Supplementary data

Interplay between the trigger loop and the F loop during RNA polymerase catalysis

Nataliya Miropolskaya, Daria Esyunina, Saulius Klimašauskas, Vadim Nikiforov, Irina Artsimovitch, Andrey Kulbachinskiy

Supplementary Text

Calculation of thermodynamic parameters of nucleotide addition by bacterial RNAP

The reaction of nucleotide addition by RNAP can be described by the following scheme:

TEC_n + NTP ⇌ TEC_n-NTP ⇌ TEC_n-NTP# → TEC_n+1 +PPi

where TEC_n-NTP is the ground state and TEC_n-NTP# is the transition state TEC-substrate complex (also called activated complex).

In accordance to the transition state theory of catalysis, the catalytic rate is then determined by the equation

\[ k_{obs} = k_B T/\hbar e^{-\Delta G#/RT} = k_B T/\hbar e^{-\Delta H#/RT-\Delta S#/R} \]  \[1\]

where \( k_B \) is the Boltzmann constant, \( T \) is the absolute temperature, \( \hbar \) is the Planck constant, \( R \) is the universal gas constant, and \( \Delta G#, \Delta H#, \) and \( T\Delta S# \) are the free energy, enthalpy and entropy of activation that correspond to the differences in these parameters between the activated and the ground state TEC-NTP complexes.

On the other hand, the temperature dependence of the catalytic rate is given by the Arrhenius equation:

\[ k_{obs} = A e^{-E_a/(RT)} \]  \[2\]

where \( E_a \) is the activation energy and \( A \) is a pre-exponential factor.

Hence, the logarithm of \( k_{obs} \) linearly depends on inverse temperature, and \( E_a \) determines the slope of the linear dependence:

\[ \ln k_{obs} = \ln A - E_a/(RT) \]  \[3\]

The \( \Delta G#, \Delta H#, \) and \( T\Delta S# \) values for the transition between the ground state and the activated enzyme-substrate complex can be then calculated using the equations which follow from equations [1] and [2] (see (23) for details):

\[ \Delta G# = RT(\ln(k_B T/\hbar) - \ln k_{obs}) \]  \[4\]

\[ \Delta H# = E_a - RT \]  \[5\]

\[ T\Delta S# = \Delta H# - \Delta G# \]  \[6\]
The higher activity of mesophilic enzymes in comparison with their thermophilic counterparts at moderate and low temperatures is usually explained by more flexible structures of their active centres, thus facilitating conformational changes required for catalysis. In thermodynamics terms, the increased flexibility of the ground state enzyme-substrate complex in mesophilic enzymes results in lowering of the heat energy required for the restructuring of the active centre during activation, *i.e.* in lowering of the activation enthalpy. On the other hand, the increased flexibility also increases the entropy of the ground state, thus making the reaction less entropically favorable (e.g., (23,25)). However, as follows from the right part of equation [1], the contribution of the $\Delta H^\#$ term to the $k_{obs}$ value is larger at lower temperatures. Thus, the lowering of the $\Delta H^\#$ value allows the mesophilic enzymes, and in particular, $Dra$ RNA, to retain a significant level of activity at lower temperatures, despite the compensatory decrease in the activation entropy.

It should be noted that the absolute values of the thermodynamic parameters of the activation are usually difficult to interpret from the structural point of view due to a complex nature of the enthalpy and entropy changes in enzymatic reactions and due to the limitations of the transition state theory (23). However, analysis of changes in these parameters between different enzyme variants is more reliable and can be used for structure-based interpretation of the reaction mechanism (23,25,33,34). Thus, we used the $\Delta G^\#$, $\Delta H^\#$ and $T\Delta S^\#$ values to compare catalysis by analyzed RNA variants with wild-type Taq RNA.

