Rational engineering of a highly active and resilient α-carbonic anhydrase from the hydrothermal vent species *Persephonella hydrogeniphila*

Colleen Varaidzo Manyumwa[1] (iD), Chenxi Zhang[1], Carsten Jers[1] and Ivan Mijakovic[1,2] (iD)

1 The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark
2 Systems and Synthetic Biology Division, Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden

Carbonic anhydrases (CAs) are ideal catalysts for carbon dioxide sequestration in efforts to alleviate climate change. Here, we report the characterisation of three α-CAs that originate from the thermophilic bacteria *Persephonella hydrogeniphila* (PhyCA), *Persephonella atlantica* (PaCA), and *Persephonella* sp. KM09-Lau-8 (PlauCA) isolated from hydrothermal vents. The three α-Cas, showing high sequence similarities, were produced in *Escherichia coli*, purified and characterised. Surprisingly, they revealed very different behaviours with regards to their thermostability profiles. PhyCA presented a more stable thermostability profile amongst the three, thus we chose it for rational engineering to improve it further. PhyCA's residue K88, a proton transfer residue in α-CAs, was mutated to His, Ala, Gln and Tyr. A 4-fold activity improvement was noted for variants K88H and K88Q at 30 °C, owing to the higher proton transfer efficiency of the replacement proton transfer residues. K88Q also proved more stable than PhyCA. K88Y did not increase activity, but notably increased thermal stability, with this enzyme variant retaining 50% of its initial activity after incubation for 1 h at 90 °C. Removal of the two main proton shuttles (variant H85A_K88A) resulted in diminished activity of the enzyme. Molecular dynamics simulations performed for PhyCA and all its variants revealed differences in residue fluctuations, with K88A resulting in a general reduction in root mean square fluctuation (RMSF) of active site residues as well as most of the CA's residues. Its specific activity and stability in turn increased compared to the wild type.

Introduction

Carbonic anhydrases (CAs) are enzymes that catalyse the reversible reaction between water and CO_2 to produce bicarbonate and hydrogen ions. This enzyme has become prominent due to its application in various methods for removal of CO_2. CAs are particularly useful for treating major sources of CO_2 such as flue gas from power plants. Since such sources generally have a high temperature [1], thermostability of CAs is often an important requirement. To date, eight different classes of CAs are known: α, β, γ, δ, ζ, η, θ, and ι classes [2–5]. The α-class is the most widely studied, comprising enzymes that are highly active and thermostable, originating from bacteria isolated from hot springs and hydrothermal vents. Extreme environments

have been relatively reliable sources of industrially important enzymes throughout the years. Organisms that have adapted to survive and thrive in these environments are known to produce enzymes with properties adapted to the extreme surroundings. Halophilic environments, for example, have been noted to harbour bacteria that can produce enzymes with high salt tolerance [6–8]. Given that heat tolerance is a desirable characteristic of enzymes in today's industrial settings, thermostable enzymes have been sought after in thermophilic bacteria isolated from high temperature habitats such as hot springs, hydrothermal vents, and even volcanic regions [9–12]. The most active CA reported to date is the SazCA from the hot spring bacterium *Sulfurihydrogenibium azorense* [13,14]. *Sulfurihydrogenibium yellowstonense* [15,16], also was isolated from a hot spring, produces a CA that is quite efficient and somewhat thermostable. From the hydrothermal vent systems, bacteria such as *Thermovibrio ammonificans* and *Persephonella marina* have been isolated and found to produce thermophilic CAs [17–21].

Proton transfer has been identified as a rate-limiting step in the hydration of CO_2 by Cas. A number of CA amino acid residues are lined up in the catalytic cavity to carry out this reaction [22,23]. The proton comes from a water molecule that binds to the Zn^{2+} ion found in the active site, which is coordinated by three His residues in α-CAs and γ-CAs. Some CAs, however, bind other metal ions such as Fe^{2+} and Mg^{2+} [24,25]. In human CAII (HCAII), the main proton transfer residue has been identified to be H64 [23,26,27]. Mutation studies have been undertaken to mutate the residue equivalent in HCAIII, originally a Lys, to a His (K64H). This enhanced activity to some extent but did not match the proton shuttling efficiency in HCAII, possibly due to structural constraints [28]. In bacterial CAs, H85 (PhyCA numbering) is the structural equivalent of H64 in HCA II. H85 and K88 are considered to be involved in proton shuttling as well as the residues Y26, N83, T196 and T197 [23,26,29,30]. Interestingly, all these residues form hydrogen bonds as both proton donors and acceptors, except for K88, which is just a donor. A degree of conservation has been observed for the position of K88 in bacterial α-CAs, with Lys and Gln being the two alternatives [31,32].

Given the discovery of thermostable CAs from *P. marina*, we decided to explore other CAs from the genus *Persephonella* in this study. One of them originates from the bacterium *Persephonella hydrogeniphila* [33], PhyCA, and was studied *in silico* previously [32]. Other CAs we considered are PaCA from *Persephonella atlantica* [34] and PlauCA from *Persephonella* sp.

KM09-Lau-8 [35]. After characterisation of the selected wild-type α-CAs, PhyCA showed a more stable thermostability profile, thus we further explored rational enzyme engineering on this CA. This was done in an attempt to (a) increase enzyme activity and/or stability and (b) find out the level of importance of the proton shuttle residue's presence in PhyCA, as a representative of bacterial α-CAs. Molecular dynamics (MD) simulations and computational analysis followed mutation studies in an endeavour to understand structural differences caused by mutations.

Results and discussion

α-CAs from genus *Persephonella* show high sequence similarity but different thermostability profiles

Persephonella sp. is a Gram-negative bacterium that is often associated with hydrothermal vent systems, and it grows at high temperatures (60–90 °C) [26]. In this study, we explored CAs originating from the genus *Persephonella*, other than *P. marina* which has already been extensively studied. Similar to *P. marina* CAs, previous *in silico* work suggested *P. hydrogeniphila* CA (PhyCA) to be stable at high temperatures [32]. Using the PhyCA sequence as query in BLASTp analysis led to identification of the CAs from *Persephonella* sp. KM09-Lau-8 (PlauCA) and *P. atlantica* (PaCA) with sequence identities of 80% and 72%, respectively. We continued to align these three enzymes to *P. marina* CAs (PmCA1 and PmCA2) as well as the well-characterised enzymes from *T. ammonificans* (TaCA) and *S. azorense* (SazCA) (Fig. 1A). The sequence alignment revealed high sequence similarity amongst the CA genes from *Persephonella* with each pair exhibiting over 70% sequence identity. PhyCA and PaCA were most similar to PmCA2, which was isolated from the Logatchev hydrothermal fields, with 89% and 84% sequence identities respectively. In the alignment, all catalytic pocket residues are completely conserved except the protein shuttling residue K88. All the enzymes had a Lys in this position, except SazCA, where a Gln is found.

