Supplementary Information

Transcriptional processing of an unnatural base pair by eukaryotic RNA polymerase II

Juntaek Oh\textsuperscript{1}, Ji Shin\textsuperscript{1,7,\#}, Ilona Christy Unarta\textsuperscript{2,\#}, Wei Wang\textsuperscript{1,8}, Aaron W. Feldman\textsuperscript{3}, Rebekah J. Karadeema\textsuperscript{3}, Liang Xu\textsuperscript{1,9}, Jun Xu\textsuperscript{1}, Jenny Chong\textsuperscript{1}, Ramanarayanan Krishnamurthy\textsuperscript{3}, Xuhui Huang\textsuperscript{2}, Floyd E. Romesberg\textsuperscript{4}, Dong Wang\textsuperscript{1,5,6,*}

\textsuperscript{1}Division of Pharmaceutical Sciences, Skaggs School of Pharmacy & Pharmaceutical Sciences; University of California, San Diego, La Jolla, California 92093, United States
\textsuperscript{2}Department of Chemistry, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
\textsuperscript{3}Department of Chemistry, The Scripps Research Institute, La Jolla, California 92037, United States
\textsuperscript{4}Synthorx, a Sanofi Company, La Jolla, CA 92037, United States
\textsuperscript{5}Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, California 92093, United States
\textsuperscript{6}Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, United States
\textsuperscript{7}Present Address: Department of Molecular Microbiology, Center for Advanced Laboratory Medicine (CALM), University of California, San Diego, La Jolla, California, 92093
\textsuperscript{8}Present Address: Advanced Medical Research Institute, Shandong University, Jinan 250012, China
\textsuperscript{9}Present Address: Department of Chemistry, Sun Yat-Sen University, Guangzhou 510275, China

* To whom correspondence should be addressed. Tel: +1 858 822 5561; Fax: +1 858 822 1953; Email: dongwang@ucsd.edu, DW: https://orcid.org/0000-0002-2829-1546

\# These authors contributed equally.

This file contains Supplementary Tables 1-3
Supplementary Tables

**Supplementary Table 1.** Kinetic parameters of incorporation and selectivity of unnatural nucleotide triphosphate over natural nucleotide triphosphate incorporation. The observed rate constant $K_{\text{obs}}$ (hr $^{-1}$) values of single nucleotide incorporation (at 1 mM) were calculated from one-phase association regression by using Prism 6 (Fig 1c). All experiments were performed independently three times. Results in Table S1 are shown as means with standard deviation errors. Discrimination power of rNaMTP in dTPT3 template against natural NTPs is defined as ratios of the $K_{\text{obs}}$ of an unnatural nucleotide triphosphate over the $K_{\text{obs}}$ of a natural nucleotide triphosphate at same concentration. For example, the discrimination of rNaMTP over ATP is calculated by $K_{\text{obs, rNaMTP}} / K_{\text{obs, ATP}}$.

| Template | Substrate | $K_{\text{obs}}$ (hr $^{-1}$) | Discrimination |
|----------|-----------|-----------------------------|----------------|
| dTPT3    | rNaMTP    | 1.3 ± 0.1                   | rNaMTP/ rNaMTP | 1  |
|          | ATP       | 0.052 ± 0.004               | rNaMTP/ATP     | 25 |
|          | UTP       | 0.047 ± 0.006               | rNaMTP/UTP     | 28 |
|          | GTP       | 0.041 ± 0.006               | rNaMTP/GTP     | 32 |
|          | CTP       | 0.002 ± 0.014               | rNaMTP/CTP     | 650|
| dNaM     | rTPT3TP   | 1.0 ± 0.1                   | rTPT3TP/ rTPT3TP | 1  |
|          | ATP       | 0.73 ± 0.05                 | rTPT3TP/ATP    | 1.4|
|          | UTP       | 0.11 ± 0.09                 | rTPT3TP/UTP    | 9.1|
|          | GTP       | 0.15 ± 0.02                 | rTPT3TP/GTP    | 6.7|
|          | CTP       | 0.072 ± 0.003               | rTPT3TP/CTP    | 14 |
Supplementary Table 2. Crystallographic data collection and refinement statistics.

|                  | dTPT3_apo | dTPT3_rNaMTP | dTPT3_rNaM |
|------------------|-----------|--------------|------------|
| PDB ID           | 7KED      | 7KEE         | 7KEF       |
| **Data collection**                                      |
| Space group      | C 1 2 1   |              |            |
| Cell dimensions  |           |              |            |
| \(a, b, c\) (Å) | 169.9 223.0 194.7 | 169.5 222.7 194.7 | 167.7 221.7 193.3 |
| \(\beta\) (°, \(\alpha, \gamma = 90°\)) | 101.6 | 101.9 | 101.8 |
| Resolution (Å)   | 71.2 – 3.6 (3.73 – 3.6)\(^a\) | 49.2 - 3.45 (3.57 - 3.45) | 48.1 - 3.89 (4.03 -3.89) |
| \(R_{merge}\)   | 0.27 (1.29) | 0.25 (1.27) | 0.35 (1.65) |
| \(I / \sigma I\) | 2.3 (0.8) | 4.24 (0.75) | 2.51 (0.45) |
| CC1/2\(^b\)     | 0.89 (0.32) | 0.95 (0.30) | 0.90 (0.32) |
| Completeness (%) | 96.9 (97.0) | 99.8 (99.3) | 98.6 (88.5) |
| Redundancy       | 3.2 (3.1) | 2.0 (2.0) | 2.0 (2.0) |

**Refinement**

|                |           |              |            |
|----------------|-----------|--------------|------------|
| No. reflections| 79871 (7823) | 92635 (9202) | 62514 (5590) |
| \(R_{work} / R_{free}\) | 0.266 / 0.306 | 0.259 / 0.302 | 0.268 / 0.325 |
| No. atoms      |           |              |            |
| Protein        | 28907     | 28997        | 29169      |
| Ligand/ion     | 9         | 42           | 33         |
| **B-factors**  |           |              |            |
| Protein        | 81.6      | 88.9         | 115.3      |
| Ligand/ion     | 95.3      | 131.7        | 116.9      |
| **R.m.s. deviations** |
| Bond lengths (Å) | 0.005       | 0.006        | 0.005       |
| Bond angles (°) | 1.04        | 1.22         | 1.05        |

\(^a\)Values in parentheses are for the highest-resolution shell.

\(^b\)CC1/2, half correlation coefficient is used to define high resolution cutoff.
**Supplementary Table 3.** Number of atoms, water molecules, sodium and chloride ions for each type of MD simulation.

| Type of Simulation | Total No. of Atoms | No. of water molecules | No. of Na\(^+\) ions | No. of Cl\(^-\) ions |
|-------------------|--------------------|------------------------|-----------------------|----------------------|
| dTPT3             |                    |                        |                       |                      |
| rNaMTP            | 440,208            | 126,375                | 505                   | 400                  |
| ATP               | 440,204            | 126,376                | 505                   | 400                  |
| GTP               | 440,205            | 126,376                | 505                   | 400                  |
| UTP               | 440,201            | 126,376                | 505                   | 400                  |
| CTP               | 440,202            | 126,376                | 505                   | 400                  |
| dNaM              |                    |                        |                       |                      |
| rTPT3TP           | 440,673            | 126,530                | 505                   | 400                  |
| ATP               | 440,673            | 126,530                | 505                   | 400                  |