# Expression and structural analysis of a A495K mutant of *Thermotoga maritima* membrane-bound pyrophosphatase

Ardit Fejzullahi

Master's thesis

University of Helsinki

January 2023

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

# Abstract

**Faculty**: Faculty of Agriculture and Forestry and Faculty of Biological and Environmental Sciences

**Degree programme**: Masters's Programme of Microbiology and Microbial Biotechnology

**Author**: Ardit Fejzullahi

**Title**: Expression and structural analysis of a A495K mutant of *Thermotoga maritima* membrane-bound pyrophosphatase

**Level**: Master's thesis

**Month and year**: January 2023

**Number of pages**: 47

**Keywords**: Membrane-bound pyrophosphatases*, Thermotoga maritima,* ion pumping, structural biology

**Supervisors:** Prof. Adrian Goldman and Dr. Keni Vidilaseris

**Where deposited**: The Helsinki University Library (https://ethesis.helsinki.fi/en/)

**Abstract**:

Membrane-bound pyrophosphatases (M-PPases) catalyse the reversible hydrolysis of pyrophosphate into two inorganic phosphate molecules. This hydrolysis is coupled to the transport of protons and/or sodium ions across the biological membrane, generating an electrochemical gradient, which can be utilized by the host organism as an energy source under different stress conditions. The essential physiological roles of M-PPases in agriculturally relevant plants and various human pathogens, such as *Bacteroides fragilis*, *Plasmodium falciparum*, *Toxoplasma gondii*, and *Trypanosoma brucei,* make them an important research area. Despite the biochemical studies conducted with several M-PPases and structural characterization of *Thermotoga maritima* and *Vigna radiata* M-PPases, there remain several questions regarding the biochemistry and catalytic mechanism of M-PPases, one being the basis of $K^+$ dependence. Mutational studies have suggested that a single residue at position 495 is the sole determinant of $K^+$ dependence in M-PPases. In $K^+$-dependent M-PPases this residue is an alanine and a $K^+$ ion binding in close proximity of it, has been shown of having stimulatory effect on $K^+$-dependent M-PPases by increasing the maximal rate of pyrophosphate hydrolysis. In $K^+$-independent M-PPases, a lysine residue resides at 495 position, where the ε-$NH3^+$ group of the lysine residue has been suggested, based on structural modelling and Ala → Lys (A495K) mutational studies, to structurally and functionally mimic the activating function of $K^+$ in $K^+$-dependent M-PPases. In order to provide structural basis to this, in this study, a M-PPase belonging to a hyperthermophilic deep-sea bacterium *T. maritima* containing a single A495K point mutation was expressed in *Saccharomyces cerevisiae*, solubilized using the high-temperature ("Hot-solve") method, purified with metal affinity chromatography and structurally characterized using X-ray crystallography. As suggested, the ε-$NH3^+$ group of the introduced lysine was shown to occupy the same space and form the same interactions at the active site as the $K^+$ ion in $K^+$-dependent M-PPases. However, these structural findings were in contradiction with the conducted hydrolytic activity assay which showed that the A495K mutation, besides abolishing $K^+$ dependence, significantly reduced the overall activity compared to the wild type by ~2-10-fold, depending on the $K^+$ concentration. These findings suggest that additional factor(s) besides the 495 residue determine the $K^+$ dependence in M-PPases.

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

# Tiivistelmä

**Tiedekunta**: Maatalous-metsätieteellinen tiedekunta ja Bio- ja ympäristötieteellinen tiedekunta

**Tutkinto-ohjelma:** Mikrobiologian ja mikrobibiotekniikan maisteriohjelma

**Tekijä**: Ardit Fejzullahi

**Työn nimi**: *Thermotoga maritiman* membraani-pyrofosfataasin A495K-mutantin ekspressio ja rakenteellinen karakterisointi

**Työn laji**: Maisterintutkielma

**Aika**: Tammikuu 2023

**Sivumäärä**: 47

**Avainsanat**: Membraani-pyrofosfataasit*, Thermotoga maritima,* ionien pumppaus, rakenteellinen biologia

**Ohjaajat:** Prof. Adrian Goldman and Dr. Keni Vidilaseris

**Säilytyspaikka**: Helsingin yliopiston kirjasto (https://ethesis.helsinki.fi/en/)

**Tiivistelmä**:

Membraani-pyrofosfataasit (M-PPaasit) katalysoivat pyrofosfaatin käänteisen hydrolyysin epäorgaanisiksi fosfaatti-ioneiksi. Tämä hydrolyysireaktio on kytkettynä protonien ja/tai natriumionien kuljetukseen solukalvon läpi luoden sähkökemiallisen gradientin, jota voidaan käyttää energialähteenä isäntäorganismin toimesta erilaisten stressiolosuhteiden aikana. M-PPaasien fysiologisesti merkittävät toiminnot maatalouden kannalta tärkeissä kasveissa sekä monissa ihmisten taudinaiheuttajissa, kuten *Bacteroides fragilis*, *Plasmodium falciparum*, *Toxoplasma gondii* ja *Trypanosoma brucei*, tekevät niistä tärkeän tutkimuskohteen. Lukuisiin M-PPaaseihin kohdistuneista biokemiallisista tutkimuksista sekä *Thermotoga maritima* ja *Vigna radiata* M-PPaasien rakenteiden karakterisoinneista riippumatta, M-PPaasien biokemiasta sekä katalyyttisestä mekanismista on edelleen avoimia kysymyksiä, joista yksi on $K^+$-ioniriippuvuus. Mutaatiotutkimukset ovat viitanneet siihen, että yksi aminohappotähde paikassa 495 on ainoa $K^+$-ioniriippuvuutta määrittelevä tekijä M-PPaaseissa. $K^+$-riippuvaisissa M-PPaaseissa tämä aminohappotähde on alaniini ja sen läheisyyteen sitoutuneella $K^+$-ionilla on todettu olevan tehostava vaikutus $K^+$-riippuvaisissa M-PPaaseissa. Sen on osoitettu kasvattavan pyrofosfaatin hydrolyysireaktion maksimaalista nopeutta. $K^+$-riippumattomissa M-PPaaseissa, paikassa 495 sijaitsee lysiini aminohappotähde, jonka ε-$NH3^+$-ryhmän on rakenteellisten mallinnusten sekä Ala → Lys (A495K) mutaatiotutkimusten perusteella ehdotettu jäljittelevän rakenteellisesti ja toiminnallisesti $K^+$ ionin aktivoivaa vaikutusta $K^+$-riippuvaisissa M-PPaaseissa. Antaakseen rakenteellista perustetta tälle väitteelle, tässä tutkimuksessa hypertermofiilisen syvänmeren bakteerin *T. maritima* M-PPaasi, joka sisältää yhden A495K pistemutaation, tuotettiin *Saccharomyces cerevisiae:ssa*, liuotettiin korkean lämpötilan (“Hot-solve”) menetelmällä, puhdistettiin metalli affiniteetti kromatografialla ja rakenteellisesti karakterisoitiin röntgenkristallografialla. Niin kuin on ehdotettu, lisätyn lysiinin ε-$NH3^+$-ryhmä todettiin asettuvan samaan kohtaan ja muodostavan samat vuorovaikutukset aktiivisessa keskuksessa kuin $K^+$-ioni $K^+$-riippuvaisissa M-PPaaseissa. Tästä huolimatta nämä rakenteelliset löydökset ovat ristiriidassa tehdyn hydrolyyttisen aktivisuusanalyysin kanssa, josta ilmeni A495K mutaation, $K^+$-riippuvuuden hävittämisen lisäksi, vähentävän kokonaisaktiivisuutta $K^+$ konsentraatiosta riippuen ~2-10-kertaisesti villityyppiin verrattuna. Nämä löydökset viittaavat siihen, että on olemassa jokin muu tekijä (tai muita tekijöitä), 495 aminohappotähteen lisäksi, joka määrittelee $K^+$-ioniriippuvuuden M-PPaaseissa.

# 1. Introduction

## 1.1 Pyrophosphatases

Inorganic pyrophosphatases (PPases) are enzymes that catalyse the reversible hydrolyses of pyrophosphate ($PP_i$) into two phosphate ($P_i$) molecules. First discovered in 1928 (Kay, 1928), they are found in all forms of life having a crucial role in controlling the cellular concentration of $PP_i$, which is formed in abundance as a by-product of at least 190 biosynthetic reactions, such as nucleic acid, polysaccharide, lipid, and protein synthesis. Thus, controlling the cellular $P_i/PP_i$ balance is essential for driving these vital biosynthetic reactions forward and for regulation of several enzymes, further highlighting the broad effect of PPases on cell metabolism (Holmes et al., 2019; Kajander et al., 2013; Shah et al., 2016).

Soluble PPases (S-PPases) are primarily responsible for regulating the intracellular $PP_i$ concentration and they are subdivided into two major evolutionary unrelated protein families, family I and family II. They are found in high abundance in all living organisms and have been shown to be crucial for life by being essential for normal cell growth and function (Chen et al., 1990; Shah et al., 2016). Membrane-bound PPases (M-PPases), another family of PPases, are yet very different from S-PPases, differing in sequence, structure, mechanism, and function to their soluble counterparts. Unlike the globular S-PPases, M-PPases are embedded in the biological membrane through 15-17 transmembrane helices (TMH). They act as primary ion-pumps by coupling the hydrolysis or synthesis of $PP_i$ into the translocation of $H^+$ and/or $Na^+$ across the membrane, generating an electrochemical gradient which can be utilized in several cellular functions, such as vacuole and acidocalcisome regulation (Holmes et al., 2019; Tsai et al., 2014). M-PPases differ functionally from S-PPases by having a minor role in controlling the intracellular $PP_i$ but broader biological functions by acting as primary ion-pumps and being crucial for cell survival under different stress conditions. In addition, M-PPases are found in most plants and protists, and in some species of bacteria and archaea, but not in multicellular animals unlike S-PPases (Baykov et al., 2013; Kajander et al., 2013).

## 1.2 Classification of membrane-bound pyrophosphatases

Phylogenetic analysis, sequence comparison and functional characterization of M-PPases have revealed three major M-PPase families: $Na^+$-regulated $H^+$-pumping, $Na^+$ and $K^+$-independent $H^+$-pumping, and $K^+$-dependent M-PPases (Artukka et al., 2018). M-PPases differ and can be alternatively classified based on their co-factor dependence and/or cation transport specificity. All M-PPases have specificity for pyrophosphate and dependence on $Mg^{2+}$ as an essential co-factor to

function. $K^+$-dependent M-PPases additionally require potassium ($K^+$) for maximal catalytic activity, and they are further divided into several subfamilies depending on their cation transport specificity: $H^+$-pumping ($H^+$-PPase), $Na^+$-pumping ($Na^+$-PPase), and Na+/H+ dual-pumping ($Na^+/H^+$-PPase) M-PPases (Artukka et al., 2018; Luoto et al., 2015). The $K^+$-dependent $H^+$-PPase subfamily consists of three lineages of $H^+$-PPases differing on the position of a semi-conserved glutamate, which is a crucial residue for regulating the ion transport specificity (Tsai et al., 2014). All three $K^+$-dependent $H^+$-PPase lineages are strictly $H^+$ transporters, whereas $K^+$-dependent $Na^+$-PPases can also transport $H^+$ under subphysiological $Na^+$ concentrations (>10 mM) (Nordbo et al., 2016). There are two $K^+$-dependent $Na^+/H^+$-PPase lineages, $Na^+$-regulated and true $Na^+/H^+$-PPases, which can transport both $Na^+$ and $H^+$ in a non-competitive manner under physiological conditions. These two lineages differ from each other based on their $H^+$-pumping activity in response to different $Na^+$ concentrations. (Artukka et al., 2018; Luoto et al., 2013; Nordbo et al., 2016). $Na^+$-PPases and $Na^+/H^+$-PPases require $Na^+$ for both $PP_i$ hydrolysis and cation transport, and they are further activated by $K^+$, acting as a non-essential activator, in contrast to $K^+$-dependent $H^+$-PPases, where $K^+$ is thought to be important for activity (Artukka et al., 2018). Phylogenetic and structural analysis suggest that $K^+$-dependent $H^+$-PPases and $Na^+/H^+$-PPases have been evolved independently from $K^+$-dependent $Na^+$-PPases through changes in transport specificity, and thus making the $K^+$-dependent $Na^+$-PPase subfamily a likely ancestral form of the $K^+$-dependent M-PPase family (Luoto et al., 2011; Luoto et al., 2013).

Unlike $K^+$-dependent M-PPases, $Na^+$-regulated, and $Na^+$ and $K^+$-independent $H^+$-PPase families do not require $Na^+$ or $K^+$ for their activity. Even though both are co-factor independent and monospecific $H^+$ transporters, they are separated by a long phylogenetic distance and differ from each other based on the effect $Na^+$ has on their $H^+$-pumping activity. $Na^+$-regulated $H^+$-PPases are thought to contain an inhibitory $Na^+$-binding site, which when saturated at high $Na^+$ concentrations, leads to the inhibition of $H^+$-PPase activity. High $K^+$ concentration has also been shown to inhibit $Na^+$-regulated $H^+$-PPases, a feature that is thought to result from weak binding of $K^+$ to this inhibitory $Na^+$-binding site. Thus, members of $Na^+$-regulated $H^+$-PPase family are alternatively called $Na^+$ and $K^+$-regulated $H^+$-PPases (Artukka et al., 2018; Luoto et al., 2015; Nordbo et al., 2016).

Overall, since the initial finding of the *Rhodospirillum rubrum* M-PPase in 1966 (Baltscheffsky et al., 1966), six types of M-PPases have been discovered: $Na^+$-regulated $H^+$-PPases, $Na^+$ and $K^+$-independent $H^+$-PPases, $K^+$-dependent $H^+$-PPases, $Na^+$-PPases, $Na^+$-regulated $Na^+/H^+$-PPases and true $Na^+/H^+$-PPases (Table 1.1) (Artukka et al., 2018). M-PPases capable of pumping $Na^+$ (including dual pumps) are exclusively found in prokaryotes, whereas $H^+$-PPases, regardless of cation

dependence, are also found in eukaryotes. These functional differences of M-PPases mentioned above reflect on their roles they have in their host organism (Shah et al., 2016).

Table 1.1: Types of M-PPases

| Type | Pumping specificity | Cation dependence | Examples |
|---|---|---|---|
| $Na^+$-PPase | $Na^+$ | $Na^+$ and $K^+$ | *Thermotoga maritima* (*Tm*-PPase) |
| True $Na^+/H^+$-PPase | $Na^+$ and $H^+$ | $Na^+$ and $K^+$ | *Bacteroides vulgatus* (*Bv*-PPase) |
| $Na^+$-regulated $Na^+/H^+$-PPase | $Na^+$ and $H^+$ | $Na^+$ and $K^+$ | *Clostridium leptum* (*Cl*-PPase) |
| $K^+$-dependent $H^+$-PPase | $H^+$ | $K^+$ | *Vigna radiata* (*Vr*-PPase) |
| $Na^+$ and $K^+$-independent $H^+$-PPase | $H^+$ | - | *Geobacter sulfurreducens* (*Gs*-PPase) |
| $Na^+$-regulated $H^+$-PPase | $H^+$ | - | *Chlorobium limicola* (*Cl*(2)-PPase) |

## 1.3 Physiological roles of M-PPases in plants, bacteria, and protozoan parasites

Unlike in the case of S-PPases, where the energy stored in a phosphoanhydride bond is lost as heat after $PP_i$ hydrolysis, M-PPases preserve part of this energy by generating an electrochemical gradient across the membrane through translocation of $H^+$ and/or $Na^+$ ions (Primo et al., 2019). This membrane potential energizes secondary transporters to carry out a wide variety of cellular functions such as maintaining ion homeostasis and transporting different metabolites against their concentration gradient. Found in plants, bacteria, and protozoan parasites, M-PPases have been shown to be vital for cell survival under stress conditions and plant maturation (Holmes et al., 2019; Vinothkumar & Henderson, 2010).

Plant $H^+$-PPases can be divided into two distinct subclasses based on their $K^+$ dependence: type I being dependent of $K^+$ on reaching maximal activity, and type II being insensitive to $K^+$. Type II $H^+$-PPases are found exclusively in the Golgi apparatus and their total amount accounts for less than 0.3% of that the type I $H^+$-PPase amount. Indeed, the majority of the research has focused on the type I $H^+$-PPases, which are found on the membrane surrounding the vacuole (tonoplast) and are therefore also known as vacuolar PPases (V-PPases) (Segami et al., 2018; Segami et al., 2010). Accounting for 10% of the total proteins in the vacuole, V-PPases have two functions in plant cells. First is the removal of excess $PP_i$, which consequently prevents the toxic build-up of cytosolic $PP_i$, and the second is the acidification of the vacuole in co-operation with vacuolar $H^+$-ATPase (V-ATPase). This dual function of V-PPases plays important roles in stress resistance and plant development (Segami

et al., 2018). All vacuolar functions, such as osmoregulation and protein storage, rely on the ability to transport massive fluxes of molecules in and out of the vacuole. V-PPase, together with V-ATPase, maintain the gradient of electrochemical difference for $H^+$ across the tonoplast, providing the necessary acidic conditions to hydrolytic enzymes and energy for the secondary active transport of various molecules, such as sugars, nutrients, ions, and harmful toxins. V-PPase activity is usually high in young growing plant tissues, where ATP availability is low due to high energy demand of dividing cells, and $PP_i$ concentration high, due to the increased macromolecule biosynthesis (Ferjani et al., 2011; Martinoia et al., 2007; Schilling et al., 2017). Besides this, genes encoding V-PPases are upregulated by abiotic and biotic stresses, such as drought (Gaxiola et al., 2001; Wang et al., 2009; Wehner et al., 2015), low nutrient availability (Yang et al., 2007; Yang et al., 2014), coldness (Carystinos et al., 1995), and disease (Hernández et al., 2016; Yang et al., 2020). Furthermore, over-expression of V-PPase in various plant species, including several crops, has been shown to enhance salt and drought tolerance compared to the control or knock-out mutants (Fuglsang et al., 2011; Gaxiola et al., 2012). It is generally thought that the stress-tolerance role of V-PPase is to compensate for the decreased V-ATPase activity under low ATP conditions. However, V-PPase has shown to also induce plant growth under nonstressed conditions (Holmes et al., 2019). Removal of cytosolic $PP_i$ and transport of plant hormone, auxin, are attributed to the developmental roles of V-PPase (Asaoka et al., 2016; Li et al., 2005). $PP_i$ can be toxic at high levels by inhibiting $PP_i$-dependent pathways, such as gluconeogenesis, and by forming insoluble complexes with $Ca^{2+}$ or $Mg^{2+}$ which can be cell-damaging. Thus, V-PPase-mediated removal of cytosolic $PP_i$ is important for plant maturation and growth (Ferjani et al., 2011; Segami et al., 2018).

