

## **Supplementary Information for**

### **Structural and functional insight into mismatch extension by human DNA polymerase $\alpha$**

Andrey G. Baranovskiy, Nigar D. Babayeva, Alisa E. Lisova, Lucia M. Morstatd, Tahir H. Tahirov

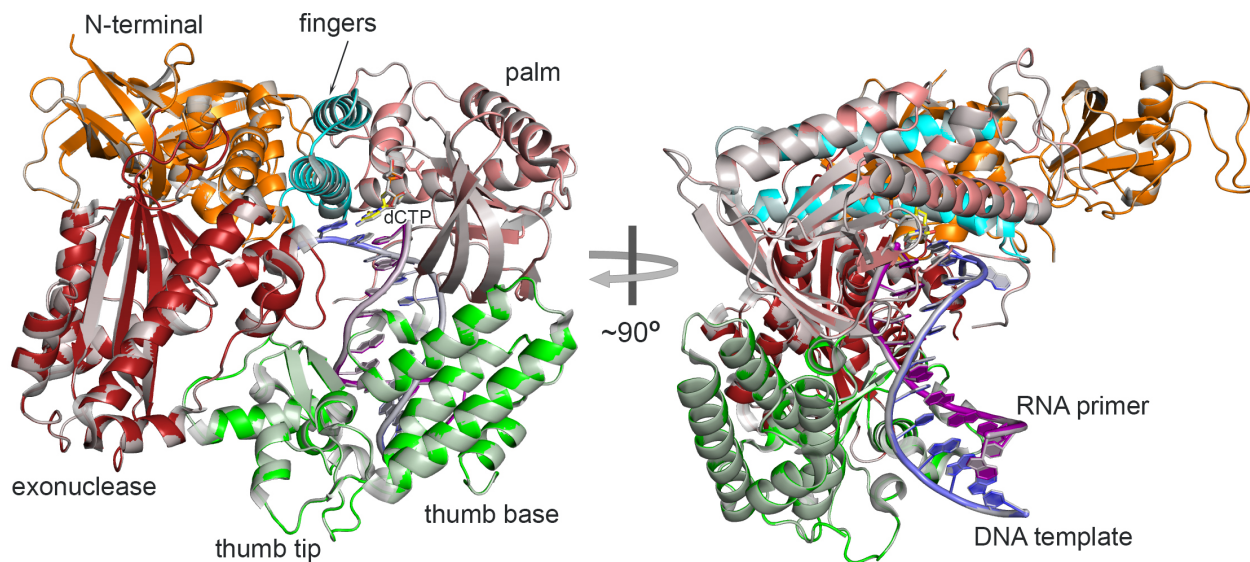
Tahir H. Tahirov

Email: [ttahirov@unmc.edu](mailto:ttahirov@unmc.edu)

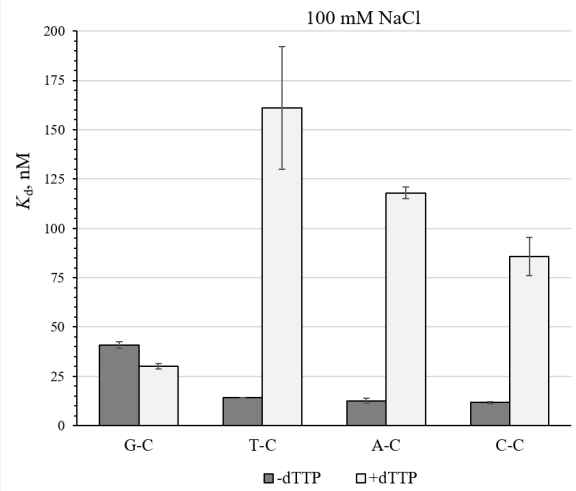
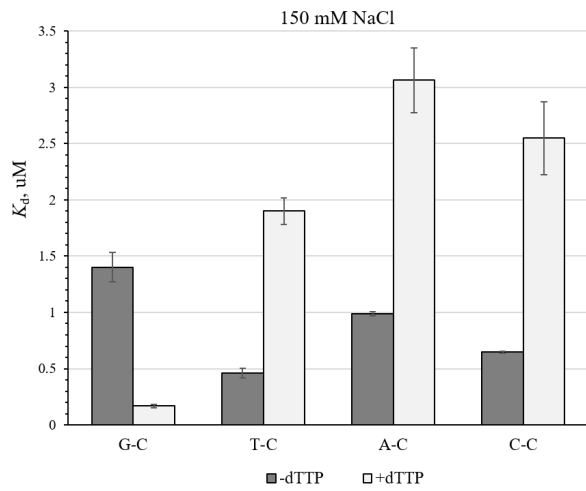
#### **This PDF file includes:**

