



The SERCA residue Glu340 mediates interdomain communication that guides Ca²⁺ transport

Maxwell M. G. Geurts, Johannes D. Clausen, Bertrand Arnou, Cédric Montigny, Guillaume Lenoir, Robin A. Corey, Christine Jaxel, Jesper V. Møller, Poul Nissen, Jens Peter Andersen, et al.

► To cite this version:

Maxwell M. G. Geurts, Johannes D. Clausen, Bertrand Arnou, Cédric Montigny, Guillaume Lenoir, et al.. The SERCA residue Glu340 mediates interdomain communication that guides Ca²⁺ transport. Proceedings of the National Academy of Sciences of the United States of America, 2020, 117 (49), pp.31114-31122. 10.1073/pnas.2014896117 . hal-03385675

HAL Id: hal-03385675

<https://hal.science/hal-03385675>

Submitted on 26 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dear Author

Please use this PDF proof to check the layout of your article. If you would like any changes to be made to the layout, you can leave instructions in the online proofing interface. First, return to the online proofing interface by clicking "Edit" at the top page, then insert a Comment in the relevant location. Making your changes directly in the online proofing interface is the quickest, easiest way to correct and submit your proof.

Please note that changes made to the article in the online proofing interface will be added to the article before publication, but are not reflected in this PDF proof.

If you would prefer to submit your corrections by annotating the PDF proof, please download and submit an annotatable PDF proof by clicking the link below.

 [Annotate PDF](#)

The SERCA residue Glu340 mediates interdomain communication that guides Ca^{2+} transport

Maxwell M. G. Geurts^{a,1}, Johannes D. Clausen^{b,c,1}, Bertrand Arnou^{b,d,1}, Cedric Montigny^d, Guillaume Lenoir^d, Robin A. Corey^a, Christine Jaxel^d, Jesper V. Møller^b, Poul Nissen^{c,e}, Jens Peter Andersen^b, Marc le Maire^d, and Maike Bublitz^{a,2}

^aDepartment of Biochemistry, University of Oxford, OX1 3QU Oxford, United Kingdom; ^bDepartment of Biomedicine, Aarhus University, 8000 Aarhus C, Denmark; ^cDepartment of Molecular Biology and Genetics, Aarhus University, 8000 Aarhus C, Denmark; ^dInstitute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Saclay, 91198 Gif-sur-Yvette, France; and ^eDanish Research Institute of Translational Neuroscience, Nordic EMBL Partnership for Molecular Medicine

Edited by Ivet Bahar, University of Pittsburgh School of Medicine, Pittsburgh, PA, and approved October 21, 2020 (received for review July 15, 2020)

The sarco(endo)plasmic reticulum Ca^{2+} -ATPase (SERCA) is a P-type ATPase that transports Ca^{2+} from the cytosol into the sarco(endo)plasmic reticulum (SR/ER) lumen, driven by ATP. This primary transport activity depends on tight coupling between movements of the transmembrane helices forming the two Ca^{2+} -binding sites and the cytosolic headpiece mediating ATP hydrolysis. We have addressed the molecular basis for this intramolecular communication by analyzing the structure and functional properties of the SERCA mutant E340A. The mutated Glu340 residue is strictly conserved among the P-type ATPase family of membrane transporters and is located at a seemingly strategic position at the interface between the phosphorylation domain and the cytosolic ends of 5 of SERCA's 10 transmembrane helices. The mutant displays a marked slowing of the Ca^{2+} -binding kinetics, and its crystal structure in the presence of Ca^{2+} and ATP analog reveals a rotated headpiece, altered connectivity between the cytosolic domains, and an altered hydrogen bonding pattern around residue 340. Supported by molecular dynamics simulations, we conclude that the E340A mutation causes a stabilization of the Ca^{2+} sites in a more occluded state, hence displaying slowed dynamics. This finding underpins a crucial role of Glu340 in interdomain communication between the headpiece and the Ca^{2+} -binding transmembrane region.

Ca^{2+} binding | P-type ATPase | SERCA | tryptophan fluorescence | molecular dynamics simulations

The Ca^{2+} -ATPase of sarco(endo)plasmic reticulum (SERCA) is an ion-translocating ATPase belonging to the P-type family of membrane transporters. It pumps cytosolic Ca^{2+} ions across the sarco- or endoplasmic reticulum (SR/ER) membrane at the expense of ATP, a vital function in all living cells, particularly in the context of muscle contraction, Ca^{2+} signaling, and cell survival (1).

The molecular Ca^{2+} transport mechanism of SERCA is based on a cyclic transition through different conformations of the 110-kDa membrane protein, allowing alternating access to the Ca^{2+} -binding sites from the cytosol and the SR/ER lumen. Binding and hydrolysis of ATP is mediated by SERCA's cytosolic "headpiece" which consists of three roughly globular domains, the nucleotide-binding (N), phosphorylation (P), and actuator (A) domains, while the Ca^{2+} -binding sites are located in the transmembrane (TM) part of the protein, which consists of 10 α -helices (M1 through M10) (Fig. 1A). ATP binding and the subsequent transient phosphorylation of a conserved aspartate residue are coupled to a sequence of binding of two Ca^{2+} ions from the cytosol, followed by occlusion (a state where the ions are shielded from access to either side of the membrane), and luminal release of these ions. This is associated with conformational transitions between so-called *E1*, *E1P*, *E2P*, and *E2* forms (Fig. 1B, "P" representing phosphorylation and brackets representing ion occlusion) (1, 2).

A number of SERCA1a crystal structures have shed light on the nature of the conformational changes associated with Ca^{2+} transport (reviewed in refs. 3–5). There are two Ca^{2+} -binding sites within the TM domain of SERCA, denoted sites I and II based on a proven sequential order of Ca^{2+} binding (6). The Ca^{2+} coordination in site I is mediated by residues in helices M5, M6, and M8, and in site II by residues in M4 and M6. Ca^{2+} binding from the cytosol to SERCA involves the characteristically kinked M1 helix (Fig. 1A): a large vectorial movement of M1 toward the luminal side of the membrane mediates Ca^{2+} access from the cytosolic side, whereas Ca^{2+} occlusion after binding depends on an M1 movement in the opposite direction, allowing the formation of a hydrophobic cluster around the kink that shields the Ca^{2+} sites. This goes along with a smaller shift of M3 relative to M5 (7, 8). The transition from the *E1P* to the *E2P* state is associated with a large rotation of the A domain, which causes a distortion of the coordination geometry at the high-affinity Ca^{2+} -binding sites and a concomitant loss of Ca^{2+} affinity, along with ADP release (9) and the formation of a luminal exit pathway for Ca^{2+} release (10).

Significance

We present a crystal structure, functional data, and molecular dynamics (MD) simulations of the sarco(endo)plasmic reticulum Ca^{2+} -ATPase (SERCA) mutant E340A. The mutation slows Ca^{2+} -binding kinetics, and the structural differences between wild type and E340A indicate that the mutation disrupts a central interdomain "communication hub" governing Ca^{2+} binding/dissociation. MD simulations reveal altered dynamics in regions mediating Ca^{2+} occlusion, a critical step in SERCA's alternating access mechanism. The mutation stabilizes a more occluded state of the Ca^{2+} sites. The strict conservation of Glu340 among P-type ATPases is the result of its critical role in interdomain communication between the cytosolic headpiece and the transmembrane domain, ensuring a delicate balance between dynamics of ion binding, occlusion, and release—key steps in the transport process.

