Check for updates

Thermostable DNA ligases from hyperthermophiles in biotechnology

Jingru Shi[1], Philippe M. Oger[2], Peng Cao[3] and Likui Zhang[1]*

[1]College of Environmental Science and Engineering, Yangzhou University, Yangzhou, China, [2]University of Lyon, INSA de Lyon, CNRS UMR, Villeurbanne, France, [3]Faculty of Environment and Life, Beijing University of Technology, Beijing, China

DNA ligase is an important enzyme ubiquitous in all three kingdoms of life that can ligate DNA strands, thus playing essential roles in DNA replication, repair and recombination *in vivo*. *In vitro*, DNA ligase is also used in biotechnological applications requiring in DNA manipulation, including molecular cloning, mutation detection, DNA assembly, DNA sequencing, and other aspects. Thermophilic and thermostable enzymes from hyperthermophiles that thrive in the high-temperature (above 80°C) environments have provided an important pool of useful enzymes as biotechnological reagents. Similar to other organisms, each hyperthermophile harbors at least one DNA ligase. In this review, we summarize recent progress on structural and biochemical properties of thermostable DNA ligases from hyperthermophiles, focusing on similarities and differences between DNA ligases from hyperthermophilic bacteria and archaea, and between these thermostable DNA ligases and non-thermostable homologs. Additionally, altered thermostable DNA ligases are discussed. Possessing improved fidelity or thermostability compared to the wild-type enzymes, they could be potential DNA ligases for biotechnology in the future. Importantly, we also describe current applications of thermostable DNA ligases from hyperthermophiles in biotechnology.

KEYWORDS

DNA ligase, ligase cycling reaction, DNA manipulation, thermostability, hyperthermophiles
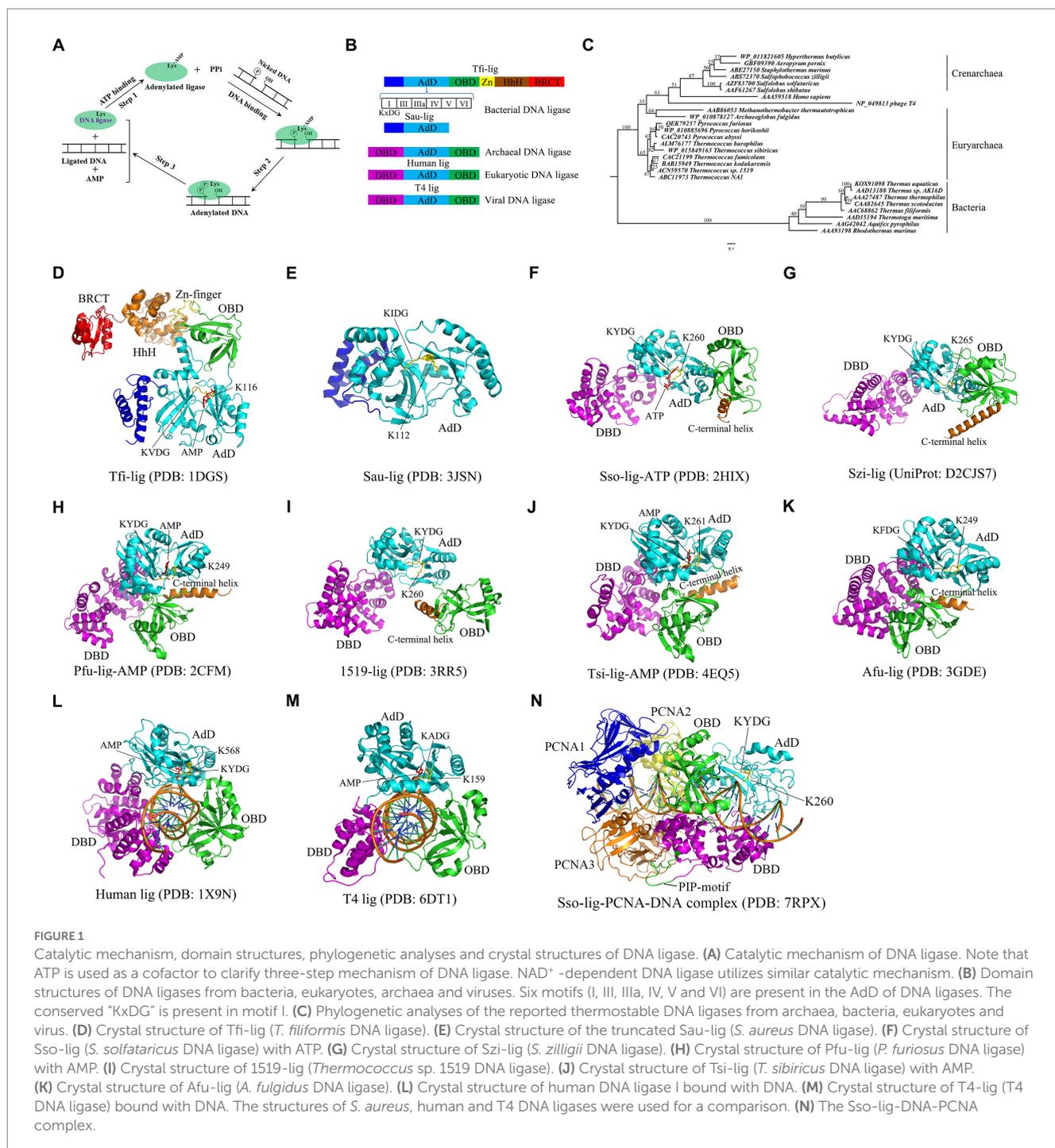
Introduction

DNA ligase is one of essential enzymes involved in DNA replication, repair and recombination *in vivo*, thus playing important roles in maintaining genomic integrity. DNA ligase is widespread in bacteria, archaea, eukaryotes, and even viruses. All the reported DNA ligases employ a common three-step mechanism to form a phosphodiester bond using either ATP or NAD+ as a high-energy cofactor (Figure 1A; Williamson and Leiros, 2019; Williamson and Leiros, 2020). The first step is a nucleophilic attack on 5′-phosphate of ATP or NAD+ cofactor by the highly conserved lysine in DNA ligase, thus yielding the ligase-AMP intermediate and releasing PPi. The second step is to transfer AMP from the adenylated ligase to the 5′-phosphate of one strand, thereby leading to the formation of the adenylated DNA intermediate. Next, the third step is to join the DNA strands by forming a phosphodiester bond between the 5′-phosphate group of the adenylated DNA strand and the 3′-hydroxyl group of the second strand, thus releasing AMP (Williamson and Leiros, 2019). Based on their cofactor utilization specificity, DNA ligases are divided into two families: ATP-dependent DNA ligases

and NAD$^+$-dependent DNA ligases. While ATP-dependent DNA ligases are typically present in archaea, eukaryotes, and viruses, including bacteriophages, NAD$^+$-dependent DNA ligases are predominately found in bacteria and some eukaryotic viruses.

