

Substitution of Deoxyinosine as Universal Base in Oligonucleotides for DNA Ligation

Zhiliang Yu

College of Biological and Environmental Engineering, Zhejiang University of Technology, Hangzhou, China
Email: zlyu@zjut.edu.cn

Received 2013

ABSTRACT

Oligonucleotides libraries have been developed for various applications, but the library size of oligonucleotide increases dramatically with the addition of oligonucleotide length. To assess the possibility of shortening library size by using universal base, deoxyinosine (dITP), the effect of single/multiple dITP substituted in oligonucleotide on ligation was investigated. It was found that different pairs with dITP had different ligation patterns and pairs with dITP at different locations also showed different ligation patterns. With the departure of substitution position from ligation site, the ligation yield increased. Single dITP substitution at ligation site did extremely hurt the ligation efficiency, except for I:C pair. On the other hand, single substitution at two bases or more apart from ligation site, there is no obvious effect on ligation. Multiple dITP substitutions can more or less affect the ligation, besides I:C pair. This research demonstrated that dITP can be applied to reduce oligonucleotide library size after substitution.

Keywords: Deoxyinosine; Substitution; Universal Base; Ligation

1. Introduction

Oligonucleotide libraries have been developed for a number of applications: sequencing by hybridization and ligation, mutant diagnostics, gene expression analysis, and identification of microorganisms. The size of a library of oligonucleotide increases dramatically with increasing length of oligonucleotide, e.g. a library of all 8mers would require 65,536 oligonucleotides. People have attempted to reduce library size by using shorter length [1]. However, it has a number of shortcomings, namely low stability of hybrids, difficulties in discriminating terminal mismatches and a wide variation in the stabilities of AT- and GC-rich duplexes. It was shown that the successful way to reduce library size and also compensate for these observed shortcomings is to add universal bases that can pair with any of the four natural bases [2]. Therefore, a universal base, which could substitute for any of the four natural bases, would be of great utility for manipulating DNA.

Inosine occurs naturally in the wobble position of the anticodon of some tRNAs, where it appears to pair with adenosine, cytidine and uridine of the codon of mRNAs [3,4]. 2'-deoxyinosine (dITP) is widely used as an ambiguous nucleoside in oligonucleotide probes for hybridization [5,6], and primers for PCR [7-9] and sequencing [10]. It was also reported that dITP can be incorporated into PCR products when dITP is added to a mixture of

three dNTPs at normal concentrations and the fourth dNTP at lower concentration [11-14]. However, to date, very little work has been reported about the universal base substitution in oligonucleotides for ligation. It has been shown that neither 3-nitropyrrole nor 5-nitropyrrole is recognized by T4 polynucleotide ligase when present at either 5'- or 3'-end of an oligonucleotide [15]. It was found that when 3-nitropyrrole is near to the 3' terminus, *i.e.* still within active site of ligase, it can be recognized by Tth DNA ligase and causes enhanced ligation fidelity. If 3-nitropyrrole was substituted at the second site from the nick no ligation observed [16]. Apart from this nitropyrrole, one of universal bases, there is no report to systematically investigate the dITP substitution for DNA ligation. Nowadays, dITP residues can be commercially and economically substituted in oligonucleotides at terminal and/or internal sites (<https://www.idtdna.com/>), allowing dramatically reducing the oligonucleotide library size. As nothing is known about the dITP substitution for ligation, in this study, the effect of single- and multiple-dITP substitution at different positions on ligation was investigated.

2. Materials and Methods

2.1. Oligonucleotides

The oligonucleotides were designed using Primer Premier 5.0 software, and synthesized by Integrated DNA

Technologies (IDT, Coralville, IA, USA). The required modification with phosphorylation and/or dITP substitution was also performed by IDT, and the oligonucleotides were delivered after PAGE or HPLC purification.

2.2. Hybridization and Ligation

Typically, oligonucleotide probes were buffered in 10 mM Tris-HCl, 1 mM EDTA, pH8.0 solution and the concentration was calculated after UV-Vis measurement using spectrophotometer (Nanodrop 1000, Wilmington, DE, USA). Unless otherwise stated, Hybridization reaction was performed in a 10 μ l reaction mixture containing 1 μ M oligonucleotide mixture after denature at 95°C for 5 min followed by slowly cooling down to 10°C for hybridization. Then DNA oligonucleotides were ligated at room temperature for 30 min after addition of ligation buffer [66 mM Tris-HCl, pH7.6/10 mM MgCl₂/1 mM dithiothreitol/7.5% PEG 6000/1 mM ATP] and Quick DNA ligase (New England Biolabs, MA, USA) in a final volume of 20 μ l. Ligation reactions were terminated by heating at 65°C for 20 min.