We synthesised 6xHis-tagged variants of CAs in *Escherichia coli* and purified the proteins in order to perform an experimental characterisation of their enzymatic activity. SDS/PAGE revealed the sizes of the proteins investigated in this study to be approximately 25 kDa (PaCA), 26 kDa (PhyCA) and 26 kDa (PlauCA), which fit with the theoretical molecular weights (Fig. S1). Specific activities of PaCA, PlauCA and PhyCA were 2913.2, 3455.8 and 2832.2
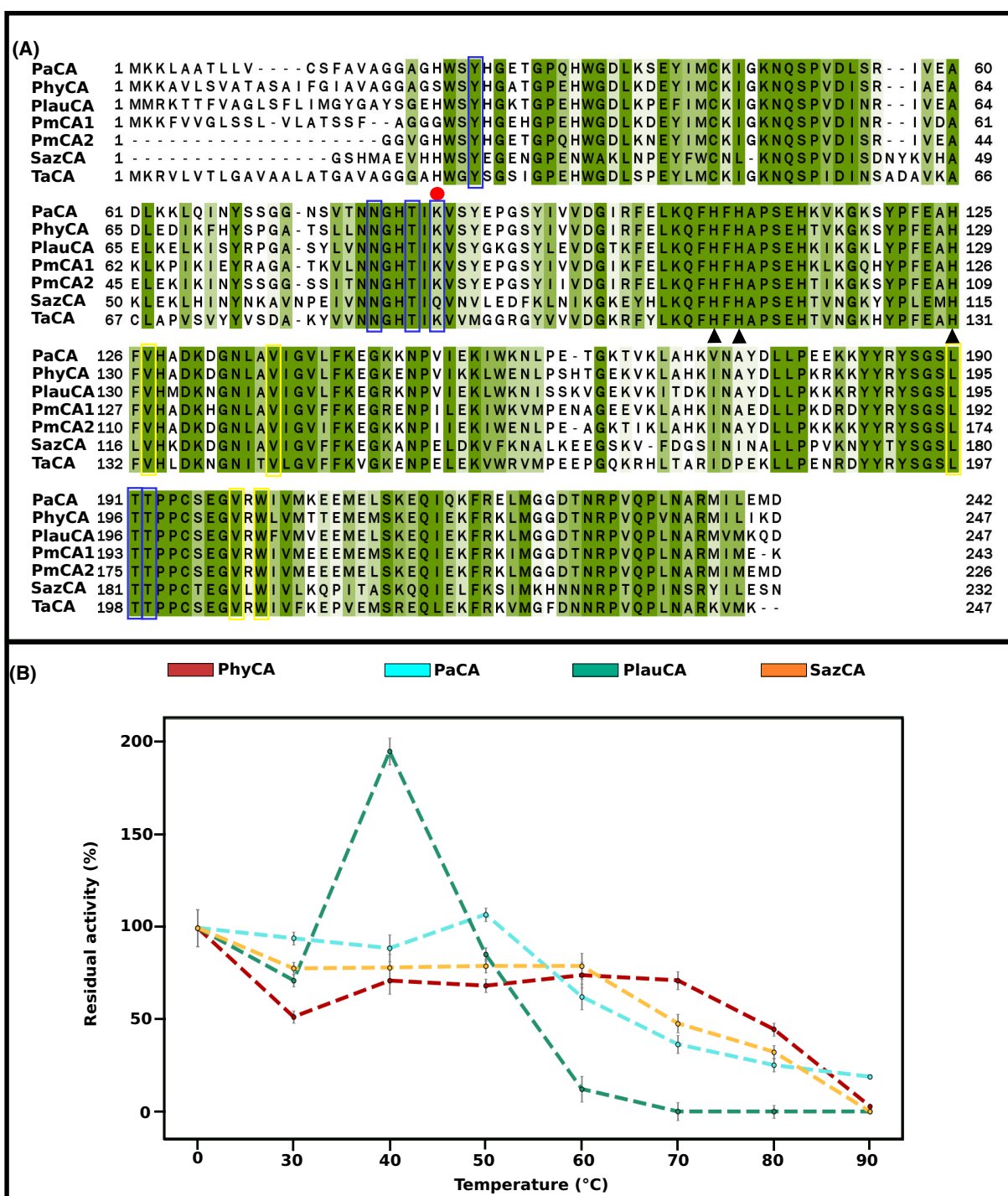
Fig. 1. (A) Sequence alignment of α-CA sequences from *P. atlantica* (NCBI Accession number: WP_200672902.1), *P. hydrogeniphila* (NCBI Accession number: WP_096999253.1), *P.* sp. KM09-Lau-8 (NCBI Accession number: WP_231475403.1), *P. marina* (PDB ID: 6IM3 and 6EKI), *S. azorense* (PDB ID: 4X5S) and *T. ammonificans* (PDB ID: 43CT). Residues with over 50% conservation are coloured green, with intensity increasing with an increase in conservation. Alignment was performed using the Tree-based Consistency Objective Function for Alignment Evaluation (T-COFFEE). The red dot shows the proton transfer residue position targeted for mutation in PhyCA. Zn^{2+} coordinating residues are shown by the black arrow heads, residues boxed in yellow form the CO_2 binding pocket and those in blue boxes are involved in proton shuttling. (B) Residual activities of PhyCA, PaCA, PlauCA and SazCA after incubation for 1 h at the respective temperatures are expressed as percentages of the initial untreated sample (0 °C). Each point is the mean ± SD of three independent reactions.

WAU·mg^{-1} respectively. To evaluate the thermostability, we incubated the proteins at various temperatures for 1 h and tested the residual activity (Fig. 1B). PlauCA exhibited low thermostability compared to PhyCA and PaCA, with only 12% residual activity at 60 °C and losing all activity by 70 °C. A comparable thermostability profile was observed for PhyCA and PaCA, but the former had lower activity at temperatures 30–50 °C. Although PaCA had 25% residual activity at 90 °C, its activity started to decrease at 60 °C whereas PhyCA appeared to maintain activity better, only showing a decrease at 80 °C. We therefore decided to proceed with engineering PhyCA.

Single mutations of proton transfer residue, K88, improve specific activity of PhyCA and variant K88Y retains almost half its initial activity after incubation at 90 °C

PhyCA was used as a case study for rational mutagenesis to investigate if its activity could be increased to match or surpass that of SazCA, the most active bacterial CA to date (see Table 1). In order to do this, we targeted the proton transfer residue K88, which was the only active site residue with variation in our alignment. We also included thermostability assays to view how single mutations of this residue would affect the thermostability profile of PhyCA.