Bacterial M-PPases are usually present in certain species that live in harsh and low-energy environmental conditions, such as obligate anaerobe *Bacteroides* species and deep-sea bacteria (e.g., *Thermotoga maritima*). Unlike in plants, all M-PPase types can be found in prokaryotes, the only group of organisms where $Na^+$-pumping M-PPases can be found (Holmes et al., 2019). Located in cell membranes, bacterial M-PPases have been suggested to cover for the energy requirement of the cell during low-energy and stress conditions by generating ion electrochemical gradient across the membrane, that in turn can be used for ATP synthesis and other essential transport processes (Serrano et al., 2007). Increased $H^+$-PPase activity was observed in *R. rubrum* under acute salt stress in aerobic conditions (López-Marqués et al., 2004). In addition, loss-of-function $H^+$-PPase showed impaired growth under low oxygen stress and in low-light conditions compared to the wild-type in *R. rubrum*, further highlighting the importance of $H^+$-PPase activity under low-energy conditions (García-Contreras et al., 2004). The $K^+$-dependent $Na^+$-PPase of *T. maritima* (*Tm*-PPase) has been solved and

used as a model in structure-function and biochemical studies of M-PPases (Kellosalo et al., 2012; Li et al., 2016; Vidilaseris et al., 2019), mainly due to its ability to withstand high temperatures without being denatured or deactivated. This high thermostability allows to isolate the protein in high-purity using high-temperature solubilization and to conduct reliable activity assays at room temperature (Kellosalo et al., 2011; López-Marqués et al., 2005). *T. maritima* is a hyperthermophilic bacterium that can be found in geothermally heated undersea vents and sea floors, with an optimum growth temperature of 80 °C (Huber et al., 1986). Even though the role of *Tm*-PPase in *T. maritima* has not been studied, it most likely has to do with the use of $PP_i$ as an alternative energy source under harsh, low-energy conditions experienced in geothermally heated deep-sea environments.

In several protozoan parasites (e.g., *Plasmodium falciparum*, *Trypanosoma brucei*, *Toxoplasma gondii, Leishmania donovani, and Leptomonas wallacei*), both $K^+$-dependent and -independent $H^+$-PPase variants can be found in the cell membrane and the Golgi apparatus, but primarily in the membrane of the acidocalcisome (Docampo & Moreno, 2008; Marchesini et al., 2000). Acidocalcisomes are dense acidic organelles serving various vital cellular functions in parasites, including storing cations and phosphorous compounds, and regulating intracellular pH, calcium homeostasis, osmotic pressure, and polyphosphate metabolism. $H^+$-PPases together with $H^+$-ATPases acidify and energize the acidocalcisome via $H^+$ translocation (Docampo et al., 2013). One study on $H^+$-PPase knockout strains of trypanosomiasis-causing *T. brucei* observed significantly slower growth rate and reduced cell density compared to the wild-type strains Loss of functional acidocalcisomes, including a 10-fold drop of stored polyphosphate, and impaired $Ca^{2+}$ transport and intracellular pH regulation, was observed (Lemercier et al., 2002). $Ca^{2+}$ transport together with polyphosphates are thought to be important for osmoregulation, which is vital for parasites due to the changing conditions (i.e., pH and osmolarity) during transitions between vectors and hosts. $Ca^{2+}$ is also an important intracellular signalling molecule in protozoan parasites and essential for invasion of host cell. Most of $Ca^{2+}$ in acidocalcisomes is bound to polyphosphates and it can be released under hypoosmotic conditions through hydrolysis of polyphosphates resulting in an increase in intracellular ionic strength (Docampo et al., 2013; Lemercier et al., 2002). $H^+$-PPase knock-out studies in human pathogens, *T. gondii* and *P. falciparum,* also showed impaired acidocalcisome acidification and loss of infectivity, respectively (Liu et al., 2014; Zhang et al., 2018).

## 1.4 The relevance of M-PPases in global food security and human disease prevention

The roles of $H^+$-PPases in plant stress-tolerance and development makes them agriculturally relevant targets for combatting increasing climate change-induced crop yield losses and global food security

problems. Manipulation of these proteins, such as overexpression of V-PPase genes in plants, has shown improvements in plant survival and growth under abiotic and biotic stress conditions (Gaxiola et al., 2001; Li et al., 2005; Shah et al., 2016; Yang et al., 2014).

M-PPases are also clinically relevant antiparasitic drug targets as they promote survival, energy efficiency and infectivity in several human parasitic pathogens causing severe diseases with high prevalence and fatality rates, such as trypanosomiasis (*Trypanosome* spp.), toxoplasmosis (*Toxoplasma gondii*), malaria (*Plasmodium* spp.), and leishmaniasis (*Leishmania* spp.) (Lemercier et al., 2002; Liu et al., 2014; Zhang et al., 2018). Expected northwards spread of these tropical diseases due to climate change, combined with rising resistance against many current antiparasitic drugs, makes treatment of these diseases a serious challenge, demanding the development of novel antiparasitic therapeutics (Mora et al., 2022; Shah et al., 2016). Bacterial M-PPases are also relevant drug-targets as they are suggested to promote survivability during low-energy and stress conditions in several human bacterial pathogens belonging to *Clostridium* spp. (e.g., *C. tetani and C. botulinum*) and *Bacteroides* spp. (e.g., *B. fragilis*) (Luoto et al., 2011; Shah et al., 2016). Despite playing an important role in human health as commensal micro-organisms, *Bacteroides* spp. are associated with invasive infections, such as bacteremia and complicated intra-abdominal, skin, and soft tissue infections. *B. fragilis* is the main culprit of these infections and one of the most antimicrobial-resistant anaerobic pathogen groups, with associated mortality rate of up to 60%, depending on the treatment (Wexler, 2007). High antibiotic resistance and emerging multi-drug resistant strains associated with high mortality rates make *Bacteroides* spp. a major health threat (Aires, 2016).

Overall, the essential physiological roles of M-PPases in agriculturally relevant plants and various human pathogens makes them an important area of research. As there are no human M-PPase analogues, they are validated drug targets against various human pathogens, and targets for modification in plants to improve tolerance against abiotic and biotic stressors (Holmes et al., 2019).

## 1.5 Structural overview of M-PPases

All M-PPases are formed into homodimers that consist of 70-81-kDa subunits, with each being arranged from 15-17 transmembrane helices (TMH) (Figure 1.1A). These TMHs split into two concentric ring structures, the inner ring (helices 5-6, 11-12, and 15-16) forming the functional core of the enzyme, and the outer ring (helices 1-4, 7-10, and 13-14) interacting with the membrane and forming the main dimeric interface (Figure 1.1B). This interface is maintained by the hydrophobic and hydrogen bond interactions between the residues of helices 10 and 15. The gated ion transport

channel formed by the inner ring TMHs contains four functional core regions: the active site, coupling funnel, ion gate and exit channel (Holmes et al., 2019; Kellosalo et al., 2012; Li et al., 2016).

Residues mentioned in following sections are numbered using the Ballesteros-Weinstein numbering system (Ballesteros & Weinstein, 1995), e.g., D5.65, where the letter corresponds to the residue in question and the number defines its location. The first number indicates the TMH on which the residue is located, and the second number defines its offset from a well conserved residue close to the centre of this TMH (assigned at position 50) (Tsai et al., 2014).

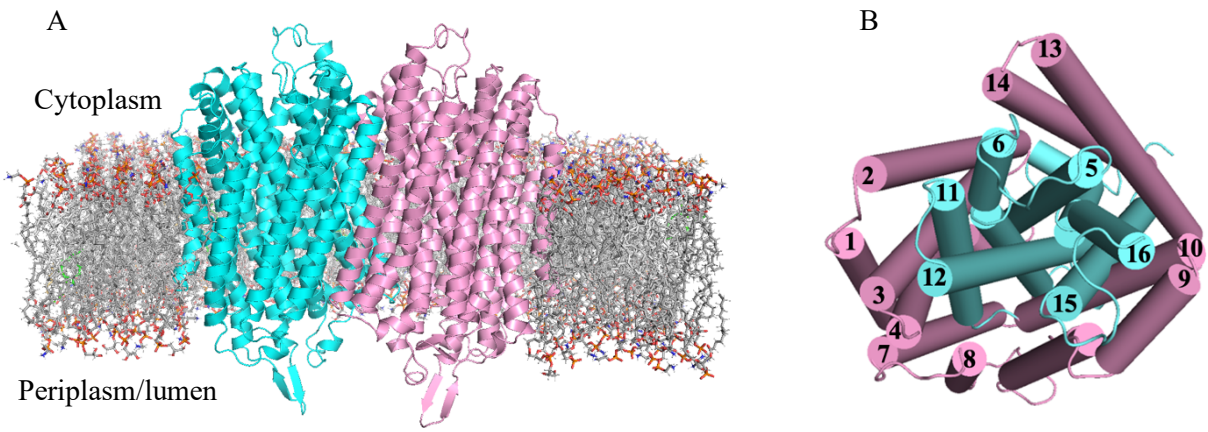


Figure 1.1: Overview of the M-PPase structure. (A) Membrane plane view of the homodimeric M-PPase embedded in the lipid bilayer. The two M-PPase subunits colored in cyan and pink. (B) The concentric ring arrangement of the 16 transmembrane helices in a M-PPase monomer. Viewed from the cytoplasmic side, the outer ring helices are colored in pink and the inner ones in cyan. PyMol (Schrödinger, LLC, 2015), *Tm*-PPase:$Mg_5$IDP structure (PDB: 5LZQ) (Li et al., 2016), and CHARMM membrane builder GUI (Jo et al., 2008; Pogozheva et al., 2022) were used in generating the figures.

## 1.6 Functional core regions of M-PPases

The active site protrudes towards the cytoplasmic side about 20 Å above the membrane surface. This funnel-shaped region is lined up with evolutionarily conserved residues that are involved in the binding of the pyrophosphate-substrate. These include 12 aspartates, three lysines, and one asparagine, which bind to the substrate directly or indirectly through coordination with up to five magnesium ions and a water molecule (Figure 1.2). These binding interactions precisely position the pyrophosphate-substrate at the active site for the hydrolysis, which is initiated by the nucleophilic water molecule, activated by two aspartates (D6.43 and D16.39, Ballesteros-Weinstein numbering system, see above) upon substrate-binding (Kellosalo et al., 2012; Shah et al., 2016). Substrate-binding also induces the closure of the active site cavity from the cytoplasmic side through reordering of the loops between TMHs 5 and 6, and TMHs 13 and 14, which ensures directional pumping

(Kellosalo et al., 2012; Lin et al., 2012). The cationic centre, located at the active site just below the substrate-binding site, is occupied by either $K^+$ or the ε-$NH_3^+$ group of a lysine residue (K12.46) in $K^+$-dependent or $K^+$-independent M-PPases respectively. The cationic centre is in direct contact with the $PP_i$ molecule and has a stimulatory effect on M-PPase activity, including a role in inter-subunit communication observed in substrate inhibition (Artukka et al., 2018; Malinen et al., 2007).

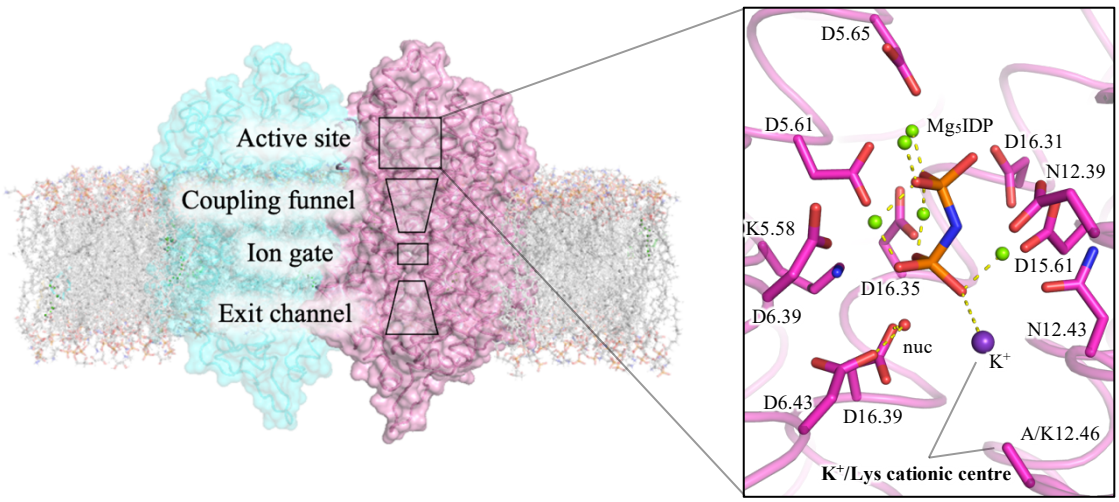


Figure 1.2: Functional sites of M-PPases focusing on the active site. Close-up figure of the active site in a substrate-bound state (*Vr*-PPase:$Mg_5$IDP, PDB: 4A01) (Lin et al., 2012), where the imidodiphosphate (IDP) substrate analogue is coordinated by 5 magnesium ions (green spheres), potassium ion (purple sphere) and the surrounding residues. Nucleophilic water molecule (nuc, red sphere) is coordinated and activated by D6.43 and D16.39 for the initiation of the hydrolysis. As in the case of $K^+$-dependent M-PPases, the cationic center is occupied by the $K^+$ and an alanine is located at the 12.46 position (labelled in bold). Dashed yellow lines indicate metal ion coordination interactions. Helix 11 has been removed for clarity.

The active site is followed by the coupling funnel, which forms an ion translocation channel between the active site and the ion gate, comprising of eight highly conserved charged residues belonging to five TMHs (Kellosalo et al., 2012) (Figure 1.3). The ion gate region is located at the centre of the protein, and it is made up of several conserved charged and hydrophilic residues. In the resting state, the gate itself consists of an ion pair (E6.57-K16.50) or an ion triplet (D6.50-K16.50-E6.53) in $K^+$-dependent $H^+$-PPases and $K^+$-dependent $Na^+$-PPases, respectively (Li et al., 2016). Substrate-binding and subsequent conformational changes in TMHs 6 and 16 lead to reorientation of these residues. In $K^+$-dependent $Na^+$-PPases, K16.50 residue swings out of the $Na^+$-binding pocket allowing the $Na^+$ to bind. D/N16.46 residue moves inward towards the D6.50, E6.53, and S6.54 gate residues, which together coordinate the bound $Na^+$ (Figure 1.3, shown in brackets). In $K^+$-dependent $H^+$-PPases, in the substrate-bound state, the E6.57-K16.50 ion pair is disrupted with E6.57 most likely being protonated, and K16.50 neutralized by the surrounding gate residues D6.50, S6.54, and N16.46.

K16.50 forms ion bridges with these residues, coordinating a nearby water molecule, which may be involved in the transport of protons from the attacking water nucleophile at the active site to the E6.57 at the ion gate via water wire (Grotthuss chain) (Kellosalo et al., 2012; Li et al., 2016; Lin et al., 2012). The position of the semi-conserved E6.53/57 is thought to be the main determinant of cation-pumping specificity, where the movement of E6.53 to E6.57 by one helical turn destroys the $Na^+$-binding site, converting $H^+$-PPase to $Na^+$-PPase (Li et al., 2016; Luoto et al., 2011). The $Na^+$ coordinated at the $Na^+$-binding site or the proton localized in E6.57 is moved towards the exit channel (Figure 1.4), which connects the ion gate to the cytosol or periplasm.

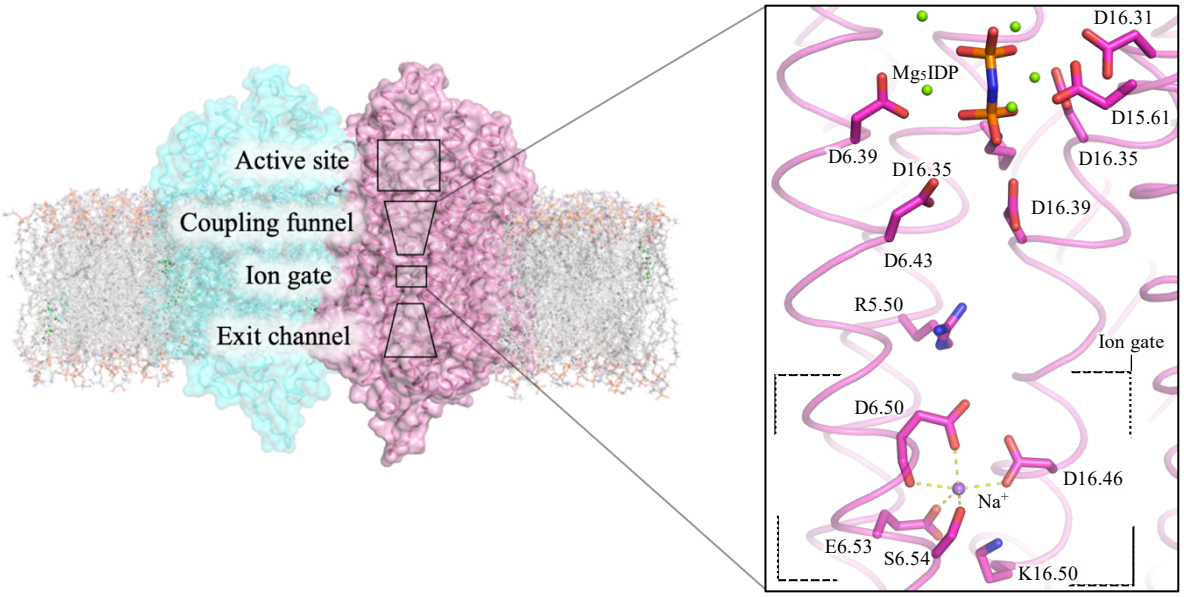


Figure 1.3: Functional sites of M-PPases focusing on the coupling funnel and ion gate. Close-up figure of the coupling funnel and ion gate (highlighted with brackets) in a substrate-bound state (*Tm*-PPase:$Mg_5IDP$, PDB: 5LZQ) (Li et al., 2016), where the key residues are displayed. Sodium ion (purple sphere) is coordinated at the ion gate by the surrounding D6.50, E6.53, S6.54 and D16.46 gate residues. Helices 7-9, and 11-12 removed for clarity.

Just above the exit channel, the hydrophobic gate is located, which is thought to prevent the cation backflow from the lumen/periplasm into the cytoplasm. High hydrophobicity and small pore diameter allow hydrophobic gate to act as a checkpoint for cation release (Figure 1.4A) (Tsai et al., 2019). A study showed that a non-phosphorous inhibitor, N-[(2-amino-6-benzothiazolyl)methyl]-1*H*-indole-2-carboxamide (ATC), binds near the exit channel of a *Tm*-PPase subunit as a dimer (Figure 1.4B) and inhibits catalytic activity in allosteric manner (Vidilaseris et al., 2019). Structural asymmetry between the subunits was further shown as the ATC-bound subunit being unable to perform hydrolysis, restricted the motions and thus catalytic activity of the unbound subunit (Vidilaseris et al., 2019). Characteristics of the hydrophobic gate, such as sequence diversity in hydrophobic gate residues and

cation pore size might be associated with cation specificity (Tsai et al., 2019). The comprehension of cation-binding and -pumping specificity is limited on the structural data of $K^+$-dependent $Na^+$-PPases and $H^+$-PPases, since, as of writing this, no structures of $K^+$-independent $H^+$-PPase nor $K^+$-dependent $Na^+/H^+$-PPase have been solved.

