Supplemental Methods, Tables, Figures

Methods

Reagents, Buffers, and Gels

Reagents used for buffers and stock solutions were purchased in the highest available grade and used as received. All solutions were prepared using 18MΩ deionized water from a Barnstead EPure system. NTPs and dNTPs (Boston Bioproducts, Thermo Fisher, New England Biolabs) used in transcription assays and PCR reactions were 99% pure and used as received. Enzymes for PCR reactions were purchased from NEB and used according to the manufacturers protocols.

Storage Buffer for core RNAP, σ^{70} and RNAP holoenzyme is 50% v/v glycerol, 0.01M Tris, 0.1 M NaCl, 0.1 mM EDTA, and 0.1 mM DTT. Transcription buffer (TB) is 40 mM Tris, 5 mM $MgCl_2$, 60 mM KCL, 1 mM DTT, 0.05 mg/ml BSA, and is adjusted to pH 8.0 at experimental temperature (19°C, 25°C, or 37°C).

2X initiation solution (IS) for rapid quench flow (RQF) transcription assays at 25 °C and 37 °C with α-^{32}P-GTP is 0.1 mg/ml heparin, 400 μM ATP, 400 μM UTP, 20 μM unlabeled GTP, and 35 nM α-^{32}P-GTP. For 25 °C and 37 °C assays with α-^{32}P-UTP, IS is 0.1 mg/ml heparin, 400 μM ATP, 400 μM GTP, 20 μM unlabeled UTP, and 35 nM α-^{32}P-UTP. IS is mixed 1:1 with pre-formed OC to reach experimental concentrations of each NTP.

Quench Solution (QS) for RQF transcription assays is 8M Urea and 15 mM EDTA in TB. QS with added dyes (QSD) for polyacrylamide gel electrophoresis (PAGE) has 0.05% xylene cyanol and 0.05% bromphenol blue in QS. TBE buffer for PAGE is 90mM Tris-borate (pH 8.3) and 2mM Na_2EDTA. All transcription gels are 20% acrylamide-(bis)acrylamide (19:1), and were made using the UreaGel system (National Diagnostics).

RNA Polymerase and λP_R Promoter DNA

RNAP core enzyme ($\alpha_2\beta\beta'\omega$) is overexpressed from pVS10 plasmid and purified via Ni affinity chromatography (1). The σ^{70} subunit is overexpressed from pIA586 and purified using Ni affinity chromatography(1). Holoenzyme is reconstituted using a 1:2 ratio of core to σ^{70}. Filter binding activity assays performed on preparations of RNAP holoenzyme used here show that 50% ± 10% of RNAP molecules form a stable open complex with the λP_R promoter at 37 °C. All RNAP concentrations reported here refer to this active fraction. λP_R promoter DNA is prepared as described previously (2). Primer sequences used to assemble the λP_R promoter DNA fragment are reported in SI Appendix Table S1.

Initiation Kinetic Assays

Initiation kinetic assays are performed and imaged as described previously (1). Assays are designed to obtain single-round synthesis of full length (FL) RNA, defined for this study as any RNA length greater than 10-mer (see main text). Briefly, 2X initiation solution (IS) is mixed 1:1 with preformed OC at time zero using a KinTek Corp. Rapid Quench Flow (RQF). Reactions are quenched with QS and RNA is separated via PAGE. Both the IS and preformed OC solutions are incubated at the experimental temperature (19 °C, 25 °C, 37 °C) for 1 h prior to mixing in the RQF at that temperature. Transcription gels are transferred to a phosphorimaging cassette and analyzed using a Typhoon 9000 phosphorimager (18 h exposure). Peak area is converted to moles of observed product as described previously (2). To obtain good labelling efficiency without the use of high concentrations of labelled NTP that cause background problems and limit the ability to detect small RNAs on the gels, we use a low concentration of ^{32}P-labelled NTP (17.5 nM) and a reduced concentration (10 μM) of the corresponding unlabeled NTP. The probability of incorporation of the radiolabeled NTP at each position is determined by the ITR sequence, and was accounted for using incorporation probabilities (see SI Appendix, Table S2) calculated for different RNA lengths as described previously (2).

Quantitative results reported here are from the average of 2-4 RQF experiments at each set of NTP concentrations investigated, and the uncertainties reported are estimates of one standard deviation from the mean.

Analysis of RNA Synthesis as a Function of Time

The amount of FL RNA and of the eight detectable short RNA species (3-mer to 10-mer) are calculated as a function of time from phosphorimager line scans of transcription gels. The only undetected short RNA is 2-mer (pppApU), which is not labelled in α-^{32}P-GTP labeled experiments and does not separate sufficiently from the broad band of unincorporated monomer to quantify in α-^{32}P-UTP labelled experiments. Nevertheless, kinetic parameters (K_m, k_{cat}) of 2-mer synthesis are obtained from fitting the kinetic data for the longer RNAs to the initiation mechanism (Fig. 3).

Analysis of Short RNAs Synthesized by Nonproductive Complexes

Kinetics of synthesis of the first short RNA of each length by nonproductive (stalled) complexes are well-described as a first order (single exponential) approach to a plateau value, as is the case for FL RNA (1). For each short RNA length, at each time point during FL RNA synthesis, the amount of short RNA synthesized by nonproductive complexes is predicted using Origin 2018b and subtracted from the observed total amount of that RNA to obtain the amount of transient short RNA synthesized by productive complexes. This dissection is shown in Figures S3-S6 for the temperatures and NTP conditions investigated. Amounts of RNA transients determined by this procedure are plotted in Fig. 3.

