Supplementary Data for

Quantitative and predictive model of kinetic regulation by *E. coli* TPP riboswitches

Sondes Guedich[1], Barbara Puffer-Enders[1], Mireille Baltzinger[2], Guillaume Hoffmann[3], Cyrielle Da Veiga[1], Fabrice Jossinet[4], Stéphane Thore[5], Guillaume Bec[1], Eric Ennifar[1], Dominique Burnouf[1*], Philippe Dumas[1*]

[1] IBMC-CNRS, Biophysique et Biologie Structurale, Architecture et Réactivité de l'ARN, Université de Strasbourg, France

[2] IBMC-CNRS, Régulations post-transcriptionnelles et nutrition, Architecture et Réactivité de l'ARN, Université de Strasbourg, France

[3] EMBL Grenoble Outstation, 71 Avenue des Martyrs, 38000 Grenoble, France

[4] IBMC-CNRS, Evolution des ARN non codants chez la levure, Architecture et Réactivité de l'ARN, Université de Strasbourg, France

[5] Université de Bordeaux, Institut Européen de Chimie et Biologie, ARNA laboratory, F33000 Bordeaux, France

***Corresponding authors:**

d.burnouf@ibmc-cnrs.unistra.fr

p.dumas@ibmc-cnrs.unistra.fr

N.B. Reference numbering corresponds to the list of references at the end of this document and is not the same as in the main text.

SD-1 K_d values of TPP as a cofactor of several enzymes

Enzyme	K_d (µM)	Organism	Remarks	Reference
1-deoxy-D-xylulose 5-phosphate synthase	**1**	*Mycobacterium tuberculosis*	K_d determined as the TPP concentration of half-maximum activity of the enzyme	Bailey et al., Glycobiology 12(2002)813
Acetolactate synthase (Acetohydroxyacid synthase)	**6**	*Pseudomonas aeruginosa*	K_d determined as the TPP concentration of half-maximum activity of the enzyme	Cho et al., Biochimie (2013) http://dx.doi.org/10.1016/j.biochi.2013.03.007
Acetolactate synthase (Acetohydroxyacid synthase)	**3.3**	*Haemophilus influenzae*	K_d determined by Trp fluorescence quenching	Choi et al., Arch. Bioch. Biophys. 466(2007)24
Phosphoketolase	**0.3**	*Lactobacillus plantarium*	K_d determined from [TPP] dependence on steady-state enzyme kinetics	Yevenes & Frey, Bioorg. Chem. 36(2008)121
Pyruvate decarboxylase	**0.45**	*Zymomonas mobilis*	K_d determined as the TPP concentration of half-maximum activity of the enzyme	Diefenbach & Duggleby, Biochem. J. 276(1991)439
Pyruvate decarboxylase	**0.35**	*Zymomonas mobilis*	K_d measurement in presence of Mg^{++}	Diefenbach & Duggleby, Biochem. J. 276(1991)439
Pyruvate decarboxylase	**2**	*Zymomonas mobilis*	Kc = 2 µM : cofactor activation constant taken as a measure of the affinity of TPP for the enzyme	Chang, Nixon & Duggleby, Biochem. 39(2000)9430
Pyruvate decarboxylase	**3.2**	*Zymomonas mobilis*	K_d determined as k_{off} / k_{on}	Chang, Nixon & Duggleby, Biochem. 39(2000)9430
Pyruvate dehydrogenase Complex	**9**	*Escherichia coli*	K_d determined by competition analysis with very affine inhibitors	Gutowski & Lienhard, JBC 251(1976)2863
Transketolase	**1**	*Escherichia coli*	Kd determined by enzymatic assay in presence of 20 µM of Mn^{2+}	Sprenger et al., Eur. J. Biochem. 230(1995)525
Transketolase	**8**	*Escherichia coli*	Kd determined in presence of 100 µM of Mg^{2+}	Sprenger et al., Eur. J. Biochem. 230(1995)525

SD-2 Temperature dependence of TPP-triggered *AThiC* folding

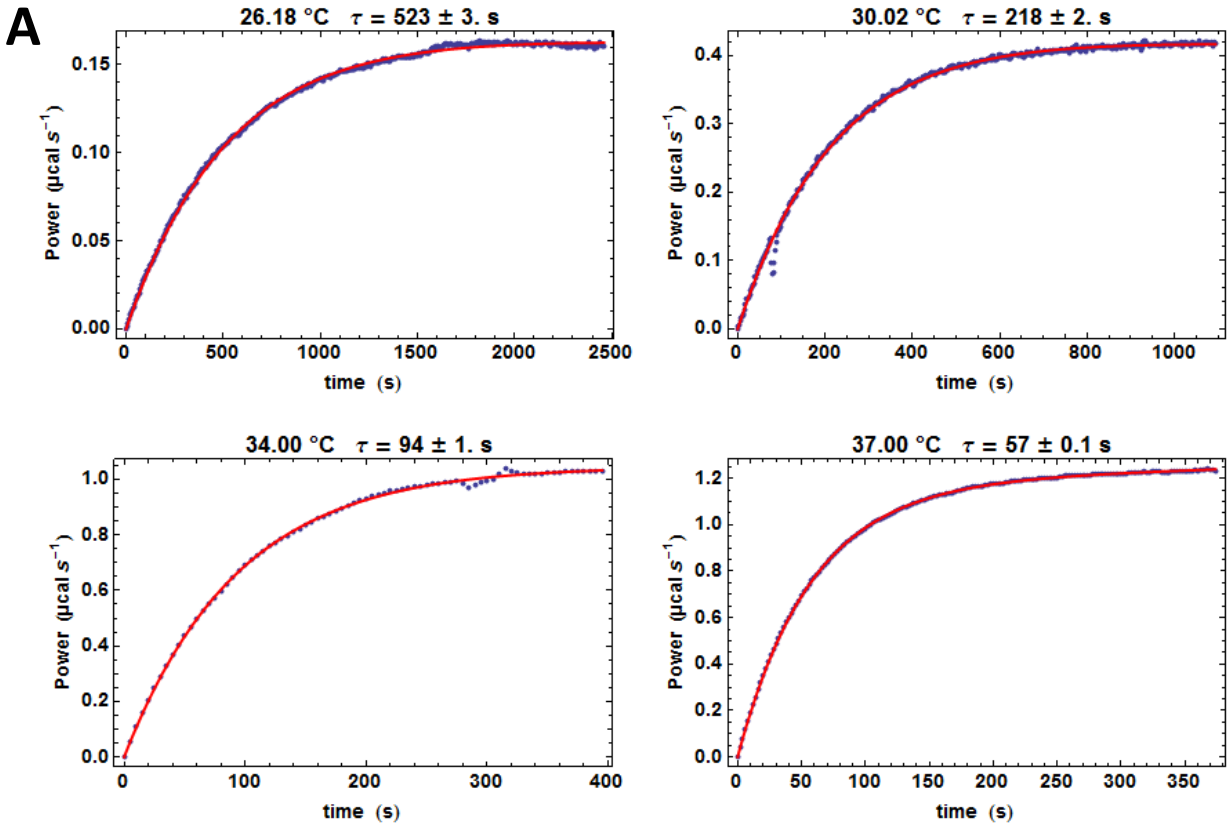


Fig. SD-2A ITC power curves at four temperatures for *AThiC*

The experimental curves were acquired in the single injection mode. The RNA and TPP concentrations in the measurement cell were 60 µM and 120 µM, respectively. The TPP was injected in 2 s. The transient part at the beginning, as well as the featureless part at the end of each injection curve, have been suppressed to keep only the part representing well the monotonous return to equilibrium following the slow component. The experimental points were fitted (red curves) with the function $\boldsymbol{P(t) = P_{max}\left(1-e^{-t/\tau}\right)+bt}$ to take into account a possible slight baseline drift (the resulting constant term $\boldsymbol{b}$ was always very small). The e.s.d of the τ values indicated at each curve were derived from the fitting of each representative part shown. These e.s.d appear to be underestimated since the variations of the τ-values resulting from modification of the representative part kept for the fitting were larger than these estimated errors. The errors taken into account for the Arrhenius plot in Fig. SD-2B are more realistic.

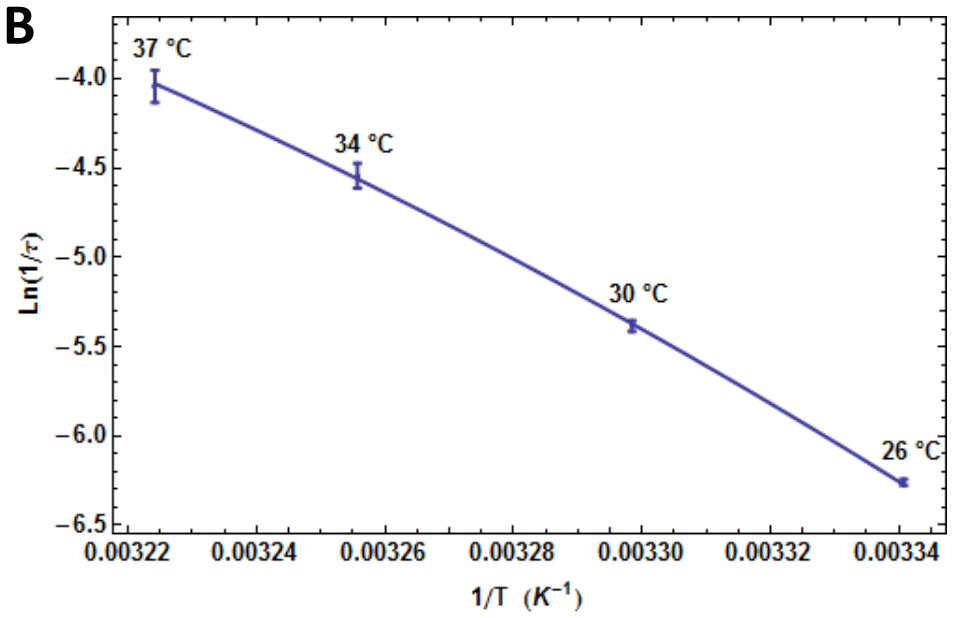


Fig. SD-2B Arrhenius plot from the results in SD-2A

The Arrhenius plot is significantly curved and the fit was performed with $\mathrm{Ln}\left(1/\tau\right)=\mathrm{Ln}\,\boldsymbol{k}_0-\left(\bar{\boldsymbol{E}}_{\boldsymbol{a}}/\boldsymbol{R}\right)\left(1/\boldsymbol{T}-1/\bar{\boldsymbol{T}}\right)-\boldsymbol{a}\left(1/\boldsymbol{T}-1/\bar{\boldsymbol{T}}\right)^2$ where $\bar{\boldsymbol{E}}_{\boldsymbol{a}}=37.8\,\mathrm{kcal\,mol^{-1}}$ is the mean activation energy between 26 and 37 °C, which corresponds to $1/\bar{\boldsymbol{T}}=\left\langle 1/\boldsymbol{T}\right\rangle$ $\left(\bar{\boldsymbol{T}}=304.7\,\boldsymbol{K},31.6\ ^{\circ}\mathrm{C}\right)$, $\boldsymbol{R}$ = 8.314 J mol^{-1} K^{-1} is the gas constant (1 cal = 4.184 J), and $\boldsymbol{k}_0$ and $\boldsymbol{a}$ are adjustable parameters. The resulting activation energy $\boldsymbol{E}_{\boldsymbol{a}}$ is thus variable with the temperature according to $\boldsymbol{E}_{\boldsymbol{a}}=\bar{\boldsymbol{E}}_{\boldsymbol{a}}+2\boldsymbol{a}\,\boldsymbol{R}\left(1/\boldsymbol{T}-1/\bar{\boldsymbol{T}}\right)$. Nonlinear regression led to $\boldsymbol{k}_0=\left(6.5\pm0.1\right)10^{-3}\,\mathrm{s^{-1}}$ (the value of $1/\tau$ at $\boldsymbol{T}=\bar{\boldsymbol{T}}$) and $\boldsymbol{a}$ = (2.8 ± 0.7)10^7 $\boldsymbol{K^2}$**.** This implies that $\boldsymbol{E}_{\boldsymbol{a}}$ varies from (44.7 ± 1.7) kcal mol^{-1} at 26 °C to (32 ± 1.7) kcal mol^{-1} at 37 °C.

SD-3 Kinetic results from •OH footprinting on *ECthiC*$_{mutP1}$ and *AThiC*

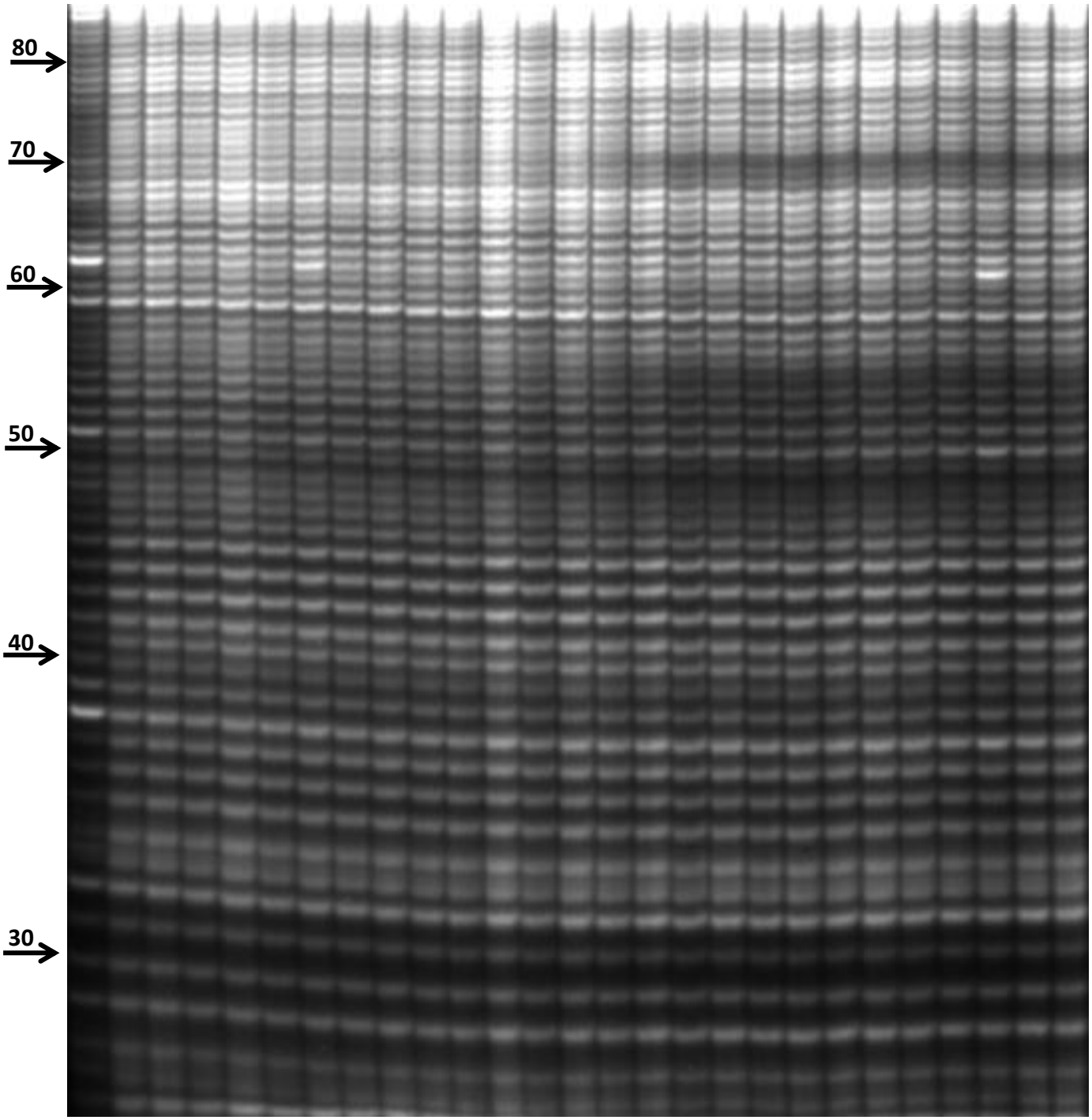


Fig. SD-3A Gel from •OH footprinting on *ECthiC*$_{mutP1}$

The two first lanes are control assays, either in the absence of H_2O_2 (lane 1) or in the absence of TPP ligand (lane 2). Conditions: 27 °C, TPP 20 µM, magnesium acetate 5 mM, potassium acetate 100 mM, Sodium cacodylate 50 mM, pH = 6.5. The lanes from the second one correspond to the times 0, 15 ms, 35, 53, 71, 89, 107, 125, 144, 167, 196, 248, 365, 662 ms and 1.5 s, 3, 6, 10, 15, 20, 25, 31, 37, 43, 50, 166 s.

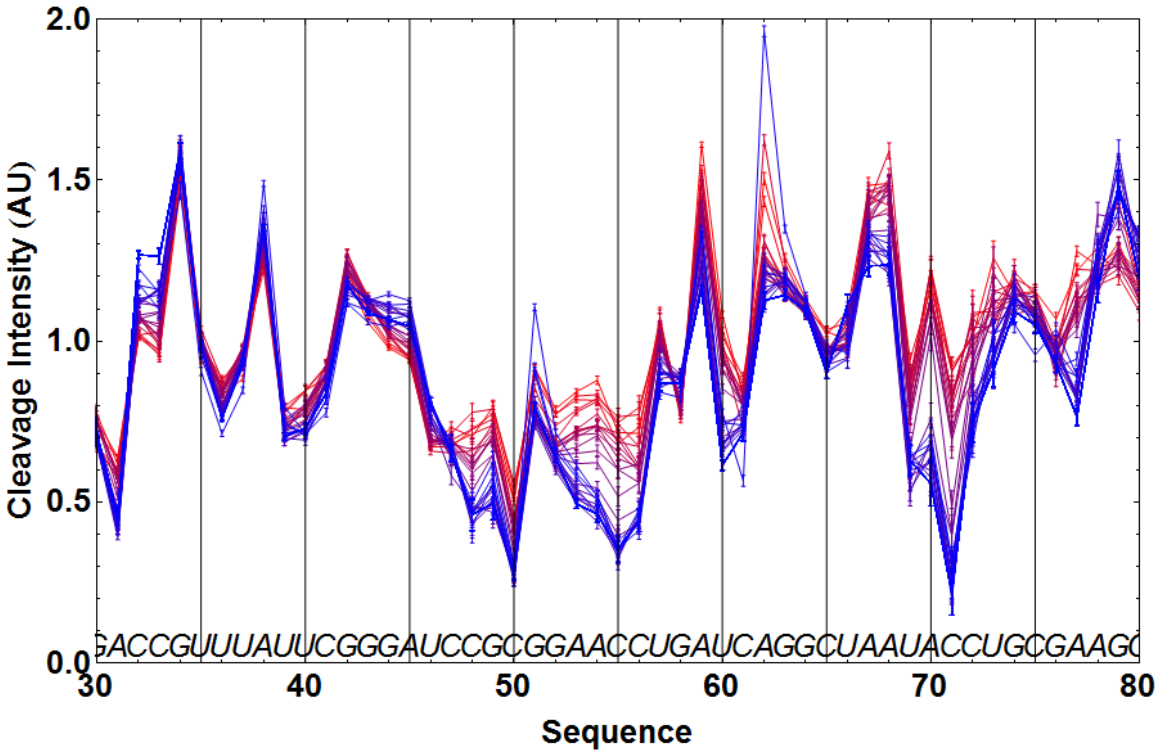


Fig. SD-3B Result of the processing of the gel in Fig. SD-3A

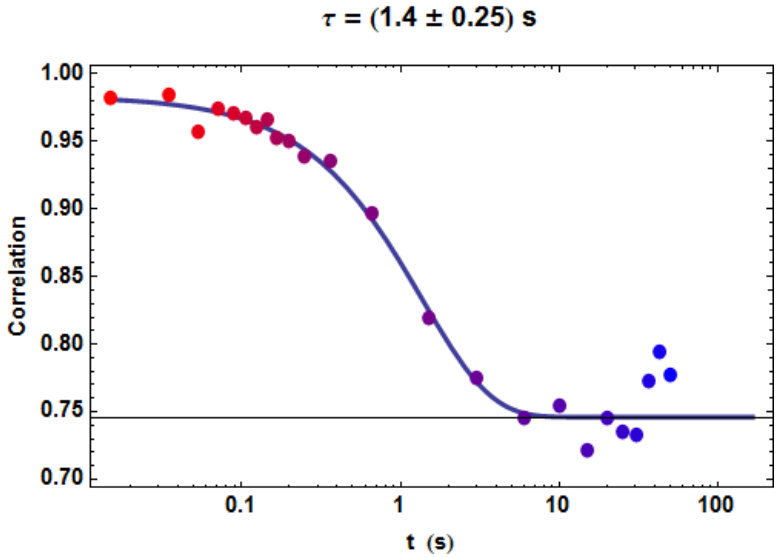


Fig. SD-3C Decorrelation of the successive cleavage curves in Fig. SD-3B

The correlation of each cleavage curve with the cleavage curve at t = 0 (pure red) is shown as a function of time. The color coding is the same as for each curve in Fig. SD-3B. The data were fully explained by an exponential fit (solid curve) with only one characteristic time τ (compare with *A. thaliana* in Fig. 3 where two distinct times were necessary). The obtained value τ = (1.4 ± 0.25 s) lies between the short and long times (0.84 and 3.2 s) obtained from equations (5a,b) with the parameters from *kinITC* in[1]. The latter values are too close to one another to be discriminated by these data.