**Changes in the TEC structure associated with the activated complex formation**

Highly positive values of the activation enthalpy ($\Delta H^\#$) and activation entropy ($T\Delta S^\#$) observed for wild-type Taq RNA suggest that transition from the ground state to the activated state during nucleotide addition by this RNA is enthalpically unfavorable but entropically favorable, implying a significant decrease in the structural ordering of the TEC during nucleotide addition. Available TEC structures suggest a possible interpretation of the activation parameters characteristic to wild-type and mutant RNA variants. The *Thermus thermophilus* ($Tth$) TEC structure with the fully folded TL that encloses the NTP substrate within the active site presumably corresponds to the activated RNA-substrate complex (1). Thus, the activated complex formation is likely connected to the TL folding and, in the activated complex, the FL harbours the tip of the folded TL in a protein pocket (Figure 1B, Figure 5 and Figure S1A). While the exact structure of the ground state complex remains unknown, it is likely related to the Stl-trapped TEC structure with a partially unfolded TL and NTP in the preinsertion site (Figure S1B). We also speculate that the unfolded TL may have a limited conformational mobility during transcription elongation and may interact with the FL, as observed in some *Saccharomyces cerevisiae* RNAPII structures (3,10), thus requiring restructuring during formation of the transition state complex. Comparison of the TEC-NTP and TEC-NTP-Stl complexes reveals that the TL folding is also accompanied by disordering of the RNA contacts with the downstream DNA. These changes include opening of the downstream $\beta$ pincer, a motion likely transmitted from the active site by the $\beta$ Fork2 element that contacts both the folded TL and the FL (Figure 5 and Figure S1) (1). In addition, the $\beta'$ Jaw domain, which interacts with the fully open TL in some RNA structures (in $Tth$ holoenzyme, Figure S1C; (29,31)), becomes disordered in the activated complex, perhaps in concert with the TL folding (Figure 5). Thus, large activation enthalpy and entropy values characteristic to Taq RNA are likely connected to the TL folding and the accompanying changes in the downstream RNA region. Other structural changes, not revealed in the available TEC structures, such as changes in the interactions with the solvent during reaction, may also significantly contribute to the catalytic parameters of nucleotide addition. These rearrangements should require breaking of many heat-driven interactions, resulting in high activation enthalpy, but are entropically favorable.
The TL deletion likely disrupts structural reorganization of the TEC associated with the TL folding, resulting in a dramatic unfavorable decrease in the activation entropy, with concomitant decrease in the activation enthalpy. In contrast to wild-type RNAP, catalysis by the ΔTL RNAP cannot be efficiently activated by temperature and, as a result, the TL deletion much more strongly affects catalysis at high temperatures, more close to the temperature optimum of transcription of Taq RNAP, than at low temperatures. This may in part explain why a similar TL deletion in a thermophilic archaean Pyrococcus furiosus RNAP had only a moderate effect on the rate of nucleotide addition by this RNAP, when measured at suboptimal temperatures (35). Thus, the TL is a specialized RNAP element that allows to dramatically accelerate catalysis in a temperature-dependent manner. The observed effects of temperature on the TL-dependent catalysis should be taken into account in future studies of the catalytic mechanisms of RNAP.

Supplementary references

34. Collins, T., Meuwis, M.A., Gerday, C. and Feller, G. (2003) Activity, stability and flexibility in glycosidases adapted to extreme thermal environments. J Mol Biol, 328, 419-428.

35. Fouqueau, T., Zeller, M.E., Cheung, A.C., Cramer, P. and Thomm, M. (2013) The RNA polymerase trigger loop functions in all three phases of the transcription cycle. Nucleic Acids Res, doi:10.1093/nar/gkt1433.
Supplementary Tables

Table S1. The rates of nucleotide addition by wild-type and mutant RNAPs at different temperatures. All \(k_{\text{obs}}\) values were measured at 1 mM UTP concentration; Stl was added to 100 \(\mu\)g/ml.

