Discovery of membrane-bound pyrophosphatase inhibitors optimized from a fragment

*Niklas G. Johansson1,#, Ainoleena Turku1,#, Keni Vidilaseris2,#, Daniel Ayuso Pérez1, Aaron Wilkinson3, Olga Lukkonen1, Matti Tamminen1, Yuezhou Zhang1, Loïc Dreano1, Evgeni Grazhdankin1, Alexandros Kiriazis1, Colin Fishwick3, Jari Yli-Kauhaluoma1, Adrian Goldman2,4, Gustav Boije af Gennäs1, Henri Xhaard1,\**

1Drug Research Program, Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, University of Helsinki, Helsinki, Finland.

2Department of Biosciences, Division of Biochemistry, University of Helsinki, Helsinki, Finland.

3School of Chemistry and Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK.

4School of Biomedical Sciences and Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK.

# Equal contribution

\*Correspondence and requests for materials should be addressed to H.X. (e-mail: henri.xhaard@helsinki.fi)

# **ABSTRACT**

Membrane-bound pyrophosphatases (mPPases) are proteins, characterized by 17 transmembrane helices, present in many parasites such as *Plasmodium* (malaria) and *Leishmania* (leishmaniasis). They have a key role in the energy homeostasis of these parasites but no human homologue, thus making them ideal as drug targets. Here we aimed to target mPPases by a multistep sequential screening process, starting with phosphate isosteres. The best compound, **§11l**, has an IC50 of 7 µM (on *Thermotoga maritima* mPPase), and its SARs were characterized but did lead to improved activity. We verified the mechanism of action of our best hits using kinetics studies; our hits act non-competitively and present thus a novel series of allosteric modulators of the mPPase activity. Altogether, this study set the ground for further tracking of mPPase as a potential drug target in the treatment of protozoan diseases.

**Keywords:** membrane-bound pyrophosphatase, protist diseases, drug design, isoxazoles

# **INTRODUCTION**

One mutual aspect of many protist parasites, for instance *Plasmodium* spp., *Toxoplasma* spp., *Trypanosoma* spp. and *Leishmania* spp., is that they possess a family of enzymes called membrane-bound pyrophosphatases(Shah *et al*., 2016). These large (70–81 kDa) homodimeric integral membrane proteins consist of 15–17 transmembrane helices per monomer and are in parasites usually located in the plasma membrane of the acidocalcisome. Their assignment is hydrolysis of pyrophosphate, a by-product from many biosynthetic pathways that may disturb physiological reactions in too high concentrations, into orthophosphates. This process generates an ion gradient coupled with the active transport of H+ or Na+ across the membranes. By maintaining the charge on the acidocalcisome and thereby its function, mPPases likely play a significant role in the lifecycle of many parasites (Docampo *et al*., 2005; McIntosh *et al*., 2001). Although mPPases are essential for parasites, especially under conditions of low-energy stress, no homologous proteins exist in vertebrates. For example, humans have other ways of dealing with pyrophosphate such as ecto-nucleotide pyrophosphatases/phosphodiesterases, inorganic pyrophosphatases and polymerases (REF).

So far, only two X-ray structures of *Thermotoga maritima*(Kellosalo *et al*., 2012; Li *et al.*, 2016; Vidilaseris *et al*., 2018)and *Vigna radiata* (Lin *et al*., 2012; Li *et al*., 2016)mPPases have been solved. In addition, the fact that mPPases do not exist in humans but nonetheless are essential for parasites makes them promising potential drug targets against protist diseases. Most of the current compounds inhibiting the mPPase functions are non-hydrolysable pyrophosphate analogues (REF), and only one mPPase inhibitor class of non-pyrophosphate origin has been reported (Vidilaseris *et al*., 2019). To avoid pyrophosphate analogues from also targeting human enzymes bearing pyrophosphate binding sites, our interest lies in the discovery of non-phosphorous organic compounds that would inhibit the activity of protist mPPases. To attain this, we used a multistep sequential compound screening process, starting with phosphate isosteres and our recently developed TmPPase based assay. (Vidilaseris *et al.*, 2018).