Supplemental Information for Figures S3-S6

Amounts of each detectable short RNA species (3-mer to 10-mer) formed prior to promoter escape at the low-UTP and high-UTP conditions at 25 °C and 37 °C are plotted vs. time (linear scale) in Figs. S3-S6. At 25 °C, high UTP, significant transient peaks in amounts of

all short RNAs except 7-mer are observed. This is consistent with previously-reported observations at 19 ºC, high UTP (1). Transient buildup of N-mer RNA is expected when the synthesis of N-mer is relatively rapid and the subsequent synthesis of (N+1)-mer is slow. At high UTP (and low GTP) the first step of initiation (pppApU synthesis) is rapid and not a bottleneck, and intermediates (5-mer, 9-mer, 10-mer) preceding steps where G is added are expected to accumulate because of the low GTP concentration. Accumulation of 3-mer, 4-mer, 6-mer and 8-mer must however indicate that the rate constants of the following steps of RNA synthesis (synthesis of 4-mer, 5-mer, 7-mer and 9-mer) are small. This was previously observed at this NTP condition at 19 ºC, where rate constants for synthesis of these four RNAs as well as 8-mer and 11-mer are smaller than for other steps of initiation. By contrast, at 25 °C and low UTP, where synthesis of pppApU is slowed by the low UTP concentration, significant transient amounts of only three RNAs (3-mer, 6-mer, 10-mer) are observed. By the same reasoning as above, the 10-mer transient indicates the rate constant for conversion to 11-mer is small, while 3-mer and 6-mer transients are interpreted in terms of slow rates for the following steps, both because the next base incorporated is U and because the rate constants of these steps are small. These results are consistent with previous findings at 19 °C (Henderson ’19).

In contrast, at the high UTP condition at 37°C only four transient RNA intermediates (5-mer, 6-mer, 9-mer, 10-mer) from productively-initiating complexes are detected (Fig. S6), and no transient RNAs are detected at the low UTP condition (Fig. S5). These observations are consistent with the proposal that an additional step at the beginning of the mechanism slows the kinetics of FL RNA synthesis at 37 ºC, especially at the low UTP condition.

Supplemental Information for Figure S7 and Table S5

Values of ln k_i for each step of initiation are plotted vs. 1/T in Fig. S7. Some 37 °C k_i values (for synthesis of 3-mer, 4-mer, 7-mer, and 8-mer) are lower bounds (as indicated in Table S4). When plotted with well-determined 19 °C and 25 °C k_i values, these 37 °C lower-

bound k_i values give linear Arrhenius plots, and we accept them as the best estimates of these 37 °C k_i values. Arrhenius activation energies $E_{Act,\,i}$ calculated from slopes of plots in Fig. S7 (slope = $-E_{act,\,i}/R$) are plotted in Fig. 4 and listed in Table S5.

Activation free energies ($\Delta G_i^{o\ddagger}$) for each step of RNA-DNA hybrid extension are calculated from the equation

$$\ln k_i = \ln k_{max} - (\Delta G_i^{o\ddagger}/RT) \qquad \text{Eq. S1}$$

where $\Delta G_i^{o\ddagger}$ is the free energy difference between the high free energy transition state of the catalytic step and the reactants (incoming NTP, pre-translocated state of the initiation complex) and k_{max} is the maximum rate constant for the hypothetical situation where $\Delta G_i^{o\ddagger} = 0$. As an estimate of the maximum second order rate constant k_{max}, we use a plausible value of a diffusion-limited (orientation-corrected) rate constant for NTP binding ($k_{max} \approx 10^3\ \mu M^{-1}\ s^{-1}$). The uncertainty in this estimate, though large, is inconsequential for the analysis reported here. We assume that k_{max} (as a diffusion-limited rate constant) is the same for each step of initiation and neglect its expected small temperature dependence.

From k_i values as a function of temperature (Fig. S7) and k_{max}, activation free energies and their enthalpic and entropic components are obtained for each initiation step:

$$\Delta G_i^{o\ddagger} = -RT \ln (k_i/k_{max}) = \Delta H_i^{o\ddagger} - T\Delta S_i^{o\ddagger}$$

where $$\Delta H_i^{o\ddagger} \approx E_A = -R\left(\frac{\partial \ln k_i}{\partial\left(\frac{1}{T}\right)}\right) \text{ and } \Delta S_i^{o\ddagger} = (\Delta H_i^{o\ddagger} - \Delta G_i^{o\ddagger})/T \qquad \text{Eq. S2}$$

Values of $\Delta G_i^{o\ddagger}$, $\Delta H_i^{o\ddagger}$, $\Delta S_i^{o\ddagger}$ and their uncertainties for each step of extension of the RNA-DNA hybrid by NTP incorporation at 19 °C are listed in Table S5 and values of $\Delta H_i^{o\ddagger}$ are interpreted in Table S6.

Table S1: Primers used for λP_R template preparation

λP_R_wt_forward (-71 to -12)	CCACGAATTCGGATAAATATCTAACACCGTGCGTGTTGACTATTTTACCTCTGGCGGTG
λP_R_wt_reverse (-24 to 31)	ACAAAACCTTCATAGAACCTCCTTACTACATGCAACCATTATCACCGCCAGAGGT
HBOT	CACCTGCACCGACAAAACCTT
HTOP	CCAGCATTCCTCCACGAATTC

Table S2: Probabilities of Incorporation of α-^{32}P-UTP and α-^{32}P-GTP

Number of UTP or GTP in Transcript	Transcript Length (UTP)	Transcript Length (GTP)	α-^{32}P-NTP Incorporation Probability
0	-	2-mer	0
1	2-mer, 3-mer	3-mer to 5-mer	0.00175
2	4-mer to 6-mer	6-mer to 9-mer	0.0035
3	7-mer to 14-mer	10-mer	0.00525
4	15-mer	11-mer to 12-mer	0.007
5	16-mer	13-mer	0.00875
6	ND[a]	14-mer to 16-mer	0.0105
10	ND[a]	31-mer	0.0175
12	31-mer	-	0.021