| RNAP  | \(k_{\text{obs}}\) (s\(^{-1}\)) | UTPL | | UTPL+Stl |
|-------|-------------------------------|------|---|-----------------|
|       | 10 °C                         | 20 °C | 30 °C | 40 °C | 20 °C | 40 °C |
| Dra   | 19.7 ± 5.2                    | 83.9 ± 30.6\(^*\) | 214 ± 50.1 | 504 ± 188\(^*\) | 849 | 321 | 33.2 | 10.1 |
| Taq WT| 0.0232 ± 0.0111               | 0.261 ± 0.021\(^*\) | 6.4 ± 3.6 | 50.0 ± 13.1\(^*\) | 0.00306 ± 0.00043 | 0.0814 ± 0.0112 |
|       | 1                             | 1     | 1     | 1     | 1     |
| F-Dra | 0.76 ± 0.13                   | 9.9 ± 2.3 | 52.9 ± 20.8 | 281 ± 185\(^*\) | 32.8 | 41.6 | 8.2 | 5.6 |
| Q1046A| 0.15 ± 0.06                   | 0.82 ± 0.18\(^*\) | 10.0 ± 1.3 | 82.8 ± 2.7\(^*\) | 6.3 | 3.1 | 1.6 | 1.7 |
| ΔFL  | 0.0012 ± 0.0007                | 0.0037 ± 0.0009 | 0.0142 ± 0.0015 | 0.0686 ± 0.0051 | 0.0000817 ± 0.0000310 | 0.0113 ± 0.0049 |
|       | 0.060                         | 0.014 | 0.0022 | 0.0014 | 0.27 | 0.14 |
| ΔTL  | 0.00000545 ± 0.00000093       | 0.0000172 ± 0.0000002 | 0.0000796 ± 0.00000034 | 0.0000314 ± 0.0000025 | 0.000005 ± 0.000005 | 0.0000622 ± 0.0000095 |
|       | 0.000024                      | 0.000066 | 0.0000124 | 0.0000063 | 0.016 | 0.0076 |
| ΔTL+ΔFL| 0.0000387 ± 0.0000009         | 0.000128 ± 0.000009 | 0.000457 ± 0.000036 | 0.00140 ± 0.00022 | 0.000425 ± 0.000089 | 0.000439 ± 0.000055 |
|       | 0.0017                        | 0.00049 | 0.000071 | 0.000028 | 0.14 | 0.054 |

Activity relative to wild-type Taq RNAP [(\(k_{\text{obs}}\))/(\(k_{\text{obs}}\)Taq WT) at the same temperature]

\(^*\) - data from Miropolskaya et al. (26)
Table S2. Activation parameters of the nucleotide addition reaction for various RNAPs at different temperatures.

