

## ARTICLE

Received 23 Jul 2015 | Accepted 18 Oct 2016 | Published 6 Dec 2016

DOI: 10.1038/ncomms13596

OPEN

# Membrane pyrophosphatases from *Thermotoga maritima* and *Vigna radiata* suggest a conserved coupling mechanism

Kun-Mou Li<sup>1,\*</sup>, Craig Wilkinson<sup>2,\*</sup>, Juho Kelloso<sup>3,†</sup>, Jia-Yin Tsai<sup>1</sup>, Tommi Kajander<sup>4</sup>, Lars J.C. Jeuken<sup>2</sup>, Yuh-Ju Sun<sup>1</sup> & Adrian Goldman<sup>2,3</sup>

Membrane-bound pyrophosphatases (M-PPases), which couple proton/sodium ion transport to pyrophosphate synthesis/hydrolysis, are important in abiotic stress resistance and in the infectivity of protozoan parasites. Here, three M-PPase structures in different catalytic states show that closure of the substrate-binding pocket by helices 5–6 affects helix 13 in the dimer interface and causes helix 12 to move down. This springs a ‘molecular mousetrap’, repositioning a conserved aspartate and activating the nucleophilic water. Corkscrew motion at helices 6 and 16 rearranges the key ionic gate residues and leads to ion pumping. The pumped ion is above the ion gate in one of the ion-bound structures, but below it in the other. Electrometric measurements show a single-turnover event with a non-hydrolysable inhibitor, supporting our model that ion pumping precedes hydrolysis. We propose a complete catalytic cycle for both proton and sodium-pumping M-PPases, and one that also explains the basis for ion specificity.

<sup>1</sup> Department of Life Science and Institute of Bioinformatics and Structural Biology, College of Life Science, National Tsing Hua University, Hsinchu 30013, Taiwan. <sup>2</sup> Astbury Centre for Structural Molecular Biology, School of Biomedical Sciences, University of Leeds, Leeds LS2 9JT, England. <sup>3</sup> Division of Biochemistry, Department of Biosciences, University of Helsinki, Helsinki, FIN-00014, Finland. <sup>4</sup> Institute of Biotechnology, University of Helsinki, PO Box 65, Helsinki, FIN-00014, Finland. \* These authors contributed equally to this work. † Present address: Institute of Biotechnology, PO Box 65, University of Helsinki, FIN-00014, Finland. Correspondence and requests for materials should be addressed to Y.-J.S. (email: yjsun@life.nthu.edu.tw) or to A.G. (email: a.goldman@leeds.ac.uk).

**M**embrane-bound pyrophosphatases (M-PPases) are enzymes that couple the synthesis or hydrolysis of pyrophosphate (PP<sub>i</sub>) to the vectorial transport of protons and/or sodium ions<sup>1</sup>. These enzymes are found in plants and in various unicellular organisms and are important for survival under different abiotic stress conditions, such as drought and cold stress in plants<sup>2</sup>. The stress resistance conveyed by M-PPases is presumably due to the capability of the enzymes to use the metabolic by-product, PP<sub>i</sub>, for creating transmembrane cation potential differences that can drive various cellular processes, including ATP synthesis and secondary transport<sup>3</sup>.

M-PPases utilize Mg<sub>2</sub>PP<sub>i</sub> as their substrate and require Mg<sup>2+</sup> for activity, binding two activating Mg<sup>2+</sup> ( $K_{\text{ds}}$  20–42 μM and 0.25–0.46 mM)<sup>2</sup> with an additional inhibitory Mg<sup>2+</sup> binding site ( $K_{\text{d}} \approx 100$  mM)<sup>4</sup> present in the active site. In addition, M-PPases can be divided into seven functional classes on the basis of their monovalent cation binding (K<sup>+</sup>-dependent and K<sup>+</sup>-independent enzymes) and pumping specificity (H<sup>+</sup>-pumps (H<sup>+</sup>-PPases), Na<sup>+</sup>-pumps (Na<sup>+</sup>-PPases) and H<sup>+</sup> and Na<sup>+</sup> co-pumps (Na<sup>+</sup>/H<sup>+</sup>-PPases)) and on the placement of a ‘semi-conserved glutamate’ essential for pumping activity<sup>3–5</sup>. K<sup>+</sup>-dependent enzymes with a conserved alanine ‘A<sup>12.46</sup>’ (*Thermotoga maritima* Na<sup>+</sup>-PPase (TmPPase) A495 and *Vigna radiata* H<sup>+</sup>-PPase (VrPPase) A537) are activated by monovalent cations, of which K<sup>+</sup> shows the highest level of activation<sup>2</sup>. K<sup>+</sup>-independent enzymes, on the other hand, contain a conserved lysine ‘K<sup>12.46</sup>’ where the N $\zeta$  amino group probably substitutes for the bound cation of the K<sup>+</sup>-dependent enzymes<sup>6</sup>. (In what follows, we refer to residues as in refs 3–5, 7, represented as: X<sup>Y,Z</sup> where X is the amino acid, Y is the helix number and Z is the residue position based on a conserved residue found within the helices of all known M-PPases: Supplementary Table 1.)

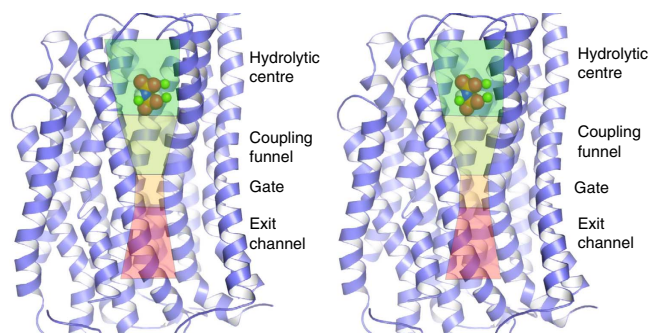
The semi-conserved glutamate is in the same position (E<sup>6.53</sup>) in the K<sup>+</sup>-independent H<sup>+</sup>-PPases as well as the (K<sup>+</sup>-dependent) Na<sup>+</sup>-PPases and Na<sup>+</sup>/H<sup>+</sup>-PPases. This position differs in the K<sup>+</sup>-dependent H<sup>+</sup>-PPases (E<sup>5.43</sup>, E<sup>6.53</sup> or E<sup>6.57</sup>)<sup>1,8</sup>.

Both Na<sup>+</sup>-PPases and Na<sup>+</sup>/H<sup>+</sup>-PPases are only active in the presence of Na<sup>+</sup> or Li<sup>+</sup> (refs 1,9), which are the two pumped alkali cations (Na<sup>+</sup> being the physiologically relevant one). In the presence of Na<sup>+</sup>, K<sup>+</sup> activates the enzyme and increases the enzyme’s affinity for Na<sup>+</sup> by about 200 fold ( $K_{\text{m}}^{\text{Na}^+}$  9–80 mM in the absence of K<sup>+</sup> and 0.036–0.45 mM in its presence<sup>1,4,9,10</sup>).

So far, the structures of two M-PPases have been solved in three different states: TmPPase in resting and product-bound states (TmPPase:Ca:Mg and TmPPase:P<sub>i2</sub>, respectively)<sup>11</sup> and VrPPase in a substrate-analogue imidodiphosphate (IDP)-bound state (VrPPase:IDP)<sup>12</sup>. These structures have shown that M-PPases are homodimers, with each protomer having 16 transmembrane helices (TMHs) and a vertically arranged, continuous active site structure with four distinctive parts: ‘hydrolytic centre’, ‘coupling funnel’, ‘ion gate’ and ‘exit channel’<sup>3</sup> (Fig. 1).

The hydrolytic centre is situated on the cytoplasmic side of the protein with conserved aspartates and lysines lining the binding pocket and coordinating the binding of Mg<sup>2+</sup>, IDP and a water molecule that is poised for nucleophilic attack<sup>12</sup>. The closure of the active site cavity by the loop between TMHs 5 and 6 (loop 5–6) upon substrate binding is necessary for directional pumping<sup>12</sup>. On the periplasmic/luminal side of the hydrolytic centre are the coupling funnel, ion gate and exit channel (Fig. 1).

Mutagenesis of key residues in TMH12 uncouples hydrolysis and pumping<sup>13,14</sup>, while mutagenesis of the gate residues E<sup>6.53/6.57</sup> (the aforementioned semi-conserved glutamate), S<sup>6.54</sup> and D/N<sup>16.46</sup> are important for sodium binding in Na<sup>+</sup>-PPases and for proton-pumping activity in H<sup>+</sup>-PPases<sup>8,12,14–16</sup>. This suggests that the sodium-binding site of Na<sup>+</sup>-PPases is found



**Figure 1 | Stereo view of the vertical channel composition of M-PPases.**

Structural overview of an M-PPase:IDP complex, showing the various sub-structures in the active site: the hydrolytic centre, coupling funnel, ion gate and exit channel. Helices 5, 9, 10 and 15 have been removed for clarity.

near these residues<sup>16</sup> with a lack of Na<sup>+</sup> affecting hydrolysis, but not Mg<sub>2</sub>PP<sub>i</sub> binding<sup>4</sup>.

