

Mesophilic enzyme function at high temperature: molecular dynamics of hyperthermophilic and mesophilic pyrophosphatases

Rupesh Agarwal^{1,2}, Utsab R. Shrestha^{*1}, Xiang-Qiang Chu⁴, Loukas Petridis^{1,3}, Jeremy C. Smith^{*1,3}

¹UT/ORNL Center for Molecular Biophysics, Biosciences Division, Oak Ridge National Laboratory, Tennessee 37831-6309; ²Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, Tennessee 37996; ³Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996; ⁴Department of Nuclear Science and Technology, Graduate School of China Academy of Engineering Physics, Beijing 100193, China.

* co-corresponding authors

Abstract

The mesophilic inorganic pyrophosphatase from *Escherichia coli* (*EcPPase*) retains function at 353 K, the physiological temperature of hyperthermophilic *Thermococcus thio还原ens*, whereas, the homolog protein from the hyperthermophilic organism (*TtPPase*) cannot function at room temperature. To explain this asymmetric behavior, we examined structural and dynamical properties of the two proteins using molecular dynamics simulations. The global flexibility of *TtPPase* is significantly higher than its mesophilic homolog at all tested temperature/pressure conditions. However, at 353 K, *EcPPase* reduces its solvent-exposed surface area and increases subunit compaction while maintaining flexibility in its catalytic pocket. In contrast, *TtPPase* lacks this adaptability and has increased rigidity and reduced protein:water interactions in its catalytic pocket at room temperature, providing a plausible explanation for its inactivity near room temperature.

Introduction

The enzymatic activity of proteins from hyperthermophilic microorganisms thriving in extreme conditions has been an active area of research for several decades^{1,2}. These microbes have an optimal temperature range for their growth and survival of about 80°C - 100°C³. In addition to extreme temperatures, these microorganisms can also withstand high hydrostatic pressures

ranging from 60 to 100 MPa (the atmospheric pressure at sea-level is 0.1 MPa)^{1,4}. Since most mesophilic proteins denature under such high temperatures⁵ and pressures⁶, it is intriguing to examine how proteins from hyperthermophilic organisms retain activity. Previous studies have suggested structural characteristics⁷⁻⁹ of proteins that enable these extremophilic microbes to thrive in severe conditions. However, there appears to be no universal adaptive mechanism, but rather a complex combination of different factors, which frequently differs according to the protein or protein family and is thus difficult to generalize^{8,10}. Furthermore, the role of dynamic characteristics such as conformational stability^{11,12} and flexibility¹³, for protein adaptability is still not well understood^{1,14}.

Recent studies have reported that the conformational sub-states of a protein are significantly perturbed by changes in temperature and pressure^{13,15-19}. Temperature enhances the internal fluctuations of a protein²⁰, and an optimum temperature may provide an appropriate balance of flexibility and rigidity required for function^{13,21,22}. Temperatures higher than the optimum can lead to loss of function through unfolding or denaturation²³. However in the case of hyperthermophilic proteins, high native-state flexibility can reduce their entropy of unfolding, thus increasing their melting temperature²⁴. Similarly, high pressure conditions can cause a protein to become inactive by the collapse of its intra-protein cavities, giving rise to an unfolded state²⁵. Pressure drives the reduction in the volume of a protein, which results in a negative entropy change of the system, which may destabilize the native state²⁶. Nonetheless, there are exceptions, and the stability and/or activity of some proteins, such as a thermolysin²⁷ and a hydrogenase from *Methanococcus jannaschii*²⁸, have been shown to increase with pressure.

Overall, a dualistic picture of protein flexibility^{24,29} and rigidity^{30,31} has been recognized as a possible factor behind thermostability of thermophilic and hyperthermophilic proteins^{13,19}. On

the one hand, flexibility is required by a protein to function. On the other hand, flexible residues trigger protein unfolding due to their large thermal fluctuations at high temperature. Hence, rigidifying flexible residues may be an effective way to improve thermostability^{30,31}. In addition to flexibility, oligomerization has been reported to be critically important for the stability of some proteins^{32,33} but the structure-based reasoning behind this stability is not understood³⁴.

We compare here a particularly interesting pair of proteins - the inorganic pyrophosphatase (PPase) from *Thermococcus thio-reducens* (*Tt*), a hyperthermophilic archaea found near hydrothermal vents of the Mid-Atlantic Ridge³⁵, with a homolog from the mesophilic bacterium, *Escherichia coli* (*Ec*)³⁶. PPase (EC 3.6.1.1) is a homo-hexameric enzyme (~120 kDa), which catalyzes the conversion of pyrophosphate into two phosphate ions. This conversion is important for many critical biochemical processes, such as the production of proteins, nucleic acid polymerization, and lipid metabolism³⁵. Although *Tt*PPase and *Ec*PPase have ~60% sequence similarity and ~40% identity, and share similar oligomeric crystal structures (**Fig. S2**), the temperatures and pressures for their optimal enzymatic activities are very dissimilar. The catalytic activity of *Tt*PPase has been reported to be maximal at ~353 K but negligible (1.3% of the maximum activity in 10 mins) near room temperature and standard atmospheric pressure (1 bar or 0.1 MPa)³⁷. In contrast, the optimal conditions for *Ec*PPase have been shown to be room temperature (298 K) and standard atmospheric pressure. Interestingly, unlike other mesophilic proteins, *Ec*PPase retains up to 95 to 100% of its enzymatic activity at 353 K for about 10 minutes, after which the activity slowly decays³⁸.

Most crystal structures of thermophilic proteins have been resolved at room temperature instead of at their native temperature or pressure conditions. Therefore, to obtain a comprehensive picture of protein structure, dynamics, and function, static structure determination (using X-ray

and neutron crystallography) must be complemented with dynamic information using various spectroscopic techniques (e.g., neutron scattering^{13,21} or NMR^{39,40}) and computer simulations^{41,42} under varying external conditions. Molecular dynamics (MD) simulation, which provides spatial and temporal information at atomic resolution⁴³ has thus become one of the most powerful methods to explore the protein energy landscapes and their flexibility-function relationships^{18,19,44} at near-native conditions and is the technique applied here.

