J Ind Microbiol Biotechnol (2015) 42:1493–1506
DOI 10.1007/s10295-015-1690-x

BIOTECHNOLOGY METHODS

Sequence and structure-based comparative analysis to assess, identify and improve the thermostability of penicillin G acylases

Priyabrata Panigrahi[1] · Deepak Chand[1] · Ruchira Mukherji[1] · Sureshkumar Ramasamy[1] · C. G. Suresh[1]

Received: 4 May 2015 / Accepted: 14 September 2015 / Published online: 29 September 2015

Abstract Penicillin acylases are enzymes employed by the pharmaceutical industry for the manufacture of semi-synthetic penicillins. There is a continuous demand for thermostable and alkalophilic enzymes in such applications. We have carried out a computational analysis of known penicillin G acylases (PGAs) in terms of their thermostable nature using various protein-stabilizing factors. While the presence of disulfide bridges was considered initially to screen putative thermostable PGAs from the database, various other factors such as high arginine to lysine ratio, less content of thermolabile amino acids, presence of proline in β-turns, more number of ion-pair and other non-bonded interactions were also considered for comparison. A modified consensus approach designed could further identify stabilizing residue positions by site-specific comparison between mesostable and thermostable PGAs. A most likely thermostable enzyme identified from the analysis was PGA from *Paracoccus denitrificans* (*Pd*PGA). This was cloned, expressed and tested for its thermostable nature using biochemical and biophysical experiments. The consensus site-specific sequence-based approach predicted *Pd*PGA to be more thermostable than *Escherichia coli* PGA, but not as thermostable as the PGA from *Achromobacter xylosoxidans*. Experimental data showed that *Pd*PGA was comparatively less thermostable than *Achromobacter xylosoxidans* PGA, although thermostability factors favored a much higher stability. Despite being mesostable, *Pd*PGA being active and stable at alkaline pH is an advantage. Finally, several residue positions could be identified in *Pd*PGA, which upon mutation selectively could improve the thermostability of the enzyme.

Keywords Thermostability · Penicillin acylase · Consensus approach · Alkalistable · Ion pairs · β-lactam antibiotics · Disulfide bridge

Introduction

Penicillin G acylase (PGA, E.C. 3.5.1.11) is an important biocatalyst, commonly employed in the pharmaceutical industry for the enzymatic deacylation of penicillin G to yield phenylacetic acid and 6-aminopenicillanic acid (6-APA). The product 6-APA is a key intermediate in the large-scale production of many semi-synthetic penicillins, which are more effective against resistant pathogens compared to natural penicillins [1]. These enzymes are also employed in resolving the racemic mixtures of chiral compounds such as secondary alcohols [10] and protection of amino and hydroxyl groups in peptide synthesis [9]. One of the classic examples where PGAs are used for peptide synthesis was in the production of artificial sweetener aspartame [9]. Other examples include the use of PGA from *Escherichia coli* (*Ec*PGA) for kinetically controlled synthesis of chiral dipeptides of phenylglycine such as D-phenylglycyl-L-phenylglycine and L-phenylglycyl-L-phenylglycine methyl esters, which further undergo cyclicization to form diketopiperazines. These diketopiperazines are used as chitinase inhibitors, food additives and synthons for preparing antiviral, fungicidal and anti-allergic compounds [31]. PGAs have also been employed for biocatalytic resolution

Electronic supplementary material The online version of this article (doi:10.1007/s10295-015-1690-x) contains supplementary material, which is available to authorized users.

✉ C. G. Suresh
cg.suresh@ncl.res.in

[1] Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Dr. Homi Bhabha Road, Pune 411008, India

of chiral amines [13], 1-phenylethylamine [12] and (R,S)-phenylglycinonitrile [6].

The enzyme is produced as an inactive precursor (signal peptide, α-chain, spacer peptide and β-chain) that undergoes post-translational removal of signal and spacer peptides, finally maturing into an active heterodimer consisting of α- and β-chains. This maturation event results in the creation of a free N-terminal serine residue in the β-chain (βSer1) which acts as nucleophile during hydrolysis of the substrate [14]. The three-dimensional structure of the active enzyme consists of the αββα Ntn hydrolase structural fold. Due to the presence of these hallmark features, PGAs are classified as members of the N-terminal Ser nucleophile (NtSn) hydrolase superfamily [20].

As the rate of any enzymatic reaction increases with temperature, the use of PGAs as commercial biocatalyst would become far more lucrative if their stability and activity at higher temperatures are improved. Several investigations in the past have been focused on enhancing the thermostability of *Ec*PGA, a widely used PGA for the manufacture of β-lactam antibiotics, through cross-linking the enzyme with glutaraldehyde [8] or carrying out site-directed mutagenesis of few carefully selected amino acid residues [22]. However, these efforts seem to result only in minor enhancement in thermostability. Hence, efforts have also been undertaken to identify novel sources of thermostable PGAs such as PGA from *Alcaligenes faecalis* (*Af*PGA) [33] and *Achromobacter xylosoxidans* (*Ax*PGA) [4] which are comparatively more stable than *Ec*PGA. In case of *Af*PGA and *Ax*PGA, the half-life of the enzymes at 55 °C has been observed to be 15 and 55 min, respectively. It has been experimentally shown that the higher thermostability of *Af*PGA can be attributed to the presence of a disulfide bond in its structure [32, 33]. Interestingly, *Ax*PGA, which is more thermostable than *Af*PGA, lacks any disulfide bond. Factors such as preference of Arg residues over Lys, decrease in the number of thermolabile amino acids and bonds, increase in proline residue content and the presence of more stable ion pairs have been suggested to play a role in the thermostability of *Ax*PGA [4]. Understanding the mechanism of increased thermostability among PGAs not only can help in the identification of novel sources of thermostable PGA, but can also aid in tailoring mesostable PGAs to improve their thermostability using protein engineering.

With the advances in high-throughput genome sequencing techniques and the availability of powerful functional annotation tools, PGAs have been predicted to occur in a wide range of microbial sources. However, the annotation of their biochemical and biophysical characteristics are currently unavailable. A preliminary computational screening of these PGAs for the identification of a putative thermostable candidate preceding their experimental characterization could prove to be beneficial with respect to both time and cost. Disulfide bond, being an experimentally proven factor for *Af*PGA thermostability, was considered as an initial criterion for the selection of putative thermostable PGA enzymes from the MEROPS [24] database. Enzymes from three different sources, namely *Sphingomonas wittichii* (*Sw*PGA), *Paracoccus denitrificans* (*Pd*PGA) and *Acinetobacter oleivorans* (*Ao*PGA), were identified as probable thermostable candidates based on the presence of the disulfide bond. These three were then compared with the already well-characterized PGAs, that is, *Ec*PGA (least thermostable/mesostable), *Af*PGA (moderately thermostable) and *Ax*PGA (most thermostable). Together, the six PGAs were first compared using a sequence-based consensus approach followed by structure-based comparative analysis, in an attempt to explore the applicability of various known protein thermostabilization mechanisms. The former analysis revealed that the three candidate enzymes could, at best, be only mesostable in nature. However, the later structural analysis revealed that *Pd*PGA might be a potentially thermostable enzyme. Since experimental evidence is lacking to verify the prediction, we have undertaken to clone the *pdpga* gene, express and purify the *Pd*PGA enzyme, followed by experimental characterization for checking its thermostability.

Materials and methods

Computational screening strategy for identifying ptPGAs (putative thermostable PGAs)

Three-dimensional structures of *Ec*PGA (PDB ID: 1GK9) and *Af*PGA (PDB ID: 3K3W) were extracted from PDB [2]. The subfamily S45.001 (penicillin G acylase precursor subfamily) of MEROPS database was selected for computational screening of putative thermostable PGAs (ptPGAs). Each of the PGA sequence from the database was aligned with *Af*PGA and only those sequences having Cys residues at equivalent positions facilitating disulfide bond formation were selected. The selected putative sequences (ptPGAs) were in their precursor form and the locations of their signal and spacer peptides were unknown. Since the active form of PGAs does not contain the signal and spacer peptides, these regions were identified by comparing the sequences with active and processed form of *Ec*PGA and *Af*PGA and were removed from ptPGAs.

