BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

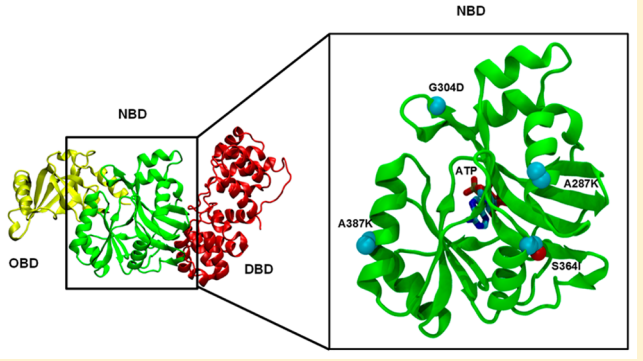
Understanding and Engineering Thermostability in DNA Ligase from *Thermococcus* sp. 1519

Hassan Pezeshgi Modarres,[†,‡] Boris D. Dorokhov,[§] Vladimir O. Popov,[∥,⊥] Nikolai V. Ravin,[§] Konstantin G. Skryabin,*[,§,⊥] and Matteo Dal Peraro*[,†,‡]

[†]Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015, Switzerland
[‡]Swiss Institute of Bioinformatics (SIB), Lausanne 1015, Switzerland
[§]Centre "Bioengineering", Russian Academy of Sciences, Moscow 117312, Russia
[∥]Bach Institute of Biochemistry, Russian Academy of Sciences, Moscow 119071, Russia
[⊥]RSC "Kurchatov Institute", Moscow 123182, Russia

Supporting Information

ABSTRACT: The physical chemical principles underlying enzymatic thermostability are keys to understand the way evolution has shaped proteins to adapt to a broad range of temperatures. Understanding the molecular determinants at the basis of protein thermostability is also an important factor for engineering more thermoresistant enzymes to be used in the industrial setting, such as, for instance, DNA ligases, which are important for DNA replication and repair and have been long used in molecular biology and biotechnology. Here, we first address the origin of thermostability in the thermophilic DNA ligase from archaeon *Thermococcus* sp. 1519 and identify thermosensitive regions using molecular modeling and simulations. In addition, we predict mutations that can enhance thermostability of the enzyme through bioinformatics analyses. We show that thermosensitive regions of this enzyme are stabilized at higher temperatures by optimization of charged groups on the surface, and we predict that thermostability can be further increased by further optimization of the network among these charged groups. Engineering this DNA ligase by introducing selected mutations (i.e., A287K, G304D, S364I, and A387K) eventually produced a significant and additive increase in the half-life of the enzyme when compared to that of the wild type.



Microorganisms are usually divided into four groups based on the temperature conditions under which they have evolved to live: psychrophilics, living in cold environments; mesophilics, living in conditions around human body temperature; thermophilics, living at temperatures between 40 and 80 °C; and hyperthermophilic, which are able to survive at temperatures higher than 80 °C.[1,2] Since 1967, when Thomas D. Brock published the first report about the discovery of microorganisms with optimal growth temperature (OGT) higher than 75 °C,[3] enzymes extracted from these microorganisms that can withstand high temperatures have been increasingly studied to understand the origin of their improved thermostability and used thereafter in biotechnological applications.[4]

Due to the decreased viscosity and increased diffusion coefficient and solubility of substrates at high temperatures, enzymes isolated from hyperthermostable organisms have, in fact, found extensive applications in various industrial domains, like genetic engineering, food processing (e.g., baking, brewing, dairy, starch hydrolysis, etc.), petroleum bioremediation, chemical processes, and the paper and pulp industries.[2−6] In addition, using enzymes that can survive at elevated temperatures usually reduces the risk of presence of common mesophiles that may contaminate the reaction environment.[4−6]

Thus, it is not surprising that numerous studies aimed at understanding the molecular origin of the thermostability of enzymes from mesophiles to hyperthermophiles have attracted the attention of the scientific community.[1,7−19]

It is now clear that nature does use a vast array of complementary strategies to enhance thermostability, such as increasing hydrogen bonding, producing better protein packing, promoting burial of hydrophobic surface area, modifying surface charges, increasing salt bridges and disulfide bonds, developing more favorable helical dipolar interactions, improving secondary structure propensity, increasing rigidity, and entropic stabilization.[1−4,7,8] For each enzyme family, decoding the strategies that hyperthermophiles use to resist high temperatures based on their sequence and structure can be

Received: September 29, 2014
Revised: February 28, 2015
Published: April 22, 2015

DOI: 10.1021/bi501227b
Biochemistry 2015, 54, 3076−3085

important to reverse-engineer these principles and in turn enhance thermostability by modifying an enzyme with the least number of mutations.[20−24]

Computational approaches can help to make such procedures faster and more efficient in order not only to produce higher thermostability but also to improve other properties like increased stability in salt solutions.[25] For instance, rational design strategies have been used to capture thermosensitive regions and propose appropriate substitutions to enhance thermostability.[26−34] In this context, molecular dynamics (MD) simulation is also a valuable means to identify thermosensitive regions in proteins,[26,29,30,32,34] by probing structure at elevated temperatures.[26−34] These insights, along with data extracted by bioinformatics analyses,[35] can be then used to engineer protein structure by optimization of unstable residues.[26−29,32,34]

Following these premises, in this work, we studied the origin of thermostability in a thermophilic DNA ligase, which is also attractive for applications in biotechnology and pharmaceutics,[36,37] and based on this analysis, we proposed mutations to enhance its thermostability. In particular, we focused on the DNA ligase (hereafter called *LigTh1519*) from the thermophilic archaeon *Thermococcus* sp. 1519 isolated from a hydrothermal vent that grows optimally at 85 °C.[38] *LigTh1519* has been recently characterized biochemically, and its structure has been solved by X-ray crystallography[38,39] (Figure 1). *LigTh1519*, as most ATP-dependent DNA ligases, is a multidomain enzyme composed of three subdomains: the DNA binding domain (DBD), nucleotide binding domain (NBD), and OB-fold domain (OBD) (Figure 1). DNA ligases catalyze phosphodiester bond formation between adjacent 3′ hydroxide and 5′ phosphate groups in nicked single-stranded DNA,[36,40−43] which may occur as a result of natural replication, in so-called Okazaki fragments, or by DNA damaging agents.[36,40−43] Moreover, while *LigTh1519* has an optimum activity above 70 °C, the half-life of this enzyme at 94 °C is only a few minutes.[38]

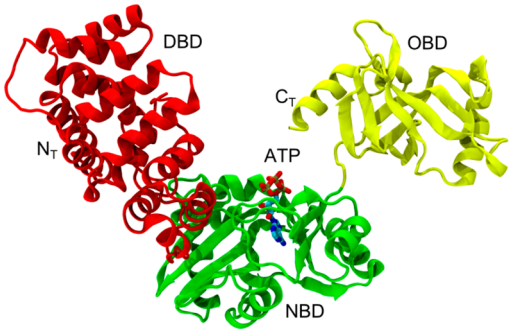


Figure 1. *LigTh1519* structure. *LigTh1519* (PDB ID: 3RR5)[39] is composed of three subdomains: the DNA binding domain (DBD) (in red), nucleotide binding domain (NBD) (in green), and OB-fold domain (OBD) (in yellow). The ATP binding site is located on the NBD. ATP is located at the binding site by superimposing the *LigTh1519* NBD onto that of the 2HIX structure, which shows a cocrystallized ATP molecule.

