Using transient equilibria (TREQ) to measure the thermodynamics of slowly assembling supramolecular systems

Christopher D. Hennecker,[1] Christophe Lachance-Brais,[1] Hanadi Sleiman,[1] and Anthony Mittermaier[1,*]

[1]Department of Chemistry, McGill University. 801 Sherbrooke St. W. Montreal, Canada H3A 0B8

[*]To whom correspondence should be addressed: anthony.mittermaier@mcgill.ca

Supramolecular chemistry involves the non-covalent assembly of monomers into materials with unique properties and wide-ranging applications. Thermal analysis is a key analytical tool in this field, as it provides quantitative thermodynamic information on both the structural stability and nature of the underlying molecular interactions. However there exist many supramolecular systems whose kinetics are so slow that the thermodynamic methods currently applied are unreliable or fail completely. We have developed a simple and rapid spectroscopic method for extracting accurate thermodynamic parameters from these systems. It is based on repeatedly raising and lowering the temperature during assembly and identifying the points of transient equilibrium as they are passed on the up- and down-scans. In a proof-of-principle application to the co-assembly of polydeoxyadenosine containing 15 adenosines (polyA) and cyanuric acid (CA), we found that roughly 30% of the CA binding sites on the polyA chains were unoccupied, with implications for high-valence systems.

Introduction

Supramolecular chemistry is emerging as a rich source of diverse materials with novel and valuable properties. Potential applications range from drug-delivery and tissue regeneration to optical sensors and organic electronics.[1] This approach involves the non-covalent self-assembly of tens to thousands of monomeric units into larger structures with emergent physical properties that derive from both the structures of the individual components and their interactions and arrangement with respect to one another.[2] Reversible assembly has some distinct advantages compared to traditional covalent synthesis. The dynamic nature of supramolecular interactions allows bonds to break and reform leading to materials with self-healing properties. Furthermore, many supramolecular systems have the ability to generate multiple morphologies and sets of physical properties from a single set of building blocks with only small modifications of the assembly conditions.[3] Nevertheless, there are unique challenges associated with this approach. Chief among these is characterizing the products of a non-covalent assembly reaction. Much of the excitement surrounding supramolecular chemistry comes from the fact that there remains much to be understood regarding the relationships between the chemical structures of the monomeric units, the supramolecular architectures, and the emerging physical properties, and there is wide possibility for new and unexpected discoveries. However, this implies that the nature of supramolecular products is difficult to predict and that rigorous structural and thermodynamic analyses are critical to advancing the field.

A variety of tools have been used to elucidate the structures produced by assembly, including atomic force, electron, and super-resolution microscopies, and solid-state NMR spectroscopy.[4-6] The stabilities of the assemblies are most commonly measured by thermal analysis. Most supramolecular structures dissociate when they are heated and reassemble when the monomer mixtures are cooled. This process can be quantified either by calorimetry[7] or by spectroscopically-detected melting and annealing.[8, 9] Detailed analyses of melting curves yield the enthalpies, ΔH, entropies ΔS, and free energies, ΔG, of assembly and shed light on the forces holding the supramolecular structures together.[10] This information is essential for determining structure/function relationships and the rational design and improvement of self-assembling systems.[11, 12] However, there exists a large class of supramolecular systems with extremely slow kinetics that only assemble or disassemble at useful rates when they are pushed far from equilibrium, i.e. under very highly stabilizing or destabilizing conditions. Common examples include amyloid fibrils, viral capsids, and a variety of self-assembling non-biological small molecules.[11-27] Interest in these kinds of slowly assembling supramolecular systems has grown in recent years, since they allow the size distributions of the resulting fibres to be tightly controlled.[24, 26-28] Current thermodynamic analyses rely on systems reaching equilibrium before the measurement is taken. In principle, this precludes thermodynamic analyses of slowly assembling systems, since equilibrium is not reached on practical timescales. Nevertheless, it is common practice in the supramolecular field to interpret non-equilibrium thermal data using equations derived for equilibrium systems, despite warnings in the literature that this is invalid.[10] Our mathematical simulations (see below) indicate this can lead to errors in reported thermodynamic parameters of >100% and equilibrium constants that differ from their true values by orders of magnitude. Thus a lack of reliable thermodynamic information for slowly-assembling systems is an impediment to the advancement of the supramolecular chemistry field.

We have developed a new experimental approach that can be performed using a standard temperature-controlled spectrophotometer and exploits transient equilibria (TREQ) to provide rigorous thermodynamic data on slowly assembling systems. Rather than waiting for the system to equilibrate (which can take days or weeks), the temperature is repeatedly raised and lowered, driving cyclic, non-

equilibrium disassembly and assembly. We find that the system briefly passes through an instant of equilibrium on each up-scan and down-scan at which the rates of assembly and disassembly are equal. The temperatures and concentration values at which these moments of equilibrium occur can be clearly identified from the spectroscopic trace, allowing the full thermodynamic melting curve to be mapped in just a few hours.

For example, we applied TREQ experiments to better understand the recently discovered co-assembly of polydeoxyadenosine (polyA) and the small molecule cyanuric acid (CA) into fibres whose biocompatibility and low cost make them promising candidates for tissue engineering and drug delivery.[29] A cross-section of the proposed structure *(Figure 1)* shows the adenosine of three different DNA strands hydrogen bonding to CA molecules in a continuous supramolecular helicene.[30] We note that the ideal helicene structure has a 1:1 ratio of dA residues and CA molecules. We recently characterized the kinetics of polyA-CA fibre assembly using non-equilibrium melting methods.[17] Equilibration of the fibres near the melting point could take up to a month of constant instrument use. Using TREQ experiments, we determined the ΔG, ΔH, and ΔS values for adding a polyA chain to the end of a growing fibre in a single 10-hour experiment. By repeating these measurements at different concentrations of CA, we determined the minimum polyA:CA ratio necessary for assembly and made the surprising discovery that about 30% of the available CA binding sites are unfilled under our conditions. These results have implications for the future development of asymmetric systems involving components of very different valences, such as polyA and CA, and demonstrate the potential of the TREQ approach for learning about slowly assembling systems.

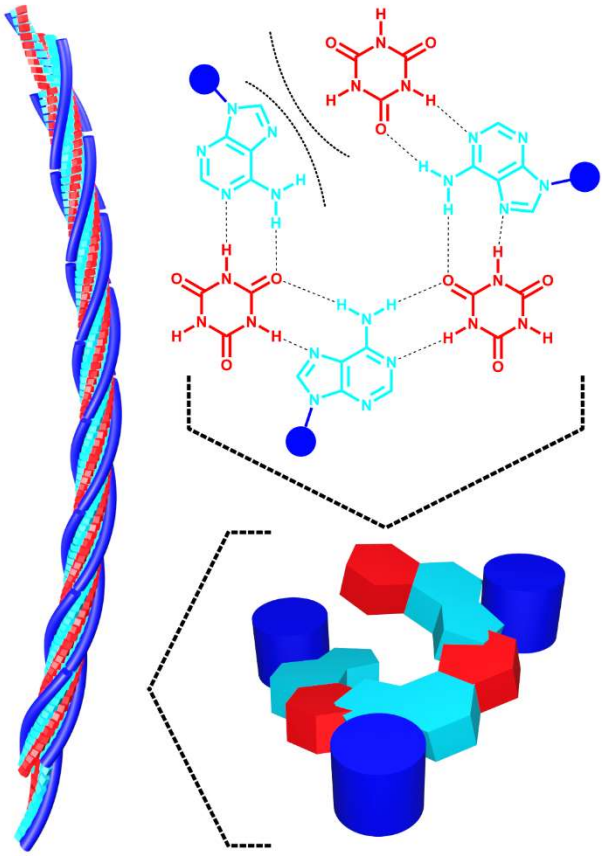

Figure 1: Proposed structure of supramolecular fibres formed from the co-assembly of poly-adenosine strands and cyanuric acid (left). A cross section of a single hexameric helicene (right).

