AIMS Biophysics, 3(1): 171-194.
DOI: 10.3934/biophy.2016.1.171
Received: 30 December 2015
Accepted: 01 March 2016
Published: 07 March 2016

http://www.aimspress.com/journal/biophysics

Review

Integral membrane pyrophosphatases: a novel drug target for human pathogens?

Nita R. Shah [1, †], Keni Vidilaseris [2, †], Henri Xhaard [3], and Adrian Goldman [1,2,*]

[1] School of Biomedical Sciences and Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK.
[2] Department of Biosciences, Division of Biochemistry, University of Helsinki, Helsinki, Finland.
[3] Faculty of Pharmacy, Centre for Drug Research, University of Helsinki, Helsinki, Finland
† Co-first authors

*** Correspondence:** Email: a.goldman@leeds.ac.uk; Tel: 0113-343-8537.

Abstract: Membrane-integral pyrophosphatases (mPPases) are found in several human pathogens, including *Plasmodium* species, the protozoan parasites that cause malaria. These enzymes hydrolyze pyrophosphate and couple this to the pumping of ions (H^+ and/or Na^+) across a membrane to generate an electrochemical gradient. mPPases play an important role in stress tolerance in plants, protozoan parasites, and bacteria. The solved structures of mPPases from *Vigna radiata* and *Thermotoga maritima* open the possibility of using structure-based drug design to generate novel molecules or repurpose known molecules against this enzyme. Here, we review the current state of knowledge regarding mPPases, focusing on their structure, the proposed mechanism of action, and their role in human pathogens. We also summarize different methodologies in structure-based drug design and propose an example region on the mPPase structure that can be exploited by these structure-based methods for drug targeting. Since mPPases are not found in animals and humans, this enzyme is a promising potential drug target against livestock and human pathogens.

Keywords: pyrophosphatase; structure-based drug design; membrane; human pathogens; malaria

Abbreviations:
PP_i: inorganic pyrophosphate;

P_i: inorganic phosphate;
PPase: pyrophosphatase;
sPPase: soluble pyrophosphatase;
mPPase: membrane-bound pyrophosphatase;
IPTG: β-D-1-thiogalactoside;
TMH: transmembrane helix;
Na^+-mPPase: Na^+-pumping membrane-bound pyrophosphatase;
Na^+/H^+-mPPase: Na^+ and H^+-pumping membrane-bound pyrophosphatase;
V-ATPase: vacuolar ATPase;
AVP1: H^+-mPPase in *Arabidopsis*;
DALYs: disability adjusted life years;
FDA: US Food and Drug Administration;
FPPS: farnesyl diphosphate synthase;
SI: selectivity indexes;
CC_{50}: 50% cytotoxicity concentration;
EC_{50}: 50% effective concentration;
VrPPase: *Vigna radiata* membrane-bound pyrophosphatase;
IDP: imidodiphosphate;
TmPPase: *Thermotoga maritima* membrane-bound pyrophosphatase;
RMSD: root mean square deviation;
GPCR: G protein-coupled receptors;
ADME: absorption; distribution; metabolism and excretion;
PAMPA: parallel artificial membrane permeability assay;
CYP: cytochrome P450;
hERG: human ether-à-go-go-related gene;
P-gp: permeability glycoprotein;
ABCC2: ATP-binding cassette sub-family C member 2;
MATE: multi-antimicrobial extrusion protein;
NMR: nuclear magnetic resonance;
SPR: surface plasmon resonance;
HTS: high throughput screening;
DUD: directory of useful decoys
MCSS: multiple copy simultaneous search;
dCTPase: deoxycytidine triphosphate pyrophosphatase 1;
IPP: isopentenyl diphosphate
dNTPs: nucleotide triphosphates;
AMDP: aminomethylenediphosphonate.

1. Introduction

Inorganic pyrophosphatases catalyse the hydrolysis of inorganic pyrophosphate (PP_i) to two inorganic phosphate (P_i) molecules. PP_i is a by-product of at least 190 cellular reactions, including the generation of the fundamental building blocks of life: polynucleotide chain synthesis and generation of aminoacyl-tRNA. Controlling the level of PP_i is crucial for driving these reactions forward, thus highlighting the importance of pyrophosphatases (PPases) [1–4]. PP_i also plays a regulatory role in the cell as an intracellular biochemical intermediate of several enzymes and *e.g.* in mammals as an inhibitor of calcification in connective tissue matrices and other extracellular fluids [2,5].

Managing the level of PP_i in the cell through PP_i hydrolysis is primarily accomplished by the ubiquitous soluble PPases (sPPases), which have k_{cat} ranging from 200 to 2000 s^{-1}, as opposed to the much slower membrane-bound PPases (mPPases) with k_{cat} of approximately 10 s^{-1} [2,6]. Chen *et al.* demonstrated that sPPase is essential in *Escherichia coli* through replacement of the chromosomal copy of *ppa*, the sPPase gene, with *ppa* under the control of a *lac* promoter. This strain grew in the presence of isopropyl β-D-1-thiogalactoside (IPTG), which induced expression of sPPase, but was unable to grow in the absence of IPTG [7]. Mitochondrial sPPase in *Saccharomyces cerevisiae* is also essential for cell viability, since this enzyme is required for mitochondrial function [8]. *E. coli* sPPase, and other bacterial sPPases, are homohexamers whereas the *S. cerevisiae* sPPase is a homodimer, though the monomeric fold of these various proteins is the same [9,10].

Unlike the globular sPPase, mPPase is embedded in the membrane *via* 15 to 17 transmembrane helices (TMHs), with no globular domain [11,12]. In addition to hydrolyzing PP_i, mPPases couple this action with the pumping of ions across the membrane to generate an electrochemical gradient [13]. This ion gradient can be utilized as a source of energy by the cell during times of stress [14]. Also unlike the sPPases, mPPases only occur in algae, plants, selected species of protozoan parasites, bacteria and archaea – but not in animals or fungi [13,15,16]. Below, we discuss the biology, function and structure of mPPases, and finish by discussing the possibility of using mPPases as a novel drug target.

2. mPPase Sub-Families

mPPase enzymes are divided into four groups based on their requirement for K^+, and the ion or ions that they pump across the membrane (summarized in Table 1). The first group consists of mPPases that translocate H^+ across the membrane in a K^+-independent manner [17]. Second, are mPPases that pump H^+ across the membrane, but require 30 to 50 mM K^+ for optimal activity [18]. Dependence on K^+ for function is due to the presence or absence of a key lysine residue in the active site; in K^+-independent mPPases the lysine side chain functionally replaces the K^+ ion. This was demonstrated with the K^+-dependent, H^+-pumping mPPase from *Carboxydotherums hydrogenoformans*: this enzyme was converted to a K^+-independent mPPase when the alanine 460 residue, which is at the position of the conserved lysine in K^+-independent mPPases, was mutated to lysine [19]. The structural basis is not yet certain, as there are no structures of a K^+-independent

mPPase, but modeling suggests that it is a direct structural replacement of the K^+ by the lysine sidechain [11].

Table 1. mPPase protein subfamilies.

Subfamily	K^+ requirement	Example organisms
H^+-pumping, K^+-independent	No	*Streptomyces coelicolor*, *Plasmodium falciparum* VP2
H^+-pumping, K^+-dependent	Yes	*Plasmodium falciparum* VP1, *Trypanosoma cruzii*, *Vigna radiata**
Na^+-pumping	Yes	*Thermotoga maritima**, *Clostridium tetani*
Na^+ and H^+-pumping	Yes	*Bacteroides vulgatus*

*The structure has been solved for the protein from these organisms [11,12].

The remaining two mPPase groups are K^+-dependent: Na^+-pumping mPPases (Na^+-mPPases) and finally, dual-pumping mPPases that translocate both Na^+ and H^+ (Na^+/H^+-mPPases) [13,20]. Phylogenetic analysis suggests Na^+-mPPases are the ancestral proteins with H^+-pumping having evolved *via* four separate evolutionary lineages [20,21]. Dual Na^+/H^+-mPPases likely evolved from Na^+-pumping proteins *via* the acquisition of key mutations that are conserved in the subfamily, *i.e. Bacteroides vulgatus* residues Thr90 (which can also be a serine in other species), Phe94, Asp146, and Met176 [13,22].

3. Biological Roles of mPPases

Though sPPases are primarily responsible for regulating intracellular PP_i levels, the ability of mPPases to translocate H^+ and/or Na^+ across a membrane against a concentration gradient enables it to play a variety of roles in different organisms. Dual-pumping Na^+/H^+-mPPases are found only in bacteria and Na^+-mPPases are found in both bacteria and archaea. H^+-pumping mPPases are the most widely distributed, being present in all three domains of life [19–21].

3.1. Plants

H^+-pumping mPPases are present in all members of the green line of evolution where the chloroplast was gained, which includes algae and land plants [15,23,24]. These enzymes couple cleavage of PP_i with pumping of H^+ across the vacuolar membrane, and together with the vacuolar ATPase (V-ATPase), acidify the vacuole, which is critical for maintaining its multiple functions, including solute and protein storage [17,25–29]. mPPases in plants play important roles in development and stress-resistance [30,31]. Overexpression of the mPPase AVP1 in *Arabidopsis* results in augmented root growth and larger leaf size and AVP1 loss of function leads to stunted growth and development [30]. Overexpression of mPPase in various species of plants also improves resistance to drought and salt-stress [32–35]. It is thus possible to develop stress-tolerant crops through overexpressing or manipulating mPPases, which could help offset the declines in

agricultural productivity and worldwide food security [36] that are expected, as global warming increases the incidence of drought [37].