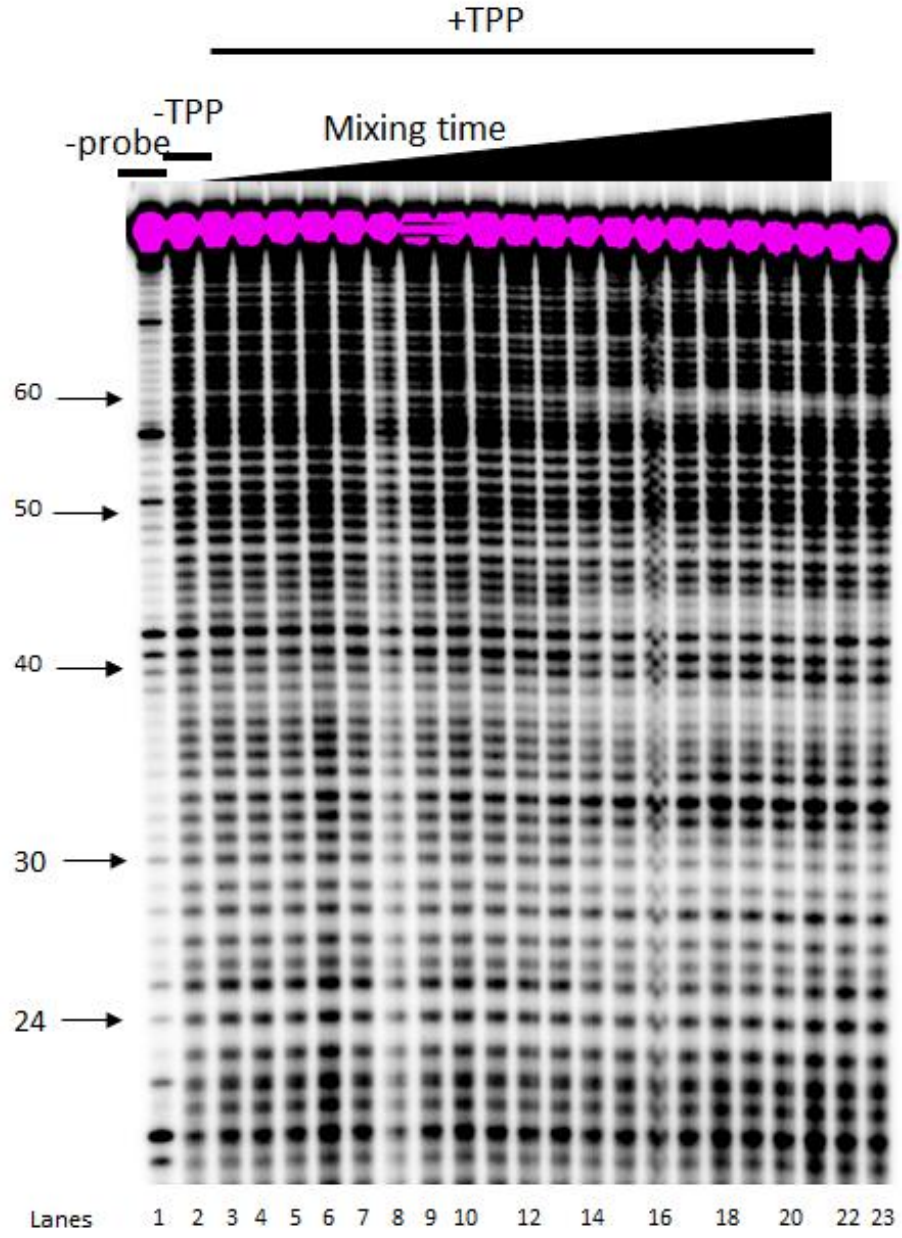


Fig. SD-3D Gel from •OH footprinting on *AThiC*.

Typical gel of •OH-radical footprinting kinetics of the TPP/aptamer complex of *A.thaliana THIC* riboswitch. The two first lanes are control assays, either in the absence of H_2O_2 (lane 1) or in the absence of TPP ligand (lane 2). Lanes 22, 23 are control experiments where RNA was renatured in the presence of TPP for 30mn at 24.5 °C. RNA and TPP concentrations were 30 nM and 20 µM, respectively, in low-salt conditions (sodium cacodylate pH 6.5, 5 mM; magnesium acetate 2 mM). Sequence numbering is indicated on the left. The times are from lane 2 to lane 23: 0, 0, 1, 2, 5, 8, 10, 13, 20, 28, 46, 95, 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 1800 s.

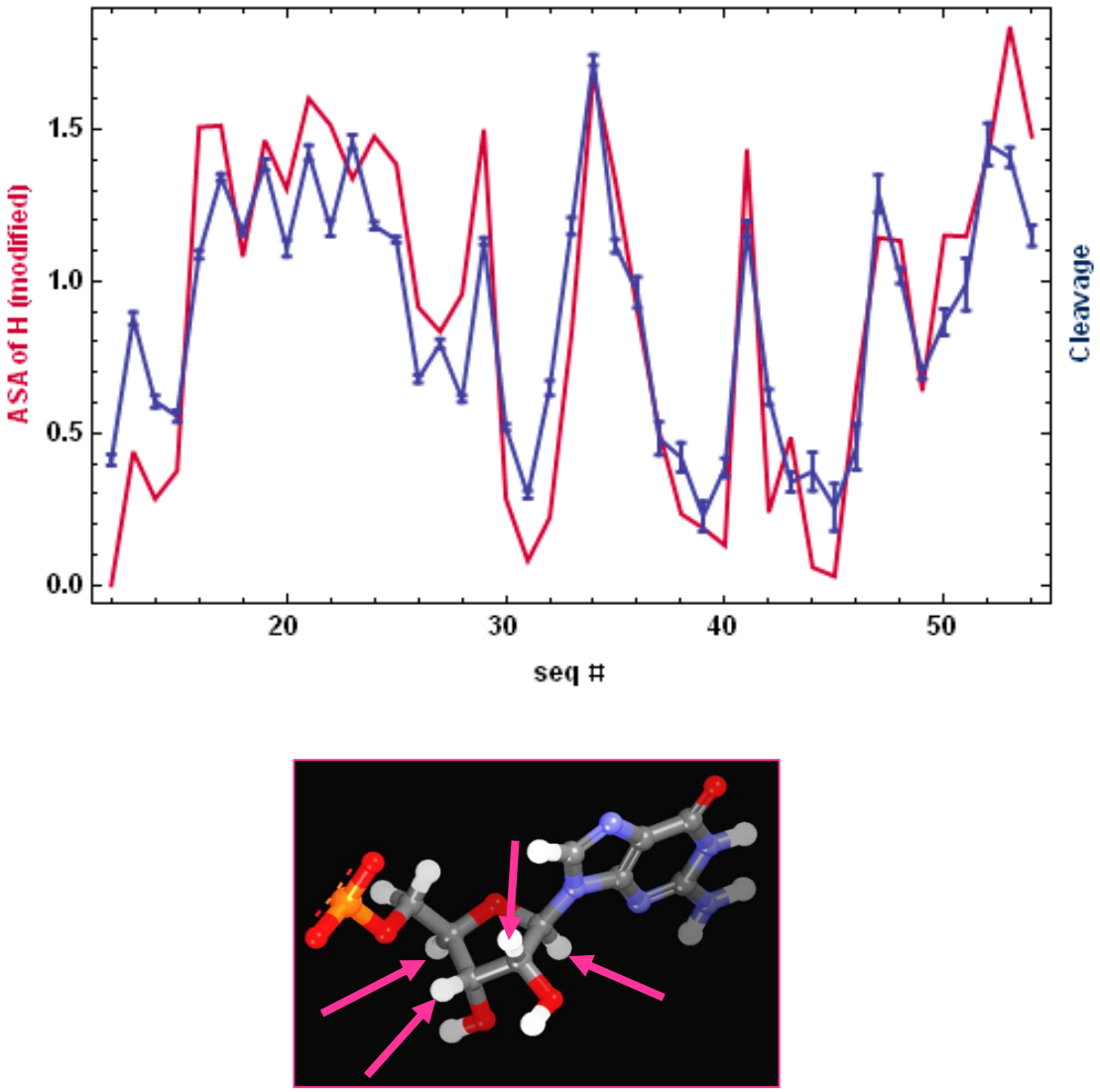


Fig. SD-3E Quality of the procedure used for gel electrophoresis quantification

The correlation between the results of hydroxyl radical cleavage quantification and the accessibility of the reactive hydrogen atoms (arrows) is shown. The blue line with error bars is the result of the quantification of a gel obtained with the THIC riboswitch aptamer from *A. thaliana* (experiment shown in Fig. 3A, main text; the cleavage, in arbitrary units, was obtained here between residues 12 and 54). The red line, corresponding to the mean accessible surface area (ASA) (in $Å^2$) of the reactive hydrogen atoms in each ribose, was calculated from the riboswitch aptamer crystal structure [2]. The modification mentioned on the left axis corresponds to a weighting of the four reactive hydrogen atoms according to their individual reactivity as described in [3]. Notably, this variable reactivity was apparent in our data. This excellent correlation was obtained by shifting the experimental cleavage curve by +1 residue, which fits with hydroxyl radical cleavage occurring within each ribose [3], and not within phosphodiester bonds. As a consequence, a cleavage within the ribose of the n^{th} residue produces a fragment of length *n*-1 on the gel.

SD-4 Kinetics of TPP binding to *AThiC* and to *ECthiC*$_{mutP1}$ by SPR

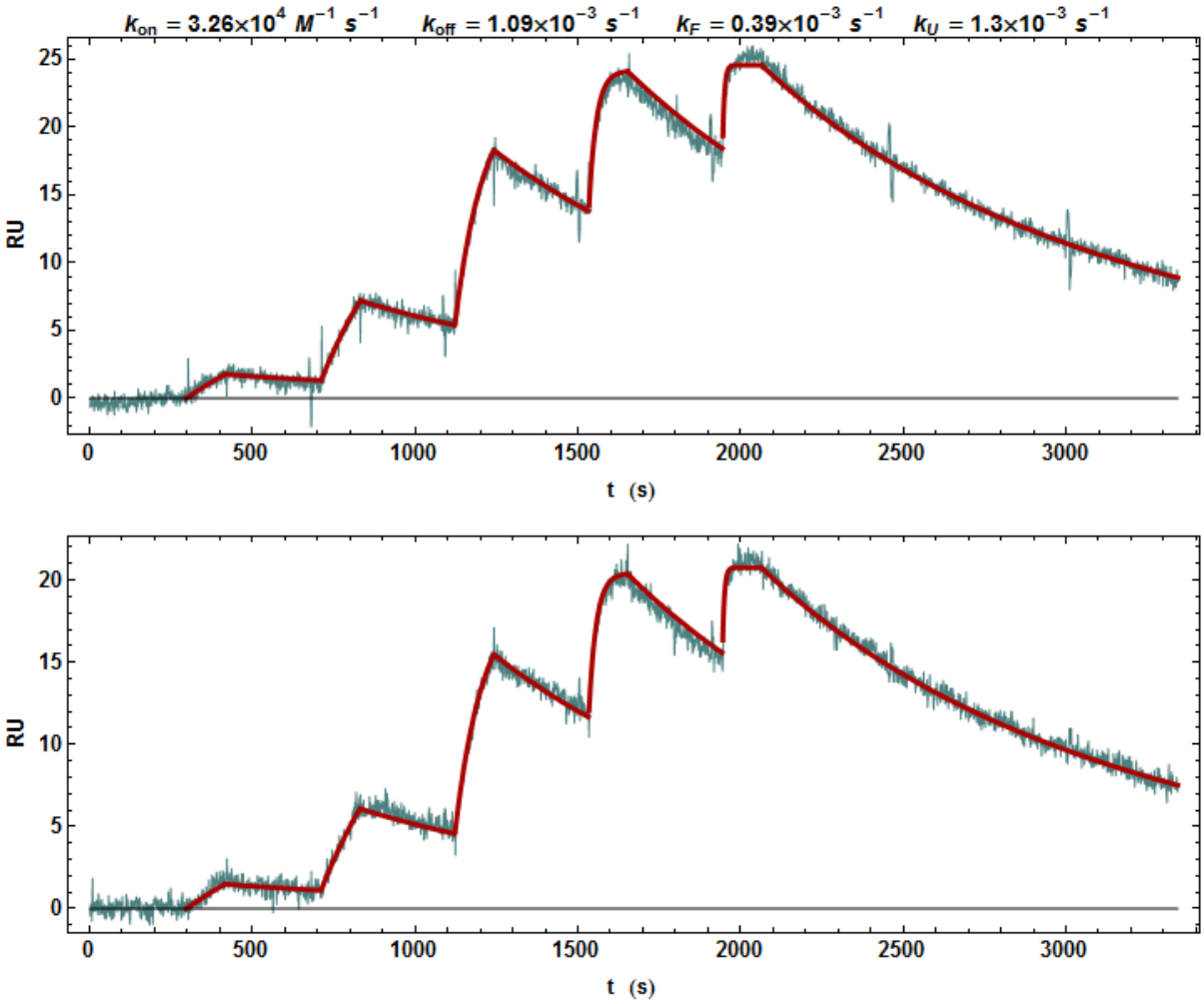


Fig. SD-4A *AThiC* SPR data (25 °C) processed with model #1

Data processing was performed by taking into account the two-step 'induced fit' mechanism of model # 1. Two independent experiments were fitted jointly (solid curves) with the same set of kinetic parameters. In each experiment the successive TPP concentrations were, from left to right, 0.020, 0.081, 0.325, 1.3 and 5.2 µM. The fit included a small drift term of the form $\exp\left(-\boldsymbol{k}_{drift}\, \boldsymbol{t}\right)$ to take into account a possible slow detachment of the RNA from the chips as well as any baseline drift. The results: $\boldsymbol{k}_{on} = \left(3.26 \pm 0.16\right) \times 10^{4}\ \mathrm{M^{-1}s^{-1}}$, $\boldsymbol{k}_{off} = \left(1.09 \pm 0.02\right) \times 10^{-3}\ \mathrm{s^{-1}}$, $\boldsymbol{k}_{F} = \left(0.39 \pm 0.03\right) \times 10^{-3}\ \mathrm{s^{-1}}$ and $\boldsymbol{k}_{U} = \left(1.30 \pm 0.08\right) \times 10^{-3}\ \mathrm{s^{-1}}$, confirm well those obtained by $^{\bullet}$OH footprinting (Fig. SD-10). The drift term was refined to a common value $\boldsymbol{k}_{drift}$ = (2.2 ± 0.4)×10^{-6} s^{-1}. The errors were estimated by repeating 3×10^5 calculations of the sum of squares of residuals with random values of the parameters around the set of parameters listed above.

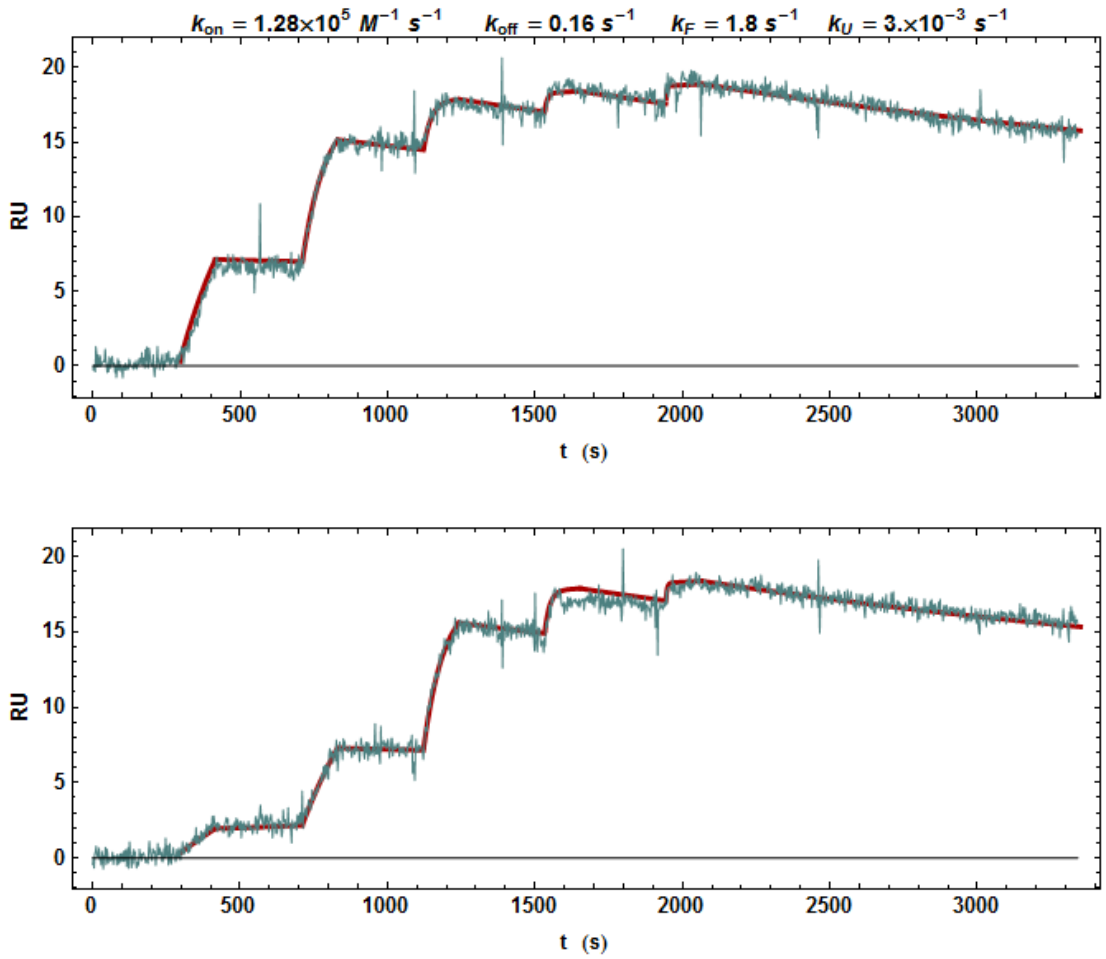


Fig. SD-4B *ECthiC$_{mutP1}$* SPR data processed with model #1

The procedure is analogous to that in Fig. SD-4A, but the successive TPP concentrations were, from left to right, 0.037, 0.11, 0.33, 1 and 3 µM for the upper curve and 0.007, 0.03, 0.125, 0.5 and 2 µM for the lower curve. The results: $k_{on} = (1.28 \pm 0.03) \times 10^5$ $M^{-1}s^{-1}$, k_{off} = (0.16 ± 0.025) s^{-1}, k_F = (1.8 ± 0.3) s^{-1}, k_U = (3 ± 0.5)×10^{-3} s^{-1}, confirm well those obtained by *kinITC* with model #1 for k_{off}, k_F and k_U (see Table 1). The drift term was refined to a common value k_{drift} = (–6.6 ± 0.8)×10^{-5} s^{-1} for the two experiments.

