Published in final edited form as:

Org Biomol Chem. 2019 February 13; 17(7): 1962–1965. doi:10.1039/c8ob01958d.

Efficiency and fidelity of T3 DNA ligase in ligase-catalysed oligonucleotide polymerisations†

Yi Lei¹, Joshua Washington¹, and Ryan Hili^{1,2}

¹Department of Chemistry, University of Georgia, Athens, GA 30602, USA.

²Department of Chemistry, University of Georgia, Athens, GA 30602, USA and Department of Chemistry and Centre for Research on Biomolecular Interactions, York University, Toronto, ON M3J 1P3, Canada.

Abstract

Ligase-catalyzed oligonucleotide polymerisations (LOOPER) can readily generate libraries of diversely-modified nucleic acid polymers, which can be subjected to iterative rounds of *in vitro* selection to evolve functional activity. While there exist several different DNA ligases, T4 DNA ligase has most often been used for the process. Recently, T3 DNA ligase was shown to be effective in LOOPER; however, little is known about the fidelity and efficiency of this enzyme in LOOPER. In this paper we evaluate the efficiency of T3 DNA ligase and T4 DNA ligase for various codon lengths and compositions within the context of polymerisation fidelity and yield. We find that T3 DNA ligase exhibits high efficiency and fidelity with short codon lengths, but struggles with longer and more complex codon libraries, while T4 DNA ligase exhibits the opposite trend. Interestingly, T3 DNA ligase is unable to accommodate modifications at the 8-position of adenosine when integrated into short codons, which will create challenges in expanding the available codon set for the process. The limitations and strengths of the two ligases are further discussed within the context of LOOPER.

Graphical Abstract

T3 and T4 DNA ligase have contrasting performance in ligase-catalysed oligonucleotide polymerisations

†Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

Conflicts of interest

There are no conflicts to declare.

Introduction

Aptamers, which are single-stranded nucleic acid polymers that bind molecular targets, continue to find broad application as diagnostic sensors and as therapeutics for human diseases.¹ Since their three-dimensional structure is governed by their nucleic acid sequence, aptamers can be readily evolved using *in vitro* selection to enable molecular recognition of their targets.^{2–4} There has been a long history of developing new methods to endow nucleic acid polymers with increased chemical diversity so that they may be evolved as protein-like polymers with improved properties for molecular recognition and catalysis. The conventional approach to include modifications throughout a single-stranded DNA polymer is *via* polymerase-catalysed primer extension using modified dNTPs.^{5–7} While one or two dNTPs are typically reserved to sequence-specifically incorporate chemical modifications during *in vitro* selections,^{8–13} up to three modifications have been successfully used during *in vitro* selections to unlock new molecular function of nucleic acid polymers.¹⁴ More recently, our lab has developed the Ligase-catalysed OligONucleotide PolymERisation (LOOPER) to incorporate greater chemical diversity into nucleic acid polymers (Figure 1).^{15–20} The method relies upon the DNA ligase-catalysed DNA-templated polymerisation of a library of short oligonucleotides, each equipped with a unique functional group. As the method employs codons, rather than single nucleotides, the number of modifications that can be incorporated throughout a nucleic acid polymer is dependent on the codon set size. Using canonical DNA, the codon set size is a maximum of 4^n , where n = codon nucleotide length.

We showed that LOOPER with T4 DNA ligase could be used to evolve a diversely-functionalized aptamer with a 256-membered codon set defined as NNNNT, where N = A, C, G, or T.¹⁷ The corresponding modified A*NNNN anticodon library, where A* denotes modified dA, was divided into 16 sub-libraries, each comprising 16 modified anticodons. This enabled the sequence-defined incorporation of 16 different unique functional groups throughout a nucleic acid polymer in a library format. This method enabled the evolution of a thrombin aptamer with excellent binding affinity ($K_D = 1.6$ nM) and novel folding in just 6 rounds of *in vitro* selection.²⁰ Importantly, the functional groups were found to be essential for binding. More recently, Liu and co-workers demonstrated LOOPER with a trinucleotide codon set using T3 DNA ligase.²¹ The codon set was derived from NNR, where R = G or A, and the anticodon library (YNN, where Y = C or T) was subdivided into eight sub-libraries, each comprising four modified anticodons. This enabled the sequence-defined incorporation of up to eight different modifications throughout a nucleic acid polymer. The system was implemented in the discovery of aptamers against PCSK9 ($K_D = 3$ nM, 15 rounds) and IL-6 ($K_D = 12–22$ nM, 7 rounds).