2.3. Polyacrylamide Gel Electrophoresis (PAGE)

After ligation, reaction solution was mixed with 6x loading dye and then electrophoresed on a 10% PAGE at 200 mV for 45 min using TBE buffer. After electrophoresis, DNA was stained with SYBR green I (Invitrogen, Carlsbad, USA) and visualized.

3. Results

3.1. Basic Description and Procedure

As shown in **Figure 1**, 23 mer single stranded oligonucleotide TP23F was designed with a 17 bases yellow region, which is complementary to 17 mer 5'-phosphorylated single stranded oligonucleotide BS17 (5'-Phos BS17), and a 6 bases green region designated by 0, -1, -2, -3, -4 and -5, respectively, from 5' to 3' according to the distance from the directly ligated site. To investigate the effect of dITP substitution on ligation, the natural bases at this green region will be substituted by dITP accordingly for ligation, unless otherwise stated. After simple hybridization between 5'-Phos BS17 and TP23F, phosphorylated hybrid with 17 bps duplex DNA flanked by a 6 bases single overhand (in green) at 5' was formed. 30 mer single stranded DNA SST30F with a 24 mer red region and also a 6mer green region, which is designed to be complementary to the green region of TP23F. As required, the natural bases at this green region of SST30F will also be modified through dITP substitution accordingly, unless otherwise stated. After hybridization of SST30F with the above duplex DNA, the longer DNA product with a 20 bps ds-region and a 24 bases ss-region

was formed through ligase treatment.

3.2. Single dITP Substitution at Different Positions for Different Natural Nucleotides

To investigate the effect of single dITP substitute on ligation, first, SST30F was such designed that its green region is poly(dC)₆, poly(dT)₆, poly(dA)₆ or poly(dG)₆ to pair with poly(dG)₆, poly(dA)₆, poly(dT)₆ or poly(dC)₆, respectively, at the green region of TP23F. After addition of 5'-Phos BS17, three natural oligonucleotides without substitution were expectedly ligated by Quick T4 ligase as shown in lane 2 of **Figure 2**. Then, the natural bases at the green region of SST30F were changed individually with single dITP substitution at different sites from position “-5” to “0” for different natural bases (C, T, A or G). Since the green region of TP23F was fixed with natural bases, after hybridization, I:G (**Figure 2(A)**), I:A (**Figure 2(B)**), I:T (**Figure 2(C)**) or I:C (**Figure 2(D)**) pairing accordingly formed at the green region where dI came from SST30F and natural bases from TP23F. As shown in **Figure 2**, single dITP can be used to substitute all four natural bases for ligation at different positions from “-5” to “0”. But ligation yields changed with the change of the distance between substituted base and ligated base. And different pairs between dITP and natural bases had different patterns of ligation yield. For I:G pair in **Figure 2(A)**, there was no obvious effect of single dITP substitution at positions from “-5” to “-2” on ligation yield and then very marginal effect after one base closer to ligated nucleotide. Finally, the ligation yield of I:G pair at position “0” dramatically decreased. Comparatively, very slight effect of dITP substitution on ligation yield occurred for I:A pair at position “-3” in **Figure 2(B)** and I:T pair at position “-2” in **Figure 2(C)**, respectively, instead of position “-1” for I:G pair. Then the ligation yield significantly dropped if dITP substitution was put at position “-2” or “-1” for I:A pair and position “-1” for I:T pair. Like for I:G pair at position “0” in **Figure 2(A)**, the ligation yields of dITP substitution at position “0” for both I:A and I:T pairs were also very weak, indicating almost zero ligation. In contrast, I:C pair in **Figure 2(D)** showed very high ligation yield no matter where the substitution is. Even substitution at position “0” had high ligation yield.

