Protein Engineering, Design & Selection, 2017, vol. 30 no. 8, pp. 559–570
doi: 10.1093/protein/gzx049
Advance Access Publication Date: 9 September 2017
Original Article

Original Article

Role of an N-terminal extension in stability and catalytic activity of a hyperthermostable α/β hydrolase fold esterase

Mrityunjay K. Singh, Santosh Shivakumaraswamy, Sathyanarayana N. Gummadi, and Narayanan Manoj*

Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences Indian Institute of Technology Madras, Chennai 600036, India

*To whom correspondence should be addressed. E-mail: nmanoj@iitm.ac.in
Edited by Dick Janssen

Received 28 June 2017; Revised 11 August 2017; Editorial Decision 14 August 2017; Accepted 25 August 2017

Abstract

The carbohydrate esterase family 7 (CE7) enzymes catalyze the deacetylation of acetyl esters of a broad range of alcohols and is unique in its activity towards cephalosporin C. The CE7 fold contains a conserved N-terminal extension that distinguishes it from the canonical α/β hydrolase fold. The hexameric quaternary structure indicates that the N-terminus may affect activity and specificity by controlling access of substrates to the buried active sites via an entrance tunnel. In this context, we characterized the catalytic parameters, conformation and thermal stability of two truncation variants lacking four and ten residues of the N-terminal region of the hyperthermostable *Thermotoga maritima* CE7 acetyl esterase (TmAcE). The truncations did not affect the secondary structure or the fold but modulated the oligomerization dynamics. A modest increase was observed in substrate specificity for acetylated xylose compared with acetylated glucose. A drastic reduction of ~30–40°C in the optimum temperature for activity of the variants indicated lower thermal stability. The loss of hyperthermostability appears to be an indirect effect associated with an increase in the conformational flexibility of an otherwise rigid neighboring loop containing a catalytic triad residue. The results suggest that the N-terminal extension was evolutionarily selected to preserve the stability of the enzyme.

Key words: access tunnel, carbohydrate esterase family 7, cephalosporin C deacetylase, N-terminal extension, thermal stability

Introduction

The carbohydrate esterase family 7 (CE7) is a member of the large and diverse α/β hydrolase superfamily. Enzymes in this family catalyze the deacetylation of acetyl esters of a broad range of alcohols including carbohydrates and aromatic groups such as *p*-nitrophenol, 4-methylumbelliferol, α/β naphthol and fluorescein derivatives. Among the five known CE families (CEs1,5,7,13,15) within this superfamily, the CE7 is unique in its activity towards cephalosporin C (CPC) and 7-amino-cephalosporanic acid (7-ACA) (Shao and Wiegel, 1995; Lorenz and Wiegel, 1997; Degrassi *et al.*, 1998, 2000; Vincent *et al.*, 2003; Hedge *et al.*, 2012; Levisson *et al.*, 2012). Crystal structures of four orthologs of the CE7 family available in the Protein Data Bank (PDB) show a highly conserved tertiary structure containing an α/β hydrolase fold making up a homohexameric assembly that represents the quaternary structure. Furthermore, the family utilizes the canonical Ser-His-Asp catalytic triad mechanism of action for the esterase activity (Vincent *et al.*, 2003; Krastanova *et al.*, 2009; Silvia *et al.*, 2011; Levisson *et al.*, 2012; Singh and Manoj, 2016a). The tertiary structure of the CE7 subunit contains three conserved structural deviations that differentiate it from the canonical α/β hydrolase fold. These include (i) an N-terminal extension containing two α-helices (αA-2, αA-1) and a β-strand (β−1),

(ii) a three-helix bundle insertion after strand β6 and, (iii) a loop insertion element called the 'β-interface loop' between strand β4 and helix αB (Fig. 1, Supplementary Fig. S1). Interestingly, all three insertion elements are also involved in inter-subunit interface interactions of the oligomeric assembly signifying their roles in shaping the CE7 enzyme properties (Vincent *et al.*, 2003). Our previous structural and biochemical studies on the *Thermotoga maritima* CE7 (TmAcE) and its variants have shown that the β-interface loop and the three-helix bundle insertion elements are crucial for the stability, oligomerization dynamics and catalytic properties of the enzyme (Singh and Manoj, 2016a, 2017).

The CE7 quaternary structure is a 'doughnut' shaped hexamer made up of a trimer of dimers and displays a narrow entrance tunnel leading to the centre of the oligomer where the six catalytic triads are located, secluded from the bulk solvent. Examination of the hexamer indicates no other tunnel or path except this, for the diffusion of substrates and products in and out of the active site. The structure also reveals that the entrance of the tunnel is 'gated' by the N-termini (Fig. 2). It is thus expected that this gating element may act to function as a 'selectivity' filter to regulate access of substrates to the active site. Although the notion of the selectivity filter in the CE7 family was first presented by Vincent *et al.* for the *Bacillus subtilis* CE7, there is lack of any experimental data to validate this feature (Vincent *et al.*, 2003). In order to gain insights into the role of the N-terminal extension in determining the catalytic and stability properties of the CE7 family, biochemical and biophysical characterization of N-terminal variants of TmAcE were carried out. CE7 enzymes are used in the conversion of 7-ACA to 3-deacetoxy-7-aminocephalosporanic acid (7-ADCA) which then serves as a template for generating derivatives containing modifications at the C-3 position (Pollegioni *et al.*, 2013). Although CE7 enzymes from different *B. subtilis* species are currently utilized, efforts towards 'rational' engineering of enzymes with desirable properties like higher activity and stability are of interest (Konecny and Sieber, 1980; Takimoto *et al.*, 1999; Sonawane, 2006; Tian *et al.*, 2014). The results of this study revealed that the absence of N-terminal residues in TmAcE variants did not affect the activity, but showed a modest effect on substrate specificity. However, the deletions significantly lowered the thermostability, suggesting that the N-terminal extension is an evolutionary adaptation with a crucial role in the maintenance of stability of the CE7 protein.

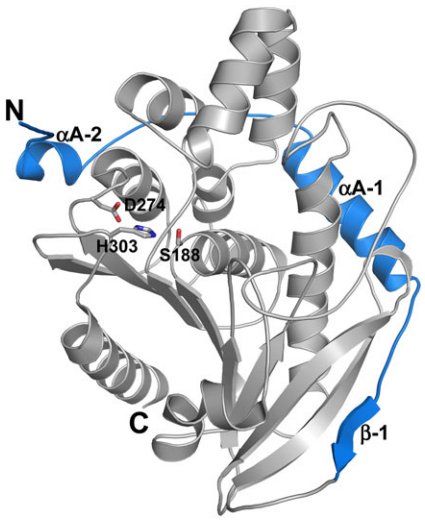

Fig. 1 The tertiary structure of TmAcE. TmAcE subunit is shown in ribbon representation. The N-terminal extension region in the CE7 subunit relative to the canonical α/β hydrolase fold is shown in blue color. The secondary structure elements in the extension region are labeled with respect to the core α/β hydrolase fold (see Supplementary Fig. S1). The catalytic triad residues Ser-His-Asp are labeled and shown in stick representation (PDB: 3M83). This figure and figures 2, and 11 are made using the PyMOL program (DeLano, 2006). For interpretation of references to colors in the legends for figures 1, 2 and 11, the reader is referred to the online version of this article.

Materials and methods

Generation of TmAcE deletion constructs

The plasmid containing the *TM0077* gene that encodes TmAcE was procured from the DNASU plasmid repository (Clone ID: TmCD00083900). The cloned gene in a pMH1 based expression vector with an ampicillin resistance marker is controlled by the

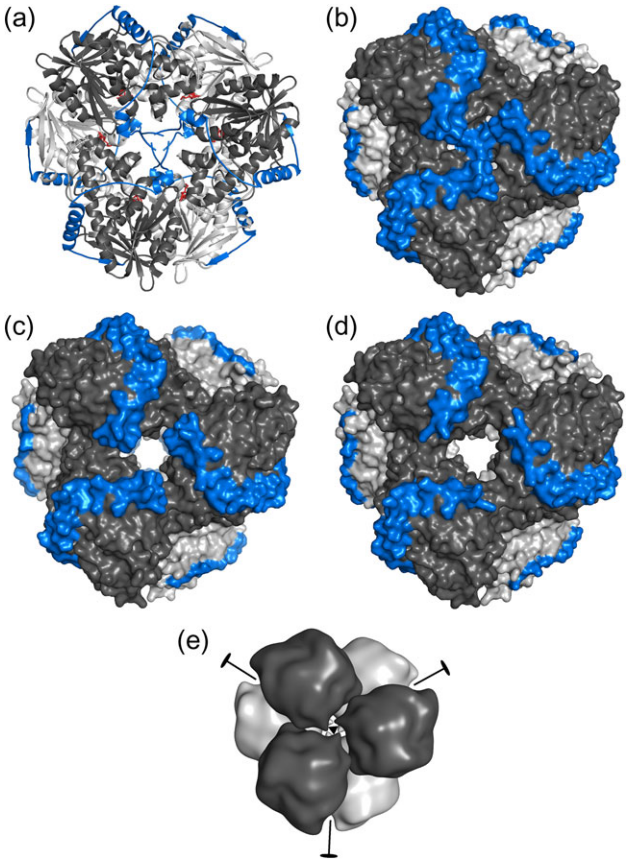

Fig. 2 Quaternary structures of TmAcE and models of variants. (**a**) Ribbon diagram of the TmAcE hexamer shown as viewed down the 3-fold axis. The N-terminal extension element is shown in blue. DEP covalently bound to the catalytic serine is shown in red sticks indicating the buried position of the catalytic cavity distal to the tunnel entrance. The N-termini of the trimer block the mouth of the tunnel that leads to the buried active site cavity (PDB: 3M83). (**b**) The corresponding surface representation of the TmAcE hexamer. (**c**) Surface diagram of model of Δ4-TmAcE showing opening up of central cavity upon deletion of first five amino acids. (**d**) Surface diagram of model of Δ10-TmAcE showing opening up of central cavity upon deletion of the first 11 amino acids. (**e**) Schematic diagram of the hexameric assembly as viewed down the 3-fold axis. The hexameric assembly is a trimer of dimers in a 'doughnut' like back-to-back arrangement of two trimers (shown in gray and black).The locations of the 2-fold axes that are perpendicular to the 3-fold axis are indicated.