The two LOOPER systems differ in many respects, which lead to unique benefits for each system: (i) longer codon sets provide access to larger number of unique modifications; (ii) larger codon sets provide greater degeneracy in the encoded modification, *i.e.*, each modification can exist at more sites in sequence space; (iii) shorter codon lengths have fewer codons, and thus fewer corresponding anticodons, and are likely to be higher fidelity; and (iv) shorter codon length have increased density of modifications. However, the most significant difference between the two systems is the use of different DNA ligases. T4 DNA ligase is a workhorse in molecular biology and is amongst the most commonly used ligases,

while T3 DNA ligase has not been as broadly used. Based on phylogeny, T3 DNA ligase is more closely related to T7 DNA ligase and shares critical sequence homology.²² In contrast, T4 DNA ligase diverged as an outgroup much earlier in evolution, and is approximately 30 % larger in size, thus the performance of T3 and T4 DNA ligase in LOOPER are expected to be different. Since only a low-throughput, single DNA template fidelity analysis was performed with T3,²¹ we sought to examine the fidelity of this ligase in a high-throughput manner to compare with T4 DNA ligase. Learning more about the strengths and limitations of T3 ligase may help guide future codon set development for LOOPER.

Results and discussion

Using a duplex DNA sequencing method that we previously developed,¹⁶ which enables the detection of errors arising from anticodon misincorporation during LOOPER, we examined the fidelity of T3 DNA ligase within the context of various codon sets. We first explored the influence on polymerization when increasing the set size of fully degenerate codon sets and their corresponding unmodified anticodons (Table 1). Remarkably, T3 DNA ligase was able to perform LOOPER along 13 repeats of NNN with 97 % fidelity and 60 % yield of full-length product, which corresponds to 96 % coupling efficiency (Table 1, entry 1). Increasing the codon set length to NNNN (11 repeats on template) dropped the fidelity to 89.2 %, but the polymerization was more efficient, resulting in 89 % of full-length product with 99% coupling efficiency (Table 1, entry 3). Interestingly, T4 DNA ligase was not able to generate observable product with either of these codons sets (Table 1, entries 2 and 4). Adjusting pH within enzyme tolerance from 6.9–8.5, Mg^{2+} concentration from 5–15 mM, and ATP concentration from 0.025–1 mM did not result in any observable yield of full-length product using these shorter anticodon libraries (see Figure S1 and S2). These data suggest that T3 DNA ligase is better able to accommodate shorter anticodon fragments during LOOPER. When performing LOOPER along a reading frame comprising eight repeats of the NNNNN codon set, T3 DNA ligase produced 65 % yield of full-length product (95 % coupling efficiency); however, it operated with just 63.8 % fidelity (Table 1, entry 5), which is likely below useful levels for *in vitro* selection, since <3 % of full-length product of an eight-codon reading frame would be without errors. This contrasts with T4 DNA ligase, which performed LOOPER with the same codon set at 81.1 % fidelity and 60 % yield (Table 1, entry 6).

We sought to explore LOOPER reaction conditions to determine if any variables could increase T3 DNA ligase fidelity for NNNNN codon sets. Thus, we adjusted pH within enzyme tolerance²² from 6.9–8.5, Mg^{2+} concentration from 5–15 mM, and ATP concentration from 0.025–1 mM. The variations had only a modest effect on yield (see Figure S3). Decreasing the pH from 8.5 to 6.9 showed the greatest effect by increasing the fidelity nearly 10 % from 57.5 % to 66.7 %. However, this level of fidelity is still considerably lower than T4 DNA ligase. Unfortunately, changing the Mg^{2+} concentration or ATP concentrations from their standard conditions resulted in decreased fidelities for T3 DNA ligase-mediated LOOPER (see Table S2). While T4 DNA ligase was unable to handle shorter anticodon lengths, it does appear to be a higher fidelity ligase during LOOPER of longer, more complex codon set libraries.