Clearly, the starting point of ion pumping is the nucleophilic water at the hydrolytic centre<sup>12</sup>, which loses a proton upon electrophilic attack on the PP<sub>i</sub>. Three different mechanisms have been proposed for how pumping is coupled to hydrolysis: substrate binding<sup>11</sup>, hydrolysis<sup>12</sup> or a ‘Mitchell-direct’ mechanism, where the proton released during hydrolysis is directly pumped<sup>17</sup>. The first two mechanisms differ in their timing: does hydrolysis precede pumping or not? The third mechanism, as it says that the proton released during hydrolysis is pumped, is not easy to reconcile with Na<sup>+</sup>-pumping enzymes. The other major question is the role of the coupling funnel: how does it connect proton abstraction from the water nucleophile to ion pumping—and indeed pumping of a different ion?

In this article we present three structures: the Na<sup>+</sup>-pumping TmPPase with IDP and Na<sup>+</sup> bound (TmPPase:IDP) at 3.5 Å resolution, the H<sup>+</sup>-pumping VrPPase with one phosphate bound (VrPPase:P<sub>i</sub>) at 3.5 Å resolution, and the newly refined structure of TmPPase with the phosphate analogue, WO<sub>4</sub> bound, at 4 Å resolution (TmPPase:WO<sub>4</sub>). (This last was previously used to solve the phase problem in the original TmPPase structure, but the structure was not refined<sup>11</sup>.) We suggest where and how TmPPase binds Na<sup>+</sup>, and how H<sup>+</sup>-pumping evolved from Na<sup>+</sup>-pumping. We develop a plausible unified mechanism for how substrate binding leads to nucleophile activation and ion pumping and provide electrometric evidence that suggests pumping precedes hydrolysis upon substrate binding. Finally, we suggest that the atomic mechanism of gating in M-PPases is similar to that of the ATPases.

## Results

**Structural overview.** The overall structures of TmPPase:IDP, TmPPase:WO<sub>4</sub> and VrPPase:P<sub>i</sub> are similar to our previous structures<sup>11,12</sup> with root mean squared deviation (r.m.s.d.) values of 0.7 Å for TmPPase:IDP–TmPPase:P<sub>i2</sub>, 1.6 Å for TmPPase:IDP–TmPPase:Ca:Mg and 1.0 Å for VrPPase:P<sub>i</sub>–VrPPase:IDP (Supplementary Fig. 1). The former was solved by molecular replacement against protein data bank (PDB): 4AV3 (ref. 11) at 3.5 Å and the latter by molecular replacement against PDB: 4A01 (ref. 12) at 3.5 Å (Table 1, see ‘Methods’ section and Table 2).

In TmPPase:IDP, an Mg<sub>5</sub>IDP-complex was fit into the positive F<sub>o</sub>–F<sub>c</sub> density in the hydrolytic centre (Supplementary Fig. 2a) and then refined. The coordination of Mg<sup>2+</sup> and IDP is almost identical to that seen in VrPPase:IDP<sup>12</sup> (Fig. 2a and Supplementary Fig. 3), even though the two enzymes pump different ions. In VrPPase:P<sub>i</sub>, the extra F<sub>o</sub>–F<sub>c</sub> electron density in the PP<sub>i</sub>-binding pocket is a phosphate ion, just 0.6 Å from the

**Table 1 | Structures and states of M-PPases.**

Protein	Resolution (Å)	No. chains/ASU	Active site contents	Closure of active site	Position of TMH12*
TmPPase:Ca:Mg <sup>†</sup>	2.6	2	CaMg	Open: 3,400 Å <sup>3</sup>	0
TmPPase:IDP <sup>§</sup>	3.5	2	Mg <sub>5</sub> IDP (Na <sup>+</sup> at ion gate)	Closed: 1,600 Å <sup>3</sup>	−2.7 Å
TmPPase:P <sub>12</sub> <sup>†</sup>	4.0	2	Mg <sub>4</sub> P <sub>12</sub>	ND <sup>  </sup>	−2.2 Å
TmPPase:WO <sub>4</sub> <sup>§</sup>	4.0	2	MgWO <sub>4</sub>	ND <sup>  </sup>	−1.0 Å
VrPPase:IDP <sup>‡</sup>	2.6	2	KMg <sub>5</sub> IDP	Closed: 1,600 Å <sup>3</sup>	−2.2 Å
VrPPase:P <sub>i</sub> <sup>§</sup>	3.5	4	Mg <sub>2</sub> P <sub>i</sub>	Open: 2,800 Å <sup>3</sup>	−0.7 Å

ASU, asymmetric unit; ND, not determined.

\*Position relative to that in TmPPase:Ca:Mg based upon K<sup>12.50</sup> C<sub>α</sub>, where a negative number represents a downwards motion.†Structures from Kellosalo *et al.*<sup>11</sup>‡Structure from Lin *et al.*<sup>12</sup>

§Structures presented in this paper.

||ND because the loops are missing, meaning that the active site volume can not be meaningfully compared. ASU, asymmetric unit.

**Table 2 | Data collection and refinement statistics.**

Crystal	TmPPase:IDP complex	TmPPase:WO <sub>4</sub>	VrPPase:P <sub>i</sub> complex
<i>Data collection</i>			
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub>	C2
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	106.4, 106.8, 161.9	83.7, 108.9, 105.8	225.7, 81.6, 264.8
β (°)	90	108.8	92.9
Source	DLS—i04	ESRF ID23-1	NSRRC—BL15A1
Wavelength	0.9795	1.214	1.000
Resolution (Å)*	30–3.49 (3.70–3.49)	25–4.0 (4.1–4.0)	30.0–3.50 (3.62–3.50)
<i>R</i> <sub>merge</sub> (%) <sup>*,†</sup>	9.3 (140.8)	5.5 (42.6)	16.8 (117.5)
<i>I</i> / <i>σ</i> <sup>*</sup>	15.83 (1.35)	14.3 (3.65)	7.5 (1.11)
Completeness (%) <sup>*</sup>	99.5 (98.6)	98.8 (92.8)	92.2 (85.1)
Redundancy <sup>*</sup>	8.8 (8.6)	3.8 (3.4)	3.5 (3.5)
<i>Refinement</i>			
Resolution (Å)	3.5	4.0	3.5
No. of reflections	23,904	29,597	54,540
<i>R</i> <sub>work</sub> (%)/ <i>R</i> <sub>free</sub> (%)	23.9/26.8	23.2/27.9	22.6/30.4
No. of atoms	10,443	9,808	21,413
Protein	10,389 (1,450 residues)	9,794 (1,383 residues)	21,385 (2,914 residues)
IDP, P <sub>i</sub> or WO <sub>4</sub>	36 (2 IDP)	10 (2 WO <sub>4</sub> )	20 (4 P <sub>i</sub> )
Mg <sup>2+</sup>	10	2	8
Na <sup>+</sup>	2	–	–
No. of chains/ASU	2	2	4
B-factors (Å <sup>2</sup> )			
All atoms	139	134.9	119.6
Protein	138.9	133.0	119.5
IDP, P <sub>i</sub> or WO <sub>4</sub>	167.8 (IDP)	198.5	159.4 (P <sub>i</sub> )
Mg	134.9	151.6	107.9
Na	112.7	–	–
r.m.s.d.			
Bond length (Å)	0.005	0.004	0.01
Bond angle (°)	1.06	1.19	1.48

\*Values in the parenthesis are for the highest resolution shell.

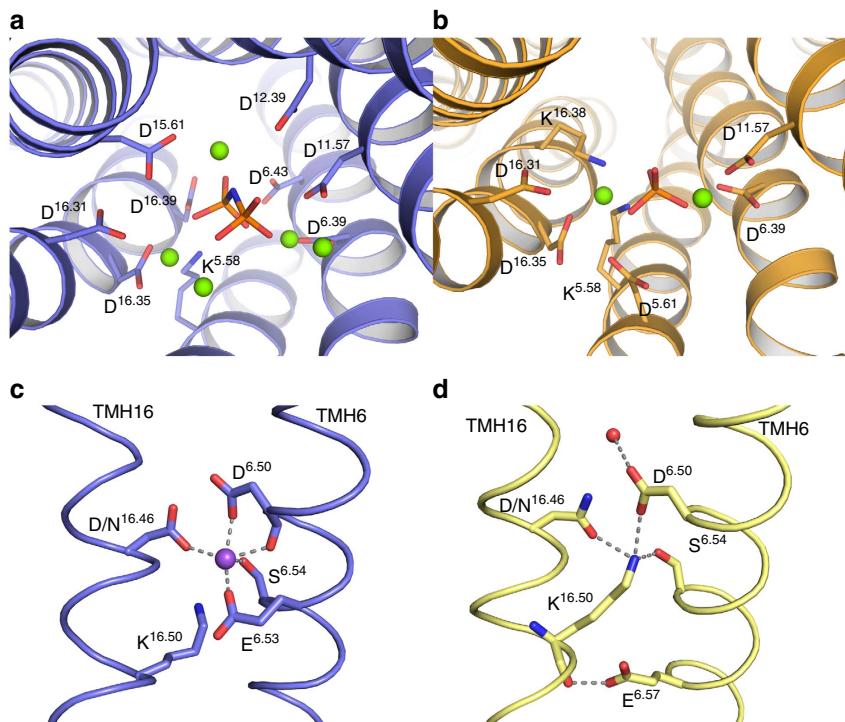
†*R*<sub>merge</sub> = Σ||*I* − <*I*>|/Σ*I*, where *I* = observed intensity, and <*I*> = average intensity from multiple observation of symmetry related reflections.

leaving group phosphate in VrPPase:IDP (Supplementary Fig. 2d). This phosphate ion is coordinated by K<sup>5.58</sup>, K<sup>16.38</sup> and two Mg<sup>2+</sup> ions, which are situated in a similar position to Mg2 and Mg3 in VrPPase:IDP and to the Ca<sup>2+</sup> and Mg<sup>2+</sup> in resting-state TmPPase (Fig. 2b).

