Best practice for improved accuracy: a critical reassessment of van't Hoff analysis of melt curves

Jacob M. Majikes^{1,*}, Michael Zwolak¹, and J. Alexander Liddle^{1,*}

¹ Microsystem and Nanotechnology Division, National Institute of Standards and Technology, Gaithersburg, MD, 20899, USA

* To whom correspondence should be addressed. Tel: +1 301 975 6050; Email: jacob.majikes@nist.gov, liddle@nist.gov.

Abstract:

Biomolecular thermodynamics, particularly for DNA, are frequently determined *via* van't Hoff analysis of optically-measured melt curves. Accurate and precise values of thermodynamic parameters are essential for the modelling of complex systems involving cooperative effects, such as RNA tertiary structure and DNA origami because the uncertainties associated with each motif in a folding energy landscape can compound, significantly reducing the power of predictive models. We follow the sources of uncertainty as they propagate through a typical van't Hoff analysis to derive best practices for melt experiments and subsequent data analysis, assuming perfect signal baseline correction. With appropriately designed experiments and analysis, a van't Hoff approach can provide surprisingly high precision, e.g., enthalpies may be determined with a precision as low as a 10⁻² kJ·mol⁻¹ for an 8 base DNA oligomer.

Statement of Significance:

Models to predict the behavior of complex cooperative nucleic acid systems, e.g., DNA origami, require high-quality thermodynamic data. van't Hoff analysis is a ubiquitous tool for extracting thermodynamic parameters from ensemble measurements. In nucleic acid systems, where specific reactions can be readily labelled and melt curves gathered with high throughput, van't Hoff analysis can rapidly extract thermodynamic data. Ensuring that this data has the necessary precision requires that numerous experimental design decisions be optimized. We identify best practices that can improve precision by several orders of magnitude for minimal additional effort for a typical van't Hoff analysis.

Introduction:

Biochemical and biomimetic systems exhibit complex responses to their environments because they sample numerous conformational states, which differ by amounts $\geq kT$ ($\approx 2.6 \text{ kJ} \cdot \text{mol}^{-1}$ or 0.6 kcal·mol⁻¹) (1). For example, proteins are able to sample states over a wide range of energies, even when well-folded, because their intrinsic energy fluctuations, resulting from their coupling with local energy fluctuations in the surrounding water, can be much larger than kT – on the order of 42 kJ·mol⁻¹ to 84 kJ·mol⁻¹ (10 kcal·mol⁻¹ to 20 kcal·mol⁻¹ or 17 ×kT to 34 ×kT)(1). Conformational changes are important in many situations, including binding of antibodies to pathogens to trigger immune response(2); ligand binding, e.g., of carbon monoxide to myoglobin in blood(3); and albumin binding, relevant in drug clearance from the body (4, 5). For any biomacromolecule, the conformation is the sum of the states of all the motifs of which it is comprised. The cooperation or competition between these motifs defines the overall conformational and energetic landscape and thus the interactions with the environment.

Designing and engineering biomolecular and biomimetic energy landscapes to yield the desired range of conformational states requires reliable and accurate predictive tools. Developing such tools is challenging for many reasons(6), not the least of which is that the uncertainties in the motif energetics compound as they compete or cooperate. This significantly raises the bar for acceptable uncertainties above what is needed to predict the behavior of those motifs in isolation.

Regardless of the technique used to obtain thermodynamic parameters, any mismatch between the assumed mechanism, or number of states, and what actually occurs in a sample can significantly

degrade predictions based on those parameters(7). It is thus important to draw a distinction between *model-form uncertainties* and *experimental uncertainties*. Experimental uncertainties, as the name suggests, come from the execution of an experiment and include variations in sample preparation, equipment noise (precision), and inaccuracy. Model-form uncertainties occur in analysis, or in prediction, when the mathematical equation used to describe a system does not reflect reality. Such uncertainties must be evaluated using uncertainty propagation for each model choice.

Examples of model-form uncertainties in nucleic acid systems might include assuming that a DNA system with multiple multivalent domains, e.g., origami, will form one or two states when in reality it might form three, four, or, if polymerization occurs, an indeterminate number(8). Similarly, it is common to assume a dsDNA strand will melt in a two-state process, but if the strand is long enough or stable enough, it may sample many partially melted states(9). Fitting the ensemble information to an incorrect model will create significant inaccuracies, and is an easy mistake to make(7).

In the experimental measurement of thermodynamics, calorimetry is the gold standard as it directly measures evolved or absorbed heat. However, all flavors of calorimetry require a large sample mass to guarantee sufficient signal-to-noise, and will detect heat from the entire sample, not just the reaction of interest. Calorimetry is also comparatively low-throughput, the thermal isolation and long equilibration times needed for heat measurements are rarely amenable to rapid data acquisition. It is also worth noting that entropy values from calorimetry are not directly measured but are estimated through input of measured enthalpies and temperatures into the van't Hoff equation.

Given the limitations of calorimetry, fitting equilibrium constants from melt curves is a common alternative for extracting thermodynamic parameters. These curves are typically derived from optical or electrochemical signals, that, with appropriate experimental design, can be made specific to the reaction of interest(10, 11) and are compatible with high-throughput measurement, e.g., covalent modification of a target structure to enable Förster resonant energy transfer (FRET)(1, 12). The benefits of specificity and ease of data collection come at the cost of more complex uncertainty propagation from the measurement, through the statistical analysis, to the extracted thermodynamic properties and subsequent predictions.

While the superficial ease with which van't Hoff analysis can be performed has made it ubiquitous,(13, 14) researchers seeking reliable results must navigate between the Scylla of reliance on simplistic approaches and the Charybdis of abandoning valuable data because of van't Hoff analysis's reputation as "in practice inaccurate"(14). We believe it timely to re-examine this topic, given the increasing popularity of high-resolution melting (HRM) DNA experiments(15, 16) and the growth of DNA nanotechnology(17–19).

In revisiting uncertainty propagation in van't Hoff analysis of melt curves, *we focus on sources of uncertainty that are common to most or all melt curve analyses,* and thus emphasize the role of experimental uncertainties over system-specific model-form uncertainties. While the rest of this article focuses on nucleic acids, with an eye towards modelling DNA nanotechnology, the insights and best practices we address are more broadly applicable.

A robust discussion of best practice may also help to reconcile the divergences in advice and experience that arise in van't Hoff analysis of nucleic acid melt curves. As the numerous tertiary structure prediction tools(20–23) and duplex thermodynamic webtools(24, 25) develop, it would be beneficial to know the degree to which minor disagreements in their results(26, 27) are definitional(28), represent meaningful distinctions, or are actually in agreement within uncertainty.

A similar discussion of van't Hoff analysis has occurred recently in the Isothermal Titration Calorimetry (ITC) community(28). ITC simultaneously generates both calorimetric thermodynamic parameters and the ensemble information necessary for van't Hoff analysis(29). Comparisons between values derived by each approach and discrepancies between them led to a reevaluation of uncertainty propagation for ITC(30–34). For a general paper that addresses experimental design we recommend work by Zhukov and Karlsson(35).

In nucleic acid systems, especially those employed in DNA nanotechnology(36, 37), one must contend with complex topologies comprising a variety of tertiary and quaternary motifs. Predicting their energy landscapes involves grappling with the uncertainty in the energetics of these motifs as well as the sheer number of possible events in the folding funnel. For example, loop entropy in DNA origami(38), is not well-captured by existing models for biological systems, and requires sufficient sample mass to confound calorimetry data with significant background signal. This makes high-quality van't Hoff analysis of optical melt curves, whose signal is only correlated to the motif of interest, highly desirable.

The energetic contributions of such a motif are extracted from melt curves by measuring some reference sequence with, and without, the motif of interest, as shown in Fig. 1 B. The energetics of this motif, with its uncertainty, are then applied to some other system of interest, shown in Fig.1 C, to predict its behavior. When predicting complex systems, the uncertainty of a single measurement will be amplified by the number of motifs and competing states or structures, significantly raising the bar for what constitutes a sufficiently small uncertainty to enable useful prediction.



