity.

The additional interface between the MTR CblBD and the MTRR CD and NADPHBD, beyond the previously reported interaction between the MTR AD and the MTRR FMNBD (Wolthers, Basran et al. 2003), is not surprising. A 2006 study demonstrated that the presence of MTRR could enhance the uptake of Cbl by *apo* MTR by stabilising MTR (Yamada, Gravel et al. 2006). Furthermore, the co-expression and purification of the MTR-MTRR complex in this study showed that both proteins can interact in their recombinant full-length form without the presence of Cbl. Additionally, the interaction between MTRR and the MTR CblBD may be independent of the conformation of MTR, whereas the FMNBD can only transfer electrons to Cbl during the activation conformation, and the novel interactions seem to not be dependent on the conformation of MTR.

6.5.3 Modelling the MTR-MMADHC complex

In addition to interacting with MTRR, MTR also interacts with MMADHC, with Cys261 of the latter reported to be essential for Cbl delivery (Li, Mascarenhas et al. 2020, Mascarenhas, Guha et al. 2023). The interaction between MTR and MMADHC has been demonstrated *in vitro* using spectrometry (Mascarenhas, Guha et al. 2023) and *in vivo* via Dual Link Proximity assays (Bassila, Ghemrawi et al. 2017). In Section 6.4.2, this study demonstrates the transfer of Cbl to MTR via $MMADHC_{\Delta N}$. Since MMADHC transports Cbl to *apo* MTR, it is reasonable to assume that this interaction occurs with MTR in the activation conformation, which corresponds to the apoenzyme, as discussed in Section 6.4.1. Furthermore, MMADHC is expected to interact near the Cbl binding pocket on the CblBD of MTR in order to deliver Cbl.

With these considerations, the MTR-MMADHC complex model was predicted using AF2 Multimer. The sequence of the full-length MTR (Q99707.2) and MMADHC (Q9H3L0) were used to generate the models, which support the interaction between MMADHC and MTR near the Cbl binding pocket (Figure 6.14). The predicted model of MMADHC shown

in the figureis identical to the crystal structure of MMADHC$_{\Delta N}$ (Ser108–Asn296) previously determined (Froese, Kopec et al. 2015, Yamada, Gherasim et al. 2015, Li, Mascarenhas et al. 2020), with the only difference being the disordered N-terminus (Stucki, Coelho et al. 2012). This disordered region does not appear to interact with MTR and is omitted in all figures for better visualisation.

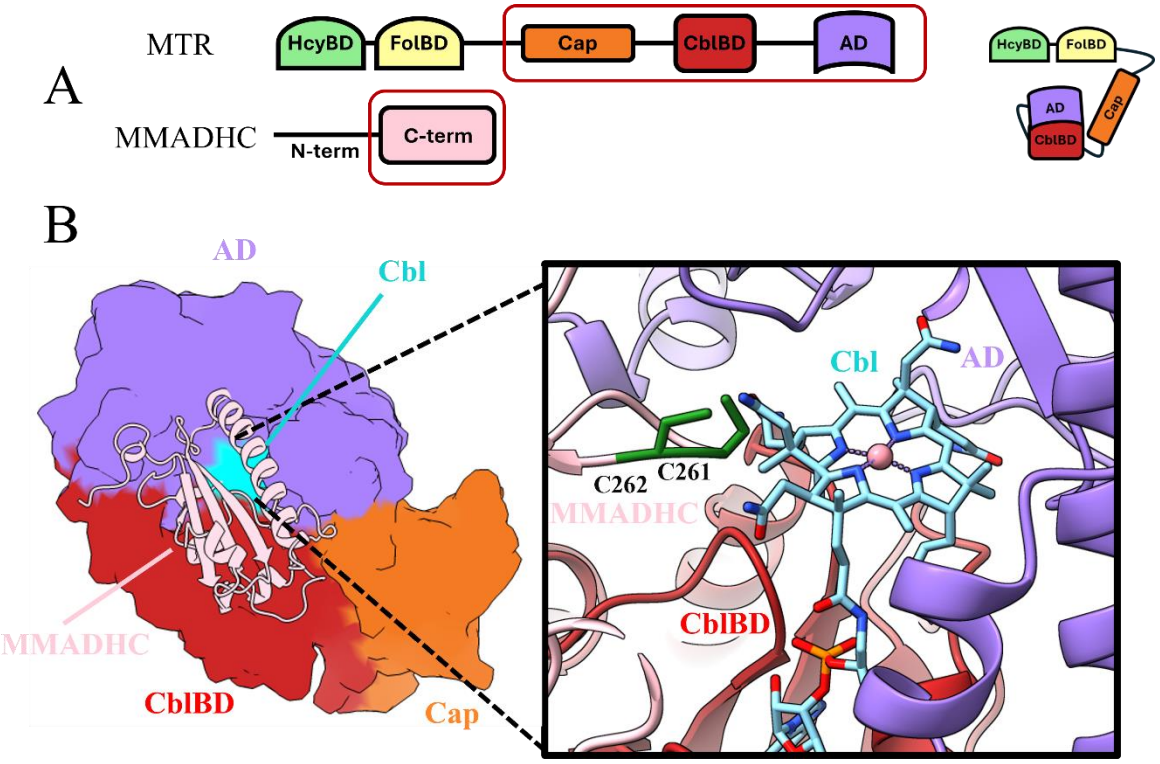

Figure 6.14: AlphaFold2 Multimer Prediction of the MTR-MMADHC Complex A. A schematic representation of the MTR and MMADHC domains used for model generation, with each domain colour-coded to match the model. A schematic of the MTR domains in the activation conformation is also provided. Red rectangles represent the domains depicted in B. **B**. The MTR-MMADHC complex model. The HcyBD and FolBD domains of MTR are omitted for clarity, although they function as a single unit in the predictions. Their positioning varies depending on the orientation of the loop connecting them to the Cap. The Cap (orange), CblBD (red), AD (purple), and Cbl (cyan) of MTR are represented as a volume, while MMADHC C-terminus (light pink) is displayed in cartoon form. The inset highlights the interaction between MMADHC and MTR. The green residues are Cys261 and Cys262 of MMADHC. The first 107 residues of MMADHC, which form a disordered loop, were also omitted for clarity in the figure. The predicted complex structure has an interface predicted template modeling (ipTM) score of 0.72.

All five models from the AF2 Multimer output show that MMADHC interacts with MTR by positioning its Cys261-Cys262 region to the Cbl-binding pocket (Figure 6.14). This positioning of MMADHC with respect to MTR is consistent with previous report that Cys261 is responsible for Cbl interaction (Mascarenhas, Guha et al. 2023). The predicted

model also presents MTR in the activation conformation, which was observed in the apoenzyme form of MTR (Chapter 5). As the interaction between MTR and MMADHC is required for Cbl delivery to *apo* MTR, the predicted structure further supports the structural evidence provided in Section 5.4.1, where *apo* MTR was found in the activation conformation.

Further analysis of the interface between MTR and MMADHC in the models was conducted using PICKLUSTER (Figure 6.15).

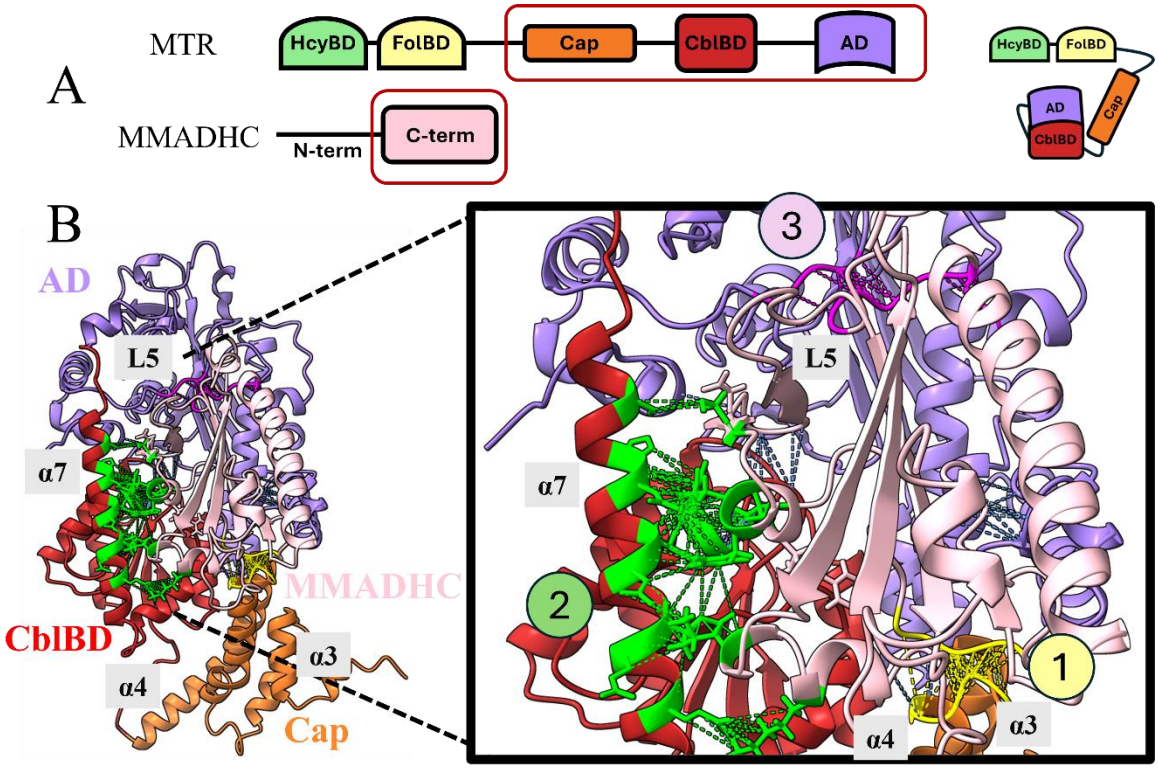

Figure 6.15: PICKLUSTER Analysis of the MTR-MMADHC Model A. A schematic representation of the MTR and MMADHC domains used for model generation, with each domain colour-coded to match the model. A diagram of the MTR domains in the activation conformation is also included. Red rectangles represent the domains depicted in B. **B**. The model of the complex displays the Cap (orange), CblBD (red), and AD (purple) of MTR, with MMADHC shown in light pink. The three interaction clusters are highlighted: yellow for the MMADHC-Cap interaction (1), green for MMADHC-CblBD (2), and pink for MMADHC-AD (3). For clarity, the HcyBD and FolBD domains of MTR, as well as the first 107 residues of MMADHC, were omitted from the figure, as these regions did not show any interactions.

PICKLUSTER revealed three main interfaces between MTR and MMADHC, each located within one of the three C-terminal domains of MTR. The first interface (labelled 1

in Figure 6.15B inset) involves MMADHC and the MTR Cap domain, where residues Tyr237, Thr238, Asn239, and Asn240 of MMADHC primarily form hydrogen bonds with Gly724, Ala725, Gly726, and Lys272 of MTR, positioned between the α3 and α4 helices of the four-helix bundle Cap.

For the second interface (labelled 2 in Figure 6.15B inset), electrostatic and hydrophobic interactions were identified between MMADHC and the long helix (α7) of MTR CblBD (suggested to interact with MTRR CD, in Section 6.5.2). The residues Leu156, Lys159, Glu247, Arg250, His251, and Ser255 from MMADHC interact with various residues in the long helix (α7), including three glutamates (Glu908, Glu911, and Glu912), two aspartates (Asp915 and Asp919), two serines (Ser864 and Ser923), and two histidines (His867 and His920).

In the third interface (labelled 3 in Figure 6.15B inset), MMADHC residues Thr187, Val188, Trp189, Ser190, and Glu191, which form a disordered loop (L5), interact with the AD of MTR through residues Arg1165, Leu1166, Tyr1227, Phe1228, and Ala1229, primarily forming hydrophobic interactions. All three interfaces depicted in the model are only feasible when MTR is in the activation conformation, in which MMADHC positions Cys261 and Cys262 towards the Cbl-binding pocket for co-factor loading. Although Cys261 and Cys262 are positioned close to Cbl in the model, as placed by AlphaFill, no direct interaction between MMADHC and Cbl ligand was predicted. This is expected, as the interaction between MTR and MMADHC occurs with MTR in its *apo* form, and therefore, no interaction between MMADHC and Cbl is anticipated.

When comparing the MTR-MTRR and MTR-MMADHC complexes, it is evident that both the FMNBD of MTRR and the MMADHC C-terminus interact with the same region in MTR (Figure 6.16). Although MTRR FMNBD and MMADHC do not interact with the same residues of MTR, they could not bind simultaneously, as they occupy the same region on MTR (Figure 6.16B).

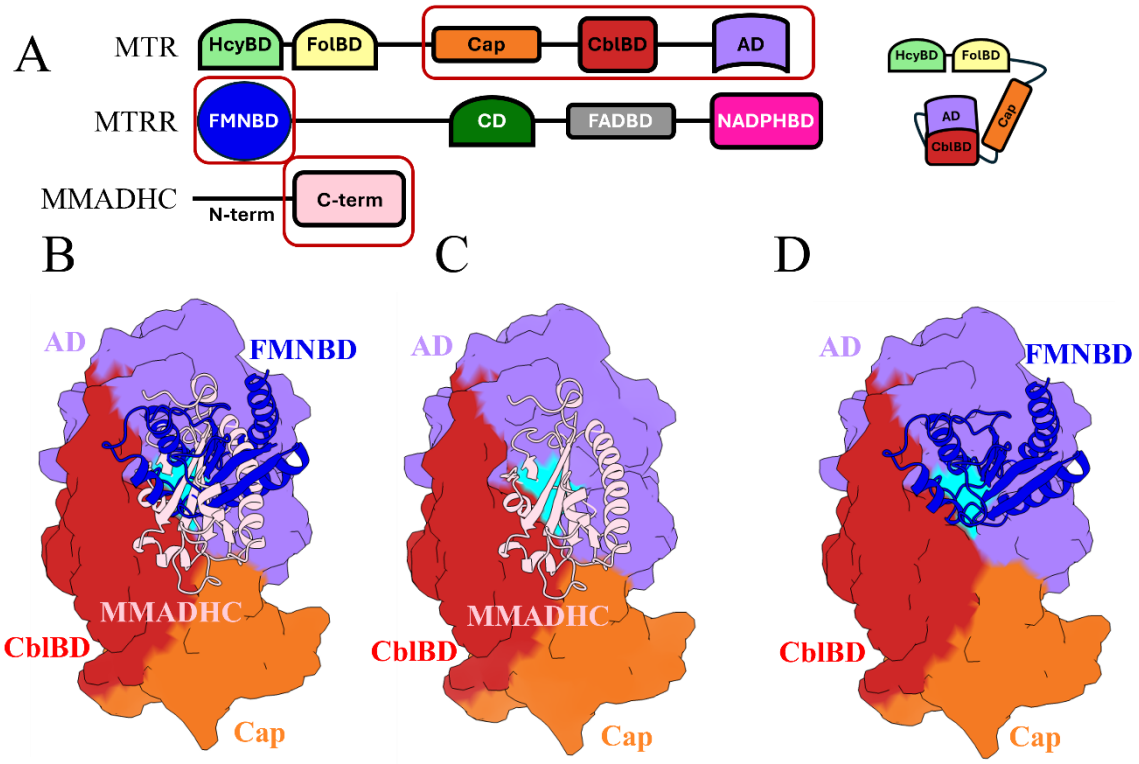

Figure 6.16: Comparison of the AlphaFold Multimer Models of the MTR-MTRR and MTR-MMADHC Complexes. A. A schematic representation of the MTR and MMADHC domains used for model generation, with each domain colour-coded to match the model. A schematic of the MTR domains in the activation conformation is also included. Red rectangles represent the domains depicted in B. **B**. Superimposition of the MTR-MTRR and MTR-MMADHC models, showing the MTR domains Cap (orange), CblBD (red), AD (purple), and Cbl (cyan) as a volume. The FMNBD of MTRR (blue) and MMADHC C-terminus (light pink) are also displayed. **C**. Model of the MTR-MTRR complex. **D**. Model of the MTR-MMADHC complex. In all models, the hinge, FADBD, CD, and NADPHBD of MTRR, the first 107 residues of MMADHC, and the HcyBD and FolBD of MTR were omitted for clarity.

The sharing of a common interface on MTR suggests that the interactions between MMADHC and MTR, and between MTRR and MTR, would be mutually exclusive in the models if both partner proteins bind to the same site on MTR. If MTRR and MMADHC interact with MTR only through the same region of MTR, the higher-order complex

composed of MTR, MTRR, MMACHC, and MMADHC observed in cells would not be possible with these models (Bassila, Ghemrawi et al. 2017).

Therefore, it would be valuable to predict a complex composed of MTR, MTRR, and MMADHC using AF2 Multimer, to determine whether MTR can interact simultaneously with both MTRR and MMADHC, as suggested by Bassila and colleagues.

6.5.4 Modelling the MTR-MTRR-MMADHC complex

MMADHC is not expected to deliver Cbl to MTR at the same time when MTRR needs to transfer electrons to re-activate Cbl. Cobalamin delivery occurs when Cbl is in the +1 oxidation state, which is reactive and can be remethylated by SAM without the aid of MTRR (Mascarenhas, Guha et al. 2023). This suggests that the interactions between MTR and MTRR, and between MTR and MMADHC, do not necessarily require a three-way complex to happen concurrently in the same spatiotemporal context. Nonetheless, evidence suggested that MTR, MTRR, and MMADHC may interact simultaneously in vivo, potentially forming a four-way complex with MMACHC (Bassila, Ghemrawi et al. 2017).

The full-length human sequences of MTR, MTRR, and MMADHC were used to generate a model of the three-way complex (Figure 6.17). All five models generated by AF2 Multimer showed virtually the same conformation, with some variation in the position of the N-half of MTR, as observed in sections 6.5.1.