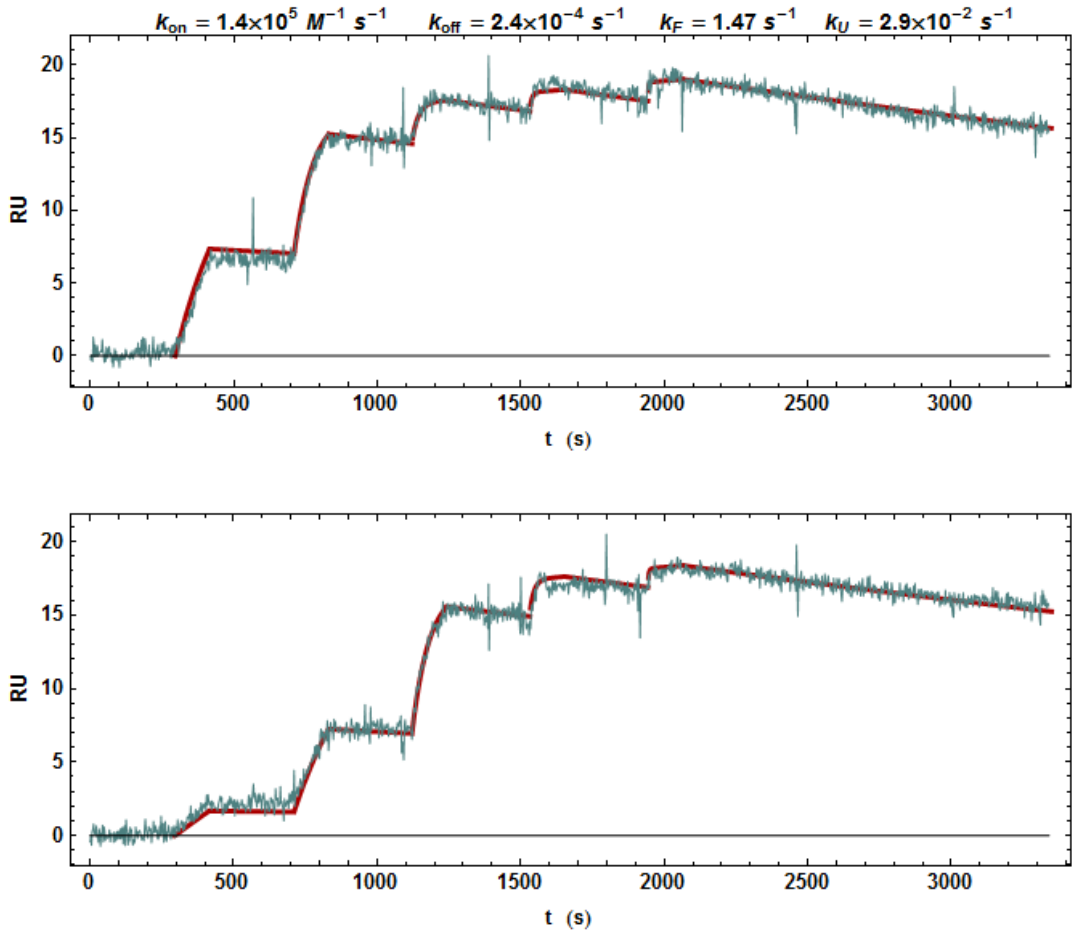
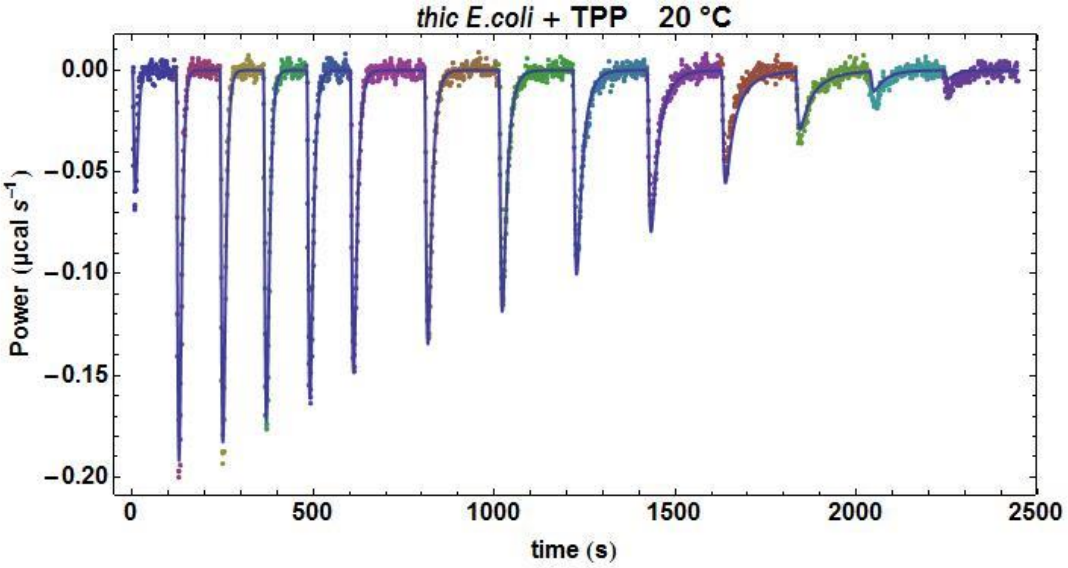
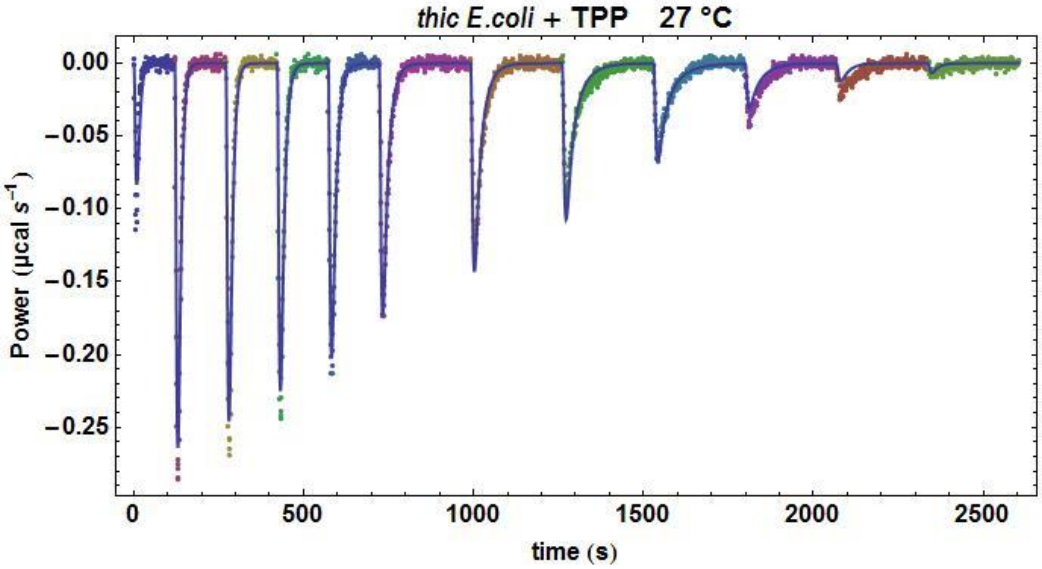


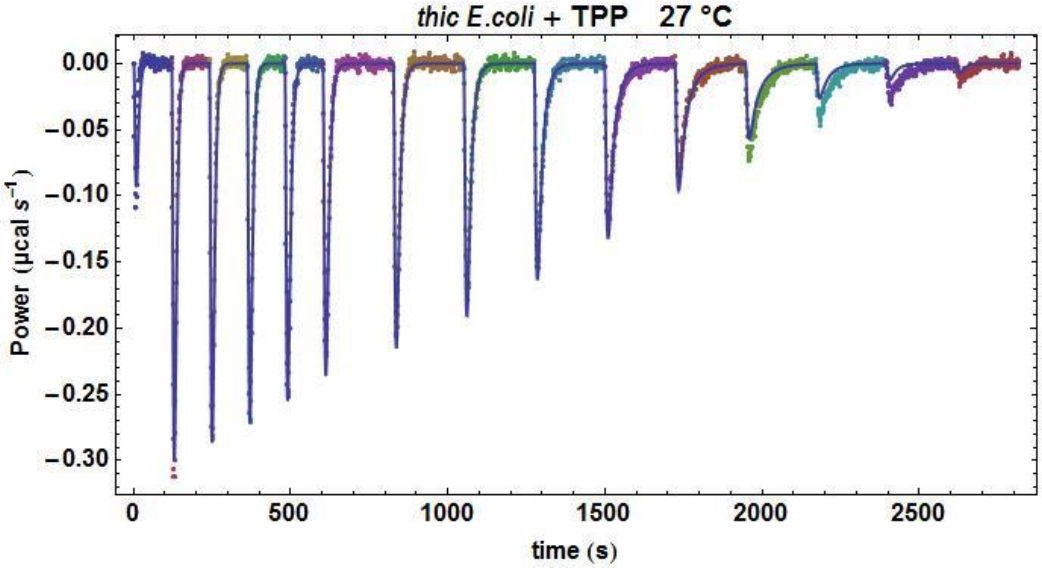
Fig. SD-4C *ECthiC$_{mutP1}$* SPR data processed with model #2

The conditions are as in Fig. SD-4B. The results: $k_{on} = (1.39 \pm 0.03) \times 10^5$ M^{-1}s^{-1}, k_{off} = (2.4 ± 0.1)×10^{-4} s^{-1}, k_F = (1.5 ± 0.35) s^{-1}, k_U = (2.9 ± 0.7)×10^{-2} s^{-1} strongly disagree with those obtained by *kinITC* with model #2 for k_F and k_U (see Table 2 in the main text). The drift terms were refined separately to k_{drift} = (–8.9 ± 0.75)×10^{-5} s^{-1} for the upper curve and (–9.5 ± 0.6)×10^{-5} s^{-1} for the lower curve.

SD-5 *kinITC* results for TPP binding to *ECthiC* and *ECthiM*







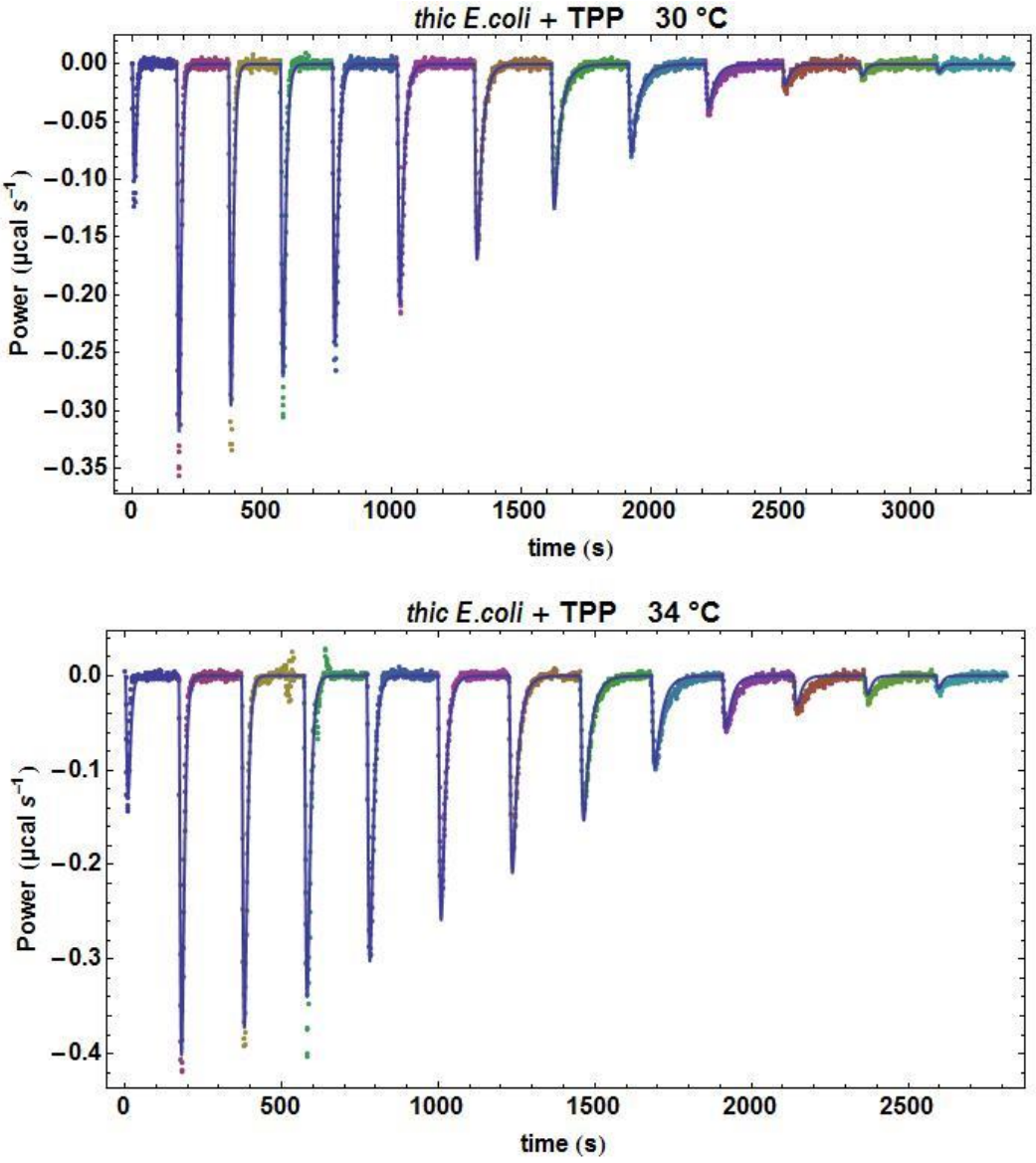


Fig. SD-5A Joint fit of injection curves at different temperatures for *ECthiC* by *kinITC*

Only the injections used in the fitting procedure for each experiment are shown. The numerical values from the fit are in SD-6A. The experiment at 27 °C was duplicated.

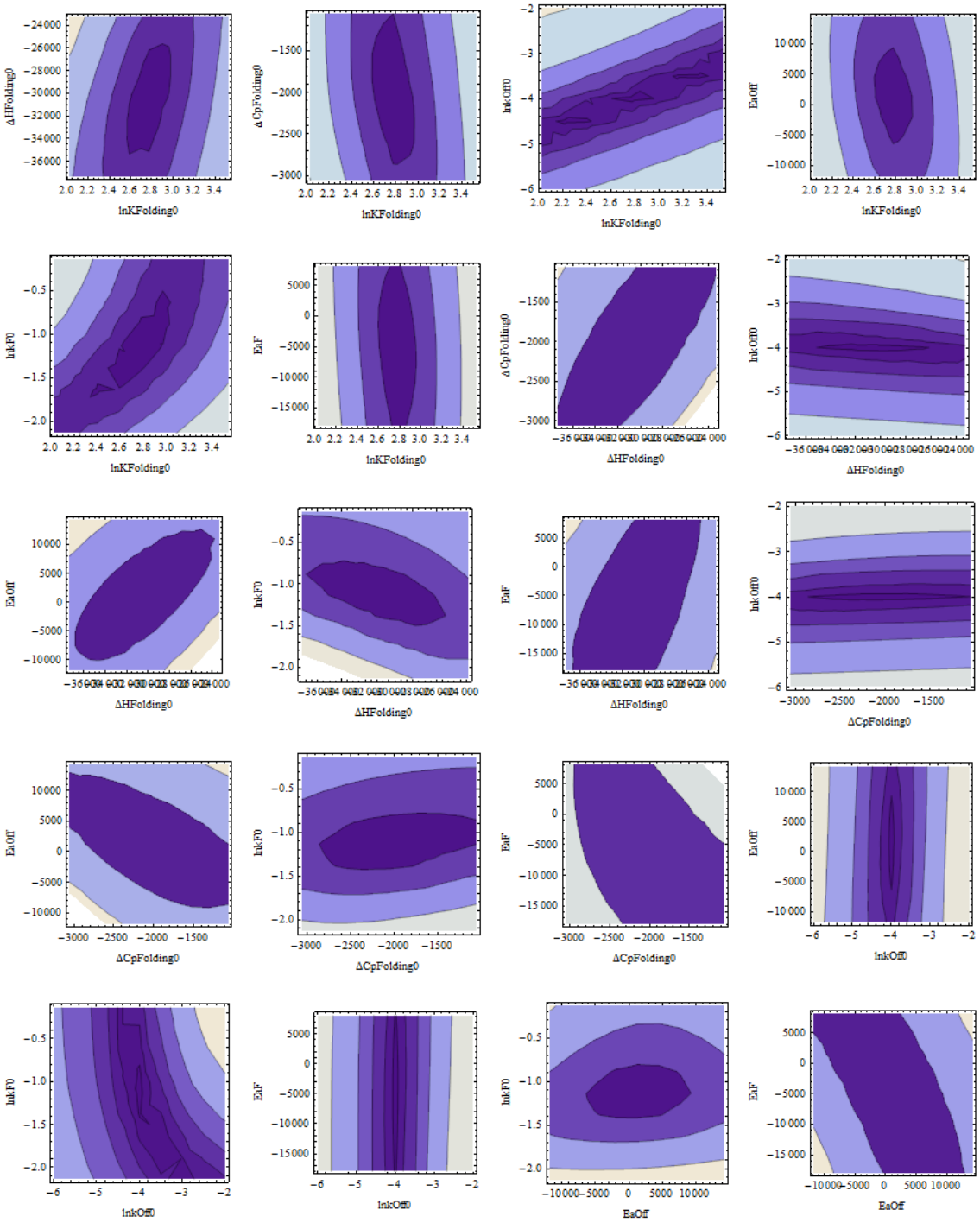
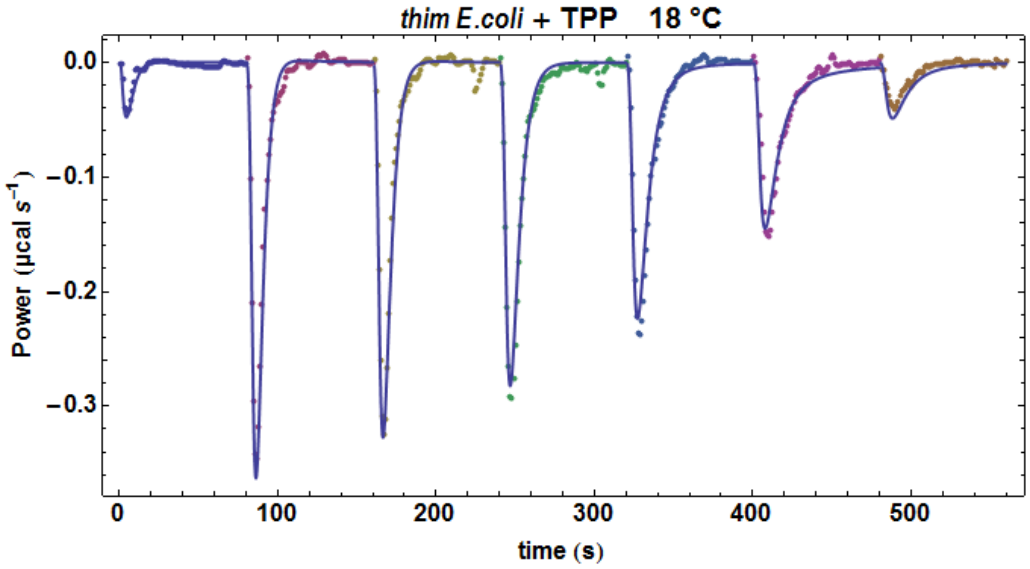
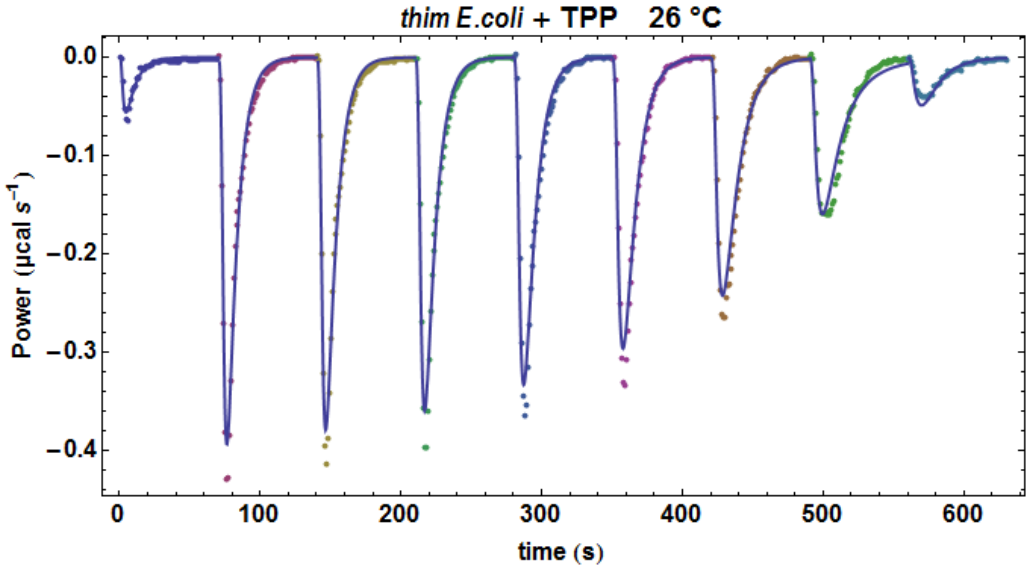
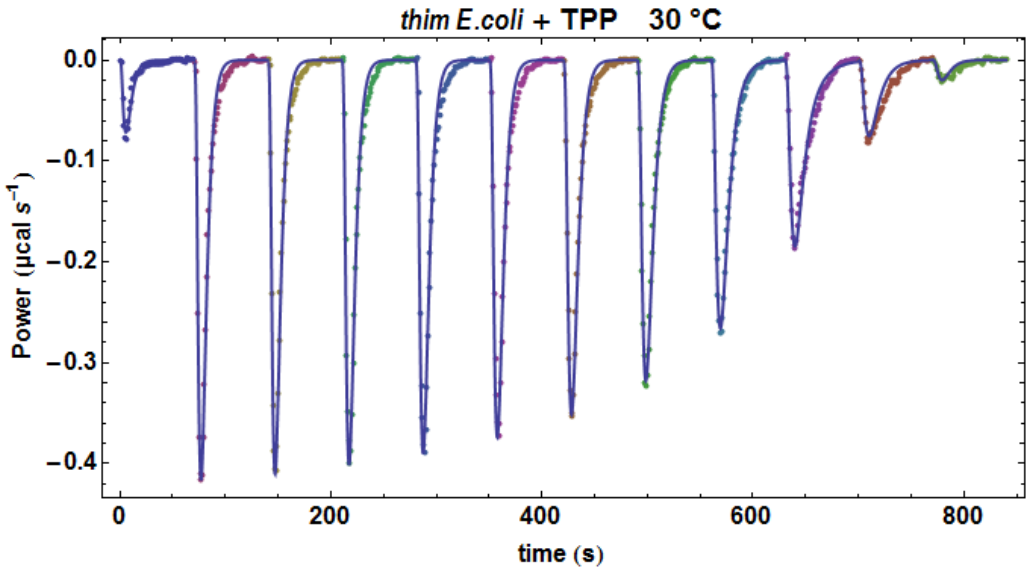