The proton transfer residue, K88 in PhyCA was mutated to Gln, which is the residue present in SazCA and then further mutated to His and Tyr, which are also residues capable of proton shuttling. In the studies of the human carbonic anhydrase III (HCAIII), proton transfer residue K64 has been declared an inefficient proton shuttle because it is too basic [22,36,37]. Its mutation to His enhanced catalytic activity although proton transfer efficiency was not at par with HCAII, which has a His residue and is known for its high catalytic activity [37]. Mutation of PhyCA's K88

to non-reactive Ala was used to investigate the effect of the absence of this proton shuttle. Since there is a second proton shuttle H85 in the catalytic site, we assumed it possible that the absence of a proton transfer residue in position 88 would not render the enzyme completely inactive. Depiction of all active sites of the variants and the wild-type (WT) PhyCA are shown in Fig. 2.

Enzymatic analysis results for variant K88A were quite unexpected. It showed a > 1.2-fold increase in activity compared to the WT at 0 °C (Table 1). Ala is a small non-reactive residue which is not capable of taking up the role of proton donation and/or accepting. The mere presence of activity in this variant is evidence that the other proton shuttle present in α-CAs (H85) is more than capable of carrying out CO_2 hydration and like in human CAs, is likely the main proton shuttle. Substitution of H85 with Ala along with the mutation K88A decreased the enzyme activity almost 35-fold. This is consistent with results obtained for HCA II, which had a 25-fold decrease in activity after introducing the H64A substitution [38]. K88 may thus not be as important a proton transfer residue but have a structural role instead. Results from testing thermostability of the variants show that although a decrease was observed at 50 and 60 °C for K88A, we detected an activation at 70 °C as also seen in K88Y (Fig. 2). Variant K88Y was less active than the other variants and the WT at lower temperature, but it surpassed all other enzymes at elevated temperatures, especially at 90 °C where it maintained almost 50% of its initial activity. Specific activities for K88A and K88Y were 4- and 10-fold higher than the WT after incubation at 90 °C respectively.

Thermostability of K88H was comparable to that of PhyCA, maintaining a steady activity up to 70 °C, followed by a decline at 80 °C and a sharp decrease at 90 °C. Specific activity for this mutant, however, was high, like that of K88Q and twice that of PhyCA (Table 1). This observation is similar to what was observed when the structural residue equivalent to K88, R67 in HCAIII was mutated to His [28]. K88Q interestingly had increased thermostability compared to PhyCA. Incubation of this variant at temperatures up to 80 °C did not affect residual activity, whereas incubation at 90 °C resulted in a drastic decrease in activity. K88Q had comparable activity to SazCA at higher temperatures, from 60 to 80 °C.

To verify that K88's contribution to proton shuttling is minor, the corresponding residue in SazCA (Q74) was mutated to a Lys and its activity and thermostability were tested. SazCA_Q74K's activity was

Table 1. Specific activities of WT CAs and their variants.

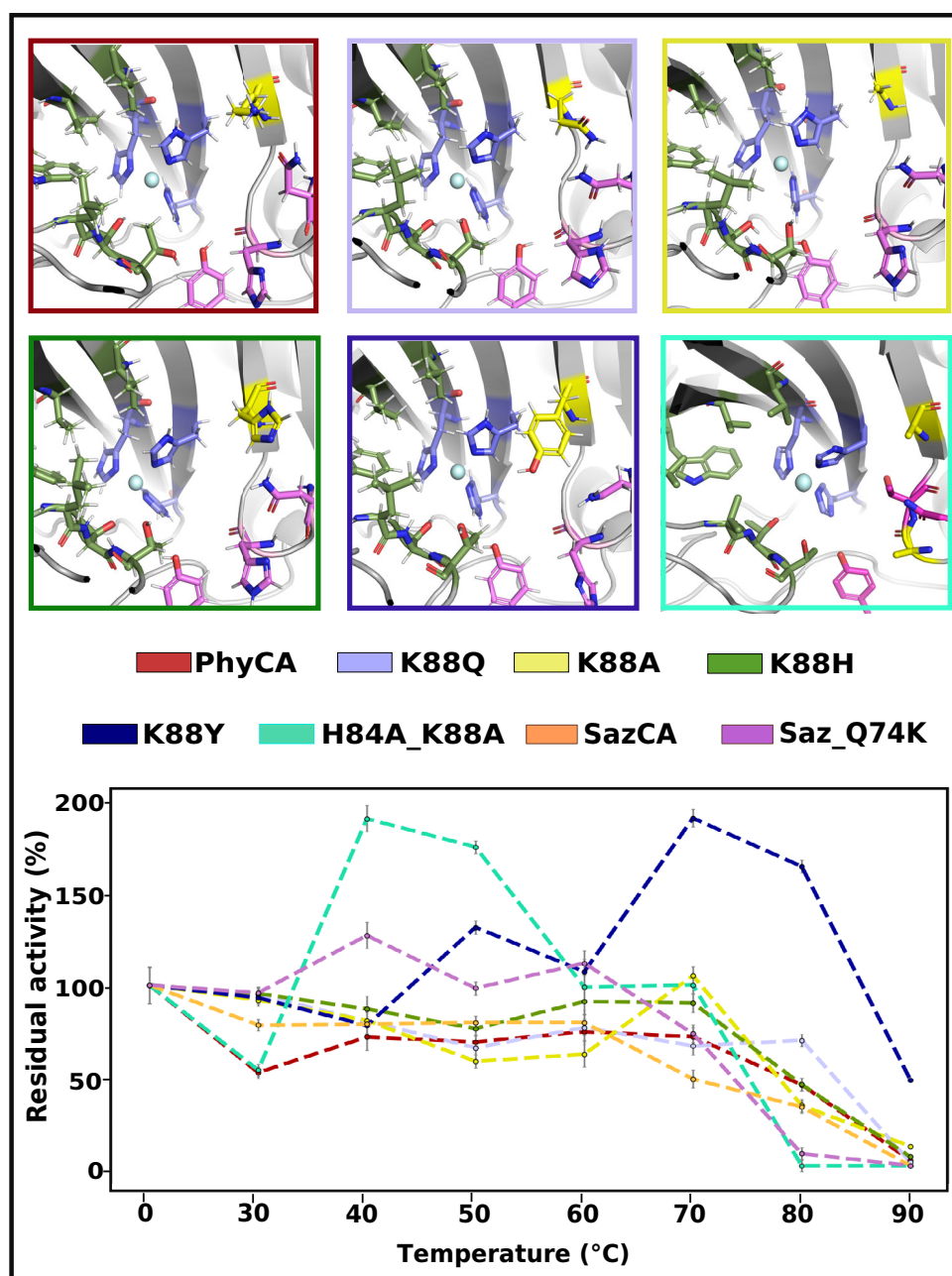| CA	Specific activity (WAU·mg^{-1}) (untreated)
PhyCA	2832.2
K88Q	6615.4
K88A	3413.8
K88H	6168.3
K88Y	1715.1
H85A_K88A	79.7
SazCA	12 027.8
SazCA_Q74K	4627.9

