Supporting Information

Noncanonical substrate preference of lambda exonuclease for 5'-nonphosphate-ended dsDNA and a mismatch-induced acceleration effect on the enzymatic reaction

Tongbo Wu¹, Yufei Yang¹,², Wei Chen¹, Jiayu Wang¹, Ziyu Yang¹, Shenlin Wang¹,², Xianjin Xiao³, Mengyuan Li³ and Meiping Zhao¹,*

¹Beijing National Laboratory for Molecular Sciences, MOE Key Laboratory of Bioorganic Chemistry and Molecular Engineering, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China
²Beijing NMR Center, Peking University, Beijing 100871, China.
³Family Planning Research Institute/Center of Reproductive Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China.
E-mail: mpzhao@pku.edu.cn

Table S1.............................................................................................................................................................................. 2
Scheme S1............................................................................................................................................................................ 4
Figure S1............................................................................................................................................................................. 6
Supplementary results and discussion.................................................................................................................................. 7
Figure S2.............................................................................................................................................................................. 7
Figure S3.............................................................................................................................................................................. 8
Figure S4.............................................................................................................................................................................. 9
Figure S5.............................................................................................................................................................................. 10
Figure S6.............................................................................................................................................................................. 11
Figure S7.............................................................................................................................................................................. 12
Table S2.............................................................................................................................................................................. 13
Table S3.............................................................................................................................................................................. 14
Figure S8.............................................................................................................................................................................. 16
Scheme S2............................................................................................................................................................................ 17
References........................................................................................................................................................................... 19
| Name                                                                 | Sequences (5'->3')b |
|----------------------------------------------------------------------|---------------------|
| **Sequences of the probe**                                          |                     |
| P1-5'OH-6FAM                                                         | TCGTCT(-FAM)CCACAGACACATACTCCA-3'BHQ-1 |
| P1'-5'OH-unlabeled                                                   | TCGTCTCCACAGACACATACTCCA |
| P2-5'PO4-6FAM                                                        | 5'PO4-TCGTCT(-FAM)CCACAGACACATACTCCA-3'BHQ-1 |
| P3-5'OH-15FAM                                                        | TCTCCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P5-5'C6-15FAM                                                        | 5'C6-TCTCCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P6-5'COOH-15FAM                                                      | 5'COOH-TCTCCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P7-5'SH-15FAM                                                        | 5'SH-TCTCCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P8-5'NH2-15FAM                                                       | 5'NH2-TCTCCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P9-5'PO4-15FAM                                                       | 5'PO4-TCTCCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P10-5'FAM                                                            | 5'FAM-TCTCCACAGACACATAC-3'BHQ-1 |
| P11-5'OH-dSpacer-15FAM                                               | 5'dSpacer-TCTCCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P12-5'C6-3FAM                                                        | 5'C6-TCT(-FAM)CCACAGACACATACTCCA-3'BHQ-1 |
| P13-5'PO4-3FAM                                                       | 5'PO4-TCT(-FAM)CCACAGACACATACTCCA-3'BHQ-1 |
| P14-5'PO4-10FAM                                                      | 5'PO4-TCTCCACAGCT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P15-5'PO4-56FAM                                                      | 5'PO4-CTGCCCTAAATCATGTTGCGTGAATCGCCATATTTAAACAAATAGGCCT(-FAM)CGCTGCCGTCGCAA-3'BHQ-1 |
| P16-5'PO4-3digoxin-15FAM                                             | 5'PO4-TCT(-digoxin)CCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P17-5'PO4-3BHQ                                                       | 5'PO4-TCT(-BHQ-1)CCACAGACACATACTCCA-3'BHQ-1 |
| P18-5'PO4-3biotin-15FAM                                              | 5'PO4-TCT(-biotin)CCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P19-5'PO4-3NH2-15FAM                                                 | 5'PO4-TCT(-NH2)CCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-1 |
| P20-5'PO4-3ROX-15FAM                                                 | 5'PO4-TCT(-ROX)CCACAGACACAT(-FAM)ACTCCATAATTAA-3'BHQ-2 |
| P21-5'PO4-3FAM                                                       | 5'PO4-TCT(-FAM)CCACAGACACAT(-ROX)ACTCCATAATTAA-3'BHQ-2 |

