BIOLOGICAL CRYSTALLOGRAPHY
ISSN 1399-0047

Structural and biophysical characterization of the α-carbonic anhydrase from the gammaproteobacterium *Thiomicrospira crunogena* XCL-2: insights into engineering thermostable enzymes for CO_2 sequestration

Received 11 February 2015
Accepted 24 June 2015

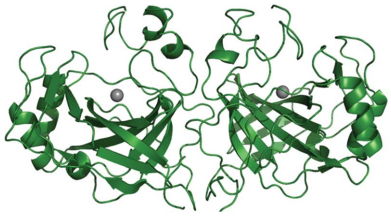
Edited by J. L. Martin, University of Queensland, Australia

‡ These authors contributed equally to this work.

Keywords: *Thiomicrospira crunogena* XCL-2; α-carbonic anhydrase; protein thermostability; biocatalyst; extremophile; CO_2 sequestration.

PDB reference: α-carbonic anhydrase, 4xz5

Supporting information: this article has supporting information at journals.iucr.org/d



Natalia A. Díaz-Torres,[a]‡ Brian P. Mahon,[a]‡ Christopher D. Boone,[a] Melissa A. Pinard,[a] Chingkuang Tu,[a] Robert Ng,[a] Mavis Agbandje-McKenna,[a] David Silverman,[b] Kathleen Scott[c] and Robert McKenna[a]*

[a]Department of Biochemistry and Molecular Biology, University of Florida College of Medicine, Gainesville, FL 32610, USA, [b]Department of Pharmacology and Therapeutics, University of Florida College of Medicine, Gainesville, FL 32610, USA, and [c]Department of Integrative Biology, University of South Florida, Tampa, FL 33620, USA. *Correspondence e-mail: rmckenna@ufl.edu

Biocatalytic CO_2 sequestration to reduce greenhouse-gas emissions from industrial processes is an active area of research. Carbonic anhydrases (CAs) are attractive enzymes for this process. However, the most active CAs display limited thermal and pH stability, making them less than ideal. As a result, there is an ongoing effort to engineer and/or find a thermostable CA to fulfill these needs. Here, the kinetic and thermal characterization is presented of an α-CA recently discovered in the mesophilic hydrothermal vent-isolate extremophile *Thiomicrospira crunogena* XCL-2 (TcruCA), which has a significantly higher thermostability compared with human CA II (melting temperature of 71.9°C *versus* 59.5°C, respectively) but with a tenfold decrease in the catalytic efficiency. The X-ray crystallographic structure of the dimeric TcruCA shows that it has a highly conserved yet compact structure compared with other α-CAs. In addition, TcruCA contains an intramolecular disulfide bond that stabilizes the enzyme. These features are thought to contribute significantly to the thermostability and pH stability of the enzyme and may be exploited to engineer α-CAs for applications in industrial CO_2 sequestration.

1. Introduction

Atmospheric concentrations of greenhouse gases (GHG) such as carbon dioxide (CO_2), chlorofluorocarbons, methane and nitrous oxide have been rising considerably owing to human-induced processes (Hansen *et al.*, 2000). One of the most abundant of the GHG is CO_2, the atmospheric accumulation of which has been correlated with a rise in global temperatures (Boone, Gill *et al.*, 2013). The burning of fossil fuels, such as coal, oil and natural gas, has sharply increased the concentration of atmospheric CO_2 and has been correlated with increased global temperatures over the past century (Intergovernmental Panel on Climate Change, 2005; Canadell *et al.*, 2007). This has promoted an extensive effort to limit CO_2 production in industrial processes to slow the rate of global climate change. However, the removal of CO_2 from industrial flue gases requires harsh chemical processes and extreme temperatures and is often cost-inefficient (Intergovernmental Panel on Climate Change, 2005; Pierre, 2012). The need for a more efficient CO_2-sequestration process has promoted developments in enzymatic CO_2 sequestration (Boone, Gill *et al.*, 2013). The use of carbonic anhydrase (CA) for CO_2

sequestration has shown promise owing to its efficiency in catalyzing CO_2 removal and its ability to be produced in large quantities (Favre *et al.*, 2009; Boron, 2010; Pierre, 2012; Boone, Habibzadegan, Gill *et al.*, 2013). However, CA must maintain its characteristic high catalytic efficiency in extreme conditions, such as high temperature, pressure and extreme pHs, in order to be a viable CO_2-sequestration agent (Savile & Lalonde, 2011; Boone, Habibzadegan, Gill *et al.*, 2013). To this end, there has been a large effort to characterize CAs from organisms that thrive in extreme environments (Alvizo *et al.*, 2014) and to engineer thermostable and pH-stable variants (Mirjafari *et al.*, 2007; Boron, 2010; Yu *et al.*, 2012; Boone, Habibzadegan, Tu *et al.*, 2013).

CAs are mainly zinc metalloenzymes that catalyze the interconversion of CO_2 and water to bicarbonate and a proton (Chegwidden *et al.*, 2000; Duda & McKenna, 2006; Domsic *et al.*, 2008). The general catalysis of CA is a metal-hydroxide 'ping-pong' mechanism composed of two independent steps (equations 1 and 2; where E is the enzyme and M^{2+} is a bivalent ion, typically Zn^{2+}),

$$\mathrm{E}M^{2+}\text{-}\mathrm{OH}^- + \mathrm{CO_2} \Leftrightarrow \mathrm{E}M^{2+}\text{-}\mathrm{HCO_3^-} \overset{\mathrm{H_2O}}{\longleftrightarrow} \mathrm{E}M^{2+}\text{-}\mathrm{H_2O} + \mathrm{HCO_3^-}, \quad (1)$$

$$\mathrm{E}M^{2+}\text{-}\mathrm{H_2O} + \mathrm{B} \Leftrightarrow \mathrm{E}M^{2+}\text{-}\mathrm{OH}^- + \mathrm{BH}^+. \quad (2)$$

The first step of catalysis (1) is initiated by nucleophilic attack on the carbon of CO_2 by the metal-bound hydroxide (EM^{2+}-OH^-) to yield bicarbonate (HCO_3^-), which is subsequently displaced by a water molecule (Lindskog, 1997; Lindskog & Silverman, 2000). The second step (2) is the removal of a proton from the now metal-bound water (EM^{2+}-H_2O) *via* an ordered water network and a residue acting as a weak base (B), which is typically a His at the opening of the active site, thus regenerating EM^{2+}-OH^- (Lindskog & Silverman, 2000; Fisher *et al.*, 2007). There are five known evolutionarily distinct classes of CAs: α, β, γ, δ and ζ (Hewett-Emmett & Tashian, 1996; Tripp *et al.*, 2001). The α-CAs are among the most efficient enzymes known, with human CA isoform II (hCA II) in particular displaying a turnover rate approaching the capture constant ($k_{cat}/K_m = 1.2 \times 10^8\ M^{-1}\ s^{-1}$). Therefore, engineering a thermostable hCA II variant or isolating thermostable isoforms from extremophilic organisms could produce a feasible agent for CO_2 sequestration.

The hydrothermal vent-dwelling chemolithoautotroph *Thiomicrospira crunogena* XCL-2 expresses an α-CA as well as two β-class CAs and one γ-class CA (Scott *et al.*, 2006; Dobrinski *et al.*, 2010). The deep-sea hydrothermal vents where *T. crunogena* is found are characterized by a harsh environment that constantly fluctuates from dilute hydrothermal fluid (2–35°C, anoxic, highly reduced, a CO_2 concentration of $\geq$1 m*M* and a pH of 5–8) to bottom water (2°C, normoxic, a CO_2 concentration of $\sim$0.02 m*M* and a pH of $\sim$7), temporally and spatially limiting the availability of nutrients to these organisms (Goffredi *et al.*, 1997; Shively *et al.*, 1998). *T. crunogena* is a sulfur-oxidizing gammaproteobacterium that is cosmopolitan; strains of this species and its congeners are inevitable isolates from deep-sea hydrothermal vents worldwide. It is equipped with a carbon-concentrating mechanism (CCM) that allows it to grow rapidly during bicarbonate and CO_2 scarcity, thus allowing the organism to survive. Previous studies have shown that CO_2 and bicarbonate transport plays an active role in CCM (Dobrinski *et al.*, 2010). Owing to this, it is presumed that the *T. crunogena* XCL-2 α-CA (TcruCA) plays a significant role in this process and has significant thermostability and pH stability in order to remain active in the temporally thermally heterogeneous environment.

We investigated the biochemical and biophysical properties of TcruCA compared with the extensively studied hCA II and in addition compare the structural properties of TcruCA with those of other α-CAs, some of which exhibit high thermostability, in order to speculate on the use of TcruCA as a biocatalyst for industrial CO_2 sequestration. We further suggest ways in which the biochemical structure of TcruCA can be utilized to engineer a novel α-CA variant that can remain highly active at extreme temperatures and pH. The results from this study will increase our repertoire of potential biocatalytic agents that can be used for industrial CO_2 sequestration.

