An intracellular pathway controlled by the N-terminus of the pump subunit inhibits the bacterial KdpFABC ion pump in high K^+ conditions

Himanshu Khandelia^{a,*}, David Stokes^b, Bjørn Panyella Pedersen^c, Vikas Dubey^a

^a*PHYLLIFE: Physical Life Science, Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, Odense 5230 M, Denmark*

^b*Skirball Institute, Dept. of Cell Biology, New York University School of Medicine, New York, NY 10016 USA*

^c*Department of Molecular Biology and Genetics, Aarhus University, 8000 Aarhus C, Denmark*

Abstract

The heterotetrameric bacterial KdpFABC transmembrane protein complex is an ion channel-pump hybrid that consumes ATP to import K^+ against its transmembrane chemical potential gradient in low external K^+ environments. The KdpB ion-pump subunit of KdpFABC is a P-type ATPase, and catalyses ATP hydrolysis. Under high external K^+ conditions, K^+ can diffuse into the cells through passive ion channels. KdpFABC must therefore be inhibited in high K^+ conditions to conserve cellular ATP. Inhibition is thought to occur via unusual phosphorylation of residue Ser162 of the TGES motif of the cytoplasmic A domain. It is proposed that phosphorylation most likely traps KdpB in an inactive E1-P like conformation, but the molecular mechanism of phosphorylation-mediated inhibition and the allosteric links between phosphorylation on the A domain and the inactivation of the pump remain unknown. Here, we employ molecular dynamics (MD) simulations of the dephosphorylated and phosphorylated versions of KdpFABC to demonstrate that phosphorylated KdpB is trapped in a conformation where the ion-binding site is hydrated by an intracellular pathway between transmembrane helices M1 and M2 which opens in response to the rearrangement of

*Corresponding author

Email address: hkhandel@sdu.dk (Himanshu Khandelia)

cytoplasmic domains resulting from phosphorylation. Cytoplasmic access of water to the ion-binding site is accompanied by a remarkable loss of secondary structure of the KdpB N-terminus and disruption of a key salt bridge between Glu87 in the A domain and Arg212 in the P domain. Our results provide the molecular basis of a unique mechanism of regulation amongst P-type ATPases, and suggest that the N-terminus has a significant role to play in the conformational cycle and regulation of KdpFABC

Keywords: Molecular Dynamics simulations, KdpFABC, P-type ATPases, Ion pumps, Membrane Proteins

1. Introduction

KdpFABC is a unique K^+ ion transporting complex expressed in bacteria under low external K^+ concentrations. Bacterial survival under such conditions hinges upon sufficient inward transport of K^+ , and the KdpFABC complex consumes ATP to transport K^+ against its concentration gradient into the bacterial cytoplasm. The protein thus maintains the requisite high intracellular concentration of K^+ in low K^+ growth conditions [1]. Under high external K^+ conditions, the pump is not required because K^+ can diffuse into the cell through constitutively expressed K^+ channels. KdpFABC is a heterotetramer, and the two main functional transport subunits are the ion-channel like subunit KdpA and the pump like subunit KdpB. The KdpA subunit most likely descends from the bacterial KscA family of ion channels, while KdpB clearly belongs to the P-type ATPase family of ion pumps which also include the Na, K ATPase, Ca^{2+} -ATPase (SERCA) and the gastric H^+ , K^+ ATPase [2]. KdpFABC is therefore a one-of-a-kind, unique channel-pump hybrid and an evolutionary outlier. In ion pumps such as SERCA, the ion-binding sites are located in the transmembrane region. The protein undergoes transitions between the so-called E1 and E2 states [3, 4], which face the two different sides of the membrane and have differential affinities for different ions, making the ion-binding site more selective towards the ion present at a lower concentration. The conformational transitions between the E1 and E2 states are driven by the binding of ions to the TM ion-binding sites, and by the ATP-driven phosphorylation and dephosphorylation of the cytoplasmic P domain which contains a canonical aspartate residue which gets phosphorylated. ATP binds to the nucleotide-binding N domain which is an internal kinase and phosphorylates the aspartate residue on the P domain. Finally,

the actuator A domain uses a TGES motif to catalyse the dephosphorylation of the aspartate residue in the second half of the cycle [5]. The dephosphorylation and phosphorylation of the P domain are accompanied by significant rearrangements in the relative orientations of the cytoplasmic domains, and this rearrangement drives reorientation of the TM helices. The rearrangement of the TM helices opens intracellular or extracellular pathways leading to the ion-binding sites, through which ions bind, or are released. In this way, ions are transported across the membrane through the TM region. For a more detailed review of the conformational changes which occur during the complete transport cycle of P-type ATPases, the reader can consult several outstanding reviews on the subject [6, 7, 5]. However, the exact mechanism of transport of K^+ through KdpFABC, and how the KdpA subunit is coupled to the KdpB subunit remains unclear. Crystal [8] and cryo-EM [9] structures show that K^+ ions bind to the KdpA subunit in its outside open state. It is not clear however, how the external gate of the KdpA subunit closes, or whether the K^+ ions are transported through the channel subunit, or are transported from the KdpA to the KdpB subunit, and then released into the cytoplasm through the KdpB subunit [10]. The latter scenario is unlikely because in the available high dimensional structures of the pump, there is no conduit of large enough diameter between the KdpA and the KdpB subunit through which a K^+ ion can travel. Solving the problem of mechanism of TM ion transport probably requires the availability of the complex in conformations different from those observed so far. The focus of this paper is the mechanism of regulation of KdpFABC in high external K^+ conditions.