| RNAP | t (°C) | ΔG# (kJ/mol) | Δ∆G# a (kJ/mol) | ΔH# (kJ/mol) | Δ∆H# a (kJ/mol) | TΔS# (kJ/mol) | TΔ∆S# a (kJ/mol) |
|------|--------|--------------|-----------------|-------------|-----------------|--------------|-----------------|
| **Dra** | | | | | | | |
| 10  | 62.20 ± 0.62 | -15.87 | 76.45 ± 5.72 | -114.60 | 14.24 ± 5.75 | -98.73 |
| 20  | 60.95 ± 0.88 | -14.07 | 76.36 ± 5.72 | -114.60 | 15.42 ± 5.79 | -100.53 |
| 30  | 60.76 ± 0.59 | -8.83 | 76.28 ± 5.72 | -114.60 | 15.52 ± 5.75 | -105.77 |
| 40  | 60.61 ± 0.97 | -6.01 | 76.20 ± 5.72 | -114.60 | 15.58 ± 5.80 | -108.59 |
| **Taq WT** | | | | | | | |
| 10  | 78.08 ± 1.13 | 0 | 191.05 ± 10.17 | 0 | 112.97 ± 10.23 | 0 |
| 20  | 75.01 ± 0.20 | 0 | 190.96 ± 10.17 | 0 | 115.95 ± 10.17 | 0 |
| 30  | 69.58 ± 1.43 | 0 | 190.88 ± 10.17 | 0 | 121.30 ± 10.27 | 0 |
| 40  | 66.62 ± 0.68 | 0 | 190.80 ± 10.17 | 0 | 124.17 ± 10.19 | 0 |
| **F-Dra** | | | | | | | |
| 10  | 69.86 ± 0.40 | -8.22 | 147.15 ± 11.88 | -43.90 | 77.29 ± 11.88 | -35.68 |
| 20  | 66.17 ± 0.75 | -8.85 | 147.06 ± 11.88 | -43.90 | 80.90 ± 11.90 | -35.05 |
| 30  | 64.27 ± 0.99 | -5.31 | 146.98 ± 11.88 | -43.90 | 82.71 ± 11.92 | -38.59 |
| 40  | 62.13 ± 0.17 | -4.50 | 146.90 ± 11.88 | -43.90 | 84.77 ± 11.88 | -39.40 |
| **Q1046A** | | | | | | | |
| 10  | 73.74 ± 0.92 | -4.34 | 155.95 ± 9.98 | -35.10 | 82.21 ± 10.03 | -30.76 |
| 20  | 72.22 ± 0.66 | -2.79 | 155.86 ± 9.98 | -35.10 | 83.64 ± 10.01 | -32.31 |
| 30  | 68.47 ± 0.32 | -1.12 | 155.78 ± 9.98 | -35.10 | 87.31 ± 9.99 | -33.98 |
| 40  | 65.31 ± 0.08 | -1.31 | 155.70 ± 9.98 | -35.10 | 90.38 ± 9.98 | -33.79 |
| **ΔFL** | | | | | | | |
| 10  | 85.13 ± 1.55 | 7.06 | 97.55 ± 6.87 | -93.50 | 12.41 ± 7.05 | -100.56 |
| 20  | 85.38 ± 0.61 | 10.36 | 97.46 ± 6.87 | -93.50 | 12.09 ± 6.90 | -103.86 |
| 30  | 85.00 ± 0.27 | 15.41 | 97.38 ± 6.87 | -93.50 | 12.38 ± 6.88 | -108.91 |
| 40  | 83.78 ± 0.19 | 17.16 | 97.30 ± 6.87 | -93.50 | 13.51 ± 6.88 | -110.66 |
| **ΔTL** | | | | | | | |
| 10  | 97.74 ± 0.01 | 19.66 | 98.45 ± 5.60 | -92.60 | 0.71 ± 5.60 | -112.26 |
| 20  | 98.48 ± 0.02 | 23.47 | 98.36 ± 5.60 | -92.60 | -0.12 ± 5.60 | -116.07 |
| 30  | 98.06 ± 0.01 | 28.47 | 98.28 ± 5.60 | -92.60 | 0.22 ± 5.60 | -121.07 |
| 40  | 97.81 ± 0.03 | 31.18 | 98.20 ± 5.60 | -92.60 | 0.39 ± 5.60 | -123.78 |
| **ΔTL+ΔFL** | | | | | | | |
| 10  | 93.12 ± 0.02 | 15.05 | 86.35 ± 4.62 | -104.70 | -6.78 ± 4.62 | -119.75 |
| 20  | 93.58 ± 0.01 | 18.57 | 86.26 ± 4.62 | -104.70 | -7.32 ± 4.62 | -123.27 |
| 30  | 93.65 ± 0.02 | 24.07 | 86.18 ± 4.62 | -104.70 | -7.47 ± 4.62 | -128.77 |
| 40  | 93.91 ± 0.05 | 27.29 | 86.10 ± 4.62 | -104.70 | -7.81 ± 4.62 | -131.99 |

* Differences in the thermodynamics parameters relative to wild-type Taq RNAP.
Table S3. Fidelity of nucleotide incorporation by wild-type and mutant Taq RNAPs. All $k_{obs}$ values were measured at 1 mM NTP concentrations at 40°C. The data for UTP are from Table S1.