araBAD promoter. The gene codes for a 325 amino acid long protein with an additional twelve residue N-terminal expression and purification tag (MGSDKIHHHHHH). Three N-terminal deletion constructs were generated by PCR using the polymerase incomplete primer extension method (Klock *et al.*, 2008; Klock and Lesley, 2009). These include, TmAcE lacking the tag (ΔHT-TmAcE); two deletion constructs lacking the tag and the residues 2–5 (Δ4-TmAcE) and the residues 2–11 (Δ10-TmAcE). The primers were as follows: ΔHT-TmAcE sense: 5′-ACATACCCATGGCCTTCTTCGATTTACCACTCGAAGAAC-3′, ΔHT-TmAcE antisense: 5′-AGGCCATGGGTATGTATATCTCC TTCTTAAAGTTAAAC-3′, Δ4-TmAcE sense: 5′-CCATGTTACC ACTCGAAGAACTGAAGAA-3′, Δ4-TmAcE antisense 5′-GTGGT AACATGGGTATGTATATCTCCTTC-3′, Δ10-TmAcE sense: 5′-T AC CCATGAAGAAATATCGTCCAGAGCGGTACGAAGAG-3′, Δ10-TmAcE antisense: 5′-CGATATTTCTTCATGGGTATGTAT ATCTCCTTCTTAAAGTTA-3′. PCR products were treated with DpnI enzyme for 4 h at 37°C for digestion of the parental template. After digestion, the PCR product was transformed into *Escherichia coli* DH5α cells. Positive clones were selected by ampicillin resistance and grown overnight in Luria Bertani (LB) broth with 0.1 mg/ml ampicillin. Plasmid purification was carried out using the Qiagen Plasmid Miniprep Kit as per the manufacturer's instructions. Deletions were confirmed by DNA sequencing of the entire protein coding region.

Protein expression and purification

The four proteins were expressed in *E. coli* BL-21 RIL (DE3) cells following an identical protocol. Plasmid transformants were grown overnight in 10 ml LB broth supplemented with 0.1 mg/ml ampicillin and 0.034 mg/ml chloramphenicol at 37°C with constant shaking at 180 rpm. This overnight culture was used to inoculate 1 l fresh LB broth with antibiotics and grown at 37°C with shaking at 180 rpm until an OD_{600} of 0.5–0.6 was reached. The cells were induced by adding L (+) arabinose to a final concentration of 0.2% (w/v). Cells were harvested by centrifugation (4500 g at 4°C) after 4 h of further incubation at 37°C at 180 rpm. About 1 g cell pellet was re-suspended in 10 ml of lysis buffer (20 mM Tris HCl, pH 8.0, 150 mM NaCl, for TmAcE and 20 mM Tris HCl, pH 8.0, 20 mM NaCl, for ΔHT-TmAcE, Δ4-TmAcE and Δ10-TmAcE). The sample was then placed on ice and lysed by sonication. Cell debris was clarified by centrifugation at 18 500 g for 45 min at 4°C. The supernatant was subjected to thermal precipitation of host cell proteins for a 30 min period in a water bath at 60°C (for TmAcE and ΔHT-TmAcE) and at 50°C (for Δ4-TmAcE and Δ10-TmAcE). The precipitated proteins were removed by centrifugation and the supernatant used for purification using affinity chromatography. In the case of TmAcE, the supernatant was loaded onto a Ni-Nitrilotriacetic acid (Ni^{2+}-NTA) sepharose column (GE Healthcare) pre-equilibrated with buffer A (20 mM Tris HCl, pH 8.0, 150 mM NaCl, 30 mM imidazole). The column was washed with 20 column volumes of buffer A. The protein was eluted using elution buffer (20 mM Tris HCl, pH 8.0, 150 mM NaCl, 300 mM imidazole). The eluted protein was then exchanged into buffer B (10 mM Tris HCl, pH 8.0, 150 mM NaCl) using a HiPrep 26/10 desalting column (GE Healthcare). For the deletion constructs, the supernatant was loaded on a 4 ml mono Q column pre-equilibrated with buffer (20 mM Tris HCl, pH 8.0, 20 mM NaCl). The column was washed with five column volumes of 20 mM Tris HCl, pH 8.0 buffer containing a gradient of NaCl from 20–100 mM. The presence of the relevant protein in the collected fractions was checked using an activity assay. Appropriate fractions were pooled and purified further by size exclusion chromatography (SEC) into the final buffer (10 mM Tris HCl, pH 8.0, 150 mM NaCl), concentrated to 25 mg/ml and stored at −80°C.

SEC was used to determine the quaternary structures of the purified proteins using a Superdex 200 10/300 GL analytical column (GE Healthcare) maintained at 4°C. The column was equilibrated with 20 mM Tris HCl, pH 8.0, 150 mM NaCl. 50 µl of protein (2.0 mg/ml) was loaded onto the column and eluted at a flow rate of 0.4 ml/min. Chromatograms were obtained by monitoring absorbance at 280 nm. The apparent molecular weights (Mr) of the samples were estimated by fitting the elution volume onto the calibration curve. (Supplementary Fig. S2).

Effect of pH and temperature on enzyme activity

All spectrophotometric measurements in this and the subsequent sections were made in a Perkin Elmer Lambda 25 UV-visible spectrometer equipped with a Peltier temperature controller and a magnetic stirrer. The effects of pH and temperature on the esterase activity of TmAcE and its variants were determined using 4-methylumbelliferyl acetate (MUFA) as the substrate. Esterase activity against MUFA was determined by monitoring the continuous release of 4-methylumbelliferone at 335 nm. The effect of pH on enzyme activity was carried out at 50°C using 50 mM buffers; citrate-sodium phosphate (pH 5.0–7.0), sodium phosphate (pH 7.5–8.0), sodium borate buffer pH (pH 8.5) and gylcine-NaOH (pH 9.0). A 400 mM stock solution of MUFA was prepared in ethylene glycol monoethyl ether. The reaction mixture (final volume 2 ml) contained 2 µg of enzyme in the buffer. The reaction was initiated by adding 20 µl of freshly prepared substrate from a 200 mM stock solution and the absorbance monitored for 60 s at 335 nm. The effect of temperature on activity was determined by measuring activity at different temperatures in the range 20–95°C, at pH 8.0. The reaction was initiated by adding 20 µl of freshly prepared substrate from the stock solution (400 mM) and the absorbance was monitored for 60 s at 335 nm. The change in buffer pH due to temperature was appropriately adjusted. Thermal deactivation studies were carried out by incubating 0.2 mg/ml of the enzyme at 60°C and 70°C for time periods up to 7 h. Aliquots of the enzyme at different time points were taken and the residual activity was measured at pH 8.0 and 50°C, using MUFA as substrate, as described above. One unit of esterase activity is defined as the amount of enzyme releasing 1 µmol of the product per min under assay conditions.

Steady state kinetic parameters

Activity assays for the four proteins were carried out following previously reported protocols (Singh and Manoj, 2016a). All measurements were made at a temperature of 50°C. Esterase activity against substrates *p*-nitrophenyl acetate (PNPA) and MUFA were determined by monitoring continuous release of products, *p*-nitrophenol at 405 nm and 4-methylumbelliferone at 335 nm (ε = 18000 $M^{-1}cm^{-1}$ and 11 040 $M^{-1}cm^{-1}$ for *p*-nitrophenol and 4-methylumbelliferone, respectively, under assay conditions). Stock solutions of PNPA and MUFA (200 mM) were prepared in ethylene glycol monoethyl ether. The reaction mixture of 2 ml contained 2 µg of enzyme in 50 mM sodium phosphate buffer, pH 8.0. The reaction was initiated by adding 20 µl of the substrate stock solution and the absorbance monitored for 60 s at the wavelength corresponding to the absorbance maxima of the product. Activities towards the two carbohydrate esters, penta-acetylated glucose (5-Ac-Glu) and tetra-acetylated xylose (4-Ac-Xy) and towards 7-aminocephalosporanic acid (7-ACA) were determined in 50 mM sodium phosphate buffer, pH 8.0. The product

(acetate) released was measured using the acetate estimation kit from Megazyme Ltd as per manufacturer's instructions. The reaction mixture in a total volume of 1 ml contained 4 µg of enzyme in 50 mM sodium phosphate buffer, pH 8.0. Stock solution of 200 mM 7-ACA was prepared by dissolving in water and lowering the pH to 4.0 prior to addition to the reaction mixture. The reaction was initiated by adding 200 µM of substrate (5-Ac-Glu, 4-Ac-Xy, 7-ACA) in buffer and allowed to proceed for 60 s in a thermomixer with continuous stirring at 600 rpm. The reaction was stopped by lowering the pH by addition of 50 µl of 1 M hydrochloric acid. A reaction time period of 10 min was used for 7-ACA. A 50 µl aliquot was taken from the reaction mixture to estimate the product released. The control blank contained all reaction components except the enzyme. Kinetic parameters of the esterase activity of TmAcE and its variants were measured by varying the substrate concentration and measuring activity. The initial rates of the reactions were measured and plotted against substrate concentrations. The saturation enzyme kinetic traces were fitted to a standard Michaelis-Menten equation using the SigmaPlot 11 program, Enzyme Kinetics Module1.3 and the values for k_{cat}, K_M were calculated. The effect of non-enzymatic hydrolysis of substrates was appropriately adjusted. All measurements were carried out in triplicates. One unit of esterase activity is defined as the amount of enzyme releasing 1 µmol of product per min under assay conditions.

