Structural characterization of the CNNM protein, a putative magnesium transporter

Rayan Fakih
Department of Biochemistry
McGill University, Montreal

December 2020

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Mg^{2+}: Magnesium cation

CNNM: Cyclin M, CBS-pair domain mediated divalent metal cation transport mediators

CBS: Cystathionine β-synthase

PRL: Phosphatase of regenerating liver

CNBH: Cyclic nucleotide binding homology (domain)

CorC/HlyC: Cobalt resistance C/Hemolysin C-like (C-terminal domain)

DUF21: Domain of unknown function 21

c(N)MP: cyclic (nucleotide) monophosphate

NMR: Nuclear magnetic resonance

SV-AUC: Sedimentation velocity analytical ultracentrifugation

RMSD: Root-mean-square deviation

AMP-PNP: β,γ-Imidoadenosine 5′-triphosphate

tCNNM: CNNM from a thermophilic bacterium

DDM: N-Dodecyl β-D-maltoside

MALDI MS: Matrix-assisted laser desorption/ionization mass spectrometry

OG488: Oregon Green 488

TIRF: Total internal reflective fluorescence (microscopy)

ABSTRACT

Mg^{2+} is a biologically important cation required for over 600 enzymatic processes, ranging from energy metabolism to cell cycle regulation. Mg^{2+} homeostasis is mediated through the action of various ion transporters. CNNMs (CBS domain metal ion transport mediators) are integral membrane proteins that mediate metal ion (e.g. Mg^{2+}) homeostasis in organisms from bacteria to humans.

The structure of the CNNMs consists of a conserved transmembrane domain (DUF21) and regulatory cytosolic cystathionine-beta-synthase (CBS-pair or Bateman) domain. Eukaryotic CNNMs, in particular, contain an additional extracellular domain, and a putative C-terminal cyclic nucleotide binding homology (CNBH) domain. While these proteins appear to control Mg^{2+} transport, with disease mutations leading to several disorders related to Mg^{2+} transport and localization, their mechanism is unknown and their status as Mg^{2+} transporters remains actively debated. No full three-dimensional structures of the proteins exist, and only one domain (the CBS-pair) has been extensively characterized thus far. The goal of my thesis research is to take a step-by-step approach in characterizing these proteins, starting with individual soluble domains and leading up to the full transmembrane protein.

Using X-ray crystallography and various *in-vitro* techniques (NMR, scanning differential fluorimetry, sedimentation velocity), we obtained a structure of the CNBH domain and established it as a dimerization interface for CNNMs. We also identified a structural interplay between the connected CNBH and CBS-pair domains by solving the structure for the full cytosolic fragment, and discovered a link between ligand binding and dimerization in the CBS-pair domain. Finally, we purified and crystallized a larger CNNM fragment, and demonstrated that it forms oligomers in solution using fluorescence microscopy. These findings provide structural insight to the largely uncharacterized CNNMs, and allow us to better understand their biological mechanisms and involvement in Mg^{2+} homeostasis and related diseases.

RESUMÉ

Mg^{2+} est un cation biologiquement important et requis pour plus de 600 processus enzymatiques, allant du métabolisme énergétique à la régulation du cycle cellulaire. L'homéostasie du Mg^{2+} est médiée par l'action de divers transporteurs d'ions. Les CNNMs (médiateurs de transport d'ions métalliques dans le domaine CBS) sont des protéines membranaires intégrales qui interviennent dans l'homéostasie des ions métalliques (ex: Mg^{2+}) dans les organismes, des bactéries aux humains.

La structure des CNNMs contient un domaine transmembranaire conservé (DUF21) et un domaine régulatoire cytosolique cystathionine-bêta-synthase (paire CBS). Les CNNM eucaryotes, en particulier, contiennent un domaine extracellulaire supplémentaire et un domaine d'homologie de liaison de nucléotide cyclique C-terminal putatif (CNBH). Alors que ces protéines semblent contrôler le transport du Mg^{2+}, avec des mutations pathologiques conduisant à plusieurs troubles liés au transport et à la localisation du Mg^{2+}, leur mécanisme est inconnu et leur statut en tant que transporteurs de Mg^{2+} reste vigoureusement débattu. Aucune structure tridimensionnelle complète des protéines n'existe, et seul la paire CBS a été caractérisé à date. Le but de notre projet est de caractériser ces protéines, commençant par les domaines solubles individuels jusqu'à la protéine transmembranaire complète.

En utilisant la cristallographie aux rayons X et diverses techniques *in vitro* (RMN, fluorimétrie différentielle, centrifugation analytique), nous obtenons une structure du domaine CNBH et l'établissons comme interface de dimérisation pour les CNNMs. Nous identifions également une interaction structurelle entre les domaines de la paire CBS et CNBH en résolvant la structure du fragment cytosolique complet. Nous introduisons également un lien entre la dimérisation dans le domaine de la paire CBS et sa liaison aux ligands. Enfin, nous purifions et cristallisons le CNNM complet et démontrons qu'il forme des oligomères en solution à l'aide de techniques de microscopie à fluorescence. Ces résultats fournissent un aperçu structurel des CNNMs largement non caractérisés, ce qui nous permet de mieux comprendre leurs mécanismes biologiques et leur implication dans l'homéostasie du Mg^{2+} et les maladies associées.

ACKNOWLEDGEMENTS

I would like thank the members of the Gehring lab who, for the past 4-5 years, have been like a family to me. I fondly look back on the many events we have organized and attended together, the many birthdays we have celebrated, and the many pastries we have baked and shared with one another. I will definitely miss our famous "cake o'clock"s at 3 in the afternoon.

To Meng and Seby, I warmly thank you for accepting me as your student, for teaching me the basics of a structural biology lab, for trusting me to work alongside you during projects, and for being patient with me whenever I felt discouraged. Seby, you are my role model as a focused and hardworking scientist. I have practically learned how to multitask in the short time of working under you. I wish you the best of luck in the future. To Guennadi, thank you for your insight and for introducing me to the more complicated tools and machines. Your laugh is absolutely memorable and brightens up everyone's mood. To Véronique, thank you for the pleasant conversations as a fellow francophone. You were a motherly figure to all of us. To George, thank you for the great times and for being a long-time brother figure. Anshu, Valeria, Cordelia, Zhidian, Jean, Molly, and Raphael, I enjoyed talking and spending time with all of you. To Robert, my student and partner in crime, I have enjoyed our time together, and mentoring you. You are a very passionate man who is always seeking out new hobbies and skills, and I wish you all the success in the future. Thank you for putting up with my newbie teaching skills and for covering for me whenever I was busy or felt unwell. It was a joy and honor watching you blossom into an independent researcher. To Kalle, thank you for accepting me in your lab and providing me with the means of pursuing my research projects. I enjoyed your enthusiasm and sense of humor. I would also like to thank you for lending an ear and encouraging me in my most vulnerable moments. I wish you and your family all the best.

I would also like to extend my gratitude towards the members of my Research Advisor Committee, Dr. Rikard Blunck, Dr. Alba Guarné, and Dr. Joaquin Ortega for their wise words and insight in helping me move forward with my project. I would like to especially thank Dr. Blunck for allowing me to work in his lab for my labelling experiments, and his student Lena Möller for helping me every step of the way. I hope you enjoyed the cakes I baked for you.

Finally, I would like to thank Dr. Dean Clodfelter from Beckman Coulter for teaching me the basics in analytical ultracentrifugation.

AUTHOR CONTRIBUTION

PhD student Seby Chen and research associate Dr. Guennadi Kozlov from the Gehring lab helped design the experiments, clone the constructs, perform the NMR experiments, and solve the crystal structures for the CNNM cyclic nucleotide binding homology domains and the cytosolic fragments. Seby Chen and I performed the differential scanning fluorimetry. I performed the analytical ultracentrifugation sedimentation velocity experiments on the soluble domains, and large scale purification and crystallization of the prokaryotic tCNNM constructs. Additionally, I helped with large scale purification and crystallization of the human CNNM cytosolic constructs. I also performed maleimide labelling of tCNNM for single molecule subunit counting experiments, which were carried out by Lena Möller from the Blunck lab. Dr. Kalle Gehring oversaw the research.

Part of the work featured in this thesis is already published under the titles “The cyclic nucleotide-binding homology domain of the integral membrane protein CNNM mediates dimerization and is required for Mg^{2+} efflux activity” [1] and “Mg^{2+}-ATP Sensing in CNNM, a Putative Magnesium Transporter” [2] by Chen *et al*, and is referenced accordingly.

CHAPTER 1. Introduction

1.1 Magnesium in biology

Magnesium (Mg^{2+}) is the second most abundant divalent cation in the body, and the most abundant in the cell. Out of the total magnesium content in the human body, nearly 52.9% is found mineralized in the bone, 27.0% in the muscle, 19.3% in the soft tissue, 0.5% in the erythrocytes, and 0.3% in the blood serum [3]. This leaves an intracellular magnesium concentration of 5-20 mM, with the majority found in complex with proteins, membranes, complex anions (e.g. (poly)phosphates, sulphates) and nucleotides (e.g. ATP, GTP), and only 1-5% being found in the ionized form (Mg^{2+}) [4]. The physiological concentration of free cytoplasmic Mg^{2+}varies between 0.5-1.0 mM [5].

Mg^{2+} is the most functionally diverse ion and is an essential element for life. It is required for the catalytic activity of over 600 enzymes, making it a vital mediator of various biochemical processes such as glycolysis (e.g. hexokinase), signalling (e.g. adenylate and guanylate kinase), lipid metabolism (e.g. lipoprotein lipase, lecithin cholesterol acyltransferase), ion transport (e.g. Na^{+}/K^{+}-ATPase, Ca^{2+}-ATPase), nucleotide synthesis (e.g. 5-phosphoribosyl-pyrophosphate synthetase), nucleic acid synthesis (DNA and RNA polymerase), etc. [4] Mg^{2+} is also required for the structural stability of nucleic acids (RNA and DNA), various non-enzymatic proteins and protein complexes (e.g. hemoglobin, calmodulin), ribosomes, porphyrin-containing molecules (e.g. chlorophyll), and mitochondrial membranes [4].

Physiologically, Mg^{2+} helps maintain the electrical potential necessary for nervous and muscular action potential activation and for signal transmission, thereby maintaining proper neurological function, muscle contraction and relaxation, and blood pressure [2]. It also regulates bone formation and resorption, blood clotting (via platelet activated thrombosis), inflammatory immune responses, neurotransmitter metabolism and release, as well as insulin sensitivity for glucose metabolism [2]. Additionally, high intracellular Mg^{2+} was proposed to be associated with cell proliferation and tumor progression, due to the cation's ability to associate with the cellular energy currency ATP as well as with many metabolically-involved enzymes [6-7].

1.2 Magnesium homeostasis

Given that Mg^{2+} is involved in various essential processes, it important that both its intracellular and extracellular concentrations remain at physiological levels for proper cell functioning [8-9]. Deviations from these physiological concentrations invariably leads to disease. Such conditions include hypertension, cardiac arrhythmia, asthma, migraines, bone disorders, and even cancer [5, 8-9].

Homeostasis is tightly regulated within the organelles, the cytoplasm, and the extracellular space through two major mechanisms [11]. First, it can be maintained through the activity of specific chaperone proteins that enter in complex with Mg^{2+} as a means of storing the ion. The second method is through the activity of transmembrane transport proteins that vary their translocation activity in accordance with Mg^{2+} concentration [5].

The chemical and biochemical properties of Mg^{2+} present the cellular system with a significant challenge when transporting the ion across biological membranes. In order for an ion transporter to successfully translocate an ion across a membrane, it must recognize the ion of interest in its hydrated form, strip off most or all of the water of hydration at a selective pore, and accommodate the bare ion before releasing it on the far side of the membrane [12]. While the bare Mg^{2+} has the smallest ionic radius among all biological cations, it has the largest hydration shell; the cation's radius increases 400-fold when bound to water [4]. This is a much larger difference than that seen with Na^{+} and Ca^{2+}, which increase 25-fold, or K^{+}, which only increases fourfold [13-14]. Mg^{2+} is also the most charge dense of all biological cations, holding the waters within its hydration shell tighter than Ca^{2+}, K^{+}, and Na^{2+} by a factor of $10^{3}-10^{4}$ [12]. Moreover, the hydrated Mg^{2+} is more rigid than the other cations, always assuming a restricted octahedral coordination geometry. Mg^{2+} also almost always prefers to coordinate with oxygen (i.e. in water molecules) [15]. The large volume change and the relative strength of interaction with the waters of hydrations thus present a challenge for a Mg^{2+} transporter far greater than any other cation transport system [12]

1.3 Magnesium transport proteins

Despite the biological abundance and importance of Mg^{2+}, there is little information regarding the proteins that transport Mg^{2+}, the mechanisms by which they do so, and their physiological roles within the cell. The most thoroughly characterized Mg^{2+} transport systems to date come from prokaryotes: CorA, MgtA/B, and MgtE [12]. CorA and MgtE regulate Mg^{2+} uptake in cells, and are regarded as the primary suppliers of Mg^{2+} in prokaryotes [16-18]. MgtA and MgtB operate as a P-type ATPase in mediating Mg^{2+} influx [19].

Recently, eukaryotic homologues of some of these prokaryotic Mg^{2+} transporters have been identified and cloned [12]. Mrs2p is a homologue of CorA present in yeast and mammalian inner mitochondrial membranes, and has been demonstrated to mediate Mg^{2+} influx in isolated yeast mitochondria driven by the inner membrane potential [12, 20]. Eukaryotic MgtE homologs are represented by the SLC41 family of solute carriers with three members (SLC41A1−3), which were found to be expressed in several tissues, most notably renal epithelial cells. SLC41A1 and SLC41A2 were shown to transport Mg^{2+} in a voltage-dependent manner regulated by extracellular magnesium. [20].

Several novel transport systems have also been identified in eukaryotes. TRPM6 and TRPM7, belonging to the transient receptor potential (TRP) transporter family, mediate transport of Ca^{2+} and Mg^{2+} in mammals [12, 21]. Claudins mediate paracellular Mg^{2+} transport; Claudin-16 allows for selective re-uptake of Mg^{2+} in the human kidney in particular [12]. MagT1 is considered a dedicated mammalian transporter Mg^{2+}, eliciting large Mg^{2+}-evoked currents with little permeability to other divalent cations in oocytes [12]. Finally, CNNM2 is linked to the transport of Mg^{2+} and other divalent cations (Co^{2+}, Mn^{2+}, Sr^{2+}, Ba^{2+}, Cu^{2+}, and Fe^{2+}) when expressed in Xenopus oocytes [12, 22]. Of these listed systems, the least studied belong to the CNNM family (originally termed ACDP, for ancient conserved domain protein).

1.4 CNNM proteins

In 2003, Wang *et al.* identified the *CNNM* gene family in the 10q23-10q24 chromosome. In most animals, including humans, this gene family comprises of four isoforms : CNNM1-4 [23]. They were initially named ACDP because all isoforms share a domain that is phylogenetically conserved from bacteria to humans [23]. The family was renamed to CNNM1-4 (for Cyclin M 1-4) due the presence of an ancient cyclin box motif and a weak sequence similarity with cyclins, although no cyclin-related function has been proven so far. Instead, the CNNM proteins were shown to reside on the plasma membrane [24], and are now referred to as "CBS domain metal transport mediators." The four CNNM isoforms share 62.8% nucleotide sequence identity and a 65.5% residue identity [23]. They also share a high homology (>50%) with bacterial Co^{2+}/Mg^{2+} efflux protein CorC [23].

While CNNM1-4 share homologous structures, they differ in function and tissue distribution. CNNM1 is mainly expressed in neural tissue and in spermatogonial stem cells in the testes [25]. It has been shown to regulate cell cycle and differentiation in spermatogonial cells, and to act as a cytosolic copper chaperone [25-26]. CNNM2-4 are more ubiquitously expressed in all cell types [23]. The biochemical function of CNNM proteins is actively debated. They are believed to either transport Mg^{2+} directly across the plasma membrane, or regulate Mg^{2+} transport through a known transporter protein [27]. CNNM2 and 4 are basolateral proteins that have been associated with magnesium efflux, and their expression in the epithelia of renal nephrons and intestinal mucosal barrier has been tied to Mg^{2+} reabsorption [8-9, 27].

CNNMs are clinically relevant proteins. In mammals, point mutations in CNNM1 are associated with the urofacial syndrome, an extremely rare inherited disorder characterized by an unusual facial expression and disorder of the urinary tract (uropathy) [28]. Mutations in CNNM2 are linked with familial dominant hypomagnesemia accompanied by low Mg^{2+} serum level and symptoms such as muscle weakness, tremor, and headaches [10]. Mutations in CNNM4 have been associated with the Jalili syndrome, characterized by a combination of autosomal recessive cone-rod dystrophy and amylogenesis imperfecta (disorder of teeth development) [29]. In 2014, the CNNMs were shown to interact physically with members of the PRL (Phosphatases of

Regenerating Liver) family to regulate Mg^{2+} transport. This interaction was shown to promote tumor formation and invasiveness in animal and cellular models by raising the intracellular magnesium levels [6, 7].

CNNMs are not limited to metazoa or eukaryotes. In fact, there are amazingly over 70,000 CNNM protein sequences known found in all kingdoms of life [30]. In *E. coli*, there are two CNNM genes of unknown function, YtfL and YfjD, and related protein CorC, that mediate sensitivity to cobalt. [31-32] Budding yeasts have a single CNNM gene, *MAM3*, which confers sensitivity to high levels of Mn^{2+}, Co^{2+}, and Zn^{2+} [33].

1.5 CNNM structure

Structurally, CNNM proteins are defined by a conserved a transmembrane domain (domain of unknown function 21 [DUF21]), followed by a cytosolic CBS-pair domain [24]. Eukaryotic CNNMs contain an additional extracellular domain at the N-terminus, and a cyclic nucleotide-binding homology (CNBH) domain at the cytosolic C-terminus [24]. In prokaryotes, the extracellular domain is absent and the C-terminal CNBH domain is replaced by a smaller Corc/HlyC-like domain (Figure 1) [30].

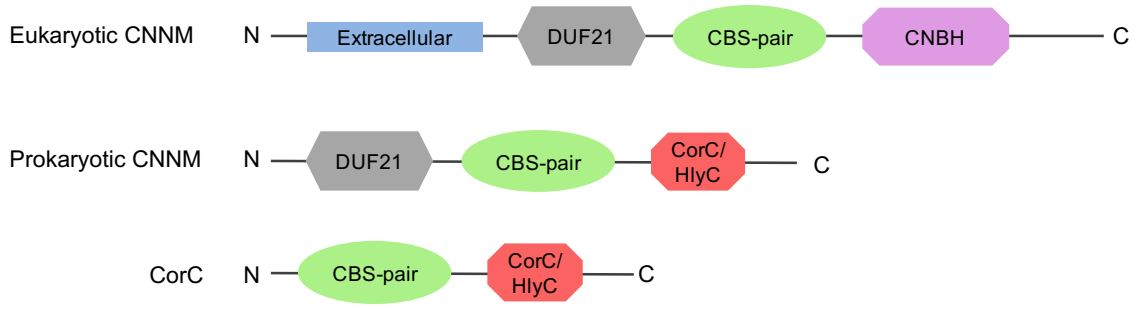

Figure 1. Domain structure of CNNM proteins All CNNM proteins contain a DUF21 transmembrane domain and a cytosolic CBS-pair domain. CorC Mg^{2+}/Co^{2+} efflux protein, a soluble bacterial CNNM ortholog, is shown for comparison.

1.5.1 CBS-pair domain

The cytosolic portion of the CNNM contains two cystathionine β-synthase (CBS) domains joined together in tandem. This domain was first discovered in the cystathionine β-synthase enzyme that protects against homocystinuria in humans by Alexander Bateman in 1997 [34]. Since then, CBS pair domains have been identified as stand-alone proteins or linked to thousands of cytosolic and transmembrane proteins. The CBS pair domain is highly conserved from bacteria to humans, and display a high level of structural conservation across species [23].

Structurally, the CBS motif is a small 60-residue element that is composed of a β_1-α_1-β_2-β_3- α_2 fold, usually with an extra α_0-helix preceding the fold [34]. Short, additional α-helical fragments within the canonical motif can also be present, such as in the case of CNNMs [34, 35]. This motif is almost always found in tandem repeats on the polypeptide chain, where they associate to form a CBS-pair domain (also called a Bateman domain). There are very few examples of CBS motifs found as a stand-alone motifs. Additionally, CBS-pair domains generally homodimerize by hydrophobic interactions of the α-helices on complementary CBS motifs to form a disc or ring-like structure composed of four CBS motifs (CBS module) (See Figure 2) [35].