**Complementary strands for P1, P1’ and P2**

| C1-PM(42)                                                 | GTTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
|-----------------------------------------------------------|--------------------------------------------|
| C1-1-mis(42)                                              | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-2-mis(42)                                              | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-3-mis(42)                                              | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-4-mis(42)                                              | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-1,3-mis(42)                                            | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-2,4-mis(42)                                            | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-1,2-mis(42)                                            | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-1,2-mis(35)                                            | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-2-mis(34)                                              | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-PM(33)                                                 | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-PM(32)                                                 | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-PM(31)                                                 | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C1-PM(30)                                                 | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |

**Complementary strands for P3, P4, P5, P6, P7, P8, P9, P10, P11 and P12**

| C2-PM(42)                                                 | GTTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
|-----------------------------------------------------------|--------------------------------------------|
| C2-PM(33)                                                 | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| C2-PM(32)                                                 | GTTTTAATATGGAGATATGGTCTGTGGAGACGAGAGTAAG |
| Complementary strands for P9, P13, P16, P17, P18, P19, P20, P21 and P23 |
|--------------------------|
| C3-PM                    |
| C3-1-mis                 |
| C3-2-mis                 |
| C3-3-mis                 |
| C3-4-mis                 |
| C3-5-mis                 |
| C3-6-mis                 |
| C3-7-mis                 |
| C3-8-mis                 |
| C3-9-mis                 |
| C3-10-mis                |
| C3-11-mis                |
| C3-12-mis                |
| C3-13-mis                |
| C3-14-mis                |

| Complementary strands for P14 |
|-------------------------------|
| C4-PM                         |
| C4-9-mis                      |

| Complementary strands for P15 |
|-------------------------------|
| C5-PM                         |
| C5-55-mis                     |

---

* Each labeled probe has a quencher at the 3' end (BHQ-1 or BHQ-2). BHQ1 is Black Hole Quencher 1 and BHQ2 is Black Hole Quencher 2. The positions of fluorophores and other internal labels in the probe sequence are indicated according to its position from the 5' end of the probe. The modifications at the 5' end of the probe are also specified in the probe name. For example, "P18-5'PO₄-3biotin-15FAM" means Probe 18 has a 5' PO₄ end and FAM is labeled at Position 15 from the 5' end. Besides, the base at Position 3 from the 5' end is labeled with biotin. FAM is 6-carboxy-fluorescein and ROX is 6-carboxy-X-rhodamine.

* All the complementary DNA sequences without specified 5' modifications have a 5'-OH end. The bases shown in bold and underlined in the complementary strands indicate that they are mismatched with the opposite bases in the probe.
Scheme S1. Sequences of the longest perfectly matched (PM) dsDNA substrates used in this work.

In the following DNA duplexes, T represents T(-FAM), indicating that FAM is labeled at dT. The label attached to T is specified after the name of each related probe. Underlined part indicates the unpaired region.

P1-5'OH-6FAM/C1-PM(42):
5'-OH-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P2-5'PO4-6FAM/C1-PM(42):
5'-PO4-TGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P3-5'OH-15FAM/C2-PM(42):
5'-OH-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P5-5'C6-15FAM/C2-PM(42):
5'C6-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P6-5'COOH-15FAM/C2-PM(42):
5'COOH-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P7-5'H5-15FAM/C2-PM(42):
5'H5-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P8-5'NH2-15FAM/C2-PM(42):
5'NH2-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P9-5'PO4-15FAM/C3-PM:
5'PO4-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P10-5'FAM/C2-PM(42):
5'FAM-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P11-5'OH-dSpacer-15FAM/C2-PM(42):
5'dSpacer-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P12-5'C6-3FAM/C2-PM(42):
5'C6-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5'-OH