**Bound sodium ion.** Positive F<sub>o</sub>–F<sub>c</sub> density was seen in TmPPase:IDP between residues D<sup>6.50</sup>, E<sup>6.53</sup>, S<sup>6.54</sup> and D<sup>16.46</sup> (Fig. 2c and Supplementary Fig. 2b). Close to this, K<sup>16.50</sup>, which forms an ion-triplet in the resting-state TmPPase structure with D<sup>6.50</sup> and E<sup>6.53</sup>, adopts a different orientation and the K<sup>16.50</sup> N<sub>ζ</sub> is disordered. Mutagenesis has shown that the gate residues E<sup>6.53</sup>, S<sup>6.54</sup> and D/N<sup>16.46</sup> are important in sodium binding in Na<sup>+</sup>-PPases<sup>16</sup>.

A sodium ion was therefore refined into this density. The resulting bond distances were between 2.39–2.49 Å (mean: 2.43 Å) and the site was penta-coordinated (by the carboxylate groups of D<sup>6.50</sup>, E<sup>6.53</sup> and D<sup>16.46</sup>, the O<sup>γ</sup> of S<sup>6.54</sup> and the main-chain carbonyl of D<sup>6.50</sup>)<sup>18</sup> (Fig. 2c). A score of 0.90 was obtained from a valence test<sup>19</sup> at this position, with scores over 0.7 indicative of sodium ions (scores were between 0.1 and 0.4 for water molecules elsewhere in the structure). All of these results indicate that the density is most likely a bound Na<sup>+</sup> and not a water molecule.

**Comparison with previous structures.** The new structures of TmPPase and VrPPase were compared with previously solved



**Figure 2 | Hydrolytic centre and ion gate of TmPPase and VrPPase.** Key residues are labelled and hydrogen bonds shown as dashed lines.

(a) The hydrolytic centre of TmPPase:IDP (blue in all figures) with bound  $Mg^{2+}$  (green), IDP (orange) and coordinating residues. (b) Hydrolytic centre of VrPPase:P<sub>1</sub> (gold) showing  $Mg^{2+}$  (green), P<sub>1</sub> (orange) and coordinating residues. (c) Ion gate region of TmPPase:IDP showing bound  $Na^{+}$  (purple) and its coordination (grey). TMH 6 and 16 are labelled. (d) Ion gate of VrPPase:IDP (yellow) showing the salt-bridge network (yellow) and coordinated waters (red). TMH 6 and 16 are labelled.

structures (PDB ID: 4AV3, 4AV6, 4A01). Notable was the volume of the hydrolytic centre, which approximately halves upon binding substrate: from 3,400 Å<sup>3</sup> in resting-state TmPPase to 1,600 Å<sup>3</sup> in the IDP-bound states of VrPPase and TmPPase, but expanding to 2,800 Å<sup>3</sup> in VrPPase:P<sub>1</sub> (Table 1). The cytoplasmic ends of the inner-ring TMHs 5, 6, 11, 12, 15 and 16 and the outer-ring TMHs 3, 4, 13 and 14 are constricted in TmPPase:IDP<sup>12</sup> (Fig. 3) and loops 5–6 and 13–14 are ordered and cover the active site cavity. The closure of the active site cavity involves the formation of an intricate salt-bridge network between the residues of TMHs 5, 12, 15 and loop 5–6 in both VrPPase:IDP<sup>12</sup> and TmPPase:IDP. Loop 13–14, which folds on top of loop 5–6, probably plays an important role in stabilizing the placement of loop 5–6, and hence the closure of the active site, as only TMHs 13 and 14 move horizontally during the conversion from the IDP- to P<sub>12</sub>-bound states (Fig. 3b).

The largest movement in the entire structure upon substrate binding is at TMH13 (5.5–8 Å at the C $\alpha$ ), apparently driven by motion at TMH5 and loop 5–6 (Supplementary Fig. 4a). The interactions of E<sup>5.71</sup>–R<sup>13.62</sup>–R/I<sup>10.33</sup>, are abolished in the absence of substrate (Supplementary Fig. 4b,c); R<sup>13.62</sup> links motions of the inner ring to the outer ring and into the neighbouring protomer. This motion is not propagated through the membrane: the membrane and periplasmic regions of TMHs 10 and 13 do not show any sign of conformational change. Nonetheless, it is this region that contributes most to the M-PPase dimer interface<sup>12</sup>.

**Analysis of helical and hydrogen bond geometry.** The region surrounding D<sup>6.43</sup> in TmPPase:IDP has approximately canonical  $\alpha$ -helical hydrogen-bonding (calculated using HBplot (VirtuaDrug, Hungary)<sup>20,21</sup>), whereas the structures are significantly bent in the resting state (Supplementary Table 2). Residues around D<sup>6.43</sup> bend on average 15° from linear as

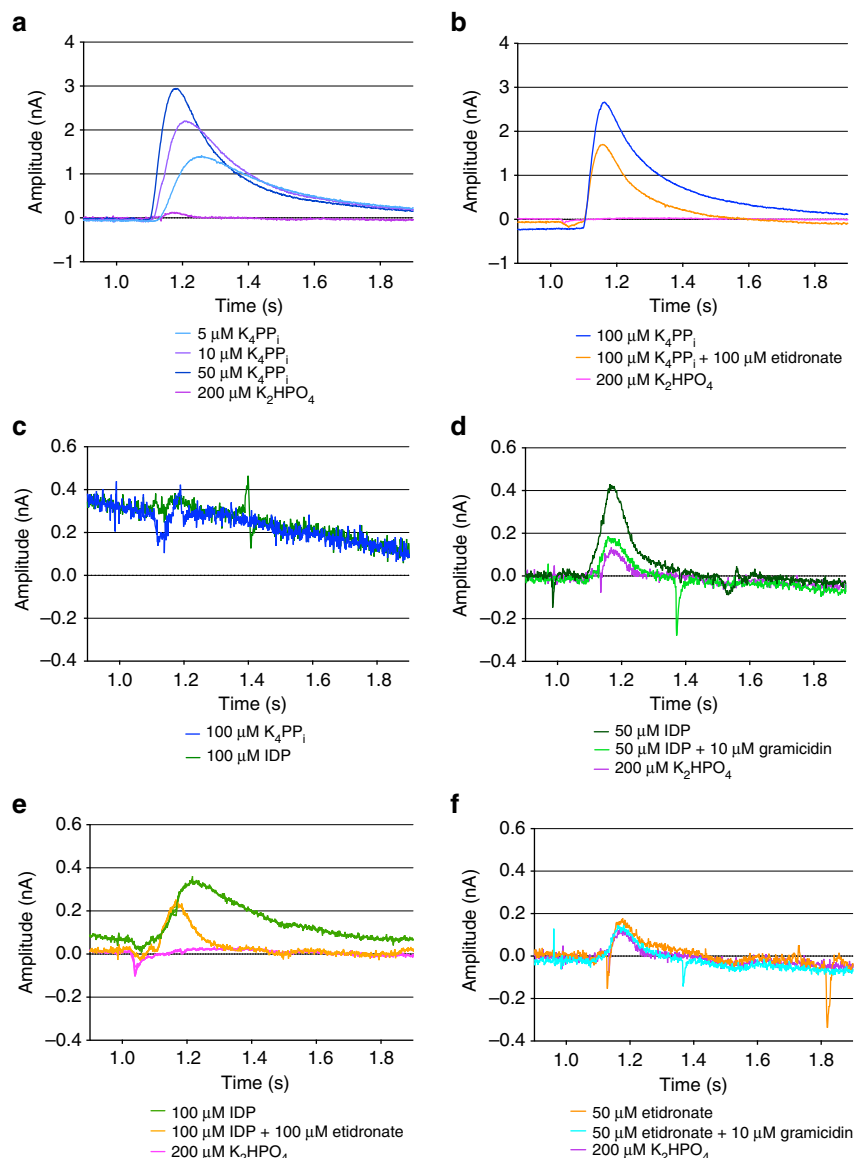
measured using Bendix<sup>22</sup> (Supplementary Fig. 5). When TMH12 moves down by 2 Å, the loss of the K<sup>12.50</sup>–D<sup>6.43</sup> ion pair leads to straightening of the 6.43 region by about 7° and a spring-loaded 3–4 Å movement of the entire D<sup>6.43</sup> carboxylate group so that it now hydrogen bonds the water nucleophile, as seen in the VrPPase:IDP structure (Supplementary Fig. 6d). Because of this conformational change, some of the hydrogen bonds are lost and others converted to a  $3_{10}$  pattern before D<sup>6.43</sup>, with a  $\pi$ -helix hydrogen-bonding pattern after it (Supplementary Table 2). In addition, the backbone carbonyl on D<sup>6.43</sup> is flipped away from the helix axis and can no longer make a hydrogen bond to the backbone amide on L/M<sup>6.47</sup> (Supplementary Table 2).