All the reported DNA ligases possess the adenylation domain (AdD) and oligomer-binding domain (OBD) (Figure 1B), which are responsible for the formation of the ligase-AMP intermediate and DNA binding, respectively. Interestingly, bacterial DNA ligases appear to harbor more conserved motifs than archaeal, eukaryotic and viral DNA ligases. For example, the DNA ligase from the hyperthermophilic

bacterium *Thermus filiform* possesses a domain containing a Zn finger motif and a HhH (Helix-hairpin-Helix) motif, and a BRCT (BRCA1 C-terminus) domain in addition to AdD and OBD (Figure 1B; Lee et al., 2000). Notably, six conserved motifs (I, III, IIIa, IV, V, and VI) are present in the AdD in DNA ligase. The "KxDG" sequence is highly conserved in the conserved motif I of DNA ligase, among which the invariable residue lysine is essential for the formation of the ligase-AMP intermediate.

In vitro, DNA ligases are used for molecular cloning, DNA assembly, DNA sequencing and mutation detection associated with genetic

FIGURE 1
Catalytic mechanism, domain structures, phylogenetic analyses and crystal structures of DNA ligase. **(A)** Catalytic mechanism of DNA ligase. Note that ATP is used as a cofactor to clarify three-step mechanism of DNA ligase. NAD$^+$-dependent DNA ligase utilizes similar catalytic mechanism. **(B)** Domain structures of DNA ligases from bacteria, eukaryotes, archaea and viruses. Six motifs (I, III, IIIa, IV, V and VI) are present in the AdD of DNA ligases. The conserved "KxDG" is present in motif I. **(C)** Phylogenetic analyses of the reported thermostable DNA ligases from archaea, bacteria, eukaryotes and virus. **(D)** Crystal structure of Tfi-lig (*T. filiformis* DNA ligase). **(E)** Crystal structure of the truncated Sau-lig (*S. aureus* DNA ligase). **(F)** Crystal structure of Sso-lig (*S. solfataricus* DNA ligase) with ATP. **(G)** Crystal structure of Szi-lig (*S. zilligii* DNA ligase). **(H)** Crystal structure of Pfu-lig (*P. furiosus* DNA ligase) with AMP. **(I)** Crystal structure of 1519-lig (*Thermococcus* sp. 1519 DNA ligase). **(J)** Crystal structure of Tsi-lig (*T. sibiricus* DNA ligase) with AMP. **(K)** Crystal structure of Afu-lig (*A. fulgidus* DNA ligase). **(L)** Crystal structure of human DNA ligase I bound with DNA. **(M)** Crystal structure of T4-lig (T4 DNA ligase) bound with DNA. The structures of *S. aureus*, human and T4 DNA ligases were used for a comparison. **(N)** The Sso-lig-DNA-PCNA complex.

diseases (Barany, 1991; Duckworth et al., 2023). The ATP-dependent bacteriophage T4 DNA ligase is one of the most commercial DNA ligases in molecular cloning since it has a strong ability to join cohesive fragments (Shi et al., 2018). Additionally, the predominant shortcoming of this ligase is that it is irreversibly inactivated when exposed to temperatures above 65°C, which confines its application in biotechnology requiring DNA ligation to low temperature. Fortunately, thermostable DNA ligases from hyperthermophiles that can ligate DNA at the elevated temperature and possess a strong thermostabililty have been characterized, which are well suited for biotechnological applications in high-temperature environments.

Hyperthermophiles thrive in high-temperature environments at a temperature range of 80°C to 100°C (Verma et al., 2022). At least one DNA ligase is encoded in the genome of each hyperthermophile. Sequence alignments show that thermostable DNA ligases from archaea and bacteria display low similarities, but they possess six conserved motifs present in all the reported DNA ligases. The phylogenetic relationship of the thermostable DNA ligases from bacteria and archaea is shown in the phylogenetic tree constructed from the DNA ligases from bacteria and archaea, human and the bacteriophage T4 (Figure 1C), demonstrating that the human and T4 DNA ligases display a closer phylogenetic relationship to thermostable archaeal DNA ligases than bacterial DNA ligases. Compared to non-thermostable DNA ligases, thermostable DNA ligases possess a higher fidelity, which can avoid errors for ligation-based molecular diagnostic techniques. Besides, thermostable DNA ligases exhibit a strong thermostability, retaining activity after multiple thermal cycles. Additionally, some thermostable DNA ligases, such as HiFi Taq DNA ligase, appear to exhibit more effective discrimination between correct and mismatched base pairs at either side of nicks. Currently, *Pyrococcus furiosus* DNA ligase, *Thermococcus* 9°N DNA ligase, and Ampligase are commercially available and active at the elevated temperature.

In this review, we summarize current progress on structural and biochemical characteristics of thermostable DNA ligases from hyperthermophilic bacteria and archaea, focusing on similarities and differences between thermostable DNA ligases and non-thermostable homologs, and between thermostable bacterial and archaeal DNA ligases. Additionally, altered thermostable DNA ligases with improved properties and their potential applications in biotechnology are discussed.

Thermostable DNA ligases from hyperthermophilic bacteria

Currently, eight DNA ligases from hyperthermophilic bacteria have been biochemically or structurally characterized from *Thermus thermophilus* HB8 (Takahashi et al., 1984), *Thermus* sp. AK16D (Tong et al., 1999), *Thermus aquaticus* (Housby et al., 2000), *Rhodothermus marinus* (Housby et al., 2000), *Thermus scotoductus* (Housby et al., 2000), *Thermus filiformi* (Lee et al., 2000), *Aquifex pyrophilus* (Lim et al., 2001), and *Thermotoga maritima* (Le et al., 2010). The common characteristic of these thermostable bacterial DNA ligases is that they utilize NAD$^+$ rather than ATP as a cofactor (Table 1).

Besides, these four DNA ligases from *T. thermophilus* HB8, *Thermus* sp. AK16D, *A. pyrophilus*, and *T. maritima* display maximum activity in the presence of Mg^{2+} or Mn^{2+}. In addition to Mg^{2+} or Mn^{2+}, Ca^{2+} can also stimulate the ligation activities of the *Thermus* sp. AK16D and *T. maritima* DNA ligases. The optimal temperatures of these thermostable bacterial DNA ligases are at a temperature range of 55°C ~ 72°C, demonstrating that they are thermophilic. Additionally, the DNA ligases from *A. pyrophilus* and *T. maritima* retain their ligation activity after heating at 95°C for 30 min and 60 min, respectively, thus showing that they are thermostable DNA ligases. However, the *R. marinus* and *T. scotoductus* DNA ligases lose their activity after heating at 91°C for 7 min and 26 min, respectively. Besides, the *T. maritima* and *Thermus* sp. AK16D DNA ligases exhibit the highest ligation activity at pH 8.0.