P13-5'PO4-3FAM/C3-PM:
5'PO4-TCGTCTCCACAGACACATACTCCA-3'BHQ-1
3'-GAATGAGACGAGGGTGCTGTGTATGAGGTTTAAATTTTG-5'-'OH

P14-5'PO_4-10FAM/C4-PM:
5'PO_4-TCTCCACAGCACATACTCCCA-3'BHQ-1
3'-GAATGAGACGAGGGTGCTGTGTATGAGGTTTAAATTTTG-5'-'OH

P15-5'PO_4-56FAM/PM:
5'PO_4-CTGCCTAAAATTACATGTTGGCAGGAATCAGCCATATTTAAACAAATTAAGCCGCATGCTGCGCTGCCA-3'BHQ-1
3'-GACGGATTTAATGTACACCGGCATCCTTTAGGGTATATTGTTTAAATCAGGAGCGACGCGGTCGGA-5'-'OH

P16-5'PO_4-3digoxin-15FAM/C3-PM: Here T represents T(-digoxin).
5'PO_4-TCACACAGACAGACTCCATAATTTA-3'BHQ-1
3'-GAATGAGGAGAGGTGGCTGTGATTGAGGTATTTAATTTTG-5'-'OH

P17-5'PO_4-3BHQ/C3-PM: Here T represents T(-BHQ-1).
5'PO_4-TCACACAGACAGACTCCATAATTTA-3'BHQ-1
3'-GAATGAGGAGAGGTGGCTGTGATTGAGGTATTTAATTTTG-5'-'OH

P18-5'PO_4-3biotin-15FAM/C3-PM: Here T represents T(-biotin).
5'PO_4-TCACACAGACAGACTCCATAATTTA-3'BHQ-1
3'-GAATGAGGAGAGGTGGCTGTGATTGAGGTATTTAATTTTG-5'-'OH

P19-5'PO_4-3NH_2-15FAM/C3-PM: Here T represents T(-NH_2).
5'PO_4-TCACACAGACAGACTCCATAATTTA-3'BHQ-1
3'-GAATGAGGAGAGGTGGCTGTGATTGAGGTATTTAATTTTG-5'-'OH

P20-5'PO_4-3ROX-15FAM/C3-PM: Here T represents T(-ROX).
5'PO_4-TCACACAGACAGACTCCATAATTTA-3'BHQ-2
3'-GAATGAGGAGAGGTGGCTGTGATTGAGGTATTTAATTTTG-5'-'OH

P21-5'PO_4-3FAM-15ROX/C3-PM: Here T represents T(-ROX).
5'PO_4-TCACACAGACAGACTCCATAATTTA-3'BHQ-2
3'-GAATGAGGAGAGGTGGCTGTGATTGAGGTATTTAATTTTG-5'-'OH
Figure S1. Chemical structures of the modifications. (a) Structures of the modifications at the 5’ end and the 3’ end. (b) Structures of the tags labeled at the internal bases of DNA strands.
Supplementary results and discussion