3.3. Multiple dITP Substitutions at “-3”, “-4” or “-5” Position for Different Natural Nucleotides

3.3.1. Triple Substitutions at “-3”, “-4” and “-5” Positions

The results in **Figure 2** indicate that for all four natural nucleotides there is no apparent effect of single substitution from position “-5” to “-3” on ligation yield. There-

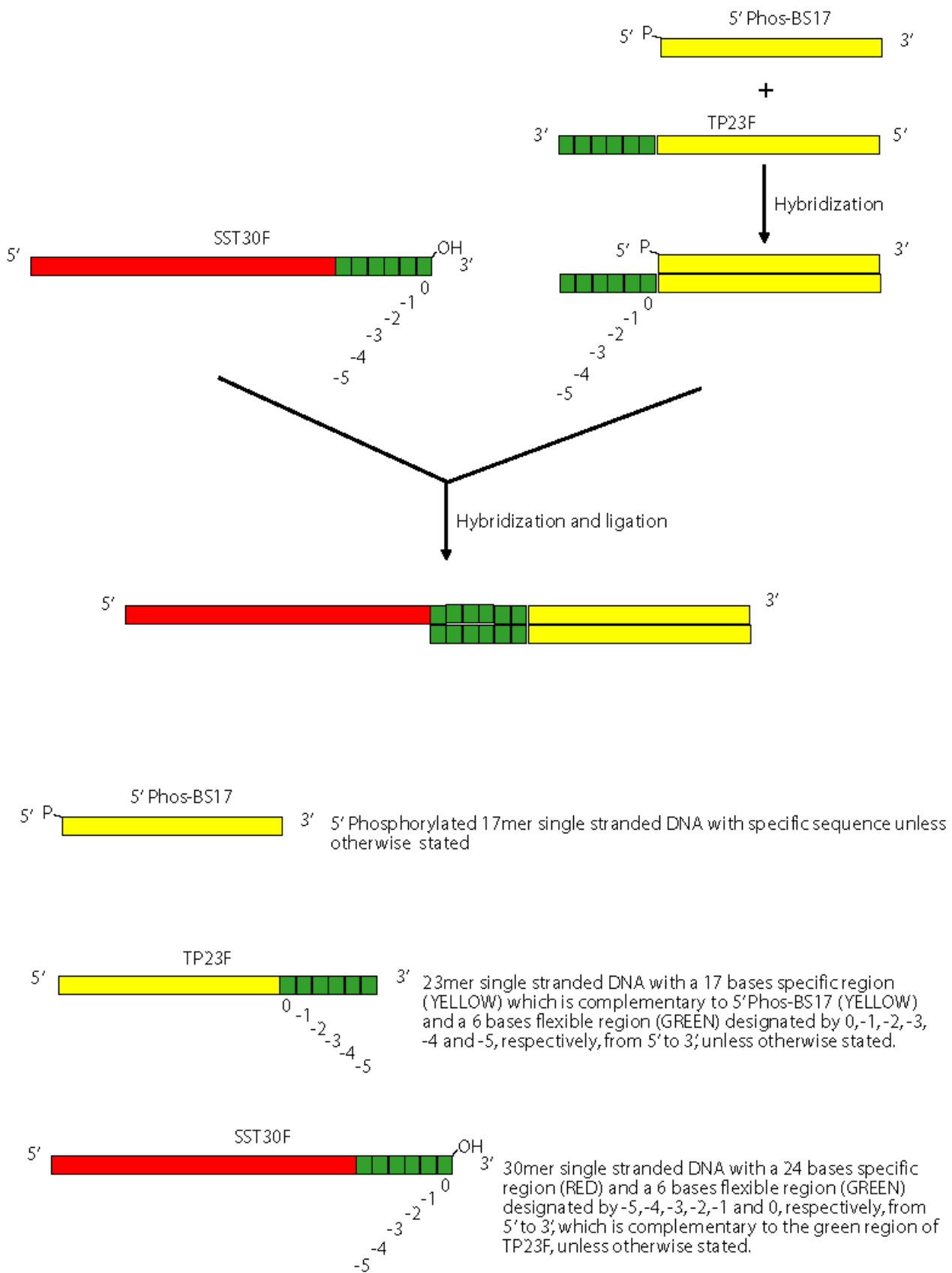


Figure 1. Schematic presentation of ligation with dITP substitution.