Author contributions: J.V.M., P.N., J.P.A., M.I.M., and M.B. designed research; M.M.G.G., J.D.C., B.A., C.M., G.L., R.A.C., C.J., and M.B. performed research; M.M.G.G., C.M., G.L., R.A.C., and M.B. analyzed data; M.B. wrote the paper; and C.M., G.L., P.N., J.P.A., and M.I.M. contributed to the final editing of the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

¹M.M.G.G., J.D.C., and B.A. contributed equally to this work.

²To whom correspondence may be addressed. Email: maike.bublitz@bioch.ox.ac.uk.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2014896117/-DCSupplemental>.

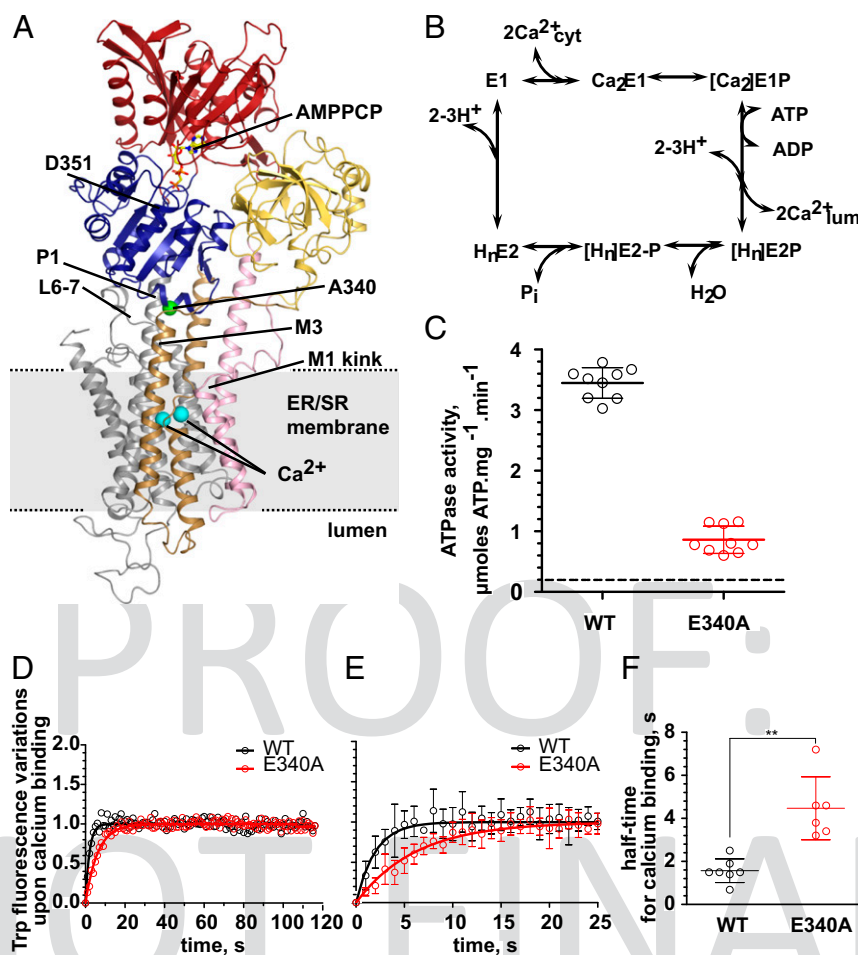


Fig. 1. SERCA reaction scheme, structure, and activity measurements. (A) Crystal structure of SERCA E340A shown as cartoon. Nucleotide-binding (N) domain is red, actuator (A) domain yellow, phosphorylation (P) domain blue, M1-2 pink, M3-4 brown, M5-10 gray. AMPPCP is shown as ball-and-stick, Ca²⁺ ions as cyan spheres. Residue 340 is indicated by a green sphere. (B) Schematic reaction cycle of SERCA. (C) Steady-state ATPase activity of purified WT SERCA (black) or E340A mutant (red) in an enzyme-coupled assay. WT: 3.45 ± 0.26 ($n = 9$); E340A: 0.86 ± 0.22 ($n = 9$); background (dashes): 0.18 ± 0.06 ($n = 18$). Three independent experiments were done from two different batches of purified protein. (D) Time course of Ca²⁺-binding transition determined by tryptophan fluorescence. Each time point represents the average of several measurements ($n = 7$ for WT and $n = 6$ for E340A from two independent experiments). (E) Magnified view of the early time points in D. Error bars correspond to the SD from the mean ($n = 7$ for WT and $n = 6$ for E340A). (F) Half-time for Ca²⁺ binding for WT SERCA and E340A. $t_{1/2}$ were determined on each individual experiment. $t_{1/2}$ are 1.6 ± 0.6 s for WT and 4.5 ± 1.5 s for E340A with $P = 0.0023$ (one-tailed paired t test).

Both the P and the A domains of the cytosolic headpiece are in contact with the 10-transmembrane helical domain, albeit by different structural elements: The A domain is connected via three loops (A-M1, M2-A, and A-M3) that are thought to relay the movements of the A domain to the M1 to M3 helices. The P domain, however, is structurally very tightly integrated into the cytosolic ends of M4 and M5. M4 is linked to the so-called P1 helix (Pro337-Cys344), a short α -helix that runs roughly parallel to the membrane surface, at the membrane-facing side of the P domain. In such a location, P1 may be a key element of inter-domain communication: it connects directly to a β -strand in the P domain ending with the phosphorylated aspartate residue, and it makes contact with the cytosolic end of M3 and to the loop between M6 and M7 (L6-7), which has been shown to play an important role in SERCA catalysis (11–17) (Fig. 1A).

To understand the Ca²⁺ transport mechanism it is mandatory to obtain information on the structural relations of residues critical in mediating the communication between the membranous Ca²⁺-binding sites and the cytosolic phosphorylation site. Glu340 is a centrally positioned residue in the P1 helix at the interface between the phosphorylation domain and the cytosolic

ends of M3 through M7 (Fig. 1A). It is almost universally conserved throughout the large P-type ATPase superfamily. With the exception of polyamine transporting pumps and some bacterial representatives that mostly have a glutamine or asparagine residue at this position, all animal P-type ATPases known to pump ions or lipids strictly possess glutamate (SI Appendix, Table S1) (18).

Mutation of Glu340 seems to affect the partial reactions involving cytoplasmic Ca²⁺ interactions (17), thus raising the question of whether this residue plays an important role in linking the P and the TM domains. Structural information supporting such a role has, however, been lacking.

In general, structural information on SERCA mutants is relatively scarce with only four published crystal structures to date (19–21). The main reason for this shortage is that purification and crystallization of recombinantly produced SERCA is challenging, with low protein yields and poor stability in the absence of native lipids compared to the native enzyme purified from skeletal muscle.

