File S1 supplementary methods, figures, and tables

MeltR Software Provides Facile Determination of Nucleic Acid Thermodynamics

Jacob P. Sieg[1,2,*], Sebastian J. Arteaga[3], Brent M. Znosko[3], Philip C. Bevilacqua[1,2,4,*]

[1]Department of Chemistry, Pennsylvania State University, University Park, PA 16802.

[2]Center for RNA Molecular Biology, Pennsylvania State University, University Park, PA 16802.

[3]Department of Chemistry, Saint Louis University, Saint Louis, MO 63103.

[4]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802.

*Correspondence should be directed to jus841@psu.edu and pcb5@psu.edu

Keywords: Absorbance-detected melting curves, thermodynamics, RNA

ORCID:

Jacob P. Sieg: 0000-0001-5414-1667

Sebastian J. Arteaga: 0000-0001-7641-7919

Brent M. Znosko: 0000-0002-6823-6218

Philip C. Bevilacqua: 0000-0001-8074-3434

Contents

Supplementary information methods

Data processing by meltR.A

Raw data preprocessing. MeltR performs the following data preprocessing steps before fitting (Figure 1B):

1.) RNA and DNA extinction coefficients are calculated from sequences using the extinction coefficient data[1] and the method recommended by Tinoco and colleagues. Extinction coefficients are most accurate for absorbances in the upper baseline where the RNA is unfolded.[2] Extinction coefficients for RNA can be calculated at wavelengths that are not 260 nm.[1] Alternatively, the user can provide custom extinction coefficients or concentrations for each non-blank sample in the data set (Figure 1B).

2.) If blanks are provided in the dataset, the background absorbance of a user-specified blank is scaled to the pathlength of the cuvette and subtracted from each curve (Figure 1B).

3.) Users can supply concentrations for non-absorbance data. If concentrations were not supplied, the total strand concentration (C_t) is calculated at a user-specified temperature where the RNA is unfolded (default = 90 °C) using the extinction coefficient(s) and Beer's law. For homoduplexes and monomolecular-self structured molecules, where there is only one nucleic acid in the sample, the C_t is the total concentration of the single strand. For heteroduplexes, the C_t is the total concentration of strand 1 plus the total concentration of strand 2, which uses the average of the two extinction coefficients (Figure 1B, Row 4, Step 2).

4.) High and low temperature data are trimmed (to the user's specification) (Figure 1B).

5.) First and second derivatives are taken using polynomial regression. First, the data are approximated using a 20th order polynomial. Then, the first and second derivatives of the polynomial are determined analytically using calculus. There are other ways to approximate first and second derivatives of melting curves, but we found that polynomial regression was more robust and precise. Also, melting curves can be approximated using lower order polynomials, but the 20th order polynomial improves the approximation of end behavior.

6.) Initial parameter estimates, which are required for non-linear regression, are calculated for each curve.

The initial values for slopes and intercepts of the baselines are estimated by fitting absorbance values that are ≥75th quantile for the upper baseline, and fitting absorbance values that are ≤25th quantile for the lower baseline to y = mx + b (Figure 1B.)

Initial values for the enthalpy ($\Delta H°$) are determined using the $T_{0.5}$ and $T_{0.75}$ (in Kelvin) from first and second derivative curves using equation 1, where C = -0.007, -0.0044, and -0.0032 kcal/mol*K for heteroduplex, homoduplex, and monomolecular self-structured models, respectively (Figure 1B).[3]

$$\Delta H° = \frac{C}{\left(\frac{1}{T_{0.5}} - \frac{1}{T_{0.75}}\right)} \quad (1)$$

Note that the value of $T_{0.5}$ is approximated by finding the maximum of the first derivative curve, and $T_{0.75}$ is approximated by finding the minimum of the second derivative, to a precision of ≤0.1 °C, where $T_{0.5}$ is the approximate melting temperature, T_m, where 50% of the nucleic acid is single stranded, and $T_{0.75}$ is the approximate temperature where 75% of the nucleic acid is single stranded.

Thermodynamic models for duplex formation. Thermodynamic parameters for helix formation are obtained using equation 2, which relates the equilibrium constant to temperature:

$$\ln(K) = \frac{\Delta S°}{\mathrm{R}} - \frac{\Delta H°}{\mathrm{R}T} \quad (2)$$

Δ*S*° is the entropy of helix formation, Δ*H*° is the enthalpy of helix formation, R is the gas constant in kcal/mol, *T* is the temperature in Kelvin, and *K* is the equilibrium constant given by equation 3, 4, and 5 for heteroduplexes, homoduplexes, and monomolecular self-structured RNA, respectively.

$$K = \frac{[AB]}{[A][B]} \quad (3)$$

$$K = \frac{[A_2]}{[A]^2} \quad (4)$$

$$K = \frac{[F]}{[U]} \quad (5)$$

For equations 3 and 4, *[A]* and *[B]* are the concentration of different strands, *[AB]* is the concentration of strand A in a duplex with strand B, and *[A_2]* is the concentration of a self-complementary strand A in a duplex with another strand of A. For equation 5, *[F]* is the concentration of a monomolecular self-structured RNA in the folded state, and *[U]* is the concentration of a monomolecular self-structured RNA in the unfolded state.

MeltR uses three methods to calculate thermodynamic parameters: (1) fitting melting curves individually, (2) fitting the T_m as a function of temperature, and (3) global fitting melting curves.

Method 1: Fitting melting curves individually. Method 1 fits absorbance as a function of temperature for each melt individually. Baselines are modeled as *y* = m*x* + b for the absorbance of the unfolded single-stranded and folded duplex/hairpin states (equation 6).

$$A = \mathrm{m}T + \mathrm{b} \quad (6)$$

The absorption of each sample as a function of temperature is fit to equation 7 using the *nls* function[4] in base R, which performs non-linear least-squares fitting.[5] DS and SS represent the nucleic acid in the double-stranded (or folded hairpin) and single-stranded (or unfolded hairpin) states, respectively.

$$\mathrm{A}(T) = (\mathrm{m}_{\mathrm{DS}}T + \mathrm{b}_{\mathrm{DS}})f_{\mathrm{DS}}(T) + (\mathrm{m}_{\mathrm{SS}}T + \mathrm{b}_{\mathrm{SS}})(1 - f_{\mathrm{DS}}(T)) \quad (7)$$

f_{DS}(T) is given by the analytic solution of the binding constant. *MeltR* uses equation 8 for heteroduplexes, equation 9 for homoduplexes, and equation 10 for monomolecular self-structured RNA.

$$f_{\mathrm{DS}}(T) = \frac{\frac{2}{K(T)C_t} + 2 - \sqrt{\left(\frac{2}{K(T)C_t} + 2\right)^2 - 4}}{2} \quad (8)$$

$$f_{\mathrm{DS}}(T) = \frac{\frac{1}{2K(T)C_t} + 2 - \sqrt{\left(\frac{1}{2K(T)C_t} + 2\right)^2 - 4}}{2} \quad (9)$$

$$f_{\mathrm{DS}}(T) = \frac{K(T)}{K(T) + 1} \quad (10)$$

where C_t is the total strand concentration and *K*(*T*) is the equilibrium constant as a function of temperature, given by equation 11, where B is 4/C_t, 1/C_t, and 1 for heteroduplexes, homoduplexes, and monomolecular self-structured RNA, respectively.

$$K(\mathrm{T}) = \mathrm{e}^{\frac{\Delta H^\circ}{\mathrm{R}}\left(\frac{1}{T_m}-\frac{1}{T}\right)+\ln(\mathrm{B})} \tag{11}$$

Note, K(T) is in terms of $\Delta H°$, T_m, and C_t, instead of $\Delta S°$, to increase the ease of estimating initial parameters for non-linear regression and to increase the robustness of the *nls* algorithm. Briefly, $\Delta S°$ is replaced with T_m and C_t by solving equation 2 for $\Delta S°$ at the T_m. Also note that $\Delta H°$ will be most accurate where K is best known, which is near the T_m. It is thus best to have the T_m be near the temperature of interest for the parameters, often 37 °C.