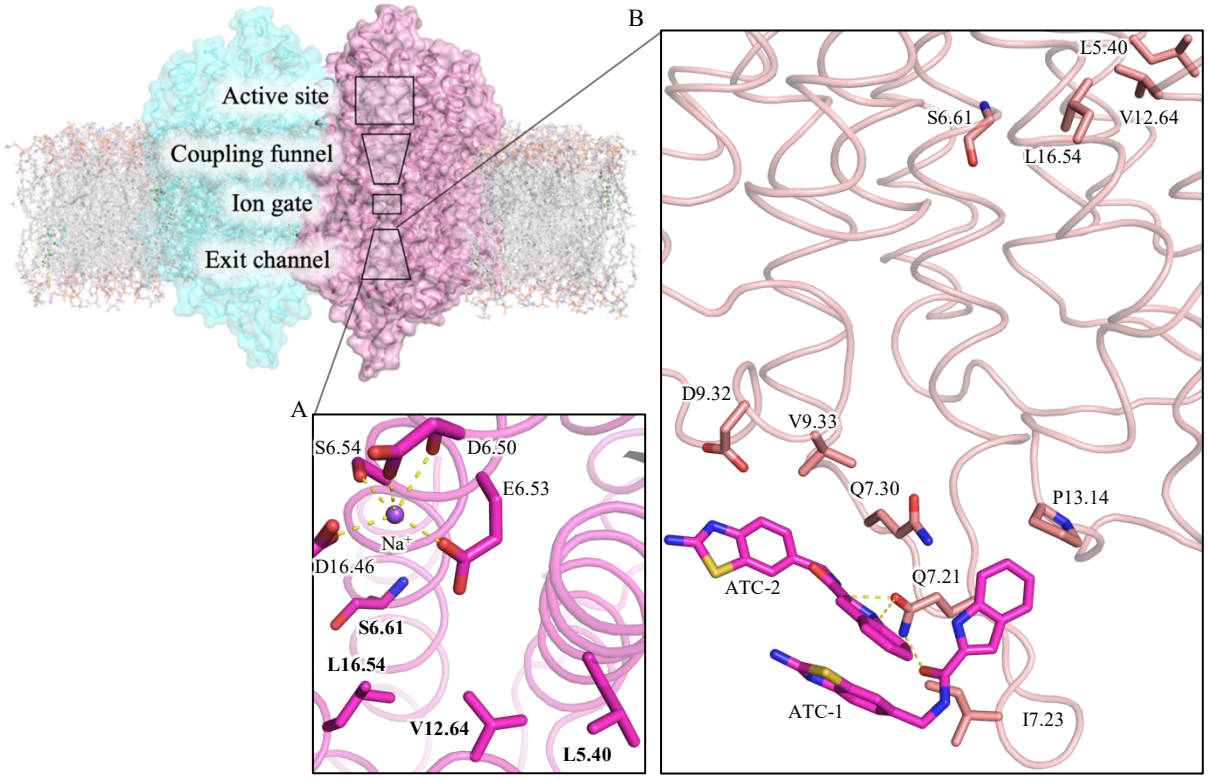


Figure 1.4: Functional sites of M-PPases focusing on exit channel. (A) Close-up figure of hydrophobic gate located below the ion gate and just above of the exit channel in a substrate-bound *Tm*-PPase (PDB: 5LZQ) (Li et al., 2016), where the key residues are displayed, including the hydrophobic gate itself consisting of S6.61, L16.64, V12.64 and L5.40 (labelled in bold). (B) Figure of two ATC inhibitor molecules bound to the end of the exit channel of a substrate-bound *Tm*-PPase (PDB: 6QXA) (Vidilaseris et al., 2019). Interactions are shown by dashed lines.

## 1.7 Cation translocation mechanism

So far, the structures of the $K^+$-dependent $Na^+$-pumping *Thermotoga maritima* M-PPase (*Tm*-PPase) and $H^+$-pumping *Vigna radiata* M-PPase (*Vr*-PPase) have been solved in several conformations by X-ray crystallography (Kellosalo et al., 2012; Li et al., 2016; Lin et al., 2012; Tsai et al., 2019; Vidilaseris et al., 2019). Although these have provided detailed structure-function insights into M-PPases, there remain open questions, such as the basis of $K^+$-dependence, and the mechanism of cation translocation and selectivity.

Structural studies of M-PPases have suggested that conformational changes induced by substrate-binding lead to the coordination of the nucleophilic water at the active site and, additionally, are related to further changes that promote cation pumping. However, the chronological order of the substrate hydrolysis and cation pumping cannot be deduced based on the currently available structural data of M-PPases captured in different catalytic states (Holmes et al., 2019). In the transport-after-hydrolysis mechanism, the conformational changes upon substrate binding leads to the activation of the nucleophilic water and substrate hydrolysis. Further structural changes induced by the hydrolysis lead to the pumping of the nucleophilic water proton or pushing of the $Na^+$ forward from the ion gate into the exit channel for pumping via a "billiard-type" mechanism (Baykov et al., 2013; Holmes et al., 2019). In contrast, in the transport-before-hydrolysis mechanism, cation pumping occurs first and it is triggered by the closure of the active site upon substrate-binding. The release of the cation from the ion gate increases the overall negative charge of the region, leading to the downward movement of TMH 12 to its original position prior of substrate-binding, and to the deprotonation of the aspartate pair D6.43 and D16.39 (Figure 1.2), which in turn activates the nucleophilic water at the active site leading to substrate hydrolysis. In the transport-before-hydrolysis mechanism, the pumped cation originates from the surrounding medium or from previous hydrolysis events (Holmes et al., 2019; Li et al., 2016). In the future, solving of M-PPase structures at a higher resolution combined with biochemical data may help to resolve these disputes surrounding the mechanism by which the cations are pumped.

## 1.8 $K^+$ dependence and the role of $K^+$/Lys cationic centre

Sequence comparisons between $H^+$-PPases revealed two residues involved in $K^+$ dependence. In $K^+$-dependent $H^+$-PPases, alanine and glycine/alanine are conserved at positions 12.46 and 12.49, respectively, whereas in $K^+$-independent $H^+$-PPases these equivalent positions are occupied by a lysine at 12.46 and a threonine at 12.49. Thus, these A12.46K and G/A12.49T substitutions were determined to be involved in the divergence of $K^+$-dependent and $K^+$-independent families (Artukka et al., 2018; Belogurov & Lahti, 2002). To investigate the role of these residues in $K^+$ requirement, site-directed mutagenesis experiments on $K^+$-dependent $H^+$-PPase of *Carboxydothermus hydrogenoformans* were conducted, which revealed that the Ala → Lys substitution at 12.46 abolishes the $K^+$ dependence. The Gly/Ala → Thr substitution at 12.49 in the other hand did not generate a $K^+$-independent protein, but instead reduced the affinity for $K^+$ by 3-fold (Belogurov & Lahti, 2002). Further systematic analysis of the cationic centre confirmed the A/K12.46 residue as the key determinant of $K^+$ dependence in M-PPases. In $K^+$-dependent M-PPases, this residue is alanine, and the cationic centre is occupied by $K^+$ (Figure 1.2). In $K^+$-independent M-PPases, this position is

occupied by the ε-$NH3^+$ group of the lysine residue, which is suggested to functionally and structurally mimic the role of $K^+$ in substrate coordination (Figure 1.2) (Artukka et al., 2018; Kellosalo et al., 2012; Lin et al., 2012). As previously mentioned, $K^+$ has a stimulatory effect on the activity of $K^+$-dependent M-PPases by increasing the maximal rate of PPi hydrolysis in all $K^+$-dependent M-PPases and binding affinity of $Na^+$ in $Na^+$-pumping M-PPases. Mutational experiments of the $K^+$/Lys cationic centre across all M-PPase subfamilies revealed that the lysine is a non-essential activator in $K^+$-independent M-PPases having similar activating function as $K^+$ in $K^+$-dependent M-PPases (Artukka et al., 2018). Furthermore, the native state of $K^+$/Lys cationic centre has been shown to be crucial in inter-subunit communication when it comes to substrate inhibition, as it was abolished by the Lys → Ala substitution in $K^+$-independent M-PPases and it was observed in $K^+$-dependent M-PPases only in the presence of $K^+$. Although the Ala → Lys substituted $K^+$-dependent M-PPases facilitated $K^+$-independence, they did not confer substrate inhibition, except one enzyme, where the introduced Lys was speculated to fulfil the structural requirements imposed by inter-subunit communication regarding substrate inhibition. Consequently, this suggests that the Lys and $K^+$ cannot perfectly replace each other, and the degree of structural and functional compensation provided by introduced Lys or $K^+$ cationic groups in Ala → Lys or Lys → Ala variants, respectively, depends on the M-PPase enzyme (Artukka et al., 2018).

Besides mutational and biochemical studies, structural modelling also suggests that the ε -$NH_3^+$ group of the introduced Lys occupies the same region at the active site as the $K^+$, therefore mimicking its activating functions (Kellosalo et al., 2012). Despite this, the structural basis of this claim is not certain yet since there are no experimentally solved structures of $K^+$-independent M-PPases, which could reveal the mechanism of $K^+$ dependence and provide additional basis for previous studies showing a single Ala → Lys substitution at 12.46 being able to abolish $K^+$ sensitivity and generate a $K^+$-independent M-PPase (Artukka et al., 2018; Belogurov & Lahti, 2002). Thus, the purpose of this thesis was to produce and characterize an Ala → Lys mutated M-PPase (A495K *Tm*-PPase) and to provide structural experimental evidence in support of the lysine at K12.46 (K495) being able to mimic the functions of $K^+$ in $K^+$-independent M-PPases. Besides $K^+$ dependence, a new higher resolution (near-atomic) structure could provide a more comprehensive understanding on the previously mentioned properties of M-PPases, such as the cation translocation mechanism.

# 2. Methods

## 2.1 Yeast culturing, transformation, and protein expression

*Saccharomyces cerevisiae* strain BJ1991 (genotype: *MATα rb1-1122 pep4-3 leu2 trp1 ura3-52 gal2*) and His-tagged A495K *Thermotoga maritima* membrane-bound pyrophosphatase (*Tm*-PPase) containing plasmid pRS1024 were provided by Dr. Keni Vidilaseris of Prof. Adrian Goldman´s group. Briefly, the N-terminally His-tagged *Tm*-PPase had been cloned into the XhoI/SpeI-cloning site of the multicopy pRS1024 vector under the control of the PMA1 promoter. Mutagenesis had been performed using inverse PCR after which the mutation had been confirmed through sequencing. Pre-culture of *S. cerevisiae* strain BJ1991 was made by inoculating a loopful of yeast cells from previously streaked 1x yeast peptone (YP) agar plates (1% yeast extract, 2% Bacto peptone, 2 % Bacto agar, 2% glucose and 100 µg/ml ampicillin) into prewarmed 7.5 ml of YP solution containing 1.5x YP, 2.7 % glucose and 100 µg/ml ampicillin. The cells were left to grow over night at 30°C with shaking (200 rpm) before transformation.

The leucine-selectable A495K *Tm*-PPase plasmid pRS1024 was transformed into yeast cells using the lithium acetate/polyethylene glycol (PEG) method (Schiestl & Gietz, 1989), by first centrifuging 1 ml of overnight yeast culture (3000 rpm, 10 sec) and resuspending the pellets in 150 µl solution containing 40% polyethylene glycol (PEG) 3350, 0.1 M lithium acetate, and 1x TE buffer (pH 7.5). 10 µg of freshly denatured fish sperm carrier DNA and 500 ng of pRS1024 were then added and vortexed thoroughly. After this, 10 µl of dimethyl sulfoxide (DMSO) was added and vortexed before incubating the transformation samples at 28 °C for 15 minutes and further heat-shocking them at 42 °C for 20 minutes. Samples were then centrifuged at 16,000 x g for 30 seconds. Cell pellets were gently resuspended in 100 µl of 1x TE (pH 7.5) before growing the transformants for 2 days at 30 °C on selective synthetic complete drop-out (SCD) medium plates containing 2% Bacto agar, 2% glucose, 100 µg/ml ampicillin, 0.67% yeast nitrogen base (YNB), and appropriate amino acids except leucine (SCD-Leu) to maintain plasmid selection. The *Tm*-PPase-pRS1024-containing cells were restreaked on SCD-Leu plates and grown for 2 days at 30 °C to obtain pure isolates. The cells (10 loopfuls/flask) were first precultured in 2 x 850 ml of SCD-Leu medium (2% glucose, 0.67% YNB, 100 µg/ml ampicillin, and supplemented with the appropriate amino acids without leucine) for 24h in 1 l Erlenmeyer flasks at 30 °C with 220 rpm shaking. This preculture was evenly transferred to 7 x 740 ml of prewarmed 1.5 x concentrated YP medium, containing 2.7% glucose and 100 µg/ml ampicillin, and further grown for 8 h in 2 l Erlenmeyer flasks with 200 rpm shaking at 30 °C before collection. Cells were collected by centrifugation at 5000 rpm (Sorvall centrifuge) for 10 min and the

cell pellet was subsequently washed with around 400 ml of ddH2O before resuspending it in 300 ml of $ddH_2O$. After resuspension, cells were transferred to 50 ml falcon tubes and pelleted by centrifugation at 4000 rpm (Sorvall centrifuge) for 3 min. Cells were frozen with liquid nitrogen and stored at -4 °C for later use.

## 2.2 Membrane extraction

The frozen cell pellet was thawed and resuspended in 140 ml of buffer A (1 ml per 2g of cell pellet) containing 0.2 M Tris (pH 7.5), 40% (v/v) glycerol, 10 mM ethylenediamine tetra acetic acid (EDTA, pH 7.5), 2 mM dithiothreitol (DTT), and 1 mM phenylmethylsulfonylfluoride (PMSF; Sigma). Cell suspension was poured to the lysis chamber of a bead beater (BeadBeater, BioSpec Products) with 165 g of cold 0.5-mm glass beads and the outside chamber was filled with ice and ice-cold water. The cells were lysed under ice by shaking 25 times for 30 seconds, interspersed with 1.5-minute cooling periods between shakings. The lysate was transferred to 50 ml falcon tubes and the glass beads were washed with 2 x 25 ml of buffer B [10 mM Tris (pH 7.5), 10% (v/v) glycerol, 5 mM EDTA (pH 7.5) and 1 mM DTT]. Tubes were filled to 50 ml with buffer B and KCl (0.329 M) and mixed before centrifuging at 4000 rpm for 10 minutes at 4 °C to separate cell debris and glass beads. The membrane fraction was collected by ultracentrifuging the supernatant at 42500 rpm (Beckman, Ti45 rotor) for 45 min. The pelleted membranes were then thoroughly resuspended in 4 ml of buffer C [50 mM 2-(N-Morpholino)ethanesulphonic acid (MES)-NaOH (pH 6.5), 20% (v/v) glycerol, 5.2 mM $MgCl_2$, 50 mM KCl, 1.33 mM DTT, 1.68 mM PMSF, 2 µg/ml pepstatin A (Sigma)]. Resuspended membranes were further homogenized with a Dounce homogenizer before transferring the membranes to 15 ml falcon tubes and freezing them with liquid nitrogen for storage at -80°C.

## 2.3 Protein solubilization

A modified version of the “hot-solve”-purification method (López-Marqués et al., 2005) was used to solubilize and purify the A495K *Tm*-PPase protein. For detergent solubilization and subsequent activity assay, the total protein concentration of extracted membranes was diluted in buffer C to 7.2 mg/ml into 300 µl fractions in 1.5 ml Eppendorf tubes. 1 µl of 0.1M $Na_2PP_i$ was immediately added to the diluted membranes. Solubilization buffer consisting of 50 mM MES-NaOH (pH 6.5), 20% (v/v) glycerol, and either 5.34% (w/v) n-dodecyl-b-D-maltopyranoside (DDM; Anatrace) or 2.67% lauryl maltose neopentyl glycol (LMNG; Anatrace) was together with diluted membranes pre-heated at 20 °C, 50 °C, 65 °C, or 75 °C for 15 minutes. After heating, 100 µl of solubilization buffer was added to each 300 µl diluted membrane aliquot so that the final amount of detergent was 1.34% (w/v) or 1.0%

(w/v) for DDM and LMNG, respectively. They were mixed before incubating at 20 °C, 50 °C, 65 °C, or 75 °C for 1.5 h. Through incubation, the solubilization samples were mixed by inversion every 5 minutes for 45 minutes. Immediately after this, the denatured proteins were removed by centrifugation at 16 100 rpm for 1 hour at 20 °C and the supernatant containing the solubilized protein was kept. Appropriate samples were taken before and after centrifugation, which were analysed immediately or frozen in liquid nitrogen and stored at -80 °C for later analysis. Image Lab software (Bio-Rad) was used to determine the detergent solubilization efficiencies of DDM and LMNG at different temperatures by comparing the relative intensities of the Western blot bands of samples taken from extracted membrane fractions (before centrifugation) and their corresponding detergent-solubilized fractions (after centrifugation) with the respect to a reference band consisting of known amount of wild-type *Tm*-PPase.

For large-scale protein solubilization, the extracted membranes were diluted with buffer C to 7.2 mg/ml into 3 X 25-30 ml fractions in 50 ml falcon tubes. 100 µl of 0.1M $Na_2PP_i$ was added to the diluted membranes which were then together with the solubilization buffer [50 mM MES (pH 6.5), 20% (v/v) glycerol, and 5.34% (w/v) DDM)] pre-warmed for 15 minutes at 75 °C. After this, 1 ml of solubilization buffer for every 3 ml of diluted membranes was added [final amount of 1.34% (w/v) DDM] and mixed before incubating at 75 °C for 1.5 h. Solubilizing membranes were mixed by inversion every 15 minutes during incubation. After this, the samples were centrifuged at 4000 rpm for 5 min at 20 °C and the supernatant was collected and left on ice to cool for 30 minutes before centrifuging again at 4000 rpm for 15 min at 4 °C. After this, the collected supernatant was immediately used for Ni-NTA affinity purification of A495K *Tm*-PPase. Appropriate samples were taken before and after the first centrifugation, which were immediately analysed or stored at -80 °C for later use.

## 2.4 Ni-NTA affinity purification

To each 40 ml of solubilized protein fraction, 32 ml of buffer C and 6 ml of 3.5M KCl (final concentration of 0.263 M) were added. These samples were then incubated at 40 °C for 1.5 hours under 50 rpm of shaking with 2 ml of nickel–nitrilotriacetic acid (Ni-NTA) beads (Qiagen), which had been prewashed with 20 ml of buffer C containing 0.2% (w/v) DDM. After this, the solubilized protein-bounded matrix was loaded into a 20 ml Econo-Pac chromatography column (Bio-Rad), which was then washed with 2 X column volume (CV) of washing buffer [20% (v/v) glycerol, 50 mM MES-NaOH (pH 6.5), 20 mM imidazole (pH 6.5), 50 mM KCl, 5 mM $MgCl_2$, 0.05% (w/v) DDM, 1 mM DDT, 1 mM PMSF, and 2 µg/ml pepstatin A], and the bound protein was eluted with 2

X CV of elution buffer [3.5% (v/v) glycerol, 50 mM MES-NaOH (pH 6.5), 400 mM imidazole (pH 6.5), 50 mM KCl, 5 mM $MgCl_2$, 0.5% (w/v) octyl glucose neopentyl glycol (OGNPG; Anatrace), 1 mM DDT, 1 mM PMSF, and 2 µg/ml pepstatin A]. All fractions were stored at 4 °C for subsequent analysis. Re-purification of the flow-through was done by re-applying the flow-through fraction to the purification column, which was first prepared by washing it with 2 X CV of buffer C. The washing and elution of the column was done similarly as in the first purification.