Inhibition assay

The effect of serine modifying reagent diethyl *p*-nitrophenyl phosphate (paraoxon) on the enzyme was determined in 50 mM sodium phosphate buffer (pH 8.0) by incubating 0.2 mg/ml of enzyme with varying concentrations of paraoxon (0.0–1.0 mM) for 1 h at 4°C, before being assayed for residual activity at 50°C. The specific activity of the treated enzyme preparation was assayed using MUFA as the substrate, as described in the earlier section. The control blank contained paraoxon-free enzyme that was incubated for 1 h at 4°C. Values are reported as the percentage residual activity relative to the control. As a measure of protein sensitivity to paraoxon, IC_{50} values for 1 h were calculated by fitting the data points to a one-phase decay exponential equation in the GraphPad Prism 5.0 program.

CD spectroscopy and differential scanning calorimetry

Circular dichroism (CD) and differential scanning calorimetry (DSC) experiments were carried out with protein samples in 20 mM sodium phosphate buffer (pH 8.0). Far-UV CD spectra were measured using a Jasco J-815 spectropolarimeter equipped with a Peltier temperature controller. Spectra in the 200–250 nm range were recorded for ΔHT-TmAcE, Δ4-TmAcE and Δ10-TmAcE using a 1 mm path-length quartz cell containing protein at a concentration of 0.25 mg/ml, whereas, the TmAcE spectra was recorded using 0.1 mg/ml protein. Data were acquired with a scan rate of 100 nm/min and response time of 1 s. The effect of temperature on ellipticity (θ) was investigated by recording the spectra in the temperature mode (5–90°C) using stepwise increase at regular intervals to achieve thermal equilibrium. For thermal unfolding monitored by CD, ellipticity was measured at 222 nm with a scan rate of 1°C/min for the temperature range 5–95°C. The final spectra represented the average of at least three independent measurements normalized with respect to a comparative blank. The melting temperature was calculated utilizing the Boltzmann sigmoid (two-state model) equation in GraphPad Prism 6.0. DSC thermograms were recorded on a MicroCal VP-DSC microcalorimeter (GE Healthcare). Instrument base line was measured from reference and sample cells containing buffer. The sample concentrations used were 1.0 and 0.5 mg/ml, for TmAcE and the three variants, respectively. The samples were degassed for 10 min and equilibrated for 15 min. The heating scan rate was 1°C/min from 20 to 120°C at a constant pressure of 21 psi. Thermograms were recorded during the heating scans and appropriately corrected. Non-linear least-squared curve fitting procedure supplied with Origin 6.0 was used to resolve the complex profiles into individual components and to calculate values for transition temperatures of the component peaks.

Results

Design of the TmAcE N-terminal variants

The tertiary structures of TmAcE and the other orthologs reveal that the N-terminal extension element circumvents the subunit and is ~46 residues long (Fig. 1) (PDBs: 5FDF, 3FCY, 1ODS, 2XLB). The hexameric quaternary structure with a 32 point group symmetry contains two trimers that pack back-to-back with their 3-fold axes aligned but staggered by ~60° (Fig. 2). The active site cavity is accessible via a ~37 Å long axial channel with two entrances, one on each side of the two trimers. Each of these entrances is ~13 Å wide and connects to a short tunnel spanning ~10 Å reaching into the active site cavities (Singh and Manoj, 2016b).The N-terminus of each subunit in the trimeric arrangement is exposed to bulk solvent and protrude into the central part of this assembly constituting a gate in the tunnel mouth that leads to the catalytic cavities buried inside the oligomer. The first seven residues from each monomer of the trimer are involved in this structural feature (Supplementary Fig. S3). It was fairly clear that the tunnel entrance may function to modulate access of substrates to the poly-specific active site of this family. This assembly also offers a likely structural explanation for the absence of activity of the CE7 family towards acetyl esters of long chain carbohydrate polymers. Available biochemical data show that hydrolysis in this family is strictly limited to smaller sugar substrates like acetylated glucose, xylose and xylooligosaccharides (xylobiose, xylotriose) (Lorenz and Wiegel, 1997; Degrassi *et al.*, 2000; Vincent *et al.*, 2003; Montoro-García *et al.*, 2011; Hedge *et al.*, 2012; Levisson *et al.*, 2012). In contrast, paralogous families like the CE1, CE5, CE13 and CE15, which are either monomeric or dimeric, contain members that show acetylxylan esterase activity in addition to activity on small generic esters (Ghosh *et al.*, 1999; Schubot *et al.*, 2001; Kodama *et al.*, 2009; Pokkuluri *et al.*, 2011).

Sequence comparisons and superposition of the structures of the CE7 members indicate that the N-terminal region is well conserved (Supplementary Figs S4 and S5), although in some crystal forms, the first two residues are disordered. The hypothesis here is that the N-terminal region can undergo reversible conformational change and thus regulates the activity of the enzyme, and that an appropriate deletion of this region should therefore open up access to the active site and alter activity and/or specificity. Such gating elements involving single residues or secondary structure elements or loops or entire domains acting as 'lids' can have direct impact on enzymatic properties (Fishelovitch, *et al.*, 2009; Gora *et al.*, 2013). To test this hypothesis in the CE7 family, a set of N-terminal deletion variants of TmAcE was generated. In the proposed gating region, residues 1–7 are in a flexible random coil conformation thereby allowing for diffusion of substrates/products. An examination of the structure showed that the region from residues 1–11 are involved in inter-subunit and intra-subunit non-bonded interactions. The residues 1–4 in each subunit are

primarily involved in multiple inter-subunit non-bonded interactions (~26 interactions/subunit < 4.0 Å) with the same set of residues from the 3-fold related subunits within the trimer. However, residues 8–11 constitute an α-helix (αA-2) providing rigidity to the region in addition to contributing to multiple intra-subunit non-covalent interactions with the residues 273–283 of an adjacent loop (Supplementary Tables SI, SII). Interestingly, this loop contains a catalytic triad residue, namely Asp274 (Fig. 1).

All published biochemical and structural studies on TmAcE were hitherto performed using a protein construct containing a N-terminal 12 residue expression and purification tag (MGSDKIHHHHHH) to enable ease of protein purification (Hedge *et al.*, 2012; Levisson *et al.*, 2012; Singh and Manoj, 2016,a, b, 2017). However, examination of the structure suggests that the tag, although flexible and disordered in the crystal structures, could create a crowding effect at the tunnel entrance and thereby affect the diffusion of substrates towards the catalytic site. In light of the above observations, three N-terminal variants of the enzyme were generated. These include the wild-type TmAcE lacking the tag (henceforth called ΔHT-TmAcE) and two deletion constructs lacking the tag and the residues 2–5 (Δ4-TmAcE) and the residues 2–11 (Δ10-TmAcE) (Fig. 2). The expanded deletion in the latter variant was chosen to include residues of the helix αA-2 to examine the effect of this secondary structure feature on the enzyme properties (Fig. 1). These variants were purified and characterized and their properties compared with that of the tagged form, namely TmAcE.

Purification and oligomeric states of TmAcE and variants

Recombinant TmAcE and its N-terminal variants were purified to homogeneity as determined by SDS-PAGE analysis (Supplementary Fig. S6). The three tag-free variants were purified according to a modified protocol involving ion-exchange and SEC. The oligomeric states of the proteins were characterized by SEC (Fig. 3). The SEC profiles for TmAcE, ΔHT-TmAcE and Δ4-TmAcE display two distinct peaks, where the dominant peak and the minor peak correspond to the hexameric (~210–220 kDa) and the dimeric species (~76–78 kDa). A SEC peak corresponding to the dimeric species has previously been reported for both recombinant ΔHT-TmAcE and TmAcE, indicating the dynamic nature of the oligomeric state (Drzewiecki *et al.*, 2010, Singh and Manoj, 2016a). In Δ10-TmAcE, a single dominant peak at ~472 kDa, probably a higher oligomeric form of the protein, is observed (Fig. 3c). Two minor peaks at ~59 kDa, corresponding to the dimeric species, and ~1153 kDa, a likely aggregate, can be observed. While the data indicates no change in the oligomeric states of ΔHT-TmAcE and Δ4-TmAcE, the significant change observed in the Δ10-TmAcE was unexpected since the deleted residues do not make significant inter-subunit interactions relative to contributions from the rest of the subunit.

Effect of pH and temperature on specific activity of purified enzymes

The specific activities of all constructs were examined as a function of pH and temperature separately, using 4-methylumbelliferyl acetate (MUFA) as the substrate. The effect of pH on activity was measured in the pH range 5.0–9.0 at 50°C. TmAcE and the variants exhibited optimal activity between pH 7.5 and 8.5, with maximum activity at pH 8.0 (Fig. 4a). Figure 4b shows the effect of temperature on the activity of the proteins in the range 20–95°C at pH 8.0. The esterase activity for TmAcE and ΔHT-TmAcE increased from 20°C until 95°C. Excessive auto-hydrolysis of the substrate at higher temperatures limited the assay to 95°C. Presumably, the optimum temperature for both constructs is close to the value of 100°C, as reported earlier for TmAcE (Levisson *et al.*, 2012). It is evident that the N-terminal tag did not significantly affect the activity-pH/temperature relationship of TmAcE. In contrast, both Δ4-TmAcE and Δ10-TmAcE display an increase in activity from 20°C and achieve a maximum value between 60 and 70°C. Thereafter, these proteins show a drastic decrease in activity from 70 to 80°C. The data suggest that the deletion variants are almost as well folded as the tagged and the tag-free forms, but their thermodynamic stabilities are seriously disrupted.