Results and Discussion

Theory

Fibre assembly can be described by kinetic schemes such as the Goldstein-Stryer (GS) cooperative kinetic model:[17, 29, 31]

$$M_1 \underset{k_{n-}}{\overset{k_{n+}[M]}{\rightleftarrows}} M_2 \quad \cdots \quad M_{s-1} \underset{k_{n-}}{\overset{k_{n+}[M]}{\rightleftarrows}} \boxed{M_s} \underset{k_{e-}}{\overset{k_{e+}[M]}{\rightleftarrows}} M_{s+1} \quad \cdots \quad M_N \underset{k_{e-}}{\overset{k_{e+}[M]}{\rightleftarrows}} M_{N+1} \qquad (1)$$

nucleus

where M_N is a fibre containing *N* monomers. Association and dissociation of monomers from short oligomers less than the critical nucleus size, *s*, are described by the nucleation rate constants k_{n+} and k_{n-} respectively, while oligomers larger than *s* are described with the elongation rate constants k_{e+} and k_{e-}. An experimental parameter of great importance is the critical monomer concentration, $[M]_c$, at which the net rate of assembly or disassembly is zero at equilibrium. For rapidly-equilibrating systems, $[M]_c$ versus T curves can be measured directly by traditional melting or reannealing experiments and analyzed to obtain the enthalpies, entropies, and equilibrium dissociation constants for a monomer adding to the end of a fibre (ΔH_e, ΔS_e, and K_e, respectively) as well as the corresponding parameters for fibre nucleation.[32] For cooperative assembly, where nucleation is far less favourable than elongation, $[M]_c \approx K_e$ and a simplified analysis is commonly used; the maximum temperature at which fibres barely begin to form is identified as the elongation temperature, T_e, this temperature can be found by either fitting the elongation process or can approximated from the assembly curve,[33, 34] while $[M]_c$ is equated to the total monomer concentration, c_T. The experiment is repeated several times at different c_T values (Fig 2a), where increasing c_T leads to an increase in T_e. A van 't Hoff plot of $\ln(c_T)$ vs $1/T_e$ is then used to extract values of ΔH_e and ΔS_e.

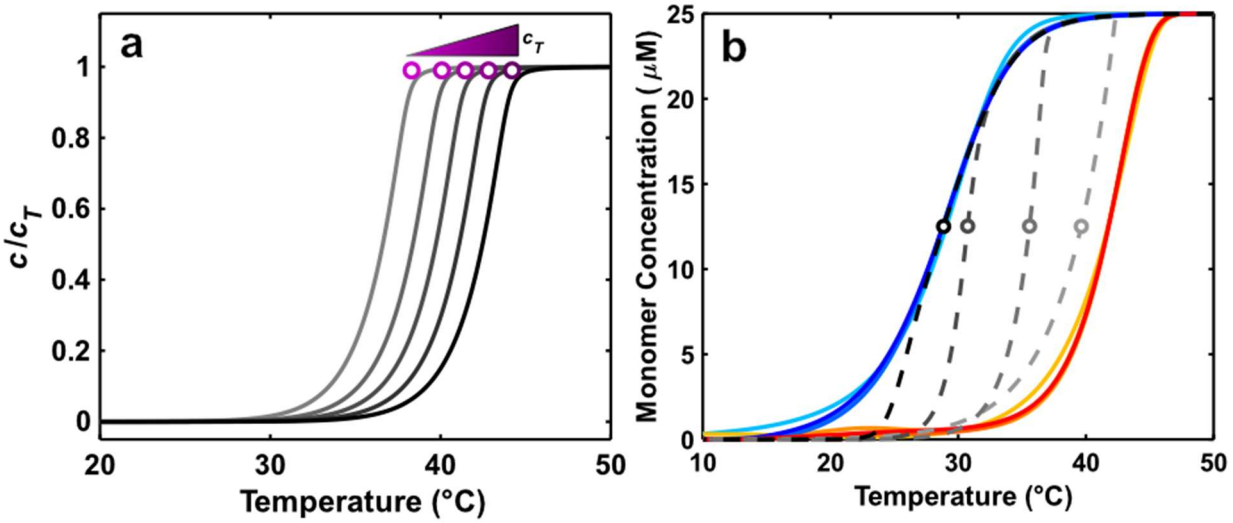

Figure 2: ***a)*** *Simulated assembly curves for different total concentrations of monomer (c_T), increasing concentrations are shown as a gradient from grey to black, T_e values are shown as points using a purple gradient.* ***b)*** *Fibre assembly/disassembly simulated using the Goldstein-Stryer model and kinetic parameters that give similar melting and annealing curves (solid lines) with drastically different equilibrium curves (dashed lines). Heating curves are shown in red/orange and cooling curves are shown in blue/cyan. The offset between heating and cooling data is due to thermal hysteresis (TH). Simulation parameters are listed in Supplementary Table 3.*

The situation is far more complicated for slowly assembling systems, such as polyA-CA fibres studied here. In these cases, the rate at which the system relaxes to equilibrium is far slower than available temperature scan rates, thus both folding (cooling) and unfolding (heating) occur out of equilibrium. The populations effectively lag behind the changing temperature such that the cooling and heating scans are offset, in a phenomenon known as thermal hysteresis (TH). Data for the up-scan lie to the right of the equilibrium $[M]_c$ vs T curve, and data for the down-scan lie to the left, as illustrated in *Figure 2b.* The resulting TH loops are rich in kinetic information, but are unsuitable for thermodynamic analyses, since the shape and location of the equilibrium curve is ill-defined, apart from the fact that it must lie somewhere between the heating and cooling scans[10, 17] To illustrate, fibres obeying the GS assembly model can have very different thermodynamic parameters and equilibrium curves and yet produce nearly superimposable thermal hysteresis data (*Figure 2b*).

