Detection of Polypeptide Conformational Transitions in Solution via Sound Velocity

Alyssa M. Blake[1,3], Graham D.B. Parkinson[2,3] and Paul S. Russo[1,2,3]*

[1] School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA, USA 30332

[2] School of Materials Science and Engineering, School, Georgia Institute of Technology, Atlanta, GA, USA 30332

[3] Georgia Tech Polymer Network, GTPN, Georgia Institute of Technology, Atlanta, GA, USA 30332
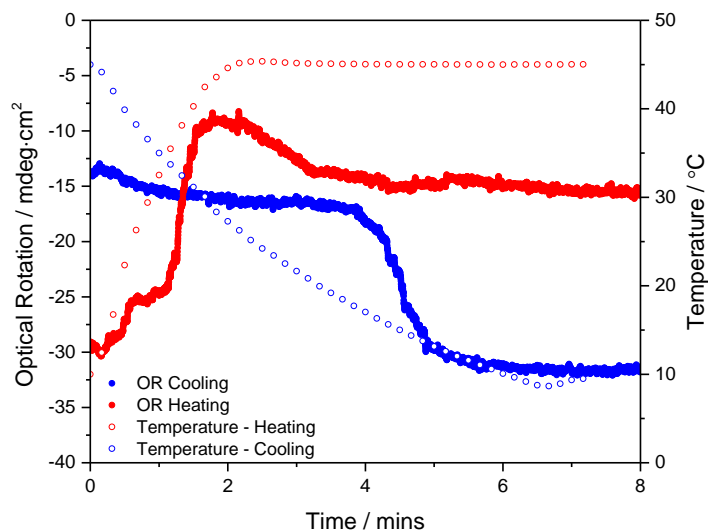
Supporting Information

Figure S1. Temperature jump control experiment using PCBL-79 in m-cresol. A heating and cooling experiment was conducted to determine how quickly the conformational transition occurs to determine the hold time for the optical rotation experiments.

Data Analysis for Optical Rotation Measurements. Analysis of optical rotation data requires the specific optical rotation to be calculated as mentioned in the main text. Figure 3 in the main text shows the specific optical rotation for the three PCBL samples. To aid in identification of the transition temperature, a derivative of the speific optical rotation was calculated and is shown in Figure S2. From the derivative curve, it is clear that the transition occurs around ~19 °C with agreement between the heating and cooling measurments. The reversibility and reproducbility of the conformational transition is clear when comparing the two temperature cycles.
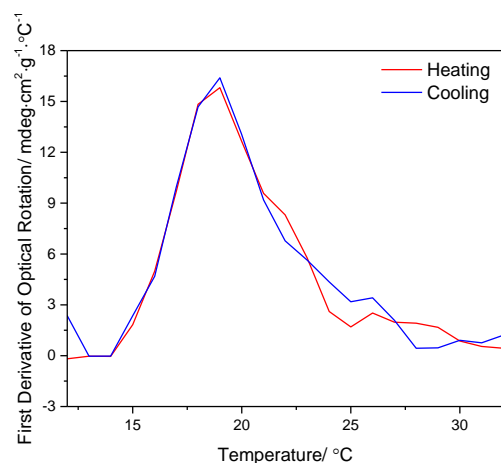
2

Figure S2. A first order derivative of the heating and cooling data for PCBL-155 to determine the transition temperature from optical rotation. A transition temperature of ~19 °C was obtained for PCBL-155.