Fig. 2. Cartoon representations of the active site of PhyCA and variants K88Q, K88A, K88H and K88Y generated in PYMOL. Stick residues in green are CO_2 binding pocket residues, in blue are the His residues coordinating Zn^{2+} metal ion, which is shown as a pale cyan sphere. In pink are the residues involved in proton transfer including H85, and in yellow is the mutation residue in position 88. Residual activities of PhyCA and its variants, as well as SazCA and its mutant, Saz_Q74K, are portrayed below the structures as percentages of the initial untreated sample (0 °C) at their respective treatment temperatures (30–90 °C). Each point is the mean ± SD of three independent reactions.

reduced to almost a third of SazCA's initial activity at 0 °C, substantiating the hypothesis that Lys reduces the enzyme's ability to shuttle protons effectively (Table 1). Like SazCA, SazCA_Q74K exhibited decreased activity after incubation at 70 and 80 °C, but inactive at 90 °C.
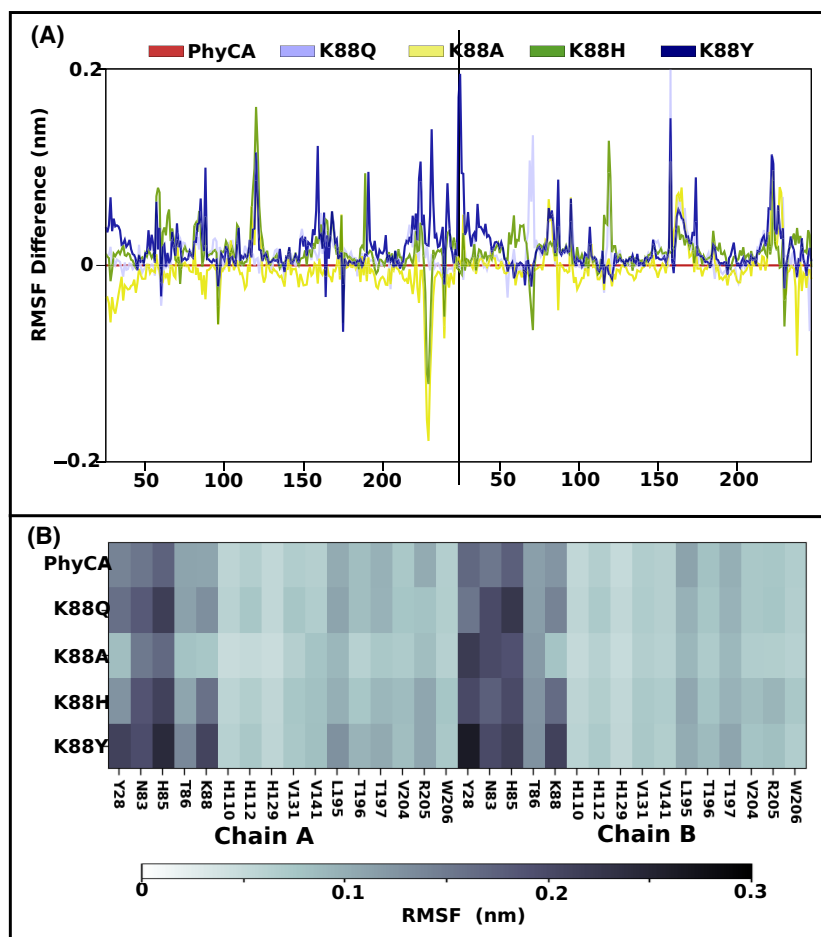
Fig. 3. Residue fluctuations in mutant structures. (A) Change in root mean square fluctuation (RMSF) of the dimeric structures of K88Q (light purple), K88A (yellow), K88H (green) and K88Y (blue) compared to PhyCA (red). (B) RMSF for catalytic pocket residues for chain A and B of PhyCA, K88Q, K88A, K88H, K88Y.

MD simulations reveal different fluctuation patterns amongst variants and interesting motions of proton shuttling residues in K88Y

We proceeded to conduct MD simulations on all variants, from which we carried out several analyses. Conformational changes of the active sites in the presence of the mutations were an important part of the analysis. All substitution residues considered for PhyCA, except Tyr, were smaller than the original Lys, thus a change in the active site opening was expected upon mutation. Root mean square deviation (RMSD) was therefore calculated for active sites to investigate the evolution of the pockets throughout the simulation (Fig. S2). Amongst all variants, K88Y showed a contrasting pattern and was thus investigated further.

Reduction in fluctuation of mutation residue in K88A may infer stability to catalytic pocket

Fluctuations of the residues for each variant were investigated using RMSF analysis (Fig. 3). For variants K88Q, K88H and K88Y, more than 78% of the residues fluctuated more than in the WT (Fig. 3A). Amongst these residues in all three variants, the substitution residue had higher RMSF values than the original K88. In both chains A and B, substitution residue Y88 was fluctuating twice as much as in WT and H85 also had higher RMSF values (Fig. 3B).

Contrary to all other variants, in K88A, most residues attained lower fluctuations, inferring structural stability brought about by the mutation. A closer look revealed a decrease in RMSF of mutated residue A88 in both chain A and B, possibly bringing

about a degree of stability to the catalytic pocket. Two residues that showed significant decreases in RMSF as well were T231, which is located behind the active site, as well as N239 found in the interface region possibly inducing stability. Stabilisation of the dimeric interface is further supported by the decrease of RMSF, though to a lesser extent, in other interface residues including G50, A62, A64, R190, S192, E202, G203, A240, M242, I243 and L244. Interface residues of PhyCA were previously described [32]. The size of substitution residue A88 also allowed for a visibly larger opening of the active site as well, explaining the increase in activity compared to WT PhyCA.