Based on SDS-PAGE electrophoresis, all fractions with purified protein and without impurities were appropriately combined and centrifuged at 4000 rpm for 10 min to remove any aggregates. After this, the protein samples were concentrated to ~ 200 µl aliquots with Amicon Ultra-4 30 kDa MWCO concentration column (Millipore), which was first washed with 1:5 diluted elution solution by centrifugation at 4000 rpm for 2 min. Concentrated protein was buffer-exchanged to the crystallization buffer [20 mM MES-NaOH (pH 6.5), 3.5% (v/v) glycerol, 50 mM KCl, 5 mM $MgCl_2$, 2 mM DTT, and 0.5% OGNG] using Micro Bio-Spin 6 columns (Bio-Rad) according to the manufacturer´s instruction. Buffer-exchanged protein samples were frozen with liquid nitrogen and stored at -80°C for subsequent activity assay and crystallization experiments.

## 2.5 Protein quantification

Total protein concentration of the membrane preparation was determined in triplicate by Bradford assay (Bradford, 1976) using bovine plasma gamma-globulin (BPGG, Sigma) as a standard. The standard curve consisted of 8 BPGG concentration levels: 0 mg/ml, 0.053 mg/ml, 0.1 mg/ml, 0.2 mg/ml, 0.4 mg/ml, 0.5 mg/ml, and 0.6 mg/ml. 20-fold, 50-fold, 100-fold and 200-fold membrane sample dilutions in $ddH_2O$ were used for the assay. First, 200 µl of Bradford reagent (Coomassie G-250 dye, Bio-Rad) was added to the wells of the 96-well plate. After this, 10 µl of each standard or membrane sample was added to the wells in triplicate and mixed. The samples were incubated at room temperature for 5 minutes before measuring the absorbance at 595 nm. Total protein concentration was calculated from the absorbance of the diluted samples using the standard curve. The concentration of the purified buffer-exchanged protein was determined in duplicate by the A280 method using a NanoDrop 2000 spectrophotometer at $OD_{280nm}$. For detergent solubilization assay, the protein concentration of the samples was determined from the Western Blot results by Image Lab software (Bio-Rad). Here, the intensity of each band was compared to the reference band of the purified *Tm*-PPase of known amount.

## 2.6 Hydrolytic activity assay

The $PP_i$ hydrolytic activity was determined using the molybdenum blue reaction method (Vidilaseris et al., 2018). For membrane preparation and small-scale solubilization assay, 40 µl reaction mixture per well containing 75 mM Tris-HCl pH 8.0, 3.5 mM $MgCl_2$, 143 mM KCl, 29 mM NaCl, 1.4 mM NaF, 12 µg/ml gramicidin D, and 5 µl of sample were added to PCR tube strips and sealed with adhesive PCR plate seal (Thermo Scientific), which was then cut to separate each strip of 8 tubes. The samples, including the appropriate controls, were pre-incubated for 5 minutes at 75 °C in a 96-well heat block to destroy any endogenous yeast pyrophosphatase activity before adding 10 µl of 2 mM sodium pyrophosphate ($Na_2PP_i$) substrate to the samples and mixing. Then, the samples were incubated for 5 minutes at 75 °C and then cooled on ice for 10 minutes. After enzyme inactivation by cooling, $Na_2PP_i$ was added to negative controls and then 60 µl of colouring solution was added to each tube to a final concentration of 84 mM ascorbic acid, 3 mM ammonium heptamolybdate, and 250 mM HCl. After 10 minutes of incubation on ice, in order to complex out the excess molybdate, 90 µl of arsenite-citrate solution was added in a fume cabinet to a final concentration of 69 mM sodium arsenite, 35 mM trisodium citrate, and 0.9 % (v/v) acetic acid. After 1 hour of incubation at room temperature, the samples were transferred to a 96 well PS microplate (Thermo Scientific) and the absorbance was measured at 860 nm on a MultiSkan Go microplate reader (Thermo Scientific). The total amount of released $P_i$ was determined by comparing the absorbance measurement at 860 nm to a $P_i$ standard curve comprising of 3 reaction mixtures with different amounts of sodium phosphate dihydrate ($NaH_2PO_4 \cdot 2H_20$): 2.5 nmol, 10 nmol, and 20 nmol. A reaction mixture without $NaH_2PO_4 \cdot 2H_20$ was used as a blank in the calculations and all measurements were done in triplicate.

For purified and detergent-solubilized A495K *Tm*-PPase and wild-type *Tm*-PPase, samples were first reconstituted into proteoliposomes by adding 10.4 ng of A495K *Tm*-PPase or 13.2 ng of *Tm*-PPase to 31.3 µl of reactivation solution [30 mg/ml soya lecithin, 3 mg/ml cholesteryl hemisuccinate, 20 mM Tris-HCl (pH 8), 1 mM DTT, and 7.2 % (w/v) DDM], which was prepared according to the protocol (Vidilaseris et al., 2018), and filling up to 50 µl with crystallization buffer [20 mM MES-NaOH (pH 6.5), 3.5% (v/v) glycerol, 50 mM KCl, 5 mM $MgCl_2$, 2 mM DTT, and 0.5% OGNG]. Four different 40-µl reaction mixtures with varying KCl concentration [75 mM Tris-HCl pH (8.0), 3.5 mM $MgCl_2$, 143 mM KCl, 29 mM NaCl, 0.2 µl reconstituted protein, and 0, 20, 50, or 100 mM KCl] were prepared for comparison of potassium dependence between A495K *Tm*-PPase and the wild-type. Specific PPase activity (µmol/min/mg) was calculated by dividing the average of released phosphate (µmol) with the reaction time (min) and the amount of loaded A495K *Tm*-PPase or WT *Tm*-PPase (mg).

## 2.7 Electrophoretic analysis and Western blotting

Solubilization and purification fraction samples were analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), which was carried out on 4–20% Mini-PROTEAN TGX Precast gel (Bio-Rad) using 1 X SDS running buffer. Samples, including standards, were mixed with 4 X SDS loading dye [50 mM Tris-HCl (pH 6.8), 2% SDS, 0.1% bromophenol blue, 10% glycerol, 2.5% 2-mercaptoethanol] in a ratio of 3:1 and incubated for 10 min at room temperature before loading them into the gel. For the detergent solubilization assay, 12 µl of each sample was loaded to the gel and run first for 20 min at 150 V and then for 40 min at 200 V, whereas for purification analysis, 30 µl of each sample was loaded to the gel and run for 1h at 120 V. 5 µl of molecular weight marker (10–250 kDa PageRuler Plus Prestained Protein Ladder, ThermoFisher Scientific) was also loaded to the gels. The gels were first washed with water for 15 minutes before staining them with Coomassie-blue (Bio-Safe) for 20 min and then destaining them with water on an orbital shaker overnight. The gels were imaged by Gel Doc EZ imager with White Light Conversion Screen (Bio-Rad) and analysed using Image Lab software (Bio-Rad).

Western blot analysis was performed on small-scale solubilization assay samples by transferring the electrophoresed protein bands onto 0.2 µm nitrocellulose membrane (Trans-Blot Turbo Transfer Pack, Bio-Rad) using the Trans-Blot Turbo Transfer System (Bio-Rad) with the settings of 2.5 A and 25 V for 3 minutes. After washing the membrane twice for 10 minutes with Tris-buffer saline (TBS) buffer [20 mM Tris-HCl and 150 mM NaCl (pH 7.5)], it was incubated for 1 hour in blocking buffer containing 3% (w/v) bovine serum albumin (BSA) in TBS-T (0.1% Tween 20 in TBS). The membrane was then incubated for 1 hour with primary polyclonal rabbit anti-mPPase antibody (Kellosalo et al., 2011) diluted 1/10000 in blocking buffer. After washing the membrane twice with TBS-T for 10 min and once with TBS buffer for 10 min, the membrane was treated with diluted horseradish peroxidase-conjugated anti-rabbit secondary antibody [1:20000 in Tween 20 containing 5% (w/v) skim milk] for 1 h. The membrane was then washed three times for 10 minutes with TBS-T before visualizing the immunoreactive bands by treating the membrane with enhanced chemiluminescent detection (ECL) reagent (Amersham, GE Healthcare) and using ChemiDoc MP Imaging system (Bio-Rad) together with Image Lab Software (Bio-Rad) for imaging and analysis, respectively. All washing/incubation steps prior to ECL treatment were performed on an orbital shaker at room temperature.

## 2.8 Protein crystallization

For crystallization, purified buffer-exchanged protein was concentrated to 7-8 mg/ml and 1 mM of substrate analogue imidodiphosphate (IDP) and/or a non-competitive inhibitor of *Tm*-PPase, ATC (Vidilaseris et al., 2019), was added to the protein solution, which was then incubated for 30 minutes on ice before centrifuging for 15 min at 16000*g* (4 °C). A495K *Tm*-PPase was crystallized in three different conformations, unbound (APO) state, IDP-bound, and IDP:ATC-bound. The initial screening of crystallization conditions for each conformation was performed according to the sparse matrix method (Jancarik & Kim, 1991) at 20 °C using the commercial MemGold HT-96 (Molecular Dimensions) screen, MRC two-well crystallization plates (Swissci) and the sitting-drop vapor-diffusion method. 50 µl of each screening solution (96 in total) was added into each reservoir well and 2 µl of protein solution into the 8-well strip caps. A mosquito robot (mosquito LCP, SPT Labtech) was used to prepare the crystallization plates by dispensing drops of protein and screening solution into the crystallization wells at ratios of 100 nl:100 nl and 200 nl:100 nl. The plates were sealed with clear plate sealer and then monitored for crystal formation for three weeks using a minstrel DT UV protein crystal imaging system (Rigaku). The crystallization conditions of the crystal hits that appeared from the screening were then optimized to get better crystals by varying PEG concentrations (Table 3.3). 24-well optimization plates were manually set up for each conformation using the sitting-drop vapor-diffusion method, by pipetting into the wells 1 µl of 8 mg/ml protein solution and 1 µl of optimized crystallization solution, and filling the reservoirs with 500 µl of appropriate optimized crystallization solution. Optimization plates were sealed with clear tape, kept at room temperature, and monitored daily under a microscope for a week. Crystals were harvested under a microscope from optimization plates and MRC two-well screening plates using CryoLoops (Molecular Dimensions). Prior to harvesting, 0.5 µl of reservoir solution (crystallization solution) was added to the well to minimize evaporation. The harvested crystals were directly frozen in liquid nitrogen and transported in a liquid nitrogen dewar to European Synchrotron Radiation Facility (ESRF), Grenoble (France) for X-ray diffraction.

## 2.9 X-ray data collection and structure solving

X-ray diffraction data were collected at ESRF on the beamline ID23-1 at 100 K cryostream on a PILATUS 6M detector with the help of Dr. Keni Vidilaseris. 2400 diffraction images were collected with an oscillation angle of 0.15 degrees and then processed using autoPROC (Vonrhein et al., 2011). AutoPROC toolbox contains X-ray Detector Software (XDS) package (Kabsch, 2010), which identifies and indexes the diffraction spots, and then integrates their intensities. Following the

autoPROC pipeline, AIMLESS program (Evans & Murshudov, 2013) was used to scale and merge the integrated diffraction intensities, where the space group is confirmed, systematic errors corrected, standard deviations of intensities validated, and reflection outliers rejected. An anisotropic cut-off (besides isotropic cut-off) and correction were alternatively applied on merged intensity data at a local mean I/σ(I) of 1.2 (default) using the STARANISO tool (Vonrhein et al., 2018) (included in autoPROC), before generating the merged reflection file used in solving the structure. Phenix.xtriage (v. 2.8.3) was used to analyse the quality of the merged data. The A495K *Tm*-PPases:$Mg_5$IDP:ATC structure was solved by molecular replacement using phenix.phaser-mr (v. 2.8.3) (Adams et al., 2010), anisotropically-corrected data and *Tm*-PPase:$Mg_5$IDP (PDB: 5LZQ) (Li et al., 2016) structure with $Mg_5$IDP removed as the search model. The model structure was built and refined in iterative cycles using phenix.refine (v. 1.19.1 – 4122) (Adams et al., 2010), interspersed with manual model adjustments using Coot (v. 0.9.3) (Emsley et al., 2010). The coordinates and geometric restraints of IDP (PDB: 2PN) and ATC (PDB: GQB) ligands were generated from SMILES strings using phenix.elbow (Moriarty et al., 2009) and then they were fitted to the electron density map using phenix.ligandfit (Adams et al., 2010) and Coot. Phenix.maps (Adams et al., 2010) was used to generate electron density maps for the figures, which were generated using PyMol 2.1.1 (Schrödinger, LLC, 2015). PyMol was also used for structural alignments when comparing the monomers/dimers within the structure or to other M-PPases. PDBe PISA server (https://www.ebi.ac.uk/pdbe/pisa/) (Krissinel & Henrick, 2007) was used for analysis of hydrogen bonds and monomer interfaces. Protein-Ligand Interaction Profiler (Adasme et al., 2021) was used to analyse hydrophobic and metal coordination interactions between the protein and ligands. MolProbity web server (http://molprobity.biochem.duke.edu/) (Davis et al., 2007) was used to evaluate the final structure. SplitPocket web server (http://pocket.med.wayne.edu/patch/) (Tseng et al., 2009) was used to analyse the volume of the hydrolytic pocket of M-PPase structures.

# 3. Results

## 3.1 Expression of A495K *Tm*-PPase in *S. cerevisiae*

*S. cerevisae* BJ1991 strain lacking two proteases, proteinase A and B, was used to express the histidine-tagged A495K *Tm*-PPase containing pRS1024 plasmid under control of constitutive *PMA1* promoter. Out of 6.88 litres of large-scale expression culture, 82.0 grams of cells were collected from which in total of 25 ml of membranes were extracted. The total protein concentration of the membrane preparation was determined, using the bovine plasma gamma-globulin (BPGG) standard curve (Figure 3.1), to be 21.0 mg/ml (Table 3.1). This amount was sufficient for the solubilization step.

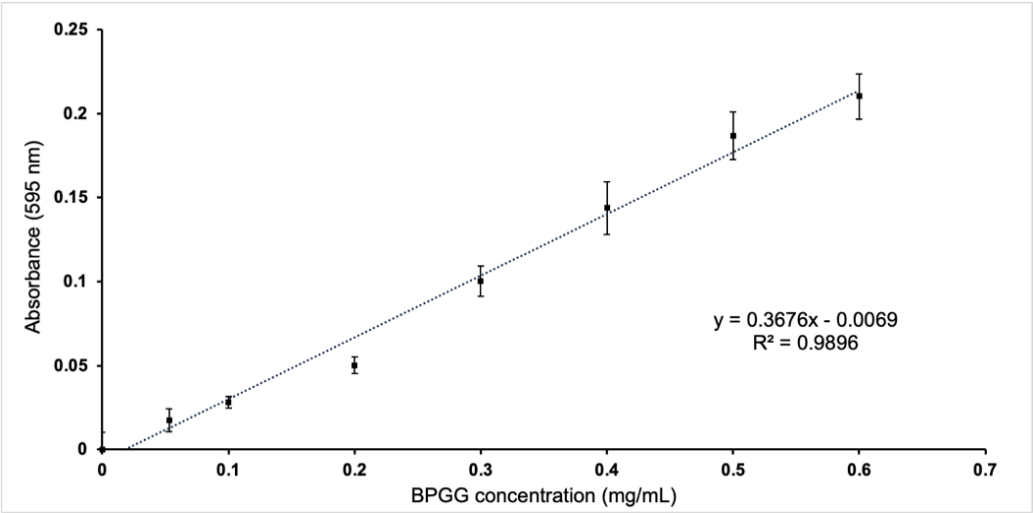


Figure 3.1: BPGG standard curve used for the Bradford protein assay. The standard curve was obtained by measuring the absorbance of BPGG standards at 595 nm. Each bar represents the mean of triplicate measurments ± S.D.

Table 3.1: Calculated protein concentrations for three sample dilutions in triplicates

| 50-fold dilution (mg/ml) | 100-fold dilution (mg/ml) | 200-fold dilution (mg/ml) | Average (mg/ml) |
|---|---|---|---|
| 17.4 | 22.2 | 23.3 | 21.0 |

## 3.2 Solubilization and activity assay of A495K *Tm*-PPase

After expressing in *S. cerevisae* BJ1991 and extracting the yeast membrane fraction, A495K *Tm*-PPase protein was solubilized using high-temperature solubilization ("hot-solve" method) (López-Marqués et al., 2005). Although high-temperature solubilization of wild-type *Tm*-PPase using DDM has been previously conducted with high solubility (Kellosalo et al., 2011), small-scale solubilization assay consisting of DDM and LMNG at different temperatures was performed due to the possible

structural effects of the point mutation (e.g., thermostability), which could affect the solubilization. Western blot of the SDS-PAGE gel (Figure 3.2A) loaded with detergent-solubilized protein samples and corresponding membrane preparation samples showed expression of a 67-kDa polypeptide, which was immunodetected using antibodies specific to a peptide conserved in M-PPases, thus confirming it being the A495K *Tm*-PPase. Comparison of the relative 67-kDa band intensities before and after solubilization showed high solubilization of over 98% for both detergents in all tested temperatures (Figure 3.2A). Since using temperatures close to the optimal temperature of *T. maritima* (~80 °C) for solubilization of *Tm*-PPase has been shown to result in high yield of detergent-solubilized protein with low content of impurities (Kellosalo et al., 2011), detergent solubilizations done at 65°C and 75°C were further visualized with Coomassie-blue staining of the SDS-PAGE gel loaded with these samples (Figure 3.2B). Both detergents, DDM and LMNG, solubilized the A495K *Tm*-PPase with high purity, as only the 67-kDa band is clearly visible in the soluble fractions.