[a] Not detected

Table S3: Lag Times and First Order Rate Constants k_{FL} for FL (> 10-mer) RNA Synthesis

	Low UTP Lag Time	High UTP Lag Time	Low UTP k_{FL}	High UTP k_{FL}
19 °C	3.6 s	2.3 s	0.062 s^{-1}	0.11 s^{-1}
25 °C	2.2 s	1.2 s	0.082 s^{-1}	0.22 s^{-1}
37 °C	2.0 s	0.35 s	0.048 s^{-1}	0.39 s^{-1}

Table S4: Composite Rate Constants k_i (μM^{-1} s^{-1}) for Each Initiation Step at 19, 25, and 37 °C [a]

Base Position i	Base Incorporated	19 °C [b] k_i (μM^{-1} s^{-1})	25 °C k_i (μM^{-1} s^{-1})	37 °C [c] k_i (μM^{-1} s^{-1})
2	U	0.63 ± 0.22	0.69 ± 0.24	1.5 ± 0.5
3	G	0.15 ± 0.08	0.49 + 0.64	≥ 2
4	U	0.012 ± 0.001	0.033 ± 0.005	≥ 0.2
5	A	0.064 ± 0.044	0.077 ± 0.040	0.11 ± 0.07
6	G	0.16 ± 0.04	0.21 ± 0.04	0.24 ± 0.08
7	U	0.017 ± 0.002	0.021 ± 0.003	≥ 0.03
8	A	0.037 ± 0.020	0.044 ± 0.022	≥ 0.07
9	A	0.052 ± 0.035	0.043 ± 0.021	0.033 ± 0.024
10	G	0.13 ± 0.03	0.22 + 0.06	0.62 ± 0.17
11	G	0.028 ± 0.003	0.049 ± 0.006	0.21 ± 0.03

[a] From fitting kinetic data at each temperature to Mechanism 2 (Fig. 3).

[b] Previously reported 19 °C ki from fitting to Mechanism 1 (1) are not significantly different.

[c] Uncertainties in 37 °C k_i values could not be determined from the fit because of the insufficient number of detectable small RNA transient intermediates (Figs. S5-S6). Uncertainties in other 37 °C k_i are estimated using the larger of the % uncertainties for that step in the 19 °C and 25 °C data.

Table S5: Arrhenius Activation Energies $E_{A,i}$ and Transition State Barriers $\Delta G_i^{o\ddagger}$, $\Delta H_i^{o\ddagger}$, $\Delta S_i^{o\ddagger}$ for Each Step of Initiation[a]

Base Position i	k_i [b] ($\mu M^{-1} s^{-1}$)	$\Delta G_i^{o\ddagger}$ [c] (kcal/mol)	$E_{A,i} \approx \Delta H_i^{o\ddagger}$ (kcal/mol)	$\Delta S_i^{o\ddagger}$ (eu)
2	0.63 ± 0.22	4.3 ± 0.3	9 ± 5	~ 16
3	0.15 ± 0.08	5.1 ± 0.8	27 ± 14	~ 75
4	0.012 ± 0.001	6.6 ± 0.1	26 ± 2	~ 66
5	0.064 ± 0.044	5.6 ± 0.3	5.1 ± 10	~ -2
6	0.16 ± 0.04	5.1 ± 0.1	7.2 ± 3	~ 7
7	0.017 ± 0.002	6.4 ± 0.1	6.8 ± 2	~ 1
8	0.037 ± 0.020	5.9 ± 0.3	6.6 ± 7	~ 2
9	0.052 ± 0.039	5.7 ± 0.3	-4.4 ± 10	~ -35
10	0.13 ± 0.03	5.2 ± 0.2	16 ± 4	~ 37
11	0.028 ± 0.003	6.1 ± 0.1	21 ± 2	~ 51

[a] Activation free energy, enthalpy, and entropy changes $\Delta G_i^{o\ddagger}, \Delta H_i^{o\ddagger}, and\ \Delta S_i^{o\ddagger}$ are defined in Eq. S1-S2.

[b] Determined at 19 °C ((1); see Table S4).

[c] Calculated from 19 °C k_i values using Eq. S1 with $k_{max} = 10^3\ \mu M^{-1} s^{-1}$

Table S6: Interpretation of Rate Constants k_i and Barriers $\Delta G_i^{o\ddagger}$, $\Delta H_i^{o\ddagger}$, $\Delta S_i^{o\ddagger}$ for Each Step of Initiation

Base Position i	k_i/k_3	Proposed Region where RNAP-DNA Contacts are Disrupted in step i [a]	Proposed Number of Bubble Bases that Stack in Step i [b]	Proposed Positions of Duplex Formation (relative to +1 TSS) [c]
2	~ 4.2	-	0	-
3	1	-	0	-
4	~ 0.1	Discriminator	0	-
5	~ 0.4	Discriminator	4	-1, -2
6	~ 1	-	4	-3, -4
7	~ 0.1	-10	4	-5, -6
8	~ 0.2	-10	4	-7, -8
9	~ 0.3	-10	6	-9, -10, -11
10	~ 0.9	-	2	-
11	~ 0.2	-35	0	-

[a] Based on analysis in (1). Reductions in k_i values for steps 3, 6, and 10 as compared to step 2 (no translocation) cannot be dissected based on activation thermodynamics (Table S5) but may include disruption of some discriminator or -10 contacts.

[b] The proposed number of bases that convert from unstacked to stacked in each step is estimated from differences in activation enthalpies (Table S5): Number of bases that stack in step i = $(\Delta H_3^{o\ddagger} - \Delta H_i^{o\ddagger})/5$, assuming -5 kcal/mol per base stacking interaction. Similar results are obtained from analysis of $(\Delta S_3^{o\ddagger} - \Delta S_i^{o\ddagger})$ assuming an entropy change of approximately -15 to -20 eu per stacking interaction.