In order to understand the molecular origin of these properties, we first studied *LigTh1519* using MD simulations at different temperatures starting from its atomistic structure.[39] Results of this initial analysis revealed the most relevant thermosensitive regions, which we targeted using a set of rational design strategies in order to predict mutations able to increase *LigTh1519*'s thermostability. These predictions were eventually produced and tested *in vitro* and showed a remarkable and additive increase of the half-life time at 94 °C for the engineered enzyme (namely, up to about 41 min when up 4 mutations were simultaneously engineered), whereas 50% inactivation of parental *LigTh1519* was observed already upon 8 min incubation at this temperature.

MATERIALS AND METHODS

Computational Procedure. The structure of *LigTh1519* DNA ligase (PDB ID: 3RR5[39]) at 3 Å resolution was used for all MD simulations and analysis.[39] Selected mutated structures (i.e., A287K, G304D, S364I, and A387K) were built based on this X-ray structure using the Modeler package.[44] The X-ray structures were protonated under pH 7 conditions and inserted in a 11.6 × 8.6 × 10.8 nm^3 box of water molecules. $MgCl_2$ was added to neutralize the system and to adjust the ion concentration to 10 mM. The final system had ∼90 000 atoms and ∼27 000 water molecules. The systems were first energy-minimized with constrained C_α atoms and then without any constraints for 1000 steps. Then, to equilibrate the system, the temperature was increased gradually up to 300 K in the NVT ensemble and was kept at 300 K for 100 ps using a 1 fs time step. Finally, NPT MD simulations were run at 300 K for 500 ps with a 2 fs time step to complete the equilibration procedure. The equilibrated structures were thus used as a starting point for production (Figures S1 and S2).

All MD simulations were performed using the NAMD simulation package[45] with the Amber99SB force field[46] and a TIP3P water model.[47] Constant temperature was imposed by using Langevin dynamics with a damping coefficient of 5.0 ps. Constant pressure of 1 atm was maintained with Langevin piston dynamics using a 200 fs decay period and 50 fs time constant.[48,49] All production MD simulations were run at 1 bar with a time step of 2 fs, using the SHAKE algorithm[50] on all bonds and PME[51] for treating electrostatic interactions. Simulations of the wild-type enzyme were run at seven different temperatures, namely, 280, 300, 320, 360, 380, 400, and 500 K, for more than 10 ns in the NPT ensemble.

Salt bridge analysis, visualizations, and rendering of figures were performed using VMD.[52] For root-mean-square fluctuation (RMSF) analysis, ProDy package[53] was used. For bioinformatics analysis, BLAST[35] was used to find homologue sequences and ClustalW 2.0[54] was used to perform multiple sequence alignments. Analysis of multiple sequence alignments were performed using Python[55] and Biopython[56] packages. To evaluate the effect of mutations on the thermodynamic properties of the enzyme, we used FoldX, which is a fast and quantitative algorithm that uses an all-atom description of protein structures to estimate interactions that are important for protein stability.[57]

Plasmids, Bacterial Strains, and Culture Conditions. For expression of recombinant *LigTh1519* ligase in *Escherichia coli*, the corresponding gene was cloned in pQE30 (Qiagen), yielding expression vector pQE30-*LigTh1519*.[38] This plasmid allowed expression of the recombinant enzyme with an N-terminal 6-histidine tag. Site-directed mutagenesis of pQE30-*LigTh1519* was performed to introduce four mutations (singularly and altogether) in *LigTh1519*: A287K (GCC → AAA), G304D (GGC → GAT), S364I (AGC → ATT), and A387K (GCG → AAA). Amino acid numbers are given according to the sequence of the native protein (GenBank ACN59570). The corresponding vectors, pQE30-*LigTh1519*-

DOI: 10.1021/bi501227b
Biochemistry 2015, 54, 3076−3085

mut, pQE30-*LigTh1519-A287K*, pQE30-*LigTh1519-G304D*, pQE30-*LigTh1519-S364I*, and pQE30-*LigTh1519-A387K* were used to express mutant enzymes, *LigTh1519*mut, in *E. coli* cells.

Rosetta-gami (DE3) (Novagen) strain was applied for expression of recombinant ligase and its mutant derivatives. The *E. coli* strains were cultured in Luria−Bertani (LB) medium at 37 °C with shaking, and ampicillin was added to the medium at final concentration of 100 mg mL^{-1}.

Expression and Purification of *LigTh1519* and Its Mutants. The cells were grown at 37 °C in 10 mL of LB medium to the midexponential growth phase (OD_{600} of 0.5). Isopropyl β-D-1-thiogalactopyranoside was added at final concentration of 1 mM to induce gene expression, followed by an additional 12 h of incubation. The cells were then harvested by centrifugation and resuspended in 1 mL of 50 mM phosphate buffer (pH 8.0) containing 300 mM NaCl and 10 mM imidazole. The cells were disrupted by treatment with lysosyme (1 mg/mL) for 30 min and sonication, followed by centrifugation (8000*g*, 30 min, 4 °C). The supernatant was then mixed with 200 μL of a 50% slurry of Ni-NTA resin (Qiagen) and incubated with gentle agitation for 30 min at 4 °C. The mixture was centrifigated for 10 s at 1000*g* (4 °C) to pellet the resin. It was then washed twice with wash buffer (50 mM phosphate buffer, pH 8.0, 300 mM NaCl, 20 mM imidazole). The bound proteins were eluted twice in 200 μL of elution buffer (50 mM phosphate buffer, pH 8.0, 300 mM NaCl, 200 mM imidazole). The eluates were combined and dialyzed against 50 mM Tris-HCl buffer (pH 7.2) containing 1 mM DTT and 1 mM EDTA at 4 °C overnight in Slide-A-Lyser MINI dialysis units 3500 MVCO (Thermo Scientific). The protein concentration was determined by the Bradford method using BSA as a standard. Protein purity was then assessed by 10% SDS-PAGE.

DNA Ligation Assay. The substrate used in the ligation assays was a 80 bp DNA duplex, which contained an internal nick. It was produced with the annealing of 35-mer oligonucleotide LA-35bt (5′-CAG AGG ATT GTT GAC CGG CCC GTT TGT CAG CAA CG-3′) and 40-mer oligonucleotide LA-40FAM (5′-CGC ACC GTG ACG CCA AGC TTG CAT TCC TAC AGG TCG ACT C-3′) to complementary 80-mer oligonucleotide LA-80 (5′-CGT TGC TGA CAA ACG GGC CGG TCA ACA ATC CTC TGG AGT CGA CCT GTA GGA ATG CAA GCT TGG CGT CAC GGT GCG CCA AC-3′). The 35-mer oligonucleotide was labeled with biotin at its 3′ end and phosphorylated at its 5′ end, and the 40-mer oligonucleotide was labeled with fluorescein 6-FAM at its 5′ end. To produce the ligase substrate, the mixture was slowly cooled to room temperature after heating at 95 °C for 5 min. The molar ratio of the components in the annealing mixture was 1:1:1.

The assay for the nick-closing activity of DNA ligase was performed as previously described[38,58] with minor modifications. Thirty picomoles of fluorescein/biotin-labeled substrate was added to the assay mixture (final volume 30 μL) containing 40 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 100 μM ATP, and DNA ligase. To determine the DNA ligase activity, we made a series of dilutions of the investigated enzyme samples and used them in test assays. DNA ligase activity assay was conducted at 55 °C for 1 h. The reactions were stopped by adding EDTA to 50 mM followed by a 10 min incubation on ice.