In addition to the changes seen in TMH 6 (above), TMH 16 also changes conformation when IDP binds (Supplementary Figs 5 and 7), with ~1.5 Å ‘corkscrew change’ in TMH16 at N/D<sup>16.46</sup>–K<sup>16.50</sup> due to the formation of a few residues of  $3_{10}$  helix at S<sup>16.44</sup>–L<sup>16.45</sup> (Supplementary Table 2). These helical movements shift the positions of D<sup>6.50</sup> and S<sup>6.54</sup> in VrPPase (Fig. 2d and Supplementary Fig. 7b) and of D<sup>6.50</sup>, E<sup>6.53</sup>, S<sup>6.54</sup> and K<sup>16.50</sup> in TmPPase (Fig. 2c and Supplementary Fig. 7a).

**Electrometric studies of proton pumping.** We studied the pumping of H<sup>+</sup> by VrPPase on the Nanion SURFE<sup>2</sup>R N1. This technique utilizes a membrane-coated, gold sensor on individual chips. The protein is first reconstituted into liposomes (see ‘Methods’ section), the liposome is then adhered to this membrane layer and the current across the liposomes is measured over time following a rapid switch between two ionically balanced buffers, with the activating buffer containing the relevant substrate or inhibitor.

A positive signal up to 3 nA is repeatedly observed with K<sub>4</sub>PP<sub>1</sub> (Fig. 4a). This signal is diminished when the inhibitor etidronate (Ethane-1-hydroxy-1,1-diphosphonate; PO<sub>3</sub>–C(CH<sub>3</sub>)(OH)–PO<sub>3</sub>)





**Figure 4 | Electrometric sensor traces for  $H^+$ -pumping.** Traces obtained for VrPPase from the Nanion SURFE<sup>2</sup>R N1 machine. **(a)** Effect of increasing concentrations of  $K_4PP_i$  and  $K_2HPO_4$  on signal amplitude of VrPPase proteoliposomes. **(b)** Effect of etidronate addition to  $K_4PP_i$ -induced signals of VrPPase proteoliposomes. **(c)** Effect of  $K_4PP_i$  and IDP on liposomes without pyrophosphatase. **(d)** Effect of IDP on amplitude in the presence and absence of gramicidin compared with  $K_2HPO_4$ . **(e)** Effect of etidronate addition to IDP-induced signals of VrPPase proteoliposomes. **(f)** Effect of etidronate on amplitude in the presence and absence of gramicidin compared with  $K_2HPO_4$ . Activating buffer containing substrate or inhibitor was added at 1 s and removed at 2 s. Two anomalous peaks (covering  $<0.01$ s) were removed from **e** for the sake of clarity.

the ion gate. In the IDP-bound states of both proteins, two notable changes occur: the salt bridge is broken and  $D/N^{16.46}$  swings into the vicinity of the other gate residues (Supplementary Fig. 7). The big difference between TmPPase and VrPPase is in the position of  $K^{16.50}$ ; indeed this is the most significant difference in the entire active site and so must be related to the central difference between the proteins—their pumping specificity. In TmPPase:IDP (Fig. 2c), it has swung away and is disordered ( $N\zeta$  B-factor =  $112.5 \text{ \AA}^2$ ), while in VrPPase:IDP,  $K^{16.50}$  forms a well-ordered ion pair with  $D^{6.50}$  (Fig. 2d). This difference is intimately related both to coupling and to ion selectivity, as we describe below.

In TmPPase, sodium binding is facilitated by the inward motion of  $D^{16.46}$  into the sodium-binding site and by the movement of TMHs 6 and 16, which is almost twice as big as that in VrPPase. These helical movements bring  $D^{6.50}$ ,  $E^{6.53}$  and  $S^{6.54}$

to the correct position to bind  $Na^+$  and move  $K^{16.50}$  away from the binding site (Supplementary Fig. 7a). The changes are, of course, driven by the downward motion of TMH 12 and the loss of the  $D^{6.43}$ – $K^{12.50}$  ion pair, and so by substrate binding and water activation (see above; Supplementary Fig. 7a). This also provides a direct link between substrate and ion binding: the  $Na^+$  site forms only upon substrate binding, and so the binding affinity of substrate increases in the presence of sodium and vice versa<sup>4</sup>.

By contrast, VrPPase has a completely different gate structure due to the shift in the position of the semi-conserved glutamate. The VrPPase:IDP has, instead of  $Na^+$ , a number of water molecules lining the channel, including one coordinated by the gate residues (Fig. 2d). These water molecules are absent in the VrPPase:P<sub>i</sub>, probably due to the low resolution of the structure (3.5 Å). Nonetheless,  $K^{16.50}$  is ordered in both states of VrPPase,

unlike in TmPPase. We propose that, in the IDP-bound state, the semi-conserved glutamate  $E^{6.57}$  at the ion gate is protonated, disrupting the  $E^{6.57}$ - $K^{16.50}$  ion pair, and causing  $K^{16.50}$  to move 'upward' (Supplementary Fig. 7). This leads to charge separation at the gate, with a full-proton localized on  $E^{6.57}$ , the ionic gate closed and  $K^{16.50}$  neutralized by  $D^{6.50}$ ,  $S^{6.54}$  and  $N^{16.46}$  (Fig. 2d) in a way that is analogous to the neutralization of the charge of  $Na^+$  in TmPPase (Fig. 2c).

The description above has another fundamental implication: in TmPPase, the positive charge to be pumped (the sodium ion) is still above  $K^{16.50}$  (on the cytoplasmic side; Fig. 2c), whereas in VrPPase, it is below (the proton localized on  $E^{6.57}$ ) and thus ready to be transferred to the exit channel. It is nonetheless likely that both states (charge above  $K^{16.50}$  and charge below  $K^{16.50}$ ) exist in both proteins. Our structures thus appear to visualize two different sub-states in the catalytic cycle.

Finally, functional experiments with *Streptomyces coelicolor* M-PPase have shown that M256<sup>6.47</sup>T and I242<sup>6.33</sup>T mutant proteins have 90–100% of wild-type hydrolysis activity, but only a 25–30% wild-type coupling ratio between hydrolysis and proton-pumping<sup>15</sup>. Even though the residues are not completely conserved (VrPPase:  $M^{6.47}$  and  $I^{6.33}$ ; TmPPase:  $L^{6.47}$  and  $I^{6.33}$ ) the fact that these are hydrophobic to polar mutations create 'loose' coupling mutations is consistent with our model that TMH6 is also essential in ensuring coupling. These mutations could perturb the movement of this helix so that hydrolysis can proceed even without ion transport to the exit channel.

The structures described in Table 1 encompass all the states for a minimal catalytic scheme for M-PPases (Fig. 5). For a pump to achieve directional flow, the sequence of events—that is, the rates of the unitary steps—must be organized such that pumping is faster than the reverse reaction. For M-PPases, it implies that, for example,  $k_3$  must be faster than  $k_{-2}$  for unitary committing steps on the enzyme and, typically, that  $k_3 < k_2 < k_1$  (Fig. 5) as has been seen, for instance, in cytochrome c oxidase<sup>24,25</sup>. This ensures that  $H^+/Na^+$  pumping is faster than opening to the cytoplasm and release of  $PP_i$  or  $P_i$ . Our static structures can only hint at the possibilities, but we believe that the key step is the breaking of the  $D^{6.43}$ - $K^{12.50}$  ion pair. This appears to release stored strain energy in the structure, as it leads to a new conformation where the  $D^{6.43}$  carboxylate group moves towards the now-bound negatively charged phosphate (Fig. 5b), the 6.43 region does not form main-chain hydrogen bonds, and the pumped ion becomes localized at the ionic gate. Consistent with this, Lee *et al.*<sup>26</sup> have shown that mutation of  $K^{12.50}$  in VrPPase abolishes both hydrolysis and pumping. The role we propose explains these results as  $K^{12.50}$  does not bind substrate, nor is it part of the ionic gate.

We therefore speculate that the pyrophosphate binding and localization of the ion is the fast step in the reaction cycle and that the reverse of this step is slow ( $k_1 > k_{-1}$ ). In addition, we propose that TmPPase:IDP and VrPPase:IDP structures represent two different states in the pumping process: the former with the pumped ion cytoplasmic (above  $K^{16.50}$ ; Fig. 5b), and the latter with it luminal (below  $K^{16.50}$ ; Fig. 5c). At this point, the protein is open neither to the cytoplasm nor to the periplasmic/luminal side; and the gate is closed. The next, presumably slower step, would involve the exit channel opening, probably through further 'downward' movement of TMH12 (ref. 11), diffusion of the ion away and closure of the gate (Fig. 5d). This would pave the way for hydrolysis as the final, slow-committing step (Fig. 5e). Only after these steps have happened can the active site reopen on the cytoplasmic side, and the two phosphates leave in an ordered manner (Fig. 5f,g; see above). This proposed mechanism thus ensures coupling between hydrolysis and pumping. The opening of the active site presumably also decreases the affinity of the

active site for product ( $MgP_i$ ), allowing the second phosphate to leave the active site, and thus returning the enzyme to the resting state (M-PPase: $Mg_2$ ; Fig. 5a). A similar kinetic scheme run in reverse would explain how M-PPases are able to synthesize pyrophosphate. This raises another important question: which comes first, hydrolysis or pumping?