**Figure S2.** (a) Typical fluorescence curves of P1-5’OH-6FAM-duplexes and P2-5’PO₄-6FAM-duplexes digested by λ exo (corresponding to Figure 1a). (b) The complete sequences of the tested P1-5’OH-6FAM/C1-PM(42) and P2-5’PO₄-6FAM/C1-PM(42) duplexes are shown, where T represents T(-FAM) and underlined part indicates the unpaired region. In the schematic structures, the mismatched bases are indicated in red in the tested mismatched DNA duplexes. Only the first fifteen base pairs from the 5’ end of the labeled probe were shown for comparison with the PM duplex. The relative rates of fluorescence increase of P2-5’PO₄-6FAM/C1-PM(42) and P2-5’PO₄-6FAM/C1-2-mis(42) duplexes were found to be 545±2 and 611±9, respectively, in comparison with that of P1-5’OH-6FAM/C1-PM(42) duplex which was set to 1.0.
Figure S3. (a) Fluorescence curves of shorter P1-5’OH-6FAM duplexes digested by λ exo. (b) The complete sequence of the P1-5’OH-6FAM/C1-PM(42) duplex is shown, where T represents T(-FAM) and underlined part indicates the unpaired region. In the schematic structures, only the first fifteen base pairs from the 5’ end of the labeled probe were shown to highlight the variation of the sequences near 5’ end in comparison with the P1-5’OH-6FAM/C1-PM(42) duplex. Other parts of all the tested duplexes are identical. The mismatched bases are indicated in red in the tested mismatched DNA duplexes.
Figure S4. The digestion rates of different 5'-OH DNA duplexes by λ exo. (a) The relative rates of fluorescence increase of P3-5’OH-15FAM-duplexes. The rate of fluorescence increase of P3-5’OH-15FAM/C2-PM(42) was set to 1.0. The complementary strands used were indicated in the legend of the figure. (b) The complete sequence of the P3-5’OH-15FAM/C2-PM(42) duplex is shown, where T represents T(-FAM) and underlined part indicates the unpaired region. In the schematic structures of P3-5’OH-15FAM and the complementary strands (C2 series), only the first fifteen base pairs from the 5' end of the labeled probe were shown to highlight the variation of the sequences near 5' end in comparison with the P3-5’OH-15FAM/C2-PM(42) duplex. Other parts of all the tested duplexes are identical. The mismatched bases are indicated in red in the tested mismatched DNA duplexes.
P3-5’OH-15FAM/C2-PM(42):
5’-OH-TCTCCACAGACACATACTCCATAATTTAA-3’BHQ-1
3’- GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5’-OH

P6-5’COOH-15FAM/C2-PM(42):
5’COOH-TCTCCACAGACACATACTCCATAATTTAA-3’BHQ-1
3’- GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5’-OH

P7-5’HS-15FAM/C2-PM(42):
5’HS-TCTCCACAGACACATACTCCATAATTTAA-3’BHQ-1
3’- GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5’-OH

P8-5’NH₂-15FAM/C2-PM(42):
5’NH₂-TCTCCACAGACACATACTCCATAATTTAA-3’BHQ-1
3’- GAATGAGAGCAGAGGTGTCTGTGTATGAGGTATTAAATTTTG-5’-OH

Figure S5. Fluorescence intensity responses of P3-5’OH-15FAM, P6-5’COOH-15FAM, P7-5’HS-15FAM and P8-5’NH₂-15FAM duplexes digested by λ exo. The complete sequences of the P3-5’OH-15FAM/C2-PM(42), P6-5’COOH-15FAM/C2-PM(42), P7-5’HS-15FAM/C2-PM(42) and P8-5’NH₂-15FAM/C2-PM(42) duplexes are shown, where T(-FAM) represents T(-FAM) and underlined part indicates the unpaired region. All the duplexes with C2-mis(42) have a 2-nt overhang structure.
Figure S6. (a) Fluorescence intensity responses of P3-5’OH-15FAM-duplexes digested by λ exo. P3-5’OH-15FAM/C2-2-mis(42) has a 2-nt overhang structure. P3-5’OH-15FAM/C2-3-mis(42), P3-5’OH-15FAM/C2-2,3-mis(42) and P3-5’OH-15FAM/C2-1,2,3-mis(42) have a 3-nt overhang structure. (b) Fluorescence intensity responses of P5-5’C6-15FAM-duplexes digested by λ exo. P5-5’C6-15FAM/C2-PM(32), P5-5’C6-15FAM/C2-PM(31), P5-5’C6-15FAM/C2-PM(30), P5-5’C6-15FAM/C2-PM(29) and P5-5’C6-15FAM/C2-PM(28) have a 0-nt, 1-nt, 2-nt, 3-nt, and 4-nt overhang structure, respectively.
**Figure S7.** The relative digestion rates of P14-5’PO₄-10FAM and P15-5’PO₄-56FAM duplexes by λ exo. The relative rate of fluorescence increase of P14-5’PO₄-10FAM/C4-PM was set to 1.0 for P14-5’PO₄-10FAM/C4-9-mis. The relative rate of fluorescence increase of P15-5’PO₄-56FAM/C5-PM was set to 1.0 for P15-5’PO₄-56FAM/C5-55-mis. The complete sequences of the P14-5’PO₄-10FAM/C4-PM and P15-5’PO₄-56FAM/C5-PM duplexes are shown, where T represents T(-FAM) and underlined part indicates the unpaired region. In the schematic structures of the variation part of the duplexes, the mismatched bases are indicated in red in the tested mismatched DNA duplexes.
Table S2. t values obtained for 2-mis duplexes of different probes compared with the P9-5’PO₄-15FAM/C3-2-mis duplex.