In both crystal [8] (Figure 1) and cryo-EM structures [9] of the KdpFABC complex, the Serine residue (Ser162) in the TGES motif is found to be phosphorylated [11], while it is well established that the preceding Glutamate residue (Glu161) catalyses the hydrolysis of the phosphorylated aspartate residue. The role of serine phosphorylation has now been traced to the inhibition of the pump when the cells are transferred to a high K^+ medium. Under high K^+ conditions, the KdpFABC pump must be inhibited [12, 13], to avoid ATP hydrolysis and waste of cellular resources when inward ion transport can proceed through passive ion channels without expending energy. Therefore, phosphorylation of the Serine residue (Ser162) is likely to trap the complex into a conformation that can no longer cycle between the E1 and E2 states, thus preventing ATP turnover. The molecular mechanism by which phosphorylation of Ser162 inhibits KdpFABC remains unknown. The Ser162Ala mutation makes the protein constitutively active, while the

Figure 1: Starting setup of the simulations performed in this work, showing the crystal structure of the KdpFABC heterotetramer complex, pdb id 5MRW. The channel KdpA subunit is coloured green while the KdpC and KdpF subunits are orange. A water molecule (red sphere) is visible in the KdpB pump canonical ion-binding site, and a K⁺ ion (purple) is resolved in the KdpA subunit. Residue Ser162 of the TGES motif in the A domain is phosphorylated. Lipid headgroups are coloured tan. For clarity, lipid tails, hydrating water molecules on both sides of the membrane and bulk ions are not shown.

65 phosphorylation mimicking mutation Ser162Asp inactivates the protein [11].
66 Like in the inactivating Glu161Gln mutation, the pump is stuck in an E1-P
67 or E2-P conformation in the Ser162Asp mutation [11]. Most importantly, the
68 pump and channel subunits become uncoupled in either of these mutations,
69 suggesting that phosphorylation (or lack thereof) of Ser162 has long-range
70 allosteric effects that influence the conformational flexibility and transitions
71 of the pump during the transport cycle. Here, we employ Molecular Dynam-
72 ics (MD) simulations of the unphosphorylated (SER) and phosphorylated
73 (PHOS) Ser162 versions of KdpFABC to show that phosphorylation drives
74 the opening of an intracellular water-filled pathway leading to the KdpB
75 ion-binding sites, through TM helices M1, M2 and M4. The access to this
76 pathway is controlled by the interactions of the N-terminus of KdpB with the
77 A domain. We propose that the opening of this previously unknown intracel-
78 lular pathway can trap KdpFABC in one of the afore-mentioned phoshy-

79 lated EP states, thus applying brakes on the pump catalytic cycle and on
80 ATP hydrolysis.

Figure 2: (A) The root mean squared deviation of the cytoplasmic domains of the protein backbone of the PHOS and the SER simulations over 10 different trajectories of 200 ns each. (B) (C) and (D) A histogram of the RMSD of the A, N and P domains. All three cytoplasmic domains deviate more from the starting structure in the PHOS simulations.

81 2. Results and Discussion

82 2.1. Cytoplasmic domains are highly mobile in phosphorylated Kdp

83 The simulations reveal distinct and reproducible differences in the be-
84 havior of KdpB that are attributable to serine phosphorylation. The first is
85 the mobility of the cytoplasmic domains of KdpB. In particular, the average
86 root mean squared deviations of each cytoplasmic domain relative to their
87 starting positions are significantly higher in PHOS simulations, which can be
88 seen in the RMSD distributions for the individual domains (Figure 2). The
89 *RMSD* fluctuates between 0.2 and 0.6 nm for both systems, suggesting an
90 overall stable tertiary structure in the simulations (Figure 2A). The higher
91 deviations from the starting structure reflect that the protein samples more
92 conformations in the PHOS simulations than in the SER simulations. The
93 possible functional relevance of this is discussed in the following sections. A
94 covariance and principal components analysis further elucidates the differing
95 dynamic behavior in PHOS and SER simulations. The first two eigenvectors