[c] These predictions assume that disruption of RNAP-strand contacts and duplex formation occur first at the downstream edge of the discriminator and propagate upstream, and that formation of one base pair has the same enthalpy change (-10 kcal mol^{-1}) as stacking of two bases.

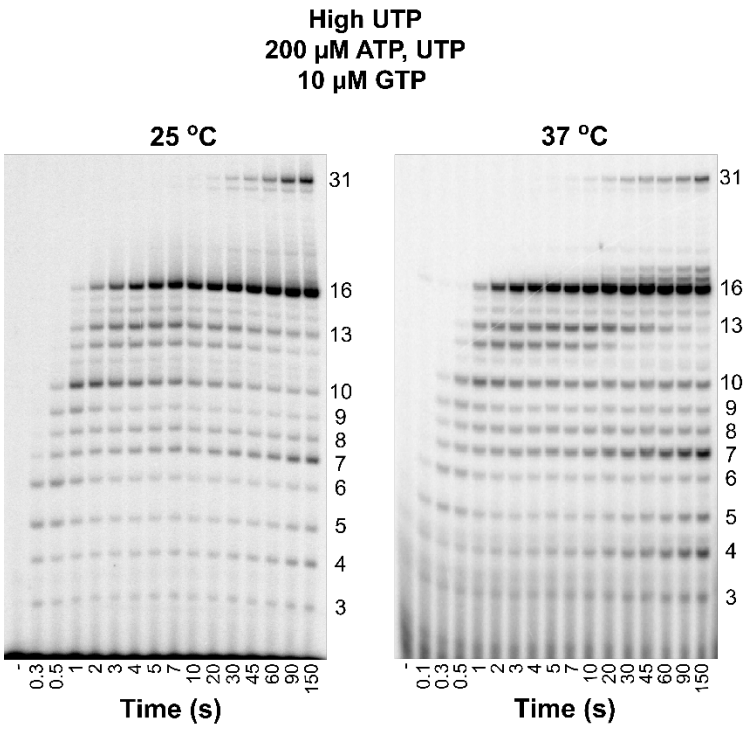

Figure S1: Step-by-Step Kinetics of Transcription Initiation from λP_R Promoter at the High-UTP Condition. Polyacrylamide gel electrophoretic separations of individual ^{32}P-labelled RNA bands (from 3-mer to 16-mer and 31-mer) observed as a function of time during productive and nonproductive initiation at 25 °C **(left)** and 37 °C **(right)**. Lanes span the time range from 0.1 or 0.3 s to 150 s after adding NTPs and heparin to OC formed by premixing RNAP and λP_R promoter DNA (See SI Methods). Gels shown are for the "high UTP" condition: 200 µM ATP and UTP, 10 µM GTP, 17.5 nM α-^{32}P-GTP. CTP is omitted causing transcription to halt at a 16-mer RNA, before read-through occurs to synthesize longer transcripts (1). Representative "low UTP" (200 µM ATP and GTP, 10 µM UTP, 17.5 nM α-^{32}P-UTP) gels at 25 °C and 37 °C are shown in the main text Fig. 2.