Fig. 1. A - The uncertainties on motif thermodynamics for loops, aptamers, etc., compound in complex systems, including nucleic acid tertiary structures. These comprise numerous motifs and often contain multiple competing states, as illustrated by two possible tertiary structures for 255 bases of the M13mp18 origin of replication. Energetic uncertainties on folding motifs are important in DNA origami where a large viral scaffold is incrementally folded into a rigid shape during the annealing process. B - Typically, the excess energetics, e.g., loop entropy, for motifs which cannot be measured independently are found by comparing separate measurements of a reference sequence (green), and reference sequence attached to the motif (black). C - Data from the motif is used to predict its effect on a new system (blue). A large uncertainty (20 %) in the energetics of the motif propagates, leading to a large spread in predicted melt temperature for the new system.

For both nanotechnology and tertiary structure prediction, it is worth mentioning the toehold-mediated catalysis method for designing a van't Hoff experiment(39, 40). While this technique also feeds

ensemble information through the van't Hoff equations, by clever design of competing strand displacement, it directly probes competition between a sequence with and without a motif of interest in a single sample, which significantly reduces uncertainty. We note that that it is best suited for motifs(40) with modest energetic contributions, e.g., dangling ends and mismatches, while melt-curve analysis is better suited to motifs whose energetic contributions are larger, e.g., aptamers and loops, but not large enough to make calorimetry a viable option.

The relationship between the energetics of a reference and motif for measurement may most easily be derived for unimolecular systems, shown in SI Section 2, and depicted in Fig. 2. Ideally, a motif should shift the melt curve as much as possible, but not beyond the measurable range in liquid water. To satisfy this constraint, in a typical unimolecular system, the motif may contribute as much as 20 % of ΔG_{total} . Larger shifts can be tolerated in bimolecular systems by changing concentrations and making use of the concentration dependence of T_m to move the melt curve back into the measurable range. *However, for systems where the energetics are not precisely known* a priori *a useful heuristic is that the energetics of the motif should be* ≈ 10 % of those of the reference. Toehold-based techniques which measure the ensemble between the reference strand with, and without, the motif directly would correspond to a ΔG_{ref} of zero(39, 40). This allows for precise evaluation of very small ΔG_{motif} , as in dangling ends, but precludes measurement of motifs with a larger energetic contribution(40).



Fig. 2. Relationship between the relative stability of the motif of interest and reference sequence. There is a clear optimal measurement zone in which the melt curve is shifted through as much of the measurable range as possible but not beyond it. The vertical dashed line denotes the end of the measurable range and the boiling point of water.

This 10 % rule of thumb is especially useful as the uncertainties on the extracted thermodynamics are proportional to the magnitude of the terms themselves. As there are two measurements whose uncertainties are added in quadrature, $\sigma_{motif} = \sqrt{\sigma_{ref}^2 + \sigma_{ref+motif}^2} = \sqrt{\sigma_{ref}^2 + 1.1^2 \sigma_{ref}^2}$, we can therefore approximate the uncertainty on the motif to be slightly greater than $\sqrt{2}\sigma_{ref}$. As the reference must be proportionally larger than the motif, i.e., $10\times$, such an experiment is expected to have $10 \cdot \sqrt{2}$ higher uncertainty than one directly measuring the motif energetics.

Evaluating these uncertainties is further complicated by the ease with which the thermodynamics are perturbed. If FRET pairs, rather than UV-vis, circular dichroism, or electrochemistry, are used to obtain

the ensemble information in Fig.1 and Fig. 2, they can shift hybridization thermodynamics in a sequence-dependent manner(41), similarly, dangling DNA bases make sequence-specific energetic contributions to hybridization(40, 42, 43).

For van't Hoff analysis of melt curves, perturbations from FRET pairs will cancel as they exist in both the reference and motif sample. However, they will contribute to uncertainty the same way any change in the reference energetics would, i.e., by increasing ΔG_{ref} . The contributions of dangling ends(44, 45), or base stacking, between the motif and reference sequence do not cancel in this way. However, any complex system of interest will have such a contribution where the motif is connected to the system. We consider this a system-specific model-form consideration and encourage the experimentalist to carefully describe the context of the measurement and the theorist to carefully consider that context when developing their predictions.

We revisit the sensitivity of van't Hoff analysis of optically measured nucleic acid melt curves to experimental uncertainties through simulation. *In doing so, we aim to develop a physical intuition of how uncertainty propagates through van't Hoff analysis of melt curves to the extracted thermodynamic parameters, and how experimental choices can exacerbate or minimize that uncertainty.*

Most of our recommendations can be inferred from the observation that for van't Hoff analysis, the measured quantity is the melt curve, rather than the thermodynamic quantities themselves. Anything which allows better sampling of the distribution of melted versus unmelted states will significantly increase the accuracy of the thermodynamic parameter extraction. We additionally observe that both neglect of ΔC_p and the use of numerical derivatives significantly reduce the quality of the analysis.

Finally, we reiterate that we leave many model-form considerations for future work. We assume a statistically rigorous signal baseline subtraction technique which accurately represents the ensemble, and do not consider the role of the variety of available approaches(46–48). Similarly, we neglect the role of least-squares algorithm settings in finding the broad and shallow minimum in fit quality as a function of ΔH° , ΔS° , and ΔC_p , nor do we consider advanced fitting techniques which fit the melt curve directly using nonlinear regression.

Van't Hoff analysis at a glance:

The goal of van't Hoff analysis is to use ensemble information to extract thermodynamic parameters for a state change. The van't Hoff relation, Eq.(1), connects the equilibrium constant, K, with the Gibbs free energy change between states, $\Delta G(T)$, gas constant, R, and temperature, T.

The concentrations of both products and reactants are found by assuming the melt curve represents the fraction of ssDNA (F_{ssDNA}), and that the transition is from fully hybridized dsDNA to fully melted ssDNA. If these assumptions are not valid, additional care must be taken in the analysis(49). Eqs. (2) and (3) describe the equilibrium constant, [K], for bimolecular and unimolecular cases, respectively. [C] is the known concentration of the component strands, which cancels out for unimolecular reactions, and σ is the uncertainty on F_{ssDNA} . To limit the number of variables being tested, we assume for bimolecular systems that [Strand1] = [Strand2] in Eq. (2).

Eq.(1):
$$[K] = e^{\frac{-\Delta G(T)}{RT}}$$

Eq.(2): $[K_{bi}] = \frac{[Products]}{[Reactants]} = \frac{[dsDNA]}{[Strand1][Strand2]} = \frac{1 - (F_{ssDNA} \pm \sigma)}{(F_{ssDNA} \pm \sigma)(F_{ssDNA} \pm \sigma)[C]}$
Eq.(3): $[K_{uni}] = \frac{[Hairpin]}{[ssDNA]} = \frac{1 - (F_{ssDNA} \pm \sigma)}{(F_{ssDNA} \pm \sigma)}$

The right side of Eq.(1) gives the thermodynamic quantities of interest, specifically the temperaturedependent change in Gibbs free energy, $\Delta G(T)$, which is further divided into temperature dependent enthalpy, $\Delta H(T)$, and entropy, $\Delta S(T)$. These in turn are described as temperature independent terms, ΔH° and ΔS° , where temperature dependence is represented by the change in heat capacity, ΔC_p , at a given reference temperature, T_{ref} , as shown in Eq.(4), Eq.(5), and Eq.(6).

Eq.(4):
$$ln([K]) = \frac{-\Delta G(T)}{RT} = -\frac{\Delta H(T)}{RT} + \frac{\Delta S(T)}{R}$$

Eq.(5): $\Delta H(T) = \Delta H^{\circ} + \Delta C_p (T - T_{ref})$
Eq.(6): $\Delta S(T) = \Delta S^{\circ} + \Delta C_p ln \left(\frac{T}{T_{ref}}\right)$

These combine to give Eq.(7), which may be fitted to extract thermodynamic quantities from ensemble information as a function of temperature.