The kinetic thermal stability of each protein was next investigated by thermal deactivation studies. Measurements of residual activity were made after incubation of the protein at 60 and 70°C, for different time periods. As shown in Fig. 5, TmAcE and ΔHT-TmAcE

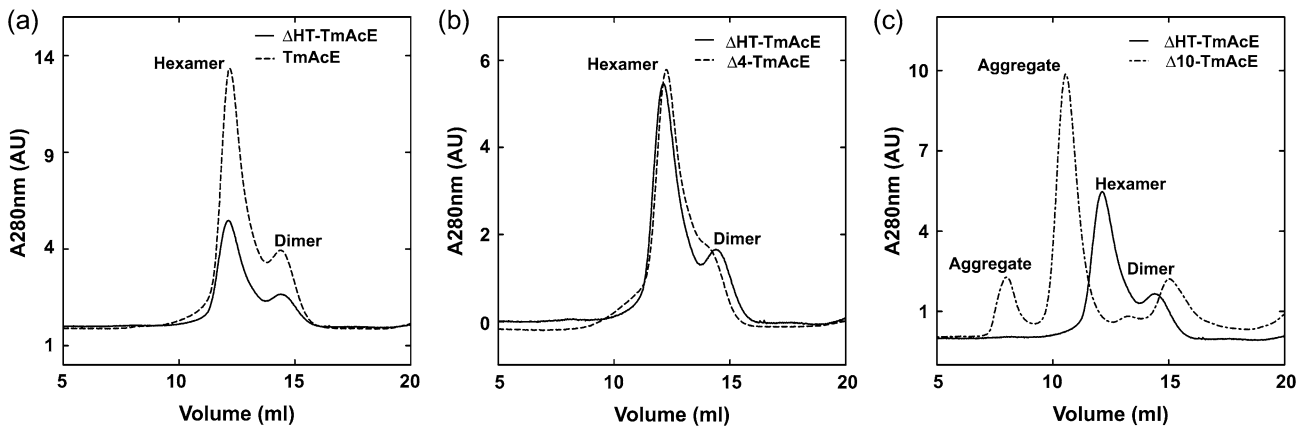

Fig. 3 Oligomeric states of TmAcE and variants. The solid line in all panels in the size exclusion chromatograms represents the native N-terminal tag-free protein (ΔHT-TmAcE), while the other proteins are represented by dashed lines. The elution profile peaks of ΔHT-TmAcE correspond to native molecular masses of 222 kDa and 74 kDa, indicating a mixture of a predominant hexameric species and a minor dimeric species, respectively. Comparison of elution profiles of (**a**) ΔHT-TmAcE and TmAcE (**b**) ΔHT-TmAcE and Δ4-TmAcE, indicate similar oligomerization states. (**c**) Comparison of elution profiles of ΔHT-TmAcE and Δ10-TmAcE. The elution peaks for Δ10-TmAcE correspond to 472 kDa for the major species and 1152 kDa and 59 kDa for the two minor species, and differ significantly from the other proteins.

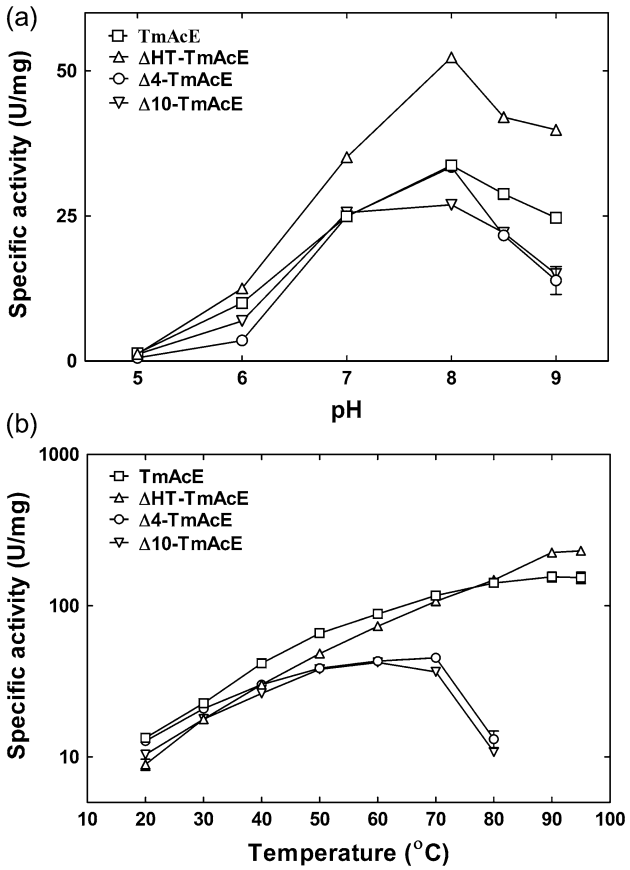

Fig. 4 Effect of pH and temperature on activity of TmAcE and variants. (**a**) pH dependence of hydrolysis of MUFA as the substrate at 50°C. (**b**) Semilog plots of temperature dependence of hydrolysis of MUFA by the four proteins measured at the optimum pH of 8.0. The optimum temperature for TmAcE and ΔHT-TmAcE is assumed to be 100°C while the values for Δ4-TmAcE and Δ10-TmAcE are 70 and 60°C, respectively.

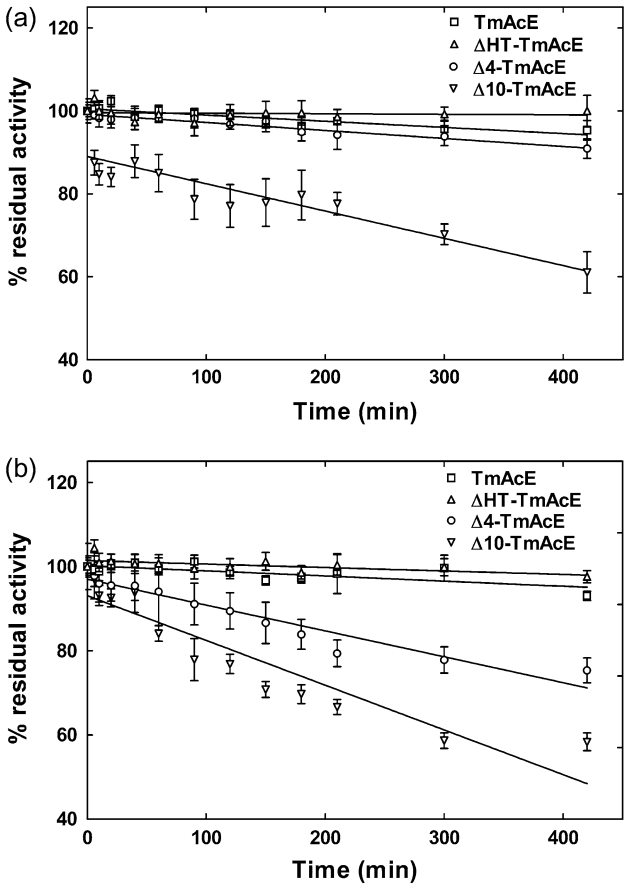

Fig. 5 Thermal deactivation kinetics of TmAcE and variants at (**a**) 60°C (**b**) 70°C. The enzyme solution was incubated at 60 and 70°C for different time periods and the residual activity of the sample after heat treatment was measured using MUFA as the substrate at pH 8.0 and 50°C.

are stable at both 60 and 70°C for 7 h, retaining ~98% of their initial activity. Δ4-TmAcE at 60°C retains greater than 90% of its initial activity after 7 h of incubation, whereas at 70°C, ~75% of residual activity is retained after the same time period. However, Δ10-TmAcE shows larger decrease in the activity, retaining ~60% of the residual activity after 7 h of incubation at both 60 and 70°C. Thus, the four and ten residue deletions have significantly disrupted both the optimum temperature and the thermal stability of the enzyme. Since these effects could likely be the result of conformational changes affecting the secondary, tertiary or quaternary structures, changes in the structure and the thermal stability of these proteins were studied by far-UV CD spectroscopy and DSC.

Secondary structure and thermostability of the N-terminal variants

Far-UV CD spectra of the proteins were measured at different temperatures (5–90°C) to evaluate the effect of temperature on the secondary structure (Fig. 6). Comparison of the spectra at 30°C showed minimal differences suggesting that the overall secondary structure content of the proteins are unaffected by the deletions. The data also show that the conformational changes in TmAcE, ΔHT-TmAcE and Δ4-TmAcE as a result of increase in temperatures, are comparable. However, Δ10-TmAcE is more thermo-sensitive compared to the others and shows loss of secondary structure at temperatures higher than 75°C. Thermal unfolding of the proteins was measured by monitoring changes in CD ellipticity at 222 nm (Fig. 7). The ellipticity values of TmAcE, ΔHT-TmAcE and Δ4-TmAcE show a moderate linear decrease with increase in temperature up to 95°C indicating that the secondary structure is not considerably affected and that the unfolding transition occurs at temperatures higher than 95°C. However, in contrast, a single transition is observed in Δ10-TmAcE, leading to a significant loss in the ellipticity at temperatures higher than 75°C. The behavior closely represents a two-state transition between a folded and an unfolded form.