Nevertheless, data for systems exhibiting pronounced thermal hysteresis have frequently been analyzed as if they were obtained at equilibrium. Heating curves are typically used together with the concentration-dependent T_e approach described above[12, 23-25], although sometimes cooling scans have been employed instead.[11, 20-22] Interestingly, in their seminal 2003 review, Mergny and Lacroix point out that "*analysis of the concentration dependency of the denaturation profile only is seriously flawed*" and urge "*great caution about conclusions reached solely by analysis of the heating curves, a recurrent theme in the literature*", when thermal hysteresis is present.[10] To gain a clearer picture of the magnitude of the problem, we simulated TH data using GS parameters matching our polyA-CA system at different values of c_T and analyzed the resulting concentration dependent T_e values. Using heating scans, the extracted value of ΔH_e was 2.6-fold too large, using cooling scans it was 2-fold too small, and K_e values were incorrect by two to seven orders of magnitude (*Supplementary Figure 1 and Supplementary Table 1*). In some studies[26, 27], different temperature scan rates produce superimposable heating data and it has been argued this validates their use in the concentration dependent T_e analysis. To test this hypothesis, we slightly modified our GS parameters to reproduce this effect and repeated the calculations. The resulting ΔH_e value was still about 1.8-fold too large (*Supplementary Figure 2 and Supplementary Table 2*). Thus, commonly used thermal melting and reannealing experiments do not provide reliable thermodynamic data for slowly assembling systems. Notably our TREQ method reproduces the thermodynamic parameters in these simulations with a high degree of accuracy (*Supplementary Figure 1/2 and Supplementary Table 1/2*).

Recent work from the Yamaguchi lab[19] has explored how the spectra of slowly equilibrating, self assembling systems respond to repeated heating and cooling cycles.[35] Depending on the starting and ending temperatures and ramp rates, a rich diversity of shapes (thermal hysteresis loops) have been observed, providing qualitative information on the underlying assembly reactions. However, to date there has not been a straightforward way to extract quantitative thermodynamic information from these data. Our new TREQ approach uniquely fills this gap. In order to illustrate the fundamental principles, we performed kinetic simulations using the GS assembly model and parameters for polyA-CA fibres *(Figure 2a, see Supplementary Information)*. The dashed black line indicates the equilibrium $[M]_c$ versus T curve,

while the simulated heating and cooling scans give the left- and right-shifted blue and red curves, respectively. Thus the location of the true $[M]_c$ equilibrium curve is obscured by the thermal hysteresis.

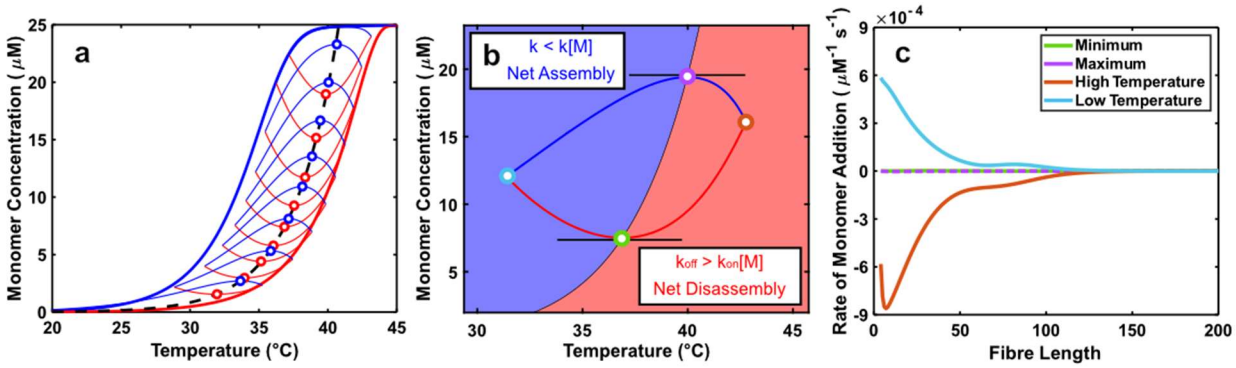

Figure 2: ***a)*** *Kinetic simulations of a typical hysteresis experiment (bold lines) and TREQ experiment (narrow lines). Cooling traces are shown in blue, heating traces are shown in red. The experiment begins by cooling from 45°C to 36°C, this is followed by the first up-scan (36°C to 44°C), a second down-scan (44°C to 35°C), a second up-scan (35°C to 43°C), a third down-scan (43°C to 34°C), etc. The equilibrium profile is shown as the dashed black line, with the extrema of each TREQ cycle shown as points.* ***b)*** *An isolated TREQ cycle: assembly occurs only in the blue shaded region; disassembly only occurs in the red shaded region. The interface of these two regions represents a system at equilibrium. Coloured points represent the position of calculated monomer flux in panel c.* ***c)*** *Calculated monomer flux of fibres for points shown in panel b, the horizontal extrema of the TREQ cycle have 100-fold less flux then either the high or low temperature values.*

The TREQ method is based on our discovery that repeatedly raising and lowering the temperature such that it repeatedly traverses the equilibrium curve reveals the precise locations of the hidden equilibria. Simulating TREQ data for polyA-CA assembly gives a series of concave-up and concave-down arcs on the heating and cooling scans, respectively (narrow red and blue curves in *Figure 2a*). Strikingly, the $[M]_c$ values (black line) pass directly through the extrema (concentration maxima and minima) of the cooling and heating arcs. Thus experimentally determined extrema can be interpreted as a set of $[M]_c(T)$ values. The physical process underlying this behaviour can be understood as follows: for cooperatively assembled fibres, such as polyA-CA, equilibrium is reached when the rate of monomer addition to the end of a fibre ($k_{e+}[M]_c$) is exactly equal to the rate of monomer dissociation from the end of a fibre (k_{e-}), such that the net rate of fibre growth is zero (thus $[M]_c \approx K_e$).[32] When $[M_1] < [M]_c$ there is net dissociation and $[M_1]$ increases with time, corresponding to the red region below the $[M]_c$ curve in *Figure 2b*. When $[M_1] > [M]_c$ there is net association and $[M_1]$ decreases with time, corresponding to the blue region above the $[M]_c$ curve. Every cooling scan starts in the red region with net dissociation (increasing $[M_1]$) and ends in the blue region with net association (decreasing $[M_1]$). As the temperature crosses the boundary where $[M_1]=[M]_c$, net fibre growth is zero, the arc is exactly horizontal, and the maximum is reached. Conversely, every heating scan starts in the blue region with decreasing $[M_1]$ and ends in the red region with increasing $[M_1]$. As the temperature crosses the $[M_1]=[M]_c$ boundary, the free monomer concentration is at a minimum. To validate this interpretation, we calculated the net rate of monomer addition to each length of fibre in the simulation. At the lower and upper limiting scan temperatures (orange and cyan), the rates of monomer addition and release are at least 100-fold greater than at the horizontal extrema of the heating and cooling arcs (green and purple) *(Figure 2c)*.