Based on the preliminary structure-activity relationships of the best hit compounds, we developed more active analogues and subsequently ended up with the present three compounds, compounds **§4**, **§11l** and **§11m**, with an IC50 < 20 µM and a MW of 210–460 Da.

a goal is to get compounds with F for structural studies and small size to carry optimization

Overall we attempted to use computational docking to guide the synthesis, however it remains speculative first as to whether the hits actually target the active site, and second the binding poses of fragment showed not robust poses. The modelling will thus not be presented in this manuscript.

# **RESULTS AND DISCUSSION**

**Initial screening**

*Focussed bioisostere screening* – In our primary screen (leading to the found inhibitors) we assessed a small library of phosphate mimics possessing properties required to hypothetically match the pyrophosphate substrate binding site (*e.g.* high polarity). Having in mind a medicinal chemistry driven optimization we screened 22 commercially purchased, low molecular weight (MW, range 128–338 Da, average 220 Da) fragments(Supporting Information, Figure S1). The library included common phosphate isosteres such as boronate, sulfone, sulfonamide, and isoxazole functionalities (Zhang *et al*., 2017). The biological activity was assessed at 100 µM using an enzymatic inhibition assay against purified and solubilized XX protein (Vidilaseris *et al*., 2018). As a result four compounds presented over 50% inhibition (Figure 1), but for only three of them this result could be replicated and refined to provide an IC50 estimate (estimates of IC50 based on single experiments in 4–6 replicates for three concentration are later referred to as IC50estimate). Compound §**1**, §**2**, and §**3** all include a heterocyclic five-membered ring and an aryl moiety; §**1** (MW of 234 Da) has a 5-phenylisoxazole core with an IC50estimate of 65 µM, §**2** (MW of 243 Da) contains a sulfonyl substituted benzo[*d*]oxazolidinone moiety with an IC50estimate of 40 µM, and §**3** (MW of 328 Da) is a sulfamoylbenzene linked thiazolone with an IC50estimate of 50 µM.



**Figure 1.** Fragment hits obtained from the initial library.

*Catalogue exploration of the initial hit* **–** Among the three best hits, isoxazole §**1** seemed synthetically most accessible and was thus selected for further exploration. First we did some preliminary predictions by docking about 50 isoxazole-3-carboxylic acids with an altered 5-position. This suggested us to supplement the series with 12 analogues (4 purchased and 8 synthesized). Unfortunately most of the expansion did not yield any huge leap in activity. We nonetheless discovered compound **§4**, which was then used as a positive control throughout the assays in addition to imidodiphosphate (IDP) early in the discovery process. Compound **§4** is rather small (MW 210 Da) with an IC50 of 17 M (6***–***8 concentrations in 3 replicates), which provides it with a high ligand efficiency (LE 0.47) and makes it promising for further structural exploration. Unfortunately, the direct availability of analogues was low and substituting compound **§4** further (4 analogues purchased) did not either provide more active compounds (Supporting Information, Figure S1)*.*