Relative activity of DNA ligase was evaluated by measuring the fraction of duplex substrate in which the single-strand break between LA-35bt and LA-40FAM was sealed. The product of the reaction, a 75 bp oligonucleotide, contains both fluorescein and biotin and may be immobilized on streptavidin particles. For each assay sample, 0.1 mg of streptavidin magnetic beads (Dynabeads M-270 Streptavidin, Invitrogen) was prepared by washing three times with 1× B&W buffer (5 mM Tris-HCl, pH 7.5, 0.5 mM EDTA, 1 M NaCl) followed by resuspending in 2× B&W buffer. Thirty microliters of ligase assay sample was mixed with an equal volume of resuspended beads, and then the mixture was incubated at room temperature for 20 min with gentle rotation to bind biotinylated DNA. The beads were separated from the unbound DNA using a magnet and washed three times with 1× B&W buffer. Finally, immobilized DNA was denatured with 0.1 N NaOH, washed, and released from the beads by incubation at 65 °C for 5 min in 10 mM EDTA with 95% formamide. Fluorescent signal was acquired with a TBS-380 mini-fluorometer (Turner BioSystems, emission: 515−575 nm, excitation: 465−485 nm) calibrated with 1 pmol of oligonucleotide LA-40FAM.

Thermostability was evaluated by this assay after incubating the enzyme (1 pmol) in 40 mM Tris-HCl buffer (pH 7.5) with 10 mM $MgCl_2$ for various lengths of time at 94 °C. (The selected temperature of 94 °C is the most relevant from a practical stand point because a promising area of application of *LigTh1519* is in ligase chain reaction (LCR). One step of LCR (like PCR) is, in fact, denaturation of DNA by heating at 94 °C.) Following incubation, the sample was cooled on ice. The dsDNA template and 100 μM ATP were added, and the assay was conducted at 55 °C for 1 h followed by stopping the reaction with 50 mM EDTA. Fluorescence signal detected from ligated oligonucleotides provides quantitative evaluation of the activity of the ligase adjusted relative to the control samples without DNA ligase. The enzyme half-inactivation time calculation was based on a logarithmic plot by determining the time required for a 2-fold loss of activity in the linear phase.

RESULTS

To understand the origin of the thermostability of *LigTh1519*, MD simulations at different temperatures were first used to identify domains with the highest thermosensitivity. Then, to determine how mesophilic sequences have evolved to stabilize these thermosensitive regions, the *LigTh1519* sequence was compared with that of mesophilic homologues. Next, comparison with hyperthermophilic species was key to proposing a set of mutations that could increase the thermostability of *LigTh1519* without decreasing its activity when tested *in vitro*.

Origin of *LigTh1519* Thermostability. The NBD contains the catalytic site, and it is known to function even without the other two domains[36,40−43] (i.e., DBD and OBD shown in red and yellow, respectively, in Figure 1). Therefore, we mainly focused on this domain to study and engineer *LigTh1519* thermostability. First, to find those regions of the protein that are most sensitive to temperature changes, MD simulations at different temperatures were run. These simulations indicated that the NBD is fairly stable overall, showing very low root-mean-square deviation (RMSD) from the reference X-ray structure (Figure S1). RMSF analysis of MD trajectories was used to detect regions more affected by structural variations upon changes in temperature (Figure 2). From the RMSF analysis, we found that the most thermosensitive region is located between residues L300 and E330 (Figure 3). In addition, residues in the regions A230−Y250 and K380−L400 were shown to be thermosensitive,

DOI: 10.1021/bi501227b
Biochemistry 2015, 54, 3076−3085

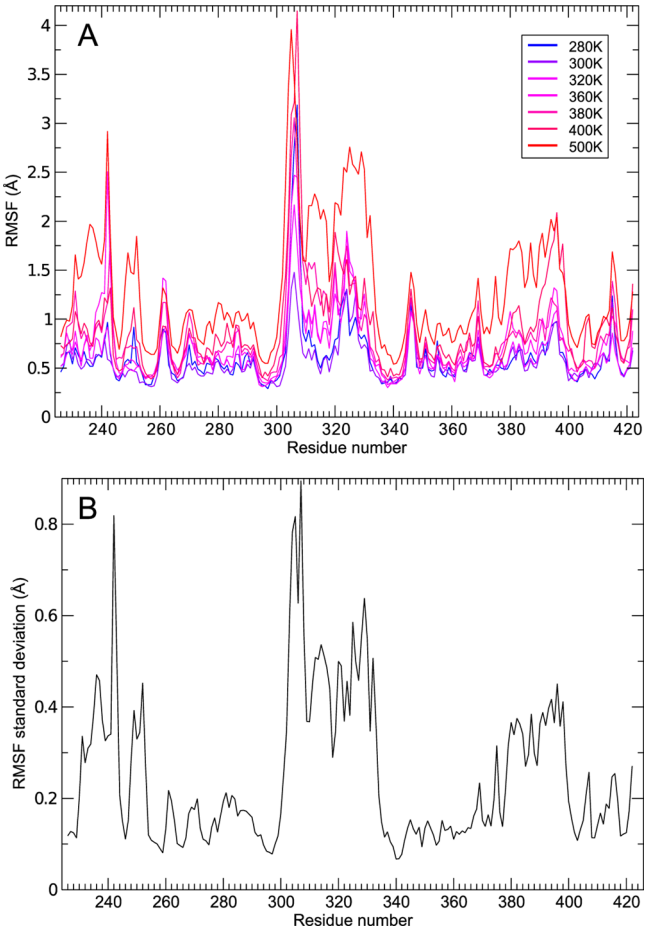


Figure 2. Thermosensitive analysis of *LigTh1519* based on MD simulations. (A) Root-mean-square fluctuations (RMSF) of NBD residues at different temperatures simulated by MD. The highest thermosensitivity is found between residues 300 and 333 in the NBD. (B) Standard deviation of RMSF values at the different temperatures examined.

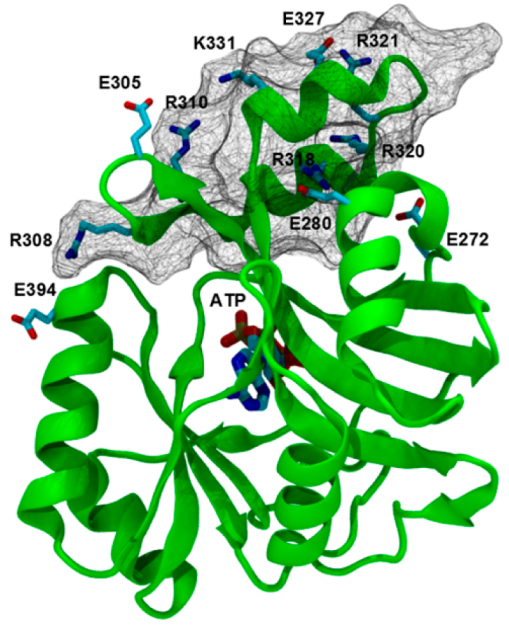


Figure 3. *LigTh1519* thermosensitive regions. Location of the most temperature-sensitive region in NBD is shown in wire-frame representation. Residues that could preserve the integrity of this thermosensitive region at high temperatures after comparison with other mesophiles are shown in licorice representation.

Table 1. Salt Bridge Network in *LigTh1519* and Substitutions in Mesophiles[a]

	272	305	308	310	320	321	327	331	394
LigTh1519	E	E	R	R	R	R	E	K	E
ZP_05975074	E	R	K	L	R	R	E	Q	A
NP_988090	E	D	K	R	R	R	K	K	S
NP_279843	E	E	D	L	R	R	A	E	E
YP_001827832	D	E	R	R	G	S	G	H	E
NP_217578	D	P	R	Q	G	R	A	T	A
NP_962051	D	P	S	Q	G	R	A	A	A
NP_625491	D	A	R	R	G	S	T	A	A
NP_828312	D	G	R	R	G	S	A	S	A

[a]Extracted from multiple sequence alignments (MSA) of *LigTh1519* with mesophilic species (see also Table S1) where the salt bridge network connectivity is indicated on the top. The relative recurrent substitutions in the network are reported for selected species.

although to a lesser extent (Figures 2 and 3). Secondary structure analysis of these thermosensitive regions indicated that there are a number of residues that are located on turns and are thus naturally flexible. However, residues located on helices (i.e., Q313−F319, E327−I332, and K380−G396) are more relevant and can have an active role in defining the temperature-dependent properties of the enzyme.