A problem arises once the concentration is reduced for the PCBL samples. The derivative curves have more noise in their baselines. Due to this increase in noise, a Logistic 5 sigmoidal fit (Origin Pro 2016) can be applied to the optical rotation data (e.g., Figure 4A-C main text) to help smooth the data as shown in Figure S3A. The first order derivative is taken of this fit to help elucidate the transition temperature for each sample. Logistic regression model is commonly used to analyze/ model a system made up of a binary response. The probability of the polypeptide system to exist as one of two classes/events (random coil or α-helix), makes it suitable to be analyzed by a logistic sigmoidal fit. The derivative curves obtained from the logistic fit for PCBL-155 are shown in Figure S3B, illustrating a transition temperature around 19 °C. Matsuoka et.al reported a transition temperature of ~26 °C for a PCBL sample with a similar molecular weight.[12] The disagreement between the values previously reported and the values in this work could arise from differences in

instrumental equipment, sample preparation, polypeptide concentration, dispersity, and solvent purity.[12] The skewness of the derivative curve highlights that the sigmoidal curve is not entirely symmetrical, in agreement with the Zimm-Bragg analysis.
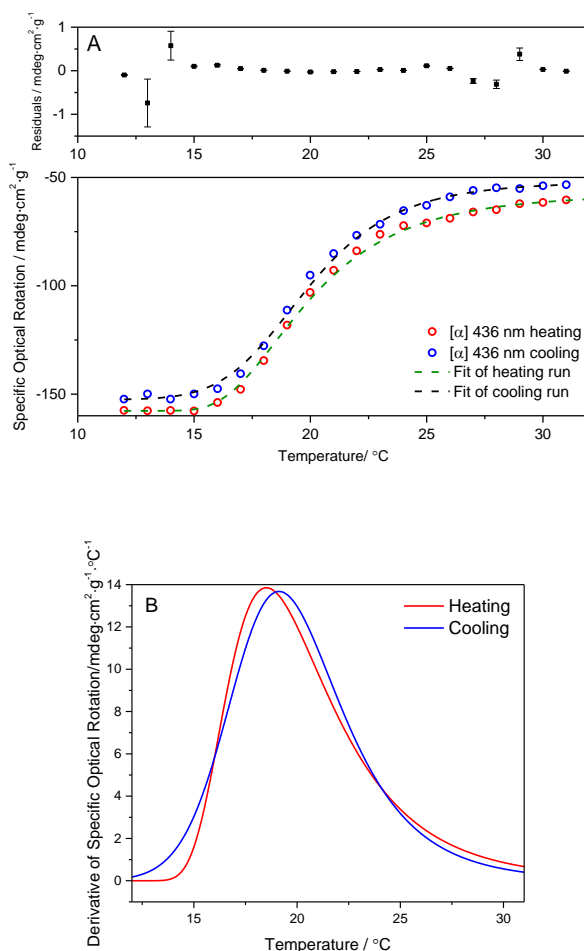

Figure S3. (A) Logistic 5 sigmoidal fit applied to the specific optical rotation data for PCBL-155. The fit helps to smooth the curve and reduce the noise of the data. (B) A first order derivative was taken of the Logistic 5 sigmoidal fit for PCBL-155 to determine the transition temperature for the optical rotation data. A transition temperature of ~19 °C was obtained for PCBL-155.

Data Analysis of Sound Velocity Measurements. Similar to optical rotation, sound velocity measurements require post-measurement data processing and analysis. A plot of sound velocity against temperature is almost linear, as shown in Figure 4A of the main text. Close inspection reveals a barely perceptible deviation from linearity near the polypeptide transition temperature according to OR measurements. We took the derivative of the raw sound velocity data to illustrate this minute deviation from linearity, resulting in Figure 4B of the main text, which indicates a change in the sample at ~18 °C. The high precision (5 significant digits) of the sound velocity measurements enables the subtle change to be detected. The difference of fit can be calculated for each temperature run by determining a baseline from the derivative. Figure 4B of the main text shows the derivative for PCBL-155. At higher tempearatures, above 24 °C, the sound velocity levels off showing no significant change. For this reason, the baseline was chosen to be 25-30 °C. These five data points were plotted and a linear trendline was generated to calculate the difference of fit for the entire temperature run from 13-30 °C.