$\Delta S°$ of helix formation is later calculated from $\Delta H°$ and T_m for each curve (Figure 1D.1).

$$\Delta S° = \frac{\Delta H°}{T_m} + \mathrm{R}\ln(\mathrm{B}) \tag{12}$$

The free energy at 37 °C ($\Delta G°_{37}$) is calculated from $\Delta H°$ and $\Delta S°$ for each curve (Figure 1D.1).

$$\Delta \mathrm{G}° = \Delta H° - T\,\Delta S° \tag{13}$$

The results for method 1 are reported by taking the mean value for each thermodynamic parameter and T_m across all samples (Figure 1D.2).

Method 2: Fitting the T_m as a function of C_t. Method 2 uses *nls* to fit the relationship between $1/T_m$ and C_t to equations 14 and 15 for heteroduplexes and homoduplexes, respectively. The T_m of monomolecular, self-structured RNA is independent of C_t, so method 2 cannot be used to provide thermodynamic parameters for this system, and *MeltR* automatically turns method 2 off. Nonetheless, it is recommended that the user conducts melts as a function of concentration to assure that the T_m is truly independent of concentration. The range of concentrations that should be tested can be estimated by entering the $\Delta H°_{\mathrm{duplex}}$ and $\Delta S°_{\mathrm{duplex}}$ for a melt into equation 15 and seeing how much the T_m would change if the system were homoduplex, which all hairpins can form at a high enough strand concentration; one should choose a range of concentrations for this experiment that would give a T_m outside the range of experimental error, typically at least 2-3 °C. Note that this range of concentrations is larger for long helices since ΔH° is in the denominator of equation 15.

$$\frac{1}{\mathrm{T_m}} = \frac{\mathrm{R}}{\Delta H°}\ln(\mathrm{C_t}) + \frac{\Delta S°}{\Delta H°} - \frac{\mathrm{R}}{\Delta H°}\ln(4) \tag{14}$$

$$\frac{1}{\mathrm{T_m}} = \frac{\mathrm{R}}{\Delta H°}\ln(C_t) + \frac{\Delta S°}{\Delta H°} \tag{15}$$

The T_m for each sample calculated from method 1 is used for method 2 by default. The propagated error in the T_m from the method 1 fits can be used as weights in the regression for method 2, propagated using equation 16. The weighted regression is turned off by default.

$$\mathrm{Weight}\ 1/T_m = \frac{{T_m}^2}{\sigma_{T_m}} \tag{16}$$

MeltR provides the user with two alternative approaches for determining the T_ms for method 2, and neither can use a weighted regression. The first uses the linear baseline parameters from method 1 to calculate the f(T) using equation 17. This is like the default and, therefore, also has a dependency on method 1.

$$f_{\mathrm{DS}}(T) = \frac{\mathrm{A} - (\mathrm{m_{SS}T} + \mathrm{b_{SS}})}{(\mathrm{m_{DS}T} + \mathrm{b_{DS}}) - (\mathrm{m_{SS}T} + \mathrm{b_{SS}})} \tag{17}$$

f(T) is approximately linear in the range of 0.4 to 0.6. Thus, MeltR fits f(T in {0.4 to 0.6}) to y = mT + b using the *lm* function in base R, which is solved at y = 0.5 to accurately determine the melting temperature for

each C_t. Alternatively, the T_ms for method 2 can be estimated using polynomial regression followed by the first derivative analysis (step 5 of data processing), which is independent of method 1.

$\Delta G°_{37}$ is calculated from $\Delta H°$ and $\Delta S°$ using equation 13. Error in $\Delta G°_{37}$ is calculated by propagating variances ($\sigma_{\Delta H°}$ and $\sigma_{\Delta S°}$) and covariance ($\sigma_{\Delta H°\Delta S°}$) in the fit terms $\Delta H°$ and $\Delta S°$ (equation 18).

$$\sigma_{\Delta G°} = \sqrt{(\sigma_{\Delta H°})^2 + (T * \sigma_{\Delta S°})^2 - 2 * T * \sigma_{\Delta H°\Delta S°}} \tag{18}$$

Method 3: Global fitting. This method is not provided by the original *MeltWin* program. Method 3 fits all melts simultaneously in a global fit using the *nls* algorithm in base R. In this global fit, equation 2 is plugged directly into equations 8-10, rather than used parametrically, so that the fit is in terms of $\Delta S°$ instead of T_m, allowing linking of individual curves. The baseline parameters are allowed to vary in each melt, but $\Delta H°$ and $\Delta S°$ are constrained to single values for all melts. For global fitting, *MeltR* uses the slopes and intercepts of the method 1 fits as initial parameter estimates for the slopes and intercepts of the global fit and uses the average of $\Delta H°$ and $\Delta S°$ from method 1 as initial parameter estimates. $\Delta G°_{37}$ and its error are calculated using equations 13 and 18, respectively.

T_m at 0.1 mM. The T_m at 0.1 mM is not a true thermodynamic parameter and is included in the output for historical reasons. It is an expectation maximized value, meaning it is estimated from a fit. The T_m at 0.1 mM is calculated using equation 19, obtained by rearranging equation 12.

$$\mathrm{T_m}^{0.1\ \mathrm{mM}} = \frac{\Delta H°}{\Delta S° - \mathrm{R} * \ln(\mathrm{B})} \tag{19}$$

The propagation of error to the T_m at 0.1 mM was calculated with equation 20.

$$\sigma_{\mathrm{T_m}^{0.1\ \mathrm{mM}}} = \sqrt{\left(\frac{\sigma_{\Delta H°}}{\Delta S° - \mathrm{R} * \ln(\mathrm{B})}\right)^2 + \left(\frac{\Delta H° \sigma_{\Delta S°}}{(\Delta S° - \mathrm{R} * \ln(\mathrm{B}))^2}\right)^2 - 2\frac{\Delta H° \sigma_{\Delta H \Delta S°}}{(\Delta S° - \mathrm{R} * \ln(\mathrm{B}))^3}} \tag{20}$$

Covariance is estimated using equation 21 for method 1, where the $\overline{\Delta \mathrm{H}°}$ and $\overline{\Delta S°}$ represent the mean of the individual fits. Covariance is extracted from the fits for methods 2 and 3.[6]

$$\sigma_{\Delta H°\Delta S°} = \frac{1}{N}\sum_{i=1}^{N} (\Delta H°_i - \overline{\Delta H°})\,(\Delta S°_i - \overline{\Delta S°}) \tag{21}$$

Data processing by the BLTrimmer

The *BLTrimmer* function explores the dependence of fit parameters on baseline trimming, starting from an existing *meltR.A* fit object. The *BLTrimmer* uses three steps: rapid testing of baseline trims, assessment, and a slower treatment of an ensemble of baselines that produces internally consistent thermodynamic parameters.

Baseline testing. For the first step, the *BLTrimmer* identifies a “no-trim range”, by default 15 to 85% folded using equations 8, 9, or 10, and a *meltR.A* fit object. Then, the *BLTrimmer* generates and fits many baseline trims by adding different amounts of data to the “no-trim range”. The user can generate baseline trims using a fixed or a floating baseline trim. The fixed method uses the same baseline length for each sample to test how baseline length effects the subsequent thermodynamic parameters. The default, floating, method allows different samples to have different baseline lengths. The floating method generates many possible trims of trims. The user can only test a small subset of these trims using a reasonable amount of computational time. Thus, a randomly selected subset of trims is tested using a fast, lightweight fit. The *BLTrimmer* fits each sample using method 1, using the global fit from the *meltR.A* object as initial guesses

for the non-linear regression step. Using these good initial guesses reduces the time the *nls* algorithm needs to find a solution. Next, the *BLTrimmer* fits each sample using method 2. Δ*H*° values from each method and baseline trim are then passed to the assessment step.

Baseline assessment. The user can choose one of three assessment methods to identify the baseline trims that produce the most internally consistent thermodynamic parameters.