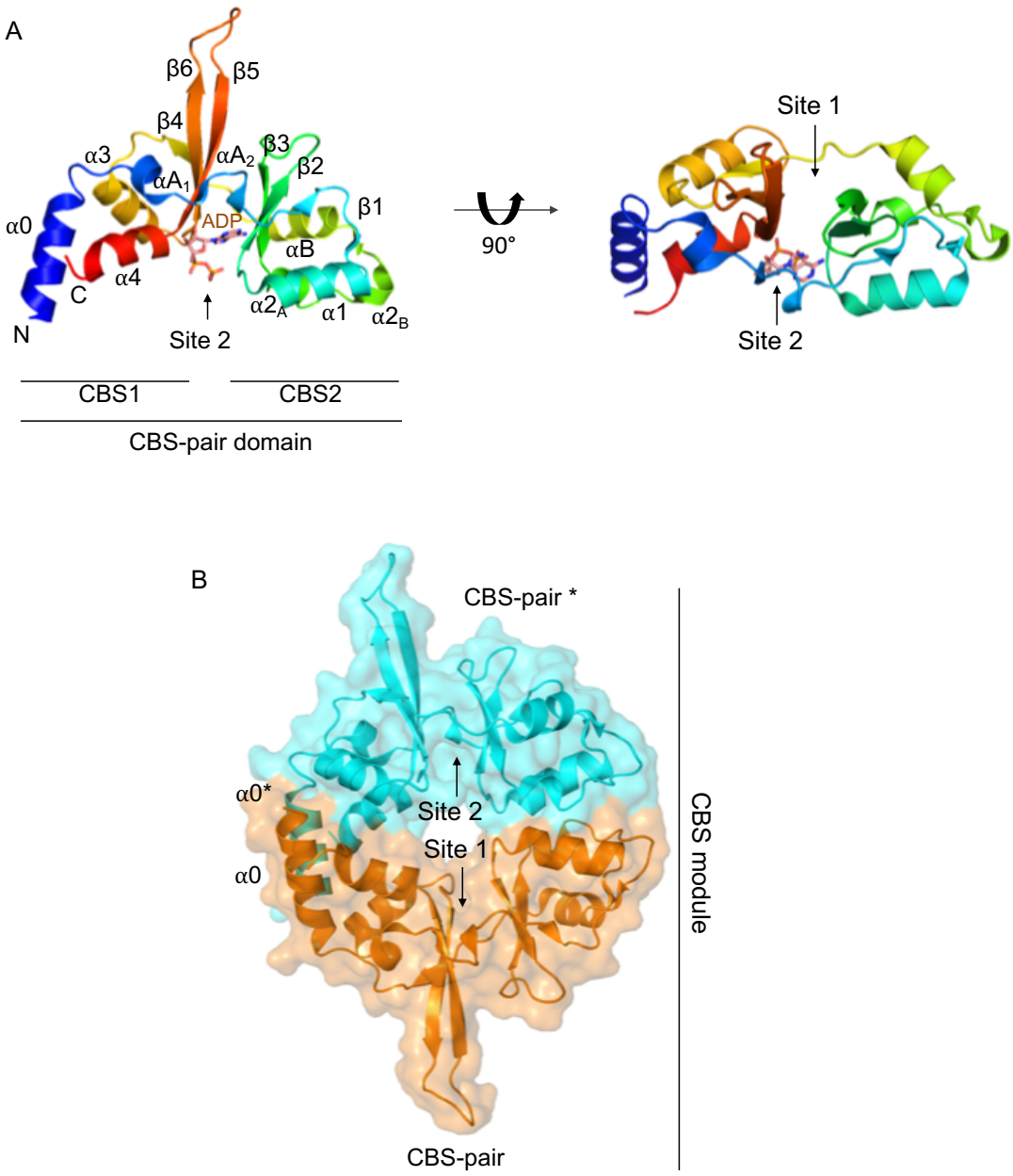

Figure 2. Structure of the CBS-pair domain Crystal structure of the CBS-pair domain of CNNM2 (PDB code 4IY0) (A) Two views of the structural elements configuring the CBS-pair domain of ADP-bound CNNM2. The structure consists of tandem repeats of the canonical β1-α1-β2-β3-α2 fold with the extra α0-helix. Additional short α-helical (αA_1, αA_2, αB) fragments are also present. (B) Dimerization of two CBS-pair domains forms a disk-like CBS module. Site 1 and Site 2 are indicated lie on opposite sides of the CBS-pair domain. Site 2 hosts adenosine ligands (e.g. ADP) in CNNM.

CBS pair domains are found in proteins with diverse functions. They are believed to act as sensors of intracellular energy status, sensing intracellular conditions through binding of

adenosine ligands (e.g.: AMP, ADP, ATP, NADH, NAD, S-adenosyl methionine) and metal ions (e.g.: Mg^{2+} in bacterial transporter MgtE) at their interface [35]. Potential ligand binding takes place at one of two cavities (Site 1 (S1) and Site 2 (S2)) at the dimer interface, formed by the walls of the β-sheets. Many studies describe the effect of point mutations in these domains and their implication in hereditary diseases, making them and their host proteins attractive targets in therapeutic research [35].

1.5.1.1 CBS-pair domain in CNNMs

The CBS-pair domain is the most extensively characterized domain in CNNMs. Its identification in the cytoplasmic region of CNNMs immediately suggested regulation of CNNM function by adenosine derivatives.

The first experimental evidence of nucleotide binding was provided by Hirata *et al*, who used surface plasmon resonance (SPR) techniques and filter-binding assays to detect binding of ATP in human CNNMs [27]. In the presence of Mg^{2+}, all of these proteins, except for CNNM3, showed saturable binding with ATP. K_d values were 43.4 ± 8.9 μM for CNNM4, 159 ± 28 μM for CNNM2, and 915 ± 389 μM for CNNM1. Interestingly, these K_d values inversely correlate with the CNNMs' capacities for Mg^{2+} efflux; CNNM4 showing the highest activity followed by CNNM2 and then CNNM1, and CNNM3 showing almost no efflux activity at all. In the absence of Mg^{2+}, no binding was observed from all of the proteins, suggesting that nucleotide interactions are Mg^{2+}-dependent [27]. Published structures of the CBS-pair domain of CNNM2 show that the domain can host ADP, AMP and ATP-Mg^{2+} at the S2 cavity (Figure 2A); Mg^{2+} favoring ATP binding by alleviating the otherwise negative charge repulsion existing between acidic residues on the protein and the polyphosphate groups of ATP [36]. The CBS-domain is crucial for CNNM function. The T568I mutation in hereditary magnesemia specifically targets ligand binding in CNNM2, abolishing Mg^{2+} efflux [27]. Additionally, deletion of the CBS-pair domain in CNNM2 and CNNM4 was also found to abrogate Mg^{2+} transport [27].

1.5.1.2 CBS-pair domain binds oncogenic PRL phosphatases

In 2014 and 2015, independent studies by Tremblay *et al* and Miki *et al* reported that a major mechanism of oncogenesis is mediated by the interaction of CNNMs with the

phosphotyrosine phosphatase subfamily PRLs [6, 7]. PRLs (phosphatase of regenerating liver) are highly oncogenic phosphotyrosine phosphatases (PTPs) that are highly over-expressed in solid tumors and haematological cancers, thus their association with and impact on CNNM function is of high biological importance [37]. The subfamily consists of three members PRL1-3 [38]. All three are highly conserved amongst themselves and are capable of binding to all four CNNM isoforms with nanomolar affinity [37]. Their dephosphorylation mechanism is poorly understood in part because they have low levels of enzymatic activity and because they have no define biochemical substrates. PRLs have a conserved cysteine residue in their catalytic domain that forms a phosphocysteine intermediate to mediate dephosphorylation [37].

The mechanism by which PRLs regulate CNNMs has yet to be elucidated, but current knowledge dictates that their association raises the intracellular Mg^{2+} levels either by promoting Mg^{2+} influx or by inhibiting Mg^{2+} efflux. Tremblay *et al* have shown that the interaction between CNNM3 and PRL2 leads to Mg^{2+} influx, which is positively correlated to tumor growth in breast cancers [6]. Meanwhile, Miki *et al* have shown that PRL3 inhibits Mg^{2+} efflux in CNNM4, and that the retention of intracellular Mg^{2+} is linked to progression of malignant tumors in colorectal cancers [7]. Having two distinct mechanisms on CNNMs is paradoxical given that PRL2-3 and the associated CNNM domains have highly conserved structures. Arjona and de Baaij cite this oddity to support their view that CNNMs are not Mg^{2+} channels bur rather Mg^{2+} channel regulators [39]. Regardless, there is agreement that PRLs increase intracellular Mg^{2+} levels and disrupt the energy metabolism, signalling, and cell cycle pathways in which magnesium participates [6, 7]. This leads to uncontrolled growth cell rates and abnormal metabolism, two hallmarks of cancer [40].

The CBS-pair domain is the site of interaction for PRLs [6, 7]. The basis of this interaction takes place at an extended loop with no defined secondary structure between the two CBS folds (Figure 3A) [6]. X-ray crystallography studies reveal that a highly conserved aspartate residue, located at the tip of the extended loop of the CBS2 fold, invades (and inhibits) the catalytic cavity of PRL, acting as a pseudosubstrate that is crucial for the CNNM·PRL interaction [41]. The extended loop is unique to the CBS-pair domains of CNNMs, and is absent in other CBS-containing proteins such as the AMP kinase (Figure 3B) [6]. Interestingly, phylogenetic

analysis reveals that the loop exists only in organisms having a PRL ortholog, suggesting co-evolution between both proteins [6].

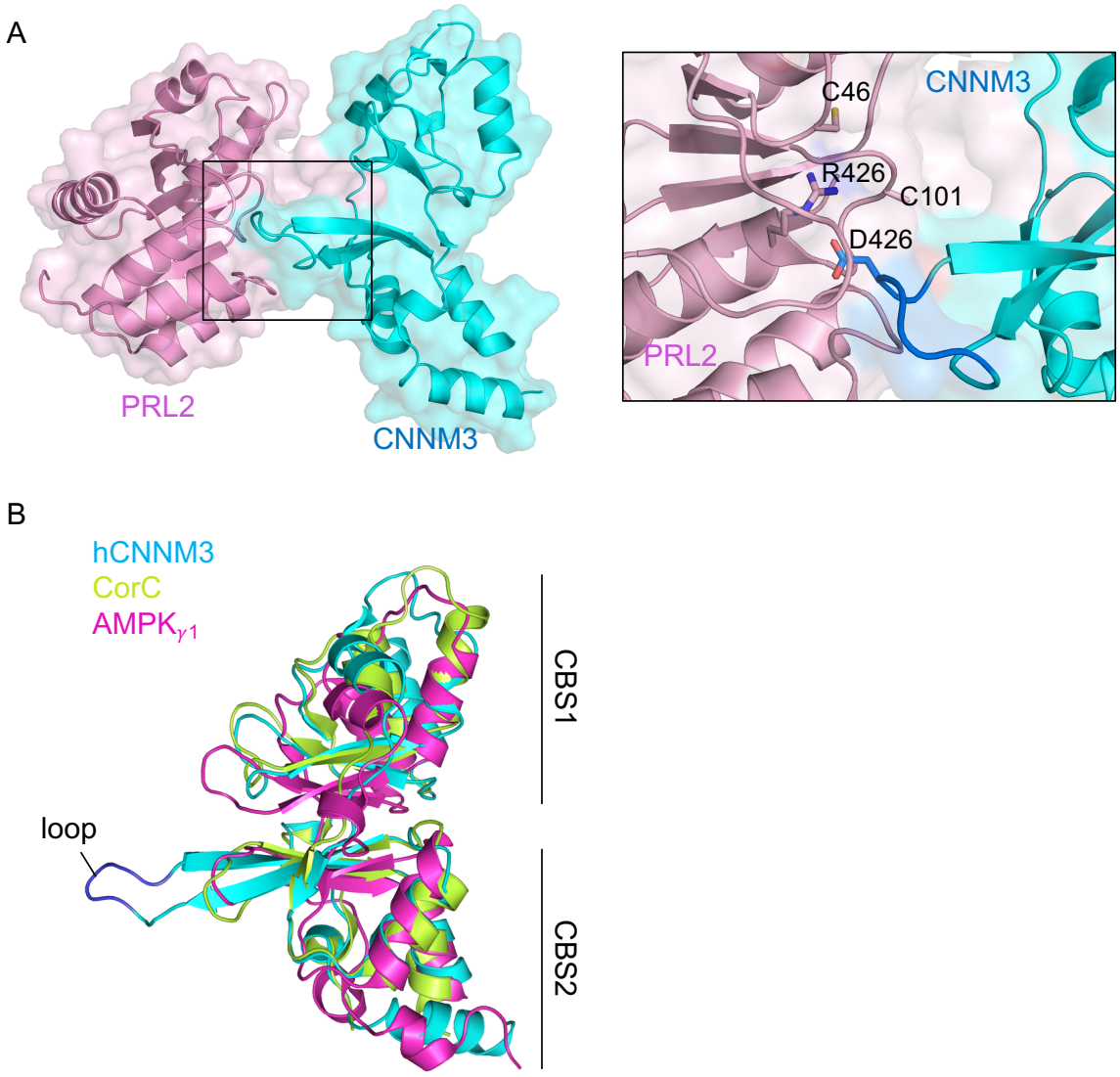

Figure 3. PRL phosphatases bind the CBS-pair domain via an extended loop (A) Crystal structure of CBS-pair domain of human CNNM3 bound to human PRL2 (PDB code 5K22). Close up view of the interaction site shows CBS2's extended loop in PRL's catalytic cavity. Conserved aspartate residue in CNNM3 and residues making up PRL's catalytic site are indicated (B) Overlay of the CBS-pair domains of human CNNM3 (blue), bacterial CNNM ortholog CorC (green), and $AMPK_{\gamma 1}$ (magenta) (PDB code 4HG0 and 2UV4). Extended loop located in human CNNM's CBS2 motif is absent in CorC and $AMPK_{\gamma 1}$.

1.5.1.3 Ligand-Induced Conformational Changes in CBS-pair domain

X-ray crystallography studies conducted by Martinez-Cruz *et al* on the CBS-pair domain of CNNM2 reveal significant conformational changes upon ligand binding.

As mentioned previously, the CBS-pair domain of CNNMs binds ATP in a Mg^{2+}-dependent manner in the S2 cavity [27, 36]. In the absence of bound nucleotide, the two CBS motifs maintain a relative orientation that is determined through a mixed network of hydrophobic, hydrogen bond (e.g. centered on T568 in CNNM2), and salt-link interactions between the residues that form the walls of the S1 and S2 cavities. These interactions cause a contortion in the CBS-pair domain, resulting in a Y-shaped "twisted" conformation in the dimer CBS module [36]. Binding of ATP- Mg^{2+} in the S2 cavity disrupts all of these interactions and establishes new ones with residues from both CBS1 and CBS2 motifs at opposite walls of the cavity. This triggers the shift of several secondary elements, among which is the α0 helix connecting the CBS-pair domain to the transmembrane domain, causing the dimer to adopt a disk-like "flat" conformation (Figure 4B) [36]. The addition of ADP and AMP to the CBS-pair domain of CNNM2 also showed a similar shift towards the "flat" conformation. Interestingly, the $CNNM2_{T568I}$ mutation in hypomagnesemia, which is unable to bind nucleotides, mimics the structural effect of nucleotide binding and irreversibly stabilizes the "flat" conformer (Figure 4C) [36]. These conformers mimic the structures of the CBS-pair domains of CNNM's bacterial ortholog CorC and bacterial Mg^{2+} transporter MgtE, which cycle between "open" and "closed" conformations depending on their ligand-bound state (AMP for CorC; Mg^{2+} and ATP for MgtE) [42-45]. Furthermore, in MgtE, this conformational change is transmitted to the transmembrane region of the transporter through the C-terminal α-helices connecting both CBS-pair domains with the transmembrane domains, reminiscent of the N-terminal α0 helix in CNNM2, thus regulating the Mg^{2+} intake (Figure 4A) [44-45].

After solving the crystal structures of the CBS-pair domain of CNNM2 in complex with PRL1, Martinez-Cruz *et al* reported a similar flattening of the CBS module, even in the absence of nucleotides [46]. In this case, the change in the relative orientation of the CBS motifs appears to be triggered by the electrostatic attraction between the surfaces of the two interacting proteins

(Figure 4D). This flattening is consistent in published structures of the CBS-pair domain of CNNM3 in complex with PRL2, and with PRL3 [41].

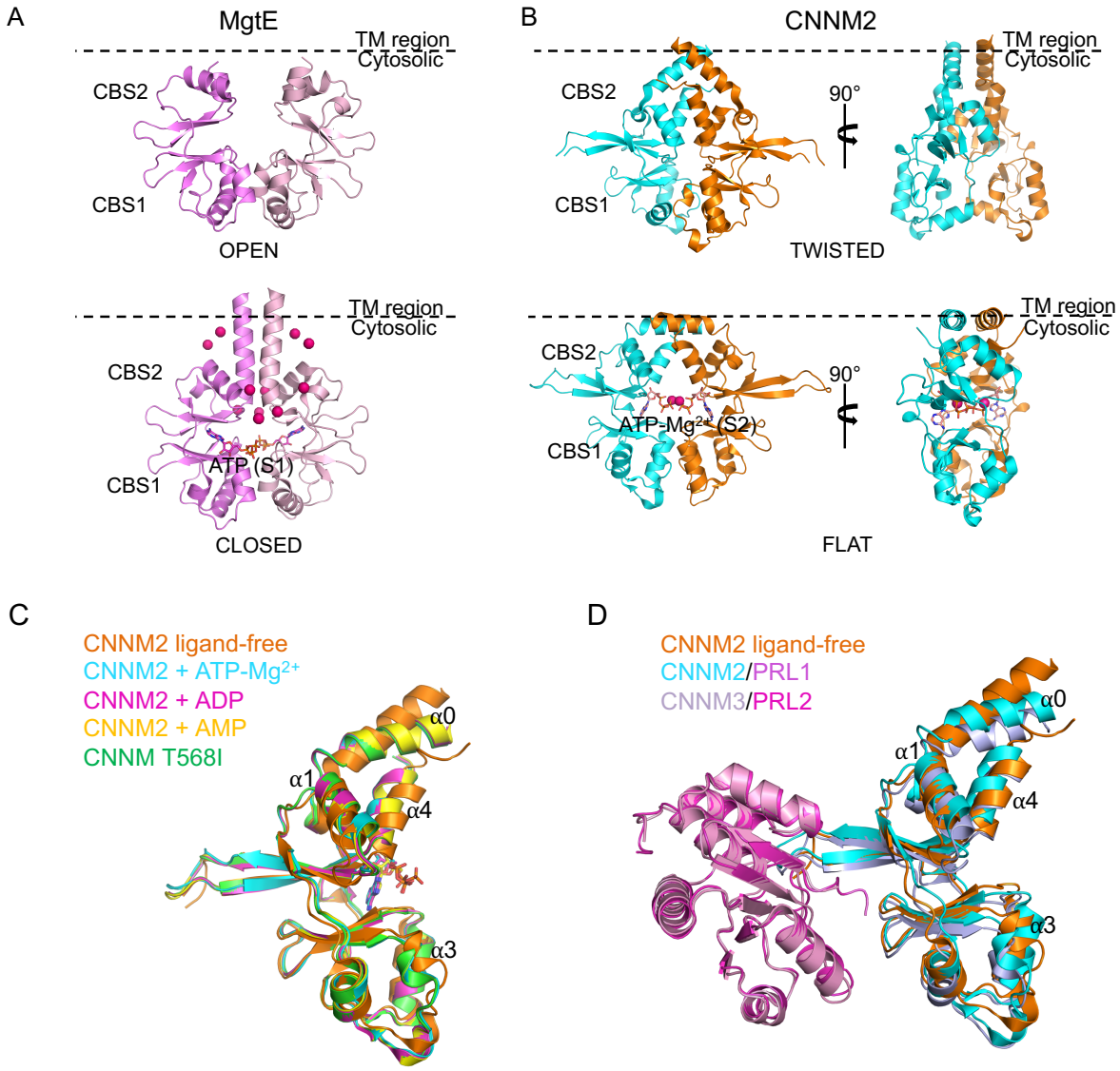

Figure 4. Ligand binding induces conformational change in the CBS-pair domain (A-B) Crystal structures of the CBS-pair domains of bacterial Mg^{2+} transporter MgtE (left) and human CNNM2 (right) ligand-free and bound to Mg^{2+} (hot pink) and ATP (PDB code 5X9H and 2YVZ). Ligand-free structures of MgtE and CNNM2 show an "open" and "twisted" dimer conformation, whereas the ligand-bound structures show a "closed" and "flat" dimer conformer. Cytosolic and transmembrane (TM) limits are defined by a broken line. (C) Overlay of the crystal structures of the CBS-pair domain of human and murine CNNM2 in the presence and absence of nucleotide (and Mg^{2+}) ligands (PDB code 4IYS, 4P1G, 4IY0, 4P1O and 4IY4) (D) Overlay of the crystal structures of the CBS-pair domain of human CNNM2 ligand free and in complex with PRL1, and of human CNNM3 in complex with PRL2 (PDB code 5MMZ and 5K22). The affected secondary structural elements (α-helices) are indicated.

Some authors postulate that, if sustained over time, the "flat" conformation of CBS-pair domain deactivates the Mg^{2+} extrusion capacity of the corresponding transporter [47]. In agreement with this hypothesis, the $CNNM2_{T568I}$ mutation, which irreversibly "locks" the CBS-pair pair domain in a "flat" conformation, abolishes basolateral Mg^{2+} efflux in hepatocytes, thereby exerting its pathological effect on humans [36]. All of these findings support the regulatory role of the CBS-pair domain in CNNMs.

1.5.2 CNBH domain

Unlike the extensively characterized CBS-pair domain, very little is known about the CNBH domain and its implications in CNNM function. When the human *CNNM* gene family was first cloned by Wang *et al*, this domain was named cyclic-nucleotide (cNMP) binding domain due to the sequence similarities it shared with the cNMP binding domains found in various ion channels (e.g. cAMP-regulated K^+ channel MloK1) and cNMP-dependent kinases (e.g. protein kinase A) [23, 47-48]. Because of this, it was speculated that the CNBH domain regulates CNNM function through cNMP binding. Otherwise, the CNBH domain does not interact with the CBS-pair domain, suggesting that these two intracellular domains act as independent modules, nor does it interact with PRLs [46].

1.5.2.1 CorC/HlyC domain

As mentioned previously, instead of a CNBH domain, prokaryotic CNNMs have a shorter CorC/HlyC domain at their C-terminus [30]. This domain is found at the C-terminus of CorC, a soluble bacterial CNNM ortholog that is involved in Mg^{2+} and Co^{2+} efflux [32]. The structure consists of a five-stranded antiparallel β-strands stacked with two α-helices (Figure 5).