In this study, we have determined the crystal structure of yeast-expressed rabbit SERCA1a mutant E340A at 3.2-Å

resolution in the $\text{Ca}_2\text{E1}$ form with bound ATP analog (AMPPCP), and we examined its functional properties, including ATPase activity and Ca^{2+} -binding kinetics. This is notably a determination of the structure of a SERCA mutant being catalytically active, i.e., capable of completing a full catalytic cycle, albeit with altered kinetics. Furthermore, molecular dynamics (MD) simulations of both wild-type (WT) and E340A structures embedded in a lipid bilayer supported our conclusions derived from structural and functional studies. Our data link the structure to Glu340's functional importance and provide insight into the interdomain communication that guides Ca^{2+} transport by SERCA.

Results

SERCA E340A Has a Lower ATPase Activity and Slower Ca^{2+} -Binding Kinetics than WT. The effect of the E340A mutation on yeast-expressed SERCA function was addressed by using an enzyme-coupled assay to estimate the specific activity of the E340A mutant in detergent. We observed that the ATPase activity of the E340A mutant is about 25% of the WT (0.86 vs. 3.45 $\mu\text{moles ATP/mg/min}$, $n = 9$) (Fig. 1C and *SI Appendix, Fig. S1*), i.e., a marked slowing of the enzyme cycle similar to, although not as extensive, as that reported in prior studies carried out in microsomal membranes (11 to 13%) (14, 17). The presence of a large excess of C_{12}E_8 in our preparation, a mild detergent that has been described to accelerate specifically the $\text{E2} \rightarrow \text{Ca}_2\text{E1}$ transition, provides a likely explanation for the less drastic effect of the E340A mutation observed here (22–24).

In order to assess whether this overall deceleration is due to a reduced rate of Ca^{2+} binding to SERCA E340A, we measured the time course of the Ca^{2+} -binding transition, by recording changes in intrinsic tryptophan fluorescence upon addition of Ca^{2+} to the E2 state of the protein. The E2 to $\text{Ca}_2\text{E1}$ transition of E340A is significantly slower than that of the WT, with an increase of $t_{1/2}$ of about threefold (Fig. 1D–F). This suggests that the effect of the E340A mutation on the overall turnover of SERCA may be caused by a delay in the conformational change associated with Ca^{2+} binding, i.e., the $\text{H}_\text{P}\text{E2} \rightarrow \text{E1} \rightarrow \text{Ca}_2\text{E1}$ transition (see further discussion below).

The Mutation E340A Causes Global Changes in SERCA's Domain Arrangement. Our crystal structure of SERCA E340A at 3.2-Å resolution (Fig. 1A and *SI Appendix, Table S2*) allows for a detailed structural comparison with the WT structure. When superposed, the WT and E340A structures deviate by a substantial rmsd of 2.3 Å over all main chain atoms. Nevertheless, when superposed on the C-terminal transmembrane helix bundle M5 to M10 (residues 750 to 994, main chain rmsd = 0.49 Å), the E340A mutant and WT have an almost identical arrangement of the remaining transmembrane helices M1 to M4 (main chain rmsd = 0.56 Å), with a small 1.6-Å bend at the cytosolic end of M3 in E340A (Fig. 2A), and a slightly more kinked helix M1 (105° in E340A vs. 109° in WT). In line with this, the Ca^{2+} -binding sites look identical in WT and E340A.

In contrast to the almost identical arrangement in the TM domain, there is a pronounced shift (approximately 10 Å) in the position of the cytosolic headpiece, which is rotated further “downward” toward the membrane surface in E340A (Fig. 2A and B and *Movie S1*). At the membrane-facing side of the P domain, the distance between Pro337 at the distal end of the P1 helix and Pro312 in M4 (the nearest residue to Pro337 at the approximate level of the membrane surface) decreases from 15.6 Å in WT to 12.3 Å in E340A. The pivot points of this headpiece movement are at the cytosolic ends of M2, M4, and M5. Whereas M2 and M4 stay very rigid throughout their entire lengths, M5 undergoes a slight bend in the region between its N-terminal end (at Phe740), which is embedded in the P domain, up till residue Gly750, which sits at the “rear side” of P1, exactly opposite

residue 340 (Fig. 2B). On the “front side” of P1, the N-terminal end of M3 moves in the opposite direction of the shifting P1 helix, bending slightly inwards with its first two N-terminal windings (Fig. 2A, *Inset* and *Movie S1*).

While the domain bodies themselves do not change between WT and E340A (main chain rmsd of 0.37 Å, 0.37 Å, and 0.32 Å for the single P, N, and A domains, respectively), their concerted movement relative to the rest of the structure in E340A is reflected by an overall main chain rmsd for the entire headpiece of 1.16 Å. The P and N domains appear to move together as one rigid body (main chain rmsd = 0.47 Å for the P/N body), whereas the A domain is displaced slightly further, with the distance between the N and the A domain increasing by approximately 3 Å in the final position.

Intriguingly, the downward rotation of the cytosolic headpiece in E340A is an extension of exactly the same trajectory as the movement of WT SERCA when it shifts from empty MgE1 to the Ca^{2+} -bound E1 form ($\text{Ca}_2\text{E1}$): when Ca^{2+} binds to WT SERCA, the entire P1 to M3 network moves upwards relative to L6-7 and M5, by about 4.5 Å, while the overall headpiece rotates down toward the membrane surface (8). In E340A, this headpiece downward rotation continues further by about 10 Å (as visualized in a superposed morph between the MgE1 [Protein Data Bank {PDB} 4H1W] and the $\text{Ca}_2\text{E1}$ WT [PDB 3N8G], and E340A [PDB 6RB2] structures; see *Movie S2*). This extension of the WT Ca^{2+} -binding trajectory is reflected in a larger main chain rmsd between the empty WT structure and the Ca^{2+} -bound states in E340A (5.6 Å) than in WT SERCA (4.7 Å).

The observation of an “overshooting” Ca^{2+} -binding movement in E340A led us to ask whether this mutation might lead to a stabilization of a Ca^{2+} occluded state, the state that follows Ca^{2+} (and nucleotide) binding in the catalytic cycle. Ca^{2+} occlusion is normally stabilized by the initiation of phosphorylation by ATP and not by Ca^{2+} binding alone (7). We hence looked at further structural deviations of SERCA E340A from WT, starting at the contacts between the three domains of the cytosolic headpiece (N, P, and A domains).

Connectivity between the N and A Domains Is Looser in E340A. While the P/N body moves as one rigid body as described above, the distance between the N and A domains increases in E340A. The main chain rmsd of 1.4 Å for the N/A body is the largest rms deviation measured for any element of the two structures, indicating that this is the largest structural change induced by the mutation. This change leads to the loss of six of eight polar contacts between the N and the A domains (Arg134 to Phe427, Arg139 to Asp426 and Glu435, Lys169 to Thr484, Thr171 to Leu578, and Lys218 to Asp422), leaving only two contacts intact between these two domains (Arg134 to Asp426 and Thr171 to Glu486), whereas the A domain's contacts to the P domain remain intact (Gly156 to Ala725 and Val726, Asn213 to Thr729) (Fig. 2C and D). This loss of N-A contacts is also evident from our MD simulations (see below and Fig. 4C).