The *T. thermophilus* DNA ligase is one of the well-characterized DNA ligases from hyperthermophilic bacteria (Takahashi et al., 1984; Luo et al., 1996; Luo and Barany, 1996; Muerhoff et al., 2004; Wang et al., 2020). Compared to the mesophilic ATP-dependent DNA ligases, the *T. thermophilus* DNA ligase possesses three distinct characteristics. Firstly, the optimal temperature of the *T. thermophilus* DNA ligase was estimated to be approximately 65°C instead of 37°C. Secondly, the enzyme possesses strong thermostability. Thirdly, the enzyme displays a higher fidelity than the T4 DNA ligase.

The crystal structure of the NAD$^+$-dependent *T. filiformi* DNA ligase was solved (Figure 1D), displaying a highly modular architecture that comprises four domains including the AdD (domain 1), the OBD (domain 2), the domain 3 harboring a zinc finger and a HhH motif, and BRCT (domain 4). In contrast to the truncated *S. aureus* DNA ligase lacking OBD (Figure 1E; Han et al., 2009), the *T. filiformi* DNA ligase harbors the OBD, which is involved in DNA binding DNA. Besides to the OBD (domain 2), the zinc finger motif (subdomain 3a) and the HhH motif domain (subdomain 3b), which are all known to bind to nucleic acids (Trasviña-Arenas et al., 2021; Hajdu et al., 2023), are present in the *T. filiformi* DNA ligase, demonstrating a unique organization and spatial arrangement of these domains. Additionally, the *T. filiformi* DNA ligase has a BRCT domain with much mobility in the open conformation, which might act as a gate to regulate DNA binding and release. Overall, the structure of the *T. filiformi* DNA ligase provides insight into the understanding of domain organization, catalytic mechanism and evolution of covalent nucleotidyl transferases, which might be generally applicable to eukaryotic DNA ligases.

Mutational analyzes show that only the AdD of the *T. filiformis* DNA ligase is responsible for forming the ligase-AMP complex (Jeon et al., 2004). The OBD, the zinc finger and HhH motif in the domain 3 of the *T. filiformis* DNA ligase are essential for forming the ligase-DNA complex. However, the BRCT domain of the *T. filiformis* DNA ligase is not essential for the enzyme activity. In contrast, substantial loss of ligation activity was observed for the mutant with the deletion of the BRCT domain of the *Thermus* sp. AK16D DNA ligase (Feng et al., 2004). However, the mutant can still form a ligase-NMP intermediate rather than NMP-DNA accumulation, suggesting that the BRCT domain is involved in the substrate adenylation step. Further mutational analysis shows that the mutation of G617 in the *Thermus* sp. AK16D DNA ligase to Ile resulted in a low ligation activity and caused this enzyme to be weakly adenylated (Feng et al., 2004), suggesting that residue G617 in the BRCT domain plays a role in the enzyme adenylation step.

TABLE 1 Comparison of biochemical characteristics of DNA ligases from hyperthermophilic bacteria and archaea.

Hyperthermophile		Optimal T[a] (°C)	Optimal pH	Divalent ion requirement	Cofactor	PDB ID	Reference
Bacteria	Thermus thermophilus HB8	65~72	N.D.[b]	Mg^{2+} and Mn^{2+}	NAD^+	N.D.	Takahashi et al. (1984)
	Thermus sp. AK16D	N.D.	N.D.	Mg^{2+}, Mn^{2+} and Ca^{2+}	NAD^+	N.D.	Tong et al. (1999)
	Thermus aquaticus	N.D.	N.D.	N.D.	N.D.	N.D.	Housby et al. (2000)
	Rhodothermus marinus	55	N.D.	N.D.	NAD^+	N.D.	Housby et al. (2000)
	Thermus scotoductus	65	N.D.	N.D.	NAD^+	N.D.	Housby et al. (2000)
	Thermus filiformis	N.D.	N.D.	N.D.	N.D.	1DGS	Lee et al. (2000)
	Aquifex pyrophilus	65	8.0~8.6	Mg^{2+} and Mn^{2+}	NAD^+	N.D.	Lim et al. (2001)
	Thermotoga maritima	60	8.0	Mg^{2+}, Mn^{2+} and Ca^{2+}	NAD^+	N.D.	Le et al. (2010)
Archaea	Methanobacterium thermoautotrophicum	70	N.D.	N.D.	N.D.	N.D.	Sriskanda et al. (2000)
	Thermococcus kodakaraensis	100	8.0	Mg^{2+}, Mn^{2+}, Sr^{2+} and Ca^{2+}	ATP and NAD^+	N.D.	Nakatani et al. (2000)
	Sulfolobus shibatae	60~80	N.D.	Mn^{2+}, Mg^{2+} and Ca^{2+}	ATP and dATP	N.D.	Lai et al. (2002)
	Aeropyrum pernix	70	7.5	Mg^{2+}, Mn^{2+}, Ca^{2+} and Co^{2+}	ADP and ATP	N.D.	Jeon and Ishikawa (2003)
	Thermococcus fumicolans	65	7.0	Mg^{2+}	ATP and NAD^+	N.D.	Rolland et al. (2004)
	Pyrococcus horikoshii	70~90	N.D.	N.D.	ATP	N.D.	Keppetipola and Shuman, 2005
	Pyrococcus furiosus	N.D.	N.D.	N.D.	N.D.	3GDE	Nishida et al. (2006)
	Thermococcus sp. NA1	80	7.5	Mg^{2+} and Zn^{2+}	ATP and NAD^+	N.D.	Kim et al. (2006)
	Sulfolobus solfataricus	N.D.	N.D.	N.D.	N.D.	2HIX, 7RPX	Pascal et al. (2006) and Sverzhinsky et al. (2022)
	Staphylothermus marinus	N.D.	N.D.	Mg^{2+} and Mn^{2+}	N.D.	N.D.	Seo et al. (2007)
	Sulfophobococcus zilligii	N.D.	N.D.	N.D.	ADP, GTP and ATP	D2CJS7[c]	Sun et al. (2008)
	Archaeoglobus fulgidus	N.D.	N.D.	N.D.	N.D.	2CFM	Kim et al. (2009)
	Thermococcus sibiricus	N.D.	N.D.	N.D.	N.D.	4EQ5	Petrova T. E. et al. (2012)
	Thermococcus sp. 1,519	70	7.0~10.5	Mg^{2+}	ATP	3RR5	Petrova T. et al. (2012)
	Hyperthermus butylicus	75	8.0	Mg^{2+} and Mn^{2+}	ADP, GTP and ATP	N.D.	Kim et al. (2013)
	Pyrococcus abyssi	70	N.D.	N.D.	N.D.	N.D.	Oscorbin et al. (2015)
	Thermococcus barophilus Ch5	65~70	6.0~9.0	Mn^{2+}, Mg^{2+} and Ca^{2+}	ATP and UTP	N.D.	Shi et al. (2019)

[a]T, Temperature.
[b]N.D., Not determined.
[c]The crystal structure of the S. zilligii DNA ligase was solved, but no PDB accession number is given.