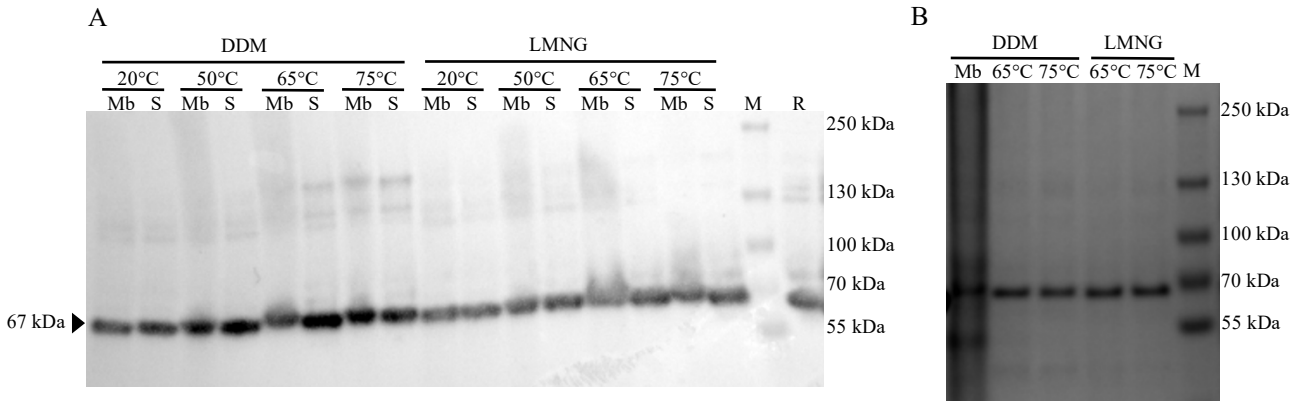


Figure 3.2: Solubilization assay of A495K *Tm*-PPase using DDM or LMNG at different temperatures. Extracted membrane fractions containing A495K *Tm*-PPase solubilized with 1.34% (w/v) DDM or 1.0% (w/v) LMNG detergents at four different temperatures. SDS-PAGE gel was loaded with 48.5 μg of total membrane protein and the A495K *Tm*-PPase was visualized with Western blotting and Coomassie-blue staining. (A) Immunodetection of A495K *Tm*-PPase. Based on the molecular weight (MW) standard (M) and using the MW analysis tools of Image Lab software, a high-density band for the expressed A495K *Tm*-PPase is seen with a MW of around 67 kDa (arrow). Detergent solubilization efficiency assessed by comparing the relative band intensities of extracted membranes (Mb) and the soluble fractions (S) with the respect to a reference band consisting of 0.1 mg of wild-type *Tm*-PPase (R). (B) Coomassie-blue staining of the soluble fractions at 65°C and 75°C and a sample of extracted membrane fraction (Mb).

After successful solubilization of A495K *Tm*-PPase was observed using either DDM or LMNG detergent, the $PP_i$ hydrolysing activity of the solubilization assay samples was assessed before choosing appropriate solubilization conditions for large-scale solubilization (Figure 3.3). No significant difference was observed in specific PPase activity between the two detergents, except at

room temperature (20 °C), where the highest activity between solubilized samples was observed to be 2.68 µmol/min/mg using DDM. In general, the solubilized samples had lower activity than the non-solubilized membrane preparation. The PPase activity of the protein samples solubilized with DDM was between 14% (at 20 °C) and 53% (at 65 °C) less than the PPase activity of non-solubilized membranes. In the case of LMNG, between 36% (50 °C) and 56% (65 °C) decrease of PPase activity was observed compared to the membrane preparation. Since PPase activity of protein sample solubilized in DDM was observed at 75 °C (1.76 µmol/min/mg) and due to high post-solubilization purity (3.2B) and previous demonstrations (Kellosalo et al., 2011; Li et al., 2016; Vidilaseris et al., 2019), DDM detergent and solubilization temperature of 75 °C were chosen for the large-scale solubilization.

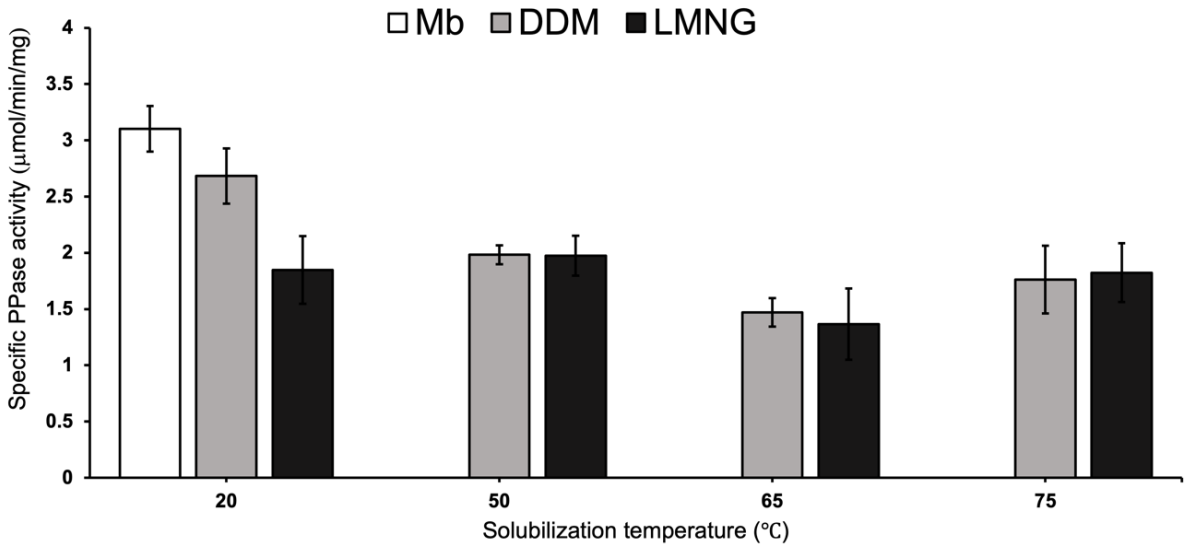


Figure 3.3: PPase activity assay of A495K *Tm*-PPase solubilized in different temperatures using DDM or LMNG. The values of specific PPase activity are per mg of A495K *Tm*-PPase protein and are shown as the mean and standard deviation of triplicate experiments. Depending on the solubilization efficiency, 8.8–15.6 mg/ml of A495K *Tm*-PPase was used in the assay. For the extracted membranes (Mb), the triplicate values of two different dilutions (2 mg/ml and 7 mg/ml) were used.

## 3.3 Large-scale solubilization and purification of A495K *Tm*-PPase

After solubilization at 75 °C using DDM and Ni-NTA purification (Figure 3.4), 12.85 mg of A495K *Tm*-PPase was obtained from 6.88 litres of cell culture (1.87 mg/L), which was sufficient for subsequent crystallization experiments, which generally require high protein concentration (8-20 mg/ml) and purity (McPherson & Cudney, 2014). The 67-kDa expression product is seen in all purification fractions including the non-solubilized extracted membranes (Figure 3.4A). As in the detergent solubilization assay, successful solubilization of the A495K *Tm*-PPase from the membranes

was observed based on the 67-kDa band intensity, which is seen to be the only clear visible band in the solubilized membrane fraction (SMF). Most of the protein eluted from the Ni-NTA matrix in the presence of 400 mM imidazole, is seen in fractions E1 and E2 (Figure 3.4A). These elution fractions including washing fraction 1 were combined for subsequent experiments, since the ~67-kDa band was observed in high intensity with no contaminating bands in all three fractions. The flow-through was re-applied to the purification column and re-purified, since it contained a lot of protein after the first purification (Figure 3.4B). After re-purification of the flow through, fractions E1.2 and E2.2 were combined for subsequent experiments.

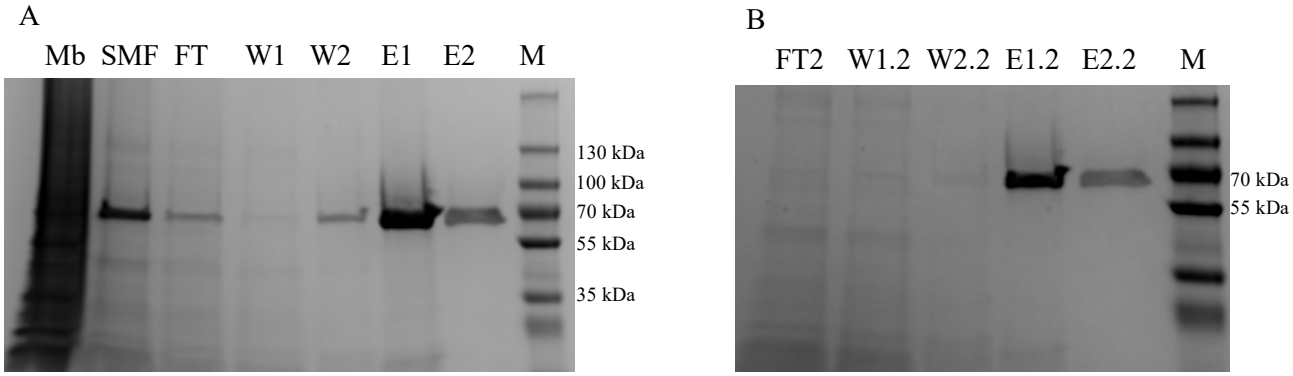


Figure 3.4: Electrophoretic analysis of A495K *Tm*-PPase Ni-NTA purification. (A) Purification of the large-scale expression. SDS–PAGE gel loaded with non-solubilized extracted membranes (Mb), solubilized membrane fraction (SMF), flow through (FT), washing fractions (W1-2), elution fractions (E1-2), and MW standard (M). 67-kDa band is seen in all lanes including in the pre-purification samples (Mb and SMF). (B) Re-purification of the flow-through (FT). SDS-PAGE gel loaded with MW standard (M), flow through (FT2), washing fractions (W1.2-W2.2) and elution fractions (E1.2-E2.2).

Pooled purification fractions were concentrated to 2-10 mg/ml and buffer-exchanged to crystallization buffer for activity assay and crystallization. Purified protein samples of A495K *Tm*-PPase and wild-type *Tm*-PPase were first reconstituted into proteoliposomes before measuring their specific PPase activity in fixed $Na^+$ and substrate ($Na_2PP_i$) concentration, but in different $K^+$ concentrations (Table 3.2, Figure 3.5). As expected, the hydrolytic activity of the wild-type *Tm*-PPase increased as the $K^+$ concentration increases up to 20 mM after which saturation begins to occur. Furthermore, the Ala → Lys substitution at 495 position rendered the activity of A495K *Tm*-PPase independent of the stimulatory effect of $K^+$, which is in line with previous studies (Artukka et al., 2018; Belogurov & Lahti, 2002). The activity of the A495K *Tm*-PPase gradually decreases as the $K^+$ concentration increases, possibly indicating displacement of the active site-bound $Mg^{2+}$ ions by $K^+$ ion as speculated in previous studies (Artukka et al., 2018). The overall PPase activity of the mutant is significantly lower than the wild-type, where even in the absence of $K^+$, the activity of the *Tm*-

PPase is nearly 2-fold higher than the mutant, most likely due to the $Na^+$ ions ability to partially replace the activator-role of $K^+$ in $Na^+$ -transporting M-PPases (Artukka et al., 2018; Malinen et al., 2007).

Table 3.2: A495K *Tm*-PPase and *Tm*-PPase activities in different $K^+$ concentrations

| | **Specific PPase activity (µmol/min/mg)** | |
|---|---|---|
| **[$K^+$] (mM)** | **A495K *Tm*-PPase** | ***Tm*-PPase** |
| 0 | 6.23 ± 0.22 | 11.5 ± 0.80 |
| 5 | 6.10 ± 0.17 | 28.8 ± 0.39 |
| 20 | 5.95 ± 0.14 | 42.6 ± 9.65 |
| 100 | 5.14 ± 0.27 | 49.0 ± 0.65 |

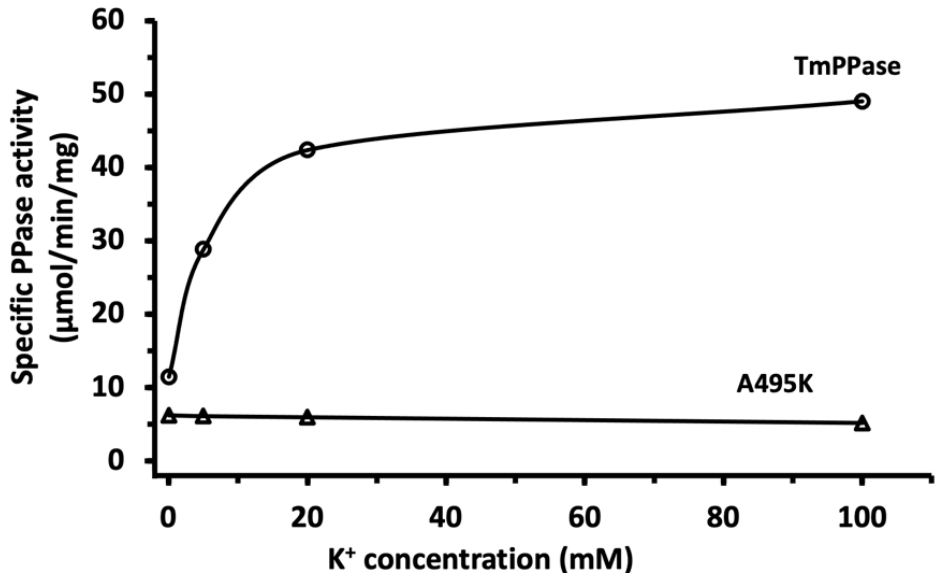


Figure 3.5: $K^+$ independence of PPase activity by A495K *Tm*-PPase compared to the wild type. The values of activity are per mg of *Tm*-PPase (dots) or A495K *Tm*-PPase (triangles) and they represent the means of three independent experiments (Table 3.2).

## 3.4 Crystallization of A495K *Tm*-PPase

*Tm*-PPase has been previously solved in various conformations, including resting state (APO) (Kellosalo et al., 2012), substrate analogue-bound state (IDP-bound) (Li et al., 2016), and substrate analogue- and inhibitor-bound state (IDP:ATC-bound) (Vidilaseris et al., 2019). Thus, A495K *Tm*-PPase was crystallized in these three different conformations. Initially, the optimum crystallization conditions were screened using commercial screening kit consisting of 96 different conditions.

Several crystallization conditions produced good quality crystals for all three conformations (Figure 3.6, Table 3.3). Based on the degree of supersaturation (e.g., solvent precipitation), size of the crystals, and crystal amount (nucleation sites), the four best crystallization conditions were chosen from the initial screening and optimized by varying the precipitant (PEGs) concentration (Table 3.3). In total, 44 crystals (14 APO, 16 IDP-bound, 14 IDP:ATC-bound) were harvested from optimization and MRC two-well screening plates, and sent to ESFR for data collection.

Table 3.3: Optimized crystallization conditions of the A495K *Tm*-PPase conformations

| Conformation | Optimized crystallization conditions | Original screening conditions |
|---|---|---|
| APO | 0.2 M sodium chloride, 0.1 M HEPES pH 7.0, and 18%, 20%, 22%, 25%, 27%, or 30% (v/v) PEG 350 MME | 0.2 M sodium chloride, 0.1 M HEPES pH 7.0, 22% (v/v) PEG 500 MME |
| IDP and IDP:ATC | 0.2 M calcium chloride, 0.1 M HEPES pH 7.0, and 28%, 30%, 33%, 35%, 37%, or 40% (v/v) PEG 400 | 0.2 M calcium chloride dihydrate, 0.1 M HEPES pH 7.0, and 33% (v/v) PEG 400 |
| IDP:ATC | 0.2 M calcium chloride, 0.1 M Tris-HCl pH 8.0, and 37%, 39%, 41%, 44%, 47%, or 49% (v/v) PEG 400 | 0.2 M calcium chloride dihydrate, 0.1 M Tris-HCl pH 8.0, 44% (v/v) PEG 400 |
| IDP:ATC | 0.1 M sodium chloride, 0.1 M MES pH 6.5, and 28%, 30%, 33%, 35%, 37%, or 40% (v/v) PEG 400, and ethylene glycol 4% (v/v) | **0.1 M sodium chloride, 0.1 M MES pH 6.5, 33% (v/v) PEG 400, and ethylene glycol 4% (v/v) *** |
| APO, IDP and IDP:ATC | 0.1 M sodium chloride, 0.1 M sodium phosphate pH 7.4, and 25%, 28%, 30%, 33%, 35%, or 37% (v/v) PEG 300 | 0.1 M sodium chloride, 0.1 M sodium phosphate pH 7.0, and 33% (v/v) PEG 300 |
| APO and IDP | 0.2 M calcium chloride, 0.1 M MES pH 6.5, and 28%, 30%, 33%, 35%, 37%, or 40% (v/v) PEG 400 | 0.2 M calcium chloride dihydrate, 0.1 M MES pH 6.5, and 33% (v/v) PEG 400 |
| APO | 0.001 M zinc sulfate, 0.05 M HEPES pH 7.8, and 20%, 23%, 25%, 28%, 30%, or 33% (v/v) PEG 600 | 0.001 M zinc sulfate heptahdyrate, 0.05 M HEPES pH 7.8, and 28% (v/v) PEG 600 |
| IDP | 0.001 M cadmium sulfate, 0.1 M MES pH 6.5, and 25%, 28%, 30%, 33%, 35%, or 37% (v/v) PEG 400 | 0.001 M cadmium chloride hemi(pentahydrate), 0.03 M magnesium chloride hexahydrate, 0.1 M MES pH 6.5, and 30% (v/v) PEG 400 |

*Condition of the solved structure


Figure 3.6: Crystals of A495K *Tm*-PPase conformations APO (A), IDP (B), and IDP:ATC (C). These crystals were obtained in the screening of crystallization conditions using a commercial screen. Crystallization conditions in (A): 0.2 M calcium chloride dihydrate, 0.1 M Tris-HCl pH 8.0, 44% (v/v) PEG 400, (B): 0.2 M calcium chloride dihydrate, 0.1 M HEPES pH 7.0, and 33% (v/v) PEG 400, and (C): 0.1 M sodium chloride, 0.1 M sodium phosphate pH 7.0, and 33% (v/v) PEG 300.

## 3.5 X-ray data collection and structure solving

IDP:ATC-bound A495K *Tm*-PPase crystallized in 0.1 M sodium chloride, 0.1 M MES pH 6.5, 33% (v/v) PEG 400, and ethylene glycol 4% (v/v) diffracted the best to 3.06 Å (Table 3.4). The collected diffraction data were indexed, integrated, scaled, merged, and corrected for anisotropy using the autoPROC pipeline (Vonrhein et al., 2011). Two A495K *Tm*-PPase dimers per asymmetric unit cell were identified with a space group of $P2_12_12_1$ as in the previously solved *Tm*-PPase:$Mg_5$IDP:ATC (PDB: 6QXA) (Vidilaseris et al., 2019) structure. Analysis of the isotropically treated data (no anisotropic correction) showed severe anisotropy, which prompted the use of anisotropically-corrected data for subsequent structure solving and refinement. The quality of this corrected data was also analysed, which besides the data being elliptically truncated and having overall low spherical completeness (62.8%) due to anisotropy, showed no data anomalies or crystal defects that could prevent structure solution. After assessing the data quality, the structure was solved by molecular replacement using the *Tm*-PPase:$Mg_5$IDP (PDB: 5LZQ) (Li et al., 2016) with $Mg_5$IDP removed as the search model. After first the round of refinement, positive ($F_o$-$F_c$) electron density was observed at 3 σ at the substrate binding site, cationic centre at position $A495^{12.46}$ (Ballesteros-Weinstein residue number superscripted), and ion gate in all 4 chains (Figure 3.7A). The $Mg_5$IDP complex was built at the substrate binding site similarly as in the search model and the sodium ion was fitted to the density at the ion gate (Figure 3.7B). Alanine at the 495 position was mutated to lysine, and manually refined to fit the positive ($F_o$-$F_c$) electron density. Positive ($F_o$-$F_c$) electron densities shaped as two pancakes were observed at 3 σ near the exit channel of the chain A (Figure 3.8A). Similar pancake-shaped positive density was also observed beside the exit channel of chain D (Figure 3.8A). ATC molecules

were fitted into these appropriately (Figure 3.8B) as seen in previously solved *Tm*-PPase:IDP:ATC (PDB: 6QXA) (Vidilaseris et al., 2019) structure. After the automated structure refinement process, including manual rebuilding and fitting of $Mg_5$IDP complex, ATC and K495 into the positive electron density ($F_o$-$F_c$), the final $R_{work}/R_{free}$ was 21.74%/27.64% with appropriate stereochemistry. 98.3% of the residues were found in the favoured region of the Ramachandran plot, with no outliers. Complete data collection and refinement statistics can be seen at Table 3.4. All diffraction data was processed to a resolution limit of 3.06 Å.