Figures S1 to S4

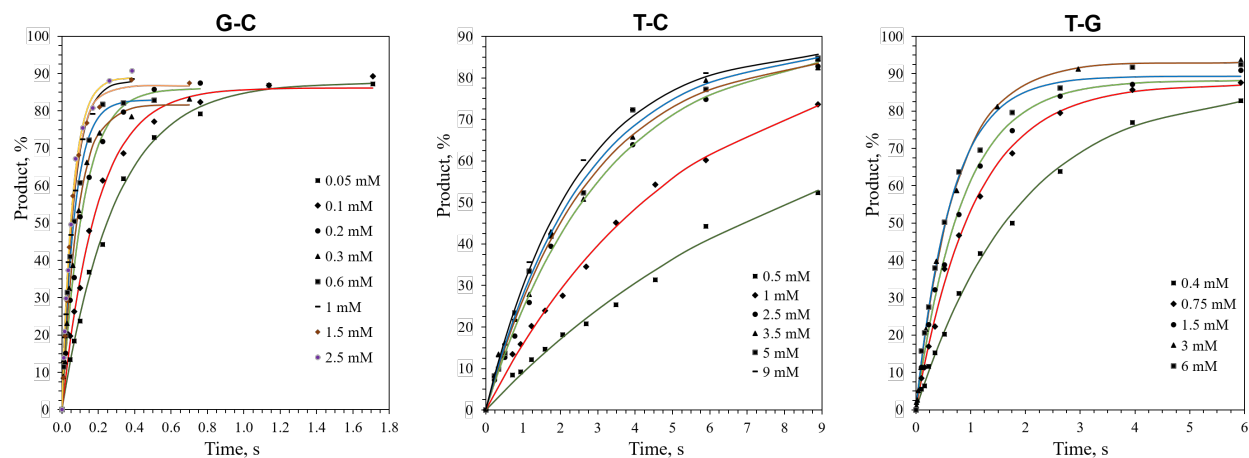
Tables S1 to S2



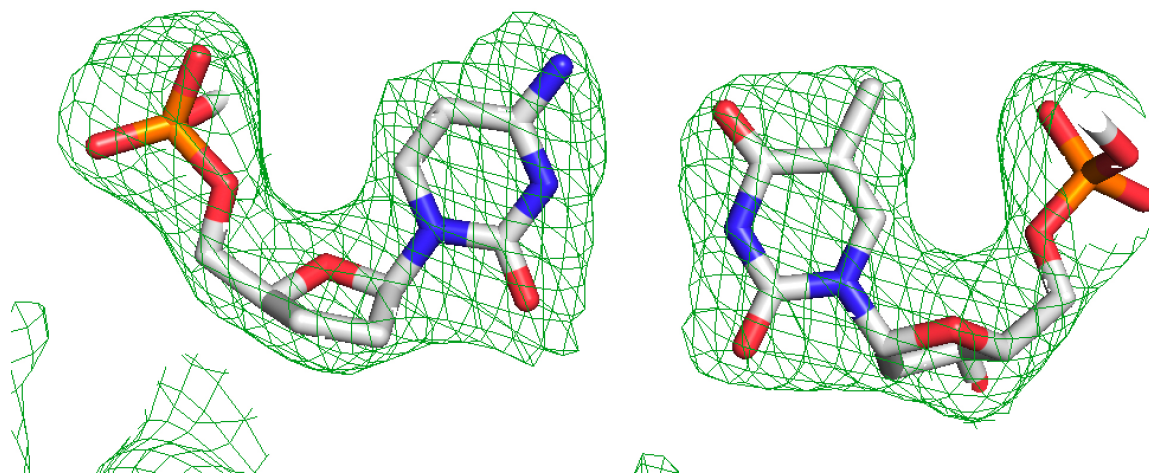
**Fig. S1. Overall view of aligned human Pol $\alpha_{CD}$  ternary complexes containing a matched and mismatched template:primer.** In the complex with a mismatched template:primer, the carbons of dCTP, DNA template, and RNA primer are colored yellow, marine, and purple, respectively. Subdomains of Pol $\alpha_{CD}$ : N-terminal, exonuclease, fingers, palm, and thumb, are colored orange, red, cyan, salmon, and green, respectively. In the complex containing a correct template:primer (PDB code 4QCL), all molecules are colored gray with 10% transparency. The complexes are aligned with RMSD of 0.075 Å for 850 C $\alpha$  atoms.