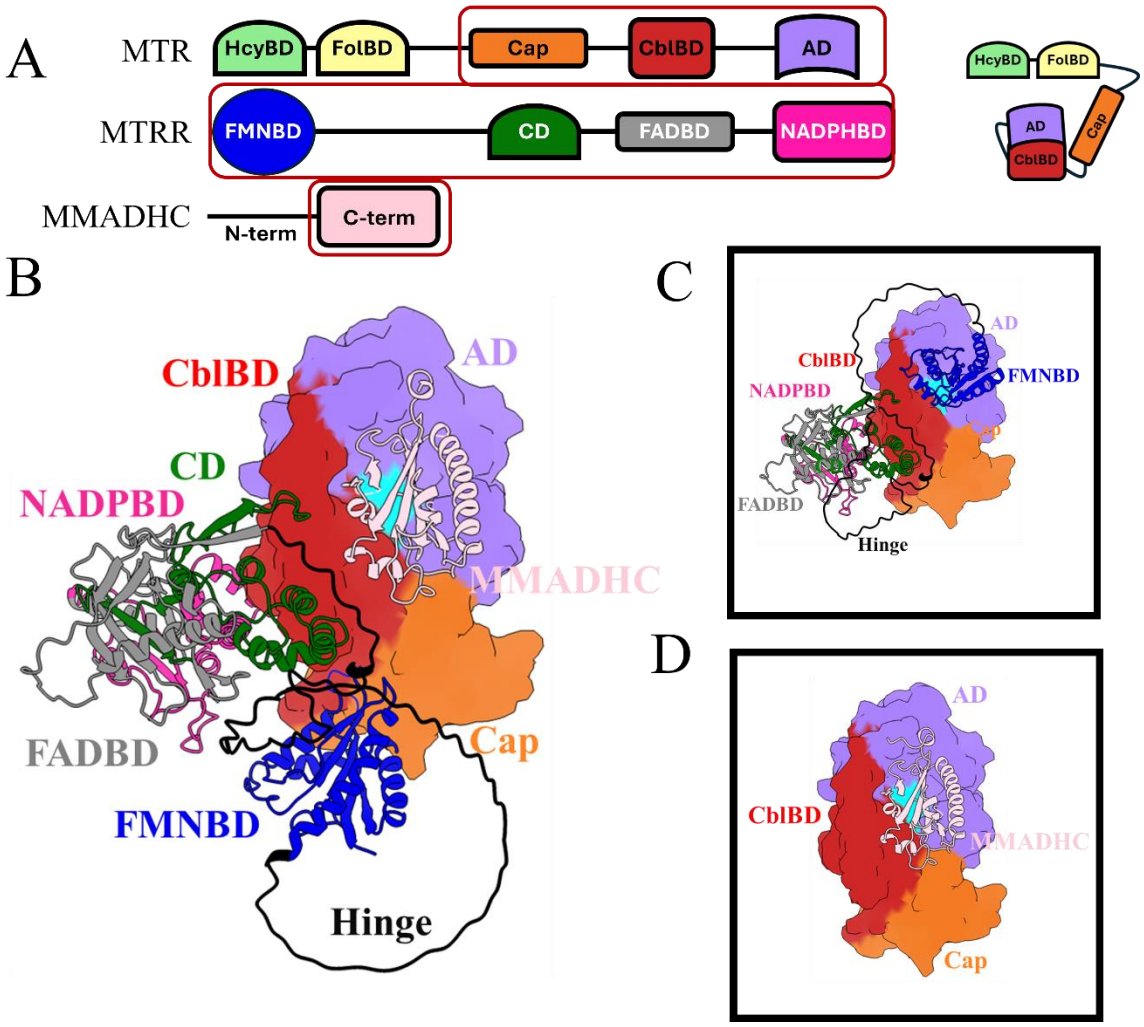

Figure 6.17: AlphaFold Multimer prediction of the complex MTR-MTRR-MMADHC. A. A schematic representation of the MTR, MTRR and MMADHC domains used for model generation, with each domain colour-coded to match the model. Red rectangles represent the domains depicted in B. **B**. Model of the complex MTR-MTRR-MMADHC, with MTR shown in volume, with the Cap (orange), CblBD (red), AD (purple), and Cbl (cyan). MMADHC C-terminus (light pink) and MTRR domains FMNBD (blue), FADBD (grey), CD (green), and NADPHBD (pink) are shown in cartoon. **C**. Model of the MTR-MTRR complex. D. Model of the MTR-MMADHC complex. In all models shown, the first 107 residues of MMADHC, and the HcyBD and FolBD of MTR were omitted for clarity. The predicted complex structure has an interface predicted template modeling (ipTM) score of 0.6.

All five generated models display the three proteins in a stoichiometric 1:1:1 complex (Figure 6.17), In the three-way model, CblBD of MTR only interacts with the CD and NADPHBD of MTRR but not the FMNBD, unlike in the MTR-MTRR model where the FMNBD, CD, and NADPHBD of MTRR interact with MTR (Section 6.5.2). The FMNBD of MTRR, connected to the rest of the protein via a long hinge, adopts different positions in all predicted three-way models due to its flexibility. However, none of the predicted models

show the FMNBD interacting with MTR in the same region as observed in the MTR-MTRR model.

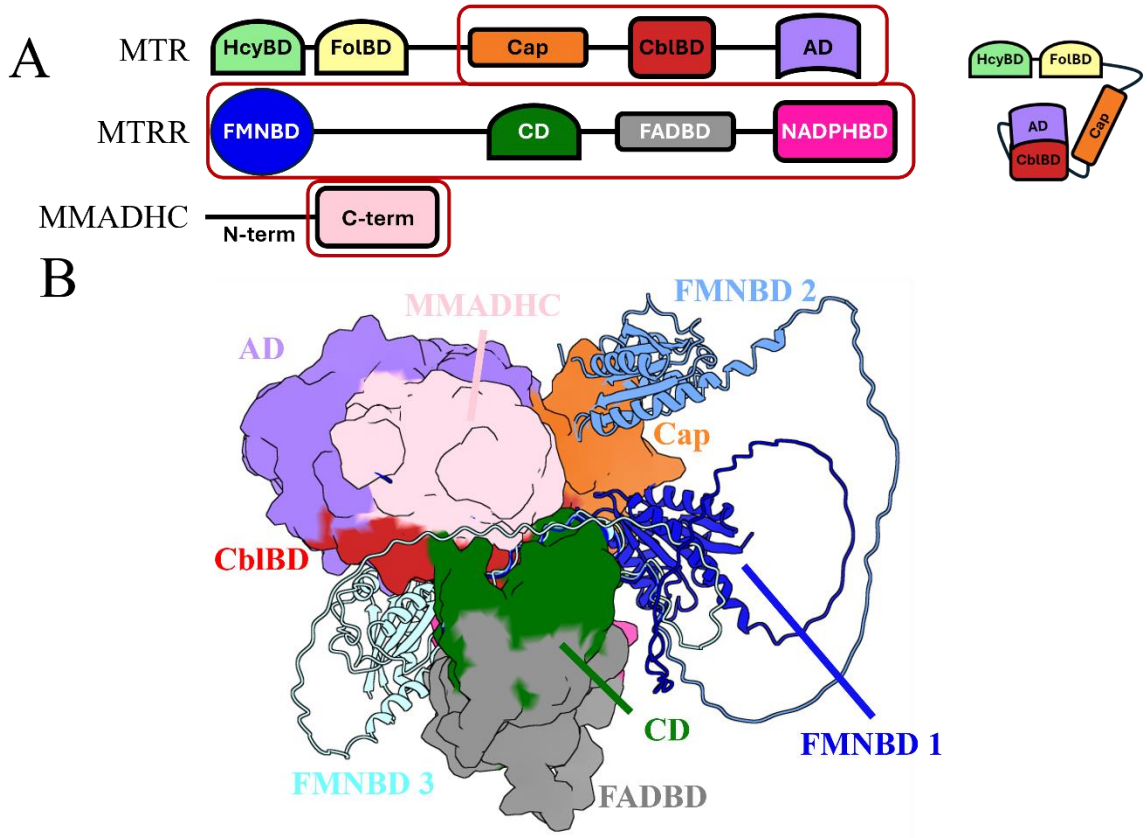

Figure 6.18: Comparison of Three AlphaFold Multimer Predictions of the MTR-MTRR-MMADHC Complex A. A schematic representation of the MTR, MTRR, and MMADHC domains used for model generation, with each domain colour-coded to match the model. Red rectangles represent the domains depicted in B. **B**. The model shows the MTR domains Cap (orange), CblBD (red), and AD (purple), the MMADHC C-terminus (light pink), and the MTRR domains FADBD (grey) and CD (green) represented as volumes. The only difference between the three volumes displayed is the positioning of the FMNBD and hinge of MTRR (shades of blue).

In the three-way model MMADHC interacts with MTR near the Cbl-binding pocket, in the same position as observed in the MTR-MMADHC model (Section 6.5.4). In all five predicted models, MTR adopts the activation conformation, with MMADHC positioned such that its Cys261 and Cys262 are near the Cbl-binding pocket, displaying three interaction regions with the Cap, CblBD, and AD of MTR, as observed in Figure 6.15. Regarding the interaction between MTR and MTRR in the three-way complex, the previously identified novel interaction (Section 6.5.2) remains unchanged, with the CD of

MTRR interacting with the long helix (α7) of the CblBD, and the NADPHBD of MTRR interacting with the CblBD through the α3 and α4 helices. The only difference between the two-way MTR-MTRR complex and the MTR-MTRR-MMADHC complex is the positioning of the FMNBD, which cannot be near the Cbl-binding pocket due to the presence of MMADHC. As a result, the FMNBD of MTRR adopts various positions across the generated models, suggesting that when MMADHC is bound to MTR, the FMNBD of MTRR does not interact with MTR. In this scenario, the interaction between MTR and MTRR occurs solely through the CD and NADPHBD of MTRR.

The three-way, MTR-MTRR, and MTR-MTRR$_{\Delta FMN}$ complex model support the interaction between CD and NADPHBD of MTRR and CblBD of MTR in addition to the FMNBD of MTRR with CblBD. The three-way complex model also supports the possibility of a supramolecular complex in vivo, as it shows that MTRR and MMADHC can interact with MTR simultaneously, given that MTRR has more than one interacting interface with MTR.

6.6 Pulldown studies of MTR complexed with MTRR and MMADHC$_{\Delta N}$

After demonstrating that AF2 Multimer can support a three-way complex formed by MTR, MTRR, and MMADHC, recombinant protein pull-downs were attempted to experimentally verify three-way complex formation. The recombinant MTR and MMADHC$_{\Delta N}$ used were the same as those in Section 6.4. Since the MTRR construct used in other parts of this study was co-expressed with MTR (Section 6.2), a new MTRR expression construct was generated for the pulldown studies here.

6.6.1 Expression and purification of full-length MTRR

Full-length MTRR, incoporated with a TEV-cleavable C-terminal His_{10} tag (Figure 6.19A), was expressed in *Sf9* insect cells. The expression followed the standard procedures outlined in Section 2.1.2, using 10 L of cells at a concentration of $1x10^6$ cells/mL. After 72 hours of expression, the cells were harvested and purified as described in Section 2.1.3, using Ni-NTA affinity purification followed by gel filtration with a 10/300 Superdex 200 chromatography column. The elution fractions were collected and analysed by SDS-PAGE (Figure 6.19B-D).

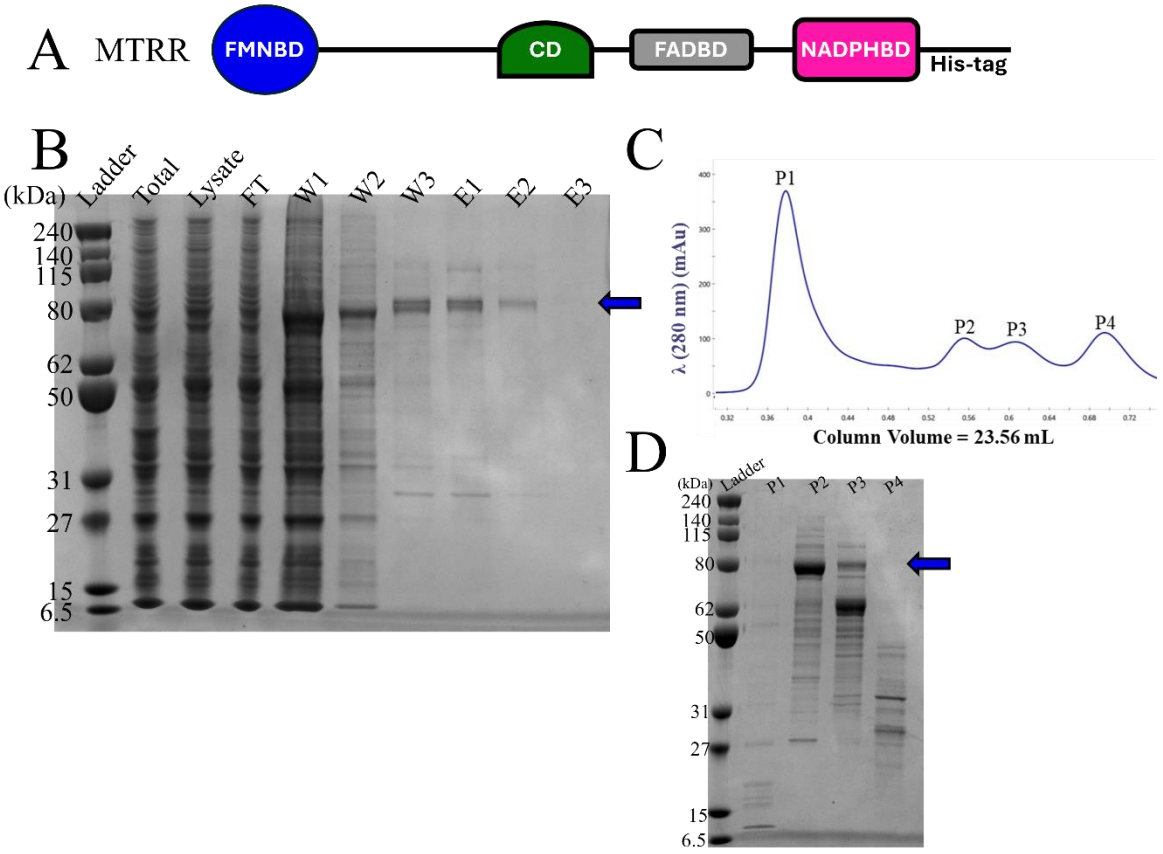

Figure 6.19: Purification of Full-Length MTRR by Ni-NTA Chromatography Followed by Gel Filtration Chromatography A. A schematic representation of the full-length MTRR domains and the location of the TEV-cleavable His_{10}-tag. **B**. SDS-PAGE of the Ni-NTA affinity purification, where each lane represents a step in the purification process. From left to right: protein standard ladder, total lysate, flow-through (FT), washes 1-3 (W1-W3), and eluates 1-5 (E1-E5). The blue arrow indicates the presence of protein corresponding to the molecular weight of MTRR (80 kDa). **C**. Chromatogram from the Superdex S200 10/300 column showing the four peaks observed at 280 nm absorbance, labelled P1-P4. **D**. SDS-PAGE analysis of samples from each peak (P1-P4), indicating the presence of full-length MTRR in peak 2 and predominantly degraded MTRR in peak 3.

As shown in Figure 6.19D, most protein in gel filtration peak P2 corresponds to full-length MTRR, while peak P3 likely contains degraded MTRR, due the intensity of the band, which can be assumed to be recombinant protein. The high flexibility of certain regions of MTRR (e.g. the long hinge connecting the FMNBD and the rest of the protein) makes it an unstable protein, as seen previously (Wolthers, Basran et al. 2003) and in the present study (Section 6.2), where no free MTRR was observed unless it was forming a complex with MTR. This instability and propensity for degradation in full-length MTRR resulted in a very low yield after purification (182 μg purified protein from 10 L cell culture). Nevertheless, it was sufficient to proceed with the pulldown experiments involving MTR and $MMADHC_{\Delta N}$.

6.6.2 Testing the three-way complex formation by pulldown

The recombinant proteins were used in a pulldown assay to test whether MTR (purified as per Section 3.2), MTRR (purified as per Section 6.6.1), and $MMADHC_{\Delta N}$ (previously purified by the group as stated in Section 6.4) can form a three-way complex. Briefly, FLAG-tagged MTR was immobilised on the ANTI-FLAG® M2 Affinity Gel, and MTRR and $MMADHC_{\Delta N}$ were added simultaneously to the MTR-loaded resin. After incubation, the resin was washed, and the bound proteins were eluted and analysed via Western Blot, using commercially available polyclonal antibodies against human MTR, MTRR, or MMADHC. The antibodies, obtained from Antibodies.com, were raised against full-length human MTR (A83766), residues 50-230 of MTRR (A91274), and full-length MMADHC (A89274).

Neither recombinant MTRR nor $MMADHC_{\Delta N}$ have a FLAG tag in their expression constructs, so their presence in the elution from the ANTI-FLAG® resin would indicate interaction with FLAG-tagged MTR. As shown in Section 6.3.2, when MTRR interacts with MTR, it is retained on the FLAG resin via the interaction of MTR with both MTRR and the resin, even though MTRR itself does not have affinity for the resin. Due to the low quantity

of proteins used in the pulldown (10 µg each), antibodies against each protein were utilised to detect them in the elution fraction by Western blotting.

The pulldown was conducted as described in Section 2.6, with appropriate controls including the incubation of each protein individually on the resin and the two-way complexes of MTR-MTRR or MTR-MMADHC$_{\Delta N}$. Additionally, two different samples of MMADHC$_{\Delta N}$ were used in the pulldown: the MMADHC$_{\Delta N}$-Cbl (holo) and the as-purified *apo* form (MMADHC$_{\Delta Napo}$). The use of both MMADHC$_{\Delta N}$ samples, with and without Cbl, aims to determine whether the presence of the cofactor influences protein-protein interactions. The pulldown was completed, and the Western Blots are shown in Figure 6.14.