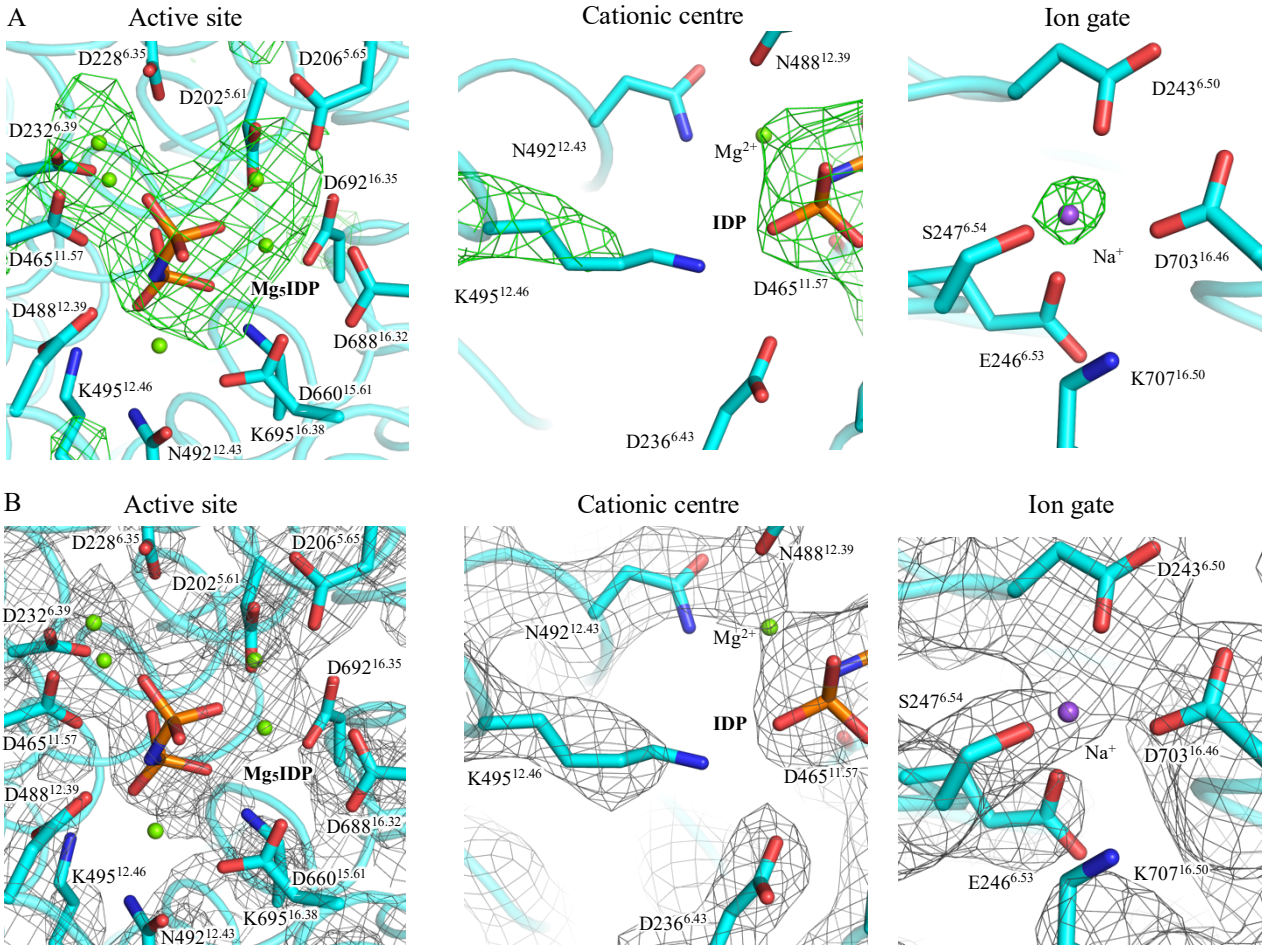


Figure 3.7: Electron density maps of functional regions in the A495K *Tm*-PPases:$Mg_5$IDP:ATC structure. (A) $F_o$-$F_c$ omit maps of active site, cationic centre, and ion gate with fitted ligands, including five $Mg^{2+}$ (green spheres), IDP and $Na^+$ (violet sphere). Positive $F_o$-$F_c$ difference density is contoured at 3σ and seen as green mesh. At the cationic centre (middle panel) lysine ($K495^{12.46}$) residue, replacing the alanine residue, is seen fitting into the $F_o$-$F_c$ density. (B) $2F_o$-$F_c$ maps of active site, cationic centre and ion gate fitted with ligands, contoured at 1σ and seen as grey mesh. All pictures were generated using the chain B of the solved A495K *Tm*-PPases:$Mg_5$IDP:ATC structure. Ballesteros-Weinstein residue number superscripted on the right of the conventional residue number.

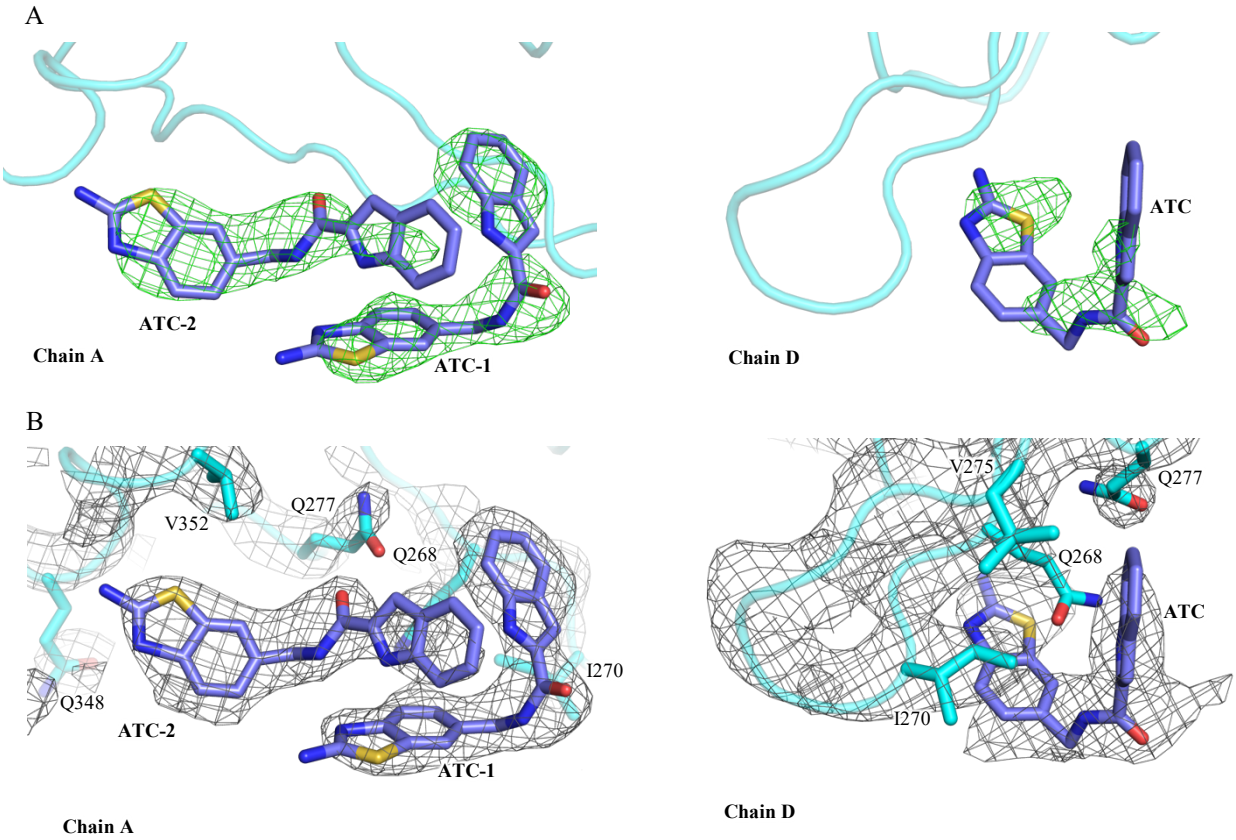


Figure 3.8: Electron density maps of ATC binding sites of the A495K *Tm*-PPases:$Mg_5$IDP:ATC structure. (A) $F_o$-$F_c$ omit maps of the fitted ATC molecules in chain A and chain D near the exit channel. Positive $F_o$-$F_c$ difference density is contoured at 3σ and seen as green mesh. (B) $2F_o$-$F_c$ maps (contoured at 1σ and seen as grey mesh) of the ATC molecules surrounded by the interacting residues.

Table 3.4: X-ray data collection and refinement statistics of A495K *Tm*-PPase:$Mg_5$IDP:ATC structure

| **Data collection statistics*** | | **Refinement statistics‡** | |
|---|---|---|---|
| Space group | P $2_12_12_1$ | $R_{work}$(%)/$R_{free}$(%) | 21.74/27.64 |
| Unit cell dimensions | | No. of atoms | 21621 |
| *a, b, c* (Å) | 99.6, 145.9, 251.3 | Protein | 21490 |
| α, β, λ (°) | 90, 90, 90 | Ligands | 181 |
| X-ray source | ESRF ID23-1 | Solvent | 2 |
| Wavelength (Å) | 0.9763 | No. of chains (ASU) | 4 |
| Resolution range (Å) | 126.20-3.06 (3.39-3.06) | Average B-factor (Å$^2$) | 65.82 |
| $CC_{1/2}$ | 0.998 (0.706) | Protein | 65.73 |
| Anisotropy direction | | Ligands/ions | 80.61 |
| Resolution, where locally-averaged value of $I/\sigma(I)$ >= 1.2 | | Solvent | 47.43 |
| Overall (Å) | 3.06 | R.M.S. Deviations | |
| along h axis (Å) | 3.48 | Bond lengths (Å) | 0.002 |
| along k axis (Å) | 4.16 | Bond angle (°) | 0.47 |
| along l axis (Å) | 3.01 | Ramachandran statistics † | |
| Total reflections | 584108 (29017) | Favoured (%) | 98.34 |
| Unique reflections | 43906 (2195) | Allowed (%) | 1.66 |
| Mean $I$ / $\sigma I$ | 10.3 (1.6) | Outliers (%) | 0 |
| Completeness (ellipsoidal) | 0.95 (0.77) | MolProbity score† | 1.84 (100$^{th}$ percentile) |
| Multiplicity | 13.3 (13.2) | | |
| Wilson B (Å$^2$) | 79.16 | | |
| $R_{merge}$ | 0.19 (1.57) | | |
| $R_{meas}$ | 0.20 (1.63) | | |
| $R_{pim}$ | 0.05 (0.45) | | |

Statistics for the highest-resolution shell are shown in parentheses.
*Data collection statistics, except Wilson B (Å$^2$), were calculated by the AIMLESS tool of the autoPROC software package after the anisotropic cut-off.
‡Refinement statistics including Wilson B (Å$^2$) of the collected data were calculated by Phenix
†Ramachandran statistics and MolProbity score were calculated by MolProbity

## 3.6 Structural overview of the A495K *Tm*-PPase:$Mg_5$IDP:ATC

The overall structure of the A495K *Tm*-PPase:$Mg_5$IDP:ATC (Figure 3.9) is very similar to the other solved M-PPase structures, with the root-mean-square deviation (r.m.s.d.) ranging from 0.30 Å (*Tm*-PPase:$Mg_5$IDP:ATC) to 1.24 Å (*Vr*-PPase:$Mg_2P_i$) based on the average distance of Cα atoms of

subunit A (Table 3.5). When aligning and comparing the two dimers, AB and CD (Figure 3.9), the r.m.s.d. per Cα atom is 0.33 Å and between individual monomers it ranges from 0.22 to 0.30 Å. As previously described (Kellosalo et al., 2012; Lin et al., 2012), there is an interface between the monomers forming the dimer (AB or CD) (Figure 3.10) and physiologically irrelevant ATC-mediated interface between the dimers AB and CD, most likely due to crystal packing (Vidilaseris et al., 2019).

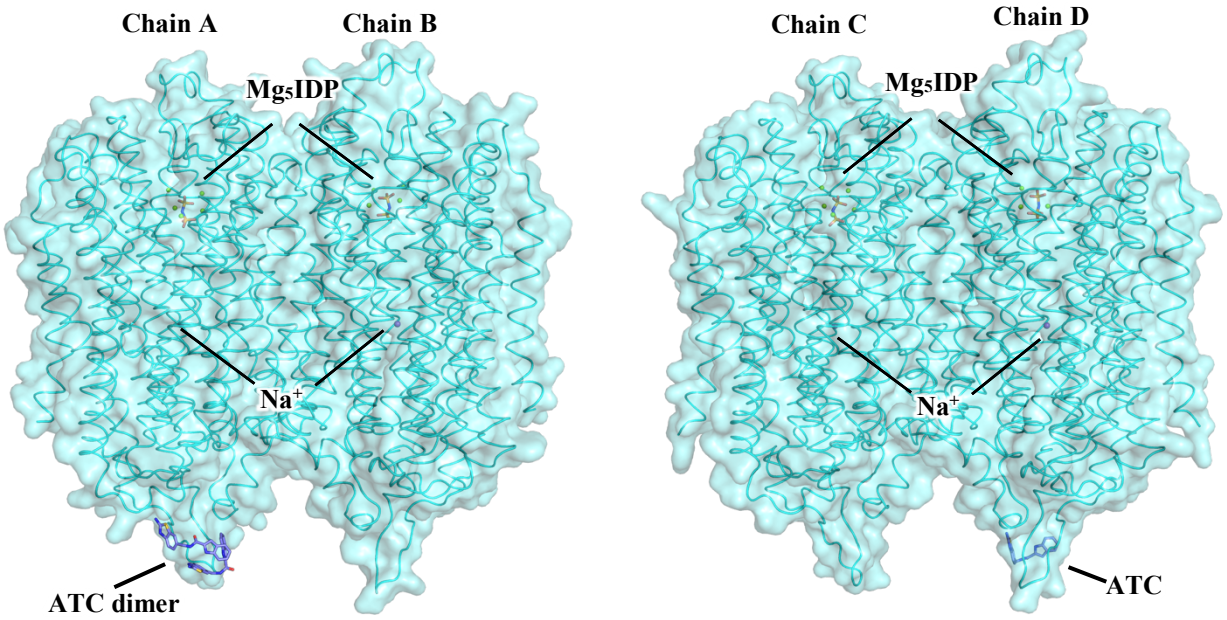


Figure 3.9: Structural overview of the A495K *Tm*-PPases:$Mg_5$IDP:ATC from the membrane plane

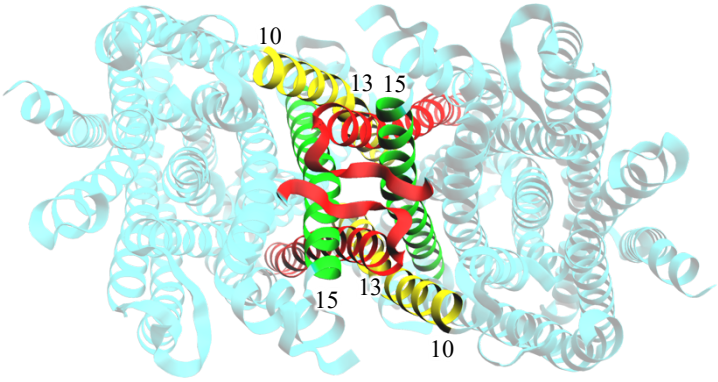


Figure 3.10: A495K *Tm*-PPases:IDP:ATC dimer interface. The AB dimer viewed from the periplasmic side with interacting transmembrane helices coloured and numbered, where yellow is helice 10, green 15, and red 13, including the antiparallel ß-sheet between helices 12 and 13.

Table 3.5: Comparison of chain A of A495K *Tm*-PPase:$Mg_5$IDP:ATC to other M-PPases

| Compared structure * | Rmsd (Å) |
|---|---|
| *Tm*-PPase:$Mg_5$IDP:ATC (6QXA) | 0.30 |
| *Tm*-PPase:$Mg_5$IDP (5LZQ) | 0.35 |
| *Tm*-PPase:$Mg_4P_{i2}$ (4AV6) | 0.86 |
| *Vr*-PPase:$Mg_5P_{i2}$ (6AFS) | 0.88 |
| *Vr*-PPase:$Mg_5$IDP (4A01) | 0.88 |
| *Tm*-PPase:CaMg (4AV3) | 1.04 |
| *Tm*-PPase:$Mg_2WO_4$ (5LZR) | 1.22 |
| *Vr*-PPase:$Mg_2P_i$ (5GPJ) | 1.24 |

*PDB ID of the compared structures shown in parentheses

Closer inspection of the active site shows IDP being coordinated by all five $Mg^{2+}$ ions and the surrounding residues, including K695$^{16.38}$, K663$^{15.64}$ and K495$^{12.46}$ (Figure 3.11A), which form hydrogen bonds with IDP. Consistent with previous structural studies (Li et al., 2016; Vidilaseris et al., 2019), the sodium ion is penta-coordinated at the ionic gate by D243$^{6.50}$, D246$^{6.53}$, S247$^{6.54}$, and D703$^{16.46}$ residues (Figure 3.11B). As previously mentioned, in the resting-state *Tm*-PPase structure with no bound $Na^+$, the K707$^{16.50}$ residue is reorientated and forms an ion-triplet with D243$^{6.50}$ and D246$^{6.53}$ which comprises the ionic gate (Li et al., 2016). During refinement, additional density was observed near the exit channel of chains A and D, to which ATCs were fitted (Figure 3.8). Two ATC molecules bind in a hydrophobic cleft formed by the connecting loops of TMH6-7, TMH8-9, and TMH12-13 located near the exit channel of chain A. ATC-1 forms hydrophobic interactions with residues Q268 and I270 of loop between TMH6 and 7 (loop6-7), whereas ATC-2 forms a hydrogen bond to Q348, part of the loop8-9 (Figure 3.11C). In chain D, one ATC molecule binds, similarly to ATC-1 in Chain A, in a cleft formed by loop6-7, forming hydrophobic interactions with I270, V275, and Q277, and a hydrogen bond with K269. Observed ATC binding-sites are consistent with previous findings (Vidilaseris et al., 2019).