Thermostable DNA ligases from hyperthermophilic archaea

Seventeen DNA ligases have been reported from hyperthermophilic archaea to date, including *Methanobacterium thermoautotrophicum* (Sriskanda et al., 2000), *Thermococcus kodakaraensis* (Nakatani et al., 2000), *Sulfolobus shibatae* (Lai et al., 2002), *Aeropyrum pernix* (Jeon and Ishikawa, 2003), *Thermococcus fumicolans* (Rolland et al., 2004), *Pyrococcus horikoshii* (Keppetipola and Shuman, 2005), *P. furiosus* (Nishida et al., 2006), *Thermococcus* sp. NA1 (Kim et al., 2006), *Sulfolobus solfataricus* (Pascal et al., 2006), *Staphylothermus marinus* (Seo et al., 2007), *Sulfophobococcus zilligii* (Sun et al., 2008), *Archaeoglobus fulgidus* (Kim et al., 2009), *Thermococcus sibiricus* (Petrova T. E. et al., 2012), *Thermococcus* sp. 1519 (Petrova T. et al., 2012), *Hyperthermus butylicus* (Kim et al., 2013), *Pyrococcus abyssi* (Oscorbin et al., 2015), and *Thermococcus barophilus* Ch5 (Shi et al., 2019). Comparison of biochemical characteristics of thermostable archaeal and bacterial DNA ligases is summarized in Table 1, highlighting their similar and different properties.

In contrast to DNA ligases from hyperthermophilic bacteria that specifically use NAD$^+$ as a cofactor, DNA ligases from hyperthermophilic archaea preferably utilize ATP as a cofactor. In addition to ATP, the DNA ligases from *T. kodakaraensis*, *T. fumicolans*, and *Thermococcus* sp. NA1 can also use NAD$^+$ for their ligation activity. Interestingly, ADP can be used as a cofactor for DNA ligases from *H. butylicus*, *S. zilligii*, and *A. pernix* to join DNA. Besides ATP and NAD$^+$, GTP is also a cofactor for the DNA ligases from *H. butylicus* and *S. zilligii*. Additionally, dATP and UTP can be used by the DNA ligase from *S. shibatae*. Interestingly, no extra ATP is needed for the recombinant *T. barophilus* Ch5 DNA ligase to ligate DNA since the enzyme is already adenylated after purification, which can reduce the cost of its application in biotechnology.

As observed for DNA ligases from hyperthermophilic bacteria, DNA ligases from hyperthermophilic archaea display maximum activity at the elevated temperature (65°C ∼ 100°C). Besides, the DNA ligases from hyperthermophilic *S. marinus*, *A. pernix*, and *T. barophilus* Ch5 retain ligation activity after heating 100°C for at least 1 h, suggesting that their thermostablilty is higher than that of DNA ligases from hyperthermophilic bacteria. In contrast, the DNA ligases from *P. abyssi* and *Thermococcus* sp. 1519 display a lower thermostability than that of the DNA ligases from *T. maritima* and *A. pyrophilus*. Moreover, the *T. kodakaraensis* DNA ligase displays a nick-sealing activity at 100°C, demonstrating that it possesses the highest thermophilicity.

The crystal structures of six thermostable DNA ligases from hyperthermophilic archaea have been solved to date, comprising three domains: AdD, OBD, and the N-terminal DNA-binding domain (DBD) (Figure 1B), which includes *S. solfataricus* (Figure 1F; Pascal et al., 2006), *S. zilligii* (Figure 1G; Supangat et al., 2010), *P. furiosus* (Figure 1H; Nishida et al., 2006), *Thermococcus* sp. 1519 (Figure 1I; Petrova T. et al., 2012), *T. sibiricus* (Figure 1J; Petrova T. E. et al., 2012), and *A. fulgidus* (Figure 1K; Kim et al., 2009). In contrast to the *T. filiformi* DNA ligase structure, thermostable DNA ligases from these hyperthermophilic archaea lack the Zn-finger motif, HhH motif, and BRCT domain, but possess the DBD. Interestingly, the N-terminal DBD is absent in bacterial DNA ligase, but only present in the eukaryotic and archaeal DNA ligases, which might participate in distorting the DNA substrate and maintaining an active conformation

of the catalytic core. Similar to the *T. filiformi* DNA ligase structure, these archaeal DNA ligases harbor the AdD and OBD, but their conformation varies. Although the structures of thermostable DNA ligases resemble the DNA-bound structures of the human DNA ligases I (Figure 1L) and T4 DNA ligase (Figure 1M; Pascal et al., 2004; Shi et al., 2018), which are distinct from the structures of bacterial DNA ligases, their domain arrangements differ substantially. The conformational flexibility might be critical for ligating DNA of thermostable DNA ligases.

The open and extended conformations of AdD and OBD were captured in the *S. solfataricus* DNA ligase, where its OBD was turned away from the AdD (Figure 1F). In contrast, an intermediate conformation is found in the *Thermococcus* sp.1519 ligase structure (Figure 1I), where its OBD was rotated anticlockwise around the AdD by ∼ 90º, which sharply contrasts with the open extended conformation in the *S. solfataricus* DNA ligase. Besides, the structures of the DNA ligases from *P. furiosus* (Figure 1H), *T. sibiricus* and *A. fulgidus* adopted a closed conformation that was yielded with a further 120º rotation of the OBD (Figures 1J,K). Interestingly, their closed conformation might be stabilized with a C-terminal helix in these structures via regulating several ionic interactions between the AdD and the OBD in *P. furiosus* DNA ligase (Figure 1H).

S. solfataricus possesses three proliferating cell nuclear antigen (PCNA 1, 2 and 3) proteins, which sharply contrasts with *P. furiosus* harboring the homotrimeric PCNA. Each subunit of the *S. solfataricus* PCNA interacts with a specific enzyme, among which the PCNA3 client enzyme is the ligase (Dionne et al., 2003). Biochemical data show that the *S. solfataricus* DNA ligase activity was stimulated by the heterotrimeric *S. solfataricus* PCNA by enhancing ligase encirclement at nicked DNA (Dionne et al., 2003; Pascal et al., 2006), which was confirmed in a negative stain electron microscopy structure of the *P. furiosus* DNA ligase in complex with a nicked, non-ligatable DNA and homotrimeric PCNA by single-particle analysis (Mayanagi et al., 2009). Recently, Sverzhinsky et al. (2022) solved the cryo-EM structure of the complex of the *S. solfataricus* DNA ligase, DNA, and the heterotrimeric PCNA (Figure 1N), demonstrating that the canonical PCNA-interacting peptide (PIP) motif in the DBD of the ligases interacts with the inter-domain connecting loop of PCNA3, which provide support for this stimulation of the ligase activity by the heterotrimeric PCNA.