Assessment method 1 tests how well the Δ*H*° values from each sample agree with each other. The normalized standard deviation ($\hat{\sigma}_{\Delta H^\circ_1}$) is calculated for each baseline trim, where $\sigma_{\Delta H^\circ_1}$ and $\mu_{\Delta H^\circ_1}$ are the standard deviation and mean Δ*H*° from the samples in method 1, respectively (equation 22).

$$\hat{\sigma}_{\Delta H^\circ_1} = \left| \frac{\sigma_{\Delta H^\circ_1}}{\mu_{\Delta H^\circ_1}} \right| \tag{22}$$

$\hat{\sigma}_{\Delta H^\circ_1}$ is then ranked into quantiles on a scale from 0 to 1. For example, if a baseline trim has the 100th smallest $\hat{\sigma}_{\Delta H^\circ_1}$ in a set of 1000 baseline trims, its quantile ($Q_{\Delta H^\circ_1}$) is 0.1. Then, trims with $Q_{\Delta H^\circ_1}$ at or below a user specified value are passed to the ensemble analysis (Figure S2A).

Assessment method 2 tests how well the Δ*H*° values from methods 1 and 2 agree using equation 23, where Δ*H*°$_2$ is produced by method 2.

$$\hat{\sigma}_{\Delta H^\circ_{1,2}} = 2 \left| \frac{\mu_{\Delta H^\circ_1} - \Delta H^\circ_2}{\mu_{\Delta H^\circ_1} + \Delta H^\circ_2} \right| \tag{23}$$

$\hat{\sigma}_{\Delta H^\circ_{1,2}}$ is then ranked into quantiles on a scale from 0 to 1 to produce $Q_{\Delta H^\circ_{1,2}}$. Then trims with $Q_{\Delta H^\circ_{1,2}}$ at or below a user specified value are passed to the ensemble analysis (Figure S2C).

Assessment method 3 combines assessment methods 1 and 2 using a statistic called the error distance ($\hat{\sigma}_D$, Figure S2E, equation 24).

$$\hat{\sigma}_D = \sqrt{Q_{\Delta H^\circ_1}{}^2 + Q_{\Delta H^\circ_{1,2}}{}^2} \tag{24}$$

$\hat{\sigma}_D$ is then ranked into quantiles on a scale from 0 to 1 to produce Q_D values. Trims with Q_D at or below a user specified value are passed to the ensemble analysis.

Ensemble analysis. Trims that pass the assessment criterion are treated as an ensemble of equally feasible trims. Each trim is passed to *meltR.A* and fit, which takes considerably longer than the lightweight analysis in the testing step because initial guesses are predicted for each trim using the protocol described for *meltR.A*. The average of resulting thermodynamic parameters is reported with 95% confidence intervals.

Data Modeling

For certain aspects of our study (e.g. to see if MeltR could reproduce thermodynamic parameters), data were modeled to produce simulated melting curves with realistic baselines and random scatter. First, a series of nine samples were generated with C_ts that evenly span C_t = 0.2/ε and 20/ε on a natural logarithm transformed scale, where ε is the molar extinction coefficient of the RNA, used as the extinction coefficient of the upper baseline. Samples were placed into virtual 1, 0.5, and 0.1 cm pathlength (ℓ) cuvettes so that the calculated absorbance was between 0.2 and 2. The extinction coefficient of the lower baseline was generated by multiplying ε by 0.6. Random baseline slopes (m) were seeded into the modeled data using a Gaussian distribution with the *rnorm* function in base R, with a mean of 0.00111 and a standard deviation of 0.00078, which were calculated from a sample of nine fits to real data. The intercept (b) was generated

for each baseline using equation 25, where ℓ is the pathlength of the cuvette and ε is the extinction coefficient.

$$b = C_t * \ell * \varepsilon - 90 * \mathrm{m} \tag{25}$$

Then, parameters were plugged into equation 7, using the Δ*H*° and Δ*S*° reported in the aforementioned papers[7,8] to generate melting curves that span 5 to 95 °C with 0.5 °C increments. Random absorbance scatter was seeded into the melting curves using the *rnorm* function with a mean of 0 and a standard deviation of 0.0005.

Comparisons between methods

Percent error, or difference, between parameters determined using different methods was calculated using a generalized formula for more than two observations in *X*, except for Figure S4E.

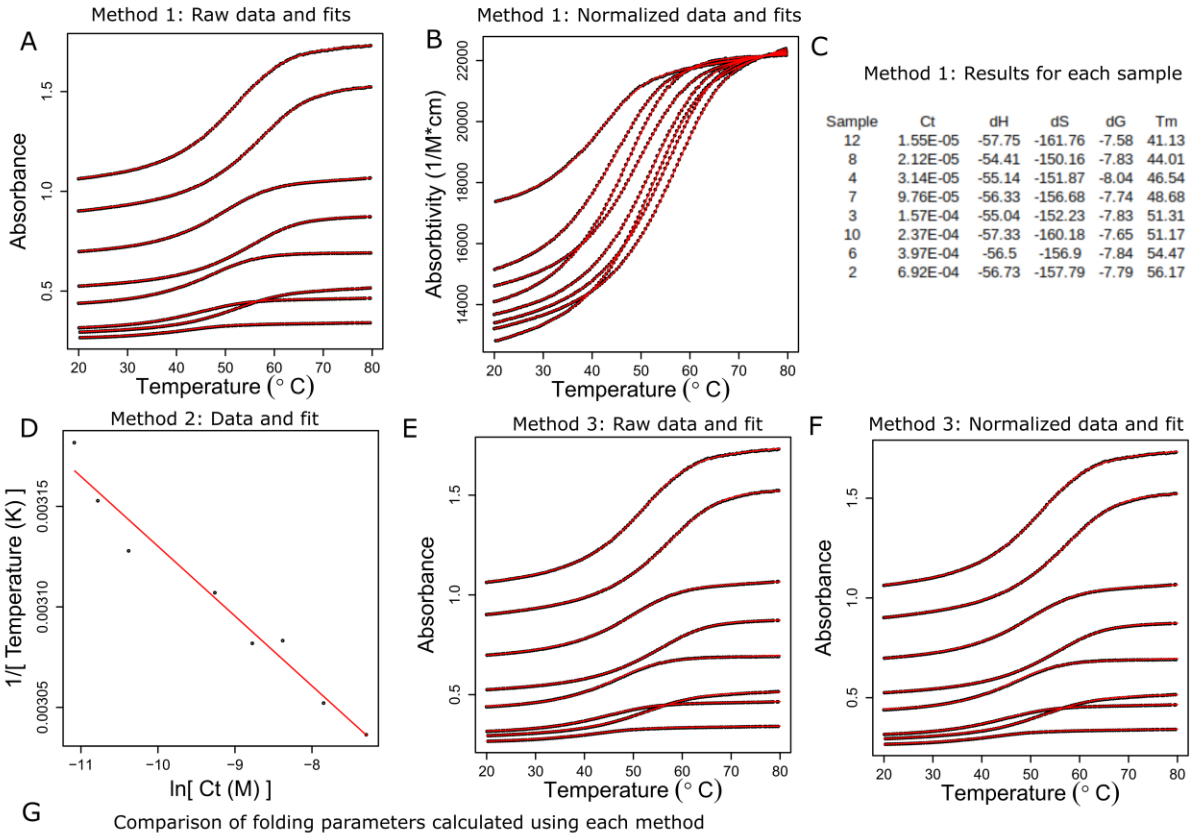
$$\text{For } X \in \{X_1, X_2, X_3, \ldots X_N\}$$

$$\%\text{error} = \left|\frac{\max(X) - \min(X)}{\mu_X}\right| \tag{26}$$

Thus, percent error is taken by subtracting the minimum value from the maximum value and dividing by the mean (equation 26). There are other formulations for calculating percent error between multiple observations, but we determined that equation 26 is the most conservative.

$$\text{Standard } \%\text{error} = \frac{X_{\text{Observed}} - X_{\text{Known}}}{X_{\text{Known}}} \tag{27}$$