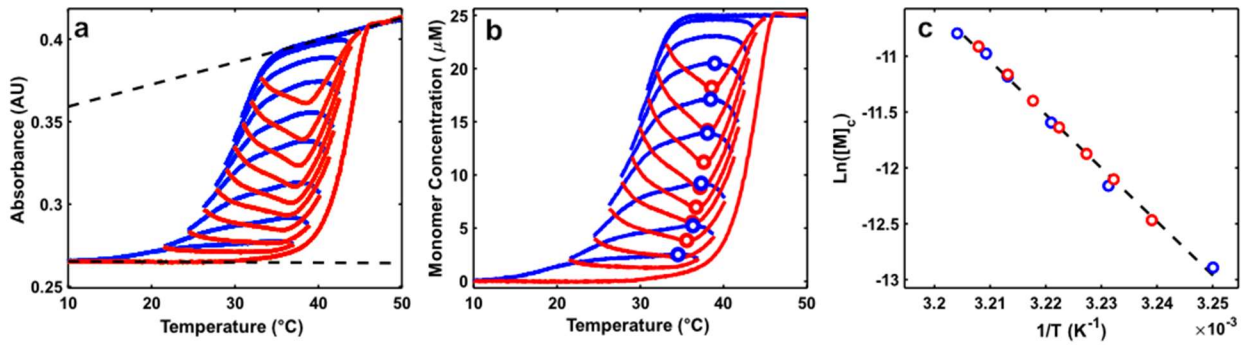

Figure 3: ***a)*** *Raw absorbance data for a 15mer polyA-CA coassembly with 25 µM polyA$_{15}$ and 15 mM CA at pH 4.5, blue lines represent cooling traces and red lines represent heating traces. Unfolded (top black line) and folded (bottom black line) are also shown.* ***b)*** *Processed TREQ data with extrema of each cycle shown as points.* ***c)*** *Van 't Hoff analysis of experimental TREQ points, plotted along with confidence interval of one standard deviation.*

Analysis of experimental TREQ data

We performed a TREQ experiment on a mixture of CA and polyA chains by repeatedly raising and lowering the temperature over a window of roughly 15°C that shifted from (29-44) to (14-33) °C in 9 cycles while we monitored the spectroscopic absorbance at 260 nm (*Figure 4a*, see *Supplementary Information* for a guide to selecting sliding T-windows). The lower and upper absorbance regions were fitted to linear baselines and assigned 100% and 0% folded, i.e. [M_1] = 0 and 25 µM, respectively. The converted data are shown in *Figure 4b*, with blue and red indicating cooling and heating, respectively, and open circles placed at the extrema. These experimental arcs have a remarkable similarity to the calculations shown in *Figure 3a*. The y- and x-values of the extrema correspond directly to critical monomer concentration, $[M]_c$, and temperature pairs. As discussed above, $[M]_c$ values are equivalent to the equilibrium dissociation constant, K_e, for adding a polyA to the end of an elongating fibre, for this system. A van 't Hoff plot of $\ln([M]_c)=\ln(K_e)$ vs 1/T is linear with a slope of $-\Delta H_e/R$ and y-intercept of $\Delta S_e/R$ *(Figure 3c)*, giving ΔH_e= 100 ± 2 kcal mol^{-1} and ΔS_e = 335 ± 7 cal mol^{-1} K^{-1}. Notably, although the values of ΔH_e and ΔS_e determined by TREQ differ from those obtained by kinetic fits to TH data by factors of 1.6 *(Supplementary Table 4)*, repeating the TH analysis with ΔH_e and ΔS_e fixed to the TREQ-derived values gives good agreement with experimental data *(Supplementary Figure 4)*, illustrating the insensitivity of the kinetic fits to these thermodynamic parameters. In general, we would strongly recommend that, even if assembly kinetics are the main interest, the combination of TREQ and thermal hysteresis experiments provide more robust solutions than thermal hysteresis alone, as TREQ resolves ambiguity in the fitted rate constants and ratios thereof.

Furthermore, the thermodynamic parameters provide a basis for comparing polyA-CA fibres to other nucleic acid structures. For example, polyA/polyT ($dA_{15}dT_{15}$) duplex dissociation is predicted to have approximately ΔH = 108 kcal mol^{-1} and ΔS = 335 cal mol^{-1} K^{-1} under similar solution conditions to those used here,[36] It is intriguingly similar to the values we measured for polyA-CA assembly (100 kcal mol^{-1} and 335 cal mol^{-1} K^{-1}). At first glance, we would have expected polyA-CA fibres to show much higher enthalpies and entropies than dAdT duplexes, since there are three strands rather than two, each dA forms twice as

many hydrogen bonds and immobilizes a CA molecule in the putative polyA-CA structure *(Figure 1).* However partial vacancy of CA binding sites may help to reconcile these observations, as elaborated below.

Stoichiometry of polyA-CA fibres.

One of the great advantages of quantitative thermodynamic data is that much can be learned about the system of interest through careful analyses of how energetic parameters vary with changing conditions. For instance, the presumptive structure of polyA-CA fibres shows that one molecule of CA is present for every deoxyadenosine residue in each polyA chain. In other words, when one of the dA_{15} polyA chains binds the end of an elongating fibre, it should be accompanied by 15 CA molecules. While equilibrium dialysis experiments are consistent with this structure,[29] they have relatively low precision and the stoichiometry is very difficult to measure with accuracy. This property is of great interest since a CA:polyA stoichiometry of less than 15 would reveal the existence of defects, which could potentially be targeted with other small molecules. Thermodynamic data can help to resolve this issue, since the apparent dissociation constant, K_e, for a polyA chain binding to the end of the fibre should vary with CA concentration in a predictable way. For instance, if a polyA chain always brings with it *c* molecules of CA, i.e.

$$M_n + M_1 + cCA \overset{K_{eq}}{\longleftrightarrow} M_{n+1} \tag{2}$$

(following the nomenclature of *Equation 1*), then the full equilibrium dissociation constant for the process is given by

$$K^{\circ}_{(T)} = \frac{[M_n][M_1][CA]^c}{[M_{n+1}]} \tag{3}$$

This is something of an over-simplification, as elaborated below, but for now it serves to illustrate the dependence of K_e on [CA]. For polyA-CA fibres, CA is always in great excess so that its concentration is effectively constant for any set of assembly conditions. The apparent polyA dissociation constant K_e is related to the full equilibrium constant according to

$$K_e = \left.\frac{[M_n][M_1]}{[M_{n+1}]}\right|_{[CA]} = K^{\circ}_{(T)}[CA]^{-c} \tag{4}$$

with the temperature dependence of the standard equilibrium constant $(K^{\circ}_{(T)})$ described by

$$K^{\circ}_{(T)} = \exp\left(-\frac{\left(\Delta H_{(T)} - T\Delta S_{(T)}\right)}{RT}\right) \tag{5}$$

Therefore, measuring K_e at a series of different CA concentrations should produce offset van 't Hoff plots where the vertical distance between each line follows the stoichiometry of CA. To proceed, we noted that stabilization of polyA-CA fibres at high [CA] is largely entropic in nature, since it is primarily driven by differences in the entropy of dilution when dissociation of a polyA chain concomitantly releases *c* molecules of CA into solution.

We repeated the TREQ experiment at four CA concentrations between 7.5 and 15 mM *(Supplementary Figure 3)*. Van 't Hoff plots of the resulting K_e values are shown in *Figure 4*. Fitting *Equation 4* to this data

set allows us to directly obtain the stoichiometry of CA. To account for the possibility of a temperature dependent enthalpy value we extracted global values of ΔH_e and ΔC_p. The heat capacity change of binding, ΔC_p, accounts for any temperature-dependent differences in the slopes of the different experiments according to:

$$\Delta H_e(T) = \Delta H_e(T_0) + \Delta C_p(T - T_0) \quad (6)$$

$$\Delta S_e(T) = \Delta S_e(T_0) + \Delta C_p \ln\left(\frac{T}{T_0}\right) \quad (7)$$

The extracted ΔC_p = -0.6 ± 0.3 kcal mol^{-1} K^{-1} indicates that the enthalpy of adding a polyA chain to a growing fibre has only a slight temperature dependence. This is perhaps unsurprising, since ΔC_p values associated with nucleic acid folding are largely sequence dependent and have been observed to vary from slightly negative to positive values.[37] The global fit was in good agreement with experimental data points *(Figure 4, Supplementary Table 5)*. Surprisingly, the extracted stoichiometry coefficient, *c* = 10.4 ± 0.6, implies that 30% of possible CA binding sites are unoccupied in polyA-CA fibres under these conditions.