Local Changes at the Mutated Residue 340. In the $\text{Ca}_2\text{E1}$ form of WT SERCA (PDB 3N8G), the Glu340 side chain is in hydrogen bonding distance to the side chain OH group of Thr247 and the main chain nitrogen of Leu249 at the N-terminal end of M3, with additional electrostatic attraction by the partial positive charge of the helix N terminus. A third interaction of Glu340 is a water-mediated contact to Arg822 in L6-7 (Fig. 3A).

Interestingly, the local conformation around the Glu \rightarrow Ala mutation (within a radius of ~ 10 Å) is very similar to WT, with measurable shifts of under 2 Å (Figs. 2A and 3A and B). While the local similarity may seem surprising at first glance, looking at the larger global structural changes described above, it becomes clear that residue A340 is located very close to the pivot point for

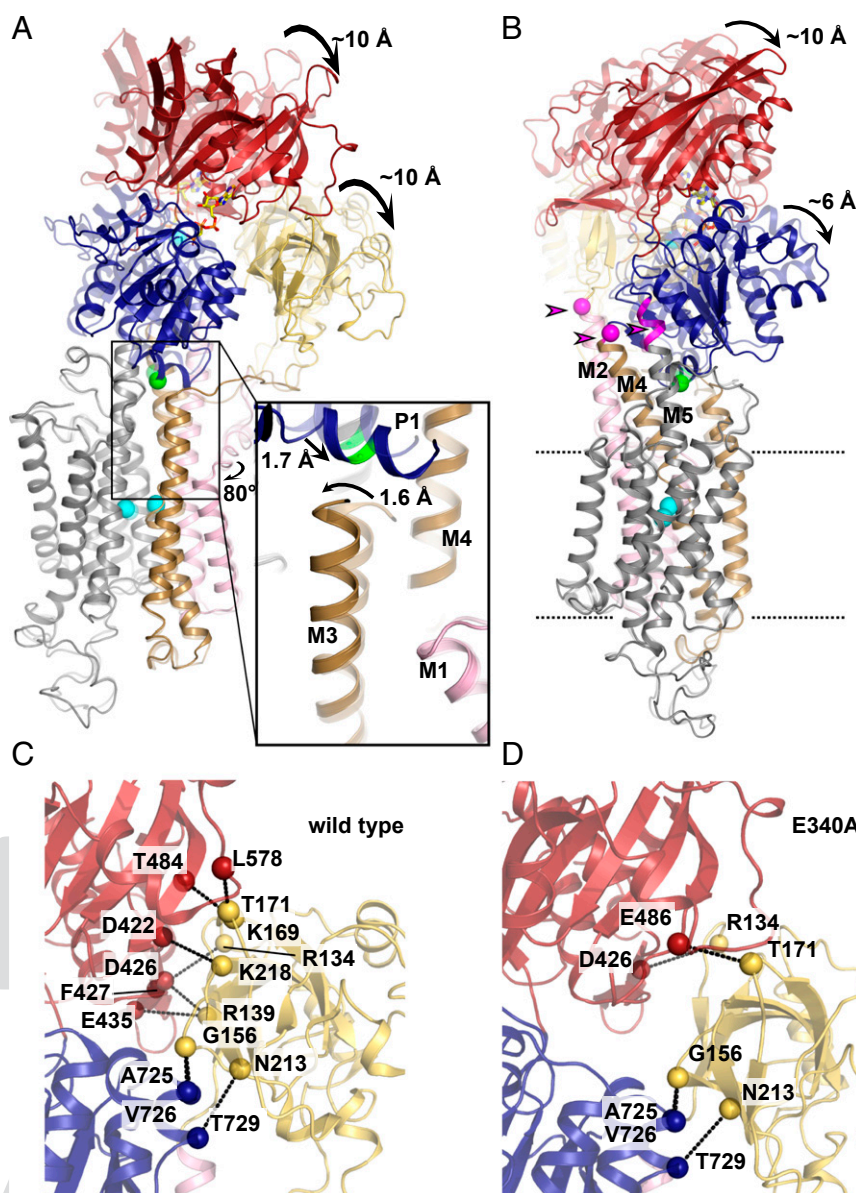


Fig. 2. Comparison of SERCA WT (PDB 3N8G) and E340A crystal structures. Coloring is as in Fig. 1. (A) Domain rearrangements in E340A. Superposition of E340A with WT (transparent). (Inset) Local changes near residue 340. (B) Side view of A, with the pivot regions at the cytosolic ends of M2, M4, and M5 highlighted in magenta and by arrowheads. (C and D) Polar contacts between the A domain (yellow) and the N (red) and P (blue) domains in WT SERCA (B) and E340A (C).

the described headpiece rotation and therefore does not move much itself.

The overall distance between P1 and M3 is slightly smaller than in WT, and the nature of the contact between these two structural elements is very different in E340A: there are no apparent polar interactions between P1 and M3, except for a distant (4.4 Å) contact between Thr341 in P1 and Thr247 in M3 (Fig. 3B). Instead, the electrostatic attraction to the partially positive M3 N terminus is replaced with hydrophobic interactions (Ala340 to Leu249, Pro337 to the methyl group of Thr247, and Thr341 to Pro248). In fact, this interaction appears to be rather tight, as shown by our MD simulations (see below and Fig. 4B).

Removal of the Glu340 side chain abolishes the contact to Arg822, and the entire L6-7 loop (Phe809 to Ser830) is displaced laterally by ~3 Å, increasing the distance between Arg822 and

the P1 helix. In order to test the hypothesis that the loss of the Arg822-Glu340 contact in E340A contributes to the functional impairment of this mutation, and to validate the structural and functional differences between WT and E340A described above, we conducted MD simulations.

Molecular Dynamics Simulations Support a More Occluded State of E340A.

In order to assess whether the differences we found between the crystal structures of WT SERCA and E340A might reflect stable conformations occurring in the native membrane environment, and to find further explanations for the observed functional impairment of the mutant, we carried out 3×200 ns of all-atom MD simulations of each structure reconstituted in a POPC lipid bilayer. As an internal control, we generated an $q:14$ in silico mutant E340A_{is} by truncation of the Glu340 side chain

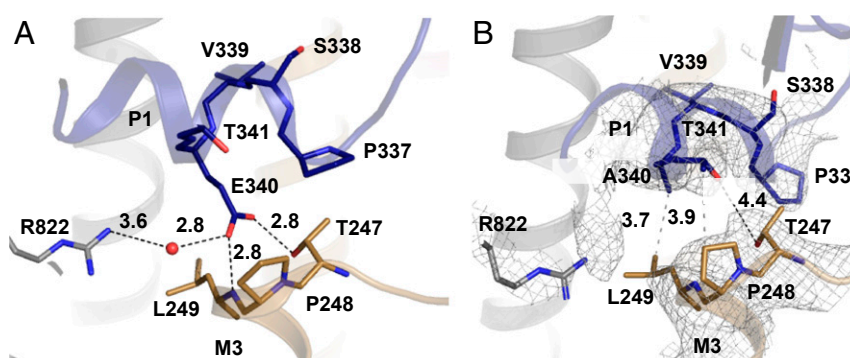


Fig. 3. Local structural changes near the E340A mutation. (A) Local hydrogen bond network (black dashes) around residue 340 in WT SERCA. (B) Local hydrogen bond network (black dashes) and hydrophobic contacts (gray dashes) around residue 340 in E340A. Gray mesh: 2mFo-DFc electron density map contoured at 1.0 σ .

from the WT Ca₂E1-AMPPCP structure and subjected it to the same simulations.