Here, we perform MD simulations to understand the effects of pressure and temperature on the structural and dynamic behavior of the two PPase homologs in their native and non-native environments. We find that *Tt*PPase possesses higher global flexibility at both native and non-native conditions than its mesophilic homolog (*Ec*PPase). However, this effect is not reflected locally in the catalytic pocket. Additionally, we determined that factors accompanying enzymatic activity of PPase are the number of hydrogen bonds and water molecules in its catalytic pocket. Furthermore, we provide potential factors behind the observed enzyme activity and/or adaptability of *Ec*PPase at high temperature and alternatively the inability of *Tt*PPase to do so at low temperature.

Results

I. Structural flexibility

The structural flexibility of proteins is associated with many biological functions such as catalytic activity, substrate binding, and molecular recognition^{45,46}. To understand the effect of temperature and pressure on the protein global flexibility, we calculated the average mean-square displacement (MSD) of the C α atoms. The MSD is often used to compare the stability of thermophilic and mesophilic protein homologs^{18,19}. Overall, the MSD results show that *Tt*PPase has higher flexibility compared to *Ec*PPase at all the conditions studied here (**Fig. 1**). This is therefore an

example of a thermophilic protein for which increased rigidity is not associated with thermostability.

For both the homologs, a two-fold increase in MSD is observed at 353 K relative to 298 K. Unlike temperature, the effect of pressure was less pronounced, an increase in pressure causing a small reduction in MSD for *Tt*PPase, and no change for *Ec*PPase. Interestingly, the MSD values for both the monomers showed an effect of pressure at 353 K, contrary to the respective hexameric forms (**Fig. 1 and S5**). These results suggest that oligomerization may confer stability to increased pressure at high temperature.

Flexibility is often correlated with activity or stability of mesophilic and thermophilic homologs^{45,46}. However, although at 298 K *Tt*PPase is more flexible than *Ec*PPase, the activity of *Tt*PPase is negligible near room temperature³⁷. Further, even though the flexibility of *Ec*PPase at 298 K is smaller than that at 353 K, *Ec*PPase is catalytically active at both temperatures. Hence, the global flexibility of the PPase does not follow the observed functional properties.

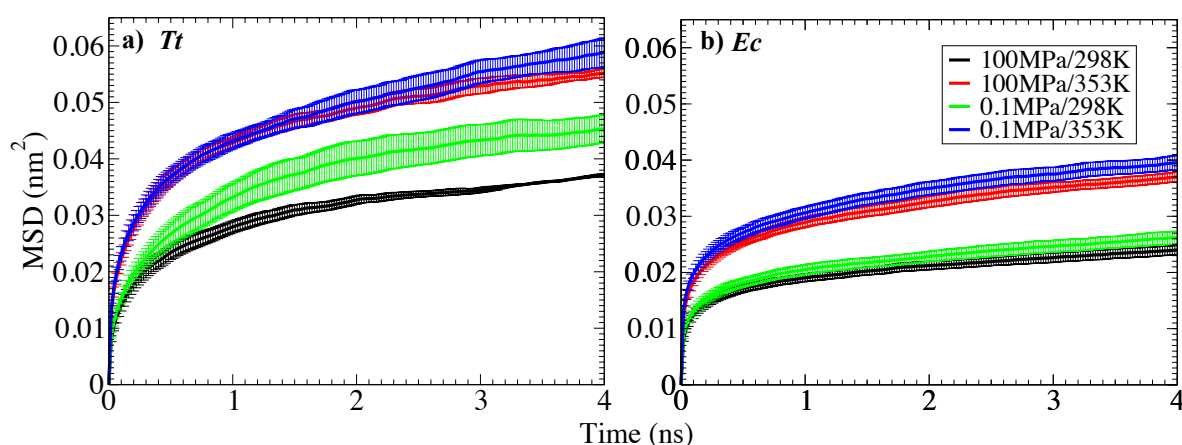


Figure 1: Mean square displacement (MSD) of C-alpha atoms for (a) hexameric *Tt*PPase and (b) hexameric *Ec*PPase. Higher values of MSD for *Tt*PPase at all conditions indicate higher structural flexibility than *Ec*PPase. Error bars shown in this and the subsequent figures are the standard error of the mean, and those not immediately visible are at most the size of the symbol.

II. Compactness

Compactness is another common adaptive mechanism of thermophilic proteins^{34,47}. To characterize the compactness of the hexameric structures over the course of the simulations, we calculated the solvent accessible surface area (SASA) of the proteins, which is inversely proportional to the number of native contacts⁴⁸. The density plots of the SASA show that *Ec*PPase has larger solvent accessibility than *Tt*PPase, irrespective of the temperature and pressure conditions studied here (**Fig. 2a and S1b**). A similar difference is found in the hexameric crystal structures (*Tt*PPase: 374 nm² and *Ec*PPase: 393 nm²; **Table S3**). However, the solvent accessibility of *Ec*PPase is reduced at 353 K compared to 298 K, whereas that of *Tt*PPase remains unaffected by the change in temperature. Pressure did not significantly affect the SASA of either homolog. The SASA probability distribution for monomeric structures showed a similar change in compactness as observed in the hexameric form (**Fig S4a**).

To further quantify compactness, we calculated the number of intra- and inter-chain hydrogen bonds (HBs) of the hexameric form. The normalized mean number of inter-chain HBs was nearly double for *Tt*PPase than for *Ec*PPase and remained constant at all conditions (**Fig. 2b**), confirming the closer packing of the subunits in *Tt*PPase than in *Ec*PPase. The mean number of intra-chain HBs did not change greatly between the two homologs (**Fig. S1a**) at all pressure and temperature conditions. These results suggest that the higher compactness of the *Tt*PPase hexamer is due to compactness between the subunits and also within each monomer.