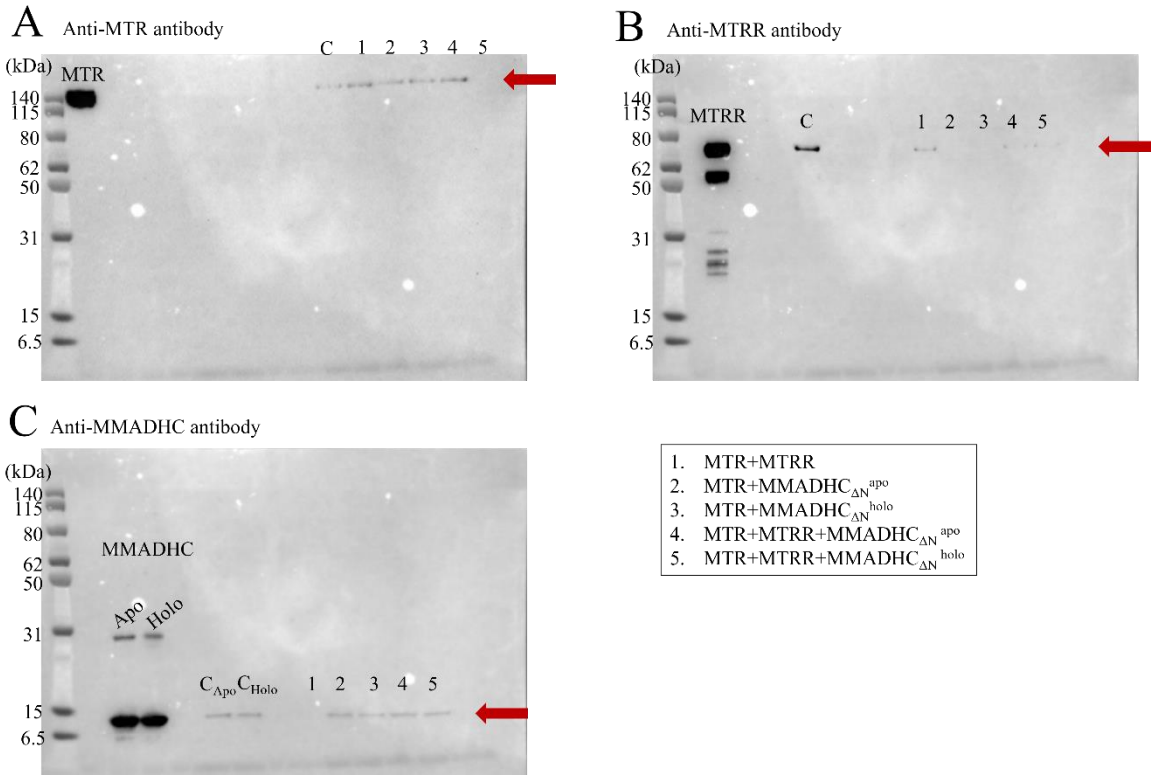

Figure 6.20: Western Blots of the Pulldown Assay Using Antibodies Against MTR, MTRR, or MMADHC. Each western blot using anti-MTR (**A**), anti-MTRR (**B**), and anti-MMADHC (**C**) antibodies. Each blot displays the protein ladder, followed by the input (labelled as the proteins name) of each protein used for the pulldown. The lane labelled C (control) on each blot represents the elution from the resin incubated only with the protein shown in the input. C$_{Apo}$ represents the control sample for MMADHC$_{\Delta Napo}$, and C$_{Holo}$ represents the control sample for MMADHC$_{\Delta N}$-Cbl. The following lanes (1-5) represent the elutions of samples composed of the following mixtures: 1. MTR+MTRR, 2. MTR+MMADHC$_{\Delta Napo}$, 3. MTR+MMADHC$_{\Delta N}$-Cbl, 4. MTR+MTRR+MMADHC$_{\Delta Napo}$, and 5. MTR+MTRR+MMADHC$_{\Delta N}$-Cbl. Representative western blots from two independent experiments for ach antibody used.

The Western blots revealed the presence of MTRR (Figure 6.20B, lanes 1,4, and 5) and MMADHC$_{\Delta N}$ (Figure 6.20C, lanes 2-5) in the elution of the pulldowns, which contains MTR (lanes 1-5); however, these proteins were also observed in the elution when they were incubated with the resin without MTR (Figure 6.20B and C, lane C). This indicates that MTRR and MMADHC$_{\Delta N}$ interacted with the FLAG resin even when it is not pre-loaded with FLAG-tagged MTR. As MTRR and MMADHC$_{\Delta N}$ were detected both in their control lanes and in the samples incubated with MTR, it is not possible to conclusively determine whether all three proteins are interacting, based on this experiment.

The two-way interactions between MTR and MTRR, and between MTR and MMADHC$_{\Delta N}$, have already been demonstrated through Native Blue-PAGE and Cbl transfer assays in Sections 6.3 and 6.4, respectively. AF3 modelling in Section 6.5 suggest a novel interaction between MTR (CblBD) and MTRR (CD and NADPHBD), in addition to the Cbl-FMNBD interaction, further supporting the presence of higher-order complex formation (Bassila, Ghemrawi et al. 2017). However, the pulldown assay aimed at demonstrating the possible formation of this complex was inconclusive.

6.7 Discussion

Advances in protein structure prediction, particularly through AlphaFold, have led to the development of AF2 Multimer, which facilitates the prediction of protein complex structures. The integration of AF2 Multimer with biochemical studies provides deeper insights into protein-protein interactions (Akdel, Pires et al. 2022, Bryant, Pozzati et al. 2022).

The interactions between MMACHC, MMADHC, MTR, and MTRR are essential for the coordination of the cytosolic branch of the intracellular cobalamin metabolism. Although individual components of this complex have been described structurally, a

comprehensive understanding of these protein interactions remains incomplete. This chapter explored these interactions using MTR as the focal point within this network, employing both biophysical and computational approaches.

6.7.1 Structural basis of Cbl delivery - MTR-MMADHC interaction

The recombinant $MMADHC_{\Delta N}$ used in this chapter's work lacks the N-terminus (residues 1-123), a mitochondrial leader sequence that is not essential for Cbl delivery to MTR (Stucki, Coelho et al. 2012). Cbl binding to MMADHC occurs through its upper axial ligand, interacting with Cys261 of MMADHC, forming a Co-S coordination and generating thiolato-cob[III]alamin (Li, Mascarenhas et al. 2020). In this study, incubating $MMADHC_{\Delta N}$ with OHCbl revealed a change in the UV-Vis spectrum of Cbl, indicating a transition from OHCbl to thiolato-cob[III]alamin, confirming Cbl binding to $MMADHC_{\Delta N}$.

The interaction between $MMADHC_{\Delta N}$ and Cbl has been previously demonstrated, and a structural depiction of this interaction was obtained via X-ray crystallography (PDB: 6X8Z) (Li, Mascarenhas et al. 2020). The absence of the N-terminal residues (1-123) in the $MMADHC_{\Delta N}$-Cbl structure confirms that this region is not necessary for Cbl loading, and AF3 predictions (Section 6.5.3) identified this region as a disordered loop, likely representing a mitochondrial leader sequence required for Cbl delivery to the mitochondrial branch of Cbl metabolism (Stucki, Coelho et al. 2012). These findings agree with observations that mutations in the N-terminal region of MMADHC result only in AdoCbl deficiency (mitochondrial branch of Cbl metabolism) but do not affect MetCbl (cytosolic branch of Cbl metabolism) (Jusufi, Suormala et al. 2014).

Further incubation of $MMADHC_{\Delta N}$-Cbl with recombinant MTR demonstrated $MMADHC_{\Delta N}$'s ability to transfer Cbl to full-length MTR *in vitro*, using gel filtration chromatography to monitor absorbance at 280 and 350 nm. Detecting the Cbl transfer from

MMADHC$_{\Delta N}$ to MTR is also possible by tracking UV-Vis spectral changes, as Cbl transitions from thiolato-cob[III]alamin to cob[II]alamin, causing a shift in spectral peaks (Mascarenhas, Guha et al. 2023). Besides the size exclusion experiments, MTR was incubated with MMADHC$_{\Delta N}$-Cbl, and the UV-Vis of the reaction was monitored. However, due to the limited availability of full-length MTR, the amount of Cbl transferred to MTR was insufficient to produce significant spectral changes.

The approach of monitoring absorbance of 280 and 350 nm confirms the presence of both Cbl and MTR in the same UV-Vis peak, but it could not verify whether MTR and MMADHC$_{\Delta N}$ were still interacting. To overcome this limitation, alternative assays such as Surface Plasmon Resonance (SPR) could be employed to measure the association/dissociation kinetics of MMADHC$_{\Delta N}$ with immobilised MTR (Douzi 2017). This binding assay could also help determine whether the interaction between these two proteins is transient and whether they can still interact in the absence of Cbl. Although chromatography did not provide comprehensive data on the MTR-MMADHC$_{\Delta N}$ interaction, it confirmed the formation of the complex and the successful transfer of Cbl to MTR.

Since the MTR-MMADHC$_{\Delta N}$ complex is formed for Cbl delivery, the interaction interface between these proteins likely involves Cys261 of MMADHC, which interacts with Cbl (Li, Mascarenhas et al. 2020), and the Cbl-binding pocket of MTR. This interface was examined by generating models using AF2 Multimer (Section 6.5.3, Figure 6.14). All models of the MTR-MMADHC complex display Cys261 and Cys262 of MMADHC oriented towards the Cbl-binding pocket of MTR, confirming the accuracy of the model, since these residues are essential for the Cbl loading.

Cys261 and Cys262 of MMADHC did not show interactions with other MTR residues, which is consistent with previous data showing that Cys261 interacts with Cbl

during the MTR interaction, while Cys262 is responsible for displacing and depositing Cbl onto MTR by forming a disulphide bond with Cys261 (Mascarenhas, Guha et al. 2023). The stabilisation of the MTR-MMADHC complex for Cbl delivery is likely mediated by three distinct interaction regions identified in AF3 generated model in this study (Figure 6.15). Among them, the primary interface is between MMADHC and the AD of MTR, occurring near Cys261 and Cys262.

Cys261 and Cys262 of MMADHC are situated between two antiparallel β-strands, a conserved region of the protein (Figure 6.21). Conservation analysis of human MMADHC, performed using ConSurf (Ashkenazy, Abadi et al. 2016), reinforces the importance of these residues in Cbl delivery and further supports the hypothesis that the AF2 Multimer model likely reflects the configuration of the MTR-MMADHC complex. The conservation analysis was conducted using 150 orthologous genes from various eukaryotic species.

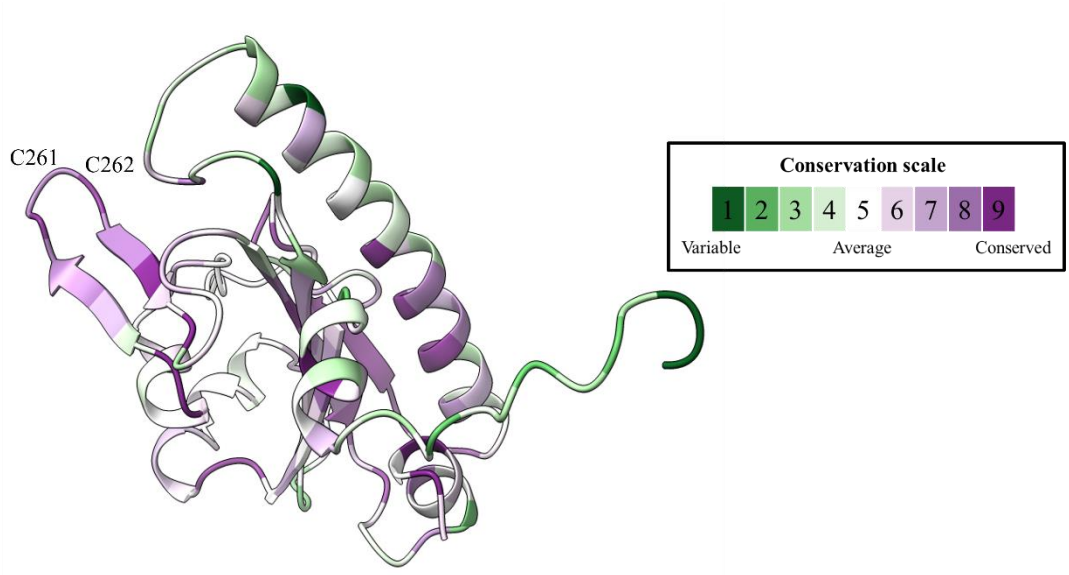

Figure 6.21: MMADHC$_{\Delta N}$ Conservation Analysis by ConSurf. The structure of MMADHC$_{\Delta N}$ is colour-coded according to the conservation scale, with the residues Cys261 and Cys262 highlighted in the structure. These residues are in a conserved region of the protein.

Besides the positioning of MMADHC relative to MTR, another crucial observation from the MTR-MMADHC model is the conformation of MTR. Given that the interaction

between these two proteins occurs for Cbl delivery, MTR is expected to be in its *apo* state. The structure of *apo* human MTR determined in this study, as described in Chapter 5 Section 5.4.1, shows that it adopts the activation conformation, in which the AD and CblBD are in close contact. In all models generated for the MTR-MMADHC complex, MTR is consistently positioned in the activation conformation. This conformation, when considered alongside the structure of *apo* MTR described in Chapter 5, further supports the reliability of the generated models.

It is logical that the conformation of MTR, when interacting with MMADHC, would be the activation conformation. This is because the activation state is the only conformation in which the Cbl ligand and Cbl-binding pocket are accessible. In other conformations, such as the catalytic or resting states, this region of MTR remains inaccessible to protect Cbl from oxidation, as shown in *E. coli* and *T. filiformis* bacterial structures (Drennan, Huang et al. 1994, Watkins, Wang et al. 2023). Additionally, when Cbl is loaded onto MTR, it undergoes a change from the +3 to the +1 oxidation state, which is highly reactive and primed for remethylation (Mascarenhas, Guha et al. 2023). In the +1 oxidation state, Cbl can be rapidly remethylated by SAM, as the AD is closely interacting with the CblBD. Moreover, the activation conformation allows MTR to accept cob[I]alamin and methylate it before it is oxidised to cob[II]alamin, thereby preventing the unnecessary need for the reactivation cycle.

To date, no interacting residues between MTR and MMADHC have been reported as key sites for missense mutations associated with pathogenic variants. While these specific residues are not linked to pathogenic *CblD* (MMADHC) or *CblG* (MTR) variants, understanding the potential interacting regions between these proteins may help to elucidate the basis of this interaction.

6.7.2 The molecular basis of the reactivation complex - MTR-MTRR interaction

The interaction between MTR and MTRR is essential for the reactivation cycle of MTR (Olteanu and Banerjee 2001). For this interaction to occur, the FMNBD of MTRR must be positioned near the Cbl ligand in the CblBD, which needs to be accessible. In the activation conformation, the Cbl is exposed to solvent, as the CblBD interacts with the AD, and the Cap is in the Cap-off state, facilitating electron transfer from the FMNBD to Cbl (Koutmos, Datta et al. 2009).

The expression of a plasmid construct containing both MTR and MTRR enabled the purification of a stable complex without the addition of Cbl (Section 6.3). The MTR-MTRR complex model generated by AF2 showed that in addition to the expected interaction between the MTRR FMNBD and MTR CblBD, which occurs in the same CblBD interface as the MMADHC-MTR interaction (Figure 6.16), the other domains of MTRR maintained a consistent position across all generated models, suggesting possible additional interactions with MTR.

The new models, generated using the $MTRR_{\Delta FMN}$ sequence, suggest that MTRR may interact with MTR CblBD through its CD and NADPHBD, interactions which have not been previously described (Section 6.5.2). These two additional interfaces with the MTR CblBD could potentially enhance the stability of the MTR-MTRR complex, and also allows it to exist in conformations of MTR other than the activation state, as it would not be dependent solely on the Cbl-binding pocket-FMNBD interaction.

In all MTR-MTRR models, the FMNBD of MTRR is positioned away from the other domains of MTRR, positioned close to CblBD of MTR. MTRR is present in most of eukaryotes, and belongs to the diflavin reductase family (Olteanu and Banerjee 2001), known to adopt closed or open conformations. These proteins can transfer electrons

sequentially from NADPHBD to FADBD, then to FMNBD, and finally to the final acceptor (Murataliev, Feyereisen et al. 2004). The electron flow within this protein family involves conformational changes between the closed state (unreactive), where all domains are juxtaposed, and the FMNBD is buried from solvent (Garcin, Bruns et al. 2004), and the open state (reactive), where FMNBD can transfer the electron to the final acceptor (Meints, Gustafsson et al. 2011).

Although a full-length structure of MTRR is not yet available, both open and closed conformations can be modelled using AF2. Of 25 models generated, three of them show the protein in what appears to be a closed conformation (Figure 6.22C), and the remaining 22 in an open conformation (Figure 6.22B).

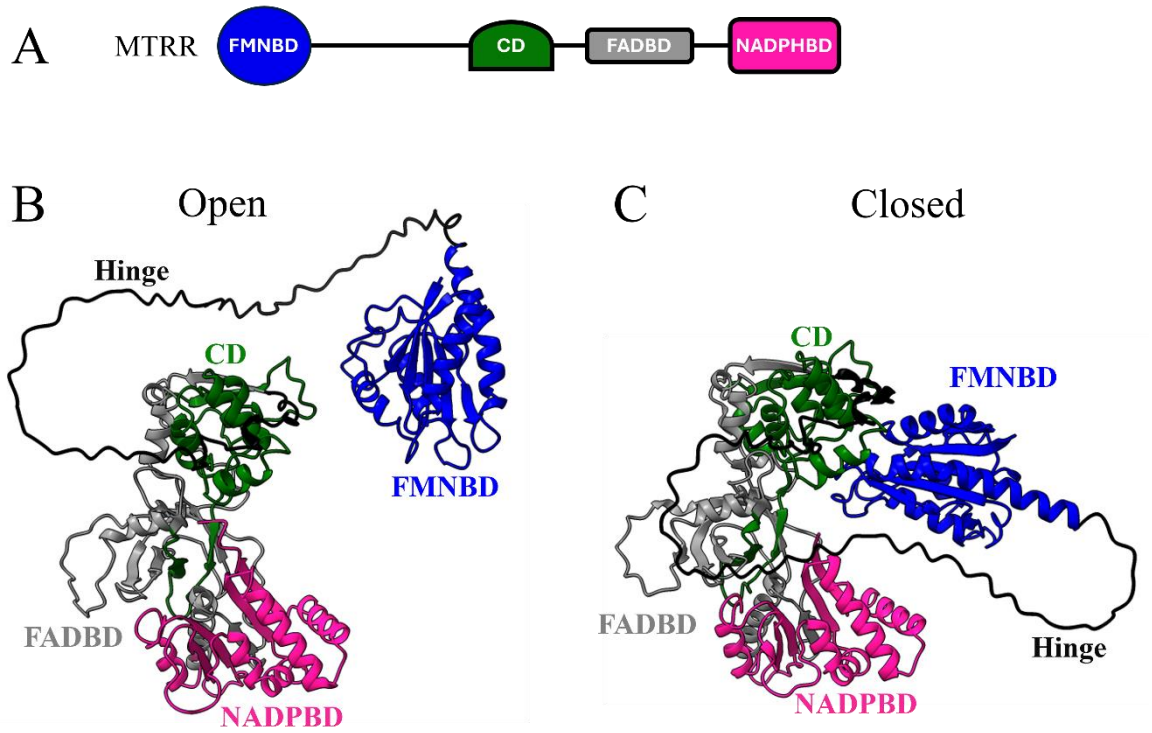

Figure 6.22: Open and Closed conformations of MTRR predicted by AlphaFold A. Schematic cartoon of the MTRR domains, showing the FMNBD (blue), CD (green), FADBD (grey), and NADPHBD (pink). **B**. Open conformation of MTRR, with the FMNBD positioned further from the other domains of the protein. **C**. Closed conformation of MTRR, showing the FMNBD near the other domains.