Master equations for high-valence assembly

The thermodynamics of multivalent supramolecular assembly can be summarized in terms of two main trends: the "principle of maximum occupancy" which refers to the tendency of systems to evolve toward the most stable state with full occupancy of binding sites, and the "entropy factor" which favours the state of the system with the largest number of product species.[38] For most of the supramolecular systems studied to date, the valency (number of binding sites per monomer) is relatively small (<6), the principle of maximum occupancy dominates, and the all sites are generally filled in the assembled materials.[39 40] However, for high-valence monomers, such as the polyA chains studied here, the entropy factor strongly opposes the principle of maximum occupancy and more complex behaviour emerges. For example, each dA_{15} chain creates an additional 15 potential CA binding sites, on average, as it adds to the end of growing fibre; one site must be created for each additional dA residue to achieve the theoretical 1:1 dA:CA stoichiometry. The number of ways to fill *c* of the 15 binding sites with *c* molecules of CA is given by the binomial coefficient[41]

$$N_c = \frac{15!}{c!\,(15-c)!} \quad (8)$$

While there is only *N*=1 way completely fill all 15 binding sites (*c*=15), there exists a total of *N*=32,766 distinct ways fill the sites with 1≤*c*≤14 molecules of CA. A simplified model of this energy diagram is seen in *Figure 5b*, where partially filled states are higher in energy but are more numerous. Therefore, even though a polyA chain with 15 bound CA molecules may represent the single lowest energy configuration, there exists such an enormous number of partly-filled configurations that these dominate, with a broad distribution of CA uptake and just 10 of the 15 sites being filled on average as seen in *Figure 5c*.

This explanation implies that polyA chains can bring a variable number of CA molecules with them when they attach to the end of a growing fibre, which is inconsistent with *Equation 4*, where the stoichiometry is fixed. To resolve this inconsistency, we developed a simple combinatorial model to describe polyA-CA fibre elongation. There is a free energy penalty for bringing an unbound polyA chain in close proximity to the end of a fibre, $\Delta G_{polyA}=\Delta H_{polyA} - T\Delta S_{polyA}$. This is compensated by energetically favourable binding of CA molecules to the newly-created 15 binding sites. All CA molecules are assumed to bind with equal free energy $\Delta G_{CA}=\Delta H_{CA}-T\Delta S_{CA}$. The total free energy change for a polyA chain binding along with a specific configuration of *c* CA molecules is $\Delta G_{polyA} + c\Delta G_{CA}$. Overall, the apparent equilibrium dissociation constant for polyA chain binding is given by[42]

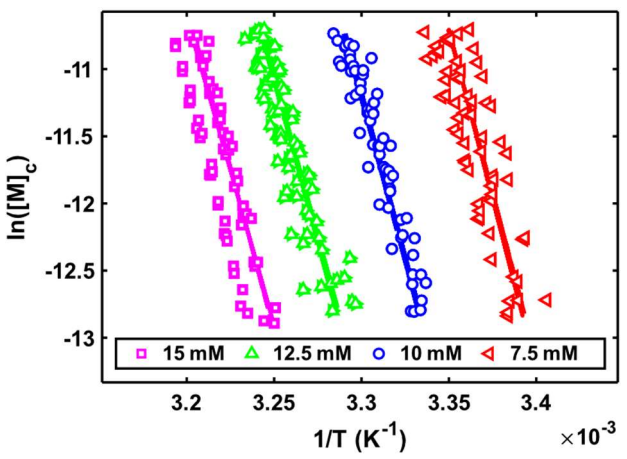

Figure 4: Van 't Hoff plot of TREQ data at different cyanuric acid concentrations. Coloured symbols represent experimental data from TREQ traces, solid-coloured lines represent a global fit of Equation 4 and dashed coloured lines represent a global fit of Equation 9. Solid and dashed lines are virtually superimposed on each other. Experimental errors are smaller than the size of the symbols.

$$(K_e)^{-1} = K_{polyA}(1 + K_{CA}[CA])^{15} \quad (9)$$

where $K_{polyA}=\exp(-\Delta G_{polyA}/RT)$ and $K_{CA}= \exp(-\Delta G_{CA}/RT)$. The average number of CA molecules can be calculated using the following equation

$$\langle c \rangle = 15\frac{K_{CA}[CA]}{1 + K_{CA}[CA]} \quad (10)$$

and the fraction of bound states with a given number of CA molecules can be calculated by

$$\theta_c = \left(\frac{15!}{c!\,(15-c)!}\right)\frac{K_{CA}[CA]^c}{(1 + K_{CA}[CA])^{15}} \quad (11)$$

We fit *Equation 9* to the TREQ data, obtaining excellent agreement, and extracting ΔH_{polyA}, ΔS_{polyA}, ΔH_{CA}, and ΔS_{CA} *(Figure 4, Supplementary Table 6).* These parameters allowed us to calculate the fractions of polyA chains with different numbers of CA molecules bound at different temperatures and [CA], providing a highly detailed description of assembly *(Figure 5c)*. Under highly stabilizing conditions of high [CA] and low temperature, the Equations predict that almost all binding sites are filled, in agreement with previous dialysis experiments.[29] Importantly, *Equation 9* and *Equation 10* explain why we observe 10 bound CA, and not more or less, even though experiments were performed at different [CA]. All experiments used 25 µM polyA, which means that we only detected K_e values between about 3 µM and 22 µM in all cases. This implies that the K_{CA}[CA] values are nearly identical in all experiments (since K_{polyA} does not change much with temperature) From *Equation 10*, this implies that <c> is very similar in all experiments, ranging from 10 to 11, and in excellent agreement with the simple fit described in the previous section.

High valence supramolecular systems have many useful properties that are only just beginning to be explored, such as the ability to self-heal, responsiveness to stimuli, and simple, inexpensive chemical

derivatization. Examples include small molecule-directed nucleic acid assembly (CA + polyadenosine or polydeoxyadeonsine[17, 29]; melamine + polythymine[43]) and non-covalent polymer crosslinking via multiple metal chelation[39, 44] or host/guest interactions[45, 46]. *Equation 9* and *Equation 10* can serve as starting points for quantitatively describing assembly in such systems, where simple probabilistic considerations ensure that some of the binding sites will remain vacant under many conditions. Furthermore, we find that TREQ-derived data are sufficient to extract the relevant thermodynamic parameters robustly, providing a new avenue for gaining insight into these complex materials.

