Electronic Supplementary Material (ESI) for Analyst.

Supplementary Information

Modified Cytosines versus Cytosine in a DNA Polymerase: Retrieving Thermodynamic and Kinetic Constants at the Single Molecule Level

Ji Yoon Lee[1] and Joon Won Park[1,2]*

[1]Department of Chemistry, Pohang University of Science and Technology, 77 Cheongam-Ro, Nam-Gu, Pohang 37673, Republic of Korea

[2]Institute of Convergence Science, Yonsei University, 50 Yonsei-ro Seodaemun-gu, Seoul, 03722, Republic of Korea

* Corresponding author: Joon Won Park (E-mail: jwpark@postech.ac.kr)

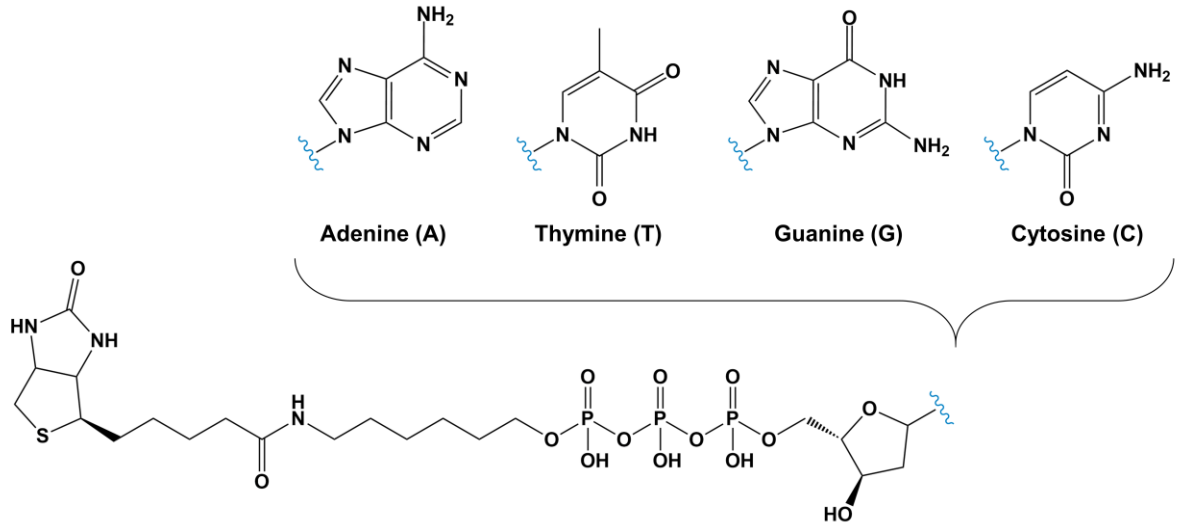


Figure S1. Structure of dNTPs (deoxynucleotide triphosphates) tethered to biotin. The dNTP (A, T, G, or C) analog was used for the substrate immobilization (Purity: ≥ 95% (HPLC)).

Name	Sequence (5’ → 3’)	Length
Primer	CCG TGC TCA TCA TAG TAA CC	20
Template (C)	TTT TTT TTT TTT TTT TTT TCG GTT ACT ATG ATG AGC ACG GGC GGC ACT GA	50
Template (5mC)	TTT TTT TTT TTT TTT TTT T(5mC)G GTT ACT ATG ATG AGC ACG GGC GGC ACT GA	50
Template (5caC)	TTT TTT TTT TTT TTT TTT T(5caC)G GTT ACT ATG ATG AGC ACG GGC GGC ACT GA	50

Table S1. Sequence of a primer (20-mer) and template DNAs (50-mer) used in the experiment. In the template sequences, the primer-binding regions are indicated in blue. Cytosine (C), 5-methylcytosine (5mC), and 5-carboxylcytosine (5caC) are in red.