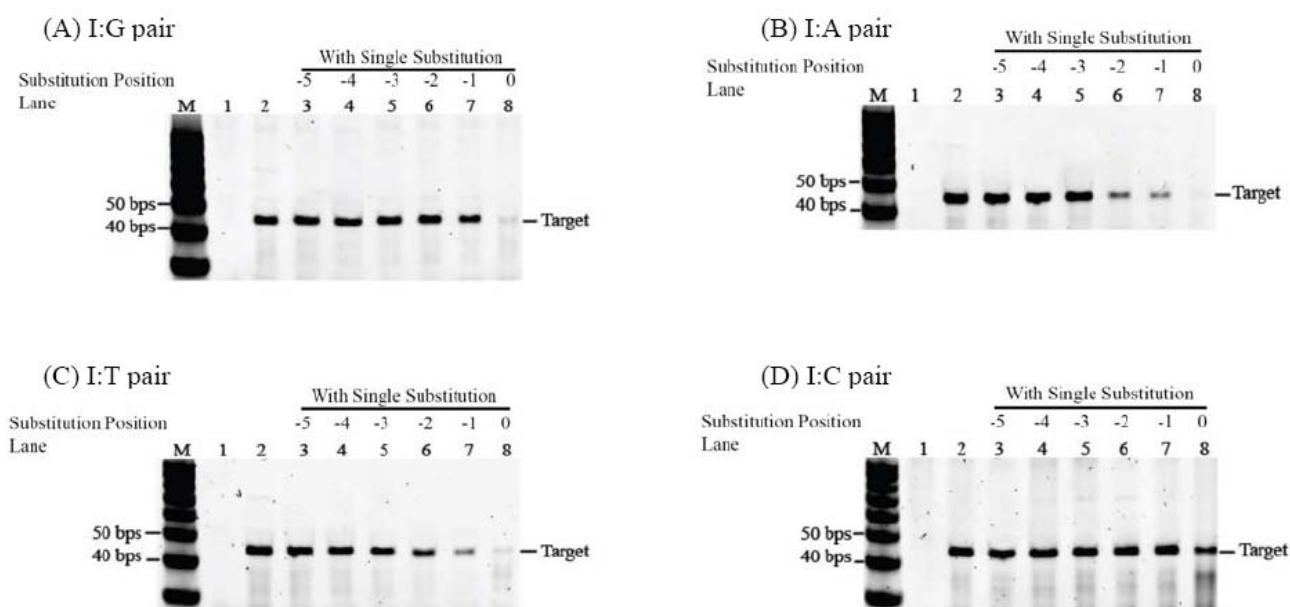


Figure 2. Ligation yields derived from I:G pair (A), I:A pair (B), I:T pair (C) and I:C pair (D), respectively, where dI is at the green segment of SST30F and paired natural bases (dNTPs) are at the green region of TP23F, at different positions from “-5” to “0” (from lane 3 to lane 8, respectively) according to the distance from the ligated base. As expected, positive control (lane 2) with all components has the desired ligated product based on the 10 bps DNA ladder (lane M), and negative control (lane 1) without DNA ligase doesn’t have the ligated DNA.

fore, the question arises: is it possible to make triple dITP substitutions at positions “-3”, “-4” and “-5” for all natural bases? To address this question, TP23F was such designed that its green region from position “-5” to “0” is poly(dI)₃(dC)₃, poly(dI)₃(dT)₃, poly(dI)₃(dA)₃ or poly(dI)₃(dG)₃ to pair with poly(dG)₆, poly(dA)₆, poly(dT)₆ or poly(dC)₆, respectively, at the green region of SST30F. After formation of hybrid between TP23F and SST30F, 3 G:Is, 3 A:Is, 3 T:Is, or 3 C:Is will accordingly be formed. As shown in **Figure 3**, both triple A:I and T:I substitutions gave similar results where positive controls without substitutions (-) showed ligated products but substituted reactions (+) didn’t have desired products. In contrast, both positive control and substituted reaction for triple C:I pairs had nice target products, indicating DNA molecules with triple dITP substitutions at positions “-3”, “-4” and “-5” can be effectively ligated by Quick DNA ligase. Interestingly, G:I pairs had similar results as C:I pairs, except that the ligated product in substituted reaction from G:I pairs migrated faster than the ligated products from C:I pairs. Actually, based on the DNA ladder (lane M) and control experiment without SST30F (data not shown), it was indicated that the ligated product in substituted reaction for triple G:I pairs was the 40 bps two-BS17-TP23F concatenated molecule with 3 C:I pairs followed by 3 I:C pairs in the 6 bps green region instead of the desired product, SST30F-BS17-TP23F, further indicating that triple C:I or I:C pairs can be effectively ligated by Quick DNA ligase.