Conformational changes in variant K88Y's catalytic pocket are observed as proton shuttles H88 and Y88 move in and out of active site

Fluctuations of Y88 and H85 in variant K88Y were assumed to be an indication of increased proton transfer directly inferring higher activity. But there could exist presently unknown elements deterring this increase. Tyr is quite a large residue and appeared as though it was slightly restricting the active site, which might explain the lower activity compared to the other variants. However, its ability to preserve activity at 90 °C clearly suggests that this mutation contributed to stability of the active site at high temperatures. The RMSF observations (Fig. 3B) plus the peculiar RMSD pattern observed in Fig. S2 warranted a closer look at the conformations visited by the proton transfer residues as well as the active site. Tyr is a flexible residue known to rotate along the C_α-C_β and C_β-C_γ bonds, exhibiting four rotamer states [39,40]. Distances between the proton transfer residues H85 and substitution residue Y88 were thus measured in VMD, and the corresponding plot is shown in Fig. 4A. We observed that both H85 and Y88 were continuously moving toward and away from Zn^{2+} in the catalytic pocket, with H85 getting as close as 0.46 nm of the Zn^{2+} metal ion and Y88 even closer at 0.41 nm. Due to these motions, the active site of K88Y visited several different conformations back and forth, unlike the WT, which maintained relatively stable active site conformations throughout the simulation (Fig. S3). Structures of K88Y's active site were extracted for scrutiny at four different points of the simulation shown by the arrows and viewed in PYMOL.

The four points were (a) at 26.85 ns where H85 and Y88 were both rotated out, (b) at 43 ns where H85 was rotated into the active site and Y88 was rotated out, (c) at 50.41 ns where both Y88 and H85

were rotated into the active site and finally, (d) at 61.54 ns where Y88 was rotated in and H85 was rotated out. Illustrations of the active site at these points are shown in Fig. 4B. The active site opening at 43 ns was wide and shallow, similar to how the WT active site appears. Opening at 61.54 ns was a bit more constricted with Y88 in and H85 out, tilting to view the inside of the catalytic pocket revealed a deeper pocket, now including the contribution of residues T231 and N232 (in orange) at the bottom of the pocket. These two residues also almost doubled in RMSF compared to PhyCA, which is the opposite of what was observed for T231 in K88A. Opening at 50.41 ns was much narrower than the previous two with both Y88 and H85 rotated in, but the sub-pocket seen at 61.54 ns appeared even deeper, with F111 (also in orange), additionally contributing to this pocket as well. Activity of K88Y appears to be a constant evolution of the active sites amongst these different conformations, possibly being a little more restrictive toward movement of substrate in and product out of the activity site, resulting in a lower turnover.

Conclusion

In this study, the purified α-CAs PaCA, PlauCA and PhyCA displayed similar activities, but PhyCA's ability to preserve activity at elevated temperatures warranted its engineering to improve it further. We improved specific activity of PhyCA by targeting proton transfer residue, K88 and in the variants K88Q and K88H, by a factor of just over 2-fold. Higher thermotolerance was also observed in variant K88Q from 70 (in WT) to 80 °C (in mutant). We confirmed the reduced catalytic activity with Lys in position 88 by introducing it to SazCA which originally has a Gln, and this variant (Saz_Q74K) consequently decreased both in activity and thermostability. PhyCA's variant K88A surprisingly had a higher activity and stability than the WT, with residue fluctuations of most of the structure, including interface residues, decreasing with this mutation. Although a reduction in activity in variant K88Y, this mutation beneficially resulted in retaining 47% of the CA's activity at 90 °C compared to 2% in the WT. Overall, the objective to rationally engineer the CA from *P. hydrogeniphila* resulted in the discovery of the variant K88Y that would be an ideal candidate for CO_2 sequestration at high temperatures. The results obtained in this study have the potential to be applied to other CAs as well to enhance catalytic efficiency and/or thermostability.
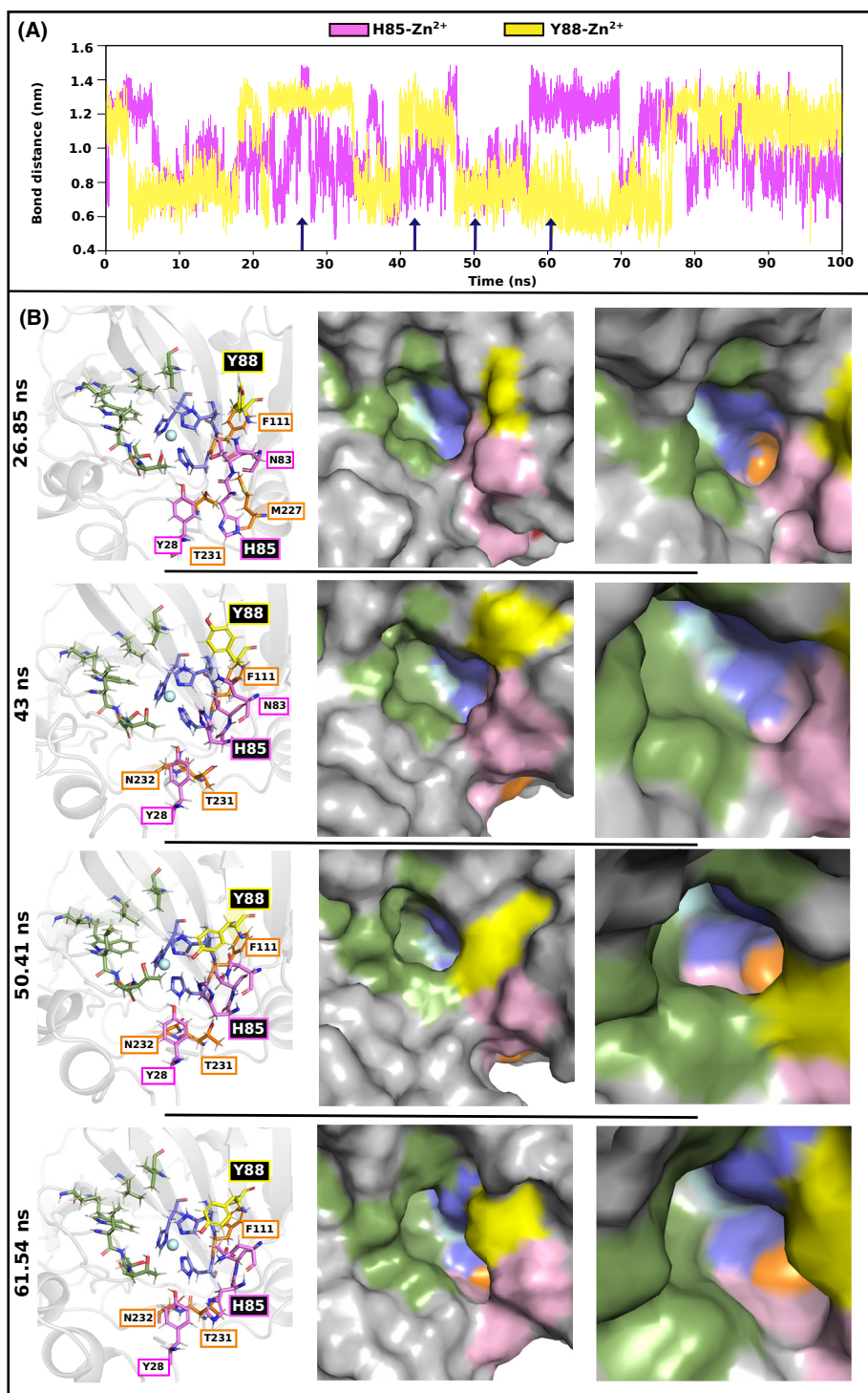
Fig. 4. Active site structure and corresponding coordinating atom bond distances from the Zn²⁺ during MD simulations of K88Y. (A) Distance between Zn²⁺ and (i) Y88 (yellow) and (ii) H85 (pink) over the course of the 100 ns trajectory. The arrows show the time frames from which structures in (B) were extracted. (B) The left column shows the stick representations of the active site residues at their different time frames indicated by the arrows. Middle column shows a zoomed-out view of the catalytic pocket, and the right column shows a zoomed in and tilted view of the inside of the catalytic pocket. Residues in green, blue and pink are the CO_2 pocket residues, Zn²⁺ coordinating residues and proton transfer residues respectively. The main proton shuttle, H85 is labelled in the box with the black background. Substitution residue Y88 is coloured yellow. All structures were generated in PYMOL.