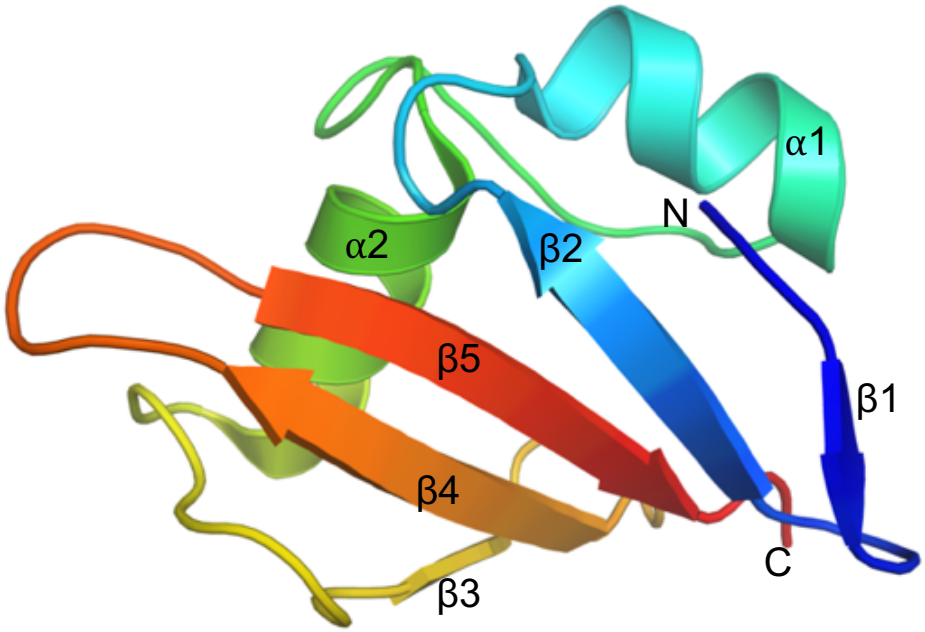

Figure 5. Crystal structure of CorC/HlyC domain Structure of the CorC/HlyC domain of CorC Mg^{2+}/Co^{2+} efflux protein from *E. Coli* (PDB code 4HG0), featuring all of the secondary structural elements that make it up.

1.5.3 DUF21 domain

DUF21 (Domain of Unknown Function 21) is a membrane spanning domain, of which almost nothing is known apart from its amino acid composition. Secondary structure prediction softwares indicate that the DUF21 domain consists of four transmembrane helices [24]. This was initially deemed problematic, as an even number of transmembrane helices would imply that the amino- and carboxy- termini of CNNM lie on the same side of the membrane. However, previous knowledge on CNNM's structure indicates that the amino- and carboxy- termini are respectively extracellular and intracellular [24]. To remedy this, de Baaij *et al* suggested a structure comparable to that of the glutamate receptor (GluR) ion channels, consisting of three full membrane-spanning helices and one re-entrant helix, which goes part way through the membrane and exits through the side it originates [24]. *In silico* analysis of the CNNM2 protein sequence reveals that the second transmembrane helix is the shortest and least hydrophobic of the four, suggesting that it may be re-entrant in the phospholipid bilayer (Figure 6) [24].

The DUF21 domains of CNNM has been speculated to be associated with Mg^{2+} transport. However, given the presence of only four transmembrane helices, it is not possible for the CNNMs to form a Mg^{2+} ion channel in their monomeric form. This discrepancy has been cited by Arjona and de Baaij to further support their theory that the CNNMs are not Mg^{2+} transporters, but rather function as Mg^{2+} sensors and regulate a known Mg^{2+} transporter [39] Because of the CBS-pair domain's propensity to dimerize and its relevant regulatory role in its host proteins, it is theorized that the DUF21 domain also dimerizes *in vivo* to translocate ions [35]. In fact, several known Mg^{2+} transporters are shown to form homomeric quaternary structures *in vivo* because they lack sufficient transmembrane helices in their monomeric state to form functional ion channels. Bacterial Mg^{2+} transporter MgtE, with five transmembrane helices, dimerizes [43-45]; CorA, with only 2 helices, forms a pentamer [49]; and TRPM7, with six helices, forms a tetramer [21]. In spite of this, it is also just as likely that DUF21 has no transportation capacity and that it simply anchors its host proteins to the plasma membrane.

The DUF21 domain of CNNMs holds clinical relevance, as point mutations in the transmembrane domain of CNNM4 are associated with the Jalili syndrome [29].

1.5.4 The extracellular domain

A bioinformatics/sequence analysis predicts that the N-terminus of eukaryotic CNNMs originally consists of five transmembrane α-helices upon translation, one of which is cleaved off by the signal peptidase complex during processing at the endoplasmic reticulum (Figure 6) [24].

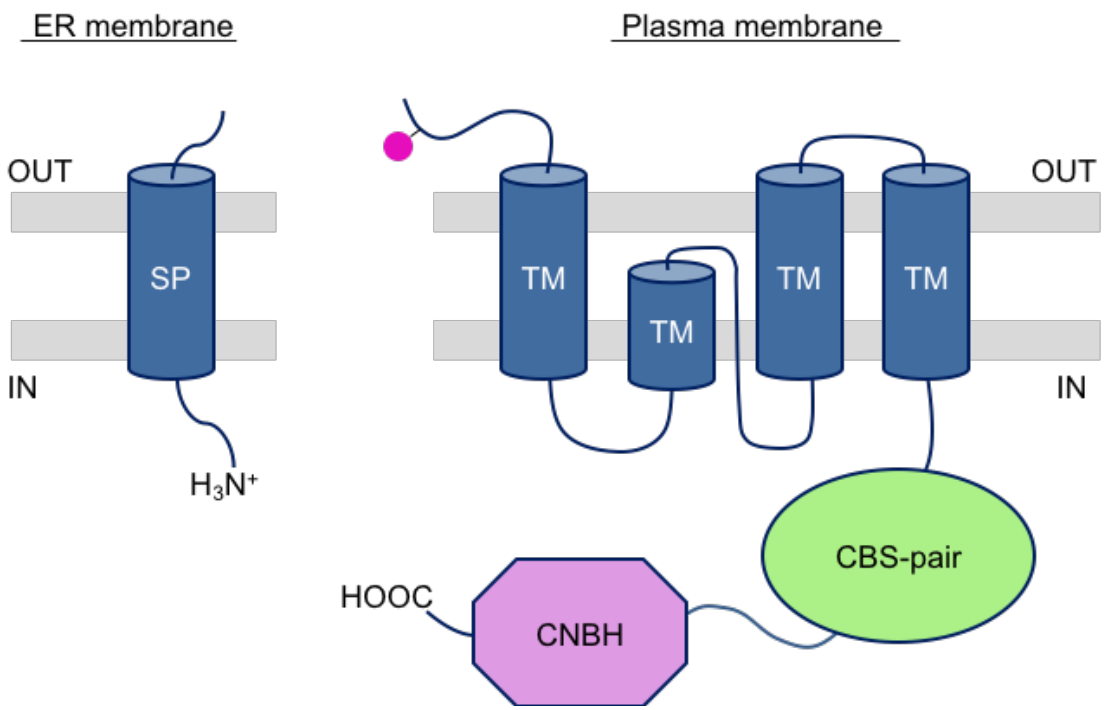

Figure 6. Topology of CNNMs Schematic of the post-translational processing of CNNM proteins from the ER to the plasma membrane after removal of the signal peptide (SP). Asparagine glycosylation is indicated by a pink dot. Transmembrane helices are labelled as TM, and cytosolic domains are labelled as CBS-pair and CNBH (cyclic nucleotide binding homology). This schematic of CNNM's membrane domain demonstrates that the second transmembrane helix is re-entrant from the cytosolic side, allowing the localization of the C-terminal domains in the cytosol.

The remaining external stretch of amino acids is glycosylated at a conserved asparagine residue (N112, N73, and N85 in CNNM2, CNNM3, and CNNM4, respectively), which is indispensable for the trafficking and stability of the host protein to the plasma membrane [24]. It is also predicted to be a β-strand enriched zone [24].

1.6 Project goals

CNNM proteins are an important and fundamental part of biology yet despite their biological relevance and ubiquity, no full three-dimensional structure is available. Their function also remains to be elucidated. Biochemically, they have been implicated in Mg^{2+} transport from bacteria to animals, but it is unclear whether CNNMs function directly as transporters or whether they regulate other transporters [39]. Our project aims to study the structures of these proteins through a step-by-step approach, starting with individual domains and leading up the full-length CNNM proteins, in order to better understand their biochemical function and regulation.

I. Our first goal is to characterize the previously uncharacterized CNBH domain of CNNMs by solving its structure with X-ray crystallography. We also study its oligomerization state using sedimentation velocity analytical ultracentrifugation, and investigate its predicted ability to bind cyclic nucleotides using differential scanning fluorometry and spectroscopic techniques (NMR).

II. After characterizing the CNBH domain on its own, we determine how it behaves in conjunction with the CBS-pair domain by solving the structure of the full cytosolic fragment with X-ray crystallography. We also investigate the link between ligand binding and dimerization in the soluble domains using sedimentation velocity analytical ultracentrifugation.

III. Finally, we endeavor to arrive at a full structure of CNNM proteins. To do so, we purify and crystallize a thermophilic eubacterial CNNM and study its oligomerization state using a fluorescence microscopy subunit counting technique.

Investigating the structure of the individual domains as well as the full-length protein in a well-controlled environment using *in vitro* techniques will allow us to gain more insight on the regulation and function of the CNNMs *in vivo*. It will also allow us to better understand their complex biological involvement in Mg^{2+} homeostasis, in diseases wherein ion transport is impaired, and in cancer.

CHAPTER 2. MATERIALS AND METHODS

2.1 Cloning of DNA constructs

All constructs were designed using UniProtKB database and the JPred secondary structure prediction server, were codon-optimized for *Escherichia coli* (Bio Basic Inc., Canada), and were verified by Sanger sequencing in the Genome Quebec sequencing facility. Human CNNM1 CNBH domain (residues 569-798), CNNM2 CNBH domain (residues 585-824), CNNM3 CNBH domain (residues 453-707), CNNM4 CNBH domain (residues 513-728) were sub-cloned into into BamHI and NotI sites of pGEX-6P-1 vector (Amersham-Pharmacia) with an N-terminal GST-tag. CBS-pair domains of human CNNM1 (residues 412-568), CNNM2 (residues 429-584), CNNM3 (residues 299-452), and CNNM4 (residues 356-511) were sub-cloned into BamHI and NotI sites of pSMT3 vector (Mossessova and Lima, 2000) with an N-terminal His6-SUMO-tag. Human CNNM2's cytosolic fragment (residues 429-817 Δ724-767) was subcloned into NdeI and XhoI sites of pET29a vector (Millipore Sigma) with a C-terminal His6-tag. Human CNNM3's cytosolic fragment (residues 299-658) was sub-cloned into BamHI and NotI sites of pGEX-6P-1 vector. Human CNNM4's cytosolic fragment (residues 356-726 Δ649-670) was sub-cloned into BamHI and NotI sites of pSMT3 vector with an N-terminal His6-SUMO-tag. CNNM constructs from thermophilic bacteria *Tediphilus thermophilus* (tCNNM) (residues 2-445), (residues 2-423), and (residues 2-332), and cytosolic fragment (residues 208-423) were sub-cloned into NdeI and XhoI sites of pET29a vector. Finally, cloning of human PRL2 construct (C95A, C96A, C119A, DCCVQ) was described previously [35]. Mutagenesis was performed using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent).

2.2 Expression and Purification

Cultures were grown at 37 °C in LB to an optical density of 0.8 and induced with 1 mM IPTG overnight at 20 °C. Cell pellet was obtained by centrifuging at 5000 g for 20 min at 4°C. The pellet was re-suspended in Buffer A (50 mM HEPES, 500 mM NaCl, 5% glycerol, 5 mM BME, pH 7.5) and lysed by sonication. Cellular debris was removed by centrifugation at 44000 g for 45 min at 4°C. For DUF21-containing constructs, the membrane fraction was obtained by ultracentrifuging the clarified lysate at 150000g for 1 hour at 4°C, manually homogenized and incubated in Buffer A containing 1% (w/v) n-dodecyl-β-maltoside (DDM) for 1 hour at 4°C in

order to solubilize the membrane proteins. Insoluble fragments were then pelleted by re-centrifuging at 150000g for 30 minutes.

For GST constructs, the supernatant was loaded onto Pierce Glutathione Agarose resin, washed with Buffer A and eluted with Buffer A containing 20 mM glutathione. The GST tag was removed by overnight incubation with PreScission Protease, leaving an N-terminal Gly-Pro-Leu-Gly-Ser extension. For 6His-containing constructs, the supernatant was loaded onto Qiagen Ni-NTA or TALON Co^{2+} resin, washed with buffer A containing 30 mM imidazole and eluted with buffer A containing 500 mM imidazole. For SUMO-containing constructs, the SUMO-tag was removed by overnight incubation with His-Ulp, leaving an N-terminal Ser extension. For DUF21-containing constructs, the wash and elution buffer were supplemented with 0.05% (w/v) DDM.

Excess nucleotides were removed from CNNM CBS-pair domain containing constructs by applying affinity-purified proteins onto an anion exchange column (BioSuite Q 13 mm AXC) equilibrated with buffer B (50 mM HEPES, 10 mM NaCl, 1 mM EDTA, pH 7.5) and eluted using a linear gradient of buffer C (50 mM HEPES, 1 M NaCl, 1 mM EDTA, pH 7.5). Proteins were then further purified by Superdex-200 (cytosolic and DUF21-containing fragments) or -75 (CNBH and CBS-pair domain) size-exclusion column (GE Healthcare) in HPLC buffer (20 mM HEPES, 100 mM NaCl, 3 mM TCEP, pH 7.5); supplemented with 0.05% (w/v) DDM for membrane proteins. The final purified protein was concentrated to around 10 mg/mL (measured by NanoDrop), and the purity verified by SDS-PAGE.

2.3 Crystallization

Crystallization trials were carried out using a sitting-drop format in 96-well Intelli-3 crystallization plates with a variety of commercially available screens (Classic Suites II from Qiagen, PEG I-II, ProComplex Suite, Nucleix, MBClassI-II, and MemGold I-II). Screening plates were set up using the NT8 robot and were incubated at a constant temperature of 20˚C. Drops consisted of 300 nl protein solution mixed with 300 nl precipitant solution and the reservoir volume was 50 µl. Improved crystals were subsequently obtained by refining the successful conditions using a hanging-drop format in 24-well VDX plates (Hampton Research)

with drops consisting of 1 µl protein and 1 µl precipitant equilibrated over a reservoir volume of 1 ml. tCNNM crystals were cryoprotected by soaking in mother liquor supplemented with 40% PEG200, picked up in a nylon loop and flash-cooled in a N_2 cold stream. Crystal diffraction was carried out using a D8 Venture X-ray diffractometer (Bruker).

2.4 Scanning differential fluorimetry assays

Each reaction contained 20 µL of solution with 25 µM CNBH domain from CNNM1-4, 1x Protein Thermal Shift TM dye (Life Technologies), HPLC buffer with and without nucleotides and $MgCl_2$. Samples were heated from 25°C to 99°C at a rate of 1°C per minute and fluorescence signals were monitored by StepOne Plus quantitative real-time PCR system (Life Technologies). Data were analyzed using Thermal Shift software (Life Technologies). The maximum change of fluorescence with respect to temperature was used to determine the melting temperature (T_m). Each sample was performed in triplicates and standard error was calculated for each Tm measured. Positive control protein (lpg1496-KLAMP1) was expressed and purified as described in [50].

2.5 2D NMR (^{1}H-^{15}N HSQC)

^{15}N-labeled human CNNM4 (513-728) was used for this purpose. Cells were grown at 37 °C in minimal (M9) media (with added 1 mg/L thiamine, 2 mM $MgSO_4$, 0.4% glucose, 0.8 mg/L $FeCl_3$, and 1 g/L $^{15}NH_4Cl$) to an optical density of 1.00, and induced with 1 mM IPTG overnight at 20°C. Proteins were purified as indicated above. ^{1}H-^{15}N heteronuclear single quantum correlation spectroscopy was performed on a Bruker 600 MHz spectrometer using 0.15 mM ^{15}N-labeled CNNM4 CNBH domain in 90% HPLC buffer and 10% D_2O. Testing for cNMP binding was carried by acquiring TROSY spectra before and after the addition of 3 mM cAMP. NMR data were acquired at 40 °C. NMR spectra were processed using NMRPipe and analyzed with SPARKY [51-52].

2.6 Analytical Ultracentrifugation (AUC)

All sedimentation velocity AUC experiments were performed at 20 °C using a Backman Coulter XL-I analytical ultracentrifuge with an An-60Ti rotor at 98,000g for 18 hours with scans performed every 60 seconds. A double-sector cell, equipped with a 12 mm Epon centerpiece and

sapphire windows, was loaded with 380 and 400 µL of sample and HPLC buffer, respectively. Experiments focused on protein concentration were monitored for absorbance with UV detection at 280 nm, whereas ligand-binding focused experiments were monitored with Rayleigh interference. The data were analyzed with Sedfit v1501b using a continuous c(s) distribution [53]. Numerical values for the solvent density, viscosity, and the partial specific volume were determined using Sednterp [54]. Buffer density and viscosity were calculated to be 1.0039 g/cm^3 and 0.01026 mPa•s, respectively (20 mM HEPES, 100 mM NaCl, 3 mM TCEP, pH 7.5). Partial specific volumes for CNBH domains of CNNM1-4 were calculated to be 0.7172 cm^3/g, 0.7419 cm^3/g, 0.7416 cm^3/g, and 0.7350 cm^3/g. Partial specific volumes for the cytosolic CNNM2-4 fragments, the CBS-pair domains of CNNM1-4, and for the CBS-pair domain of CNNM4 in complex with PRL2 were calculated to be 0.7472, 0.7408, 0.7327, 0.7435, 0.7448, 0.7396, 0.7414, and 0.7391 cm^3/g, respectively. The frictional ratios of human CNNM2-3 CNBH domains, and of human CNNM4 CBS-pair domain by itself and in complex with PRL2 were calculated using US-SOMO to be 1.29, 1.21, 1.25 and 1.33, respectively [55]. The frictional ratio (f/f_0) values for the CNBH domains of CNNM1 and CNNM4 was estimated to be 1.29 based on similarity to CNNM2, and that of the CBS-pair domains of human CNNM2-4 was estimated to be 1.25 based on similarity to CNNM4. Default f/f_0 values were used for the cytosolic fragments. c(s) distribution graphs were plotted using GUSSI [56].

2.7 Maleimide labelling of tCNNM

For mass spectrometry experiments, the cytosolic fragment was reacted with N-ethylmaleimide at room temperature for 2 hours in HPLC buffer, using a 10 :1 molar ratio of reagent to protein. The reaction was quenched with the addition of DTT at a 10-fold molar excess to the labelling reagent. Excess maleimide and reducing agent were then removed by applying the sample on a HiPrep™ 26/10 Desalting column in HPLC buffer. For subunit counting experiment, full length protein was reacted with Oregon Green 488 (OG488) maleimide at a 10-fold molar excess overnight at 4°C in HPLC buffer with 0.05% (w/v) DDM, in the dark. Excess dye was removed by loading the sample on a Superose® 6 10/300 GL column, packed with HisPur™ Cobalt Superflow resin, in HPLC buffer supplemented with 0.05% (w/v) DDM and 30 mM imidazole, and eluted using a linear gradient of HPLC bufer supplemented with 0.05% (w/v) DDM and 500 mM imidazole. A UV detector of 500 nm was used to monitor the levels of OG488 during the

loading and washing stages, and promptly turned off during the elution step. Labelling efficiency was verified by NanoDrop by measuring the absorbance intensities at 280 nm and 488 nm.

2.8 Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS)

N-ethylmaleimide labeled proteins were extracted with the chloroform/methanol/water method [57] and dissolved in matrix solution (50:50 H_2O and ACN with 0.1% TFA) to a final concentration of 100 pmol/µl. The sample was mixed with 10 mg/ml of sinapinic acid matrix at sample to matrix ratios of 1:10. 1ul of the final mix was applied to a MALDI target plate and allowed to air dry before running on a Microflex LT MALDI-TOF (time of flight) with FlexControl. The resulting spectra were analyzed with flexAnalysis [58].

2.9 Subunit counting of tCNNM

OG488-labeled proteins were attached to a coverslip surface coated with Ni-NTA resin, washed thoroughly with HPLC buffer supplemented with 0.05% DDM and 30 mM imidazole, and imaged using total internal reflection (TIRF) microscopy as described in [59]. Images were recorded with a backlit EMCCD camera (iXon+ 860BV, Andor Technology, South Windsor, CT) at a sampling rate of 20 Hz. Image analysis, step detection, and subunit counting were performed using PIF (Progressive Idealization and Filtering) software [59].

CHAPTER 3. RESULTS

3.1 Structural and functional characterization of the CNBH domain

Among the characterized soluble domains of CNNMs, very little is known about the CNBH domain of the CNNM proteins beyond its sequence homology with the cyclic nucleotide binding domains present in kinases and ion channels [23, 47-48]. As such, in order to better understand its role in CNNM's biochemical function, we endeavoured to characterize this domain first.

3.1.1 Structure of the CNBH domain

In order to gain insight into the CNBH domain of CNNMs, we sought out to determine its three-dimensional structure via X-ray crystallography.

Crystals for human CNNM3 CNBH domain were first obtained by Dr. Guennadi Kozlov, and diffracted to 3.0 Å. As there were no available homologous structures at the time, the structure was solved with single-wavelength anomalous dispersion (SAD) using crystals from selenomethionine-labelled proteins for experimental phasing [1]. The structure of the selenomethionine-labeled crystal was solved at 1.9 Å (Figure 7A). Soon after, crystals of CNNM2 (residues 585-822, Δ724-767) were obtained, and the structure was solved by molecular replacement by Seby Chen [1]. Both CNNM2 and CNNM3 CNBH crystals contained two molecules in their asymmetric unit. The resulting CNBH domain structures were also very similar, with an RMSD value of 0.6 Å for 121 Cα atoms (Figure 7B).