Although identifying the existence of a transition point by taking a derivative was effective, a different procedure was used to display the transition as a continuous change. An example for calcluating the difference of fit for the sound velocity of PCBL-155 follows. Using the equation,

$$y = a + bx \tag{2}$$

where b is the slope and a is the y-intercept obtained from the linear fit equation generated from a five-data point baseline, the extrapolated sound velocity can be calculated for each tempearture between 13-30 °C. This will be denoted as the sound velocity fit data (SV_{calc}). Next, the difference between the SV_{calc} data and the experimental sound velocity data can be calculated ($SV-SV_{calc}$).

5

The difference is then plotted as a function of temperature producing the sound velocity (SV-SV$_{calc}$) transition curve.
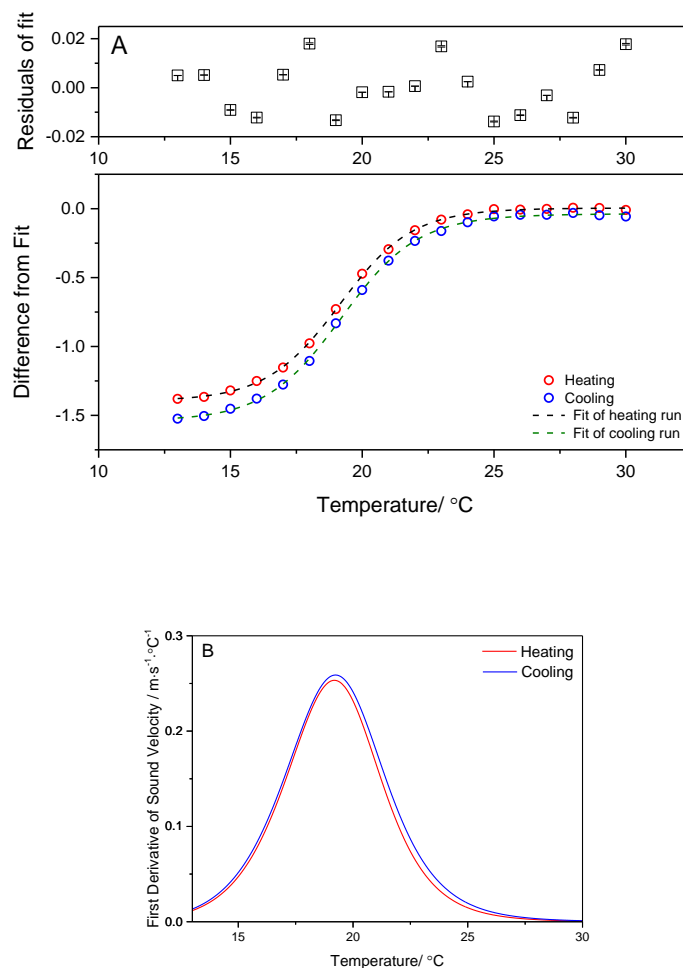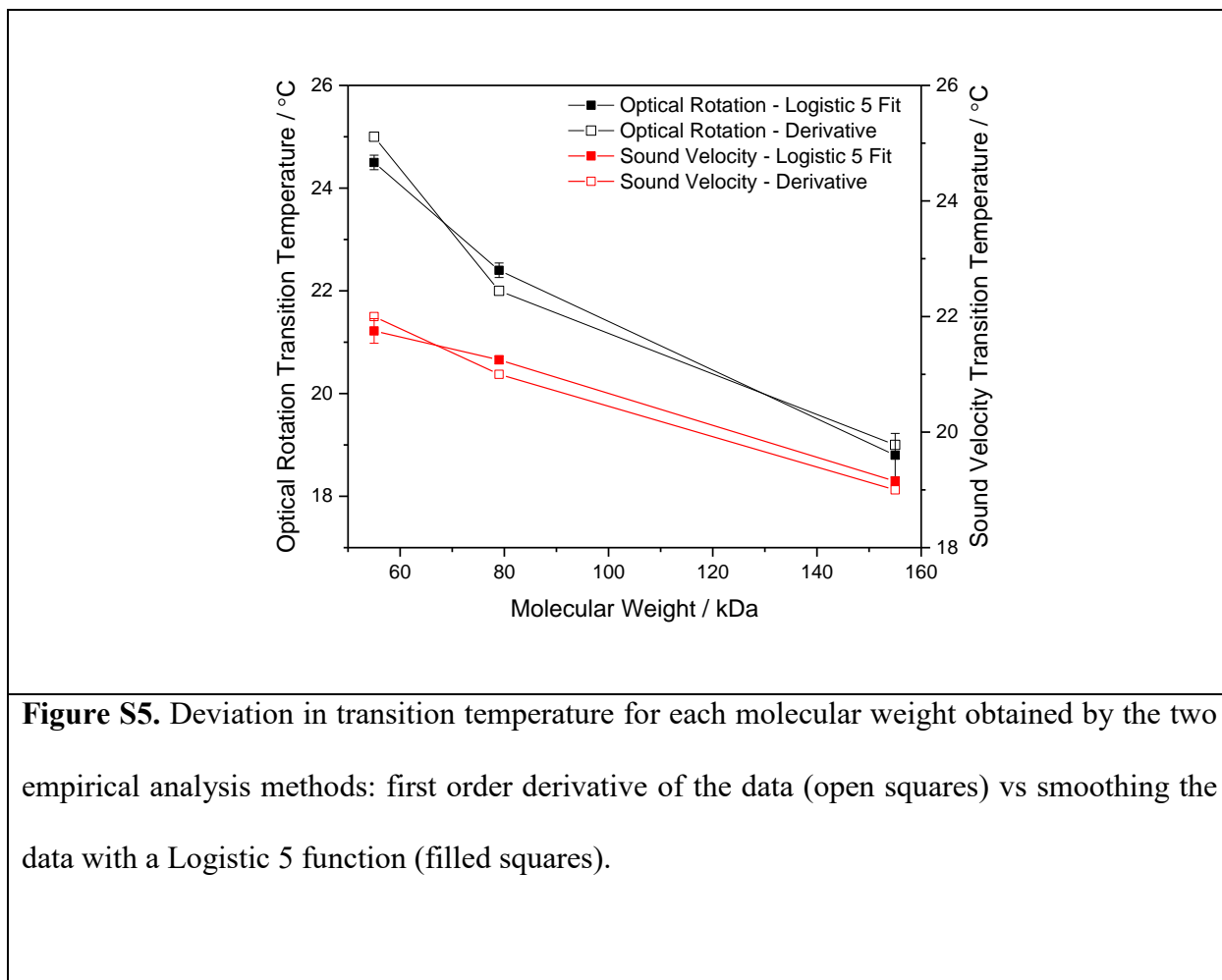