Materials and methods

Sequence identification and retrieval

The sequence for *P. hydrogeniphila* α-CA (PhyCA, NCBI Accession number: WP_096999253.1) was acquired from NCBI [41] and submitted to NCBI BLASTp to search for other α-CAs from the genus *Persephonella*. From this analysis, two proteins originating from the bacteria *Persephonella* sp. KM09-Lau-8, (PlauCA, Accession number WP_231475403.1) and *P. atlantica* (PaCA, Accession number: WP_200672902.1) were selected. Sequence alignments were performed using T-COFFEE [42] for retrieved CAs PhyCA, PaCA and PlauCA as well sequences from *Persephonella marina* PmCA1 (PDB ID: 6IM3), PmCA2 (PDB ID: 6EKI). TaCA (PDB ID: 43CT) from *T. ammonificans* which is in the same order as *Persephonella* as well SazCA (PDB ID: 4X5S), one of the most efficient CAs currently from and *S. azorense*, was also included in the sequence alignment for comparison.

DNA manipulations and strain construction

Genes encoding the mature part of PaCA, PlauCA, PhyCA and SazCA were codon-optimised for *E. coli* and synthesised by Integrated DNA Technologies (IDT) (Table S1). The recombinant plasmids pETM10-PhyCA-His, pETM10-PaCA-His, pETm10-PlauCA-His and pETM10-SazCA-His, were constructed using the Gibson Assembly method [43]. The vector backbone was PCR-amplified using relevant primers and pETM10 vector as template. Variants of PhyCA were obtained by site-directed mutagenesis. All oligonucleotide primers were procured from IDT and are listed in Tables S2 and S3.

E. coli strains NM522 and BL21(DE3) were used for cloning and expression respectively. The strains containing the plasmid were grown overnight in LB medium (1% (w/v) peptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl) supplemented with 50 μg·mL^{-1} kanamycin.

For protein synthesis, the *E. coli* BL21(DE3) with relevant plasmids were cultured in LB medium to an OD$_{600}$ of 0.5–0.6 at which point IPTG (final concentration 1 mM) was added to induce protein expression. Cultures were grown for a further 3 h after which cells were harvested by centrifugation at 6000 *g* for 10 min. The cell pellet was resuspended in lysis buffer (25 mM Tris/HCl (pH 8.0), 100 mM NaCl, 10% glycerol, 0.2 mg·mL^{-1} lysozyme) prior to sonication using a probe sonicator (3 min, 70% amplitude). The lysed cells were centrifuged at 6000 *g* for 15 min and the supernatant was collected. This was loaded onto an Ni-NTA column equilibrated with an equilibration buffer (20 mM Tris, 0.5 M NaCl, pH 7.5), washed with washing buffer (20 mM Tris, 0.5 M NaCl, 10 mM imidazole, pH 7.5) and eluted with elution buffer (20 mM Tris, 0.5 M NaCl, 250 mM imidazole, pH 7.5). PD-10 columns (Cytiva, Muskegon, Michigan, USA) were equilibrated with the equilibration buffer and used to remove the imidazole from protein samples. Protein concentration was measured using the Bradford's assay [44] using bovine serum albumin (BSA) as a standard.

CO$_2$ hydration assay

CO$_2$ hydration activity was measured by the Wilbur-Anderson method [45]. Dry ice was bubbled into deionised water for approximately 30 min to attain a CO$_2$-saturated solution as the substrate. Tris/Cl buffer (50 mM, pH 8.4) was mixed with 0.04% bromothymol blue indicator to record the change in pH from 8.4 (blue) to 6.4 (yellow). All reactions were performed on ice. To start the reaction, 3 mL CO$_2$ saturated solution was added to 2 mL of coloured buffer and 50 μL enzyme solution (50 μL Tris/Cl buffer for the uncatalyzed reaction). Time for the uncatalyzed reaction was noted as T_0 and the enzyme-driven reaction was T. Activity was calculated in Wilbur-Anderson units (WAU) using the formula $(T_0 - T)/T$. Specific activity (WAU·mg^{-1}) was determined by dividing WAU with protein concentration.

Thermostability assays

Thermostability was investigated by incubating samples at individual temperatures from 30 to 90 °C in 10 °C increments for 1 h. Samples were cooled on ice for an hour and then assayed on ice as described in the previous section (CO$_2$ hydration assay). Residual activity was evaluated using the CO$_2$ hydration assay.

Molecular dynamics simulations

Dimeric structures for PhyCA and its variants were modelled using swiss-model [46] using PmCA2 (PDB ID: 6EKI) as a template. Following validation using z-DOPE (Discrete Optimized Protein Energy) and verify3d [47], the structures were protonated using h++ [48]. ambertools22 [49,50] was applied for the inference of metal parameters to Zn^{2+} coordinating residues and those immediately surrounding them [27]. MD simulations were performed at 363 K (90 °C) using gromacs v2020.6 [51] following minimisation, canonical ensemble and isothermal-isobaric equilibration consecutively at 363 K. gromacs was used to calculate root mean square deviation (RMSD) to observe changes in conformations, root mean square fluctuation (RMSF) to investigate individual residue fluctuations and radius of gyration (Rg) to understand changes in protein compactness over the course of the simulation. visual molecular dynamics software (vmd) [52] was used to measure atomic distances between proton transfer residues and Zn^{2+} metal ion in the active site and their evolution throughout the trajectory and pymol [53] was used to view structures.