MTRR exhibits a slower conformational switch between open and closed states compared to other diflavin reductases, such as neuronal nitric oxide synthase and endothelial nitric oxide synthase (Haque, Bayachou et al. 2014). This slower rate is primarily attributed to the size of the hinge. The hinge in MTRR is significantly longer among the diflavin reductases (90 residues vs 14-25 residues in the other enzymes), which necessitates larger structural movements to transition between closed and open conformations. Additionally, MTR undergoes inactivation less frequently than other final acceptors, thus MTRR needs to perform the electron transfer less frequently (Jarrett, Huang et al. 1998).

The longer hinge of MTRR is essential to facilitate the interaction between its FMNBD and MTR, particularly given the novel interactions identified between MTR and MTRR CD and NADPHBD (Section 6.5.2). Since MTRR predominantly exists in the open conformation (approximately 72% in solution) (Haque, Bayachou et al. 2014), stabilisation of the MTR-MTRR complex would initially occur via interactions between the MTR CblBD and the MTRR CD and NADPHBD, with the MTRR FMNBD engaging with MTR CblBD only during the reactivation cycle to donate an electron to Cbl. The pre-existing interaction of MTRR CD and NADPHBD with MTR before the reduction of cob[II]alamin suggests that this reduction can be executed more efficiently when required, as the MTR-MTRR complex is already stabilised by the interaction of CblBD with CD and NADPHBD, thereby enhancing the efficiency of the reactivation cycle.

The interaction between the CD and NADPHBD of MTRR with the CblBD of MTR enables the MTR-MTRR complex to exist beyond the activation conformation of MTR. This stable complex agrees with biophysical data, which shows that MTRR can stabilise *apo* MTR and increase Cbl incorporation into the enzyme. This is feasible only if the MTR Cbl-binding pocket is not obstructed by the MTRR FMNBD, indicating the necessity for another

interaction region between the two proteins to maintain integrity of the complex (Yamada, Gravel et al. 2006).

MTRR is also known to be a selective partner of human MTR since it does not interact with bacterial MTR, despite their structural homology (Yamada, Gravel et al. 2006). To investigate the lack of cross-reactivity between human MTRR and bacterial MTR, the conservation of the CblBD was analysed alongside the CD and NADPHBD of MTRR using ConSurf (Figure 6.23), which compared 150 genes orthologues for each MTR and MTRR.

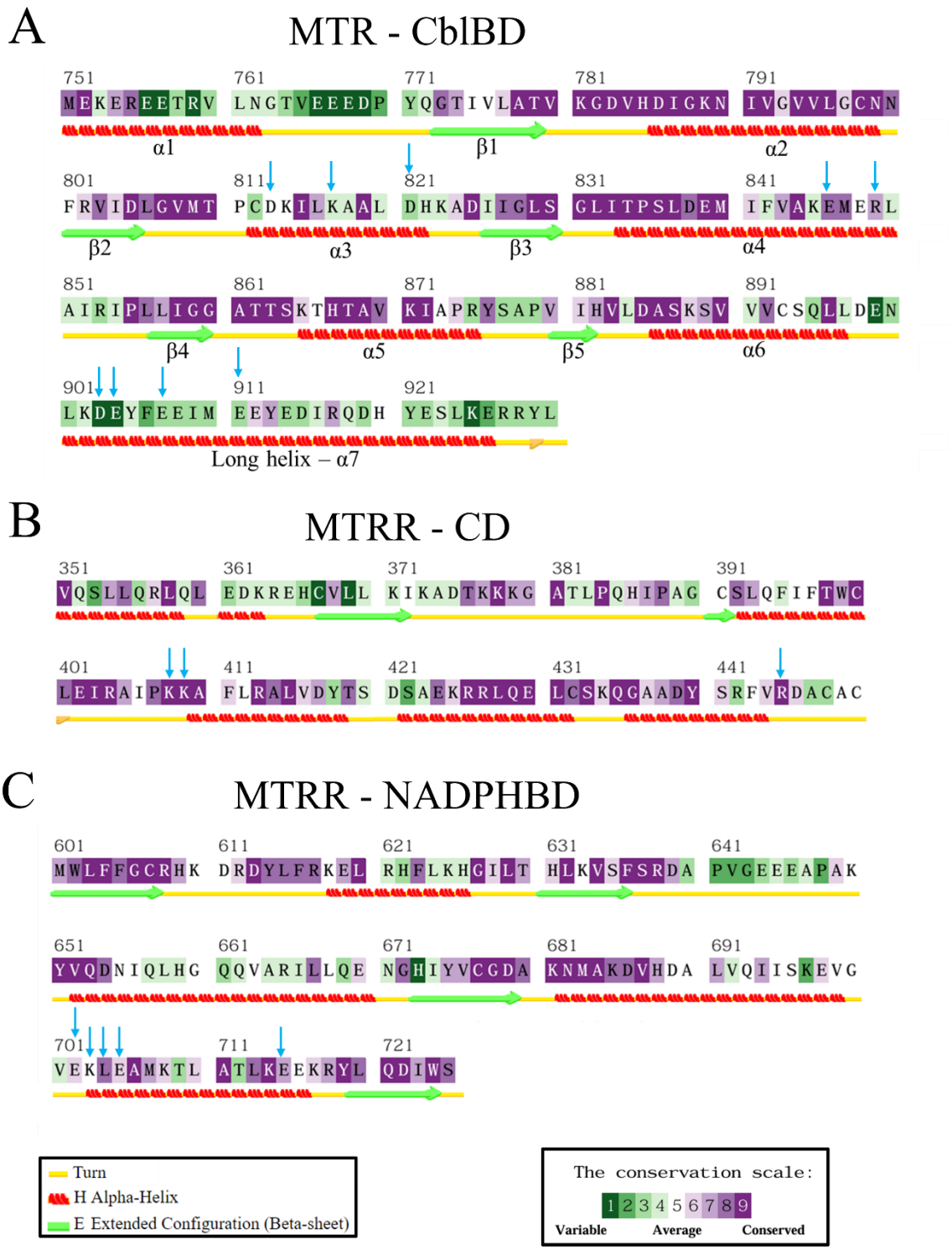

Figure 6.23: Conservation analysis of MTR CblBD and MTRR CD and NADPHBD performed on ConSurf, including the secondary structure of the residues A. Conservation analysis of the MTR CblBD, with blue arrows pointing to the residues shown to interact with the CD and NADPHBD of MTRR. **B**. Conservation of the residues in the CD of MTRR, with blue arrows indicating the residues that interact with the CblBD. **C**. Conservation of the residues in the NADPHBD of MTRR, with arrows indicating the residues interacting with the CblBD of MTR.

Although human and bacterial MTR are structurally similar, (Sections 1.6.1 and 1.6.2), the specific MTR residues involved in the interaction with the CD and NADPHBD of MTRR are not conserved between human and bacteria (Figure 6.23A). Additionally, the residues of MTRR involved in this interaction exhibit some degree of conservation among eukaryotes (Figure 6.23B and C). This suggests that the selectivity of MTRR to human MTR (i.e., it does not reactivate bacterial MTR) may be due to subtle changed residues in MTR that shows no conservation amongst bacterial MTR, as most of the human CblBD is conserved and structurally similar to its bacterial orthologues.

The regions of the MTR CblBD that do not display high identity with the bacterial orthologues, may relate not only to the specificity of MTRR activity but also to MMADHC binding. The long helix α7 in the CblBD, which serves as an interface for both MTRR (Section 6.5.2) and MMADHC (Section 6.5.3), is the least conserved region of MTR when compared to bacterial MTR. This lack of conservation between prokaryotes and eukaryotes, coupled with the long helix's interactions with both MTRR and MMADHC, implies that bacterial MTR may not interact with these proteins, making the MTR-MTRR and MTR-MMADHC complexes likely specific to the protein in chordates, given the absence of MTRR and MMADHC genes in prokaryotes. Therefore, future studies on these complexes should consider that bacterial MTR may not serve as a suitable model in determining activity or interaction with MTRR or MMADHC.

Although the human MTR long helix α7 lies between the CblBD and AD and displays low conservation with MTR bacterial orthologues, its residue His920 (Figure 6.13) has been linked to a pathogenic variant of *CblG* defect (Section 1.3.3) (Leclerc, Campeau et al. 1996). A patient with the His920Asp missense mutation exhibited symptoms of homocystinuria, which improved with appropriate treatment, including OHCbl, betaine, folic acid, and methionine supplementation. However, this patient also carried a Ser450Asp

mutation in the FolBD, so it remains unclear whether His920 alone can be attributed to the pathogenicity (Kripps, Sremba et al. 2022). No residues at the interaction interface between MTRR and MTR have been associated with pathogenic or likely pathogenic *CblE* defect due to MTRR variants (Section 1.3.3).

A construct of $MTRR_{\Delta FMN}$ was generated for expression in *E. coli* for biochemical testing with MTR to verify whether these proteins interact, as predicted by AF2 Multimer. This construct would be used to assess if the interaction between MTR CblBD with MTRR CD and NADPHBD occurs. However, despite intense efforts, the $MTRR_{\Delta FMN}$ protein could not be purified due to its instability in solution. Even with the lack of biophysical data, the model generated suggests that the interaction between the CblBD of MTR and the CD and NADPHBD of MTRR can exist. This hypothesis is supported by biophysical data indicating that MTRR may stabilise MTR, as well as by the conservation analysis of both proteins, as well as the interactional analysis of the long helix of MTR and MTRR.

6.7.3 Structural Insights into the Ternary Complex – MTR-MTRR-MMADHC

Evidence for the existence of a possible supramolecular complex was investigated through pulldown assays and AF2 Multimer complex structure prediction, using MTR, MTRR, and MMADHC. MMACHC was not included in the pulldown or structure prediction experiments because, while previous studies suggest its involvement in the complex, Cbl transfer to MTR *in vitro* is facilitated by MMADHC. MMACHC is believed to be part of the complex via its interaction with MMADHC, not MTR (Froese, Kopec et al. 2015, Mascarenhas, Guha et al. 2023). Here, MTR was used as the central protein in the three-way complex, as its interactions with both MTRR and MMADHC are already established.

Modelling the two-way complexes MTR-MTRR and MTR-MMADHC suggested that MTR interacts with these proteins in the activation conformation, with the Cbl-binding pocket in CblBD serving as a crucial interface for this interaction, as it is the target for both MTRR and MMADHC (Figure 6.16). Although the two-way models indicated that MTRR and MMADHC interact with MTR in the same region (i.e. near the Cbl molecule), the MTR-MTRR-MMADHC model supported a previously unknown interface between MTR and MTRR. This interface occurs through the CD and NADPHBD of MTRR with the CblBD of MTR (Figure 6.12), allowing the formation of the three-way complex, with MMADHC occupying the Cbl-binding pocket, suggesting no other major interaction region between MMADHC and MTR.

The biological relevance of the putative supramolecular complex comprising MTR, MTRR, and MMADHC remains unclear. MTRR and MMADHC do not need to simultaneously perform their respective roles in Cbl delivery and reactivation, meaning both do not need to interact with MTR at the same time. The formation of this complex could be related to the biology of MTRR, discussed in Section 6.7.2. The elongated hinge and slower electron transfer rate of MTRR (Jarrett, Huang et al. 1998) suggest that initial binding of MTRR to MTR through the CD and NADPHBD facilitates more efficient reduction of Cbl when required. Additionally, evidence that MTRR stabilises *apo* MTR (Yamada, Gravel et al. 2006) suggests that MTR may be "primed" by MTRR before Cbl delivery by MMADHC, with MMADHC interacting with the MTR-MTRR complex to form a supramolecular complex upon Cbl delivery.

To date, the only biochemical evidence for the MTR-MTRR-MMADHC complex has been demonstrated *in vivo* through proximity ligation assays. A pulldown assay therefore is performed here using MTR, MTRR, and $MMADHC_{\Delta N}$ to explore the existence of this complex. MTR was used as bait due to the presence of a FLAG tag, which is absent in the

other two proteins. Although western blot analysis of the elution showed the presence of all three proteins (Figure 6.20), MMADHC and MTRR were also detected in the controls, indicating their attachment to the empty resin.

Although it may seem contradictory that both the FMNBD of MTRR and MMADHC interact with the same region of MTR yet can still form a three-way complex, this can be explained by novel interactions between the CblBD of MTR with the CD and NADPHBD of MTRR. At the moment when MMADHC is interacting with *apo* MTR, it is likely stabilised by MTRR only through its CD and NADPHBD, since the interaction of MMADHC with MTR indicates that MTR is in the *apo* state receiving Cbl (Figure 6.17B), and the electron from the FMNBD is not necessary. Upon the dissociation of MMADHC, MTR can remethylate the delivered cob[I]alamin and start the catalytic cycle. When Cbl is oxidised into cob[II]alamin, the FMNBD of MTRR can then interact with Cbl (Figure 6.8). In summary, the MTR–MTRR–MMADHC three-way complex likely represents the MTRR-stabilised *apo* MTR receiving Cbl from MMADHC, which does not need to occur simultaneously with the interaction of MTRR's FMNBD with MTR, thereby allowing both interactions to take place. After MMADHC is released from MTR and it needs to undergo the reactivation cycle, the FMNBD of MTRR can interact with the CblBD, which is no longer interacting with MMADHC.

The pulldown method used in Section 6.6 was insufficient to confirm the existence of the MTR-MTRR-MMADHC complex due to a failure in the negative control. To address this issue, alternative pulldown approaches could be employed, such as removing the tag from two proteins and using the His-tag on the third protein as bait on a nickel resin, repeating the process with each protein serving as bait in turn. Although these assays would provide additional evidence of the interactions, they would require large amounts of protein, which is impractical given the currently low expression levels of MTR and MTRR.

Another approach to investigate the dynamics of three-way complex formation could involve using a surface plasmon resonance (SPR) system. SPR would not only reduce the amount of protein needed but also provide insights into the association rates of MTRR and MMADHC with MTR and help to determine the transient nature of MMADHC's binding to MTR, as the stability of the MTR-MTRR complex was previously established (Section 6.3). This type of experiment has been used to show that the C-terminus of MMADHC is responsible for interacting with MMACHC (Deme, Miousse et al. 2012).

Further exploration of complex formation could involve the use of a construct of $MTRR_{\Delta FMN}$, which may offer insights into the predicted interaction between MTR and $MTRR_{\Delta FMN}$ (Figure 6.11). In addition to biophysical assays, further experimental structures could clarify the accuracy of the AF3 predicted models. The stable MTR-MTRR complex (Section 6.2), the MTR-$MMADHC_{\Delta N}$ complex (Section 6.4), and the combination of recombinant MTR, MTRR, and $MMADHC_{\Delta N}$ have been attempted for grid preparation and cryo-EM. Although datasets were collected (not shown), no 2D classes corresponding to the complexes were identified. The AF3 predicted models indicated flexibility in several regions (e.g., between the N- and C-terminal of MTR, the hinge of MTRR, and the FMNBD of MTRR in the three-way complex), which may have contributed to the low quality of the datasets. To address this, using a truncated $MTRR_{\Delta FMN}$ could reduce flexibility, and testing different grid types could improve the quality of the cryo-EM dataset.

7. Conclusions, Perspective and Future Work

7.1 Key findings in this work

Methionine synthase (MTR) is the end-user of cobalamin (Cbl) within the cytosolic branch of the intracellular Cbl metabolism. By using Cbl as a co-factor, it remethylates homocysteine (Hcy) to methionine (Met) by using the methyl group from methyltetrahydrofolate (MeTHF) (Banerjee and Matthews 1990). Met is an essential amino acid, crucial for protein synthesis and the production of S-adenosylmethionine (SAM), a universal methyl donor involved in various methylation reactions, which are essential for cellular function and gene expression. Therefore, MTR acts as a link between the folate and Met/SAM cycles, playing an important role in many methylation processes and amino acid metabolism (Froese, Fowler et al. 2019). The deficiency in MTR leads to elevated levels of Hcy and decreased levels of Met, which can cause encephalopathy, homocysteinemia, seizures, megaloblastic anaemia, and other symptoms.

Despite being important in the human metabolism, the multi-domain and flexible nature of MTR are obstacles in obtaining structural data. For instance, there is no structural data on the Cap or CblBD of the human MTR, or the C-terminal half of the protein as it is seen for the bacterial orthologue. As a result, much of the current structural and biochemical understanding of this protein has been derived from data of its bacterial orthologue (Koutmos, Datta et al. 2009). Although the bacterial and human MTR share similarities, it is important to study the human protein, once the findings for the bacterial MTR might not be totally translated to the human.

This thesis outlines my efforts and accomplishment to obtain recombinant full-length human MTR suitable for structural and interactional studies. Chapter 3 describe the expression, purification, and biochemical characterisation of recombinant MTR. The

findings confirm that although a significant portion of the purified protein forms aggregates, part of it remains homogeneous, monomeric, intact, and in its apoenzyme form. The *apo* form of the recombinant MTR broads the possibilities on ligands combinations to be used on the protein for structural and biophysical assays. Furthermore, the development of a miniaturised activity assay, based upon previous method (Mascarenhas, Gouda et al. 2022) that was low-throughput, demonstrate that the protein is performing its catalytic cycle. Negative staining electron microscopy (EM) provided visual confirmation that the sample would be suitable for further structural studies using cryo-EM.