Fig. SD-5B *kinITC* residual maps for *ECthiC*

All 2D sections of the function to be minimized, the sum-of-squares of residuals (SOS), show a minimum. For all sections the minimum value of the SOS (deep-blue shading) = 0.645, the 2nd contour is at 1.32, and the outermost contour of the top-left section is at 5.48.


thim E.coli + TPP 18 °C
Power (µcal s^{-1})
0.0
−0.1
−0.2
−0.3
0
100
200
300
400
500
time (s)


thim E.coli + TPP 26 °C
Power (µcal s^{-1})
0.0
−0.1
−0.2
−0.3
−0.4
0
100
200
300
400
500
600
time (s)


thim E.coli + TPP 30 °C
Power (µcal s^{-1})
0.0
−0.1
−0.2
−0.3
−0.4
0
200
400
600
800
time (s)

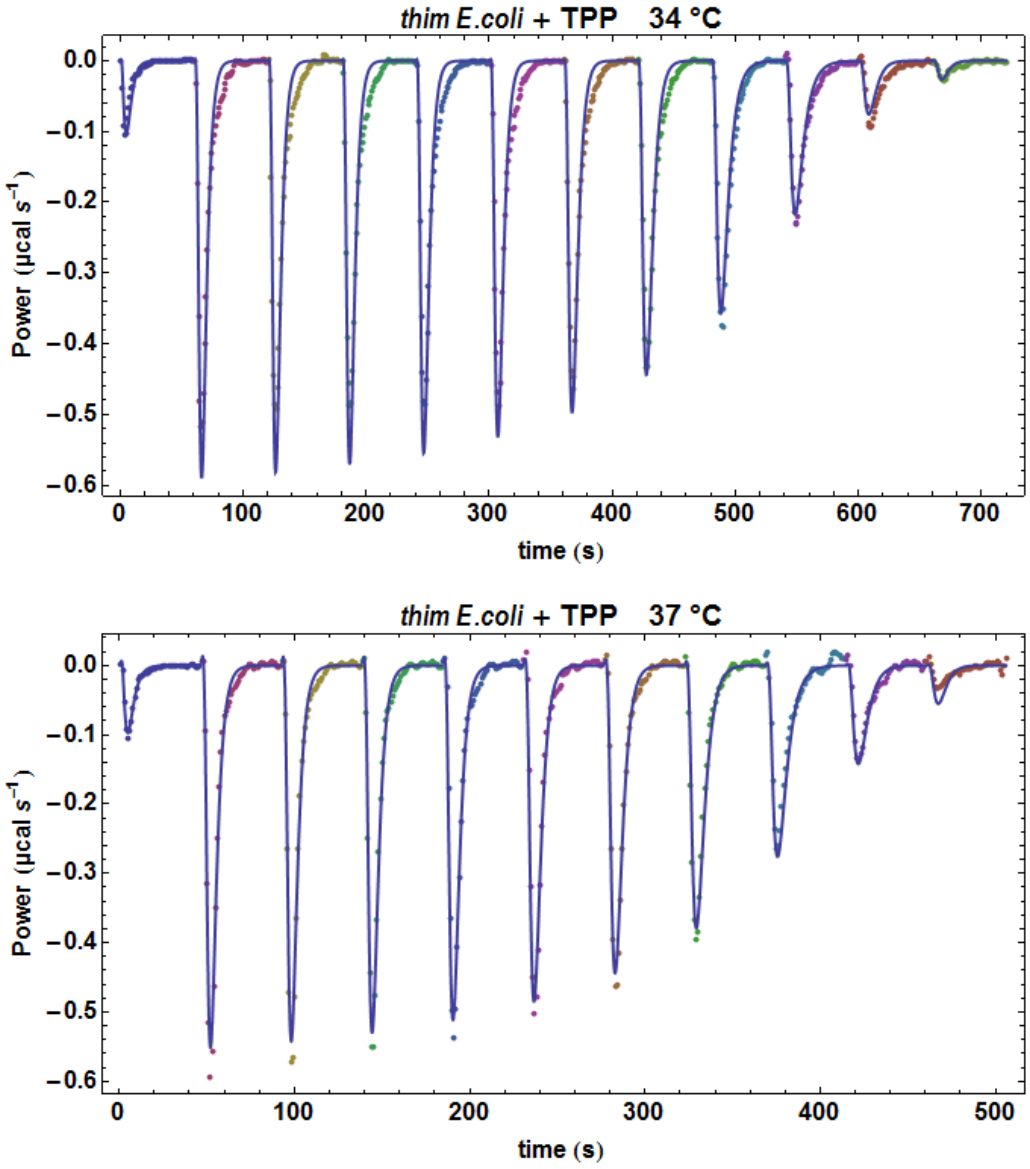


Fig. SD-5C Joint fit of injection curves at different temperatures for *ECthiM* by *kinITC*

Only the injections used in the fitting procedure for each experiment are shown. The numerical values from the fit are in SD-6B. As already noticed in [1], apart for the slowest kinetics at 18 and 26 °C, there is a slight, but systematic, lack of fit occurring in a time-range centered around 30 s after the beginning of each injection (particularly for 34 °C). Since this does not appear for *ECthiC* with a slower kinetics (Fig. SD-5A), we hypothesize that this results from imperfect account for the instrument response time.

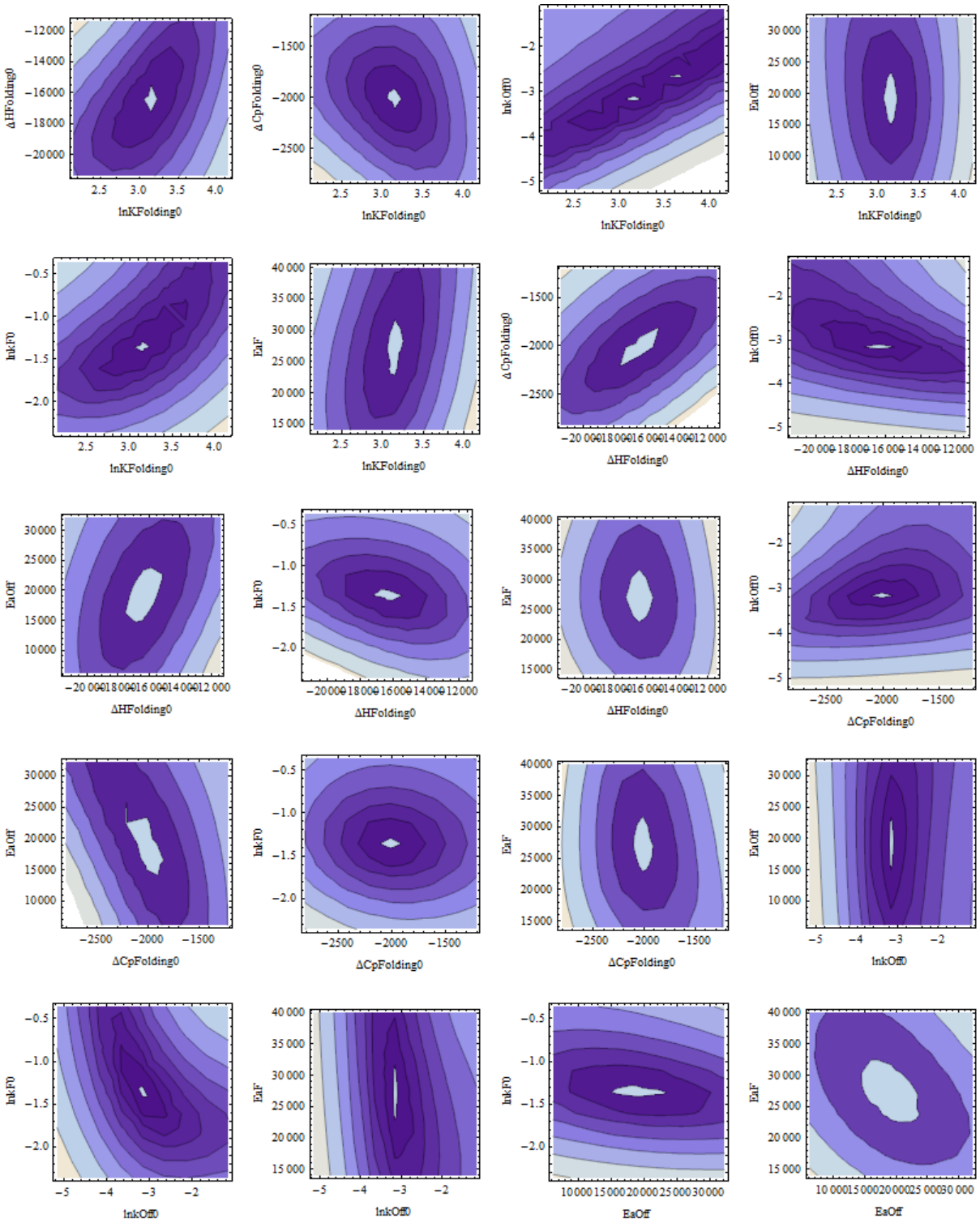


Fig. SD-5D *kinITC* residual maps for *ECthiM*.

All 2D sections of the function to be minimized, the sum-of-squares of residuals (SOS), show a well-defined minimum. For all sections the minimum value of the SOS (light-blue shading at the center) = 0.722, the 2nd contour is at 1.09, and the outermost contour of the top-left section is at 7.95.

SD-6A *ECthiC* binding parameters from *kinITC* used with model #1

The thermodynamic values marked with (ITC) correspond to overall values measured by ITC: ΔH^0(ITC) = ΔH^0(Binding) + ΔH^0(Folding); K_d^0(ITC) = K_d^0 /K^0(Folding). The superscript *0* indicates that the values correspond to the arbitrary reference temperature T^0 = 30 °C. The terms α and β refer to $\partial \Delta C_p$(ITC)/∂T and $\partial \Delta C_p$(Folding)/∂T , respectively. The standard deviations were estimated from the Hessian matrix of the sum of squares of the residuals (giving the curvatures of the SOS function at its minimum). Analogous values for $ECthiC_{mutP1}$ are in Supplementary Information of reference [1].

Thermodynamic results

ΔH^0(ITC)= –(28.2 ± 0.3) kcal mol^{-1} ΔC^0_p (ITC)= –(2142 ± 44) cal mol^{-1} K^{-1}

α = (–139 ± 4.5) cal mol^{-1} K^{-2}

ΔH^0(Folding) = –(30.3 ± 0.9) kcal mol^{-1} ΔC^0_p(Folding) = –(2055 ± 54) cal mol^{-1} K^{-1}

β = – (43 ± 1.5) cal mol^{-1} K^{-1}

K_d^0 (ITC) = (55 ± 7) nM K_d^0 = (0.9 ± 0.16) µM

K^0(Folding) = 16 ± 1

Kinetic results

k_{on}^0 = (21 ± 5.7) × 10^3 M^{-1} s^{-1} $\Delta H_{on}^{\ddagger}$ = $\Delta H_{off}^{\ddagger}$ + ΔH^0(Binding) = (3.3 ± 0.9) kcal mol^{-1}

k_{off}^0 = (0.018 ± 0.0017) s^{-1} $\Delta H_{off}^{\ddagger}$ = (1.2 ± 0.04) kcal mol^{-1}

k_F^0 = (0.32 ± 0.01) s^{-1} $\Delta H_F^{\ddagger}$ = (–4.9 ± 0.2) kcal mol^{-1}

k_U^0 = (0.02 ± 0.002) s^{-1} $\Delta H_U^{\ddagger}$ = $\Delta H_F^{\ddagger}$ – ΔH^0(Folding) = (25.4 ± 1) kcal mol^{-1}

SD-6B *ECthiM* binding parameters from *kinITC* used with model #1

The notations are as in SD-6A with the same reference temperature T^0 = 30 °C. The values of ΔH^0(ITC) and K_d^0(ITC) agree well with those in reference [4].

Thermodynamic results

ΔH^0(ITC)= –(18.8 ± 0.3) kcal mol^{-1} ΔC^0_p (ITC)= –(693 ± 13) cal mol^{-1} K^{-1}

α = (21.2 ± 0.6) cal mol^{-1} K^{-2}

ΔH^0(Folding) = –(16.4 ± 0.3) kcal mol^{-1} ΔC^0_p(Folding) = –(2012 ± 52) cal mol^{-1} K^{-1}

β = – (95.3 ± 3) cal mol^{-1} K^{-1}

K_d^0 (ITC) = (19.1 ± 2.5) nM K_d^0 = (0.45 ± 0.1) µM

K^0(Folding) = 23.3 ± 2.1

Kinetic results

k_{on}^0 = (9.5 ± 2.9) × 10^4 M^{-1} s^{-1} $\Delta H_{on}^{\ddagger} = \Delta H_{off}^{\ddagger} + \Delta H^0$(Binding) = (16.7 ± 0.7) kcal mol^{-1}

k_{off}^0 = (0.042 ± 0.0035) s^{-1} $\Delta H_{off}^{\ddagger}$ = (19.2 ± 0.6) kcal mol^{-1}

k_F^0 = (0.26 ± 0.01) s^{-1} $\Delta H_F^{\ddagger}$ = (27.0 ± 0.75) kcal mol^{-1}

k_U^0 = (0.011 ± 0.0015) s^{-1} $\Delta H_U^{\ddagger} = \Delta H_F^{\ddagger} - \Delta H^0$(Folding) = (43.4 ± 0.8) kcal mol^{-1}

SD-7 Thermodynamic and kinetic results from *kinITC* for *ECthiC*

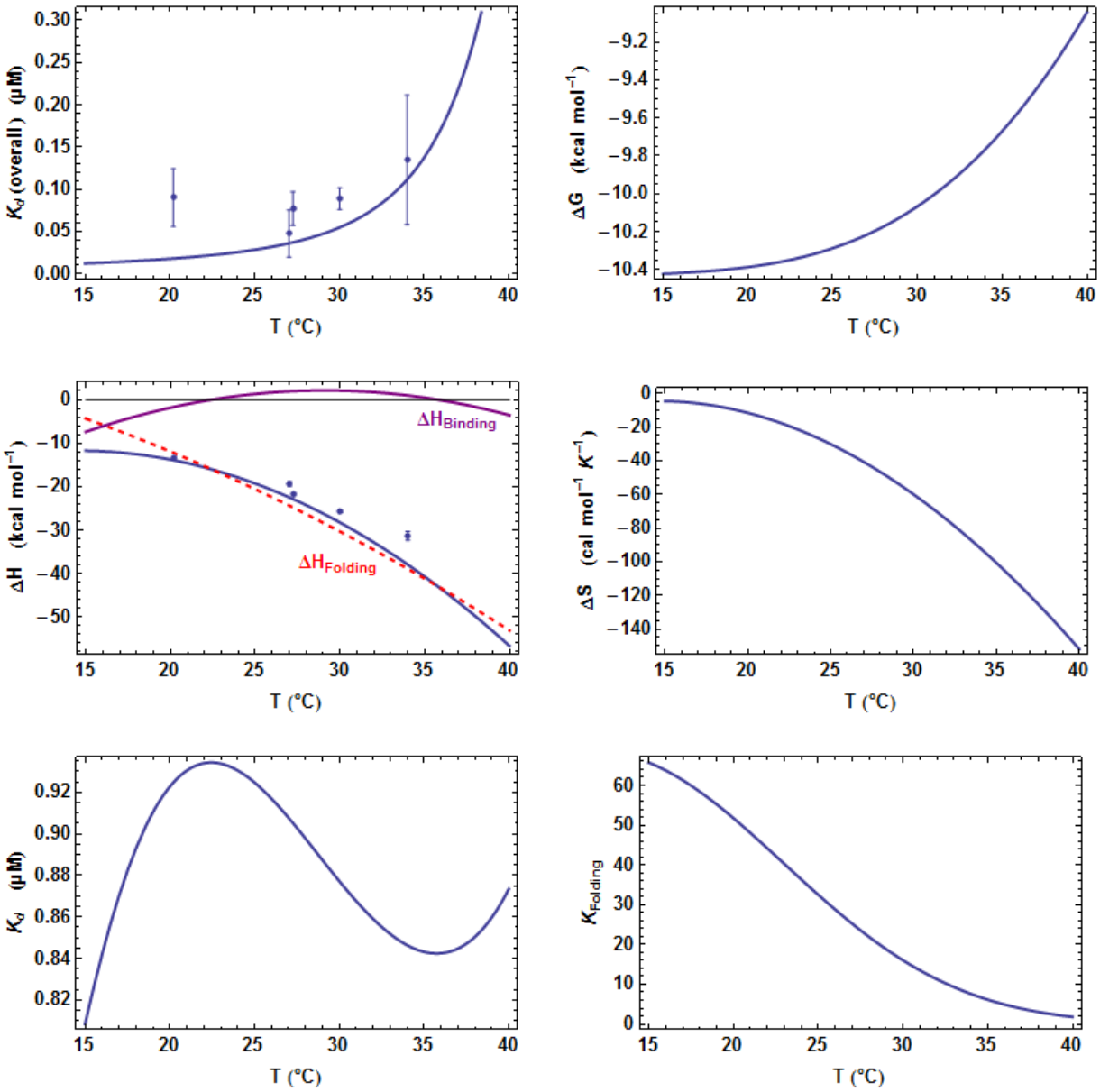


Fig. SD-7A Thermodynamic results from *kinITC* for *ECthiC*

The evolution with the temperature of all thermodynamic parameters of the induced-fit mechanism (equations 1a,b) was obtained with *kinITC* used with model #1. The figures for ΔG and ΔS refer to the complete reaction involving TPP binding and RNA folding. The evolution of K_d(overall) is shown along with the individual values obtained by normal use of ITC at five different temperatures (blue dots). The discrepancy with the continuous curve from *kinITC* is illustrative of the overall level of errors. The figure for ΔH collects $\Delta H_{Binding}$, $\Delta H_{Folding}$ (red dashed curve) and $\Delta H_{ITC} = \Delta H_{Binding} + \Delta H_{Folding}$ shown in blue with the individual values obtained by normal use of ITC (as for $\boldsymbol{K_d}$(overall)). Note the extrema of $\boldsymbol{K_d}$ reached at the temperatures where $\Delta H_{Binding} = 0$ in agreement with the van 't Hoff equation linking $\boldsymbol{K_d}$(intial binding) and $\Delta H_{Binding}$. Such an unusual feature should not be taken at face value and is merely representative of the fact that $\boldsymbol{K_d}$ does not vary significantly in the temperature range considered.