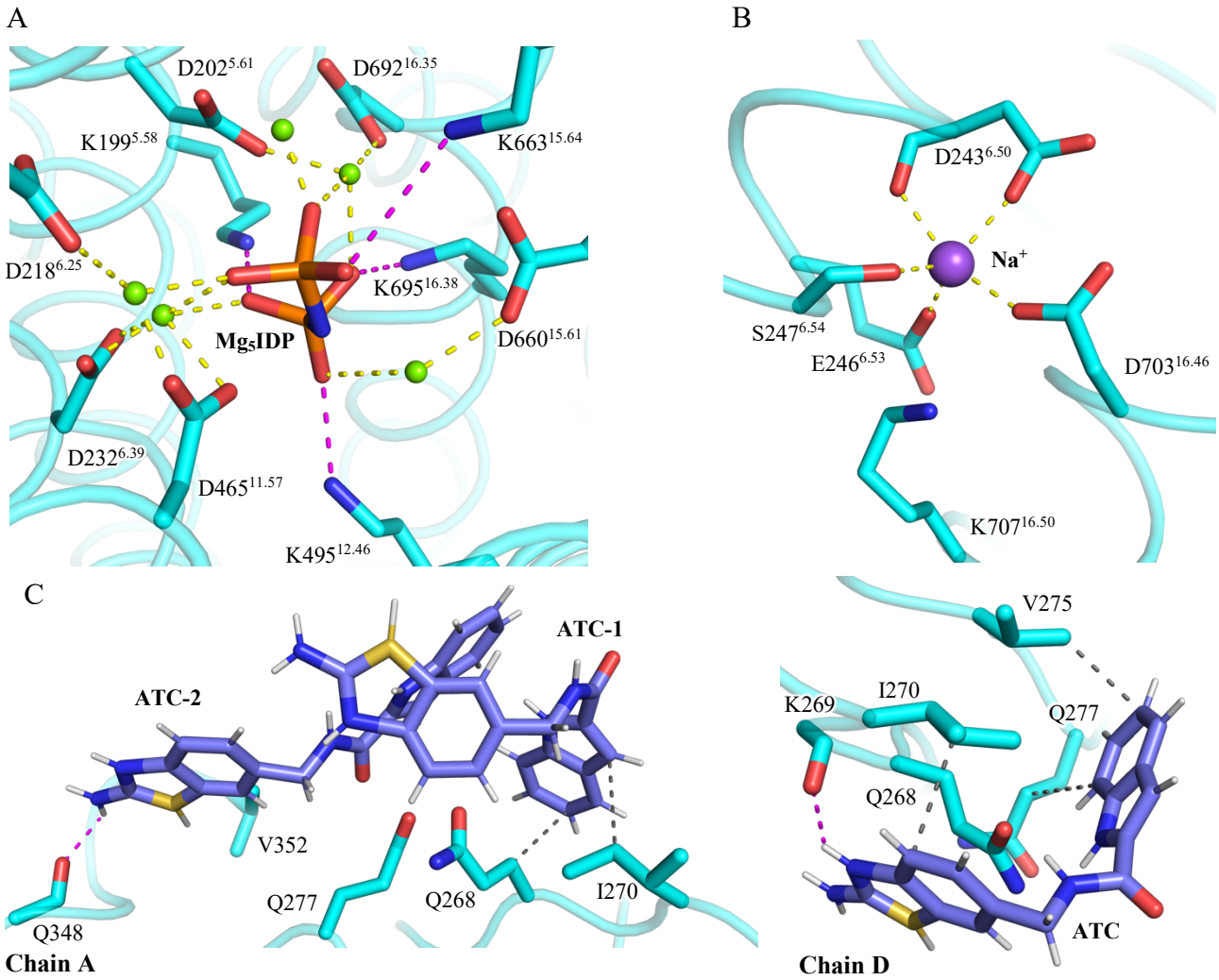

Figure 3.11: Structural overview of A495K *Tm*-PPases:$Mg_5$IDP:ATC active site, ion gate and ATC binding sites. (A) Active site where IDP is coordinated by $Mg^{2+}$ ions (green spheres) and surrounding residues. (B) Ion gate with $Na^+$ ion (purple sphere) coordinated by the oxygens of two aspartate residues, one glutamate and one serine. (C) ATC binding sites in chain A and chain D with interacting residues shown. Purple, yellow and grey dashed lines represent hydrogen bonds, metal ion coordination and hydrophobic interactions, respectively. Figures A and B generated from chain B.

## 3.7 Structural characteristics of $K^+$ -independence in A495K *Tm*-PPase:$Mg_5$IDP:ATC

Positive ($F_o$-$F_c$) electron density was observed at the cationic centre, emanating from A495$^{12.46}$ residue and extending towards the IDP-substrate (Figure 3.7A). A495$^{12.46}$ was mutated to a lysine (K495$^{12.46}$) which was fitted into the positive density. The $F_o$-$F_c$ electron density map of K495$^{12.46}$ was poorer in chains A, C and D compared to B, making it a more reliable chain for analysing the cationic centre. K495$^{12.46}$ and the rest of the side chains at the active site coordinating IDP-substrate are well defined in the $2F_o$-$F_c$ electron density map (Figure 3.7B). The orientation of these side chains and the coordination of the $Mg_5$IDP is almost identical to other M-PPases, with the r.m.s.d. per Cα atom of the residues at the cationic centre (all 10 residues from chain B seen in Figure 3.12A) being 0.29 Å, 0.31 Å and 0.31 Å compared to *Tm*-PPase:$Mg_5$IDP (PDB: 5LZQ), *Tm*-PPase:$Mg_5$IDP:ATC (PDB: 6QXA), and *Vr*-PPase:$Mg_5$IDP (PDB: 4A01), respectively. However, the hydrolytic pocket volume of the active site, 407 Å$^3$, is smaller compared to these structures, where the volume ranges from 522 (PDB: 4A01) to 751 Å$^3$ (PDB: 6QXA). Deletion of $K^+$ and replacement of A495 of *Tm*-PPase:$Mg_5$IDP:ATC (PDB: 6QXA) into K495 by manual modelling in Coot led in decrease of the hydrolytic volume to 474 Å$^3$, suggesting that the mutated lysine (A495K) is the reason for the lower hydrolytic volume observed in A495K *Tm*-PPases:$Mg_5$IDP:ATC. The K495$^{12.46}$ reaches towards the hydrolytic pocket where the ε-$NH_3^+$ group coordinates the IDP-substrate by forming a hydrogen bond with the electrophilic phosphate group of the IDP (Figure 3.12A). When comparing to the wild-type $K^+$-dependent variant, *Tm*-PPases:$Mg_5$IDP:ATC (Figure 3.12B), the ε-$NH_3^+$ group of the K495$^{12.46}$ is seen to be occupying the same position and forming the same interactions as the $K^+$ (Figure 3.12C).

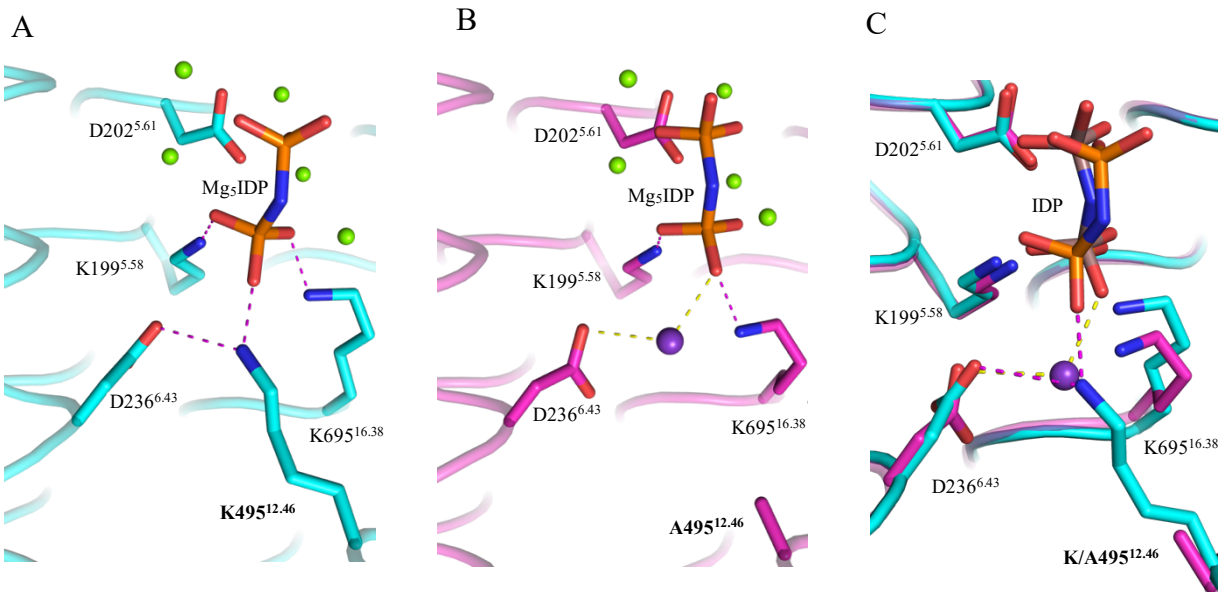


Figure 3.12: Overview of the cationic centre of A495K *Tm*-PPases:$Mg_5$IDP:ATC. (A) The $K495^{12.46}$ cationic centre (labelled in bold) of A495K *Tm*-PPases:$Mg_5$IDP:ATC coordinating the IDP together with other surrounding residues including the $Mg^{2+}$ metal cage (green spheres). (B) The $K^+$ cationic centre of wild-type *Tm*-PPases:$Mg_5$IDP:ATC (PDB: 6QXA) (Vidilaseris et al., 2019) with signature $A495^{12.46}$ residue (labelled in bold). (C) $K495^{12.46}$(Lys)/$K^+$ cationic centre of structurally aligned A495K *Tm*-PPases:$Mg_5$IDP:ATC and *Tm*-PPases:$Mg_5$IDP:ATC. Phosphorous atoms in IDP of wild-type *Tm*-PPase coloured in brown for clarity. Hydrogen bonds shown by purple dashed lines and coordination interactions of $K^+$ (purple sphere) shown by yellow dashed lines

# 4. DISCUSSION

## 4.1 The production of A495K *Tm*-PPase protein for activity assay and crystallization experiments

Two important requirements of X-ray crystallography are high protein purity and concentration, as obtaining well diffracting protein crystals requires protein homogeneity and bringing the protein to supersaturation. Besides this, another challenge is to identify the optimum crystallization conditions, often requiring a combination matrix of different crystallization conditions to be tested and thus large amounts of purified protein. Thus, for these two requirements, a suitable expression system is crucial (Abts et al., 2012; McPherson & Cudney, 2014). The suitability of *S. cerevisae* for heterologous expression and characterization of several membrane-bound pyrophosphatases of different organisms has been demonstrated (Kellosalo et al., 2012; Kellosalo et al., 2011; Li et al., 2016; López-Marqués et al., 2005; Tsai et al., 2019; Vidilaseris et al., 2019), and hence it was used as an expression host in this work. The BJ1991 strain lacking genes for vacuolar aspartyl protease (proteinase A; *pep4-3*) and proteinase B (*rb1-1122*) in particular has been previously used successfully for expression and structural characterization of *Tm*-PPase protein (Kellosalo et al., 2012; Li et al., 2016; Tsai et al., 2019; Vidilaseris et al., 2019) and thus was chosen for expression of A495K *Tm*-PPase. Using above mentioned yeast strain and methods for expression and purification of A495K *Tm*-PPase, a total yield of 1.87 mg/l was obtained, which is less than previously reported (2.5 mg/l) when using the same strain and methods (Kellosalo et al., 2011). One likely reason for the lower yield is the loss of protein during purification, where some of it was lost in the column flow through even after repurification of the flow through, as seen in Figure 3.4B. This was thought to be due to the insufficient amount of added Ni-NTA beads, which would have resulted in free unbound protein capable of flowing through the column. However, this amount, combined with successful protein purification (Figure 3.4), was adequate for the activity assay and subsequent crystallization experiments, which resulted in verifying the $K^+$-independent $PP_i$ hydrolysis of A495K *Tm*-PPase and, more importantly, solving the structure of A495K *Tm*-PPases:$Mg_5$IDP:ATC.

## 4.2 A495K *Tm*-PPases:$Mg_5$IDP:ATC structure

A495K *Tm*-PPases:$Mg_5$IDP:ATC structure was solved at a resolution of 3.06Å. Reliable interpretation of the functional regions, including active site, ionic gate and the cationic centre, were able to be drawn from the electron density maps generated from the obtained diffraction data (Figure 3.7). Ligands, including $Mg_5$IDP and ATC, and the sidechains of residues surrounding them were able to be built into the electron density maps, which showed sufficient details (Figure 3.7 and 3.8).

Structural comparison between the solved A495K *Tm*-PPases:$Mg_5$IDP:ATC structure and other M-PPases, especially *Tm*-PPases, revealed high similarity in the overall structure (Table 3.5). Besides the A495K mutation, the active site of the A495K *Tm*-PPases:$Mg_5$IDP:ATC was seen to be almost identical to other M-PPases complexed with $Mg_5$IDP. The structure was solved bound to ATC, an allosteric inhibitor of *Tm*-PPases discovered by Vidilaseris et al., 2019. As described, the ATC dimer bound into the chain A "locks" and prevents the full motion in chain B, inhibiting hydrolysis in an asymmetric matter. Although no activity assay was performed in the presence of ATC, the structural evidence supports the previous findings (Vidilaseris et al., 2019) as the ATC is seen to bind as a dimer near the exit channel of chain A, presumably preventing the motion of chain B, where the $Mg_5$IDP is seen bound (Figure 3.9).

Unfortunately, no mechanistic insight into cation translocation mechanism was provided by the solved A495K *Tm*-PPases structure, as it was only solved in one conformational state ("locked" by ATC and substrate-bound) at a resolution of 3.06 Å. Solving the A495K *Tm*-PPase structure in additional conformational states, e.g., IDP-bound (substrate-analogue-bound), $P_i$-bound (product-bound), or resting, at high-resolution, where for instance water molecules could have been reliably modelled (<2.5 Å), would have allowed for structural comparison between the different catalytic states, possibly providing more detailed structure-function insights into M-PPases. However, new structural information on $K^+$ dependence and the regarding role of the $K^+$/Lys cation centre was provided.

## 4.3 Contradiction between the structural and biochemical data regarding $K^+$ dependence

A495K *Tm*-PPase exhibited insensitivity to $K^+$ and significantly decreased catalytic activity (Figure 3.5), which is in line with previous Ala → Lys mutational studies (Artukka et al., 2018; Belogurov & Lahti, 2002). Artukka et al. (2018) suggested there to be two co-existing, $K^+$-bound and $K^+$-unbound, pathways in $PP_i$ hydrolysis and that in many Ala → Lys-substituted M-PPases, the $PP_i$ hydrolysis occurs mainly via the low-activity $K^+$-unbound pathway, which could also be the case for the A495K *Tm*-PPase. Furthermore, the degree of catalytic compensation provided by the introduced lysine varied between different M-PPases from 4% to 50% of the wild-type activity, suggesting that the positioning of the introduced lysine, in respect to the authentic $K^+$ position, dictates the level of compensation (Artukka et al., 2018). However, this claim and the observed overall decreased $PP_i$

hydrolysis (Figure 3.5) are in contradiction with the observed structural data (Figure 3.12). The cationic centre of the solved structure suggests an almost identical structural compensation of $K^+$ by the introduced ε-$NH_3^+$ group of the $K495^{12.46}$, as it is seen to be occupying the same position and forming the same interactions as the $K^+$ in wild type *Tm*-PPase (Figure 3.12C). The enzyme not being active in the absence of $K^+$, despite the almost perfect structural compensation, implies that the $K^+$/Lys cationic centre is not the sole determinant of $K^+$ dependence. Another factor determining $K^+$ dependence could be the residue at the position 12.49 which, besides A/K12.46, is different between $K^+$-dependent and -independent M-PPases (Belogurov & Lahti, 2002). The Gly/Ala12.49 conserved among the $K^+$-independent is crucial for $K^+$-binding affinity (Belogurov & Lahti, 2002), but due to lack of structural evidence it is not evident what kind of effects changes at 12.49 position has on the cationic centre and what is its role in $PP_i$ hydrolysis. Indeed, in the future, structural characterization of a double mutant variant ($A^{12.46}K$ and $T^{12.49}A$), or even better, solving a $K^+$-independent M-PPase structure at high-resolution, could provide more information on $K^+$ dependence and on the role of $G/A/T^{12.49}$ residue.

In conclusion, A495K *Tm*-PPase was successfully expressed and purified, which allowed to conduct the hydrolytic activity assay and solve the structure of A495K *Tm*-PPases:$Mg_5$IDP:ATC, which gave new insights on $K^+$ dependence of M-PPases.

## 5. Acknowledgments

I would like to thank my supervisor Prof. Adrian Goldman for his time, thoughtful advice and for providing me the opportunity to carry out my master´s thesis work in his lab. I would also like to thank my co-supervisor Dr. Keni Vidilaseris for his time, exceptional mentoring, and guidance throughout the experimental part of the work and during the analysis of the final structural results. Finally, I would like to thank all the members of Prof. Adrian Goldman´s Helsinki group who always had the time to help me in the lab when I needed.