Next, thermally induced unfolding of the proteins were investigated using DSC. Since the denaturation process was irreversible for all four proteins, the reported transition temperatures of the unfolding process are apparent values and are used here only for relative stability comparisons (Barone *et al.*, 1993). Inspection of the thermograms of TmAcE, ΔHT-TmAcE and Δ4-TmAcE indicates that the unfolding process has multiple transitions. Deconvolution of the DSC data calculated using an independent two-state transition model indicates two transitions for TmAcE, ΔHT-TmAcE and Δ4-TmAcE (Fig. 8). The numerical values of the thermal midpoints, $T_{m,i}$, for the two transitions are ~99 and 103°C for TmAcE and ΔHT-TmAcE. In the case of Δ4-TmAcE, two transitions at ~84 and 99°C, are observed. The multiple transitions probably occur due to independent

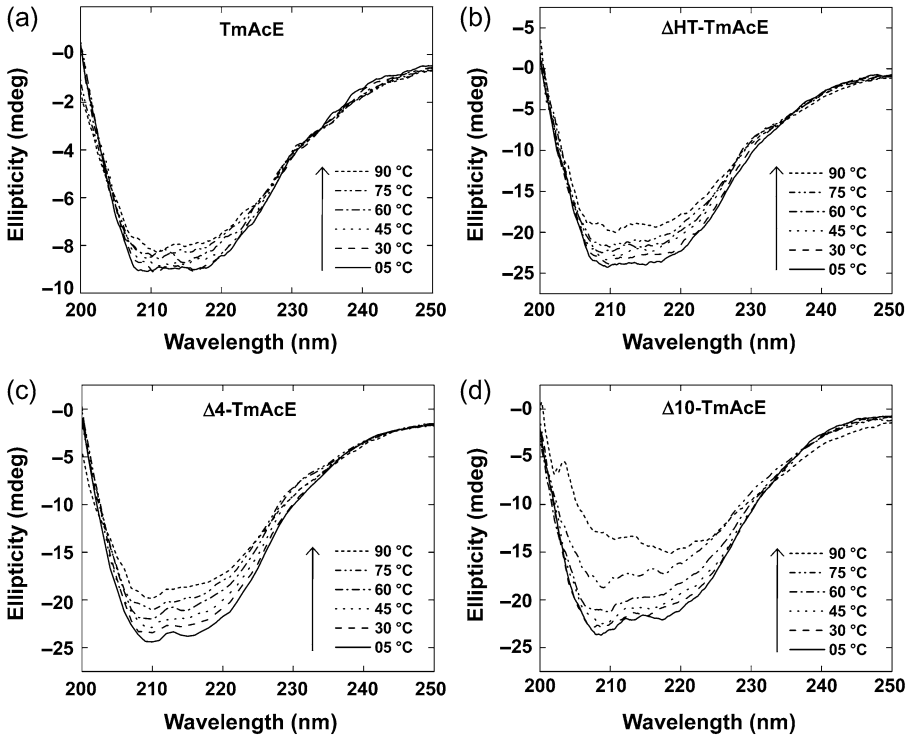

Fig. 6 Effect of temperature on secondary structure conformation of TmAcE and variants. Far-UV CD spectra of (**a**) TmAcE (**b**) ΔHT-TmAcE (**c**) Δ4-TmAcE (**d**) Δ10-TmAcE. Spectra were recorded at different temperatures in the range 5–90°C. Δ10-TmAcE shows significant loss of secondary structure at temperatures higher than 75°C, whereas the spectra for TmAcE, ΔHT-TmAcE and Δ4-TmAcE are similar, with little loss of secondary structure up to 95°C.

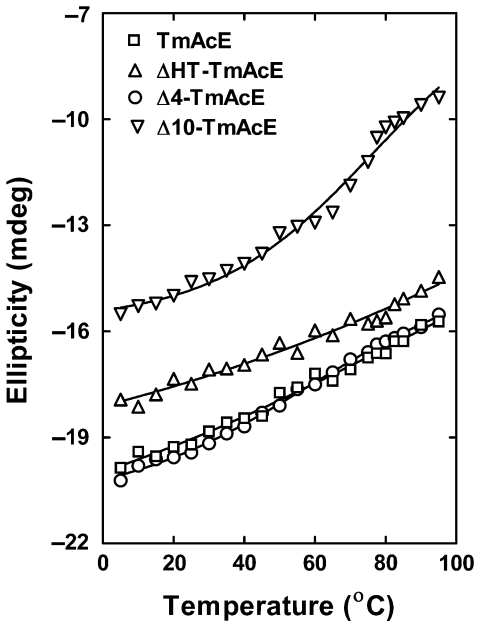

Fig. 7 Thermal stability of TmAcE and variants. Thermal unfolding curves of TmAcE and variants at pH 8.0 as obtained by recording the change in ellipticity at 222 nm for temperatures in the range 5–95°C. TmAcE, ΔHT-TmAcE and Δ4-TmAcE show a linear reduction in the CD signal strength indicating a transition from one stable state to another, with unfolding occurring at a temperature higher than 95°C. A thermal unfolding transition for Δ10-TmAcE at ~77°C was calculated using the Boltzmann sigmoid equation, assuming a two-state model.

dissociation of the structural components in the dynamic oligomer that includes a mixture of hexameric and dimeric and/or trimeric species in these proteins. Interestingly, the thermal unfolding profile of Δ10-TmAcE indicates a single component transition that fits a simple two-state model of unfolding with an apparent T_m at 81°C (Fig. 8d). In all samples, the DSC-derived denaturing temperatures are in close agreement with the melting temperatures obtained from the CD studies. The results so far show that deletion of the tag from TmAcE does not affect thermal stability or secondary structure of the protein. Furthermore, Δ4-TmAcE shows a marginal decrease in T_m whereas the decrease is larger in the case of Δ10-TmAcE. The single component conformational transition observed in Δ10-TmAcE is consistent with the distinctive oligomeric composition observed in the SEC data.

Inhibition kinetics

The catalytic site of TmAcE contains the classical Ser-His-Asp triad, with Ser188 acting as the nucleophile (Fig. 1). Previous inhibition kinetics studies have established that paraoxon (diethyl 4-nitrophenyl phosphate) is a competitive irreversible inhibitor that binds to TmAcE with a dissociation constant of ~0.5 mM. The crystal structure of the paraoxon bound form (TmAcE-DEP, PDB: 3M83) shows that the diethyl-phosphate moiety is indeed covalently bound to Ser188 (Levisson *et al.*, 2012). In order to examine whether the N-terminal region affects the access of paraoxon for the active site, inhibition kinetics studies were performed. The enzyme was incubated with varying concentrations of inhibitor and the specific activity measured using MUFA as the substrate. The IC_{50} values for inhibition ranged from 68 to 106 μM (Fig. 9). The decrease in

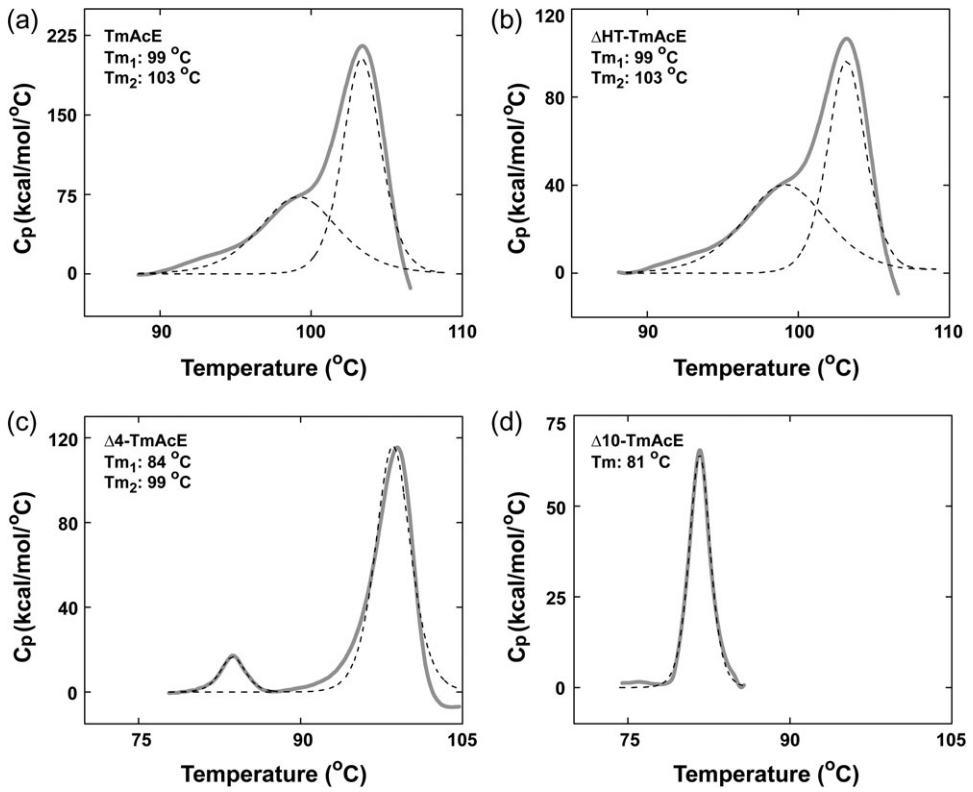

Fig. 8 Thermal unfolding of TmAcE and variants as measured by DSC. DSC thermograms of each protein (solid gray line) and the subsequent deconvolution analysis of thermal unfolding transitions (broken line) are shown. (**a**) Two transitions in TmAcE occur at 99°C and 103°C. (**b**) Two transitions in ΔHT-TmAcE occur at 99 and 103°C. (**c**) Two transitions in Δ4-TmAcE occur at 84 and 99°C. (**d**) A single thermal unfolding transition in Δ10-TmAcE occurs at 81°C.

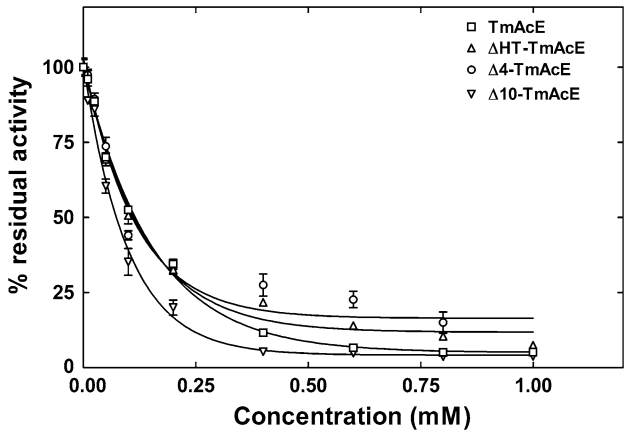

Fig. 9 Kinetics of inhibition of esterase activity of TmAcE and variants by paraoxon. Activity of paraoxon treated enzyme samples as a function of paraoxon concentration using MUFA as a substrate is shown. The activity values are shown relative to the specific activity of the untreated sample. IC_{50} values were obtained from fitting data to a one-phase decay exponential equation. The IC_{50} values are 106 ± 2.5 μM (TmAcE), 91 ± 3.7 μM (ΔHT-TmAcE), 80 ± 7.0 μM (Δ4-TmAcE) and 68 ± 3.5 μM (Δ10-TmAcE).