Supplementary Figures and Tables



Sample	Ct	dH	dS	dG	Tm
12	1.55E-05	-57.75	-161.76	-7.58	41.13
8	2.12E-05	-54.41	-150.16	-7.83	44.01
4	3.14E-05	-55.14	-151.87	-8.04	46.54
7	9.76E-05	-56.33	-156.68	-7.74	48.68
3	1.57E-04	-55.04	-152.23	-7.83	51.31
10	2.37E-04	-57.33	-160.18	-7.65	51.17
6	3.97E-04	-56.5	-156.9	-7.84	54.47
2	6.92E-04	-56.73	-157.79	-7.79	56.17

Method	dH	SE.dH	dS	SE.dS	dG	SE.dG	Tm_at_0.1mM	SE.Tm_at_0.1mM
1 Individual fits	-56.15	1.18	-155.95	4.15	-7.79	0.14	49.11	3.74
2 Tm versus ln[Ct]	-57.35	4.46	-159.65	13.83	-7.83	0.18	49.1	0.37
3 Global fit	-56.33	0.2	-156.33	0.63	-7.84	0.01	49.4	0.03

Figure S1 The *meltR.A* function generates figures and tables for assessing results. Black circles represent data points, and red lines represent fits. Data were collected on self-complementary 5'CGCGCG3' RNA.[9] **(A)** Raw data fit using method 1. **(B)** Normalized data fit using method 1. Data were normalized to calculate the absorptivity at each temperature by dividing the absorbance by C_t (M) and the pathlength of the cuvette (cm). **(C)** Table summarizing the individual fits for method 1. Sample is the sample identifier, and data are presented in order of increasing C_t, the total strand concentration (M). dH is the $\Delta H°$ (kcal/mol). dS is the $\Delta S°$ (cal/mol•K). dG is the $\Delta G°_{37}$ (kcal/mol). T_m is the T_m of the duplex (°C). All parameters are for folding rather than unfolding. **(D)** Plot of $1/T_m$ versus the natural logarithm of the C_t, used for method 2. **(E)** Raw data fit using method 3, global fitting. **(F)** Normalized data fit using method 3. **(G)** Table summarizing results from the three methods. SE stands for the standard error, calculated from the regression models, represents the precision of the parameter given the model and data. Note, values are rounded to the hundredths place by MeltR. Trailing zeros are often removed by R or Excel.

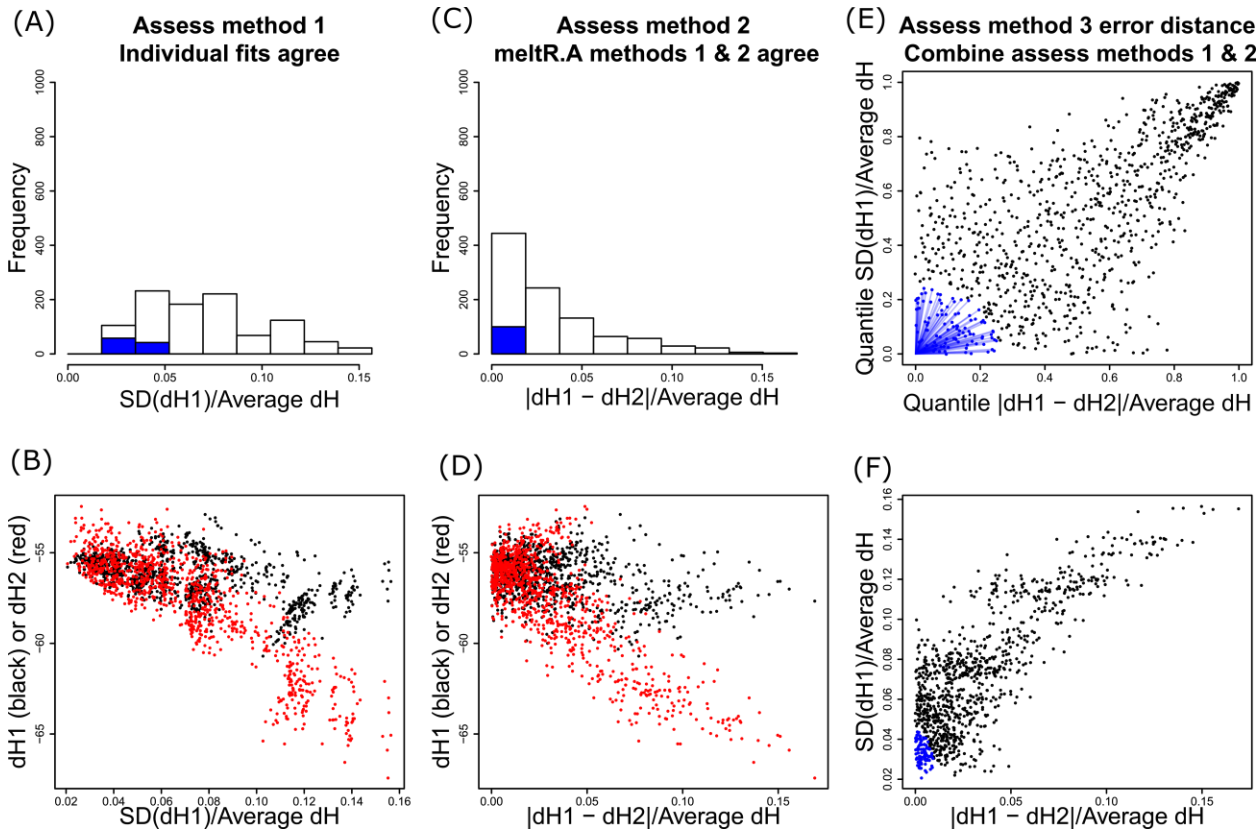


Figure S2 The *BLTrimmer* generates a *.pdf file for assessing results. **(A)** Assess method 1, individual fits agree with each other. Histogram of the standard deviation of Δ*H*° values determined by fitting 1000 baseline trims with method 1. Grey bars represent the results of all 1000 baseline trims. Blue bars represent the optimum 100 baseline trims with Q_E values ≤10% quantile. **(B)** Distribution of Δ*H*° values from fitting 1000 trims with methods 1 and 2 versus the standard deviation in the method 1 fits. Each point comes from a single trim. A black point represents the Δ*H*° from fitting a baseline trim with method 1 and is associated with a red point from fitting the same baseline trim with method 2. **(C)** Assess method 2, methods 1 and 2 agree with each other. Histogram of the Δ*H*° agreement between methods 1 and 2 determined by fitting 1000 baseline trims. Grey bars represent the results of all 1000 baseline trims. Blue bars represent the optimum 100 baseline trims with Q_E values ≤10% quantile. **(D)** Distribution of Δ*H*° values from fitting 1000 baseline trims with methods 1 and 2 versus Δ*H*° agreement between methods 1 and 2. Each point comes from a baseline trim. A black point represents the Δ*H*° from fitting a baseline trim with method 1 and is associated with a red point from fitting the same baseline trim with method 2. **(E)** Assess method 3, error distance, combines agreement between individual fits in method 1 with agreement between methods 1 and 2. Ranked agreement of Δ*H*° values for method 1 versus the ranked agreement between Δ*H*° values from methods 1 and 2. Error distance (Q_E) is the Pythagorean distance between a point and the origin of the plot. Each point represents a fit to a single baseline trim. Blue points represent the optimum 100 baseline trims with Q_E values ≤10% quantile. Blue lines represent the Q_E. **(F)** The standard deviation of ΔH° values produced by method 1 versus Δ*H*° agreement between methods 1 and 2. Each point represents a fit to a single baseline trim. Blue points represent the optimum 100 baseline trims with Q_E values ≤10% quantile.