3.2. Protozoan parasites

Protozoan parasites are unicellular eukaryotes that are responsible for numerous human diseases, including malaria (caused by *Plasmodium spp*.), toxoplasmosis (caused by *Toxoplasma gondii*), trypanosomiasis (caused by *Trypanosoma spp*.) and leishmaniasis (caused by *Leishmania spp*.) [38]. *Plasmodium spp*. and *T. gondii* are protozoan parasites from phylum apicomplexa, whereas *Trypanosoma spp*. and *Leishmania spp*. are from class kinetoplastida. Apicomplexans and kinetoplastids differ in two ways. Apicomplexan parasites have a unique organelle called the apicoplast at their cell tip that is essential for penetrating the host cell, and, even though they do not have flagella or cilia, most apicomplexan cells are motile, as they utilize a gliding mechanism [39]. In contrast, kinetoplastids have a single flagellum at the posterior end of their cell for motility and a kinetoplast, located near the basal body of the flagellum, which is a mitochondrial genome consisting of thousands of molecules of topologically interlocked circular DNA. Both apicomplexans and kinetoplastids have H^+-pumping mPPases, not found in humans, that are important for their survival during their life cycle (discussed below).

Malaria, toxoplasmosis, and leishmaniasis have disability adjusted life years (DALYs) in the millions with far-reaching global distributions [40]. African trypanosomiasis (*Trypanosoma spp*.) is similar: it blights the lives of ~60 million people in sub-Saharan Africa with ~0.6 million DALYs in 2010 [40]. Of these, malaria has the greatest worldwide impact, with the World Health Organization estimating 198 million cases and 584 000 deaths due to malaria in 2013 alone [41]. For the purpose of this review, we will use malaria as the prime example to illustrate the various transitions that occur throughout the protozoan parasite life cycle, though similar transitions occur in the other parasite species as well. Malaria is caused by several *Plasmodium* species, with *Plasmodium falciparum* being the most prevalent and virulent, and these unicellular parasites are transmitted between humans by *Anopheles* mosquito vectors [41,42]. To date, there are no effective, FDA-approved vaccines against malaria, presumably due to the complexities of malarial immune resistance [43]. Despite repeated exposure, most individuals living in areas where malaria is endemic are not immune to *Plasmodium* infection, though they appear asymptomatic and therefore remain untreated [43,44]. This population of individuals thus acts as a reservoir for the malaria parasite. Fighting malaria is further complicated by the emergence of *P. falciparum* strains, especially in south-east Asia [45], that are resistant to most of the currently-available drugs. The key to tackling the global malaria problem may lie in exploiting the transition between key stages of the complex life cycle of this host-dependent parasite.

The *Plasmodium* parasite has three major life cycle stages: the human liver stage, the human blood stage, and the mosquito stage [42,43]. The human liver stage begins when a female *Plasmodium*-infected *Anopheles* mosquito transfers the elongated and motile form of the parasite, the sporozoite, from the saliva into the human dermis while it feeds on human blood. These sporozoite parasites reach the bloodstream and travel to the liver, where they invade hepatocytes and replicate [46,47]. Inside the hepatocytes, the sporozoites differentiate into merozoites, an asexual

form of the parasite, which are released into the bloodstream, thereby beginning the human blood stage [47,48]. The merozoites, which are the main pathogenic form, then go on to infect erythrocytes, in which they replicate and produce more merozoite cells that are again released into the bloodstream. During the human blood stage, some merozoites also differentiate into male or female gametocytes, the non-pathogenic, sexual form of the parasite. The mosquito stage of the life cycle is initiated when mosquitoes take up the gametocyte form of *Plasmodium* when they feed on infected human blood. In the mosquito, the gametocytes combine in the midgut and produce new, asexual sporozoites, which replicate and spread to the salivary glands, where the parasitic cells can begin the cycle again when the infected mosquito feeds on and infects another human [42,43].

During the parasitic life cycle, the unicellular protozoan cell experiences transitions between environments with different levels of osmotic pressure, from intracellular to extracellular environments and from human to vector host environments. Even within the vector insect gut alone, the osmolarity can vary from 320 to 1000 mosmol/kg, depending on the feeding activities of the insect [49]. To survive the changes in environment, the parasites must quickly adjust their internal osmotic pressure [50]. One major mechanism for dealing with this osmotic stress involves a small organelle known as the acidocalcisome [51]. The acidocalcisome acts as an acidic storage compartment for numerous ions, including Ca^{2+}, polyphosphate and PP_i [52–55]. During osmotic stress, parasites are able to rapidly hydrolyze or synthesize polyphosphate to adapt to different osmotic environments and to release the energy stored in the polyphosphate molecule [56]. Polyphosphate chelates Ca^{2+} ions, and therefore the acidocalcisome likely plays a role in regulating Ca^{2+} levels in the protozoan cell [57]. Maintaining a low pH within the acidocalcisome is crucial for acidocalcisome function, and loss of acidocalcisome acidity leads to a 10 fold decrease in polyphosphate levels in the acidocalcisome, a loss of intracellular pH regulation, and decreased growth rate and final cell density in the protozoan parasite *Trypanosoma brucei* (the cause of African sleeping sickness) [58]. A high concentration of H^+ within the acidocalcisome is also important for regulating Ca^{2+} levels, since one mechanism of Ca^{2+} uptake into the acidocalcisome is the exchange of H^+ for Ca^{2+} by Ca^{2+}-ATPase [59]. Fusion of the acidocalcisome with the contractile vacuole complex and translocation of aquaporin, which is a water channel, to the contractile vacuole also plays a significant role in osmoregulation in protozoans. This cyclic-AMP- and microtubule-dependent fusion mediates the regulatory volume decrease process in *Trypanosoma cruzi* under hypo-osmotic stress conditions [60].

3.3. Are mPPases validated drug targets?

Just as in plant vacuoles, there are two types of proton pumps in protozoan acidocalcisomes: H^+-ATPases and H^+-pumping mPPases [61–63]. Knocking down H^+-pumping mPPase expression in *T. brucei* prevents acidocalcisome acidification, causes a 90% reduction in polyphosphate, and prevents stabilization of intracellular pH on exposure to external basic pH; thus, H^+-ATPase activity alone is *not* sufficient for acidocalcisome activity and a normal parasite [58]. In *T. gondii,* knockout mutants of mPPase display a reduction in the virulence of tachyzoites in mice; these are the fast-growing, asexual form of the parasite [64]. In addition, Rodrigues *et al.* showed that bisphosphonate derivatives (mPPase inhibitors) were able to significantly inhibit intracellular proliferation of *T.*

gondii in human foreskin fibroblasts without affecting the host cells [65]. Bisphosphonate derivatives also inhibit the growth of *P. falciparum*, *T. brucei*, *T. cruzi*, and *Leishmania donovani* with nanomolar or low-micromolar IC_{50} values [65–67]. However the mechanism of action of these bisphosphonates is complex, as they inhibit not only mPPase, but also farnesyl pyrophosphate synthase (FPPS), which occurs in mammalian cells as well [67]. Indeed, bisphosphonates are used to treat bone diseases [68]. Despite the fact that bisphosphonates are used to inhibit human enzymes, a recent study by Yang *et al.* suggests several lipophilic bisphosphonate molecules have the potential to be used against *T. brucei* in humans without harmful side effects [39,69,70]. Bisphosphonate compounds were screened for cytotoxicity in HEK293T cells and several of these molecules exhibited selectivity indexes (SI) (HEK293T cell 50% cytotoxicity concentration (CC_{50})/50% effective concentration (EC_{50}) against *T. brucei*) of greater than 50 [39]. Therefore, though most bisphosphonate compounds are *not* potential candidates as anti-parasitic drugs due to cytotoxicity against human cells, modification of these compounds could lead to decreased cytotoxicity while inhibiting parasitic enzymes.

Together, these studies highlight the importance of PP_i-driven proton transport by mPPases in protozoan parasites. They clearly indicate mPPases are viable potential drug targets against human protozoan parasites. By targeting the protozoan mPPase and thus interfering with pH regulation and the ability to adapt to varying osmolarity, protozoan pathogens may no longer survive the transition to different host environments.

3.4. mPPases in bacterial pathogens

mPPases are only present in select bacterial species; even within a single genus, some species have an mPPase gene whereas others do not. This is exemplified by the genus *Clostridium*, in which several species, such as *Clostridium leptum* and *Clostridium botulinum*, have Na^+/H^+ mPPases whereas other species, such as *Clostridium difficile* and *Clostridium perfringens*, lack mPPase genes entirely [13,21] (Shah, N. R. and Goldman, A. unpublished data). Bacterial species that possess mPPases that are of great importance to human health are *B. vulgatus* and *Bacteroides fragilis*, along with several other members of genus *Bacteroides* [13]. In healthy individuals, *Bacteroides* form a mutualistic relationship with the human host as part of the intestinal microbiota [71–74]. However, when these bacteria escape the gut environment, they can cause bacteraemia and abscesses in various places in the human body, including the brain [75–77]. *Bacteroides* are the most commonly isolated anaerobic pathogen in humans, and bacteremia caused by *Bacteroides* species has an associated mortality rate of over 19% [77,78]. Moreover, they have the highest resistance rates of all anaerobic pathogens, likely due to a wide variety of antibiotic resistance mechanisms [77,79,80], meaning that novel, *Bacteroides*-specific drugs would be of great use.