**Figure 2***.* Exploring the aryl group changes. / Compound **§4** and its analogue **§5**

**Systematic exploration of fragment hit 1**

*Aryl substitutions* – In order to complete the former set of compounds, we continued to explore further variations of the phenyl moiety by synthesizing 10 5-arylisoxazole-3-carboxylic acid derivatives. Compounds (Supporting Information, Figure S1) were either synthesized from obtainable building blocks or directly purchased from commercial vendors. As illustrated in Scheme 1, the isoxazole series was constructed utilizing three different approaches. The isoxazole ring can *e.g.* be formed from substituted alkynes[REF] **§6** and ketones[REF] **§7**. However, to easily be able to vary the substituents of the 5-position, we were initially tempted to use isoxazoles functionalized with a synthetic handle, such as a halogen or a boronic acid/ester group that further could be used in Suzuki coupling reactions. As our obtained yields were rather low, we quite quickly restructured our approach utilizing the method used in making the key *N*-methyliminodiacetic acid (MIDA) boronate ester **§8**[Grob]. Ethyl 5-arylisoxazole-3-carboxylates **§9** were thus made from arylalkynes **§6** with ethyl 2-chloro-2-(hydroxyimino)acetate (forming nitrile oxide *in situ*) in a one-step [3+2] cycloaddition . Additionally we tried aldol condensation of arylketones **§7** with diethyl oxalate, followed by ring cyclization of the condensation products and hydroxylamine into ethyl 5-arylisoxazole-3-carboxylates **§9**. All the formed ethyl esters **§9a–o** were subsequently hydrolyzed to their corresponding carboxylic acids **§10a–o** with lithium hydroxide. Some of the obtained acids (**§10i–k**) were furthermore activated as acid chlorides before esterification to 2-bromophenyl 5-arylisoxazole-3-carboxylates **§11i–m**.

**Scheme 1. Synthesis of 3,5-disubstituted isoxazolesa** aReagents and conditions: (a) Ethyl-2-chloro-2-(hydroxyimino)acetate, NaHCO3, MeCN, rt; (b) NaOEt, diethyl oxalate, EtOH:Et2O, rt,; (c) NH2OH∙HCl, EtOH, reflux; (d)[Grob] ArBr, X Phos Pd G1, K2CO3, KO*t*-Bu, MeCN:*i*-PrOH, MW, 120 °C, 20 min; (e) LiOH, EtOH:H2O, rt,, then acidification (H+); (f) (i) (COCl)2, DMF, DCM, rt, (ii) *o*-Bromophenol, Et3N, 0 °C → rt.

We first demonstrated that **§10a**, with an unsubstituted phenyl ring, possessed no inhibitory activity towards TmPPase. Substitution of the phenyl ring with a polar nitro (as in **1)**, charged carboxylic acid (as in **§10c)**, or aliphatic *t*-butyl (as in **§10m**) moiety led to activities between 40 and 60 M. Favourable modifications might origin from compensatory interactions or a property of the fragment to flip around and are thus difficult to rationalize. Generally disubstitutions showed no superiority to monosubstituents, yet the 3,5-di-*tert*-butylphenyl substituted analogue **§10m** wasthe most active compound in these series (with an IC50estimate of 47 M). Compound **§10m** represents a 4-fold gain in activity compared to its 3,5-dimethylphenyl substituted analogue **§10l**. This led us to hypothesize that the hydrophobic bulk of compound **§10m** is well positioned in the binding site and acts specifically, particularly taking in consideration the many other lipophilic, nonetheless quite inactive compounds that were tested (Supporting Information, Figure S1). Interestingly, 3,5-bis(trifluoromethyl) substitution (as in **§10n**) abolished the activity completely (Supporting Information, Figure S1).

Monosubstitution at the *ortho* position, as in the original hit compound **1**, was thus studied by five compounds (Supporting Information, Figure S1), and no other substitutions than the nitro group in **1** were tolerated in this position. Of the *meta* substitutions, carboxyl and methoxy/alkyl groups were tolerated, as in compounds **§10b** and **X**. *Para* substitutions led to inhibition only in one case/two cases: nitrile substituted **§10h** and nitro? showing TmPPase inhibition (IC50estimate58 M). As part of the aryl exploration, molecular modelling suggested two positively charged interaction sites in the orthosteric cavity (amino acids xx, xx,xx and xx, xx,xx). We thus made an attempt to bridge these interactions sites, but unfortunately none of the compounds/dimers **§9o** and **§10o** had an IC50 below 100 M. conclusion?