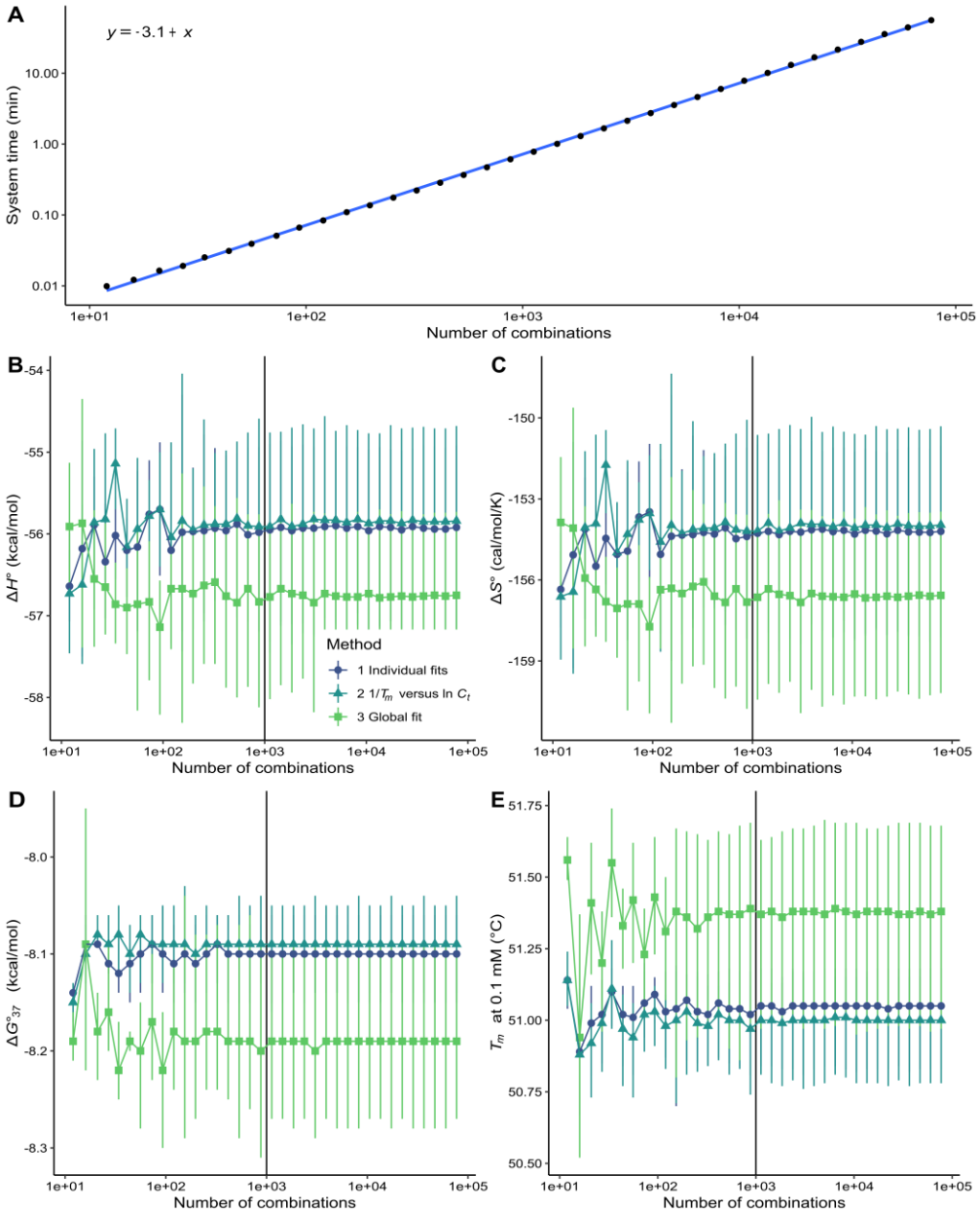


Figure S3 1000 randomly generated baseline trims provide accurate thermodynamic parameters and is computationally achievable. **(A)** System time for the *BLTrimmer* as a function of the number of baseline trims (combinations) evaluated in each run. **(B-E)** Convergence of thermodynamic parameters Δ*H*°, Δ*S*°, Δ*G*°$_{37}$, and T_m as a function of the number of baseline trims evaluated in a *BLTrimmer* analysis. Points represent the mean value, and error bars represent 95% confidence intervals. The black vertical lines represent the default setting of 1000 trims.

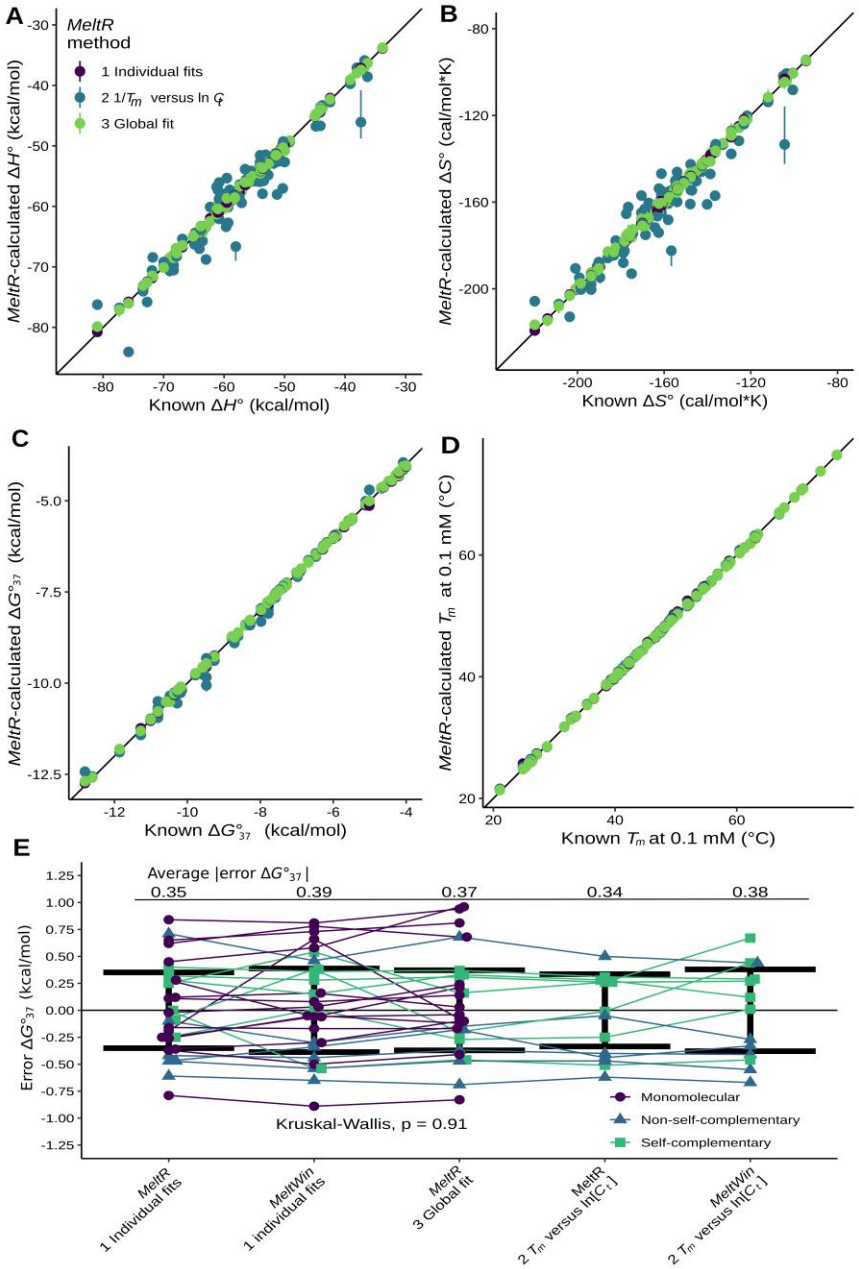


Figure S4 *MeltR* accurately fits modeled data with no user intervention. **(A)** Δ*H*° determined using *meltR.A* followed by the *BLTrimmer* versus the known Δ*H*° for 96 two-state folding helices. Vertical error bars represent 95% confidence intervals. **(B)** Δ*S*° determined using *meltR.A* followed by the *BLTrimmer* versus the known Δ*S*°. Error bars are the same as in panel A. **(C)** Δ*G*°$_{37}$ determined using *meltR.A* followed by the *BLTrimmer* versus the known Δ*G*°$_{37}$. Error bars are the same as in panel A. **(D)** T_m at 0.1 mM determined using *meltR.A* followed by the *BLTrimmer* versus the known T_m at 0.1 mM. Error bars are the same as in panel A. **(E)** Agreement between *MeltR* and *MeltWin* methods and the consensus nearest neighbor model for 28 different datasets.[7,9–12] Each point represents a fit to a data set collected on one helix. A Kruskal-Wallis one-way test is a non-parametric analysis of variance (ANOVA) test to determine the likelihood that observations originated from the same distribution. Thus, there is a 91% probability that the errors produced by each program or method can be explained by a single distribution. Statistical tests were performed using *ggpubr*.[13] The black error bars represent the average absolute error in the Δ*G*°$_{37}$ across the dataset and the magnitude of the error bars is reported above the graph.