Generality of the Method

Our aim for the TREQ method is that it can be used as a general tool to determine the thermodynamic parameters of supramolecular assembly when standard thermal melting and annealing experiments are unsuitable for thermodynamic analysis. Towards this end, we have also tested the method on a tetrameric intermolecular guanine quadruplex (G4) in aqueous buffer, and zinc-porphyrin self-assembly in mixture of methylcyclohexane and chloroform. In both cases, we obtained series of concave-up and concave-down arcs, similar to those of the polyA-CA fibres *(Supplementary Figure 7)*. In parallel, we used computer simulations to model the TREQ experiment for different types of self assembling systems and observed two patterns of behaviour: either all the extrema aligned with the equilibrium curve or the maxima for the cooling curves and minima for the heating curves were offset from one another *(Supplementary Figure 7)*. This provides a useful guide for interpreting TREQ data on new systems of interest: when the extrema align, they can be used to trace out the equilibrium curve (as for polyA-CA fibres and the intermolecular G4). When they are offset, they cannot be directly equated to equilibrium temperature/concentration pairs (as for the zinc porphyrin system), although the data are still information-rich, as detailed in the *Supplementary Information*. Fortunately, many slowly assembling supramolecular structures are amenable to the TREQ approach and, in these cases, it provides thermodynamic information that is not readily available from other sources. For example, a polyA-CA $[M]_c$ dataset similar to the one reported here would require a scan rate of <0.001 K min^{-1} in traditional melting measurements, leading to experiments on the impractically long timescale of a month. Our study demonstrates how the ready availability of high-quality thermodynamic dynamic data can lead to new insights, such as the prevalence of unfilled CA binding sites in polyA-CA fibres, and provides an opportunity to test theoretical developments, such as our master equation for high-valence assembly. These advances would not have been realistically possible for polyA-CA structures using previously existing methods.

A large number of slowly assembling supramolecular systems have been described in the literature, with only a subset referenced in this study.[11-27] This field is expected to expand in the coming years, since slow, nonequilibrium, nucleated assembly is a living polymerization process. The advantages of living polymers in supramolecular chemistry are an area of active research, with benefits already evident in the level of control they give over fibre length and monodispersity.[24, 26-28] Notably, thermodynamic information for slowly assembling systems is either completely lacking or determined using methods that we and others[10] have shown to be unreliable for such systems. We believe that the TREQ method presented here is a big step towards filling this gap in our knowledge. It can be applied to a wide variety of systems using common benchtop laboratory equipment and measurement times are on the order of 10 hours. The experiments are straightforward to set up and a typical analysis (eg van 't Hoff plot), can be performed entirely using standard spreadsheet software *(see Supplementary Information)*. We believe that the TREQ method will prove generally useful to the supramolecular chemistry community.

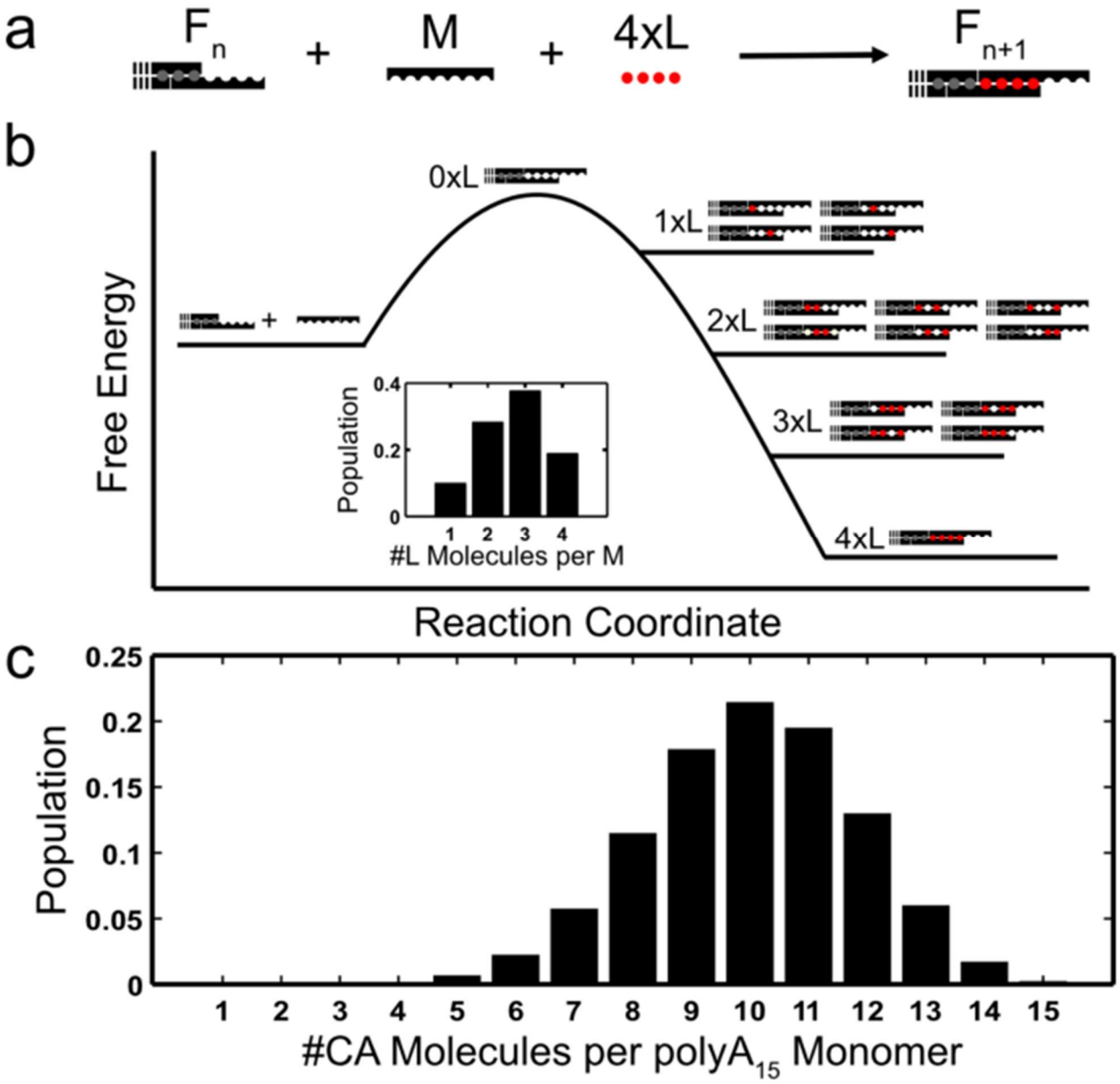

Figure 5*: **a)** A simple constant stoichiometry assembly model where the end of a growing fibre (F_n) assembles with one monomer M and 4 ligand molecules L to create a fibre of length n+1 (F_{n+1}). **b)** A free energy diagram of a variable stoichiometry assembly model where the end of a growing fibre (F_n) can assembly with a monomer M and any number of ligand molecules L up to a maximum of 4, the insert represents the populations of each stoichiometry. **c)** The populations of each stoichiometry for the self assembly of polyA-CA fibres at 25˚C with a concentration of 12.5mM CA.*

Online Methods

Materials

Cyanuric acid (CA), tris(hydroxymethyl)aminomethane (Tris), magnesium chloride hexahydrate ($MgCl_2 \cdot 6$ H2O), sodium chloride (NaCl), glacial acetic acid and urea were used as purchased from Sigma-Aldrich. Boric acid was obtained from Fisher Scientific and used as supplied. Acrylamide/bis-acrylamide (40% 19:1) solution, ammonium persulfate and tetramethylethylenediamine (TEMED) were used as purchased from BioShop Canada Inc.