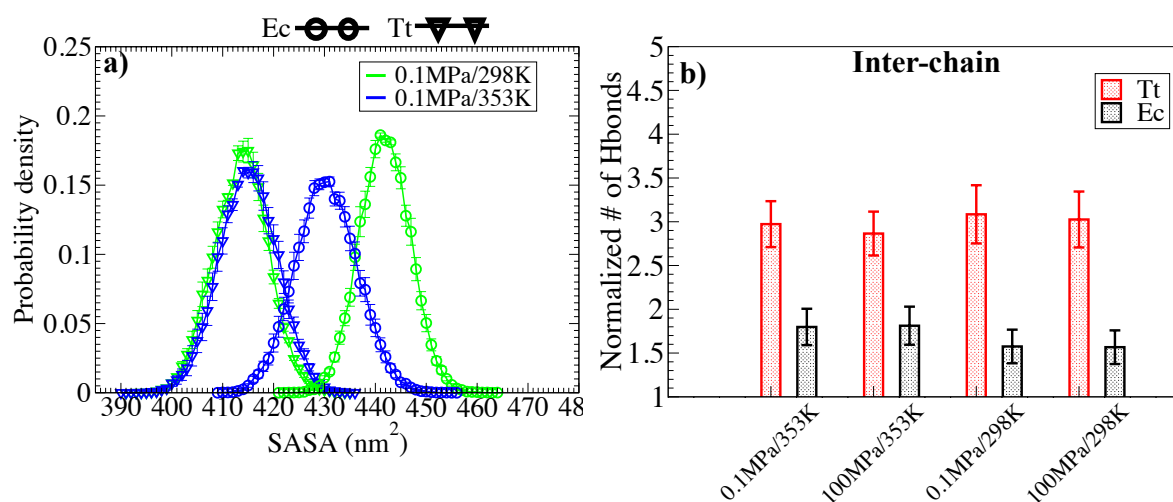


Figure 2: a) Probability distribution for solvent accessible surface area (SASA) of the hexameric forms for *TtPPase* (triangle) and *EcPPase* (circle); b) The normalized number of interchain hydrogen bonds.

III. Protein cavities

Intra-protein cavities have been recognized to be important for the stability and function of proteins^{49,50}, and water molecules buried within these cavities have been reported to be influential in temperature and pressure-mediated unfolding^{25,51}. We calculated the number of water molecules enclosed in the completely buried cavities of both proteins. This was found to be significantly higher for *EcPPase* than *TtPPase* (**Fig. 3**) and not significantly affected by the temperature and pressure. These results are consistent with the inter-chain Hbond results discussed earlier and with calculations that show that the crystal structure of *TtPPase* has smaller buried cavity volume (303 water molecules) than *EcPPase* (427 water molecules) (**Table S3**).

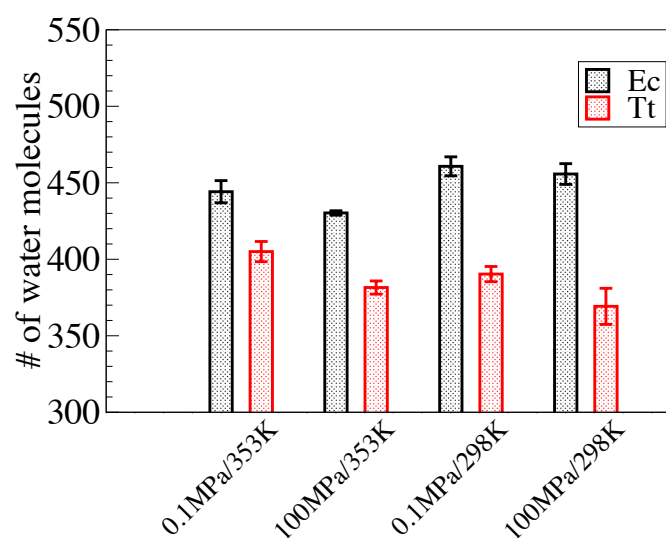


Figure 3: Number of water molecules in completely buried cavities of hexameric form at different temperature and pressure conditions. *Tt*PPase (black shaded bars) vs. *Ec*PPase (red shaded bars).

IV. Flexibility in the catalytic pocket and its interaction with water

The precise 3D arrangement of residues and their interaction with water molecules in an enzyme's catalytic pocket determines the function^{52,53}. Previous studies have identified conserved residues in the catalytic pocket of PPase enzymes (**Table S1 and Figs. S2a and S2b**) that are critical for catalysis and are also required for coordinating the divalent ions^{54,55}. The catalytically inactive *Tt*PPase at 298 K has ~60% more HBs (~15 HBs i.e., 2.5 per monomer) between the catalytic residues than the catalytically-active forms of PPase (*Tt*PPase at 353 K and *Ec*PPase at 298 and 353 K) (**Figs. 4 a, c**). These results, for the hexameric form were also found in the simulations of the monomeric form (**Fig. S4b**). We did not observe any effect of pressure on the number of HBs between the residues in the catalytic pocket (**Figs. 4 a, c**).

Water molecules in the catalytic pockets can be critical for enzymatic function^{56,57,58}. Interestingly, the catalytic pockets of both homologs have ~110 protein:water HBs at 353 K, consistent with the observed catalytic activity of *Ec*PPase at this elevated temperature³⁸. However,

at 298 K, *Ec*PPase significantly increases the number of protein:water HBs, to around ~160 HBs, whereas *Tt*PPase does this to a much lesser degree, to ~130 HBs (**Figs. 4 b, d**).

We also quantified the flexibility and solvent exposure of the catalytic pocket by calculating the root mean square fluctuation (RMSF) of these conserved residues. The RMSF of these residues in *Ec*PPase varies weakly with temperature, whereas *Tt*PPase has significantly lower fluctuations at 298 K compared to 353 K (**Fig. S3**). *Ec*PPase therefore maintains its local flexibility of the catalytic pocket at both temperatures whereas *Tt*PPase becomes rigid at the lower temperature.