2. Materials and methods

2.1. Protein expression and purification

TcruCA was expressed in recombinant *Escherichia coli* BL21(DE3) cells as previously reported (Forsman *et al.*, 1988; Dobrinski *et al.*, 2010). In brief, *E. coli* cells containing the plasmid encoding TcruCA were grown in Luria broth supplemented with 50 µg ml^{-1} kanamycin to an OD_{600} of 0.6–1.0, at which point expression of TcruCA was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (to a final concentration of 0.1 mg ml^{-1}) for $\sim$4 h at 37°C. The cells containing TcruCA were harvested and enzymatically lysed. TcruCA was purified by affinity chromatography using an agarose resin coupled to the inhibitor *p*-(aminomethyl)-benzenesulfonamide (*p*-AMBS; Sigma). The protein was eluted with sodium azide, buffer-exchanged against 50 m*M* Tris–HCl pH 7.8, 100 m*M* NaCl and concentrated using centrifugation. The final protein concentration was determined by UV–Vis spectroscopy at 280 nm using an extinction coefficient of 42 985 M^{-1} cm^{-1} (calculated from the amino-acid sequence; Gill & von Hippel, 1989). The purity was estimated by SDS–PAGE and Coomassie Blue staining (Díaz-Torres *et al.*, 2012).

2.2. Size-exclusion chromatography

The molecular weight and oligomeric state of TcruCA were determined using size-exclusion chromatography (SEC) *via* a GE Healthcare ÄKTA Fast Protein Liquid Chromatography (FPLC) purification system equipped with a prepacked HiPrep 16/60 Sephacryl S-200 gel-filtration column (exclusion range 5–250 kDa; GE Healthcare Biosciences AB, Sweden). The column was equilibrated with 50 m*M* Tris–HCl pH 7.0,

100 m*M* NaCl buffer. The protein was eluted at a flow rate of 0.1 ml min^{-1}, collecting 1 ml fractions. Protein fractions were detected by absorbance at 280 nm, and data acquisition and processing were performed using the *UNICORN* FPLC software. The Sephacryl S-200 column was calibrated using Gel Filtration Standard molecular-weight markers from Bio-Rad (Supplementary Fig. S2; catalog No. 151-1901) following the instructions provided. The molecular weight of monomeric TcruCA was estimated to be 33 kDa by SDS–PAGE analysis.

2.3. Crystallization and X-ray data collection

Crystals of TcruCA were obtained using sitting-drop vapour diffusion as described by Diáz Torres *et al.* (2012), with a precipitant solution consisting of 2%(*v*/*v*) Tacsimate pH 4.0, 0.1 *M* sodium acetate trihydrate pH 4.6, 16%(*w*/*v*) PEG 3350 at 17°C. Small crystals (0.05 × 0.05 × 0.01 mm) were observed after 10 d. The crystals were cryoprotected with 20% glycerol prior to vitrification in liquid nitrogen.

X-ray diffraction data were collected on beamline F1 at the Cornell High Energy Synchrotron Source (CHESS) using a wavelength of 0.9177 Å. The data sets were collected using an ADSC Quantum 270 CCD detector at a crystal-to-detector distance of 300 mm with a 1° oscillation angle and an exposure time of 1 min per image. A total of 360 images were collected. The data were indexed, integrated and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). However, radiation damage was observed within the later collected images, and so the data set used for the structure determination was truncated to only the first 180 images. The data were scaled in the monoclinic space group *C*2 (unit-cell parameters $a = 127.1$, $b = 102.2$, $c = 105.0$ Å, $\beta = 127.3°$) to 2.6 Å resolution with a completeness of 99.8% and an $R_{\rm merge}$ of 10.0% (Table 1).

2.4. Structure determination and refinement

The structure of TcruCA was determined using molecular replacement (MR) with the high-resolution structure of hCA II (PDB entry 3ks3; Avvaru *et al.*, 2010) as a search model. MR solutions were calculated using *PHENIX* (McCoy *et al.*, 2007; Adams *et al.*, 2011). The starting phases of the TcruCA model yielded a unique solution comprising of four molecules in the asymmetric unit (Supplementary Fig. S1). Each chain contained 230 amino acids, corresponding to residues 75–304 of full-length TcruCA. Refinement was also completed using *PHENIX* (Adams *et al.*, 2011). Each refinement was performed with 5% of the unique reflections excluded for the calculation of $R_{\rm free}$ (Brünger, 1992). Manual refitting of the model between each refinement was performed using *Coot* (Emsley & Cowtan, 2004). Superimpositions of the C^{α} atoms of chain *A* onto chains *B*, *C* and *D* (r.m.s.d. values of 0.23, 0.23 and 0.24 Å, respectively) indicated there were no major structural perturbations between the main chains of each monomer. Therefore, noncrystallographic symmetry (NCS) operators were employed for the remainder of the refinement. The final model of TcruCA was refined to an $R_{\rm cryst}$ of 17.1% and an $R_{\rm free}$ of 24.0% (Table 1). The model geometries and statistics were assessed by *PROCHECK* (Laskowski *et al.*, 1993). All figures were produced using *PyMOL* (Schrödinger).

Table 1
X-ray crystallographic statistics for TcruCA.

Values in parentheses are for the highest resolution shell.

PDB code	4xz5
Space group	*C*2
Unit-cell parameters (Å, °)	$a = 127.1$, $b = 102.2$, $c = 105.0$, $\beta = 127.3$
Resolution (Å)	20.0–2.60 (2.69–2.60)
Total reflections	120808
Unique reflections	32859
$R_{\rm merge}$† (%)	10.0 (40.5)
$CC_{1/2}$	0.95 (0.85)
$\langle I/\sigma(I)\rangle$	9.2 (2.5)
Multiplicity	3.7 (3.5)
Completeness (%)	99.8 (99.1)
$R_{\rm cryst}$‡ (%)	17.1 (21.5)
$R_{\rm free}$§ (%)	24.0 (30.6)
$V_{\rm M}$ (Å^3 Da^{-1})	2.04
Residue Nos.	75–304
No. of protein atoms	7251 [all chains]
No. of water molecules	37
No. of Zn atoms	4
No. of bicarbonate molecules	4
Ramachandran statistics (%)	
Favored	97.7
Allowed	1.86
Generously allowed	0.44
Average *B* factors (Å^2)	
Main chain	32.6
Side chain	37.6
Water	24.3
Bicarbonate	30.0
R.m.s.d.	
Bond lengths (Å)	0.009
Angles (°)	1.3

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl) \times 100$. ‡ $R_{\rm cryst} = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}| \times 100$. § $R_{\rm free}$ is calculated in the same way as $R_{\rm cryst}$ except for using data omitted from refinement (5% of reflections for all data sets).

2.5. ^{18}O-exchange kinetic analysis

The kinetic rates of TcruCA were obtained by measurement of the depletion of ^{18}O from species of CO_2 at chemical equilibrium by membrane-inlet mass spectrometry as described by Tu *et al.* (1989, 2009). This occurs *via* the continuous measurement of the various isotopic species of CO_2 provided by CO_2 diffusing across a dissolved gas-permeable membrane; the membrane is submerged in the reaction solution and connected by glass tubing to a mass spectrometer (Extrel EXM-200; Tu *et al.*, 2009). The catalyzed and uncatalyzed exchange of ^{18}O between CO_2 and water at chemical equilibrium were measured in the absence of buffer at a total substrate concentration of 25 m*M*. The reaction solution was maintained at 25°C, and the ionic strength of the solution was normalized at 0.2 *M* by the addition of Na_2SO_4. CA catalysis of isotope-labeled substrate is described by the following two-step process: the first stage of catalysis is the dehydration of ^{18}O-labeled HCO_3^- (3). During the second catalytic stage, the zinc-bound ^{18}O-labeled hydroxide is protonated, forming $H_2^{18}O$, which is then released into solution (4),

$$HCOO^{18}O^- + EZnH_2O \overset{-H_2O}{\longleftrightarrow} EZnHCOO^{18}O^- \longleftrightarrow COO + EZn^{18}OH^-, \quad (3)$$

$$H^+His64\text{-}EZn^{18}OH \longleftrightarrow His64\text{-}EZnH_2^{18}O \overset{H_2O}{\longleftrightarrow} His64\text{-}EZnH_2O + H_2^{18}O. \quad (4)$$

The catalytic rate for the interconversion of CO_2 and HCO_3^- at chemical equilibrium is described by R_1 (5). k_{cat}^{ex} is the rate constant for maximal interconversion of substrate to product, which in this case is CO_2 and HCO_3^-. $K_{eff}^{CO_2}$ is the effective binding constant of substrate (either CO_2 or HCO_3^-) to the enzyme (CA). Therefore, the ratio $k_{cat}^{ex}/K_{eff}^{CO_2}$ is equivalent to the catalytic efficiency (k_{cat}/K_m) of hydration obtained under steady-state conditions (Simonsson *et al.*, 1979),

$$\frac{R_1}{[E]} = \frac{k_{cat}^{ex}[CO_2]}{K_{eff}^{CO_2} + [CO_2]}. \quad (5)$$

The rate of proton transfer, which in this case is obtained by measuring the diffusion of ^{18}O-labeled water from the enzyme to the solvent, is given by R_{H_2O}. R_{H_2O} is dependent on the donation of protons to the ^{18}O-labeled zinc-bound hydroxide by His64 as a second independent step of the catalysis by hCA II, or its equivalent in other isozymes, as shown in (4) (Simonsson *et al.*, 1979; Tu *et al.*, 1989). In (6), k_B is the rate constant for proton transfer to the zinc-bound hydroxide and $(K_a)_{His64}$ and $(K_a)_{ZnH_2O}$ are the ionization constants of the proton donor His64 and the zinc-bound water molecule, respectively,

$$R_{H_2O}[E] = \frac{k_B}{\left[1 + \frac{(K_a)_{His64}}{[H^+]}\right]\left[1 + \frac{[H^+]}{(K_a)_{ZnH_2O}}\right]}. \quad (6)$$

(5) and (6) were fitted to the data by using nonlinear least-squares methods in *EnzFitter* (Biosoft).