Eq.(7):
$$ln([K]) = -\frac{\Delta H^{\circ}}{RT} + \frac{\Delta S^{\circ}}{R} - \frac{\Delta C_p}{R} \left(\left(1 - \frac{T_{ref}}{T}\right) - ln\left(\frac{T}{T_{ref}}\right) \right)$$

Typical analysis of optical data involves plotting $\ln([K])$ vs 1/T, and fitting the ΔH° , ΔS° , and ΔC_p as the slope, the intercept at infinite temperature, and the curvature respectively. From Eq.(7) it is not straightforward to determine how uncertainties on T and on F_{ssDNA} will propagate through to ΔH° , ΔS° , and ΔC_p . Variations in concentration due to uncertainties in pipetted volume have a negligible effect on the extracted parameters as they typically occur in single-digit percentages, and the concentration terms are inside a logarithm, shown further in SI Section 11. We therefore neglect pipetting uncertainty, however, for samples prepared *via* serial dilution, uncertainties compound, and may become non-negligible.

Three important insights from Eq.(2) and Eq.(3) may be obtained at a glance. First, we see that as F_{ssDNA} approaches the values of 0 and 1, the uncertainty, σ , will dominate [K]. This is why many approaches to van't Hoff analysis only use the melt temperature, T_m , defined as the temperature where $F_{ssDNA} = 0.5$. At the T_m , uncertainty in [K] is at its lowest. A more phenomenological explanation is to say that the signal-to-noise ratio of the measured reaction is highest in the center of the sigmoid and lowest on the plateaus.

The second insight is that noise in F_{ssDNA} propagate asymmetrically into $\ln([K])$. By examining Eq.(2), we can see that, at the plateau where $F_{ssDNA} = 1$, noise creates only small perturbations in [K], while at the plateau where $F_{ssDNA} = 0$ the small denominator amplifies the effects of noise and thus perturbations in [K]. This can distort the fit, particularly if normalization or selection of fractional data range is sub-optimal, discussed in section 15 of the SI.

The third insight is the statistical nature of entropy-enthalpy compensation. The relatively narrow range of accessible 1/T values, coupled with the inevitable experimental uncertainties in $\ln([K])$ means that a shift in fitted slope, ΔH° , will result in a compensating fit in intercept, ΔS° , as illustrated in Fig. 3. While entropy-enthalpy compensation may exist as a result of genuine physical phenomena, one should be cautious in inferring its existence from thermodynamic data obtained *via* van't Hoff analysis(50, 51).



Fig. 3. Illustration of extraction of entropy and enthalpy from a van't Hoff plot, and mapping of uncertainty in K on to statistical compensation in ΔH° and ΔS° . The uncertainty envelope in *K* is narrowest at the T_m . The red and blue lines illustrate compensating fits that could appear equally valid in the presence of experimental uncertainty. For graphic simplicity the effects of ΔC_p , are neglected in this illustration.

While Fig. 3 neglects curvature in $\Delta H(T)$ and $\Delta S(T)$, i.e., ΔC_p , for visual clarity, it is of sufficient import to address directly. While the choice of T_{ref} is arbitrary (typically standard temperature, 25 °C, or physiological temperature, 37 °C), ΔC_p describes the curvature in the van't Hoff plot (example in SI Section 1) and the T_m is the center of that curvature for any given melt experiment. The uncertainty in ΔC_p therefore propagates from that data center of mass at the T_m on to both $\Delta H(T)$ and $\Delta S(T)$ as a function of $(T-T_m)$, as shown in Eq.(8) and Eq.(9) (assuming no correlation in uncertainty between ΔH° and ΔC_p , and no uncertainty in temperature) and which is readily derived from Eq.(5).

Eq.(8):
$$\Delta H(T) = \Delta H^{\circ}_{T_{ref}=T_m} + \Delta C_p(T - T_m)$$

Eq.(9): $\sigma_{\Delta H(T)} = \sqrt{\sigma_{\Delta H^{\circ}_{T_{ref}=T_m}}^2 + (\sigma_{\Delta C_p}(T - T_m))^2}$

The initial T_m must be reported in order to recapitulate temperature dependent uncertainties and evaluate the excess thermodynamics of arbitrary motifs, since their temperature dependent uncertainties become more complicated, as the $\sigma_{\Delta H(T)}$ from two separate measurements (Fig.1 b) are added in quadrature.

The uncertainty in the ΔH° and ΔS° at a reference temperature, whether standard or physiological, will be amplified by the uncertainty in ΔC_p by $(T_{ref} - T_m)$. Given that the T_m changes as the phenomenon of interest shifts the energetics, *this can cause non-negligible systematic trends in uncertainty*. Predictions of uncertainty over a broad temperature range require both the uncertainty on ΔC_p and the value of T_m .

It is not possible to simply define what constitutes a sufficiently precise uncertainty. A motif with some kiloJoules per mole of uncertainty in its enthalpy might contribute more or less to the envelope of predicted behavior depending on whether that motif contributes an entropy change and whether compensation is accounted for. Similarly, a measurement that is good enough to predict behavior close to the T_m might fail entirely at much higher or lower temperatures. What constitutes an acceptably accurate and precise measurement of thermodynamics depends intrinsically on how one wishes to use it and should always be evaluated on a case-by-case basis.

Methods Simulations: We apply a Monte Carlo method: equilibrium melt curves are calculated from input thermodynamic quantities, noise is then applied to the melt curves to generate pseudo-experimental data, which is then analyzed. This process is automated and can rapidly compare the

precision of extracted thermodynamics for different experimental designs and analysis styles and is discussed in depth in SI section 4 and 5.

Despite the speed and simplicity of each simulation, the size of parameter space in van't Hoff analysis is vast, especially when including experimental choices. Parameters worth considering include values for ΔH° , ΔS° , and ΔC_p , number of sample replicates, concentration, T_m ranges, method for determining T_m , criteria for what fraction of the melt curve to analyze, whether analyzing a unimolecular (hairpin) or bimolecular reaction, etc. These multifarious choices present both a logistical hurdle and a complex data presentation/communication problem.

As our emphasis in this work is on general lessons for experimental design, we bound this space whenever possible. We separately vary noise and accuracy in temperature (x-axis) and noise only on the readout (y-axis). We assume all melt curves report from 0 % to 100 % completion and that any y-axis shifts are removed in normalization. We fix the density of measurements on the x-axis to the temperature imprecision, i.e., an instrument with 0.1 °C temperature imprecision cannot reasonably measure optical intensity at temperature intervals smaller than 0.1 °C.

Finally, we do not address the complex issue of optical melt curve baseline correction here, as it is sufficiently nuanced and signal-dependent that it merits its own in-depth evaluation(46, 47, 52).

Default simulation inputs: We choose a 10 base sequence as being sufficiently short for accurate van't Hoff analysis. We assume measurement uncertainties consistent with those of a typical laboratory PCR system. Our simulation inputs are therefore as follows:

Thermodynamic: Bimolecular reaction, $\Delta H^{\circ} = -292.0 \text{ kJ} \cdot \text{mol}^{-1}$ (-69.8 kcal·mol⁻¹), $\Delta S^{\circ} = -815.9 \text{ J} \cdot \text{mol}^{-1}$. ¹.K⁻¹ (-195 cal·mol⁻¹·K⁻¹), $\Delta C_p = -4.18 \text{ kJ} \cdot \text{mol}^{-1}$.K⁻¹ (-1.0 kcal·mol⁻¹·K⁻¹)

Reference temperature: T_{ref} = 37 °C, physiological temperature, unless otherwise stated

Model uncertainties: 0.1 % readout noise (y-axis), 0.1 °C temperature imprecision, no temperature inaccuracies

Analysis: Fractional data range 0.8, T_m range of ± 10 °C associated with varying concentration for multimelt curve experiments and default concentration of 40 µmol·L⁻¹. When varying concentration both [ssDNA] are set to be equimolar.

Replicates: 2 000 replicates of each data point were performed. This is significantly more than could be performed experimentally and ensures accurate sampling of the uncertainty on the uncertainty envelope of the extracted thermodynamic parameters.