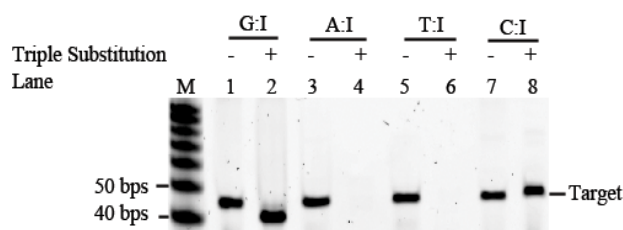


Figure 3. Effect of triple dITP substitutions on ligation.

3.3.2. Double Substitutions at “-3”, “-4” or “-5” Position

Since triple substitutions for A:I, T:I and G:I pairs cannot work using Quick DNA ligase, except for C:I pairs, double substitutions at position “-3”, “-4” or “-5” for A:I, T:I and G:I pairs may be achievable. Therefore TP23F was designed in the same way as in triple substitutions experiments, except that only two dITPs were put at two out of three positions from “-3” to “-5” instead of three dITPs. The results in **Figure 4** showed that the combination of double substitutions at position “-4” and “-5” out of three combinations for all three types of pairs had the highest ligation yield, but compared to the ligation without any substitution, the ligation yield of this combination at position “-4” and “-5” was still much lower. This observation tells us two things. First, the distance of first substituted nucleotide away from the ligated site is critical, which totally agrees with the findings of single substitution experiments; second, compared to single substitution or triple substitutions, after one more or less subs-

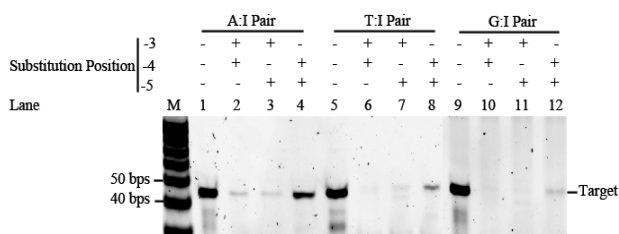


Figure 4. Effect of double dITP substitutions on ligation.

titution, the yield drops or enhances obviously, respectively, indicating that substitution number is also crucial for ligation yield with substitution. Moreover, the overall order of ligation yield with double substitutions is A:I > T:I > G:I.

3.4. Recognition of dITP by Different Ligases

To investigate the feasibility of ligation with substitution at position “0” by different ligases, the widely used ligases including Quick T4 ligase, normal T4 DNA ligase, *E.coli* DNA ligase and *Taq* DNA ligase were applied to ligate I:A pair at position “0”. As shown in **Figure 5**, different ligases had different ligation efficiencies for ligation reactions without substitution; Quick T4 ligase showed the highest yield followed by normal T4 DNA ligase and *E.coli* DNA ligase; *Taq* DNA ligase almost gave no ligation due to its short overhang. In contrast, all ligases tested in this study gave similar results with very low ligation yield after dITP substitution. Since there is no difference in ligation yield with substitution among 4 ligases, and Quick T4 ligase gave the highest ligation yield without substitution, Quick T4 ligase was chosen for subsequent further tests.

3.5. Increasing of Ligation Yield by Adding dITP

To investigate the feasibility of increasing the ligation yield by adding dITP, TP23F was such designed that its green region is poly(dT)₆ to be complementary to poly(dA)₆ at the green region of SST30F. Since there is no substitution, after ligation, it expectedly gave very nice ligated band as shown in lane 2 of **Figure 6**. If poly(dT)₆ of TP23F at the green region was shortened to poly(dT)₄ with two bases reduction, accordingly the ligation yield dropped remarkably as shown in lane 3 of **Figure 6**, which is presumably due to the decrease of thermal stability. Then, if two dITPs were added at 3'-end of TP23F to enlarge its green region to be 6 bases again with a sequence of poly(dT)₄(dI)₂, the ligation yield was dramatically recovered as shown in lane 4 of **Figure 6**. However, if two dCs instead of two dITPs were used, the ligation yield further decreased as shown in lane 5 of **Figure 6**, indicating that mismatch introduction can seriously hurt the ligation. All these findings indicate two things: 1) dITP can be used to increase the effective oligonucleo-