To quantify the conformational downward shift of the cytosolic headpiece relative to the membrane, we determined the angle between three points; Leu13 in the A domain and Thr86

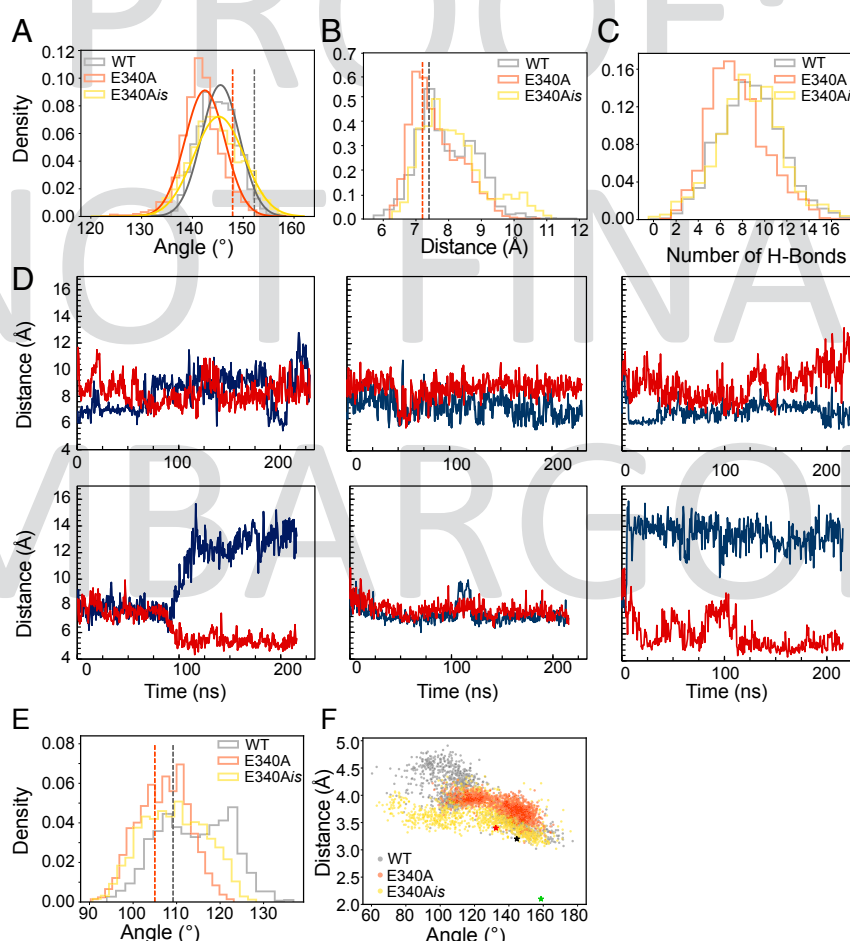


Fig. 4. Molecular dynamics simulations of SERCA WT (PDB 3N8G), E340A and modeled mutant E340A_{is}. WT is shown in gray, E340A in orange, and E340A_{is} in yellow. Histograms give density as an arbitrary unit on the y axis. Dotted lines and asterisks refer to reference values in respective crystal structures. (A) Angle histogram between the soluble headpiece and the transmembrane domain, measured between Leu13 in the A domain and Thr86 and Leu98 in M2. (B) Distance histogram between C α atoms of Leu249 at the tip of M3 and residue 340. (C) Histogram of hydrogen bonds between the A-domain and the N/P-domain body. (D) Distance traces between Arg822-N ϵ and Glu/Ala340-C α (blue) and between Leu249-C α and Pro824-C α (red). Left to Right: Simulation runs 1 to 3. Top traces, WT; Bottom traces E340A. (E) Histogram of kink angles in M1, measured between C α atoms of Trp50, Arg63, and Val74. (F) Geometry at the catalytic site. Angles between the O- and P γ -atoms of the terminal phosphoanhydride bond of ATP and the closest carboxyl oxygen (O δ) of Asp351 versus the distance between O δ and P γ . Green asterisk: 1T5T, ADP-AlFx complex.

and Leu98 in M2 (*SI Appendix, Fig. S2A*). When comparing the two crystal structures, the angle in E340A is 5° more acute (147°) than in WT (152°). During the course of the simulations, both structures relax toward more acute angles, but the difference between them stays about the same, with E340A having a median angle of 142° and WT of 146° (Fig. 4A). This confirms that the difference in headpiece angle is a true feature of E340A that persists in a dynamic membrane environment.

Coherent with the above finding, the P1 helix remains closer to the tip of M3 in E340A, as shown by the distance between the C α atoms of residue 340 and Leu249 at the tip of M3 (Fig. 4B and *SI Appendix, Fig. S2B*). The WT structure exhibits a bimodal distribution of distances with peaks at 7.5 Å and 8.8 Å, whereas E340A has a prominent peak at a shorter distance than WT (6.9 Å).

The loss of connectivity between the A domain and the N/P body described above is also apparent from the MD simulations. The number of H bonds between these elements remains reduced in E340A throughout the simulation (Fig. 4C).

In the three analyses described above, the *in silico* mutant E340A_{is} gives similar distributions to WT, likely because the rather large domain movements are too slow to be seen after the short simulation time of 200 ns. When looking at more local dynamics, however, E340A_{is} does behave more similarly to E340A, as seen in the analyses below.

The crystal structure suggests that Glu340 can engage in up to three hydrogen bonds: with Thr247, the backbone of Leu249, and a water-mediated contact to Arg822 (Fig. 3A). The link to Arg822 in L6-7 is completely lost in E340A (Fig. 3B). Despite the loss of this ionic contact, there is only a slight displacement of L6-7, which is held in place by contacts to L8-9 and other regions of the P domain. Accordingly, we could not detect an increased motility of L6-7 in our MD simulations (*SI Appendix, Fig. S3A*). Nevertheless, our simulation data show that—in the absence of the Glu340 side chain—Arg822 can swing out of the site between P1 and M3 (*SI Appendix, Fig. S3B*). This is measurable as a sudden increase in the distance between Arg822 N ϵ and Ala340 C β from about 7 Å to 13 Å in E340A and to more than 14 Å in E340A_{is} (Fig. 4D and *SI Appendix, Fig. S3C*). In WT SERCA, Arg822 does not swing out spontaneously, and the distance stays at 7 to 10 Å (Fig. 4D). Importantly, once moved to the “out” position, Arg822 does not return to the “in” position over the course of any of our simulations of E340A and E340A_{is} (Fig. 4D and *SI Appendix, Fig. S3C*). The outward swing of Arg822 does, however, not lead to a loss of contact between P1, M3, and L6-7, as might be expected at first glance. On the contrary, concomitant with the outward swing of Arg822 from the wedge between P1 and M3, the space is taken up by Leu249 at the tip of M3, which closes in on the gap and approaches Pro824 in the L6-7 loop (Fig. 4B and *SI Appendix, Fig. S3D*), reducing the distance between the two C α atoms from 10 Å (E340A_{is}) or 7.5 Å (E340A) to 5 Å (Fig. 4D and *SI Appendix, Fig. S3C*). This approach of M3 toward L6-7 is mutually exclusive with the in position of Arg822, as seen by a conspicuous correlation of the distance plots between Arg822 and Ala340 and between Leu249 and Pro824. This correlation is also evident in the *in silico* mutant E340A_{is} but absent from the WT (Fig. 4D and *SI Appendix, Fig. S3C*), suggesting that the changes in the contact network around P1 play a central role in mediating the observed structural changes in E340A.