In the absence of three-dimensional structures of *Ax*PGA and ptPGAs, high-resolution crystal structure of *Ec*PGA was used for building three-dimensional homology models using Prime 3.0 (Version 3.1, Schrödinger, LLC, New York, NY, 2012) and validated by model validation programs (Table S2). The effect of temperature on

the stability of PGA enzymes was studied by explicit solvent molecular dynamics simulations at three temperature scales (330, 400 and 500 K, respectively) for a timescale of 15 ns using OPLS-AA force field in Gromacs 4.5 [23]. The Tip4p solvent model was used for modeling the solvent and appropriate Na^+ and Cl^- ions were added to neutralize the charge in the system. The system was first subjected to steepest descent followed by conjugate gradient energy minimization. The resulting system was equilibrated in NVT ensemble for 100 ps at the respective temperatures using V-rescale temperature coupling. The system was further equilibrated with NPT ensemble for 100 ps at 1 atm pressure and respective temperatures, using Parrinello-Rahman pressure coupling. The equilibrated system was finally subjected to molecular dynamics simulation using leapfrog integrator.

Multiple sequence alignment of PGA sequences was carried out using ClustalX [29]. Intramolecular interactions in PGA enzyme structures were identified using iCAPS module of the iRDP Web server (http://irdp.ncl.res.in) using default parameters.

Cloning, expression and purification of PdPGA

The *pdpga* gene, encoded in the plasmid of *Paracoccous dentrificans PD1222*, was cloned in the pETKat vector (Gift from Dr. KatrinTiemann, Clemons Lab, Caltech). The insert has been amplified using gene-specific primers (forward primer: GAA AAC CTG TAC TTC CAG AGC ATG GGC ACC CAG GTC GAG and reverse primer: CCC TGA AAC AAG ACT TCC AAC CGC GGA ACG GCA AGG GTT T) and inserted into pETKat vector by PIPE cloning protocol [15]. In short, the pETKat vectors were PCR amplified with vector PIPE-for (5′-TTGGAAGTCTTGTTTCAGGGACCA-3′) and vector PIPE-rev (5′-GCTCTGGAAGTACAGGTTTTCACC-3′). 1.2 µl of insert and vector was mixed keeping on ice and 50 µl NovaBlue (Invitrogen) competent cells was added. This pETKat vector contains a suicide cassette, derived from pDest53 (Invitrogen) for better cloning efficacy. The sequence-confirmed clones were expressed in *E. coli* BL21-Gold(DE3) strain. The cell culture was incubated in a shaker incubator at 37 °C. Once the OD_{600} of the culture reached 0.8, the cells were induced with 0.75 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) and grown at 16 °C for 16–18 h. The cells were harvested by centrifugation at 5000 rpm for 30 min at 4 °C. The cell pellet obtained was resuspended in minimum volume of lysis buffer (25 mM Tris–HCl pH 7.5, 100 mM NaCl and 20 mM imidazole) and subjected to sonication followed by centrifugation at 12,000 rpm for 45 min at 4 °C. The C-terminal His-tagged *Pd*PGA protein was purified by using Ni–NTA affinity chromatography. The supernatant was loaded on to a column containing Ni^{+2}-Sepharose beads pre-equilibrated with equilibration buffer (25 mM Tris–HCl pH 7.5, 100 mM NaCl and 20 mM imidazole). The matrix was then washed with equilibration buffer followed by removal of non-specific and weakly bound proteins by washing with wash buffer (25 mM Tris–HCl pH 7.5, 100 mM NaCl and 50 mM imidazole). The matrix-bound *Pd*PGA enzyme was eluted by passing the elution buffer (25 mM Tris–HCl pH 7.5, 100 mM NaCl and 500 mM imidazole). The high concentration of imidazole in the elution fraction was removed by passing it through a PD10 desalting column and exchanged with glycine-NaOH buffer, pH 10. The protein solution obtained was finally subjected to gel-filtration chromatography using S200 column, pre-equilibrated with glycine–NaOH buffer, pH 10. Fractions containing PGA activity were pooled and analyzed by running 12 % SDS-PAGE and Western blot.

Penicillin G acylase activity

The enzyme activity of *Pd*PGA was determined by allowing the purified enzyme preparation (0.04 mg/ml) to act on its substrate penicillin G (20 mg/ml) in 50 mM phosphate buffer, pH 7.5, at 40 °C for 10 min. The reaction was quenched by adding 1 ml citrate phosphate buffer (300 mM citric acid in 50 mM phosphate buffer, pH 2.5). The product 6-APA released was estimated spectrophotometrically at 415 nm after reacting it with 2 ml of 0.6 % (w/v) *p*-dimethylaminobenzaldehyde in methanol. Readings of appropriate blanks and controls were used to subtract contribution due to any degraded products. The optimum pH of the *Pd*PGA activity was estimated between pH 3 and 12 under standard conditions, while optimal temperature for *Pd*PGA activity was determined over the temperature range of 25–60 °C in 50 mM buffer of pH 10. For studying the thermal stability of *Pd*PGA, enzyme samples were heated in the range of temperatures 35–50 °C for up to 30 min, followed by estimation of enzyme activity. The pH stability profile of *Pd*PGA was estimated by incubating the enzyme in buffers within pH range 3–12 for up to 3 h, followed by estimation of enzyme activity.

Thermal unfolding using steady-state fluorescence

The intrinsic fluorescence of *Pd*PGA was recorded using a Perkin Elmer LS50 fluorescence spectrophotometer attached to a Julabo F20 water bath. The enzyme (0.04 mg/ml) was excited at 295 nm followed by measurement of the emission spectra between 310 and 400 nm, while keeping the speed at 100 nm min^{-1} and slit width at 7 nm. Background emission from the buffer was subtracted. Thermal unfolding of *Pd*PGA was monitored by incubating the enzyme between 25 and 70 °C for 10 min followed by measurement of emission spectra. Protein aggregation due

to high temperature was monitored in the same range (25 and 70 °C) by Rayleigh scattering measurements on the same instrument.

Thermal unfolding monitored using CD spectroscopy

CD measurements of the purified *Pd*PGA were recorded using a Jasco J-815-150S (Jasco, Tokyo, Japan) spectropolarimeter connected to a Peltier CDFL cell circulating water bath. Far-UV spectra were recorded in a rectangular quartz cell of 1-mm path length in the range of 200–250 nm at a scan speed 100 nm/min with a response time of 1 s and slit width 1 nm. Purified *Pd*PGA at a concentration of 0.05 mg/ml was used for all measurements. Each spectrum was recorded as an average of five scans. Thermal denaturation studies of *Pd*PGA were carried out by incubating the enzyme at temperatures ranging between 25 and 70 °C for 10 min. Results were expressed as mean residue ellipticity (MRE) in deg cm^2/dmol defined as

$$\mathrm{MRE} = M\theta_{\lambda}/10dcr,$$

where M is the molecular weight of the protein, θ_{λ} is CD in millidegree, d is the path length in cm, c is the protein concentration in mg/ml and r is the average number of amino acid residues in the protein. The relative content of various secondary structure elements was calculated using the CDPro software (http://lamar.colostate.edu/~sreeram/CDPro/main.html). Low NRMSD values were observed for analysis with CONTINLL.

Results and discussions

The stability of an enzyme in various environmental conditions such as variant temperature and pH is an inherent property which depends not only on the three-dimensional structure, but also on its amino acid sequence. Thermostable enzymes offer industrial advantage over mesostable enzymes in terms of their storage, reaction rates and reduced microbial contaminations. Screening for a thermostable candidate in case of PGAs is a challenging problem both due to large size of the enzyme (750–850 residues) and its complex maturation process. While several molecular determinants have been proposed which contribute to thermodynamic stabilities of protein [34], the absence of a single dominant mechanism makes identification of a potentially stable enzyme candidate a complex task. In this work presented here, we have carried out both sequence- and structure-based thermostability analyses toward identification of potential thermostable PGA and experimentally verified using a candidate enzyme.

Putative thermostable PGAs (ptPGAs): screening and molecular modeling

The presence of a disulfide bond has been experimentally shown to provide thermostability in the case of *Af*PGA. The subfamily S45.001 of the MEROPS database consists of several PGA sequences of which *Sw*PGA, *Pd*PGA and *Ao*PGA (NCBI GI numbers 148553843, 119378111 and 299769954, respectively) were selected due to the presence of Cys residues at positions equivalent to that of *Af*PGA, which could lead to eventual formation of disulfide bond and possibly contribute toward thermostability of the enzymes. The selected three PGAs would be referred together as ptPGAs (putative thermostable PGAs). Although other PGA sequences in the MEROPS database were found to contain Cys residues, these were mainly found to occur in the signal peptide which would be cleaved off during post-translational processing (Table 1). The sequence renumbering of the mature form of PGA enzymes after removal of signal and spacer peptide is given in Table S1.