IC_{50} value is directly correlated to the size of the deletion. The tagged TmAcE was the least sensitive to inhibition and the Δ10-TmAcE showed the most sensitivity, suggesting that diffusion of the inhibitor into the active site is probably affected by the N-terminal region.

Substrate specificity and kinetic parameters

A key issue in this study was to explore the effect of the deletions on the activity and the substrate specificity of TmAcE. These properties were examined by steady state kinetic studies using five substrates. These include three generic substrates, MUFA, PNPA and 7-ACA and two specific carbohydrate substrates, glucose penta-acetate (5-Ac-Glu), xylose tetra-acetate (4-Ac-Xy). Enzyme kinetic parameters of the samples were measured at pH 8.0 and 50°C for all substrates. All four proteins displayed Michaelis-Menten kinetics with all substrates tested and the measured kinetic parameters are summarized in Table I. The highest catalytic efficiency (k_{cat}/K_M) for the four proteins is observed for the carbohydrate esters (4-Ac-Xy, 5-Ac-Glu) and is predominantly an attribute of high turnover numbers (k_{cat}). The values for the generic esters are far lower (~5.8–18-fold) in comparison. All four proteins showed the least specific activity and specificity towards 7-ACA (~9–320-fold lower), largely due to poor affinity (K_M) for the substrate.

Among the four constructs, the tag-free native form (ΔHT-TmAcE) exhibits the highest catalytic efficiencies for the generic esters, ~40–90% higher compared to the others, with variations observed both in the V_{max} and K_M values. However, among the two carbohydrate esters, the four proteins display minimal differences in the specificity constants for 5-Ac-Glu. Interestingly, Δ10-TmAcE displayed a 1.4 to 1.8-fold higher catalytic efficiency for the 4-Ac-Xy substrate compared to the others (Fig. 10). This is largely the result of a lower K_M value (2.3–3.2-fold) for this substrate compared to the others.

Table I. Steady state kinetic parameters of TmAcE and N-terminal variants[a]

Substrate	Enzyme	V_{max} (U/mg)	K_M (μM)	$k_{cat} \times 10^2$ (min^{-1})	$k_{cat}/K_M \times 10^4$ ($s^{-1}M^{-1}$)
PNPA	TmAcE	95.2 ± 1.2	175.1 ± 6.9	36.2 ± 0.4	34.4 ± 0.9
	ΔHT-TmAcE	155.3 ± 1.9	193.4 ± 6.7	57.7 ± 0.7	49.7 ± 1.1
	Δ4-TmAcE	79.7 ± 2.1	230.6 ± 19.4	29.2 ± 0.8	21.1 ± 1.9
	Δ10-TmAcE	80.0 ± 0.9	309.0 ± 9.3	28.7 ± 0.3	15.5 ± 0.8
MUFA	TmAcE	67.4 ± 0.7	147.6 ± 4.2	25.6 ± 0.2	28.9 ± 0.5
	ΔHT-TmAcE	101.8 ± 2.1	197.2 ± 10.2	37.8 ± 0.7	32.0 ± 0.1
	Δ4-TmAcE	65.2 ± 1.7	174.8 ± 14.0	23.9 ± 0.6	22.8 ± 1.9
	Δ10-TmAcE	53.5 ± 1.4	187.3 ± 12.8	19.3 ± 0.5	17.1 ± 1.3
5-Ac-Glu	TmAcE	502.3 ± 22.3	207.4 ± 23.2	193.9 ± 8.6	155.8 ± 18.8
	ΔHT-TmAcE	560.4 ± 25.2	223.6 ± 24.7	208.2 ± 9.4	155.5 ± 18.5
	Δ4-TmAcE	483.4 ± 17.9	202.9 ± 19.1	177.3 ± 6.5	145.6 ± 14.7
	Δ10-TmAcE	489.0 ± 20.7	206.9 ± 22.1	175.9 ± 7.4	141.8 ± 16.3
4-Ac-Xy	TmAcE	436.2 ± 14.1	133.7 ± 13.1	168.4 ± 5.4	208.3 ± 21.6
	ΔHT-TmAcE	526.5 ± 22.4	189.0 ± 21.0	195.6 ± 8.3	172.5 ± 21.6
	Δ4-TmAcE	436.2 ± 13.3	133.8 ± 12.1	160.0 ± 5.0	199.4 ± 19.0
	Δ10-TmAcE	283.2 ± 6.7	59.3 ± 5.6	101.9 ± 2.4	286.6 ± 18.8
7-ACA	TmAcE	38.7 ± 3.6	1383.5 ± 238.1	14.9 ± 1.4	1.8 ± 0.4
	ΔHT-TmAcE	38.6 ± 3.6	1135.2 ± 204.1	14.3 ± 1.3	2.1 ± 0.4
	Δ4-TmAcE	33.0 ± 2.2	1119.9 ± 155.1	12.1 ± 0.8	1.8 ± 0.3
	Δ10-TmAcE	25.9 ± 2.6	1562.7 ± 275.8	9.3 ± 0.9	0.9 ± 0.2

[a]Reactions were carried out in triplicates as described in the materials and methods section. Errors shown are standard errors of three independent measurements.

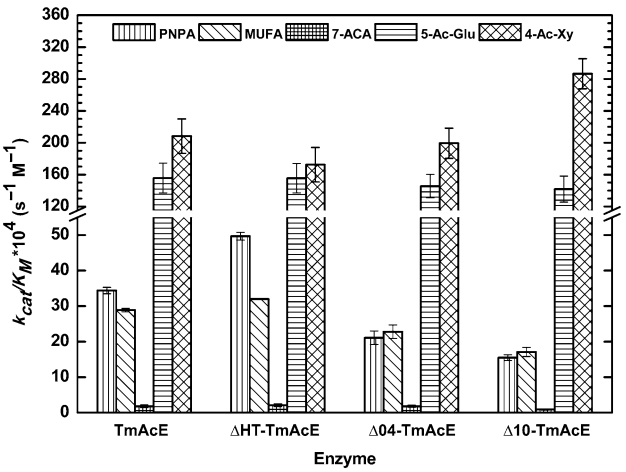

Fig. 10 Comparison of kinetic parameters of TmAcE and variants for different substrates. The parameters were obtained from steady state kinetics experiments performed at pH 8.0, 50°C and corresponding constants were calculated by fitting the data to Michaelis-Menten plots. The histograms represent the catalytic efficiency (k_{cat}/K_M) of the four proteins for different substrates and indicate substrate preferences.

Discussion

Since the identification of hyperthermophilic proteins, numerous studies have been carried out to unravel the structural basis for thermostability. In fact, proteins from *T. maritima* have been frequently adopted as model systems for these studies (Szilágyi and Závodszky, 2000; Robinson-Rechavi *et al.*, 2006). The molecular mechanisms for thermostability of proteins are diverse and protein dependent. Nonetheless, some general features are known to determine higher stability of thermophilic proteins compared to their mesophilic homologs. The major factors include larger number of non-covalent interactions, higher order oligomerization, metal coordination interactions, post-translational modifications, disulfide bonds and better conformational parameters that include enhanced core packing, lower entropy of unfolding and higher secondary structure content (Russell and Taylor, 1995; Petsko, 2001; Tanaka *et al.*, 2004; Li *et al.*, 2005; Razvi and Scholtz, 2006; Sadeghi *et al.*, 2006; Devenish and Gerrard, 2009; Nishi *et al.*, 2013; Reed *et al.*, 2013; Marsh and Teichmann, 2014).

The unusual ability to deacetylate CPC and 7-ACA make the CE7 family an industrially important class of enzymes. The study reported here was carried out to decipher the role of the CE7 N-terminus in the catalytic properties and the thermal stability. The N-terminal extension region is an evolutionarily conserved structural feature that is unique to the CE7 family among the paralogous CE families in the large and divergent α/β hydrolase superfamily. This observation is consistent with numerous comparative studies on the molecular evolution of protein superfamilies which have revealed that independent deletion and/or insertion events at multiple positions on a core conserved fold are pervasive adaptations across paralogous families. The consequential conformational changes arising out of such variability result in the evolution of divergent oligomerization dynamics, stability and biochemical activity (Borchert *et al.*, 1994; D'Alessio, 1999; Schmitz *et al.*, 2008; Hashimoto and Panchenko, 2010; Hashimoto *et al.*, 2010; Mehl *et al.*, 2011; Thoms *et al.*, 2011; Reich *et al.*, 2014). In the CE7 hexamer, the buried active sites are accessible only through an axial tunnel that runs along the 3-fold axis. In contrast to the CEs with exposed active sites, the oligomeric assembly in CE7 probably results in more protein-substrate interactions as the substrate diffuses into the active site. The N-terminal residues occluding the entrance mouths of this tunnel are exposed to the bulk solvent and was proposed to act as a 'selectivity' filter that could regulate substrate access to the active sites based on size limitation (Vincent *et al.*, 2003). Similar strategies for compartmentalization of active sites by oligomerization and/or in combination with specific structural features that act as 'gates' to control ligand access though a tunnel have been observed in several

enzymes. These tunnels and the gating elements are known to affect structure-function relationships, including substrate specificity and selectivity, kinetic parameters, cofactor transport and product egress (reviewed in Gora *et al.*, 2013; Kingsley and Lill, 2015). For instance, in the broad specificity cytochrome P450 superfamily, a common dynamic gated channel mechanism controls the diverse range of substrates by kinetically modulating enzyme-substrate specificity (Winn *et al.*, 2002). On the other hand, the tunnel gate in NiFe hydrogenases can distinguish between comparable molecules such as O_2 and CO (Liebgott *et al.*, 2010). Deciphering and understanding the structural basis of enzyme gates can thus be used to direct rational strategies for gate engineering to generate enzymes with desirable catalytic properties. A structurally analogous example to the CE7 family is that of the hexameric ATP-dependent Lon protease with a sequestered active site accessible through an axial tunnel. Here, two conserved loops serve as a gating mechanism to regulate substrate access to the active site. Mutation of a Phe residue to Ala in one surface exposed loop resulted in almost complete loss of proteolytic activity (Cha *et al.*, 2010).