With the aim of determining if optimization of the salt bridge network is implicated in conferring more elevated thermostability, we next compared salt bridge forming residue pairs on *LigTh1519* with those of homologous mesophilic DNA ligases, mostly focusing on those pairs located in the detected temperature-sensitive regions (Figures 2 and 3). In order to do that, multiple sequence alignment (MSA) was performed with eight mesophilic homologue sequences extracted from the NCBI database[59] (Table S1). The salt bridge analysis and sequence comparison results are summarized in Table 1. In addition, Figure 3 shows the detected thermostabilizing salt bridge network for *LigTh1519* in its structure as well as its location in the detected thermosensitive region. Taken together, we found that an extended network of salt bridges (i.e., E280−R318, E272−R320, E305−K331, E305−R310, E327−R321, E327−K331, and E394−R308, Figure 3) is responsible for connecting the thermosensitive region (residues L300−E330) to the rest of the NBD.

Thermostability Engineering in *LigTh1519*. Using BLAST, we found ∼1000 homologous sequences for *LigTh1519* DNA ligase. They were classified as 6 psychrophilic, 883 mesophilic, 56 thermophilic, and 55 hyperthermophilic sequences. As *LigTh1519* DNA ligase is a thermophilic protein, and our goal was to increase thermostability, only hyperthermophilic sequences were chosen for multiple sequence alignment (Table S2). The distribution of amino acids at each position on the alignment was used to find key changes that function to enhance thermostability of this family of enzymes. This led us to focus on hydrophobic (or hydrophilic) residues on *LigTh1519* for which their counterparts in hyperthermophiles are hydrophilic (or hydrophobic) with more than 75% probability. Then, to find the most appropriate substitutions leading to enhanced thermostability, we analyzed the amino

DOI: 10.1021/bi501227b
Biochemistry 2015, 54, 3076−3085

Table 2. Predicted Thermostable Mutation for *LigTh1519*[a]

residue	HB	HL	distribution (%)	mutant candidates	selected
A287	0.11	0.89	K: 35.7; E: 28.6; A: 10.7; D: 8.9; N: 7.1; S: 3.6; R: 3.6; T: 1.8	K (ΔG = −0.65 kcal/mol) E (ΔG = −0.03 kcal/mol)	K
G304	0.18	0.82	D: 57.1; G: 14.3; K: 10.7; E: 5.4; N: 5.4; H: 3.6; I: 1.8; V: 1.8	D (ΔG = −0.61 kcal/mol)	D
S364	0.93	0.07	I: 57.1; V: 25.0; L: 7.1; A: 3.6; S: 3.6; T: 3.6T	I (ΔG = −1.06 kcal/mol)	I
A387	0.16	0.75	K: 21.4; E: 21.4; S: 10.7; Q: 8.9; A: 7.1; V: 7.1; N: 5.4; T: 3.6	K (ΔG = 0.04 kcal/mol) E (ΔG = −0.06 kcal/mol)	K

[a]In the first column, only hydrophobic (or hydrophilic) residues for which their counterparts in hyperthermophiles are hydrophilic (or hydrophobic) with more than 75% probability are shown. Percentage of hydrophobic or hydrophilic residues is indicated as HB and HL, respectively, in the second and third columns. The forth column reports the residue distribution at the selected site. Selected candidate(s) for mutation and the corresponding calculated free energy of mutation are shown in the fifth column.

acid composition in the alignment at selected positions, and, finally, amino acids with highest frequency at each position were chosen (Table 2). Following this screening protocol, the best candidate positions for increasing thermostability, obtained by comparing the target sequence with 55 hyperthermophilic sequences, are shown in Table 2. A287, G304, and A387 were hydrophobic residues for which their counterparts on the alignment were mainly hydrophilic (with a probability of hydrophilic amino acids of 89, 82, and 75%, respectively). On the other hand, S364 was the only hydrophilic amino acid with a probability of a hydrophobic counterpart occurring on the alignment as high as 93% (Table 2). In addition, to predict the effect of these mutations on the protein stability in terms of the relative change in free energy, the FoldX force field was used to estimate the $\Delta\Delta G$ contribution for the proposed mutations (Table 2).

Among the proposed position sites, we found that the most probable substitution for G304 is aspartate with 57.1% (Table 2). G304 is located on a turn facing an adjacent β-strand where two arginines (R308 and R310) are located. Moreover, K331 is located on a α-helix next to G304; thus, G304D appears to be a convenient choice to create a strong network of salt bridges in that region (Figure 3 and 4). S364 is another residue that is often mutated in hyperthermophiles' DNA ligases, with mutation to an isoleucine being the most probable option (namely, 57.1%). The molecular reason for this effect is likely to be found in the possibility that I364 joins and further stabilizes the hydrophobic core composed of V294, V296, I341, V344, L349, L361, and V365 (Figure 3).

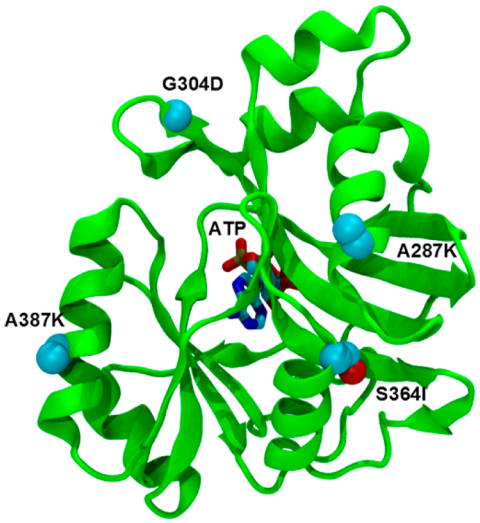


Figure 4. Thermostabilizing mutations in *LigTh1519*'s structure. Residues predicted to increase thermostability are shown in space filling representation. All of the proposed mutations are far from the ATP binding site and its interacting residues, which are located on the opposite face of the NBD.

Among other possible mutation sites, there are candidate substitutions that have a similar probability to occur but completely different chemical properties. For example, in the case of A287, two equally good substitutions are possible: lysine (35.7%) and glutamate (28.6%). To select the most appropriate mutations in these cases, the FoldX force field was used to rank the most suitable substitutions by accounting for interactions with neighboring residues. In particular, A287 is located on a helix (Figure 4), where it may form a salt bridge with E366. Thus, a lysine substitution is more likely to enhance thermostability, as confirmed by FoldX calculations, which indicate that A287K is energetically more favorable than A287E (ΔG_{A287K} = −0.65 kcal/mol vs ΔG_{A287E} = −0.03 kcal/mol, Table 2). A387 is another of these cases, with lysine and glutamate being equally possible mutations (both with 21.4%). However, in this case, FoldX calculations could not help to discriminate between these two possible substitutions, as it showed no difference between predicted values of ΔG_{A387K} and ΔG_{A387E} (Table 2). Because both methods, MSA analysis and FoldX calculations (Table S3), as well as MD analysis of the two mutants (see Discussion), showed that there is an equal chance for lysine and glutamate to stabilize the structure, we picked lysine for subsequent experimental evaluation. On the basis of this analysis, we therefore predicted a list of point mutations, namely, A287K, G304D, S364I, and A387K, to be tested experimentally for their ability to enhance thermostability.