2.6. Differential scanning calorimetry

Differential scanning calorimetry (DSC) experiments were performed to assess the thermostability of TcruCA using a VP-DSC microcalorimeter (Microcal Inc., North Hampton, Massachusetts, USA) with a cell volume of ∼0.6 ml. TcruCA

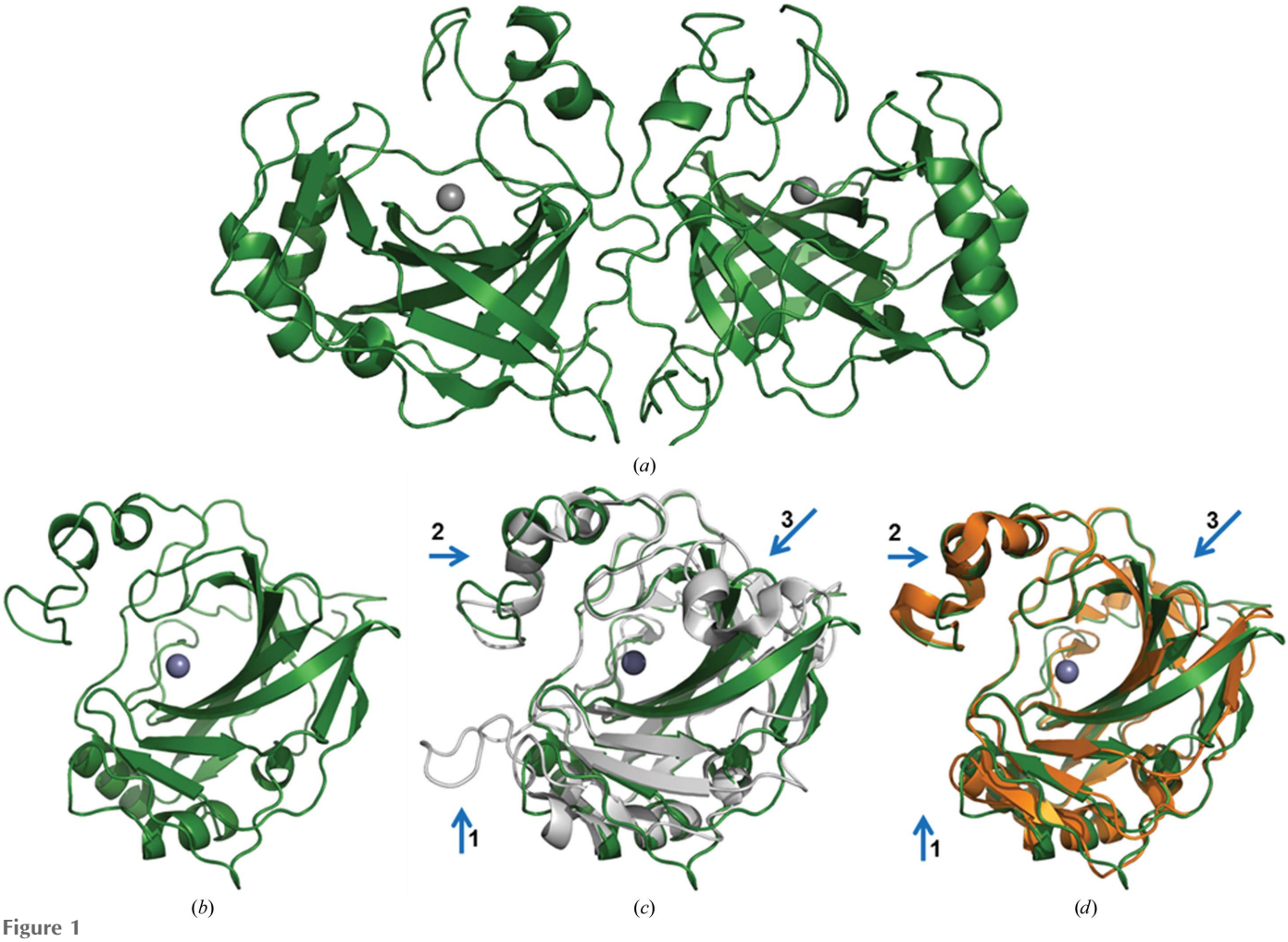


Figure 1
(*a*) Ribbon diagram of the TcruCA homodimer. Zinc is depicted as a gray sphere. (*b*) Overall structure of the TcruCA monomer (green), with zinc depicted as a gray sphere. (*c*) Overlay of the TcruCA monomer (green) with hCA II (gray). Surface regions pertaining to structural variation are labeled 1–3. (*d*) Overlay of the TcruCA monomer with SspCA (orange). The structural overlay was performed using *PyMol*.

samples were extensively buffer-exchanged into solutions with pH values from 4 to 9 increasing by one pH unit. Phosphate–citrate buffer was used for pH values 4–6 and 50 m*M* Tris–HCl, 100 m*M* NaCl buffer was used for pH values 7–9, all of which had a protein concentration of 30 µ*M*. The samples were degassed at 16°C for 20 min prior to data collection. DSC scans were collected from 30 to 90°C at a scan rate of 60°C h^{-1}. Scans for all the samples were performed in triplicate. A reference scan for each pH value, where buffer was placed in both the reference and the sample cells, was also performed in order to subtract the contributing heat capacity of the buffer solution. After subtracting the reference and adjusting the baseline for each scan, the calorimetric enthalpies of unfolding were calculated by integrating the area under the peaks in each thermogram. The thermograms were fitted to a non-two-state reversible unfolding model to obtain van't Hoff enthalpies of unfolding (ΔH_v; Bruylants *et al.*, 2005). The melting temperature (T_m) values of the TcruCA samples at different pH values and of hCA II were obtained from the midpoints of the thermograms.

3. Results and discussion

3.1. Overall structure of TcruCA

The TcruCA crystals exhibited *C*2 symmetry with four monomers in the asymmetric unit (Supplementary Fig. S1). Specifically, the asymmetric unit contained two dimers. Calculations of the Matthews coefficient with four monomers in the asymmetric unit resulted in a value of 2.04 Å^3 Da^{-1} with ∼40% solvent content, which correlates to what was observed in our structure solution (Table 1). Size-exclusion and DSC (data not shown) analysis indicate that TcruCA exists predominantly as a dimer in solution and, in addition to the crystallographic information, indicates the enzyme is a homodimer (Fig. 1*a*). A region of 52 amino acids at the N-terminus is not observed in the TcruCA structure. This region was previously predicted to correspond to a likely signal peptide in TcruCA (Díaz Torres *et al.*, 2012).

Table 2
Sequence and structural alignments of TcruCA.

α-CA	Sequence identity (%)	No. of residues aligned	C^α r.m.s.d. (Å)
hCA II	32.5	228	1.4
hCA VI	32.5	206	1.7
hCA IX	35.5	214	1.4
hCA XII	33.2	217	1.4
SspCA	41.0	227	1.3
Cr-αCA1	33.5	221	1.6
AoCA	29.5	207	1.6

The TcruCA monomer has the signature secondary structure typical of an α-CA fold, with helical and loop structures present towards the surface and a conical active-site cavity comprised of mostly β-structure (Fig. 1*b*). Overlay of TcruCA with hCA II shows there is structural conservation between them which is highly prevalent within the active sites (C^α r.m.s.d. of 1.4 Å). Based on sequence alignments of TcruCA (ExPasy *SIM* alignment software; Huang & Miller, 1991) with several human and microbial CAs, it is predicted that this conservation in active-site structure is present in several isoforms and it is summarized in Table 2, where hCA, SspCA, Cr-αCA1 and AoCA represent the α-CAs from human, the bacterium *Sulfurihydrogenibium yellowstonense* YO3AOPO1, *Chlamydomonas reinhardtii* and *Aspergillus oryzae*, respectively. The majority of structural variation between TcruCA and hCA II occurs on the enzyme surface and can be globally divided into three regions (Fig. 1*c*). TcruCA displays a more compact overall structure because of truncations of surface loops in region 1 and the absence of an α-helix in region 3 (Fig. 1*c*). Region 2 shows alterations in secondary structure most likely owing to amino-acid differences between hCA II and TcruCA (Fig. 1*c*). SspCA was selected for structural comparison as it is currently the most thermostable α-CA and is a very attractive target for development into a biocatalytic agent for industrial carbon sequestration (Capasso *et al.*, 2012). Comparisons of TcruCA with SspCA (Capasso *et al.*, 2012; Table 2; C^α r.m.s.d. = 1.3 Å) exhibited more conservation in surface regions relative to TcruCA, as SspCA also has a more compact structure compared with hCA II. Specifically, TcruCA and SspCA share the same loop deletions in region 1 and loss of secondary structure in region 3 (Fig. 1*d*). Furthermore, there appears to be greater