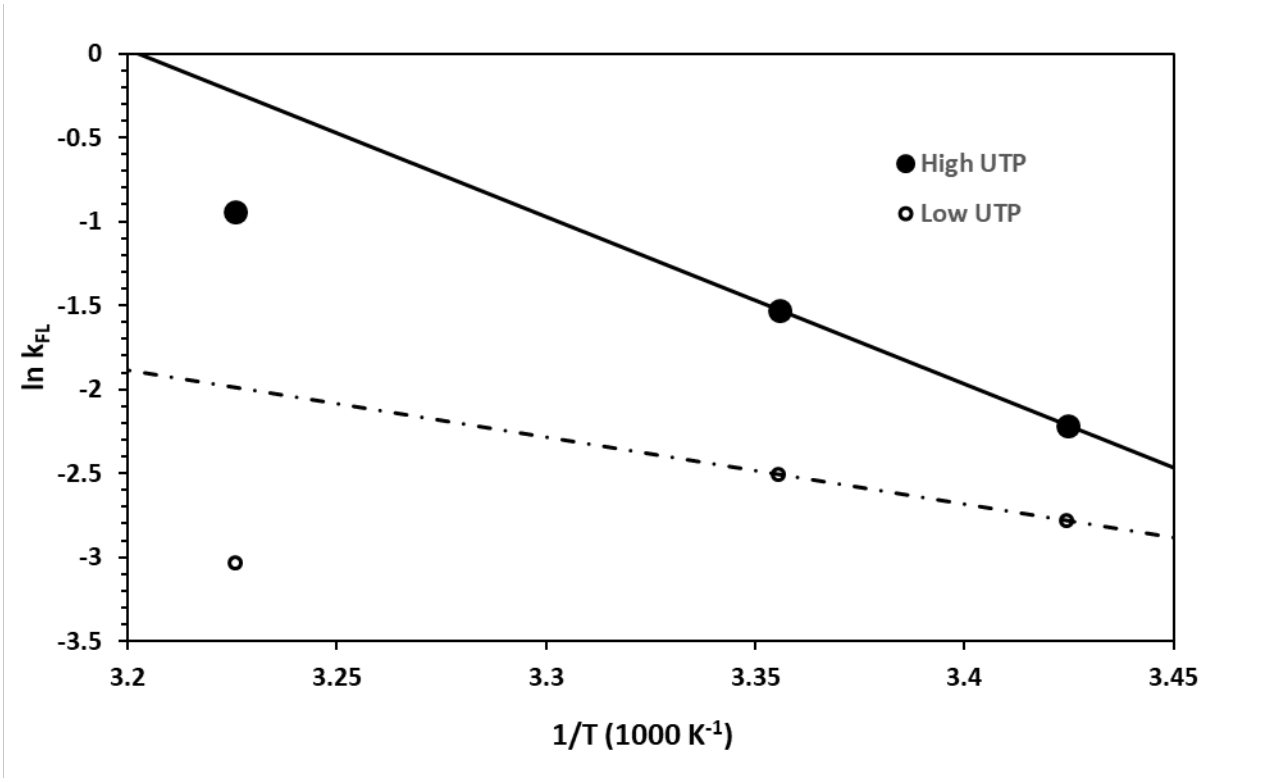

Figure S2: Arrhenius Plots of lnk_{FL} vs 1/T. Natural logarithms of first order rate constants for full length RNA production (k_{FL}, in s^{-1}) are plotted vs inverse temperature (K^{-1}) for high UTP (200 µM UTP, 10 µM GTP; filled circles) and low UTP (10 µM UTP, 200 µM GTP; open circles). At both NTP conditions (especially low UTP), the 37 °C rate constant is significantly smaller than predicted assuming a linear Arrhenius relationship defined by the 19 °C and 25 °C rate constants (predictions shown).

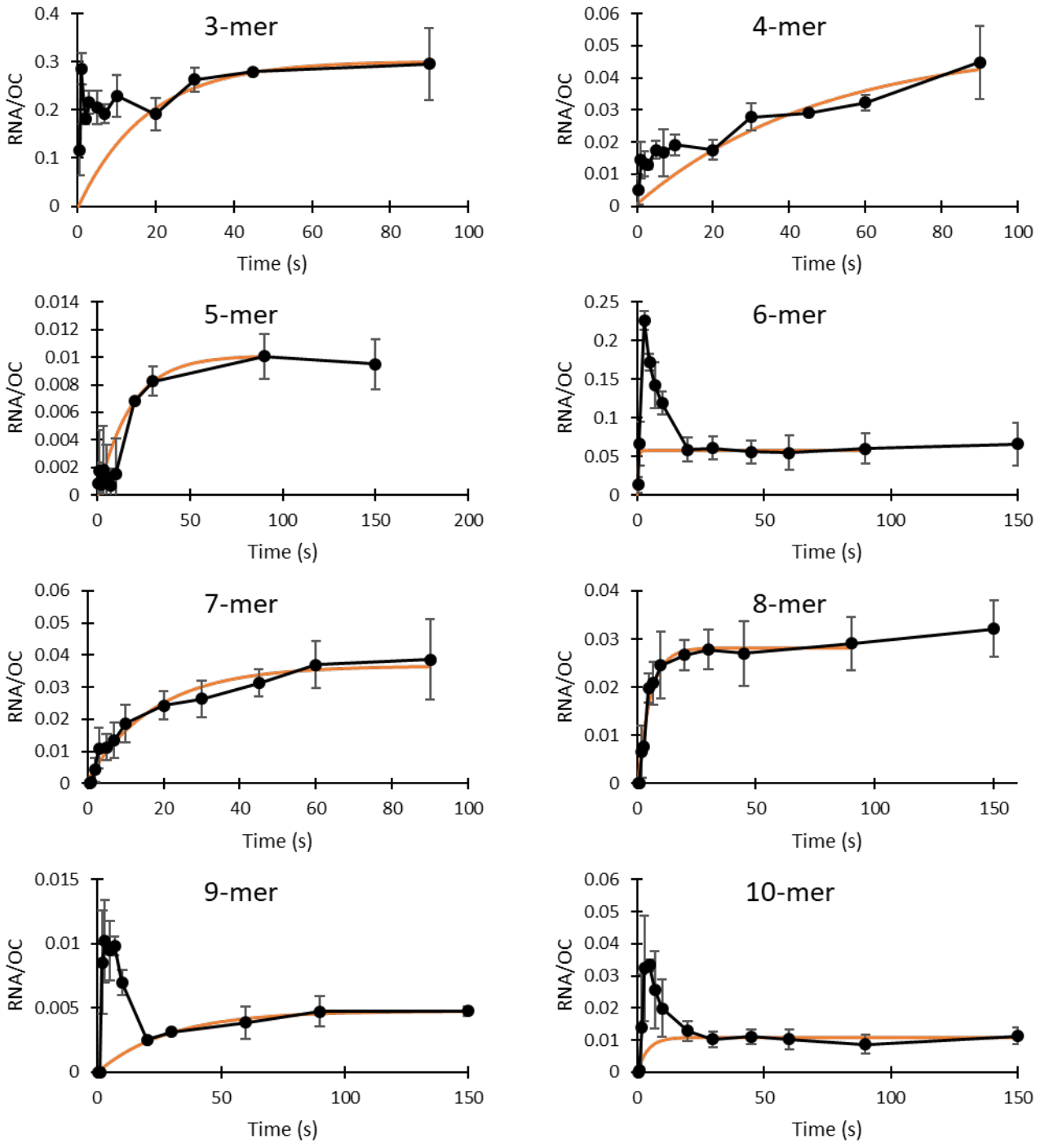

Figure S3: Average Amounts of Short RNAs vs Time at 25 °C, Low UTP. Amounts of 3-mer to 10-mer RNA per OC (averages of 4 independent experiments) are shown with one standard deviation error bars (black). Data for all detectable lengths up to the point of escape are shown. NTP concentrations are 200 µM ATP and GTP, 10 µM UTP. For each RNA length, the predicted contribution from the initial round of synthesis by non-productive complexes is also shown (orange).