Chapter 4 details the preparation of cryo-EM grids, data collection, and processing. The *apo* state of MTR facilitates data collection with various ligand combinations, allowing investigation of ligand-induced conformational changes. The grids preparation and data collection were performed at the York Structural Biology Laboratory, at the University at York. The need to travel to York to prepare the grids and access the Glacios microscope means that grid optimisation has become a crucial and time-consuming bottleneck in the acquisition of high-quality cryo-EM datasets. Although the distance and limited availability of the microscope present challenges, the importance of optimising the sample and grid type cannot be overstated, as it directly impacts the quality of the data. Through diligent planning and effort, I was able to overcome these difficulties and managed to find optimal ligand combinations and grid types that led to high-resolution density maps. Throughout the analysis of the datasets, it became clear that no particles containing all five domains of MTR were observed. The two halves of MTR are seen independently due to the flexibility of the linker connecting them (Linker 1) and the fact that no domains of the C-half were interacting with the N-half. The absence of particles corresponding to full-length MTR was not surprising, as it is known to be a dynamic protein, and the most common strategy to study its structure is to express its domains or regions separately.

Chapter 5 presents the analysis and interpretation of the cryo-EM data, showing that human *apo* MTR adopts the activation conformation, consistent with previous observations for *T. thermophilus* MTR (Mendoza, Purchal et al. 2023). According to my dataset on MTR+MetCbl, upon MetCbl binding, MTR undergoes a conformational shift to the resting conformation, which is the conformation it acquires prior to undergoing the catalytic cycle. Analysis of the high-resolution density map of the N-terminal half bound to MeTHF reveals a configuration of the HcyBD and FolBD as a rigid body, nearly identical to that previously crystallised in the group for human MTR (PDB: 4CCZ). Additionally, three structures of the C-terminal half in the activation conformation are determined at ~3.0 Å resolution—one without any ligand, and two with Cbl bound either in the His-on or His-off state.

Further analysis of these structures of the C-terminal half of human MTR highlights the role of His785 from CblBD in coordinating the lower axial position of Cbl, similar to what was previously observed in bacterial MTR. It is also found that Tyr1177 from the AD, which undergoes conformational changes in bacterial MTR, assumes the same position in the human protein regardless of the presence of Cbl or the His-on/off state. In bacterial MTR, the conserved tyrosine undergoes a conformational change during the reactivation of Cbl and is placed closer to the Cbl ligand in the His-off state than in the His-on. These findings suggest that while some observations from bacterial MTR, such as the function of His785, can be applicable to the human protein, not all residues behave similarly, such as Tyr1177, indicating potential differences in the reactivation of Cbl between bacterial and human MTR (see next section 7.2).

Chapter 6 investigates the interaction of MTR with MTRR and MMADHC, using both biophysical (pulldown and spectroscopic assays) and computational (AlphaFold Multimer for model generation and conservation analysis) methods. The pulldown approach demonstrates that MTR and MTRR form a stable complex, confirmed through Blue Native-

PAGE and pulldown assays. The interaction between MTR and MMADHC is confirmed spectroscopically via monitoring Cbl transfer from MMADHC to MTR, although the formation of a stable complex could not be conclusively observed. Additionally, pulldown experiments to assess the formation of a stable complex between the recombinant proteins do not yield conclusive results to date. AF Multimer modelling reveals a novel interaction between MTR and MTRR, involving domains from both proteins (i.e. CblBD from MTR and CD and NADPHBD from MTRR) that are not known to interact with each other before this study. Altogether these results suggest the potential formation of a supramolecular complex involving MTR, MTRR, and MMADHC, which remains to be verified in future through biophysical assays using recombinant proteins (section 7.3 and 7.4).

7.2 Relevance of the current data to the field of the biology of MTR

7.2.1 Insights into the reactivation cycle of human MTR

The reactivation of MTR occurs when Cbl is oxidised into cob[II]alamin, and in humans, the protein MTRR is responsible for transferring an electron to Cbl to allow its remethylation. Although the bacterial MTR also undergoes inactivation for the same reason, the reactivation is performed with the aid of the protein flavodoxin. Based on the C-terminal half structures determined for MTR in Chapter 5, a revised paradigm for the remethylation of cobalamin (Cbl) during the reactivation cycle of human MTR is proposed. The unchanged position of Tyr1177, and the movements observed in the Cbl ligand between the His-on and His-off states, provide key evidence that, in human MTR, the transition from the His-on to His-off state induces a shift in the coordination state of Cbl from 5-coordinate to 4-coordinate. This stabilises the Cbl molecule, preparing it for reductive methylation upon binding with MTRR and SAM. These findings indicate that the primary difference between human and bacterial MTR is that, in bacteria, the movement of the conserved tyrosine

residue allows the reductive methylation of Cbl. In contrast, in the human enzyme, the transition from the His-on to the His-off state is sufficient to enable Cbl to be reductively methylated.

Although the overall mechanism of reactivation appears similar to that observed in bacterial MTR, these structural differences, which were previously unrecognised due to the absence of C-terminal half structures of human MTR, challenge the broad assumption that the biology of bacterial MTR can be directly translated to the human protein. This underscores the need for further studies on human MTR to fully understand its structure and function. These differences in the reactivation mechanism of human MTR might be related to evolutionary adaptations to the different physiological environments in humans compared to bacteria. It is well known that the reactivation of MTR occurs through its interaction with different proteins in humans and bacteria, which may indicate enhanced efficiency or regulation of the reactivation process. This highlights the exceptional importance of studying the features of human MTR independently of its bacterial orthologue.

7.2.2 Mapping protein-protein interfaces of MTR

The models predicted using AF Multimer not only confirm the previously observed interactions of MTR with MTRR, and of MTR with MMADHC (Yamada, Gravel et al. 2006, Mascarenhas, Guha et al. 2023), it also unveils a previously unrecognised interaction, involving the CblBD of MTR with the CD and NADPHBD of MTRR. This interaction suggests that the connection between these two proteins extends beyond the reactivation of MTR, aligning with biochemical data that indicates MTRR may enhance MTR activity and stabilise its apo form (Zhang, Liu et al. 2020). Furthermore, this interaction could facilitate the simultaneous binding of MTRR and MMADHC to MTR, as both proteins have been shown to interact with the same region of MTR (the Cbl-binding pocket). The concurrent binding of MMADHC and MTRR to MTR may not seem obvious because they are required

at different times—Cbl loading into the *apo* protein and reactivation of Cbl, respectively—but it could represent a means by which the *apo* MTR is stabilised by MTRR prior to receiving Cbl from MMADHC. The simultaneous interaction of MTRR and MMADHC with MTR also supports the hypothesis of a putative supramolecular complex involving these three proteins along with MMACHC (Bassila, Ghemrawi et al. 2017). Gaining further insights into the structural basis of these interactions could advance our understanding of how they occur and how genetic mutations might impact these interactions in disorders of intracellular cobalamin metabolism.

7.3 Limitations of the current study

Although this thesis has provided valuable insights into the structure and biochemistry of MTR and its complex formation with MTRR and MMADHC, certain limitations need to be acknowledged.

While in Chapter 3, the expression and purification protocols described for the human full-length MTR yielded a high-quality protein that exhibits enzymatic activity and is suitable for cryo-EM, the 96-well plate activity assay developed has a sensitivity limitation. Although the assay is sufficient to determine whether the protein is active or not, it is not sensitive for kinetic studies, which can indicate the K_m and V_{max} of MTR under various conditions for example. The 96-well assay is not sensitive enough because of the low amount of MeTHF used results in a low number of THF molecules (which are used in the readout of the assay), thereby it impacts in the sensitivity of the assay.

Although the cuvette method, which requires a large volume of reagents and protein (1 mL final volume), can be employed for kinetic studies—as demonstrated for *T. thermophilus*, where SAM, MeTHF, and Hcy were tested in a dose–response manner (Mendoza, Purchal et al. 2023) to obtain similar data for human MTR, a highly sensitive

miniaturised assay is needed due to limited protein availability. One possibility is to use an enzyme-coupled assay, where an enzyme generates products that can be detected via fluorescence, offering high sensitivity. An example is the optimisation of an activity assay in which the enzyme methionine gamma-lyase is used to convert the methionine produced by MTR into methanethiol, α-ketobutyrate, and ammonia (Sato and Nozaki 2009). A fluorogenic probe for free thiols (methanethiol) can then be used to quantify the amount of methionine produced in the reaction (Liu, Huo et al. 2016). Once optimised, this assay can be used to test different substrate concentrations in MTR activity assays for kinetic characterisation. Furthermore, it can be used to study protein mutants, investigate the activity of MTR complexed with MTRR, or test different types of Cbl.

During the processing of the cryo-EM data, MTR was observed as two halves due to its conformation and flexibility. As mentioned in Chapter 5, MTR has two flexible linkers in particular that allow it to undergo conformational changes during its catalytic and reactivation cycles. Although Linker 1 is responsible for MTR not being seen as a single particle, in the reactivation conformation Linker 2 does not interfere, since the two domains it connects (CblBD and AD) are interacting with each other. This implies that the interaction of certain domains can overcome the problem of the linker flexibility and might allow the visualisation of the full-length or near full-length MTR. For instance, during the catalytic cycle, the CblBD interacts with HcyBD and FolBD in the N-half, and if either of these two interactions can be captured, it might be possible to see most of the protein because it will reduce the flexibility between the halves of the protein. Although this would allow the visualisation of novel conformations, the AD would no longer be associated with the CblBD and would likely not be seen in those particles.

Because of its flexibility, the visualisation of full-length MTR is challenging and could be achieved using, for example, cross-linking strategies or constructs with shorter

linkers, but this would impact the native biology of the protein. Although achieving the visualisation of all five domains of human MTR in a single structure might sound groundbreaking, it should be considered that this could involve changing the endogenous flexibility and dynamics of MTR, which is one of the most interesting and least understood features of the protein.

Using the sample of apo MTR for cryo-EM was a determining factor in obtaining high-resolution maps that allowed the observation of details in MTR structure, such as the positions of residues His785 and Tyr1177. The optimisation of the grids' preparation described in Chapter 4 was sufficient to generate valuable data for MTR using two ligand combinations (MTR+MeTHF and MTR+MeTHF+SAM+OHCbl), but the grids of the samples containing apo MTR and MTR+MetCbl did not result in any high-resolution maps due to the quality of the dataset. Although the datasets from the latter samples did not result in any high-resolution density maps, the 2D classes generated were sufficient to infer the conformation of MTR in both cases.

Further rounds of optimisation, such as using different buffers, adding detergents, or using different grid types, may help to improve these datasets sufficiently to generate a density map. Besides that, the use of nanobodies could further stabilise specific flexible regions of proteins and allow the acquisition of an improved dataset (Wu and Rapoport 2021). Furthermore, these optimisations can also assist in the data acquisition of other samples that did not even generate 2D classes, such as the complexes MTR–MTRR and MTR–MMADHC.

Finally, while the interaction studies in Chapter 6 show the interaction of MTR with MTRR and MMADHC, and the AF2 Multimer models offer new insights into the interaction of these three proteins and the formation of the MTR–MTRR–MMADHC complex, more

biochemical data is still needed to validate these models. In particular, using a construct of $MTRR_{\Delta FMN}$ would confirm the novel interaction found between the CblBD of MTR and the CD and NADPHBD of MTRR.

The biophysical study of interactions between MTR, MTRR, and MMADHC could be performed using Surface Plasmon Resonance (SPR). This technique allows for the real-time analysis of molecular interactions, taking advantage of the FLAG tag present exclusively in the recombinant MTR used throughout this work (Takeda, Goshima et al. 2010). This highly sensitive method can be applied to investigate the association and dissociation rates of both MTRR and MMADHC with MTR, using low amounts of protein since protein availability is an issue with human MTR. Because it can detect even weak interactions, valuable information can be obtained, such as differences in the interaction between MTR and *apo* or Cbl-loaded MMADHC, or insights into the formation and dynamics of higher-order complexes involving MTR, MTRR, and MMADHC.

7.4 Laying the groundwork for a future direction – CRISPR-mediated endogenous tagging

The experiments I conducted in this thesis have focused primarily on the use of recombinant proteins. Although recombinant expression of proteins can be a good strategy for acquiring structural and biophysical data and for studying protein–protein interactions, endogenous proteins may offer more relevant insights. For example, obtaining high-quality cryo-EM data for dynamic complexes can be challenging when using recombinant proteins, and employing the endogenous complex can be an effective alternative (Yan, Hang et al. 2015). Furthermore, the endogenous approach can reveal novel participants of the MTR interactome.

To further comprehend the interactions between MTR and other proteins, especially those involved in intracellular Cbl metabolism (such as MTRR, MMADHC, and MMACHC), an endogenous approach may provide new insights into these interactions. The rationale for this strategy involves using CRISPR to knock in an affinity tag (e.g., FLAG) into the MTR gene in HEK293 cells to "fish out" interacting proteins (Auer, Duroure et al. 2014).

During the initial development of my DPhil project, aiming to scope out possible directions for generating thesis data, I have attempted in creating a monoclonal cell line in which the *MTR* gene is edited so that the protein is translated with an affinity tag. This avenue was put aside, to make way for recombinant protein work and cryo-EM studies that progressed well at the time. Nevertheless, herein I provide a brief summary of what has been achieved, to illustrate that this is a feasible approach to gather more data and reveal more information on the human MTR. This research direction is now being continued by another PhD student in the group.

The knock-in approach I used focused on the homology-directed repair (HDR) method, which differs from non-homologous end joining (NHEJ) because the donor DNA to be inserted into the genome has long homologous arms that increase the probability of successful DNA knock-in (He, Tan et al. 2016). The overall flow for monoclonal cell line generation starts with the design of sgRNAs targeting specific regions of the *MTR* gene. Targeting regions are more limited in CRISPR knock-in because they must be located in an exon (since it needs to be expressed) and, in this case, in regions of the protein that will not interfere with its activity or interactions with other proteins.

The sequence of the *MTR* gene was used to design sgRNAs targeting key regions of the protein (N-terminal, C-terminal, and the linker between the protein halves) using the

sgRNA design tool on Benchling (https://benchling.com/). Different regions of the protein were selected because sgRNAs not always are capable of cleaving the genome in some regions, so having more options of regions increase the chance of finding a promising sgRNA. The 20-nucleotide dsDNA sequences were cloned into a vector (eSpCas9(1.1) Dual sgRNA, Addgene #86613) for Cas9 expression (Agudelo, Duringer et al. 2017).

As described in Section 2.8.1, the sgRNAs (1-14) were cloned into the vector, and cloning was confirmed by PCR amplification of the sgRNA sequences in the plasmids (Figure 7.1). sgRNA 11, which was correctly cloned into the plasmid (Figure 7.1A), was tested in a surveyor assay to check if it could cleave the MTR gene (Figure 7.1B). Briefly, the surveyor assay consists of transfecting HEK293 cells with the plasmid expressing sgRNA 11, extracting the genomic DNA, amplifying the MTR gene, and using T7 endonuclease to digest the cleaved DNA (which indicates that the sgRNA worked). When the sample is loaded on an agarose gel alongside a control sample, it is possible to visualise that the DNA band is less intense and that lower bands are present, indicating that the DNA was correctly cleaved.

The sgRNA 11, targeting the region of residue Gln6 of MTR, successfully cleaved the MTR gene (Figure 7.1B) and is a good candidate for generating a cell line expressing endogenous N-terminal FLAG-tagged MTR. To achieve this, future next steps would involve designing a donor DNA complementary to the region targeted by sgRNA 11 and performing the transfection (Figure 7.1C). The donor template, containing the sequence of the FLAG peptide and homologous arms for the cleavage region of the sgRNA 11 will correctly insert the tag in the N-terminus of MTR. The dual sgRNA plasmid allows the expression of simultaneously the sgRNA targeting MTR and the gene ATP1A1, and the addition of a donor template for each cleaved region allows the insertion of the tag and generation of ouabain-resistant cells, respectively. After cell selection using the resistance to

ouabain, generation of monoclonal cell lines by single cell dilution, and confirmation that the tag has been correctly inserted by western blotting and DNA sequencing, the cell line can be used to study the interacting proteins of MTR, using the FLAG tag for purification.

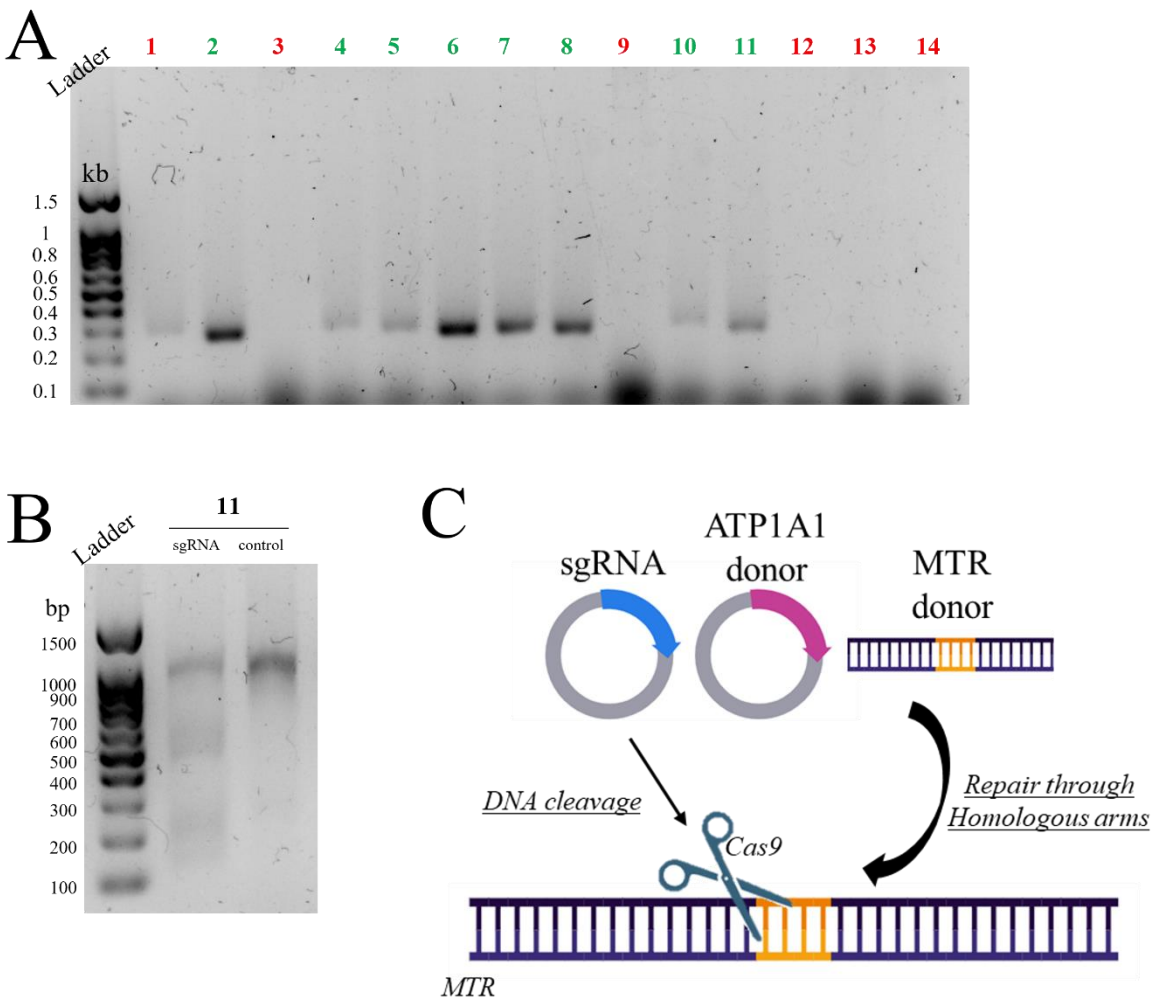

Figure 7.1: Cloning and Surveyor Assay of sgRNAs Targeting the MTR Gene. A. PCR amplification of the cloned sgRNAs (1–14), highlighting those successfully cloned into the vector (green) and those that were not cloned (red). **B.** Surveyor assay using the vector expressing sgRNA 11, showing a less intense band in the treated sample compared to the control. **C.** Diagram displaying the three-plasmid transfection, in which the sgRNA plasmid expresses the Cas9 with either sgRNA targeting the *MTR* or *ATP1A1* gene. The second plasmid contains the repair template for the *ATP1A1* gene and confers resistance to ouabain. The MTR donor, transfected as a double strand DNA, inserts the sequence of FLAG peptide into the N-terminus of MTR.