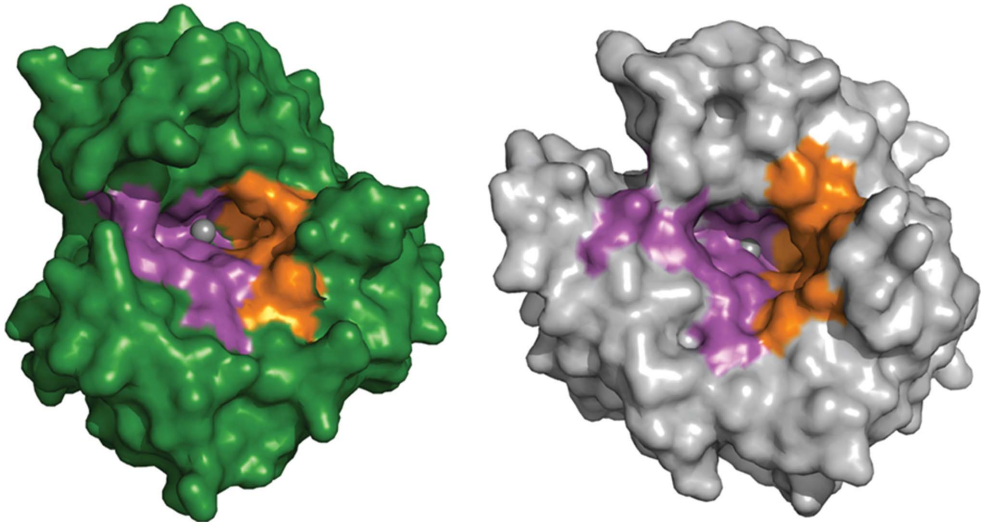


Figure 2
Active site of TcruCA (green) and hCA II (gray). Hydrophobic (orange) and hydrophilic (purple) regions of the TcruCA active site (*a*) are highlighted and show a similar arrangement compared with those of hCA II (*b*).

Table 3
Catalytic parameters of TcruCA and hCA II.

Parameter	TcruCA	hCA II†
$k_{cat}^{ex}/K_{eff}^{CO_2}$ (μM^{-1} s^{-1})	11.0 ± 0.1	120
k_B (μs^{-1})	0.30 ± 0.05	0.8
pK_{a,ZnH_2O}‡	6.6 ± 0.1	6.9
pK_{a,ZnH_2O}§	6.4 ± 0.2	6.8
$pK_{a,PSR}$§	6.4 ± 0.2	7.2

† Metcalf *et al.* (1996). Standard errors are no larger than 20%. ‡ Determined by ^{18}O-exchange from calculation of $k_{cat}^{ex}/K_{eff}^{CO_2}$ in the hydration of CO_2. § Determined from calculation of R_{H_2O}/[E].

structural overlap observed in region 2 (Fig. 1*d*). These results may indicate strong evolutionary linkages between the microbial CAs compared with those found in mammals.

The overall characteristics of the active site of TcruCA exhibits defined hydrophobic and hydrophilic pockets similar to all other α-CAs (Fig. 2*a*; Pinard *et al.*, 2015). This is especially apparent when compared with the active site of hCA II (Fig. 2*b*). Specific residues that define the active site of TcruCA, including the positions of the observed zinc-bound H_2O/OH^- and 'deep water' (DW), are outlined in Fig. 3(*a*). Comparison of these residues to those in hCA II shows that the majority are conserved, including the proton-shuttling residue His140 (64) (hCA II numbering is given in parentheses throughout this manuscript) and the zinc-coordinating residues His165 (94), His167 (96) and His184 (119) (Fig. 3*b*; Pinard *et al.*, 2015). In addition, we have observed a bicarbonate ion bound to the catalytic zinc in all chains. Interestingly, the only difference between the active-site residues occurs outwards from the catalytic zinc in the hydrophilic pocket at positions 141 (65), 143 (67) and 145 (69), and a single residue in the hydrophobic pocket at position 162 (91). A similar result is observed when comparing the active-site residues of TcruCA with SspCA, which shows conservation between each residue except for those in the hydrophobic pocket at positions 162 (91), 199 (146) and 194 (141) (Fig. 3*c*). However, despite the disparity between amino-acid identities between these positions, each maintains its primary residue property such that there is no replacement of hydrophobic by hydrophilic residues or *vice versa*, with the exception of position 162 (91) between TcruCA and SspCA (Leu and Lys in TcruCA and SspCA, respectively). The structural conservation observed between TcruCA and other CAs indicates it is a suitable 'template' to engineer a thermostable α-CA.

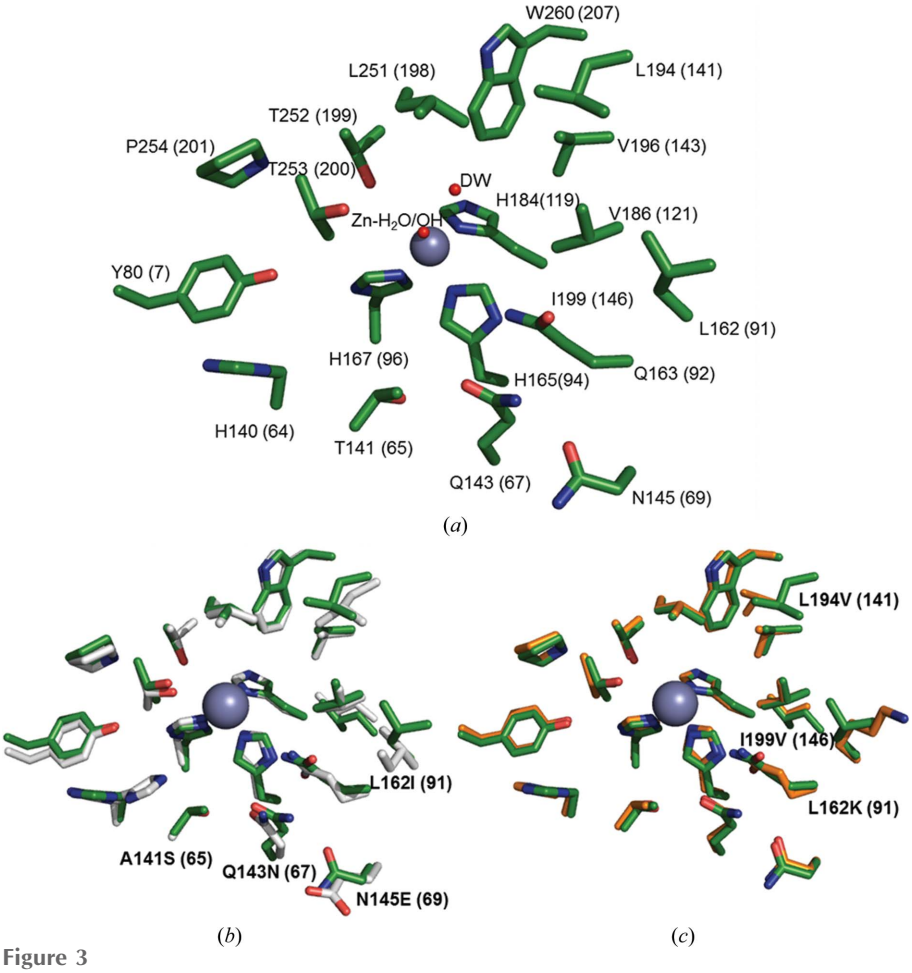


Figure 3
Active site of TcruCA (green). (*a*) Residues listed as important for TcruCA catalysis (shown as sticks) with zinc-bound solvent shown for clarity. Residues are numbered relative to their position in the full-length chain of TcruCA, with the corresponding positions in hCA II shown in parentheses. (*b*) Overlay of the active-site residues of TcruCA (green) and hCA II (gray). Residues that differ between TcruCA and hCA II are labeled, with the hCA II numbering given in parentheses. (*c*) Overlay of active-site residues of TcruCA and SspCA (orange). Residues that differ between TcruCA and SspCA are labeled, with the hCA II numbering given in parentheses.

3.2. TcruCA catalytic activity

The catalytic activity of TcruCA was assayed and compared with that of hCA II. The first rate obtained, R_1 (5), was used to calculate $k_{cat}^{ex}/K_{eff}^{CO_2}$, which is equivalent to the catalytic efficiency. The pH profile of TcruCA for this rate constant (Supplementary Fig. S3) was fitted to a single ionization event, giving a maximal k_{cat}/K_m of 11 ± 1 μM^{-1} s^{-1}, which is tenfold lower than the catalytic efficiency of hCA II (120 μM^{-1} s^{-1}). The kinetic pK_a determined from the pH profile of $k_{cat}^{ex}/K_{eff}^{CO_2}$, which is an estimated value of the pK_a of the zinc-bound water (6.6 ± 0.1), was lower by approximately one unit (Table 3), which would cause a decrease in the nucleophilicity of the zinc-bound hydroxide. The reduced catalytic efficiency observed in TcruCA could be owing to the displacement of one of the valines in the hydrophobic pocket that binds the substrate CO_2 (Fisher *et al.*, 2007). Val260 (207) is displaced ∼1.2 Å towards the CO_2-binding site, which could sterically inhibit the proper binding orientation for CO_2, thus redu-

cing the catalytic efficiency. Interestingly, there is also an approximately tenfold difference in catalytic efficiency between TcruCA and SspCA [k_{cat}/K_m = 11 μM^{-1} s^{-1} (Table 3) for TcruCA and 111 μM^{-1} s^{-1} for SspCA (Vullo *et al.*, 2012)]. This might be owing to a substitution at position 162 (91), where Leu in TcruCA is replaced by Lys in SspCA. However, on comparing the active sites of human isoforms which contain a Lys in the same position (hCA IV, VA, VB and VII) with those with a Leu in the same position (hCA IX) there is no observable relationship between their catalytic efficiencies and the characteristics of this residue alone (Pinard *et al.*, 2015).