of the covariance matrix of atomic fluctuations represent the largest deviations in the dataset (Figure 3) and are used to analyze the global collective motions. For this analysis, each protein conformation sampled during the simulations was projected upon the two dimensional space formed by the first two eigenvectors to obtain a 2D-scatter plot (Figure 3A). In addition, a histogram representing the 1D motion along each eigenvector is shown along the margin of each axis. This projection map reveals several clouds which correspond to distinct protein conformations sampled during the simulations. These clouds are likely to represent free energy minima in the protein conformational landscape which are clearly different for PHOS (blue) and SER (orange) simulations. The analysis is similar to what we performed earlier in the context of a different membrane protein [14]. More specifically, this analysis reveals novel conformations, labeled C and D, that are sampled in the PHOS simulations, whereas the SER simulations remain within the larger cloud of conformations labeled B. A *k-means* clustering algorithm was used to extract representative conformations from each conformational cluster, which are illustrated in panels B, C and D of Figure 3). In conformations C and D, sampled only in the PHOS simulations, the A-domain (yellow) uncouples from the P-domain (blue), and collapses towards the membrane.

Figure 3: Principle component analysis of conformations sampled during the simulations. (A) Projection of individual frames from the simulation onto the first two eigenvectors (PC1 and PC2) from the covariance matrix. The prevalent conformation represented by the cloud labeled B is sampled by both PHOS and SER simulations. However, the PHOS simulation has sampled novel conformations represented by the peripheral clouds labeled C and D. Projections onto each individual axis are shown along the top and right margins. (B-D) Structures extracted by *k-means* analysis of panel A from clouds B, C and D, respectively. Cytoplasmic domains are colored as follows: A-domain yellow, N-domain red, P-domain blue.

115 Trajectories projected upon the first eigen vectors for the PHOS and SER
116 simulations are shown in Movies S1 and S2.

117 2.2. Phosphorylation opens an intracellular water pathway leading to the 118 *KdpB* ion-binding site

119 The second key difference involves the canonical ion-binding site of P-
120 type ATPases formed near Pro264 on the M4 helix. This site is occupied
121 by a water molecule in the crystal structure of KdpB (Figure 1). In the
122 PHOS simulations, this water molecule escaped almost immediately, whereas
123 it remains bound for the duration of the SER simulations.

Figure 4: Water accessibility of the canonical ion binding site in KdpB. (A) Radial distribution functions between water and the center of mass of residues Pro264, Thr265, Asp583 and Lys586 from KdpB, which surround the canonical ion binding site. The radial distribution function corresponds to a normalized histogram of distances between two atom selections, averaged over the time of the simulation. (B) Initial configuration of the canonical site for both the SER and PHOS simulations with a water molecule (blue) bound at the location seen in the X-ray structure. (C) and (D) Final frames from PHOS and SER simulations respectively. The canonical ion-binding site is significantly more hydrated in the PHOS simulations with water penetrating the site from the cytoplasmic side of the membrane (top of the panel).

124 This escape is facilitated by the opening of an aqueous channel in KdpB
125 from the cytoplasmic side of the membrane (Figure 4, and supplementary
126 movie S3). As a result, the Pro264-bound water molecule is frequently ex-
127 changed with a different water molecule in the aqueous channel. In contrast,

128 water never reaches the canonical binding site and the initial water molecule
 129 remained occluded in any of the 10 SER simulations. (Figure 4). To quan-
 130 tify the penetration of water into this site, the radial distribution function
 131 $g(r)$, which is a normalised histogram of pair-wise residue distances, is calcu-
 132 lated for all water molecules relative to the center of mass defined by Pro264,
 133 Thr265, Asp583, and Lys586, which surround the site and coordinate the wa-
 134 ter molecule in the crystal structure (Figure 4A). These data indicate that
 135 this intramembrane site becomes highly hydrated when Ser162 on KdpB is
 136 phosphorylated, but not for the unphosphorylated Kdp complex.

137 The water pathway opens between helices M1 and M2, and in the case of
 138 the SER simulations, the pathway is closed by a salt bridge between Glu87 on
 139 the A-domain and Arg212 on the P-domain (Figure 5D). The salt bridge is
 140 broken in the PHOS simulations as a result of the large scale motions of the
 141 cytoplasmic domains, and in this way separating helices M1 and M2, and
 142 opening the water channel. An intracellular pathway between helices M1,
 143 M2 and M4 is also postulated for the sarcoplasmic reticulum Ca^{2+} -ATPase
 144 [15], and was observed in the first crystal structure of SERCA [16], as well as
 145 in inhibitor-bound structures of SERCA [17]. The M1 helix probably plays
 146 a role in the occlusion of the bound Ca^{2+} in SERCA [18, 19]. Palytoxin,
 147 which converts the Na^+ , K^+ ATPase to a channel-like protein, opens a con-
 148 tinuous ion-pathway which also passes between M1, M2, M4 and M6 on the
 149 cytoplasmic side [20]. Whether or not the observed pathway leading to the
 150 ion-binding sites in our simulations of the KdpFABC comprise of a functional
 151 pathway for K^+ release into the cytoplasm remains to be seen. However, it
 152 is clear that similar pathways are linked to functional roles in other P-type
 153 ATPases. In the case of KdpFABC, the pathway could only serve to trap
 154 the protein into an inactive state, and not be responsible for ion release into
 155 the cytoplasm. The pathway observed in the simulations is not seen in any
 156 of the available structures of KdpFABC to date. It is possible that crystal
 157 and cryo-EM structures only samples a subset of the large conformational
 158 landscape sampled by the phosphorylated pump. So far, no structure of the
 159 active dephosphorylated pump has emerged.