Table 1 PGAs from S45.001 family of the MEROPS database containing Cys residues are listed

MEROPS ID	Organism[b]	Cys residue position[a]
MER059680	*Sphingomonas wittichii (SwPGA)*	6, 27, **766, 799**
MER074190	*Paracoccus denitrificans (PdPGA)*	21, 675, **733, 766**
MER219436	*Acinetobacter oleivorans (AoPGA)*	27, **765, 798**
MER312023	*Acinetobacter baumannii*	27, **765, 798**
MER081546	*Acinetobacter baumannii*	52, **790, 823**
MER285059	*Acinetobacter calcoaceticus*	27, **765, 798**
MER238699	*Achromobacter piechaudii*	14
MER107787	*uncultured γ proteobacterium*	25
MER312015	*Marinobacterium stanieri*	242
MER087144	*Serratia proteamaculans*	19
MER107793	*Achromobacter sp. CCM 4824*	16, 782
MER311995	*Cupriavidus basilensis*	12, 358
MER238517	*Enterobacter cloacae*	18, 724
MER003307	*Kluyvera citrophila*	18, 19
MER169879	*Luminiphilus syltensis NOR5-1B*	7, 19, 211
MER311986	*Shigella sp. D9*	11, 18, 19

[a] The Cys positions that are in bold form disulfide bond, while the underlined ones are removed with signal peptide during post-translational processing. Residue numbers are according to their positions in the precursor sequence (with signal and spacer peptide)

[b] PGA sequences from *Acinetobacter* genus are more than 94 % identical to each other; therefore, only one of them (*Acinetobacter oleivorans*; *Ao*PGA) was used in the analysis

Table 2 List of 11 experimentally characterized mutations known to enhance thermostability in *Ec*PGA

Stabilizing mutations (Polizzi et al. 2006)	Corresponding residues present in other PGAs				
	*Ax*PGA	*Af*PGA	*Sw*PGA	*Pd*PGA	*Ao*PGA
β84A → P	P	P	P	P	P
β400 V → L	L	L	L	L	L
α25 W → Y	Y	Y	Y	H	Y
β359 V → L	L	L	V	L	L
α150T → N	N	N	N	S	S
β311T → P/β312Q → A	P/A	P/A	P/A	D/P	N/A
β100L → E	E	E	D	T	E
α80A → R	R	R	A	S	Q
β305A → S	R	K	S	D	Q
β348 N → D	D	K	A	Q	R
Total number of conserved stabilizing mutations	9	8	5	3	5

The amino acid residues present at these 11 sites among other PGAs are listed

In the absence of three-dimensional structures for the enzymes *Ax*PGA and ptPGAs, 3D homology models were built using *Ec*PGA structure as template. Sequence comparison of the above-mentioned PGAs revealed that the α-chain of target sequences were more than 41 % identical (62 % similar) to *Ec*PGA, while the β-chains were more than 34 % identical (51 % similar). The key residues involved in catalysis such as βSer1, βGln23, βAla69, βAsn241 and βArg263 of *Ec*PGA are well conserved among all [7, 18]. Similarly the residues involved in the binding of penicillin G, namely, αPhe146, βIle177, βPro49 and βTrp154, were also found to be conserved in all enzymes [18, 19]. Other residues in the penicillin side chain-binding pocket such as βVal56, βThr32 and βPhe24 were observed to be partially conserved. Thus, the tertiary structure models built for *Ax*PGA and ptPGAs were of high quality as validated by model validation programs (Table S2).

Site-specific sequence-based consensus approach for thermostability analysis

Rational protein engineering or directed evolution methods have shown that careful selection of sites for mutagenesis could result in an enhancement of thermostability among mesostable enzymes [3]. In case of larger proteins such as PGA, identification of potential sites by these methods remains a challenging problem due to wide range of possible substitutions. The consensus method has been proven to be successful in such cases where the potential mutation sites are identified solely based on sequence comparison [16, 27]. Polizzi et al. have combined the sequence-based consensus approach with the structural information by examining 21 amino acid positions in *Ec*PGA in an effort to increase its thermostability by site-directed mutagenesis experiments [22]. Approximately, 50 % (11) of the selected mutations were found to show positive effect toward thermostability. These 11 mutation positions on analysis in *Ax*PGA, *Af*PGA and ptPGAs revealed the natural presence of some of these stabilizing mutations (Table 2). Of these 11, a total of 9 stabilizing mutations were observed to be conserved in *Ax*PGA (most thermostable), while *Af*PGA, *Sw*PGA, *Pd*PGA and *Ao*PGA showed 8, 5, 3 and 5 conserved mutations, respectively. Natural occurrence of these stabilizing amino acid residues in ptPGAs hinted at their higher thermostability as compared to *Ec*PGA, but lower compared to that of *Ax*PGA and *Af*PGA.

In an effort to identify more such stabilizing sites, a modified consensus approach was designed which involved position-wise comparison of mesostable PGAs with their thermostable homologs. Mesostable *Ec*PGA and *Kc*PGA (PGA from *Kluyvera cryocrescens*; Genbank accession AID61747) sequences were aligned with thermostable *Af*PGA and *Ax*PGA using multiple sequence alignment. It was assumed that a residue position conserved among thermostable enzymes (*Af*PGA and *Ax*PGA) while variable among mesostable homologs (*Ec*PGA and *Kc*PGA) could be considered as a potential site selected by nature for protein thermostabilization. While Ala was found at the α80 position and Leu at β100 in both *Ec*PGA and *Kc*PGA, in case of the thermostable *Af*PGA and *Ax*PGA, at this position charged residues Arg and Glu, respectively, were found to be conserved since these residues are better suited for higher temperatures (Table 3). The possible reason could be the observed ion-pair network involving these residues in the *Af*PGA structure. Similarly, the β311 and β312 positions of both thermostable *Af*PGA and *Ax*PGA contains the conserved Pro and Ala residues, respectively, whereas the mesostable *Kc*PGA possesses Ala and Glu residues and *Ec*PGA has Thr and Gln at the corresponding positions

Table 3 List of 24 sites identified by sequence-based consensus approach

EcPGA Resno[a]	Mesostable		Thermostable		Putative thermostable[b]		
	*Ec*PGA	*Kc*PGA	*Af*PGA	*Ax*PGA	*Pd*PGA	*Sw*PGA	*Ao*PGA
α79	R	R	Q	Q	R	A	K
α80	A	A	R	R	***S***	A	***Q***
α108	N	N	R	R	L	L	L
α121	T	T	D	D	D	E	D
β98	K	K	T	T	E	Q	P
β100	L	L	E	E	T	D	E
β112	Q	Q	A	A	P	S	A
β129	T	T	F	F	I	F	W
β133	T	T	Q	Q	T	Q	N
β218	K	K	L	L	***Q***	L	***Q***
β234	S	S	Q	Q	Q	G	K
β280	D	D	Q	Q	***D***	D	Q
β308	S	A	Q	Q	E	G	Q
β311	T	A	P	P	D	P	N
β312	Q	E	A	A	P	A	A
β313	S	N	D	D	Q	H	S
β337	K	K	G	G	R	–	K
β404	K	K	A	A	R	Q	Q
β432	E	Q	A	A	A	A	T
β436	K	K	Q	Q	***D***	***A***	K
β443	S	T	A	A	A	D	N
β457	N	N	K	K	S	L	K
β519	K	K	P	P	P	P	V
β544	E	D	R	R	E	E	D
Total number of sites where different residues are present than observed in thermostable *Af* PGA/*AxPGA*					19	17	17
Total number of sites predicted to be thermostabilizing					4	1	2

At these specific sites, mesostable *Ec*PGA/*Kc*PGA have variable residues compared to the conserved residues of thermostable *Af*PGA/*Ax*PGA. The residues at these 24 sites among the putative thermostable PGAs are listed

[a] The residue positions in bold correspond to the sites for which experimental evidence is available as stabilization sites

[b] The bold and italic residues are the sites having higher potential of thermostabilization as predicted by bidirectional stability prediction analysis using FoldX and I-Mutant. The hyphen in *Sw*PGA column represents the gap in the sequence alignment

(Table 3). This was also substantiated by the mutational experiments carried out by Polizzi et al., wherein *Ec*PGA showed that the α80A → R substitution resulted in a 2.7-fold increase of half-life at 50 °C, and the β100L → E mutation, where uncharged non-polar residue Leu was substituted with negatively charged residue Glu, showed a 1.2-fold increase in half-life at 50 °C [22]. The *Ec*PGA double mutant (β311T → P/β312Q → A) showed enhanced enzyme half-life by nearly twofold [22].