1xTBE (Tris-boric acid-EDTA) buffer was composed of 45 mM Tris, 45 mM boric acid and 2 mM EDTA at pH 8.3. 1xAcMg buffer was composed of 40 mM acetic acid, 7.6 mM $MgCl_2 \cdot 6$ H2O, with pH adjusted to 4.5. Buffers and samples were prepared with Milli-Q water.

d(A_{15}) oligonucleotides were synthesized on a Mermade-12 synthesizer, purified by denaturing polyacrylamide gel electrophoresis (PAGE 20%, 1xTBE running buffer, 8 M urea) and desalted with Gel-Pak desalting columns from Glen Research. Purity of the strand was confirmed by HRMS (Calculated mass: 4635.18; Observed mass: 4634.28).

Stock solutions of 20 mM CA were prepared by dissolution in 100 mL of Milli-Q water in a volumetric flask and adjusted with acetic acid to pH 4.5. To properly dissolve and degas the solutions, they were heated at 65 °C and sonicated, then cooled down to room temperature before being used.

Samples of 100 µL of polyA_{15} (25 µM) and CA (7.5, 10.0, 12.5 and 15.0 mM) in pH 4.5 $Mg(OAc)_2$ buffer (7.6 mM) were made in quadruplicates. A thin layer (~30 µL) of silicon oil was applied on top to prevent evaporation during experiments.

Instrumentation

UV-Vis absorbance-based quantification of d(A_{15}) was performed on a Nanodrop Lite spectrophotometer from Thermo Scientific. DNA purification by PAGE was carried out on a 20 x 20 cm vertical acrylamide Hoefer 600 electrophoresis unit.

UV-Vis absorbance studies were performed using a 1.0 mm quartz cuvette and monitored at 260 nm on an Agilent Cary 300 Series UV-Vis Spectrophotometer equipped with a Peltier temperature controller and water recirculator. A variable temperature range which started from 50-40 °C and went down to 10-4 °C was scanned at a rate of 0.5 °C/min and with a max equilibration time of 30 min. Argon gas and drierite were used to dry the chamber at temperatures below 10 °C.

Data availability

The MATLAB code for performing the analyses herein is available from the corresponding author upon request.

Code availability

The data for performing the analyses herein are available from the corresponding author upon request.

References

1. Busseron, E.; Ruff, Y.; Moulin, E.; Giuseppone, N., Supramolecular self-assemblies as functional nanomaterials. *Nanoscale* **2013,** *5* (16), 7098-7140.
2. Elemans, J. A.; Rowan, A. E.; Nolte, R. J., Mastering molecular matter. Supramolecular architectures by hierarchical self-assembly. *Journal of Materials Chemistry* **2003,** *13* (11), 2661-2670.
3. Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S., Dynamic Combinatorial Chemistry. *Chemical Reviews* **2006,** *106* (9), 3652-3711.
4. Medrano, M.; Fuertes, M. A. n.; Valbuena, A.; Carrillo, P. J.; Rodríguez-Huete, A.; Mateu, M. G., Imaging and quantitation of a succession of transient intermediates reveal the reversible self-assembly pathway of a simple icosahedral virus capsid. *Journal of the American Chemical Society* **2016,** *138* (47), 15385-15396.
5. Pinotsi, D.; Buell, A. K.; Galvagnion, C.; Dobson, C. M.; Kaminski Schierle, G. S.; Kaminski, C. F., Direct Observation of Heterogeneous Amyloid Fibril Growth Kinetics via Two-Color Super-Resolution Microscopy. *Nano Letters* **2014,** *14* (1), 339-345.
6. Rennella, E.; Sekhar, A.; Kay, L. E., Self-assembly of human Profilin-1 detected by Carr–Purcell–Meiboom–Gill nuclear magnetic resonance (CPMG NMR) spectroscopy. *Biochemistry* **2017,** *56* (5), 692-703.
7. Bellot, M.; Bouteiller, L., Thermodynamic description of bis-urea self-assembly: competition between two supramolecular polymers. *Langmuir* **2008,** *24* (24), 14176-14182.
8. Pasternack, R. F.; Goldsmith, J. I.; Szép, S.; Gibbs, E. J., A spectroscopic and thermodynamic study of porphyrin/DNA supramolecular assemblies. *Biophysical journal* **1998,** *75* (2), 1024-1031.
9. De Greef, T. F.; Smulders, M. M.; Wolffs, M.; Schenning, A. P.; Sijbesma, R. P.; Meijer, E., Supramolecular polymerization. *Chemical Reviews* **2009,** *109* (11), 5687-5754.
10. Mergny, J.-L.; Lacroix, L., Analysis of thermal melting curves. *Oligonucleotides* **2003,** *13* (6), 515-537.
11. Greciano, E. E.; Alsina, S.; Ghosh, G.; Fernández, G.; Sánchez, L., Alkyl Bridge Length to Bias the Kinetics and Stability of Consecutive Supramolecular Polymerizations. *Small Methods* **2020,** *4* (2), 1900715.
12. Osypenko, A.; Moulin, E.; Gavat, O.; Fuks, G.; Maaloum, M.; Koenis, M. A. J.; Buma, W. J.; Giuseppone, N., Temperature Control of Sequential Nucleation–Growth Mechanisms in Hierarchical Supramolecular Polymers. *Chemistry – A European Journal* **2019,** *25* (56), 13008-13016.
13. Singh, S.; Zlotnick, A. J. J. o. B. C., Observed hysteresis of virus capsid disassembly is implicit in kinetic models of assembly. **2003,** *278* (20), 18249-18255.
14. Sambe, L.; de La Rosa, V. R.; Belal, K.; Stoffelbach, F.; Lyskawa, J.; Delattre, F.; Bria, M.; Cooke, G.; Hoogenboom, R.; Woisel, P., Programmable Polymer-Based Supramolecular Temperature Sensor with a Memory Function. *Angewandte Chemie International Edition* **2014,** *53* (20), 5044-5048.
15. Mizuno, K.; Boudko, S. P.; Engel, J.; Bächinger, H. P., Kinetic hysteresis in collagen folding. *Biophysical journal* **2010,** *98* (12), 3004-3014.
16. Dastan, A.; Frith, W. J.; Cleaver, D. J., Thermal Hysteresis and Seeding of Twisted Fibers Formed by Achiral Discotic Particles. *The Journal of Physical Chemistry B* **2017,** *121* (42), 9920-9928.
17. Harkness V, R. W.; Avakyan, N.; Sleiman, H. F.; Mittermaier, A. K., Mapping the energy landscapes of supramolecular assembly by thermal hysteresis. *Nature Communications* **2018,** *9* (1), 3152.
18. Harkness, R. W.; Hennecker, C.; Grün, J. T.; Blümler, A.; Heckel, A.; Schwalbe, H.; Mittermaier, A. K., Parallel reaction pathways accelerate folding of a guanine quadruplex. *Nucleic acids research* **2021,** *49* (3), 1247-1262.