160 2.3. Hydration of the ion-binding site is accompanied by changes in the con- 161 formations of the N-terminus

162 The opening of the intracellular pathway leading to the ion-binding sites
 163 is accompanied by remarkable changes in the secondary structure of the N-
 164 terminus of the KdpB. In the SER simulations, residues 11 through 23 remain

Figure 5: Interactions between M1 and M2 control entry of water into the canonical ion binding site. (A) Snapshot from the PHOS simulations shows that Glu87 and Arg212 are displaced and a water pathway is detected by CAVER (green surface) that leads to the canonical binding site. (B) Snapshot from the SER simulations shows the proximity of Glu87 on M1 and Arg212 on M2. The canonical binding site is below and no water pathway is present. (C) Radial distribution function between the side-chain center of masses of Glu87 and Arg212. Higher peaks for SER simulations indicate stronger interactions between these residues in this system, whereas the lower peaks for PHOS simulations indicate disruption of these interactions. (D) The helical content of the linker between M2 and the A-domain (residues 68-92) is significantly lower for the PHOS simulations indicating unwinding of this structural element.

165 helical most of the time during the simulations. In the PHOS simulations,
 166 however, the helical content of these residues is nearly completely lost (Fig-
 167 ure 5D). The conformations of the N-terminus relative to the ion-binding site
 168 are shown in (Figure 1). In the SER simulations, the helical N-terminus is
 169 linked to the M1 via a random-coil linker comprising of residues 26 through
 170 34, with 27 and 34 being helix-breaking Proline residues. Part of the N-
 171 terminus remains membrane-bound, but the rest projects into the cytoplasm
 172 to interact closely with residues 83 through 98 on the M2 helix which also
 173 protrude into the cytoplasm (Figure 6C). In the PHOS simulations however,
 174 the N-terminus is fully unfolded, and collapses parallel to the membrane. In
 175 addition, the M2 helix also unfolds near residues 87 through 92, breaking
 176 the salt bridge between Glu87 on M2 and Arg212 on top of helix M4. These
 177 data are depicted in Figure 6, where it is clear that the distance between
 178 the N-terminus and the residues 87 through 92 fluctuates significantly, and

179 is higher in the PHOS simulations, compared to the SER simulations. These
180 data, combined with the earlier claim that the intracellular pathway between
181 helices M1, M2 and M4 could play a functional role, suggest that the N-
182 terminus of KdpFABC has an important role to play in the conformational
183 cycle of the pump, and may even regulate the release of the K^+ ion into
184 the cytoplasm. The N-terminus was previously shown to regulate the E1-E2
185 conformational equilibrium of the Na^+ , K^+ , ATPase by interacting with the
186 cytoplasmic domains [21, 22], and it has since been hypothesized that the
187 N-terminus may play a role in the regulation or the conformations of both
188 the Na^+ , K^+ , ATPase and the H^+ , K^+ , ATPase [23].

189 Taken together, there are several dynamic phenomena that occur in the
190 PHOS simulations, and which are absent in the SER systems, which lead to
191 the formation of the water pathway leading to the ion-binding sites. These
192 changes are: the unfolding of the N-terminus, the unfolding of the M2 helix
193 between residues 86 and 91, the large-scale movement of the cytoplasmic
194 domains and the broken salt bridge between Glu87 and Arg212. We are
195 reluctant to ascribe any of these as a single potential decisive step which
196 lead to hydration of the binding site. Current models of allostery rather
197 emphasize the dynamic nature of many small conformational changes which
198 add up to altered dynamics of a conformational ensemble of a protein upon
199 binding of an allosteric ligand. We advocate this ensemble view of allostery
200 according to which reactions such as phosphorylation can regulate protein
201 function allosterically without gross conformational changes.

Figure 6: (A) The average distances between the center of mass of the N-terminus (residues 9 through 25) and the cytoplasmic portion of the M2 helix (residues 87 through 92) over all 10 simulations. Except in the first simulation replica, the distance increases and fluctuates more for all PHOS replicas compared to the SER replicas. (B) The histogram of distances shown in A. C and D: The interactions between the N-terminus (green) and residues 83 through 98 (orange) in the A domain in the SER and PHOS simulations respectively. The N-terminus is fully unfolded in the PHOS simulations, but remains helical in the SER simulations. The color coding for the cytoplasmic domains is the same as Figure 1.