A total of 24 stabilization sites were identified where mesostable PGAs have different residues compared to the conserved residues in thermostable PGAs (Table 3). Sites were selected such that the residues are surface exposed and lie outside a 10 Å radius cutoff from the active site, so that substitution of these residues would have minimum effect on the catalytic efficiency of the enzyme. Based on residue preference at these 24 sites, ptPGAs were assigned to either mesostable or thermostable groups. Analysis revealed that 19 sites in *Pd*PGA and 17 in each case of *Sw*PGA and *Ao*PGA were found to possess different residues than observed in the case of thermostable *Ax*PGA/*Af*PGA (Table 3). At the α80 position, among ptPGAs, *Sw*PGA was found to contain a non-polar Ala, while *Pd*PGA and *Ao*PGA contain uncharged polar Ser and Gln, respectively, in lieu of the positively charged Arg residues found in *Af*PGA/*Ax*PGA. Similarly, at the β100 position,

ptPGAs possesses different residues than those observed in *Ax*PGA (Table 3). At positions β311/β312, only *Sw*PGA was found to maintain residues similar to *Af*PGA/*Ax*PGA (P/A), while others have different residues (D/P in *Pd*PGA and N/A in *Ao*PGA). Thus, more than 70 % of the concerned sites in ptPGAs behave more similar to mesostable enzymes.

The 24 identified sites were then analyzed for their effect on stability from a structural point of view. For the *Af*PGA–*Sw*PGA pair of enzymes, stabilizing residues present in *Af*PGA at the 24 sites were introduced in *Sw*PGA and similarly the reverse process was carried out in *Af*PGA by introducing the residues of *Sw*PGA. Sites where mutations led to an enhancement in thermostability of *Sw*PGA and the same reverse mutations in *Af*PGA resulting in destabilization could have high probability of enhancing the thermostability. The mutational effect on protein stability was estimated using FoldX [11] and I-Mutant [5] which were based on empirical potential energy function and support vector machine, respectively. These tools not only have scores for favorable interactions, but also have penalties for unfavorable clashes. A mutation is considered to increase the stability if both FoldX and I-Mutant predicts so, otherwise not. Similar studies were also carried out for *Af*PGA–*Pd*PGA and *Af*PGA–*Ao*PGA pairs of enzymes. Of the 24 sites, a total of one (β429A → Q), four (α78S → R, β217Q → L, β278D → Q and β422D → Q) and two (α78Q → R and β218Q → L) sites, respectively, in case of *Sw*PGA, *Pd*PGA and *Ao*PGA were identified which showed higher potential for increasing thermostability (Table 3 and Table S6). Of these, the positions corresponding to α80, β218 and β436 in *Ec*PGA were found to be the most potent sites which upon mutation could increase thermostability in two ptPGAs. Of these three sites, α80 position has also been shown to be a prime site for thermostabilization by Polizzi et al. Overall, the stability analysis helped to narrow down the choices for potential substitution sites, aimed at enhancement of thermostability, to three possible positions for the selected ptPGAs.

Structure-based approach of thermostability analysis

In this approach, the modeled structures of *Sw*PGA, *Pd*PGA and *Ao*PGA were compared with the crystal structure of *Ec*PGA (less thermostable), *Af*PGA (moderately thermostable) and the modeled structure of *Ax*PGA (most thermostable) in terms of several factors known to influence the thermodynamic stability of proteins, such as disulfide bridges, ion pairs, hydrogen bonds, aromatic–aromatic, aromatic–sulfur and cation–pi interactions, entropic stabilization due to proline residues and reduction of deamidation damages. The effect of temperature on structural stability of enzymes was monitored for all the six PGAs by subjecting them to molecular dynamics simulations at 330, 400 and 500 K temperatures. The average RMSD values of Cα atoms of the enzymes during all the three temperature-dependent simulations were monitored (Table S3). At every 10 ps intervals, enzyme conformations were extracted and structural parameters were monitored. Below is a description of the comparative analysis of the six PGA enzymes considered.

Stabilization by disulfide bridges

Disulfide bridges are known to stabilize protein structure by reducing the conformational entropy of its unfolded state [17]. *Af*PGA contains one disulfide bond between residues βCys492 and βCys525 in its structure [32]. Both *Ec*PGA and *Ax*PGA lack this disulfide bridge; in the place of residues βCys492 and βCys525 of *Af*PGA, the residues βTrp500 and βArg533 were observed in *Ec*PGA, while βVal500 and βSer533 were present in *Ax*PGA. The modeled structures of ptPGAs showed potential to form disulfide bond between a pair of cysteine residues of β-chain (*Sw*PGA: βCys489–βCys522, *Pd*PGA: βCys483–βCys516 and *Ao*PGA: βCys492–βCys525; Fig. S1). Previous studies have shown that the contribution of entropy to the free energy of stabilization increases proportionately with the number of residues separating the two cysteine residues of the disulfide bond [17, 21]. The number of residues separating the two Cys residues in all ptPGAs is the same as that in *Af*PGA (32 residues), suggesting a similar entropic effect. Thus, the presence of disulfide bridges in ptPGAs is expected to contribute positively to their thermostability, as seen in the case of *Af*PGA.

Preference of Arg over Lys: a mechanism for thermostabilization

*Ax*PGA has been observed to be more thermostable than *Af*PGA, though it lacks a disulfide bond suggesting the higher influence of other factors toward its thermostability. Preference of Arg over Lys residue and higher number of ion pairs are proposed to be the contributing thermostability factors in *Ax*PGA [4]. Arg residues are favored at higher temperatures more than Lys, due to resonance stabilization of their side chain along with higher surface area for charged interactions. Thermostable proteins are known to maintain the value of their Arg/Lys ratio greater than one [34]. Although the total content of Arg and Lys residues together was found to be almost equal across all six PGAs, Arg/Lys ratios of individual enzymes were found to differ considerably, correlating well with their thermostable nature. While in less thermostable *Ec*PGA, the ratio was observed to be 0.78 (Lys preferred over Arg), the ratio is changed to 1.6 (Arg preferred over Lys) in moderately

thermostable *Af*PGA, and in case of most thermostable *Ax*PGA, the ratio was found to be the highest (2.29). Among the selected ptPGA enzymes, *Sw*PGA and *Pd*PGA have Arg/Lys ratios 1.26 and 3.23, respectively, suggesting an increase in thermal stability, while in *Ao*PGA this ratio was found to be the least (0.38). Thus, the overall analysis suggested that Arg/Lys ratio contributed maximally toward thermostability in case of the enzyme *Pd*PGA.

Presence of higher ion pair networks: a stabilizing mechanism

Thermostable enzymes *Af*PGA and *Ax*PGA were found to maintain comparatively higher percentage of ion pairs and ion pair networks in their tertiary structures than mesostable *Ec*PGA. During the course of simulations at 330 K, 400 K and 500 K, at every 10 ps interval, enzyme conformations were extracted and the total number of ion pairs (IPs) and percentage of ion pairs that are involved in network formation (IPnets) were estimated. The average values of the number of IPs and percentage of IP networks of *Ec*PGA were considered as reference, and relative percentages were calculated for other PGAs with respect to *Ec*PGA. A relative percentage value >0 shows higher percentage of IPs and IPnets as compared to *Ec*PGA. Three ranges of electrostatics interactions were estimated, that is: short range, medium range and long range, based on the distance cutoffs between the acidic and basic residue side chains as 4, 6 and 8 Å, respectively.

Figure 1 shows the relative percentage values of IPs and IPnets of *Af*PGA, *Ax*PGA, *Sw*PGA, *Pd*PGA and *Ao*PGA compared to *Ec*PGA monitored at 330, 400 and 500 K of simulations. At 330 K, both thermostable *Af*PGA and *Ax*PGA were found to contain higher percentage (17.48 and 13.03 % higher, respectively) of *short-range* IPs compared to *Ec*PGA. Similar observations were recorded for *Sw*PGA and *Pd*PGA which maintained 8.27 and 30.61 % more *short-range* IPs as compared to *Ec*PGA. *Ao*PGA was found to contain the least percentage of *short-range* IPs (22.53 % less as compared to *Ec*PGA) among all PGAs. At 400 and 500 K, *Pd*PGA was observed to dominate all PGAs in terms of *short-range* IPs. Similarly, *Pd*PGA was also observed to dominantly maintain *medium-* and *long-range* IPs at 330, 400 and 500 K.