Figure S4. (A) Logistic 5 sigmoidal fit applied to both heating and cooling sound velocity data for PCBL-155. Smoothing of the data helps to reduce the noise that can interfere with obtaining a transition temperature. (B) A first order derivative was taken of the fit to determine the transition temperature for the sound velocity data and was determined to be ~19 °C.

Similar to the OR, once the concentration of PCBL is decreased, the sound velocity data and its respective derivative become noisy. To counteract this increase in noise level, the Logistic 5 fit was applied to the sound velocity data in the same way it was applied to the OR data. Once the fit was obtained, the first order derivative was taken to determine the transition temperature. Figure S4 shows the SV data with the Logistic fit applied and then the derivative obtained from the fit.

Figure S5. Deviation in transition temperature for each molecular weight obtained by the two empirical analysis methods: first order derivative of the data (open squares) vs smoothing the data with a Logistic 5 function (filled squares).
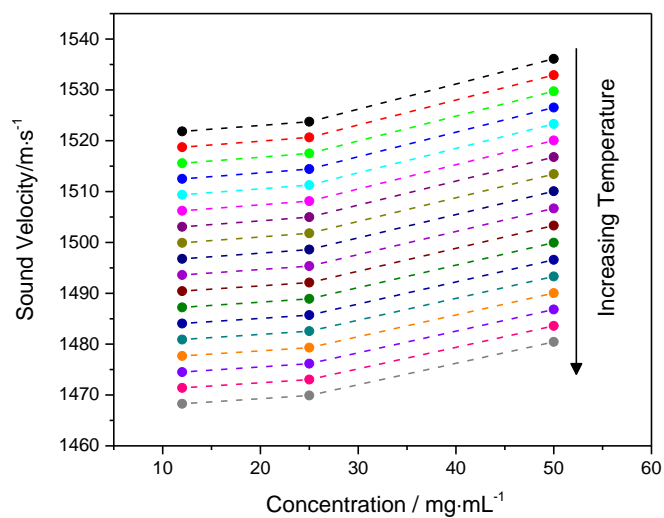
Sound Velocity Trends

Figure S6. Sound velocity increases with concentration at any temperature measured.
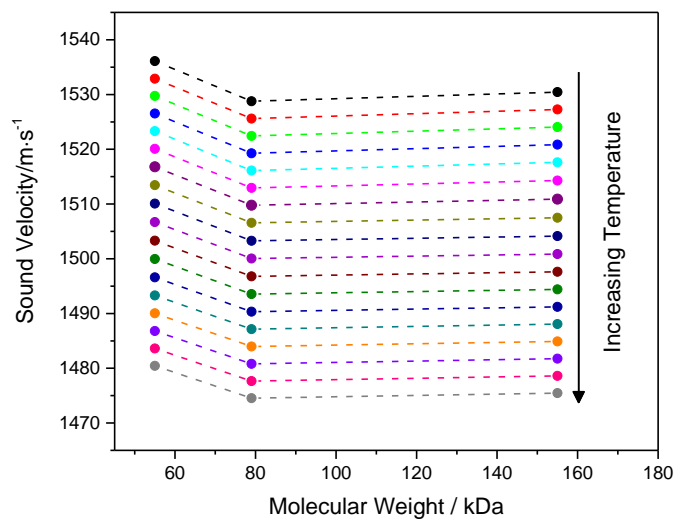

Figure S7. Sound velocity decreases moderately with molar mass at any temperature measured.

Zimm-Bragg Simulations and Fitting

Figures S8 and S9 overlay output from the MATLAB code, representing Equation 3 in the main text, onto data converted to f_H from sound velocity or optical rotation signals. In a typical manual fitting session, we began with $\Delta H = 0.93$ kcal/mol, $\sigma = 6.25 \times 10^{-6}$, and $\Delta S = (1/300)$ kcal·mol^{-1}·K^{-1}. Small adjustments were sometimes made to increase or decrease the span of the SV or OR signals, up to 5% (program variable smush). Only in one case was subtraction a linear offset deemed necessary (about 2%) to ensure a zero signal at low temperatures (program variable offset). Adjustments to match the theory to the data were made, as follows:

- Increase ΔH: flattens low-T part of curve; pushes transition to higher temperatures.

- Decrease σ: sharpens the upturn near the critical temperature and steepens the rapidly rising part of the curve.

- Decrease parameter "smush" (default value = 1). This parameter was used to scale the input data to $f_H \approx 0.9$ at 30°C, as suggested by Matsuoka et al.[1]