Another possibility for the decreased k_{cat}/K_M is a loss of the region 1 loop (residues 230–240 in hCA II), which has been associated with a reduction in catalytic efficiency (Boone *et al.*, 2015). In hCA II this region contains four Glu residues, which may establish an internal buffering mechanism, similar to the predicted function of the proteoglycan domain of hCA IX (Alterio *et al.*, 2009), which may enhance substrate/product entry/exit from the active site. However, more data are needed to justify this observation.

The second rate obtained from the ^{18}O-exchange kinetic assays is R_{H_2O} (6), the rate of release of $H_2^{18}O$ from the active site. The pH profiles for R_{H_2O}/[E] have a characteristic bell-shaped curve for most of the pH range covered in the studies (Supplementary Fig. S3), which is attributed to the transfer of a proton from the proton-shuttling residue histidine to the zinc-bound hydroxide (4). The fit of (6) to the data yielded a rate constant for intramolecular proton transfer (k_B) of 0.30 ± 0.05 μs^{-1}. The rate constant k_B of TcruCA was noticeably closer to that of hCA II than k_{cat}/K_m (Table 3), suggesting that TcruCA is most likely to use the same proton-shuttling residue at His140 (64) for proton transfer. The pK_a of the zinc-bound water and the proton-shuttle residue in TcruCA are both 6.4 ± 0.1, which contributes slightly to a reduction in proton transfer compared with hCA II. It is most likely to be residues in the TcruCA hydrophilic pocket, Thr141 (65) and Gln143 (67), which are replaced by an Ala and Asn, respectively, in hCA II (Fig. 3*b*), that influence the decrease in proton transfer, as this would alter the ideal bond distances between these active-site residues and the ordered water network of the active site of the enzyme that are necessary for ideal proton transfer.

3.3. TcruCA thermostability

The thermostability of TcruCA was determined at pH 5 to pH 9 by DSC and compared with that of hCA II at pH 8. The melting temperature (T_m) of hCA II and the thermal inactivation temperature is 59.5 ± 0.5°C, and is observed as a single endothermic peak at the midpoint of the DSC curves representative of the main unfolding transition (Mikulski *et al.*, 2011). Unlike the scans collected for hCA II, the thermograms for TcruCA presented two transition peaks: one between 58 and 62°C and the other between 68 and 72°C, dependent on the pH of the study (Table 4). A reverse scan performed separately displayed only one peak around 59°C (data not shown), suggesting that the first transition is reversible, a result similar to that observed previously for NAD^+ synthetase (Yang *et al.*, 2004). This result, coupled with size-exclusion and X-ray crystallographic data, further implies that TcruCA is most likely to exist as a dimer in solution. Therefore, the first transition observed in the thermograms of TcruCA is presumed to be the dissociation of dimer complexes. After fitting the data to a non-two-state transition, the calorimetric enthalpies (ΔH), the van't Hoff enthalpies (ΔH_v) and the T_m were calculated more readily (Table 4). The melting temperatures of the two peaks were plotted as a function of pH separately (Supplementary Fig. S4) to determine which transition had a higher pH sensitivity in terms of thermostability. The T_m for the first transition increased by roughly 1–2°C for each pH interval, while the T_m for the second peak increased approximately 3° over the pH range from 5 to 9. The pH profile for the T_m of both transitions (more so for the first transition) demonstrated that there is a direct relationship between pH and temperature for TcruCA, as the conformational stability of the enzyme was enhanced as the pH was increased. It should be mentioned that it is not clear whether this observation is unique to TcruCA and what elevated pH results in a decrease in the thermostability of the enzyme. However, this does suggest a potential for the enzyme to maintain its activity for industrial CO_2-sequestration processes in solutions of pH >8.

The scans collected at pH 5–9 demonstrated that the T_m of the second transition (predicted monomeric state) increased from 68.1 to 71.9°C as the samples were switched to more basic buffers. Compared with hCA II at pH 8.0, our results show that TcruCA has a T_m of 71.9 ± 0.4°C, whereas hCA II has a T_m of 55.96 ± 0.05°C. The 15.94°C difference between the T_m values demonstrates that TcruCA has greater thermal conformational stability than hCA II. Furthermore, this relationship between pH and temperature is not observed for hCA II (data not shown), where the conformational stability of hCA II does not increase beyond a pH of 7, similar to that observed in bovine CA II (Matulis *et al.*, 2005). This might imply that TcruCA remains stable at pH values of >9. Thermostability has long been associated with structural and

Table 4
Thermodynamic parameters for the unfolding of TcruCA.

	First transition (dimer)			Second transition (monomer)		
pH	T_m (°C)	ΔH (kcal mol^{-1})	ΔH_v (kcal mol^{-1})	T_m (°C)	ΔH (kcal mol^{-1})	ΔH_v (kcal mol^{-1})
5.0	57.7 ± 0.7	130 ± 33	91 ± 17	68.1 ± 0.8	193 ± 34	75 ± 12
6.0	58.3 ± 0.4	85 ± 15	123 ± 20	69.5 ± 0.4	234 ± 18	69 ± 7
7.0	59.8 ± 0.5	211 ± 30	104 ± 15	69.9 ± 0.6	189 ± 30	102 ± 23
8.0	60.1 ± 0.3	129 ± 12	109 ± 11	71.9 ± 0.4	128 ± 13	103 ± 13
9.0	62.1 ± 0.4	242 ± 25	98 ± 11	71.0 ± 0.5	119 ± 24	132 ± 28
hCA II†	—	—	—	55.95 ± 0.05	170 ± 4	181 ± 5

† Values for hCA II are at pH 8.0.

conformational stability of proteins, along with amino-acid composition, hydrogen bonding and solvent accessibility (Vogt *et al.*, 1997; Pechkova *et al.*, 2007). Therefore, our results indicate that TcruCA is more thermodynamically stable and has greater conformational stability than hCA II.

Structural variations between hCA II and TcruCA could potentially impact the differences observed in thermostability. As mentioned previously, TcruCA exhibits a more compact structure compared with hCA II (Fig. 1*b*), which is predicted to enhance the folding stability of TcruCA. Interestingly, a similar compact structure is observed in SspCA, which has been shown to maintain catalytic activity up to 100°C (Capasso *et al.*, 2012).

In addition to its compact structure, it is predicted that the intermolecular disulfide bridge between Cys99 (21) and Cys256 (203) in TcruCA contributes significantly to the thermostability of the enzyme (Supplementary Fig. S5*a*). This disulfide bond bridges a helix at the N-terminus with a loop containing Thr252 (199), which is an essential residue for CO_2 hydration (Mehler *et al.*, 2002). Protein-folding studies of hCA II have demonstrated that the conformation of the (198–206) loop is strained at Pro(202) owing to the adoption of a *cis*