Since interaction networks are comparatively stable and energetically more favorable than isolated interactions, PGA enzymes were also compared in terms of percentage of occurrence of ion pair networks (IPnets). At 330 K, compared to *Ec*PGA, thermostable *Af*PGA and *Ax*PGA were found to maintain slightly higher percentage (4.7 and 1.6 %, respectively) of *short-range* IPnets. Among the ptPGAs, *Sw*PGA and *Ao*PGA maintained a lower percentage (9.42 and 43.35 % lower, respectively) of IPnets. Notably, *Pd*PGA was found to contain the highest ion pair network percentage (22.42 % higher than *Ec*PGA) among all PGA enzymes. Interestingly at 400 K and 500 K, *Pd*PGA was again found to dominate all PGAs with respect to IPnet values. When medium- and long-range ionic interactions were monitored, *Pd*PGA consistently dominated all PGAs in terms of IPnet values at 300, 400 and 500 K (Fig. 1). The observation of larger number of IPs and IPnets (Fig. S2) along with the possible existence of a disulfide bridge suggests a possibility of higher thermostability in case of the *Pd*PGA enzyme.

Loop stabilization by proline residues

Proline residues being conformationally rigid are known to provide stability by entropic effect [35]. Many thermophilic and hyperthermophilic proteins have shown to adopt this stabilization mechanism by maintaining higher proline content as compared to their mesophilic homologs. A positive correlation between proline content and thermostability was observed among the PGA enzymes under study. The known thermostable *Af*PGA (6.02 %) and *Ax*PGA (6.51 %) enzymes were observed to have higher proline content compared to mesostable *Ec*PGA (5.22 %). All the selected ptPGA enzymes were found to have slightly higher proline content compared to *Ec*PGA (*Sw*PGA: 5.95 %, *Pd*PGA: 5.89 % and *Ao*PGA: 5.72 %).

Introduction of proline residues at the second position of β-turns (Bt2P) has been considered as one of the protein engineering strategies for enhancement of protein thermostability [28]. The average number of Bt2P residues among the six PGA enzymes was monitored during all temperature scales of simulation (Fig. S3) and relative percentage values were compared. At 330 K, thermostable *Af*PGA and *Ax*PGA were observed to maintain higher average percentage of Bt2P residues (36.99 and 27.05 % higher, respectively) than *Ec*PGA. Among the ptPGAs, the behavior of *Pd*PGA was found to be similar to that of thermostable PGAs having higher average Bt2P percentages (28.87 % higher) than *Ec*PGA. *Pd*PGA was also found to maintain stable β-turns even at higher temperatures of 400 and 500 K (Fig. S3) suggesting the loop stabilization by proline residues as another possible contributing factor toward *Pd*PGA thermostability.

Lesser number of thermolabile residues

Asparagine (Asn) and glutamine (Gln) are considered as thermolabile amino acids since they undergo deamidation at higher temperatures. Spontaneous deamidation of Asn leads to formation of aspartic or iso-aspartic residues resulting in important functional and biological damage in peptides and protein structures. In acid–base

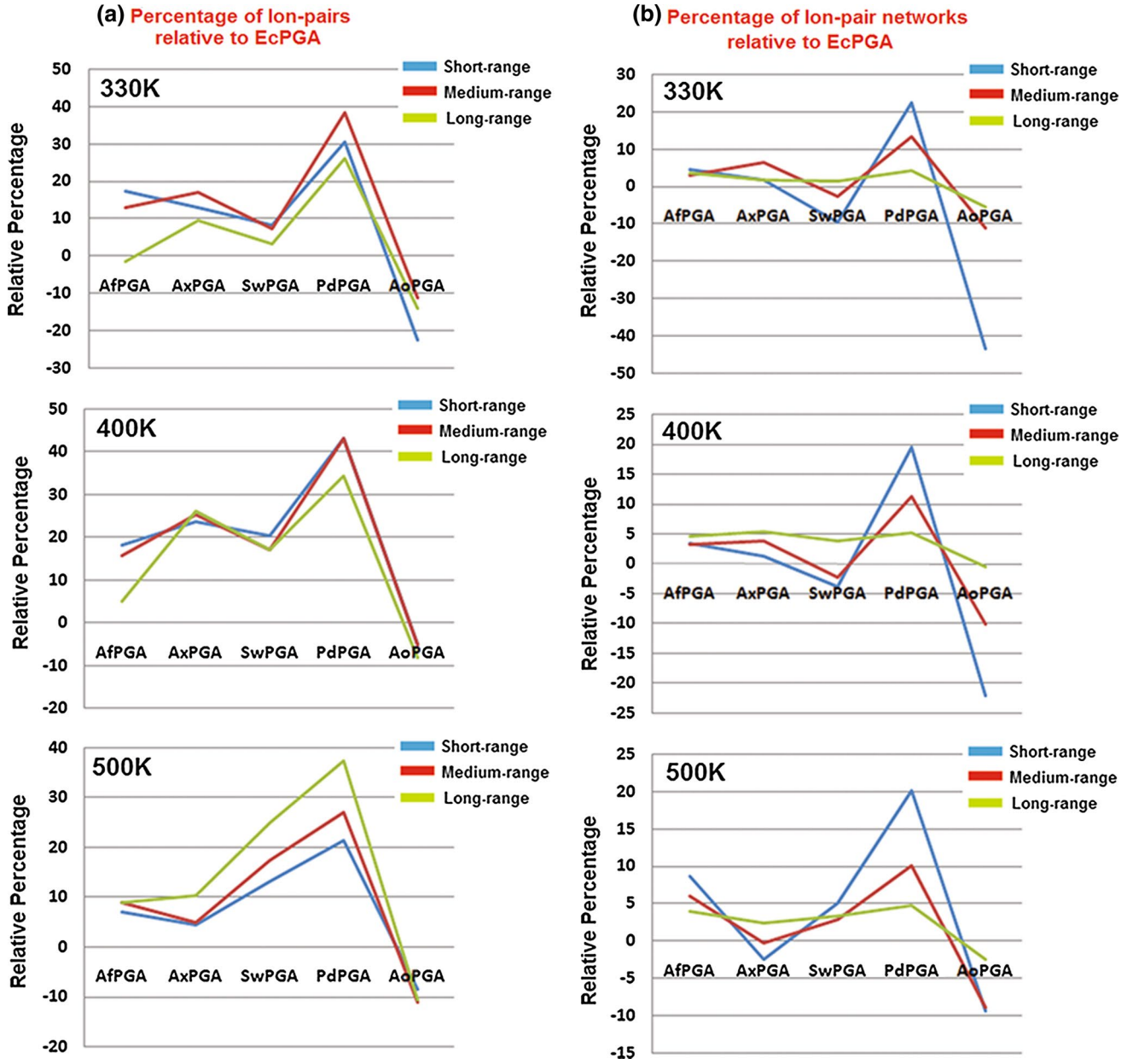

Fig. 1 *Line plots* depicting the average percentages of ion pairs (*panel a*) and ion pair networks (*panel b*) of *Af*PGA, *Ax*PGA, *Sw*PGA, *Pd*PGA and *Ao*PGA relative to *Ec*PGA (relative percentage values in *y* axis), during 330 K (*top panel*), 400 K (*middle panel*) and 500 K (*bottom panel*) of molecular dynamics simulations. The *blue, red and green lines* in the plots correspond to short-range, medium-range and long-range electrostatic interactions estimated using 4, 6 and 8 Å, respectively, the distance cutoffs between the acidic and basic residues

mechanism of deamidation, residues like Ser and Thr are thought to act as acid groups, which protonate the leaving side chain amide group of Asn. Therefore, thermophilic proteins tend to have less of uncharged polar amino acid content (Asn+Gln+Ser+Thr (NQST) content) to reduce heat-induced damage [25]. Both *Af*PGA and *Ax*PGA have lower percentage of NQST residues (23.6 and 18.22) than *Ec*PGA (24.5). Among the ptPGAs, *Sw*PGA and *Pd*PGA were found to behave similar to thermostable PGAs by their low NQST content as compared to *Ec*PGA (18.8 and 18.24 % respectively). In contrast, *Ao*PGA has the highest NQST content among all PGA enzymes (25.3 %).

Asn–Gly bonds are considered thermolabile in nature, since they undergo deamidation at higher temperature through β-aspartyl shift mechanism [25]. Based on the predicted coefficient of deamidation values, *Ec*PGA was found to contain five thermolabile Asn–Gly bonds (Table S4). Of these, four bonds were observed to be substituted either at

Asn or Gly position among the thermostable *Af*PGA and *Ax*PGA. Among ptPGAs, a total of two, three and three such bonds are substituted among *Sw*PGA, *Pd*PGA and *Ao*PGA, respectively. The reduction in thermolabile amino acid content and thermolabile bond substitutions could be considered as another stabilization strategy in the PGA enzyme family.