Experimental Validation and Characterization of *LigTh1519* Mutants. The proposed mutations were introduced into the gene encoding *LigTh1519* ligase by site-directed mutagenesis. The mutations were introduced either separately or altogether to more precisely dissect their contribution on thermostability. The recombinant proteins, unmodified *LigTh1519*, single mutants (*LigTh1519-A287K*, *LigTh1519-G304D*, *LigTh1519-S364I*, and *LigTh1519-A387K*), and a mutant with all 4 combined mutations, *LigTh1519*mut (i.e., A287K, G304D, S364I, and A387K), were expressed in *E. coli* and purified by metal affinity chromatography on Ni-NTA resin (Qiagen). The purification of the enzymes was monitored by SDS-PAGE, which revealed a homogeneous 64.8 kDa major protein band, the size expected for the fusion product comprising the 63.38 kDa *LigTh1519* DNA ligase protein, and a 1.4 kDa peptide corresponding to the N-terminal His-tag.

DOI: 10.1021/bi501227b
Biochemistry 2015, 54, 3076−3085

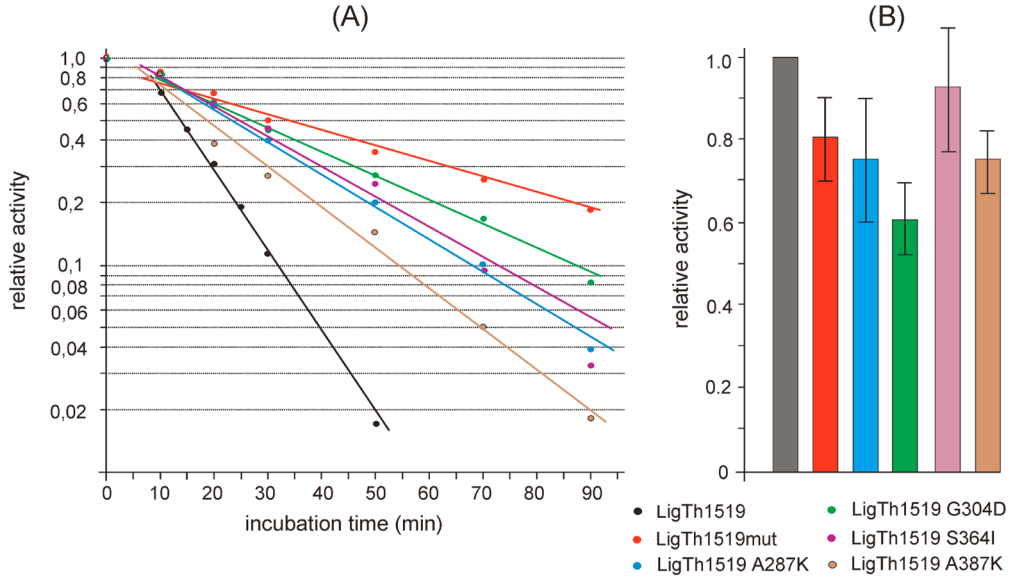


Figure 5. Thermostability and activity of unmodified and mutated *LigTh1519*. (A) Thermostability was evaluated after incubating the enzyme for various durations at 94 °C. The residual ligase activity of the treated samples was then determined by the aforementioned nick-closing activity assay. (B) Relative activities of equal amounts of purified enzymes were determined by the nick-closing activity assay. Data represent the means of six experiments (serial dilutions of enzymes), and error bars represent standard deviation.

The activity of *LigTh1519* ligase and all mutants was quantitatively determined by a nick-closing activity assay utilizing a fluorescein/biotin substrate (Figure 5A). It was found that the half-inactivation time was 41 min for the enzyme carrying all 4 combined mutations, whereas the unmodified enzyme lost half of its activity upon 8 min incubation at 94 °C. Enzymes with single mutations exhibited moderate, but non-negligible, increases in thermostability in comparison to that of *LigTh1519*mut, with half-inactivation times at 94 °C of about 15 min for *LigTh1519-A387K*, 20 min for *LigTh1519-A287K* and *LigTh1519-S364I*, and 25 min for *LigTh1519-G304* (Figure 5A), hinting at a quasi-additive effect of the individual mutations on thermostability. The thermostability data of *LigTh1519* at 94 °C are in agreement with that reported in a previous study.[38]

It is known that thermostable enzymes often have lower specific activity than their mesophilic counterparts since the increased rigidity of the structure may reduce catalytic efficiency. In particular, a *LigTh1519* mutant was designed to result in more salt bridges to stabilize the enzyme's structure. Therefore, we evaluated the relative activities of *LigTh1519* ligase and its mutants in nick-closing assay. Equal amounts of purified enzymes (five different concentrations) were assayed, and it was found that *LigTh1519*mut retained about 80% activity (Figure 5B). Activities of the other individual mutants were in the range of 70−90% of that of wild type except for *LigTh1519-G304*, which exhibited about 60% of the maximal activity (Figure 5B).

DISCUSSION

As high temperature is a favorable condition for many industrial processes, the issue of enzymatic thermostability is a sensitive topic in this domain. Thus, hyperthermophilic enzymes isolated from hyperthermophilic species provide a naturally evolved and efficient tool for a variety of industrial and biotechnological applications. Nonetheless, the ability of computational and experimental techniques to further enhance thermostability would be beneficial to meet a wider array of industrial demands even when naturally isolated enzymes for such purposes are lacking.

Among isolated enzymes, DNA ligase has been long used in biology and biotechnology, specifically for PCR and LCR.[60,61] It was shown, for instance, that amplification of long DNA fragments could be enhanced if a thermostable DNA ligase is used to add a ligation step in PCR cycles. On the basis of this evidence, a new design was introduced for PCR amplification of circular plasmids using a thermostable DNA ligase.[60,61] Although there have been some reports on the isolation and purification of DNA ligases from hyper/thermophilic species,[39,58] to our knowledge, the work presented here is the first that addresses the thermostability of a thermophilic DNA ligase with the aim of enhancing its thermophilic properties via computational-based protein engineering.

To study and enhance the thermal stability of an enzyme, deep knowledge of its unfolding mechanism is needed. Mechanically, weak points in a protein's structure are the most prone to initiating and leading to unfolding at high temperatures. Thus, the determination and optimization of these weak regions have been the focus of several studies analyzing the thermostability of proteins.[26−34] B-factor analysis of crystallographic structures has been shown to be useful to identify flexible residues that can be targeted for thermostability engineering.[27,62−67] The dynamic information included in B-factors may, however, be affected by artifacts of crystal packing and does not always reflect the native dynamics in solution. Nonetheless, the integration of different methods has been shown to increase the chance of finding mutations that lead to more thermostable proteins.[63,65] Thus, RMSF analysis extracted from MD simulations can be used as an alternative to B-factor analysis to find the most flexible residues while at the same time probing the effect of multiple temperatures and regions that are more thermosensitive.[26,30] Here, we used this strategy with *LigTh1519* and found that a thermosensitive region is located in the region between L300 and E330 of the NBD (Figures 2 and 3).