B. vulgatus has a dual-pumping Na^+/H^+-mPPase embedded in the inner membrane that hydrolyzes PP_i at the cytoplasmic face of the membrane and pumps ions across the inner membrane to the periplasm against a concentration gradient [13]. As in plants, overexpression of mPPase in bacteria leads to greater resistance to stressors, including heat, hydrogen peroxide, and high concentrations of NaCl [81]. Though mPPases are not essential for bacterial survival, inhibiting this enzyme will likely diminish the viability of *Bacteroides* cells, especially under stressful conditions.

Furthermore, in principle a molecule that could force the mPPase ion gate open, rather than just inhibiting the enzyme, would lead to a leaky inner membrane, collapse the electrochemical gradient driving ATP synthesis, and so cause bacterial cell death. Specifically targeting the mPPase in *B. vulgatus* would have the added bonus of only affecting bacteria that possess mPPases, thereby limiting disturbance of the gut flora, which can lead to antibiotic-associated diarrhea in 15-25% of antibiotic-treated patients. Colonization of the gut by the microbiota also plays a role in protection against pathogens, so targeting just a fraction of the microbiota could be beneficial [82]. For instance, a significant risk of antibiotic treatment is overgrowth of *C. difficile* that can lead to recurring *C. difficile*-induced colitis [83]. Targeting mPPase-containing organisms may not carry the same risk, as much of the microbiota would be preserved [84].

4. Structure and Function of mPPases

4.1. Overall structure of mPPases

Until now, only the mPPase structure from two species have been solved – one from mung bean (VrPPase) in the inhibitor (imidodiphosphate, IDP) bound state (VrPPase:IDP:Mg_5) [12] and one from *T. maritima* (TmPPase) in the resting (TmPPase:Ca:Mg) and product bound states (TmPPase:P_{i2}:Mg_4) [11]. VrPPase is a proton translocating enzyme and TmPPase is a sodium-translocating enzyme. Both structures were solved by X-ray crystallography. The overall structures of both enzymes are very similar (Cα's root mean square deviation (RMSD) 1.57 Å), with complete conservation of the position and identity of all catalytic residues [6]. Both structures consist of a two-fold symmetric homodimer in which each monomer has 16 TMHs (Figure 1A) [11,12]. In the TmPPase structure, the dimeric interface is mainly formed *via* hydrophobic interaction of TMH10, 13, and 15 and some hydrogen bonds [11] as well as an ion pair formed by the C-terminal carboxylate. In VrPPase, in addition to hydrophobic interactions and hydrogen bonds, the dimeric interface contains two salt bridges [12].

The TMHs of each monomer forms a rose-like arrangement consisting of two rings of helices (Figure 1B). The inner ring consists of six TMHs (TMH5-6, TMH11-12, and TMH15-16) and the outer ring consists of the other ten (TMH1-4, TMH7-10, and TMH13-14). Viewed from the cytoplasmic side, both rings are arranged sequentially counterclockwise with the exception of TMH1 [12]. However, the helices in the two rings have opposite tilts (Fig 1B), such that the inner ring of helices appears to be tilted "clockwise", and the outer ring, except for TMH1, "counterclockwise", viewed from the cytoplasm (Figure 1B). Viewed parallel to the membrane (data not shown), this corresponds to the inner ring helices being tilted on average -20° (counterclockwise) from the vertical, and the outer ring helices, except for TMH1, being tilted on average +19° (clockwise). This crossing angle between the helices is thus about 40°, similar to a classic "ridges-in-grooves" arrangement.

TMHs of the inner ring form the active site for the hydrolytic reaction and ion pumping (Figure 1C). The inner ring of helices is comprised of four distinct areas: (i) the hydrolytic center, (ii) the coupling funnel, (iii) the gates below the membrane surface, and (iv) the exit channel [11]. The hydrolytic center, about 20 Å above the membrane surface, is a funnel-shaped pocket with highly

conserved residues consisting of 12 acidic residues ($D^{5.61}$, $D^{5.65}$, $E^{5.76}$, $D^{5.77}$, $D^{6.35}$, $D^{6.39}$, $D^{6.43}$, $D^{11.57}$, $D^{15.61}$, $D^{16.31}$, $D^{16.35}$, $D^{16.39}$), three basic residues ($K^{5.58}$, $K^{15.64}$, $K^{16.38}$) and an asparagine ($N^{12.53}$) [11,12] (Ballesteros-Weinstein residue numbering as described in Tsai *et al.* [22]). These residues are important for substrate binding either directly or indirectly *via* Mg^{2+} or water. The coupling funnel consists of an ionic network of eight conserved charged residues on TMH5-6, TMH11-12, and TMH16 ($R^{5.50}$, $K^{5.58}$, $D^{6.39}$, $D^{6.50}$, $D^{11.50}$, $K^{12.50}$, $K^{16.38}$, $D^{16.39}$) and mutation of any of these residues has major effects on function [85]. Below the coupling funnel lie the ionic and hydrophobic gates. The ionic gate is comprised of the conserved-charged, and hydrophilic residues in TMH 5, 6, and 16 ($S^{5.43}$, $D^{6.50}$, $E^{6.53}$, $S^{6.53}$, $S^{6.54}$, $D^{16.46}$, $K^{16.50}$). The hydrophobic gate is located below the ionic gate, and is formed by semi-conserved non-polar residues from four TMHs of the inner barrel (TMH5, 6, 12, and 16). This gate likely has the essential function of preventing ion back-flow from the lumen/periplasm to the cytoplasm (6). Below the hydrophobic gate lies the exit channel, which facilitates ion release to the lumen/periplasmic region. Among mPPases, the residues in the exit channel are not conserved, which indicates this region does not play a significant role in the ion release mechanism.

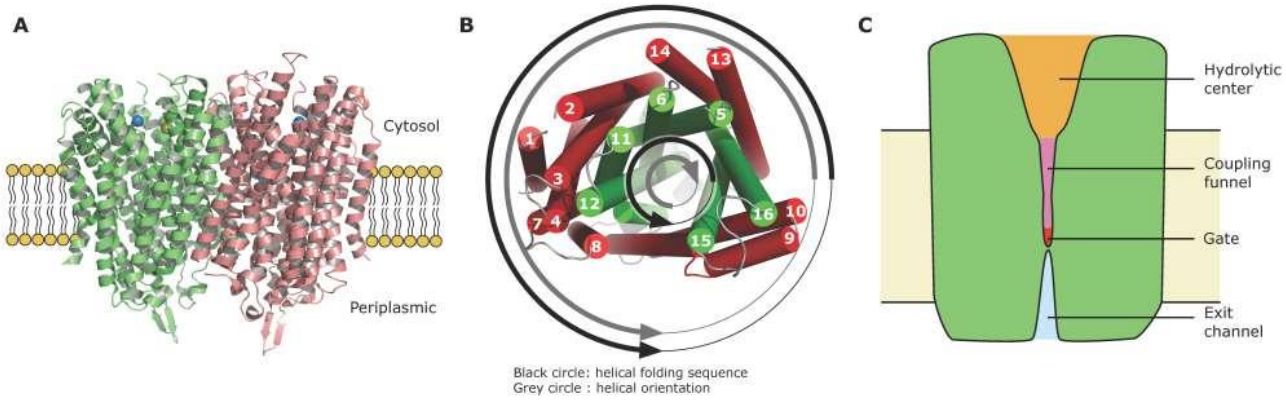

Figure 1. Structure of membrane bound pyrophosphatase. (A) The dimeric structure of TmPPase [11]. The different colors (green and pink) represent the two subunits. The blue and yellow spheres show the position of Mg^{2+} and Ca^{2+} ions in the hydrolytic center, respectively. (B) Inner (green) and outer (red) ring of TMHs drawn as cylindrical cartoons. Black circle represents helical folding sequence and grey circle represents the helical orientation. (C) Scheme of mPPase structure showing the hydrolytic center on the top (orange), the coupling funnel (pink), the gate (red), and the closed exit channel in the bottom (light blue).

4.2. Hydrolysis and pumping mechanism

Upon substrate binding, the most prominent conformational changes occur in TMH6, 11, and 12 (Figure 2A). In the resting state structure of TmPPase, the funnel-shaped pocket is open to the cytoplasmic environment with a pocket volume of 2581 $Å^3$ (Figure 2B, left panel). The conserved

loop 5-6 cannot be observed in this structure due to a lack of electron density, which is probably due to the flexibility of the loop. Upon substrate binding, the overall TMH structure in the cytoplasmic side constricts *via* TMH movement toward the center of the binding pocket (Figure 2A-B). This constriction leads to reduction of the pocket volume to 602 $Å^3$ (1,521 $Å^3$ [12]) based on the VrPPase:IDP:Mg_5 structure (Figure 2B, middle panel). In this structure, loop 5-6 is clearly ordered (Figure 2A, arrowhead). It blocks the active site and $E268^{5.75}$ binds the IDP indirectly *via* two water molecules [12]. The substrate is located and coordinated for enzymatic hydrolysis by five Mg^{2+} ions, one K^+, and the side chains of seven residues ($K250^{5.58}$, $D253^{5.61}$, $D691^{15.61}$, $K694^{15.64}$, $D727^{16.35}$, $K730^{16.38}$, and $D731^{16.39}$). The K^+ ion increases the k_{cat} over three-fold [86]. After hydrolysis and ion pumping has occurred, the capping of the funnel opens again, as seen in the TmPPase:Pi_2:Mg_4 structure (Figure 2B, right panel). In this conformation, the helices are still constricted toward the center of the binding site and there is a further decrease of the pocket volume to 407 $Å^3$. However in this structure, loop 5-6 is no longer observed, either due to conformational changes after hydrolysis that cause it to become flexible (to facilitate the pocket opening for phosphate release) or due to the low resolution of the structure (4.0 Å) [11].