Note that the amino acid numbering follows that of Phy-CA's sequence unless explicitly stated otherwise.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

Author contributions

Conceptualisation, IM and CVM; methodology, CVM, CJ and CZ; validation, CVM and CZ; formal analysis, CVM and CZ; investigation, CVM, CJ and CZ.; resources, IM; writing—original draft preparation, CVM; writing—review and editing, IM, CJ, CVM and CZ; visualisation, CVM and CZ; supervision, IM and CJ; project administration, IM and CJ; funding acquisition, IM. All authors have read and agreed to the published version of the manuscript.

Peer review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/febs.17346.

Data availability statement

The data that support the findings of this study are available in this published article and in the Supporting Information. Protein models are available from the corresponding author upon reasonable request.

References

1 Bose H & Satyanarayana T (2017) Microbial carbonic anhydrases in biomimetic carbon sequestration for mitigating global warming: prospects and perspectives. *Front Microbiol* **8**, 1615.

2 DiMario RJ, Machingura MC, Waldrop GL & Moroney JV (2018) The many types of carbonic anhydrases in photosynthetic organisms. *Plant Sci* **268**, 11–17.

3 Kikutani S, Nakajima K, Nagasato C, Tsuji Y, Miyatake A & Matsuda Y (2016) Thylakoid luminal θ-carbonic anhydrase critical for growth and photosynthesis in the marine diatom *Phaeodactylum tricornutum. Proc Natl Acad Sci U S A* **113**, 9828–9833.

4 Tan SI, Han YL, Yu YJ, Chiu CY, Chang YK, Ouyang S, Fan KC, Lo KH & Ng IS (2018) Efficient carbon dioxide sequestration by using recombinant carbonic anhydrase. *Process Biochem* **73**, 38–46.

5 Capasso C & Supuran CT (2015) An overview of the alpha-, beta-and gamma-carbonic anhydrases from bacteria: can bacterial carbonic anhydrases shed new light on evolution of bacteria? *J Enzyme Inhib Med Chem* **30**, 325–332.

6 de Lourdes MM, Pérez D, García MT & Mellado E (2013) Halophilic bacteria as a source of novel hydrolytic enzymes. *Life* **3**, 38–51.

7 Marhuenda-Egea FC & Bonete MJ (2002) Extreme halophilic enzymes in organic solvents. *Curr Opin Biotechnol* **13**, 385–389.

8 Qiu J, Han R & Wang C (2021) Microbial halophilic lipases: a review. *J Basic Microbiol* **61**, 594–602.

9 Kambourova M (2018) Thermostable enzymes and polysaccharides produced by thermophilic bacteria isolated from Bulgarian hot springs. *Eng Life Sci* **18**, 758–767.

10 Mohammad BT, Al Daghistani HI, Jaouani A, Abdel-Latif S & Kennes C (2017) 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.

11 Takano Y, Edazawa Y, Kobayashi K, Urabe T & Marumo K (2005) Evidence of sub-vent biosphere: enzymatic activities in 308°C deep-sea hydrothermal systems at Suiyo seamount, IzuBonin arc, Western Pacific Ocean. *Earth Planet Sci Lett* **229**, 193–203.

12 Seo M-J, Lee B-S, Pyun Y-R & Park H (2011) Isolation and characterization of N-acylhomoserine lactonase from the thermophilic bacterium, Geobacillus caldoxylosilyticus YS-8. *Biosci Biotechnol Biochem* **75**, 1789–1795.

13 De Simone G, Monti SM, Alterio V, Buonanno M, De Luca V, Rossi M, Carginale V, Supuran CT, Capasso C & Di Fiore A (2015) Crystal structure of the most catalytically effective carbonic anhydrase enzyme known, SazCA from the thermophilic bacterium *Sulfurihydrogenibium azorense. Bioorg Med Chem Lett* **25**, 2002–2006.

14 Aguiar P, Beveridge T & Reysenbach A-L (2004) *Sulfurihydrogenibium azorense*, sp. nov., a thermophilic hydrogen-oxidizing microaerophile from terrestrial hot springs in the Azores. *Int J Syst Evol Microbiol* **54**, 33–39.

15 Capasso C, De Luca V, Carginale V, Cannio R & Rossi M (2012) Biochemical properties of a novel and highly thermostable bacterial α-carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* YO3AOP1. *J Enzyme Inhib Med Chem* **27**, 892–897.

16 Di Fiore A, Capasso C, De Luca V, Monti SM, Carginale V, Supuran CT, Scozzafava A, Pedone C, Rossi M & De Simone G (2013) X-ray structure of the first extremo-α-carbonic anhydrase', a dimeric enzyme from the thermophilic bacterium *Sulfurihydrogenibium yellowstonense* YO3AOP1. *Acta Crystallogr D* **69**, 1150–1159.

17 Vetriani C, Speck MD, Ellor SV, Lutz RA & Starovoytov V (2004) *Thermovibrio ammonificans* sp. nov., a thermophilic, chemolithotrophic, nitrate-ammonifying bacterium from deep-sea hydrothermal vents. *Int J Syst Evol Microbiol* **54**, 175–181.

18 James P, Isupov MN, Sayer C, Saneei V, Berg S, Lioliou M, Kotlar HK & Littlechild JA (2014) The structure of a tetrameric α-carbonic anhydrase from *Thermovibrio ammonificans* reveals a core formed around intermolecular disulfides that contribute to its thermostability. *Acta Crystallogr D* **70**, 2607–2618.

19 Fredslund F, Borchert MS, Poulsen JC, Mortensen SB, Perner M, Streit WR & Leggio LL (2018) Structure of a hyperthermostable carbonic anhydrase identified from an active hydrothermal vent chimney. *Enzyme Microb Technol* **114**, 48–54.

20 Kim S, Sung J, Yeon J, Choi SH & Jin MS (2019) Crystal structure of a highly thermostable α-carbonic anhydrase from *Persephonella marina* EX-H1. *Mol Cells* **42**, 460–469.

21 Götz D, Banta A, Beveridge TJ, Rushdi AI, Simoneit BR & Reysenbach AL (2002) *Persephonella marina* gen. Nov., sp. nov. and *Persephonella guaymasensis* sp. nov., two novel, thermophilic, hydrogen-oxidizing microaerophiles from deep-sea hydrothermal vents. *Int J Syst Evol Microbiol* **52**, 1349–1359.

22 Silverman DN, Lindskog S, Wynns GC, Norman AB & Laipis PJ (1993) Rate-equilibria relationships in intramolecular proton transfer in human carbonic anhydrase III. *Biochemistry* **32**, 10757–10762.

23 Silverman DN & Lindskog S (1988) The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water. *Acc Chem Res* **21**, 30–36.