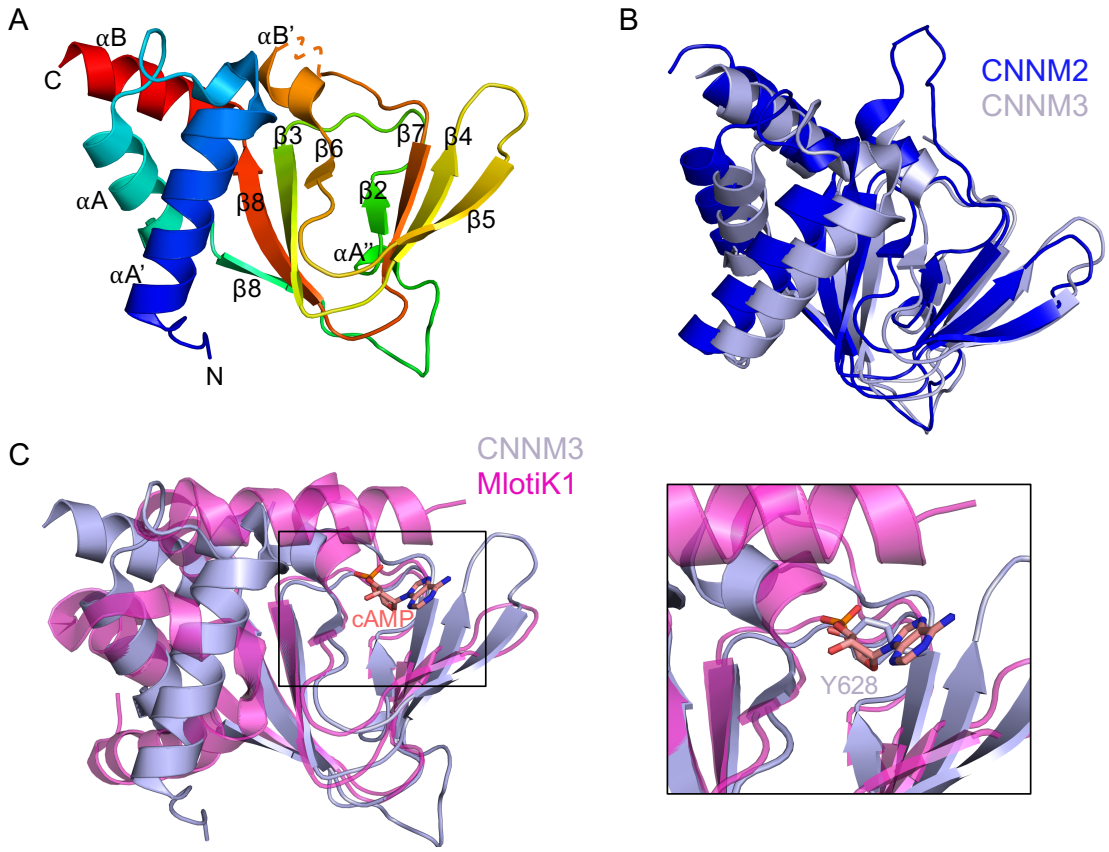

Figure 7. Structure of CNNM CNBH domain (A) Solved crystal structure of human CNNM3's CNBH domain. The structural elements making up the domain are indicated (B) Overlay of structures of human CNNM2 and CNNM3 (PDB code 6DFD and 6DJ3) (C) Overlay of human CNNM3 CNBH with cAMP-binding domain of cAMP-bound *Mesorhizobium loti* K^+ channel (PDB code 1VP6). Close-up of CNNM3 reveals a tyrosine residue (Tyr628) occupies the putative cAMP binding site.

The solved CNBH domain structures were similar to the structures of cyclic nucleotide-binding domains, containing an eight-stranded antiparallel β-roll that is capped by an α- helical bundle on the side [47]. The CNNM CNBH domain contains an additional single-turn α-helix (αA″) located between the first two β-strands (β1 and β2) (Figure 7A-B). For additional structural comparison, we used the cAMP-bound state for cyclic nucleotide-binding domain of the bacterial K^+ channel from *Mesorhizobium loti* (MlotiK1). Overall structural similarity of human CNNM3 CNBH domain to the MlotiK1 cAMP-binding domain is very high (Z-score 13.5 on DALI database), particularly in the β-roll region, which is the site of nucleotide binding [1]. While differences largely include slightly different orientations of helices and varied loop conformation, the main difference is the absence of an additional helix (αC) in CNNM3 (Figure

7C, left). Conventionally, in cyclic nucleotide-binding domains, this C-terminal helix closes on the bound nucleotide, providing additional contacts and increased affinity [60]. Furthermore, a structural alignment of the CNNM3 CNBH structure with the nucleotide-bound form of MlotiK1 reveals that CNNM3 contains a tyrosine residue (Y628) in the middle of the putative nucleotide-binding site that would sterically clash with a bound ligand (Figure 7C, right). This residue is an alanine in MlotiK1 but a tyrosine or phenylalanine in all four CNNM isoforms [1]. All of these observations suggest that the CNBH domains of CNNMs do not bind cyclic nucleotides.

3.1.2 The CNBH domain of CNNM does not bind cNMP

To test CNNM CNBH's ability to bind cyclic nucleotides (or lack thereof), we ran scanning differential fluorimetry experiments on the purified CNBH domains of CNNM1-4 in the presence and absence of cyclic nucleotides (cAMP and cGMP) and $MgCl_2$. The KLAMP1 domain from Legionella effector protein Ipg1496 (indicated as Ipg1496_KLAMP1), which binds cAMP, was used as a positive control [50]. The melting temperatures are recorded in Table I below.

Table I. Melting temperatures of CNBH domains of ion transporters

Protein	Condition	Tm (°C)				
		Replicate 1	Replicate 2	Replicate 3	Mean	Standard Error
CNNM1	No addition	48.20	48.02	48.11	48.11	0.05
	cAMP	48.15	48.01	48.13	48.10	0.05
	cAMP+ $MgCl_2$	48.16	48.15	48.10	48.14	0.02
	cGMP	48.17	48.04	48.01	48.08	0.05
	cGMP+ $MgCl_2$	48.16	48.18	48.29	48.21	0.04
CNNM2	No addition	50.23	49.94	50.10	50.09	0.09
	cAMP	50.12	50.14	50.19	50.15	0.02
	cAMP+ $MgCl_2$	50.32	50.18	50.07	50.19	0.07
	cGMP	50.34	50.21	50.13	50.23	0.06
	cGMP+ $MgCl_2$	50.39	50.30	50.32	50.34	0.03
CNNM3	No addition	49.49	49.48	49.43	49.47	0.02
	cAMP	49.43	49.38	49.43	49.41	0.02
	cAMP+ $MgCl_2$	49.35	49.69	49.32	49.46	0.12
	cGMP	49.54	49.32	49.43	49.43	0.07
	cGMP+ $MgCl_2$	49.38	49.35	49.67	49.47	0.10
CNNM4	No addition	52.46	52.53	52.36	52.45	0.05
	cAMP	52.56	52.50	52.56	52.54	0.02
	cAMP+ $MgCl_2$	52.50	52.39	52.40	52.43	0.03
	cGMP	52.65	52.64	52.53	52.61	0.04
	cGMP+ $MgCl_2$	52.58	52.64	52.63	52.62	0.02
Ipg1496_KLAMP1	No addition	74.08	74.01	74.05	74.05	0.02
	cAMP	74.67	74.67	74.82	74.72	0.05

Protein stability typically increases upon binding of a ligand, which results in a higher melting (denaturation) temperature [61]. To gauge the significance of the melting temperature variation, a statistical two-tailed T-test was performed for the treated CNBH domains, with α-

value set at 0.05 [62]. Because four CNNM isoforms are being compared, the Bonferroni correction is applied for their p-values. [63].

Table II. Two-tailed T-test on melting temperatures of CNBH domains

Protein	Condition	P-value	Corrected P-value
CNNM1	No addition	-	-
	cAMP	0.854	0.999
	cAMP+ $MgCl_2$	0.654	0.999
	cGMP	0.635	0.999
	cGMP+ $MgCl_2$	0.203	0.412
CNNM2	No addition	-	-
	cAMP	0.526	0.999
	cAMP+ $MgCl_2$	0.418	0.999
	cGMP	0.259	0.999
	cGMP+ $MgCl_2$	0.0490	0.196
CNNM3	No addition	-	-
	cAMP	0.0993	0.397
	cAMP+ $MgCl_2$	0.917	0.999
	cGMP	0.609	0.999
	cGMP+ $MgCl_2$	0.999	0.999
CNNM4	No addition	-	-
	cAMP	0.166	0.664
	cAMP+ $MgCl_2$	0.758	0.999
	cGMP	0.0664	0.266
	cGMP+ $MgCl_2$	0.0341	0.136
Ipg1496_KLAMP1	No addition	-	-
	cAMP	0.000237	-

It can be observed that the KLAMP1 positive control, which binds cAMP, shows a small but significant (p value = 0. 000237) increase in stability [50]. None of the melting temperatures of the CNBH domains showed a significant change with cAMP or cGMP (±$MgCl_2$) addition after application of the Bonferroni correction for multiple comparisons (Tables I-II and Figure 8A).

Ligand binding in CNBH human CNNM4 was also tested using ^{1}H-^{15}N HSQC NMR, with spectra recorded in the presence and absence of cAMP (Figure 8B).

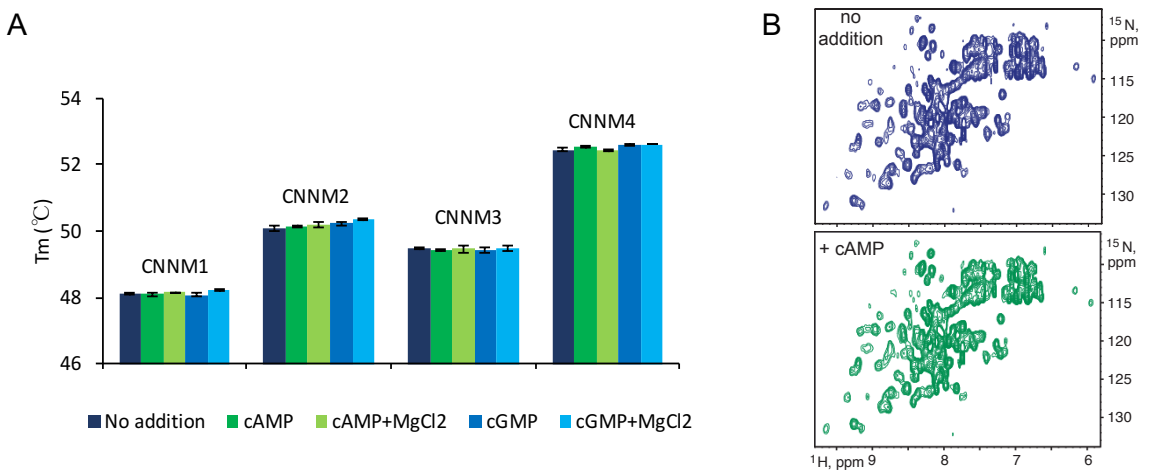

Figure 8. The CNBH domain of CNNMs does not bind cNMPs. (A) Melting temperatures of CNBH of human CNNM1-4 in the presence and absence of cyclic nucleotides and $MgCl_2$ measured by differential scanning fluorimetry. Standard errors are indicated by ticks. (B) ^{1}H-^{15}N NMR spectra for human CNNM4 with and without added 3 mM cAMP.

NMR is very sensitive to molecular interactions, and is capable of detecting of low-affinity interactions even in the millimolar range. The ^{1}H-^{15}N correlation spectrum of the ligand-free CNBH of CNNM4 shows good dispersion signals, characteristic of a well-folded protein [64]. However, no significant spectral changes were observed upon the addition of 3 mM of cAMP (Figure 8B).

The lack of shifts in the melting temperatures (scanning fluorimetry) and chemical signals (NMR), together with the absence of structural features required to bind cNMPs, indicate that the CNNM CNBH domain does not bind cyclic nucleotides.

3.1.3 The CNBH domain as a dimerization interface

The crystal structure of the CNBH domains of human CNNM2 and CNNM3 shows two molecules in the asymmetric unit. To verify whether dimerization occurs in solution, sedimentation velocity analytical centrifugation (SV-AUC) experiments were run with increasing concentrations of CNNM1-4 CNBH domains. The resulting distribution graphs are presented in Figure 9 and the values for the molecular weights recorded in Table III.

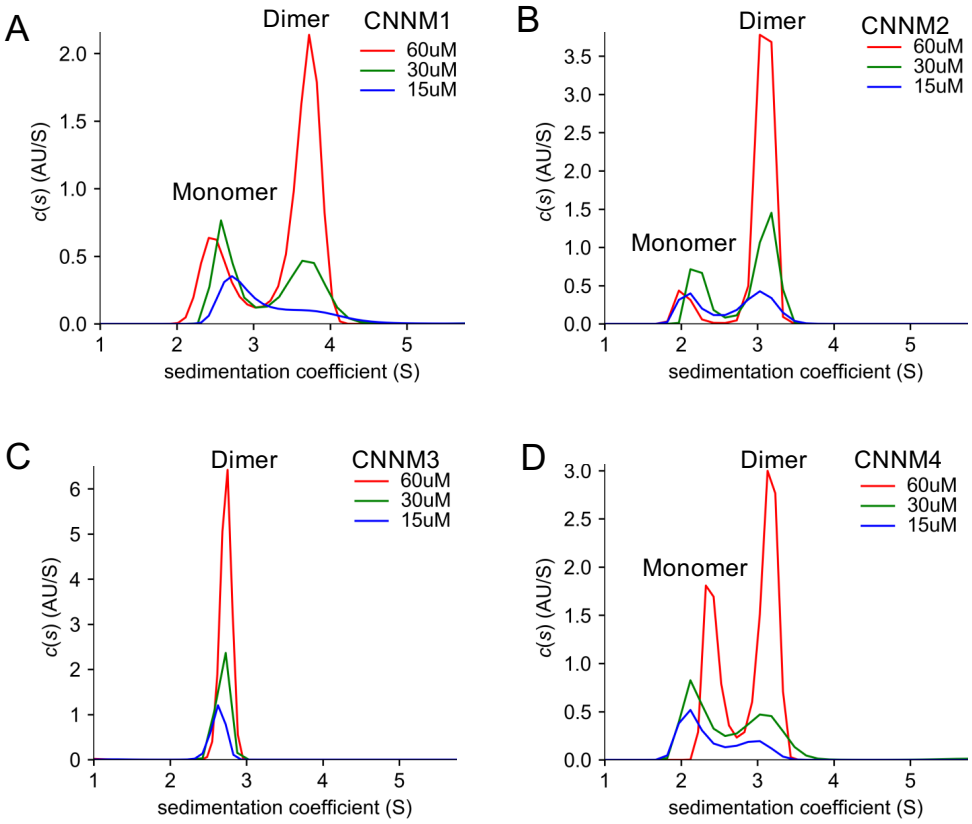

Figure 9. CNNM CNBH domain dimerizes in solution Sedimentation velocity distribution profiles for CNBH domain of (A) human CNNM1 (B) human CNNM2 (C) human CNNM3 and (D) human CNNM4 at varied concentrations. Monomer and dimer peaks are indicated.

Table III. SV-AUC sedimentation coefficients and estimated molecular weights for CNNM CNBH domain

		Sedimentation coefficient (S)		Estimated MW (kDa)		RMSD	Theoretical MW (kDa)	
CNBH domain	[Protein] (uM)	Peak 1	Peak 2	Peak 1	Peak 2		Monomer	Dimer
CNNM1	60	2.52	3.68	28.1	52.9	0.0051	27.0	54.0
	30	2.62	3.68	30.2	46.7	0.0049		
	15	2.75	3.70	31.1	43.1	0.0041		
CNNM2	60	2.65	3.10	22.3	40.7	0.0053	22.7	45.4
	30	2.19	3.11	24.4	40.9	0.0047		
	15	2.11	3.04	25.8	41.4	0.0028		
CNNM3	60	2.76	-	44.2	-	0.0050	23.9	47.8
	30	2.70	-	37.1	-	0.0031		
	15	2.67	-	35.8	-	0.0025		
CNNM4	60	2.38	3.15	25.5	37.3	0.0053	25.0	50.0
	30	2.25	3.11	25.5	40.2	0.0043		
	15	2.24	3.06	25.3	35.6	0.0021		

Analysis of the distribution plots and molecular weight values reveals that the CNBH domain CNNM3 forms a singular dimer species (Figure 9C), whereas CNNM1, CNNM2, and CNNM4 show a mixture of both monomer and dimer species that varied with protein concentration (Figure 9A, B and D). Thus, CNNM3 shows the highest propensity to dimerize, followed by CNNM1-2, and CNNM4 has the lowest dimerization tendency.

This was further investigated by examining the dimerization interface of CNBH (Figure 10 A-B).

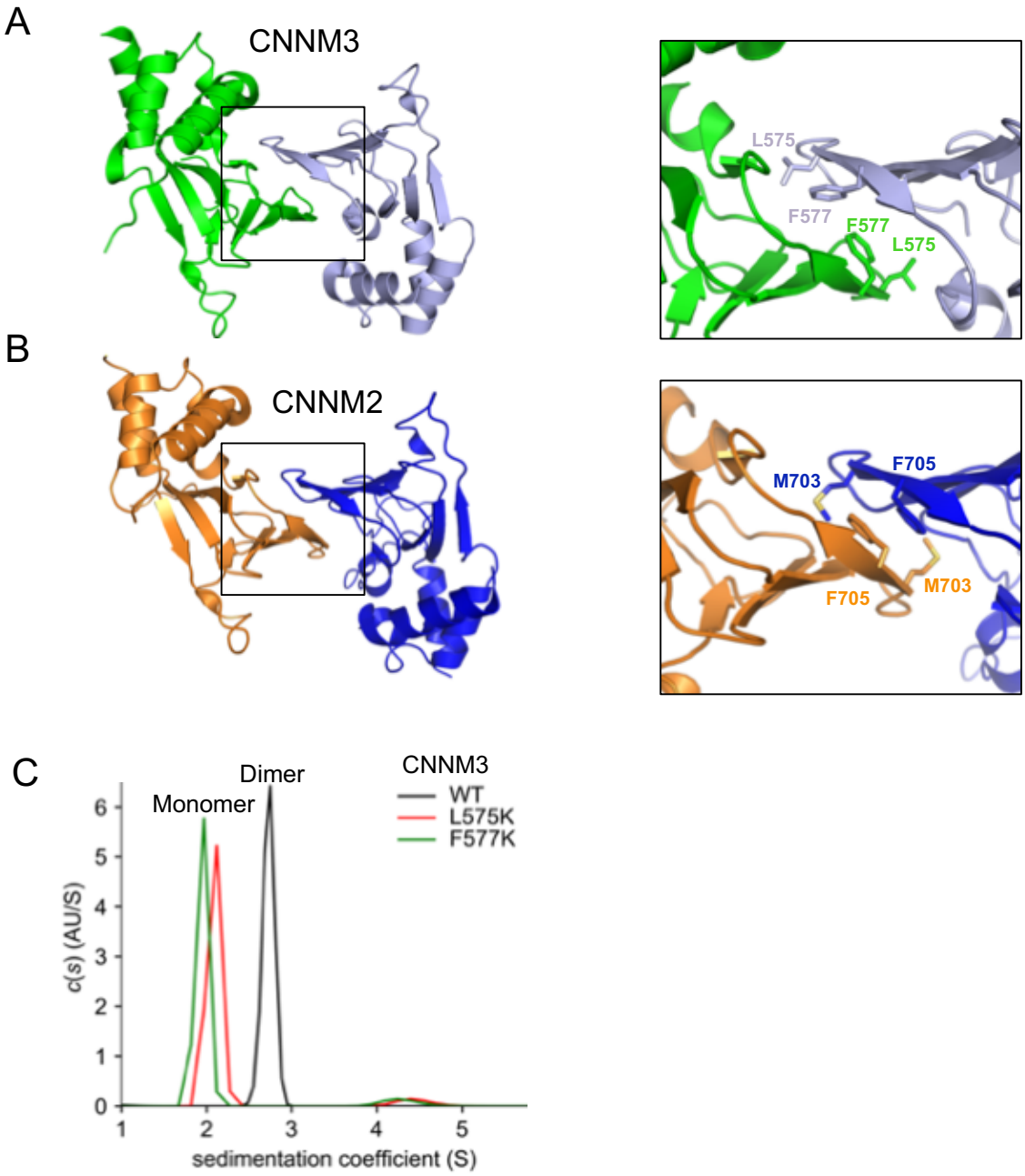

Figure 10. Dimerization interface of CNNM CNBH domain (A-B) Crystal structure of human CNNM2 and CNNM3 CNBH dimers, and close-up of dimerization interface. Residues involved in dimerization are indicated. (C) Sedimentation velocity distribution curves of CNBH domain of CNNM3 wildtype (WT) and mutants of the dimer interface. Monomer and Dimer peaks are indicated.

Crystals structures of CNNM2 and CNNM3 show that dimerization takes place at the conserved β-roll region, where the putative cyclic nucleotide binding site is located [1, 60], and is primarily a result of hydrophobic interaction between residues in β4 and β5. In CNNM3, these residues are leucine (L575) and phenylalanine (F577). In CNNM2, they are methionine (M703) and phenylalanine (F705). Mutation of either residues in CNNM3 for a basic lysine residue (L575K and F577K) causes a shift from the singular dimer peak to a monomer in SV-AUC, showing that dimerization is disrupted (Figure 10 C).

These results indicate that the conserved β-roll region meant for binding cNMPs behaves as a dimerization surface. As such, the CNBH domain in CNNMs participates in dimerization rather than in the binding of cyclic nucleotides.

3.2 Structural characterization of the full cytosolic domain

Because of the lack of direct interaction between the two cytosolic domains, the CBS-pair and CNBH domains are thought to be independent of one another [46]. Thus, following the characterization of the CNBH domain, we sought out to determine how it behaved in conjunction with the CBS-pair domain by studying the full cytosolic domain. Further biophysical studies on the CBS-pair domain were also conducted along the way to establish a link between ligand binding and domain dimerization.