Comparing the dynamics of the transmembrane segments throughout the simulations, we found a remarkable difference in the kinked region of M1 between WT and both E340A and E340A_{is}: M1 is rigidly kinked in E340A, while in WT the kink region displays more dynamics and can straighten to some extent (Fig. 4E and *SI Appendix, Fig. S4*). This finding is particularly interesting in light of prior studies showing an involvement of M1 with Ca²⁺ binding and occlusion (7, 8).

In order to further probe our simulation data with respect to Ca²⁺ occlusion in E340A, we analyzed the nucleotide-binding site. Ca²⁺ occlusion is coupled to phosphorylation of Asp351, and a more occluded conformation of SERCA would be expected to be more prone to being phosphorylated. We hence compared the distances between the ATP γ -phosphate and the respective closer of the two carboxyl oxygen atoms of Asp351, along with the angles between the carboxyl oxygen and the O-P bond of the terminal phosphoanhydride in ATP (*SI Appendix, Fig. S5*). For comparison, the crystal structure in which the fully occluded transition state of phosphorylation has been trapped with ADP-AlF_x (7) the distance and angle between the Asp351 O₈ and the aluminum atom are 2.1 Å and 159°, respectively. As seen in Fig. 4F, the WT Ca₂E1-AMPPCP structure displays a large spread of angles (between 60° and 180°) and distances (between 3 and 5 Å), many of which are not compatible with an in-line associative hydrolysis mechanism. In E340A, the distances take up a much narrower spread mostly between 3.5 and 4 Å, and the angles are more restricted toward the catalytically competent obtuse region (between 110° and 170°). E340A_{is} displays an intermediate distribution. Our data are therefore consistent with a more occluded state in both the ion-binding and the nucleotide-binding regions of E340A.

Discussion

We have addressed the molecular basis for intramolecular communication in SERCA by analyzing the structure and functional properties of the SERCA mutant E340A. The Glu340 residue is strongly conserved among P-type ATPase membrane transporters and is located at a central position of the ATPase, at the interface between the phosphorylation domain and transmembrane segment, where the Ca²⁺-binding sites are located.

The structure of the mutant differs from the equivalent wild-type structure in: 1) a lowered cytosolic headpiece relative to the membrane, 2) altered connectivity between the nucleotide-binding and the actuator domains, and 3) a shift in the Ca²⁺-gating transmembrane helix M3. MD confirms the structural changes observed and moreover reveals altered dynamics in regions involved in Ca²⁺ occlusion, in particular a rigidly kinked M1. These structural findings point to a stabilization of a more occluded conformation and are discussed below in relation to the functional properties of the E340A mutant.

The downward movement of the cytosolic headpiece toward the membrane in E340A is an extension of exactly the same trajectory as the movement of WT SERCA when it shifts from empty MgE1 to the Ca²⁺-bound E1 form (Ca₂E1). Hence, the loss of Glu340 allows an overshoot of this closure movement after ion binding, which is pivoting around P1. At the same time, within the headpiece, the A domain gets displaced and loses more than half of its contacts with the N domain. The exact functional implication of this effect is not entirely clear. Crystal structures, X-ray solution scattering, and our MD simulations show that the A domain undergoes the largest and most diverse movements during the catalytic cycle (9, 10, 20, 25), and through its direct linkage to TM helices M1 to M3 it directly affects the geometry of the Ca²⁺-binding sites. On the other hand, its position relative to the N and P domains dictates whether the site of ATP hydrolysis can adopt a catalytically competent conformation or not. Hence, any alteration of the dynamics of the A domain will coactively alter SERCA catalysis.

The striking similarity of the Ca²⁺-binding sites and their immediate surroundings in the WT and E340A crystal structures indicates that the differences in the Ca²⁺ binding and dissociation behavior are not caused by a perturbation of the geometry at the Ca²⁺ sites in the TM domain but rather by altered kinetics of the actual binding process. One of the few structural differences in the transmembrane region is a small inward shift of the tip of M3 in E340A. Interestingly, a mutation of Leu249 at the tip of

807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826

834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868

M3 to alanine leads to increased rates of both Ca^{2+} binding and dissociation (26). Furthermore, in the SERCA Ca^{2+} -binding site mutant E309Q, which is defective in both Ca^{2+} binding and occlusion, the tip of M3 is outward shifted (20), highlighting the immediate influence of small changes to the spacing between P1 and M3 for the ion-binding dynamics in SERCA.

The loss of the electrostatic interaction between Glu340 and Arg822 in L6-7 appears to have only a small structural effect at first glance. However, prior studies have implicated an important role of L6-7 in SERCA catalysis (16, 17, 26). For example, the D813A/D818A mutation leads to a dramatic loss of Ca^{2+} affinity (16). There are also conspicuous similarities between the behavior of the mutant E340A and some L6-7 mutants, including R822A: a slowing of the Ca^{2+} -binding transition from E2 and of Ca^{2+} dissociation (17). The loop L6-7 has moreover been suggested to be part of an ion access pathway in both SERCA and the Na^+/K^+ -ATPase (12, 13, 15, 27), and/or to contribute to the coordination of events between the cytosolic and transmembrane domains (14, 16, 17). A very recent study on SERCA2b has found that small, local conformational changes in the area around the L6-7-P1 interaction—in this case induced by long-range effects from changes to the luminal region of the M domain—are accompanied by changes to the entire headpiece arrangement (28). In light of our finding that the E340A mutation allows for an irreversible swinging out of Arg822, which is correlated to a concomitant inward movement of M3, the reverse conclusion must be that the interaction network between Arg822, Glu340, and Leu249 is critical for maintaining a proper architecture of the Ca^{2+} entry and exit pathways.

Finally the “rigidification” of the M1 kink which we see in our MD simulations is also in line with a shift of E340A toward a slightly more occluded state than WT. The kink at Leu60 in M1 permits the conserved residue Phe57 to engage in a hydrophobic cluster that shields off the Ca^{2+} -binding sites as a prerequisite for ion occlusion. Notably, a straight M1 helix has been seen in nucleotide-free inward-open E1 structures (6, 20), and is associated with an unoccluded $\text{Ca}_2\text{E1}$ state of SERCA. Most notably, when probing the reported interrelation of ion occlusion and phosphorylation (7), we find the geometry of the phosphorylation site to be more narrowly clustered around catalytically competent values.