The transition temperature was taken as the value when $f_H = \frac{1}{2}$. Due to the asymmetry of the sigmoidal curves, which reflects the high cooperativity of the transition, this slightly exceeds the inflection point identified by the empirical method.
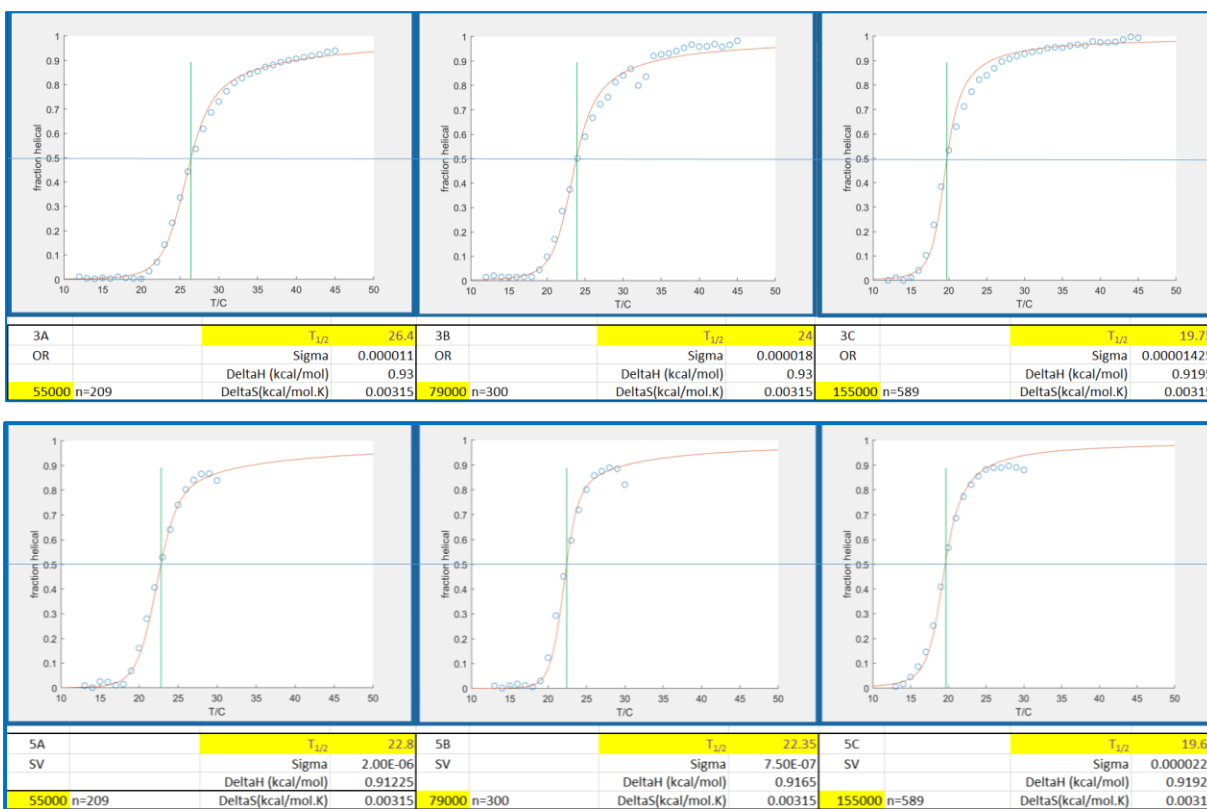
3A		$T_{1/2}$	26.4
OR		Sigma	0.000011
		DeltaH (kcal/mol)	0.93
55000	n=209	DeltaS(kcal/mol.K)	0.00315

3B		$T_{1/2}$	24
OR		Sigma	0.000018
		DeltaH (kcal/mol)	0.93
79000	n=300	DeltaS(kcal/mol.K)	0.00315

3C		$T_{1/2}$	19.75
OR		Sigma	0.00001425
		DeltaH (kcal/mol)	0.9195
155000	n=589	DeltaS(kcal/mol.K)	0.00315

5A		$T_{1/2}$	22.8
SV		Sigma	2.00E-06
		DeltaH (kcal/mol)	0.91225
55000	n=209	DeltaS(kcal/mol.K)	0.00315

5B		$T_{1/2}$	22.35
SV		Sigma	7.50E-07
		DeltaH (kcal/mol)	0.9165
79000	n=300	DeltaS(kcal/mol.K)	0.00315