Analysis of other non-bonded interactions

In addition to disulfide bond and ion pairs, PGA enzymes were also compared in terms of other non-bonded interactions such as hydrogen bonds, aromatic–aromatic, aromatic–sulfur and cation–pi interactions. The interactions were monitored at all temperatures of molecular dynamics simulations (Fig. S4). Since no direct correlations were observed for these interactions between mesostable and thermostable PGA enzymes, it was difficult to assess the degree of contribution of these factors toward PGA thermostability.

In summary, the screened ptPGA enzymes were compared with the experimentally characterized *Ec*PGA, *Af*PGA, and *Ax*PGA (in increasing order of thermostability) first in terms of their sequence followed by their structural comparison. In the sequence-based consensus approach, PGAs were compared in terms of their residue preference at 24 thermostabilization sites, while in the structure-based approach enzymes were compared in terms of various structural features known to contribute to protein thermostability. The *sequence-based* analysis revealed that ptPGA enzymes could have higher thermostability as compared to *Ec*PGA while also highlighting few additional potential sites which on mutation could improve their thermostability. The structural analysis of the enzymes emphasized that, of all the ptPGAs, *Pd*PGA could be equally thermostable as *Af*PGA and *Ax*PGA owing to the presence of a disulfide bond, high Arg/Lys ratio, higher number of stable ion pair networks, greater proline content, higher percentages of proline residues in beta-turns and lower content of thermolabile residues and bonds. Finally, the experimental verification of the structure–function relationship and stability of the *Pd*PGA enzyme was undertaken.

Experimental characterization of a selected enzyme PdPGA

Purification of PdPGA

The *Pd*PGA gene (Genbank Accession: ABL72874.1) is encoded in the plasmid of *Paracoccus denitrificans PD1222*, a Gram-negative soil bacteria capable of reducing nitrate by denitrification. The precursor protein consists of 790 amino acid residues, which upon maturation by removal of signal and spacer peptide forms the active enzyme. The *Pd*PGA expressed was purified from *E. coli* BL21-Gold (DE3) cells by a three-step purification protocol of Ni–NTA affinity chromatography followed by desalting and size exclusion chromatography. The SDS-PAGE and Western blot confirmed the expression of *Pd*PGA. The purified enzyme showed two bands on SDS-PAGE, corresponding to α and β chains, along with the detection of β-chain by Western blot with monoclonal anti-His antibodies (Fig. S5). The purified *Pd*PGA fraction showed activity toward the substrate penicillin G.

Alkaline stability of PdPGA

The optimum pH of *Pd*PGA enzyme activity was observed to be pH 10 (Fig. 2a). The pH stability profile of *Pd*PGA showed that the enzyme was stable over a wide range of pH 5–11 (Fig. 2c). The enzyme was found to be most stable in alkaline pH 10. The enzyme retained almost 60 % of its activity at pH 10 even after 3 h of incubation. The enzyme was also found to be stable at pH 11, retaining 50 % of its activity after 3 h of incubation. However, it was found to be unstable at pH 12. Toward the acidic pH scale, *Pd*PGA was found to be stable till pH 5 with 50 % of enzyme activity being retained after 3 h of incubation. At pH 4, 70 % reduction of enzyme activity was observed after 3 h of incubation.

Temperature stability profile of PdPGA

The optimum temperature for substrate hydrolysis by *Pd*PGA was observed to be 45 °C (Fig. 2b). The temperature stability profile of *Pd*PGA was monitored up to 50 °C (Fig. 2d). At 50 °C upon 10 min incubation, complete loss of enzyme activity was observed. However, at 45 °C, the enzyme was found to be stable for 20 min. Overall, the enzyme was found to be mesostable and had lesser thermostability compared to both *Af*PGA and *Ax*PGA which are known to be stable beyond 50 °C. Loss of stability beyond 45 °C could be either due to global structural changes of the enzyme or due to local structural changes near the active site. To understand this, thermal unfolding of *Pd*PGA studies using both steady-state fluorescence and CD spectroscopy were carried out.

Probing structural changes using fluorescence

Compared to tyrosine and phenylalanine residues, tryptophan (Trp) residues exhibit a much stronger fluorescence on excitation. Since Trp fluorescence is dependent on its microenvironment, even minor changes in this microenvironment can lead to changes in the Trp fluorescence spectra. *Pd*PGA has 20 Trp residues of

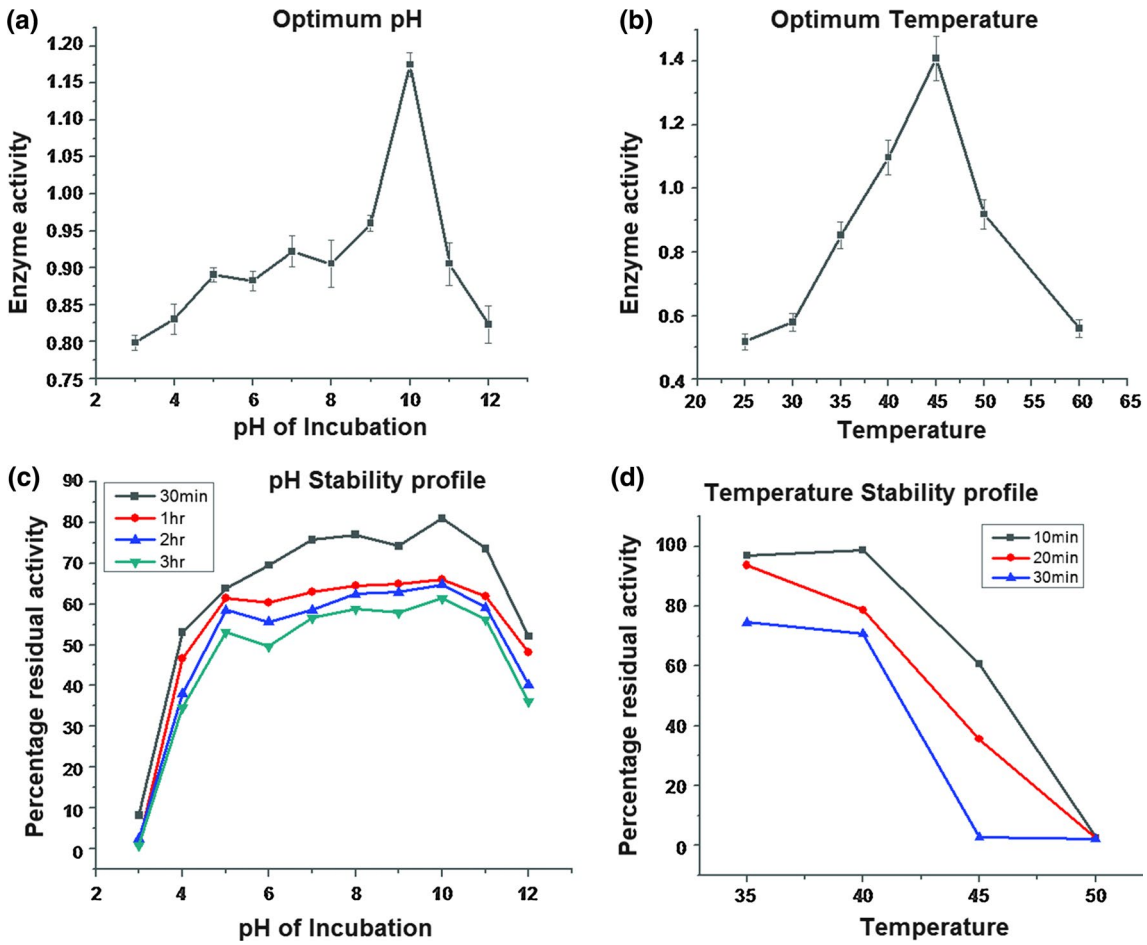

Fig. 2 a, b The optimum pH and temperature of *Pd*PGA activity. Optimum pH of *Pd*PGA was determined using buffers in the pH range of 3–12 while incubating at 40 °C. Buffers used were 25 mM glycine–HCl for pH 1–3, acetate for pH 4–5, phosphate for pH 6–7, Tris–HCl for pH 8–9 and glycine–NaOH buffer for pH 10–12. Optimum temperature for *Pd*PGA activity was determined using 0.05 M buffer of pH 10 over a temperature range of 25–60 °C. **c, d** The stability profile of *Pd*PGA at various pH values and temperatures, respectively, estimated at regular time intervals

which 5 residues were found to be within 10 Å of its N-terminal catalytic residue βSer1. Thus, heat-induced conformational changes in *Pd*PGA could be easily monitored by recording Trp fluorescence spectra at various temperatures. The λ_{max} of the intrinsic fluorescence spectrum of *Pd*PGA was near 338 nm between 25 and 40 °C (Fig. 3a). However, a redshift to 345 nm at 45 °C indicated a possible alteration of Trp microenvironment to a partially hydrophilic nature due to conformational changes in the protein upon heat treatment (Fig. 3b). These observations could be correlated with the observed 40 % decrease in enzyme activity at 45 °C as well as the complete loss of activity at 50 °C. Beyond 50 °C further redshift was observed till 350 nm at 70 °C. Rayleigh light scattering experiment carried out using the above experimental setup was used to study protein aggregation due to heat treatment (Fig. 3c). A gradual increase in scattering intensity after 45 °C indicated protein aggregation due to heat-induced denaturation resulting in loss of enzyme activity.