Reporting uncertainties: We have found it simplest to report relative rather than absolute uncertainties. We choose to divide our uncertainty into the Relative Standard Deviation (RStD) to represent variance and precision, Relative Systematic Deviation (RSyD) to represent accuracy, and Relative Total Deviation (RTD) to represent total uncertainty. We use Relative Deviation (RD) for axis labels when these are plotted together. These are shown respectively in Eq.(10), Eq.(11), and Eq.(12), where *i* is the *i*th of *n* replicates, $\widehat{\Delta H}^{\circ}$ is the average extracted enthalpy, and $\Delta H^{\circ}_{expected}$ is the input enthalpy.

Eq.(10):
$$RStD_{\Delta H^{\circ}} = \frac{1}{\Delta H^{\circ}_{expected}} \sqrt{\frac{1}{n} \sum_{i}^{n} (\Delta H^{\circ}_{i} - \widehat{\Delta H^{\circ}})^{2}}$$

Eq.(11): $RSyD_{\Delta H^{\circ}} = \frac{1}{\Delta H^{\circ}_{expected}} \sum_{i}^{n} (\Delta H^{\circ}_{i} - \Delta H^{\circ}_{expected})/n$

Eq.(12):
$$RTD_{\Delta H^{\circ}} = \sqrt{RStD_{\Delta H^{\circ}}^{2} + RSyD_{\Delta H^{\circ}}^{2}}$$

The RStD, RSyD, and RTD are defined similarly for ΔS° , and ΔC_{p} .

We plot RTD and RSyD as a function of any given source of uncertainty. The RTD provides a measure of what might be expected for a typical analysis, which would converge towards the RSyD if an infinite number of samples were analyzed under the same conditions to remove all uncertainty due to random variation.

While the use of relative uncertainties allows for easier generalization of trends and comparison between noise conditions and experimental systems, it can complicate comparison between measurements. In such cases, we may report the total uncertainty in real units, which we will denote as TD, SyD, and StD.

To facilitate evaluation of how 'good' a measurement or experimental design must be, we choose a benchmark uncertainty, which represents the RTD on a single measurement, propagated to the uncertainty on the motif of interest, that would shift the predicted melt curve by 0.25 °C. This condition is satisfied when the relative uncertainty is less than 0.0006 RTD_{ΔH°}, 0.00069 RTD_{ΔS°} or 0.176 kJ·mol⁻¹ and 0.57 J·mol⁻¹·K⁻¹, respectively (derived in SI Section 3). This benchmark is shown as a downward arrow on the y-axis in Figs. 8 and 9, and sets a high bar for what constitutes an acceptable uncertainty. However, it is necessary for quantitative prediction of complex nucleic acid tertiary structure or DNA nanostructures which exhibit multiple layers of cooperativity and competition. As we will show, it is also achievable.

We note that the uncertainties on our uncertainties (RStD, RSyD, & RTD), as quantified by box averaging, are sufficiently low to fit within the data markers, and are thus not plotted.

A note on notation: It is rare to delve into thermodynamics in any depth without being pursued by the spectre of sign errors. In nucleic acid systems, thermodynamics may be referenced either to melting or hybridization. While experimental work often reports in reference to melting, and predictive work often reports in reference to hybridization, this is by no means universal. *In this work, we use and report the energetics of hybridization*

Results & Discussion

Common experimental designs, van't Hoff analyses, and resultant features: There are two common setups of a van't Hoff melt-curve experiment. The first, shown in Fig. 4 A, involves measurement of the melt curve at only one concentration and uses some fraction of the data points around the T_m . This is necessary for unimolecular interactions where varying the concentration will not shift the T_m and allows sampling in $\ln([K])$ vs. 1/T space.



Fig. 4. A - Single curve, full-width analysis illustration and resulting uncertainty as a function of fractional data range for 0.1 %, 1.0 %, and 2.0 % read noise. The upper bound of each shaded region represents the total uncertainty, RTD, and the lower bound represents systematic contributions, RSyD, and thus the smallest possible uncertainty. B - Concentration variation analysis illustration, the table shows experimentally achievable T_m for example oligomers of different lengths of base pairs (bp), and corresponding plots show relative and real-unit uncertainties as a function of the achievable T_m range. It should be noted that a 100 bp strand would be highly unlikely to melt in a two-state reaction.

For this approach, it is important to note that the uncertainties in Eq.(2) and Eq.(3) are magnified at values of F_{ssDNA} close to 0 or 1. In short, the further a data point is away from the T_m the more noise, and the less signal, it contributes to the fitting. There is some ideal fractional data range of the melt curve for which any additional data points will degrade the quality of the analysis. This optimal fractional data range depends on the steepness of the curve, density of data points in T, and read noise of the equipment, as depicted in Fig. 4 A. For our default analysis condition, we choose a fractional data range of 0.8 of the melt curve around T_m , as this is close to the optimum value for the three different read noise conditions shown in Fig. 4 B. This is not a universal, optimal, fractional data range, which we hope to pursue analytically in later work. Rather, this is a reasonable rule of thumb, in line with our understanding of typical practice for van't Hoff analysis of melt curves. This range will avoid contributing more noise than signal except in only the most egregious cases and readily captures the lowest uncertainty for $\ln[[K]]$ in the melt curve.

In the second, and more common experimental setup, illustrated in Fig. 4 B, the concentrations of one or both the ssDNA reactants are varied, while only the T_m is used for the analysis. This minimizes the effect of the uncertainty in F_{ssDNA} , in Eq.(2), but at the cost of much additional experimental work as most of the data available from the melt curve is discarded.

The range over which T_m can be varied by changing concentrations is bounded by solubility and sensitivity limits for the maximum and minimum achievable concentrations, respectively. As longer oligomers experience less of a T_m shift with concentration (see Fig. 4 B) they can only sample a smaller ΔT_m range, resulting in a less precise van't Hoff analysis. The RTD in ΔH° for the example oligomers are plotted as a function of ΔT_m , where dashed RTD lines indicate ΔT_m values outside of the experimentally achievable range. The RTD for longer oligomers at the same ΔT_m is slightly lower than for short oligomers, as the change in concentration for the former is larger at that same ΔT_m , slightly increasing the sampled space in $\ln([K])$. When reported in real units, van't Hoff analysis is far more accurate for weaker interactions, even before accounting for the larger achievable ΔT_m range accessible with them.

It is worth considering how the shape and position of the curves changes the quality of the parameter extraction. This will also provide a good example of temperature dependent uncertainties. Fig. 5 A and B, shows two sets of systems in which the thermodynamics were varied in ways not representative of real DNA strands, in order to independently vary the width and position of the melt curves. In the first, we varied the width of the melt curve by changing the input ΔH° and choosing a value of ΔS° to keep T_m constant, the ΔC_p was scaled with ΔH° to minimize shifts in T_m , and to prevent the curve from exhibiting cold denaturation. In the second, we varied the T_m by keeping ΔS° constant and varying ΔH° . Details for both sets of simulations are given in the SI Section 5. All of these melt curves were analyzed using the default fractional data range of 0.8.



Fig. 5. A - Example illustrating the effect of melt-curve width on uncertainty where the upper line represents the uncertainty on a typical experiment, RTD, while the lower line of a pair represents the systematic contributions, RSyD. B - Large changes in T_m have minimal effects on the RTD and RSyD. Both A and B are evaluated at $T_{ref} = T_m$ as opposed to $T_{ref} = 37$ °C to allow comparison before the projection of uncertainty in ΔC_p across T_m - T_{ref} .

As shown in Fig. 5 A, steeper curves result in higher RTD and RSyD. Additionally, it should be noted that, because narrower widths correspond to larger thermodynamic parameters, the uncertainty in real units is also larger.