# 6. References


Abts, A., Schwarz, C., Tschapek, B., Smits, S., & Schmitt, L. (2012). Rational and Irrational Approaches to Convince a Protein to Crystallize. *InTech*. https://doi.org/10.5772/28014

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., … Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica Section D Biological crystallography*, 66, 213–221. https://doi.org/10.1107/s0907444909052925

Adasme, M. F., Linnemann, K. L., Bolz, S. N., Kaiser, F., Salentin, S., Haupt, V. J., & Schroeder, M. (2021). PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA. *Nucleic Acids Research*, 49(W1), W530–W534. https://doi.org/10.1093/nar/gkab294

Aires, J. (2016). Antimicrobial Resistance and Drug Efflux Pumps in *Bacteroides*. In X.-Z. Li, C. A. Elkins, & H. I. Zgurskaya (Eds.), Efflux-Mediated Antimicrobial Resistance in Bacteria: Mechanisms, Regulation and Clinical Implications. *Springer International Publishing*, 515–526. https://doi.org/10.1007/978-3-319-39658-3_20

Artukka, E., Luoto, H. H., Baykov, A. A., Lahti, R., & Malinen, A. M. (2018). Role of the potassium/lysine cationic center in catalysis and functional asymmetry in membrane-bound pyrophosphatases. *The Biochemical Journal*, 475(6), 1141–1158. https://doi.org/10.1042/BCJ20180071

Asaoka, M. M., Segami, S., Ferjani, A., & Maeshima, M. (2016). Contribution of $PP_i$-Hydrolyzing Function of Vacuolar $H^+$-Pyrophosphatase in Vegetative Growth of *Arabidopsis*: Evidenced by Expression of Uncoupling Mutated Enzymes. *Frontiers Plant Science*, 7, 415. https://doi.org/10.3389/fpls.2016.00415

Ballesteros, J. A., & Weinstein, H. (1995). Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods in Neurosciences*, 25, 366–428. https://doi.org/https://doi.org/10.1016/S1043-9471(05)80049-7

Baltscheffsky, H., Von Stedingk, L. V., Heldt, H. W., & Klingenberg, M. (1966). Inorganic pyrophosphate: formation in bacterial photophosphorylation. *Science*, 153(3740), 1120–1122. https://doi.org/10.1126/science.153.3740.1120

Baykov, A. A., Malinen, A. M., Luoto, H. H., & Lahti, R. (2013). Pyrophosphate-fueled $Na^+$ and $H^+$ transport in prokaryotes. *Microbiology and Molecular Biology Reviews*, 77(2), 267–276. https://doi.org/10.1128/MMBR.00003-13

Belogurov, G. A., & Lahti, R. (2002). A lysine substitute for $K^+$: A460K mutation eliminates $K^+$ dependence in $H^+$-pyrophosphatase of *Carboxydothermus hydrogenoformans*. *Journal of Biological Chemistry*, 277(51), 49651–49654. https://doi.org/https://doi.org/10.1074/jbc.M210341200

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72, 248–254. https://doi.org/10.1006/abio.1976.9999

Carystinos, G. D., MacDonald, H. R., Monroy, A. F., Dhindsa, R. S., & Poole, R. J. (1995). Vacuolar $H^+$-Translocating Pyrophosphatase Is Induced by Anoxia or Chilling in Seedlings of Rice. *Plant Physiology*, 108(2), 641–649. https://doi.org/10.1104/pp.108.2.641

Chen, J., Brevet, A., Fromant, M., Lévêque, F., Schmitter, J. M., Blanquet, S., & Plateau, P. (1990). Pyrophosphatase is essential for growth of *Escherichia coli*. *Journal of Bacteriology*, 172(10), 5686–5689. https://doi.org/10.1128/jb.172.10.5686-5689.1990

Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., … Richardson, D. C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Reseach*, 35(Web Server issue), W375–383. https://doi.org/10.1093/nar/gkm216

Docampo, R., Jimenez, V., Lander, N., Li, Z. H., & Niyogi, S. (2013). New insights into roles of acidocalcisomes and contractile vacuole complex in osmoregulation in protists. *International Review of Cell and Molecular Biology*, 305, 69–113. https://doi.org/10.1016/b978-0-12-407695-2.00002-0

Docampo, R., & Moreno, S. N. (2008). The acidocalcisome as a target for chemotherapeutic agents in protozoan parasites. *Current Pharmaceutical Design*, 14(9), 882–888. https://doi.org/10.2174/138161208784041079

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D Biological Crystallography*, 66, 486–501. https://doi.org/10.1107/s0907444910007493

Evans, P. R., & Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta Crystallographica Section D Biological Crystallography*, 69, 1204–1214. https://doi.org/10.1107/s0907444913000061

Ferjani, A., Segami, S., Horiguchi, G., Muto, Y., Maeshima, M., & Tsukaya, H. (2011). Keep an eye on $PP_i$: the vacuolar-type $H^+$-pyrophosphatase regulates postgerminative development in *Arabidopsis*. *Plant Cell*, 23(8), 2895–2908. https://doi.org/10.1105/tpc.111.085415

Fuglsang, A. T., Paez-Valencia, J., & Gaxiola, R. A. (2011). Plant Proton Pumps: Regulatory Circuits Involving $H^+$-ATPase and $H^+$-PPase. In M. Geisler & K. Venema (Eds.), Transporters and Pumps in Plant Signaling. Springer Berlin Heidelberg, 7, 39–64 https://doi.org/10.1007/978-3-642-14369-4_2

García-Contreras, R., Celis, H., & Romero, I. (2004). Importance of Rhodospirillum rubrum $H^+$-Pyrophosphatase under Low-Energy Conditions. *Journal of Bacteriology*, 186(19), 6651–6655. https://doi.org/doi:10.1128/JB.186.19.6651-6655.2004

Gaxiola, R. A., Li, J., Undurraga, S., Dang, L. M., Allen, G. J., Alper, S. L., & Fink, G. R. (2001). Drought- and salt-tolerant plants result from overexpression of the AVP1 $H^+$-pump. *Proceedings of the National Academy of Sciences*, 98(20), 11444–11449. https://doi.org/10.1073/pnas.191389398

Gaxiola, R. A., Sanchez, C. A., Paez-Valencia, J., Ayre, B. G., & Elser, J. J. (2012). Genetic manipulation of a "vacuolar" H(+)-PPase: from salt tolerance to yield enhancement under phosphorus-deficient soils. *Plant Physiology*, 159(1), 3–11. https://doi.org/10.1104/pp.112.195701

Hernández, A., Herrera-Palau, R., Madroñal, J. M., Albi, T., López-Lluch, G., Perez-Castiñeira, J. R., … Serrano, A. (2016). Vacuolar $H^+$-Pyrophosphatase AVP1 is Involved in Amine Fungicide Tolerance in *Arabidopsis thaliana* and Provides Tridemorph Resistance in Yeast. *Frontiers in Plant Science*, 7, 85. https://doi.org/10.3389/fpls.2016.00085

Holmes, A. O. M., Kalli, A. C., & Goldman, A. (2019). The Function of Membrane Integral Pyrophosphatases From Whole Organism to Single Molecule. *Frontiers in Molecular Biosciences*, 6, 132. https://doi.org/10.3389/fmolb.2019.00132

Huber, R., Langworthy, T. A., König, H., Thomm, M., Woese, C. R., Sleytr, U. B., & Stetter, K. O. (1986). *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. *Archives of Microbiology*, 144(4), 324–333. https://doi.org/10.1007/BF00409880

Jancarik, J., & Kim, S.-H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of Applied Crystallography*, 24(4), 409–411. https://doi.org/https://doi.org/10.1107/S0021889891004430

Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of Computational Chemistry*, 29(11), 1859–1865. https://doi.org/10.1002/jcc.20945

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta crystallographica Section D Biological Crystallography*, 66, 133–144. https://doi.org/10.1107/s0907444909047374

Kajander, T., Kellosalo, J., & Goldman, A. (2013). Inorganic pyrophosphatases: one substrate, three mechanisms. *FEBS Letters*, 587(13), 1863–1869. https://doi.org/10.1016/j.febslet.2013.05.003

Kay, H. D. (1928). The phosphatases of mammalian tissues: Pyrophosphatase. *Biochemical Journal*, 22(6), 1446–1448. https://doi.org/10.1042/bj0221446

Kellosalo, J., Kajander, T., Kogan, K., Pokharel, K., & Goldman, A. (2012). The Structure and Catalytic Cycle of a Sodium-Pumping Pyrophosphatase. *Science*, 337(6093), 473–476. https://doi.org/doi:10.1126/science.1222505

Kellosalo, J., Kajander, T., Palmgren, M. G., Lopéz-Marqués, R. L., & Goldman, A. (2011). Heterologous expression and purification of membrane-bound pyrophosphatases. *Protein Expression and Purification*, 79(1), 25–34. https://doi.org/https://doi.org/10.1016/j.pep.2011.05.020

Krissinel, E., & Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology*, 372(3), 774–797. https://doi.org/https://doi.org/10.1016/j.jmb.2007.05.022

Lemercier, G., Dutoya, S., Luo, S., Ruiz, F. A., Rodrigues, C. O., Baltz, T., … Bakalara, N. (2002). A vacuolar-type $H^+$-pyrophosphatase governs maintenance of functional acidocalcisomes and growth of the insect and mammalian forms of *Trypanosoma brucei*. *Journal of Biological Chemistry*, 277(40), 37369–37376. https://doi.org/10.1074/jbc.M204744200

Li, J., Yang, H., Peer, W. A., Richter, G., Blakeslee, J., Bandyopadhyay, A., … Gaxiola, R. (2005). *Arabidopsis* $H^+$-PPase AVP1 regulates auxin-mediated organ development. *Science*, 310(5745), 121–125. https://doi.org/10.1126/science.1115711

Li, K.-M., Wilkinson, C., Kellosalo, J., Tsai, J.-Y., Kajander, T., Jeuken, L. J. C., … Goldman, A. (2016). Membrane pyrophosphatases from *Thermotoga maritima* and *Vigna radiata* suggest a conserved coupling mechanism. *Nature Communications*, 7(13596), 1-11. https://doi.org/10.1038/ncomms13596

Lin, S. M., Tsai, J. Y., Hsiao, C. D., Huang, Y. T., Chiu, C. L., Liu, M. H., … Sun, Y. J. (2012). Crystal structure of a membrane-embedded $H^+$-translocating pyrophosphatase. *Nature*, 484(7394), 399–403. https://doi.org/10.1038/nature10963

Liu, J., Pace, D., Dou, Z., King, T. P., Guidot, D., Li, Z. H., … Moreno, S. N. (2014). A vacuolar-$H^+$-pyrophosphatase (TgVP1) is required for microneme secretion, host cell invasion, and extracellular survival of *Toxoplasma gondii*. *Molecular Microbiology*, 93(4), 698–712. https://doi.org/10.1111/mmi.12685

López-Marqués, R. L., Pérez-Castiñeira, J. R., Buch-Pedersen, M. J., Marco, S., Rigaud, J.-L., Palmgren, M. G., & Serrano, A. (2005). Large-scale purification of the proton pumping pyrophosphatase from *Thermotoga maritima*: A "Hot-Solve" method for isolation of recombinant thermophilic membrane proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1716(1), 69–76. https://doi.org/https://doi.org/10.1016/j.bbamem.2005.08.004

López-Marqués, R. L., Pérez-Castiñeira, J. R., Losada, M., & Serrano, A. (2004). Differential regulation of soluble and membrane-bound inorganic pyrophosphatases in the photosynthetic bacterium *Rhodospirillum rubrum* provides insights into pyrophosphate-based stress bioenergetics. *Journal of Bacteriology*, 186(16), 5418–5426. https://doi.org/10.1128/jb.186.16.5418-5426.2004

Luoto, H. H., Belogurov, G. A., Baykov, A. A., Lahti, R., & Malinen, A. M. (2011). $Na^+$-translocating membrane pyrophosphatases are widespread in the microbial world and evolutionarily precede $H^+$-translocating pyrophosphatases. *Journal of Biological Chemistry*, 286(24), 21633–21642. https://doi.org/10.1074/jbc.M111.244483

Luoto, H. H., Nordbo, E., Baykov, A. A., Lahti, R., & Malinen, A. M. (2013). Membrane $Na^+$-pyrophosphatases can transport protons at low sodium concentrations. *Journal of Biological Chemistry*, 288(49), 35489–35499. https://doi.org/10.1074/jbc.M113.510909

Luoto, H. H., Nordbo, E., Malinen, A. M., Baykov, A. A., & Lahti, R. (2015). Evolutionarily divergent, $Na^+$-regulated $H^+$-transporting membrane-bound pyrophosphatases. *Biochemical Journal*, 467(2), 281–291. https://doi.org/10.1042/BJ20141434

Malinen, A. M., Belogurov, G. A., Baykov, A. A., & Lahti, R. (2007). $Na^+$-pyrophosphatase: a novel primary sodium pump. *Biochemistry*, 46(30), 8872–8878. https://doi.org/10.1021/bi700564b

Marchesini, N., Luo, S., Rodrigues, C. O., Moreno, S. N., & Docampo, R. (2000). Acidocalcisomes and a vacuolar $H^+$-pyrophosphatase in malaria parasites. *Biochemical Journal*, 347*,* 243–253.

Martinoia, E., Maeshima, M., & Neuhaus, H. E. (2007). Vacuolar transporters and their essential role in plant metabolism. *Journal of Experimental Botany*, 58(1), 83–102. https://doi.org/10.1093/jxb/erl183

McPherson, A., & Cudney, B. (2014). Optimization of crystallization conditions for biological macromolecules. *Acta Crystallographica Section F Structural Biology Communications*, 70, 1445–1467. https://doi.org/10.1107/s2053230x14019670

Mora, C., McKenzie, T., Gaw, I. M., Dean, J. M., von Hammerstein, H., Knudson, T. A., … Franklin, E. C. (2022). Over half of known human pathogenic diseases can be aggravated by climate change. *Nature Climate Change*, 12(9), 869–875. https://doi.org/10.1038/s41558-022-01426-1

Moriarty, N. W., Grosse-Kunstleve, R. W., & Adams, P. D. (2009). electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallographica Section D Biological Crystallography*, 65, 1074–1080. https://doi.org/10.1107/s0907444909029436

Nordbo, E., Luoto, H. H., Baykov, A. A., Lahti, R., & Malinen, A. M. (2016). Two independent evolutionary routes to $Na^+/H^+$ cotransport function in membrane pyrophosphatases. *Biochemical Journal*, 473(19), 3099–3111. https://doi.org/10.1042/BCJ20160529

Pogozheva, I. D., Armstrong, G. A., Kong, L., Hartnagel, T. J., Carpino, C. A., Gee, S. E., … Im, W. (2022). Comparative Molecular Dynamics Simulation Studies of Realistic Eukaryotic, Prokaryotic, and Archaeal Membranes. *Journal of Chemical Information and Modeling*, 62(4), 1036–1051. https://doi.org/10.1021/acs.jcim.1c01514

Primo, C., Pizzio, G. A., Yang, J., Gaxiola, R. A., Scholz-Starke, J., & Hirschi, K. D. (2019). Plant proton pumping pyrophosphatase: the potential for its pyrophosphate synthesis activity to modulate plant growth. *Plant Biology,* 21(6), 989–996. https://doi.org/10.1111/plb.13007

Schiestl, R. H., & Gietz, R. D. (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Current Genetic*, 16(5-6), 339–346. https://doi.org/10.1007/bf00340712

Schilling, R. K., Tester, M., Marschner, P., Plett, D. C., & Roy, S. J. (2017). AVP1: One Protein, Many Roles. *Trends in Plant Science*, 22(2), 154–162. https://doi.org/10.1016/j.tplants.2016.11.012

Schrödinger, LLC. (2015). The PyMOL Molecular Graphics System, Version 2.0 (2.1.1). https://www.pymol.org/

Segami, S., Asaoka, M., Kinoshita, S., Fukuda, M., Nakanishi, Y., & Maeshima, M. (2018). Biochemical, Structural and Physiological Characteristics of Vacuolar $H^+$-Pyrophosphatase. *Plant and Cell Physiology*, 59(7), 1300–1308. https://doi.org/10.1093/pcp/pcy054

Segami, S., Nakanishi, Y., Sato, M. H., & Maeshima, M. (2010). Quantification, organ-specific accumulation and intracellular localization of type II $H^+$-pyrophosphatase in *Arabidopsis thaliana*. *Plant and Cell Physiology*, 51(8), 1350–1360. https://doi.org/10.1093/pcp/pcq096

Serrano, A., Pérez-Castiñeira, J. R., Baltscheffsky, M., & Baltscheffsky, H. (2007). $H^+$-PPases: yesterday, today and tomorrow. *IUBMB Life*, 59(2), 76–83. https://doi.org/https://doi.org/10.1080/15216540701258132

Shah N. R., Vidilaseris Keni., Xhaard H. & Goldman A. (2016). Integral membrane pyrophosphatases: a novel drug target for human pathogens? *AIMS Biophysics*, 3(1), 171–194. https://doi.org/10.3934/biophy.2016.1.171

Tsai, J.-Y., Tang, K.-Z., Li, K.-M., Hsu, B.-L., Chiang, Y.-W., Goldman, A., & Sun, Y.-J. (2019). Roles of the Hydrophobic Gate and Exit Channel in *Vigna radiata* Pyrophosphatase Ion Translocation. *Journal of Molecular Biology*, 431(8), 1619–1632. https://doi.org/https://doi.org/10.1016/j.jmb.2019.03.009

Tsai, J. Y., Kellosalo, J., Sun, Y. J., & Goldman, A. (2014). Proton/sodium pumping pyrophosphatases: the last of the primary ion pumps. *Current Opinion in Structural Biology*, 27, 38–47. https://doi.org/10.1016/j.sbi.2014.03.007

Tseng, Y. Y., Dupree, C., Chen, Z. J., & Li, W. H. (2009). SplitPocket: identification of protein functional surfaces and characterization of their spatial patterns. *Nucleic Acids Research*, 37(Web Server issue), W384–389. https://doi.org/10.1093/nar/gkp308

Vidilaseris, K., Kellosalo, J., & Goldman, A. (2018). A high-throughput method for orthophosphate determination of thermostable membrane-bound pyrophosphatase activity. *Analytical Methods*, 10(6), 646–651. https://doi.org/10.1039/C7AY02558K

Vidilaseris, K., Kiriazis, A., Turku, A., Khattab, A., Johansson, N. G., Leino, T. O., … Goldman, A. (2019). Asymmetry in catalysis by *Thermotoga maritima* membrane-bound pyrophosphatase demonstrated by a nonphosphorus allosteric inhibitor. *Science Advances,* 5(5), eaav7574. https://doi.org/doi:10.1126/sciadv.aav7574

Vinothkumar, K. R., & Henderson, R. (2010). Structures of membrane proteins. *Quarterly Reviews of Biophysics*, 43(1), 65–158. https://doi.org/10.1017/s0033583510000041

Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., … Bricogne, G. (2011). Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica Section D Biological Crystallography*, 67, 293–302. https://doi.org/10.1107/s0907444911007773

Vonrhein, C., Tickle, I. J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A. J., & Bricogne, G. (2018). Advances in automated data analysis and processing within autoPROC, combined with improved characterisation, mitigation and visualisation of the anisotropy of diffraction limits using STARANISO. *Acta Crystallographica Section A Foundations and Advances*, 74, a360. https://doi.org/10.1107/s010876731809640x

Wang, Y., Xu, H., Zhang, G., Zhu, H., Zhang, L., Zhang, Z., … Ma, Z. (2009). Expression and responses to dehydration and salinity stresses of V-PPase gene members in wheat. *Journal of Genetics and Genomics*, 36(12), 711–720. https://doi.org/10.1016/s1673-8527(08)60164-2

Wehner, G. G., Balko, C. C., Enders, M. M., Humbeck, K. K., & Ordon, F. F. (2015). Identification of genomic regions involved in tolerance to drought stress and drought stress induced leaf senescence in juvenile barley. *BMC Plant Biology*, 15, 125. https://doi.org/10.1186/s12870-015-0524-3

Wexler, H. M. (2007). *Bacteroides*: the good, the bad, and the nitty-gritty. *Clinical Microbiology Reviews*, 20(4), 593–621. https://doi.org/10.1128/cmr.00008-07

Yang, H., Knapp, J., Koirala, P., Rajagopal, D., Peer, W. A., Silbart, L. K., … Gaxiola, R. A. (2007). Enhanced phosphorus nutrition in monocots and dicots over-expressing a phosphorus-responsive type I $H^+$-pyrophosphatase. *Plant Biotechnology Journal*, 5(6), 735–745. https://doi.org/10.1111/j.1467-7652.2007.00281.x

Yang, H., Zhang, X., Gaxiola, R. A., Xu, G., Peer, W. A., & Murphy, A. S. (2014). Over-expression of the *Arabidopsis* proton-pyrophosphatase AVP1 enhances transplant survival, root mass, and fruit development under limiting phosphorus conditions. *Journal of Experimental Botany*, 65(12), 3045–3053. https://doi.org/10.1093/jxb/eru149

Yang, J., Zhang, T., Li, J., Wu, N., Wu, G.,Yang, J., … Chen, J. (2020). *Chinese wheat mosaic virus*-derived vsiRNA-20 can regulate virus infection in wheat through inhibition of vacuolar- $H(^+)$ -PPase induced cell death. *New Phytologist*, 226(1), 205–220. https://doi.org/10.1111/nph.16358

Zhang, M., Wang, C., Otto, T. D., Oberstaller, J., Liao, X., Adapa, S. R., … Adams, J. H. (2018). Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science*, 360(6388). https://doi.org/10.1126/science.aap7847