5C		$T_{1/2}$	19.63
SV		Sigma	0.0000225
		DeltaH (kcal/mol)	0.91925
155000	n=589	DeltaS(kcal/mol.K)	0.00315

Figure S8. MATLAB-simulated helix vs. temperature curves using Equation 2 of the main text are overlaid on data measured different molecular weights, both for optical rotation (top) and sound velocity (bottom) experiments. Parameters used are shown in tables.
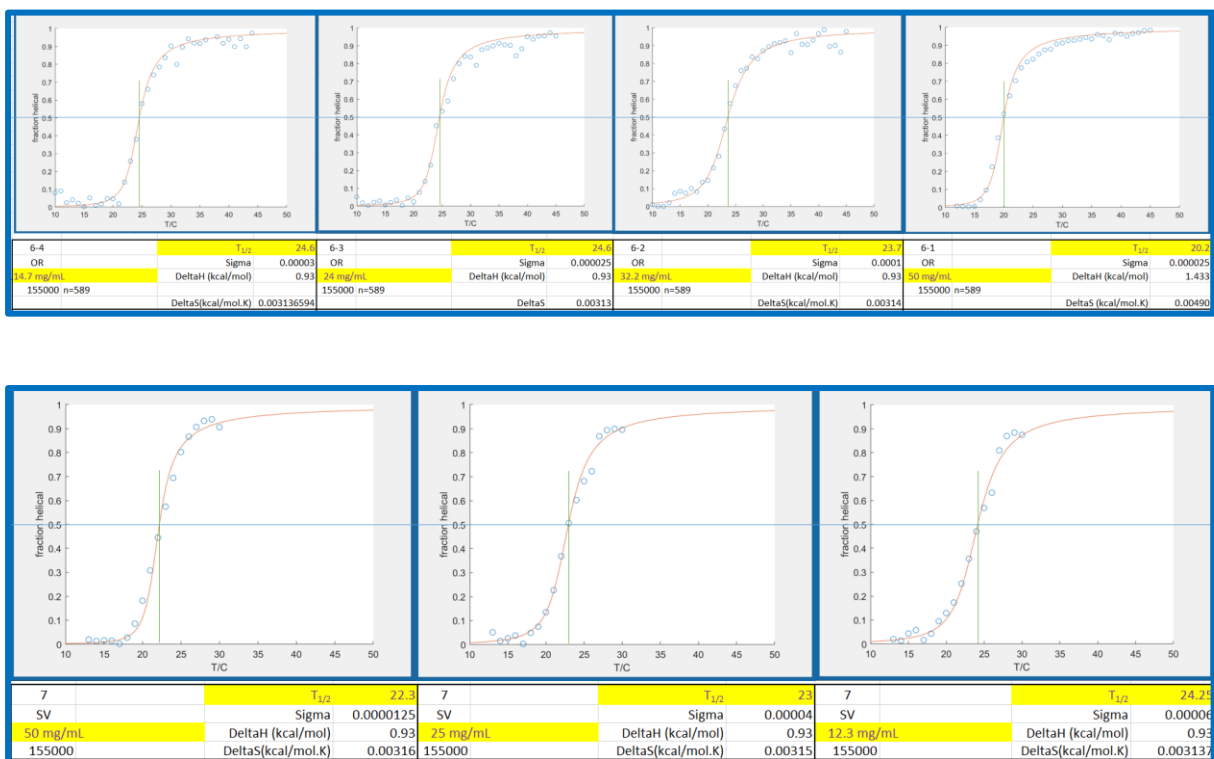
Figure S9. MATLAB-simulated helix vs. temperature curves using Equation 2 of the main text are overlaid on data measured different concentrations for PCBL-155, both for optical rotation (top) and sound velocity (bottom) experiments. Parameters used in the manual simulation are shown in tables. The sample at 32.2 mg/mL was not measured by SV.