202 To conclude, we have carried out MD simulations of phosphorylated and
 203 dephosphorylated versions of the bacterial KdpFABC complex to propose
 204 that the phosphorylation inhibits the protein by opening an intracellular
 205 water pathway leading to the ion-binding sites in the KdpB pump subunit.
 206 We also hypothesize that the N-terminal residues guard the entrance to the
 207 ion-binding site in the absence of Ser162 phosphorylation, and in this way, the
 208 KdpFABC N-terminus may have a significant role to play in the regulation
 209 of the pump. Since it can control the entry of water (and ions) to the pump
 210 subunit ion-binding site, the N-terminus is also likely to be involved in the
 211 regulation and kinetics of the overall conformational cycle of the pump. Our
 212 hypotheses can be experimentally validated or refuted by simple mutational
 213 experiments such as the measurement of pump activity and ATP turnover
 214 in a mutant where the N-terminus is truncated. Our simulations are not
 215 very long by current standards. However, all simulation replicas point to
 216 the same observation with respect to the hydration of the ion-binding site,
 217 lending confidence to our sampling protocols. Simulations using the cryo-EM
 218 structures as starting models will be implemented in the future, although
 219 these structures have a lower resolution than the crystal structure used as a
 220 starting model in the current simulations.

221 3. Materials and Methods

222 To set up the system for simulation, a bilayer containing 550 lipid molecules
 223 was constructed using CHARMM-GUI with POPC and POPE in a 4:1 ratio,
 224 as a simple mimic of the bacterial membrane. The published crystal struc-
 225 ture for KdpFABC (PDB: 5MRW), including the bound potassium ion as
 226 well as the protein, was introduced in the bilayer with the help of the OPM
 227 server and the system was hydrated with 75,000 water molecules. We used
 228 the crystal structure because it had the highest resolution. 150mM of salt
 229 (KCl) was added to the aqueous phase. The protonation states of the titrat-
 230 able amino acids were initially set by PROPKA [24]. Several residues had
 231 a pKa higher than 7, and were therefore protonated. However, a short 50
 232 ns simulation run showed that all acidic amino acids had a pKa values well
 233 below 7.0. With this in mind, we kept all acidic amino acids deprotonated
 234 during the production simulations. Ten independent simulations of 200 ns
 235 each were launched for KdpFABC with and without serine phosphorylation,
 236 referred to as PHOS and SER simulations respectively.

237 All-atom MD simulations were performed under periodic boundary con-
 238 ditions using GROMACS version 2016.3 [25, 26, 27, 28, 29]. 10 replicates
 239 with different initial velocities were simulated for both the PHOS and the
 240 SER simulations. We used the TIP3P water model with Lennard-Jones in-
 241 teractions on hydrogen atoms. A 1.2 *nm* cutoff was used for non-bonded
 242 neighbour list, and was updated every 10 steps. The van der Waals interac-
 243 tions were switched off between 1.0 to 1.2 *nm*. Electrostatic interactions were
 244 treated with the particle mesh Ewald (PME) method [30, 31]. All systems
 245 were minimised with 5,000 steps using steepest decent algorithm. For each
 246 simulation, the converged system was equilibrated for 5 ns with backbone
 247 restraints on the protein followed by a production run of 200 ns. The tem-
 248 perature of the system was kept at 310K with the Nosé-Hoover thermostat
 249 [32, 33] after an equilibration run which was performed with the Berendsen
 250 thermostat [34]. The pressure was kept at 1 bar with the Parrinello-Rahman
 251 barostat [35] along with a semi-isotropic pressure coupling scheme. The Lin-
 252 ear Constraint Solver (LINCS) [36] algorithm was used to constrain all bonds
 253 containing hydrogen. A 2 fs time step was used and trajectories were sam-
 254 pled every 20 ps. The data analysis was carried out using GROMACS and
 255 homemade scripts. The analysis was performed on the concatenated trajec-
 256 tories from the 10 individual simulations. Snapshots were rendered using
 257 Visual Molecular Dynamics (VMD)[37]. Data analysis was performed on a
 258 single concatenated trajectory for each of the three systems, unless other-
 259 wise mentioned. Trajectories for the center of mass of different domains were
 260 extracted by MDAnalysis tools [38, 39] and distance and angular data were
 261 calculated by homemade scripts.