| Probe                                | X    | S   | t    |
|--------------------------------------|------|-----|------|
| P9-5’PO₄-15FAM                       | 1.13 | 0.05| N.A. ¤ |
| P13-5’PO₄-3FAM                       | 1.43 | 0.06| 6.65 †|
| P17-5’PO₄-3BHQ                       | 1.56 | 0.09| 7.23 |
| P16-5’PO₄-3digoxin-15FAM             | 1.18 | 0.06| 1.11 |
| P18-5’PO₄-3biotin-15FAM              | 1    | 0.06| -2.88|
| P19-5’PO₄-3NH₂-15FAM                 | 1.14 | 0.04| 0.27 |
| P20-5’PO₄-3ROX-15FAMª                | 1.39 | 0.03| 7.72 |
| P20-5’PO₄-3ROX-15FAMª                | 1.55 | 0.02| 13.51|
| P21-5’PO₄-3FAM-15ROXᶜ                | 1.51 | 0.06| 8.43 |
| P21-5’PO₄-3FAM-15ROXᵈ                | 1.43 | 0.01| 10.19|

a, d The rate of fluorescence increase was calculated based on the signal from ROX channel.
b, c The rate of fluorescence increase was calculated based on the signal from FAM channel.
¢ Not applicable.
† The t value for P=0.01 at n’=5 is 4.032. The t values that are of significance were marked in bold.

The t value was calculated using

\[
t = \frac{X_1 - X_2}{S_{X_1X_2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad S_{X_1X_2} = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 1},
\]

where \(X_1\) is the average value of the relative rate of fluorescence increase of P9-5’PO₄-15FAM/C3-2-mis and \(X_2\) is that of other 2-mis substrates. \(S_1\) (or \(S_2\)) is the standard deviation of \(X_1\) (or \(X_2\)). The values of \(X\) and \(S\) were obtained in Table S2. \(n\) stands for the number of replicate experiments and \(n = 3\) for all the substrates.
Table S3. Enzyme kinetics of the reactions between λ exo and different DNA duplexes.