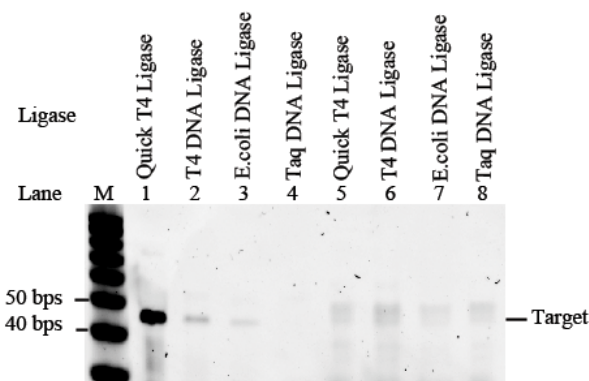


Figure 5. Recognition of dITP substituted in oligonucleotide by ligase.

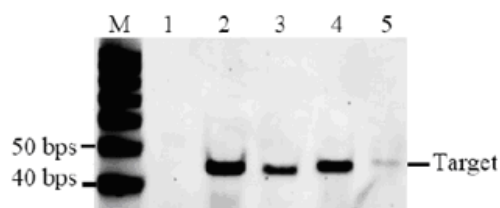


Figure 6. Increase of ligation yield after dITP substitution.

tide size for ligation; 2) multiple dITPs can be substituted within the certain bases from the ligated site as found above.

4. Conclusions

4.1. For Single Substitution

1) Different pairs with dITP had different ligation patterns: I:C pair had the highest ligation efficiency, followed by I:G pair, I:T pair and I:A pair.

2) Substitution at different locations have different ligation patterns: dITP substitution directly at ligation site did extremely hurt ligation yields for I:G pair, I:T pair and I:A pair, except for I:C pair. With the departure of substitution position from ligation site, the ligation increased. dITP substitution at two bases or more apart from ligation site had no effect on ligation.

4.2. For Multiple Substitutions

1) Triple substitutions: There was no ligation yield with triple substitutions at positions “-3”, “-4” and “-5” for I:G pair, I:T pair and I:A pair. However, ligation with triple substitutions for I:C pair still worked well.

2) Double substitutions: The combination of double substitutions at position “-4” and “-5” out of three combinations for all three types of pairs had the highest ligation yield, but compared to the ligation without any substitution, the ligation yield of this combination was still much lower. Moreover, the overall order of ligation yield with double substitutions is A:I > T:I > G:I.

4.3. All These Findings Indicate

1) dITP can be used to increase the effective oligonucleotide size for ligation;

2) Multiple dITPs can substitute the certain bases from the ligated site as found above.

5. Acknowledgements

This work was supported by Natural Science Foundation of Zhejiang Province, China (Y5100153), Welfare Technology Applied Research Project of Zhejiang Province, China (2011C23007), and Natural Science Foundation of ZJUT (20100213) to Z.L. Yu.