Thermostable DNA ligases: ligase chain rection for detection of single nucleotide polymorphism

Like PCR, ligase chain rection (LCR) mediated by thermostable DNA ligase can be used to amplify DNA to detect a single base substitution since this ligase can specifically seal two adjacent oligonucleotides complementary to a DNA template strand containing a perfect base pair rather than a single base mismatch (Gibriel and Adel, 2017; Liang et al., 2022). LCR is usually performed by a thermal cycling of denaturing the target dsDNA at 95°C, annealing the probes with the corresponding denatured template at 55°C, and ligating the nicks created between the two probes at 66°C by thermostable DNA ligase (Figure 2A). If the single base mismatch is present in the target dsDNA, no ligation product would be observed. Thus, LCR is a powerful method for the detection of single base mutation in single nucleotide polymorphism (SNP).
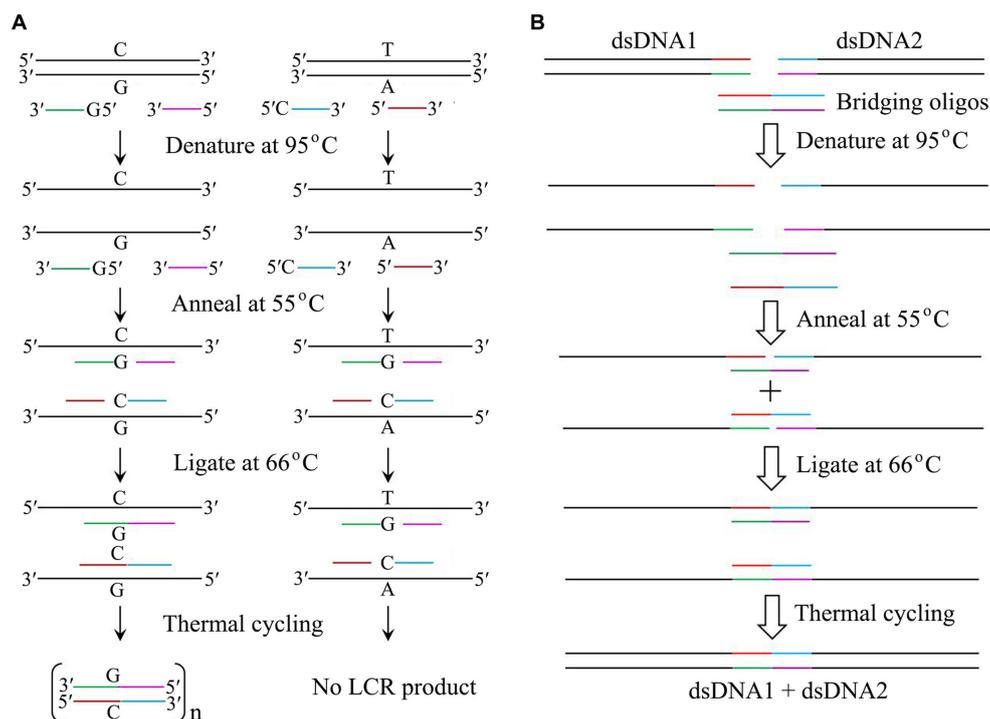
FIGURE 2
Application of thermostable DNA ligases from hyperthermophiles. **(A)** Mutation detection by LCR catalyzed by thermostable DNA ligases. Note that both probes of 5′-G-oligo and 5′-C-oligo are phosphorylated. **(B)** DNA assembly by LCR catalyzed by thermostable DNA ligases. The details are in text.

Thermostable DNA ligase is a cornerstone for LCR due to its high fidelity and strong thermostability, thus allowing the useful DNA-based diagnostic tests for inherited diseases to be performed in clinical laboratories (Liu et al., 2020; Malik et al., 2021; Zhang et al., 2021; Li et al., 2022). The thermostable DNA ligases from *T. thermophilus* and *T. aquaticus* display the higher ligation fidelity than T4 DNA ligase (Barany, 1991; Luo et al., 1996), and have been commercially used for LCR to date. Currently, LCR mediated by thermostable DNA ligase has been employed to detect point mutations associated with genetic diseases.

Besides LCR, ligase detection reaction (LDR), ligation amplification reaction (LAR), and gap-LCR have also been developed with thermostable DNA ligases (Wu and Wallace, 1989; Abravaya et al., 1995; Li et al., 2022). Additionally, thermostable DNA ligases can also be used for rolling-circle amplification (RCA) in the presence of padlock probes, which is an alternative method for detecting SNP (Yan et al., 2023). Modification of LCR, gap-LCR was developed by filling in a gap between annealed probes by DNA polymerase to reduce the background generated by target-independent, blunt-end ligation. To date, it has been confirmed that gap-LCR is useful for detection of a mutation in the reverse transcriptase gene of HIV that confers AZT resistance (Wu and Wallace, 1989; Abravaya et al., 1995).

Thermostable DNA ligases: LCR for DNA assembly

Since establishing a highly efficient DNA assembly method is significantly essential for automated high-throughput DNA assembly,

De Kok et al. (2014) reported a one-step, scarless DNA assembly via LCR involving a thermostable ligase by using single-stranded bridging oligos complementary to the ends of target DNA. DNA assembly is usually completed by a thermal cycling of denaturing two substrates and two bridging oligos at 95°C, annealing the bridging oligo with the target dsDNA at 55°C, and ligating the nicks created between the bridging oligo and the target dsDNA at 66°C by thermostable DNA ligase (Figure 2B). Compared to Gibson isothermal assembly (Gibson et al., 2009), DNA assembly mediated by LCR is a rapid and reliable method with a cheap, rapid, and convenient workflow. Therefore, LCR has become a powerful method for both manual and automated high-throughput DNA assembly.

Recently, an effective DNA assembly method was reported using thermostable exonuclease and ligase "(DATEL)" including *T. aquaticus* DNA ligase, thermal exonucleases (*T. aquaticus* and *P. furiosus* DNA polymerases; Kang et al., 2018). Due to rapid assembly (2~10 DNA fragments per 1~2 h), high accuracy (between 74 and 100%) and the simple operation system, DATEL displays a great potential for high-throughput assembly of DNA fragments, which will greatly promote the rapid development of metabolic engineering and synthetic biology.

Thermostable DNA ligases: directed evolution

Directed evolution is an important method to engineer enzyme mutants. The thermostable DNA ligase from *T. maritima* can be used to mediate PCR production of circular plasmid (PPCP) that is catalyzed by the *T. aquaticus* DNA polymerase (Le et al., 2013). This

PPCP method has an important application in directed evolution by allowing one-step construction of mutagenesis libraries through *in situ* error-prone PCR. For example, random mutagenesis libraries of a xylanase gene and two cellulase genes have been created by this PPCP method mediated by the *T. maritima* DNA ligase. Therefore, *in situ* error-prone PPCP mediated by thermostable DNA ligase is useful for generating random mutagenesis libraries for directed evolution.