262 Covariance and principal components analysis (PCA) was performed us-
 263 ing GROMACS. After removing center of mass translation and rotation, the
 264 covariance analysis was performed for the protein backbone atoms. First, a
 265 covariance matrix C is extracted from atomic fluctuations

$$C = \langle (\mathbf{x}(t) - \langle \mathbf{x} \rangle)(\mathbf{x}(t) - \langle \mathbf{x} \rangle)^T \rangle \quad (1)$$

266 where $\mathbf{x}$ is a $3N$ -dimensional column vector describing the coordinates of
 267 the N protein backbone atoms, and $\mathbf{x}(t)$ are the positions at time t . The
 268 triangular brackets indicate an ensemble average. C is a $3N \times 3N$ symmetric
 269 matrix, and is then diagonalised by a coordinate tranformation R

$$C = R\Lambda R^T \quad (2)$$

$$\Lambda = R^T C R \quad (3)$$

270 This orthogonal transformation transforms C into a diagonal matrix $\Lambda =$
 271 $diag(\lambda_1, \lambda_2, \dots, \lambda_{3N})$, which contains the eigenvalues λ_i of the covariance matrix
 272 C . The i th column of R contains the i th eigenvector $\mathbf{r}_i$ with the corresponding
 273 eigenvalue λ_i .

274 The global collective motions of the trajectory can then be obtained by
 275 projecting the trajectory ensemble onto individual eigenvectors, to obtain the
 276 principal components $\mathbf{p}_i, i = 1, 2, \dots, N$ by taking an internal product between
 277 the transpose of the eigenvector and the atomic fluctuation.

$$p_i(t) = \mathbf{r}_i \cdot (\mathbf{x}(t) - \langle \mathbf{x} \rangle) \quad (4)$$

278 Each trajectory snapshot, thus projected on two different eigenvectors, yields
 279 one point on the point cloud distributions shown in Figure 3.

280 4. Acknowledgements

281 HK and VD are funded by Lundbeckfonden grant numbers R82-2011-
 282 7280. The simulations were performed on the ABACUS 2.0 supercomputer
 283 at the Danish e-Science Center at SDU, the Piz Daint supercomputer un-
 284 der grant number pr98 and on the Kay supercomputer under the PRACE
 285 DECI grant number ANNEX. The financial support by the Novo Nordisk
 286 Foundation grant (number NF18OC0032608) for the ROBUST-project is also
 287 acknowledged. DS acknowledges funding from the NIH grant number R01
 288 GM108043

289 The authors thank Wojciech Kopec for useful discussions.

290 5. References

- 291 [1] A. Ballal, B. Basu, S. K. Apte, The kdp-atpase system and its regu-
 292 lation, J Biosci 32 (2007) 559–68. URL: [https://www.ncbi.nlm.nih.](https://www.ncbi.nlm.nih.gov/pubmed/17536175)
 293 [gov/pubmed/17536175](https://www.ncbi.nlm.nih.gov/pubmed/17536175). doi:10.1007/s12038-007-0055-7.
- 294 [2] M. Bramkamp, K. Altendorf, J.-C. Greie, Common patterns and
 295 unique features of p-type ATPases: a comparative view on the
 296 KdpFABC complex from Escherichia coli (review), Molecular Mem-
 297 brane Biology 24 (2007) 375–386. URL: [https://doi.org/10.1080/](https://doi.org/10.1080/09687680701418931)
 298 [09687680701418931](https://doi.org/10.1080/09687680701418931). doi:10.1080/09687680701418931.

- 299 [3] R. L. Post, S. Kume, T. Tobin, B. Orcutt, A. K. Sen, Flexibil-
300 ity of an active center in sodium-plus-potassium adenosine triphos-
301 phatase, *Journal of General Physiology* 54 (1969) 306–326. URL: <https://doi.org/10.1085/jgp.54.1.306>. doi:10.1085/jgp.54.1.306.
302
- 303 [4] R. W. Albers, Biochemical aspects of active transport, *An-*
304 *ual Review of Biochemistry* 36 (1967) 727–756. URL: <https://doi.org/10.1146/annurev.bi.36.070167.003455>. doi:10.1146/annurev.
305 bi.36.070167.003455.
306
- 307 [5] J. V. MÅyler, C. Olesen, A. M. Winther, P. Nissen, The sarcoplasmic
308 ca^{2+} -atpase: design of a perfect chemi-osmotic pump, *Quarterly reviews*
309 *of biophysics* 43 (2010) 501–66. doi:10.1017/S003358351000017X.
- 310 [6] M. G. Palmgren, P. Nissen, P-type atpases, *Annual review of*
311 *biophysics* 40 (2011) 243–66. URL: <http://www.ncbi.nlm.nih.gov/pubmed/21351879>. doi:10.1146/annurev.biophys.093008.131331.
312
- 313 [7] C. Toyoshima, R. Kanai, F. Cornelius, First crystal structures of
314 na^{+} , k^{+} -ATPase: New light on the oldest ion pump, *Structure* 19
315 (2011) 1732–1738. URL: <https://doi.org/10.1016/j.str.2011.10.016>. doi:10.1016/j.str.2011.10.016.
316
- 317 [8] C. S. Huang, B. P. Pedersen, D. L. Stokes, Crystal structure
318 of the potassium-importing kdpfabc membrane complex, *Nature*
319 546 (2017) 681–685. URL: <https://www.ncbi.nlm.nih.gov/pubmed/28636601>. doi:10.1038/nature22970.
320
- 321 [9] C. Stock, L. Hielkema, I. Tascon, D. Wunnicke, G. T. Oostergetel,
322 M. Azkargorta, C. Paulino, I. Hanelt, Cryo-em structures of kdpfabc
323 suggest a $k(+)$ transport mechanism via two inter-subunit half-channels,
324 *Nat Commun* 9 (2018) 4971. URL: <https://www.ncbi.nlm.nih.gov/pubmed/30478378>. doi:10.1038/s41467-018-07319-2.
325
- 326 [10] B. P. Pedersen, D. L. Stokes, H.-J. Apell, The KdpFABC com-
327 plex – k^{+} transport against all odds, *Molecular Membrane Biology*
328 35 (2019) 21–38. URL: <https://doi.org/10.1080/09687688.2019.1638977>. doi:10.1080/09687688.2019.1638977.
329