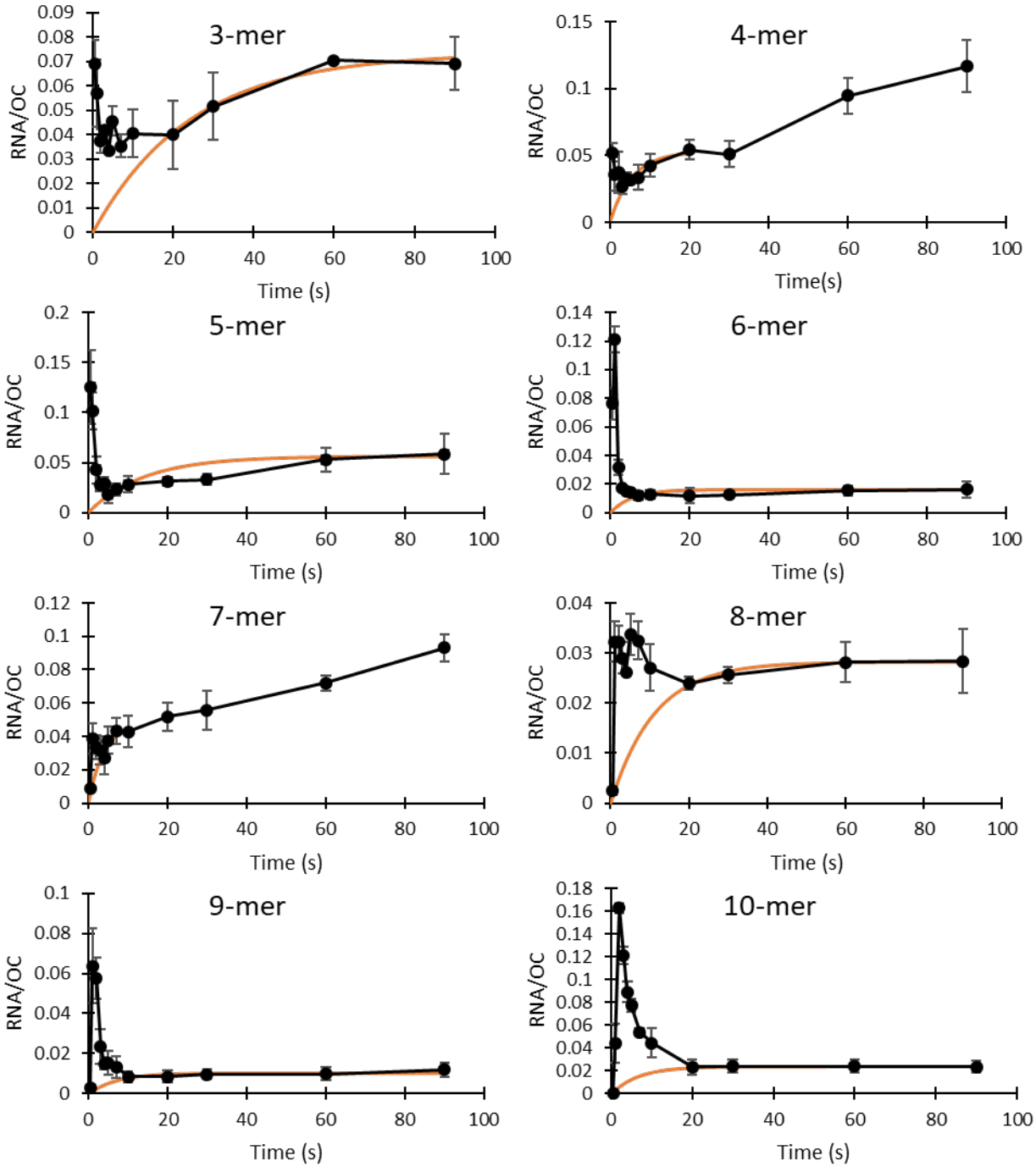

Figure S4: Average Amounts of Short RNAs vs Time at 25 °C, High UTP. Amounts of 3-mer to 10-mer RNA per OC (averages of 4 independent experiments) are shown with one standard deviation error bars (black). Data for all detectable lengths up to the point of escape are shown. NTP concentrations are 200 µM ATP and UTP, 10 µM GTP. For each RNA length, the predicted contribution from the initial round of synthesis by non-productive complexes is also shown (orange).