Circular dichroism study

The change in secondary structure of *Pd*PGA in the course of thermal denaturation was observed in far-UV CD analysis. Figure 3d shows the CD spectra of *Pd*PGA from 25 to 70 °C. Gradual reduction in ellipticity above 45 °C correlated with the loss of activity as well as altered Trp environment. The CD-pro analysis showed a slight decrease in α-helical content, but a significant decrease in β-sheet content of enzyme beyond 45 and 50 °C compared to 40 °C (Table S5). The decrease in ordered secondary structure content led to an increase in turns and unordered regions. This observation can be linked with the possible unfolding of the αββα core Ntn-hydrolase fold which contains the catalytic site, thus leading to loss of enzyme activity.

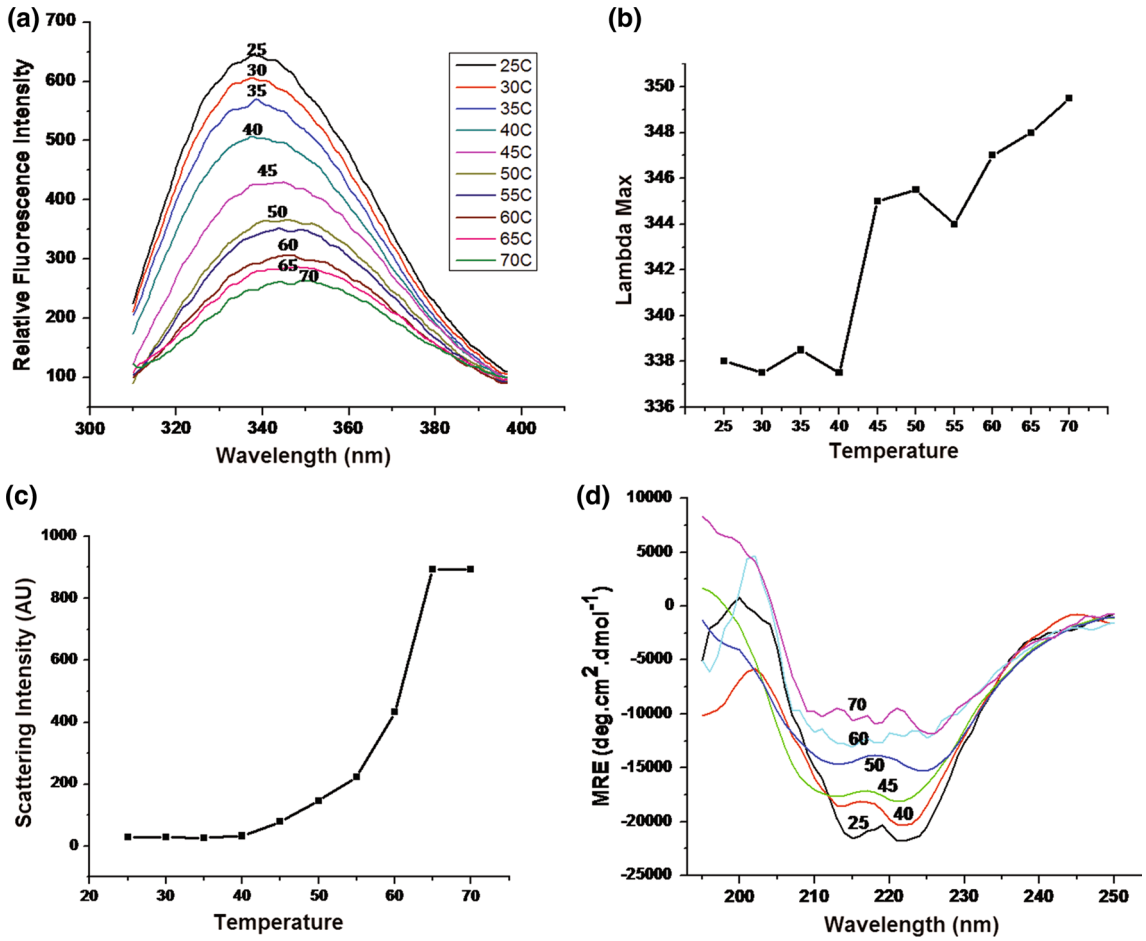

Fig. 3 **a** Trp fluorescence spectra of *Pd*PGA subjected to various temperatures from 25 to 70 °C. **b** *Line plot* showing the increase of λ_{max} of fluorescence spectra (*redshift*) of *Pd*PGA at higher temperatures. **c** *Line plot* depicting the scattering intensities of *Pd*PGA at various temperatures. **d** Far-UV CD spectra of *Pd*PGA subjected to various heat treatments

Although *Pd*PGA contains a disulfide bond similar to the thermostable *Af*PGA and has higher preference for Arg over Lys, high density of ion pair networks, lower content of thermolabile residues and higher percentage of Bt2P residues like in *Ax*PGA, its thermal behavior is closer to mesostable *Ec*PGA. Thus, we may infer that either the observed factors do not contribute significantly to *Pd*PGA thermostability or that their cumulative effect is not sufficient. Admittedly, identification of the true basis of thermostability is highly complex, since protein thermostability is due to cooperative optimization of several factors rather than by a single prevailing factor. Furthermore, the extent of contribution of factors differs between different taxons, different organisms within a given taxon and even between different proteins from the same organism, thus making generalization a complicated task [30]. Among the factors analyzed here, the presence of disulfide bond was indeed found to be a positive contributor to *Pd*PGA thermostability. When *Pd*PGA was treated with 50 and 100 mM of reducing agent dithiothreitol (DTT), 40 and 80 % decrease in activity was observed compared to the activity of the native *Pd*PGA enzyme at 40 °C. In case of *Af*PGA, the reduction of disulfide bridge with 10 mM DTT has been reported to decrease approximately 50 % of enzyme activity [33]. Since the two Cys residues connect the layers of αββα Ntn hydrolase fold (Fig. S6), the possible role of disulfide bond could be to maintain the structural integrity of the Ntn hydrolase fold. The high-density ion pair networks observed in *Pd*PGA might contribute to the pH stability of *Pd*PGA. The enzyme is stable in a broad range of pH (pH 5–11), particularly in alkaline condition (pH 10). Even its optimum pH is at pH 10. Enzymes stable in a broad range of pH are not only useful during normal handling, but also in industrial applications.

In a recent attempt toward making an alkaline stable PGA enzyme, Suplatov et al. 2014 have carried out βAsp484 → Asn mutation which showed a ninefold increase in the stability of *Ec*PGA at pH 10 [26].

Interestingly, it was observed that at the position corresponding to βAsp484 of *Ec*PGA, it is indeed an Asn residue present in *Pd*PGA, which could act as a major contributing factor toward its alkaline stability and optimum pH condition as 10. In our approach toward prediction of thermostability, consensus site-specific sequence-based approach gave a more realistic estimate of *Pd*PGA thermostability as compared to the known thermostability factors studied structurally. The computational analysis carried out in this study is based on only two experimentally determined three-dimensional structures, *Ec*PGA and *Af*PGA. Determination of more crystal structures, of other family members including *Ax*PGA and *Pd*PGA, could improve the prediction method and correlate the extent of stabilizing factors effecting thermostability with corresponding experimental results.