**Table 1.** Selected examples of the compounds of the isoxazole series

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | | | | |  |
| **Compound** | **R1** | **R2** | **R3** | **R4** | **R5** | **Activity [µM]** |
| **1 (mPP-0007)** | NO2 | H | H | H | H | 63 |
| **§10a (mPP-0149)** | H | H | H | H | H | > 500 |
| *Polar or charged substitutions* | | | | | | |
| **§10b (mPP-0023)** | H | COOH | H | H | H | 80 |
| **§10h (mPP-00147)** | H | H | CN | H | H | 58 |
| **§10o (mPP-0153)** | H |  | H | H | H | 118 |
| *Aliphatic substitutions* | | | | | | |
| **§10l (mPP-0026)** | H | Me | H | Me | H | 219 |
| **§10m (mPP-0031)** | H | *t*-Bu | H | *t*-Bu | H | 47 |

*Alternative 5-heterocycles* **–** As a complement to the aryl group substitutions of the isoxazole core, we studied whether the exchange of the whole aryl group would benefit the compound binding. Changing the phenyl group to a 2-methylthiazol-4-yl group as in **§4** (Figure 2) was indeed successful, however the activity diminished completely for **§5** where the substituent was switched to a 1,5-dimethyl-1*H*-pyrazol-4-yl group.

The structural replacements of the 5-arylisoxazole-3-carboxylic acid core were tolerated but did not yield significantly better activities. We thus followed up investigating more subtle modifications of these compounds. Replacing the isoxazole ring (as in **§10a)** either with pyrazole (as in **§12**) or thiazole (as in **§13**) led to at least 5-fold or 25-fold increase in the activity, respectively (Figure 3). Something of the *para*-nitro compounds **if fitting when tested**.



**Figure 3.** Isoxazole ring modifications.

*Using 3,5-di-tert-butylphenyl moiety as a new fragment core* **–** Initially the presence of an isoxazole core has been central to the compound design. Considering that the 3,5-di-*tert*-butylphenyl moiety could be the driver towards active compounds, especially as the preliminary SAR data showed that the 3,5-dimethylphenyl analogue **§10l** had a reduced IC50estimate when compared to the 3,5-di-*tert*-butylphenyl analogue **§10m** (47 µM for **§10m** against 219 µM for **§10l**; Table 1), we decided to evaluate the role of the isoxazole-5-carboxylic acid itself. Hence we studied a set of 16 compound analogues carrying only a 3,5-di-*tert*-butylphenyl substituent without the isoxazole-5-carboxylic acid moiety (Figure 4; Supporting Information, Figure S1). As a result seven of these derivatives showed TmPPase inhibition of which compound **§16** was the best one with an IC50estimate of 22 µM.



**Figure 4.**Examples of the active 3,5-di-*tert*-butylphenyl analogues.

**Growing the fragment 1**

*Growing with carboxyl to ester substitutions* **–** The initial modifications of the 5-arylisoxazole-3-carboxylic acid core led to the discovery of seven compounds with IC50estimates’sof 47–220 µM, but no gain in affinity compared to the initial hit **1**. We thus turned to explore the other side of the isoxazole molecule by doing some acid modifications. As a first attempt we tested a set of ethyl ester synthesis intermediates. Generally, these esters lost activity compared to their carboxylic acid analogues. Compound **§9m** had a 2-fold drop in the activity in comparison to its acid analogue **§10m**, whereas compound **§9g** (*m*eta methoxy substituted ester, 179 µM) showed a 2-fold gain in activity to its corresponding acid **§10g**.

**Table 2.** Selected examples of the compounds of the isoxazole series

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | | | | |  |
| **Compound** | **R1** | **R2** | **R3** | **R4** | **R5** | **Activity [µM]** |
| *Ethyl ester substitutions* | | | | | | |
| **§9a (mPP-0148)** | H | H | H | H | Et | > 500 |
| **§9f (mPP-0162)** | H | OMe | H | H | Et | 219 |
| **§9g (mPP-0160)** | H | H | OMe | H | Et | 179 |
| **§9m (mPP-0154)** | H | *t*-Bu | H | *t*-Bu | Et | 76 |
| **§9o (mPP-0152)** | H |  | H | H | Et | > 500 |