The RTD $_{\Delta H^{\circ}}$ of systems with identical melt-curve widths but different T_m , when evaluated at $T_{ref} = T_m$, as plotted in Fig. 5 B, yields essentially identical values of RTD $_{\Delta H^{\circ}}$. Comparison with Fig. 6 A, which shows the temperature-dependent RTD $_{\Delta S(T)}$ and RSyD $_{\Delta S(T)}$, in which the uncertainty on ΔC_p is projected across T onto $\Delta S(T)$, illustrates how important this effect is. It should be noted that for each of these curves the input ΔS° is identical. The difference in size and position of the uncertainty envelope is solely due the difference in T_m and compensation of uncertainty between the varied input ΔH° and ΔS° . In Fig. 6 B the RTD of ΔH° and ΔS° are plotted as a function of T_m at $T_{ref} = 37 \,^{\circ}$ C, as opposed to $T_{ref} = T_m$ in

Fig. 5 B. As one might expect from Fig. 3, the uncertainty on the thermodynamics increases the further the T_m is from T_{ref} , resulting in a systematic trend in RTD. As many phenomena of interest to the experimentalist systematically shift the T_m , care must be taken when fitting across numerous experiments weighted by their uncertainties, as it is possible to accidentally introduce systematic errors.

The T_m variation example system in Fig. 5 B and Fig. 6 A also serves to illustrate the uncertainty propagation for the measurement type shown in Fig. 1, by extracting the excess ΔH° between the two most disparate curves in Fig. 5 B. The temperature dependent uncertainties, Eq.(9), for the reference and reference+motif are added in quadrature and plotted in Fig. 6 C. This shows that the temperature-dependent uncertainty of the motif is larger than for either constituent measurement and has a shallower minimum between the data-centers-of-mass for each constituent measurement. This is reported in real units, as relative uncertainties cannot be added if they are relative to different quantities.



Fig. 6. A - Uncertainty on $\Delta S(T)$ for three of the T_m variation examples in Fig. 5 B, where the upper line bounding a shaded region represents the typical experiment uncertainty, RTD, and the lower line represents the systematic contributions, RSyD. B -Data from Fig. 5 B, evaluated at T_{ref} = 37 °C, showing the projection of uncertainty in ΔC_p on to ΔS° and ΔH° for RTD, upper curves and RSyD, lower curves. As this plot shows both ΔS° and ΔH° we did not shade the uncertainty envelope. C - Uncertainty (TD) in real units for the excess ΔH° calculated from the ΔH° = -300 kJ·mol⁻¹ and -275 kJ·mol⁻¹ melt curves as the 'reference+motif' and 'reference', respectively. Note that the extracted excess ΔH° is 25.1 kJ·mol⁻¹, with uncertainty of $(\pm 31, \pm 18, \text{ and } \pm 11) \text{ kJ} \cdot \text{mol}^{-1}$ at 4 °C, 25 °C, and 37 °C respectively.

These differences are useful to consider when interpreting studies which neglect ΔC_p . The quality of fit, and exact value of ΔH° and ΔS° , as evaluated at $T_{ref} = T_m$ only change subtly when ΔC_p is not included, shown in SI section 6. As demonstrated in the preceding discussion, ignoring the temperature dependent uncertainty on the fitted parameters introduces systematic errors when comparing the energetics of systems with different T_m .

For the experimentalist, we highlight the following as key features of van't Hoff analysis: first, that sampling too much, or too little, of the melt curve can significantly reduce accuracy, and that the optimum sampling will be system dependent; second, melt curve width and accessible concentration variation are important and, where possible, reference strands with smaller energetics should be used; third, the temperature dependence of fitting uncertainty should be considered.

While it may seem counterintuitive that a larger energetic change is harder to measure than a smaller one, it is critical to emphasize that *in van't Hoff analysis one is not measuring the thermodynamics but measuring the melt curve and extracting the thermodynamics*.

Individual melt curves, instrumental uncertainties: Before comparing full simulations in which numerous replicates are measured and analyzed, it is informative to observe the basic behavior of the van't Hoff analysis of an individual melt curve as a function of the different sources of uncertainty. To do so, we separately varied readout noise, temperature imprecision, and temperature accuracy using the default energetics and a bimolecular reaction, with a 1:1 ratio between the ssDNA strands and at a concentration of 40 μ mol·L⁻¹.

Examples of the noise applied to the melt curve, and its effect on the extracted thermodynamic parameters in real units, and their uncertainties in relative terms are shown in Fig. 7. Each data point in the middle rows of plots in Fig. 7 represent a single Monte Carlo simulation of van't Hoff analysis. In the bottom two rows, the RTD and RSyD are plotted with and without data markers respectively.

While the red melt curves at the top of Fig. 7 may not appear excessively noisy, the corresponding ΔH° , ΔS° , and ΔC_p plots reveal significant variation. As the temperature imprecision determines both the minimum temperature increment and data density along the x-axis, its effect on analysis precision is particularly pernicious, damaging it twofold.

In short, a melt curve which "looks good" may still be of insufficient quality for accurate parameter extraction.

A second feature of note is the correspondence between the cartoon in Fig. 3 and the middle rows of Fig. 7. These simulations replicate the appearance of entropy/enthalpy compensation(50, 51) solely through the effect of experimental uncertainty on the fit in $\ln([K])$ vs 1/T. Similarly, there is a temperature dependent compensation between either ΔH° or ΔS° and ΔC_p , as shown in the SI section 14, when $T_{ref} = T_m$, the distribution in ΔC_p appears more random. This makes sense as the variation in ΔC_p propagates onto ΔH° as a function of $(T_{ref} - T_m)$.

Generally, the uncertainties in the bottom two rows match our expectations for the propagation of equipment uncertainties through the analysis. It is important to note that for ΔH° and ΔS° the RTD is orders of magnitude higher than our benchmark uncertainty – this is unsurprising, as it is unlikely that a single replicate of an experiment would provide sufficiently precise thermodynamic information.



Fig. 7. Single curve van't Hoff analysis. Top – Effect on melt curve of readout noise (left), temperature noise (center), and temperature accuracy (right). The black curve is the ground truth, while the red points represent the simulated measurements. Middle – scatter plots of the extracted ΔH° and ΔS° (upper row) and the ΔH° and ΔC_{p} (lower row) for individual simulations, where black, blue, and red indicate varying readout noise (0.0125 %, 1 %, 2.5 %), temperature imprecision (0.025 °C, 0.5 °C, 1 °C), and temperature accuracy (0.01 °C, 1 °C, 5 °C), respectively. Bottom – uncertainty where the top curve represents a typical single curve measurement, RTD, and the lower represents the limit imposed by systematic effects, RSyD, for ΔH° and ΔS° (upper row) and for ΔC_{p} (lower row). All simulation inputs not being explicitly varied in the plot were set to default conditions, including $T_{ref} = 37 °C$.

Full experiments, choices in experiment and analysis: In actual experiments, van't Hoff analysis is generally applied to many replicate measurements of a system of interest. As discussed briefly above, the two most common experimental setups and analytical approaches for van't Hoff analysis are to vary the concentration and only use the T_m , or to average melt curves from many replicates of a single concentration and use some fraction of that averaged curve. Here we present a third, hybrid approach which combines the multiple concentrations of the first with the more complete data utilization of the second.

These approaches will be referred to as the "Avg. Replicate, Full Curve", "Varied Conc., T_m Only", and "Varied Conc., Full Curve", respectively.

Examples of both the melt curves and van't Hoff plots of these three approaches are presented in the first two rows of Fig. 8. As is immediately apparent, the hybrid analysis has two advantages. First, the increase in data density makes identification and quantification of curvature in the van't Hoff plot, i.e., ΔC_p , far easier. Second, it provides redundancy in $\ln([K])$ vs 1/T space where data from multiple curves overlap. Errors in fluorescence, ultraviolet-visible absorbance, or circular dichroism baseline correction may be revealed as mismatches between the ends of overlapping curve segments.

The bottom third of Fig. 8 shows the $\text{RTD}_{\Delta \text{H}^\circ}$ for these three styles of experiment as a function of the number of samples measured in a single experiment, and as a function of the fractional data range. In the former, the concentration for each sample in an experiment was set such that each experiment sampled the same range of ΔT_m (10 °C), such that individual sample T_m values were evenly spaced across the ΔT_m range.

Fig. 8. A – melt curve and van't Hoff plots for three different types of van't Hoff exper'iment: 'Average of replicates, full curve', 'Varied concentrations, T_m only', and 'Varied concentrations, full curve'. B – uncertainty on ΔH° as a function of the number of samples measured for each type with the same instrumental uncertainties (left), and as a function of fractional data range for the two relevant types (right). The upper line bounding a shaded region represents the uncertainty on a typical experiment, RTD, while the lower line represents the systematic contributions, RSyD. The small arrows on the y-axis indicate the benchmark acceptable uncertainty. All simulations were run with inputs set to default conditions.