DOI: 10.1021/bi501227b
Biochemistry 2015, 54, 3076−3085

To make thermosensitive regions that can withstand harsh conditions, nature uses different strategies. Several studies have addressed the differences between mesophilic and thermophilic homologous proteins.[1−4,7,8] On the basis of a number of reports, one of the most common avenues taken by nature consists of altering charged residues on a protein's surface to increase the number of salt bridges.[1−4,7,8] Our study, based on multiple sequence alignments, confirmed that in *LigTh1519* this network is significantly more populated than it is in mesophilic counterparts of *LigTh1519*. In particular, there is a strong salt bridge network that preserves the connection of the thermosensitive region from residues L300 to E330 (Figure 3) at the NBD, presumably favoring the integrity of the domain at more elevated temperatures. Specifically, E280 and R318, which are conserved in all of the mesophilic sequences, form a salt bridge that defines a minimal scaffold that helps to preserve the connection of this domain to the rest of the enzyme. Then, the addition of more salt bridges in thermophilic homologues, such as *LigTh1519*, to this minimum scaffold can enhance the strength of this network, preserving the integrity of protein subdomains against thermal perturbations. In fact, in *LigTh1519*, a series of additional salt bridges (i.e., E272−R320, E305−K331, E305−R310, E327−R321, E327−K331, and E394−R308) is observed between this subdomain and the rest of the NBD, keeping the structure tight at that location (Table 1 and Figure 3).

In particular, MD simulations of the individual mutants and *LigTh1519*mut system revealed that G304D in this region of the NBD showed a significant change in flexibility (Figure S3), likely due to the production of a tighter salt bridge network (e.g., D304 interacts with R310 and R308 with 53% occupancy). This is also consistent with the fact that this substitution is readily picked up by our bioinformatics analysis and that this mutant is the one that individually seems to affect the half-life at 94 °C more significantly (Figure 5A). For the other two substitutions of this kind at positions 287 and 387, MD analysis of the single and combined mutations showed a less pronounced variation in flexibility (if not an increase, as for K387, Figure S3), likely reflecting the fact that the newly formed salt bridge network is not as tight as that in the G304D mutant (35% occupancy for K287 and 35% for K387) and that they are located in areas of the enzyme that are more solvent-exposed and less connected with the rest of the enzyme in comparison to G304D (Figure 3). This feature is qualitatively consistent with the fact that these 2 mutations seem to individually account for the overall thermostabilization of *LigTh1519* to a lesser extent (Figure 5A).

The identification of the thermosensitive regions and their stabilizing factors could be used as a lead; however, it cannot provide exhaustive information for proposing mutations to enhance thermostability. Although MD simulation is useful to locate candidate positions on the structure that are more temperature-sensitive and can be targeted for thermostability engineering (as for position 304 in our case), it cannot directly propose appropriate substitutions to enhance thermostability. To predict suitable substitutions, more general strategies based on comparison with sequences from highly thermostable proteins are useful. Comparison with sequences from highly thermostable proteins also reduces the probability of perturbing key conserved residues, such as, for instance, those directly involved in the enzyme's catalytic activity. In *LigTh1519*, five different highly conserved motifs are, in fact, located in the NBD that form the catalytic site. These motifs are essential for the three consecutive steps of DNA ligation (Figure 1), and substitutions in these regions can significantly impact the enzymatic activity. This could explain the fact that, while it enhances thermostability to a greater extent among the single mutants tested (Figure 5A), G304D affects the ligase's enzymatic activity the most, likely due to its closer proximity to the active site (Figures 3 and 5B).

Specifically, thermostability analysis of *LigTh1519* by MD simulations at different temperatures showed that the most thermally sensitive regions are located only in the NBD (Figure 2), whereas in the DBD and OBD, RMSF peaks are mainly located on coils and turns. It is also noteworthy that multiple sequence alignment confirms this point by suggesting mutations that are mostly located in the NBD (residues M225−K420, Table 1), indicating that the detected thermosensitive region has been targeted for thermostabilization by nature. However, as the mutation candidates are based on major differences among amino acid content, there is little chance of proposing any mutation in more conserved motifs, such as the catalytic site (Figure 4). This safe selection of mutation sites could be the main reason why our engineered DNA ligases did not show a significant decrease in their catalytic activity. In addition, using MSA, we detected and proposed mutations that were impossible to detect only by using RMSF analysis. Among the proposed mutations (Table 2), G304 is immediately picked up by the MD-based analysis (Figure 2), whereas on the other hand, S364 emerged only from multiple sequence alignments and was almost impossible to detect by using only RMSF analysis. While sequence comparison could be used to propose mutations that are able to optimize the hydrophobic core, thermosensitivity analysis using MD simulations was not very informative regarding this kind of interaction. All together, these specific results show that MD simulations and sequence comparisons have to be used in synergy to obtain deeper insights into enzyme stability and thermostability engineering.

In addition, using a sequence comparison approach does not always result in the identification of a single mutation for each candidate position in the sequence. Here, we used FoldX to rank the proposed mutations, a strategy that turned out to give reasonable estimations given the intrinsic error associated with the calculations.[68] In the event that energy calculations could not unequivocally select a suitable mutation, as for A387 in our case, MD simulations for both of the candidate mutants, K387 and E387, revealed very similar flexibility behavior resulting from the multiple and equal possibilities of charged residues in this region forming salt bridges with the adjacent residues from the same helix (Figure 3). In particular, in MD, K387 can form salt bridges with E383 and E384, whereas the E387 mutant can do the same with the basic residues K390 and R391. On the basis of this evidence, we predict that K387 and E387 can have an equal role in increasing thermostability. From the thermostability assessment of individual mutations (Figure 5A), both mutants are expected to have a similar, moderate effect (∼15 min) on the global enhancement of the half-inactivation time at 94 °C.

In summary, a comparative *in silico* and *in vitro* analysis of the individual and combined impact of A287K, G304D, S364I, and A387K mutations allowed us to better dissect their contribution to *LigTh1519*'s thermostability. We found that each of them represents relevant sites, invariantly producing a non-negligible, although moderate, contribution to thermostability. While the wild-type ligase has a half-life time at 94 °C of 8 min, the

DOI: 10.1021/bi501227b
Biochemistry 2015, 54, 3076−3085

A387K mutant reaches 15 min, S364I and A287K reach 20 min, and G304D reaches 25 min (Figure 5A). A precise analysis to disentangle the contribution of each single mutation remains difficult, mainly because their nontrivial, reciprocal spatial and dynamic relationship; however, we can conclude that each mutation seems to have a quasi-additive effect on the overall half-life time at 94 °C, resulting in a total increase of 41 min for the enzyme carrying the 4 combined mutations (*LigTh1519*-mut).

CONCLUSIONS

In conclusion, in this study, we used computational techniques to understand more deeply the origin of thermostability in *LigTh1519* DNA ligase as a thermostable enzyme. We engineered it to achieve higher thermal stability using rational design together with what we learned from the origin of the thermostability of the wild type. Our results showed that an extended salt bridge network in the NBD plays a crucial role in preserving the stability of *LigTh1519* at high temperatures over ∼70 °C by keeping the thermosensitive region connected to the enzyme's body. Finally, we suggested and tested four mutations, A287K, G304D, S364I, and A387K, with regard to their ability to enhance the thermostability of the enzyme. Structural and dynamic analyses showed that A287K, G304D, and A387K mutations play a role in thermostability through electrostatically optimizing the surface, whereas S364I contributes to optimizing the hydrophobic core of the enzyme. Experimental results proved that the proposed substitutions enhance thermostability of the protein by increasing in an additive fashion the half-inactivation time at 94 °C from 8 to 41 min without significantly reducing the specific catalytic activity of the enzyme.

The findings reported here emphasize the importance and capability of molecular simulations combined with bioinformatics analyses to reveal sources of stability/instability and to predict mutations able to enhance thermostability without affecting enzymatic activity. Therefore, the engineered hyperthermostable DNA ligase derived in this work could be used to optimize further PCR protocols.[60,61] Furthermore, the approach presented here and applied to *LigTh1519* could be similarly adopted for other enzymes and can contribute to systematically producing more hyperthermostable enzymes for use in industrial settings.