We next examined how chemical modifications on anticodons influence LOOPER efficiency and fidelity using T3 DNA ligase. The T3 DNA ligase-catalysed LOOPER system was reported with the Y*NN anticodon set, where the Y* (C or T) carried the chemical modifications; this differs from the T4 DNA ligase system, where modifications are carried on the adenosine group of A*NNNN anticodons. Since the reported²¹ C and T modifications used during T3 DNA-ligase-catalysed LOOPER were chemically synthesized, we decided to evaluate LOOPER with modified nucleobases that were commercially available, as these are more likely to be used by practitioners of LOOPER (Figure 2). Fidelities and efficiencies for various codon sets were surveyed (Table 2). For LOOPER with YNN anticodons, the presence of a modification on Y did not influence fidelity or yield (Table 2, entries 1–2). The YNNN and YNNNN anticodon sets were also observed to be neutral towards modifications of Y (Table 2, entries 5–4 and 7–8). As expected, the longer codons, which have more complex libraries were challenging for T3 ligase. The YNNNN anticodon set, which comprises 512 members, resulted in only 74.3% fidelity when modified – this level of fidelity is unlikely to be practical for *in vitro* selection, since less than 9 % of the modified DNA product will have been correctly translated from the template genotype with a 40 nt reading frame.

Interestingly, T3 DNA ligase was unable to accommodate modifications at the 8-position of adenosine when used in trinucleotide or tetranucleotide anticodons within a library context, resulting in no observable yield of the full-length product (Table 2, entries 3 and 6). Adjustment of LOOPER variables including pH from 6.9–8.5, Mg²⁺ concentration from 5–15 mM, and ATP concentration from 0.025–1 mM failed to result in any detectable yield of full-length product for A*NN. Due to the poor annealing of shorter anticodons we also explored the use of molecular crowding reagent, PEG 6000 at 5 % w/v; however, this also failed to promote LOOPER (Figure S4). Adjustment of LOOPER conditions for the A*NNN system yielded modest increases. At pH 8.5 or in the presence with the inclusion of PEG up to 5% w/v resulted in 20 % yield of full-length product (Figure S5). While this low yield will likely preclude its use in library preparation and *in vitro* selection due to strong codon bias, we decided to sequence the product to determine the innate fidelity of T3 DNA ligase for this type of modification on challenging anticodons. Consistent with the unmodified ANNN anticodon library (Table 2, entry 7), a high fidelity of 97.2% was observed for A*NNN anticodon set at pH 8.5 (Table 2, entry 8).

Since unmodified dA is readily accepted by T3 DNA ligase as a substrate for LOOPER (Table 2, entry 7), the switch from an *anti* to a *syn* conformation about the glycosidic bond that results from a modification at the 8-position of the adenine ring is potentially a contributing factor to loss of efficiency during polymerisation (Figure 3).¹⁶ When adopting the *syn* dA conformer, the adenosine cannot engage in Watson-Crick base pairing with thymidine – this loss of hydrogen bonding may be partially responsible for the marked decreased efficiency of handling trinucleotide and tetranucleotide building blocks that contain a modified dA. The presence of a modification at the 5-position of either dC or dT did not affect the fidelity of the LOOPER, irrespective of the codon length. This is in stark contrast to dA, where a modification at the 8-position drastically improves the fidelity of the system for both T3 and T4 DNA ligases (Table 2, entries 7–8, 11–14). Due to the fact that pyrimidines adopt an *anti* conformation about the glycosidic bond, even if the 5-position is