Although four CE7 orthologs have been characterized so far, information about the role of the N-terminal region is unavailable. Hence, deletion variants of this region of TmAcE were generated and characterized for their *in vitro* catalytic properties and stability. The protein purification process shows that deletion of the 4 residues that constitute the 'selectivity' filter and the expanded 10 residue region spanning an α-helix, do not affect the expression of the protein. The presence or absence of the N-terminal 12 residue expression and purification tag had little or no effect on oligomerization, pH and temperature optimum values, thermal stability and inhibitor IC_{50} properties. Moreover, SEC studies show that the expected mixture of a minor dimeric and a major hexameric state observed in TmAcE and ΔHT-TmAcE is maintained in Δ4-TmAcE, whereas the Δ10-TmAcE appears to largely contain a dodecameric oligomeric state. The 10 residue deletion may have resulted in the exposure of a hydrophobic surface on the quaternary structure, leading to formation of a higher oligomer (Rezaei-Ghaleh *et al.*, 2008; Philo and Arakawa, 2009). The results strongly imply that the N-terminal does have a function in preserving the optimal dynamic quaternary assembly of the native enzyme.

The identical pH optimum profile for specific activity of all constructs indicates that the optimum pH is primarily a function of catalytic residues. Kinetic studies showed that the deletion of the tag (ΔHT-TmAcE) resulted in a minor increase (~10–40%) in catalytic efficiency (k_{cat}/K_M) only for the generic esters. The data indicate that the tag is conformationally flexible and that these residues are unlikely to have any interactions with the native subunit, consistent with the observation of a disordered tag in all available crystal structures. Contrary to our expectation that engineering an open access tunnel will lead to enhancement of activity and/or modification of substrate selectivity, the deletion of the N-terminal region had only modest effects on the catalytic parameters. Evidently, other factors, for instance, interactions of residues lining the entire tunnel to the catalytic site may also be crucial. All four proteins showed the highest catalytic efficiency for the two carbohydrate esters indicating that the overall enzyme specificity remains unchanged. Furthermore, the absolute values of the specificity constant for acetylated glucose are also very similar across the four proteins. Interestingly, a noticeable increase in catalytic efficiency (1.8-fold) is observed in the Δ10-TmAcE for the xylose derivative. In this case, although the k_{cat} value has decreased in comparison to the other proteins, the K_M value shows a larger decrease. This suggests that the expanded 10 residue deletion may have caused some distinct changes to the microenvironment of the access tunnel and/or the active site region, which is reflected as a higher affinity (K_M) value specifically for the acetylated xylose rather than acetylated glucose. This finding is particularly pertinent considering that the putative physiological role of TmAcE is to carry out the intracellular deacetylation of acetylated xylooligosaccharides as a component of the xylan utilization pathway in *T. maritima* (Drzewiecki *et al.*, 2010). Thus, one premise, albeit speculative, is that the effect of the 'gate' modifications on substrate specificity and selectivity may be more prominent if kinetic measurements are performed with the physiologically relevant acetylated xylooligosaccharides.

The optimum temperature of ~100°C for TmAcE and ΔHT-TmAcE is significantly higher than those reported so far for CE7 members, including the *Thermoanaerobacterium* CE7 that has a temperature optimum of 80°C (Shao and Wiegel, 1995). However, the removal of either 4 or 10 residues from the N-terminus drastically affected the temperature optimum for maximal activity, which are lower than the tag-free full length protein by 30 and 40°C, respectively. The sharp decrease suggests that the N-terminus plays a crucial role in maintaining the conformational stability of the protein at high temperatures. Kinetic thermal stability studies confirm that the Δ4-TmAcE and Δ10-TmAcE variants are far more susceptible to denaturation at higher temperatures, with the Δ10-TmAcE variant displaying the most sensitivity. Biophysical characteristics explored using CD spectra and DSC thermograms indicate that TmAcE, ΔHT-TmAcE and Δ4-TmAcE followed the same pattern of conformational changes as a function of temperature. The absence of sharp transitions in these three proteins with increasing temperature probably arise from transitions from one stable state to another or due to structural contraction of the protein that may help in retaining structure and attain hyperthermophilicity (Giordano *et al.*, 2005). The DSC unfolding thermograms indicate multi-phasic unfolding behavior with at least two intermediates, most likely related to the transitions from the hexameric to dimeric/trimeric form and further to the monomeric form which then unfolds at a higher temperature. This complex behavior is consistent with the dynamic oligomeric nature that is shared by the three proteins. However, Δ10-TmAcE shows a sharp two-state transition in the unfolding process in both the CD and DSC analyses, with apparent T_m values of 81°C. Apparently, the dissociation in the Δ10-TmAcE oligomer is so rapid that it does not have a limiting effect on protein denaturation. Alternatively, the dissociation and unfolding temperatures of the oligomer components may overlap (Santiago *et al.*, 2010). The evidence suggests that deletion of 10 residues has resulted in disruption of specific protein–protein interactions that affected the oligomeric assembly and thereby, its thermostability and thermal denaturation mechanism.

Protein folds in general are thermodynamically marginally stable (Taverna and Goldstein, 2002; Goldstein, 2011). The resulting activity-stability tradeoff thus limits efforts to enhance catalytic properties of enzymes by rational protein engineering since a favorable mutation may result in lower stability of the protein (Tokuriki *et al.*, 2008; Tokuriki and Tawfik, 2009; Soskine and Tawfik, 2010). Considering the crystal structure of TmAcE and the biophysical characteristics of its variants, it appears that the loss of favorable interactions between the N-terminal region and the rest of the protein in the variants has probably affected the oligomeric assembly causing the drastic loss in activity at temperatures higher than 60°C. Alternatively, given the minor changes in the secondary structures at high temperatures and the T_m value of 99°C for Δ4-TmAcE,

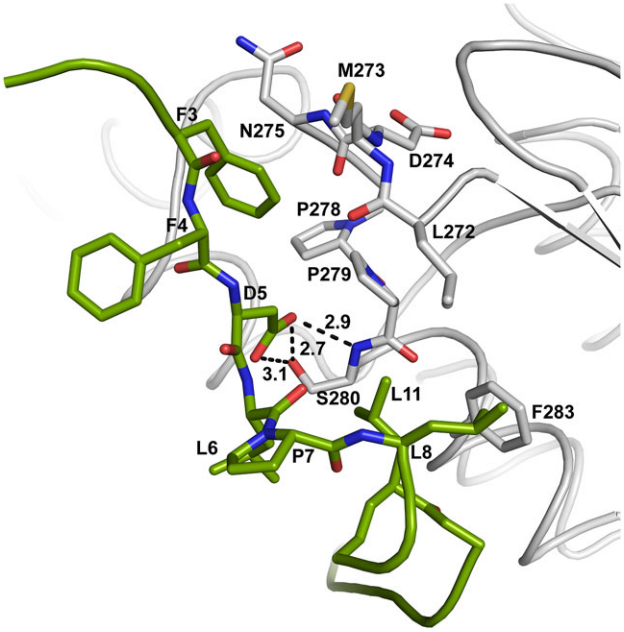

Fig. 11 Non-bonded interactions between the N-terminal extension region (green) and the neighboring protein segment containing residues 272–280 (gray) (PDB: 3M83). Asp274 corresponds to the aspartate of the Ser-His-Asp catalytic triad present in the CE7 family. The relevant residues are shown in stick representation. Hydrogen bonds are shown as dashed lines. The distances shown are in Å. The interactions between the N-terminal region and the loop (272–280) stabilizes the conformation of the catalytic aspartate and are required for the maintenance of catalytic activity at high temperatures.

the more plausible explanation for the decrease in the optimum temperature is that a restricted stability loss of the region around the active site is responsible for the effect. Evaluation of the non-bonded interactions of the relevant stretch of residues in the crystal structure of TmAcE provides a justification. More specifically, the first four residues are primarily involved in inter-subunit interactions whereas the residues 5–11, encompassing α-helix αA-2 participate in intra-subunit interactions, including three hydrogen bonds, with the neighboring loop (residues 273–283) that contains the catalytic triad residue, Asp274 (Fig. 11). Furthermore, the length and conformation of this loop are well conserved across the CE7 family. Apparently, these interactions together provide localized stabilization of the geometry of the catalytic triad that is probably lost at temperatures higher than 60°C. Structural and computational studies have shown that the geometry of the catalytic triad in serine esterases is typically preorganized and rigid (Smith *et al.*, 2008). It is noteworthy that the lack of four residues results in an almost equivalent degree of loss of optimum temperature as that observed for the ten residue deletion, thus, demonstrating the critical role of the N-terminal region in shaping the properties of the enzyme. In the absence of the N-terminal extension we can expect the stability of the protein to be severely compromised at the optimum growth temperature of 80°C for *T. maritima*. It is probable that the stability of CE7 enzymes from mesophilic organisms would suffer similar reduction from their respective optimum temperatures.

Conclusions

In CE7 enzymes, oligomerization is important for sequestering the active site which is otherwise promiscuous. Although the direct role of the N-terminal extension region as a selectivity filter to regulate substrate access remains unclear, the first ten residues appear to influence the substrate specificity of TmAcE. This region, however, is critical for the maintenance of the competent oligomeric form and the thermodynamic and kinetic stability. The loss of hyperthermostability arising out of deletion of this region appears be an indirect effect of an increase in the conformational flexibility of an otherwise rigid neighboring loop containing a catalytic triad residue. Together, the data suggest that the N-terminal extension in the CE7 fold evolved primarily to modulate the substrate specificity, dynamics of the oligomeric assembly and the thermal stability of the enzyme.

Supplementary Data

Supplementary data are available at *Protein Engineering, Design and Selection* online.

Acknowledgements

We thank Infant Sagayaraj for preparing the figures. MKS received fellowship support from the CSIR, Government of India. Infrastructure support from the DST-FIST Facility, IIT Madras, is acknowledged.