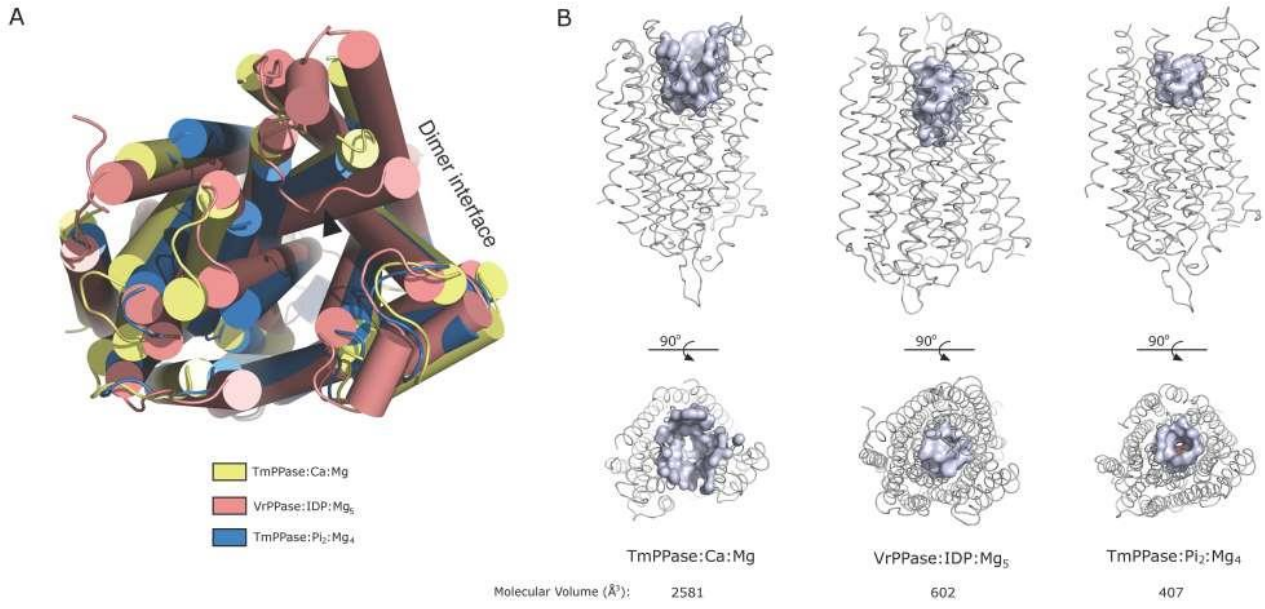

Figure 2. Movement of transmembrane helices in mPPase upon substrate binding. (A) Superposition of TmPPase structure in resting state (pale yellow), VrPPase structure in inhibitor bound state (salmon), and TmPPase structure in product bound state (sky blue). All structures were drawn as cylindrical cartoons. The arrowhead points at the loop 5-6 in the inhibitor bound state of VrPPase. (B) Cartoon representation of three mPPase structures showing the hydrolytic pocket (blue white) of each structure viewed parallel (upper panel) and perpendicular (lower panel) to the membrane. The molecular volume of each pocket was determined using the Splitpocket server (http://pocket.med.wayne.edu/patch/) [130,131].

Though there are some conformational changes in the TMHs on the cytoplasmic side, the overall TMH conformation on the lumen/periplasmic side of the enzyme is quite rigid. In the structures of both mPPase species, the gate and exit channel are closed, so the exact mechanism of H^+ or Na^+ pumping is still not clear [11,12]. Based on the structure of VrPPase:IDP:Mg_5, Lin *et al.* proposed a "Grotthuss-chain" mechanism to explain the proton translocation process from the cytoplasm to the vacuolar lumen. In this mechanism, the binding of PP_i to the funnel-shaped pocket induces the nucleophilic attack of the water molecule, which interacts with $D287^{6.39}$ and $D73^{16.39}$ *via* hydrogen bonds, to one of the phosphate groups of PP_i; this results in the breaking of the phosphoanhydride bond and produces two P_i and one H^+ ion. The H^+ ion is then translocated to the exit channel in the lumen through a Grotthuss-chain formed by a series of hydrogen bond-forming residues in the coupling funnel.

Since TmPPase transports Na^+ ion instead of H^+, a Grotthuss-chain mechanism is not possible. Therefore, Kellosalo *et al*. proposed a "binding change" mechanism [11]. In comparing the different structures of TmPPase and VrPPase, there is a 2 Å downward motion of TMH12 upon closing of the funnel-shaped pocket due to substrate binding. The movement of TMH12 is facilitated by the presence of a conserved flat hydrophobic surface ($A^{12.53}$, $I^{12.54}$, $A^{12.57}$, $A/I^{12.58}$) in the vicinity of the coupling funnel, which acts as a lubricant [11]. This movement probably drives opening of the gate and exit channel [11,87]. Mutation of $I545A^{12.54}$ and other residues in this region on VrPPase results in normal hydrolysis activity but a loss in H^+-pumping, which means there is uncoupling between PP_i hydrolysis and proton translocation [87]. Based on these data, Kellosalo *et al.* [11] proposed that PP_i binding to the active site induces a larger-scale conformational change than that seen in the static structures. In addition to the closing of the loop between TMH5 and 6, TMH12 would move down further than the 2 Å observed (see above); opening the gate and the exit channel and resulting in the diffusion of an ion to the extracellular medium. This transport increases the overall negative charge in the active site due to loss of the Na^+ (H^+ in the proton pumping enzymes) and in particular increase the negative charge on residues $D^{16.39}$ and $D^{6.43}$ (the residues that bind the nucleophilic water molecule). This would lead to proton abstraction and catalysis. After hydrolysis, the electrophilic phosphate dissociates from the enzyme, followed by the leaving-group phosphate. This mechanism has the advantage over the "Grotthus-chain" mechanism because it can explain how H^+-pumping and Na^+-pumping mPPases nonetheless have structurally-conserved active sites [11].

The importance of mPPase and the availability of structural data give this enzyme great practical interest for structure-based drug design methods to fight various human pathogens.

5. Structure-Based Drug Design for mPPase

Membrane proteins comprise about 25% of all proteins in the cell and more than 60% of them are drug targets [88,89]. Among drugs available in the DrugBank database [90] that target human membrane proteins, ~36% target G protein-coupled receptors (GPCRs) and ~8% target ligand-gated ion channels [89]. However, only a small fraction of drugs (~1%) in the database target parasitic organisms [89]. Therefore, since parasitic diseases have worldwide health, social, and economic impacts [40], new drugs targeting human parasites are urgently needed.

5.1. General requirements for drug discovery

For a compound to be a potential drug candidate, there needs to be evidence that it has a disease-modifying action. It is thus clearly important whether the target has been validated or not: as discussed above, mPPases are validated targets, both by RNAi experiments [58] and because of the effect of bisphosphonates [65–67]. The further research on the effect of small molecules that is nonetheless required would benefit from chemical probes (see below). Another important step is to establish a proof-of-concept that potent mPPase inhibitors would have an effect *ex vivo* and in animal models. Promising lead compounds targeted at the mPPase can be, for example, further explored using *P. falciparum* and *L. major* in *ex vivo* and mouse models.

A second critical aspect required for a clinical candidate molecule is to achieve a concentration-time profile in the body that is adequate for the desired efficacy and safety. This is done by optimizing, *via* synthesis of analogues, the compound profile with respect to its absorption, distribution, metabolism and excretion (ADME) properties. The most common tests are compound permeability (membrane permeation assays such as Caco2 or parallel artificial membrane permeability assay (PAMPA)), solubility and plasma protein binding (which will be determinants in compound bioavailability, and therefore the dose taken; higher doses means higher probability of adverse effects), and metabolic profile (cytochrome P450 (CYP) inhibition). In the case of mPPases, compounds would be orally absorbed and the site of action would be the blood. A favorable ADME profile is usually sought during the early stages of the project when compound libraries are prepared, by avoiding compounds likely to have poor absorption or permeability. The first published and best-known set of rules for design is Lipinski's "rule of five", meant to avoid poor absorption or permeability: no more than five hydrogen bond donors, no more than ten hydrogen bond acceptors, molecular mass less than 500 Da, and calculated Log P (lipophilicity of the compound) less than 5 [91].

The third important aspect is toxicity and adverse effects. The commonly tested toxicities are cardiac toxicity (human ether-à-go-go-related gene (hERG) channel inhibition) as well as hepatotoxicity and liver injury; the latter is, in some cases, a consequence of drug-drug interactions at transporters (permeability glycoprotein (P-gp), but also ATP-binding cassette sub-family C member 2 (ABCC2) and multi-antimicrobial extrusion protein (MATE)) and metabolic enzymes such as CYPs. Promiscuous binding, which may lead to adverse effects, is further tested against panels of common targets (GPCRs, kinases). In addition, tests are usually conducted on counter-targets, such as phylogenetically related proteins and proteins binding chemically similar ligands. For mPPases, mammals do not contain homologous proteins, making this enzyme a feasible target for drug design. Nonetheless, care should be taken because compounds targeted at phosphate sites are likely to bind promiscuously to other (pyro)phosphate binding proteins – such as FPPS (see above). The compounds developed should be designed to be specific by taking advantage of unique regions and aspects of the binding pocket.