modified, C/T is engaged in hydrogen bonding to the template in both the modified and unmodified building blocks. This contrasts with dA when it is modified at position 8. Such a modification causes a conformational switch from *anti* to *syn*, resulting in loss of hydrogen bonding and greater reliance on the other four nucleotides for hybridization to the template; fidelity would be expected to rise under such a circumstance. What is especially intriguing is that the modified A*NNNN anticodon has a significantly higher fidelity than the tetranucleotide NNNN, despite having the same codon set size (Table 1, entry 3 vs. Table 2, entry 13). As expected, the fidelity of unmodified ANNNN drops compared with NNNN, since a single mismatch error in the pentanucleotide would result in 20% loss of matching, versus the NNNN, where a single nucleotide mismatch would result in a 25% loss of matching. These data suggest that the adenosine modifications contribute to an increase in fidelity by a more complex mechanism that is currently not fully understood. With the recently solved crystal structure of T4 DNA ligase,²³ insights into the enzyme structure may serve to guide future development of LOOPER anticodon sets.

Conclusions

The T3 DNA ligase is a very promising enzyme for LOOPER-mediated synthesis of modified DNA. Its strengths over T4 DNA ligase at handling shorter codon lengths more efficiently enable higher densities of modifications to be sequence-specifically installed on DNA. However, this occurs at the expense of fewer displayed functional groups (number of unique functional groups) and less sequence degeneracy (sequence space on DNA where a specific R group may be displayed). The fidelity of T4 DNA ligase for longer more complex codon sets, such as NNNNN, is of particular note. The weakness of T3 DNA ligase at handling modified dA in smaller codon lengths is likely a result of *anti*→*syn* conformational switch about the glycosidic bond of dA that is poorly tolerated; this creates challenges for the use of T3 DNA ligase in LOOPER with fully modified NNN or NNNN libraries in its current optimized system. Achieving this would theoretically enable sampling of the complete sequence landscape. To this end, the development of high fidelity LOOPER that is compatible with a fully degenerate codon set (4ⁿ) is forthcoming.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health (R21CA207711) and the Natural Sciences and Engineering Research Council of Canada (NSERC). We would like to thank the Georgia Genomics Facility for DNA sequencing services and the PAMS core facility at the University of Georgia for their help in the characterisation of oligonucleotides.

Notes and references

1. Dunn MR, Jimenez RM and Chaput JC, Nat. Rev. Chem 2017, 1, 0076 DOI: 10.1038/s41570-017-0076
2. Tuerk C and Gold L, Science, 1990, 249, 505. [PubMed: 2200121]
3. Robertson DL and Joyce GF, Nature, 1990, 344, 467. [PubMed: 1690861] Ellington AD and Szostak JW, Nature, 1990, 346, 818. [PubMed: 1697402]

4. Meek KN, Rangel AE and Heemstra JM, *Methods*, 2016, 106, 29. [PubMed: 27012179]
5. Diafa S and Hollenstein M, *Molecules*, 2015, 20, 16643. [PubMed: 26389865]
6. Lei Y, Kong D and Hili R, *ACS Comb. Sci*, 2016, 18, 355. [PubMed: 27275512]
7. Latham JA, Johnson R and Toole JT, *Nucl. Acids Res*, 1994, 22, 2817. [PubMed: 7519769]
8. MacPherson IS, Temme JS, Habeshian S, Felczak K, Pankiewicz K, Hedstrom L and Krauss IJ, *Angew. Chem., Int. Ed*, 2011, 50, 11238.
9. Ohloff JC, Gelinas AD, Jarvis TC, Ochsner OA, Schneider DJ and Gold L *Mol. Ther. Nucleic Acids*, 2014, 10, e201.
10. Tolle F, Brändle GM, Matzner D and Mayer G, *Angew. Chem., Int. Ed*, 2015, 54, 10971.
11. Zhou C, Avins JL, Klauser PC, Brandsen BM, Lee Y, and
12. Silverman SKJ, *J. Am. Chem. Soc*, 2016, 138, 2106. [PubMed: 26854515]
13. Gawande BN, Rohloff JC, Carter JD, von Carlowitz I, Zhang C, Schneider DJ and Janjic N, *Proc. Natl. Acad. Sci. U. S. A*, 2017, 114, 2898. [PubMed: 28265062]
14. Wang Y, Liu E, Lam CH and Perrin DM, *Chem. Sci*, 2018, 9, 1813. [PubMed: 29675226]
15. Guo C, Watkins CP and Hili R, *J. Am. Chem. Soc*, 2015, 137, 11191–11196. [PubMed: 26274786]
16. Lei Y, Kong D and Hili R, *ACS Comb. Sci*, 2015, 17, 716. [PubMed: 26513677]
17. Kong D, Lei Y, Yeung W and Hili R, *Angew. Chem., Int. Ed*, 2016, 55, 13164.
18. Guo C and Hili R, *Bioconjugate Chem*, 2017, 28, 314.
19. Lei Y and Hili R, *Org. Biomol. Chem*, 2017, 15, 2349. [PubMed: 28244520]
20. Kong D, Yeung W and Hili R, *J. Am. Chem. Soc* 2017, 139, 13977. [PubMed: 28938065]
21. Chen Z, Lichtor PA, Berliner AP, Chen JC and Liu DR, *Nat. Chem* 2018, 10, 420–427. [PubMed: 29507367]
22. Cai L, Hu C, Shen S, Wang W, Huang W, *J. Biochem*, 2004, 135, 397. [PubMed: 15113838]
23. Shi K et al. *Nucl. Acids Res* 2018, In press, DOI: 10.1093/nar/gky776.