| Duplexes                      | Mismatch acceleration | $K_m$ (nmol·L$^{-1}$) | $k_{cat}$ (s$^{-1}$) | $k_{cat}/K_m$ (nmol·L$^{-1}$·s$^{-1}$) |
|-------------------------------|-----------------------|-----------------------|----------------------|---------------------------------------|
| P13-5’PO$_4$-3FAM/C3-PM       | N.A.$^b$              | 9.0±1.8               | 1.34±0.09            | 0.15±0.03                             |
| P13-5’PO$_4$-3FAM/C3-2-mis    | Yes                   | 44±12                 | 4.66±0.76            | 0.11±0.03                             |
| P17-5’PO$_4$-3BHQ/C3-PM       | N.A.                  | 8.6±0.8               | 0.56±0.02            | 0.07±0.01                             |
| P17-5’PO$_4$-3BHQ/C3-2-mis    | Yes                   | 39±9                  | 3.45±0.30            | 0.09±0.02                             |
| P9-15FAM/C3-PM                | N.A.                  | 8.4±0.8               | 0.293±0.007          | 0.035±0.003                           |
| P9-15FAM/C3-2-mis             | No                    | 10.4±2.1              | 0.33±0.02            | 0.032±0.007                           |
| P9-15FAM/C3-14-mis            | Yes                   | 23.0±2.4              | 0.67±0.03            | 0.029±0.003                           |
| P16-5’PO$_4$-3digoxin-15FAM/C3-PM | N.A.            | 9.8±1.6               | 0.26±0.01            | 0.026±0.004                           |
| P16-5’PO$_4$-3digoxin-15FAM/C3-2-mis | No            | 10.6±2.2              | 0.30±0.01            | 0.028±0.006                           |
| P16-5’PO$_4$-3digoxin-15FAM/C3-14-mis | Yes          | 31.1±2.6              | 0.78±0.03            | 0.025±0.002                           |
| P19-5’PO$_4$-3NH$_2$-15FAM/C3-PM | N.A.            | 7.2±0.2               | 0.16±0.01            | 0.021±0.002                           |
| P19-5’PO$_4$-3NH$_2$-15FAM/C3-2-mis | No             | 10.2±0.7              | 0.24±0.02            | 0.024±0.003                           |
| P19-5’PO$_4$-3NH$_2$-15FAM/C3-14-mis | Yes           | 60.8±9.2              | 2.45±0.25            | 0.040±0.007                           |

$^a$The mismatched duplexes that showed significant difference in the digestion rate from the corresponding PM-duplexes in student’s t test (Table S3) were regarded to have mismatch acceleration effect. Otherwise the mismatched duplexes were determined to have negligible mismatch acceleration effect.

$^b$Not applicable for the PM-duplexes.

As shown in Table S3, the $K_m$ values of the PM-duplexes are generally in the range from 7 to 11 nmol/L, while those of the duplexes with mismatched base pairs are quite different. For those mismatched duplexes that showed an acceleration effect with the tag (P13-5’PO$_4$-3FAM/C3-2-mis, P17-5’PO$_4$-3BHQ/C3-2-mis, P9-5’PO$_4$-15FAM/C3-14-mis, P16-5’PO$_4$-3digoxin-15FAM/C3-14-mis and P19-5’PO$_4$-3NH$_2$-15FAM/C3-14-mis), the $K_m$ values all notably increased in comparison with those of the PM-duplexes. Whereas for mismatched duplexes that had no acceleration effects with the tag (P9-5’PO$_4$-15FAM/C3-2-mis, P16-5’PO$_4$-3digoxin-15FAM/C3-2-mis and P19-5’PO$_4$-3NH$_2$-15FAM/C3-2-mis), the $K_m$ values only slightly increased. These results suggested that the apparent binding affinity had changed for those mismatched duplexes which were digested faster than the PM-duplexes. The mismatched duplexes that showed an acceleration effect with the tag had larger $k_{cat}$ values than the PM-duplexes, while those having no acceleration effects with the tag displayed almost the same $k_{cat}$ values as those of the PM-duplexes. These results were consistent with the digestion rates obtained by the fluorescence analysis. The $k_{cat}/K_m$ values of P13-5’PO$_4$-3FAM/C3-PM, P13-5’PO$_4$-3FAM/C3-2-mis, P17-5’PO$_4$-3BHQ/C3-PM and P17-5’PO$_4$-3BHQ/C3-2-mis were very close to each other and the $k_{cat}/K_m$
values of P9-5’PO₄-15FAM/C3-PM, P9-5’PO₄-15FAM/C3-2-mis, P9-5’PO₄-15FAM/C3-14-mis, P16-5’PO₄-3digoxin-15FAM/C3-PM, P16-5’PO₄-3digoxin-15FAM/C3-2-mis, P16-5’PO₄-3digoxin-15FAM/C3-14-mis, P19-5’PO₄-3NH₂-15FAM/C3-PM, P19-5’PO₄-3NH₂-15FAM/C3-2-mis, and P19-5’PO₄-3NH₂-15FAM/C3-14-mis were at similar levels.