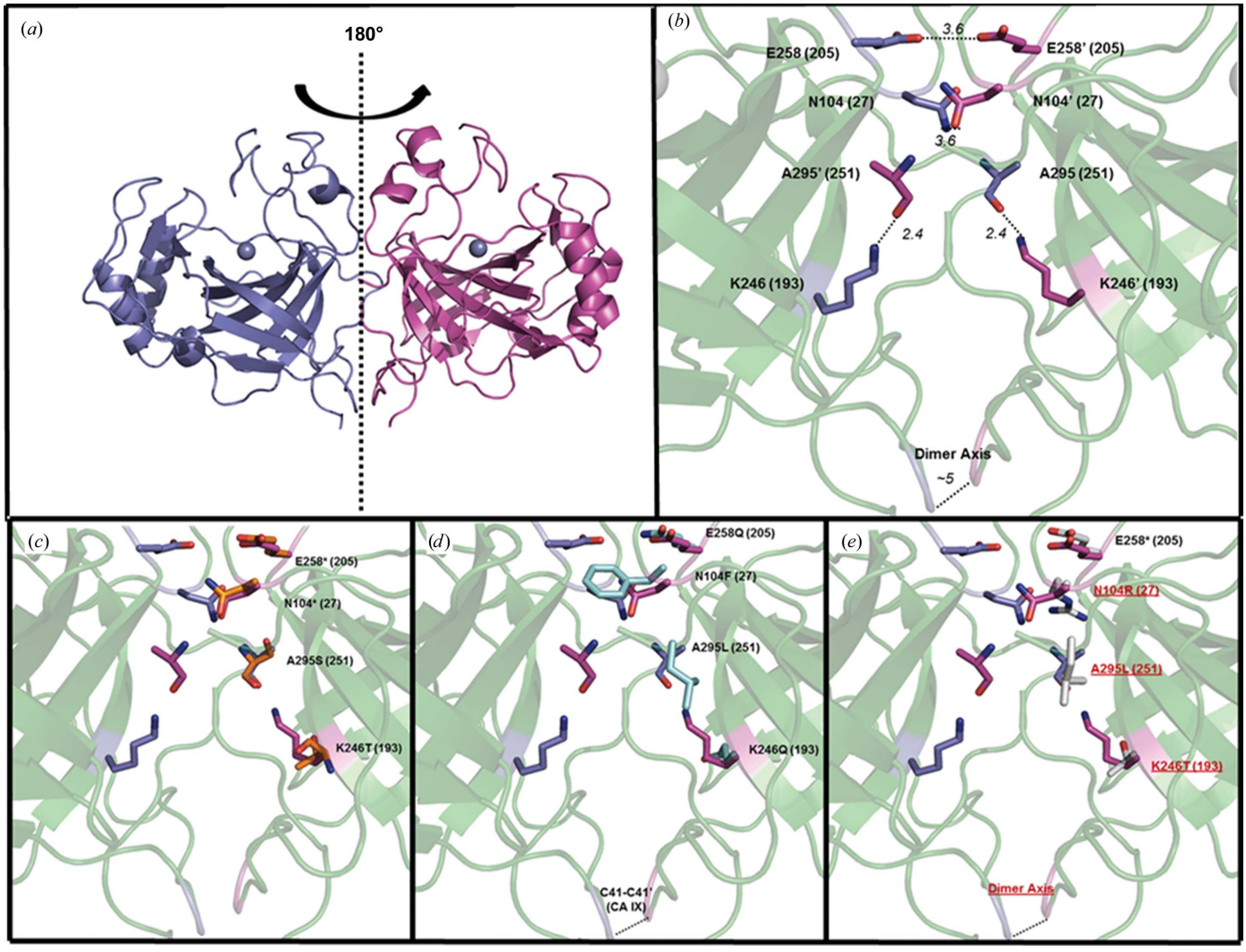


Figure 4
Dimer of TcruCA. (*a*) TcruCA homodimer with highlighted monomeric chains *A* (purple) and *B* (pink) shown as a ribbon diagram. The TcruCA dimer is shown as a 180° rotation between monomeric units. (*b*) Highlighted residues predicted to be responsible for stabilizing the dimer interface between TcruCA monomers. Residues are shown as sticks from chains *A* (purple) and *B* (pink), and are labeled according to their position in full-length TcruCA (hCA II residue positions are shown in parentheses). The dimer axis of TcruCA monomers is shown as a ribbon with distances in Å between C^{α} atoms from chains *A* (purple) and *B* (pink) shown. Possible hydrogen-bond interactions between residues are shown as dotted lines with distances in Å. (*c*) Overlay of the TcruCA (chains *A* and *B* in purple and pink, respectively) dimer interface with that of SspCA (orange; one chain only). Residues that vary between interfaces are labeled, with hCA II numbering given in parentheses. (*d*) Overlay of the TcruCA (chains *A* and *B* in purple and pink, respectively) dimer interface with that of hCA IX (cyan; one chain only). Varying residues between each interface are labeled, with hCA II numbering given in parentheses. Also highlighted is the presence of an intermolecule disulfide bond in hCA IX located in the dimer interface. Residue positions are given relative to hCA II numbering. (*e*) Overlay of the TcruCA (chains *A* and *B* in purple and pink, respectively) dimer interface with that of hCA II (gray; one chain only). Conserved residues (*) and those that vary between TcruCA and hCA II that can be exploited to engineer a dimeric hCA II variant are highlighted.

Table 5
Comparison of the TcruCA dimeric interface.

All data were calculated using *PDBePISA* (Krissinel & Henrick, 2007).

Dimeric α-CA	PDB code	Interface area† (Å^2)	No. of interfacing residues‡	ΔG_{solv}§ (kcal mol^{-1})	ΔG_{solv} *P*-value¶	No. of hydrogen bonds
TcruCA	4xz5	898	63 (31, 32)	−9.7	0.123	18
hCA VI	3fe4	869	54 (28, 26)	−9.8	0.552	11
hCA IX	3iai	802	54 (27, 27)	−9.0	0.227	2
hCA XII	1jcz	1130	72 (35, 37)	+1.1	0.783	17
SspCA	4g7a	1162	67 (35, 32)	−12.8	0.050	11
Cr-αCA1	3b1b	1493	84 (42, 42)	−17.7	0.068	21
AoCA	3q31	1812	93 (46, 47)	−18.4	0.189	23

† Calculated as the difference in the total accessible surface areas of isolated and interfacing structures divided by two. ‡ Values in parentheses correspond to the numbers of residues from each monomer. § Solvation free-energy gain upon formation of the interface. $\Delta G_{solv} < 0$ corresponds to positive protein affinity by hydrophobic interfaces (excludes contributions from salt bridges and hydrogen bonds). ¶ Measure of the likelihood of the interface forming as a result of crystal packing. Note that ΔG_{solv} *P*-values of >0.5 indicate that the interface is likely to be an artifact of crystal packing and not biological in nature.

Table 6
Predicted hydrogen bonds in dimer stabilization.

Monomer *A*	Monomer *B*	Distance (Å)
Asn104 O$^{\delta1}$	Ser248 O$^{\gamma}$	3.6
Ala115 O	Ser120 O$^{\gamma}$	3.9
Gly117 O	Thr119 N	3.1
Thr119 O$^{\gamma1}$	Asp304 N	3.0
Ser120 O$^{\gamma}$	Gly117 N	3.3
Glu258 O$^{\epsilon1}$	Asn104 N$^{\delta2}$	3.6
Ala295 O	Lys246 N	2.5
Asp304 O$^{\delta1}$	Ser120 O$^{\gamma}$	2.7
Asp304 O$^{\delta1}$	Ser120 N	3.8
Asp304 O$^{\delta1}$	Thr119 O$^{\gamma1}$	2.3
Ser248 O$^{\gamma}$	Asn104 O$^{\delta1}$	3.2
Thr119 N	Gly117 O	3.3
Asp304 N	Thr119 O$^{\gamma1}$	3.0
Gly117 N	Ser120 O$^{\gamma}$	3.4
Asn104 N$^{\delta2}$	Glu258 O$^{\epsilon1}$	3.4
Lys246 N	Ala295 O	2.4
Ser12 O$^{\gamma}$	Asp304 O$^{\delta2}$	3.2
Thr119 O$^{\gamma1}$	Asp304 O$^{\delta2}$	3.1

conformation that could potentially be stabilized by interactions with the residues in the (123–139) helix (Fransson *et al.*, 1992; Jonasson *et al.*, 1997). However, in TcruCA the segment corresponding to the (123–139) helix in hCA II is instead replaced by a short loop. Therefore, the presence of the disulfide bond could potentially stabilize the loop with Thr252 (199) and Pro254 (202) and thus contribute to the increased stability of the enzyme.

The same intramolecular disulfide bond between Cys99 (21) and Cys256 (203) in TcruCA has been observed in other α-CA monomers, such as hCA IV, hCA VI, hCA IX, hCA XII, NgA (*Neisseria gonorrhoeae* α-carbonic anhydrase), AoCA and SspCA (Kannan *et al.*, 1972; Stams *et al.*, 1996; Whittington *et al.*, 2001; Alterio *et al.*, 2009; Cuesta-Seijo *et al.*, 2011; Capasso *et al.*, 2012). In addition, the position of this disulfide bridge is conserved between these enzymes (Supplementary Figs. S5*b* and S5*c*). The presence of the disulfide bond in hCA IV was proved to confer decreased susceptibility of the protein to denaturation by 5% SDS (Whitney & Briggle, 1982; Waheed *et al.*, 1996). Guanidine–HCl (GdnHCl)-induced denaturation experiments performed with NgCA also demonstrated that the inactivation of the enzyme occurred at lower concentrations of GdnHCl upon reduction of the disulfide bond, decreasing from 2.1 *M* GdnHCl to 1.2 *M* GdnHCl, which is comparable to the inactivation of hCA II by GdnHCl (at 0.9 *M* GdnHCl) (Mårtensson *et al.*, 1992; Elleby *et al.*, 2001). The reduction of the disulfide bond in NgCA led to decreased conformational stability, making it more sensitive to denaturation/inactivation by GdnHCl. It was inferred in subsequent papers regarding other α-CAs that the presence of this disulfide bond accounted for increased conformational stability, including the incorporation of the same disulfide bond into an hCA II variant (Boone, Habibzadegan, Gill *et al.*, 2013). Therefore, previous studies as well as our data strongly suggest that this disulfide bond is a major contributor to the increase in thermal and conformational stability of TcruCA.

3.4. Utilizing the TcruCA dimer interface for thermostable enzyme engineering

As mentioned previously, TcruCA is predicted to exist as a homodimer. Structural analysis of the homodimeric complex shows that each monomeric subunit of TcruCA appears as a 180° rotation about the dimeric interface, as has been observed in other dimeric α-CAs (Whittington *et al.*, 2001; Alterio *et al.*, 2009; Vullo *et al.*, 2012), with parallel facing active-site openings (Fig. 4*a*). Residues that contribute to the biological dimeric interface were predicted using the *PDBePISA* software (Krissinel & Henrick, 2007). The results from *PDBePISA* analysis indicate that 63 residues form the dimeric interface of TcruCA, with a total of 18 predicted hydrogen bonds contributing to the stabilization of the dimer (Table 5). In addition, the ΔG_{solv} values for TcruCA and known biological dimers of hCA VI, hCA IX, Cr-αCA1, *Ao*CA and SspCA were comparable and indicative of favorable dimer formation (Table 4). Interestingly, dimeric interaction of hCA IX, which has a predicted 54 interfacing residues, has only two predicted hydrogen bonds contributing to dimer stabilization. However, unlike TcruCA and other prokaryotic α-CAs, hCA IX has an intermolecular disulfide bond that stabilizes the complex *in vivo* (Alterio *et al.*, 2009). Therefore, it is predicted that the majority of interactions that stabilize the dimeric complex of hCA IX can be attributed to both hydrophobic interactions and the intermolecular disulfide bridge between monomers, whereas the various microbial α-CA dimers, including TcruCA, are stabilized *via* mostly hydrogen-bonding interactions (Table 6). It also should be noted that despite the prediction that hCA VI and XII form dimers, ΔG_{solv} *P*-values for each enzyme indicate it is difficult to discern whether these interactions are biological or an artifact of crystal packing (Table 5). Interestingly, however, CA XII is a known biological dimer (Whittington *et al.*, 2001), therefore using this value alone cannot determine the exact oligomeric state.