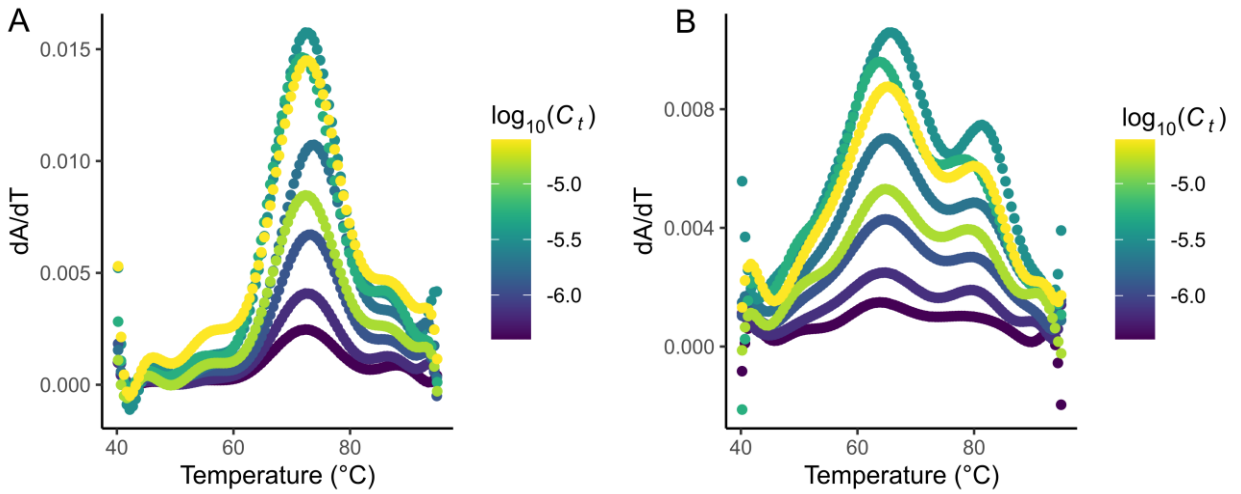


Figure S5 First derivative analysis provided by *MeltR* for long RNA. **(A)** First derivative of melting curves collected on the cleaved-CPEB3 (68 nt) ribozyme. **(B)** First derivative of melting curves collected on a guanine riboswitch aptamer (73 nt).

Table S1 Thermodynamic parameters from fitting raw absorbance melting curves with *MeltR* and *MeltWin*. All parameters are for folding rather than unfolding. Errors are reported as 95% confidence intervals and standard error for *MeltR* and *MeltWin*, respectively.