Conclusion

The current study highlights the complexity of sequence–structure–function relationship in the PGA family of enzymes. Although several mechanisms for protein structure stabilization were considered for the analysis, the lack of any dominant factor or universal strategy makes it difficult to screen novel candidates based on a single strategy. The consensus sequence-based analysis classified the selected ptPGAs to be mesostable. Extensive structural analysis carried out on the ptPGAs and their comparison with the other PGA enzymes singled out *Pd*PGA as the most potent candidate for the exploration of its thermostable nature. Experimental characterization revealed that *Pd*PGA was comparatively less stable than the known thermostable *Af*PGA and *Ax*PGA enzymes. However, the enzyme was unique in terms of its maximal activity as well as stability under alkaline condition. Although the experimental characterization of *Pd*PGA thermostability limits its application potential as an improved industrial biocatalyst, the sustained effort toward increasing its thermal stability through substitution mutation of the identified sites, could transform this uniquely alkalistable PGA into an extremely viable industrial biocatalyst.

Acknowledgments PP, DC and RM thank the Council of Scientific and Industrial Research, New Delhi, for Senior Research Fellowship. SK is DST Ramanujan Fellow. This research was carried out under the CSIR-NCL's Centre of Excellence in Scientific Computation (CoESC).

References

1. Arroyo M, de la Mata I, Acebal C, Castillon MP (2003) Biotechnological applications of penicillin acylases: state-of-the-art. Appl Microbiol Biotechnol 60:507–514. doi:10.1007/s00253-002-1113-6
2. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242
3. Blundell TL, Elliott G, Gardner SP, Hubbard T, Islam S, Johnson M, Mantafounis D, Murray-Rust P, Overington J, Pitts JE, Sali A, Sibanda BL, Singh J, Sternberg MJE, Sutcliffe MJ, Thornton JM, Travers P (1989) Protein Engineering and Design. Philosophical Trans Roy Soc Lond B Biol Sci 324:447–460
4. Cai G, Zhu S, Yang S, Zhao G, Jiang W (2004) Cloning, overexpression, and characterization of a novel thermostable penicillin G acylase from Achromobacter xylosoxidans: probing the molecular basis for its high thermostability. Appl Environ Microbiol 70:2764–2770
5. Capriotti E, Fariselli P, Casadio R (2005) I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res 33:W306–W310
6. Chilov GG, Moody HM, Boesten WHJ, Å vedas VK (2003) Resolution of (RS)-phenylglycinonitrile by penicillin acylase-catalyzed acylation in aqueous medium. Tetrahedron: Asymmetry 14:2613-2617. http://dx.doi.org/10.1016/S0957-4166(03)00523-8
7. Duggleby HJ, Tolley SP, Hill CP, Dodson EJ, Dodson G, Moody PC (1995) Penicillin acylase has a single-amino-acid catalytic centre. Nature 373:264–268. doi:10.1038/373264a0
8. Erarslan A, Kocer H (1992) Thermal inactivation kinetics of penicillin G acylase obtained from a mutant derivative of Escherichia coli ATCC 11105. J Chem Technol Biotechnol 55:79–84
9. Fuganti C, Grasselli P, Casati P (1986) Immobilized penicillin acylase: application to the synthesis of the dipeptide aspartame. Tetrahedron Lett 27:3191–3194
10. Fuganti C, Grasselli P, Seneci PF, Servi S, Casati P (1986) Immobilized benzylpenicillin acylase: application to the synthesis of optically active forms of carnitin and propranalol. Tetrahedron Lett 27:2061–2062
11. Guerois R, Nielsen JE, Serrano L (2002) Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J Mol Biol 320:369–387
12. Guranda DT, Khimiuk AI, van Langen LM, van Rantwijk F, Sheldon RA, Å vedas VK (2004) An 'easy-on, easy-off' protecting group for the enzymatic resolution of (±)-1-phenylethylamine in an aqueous medium. Tetrahedron: Asymmetry 15:2901-2906. http://dx.doi.org/10.1016/j.tetasy.2004.06.051
13. Ismail H, Lau RM, van Langen LM, van Rantwijk F, Svedas VK, Sheldon RA (2008) A green, fully enzymatic procedure for amine resolution, using a lipase and a penicillin G acylase. Green Chem 10:415–418. doi:10.1039/b714088f
14. Kasche V, Lummer K, Nurk A, Piotraschke E, Rieks A, Stoeva S, Voelter W (1999) Intramolecular autoproteolysis initiates the maturation of penicillin amidase from Escherichia coli. Biochim. et Biophys. Acta 1433:76–86
15. Klock HE, Lesley SA (2009) The Polymerase Incomplete Primer Extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. Methods Mol Biol 498:91–103. doi:10.1007/978-1-59745-196-3_6
16. Lehmann M, Loch C, Middendorf A, Studer D, Lassen SF, Pasamontes L, van Loon AP, Wyss M (2002) The consensus concept for thermostability engineering of proteins: further proof of concept. Protein Eng 15:403–411
17. Matsumura M, Signor G, Matthews BW (1989) Substantial increase of protein stability by multiple disulphide bonds. Nature 342:291–293. doi:10.1038/342291a0
18. McVey CE, Walsh MA, Dodson GG, Wilson KS, Brannigan JA (2001) Crystal structures of penicillin acylase enzyme-substrate complexes: structural insights into the catalytic mechanism. J Mol Biol 313:139–150. doi:10.1006/jmbi.2001.5043
19. Oh B, Kim K, Park J, Yoon J, Han D, Kim Y (2004) Modifying the substrate specificity of penicillin G acylase to cephalosporin

acylase by mutating active-site residues. Biochem Biophys Res Commun 319:486–492. doi:10.1016/j.bbrc.2004.05.017
20. Oinonen C, Rouvinen J (2000) Structural comparison of Ntn-hydrolases. Protein Sci 9:2329–2337. doi:10.1110/ps.9.12.2329
21. Pace CN, Grimsley GR, Thomson JA, Barnett BJ (1988) Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. J Biol Chem 263:11820–11825
22. Polizzi KM, Chaparro-Riggers JF, Vazquez-Figueroa E, Bommarius AS (2006) Structure-guided consensus approach to create a more thermostable penicillin G acylase. Biotechnol J 1:531–536. doi:10.1002/biot.200600029
23. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B, Lindahl E (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 29:845–854. doi:10.1093/bioinformatics/btt055
24. Rawlings ND, Barrett AJ, Bateman A (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res 40:D343–D350. doi:10.1093/nar/gkr987
25. Robinson NE (2002) Protein deamidation. Proc Natl Acad Sci USA 99:5283–5288. doi:10.1073/pnas.082102799
26. Suplatov D, Panin N, Kirilin E, Shcherbakova T, Kudryavtsev P, Svedas V (2014) Computational design of a pH stable enzyme: understanding molecular mechanism of penicillin acylase's adaptation to alkaline conditions. PLoS One 9:e100643. doi:10.1371/journal.pone.0100643
27. Suplatov D, Voevodin V, Svedas V (2015) Robust enzyme design: bioinformatic tools for improved protein stability. Biotechnol J 10:344–355. doi:10.1002/biot.201400150
28. Suzuki Y, Oishi K, Nakano H, Nagayama T (1987) A strong correlation between the increase in number of proline residues and the rise in thermostability of five Bacillus oligo-1,6-glucosidases. Appl Microbiol Biotechnol 26:546–551. doi:10.1007/bf00253030
29. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882
30. Trivedi S, Gehlot HS, Rao SR (2006) Protein thermostability in Archaea and Eubacteria. Genet Mol Res 5:816–827
31. van Langen LM, van Rantwijk F, Å vedas VK, Sheldon RA (2000) Penicillin acylase-catalyzed peptide synthesis: a chemoenzymatic route to stereoisomers of 3,6-diphenylpiperazine-2,5-dione. Tetrahedron: Asymmetry 11:1077-1083. http://dx.doi.org/10.1016/S0957-4166(00)00027-6
32. Varshney NK, Suresh Kumar R, Ignatova Z, Prabhune A, Pundle A, Dodson E, Suresh CG (2012) Crystallization and X-ray structure analysis of a thermostable penicillin G acylase from Alcaligenes faecalis. Acta Crystallographica Section F 68:273–277. doi:10.1107/S1744309111053930
33. Verhaert RM, Riemens AM, van der Laan JM, van Duin J, Quax WJ (1997) Molecular cloning and analysis of the gene encoding the thermostable penicillin G acylase from Alcaligenes faecalis. Appl Environ Microbiol 63:3412–3418
34. Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65:1–43. doi:10.1128/MMBR.65.1.1-43.2001
35. Watanabe K, Masuda T, Ohashi H, Mihara H, Suzuki Y (1994) Multiple proline substitutions cumulatively thermostabilize Bacillus cereus ATCC7064 oligo-1,6-glucosidase. Irrefragable proof supporting the proline rule. Eur J Biochem 226:277–283