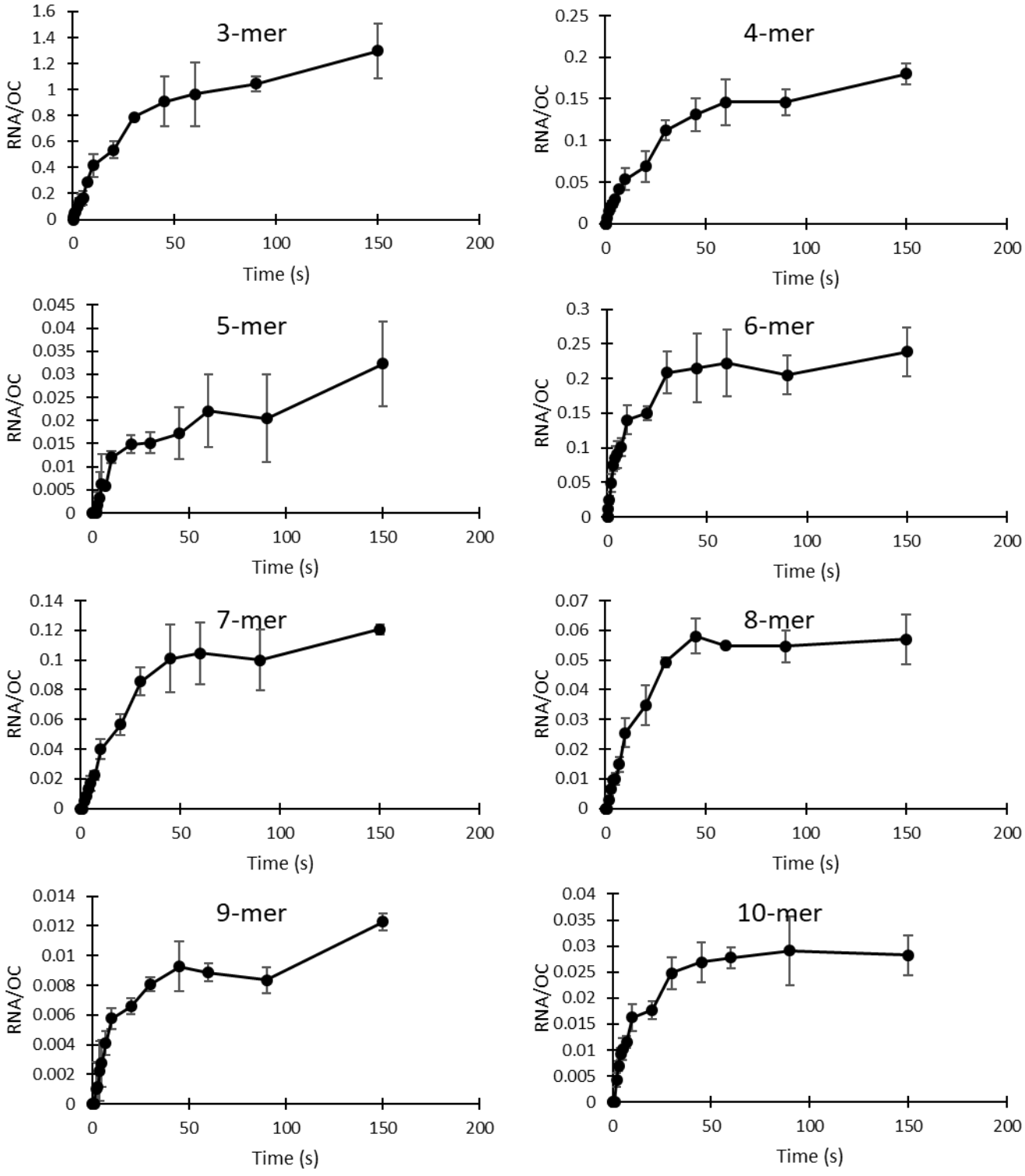

Figure S5: Average Amounts of Short RNAs vs Time at 37 °C, Low UTP. Amounts of 3-mer to 10-mer RNA per OC (averages of 3 independent experiments) are shown with one standard deviation error bars (black). Data for all detectable lengths up to the point of escape are shown. NTP concentrations are 200 µM ATP and GTP, 10 µM UTP. Transient concentrations of short RNA synthesized by productive OC are too small to detect at 37 °C, low UTP, and therefore these curves show the initial and subsequent rounds of RNA synthesis by the non-productive fraction of OCs.