Helix	MeltR ΔH° Method 1[a] (kcal/mol)	MeltR ΔH° Method 2[b] (kcal/mol)	MeltR ΔH° Method 3[c] (kcal/mol)	MeltWin ΔH° Method 1[a] (kcal/mol)	MeltWin ΔH° Method 2[b] (kcal/mol)	MeltR ΔS° Method 1[a] (cal/mol*K)	MeltR ΔS° Method 2[b] (cal/mol*K)	MeltR ΔS° Method 3[c] (cal/mol*K)	MeltWin ΔS° Method 1[a] (cal/mol*K)	MeltWin ΔS° Method 2[b] (cal/mol*K)	MeltR ΔG°_{37} Method 1[a] (kcal/mol)	MeltR ΔG°_{37} Method 2[b] (kcal/mol)	MeltR ΔG°_{37} Method 3[c] (kcal/mol)	MeltWin ΔG°_{37} Method 1[a] (kcal/mol)	MeltWin ΔG°_{37} Method 2[b] (kcal/mol)	MeltR T_m[d] Method 1[a] (°C)	MeltR T_m[d] Method 2[b] (°C)	MeltR T_m[d] Method 3[c] (°C)	MeltWin T_m[d] Method 1[a] (°C)	MeltWin T_m[d] Method 2[b] (°C)
Self-complementary, homoduplex sequences[e]																				
5′-ACCGGU-3′ 3′-UGGCCA-5′	-55 (-58 to -53)	-58 (-62 to -54)	-56 (-61 to -52)	-54 (4)	-54 (3)	-155 (-163 to -147)	-163 (-176 to -150)	-157 (-174 to -144)	-152 (12)	-151 (9)	-7.07 (-7.12 to -7.01)	-7.12 (-7.18 to -7.07)	-7.09 (-7.20 to -6.96)	-7.11 (0.20)	-7.12 (0.06)	45.0 (44.7 to 45.4)	44.9 (44.5 to 45.3)	45.1 (44.2 to 45.7)	45.4	45.6
5′-AGCCGGCU-3′ 3′-UCGGCCGA-5′	-75 (-76 to -74)	-81 (-83 to -79)	-80 (-82 to -78)	-86 (6)	-86 (3)	-198 (-200 to -196)	-215 (-222 to -209)	-213 (-220 to -206)	-233 (19)	-232 (8)	-13.40 (-13.47 to -13.33)	-13.89 (-14.12 to -13.70)	-13.85 (-14.07 to -13.66)	-13.94 (0.57)	-13.86 (0.23)	72.7 (72.6 to 72.8)	72.1 (72.1 to 72.3)	72.4 (71.8 to 72.9)	69.8	69.7
5′-CCAUGG-3′ 3′-GGUACC-5′	-60 (-61 to -59)	-60 (-61 to -59)	-60 (-61 to -59)	-60 (2)	-60 (1)	-172 (-175 to -169)	-172 (-175 to -169)	-171 (-175 to -167)	-174 (6)	-172 (3)	-6.78 (-6.80 to -6.74)	-6.78 (-6.80 to -6.75)	-6.80 (-6.84 to -6.76)	-6.52 (0.05)	-6.52 (0.01)	42.8 (42.7 to 42.9)	42.8 (42.7 to 42.9)	42.9 (42.8 to 43.0)	41.4	41.4
5′-CGCGCG-3′ 3′-GCGCGC-5′	-56 (-57 to -55)	-56 (-57 to -55)	-57 (-58 to -56)	-56 (1)	-52 (4)	-154 (-158 to -151)	-154 (-158 to -150)	-157 (-161 to -154)	-154 (4)	-143 (11)	-8.14 (-8.19 to -8.11)	-8.13 (-8.18 to -8.10)	-8.22 (-8.28 to -8.14)	-7.85 (0.17)	-7.72 (0.16)	51.3 (51.1 to 51.4)	51.3 (51.0 to 51.4)	51.5 (51.2 to 51.8)	49.6	49.7
5′-GAUAUAUC-3′ 3′-CUAUAUAG-5′	-73 (-74 to -71)	-64 (-66 to -63)	-73 (-74 to -71)	-74 (4)	-66 (3)	-218 (-222 to -213)	-190 (-196 to -186)	-217 (-222 to -212)	-221 (14)	-195 (8)	-5.37 (-5.4 to -5.34)	-5.46 (-5.49 to -5.43)	-5.40 (-5.44 to -5.34)	-5.41 (0.06)	-5.48 (0.04)	35.7 (35.6 to 35.9)	36.0 (35.8 to 36.1)	35.8 (35.5 to 36.0)	35.9	36.1
5′-GCAAUUGC-3′ 3′-CGUUAACG-5′	-78 (-79 to -76)	-76 (-77 to -74)	-81 (-85 to -79)	-79 (4)	-78 (2)	-222 (-226 to -217)	-216 (-221 to -211)	-233 (-244 to -228)	-230 (12)	-226 (5)	-8.62 (-8.68 to -8.54)	-8.54 (-8.58 to -8.47)	-8.74 (-8.86 to -8.68)	-8.15 (0.12)	-8.09 (0.05)	49.3 (49.1 to 49.3)	49.2 (49.1 to 49.3)	49.2 (49.1 to 49.3)	47.0	46.9
5′-UAUAUAUA-3′ 3′-AUAUAUAU-5′	-62 (-66 to -60)	-62 (-65 to -59)	-64 (-67 to -62)	-63 (2)	-63 (3)	-194 (-207 to -185)	-193 (-204 to -183)	-200 (-208 to -194)	-196 (7)	-194 (9)	-2.15 (-2.33 to -1.86)	-2.17 (-2.36 to -1.92)	-2.10 (-2.20 to -2.0)	-2.27 (0.09)	-2.30 (0.17)	20.4 (20.1 to 20.5)	20.4 (20.1 to 20.5)	20.7 (20.5 to 20.8)	21.1	21.1
Non-self-complementary, heteroduplex sequences																				
5′-CGAAAGGU-3′ 3′-GCUUUCCA-5′	-68 (-69 to -67)	-68 (-70 to -66)	-67 (-69 to -66)	-67 (3)	-66 (1)	-188 (-192 to -185)	-188 (-193 to -183)	-185 (-191 to -181)	-185 (9)	-181 (5)	-9.78 (-9.85 to -9.71)	-9.76 (-9.85 to -9.68)	-9.73 (-9.82 to -9.66)	-9.80 (0.09)	-9.77 (0.06)	52.5 (52.3 to 52.7)	52.5 (52.3 to 52.6)	52.5 (52.3 to 52.8)	52.9	53.0
5′-CGUUGC-3′ 3′-GCAACG-5′	-51 (-53 to -50)	-49 (-52 to -47)	-52 (-54 to -50)	-51 (4)	-48 (2)	-143 (-150 to -139)	-135 (-146 to -128)	-143 (-152 to -139)	-141 (11)	-132 (6)	-7.02 (-7.07 to -6.95)	-6.96 (-7.01 to -6.88)	-7.06 (-7.14 to -6.96)	-7.21 (0.09)	-7.18 (0.02)	39.9 (39.5 to 40.3)	39.8 (39.1 to 40.2)	40.2 (39.7 to 40.6)	41.2	41.2
5′-CUGAGUC-3′ 3′-GACUCAG-5′	-64 (-65 to -63)	-64 (-65 to -63)	-63 (-64 to -63)	-63 (2)	-64 (2)	-178 (-181 to -175)	-178 (-181 to -175)	-175 (-178 to -173)	-175 (6)	-176 (5)	-9.08 (-9.11 to -9.05)	-9.08 (-9.12 to -9.05)	-9.08 (-9.14 to -9.04)	-9.15 (0.03)	-9.16 (0.05)	49.8 (49.7 to 50.0)	49.8 (49.7 to 49.9)	50.0 (49.8 to 50.2)	50.4	50.4
F-CGAAAGGU-3′[f] Q-GCUUUCCA-5′	-84 (-88 to -80)	-83 (-88 to -79)	-81 (-85 to -78)	-81 (2)	-81 (7)	-226 (-237 to -215)	-225 (-238 to -212)	-217 (-228 to -210)	-218 (5)	-21 7 (20)	-13.7 5 (-14.06 to -13.49)	-13.7 2 (-14.04 to -13.43)	-13.46 (-13.7 to -13.26)	-13.63 (0.09)	-13.62 (0.54)	66.3 (66.0 to 66.7)	66.3 (65.9 to 66.7)	66.1 (65.6 to 66.6)	66.7	66.7
F-CGUUGC-3′[f] Q-GCAACG-5′	-58 (-61 to -56)	-62 (-68 to -51)	-60 (-62 to -58)	-61 (3)	-62 (4)	-156 (-163 to -150)	-168 (-186 to -133)	-161 (-166 to -154)	-164 (10)	-167 (11)	-10.13 (-10.32 to -9.98)	-10.34 (-10.68 to -9.93)	-10.17 (-10.32 to -10.01)	-10.38 (0.16)	-10.4 (0.21)	57.4 (56.9 to 58.6)	57.2 (56.6 to 59.0)	57.0 (56.7 to 57.3)	57.7	57.6
F-CUGAGUC-3′[f] Q-GACUCAG-5′	-75 (-78 to -74)	-75 (-78 to -74)	-75 (-77 to -73)	-74 (2)	-75 (6)	-200 (-209 to -196)	-200 (-208 to -195)	-201 (-205 to -194)	-197 (7)	-198 (18)	-13.17 (-13.45 to -13.04)	-13.17 (-13.43 to -13.01)	-13.25 (-13.4 to -13.06)	-13.19 (0.28)	-13.21 (0.51)	67.0 (66.9 to 67.2)	67.0 (66.8 to 67.2)	67.3 (67.0 to 67.7)	67.5	67.5
Monomolecular, self-structured helices (tetraloop hairpins)[e]																				
5′-GCCGUGAGGC-3′	-41 (-42 to -39)		-40 (-42 to -38)	-38 (2)		-118 (-122 to -115)		-116 (-123 to -110)	-112 (6)		-3.95 (-4.05 to -3.85)		-3.80 (-4.05 to -3.66)	-3.73 (0.16)		70.3 (70.0 to 70.5)		70.3 (70.0 to 70.5)	70.5	
5′-GCCUAACGGC-3′	-34 (-35 to -33)		-31 (-34 to -29)	-33 (1)		-101 (-104 to -98)		-92 (-99 to -84)	-97 (4)		-2.94 (-3.05 to -2.85)		-2.66 (-2.81 to -2.43)	-2.88 (0.12)		66.1 (65.5 to 66.7)		66.1 (65.5 to 66.7)	66.8	
5′-GCGGCAACGC-3′	-35 (-37 to -33)		-37 (-40 to -34)	-34 (3)		-106 (-112 to -100)		-111 (-122 to -103)	-104 (9)		-2.24 (-2.32 to -2.15)		-2.26 (-2.40 to -2.15)	-2.25 (0.27)		58.0 (57.3 to 58.5)		58.0 (57.3 to 58.5)	58.8	
5′-GCUGAAAGGC-3′	-3 7 (-39 to -36)		-37 (-42 to -34)	-34 (4)		-116 (-123 to -111)		-116 (-129 to -105)	-106 (12)		-1.47 (-1.53 to -1.35)		-1.41 (-1.53 to -1.27)	-1.30 (0.05)		49.5 (47.9 to 50.1)		49.5 (47.9 to 50.1)	49.3	
5′-GCUGAGAGGC-3′	-35 (-37 to -33)		-36 (-40 to -33)	-35 (3)		-110 (-117 to -103)		-113 (-127 to -101)	-110 (11)		-1.23 (-1.27 to -1.20)		-1.13 (-1.21 to -1.07)	-1.27 (0.07)		48.0 (47.6 to 48.6)		48.0 (47.6 to 48.6)	48.6	