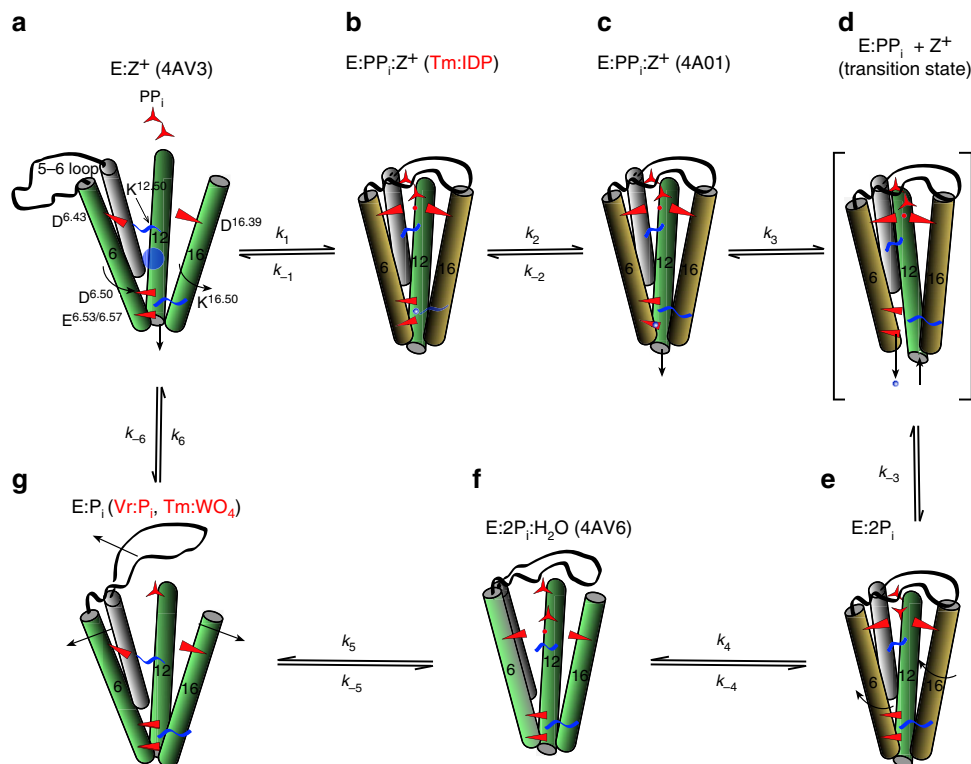
The best possible explanation of the 0.4 nA signal from IDP (Fig. 4d) from the Nanion SURFE<sup>2</sup>R results is that this non-hydrolyzable analogue causes a single-turnover event as substrate hydrolysis is essential for continued ion pumping; without it, product release cannot occur. This would prove that substrate binding drives pumping, the model we prefer, as pumping can occur without hydrolysis.

This is supported by three lines of argument. First, gramicidin forms a specific channel for monovalent cations. Any reduction in the proton-driven signal from IDP upon adding gramicidin must therefore be due to a collapse of the monovalent cation gradient. As there is a clear and reproducible reduction upon adding gramicidin (Fig. 4d), IDP must cause proton pumping. Second, we have a clear measure of the component of the current due to binding of  $Mg^{2+}$  to the protein. The residual signal (0.1–0.2 nA), seen when etidronate or  $KH_2PO_4$  (Fig. 4) bind, can not be collapsed by gramicidin (Fig. 4f), and so is due to this other component. The Nanion SURFE<sup>2</sup>R data thus also suggest that IDP and etidronate bind differently; IDP binds like substrate and is able to drive the conformational changes required for ion-pumping, while etidronate is not able to bind in the same manner. This is so even though etidronate is a somewhat better inhibitor of VrPPase<sup>27</sup> (IDP  $K_i = 12 \mu M$ , Etidronate  $K_i = 6.5 \mu M$ ). Modelling etidronate into the active site indicates the following: it prevents proper closure of the active site due to steric hindrance with  $D^{15.61}$  and one of the coordinating magnesium ions (Supplementary Fig. 8). Third, Fig. 4a,d provides further evidence that there is coupling between substrate hydrolysis and ion pumping in the catalytic cycle of M-PPases, since the signal induced by  $PP_i$  (1–3 nA) is up to ten times greater than that induced by IDP (0.4 nA). Conversely, we suggest that etidronate prevents the formation of vital interactions in the hydrolytic centre or closure of the active site, so neither pumping nor hydrolysis can occur. This again emphasizes the importance of substrate specificity to enzyme activation.

It cannot be fully excluded that the single-turnover signals are due to alternative transport events. We again note that the gramicidin control (Fig. 4d) strongly suggests that IDP, contrary to etidronate, induces a transmembrane charge transfer. If this charge transfer is not connected to the pumped proton as we propose here, alternative transport mechanisms could be envisioned in which the proton is transported after  $PP_i$ . This would be similar to the proposal in Lin *et al.*<sup>12</sup>, but not similar to the established mechanism in rotary ATPases<sup>28,29</sup>, nor to the general behaviour of non-control point enzymes, where the free energy difference between substrate and product is expressed as the free energy difference between the binding of substrate versus the binding of product<sup>30</sup>.

Recent mutagenesis studies<sup>14</sup> showed that the  $I^{12.54}A$  and  $L^{12.64}A$  (Supplementary Fig. 6c) mutants of VrPPase are decoupled: they do not pump but retain half of the hydrolysis activity of wild-type. This corroborates our hypothesis, as these mutants would change the molecular smoothness of the lower-end of TMH12, thus affecting its downward motion<sup>11</sup>.

There is neither sequence, structure nor overall mechanistic similarity between M-PPases and the pumping ATPases<sup>11,12</sup>. We believe, however, that there is nonetheless distant mechanistic similarity at the ion gate. The protonated-gate state of VrPPase:IDP (Fig. 6a) and the ion-bound state of TmPPase:IDP (Fig. 6b) are analogous to the ATP-analogue binding E1-state of



**Figure 5 | Complete kinetic scheme of M-PPase catalytic cycle.** (a–g) All the proposed conformational states of the M-PPase catalytic cycle are shown in order, including the transition state for pumping (d) and a ‘relaxed product’ state (f) analogous to yeast pyrophosphatase<sup>46</sup>. The structures used in the manuscript are shown, with the new structures marked in red; for the rest, the PDB code is shown. For clarity, only key helices involved in formation of the hydrolytic centre and ion channel are displayed. Aspartate and lysine residues are shown in red and blue respectively and labelled on state (a). Changes to helix position and rotation are denoted with arrows and by changes in the shading of the helices. The pumped ion ( $\text{Na}^+/\text{H}^+$ ) is represented as a blue sphere in states (b–d) and as a large blue sphere in state (a) to denote the unknown position of the pumped ion in the channel of the resting state. The nucleophilic water is represented as a red sphere in states (b–d).

P-type  $\text{H}^+$ -ATPase<sup>31</sup>, where the protonated acceptor/donor residue D684 forms hydrogen bonds with N106 inside an occluded binding pocket (Fig. 6c). Similarly, in VrPPase:IDP, protonated  $E^{6.57}$  makes hydrogen bonds with the carbonyl group of  $K^{16.50}$  and  $\text{O}\gamma$  of  $S^{5.43}$  but not to  $K^{16.50}$   $\text{N}\zeta$  (Figs 2d and 6a)<sup>31</sup>. In addition, the movement of  $K^{16.50}$  in and out of the vicinity of the proton donor/acceptor in M-PPases is like that of R210 in the F-ATPase  $\alpha$ -subunit<sup>29</sup> and R655 in P-type  $\text{H}^+$ -ATPase<sup>31</sup>, where its movement is associated with ion binding and release (Fig. 6). This alternating motion may be a unifying mechanistic principle for primary ion pumps.

Phylogenetic analyses show that  $\text{Na}^+$ -PPases were the first M-PPases to evolve and that the evolution of  $\text{H}^+$ -PPases has occurred at least three times, leading to different  $\text{H}^+$ -PPases with the semi-conserved glutamate in different positions ( $E^{5.43}$ ,  $E^{6.53}$  or  $E^{6.57}$ )<sup>8</sup>. The only major difference between the gate structure of sodium-pumping TmPPase and proton-pumping VrPPase is the shift in the position of the semi-conserved glutamate from  $E^{6.53}$  to  $E^{6.57}$ , respectively. While changing the position of the semi-conserved glutamate through mutagenesis has so far led only to protein inactivation<sup>8</sup>, our structures indicate why the  $E^{6.53}$  to  $E^{6.57}$  change led to the evolution of plant-type  $\text{H}^+$ -PPases from  $\text{Na}^+$ -PPases<sup>11</sup>.