MATLAB code

MATLAB Version: R2017b

- **Purpose:** imports one dataset already converted to f_H vs. T (centigrade) by associating the low-T part of an OR or SV curve with 0% helix and the high-T part with 100% helix, permits adjustment of ΔH, ΔS, σ, scaling and offset of the curves.

- **Speed:** About 1 s (Laptop, Windows 10 64-bit, Intel i7 @ 2.9 GHz, 16 Gb RAM)

```matlab
% ZimmBraggImportOneFhTrace.m
% Text after green percent signs is comments.
% Generates Zimm-Bragg total helix fraction for 2 x 2 statistical weight matrix
% (nearest neighbor only; μ = 1 in Zimm-Bragg notation). This code plays fast
% and loose with n; assume it's same as DP.
%% Import the data from Excel...this code generated by import wizard.
[~, ~, raw] = xlsread('C:path.filename.xlsx','SheetName','j30:k63');
%The data should already look like a sigmoidal curve; prepare in Excel.
%The last bit—j30:k63—is top left ColumnRow and bottom right ColumnRow
%containing the experimental data points in two adjacent columns.
%% Create output variable
data = reshape([raw{:}],size(raw));
%% Allocate imported array to column variable names
Tex = data(:,1); %experimental temp in centigrade
smush=1;%Use sparingly (e.g., 0.98 or 1.02) to adjust span of SV or OR data
offset=0.0;%Subtract or add a constant; rarely needed.
fhex = (data(:,2)-offset)*smush; %experimental fh values
%% Clear temporary variables
clearvars data raw;
%%
n=589;%we do not distinguish between DP and n
Rgas=0.001987 %kcal/mol.K
DeltaH=.93 %kcal/mol…Matsuoka's suggestion
%DeltaS=DeltaH/295 %kcal/mol.K  used sometimes...gives a value ~0.003
DeltaS = 0.00315; %selected because it's about DeltaH/300K
Sigma=6.25e-6; %Matsuoka's suggestion
%Sigma=.00001425; comment out above line if you activate this one
ZeroMatrix=[0 0; 0 0]
for i=1:100 %begin simulation
 TC(i) = 9+i*0.5;%starts at 9.5C, then every 0.5 degree for 100 points
 TK(i)=TC(i)+273.15
 s(i)=exp(-(DeltaH-TK(i)*DeltaS)/(Rgas*TK(i)));
 SigS=Sigma*s(i);
 U = [1 SigS;   1   s(i) ];% 2 x 2 statistical weight matrix
 Jleft=[1 0];%left extractor, Eq. 2 main text
 Jright=[1 1]';%right extractor, Eq. 2 main text
 z=Jleft*U^n*Jright;%n=how many assignable bonds...about DP
 Us = [0 SigS;0 s(i)];%the 2 x 2 matrix Us in Eq. 2 main text
 Jleftfh=[1 0 0 0];%left extractor for the composite matrix
 Jrightfh=[0 0 1 1]';%right extractor for the composite matrix
 Ucomposite=[U Us;ZeroMatrix U];%4 x 4 composite matrix; Eq. 3 main text
 fh(i)=Jleftfh*(Ucomposite^n)*Jrightfh/(n*z);
end
s;%remove ; to print s if you want
fh;
figure;
scatter(Tex,fhex)
hold on
plot(TC,fh)
axis([10 50 0 1])
xlabel('T/C')
ylabel ('fraction helical')
DeltaH
DeltaS
Sigma
xlswrite('fh_ManualFitResults.xlsx',[TC(:),fh(:)]);%write result to Excel
```

Reference

1.   Matsuoka, M.; Norisuye, T.; Teramoto, A.; Fujita, H., Solution Properties of Synthetic Polypeptides. XV. Helix-Coil Transition in Poly(ε-carbobenzoxy L-lysine). *Biopolymers* **1973,** *12*, 1515-1532.   doi/10.1002/bip.1973.360120707