It is unsurprising that Fig. 8 would indicate that precision in the "Varied Conc., T_m Only" is most sensitive to the number of samples. It uses the fewest data points of all three approaches, so it is furthest from the point of diminishing returns in data density. Similarly, the precision of the "Avg. Replicate, Full Curve" has a weak dependence on the number of samples, *N*, as averaging will only reduce the uncertainty on the noise in the melt curve by \sqrt{N} . From Fig. 7, one would anticipate averaging curves to result in only moderate improvements in accuracy compared to measuring additional concentrations as the latter will increase the range of the measured data in 1/T. Finally, as one might anticipate from data density in the top plots of Fig. 8, *the "Varied Conc., Full Curve" achieves the best precision for the least experimental effort*.

For the two "Full Curve" methods, the behavior of RTD as a function of the fractional data range is not dissimilar to that presented in Fig. 4. The precision for the "Varied Conc., Full Curve" approach is less sensitive than that of the "Avg. Replicate, Full Curve", which is unsurprising as overlap between melt-curves in $\ln([K])$ and 1/T can mitigate the effect of the asymmetric propagation of noise in F_{ssDNA} on to $\ln([K])$.

Full experiments, instrumental uncertainties: Fig. 9 and Fig. 10 show the uncertainty as a function of equipment uncertainties identical to those used in Fig. 7 for the three experimental approaches shown in Fig. 8. As the $\text{RTD}_{\Delta H^{\circ}}$ and $\text{RTD}_{\Delta S^{\circ}}$ are nearly identical, we choose to present only $\text{RTD}_{\Delta H^{\circ}}$ and $\text{RTD}_{\Delta Cp}$ to reduce visual clutter. As in the initial

assessment of analysis approaches, the "Varied Conc., Full Curve" method is approximately $30 \times to 40 \times more$ precise than either of the other two.

One reason for this lower uncertainty is that the higher data density allows for a much better fit of the curvature of the van't Hoff plot, thus reducing errors in ΔC_p , Fig. 10 A, which, in turn reduces uncertainties when extrapolating other parameters to a given reference temperature, e.g., physiological temperature, as illustrated in Fig. 9. The various levels in ΔC_p precision for the different experimental approaches, shown Fig. 10, are consistent with observations of ITC sensitivity(35), in that the analysis is also sensitive to the temperature range over which a transition is measured.

The effects of averaging multiple melt curves on $\text{RTD}_{\Delta \text{H}^\circ}$, detailed in section 4 of the SI, shown in Fig. 9 are consistent with the expected $1/\sqrt{N}$ scale reduction in noise associated with averaging *N* samples in a full-curve analysis for a single melt curve like the one shown in Fig. 7.

Fig. 9. Comparison of the $\text{RTD}_{\Delta H^\circ}$ evaluated at physiological temperature for different types of van't Hoff analysis as a function of read noise, temperature imprecision, and temperature accuracy under default conditions. The top line of each shaded region represents the uncertainty on a typical experiment, RTD, while the lower line of a pair represents the systematic contributions, RSyD. The small arrows on the Y axis indicate the benchmark uncertainty.

The "Varied Conc., Full Curve" approach performs better than the benchmark for all the equipment uncertainty conditions. For a relatively efficient FRET pair and typical real time PCR equipment, read noise and temperature imprecision of < 0.1 % can be attained.

Fig. 10. Comparison of $\text{RTD}_{\Delta Cp}$ between van't Hoff approaches as a function of read noise, temperature imprecision, and temperature accuracy under default conditions. The upper line bounding a shaded region

represents the uncertainty on a typical experiment, RTD, while the lower line of a pair represents the systematic contributions, RSyD and thus the smallest possible uncertainty.

The results presented here are consistent with the observation that measurements with more data points, and data points over a broader range in a fitting space, i.e., the hybrid approach, will be less sensitive to sources of uncertainty.

Full experiments, numerical derivatives and T_m: It is common in the literature, particularly with experimental data gathered using intercalating dyes, to use a numerical derivative $(\Delta F_{ssDNA}/\Delta T)$ of the raw melt curve, typically without baseline correction but with some level of smoothing, to determine T_m , as shown in Fig. 11 A. Here, Δ denotes the change in the fraction of ssDNA over a corresponding temperature interval to obtain a numerical derivative, rather than as the change between two states for thermodynamic properties.

This choice has serious limitations and significant drawbacks. The least of which is the validity of the claim that taking the numerical derivative of a melt curve 'removes' the temperature dependent baseline of its fluorescence reporter – this is only true if the baseline is perfectly linear and its contribution is added to the signal, rather being multiplied by the sigmoid. If the baseline is multiplicative to the signal, by the chain rule of differentiation, the T_m will be offset in proportion to the magnitude of the baseline temperature dependence. For many applications this may be a small difference, but the decision to ignore it should be made only after careful consideration. This is discussed further in SI Section 13.

A more serious flaw of this approach, and a subject of previous work(28), is in the definition of T_m . For thermodynamic purposes, the T_m is defined as $F_{ssDNA} = 0.5$, or the halfway point of the reaction. However, the peak of $\Delta F_{ssDNA}/\Delta T$ only gives $F_{ssDNA} = 0.5$ if the melt curve is odd-symmetric about the T_m , which is generally not the case, especially for bimolecular reactions.

For the more mathematically inclined, the melt curve is a cumulative distribution function (CDF) of the density of hybridized versus melted states, while $\Delta F_{ssDNA}/\Delta T$ is the probability density function (PDF). The peak of $\Delta F_{ssDNA}/\Delta T$ gives the mode of this CDF, while the T_m is defined as the median. The median and mode are *only* the same if the CDF is odd symmetric. For a rigorous consideration of this problem, albeit neglecting the complication of a non-zero ΔC_p , we direct the reader work by Owczarzy(28).

Finally, and most importantly, *numerical derivatives inherently increase noise*. While smoothing algorithms are often used, this merely sweeps the uncertainty propagation out of sight. *In short, using numerical derivatives to find the* T_m *results in an estimated value that has higher uncertainty and is inherently inaccurate.*

A preferable approach would be to fit a linear or quadratic function to a limited number of points about the T_m and interpolate that fitted function to $F_{ssDNA} = 0.5$. While this approach is susceptible to its own artifacts, especially if too many data points are fitted, it is more physically representative of the T_m and less likely to amplify noise in the data.

To show this directly, we plot $\Delta F_{ssDNA}/\Delta T$ and T_m by both interpolation and differentiation for several different melt curves in Fig. 11 A. These simulations were run under default instrumental uncertainties. The Savistky-Golay smoothing for $\Delta F_{ssDNA}/\Delta T$, and the number of data point used to interpolate across the linear fit, to $F_{ssDNA} = 0.5$, were both set at 11 to ensure a fair comparison where both approaches use the same amount of data. Plotting the known T_m against the extracted value, Fig. 11 B, clearly shows that the $\Delta F_{ssDNA}/\Delta T$ technique is both less accurate than linear interpolation and has a large systematic inaccuracy. It

should be noted that a less-than-perfect signal baseline correction would also degrade the identification of T_m by interpolation.

Fig. 11. A - numerical derivative of F_{ssDNA} with Savitsky-Golay smoothing. B - Extracted T_m versus true T_m value for both methods. Uncertainties lie within the symbol for the y-value interpolation. C uncertainty on ΔH^0 for the "Varied Concentration T_m Only" approach for both methods of T_m identification. The upper bound of each shaded region represents the uncertainty on a typical experiment, RTD, while the lower bound represents the systematic contributions, RSyD.

Fig. 11 C shows how the use of numerical derivatives affects the RTD_{Δ H^o} from the "Varied Conc., T_m Only" approach. Unsurprisingly, the use of a numerical derivative, $\Delta F_{ssDNA}/\Delta T$, results in an extraction of Δ H^o that is at 100 × to 500 × less precise than interpolation.