The $k_{cat}/K_m$, defined as the specificity constant, is regarded as an approximate measurement of enzyme efficiency. For those mismatched duplexes having an acceleration effect with the tag, the $k_{cat}$ and $K_m$ values increased proportionally compared to the PM-duplexes. This is a characteristic of non-productive binding of the substrates to an enzyme (S1).

According to the Michaelis-Menten equation, the reaction of λ exo and the substrate can be described as equation 1:

\[
E + S \xrightarrow{k_{cat}} ES \xrightarrow{k_{2}} E + P \quad (Eq. 1).
\]

If the substrate binds with λ exo but does not lead to production, it can be described as equation 2:

\[
E + S \xrightarrow{k_e} SE \quad (Eq. 2).
\]

The rate of the reaction without non-productive complexes can be described as the normal form of the Michaelis-Menten equation:

\[
v = \frac{k_{cat}e_0s}{K_m + s} \quad (Eq. 3).
\]

where $e_0$ is the initial concentration of the enzyme and $s$ is the initial concentration of the substrate. While the rate of reaction with non-productive complexes can be described as:

\[
v = \frac{k_2}{1 + \frac{K_m}{K_e}} \frac{k_{cat}e_0s}{K_m + s} \quad (Eq. 4)
\]

Comparing equation 3 and 4, the apparent $K'_m$ and $k'_{cat}$ can be described as:

\[
K'_m = \frac{K_m}{1 + \frac{K_m}{K_e}} \quad (Eq. 5)
\]

and

\[
k'_{cat} = \frac{k_{cat}}{1 + \frac{K_m}{K_e}} \quad (Eq. 6).
\]

Thus, $K'_m/k'_{cat} = K_m/k_{cat}$ and the $K'_m$ and $k'_{cat}$ will decrease proportionally when the non-productive binding occurs.
Figure S8. (a) Polyacrylamide gel electrophoresis of the protein λ exo WT-2 after expression and purification. Lane 1: *E. coli* BL21(AI) with λ exo expression. Lane 2: precipitate of the *E. coli* BL21(AI) after ultrasonication and centrifugation. Lane 3: supernatant of the *E. coli* BL21(AI) after ultrasonication and centrifugation. Lane 4: eluate after the supernatant was loaded on the Ni-NTA column. Lane 5: elution with Buffer A (50mM NaH$_2$PO$_4$, 300mM NaCl, pH 8.0). Lane 6: elution with Buffer B (50mM NaH$_2$PO$_4$, 300mM NaCl, 20 mM imidazole, pH 7.0). Lane 7: elution with Buffer B for the second time. Lane 8: elution with Buffer C (50mM NaH$_2$PO$_4$, 300mM NaCl, 200mM imidazole, pH 7.0). Lane 9: λ exo purchased from NEB (WT-1). Lane 10: protein marker. (b) Fluorescence intensity responses of P12-5’PO$_4$-3FAM/C3-PM digested by different λ exo variants.
Scheme S2. (a) Schematic depiction of the non-productive binding complex (Complex 1) and the productive binding complex (Complex 2). λ exo was drawn as a toroid in green. (b) Schematic illustration of the interactions between λ exo and PM or 2-mis DNA duplexes with the internal label located at different positions. The relative digestion rates of different complexes were compared.

According to the reported digestion mechanism of 5’-PO₄ dsDNA by λ exo, we inferred that the initial binding complex formed by the enzyme and the DNA substrate was a natural non-productive binding complex (Complex 1), in which the 5’ end of the duplex has not been unwound so that no production occurs. The binding complex in the second and third step should be a natural productive binding complex (Complex 2), in which the 5’ end of the substrate strand has entered the reaction center of λ exo and the phosphodiester bond is ready to be hydrolysed (Scheme S2a). During the digestion process, Complex 1 needs to step over the energy penalty of unwinding the first two nucleotides at the 5’ end to transform to Complex 2. Recently there was a hypothesis that when digested by one subunit of λ exo, the 5’ end of the substrate strand might occasionally slip out and rebind to another subunit, which consumed significantly more time (S2). This hypothesis gave us a hint that Complex 2 might also transform back to Complex 1. So there was an equilibrium between Complex 1 and Complex 2, and the non-productive binding complex...
always accounted for a fraction. In the form of Complex 2, the PM-duplex has already been unwound just as the mismatched duplex, so the difference between the mismatched duplex and PM-duplex should mainly come from the unwinding period when Complex 1 is transformed to Complex 2.