3.2.1 Structure of the cytosolic fragment

Crystal structures of the cytosolic fragments of ligand-free human CNNM3 and CNNM2 bound to a non-hydrolyzable ATP analog (β,γ-Imidoadenosine 5′-triphosphate; AMP-PNP) were solved by Seby Chen using previously determined structures of the CBS-pair and CNBH

domains [2]. Additional densities also reveal a long linker domain interconnecting both independent units (Figure 11 A-B)

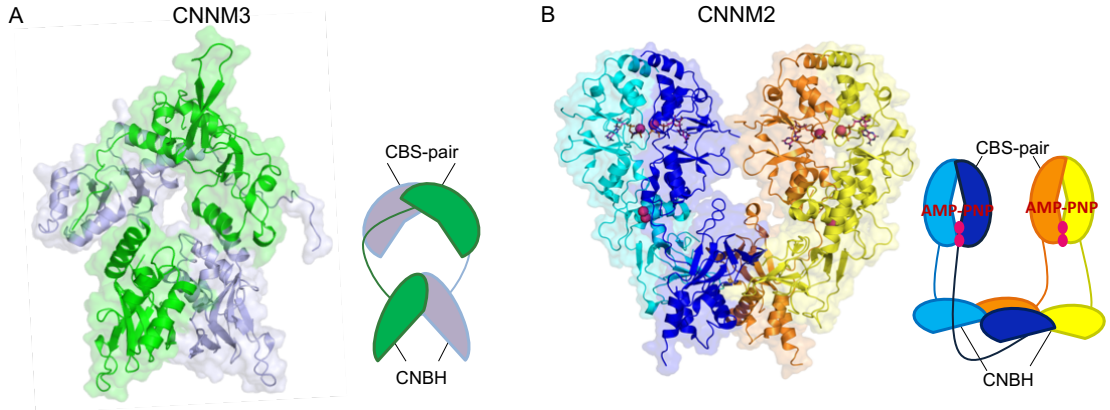

Figure 11. Structure of CNNM cytosolic fragment (A-B) Solved crystal structures of cytosolic fragments of human CNNM2, AMP-PNP-Mg^{2+} bound, and ligand-free CNNM3 (PDB code 6MN6 and 6N7E). Simplified caricatures depict an open V-shaped conformation for CNNM3, and a closed "flat" conformation for CNNM2.

The two structures show significant differences, presumably due to Mg^{2+} and nucleotide binding in CNNM2. AMP-PNP-Mg^{2+} bound CNNM2 shows two sets of crossed dimers, with the dimer interfaces between the CNBH and CBS displaying similar conformations as those observed in the isolated domains [1, 36]. In contrast, the ligand-free CNNM3 shows an open conformation, with a weak dimerization between both cytosolic fragments.

In both structures, the CBS-pair domains form dimers, but the angle between both protomers differs by 90° (Figure 12A). The CBS-pair domain of ligand-free CNNM3 shows a very distinctive "open" V-shaped conformation, which differs greatly from the previously reported "twisted" conformation observed in the isolated CBS-pair domain structure of ligand-free CNNM2. Meanwhile, AMP-PNP-Mg^{2+} bound CNNM2's CBS-pair domain forms a flat disk-like dimer, similar to the structure of the isolated CBS-pair domain of CNNM2 with ATP-Mg^{2+} bound [36]. An overlay of both structures (Figure 12B) reveals significant shifts in the α0 and α4 helices, with notable shifts in α1 and α3 as well, as seen in isolated CBS-pair domain structures [36].

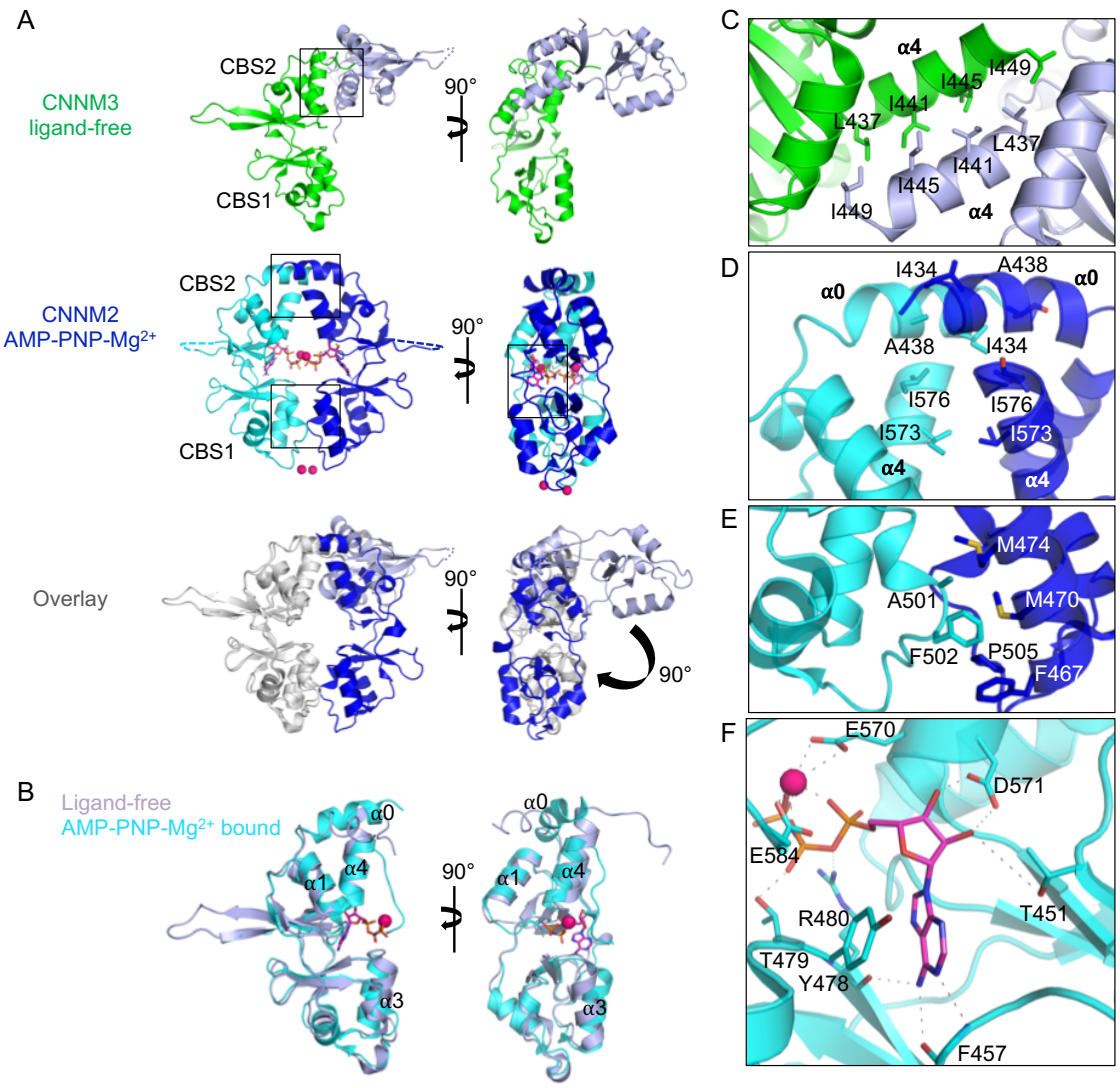

Figure 12. CBS modules undergo conformational changes upon AMP-PNP-Mg^{2+} binding
(A) Structural comparison of the CNNM CBS modules in ligand-free (CNNM3) and AMP-PNP-Mg^{2+} bound (CNNM2) states (B) Overlay of the structures of the CBS-pair domain of CNNM2 and CNNM3, showing the shifts in secondary structural elements (α-helices) between the ligand-free and ligand-bound states (C) Close-up view of the dimerization interface of the ligand-free CBS-pair domain of CNNM3. (D-E) Close-up view of the dimerization interface of the AMP-PNP-Mg^{2+} bound CBS-pair domain of CNNM2. Residues involved in hydrophobic contacts are labelled. (F) Close-up of AMP-PNP bound to the canonical nucleotide binding site of CNNM2. The residues involved in ligand binding are labelled.

The loose dimerization of the ligand-free CBS-pair domain is mediated through hydrophobic contacts between the leucine and isoleucine residues in the α4 helix in the CBS2 motif (Figure 12C). The more rigid dimerization of the AMP-PNP-Mg^{2+} bound CBS-pair domain is mediated through hydrophobic contacts in the α0 and α4 helices in the CBS2 motif

(Figure 12D), and through hydrophobic contacts in the α helices in the CBS1 motif (Figure 12E). The AMP-PNP analog binds to the cavity formed by the CBS2 motif (Figure 12F), which corresponds to the canonical binding site of the natural adenosine ligands ATP/ADP/AMP [34]. The adenine is stabilized by hydrophobic stacking with an aromatic residue (Y478) and by hydrogen bonding with the backbones of neighbouring residues (F457 and R480). The ribose sugar hydrogen-bonds with D571, T451, and T568's side chains. The imidotriphosphate is stabilized by hydrogen bonds with the side chains of R480 and T479, and by coordinating with Mg^{2+}. Finally, Mg^{2+} is also coordinated by the carboxylates of E570 and E584 [2, 36].

While the CBS-pair and CNBH domains do not physically interact with each other, the conformational changes observed in the CBS modules translated to the CNBH dimers, also showing a 90° difference between both units (Figure 13A).

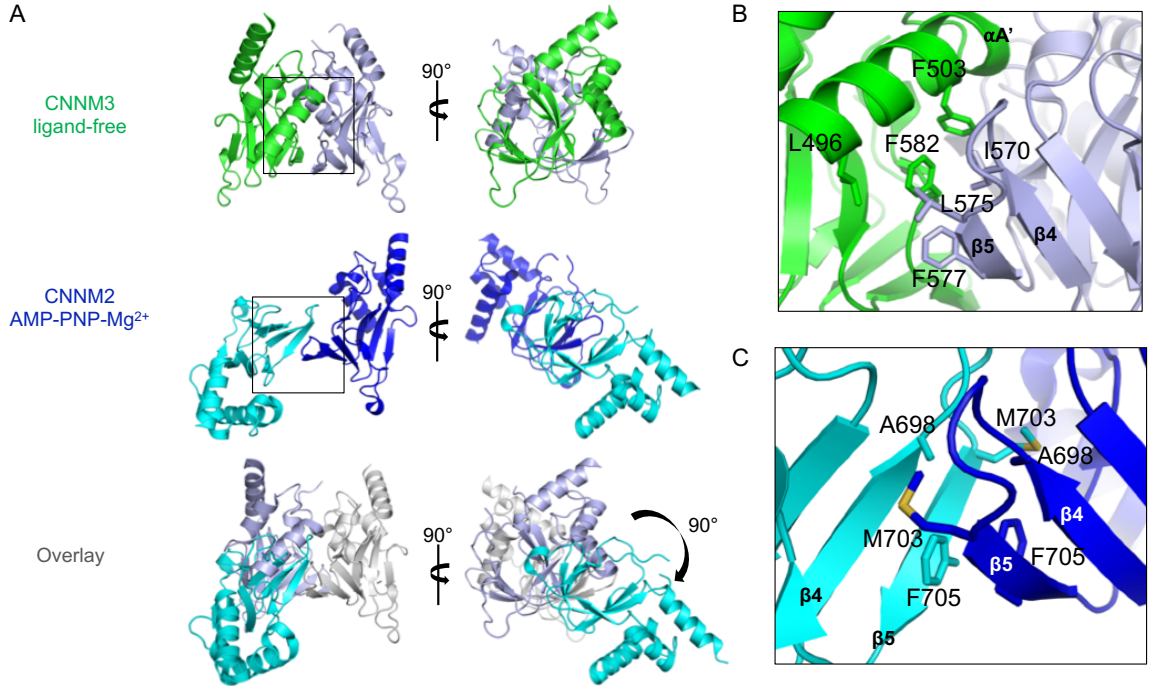

Figure 13. Conformational changes in the CBS-pair domain translates to the CNBH domain (A) Structural comparison of the CNBH domain of CNNMs in ligand-free (CNNM3) and ligand bound (CNNM2) states (B) Close-up view of the dimerization interface of the CNBH domain in ligand-free CNNM3. Dimerization is mediated through hydrophobic contacts of β-roll of one protomer and αA' helix of the other protomer. (C) Close-up view of the dimerization interface of the CNBH domain in AMP-PNP-Mg^{2+} bound CNNM2. Dimerization is mediated through hydrophobic contacts of β-roll of both protomers.

The CNBH domain of CNNM3 presents a new unreported V-shaped dimer conformation formed by hydrophobic contact between the residues (F577, L575 and I570) in the β4 and β5 elements of one protomer to the αA' residues (L496, F503 and F582) of another protomer (Figure 13B). In contrast, the CNBH domain of CNNM2 dimerizes in a flat conformation previously seen in the isolated CNBH structures of CNNM3 and CNNM2 [1]. Similarly, dimerization is mediated by hydrophobic contacts between the M703 and F705 residues in the β4 and β5 elements of each protomer (Figure 13C).

These results present the different conformations the CNNM cytosolic fragments, and their individual elements, can take in the presence and absence of a bound ligand, and demonstrate an significant structural interplay between the CBS-pair and CNBH domains.

3.2.2 The cytosolic domain dimerizes in solution

As seen above, the crystal structure of the CNNM3 cytosolic fragment shows a dimeric structure, whereas that of CNNM2 shows tetrameric structure (Figure 11). To verify the possibility of the proteins forming tetramers, the oligomerization states of the cytosolic fragments for human CNNM2-4 were studied using SV-AUC experiments with increasing protein concentrations. The resulting sedimentation distribution curves and molecular weights are shown in Figure 14 and Table IV respectively.

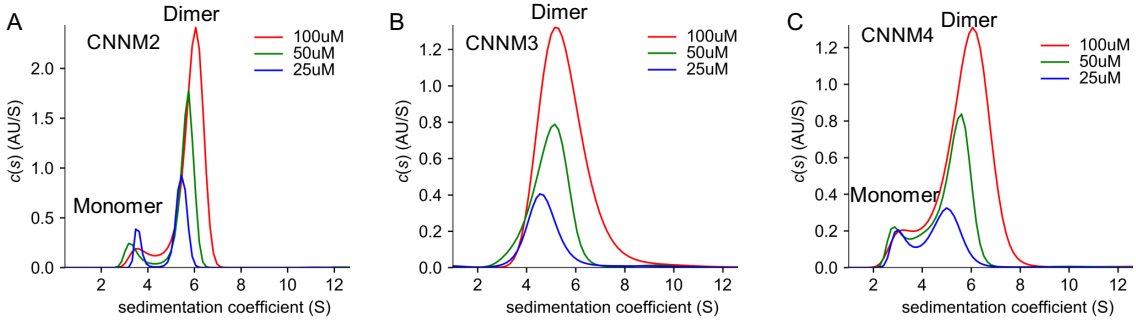

Figure 14. CNNM cytosolic fragments dimerize in a concentration-dependent manner SV-AUC distribution profile for cytosolic fragments of (A) CNNM2 (B) CNNM3 and (C) CNNM4 with varying protein concentrations. Monomer and dimer peaks are indicated.

Table IV: SV-AUC sedimentation coefficients and estimated molecular weights for human CNNM cytosolic fragments

		Sedimentation coefficient (S)		Estimated MW (kDa)		RMSD	Theoretical MW (kDa)	
Cytosolic domain	[Protein] (uM)	Peak 1	Peak 2	Peak 1	Peak 2		Monomer	Dimer
CNNM2	100	3.62	5.94	44.7	87.0	0.0051	40.7	81.4
	50	3.23	5.65	41.4	85.5	0.0046		
	25	3.55	5.43	43.1	83.7	0.0038		
CNNM3	100	-	5.35	-	84.0	0.0053	41.1	82.2
	50	-	4.99	-	80.7	0.0047		
	25	-	4.57	-	76.9	0.0029		
CNNM4	100	3.29	5.98	42.2	88.9	0.0049	40.2	80.4
	50	2.99	5.41	40.4	84.6	0.0039		
	25	3.08	4.99	39.0	81.3	0.0030		

The cytosolic fragments of CNNM2 and CNNM4 sediment as a mixture of monomer and dimers in a concentration dependent manner (Figure 14A and 14C), whereas that of CNNM3 sediments as a singular dimer (Figure 14B). No tetrameric peak was observed in any of the distribution curves.

To verify whether or not tetramerization is a result of ligand-binding, the cytosolic domain of CNNM4, which has the highest reported binding affinity for ATP-Mg^{2+} [27], was run using Rayleigh interference detection-based SV-AUC in the presence of ligands (i.e. ATP, Mg^{2+}, ATP-Mg^{2+}). The resulting sedimentation distribution curves and molecular weights are shown in Figure 15 and Table V respectively.

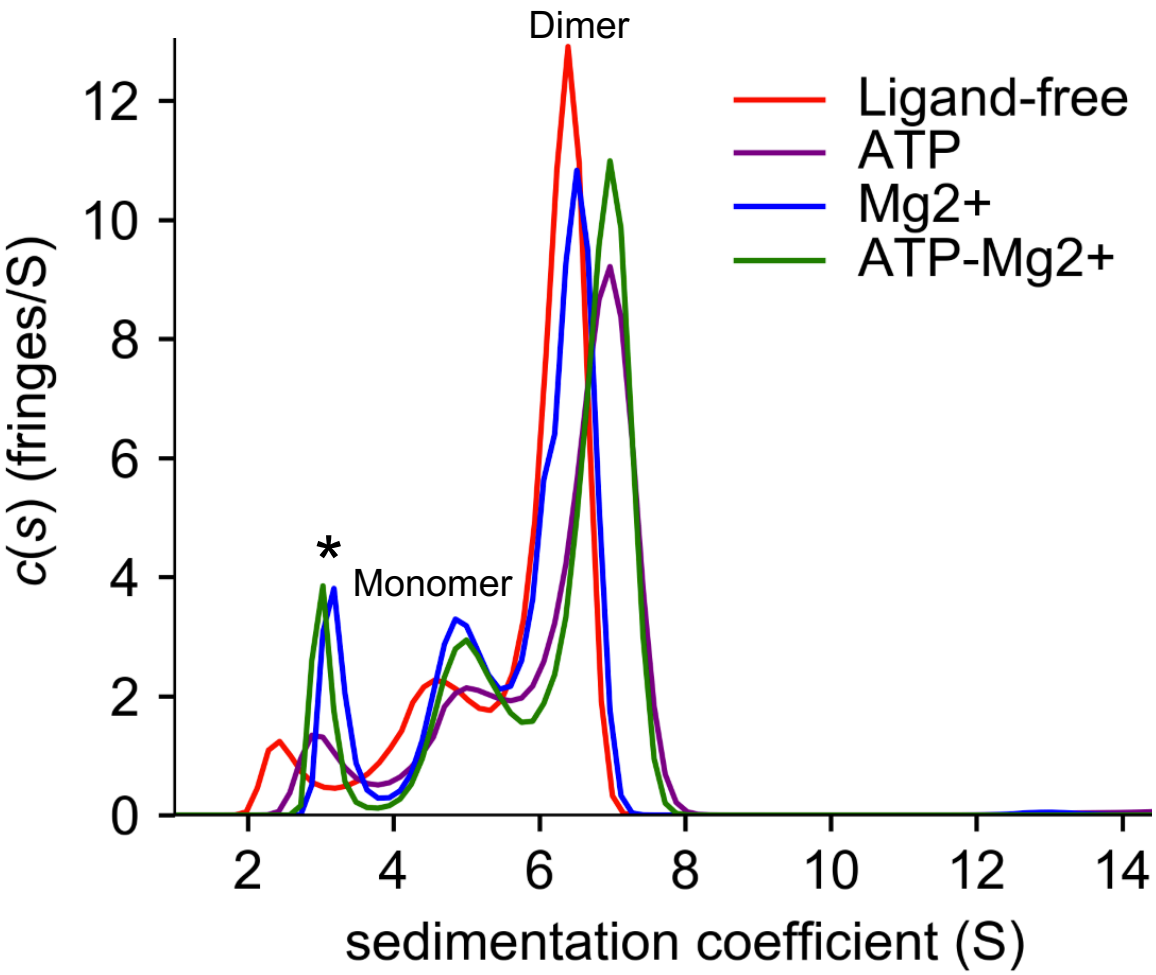

Figure 15. CNNM4 cytosolic fragment dimerization is independent of ligand binding. Sedimentation velocity distribution profiles of human CNNM4 cytosolic fragment ligand-free, with ATP, Mg^{2+} (10 mM $MgCl_2$), and ATP-Mg^{2+} (1 mM ATP,10 mM $MgCl_2$). Monomer and dimer peaks are indicated accordingly. The asterisk (*) indicates an artifact peak resulting from Rayleigh interference detection.

Table V: SV-AUC sedimentation coefficients and estimated molecular weights for human CNNM4 cytosolic fragment

		Sedimentation coefficient		Estimated MW (kDa)		RMSD	Theoretical MW (kDa)	
Protein	Condition	Peak 1	Peak 2	Peak 1	Peak 2		Monomer	Dimer
CNNM4 cytosolic fragment	No addition	4.51	6.47	37.6	76.5	0.0042	40.2	80.4
	ATP	5.01	6.95	44.3	83.9	0.0049		
	Mg^{2+}	4.98	6.50	43.9	79.1	0.0052		
	ATP-Mg^{2+}	5.00	6.96	44.1	85.0	0.0050		

It can be observed that addition of ligands to the cytosolic fragment of CNNM4 produces a similar monomer-dimer mixture (Figure 15 and Table V).Thus, the cytosolic domain primarily dimerizes in solution, regardless of ligand binding. The tetramer observed in the structure above is likely a result of crystal packing. Interestingly, the above sedimentation distribution profiles are similar to the ones produced in isolated CNBH domain (Figure 9B-D), reflecting the involvement of said domain in dimerization of the cytosolic fragment [1].

Thus, the cytosolic fragments form dimers in solution.

3.2.3 Ligand binding promotes CBS-pair domain dimerization in CNNMs

As seen above, the CBS-pair domain can either form a loose or rigid dimer conformer depending on its ligand-bound state (Figure 12). To investigate a possible link between ligand-binding and domain dimerization, SV-AUC experiments were done on the isolated CBS-pair domains of human CNNM1-4 with and without the addition of ATP-Mg^{2+}. The resulting sedimentation distribution curves and estimated molecular weights are shown in Figure 16 and Table VI respectively.