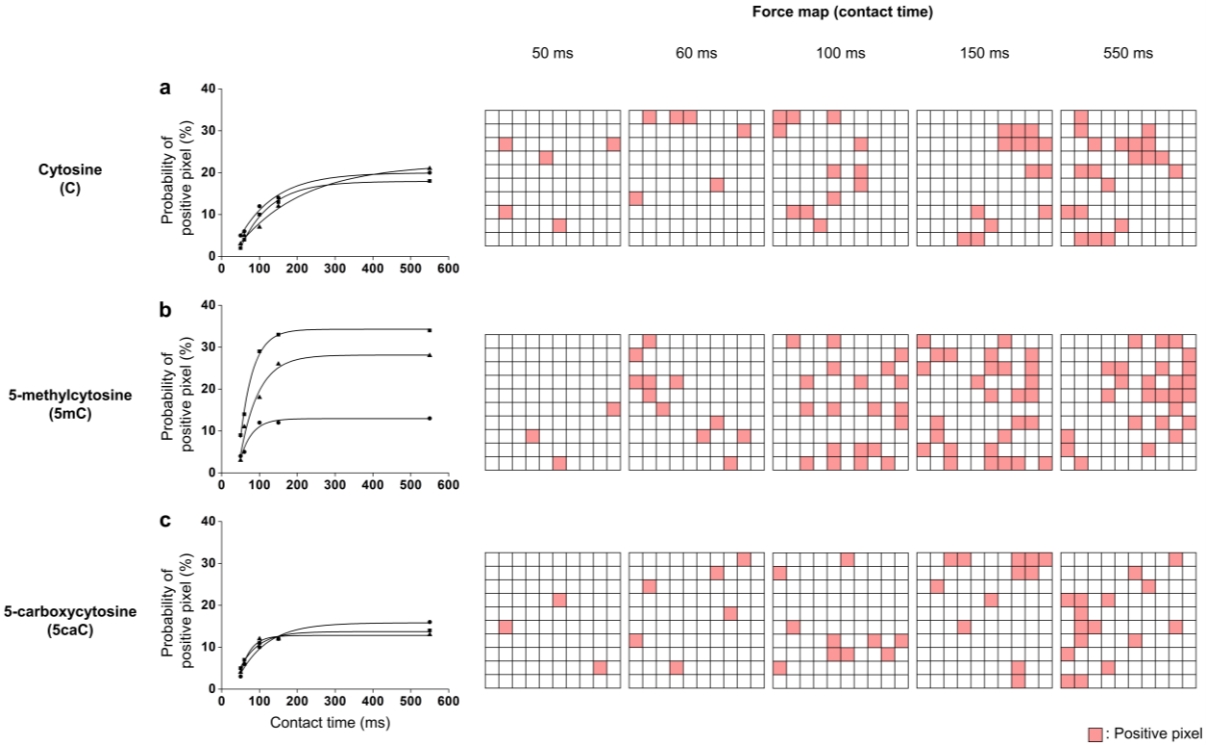


Figure S2. Change in the specific binding probability between the p125/DNA duplex complex and the immobilized dGTP depending on the contact time. We collected force-distance curves five times per pixel in the area of 1.0 × 1.0 μm^2 (the total number of pixels: 10**0**), and a pixel showing the specific force-distance curves no less than two times was set as a 'positive pixel' (red pixels in the above maps). The positive pixel probability and the force map depend on the contact time for C (a), 5mC (b), and 5caC (c). The probability was calculated by dividing the number of positive pixels by 100. For the above experiment, the applied force was 100 pN, the initial contact time was found to be 50 ms, and additional contact times of 10, 50, 100, or 500 ms were given. Therefore, the total contact times were 50, 60, 100, 150, and 550 ms.