We propose that moving the semi-conserved glutamate down one helical turn to  $E^{6.57}$  changes the pumping specificity first, by destroying the  $\text{Na}^+$ -binding site in the IDP-bound state and, second, by creating a proton acceptor at the end of the cytoplasmic Grotthuss chain<sup>32</sup>. In the new position, it interacts with  $K^{16.50}$  in the same way as the D684–R655 pair in

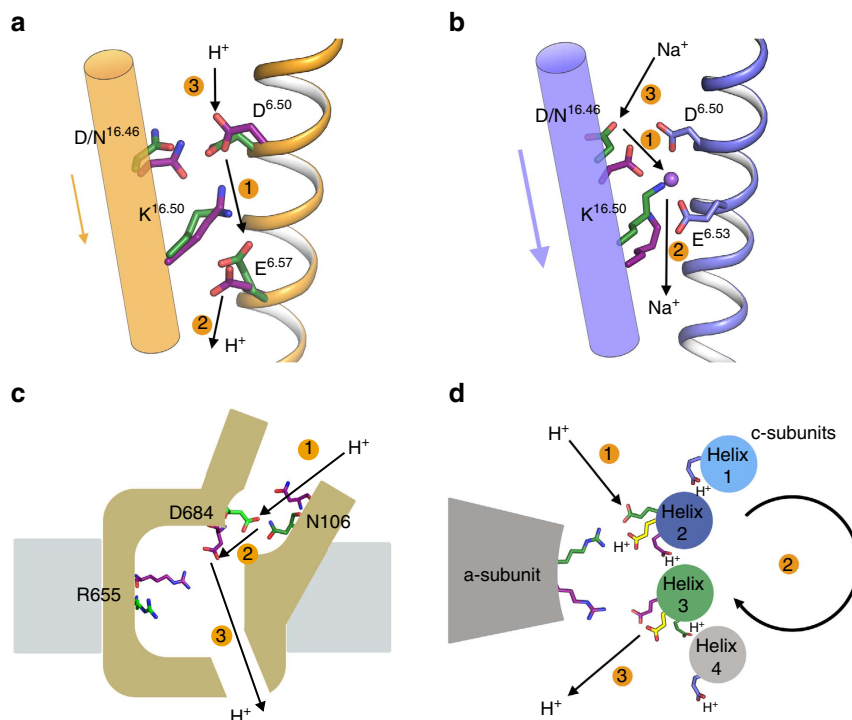
$\text{H}^+$ -ATPase (see above). As a result,  $K^{16.50}$   $\text{N}\zeta$  adopts the same conformation in the IDP-bound  $\text{H}^+$ -PPases (Supplementary Fig. 7b) as it does in the resting state of  $\text{Na}^+$ -PPases (Supplementary Fig. 7a), leading to the change in pumping specificity. The gate residues, which in  $\text{Na}^+$ -PPases would bind  $\text{Na}^+$ , neutralize  $K^{16.50}$ : slight changes in the order of conformational steps lead to a profound change in the ion pumped.

The model also explains the evolution of  $\text{Na}^+/\text{H}^+$ -pumping M-PPases (see above). If this model is correct, the switch between pumping  $\text{Na}^+$  and  $\text{H}^+$  only requires repositioning TMH6 so that the  $\text{Na}^+$  site does not form, but that  $E^{6.53}$  can act as the end of the Grotthuss chain. This could be driven by the allosteric motion driven by TMH13, as described above. We are exploring these ideas.

## Methods

**Expression and purification of VrPPase.** *Saccharomyces cerevisiae* strain BJ2168 was transformed with the galactose-inducible vector pYES2 containing C-terminally His<sub>6</sub>-tagged VrPPase by the lithium acetate transformation method<sup>33</sup>. All yeast cells were collected by centrifuging at  $\times 2,300$  g at 4 °C for 10 min after 3-day induction. Next, the yeast cells were resuspended in 5 mM Tris/Ethylene glycol tetraacetic acid (EGTA) (pH 7.6), 10% (w/v) glycerol, 0.6% (w/v) Tris (base), 0.6% (w/v) ascorbate, 1.5% (w/v) Polyvinylpyrrolidone ( $M_r$  40,000; PVP40000), 1 mM Phenylmethylsulfonyl fluoride (PMSF) and  $1 \mu\text{g ml}^{-1}$  pepstatin A, and disrupted by ultra-sonication. The homogenate was then centrifuged differentially at 2,300g for 10 min and 126,000g for 35 min at 4 °C. After ultracentrifugation, the membrane vesicles were resuspended in extraction buffer (25 mM MOPS/KOH (pH 7.0), 400 mM KCl, 4 mM  $\text{MgCl}_2$ , 20% (w/v) glycerol and 1 mM PMSF) and solubilized by using *n*-dodecyl- $\beta$ -D-maltopyranoside at a protein-to-detergent ratio of 1:2.5. The suspension was gently stirred at 4 °C for 1 h followed by





**Figure 6 | Comparison of M-PPases with ATPases, demonstrating the formal similarity and the coupling of Lys/Arg motion to ion pumping.**

(a) Proton pumping and peristaltic motion of K<sup>16.50</sup> in VrPPase and (b) sodium pumping and peristaltic motion of K<sup>16.50</sup> in TmPPase. (c) Structural overview and hypothesized proton transport mechanism of P-type ATPases<sup>31</sup>. (d) Proton pumping and rotary motion in F-ATPase<sup>29</sup>: (based on figure 6 from reference 29) in the hydrolysis direction. Colour scheme: green, initial residue states; yellow, intermediate states; magenta, final states; blue before and after (F-ATPase). D<sup>6.50</sup> and E<sup>6.53</sup> in (b) and the helices in (a,b) are coloured as in Figs 2 and 3 to emphasize the fact that the motions there are relatively small. Numbers represent order of ion progression through the respective ion channels.

centrifugation at 126,000g for 35 min. After ultracentrifugation, the solubilized VrPPase was loaded onto a Ni<sup>2+</sup>-NTA column and eluted with a buffer containing 25 mM MOPS/KOH (pH 7.1), 400 mM KCl, 4 mM MgCl<sub>2</sub>, 20% (w/v) glycerol, 1 mM PMSF, 0.15% (w/v) *n*-decyl-β-D-maltopyranoside and 250 mM imidazole. The purified VrPPase was exchanged to the crystallization buffer (25 mM 4-morpholineethanesulfonic acid (MES) (pH 6.5), 400 mM KCl, 4 mM MgCl<sub>2</sub>, 20% (w/v) glycerol, 0.15% (w/v) *n*-decyl-β-D-maltopyranoside) and concentrated to 10 mg ml<sup>-1</sup> for crystallization.

**Expression and purification of TmPPase.** The expression and purification of TmPPase followed a modified version of the 'hot-solve'-purification method<sup>34</sup>. BJ1991-strain (Mat(Δ) leu2 trp1 ura3-251 prb1-1122 pep4-3 gal2) *Saccharomyces cerevisiae* cells were transformed with plasmid pRS1024, which contained a LEU2-gene and N-terminally His<sub>6</sub>-tagged TmPPase under a constitutive PMA1-promoter<sup>34</sup>. The cells were first grown in selective synthetic complete drop-out (SCD)-media to the end of the exponential growth phase and then 250 ml of this culture was used to inoculate 740 ml of 1.5 × concentrated YEP-media with 2% glucose. The cells were grown in this for 8 h at 30 °C after which they were collected, lysed via sonication with 0.2 mm glass beads and their membranes were extracted using ultracentrifugation (100,000g for 45 min) with pellets resuspended in resuspension buffer: 50 mM MES-NaOH pH 6.5, 20% (v/v) glycerol, 50 mM KCl, 5.2 mM MgCl<sub>2</sub>, 0.33 mM Na<sub>2</sub>PP<sub>i</sub>, 1.33 mM Dithiothreitol (DTT), 2 μg ml<sup>-1</sup> (w/v) pepstatin-A (Sigma) and 0.334 mM PMSF (Sigma)<sup>34</sup>. For solubilization, the extracted membranes were diluted into 30 ml fractions containing 7.2 mg ml<sup>-1</sup> of membranes in resuspension buffer and heated at 75 °C for 15 min. After this each membrane aliquot was mixed with 10 ml of pre-heated solubilization buffer containing 50 mM MES-NaOH pH 6.5, 20% (v/v) glycerol and 5.334% (w/v) *n*-dodecyl-β-D-maltopyranoside (Anatrace) and incubated at 75 °C for 1.5 h. The denatured proteins were removed by immediately centrifuging at 3,300g for 5 min at 22 °C and then at 3,300g for 15 min at 4 °C. KCl was added to a final concentration of 0.3 M together with 2 ml of Ni-NTA matrix (Qiagen) to each 40 ml fraction of solubilized protein. The protein was bound to the Ni-NTA matrix at 42 °C for 1.5 h after which the matrix was loaded into a column. The column was washed with 1.5 × column volume (CV) of washing buffer (50 mM MES-NaOH pH 6.5, 20% (v/v) glycerol, 50 mM KCl, 20 mM imidazole pH 6.5, 5 mM MgCl<sub>2</sub>, 1 mM DTT, 2 μg ml<sup>-1</sup> (w/v) pepstatin-A, 0.2 mM PMSF and 0.5% octyl glucose neopentyl glycol (OGNPG, Anatrace) and the protein was eluted with 2 × CV of elution buffer (50 mM MES-NaOH pH 6.5, 3.5% (v/v) glycerol, 50 mM KCl, 400 mM imidazole pH 6.5, 5 mM MgCl<sub>2</sub>, 1 mM DTT, 2 μg ml<sup>-1</sup> (w/v) pepstatin-A, 0.2 mM PMSF and 0.5% OGNPG).