Altered thermostable DNA ligases with the improved properties

Protein engineering is a powerful method for constructing enzyme mutants with the enhanced catalytic properties. The *P. furiosus* DNA ligase is the first reported thermostable DNA ligase that has been altered (Nishida et al., 2006). Besides, DNA ligases from *Thermus* sp. AK16D, *Thermococcus* sp. 1,519 and *T. thermophilus* were also targeted for engineering mutants with better properties (Luo et al., 1996; Feng, 2007; Modarres et al., 2015), such as increased fidelity and thermostability.

Compared to the wild-type protein, the D540R mutant of the *P. furiosus* DNA ligase displayed the nick-joining activity at a broaden temperature range (20°C~80°C) (Tanabe et al., 2014). Besides, the biophysics experiments confirmed that the D540R mutant exhibits the increased binding of the nicked DNA substrate and formation of the covalent ligase-AMP intermediate (Tanabe et al., 2012). Additionally, a further increased nick-sealing activity was observed for the engineered *P. furiosus* DNA ligase mutants harboring the D540R mutation plus K554A/K558A mutations or the D540R mutation and a deletion of the final four residues of the C-terminal helix.

Using molecular modeling and simulations, Modarres et al. (2015) predicted the substitutions that can increase thermostability of the DNA ligase from *Thermococcus* sp. 1,519 and identified its thermosensitive regions. By optimizing the charged groups on the surface in the thermosensitive regions by introducing the selected mutations (A287K, G304D, S364I, and A387K), the remarkable and additive increase of thermostability of the enzyme was observed relative to the wild-type protein. Thus, the altered *Thermococcus* sp. 1519 DNA ligase with increased thermostability might provide potential DNA ligases used for biotechnology.

LCR and LDR are the powerful methods for the detection of SNPs, which require a thermostable DNA ligase with high fidelity. Notably, the mutations at residue K294 in the *T. thermophilus* DNA ligase to be Arg and Pro increased 4-fold fidelity and 11-fold fidelity in addition to still retaining the nick-sealing activities (Luo et al., 1996), respectively. Additionally, the mutations of D286E/G287A/V289I/K291R resulted in the enhanced ligation fidelity of the *Thermus* sp. AK16D DNA ligase (Feng, 2007). Thus, these altered DNA ligase mutants with high-fidelity are potential targets for LCR and LDR.

Conclusion and future directions

Thermostable DNA ligases from hyperthermophiles display structural and biochemical properties distinct from those of non-thermostable orthologs. Several DNA ligase mutants from *P. furiosus*, *Thermococcus* sp. 1519, and *T. thermophilus* with improved properties have been engineered to date. However, the thermostable DNA ligase mutants with superior properties need to be engineered by directed evolution to meet the demands of biotechnological development.

Currently, no crystal structure of the complex of thermostable DNA ligase with DNA has been solved. It will be worthwhile to solve the structures of the thermostable DNA ligase complexes with DNA, which might provide insight into understanding structural and functional similarities and differences between thermostable DNA ligases and non-thermostable orthologs. Additionally, the relationship between protein dynamics and catalysis of thermostable DNA ligases at the elevated temperature needs to be investigated.

Author contributions

JS made Table 1 and wrote this manuscript; PO wrote and revised this manuscript; PC made Figure 2; LZ designed, wrote, and revised this manuscript. All authors approved it for publication.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abravaya, K., Carrino, J. J., Muldoon, S., and Lee, H. H. (1995). Detection of point mutations with a modified ligase chain reaction (gap-LCR). *Nucleic Acids Res.* 23, 675–682. doi: 10.1093/nar/23.4.675

Barany, F. (1991). Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc. Natl. Acad. Sci. U. S. A.* 88, 189–193. doi: 10.1073/pnas.88.1.189

De Kok, S., Stanton, L. H., Slaby, T., Durot, M., Holmes, V. F., Patel, K. G., et al. (2014). Rapid and reliable DNA assembly via ligase cycling reaction. *ACS Synth. Biol.* 3, 97–106. doi: 10.1021/sb4001992

Dionne, I., Nookala, R. K., Jackson, S. P., Doherty, A. J., and Bell, S. D. (2003). A heterotrimeric PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol. Cell* 11, 275–282. doi: 10.1016/s1097-2765(02)00824-9

Duckworth, A. T., Bilotti, K., Potapov, V., and Lohman, G. J. S. (2023). Profiling DNA ligase substrate specificity with a pacific biosciences single-molecule real-time sequencing assay. *Curr. Protoc.* 3:e690. doi: 10.1002/cpz1.690

Feng, H. (2007). Mutational analysis of bacterial NAD+-dependent DNA ligase: role of motif IV in ligation catalysis. *Acta Biochim. Biophys. Sin. Shanghai* 39, 608–616. doi: 10.1111/j.1745-7270.2007.00313.x

Feng, H., Parker, J. M., Lu, J., and Cao, W. (2004). Effects of deletion and site-directed mutations on ligation steps of NAD+-dependent DNA ligase: a biochemical analysis of BRCA1 C-terminal domain. *Biochemistry* 43, 12648–12659. doi: 10.1021/bi049451c

Gibriel, A. A., and Adel, O. (2017). Advances in ligase chain reaction and ligation-based amplifications for genotyping assays: detection and applications. *Mutat. Res. Rev. Mutat. Res.* 773, 66–90. doi: 10.1016/j.mrrev.2017.05.001

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318

Hajdu, B., Hunyadi-Gulyás, É., Kato, K., Kawaguchi, A., Nagata, K., and Gyurcsik, B. (2023). Zinc binding of a Cys2His2-type zinc finger protein is enhanced by the interaction with DNA. *J. Biol. Inorg. Chem.* 28, 301–315. doi: 10.1007/s00775-023-01988-1

Han, S., Chang, J. S., and Griffor, M. (2009). Structure of the adenylation domain of NAD(+)-dependent DNA ligase from *Staphylococcus aureus. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 65, 1078–1082. doi: 10.1107/S1744309109036872

Housby, J. N., Thorbjarnardóttir, S. H., Jónsson, Z. O., and Southern, E. M. (2000). Optimised ligation of oligonucleotides by thermal ligases: comparison of *Thermus scotoductus* and *Rhodothermus marinus* DNA ligases to other thermophilic ligases. *Nucleic Acids Res.* 28, 10e–110e. doi: 10.1093/nar/28.3.e10

Jeon, S. J., and Ishikawa, K. (2003). A novel ADP-dependent DNA ligase from *Aeropyrum pernix* K1. *FEBS Lett.* 550, 69–73. doi: 10.1016/s0014-5793(03)00821-4

Jeon, H. J., Shin, H. J., Choi, J. J., Hoe, H. S., Kim, H. K., Suh, S. W., et al. (2004). Mutational analyses of the thermostable NAD+-dependent DNA ligase from *Thermus filiformis. FEMS Microbiol. Lett.* 237, 111–118. doi: 10.1016/j.femsle.2004.06.018