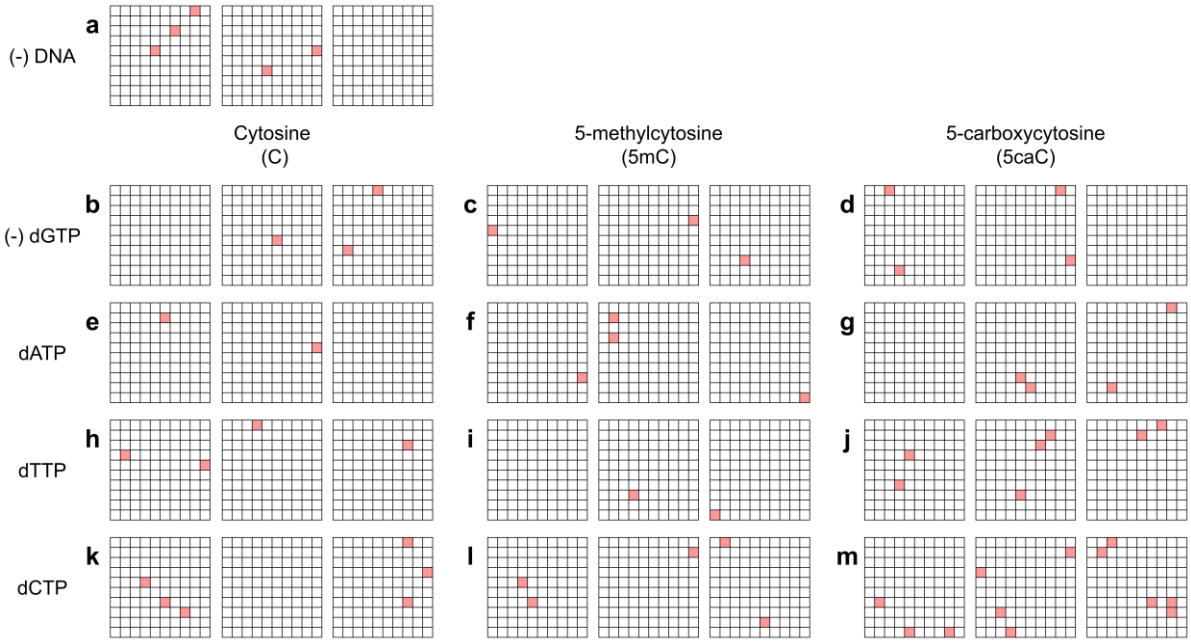


Figure S3. Control experiments to evaluate the reliability of the specific binding interaction between the p125/DNA duplex complex and the immobilized dGTP. We collected force-distance curves five times per pixel in the area of 1.0 × 1.0 μm^2 (total number of the pixels: 100). A pixel showing the specific force-distance curves no less than two times was set as a 'positive one' (colored red in the above maps); (a), interaction between p125 and the immobilized dGTP in the absence of the DNA duplex, (b-d), interaction between the p125/DNA duplex complex in the absence of the immobilized dGTP, (e-g), interaction between the p125/DNA duplex complex and the immobilized dATP, (h-j), interaction between the p125/DNA duplex complex and the immobilized dTTP, (k-m), interaction between the p125/DNA duplex complex and the immobilized dCTP. The maps on the left (b,e,h, and k) are for C, the maps on the middle (c, f, i, and l) are for 5mC, and the maps on the right (d, g, j, and m) are for 5caC.

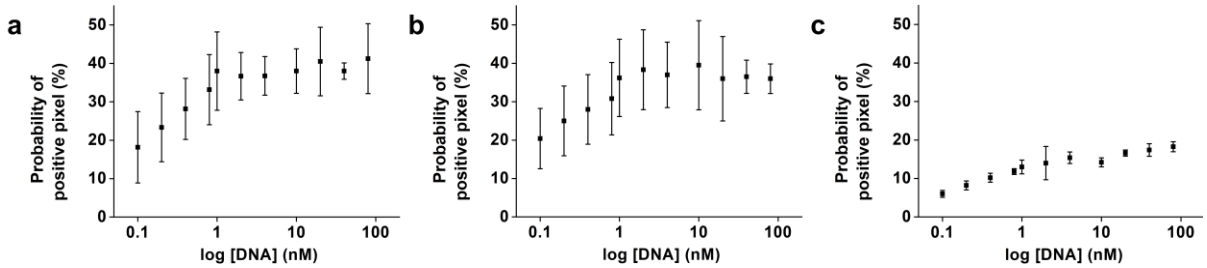
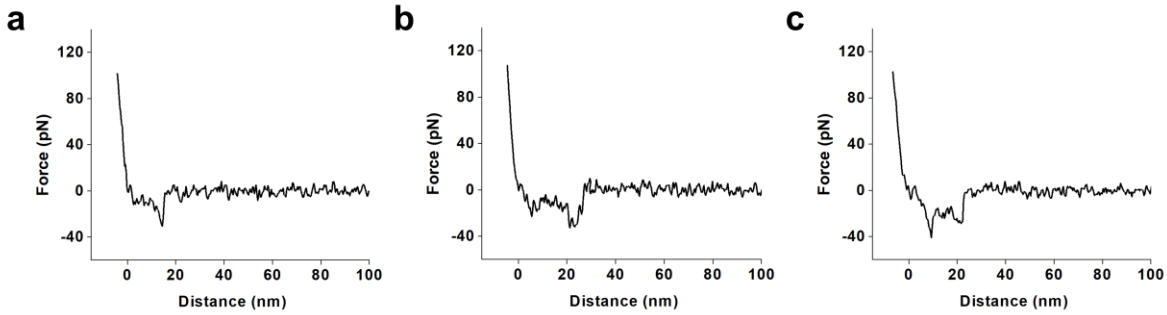


Figure S4. Change in the number of positive pixels according to the added DNA duplex concentration. The number of positive pixels was counted by adding corresponding DNA to the reaction solution. The DNA concentration was changed from 0.1 nM to 80 nM, and was expressed on a log scale. (a) for C, (b) for 5mC, and (c) for 5caC.



d

	K_{DNA}			K_{dGTP}		
	C (n=281)	**5mC** (n=452)	**5caC** (n=114)	**C** (n=296)	**5mC** (n=263)	**5caC** (n=273)
Single peak (%)	**83.6**	**84.7**	**82.4**	**80.7**	**84.4**	**96.3**
Double peaks (%)	**11.7**	**10.6**	**15.8**	**15.2**	**8.8**	**3.3**
Multiple peaks (%)	**4.6**	**4.6**	**1.8**	**4.1**	**6.8**	**0.4**

Figure S5. Analysis of the observed specific unbinding force curve. The peaks of the force curves were divided into three types: single (a), double (b), and multiple (c), and the ratio of each peak was analyzed. The force curves of the positive pixels in Figures 3b and 4b were analyzed. For Fig. 3b, 712 single peaks were observed out of the total peaks (n=847) (84%), and for Fig. 4b, 724 single peaks were observed out of the total peaks (n=832) (87%). (d) Percentages of the peak type for the unbinding force curves of each set. Out of the total peaks (n=1,679), 1,436 single peaks were observed (86%).