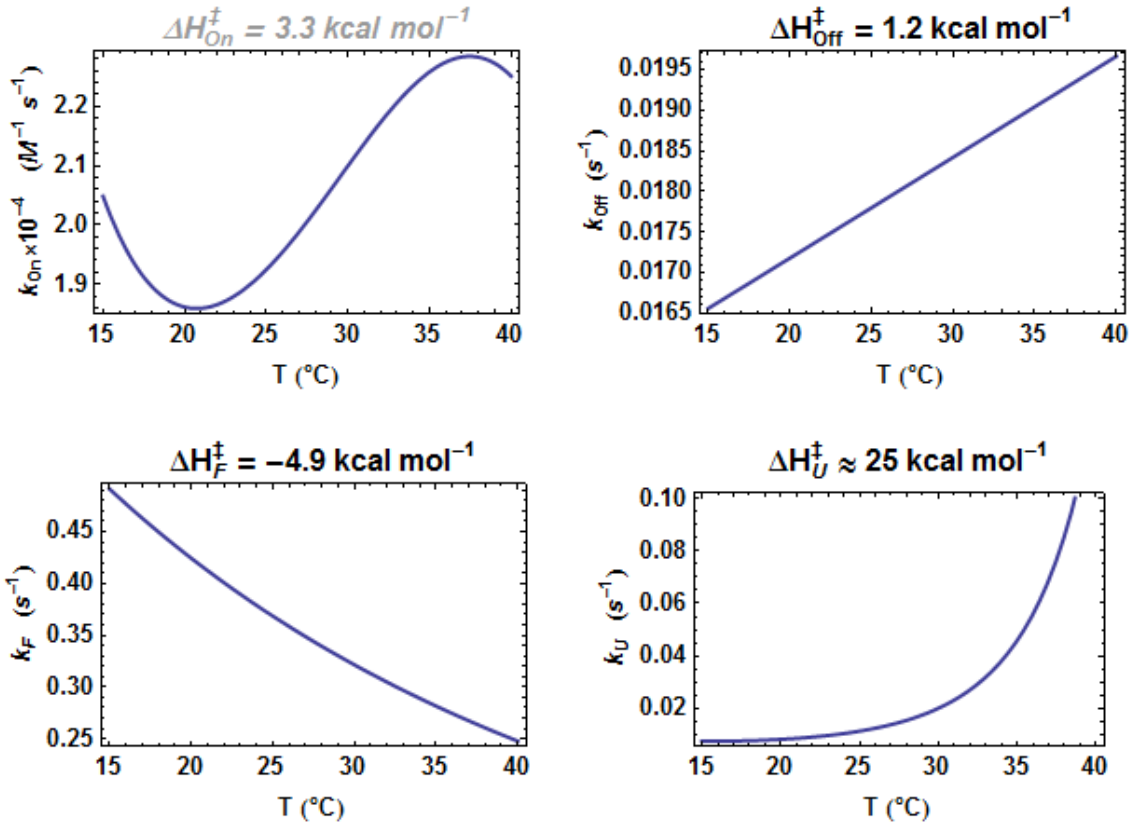


Fig. SD-7B Kinetic results from *kinITC* for *ECthiC*

The evolution with the temperature of all kinetic parameters in equations 1a,b was obtained with *kinITC* used with model #1. The activation energies $\Delta H^{\ddagger}_{off}$ and $\Delta H^{\ddagger}_{F}$ were obtained as refinable parameters in *kinITC*, and $\Delta H^{\ddagger}_{on}$ and $\Delta H^{\ddagger}_{U}$ were deduced from $\Delta H^{0}_{Binding} = \Delta H^{\ddagger}_{on} - \Delta H^{\ddagger}_{off}$ and $\Delta H^{0}_{Folding} = \Delta H^{\ddagger}_{F} - \Delta H^{\ddagger}_{U}$, respectively. The undulation of k_{on} should not be taken at face value too, since it simply reflects the undulation of $\boldsymbol{K}_d$ in Fig. SD-7A and the fact that k_{off} is almost constant in the temperature range considered. The amplitudes of undulations are thus likely representative of the overall level of errors in the *kinITC* procedure. For this reason $\Delta H^{\ddagger}_{on}$ = 3.3 kcal mol^{-1} is shown in gray.

SD-8 Kinetic equations governing the evolution of RNA concentrations

Although most of the points described in this section are known, we recall them in a comprehensive way to introduce clearly the original features. We note $\boldsymbol{L}_{\text{tot}}$ and $\boldsymbol{R}_{\text{tot}}$ the total TPP and RNA concentrations. If the ligand is in large excess $\left(\boldsymbol{L}_{\text{tot}} >> \boldsymbol{R}_{\text{tot}} \right)$, kinetic model #1 leads from equations (1a,b) to the pseudo first-order equations:

$$\frac{d\boldsymbol{R}_1}{\boldsymbol{dt}} = \boldsymbol{k_{on}} \boldsymbol{L}_{\text{tot}} \boldsymbol{R}_0 - \left(\boldsymbol{k_{off}} + \boldsymbol{k_F} \right) \boldsymbol{R}_1 + \boldsymbol{k_U} \boldsymbol{R}_2 \qquad \text{(S8.1a)}$$

$$\frac{d\boldsymbol{R}_2}{\boldsymbol{dt}} = \boldsymbol{k_F} \boldsymbol{R}_1 - \boldsymbol{k_U} \boldsymbol{R}_2 \qquad \text{(S8.1b)}$$

with $\boldsymbol{R}_0, \boldsymbol{R}_1, \boldsymbol{R}_2$ being short notations for the time-dependent concentrations $[\boldsymbol{R}_0](\boldsymbol{t}), [\boldsymbol{R}_1](\boldsymbol{t}), [\boldsymbol{R}_2](\boldsymbol{t})$. By considering the conservation equation $\boldsymbol{R}_0 = \boldsymbol{R}_{\text{tot}} - \boldsymbol{R}_1 - \boldsymbol{R}_2$, equation (S8.1a) can be expressed in terms of $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ only, which yields a classical system of two coupled linear differential equations in $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$. It is common to write this system under the matrix form:

$$\begin{bmatrix} \dot{\boldsymbol{R}}_1 \\ \dot{\boldsymbol{R}}_2 \end{bmatrix} = \begin{bmatrix} -\left(\boldsymbol{k_{on}} \boldsymbol{L}_{\text{tot}} + \boldsymbol{k_{off}} + \boldsymbol{k_F} \right) & \boldsymbol{k_U} - \boldsymbol{k_{on}} \boldsymbol{L}_{\text{tot}} \\ \boldsymbol{k_F} & -\boldsymbol{k_U} \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_1 \\ \boldsymbol{R}_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{k_{on}} \boldsymbol{L}_{\text{tot}} \boldsymbol{R}_{\text{tot}} \\ 0 \end{bmatrix} \qquad \text{(S8.1c)}$$

where the dots on $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ stand for time derivatives. An analogous matrix equation would be obtained for kinetic model #2. It is well known that the general solution of such a system involves linear combinations of two exponential functions:

$$\boldsymbol{R}_1(\boldsymbol{t}) = \tilde{\boldsymbol{R}}_1 + \boldsymbol{f}_1 \exp(-\boldsymbol{k_{fast}}\, \boldsymbol{t}) + \boldsymbol{s}_1 \exp(-\boldsymbol{k_{slow}}\, \boldsymbol{t}) \qquad \text{(S8.2a)}$$

$$\boldsymbol{R}_2(\boldsymbol{t}) = \tilde{\boldsymbol{R}}_2 + \boldsymbol{f}_2 \exp(-\boldsymbol{k_{fast}}\, \boldsymbol{t}) + \boldsymbol{s}_2 \exp(-\boldsymbol{k_{slow}}\, \boldsymbol{t}) \qquad \text{(S8.2b)}$$

where $\tilde{\boldsymbol{R}}_1$ and $\tilde{\boldsymbol{R}}_2$ are constant terms corresponding respectively to the concentrations of $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ at equilibrium (*i.e.* the concentrations such that $\boldsymbol{dR}_1 / \boldsymbol{dt} = \boldsymbol{dR}_2 / \boldsymbol{dt} = 0$), where the coefficients $\boldsymbol{f_i}, \boldsymbol{s_i}$ (i=1,2) are time-independent parameters and where $\boldsymbol{k_{slow}}, \boldsymbol{k_{fast}}$ (with $\boldsymbol{k_{slow}} < \boldsymbol{k_{fast}}$) are kinetic parameters governing, respectively, a 'slow' and a 'fast' component

of concentration variations or, equivalently, a 'long' and a 'short' relaxation time $\left(\tau_{long} = \boldsymbol{k}_{slow}^{-1}, \tau_{short} = \boldsymbol{k}_{fast}^{-1}\right)$. In the following, we give explicit expressions of all terms necessary for computing $\boldsymbol{R}_2(\boldsymbol{t})$ from the kinetic parameters, since $\boldsymbol{R}_2(\boldsymbol{t})$ is essential to express the probability of a riboswitch like *ECThiC* and *ECthiM* to be in the ON- or OFF-state.

• **Determination of $\tilde{\boldsymbol{R}}_1, \tilde{\boldsymbol{R}}_2$ (model #1)**

By solving $\boldsymbol{dR}_1 / \boldsymbol{dt} = 0$ and $\boldsymbol{dR}_2 / \boldsymbol{dt} = 0$ in equations (S8.1a,b), it is obtained:

$$\tilde{\boldsymbol{R}}_1 = \frac{\boldsymbol{k}_{on}\, \boldsymbol{k}_U\, \boldsymbol{L}_{\text{tot}}\, \boldsymbol{R}_{\text{tot}}}{\boldsymbol{k}_{off}\, \boldsymbol{k}_U + \boldsymbol{k}_{on}\left(\boldsymbol{k}_F + \boldsymbol{k}_U\right)\boldsymbol{L}_{\text{tot}}} \qquad \tilde{\boldsymbol{R}}_2 = \frac{\boldsymbol{k}_{on}\, \boldsymbol{k}_F\, \boldsymbol{L}_{\text{tot}}\, \boldsymbol{R}_{\text{tot}}}{\boldsymbol{k}_{off}\, \boldsymbol{k}_U + \boldsymbol{k}_{on}\left(\boldsymbol{k}_F + \boldsymbol{k}_U\right)\boldsymbol{L}_{\text{tot}}} \qquad \text{(S8.3a)}$$

The latter are transformed with equilibrium constants only ($\boldsymbol{K}_d = \boldsymbol{k}_{off} / \boldsymbol{k}_{on}$, $\boldsymbol{K}_F = \boldsymbol{k}_F / \boldsymbol{k}_U$) as:

$$\tilde{\boldsymbol{R}}_1 = \boldsymbol{R}_{\text{tot}}\left(1 + \boldsymbol{K}_F + \boldsymbol{K}_d \boldsymbol{L}_{\text{tot}}^{-1}\right)^{-1} \qquad \tilde{\boldsymbol{R}}_2 = \boldsymbol{R}_{\text{tot}} \boldsymbol{K}_F \left(1 + \boldsymbol{K}_F + \boldsymbol{K}_d \boldsymbol{L}_{\text{tot}}^{-1}\right)^{-1} \qquad \text{(S8.3b)}$$

To express $\boldsymbol{P}_{ON}(\boldsymbol{L}_{\text{tot}})$, the probability of a riboswitch to be in the ON-state as a function of $\boldsymbol{L}_{\text{tot}}$, it will be necessary to express $\tilde{\boldsymbol{R}}_2$ in molar fraction of the total RNA species rather than in concentration, and it will thus be used:

$$\tilde{\boldsymbol{R}}_2 = \boldsymbol{K}_F \left(1 + \boldsymbol{K}_F + \boldsymbol{K}_d \boldsymbol{L}_{\text{tot}}^{-1}\right)^{-1} \qquad \text{(S8.3c)}$$

• **Determination of $\boldsymbol{k}_{slow}, \boldsymbol{k}_{fast}$ (models #1 and #2)**

Practically, $-\boldsymbol{k}_{slow}$ and $-\boldsymbol{k}_{fast}$ are obtained as the eigenvalues of the square matrix (equation S8.1c) for kinetic model #1. Calculation of these eigenvalues yields for both models (see reference [5], pp. 118-119, for an example corresponding to model #1):

$$\boldsymbol{k}_{fast \text{ or } slow} = \frac{1}{2}\left[\boldsymbol{k}_U + \boldsymbol{k}_F + \boldsymbol{k}_{on}\left(\boldsymbol{L}_{\text{tot}} + \boldsymbol{K}_d\right)\right]\left(1 \pm \sqrt{1 - \frac{4\boldsymbol{k}_{on}\boldsymbol{k}_U\, \boldsymbol{K}_{\text{model}}}{\left[\boldsymbol{k}_U + \boldsymbol{k}_F + \boldsymbol{k}_{on}\left(\boldsymbol{L}_{\text{tot}} + \boldsymbol{K}_d\right)\right]^2}}\right) \qquad \text{(S8.4a)}$$

where $\boldsymbol{k}_{fast}$ and $\boldsymbol{k}_{slow}$ correspond, respectively, to the plus and minus sign within the parenthesis, and where the term $\boldsymbol{K}_{\text{model}}$ is model-dependent according to:

$$K_{\text{model \#1}} = K_d + (1 + K_F) L_{\text{tot}} \quad \text{(S8.4b)}$$

$$K_{\text{model \#2}} = K_d + K_F (L_{\text{tot}} + K_d) \quad \text{(S8.4c)}$$

• **Determination of f_2, s_2 (model #1)**

Only f_2 and s_2 for model #1 allowing to express the probability of the *ECthiC* and *ECthiM* riboswitches to be in the ON- or OFF-state are of interest in this work. Following Fersht [5], we introduce the short notations: $p = k_F + k_U + k_{off} + k_{on} L_{\text{tot}}$, $r = k_{off} k_U + k_{on} (k_F + k_U) L_{\text{tot}}$ and $q = (p^2 - 4r)^{1/2}$, which leads to the condensed forms $k_{fast} = (p + q)/2$ and $k_{slow} = (p - q)/2$ from equation (S8.4a). If no ligand is bound at *t* = 0, *i.e.* $R_1(0) = R_2(0) = 0$, which is the only situation of interest in this study, it may be shown (with the help of a software with symbolic capabilities like *Mathematica*):

$$f_2 = k_F k_{on} L_{\text{tot}} k_{slow} (q r)^{-1} > 0 \quad \text{(S8.5a)}$$

$$s_2 = -k_F k_{on} L_{\text{tot}} k_{fast} (q r)^{-1} < 0 \quad \text{(S8.5b)}$$

The remarkably simple relationship (valid for any value of L_{tot}) may then be derived:

$$\frac{f_2}{s_2} = -\frac{k_{slow}}{k_{fast}} \quad \text{(S8.5c)}$$

We are not aware of whether or not equations S8.5a-c have been mentioned elsewhere. Equation (S8.5c) means that the faster the fast component in comparison of the slow component, the smaller its amplitude in comparison of that of the slow component. It results that $f_2 \exp(-k_{fast}\, t)$ **becomes rapidly negligible in comparison of** $s_2 \exp(-k_{slow}\, t)$ due to both $k_{fast} > k_{slow}$ and $|f_2| < |s_2|$. Numerical calculations showed that, for *ECthiM* and *ECthiC*, the maximum possible value of k_{slow}/k_{fast} is 0.6 and that it is less, or much less, than 0.25 at all temperatures in the ligand concentration range of the ON/OFF-switch. This is at the basis of extremely useful simplifications to obtain an explicit expression for the ligand concentration of the ON/OFF switch (SD-12).

SD-9 Evolution of the kinetic parameter k_{slow} with the ligand concentration

In all situations of interest, $k_{slow}\left(L_{tot}\right)$ (equation S8.4a) increases and reaches a plateau for sufficiently large values of the ligand concentration L_{tot} (Fig. SD-9). It is important to know what 'sufficiently large' means since a kinetic regulation is possible only when $k_{slow}\left(L_{tot}\right)$ varies significantly with L_{tot} (*i.e.* roughly on the left side of the vertical line in Fig. SD-9). The results will be given without proofs. Here, we first consider situations where $k_{off} \ll k_F + k_U$, which is relevant for *ECthiC* and *ECthiM*. This hypothesis will then be released to obtain one result of interest for *AThiC*.

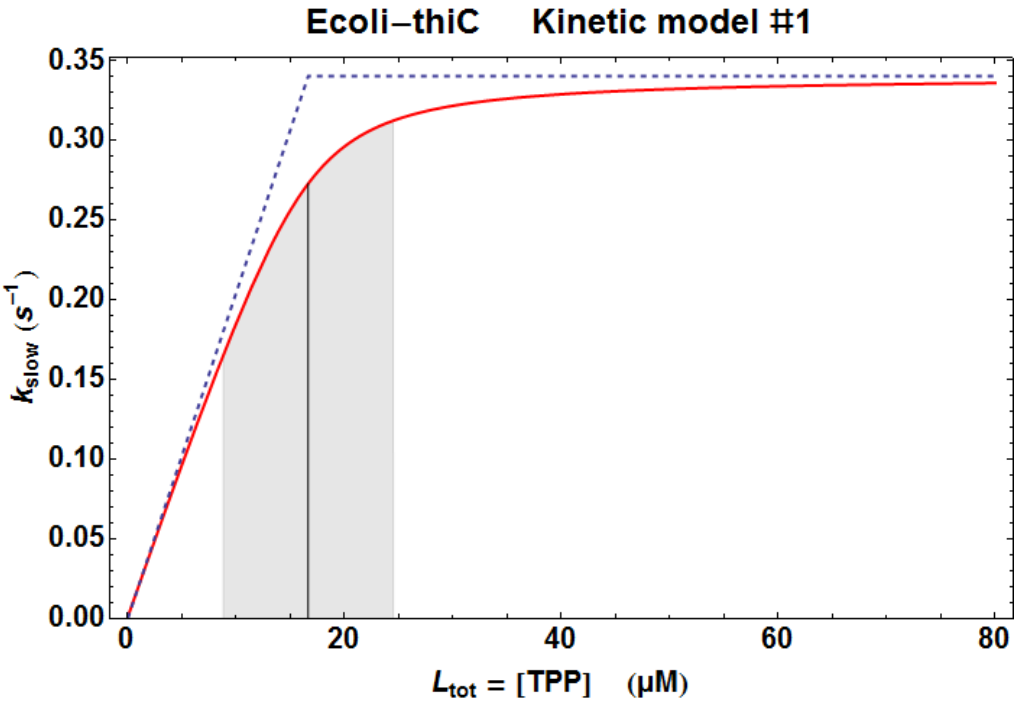


Fig. SD-9 Essential features of the evolution of k_{slow} *vs.* L_{tot} (equations S8.4a,b).

The function $k_{slow}\left(L_{tot}\right)$ (Fig. SD-9, red curve) was calculated with model #1 ('induced fit') with the kinetic parameters at 30 °C for *ECthiC* (SD-6A). In the general case (respecting the hypothesis $k_{off} \ll k_F + k_U$), the vertical line defined by the intersection of the two dashed lines and marking the transition zone (grey dashing) between possible and impossible kinetic regulation lies at $L_{tot} = \left(k_F + k_U\right)/k_{on}$ (which is valid for both kinetic models). The plateau (horizontal dashed line) lies at $k_F + k_U$ for model #1 and at k_F for model #2**.** Accordingly,

the linear dependence of k_{slow} *vs.* L_{tot} (ascending dashed line) is $k_{slow} = k_{on} L_{tot} + k^0_{slow}$ for model #1 and $k_{slow} = \left[k_{on} k_F / \left(k_F + k_U \right) \right] L_{tot} + k^0_{slow}$ for model #2. A clear experimental example representative of k_{slow} for model #2 may be seen in reference [6].