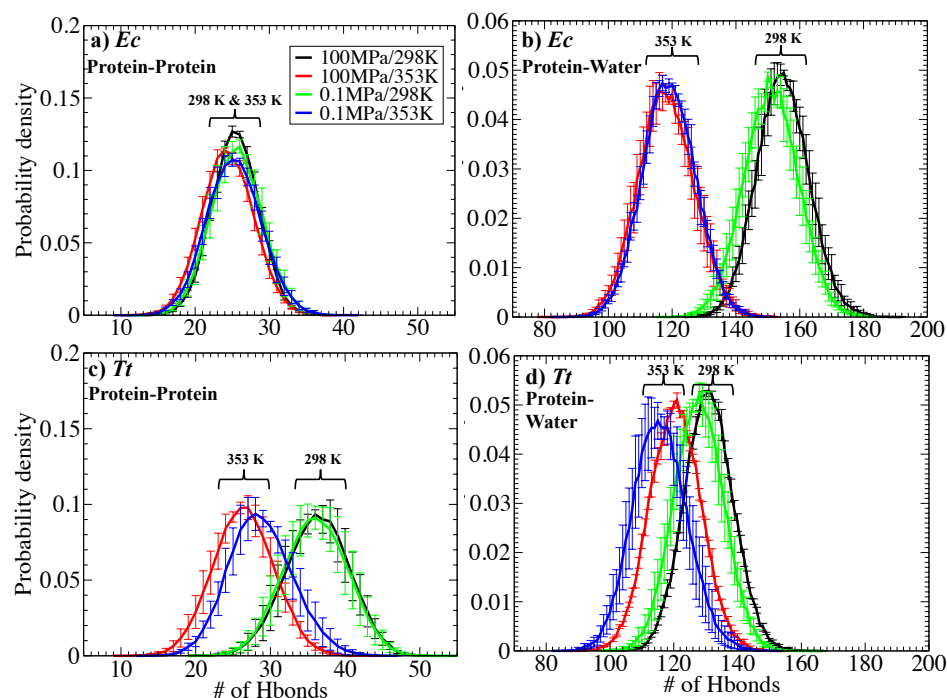


Figure 4: Probability distributions of the number of hydrogen bonds between the residues in the catalytic pocket: hexameric *Ec*PPase (a) and hexameric *Tt*PPase (c). Probability distributions of the number of hydrogen bonds between the catalytic pocket residues and water molecules in the catalytic pocket: hexameric *Ec*PPase (b) and hexameric *Tt*PPase (d).

Discussion and Conclusions

In this work, we compare structural and dynamic properties of hyperthermophilic and mesophilic PPases using MD simulations mimicking deep-sea and ambient conditions. The results indicate that *Tt*PPase has been designed to function at high temperatures with a more compact structure and reduced number of intra-protein hydrogen bonds in the catalytic pocket. However, *Tt*PPase does not maintain both of these properties at room temperature, where it cannot catalyze the enzymatic reaction. Interestingly, we found *Ec*PPase adapts and retains its activity at high temperature by incorporating similar strategies used by *Tt*PPase to function at high temperature: maintaining hydrogen bonds in the catalytic pocket and increasing its compactness.

The overall structural flexibility of a protein is sometimes assumed to be relevant for its enzymatic activity⁴⁵ and the MSD is often used to quantify this. A general framework for understanding the stability and function of hyperthermophilic proteins in their native conditions has been proposed based upon the hypothesis that enhanced rigidity underlies increased thermal stability²⁴. However, other experimental and computational studies have reported that hyperthermophilic proteins have larger conformational flexibility than their mesophilic homologs^{19,59}. Here, we also observe higher overall flexibility of *Tt*PPase than *Ec*PPase at all temperature/pressure conditions. Indeed, even at room temperature, where *Tt*PPase is enzymatically inactive, it is more flexible than *Ec*PPase. Thus, differences in the overall flexibility of PPase are not directly associated with differences its enzymatic activity.

Although the crystal structures of both homologs were not resolved at their native conditions, the RMSD between them is small (0.18 nm) and both homologs have similar radii of gyration (2.9 nm). However, the buried cavity size and solvent accessibility of the crystal structure of *Ec*PPase are greater than for *Tt*PPase. Likewise, from the simulations the solvent-accessible

surface area and number of inter-chain Hbonds (both of which are related to compactness) of *Tt*PPase indicate that it is more compact than *Ec*PPase at all the conditions investigated here. The main contributing factors to the compactness of *Tt*PPase are inter-subunit compaction as shown by increased inter-chain H-bonding, together with the reduced solvent-exposed surface area of each monomer as shown by the simulations of the monomeric forms. These results are not surprising since compactness is a known adaptive factor in common between many thermophilic proteins.

Moreover, a decrease in the number of water molecules in the buried cavities of *Tt*PPase compared to *Ec*PPase was observed at all conditions. These results are consistent with previous work⁶⁰ which showed that proteins that are active at extremely low temperatures (psychrophilic proteins) have a comparatively larger average cavity size. Similarly, a reduction in buried cavity volume has been reported as an adaptation (thermostability) mechanism at high-temperature conditions because large water-filled cavities are known to be a driving factor for protein denaturation at high temperature and pressure⁶¹.

Interestingly, *Ec*PPase exhibited a change in its compactness with a change in temperature, becoming more compact at higher temperature than at its natural temperature. In contrast, the compactness of *Tt*PPase remains the same at room and high temperatures. Additionally, we also studied the structure and dynamics of the catalytic pocket. Here, at 298 K we observed an increase in rigidity at 298 K compared to 353 K, as shown by the lower RMSF of catalytic pocket residues and a greatly increased number of H-Bonds between the catalytic pocket residues. In contrast, we observed similar catalytic pocket residue H-bonds in *Ec*PPase as for *Tt*PPase at 353 K, and this number remained unaltered upon change in temperature. This local flexibility (based upon the number of HBs and RMSF) of the catalytic pocket agrees with the existing “corresponding states”

hypothesis, according to which a thermophilic protein has a more rigid catalytic pocket than its mesophilic homolog at room temperature⁶². However, both of these homologs should exhibit similar flexibility at their respective functional temperatures⁶². Furthermore, we see that the trend observed in global flexibility measured by the MSD is not reflected locally in the catalytic pocket.