Quality of van't Hoff analysis in real units: So far, we have used relative uncertainty (RD) to minimize visual clutter. Here, we present some examples in real units to provide concrete examples and further illustrate the value of the simulation approach in guiding experimental design.

In previous experimental work(38), we observed an ΔH° of -229.5 ± 3.2 kJ·mol⁻¹ (54.9 ± 0.76 kcal·mol⁻¹) for a unimolecular looping event. A set of simulations using the experimental mean value of ΔH° as the input, discussed in detail in SI Section 9, yielded -231.0 ± 0.98 kJ·mol⁻¹ (55.2 ± 0.23 kcal·mol⁻¹). It is encouraging that these Monte Carlo simulations return comparable uncertainties.

A typical Monte Carlo simulation takes minutes to run, enabling an experimentalist to rapidly explore how modifications to an experimental design are likely to improve or degrade its results. For example, our experiments above would have realized a three-fold improvement in precision and reduced the systematic inaccuracy by half had they been performed with a 0.1 °C temperature increment rather than 0.61 °C. In the case of readout noise, the benefit of reducing the noise is less significant. Our simulations (Fig. 7) show a relatively weak, close-to-linear dependence of uncertainty on read noise, in good agreement with the relative error predicted analytically by Xi *et al.*(53). As detailed in SI section 12, our test case for use of more precise and accurate equipment yields thermodynamic parameters approximately $2 \times$ more precise. Insights of this nature can be valuable in designing experiments to achieve a desired precision while ensuring the best return on equipment time, person-hours, and reagents.

The benefits of using better equipment can be dwarfed by other experimental design choices. For example, the thermodynamic parameter extraction uncertainties for a random 8 base oligomer as compared to those for a random 30 base oligomer are 300 to $600 \times$ smaller, as detailed in SI section 10.

The same care and attention required to produce high-quality data from melt-curve experiments should also be exercised in choosing to do a melt-curve experiment in the first place. Motif energetics may be determined most effectively using toe-hold-mediated catalysis, melt-curve methods, or calorimetry, depending on their magnitude and the presence of confounding factors.

Conclusions

As we have shown, van't Hoff analysis of melt-curve data is much more sensitive to experimental design and the thermodynamics of the measured system than it is to equipment uncertainties. However, for any given combination of equipment and experimental design, the uncertainty in the analysis comes predominantly from random, equipment-related effects, rather than systematic effects. We caution the reader that this conclusion does not reflect the effects of model-form errors, which will cause large systematic inaccuracies. By avoiding the common pitfalls and applying the experimental design principles that we have described here, melt-curve experiments can deliver thermodynamic information that can be up to four orders of magnitude more precise¹ for the same level of experimental effort.

¹ We observed improvements in precision of $100 \times to 500 \times$ when using interpolation versus a numerical derivative, $300 \times to 600 \times by$ using an 8 base oligomer versus a 30 base oligomer reference sequence, $30 \times to 40 \times by$ using the "Varied concentrations, full curve" approach versus either the "Varied Conc., T_m Only" or "Avg. Replicate, Full Curve", approaches, $5 \times by$ using more concentrations than replicates, and $3 \times by$ increasing data density in T (see SI sections 9, 10, & 12). The effects of these choices will vary by system, but

Further improvements in thermodynamic predictions would be enabled by reliable models for per-base ΔC_p . This would necessarily include accounting for systematic uncertainties that arise because variations in GC content cause variations in T_m , which, as we have discussed, change the propagation of uncertainty in ΔC_p as a function of $(T_m - T_{ref})$.

We reiterate that van't Hoff analysis does not directly measure thermodynamics, rather the melt curves are measured, and thermodynamics are extracted. In addition, as noted above, the uncertainty of van't Hoff analysis degrades the further T_m diverges from T_{ref} . As an overarching principle, one should maximize the amount of $\ln([K])$ vs. 1/T space one samples, while minimizing unnecessary propagation of noise, e.g., from the melt curve plateaus. This principle leads to the following best practices:

- Avoid using numerical derivatives.
- Use the smallest feasible reference sequence.
- Focus on acquiring data at more concentrations versus more replicates.
- Maximize the data density across the melt transition.
- Do not neglect ΔC_p without careful consideration.

We again note that the fitting landscape in terms of ΔH° , ΔS° , and ΔC_p is shallow, i.e., there are many slightly different sets of thermodynamic values that yield similar fit quality. In combination with the entropy-enthalpy compensation artifact, this shallow fitting landscape means that results can vary with fitting algorithm settings and numerical precision. We plan to investigate these factors in future work.

Finally, we emphasize that the best practices discussed here may be usefully applied to ensemble data of many kinds of reversible biomolecular interactions, e.g., protein binding.

Additional Information:

Supplementary Information is available for this paper. Correspondence and requests for materials should be addressed to Jacob Majikes or J. Alexander Liddle.

Declaration of Interests:

The authors declare no competing interests.

Author Contributions:

JM and AL designed the research. JM carried out all Monte Carlo simulations. MZ checked the simulations and helped select reporting notation that would directly convey our results in a compact, visual manner. JM and AL wrote the article.

References

- 1. Wei, G., W. Xi, R. Nussinov, and B. Ma. 2016. Protein Ensembles: How Does Nature Harness Thermodynamic Fluctuations for Life? The Diverse Functional Roles of Conformational Ensembles in the Cell. *Chem. Rev.* 116:6516–6551.
- 2. Piepenbrink, K.H., B.E. Gloor, K.M. Armstrong, and B.M. Baker. 2009. Methods for quantifying T cell receptor binding affinities and thermodynamics. 1st ed. Elsevier Inc.
- 3. Frauenfelder, H., S.G. Sligar, and P.G. Wolynes. 1991. The Energy Landscapes and

the best practice we suggest could yield several orders of magnitude more precision than a concatenation of commonly-observed sub-optimal choices.

Motions of Proteins. Science (80-.). 254.

- 4. Holdgate, G.A., and W.H.J. Ward. 2005. Measurements of binding thermodynamics in drug discovery. *Drug Discov. Today.* 10:1543–1550.
- 5. Suryawanshi, V.D., L.S. Walekar, A.H. Gore, P. V. Anbhule, and G.B. Kolekar. 2016. Spectroscopic analysis on the binding interaction of biologically active pyrimidine derivative with bovine serum albumin. *J. Pharm. Anal.* 6:56–63.
- 6. Ouldridge, T.E. 2018. The importance of thermodynamics for molecular systems, and the importance of molecular systems for thermodynamics. *Nat. Comput.* 17:3–29.
- 7. Winzor, D.J., and C.M. Jackson. 2006. Interpretation of the temperature dependence of equilibrium and rate constants. *J. Mol. Recognit.* 19:389–407.
- 8. Majikes, J.M., P.N. Patrone, A.J. Kearsley, M. Zwolak, and J.A. Liddle. 2021. Failure mechanisms in DNA self-assembly: Barriers to single-fold yield. *ACS Nano*. 15:3284–3294.
- 9. Haq, I., B.Z. Chowdhry, and J.B. Chaires. 1997. Singular value decomposition of 3-d DNA melting curves reveals complexity in the melting process. *Eur. Biophys. J.* 26:419–426.
- 10. Schroeder, S.J., and D.H. Turner. 2009. Optical Melting Measurements of Nucleic Acid Thermodynamics. In: Methods in Enzymology. Elsevier Inc. pp. 371–387.
- Marky, L.A., and K.J. Breslauer. 1987. Calculating Thermodynamic Data for Transitions of any Molecularity from Equilibrium Melting Curves. *Biopolymers*. 26:1601–1620.
- 12. You, Y., A. V Tataurov, and R. Owczarzy. 2011. Measuring Thermodynamic Details of DNA Hybridization Using Fluorescence. *Biopolymers*. 95:472–486.
- 13. Howard, K.P. 2000. Thermodynamics of DNA Duplex Formation: A Biophysical Chemistry Laboratory Experiment. *Lab.* 77:77–79.
- 14. Atkins, P., J. De Paula, and J. Keeler. 2017. Atkin's Physical Chemistry, 11th Ed. .
- 15. Reed, G.H., J.O. Kent, and C.T. Wittwer. 2007. High-resolution DNA melting analysis for simple and efficient molecular diagnostics. *Pharmacogenomics*. 8:597–608.
- Böhme, K., P. Calo-Mata, J. Barros-Velázquez, and I. Ortea. 2019. Review of Recent DNA-Based Methods for Main Food-Authentication Topics. J. Agric. Food Chem. 67:3854–3864.
- 17. Ramezani, H., and H. Dietz. 2019. Building machines with DNA molecules. *Nat. Rev. Genet.* 12.
- 18. Linko, V. 2020. At the dawn of applied DNA nanotechnology. *Molecules*. 25:2–5.
- 19. Deluca, M., Z. Shi, C.E. Castro, and G. Arya. 2020. Dynamic DNA nanotechnology: Toward functional nanoscale devices. *Nanoscale Horizons*. 5:182–201.
- Zuker, M., D.H. Mathews, and D.H. Turner. 1999. Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide. In: Barciszewski J, BFC Clark, editors. RNA Biochemistry and Biotechnology. Dordrecht: Springer Netherlands. pp. 11–43.
- 21. Markham, N.R., and M. Zuker. 2008. UNAFold. In: Keith JM, editor. Bioinformatics: Structure, Function and Applications. Totowa, NJ: Humana Press. pp. 3–31.
- 22. Markham, N.R., and M. Zuker. 2005. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res.* 33:W577–W581.
- 23. Ding, Y., and C.E. Lawrence. 2003. A statistical sampling algorithm for RNA secondary structure prediction. *Nucleic Acids Res.* 31:7280–7301.
- Owczarzy, R., A. V. Tataurov, Y. Wu, J.A. Manthey, K.A. McQuisten, H.G. Almabrazi, K.F. Pedersen, Y. Lin, J. Garretson, N.O. McEntaggart, C.A. Sailor, R.B. Dawson, and A.S. Peek. 2008. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res.* 36.