Equation S8.4a does not extrapolate to $k_{slow} = 0$ for $L_{tot} \to 0$ (*i.e.* $k^0_{slow} \neq 0$), which is not visible in Fig. SD-9 because, in that particular case, the extrapolated value k^0_{slow} is negligible. However, experimental data in the literature, *e.g.* Fig. 8b in reference[6] and Fig. 4a in reference[7], show well that $k^0_{slow} \neq 0$. The theoretical extrapolated value is $k^0_{slow} = 1/2 \left(k_{Sum} - \sqrt{k_{Sum}{}^2 - 4 k_U k_{off}} \right)$ for model #1 (with $k_{Sum} = k_F + k_U + k_{off}$) and, for model #2, k^0_{slow} is equal to the smaller of the two values k_{off} and $k_F + k_U$. Since it is assumed here that $k_{off} << k_F + k_U$, a good approximation of $1/2 \left(k_{Sum} - \sqrt{k_{Sum}{}^2 - 4 k_U k_{off}} \right)$ for model #1 may be shown to be k_{off} / K_F and, for model #2, $\min(k_{off}, k_F + k_U) = k_{off}$. In conclusion, the linear regimes at the basis of kinetic regulation are given by:

Model #1: $k_{slow} \left(L_{tot} \right) \approx k_{off} / K_F + k_{on} L_{tot}$ (S9.1)

Model #2: $k_{slow} \left(L_{tot} \right) \approx k_{off} + \left[k_{on} k_F / \left(k_F + k_U \right) \right] L_{tot}$ (S9.2)

Equation S9.1 is applicable to *ECthiC* and *ECthiM* and will be used in SD-12.

Finally, to fit with the values of the kinetic parameters obtained for *AThiC*, we release the hypothesis $k_{off} << k_F + k_U$. For model #1, the slope of the linear portion of $k_{slow} \left(L_{tot} \right)$ decreases from k_{on} for $k_{off} \ll k_F + k_U$ (as assumed previously), to $1/2\, k_{on}$ when $k_{off} = k_F + k_U$. For *AThiC*, it was precisely obtained $k_{off} \approx k_F + k_U$ with the values $k_{off} = 1.1 \times 10^{-3}\, s^{-1} \approx k_F + k_U = 1.7 \times 10^{-3}\, s^{-1}$ obtained by SPR at 25 °C (Fig. SD 4A). The vertical line separating the two regimes in Fig. SD-9 is then displaced from $\left(k_F + k_U \right) / k_{on}$ to *ca.* $2 \left(k_F + k_U \right) / k_{on}$. With $k_{on} = 3.3 \times 10^4\, M^{-1} s^{-1}$ (Fig. SD 4A), this corresponds to a low TPP concentration (around 0.1 µM). This is in agreement with the fact that no significant variation of k_{slow} was observed for *AThiC* with TPP concentrations varying from 2 µM to higher concentrations and with the absence of kinetic regulation for *AThiC*.

SD-10 Quantification of the kinetic curves from hydroxyl radical footprinting

Hydroxyl-radical footprinting provides us with a kinetic cleavage curves for each residue (as in Fig. 3B). Our goal here is of obtaining as much information as possible from these cleavage curves by making use of a kinetic model giving the evolution of the concentrations of the different RNA species $\boldsymbol{R}_0(\boldsymbol{t}), \boldsymbol{R}_1(\boldsymbol{t}), \boldsymbol{R}_2(\boldsymbol{t})$ (equations 1a,b). The TPP being in large excess, its concentration may be assumed to be constant.

The variation of the overall cleavage $\boldsymbol{C_n(t)}$ observed at the n^{th} residue is obtained by summing the specific cleavage of each RNA species at this position: $\boldsymbol{dC_n/dt} = \sum \alpha_i^n \boldsymbol{R_i(t)}$. In this sum, the α_i^n are constant terms quantifying, for a given $^{\bullet}$OH concentration, the propensity of the species $\boldsymbol{R_i}$ to be cleaved at the n^{th} residue (there are thus three α_i^n for each residue #*n*). Since each function $\boldsymbol{R_i(t)}$ is known from the kinetic model (equation (4) in the main text or S8.2 in this document), $\boldsymbol{C_n(t)}$ is readily obtained by integration of $\boldsymbol{dC_n/dt} = \sum \alpha_i^n \boldsymbol{R_i(t)}$. In fact, because the time interval $\delta \boldsymbol{t}$ left for the cleavage is very short ($\delta \boldsymbol{t}$ = 15 ms), $\boldsymbol{R_i(t)}$ may be considered constant during the cleavage if $\boldsymbol{k_{fast}}\, \delta \boldsymbol{t} \ll 1$, and it results $\boldsymbol{C_n(t)} \approx \delta \boldsymbol{t} \sum_i \alpha_i^n \boldsymbol{R_i(t)}$ (which is equivalent to the form considered *a priori* in reference [8]). The three α_i^n for each residue have to be determined to fit all cleavage curves as those in Fig. 3B. However, in some way, the real unknowns in this method are the parameters of the kinetic model, not the α_i^n which are obtained 'for free' by linear regression when the kinetic model has allowed obtaining the three functions $\boldsymbol{R_i(t)}$. For considering the results as physically meaningful and not merely as corresponding to a numerical fit without significance, all α_i^n have to be positive or, at most, slightly negative for few of them. This was always the case. This global processing method is interesting for two reasons. First, supplemented with $\boldsymbol{K_d}$(overall) obtained by fully independent means like ITC or SPR (which constrained the kinetic parameters), it allowed fitting at once **all** cleavage curves with very few kinetic parameters, namely the three parameters $\boldsymbol{k_{on}}$ (or $\boldsymbol{k_{off}}$), and $\boldsymbol{k_F}$ and $\boldsymbol{k_U}$. In practice, it appeared that the exact value of $\boldsymbol{K_F} = \boldsymbol{k_F}/\boldsymbol{k_U}$ had little influence and, by fixing it in a first

step to some reasonable value, only two free parameters were sufficient: k_{on} (or k_{off}), and the sum $k_F + k_U$. The second interest of this method is of yielding rich kinetic and spatial information about TPP binding through the set of α_i^n, which allowed us to determine the primary site of TPP interaction.

An illustration of the quality of fit of the $^{\bullet}$OH cleavage curves (Fig. 3B) by this kinetic-model-based method is given in Fig. SD-10.

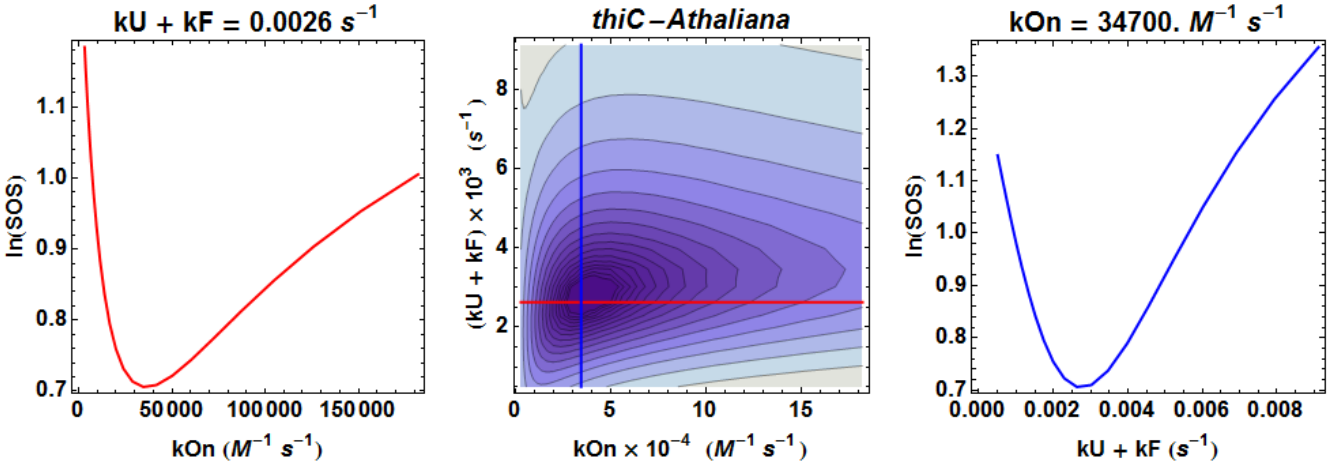


Fig. SD-10 Quality of fit of the $^{\bullet}$OH cleavage curves by the kinetic-model-based method

The central figure represents the logarithm of the residual sum of squares (SOS) as a function of the two essential parameters, k_{on} and $k_U + k_F$, used in the minimization (the overall K_d value obtained by ITC was imposed in the search for k_{on} and $k_U + k_F$). The left- and right-hand curves represent the variation of $\ln(SOS)$ along the horizontal red line ($k_U + k_F$ constant) and the vertical blue line (k_{on} constant), respectively. This shows that the minimum is very well defined although only two really free parameters were used for fitting at once all experimental kinetic curves. The fit resulting from this kinetic-model-based method, along with a bi-exponential fit derived from equation (4), are shown in Fig. 3B for a subset of residues.

SD-11 Primary and late TPP/RNA interaction for *ECthiC*$_{mutP1}$

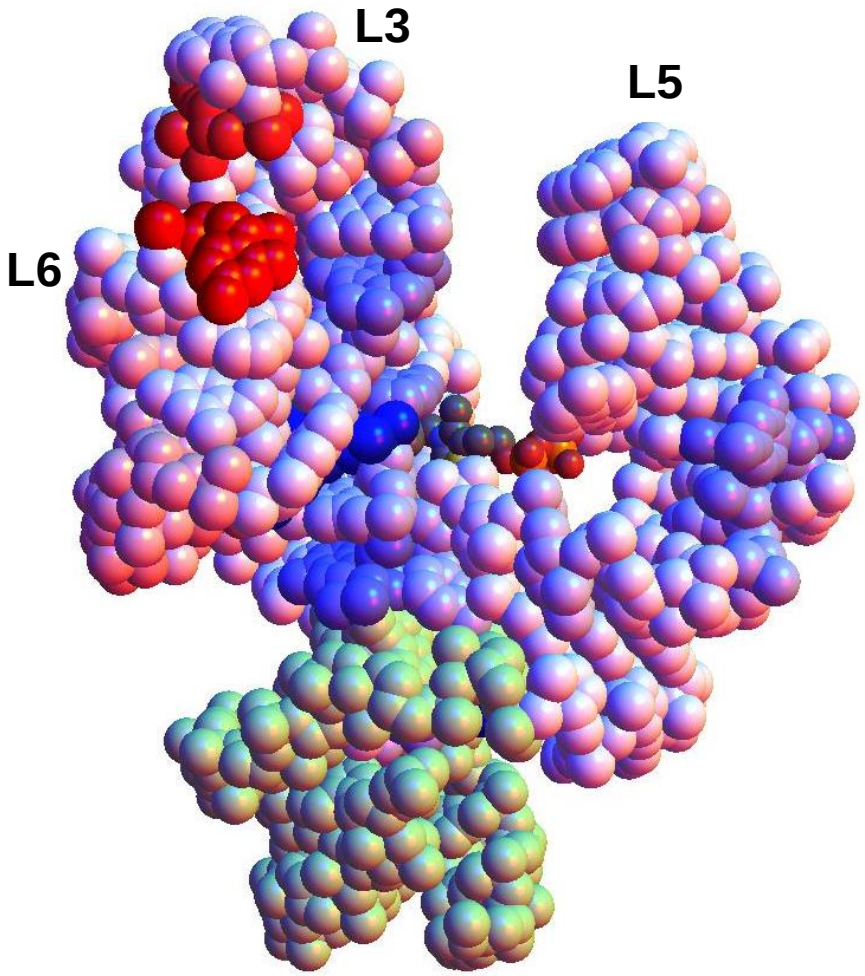


Fig. SD-11A Color coding of early TPP/RNA interaction for *ECthiC*$_{mutP1}$

The 3D structure of the model (see below) has been opened to visualize the interaction of the TPP-pyrimidine moiety with the RNA. The colors evolve from pure blue (minimum negative value) to pure red (maximum positive value) for the normalized difference $\left(\alpha_1^n - \alpha_0^n\right)/\sigma\left(\alpha_1^n - \alpha_0^n\right)$. The greenish colour corresponds to the 5'-end and 3'-end residues for which no accurate results were obtained. In this case, the most negative value is still located where the TPP-pyrimidine moiety interacts with the riboswitch (compare with Fig. 6B), but there are also some indications of pyrophosphate interaction site in the P4/P5 junction (Fig. 1D). The orientation in this figure is 180 ° away from the orientation in Fig. 6B, which was necessary for good visibility.

The 3D model of *ECthiC*$_{mupP1}$ was constructed using an interactive two-step process. In a first step, a structure-based sequence alignment was done between the *E. coli* orthologous sequence and the X-ray structure available for *A. thaliana* from reference [2] (PDB-id 2CKY). The alignment was built semi automatically with the last version of Assemble in reference [9] and then used to derive a core model for *ECthiC*. In a second step, single base substitutions, addition of the P6L6 domain (Fig. 2D), along with the geometric refinement of the final model, were all done in Assemble.

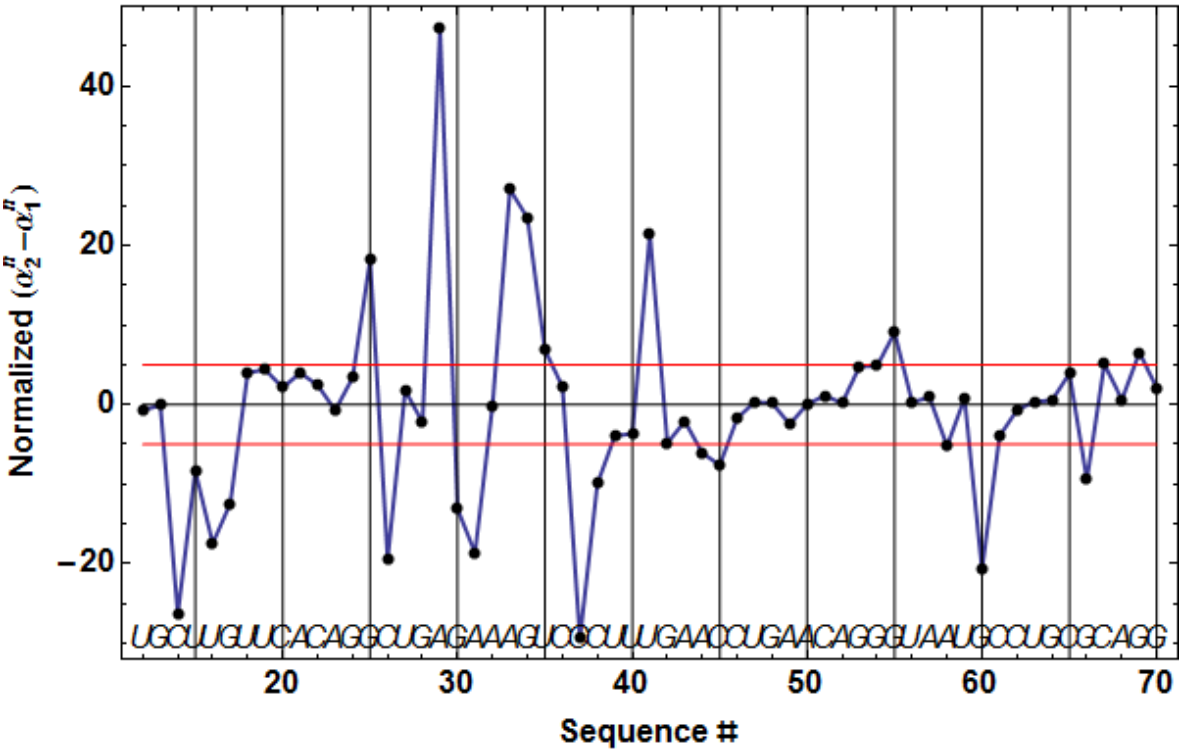


Fig. SD-11B TPP/RNA late interaction for *AThiC*

Evolution of the normalized difference $\left(\alpha_2^n - \alpha_1^n\right) / \sigma\left(\alpha_2^n - \alpha_1^n\right)$ between the coefficients α_1^n and α_2^n representative of the cleavage propensity of residue #*n*, respectively in the early formed complex R_1 and the fully folded complex R_2 (equations 1a,b). A large negative (resp. positive) value of the difference is the mark of late protection against cleavage (resp. late enhanced cleavage) after TPP addition. The red lines mark the positive and negative limits at a 5σ threshold.

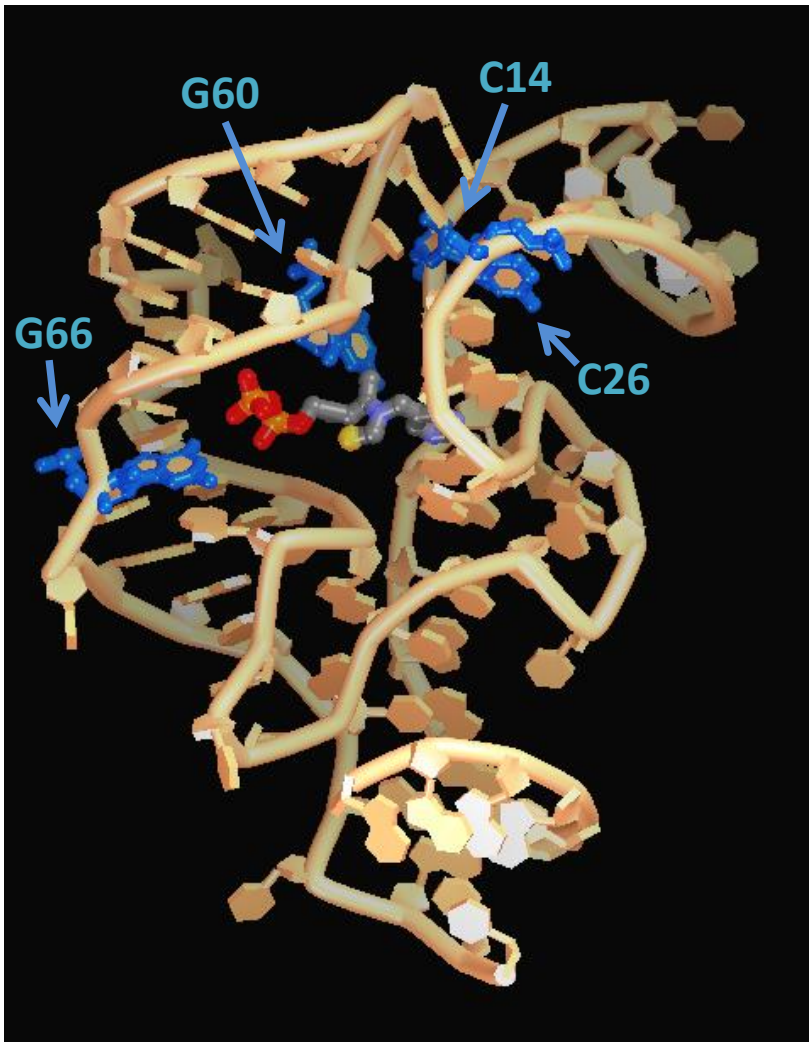


Fig. SD-11C Visualization of the TPP/RNA late interaction for *AThiC*

Selected residues with large negative difference values $\left(\alpha_2^n - \alpha_1^n\right)/\sigma\left(\alpha_2^n - \alpha_1^n\right)$ in Fig. SD-11B are located in the 3D structure. These residues are thus protected against cleavage during the final folding/closing step of the RNA. Residues C14 and C26 in P3 are protected by the closing of the apical loop L5, residue G60 is protected by the formation of a close contact with the TPP, and residue G66, which shows a less important protection, interacts through its base with a pyrophosphate-bound magnesium cation [2]. This contact, in itself, is not susceptible to protect the H atoms of the ribose against $^{\bullet}$OH-mediated cleavage and, therefore, the enhanced protection should result from a change of conformation induced by the interaction with the pyrophosphate. The figure was made with Accelrys Discovery Studio Visualizer.

SD-12 ON-state probability for a kinetically-regulated riboswitch

Our goal in this section is to relate the measured kinetic parameters to the ligand concentration of the ON/OFF switch. We recall that the expression of the probability for the riboswitch to be in the ON state is (equation 3 in the main text):

$$P_{\mathrm{ON}} = 1 - \tilde{R}_2 - f_2 \exp(-k_{fast}\, t_S) - s_2 \exp(-k_{slow}\, t_S) \qquad \text{(S12.1)}$$

and, as shown in SD-8, that $\tilde{R}_2$, f_2, s_2, k_{slow} and k_{fast} can all be expressed in terms of the overall ligand concentration L_{tot} and of the kinetic parameters k_{on}, k_{off}, k_F and k_U defined in equations (1a,b). Therefore, for a given riboswitch, P_{ON} is a function of L_{tot} and of the pausing time t_S, which is illustrated in Fig. SD-12A for three values of t_S.