*Growing with sulfonamide derivatives –* Two of the initial hits (compounds **2** and **3**, Figure 1) include a sulfon or sulfonamide group, indicating that these functional groups might be favourable for compound binding. For that reason we did a catalogue exploration approach of adding a sulfonamide moiety to the *meta* (R4*)* position of the 5-arylisoxazole-3-carboxylate core. All seven compounds in this screen were direct analogues of compound **§9g**. Two compounds showed an improvement of the IC50estimate’s when compared to the parent compound (Supporting Information, Figure S1), **§19** being the best one with 6-fold IC50estimate improvement (Figure 3). but less efficient



**Figure 5.** The benefit of the sulfonamide addition.

*Growing with 2-bromoester substitutions* – Among the analogues purchased during the initial round of the structural modifications, we had two separate subsets of six compounds, which indicated the benefit of bromophenyl esters of the isoxazole-3-carboxylate core, especially 2-bromophenyl carboxylates. Unfortunately, neither of these series met the final purity requirements (Supporting Information, Figure S1). This led us to investigate 2-bromophenyl carboxylate substitutions with hit compounds **§10m**, **§4** and **§13**. Also compound **§10l**, an almost identical analogue of the parent compound in one of the subsets (3,5-dimethylphenyl of **§10l** vs. 3,4-dimethylphenyl of the purchased parent compound) was studied. The 2-bromophenyl carboxylate modifications were indeed successful for compounds **§10l** and **§10m** (30-fold and 4.5-fold increase in the activity, respectively), but not for **§20** and **§21** (Figure 6).

What about the para nitro ones?



**Figure 6.** Bromoester analogues of the compounds **§10l**, **§10m**, **§4** and **§13**.

In order to gain a higher ligand efficiency for compounds **§11l** and **§11m** we tried to remove some superfluous ligands. Although we successfully managed to remove one of the alkyl chains (as in **§11j** and **§11k**) with only slight change in the inhibition, the removal of the bromine (as in **§22**) caused more severe effects on the inhibition. Additionally, we tried to make compound **§11l** more stable by changing the ester bond into the corresponding amide **§23**, but unfortunately also this modification only led loss of inhibition.

**Table 2.** Bromoester analogues of the compounds **§10l** and **§10m.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | | | | |  |
| **Compound** | **R1** | **R2** | **R3** | **R4** | **R5** | **Activity [µM]** |
| Carboxy substitutions | | | | | | |
| **§10l (mPP-0026)** | H | Me | H | Me | OH | 219 |
| **§11l (mPP-0224)** | H | Me | H | Me |  | 7 |
| **§10m (mPP-0031)** | H | *t*-Bu | H | *t*-Bu | OH | 47 |
| **§11m (mPP-0223)** | H | *t*-Bu | H | *t*-Bu |  | 10 |
| **§10j**  **6-mono (mPP)(NJ1-88)** | H | Me | H | H | OH | XX |
| **§11j**  **26-mono (mPP) (NJ2-2)** | H | Me | H | H |  | XX |
| **§10k**  **9-mono (mPP) (NJ1-89)** | H | *t*-Bu | H | H | OH | XX |
| **§11k**  **27-mono (mPP) (NJ2-3)** | H | *t*-Bu | H | H |  | XX |
| **§22**  **6-noBr (mPP) (NJ2-7)** | H | Me | H | Me |  | XX |
| **§23**  **6-amide (mPP) (OL-2rc)** | H | Me | H | Me |  | XX |

Accurate IC50 curves of **§4**, **§11l**, **§11m** (in TmPPase) Sup.Info. Fig. S3?

Kinetics of **§4 (**and **§11l** and **§11m** if allowed)

# **CONCLUSIONS**

In this manuscript we report non-phosphorus compounds inhibiting the TmPPase, the best of which (**§11l**) has IC50 of 7 µM. Even if the current compounds have too weak binding constants to obtain biological activities in animal, the hits are still small, almost fragments, which gives a lot of leverage to grow them. Altogether the SARs of the found inhibitors and the validation of on-site action using kinetics contributes to the demonstration that mPPases are a valid and druggable targets that could be pursued against parasitic diseases.