24 Sridharan U, Willard BB, Gustafsson TN, Högbom M, Stålhammar H, Meuller J, Sandalova T, Murshudov GN, Lindskog S, Svärd SG *et al.* (2021) Structural and functional characterization of a putative carbonic anhydrase from *Geobacillus kaustophilus* reveals its cambialistic function. *Biochem Biophys Res Commun* **547**, 96–101.

25 Tripp BC, Bell CB 3rd, Cruz F, Krebs C & Ferry JG (2004) A role for iron in an ancient carbonic anhydrase. *J Biol Chem* **279**, 6683–6687.

26 Tu C, Silverman DN, Forsman C, Jonsson BH & Lindskog S (1989) Role of histidine 64 in the catalytic mechanism of human carbonic anhydrase II studied with a site-specific mutant. *Biochemistry* **28**, 7913–7918.

27 Sanyanga TA, Nizami B & Tastan Bishop Ö (2019) Mechanism of action of non-synonymous single nucleotide variations associated with α-carbonic anhydrase II deficiency. *Molecules* **24**, 3987.

28 An H, Tu C, Ren K, Laipis PJ & Silverman DN (2002) Proton transfer within the active-site cavity of carbonic anhydrase III. *Biochim Biophys Acta Proteins Proteomics* **1599**, 21–27.

29 Boone CD, Pinard M, McKenna R & Silverman D (2014) Catalytic mechanism of alpha-class carbonic anhydrases: CO_2 hydration and proton transfer. In Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications. Subcellular Biochemistry (Frost SC & McKenna R, eds), Vol. **75**, pp. 31–52. Springer, Dordrecht.

30 Mikulski RL & Silverman DN (2010) Proton transfer in catalysis and the role of proton shuttles in carbonic anhydrase. *Biochim Biophys Acta Proteins Proteomics* **1804**, 422–426.

31 Smith KS & Ferry JG (2000) Prokaryotic carbonic anhydrases. *FEMS Microbiol Rev* **24**, 335–366.

32 Manyumwa CV, Emameh RZ & Tastan Bishop Ö (2020) Alpha-carbonic anhydrases from hydrothermal vent sources as potential carbon dioxide sequestration agents: in silico sequence, structure and dynamics analyses. *Int J Mol Sci* **21**, 8066.

33 Nakagawa S, Takai K, Horikoshi K & Sako Y (2003) *Persephonella hydrogeniphila* sp. nov., a novel thermophilic, hydrogen-oxidizing bacterium from a deep-sea hydrothermal vent chimney. *Int J Syst Evol Microbiol* **53**, 863–869.

34 François DX, Barbeyron T, L'Haridon S, Corre E, Cambon-Bonavita M-A & Godfroy A (2021) *Persephonella atlantica* sp. nov.: how to adapt to physico-chemical gradients in high temperature hydrothermal habitats. *Syst Appl Microbiol* **44**, 126176.

35 Reysenbach AL, Huntemann M, Han J, Chen A, Kyrpides N, Mavromatis K, Markowitz V, Palaniappan K, Ivanova N, Schaumberg A *et al.* (2015) *Persephonella* sp. KM09-Lau-8 strain KM09 Lau8, whole genome shotgun sequencing project. Genbank: NZ_BBCJ00000000.

36 Elder I, Fisher Z, Laipis PJ, Tu C, McKenna R & Silverman DN (2007) Structural and kinetic analysis of proton shuttle residues in the active site of human carbonic anhydrase III. *Proteins* **68**, 337–343.

37 Jewell DA, Tu C, Paranawithana SR, Tanhauser SM, LoGrasso PV, Laipis PJ & Silverman DN (1991) Enhancement of the catalytic properties of human carbonic anhydrase III bysite-directed mutagenesis. *Biochemistry* **30**, 1484–1490.

38 Fisher Z, Hernandez Prada JA, Tu C, Duda D, Yoshioka C, An H, Govindasamy L, Silverman DN & McKenna R (2005) Structural and kinetic characterization of active-site histidine as a proton shuttle in catalysis by human carbonic anhydrase II. *Biochemistry* **44**, 1097–1105.

39 Amaro M, Kubiak-Ossowska K, Birch DJ & Rolinski OJ (2013) Initial stages of beta-amyloid Aβ_{1-40} and Aβ_{1-42} oligomerization observed using fluorescence decay and molecular dynamics analyses of tyrosine. *Methods Appl Fluoresc* **1**, 015006.

40 Mancini O, Rolinski OJ, Kubiak-Ossowska K & Mulheran PA (2018) Tyrosine rotamer states in beta amyloid: signatures of aggregation and fibrillation. *ACS Omega* **3**, 16046–16056.

41 Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Federhen S *et al.* (2010) Database resources of the national center for biotechnology information. *Nucleic Acids Res* **39**, D38–D51.

42 Notredame C, Higgins DG & Heringa J (2000) T-coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205–217.

43 Gibson DG, Benders GA, Axelrod KC, Zaveri J, Algire MA, Moodie M, Montague MG, Venter JC, Smith HO & Hutchison CA 3rd (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343–345.

44 Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.

45 Wilbur KM & Anderson NG (1948) Electrometric and colorimetric determination of carbonic anhydrase. *J Biol Chem* **176**, 147–154.

46 Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296–W303.

47 Eisenberg D, Lüthy R & Bowie JU (1997) VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol* **277**, 396–404.

48 Gordon JC, Myers JB, Folta T, Shoja V, Heath LS & Onufriev A (2005) H++: a server for estimating pKa values and adding missing hydrogens to macromolecules. *Nucleic Acids Res* **33**, W368–W371.

49 Case DA, Aktulga HM, Belfon K, Ben-Shalom IY, Berryman JT, Brozell SR, Cerutti DS, Cheatham TE III, Cisneros GA, Cruzeiro VWD *et al.* (2022) AmberTools22. University of California, San Francisco, CA.

50 Case D, Cerutti DS, Cheatham TE III, Darden TA, Duke RE, Giese TJ, Gohlke H, Goetz AW, Greene D, Homeyer N *et al.* (2017) AMBER 2017. University of California, San Francisco, CA.

51 Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B & Lindahl E (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25.

52 Humphrey W, Dalke A & Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* **14**, 33–38.

53 DeLano WL (2002) PyMOL: An open-source molecular graphics tool. *CCP4 Newsl Protein Crystallogr* **40**, 82–92.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. SDS/PAGE of CAs.

Fig. S2. Root mean square deviation (RMSD) of active sites.

Fig. S3. Active site structure and corresponding coordinating atom bond distances from the Zn^{2+} during MD simulations of PhyCA WT.

Table S1. Protein sequence of PhyCA, PaCA and PlauCA.

Table S2. Sequence of primers used for recombinant CA construction.

Table S3. Sequence of primers used for PhyCA site-directed mutation.