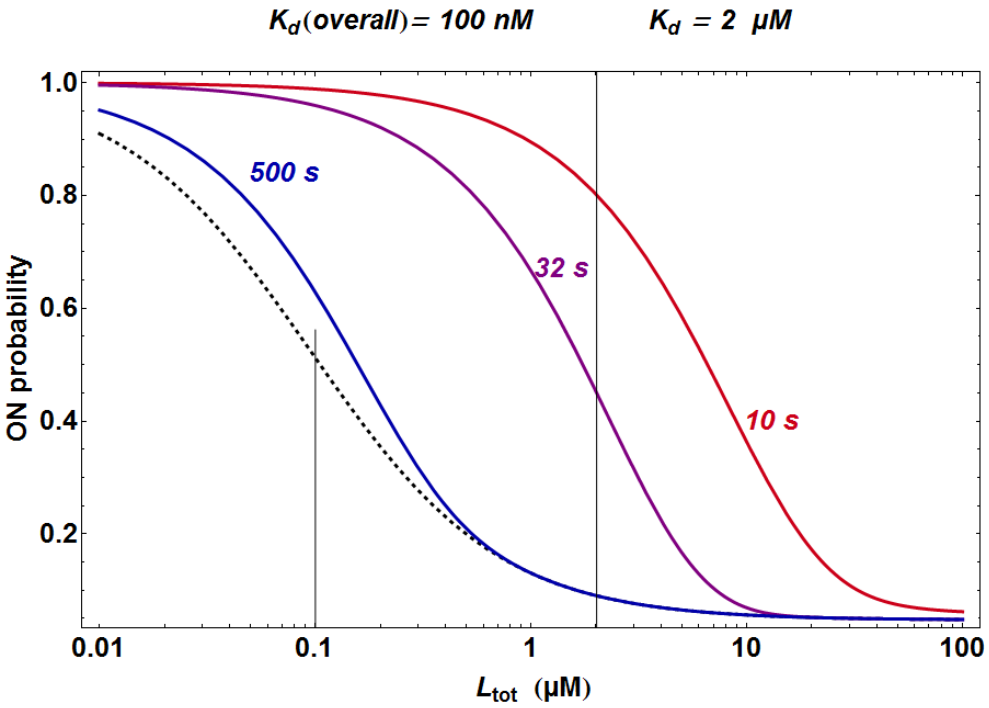


Fig. SD-12A Influence of the 'pausing time' t_S on $P_{\mathrm{ON}}(L_{\mathrm{tot}})$

The function $P_{\mathrm{ON}}(L_{\mathrm{tot}})$ was calculated with the indicated values of K_d(overall) and K_d and for three values of t_S (10, 32 and 500 s). When t_S becomes large, as for $t_s = 500\,s$, the two exponential terms in equation (S12.1) become negligible and the kinetic regulation curve becomes close to the 'thermodynamic' regulation curve (dashed curve) defined by $P_{\mathrm{ON}} = 1 - \tilde{R}_2$. The two curves would coincide exactly for $t_s \to \infty$. The full-height vertical bar marks the K_d of the initial step (2 μM). It also marks the inflection point of the curve corresponding to $t_s = 32\,s$ since k_{off} was chosen according to $k_{\mathrm{off}} = t_S^{-1} = 1/32\,s^{-1}$. The short vertical segment marks K_d(overall) = 100 nM, which corresponds to the inflection point of the 'thermodynamic' regulation curve. The other values for the calculation were $k_{\mathrm{on}} = 1.57\ 10^4\ \mathrm{M}^{-1}\,s^{-1}$, $k_{\mathrm{F}} = 0.45\ s^{-1}$ and $k_{\mathrm{U}} = 0.02\ s^{-1}$.

Figure SD-12A illustrates how the 'thermodynamic regulation' (dashed curve) would be efficient around the ligand concentration defined by K_d(overall) (vertical segment marking the inflection point of the dashed curve), whereas the 'kinetic regulation' operates at higher concentrations and depends strongly on t_S. The full-height vertical line marks the value of K_d (2 µM in this example) and, therefore, at some value of t_S (32 s in this example), the inflection point of the resulting curve lies at $L_{tot} = K_d$ and the kinetic regulation is thus most efficient around $L_{tot} = K_d$. It is determined in the following what is the theoretical expression for this peculiar value of t_S.

Our goal is thus to obtain a quantitative estimate of the position of the inflection point marking the ON/OFF switch as a function of all measured kinetic parameters and of t_S. Equation (S12.1) does not allow obtaining such a result in any practical way. However, as explained in SD-8 (equation S8.5c *et seq.*) the term $f_2 \exp(-k_{fast}\, t_S)$ becomes rapidly negligible in comparison of $s_2 \exp(-k_{slow}\, t_S)$. One may then consider the approximation $P_{ON} \approx -s_2 \exp(-k_{slow}\, t_S)$, which is only valid for sufficiently large concentrations of the ligand (*i.e.* $L_{tot} >> K_d$(overall)) and derives from neglecting $-f_2 \exp(-k_{fast}\, t_S)$ and the purely thermodynamic term $1-\tilde{R}_2$, which either becomes constant or negligible (dashed curve in Fig. SD-12A) and, therefore, does not influence significantly the second derivative of $P_{ON}(L_{tot})$ allowing to locate the inflexion point. Examination of the orders of magnitude of the different terms defined in SD-8 led to $4r/p^2 < 1$ in all situations, and even $4r/p^2 \ll 1$ for low ligand concentrations. It is useless (and would be cumbersome) to reproduce the successive estimations obtained by standard approximation methods deriving from the low value of $4r/p^2$; these allowed us to derive $-s_2 \approx \tilde{R}_2$. Numerical calculations confirmed fully this approximation and showed that $-s_2 \approx 1$ and varies almost linearly around 1 when $L_{tot} \approx K_d$ (blue curves and short vertical bar in Fig. SD-12B).

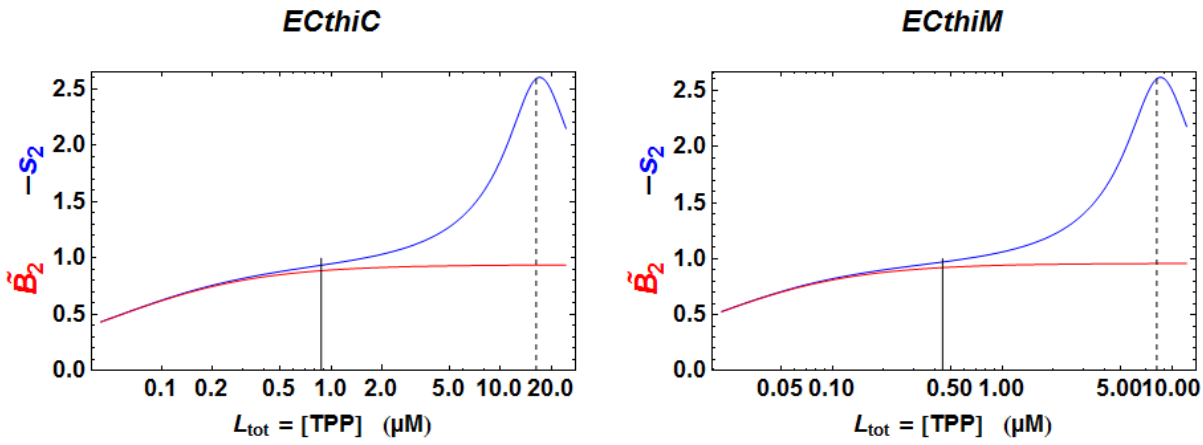


Fig. SD-12B Illustration of the approximations used to characterize the ON/OFF switch

The short vertical bar lies at $L_{tot} = K_d$. The quality of the approximation $-s_2 \approx \tilde{R}_2$ appears well for $L_{tot} \leq K_d$ as the blue and red curves become indistinguishable. The dashed vertical bar lies at $L_{tot} = (k_F + k_U)/k_{on}$, which is the concentration separating the linear regime from the plateau (Fig. SD-9). Therefore, $L_{tot} = K_d$ lies deeply within the 'kinetic regulation' range.

We have thus obtained the following important result: To determine the position of the critical inflection point in the regulation curve, it is legitimate to ignore the coefficient $-s_2$ and to consider the greatly simplifying approximation $P_{ON} \approx \exp(-k_{slow}\, t_S)$ (Fig. SD-12C). To make use of it, one has to express k_{slow} as a function of L_{tot}. Since, as required for an efficient kinetic regulation, the switching point defined by the ligand concentration $L_{tot} \approx K_d$ lies always in the linear range of $k_{slow}(L_{tot})$ (see legend of Fig. SD-12B), we may use the linear approximation $k_{slow} \approx k_{off}/K_F + k_{on}\, L_{tot}$ (equation S9.1). It is then easily obtained that the inflection point lies closely at $L_{tot} \approx (k_{on}\, t_S)^{-1}$.

A very important conclusion derives from this: If the pausing time t_S is close to k_{off}^{-1}, the kinetic regulation occurs around the concentration $L_{tot} = k_{off}/k_{on} = K_d$, whereas a classical thermodynamic regulation would occur around $L_{tot} = K_d$ (overall). Numerical calculations fully confirmed this result (Fig. SD-12C) and the level of the approximations used to obtain it appears through the small empirical correction $t_S = 1.2\, k_{off}^{-1}$ as a better estimate to obtain a kinetic regulation at $L_{tot} \approx K_d$.

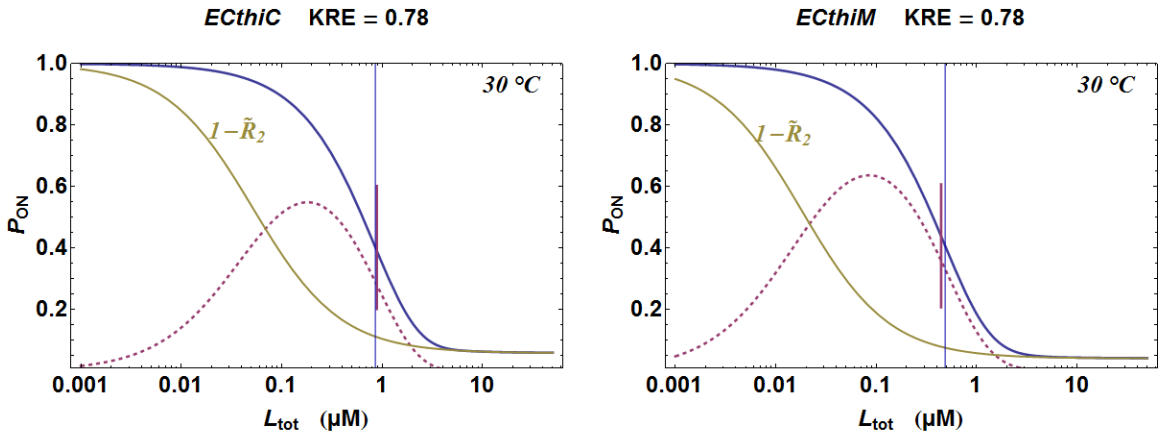


Fig. SD-12C Exact and approximate theoretical positions of the ON/OFF switch

The sigmoid curve of the exact dependence of P_{ON} as a function of the ligand concentration (equation S12.1 with $t_S = 1.2\,k_{off}^{-1}$) is shown as the solid blue curve. The curve marked $1-\tilde{R}_2$ corresponds to a hypothetical 'thermodynamic regulation' around K_d (overall). The bell-shaped dashed curve corresponds to the approximation $P_{ON} \approx \exp(-k_{slow}\, t_S)$, which is only valid at sufficiently large ligand concentrations and **only for determining the position of the inflection point**. The full-height vertical line marks the exact position of this inflection point on the blue curve and the short purple vertical bar marks its position from the approximation.

The indicated values of the 'kinetic regulation efficiency' (KRE) (see SD-14) are exact values derived from $\left|dP_{ON}/d\left(\log_{10} L_{tot}\right)\right|$ at the inflection point; the latter agree reasonably well with the theoretical approximation $\ln(10)\,e^{-(1+1/K_F)}$ equal to 0.796 for *ECthiC* and 0.811 for *ECthiM*.

SD-13 Explanation of an apparent contradiction with previous work

It is necessary to discuss our conclusion that a kinetic regulation results from $\boldsymbol{t_S \approx k_{off}^{-1}}$ in view of one in reference [10] which stated instead that it results from $\boldsymbol{t_S}$ being significantly smaller, or much less, than $\boldsymbol{k_{off}^{-1}}$. The two conclusions are only in apparent contradiction. Indeed, it was shown in [10] that the binding of the ligand (the adenine base or its fluorescent analogues) was well described by a single-step mechanism, hence the $\boldsymbol{k_{off}}$ under consideration in [10] cannot be identified to our $\boldsymbol{k_{off}}$ pertaining only to the first kinetic step of a composite mechanism (equations 1a,b). However, as mentioned in reference [1], an apparent 'overall' $\boldsymbol{k_{off}}$ for this composite mechanism comparable to $\boldsymbol{k_{off}}$ used in [10] can be estimated by consideration of the steady-state approximation as $\boldsymbol{k_{off}}\,\text{(overall)} \approx \boldsymbol{k_{off}}\ \boldsymbol{k_U} / \left(\boldsymbol{k_{off}} + \boldsymbol{k_F} \right)$. For both *EcthiC* and *ECthiM* one has $\boldsymbol{k_{off}} << \boldsymbol{k_F}$, which implies $\boldsymbol{k_{off}}\,\text{(overall)} \approx \boldsymbol{k_{off}} / \boldsymbol{K_F}$ with, respectively, $\boldsymbol{K_F} \approx 15$ (Fig. SD-7A) and $\boldsymbol{K_F} \approx 20$ (Fig. 4A) at 30 °C. As a consequence, using $\boldsymbol{t_S} \approx 1/\boldsymbol{k_{off}}\,\text{(overall)}$ would yield $\boldsymbol{t_S} \approx 750\,\text{s}$ (*ECthiC*) and $\boldsymbol{t_S} \approx 600\,\text{s}$ (*ECthiM*) at 30 °C. With such large $\boldsymbol{t_S}$ values, equation S12.1 would indeed produce 'kinetic' regulation curves very close to the thermodynamic regulation dashed curve in Fig. SD-12A. Our conclusion that a kinetic regulation results from $\boldsymbol{t_S \approx k_{off}^{-1}}$ is thus perfectly compatible with $\boldsymbol{t_S}$ having to be significantly smaller, or much less, than $1/\boldsymbol{k_{off}}\,\text{(overall)}$, as concluded in [10].

SD-14 Comparison of kinetic and thermodynamic regulation efficiencies

The efficiency of regulation can be defined as the slope $\left|\boldsymbol{dP_{ON}}/\boldsymbol{d}\left(\log_{10}\boldsymbol{L}_{\text{tot}}\right)\right|$ at the inflection point since the steeper this slope, the sharper the regulation. In the frame of this work, one may define a theoretical 'Kinetic Regulation Efficiency' (KRE) by using $\boldsymbol{P}_{\text{ON}} \approx \exp(-\boldsymbol{k_{slow}}\,\boldsymbol{t_S})$ with $\boldsymbol{t_S} \approx \boldsymbol{k_{off}^{-1}}$ (as explained in SD-12) and $\boldsymbol{k_{slow}} \approx \boldsymbol{k_{off}}/\boldsymbol{K_F} + \boldsymbol{k_{on}}\,\boldsymbol{L}_{\text{tot}}$ (equation S9.1). With $\log_{10}\boldsymbol{L}_{\text{tot}}$ as the variable, it can be obtained for any riboswitch ensuring a kinetic regulation according to model #1:

$$\text{KRE} = \left|\boldsymbol{dP_{ON}}/\boldsymbol{d}\left(\log_{10}\boldsymbol{L}_{\text{tot}}\right)\right| = \ln(10)\,\boldsymbol{e}^{-(1+1/\boldsymbol{K_F})} \qquad \text{(S14.1)}$$

which simplifies to $\text{KRE} \approx \text{KRE}_{\text{max}} = \ln(10)\,\boldsymbol{e}^{-1} = 0.847$ when $\boldsymbol{K_F} = \boldsymbol{k_F}/\boldsymbol{k_U} \gg 1$. One thus obtains the interesting result that the KRE is close to 0.847 (Fig. SD-12C). As a rule of thumb, this means that a four-fold increase of the ligand concentration around the switching point *(i.e.* from $\boldsymbol{L}_{\text{tot}} \approx \boldsymbol{K_d}/2$ to $\boldsymbol{L}_{\text{tot}} \approx 2\boldsymbol{K_d}$) always makes $\boldsymbol{P}_{\text{ON}}$ drop by *ca.* 1/2 (from ~0.65 to ~0.15).

In a situation of classical 'thermodynamic regulation' with no distinction between $\boldsymbol{K_d}$ and $\boldsymbol{K_d}\,(\text{overall})$, and always with the same assumption of an excess of ligand, the fraction of ligand-bound RNA is given by $\boldsymbol{K_d}\left(\boldsymbol{K_d} + \boldsymbol{L}_{\text{tot}}\right)^{-1}$, which implies $\boldsymbol{P}_{\text{ON}} = 1 - \boldsymbol{K_d}\left(\boldsymbol{K_d} + \boldsymbol{L}_{\text{tot}}\right)^{-1}$. The inflection point (with $\log_{10}\boldsymbol{L}_{\text{tot}}$ as the variable) lies at $\boldsymbol{L}_{\text{tot}} = \boldsymbol{K_d}$, and calculation of the slope at this concentration shows that the 'Thermodynamic Regulation Efficiency' (TRE) has a universal value:

$$\text{TRE} = \left|\boldsymbol{dP_{ON}}/\boldsymbol{d}\left(\log_{10}\boldsymbol{L}_{\text{tot}}\right)\right| = \ln(10)/4 \approx 0.576 \qquad \text{(S14.2)}$$

The definition of KRE and TRE with $\left|\boldsymbol{dP_{ON}}/\boldsymbol{d}\left(\log_{10}\boldsymbol{L}_{\text{tot}}\right)\right|$ is justified for the ease of thinking to orders of magnitude in terms of decades. Defining KRE and TRE with $\left|\boldsymbol{dP_{ON}}/\boldsymbol{d}\left(\ln\boldsymbol{L}_{\text{tot}}\right)\right|$ would lead to the simpler results $\text{KRE}_{\text{max}} = 1/\boldsymbol{e}$ and $\text{TRE} = 1/4$ without any change in the previous conclusions.

SD-15 *In vitro* transcription

E. coli thiC:

5’-taatttaaaatttatcaaaaagagtaTTGACTtaaagtctaacctatagGATACTtacagccAtggagagggagacccc

taatt*tcttgtcggagtgccttaactggctgagaccgtttattcgggatccgcggaacctgatcaggctaatacctgcgaagggaaca*

*aga*gttaatctgctatcgcatcgcccctgcggcgatcgtctcttgcttcatccgtcgtctgacaagccacgtccttaacttttt**ggaa**tgag

ct ATG tct gca aca aaa ctg acc cgc cgc gaa caa cgc gcc cgg gcc caa cat ttt atc gac acc

E. coli thiM:

5’-taatttaaaatttatcaaaaagagtaTTGACTtaaagtctaacctatagGATACTtacagccAtggagagggagaccccg

caaccaaac*gactcggggtgcccttctgcgtgaaggctgagaaatacccgtatcacctgatctggataatgccagcgtagggaagt*

*c*acggaccaccaggtcattgcttcttcacgttatggc**aggag**caaact ATG caa gtc gac ctg ctg ggt tca gcg caa tct

gcg cac gcg tta cac ctt ttt cac caa cat tcc

SD-15A Template sequences for *in vitro* transcription

The full sequences of *E.coli thiC* and *thiM* riboswitches and the twenty first codons of the downstream protein sequence were cloned under the T7A1promoter. -35 and -10 boxes of the promoter are in capital. The sequences initially transcribed for labelling with GTP^{32} are underlined, the capital A being the first ribonucleotide incorporated. The aptamer sequences are in italic, the initiator codons are highlighted in red and the Shine-Dalgarno (SD) sequences are in bold.