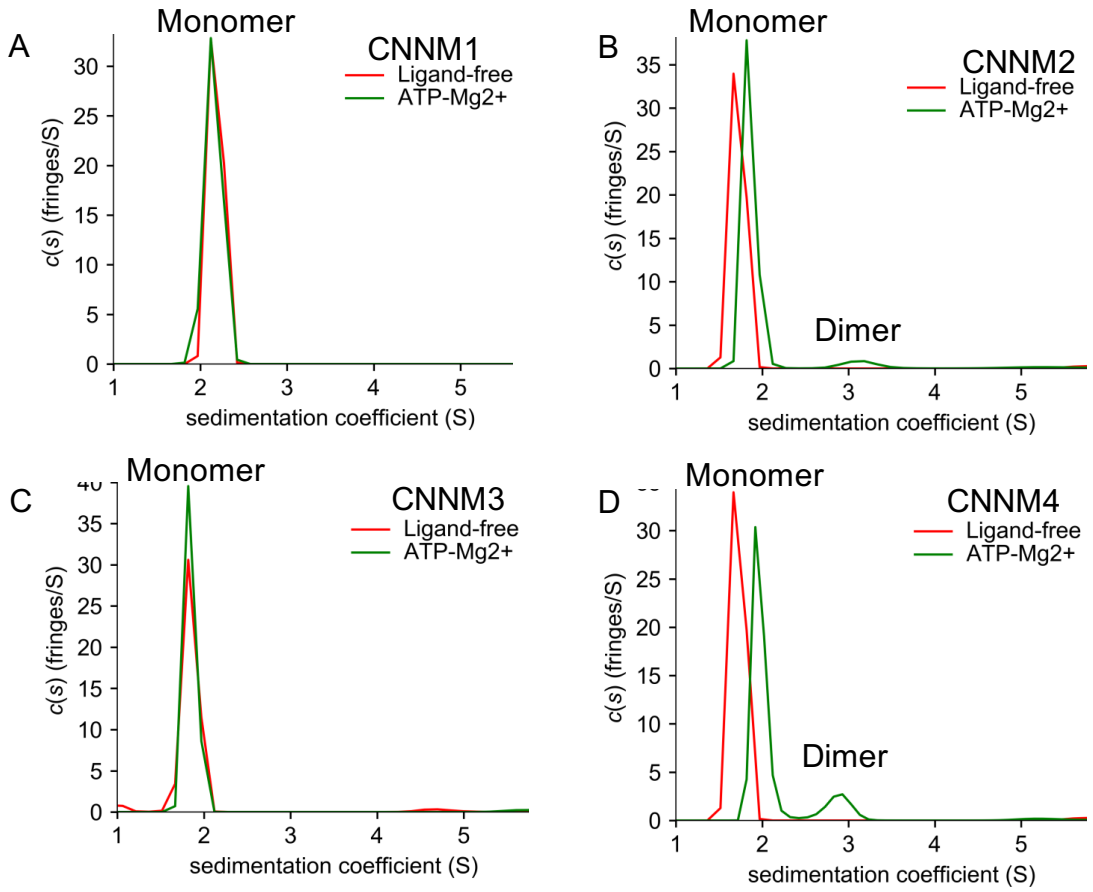

Figure 16. ATP-Mg^{2+} binding promotes CBS-pair domain dimerization Sedimentation velocity distribution profiles of CBS-pair domain of (A) human CNNM1 (B) human CNNM2 (C) human CNNM3 and (D) human CNNM4 with and without ATP-Mg^{2+} (1mM ATP, 10mM $MgCl_2$). Monomer and dimer peaks are indicated.

Table VI. SV-AUC sedimentation coefficients and estimated molecular weights for CBS-pair domain of human CNNM1-4

		Sedimentation coefficient		Estimated MW (kDa)		RMSD	Theoretical MW (kDa)	
CBS-pair domain	Condition	Peak 1	Peak 2	Peak 1	Peak 2		Monomer	Dimer
CNNM1	No addition	2.09	-	18.8	-	0.0051	18.0	36.0
	ATP-Mg^{2+}	2.10	-	18.9	-	0.0048		
CNNM2	No addition	1.83	-	17.7	-	0.0039	17.9	35.9
	ATP-Mg^{2+}	1.89	3.09	18.3	41.8	0.0046		
CNNM3	No addition	1.86	-	18.7	-	0.0049	17.6	35.2
	ATP-Mg^{2+}	1.85	-	18.3	-	0.0051		
CNNM4	No addition	1.73	-	18.3	-	0.0034		
	ATP-Mg^{2+}	1.96	2.87	19.1	42.1	0.0040	18.1	36.2

In the absence of ligands, all CBS-pair domain isoforms sediment as monomers. When ATP-Mg^{2+} is added, a weak dimer species is detected in the distribution profile for CNNM2, and a slightly more intense one for CNNM4 (Figure 16B and D). In contrast, CNNM1 and CNNM3 still sediment as monomers (Figure 16A and C). These results are interestingly reflective of the SPR and filter-binding assays previously conducted by Hirata *et al*. CNNM4 binds ATP-Mg^{2+} with the highest affinity followed by CNNM2, while CNNM1 and CNNM3 show very weak binding and no binding, respectively [27].

Using CNNM4, further SV-AUC experiments were run by varying the adenosine moieties used (ADP, AMP). The CBS-pair domain of CNNM4 was also run in the presence of PRL2, as the CBS-PRL complex was previously reported to produce crystals of the CBS dimer in the rigid "flat" conformer [46]. Additionally, we tested two disease-causing mutants which target ATP-binding, T495I and P409L (T568 and P482 in CNNM2) [2, 27]. The resulting sedimentation distribution curves and estimated molecular weights are shown in Figure 17 and Table VII respectively.

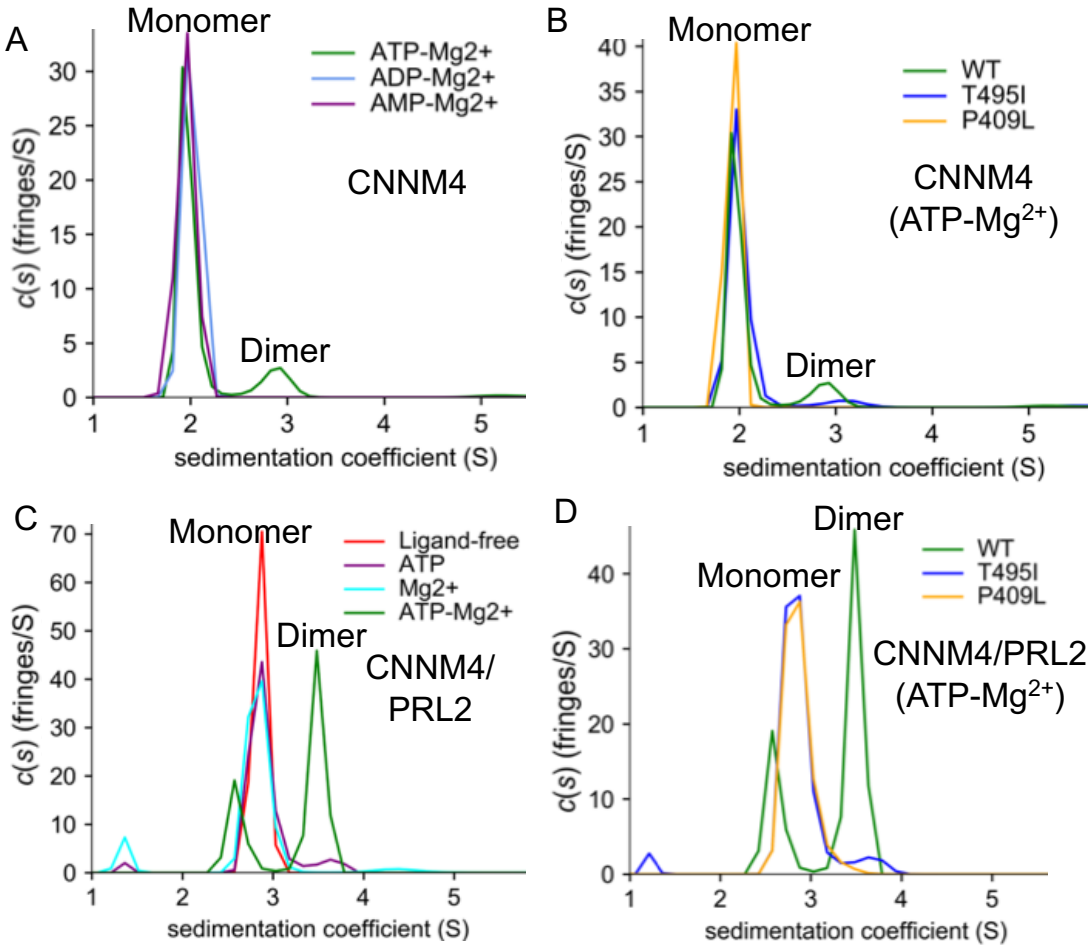

Figure 17. CBS-pair domain dimerization depends on ATP-Mg^{2+} binding and is enhanced by PRL binding Sedimentation velocity distribution profiles of human CNNM4 CBS-pair domain (A) in the presence of ATP-Mg^{2+}, ADP-Mg^{2+}, and AMP-Mg^{2+} (1 mM ADP, 1 mM AMP, 10 mM $MgCl_2$) (B) with ATP-Mg^{2+} using CNNM4 mutants (C) in complex with PRL2 in the presence of Mg^{2+}, ATP, and ATP- Mg^{2+} (1 mM ATP, 10 mM $MgCl_2$) (D) in complex with PRL2 in the presence of ATP-Mg^{2+}, using CNNM4 mutants. Monomer and dimer peaks are indicated.

Table VII. SV-AUC sedimentation coefficients and estimated molecular weights for CNNM CBS-pair domain and PRL-CBS complex

		Sedimentation coefficient		Estimated MW (kDa)		RMSD	Theoretical MW (kDa)	
Protein	Condition	Peak 1	Peak 2	Peak 1	Peak 2		Monomer	Dimer
CNNM4	No addition	1.73	-	18.3	-	0.0034		
	ATP-Mg^{2+}	1.96	2.87	19.1	42.1	0.0040	18.1	36.2
	ADP-Mg^{2+}	2.02	-	17.8	-	0.0041		
	AMP-Mg^{2+}	1.97	-	18.3	-	0.0054		
CNNM4 T495I	ATP-Mg^{2+}	1.97	-	18.5	-	0.0050	18.1	36.2
CNNM4 P409L	ATP-Mg^{2+}	1.99	-	18.9	-	0.0049	18.1	36.2
CNNM4/ PRL2	No addition	2.85	-	38.2	-	0.0034		
	ATP- Mg^{2+}	2.61	3.61	39.9	73.1	0.0040	38.8	77.7
	Mg^{2+}	2.83	-	38.9	-	0.0054		
	ATP	2.85	-	37.7	-	0.0041		
CNNM4 T495I/ PRL2	ATP-Mg^{2+}	2.83	-	36.9	-	0.0055	38.8	77.7
CNNM4 P409L/ PRL2	ATP-Mg^{2+}	2.84	-	39.4	-	0.0051	38.8	77.7

The dimer moiety of the CBS-pair domain of CNNM4 is only observed in the presence of ATP-Mg^{2+}, not ADP-Mg^{2+} nor AMP-Mg^{2+} (Figure 17A). Disrupting ATP binding in this domain is also shown to disrupt dimerization, as all mutants sediment as a singular monomer peak (Figure 17B).

Binding of PRL greatly enhances dimerization in the presence of ATP-Mg^{2+}, while addition of ATP or Mg^{2+} alone did not lead to dimer formation (Figure 17C). Similarly, disease-causing mutations in the CBS-pair domain that prevent ATP-Mg^{2+} binding also prevented dimer formation (Figure 17D).

These results establish an important link between ligand-binding (ATP-Mg^{2+} and PRL) and CBS-pair domain dimerization.

3.3 Structural characterization of full-length CNNM

This ultimate goal of our project is to arrive at the full structure of the CNNM, including the elusive DUF21 transmembrane domain. Due to the difficulty of expressing the mammalian membrane proteins, we turned to prokaryotic CNNMs from thermophilic species for their crystallization stability and ability to express in large scale in *E.coli*, [65, 66]. We also investigate the hypothesis regarding CNNMs' ability to form dimers using microscopic techniques.

3.3.1 Purification and crystallization of membrane-containing CNNM

Full length CNNM from a thermophilic bacteria (tCNNM) (residues 2-445) was purified along with its truncations (residues 2-423, and 2-332) in the presence of DDM for solubility. Constructs (2-445) and (2-423) consist of the DUF21 domain, the CBS-pair domain, and the C-terminal CorC/HlyC domain, whereas construct (2-332) has its CorC/HlyC domain removed (ΔCTD).

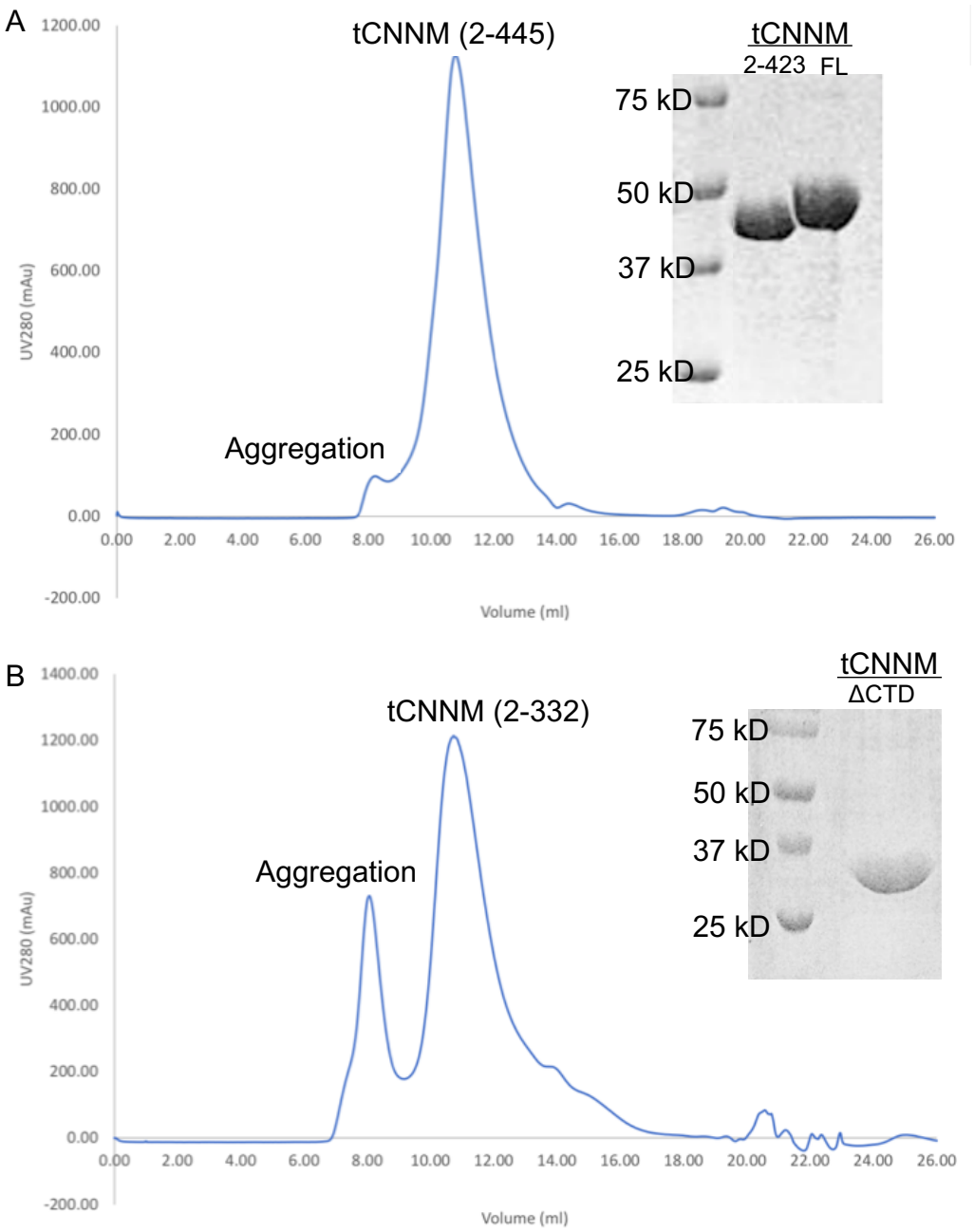

Figure 18. Purification profile of thermophilic CNNM Size exclusion chromatography profiles and SDS-PAGE gels for (A) full length (FL) tCNNM and (B) tCNNM truncated at its C-terminal CorC/HlyC domain (ΔCTD). Aggregation and peaks of interest are labelled accordingly on the graph.

The proteins behave well in the DDM detergent. Full length tCNNM, in particular, shows minimal aggregation in the elution profile (Figure 18A). The constructs yields up to 3 mg of protein per liter of culture, and is soluble up to 40 mg/ml. The C-terminal deletion construct appears to be more prone to aggregation than its full length counterpart (Figure 18B). This

protein yields about 1 mg of protein per liter of culture, and is soluble up to 15 mg/ml. SDS-PAGE gels of all constructs show high purity, demonstrating that large scale expression and purification of these membrane proteins is indeed possible.

Proteins were concentrated to a final concentration of at least 10 mg/ml and set up for sitting drop vapor diffusion crystallization. Only tCNNM$_{\Delta CTD}$ produced preliminary crystals, and under low ionic conditions (0.1 M Tris pH 8.0, 44% PEG200). Crystals were verified with SDS-PAGE before conducting optimization trials using the hanging drop method. Optimized crystals were then tested for diffraction using a home source (See Figure 19).

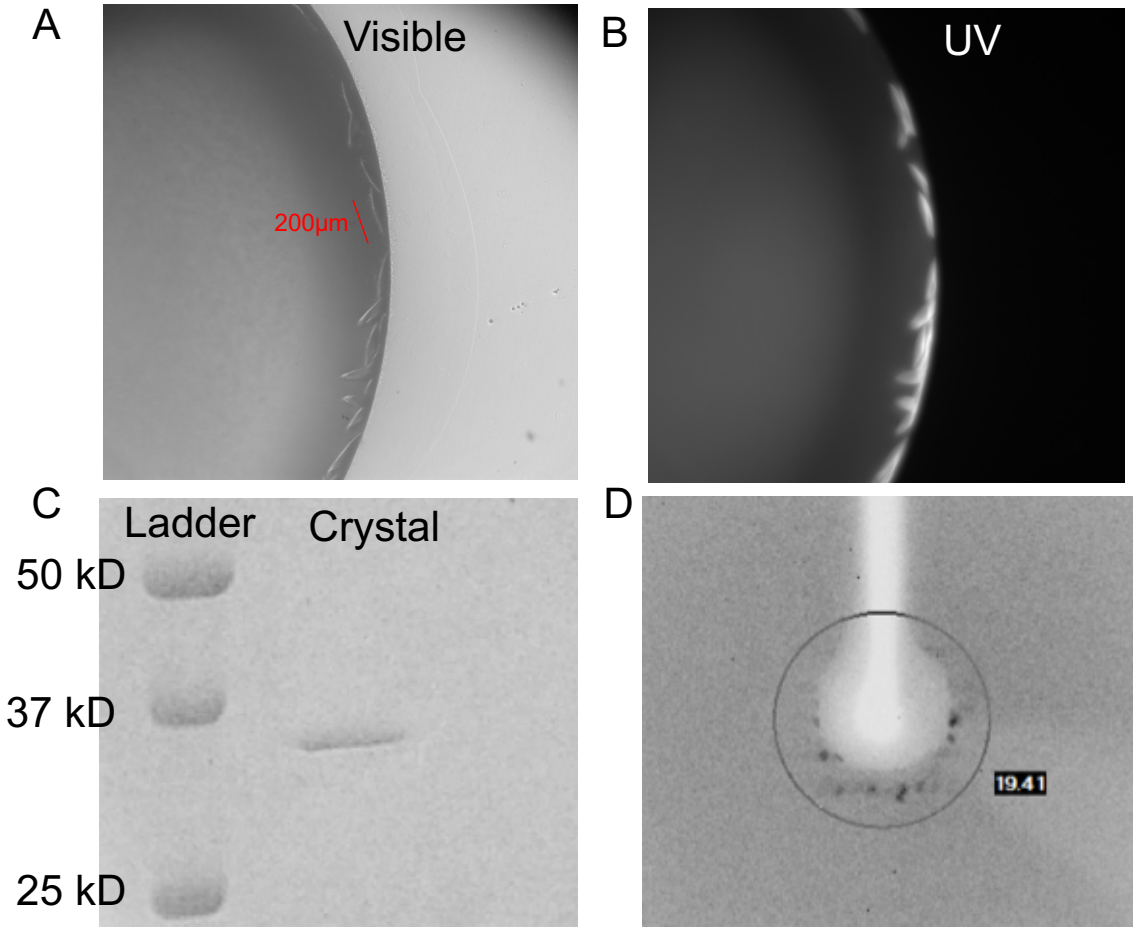

Figure 19. Crystallization and diffraction of tCNNM$_{\Delta CTD}$ crystals (A-B) Optimized crystal drop of 15 mg/ml tCNNM$_{\Delta CTD}$ crystallized in 36% PEG200 and 0.1M Tris pH 8.5, imaged under visible and UV light. Crystal size is indicated in red. UV light screening confirms the presence of (tryptophan-containing) protein crystals. Crystal size is indicated. (C) SDS-PAGE gel of crystal confirms crystallized tCNNM$_{\Delta CTD}$ construct (D) Diffraction pattern of protein crystal. Crystal diffracts to 19.41Å.

Preliminary optimization trials gave crystals that were just barely large enough (200 μm) to get scooped up with a nylon loop and measured for diffraction (Figure 19A). Even then, crystals diffracted to 19.41Å (Figure 19D), leaving hope for future optimizations (e.g. by further screening with crystallization additives) [66].

3.3.2. CNNM oligomerizes in solution

One of the reasons CNNM's status as an ion transporter is brought to question is the small number of helices that make up the transmembrane domain in the monomeric protein. Because of the CBS-pair domain's physiological propensity for dimerization, it is speculated that CNNMs also dimerize to mediate ion transport [35, 39]. Thus, in addition to solving the full structure, we would like to determine the *in vitro* oligomerization state of the CNNM protein to gain insight on its *in vivo* mechanism.