How Does the Structural Stabilization of the Occluded State Manifest Itself in Functional Terms? Our tryptophan fluorescence data obtained with the detergent solubilized yeast enzyme used for crystallization show that the E340A mutation caused a pronounced slowing of the reaction sequence $\text{H}_n\text{E2} \rightarrow \text{E1} \rightarrow \text{Ca}_2\text{E1}$ consisting of proton release, Ca^{2+} binding, and the associated conformational changes. This effect explains the reduced ATPase activity of the E340A mutant relative to WT, and the slowing agrees with the change in the dynamics of Ca^{2+} gating suggested from the structural changes in relation to the Ca^{2+} -binding sites. It should be noted that yeast-expressed SERCA has been demonstrated to be as active and stable as SERCA extracted from the SR, both in membranes and in detergent, when measured under the appropriate experimental conditions, making the comparison straightforward and reliable (29–31).

The E340A mutant has previously been analyzed functionally following expression in COS cells (17), showing effects qualitatively similar to the present results. In accordance with our finding of a more occluded E340A structure, the previous functional analysis in COS cell microsomes (17) showed that the E340A mutation not only slows down Ca^{2+} binding, but also the reverse reaction $\text{Ca}_2\text{E1} \rightarrow \text{E1}$ (Ca^{2+} dissociation from the high-affinity sites toward the cytosol). The mutation was moreover found not to affect SERCA’s apparent vanadate affinity (which is typically affected by E2 to E1 shifts), suggesting that the slowing of $\text{H}_n\text{E2} \rightarrow \text{E1} \rightarrow \text{Ca}_2\text{E1}$ is due to inhibition of the latter

part of this reaction sequence (i.e., actual Ca^{2+} binding). The mutation also had no significant effect on the remaining partial reaction steps of the cycle including the phosphorylation ($\text{Ca}_2\text{E1} \rightarrow \text{Ca}_2\text{E1P}$), the rate-limiting $\text{Ca}_2\text{E1P} \rightarrow \text{E2P}$ transition, and the dephosphorylation $\text{E2P} \rightarrow \text{E2}$ (17). Hence, the kinetic constraints observed for the E340A mutant are consistent with the hypothesis that perturbation of the hydrogen-bonding network around Glu340 causes a slowing, in both directions, of the structural changes associated with the binding of the second Ca^{2+} ion to the CaE1 state and the subsequent occlusion step.

In conclusion, our data suggest that the E340A mutation stabilizes a conformational state closer to occlusion and phosphorylation than the WT structure. This slows down Ca^{2+} entry and release but favors phosphorylation by ATP. Hence, the role of Glu340 in WT SERCA is to maintain the necessary structural flexibility for rapid Ca^{2+} exchange at the binding sites and to relay this flexibility to the site of phosphorylation.

This key function provides an explanation for the evolutionary acquisition and strict conservation of this glutamate residue throughout the entire P-type ATPase family, irrespective of their substrate specificity.

Materials and Methods

Chemicals. Octaethylene glycol monododecyl ether (C_{12}E_8) was purchased from Nikkol Chemical (BL-8SY), and n-dodecyl β -D-maltopyranoside (DDM) was from Anatrace (D310). Streptavidin sepharose high-performance resin was provided by GE Healthcare (17-5113-01). Thapsigargin (TG stock solution was prepared at 1 mg/mL in DMSO, i.e., about 1.5 mM) was from VWR International (586005). All other chemical products were purchased from Sigma. Sequence for SERCA1a heterologous expression was from rabbit.

Cloning, Expression, and Purification. SERCA1a E340A cDNA (SERCA1a[E340A]) was recovered from pMT2 vector initially used for expression in COS cells (17) and cloned into the yeast expression plasmid pYeDP60, with a C-terminal biotin acceptor domain (32), resulting in pYeDP60_SERCA1a[E340A]-Bad. The construct was checked by sequencing (Eurofins MWG). The *Saccharomyces cerevisiae* yeast strain W303.1b/Gal4 (α , leu2, his3, trp1::TRP1-GAL10-GAL4, ura3, ade2-1, canr, cir) was the same as previously described (32). Transformation was performed according to the lithium acetate/single-stranded carrier DNA/PEG method (33).

Expression and purification of the E340A mutant SERCA1a was done as previously described for the WT enzyme (32). Briefly, after overexpression in the yeast *S. cerevisiae*, light membranes were prepared and solubilized by DDM for subsequent purification by streptavidin affinity chromatography (30, 32, 34). Purified WT SERCA1a was recovered in a buffer containing 50 mM Mops-Tris pH 7, 100 mM KCl, 5 mM MgCl_2 , 2.1 mM CaCl_2 , 40% glycerol (vol/vol), and 0.5 mg/mL DDM, together with some thrombin remaining from the elution procedure. The protein concentration in the purified fraction was typically in the 0.05- to 0.15-mg/mL range, depending on the batch. Further purification for crystallization trials was carried out as described previously for the SERCA E309Q mutant (20), involving exchange of DDM with C_{12}E_8 and relipidation of the purified SERCA with DOPC.

ATPase Activity Measurements. Steady-state ATPase activity measurements were performed using an enzyme-coupled assay by measuring the rate of NADH oxidation (followed by absorbance at 340 nm) in the presence of 0.02 mg/mL lactate dehydrogenase, 0.04 mg/mL pyruvate kinase, 1 mM phosphoenolpyruvate, 0.3 to 0.4 mM NADH, 5 mM MgATP (31) and in the additional presence of 1 mg/mL C_{12}E_8 and 50 μM free Ca^{2+} to limit time-dependent inactivation of the SERCA1a (22), at 30 °C and pH 7.5.

Tryptophan Fluorescence. Intrinsic fluorescence was measured with a Fluorolog spectrofluorimeter (Horiba), in a temperature-regulated and continuously stirred 2-mL quartz cuvette. Excitation and emission wavelengths were set at 295 and 320 nm, with bandwidth of 2 and 5 nm, respectively. Integration time for recording of the signal was 2 s. SERCA1a intrinsic fluorescence changes were measured with purified WT or purified E340A mutant suspended at a protein concentration of about 10 $\mu\text{g/mL}$ in a buffer containing 50 mM Mes-Tris pH 6.5, 5 mM MgCl_2 , 20% glycerol (vol/vol) and 2 mg/mL DDM, at 20 °C. Initial Ca^{2+} concentration was adjusted to 105 μM on top of the contaminating Ca^{2+} (3 to 5 μM) already present in the buffer. This was followed by the addition of 5 mM EGTA (EG), reducing $[\text{Ca}^{2+}]_{\text{free}}$ to

about 100 nM. An extra addition of 12.5 mM CaCl_2 (Ca) allowed to reach a final $[\text{Ca}^{2+}]_{\text{free}}$ of about 7.6 mM to recover the initial level of fluorescence. Data were fitted with the GraphPad Prism program with an exponential two-phase association law on the first 60 s of the raw data following Ca^{2+} addition.

Crystallization. Purified SERCA E340A at a concentration of 8 to 10 mg/mL was supplemented with Ca^{2+} , Mg^{2+} , and AMPPCP to 10, 3, and 1 mM, respectively, and crystallization trials were carried out using the vapor diffusion technique. The crystallization drops were prepared as 1- μL sitting drops, with each drop containing a 1:1 mixture of the concentrated, DOPC-relipidated, and C_{12}E_8 -solubilized Ca^{2+} -ATPase solution and the reservoir solution. Diffraction-quality crystals were obtained with reservoir solutions containing 200 mM lithium acetate, 18 to 22% PEG6K, 6 to 9% glycerol, and 3% tert-butanol. The E340A crystals varied markedly in appearance from crystals obtained with WT under similar conditions. Hence, the E340A crystals displayed a rectangular stick morphology, whereas WT crystals were diamond shaped (compare figure S2 in ref. 20).