SD-15B Kinetics of *in-vitro* transcription of *E. coli thiC* and *thiM* riboswitches

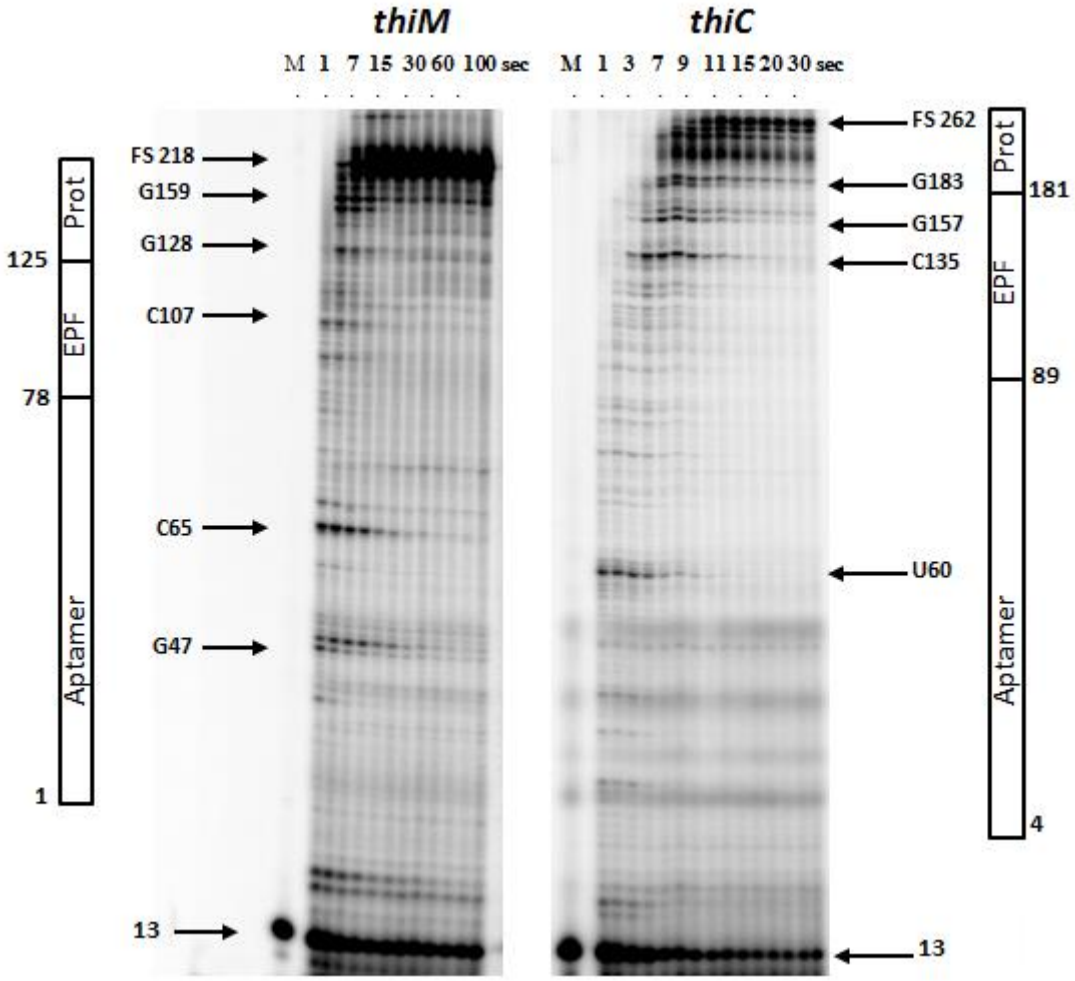


Fig. SD-15B Gel analysis of the kinetics of *in vitro* transcription of *E. Coli thiM* and *thiC* riboswitches (sequences in SD-15A) by the *E.coli* RNA polymerase at 37 °C. Lanes M show the labelling step transcription products before the addition of the four NTPs. Pause sites are indicated by arrows. The aptamer and expression platform domains of each riboswitch and the N-terminal part of the thiM and thiC proteins are indicated on each side.

SD-16 RNA polymerase pausing during *ECthiM* template transcription and translation regulation mechanism

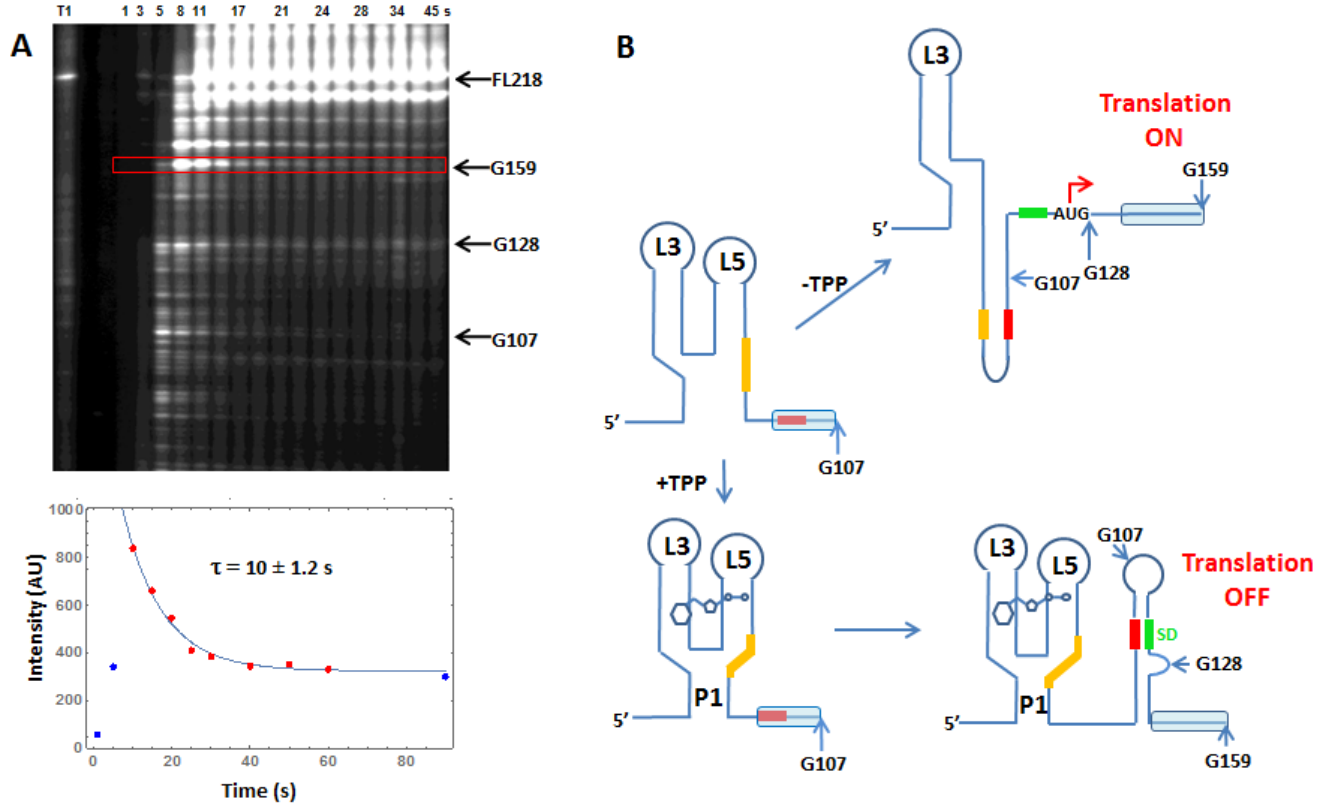


Fig. SD-16 RNA polymerase pause sites observed during *ECthiM* template transcription

A: The transcribed sequence is in SD-15A. Top: pause sites positions are indicated on the right. FL: full length transcript. The left lane is an RNAse-T1 ladder. Bottom: quantitative analysis of the duration of pausing at G159. Transcription experiments were performed at 37 °C in presence of NusA (0.3 µM).

B: Regulation model for the *ECthiM* riboswitch. At low TPP concentration, the switching sequence (orange) pairs with the anti-Shine-Dalgarno sequence (red). This pairing might be facilitated by the pausing of the polymerase at G128 that hides the Shine-Dalgarno (SD) sequence (green). When polymerization resumes, the SD sequence is free to bind the 30S ribosomal subunit to initiate translation. At high TPP concentration, the switching sequence contributes to the aptamer structure. As polymerization resumes, the anti-SD sequence pairs with SD, and the polymerase pausing at position G159 contributes to stabilize the hairpin that locks the SD sequence so that translation is OFF. The RNA polymerase is represented as a blue rectangle. Adapted from [11; 12].

SD-17 ITC results for *AThiC*

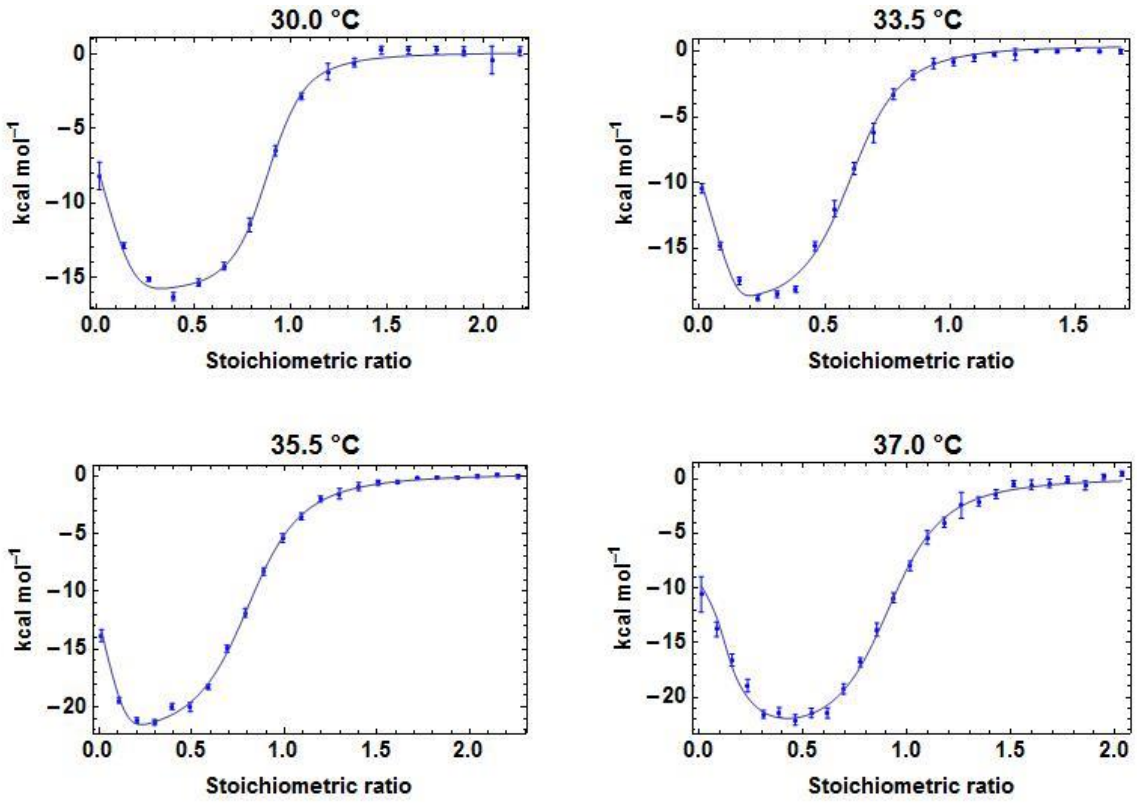


Fig. SD-17A ITC processing for *AThiC*

The results show unambiguously atypical titration curves corresponding to two distinct RNA species, which were not detected, neither by •OH radical footprinting nor by SPR. The fitting of each data set (continuous curves) yielded K_d and ΔH values for each RNA species. Importantly, it is impossible to fit these data by assuming simply that the two RNA species are engaged into equilibrium prior their binding to TPP. The results of the fitting are shown in **Fig. SD-17BC.**

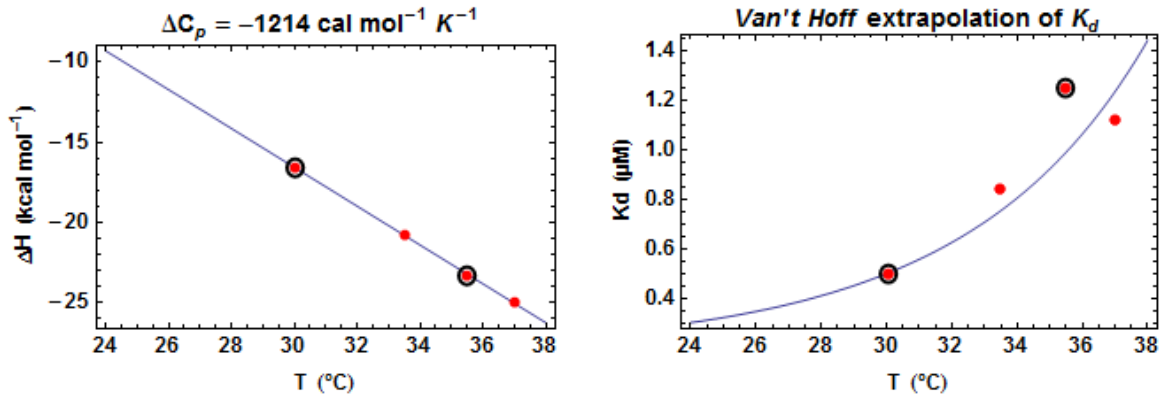


Fig. SD-17B Temperature dependence of ΔH and K_d for the major species of *ATthiC*

The results for the four temperatures are shown. The ΔH values yield a very well-defined ΔC_p and the continuous curve for the K_d is the theoretical extrapolation of the K_d at 30 °C from van't Hoff equation. Note that these these K_d values obtained by ITC correspond to K_d(overall). The agreement with the experimental points is reasonable. (The black circles mark experiments made the same day with the same batch of RNA, which is highlighted for comparison with Fig. SD-17C).

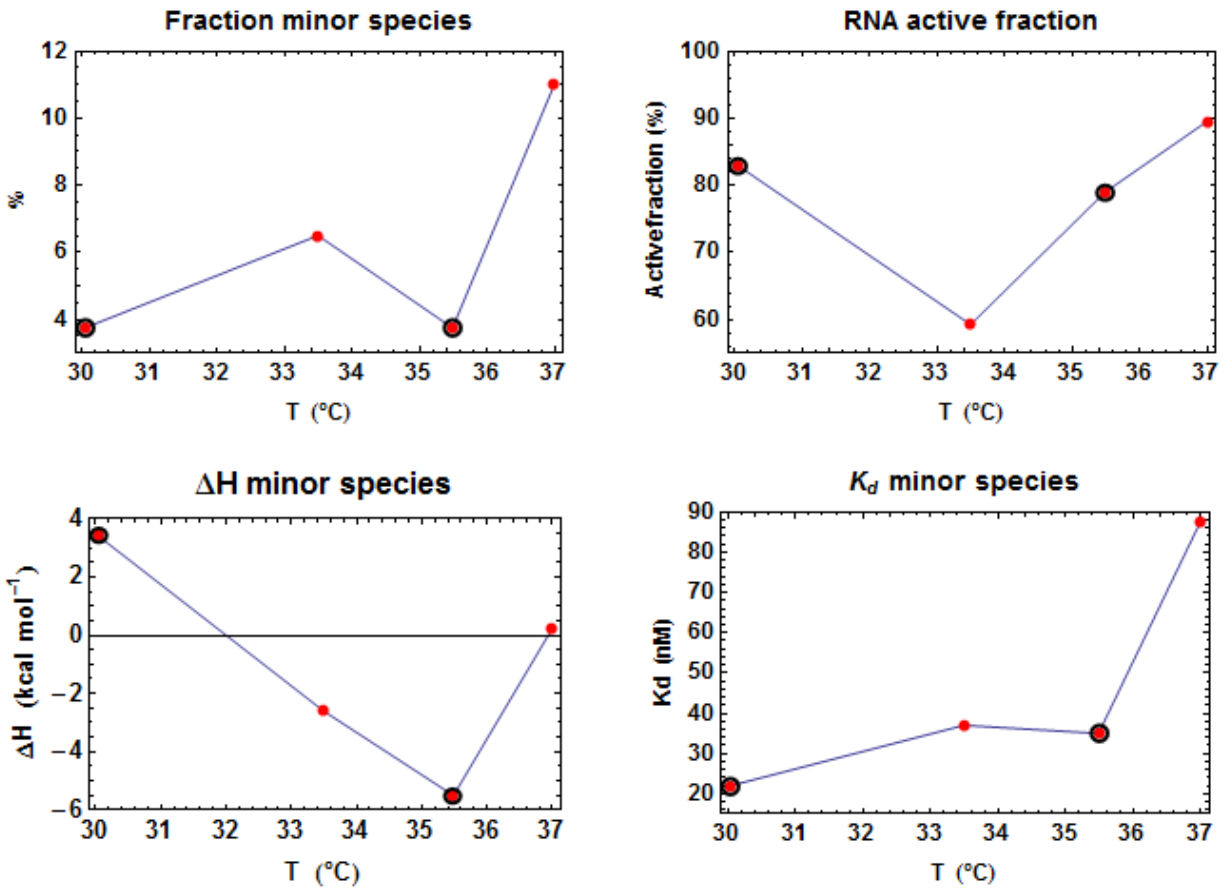


Fig. SD-17C ΔH, K_d and relative fraction for the minor species of *AThiC*

The results show an erratic distribution of the results for the minor species in comparison of the results for the major species. The black circles mark experiments made the same day with the same batch of RNA, which highlights that the fraction and the K_d of this minor species, as well as the overall active fraction of the RNA (including both the major and minor species), are most likely influenced by slight differences during the preparation of the RNA. We are not able to propose a clear interpretation, neither for the higher affinity of TPP for this minor species, nor for the fact that the fraction and the K_d of this minor species seem correlated in such a way that the ratio fraction/ K_d is almost constant.

References

1. Burnouf, D., Ennifar, E., Guedich, S., Puffer, B., Hoffmann, G., Bec, G., Disdier, F., Baltzinger, M. & Dumas, P. (2012). kinITC: a new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *J. Am. Chem. Soc.* **134**, 559-65.
2. Thore, S., Leibundgut, M. & Ban, N. (2006). Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science* **312**, 1208-11.
3. Balasubramanian, B., Pogozelski, W. K. & Tullius, T. D. (1998). DNA strand breaking by the hydroxyl radical is governed by the accessible surface areas of the hydrogen atoms of the DNA backbone. *Proc. Natl. Acad. Sci. U S A* **95**, 9738-43.
4. Kulshina, N., Edwards, T. E. & Ferre-D'Amare, A. R. (2009). Thermodynamic analysis of ligand binding and ligand binding-induced tertiary structure formation by the thiamine pyrophosphate riboswitch. *RNA* **16**, 186-96.
5. Fersht, A. (1977). *Enzyme structure and mechanism*, W.H. Freeman & Co., Reading & San Franciso.
6. Gilbert, S. D., Stoddard, C. D., Wise, S. J. & Batey, R. T. (2006). Thermodynamic and Kinetic Characterization of Ligand Binding to the Purine Riboswitch Aptamer Domain. *Journal of Molecular Biology* **359**, 754-768.
7. Lang, K., Rieder, R. & Micura, R. (2007). Ligand-induced folding of the thiM TPP riboswitch investigated by a structure-based fluorescence spectroscopic approach. *Nucleic Acids Res.* **35**, 5370-5378.
8. Martin, J. S., Simmons, K. & Laederach, A. (2009). Exhaustive Enumeration of Kinetic Model Topologies for the Analysis of Time-Resolved RNA Folding. *Algorithms* **2**, 200-214.
9. Jossinet, F., Ludwig, T. E. & Westhof, E. (2010). Assemble: an interactive graphical tool to analyze and build RNA architectures at the 2D and 3D levels. *Bioinformatics* **26**, 2057-9.
10. Wickiser, J. K., Cheah, M. T., Breaker, R. R. & Crothers, D. M. (2005). The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry* **44**, 13404-14.
11. Winkler, W., Nahvi, A. & Breaker, R. R. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952-6.
12. Rentmeister, A., Mayer, G., Kuhn, N. & Famulok, M. (2007). Conformational changes in the expression domain of the Escherichia coli thiM riboswitch. *Nucleic Acids Research* **35**, 3713-3722.