- 330 [11] M. E. Sweet, X. Zhang, H. Erdjument-Bromage, V. Dubey, H. Khan-
331 delia, T. A. Neubert, B. P. Pedersen, D. L. Stokes, Serine phos-
332 phorylation regulates the p-type potassium pump KdpFABC, *eLife* 9
333 (2020). URL: <https://doi.org/10.7554/elife.55480>. doi:10.7554/
334 elife.55480.
- 335 [12] D. B. Rhoads, L. Laimins, W. Epstein, Functional organization of the
336 kdp genes of *escherichia coli* k-12, *J Bacteriol* 135 (1978) 445–52. URL:
337 <https://www.ncbi.nlm.nih.gov/pubmed/355227>.
- 338 [13] A. J. Roe, D. McLaggan, C. P. O’Byrne, I. R. Booth, Rapid
339 inactivation of the *escherichia coli* kdp k⁺ uptake system by
340 high potassium concentrations, *Mol Microbiol* 35 (2000) 1235–
341 43. URL: <https://www.ncbi.nlm.nih.gov/pubmed/10712703>. doi:10.
342 1046/j.1365-2958.2000.01793.x.
- 343 [14] V. Dubey, B. Bozorg, D. Wüstner, H. Khandelia, Cholesterol binding
344 to the sterol-sensing region of niemann pick c1 protein confines dynam-
345 ics of its n-terminal domain, *PLOS Computational Biology* 16 (2020)
346 e1007554. URL: <https://doi.org/10.1371/journal.pcbi.1007554>.
347 doi:10.1371/journal.pcbi.1007554.
- 348 [15] M. Bublitz, M. Musgaard, H. Poulsen, L. Thøgersen, C. Olesen,
349 B. Schiøtt, J. P. Morth, J. V. Møller, P. Nissen, Ion pathways in the
350 sarcoplasmic reticulum ca²⁺-ATPase, *Journal of Biological Chemistry*
351 288 (2013) 10759–10765. URL: [https://doi.org/10.1074/jbc.r112.](https://doi.org/10.1074/jbc.r112.436550)
352 436550. doi:10.1074/jbc.r112.436550.
- 353 [16] C. Toyoshima, M. Nakasako, H. Nomura, H. Ogawa, Crystal structure
354 of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution, *Nature*
355 405 (2000) 647–655. URL: <https://doi.org/10.1038/35015017>.
356 doi:10.1038/35015017.
- 357 [17] M. Takahashi, Y. Kondou, C. Toyoshima, Interdomain communica-
358 tion in calcium pump as revealed in the crystal structures with trans-
359 membrane inhibitors, *Proceedings of the National Academy of Sci-*
360 *ences* 104 (2007) 5800–5805. URL: [https://doi.org/10.1073/pnas.](https://doi.org/10.1073/pnas.0700979104)
361 0700979104. doi:10.1073/pnas.0700979104.