Monomolecular, self-structured helices (triloop hairpins)[e]												
5′-GCCGUUUCGGC-3′	-46 (-51 to -42)	-46 (-50 to -38)	-43 (4)	-134 (-150 to -121)	-133 (-146 to -112)	-126 (11)	-4.59 (-5.12 to -4.21)	-4.34 (-4.93 to -3.16)	-4.40 (0.17)	71.0 (69.6 to 71.9)	71.0 (69.6 to 71.9)	72.0
5′-GCCUUUUAGGC-3′	-35 (-38 to -32)	-37 (-40 to -28)	-36 (2)	-102 (-109 to -94)	-107 (-117 to -82)	-106 (5)	-3.47 (-3.65 to -3.25)	-3.54 (-3.9 to -2.87)	-3.45 (0.15)	71.0 (70.2 to 71.8)	71.0 (70.2 to 71.8)	69.7
5′-GGCAAAAUGCC-3′	-31 (-33 to -29)	-32 (-36 to -29)	-30 (2)	-92 (-100 to -86)	-95 (-108 to -86)	-89 (7)	-2.33 (-2.48 to -2.14)	-2.29 (-2.40 to -2.21)	-2.36 (0.17)	62.1 (60.7 to 63.3)	62.1 (60.7 to 63.3)	63.5
5′-GGCCACAGGCC-3′	-39 (-43 to -35)	-36 (-44 to -27)	-38 (2)	-111 (-123 to -99)	-104 (-127 to -78)	-110 (5)	-4.32 (-4.75 to -3.79)	-3.90 (-4.96 to -2.73)	-4.16 (0.17)	75.6 (73.7 to 78.0)	75.6 (73.7 to 78.0)	74.9
5′-GGCGAGACGCC-3′	-36 (-38 to -35)	-37 (-39 to -36)	-37 (3)	-107 (-111 to -104)	-110 (-116 to -105)	-109 (8)	-3.15 (-3.23 to -3.07)	-2.89 (-3.16 to -2.71)	-3.11 (0.25)	66.2 (65.0 to 67.0)	66.2 (65.0 to 67.0)	65.4
5′-GGCUAACGGCC-3′	-34 (-35 to -32)	-33 (-39 to -31)	-35 (1)	-100 (-104 to -94)	-99 (-115 to -91)	-102 (4)	-2.77 (-2.87 to -2.66)	-2.81 (-3.01 to -2.65)	-2.91 (0.08)	64.7 (63.7 to 65.7)	64.6 (63.7 to 65.7)	65.6
Monomolecular, self-structured helices (pentaloop hairpins)[e]												
5′-CGAGGUAAGCUCG-3′	-34 (-36 to -32)	-35 (-38 to -33)	-36 (2)	-102 (-109 to -96)	-107 (-116 to -101)	-107 (7)	-2.14 (-2.38 to -1.92)	-2.21 (-2.52 to -2.02)	-2.48 (0.13)	57.9 (56.9 to 58.9)	57.9 (56.9 to 58.9)	60.2
5′-CGUGUUCGUCACG-3′	-44 (-46 to -43)	-45 (-47 to -42)	-44 (4)	-133 (-138 to -128)	-136 (-140 to -126)	-131 (14)	-3.13 (-3.26 to -3.03)	-3.17 (-3.27 to -3.00)	-3.23 (0.11)	60.4 (60.2 to 60.5)	60.4 (60.2 to 60.6)	61.6
5′-GACGCUCAACGUC-3′	-38 (-39 to -37)	-38 (-40 to -35)	-33 (4)	-113 (-116 to -109)	-112 (-121 to -105)	-99 (14)	-2.77 (-2.83 to -2.71)	-2.66 (-2.8 to -2.47)	-1.92 (0.08)	61.4 (60.9 to 61.9)	61.4 (60.9 to 61.9)	56.3
5′-GCACUGUUCGUGC-3′	-44 (-45 to -42)	-43 (-45 to -41)	-44 (3)	-128 (-132 to -123)	-125 (-133 to -120)	-131 (10)	-3.75 (-3.86 to -3.59)	-3.59 (-3.83 to -3.39)	-3.67 (0.05)	66.1 (65.9 to 66.3)	66.1 (65.9 to 66.3)	65.1

[a]Determined with method 1, fitting curves individually and averaging the results.
[b]Determined with method 2, $1/T_m$ versus natural logarithm of C_t analysis.
[c]Determined with method 3, globally fitting all curves with the same non-linear model.
[d]Expectation maximized T_m at a C_t of 0.1 mM.
[e]Homoduplex,[9] pentaloop hairpin,[12] tetraloop hairpin,[11] and triloop hairpin[10] datasets were compiled from a published source.
[f]F represents a 5′-FAM label, and Q represents a 3′-BHQ1 label.
[g]Underline represents loop residues in a self-structured hairpin loop.

Table S2 Thermodynamic parameters from fitting raw absorbance melting curves of long RNA with *MeltR* and *MeltWin*. All parameters are for folding rather than unfolding.

RNA	Method	Δ*H*° (kcal/mol)	Δ*H*° Error[a] (kcal/mol)	Δ*S*° (kcal/mol)	Δ*S*° Error[a] (kcal/mol)	Δ*G*°$_{37}$ (kcal/mol)	Δ*G*°$_{37}$ Error[a] (kcal/mol)	T_m at 0.1 mM (°C)	T_m at 0.1 mM error[a] (°C)
CPEB3	*MeltR* 1 Individual fits	-75	-77 to -73	-216	-223 to -211	-7.58	-7.81 to -7.40	71.9	71.7 to 72.0
CPEB3	*MeltR* 3 Global fit	-73	-75 to -71	-212	-217 to -206	-7.39	-7.63 to -7.17	71.9	71.7 to 72.0
CPEB3	*MeltWin* 1 Individual fits	-73	2	-212	5	-7.39	0.10	71.8	

[a]Errors are reported as 95% confidence intervals and standard error for *MeltR* and *MeltWin*, respectively. Note, *MeltWin* does not supply uncertainty for the T_m at 0.1 mM.

Supplementary References

(1) *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*; Cleveland: CRC Press, 1975.
(2) Puglisi, J. D.; Tinoco, I. Absorbance Melting Curves of RNA. In *Methods in Enzymology*; RNA Processing Part A: General Methods; Academic Press, 1989; Vol. 180, pp 304–325. https://doi.org/10.1016/0076-6879(89)80108-9.
(3) Gralla, J.; Crothers, D. M. Free Energy of Imperfect Nucleic Acid Helices: III. Small Internal Loops Resulting from Mismatches. *J. Mol. Biol.* **1973**, *78* (2), 301–319. https://doi.org/10.1016/0022-2836(73)90118-6.
(4) *Nonlinear Regression with R*; Ritz, C., Streibig, J. C., Eds.; Springer New York: New York, NY, 2009. https://doi.org/10.1007/978-0-387-09616-2.
(5) R Core Team. R: A Language and Environment for Statistical Computing. *R* **2022**.
(6) Bevington, P. *Data Reduction and Error Analysis for the Physical Sciences*; McGraw-Hill, 1969.
(7) Xia, T.; SantaLucia, J.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner, D. H. Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson−Crick Base Pairs. *Biochemistry* **1998**, *37* (42), 14719–14735. https://doi.org/10.1021/bi9809425.
(8) Antao, V. P.; Tinoco, I., Jr. Thermodynamic Parameters for Loop Formation in RNA and DNA Hairpin Tetraloops. *Nucleic Acids Res.* **1992**, *20* (4), 819–824. https://doi.org/10.1093/nar/20.4.819.
(9) Adams, M. S.; Znosko, B. M. Thermodynamic Characterization and Nearest Neighbor Parameters for RNA Duplexes under Molecular Crowding Conditions. *Nucleic Acids Res.* **2019**, *47* (7), 3658–3666. https://doi.org/10.1093/nar/gkz019.
(10) Thulasi, P.; Pandya, L. K.; Znosko, B. M. Thermodynamic Characterization of RNA Triloops. *Biochemistry* **2010**, *49* (42), 9058–9062. https://doi.org/10.1021/bi101164s.
(11) Sheehy, J. P.; Davis, A. R.; Znosko, B. M. Thermodynamic Characterization of Naturally Occurring RNA Tetraloops. *RNA* **2010**, *16* (2), 417–429. https://doi.org/10.1261/rna.1773110.
(12) Saon, Md. S.; Znosko, B. M. Thermodynamic Characterization of Naturally Occurring RNA Pentaloops. *RNA* **2022**, *28* (6), 832–841. https://doi.org/10.1261/rna.078915.121.
(13) Kassambara, A. Ggpubr: “ggplot2” Based Publication Ready Plots. *R Package* **2020**, *0.4.0*.