REFERENCES

- [1] A. V. Fotin, A. L. Drobyshev, A. N. Perov and A. D. Mirzabekov, "Parallel Thermodynamic Analysis of Duplexes on Oligodeoxyribonucleotide Microchips," *Nucleic Acids Research*, Vol. 26, No. 6, 1998, pp. 1515-1521. <http://dx.doi.org/10.1093/nar/26.6.1515>
- [2] S. Parinov, V. Barsky, G. Yershov, E. Kirillov, E. Timofeev, A. Belgovskiy and A. Mirzabekov, "DNA Sequencing by Hybridization to Microchip Octa- and Decanucleotides Extended by Stacked Pentanucleotides," *Nucleic Acids Research*, Vol. 24, No. 15, 1996, pp. 2998-3004. <http://dx.doi.org/10.1093/nar/24.15.2998>
- [3] F. H. Crick, "Codon-Anticodon Pairing: the Wobble Hypothesis," *Journal of Molecular Biology*, Vol. 19, No. 2, 1966, pp. 548-555. [http://dx.doi.org/10.1016/S0022-2836\(66\)80022-0](http://dx.doi.org/10.1016/S0022-2836(66)80022-0)
- [4] M. D. Topal and J. R. Fresco, "Base Pairing and Fidelity in Codon-anticodon Interaction," *Nature*, Vol. 263, No. 5575, 1976, pp. 289-293. <http://dx.doi.org/10.1038/263289a0>
- [5] E. Ohtsuka, S. Matsuki, M. Ikehara, Y. Takahashi and K. Matsubara, "An Alternative Approach to Deoxyoligonucleotides as Hybridization Probes by Insertion of Deoxyinosine at Ambiguous Codon Positions," *Journal of Biological Chemistry*, Vol. 260, No. 5, 1985, pp. 2605-2608.
- [6] Y. Kawase, S. Iwai, H. Inoue, K. Miura and E. Ohtsuka, "Studies on Nucleic Acid Interactions. I. Stabilities of Mini-Duplexes (dG2A4XA4G2-dC2T4YT4C2) and Self-Complementary d(GGGAAXYTTCCC) Containing Deoxyinosine and Other Mismatched Bases," *Nucleic Acids Research*, Vol. 14, No. 19, 1986, pp. 7727-7736. <http://dx.doi.org/10.1093/nar/14.19.7727>
- [7] R. V. Patil and E. E. Dekker, "PCR Amplification of an *Escherichia coli* Gene Using Mixed Primers Containing Deoxyinosine at Ambiguous Positions in Degenerate Amino Acid Codons," *Nucleic Acids Research*, Vol. 18, No. 10, 1990, p. 3080. <http://dx.doi.org/10.1093/nar/18.10.3080>
- [8] H. Liu and R. Nichols, "PCR Amplification Using Deoxyinosine to Replace An Entire Codon and at Ambiguous Positions," *Biotechniques*, Vol. 16, No. 1, 1994, pp. 24-26.
- [9] D. R. Kilpatrick, B. Nottay, C. F. Yang, S. J. Yang, M. N. Mulders, B. P. Holloway, *et al.*, "Group-Specific Identification of Polioviruses by PCR Using Primers Containing Mixed-Base or Deoxyinosine Residue at Positions of Codon Degeneracy," *Journal of Clinical Microbiology*, Vol. 34, No. 12, 1996, pp. 2990-2996.
- [10] D. Loakes, D. M. Brown, S. Linde and F. Hill, "3-Nitropyrrole and 5-Nitroindole as Universal Bases in Primers for DNA Sequencing and PCR," *Nucleic Acids Research*, Vol. 23, No. 13, 1995, pp. 2361-2366. <http://dx.doi.org/10.1093/nar/23.13.2361>
- [11] H. Dierick, M. Stul, W. De Kelder, P. Marynen and J. J. Cassiman, "Incorporation of dITP or 7-Deaza dGTP during PCR Improves Sequencing of the Product," *Nucleic Acids Research*, Vol. 21, No. 18, 1993, pp. 4427-4428. <http://dx.doi.org/10.1093/nar/21.18.4427>
- [12] S. L. Turner and F. J. Jenkins, "Use of Deoxyinosine in PCR to Improve Amplification of GC-rich DNA," *Biotechniques*, Vol. 19, No. 1, 1995, pp. 48-52.
- [13] O. P. Kuipers, "Random Mutagenesis by Using Mixtures of dNTP and dITP in PCR," *Methods in Molecular Biology*, Vol. 57, 1996, pp. 351-356.
- [14] A. Kobayashi, M. Kitaoka and K. Hayashi, "Analyses of PCR Products Using DNA Templates Containing A Consecutive Ceoxyinosine Sequence," *Nucleic Acids Symposium Series (Oxford)*, Vol. 48, 2004, pp. 225-226. <http://dx.doi.org/10.1093/nass/48.1.225>
- [15] D. Loakes, A. Van Aerschot, D. M. Brown and F. Hill, "Enzymatic Recognition of Acyclic Universal Base Analogues in Oligonucleotides," *Nucleosides & Nucleotides*, Vol. 15, No. 11-12, 1996, pp. 1891-1904. <http://dx.doi.org/10.1080/07328319608002740>
- [16] J. Luo, D. E. Bergstrom and F. Barany, "Improving the Fidelity of *Thermus thermophilus* DNA Ligase," *Nucleic Acids Research*, Vol. 24, No. 15, 1996, pp. 3071-3078. <http://dx.doi.org/10.1093/nar/24.15.3071>