To study CNNM's oligomerization in solution, we resorted to the novel method single-molecule total internal reflection fluorescence (TIRF) microscopy. Very briefly, this technique consists of counting the number of fluorescently-labelled protein subunits localized on a surface (e.g. surface of a cell or resin-coated coverslip) during a series of photobleaching "steps." The number of photobleaching "steps" required to attenuate a fluorescent signal will correspond to the number of subunits of protein localized in one spot. The subunit counts measured from multiple spots are then compiled and plotted as a probability curve to determine the most likely oligomerization state of a protein [59].

Our approach consisted of labelling the prokaryotic CNNM using a fluorescent probe. The tCNNM sequence contains only one cysteine residue (C411) in its C-terminal cytosolic region, making it a desirable candidate for maleimide labelling. Prior to performing labelling, we assessed the availability of the cysteine thiol (-SH) group, as the residue was located directly inside of the CorC/HlyC domain (Figure 20A). Secondary structure prediction server JPred predicts C411 to be localized at the end of a β-strand (β4), and sequence alignment with the CorC/HlyC domain of CorC transporter in *E.coli* further cements this (Figure 20B and C). To finalize the assessment of the cysteine's accessibility, the cytosolic fragment of tCNNM

(residues 208-423) was labelled with a non-fluorescent maleimide reagent, N-ethylmaleimide (NEM), and analyzed with mass spectrometry.

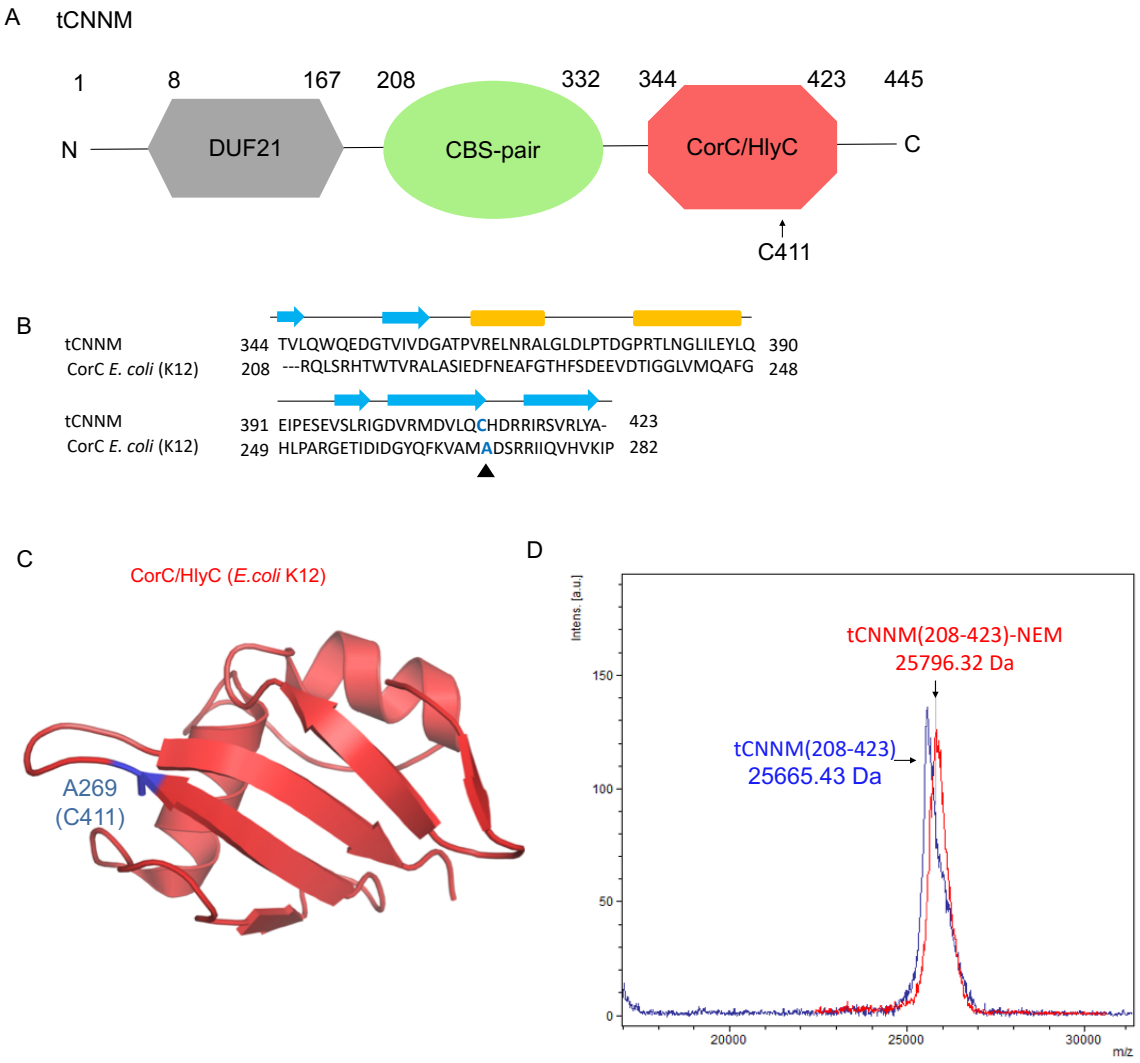

Figure 20. Lone cysteine residue in tCNNM is accessible for maleimide labelling
(A) Domain map of tCNNM. Lone cysteine 411 (C411) residue is indicated with an arrow. (B) Sequence alignment between CorC/HlyC domain of tCNNM and CorC efflux protein from *E. coli K12*. Secondary structural elements are indicated by a blue arrow (β-sheet) and an orange rectangle (α-helix). C411 and aligned residue (A269) are highlighted in blue and indicated with an arrowhead. (C) Structure of CorC/HlyC domain of CorC efflux protein from *E. coli K12*. The aligned residue (A269) is highlighted in blue. (D) MALD-MS spectrum of tCNNM (208-423) with (red) and without (blue) added NEM. Estimated masses are indicated next to each peak.

The estimated mass difference between the labelled and unlabelled sample was of 130 Da, which corresponds approximately to the mass of the NEM probe (Figure 20D). Therefore,

the cysteine residue was deemed accessible for labelling with a maleimide derivative conjugated to a fluorophore.

Full-length tCNNM was fluorescently labelled with a maleimide reagent conjugated to Oregon Green 488 (OG488). Following the removal of excess dye, the labelling efficiency was measured spectrophotometrically with a NanoDrop, comparing the absorbance intensity at 280 nm (protein) to that at 488 nm (OG488), to be about 99.6%. The protein was bound to a coverslip coated with nickel resin (for binding of polyhistidine-tagged protein) and underwent single-molecule TIRF microscopy. Conduction of the experiment and data analysis was done with the help of members from the Blunck laboratory (Université de Montreal). Preliminary step histograms and subunit number probability analyses are displayed in Figure 21 below.

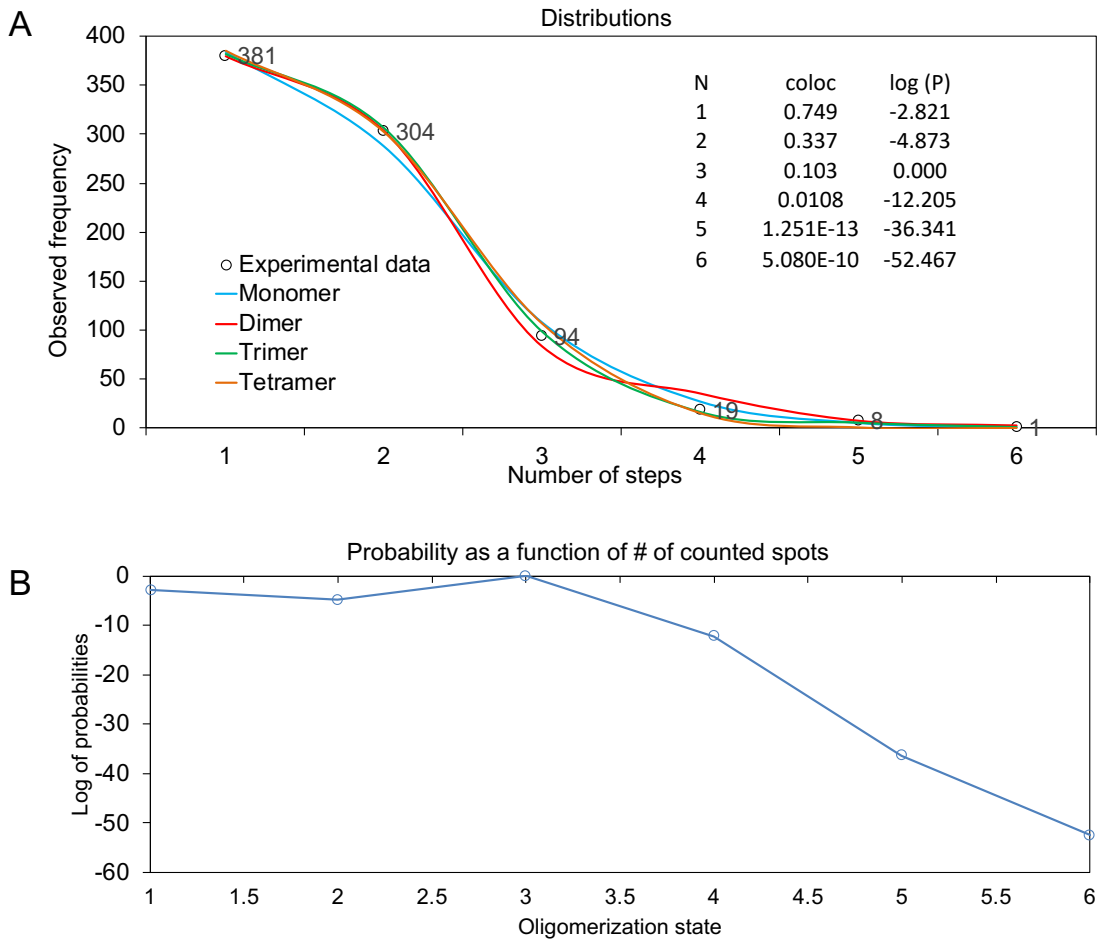

Figure 21. Subunit counting analysis of full length tCNNM (A) Distribution curve for probable oligomer fits (colored) overlaid with experimental data (open dots) featuring the number of steps counted per spot. Colocalization degree of subunits (coloc), probability (P) and log of probability (log(P)) for each oligomer state (N) are indicated on the side (B) Log of probability graphs of the oligomerization state of the protein as a function of the number of counted spots.

As seen from the experimental data presented in the distributions graph, there are a lot of counts of 1 to 3 photobleaching steps, indicating that 99% of the spots consist of one to three colocalized subunits. When comparing to the curve fittings of each oligomerization state (monomer to tetramer), the experimental data is shown to fit best with the curve corresponding to the trimer (Figure 21A). This sets the probability of the trimer state being the highest, followed by the monomeric and dimeric state (Figure 21B).

Therefore, according to these preliminary measurements, tCNNM has the highest probability of forming a trimer, followed by a monomer and a dimer.

CHAPTER 4. Discussion

Mg^{2+} is a biologically important cation required for over 600 enzymatic processes, ranging from energy metabolism to cell cycle regulation [4]. It also stabilizes many macromolecular complexes including mitochondrial membranes, ribosomes, and nucleic acids [4]. Because the mechanisms governing Mg^{2+} homeostasis are important, a particular focus is being placed on studying the structure, function, and regulation of Mg^{2+} ion transporters. The CNNMs are a large family of proteins involved in mediating ion (Mg^{2+}) transport in all kingdoms of life, from bacteria to humans [27, 30-33], but their status as Mg^{2+} transporters is still actively debated and their complete structure remains unknown [39]. This thesis aims to expand the current structural knowledge of these proteins by studying each of the domains that constitute them. Here, we characterized the structure of the C-terminal CNBH domain and the cytosolic fragment, as well as crystallized and investigated the *in vitro* stoichiometry of the full-length transmembrane protein.

4.1 CNBH domain mediates dimerization instead of cNMP binding

Our first goal was to characterize the CNBH domain and investigate the hypothesis surrounding the domain's ability to bind cyclic nucleotides to regulate CNNM function, as predicted from its homology to the cNMP binding domains of previously characterized ion channels [23].

We successfully solved the structure of the CNBH domains of human CNNM2 and CNNM3 with X-ray crystallography, and demonstrated that it shares many of the structural elements present in most cNMP-binding motifs [23, 47]. However, an overlay of the CNNM3 CNBH domain with the previously characterized cAMP-binding domain of bacterial K^+ channel MloK1 reveals the absence of a C-terminal α-helix required for structurally stabilizing the cNMP ligands (Figure 7C). Additionally, the putative binding site for cNMP ligands in CNBH, located in the conserved β-roll region, is obstructed by a bulky aromatic residue (Y628 in CNNM3) [1]. *In vitro* experiments (scanning differential fluorometry and ^{1}H-^{15}N HSQC NMR) on purified CNNM1-4 CNBH showed no significant shift in the melting temperatures nor in the spectral

signals of the proteins (Figure 8 and Tables I-II) when cAMP and cGMP were added. The absence of structural elements required to bind cNMP, together with the lack of significant change in the melting temperature and NMR spectral shifts strongly suggest that the CNBH domain does not bind cNMPs.

Instead, the conserved β-roll region was shown to behave as a dimerization interface. Three-dimensional structures of the CNBH domains of CNNM2 and CNNM3 demonstrate contacts taking place between conserved hydrophobic residues in the β4 and β5 elements (M703 and F705 in CNNM2, L575 and F577 in CNNM3) which, when mutated to basic lysine residues, cause the domain to sediment as a monomer with analytical ultracentrifugation (Figure 10). Additionally, each CNNM isoform shows a different propensity for CNBH dimerization. CNNM3 sediments only as a dimer. CNNM1-2 show intermediate dimerization. CNNM4 dimerizes the least (Figure 9). This propensity to dimerize may be linked to the residues responsible for forming contacts between the protomers. CNNM1 and CNNM3, which have a higher likelihood of dimerization, form hydrophobic contacts through leucine and phenylalanine residues. CNNM2 and CNNM4, which are less likely to dimerize, form contacts through methionine and phenylalanine residues [1]. Leucine is more hydrophobic than methionine according to the Kyte-Doolittle scale, and thus may form more stable contacts during CNBH dimerization [67]. Further differences in CNBH dimerization between CNNM2 and CNNM4, and between CNNM1 and CNNM3, may be due to additional contacts between residues found within and/or outside of the β-roll interaction surface, as later seen in Figure 10 (e.g. alanine in CNNM2 and isoleucine in CNNM3) [2]. Further studies are required to investigate these differences.

Interestingly, the propensity for CNBH to dimerize is inversely correlated to each isoform's ability to mediate Mg^{2+} efflux; CNNM4 shows the highest activity, CNNM2 shows intermediate activity, CNNM1 shows weak activity, and CNNM3 shows no efflux activity at all [27]. Dimerization of the cNMP binding motif as a means of mediating transporter activity has been previously described in the voltage-gated KCNH (ether-à-go-go K+) channel [68]. Similar to CNNM's CNBH domain, the cNMP binding domain of KCNH cannot bind cyclic nucleotides

due to a structural obstruction of the β-roll binding pocket [68]. Instead, it dimerizes through its β-roll region to mediate channel activity in a cNMP-independent manner [69].
Thus, the CNBH domain appears to modulate CNNM function through dimerization rather than by binding to cyclic nucleotides.

4.2 CNBH domain and CBS-pair domain interact through interdomain loop

As there are no previously reported interactions between the CNBH domain and the CBS-pair domain, they are presumed to behave independently of one another [46]. To that effect, soon after characterizing the CNBH domain on its own, we endeavored to study the full cytosolic fragment's structure to investigate the possibility of an interplay between the two cytosolic domains.

We successfully solved the structures of the cytosolic fragments of human CNNM2 in the ligand (AMP-PNP-Mg^{2+}) bound state and of ligand-free human CNNM3 using X-ray crystallography, and demonstrate important structural differences between them. The CBS module in AMP-PNP-Mg^{2+} bound CNNM2 adopts a similar "flat" conformer as previously reported ATP-Mg^{2+} bound form of isolated CBS-pair domain of CNNM2 (Figure 12A) [2, 36]. In contrast, ligand-free CNNM3 presents a new loosely bound V-shaped dimer which differs from the "twisted" conformer observed in the isolated CNNM2 CBS-pair domain, but is observed in chloride channels [2, 70]. Interestingly, the conformational differences observed in the CBS-pair domain of CNNM2 and CNNM3 are also translated to their CNBH domains, to which they are connected via interdomain linkers. The CNBH domain from CNNM2's cytosolic fragment shares the same "flat" dimer conformation observed in the isolated CNBH domain (Figures 10B and 13A) [1]. This gives the AMP-PNP-Mg^{2+} bound CNNM2 cytosolic fragment an overall "closed" conformation (Figure 11A) [2]. Meanwhile, the CNBH domain from CNNM3's cytosolic fragment adopts a novel V-shaped dimer formed by different hydrophobic contacts, thereby giving the ligand free CNNM3 construct an overall "open" conformation (Figures 11A and 13) [2].

The CBS-pair domain is regarded as an intracellular sensor for nucleotides and metal ions in its host proteins [35]. This is the case for bacterial Mg^{2+} MgtE, wherein the CBS-pair domain acts as a sensor for intracellular Mg^{2+} through conformational changes. Under low Mg^{2+} conditions, the dimer interface between the paired CBS-pair domain of MgtE is loosened by the repulsive forces of conserved acidic residues in the CBS2 motifs, and the resulting CBS module adopts an "open" conformation. Meanwhile, under high Mg^{2+} conditions, Mg^{2+} binds to these acidic residues, resulting in a "closed" dimeric assembly of the CBS module [44-45]. Additionally, models of the MgtE gating mechanism depict that the connected cytosolic N-terminal domain (N-domain) is also involved in these regulatory conformational changes: the N-domain closes on the CBS-pair domain under high intracellular Mg^{2+} and forms a compact, globular structure, but swings away from it under low Mg^{2+} conditions [45]. These structural changes are transmitted to the transmembrane domain, thereby modulating the ion-conducting pore [43-45].

While no physical contact exists between the two cytosolic domains in CNNMs [46], it is still likely that they both play a similar regulatory role with an open-to-closed conformational change driven by ligand (ATP-Mg^{2+}) binding. Therefore, our findings support the existence of a structural interplay between the CNNM CNBH and CBS-pair domains.

4.3 CBS-pair domain dimerization linked to ligand binding

The isolated CNNM CBS module forms either a compact "flat" or "twisted" dimer assembly under ligand (e.g. ATP-Mg^{2+} or PRL) bound and ligand-free conditions, respectively [36, 41, 46]. The novel loosely-bound "open" V-shaped dimer discovered in the cytosolic fragment of ligand-free CNNM3 thus raises questions about a possible link between ligand-binding and the dimerization state of the CBS-pair domain [2].

SV-AUC reveals that the CBS-pair domain of all CNNM isoforms sediment as monomers in the ligand-free state. A dimeric species is detected in CNNM4 and CNNM2 (albeit weaker) when ATP-Mg^{2+} is added, but CNNM1 and CNNM3 still sediment as monomers (Figure 16 and Table VI). These observations directly correlate to the binding affinities measured for each CNNM isoform measured by Hirata *et al*: CNNM4 binds ATP-Mg^{2+} with the highest

affinity, CNNM2 shows almost a 4-fold weaker affinity, CNNM1 binds ATP-Mg^{2+} with weak near-micromolar affinity, and CNNM3 does not bind ATP-Mg^{2+} [27]. Interestingly, dimerization in CNNM4 was only observed with ATP-Mg^{2+}, not with ATP or Mg^{2+} alone, nor with other molecules such as AMP-Mg^{2+} or ADP-Mg^{2+} (Figure 17 and Table VII) [2]. The first set of observations is consistent with the SPR and filter binding experiments previously conducted by Hirata *et al*, which revealed that none of the CNNM isoforms bind ATP in the absence of added Mg^{2+} [27]. Corral-Rodríguez *et al* have also demonstrated that Mg^{2+} favors ATP binding using NMR, speculating that it alleviates the negative charges the acidic residues in the CBS-pair domain's binding pocket and the β- and γ-phosphates of ATP [36]. Thus, ATP binding in CNNM is Mg^{2+}-dependent. The second set of observations is consistent with radioactivity binding assays conducted by Hirata *et al*, which showed a very weak micromolar affinity of CNNM2 to AMP-Mg^{2+}, suggesting a requirement for the terminal phosphate to bind to the CBS-pair domain [27]. This, however, contrasts the findings from Corral-Rodríguez *et al*, wherein the CNNM2 CBS-pair domain was shown to bind AMP and ADP in protein crystals and in NMR [36].

Addition of PRL2 enhanced dimerization of the CBS-pair domain in the presence of ATP-Mg^{2+} (Figure 17C and Table VII). PRL binding induces a "flat" CBS-pair dimer assembly similar to nucleotide binding in crystal structures [46]. Thus, the two elements likely produce a cooperative effect which further stabilizes the dimer species in solution. In contrast, disease-causing mutations which target ATP binding in CNNM4 were found to sediment as sole monomeric species, even when bound to PRL (Figure 17D and Table VII) [2], further cementing the link between ligand binding and domain dimerization.

Going by the speculation that the "flat" dimer is more stable than its "twisted" or "open" counterpart due to extended hydrophobic contacts between complementary CBS1 and CBS2 motifs [2, 36], it is curious that the PRL-CBS complex sediments only as a dimer in the presence of ATP-Mg^{2+}. Similarly, the disease mutation T495I's CNNM2 counterpart (T568I), while disrupting nucleotide binding, also adopts this characteristic "flat" assembly of the CBS module [36]. However, its CNNM4 homolog sediments as a monomer (Figure 17B). It is crucial to note that these crystals were obtained under high protein concentrations (>10 mg/ml) which would favor protein-to-protein contact [36, 65], compared to the lower concentrations used in our

analytical ultracentrifugation experiments (≤1mg/ml). This introduces a concentration-dependent element to the dimerization dynamic which goes beyond the abilities of sedimentation velocity, and settles in the field of sedimentation equilibrium [71]. As such, the CBS-pair domain may constantly exchange between the monomer and dimer state in solution, with increase in protein concentration or ligand addition potentially shifting the equilibrium towards a (detectable) dimer state. Further studies are required to define this complex dynamic.