The above picture is further supported by calculations of the number of H-bonds of the catalytic pocket residues with water molecules. In the case of *EcPPase*, we observe a dramatic decrease in the number of Hbonds with water at 353 K when compared to 298 K, whereas there is a negligible change for *TtPPase*. This result is consistent with the compactness results showing that *EcPPase* has the capability to adapt to a high temperature whereas *TtPPase* lacks the ability to adapt to low temperature environment.

Based on these results, we suggest that at lower temperatures, both of these homologs would need comparatively more water molecules in their catalytic pockets and a larger exposed surface area to function and that the opposite is needed at higher temperatures. *EcPPase* is able to make the aforementioned changes to its structure and hence is able to adapt to high temperature/pressure but *TtPPase* fails to do so at low temperature. Although the overall structure of *TtPPase* barely responds to change in temperature/pressure, the catalytic pocket becomes more rigid at 293 K than at 353 K due to an increase in the number of intra-protein HBs. This rigidity may explain its enzymatic inactivity at room temperature. In contrast, *EcPPase* adapts to high temperature by reducing solvent-exposed surface and adopting a more compact oligomeric structure. In addition, *EcPPase* preserves the number of hydrogen bonds within the residues in the catalytic pocket at all the conditions and reduces its interactions with water molecules at higher temperature. This provides a possible explanation behind its activity even at a high temperature.

We observed an effect of pressure on the MSD of monomeric form, but not for hexameric form. However, the SASA and the catalytic pocket behavior of the monomeric and hexameric forms are relatively unaffected. Consequently, in the scope of our study, we suggest that the selective pressure behind oligomerization in the case of PPase remains unclear.

Finally, our results provide a structural and dynamic basis for the experimentally observed activity of *Ec*PPase at both high and room temperature. Such an intriguing ability of adaptation of mesophilic PPase could be possibly explained by existing evolutionary theory since hyperthermophilic archaea are thought to be a universal ancestor⁶³. Indeed, recent work⁶⁴ has suggested that thermophilic proteins (isolated from thermophilic archaea *Pyrococcus furiosus*) were “*de novo*” designed in a hot environment and then used a “structure-based mechanism” to adapt to a mesophilic environment on recolonization. In the structure-based mechanism adaptation to high temperatures is through compaction, which here is mostly inter-subunit, with the monomer structures relatively unchanged. Phylogenetic analyses⁶⁵ show that *T. thioreducens* (also hyperthermophilic archaea) shares the same clade as *P. furiosus*; hence, PPase homologs studied here may have evolved according to the structure-based adaptation theory. Thus, *Ec*PPase may have been evolved from its ancestral counterpart *Tt*PPase to recolonize at room temperature and may therefore be able to adapt when introduced to a higher temperature. Moreover, since thermophiles do not live at room temperature, there is no selective need for high activity at room temperature.

Methods

All-atom MD simulations of monomeric and hexameric *E. coli* PPase (*Ec*PPase) and *Thermococcus thioreducens* PPase (*Tt*PPase) were carried using GROMACS (version 2016.3)

suite^{66,67}. Analysis of the MD trajectories was performed using python package pytraj⁶⁸ and GROMACS analysis tools. Hexameric *Ec*PPase (PDB ID 1I6T³⁶) and *Ti*PPase (PDB ID 3Q5V³⁵) were solvated in a TIP3P cubic box of water (~150,000 atoms). The crystal structures for both the proteins were resolved at the same conditions (298 K and 0.1 MPa) and are also in the same space group (H32). The monomeric structures for both the proteins were modeled by taking only one chain from the hexameric structures. These monomeric models were then solvated in a TIP3P cubic box (~36,000 atoms).

Simulation details

All simulations were performed using the AMBER99SB potential⁶⁶. Each protein was simulated at four different conditions: two temperatures: 298 K (ambient for *Ec*PPase) and 353 K (ambient for *Ti*PPase); and two pressures: 0.1 MPa (ambient for *Ec*PPase) and 100 MPa (ambient for *Ti*PPase). All systems were prepared in a three-step process: initial energy minimization, NVT equilibration and a NPT production run. Energy minimization was performed with the steepest-descent algorithm^{66,67} to a tolerance of 1000 kJ/ (mol·nm). NVT equilibration was performed for 1 ns at each temperature fixed with the V-Rescale thermostat⁶⁹. Following this step, the NPT ensemble protocol was followed to generate production runs (500 nanoseconds for the hexamer and 1 microsecond for the monomers) with the V-Rescale thermostat⁶⁹ used to maintain the required temperature and Parrinello-Rahman⁷⁰ to maintain the pressure at 0.1 MPa or 100 MPa. Since the compressibility of water changes drastically with pressure, the compressibility of water was taken to be $4.5 \times 10^{-5} \text{ bar}^{-1}$ at 0.1 MPa and $3.5 \times 10^{-5} \text{ bar}^{-1}$ at 100 MPa, based on previous experimental work⁷¹. Particle Mesh Ewald⁷² was used for long-range electrostatics with a short-range electrostatics and van der Waals cutoff of 1 nm. Three independent runs were performed for

each simulation (**Table S2**). The simulations were considered converged when the fluctuations in the root mean square deviation of C-alpha atoms ($\text{RMSD}_{\text{C}\alpha}$) reached a plateau with time.

Analysis details

Each of the metric was calculated from the last 300 ns of each simulation trajectory. The average value of each metric from 3 independent trajectories starting with different velocities has been reported, where the standard error of the mean has been used as the error bar.