5.2. Experimental screening

At the moment only bisphosphonates are known to inhibit the mPPase. The first step to start in any drug development project is to discover novel hit compounds, and in order to do so, a robust assay is needed to complement the virtual screening (below). In our case we have at our disposal an assay to determine the orthophosphate released from the mPPase activity based on the molybdenum blue reaction [92], which we adapted for high throughput screening (Vidilaseris, K. unpublished).

Hit discovery usually starts with a screening campaign. The libraries may be small to medium size to large size. A current trend in screening is towards miniaturization and automation of screening assays to reduce costs. An important factor is the druggability of the target, *i.e.* our ability to modulate it with small molecules [93,94]. In addition, the response towards several analogues is useful to see if trends in the structure-activity relationships may be discerned during screening.

Fragment-based drug discovery has received a lot of attention since the seminal paper of Shuker and coworkers [95]. The strategy is built around assembling small fragments, whose binding is identified by nuclear magnetic resonance (NMR), surface plasmon resonance (SPR), X-ray crystallography (soaking of fragment libraries into crystals), or by other biochemical methods. The crystallographic approaches require crystals that diffract to 2.5 Å or better, which is higher-resolution than all except one of the mPPase structures. For GPCRs, another class of membrane proteins, fragment based discovery has proven useful [96] even if NMR screening has required working with protein that is thermostabilized, solubilized and attached to a resin [97]. Such a screening campaign can benefit from the use of computational methods, as described below. One obstacle to the use of such methods is that virtual screening methods need known inhibitors as a control; these do not currently exist for mPPases, as bisphosphonates are so unique they do not work well as a positive control in virtual screening.

5.3. Computational screening

Computational approaches have become ever more important in modern drug discovery [98]. In virtual screening, each molecule from a virtual chemical library is docked into the binding pocket and then evaluated by estimating the binding affinity and the fitness within the pocket. Then, the obtained score is used to rank the compounds and select the best hits for experimental testing and biological activity assays. This method can improve the hit rate since only the compounds that are most likely to bind are screened with *in vitro* assays. It is therefore quicker and cheaper than experimental high throughput screening (HTS) for producing promising binders [99], and provides a useful first step in an experimental high-throughput pipeline. There are several phases in virtual screening: (i) preparation of the target and database, (ii) docking, scoring, and validation, and (iii) ranking the compounds for experimental testing [99].

Docking-based virtual screening methods are used to rank libraries based on docking scores. Docking screening usually uses a flexible ligand-rigid protein approach. Decoy compounds are used as controls similar to known active compounds; these can be selected from the Directory of Useful Decoys (DUD) for example [100]. The controls are used to find the optimal parameters to retrieve known compounds with a high rank, better than the inactive decoys. To evaluate the best set of

possible solutions from a sampling process, a scoring function is applied. Algorithms commonly used for this purpose include force-field-based scoring functions (calculating the binding energy as the sum of non-bonded interactions between the ligand and the binding site) [101], empirical functions (breaking down the binding energy into several components, multiplying them by a certain coefficient, and recombining them to produce a final score) [102], and knowledge-based scoring functions (statistical analysis of solved ligand-protein complex structures) [103]. DOCK [104], AutoDock [105], and GOLD [106] use force-field-based scoring functions, LUDI [107] and $PLANTS_{PLP}$ [108] use empirical scoring functions, and SMoG2001 [109] uses knowledge-based scoring functions.

Pharmacophore-based virtual screening methods find compounds that match a set of features with a certain spatial organization. It does not provide a ranking (in its simplest form) but rather a yes/no classification. The features are mapped from existing compounds (shared features) or directly set from the binding site. Pharmacophore modeling usually generates conformations on-the-fly. The control group should possess all the features used to screen the compounds. The screening performance is usually assessed from the ability of the pharmacophore to retrieve known ligands and decoys (true positives and negatives, false positives and negatives); a standard method is to use receiver operating characteristic curves [110].

An alternative approach, which overcomes the limitations of virtual libraries, is computational *de novo* design. This can produce unlimited innovation in generating novel small molecule ligands that bind protein targets. Based on initial constraints generated from the input, this method is divided into two strategies: ligand-based (if there is no structure available) and receptor-based [111]. We discuss only the latter. In the receptor-based strategy, the binding site on the target protein is identified and used to generate shape constraints for ligand docking and ligand-receptor interactions (hydrogen bond, electrostatic, and hydrophobic interactions). Methods that can be used to derive interaction sites include rule-based and grid-based methods [111]. The outcome of both is a map of the protein target displaying sites where a ligand might interact favorably. The other method, Multiple Copy Simultaneous Search (MCSS), can determine energetically favorable ligand positions and predict the orientations of the ligand functional groups in the binding site [112]. After randomized positioning of multi-copies of functional groups in the binding pocket, force field minimization is performed and the fragments with the interaction energy below a set threshold are chosen as starting fragments for filtering and further assembly.

For ligand assembly within the binding site, single atoms or fragments can be used as the basic building blocks. The former approach generates more structural variety but increases the number of probable solutions, leads to longer processing times and so makes it more difficult to find suitable candidates. The fragment based-building block method is thus more common nowadays as it reduces the size of the search space significantly, making it faster to find solutions with chemically stable, easy-to-make, drug-like compounds. Further ligand assembly can be performed using *e.g.* the linking method or the growing method. In the linking method, the fragments are placed in key interaction sites of the binding site and then connected to each other using linkers to create complete molecules. The growing method starts with the placement of a single fragment at a key interaction site of the binding site. The fragment is then grown to introduce proper interactions between the ligand and both the key interaction sites and regions between key interaction sites in the binding site.

LigBuilder [113] and SPROUT [114] are two programs that use both methods for ligand assembly, the latter of which has been used for *de novo* ligand design to transmembrane regions of membrane-bound proteins [115].

5.4. Hit expansion and lead optimisation

Once initial hits have been identified by experimental or virtual screening (above), the library must be expanded (hit expansion). This can be done by building focused compound libraries, for example enriched in phosphate bioisosteres. As reviewed elsewhere [116,117], such compounds include phosphonates, carboxylates and malonates, sulfamates, squaramides and squaric acids, boron-containing fragments, and phosphorothioates. Hit expansion usually starts with the testing of analogues or compounds sharing similar properties. The binding site structure can be used to guide the synthesis from X-ray crystallography or from docking studies, *i.e*. structure-based compound design. The most promising compounds are then optimized through an extensive synthesis program, usually investigating the effect of bioiosteric replacement or changes in the main scaffold (scaffold hopping) combined with a battery of tests *in vitro* and *in vivo*. Current best practice includes ADME testing [91] as early as possible, as a "mistake" of the field has been a too early strong focus on potency [118]. Hit compounds will be a starting point, but they can also be used as chemical probes to verify the tractability of the target [119]. Chemical probes may also be critical for the development of cheaper or more robust screening assays.

Two examples exemplify the possibilities for discovering inhibitors of the mPPases: FPPS and the deoxycytidine triphosphate pyrophosphatase 1 (dCTPase). FPPS is an important drug target that, like mPPase, binds bisphosphonates. Inhibitors of FPPS have been derived from natural ligands [120] and have been pursued from NMR screening of 400 fragments followed by X-ray crystallography [121]. Another approach for this enzyme used virtual screening of a small library of one thousand compounds followed by two rounds of analogue similarity search of the hits, identifying new classes of FPPS inhibitors, diterpenoids and sesquiterpenoids. This was followed by X-ray crystallography to show that these compounds bind to the isopentenyl diphosphate (IPP) site [122] or to a snapshot conformation found in molecular dynamics simulations [123]. Overall, these studies have identified that FPPS can be inhibited by binding at three different sites: the acid-rich catalytic site involved in the binding of three Mg^{2+}, the homoallylic site IPP site, a highly positively charged binding pocket and a modulatory site. dCTPase, which regulates the intracellular nucleotide pool through hydrolytic degradation of canonical and noncanonical nucleotide triphosphates (dNTPs), is highly expressed in many cancers. For dCTPase, a collection of 5,500 compounds was screened using a classical high-throughput-adapted Malachite Green assay [124], a test for the colorimetric detection of inorganic phosphate and the hits optimized using a classical medicinal chemistry approach [125].

5.5. Tractability of the mPPase binding site

The importance of mPPases, the availability of its structure and the fact that mammals do not contain homologous proteins make this enzyme a feasible target for drug design. Though the

structure of an mPPase from a pathogenic organism has not yet been solved, structure-based drug design against this group of enzymes is possible using the available solved structure from *T. maritima* and mung bean, since they have high sequence identity (40 - 43% and 44 - 53% between mPPases from pathogens *P. falciparum*, *T. gondii*, *T. brucei*, *Leishmania mexicana*, and *B. vulgatus* with TmPPase and VrPPase, respectively) with the highest sequence identity in the catalytic regions. Moreover, even though TmPPase and VrPPase translocate different ions and have 42% sequence identity, the active site of both structures are very similar with a root mean square deviation/Cα of 0.84 Å (based on the structure of VrPPase:IDP:Mg_5 and TmPPase:P_{i2}:Mg_4), despite the fact that one is in the resting state and the other has inhibitor bound. This demonstrates the low flexibility of mPPase in the active site, even between different conformations and different species. mPPases from pathogens will also have a similar structure.