Examination of the dimeric interface of TcruCA indicates that potential key interactions form between residues 114 and 122 (37 and 43). This is indicative of the results from *PDBe-*

PISA analysis (not shown). Interestingly, the same region is predicted (*via PDBePISA* analysis) to be important for dimerization in both hCA IX and SspCA, and is also the location of the intermolecular disulfide bond found in hCA IX. As such, this region has been defined as the dimer axis (Fig. 4*b*). The TcruCA dimer is stabilized by several hydrogen bonds, including those found between adjacent Glu258 (205) residues (chains *A* and *B*), and interactions between the main chains of the Ala296 (251) residues and the side-chain amines of Lys246 (193) (Fig. 4*b*). In addition, we observe weak interactions between adjacent Asn104 (27) residues (chains *A* and *B*); however, the distances indicate that this may not be owing to hydrogen bonding (Fig. 4*b*). Overlay of these residue positions with those in SspCA, hCA IX and hCA II indicates that the position of Glu258 (205) is conserved in each isoform (Figs. 4*c* and 4*d*). Despite the clear interaction of this residue between monomers, the fact that it is conserved in hCA II, which does not dimerize, suggests that it alone is not sufficient to stabilize an oligomeric state. In addition, the crystallization conditions for both TcruCA and SspCA were acidic (pH of <5.5), which would indicate that this residue is deprotonated and thus the interaction formed at the dimer interface of both TcruCA and SspCA may be an artifact of crystal packing. Asn104 (27) is conserved between TcruCA and SspCA and clearly forms an interaction between monomers (Fig. 4*c*). This residue is not conserved, however, in hCA IX (Fig. 4*d*) and hCA II (Fig. 4*e*). This suggest that Asn104 (27) might have a significant influence on the dimerization of CA monomers. A predicted strong hydrogen bond (2.4 Å) is formed between Lys246 (193) of chain *A* and the carboxyl of Ala296′ (251′) of chain *B* (Fig. 4*d*). It is possible that owing to the strength of this hydrogen bond and the 'criss-crossing' arrangement of interactions between chain *A* and chain *B* of TcruCA monomers, these interaction would have a substantial impact on stabilizing the dimerization complex. Overlay of this residue position in SspCA, hCA II and hCA IX shows that this interaction is not conserved. Instead, this interaction is abolished by either replacement of Lys246 (193) with a Thr, which does not extend enough at its side chain to form any interaction (SspCA; Fig. 4*c*), or the replacement of Ala295 (251) with the more bulky Leu (hCA IX and hCA II; Figs. 4*b* and 4*c*, respectively), which would eliminate any possibility of hydrogen-bond formation. Despite this interaction not being present in SspCA and hCA IX, dimerization still occurs in both isoforms. It is likely that this is attributable to different hydrogen-bonding network arrangements in SspCA and the presence of the disulfide linkage and hydrophobic interactions in hCA IX, both of which are not present in hCA II.

The dimeric nature of TcruCA appears to greatly enhance the thermostability of the enzyme, as observed in the DSC data (Table 4). The same has been observed regarding other proteins or enzymes, where dimeric complexes favor a more stable state *versus* the monomeric unit (Anosike *et al.*, 1975). The same can be said regarding the stability of protein crystals, which can remain preserved in their crystalline form for very long periods of time (Shenoy *et al.*, 2001). These criteria alone would potentially make TcruCA, or any other dimeric α-CA, a good candidate for industrial use owing to its attributed thermostability. To date, protein engineering of a thermostable hCA II variant has been focused on alterations in the monomeric structure of the enzyme (Boone, Habibzadegan, Gill *et al.*, 2013; Boone, Habibzadegan, Tu *et al.*, 2013; Boone *et al.*, 2015). Although this has been successful in increasing the overall conformational stability of the enzyme, none of these variants have been utilized for industrial applications. Engineering of a thermostable hCA II dimer may be useful in these processes. It is possible that engineering a 'TcruCA mimic' using hCA II with the aforementioned sites [specifically the residues at positions 258 (205), 104 (27), 295 (251) and 246 (193)] would influence the formation of an hCA II dimer. Furthermore, it is observed that there is an overlap in the residues found in the dimer axis of TcruCA and hCA II (not shown). As such, it may also be possible to alter residues in this region to influence the formation of hydrogen bonding between monomers of hCA II in order to influence dimerization. This would create a dimeric interface similar to that observed in the microbial α-CAs. However, since this interface is constructed of mostly hydrogen-bond interactions it may be significantly sensitive to extreme pH changes and hence may not be useful for industrial CO_2 sequestration (Maeda *et al.*, 1995; Ochoa-González *et al.*, 2014). To combat this, it may be beneficial to engineer an hCA II dimeric interface consisting of combinations of attributes from TcruCA and hCA IX such that the key residues highlighted in Fig. 4(*e*) can be incorporated into the interface, along with the hCA IX intermolecular disulfide found in the dimer axis (Fig. 4*d*). Previous studies have shown that hCA II can be engineered into a dimer *via* an intermolecular disulfide (Mack *et al.*, 2011); however, this study did not speculate on the potential increased stability of the hCA II dimer. The presence of an engineered disulfide will potentially keep the hCA II dimer intact at more extreme pH levels (pH of ∼3.0; He *et al.*, 2006). These notions can be taken a step further by being employed to generate a dimer of previously discovered thermostable hCA II variants.

4. Conclusions

TcruCA is a dimeric α-CA (Fig. 1*a*) that is active and appears to be thermostable over a wide range of pH values. The overall structure of TcruCA exhibits a typical α-CA fold and reveals a conserved active-site structure relative to other isoforms. It is likely that this region is the highest contributor to the ∼30–40% sequence identity of TcruCA to other isoforms (Table 2). Like other microbial α-CAs, TcruCA exhibits a more compact structure relative to hCA II, with deletions in surface loops and helical structure (Figs. 1*b* and 1*c*). TcruCA also contains an intramolecular disulfide bond that has been observed in several α-CAs (Supplementary Fig. S5). The compact structure, the presence of an intramolecular disulfide bond and the dimeric arrangement of TcruCA are proposed to contribute significantly to the thermostability of the enzyme (Boone, Habibzadegan, Tu *et al.*, 2013; Boone *et al.*, 2015). Similar to other α-CAs, the active site of TcruCA is subdivided into a hydrophilic and a hydrophobic pocket (Fig. 2). Examination of

the active site reveals that the residues are highly conserved compared with hCA II and SspCA, with residue variations occurring mostly in the hydrophilic cleft (Fig. 3). As such, TcruCA exhibits a reduced catalytic efficiency (approximately tenfold) compared with hCA II, which is most likely attributable to these residue substitutions (Table 3). Furthermore, the reduced catalytic efficiency might also be a result of the loop deletion in region 1 (corresponding to residues 230–240 in hCA II), which has recently been suggested to be important for catalysis (Boone *et al.*, 2015).

The dimer of TcruCA is formed predominantly through hydrogen bonding and interactions *via* residues with hydrophilic side chains. This is comparable to other microbial α-CAs that typically form dimeric complexes, such as SspCA. This observation is alternative to hCA IX, which contains only two hydrogen bonds in the dimer interface. The hCA IX dimer is instead stabilized by hydrophobic interactions and the presence of an intermolecular disulfide bridge. As such, it is possible that utilizing the predicted key interface interactions observed in both TcruCA and hCA IX in order to engineer a hCA II dimer might produce a highly thermostable and pH-stable CA variant that can be implemented for industrial CO_2 sequestration. This includes implementing residues to facilitate hydrogen bonding at positions 104 (27), 296 (251) and 246 (193) and an intermolecular disulfide at position (41) (Fig. 4*e*). The successful construction of a dimeric hCA II variant may lead to a novel biocatalytic agent for industrial CO_2 sequestration.

Acknowledgements

This work was supported in part by an NIH GM25154 and NSF-MCB-0643713 grant. The authors would like to thank the staff at the Cornell High Energy Synchrotron Source (CHESS) for assistance during X-ray diffraction data collection and the Center of Structural Biology (CSB) for support of the X-ray facility at the University of Florida.