The crystal structure RMSD comparison was performed using the SuperPose webserver⁷³. The number of hydrogen bonds was computed using the “hbond” utility in GROMACS using a donor–acceptor cutoff distance of 0.32 nm and a cutoff angle of 20°. The solvent accessible surface area (SASA) and volume were calculated using a 0.14 nm probe size for the whole protein. The volume of buried cavities was determined using the “trj_cavity” module⁷⁴ in GROMACS. The volume of the completely buried cavities (CBC) was calculated with a 1.4 Å grid spacing (-spacing) and degree of buriedness of 6 Å. The number of water molecules in the cavities was estimated by dividing the total cavity volume by the volume of a buried water molecule near the protein surface (2.29 nm³) as reported in previous work⁷⁵. Intra and interchain hydrogen bonds were calculated using the pytraj python package⁶⁸. Inter chain hydrogen bonds were normalized by dividing it by the number of chains (n=6).

Acknowledgments

We would like to thank Dr. Micholas Dean Smith and Dr. Omar Demerdash at ORNL for their valuable discussions. This research made use of the Compute and Data Environment for Science (CADES) computing resource at the Oak Ridge National Laboratory (ORNL), which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. This research was supported by the Genomic Science Program, Office of

Biological and Environmental Research, U. S. Department of Energy (DOE), under Contract FWP ERKP752 and of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U. S. Department of Energy under Contract No. DE-AC02-05CH11231.

References

- 1 Vieille, C. & Zeikus, G. J. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* **65**, 1-43, doi:10.1128/MMBR.65.1.1-43.2001 (2001).
- 2 Rigoldi, F., Donini, S., Redaelli, A., Parisini, E. & Gautieri, A. in *APL Bioeng* Vol. 2 (2018).
- 3 Mohammad, B. T., Al Daghistani, H. I., Jaouani, A., Abdel-Latif, S. & Kennes, C. Isolation and Characterization of Thermophilic Bacteria from Jordanian Hot Springs: *Bacillus licheniformis* and *Thermomonas hydrothermalis* Isolates as Potential Producers of Thermostable Enzymes. *Int J Microbiol* **2017**, 6943952, doi:10.1155/2017/6943952 (2017).
- 4 Yayanos, A. A. Piezophiles. *e LS* (2001).
- 5 Daniel, R. M., Dines, M. & Petach, H. H. The denaturation and degradation of stable enzymes at high temperatures. *Biochem J* **317** (Pt 1), 1-11, doi:10.1042/bj3170001 (1996).
- 6 Roche, J. & Royer, C. A. Lessons from pressure denaturation of proteins. *J R Soc Interface* **15**, doi:10.1098/rsif.2018.0244 (2018).
- 7 Vogt, G., Woell, S. & Argos, P. Protein thermal stability, hydrogen bonds, and ion pairs. *J Mol Biol* **269**, 631-643, doi:10.1006/jmbi.1997.1042 (1997).
- 8 Pucci, F. & Rooman, M. Physical and molecular bases of protein thermal stability and cold adaptation. *Curr Opin Struct Biol* **42**, 117-128, doi:10.1016/j.sbi.2016.12.007 (2017).
- 9 Salminen, T. *et al.* An unusual route to thermostability disclosed by the comparison of *Thermus thermophilus* and *Escherichia coli* inorganic pyrophosphatases. *Protein science : a publication of the Protein Society* **5**, 1014-1025, doi:10.1002/pro.5560050604 (1996).
- 10 Folch, B., Dehouck, Y. & Rooman, M. in *Biophysical journal* Vol. 98 667-677 (2010).
- 11 Szilágyi, A. & Závodszky, P. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure* **8**, 493-504, doi:10.1016/s0969-2126(00)00133-7 (2000).
- 12 Argos, P. *et al.* Thermal stability and protein structure. *Biochemistry* **18**, 5698-5703, doi:10.1021/bi00592a028 (1979).
- 13 Shrestha, U. R. *et al.* Effects of pressure on the dynamics of an oligomeric protein from deep-sea hyperthermophile. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 13886-13891, doi:10.1073/pnas.1514478112 (2015).
- 14 Brock, T. D. & Freeze, H. *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. *J Bacteriol* **98**, 289-297 (1969).
- 15 Librizzi, F., Carrotta, R., Peters, J. & Cupane, A. The effects of pressure on the energy landscape of proteins. *Sci Rep* **8**, 2037, doi:10.1038/s41598-018-20417-x (2018).
- 16 Mallamace, F. *et al.* Energy landscape in protein folding and unfolding. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 3159-3163, doi:10.1073/pnas.1524864113 (2016).
- 17 Frauenfelder, H. *et al.* Proteins and Pressure. *Journal of Physical Chemistry* **94**, 1024-1037, doi:DOI 10.1021/j100366a002 (1990).
- 18 Meinhold, L., Smith, J. C., Kitao, A. & Zewail, A. H. Picosecond fluctuating protein energy landscape mapped by pressure temperature molecular dynamics simulation.