Some PPase inhibitors, such as IDP and bisphosphonate derivatives, inhibit mPPase more than sPPases [6,69]. This is because, even though both sPPase and mPPase hydrolyze PP_i, their active site and hydrolytic mechanism are completely different; it is therefore possible to develop drugs that specifically target mPPase but not sPPase. The structural differences and mechanisms of both enzymes were comprehensively explained in Kajander *et al*. [6].

The first step in designing drugs against mPPase is identifying ligand-binding sites. The most obvious is the catalytic pocket on the cytoplasmic side of the membrane: competitive inhibitors, like IDP and bisphosphonates, inhibit mPPase activity by competitively binding there, and so these molecules can be used as initial leads for finding other compounds with drug-like properties that specifically target mPPase. In the structure of VrPPase with inhibitor bound (VrPPase:IDP:Mg_5) [12], IDP is bound in the acidic environment of a tunnel-like pocket, specifically to the side chain of residues $K250^{5.58}$, $D253^{5.61}$, $D691^{15.61}$, $K694^{15.64}$, $D727^{16.35}$, $K730^{16.38}$, and $D731^{16.39}$. Upon IDP binding the overall hydrolytic pocket is constricted compared to the open conformation of the TmPPase in the resting state (Figure 2). Inhibition assays of VrPPase show that aminomethylenediphosphonate (AMDP) has a six times greater inhibitory effect on enzyme activity than IDP [126], but the opposite is true for the sPPase from *S. cerevisiae* (ScPPase) (53). Analysis of the ScPPase structures in the resting (pdb code: 1HUJ) and PP_i bound states (pdb code: 1E6A) showed that the volume of the pocket is smaller than in mPPase [127]. The greater inhibition of mPPase by AMDP than IDP might be due to the larger pocket size, especially in the resting state structure. Therefore, in order to achieve better fitting and specificity for binding to the hydrolytic pocket of mPPase, *de novo* design by growing/linking methods can be employed based on the AMDP molecule. The initial hits identified from the structure-based design approaches described above could be further refined using a combination of compound synthesis and screening to increase on-target potency and selectivity against mPPase from human pathogens. The best inhibitor could also be co-crystallized with the mPPase for structure determination to analyze the binding mode of the inhibitor. Developing or identifying new compounds that bind selectively to the active site of mPPases and inhibit activity should lead to novel treatment options against several human pathogens.

6. Conclusion and Perspectives

A major hurdle for drug discovery against human pathogens is ensuring minimal toxicity towards the host; even antimicrobials that have been approved for use in humans are continuously monitored for evidence of toxicity [128,129]. An important strategy to minimize the chance of harming human cells is to target biological components that are not present in animals. In this respect, since animals lack an mPPase gene, mPPases are a credible, novel target for drug design against human pathogens that express this enzyme, including several species of protozoan parasites and prokaryotic *Bacteroides* species. Of the mPPase-expressing human pathogens, malaria, which is caused by protozoan parasites, is arguably the most devastating globally, especially considering that 78% of deaths due to malaria occur in children under the age of 5 [45].

mPPases play a critical role in stress tolerance and, in protozoa particularly, in adapting to osmotic stress. mPPases appear to be crucial for the various transitions in environment and protozoan cell type that occur over the complex parasitic life cycle. The solved structures of two mPPases, the H^+-pumping, K^+-dependent *V. radiata* mPPase and the Na^+-pumping *T. maritima* mPPase, are a good basis for structure-based drug design, either to develop novel molecules or repurpose known molecules against mPPases. In addition to the static structural information gained through crystallographic means, complementary dynamics experiments and modeling can contribute to the drug-discovery process by providing further information regarding conformational changes during the catalytic cycle of the enzyme.

However, additional mPPase structures, particularly of a protozoan H^+-pumping mPPase and a Na^+/H^+ dual-pumping mPPase, as found in *Bacteroides* species, would provide valuable information for specifically targeting human pathogens that possess these types of enzymes. Although there are many questions still to be answered in the field, the current state of structural knowledge and the experiments underway in our and other laboratories around the world mean that mPPase is an intriguing candidate for developing precision drugs – ones with a unique target, and so with a reduced chance of side-effects.

Conflict of Interest

All authors declare no conflict of interest in this paper.

References

1. Cooperman BS, Baykov AA, Lahti R (1992) Evolutionary conservation of the active site of soluble inorganic pyrophosphatase. *Trends Biochem Sci* 17: 262–266.
2. Lahti R (1983) Microbial inorganic pyrophosphatases. *Microbiol Rev* 47: 169–178.
3. Klemme JH (1976) Regulation of intracellular pyrophosphatase-activity and conservation of the phosphoanhydride-energy of inorganic pyrophosphate in microbial metabolism. *Z Naturforsch C* 31: 544–550.
4. Heinonen JK (2001) Biological role of inorganic pyrophosphate: Springer Science & Business Media.

5. Terkeltaub RA (2001) Inorganic pyrophosphate generation and disposition in pathophysiology. *Am J Physiol Cell Physiol* 281: C1–C11.
6. Kajander T, Kellosalo J, Goldman A (2013) Inorganic pyrophosphatases: one substrate, three mechanisms. *FEBS Lett* 587: 1863–1869.
7. Chen J, Brevet A, Fromant M, et al. (1990) Pyrophosphatase is essential for growth of *Escherichia coli*. *J Bacteriol* 172: 5686–5689.
8. Lundin M, Baltscheffsky H, Ronne H (1991) Yeast PPA2 gene encodes a mitochondrial inorganic pyrophosphatase that is essential for mitochondrial function. *J Biol Chem* 266: 12168–12172.
9. Kankare J, Salminen T, Lahti R, et al. (1996) Structure of *Escherichia coli* inorganic pyrophosphatase at 2.2 A resolution. *Acta Crystallogr D* 52: 551–563.
10. Arutiunian E, Terzian S, Voronova A, et al. (1981) X-Ray diffraction study of inorganic pyrophosphatase from baker's yeast at the 3 angstroms resolution (Russian). *DoklAkadNauk Sssr* 258: 1481.
11. Kellosalo J, Kajander T, Kogan K, et al. (2012) The structure and catalytic cycle of a sodium-pumping pyrophosphatase. *Science* 337: 473–476.
12. Lin SM, Tsai JY, Hsiao CD, et al. (2012) Crystal structure of a membrane-embedded H^+-translocating pyrophosphatase. *Nature* 484: 399–403.
13. Luoto HH, Baykov AA, Lahti R, et al. (2013) Membrane-integral pyrophosphatase subfamily capable of translocating both Na^+ and H^+. *P Natl Acad Sci U S A* 110: 1255–1260.
14. Garcia-Contreras R, Celis H, Romero I (2004) Importance of *Rhodospirillum rubrum* H^+-pyrophosphatase under low-energy conditions. *J Bacteriol* 186: 6651–6655.
15. Serrano A, Pérez-Castiñeira JR, Baltscheffsky M, et al. (2007) H^+-PPases: yesterday, today and tomorrow. *IUBMB Life* 59: 76–83.
16. Baykov AA, Malinen AM, Luoto HH, et al. (2013) Pyrophosphate-fueled Na^+ and H^+ transport in prokaryotes. *Microbiol Mol Biol Rev* 77: 267–276.
17. Taiz L (1992) The Plant Vacuole. *J Exp Biol* 172: 113–122.
18. Maeshima M, Yoshida S (1989) Purification and properties of vacuolar membrane proton-translocating inorganic pyrophosphatase from mung bean. *J Biol Chem* 264: 20068–20073.
19. Belogurov GA, Lahti R (2002) A lysine substitute for K^+. A460K mutation eliminates K^+ dependence in H^+-pyrophosphatase of *Carboxydothermus hydrogenoformans*. *J Biol Chem* 277: 49651–49654.
20. Luoto HH, Belogurov GA, Baykov AA, et al. (2011) Na^+-translocating Membrane Pyrophosphatases Are Widespread in the Microbial World and Evolutionarily Precede H^+-translocating Pyrophosphatases. *J Biol Chem* 286: 21633–21642.
21. Luoto HH, Nordbo E, Malinen AM, et al. (2015) Evolutionarily divergent, Na^+-regulated H^+-transporting membrane-bound pyrophosphatases. *Biochem J* 467: 281–291.
22. Tsai JY, Kellosalo J, Sun YJ, et al. (2014) Proton/sodium pumping pyrophosphatases: the last of the primary ion pumps. *Curr Opin Struct Biol* 27: 38–47.
23. Drozdowicz YM, Rea PA (2001) Vacuolar H^+ pyrophosphatases: from the evolutionary backwaters into the mainstream. *Trends Plant Sci* 6: 206–211.