19. Yamaguchi, M., Thermal Hysteresis Involving Reversible Self-Catalytic Reactions. *Accounts of Chemical Research* **2021**, 10226-10234.
20. Fukushima, T.; Tamaki, K.; Isobe, A.; Hirose, T.; Shimizu, N.; Takagi, H.; Haruki, R.; Adachi, S.-i.; Hollamby, M. J.; Yagai, S., Diarylethene-Powered Light-Induced Folding of Supramolecular Polymers. *Journal of the American Chemical Society* **2021,** *143* (15), 5845-5854.
21. Kar, H.; Ghosh, G.; Ghosh, S., Solvent Geometry Regulated Cooperative Supramolecular Polymerization. *Chemistry – A European Journal* **2017,** *23* (44), 10536-10542.
22. Fernández, Z.; Fernández, B.; Quiñoá, E.; Freire, F., The Competitive Aggregation Pathway of an Asymmetric Chiral Oligo(p-phenyleneethynylene) Towards the Formation of Individual P and M Supramolecular Helical Polymers. *Angewandte Chemie International Edition* **2021,** *60* (18), 9919-9924.
23. Xu, F.; Pfeifer, L.; Crespi, S.; Leung, F. K.-C.; Stuart, M. C. A.; Wezenberg, S. J.; Feringa, B. L., From Photoinduced Supramolecular Polymerization to Responsive Organogels. *Journal of the American Chemical Society* **2021,** *143* (15), 5990-5997.
24. Sarkar, S.; Sarkar, A.; George, S. J., Stereoselective Seed-Induced Living Supramolecular Polymerization. *Angewandte Chemie International Edition* **2020,** *59* (45), 19841-19845.
25. Haedler, A. T.; Meskers, S. C. J.; Zha, R. H.; Kivala, M.; Schmidt, H.-W.; Meijer, E. W., Pathway Complexity in the Enantioselective Self-Assembly of Functional Carbonyl-Bridged Triarylamine Trisamides. *Journal of the American Chemical Society* **2016,** *138* (33), 10539-10545.
26. Ogi, S.; Stepanenko, V.; Sugiyasu, K.; Takeuchi, M.; Würthner, F., Mechanism of Self-Assembly Process and Seeded Supramolecular Polymerization of Perylene Bisimide Organogelator. *Journal of the American Chemical Society* **2015,** *137* (9), 3300-3307.
27. Wang, H.; Zhang, Y.; Chen, Y.; Pan, H.; Ren, X.; Chen, Z., Living Supramolecular Polymerization of an Aza-BODIPY Dye Controlled by a Hydrogen-Bond-Accepting Triazole Unit Introduced by Click Chemistry. *Angewandte Chemie* **2020,** *132* (13), 5223-5230.
28. Mukhopadhyay, R. D.; Ajayaghosh, A., Living supramolecular polymerization. *Science* **2015,** *349* (6245), 241-242.
29. Avakyan, N.; Greschner, A. A.; Aldaye, F.; Serpell, C. J.; Toader, V.; Petitjean, A.; Sleiman, H. F., Reprogramming the assembly of unmodified DNA with a small molecule. *Nature Chemistry* **2016,** *8*, 368.
30. Alenaizan, A.; Fauché, K.; Krishnamurthy, R.; Sherrill, C. D., Noncovalent Helicene Structure between Nucleic Acids and Cyanuric Acid. *Chemistry–A European Journal* **2021,** *27* (12), 4043-4052.
31. Goldstein, R. F.; Stryer, L., Cooperative polymerization reactions. Analytical approximations, numerical examples, and experimental strategy. *Biophysical Journal* **1986,** *50* (4), 583-599.
32. Zhao, D.; Moore, J. S., Nucleation–elongation: a mechanism for cooperative supramolecular polymerization. *Organic & Biomolecular Chemistry* **2003,** *1* (20), 3471-3491.
33. Jonkheijm, P.; van der Schoot, P.; Schenning, A. P.; Meijer, E., Probing the solvent-assisted nucleation pathway in chemical self-assembly. *Science* **2006,** *313* (5783), 80-83.
34. van der Schoot, P., Supramolecular polymers. *Taylor & Francis Group, London* **2005**.
35. Arisawa, M.; Iwamoto, R.; Yamaguchi, M., Unstable and Stable Thermal Hysteresis Under Thermal Triangle Waves. *ChemistrySelect* **2021,** *6* (18), 4461-4465.
36. Markham, N. R.; Zuker, M., DINAMelt web server for nucleic acid melting prediction. *Nucleic acids research* **2005,** *33* (suppl_2), W577-W581.
37. Mikulecky, P. J.; Feig, A. L., Heat capacity changes associated with nucleic acid folding. **2006,** *82* (1), 38-58.
38. Self-Processes — Programmed Supramolecular Systems. In *Supramolecular Chemistry*, 1995; pp 139-197.
39. Hamacek, J. J. M., Self-Assembly Principles of Helicates. Wiley Online Library: 2013; pp 91-124.
40. Paneerselvam, A. P.; Mishra, S. S.; Chand, D. K. J. J. o. C. S., Linear and circular helicates: A brief review. **2018,** *130* (7), 96.

41. Hamacek, J.; Borkovec, M.; Piguet, C. J. C. A. E. J., A Simple Thermodynamic Model for Quantitatively Addressing Cooperativity in Multicomponent Self-Assembly Processes—Part 1: Theoretical Concepts and Application to Monometallic Coordination Complexes and Bimetallic Helicates Possessing Identical Binding Sites. **2005,** *11* (18), 5217-5226.
42. Dill, K. A.; Bromberg, S.; Stigter, D., *Molecular driving forces: statistical thermodynamics in biology, chemistry, physics, and nanoscience*. Garland Science: 2010.
43. Li, Q.; Zhao, J.; Liu, L.; Jonchhe, S.; Rizzuto, F. J.; Mandal, S.; He, H.; Wei, S.; Sleiman, H. F.; Mao, H., A poly (thymine)–melamine duplex for the assembly of DNA nanomaterials. *Nature Materials* **2020,** *19* (9), 1012-1018.
44. Lehn, J.-M.; Rigault, A.; Siegel, J.; Harrowfield, J.; Chevrier, B.; Moras, D., Spontaneous assembly of double-stranded helicates from oligobipyridine ligands and copper (I) cations: structure of an inorganic double helix. *Proceedings of the National Academy of Sciences* **1987,** *84* (9), 2565-2569.
45. Mulder, A.; Huskens, J.; Reinhoudt, D. N., Multivalency in supramolecular chemistry and nanofabrication. *Organic & biomolecular chemistry* **2004,** *2* (23), 3409-3424.
46. Lin, M.; Dai, Y.; Xia, F.; Zhang, X., Advances in non-covalent crosslinked polymer micelles for biomedical applications. *Materials Science and Engineering: C* **2020**, 111626.

Acknowledgements

The authors thank Professor E. W. Meijer for supplying the zinc porphyrin sample as well as for the many fruitful discussions of the work. C.D.H. was supported by an NSERC postgraduate scholarship. C.L.B. was supported by an NSERC Vanier Canada Graduate Scholarship. H.S. would like to thank NSERC, CIHR, CFI, and the Canada Research Chairs Program for funding. H.S. is a Cottrell Scholar of the Research Corporation. A.M. was supported by NSERC (Discovery grant number 327028–09).

Author Contributions

H.S. and A.M. conceived the project. C.L.B. acquired all experimental data. C.D.H. performed simulations, data analysis and interpreted the results with A.M. A.M. and H.S. supervised the projects. C.D.H. and A.M. wrote the paper.

Competing Interests

The Authors declare no competing interests.