- Proceedings of the National Academy of Sciences of the United States of America* **104**, 17261-17265, doi:10.1073/pnas.0708199104 (2007).
- 19 Meinhold, L. *et al.* Protein dynamics and stability: the distribution of atomic fluctuations in thermophilic and mesophilic dihydrofolate reductase derived using elastic incoherent neutron scattering. *Biophys J* **94**, 4812-4818, doi:10.1529/biophysj.107.121418 (2008).
- 20 Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964-972, doi:10.1038/nature06522 (2007).
- 21 Shrestha, U. R. *et al.* Quasi-elastic Neutron Scattering Reveals Ligand-Induced Protein Dynamics of a G-Protein-Coupled Receptor. *J Phys Chem Lett* **7**, 4130-4136, doi:10.1021/acs.jpcclett.6b01632 (2016).
- 22 Ferrand, M., Dianoux, A. J., Petry, W. & Zaccai, G. Thermal motions and function of bacteriorhodopsin in purple membranes: effects of temperature and hydration studied by neutron scattering. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 9668-9672, doi:10.1073/pnas.90.20.9668 (1993).
- 23 Borzova, V. A. *et al.* Kinetics of Thermal Denaturation and Aggregation of Bovine Serum Albumin. *PLoS One* **11**, e0153495, doi:10.1371/journal.pone.0153495 (2016).
- 24 Karshikoff, A., Nilsson, L. & Ladenstein, R. Rigidity versus flexibility: the dilemma of understanding protein thermal stability. *FEBS J* **282**, 3899-3917, doi:10.1111/febs.13343 (2015).
- 25 Roche, J. *et al.* Cavities determine the pressure unfolding of proteins. *Proc Natl Acad Sci U S A* **109**, 6945-6950, doi:10.1073/pnas.1200915109 (2012).
- 26 Doster, W. & Gebhardt, R. High pressure - unfolding of myoglobin studied by dynamic neutron scattering. *Chemical Physics* **292**, 383-387, doi:10.1016/S0301-0104(03)00064-8 (2003).
- 27 Fukuda, M. & Kunugi, S. Pressure dependence of thermolysin catalysis. *Eur J Biochem* **142**, 565-570, doi:10.1111/j.1432-1033.1984.tb08323.x (1984).
- 28 Konisky, J., Michels, P. C. & Clark, D. S. Pressure stabilization is not a general property of thermophilic enzymes: the adenylate kinases of *Methanococcus voltae*, *Methanococcus maripaludis*, *Methanococcus thermolithotrophicus*, and *Methanococcus jannaschii*. *Appl Environ Microbiol* **61**, 2762-2764 (1995).
- 29 Hernandez, G., Jenney, F. E., Jr., Adams, M. W. & LeMaster, D. M. Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 3166-3170, doi:10.1073/pnas.040569697 (2000).
- 30 Yu, H. & Huang, H. Engineering proteins for thermostability through rigidifying flexible sites. *Biotechnol Adv* **32**, 308-315, doi:10.1016/j.biotechadv.2013.10.012 (2014).
- 31 Zhang, X. F. *et al.* A general and efficient strategy for generating the stable enzymes. *Sci Rep* **6**, 33797, doi:10.1038/srep33797 (2016).
- 32 Ogasahara, K., Khechinashvili, N. N., Nakamura, M., Yoshimoto, T. & Yutani, K. Thermal stability of pyrrolidone carboxyl peptidases from the hyperthermophilic Archaeon, *Pyrococcus furiosus*. *Eur J Biochem* **268**, 3233-3242, doi:10.1046/j.1432-1327.2001.02220.x (2001).
- 33 Thoma, R., Hennig, M., Sterner, R. & Kirschner, K. Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*. *Structure* **8**, 265-276, doi:10.1016/s0969-2126(00)00106-4 (2000).

- 34 Kumar, S., Tsai, C. J. & Nussinov, R. Factors enhancing protein thermostability. *Protein Eng* **13**, 179-191, doi:10.1093/protein/13.3.179 (2000).
- 35 Hughes, R. C. *et al.* Inorganic pyrophosphatase crystals from *Thermococcus thioireducens* for X-ray and neutron diffraction. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **68**, 1482-1487, doi:10.1107/S1744309112032447 (2012).
- 36 Samygina, V. R. *et al.* The structures of *Escherichia coli* inorganic pyrophosphatase complexed with Ca(2+) or CaPP(i) at atomic resolution and their mechanistic implications. *J Mol Biol* **314**, 633-645, doi:10.1006/jmbi.2001.5149 (2001).
- 37 Chu, X. Q. *et al.* Dynamic behavior of oligomeric inorganic pyrophosphatase explored by quasielastic neutron scattering. *J Phys Chem B* **116**, 9917-9921, doi:10.1021/jp303127w (2012).
- 38 Josse, J. Constitutive inorganic pyrophosphatase of *Escherichia coli*. 1. Purification and catalytic properties. *The Journal of biological chemistry* **241**, 1938-1947 (1966).
- 39 Cavalli, A., Salvatella, X., Dobson, C. M. & Vendruscolo, M. Protein structure determination from NMR chemical shifts. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 9615-9620, doi:10.1073/pnas.0610313104 (2007).
- 40 Markwick, P. R., Malliavin, T. & Nilges, M. Structural biology by NMR: structure, dynamics, and interactions. *PLoS Comput Biol* **4**, e1000168, doi:10.1371/journal.pcbi.1000168 (2008).
- 41 Hollingsworth, S. A. & Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **99**, 1129-1143, doi:10.1016/j.neuron.2018.08.011 (2018).
- 42 Karplus, M. & Kuriyan, J. Molecular dynamics and protein function. *Proc Natl Acad Sci U S A* **102**, 6679-6685, doi:10.1073/pnas.0408930102 (2005).
- 43 Ulmschneider, J. P. & Ulmschneider, M. B. Molecular Dynamics Simulations Are Redefining Our View of Peptides Interacting with Biological Membranes. *Acc Chem Res* **51**, 1106-1116, doi:10.1021/acs.accounts.7b00613 (2018).
- 44 Tavernelli, I., Cotesta, S. & Di Iorio, E. E. Protein dynamics, thermal stability, and free-energy landscapes: a molecular dynamics investigation. *Biophysical journal* **85**, 2641-2649, doi:10.1016/S0006-3495(03)74687-6 (2003).
- 45 Henzler-Wildman, K. A. *et al.* A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913-916, doi:http://www.nature.com/nature/journal/v450/n7171/suppinfo/nature06407_S1.html (2007).
- 46 Henzler-Wildman, K. A. *et al.* Intrinsic motions along an enzymatic reaction trajectory. *Nature* **450**, 838-844, doi:10.1038/nature06410 (2007).
- 47 Robinson-Rechavi, M., Alibes, A. & Godzik, A. Contribution of electrostatic interactions, compactness and quaternary structure to protein thermostability: lessons from structural genomics of *Thermotoga maritima*. *J Mol Biol* **356**, 547-557, doi:10.1016/j.jmb.2005.11.065 (2006).
- 48 Bogatyreva, N. S. & Ivankov, D. N. [The relationship between protein accessible surface area and number of native contacts in its structure]. *Mol Biol (Mosk)* **42**, 1048-1055 (2008).
- 49 Hubbard, S. J. & Argos, P. Cavities and packing at protein interfaces. *Protein science : a publication of the Protein Society* **3**, 2194-2206, doi:10.1002/pro.5560031205 (1994).
- 50 Sonavane, S. & Chakrabarti, P. in *PLoS Comput Biol* Vol. 4 (2008).