The complexes obtained from such endogenous isolation and purification of MTR has huge potential for subsequent investigations. Firstly they can be further analysed using

mass spectrometry (MS) to identify the interacting partners (Gingras, Gstaiger et al. 2007). Secondly, depending on the purity of the sample, the endogenous complexes can also be used for cryo-EM data collection. Hopefully at least the complexes MTR-MTRR and MTR-MTRR-MMACHC-MMADHC will be identified and can be tested on grids. Furthermore, the MS data can be used to reconstruct the complexes recombinantly for further biochemical and structural analysis.

Although there are considerable bottlenecks in the CRISPR knock-in technique, such as obtaining an sgRNA that efficiently cleaves the DNA and the selection of correctly edited monoclonal cell lines, the technique is promising and has been performed previously in the group (Kelly, Tranter et al. 2022). Once completed, the edited cell line can provide novel structural, interactional, and biophysical insights into MTR by utilising the endogenous protein.

The elucidation of protein structures is a crucial step in advancing drug development targeting these proteins. Indeed, certain drugs for HIV treatment were developed using the crystal structure of a protease (Lapatto, Blundell et al. 1989, Miller, Schneider et al. 1989). The description of the human MTR structure presented in this work paves the way for further opportunities in drug discovery.

Chapter 6, which provides insights into the interactions of MTR with MTRR and MMADHC, highlights the potential for utilising pharmacological chaperones. These chaperones could prove beneficial in restoring the interactions of MTR mutants with other proteins, thereby influencing intracellular Cbl metabolism. Certain pharmacological chaperones have been demonstrated to rescue the activity of misfolded proteins (Hawtin 2006), a finding that could be relevant to specific MTR mutations affecting protein-protein interactions. The detailed description of the molecular interface between MTR and its

interacting partners may also aid in the computational design of such pharmacological chaperones (Grasso, Galderisi et al. 2023). Furthermore, the high-resolution structure of full-length MTR, as described in Chapter 5, could be utilised in the design of small molecules aimed at restoring the activity of mutant proteins that can result in metabolic diseases (Qiao, Molina et al. 2006).

In summary, this study of the fundamental structural biology of MTR represents an initial step towards a more applied approach, focusing on the discovery of ligands to address metabolic diseases caused by alterations in the native structure of the protein.

Since the first structure of MTR was published 30 years ago (Drennan, Huang et al. 1994), this is the first time that human MTR has been observed in its full length, providing a high-resolution structure of the C-terminal half domains (Cap, CblBD, and AD), thereby revealing novel insights into the reactivation cycle of human MTR. Until now, the *E. coli* MTR has been used as a model to infer the mechanism of reactivation in human MTR (Koutmos, Datta et al. 2009). Although bacterial and human MTR are structurally similar, differences such as the partner protein involved (MTRR in humans, flavodoxin in bacteria) and the mechanism of cobalamin loading (via MMADHC in humans, still uncertain in bacteria) must be considered when translating findings from one protein to the other. More than 80 litres of Sf9 cells, nearly 50 grids for cryo-EM data collection, and almost 100,000 micrographs analysed in total were necessary to address a multi-decade knowledge gap in the structural biology of human MTR.

8. References

Abramson, J., J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard and J. Bambrick (2024). "Accurate structure prediction of biomolecular interactions with AlphaFold 3." Nature: 1-3.

Adams, J., S. K. Ross, L. Mervyn, K. Boddy and P. King (1971). "Absorption of cyanocobalamin, coenzyme B12, methylcobalamin, and hydroxocobalamin at different dose levels." Scandinavian journal of gastroenterology **6**(3): 249-252.

Agudelo, D., A. Duringer, L. Bozoyan, C. C. Huard, S. Carter, J. Loehr, D. Synodinou, M. Drouin, J. Salsman and G. Dellaire (2017). "Marker-free coselection for CRISPR-driven genome editing in human cells." Nature methods **14**(6): 615-620.

Akdel, M., D. E. Pires, E. P. Pardo, J. Jänes, A. O. Zalevsky, B. Mészáros, P. Bryant, L. L. Good, R. A. Laskowski and G. Pozzati (2022). "A structural biology community assessment of AlphaFold2 applications." Nature Structural & Molecular Biology **29**(11): 1056-1067.

Alfonso-Prieto, M., X. Biarnés, M. Kumar, C. Rovira and P. M. Kozlowski (2010). "Reductive cleavage mechanism of Co− C bond in cobalamin-dependent methionine synthase." The Journal of Physical Chemistry B **114**(40): 12965-12971.

Alonso, H., P. L. Cummins and J. E. Gready (2009). "Methyltetrahydrofolate: corrinoid/iron–sulfur Protein Methyltransferase (MeTr): Protonation State of the Ligand and Active-Site Residues." The Journal of Physical Chemistry B **113**(44): 14787-14796.

Ashkenazy, H., S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko and N. Ben-Tal (2016). "ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules." Nucleic acids research **44**(W1): W344-W350.

Auer, T. O., K. Duroure, A. De Cian, J.-P. Concordet and F. Del Bene (2014). "Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair." Genome research **24**(1): 142-153.

Augustijn, D., S. Mahapatra, W. Streicher, H. Svilenov, A. Kulakova, C. Pohl and Å. Rinnan (2019). "Novel non-linear curve fitting to resolve protein unfolding transitions in intrinsic fluorescence differential scanning fluorimetry." European Journal of Pharmaceutics and Biopharmaceutics **142**: 506-517.

Baek, M., F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch and R. D. Schaeffer (2021). "Accurate prediction of protein structures and interactions using a three-track neural network." Science **373**(6557): 871-876.

Bai, X.-C., G. McMullan and S. H. Scheres (2015). "How cryo-EM is revolutionizing structural biology." Trends in biochemical sciences **40**(1): 49-57.

Bandarian, V., M. L. Ludwig and R. G. Matthews (2003). "Factors modulating conformational equilibria in large modular proteins: a case study with cobalamin-dependent methionine synthase." Proceedings of the National Academy of Sciences **100**(14): 8156-8163.

Bandarian, V., K. A. Pattridge, B. W. Lennon, D. P. Huddler, R. G. Matthews and M. L. Ludwig (2002). "Domain alternation switches B12-dependent methionine synthase to the activation conformation." Nature structural biology **9**(1): 53-56.

Banerjee, R. (1999). Chemistry and Biochemistry of B12, John Wiley & Sons.

Banerjee, R. V., V. Frasca, D. P. Ballou and R. G. Matthews (1990). "Participation of cob (I) alamin in the reaction catalyzed by methionine synthase from Escherichia coli: a steady-state and rapid reaction kinetic analysis." Biochemistry **29**(50): 11101-11109.

Banerjee, R. V. and R. G. Matthews (1990). "Cobalamin-dependent methionine synthase." The FASEB journal **4**(5): 1450-1459.

Bassila, C., R. Ghemrawi, J. Flayac, D. S. Froese, M. R. Baumgartner, J.-L. Guéant and D. Coelho (2017). "Methionine synthase and methionine synthase reductase interact with MMACHC and

with MMADHC." Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease **1863**(1): 103-112.
Beedholm-Ebsen, R., K. van de Wetering, T. Hardlei, E. Nexø, P. Borst and S. K. Moestrup (2010). "Identification of multidrug resistance protein 1 (MRP1/ABCC1) as a molecular gate for cellular export of cobalamin." Blood, The Journal of the American Society of Hematology **115**(8): 1632-1639.
Bendory, T., A. Bartesaghi and A. Singer (2020). "Single-particle cryo-electron microscopy: Mathematical theory, computational challenges, and opportunities." IEEE signal processing magazine **37**(2): 58-76.
Bepler, T., A. Morin, M. Rapp, J. Brasch, L. Shapiro, A. J. Noble and B. Berger (2019). "Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs." Nature methods **16**(11): 1153-1160.
Boyes, B. E., D. G. Walker and P. L. McGeer (1988). "Separation of large DNA restriction fragments on a size-exclusion column by a nonideal mechanism." Analytical biochemistry **170**(1): 127-134.
Broderick, J. B., B. R. Duffus, K. S. Duschene and E. M. Shepard (2014). "Radical S-adenosylmethionine enzymes." Chemical reviews **114**(8): 4229-4317.
Bryant, P., G. Pozzati and A. Elofsson (2022). Improved prediction of protein-protein interactions using AlphaFold2. Nat. Commun. 13, 1265.
Butola, L. K., P. K. Kute, A. Anjankar, A. Dhok, N. Gusain and A. Vagga (2020). "Vitamin B12-do you know everything." Journal of Evolution of Medical and Dental Sciences **9**(42): 3139-3147.
Chen, Z., S. Chakraborty and R. Banerjee (1995). "Demonstration that mammalian methionine synthases are predominantly cobalamin-loaded." Journal of Biological Chemistry **270**(33): 19246-19249.
Cheng, J., G. Novati, J. Pan, C. Bycroft, A. Žemgulytė, T. Applebaum, A. Pritzel, L. H. Wong, M. Zielinski and T. Sargeant (2023). "Accurate proteome-wide missense variant effect prediction with AlphaMissense." Science **381**(6664): eadg7492.
Cheng, Y. (2018). "Single-particle cryo-EM—How did it get here and where will it go." Science **361**(6405): 876-880.
Claasen, M., Z. Kofinova, M. Contino and W. B. Struwe (2024). "Analysis of Protein Complex Formation at Micromolar Concentrations by Coupling Microfluidics with Mass Photometry." Journal of Visualized Experiments: Jove(203).
Coelho, D., J. C. Kim, I. R. Miousse, S. Fung, M. du Moulin, I. Buers, T. Suormala, P. Burda, M. Frapolli and M. Stucki (2012). "Mutations in ABCD4 cause a new inborn error of vitamin B12 metabolism." Nature genetics **44**(10): 1152-1155.
Coelho, D., T. Suormala, M. Stucki, J. P. Lerner-Ellis, D. S. Rosenblatt, R. F. Newbold, M. R. Baumgartner and B. Fowler (2008). "Gene identification for the cblD defect of vitamin B12 metabolism." New England Journal of Medicine **358**(14): 1454-1464.
Craig, W. J. and A. R. Mangels (2009). "Position of the American Dietetic Association: vegetarian diets." Journal of the American dietetic association **109**(7): 1266.
Croll, T. I. (2018). "ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps." Acta Crystallographica Section D: Structural Biology **74**(6): 519-530.
Datta, S., M. Koutmos, K. A. Pattridge, M. L. Ludwig and R. G. Matthews (2008). "A disulfide-stabilized conformer of methionine synthase reveals an unexpected role for the histidine ligand of the cobalamin cofactor." Proceedings of the National Academy of Sciences **105**(11): 4115-4120.
Deme, J. C., M. A. Hancock, X. Xia, C. A. Shintre, M. Plesa, J. C. Kim, E. P. Carpenter, D. S. Rosenblatt and J. W. Coulton (2014). "Purification and interaction analyses of two human lysosomal vitamin B12 transporters: LMBD1 and ABCD4." Molecular membrane biology **31**(7-8): 250-261.

Deme, J. C., I. R. Miousse, M. Plesa, J. C. Kim, M. A. Hancock, W. Mah, D. S. Rosenblatt and J. W. Coulton (2012). "Structural features of recombinant MMADHC isoforms and their interactions with MMACHC, proteins of mammalian vitamin B12 metabolism." Molecular genetics and metabolism **107**(3): 352-362.
den Boer, M. A., S.-H. Lai, X. Xue, M. D. van Kampen, B. Bleijlevens and A. J. Heck (2021). "Comparative analysis of antibodies and heavily glycosylated macromolecular immune complexes by size-exclusion chromatography multi-angle light scattering, native charge detection mass spectrometry, and mass photometry." Analytical Chemistry **94**(2): 892-900.
Dixon, M. M., S. Huang, R. G. Matthews and M. Ludwig (1996). "The structure of the C-terminal domain of methionine synthase: presenting S-adenosylmethionine for reductive methylation of B12." Structure **4**(11): 1263-1275.
Doerr, A. (2016). "Single-particle cryo-electron microscopy: a brief overview of how to solve a macromolecular structure using single-particle cryo-electron microscopy (cryo-EM)." Nature methods **13**(1): 23-24.
Douzi, B. (2017). "Protein–protein interactions: surface plasmon resonance." Bacterial Protein Secretion Systems: Methods and Protocols: 257-275.
Drennan, C. L., S. Huang, J. T. Drummond, R. G. Matthews and M. L. Ludwig (1994). "How a protein binds B12: a 3.0 Å X-ray structure of B12-binding domains of methionine synthase." Science **266**(5191): 1669-1674.
Drummond, J. T., S. Huang, R. M. Blumenthal and R. G. Matthews (1993). "Assignment of enzymic function to specific protein regions of cobalamin-dependent methionine synthase from Escherichia coli." Biochemistry **32**(36): 9290-9295.
Drummond, J. T., J. Jarrett, J. C. Gonzalez, S. Huang and R. G. Matthews (1995). "Characterization of nonradioactive assays for cobalamin-dependent and cobalamin-independent methionine synthase enzymes." Analytical biochemistry **228**(2): 323-329.
Duong-Ly, K. C. and S. B. Gabelli (2014). Gel filtration chromatography (size exclusion chromatography) of proteins. Methods in enzymology, Elsevier. **541:** 105-114.
Evans, J. C., D. P. Huddler, M. T. Hilgers, G. Romanchuk, R. G. Matthews and M. L. Ludwig (2004). "Structures of the N-terminal modules imply large domain motions during catalysis by methionine synthase." Proceedings of the National Academy of Sciences **101**(11): 3729-3736.
Evans, J. C., D. P. Huddler, J. Jiracek, C. Castro, N. S. Millian, T. A. Garrow and M. L. Ludwig (2002). "Betaine-homocysteine methyltransferase: zinc in a distorted barrel." Structure **10**(9): 1159-1171.
Evans, R., M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell and J. Yim (2021). "Protein complex prediction with AlphaFold-Multimer." biorxiv: 2021.2010. 2004.463034.
Fabrini, R., A. De Luca, L. Stella, G. Mei, B. Orioni, S. Ciccone, G. Federici, M. Lo Bello and G. Ricci (2009). "Monomer− dimer equilibrium in glutathione transferases: A critical re-examination." Biochemistry **48**(43): 10473-10482.
Faruqi, A. and R. Henderson (2007). "Electronic detectors for electron microscopy." Current opinion in structural biology **17**(5): 549-555.
Feathers, J. R., K. A. Spoth and J. C. Fromme (2021). "Experimental evaluation of super-resolution imaging and magnification choice in single-particle cryo-EM." Journal of structural biology: X **5**: 100047.
Fedosov, S. N., N. U. Fedosova, B. Kräutler, E. Nexø and T. E. Petersen (2007). "Mechanisms of discrimination between cobalamins and their natural analogues during their binding to the specific B12-transporting proteins." Biochemistry **46**(21): 6446-6458.
Fettelschoss, V., P. Burda, C. Sagné, D. Coelho, C. De Laet, S. Lutz, T. Suormala, B. Fowler, N. Pietrancosta and B. Gasnier (2017). "Clinical or ATPase domain mutations in ABCD4 disrupt the interaction between the vitamin B12-trafficking proteins ABCD4 and LMBD1." Journal of Biological Chemistry **292**(28): 11980-11991.