# **EXPERIMENTAL SECTION**

In the initial screen the purity of commercial analogues was tested only for the four best hits, whereas all compounds were purity tested in the subsequent screens. Compound purity was (predominantly) above 95 % as determined by LC-MS and characterized by HRMS. Full descriptions of biological, computational and synthetic methods can be found in the Supporting Information.

## Medicinal chemistry

*Computational methods.* Commercial analogues of the hit compounds hits were retrieved by screening of the ZINC12 database (clean drug-like subset; downloaded on 27.04.2016) (Irwin et al. 2012) using a KNIME (Berthold et al. 2007) workflow connected to the RDKit nodes (Landrum 2016) (available as Supporting Information SX). The SAR analyses were conducted using Schrödinger Canvas software (Schrödinger 2016).

*Synthetic methods.* An overview of all compounds in this study is presented in the Supporting Information, Figure S1. Compounds **§1–§5** and **§15–§19** were obtained from commercial suppliers and tested without further purification. Synthetic procedures and characterization data for **§4**, **§11l** and **§11m** are shown below (data for the remaining compounds are presented in the Supporting Information).

**5-(2-Methylthiazol-4-yl)isoxazole-3-carboxylic acid (§4).** Purchased from Sigma (*CDS018284*). MTI61 1H NMR (400.15 MHz, DMSO-*d6*) *δ* = 8.28 (s, 1H), 7.08 (s, 1H), 2.73 (s, 3H). 13C NMR (100.63 MHz, DMSO-*d6*) *δ* = 167.8, 166.3, 160.7, 157.7, 141.3, 120.7, 101.2, 18.8. HRMS-ESI (*m*/*z*): calculated for C8H7N2O3S [M+H]+ 211.0177; found, 211.0179.

*Esterification of compounds* ***§11i*–*m*** *and* ***§20*** *(*Supporting Information, General procedure D). To the carboxylic acid (1 equiv) was added a solution of oxalyl chloride (1.1 equiv) in DCM (1 mL) and one drop of anhydrous DMF. The mixture was stirred for 30 min, bubbled with argon and cooled to 0 °C in an ice bath. To the cooled acid chloride solution was slowly added 2-bromophenol (1.1 equiv) in DCM (0.75 mL), followed by triethylamine (1.1 equiv) in DCM (0.75 mL). The mixture was stirred for 20 min in the ice bath and another 30 min at room temperature (~1 hour since the phenol addition). The solvent was evaporated and the residue was purified by Biotage Isolera flash chromatography with TLC gradient estimation and a 10 g SNAP KP-Sil cartridge with hexane:ethyl acetate (1:0 → 0:1) as the eluent.

**2-Bromophenyl 5-(3,5-dimethylphenyl)isoxazole-3-carboxylate (§11l).** General procedure D was used to give **§11l** as a white solid (9.7 mg, 57%). NJ1-58 1H NMR (400.15 MHz, CDCl3) *δ* = 7.68 (dd, *J* = 8.2, 1.4 Hz, 1H), 7.48 (s, 2H), 7.44–7.38 (m, 1H), 7.32 (dd, *J* = 8.2, 1.4 Hz, 1H), 7.24–7.18 (m, 1H), 7.14 (s, 1H), 7.05 (s, 1H), 2.40 (s, 6H). 13C NMR (100.63 MHz, CDCl3) *δ* = 172.9, 157.7, 156.0, 147.7, 139.1, 133.8, 132.9, 128.8, 128.2, 126.4, 123.9, 123.7, 116.1, 100.3, 21.4. HRMS-ESI (*m*/*z*): calculated for C18H15BrNO3 [M+H]+ 372.0235; found, 372.0233.

**2-Bromophenyl 5-(3,5-di-*tert*-butylphenyl)isoxazole-3-carboxylate (§11m).** General procedure D was used to give **§11m** as a colorless oil that later on turned into a white solid (14 mg, 82%). NJ1-59 1H NMR (400.15 MHz, CDCl3) *δ* =7.70–7.66 (m, 3H), 7.59–7.57 (m, 1H), 7.44–7.38 (m, 1H), 7.32 (dd, *J* = 8.1, 1.6 Hz, 1H), 7.24–7.19 (m,1H), 7.08 (s, 1H), 1.39 (s, 18H). 13C NMR (100.63 MHz, CDCl3) *δ* = 173.5, 157.8, 156.0, 152.1, 147.8, 133.8, 128.8, 128.2, 126.0, 125.5, 123.8, 120.5, 116.2, 100.2, 35.2, 31.5. HRMS-ESI (*m*/*z*): calculated for C24H27BrNO3 [M+H]+ 456.1174; found, 456.1171.