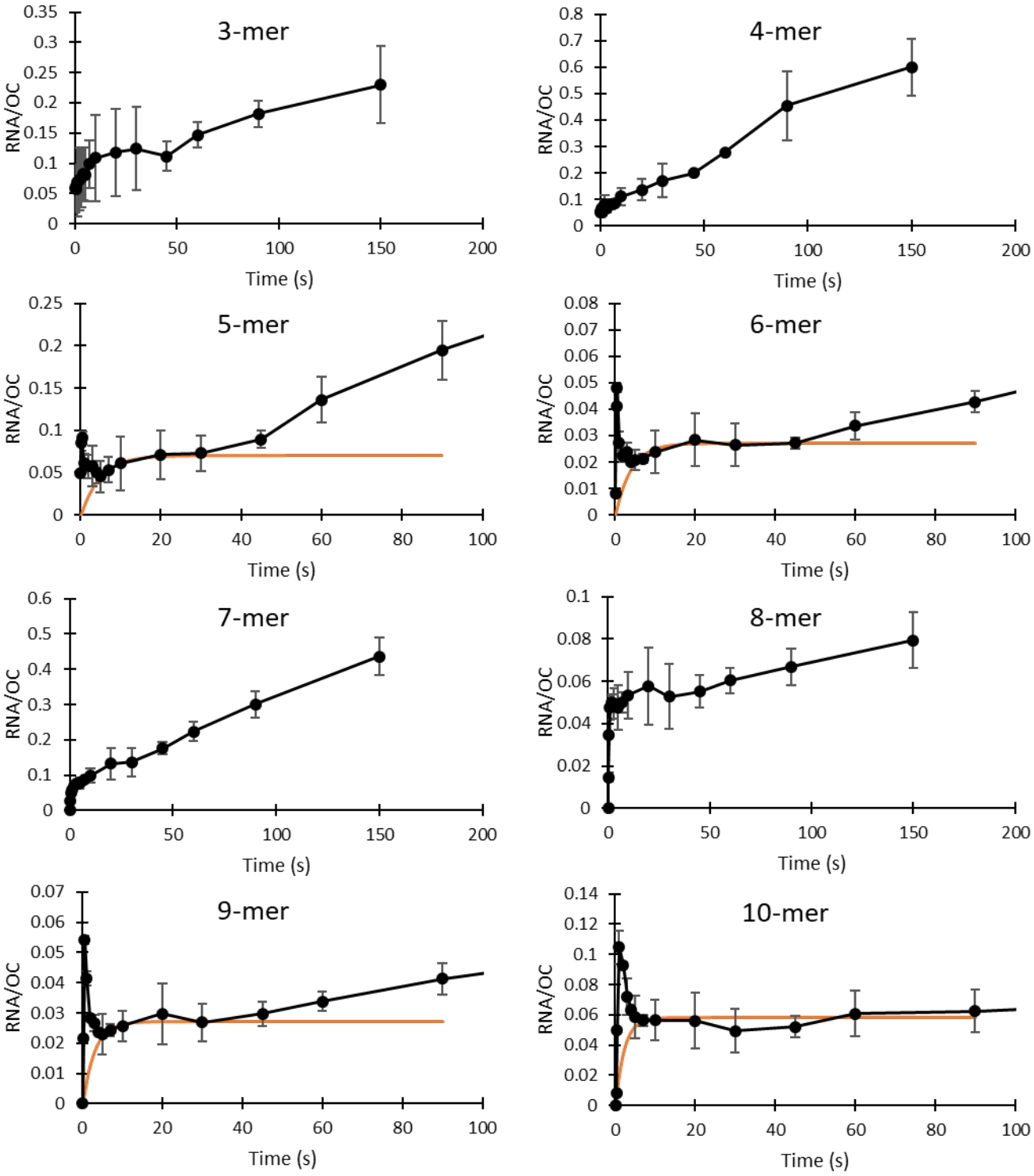

Figure S6: Average Amounts of Short RNAs vs Time at 37 °C, High UTP. Amounts of 3-mer to 10-mer RNA per OC (averages of 2 independent experiments) are shown with error bars representing the range (black). Data for all detectable lengths up to the point of escape are shown. NTP concentrations are 200 µM ATP and UTP, 10 µM GTP. For each RNA length where a transient concentration of RNA from synthesis by productive complexes is observed, the predicted contribution from the initial round of synthesis by non-productive complexes is also shown (orange).

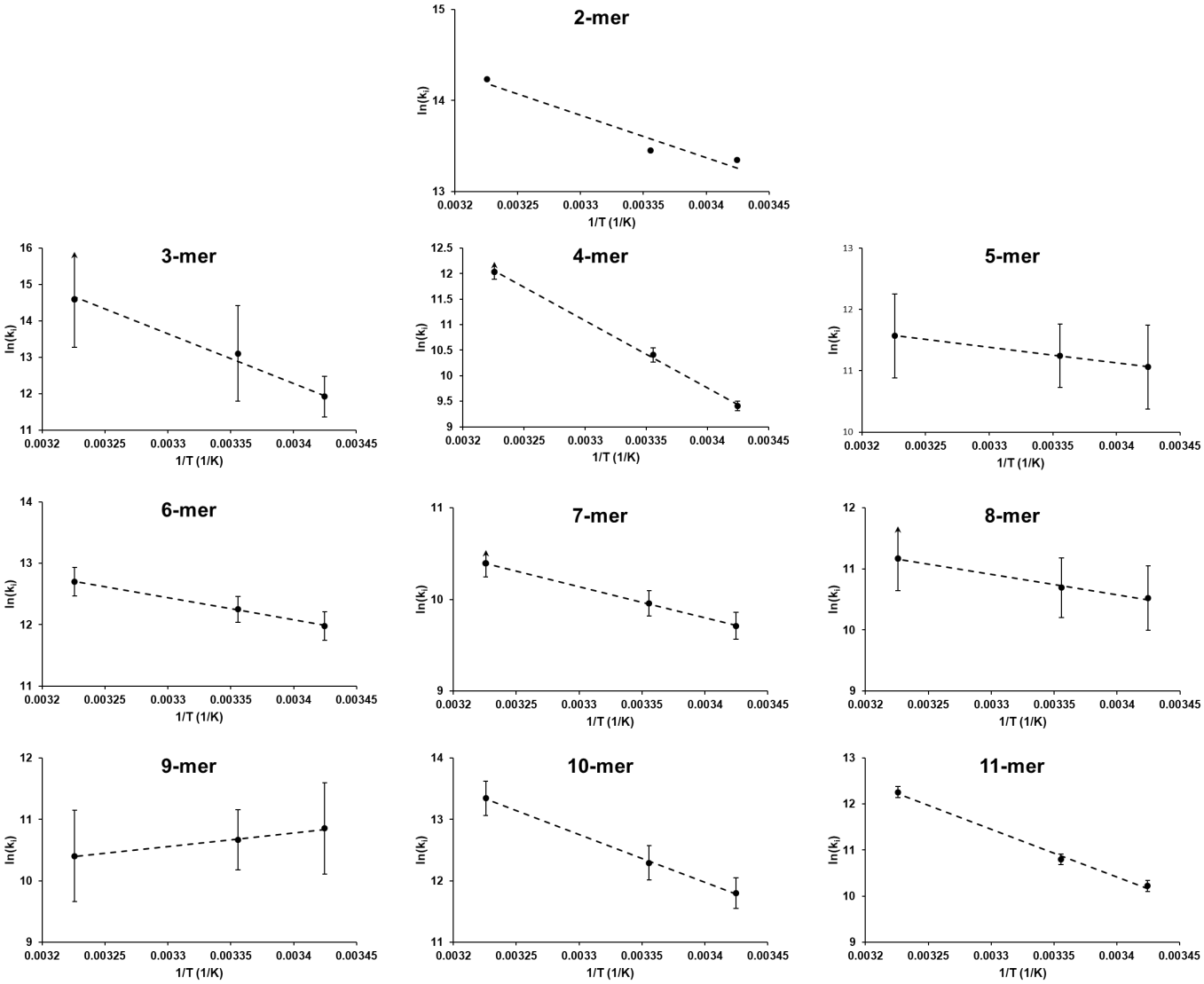

Figure S7: Temperature Dependences of Composite Initiation Rate Constants k_i: Arrhenius Plots
For each step i, values of $\ln k_i$ (with k_i in $M^{-1}s^{-1}$) are plotted vs 1/T. Uncertainties are listed in Table S4. Slopes of linear fits yield Arrhenius activation energies $E_{act,\,i}$ (slope = $-E_{act,i}/R$), reported in Table S5.

Supplemental References

1. Henderson KL*, et al.* (2019) RNA Polymerase: Step-by-Step Kinetics and Mechanism of Transcription Initiation. *Biochemistry* 58:2339-2352.
2. Henderson KL*, et al.* (2017) Mechanism of transcription initiation and promoter escape by *E. coli* RNA polymerase. *Proceedings of the National Academy of Sciences* 114(15):E3032-E3040.