- [18] A. P. Einholm, J. P. Andersen, B. Vilsen, Roles of transmembrane segment m1 of Na^+ , K^+ -ATPase and Ca^{2+} -ATPase, the gatekeeper and the pivot, *Journal of Bioenergetics and Biomembranes* 39 (2007) 357–366. URL: <https://doi.org/10.1007/s10863-007-9106-x>. doi:10.1007/s10863-007-9106-x.
- [19] M. M. G. Geurts, J. D. Clausen, B. Arnou, C. Montigny, G. Lenoir, R. A. Corey, C. Jaxel, J. V. Møller, P. Nissen, J. P. Andersen, M. le Maire, M. Bublitz, The SERCA residue glu340 mediates interdomain communication that guides Ca^{2+} transport, *Proceedings of the National Academy of Sciences* (2020) 202014896. URL: <https://doi.org/10.1073/pnas.2014896117>. doi:10.1073/pnas.2014896117.
- [20] A. Takeuchi, N. Reyes, P. Artigas, D. C. Gadsby, The ion pathway through the opened Na^+ , K^+ -ATPase pump, *Nature* 456 (2008) 413–416. URL: <https://doi.org/10.1038/nature07350>. doi:10.1038/nature07350.
- [21] N. Boxenbaum, S. E. Daly, Z. Z. Javaid, L. K. Lane, R. Blostein, Changes in steady-state conformational equilibrium resulting from cytoplasmic mutations of the Na , K -ATPase α -subunit, *Journal of Biological Chemistry* 273 (1998) 23086–23092. URL: <https://doi.org/10.1074/jbc.273.36.23086>. doi:10.1074/jbc.273.36.23086.
- [22] Q. Jiang, A. Garcia, M. Han, F. Cornelius, H.-J. Apell, H. Khandelia, R. J. Clarke, Electrostatic stabilization plays a central role in autoinhibitory regulation of the Na^+ , K^+ -ATPase, *Biophysical Journal* 112 (2017) 288–299. URL: <https://doi.org/10.1016/j.bpj.2016.12.008>. doi:10.1016/j.bpj.2016.12.008.
- [23] D. Diaz, R. J. Clarke, Evolutionary analysis of the lysine-rich n-terminal cytoplasmic domains of the gastric H^+ , K^+ -ATPase and the Na^+ , K^+ -ATPase, *The Journal of Membrane Biology* 251 (2018) 653–666. URL: <https://doi.org/10.1007/s00232-018-0043-x>. doi:10.1007/s00232-018-0043-x.
- [24] C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski, J. H. Jensen, Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values, *Journal of Chemical Theory and*

- 395 Computation 7 (2011) 2284–2295. URL: [https://doi.org/10.1021/](https://doi.org/10.1021/ct200133y)
396 [ct200133y](https://doi.org/10.1021/ct200133y). doi:10.1021/ct200133y.
- 397 [25] H. J. Berendsen, D. van der Spoel, R. van Drunen, Gromacs: a message-
398 passing parallel molecular dynamics implementation, Computer physics
399 communications 91 (1995) 43–56.
- 400 [26] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J.
401 Berendsen, Gromacs: fast, flexible, and free, Journal of computational
402 chemistry 26 (2005) 1701–1718.
- 403 [27] B. Hess, C. Kutzner, D. Van Der Spoel, E. Lindahl, Gromacs 4: algo-
404 rithms for highly efficient, load-balanced, and scalable molecular simu-
405 lation, Journal of chemical theory and computation 4 (2008) 435–447.
- 406 [28] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov,
407 M. R. Shirts, J. C. Smith, P. M. Kasson, D. Van Der Spoel, B. Hess,
408 E. Lindahl, Gromacs 4.5: a high-throughput and highly parallel open
409 source molecular simulation toolkit, Bioinformatics 29 (2013) 845–854.
- 410 [29] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess,
411 E. Lindahl, Gromacs: High performance molecular simulations through
412 multi-level parallelism from laptops to supercomputers, SoftwareX 1
413 (2015) 19–25.
- 414 [30] T. Darden, D. York, L. Pedersen, Particle mesh ewald: An $n \cdot \log(n)$
415 method for ewald sums in large systems, The Journal of chemical physics
416 98 (1993) 10089–10092.
- 417 [31] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Ped-
418 ersen, A smooth particle mesh ewald method, The Journal of chemical
419 physics 103 (1995) 8577–8593.
- 420 [32] S. Nosé, A unified formulation of the constant temperature molecular
421 dynamics methods, The Journal of chemical physics 81 (1984) 511–519.
- 422 [33] W. G. Hoover, Canonical dynamics: Equilibrium phase-space distribu-
423 tions, Phys. Rev. A 31 (1985) 1695–1697.
- 424 [34] H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola, J. Haak,
425 Molecular dynamics with coupling to an external bath, The Journal of
426 chemical physics 81 (1984) 3684–3690.

- 427 [35] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A
428 new molecular dynamics method, *Journal of Applied physics* 52 (1981)
429 7182–7190.
- 430 [36] B. Hess, H. Bekker, H. J. Berendsen, J. G. Fraaije, Lincs: a linear
431 constraint solver for molecular simulations, *Journal of computational*
432 *chemistry* 18 (1997) 1463–1472.
- 433 [37] W. Humphrey, A. Dalke, K. Schulten, Vmd: visual molecular dynamics,
434 *Journal of molecular graphics* 14 (1996) 33–38.
- 435 [38] R. J. Gowers, M. Linke, J. Barnoud, T. J. Reddy, M. N. Melo, S. L.
436 Seyler, D. L. Dotson, J. Domanski, S. Buchoux, I. M. Kenney, O. Beck-
437 stein, Mdanalysis: a python package for the rapid analysis of molecular
438 dynamics simulations, in: *Proceedings of the 15th Python in Science*
439 *Conference*, volume 98, 2016.
- 440 [39] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O. Beckstein, Md-
441 analysis: A toolkit for the analysis of molecular dynamics simulations,
442 *Journal of Computational Chemistry* 32 (2011) 2319–2327.