Conflict of interest

The authors have no conflict of interest to declare.

Funding

This work was supported by the Department of Biotechnology, Ministry of Science and Technology, Government of India (BT/PR13805/PID/06/550/2010).

Authors' contributions

M.K.S. and S.S. performed the experiments. S.N.G. participated in the data analysis. N.M. conceived and designed the study, analyzed the data and wrote the manuscript.

References

Barone,G., Del Vecchio,P., Fessas,D., Giancola,C. and Graziano,G. (1993) *Thermochim. Acta*, **227**, 185–195.
Borchert,T.V., Abagyan,R., Jaenicke,R. and Wierenga,R.K. (1994) *Proc. Natl. Acad. Sci.*, **91**, 1515–1518.
Cha,S.S., An,Y.J., Lee,C.R. *et al* (2010) *EMBO J.*, **29**, 3520–3530.
D'Alessio,G. (1999) *Prog. Biophys. Mol. Biol.*, **72**, 271–298.
Degrassi,G., Kojic,M., Ljubijankic,G. and Venturi,V. (2000) *Microbiology*, **146**, 1585–1591.
Degrassi,G., Okeke,B.C., Bruschi,C.V. and Venturi,V. (1998) *Appl. Environ. Microbiol.*, **64**, 789–792.
DeLano,W.L. (2006). The PyMOL Molecular Graphics System, http://www.pymol.org.
Devenish,S.R.A. and Gerrard,J.A. (2009) *Org. Biomol. Chem.*, **7**, 833–839.
Drzewiecki,K., Angelov,A., Ballschmiter,M., Tiefenbach,K., Sterner,R. and Liebl,W. (2010) *Microb. Biotechnol.*, **3**, 84–92.
Fishelovitch,D., Shaik,S., Wolfson,H.J. and Nussinov,R. (2009) *J. Phys. Chem. B*, **113**, 13018–13025.
Ghosh,D., Erman,M., Sawicki,M. *et al* (1999) *Acta Crystallogr. D Biol. Crystallogr.*, **55**, 779–784.
Giordano,A., Febbraio,F., Russo,C. and Rossi,M. (2005) *Biochem. J.*, **388**, 657–667.
Goldstein,R.A. (2011) *Proteins*, **79**, 1396–1407.

Gora,A., Brezovsky,J. and Damborsky,J. (2013) *Chem. Rev.*, **113**, 5871–5923.
Hashimoto,K., Madej,T., Bryant,S.H. and Panchenko,A.R. (2010) *J. Mol. Biol.*, **399**, 196–206.
Hashimoto,K. and Panchenko,A.R. (2010) *Proc. Natl. Acad. Sci.*, **107**, 20352–20357.
Hedge,M.K., Gehring,A.M., Adkins,C.T., Weston,L.A., Lavis,L.D. and Johnson,R.J. (2012) *Biochim. Biophys. Acta (BBA)*, **1824**, 1024–1030.
Kingsley,L.J. and Lill,M.A. (2015) *Proteins.*, **83**, 599–611.
Klock,H.E., Koesema,E.J., Knuth,M.W. and Lesley,S.A. (2008) *Proteins Struct. Funct. Bioinforma.*, **71**, 982–994.
Klock,H.E. and Lesley,S.A. (2009) In Doyle,S.A. (ed), *Methods in Molecular Biology: High Throughput Protein Expression and Purification*. Humana Press, Totowa, NJ, pp. 91–103.
Kodama,Y., Masaki,K., Kondo,H., Suzuki,M., Tsuda,S., Nagura,T., Shimba,N., Suzuki,E. and Iefuji,H. (2009) *Proteins Struct. Funct. Bioinforma.*, **77**, 710–717.
Konecny,J. and Sieber,M. (1980) *Biotechnol. Bioeng.*, **22**, 2013–2029.
Krastanova,I., Cassetta,A., Wiegel,J. and Lamba,D. (2009). 10.2210/PDB3FCY/PDB.
Levisson,M., Han,G.W., Deller,M.C. *et al* (2012) *Proteins Struct. Funct. Bioinforma.*, **80**, 1545–1559.
Li,W.F., Zhou,X.X. and Lu,P. (2005) *Biotechnol. Adv.*, **23**, 271–281.
Liebgott,P.P., Leroux,F., Burlat,B. *et al* (2010) *Nat. Chem. Biol.*, **6**, 63–70.
Lorenz,W.W. and Wiegel,J. (1997) *J. Bacteriol.*, **179**, 5436–5441.
Marsh,J.A. and Teichmann,S.A. (2014) *PLoS Biol.*, **12**, e1001870.
Mehl,A.F., Nalin,U.G., Ahmed,Z., Wells,A. and Spyratos,T.D. (2011) *Int. J. Biol. Macromol.*, **48**, 627–633.
Montoro-García,S., Gil-Ortiz,F., García-Carmona,F., Polo,L.M., Rubio,V. and Sánchez-Ferrer,Á. (2011) *Biochem. J.*, **436**, 321–330.
Nishi,H., Hashimoto,K., Madej,T. and Panchenko,A.R. (2013) *Prog. Mol. Biol. Transl. Sci.*, **117**, 3–24.
Petsko,G.A. (2001) *Methods Enzym.*, **334**, 469–478.
Philo,J.S. and Arakawa,T. (2009) *Curr. Pharm. Biotechnol.*, **10**, 348–351.
Pokkuluri,P.R., Duke,N.E.C., Wood,S.J., Cotta,M.A., Li,X., Biely,P. and Schiffer,M. (2011) *Proteins Struct. Funct. Bioinforma.*, **79**, 2588–2592.
Pollegioni,L., Rosini,E. and Molla,G. (2013) *Appl. Microbiol. Biotechnol.*, **97**, 2341–2355.
Razvi,A. and Scholtz,J.M. (2006) *Protein Sci.*, **15**, 1569–1578.
Reed,C.J., Lewis,H., Trejo,E., Winston,V. and Evilia,C. (2013) *Archaea*, **2013**, 373275.
Reich,S., Kress,N., Nestl,B.M. and Hauer,B. (2014) *J. Struct. Biol.*, **185**, 228–233.
Rezaei-Ghaleh,N., Ramshini,H., Ebrahim-Habibi,A., Moosavi-Movahedi,A.A. and Nemat-Gorgani,M. (2008) *Biophys. Chem.*, **132**, 23–32.
Robinson-Rechavi,M., Alibés,A. and Godzik,A. (2006) *J. Mol. Biol.*, **356**, 547–557.
Russell,R.J.M. and Taylor,G.L. (1995) *Curr. Opin. Biotechnol.*, **6**, 370–374.
Sadeghi,M., Naderi-Manesh,H., Zarrabi,M. and Ranjbar,B. (2006) *Biophys. Chem.*, **119**, 256–270.
Santiago,P.S., Carvalho,J.W.P., Domingues,M.M., Santos,N.C. and Tabak,M. (2010) *Biophys. Chem.*, **152**, 128–138.
Schmitz,K.R., Liu,J., Li,S., Setty,T.G., Wood,C.S., Burd,C.G. and Ferguson,K.M. (2008) *Dev. Cell.*, **14**, 523–534.
Schubot,F.D., Kataeva,I.A., Blum,D.L., Shah,A.K., Ljungdahl,L.G., Rose,J.P. and Wang,B.C. (2001) *Biochemistry*, **40**, 12524–12532.
Shao,W. and Wiegel,J. (1995) *Appl. Environ. Microbiol.*, **61**, 729–733.
Silvia,M.G., Fernando,G.O., Francisco,G.C., Luis,M.P., Vicente,R. and Alvaro,S.F. (2011) *Biochem. J.*, **436**, 321–330.
Singh,M.K. and Manoj,N. (2016a) *J. Struct. Biol.*, **194**, 434–445.
Singh,M.K. and Manoj,N. (2016b) *Biochem. Biophys. Res. Commun.*, **476**, 63–68.
Singh,M.K. and Manoj,N. (2017) *Proteins Struct. Funct. Bioinforma*, **85**, 694–708.
Smith,A.J.T., Müller,R., Toscano,M.D., Kast,P., Hellinga,H.W., Hilvert,D. and Houk,K.N. (2008) *J. Am. Chem. Soc.*, **130**, 15361–15373.
Sonawane,V.C. (2006) *Crit. Rev. Biotechnol.*, **26**, 95–120.
Soskine,M. and Tawfik,D.S. (2010) *Nat. Rev. Genet.*, **11**, 572–582.
Szilágyi,A. and Závodszky,P. (2000) *Structure*, **8**, 493–504.
Takimoto,A., Yagi,S. and Mitsushima,K. (1999) *J. Biosci. Bioeng.*, **87**, 456–462.
Tanaka,Y., Tsumoto,K., Yasutake,Y., Umetsu,M., Yao,M., Fukada,H., Tanaka,I. and Kumagai,I. (2004) *J. Biol. Chem.*, **279**, 32957–32967.
Taverna,D.M. and Goldstein,R.A. (2002) *Proteins*, **46**, 105–109.
Thoms,S., Hofhuis,J., Thöing,C., Gärtner,J. and Niemann,H.H. (2011) *J. Struct. Biol.*, **175**, 362–371.
Tian,Q., Song,P., Jiang,L., Li,S. and Huang,H. (2014) *Appl. Microbiol. Biotechnol.*, **98**, 2081–2089.
Tokuriki,N., Stricher,F., Serrano,L. and Tawfik,D.S. (2008) *PLoS Comput. Biol.*, **4**, e1000002.
Tokuriki,N. and Tawfik,D.S. (2009) *Curr. Opin. Struct. Biol.*, **19**, 596–604.
Vincent,F., Charnock,S.J., Verschueren,K.H.G. *et al* (2003) *J. Mol. Biol.*, **330**, 593–606.
Winn,P.J., Ludemann,S.K., Gauges,R., Lounnas,V. and Wade,R.C. (2002) *Proc .Natl. Acad. Sci.*, **99**, 5361–5366.