As show in Scheme S2b, for a duplex containing a mismatched base pair but without an adjacent tag at 3’ side (Substrate 1), though the 5’ end was roughly unwound and thus reduced the energy penalty, the unpaired end was free and might not bind to the reaction center tightly. The net effect of the mismatch itself only had a little influence on the binding complex states, so it could only slightly increase the digestion rate in comparison with the corresponding PM-duplex (Substrate 2), such as the case of the P9-5’PO₄-15FAM/C3-2-mis duplex. When a FAM (or ROX or BHQ-1) tag was labeled at the 3’ side of the mismatched base pair (in Substrate 3), the tag and mismatched base pair might have a synergistic effect which could promote the transformation from Complex 1 to Complex 2. The presence of a mismatched base pair made the 5’ end of the substrate more flexible than the PM substrate. Thus, when the mismatched duplex forms Complex 2 with λ exo, it was easier for the FAM (or ROX or BHQ-1) tag to get close to the π-π stacking interaction area of the active subunit of λ exo. These additional interactions might help stabilize the Complex 2, resulting in higher yield of the products. For a duplex with an internal tag but without a mismatched base at the 5’ upstream position of the labeled base (Substrate 4), it would first form Complex 1 with λ exo. Then the resulting complex went through the unwinding process and became Complex 2; thus, the digestion rate was slower than that of Substrate 2.

The interactions between the labeled tags and λ exo were largely affected by the structures of the tags. Some of the tested tags (such as digoxin, biotin, and amino) did not interact with the amino acid residues in the active center of λ exo, so the presence of a mismatched base pair at the 5’ side of these tags had little effect on stabilization of the productive Complex 2; and the acceleration effect on the digestion reaction was the same as that of Substrate 1. It’s worth to mention that above-discussed properties were not limited to the tags labeled at Position 3 which was very close to the 5’ end. They had also been observed for P14-5’PO₄-10FAM/C4-9-mis and P15-5’PO₄-56FAM/C5-55-mis duplexes, indicating that a combination of the mismatched base pair and the tag with a conjugated structure at 3’ side not only influenced the initial steps of the digestion reaction but also affected the digestion process during the reactions.

The above mechanism could also explain the kinetic constants obtained for the substrates with blunt and recessed 5’ phosphorylated ends. It was reported that the $K_m/k_{cat}$ ratios of these two kinds of substrates were the same, but the $K_m$ and $k_{cat}$ values of the substrates with recessed 5’ phosphorylated end (recessed 5’-PO₄ substrate) were all larger than those of the substrate with blunt 5’ phosphorylated end (blunt 5’-PO₄ substrate) (S3). That was most likely because the recessed 5’-PO₄ substrate had a 3’ overhang which confined the recession of λ exo and promoted the transformation to Complex 2. Compared to the blunt 5’-PO₄ substrate, the recessed 5’-PO₄ substrate formed more productive-binding complex with λ exo, thus resulting in the same $K_m/k_{cat}$ and larger $K_m$ and $k_{cat}$ values.
References

S1. Fersht A (1984) Enzyme Structure and Mechanism (Second Edition) (W. H. Freeman).

S2. Pan X, Yan J, Patel A, Wysocki VH, & Bell CE (2015) Mutant poisoning demonstrates a nonsequential mechanism for digestion of double-stranded DNA by lambda exonuclease trimers. Biochemistry-USS 54(3):942-951.

S3. Mitsis PG & Kwagh JG (1999) Characterization of the interaction of lambda exonuclease with the ends of DNA. Nucleic Acids Research 27(15):3057-3063.