- 25. Kibbe, W.A. 2007. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.* 35:W43-6.
- 26. Tulpan, D., M. Andronescu, and S. Leger. 2010. Free energy estimation of short DNA duplex hybridizations. *BMC Bioinformatics*. 11.
- Chavali, S., A. Mahajan, R. Tabassum, S. Maiti, and D. Bharadwaj. 2005. Oligonucleotide properties determination and primer designing: A critical examination of predictions. *Bioinformatics*. 21:3918–3925.
- 28. Owczarzy, R. 2005. Melting temperatures of nucleic acids: Discrepancies in analysis. *Biophys. Chem.* 117:207–215.
- 29. Vieregg, J.R. 2010. Nucleic Acid Structural Energetics. Encycl. Anal. Chem.
- 30. Tellinghuisen, J. 2004. Statistical Error in Isothermal Titration Calorimetry. *Methods Enzymol.* 383.
- 31. Tellinghuisen, J. 2006. Van't Hoff analysis of K° (T): How good...or bad? *Biophys. Chem.* 120:114–120.
- 32. Horn, J.R., J.F. Brandts, and K.P. Murphy. 2002. van't Hoff and Calorimetric Enthalpies II: Effects of Linked Equilibria. *Biochemistry*. 41:7501–7507.
- Horn, J.R., D. Russell, E.A. Lewis, and K.P. Murphy. 2001. van't Hoff and Calorimetric Enthalpies from Isothermal Titration Calorimetry: Are There Significant Discrepancies? †. *Biochemistry*. 40:1774–1778.
- 34. Chaires, J.B. 1997. Possible origin of differences between van't Hoff and calorimetric enthalpy estimates. *Biophys. Chem.* 64:15–23.
- 35. Zhukov, A., and R. Karlsson. 2007. Statistical aspects of van't Hoff analysis : a simulation study. *J. Mol. Recognit.* 20:379–385.
- Dannenberg, F., K.E. Dunn, J. Bath, A.J. Turberfield, T.E. Ouldridge, M. Kwiatkowska, A.J. Turberfield, and T.E. Ouldridge. 2015. Modelling DNA origami self-assembly at the domain level. *J. Chem. Phys.* 143:1–19.
- 37. Arbona, J.-M., J.-P. Aimé, and J. Elezgaray. 2013. Cooperativity in the annealing of DNA origamis. *J. Chem. Phys.* 138:015105.
- Majikes, J.M., P.N. Patrone, D. Schiffels, M. Zwolak, J. Kearsley, S.P. Forry, and J.A. Liddle. 2020. Revealing thermodynamics of DNA origami folding via affine transformations. *Nucleic Acids Res.* 1–13.
- 39. Wang, C., J.H. Bae, and D.Y. Zhang. 2016. Native characterization of nucleic acid motif thermodynamics via non-covalent catalysis. *Nat. Commun.* 7:1–11.
- 40. Bae, J.H., J.Z. Fang, and D.Y. Zhang. 2020. High-throughput methods for measuring DNA thermodynamics. *Nucleic Acids Res.* 48:E89.
- 41. Moreira, B.G., Y. You, and R. Owczarzy. 2015. Cy3 and Cy5 dyes attached to oligonucleotide terminus stabilize DNA duplexes: Predictive thermodynamic model. *Biophys. Chem.* 198:36–44.
- 42. Di Michele, L., B.M. Mognetti, T. Yanagishima, P. Varilly, Z. Ruff, D. Frenkel, and E. Eiser. 2014. Effect of inert tails on the thermodynamics of DNA hybridization. *J. Am. Chem. Soc.* 136:6538–6541.
- 43. Ohmichi, T., S. ichi Nakano, D. Miyoshi, and N. Sugimoto. 2002. Long RNA dangling end has large energetic contribution to duplex stability. *J. Am. Chem. Soc.* 124:10367–10372.
- 44. Ohmichi, T., S. Nakano, D. Miyoshi, and N. Sugimoto. 2002. Long RNA Dangling End Has Large Energetic Contribution to Duplex Stability. *J. Am. Chem. Soc.* 124:10367–10372.
- 45. Di Michele, L., B.M. Mognetti, T. Yanagishima, P. Varilly, Z. Ruff, D. Frenkel, and E. Eiser. 2014. Effect of Inert Tails on the Thermodynamics of DNA Hybridization. *J. Am. Chem. Soc.* 136:6538–6541.

- 46. Patrone, P., A. Kearsley, J. Majikes, and J.A. Liddle. 2020. Analysis and Uncertainty Quantification of DNA Fluorescence Melt Data: Applications of Affine Transformations. *Anal. Biochem.*
- 47. Patrone, P.N., E.L. Romsos, M.H. Cleveland, P.M. Vallone, and A.J. Kearsley. 2020. Affine analysis for quantitative PCR measurements. *Anal. Bioanal. Chem.* 412:7977–7988.
- 48. Sanford, L.N., J.O. Kent, and C.T. Wittwer. 2013. Quantum Method for Fluorescence Background Removal in DNA Melting Analysis. *Anal. Chem.* 9907–9915.
- Majikes, J.M., P.N. Patrone, A.J. Kearsley, M. Zwolak, and J.A. Liddle. 2021. Failure Mechanisms in DNA Self-Assembly: Barriers to Single-Fold Yield. ACS Nano. 15:3284–3294.
- 50. Sharp, K. 2001. Entropy--enthalpy compensation: Fact or artifact? *Protein Sci.* 10:661–667.
- 51. Krug, R.R., W.G. Hunter, and R.A. Grieger. 1976. Enthalpy-Entropy Compensation. 1. Some Fundamental Statistical Problems Associated with the Analysis of van't Hoff and Arrhenius Data. *J. Phys. Chem.* 80:2335–2341.
- 52. Palais, R., and C.T. Wittwer. 2009. Chapter 13 Mathematical Algorithms for High-Resolution DNA Melting Analysis. 1st ed. Elesvier Inc.
- 53. Xia, T., J. SantaLucia, M.E. Burkard, R. Kierzek, S.J. Schroeder, X. Jiao, C. Cox, and D.H. Turner. 1998. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson - Crick base pairs. *Biochemistry*. 37:14719–14735.