References

Adams, P. D. *et al.* (2011). *Methods*, **55**, 94–106.
Alterio, V., Hilvo, M., Di Fiore, A., Supuran, C. T., Pan, P., Parkkila, S., Scaloni, A., Pastorek, J., Pastorekova, S., Pedone, C., Scozzafava, A., Monti, S. M. & De Simone, G. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 16233–16238.
Alvizo, O., Nguyen, L. J., Savile, C. K., Bresson, J. A., Lakhapatri, S. L., Solis, E. O. P., Fox, R. J., Broering, J. M., Benoit, M. R., Zimmerman, S. A., Novick, S. J., Liang, J. & Lalonde, J. J. (2014). *Proc. Natl Acad. Sci. USA*, **111**, 16436–16441.
Anosike, E. O., Moreland, B. H. & Watts, D. C. (1975). *Biochem. J.* **145**, 535–543.
Avvaru, B. S., Kim, C. U., Sippel, K. H., Gruner, S. M., Agbandje-McKenna, M., Silverman, D. N. & McKenna, R. (2010). *Biochemistry*, **49**, 249–251.
Boone, C. D., Gill, S., Habibzadegan, A. & McKenna, R. (2013*a*). *Int. J. Chem. Eng.* **2013**, 813931.
Boone, C. D., Habibzadegan, A., Gill, S. & McKenna, R. (2013*b*). *Biomolecules*, **3**, 553–562.
Boone, C. D., Habibzadegan, A., Tu, C., Silverman, D. N. & McKenna, R. (2013). *Acta Cryst.* D**69**, 1414–1422.
Boone, C. D., Rasi, V., Tu, C. & McKenna, R. (2015). *FEBS J.* **282**, 1445–1457.
Boron, W. F. (2010). *Biochim. Biophys. Acta*, **1804**, 410–421.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Bruylants, G., Wouters, J. & Michaux, C. (2005). *Curr. Med. Chem.* **12**, 2011–2020.
Canadell, J. G., Le Quéré, C., Raupach, M. R., Field, C. B., Buitenhuis, E. T., Ciais, P., Conway, T. J., Gillett, N. P., Houghton, R. A. & Marland, G. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 18866–18870.
Capasso, C., De Luca, V., Carginale, V., Cannio, R. & Rossi, M. (2012). *J. Enzyme Inhib. Med. Chem.* **27**, 892–897.
Chegwidden, W. R., Carter, N. D. & Edwards, Y. H. (2000). *The Carbonic Anhydrases: New Horizons*. Basel: Birkhäser Verlag.
Cuesta-Seijo, J. A., Borchert, M. S., Navarro-Poulsen, J.-C., Schnorr, K. M., Mortensen, S. B. & Lo Leggio, L. (2011). *FEBS Lett.* **585**, 1042–1048.
Díaz Torres, N., González, G., Biswas, S., Scott, K. M. & McKenna, R. (2012). *Acta Cryst.* F**68**, 1064–1066.
Dobrinski, K. P., Boller, A. J. & Scott, K. M. (2010). *Appl. Environ. Microbiol.* **76**, 3561–3567.
Domsic, J. F., Avvaru, B. S., Kim, C. U., Gruner, S. M., Agbandje-McKenna, M., Silverman, D. N. & McKenna, R. (2008). *J. Biol. Chem.* **283**, 30766–30771.
Duda, D. M. & McKenna, R. (2006). *Handbook of Metalloproteins*, New York: John Wiley & Sons.
Elleby, B., Chirica, L. C., Tu, C., Zeppezauer, M. & Lindskog, S. (2001). *Eur. J. Biochem. FEBS*, **268**, 1613–1619.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Favre, N., Christ, M. L. & Pierre, A. C. (2009). *J. Mol. Catal. B Enzym.* **60**, 163–170.
Fisher, S. Z., Tu, C., Bhatt, D., Govindasamy, L., Agbandje-McKenna, M., McKenna, R. & Silverman, D. N. (2007). *Biochemistry*, **46**, 3803–3813.
Forsman, C., Behravan, G., Osterman, A., Jonsson, B. H., Enzell, C. R., Berg, J., Bartók, M., Pelczer, I. & Dombi, G. (1988). *Acta Chem. Scand.* **42**b, 314–318.
Fransson, C., Freskgård, P. O., Herbertsson, H., Johansson, A., Jonasson, P., Mårtensson, L. G., Svensson, M., Jonsson, B. H. & Carlsson, U. (1992). *FEBS Lett.* **296**, 90–94.
Gill, S. C. & von Hippel, P. H. (1989). *Anal. Biochem.* **182**, 319–326.
Goffredi, S. K., Childress, J. J., Desaulniers, N. T., Lee, R. W., Lallier, F. H. & Hammond, D. (1997). *J. Exp. Biol.* **200**, 883–896.
Göran Mårtensson, L., Jonsson, B. H., Andersson, M., Kihlgren, A., Bergenhem, N. & Carlsson, U. (1992). *Biochim. Biophys. Acta*, **1118**, 179–186.
Hansen, J., Sato, M., Ruedy, R., Lacis, A. & Oinas, V. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 9875–9880.
He, H. T., Gürsoy, R. N., Kupczyk-Subotkowska, L., Tian, J., Williams, T. & Siahaan, T. J. (2006). *J. Pharm. Sci.* **95**, 2222–2234.
Hewett-Emmett, D. & Tashian, R. E. (1996). *Mol. Phylogenet. Evol.* **5**, 50–77.
Huang, X. & Miller, W. (1991). *Adv. Appl. Math.* **12**, 337–357.
Intergovernmental Panel on Climate Change (2005). *IPCC Special Report on Carbon Dioxide Capture and Storage.* Cambridge University Press.
Jonasson, P., Aronsson, G., Carlsson, U. & Jonsson, B. H. (1997). *Biochemistry*, **36**, 5142–5148.
Kannan, K. K., Liljas, A., Waara, I., Bergstén, P. C., Lövgren, S., Strandberg, B., Bengtsson, U., Carlbom, U., Fridborg, K., Järup, L. & Petef, M. (1972). *Cold Spring Harb. Symp. Quant. Biol.* **36**, 221–231.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lindskog, S. (1997). *Pharmacol. Ther.* **74**, 1–20.
Lindskog, S. & Silverman, D. N. (2000). *EXS*, pp. 175–195.
Mack, E. T., Snyder, P. W., Perez-Castillejos, R. & Whitesides, G. M. (2011). *J. Am. Chem. Soc.* **133**, 11701–11715.
Maeda, K., Owada, M., Kimura, N., Omata, K. & Karube, I. (1995). *Energy Convers. Manage.* **36**, 717–720.

Matulis, D., Kranz, J. K., Salemme, F. R. & Todd, M. J. (2005). *Biochemistry*, **44**, 5258–5266.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Mehler, E. L., Fuxreiter, M., Simon, I. & Garcia-Moreno, E. B. (2002). *Proteins*, **48**, 283–292.
Metcalf, W. W., Jiang, W., Daniels, L. L., Kim, S.-K., Haldimann, A. & Wanner, B. L. (1996). *Plasmid*, **35**, 1–13.
Mikulski, R., Domsic, J. F., Ling, G., Tu, C., Robbins, A. H., Silverman, D. N. & McKenna, R. (2011). *Arch. Biochem. Biophys.* **516**, 97–102.
Mirjafari, P., Asghari, K. & Mahinpey, N. (2007). *Ind. Eng. Chem. Res.* **46**, 921–926.
Ochoa-González, R., Díaz-Somoano, M. & Martínez-Tarazona, M. R. (2014). *J. Hazard. Mater.* **276**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pechkova, E., Sivozhelezov, V. & Nicolini, C. (2007). *Arch. Biochem. Biophys.* **466**, 40–48.
Pierre, A. C. (2012). *ISRN Chem. Eng.* **2012**, 753687.
Pinard, M. A., Mahon, B. P. & McKenna, R. (2015). *Biomed Res. Int.* **2015**, 453543.
Savile, C. K. & Lalonde, J. J. (2011). *Curr. Opin. Biotechnol.* **22**, 818–823.
Scott, K. M. *et al.* (2006). *PLoS Biol.* **4**, e383.
Shenoy, B., Wang, Y., Shan, W. & Margolin, A. L. (2001). *Biotechnol. Bioeng.* **73**, 358–369.
Shively, J. M., van Keulen, G. & Meijer, W. G. (1998). *Annu. Rev. Microbiol.* **52**, 191–230.
Simonsson, I., Jonsson, B. H. & Lindskog, S. (1979). *Eur. J. Biochem.* **93**, 409–417.
Stams, T., Nair, S. K., Okuyama, T., Waheed, A., Sly, W. S. & Christianson, D. W. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 13589–13594.
Tripp, B. C., Smith, K. & Ferry, J. G. (2001). *J. Biol. Chem.* **276**, 48615–48618.
Tu, C., Mikulski, R., Swenson, E. R. & Silverman, D. N. (2009). *Free Radic. Biol. Med.* **46**, 14–19.
Tu, C. K., Silverman, D. N., Forsman, C., Jonsson, B. H. & Lindskog, S. (1989). *Biochemistry*, **28**, 7913–7918.
Vogt, G., Woell, S. & Argos, P. (1997). *J. Mol. Biol.* **269**, 631–643.
Vullo, D., De Luca, V., Scozzafava, A., Carginale, V., Rossi, M., Supuran, C. T. & Capasso, C. (2012). *Bioorg. Med. Chem. Lett.* **22**, 6324–6327.
Waheed, A., Okuyama, T., Heyduk, T. & Sly, W. S. (1996). *Arch. Biochem. Biophys.* **333**, 432–438.
Whitney, P. L. & Briggle, T. V. (1982). *J. Biol. Chem.* **257**, 12056–12059.
Whittington, D. A., Waheed, A., Ulmasov, B., Shah, G. N., Grubb, J. H., Sly, W. S. & Christianson, D. W. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 9545–9550.
Yang, Z. W., Tendian, S. W., Carson, W. M., Brouillette, W. J., Delucas, L. J. & Brouillette, C. G. (2004). *Protein Sci.* **13**, 830–841.
Yu, Y., Chen, B., Qi, W., Li, X., Shin, Y., Lei, C. & Liu, J. (2012). *Microporous Mesoporous Mater.* **153**, 166–170.