24. Björn LO (2015) The evolution of photosynthesis and its environmental impact. *Photobiology*: Springer. pp. 207–230.
25. Kriegel A, Andres Z, Medzihradszky A, et al. (2015) Job Sharing in the Endomembrane System: Vacuolar Acidification Requires the Combined Activity of V-ATPase and V-PPase. *Plant Cell* 27: 3383–3396
26. Nakayasu T, Kawauchi, K., Hirata, H., et al. (1999) Cycloprodigiosin hydrochloride inhibits acidification of the plant vacuole. *Plant Cell Physiol* 40: 143–148.
27. Bethmann B, Thaler M, Simonis W, et al. (1995) Electrochemical Potential Gradients of H^+, K^+, Ca^{2+}, and Cl^- across the Tonoplast of the Green Alga *Eremosphaera Viridis*. *Plant Physiol* 109: 1317–1326.
28. Schumaker KS, Sze H (1990) Solubilization and reconstitution of the oat root vacuolar H^+/Ca^{2+} exchanger. *Plant Physiol* 92: 340–345.
29. Jiang L, Phillips TE, Hamm CA, et al. (2001) The protein storage vacuole: a unique compound organelle. *J Cell Biol* 155: 991–1002.
30. Li J, Yang H, Peer WA, et al. (2005) Arabidopsis H^+-PPase AVP1 regulates auxin-mediated organ development. *Science* 310: 121–125.
31. Sabatini S, Beis D, Wolkenfelt H, et al. (1999) An auxin-dependent distal organizer of pattern and polarity in the Arabidopsis root. *Cell* 99: 463–472.
32. Brini F, Hanin M, Mezghani I, et al. (2007) Overexpression of wheat Na+/H+ antiporter TNHX1 and H+-pyrophosphatase TVP1 improve salt- and drought-stress tolerance in Arabidopsis thaliana plants. *J Exp Bot* 58: 301–308.
33. Lv SL, Lian LJ, Tao PL, et al. (2009) Overexpression of *Thellungiella halophila* H^+-PPase (TsVP) in cotton enhances drought stress resistance of plants. *Planta* 229: 899–910.
34. Li X, Guo C, Gu J, et al. (2014) Overexpression of VP, a vacuolar H^+-pyrophosphatase gene in wheat (*Triticum aestivum* L.), improves tobacco plant growth under Pi and N deprivation, high salinity, and drought. *J Exp Bot* 65: 683–696.
35. Park S, Li J, Pittman JK, et al. (2005) Up-regulation of a H^+-pyrophosphatase (H^+-PPase) as a strategy to engineer drought-resistant crop plants. *P Natl Acad Sci U S A* 102: 18830–18835.
36. Battisti DS, Naylor RL (2009) Historical warnings of future food insecurity with unprecedented seasonal heat. *Science* 323: 240–244.
37. Zhang H, Shen G, Kuppu S, et al. (2011) Creating drought- and salt-tolerant cotton by overexpressing a vacuolar pyrophosphatase gene. *Plant Signal Behav* 6: 861–863.
38. Yaeger R (1996) Protozoa: structure, classification, growth, and development.
39. Kappe SH, Buscaglia CA, Bergman LW, et al. (2004) Apicomplexan gliding motility and host cell invasion: overhauling the motor model. *Trends Parasitol* 20: 13–16.
40. Murray CJL, Vos T, Lozano R, et al. (2012) Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 380: 2197–2223.
41. World Health Organization (2013) WHO Global malaria report 2013. Geneva, Switzerland: World Health Organization.
42. Greenwood BM, Fidock DA, Kyle DE, et al. (2008) Malaria: progress, perils, and prospects for eradication. *J Clin Invest* 118: 1266–1276.

43. Crompton PD, Moebius J, Portugal S, et al. (2014) Malaria immunity in man and mosquito: insights into unsolved mysteries of a deadly infectious disease. *Annu Rev Immunol* 32: 157–187.
44. Smith T, Schellenberg JA, Hayes R (1994) Attributable fraction estimates and case definitions for malaria in endemic areas. *Stat Med* 13: 2345–2358.
45. World Health Organization (2014) WHO global malaria report 2014. Geneva, Switzerland: World Health Organization.
46. Shortt HE, Fairley NH, Covell G, et al. (1951) The pre-erythrocytic stage of *Plasmodium falciparum*. *Trans R Soc Trop Med Hyg* 44: 405–419.
47. Mota MM, Pradel G, Vanderberg JP, et al. (2001) Migration of *Plasmodium* sporozoites through cells before infection. *Science* 291: 141–144.
48. Sturm A, Amino R, van de Sand C, et al. (2006) Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science* 313: 1287–1290.
49. Kollien A, Meyer H, Großpietsch T, et al. (1998) Correlation of the development of *Trypanosoma cruzi* with the physiology in the rectum of the vector. *Parasitol Int* 47: 143.
50. Kollien A, Schaub G (2000) The development of *Trypanosoma cruzi* in triatominae. *Parasitol Today* 16: 381–387.
51. Docampo R, Jimenez V, Lander N, et al. (2013) New insights into roles of acidocalcisomes and contractile vacuole complex in osmoregulation in protists. *Int Rev Cell Mol Biol* 305: 69–113.
52. Docampo R, Scott DA, Vercesi AE, et al. (1995) Intracellular Ca^{2+} storage in acidocalcisomes of *Trypanosoma cruzi*. *Biochem J* 310: 1005–1012.
53. Garcia CR, Ann SE, Tavares ES, et al. (1998) Acidic calcium pools in intraerythrocytic malaria parasites. *Eur J Cell Biol* 76: 133–138.
54. Scott DA, Moreno S, Docampo R (1995) Ca^{2+} storage in *Trypanosoma brucei*: the influence of cytoplasmic pH and importance of vacuolar acidity. *Biochem J* 310: 789–794.
55. Moreno S, Zhong L (1996) Acidocalcisomes in *Toxoplasma gondii* tachyzoites. *Biochem J* 313: 655–659.
56. Ruiz FA, Rodrigues CO, Docampo R (2001) Rapid Changes in Polyphosphate Content within Acidocalcisomes in Response to Cell Growth, Differentiation, and Environmental Stress in *Trypanosoma cruzi*. *J Biol Chem* 276: 26114–26121.
57. Ruiz FA, Marchesini N, Seufferheld M, et al. (2001) The polyphosphate bodies of *Chlamydomonas reinhardtii* possess a proton-pumping pyrophosphatase and are similar to acidocalcisomes. *J Biol Chem* 276: 46196–46203.
58. Lemercier G, Dutoya S, Luo S, et al. (2002) A vacuolar-type H^+-pyrophosphatase governs maintenance of functional acidocalcisomes and growth of the insect and mammalian forms of *Trypanosoma brucei*. *J Biol Chem* 277: 37369–37376.
59. Vercesi AE, Moreno S, Docampo R (1994) Ca^{2+}/H^+ exchange in acidic vacuoles of *Trypanosoma brucei*. *Biochem J* 304: 227–233.
60. Rohloff P, Montalvetti A, Docampo R (2004) Acidocalcisomes and the contractile vacuole complex are involved in osmoregulation in *Trypanosoma cruzi*. *J Biol Chem* 279: 52270–52281.
61. Scott DA, de Souza W, Benchimol M, et al. (1998) Presence of a plant-like proton-pumping pyrophosphatase in acidocalcisomes of *Trypanosoma cruzi*. *J Biol Chem* 273: 22151–22158.

62. Marchesini N, Luo S, Rodrigues C, et al. (2000) Acidocalcisomes and a vacuolar H^+-pyrophosphatase in malaria parasites. *Biochem J* 347: 243–253.
63. Lu H-G, Zhong L, de Souza W, et al. (1998) Ca^{2+} content and expression of an acidocalcisomal calcium pump are elevated in intracellular forms of *Trypanosoma cruzi*. *Mol Cell Biol* 18: 2309–2323.
64. Liu J, Pace D, Dou Z, et al. (2014) A vacuolar H^+ pyrophosphatase (TgVP1) is required for microneme secretion, host cell invasion, and extracellular survival of *Toxoplasma gondii*. *Molecular microbiology* 93: 698–712.
65. Rodrigues CO, Scott DA, Bailey BN, et al. (2000) Vacuolar proton pyrophosphatase activity and pyrophosphate (PPi) in *Toxoplasma gondii* as possible chemotherapeutic targets. *Biochem J* 349 Pt 3: 737–745.
66. Martin MB, Grimley JS, Lewis JC, et al. (2001) Bisphosphonates Inhibit the Growth of *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania donovani*, *Toxoplasma gondii*, and *Plasmodium falciparum*: A Potential Route to Chemotherapy. *J Med Chem* 44: 909–916.
67. Martin MB, Sanders JM, Kendrick H, et al. (2002) Activity of bisphosphonates against *Trypanosoma brucei rhodesiense*. *J Med Chem* 45: 2904–2914.
68. Rodan GA, Martin TJ (2000) Therapeutic approaches to bone diseases. *Science* 289: 1508–1514.
69. Baykov AA, Dubnova EB, Bakuleva NP, et al. (1993) Differential sensitivity of membrane-associated pyrophosphatases to inhibition by diphosphonates and fluoride delineates two classes of enzyme. *FEBS Lett* 327: 199–202.
70. Smirnova IN, Kudryavtseva NA, Komissarenko SV, et al. (1988) Diphosphonates are potent inhibitors of mammalian inorganic pyrophosphatase. *Arch Biochem Biophys* 267: 280–284.
71. Holdeman L, Good I, Moore W (1976) Human fecal flora: variation in bacterial composition within individuals and a possible effect of emotional stress. *Appl Environ Microbiol* 31: 359–375.
72. Salyers A (1984) Bacteroides of the human lower intestinal tract. *Annu Rev Microbiol* 38: 293–313.
73. Xu J, Gordon JI (2003) Honor thy symbionts. *P Natl Acad Sci USA* 100: 10452–10459.
74. Bäckhed F, Ley RE, Sonnenburg JL, et al. (2005) Host-bacterial mutualism in the human intestine. *Science* 307: 1915–1920.
75. Brook I, Johnson N, Overturf GD, et al. (1977) Mixed bacterial meningitis: A complication of ventriculo-and lumboperitoneal shunts: Report of two cases. *J Neurosurg* 47: 961–964.
76. Odugbemi T, Jatto S, Afolabi K (1985) *Bacteroides fragilis* meningitis. *J Clin Microbiol* 21: 282–283.
77. Wexler HM (2007) *Bacteroides*: the good, the bad, and the nitty-gritty. *Clin Microbiol Rev* 20: 593-621.
78. Goldstein EJ (1996) Anaerobic bacteremia. *Clin Infect Dis* 23: S97–S101.
79. Hecht DW (2004) Prevalence of antibiotic resistance in anaerobic bacteria: worrisome developments. *Clin Infect Dis* 39: 92–97.
80. Nguyen MH, Victor LY, Morris AJ, et al. (2000) Antimicrobial resistance and clinical outcome of *Bacteroides* bacteremia: findings of a multicenter prospective observational trial. *Clin Infect Dis* 30: 870–876.