ASSOCIATED CONTENT

Supporting Information

Figure S1: RMSD of the NBD at different temperatures during MD simulations. Figure S2: RMSD of the NBD for MD simulations of the wild type, the 4 predicted mutations alone, and the four mutants combined. Figure S3: RMSF of the NBD for MD simulations of the wild type, the 4 predicted mutations alone, and the four mutants combined. Table S1: List of accession codes of mesophilic homologues. Table S2: List of accession codes of sequences belonging to hyperthermophilic mesophilic species. Table S3: Nonbonded interaction energy for residues at positions 287 and 387 for the wild type and mutant systems. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/bi501227b.

AUTHOR INFORMATION

Corresponding Authors

*(K.G.S.) E-mail: skryabin@biengi.ac.ru; Phone: 007 4991357319.

*(M.D.P.) E-mail: matteo.dalperaro@epfl.ch; Phone: 0041 216931861.

Funding

This work was supported by the Scientific and Technological Cooperation Programme Switzerland-Russia [IZLRZ3_128371]; the Ministry of Education and Sciences of Russia [14.120.14.6150-NSh]; the Russian Science Fund grant 14-24-00172 (project of V.O.P.); and by the program "Molecular and Cellular Biology" of the Russian Academy of Science (project of N.V.R.).

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Vieille, C., and Zeikus, G. J. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev. 65*, 1−43.

(2) Bouzas, T. D., Barros-Velazquez, J., and Villa, T. G. (2006) Industrial applications of hyperthermophilic enzymes: a review. *Protein Pept. Lett. 13*, 645−651.

(3) Brock, T. D. (1967) Life at High Temperatures. *Science 158*, 1012−1019.

(4) Haki, G. D., and Rakshit, S. K. (2003) Developments in industrially important thermostable enzymes: a review. *Bioresour. Technol. 89*, 17−34.

(5) Demirjian, D. C., Moris-Varas, F., and Cassidy, C. S. (2001) Enzymes from extremophiles. *Curr. Opin. Chem. Biol. 5*, 144−151.

(6) Zamost, B. L., Nielsen, H. K., and Starnes, R. L. (1991) Thermostable enzymes for industrial applications. *J. Ind. Microbiol. 8*, 71−81.

(7) Sawle, L., and Ghosh, K. (2011) How do thermophilic proteins and proteomes withstand high temperature? *Biophys. J. 101*, 217−227.

(8) Petsko, G. A. (2001) Structural basis of thermostability in hyperthermophilic proteins, or "there's more than one way to skin a cat". *Methods Enzymol. 334*, 469−478.

(9) Kumar, S., and Nussinov, R. (2001) How do thermophilic proteins deal with heat? *Cell. Mol. Life Sci. 58*, 1216−1233.

(10) Razvi, A., and Scholtz, J. M. (2006) Lessons in stability from thermophilic proteins. *Protein Sci. 15*, 1569−1578.

(11) Zhou, H. X. (2002) Toward the physical basis of thermophilic proteins: linking of enriched polar interactions and reduced heat capacity of unfolding. *Biophys. J. 83*, 3126−3133.

(12) Das, R., and Gerstein, M. (2000) The stability of thermophilic proteins: a study based on comprehensive genome comparison. *Funct. Integr. Genomics 1*, 76−88.

(13) Gianese, G., Bossa, F., and Pascarella, S. (2002) Comparative structural analysis of psychrophilic and meso- and thermophilic enzymes. *Proteins 47*, 236−249.

(14) Zeldovich, K. B., Berezovsky, I. N., and Shakhnovich, E. I. (2007) Protein and DNA sequence determinants of thermophilic adaptation. *PLoS Comput. Biol. 3*, 62−72.

(15) Ladenstein, R., and Antranikian, G. (1998) Proteins from hyperthermophiles: stability and enzymatic catalysis close to the boiling point of water. *Adv. Biochem. Eng./Biotechnol. 61*, 37−85.

(16) Adams, M. W. W. (1993) Enzymes and proteins from organisms that grow near and above 100 degree C. *Annu. Rev. Microbiol. 47*, 627−658.

(17) Paiardini, A., Sali, R., Bossa, F., and Pascarella, S. (2008) "Hot cores" in proteins: comparative analysis of the apolar contact area in structures from hyper/thermophilic and mesophilic organisms. *BMC Struct. Biol. 8*, 8−14.