- 51 Cioni, P. Role of protein cavities on unfolding volume change and on internal dynamics under pressure. *Biophysical journal* **91**, 3390-3396, doi:10.1529/biophysj.106.085670 (2006).
- 52 Dundas, J., Adamian, L. & Liang, J. Structural signatures of enzyme binding pockets from order-independent surface alignment: a study of metalloendopeptidase and NAD binding proteins. *J Mol Biol* **406**, 713-729, doi:10.1016/j.jmb.2010.12.005 (2011).
- 53 Kumar, P. *et al.* Low-Barrier and Canonical Hydrogen Bonds Modulate Activity and Specificity of a Catalytic Triad. *Angew Chem Int Ed Engl* **58**, 16260-16266, doi:10.1002/anie.201908535 (2019).
- 54 Yang, L., Liao, R. Z., Yu, J. G. & Liu, R. Z. DFT study on the mechanism of Escherichia coli inorganic pyrophosphatase. *J Phys Chem B* **113**, 6505-6510, doi:10.1021/jp810003w (2009).
- 55 Cooperman, B. S., Baykov, A. A. & Lahti, R. Evolutionary conservation of the active site of soluble inorganic pyrophosphatase. *Trends Biochem Sci* **17**, 262-266, doi:10.1016/0968-0004(92)90406-y (1992).
- 56 Kurkal, V. *et al.* Enzyme activity and flexibility at very low hydration. *Biophysical journal* **89**, 1282-1287, doi:10.1529/biophysj.104.058677 (2005).
- 57 Kurkal, V. *et al.* Low frequency enzyme dynamics as a function of temperature and hydration: A neutron scattering study. *Chemical Physics* **317**, 267-273, doi:10.1016/j.chemphys.2005.05.019 (2005).
- 58 Carugo, O. Structure and function of water molecules buried in the protein core. *Curr Protein Pept Sci* **16**, 259-265, doi:10.2174/1389203716666150227162803 (2015).
- 59 Liu, Z. *et al.* Entropic contribution to enhanced thermal stability in the thermostable P450 CYP119. *Proc Natl Acad Sci U S A* **115**, E10049-E10058, doi:10.1073/pnas.1807473115 (2018).
- 60 Paredes, D. I., Watters, K., Pitman, D. J., Bystroff, C. & Dordick, J. S. Comparative void-volume analysis of psychrophilic and mesophilic enzymes: Structural bioinformatics of psychrophilic enzymes reveals sources of core flexibility. *BMC Struct Biol* **11**, 42, doi:10.1186/1472-6807-11-42 (2011).
- 61 Xue, M. *et al.* How internal cavities destabilize a protein. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 21031-21036, doi:10.1073/pnas.1911181116 (2019).
- 62 Závodszy, P., Kardos, J., Svingor, Á. & Petsko, G. A. in *Proceedings of the National Academy of Sciences of the United States of America* Vol. 95 7406-7411 (1998).
- 63 Di Giulio, M. The universal ancestor and the ancestor of bacteria were hyperthermophiles. *J Mol Evol* **57**, 721-730, doi:10.1007/s00239-003-2522-6 (2003).
- 64 Berezovsky, I. N. & Shakhnovich, E. I. Physics and evolution of thermophilic adaptation. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 12742-12747, doi:10.1073/pnas.0503890102 (2005).
- 65 Catchpole, R. & Forterre, P. Positively twisted: The complex evolutionary history of Reverse Gyrase suggests a non-hyperthermophilic Last Universal Common Ancestor. doi:10.1101/524215 (2019).
- 66 Abraham, M., van der Spoel, D., Lindahl, E. & Hess, B.
- 67 Abraham, M. J. & Gready, J. E. Optimization of parameters for molecular dynamics simulation using smooth particle-mesh Ewald in GROMACS 4.5. *Journal of computational chemistry* **32**, 2031-2040, doi:10.1002/jcc.21773 (2011).

- 68 Roe, D. R. & Cheatham, T. E., 3rd. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* **9**, 3084-3095, doi:10.1021/ct400341p (2013).
- 69 Bussi, G., Donadio, D. & Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
- 70 Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. doi:1.328693 (1998).
- 71 Fine, R. A. & Millero, F. J. Compressibility of Water as a Function of Temperature and Pressure. *Journal of Chemical Physics* **59**, 5529-5536, doi:Doi 10.1063/1.1679903 (1973).
- 72 Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *Journal of Chemical Physics* **98**, doi:1.464397 (1993).
- 73 Maiti, R., Van Domselaar, G. H., Zhang, H. & Wishart, D. S. SuperPose: a simple server for sophisticated structural superposition. *Nucleic acids research* **32**, W590-594, doi:10.1093/nar/gkh477 (2004).
- 74 Paramo, T., East, A., Garzon, D., Ulmschneider, M. B. & Bond, P. J. Efficient Characterization of Protein Cavities within Molecular Simulation Trajectories: trj_cavity. *J Chem Theory Comput* **10**, 2151-2164, doi:10.1021/ct401098b (2014).
- 75 Gerstein, M. & Chothia, C. Packing at the protein-water interface. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 10167-10172, doi:10.1073/pnas.93.19.10167 (1996).