**Crystallization and structure determination of VrPPase.** Crystallization trials were carried out using the hanging-drop vapour-diffusion method<sup>35</sup>. The VrPPase:P<sub>i</sub> complex crystals were obtained in 2 days over a reservoir solution containing 100 mM sodium/potassium phosphate (pH 6.2), 200 mM NaCl and 37% (w/v) PEG600 at 20 °C using 0.5 μl protein drops mixed 1:1 with the reservoir solution. The X-ray diffraction data were collected at the BL44XU beamline at Spring8, Japan and BL15A1 beamline at the NSRRC, Hsinchu, Taiwan. Data were processed by HKL2000 (ref. 36). The VrPPase:P<sub>i</sub> complex crystals belonged to the monoclinic space group C2 with the cell parameters *a* = 225.7 Å, *b* = 81.6 Å, *c* = 264.8 Å and β = 92.9° (Table 2). The Matthews' coefficient was calculated to be 3.8 Å<sup>3</sup> Da<sup>-1</sup>, corresponding to a solvent content of 67.7% with four subunits per asymmetric unit<sup>37</sup>. The crystal structure was determined by molecular replacement with MOLREP<sup>38</sup> using the VrPPase:IDP complex (PDB: 4A01)<sup>12</sup> as a search model. There were four molecules as two dimers (AB and CD dimers) per asymmetric unit. The entire crystal structure was completed manually with COOT<sup>39</sup> and refined by REFMAC5 (ref. 40). The stereochemistry of protein residues and the secondary structural features were evaluated by PROCHECK<sup>41</sup>. The VrPPase:P<sub>i</sub> complex with two dimers per asymmetric unit contains 21,413 protein atoms, four P<sub>i</sub> and eight Mg<sup>2+</sup> ions. The structure was refined to an *R*-factor of 22.6% and an *R*<sub>free</sub> of 30.4% at 3.5 Å. The data collection and refinement statistics are listed in Table 2 and representative electron density for the P<sub>i</sub> in Supplementary Fig. 2d.

**Crystallization and structure determination of TmPPase.** For use in crystallization, the purified TmPPase was concentrated with Amicon Ultra 50,000 MWCO concentrator (Millipore), buffer exchanged to crystallization buffer (50 mM MES-NaOH pH 6.5, 3.5% (v/v) glycerol, 50 mM KCl, 5 mM MgCl<sub>2</sub>, 4 mM Na<sub>2</sub>IDP (Sigma-Aldrich: 10631), 2 mM DTT and 0.5% OGNPG) with Micro Bio-Spin 6 column (Bio-Rad) and diluted to a concentration of 10 mg ml<sup>-1</sup>. The crystallization of TmPPase:IDP was carried out using the vapour-diffusion method<sup>34</sup> in a 1 μl + 1 μl drop on a 24-well plate at room temperature using a well solution containing 32% PEG 400, 0.1 M Tris pH 8.5, 0.1 M MgCl<sub>2</sub>, 0.1 M NaCl and 2 mM DTT. The crystal was frozen directly from the mother liquor and X-ray diffraction data (Table 2) were collected at beamline i04 of Diamond Light Source, UK. Data for the TmPPase:WO<sub>4</sub> structure were collected as part of a previous study<sup>11</sup>. TmPPase crystals produced as above were soaked in 10 mM Na<sub>2</sub>WO<sub>4</sub> overnight, where they had only been used for phasing the resting TmPPase:CaMg structure in the previous study<sup>11</sup>. Both data sets were processed with X-ray Detector Software (XDS)<sup>42</sup> and the structures were solved by molecular

replacement with Phaser<sup>43</sup> using the resting-state TmPPase (4AV3)<sup>11</sup> as the search model. Model refinement was carried out in Phenix.refine (v. 1.9–1692)<sup>44</sup>, with manual alterations being made to the model in COOT<sup>39</sup>. Refinement was carried out at 3.5 Å and 4 Å for TmPPase:IDP and TmPPase:WO<sub>4</sub>, respectively, using reference restraints from the high resolution (2.6 Å) TmPPase structure (4AV3), as well as secondary structure restraints. These were used alongside Torsion-angle NCS restraints and tighter stereochemical weighting to maintain realistic geometry and to prevent over-fitting the data at these resolutions. In both models, the density for the helices was clear and unambiguous. There was additional density for the loop regions in the TmPPase:IDP structure, including the TMH 5–6 loop region. These loop regions were omitted from the TmPPase:WO<sub>4</sub> structure due to the high degree of disorder in these regions. In both instances, many side chains were also omitted from the model due to a lack of density. During refinement, we saw strong positive (F<sub>o</sub>–F<sub>c</sub>) density at 4.7σ around the ion gate in protomer A. This was attributed to a sodium ion based upon the level of density, the optimal interaction distances with the surrounding residues and previous studies showing that mutating these residues affected sodium binding<sup>8,9,16</sup>. Following refinement, the R-factors for TmPPase:IDP were: 24.8%/27.5% with 95.4% of residues in the preferred region of the Ramachandran plot. For the TmPPase:WO<sub>4</sub> structure, the R-factors were: 23.2%/27.9% with 93.4% of residues within the preferred region of the Ramachandran plot. Data collection and refinement statistics are listed in Table 2 with representative electron density in Supplementary Fig. 2.

**Structural analysis of TmPPase and VrPPase.** We analysed all of the structures, TmPPase:IDP, TmPPase:WO<sub>4</sub> and VrPPase:P<sub>i</sub> (new) with the previously solved structures, TmPPase:Ca:Mg, TmPPase:2P<sub>i</sub> and VrPPase:IDP (previously solved) with two pieces of software. Hydrogen bonding patterns of residues were determined and compared using the hydrogen bond plot tool HBplot (VirtuaDrug, Hungary)<sup>20,21</sup>. This information was used to determine the classification of helix ( $\alpha$ ,  $\pi$  or  $3_{10}$ ) and was used to compare the structures. Second, the helical geometry was studied using the Bendix plugin for the Visual Molecular Dynamics software<sup>22</sup>. Using this software, the bend angle of each helix was calculated and compared between the structures of M-PPases.

**Reconstitution of VrPPase into liposomes.** The reconstitution protocol and subsequent activity assay used were adapted from the study by Liu *et al.*<sup>45</sup>. Purified VrPPase protein (see above) was diluted to 50 μg ml<sup>-1</sup> in VrPPase reconstitution buffer (0.25 M sorbitol, 1 mM MgSO<sub>4</sub>, 0.1 mM EGTA, 2 mM DTT, 10 mM Tricine-Na (pH 7.5)). 1 ml of diluted protein was mixed with 15 μl of 120 mg ml<sup>-1</sup> phosphatidylcholine from soybean (dissolved in VrPPase reconstitution buffer). SM-2 Bio-beads were added at 0.25 mg μl<sup>-1</sup> and the samples were incubated for 1 h at 4 °C with gentle agitation. The samples were run through Biorad micro Bio-spin 6 columns, equilibrated with VrPPase reconstitution buffer to remove residual NaCl and glycerol. Resulting proteoliposomes were frozen in liquid nitrogen and stored at –80 °C until needed for measurements. To ensure the protein was still active, the P<sub>i</sub> hydrolytic activity of the reconstituted protein was measured using the ammonium molybdate method for phosphate release<sup>45</sup>. The resulting activities are shown in Supplementary Fig. 9.

**Nanion SURFE<sup>2</sup>R measurements of VrPPase.** Measurements were carried out using the SURFE<sup>2</sup>R N1 machine from Nanion Technologies (Munich, Germany). Sensors were prepared according to the ‘SURFE<sup>2</sup>R sensor preparation’ protocol (Nanion’s standard procedures) using their sensor prep A2 and B solutions. 0.4 μg of sonicated VrPPase proteoliposomes (UP50H compact lab homogenizer (Hielscher), 1 mm diameter tip: 20 % amplitude, 10 pulses, pulse time = 0.5 s) were combined with VrPPase SURFE<sup>2</sup>R buffer (100 mM HEPES-KOH (pH 8), 50 mM KCl, 5 mM MgCl<sub>2</sub>, 0.16 mM EGTA) to a total volume of 50 μl and applied directly to each sensor. Sensors were spun at 2,500g for 30 min and stored at 4 °C for at least 1 h before measurements. Sensors were rinsed between measurements with 2 × VrPPase SURFE<sup>2</sup>R buffer C (100 mM HEPES-KOH, pH 8). Measurements were taken over 3 s (procedure BAB) with each buffer (VrPPase SURFE<sup>2</sup>R buffers A (activating buffer) and B (non-activating buffer)) added at 1 s intervals, the resulting charge differences were measured. Various compounds were added to each buffer based upon the experiments being performed (K<sub>4</sub>PP<sub>i</sub>, Na<sub>4</sub>PP<sub>i</sub>, K<sub>2</sub>HPO<sub>4</sub>, Na<sub>2</sub>HPO<sub>4</sub>, IDP, etidronate). Experiments were attempted using reconstituted TmPPase, but these were unsuccessful as TmPPase is not active enough at room temperature.

**Data availability.** The atomic coordinates and structure factors of the TmPPase:IDP, TmPPase:WO<sub>4</sub> and VrPPase:P<sub>i</sub> complexes have been deposited in the Protein Data Bank, www.rcsb.org (PDB IDs 5LZQ, 5LZR and 5GPJ). All other data that support the findings in this study are available from the corresponding authors upon reasonable request.