81. Yoon H-S, Kim S-Y, Kim I-S (2013) Stress response of plant H^+-PPase-expressing transgenic *Escherichia coli* and *Saccharomyces cerevisiae*: a potentially useful mechanism for the development of stress-tolerant organisms. *J Appl Genet* 54: 129–133.
82. Rafii F, Sutherland JB, Cerniglia CE (2008) Effects of treatment with antimicrobial agents on the human colonic microflora. *Ther Clin Risk Manag* 4: 1343.
83. Owens RC, Donskey CJ, Gaynes RP, et al. (2008) Antimicrobial-associated risk factors for *Clostridium difficile* infection. *Clin Infect Dis* 46: S19–S31.
84. Hopkins M, Macfarlane G (2002) Changes in predominant bacterial populations in human faeces with age and with *Clostridium difficile* infection. *J Med Microbiol* 51: 448–454.
85. Gaxiola RA, Palmgren MG, Schumacher K (2007) Plant proton pumps. *FEBS Lett* 581: 2204–2214.
86. Maeshima M (2000) Vacuolar H^+-pyrophosphatase. *Biochim Biophys Acta* 1465: 37–51.
87. Asaoka M, Segami S, Maeshima M (2014) Identification of the critical residues for the function of vacuolar H^+-pyrophosphatase by mutational analysis based on the 3D structure. *J Biochem* 156: 333–344.
88. Terstappen GC, Reggiani A (2001) *In silico* research in drug discovery. *Trends Pharmacol Sci* 22: 23–26.
89. Rask-Andersen M, Almen MS, Schioth HB (2011) Trends in the exploitation of novel drug targets. *Nat Rev Drug Discov* 10: 579–590.
90. Wishart DS, Knox C, Guo AC, et al. (2006) DrugBank: a comprehensive resource for *in silico* drug discovery and exploration. *Nucleic Acids Res* 34: D668–672.
91. Lipinski CA, Lombardo F, Dominy BW, et al. (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliver Rev* 23: 3–25.
92. Biologic E, Foa P, Zak B (1967) Microdetermination of inorganic phosphate, phospholipids, and total phosphate in biologic materials. *Clin Chem* 13: 326–332.
93. Hajduk PJ, Huth JR, Fesik SW (2005) Druggability indices for protein targets derived from NMR-based screening data. *J Med Chem* 48: 2518–2525.
94. Hajduk PJ, Huth JR, Tse C (2005) Predicting protein druggability. *Drug Discov Today* 10: 1675–1682.
95. Shuker SB, Hajduk PJ, Meadows RP, et al. (1996) Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 274: 1531–1534.
96. Andrews SP, Brown GA, Christopher JA (2014) Structure-Based and Fragment-Based GPCR Drug Discovery. *ChemMedChem* 9: 256–275.
97. Congreve M, Rich RL, Myszka DG, et al. (2011) 5 Fragment Screening of Stabilized G-Protein-Coupled Receptors Using Biophysical Methods. *Methods Enzymol* 493: 115.
98. Hawkins PCD (2006) A comparison of structure-based and shape-based tools for virtual screening. *Abstr Pap Am Chem S* 231.
99. Lyne PD (2002) Structure-based virtual screening: an overview. *Drug Discov Today* 7: 1047–1055.
100. Huang N, Shoichet BK, Irwin JJ (2006) Benchmarking sets for molecular docking. *J Med Chem* 49: 6789–6801.

101. Aqvist J, Luzhkov VB, Brandsdal BO (2002) Ligand binding affinities from MD simulations. *Acc Chem Res* 35: 358–365.
102. Bohm HJ (1998) Prediction of binding constants of protein ligands: A fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *J Comput Aid Mol Des* 12: 309–323.
103. Muegge I, Martin YC (1999) A general and fast scoring function for protein-ligand interactions: A simplified potential approach. *J Med Chem* 42: 791–804.
104. Kuntz ID, Blaney JM, Oatley SJ, et al. (1982) A geometric approach to macromolecule-ligand interactions. *J Mol Biol* 161: 269–288.
105. Morris GM, Goodsell DS, Halliday RS, et al. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19: 1639–1662.
106. Verdonk ML, Cole JC, Hartshorn MJ, et al. (2003) Improved protein-ligand docking using GOLD. *Proteins* 52: 609–623.
107. Bohm HJ (1992) The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J Comput Aid Mol Des* 6: 61–78.
108. Korb O, Stutzle T, Exner TE (2009) Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J Chem Inf Model* 49: 84–96.
109. Ishchenko AV, Shakhnovich EI (2002) SMall molecule growth 2001 (SMoG2001): An improved knowledge-based scoring function for protein-ligand interactions. *J Med Chem* 45: 2770–2780.
110. Triballeau N, Acher F, Brabet I, et al. (2005) Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *J of M Chem* 48: 2534–2547.
111. Schneider G, Fechner U (2005) Computer-based de novo design of drug-like molecules. *Nat Rev Drug Discov* 4: 649–663.
112. Miranker A, Karplus M (1991) Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins* 11: 29–34.
113. Wang R, Gao Y, Lai L (2000) LigBuilder: A Multi-Purpose Program for Structure-Based Drug Design. *J Mol Model* 6: 498–516.
114. Gillet V, Johnson AP, Mata P, et al. (1993) SPROUT: a program for structure generation. *J Comput Aid Mol Des* 7: 127–153.
115. Lolicato M, Bucchi A, Arrigoni C, et al. (2014) Cyclic dinucleotides bind the C-linker of HCN4 to control channel cAMP responsiveness. *Nat Chemical Biol* 10: 457–462.
116. Elliott TS, Slowey A, Ye Y, et al. (2012) The use of phosphate bioisosteres in medicinal chemistry and chemical biology. *MedChemComm* 3: 735–751.
117. Rye C, Baell J (2005) Phosphate isosteres in medicinal chemistry. *Curr Med Chem* 12: 3127–3141.
118. Manly CJ, Chandrasekhar J, Ochterski JW, et al. (2008) Strategies and tactics for optimizing the Hit-to-Lead process and beyond—A computational chemistry perspective. *Drug Discov Today* 13: 99–109.

119. Garbaccio RM, Parmee ER (2016) The Impact of Chemical Probes in Drug Discovery: A Pharmaceutical Industry Perspective. *Chem Biol* 23: 10–17.
120. Zhang Y, Zhu W, Liu Y-L, et al. (2013) Chemo-immunotherapeutic antimalarials targeting isoprenoid biosynthesis. *ACS Med Chem Lett* 4: 423–427.
121. Jahnke W, Rondeau J-M, Cotesta S, et al. (2010) Allosteric non-bisphosphonate FPPS inhibitors identified by fragment-based discovery. *Nat Chem Biol* 6: 660–666.
122. Liu Y-L, Lindert S, Zhu W, et al. (2014) Taxodione and arenarone inhibit farnesyl diphosphate synthase by binding to the isopentenyl diphosphate site. *P Nat Acad Sci U S A* 111: E2530–E2539.
123. Lindert S, Zhu W, Liu Y-L, et al. (2013) Farnesyl diphosphate synthase inhibitors from *in silico* screening. *Chem Biol Drug Des* 81: 742–748.
124. Itaya K, Ui M (1966) A new micromethod for the colorimetric determination of inorganic phosphate. *Clin Chim Acta* 14: 361–366.
125. Llona-Minguez S, Höglund A, Jacques S, et al. (2016) Discovery of the first potent and selective inhibitors of the human dCTP pyrophosphatase 1 (dCTPase). *J Med Chem* 59: 1140–1148.
126. Zhen RG, Baykov AA, Bakuleva NP, et al. (1994) Aminomethylenediphosphonate: A Potent Type-Specific Inhibitor of Both Plant and Phototrophic Bacterial H^+-Pyrophosphatases. *Plant Physiol* 104: 153–159.
127. Heikinheimo P, Tuominen V, Ahonen AK, et al. (2001) Toward a quantum-mechanical description of metal-assisted phosphoryl transfer in pyrophosphatase. *P Natl Acad Sci U S A* 98: 3121–3126.
128. Falagas ME, Kasiakou SK (2006) Toxicity of polymyxins: a systematic review of the evidence from old and recent studies. *Crit Care* 10: R27.
129. Avent M, Rogers B, Cheng A, et al. (2011) Current use of aminoglycosides: indications, pharmacokinetics and monitoring for toxicity. *Intern Med J* 41: 441–449.
130. Tseng YY, Dupree C, Chen ZJ, et al. (2009) SplitPocket: identification of protein functional surfaces and characterization of their spatial patterns. *Nucleic Acids Res* 37: W384–W389.
131. Tseng YY, Li WH (2009) Identification of protein functional surfaces by the concept of a split pocket. *Proteins* 76: 959–976.

AIMS Press