Figure 1. Synthesis of diversely-modified DNA polymers by LOOPER. *Inset:* Thrombin aptamer evolved using LOOPER and SELEX.

Figure 2.
Commercially available amine-containing nucleobase modifiers derivatized with by amide coupling.

Figure 3.
Anti and *syn* conformations in dA modified at the 8-position.

Table 1Yield and Fidelity for T3 and T4 DNA Ligase-mediated LOOPER ^a

Entry	DNA ligase	Anticodon	Set size	Yield ^b	Reads ^c	Fidelity ^d
1	T3	NNN	64	60%	460,148	97.0 %
2	T4	NNN	64	<5 %	N/A	N/A
3	T3	NNNN	256	89 %	162,700	89.2 %
4	T4	NNNN	256	<5%	N/A	N/A
5	T3	NNNNN	1024	65%	60,208	63.8%
6	T4	NNNNN	1024	60%	33,376	81.1 %
7	T3	NNNNN	1024	64%	144,712	66.9 % ^e

^aFidelity was determined by Illumina sequencing on a MiSeq (Paired-end 150). 35 amols were used for PCR prior to sequencing to ensure sufficient number of duplex pairs.

^bYield was determined by PAGE analysis.

^cNumber of codon reads from sequencing data.

^dFidelity was calculated for pentanucleotide incorporation; the fidelity is considerably higher if evaluated at the single nucleotide level.

^ePerformed at pH 6.9.

Table 2Influence of modifications on LOOPER with T3 and T4 DNA ligases ^a

Entry	DNA ligase	Anticodon	Set size	Yield ^b	Reads ^c	Fidelity ^d
1	T3	YNN	32	66%	332,256	97.7 %
2	T3	Y*NN	32	67%	115,974	97.4 %
3	T3	A*NN	16	<5%	N/A	N/A
4	T3	YNNN	128	94%	211,693	93.7%
5	T3	Y*NNN	128	93%	107,084	92.1 %
6	T3	A*NNN	64	<5%	N/A	N/A
7	T3	ANNN	64	90%	50,712	93.5% ^e
8	T3	A*NNN	64	20%	257,721	97.2% ^e
9	T3	YNNNN	512	96%	89,382	72.6%
10	T3	Y*NNNN	512	74%	246,600	74.3%
11	T3	ANNNN	256	86%	16,370	82.0%
12	T4	ANNNN	256	87%	1536	86.7%
13	T3	A*NNNN	256	74%	63,287	96.8%
14	T4	A*NNNN	256	75%	31,936	95.1%

^aFidelity was determined by Illumina sequencing on a MiSeq (Paired-end 150). 35 amols were used for PCR prior to sequencing to ensure sufficient number of duplex pairs.

^bYield was determined by PAGE analysis.

^cNumber of codon reads from sequencing data.

^dFidelity was calculated for pentanucleotide incorporation; the fidelity is considerably higher if evaluated at the single nucleotide level.

^ePerformed at pH 8.5