Fick, R. J., M. C. Clay, L. Vander Lee, S. Scheiner, H. Al-Hashimi and R. C. Trievel (2018). "Water-mediated carbon–oxygen hydrogen bonding facilitates S-adenosylmethionine recognition in the reactivation domain of cobalamin-dependent methionine synthase." Biochemistry **57**(26): 3733-3740.
Fleischhacker, A. S. and R. G. Matthews (2007). "Ligand trans influence governs conformation in cobalamin-dependent methionine synthase." Biochemistry **46**(43): 12382-12392.
Flower, T. G. and J. H. Hurley (2021). "Crystallographic molecular replacement using an in silico-generated search model of SARS-CoV-2 ORF8." Protein Science **30**(4): 728-734.
Fofou-Caillierez, M. a. B., N. T. Mrabet, C. Chéry, N. Dreumont, J. Flayac, M. Pupavac, J. Paoli, J.-M. Alberto, D. Coelho and J.-M. Camadro (2013). "Interaction between methionine synthase isoforms and MMACHC: characterization in cblG-variant, cblG and cblC inherited causes of megaloblastic anaemia." Human molecular genetics **22**(22): 4591-4601.
Forny, P., X. Bonilla, D. Lamparter, W. Shao, T. Plessl, C. Frei, A. Bingisser, S. Goetze, A. van Drogen and K. Harshman (2023). "Integrated multi-omics reveals anaplerotic rewiring in methylmalonyl-CoA mutase deficiency." Nature Metabolism **5**(1): 80-95.
Frasca, V., R. V. Banerjee, W. R. Dunham, R. H. Sands and R. G. Matthews (1988). "Cobalamin-dependent methionine synthase from Escherichia coli B: electron paramagnetic resonance spectra of the inactive form and the active methylated form of the enzyme." Biochemistry **27**(22): 8458-8465.
Frazer, J., P. Notin, M. Dias, A. Gomez, J. K. Min, K. Brock, Y. Gal and D. S. Marks (2021). "Disease variant prediction with deep generative models of evolutionary data." Nature **599**(7883): 91-95.
Froese, D. S., B. Fowler and M. R. Baumgartner (2019). "Vitamin B12, folate, and the methionine remethylation cycle—biochemistry, pathways, and regulation." Journal of inherited metabolic disease **42**(4): 673-685.
Froese, D. S. and R. A. Gravel (2010). "Genetic disorders of vitamin B12 metabolism: eight complementation groups–eight genes." Expert reviews in molecular medicine **12**: e37.
Froese, D. S., J. Kopec, F. Fitzpatrick, M. Schuller, T. J. McCorvie, R. Chalk, T. Plessl, V. Fettelschoss, B. Fowler and M. R. Baumgartner (2015). "Structural insights into the MMACHC-MMADHC protein complex involved in vitamin B12 trafficking." Journal of Biological Chemistry **290**(49): 29167-29177.
Fujii, K., J. Galivan and F. Huennekens (1977). "Activation of methionine synthase: further characterization of the flavoprotein system." Archives of biochemistry and biophysics **178**(2): 662-670.
Garcin, E. D., C. M. Bruns, S. J. Lloyd, D. J. Hosfield, M. Tiso, R. Gachhui, D. J. Stuehr, J. A. Tainer and E. D. Getzoff (2004). "Structural Basis for Isozyme-specific Regulation of Electron Transfer in Nitric-oxide Synthase*[boxs]." Journal of biological chemistry **279**(36): 37918-37927.
Genz, L. R., T. Mulvaney, S. Nair and M. Topf (2023). "PICKLUSTER: a protein-interface clustering and analysis plug-in for UCSF ChimeraX." Bioinformatics **39**(11): btad629.
Gherasim, C., L. Hannibal, D. Rajagopalan, D. W. Jacobsen and R. Banerjee (2013). "The C-terminal domain of CblD interacts with CblC and influences intracellular cobalamin partitioning." Biochimie **95**(5): 1023-1032.
Gille, D. and A. Schmid (2015). "Vitamin B12 in meat and dairy products." Nutrition reviews **73**(2): 106-115.
Gingras, A.-C., M. Gstaiger, B. Raught and R. Aebersold (2007). "Analysis of protein complexes using mass spectrometry." Nature reviews Molecular cell biology **8**(8): 645-654.
Glaeser, R. M. (2016). "How good can cryo-EM become?" Nature methods **13**(1): 28-32.
Glaeser, R. M. (2018). "Proteins, interfaces, and cryo-EM grids." Current opinion in colloid & interface science **34**: 1-8.
Glaeser, R. M., E. Nogales and W. Chiu (2021). Single-particle Cryo-EM of biological macromolecules, IOP publishing.

Gonen, S. (2021). "Progress towards cryoEM: negative-stain procedures for biological samples." cryoEM: Methods and Protocols: 115-123.
Gouda, H., Z. Li, M. Ruetz and R. Banerjee (2023). "Coordination Chemistry Controls Coenzyme B12 Synthesis by Human Adenosine Triphosphate: Cob (I) alamin Adenosyltransferase." Inorganic Chemistry **62**(32): 12630-12633.
Goulding, C. W., D. Postigo and R. G. Matthews (1997). "Cobalamin-dependent methionine synthase is a modular protein with distinct regions for binding homocysteine, methyltetrahydrofolate, cobalamin, and adenosylmethionine." Biochemistry **36**(26): 8082-8091.
Grasso, D., S. Galderisi, A. Santucci and A. Bernini (2023). "Pharmacological chaperones and protein conformational diseases: approaches of computational structural biology." International Journal of Molecular Sciences **24**(6): 5819.
Green, R., D. Jacobsen, S. Van Tonder, M. Kew and J. Metz (1982). "Absorption of biliary cobalamin in baboons following total gastrectomy." The Journal of Laboratory and Clinical Medicine **100**(5): 771-777.
Guéant, J.-L., R.-M. Guéant-Rodriguez, V. J. Kosgei and D. Coelho (2022). "Causes and consequences of impaired methionine synthase activity in acquired and inherited disorders of vitamin B12 metabolism." Critical reviews in biochemistry and molecular biology **57**(2): 133-155.
Hamdane, D., C. Xia, S.-C. Im, H. Zhang, J.-J. P. Kim and L. Waskell (2009). "Structure and function of an NADPH-cytochrome P450 oxidoreductase in an open conformation capable of reducing cytochrome P450." Journal of Biological Chemistry **284**(17): 11374-11384.
Han, B.-G., A. Avila-Sakar, J. Remis and R. M. Glaeser (2023). "Challenges in making ideal cryo-EM samples." Current Opinion in Structural Biology **81**: 102646.
Haque, M. M., M. Bayachou, J. Tejero, C. T. Kenney, N. M. Pearl, S. C. Im, L. Waskell and D. J. Stuehr (2014). "Distinct conformational behaviors of four mammalian dual-flavin reductases (cytochrome P450 reductase, methionine synthase reductase, neuronal nitric oxide synthase, endothelial nitric oxide synthase) determine their unique catalytic profiles." The FEBS journal **281**(23): 5325-5340.
Hawtin, S. R. (2006). "Pharmacological chaperone activity of SR49059 to functionally recover misfolded mutations of the vasopressin V1a receptor." Journal of Biological Chemistry **281**(21): 14604-14614.
He, X., C. Tan, F. Wang, Y. Wang, R. Zhou, D. Cui, W. You, H. Zhao, J. Ren and B. Feng (2016). "Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair." Nucleic acids research **44**(9): e85-e85.
Henderson, R. (2013). "Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise." Proceedings of the National Academy of Sciences **110**(45): 18037-18041.
Hodgkin, D. C., J. Kamper, M. Mackay, J. Pickworth, K. N. Trueblood and J. G. White (1956). "Structure of vitamin B12." Nature **178**(4524): 64-66.
Hodgkin, D. M. C., J. Kamper, J. Lindsey, M. MacKay, J. Pickworth, J. Robertson, C. Brink-Shoemaker, J. G. White, R. Prosen and K. Trueblood (1957). "The structure of vitamin B12. I. An outline of the crystallographic investigation of vitamin B12." Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **242**(1229): 228-263.
Hoover, D. M., J. T. Jarrett, R. H. Sands, W. R. Dunham, M. L. Ludwig and R. G. Matthews (1997). "Interaction of Escherichia coli cobalamin-dependent methionine synthase and its physiological partner flavodoxin: binding of flavodoxin leads to axial ligand dissociation from the cobalamin cofactor." Biochemistry **36**(1): 127-138.
Huemer, M. and M. R. Baumgartner (2019). "The clinical presentation of cobalamin-related disorders: from acquired deficiencies to inborn errors of absorption and intracellular pathways." Journal of inherited metabolic disease **42**(4): 686-705.

Huemer, M., D. Diodato, B. Schwahn, M. Schiff, A. Bandeira, J.-F. Benoist, A. Burlina, R. Cerone, M. L. Couce and A. Garcia-Cazorla (2017). "Guidelines for diagnosis and management of the cobalamin-related remethylation disorders cblC, cblD, cblE, cblF, cblG, cblJ and MTHFR deficiency." Journal of inherited metabolic disease **40**: 21-48.
Jarrett, J. T., C. Y. Choi and R. G. Matthews (1997). "Changes in protonation associated with substrate binding and cob (I) alamin formation in cobalamin-dependent methionine synthase." Biochemistry **36**(50): 15739-15748.
Jarrett, J. T., C. L. Drennan, M. Amaratunga, J. D. Scholten, M. L. Ludwig and R. G. Matthews (1996). "A protein radical cage slows photolysis of methylcobalamin in methionine synthase from Escherichia coli." Bioorganic & Medicinal Chemistry **4**(8): 1237-1246.
Jarrett, J. T., C. W. Goulding, K. Fluhr, S. Huang and R. G. Matthews (1997). Purification and assay of cobalamin-dependent methionine synthase from Escherichia coli. Methods in enzymology, Elsevier. **281:** 196-213.
Jarrett, J. T., D. M. Hoover, M. L. Ludwig and R. G. Matthews (1998). "The mechanism of adenosylmethionine-dependent activation of methionine synthase: a rapid kinetic analysis of intermediates in reductive methylation of cob (II) alamin enzyme." Biochemistry **37**(36): 12649-12658.
Jarrett, J. T., S. Huang and R. G. Matthews (1998). "Methionine synthase exists in two distinct conformations that differ in reactivity toward methyltetrahydrofolate, adenosylmethionine, and flavodoxin." Biochemistry **37**(16): 5372-5382.
Jayaraman, M., P. M. Buck, A. A. Ignatius, K. R. King and W. Wang (2014). "Agitation-induced aggregation and subvisible particulate formation in model proteins." European Journal of Pharmaceutics and Biopharmaceutics **87**(2): 299-309.
Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek and A. Potapenko (2021). "Highly accurate protein structure prediction with AlphaFold." Nature **596**(7873): 583-589.
Jusufi, J., T. Suormala, P. Burda, B. Fowler, D. S. Froese and M. R. Baumgartner (2014). "Characterization of functional domains of the cblD (MMADHC) gene product." Journal of inherited metabolic disease **37**: 841-849.
Kelly, J. J., D. Tranter, E. Pardon, G. Chi, H. Kramer, L. Happonen, K. M. Knee, J. M. Janz, J. Steyaert and C. Bulawa (2022). "Snapshots of actin and tubulin folding inside the TRiC chaperonin." Nature structural & molecular biology **29**(5): 420-429.
Kim, J., C. Gherasim and R. Banerjee (2008). "Decyanation of vitamin B12 by a trafficking chaperone." Proceedings of the National Academy of Sciences **105**(38): 14551-14554.
Kim, J., L. Hannibal, C. Gherasim, D. W. Jacobsen and R. Banerjee (2009). "A human vitamin B12 trafficking protein uses glutathione transferase activity for processing alkylcobalamins." Journal of Biological Chemistry **284**(48): 33418-33424.
Kirkland, E. J. (1998). Advanced computing in electron microscopy, Springer.
Kitai, K., K. Kawaguchi, T. Tomohiro, M. Morita, T. So and T. Imanaka (2021). "The lysosomal protein ABCD4 can transport vitamin B12 across liposomal membranes in vitro." Journal of Biological Chemistry **296**.
Klebl, D. P., R. W. Kay, F. Sobott, N. Kapur and S. P. Muench (2022). "Towards sub-millisecond cryo-EM grid preparation." Faraday Discussions **240**: 33-43.
Koutmos, M., S. Datta, K. A. Pattridge, J. L. Smith and R. G. Matthews (2009). "Insights into the reactivation of cobalamin-dependent methionine synthase." Proceedings of the National Academy of Sciences **106**(44): 18527-18532.
Koutmos, M., C. Gherasim, J. L. Smith and R. Banerjee (2011). "Structural basis of multifunctionality in a vitamin B12-processing enzyme." Journal of Biological Chemistry **286**(34): 29780-29787.

Koutmos, M., R. Pejchal, T. M. Bomer, R. G. Matthews, J. L. Smith and M. L. Ludwig (2008). "Metal active site elasticity linked to activation of homocysteine in methionine synthases." Proceedings of the National Academy of Sciences **105**(9): 3286-3291.
Kozlowski, P. M., J. Kuta and W. Galezowski (2007). "Reductive cleavage mechanism of methylcobalamin: elementary steps of Co− C bond breaking." The Journal of Physical Chemistry B **111**(26): 7638-7645.
Kripps, K. A., L. Sremba, A. A. Larson, J. L. Van Hove, H. Nguyen, E. L. Wright, D. M. Mirsky, D. Watkins, D. S. Rosenblatt and D. Ketteridge (2022). "Methionine synthase deficiency: variable clinical presentation and benefit of early diagnosis and treatment." Journal of inherited metabolic disease **45**(2): 157-168.
Lai, S.-C., Y. Nakayama, J. M. Sequeira, B. J. Wlodarczyk, R. M. Cabrera, R. H. Finnell, T. Bottiglieri and E. V. Quadros (2013). "The transcobalamin receptor knockout mouse: a model for vitamin B12 deficiency in the central nervous system." The FASEB Journal **27**(6): 2468.
Langmore, J. P. and M. F. Smith (1992). "Quantitative energy-filtered electron microscopy of biological molecules in ice." Ultramicroscopy **46**(1-4): 349-373.
Lapatto, R., T. Blundell, A. Hemmings, J. Overington, A. Wilderspin, S. Wood, J. R. Merson, P. J. Whittle, D. E. Danley and K. F. Geoghegan (1989). "X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes." Nature **342**(6247): 299-302.
Leclerc, D., E. Campeau, P. Goyette, C. Adjalla, B. Christensen, M. Ross, P. Eydoux, D. Rosenblatt, R. Rozen and R. Gravel (1996). "Human methionine synthase: cDNA cloning and identification of mutations in patients of the cblG complementation group of folate/cobalamin disorders." Human molecular genetics **5**(12): 1867-1874.
Ledley, F. D. and D. S. Rosenblatt (1997). "Mutations in mut methylmalonic acidemia: clinical and enzymatic correlations." Human mutation **9**(1): 1-6.
Lempp, T. J., T. Suormala, R. Siegenthaler, E. R. Baumgartner, B. Fowler, B. Steinmann and M. R. Baumgartner (2007). "Mutation and biochemical analysis of 19 probands with mut0 and 13 with mut− methylmalonic aciduria: identification of seven novel mutations." Molecular genetics and metabolism **90**(3): 284-290.
Lerner-Ellis, J. P., N. Anastasio, J. Liu, D. Coelho, T. Suormala, M. Stucki, A. D. Loewy, S. Gurd, E. Grundberg and C. F. Morel (2009). "Spectrum of mutations in MMACHC, allelic expression, and evidence for genotype–phenotype correlations." Human mutation **30**(7): 1072-1081.
Li, X., P. Mooney, S. Zheng, C. R. Booth, M. B. Braunfeld, S. Gubbens, D. A. Agard and Y. Cheng (2013). "Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM." Nature methods **10**(6): 584-590.
Li, Z., R. Mascarenhas, U. T. Twahir, A. Kallon, A. Deb, M. Yaw, J. Penner-Hahn, M. Koutmos, K. Warncke and R. Banerjee (2020). "An interprotein Co–S coordination complex in the B12-trafficking pathway." Journal of the American Chemical Society **142**(38): 16334-16345.
Li, Z., R. Mascarenhas, U. T. Twahir, A. Kallon, A. Deb, M. Yaw, J. Penner-Hahn, M. Koutmos, K. Warncke and R. Banerjee (2020). "An Interprotein Co–S Coordination Complex in the B_{12}-Trafficking Pathway."
Liptak, M. D., S. Datta, R. G. Matthews and T. C. Brunold (2008). "Spectroscopic study of the cobalamin-dependent methionine synthase in the activation conformation: effects of the Y1139 residue and S-adenosylmethionine on the B12 cofactor." Journal of the American Chemical Society **130**(48): 16374-16381.
Liptak, M. D., A. S. Fleischhacker, R. G. Matthews, J. Telser and T. C. Brunold (2009). "Spectroscopic and Computational Characterization of the Base-off Forms of Cob (II) alamin." The Journal of Physical Chemistry B **113**(15): 5245-5254.
Liu, T., F. Huo, J. Li, J. Chao, Y. Zhang and C. Yin (2016). "A fast response and high sensitivity thiol fluorescent probe in living cells." Sensors and Actuators B: Chemical **232**: 619-624.
Manoli, I., J. L. Sloan and C. P. Venditti (2016). "Isolated methylmalonic acidemia."

Marcus, R. A. and N. Sutin (1985). "Electron transfers in chemistry and biology." Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics **811**(3): 265-322.
Mascarenhas, R., H. Gouda, M. Ruetz and R. Banerjee (2022). Human B12-dependent enzymes: Methionine synthase and Methylmalonyl-CoA mutase. Methods in enzymology, Elsevier. **668:** 309-326.
Mascarenhas, R., A. Guha, Z. Li, M. Ruetz, S. An, J. Seravalli and R. Banerjee (2023). "Cobalt–Sulfur Coordination Chemistry Drives B12 Loading onto Methionine Synthase." Journal of the American Chemical Society **145**(45): 24678-24689.
Mascarenhas, R., M. Ruetz, L. McDevitt, M. Koutmos and R. Banerjee (2020). "Mobile loop dynamics in adenosyltransferase control binding and reactivity of coenzyme B12." Proceedings of the National Academy of Sciences **117**(48): 30412-30422.
Masrati, G., M. Landau, N. Ben-Tal, A. Lupas, M. Kosloff and J. Kosinski (2021). "Integrative structural biology in the era of accurate structure prediction." Journal of molecular biology **433**(20): 167127.
Matte, J., M. Britten and C. Girard (2014). "The importance of milk as a source of vitamin B12 for human nutrition." Animal Frontiers **4**(2): 32-37.
Matthews, R. G., M. Koutmos and S. Datta (2008). "Cobalamin-dependent and cobamide-dependent methyltransferases." Current Opinion in Structural Biology **18**(6): 658-666.
McClure, S. M., P. L. Ahl and J. T. Blue (2018). "High throughput differential scanning fluorimetry (DSF) formulation screening with complementary dyes to assess protein unfolding and aggregation in presence of surfactants." Pharmaceutical research **35**: 1-10.
McCorvie, T. J., D. Ferreira, W. W. Yue and D. S. Froese (2023). "The complex machinery of human cobalamin metabolism." Journal of Inherited Metabolic Disease **46**(3): 406-420.
Meints, C. E., F. S. Gustafsson, N. S. Scrutton and K. R. Wolthers (2011). "Tryptophan 697 modulates hydride and interflavin electron transfer in human methionine synthase reductase." Biochemistry **50**(51): 11131-11142.
Mendoza, J., M. Purchal, K. Yamada and M. Koutmos (2023). "Structure of full-length cobalamin-dependent methionine synthase and cofactor loading captured in crystallo." Nature Communications **14**(1): 6365.
Mera, P. E. and J. C. Escalante-Semerena (2010). "Dihydroflavin-driven adenosylation of 4-coordinate Co (II) corrinoids: are cobalamin reductases enzymes or electron transfer proteins?" Journal of Biological Chemistry **285**(5): 2911-2917.
Miller, M., J. Schneider, B. K. Sathyanarayana, M. V. ToTH, G. R. Marshall, L. Clawson, L. Selk, S. B. Kent and A. Wlodawer (1989). "Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution." Science **246**(4934): 1149-1152.
Mirdita, M., K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov and M. Steinegger (2022). "ColabFold: making protein folding accessible to all." Nature methods **19**(6): 679-682.
Mooij, W. T., E. Mitsiki and A. Perrakis (2009). "ProteinCCD: enabling the design of protein truncation constructs for expression and crystallization experiments." Nucleic acids research **37**(suppl_2): W402-W405.
Moreau, M. J., I. Morin, S. P. Askin, A. Cooper, N. J. Moreland, S. G. Vasudevan and P. M. Schaeffer (2012). "Rapid determination of protein stability and ligand binding by differential scanning fluorimetry of GFP-tagged proteins." RSC advances **2**(31): 11892-11900.
Mori, K. and T. Toraya (1999). "Mechanism of reactivation of coenzyme B12-dependent diol dehydratase by a molecular chaperone-like reactivating factor." Biochemistry **38**(40): 13170-13178.
Murataliev, M. B., R. Feyereisen and F. A. Walker (2004). "Electron transfer by diflavin reductases." Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics **1698**(1): 1-26.
Nakane, T., A. Kotecha, A. Sente, G. McMullan, S. Masiulis, P. M. Brown, I. T. Grigoras, L. Malinauskaite, T. Malinauskas and J. Miehling (2020). "Single-particle cryo-EM at atomic resolution." Nature **587**(7832): 152-156.