## References

- Luoto, H. H., Baykov, A. A., Lahti, R. & Malinen, A. M. Membrane-integral pyrophosphatase subfamily capable of translocating both Na<sup>+</sup> and H<sup>+</sup>. *Proc. Natl Acad. Sci. USA* **110**, 1255–1260 (2013).
- Maeshima, M. Vacuolar H<sup>+</sup>-pyrophosphatase. *Biochim. Biophys. Acta* **1465**, 37–51 (2000).
- Tsai, J.-Y., Kellosalo, J., Sun, Y.-J. & Goldman, A. Proton/sodium pumping pyrophosphatases: the last of the primary ion pumps. *Curr. Opin. Struct. Biol.* **27**, 38–47 (2014).
- Malinen, A. M., Baykov, A. A. & Lahti, R. Mutual effects of cationic ligands and substrate on activity of the Na<sup>+</sup>-transporting pyrophosphatase of *Methanosarcina mazei*. *Biochemistry* **47**, 13447–13454 (2008).
- Luoto, H. H., Nordbo, E., Malinen, A. M., Baykov, A. A. & Lahti, R. Evolutionarily divergent, Na<sup>+</sup>-regulated H<sup>+</sup>-transporting membrane-bound pyrophosphatases. *Biochem. J.* **467**, 281–291 (2015).
- Belogurov, G. A. & Lahti, R. A lysine substitute for K<sup>+</sup>. A460K mutation eliminates K<sup>+</sup> dependence in H<sup>+</sup>-pyrophosphatase of *Carboxydotherrmus hydrogenoformans*. *J. Biol. Chem.* **277**, 49651–49654 (2002).
- Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).
- Luoto, H. H., Belogurov, G. A., Baykov, A. A., Lahti, R. & Malinen, A. M. Na<sup>+</sup>-translocating membrane pyrophosphatases are widespread in the microbial world and evolutionarily precede H<sup>+</sup>-translocating pyrophosphatases. *J. Biol. Chem.* **286**, 21633–21642 (2011).
- Belogurov, G. A. *et al.* Membrane-bound pyrophosphatase of *Thermotoga maritima* requires sodium for activity. *Biochemistry* **44**, 2088–2096 (2005).
- Malinen, A. M., Belogurov, G. A., Baykov, A. A. & Lahti, R. Na<sup>+</sup>-Pyrophosphatase: a novel primary sodium pump. *Biochemistry* **46**, 8872–8878 (2007).
- Kellosalo, J., Kajander, T., Kogan, K., Pokharel, K. & Goldman, A. The structure and catalytic cycle of a sodium-pumping pyrophosphatase. *Science* **337**, 473–476 (2012).
- Lin, S. M. *et al.* Crystal structure of a membrane-embedded H<sup>+</sup>-translocating pyrophosphatase. *Nature* **484**, 399–403 (2012).
- Hirono, M. & Maeshima, M. Functional enhancement by single-residue substitution of *Streptomyces coelicolor* A3(2) H<sup>+</sup>-translocating pyrophosphatase. *J. Biochem.* **146**, 617–621 (2009).
- Asaoka, M., Segami, S. & Maeshima, M. Identification of the critical residues for the function of vacuolar H<sup>+</sup>-pyrophosphatase by mutational analysis based on the 3D structure. *J. Biochem.* **156**, 333–344 (2014).
- Hirono, M., Nakanishi, Y. & Maeshima, M. Essential amino acid residues in the central transmembrane domains and loops for energy coupling of *Streptomyces coelicolor* A3(2) H<sup>+</sup>-pyrophosphatase. *Biochim. Biophys. Acta* **1767**, 930–939 (2007).
- Luoto, H. H., Nordbo, E., Baykov, A. A., Lahti, R. & Malinen, A. M. Membrane Na<sup>+</sup>-pyrophosphatases can transport protons at low sodium concentrations. *J. Biol. Chem.* **288**, 35489–35499 (2013).
- Baykov, A. A., Malinen, A. M., Luoto, H. H. & Lahti, R. Pyrophosphate-fueled Na<sup>+</sup> and H<sup>+</sup> transport in prokaryotes. *Microbiol. Mol. Biol. Rev.* **77**, 267–276 (2013).
- Harding, M. M. Metal-ligand geometry relevant to proteins and in proteins: sodium and potassium. *Acta Crystallogr. D* **58**, 872–874 (2002).
- Nayal, M. & Di Cera, E. Valence screening of water in protein crystals reveals potential Na<sup>+</sup> binding sites. *J. Mol. Biol.* **256**, 228–234 (1996).
- Bikadi, Z., Demko, L. & Hazai, E. Functional and structural characterization of a protein based on analysis of its hydrogen bonding network by hydrogen bonding plot. *Arch. Biochem. Biophys.* **461**, 225–234 (2007).
- McDonald, I. K. & Thornton, J. M. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793 (1994).
- Dahl, A. C., Chavent, M. & Sansom, M. S. Bendix: intuitive helix geometry analysis and abstraction. *Bioinformatics* **28**, 2193–2194 (2012).
- Steitz, T. A., Shoham, M. & Bennett, Jr W. S. Structural dynamics of yeast hexokinase during catalysis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **293**, 43–52 (1981).
- Siegbahn, P. E. & Blomberg, M. R. Energy diagrams and mechanism for proton pumping in cytochrome *c* oxidase. *Biochim. Biophys. Acta* **1767**, 1143–1156 (2007).
- Wikström, M. & Verkhovsky, M. I. Mechanism and energetics of proton translocation by the respiratory heme-copper oxidases. *Biochim. Biophys. Acta* **1767**, 1200–1214 (2007).
- Lee, C. H. *et al.* Identification of essential lysines involved in substrate binding of vacuolar H<sup>+</sup>-pyrophosphatase. *J. Biol. Chem.* **286**, 11970–11976 (2011).
- Baykov, A. A. *et al.* Differential sensitivity of membrane-associated pyrophosphatases to inhibition by diphosphonates and fluoride delineates two classes of enzyme. *FEBS Lett.* **327**, 199–202 (1993).
- Stock, D., Gibbons, C., Arechaga, I., Leslie, A. G. & Walker, J. E. The rotary mechanism of ATP synthase. *Curr. Opin. Struct. Biol.* **10**, 672–679 (2000).
- Pogoryelov, D. *et al.* Microscopic rotary mechanism of ion translocation in the F<sub>0</sub> complex of ATP synthases. *Nat. Chem. Biol.* **6**, 891–899 (2010).

30. Albery, W. J. & Knowles, J. R. Evolution of enzyme function and the development of catalytic efficiency. *Biochemistry* **15**, 5631–5640 (1976).
31. Pedersen, B. P., Buch-Pedersen, M. J., Morth, J. P., Palmgren, M. G. & Nissen, P. Crystal structure of the plasma membrane proton pump. *Nature* **450**, 1111–1114 (2007).
32. Agmon, N. The Grotthuss mechanism. *Chem. Phys. Lett.* **244**, 456–462 (1995).
33. Hsu, S. H. *et al.* Purification, characterization, and spectral analyses of histidine-tagged vacuolar H<sup>+</sup>-pyrophosphatase expressed in yeast. *Bot. Stud.* **50**, 291–301 (2009).
34. López-Marqués, R. L. *et al.* Large-scale purification of the proton pumping pyrophosphatase from *Thermotoga maritima*: a ‘Hot-Solve’ method for isolation of recombinant thermophilic membrane proteins. *Biochim. Biophys. Acta* **1716**, 69–76 (2005).
35. McPherson, A. Current approaches to macromolecular crystallization. *Eur. J. Biochem.* **189**, 1–23 (1990).
36. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data in oscillation mode. *Methods Enzymol.* **326**, 276–307 (1997).
37. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D* **59**, 1131–1137 (2003).
38. Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022–1025 (1997).
39. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
40. Vagin, A. A. *et al.* REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D* **60**, 2184–2195 (2004).
41. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 (1996).
42. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
43. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
44. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
45. Liu, T. H. *et al.* The proximity between C-termini of dimeric vacuolar H<sup>+</sup>-pyrophosphatase determined using atomic force microscopy and a gold nanoparticle technique. *FEBS J.* **276**, 4381–4394 (2009).
46. Heikinheimo, P. *et al.* Toward a quantum-mechanical description of metal-assisted phosphoryl transfer in pyrophosphatase. *Proc. Natl Acad. Sci. USA* **98**, 3121–3126 (2001).

## Acknowledgements

This research was supported by grants to A.G. from the BBSRC (ALERT-13 and BB/M021610/1), the Royal Society (Wolfson award), The Wellcome Trust (091322/2/10/2)

and EDICT (201924), to A.G. and T.K. from the Academy of Finland (12522061 and 256049), to A.G. and J.K. from the Sigrid Juselius foundation, to C.W. (Leeds 110 anniversary fellowship), to L.J.C.J. from the European Research Council ((FP/2007–2013)/ERC Grant no. 280518), to Y.-J.S. from the Ministry of Science and Technology, Taiwan (104-2311-B-007-006-MY3 and 104-2627-M-007-007) and Academia Sinica, Taiwan (AS-103-TP-B11). We thank Seija Mäki and the Biocenter Finland Crystallisation Facility for dedicated support, Riina Honkanen for producing the protein used in this study and Matthias Gantner for assistance using the Nanion SURFE<sup>2</sup>R N1. We thank Diamond Light Source for access to beam line i04 (mx8030-6), the National Synchrotron Radiation Research Center (NSRRC) for access to beam line 15A1 and Spring-8 for access to beam line 44XU.

## Author contributions

Y.-J.S. and A.G. designed the research; K.-M.L., C.W., J.K., J.-Y.T., T.K. and L.J.C.J. performed the experiments; C.W., K.-M.L., J.K. and T.K. performed model building and refinement; C.W., K.-M.L., J.K. and A.G. analysed the structures; and all authors wrote the paper.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Li, K.-M. *et al.* Membrane pyrophosphatases from *Thermotoga maritima* and *Vigna radiata* suggest a conserved coupling mechanism. *Nat. Commun.* **7**, 13596 doi: 10.1038/ncomms13596 (2016).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016