**Fig. S2. Effect of salt, incoming dNTP, and a 3'-terminal mismatch on human Pol $\alpha$  affinity to DNA.** Data are presented as a bar graph showing the mean  $\pm$ SD. The 3'-terminal base-pair or mismatch is indicated below the bar.



**Fig. S3. Pre-steady-state kinetics of cognate and mismatched primer extension at different dTTP concentration.** Percent of extended primer was plotted against time and the data were fit to a single-exponential equation. The 3'-terminal base-pair or mispair is indicated above the graphs.



**Fig. S4.** The  $2F_o - F_c$  Fourier map for the T-C mismatch at contour level of  $1.5 \sigma$ .

**Table S1. Summary of data collection and refinement statistics**

| <b>Data collection</b>            |                                    |
|-----------------------------------|------------------------------------|
| Space group                       | $P 3_221$                          |
| Cell dimensions: $a, c$ (Å)       | 140.285, 181.066                   |
| Resolution (Å)                    | 50 – 2.9 (2.95 – 2.9) <sup>a</sup> |
| $R_{\text{merge}}$                | 0.079 (0.437)                      |
| $I/\sigma I$                      | 29.3 (3.2)                         |
| Completeness (%)                  | 98.7 (97.7)                        |
| Unique reflections                | 45582 (2211)                       |
| Redundancy                        | 5.5 (3.7)                          |
| <b>Refinement</b>                 |                                    |
| Resolution (Å)                    | 44.51 – 2.9 (3.08 – 2.9)           |
| No. reflections                   | 45573 (6973)                       |
| $R_{\text{work}}/R_{\text{free}}$ | 0.195/0.228 (0.316/ 0.324)         |
| No. atoms                         |                                    |
| Protein                           | 6948                               |
| DNA/RNA                           | 494                                |
| Ligand/ion                        | 55                                 |
| Water                             | 110                                |
| Mean B-factors (Å <sup>2</sup> )  | 51.1                               |
| R.m.s deviations                  |                                    |
| Bond lengths (Å)                  | 0.008                              |
| Bond angles (°)                   | 1.4                                |
| Ramachandran plot (%)             |                                    |
| Core                              | 87.5                               |
| Allowed                           | 11.1                               |
| Generously allowed                | 1.3                                |
| Disallowed                        | 0.1                                |

<sup>a</sup> Numbers in parentheses refer to the highest-resolution shell.

**Table S2. Oligonucleotides used in this study**

| Sequence                                  | Description                  | Application     | Length |
|---|------------------------------|-----------------|--------|
| 5'-AATGTTTCTAGGCAGCTCGGAGTCC <sup>a</sup> | Template                     | kinetic studies | 25     |
| 5'-AATGTTTCTATGCAGCTCGGAGTCC              | Template (T-C, T-G mismatch) |                 | 25     |
| 5'-/Cy3/GGACTCCGAGCTGCC                   | Primer                       |                 | 15     |
| 5'-/Cy5/GGACTCCGAGCTGCG                   | Primer (T-G mismatch)        |                 | 15     |
| 5'-/BiotinTEG/AATACATAAGCGCTCCAGGCAAT     | Template                     | binding studies | 23     |
| 5'-/BiotinTEG/AATACATAATCGCTCCAGGCAAT     | Template (T-C mismatch)      |                 | 23     |
| 5'-/BiotinTEG/AATACATAACCGCTCCAGGCAAT     | Template (C-C mismatch)      |                 | 23     |
| 5'-/BiotinTEG/AATACATAAACGCTCCAGGCAAT     | Template (A-C mismatch)      |                 | 23     |
| 5'-GCCTGGAGCG/ddC/                        | primer                       |                 | 11     |
| 5'-ATAGTCGCTCCAGGC                        | template                     | crystallization | 15     |
| 5'-r(GCCUGGAGCG)/ddC/                     | primer                       |                 | 11     |

<sup>a</sup> The region complementary to a primer is underlined.