Data Collection, Processing, and Refinement. Crystals were flash frozen and tested at the microfocus beam line ID23-2 at ESRF. The best dataset diffracted to 3.2-Å resolution. In accordance with the variant crystal forms, also the space group and unit cell parameters were different from WT SERCA. Hence, WT crystallizes in space group C2 with unit cell parameters (163 Å, 76 Å, and 151 Å; 90°, 109°, and 90°), and E340A in space group $\text{P}2_12_12$ with unit cell parameters (126 Å, 232 Å, and 50 Å; 90°, 90°, and 90°), which are hitherto unknown parameters for SERCA. Phasing was done by molecular replacement with the $\text{Ca}_2\text{E1}$ -AMPPCP form (1T55). Initial rigid-body refinement with each domain as a single rigid body was followed by all-atom and TLS refinement in PHENIX (35).

Molecular graphics figures were prepared with PyMOL (36) and morphs with Morphinator (37).

Molecular Dynamics Simulations. Simulations were built using the coordinates from PDB 3N8G (WT) and PDB 6RB2 (E340A). The protein atoms were described using the CHARMM36 force field (38), and built into a lipid bilayer composed of POPC molecules and solvated with TIP3P water and Na^+ and Cl^- to 150 mM in a $120 \times 120 \times 180$ Å box. Systems were built using CHARMM-GUI (39, 40). Generation of the in silico mutant E340A_{is} was done with CHARMM-GUI.

The systems were energy minimized using the steepest descent method, then equilibrated with positional restraints on heavy atoms for 100 ps in the NPT ensemble at 310 K with the V-rescale thermostat and semiisotropic Parrinello-Rahman pressure coupling (41, 42). Production simulations were run in triplicate without positional restraints, with 2-fs time steps for a minimum of 200 ns.

All simulations were run using GROMACS 2019 (43). Simulations were analyzed using GROMACS tools and images were made using VMD (44). Graphs were plotted using Matplotlib (45).

Data Availability. X-ray crystallography data have been deposited in the Protein Data Bank under accession no. 6RB2.

ACKNOWLEDGMENTS. This work was supported by the French Infrastructure for Integrated Structural Biology (ANR-10-INSB-05), and the CNRS, the Agence Nationale de la Recherche and the Domaines d'Intérêt Majeur Maladies Infectieuses Région Ile de France (to M.I.M., C.J., G.L., and C.M.), the Center for Membrane Pumps in Cells and Disease of the Danish National Research Foundation (to J.V.M., P.N., J.P.A., and M.B.), the Lundbeck Foundation (to P.N.), the Wellcome Trust (ref. 220063/Z/20/Z to M.M.G.G.), and the Danish Council for Independent Research (to J.P.A.). We are grateful to Anna Marie Nielsen for technical support and Jesper L. Karlén and Mark Sansom for guidance and support regarding scientific computing as well as access to facilities.

993
994
995 Q:25
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054

35. P. D. Adams *et al.*, PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

36. L. L. C. Schrodinger, The PyMOL Molecular Graphics System (Version 1.3r1, 2010).

37. J. L. Karlsen, M. Bubltz, How to compare, analyze, and morph between crystal structures of different conformations: The P-Type ATPase example. *Methods Mol. Biol.* **1377**, 523–539 (2016).

38. R. B. Best *et al.*, Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).

39. J. Lee *et al.*, CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).

40. S. Jo, T. Kim, W. Im, Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS One* **2**, e880 (2007).

41. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).

42. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).

43. H. J. C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).

44. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38, 27–28 (1996).

45. J. D. Hunter, Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 99–104 (2007).

PROOF:
NOT FINAL
EMBARGOED

1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074

1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

1

- Q: 1_Please review 1) the author affiliation and footnote symbols, 2) the order of the author names, and 3) the spelling of all author names, initials, and affiliations and confirm that they are correct as set.
- Q: 2_Please review the author contribution footnote carefully. Ensure that the information is correct and that the correct author initials are listed. Note that the order of author initials matches the order of the author line per journal style. You may add contributions to the list in the footnote; however, funding may not be an author's only contribution to the work.
- Q: 3_Please note that the spelling of the following author names in the manuscript differs from the spelling provided in the article metadata: Cedric Montigny, Christine Jaxel, Marc le Maire, and Jesper V. Møller. The spelling provided in the manuscript has been retained; please confirm.
- Q: 4_Please review your open access and license selection and confirm that it is correct.
- Q: 5_Certain compound terms are hyphenated when used as adjectives and unhyphenated when used as nouns. This style has been applied consistently throughout where (and if) applicable.
- Q: 6_If you have any changes to your Supporting Information (SI) file(s), please provide revised, ready-to-publish replacement files without annotations.
- Q: 7_Please replace CEA and EMBL in the affiliations with their spelled out names.
- Q: 8_Please check affiliation "e" as provides and make certain the last mentioned location is the main institution. Please also include the missing city, state, postal code, and country.
- Q: 9_Please note that because original ref. 26 was a duplicate of ref. 20 and because original ref. 30 was a duplicate of ref. 6, all original refs. from 26 through 47 have been renumbered in the reference list and within the text. Please check the extensive renumbering as a result of the duplicates.
- Q: 10_Please check "A-M1, M2-A, and A-M3" to be sure they appear as intended. Is "M2-A" intended to be different from "A-M1 and A-M3?"
- Q: 11_Claims of priority or primacy are not allowed, per PNAS policy (<https://www.pnas.org/page/authors/format>); therefore, the term "the first" has been deleted: "This is notably a determination of the structure . . ."
- Q: 12_Please note if these PDB accession nos. are new as a result of this work, please include them in the Data Availability section below.
- Q: 13_Please confirm "Protein Data Bank" as the definition of PDB at first mention.
- Q: 14_Please include the definition of POPC here at first mention.
- Q: 15_Please include the definition of DMSO.
- Q: 16_Please define PEG.
- Q: 17_Please define DOPC.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

2

Q: 18_Please define EGTA.

Q: 19_Please define ESRF.

Q: 20_Please confirm or otherwise revise as needed the change from "variant crystals form" to "variant crystal forms."

Q: 21_Please define TLS.

Q: 22_Please indicate whether the data have been deposited in a publicly accessible database before your page proofs are returned. The data must be deposited BEFORE the paper can be published. Please also confirm that the data will be accessible upon publication.

Q: 23_All data related to this paper that are not within the main text or SI should be cited in the data availability statement, per PNAS policy (<https://www.pnas.org/page/authors/journal-policies#xi>). In addition, all data cited in the text, including newly created data, should be cited in the reference list. For each new reference, please provide the following information: 1) author names, 2) data/page title, 3) database name, 4) a direct URL to the data, 5) the date on which the data were accessed or deposited (not the release date), and 6) where the new reference citation should be added in the main text and/or data availability statement.

Q: 24_Please provide the editor(s) name for ref. 36.

Q: 25_Please provide the publisher name and location for ref. 36.

Q: 26_Please include mention of the ** in the legend for Fig. 1*F*.