Nielsen, M. J., M. R. Rasmussen, C. B. Andersen, E. Nexø and S. K. Moestrup (2012). "Vitamin B12 transport from food to the body's cells—a sophisticated, multistep pathway." Nature reviews Gastroenterology & hepatology **9**(6): 345-354.
Niesen, F. H., H. Berglund and M. Vedadi (2007). "The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability." Nature protocols **2**(9): 2212-2221.
Obeid, R., S. N. Fedosov and E. Nexo (2015). "Cobalamin coenzyme forms are not likely to be superior to cyano-and hydroxyl-cobalamin in prevention or treatment of cobalamin deficiency." Molecular nutrition & food research **59**(7): 1364-1372.
Olinares, P. D. B., J. Y. Kang, E. Llewellyn, C. Chiu, J. Chen, B. Malone, R. M. Saecker, E. A. Campbell, S. A. Darst and B. T. Chait (2021). "Native mass spectrometry-based screening for optimal sample preparation in single-particle Cryo-EM." Structure **29**(2): 186-195. e186.
Olteanu, H. and R. Banerjee (2001). "Human methionine synthase reductase, a soluble P-450 reductase-like dual flavoprotein, is sufficient for NADPH-dependent methionine synthase activation." Journal of Biological Chemistry **276**(38): 35558-35563.
Olteanu, H. and R. Banerjee (2003). "Redundancy in the pathway for redox regulation of mammalian methionine synthase: reductive activation by the dual flavoprotein, novel reductase 1." Journal of Biological Chemistry **278**(40): 38310-38314.
Padovani, D., T. Labunska, B. A. Palfey, D. P. Ballou and R. Banerjee (2008). "Adenosyltransferase tailors and delivers coenzyme B12." Nature chemical biology **4**(3): 194-196.
Park, K. and T. C. Brunold (2013). "Combined spectroscopic and computational analysis of the vibrational properties of vitamin B12 in its Co3+, Co2+, and Co1+ oxidation states." The Journal of Physical Chemistry B **117**(18): 5397-5410.
Peet, M. J., R. Henderson and C. J. Russo (2019). "The energy dependence of contrast and damage in electron cryomicroscopy of biological molecules." Ultramicroscopy **203**: 125-131.
Penczek, P. A. (2020). "Reliable cryo-EM resolution estimation with modified Fourier shell correlation." IUCrJ **7**(6): 995-1008.
Perry, C. B. and H. M. Marques (2004). "Fifty years of X-ray crystallography of vitamin B12 and its derivatives." South African journal of science **100**(7): 368-380.
Pettersen, E. F., T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris and T. E. Ferrin (2021). "UCSF ChimeraX: Structure visualization for researchers, educators, and developers." Protein science **30**(1): 70-82.
Plessl, T., C. Bürer, S. Lutz, W. W. Yue, M. R. Baumgartner and D. S. Froese (2017). "Protein destabilization and loss of protein-protein interaction are fundamental mechanisms in cblA-type methylmalonic aciduria." Human mutation **38**(8): 988-1001.
Pollock, R. J. and S. Kaufman (1978). "Dihydropteridine reductase may function in tetrahydrofolate metabolism." Journal of Neurochemistry **31**(1): 115-123.
Punjani, A. and D. J. Fleet (2023). "3DFlex: determining structure and motion of flexible proteins from cryo-EM." Nature Methods **20**(6): 860-870.
Punjani, A., J. L. Rubinstein, D. J. Fleet and M. A. Brubaker (2017). "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination." Nature methods **14**(3): 290-296.
Pupavac, M., D. Watkins, F. Petrella, S. Fahiminiya, A. Janer, W. Cheung, A. C. Gingras, T. Pastinen, J. Muenzer and J. Majewski (2016). "Inborn error of cobalamin metabolism associated with the intracellular accumulation of transcobalamin-bound cobalamin and mutations in ZNF143, which codes for a transcriptional activator." Human mutation **37**(9): 976-982.
Qiao, Y., H. Molina, A. Pandey, J. Zhang and P. A. Cole (2006). "Chemical rescue of a mutant enzyme in living cells." Science **311**(5765): 1293-1297.
Quadros, E. V. (2010). "Advances in the understanding of cobalamin assimilation and metabolism." British journal of haematology **148**(2): 195-204.

Quintana, A. M., H.-C. Yu, A. Brebner, M. Pupavac, E. A. Geiger, A. Watson, V. L. Castro, W. Cheung, S.-H. Chen and D. Watkins (2017). "Mutations in THAP11 cause an inborn error of cobalamin metabolism and developmental abnormalities." Human molecular genetics **26**(15): 2838-2849.
Ran, F., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott and F. Zhang (2013). "Genome engineering using the CRISPR-Cas9 system." Nature protocols **8**(11): 2281-2308.
Raynal-Ljutovac, K., G. Lagriffoul, P. Paccard, I. Guillet and Y. Chilliard (2008). "Composition of goat and sheep milk products: An update." Small ruminant research **79**(1): 57-72.
Reboul, C. F., M. Eager, D. Elmlund and H. Elmlund (2018). "Single-particle cryo-EM—Improved ab initio 3D reconstruction with SIMPLE/PRIME." Protein Science **27**(1): 51-61.
Rosenblatt, D. S., A. Hosack, N. V. Matiaszuk, B. A. Cooper and R. Laframboise (1985). "Defect in vitamin B12 release from lysosomes: newly described inborn error of vitamin B12 metabolism." Science **228**(4705): 1319-1321.
Roth, J. R., J. Lawrence and T. Bobik (1996). "Cobalamin (coenzyme B12): synthesis and biological significance." Annual review of microbiology **50**(1): 137-181.
Ruetz, M., G. C. Campanello, L. McDevitt, A. L. Yokom, P. K. Yadav, D. Watkins, D. S. Rosenblatt, M. D. Ohi, D. R. Southworth and R. Banerjee (2019). "Allosteric regulation of oligomerization by a B12 trafficking G-protein is corrupted in methylmalonic aciduria." Cell chemical biology **26**(7): 960-969. e964.
Ruetz, M., A. Shanmuganathan, C. Gherasim, A. Karasik, R. Salchner, C. Kieninger, K. Wurst, R. Banerjee, M. Koutmos and B. Kräutler (2017). "Antivitamin B12 Inhibition of the Human B12-Processing Enzyme CblC: Crystal Structure of an Inactive Ternary Complex with Glutathione as the Cosubstrate." Angewandte Chemie International Edition **56**(26): 7387-7392.
Russo, C. J. and L. A. Passmore (2016). "Ultrastable gold substrates: properties of a support for high-resolution electron cryomicroscopy of biological specimens." Journal of structural biology **193**(1): 33-44.
Salbaum, J. M. and C. Kappen (2012). "Genetic and epigenomic footprints of folate." Progress in molecular biology and translational science **108**: 129-158.
Saldaño, T., N. Escobedo, J. Marchetti, D. J. Zea, J. Mac Donagh, A. J. Velez Rueda, E. Gonik, A. García Melani, J. Novomisky Nechcoff and M. N. Salas (2022). "Impact of protein conformational diversity on AlphaFold predictions." Bioinformatics **38**(10): 2742-2748.
Salerno, E. V., N. A. Miller, A. Konar, Y. Li, C. Kieninger, B. Kräutler and R. J. Sension (2020). "Ultrafast Excited State Dynamics and Fluorescence from Vitamin B12 and Organometallic [Co]–C≡ C–R Cobalamins." The Journal of Physical Chemistry B **124**(30): 6651-6656.
Sato, D. and T. Nozaki (2009). "Methionine gamma-lyase: The unique reaction mechanism, physiological roles, and therapeutic applications against infectious diseases and cancers." IUBMB life **61**(11): 1019-1028.
Scarff, C. A., M. J. Fuller, R. F. Thompson and M. G. Iadanza (2018). "Variations on negative stain electron microscopy methods: tools for tackling challenging systems." JoVE (Journal of Visualized Experiments)(132): e57199.
Schägger, H. and G. von Jagow (1991). "Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form." Analytical biochemistry **199**(2): 223-231.
Scheres, S. H. (2012). "RELION: implementation of a Bayesian approach to cryo-EM structure determination." Journal of structural biology **180**(3): 519-530.
Schneider, Z. and A. Stroinski (1987). "Biosynthesis of vitamin B12." Comprehensive B **12**: 93-110.
Skiniotis, G. and D. R. Southworth (2015). "Single-particle cryo-electron microscopy of macromolecular complexes." Journal of Electron Microscopy **65**(1): 9-22.
Stabler, S. P. (2020). Vitamin B12. Present knowledge in nutrition, Elsevier**:** 257-271.

Stucki, M., D. Coelho, T. Suormala, P. Burda, B. Fowler and M. R. Baumgartner (2012). "Molecular mechanisms leading to three different phenotypes in the cblD defect of intracellular cobalamin metabolism." Human molecular genetics **21**(6): 1410-1418.
Suormala, T., M. R. Baumgartner, D. Coelho, P. Zavadakova, V. Kozich, H. G. Koch, M. Berghaüser, J. E. Wraith, A. Burlina and A. Sewell (2004). "The cblD defect causes either isolated or combined deficiency of methylcobalamin and adenosylcobalamin synthesis." Journal of Biological Chemistry **279**(41): 42742-42749.
Swaffield, J. C. and S. A. Johnston (1996). "Affinity purification of proteins binding to GST fusion proteins." Current protocols in molecular biology **33**(1): 20.22. 21-20.22. 10.
Takeda, H., N. Goshima and N. Nomura (2010). "High-throughput kinase assay based on surface plasmon resonance." Surface Plasmon Resonance: Methods and Protocols: 131-145.
Tanner, S. M., M. Aminoff, F. A. Wright, S. Liyanarachchi, M. Kuronen, A. Saarinen, O. Massika, H. Mandel, H. Broch and A. De la Chapelle (2003). "Amnionless, essential for mouse gastrulation, is mutated in recessive hereditary megaloblastic anemia." Nature genetics **33**(3): 426-429.
Terwilliger, T. C., R. W. Grosse-Kunstleve, P. V. Afonine, N. W. Moriarty, P. H. Zwart, L.-W. Hung, R. J. Read and P. D. Adams (2008). "Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard." Acta Crystallographica Section D: Biological Crystallography **64**(1): 61-69.
Tordai, H., O. Torres, M. Csepi, R. Padányi, G. L. Lukács and T. Hegedűs (2024). "Analysis of AlphaMissense data in different protein groups and structural context." Scientific Data **11**(1): 495.
Tyanova, S., T. Temu, P. Sinitcyn, A. Carlson, M. Y. Hein, T. Geiger, M. Mann and J. Cox (2016). "The Perseus computational platform for comprehensive analysis of (prote) omics data." Nature methods **13**(9): 731-740.
Vargo, M. A., L. Nguyen and R. F. Colman (2004). "Subunit interface residues of glutathione S-transferase A1-1 that are important in the monomer– dimer equilibrium." Biochemistry **43**(12): 3327-3335.
Wagner, T., F. Merino, M. Stabrin, T. Moriya, C. Antoni, A. Apelbaum, P. Hagel, O. Sitsel, T. Raisch and D. Prumbaum (2019). "SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM." Communications biology **2**(1): 218.
Watanabe, F. and T. Bito (2018). "Vitamin B12 sources and microbial interaction." Experimental Biology and Medicine **243**(2): 148-158.
Watkin, D. and D. S. Rosenblatt (1989). "Functional methionine synthase deficiency (cblE and cblG): clinical and biochemical heterogeneity." American journal of medical genetics **34**(3): 427-434.
Watkins, D., M. Ru, H.-Y. Hwang, C. D. Kim, A. Murray, N. S. Philip, W. Kim, H. Legakis, T. Wai and J. F. Hilton (2002). "Hyperhomocysteinemia due to methionine synthase deficiency, cblG: structure of the MTR gene, genotype diversity, and recognition of a common mutation, P1173L." The American Journal of Human Genetics **71**(1): 143-153.
Watkins, M. B., H. Wang, A. Burnim and N. Ando (2023). "Conformational switching and flexibility in cobalamin-dependent methionine synthase studied by small-angle X-ray scattering and cryoelectron microscopy." Proceedings of the National Academy of Sciences **120**(26): e2302531120.
Weissmann, F., G. Petzold, R. VanderLinden, P. J. Huis In't Veld, N. G. Brown, F. Lampert, S. Westermann, H. Stark, B. A. Schulman and J.-M. Peters (2016). "biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes." Proceedings of the National Academy of Sciences **113**(19): E2564-E2569.
Wittig, I., H.-P. Braun and H. Schägger (2006). "Blue native PAGE." Nature protocols **1**(1): 418-428.

Wolthers, K. R., J. Basran, A. W. Munro and N. S. Scrutton (2003). "Molecular dissection of human methionine synthase reductase: determination of the flavin redox potentials in full-length enzyme and isolated flavin-binding domains." Biochemistry **42**(13): 3911-3920.
Wolthers, K. R., X. Lou, H. S. Toogood, D. Leys and N. S. Scrutton (2007). "Mechanism of coenzyme binding to human methionine synthase reductase revealed through the crystal structure of the FNR-like module and isothermal titration calorimetry." Biochemistry **46**(42): 11833-11844.
Wolthers, K. R. and N. S. Scrutton (2007). "Protein interactions in the human methionine synthase− methionine synthase reductase complex and implications for the mechanism of enzyme reactivation." Biochemistry **46**(23): 6696-6709.
Wolthers, K. R. and N. S. Scrutton (2009). "Cobalamin uptake and reactivation occurs through specific protein interactions in the methionine synthase–methionine synthase reductase complex." The FEBS Journal **276**(7): 1942-1951.
Wolthers, K. R., H. S. Toogood, T. A. Jowitt, K. R. Marshall, D. Leys and N. S. Scrutton (2007). "Crystal structure and solution characterization of the activation domain of human methionine synthase." The FEBS Journal **274**(3): 738-750.
Wu, M. and G. C. Lander (2020). "How low can we go? Structure determination of small biological complexes using single-particle cryo-EM." Current opinion in structural biology **64**: 9-16.
Wu, X. and T. A. Rapoport (2021). "Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies)." Proceedings of the National Academy of Sciences **118**(41): e2115001118.
Yamada, K., C. Gherasim, R. Banerjee and M. Koutmos (2015). "Structure of human B12 trafficking protein CblD reveals molecular mimicry and identifies a new subfamily of nitro-FMN reductases." Journal of Biological Chemistry **290**(49): 29155-29166.
Yamada, K., R. A. Gravel, T. Toraya and R. G. Matthews (2006). "Human methionine synthase reductase is a molecular chaperone for human methionine synthase." Proceedings of the National Academy of Sciences **103**(25): 9476-9481.
Yamada, K. and M. Koutmos (2018). "The folate-binding module of Thermus thermophilus cobalamin-dependent methionine synthase displays a distinct variation of the classical TIM barrel: a TIM barrel with atwist'." Acta Crystallographica Section D: Structural Biology **74**(1): 41-51.
Yamada, K., S. Yamada, T. Tobimatsu and T. Toraya (1999). "Heterologous high level expression, purification, and enzymological properties of recombinant rat cobalamin-dependent methionine synthase." Journal of Biological Chemistry **274**(50): 35571-35576.
Yan, C., J. Hang, R. Wan, M. Huang, C. C. Wong and Y. Shi (2015). "Structure of a yeast spliceosome at 3.6-angstrom resolution." Science **349**(6253): 1182-1191.
Yip, K. M., N. Fischer, E. Paknia, A. Chari and H. Stark (2020). "Atomic-resolution protein structure determination by cryo-EM." Nature **587**(7832): 157-161.
Yu, H.-C., J. L. Sloan, G. Scharer, A. Brebner, A. M. Quintana, N. P. Achilly, I. Manoli, C. R. Coughlin, E. A. Geiger and U. Schneck (2013). "An X-linked cobalamin disorder caused by mutations in transcriptional coregulator HCFC1." The American Journal of Human Genetics **93**(3): 506-514.
Zhang, J., G.-c. Liu, X.-l. Dai, J. Wang, M.-h. Jin, N.-n. Mi and S.-q. Wang (2020). "The N-terminus of MTRR plays a role in MTR reactivation cycle beyond electron transfer." Bioorganic Chemistry **100**: 103836.
Zheng, S. Q., E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng and D. A. Agard (2017). "MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy." Nature methods **14**(4): 331-332.
Zheng, Y. and L. C. Cantley (2019). "Toward a better understanding of folate metabolism in health and disease." Journal of Experimental Medicine **216**(2): 253-266.

Zivanov, J., T. Nakane, B. O. Forsberg, D. Kimanius, W. J. Hagen, E. Lindahl and S. H. Scheres (2018). "New tools for automated high-resolution cryo-EM structure determination in RELION-3." elife **7**: e42166.
Zivanov, J., T. Nakane and S. H. Scheres (2020). "Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1." IUCrJ **7**(2): 253-267.