DOI: 10.1021/bi501227b
Biochemistry 2015, 54, 3076−3085

(18) Greaves, R. B., and Warwicker, J. (2007) Mechanisms for stabilisation and the maintenance of solubility in proteins from thermophiles. *BMC Struct. Biol. 7*, 7−18.
(19) Chakravarty, S., and Varadarajan, R. (2002) Elucidation of factors responsible for enhanced thermal stability of proteins: a structural genomics based study. *Biochemistry 41*, 8152−8161.
(20) Chen, J. M., Lu, Z. Q., Sakon, J., and Stites, W. E. (2000) Increasing the thermostability of staphylococcal nuclease: implications for the origin of protein thermostability. *J. Mol. Biol. 303*, 125−130.
(21) Dumon, C., Varvak, A., Wall, M. A., Flint, J. E., Lewis, R. J., Lakey, J. H., Morland, C., Luginbuhl, P., Healey, S., Todaro, T., DeSantis, G., Sun, M., Parra-Gessert, L., Tan, X. Q., Weiner, D. P., and Gilbert, H. J. (2008) Engineering hyperthermostability into a GH11 xylanase is mediated by subtle changes to protein structure. *J. Biol. Chem. 283*, 22557−22564.
(22) Tian, J. A., Wang, P., Gao, S., Chu, X. Y., Wu, N. F., and Fan, Y. L. (2010) Enhanced thermostability of methyl parathion hydrolase from *Ochrobactrum* sp. M231 by rational engineering of a glycine to proline mutation. *FEBS J. 277*, 4901−4908.
(23) Watanabe, K., Chishiro, K., Kitamura, K., and Suzuki, Y. (1991) Proline residues responsible for thermostability occur with high-frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase from *Bacillus thermoglucosidasius* Kp1006. *J. Biol. Chem. 266*, 24287−24294.
(24) Pace, C. N. (2000) Single surface stabilizer. *Nat. Struct. Biol. 7*, 345−346.
(25) Vazquez-Figueroa, E., Yeh, V., Broering, J. M., Chaparro-Riggers, J. F., and Bommarius, A. S. (2008) Thermostable variants constructed via the structure-guided consensus method also show increased stability in salts solutions and homogeneous aqueous-organic media. *Protein Eng., Des. Sel. 21*, 673−680.
(26) Huang, X., Gao, D., and Zhan, C. G. (2011) Computational design of a thermostable mutant of cocaine esterase via molecular dynamics simulations. *Org. Biomol. Chem. 9*, 4138−4143.
(27) Joo, J. C., Pack, S. P., Kim, Y. H., and Yoo, Y. J. (2011) Thermostabilization of *Bacillus circulans* xylanase: computational optimization of unstable residues based on thermal fluctuation analysis. *J. Biotechnol. 151*, 56−65.
(28) Bradley, E. A., Stewart, D. E., Adams, M. W. W., and Wampler, J. E. (1993) Investigations of the thermostability of rubredoxin models using molecular-dynamics simulations. *Protein Sci. 2*, 650−665.
(29) Meharenna, Y. T., and Poulos, T. L. (2010) Using molecular dynamics to probe the structural basis for enhanced stability in thermal stable cytochromes P450. *Biochemistry 49*, 6680−6686.
(30) Kundu, S., and Roy, D. (2010) Structural study of carboxylesterase from hyperthermophilic bacteria *Geobacillus stearothermophilus* by molecular dynamics simulation. *J. Mol. Graphics Modell. 28*, 820−827.
(31) Kundu, S., and Roy, D. (2009) Comparative structural studies of psychrophilic and mesophilic protein homologues by molecular dynamics simulation. *J. Mol. Graphics Modell. 27*, 871−880.
(32) Rahman, M. B. A., Karjiban, R. A., Salleh, A. B., Jacobs, D., Basri, M., Chor, A. L. T., Wahab, H. A., and Rahman, R. N. Z. R. A. (2009) Deciphering the flexibility and dynamics of *Geobacillus zalihae* strain T1 lipase at high temperatures by molecular dynamics simulation. *Protein Pept. Lett. 16*, 1360−1370.
(33) Spiwok, V., Lipovova, P., Skalova, T., Duskova, J., Dohnalek, J., Hasek, J., Russell, N. J., and Kralova, B. (2007) Cold-active enzymes studied by comparative molecular dynamics simulation. *J. Mol. Model. 13*, 485−497.
(34) Purmonen, M., Valjakka, J., Takkinen, K., Laitinen, T., and Rouvinen, J. (2007) Molecular dynamics studies on the thermostability of family 11 xylanases. *Protein Eng., Des. Sel. 20*, 551−559.
(35) Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol. 215*, 403−410.
(36) Wilkinson, A., Day, J., and Bowater, R. (2001) Bacterial DNA ligases. *Mol. Microbiol. 40*, 1241−1248.
(37) Maunders, M. J. (1993) DNA and RNA ligases (EC 6.5.1.1, EC 6.5.1.2, and EC 6.5.1.3). *Methods Mol. Biol. 16*, 213−230.
(38) Smagin, V. A., Mardanov, A. V., Bonch-Osmolovskaia, E. A., and Ravin, N. V. (2008) [Isolation and characteristics of new thermostable DNA ligase from archaea of the genus Thermococcus]. *Prikl Biokhim Mikrobiol. 44*, 523−528.
(39) Petrova, T., Bezsudnova, E. Y., Boyko, K. M., Mardanov, A. V., Polyakov, K. M., Volkov, V. V., Kozin, M., Ravin, N. V., Shabalin, I. G., Skryabin, K. G., Stekhanova, T. N., Kovalchuk, M. V., and Popov, V. O. (2012) ATP-dependent DNA ligase from *Thermococcus* sp 1519 displays a new arrangement of the OB-fold domain. *Acta Crystallogr., Sect. F 68*, 1440−1447.
(40) Tomkinson, A. E., Vijayakumar, S., Pascal, J. M., and Ellenberger, T. (2006) DNA ligases: structure, reaction mechanism, and function. *Chem. Rev. 106*, 687−699.
(41) Doherty, A. J., and Suh, S. W. (2000) Structural and mechanistic conservation in DNA ligases. *Nucleic Acids Res. 28*, 4051−4058.
(42) Ellenberger, T., and Tomkinson, A. E. (2008) Eukaryotic DNA ligases: structural and functional insights. *Annu. Rev. Biochem. 77*, 313−338.
(43) Martin, I. V., and MacNeill, S. A. (2002) ATP-dependent DNA ligases. *Genome Biol. 3*, reviews3005.
(44) Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U., and Sali, A. (2006) Comparative protein structure modeling using Modeller, in *Current Protocols in Bioinformatics* (Baxevanis, A. D., Ed.) Chapter 5, Unit 5.6, Wiley, New York.
(45) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781−1802.
(46) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins 65*, 712−726.
(47) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(48) Martyna, G. J., Tobias, D. J., and Klein, M. L. (1994) Constant pressure molecular dynamics algorithms. *J. Chem. Phys. 101*, 4177−4189.
(49) Feller, S. E., Zhang, Y. H., Pastor, R. W., and Brooks, B. R. (1995) Constant pressure molecular dynamics simulation—the Langevin piston method. *J. Chem. Phys. 103*, 4613−4621.
(50) Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys. 23*, 327−341.
(51) Darden, T., Perera, L., Li, L., and Pedersen, L. (1999) New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure 7*, R55−60.
(52) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graphics Modell. 14*, 33−38.
(53) Bakan, A., Meireles, L. M., and Bahar, I. (2011) ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics 27*, 1575−1577.
(54) Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and clustal X version 2.0. *Bioinformatics 23*, 2947−2948.
(55) Van Rossum, G., and Drake, F. L. *Python Language Reference Manual*, Network Theory Limited, Bristol, 2003.
(56) Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., and de Hoon, M. J. L. (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics 25*, 1422−1423.

DOI: 10.1021/bi501227b
Biochemistry 2015, 54, 3076−3085

(57) Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005) The FoldX web server: an online force field. *Nucleic Acids Res. 33*, W382−W388.
(58) Lauer, G., Rudd, E. A., Mckay, D. L., Ally, A., Ally, D., and Backman, K. C. (1991) Cloning, nucleotide-sequence, and engineered expression of *Thermus thermophilus* DNA ligase, a homolog of *Escherichia coli* DNA ligase. *J. Bacteriol. 173*, 5047−5053.
(59) Geer, L. Y., Marchler-Bauer, A., Geer, R. C., Han, L. Y., He, J., He, S. Q., Liu, C. L., Shi, W. Y., and Bryant, S. H. (2010) The NCBI BioSystems database. *Nucleic Acids Res. 38*, D492−D496.
(60) Le, Y. L., Peng, J. J., Pei, J. J., Li, H. Z., Duan, Z. Y., and Shao, W. L. (2010) Properties of an NAD^+-dependent DNA ligase from the hyperthermophile *Thermotoga maritima* and its application in PCR amplification of long DNA fragments. *Enzyme Microb. Technol. 46*, 113−117.
(61) Yilin, L., Huayou, C., Robert, Z., David, W. J. H., and Weilan, S. (2013) Thermostable DNA ligase-mediated PCR production of circular plasmid (PPCP) and its application in directed evolution via in situ error-prone PCR. *DNA Res. 20*, 375−382.
(62) Kim, S. J., Lee, J. A., Joo, J. C., Yoo, Y. J., Kim, Y. H., and Song, B. K. (2010) The development of a thermostable CiP (*Coprinus cinereus* peroxidase) through in silico design. *Biotechnol. Prog. 26*, 1038−1046.
(63) Kim, H. S., Quang, A. T. L., and Kim, Y. H. (2010) Development of thermostable lipase B from Candida antarctica (CalB) through in silico design employing B-factor and RosettaDesign. *Enzyme Microb. Technol. 47*, 1−5.
(64) Parthasarathy, S., and Murthy, M. R. N. (2000) Protein thermal stability: insights from atomic displacement parameters (B values). *Protein Eng. 13*, 9−13.
(65) Blum, J. K., Ricketts, M. D., and Bommarius, A. S. (2012) Improved thermostability of AEH by combining B-FIT analysis and structure-guided consensus method. *J. Biotechnol. 160*, 214−221.
(66) Zhang, J. H., Lin, Y., Sun, Y. F., Ye, Y. R., Zheng, S. P., and Han, S. Y. (2012) High-throughput screening of B factor saturation mutated *Rhizomucor miehei* lipase thermostability based on synthetic reaction. *Enzyme Microb. Technol. 50*, 325−330.
(67) Siglioccolo, A., Gerace, R., and Pascarella, S. (2010) "Cold spots" in protein cold adaptation: Insights from normalized atomic displacement parameters (B′-factors). *Biophys. Chem. 153*, 104−114.
(68) Christensen, N. J., and Kepp, K. P. (2012) Accurate stabilities of laccase mutants predicted with a modified FoldX protocol. *J. Chem. Inf. Model. 52*, 3028−3042.

DOI: 10.1021/bi501227b
Biochemistry 2015, 54, 3076−3085