Kang, Z., Ding, W., Jin, P., Du, G., and Chen, J. (2018). DNA assembly with the DATEL method. *Methods Mol. Biol.* 1772, 421–428. doi: 10.1007/978-1-4939-7795-6_24

Keppetipola, N., and Shuman, S. (2005). Characterization of a thermophilic ATP-dependent DNA ligase from the euryarchaeon *Pyrococcus horikoshii. J. Bacteriol.* 187, 6902–6908. doi: 10.1128/JB.187.20.6902-6908.2005

Kim, D. J., Kim, O., Kim, H. W., Kim, H. S., Lee, S. J., and Suh, S. W. (2009). ATP-dependent DNA ligase from *Archaeoglobus fulgidus* displays a tightly closed conformation. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 65, 544–550. doi: 10.1107/S1744309109017485

Kim, Y. J., Lee, H. S., Bae, S. S., Jeon, J. H., Yang, S. H., Lim, J. K., et al. (2006). Cloning, expression, and characterization of a DNA ligase from a hyperthermophilic archaeon *Thermococcus* sp. *Biotechnol. Lett.* 28, 401–407. doi: 10.1007/s10529-005-6070-6

Kim, J. H., Lee, K. K., Sun, Y., Seo, G. J., Cho, S. S., Kwon, S. H., et al. (2013). Broad nucleotide cofactor specificity of DNA ligase from the hyperthermophilic crenarchaeon *Hyperthermus butylicus* and its evolutionary significance. *Extremophiles* 17, 515–522. doi: 10.1007/s00792-013-0536-6

Lai, X., Shao, H., Hao, F., and Huang, L. (2002). Biochemical characterization of an ATP-dependent DNA ligase from the hyperthermophilic crenarchaeon *Sulfolobus shibatae. Extremophiles* 6, 469–477. doi: 10.1007/s00792-002-0284-5

Le, Y., Chen, H., Zagursky, R., Wu, J. H., and Shao, W. (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. doi: 10.1093/dnares/dst016

Le, Y., Peng, J., Pei, J., Li, H., Duan, Z., and Shao, W. (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. doi: 10.1016/j.enzmictec.2009.08.015

Lee, J. Y., Chang, C., Song, H. K., Moon, J., Yang, J. K., Kim, H. K., et al. (2000). Crystal structure of NAD(+)-dependent DNA ligase: modular architecture and functional implications. *EMBO J.* 19, 1119–1129. doi: 10.1093/emboj/19.5.1119

Li, Y., Wang, X., Wang, M., Liu, M., Wang, H., Xia, W., et al. (2022). Advances in ligase-based nucleic acid amplification technology for detecting gene mutations: a review. *Mol. Cell. Biochem.* 1-11, 1–11. doi: 10.1007/s11010-022-04615-w

Liang, H., Chen, Z., and Fang, G. (2022). A depth-first search algorithm for oligonucleotide design in gene assembly. *Front. Genet.* 13:1023092. doi: 10.3389/fgene.2022.1023092

Lim, J. H., Choi, J., Han, S. J., Kim, S. H., Hwang, H. Z., Jin, D. K., et al. (2001). Molecular cloning and characterization of thermostable DNA ligase from *Aquifex pyrophilus*, a hyperthermophilic bacterium. *Extremophiles* 5, 161–168. doi: 10.1007/s007920100187

Liu, Z. J., Yang, L. Y., Wei, Q. X., Ye, C. L., Xu, X. W., Zhong, G. X., et al. (2020). A novel ligase chain reaction-based electrochemical biosensing strategy for highly sensitive point mutation detection from human whole blood. *Talanta* 216:120966. doi: 10.1016/j.talanta.2020.120966

Luo, J., and Barany, F. (1996). Identification of essential residues in *Thermus thermophilus* DNA ligase. *Nucleic Acids Res.* 24, 3079–3085. doi: 10.1093/nar/24.15.3079

Luo, J., Bergstrom, D. E., and Barany, F. (1996). Improving the fidelity of *Thermus thermophilus* DNA ligase. *Nucleic Acids Res.* 24, 3071–3078. doi: 10.1093/nar/24.15.3071

Malik, A., Aldakheel, F., Rabbani, S., Alshehri, M., Chaudhary, A. A., Alkholief, M., et al. (2021). LCR based quick detection of hotspot G1896A mutation in patients with different spectrum of hepatitis B. *J. Infect. Public Health* 14, 651–654. doi: 10.1016/j.jiph.2021.01.013

Mayanagi, K., Kiyonari, S., Saito, M., Shirai, T., Ishino, Y., and Morikawa, K. (2009). Mechanism of replication machinery assembly as revealed by the DNA ligase-PCNA-DNA complex architecture. *Proc. Natl. Acad. Sci. U. S. A.* 106, 4647–4652. doi: 10.1073/pnas.0811196106

Modarres, H. P., Dorokhov, B. D., Popov, V. O., Ravin, N. V., Skryabin, K. G., and Peraro, M. D. (2015). Understanding and engineering thermostability in DNA ligase from *Thermococcus* sp. 1519. *Biochemistry* 54, 3076–3085. doi: 10.1021/bi501227b

Muerhoff, A. S., Dawson, G. J., and Desai, S. M. (2004). A non-isotopic method for the determination of activity of the thermostable NAD-dependent DNA ligase from *Thermus thermophilus* HB8. *J. Virol. Methods* 119, 171–176. doi: 10.1016/j.jviromet.2004.04.003

Nakatani, M., Ezaki, S., Atomi, H., and Imanaka, T. (2000). A DNA ligase from a hyperthermophilic archaeon with unique cofactor specificity. *J. Bacteriol.* 182, 6424–6433. doi: 10.1128/JB.182.22

Nishida, H., Kiyonari, S., Ishino, Y., and Morikawa, K. (2006). The closed structure of an archaeal DNA ligase from *Pyrococcus furiosus. J. Mol. Biol.* 360, 956–967. doi: 10.1016/j.jmb.2006.05.062

Oscorbin, I. P., Boyarskikh, U. A., Zakabunin, A. I., Khrapov, E. A., and Filipenko, M. L. (2015). DNA-binding domain of DNA ligase from the thermophilic archaeon *Pyrococcus abyssi*: improving long-range PCR and neutralization of heparin's inhibitory effect. *Appl. Biochem. Biotechnol.* 176, 1859–1869. doi: 10.1007/s12010-015-1683-2

Pascal, J. M., O'Brien, P. J., Tomkinson, A. E., and Ellenberger, T. (2004). Human DNA ligase I completely encircles and partially unwinds nicked DNA. *Nature* 432, 473–478. doi: 10.1038/nature03082

Pascal, J. M., Tsodikov, O. V., Hura, G. L., Song, W., Cotner, E. A., Classen, S., et al. (2006). A flexible interface between DNA ligase and PCNA supports conformational switching and efficient ligation of DNA. *Mol. Cell* 24, 279–291. doi: 10.1016/j.molcel.2006.08.015