These experiments were also conducted using the isolated CBS-pair domain, and may not reflect the domain's physiological behavior in the full CNNM protein. For instance, the previously characterized CNBH domain provides a scaffold for dimerization, as shown by the cytosolic fragment's strong propensity to dimerize in solution (Figure 14 and Table IV), regardless of ligand-binding conditions (Figure 15 and Table V). As such, the CBS-pair domain would already exist as a dimer physiologically due to an increase in the local concentration [72]. Furthermore, should the CNNM protein indeed dimerize through its DUF21 transmembrane domain, this would further contribute to the CBS-pair domain's dimerization as well, regardless of ligand binding [72]. To that effect, and to reiterate the point above, the dimer peaks detected in the isolated CBS-pair domain's SV-AUC distribution graphs are likely dimers that have been stabilized by extensive contacts following ligand binding, as previously demonstrated [36, 41, 46]. Therefore, our findings would be linking stable CBS-pair domain dimer formation with ligand-binding instead.

4.4 The CNNM protein forms oligomers

Because only four helices make up the CNNM transmembrane domain, the proteins' status as divalent ion transporters is actively contested [39]. However, the CNNM proteins' ubiquity and vast involvement in divalent metal transport leads authors to stipulate that the CNNM proteins self-assemble, thereby increasing the number of helices to form a functional ion-conducting pore [24]. Thus, we resolved to characterize the structure of the full membrane protein as well as define its oligomerization state in solution to investigate this hypothesis.

We have demonstrated the possibility of expressing and purifying prokaryotic isoforms of the protein in large scale and producing preliminary diffracting crystals, providing the

possibility for a full structural characterization in the near future (Figure 18 and 19). The relatively high yield provided by the full length protein also permitted us to conduct labelling experiments for single molecule TIRF microscopy to determine its *in vitro* protein stoichiometry. Several counts of one to three photobleaching "steps" were measured under the fluorescent microscope, indicating a propensity for the tCNNM subunits to colocalize and assemble. However, assuming 100% labelling, the experimental "step" data was shown to fit best with the hypothetical trimer curve (Figure 21). This leads to a higher probability for tCNNM assembling as a trimer, compared to the potential monomer and dimer assemblies. These results contrast with previous findings on the general propensity of the CBS-pair domain to form dimers [30, 35]. These results also contrast with our own findings that the cytosolic fragments of CNNM, individual or connected, also display a tendency for dimerization (Figure 9, 14-17). A likely source of error is relatively high concentration of purified protein bound to the coverslip in the experiment (30 uM), providing a lot of background fluorescence and making it more difficult for individual protein complexes to be resolved. Thus, an alternative course of action would be to express the proteins in *E. coli* cells labelled with Green Fluorescent Protein (GFP) as described in the protocol of McGuire, H. *et al*, allowing for minimal surface expression levels [59]. Moreover, the experiment itself both relies on and is sensitive to fluorescence bleaching. If we were to continue these experiments, additional precautions would be taken to ensure none of the dye molecules were bleached before subunit counting [59]. Finally, solving the full structure of the protein with X-ray crystallography will also determine the oligomerization state of the protein [59].

Regardless, our findings provide preliminary evidence of self-assembly for the CNNM proteins.

4.5 Concluding remarks

In this thesis, we provide one the first structural characterizations of the CNBH domain and of the full cytosolic fragments for eukaryotic CNNMs [1, 2]. We also provide the first experimental evidence of purification and successful crystallization for the transmembrane protein.

We establish the CNBH domain's role as a dimerization surface for CNNMs rather than a cNMP binding site, and suggest a regulatory role similar to that of the cNMP binding domain of the KCNH (ether-à-go-go K^+) channel [68-69]. Our crystal structures for the full cytosolic domain not only reveal a novel "open" conformation in the CBS module for the CNNM proteins, but suggest a structural interplay between both CBS-pair and the CNBH domains in the form of large scale "open-to-closed" changes depending on their environment [2]. To this, we propose the existence of a modulating mechanism similar to that of bacterial Mg^{2+} transporter MgtE [43-45], wherein the cytosolic domains adopt an "open" configuration under a low Mg^{2+} environment, and a "closed" one under high Mg^{2+} conditions. The conformational changes translate to the DUF21 domain through the $\alpha 0$ helices of the CBS-pair domain, modulating the gating of the ion-conducting pore [36].

Our SV-AUC results also suggest a link between ligand-binding and dimerization in the CBS-pair domain [2]. In fact, dimerization is prevalent in all of the cytosolic domains and, potentially, in the transmembrane protein as well. In spite of the fluorimetry results, we expect the full length protein to exist in a dimeric state. These findings support speculations that the CNNM protein adopt an oligomeric structure, thereby assembling sufficient transmembrane helices to form a functional ion-conducting pore, as observed in the Mg^{2+} transporters MgtE, CorA, and TRPM7 [21, 43-45, 49].

Thus, our findings provide important structural insight to the largely uncharacterized CNNM proteins, helping us better understand their biochemical behaviors. The availability of purified membrane protein also leaves room for future functional assays by reconstitution into lipid membranes for either vesicular ion transport or patch clamping experiments [45]. These studies will allow us to establish the proteins' definitive role in ion transport, and help us better understand their involvement in magnesium homeostasis, ion transport-associated diseases, and cancer.

REFERENCES

1. Chen, Y.S., Kozlov, G., Fakih, R., Funato, Y., Miki, H., and Gehring, K. (2018) The cyclic nucleotide–binding homology domain of the integral membrane protein CNNM mediates dimerization and is required for Mg^{2+} efflux activity. *J Biol Chem.* **293**, 19998–20007.
2. Chen, Y.S., Kozlov, G., Fakih, R., Yang, M., Zhang, Z., Kovrigin, E.L., and Gehring K. (2020) Mg^{2+}-ATP Sensing in CNNM, a Putative Magnesium Transporter. *Structure* **28**, 324–335
3. Elin, R.J. (1988) Magnesium metabolism in health and disease. *Disease-a-month : DM*. **34**, 161-218
4. Jahnen-Dechent, W., and Ketteler, M. (2012) Magnesium basics. *Clinical kidney journal*. **5**, i3-i14
5. Romani, A.M. (2011) Cellular magnesium homeostasis. *Archives of biochemistry and biophysics*. **512**, 1-23
6. Hardy, S., Uetani, N., Wong, N., Kostantin, E., Labbé, D.P., Bégin, L.R., Mes-Masson, A., Miranda-Saavedra, D., and Tremblay, M.L. (2014) The protein tyrosine phosphatase PRL-2 interacts with the magnesium transporter CNNM3 to promote oncogenesis. *Oncogene*, **34(8)**, 986-95.
7. Funato, Y., Yamazaki, D., Mizukami, S., Du, L., Kikuchi, K., and Miki, H. (2014) Membrane protein CNNM4-dependent Mg^{2+} efflux suppresses tumor progression. *J Clin Invest.* **124**, 5398-5410
8. Yamazaki, D., Funato, Y., Miura, J., Sato, S., Toyosawa, S., Furutani, K., Kurachi, Y., Omori, Y., Furukawa, T., Tsuda, T., Kuwabata, S., Mizukami, S., Kikuchi, K., and Miki, H. (2013) Basolateral Mg2+ extrusion via CNNM4 mediates transcellular Mg2+ transport across epithelia: a mouse model. *PLoS genetics.* **9**, e1003983
9. Chamnongpol, S., and Groisman, E.A. (2002) Mg^{2+} homeostasis and avoidance of metal toxicity. *Mol Microbiol.* **44**, 561-571
10. Stuiver, M., Lainez, S., Will, C., Terryn, S., Günzel, D., Debaix, H., Sommer, K., Kopplin, K., Thumfart, J., Kampik, N.B. *et al*. (2011) CNNM2, encoding a basolateral protein required for renal Mg^{2+} handling, is mutated in dominant hypomagnesemia. *Am. J. Hum. Genet* **88**, 333–343

11. Groisman, E.A., Hollands, K., Kriner, M.A., Lee, E.J., Park, S.Y., and Pontes, M.H. (2013) Bacterial Mg^{2+} homeostasis, transport, and virulence. *Annual review of genetics.* **47**, 625-646

12. Moomaw, A.S., and Maguire, M.E. (2008) The unique nature of Mg^{2+} channels. *Physiology (Bethesda, Md.)*. **23**, 275-285

13. Workinger, J.L., Doyle, R.P., and Bortz, J. (2018) Challenges in the Diagnosis of Magnesium Status. *Nutrients.* **10(9)**, 1202.

14. Mahler, J. and Persson, I. (2012) A Study of the Hydration of the Alkali Metal Ions in Aqueous Solution. *Inorg. Chem. 2012*, **51**, 425−438

15. Black, C.B., Huang, H.W., and Cowan, J. A. (1994) Biological coordination chemistry of magnesium, sodium, and potassium ions. Protein and nucleotide binding sites. *Coordination Chemistry Reviews*, **1351136**, 165-202

16. Hmiel, S.P., Snavely, M.D., Miller, C.G., and Maguire, M.E. (1986). Magnesium transport in Salmonella typhimurium: characterization of magnesium influx and cloning of a transport gene. *Journal of Bacteriology*. **168 (3)**, 1444–50.

17. Townsend, D.E., Esenwine, A.J., George, J., Bross, D., Maguire, M.E., and Smith, R.L. (1995). Cloning of the mgtE Mg^{2+} transporter from Providencia stuartii and the distribution of mgtE in gram-negative and gram-positive bacteria. *Journal of Bacteriology.* **177 (18)**, 5350–4.

18. Kersey, C.M., Agyemang, P.A., Dumenyo, C.K. (2012) CorA, the magnesium/nickel/cobalt transporter, affects virulence and extracellular enzyme production in the soft rot pathogen Pectobacterium carotovorum. *Mol Plant Pathol* **13(1)**, 58–71

19. Tao, T., Snavely, M.D., Farr, S.G., and Maguire, M.E. (1995) Magnesium transport in Salmonella typhimurium: mgtA encodes a P-type ATPase and is regulated by Mg^{2+} in a manner similar to that of the mgtB P-type ATPase. *J Bacteriol.* **177**, 2654–2662.

20. Romani, A. (2007) Regulation of magnesium homeostasis and transport in mammalian cells. *Archives of Biochemistry and Biophysics* **458**, 90–102

21. Duan, J., Li, Z., Li, J., Hulse, R.E., Santa-Cruz, A., Valinsky, W.C., Abiria S.A., Krapivinsky, G., Zhang, J. and Clapham, D. E. (2018) Structure of the mammalian TRPM7, a magnesium channel required during embryonic development. *PNAS* , **115 (35)** E8201-E8210

22. Goytain, A., and Quamme, G.A. (2005) Functional characterization of ACDP2 (ancient conserved domain protein), a divalent metal transporter *Physiol. Genomics* **22**, 382-389

23. Wang, C.Y., Shi, J.D., Yang, P., Kumar, P.G., Li, Q.Z., Run, Q.G., Su, Y.C., Scott, H.S., Kao, K.J., She, and J.X. (2003) Molecular cloning and characterization of a novel gene family of four ancient conserved domain proteins (ACDP). *Gene.* **306**, 37-44.

24. de Baaij, J.H., Stuiver, M., Meij, I.C., Lainez, S., Kopplin, K., Venselaar, H., Müller, D., Bindels, R.J., and Hoenderop, J.G. (2012) Membrane topology and intracellular processing of cyclin M2 (CNNM2). *J Biol Chem.* **287(17)**, 13644-55.

25. Chandran, U., Indu, S., Kumar, A.T.R., Devi, A.N., Khan, I., Srivastava, D., and Kumar, P.G. (2016) Expression of Cnnm1 and Its Association with Stemness, Cell Cycle, and Differentiation in Spermatogenic Cells in Mouse Testis1. *Biology of Reproduction.* **95**, 7, 1-12-17, 11-12

26. Alderton, A., Davies, P., Illman, K., Brown, and D.R. (2007) Ancient conserved domain protein-1 binds copper and modifies its retention in cells. *J Neurochem.* **103(1)**, 312-21.

27. Hirata, Y., Funato, Y., Takano, Y., and Miki, H. (2014) Mg^{2+}-dependent interactions of ATP with the cystathionine-β-synthase (CBS) domains of a magnesium transporter. *J Biol Chem.* **289**, 14731-14739

28. Stamatiou, K.N., and Karakos, C.D. (2010) Urofacial syndrome: A subset of neurogenic bladder dysfunction syndromes? *Indian journal of urology : IJU : journal of the Urological Society of India.* **26**, 582-584

29. Polok, B., Escher, P., Ambresin, A., Chouery, E., Bolay, S., Meunier, I., Nan, F., Hamel, C., Munier, F.L., Thilo, B., Mégarbané, A., Schorderet, D.F. (2009) Mutations in CNNM4 cause recessive cone-rod dystrophy with amelogenesis imperfecta. *Am J Hum Genet.* **84(2)**, 259-65.

30. InterPro. DUF21 Sequence Scan. Available at: http://pfam.xfam.org/family/DUF21. Accessed: August 10, 2020.
31. Chua, M.D., Liou, C.H., Bogdan, A.C., Law, H.T., Yeh, K.M., Lin, J.C., Siu, L.K., and Guttman, J.A. (2018) Klebsiella pneumoniae disassembles host microtubules in lung epithelial cells. *Cellular microbiology*. e12977
32. Armitano, J., Redder, P., Guimarães, V.A., and Linder, P. (2016) An Essential Factor for High Mg(2+) Tolerance of Staphylococcus aureus. *Frontiers in Microbiology* **7**, 1888
33. Yang, M., Jensen, L.T., Gardner, A.J., and Culotta, V.C. (2005) Manganese toxicity and Saccharomyces cerevisiae Mam3p, a member of the ACDP (ancient conserved domain protein) family. *Biochem J.* **386**, 479-487
34. Bateman A (1997) The structure of a domain common to archaebacteria and the homocystinuria disease protein. *Trends Biochem Sci* **22**, 12-13
35. Ignoul, S., and Eggermont, J. (2005) CBS domains: structure, function, and pathology in human proteins. American journal of physiology. *Cell physiology*. **289**, C1369-1378
36. Corral-Rodríguez M.A., Stuiver M., Abascal-Palacios G., Diercks T., Oyenarte I., Ereño-Orbea J., Ibáñez de Opakua A., Blanco F.J., Encinar J.A., Spiwok V., et al. (2014) Nucleotide binding triggers a conformational change of the CBS module of the magnesium transporter CNNM2 from a twisted towards a flat structure. *Biochem. J.* 464, 23–34.
37. Gulerez, I., Funato, Y., Wu, H., Yang, M., Kozlov, G., Miki, H., and Gehring, K. (2016) Phosphocysteine in the PRL-CNNM pathway mediates magnesium homeostasis. *EMBO reports*. **17**, 1890-1900
38. Kostantin, E., Hardy, S., Valinsky, W.C., Kompatscher, A., de Baaij, J.H., Zolotarov, Y., Landry, M., Uetani, N., Martínez-Cruz, L.A., Hoenderop, J.G., Shrier, A., and Tremblay, M.L. (2016) Inhibition of PRL-2·CNNM3 Protein Complex Formation Decreases Breast Cancer Proliferation and Tumor Growth. *J Biol Chem*. **291**, 10716-10725
39. Arjona, F.J., and de Baaij, J.H.F. (2018) CrossTalk opposing view: CNNM proteins are not Na(+) /Mg(2+) exchangers but Mg(2+) transport regulators playing a central role in transepithelial Mg(2+) (re)absorption. *The Journal of physiology*. **596**, 747-750

40. Hanahan, D., and Weinberg, Robert A. (2011) Hallmarks of Cancer: The Next Generation. *Cell.* **144**, 646-674

41. Zhang, H., Kozlov, G., Li, X., Wu, H., Gulerez, I., and Gehring, K. (2017) PRL3 phosphatase active site is required for binding the putative magnesium transporter CNNM3. *Scientific Reports* **7(48)**, 1-9

42. Ereno-Orbea, J., Oyenarte, I., and Martínez-Cruz, L.A. (2013) CBS domains: ligand binding sites and conformational variability. *Arch. Biochem. Biophys* **540**, 70–81

43. Hattori, M., Tanaka, Y., Fukai, S., Ishitani, R., and Nureki, O. (2007) Crystal structure of the MgtE Mg^{2+} transporter. *Nature* **448**, 1072–1075

44. Hattori, M., Iwase, N., Furuya, N., Tanaka, Y., Tsukazaki, T., Ishitani, R., Maguire, M.E., Ito, K., Maturana, A., and Nureki, O. (2009) Mg^{2+} -dependent gating of bacterial MgtE channel underlies Mg^{2+} homeostasis. *EMBO J.* **28**, 3602–3612

45. Ishitani, R., Sugita, Y., Dohmae, N., Furuya, N., Hattori, M., and Nureki, O. (2008) Mg^{2+} -sensing mechanism of Mg^{2+} transporter MgtE probed by molecular dynamics study. *Proc. Natl. Acad. Sci. U.S.A* **105**, 15393–15398

46. Giménez-Mascarell, P., Oyenarte, I., Hardy, S., Breiderhoff, T., Stuiver, M., Kostantin, E., Diercks, T., Pey, A.L., Ereño-Orbea, J., Martínez-Chantar, M.L., *et al*. Structural Basis of the Oncogenic Interaction of Phosphatase PRL-1 with the Magnesium Transporter CNNM2. *J. Biol. Chem.* **292**, 786–801.

47. Clayton G. M., Silverman W. R., Heginbotham L., and Morais-Cabral J. H. (2004) Structural basis of ligand activation in a cyclic nucleotide regulated potassium channel. *Cell* **119**, 615–627

48. Smith, F.D., Esseltine, J.L., Nygren, P.J., Veesler, D., Byrne, D.P., Vonderach, M., *et al* (2017) Local protein kinase A action proceeds through intact holoenzymes. *Science* **356 (6344)**, 1288–1293.

49. Maguire, M.E. (2006) The Structure of the CorA Magnesium Transporter, a Divalent Cation Channel. *Curr Opin Struct Biol.* **4**, 432–438.

50. Wong, K., Kozlov, G., Zhang, Y., and Gehring, K. (2015) Structure of the Legionella effector, lpg1496, suggests a role in nucleotide metabolism. *J Biol Chem* **290**, 24727-24737

51. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293

52. Goddard, T., and Kneller, D. (2008) SPARKY 3. *University of California, San Francisco*

53. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J*. **78**, 1606-1619

54. Laue, T. (1993) Computer-aided interpretation of analtical sedimentation data for proteins. *Analytical ultracentrifugation in biochemistry & polymer science*, 90-125

55. Brookes, E., Demeler, B., and Rocco, M. (2010) Developments in the US-SOMO bead modeling suite: new features in the direct residue-to-bead method, improved grid routines, and influence of accessible surface area screening. *Macromol Biosci* **10**, 746-753

56. Brautigam, C.A. (2015) Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Methods Enzymol* **562**, 109-133

57. Wessel, D., and Flügge, U.I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids *Anal. Biochem.* **138** 141–143.

58. Hsueh, P.R., Lee, T.F., Du, S.H., Teng, S.H., Liao, C.H., Sheng, W.H., et al. Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometry system for identification of Nocardia, Rhodococcus, Kocuria, Gordonia, Tsukamurella, and Listeria species. *J Clin Microbiol* **52(7)**, 2371–2379.

59. McGuire, H., Aurousseau, M.R.P., Bowie, D., and Blunck, R. (2012) Automating Single Subunit Counting of Membrane Proteins in Mammalian Cells. *J Biol Chem.* **287(43)**, 35912–35921

60. Pessoa J., Fonseca F., Furini S., and Morais-Cabral J. H. (2014) Determinants of ligand selectivity in a cyclic nucleotide-regulated potassium channel. *J. Gen. Physiol.* **144**, 41–54

61. Lo M. C., Aulabaugh A., Jin G., Cowling R., Bard J., Malamas M., and Ellestad G. (2004) Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem*. **332**, 153–159

62. Mishra P., Singh U., Pandey C.M., Mishra P., and Pandey G. (2019) Application of student's t-test, analysis of variance, and covariance. *Ann Card Anaesth.* **22(4)**, 407-411.

63. Curtin F., and Schulz P. (1998) Multiple correlations and Bonferroni's correction. *Biol Psychiatry* **44(8)**, 775-7.

64. Yao, J., Dyson, H. J., and Wright, P.E. (1997) Chemical shift dispersion and secondary structure prediction in unfolded and partly folded proteins. *FEBS Letters* **419**, 285-289

65. Price, W.N., Chen,Y., Handelman, S.K., Neely, H., Manor, P., Karlin, R., *et al*. (2009) Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. *Nature Biotechnology* **27**, 51–57

66. Wagner, S., Bader, M.L., Drew, D., and de Gier, J.W. (2006) Rationalizing membrane protein overexpression. *Trends Biotechnol.* **24(8)**, 364-71.

67. Kyte, J., and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* **157(1)**, 105-32.

68. Brelidze, T. I., Carlson, A. E., Sankaran, B., and Zagotta, W. N. (2012) Structure of the carboxy-terminal region of a KCNH channel. *Nature* **481**, 530-533

69. Haitin, Y., Carlson, A. E., and Zagotta, W. N. (2013) The structural mechanism of KCNH-channel regulation by the eag domain. *Nature* **501**, 444-448

70. Feng, L., Campbell, E.B., Hsiung, Y., and MacKinnon, R. (2010). Structure of a eukaryotic CLC transporter defines an intermediate state in the transport cycle. *Science* **330**, 635–641.

71. Cole, J.L., Lary, J.W., Moody,T., and Laue, T.M. (2008) Analytical Ultracentrifugation: Sedimentation Velocity and Sedimentation Equilibrium. *Methods Cell Biol.* **84**, 143–179.

72. Grasierger, B., Minton, A.P., Delis, C., and Metzger, H. (1986) Interaction between proteins localized in membranes. *Proc. Natl. Acad. Sci. USA* **83**, 6258-6262.