| RNAP        | $k_{obs}$ (s$^{-1}$), 40 °C |       |       |       |
|-------------|----------------------------|-------|-------|-------|
|             | UTP                        | CTP   | dTTP  |       |
| Taq WT      | 50.0 ± 13.1                 | 0.0108 ± 0.0046 | 0.00350 ± 0.00127 | 0.00022 | 0.000070 | 1 | 1 |
| F-Dra       | 281 ± 18.5                  | 0.0307 ± 0.0036 | 0.0123 ± 0.0018 | 0.00011 | 0.000044 | 0.5 | 0.63 |
| ΔFL         | 0.0686 ± 0.0051             | 0.000049 ± 0.000019 | 0.000337 ± 0.000259 | 0.00071 | 0.0049 |       |       |
| ΔTL         | 0.000314 ± 0.000025         | 0.0000015 ± 0.0000006 | 0.000025 ± 0.000005 | 0.0047 | 0.080 | 3.3 | 70.3 |
| ΔTL+ΔFL     | 0.00140 ± 0.00022           | 0.0000065 ± 0.0000033 | 0.000082 ± 0.000006 | 0.0046 | 0.059 | 21.3 | 840 |

Fidelity of NTP incorporation | $(k_{obs \text{ Incorrect}}/k_{obs \text{ Correct}})$ under the same conditions |
Changes in fidelity relative to wild-type Taq RNAP | $(\text{Fidelity}/\text{Fidelity Taq WT})$ in the same reaction |
**Table S4.** The rates of RNA cleavage by various RNAPs at different conditions.

| RNAP   | $k_{obs}$ (s$^{-1}$) |      |      |      |
|--------|----------------------|------|------|------|
|        | 40°C                 | 20°C |      |      |
|        | -GreA               | -GreA+GreA |
| $Dra$  | 0.164 ± 0.055        | 0.022 ± 0.003 | 1.3 | 8.4 |
| $Taq$  | 0.129 ± 0.015        | 0.00264 ± 0.00062 | 0.203 ± 0.020 |
|        | 1                    | 1         | 1    | 1    |
| $F-Dra$| 0.0214 ± 0.0008      | 0.224 ± 0.032 | 8.1 | 1.1 |
| $\Delta$FL | 0.000331 ± 0.0000012 | 0.184 ± 0.004 | 0.13 | 0.9 |
| $\Delta$TL | 0.0000051 ± 0.0000014 | 0.0333 ± 0.0007 | 0.0019 | 0.16 |
| $\Delta$TL+$\Delta$FL | 0.000135 ± 0.0000044 | 0.00721 ± 0.00056 | 0.0051 | 0.035 |

Activity relative to $Taq$ RNAP [(k$_{obs}$)/(k$_{obs}$ Taq WT) under the same conditions]
Figure S1. Structures of the RNAP active centre with different conformations of the TL. (A) TEC with NTP bound in the insertion site (1). The TL is in a fully folded conformation. (B) Stl-inhibited TEC with NTP in the pre-insertion site (1). The TL is partially unfolded. (C) Holoenzyme RNAP with the open TL (29). The color codes for the RNAP elements and the reaction components are shown at the bottom of the figure. In the NTP insertion complex (A) the jaw domain (semitransparent) is disordered. Residues Q1046, H1242 and E445 in the FL, TL and Fork2 are shown as CPK models. For each structure, two views are shown: the main channel view at the left, and the secondary channel view at the right. The arrow above the right view in (A) indicates the direction of view in Figure 5.
Figure S2. Temperature dependence of the differences in the activation parameters between various RNAPs and wild-type Taq RNAP. (A) The $\Delta \Delta G^\#$ values. (B) The $\Delta \Delta H^\#$ values and $T \Delta \Delta S^\#$ values. The $\Delta \Delta H^\#$ values are constant and are equal to the differences in the activation energies between corresponding RNAPs and wild-type Taq RNAP.
Figure S3. The structure of the active centre during RNA cleavage. (A) The structure of backtracked TEC of *S. cerevisiae* RNAPII (10). The color code corresponds to Figure 1 and Figure S1. The unpaired 3\(^\prime\)-segment of RNA is shown in black. The TL in the “trapped” conformation directly contacts the FL (10). Residue H1085 in the TL corresponds to H1242 in *Tth* RNAP. (B) The structure of the TEC of *Tth* RNAP in complex with Gfh1 (12). The N-terminal coiled-coil domain of Gfh1 (shown in blue) interacts with both the TL and the FL. The TL is in an unfolded conformation.