Petrova, T., Bezsudnova, E. Y., Boyko, K. M., Mardanov, A. V., Polyakov, K. M., Volkov, V. V., et al. (2012). ATP-dependent DNA ligase from *Thermococcus* sp. 1519 displays a new arrangement of the OB-fold domain. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 68, 1440–1447. doi: 10.1107/S1744309112043394

Petrova, T. E., Bezsudnova, E. Y., Dorokhov, B. D., Slutskaya, E. S., Polyakov, K. M., Dorovatovskiy, P. V., et al. (2012). Expression, purification, crystallization and preliminary crystallographic analysis of a thermostable DNA ligase from the archaeon *Thermococcus sibiricus. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 68, 163–165. doi: 10.1107/S1744309111050913

Rolland, J. L., Gueguen, Y., Persillon, C., Masson, J. M., and Dietrich, J. (2004). Characterization of a thermophilic DNA ligase from the archaeon *Thermococcus fumicolans. FEMS Microbiol. Lett.* 236, 267–273. doi: 10.1016/j.femsle.2004.05.045

Seo, M. S., Kim, Y. J., Choi, J. J., Lee, M. S., Kim, J. H., Lee, J. H., et al. (2007). Cloning and expression of a DNA ligase from the hyperthermophilic archaeon *Staphylothermus marinus* and properties of the enzyme. *J. Biotechnol.* 128, 519–530. doi: 10.1016/j.jbiotec.2006.09.024

Shi, K., Bohl, T. E., Park, J., Zasada, A., Malik, S., Banerjee, S., et al. (2018). T4 DNA ligase structure reveals a prototypical ATP-dependent ligase with a unique mode of sliding clamp interaction. *Nucleic Acids Res.* 46, 10474–10488. doi: 10.1093/nar/gky776

Shi, H., Huang, Y., Gan, Q., Rui, M., Chen, H., Tu, C., et al. (2019). Biochemical characterization of a thermostable DNA ligase from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5. *Appl. Microbiol. Biotechnol.* 103, 3795–3806. doi: 10.1007/s00253-019-09736-9

Sriskanda, V., Kelman, Z., Hurwitz, J., and Shuman, S. (2000). Characterization of an ATP-dependent DNA ligase from the thermophilic archaeon *Methanobacterium thermoautotrophicum. Nucleic Acids Res.* 28, 2221–2228. doi: 10.1093/nar/28.11.2221

Sun, Y., Seo, M. S., Kim, J. H., Kim, Y. J., Kim, G. A., Lee, J. I., et al. (2008). Novel DNA ligase with broad nucleotide cofactor specificity from the hyperthermophilic crenarchaeon *Sulfophobococcus zilligii*: influence of ancestral DNA ligase on cofactor utilization. *Environ. Microbiol.* 10, 3212–3224. doi: 10.1111/j.1462-2920.2008.01710.x

Supangat, S., An, Y. J., Sun, Y., Kwon, S. T., and Cha, S. S. (2010). Purification, crystallization and preliminary crystallographic analysis of a multiple cofactor-dependent DNA ligase from *Sulfophobococcus zilligii. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 66, 1583–1585. doi: 10.1107/S1744309110034135

Sverzhinsky, A., Tomkinson, A. E., and Pascal, J. M. (2022). Cryo-EM structures and biochemical insights into heterotrimeric PCNA regulation of DNA ligase. *Structure* 30, 371–385.e5. doi: 10.1016/j.str.2021.11.002

Takahashi, M., Yamaguchi, E., and Uchida, T. (1984). Thermophilic DNA ligase. purification and properties of the enzyme from *Thermus thermophilus* HB8. *J. Biol. Chem.* 259, 10041–10047. doi: 10.1016/S0021-9258(18)90924-5

Tanabe, M., Ishino, S., Ishino, Y., and Nishida, H. (2014). Mutations of Asp540 and the domain-connecting residues synergistically enhance *Pyrococcus furiosus* DNA ligase activity. *FEBS Lett.* 588, 230–235. doi: 10.1016/j.febslet.2013.10.037

Tanabe, M., Ishino, S., Yohda, M., Morikawa, K., Ishino, Y., and Nishida, H. (2012). Structure-based mutational study of an archaeal DNA ligase towards improvement of ligation activity. *Chembiochem* 13, 2575–2582. doi: 10.1002/cbic.201200336

Tong, J., Cao, W., and Barany, F. (1999). Biochemical properties of a high fidelity DNA ligase from *Thermus* species AK16D. *Nucleic Acids Res.* 27, 788–794. doi: 10.1093/nar/27.3.788

Trasviña-Arenas, C. H., Demir, M., Lin, W. J., and David, S. S. (2021). Structure, function and evolution of the Helix-hairpin-Helix DNA glycosylase superfamily: piecing together the evolutionary puzzle of DNA base damage repair mechanisms. *DNA Repair* 108:103231. doi: 10.1016/j.dnarep.2021.103231

Verma, D., Kumar, V., and Satyanarayana, T. (2022). Genomic attributes of thermophilic and hyperthermophilic bacteria and archaea. *World J. Microbiol. Biotechnol.* 38:135. doi: 10.1007/s11274-022-03327-z

Wang, J., Zhu, J., Wang, C., Zhou, G., Yu, X., Fan, H., et al. (2020). *Thermus thermophilus* DNA ligase connects two fragments having exceptionally short complementary termini at high temperatures. *Biochemistry* 59, 400–406. doi: 10.1021/acs.biochem.9b00866

Williamson, A., and Leiros, H. S. (2019). Structural intermediates of a DNA-ligase complex illuminate the role of the catalytic metal ion and mechanism of phosphodiester bond formation. *Nucleic Acids Res.* 47, 7147–7162. doi: 10.1093/nar/gkz596

Williamson, A., and Leiros, H. S. (2020). Structural insight into DNA joining: from conserved mechanisms to diverse scaffolds. *Nucleic Acids Res.* 48, 8225–8242. doi: 10.1093/nar/gkaa307

Wu, D. Y., and Wallace, R. B. (1989). The ligation amplification reaction (LAR)-amplification of specific DNA sequences using sequential rounds of template-dependent ligation. *Genomics* 4, 560–569. doi: 10.1016/0888-7543(89)90280-2

Yan, Y., Chang, D., Xu, Y., Chang, Y., Zhang, Q., Yuan, Q., et al. (2023). Engineering a ligase binding DNA aptamer into a templating DNA scaffold to guide the selective synthesis of circular DNAzymes and DNA aptamers. *J. Am. Chem. Soc.* 145, 2630–2637. doi: 10.1021/jacs.2c12666

Zhang, W., Liu, K., Zhang, P., Cheng, W., Zhang, Y., Li, L., et al. (2021). All-in-one approaches for rapid and highly specific quantification of single nucleotide polymorphisms based on ligase detection reaction using molecular beacons as turn-on probes. *Talanta* 224:121717. doi: 10.1016/j.talanta.2020.121717