## Biological testing

*Enzymatic assay.* As previously described.(REF)

*Aggregation tests.* To evaluate potential aggregation of samples in the TmPPase based assay, aggregate formation was studied at five concentrations (50 µM, 25 µM, 10 µM, 5 µM, and 1 μM) for compounds **§4**, **§11l** and **§11m** (Supporting Information Figure S2).

# **ASSOCIATED CONTENT**

## Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: XX.XXXX/acs.jmedchem.XXXXXXX.

Figure S1: All compounds tested for this story (structures in 2D + IC50estimate OR pre-IC50)

Figure S2: Aggregation data for compounds **§4**, **§11l** and **§11m**.

Figure S3. Accurate IC50 curves of **§4**, **§11l**, **§11m** (in TmPPase)

Full experimental procedures and characterization data (NMR, MS)

KNIME workflow

SMILES strings?

# **AUTHOR INFORMATION**

## Corresponding Author

\*E-mail: henri.xhaard@helsinki.fi. Tel. +358294159190.

## ORCID

## Author contributions

N.G.J., A.T., and K.V contributed equally. All authors contributed in the writing and editing of the manuscript.

Computational part (design, modelling and analysis) of the study was performed by A.T., E.G., L.D. and H.X.

Pharmacological part (design, biological testing and crystallisation) of the study was performed by K.V. and A.G.

Synthetical part (design, synthesis and characterization) of the study was performed by N.G.J., D.A.P, O.L. A.W., M.T., Y.Z., A.K., C.F., G.B.G. and J.Y.-K.

*OR more detailed*

All authors contributed in the writing and editing of the manuscript.

Conceptualization, K.V., A.G., H.X., J.Y.-K., G.B.G.; Methodology, N.G.J., A.T., K.V., A.G., A.K., E.G.; Investigation, N.G.J., A.T., K.V.; Resources, A.G., H.X., J.Y.-K.; Writing – Original draft, H.X., N.G.J., A.T., K.V.; Writing-Review and Editing, N.G.J., A.T., K.V., H.X., A.K., G.B.G., J.Y.-K., A.G.; Visualization, N.G.J., A.T., K.V.; Supervision and Project Administration, H.X., A.G., G.B.G., J.Y.-K.; Funding Acquisition, A.G., K.V., H.X., G.B.G..

## Funding sources

This work is supported by the grants from

Jane and Aatos Erkko Foundation and the X BBSRC (BB/\*\*) X to A.G.

The Academy of Finland (No. 308105) to K.V.

The Academy of Finland (No. 310297) to H.X.

University of Helsinki 3-year grant to G.B.G.

## Conflict of Interest

All authors declare no conflict of interest in this paper.

# **ACKNOWLEDGMENTS**

The authors wish to acknowledge: CSC – IT Center for Science and the Drug Discovery and Chemical Biology (DDCB) network for computational resources; Nina Sipari (from the Viikki Metabolomics Unit, University of Helsinki) for helping with the mass spectrometry; XXX

the Magnus Ehrnrooth Foundation, Otto A. Malm Foundation, the Gust. Komppa Fund of the Alfred Kordelin Foundation and the doctoral programme in chemistry and molecular sciences (CHEMS).

# **ABBREVIATIONS**

membrane-bound pyrophosphatase (mPPase)

imidodiphosphate (IDP)

~~inorganic pyrophosphate (PP~~~~i~~~~)~~

molecular weight (MW)

~~transmembrane helix (TMH)~~

# **REFERENCES**

XXX

# **TOC** (needs to be updated)

